Incentive Mechanism for Redactable Blockchain Governance: An Evolutionary Game Approach

Jiaxiang Sun, Rong Zhao, Haoran Yin, and Wei Cai, *Senior Member, IEEE*

***Abstract*—Blockchain technology has garnered significant attention in recent years due to its capacity to offer secure and transparent transactional systems. However, the technology's inherent immutability can present challenges in specific scenarios. While earlier research has concentrated on the development of redactable blockchain, existing solutions have primarily focused on the modification mechanism, often overlooking the critical element of an incentive mechanism for governance, which is paramount for ensuring the security of redactable blockchain. Some previous researches have explored the design of incentive mechanisms, but these studies exhibit certain shortcomings. To promote active participation, we have designed an incentive mechanism rooted in evolutionary game theory for stakeholders in redactable blockchain, aiming to facilitate the governance of redactable blockchain. Furthermore, we have conducted a comprehensive simulation founded on game-theoretic analysis. The results substantiate the effectiveness of our redactable blockchain incentive mechanism in achieving its intended objectives.**

***Index Terms*—Blockchain, blockchain governance, evolutionary game theory, incentive mechanism, redactable blockchain.**

I. Introduction

BLOCKCHAIN is an emerging distributed digital ledger technology that has attracted significant attention due to its unique features of transparency, decentralization, security, and immutability. Originally designed for cryptocurrencies, blockchain's capabilities extend to many other applications, such as gaming, supply chain monitoring, and digital twin.

Immutability is a crucial feature of blockchain, which ensures that the records on the blockchain cannot be modified or deleted. However, the immutability of blockchain can lead to problematic consequences in certain circumstances. In blockchain systems that enable smart contracts, such as Ethereum, immutability prevents vulnerabilities in smart contracts from being fixed. Instead, the contracts must be redeployed, and any losses resulting from vulnerabilities cannot be recovered. For example, the Decentralized Autonomous Organization (DAO) attack on Ethereum exploited a reentrancy vulnerability in a smart contract, causing a loss of 3.6 million Ether (ETH) [1]. To recover from this loss, Ethereum performed a hard fork and rolled back the state. However, in today's blockchain ecosystem, where multiple assets are supported on different chains and cross-chain services are developing, reaching a consensus on hard forks and state rollbacks is challenging. Immutability in blockchain systems, even those without smart contract support (e.g., Bitcoin), can lead to certain issues. Malicious users can exploit this feature by paying a small fee to publish illegal and harmful content on the blockchain through special transactions. The International Criminal Police Organization (INTERPOL) has reported that the Bitcoin blockchain contains arbitrary content in the form of illegal material, including child pornography, copyrighted material, and sensitive information [2]. Miners who process these transactions face potential legal risks. Current blockchain platforms are unable to comply with the European Union's GDPR and other regulations that require users' data to be forgotten, meaning the ability to delete users' data.

While the immutability of the blockchain presents certain challenges, its decentralized and transparent nature enables a wide range of promising applications. For example, in the industrial sector, blockchain technology can significantly bolster trust within social manufacturing environments [3]. Furthermore, it offers a robust and transparent framework for decentralized autonomous manufacturing systems [4], [5]. Moreover, in digital twin, blockchain technology can be used to synchronize data with trustworthiness and traceability, thus enhancing the credibility of digital twins. To support the development of these applications, the design of a redactable blockchain protocol is crucial. Redactable blockchains can be beneficial for certain situations where data privacy regulations, such as the General Data Protection Regulation (GDPR), mandate the capability to modify or remove personal information. They can also be valuable in cases where mistakes need to be rectified after the fact without compromising the overall integrity of the ledger, particularly in industrial area and digital twin. Several studies have focused on redactable blockchain protocols in the literature. Ateniese et al. [6] introduced a redactable blockchain that utilizes chameleon hashes, while Deuber et al. [7] proposed redactable blockchain

Manuscript received 26 October 2023; revised 24 February 2024; accepted 21 April 2024. This work was supported in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515012323; in part by the Shenzhen Science and Technology Program under Grant JCYJ20210324124205016; in part by the Open Topics of Key Laboratory of Blockchain Technology and Data Security, The Ministry of Industry and Information Technology of the People's Republic of China; and in part by the CUHK(SZ)-White Matrix Joint Metaverse Laboratory. *(Corresponding author: Wei Cai.)*

The authors are with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China (e-mail: jiaxiangsun@link.cuhk.edu.cn; rongzhao@link.cuhk.edu.cn; haoranyin@link.cuhk.edu.cn; caiwei@cuhk.edu.cn).

Digital Object Identifier 10.1109/TCSS.2024.3398044

protocols utilizing consensus-based voting. Li et al. [8] proposed an instant redactable blockchain that allows for immediate modifications by utilizing committee selection. Committee members (CMs) are required to allocate resources for the editing and voting processes without receiving any direct benefits in return. This lack of compensation may discourage rational CMs from participating in the editing process, which is detrimental to the governance of redactable blockchain. Thus, an incentive mechanism is essential to motivate CMs to actively participate in the consensus-building process and to compensate them for their expenses related to voting and editing, thereby enhancing the governance of redactable blockchain. For instance, incentive mechanism can encourage participants to maintain and update the industrial blockchain responsibly, fostering a self-regulating ecosystem. Redactable blockchain with incentive mechanism not only bolsters trust among stakeholders in complex industrial environments [3] but also aligns with the dynamic nature of manufacturing system [4], [5].

However, traditional blockchain incentive mechanisms face several challenges in redactable blockchain. First, redactable blockchains introduce a diverse array of participant roles beyond the common users and miners typically found in traditional blockchain systems. Notably, CMs and users who propose edit requests (ERs) assume pivotal roles in the ecosystem, presenting new challenges related to balancing rights and responsibilities. In redactable blockchain, CMs may have varying preferences when it comes to ERs, influenced by their perspectives, interests, and objectives. The approval of their preferred ERs provides them with utility, which could be in the form of financial gains or enhanced reputation. This utility significantly impacts their willingness to actively participate in the blockchain system. Furthermore, traditional blockchain incentive mechanisms often overlook the evaluation of the behavior of ER submitters. However, these individuals hold a crucial role in the system as they drive changes by submitting ERs. Second, the editing and voting processes themselves impose additional resource and time demands. Participants in redactable blockchains may find themselves engaged in a broader spectrum of tasks beyond the traditional roles of mining and validation. Consequently, incentive mechanisms must consider and compensate for these direct operational costs. Wang et al. [9] made an attempt to address this issue by introducing a contract-based incentive mechanism for redactable proof of stake (PoS) blockchains. However, their incentive mechanism setting is not well suited, as it mandates miners to bear additional expenses to carry out the redaction process. Furthermore, their incentive structure contains a potential loophole, which could potentially enable collusion between specific users and CMs, thereby introducing a vulnerability to conspiracy attacks. Therefore, it is imperative to conduct further research to thoroughly explore and comprehensively address this aspect.

Evolutionary game theory, a branch of game theory, examines the evolution of strategies employed by agents over time. It is particularly valuable in situations where interaction outcomes are influenced by the prevalence of strategies within a population. Given that decisions in blockchain are commonly determined by the preference of the majority, evolutionary game theory is well suited for investigating this domain. In the blockchain ecosystem, evolutionary game theory can be used to analyze strategic interactions [10], enhance security [11], promote cooperation [12], and design incentive mechanism [13]. Thus, evolutionary game theory is an effective approach to addressing current challenges.

This article presents a novel incentive mechanism for stakeholders in redactable blockchain to address the aforementioned challenges comprehensively. We integrate this incentive mechanism into the redactable blockchain protocol, promoting governance and enhancing security in the system. Additionally, we develop an evolutionary game theory framework for redactable blockchain, which allows us to analyze the strategic interactions of participants within the proposed system. Our analyses demonstrate the effectiveness of the incentive mechanism, highlighting its potential to bolster the security of redactable blockchain significantly.

This article makes three contributions to advance the field of redactable blockchain incentive mechanism.

1) *An Incentive Mechanism for CM in Redactable Blockchain:* We propose an innovative incentive mechanism designed specifically for CM in redactable blockchain, addressing the lack of incentives in most prior research. Specifically, our approach aims to encourage CMs, tasked with the critical role of voting on ERs, to engage actively in the decision-making process regarding these requests. This approach promotes efficient governance of redactable blockchain.
2) *A Game-Theoretic Analysis of Redactable Blockchain Incentive Mechanism:* Leveraging evolutionary game theory, we conduct a comprehensive analysis of the proposed redactable blockchain incentive mechanism. This approach enables us to examine the strategic interactions between users submitting ERs and CMs responsible for voting on these requests, providing valuable insights into their behavior and decision-making processes.
3) *A Simulation of Redactable Blockchain Incentive Mechanism:* We perform a simulation based on the game-theoretic analysis to validate the results and visualize the evolutionary process of the system. The simulation serves to demonstrate the outcomes of our game theory analysis.

