NEWS TRACING SYSTEM USING BLOCKCHAIN

Akshada Babar  
Department of CSE  
Dr.D.Y.Patil Institute of Engineering, Management and Research.

Prof. Nalini Jagtap  
Assistant Professor  
Department of CSE  
Dr.D.Y.Patil Institute of Engineering, Management and Research.

Akshata Mithari  
Department of CSE  
Dr.D.Y.Patil Institute of Engineering, Management and Research.

Aakash Shukla  
Department of CSE  
Dr.D.Y.Patil Institute of Engineering, Management and Research.

Prachi Chaudhari  
Department of CSE  
Dr.D.Y.Patil Institute of Engineering, Management and Research.

Abstract—The current pandemic situation has made us realize how fake news can have detrimental impacts on society. Although we can't ignore the benefits of social media, we surely can't overlook the misuse of these mediums to spread rumors and fake articles. These posts/articles/messages can be commonly called "Fake News" and not only create misinformation but also create panic among the public. Blockchain is a disruptive technology that is being used to make innovative solutions in the sectors of food, fashion, supply chain, and banking. The decentralized and transparent nature of blockchain can not only be used in the above sectors but it can also be used to tackle the problem of fake news. Several solutions have been proposed to combat the problem using blockchain. This paper will propose one such solution to combat the fake news problem that can be used by all the news agencies. The system is called the "News Tracking and Tracing System" and is built using Ethereum.

Keywords—Blockchain, Blockchain to curb fake news, Ethereum, Fake News, Fake News Detection using blockchain, IPFS, smart contracts, Solidity, traceability, Tracing Fake News.

I. INTRODUCTION

‘Who exactly is responsible for these fake news?’ is a question that cannot be answered because with the recent availability of mediums like Twitter, Facebook, Whatsapp and Instagram, everyone has become a Journalist! Journalists, in the sense, can make articles that contain rumors and spread them on a vast scale. Although the general public is responsible for the spread of misinformation, the role of journalists, news agencies, and government cannot be ignored. They are equally responsible for misleading the public. According to an article by Peter Vanderwicken in the Harvard Business Review, "The news media and the government have created a charade that serves their interests but misleads the public. Officials oblige the media's need for drama by fabricating crises and stage-managing their responses, thereby enhancing their prestige and power. Journalists dutifully report those fabrications. Both parties know the articles are self-aggrandizing manipulations and fail to inform the public about the more complex but boring issues of government policy and activity [1]." This means that the government is using the News agencies and manipulating the news content for certain political and financial benefits. On the other hand, the news agencies are manipulating the news to make it more interesting to attract public attention. The focus has shifted from authentic content to making the news more attractive. A very recent example of this can be seen in China, according to the New York Times article [2], the Chinese government is hiding the real facts after the opening of the lockdown in Wuhan by misinforming them and creating a positive image about the current situation. News agencies are supporting the government and
publishing these articles. Hence, this clearly explains how news agencies can unintentionally or intentionally spread fake news and misguide the public. Hence, we have built a system using blockchain that can be used by the news agencies to create news articles and then publish them, the system will track every point in the process of news content creation. If any entity in the news agency is trying to manipulate the news content, then that entity can be easily tracked by the crowd auditor. Hence, the spread of fake news can be stopped right at the root of its creation and the culprit can be found out easily. the system prototype has been inspired by [3].

