How to Eliminate Fake News
: Utilizing a Blockchain-based Smart Contract

Ga-Young Jang[1]

and

Hyoung-Goo Kang[2]

Abstract

This paper introduces an algorithmic model designed to combat the spread of fake news, which poses significant threats to journalism, democracy, the economy, and social welfare. The model posits that both news providers and readers play a role in the creation and dissemination of fake news, driven by incentives to maximize utility and financial gains. Using a systematic seven-step framework, the proposed algorithm integrates lump-sum transfers and an incentive-based payment scheme to promote the dissemination of truthful information while discouraging the propagation of misinformation. Furthermore, the study demonstrates how this algorithm can be implemented using blockchain-based smart contracts, capitalizing on blockchain's core attributes of irreversibility, transparency, and traceability. This research offers meaningful insights into the complex effects of fake news and presents a practical technological solution to mitigate its impact.

Keywords: algorithm, blockchain, fake news, technology

[1] California Miramar University; 3550 Camino Del Rio #208, San Diego, CA 92108; kathy.g.jang@susfinresearch.com; 1-858-209-3515

[2] Corresponding author. Hanyang University Business School; 222 Wangsimni-ro, Seongdong-gu, Seoul, Korea 04763; hyoungkang@hanyang.ac.kr; 82-2-2220-2883

INTRODUCTION

In today's information-saturated world, individuals are inundated with data. Each minute witnesses the exchange of 98,000 tweets, 160 million emails, and the upload of 600 videos to YouTube (Marr, 2018). Moreover, the ease of self-publishing has led countless individuals to share their opinions on blogs and other media platforms. This surge in information volume makes it increasingly challenging for individuals and companies to discern facts from fiction as they strive to obtain authentic information to maximize utility and financial returns.

In 2025, the Davos survey identified misinformation and disinformation as the top-ranked risk over a two-year horizon and the fifth-ranked risk over a ten-year horizon. These findings underscore the significant concern surrounding not only the prevalence of fake news but also its far-reaching and often unquantifiable impact. Evidence of the pervasive effects of fake news is readily apparent. For example, on November 10, 2022, a fake tweet from a parody Eli Lilly account falsely announced the provision of free insulin. Although Eli Lilly swiftly issued a denial, the misinformation triggered a market reaction, causing the company's stock to plummet and wiping approximately $22 billion off its market value the next day. Similarly, during the 2016 U.S. presidential election, fake news stories about candidates proliferated across social media platforms. Despite being debunked afterward, these stories significantly shaped public opinion and political discourse (Allcott & Gentzkow, 2017).

Some people blame technology, such as the internet, for rapidly disseminating fake news and amplifying the sheer volume of false information available. Indeed, there is merit to this claim, as the swift spread of information—whether true or false—would not be feasible without such technological advancements. Concurrently, others are leveraging advanced technologies, including artificial intelligence and machine learning, to combat misinformation (Cassauwers, 2019). Regardless of the approach, the consensus is clear: fake news is harmful to society, with its detrimental effects extending beyond politics, as previously discussed, to areas such as the economy, social welfare, and the financial sector (Bakir & McStay, 2018; Kogan, Moskowitz, & Niessner, 2023; Vosoughi, Roy, & Aral, 2018). Addressing the problem requires a collaborative effort, including improved media literacy, regulatory measures, and better misinformation detection tools.

Then, a critical question follows: can we utilize technology to eliminate fake news? Extant studies explore various approaches, including algorithmic methods (Shu, Sliva, Wang, Tang, & Liu, 2017), adherence to codes of conduct (Painter & Hodges, 2010), the role of fact-checking institutions (Allcott & Gentzkow, 2017), as well as advancements in machine learning (Figueira & Oliveira, 2017; Conroy, Rubin, & Chen, 2015; Rubin, Conroy, Chen, & Cornwell, 2016) and deep learning algorithms (Ruchansky, Seo, & Liu, 2017). However, the existing literature often overlooks the combined importance of cross-verification mechanisms and the alignment of incentives for both news providers and consumers, such as journalists (i.e., news producers) and readers.

In particular, providing appropriate incentives is essential to encourage specific behaviors, such as truth-telling. Algorithms that fail to account for these incentives are likely to be ineffective. If individuals can discern how a fake news detection algorithm functions, they may manipulate their behavior to bypass it. Therefore, the development of algorithmic solutions must include robust incentive mechanisms to motivate journalists and readers to voluntarily prioritize and share truthful information. Despite its importance, this aspect remains insufficiently explored in the existing literature.

To address this gap in the literature, we propose an algorithm designed to eliminate fake news while incorporating mechanisms to provide appropriate incentives for journalists and readers. The algorithm follows a seven-step process, beginning with the categorization of news articles by topics, tone, and viewpoints, followed by an assessment of the content's veracity. It then facilitates payments and imposes lump-sum taxes to encourage truth-telling. This systematic approach enables the platform to compare and cross-check news articles and associated data, implementing incentive compatibility constraints to promote adherence to truthful reporting.