The article is organized as follows. In Section II, we review the relevant literature. Section III provides preliminaries of redactable blockchain. In Section IV, we present our incentive mechanism for redactable blockchain and conduct an analysis using evolutionary game theory. Section V includes the simulation of the proposed incentive mechanism. Section VI discusses the future research directions for redactable blockchains. Finally, Section VII concludes the work.

II. Related Work

A. Redactable Blockchain

Redactable blockchain have gained significant attention in recent years, and numerous technical implementations have

been proposed. These implementations primarily fall into two categories: chameleon hash-based and nonchameleon hash-based methods. Ateniese et al. [6] introduced a redactable blockchain using chameleon hashes. In their approach, trusted entities possessing the trapdoor can efficiently compute collisions and maintain connectivity after modifying block data. Zhang et al. [14] proposed threshold trapdoor chameleon hash, where each node in the blockchain holds a portion of the trapdoor and modifications can only be made with the signature agreement of a certain number of nodes. Jia et al. [15] proposed a decentralized chameleon hash, where multiple nodes collaborate to calculate the trapdoor for modification. Voting mechanisms are commonly employed in nonchameleon hash-based methods. Studies such as Deuber et al. [7] and Marsalek and Zefferer [16] have proposed redactable blockchain protocols that employ a consensus-based voting process for modifications. Miners include information regarding proposed edits within the block, effectively casting their votes. However, the utilization of on-chain voting, presents a challenge of slow modification. To overcome this challenge, Li et al. [8] proposed a solution by separating the voting process from the underlying consensus protocol. This separation allows for the election of a committee responsible for deciding whether to approve a modification, resulting in faster voting modifications.

Many existing studies primarily concentrate on establishing editing rules for blockchain blocks without adequately addressing the critical issue of redactable blockchain governance. While these redactable blockchains have been demonstrated to be secure, ensuring accuracy, consensus, and validity with a high probability [6], [7], [8], [14], [15], [16], they still face governance-related security challenges. In redactable blockchain systems, incentive mechanisms can play a crucial role in enhancing security. Without incentive mechanisms, redactable blockchains are vulnerable to issues such as dishonest behavior and collusion attacks. Stakeholders in redactable blockchain may become dishonestly, such as engaging in illicit modifications to manipulate records to their advantage. By implementing incentive mechanisms, these challenges can be mitigated, as they incentivize honest behavior and penalize dishonest actions. In addition to discouraging dishonest behavior, incentive mechanisms are essential for fostering active participation in redactable blockchains. Active participation is vital for the effective functioning of the redactable blockchain system. Incentive mechanisms play a key role in promoting this engagement by incentivizing participants for their contributions toward maintaining the network's health and integrity.

While the blockchain domain has seen the emergence of some incentive mechanisms, such as the delegated proof of stake (DPoS) blockchain [17] and others, it is important to note that these incentive mechanisms exhibit limited universality [9]. Due to the diversity of roles, the associated operational costs, and the inherent consensus complexities in redactable blockchains, traditional blockchain incentive mechanisms are not applicable to this context. Wang et al. [9] were pioneers in addressing this issue by introducing a contract-based incentive mechanism specifically tailored for redactable PoS blockchains. However, their incentive mechanism's configuration has shown its limitations, as it imposes additional costs on validators when executing the redaction process. Furthermore, the incentive structure does not account for potential collusion between certain users and CMs, thereby introducing a vulnerability to conspiracy attacks. Consequently, there is a pressing need for further research to conduct a comprehensive investigation and formulate robust solutions to effectively address this aspect.

B. Evolutionary Game Theory

Evolutionary game theory is a branch of game theory that combines concepts from evolutionary biology and rationalistic economics to study the evolution of strategies in competitive situations. While traditional game theory assumes complete rationality and focuses on static equilibrium, evolutionary game theory takes into account the bounded rationality of agents and explores the dynamics of strategy evolution over time [18]. An important concept in evolutionary game theory is the evolutionary stable strategy (ESS), which refers to a strategy that is resistant to invasion by any alternative strategy and persists over time [19]. Replicator dynamics (RD) is another method used to analyze evolutionary game theory [20]. RD provides a mathematical framework for studying how the frequencies of different strategies in a population change over time.

Numerous scholars have utilized evolutionary game theory to investigate different facets of blockchain networks. For instance, Liu et al. [10] employed evolutionary game theory to analyze the process of mining pool selection in blockchain networks, revealing the conditions for the network to admit a unique evolutionary stable state. Ni et al. [21] applied evolutionary game theory to study the consensus-formation process in scalable blockchain with shared consensus, exploring the consensus-formation mechanism in a permissionless network of blockchain. Motepalli and Jacobsen [13] developed a reward mechanism for blockchain networks using evolutionary game theory, highlighting the importance of penalties in maintaining ledger integrity. Zhang and Wu [12] utilized evolutionary game theory to examine the cooperation mechanism among participants in blockchain networks, emphasizing the need to reduce profit expectations of malicious participants for maintaining ledger integrity. However, to the best of our knowledge, no studies have yet applied evolutionary game theory to analyze redactable blockchain.

III. Preliminaries

Voting-based redactable blockchain is a category of redactable blockchain, as proposed by Deuber et al. [7]. This mechanism utilizes voting to approve ERs within the blockchain network. When a user submits an ER, miners or nodes cast their votes on it. If the votes in favor of an ER exceed a predefined threshold, typically set at $1/2$, the ER is approved. To preserve the integrity of the blockchain, an additional field is added to the block header, storing the original Merkle root, which allows for the calculation of the block hash after modification. Notably, Li et al. introduced a committee selection process to the redactable blockchain to achieve instant redaction [8]. In redactable blockchain, there

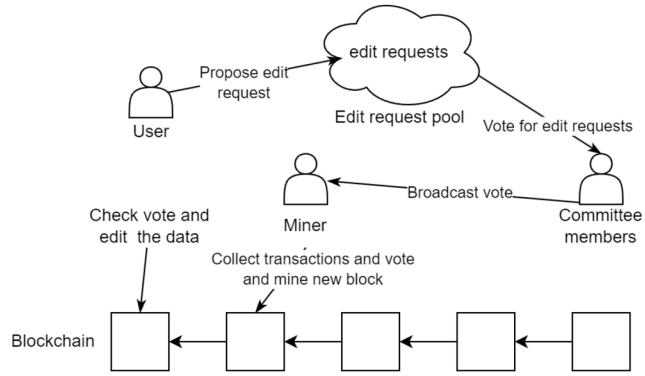


Fig. 1. Redactable blockchain system.

are three types of stakeholders. The first type comprises users who utilize the blockchain and submit ERs. The second type includes miners in proof-of-work (PoW) or validator nodes in PoS systems, who are tasked with proposing new blocks. Within the group of miners or nodes, certain individuals are chosen to serve as CMs. These CMs hold the responsibility of voting on the ERs.

As Fig. 1 shows, the redaction process comprises five steps as follows: 1) a user submits an ER; 2) CMs are selected from miners or nodes within the blockchain; 3) CMs vote on the ER and broadcast it to the blockchain network; 4) a miner or leader in the blockchain proposes a new block that includes the votes; and 5) miners or nodes check votes and edit the data [8].

The vote validation process operates as follows: 1) upon the arrival of a new block, the miner or node initially verifies the block's validity; 2) subsequently, the miner or node identifies the votes within the block and validates whether they originate from CMs by verifying signatures; and 3) finally, the miner or node computes the total votes for an ER, and if approval votes exceed half of the total votes, then the miner or node will proceed to modify the data in accordance with said ER [8].

The process for selecting CMs in a blockchain-based system varies depending on the consensus mechanism employed. In PoS blockchain, a verifiable random function-based algorithm is used to select CMs based on their stake. Each unit of stake is treated as a subuser, and the algorithm randomly selects a hash value to determine the number of subusers a participant has in the committee. On the other hand, PoW blockchain selects CMs by collecting sufficient PoW puzzle solutions, which are proportional to the computational power of each participant. It is important to note that CMs are only selected at the beginning of each voting period, and the selection process is repeated for each period. More specific details regarding the design of this process can be found in [8].

To clarify what content can be edited in a redactable blockchain, we establish the following guidelines to define the scope of permissible edits.

