Mitigating Misinformation Spread on Blockchain Enabled Social Media Networks

Rui Luo and Vikram Krishnamurthy, Fellow, IEEE

Abstract—We construct a blockchain-enabled social media network to mitigate the spread of misinformation. We derive the information transmission-time distribution by modeling the misinformation transmission as double-spend attacks on blockchain. This distribution is then incorporated into the SIR model, which substitutes the single rate parameter in the traditional SIR model. Then, on a multi-community network, we study the propagation of misinformation numerically and show that the proposed blockchain enabled social media network outperforms the baseline network in flattening the curve of the infected population.

Index Terms—Blockchain, double-spend attack, proof-of-work, misinformation propagation, SIR model, social media networks.

1 INTRODUCTION

The spread of misinformation puts the information integrity and trust relationships in social networks at risk. Recent advancements in blockchain technology have opened up new possibilities for enabling decentralized trust in a peer-to-peer network [1]. In this paper, we examine how blockchain technology can be utilized to mitigate the spread of misinformation in decentralized social media networks.

Why Distributed Ledger in Social Networks: A decentralized social network does not have a central proprietary authority that stores and controls all the data available to the users. Instead, data is stored at multiple nodes in the network. An important aspect of decentralized social media networks is that no single entity has control over what can be published; hence there is greater freedom of expression. Since there is no centralized control or moderator of content, such networks are subject to misinformation (fake news) and inappropriate content. This motivates our current paper: we propose a blockchain protocol that can mitigate the spread of misinformation in decentralized social media networks.

To prevent fraudulent transactions, blockchain based cryptocurrencies use an immutable, distributed ledger. A blockchain enabled social media network, on the other hand, uses a distributed ledger to store data. This implies that every interaction is irrevocably recorded in the blockchain and safeguarded by end-to-end encryption, and the blockchain finds consensus among all members of a social media network in a decentralized manner. With these desirable features, blockchain technology has been emphasized as a possible counter-measure to misinformation [2] and there have been recent real-world implementations [3].

Modeling Misinformation as Double-Spend Attacks in a Blockchain Enabled Network: Misinformation occurs when the truth or useful information is purposefully altered to mislead other users. Similarly in blockchain, a double-spend attack occurs when the same digital token is involved in multiple transactions. We will model information exchange in social media networks as transactions on blockchains, where misinformation is represented as double-spend attack.

Our proposed blockchain enable social media network has two components. First, the blockchain protocol mitigates the spread of misinformation as follows. New social media postings require approval by miners and further confirmation by successive blocks. The confirmation time is longer for posts that are deemed to contain misinformation. Second, we model the information diffusion in the social network as an epidemic model where the blockchain confirmation time affects the information transmission rate. Epidemic models characterize the reproduction number, which is a good proxy for the engagement rate and predictor for epidemic-like information spreading [4], and compartments in epidemic models specify different stages of being affected by misinformation and how the populations evolve.

Main Results and Organization: (1) The main idea of this paper is to construct a blockchain protocol that examines social media postings and mitigates misinformation by exploiting the blockchain transaction confirmation methodology. The social media postings are processed as blockchain transactions, with a higher proportion of honest miners in the network ensuring that misinformation takes longer to propagate.

(2) Section 2 details our model of the blockchain enabled social media network protocol. First we model the misinformation propagation in social media networks as double-spend attacks on blockchain. Then in Section 2.2 we devise an integral equation for the SIR model that includes the

---

1. Examples include Steemit, Sola, Civil, onG.social, and Sapien. See Appendix C for details.
misinformation transmission time. The transmission-time distribution is derived in Section 2.3 by creating two Poisson processes for blocks mined by dishonest and honest miners. The discrete difference equations for the SIR model are formulated in Section 2.4 which considers the different time scales of information propagation and block mining.

Section 3 constructs the multi-community network and proposes the stochastic simulation framework (Algorithm 2).

Section 4 presents simulation results of the misinformation SIR dynamics on a three-community network and illustrates the advantages of blockchain in combating misinformation. We also show that the resulting blockchain enabled social media network results in communities that are more resilient to misinformation propagation than others. Parameters of the SIR model are estimated from Twitter hashtags datasets [5].

1.1 Related Work

Previous works can be considered under two categories. We summarize these in the context of our paper.

1. **SIR epidemic model with heterogeneous parameters**: The classical SIR model assumes the same susceptibility for all individuals and the same infectivity for epidemic (information) spread in social media networks. Previous studies have analyzed the effect of the susceptibility heterogeneity by either dividing individuals into different groups or constructing a distribution of susceptibility. Baqaee [6] considers a five-population SIR model where sub-populations correspond to age groups and the interactions between age groups are calibrated using survey data. Gou and Jin [7] generalized the SIR model by considering both the heterogeneity of degree, and heterogeneity of susceptibility and recovery rate. They found that given the mean of the distribution of susceptibility, increasing the variance may block the spread of epidemics. Lachiany and Louzoun [8] studied the variability in infection rates which explains the discrepancy between the observation and the predicted number of infected individuals using the typical SIR model. Smilkove et al. [9] studied a SIR model with degree-correlated heterogeneous susceptibility and found that positive correlation between a node’s degree and susceptibility leads to a vulnerable model to the spread of disease. Stanoev et al. [10] considered concurrent spread of an arbitrary number of contagions in a network. An individual can be affected through multiple channels with its neighbors with different contact rates.

