mathematics

MDPI

Article

Achieving Anonymous and Covert Reporting on Public Blockchain Networks

Liehuang Zhu [1], Jiaqi Zhang [2], Can Zhang [1], Feng Gao [1], Zhuo Chen [1] and Zhen Li [2,3,*]

[1] School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China
[2] School of Computer Science & Technology, Beijing Institute of Technology, Beijing 100081, China
[3] Southeast Institute of Information Technology, Beijing Institute of Technology, Putian 351100, China
* Correspondence: zhen.li@bit.edu.cn

Abstract: Reporting helps to combat illegal activities and deters lawbreakers and potential lawbreakers. From ancient times to the present, public authorities have usually rewarded effective reporting information to build harmonious societies. In this process, protecting the privacy of the whistleblower is a very important issue. Existing blockchain-based anonymous reporting solutions help solve the problem of insufficient anonymity in traditional reporting solutions, but they do not address the issue of hiding the reporting behavior. The disclosure of reporting behavior may alert offenders in advance and negatively impact case handling. This paper proposes an anonymous and covert reporting scheme and rewarding mechanism based on blockchain, which realizes the covertness of the reporting behavior while protecting the privacy of the whistleblower. The proposed scheme uses ring signature and derived address technology to ensure anonymity and achieves covertness by embedding information in the ring signature based on the idea of covert communication. Theoretical analysis proves that the proposed scheme has covertness, anonymity, and unforgeability properties. Experiments show that the proposed scheme takes only 0.08 s to upload data and 0.07 s to verify while achieving covertness.

Keywords: blockchain network; covert communication; anonymous reporting; ring signature; information embedding; smart contract

MSC: 68P25; 94A60

Citation: Zhu, L.; Zhang, J.; Zhang, C.; Gao, F.; Chen, Z.; Li, Z. Achieving Anonymous and Covert Reporting on Public Blockchain Networks. *Mathematics* **2023**, *11*, 1621. https://doi.org/10.3390/math11071621

Academic Editor: Antanas Cenys

Received: 14 February 2023
Revised: 22 March 2023
Accepted: 22 March 2023
Published: 27 March 2023

1. Introduction

Reporting is a basic right of citizens. In order to crack down on people suspected of violating discipline, law, or crime, citizens will report to relevant agencies or organizations. To obtain more information, strengthen the investigation, and reduce illegal acts, public authorities will also establish a reward and reporting system. For example, the vast majority of tax violation cases investigated and handled by tax authorities each year are handled through reporting, and reporting is therefore regarded as an important source of information for tax authorities to investigate and deal with tax violation cases. The reporting system can effectively solve the problem of information asymmetry, and it can also increase the illegal cost of the offenders and deter the offenders and potential offenders.

Because of the importance of reporting, the offender will block the reporting process and even retaliate against the whistleblower. These retaliatory actions will greatly affect the whistleblower's willingness to report, which is not conducive to cracking down on illegal and criminal acts. How to protect the privacy of the whistleblower and hide the identity of the whistleblower is an important issue.

In addition, for some offenders, even if they do not know the specific whistleblower but only know that there is a reporting behavior, they can also carry out destructive actions such as retaliation. For example, a fugitive who is wanted for a reward may learn that his whereabouts have been exposed based on the reporting behavior, so as to predict the

actions of the police, or authorities with certain powers and information resources can make preparations in advance according to the reporting behavior and manage their relationship network to conceal his illegal behavior. During the process of reporting, if not only the privacy of the whistleblower can be protected, but also the reporting information or even the reporting behavior itself can be hidden, and, on this basis, the whistleblower can be reasonably rewarded, then such a system can effectively improve the grasp of information, the willingness of citizens to report, and the crackdown on crimes.

Blockchain networks use modern cryptographic algorithms to provide decentralized and unforgeable information trading platforms. In a blockchain network, reporting can then be considered as a transaction sent from the address of the whistleblower to the address of an authority. If it can be ensured that the address of the whistleblower is indistinguishable from the addresses of other users who are not involved and that the transaction to report is indistinguishable from other ordinary transactions, it can be said that the proposed scheme has achieved anonymity and a covertness of reporting. Using blockchain, researchers have proposed some solutions, such as using ring signature technology to achieve anonymity in the reporting process [1] or using blockchain mechanisms and smart contracts to ensure security [2]. However, existing schemes mainly focus on the anonymity of the reporting schemes, [1] they require an off-chain channel, and the reporting transactions in [2,3] have distinct characteristics (this paper will discuss them in detail in Section 2.2). To better protect privacy, it is equally important to ensure the covertness of reporting behavior. Researchers have proposed a number of blockchain-based covert communication schemes [4–6]. By introducing the idea of covert communication into anonymous reporting systems, the privacy of the whistleblower can be largely enhanced.

The contributions of this paper can be summarized as follows:

- This paper proposes the architecture of a blockchain-based anonymous reporting system that supports the hiding of reporting behaviors. This paper considers the scenario of using smart contracts to collect data in smart city management. Through smart contracts, users can upload normal data and whistleblowers can report secretly, which ensures that reporting behaviors are hidden throughout the whole process.
- Based on the above architecture, this paper designs an anonymous and covert reporting scheme based on blockchain. More specifically, the ring signature is used to realize the anonymity of the whistleblower during the reporting process, the information hiding mechanism is used to hide reporting information, and the derived address is used to realize the anonymity of the whistleblower during the rewarding process and the unlinkability between the blockchain addresses owned by the whistleblower.
- Through theoretical analysis, it is proved that the proposed scheme has covertness, anonymity, and unforgeability properties that satisfy the security goals. This paper also uses several experiments to prove that the scheme has a covertness and acceptable efficiency that meet the actual scenario.

The structure of this paper is as follows: In Section 2, this paper will introduce the primary knowledge and related work that support the solution of this paper. In Section 3, this paper will propose the system model and security model of the scheme. In Section 4, the algorithm of this scheme will be introduced in detail. Section 5 provides the security proof and relevant experimental analysis. Finally, Section 6 gives the conclusion of this paper.

2. Related Works

This section first introduces the relevant knowledge of the blockchain and then respectively introduces the anonymous reporting system and the covert communication mechanism based on the blockchain. Based on this, the anonymous and covert reporting scheme in this paper is introduced.

2.1. Blockchain

In 2008, Satoshi Nakamoto first proposed the concept of "blockchain" [7]. Blockchain is a decentralized, distributed public digital ledger, composed of records called blocks, used to record transactions of multiple computers. Each block contains the cryptographic hash of the previous block, the corresponding timestamp, and transaction information. Blockchain has the characteristics of decentralization, independence, openness, and security. Blockchain technology does not rely on third-party institutions, and there is no central node. Based on consensus specifications and protocols, all nodes can automatically and safely verify and send transactions in the system, without the participation of any third parties. The data on the blockchain are open to everyone, and anyone can query blockchain data or develop related applications through the public interface. The structure of the blockchain network guarantees that, unless an attacker can control more than 50% of the hash power, the attacker cannot manipulate or modify network data. In 2013, Vitalik Buterin proposed the concept of Ethereum [8]. Ethereum uses a Turing complete scripting language to create applications, namely smart contracts, which can provide diversified services.

The identity information of nodes in the blockchain network does not need to be disclosed or verified, which theoretically protects the anonymity of network members. However, in fact, the blockchain led by Bitcoin can only provide limited anonymity, and attackers can reduce the size of the anonymity set through methods such as address clustering or fund flow correlation [9]. To improve security and privacy, Monero was proposed in 2014 [10]. Monero uses the RING-CT protocol, which mainly includes one-time privacy address technology and ring signature technology [11]. Monero uses encryption technology to shield the sending address, receiving address, and transaction amount so that the addresses of both parties in the transaction cannot be associated with the user's real identity. In addition, one-time privacy address technology is used to reduce the correlation between different transactions, and ring signature technology is used to protect the anonymity of transaction senders.

2.2. Blockchain-Based Anonymous Reporting System

One of the most important problems of traditional anonymous reporting models is that the anonymity of the whistleblowers cannot be guaranteed. Once personal information is leaked, it may bring danger to personal and family property safety and even life safety, so the relevant persons who have evidence or clues may give up. This problem can be regarded as a management problem, but it can also be solved through technical means to a certain extent.