Additionally, this study presents an innovative solution that employs an algorithm based on blockchain technology. We refer to this as a blockchain-based smart contract designed to incentivize all stakeholders to understand the algorithm and to act truthfully. In our proposed solution, stakeholders would voluntarily deter the spread of fake news and promote the dissemination of authentic news. Achieving such consensus among stakeholders is crucial in any mechanism against fake news, as it aligns with the principles of freedom of speech. The smart contract leverages the inherent advantages of blockchain systems, namely transparency, traceability, and irreversibility. This paper demonstrates how to harness these benefits to implement an effective algorithm to combat fake news.

The contributions of this paper to the literature are several. First, it enhances the discourse on algorithmic methodologies for addressing a critical socio-political issue with far-reaching implications for the economic and financial well-being of both individuals and institutions. Second, by proposing a novel solution to mitigate pressing challenges such as misinformation, which threatens journalism, democratic processes, financial stability, and societal cohesion, this work adds valuable insights to the blockchain research domain. Lastly, The algorithm introduced in this study offers broader utility across fields where ensuring truth and authenticity is of paramount importance, thereby extending its relevance beyond the immediate context of this research.

The remainder of the paper is structured as follows: Section II introduces the background while Section III describes the algorithmic model. Section IV explains the applications of the model and the smart contract. Section V discusses the implications, limitations, and suggestions for future research and Section VI concludes.

BACKGROUND

Multifaceted Effects of Fake News

The effects of fake news extend beyond mere misinformation; they pose a significant threat to the politics, societal values, economic and financial welfare, and democracy. For example, fake news can erode public trust in the media and other critical institutions, leading to widespread cynicism and disengagement from civic duties. Fake news can also influence voters by spreading false information about candidates, policies, or voting processes. When individuals can no longer distinguish between credible news sources and fabrications, they may become skeptical of all information, undermining informed decision-making processes that are fundamental to a functioning democracy (Lazer et al., 2018).

Moreover, fake news often exploits and exacerbates societal divisions by reinforcing existing biases and spreading inflammatory content. This can lead to increased polarization, where different segments of society become more entrenched in their beliefs and less willing to engage in constructive dialogue (Sunstein, 2001). Such division can weaken social cohesion and create an environment where extremist views flourish, potentially leading to real-world conflicts (Benkler, Faris, & Roberts, 2018; Vosoughi, Roy, & Aral, 2018).

In addition, the economic and financial implications of fake news are profound. Businesses and financial markets rely on accurate information to make decisions. Misinformation can distort market behavior, lead to poor investment decisions, and cause significant financial losses (Gupta, Lamba, Kumaraguru, & Joshi, 2020). For instance, a fake news article about a major company's financial health can lead to stock market volatility, affecting investors and the broader economy (Arcuri, Gandolfi, & Russo, 2023; Clarke, Chen, Du, & Hu, 2020; Siering, Clapham, Engel, & Gomber, 2017). Fake news also fosters fraudulent schemes, such as fake investment opportunities or Ponzi schemes, resulting in substantial financial losses for individuals (Kogan et al., 2023).

Last but not least, the propagation of fake news poses a threat to public safety and social welfare. During the COVID-19 pandemic, false information about the virus, its treatment, and vaccines led to widespread confusion and vaccine hesitancy (Loomba et al., 2021). This misinformation not only undermined public health efforts but also resulted in preventable illnesses and deaths (Brennen et al., 2020). Fake news can also fuel discrimination, hate, and social unrest. False narratives about racial or religious groups can lead to violence and hate crimes, undermining social harmony and safety (Guess, Nyhan, & Reifler, 2018).

Overall, the concern over fake news is not just about the integrity of information but about the broader impacts on trust, social cohesion, economic stability, and even public health. As such, addressing the proliferation of fake news is crucial for upholding our societal values and ensuring the well-being of our communities.

Challenges in Addressing Fake News

Despite the widespread and multifaceted impacts of fake news, distinguishing or eliminating fake news from authentic news remains a formidable challenge for numerous reasons. First, false information or propaganda is often presented as genuine news. Second, fake news is a social construct, and identifying it is inherently contentious. Assessments of whether an article contains misinformation frequently depend on the socio-political perspectives of the audience. Third, producers of news, such as journalists, independent writers, and individuals on social media, may strategically create fake news, making it difficult to detect. Fourth, readers may consume, believe, and further disseminate fake news (Lee, 2013; Allcott & Gentzkow, 2017).

In light of these challenges, our research proposes an algorithmic solution to eliminate fake news by (i) effectively detecting it and (ii) addressing the incentive mechanisms that serve as key drivers behind its production.

Model

Description of the Algorithm

The algorithm operates through a structured process designed to maximize the dissemination of truthful information and minimize the spread of fake news. Table 1 illustrates the process.

###Insert Table 1 about here###

Our proposed algorithm is a comprehensive approach to eliminating fake news by integrating systematic evaluation, incentive alignment, and enforcement mechanisms. It emphasizes collaborative verification, involving journalists, readers, and editorial oversight to assess the truthfulness of news content. The process not only ensures rigorous truth-checking through cross-cluster assessments but also incorporates financial incentives and penalties to promote accuracy and deter manipulation. By addressing the dual challenges of detecting fake news and motivating truthful reporting, the algorithm aims to create a sustainable ecosystem where truthful information is prioritized and rewarded, while misinformation is systematically minimized and penalized.

The Assumption of Trust

Throughout the seven-step process, trust is the most crucial element. Both readers and news clusters must trust that the editor will strictly adhere to the established rules for news exposure and resource allocation. However, trust in traditional media and institutions has been eroding due to the prevalence of fake news (Cassauwers, 2019).