1) Edits are exclusively limited to the data contained within the block body.
2) Elements directly associated with payments or states are strictly prohibited from edit.
3) Elements indirectly associated with payments or states are similarly disallowed from edit.
4) Elements whose modification could potentially instigate conflicts within the blockchain are expressly forbidden from edit.
5) Elements pertaining to votes are also unequivocally restricted from edit.

We can provide the definitions of honest and malicious users within the context of the redactable blockchain system. A user is classified as malicious if they propose ERs that violate the established guidelines for editing or if they submit a large number of ERs with the intent of launching an attack on the system. Conversely, a user is deemed honest if their ERs adhere to the guidelines and do not seek to disrupt the system's integrity.

Similarly, we can establish the definitions of honest and malicious CMs. A CM is considered honest when they vote in accordance with the provided guidelines for editing. On the other hand, a CM is labeled as malicious if they engage in a "lazy vote" by skipping the voting process altogether, or if they intentionally vote against the established guidelines.

IV. Redactable Blockchain Incentive Mechanism

A. Incentive Mechanism

To enhance the security and effectiveness of the redaction process, we have carefully designed an incentive mechanism that tackles three fundamental challenges. First, we aim to balance the rights and responsibilities of the diverse participant roles. This is achieved by providing rewards to CMs who evaluate ERs, while users are required to pay a handling fee upon submission to discourage frivolous requests. CMs receive a share of these fees for requests they have voted on correctly. Second, we aim to encourage active participation in voting and consensus building. This is accomplished by tying fees to votes cast in agreement with the final outcome. In doing so, the mechanism incentivizes CMs to thoroughly evaluate requests and actively contribute to building consensus. The comparison between our work and other works is shown in Table I. Wang et al. [9] utilize contract theory to establish an incentive mechanism between validators and CMs, focusing on incentives. In contrast, we propose an incentive mechanism between users and CMs, employing evolutionary game theory for analysis, taking into account both incentives and collusion.

As Fig. 2 shows, the incentive mechanism operates in three phases.

1) *Submission Phase*: Initially, users desiring to submit ER are required to pay a fee, which acts as a barrier against spam and ensures that the submitter has skin in the game. This fee has a minimum threshold. All users who wish to submit ERs must pay a fee that exceeds the minimum threshold, with one fee required for each ER. In the real-world implementation of redactable blockchain, a more sophisticated mechanism for determining fees, such as basing the fee amount on the complexity of the ER, can be employed. This balances the rights and responsibilities between common users and those who propose ERs.

TABLE I
COMPARISON OF REDACTABLE BLOCKCHAINS

	[6], [14], [15]	[7], [16]	[8]	[9]	Ours
Implementation method	Chameleon hash	Voting mechanism	Voting mechanism	Voting mechanism	Voting mechanism
Consensus mechanism	All consensus	PoW	PoW, PoS	PoS	PoW, PoS
Incentive mechanism	None	None	None	Validator-CM, considering incentive	User-CM, considering incentive and collusion

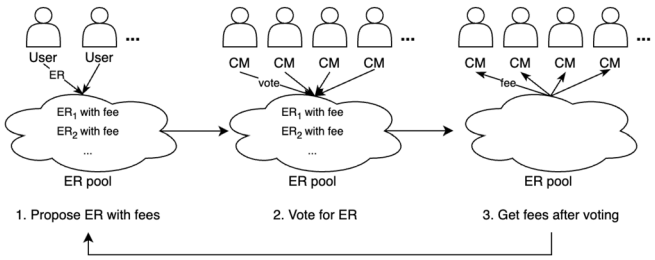


Fig. 2. Incentive mechanism in redactable blockchain.

2) *Voting Phase*: Upon successful submission, the requests are added to the ER pool, and CMs commence a voting period. CMs, who are selected based on a meritocratic system, have differential voting weight commensurate with their stake or contribution in the network. They should refer to the guidelines of permissible edits and make decisions based on the content of ER. The CMs broadcast their votes to miners or nodes, and these votes, whether approval votes or disapproval votes, will be recorded on the blockchain. A request is accepted if the total weight of approval votes exceeds more than half of the total vote weight.
3) *Reward Allocation Phase*: Once the voting period concludes, the handling fees collected during the submission phase are distributed among the CMs. These fees are specifically allocated to those who vote in alignment with the final decision on each request. To provide an example, when an ER is approved, CMs who vote in favor are entitled to a portion of the fee proportional to their voting weight in the overall approval weight. CMs who vote against this ER will receive no benefits. Conversely, if an ER is rejected, CMs who voted against it receive a share of the fee proportional to their voting weight. CMs who vote in favor will receive no compensation. Thus, CMs whose votes do not align with the final outcome incur costs for their votes but receive no benefit.

This incentive mechanism is intentionally designed to create a balanced system of rewards and risks, effectively motivating both users and committee members to take actions that bolster the integrity of the redactable blockchain.

B. Evolutionary Game Model for Redactable Blockchain

We will focus on the game played between users and committer members in a redactable blockchain. A normal form evolutionary game model, denoted as $G(N, M, U)$, comprises three components: a set of players represented by N, the strategies available to each player represented by M, and the corresponding payoffs for each strategy represented by U.

In this game, we consider a two-person game between user and CM. We define the player set as $N = \{\text{User, Committee member}\}$.

To analyze the game, we consider the strategy set for both the user and CM. The user's strategies are represented by $M_{\text{user}} = \{\text{Honest}, \text{Malicious}\}$. The "Honest" strategy entails proposing ERs according to the guidelines, while the "Malicious" strategy involves proposing ERs that violate the guidelines. Similarly, the CM's strategies are represented by $M_{\text{committee member}} = \{\text{Honest}, \text{Malicious}\}$. The "Honest" strategy involves allocating resources to maintain the ER pool, broadcasting ERs, and voting for ERs in line with the guidelines. The "Malicious" strategy, on the other hand, involves voting for ERs that go against the guidelines.

In the editing phase, each user submitting an ER is required to pay a fee. The fee of ER is denoted as f. These fees are apportioned according to the voting weight of the CM and the final decision on the ER. Only CMs whose vote concurs with the final decision qualify to receive part of the fee, corresponding to their weight. Aggregate weights of honest and malicious CMs are symbolized by w_h and w_m respectively, with $w_h + w_m = 1$. The outcome of the vote is defined by the function $g(x)$

$$g(x) = \begin{cases} 0, & \text{if } x \leq \dfrac{1}{2} \\ 1, & \text{if } x > \dfrac{1}{2}. \end{cases} \tag{1}$$

This function $g(x)$ is Heaviside step function. When $w_h > 1/2$, $g(w_h) = 1, g(w_m) = 0$ indicating that honest ERs are approved, malicious ones are rejected, and honest CMs are rewarded. Conversely, when $w_h \leq 1/2$, $g(w_h) = 0, g(w_m) = 1$ implies approval of malicious requests, rejection of honest ones, and rewards for malicious members.

Utilizing the Heaviside step function to express the reward function would result in discontinuous payoff functions, a characteristic that can only be observed in a hypothetical error-free environment [22]. To obtain a continuous approximation of the function, we will introduce a one-parameter smoothing function denoted as $g_k(x)$, where k is a smoothing parameter. The definition of the smoothing function is as follows:

$$g_k(x) = \frac{1}{1 + e^{-k(x-\frac{1}{2})}}. \tag{2}$$

This specific smoothing function will be employed whenever explicit smoothing functions are needed.

We now define the reward and payoff for the user and the CM. The reward for a honest user is the value of the approved

TABLE II
PARAMETERS IN MODEL

Variable	Description
w	Voting weight of each CM
w_h	Aggregate weight of honest CM
w_m	Aggregate weight of malicious CM
f	Fees for ER
c	Voting cost per CM
p	Preference of CM for ERs
v_h	Value derived from approval of honest ERs
v_m	Value derived from approval of malicious ERs
l	Loss incurred by CM from approved malicious ERs
b	Bribe offered by malicious user to malicious committee members
x, y	Probability of user and CM acting honestly

honest ERs, represented as $g(w_h)v_h$. For a malicious user, the reward is the value of the approved malicious ERs, represented as $g(w_m)v_m$. Thus, the payoff for a user, whether they adopt an honest or malicious strategy, is calculated as the reward minus the cost, expressed as $g(w_h)v_h - f$ or $g(w_m)v_m - f$ respectively. For a CM with weight w, the reward for following an honest strategy during an editing process is $g(w_h)(w/w_h)f$, while for the same weight but adopting a malicious strategy, the reward is $g(w_m)(w/w_m)f$. The cost to participate in the voting process is denoted by c. The preference of CM for ERs is defined as p. Hence, the payoff for a CM, whether they choose an honest or malicious strategy, is computed as the reward minus cost, represented by $g(w_h)(w/w_h)f + g(w_h)p - c - g(w_m)l$ or $g(w_m)(w/w_m)f - g(w_m)p - c - g(w_m)l$, respectively, where l represents the loss incurred by CMs due to approved malicious ERs or rejected honest ERs.