The existing anonymous reporting system often claims that it can protect the identity of the whistleblower, but the centralized structure of the existing reporting system makes people generally worry about its security and reliability. Blockchain has the characteristics of decentralization. Researchers also use this feature in many systems that need to ensure anonymity. For example, in 2018, Lu et al. [12] proposed an anonymous and private decentralized crowdsourcing system called ZebraLancer. Lu et al. [13] proposed an anonymous reputation system based on blockchain. In 2019, Yao et al. [14] proposed a blockchain-assisted lightweight anonymous authentication mechanism for distributed VFS. Similarly, blockchain technology can also be used to construct an anonymous reporting system. Wang et al. [2] proposed a blockchain-based anonymous reporting and anonymous rewarding scheme in 2018, which can ensure the anonymity of the whistleblower during the reporting and rewarding process. However, in that system, the whistleblower's reporting process is carried out through off-chain channels, and it is easy for attackers to eavesdrop on the whistleblower's information through off-chain channels. In 2019, Zou et al. [1] proposed an anonymous reporting system called ReportCoin. ReportCoin guarantees the reliability of the reporting information and the privacy of users during the entire reporting process. In that system, the reporting behavior is regarded as a transaction, which means that the reporting behavior is public, which is contrary to the security goals proposed in this paper. In addition, in 2022, Zhang et al. [3] introduced a trust currency called TCoin

used in VANETs for anonymous reporting. That scheme focuses more on the reliability of the reporting information and does not consider the covertness of the reported behavior as one of the security goals.

The above reporting systems can guarantee the personal information of the whistleblower and can reward the whistleblowers who provide effective information, thereby enhancing the value of the anonymous reporting system. However, the above systems are more focused on ensuring the anonymity of the whistleblower and are not trying to hide the reporting behaviors. In actual scenarios, this assumption may not be strong enough to complete the design goals of the reporting system. Therefore, to solve the above problems, this paper proposes an anonymous reporting system that can hide the reporting behavior and realizes the security of the reporting behavior in the whole process.

2.3. Covert Communication Mechanism

Covert communication technology enables information to be transmitted imperceptibly through open channels. Covert communication technology was proposed in the early 1980s [15,16] and has a wide range of application scenarios. Many public channels can be used for covert communication. In general, covert communication channels can be divided into storage channels and timing channels. Roughly speaking, most covert communication schemes use the storage covert channel, such as the covert channels constructed by public-key cryptography, for example, ECDSA-based covert communication channels [17], EdDSA-based covert communication channels [18], covert communication channels based on ring signatures [19], and so on. With the development of the Internet, there have been many kinds of research on the use of network streams for covert communication. For example, in [20,21], the authors use unused fields in the network protocol header to achieve covert communication, and, in [22], use packet attributes such as packet rate, packet length, etc. These methods have their advantages and disadvantages. Generally speaking, covert communication schemes based on storage channels are more stable and have larger channel capacities, while schemes based on timing channels, such as [22], achieve stronger covertness.

Considering the characteristics of blockchain, such as unforgeability and decentralization, many covert communication schemes based on blockchain have also been proposed. Blockchain-based covert communication schemes have some similarities with network-flow-based covert communication schemes, while the schemes that take advantage of the blockchain can achieve higher stability and covertness. In 2018, Partala [23] proposed a blockchain-based covert communication scheme called BLOCCE. In 2019, Li et al. [24] proposed a covert communication scheme based on a time covert channel, which has a higher robustness than traditional schemes. In 2020, Cao et al. [5] proposed a hidden data embedding scheme based on the hash chain and further proposed a hidden data embedding scheme based on the elliptic curve Diffie–Hellman chain to enhance the security of the former. Gao et al. [25] designed a blockchain covert data transmission scheme using kleptography technology to achieve a high covertness and high-performance data transmission under open network conditions. In 2021, Qin et al. [6] proposed a covert communication model based on the parity of the blockchain transaction address. By modulating the parity of the transaction address, the sender can secretly transmit the information to the receiver by adding the address. In addition, in 2022, Zhang et al. [26] designed an index matrix of address interaction for group covert communication, using the address interaction relationship and transaction amount to hide secret messages alternately. By using generative adversarial networks and IPFS, She et al. [27] proposed a blockchain-based covert communication model for hiding sensitive documents and sender identity.

Although the covert communication schemes above achieve remarkable covertness, there still exist some security risks in anonymity. Specifically, these blockchain-based covert communication schemes often face the problems of address association and transaction association. On the blockchain, an attacker can use the data on the chain to infer different addresses owned by the same user, such as using timestamps [28] or clustering

methods [29]. If one of the addresses that belong to the sender is found to be used for covert communication, the other address may also be used for covert communication. Since sending transactions requires funds, which can only be transferred from another address for the newly generated address, the flow of funds can also be regarded as a kind of address association. Therefore, for a reporting system, it is not enough to only consider the covertness of the reporting process. The attacker can use on-chain information to reduce the anonymity set, which may lead to the disclosure of the sender's identity.

Covert communication technology can be used to hide the behavior of information transmission in public channels; that is, in the view of the third party, the behavior of participating in covert communication is indistinguishable from normal communication behavior. Therefore, to solve the problem of the disclosure of reporting information in the existing anonymous reporting system, this paper introduces covert communication technology into the anonymous reporting system and proposes the blockchain-based anonymous and covert reporting scheme. The proposed scheme realizes the anonymity of the whistleblower and the covertness of the reporting behavior during the reporting process.

3. Problem Formalization

This section presents the system model and security model of this scheme. The system model describes the entities in the system and their actual work, and the security model describes the security goals expected to be achieved by this scheme.

3.1. System Model

There are three entities in the reporting scheme: the user, authority, and blockchain network. Their roles in the scheme are as follows:

- User: The user sends data to the smart contract. Generally, users will send data to the smart contract following official regulations and the requirements of the contract. When a user wants to report some illegal or criminal behaviors, the user will embed the reporting information in the sent data. At this time, the user expects to be able to report a known criminal behavior without revealing his identity and be able to receive corresponding rewards.
- Authority: In order to collect user data and fight against illegal and criminal acts, the authority deploys the smart contract in the blockchain network to collect user data, and the reporting information can also be transmitted covertly by using the contract. The authority checks whether it contains covert information while checking data. In addition, the authority will reward valuable information (including normal information or reporting information).
- Blockchain network: The blockchain network is the platform used by this scheme. Users upload data and obtain rewards by calling contracts and receiving transactions in the blockchain network.

The system model is shown in Figure 1. The whole system consists of three processes: *Initialization*, *Data Upload*, and *Reward*. In the *Initialization* phase, the authority establishes a blockchain network and deploys smart contracts on the blockchain network. On the surface, smart contracts are used for user data collection or other public services. They will also be used to make anonymous reports. Users connected to the blockchain network can use the smart contract to send normal data as well as reporting information. Each user in the blockchain network can obtain the public parameters. In the *Data Upload* phase, users send normal data required by the authority to the smart contract, and the whistleblower embeds the report materials into the normal data and uploads them to the smart contract. After the authority checks whether the data transmission is integrated, it will also check whether there is covert information embedded in it. In the *Reward* phase, the authority rewards whistleblowers who send valuable information or other ordinary users who send valuable data. This paper adopts the technology of ring signature and derived address to realize the anonymity of the whistleblower and embeds covert information using ring signature parameters.

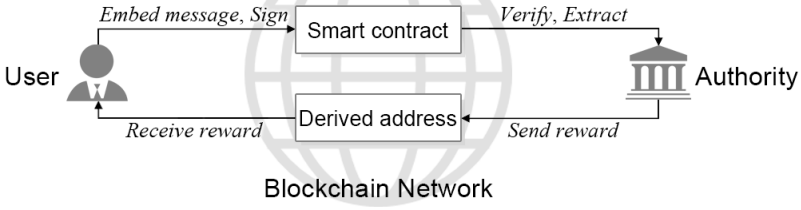


Figure 1. Structure diagram of the anonymous and covert reporting scheme.

Formally, the blockchain-based anonymous and covert reporting scheme includes the following six algorithms:

- **Setup**$(1^\lambda) \rightarrow (W, w, \mathcal{K}, \{k_1, k_2, \ldots, k_n\}, pp)$: Initialization. Given the security parameter 1^λ, the algorithm outputs the parameters used in the scheme, including the public key W and private key w of the authority, the public parameter pp, and the users' public-key group $\mathcal{K} = \{K_1, K_2, \ldots, K_n\}$ and private key $= \{k_1, k_2, \ldots, k_n\}$, where n represents the number of users.
- **Upload**$(data, Msg, \mathcal{K}, k_j, pp) \rightarrow (\sigma, \tilde{K}, A, B)$: Data upload. Given the users' public-key group, the user's private key k_j, the public parameter pp, a piece of information Msg, and the $data$ to be transmitted, the algorithm outputs a set of data for sending to the smart contract, including a ring signature σ and two public keys A, B. When the user carries out the normal data upload process, the information Msg is empty, but if the user wants to make a report, the Msg is the user's report material.
- **Verify**$(\sigma, \tilde{K}, \mathcal{K}, data, pp) \rightarrow \{0, 1\}$: Signature verification. Given the data in the smart contract, the authority performs signature verification. If the signature verification passes, the algorithm outputs 1, otherwise it outputs 0.
- **Extract**$(\sigma, w, pp) \rightarrow Msg'$: Decryption. The algorithm attempts to extract the ring signature σ with the authority private key w to obtain the report material Msg'.
- **Reward**$(A, B, pp) \rightarrow P$: Send rewards. Given two public keys A, B and the public parameter pp of the blockchain, if the data or the reporting information are valuable, the algorithm calculates a derived address P based on this data and sends a transaction to this address as a reward for reporting.
- **Gain**$(a, b, pp) \rightarrow (x, P')$: Receive rewards. The user queries the blocks on the blockchain, uses his private random numbers a, b to calculate the private key X and the corresponding address p', and tries to receive rewards. Thereafter, the user can use the new derived address to send transactions or transmit data.