2. **Blockchain enabled network models**: In social media networks, blockchain technology has been used to improve data privacy and resilience to misinformation. Chen [11] studied how blockchain can slow down the spread of rumor on social media networks. They explored decentralized contracts and virtual information credits for secure and trustful peer-to-peer information exchange. Fu and Fang [12] employed blockchain to construct a decentralized personal data management system that ensures users own and control their data without authentication from a third party. Tessone et al. [13] proposed a minimalistic stochastic model to understand the dynamics of blockchain enabled consensus on a network. Barański [14] studied mitigating the content poisoning attacks on information-centric networks.

| Symbols | Description |
|---------|-------------|
| $G$     | the undirected graph representing the social media network |
| $V$     | the set of nodes |
| $E$     | the set of edges |
| $A$     | the adjacency matrix of $G$ |
| $S(t)$  | the number of susceptible nodes at time $t$ |
| $I(t)$  | the number of infected nodes at time $t$ |
| $R(t)$  | the number of recovered nodes at time $t$ |
| $S(t)$  | the set of susceptible node and transmission time pairs at time $t$ |
| $I(t)$  | the set of infected nodes at time $t$ |
| $R(t)$  | the set of recovered nodes at time $t$ |
| $B$     | the probability matrix of the stochastic block model (SBM) |
| $V_m$   | the community $m$, which is a subset of $V$ |
| $\beta_m$ | the contact rate of nodes in community $m$ |
| $\alpha_m$ | the recovery rate of nodes in community $m$ |
| $T_{S\rightarrow I}$ | the transmission time |
| $P_{T_{S\rightarrow I}}(t)$ | the CDF of $T_{S\rightarrow I}$ |
| $P_{T_{S\rightarrow I}}(t)$ | the number of leading blocks needed to confirm a previous block |
| $T_k$ | the first time that blocks mined by dishonest miners outnumbers blocks mined by honest miners by $k$ blocks |
| $N_k$ | the total amount of blocks at $T_k$ |
| $N_1(s)$ | the number of blocks mined by dishonest miners by time $s$ |
| $N_2(s)$ | the number of blocks mined by honest miners by time $s$ |
| $N(s)$ | the total number of mined blocks by time $s$ |
| $g(j)$ | the PMF of $N_k$ |
| $G(j)$ | the CDF of $N_k$ |
| $P(T_k \leq s)$ | the CDF of the transmission time |
| $\mu_1$ | the computing power of dishonest miners |
| $\mu_2$ | the computing power of honest miners |
| $l$ | the ratio of the time unit in the network and the time unit on the blockchain |
| $c_i$ | the community label of node $i$ |
| $x_i(t)$ | the state of node $i$ at time $t$ |
| $N_{i,m}(t)$ | the set of node $i$'s neighbors from community $m$ that are infected at time $t$ |

The author proposed a blockchain enabled Proof-of-Time authentication mechanism which improves the network’s resilience. Saad et al. [15] proposed a high-level overview of a blockchain enabled framework for misinformation prevention and highlight the various design issues and consideration of such a blockchain enabled framework for tackling misinformation. Qayyum et al. [2] considered a blockchain implementation of news feed to distinguish facts from fiction.

Our work differs from existing works in that

- We explicitly model misinformation transmission as double-spend attacks on blockchain. With new information
posting decided by a blockchain transaction approval process, we analytically derive the distribution that misinformation propagates in the social media network.

- In our blockchain enabled social media network, users themselves take charge of their social interactions without relying on service providers (e.g., central authority or official publishers) – that is, users function as blockchain miners for information exchange in the social media network, resulting in a decentralized autonomous organization.
- In our numerical studies with parameters estimated from real world Twitter datasets, we show that the blockchain enabled social media network slows down the misinformation propagation and reduces the number of infected users.

## 2 PROPAGATION OF MISINFORMATION ON THE BLOCKCHAIN

In this section, we construct a SIR based model for misinformation propagation in blockchain enabled social media networks. We first introduce the blockchain protocol where social media network messages are checked by miners. Then we use the SIR model to study the propagation dynamics of misinformation. Specifically, we decompose the transmission rate into a product of the contact rate and the infectivity. We characterize the infectivity using a transmission-time distribution and derive the integral equation of the SIR model. Our key idea is to relate the transmission-time distribution with the time it takes for double-spend attacks to succeed on the blockchain.

### 2.1 Blockchain Protocol for Misinformation Propagation

This subsection describes the social media network’s blockchain system, including how miners approve and encapsulate information into blocks, as well as how blocks are confirmed and information is propagated.

Our proposed blockchain enabled social media network utilizes a distributed ledger to ensure that users equally have full control of their social interactions, i.e., what they post and what they see.

Specifically, every active user can become a miner, who checks the authenticity of other users’ messages. Miners utilize their computing power to search the knowledge base or conduct data mining and text analysis to decide if a message is truth or misinformation. The approved message will be encrypted and added to the blockchain as a block and the first miner to approve it will be rewarded virtual credit in the social media network. If a particular number of subsequent blocks are attached after it, the block is confirmed on the blockchain. We categorize miners into two groups: honest miners who approve true news, and dishonest miners who approve misinformation. Concretely, the blockchain protocol has the following rules:

1) Each miner has equal probability to be the first one to approve the message.
2) A new block will attach to an existing block of the same type, i.e., a new block created by a dishonest miner will attach to a block containing misinformation. If there is no existing blocks to attach to, a new block will become a genesis block.