The following analysis investigates various scenarios in the payoff matrix based on the strategies adopted by both the user and the CM.

If the user chooses an honest strategy and the CM follows suit, they both receive rewards of $g(w_h)v_h - f$ and $g(w_h)(w/w_h)f + g(w_h)p - c - g(w_m)l$, respectively. On the other hand, if the CM opts for a malicious strategy, their payoff amounts to $g(w_m)(w/w_m)f - g(w_m)p - c - g(w_m)l$, while the user's payoff remains at $g(w_h)v_h - f$. Now let us consider the scenario where the user chooses the malicious strategy. If the CM selects the honest strategy, the user receives a payoff of $g(w_m)v_m - f$, and the CM receives a payoff of $g(w_h)(w/w_h)f + g(w_h)p - c - g(w_m)l$. If both the user and the CM opt for the malicious strategy, we assume a situation where users attempt to bribe CMs. The malicious user promises that if the malicious request is passed, an additional bribe denoted as b will be given to all CMs who voted in favor of the request. In this case, the payoff for the user and the CM will be $g(w_m)(v_m - b) - f$ and $g(w_m)(w/w_m)(f + b) - g(w_m)p - c - g(w_m)l$, respectively. The corresponding descriptions for parameters in this section are shown in Table II. Based on the above analysis, we have constructed the payoff matrix, which is presented in Table III.

To simplify further analysis, several assumptions about this system are made as follows.

1) The game consists of only two players: the user and the CM. Both players have bounded rationality and possess the ability to learn from each other to enhance their strategies.
2) Each player has two options to choose from: an honest strategy with a probability of x (for the user) or y (for the CM), and a malicious strategy with a probability of $1 - x$ (for the user) or $1 - y$ (for the CM).
3) Assuming uniformity among users and CMs, we can apply the law of large numbers, thus postulating that the distribution of choices between honest and malicious strategies mirrors the probabilities of CMs choosing similarly. Here, the fraction of CMs opting for honesty, denoted as y, equates to $w_h = y$. Conversely, the proportion choosing a malicious strategy is $1 - y$, i.e., $w_m = 1 - y$.

C. Expected Payoff of Game Players

The expected payoffs for both users employing honest and malicious strategies are derived from the given payoff matrix. For a user adhering to an honest strategy, the expected payoff is represented as

$$U_{\text{honest user}} = g(y)v_h - f. \tag{3}$$

In contrast, when opting for a malicious strategy, the user's expected payoff transforms into

$$U_{\text{malicious user}} = g(1-y)v_m - f - (1-y)g(1-y)b. \tag{4}$$

The user's average expected payoff is computed as

$$\begin{aligned}\bar{U}_{\text{user}} &= xU_{\text{honest user}} + (1-x)U_{\text{malicious user}} \\ &= xg(y)v_h + (1-x)(g(1-y)v_m - (1-y)g(1-y)b) - f.\end{aligned} \tag{5}$$

For a CM, the expected payoff when adopting an honest strategy is expressed as

$$U_{\text{honest committee member}} = g(y)\frac{w}{y}f + g(y)p - c - g(1-y)l. \tag{6}$$

On the other hand, when opting for a malicious strategy, the expected payoff is given by

$$\begin{aligned}U_{\text{malicious committee member}} = g(1-y)\frac{w}{1-y}f - g(1-y)p - c \\ + (1-x)g(1-y)\frac{w}{1-y}b - g(1-y)l.\end{aligned} \tag{7}$$

The average expected payoff for the CM is then computed as

$$\begin{aligned}\bar{U}_{\text{committee member}} &= yU_{\text{honest committee member}} \\ &\quad + (1-y)U_{\text{malicious committee member}} \\ &= wf + yg(y)p - (1-y)g(1-y)p \\ &\quad + (1-x)g(1-y)wb - c - g(1-y)l.\end{aligned} \tag{8}$$

TABLE III
PAYOFF MATRIX

		Committee Member	
		Honest (y)	Malicious ($1-y$)
User	Honest (x)	$g(w_h)v_h - f, g(w_h)\frac{w}{w_h}f + g(w_h)p - c - g(w_m)l$	$g(w_h)v_h - f, g(w_m)\frac{w}{w_m}f - g(w_m)p - c - g(w_m)l$
	Malicious($1-x$)	$g(w_m)v_m - f, g(w_h)\frac{w}{w_h}f + g(w_h)p - c - g(w_m)l$	$g(w_m)(v_m - b) - f, g(w_m)\frac{w}{w_m}(f+b) - g(w_m)p - c - g(w_m)l$

D. ESS Analysis Between User and Committee Member

The RD function is a valuable tool in evolutionary game theory for examining the frequency of a specific strategy choice. It is represented by a dynamic differential equation that captures the changes in the frequency of a given strategy within a population. The equation can be expressed as $(dx/dt) = x(U_i - \bar{U})$, where x denotes the frequency of a player selecting strategy i, U_i represents the expected payoff associated with strategy i, and $\bar{U}$ signifies the average payoff across all players.

By incorporating (3)–(5) into the dynamic equation, the dynamic equation for the user arises

$$\begin{aligned} F(x) &= \frac{dx}{dt} \\ &= x(U_{\text{honest user}} - \bar{U}_{\text{user}}) \\ &= x(1-x)(g(y)v_h - g(1-y)v_m + (1-y)g(1-y)b). \end{aligned} \tag{10}$$

Similarly, a dynamic equation for the CM is derived

$$\begin{aligned} F(y) &= \frac{dy}{dt} \\ &= y(U_{\text{honest committee member}} - \bar{U}_{\text{committee member}}) \\ &= (1-y)g(y)wf - yg(1-y)wf \\ &\quad + y(1-y)p - (1-x)yg(1-y)wb. \end{aligned} \tag{11}$$

The dynamic equations yield five equilibrium points by solving the expressions $F(x) = 0$ and $F(y) = 0$: $E_1(0,0)$, $E_2(0,1)$, $E_3(1,0)$, $E_4(1,1)$, and $E_5(x^*, y^*)$, where $x^* = 1 - (v_m - wfb/wb^2)$, $y^* = 1 - (v_m/b)$, and $0 < x^* < 1, 0 < y^* < (1/2)$. However, it is important to note that these equilibrium points may not necessarily be the evolutionary stability strategy (ESS). To determine their stability, we analyze the Jacobian matrix, which plays a crucial role in assessing the system's evolutionary stability.

The ESS of the RD system is determined by key indicators such as the trace ($\text{tr}(J)$) and the determinant ($\det(J)$) of the Jacobian matrix. If $\det(J) > 0$ and $\text{tr}(J) < 0$, the local equilibrium point (LEP) qualifies as an ESS. Conversely, if $\det(J) < 0$, a saddle point is formed. Additionally, if $\det(J) > 0$ and $\text{tr}(J) > 0$, the LEP becomes an unstable point [23]. By evaluating partial derivatives of the dynamic equations with respect to x and y, the Jacobian matrix, denoted by (9), shown at the bottom of the page, can be derived.

TABLE IV
DETJ AND TRJ FOR THE LEP

LEP	det$J$	tr$J$
$E_1(0,0)$	$-(b-v_m)(wf+wb-p)$	$(b-v_m)+(wf+wb-p)$
$E_2(0,1)$	$-v_h(wf+p)$	v_h+wf+p
$E_3(1,0)$	$(b-v_m)(wf-p)$	$-(b-v_m)+(wf-p)$
$E_4(1,1)$	$v_h(wf+p)$	$wf+p-v_h$

To simplify the analysis, we make the following assumption: $b < v_m$. This assumption is based on the rationale that rational users would not offer a bribe that exceeds the value they gain from the ERs. Under this assumption, it is noted that the point $E_5(x^*, y^*)$ does not exist as $y^* < 0$. Consequently, our analysis will focus on the remaining four points.

The determinant [$\det(J)$] and trace [$\text{tr}(J)$] for the four LEPs can be computed by substituting these points into the Jacobian matrix. The results are presented in Table IV.

Given these analysis and assumption, we can make a further analysis on the system stability.

Proposition 1: The RD system exhibits four local equilibrium points: $E_1(0,0)$, $E_2(0,1)$, $E_3(1,0)$, and $E_4(1,1)$.

Proof: An equilibrium point (x, y) is obtained when $F(x) = 0$ and $F(y) = 0$ [24]. All four points are attainable, thus establishing the validity of Proposition 1.