3.2. Security Model

This paper assumes that there exist external attackers in the proposed scheme. Attackers can analyze data on the chain, try to discover possible reporting behaviors, and use attack methods, such as address association, to destroy the anonymity of reporting. In addition, external attackers can also send invalid data on the chain to interfere.

In order to resist the attacks of the above-mentioned attackers and realize a secure reporting scheme, the following characteristics need to be met:

- Anonymity. The whistleblower should remain anonymous during the entire reporting process, and no adversary can distinguish between the whistleblower and normal users in the reporting process. In the rewarding process, the identity of the whistleblower remains anonymous, and no one can infer the identity of the whistleblower from the rewarding information.
- Covertness. The reporting data and reporting behavior should be indistinguishable from the normal uploaded data and the normal data uploading behavior of the system. Anyone other than the authority cannot confirm whether any user has made a report.

- Unforgeability. The report information sent by the whistleblower cannot be deleted, modified, or overwritten by anyone. No one can pretend to be a real whistleblower to receive rewards.

The formal definitions of the above characteristics are given below:

Definition 1. *(Correctness.) Given the security parameter λ, for any data generated by the* **Setup** *algorithm, as long as the user honestly executes the reporting algorithm, the probability of the authority passing the signature verification and successfully decrypting the covert information is 1; as long as the authority honestly executes the reward algorithm, the probability that the user can obtain the derived address private key and successfully receive the reward is 1. That is, for all* $\lambda \in \mathbb{N}^*$

$$\Pr\begin{bmatrix}\mathbf{Setup}(1^\lambda) \to (W, w, \mathcal{K}, \{k_1, k_2, \ldots, k_n\}, pp),\\ \mathbf{Upload}(data, Msg, \mathcal{K}, k_j, pp) \to (\sigma, \tilde{K}, A, B),\\ \mathbf{Extract}(\sigma, w, pp) \to Msg' :\\ \mathbf{Verify}(\sigma, \tilde{K}, \mathcal{K}, data, pp) = 1 \wedge\\ Msg' = Msg\end{bmatrix} = 1, \tag{1}$$

$$\Pr\begin{bmatrix}\mathbf{Setup}(1^\lambda) \to (W, w, \mathcal{K}, \{k_1, k_2, \ldots, k_n\}, pp),\\ \mathbf{Upload}(data, Msg, \mathcal{K}, k_j, pp) \to (\sigma, \tilde{K}, A, B),\\ \mathbf{Reward}(A, B, pp) \to P,\\ \mathbf{Gain}(a, b, pp) \to (x, P') :\\ P = P'\end{bmatrix} = 1. \tag{2}$$

Definition 2. *(Covertness.) Given the security parameter λ and a probabilistic polynomial time adversary $\mathcal{A}$, define the following experiment* $\mathrm{PrivK}_{\mathcal{A}}(\lambda)$*:*

- $\mathbf{Setup}(1^\lambda)$ *outputs public parameters, $\mathcal{A}$ generates a set of data and Msg.*
- *Randomly select a bit $b \leftarrow \{0, 1\}$, calculate $c_b = (\sigma_b, \tilde{K}_b, A_b, B_b) \leftarrow$* **Upload** $(data, Msg, \mathcal{K}, k_j, pp)$ *and $c_{1-b} = (\sigma_{1-b}, \tilde{K}_{1-b}, A_{1-b}, B_{1-b}) \leftarrow$* **Upload** $(data, \mathcal{K}, k_j, pp)$*, and send the obtained c_0 and c_1 to $\mathcal{A}$.*
- *$\mathcal{A}$ outputs a bit b'.*
- *If $b = b'$, the experiment outputs 1, otherwise it outputs 0.*

If $\mathrm{PrivK}_{\mathcal{A}}(\lambda) = 1$*, it can be said that $\mathcal{A}$ succeeded. It can be said that a scheme achieves covertness if*

$$\Pr[\mathrm{PrivK}_{\mathcal{A}}(\lambda) = 1] = \frac{1}{2} + \mathrm{negl}(\lambda). \tag{3}$$

Definition 3. *(Unforgeability.) Given the safety parameter λ, the probability for the PPT adversary $\mathcal{A}$ to forge reporting information and falsely claim rewards is* $\mathrm{negl}(\lambda)$*. That is*

$$\Pr\begin{bmatrix}\mathbf{Setup}(1^\lambda) \to (W, w, \mathcal{K}, \{k_1, k_2, \ldots, k_n\}, pp),\\ \mathcal{A}(data', , pp) \to (\sigma', \tilde{K}', A', B') :\\ \mathbf{Verify}(\sigma', \tilde{K}', data', pp) = 1\end{bmatrix} = \mathrm{negl}(\lambda). \tag{4}$$

$$\Pr\begin{bmatrix}\mathbf{Setup}(1^\lambda) \to (W, w, \mathcal{K}, \{k_1, k_2, \ldots, k_n\}, pp),\\ \mathbf{Upload}(data, Msg, \mathcal{K}, k_j, pp) \to (\sigma, \tilde{K}, A, B),\\ \mathbf{Reward}(A, B, pp) \to P,\\ \mathcal{A}(A, B, pp) \to (x, P') :\\ P = P'\end{bmatrix} = \mathrm{negl}(\lambda). \tag{5}$$

As in many existing blockchain-based covert communication models [5,24], this paper assumes that the attacker does not have access to the flow information of the network. If an attacker has access to the network flow, techniques from anonymous networks such as Tor can be used to ensure anonymity. This paper focuses on the anonymity of the addresses

on blockchain networks; flow analysis is not the focus of this paper and it will be put to future works.

4. The Proposed Anonymous and Covert Reporting Scheme

This section gives the specific algorithms of the anonymous and covert reporting scheme based on the blockchain. The proposed scheme includes six algorithms: *Initialization*, *Data Upload*, *Verification*, *Extraction*, *Sending Reward*, and *Receiving Reward*. The authority chooses a blockchain network and generates a private key–public key pair (w, W), where the public key W corresponds to the authority's address on the blockchain. Next, the authority deploys a smart contract to collect user data on the blockchain. Any user in the network can send data to the smart contract, but only the authority can read the content in the smart contract. After the user joins the blockchain network, it initializes, sends data to the smart contract, generates a ring signature on each data, and attaches public keys A, B for rewards and transactions. When the user wants to make a report, the user embeds the report material into the ring signature. The authority checks the data in the smart contract, verifies the signature, and checks whether it contains covert information. If ordinary data or covert reporting materials are valuable, the authority will use A and B transmitted by the user to reward. Other users cannot confirm the identity of the person receiving the reward, and only the true whistleblower can obtain the reward.

The main notations used in this paper and their descriptions are given in Table 1. The specific algorithms are as follows.

Table 1. Main notations and descriptions.

Notations	Descriptions
$E(\mathbb{F}_q)$	Elliptic curve over the finite field of order q
G	The generator of E
λ	The security parameter
W, w	The public and private key of the authority
$\mathcal{K}$	Users' public key group
n	Number of elements in $\mathcal{K}$
I_j	The user whose public key is the j-th element of $\mathcal{K}$
K_j, k_j	The public and private key that belongs to I_j
H_n, H_p	Hash functions
σ	Ring signature
A, B	Public keys
a, b	Private random numbers corresponding to A, B
$\tilde{K}$	Key image
$Data$	Data uploaded to the smart contract
Msg	Information that requires covert transmission
pp	The public parameter of the blockchain

4.1. Initialization

Given security parameter 1^λ, let E be an elliptic curve defined on the finite field $\mathbb{F}_q$, where q is the order of the elliptic curve, which is a large prime number close to 2^{256}. The generator of the elliptic curve $E(\mathbb{F}_q)$ is G. The authority chooses a random number $w \in \mathbb{F}_q$ as the private key and calculates $W = wG$ as the public key. The public key W corresponds to the authority's address on the blockchain. The authority deploys a smart contract by using a secret address to collect data on the blockchain. The authority sets the smart contract so that any user on the blockchain network can send data to this smart contract, but only one who knows the secret address can obtain the data on it, which contains all the data uploaded by users. In addition to the data themselves, the data sent to the smart contract require a ring signature and two different public keys A, B. In practical applications, this smart contract can be used in various scenarios such as data collection and government applications. The data upload process described in Section 4.2 is implemented based on this smart contract.