3) When two blocks are created simultaneously, there will be a fork which results in two diverged chains.
4) When the cumulative successive blocks of one block reach a predefined number $k$, then the block is confirmed and the encapsulated message is published. The shorter forked chain will be aborted.

Based on the blockchain protocol, misinformation is propagated from an infected user to another user if a dishonest miner approves the news followed by $k$ other approvals of (other) misinformation. A schematic of the proposed blockchain enabled social media network is shown in Fig. 1.

**Fig. 1: SIR dynamics in blockchain enabled social media network.**

- **Social Network:** The top panel shows the decentralized social media network proposed in this paper. Red nodes ($u_2, u_4$) and blue nodes ($u_1, u_3, u_5$) represent infected and susceptible users, respectively. Messages are numbered in temporal order ($m_1, \cdots, m_6$) and represented as directed edges with red ones representing misinformation. Solid edges denote messages that are approved by miners and encapsulated in the blockchain while dashed ones denote messages that are disapproved and aborted.

- **Blockchain:** The middle panel exemplifies a double-spend attack in blockchain. The original information (the “i” icon) is turned into truthful message $m_1$ (blue) and misinformation $m_2$ (red), which both get approved (by $u_3$ and $u_4$ respectively as shown in the bottom half of the blocks) and encapsulated into blocks. The chain “$m_2 - m_3 - m_5$” outnumbers the other chain “$m_1$” by 2, which leads to $m_2$ getting confirmed and published in the social media network, whereas $m_1$ getting aborted.

- **SIR Model:** The bottom panel shows the three compartments ($S$-susceptible, $I$-infected, $R$-recovered) in SIR model. Because the misinformation $m_3$ is published by the blockchain, $u_1$ becomes infected by $u_2$ and enters $I$ from $S$. When a new block is created by a dishonest miner, the misinformation propagates in the social media network.
In Fig. 1, in Section 2.2, we derive the misinformation transmission-time distribution. Using this distribution we will show how the blockchain enabled social media network slows down the misinformation propagation following a blockchain transaction approval process. A numerical evaluation is provided in Section 4.

2.2 Transmission-time Distribution in SIR Model

In epidemic modeling, the transmission rate is the number of people infected in a given amount of time. In a classical SIR model, the transmission rate is defined as the product of contact rate and transmission probability. This measures the number of contacts an individual makes per unit of time and the probability that a contact results in a susceptible individual becoming infected, respectively.

The classical SIR model [16] has dynamics determined by the following ordinary differential equation:

\[ \dot{S}(t) = -\beta S(t)I(t) \]  

(1)

Here \( S(t) \), \( I(t) \) denote the number of susceptible users and infected users respectively; \( N \) denotes the total population size which is constant \( \beta > 0 \) denotes the contact rate. The number of new infections \( \beta S(t)I(t) \) is given by mass action incidence, i.e., a user is assumed to make \( \beta N \) contacts in unit time.

Eq. (1) assumes that a susceptible user get infected immediately after contacting with an infected user. It is unable to reflect true scenarios in which users become infected after a period of time – The effect of an interaction may lead to infection in the future. To account for this, we incorporate the transmission time \( T_{S\rightarrow I} \) into the model, which refers to how long it takes for a contact between a susceptible user and an infected user to turn into an infection. This construction is similar to the model with arbitrarily distributed disease stages proposed in [18].

Let \( P_{T_{S\rightarrow I}}(t) \) denote the cumulative distribution function (CDF) of the transmission time,

\[ P_{T_{S\rightarrow I}}(t) = P(T_{S\rightarrow I} \leq t), \]  

(2)

which is the probability that the transmission time lasts no longer than \( t \). The derivative \( P_{S}(t) \) is the probability density function (PDF) of the transmission time, i.e., the probability that a susceptible individual becomes infected at time \( t \) since the contact. The CDF and PDF have the following properties:

\[ P_{T_{S\rightarrow I}}(0) = 0, \quad P_{T_{S\rightarrow I}}(\infty) = 1, \]
\[ P_{T_{S\rightarrow I}}(t) \geq 0, t \geq 0. \]  

(3)

Based on the transmission-time distribution \( P_{T_{S\rightarrow I}}(t) \), we obtain the following integral equation for the susceptible population:

\[ \dot{S}(t) = -\int_0^t \beta S(\tau)I(\tau)P_{T_{S\rightarrow I}}(t-\tau)d\tau \]  

(4)

2. Also referred to as infection rate, transmission coefficient, or \( \beta \) in the literature.

3. We neglect the creation/deletion of user accounts during the misinformation transmission [4] – the speed of misinformation spread is much faster than the life cycle of social media.

4. See Section 2.1 in [17].

where \( \beta S(\tau)I(\tau) \) is the expected number of contacts between the susceptible and the infected population at time \( \tau \); \( P_{T_{S\rightarrow I}}(t-\tau) \) denotes the probability that the transmission time is \( t-\tau \).

