Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 44 (2023), No. 3, pp. 515–527
DOI : 10.47974/JIOS-1411

FakeSpotter: A blockchain-based trustworthy idea for fake news detection in social media

Sakshi Kalra *
Yash Bansal
Yashvardhan Sharma
Department of CSIS
BITS Pilani, Pilani Campus
Pilani 333031
Rajasthan
India

Gajendra Singh Chauhan
Department of HSS
BITS Pilani, Pilani Campus
Pilani 333031
Rajasthan
India

Abstract

Social media encourages information sharing without a physical barrier making it the perfect platform for learning and communication. In the meantime, it acts as a means of quickly disseminating misleading information. Researchers are battling fake news using strategies like detection, verification, mitigation, and analysis because of significant social concerns. It can be hard to tell the difference between true and false information. In the area of knowledge verification, various machine and deep learning-based approaches have been used to identify false data. However, there are some drawbacks of using AI-powered technologies, including data dependency, security concerns when applying AI-powered methods in the real world, and gaining user trust. In order to address the issues with AI-powered technologies, a blockchain-based idea (FakeSpotter) is put forth in this work. We offer an idea i.e.based on blockchain that utilizes crowdsourcing to determine whether or not content is fake. We attempt to use Blockchain technology's features correctly and completely to create a secure system with no authoritative control over information dissemination. In

* *E-mail:* p20180437@pilani.bits-pilani.ac.in

© Taru Publications

this attempt, we aim to build a system that is not reliant on pre-defined datasets and discuss the initiatives taken in the fight against disinformation.

Subject Classification: *68T50.*

Keywords: *Blockchain, Ethereum, Smart contracts, Crowd annotators, Fake news detection, Social media analytics.*

1. Introduction

Today, we spend a growing amount of time on social networking platforms. On the one hand, these channels make getting and sharing news easier, but they are also often used to spread false information. The term "fake news" refers to information that is purposefully false or misleading and is disseminated via conventional print media and online social media. It is made and spread with the goal of misleading people to get a financial or political advantage. False news that is meant to confuse people is not the same as satire or parody, which are meant to make people laugh instead of confuse them. Fake news often uses catchy headlines or completely made-up stories to get people to read it, share it online, and click on it on the Internet. In the second case, it is similar to "clickbait" headlines on the internet in that it depends on the advertising money it brings in, no matter how true the information is. It can potentially affect how people make economic and political decisions significantly.

1.1 *Initiatives in the Fight Against Disinformation*

In today's society, the proliferation of fake news on the internet poses a serious concern. Individuals increasingly rely on digital channels to exchange information and get news. They frequently lack the ability to distinguish between authentic and fake news. Helping people spot fake news is increasingly important for preventing harmful effects on public opinion and people's decisions. One of the most commonly cited examples of how fake news may sway public opinion is the 2016 US presidential election. Numerous studies, such as the one in [1] show how fake news dramatically affects political discourse and election outcomes.

Evaluating the validity of many news sources in the current context is challenging. Anybody cannot verify true news. There is a chance that the organization will be biased and may have a preset agenda if this duty is handed to a group. Several international organizations have taken the following actions to stop the dissemination of false information: Facebook

has added a feature to the site that allows users to report false information. To avoid losing the public's trust, they have clarified through public announcements that they are looking into the capability to identify fake articles using an automated method [2]. Similarly, to this, centralized organizations like Meta and Google can decide how to manage the activities on these platforms, including by deleting user accounts and removing posts without warning or giving any justification.

1.2 *Technical Initiatives in the Fight Against Disinformation*

From a technical point of view, it is interesting that many internet tools have been made to find fake news. Tech and media companies are driving the development of this technology to fight against false information. Facebook is one example, which uses its multimodal content analysis tools and 35,000 human executives to combat fake news. This set of AI-driven technologies has been used to find offensive or fake information about the coronavirus. Facebook says that this supervised classifier-based method gets around the problems of similar technologies when dealing with images that have been changed using common adversarial techniques. Aside from these big companies, other companies, like DeepTrace in the Netherlands, are also working on anti-fake technology. This company's mission is to develop a deep-fake antivirus.