User I_j chooses a random number $k_i \in \mathbb{F}_q$ as the private key, calculates $K_i = k_i G$ as the public key, and the public key K_i corresponds to the user's address on the blockchain.

Let H_n and H_p be two different hash functions. $H_n : \{0,1\}^* \to \mathbb{F}_q$ is defined as a hash function mapping from an arbitrary string to a finite field $\mathbb{F}_q$ and $H_p : \{0,1\}^* \to E(\mathbb{F}_q)$ is defined as a hash function that maps from any string to the set of all points in the elliptic curve $E(\mathbb{F}_q)$.

The algorithm outputs the public and private key of authority W and w and the public and private keys of each user $\{K_1, K_2, \ldots, K_n\}$ and $\{k_1, k_2, \ldots, k_n\}$, composes the public key into an output public-key group $\mathcal{K} = \{K_1, K_2, \ldots, K_n\}$, and outputs the public parameter hash functions H_n and H_p.

4.2. Data Upload

Both the uploading process of user common data and the secret reporting process adopt data-uploading algorithms. The input of the algorithm is the *data* transmitted by the user I_j to the smart contract and the public-key group $\mathcal{K} = \{K_1, K_2, K_3, \ldots, K_j, \ldots, K_n\}$ on the blockchain and the public parameter pp of the system, where K_i is a point on the elliptic curve. When the user wants to make a secret report, the report material Msg is also used as the input of the algorithm. Regardless of whether the user makes a report, according to the information embedding rules, it is necessary to ensure that the length of the public key group $n > 4$ and $2 < j < n - 1$. When the user reports, since the information embedding amount of each group of data in the scheme is up to 64 bytes, the length of the reporting material Msg should be less than this length.

The algorithm calculates key image

$$\tilde{K} = k_j H_p(\mathcal{K}). \tag{6}$$

Generate a random number $\alpha \in \mathbb{F}_q$ and calculate

$$c_{j+1} = H_n(\mathcal{K}, \tilde{K}, data, \alpha G, \alpha H_p(\mathcal{K})). \tag{7}$$

Then, the algorithm generates a group of numbers $\{r_1, r_2, \ldots, r_n\}(r_i \in \mathbb{F}_q)$. When the user does not report but sends normal data to the smart contract, each r_i is a random number in the finite field $\mathbb{F}_q$. When the user reports, define coding methods $S \leftarrow \mathbf{Encode}(\mathcal{S})$ and $\mathcal{S} \leftarrow \mathbf{Decode}(S)$, which map a string $\mathcal{S}$ to a point S on the elliptic curve (and the reverse). These methods are usually based on the Koblitz method [30]. The basic idea is to obtain points on elliptic curves by calculating ordinate coordinates and plain data as horizontal coordinates. Section 3 of [30] details this encoding method. Then, the algorithm follows the steps below to generate $\{r_1, r_2, \ldots, r_n\}$:

- Fill random characters after the reporting material string Msg to obtain an Msg' with a 64-byte length. Then, calculate $M \leftarrow \mathbf{Encode}(Msg')$.
- Generate a random number $r \in E(\mathbb{F}_q)$ and calculate

$$C_1 = rG, \qquad C_2 = M + rW. \tag{8}$$

- Calculate $c_1 \leftarrow \mathbf{Decode}(C_1)$ and $c_2 \leftarrow \mathbf{Decode}(C_2)$ (c_1 and c_2 are 64-byte strings) and split each into two 32-byte data

$$c_1 = r_1 \parallel r_2, \qquad c_2 = r_{n-1} \parallel r_n. \tag{9}$$

Let the other r_i be the random numbers in the finite field $\mathbb{F}_q$. In the end, there is $[r_i] = \{r_1, r_2, \ldots, r_j, \ldots, r_{n-1}, r_n\}$.

Using the generated $[r_i]$, from $i = j + 2$ to $i = n$, calculate

$$D_i = r_{i-1}G + c_{i-1}K_{i-1}, \tag{10a}$$
$$E_i = r_{i-1}H_p(\mathcal{K}) + c_{i-1}\tilde{K}, \tag{10b}$$
$$c_i = H_n(\mathcal{K}, \tilde{K}, data, D_i, E_i). \tag{10c}$$

In the same way, calculate

$$D_1 = r_nG + c_nK_n, \tag{11a}$$
$$E_1 = r_nH_p(\mathcal{K}) + c_n\tilde{K}, \tag{11b}$$
$$c_1 = H_n(\mathcal{K}, \tilde{K}, data, D_1, E_1). \tag{11c}$$

Then, from $i = 2$ to $i = j$, use (10) to calculate c_j. At last, calculate

$$r_j = \alpha - c_jk_j. \tag{12}$$

Note that r_j is calculated by (12), not randomly generated, so r_j cannot be used to store covert information. Therefore, in the algorithm given in this section, the proposed scheme uses r_1, r_2, r_{n-1} and r_n to store covert information and lets $2 < j < n - 1$. This paper will discuss the relationship between the location of covert information and the anonymity of the proposed scheme in detail in Section 5.2.

In this way, the algorithm obtains the ring signature $\sigma = (c_1, r_1, r_2, \ldots, r_n)$. Then, the algorithm generates two random numbers a and b, $a, b \in \mathbb{F}_q$, and calculates $A = aG$ and $B = bG$. A and B are used for anonymous rewarding. Finally, the algorithm outputs the ring signature σ, key image $\tilde{K}$, and two public keys A, B. The user sends the output of the algorithm $(\sigma, data, \mathcal{K}, \tilde{K}, A, B)$ to the smart contract.

4.3. Signature Verification

The authority regularly checks the smart contract, recovers the data stored in the contract during this time, and obtains multiple messages $(\sigma, data, \mathcal{K}, \tilde{K}, A, B)$. In order to verify the correctness of the message, the ring signature needs to be checked. The input of the verification algorithm is the ring signature σ, $data$, public-key group $\mathcal{K}$, key image $\tilde{K}$, and public parameter pp of the system. The algorithm first computes

$$D_2' = r_1G + c_1K_1, \tag{13a}$$
$$E_2' = r_1H_p(\mathcal{K}) + c_1\tilde{K}, \tag{13b}$$
$$c_2' = H_n(\mathcal{K}, \tilde{K}, data, D_2', E_2'). \tag{13c}$$

Then, from $i = 3$ to $i = n$, compute

$$D_i' = r_{i-1}G + c_{i-1}'K_{i-1}, \tag{14a}$$
$$E_i' = r_{i-1}H_p(\mathcal{K}) + c_{i-1}'\tilde{K}, \tag{14b}$$
$$c_i' = H_n(\mathcal{K}, \tilde{K}, data, D_i', E_i'). \tag{14c}$$

At last, compute

$$D_1' = r_nG + c_n'K_n, \tag{15a}$$
$$E_1' = r_nH_p(\mathcal{K}) + c_n'\tilde{K}, \tag{15b}$$
$$c_1' = H_n(\mathcal{K}, \tilde{K}, data, D_1', E_1'). \tag{15c}$$

Check whether c_1' equals c_1 of the ring signature σ. If $c_1' = c_1$ holds, the ring signature is valid, the algorithm outputs 1, and the corresponding message can be used. Otherwise, if there is an error in data transmission or the data have been tampered with, the algorithm will output 0, and the message will be ignored.

4.4. Extraction

If the signature is valid, the authority will check whether there exists covert information in the message. The inputs of the extraction algorithm are the ring signature σ, the authority's private key w, and the public parameter pp of the system. Obtain $[r_i]$ from the ring signature σ, then combine r_1 and r_2 and r_{n-1} and r_n to obtain c_1' and c_2':

$$c_1' = r_1 \parallel r_2, \qquad c_2' = r_{n-1} \parallel r_n. \tag{16}$$

Map c_1' and c_2' to the points C_1' and C_2' on the elliptic curve, then calculate

$$M' = C_2' - wC_1'. \tag{17}$$

At last, calculate $Msg' \leftarrow \mathbf{Decode}(M')$. If the decoding method fails or the Msg' is a random string, it indicates that data were uploaded by a normal user and do not contain reporting information. Otherwise, the meaningful part of the Msg' is the reporting information.

4.5. Send Reward

If the normal information uploaded by the user or the information reported by the whistleblower is valuable, the user or the whistleblower will be rewarded by the authority. The A and B sent by the user to the smart contract and the public parameter pp of the system are the input of the algorithm. The algorithm computes

$$P = H_n(wA)G + B. \tag{18}$$

Then, convert the point P on the elliptic curve to an address on the blockchain and send a transaction to that address.