Proposition 2: The point $E_1(0,0)$ is an ESS if the conditions given by the following hold:

$$\begin{cases} wf + wb - p > 0 \\ b - v_m + wf + wb - p < 0. \end{cases} \tag{12}$$

Proof: When $wf + wb - p > 0$ and $b - v_m + wf + wb - p < 0$, the sign of det$J$ at $E_1(0,0)$ is positive, and the sign of tr$J$ at $E_1(0,0)$ is negative. This establishes $E_1(0,0)$ as an ESS, thereby confirming the assertion in Proposition 2.

In this scenario, the financial incentives offered to CM outweigh his personal preference regarding ER. As a result, CMs

$$J = \begin{bmatrix} (1-2x)(g(y)v_h - g(1-y)v_m + (1-y)g(1-y)b) & x(1-x)(g'(y)v_h + g'(1-y)v_m - g(1-y)b - (1-y)g'(1-y)b) \\ yg(1-y)wb & -g(y)wf + (1-y)g'(y)wf - g(1-y)wf + yg'(1-y)wf \\ & +(1-2y)p - (1-x)g(1-y)wb + (1-x)yg'(1-y)wb \end{bmatrix} \tag{9}$$

may act in opposition to their preferences, leading them to approve malicious ER submitted by user. The conditions create a situation where both the user and CM engage in malicious strategies, forming a collusion between them.

Proposition 3: The point $E_2(0,1)$ consistently represents a saddle point, implying that it cannot function as an ESS. It is unlikely that the user would choose the malicious strategy while the CM selects the honest strategy.

Proof: The sign of det$J$ at $E_2(0,1)$ is always negative. Consequently, $E_2(0,1)$ invariably qualifies as a saddle point, confirming the assertion in Proposition 3.

In this scenario, if the CM chooses to act honestly, the user has no incentive to behave maliciously. If the user acts maliciously, the CM stands to gain more by also engaging in malicious behavior. Consequently, the strategy profile $E_2(0,1)$ is not an ESS.

Proposition 4: The point $E_3(1,0)$ operates as an ESS if the conditions given by the following hold:

$$\begin{cases} wf - p < 0 \\ -(b - v_m) + (wf - p) < 0. \end{cases} \tag{13}$$

Proof: Given $wf - p < 0$ and $-(b - v_m) + (wf - p) < 0$, the sign of det$J$ at $E_3(1,0)$ is positive, and the sign of tr$J$ at $E_3(1,0)$ is negative. As a result, $E_3(1,0)$ qualifies as an ESS, validating the statement in Proposition 4.

In this scenario, the CM's preference for the ERt outweighs the financial incentives they receive. Consequently, while the user is likely to submit an honest ER, the CM may choose to disapprove it based on their preferences.

Proposition 5: In scenarios where $wf + p - v_h < 0$, $E_4(1,1)$ emerges as an ESS, indicating the adoption of honest strategies by both the user and the CM.

$$\{wf + p - v_h < 0. \tag{14}$$

Proof: When $wf + p - v_h < 0$, the sign of det$J$ at $E_4(1,1)$ is positive, and the sign of tr$J$ at $E_4(1,1)$ is negative. Thus, $E_4(1,1)$ serves as an ESS, confirming the statement in Proposition 5.

In this scenario, both the user and the CM stand to gain more by acting honestly. The user will submit honest ERs, and the CM will vote truthfully. This scenario is essential for maintaining the integrity of the redactable blockchain system.

Based on the analysis provided, it can be concluded that $E_1(0,0)$, $E_3(1,0)$, and $E_4(1,1)$ can serve as ESS under certain conditions. However, $E_2(0,1)$ is not considered an ESS. We are particularly interested in the local equilibrium points $E_1(0,0)$ and $E_4(1,1)$. These points represent scenarios where the user and the CM either engage in a collusion by choosing a malicious strategy, or they both adopt an honest strategy to collectively enhance the system. Therefore, we will focus on the following condition where $E_1(0,0)$ and $E_4(1,1)$ can be identified as the ESS:

$$\begin{cases} wf + wb - p > 0 \\ b - v_m + wf + wb - p < 0 \\ wf + p - v_h < 0. \end{cases} \tag{15}$$

TABLE V
PARAMETER VALUES

w	f	b	p	v_h	v_m	l	c
0.1	2	1	0.25	3	2	0.5	0.1

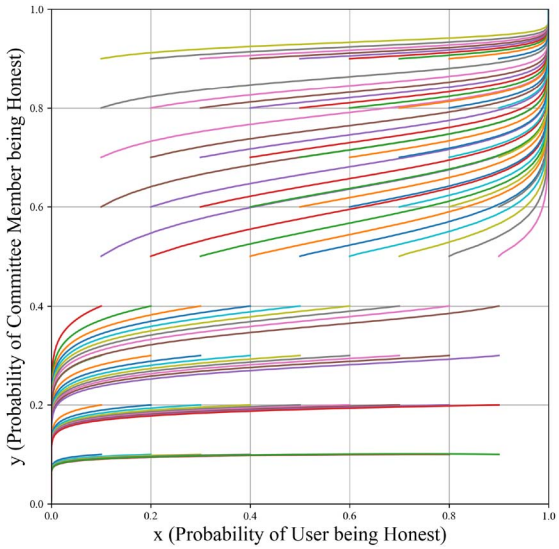


Fig. 3. Evolutionary trajectory of the system.

V. NUMERICAL SIMULATION

Numerical simulation plays a crucial role in game theory research, offering valuable insights and explanations to complement analytical findings. In this section, our emphasis shifts to the numerical simulation of the two previously mentioned cases using Python 3.10. Our goal is to investigate the impact of initial conditions on final outcomes and to explore how various parameters influence the evolutionary trajectory. The parameter values are shown in Table V.

A. Influence of Initial Conditions

As shown in Fig. 3, the evolutionary trajectory of converges toward $E_1(0,0)$ and $E_4(1,1)$. The ultimate result is contingent upon the initial condition of honest CMs. When the initial fraction of honest CMs exceeds $1/2$, the system tends to $E_4(1,1)$, indicating that both the user and CMs opt for the honest strategy. Conversely, if the initial fraction of honest CMs is less than $1/2$, the system gravitates toward $E_1(0,0)$, where both the user and CMs choose the malicious strategy.

In our further exploration of the dynamic evolutionary process involving the user and the CM, we have examined various initial conditions. Specifically, we have set the initial conditions as $(0.4, 0.4)$ and $(0.6, 0.6)$, and conducted simulations. The results are visually presented in Fig. 4. As shown in Fig. 4(a), when the initial probabilities for both the user and the CM to choose the honest strategy are set at $(0.4, 0.4)$, they ultimately both converge on adopting the malicious strategy. Conversely, as demonstrated in Fig. 4(b), when the initial probabilities are adjusted to $(0.6, 0.6)$, both the user and the CM eventually settle

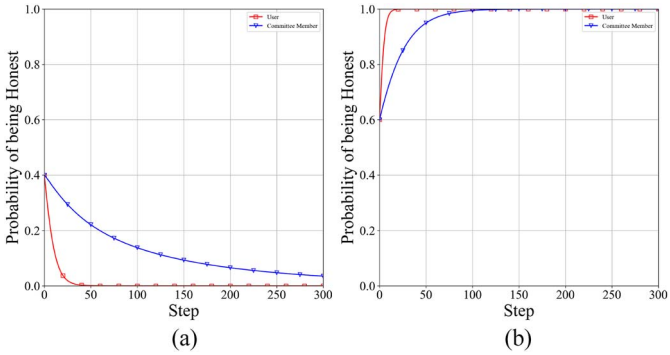


Fig. 4. Dynamic evolutionary process under different initial conditions. (a) $(0.4, 0.4)$. (b) (0.6,0.6).

on the honest strategy. These observed outcomes align with the expected evolutionary trajectory of the system.

In both of the described scenarios, the final outcomes are intricately linked to the initial conditions governing the CM's behavior. When a significant majority of the CMs adhere to honest strategies, the ultimate result tends to converge toward $E_4(1, 1)$. Under such circumstances, both the user and the CM are inclined to select honest strategies as well. Conversely, if a substantial portion of the CMs exhibit a propensity for adopting malicious strategies, the CM is more likely to follow suit by choosing a malicious strategy. Therefore, it becomes crucial to enhance the likelihood of the user selecting an honest strategy from the outset. The initial decision made by the CM plays a pivotal role in shaping the overall security of the system.

Moreover, it's important to note that the user's choice of strategy is not solely contingent on the initial decision of the CM but also relies on the specific relationships between the involved parameters. Consequently, in the upcoming section, we will delve into an exploration of how different parameters affect the system's dynamics and influence the user's strategic choices.