To address this issue, we propose the implementation of a blockchain-based smart contract. A blockchain-based smart contract can be trusted to enforce the pre-arranged mechanism due to its characteristics of irreversibility, traceability, and transparency (Nakamoto, 2008). While it is naive to expect technology alone to solve the problem of fake news, it can significantly enhance transparency around false claims and misinformation.

Furthermore, combining cryptocurrency with the smart contract allows the editor to reduce transaction costs associated with the potential complexity of the allocation mechanism, as well as with the possible frequent, albeit small, transactions (Delmolino, Arnett, Kosba, Miller, & Shi, 2016). The discussion section will further elaborate on the detailed application of the blockchain-based smart contract.

Model Specification

Consider a topic with N distinct views or interpretations. Each view forms a separate news cluster, meaning there are N news clusters related to the topic. Journalists within the same cluster (i.e., a set of news articles with similar views) produce comparable tones or interpretations. Thus,

N also represents the number of different tones or perspectives on the topic. The information derived from the N news clusters typically forms panel data, as these clusters may include past articles if journalists cover the topic over an extended period.

A news cluster may be motivated to produce fake news to promote its particular interpretation of the topic. Each news cluster is aware of the type (r_i) of its articles and the extent to which the content is genuine or fabricated. To move beyond a binary classification of news (genuine versus fake), we define the type r_i as the amount of truth in an article generated by the i-th news cluster. This approach allows for a more realistic representation, as the tone and interpretation of articles within the same topic can vary significantly across clusters.

A news cluster must be selective in presenting facts and opinions, as it is often impractical to include every relevant piece of data in an article, particularly on controversial topics. Journalists must choose the facts and opinions they deem most important. As Francesco Nucci, the applications research director at the Engineering Group in Italy, stated, "Stories are often not entirely true, but also not entirely false" (Cassauwers, 2019). Consequently, the amount of truth, or the "type" of news, varies across different articles.

Given the existence of N news clusters, the types of news across these clusters can be represented by an N-dimensional vector. For simplicity, we do not distinguish between the types of news and the clusters themselves in this model. Hence,

$$\vec{r} : N - dimensional\ vector\ of\ true\ types\ of\ news$$

$$r_i \in \vec{r} : True\ type\ of\ i^{th}\ news\ cluster.$$

A news cluster then reports its content type to an editor who manages a news platform (e.g., a newspaper, media channel, or other outlets). For instance, a journalist informs the editor about their topic and specifies the degree of truth in their article. We use an N-dimensional vector that encapsulates the reported types to represent this information. Hence,

$$\tilde{r}_i \in \tilde{r} : \text{Reported type of } i^{th} \text{ news cluster}.$$

On a news platform, an editor determines the exposure level of an article reported by news cluster i, denoted by w_i, which represents the amount of exposure granted to that cluster. Additionally, the editor may impose a lump-sum tax m_i, which signifies a payment made by journalists to the editor for the news cluster i. Based on these considerations, the following notations can be established. Therefore,

$$w_i \in \vec{w} : N - \text{dimensional vector of exposures on each news cluster}$$

$$m_i \in \vec{m} : N - \text{dimensional vector of lump-sum tax for news clusters}.$$

A news cluster has preferences regarding the exposure levels of all news clusters and the lump-sum taxes imposed. For instance, a particular news cluster may assign positive weights to arguments that align with its views while assigning negative weights to opposing perspectives. Since each of the N clusters has relative preferences, the entire preference structure for a topic can be represented by an $N \times N$ matrix, denoted as u. The notations are as follows:

$$\vec{u}_i^T \vec{w} - m_i : \text{Utility of } i^{th} \text{ news cluster}$$

$$u_{j,i} \in \vec{u}_i : \text{Utility coefficient of } i^{th} \text{ news cluster on } j^{th} \text{ news cluster}$$

$$\vec{u}_{-i} : i^{th} \text{ element is set to zero in } \vec{u}_i$$

$$\vec{u}_i \in \boldsymbol{u} : N * N \text{ utility matrix}.$$

Then, the view of a news platform becomes

$$r_{view} = \vec{w}^T \vec{r}.$$

Assuming the editor is honest, their goal would be to ensure that the news exposed reflects as much truth as possible (i.e., the editor increases w_i if $r_i > 0$) while minimizing other dimensions of the news. On the other hand, each news cluster seeks to maximize the exposure of viewpoints that align with its own while minimizing the taxes imposed. Consequently, this scenario involves information asymmetry and a conflict of interest between the editor and the news clusters. This dynamic can be represented through an optimization problem from the perspective of the editor as follows:

$$\max_{\vec{w},\vec{m}} : \mathrm{E}[\vec{w}^T \vec{r} | \tilde{r}^*] \text{ [Goal Function (GF)]}$$

$$\text{Subject to}: \tilde{r}_i^* = \arg\max_{\tilde{r}_i} : \mathrm{E}\left[\vec{u}_i^T \vec{w} - m_i\right] \text{ for all i [Incentive Compatibility (IC)]}$$

$$\vec{w}^T \vec{1} \leq w_0 \text{ [Resource Constraint (RC)]}$$

$$\mathrm{E}\left[\vec{u}_i^T \vec{w} - m_i\right] \geq 0 \text{ for all i [Participation Condition (PC)]}$$

$$w_i \geq 0.$$

The equation can be interpreted as follows. First, the goal function (GF) indicates that the editor aims to maximize the exposure of the expected amount of truth, conditional on the reports provided by other clusters. Second, the incentive compatibility (IC) condition asserts that a cluster acts strategically when reporting its type. Each cluster selects its report to increase the exposure of its preferred news while suppressing alternative perspectives and minimizing lump-sum taxes. As a result, the editor must account for this strategic behavior when forming conditional expectations about the truthfulness of the news and determining exposure levels (w) and taxes (m).