Note that in the classical SIR model, \( P_{T_{S\rightarrow I}}(t) \) is a unit step function at 0 and \( P_{T_{S\rightarrow I}}(t) \) is Dirac delta function \( \delta(t) \). Thus Eq. (4) becomes

\[ \dot{S}(t) = -\int_0^t \beta S(\tau)I(\tau)\delta(t-\tau)d\tau = -\beta S(t)I(t) \]  

(5)

which is exactly the differential equation (1) of the classical SIR model.

In the subsection below, we will incorporate the SIR model with the transmission-time distribution specified by the blockchain double-spend attack.

2.3 Transmission-Time Distribution with Blockchain

We now explain how the duration of a blockchain double-spend attack can be used to compute the misinformation transmission-time distribution.

A double-spend attack is defined as fraudulent transactions with users spending the same digital token more than once [19]. To launch a double-spend attack in the blockchain enable social media network, a dishonest miner or mining pool reverses a previous block (containing misinformation) and attempts to rapidly approve all following blocks (containing misinformation) in order to build a longer chain than that created by the collective honest miners, as shown in the middle panel of Fig. 1. The blocks that come after the previous block serve as confirmations. Once a set amount of confirmations are received, the previous block’s misinformation will be confirmed and published in the social media network. The time it takes for double-spend attacks to succeed is essentially the transmission-time distribution discussed in Section 2.2.

In what follows, we derive the transmission-time distribution using the first-hitting-time of a blockchain approval process [20]. Let \( \{N_1(s), s \geq 0\} \) and \( \{N_2(s), s \geq 0\} \) denote the blocks mined by dishonest and honest miners, which are independent Poisson processes with respective rates \( \mu_1 \) and \( \mu_2 \). The misinformation transmission-time is the first time that \( N_1 \) is \( k \) greater than \( N_2 \), which is defined as \( T_k \).

\[ T_k = \inf\{s \geq 0 : N_1(s) = N_2(s) + k, k > 0. \]  

(6)

Consider the mining process as multiple rounds of games between dishonest miners and honest miners, and in each round the winning(losing) team increase(decrease) their score by 1. We represent dishonest miners’ score after \( n \) rounds as a random walk \( S_n = \sum_{i=1}^{n} X_i \), where \( X_i, i \geq 1 \) are independent and that \( P(X_i = 1) = p = \frac{\mu_1}{\mu_1 + \mu_2} = 1 - P(X_i = -1) \). From Eq. (6) dishonest miners score \( k \) for the first time at \( T_k \), i.e.,

\[ N_k = \min\{n : S_n = k\}, \]  

(7)

where \( N_k = \infty \) if \( S_n < k \) for all \( n \).

5. See https://en.bitcoin.it/wiki/Confirmation
Lemma 1. \cite{20}

\[
P(N_k = k + 2i) = \frac{k}{k + 2i} \binom{k + 2i}{k + i} p^{k+i}(1-p)^i, \quad i = 0, 1, \ldots
\]

(8)

Proof. See Appendix A

Define \( g(i) = P(N_k = k + 2i), G(i) = P(N_k \leq k + 2i) \) as the probability mass function (PMF) and the CDF of \( N_k \) respectively. Lemma 1 shows that

\[
g(i+1) = \frac{(k+2i)(k+2i+1)}{(k+i+1)(i+1)} p^{i+1}, \quad i \geq 0
\]

(9)

which can be used to recursively compute \( g(i) \) and \( G(i) \).

Next, using law of total probability, we show the PMF and CDF of \( T_k \) in the following Theorem 1.

Theorem 1. The probability that misinformation is confirmed and published in the social media network (i.e., the corresponding block receives \( k \) confirmations in the blockchain) at time \( s \), i.e., \( P(T_k = s) \), is the sum of the probabilities of the joint events that (1) a total of \((j + k)\) blocks being mined; and (2) dishonest miners mining the \((j + k)\)-th block, resulting in its chain outnumbering honest miners’ chain by \( k \) blocks for the first time, where \( j = 0, 1, \ldots \):

\[
P(T_k = s) = \sum_{j=0}^{\infty} g(j/2) e^{-\mu s} (\mu s)^{j+k}/(j+k)!
\]

(10)

Similarly the CDF of \( T_k \) is:

\[
P(T_k \leq s) = \sum_{j=0}^{\infty} G(j/2) e^{-\mu s} (\mu s)^{j+k}/(j+k)!
\]

(11)

where \( \mu = \mu_1 + \mu_2 \) and \([x]\) denotes the largest integer less or equal to \( x \).

Proof. See Appendix B

Assume that \( \mu_1 < \mu_2 \), i.e., the computing power of the dishonest miners is less than that of the honest miners. Then the probability that the simple random walk with \( p = \frac{\mu_1}{\mu_1 + \mu_2} \) ever goes up \( k \) is equal to \((\frac{\mu_1}{\mu_2})^k \) \cite{20}. Consequently,

\[
P(N_k < \infty) = P(T_k < \infty) = \left(\frac{\mu_1}{\mu_2}\right)^k
\]

(12)

To simulate the transmission time from the CDF \( P(T_k \leq s) \mid T_k < \infty \), we propose Algorithm 1 which uses inverse transform method.

2.4 Discrete Difference Equations for SIR Model

In this subsection, we discretize the model in Eq. (4) and include the transmission-time (Section 2.2) which results in the discrete-time difference equations for SIR model.