II. RELATED WORK

According to Xinyi Zhou, Reza Z. et. al (2018) [4], the detection mechanisms can be broadly classified into Knowledge- based detection, Style based detection, Propagation based detection and Credibility based detection. Knowledge based detection uses fact checking mechanisms whereas style based detection uses different deception theories to understand the intention of the news article. On the other hand, propagation based deception tries to explore the nature of propagation of the article and credibility based detection refers to using the article source to determine its authenticity. Using blockchain to trace the source of the fake news and then determine its authenticity can be partially termed as a credibility based detection method. According to Wenqian Shang et. al. (2018) there are many systems proposed in the past for fake news tracing and detection using blockchain. [5] proposes a blockchain based system that is using the method of source evaluation and it ensures that the news article has not been tampered. Also, it uses multi-node verification to make sure that only authentic news articles are circulated over the internet. This system is interfaced with an online news reading application. The block described by this system contains a block header and a block body. The block header contains the previous block hash value, current block hash value and timestamp. The block body contains the transaction data. These blocks are connected to each other to form a chain and hence, the source of the news can be traced easily using this chain. The readers can trace back to the source to evaluate the article using the blockchain network. As stated in Wee Jing Tee et al (2018) [6], blockchain has a high potential to solve many critical problems. According to the research in this paper, blockchain technology will evolve over the next few years, and so will AI. Both these technologies in combination can be used to combat the fake news problem. The paper provides an extensive review on the use of blockchain for solving the fake news problem. The dummy blockchain systems in the past have proposed maintaining a blockchain for every individual using the social media which is not yet possible and hence, this prototype has been modified in Muhammad S. et al (2019) [7], this research proposed a system to detect, curb and trace the fake news. The system contains a blockchain network that will store all the transactions and the users of the social media network need not maintain a personal blockchain for the effective working of this system. The network will be maintained by the News Agencies and Social Media platforms in collaboration. The users can identify the fake news, malicious users producing misinformation will be blocked by the system. The transaction throughput is increased by using the BFT consensus mechanism. A different perspective of handling the image tampering can be seen in the Safi ur Rehman et al (2020) [8], it explains the usage of hamming distance to find out the similarity between the images. In this paper the authors describe an image tampering detection system to be integrated with the provenator application in order to increase the effectiveness of the Provenator. The system uses the method of hamming distance detection to find out the similarity between the images, if the hamming distance is above a certain threshold then the images are considered as similar distance or else they have been tampered with. Finally in Zonyin Shae et al (2019) [9], focus has shifted to combining AI with blockchain to implement a system to detect and trace fake news. The existing systems describe a prototype that can be interfaced with the social network and curb the spread of fake news. These prototypes concentrate on the spread of misinformation on the social media sites. Although the majority of fake news is spread on these sites, it can not be ignored that the news agencies are responsible to some extent to produce misinformation. As stated in the article[10], the way of presenting the news has changed. People have blind faith in the articles put out by these news agencies and hence, it is necessary that they have a transparent system that ensures the news processing system is unbiased and ethical. In this paper, we propose a system that can be implemented by the news agencies to publish news that is tamper proof and original.

III. BLOCKCHAIN

Before explaining the architecture of the blockchain system we will understand some blockchain terminologies and the different components used in the application. 'Blockchain' as the name suggests is a chain of blocks, these blocks contain transactions and each block is stored at a certain node somewhere[11]. These nodes are connected to form a network which is called the 'Blockchain Network'. Creating a block and deploying the block at a certain node over the network is called 'Mining the Block'. Mining is a very expensive task as it requires a lot of resources. Hence, usually, the miners are paid a fee for mining the block. Once a block is created, it can not be undone, that is why blockchain has the property of immutability.
Smart Contract: Certain rules needed to be checked before uploading the data to the block. These rules are written in a Smart Contract. So basically, the smart contract controls the entire transactions on the blockchain network. The code written in the blockchain is triggered when certain conditions in the contract are reached. This smart contract is created by consent from all the members taking part in the network and once deployed, it cannot be changed, deleted, or modified in any way. Hence, the smart contract must be properly tested before the deployment of the network.

IV. PREREQUISITES

Ethereum[12]: Ethereum is a distributed public blockchain network. It can be used to program a decentralized application and deploy it on the blockchain network. The miners are paid in ether. There is another token that is used for carrying out the transactions that are gas. A certain gas amount is paid for each transaction.

Geth: Geth is an Ethereum node that is implemented in the Go language. By installing geth we become a part of the ethereum public network. Geth will adjust to the live blockchain or create its own. Hence, we can now create software that uses the Ethereum Virtual Network.

Ganache[13]: Ganache is used for Ethereum development, it is a part of the truffle suite and it provides us with a private local blockchain network. We get 10 addresses and we can perform any actions that we would on the main chain here.

Truffle: Truffle is a framework used for the development and testing of the contracts. Truffle allows us to code smart contracts and tests them. We can deploy the smart contract to a blockchain network which will make it much easier.