4.6. Receive Reward

After sending the information, the user adopts a rewarding algorithm. The input of the algorithm is the public keys A and B that are output by the data upload algorithm and the public parameter pp of the system. Compute

$$x = H_n(aW) + b. \tag{19}$$

From this calculation, the address P' on the blockchain corresponding to the private key x is obtained. Thereafter, the user periodically monitors the blocks on the blockchain. If the output address P of the newly generated transaction on the blockchain satisfies $P = P'$, this means that the address is a derived address generated by the user's private key, and the user can use the private key x calculated by the algorithm to receive this reward, then the user can use this address to transmit data or send transactions. Later, when the user makes a report, the derived address P is used to send data or report, and the original address can also send normal data.

5. Security Analysis and Experiments

This section will analyze the correctness and security of the proposed scheme and show its actual performance through experiments.

5.1. Security Analysis

The proposed scheme uses ring signature and derived address technology to achieve the security goal. The ring signature is used to ensure the anonymity of the report and the unforgeability of the reporting information and to embed the information to achieve the covertness of the scheme. The derived address is used to ensure that only true whistleblowers can receive rewards.

Theorem 1. *(Correctness.) The scheme meets the correctness proposed in definition 1; that is, if the user honestly implements the above scheme with the authority, the user's ring signature can pass the signature verification, the reporting information can be correctly decrypted by the authority, and the reward sent by the authority can be received by the user.*

Proof. According to the generation process and verification process of the ring signature, if the signature passes the verification, then

$$H_n(\mathcal{K}, \tilde{K}, data, D_1', E_1') = c_1' = c_1 = H_n(\mathcal{K}, \tilde{K}, data, D_1, E_1). \tag{20}$$

By the collision resistance of hash function, there is

$$r_n G + c_n' K_n = D_1' = D_1 = r_n G + c_n K_n, \tag{21a}$$
$$r_n H_p(\mathcal{K}) + c_n' \tilde{K} = E_1' = E_1 = r_n H_p(\mathcal{K}) + c_n \tilde{K}. \tag{21b}$$

That is, when $c_n' = c_n$ is satisfied, the ring signature can be verified successfully. Similarly, when $c_{j+1}' = c_{j+1}$ is satisfied, the ring signature can pass the verification. So, there is

$$c_{j+1} = H_n(\mathcal{K}, \tilde{K}, data, \alpha G, \alpha H_p(\mathcal{K})), \tag{22a}$$
$$\begin{aligned} c_{j+1}' &= H_n(\mathcal{K}, \tilde{K}, data, D_{j+1}', E_{j+1}') \\ &= H_n(\mathcal{K}, \tilde{K}, data, r_j G + c_j' K_j, r_j H_p(\mathcal{K}) + c_j' \tilde{K}) \\ &= H_n(\mathcal{K}, \tilde{K}, data, (\alpha - c_j k_j) G + c_j' K_j, (\alpha - c_j k_j) H_p(\mathcal{K}) + c_j' \tilde{K}) \\ &= H_n(\mathcal{K}, \tilde{K}, data, \alpha G + (c_j' - c_j) K_j, \alpha H_p(\mathcal{K}) + (c_j' - c_j) \tilde{K}). \end{aligned} \tag{22b}$$

So, when $c_j' = c_j$, $c_{j+1}' = c_{j+1}$ holds. In conclusion, the ring signature is verified successfully if and only if $c_j' = c_j$. From the proof above, similarly, when $c_2' = c_2$, the ring signature will pass the verification. That is

$$c_2' = H_n(\mathcal{K}, \tilde{K}, data, r_1 G + c_1 K_1, r_1 H_p(\mathcal{K}) + c_1 \tilde{K}) = c_2. \tag{23}$$

Therefore, if the user and the authority follow the proposed scheme honestly, the ring signature can be verified successfully.

In the embedding process of the report material, there is

$$c_1 = r_1 \parallel r_2 = c_1', \tag{24a}$$
$$c_2 = r_{n-1} \parallel r_n = c_2'. \tag{24b}$$

So, there is $C_1' = C_1$, $C_2' = C_2$, and

$$M' = C_2' - wC_1' = C_2 - wC_1 = M + rW - wrG = M. \tag{25}$$

The reporting information can be correctly decrypted by the authority.

In the rewarding process, there is

$$P = H_n(wA)G + B = (H_n(aW) + b)G = xG. \tag{26}$$

P is the public key corresponding to the private key x, and the user can obtain the private key of the address corresponding to P. □

Theorem 2. *(Anonymity.) In the reporting phase, the probability that the authority and third-party attackers can successfully link the ring signature in the smart contract with the signer's public key is* $1/(n-4)$. *In the rewarding phase, neither the authority nor third-party attackers can link different addresses that may belong to the same user.*

Proof. The anonymity of the reporting phase is guaranteed by the ring signature [11]. The ring signature is a digital signature scheme with unforgeability and unconditional anonymity, and its security can be reduced to the difficult problems of elliptic curves. The signer uses his public key and other users' public keys to form a public-key group and uses the public-key group to sign so that no one except the signer is able to know which public key in the public-key group the signer is using. So, in the normal ring signature scheme, the anonymity size is equal to the size of the public-key group. In the proposed scheme, according to the information embedding scheme, r_1, r_2, r_{n-1}, and r_n may store covert information, so $j \neq 1, 2, n-1, n$; that is, the user's public key is in the public-key group except the four at the beginning and end. Due to the characteristics of the ring signature, the probability of the user's public key at each place is equal. Therefore, the probability that the authority or third-party attacker can link the ring signature in the smart contract with the signer's public key, which is the probability of guessing the key used by the signer successfully, is $1/(n-4)$. Therefore, anonymity can be improved simply by increasing the length of the public-key group, but this method will sacrifice efficiency to a certain extent. This is a trade-off between efficiency and privacy, and it will be evaluated in future work.

The anonymity of the rewarding phase is guaranteed by the derived address. a and b selected by the user are random numbers in the finite field, and the corresponding A and B are random public keys. The adversaries cannot distinguish between the derived address and a random address without knowing the private random numbers a and b. Take Ethereum as an example; according to Etherscan, more than 160 million addresses are used in the network. Therefore, the anonymous set size in the rewarding phase can reach 100 million theoretically; that is,

$$\Pr\begin{bmatrix}\mathbf{Setup}(1^\lambda) \rightarrow (W, w, \mathcal{K}, \{k_1, k_2, \ldots, k_n\}, pp),\\ \mathbf{Upload}(data, Msg, \mathcal{K}, k_j, pp) \rightarrow (\sigma, \tilde{K}, A, B),\\ \mathbf{Reward}(A, B, pp) \rightarrow P,\\ \mathcal{A}(A, B, P, pp) \rightarrow K' :\\ K' = K_j\end{bmatrix} \leq \frac{1}{N} = \mathrm{negl}(\lambda). \tag{27}$$

N is the total number of addresses in the blockchain network. Hence, no third-party attackers can link the newly generated address with a user on the blockchain, and they cannot confirm the user's true identity.

The whistleblower will use the new derived address the next time he reports, which ensures that each report adopts a different address. For the same user, except for the initially registered address, other addresses are calculated based on random numbers a and b; there is no association between different addresses, which reduces the possibility of attackers attacking from the perspective of address association. Attackers, without knowing a and b, cannot link different addresses that may belong to the same user. □

Theorem 3. *(Covertness.) The proposed scheme satisfies the covertness proposed in Definition 2. For a probabilistic polynomial time adversary $\mathcal{A}$, the data embedded with covert information are indistinguishable from the data without embedded information.*

Proof. Data embedding is only related to the generation process of the ring signature σ. Specifically, it is only related to r_1, r_2, r_{n-1}, and r_n in σ. As long as it is proved that these four values in ring signatures that have embedded covert messages and values in those that do not contain any message are indistinguishable from the PPT adversary $\mathcal{A}$, the covertness of the proposed scheme can be proved.

When the covert information is not embedded, r_1, r_2, r_{n-1}, and r_n are random numbers. When embedding covert information, r_1, r_2, r_{n-1}, and r_n are the numbers where elliptic curve points C_1 and C_2 are encoded and split, $C_1 = rG$ and $C_2 = M + rW$, where r is a random number, so C_1 and C_2 are also random points on the elliptic curve. Due to the

security of elliptic curve encryption, the PPT adversary $\mathcal{A}$ cannot distinguish between the encrypted data r_1, r_2, r_{n-1}, and r_n and random data.