Third, the resource constraint (RC) reflects the limitations on how much news can be exposed on a platform. Typically, a news platform cannot allocate excessive space to a single topic. For instance, a platform cannot cover the same topic across numerous sections.

Fourth, the participation condition (PC) assumes journalists will abandon the platform if their expected utility is negative.

Finally, the last constraint ensures that exposing a "negative quantity" of news on the platform is physically infeasible. The following subsection simplifies this problem and proposes a potential solution.

The Simplified Setting

We simplify the problem as follows. First, we assume that the preferences are certain, meaning that the utility coefficients $u_{i,\,i}$ are known for all i and j. Consequently, we assume that information asymmetry exists only regarding whether a news article is true or false (i.e., the type r_i). Still, there is no information asymmetry about the extent to which the i^{th} news cluster supports the j^{th} news cluster; for example, we know the political leanings of the clusters, such as which are liberal or conservative. Big data can be used to identify these preferences more accurately.

Second, we apply Myerson's (1981) revelation principle, which asserts that a signaling game can be transformed into a direct-revelation mechanism where participants truthfully report their types. With this, the maximization problem is formulated as follows:

$$\max_{\vec{w},\vec{m}} : E[\vec{w}^T\vec{r}] \text{ [Goal Function (GF)]}$$

$$Subject\ to\ :\ r_i = \arg\max_{\tilde{r}_i} : E[\vec{u}_i^T\vec{w} - m_i]\ for\ all\ \ i \text{ [Incentive Compatibility (IC)]}$$

$$\vec{w}^T\vec{1} \leq w_0 \text{ [Resource Constraint (RC)]}$$

$$E[\vec{u}_i^T\vec{w} - m_i] \geq 0 \; for \; all \; i \text{ [Participation Condition (PC)]}$$

$$w_i \geq 0.$$

In this simplified setting, the goal function is no longer based on a conditional expectation due to the condition of truth-telling incentive compatibility. Moreover, the incentive compatibility condition now requires that a cluster reports its type truthfully, as doing so aligns with its best interests. The resource constraint (RC), the participation condition (PC), and the nonnegativity condition for the levels of exposure remain the same as in the original problem. Based on this, we propose the following solution.

$$[\text{S01}]: m_i^* = \vec{u}_{-i}^T\vec{w}^*$$

$$[\text{S02}]: \hat{r}_i = E[r_i|\vec{r}_{-i}] \; for \; all \; i$$

$$[\text{S03}]: w_i = w^*(\hat{r}_i) \leq \frac{w_o}{N} \; for \; all \; i$$

$$[\text{S04}]: w^*(\cdot) = \arg\max_{w(\cdot)} : E[\vec{w}^T\vec{r}] \; subject \; to \; [\text{S01}], [\text{S02}[, and \; [\text{S03}].$$

The intuitions for the solutions are as follows. The lump-sum tax in [S01] induces a cluster to focus only on the level of exposure for its article. The notation is as follows:

$$\vec{u}_i^T\vec{w}^* - m_i^* = (\vec{u}_i - \vec{u}_{-i})^T\vec{w}^*.$$

The next step, [S02], indicates that the editor only evaluates the signal from a "target" cluster in the context of signals from other clusters. Since the editor does not use the signal from the cluster to assess the cluster itself, a cluster's knowledge of its potential deception does not influence the editor's estimation of its type. This reduces the incentive to deceive. For example, in [S02], the editor assesses whether a cluster tends to report idiosyncratic stories compared to the stories reported by other clusters (e.g., following a majority rule).

Additionally, [S03] specifies that a cluster's level of exposure depends solely on the estimated type, which, as per [S02], is influenced by the reports from other clusters. Since the transfer mechanism in [S01] ensures that a cluster focuses solely on its exposure, the combination of [S02] and [S03] promotes truth-telling behavior. In [S03], the exposure level $w(\cdot)$ cannot exceed $\frac{w_o}{N}$, as the total exposure should not exceed w_o, even when every cluster produces the maximum truth.

To simplify the problem further and develop an easily implementable algorithm, we introduce a hurdle rate, r_h. We now focus on symmetric solutions. Based on this simplification, we propose the following implementable solutions:

$$[\text{S03}']: \mathit{If}\ \ \hat{r}_i \geq r_h, \mathit{then}\ \ w_i^* = \frac{w_o}{N}. \mathit{Else}\ \ , w_i^* = 0$$

$$[\text{S04}']: r_h = \arg\max_{r_h}: E[\vec{w}^T\vec{r}], \mathit{subject}\ \ \ \mathit{to}\ \ [\text{S01}], [\text{S02}], \mathit{and}\ \ [\text{S03}'].$$

This algorithm is based on a realistic assumption that an editor will reject any article containing fake news beyond a specified hurdle rate (r_h) in expectation. The editor treats all articles equally, provided their expected truth exceeds the hurdle rate. Step [S03′] formalizes this rule: if an article from a cluster has an estimated truth more significant than the hurdle rate, it receives the same exposure as other articles ($\frac{w_o}{N}$); otherwise, its exposure is zero.