Metamask: Metamask allows us to run the Ethereum Dapps to run right in the browser. It is a wallet that will store our ether and it is a chrome extension hence, it will allow us to make ethereum transactions through regular websites.

IPFS: IPFS stands for Interplanetary File System[14] and it is used to store data in a decentralized manner. In IPFS, the data to be stored is divided into small blocks, each block has its cryptographic hash and then these files are stored on different nodes. After storing the file completely a file hash is returned by IPFS. This file hash is used to get back to the file that is stored over the IPFS.

To build the system, we have to make sure that Node.js has been installed in the system. The above prerequisites are used to build the blockchain system. That is these are the tools and technologies used to build the blockchain application.
V. SYSTEM DESIGN

Each participating entity will have interactions with the smart contract. They will communicate with each other through transactions. The proposed model will consist of four entities:

**Journalist:** A journalist is the one who collects, writes, and reports the news for the public. In our system, the journalist will be the entity who is responsible for collecting the news and then creating a post in the system.

**Editor:** A editor is the one who tweaks and censors the news articles to make it presentable for the general public. In the system, he can view all the uploads made by the journalist and then modify the content to make it suitable for presentation.

**News Agency:** A news agency is an institution under which the journalists and editors work. The news agency will be on the highest level in the hierarchy, the news agency is responsible for publishing the articles created by the journalist and modified by the editor.

**Crowd Auditor:** Crowd Auditors are some people from outside the news agency and they will filter out the fake and authentic news articles. The crowd Auditor can see the posts approved by the news agency. The crowd auditors can trace the post and then report them as fake or authentic.

The system will display the count of the votes that reported the article as fake and also the count of the votes that reported the article as true. The system will be completely transparent as every post can be traced back to the origin. If any modification is made by any entity, then the culprit can be found out.

### A. WORKING OF THE SYSTEM

The journalist will have to log in to his Metamask wallet. After the journalist logs in the wallet, he can see his Home page. He can create a post, by entering the article name and he can upload a file related to the article. He can upload the post by clicking on the "upload" button. The file uploaded by the journalist will be stored on the IPFS and the file hash will be returned by IPFS API. The journalist will have to pay the gas fee, after confirming the transaction on metamask and all the post content and the file hash will be stored on the blockchain network. The Editor (after logging into his metamask wallet) can see his Home Page on which all the posts uploaded by the Journalist along with the related file will be displayed. The editor can modify the post, he can change the name of the post, file uploaded, etc. The modified post will be displayed on the News Agency's home page when he will log in to his metamask account. The Newsagency can see a warning line above each post that has been modified by the editor. The news agency has the option to trace the post to the origin on the click of a button. When clicked, the news agency can trace back the post to its origin which will display all the modifications done on the original content. The news agency can approve or disapprove the post on the click of a button. If the post is disapproved, no further action will be taken. If the post is approved, then the news agency will by default tip an amount of 1 ether to the editor and the journalist. The
necessary contract interactions will be stored in the respective metamask wallets. The posts that have been approved by the news agency will be displayed on the homepage of the crowd auditors. The crowd auditors will also have the feature of tracing the post to display all of its modifications. A warning line for the modified posts will be displayed on their homepage as well. The crowd auditors can report the post as fake or true. Reporting the post as fake will increase the “fake count” of the post and reporting the post as true will increase the “true count” of the post. Additionally, the system will have a threshold limit for the fake count. If the fake count has crossed above the threshold, then the post will be removed from the crowd auditors page and will be added to the list of “Fake Articles.”. There will be another threshold for the number of fake articles a journalist, editor, and news agency can upload. If the threshold is crossed, the respective entity will be blocked from using the system. The “Fake Articles” will be compiled to create a database that can be used in the future for fact-checking other news articles.

B. SYSTEM ARCHITECTURE

The user will use the Metamask wallet to store his ethers, the ganache will be used to implement a local blockchain network. The proposed system will be developed using the truffle framework. The smart contract will be built using the Solidity Programming language. The tests will be written using javascript. The front end will be programmed using React. The entire transaction will be executed by using the truffle framework and ganache local blockchain network. The data will be stored on the IPFS which will return a file hash. This file hash will be stored on the blockchain network.