1.3 *Blockchain: A Secured Solution to Overcome the AI-based Problems*

Artificial Intelligence has given us some of the best tools for spotting fake information, but it can also construct false information for social networking websites. Unfortunately, AI algorithms have lately been less effective at detecting fake news than they are at producing it [3]. Also, AI-based techniques can help stop the spread of false information, but they aren't foolproof. As a result, it might occasionally commit a significant number of false positives and false negatives. Because of this, it's important to find a different, strong, and effective way to deal with this problem. Creating blockchain-based solutions has the ability to alter how information is created and shared. Blockchain can help build a trustworthy social system by making transactions trustworthy, unchangeable, reliable, verifiable, and clear. In Peer2Peer communication, blockchain technology can be used to redesign social networking platforms so that users again have control over shared posts. A system like this could help halt the dissemination of false information while increasing trust in the system itself. We aim to propose an idea that will provide a trustworthy platform

that is by the people, of the people, and for the people, using the essential properties of democracy to create a trusted social network.

2. Related Work

Researchers have been working for years on ways to use artificial intelligence to look at the content and context of data posted by users. The use of blockchain-based technologies for FND has only been attempted a very small number of times.

2.1 *Analysis Using Artificial Intelligence-Powered Technologies*

The most recent developments in this digital revolution are artificial intelligence (AI) [5] and machine learning (ML) [6]. It is used to mimic human behavior. An excellent example of this use is the work of [6], which examined the performance of classifying Twitter misinformation using five base classifiers (Linear SVM, RF, LR, NB, and KNN) and compared it to the techniques of TF and TF-IDF feature extraction with N-grams of various lengths. A study was conducted by [6] in which they assessed if a Twitter account was human or a bot and its gender. A linear SVM classifier was created to evaluate words and character grams. A stochastic gradient descent classifier was employed to examine tweets, words, and character grams. English and Spanish tweets were used as data. The results showed that bot detection was 0.906 for English bots and 0.856 for Spanish bots. Deep learning (DL) [7], a new idea in the research world, has demonstrated to be better at identifying fake news than traditional machine learning (ML) techniques. DL's advantages over ML include automatic feature extraction and little data preprocessing. But the Neural Network's lack of transparency in decision-making conceals its credibility.

2.2 *Analysis Using Blockchain-Powered Technologies*

There is a dearth of research on applying blockchain technology to address social networking security issues. Only a few small projects have tried to use blockchain to deal with the problem of false information specifically. On the other hand, the literature didn't show many blockchain-based ways to deal with specific problems in social networks. A Blockchain method called U-share was developed by [8] to let users track and manage all of the uploaded content they share on social media. It could be thought of as a permission blockchain because people in the network need to be named for shared data to be consistent, accountable, and traceable. It was

suggested that a group of third parties verify each new user joining the network and give them a unique name.

A novel social network model called Tweet-chain by [9] is based on Blockchain and uses the Proof-of-Concept protocol to manage public posts to ensure social transaction security rather than proof of work. You can use it without making any assumptions about the social network provider's reliability. A Blockchain-based social media notarization service is presented by [10]. The main idea is that the official service providers should be the ones to sign the authenticated data with their private keys. Users hardly have a chance to avoid taking ownership of their content under this system. This will stop people from posting fake content on social media in the long run. Fakechain is another blockchain consensus system created by [11] that uses data mining algorithms to authenticate shared information. It involves using a centralized blockchain where news sources are full nodes and are in charge of reading, writing, and mining, while regular users are light clients who only have access to the information in the blockchain. The idea is to re-engineer social networks as decentralized networks where users are represented as peers. It employs a variety of metrics to assess the trustworthiness of online newsworthy information sources[12].