First, we connect the PMF and CDF in Eq. (10,11) to the transmission-time distribution introduced in Section 2.2

\[
Pr_{T_k < t} (t) = P(T_k \leq t)
\]

(13)

Algorithm 1 Inverse sampling of the transmission time

Input: The transmission-time distribution \( P(T_k \leq s) \mid T_k < \infty \) with parameters \( k, \mu_1, \mu_2 \).

Output: A sampled transmission time \( s \).

1. Sample \( U_1 \sim \text{Unif}(0, 1) \)
2. if \( U_1 \leq \left(\frac{\mu_1}{\mu_2}\right)^k \) then
3. Sample \( U_2 \sim \text{Unif}(0, 1) \)
4. Compute the transmission time \( s = F_{T_k}^{-1}(U_2) \), where \( F_{T_k}(s) = P(T_k \leq s) \mid T_k < \infty \)
5. else Set the transmission time \( s = \infty \)
6. end if

The integral equation \( \dot{S}(t) \) can be rewritten as

\[
\dot{S}(t) = -\int_0^t \beta S(t) I(t) P(T_k = t - \tau) d\tau
\]

\[= -\int_0^t \beta S(t) I(t) \sum_{j=0}^{\infty} g(j/2) e^{-\mu(t-\tau)} (\mu(t-\tau))^{j+k}/(j+k)! d\tau
\]

(14)

where \([x]\) denotes the largest integer less or equal to \( x \) and \( g(j/2) \) can be computed recursively using Eq. 9.

Now we consider Eq. (14) in the discrete time, i.e., \( t = 0, 1, \ldots \). Note that the time unit in discrete time \( t \) is the time interval of information propagation among social media network users, while the time unit in discrete time \( s \) (Eq. (10,11)) is the time interval of block mining on the blockchain. We assume that the rate of block mining is faster than that of information propagation. The aim is to provide the user with a feeling of system reacting instantaneously and swiftly. Indeed, the blockchain technology is capable to meet this requirement, with Stellar and Solana achieving 2–4 seconds per block \cite{6}. On the other hand, the information propagation on Twitter or Facebook may take minutes.

Thus we define the ratio \( l = \frac{\Delta t}{\Delta \tau} \) which represents the different time scales of information propagation and block mining. Eq. (14) can be rewritten as follows

\[
Pr_{T_k < t} (t) = P(T_k \leq ls)
\]

(15)

Then we time discretize the integral equation (14) resulting in the discrete time difference equation

\[
S(t+1) = S(t) - \sum_{i=0}^{l-1} \beta S(i+1) \sum_{j=0}^{l-1} P(T_k = (t+1-i)l-j)
\]

(16)

Similarly, the discrete-time difference equations for the infected and recovered population are

\[
I(t+1) = I(t) + \sum_{j=0}^{l-1} \beta S(t+1) \sum_{i=0}^{l-1} P(T_k = (t+1-i)l-j) - \alpha I(t)
\]

R(t+1) = R(t) + \alpha I(t)

(17)

(18)

Here \( \alpha > 0 \) is the recovery rate denoting the probability that infected individuals recover and remain permanently immune to the infection.

6. https://alephzero.org/blog/what-is-the-fastest-blockchain-and%2Dwhy-analysis-of-43-blockchains/
3 SIR Model on a Multi-community Network

In this section, we generalize the SIR dynamics to a multi-community network. We model the network using a stochastic block model (SBM), where users are partitioned into three communities representing different contact rates and recovery rates. We use stochastic simulation based on Section 2 to analyze how the community structure and rate heterogeneity affect misinformation propagation in the network.

3.1 A Two-community SBM Network

Consider a social media network represented by an undirected graph $G = (V, E)$, where $V$ is the node set representing the users and $E$ is the edge set representing the users’ friendship. The adjacency matrix $A = [a_{ij}]_{N \times N}$ is a binary valued matrix where
\[
a_{ij} = \begin{cases} 
1, & \text{nodes } i \text{ and } j \text{ are connected} \\
0, & \text{otherwise}
\end{cases} (19)
\]

We partition $V$ into two disjoint communities, i.e., $V = V_1 \cup V_2$ and $V_1 \cap V_2 = \emptyset$, where $V_1 (V_2)$ represents the community with strong (weak) resilience to misinformation. Just as population in different age groups have different infection fatality rate to the epidemic [21], users in different communities have different contact rates and recovery rates to misinformation, i.e., contact rates $\beta_1 < \beta_2$ and recovery rates $\alpha_1 > \alpha_2$, where $\beta_m$ and $\alpha_m$ denote the rates for users in community $V_m$, $m \in \{1, 2\}$.

Let $c_i$ denotes node $i$’s community label. $c_i = 1$ if $i \in V_1$ and $c_i = 2$ if $i \in V_2$. We also define the probability matrix $P$ as a symmetric $2 \times 2$ matrix which specifies the probability that two nodes connect based on their communities. That is, the probability that nodes $i$ and $j$ are connected is $P_{c_i c_j}$.

The above two-community SBM model captures the community structure of social media networks and the variation of contact rates and recovery rates for different communities. This idea is similar to that in [9] that individuals are more likely to connect with others that have similar susceptibility.