After receiving the ring signature, the authority computes $M = C_2 - wC_1$, where w is the authority's private key. Since the adversary $\mathcal{A}$ does not know the authority's private key, the $\mathcal{A}$ cannot compute M from C_1 and C_2. According to the security of elliptic curve encryption, the probabilistic polynomial time adversary $\mathcal{A}$ cannot obtain M from C_1 and C_2, so the $\mathcal{A}$ cannot speculate and decipher whether the random number contains covert information. In conclusion, the scheme meets the covertness proposed in Definition 2.

In actual scenarios, nodes in the blockchain network can send data to the smart contract at any time, and the authority can also reward users at any time. Regardless of whether the user uploads normal data or the whistleblower makes a report, the authority will reward, based on the original meaning of the data themselves. For example, in a tax system, the authority can refund tax based on data uploaded by users. When a user reports, he only needs to use the information embedding algorithm to modify the random number used in the ring signature while sending the data normally, and the modified result is indistinguishable from the random value. In this case, the adversary $\mathcal{A}$ cannot distinguish whether the user has reported through transaction data on the blockchain, and the user's reporting behavior can be hidden. □

Theorem 4. *(Unforgeability.) The probability of any probability polynomial time adversary $\mathcal{A}$ forging reporting information and falsely claiming rewards is negligible.*

Proof. The unforgeability of the proposed scheme mainly comes from the unforgeability of blockchain. The PPT adversary forges reporting information on the premise of changing the data uploaded by the user to the smart contract. Data sent by users will be written directly into the smart contract. The way that the adversary $\mathcal{A}$ can tamper with the reported information without being discovered is to occupy more than 50% of the hash power of the blockchain network, or the nodes owned by the adversary $\mathcal{A}$ can cover the path of the user's outgoing data. It is believed that, in actual scenarios, the adversary $\mathcal{A}$ cannot control more than half of the nodes in the network, and as the network expands, the probability that the adversary $\mathcal{A}$ can completely cover the path of the user's outgoing data will rapidly decrease, which can be regarded as negligible. In addition, due to the covertness of the reporting information, the attacker cannot confirm whether the user has sent the reporting information or whether the information needs to be tampered with or shielded, and this attack method has no practical significance. Therefore, the probability that any PPT adversary $\mathcal{A}$ can clear or modify the data sent by the user is negligible.

The unforgeability of the rewarding phase also comes from the security of elliptic curve encryption. Since the adversary $\mathcal{A}$ does not know the authority's private key w, it is impossible to establish a connection between the derived address P and the A and B uploaded by the user. Without knowing the random numbers a and b corresponding to A and B, the adversary $\mathcal{A}$ cannot calculate the private key x corresponding to the address. Due to the security of elliptic curve encryption, the probability of the occurrence of these two events is negligible. Therefore, the probability that the PPT adversary $\mathcal{A}$ receives a reward that does not belong to him is negligible. In summary, the scheme proposed in this paper has unforgeability. □

As shown in Table 2, existing anonymous reporting schemes (e.g., [1,12]) do not consider the covertness of the reporting behavior as their design goal, and their reporting behavior is public in the system. The proposed scheme can additionally guarantee the covertness of the reporting behavior. Through encryption and message embedding, the reporting information behaves indistinguishably from random numbers. There is no way for any third party to distinguish the normal act of uploading data from the act of reporting. In addition, existing covert communication systems (e.g., [5,6]), as mentioned in Section 2.3, have the problem of address association and transaction association, which means that anonymity cannot be guaranteed. The proposed scheme is able to avoid these

two problems. In the reporting system of this paper, the user obtains the initial address through registration and then sends data to the smart contract. After that, the addresses adopted by the user are all calculated through elliptic curve calculations using A and B randomly selected by the user. The authority's private key w is also used during calculation, so third-party attackers cannot obtain the successive address adopted by the user through A and B and cannot know which user the new address corresponds to, so the former address and the new address obtained by the user cannot be associated to the same user. Though the new address is obtained through a transaction, this transaction is one-way, the two addresses do not appear in the same transaction, and the authority will not send any more transactions to this new address; therefore, the flow of funds cannot determine which addresses are related either. Generally speaking, the proposed scheme achieves anonymity, covertness, and unforgeability.

Table 2. Comparison with existing anonymous reporting schemes and covert communication schemes.

Schemes	Unforgeability	Anonymity	Covertness
BB2AR [12]	√	√	×
ReportCoin [1]	√	√	×
HC-CDE [5]	√	×	√
NNCCM [6]	√	×	√
The proposed scheme	√	√	√

5.2. Experiment

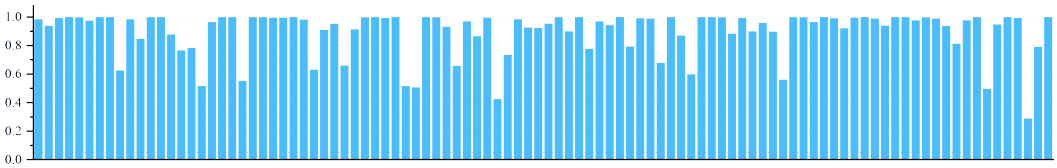
In order to show the practical performance of the proposed scheme, this paper implements the scheme. The experiment was run on a PC and the environment of the experiment is as follows:

- CPU: Intel Core i7-10875H @ 2.30 Ghz;
- OS: Windows 10 20H2;
- Memory: 16.0 GB.

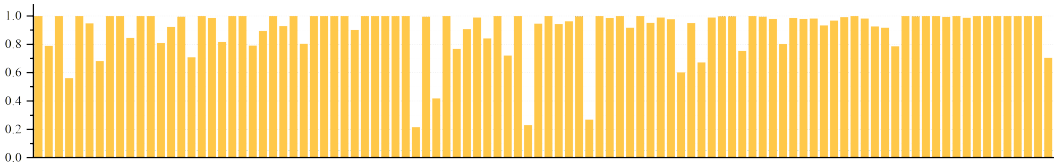
Rinkeby was used as the experimental environment of the blockchain and the smart contract was deployed on the node that represents the authority. The elliptic curve used in the experiment is consistent with that of Ethereum: both are SECP256k1 curves, and the order q is a 256-bit prime number. Keccak256 was used as the hash function in the experiment, which is also the hash function used by Ethereum. Python 3.7.6 was used to implement the proposed scheme, and libraries used in the experiment are ecdsa version 0.17.0, pycryptodome version 3.9.7, and pysha3 version 1.0.2.

In the experiment, the user node was used to send 100 transactions in which covert information was embedded (called a covert transaction) and 100 transactions without embedding any information (called a normal transaction). One transaction without embedded covert information was selected as the base transaction, and the KS and KLD values of other transactions were calculated. Figure 2 shows the KS and KLD values of these transactions. The Kolmogorov–Smirnov (KS) test is a test method that compares a frequency distribution $f(x)$ with a theoretical distribution $g(x)$ or the distribution of two observations. In the KS test, the p value is usually used to judge whether there is a difference in the distribution of the two samples tested, and p is the concomitant probability. When $p > 0.05$, the two samples are considered to conform to the same distribution. The calculated minimum p-value for the normal transactions is 0.289, and the average p-value is 0.900, which is much higher than 0.05, indicating that the KS test can be used to test the covertness of covert transactions. The minimum p-value for the covert transactions is 0.215, and the average p-value is 0.913, which is also much higher than 0.05, indicating that the difference between covert transactions and normal transactions is not significant, and the proposed scheme can resist the KS test. The Kullback–Leibler Divergence (KLD) test is a method to detect the difference between two probability distributions. KLD is used to measure the distance between two random distributions. When the distributions are the same, the KLD value is

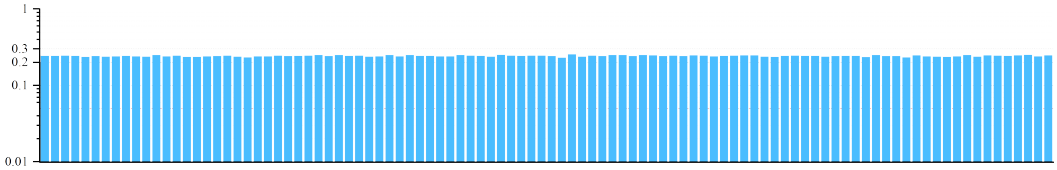
0; when the distribution difference becomes larger, the KLD value will also become larger. This paper calculates the KLD values between the base transaction and other normal transactions: the calculated maximum KLD value is 0.256 and the minimum is 0.230. The value 0.3 was set as the KLD threshold and the KLD values between the base transaction and covert transactions were calculated. If the maximum KLD of covert transactions is smaller than the threshold, it is said that the difference between the two kinds of transactions is not significant. The calculated maximum KLD value of the covert transactions is 0.256, which is less than the set KLD threshold, meaning that the proposed scheme can resist the KLD test. The above KS and KLD experiments prove that the proposed scheme achieves covertness, and there is strong indistinguishability between the data packets embedded with covert information and the data packets without any information embedded.