3. Blockchain Trilemma

The blockchain is just a group of interconnected blocks. The "blocks" are made up of digital data that is saved in a public database. In the world of cryptocurrencies, blocks store information about participants and transaction data, like the date, time, and amount of the most recent purchase. Blocks use a distinct hash code to distinguish information rather than duplicate it. Anyone can set up a node on a blockchain that copies every piece of information that the other nodes must concur upon. So, the other users and the people who made the application are getting paid. This keeps the user's information private and makes it possible for the apps to be really shared. In short, a public ledger is decentralized and doesn't have a single point of control. These characteristics of blockchain are used to make systems that people can trust. Blockchain are frequently compelled to make trade-offs that prohibit them from fulfilling all three of the following goals[1]: **Decentralized:** Created a blockchain system without

[1] https://medium.com/certik/the-Blockchain-trilemma-decentralized-scalable-and-secure-e9d8c41a87b3

relying on a single point of control. **Scalability:** A blockchain system's capacity to manage an increasing number of transactions. **Security:** The capacity of the Blockchain system to function as intended and defend itself against assaults, errors, and other unforeseen issues. Decentralized systems are important because they let anyone use and make things on the platform without asking for permission. This is one of the most critical parts of a trust-based FND system, but it is also one that is frequently ignored in other research.

4. Proposed Idea (Blockchain Based Fake News Detection)

4.1 *Different Blockchain Users*

In our proposed idea i.e. based on blockchain, there are five different types of users: News Agencies, Journalists, Editors, Crowd Annotators, and Expert Annotators, as shown in Figure 1.

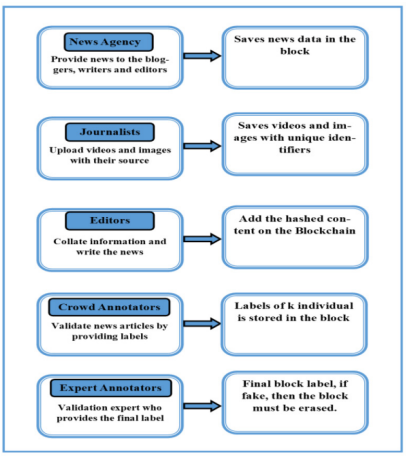


Figure 1

Different Blockchain Users

All of these users must register on a portal using their individual data, which is hashed and kept in a block. The news agencies will set up an account on the portal with their name, address, news organization license, and domain names. Once the news is saved on the blockchain, these agencies can send it to bloggers, writers, and publishers. A journalist is a person who takes news from news agencies and adds pictures and videos to it. The content is then hashed and pushed into the blockchain. The block includes details about the added media source. The journalists

must also sign up on the portal by giving their name, email address, phone number, and proof that they are allowed to work. Before the news is published, it is edited by an editor who gets it from journalists and news organizations. They have pushed a block that includes the URL of the website and the actual news that was posted there. The editors must sign up for the portal with their name, email address, phone number, job title, and license. The crowd annotators provide labels to the news as to whether it is fake or not. The news block that the editor creates is sent to at least k crowd annotators. We receive labels from the editors and add them to the block. Crowd annotators have names, email ids, bank information, phone numbers, and credit scores, among other things. Some authority figure is needed to build trust because it ensures that things can be checked and that people are responsible and answerable. At the same time, people should not feel alienated from the system and need the power to hold the authorities accountable for their actions. Hence, we propose an idea where these authorities are defined as "experts" who can add supporting documents or contradicting statements to existing news items. These experts can potentially be fact-checking websites or verified social media influencers. The expert annotators sign up on the portal with their name, email address, phone number, and bank information. They are the ones who provide the final label for each news block. In our proposed idea, we tried to figure out how to find fake news by modeling it as a binary classification problem with only two possible answers: real and fake.