B. Influence of Parameters on the Evolutionary Outcome

This section investigates the impact of system parameters on the evolutionary outcomes of the system. Six crucial parameters, namely CM vote weight (w), ER fees (f), bribes offered by malicious users to corrupt CMs (b), CM preferences for ERs (p), the value derived from approving honest ERs (v_h), and the value obtained from approving malicious ERs (v_m), collectively influence the final outcome. Our analysis primarily focuses on understanding how changes in these parameters affect the behavior of both users and CMs.

Initially, we explore the influence of these parameters on the final evolutionary outcome under two different initial probability scenarios: $(0.4, 0.4)$ and $(0.6, 0.6)$. The results of this analysis are presented in Fig. 5, which provides insights into the influence of these critical parameters on the system's outcome.

Fig. 5(a) and 5(g) illustrates the impact of the CM's vote weight on the system's final outcome for initial probabilities of $(0.4, 0.4)$ and $(0.6, 0.6)$, with w values ranging from 0.02 to 0.4. When the initial probabilities are $(0.4, 0.4)$ and $w = 0.02$, both users and CMs tend to choose honest strategies. However, as the CM vote weight (w) increases, their decisions shift toward malicious strategies. In contrast, when the initial probabilities are $(0.6, 0.6)$, both users and CMs consistently opt for honest strategies, regardless of changes in w. This highlights the influence of the CM vote weight on the system's outcome. When the vote weight is relatively small, the system leans toward honesty, even in scenarios where the initial majority is malicious.

Referring to Fig. 5(b) and 5(h), a similar trend can be observed, specifically focusing on ER fees (f) ranging from 0.2 to 4. In the case where the initial probabilities are set to $(0.4, 0.4)$ and $f = 0.2$, both users and CMs adopt honest decision-making strategies. However, as the fee value (f) increases, their preferences shift toward adopting malicious strategies. Conversely, when the initial probabilities are $(0.6, 0.6)$, both users and CMs consistently adhere to honest strategies, regardless of variations in f. This emphasizes the influence of fees on the overall outcome of the system. Even in situations where the initial majority exhibits malicious tendencies, the system tends to lean toward honesty when the fee is relatively small.

The analysis of the bribe parameter (b) within the range of 0.2 to 4 exhibits distinct trends, as illustrated in Fig. 5(c) and 5(i). When the initial probabilities are set to $(0.4, 0.4)$, an increase in the bribe amount leads to a noticeable shift in user behavior, transitioning from a malicious strategy to an honest one. Simultaneously, CMs initially tend to adopt a malicious strategy but subsequently transition back to choosing an honest strategy as the bribe amount becomes excessively large. Conversely, when the initial probabilities are $(0.6, 0.6)$, changes in the bribe parameter do not have a significant impact on the evolutionary outcome. These findings emphasize the influence of bribes on the overall system outcome. If the bribe required is too substantial, users will opt not to offer a bribe.

The analysis of the influence of the preference parameter (p) within the range of 0.05 to 1 reveals distinct trends, as illustrated in Fig. 5(d) and 5(j). When the initial probabilities are set at $(0.4, 0.4)$, an increase in the preference amount leads to a noteworthy shift in the choices made by both users and CMs, transitioning from a malicious strategy to an honest one. However, when the initial probabilities are $(0.6, 0.6)$, the preference parameter does not demonstrate any discernible impact on the evolutionary process. Consequently, it can be inferred that a higher preference value increases the likelihood of both CMs and users adopting an honest strategy.

Fig. 5(e) and 5(k) depicts the influence of the v_h parameter on the evolutionary result, with v_h values ranging from 0.6 to 12. From these figures, it is evident that the parameter v_h does not exert a significant influence on the evolutionary outcome under both conditions.

Examining Fig. 5(f) and 5(l), we delve into the influence of the values derived from the approval of malicious ERs, with v_m values ranging from 0.4 to 8. It becomes apparent that when the initial probabilities are set at $(0.4, 0.4)$, an increase in v_m tends to induce users and CMs to shift from honest strategies to malicious ones. In contrast, when the initial probabilities

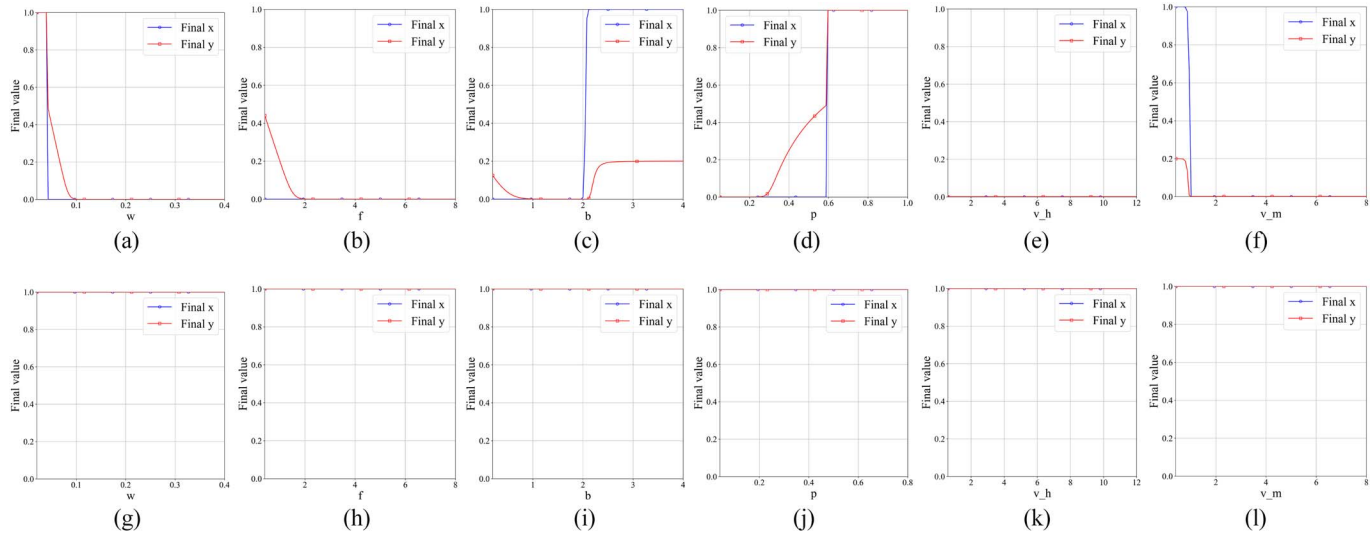


Fig. 5. Influence of parameters. (a) w in $(0.4, 0.4)$. (b) f in $(0.4, 0.4)$. (c) b in $(0.4, 0.4)$. (d) p in $((0.4, 0.4))$. (e) v_h in $(0.4, 0.4)$. (f) v_m in $(0.4, 0.4)$. (g) w in $(0.6, 0.6)$. (h) f in $(0.6, 0.6)$. (i) b in $(0.6, 0.6)$. (j) p in $(0.6, 0.6)$. (k) v_h in $(0.6, 0.6)$. (l) v_m in $(0.6, 0.6)$.