The key question is how strictly the editor should evaluate an article. In other words, how high should the hurdle rate be? A high hurdle rate would make the editor more conservative (strict), likely rejecting articles that do not meet the threshold and avoiding fake news. As a result, the platform may exclude idiosyncratic or "surprising" news, leading to a higher rate of Type I errors (false negatives). Conversely, a low hurdle rate would make the editor more willing to publish diverse articles. However, this could increase the likelihood of allowing fake news, leading to more Type II errors (false positives).

Step [S04′] demonstrates that the ideal hurdle rate should maximize the exposure of truthful content in line with the expectations defined under the constraints [S01], [S02], and [S03].

Readers' Participation

Readers provide valuable comments and feedback on news articles, and in today's interactive media landscape, they serve as a critical source of collective intelligence. However, they can also contribute to the spread of fake news. This raises two critical questions: How can we effectively use readers' comments in our mechanism design? And how can we incentivize readers to generate informative, truthful comments rather than to support fake news?

To address the first question, we assume that an editor can utilize reader comments to help form conditional expectations about the types of news clusters. To address the second question, we propose that an editor design an incentive structure for readers that mirrors the incentive structure for news clusters. Figure 1 illustrates the process of constructing this incentive scheme.

###Insert Figure 1 about here###

Based on these principles, we propose the following implementable solution incorporating reader participation:

$$[E01]: [r_h, v_h] = \arg\max_{r_h, v_h}: E[\vec{w}^T \vec{r}], \textit{subject to the constraints below}.$$

$$[\text{J01}]: m_i = \vec{u}_{-i}^T \vec{w}$$

$$[\text{J02}]: \hat{r}_i = E[r_i | \vec{r}_{-i}, v_{-i}]$$

$$[\text{J03}]: \mathit{If} \;\; \hat{r}_i \geq r_h, \mathit{then} \;\; w_i = \frac{w_i}{N}. \mathit{Else} \;\; , w_i = 0$$

$$[\text{R01}]: m_{i,v} = \vec{u}_{-i,v}^T \vec{w}.$$

$$[\text{R02}]: \hat{v}_{j,i} = E\left[v_{j,i} \middle| \vec{v}_{j,-i}, \vec{r}_{-j}\right]$$

$$[\text{R03}]: \mathit{If} \;\; \hat{v}_{j,i} \geq v_h, \mathit{then} \;\; q_{j,i} = \frac{w_{j,i}}{w_0} \cdot \frac{q_0}{N_v}. \mathit{Else} \;\; , q_{j,i} = 0.$$

The solution is divided into three components: one for the editor ([E01]), one for the news clusters ([J01] - [J03]), and one for the readers ([R01] - [R03]). In [E01], the editor sets hurdle rates for both news clusters (r_h) and readers (v_h) to maximize the exposure of truthful content. We denote the matrix **v** to represent the structure of reader comments as follows:

$$N_v: \mathit{Number\ of\ readers}$$

$$v: N * N_v \ \mathit{dimensional\ matrix\ for\ reader\ comments}$$

$$\vec{v}_i \in v: N \ \mathit{dimensional\ column\ vector}$$

$$i^{th} \ \mathit{reader's\ true\ evaluation\ for\ N\ news\ (clusters)}$$

$$v_{j,i} \in \vec{v}_i: i^{th} \ \mathit{reader's\ true\ evaluation\ for}\ j^{th} \ \mathit{news\ (cluster)}$$

$$\tilde{v}_{j,i} \in \vec{v}_i: i^{th} \ \mathit{reader's\ reported\ evaluation\ (possibly\ fake)}$$

$$\vec{v}_{j,-i}: N_v - 1 \ \mathit{vector\ of\ the\ comments\ about}\ j^{th} \ \mathit{news\ cluster\ except\ reader}\ i's.$$

Blocks [J01] - [J03] aim to encourage news clusters to report the truth based on three key principles: focusing on their report only ([J01]), estimating the report's type in the context of other clusters' reports ([J02]), and exposing news with a higher truth content more significantly ([J03]).

Blocks [R01] - [R03] are designed to motivate readers to contribute truthful comments. Block [R01] ensures that readers focus on the content of an article when commenting. Block [R01] also introduces a lump-sum tax for readers. In [R02], the editor can assess the history of a reader's comments, including whether those comments align with others and with estimated types. The related notations are as follows:

$$\vec{u}_{i,v}^{T}\vec{w} - m_{i,v}: Utility \ of \ i^{th} \ reader \ when \ commenting \ on \ v's \ news \ cluster$$

$$u_{j,i,v} \in \vec{u}_{i,v}: Utility \ coefficient \ of \ i^{th} \ reader \ on \ j^{th} \ news \ cluster$$

$$\vec{u}_{-i,v}: i^{th} \ element \ is \ set \ to \ zero \ in \ \vec{u}_{i,v}$$

$$\vec{u}_{i,v} \in u_v: N * N_v \ utility \ matrix.$$

Finally, we assume that, like the editor, a reader can prefer exposure to truthful news. However, this does not substantially alter the setting because it reduces the conflict of interest between the editor and the reader, thereby making the mechanism more robust. Block [R02] indicates that the editor evaluates the reader i's comment on the j^{th} cluster in the context of comments from other readers ($(-i)$'s) about the same cluster. Block [R03] specifies that the editor has q_0 amount of "points" available for distribution to readers. These points are something of value to readers, and the allocation is determined by whether a reader's comment meets or exceeds the threshold for truth (hurdle rate v_h). Since any reader's comment on any article can include enough truth to earn points, [R03] demonstrates the editor's capacity to allocate these points.