3.2 Stochastic Simulation of Misinformation Propagation

In this subsection, we use agent based modeling to construct the SIR dynamics on a multi-community social media network. The model that we propose fully exploits the graph structure compared with the mean field approximation in Eq. 1.

In the SIR model, a user can be in one of the three states {susceptible, infected, recovered}. We use $S(t)$, $I(t)$, $R(t)$ to represent the set of nodes in each state at time $t \in \{0, 1, \cdots\}$. The time is discrete and the unit of time can be set in accordance with the propagation speed of misinformation. We set it to minute in the numerical simulation of Section 4. We also extend the two-community model to the multi-community setting with $M$ communities.

We define $N_{i, m}(t), m \in \{1, \cdots, M\}$ as the set of node $i$’s neighbors from community $V_m$ that are infected at time $t$, i.e.,
\[
N_{i, m}(t) = \{j | A_{ij} = 1 \land j \in I(t) \land c_j = m\}. (20)
\]

First consider a social media network without the blockchain. The CDF of the transmission time is a unit step function at 0, i.e., a susceptible user is infected as soon as she contacted an infected user as shown in (5). Define $x_i(t) \in \{S, I, R\}$ as the state of user $i$ at time $t$. Then the transition matrix is:
\[
P(x_i(t+1)|x_i(t)) = \begin{bmatrix}
P_{S \rightarrow I} & 1 - P_{S \rightarrow I} & 0 \\
0 & P_{I \rightarrow R} & 1 - P_{I \rightarrow R} \\
0 & 0 & 1
\end{bmatrix}. (21)
\]

where
\[
P_{S \rightarrow I} = \prod_{m \in M} (1 - \beta_m) |N_{i, m}(t)|
\]
\[
P_{I \rightarrow R} = \alpha_{c_i}
\]

In comparison, the stochastic simulation on the multi-community network with blockchain is shown in Algorithm 6. For each infected neighbor $j$ of a susceptible user $i$, $j$ contacted $i$ with probability $\beta_j$, which is the contact rate of $j$’s community. If they contacted, $i$ will not become infected immediately as in the typical SIR model; Instead, a transmission time will be sampled from the distribution specified in Section 2.3 indicating how long it takes for $i$ to be infected due to the contact with $j$. For each of $i$’s contact with her infected neighbors, a transmission time is sampled, and the minimum of them will be the time for $i$ to get infected.

In Section 4, we will use numerical simulation to compare the results of misinformation transmission in social media networks with and without the blockchain protocol.

4 Numerical Illustration of the Blockchain Enabled Social Network

In this section, we simulate the SIR dynamics of misinformation transmission on a three-community social media network (Section 3.1) according to Algorithm 2. We compare the simulation results with and without the blockchain in the network. The results demonstrate that blockchain can flatten the curve of the population affected by misinformation in social media networks. We also show that with the same proportion of honest miners in the blockchain system, some communities are more resilient to misinformation due to higher recovery rate and lower contact rate. This indicates the Cannikin Law [22] (wooden bucket theory) when applying the distributed ledger to a multi-community network – The minimum proportion of honest miners to avoid all the users getting infected is determined by the community most fragile to misinformation.

Parameter Estimation from Twitter Datasets: We use Twitter datasets of trending hashtags [5] to obtain realistic SIR model parameters. The data is collected by querying certain hashtags to see what information is spreading virally across the site. The SIR model, which is used in our study on misin-
Fig. 2: This figure shows the SIR dynamics in the social media network with and without the blockchain protocol. The top subfigure shows the infected population $I(t)$ while the bottom subfigure shows the effective infection rate $\lambda(t)$. The blockchain enabled social media network (red star) flattens the curve of the infected population and has a smaller effective infection rate compared with the network without blockchain protocol (blue up triangle). The main takeaway is that the blockchain enabled network has stronger resilience to misinformation.

Fig. 3: The figure shows the SIR dynamics in each community of the network with the blockchain protocol. The curves represent the susceptible proportion $s(t)$ (blue), the infected proportion $i(t)$ (red), and the recovered proportion $r(t)$ (green). Community 1 has the strongest resilience to misinformation while community 3 has the weakest resilience. All users in community 3 get infected in the end. The key takeaway from this figure is as follows: Communities will react differently to the dissemination of misinformation when the distributed ledger is used to a multi-community network. To avoid infecting all users in the community most vulnerable to misinformation, a minimal proportion of honest miners $\mu_2$ can be chosen accordingly.

Formulation transmission, is also suitable here for measuring the trendiness of a topic.

**Model Parameter:** In the simulation, the number of users in the three community are 100, 60, 40 respectively, and the $3 \times 3$ probability matrix is

$$B = \begin{bmatrix} 0.2 & 0.015 & 0.012 \\ 0.015 & 0.2 & 0.02 \\ 0.012 & 0.02 & 0.2 \end{bmatrix}$$

which corresponds to the situation where users have more intra-community connections than inter-community ones. The three communities have different contact rates 0.02, 0.1, 0.5 and recovery rates 0.1, 0.06, 0.03. Because the network maintains one unique blockchain, the three communities share the same transmission-time distribution

$$P(T_k \leq s|T_k < \infty)\text{.}$$

We set $k = 2$, $\frac{\mu_1}{\mu_1 + \mu_2} = 0.3$, $l = 5$.