C. SMART CONTRACT

The UML diagram below shows the functions that the smart contract will have. Also, there will be different structures of events created. The smart contract will contain a structure called “Post” [16], the post will contain the following variables - id: to store the id of the post, article name: to store the article name, content: to store the content of the post, file hash: the hash of the file that has been uploaded on IPFS, fakeCount, trueCount: to store the number of votes, approval status: to store the approval status of the post, address: the address of the author who has written the post, proved: store the id of the previous post, tip amount: to store the tip amount for each post. The smart contract will also contain functions like createPost, tipPost, modifyPost, reportTrue, reportFake, getHash, etc.

The function ‘createPost’ is used to store the post in the blockchain network. It will accept the parameters like articleName, content and filehash. The function will emit an event createPost and the data will be stored on the blockchain network by truffle and ganache. The createPost function can be defined as follows[16]:

---

FIGURE 3. System Architecture
Algorithm 1: Journalist creates a post/news article.

**Input:** articleName is the name of the news article. Content is the content of the news article. Filehash is the hash of the file. Address of the Journalist/Editor.

1. Check whether the articleName is not empty and less than the maximum limit.
2. Check whether the content is not empty and less than the maximum limit.
3. Increment the postCount
4. //create a post with all the initial values for the variables //like prevId = postCount, address = msg.sender, //fakeCount=0, trueCount =0, approvalStatus = false.
   post[postCount]=Post(articleName, content, filehash)
5. //emit the event to create a Post emit postCreated(articleName, content, fileHash)

Algorithm 2: Editor modifies the post.

**Input:** articleName is the name of the news article. Content is the content of the news article. Filehash is the hash of the file.

6. end

The function ‘modifyPost’ is used to modify the post and connect the previous post to the modified one. It will accept parameters like articleName, content and filehash. The function will emit an event modifyPost and the data will be stored on the blockchain network. The modifyPost function can be defined as follows:
that has been submitted to IPFS for storage.

Address of the Editor.

1. Check whether the articleName is not empty and less than the maximum limit.
2. Check whether the content is not empty and less than the maximum limit.
3. Fetch the post with the current id.
curentPost = posts[id]
4. // update the prevId of the current Post
currentPost.prevId = id
5. //increment the id
id++
6. //create a post with all the initial values for the variables
//like address = msg.sender, //fakeCount=0 , trueCount =0, approvalStatus = false.
post[postCount]=Post(articleName, content, filehash)
7. //emit the event to modify a Post
emit modifyPost(articleName, content, fileHash)
8. end

The function ‘tipPost’ is used to tip the Post after approving the post. This function will accept the post id as a parameter and will emit the event ‘postTipped’. The tipPost function will be defined as follows[16]:

Algorithm 3: News Agency tips the post.

Input: id is the id of the news article.
       Address of the News Agency.
       Address of the Editor.
1. Check whether the id is less than the post count i.e is valid.
2. Fetch the post with the current id. currentPost = posts[id]
3. Fetch the post author
   address payable author = currentPost.author
4. Transfer ether from the current account to the author.
   address(author).transfer(msg.value)
5. Increment the tip amount
   currentPost.tipAmount = currentPost.tipAmount + msg.value
6. Emit the event postTipped to store the data on blockchain network
   emit postTipped(variables..)
7. end

The function ‘reportFake’ will report the post as fake and will accept the post id as a parameter. It will emit the event reportFake. The function will be as follows:

Algorithm 3: Crowd Auditor reports the post as fake.

Input: id is the id of the news article.
       Address of the Crowd Auditor
       Address of the News Agency.
1. Check whether the id is less than the post count i.e is valid.
2. Fetch the post with the current id. currentPost = posts[id]
3. Increase the fakecount of the post fakeCount++
4. Check if the fake count is below the threshold limit
   if(fakeCount< threshold ){
      emit reportFake (variables..)
   }
6. Else{
   a. Block the post.
   b. Fetch the post author
      address author =currentPost.author
   c. Check the fakecount of the author and block him if the number of fake posts are high
   d. if(author.fakeCount < threshold) {
      return
   }
   Else
The function ‘reportTrue’ will report the post as true and will accept the post id as a parameter. It will emit the event reportTrue. The function will be as follows:

Algorithm 4: Crowd Auditor reports the post as true.