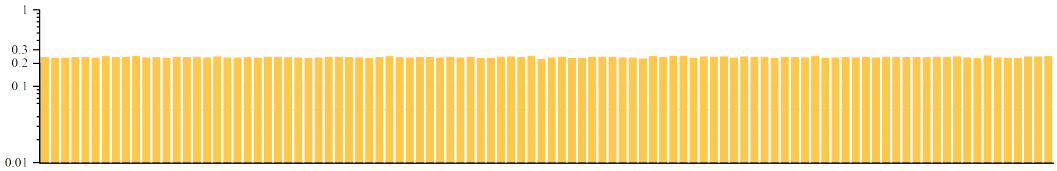
(a) KS test results of normal transactions



(b) KS test results of covert transactions



(c) KLD test results of normal transactions



(d) KLD test results of covert transactions

Figure 2. KS and KLD test results of normal transactions and covert transactions.

The algorithms of reporting, verification, and decryption in the scheme are all completed off-chain, and their running times are shown in Table 3 (let the length of the public-key group $n = 10$).

Table 3. Runtime of the sign, report, verification, and decryption algorithms.

	Sign	Report	Verify	Decrypt
Embedding message	71.94 ms	79.31 ms	67.05 ms	2.00 ms
Not embedding message	67.76 ms	74.99 ms	67.06 ms	1.93 ms

In Table 3, the running time of the reporting algorithm includes data generation, processing, and signing but does not include the process of uploading to the smart contract,

because the time for data uploading is only relevant to the state of the blockchain network. It can be seen that, when the length of the public-key group is 10, the running time of each algorithm in this paper's scheme does not exceed 0.1 s. The signing algorithm with embedded covert information consumes only 4.18 ms more on average than when no steganography is embedded. About 90% of the time in the reporting algorithm is spent in the signing process, and the time used for decryption is small compared with the time used for signature verification. Most of the time is spent on the generation and verification of the ring signature in the algorithm. When embedding information, the reporting message needs to be encrypted and processed in the signing algorithm, and it consumes a little more time compared with the time when no information is embedded, but it takes very little in the entire signing algorithm. Regardless of whether the information is embedded or not, the time cost of the rest of the algorithm is the same. It should be noted that, since the authority does not know whether the covert information is embedded in the data, the time cost of the decryption process cannot be omitted.

In the anonymous and covert reporting scheme proposed in this paper, the user and the authority do not need to interact in real-time but only need to ensure that the authority can receive the data sent by the user and the user can receive the transaction sent by the authority. Since the interactions between the user and the authority in the data uploading and rewarding process are based on sending transactions in the blockchain network, the cost and efficiency of the proposed scheme are directly related to the blockchain network used. In the Rinkeby network used in the experiment, it takes about 0.0004 RIN (about USD 1.3) to send data to the smart contract, and it takes about 15 s for the transaction to be packaged into a block.

The efficiency of the entire algorithm is also related to the length of the public-key group used in the ring signature, as shown in Figure 3. It can be seen that the algorithm for generating and verifying signatures has a linear relation with the number n of public keys in the public-key group. The more public keys in the public-key group, the longer it takes to generate the signature. Use $|\mathbb{F}_q|$ to represent the length of the elements in $\mathbb{F}_q$, then the length of the ring signature σ is $(n+1)|\mathbb{F}_q|$, and the length of the data uploaded to the smart contract is $(2n+4)|\mathbb{F}_q| + |data|$, including the ring signature, public-key group, key image, elliptic curve random point A, B, and data themselves. When the length of the uploaded data is the same, the length of the total uploaded data also has a linear relation with the number of public keys in the public-key group. Generally speaking, the more public keys used in the signing process, the larger the anonymity set, and the scheme has higher anonymity, but it will bring about a decrease in efficiency. In addition, since the proposed scheme only embeds covert information in the four random numbers at the beginning and end of the ring signature, the decryption time is independent of the length of the public-key group. As shown in Figure 4, no matter how the length of the public-key group changes, the time required for decryption remains essentially the same as long as the number of random numbers used for embedding remains the same.

Figure 4 shows the relation between the anonymity set size, channel capacity, public-key group length, and the length of the public-key group occupied by covert information. Anonymity size is related to the number of random numbers not occupied by the covert information embedding algorithm. When the length of the public-key group is constant, the greater the volume of data to transfer per transaction means the smaller the size of the anonymity set and the lower the security level of the algorithm. However, the channel capacity of the scheme depends on the number of random numbers used to embed covert information. As shown in Figure 4b, using 4 random numbers to store covert information (which is used in this paper) can transmit 64 bytes of data, 6 random numbers can transmit 128 bytes, 8 random numbers can transmit 192 bytes, and so on. In the actual application process, to improve communication efficiency while ensuring certain anonymity, it can be considered to use more random numbers to store covert information when the length of the public-key group is large.

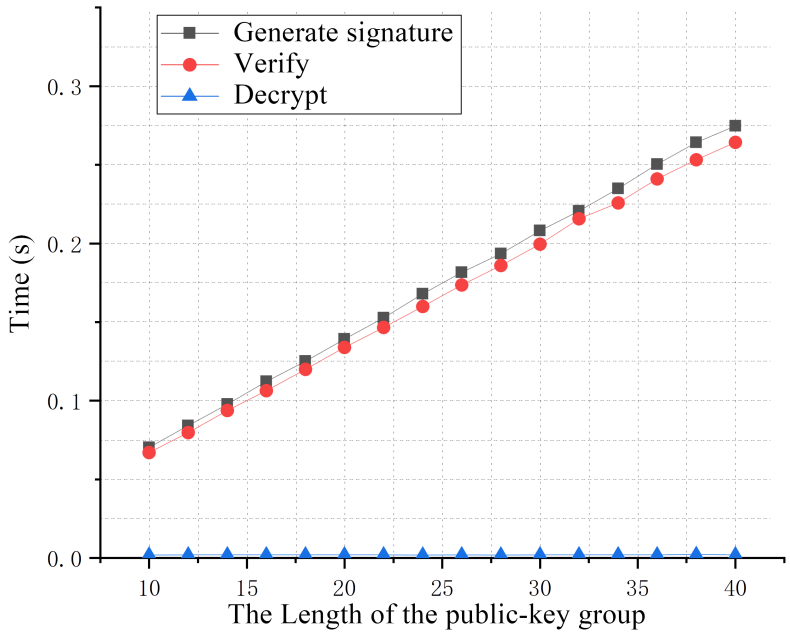


Figure 3. The relation between the length of the public-key group and the running time of the algorithm.

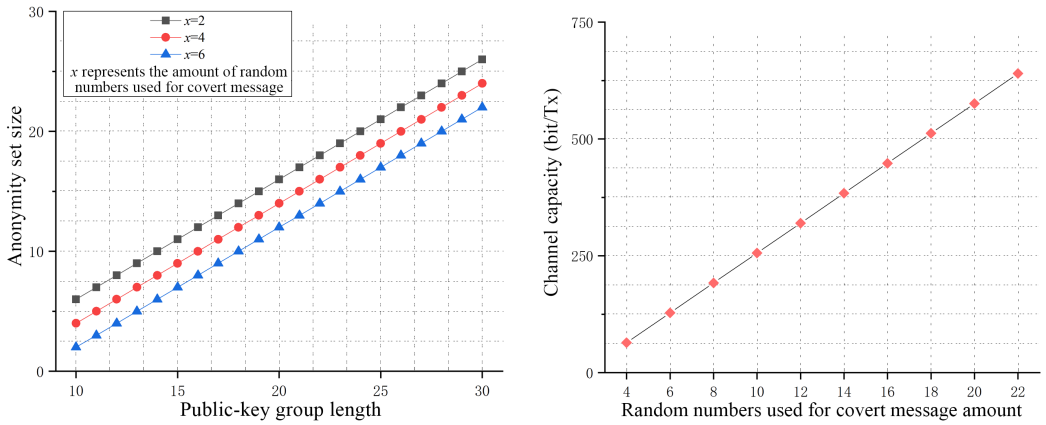


Figure 4. (**a**) shows the relation between the length of the public-key group and the size of the anonymity set when the degree of embedding of the covert information is the same (i.e., the same channel capacity), and (**b**) shows the relation between the number of random numbers occupied by the covert information and the channel capacity.

In summary, the results of the KS and KLD tests show that the proposed scheme has a good covertness. The normal data sent by the user and the data embedded with covert information can be indistinguishable, which ensures that reporting behaviors are hidden throughout the whole process. Based on the analysis of the running time of the algorithm, the running time of the proposed scheme (regardless of the time cost on the chain) is only milliseconds, and each data upload only needs to pay the transaction fee of the blockchain, which has an acceptable running efficiency and cost in actual scenarios. By simply modifying the information embedding algorithm, the relation between the size of the anonymity set and the length of information embedded can be flexibly adjusted to improve the efficiency of information transmission. By introducing covert communication technology into the anonymous reporting scheme, the proposed scheme realizes the covertness of reporting behavior in the whole process.