**Evaluation Metrics:** We use the effective infection rate [23] to quantify how the network is affected by misinformation. The effective infection rate is $\lambda(t) = \frac{\beta(t)}{\alpha(t)}$, i.e., the ratio of the empirical contact rate and the empirical recovery rate.

On a network, new infections at time $t$ can be modeled as

$$S(t) - S(t + 1) = \beta(t) \frac{kS(t)I(t)}{N}$$

where $k$ is the average node degree, $\frac{kS(t)I(t)}{N}$ refers to the total number of connections between infected and susceptible users. Then the empirical contact rate $\beta(t)$ measures the likelihood of a contact, which quantifies the "infectivity" of
the misinformation and can be computed as
\[ \beta(t) = -N \frac{S(t+1) - S(t)}{kS(t)I(t)} \] (25)

Similarly, the empirical recovery rate \( \alpha(t) \) quantifies how likely an infected user gets recovered from the misinformation
\[ \alpha(t) = \frac{R(t+1) - R(t)}{I(t)} \] (26)

Results:
- **Blockchain Improves a Network’s Resilience:** Figure 2 compares the population dynamics and the effective infection rate in the social media network with and without the blockchain protocol. In the network with blockchain, \( I(t) \) curve of \( I(t) \) is flattened, with lower number of infected users and the infections distributed along a longer time. The effective infection rate is also smaller. This suggests that the blockchain enabled social media network slows down the misinformation transmission, and also lowers the number of the infected population.

- **Different Communities Behave Differently to Misinformation:** Figure 5 compares the performance of the three communities in the network with blockchain. Specifically, Fig. 5 displays the ratio of each population, with \( s(t) = \frac{S(t)}{N(t)} \) representing the ratio of susceptible individuals in the community. \( i(t) \) and \( r(t) \) are defined similarly. Due to the pre-defined parameters, community 1 has the strongest resilience to misinformation: the peak value of infected portion is less than 0.2, and around 0.4 of the total population is unaffected in the end. On the other hand, community 3 has the weakest resilience: the peak value of infected portion is higher than 0.7, all the population are affected in the end, and the misinformation exist for around 110 minutes, much longer than that of community 1.

5 CONCLUSIONS

We constructed a blockchain enabled social media network which examines social media postings using a blockchain transaction confirmation method. The social media postings are processed as blockchain transactions, with a higher proportion of honest miners in the network ensuring that misinformation takes longer to propagate. We modeled misinformation as double-spend attacks in the blockchain and derive its transmission-time distribution. This distribution is incorporated into the SIR model. Then we exploited stochastic simulation algorithm of misinformation dynamics on a multi-community network. In numerical experiments, we demonstrated that the proposed blockchain enabled social media network has stronger resilience to misinformation than the network without blockchain. We also compared the performance of different communities in the network.

APPENDIX A
PROOF OF LEMMA 1
To prove Lemma 1, we use the Bertrand ballot theorem.

**Lemma 2.** (Bertrand’s ballot theorem [24]) In a random permutation of \( n \) values +1 and \( m \) values -1, where \( n > m \), the probability that for every \( i = 1, \ldots, n + m \), the first \( i \) elements of the permutation always contain more values +1 than -1 is \( \frac{n-m}{n+m} \).

Now we derive the distribution of \( N_k \). \( N_k = k + 2i \Rightarrow S_{k+2i} = k \) and for \( j < k + 2i \), \( S_j < k \). Recall that \( S_{k+2i} = \sum_{j=1}^{k+2i} X_j \) where \( X_j = 1 \) with probability \( p \). Then in the ordered sequence \( X_{k+2i}, X_{k+2i-1}, \ldots, X_1 \), the cumulative number of 1 is always greater than that of -1. Therefore,
\[
P(N_k = k + 2i) = P(S_{k+2i} = k)P(N_k = k + 2i|S_{k+2i} = k) = \left(\frac{k+2i}{k+i}\right) p^{k+1}(1-p)^i \frac{k}{k+2i}
\] (27)

where \( P(N_k = k + 2i|S_{k+2i} = k) = \frac{k}{k+2i} \) is derived from the Bertrand ballot lemma (Lemma 2 above).
APPENDIX B

PROOF OF THEOREM 1

Let $N(s) = N_1(s) + N_2(s)$ denote the sum of blocks mined by both dishonest miners and honest miners. Because $N_1(s)$ and $N_2(s)$ operate independently, $\{N(s), s \geq 0\}$ is a Poisson process with rate $\mu = \mu_1 + \mu_2$. From the definition we note that $T_k$ is distributed as the time of the $N_k$-th event of $N(s)$, i.e.,

$$P(T_k = s) = P(N_k = N(t))$$

Conditioning on $N(t) = k + j$ and using the fact that $N(t)$ and $N_k$ are statistically independent yields

$$P(T_k = s) = \sum_{j=0}^{\infty} P(N_k = k + j) \frac{e^{-\mu t}(\mu t)^{k+j}}{(k+j)!}$$

Similarly,

$$P(T_k \leq s) = \sum_{j=0}^{\infty} P(N_k \leq k + j) \frac{e^{-\mu t}(\mu t)^{k+j}}{(k+j)!}$$

which yields $[11]$.