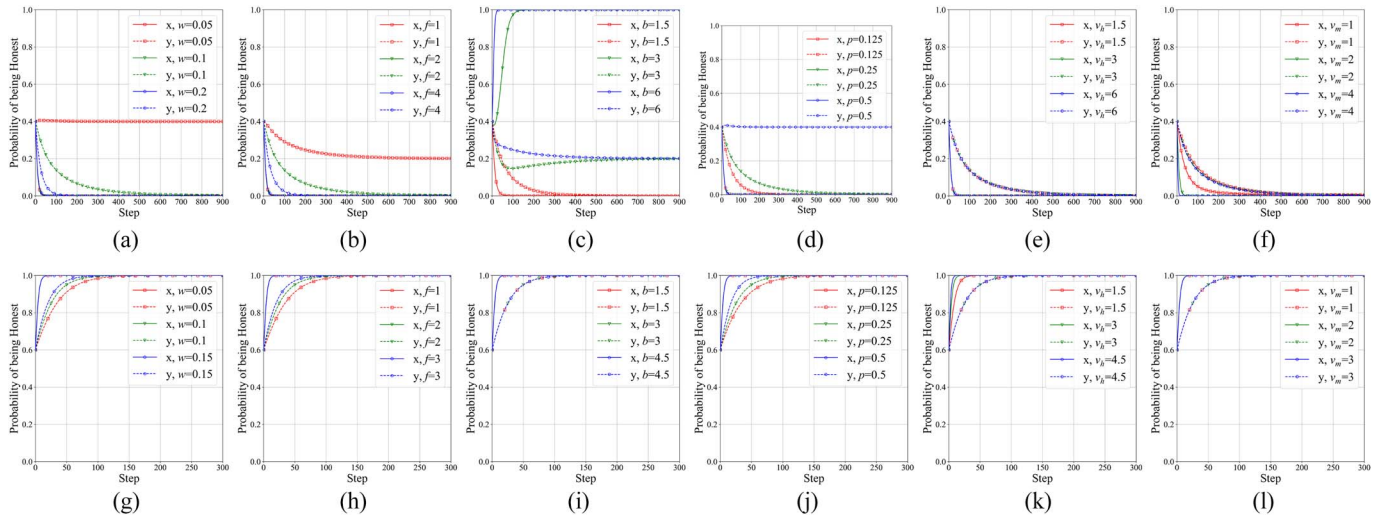


Fig. 6. Sensitivity of parameters. (a) w in $(0.4, 0.4)$. (b) f in $(0.4, 0.4)$. (c) b in $(0.4, 0.4)$. (d) p in $((0.4, 0.4))$. (e) v_h in $(0.4, 0.4)$. (f) v_m in $(0.4, 0.4)$. (g) w in $(0.6, 0.6)$. (h) f in $(0.6, 0.6)$. (i) b in $(0.6, 0.6)$. (j) p in $(0.6, 0.6)$. (k) v_h in $(0.6, 0.6)$. (l) v_m in $(0.6, 0.6)$.

are $(0.6, 0.6)$, variations in v_m do not impact the evolutionary outcome. Hence, it is evident that a higher value derived from malicious ERs prompts users to choose a malicious strategy and potentially offer a bribe.

C. Sensitivity Analysis on the Parameters

This section examines the impact of system parameters on the evolutionary outcomes of the system. The six parameters mentioned earlier will be examined in this context. Specifically, we will investigate their evolutionary process considering two initial probability scenarios: $(0.4, 0.4)$ and $(0.6, 0.6)$. The findings of this analysis are depicted in Fig. 6, which provides valuable insights into the system's sensitivity to these crucial parameters.

Fig. 6(a) and 6(g) demonstrates the system's sensitivity to CM's vote weight for initial probabilities of $(0.4, 0.4)$ and $(0.6, 0.6)$, with w values of 0.05, 0.1, and 0.15. When initial probabilities are $(0.4, 0.4)$, both user and CM tend toward malicious strategies. As w increases, the CM's adoption of malicious strategies accelerates, but the user's evolution remains unaffected. In contrast, with initial probabilities at $(0.6, 0.6)$, both opt for honest strategies, and higher w speeds up the CM's selection of honest strategies. The user's evolution is impervious to changes in w. These findings emphasize the influence of a CM's vote weight on their decision-making speed.

Moving to Fig. 6(b) and 6(h), we observe a similar sensitivity analysis, this time focused on fees for all ERs (f). These figures yield results akin to the vote weight sensitivity. The user's choice remains unaltered with varying f. As f increases, the

CM's convergence to a final strategy accelerates. In other words, higher fees expedite the CM's decision-making process.

The sensitivity analysis of the bribe parameter (b) leads to distinct observations, as depicted in Fig. 6(c) and 6(i). When the initial probabilities are set to $(0.4, 0.4)$, an increase in the bribe amount results in a notable shift in the user's choice, transitioning from a malicious strategy to an honest one. At the same time, the CM's choice shifts from malicious to probabilistic honest. However, when the initial probabilities are $(0.6, 0.6)$, the bribe parameter does not exhibit any discernible impact on the evolutionary process. In scenarios where the majority of the CMs are honest, bribes do not affect the decision-making process. However, when a majority of CMs are dishonest, higher bribe amounts prompt both users and CMs to switch to honesty.

The sensitivity analysis of the preference parameter (p) is illustrated in Fig. 6(d) and 6(j). When the initial probabilities are set at $(0.4, 0.4)$, an increase in the preference parameter results in a noticeable shift in the CM's choice, transitioning from a malicious strategy to a probabilistic honest strategy. Conversely, when the initial probabilities are $(0.6, 0.6)$, the preference parameter accelerates the CM's adoption of an honest strategy and has no effect on the user's choice. Hence, a higher preference parameter can make the CM choose honest strategy in a malicious initial condition and prompt the CM to choose the honest strategy more quickly in an honest initial condition.

In Fig. 6(e) and 6(k), we illustrate the sensitivity to the values derived from the approval of honest ERs. It is evident that when the initial probabilities are set at $(0.4, 0.4)$, the parameter v_h exhibits no significant influence on the evolutionary trajectory. Conversely, when the initial probabilities are $(0.6, 0.6)$, variations in the value of v_h do impact the user's choices, although they have no discernible effect on the CM's choices. Notably, an increase in the value of v_h tends to prompt users to adopt the honest strategy more rapidly. When the majority is malicious, values from approving honest edits do not matter. Yet, in an honest-majority situation, higher values speed up user decisions.

Examining Fig. 6(f) and 6(l), we can explore the sensitivity to values derived from the approval of malicious ERs. It becomes evident that when the initial probabilities are set to (0.4, 0.4), an increase in v_m leads users to adopt a malicious strategy more quickly. However, this parameter variation does not influence the decisions of the CMs. On the other hand, when the initial probabilities are set to (0.6, 0.6), changes in v_m do not impact the evolutionary path. The approval of malicious edits has no effect on the choice of the CMs. In scenarios where the majority of CMs are honest, the values derived from approving malicious edits have no impact. However, in situations where a majority of CMs are dishonest, higher values of v_m accelerate user decisions.

D. Discussion

Our simulation-based analysis provides valuable insights into the behaviors of users and CMs within the system. A key observation from our analysis is that the initial choice probabilities assigned to CMs exert a significant influence on the system's evolutionary trajectories. More specifically, the system exhibits a stronger proclivity for honesty when CMs have a higher initial probability of choosing honest strategies. This finding underlines the importance of setting initial conditions optimally to promote integrity and discourage malicious behaviors within the system.

Our findings also illustrate the influence that parameters such as vote weight, fees, bribes, preference, and the values derived from honest and malicious ERs, have on the evolutionary outcomes. For instance, smaller vote weights and lower fees tend to favor honest strategies. Additionally, high bribes discourage malicious behavior, while high preference values increase the likelihood of adopting honest strategies. Interestingly, an increase in the value derived from approving malicious ERs may incentivize malicious behavior, while a high value for approving honest requests pushes the system toward honesty. These findings are instrumental for designing effective incentive mechanisms within redactable blockchain systems, emphasizing the need to consider these parameters carefully.

VI. Future Research Directions

A. Scalability in Redactable Blockchain

Scalability represents a critical concern within blockchain technology, pertaining to the capacity of a blockchain network to accommodate a growing number of users. This issue retains its significance in the context of redactable blockchains, where the scalability challenges are compounded by the inclusion of redaction capabilities. Unlike traditional blockchains, redactable blockchains incorporate mechanisms that allow for the alteration of data postconfirmation, which introduces additional layers of complexity. The impact of these redaction mechanisms on the scalability of blockchain networks presents a compelling area for investigation. Moreover, the role of incentive structures in redactable blockchains and their influence on scalability merits thorough exploration. Addressing these concerns is pivotal for advancing the deployment of redactable blockchains, making scalability a paramount research focus.

B. Interoperability

Interoperability is a crucial aspect in blockchain technology. Further research should be conducted on the interoperability between redactable blockchain and traditional blockchain. To facilitate the transition of data from a traditional blockchain to a redactable blockchain and ensure data compliance within the blockchain ecosystem, we introduce a preliminary approach for blockchain interoperability. This approach involves establishing a bridge to enable the transfer of data from the traditional blockchain to the redactable blockchain. The bridge facilitates the redactable blockchain in accessing data from the traditional blockchain, allowing for specific information to be modified to adhere to legal requirements. For transferring data from a redactable blockchain back to a traditional blockchain, a more

advanced framework is essential. The process of realizing this mechanism necessitates further investigation.

C. Long-Term Data Management in Redactable Blockchain

The inherent limitation in storage capacity of blockchain networks necessitates innovative solutions to enhance their utility. Various strategies, such as leveraging the InterPlanetary File System (IPFS) or cloud-based services, have been proposed to extend the storage capabilities of blockchain. However, the integration of redactable blockchains, which allow for post-confirmation data modification, with these expanded storage solutions introduces significant challenges. Furthermore, the practical deployment of redactable blockchains, especially in real-world applications, demands the establishment of robust data lifecycle management, access control measures, and retention policies. These requirements become even more complex when integrating IPFS and cloud storage solutions, as they must ensure data integrity and security while enabling redaction capabilities. Addressing these challenges is crucial for the successful implementation and adoption of redactable blockchains in practical applications.