4.2 *Labeling Algorithm*

The proposed idea has developed a solution in which experts from different fields on the blockchain based platform give their votes for any news article. Each expert has to provide certain evidence, proof, or justification for their votes, as shown in Figure 2. The identities of these experts are kept secret, but each user can see their votes and the reasons why they voted the way they did. Experts, on the other hand, have both read and write permission. Bloom filters can be implemented to efficiently calculate whether a user trying to cast a vote is an expert in the domain or not. The crowd annotators and expert annotators are the ones who provide the labels for the news blocks. Once the editors have added the news block, it will be reviewed by at least k crowd annotators. In our further discussion, we use the value of k as 10. Let there be n crowd annotators, from i = 1 to n. We randomly select k crowd annotators and ask them to flag the news block as real or fake. All the crowd annotators have a credibility score

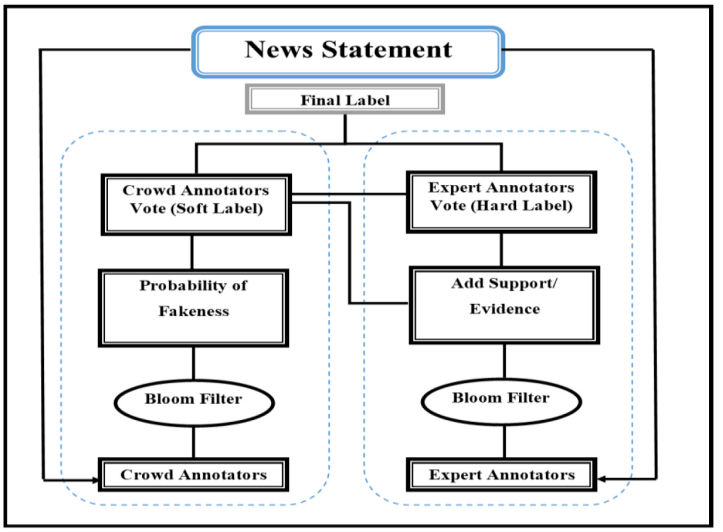


Figure 2

Labeling Algorithm

attached to them. We initialize the credibility score for each annotator as 5. The label values are 0 for real and 1 for fake. We define 2 types of final labels, which we obtain for each news block. The soft label is the label that is obtained by taking into consideration the labels obtained by k crowd annotators. The probability calculated is given by equation 1:

$$P(\text{Fake} \mid \text{Crowd Annotators Labels}) = \sum_{i=1}^{k} (c_i * \text{Label}_i) / \sum c_i \qquad (1)$$

If [P (Fake| Crowd Annotators Labels) > 0.5], we label the news as fake; otherwise, the overall label is 1. We make the soft label because it quickly gives us a good idea of the correct label. The label generated by the expert is the actual label or the hard label. Since the hard label might not be available for a while, the soft label gives a good idea of the final label and can be used to flag content. The credibility score of the crowd annotators also goes up or down based on how accurate the label they gave was and whether or not it was compared to the label the expert got. In the event that the label given by the crowd annotator matches that of the expert, his credibility increases by a factor of 0.05. In the case of a mismatch, the credibility score decreases by a factor of 0.05. If the credibility score exceeds 9, the annotator is now considered an expert and will be paid for flagging fake content. So a crowd annotator has to get at least 80 correct labels with no incorrect labels to become an expert.

5. Various Essential Platforms Involved in the Proposed Blockchain Based Idea

The main goal of the proposed idea is to give each crowd annotator a rating and set it to a certain value at the start. A weighted voting system will use these ratings as weights to decide whether or not a news claim is true. Given any news statement that enters the system, users are allowed to vote on it, and based on the final verdict, the ratings of each user that voted are affected. Establishment of the final label of the news statement requires a decision function. When setting up such a function, there are a number of things to think about and challenges to overcome. Figure 3 displays the essential platforms that can be involved in the proposed blockchain based idea.