DISCUSSION

Using a Smart Contract

We propose a solution involving a blockchain-based smart contract to address the concern about trust that enables the algorithm to work correctly. Why blockchain and why a smart contract? A smart contract is written in computer code that enforces an agreement programmatically when an event satisfies pre-specified conditions. Blockchain technologies can help improve these smart contracts' transparency, traceability, and irreversibility. The following points address why implementing a blockchain-based smart contract is appropriate for implementing the solution.

First, a blockchain-based smart contract can ensure irreversibility, which is essential in the solution (*see* [J02] and [R02]). When an editor evaluates a report from a cluster or a comment from a reader, the editor is not supposed to use "direct" information from the cluster or the reader. Instead, the editor should use "indirect" information from other clusters or readers. However, an editor might be incentivized to break this commitment if they know the participants only reveal the truth. If clusters or readers doubt the editor's commitment, they will deviate from the solution, undermining the entire mechanism. Similarly, [J03] and [R03] must be irreversible in enforcing strict adherence to the predetermined payment protocol for readers and news clusters.

Second, the solution requires transparency. The allocation of points or news exposure depends on a series of decision-making processes. The solution requires that participants understand the elements of the decision-making process, including hurdle rates, point allocation, maximization, and conditional expectations about signals. In particular, the editor's decision-making process must be transparent to maintain trust in the system. A blockchain system can provide this transparency by tracking and storing transactions on a network synchronizing information across nodes. If readers and news clusters participate in the network as nodes, they can synchronize, create, store, and broadcast information along distributed ledgers. Data manipulation is complicated once stored in a blockchain network, especially in a public blockchain where no single entity controls the network. This transparency and traceability are why global non-governmental organizations use blockchain systems to build trust and prevent corruption.

Third, traceability is crucial in our mechanism. Traceability refers to the ability to identify information at a low cost. With blockchain, every transaction is traceable because distributed ledgers synchronize transactions in near real-time and permanently. In a programmable open source blockchain, the design of a decision-making process is transparent and traceable whenever updated. Our solution requires the editor to trace the history of signals from readers and clusters to form conditional expectations and undertake complex resource allocation. The editor must also trace participant preferences to implement lump-sum taxes, as described in [J01] and [R01]. If participant preferences are not directly observable, the editor may need to accumulate big data to estimate preferences by analyzing participant behavior over time. Table 2 illustrates a sample implementation strategy using a smart contract algorithm.

Insert Table 2 about here

The final optional step strengthens the truth-telling condition (i.e., IC) on the Folk theorem. Although not explicitly mentioned in the above algorithm, an editor must collect "big data" about news clusters and readers to identify their preferences. This data can include background checks on news clusters and reader surveys.

Our algorithm requires resources to allocate incentive payments and lump-sum taxes. These resources may come from advertisers, reader donations, and/or subscription fees. The work of Chyi (2005) provides a good explanation of the subscription model in media. To make a subscription model work, it is essential to build a strong community. Blockchain can be instrumental in building this type of community (Chen 2018).

Limitations and Challenges

Implementing blockchain technology in media and journalism requires significant technical expertise and infrastructure. Developing and maintaining a blockchain system requires skilled professionals and robust technological support. These requirements could pose barriers for smaller news organizations with limited resources.

Furthermore, integrating blockchain and smart contracts in media could fundamentally reshape the traditional roles of journalists and editors. Moving towards a decentralized model emphasizes collective credibility and quality assurance rather than hierarchical oversight. This shift could democratize media production and consumption, but it also requires a cultural change within organizations and among media professionals.

In addition, while blockchain offers enhanced transparency and traceability, it questions data privacy and security. Ensuring that sensitive information is protected while maintaining the openness of the blockchain is a complex challenge. Implementing measures such as encryption and secure data handling protocols will be crucial.

Last but not least, the algorithm proposed requires resources to allocate incentive payments and lump-sum taxes. Identifying sustainable funding sources, such as advertisers, reader donations, or subscription fees. The work of Chyi (2005) on the subscription model in media highlights the importance of building a strong community to support this funding model. However, building and maintaining such a community requires ongoing effort and investment.

In sum, while blockchain-based smart contracts offer promising solutions for addressing trust, transparency, and traceability in media, it is crucial to consider these limitations and challenges. Technical, security, and structural issues must be addressed to ensure a balanced and effective implementation. By carefully weighing these factors, media organizations can harness the potential of blockchain technology to foster a more participatory and accountable media landscape.