Input: id is the id of the news article, Address of the Crowd Auditor Address of the News Agency.

1. Check whether the id is less than the post count i.e is valid.
2. Fetch the post with the current id. curentPost = posts[id]
3. Increase the trueCount of the post post.trueCount++
4. Emit reportTrue(variables...)
5. end

All the functions of the smart contract are summarised and with the help of the front end, we can invoke these functions to implement the system.

VI. POST TRACEABILITY

The posts can be traced by the news agency and the crowd auditors to its original post. Let's consider an example to understand this functionality. If a journalist 'A' uploads an article on the network, and the article is viewed by the editor. The editor modified the article file and manipulated the article content in such a way that the post becomes a 'misrepresentation' of the original post. The editor also changes the article name to make it more interesting similar to a clickbait. This article is received by the News Agency and the News Agency ignores the facts and the misrepresentation and approves the post. The post further is received by the Crowd Auditors, the crowd auditors receive a warning that the post has been modified. The crowd auditors can trace the post on a click of a button. This will display a table containing the original post and the original file uploaded by the journalist, the file uploaded by the editor after the modifications, and the address of the respective journalist, Editor, and the News Agency who has modified the post. Hence, after considering the original post and the modified posts, the crowd auditor can report the news as Fake as the modification is misrepresenting the original news content. The journalist's fakeCount will be increased, the posts fakeCount will be increased and also the respective fakeCount's of the editor and news Agency will be increased. The smart contract will further check if these fakeCounts are above the threshold. If they are found to be above the threshold then the respective entity (post, journalist, editor, news agency) will be blocked. Hence, the system will be self-maintaining itself by removing fake articles and blocking the entities that produce fake articles.

VII. LIMITATIONS

The system has certain limitations that can be improved in the future scope. The system is dependant on humans to a certain extent, for eg. if the crowd auditor doesn't do his work properly and reports the post as true, the system cannot remove fake articles and block the entities producing fake news. Hence, the system needs to implement a strong filter to decide which crowd auditors to let in. The current system can only trace the articles that are created in the system, hence, the system is incapable to trace an article from the internet. Maintaining the state of the contract every time a transaction is created is difficult. The number of resources used will be high and adding data to the block will be expensive as a certain gas fee has to be paid for each transaction according to Qayyum A et al (2019)[17]. The system has not yet been deployed on a real blockchain network and hence, the complexities after this deployment are yet unknown.

VIII. FUTURE SCOPE

The current limitations can be overcome by using Big data and Machine Learning. The current systems that are used for fact-checking and fake news detection can be used along with this system prototype to develop a better prototype. This new system can detect fake news over the internet and using big data we can trace the news articles to its origin. The current system prototype will compile the fake articles and create a database of the same. This database can be used for fact-checking in the future[18]. Integrating this prototype with social media, we can implement a functionality to block and trace all the posts to the original post, hence, the spread of fake news through social media can be blocked and curbed.

IX. CONCLUSION

Overall, the system will help the news agencies in producing authentic articles. As these articles can not be tampered with, their authenticity will stay preserved. The system will be transparent and governed by the smart contract. As the system is implemented using blockchain, it has all the advantages of blockchain i.e transparency, immutability, and security. The system will be secure as all the data will be decentralized through IPFS. The system will encourage the journalists, editors and the news agencies to publish authentic articles. Hence, this system
can prove to be a small contribution to curbing the fake news problem

X. ACKNOWLEDGEMENT
We would like to acknowledge our project guide Prof. Nalini Jagtap for guiding us throughout the research process, development and testing of the application.