Ganache Ethereum Platform: In this suggested idea, the Ethereum chain may be used as the essential platform for using blockchain technology. Each user on the network has an associated private key, public key, and address. The private key tells a cryptocurrency user that the funds in a certain address belong to them. The address, which can be used in transactions, is a shorter, representative form of the public key. Ethereum resembles a large, slow, but incredibly safe computer. Ethereum is decentralized because thousands of nodes check the execution of each function. This also makes its data unchangeable and hard to censor. When you run a function, every node on the network has to run the same function to ensure its output is correct. So, it's important to remember that a decentralized app's presence on Ethereum does not always mean it is decentralized. Instead, you should look at the whole source code to make sure that it is not limited in any way by the owner that might look suspicious. As a developer, you need to find the right balance between

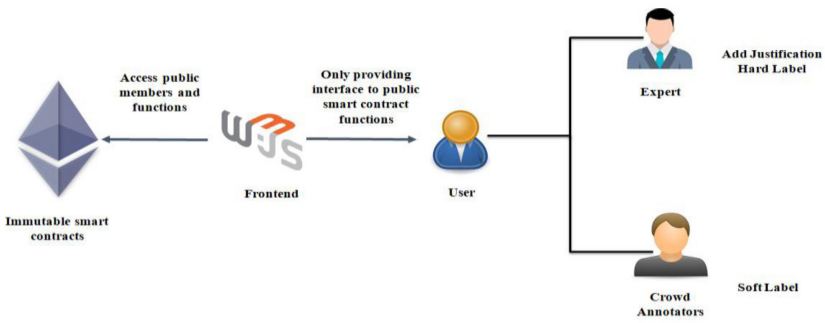


Figure 3

Essential Platforms Involved in the Implementation

keeping control of a decentralized app to fix any problems and making a platform without an owner that your users can trust to keep their data safe. The Ganache Ethereum blockchain emulation platform may be utilized for the proposed idea, which lets you engage with smart contracts on your own private blockchain. You may use it to create your own Ethereum blockchain and run tests on your Solidity contracts. It enables sophisticated mining automation and shows blockchain log data.

Smart Contracts: Smart contracts are simply the algorithms that are kept on a blockchain and run when certain criteria are satisfied. They are often used to automate the way an agreement is carried out so that everyone is sure of the result right away and no time is wasted waiting for a third party to decide. They can also set up a workflow to run automatically when certain conditions are met. Smart contracts need important decision-making functions to ensure the code, security, and decentralization can't be changed. But scaling is hard because having thousands of computers validate huge function outputs costs a lot of money. In the proposed idea, many functions related to users have been added, such as letting both users and experts make votes and letting experts explain why they voted the way they did. But there are some challenges related to the implementation of smart contracts: In an Ethereum smart contract, functions must be called by a user; they can't be called automatically when a condition is met (there must be a private key). The decision function should be set up, so we never get stuck in a loop that goes on forever. The system should not be vulnerable to attack or failure by multiple fake accounts.

6. Benefits of the Proposed Blockchain Based Idea (FakeSpotter)

Numerous advantages come with the use of a blockchain-based strategy which are listed below:

Immutability: After being added to the network, it offers content immutability. As the hash code will change if the editors try to change the media, it will be obvious that it has been tampered with. This eliminates the problem of journalists tampering with the media. Due to immutability, a person can't change their credibility scores to become an expert because doing so would be financially beneficial.

Determines the Best Real-World Fake News Detection Method: Every news piece of content added to the block now has a source, which makes it easier to check news articles. Crowd annotators may use analytical models that use deep learning or machine learning based techniques or do

content-based research to determine whether a news article is real. So, by using this method, we can actually find out which method works best in the real world to spot fake news.

Offence Free Network: Once it has been confirmed that the news article is fake, the point at which fake content was added can be determined. The offender is given a warning at this point and maybe kicked off the network if the offense is committed a certain number of times in a row. So, the blockchain network makes data unchangeable and easy to track. This brings us to the point in the process where data is changed, and accountability is needed.

Way to Dataset Generation: The system also gives you a unique and complete set of data that is created in real-time. This dataset is updated all the time, and since we can add news from a wide range of domains, we can build a large dataset with a wide range of news types that are currently needed for deep learning-based models to detect fake news effectively. This way of finding fake news will also make crowd annotators better at what they do and can be used to train experts. This system can also be made easier so that people can upload content from the web that they don't trust and have it checked by both the crowd and experts.