Extension of the literature

This study draws on and extends the existing literature in several ways. First, lump-sum taxes and conditional expectations on the signals of others are special cases of the Vickrey-Clarke-Groves (VCG) mechanism (Vickrey 1961; Clarke 1971; Groves 1973; Cohen and Loeb 1984). Second, this study applies the revelation principle of Myerson (1981), which implies that any signaling game can be transformed into a truth-telling direct revelation mechanism. Third, the basics of the proposed model rely on the implications of prior studies regarding the correlated mechanism design (McAfee, McMillan, and Reny 1989; McAfee and Reny 1992).

In our model, all information is correlated, whether from readers or news clusters, because the stakeholders deal with the same topic. Prior studies find that the efficiency of a correlated mechanism increases with correlation. Accordingly, our suggested algorithm against fake news can be improved if the correlation between information increases. An editor must use sophisticated classification methods to establish and define news clusters and readers around similar topics to increase the correlation. Furthermore, collecting big data about the information and behavior of readers and news clusters would also be essential.

Lastly, the proposed algorithm can be utilized in various fields or academic studies of innovation and technology, finance, journalism, and sociology. Since the problem of fake news is a social phenomenon threatening the basics of our society, including the economy, journalism, democracy, and even national security (Marchi, 2012; Persily, 2017), significant efforts to address this issue are deemed necessary.

Suggestions for Future Research

Following our insights and model, future research can explore alternative consensus mechanisms beyond proof-of-work to reduce the environmental impact of blockchain technology. Investigating proof-of-stake, proof-of-authority, and other innovative mechanisms can provide insights into more sustainable options for implementing blockchain in media and journalism.

Furthermore, examining how blockchain technology can be integrated with existing media platforms and systems is crucial. Research should focus on developing hybrid models that combine traditional media practices with blockchain features to enhance trust and transparency without disrupting current workflows.

Given data privacy and security concerns, future research can also focus on developing robust solutions that balance transparency with confidentiality. This includes studying encryption methods, privacy-preserving data-sharing techniques, and regulatory frameworks that protect user data while maintaining the benefits of blockchain.

Various case studies can also emerge from our study. For instance, future research could apply the proposed algorithm in conjunction with blockchain-based smart contracts to explore their potential impact on mitigating the effects of fake news in specific contexts, such as the stock market.

Lastly, future research can examine the regulatory and policy implications of adopting blockchain in media. Understanding the legal challenges, compliance requirements, and potential regulatory frameworks will help create an environment conducive to innovation while ensuring accountability and ethical standards.

Conclusion

How can we stop the propagation of fake news? Fake news is intentionally and strategically disguised as authentic, with some journalists and readers willing to spread it to maximize their own goals. The socially constructed nature of fake news makes it difficult to detect and discourage. Therefore, in seeking a solution, this study focuses on the incentives of readers and journalists. Specifically, the proposed algorithm is designed to discourage readers and journalists from manufacturing and spreading fake news while encouraging them to spread authentic news. In this algorithm, readers and news clusters would voluntarily impede fake news.

However, designing an algorithm incorporating the nexus of contracts between editors, journalists, and readers is challenging. The contracts further specify a protocol for lump-sum taxes, incentive payments, and methods of evaluating the extent of falsification using conditional expectations on large datasets. Moreover, the suggested algorithm requires trust and decentralized consensus building. As a solution, we propose exploiting the advantages of a blockchain system—namely, irreversibility, transparency, and traceability. We demonstrate that a blockchain-based smart contract can implement the proposed algorithm effectively.

Our results complement other strategies for eradicating fake news, such as codes of conduct, standards, fact-checking organizations, and deep learning. By leveraging blockchain technology, we aim to create a more robust and transparent system that aligns the incentives of all participants—editors, journalists, and readers—to foster the dissemination of authentic news and curb the spread of false information. Future research should continue to explore blockchain integration with existing media platforms, the economic models for sustainability, and the broader impacts on journalistic practices to refine and enhance this approach.

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. Journal of economic perspectives, 31(2), 211-36.

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press.

Arcuri, M. C., Gandolfi, G., & Russo, I. (2023). Does fake news impact stock returns? Evidence from US and EU stock markets. *Journal of Economics and Business*, 125, 106130.

Brennen, J. S., Simon, F., Howard, P. N., & Nielsen, R. K. (2020). Types, Sources, and Claims of COVID-19 Misinformation. Reuters Institute.

Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, *6*(2), 154-175.

Cassauwers, T. (2019, April 15). Can artificial intelligence help end fake news?. *Horizon*. https://ec.europa.eu/research-and-innovation/en/horizon-magazine/can-artificial-intelligence-help-end-fake-news.

Chen, Y. (2018). Blockchain tokens and the potential democratization of entrepreneurship and innovation. Business Horizons, 61(4), 567-575.

Chyi, H. I. (2005). Willingness to pay for online news: An empirical study on the viability of the subscription model. *Journal of Media Economics*, 18(2), 131-142.

Clarke. E. H. (1971). Multipart pricing of public goods. *Public Choice*. 11(1): 17-33.

Clarke, J., Chen, H., Du, D., & Hu, Y. J. (2020). Fake news, investor attention, and market reaction. *Information Systems Research*, *32*(1), 35-52.

Cohen, S. I., & Loeb, M. (1984). The Groves scheme, profit sharing and moral hazard. *Management Science*, 30(1), 20-24.

Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology, 52(1),* 1-4.

Delmolino, K., Arnett, M., Kosba, A., Miller, A., & Shi, E. (2016, February). Step by step towards creating a safe smart contract: Lessons and insights from a cryptocurrency lab. In *the International Conference on Financial Cryptography and Data Security (pp. 79-94).* Springer, Berlin, Heidelberg.

Figueira, A., & Oliveira, L. (2017). The current state of fake news: challenges and opportunities. *Procedia Computer Science*, 121, 817-825.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council*, *9*(3), 4.

Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013, May). Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 729-736).

Groves. T. (1973). Incentives in teams. Econometrica. 41(4): 617-631.

Kogan, S., Moskowitz, T. J., & Niessner, M. (2023). Social media and financial news manipulation. *Review of Finance*, *27*(4), 1229-1268.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094-1096.

Lee, A. M. (2013). News audiences revisited: Theorizing the link between audience motivations and news consumption. *Journal of Broadcasting & Electronic Media*, 57(3), 300-317.

Loomba, S., De Figueiredo, A., Piatek, S. J., De Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, *5*(3), 337-348.

Marchi, R. (2012). With Facebook, blogs, and fake news, teens reject journalistic "objectivity". *Journal of Communication Inquiry*, 36(3), 246-262.

Marr, B. (2018). How much data do we create every day? The mind-blowing stats everyone should read. *Forbes*, *21*, 1-5.

McAfee. R. P. & Reny. P. J. (1992). Correlated information and mechanism design. *Econometrica*. 60(2): 395-421.

McAfee. R. P. McMillan. J. & Reny. P. J. (1989). Extracting the surplus in the common-value auction. *Econometrica*. 57(6): 1451-1459.

Myerson. R. B. (1981). Optimal auction design. *Mathematics of Operations Research*. 6(1): 58-73.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Satoshi Nakamoto*.

Painter, C., & Hodges, L. (2010). Mocking the news: How The Daily Show with Jon Stewart holds traditional broadcast news accountable. *Journal of Mass Media Ethics*, 25(4), 257-274.

Persily, N. (2017). The 2016 US Election: Can democracy survive the internet?. *Journal of Democracy*, 28(2), 63-76.

Rubin, V., Conroy, N., Chen, Y., & Cornwell, S. (2016, June). Fake news or truth? using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection* (pp. 7-17).

Ruchansky, N., Seo, S., & Liu, Y. (2017, November). Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 797-806). ACM.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36.

Siering, M., Clapham, B., Engel, O., & Gomber, P. (2017). A taxonomy of financial market manipulations: establishing trust and market integrity in the financialized economy through automated fraud detection. *Journal of Information Technology*, *32*(3), 251-269.

Sunstein, C. R. (2001). *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton, NJ: Princeton University Press.

Vickrey. W. (1961). Counterspeculation. auctions. and competitive sealed tenders. *Journal of Finance*. 16(1): 8-37.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-1151.

Figure 1. The process of constructing an incentive structure.

Step 1. Form clusters based on views expressed about types of news and opinion.

Step 2. Compute the expected value for the amount of truth in a reader cluster's comments on other reader clusters' views and news.

Step 3. Compare the expected value and a hurdle rate for readers.

Step 4. Use a lump-sum tax and incentive payment structure similar to these structures for news clusters. To account for reader participation in the consuming of news, a solution similar to the previous implementable solution emerges.

Table 1. Algorithmic Process to Eliminate Fake News

Step	Title	Description	Key Actions
1	Categorization	Journalists' articles are grouped into clusters based on tones and viewpoints.	Articles with similar tone and subject matter are categorized together.
2	Truth Metrics Reporting	Each cluster reports the truth level of their articles.	Metrics include confidence intervals, standard errors, and relevant dimensions.
3	Reader Comments	Selected readers provide comments and rate the truthfulness of articles.	Readers assess and share their perceptions of the article's truth.
4	Truth Assessment	The editor evaluates the truth level of articles and comments without using information from the same cluster.	Truth is assessed using data from other correlated clusters and comments from other readers.
5	Exposure and Payments	Articles are assigned exposure levels, and payments are distributed.	Articles and readers meeting truth thresholds receive exposure and payments.
6	Lump-Sum Taxes	Taxes are imposed on clusters and readers to incentivize quality.	Participants are encouraged to prioritize accuracy and truthfulness.
7	Penalties for Falsification	Penalties are enforced for producing fake news or comments.	Responsible clusters or readers are temporarily prohibited from contributing.

Table 2. An example of the implementation strategy with a smart-contract algorithm.

Step	Agent	Description
1	Editor	Develop an irreversible smart contract, which specifies the conditional expectation function, hurdle rates, allocation rules (e.g., exposures for news clusters and points for readers), and lump-sum taxes.
2	Writers	Produce articles. Article writers (journalists) also report the amount of truth or their level of confidence about facts.
3	Readers	Comment on articles and report amount of truth and reports this estimation to the editor.
4	Editor	Compute the conditional expectations regarding news clusters and readers.
5	Platform	Allocate lump-sum taxes, exposures, and points.
6 (optional)	Platform	Later, articles turn out to be comprised of news that is either true or false. The platform forbids clusters or readers from producing or commenting on the news for a sufficient duration of time if the clusters or readers have produced fake news or comments.