APPENDIX C

LIST OF BLOCKCHAIN ENABLED SOCIAL MEDIA NETWORKS

Real world examples of blockchain enabled social media networks include

- Steemit (https://steemit.com/) is built on the Steem blockchain, a decentralized reward platform for publishers to monetize content and grow community. Those who hold more Steem tokens have more decision power on community matters and reward distributions.

- Sola (https://sola.ai/) is a hybrid of media and social network which uses AI algorithms to feed quality content to the most interested users.

- Civil (https://civil.co/) is a community-owned network of journalists who use blockchain to establish transparency and trust. On Civil’s network, independent journalists create newsrooms where they add and share their content.

- onG.social (https://ong.social/) is a blockchain based social dashboard which supports community building and social interaction with cryptocurrency rewards. It runs on two blockchains, Ethereum and Wavesplatform.

- Sapien (https://www.sapien.network/) is a social news platform built on the Ethereum blockchain that gives users control of data and content.

REFERENCES

[1] J. Huang, L. Tan, S. Mao, and K. Yu, “Blockchain network propagation mechanism based on p4p architecture,” Security and Communication Networks, vol. 2021, 2021.

[2] A. Quyyum, J. Qadir, M. U. Januja, and F. Sher, “Using blockchain to rein in the new post-truth world and check the spread of fake news,” IT Professional, vol. 21, no. 4, pp. 16–24, 2019.

[3] S. Paul, J. I. Joy, S. Sarker, S. Ahmed, A. K. Das et al., “Fake news detection in social media using blockchain,” in 2019 7th International Conference on Smart Computing & Communications (ICSCC). IEEE, 2019, pp. 1–5.

[4] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, and A. Scala, “The covid-19 social media infodemic,” Scientific Reports, vol. 10, no. 1, pp. 1–10, 2020.

[5] J. Skaza and B. Blais, “Modeling the infectiousness of twitter hashtags,” Physica A: Statistical Mechanics and its Applications, vol. 465, pp. 289–296, 2017.

[6] D. Baqae, E. Farhi, M. J. Mina, and J. H. Stock, “Reopening scenarios,” National Bureau of Economic Research, Tech. Rep., 2020.

[7] W. Gou and Z. Jin, “How heterogeneous susceptibility and recovery rates affect the spread of epidemics on networks,” Infectious Disease Modelling, vol. 2, no. 3, pp. 353–367, 2017.

[8] M. Lachiany and Y. Louzoun, “Effects of distribution of infection rate on epidemic models,” Physical Review E, vol. 94, no. 2, p. 022409, 2016.

[9] D. Smilkov, C. A. Hidalgo, and L. Kocarev, “Beyond network structure: How heterogeneous susceptibility modulates the spread of epidemics,” Scientific reports, vol. 4, no. 1, pp. 1–7, 2014.

[10] A. Stanoev, D. Tepevski, and L. Kocarev, “Modeling the spread of multiple concurrent contagions on networks,” PloS one, vol. 9, no. 6, p. e95669, 2014.

[11] Y. Chen, Q. Li, and H. Wang, “Towards trusted social networks with blockchain technology,” arXiv preprint arXiv:1801.02796, 2018.

[12] D. Fu and L. Fang, “Blockchain-based trusted computing in social network,” in 2016 2nd IEEE International Conference on Computer and Communications (ICCC). IEEE, 2016, pp. 19–22.

[13] C. Tessone, P. Tasca, and F. Iannelli, “Stochastic modelling of blockchain consensus,” Available at SSRN 3865040, 2021.

[14] S. Bararisi, “Application of blockchain and infection processes in graphs to mitigate content poisoning attacks in information-centric networks,” Ph.D. dissertation, 10 2020.

[15] M. Saad, A. Ahmad, and A. Mohaisen, “Fighting fake news propagation with blockchains,” in 2019 IEEE Conference on Communications and Network Security (CNS). IEEE, 2019, pp. 1–4.

[16] W. O. Kermack and A. G. McKendrick, “Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity,” Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character, vol. 138, no. 834, pp. 55–83, 1932.

[17] F. Brauer, C. Castillo-Chavez, and Z. Feng, Mathematical models in epidemiology. Springer, vol. 32.

[18] Z. Feng, D. Xu, and H. Zhao, “Epidemiological models with non-exponentially distributed disease stages and applications to disease control,” Bulletin of mathematical biology, vol. 69, no. 5, pp. 1511–1536, 2007.

[19] C. Pinzón and C. Rocha, “Double-spend attack models with time advantage for bitcoin,” Electronic Notes in Theoretical Computer Science, vol. 329, pp. 79–103, 2016.

[20] M. Brown, E. Peköz, and S. Ross, “Blockchain double-spend attack duration,” Probability in the Engineering and Informational Sciences, vol. 35, no. 4, pp. 858–866, 2021.

[21] D. Acemoglu, V. Chernoizhukov, I. Werning, and M. D. Whinston, “Optimal targeted lockdowns in a multigroup sir model,” American Economic Review: Insights, vol. 3, no. 4, pp. 487–502, 2021.

[22] D. Tilman, Resource Competition and Community Structure.(MPB-17), Princeton university press, 2020.

[23] P. Van Mieghem, “The viral conductance of a network,” Computer Communications, vol. 35, no. 12, pp. 1494–1506, 2012.

[24] W. Feller, An introduction to probability theory and its applications, vol 2. John Wiley & Sons, 2008.