XI. REFERENCE
[1] Vanderwicken P., Why News is not the Truth, May- June 1995. Accessed on: April 10, 2020 [Online]. Available: https://hbr.org/1995/05/why-the-news-is-not-the-truth
[2] Hernandez J. As China Cracks Down on Coronavirus Coverage, Journalists Fight Back. March 14 2020. Accessed on: April 15, 2020. [Online]. Available: https://www.nytimes.com/2020/03/14/business/media/coronavirus-china-journalists.html
[3] Takyar A. Blockchain in fake news - Transforming News Industry. Accessed on: July 27, 2019. [Online]. Available: https://www.leewayhertz.com/blockchain-fake-news/
[4] Xinyi Zhou, Reza Z.. Fake News: A Survey of Research, Detection Methods, and Opportunities Zhou, Xinyi & Zafarani, Reza, “Fake News: A Survey of Research, Detection Methods, and Opportunities”, arXiv.org. Dec 2018. 1812.00315. Accessed on: Oct 17, 2019.[Online] Available:https://arxiv.org/abs/1812.00315
[5] Wenqian Shang,Mengyu Liu, Weiguo Lin, Minzheng Jia. "Tracing the Source of News Based on Blockchain," 2018 IEE/ACIS 17th International Conference on Computer and Information Science (ICIS), Singapore, 2018, pp. 377-381, doi: 10.1109/ICIS.2018.8466516.
[6] Wee Jing Tee, and Raja K. M.. "Trust Network, Blockchain and Evolution in Social Media to Build Trust and Prevent Fake News," 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), Subang Jaya, Malaysia, 2018, pp. 1-6, doi: 10.1109/ICACCAF.2018.8776822.
[7] Muhammad S.,Ahmad A.,Mohaisen A.. "Fighting Fake News Propagation with Blockchains," 2019 IEEE Conference on Communications and Network Security (CNS), Washington DC, DC, USA, 2019, pp. 1-4, doi: 10.1109/CNS.2019.8802670.
[8] Safi ur Rehman, Dr. Muhammad U.S Khan, Dr. Mazhar Ali. "Blockchain-Based Approach for Proving the Source of Digital Media," 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2020, pp. 1-6, doi: 10.1109/iCoMET48670.2020.9073820.
[9] Zonyin Shae, Jeffrey J.P Tsai. "AI Blockchain Platform for Trusting News," 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), Dallas, TX, USA, 2019, pp. 1610-1619, doi: 10.1109/ICDCS.2019.00160.
[10] Fottrell Q., “American journalism is suffering from ‘truth decay’ — the media have become more biased over the last 30 years, RAND study says”, MarketWatch. May 28, 2019. Accessed on: Nov 9, 2019. [Online]. Available on: https://www.marketwatch.com/story/theres-a-truth-decay-in-american-journalism-study-says-media-has-become-more-biased-over-the-last-30-years-2019-05-15
[11] Jimi S. How does Blockchain work in 7 steps- A clear and Simple explanation. May 6 2018. Accessed on: September 8, 2020. [Online]. Available: https://blog.goodaudience.com/blockchain-for-beginners-what-is-blockchain-519db8c6677a
[12] "Ethereum.org". Accessed on: October 20, 2019. Available: https://ethereum.org/
[13] “Ethereum Blockchain Development with Ethereum and Dapps”. Accessed on: November 7, 2019. [Online] Available: http://myhsts.org/tutorial-learn-about-ethereum-blockchain-development-with-dapps-and-ethereum-vm.php
[14] Curran B. What is IPFS? Inter Planetary File System: Complete Beginner’s Guide. March 3, 2020. Accessed on: March 15, 2020. [Online]. Available on: https://blockonomi.com/interplanetary-file-system/
[15] Khaled S., Jayaraman R., Nizamuddin N., Mohammed A Omar. "Blockchain-Based Soybean Traceability in Agricultural Supply Chain,” in IEEE Access, vol. 7, pp. 73295-73305, 2019, doi: 10.1109/ACCESS.2019.2918000.
[16] McCubbin G. Blockchain Tutorial for Developers: Step by Step guide. April 8, 2020. Accessed on: May 10, 2020. [Online] Available on: https://www.dappuniversity.com/articles/blockchain-tutorial
[17] Qayyum A. Qadir J., Muhammad U. J. and Sher F., "Using Blockchain to Rein in the New Post-Truth World and Check the Spread of Fake News," in IT Professional, vol. 21, no. 4, pp. 16-24, 1 July-Aug. 2019, doi: 10.1109/MITP.2019.2910503.
[18] Thorne J., Vlachos A. Automated Fact Checking: Task formulations, methods and future directions. Sept 5, 2018. Accessed on: October 23, 2019. [Online]. Available on: https://arxiv.org/pdf/1806.07687.pdf