7. Conclusion and Future Work

In this work, a blockchain-based idea for using crowdsourced signals to spot fake news has been proposed. The solution rewards correctly flagging content with the opportunity to become a paid expert annotator. In this proposed idea, the crowd annotators can flag the content by using the expert annotator's knowledge of why the content is important and the supporting documents. With online training, the model will reach a point where the user has a credibility score of more than 9 and can be thought of as an expert. The system will also help more people learn how to spot fake news, which is currently a niche skill with very few people who know how to do it well. The system mayprovide an original, extensive dataset that is generated in real-time. The next step would be to add the benefits of Natural Language Processing based techniques so that users can find articles about the given statement that help them make an informed choice before voting. The incorporation of Deep Learning based models to give users an auto-generated probability of fakeness. Crowd annotators may use analytical models that use deep learning or machine learning techniques or do content-based research to determine whether a news

story is real. The next idea might be to give more weight to expert validators who are experts in the field of news statements.

References

[1] D. J. Flynn, B. Nyhan, and J. Reifler, "The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics," *Polit. Psychol.*, vol. 38, pp. 127–150 (2017), doi: 10.1111/pops.12394.

[2] A. C. Jurno, "Facebook out of control? A technically-mediated autonomous technology," *Rev. Bras. Inovação*, vol. 18, no. 2, pp. 363–386 (2019), doi: 10.20396/rbi.v18i2.8653696.

[3] J. Paschen, "Investigating the emotional appeal of fake news using artificial intelligence and human contributions," *J. Prod. Brand Manag.*, vol. 29, no. 2, pp. 223–233 (2020), doi: 10.1108/JPBM-12-2018-2179.

[4] V. R. Suri, B., Taneja, S., Aggarwal, S., & Sharma, "Fake news detection tool (FNDT): Shield against sentimental deception," *J. Inf. Optim. Sci.*, vol. 41, no. 6, pp. 1513–1524 (2020), [Online]. Available: https://doi.org/10.1080/02522667.2020.1802125.

[5] S. K. Malhotra, P., & Malik, "Fake news detection using supervised machine learning techniques.," *J. Inf. Optim. Sci.*, vol. 43, no. 1, pp. 7–15 (2022).

[6] A. Giachanou and B. Ghanem, "Bot and gender detection using textual and stylistic information notebook for PAN at CLEF 2019," *CEUR Workshop Proc.*, vol. 2380, no. September, pp. 9–12 (2019).

[7] K. N. Patil, S. P., & Jariwala, "Design of a machine learning layer to optimize forgery detection in videos," *J. Interdiscip. Math.*, vol. 23, no. 1, pp. 43–52 (2020), [Online]. Available: https://doi.org/10.1080/09720502.2020.1721653

[8] A. Chakravorty and C. Rong, "Ushare: User controlled social media based on blockchain," *Proc. 11th Int. Conf. Ubiquitous Inf. Manag. Commun. IMCOM 2017* (2017), doi: 10.1145/3022227.3022325.

[9] F. Buccafurri, G. Lax, S. Nicolazzo, and A. Nocera, "Tweetchain: An alternative to blockchain for crowd-based applications," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bio-*

informatics), vol. 10360 LNCS, pp. 386–393 (2017), doi: 10.1007/978-3-319-60131-1_24.

[10] G. Song, S. Kim, H. Hwang, and K. Lee, "Blockchain-based Notarization for Social Media," *2019 IEEE Int. Conf. Consum. Electron. ICCE 2019*, pp. 4–5 (2019), doi: 10.1109/ICCE.2019.8661978.

[11] I. S. Ochoa, G. de Mello, L. A. Silva, A. J. P. Gomes, A. M. R. Fernandes, and V. R. Q. Leithardt, FakeChain: A Blockchain Architecture to Ensure Trust in Social Media Networks, vol. 1010, no. January 2022. Springer International Publishing (2019). doi: 10.1007/978-3-030-29238-6_8.

[12] M. Torky, E. Nabil, and W. Said, "Proof of credibility: A blockchain approach for detecting and blocking fake news in social networks," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 12, pp. 321–327 (2019), doi: 10.14569/ijacsa.2019.0101243.