6. Conclusions

Reporting plays a very important role in the process of combating crimes. Blockchain-based reporting systems achieve strong anonymity and guarantee the privacy of whistleblowers. However, the existing reporting systems do not consider hiding the reporting process, which brings some security risks. By using ring signature, derived address, and public-key cryptographic encryption technology on the blockchain, the proposed scheme embeds data into ring signatures by public-key encryption to ensure covertness, guarantee anonymity by the ring signature and derived address, and propose an anonymous covert reporting scheme. Through theoretical and experimental analyses, it is proved that the proposed scheme achieves anonymity, covertness, and unforgeability and has the operational efficiency and cost to meet the actual requirement.

Although the proposed scheme achieves the security goal, it can be further improved in the dimension of reducing transmission costs. Specifically, embedding more data in a single transaction will inevitably increase the length of the public-key group in the ring signature, which will increase the computation and bring more expensive transaction transmission costs. In addition, how to ensure the reliability of anonymous reporting information is also an important issue. We are actively seeking better solutions to these problems in the future.

Author Contributions: Conceptualization, L.Z.; Methodology, L.Z.; Software, J.Z. and Z.C.; Validation, C.Z., F.G. and Z.L.; Formal analysis, J.Z. and C.Z.; Investigation, J.Z., F.G. and Z.L.; Writing—original draft, J.Z., C.Z. and F.G.; Writing—review & editing, L.Z., Z.C. and Z.L.; Visualization, Z.C. and Z.L.; Supervision, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Natural Science Foundation of China (Grants No. U1836212, No. 61872041, and No. 62172040) and National Key Research and Development Program of China (Grants No. 2021YFB2701200 and No. 2022YFB2702402).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zou, S.; Xi, J.; Wang, S.; Lu, Y.; Xu, G. Reportcoin: A novel blockchain-based incentive anonymous reporting system. *IEEE Access* **2019**, *7*, 65544–65559. [CrossRef]
2. Wang, H.; He, D.; Liu, Z.; Guo, R. Blockchain-based anonymous reporting scheme with anonymous rewarding. *IEEE Trans. Eng. Manag.* **2019**, *67*, 1514–1524. [CrossRef]
3. Zhang, L.; Xu, J. Blockchain-based anonymous authentication for traffic reporting in VANETs. *Connect. Sci.* **2022**, *34*, 1038–1065. [CrossRef]
4. Chen, Z.; Zhu, L.; Jiang, P.; Zhang, C.; Gao, F.; He, J.; Xu, D.; Zhang, Y. Blockchain Meets Covert Communication: A Survey. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 2163–2192. [CrossRef]
5. Cao, H.; Yin, H.; Gao, F.; Zhang, Z.; Khoussainov, B.; Xu, S.; Zhu, L. Chain-based Covert Data Embedding Schemes in Blockchain. *IEEE Internet Things J.* **2020**, *9*, 14699–14707. [CrossRef]
6. Qin, J.; Luo, Y.; Xiang, X.; Tan, Y. A Novel Network Covert Channel Model Based on Blockchain Transaction Parity. In Proceedings of the International Conference on Artificial Intelligence and Security, Dublin, Ireland, 19–23 July 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 54–63.
7. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. *Decentralized Bus. Rev.* **2008**, 21260.
8. Wood, G. Ethereum: A secure decentralised generalised transaction ledger. *Ethereum Proj. Yellow Pap.* **2014**, *151*, 1–32.
9. Harrigan, M.; Fretter, C. The Unreasonable Effectiveness of Address Clustering. In Proceedings of the 2016 Intl IEEE Conferences on Ubiquitous Intelligence Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), Toulouse, France, 18–21 July 2016; pp. 368–373. [CrossRef]
10. Kurt, M. Alonso and Jordi Herrera Joancomartí. Monero-Privacy in the Blockchain. Cryptology ePrint Archive, Report 2018/535. 2018. Available online: https://eprint.iacr.org/2018/535 (accessed on 21 March 2023). [CrossRef]
11. Rivest, R.L.; Shamir, A.; Tauman, Y. How to leak a secret. In Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security, Gold Coast, Australia, 9–13 December 2001; Springer: Berlin/Heidelberg, Germany, 2001; pp. 552–565.

12. Lu, Y.; Tang, Q.; Wang, G. Zebralancer: Private and anonymous crowdsourcing system atop open blockchain. In Proceedings of the 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS), Vienna, Austria, 2–6 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 853–865.
13. Lu, Z.; Wang, Q.; Qu, G.; Liu, Z. Bars: A blockchain-based anonymous reputation system for trust management in vanets. In Proceedings of the 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE), New York, NY, USA, 1–3 August 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 98–103.
14. Yao, Y.; Chang, X.; Mišić, J.; Mišić, V.B.; Li, L. BLA: Blockchain-assisted lightweight anonymous authentication for distributed vehicular fog services. *IEEE Internet Things J.* **2019**, *6*, 3775–3784.
15. Simmons, G.J. The prisoners' problem and the subliminal channel. In Proceedings of the Advances in Cryptology, Paris, France, 9–11 April 1984; Springer: Berlin/Heidelberg, Germany, 1984; pp. 51–67. [CrossRef]
16. Johnson, N.F.; Jajodia, S. Exploring steganography: Seeing the unseen. *Computer* **1998**, *31*, 26–34.
17. Bohli, J.M.; Vasco, M.I.G.; Steinwandt, R. A subliminal-free variant of ECDSA. In Proceedings of the International Workshop on Information Hiding, Alexandria, VA, USA, 10–12 July 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 375–387. [CrossRef]
18. Hartl, A.; Annessi, R.; Zseby, T. A subliminal channel in EdDSA: Information leakage with high-speed signatures. In Proceedings of the 2017 International Workshop on Managing Insider Security Threats, Dallas, TX, USA, 30 October–3 November 2017; pp. 67–78.
19. Dong, Q.; Li, X.; Liu, Y. Two extensions of the ring signature scheme of Rivest–Shamir–Taumann. *Inf. Sci.* **2012**, *188*, 338–345.
20. Ahsan, K.; Kundur, D. Practical data hiding in TCP/IP. In Proceedings of the Workshop on Multimedia Security at ACM Multimedia, French Riviera, France, 6 December 2002; Volume 2, pp. 1–8. [CrossRef]
21. Lucena, N.B.; Lewandowski, G.; Chapin, S.J. Covert channels in IPv6. In Proceedings of the International Workshop on Privacy Enhancing Technologies, Cavtat, Croatia, 30 May–1 June 2005; Springer: Berlin/Heidelberg, Germany, 2005; pp. 147–166.
22. Perkins, M.C. Hiding out in Plaintext: Covert Messaging with Bitwise Summations. Master's Thesis, Iowa State University, Ames, IA, USA, 2005. [CrossRef]
23. Partala, J. Provably secure covert communication on blockchain. *Cryptography* **2018**, *2*, 18.
24. Yanfeng, L.; Liping, D.; Jingzheng, W.; Qiang, C.; Xuehua, L.; Bei, G. Research on a new network covert channel model in blockchain environment. *J. Commun.* **2019**, *40*, 67. [CrossRef]
25. Gao, F.; Zhu, L.; Gai, K.; Zhang, C.; Liu, S. Achieving a covert channel over an open blockchain network. *IEEE Netw.* **2020**, *34*, 6–13. [CrossRef]
26. Zhang, P.; Cheng, Q.; Zhang, M.; Luo, X. A Group Covert Communication Method of Digital Currency Based on Blockchain Technology. *IEEE Trans. Netw. Sci. Eng.* **2022**, *9*, 4266–4276. [CrossRef]
27. She, W.; Huo, L.J.; Liu, W.; Zhang, Z.H.; Song, X.; Tian, Z. A Blockchain-Based Covert Communication Model for Hiding Sensitive Documents And Sender Identity. *Acta Electonica Sin.* **2022**, *50*, 1002. [CrossRef]
28. Monaco, J.V. Identifying bitcoin users by transaction behavior. In Proceedings of the Biometric and Surveillance Technology for Human and Activity Identification XII. International Society for Optics and Photonics, Baltimore, MD, USA, 22 April 2015; Volume 9457, p. 945704. [CrossRef]
29. Zheng, B.; Zhu, L.; Shen, M.; Du, X.; Guizani, M. Identifying the vulnerabilities of bitcoin anonymous mechanism based on address clustering. *Sci. China Inf. Sci.* **2020**, *63*, 1–15.
30. Koblitz, N. Elliptic curve cryptosystems. *Math. Comput.* **1987**, *48*, 203–209.