D. Real-World Deployment of Redactable Blockchain

The concept of a redactable blockchain, while currently confined to academic research, has yet to be operationalized in real-world settings. The deployment of redactable blockchain technology in practical environments is essential, as it would provide access to authentic data for the system's operation, enable the evaluation of its performance in real-world scenarios, and facilitate empirical case studies. Such deployment is a critical step in advancing the redactable blockchain from theoretical exploration to tangible application, thereby promoting its adoption and informing further development.

VII. Conclusion

This article introduces a novel contribution to the redactable blockchain field by proposing an incentive mechanism grounded in evolutionary game theory. By leveraging game theory principles, we have adeptly formulated and integrated an incentive system that fosters honest behavior while discouraging malicious actions within the network. Furthermore, we conduct simulations rooted in game-theoretic analysis to evaluate the effectiveness of our incentive mechanism. The outcomes of these simulations provide strong evidence supporting the efficiency and practicality of our incentive system, highlighting its potential impact on future research and development within the redactable blockchain domain. Additionally, our study enhances the understanding of how individual behaviors within the blockchain influence collective outcomes, which is crucial for the social dynamics of blockchain communities. The evolutionary game theory approach adopted in our research sheds light on the interaction dynamics among participants in the redactable blockchain and offers insights into the application of similar mechanisms for managing social interactions and promoting positive conduct in blockchain-based communities.

References

[1] D. Siegel, "The DAO Attack: Understanding what happened–CoinDesk," Tech. rep. Coindesk, 2016. Accessed Jan. 03, 2022. [Online]. https://www.coindesk.com/learn/2016/06/25/understanding-the-dao-attack/

[2] M. Schellekens, "Does regulation of illegal content need reconsideration in light of blockchains?" *Int. J. Law Inf. Technol.*, vol. 27, no. 3, pp. 292–305, 2019.

[3] J. Leng et al., "Makerchain: A blockchain with chemical signature for self-organizing process in social manufacturing," *J. Cleaner Prod.*, vol. 234, pp. 767–778, 2019.

[4] J. Leng, W. Sha, Z. Lin, J. Jing, Q. Liu, and X. Chen, "Blockchained smart contract pyramid-driven multi-agent autonomous process control for resilient individualised manufacturing towards industry 5.0," *Int. J. Prod. Res.*, vol. 61, no. 13, pp. 4302–4321, 2023.

[5] J. Leng et al., "ManuChain II: Blockchained smart contract system as the digital twin of decentralized autonomous manufacturing toward resilience in industry 5.0," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 53, no. 8, pp. 4715–4728, Aug. 2023.

[6] G. Ateniese, B. Magri, D. Venturi, and E. Andrade, "Redactable blockchain—or—rewriting history in bitcoin and friends," in *Proc. IEEE Eur. Symp. Secur. Privacy (EuroS&P)*, Apr. 2017, pp. 111–126.

[7] D. Deuber, B. Magri, and S. A. K. Thyagarajan, "Redactable Blockchain in the Permissionless Setting," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2019, pp. 124–138.

[8] X.-Y. Li, J. Xu, L.-Y. Yin, Y. Lu, Q. Tang, and Z.-F. Zhang, "Escaping from consensus: Instantly redactable blockchain protocols in permissionless setting," *IEEE Trans. Dependable Secure Comput.*, vol. 20, no. 5, pp. 3699–3715, Sep./Oct. 2022.

[9] Y. Wang, Y. Wu, and J. Lai, "Contract-based incentive mechanism for redactable proof-of-stake blockchains," *Secur. Commun. Netw.*, vol. 2023, May 2023, Art. no. e6403686.

[10] X. Liu, W. Wang, D. Niyato, N. Zhao, and P. Wang, "Evolutionary game for mining pool selection in blockchain networks," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 760–763, Oct. 2018.

[11] X. Liu, Z. Huang, Q. Wang, and B. Wan, "An evolutionary game theory-based method to mitigate block withholding attack in blockchain system," *Electronics*, vol. 12, no. 13, 2023, Art. no. 2808.

[12] J. Zhang and M. Wu, "Cooperation mechanism in blockchain by evolutionary game theory," *Complexity*, vol. 2021, Nov. 2021, Art. no. e1258730.

[13] S. Motepalli and H.-A. Jacobsen, "Reward mechanism for blockchains using evolutionary game theory," in *Proc. 3rd Conf. Blockchain Res. Appl. Innovative Netw. Services (BRAINS)*, Sep. 2021, pp. 217–224.

[14] J. Zhang et al., "Serving at the edge: a redactable blockchain with fixed storage," in *Proc. Int. Conf. Web Inf. Syst. Appl.*, Cham, Switzerland: Springer, 2020, pp. 654–667.

[15] M. Jia et al., "Redactable blockchain from decentralized Chameleon hash functions," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 2771–2783, 2022.

[16] A. Marsalek and T. Zefferer, "A correctable public blockchain," in *Proc. 18th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./13th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2019, pp. 554–561.

[17] R. Han, Z. Yan, X. Liang, and L. T. Yang, "How can incentive mechanisms and blockchain benefit with each other? a survey," *ACM Comput. Surveys*, vol. 55, no. 7, pp. 1–38, 2022.

[18] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, 1973.

[19] J. Hofbauer and W. H. Sandholm, "Stable games and their dynamics," *J. Econ. Theory*, vol. 144, no. 4, pp. 1665–1693, 2009.

[20] R. Cressman, *The Stability Concept of Evolutionary Game Theory: A Dynamic Approach*, vol. 94. Berlin, Germany: Springer Science & Business Media, 2013.

[21] Z. Ni, W. Wang, D. I. Kim, P. Wang, and D. Niyato, "Evolutionary game for consensus provision in permissionless blockchain networks with shards," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[22] S. Iyer, J. Reyes, and T. Killingback, "An application of evolutionary game theory to social dilemmas: the traveler's dilemma and the minimum effort coordination game," *PLoS One*, vol. 9, no. 4, 2014, Art. no. e93988.

[23] D. Friedman, "Evolutionary games in economics," *Econometrica: J. Econometric Society*, vol. 59, no. 3, pp. 637–666, May 1991.

[24] M. I. Weinstein, "Lyapunov stability of ground states of nonlinear dispersive evolution equations," *Commun. Pure Appl. Math.*, vol. 39, no. 1, pp. 51–67, 1986.

Jiaxiang Sun received the B.Sc. degree in environmental science and computer science and technology from Peking University, Beijing, China, in 2022. He is currently working toward the M.Phil. degree in computer and information engineering from the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China, under the supervision of Prof. Wei Cai.

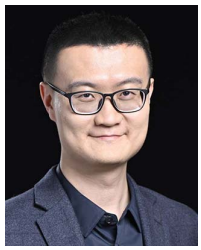
He is working with Human-Crypto Society Laboratory, Shenzhen, China. His research interests include blockchain and game theory.

Rong Zhao received the B.Eng. degree in communication engineering from The Southern University of Science and Technology, Shenzhen, China, in 2018. He is currently working toward the Ph.D. degree in computer and information engineering with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China.

He is working as a Research Assistant with Human-Crypto Society Laboratory, Shenzhen, China. His research interests include blockchain, game theory, and digital twin.

Haoran Yin received the B.Eng. degree in computer science and technology and the B.Econ. degree in financial engineering from Sichuan University, Chengdu, China, in 2022. He is currently working toward the Ph.D. degree in computer and information engineering with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China, working under the supervision of Prof. Wei Cai.

His research interests include blockchain, game, and Web3.



Wei Cai (Senior Member, IEEE) received the B.Eng. degree in software engineering from Xiamen University, Xiamen, China, in 2008, the M.S. degree in electrical engineering and computer science from Seoul National University, Seoul, Korea, in 2011, and the Ph.D. degree in electrical and computer engineering from The University of British Columbia (UBC), Vancouver, BC, Canada, in 2016.

From 2016 to 2018, he was a Postdoctoral Research Fellow with UBC. He is currently an Assistant Professor in computer engineering with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China. He is serving as the Director of the Human-Crypto Society Laboratory, Shenzhen, China, as well as the Director of the CUHK(SZ)-White Matrix Joint Metaverse Laboratory, Shenzhen, China. He has co-authored more than 100 journal and conference papers in the areas of distributed/decentralized systems and interactive multimedia.

Dr. Cai is an Associate Editor for IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, IEEE TRANSACTIONS ON CLOUD COMPUTING, and *ACM Transactions on Multimedia Computing, Communications and Applications*, and a Program Co-Chair for ACM NOSSDAV'23. He was a recipient of six best paper awards. He is a member of the ACM.