Blockchain Meets Covert Communication: A Survey

Zhuo Chen, Liehuang Zhu, *Senior Member, IEEE,* Peng Jiang*, *Member, IEEE,* Can Zhang, Feng Gao, Jialing He, Dawei Xu, Yan Zhang, *Fellow, IEEE*

***Abstract*—Covert communication enables covert information transmission in an undetectable way to prevent the exposure of communication behaviors. Blockchain-based covert communication breaks through the limitations on concealment, reliability and anti-traceability, and has shown promising application prospects in both sensitive data transmission and botnets. Although there are studies on blockchain-based covert communication, it still lacks a systematic investigation. In this paper, we conduct a comprehensive study on channel building and survey its core technologies by information embedding, transaction filtering, and transaction obfuscation. We also summarize evaluation metrics to better analyze blockchain-based covert channels. Privacy aspects are also discussed. Finally, we suggest seven future directions to stir research efforts into this area.**

***Index Terms*—Blockchain, Covert Communication, Covert Channel, Bitcoin, Ethereum.**

I. Introduction

EXPOSING communication behaviors may cause severe damages, such as the leakage of sensitive data and the discovery of cyber-espionage groups [1], [2]. Covert communication [3] can effectively hide communication behaviors based on steganography and information hiding techniques. Covert channels [4] are needed as the media to transmit covert information in the covert communication process and established in public channels. With developments of computer networks, network protocols are used to construct covert channels, i.e., network-based covert channels (covert channels for short). Traditional Covert channels usually hide covert information through the network protocol composed of protocol data units (PDUs). However, these traditional covert channels cannot guarantee concealment, reliability, and anti-traceability, and have the following drawbacks:

- *Simple to detect:* The network protocol analyzer like Wireshark [5] easily monitors PDUs. Insufficient storage fields of PDUs limit the way to embed covert information. Statistical analysis methods such as Kullback Leibler Divergence (KLD) and Kolmogorov-Smirnov (K-S) tests can easily detect PDU carrying specific covert information [6], [7], [8], [9], [10]. Therefore, traditional covert channels tend to lack concealment.
- *Effortless to destroy:* Traditional covert channels cannot guarantee reliability since everyone who has access to the network can easily restrict [11], [12], [13], [14], [15] or eliminate [16], [17] these channels. Specifically, modifying the inter-packet delay (IPD) between PDUs could greatly reduce the channel capacity and restrict corresponding channels ultimately; Normalizing redundant or reserved fields would eliminate covert information.
- *Easy to trace:* PDU explicitly records the IP address to represent communication parties' identities. Once covert channels are discovered, communication parties' identities are directly revealed. Therefore, communication parties' identities are easy to trace.

Blockchain [18] is a distributed ledger technique characterized by flooding propagation, decentralization, and anonymity. It offers natural shelter for covert channels. In general, blockchain-based covert channels [19] hide covert information through transactions and have the following advances:

- *Concealed channels:* Flooding propagation permits non-directional sending and receiving. For non-directional sending, the sender does not specify the receiver to transmit special transactions that carry covert information. For non-directional receiving, the receiver receives all on-chain transactions rather than only special transactions. Due to such non-directional communications, covert channels are more concealed.
- *Reliable communications:* The decentralization property determines that no one can modify special transactions. That is, special transactions created by the sender are the same as those received by the receiver and covert information cannot be eliminated. Covert channels are hence reliable.
- *Anti-traceable identities:* Unlike PDU containing IP addresses, transactions only contain anonymous addresses. It is difficult to link these addresses to the sender. Even though blockchain-based covert channels are exposed, communication parties' identities are untraceable.

However, blockchain's inherent features like low throughput, flooding propagation, openness, and transparency incur new challenges and impede the construction of blockchain-based covert channels:

- *How to communicate efficiently:* It takes a long time to confirm a transaction due to low throughput. Only when special transactions are confirmed, the receiver can extract covert information, which increases transmission delay.

Zhuo Chen, Liehuang Zhu, Peng Jiang, Can Zhang, and Feng Gao are with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China (e-mail: chen.zhuo@bit.edu.cn; liehuangz@bit.edu.cn; pennyjiang0301@gmail.com; canzhang@bit.edu.cn; gaofengbit@foxmail.com).

Jialing He is with the College of Computer Science, Chongqing University, Chongqing, China (e-mail: hejialing@cqu.edu.cn).

Dawei Xu is with the School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China, and also with the College of Cybersecurity, Changchun University, 6543 Weixing Road, Changchun 130022, China (e-mail: xudw@ccu.edu.cn).

Yan Zhang is with Department of Informatics, University of Oslo, Norway, and also with Simula Metropolitan Center for Digital Engineering, Norway (e-mail: yanzhang@ieee.org).

*Corresponding author.

Besides, insufficient bits that can be directly modified lead to the restricted embedding capacity in a transaction. The amount of covert information that can be included in each special transaction is rather small. Thus, it is challenging to achieve covert communication efficiently.

- *How to identify special transactions:* To guarantee the anonymity of the receiver, the sender sets output addresses of special transactions as his/her own addresses rather than the receiver's addresses. In addition, the sender would not share extra transaction information, such as input addresses and transaction amounts, with the receiver. Meanwhile, transactions are transmitted through flooding propagation instead of peer-to-peer communication. Overall, it is difficult for the receiver to identify special transactions.
- *How to conceal special transactions:* Covert communication generally transmits encrypted information, which may expose special transactions. For example, directly embedding the ciphertext into transactions' amounts would make these amounts random. However, the unmodified transactions' amounts do not possess randomness. The noticeable gap between special transactions' amounts and unmodified transactions' amounts exposes special transactions. Moreover, the difference is markedly visible because of the openness and transparency of transactions, thereby making it difficult to conceal special transactions.

A. Related Work

We present related work in this section. We investigate the existing work from two perspectives: 1) Existing work on blockchain, covert communication, and communications in blockchain; 2) Existing work on combining blockchain and covert communication.

Several researchers have conducted surveys on blockchain, covert communication, and communications in blockchain, while most of them do not particularly focus on the covert communication in blockchain. *For surveys on blockchain,* Monrat *et al.* [18] presented a comparative study of the tradeoffs of blockchain. Bodkhe *et al.* [20] gave a systematic review of various blockchain-based solutions in Industry 4.0-based applications. Besides, several studies on the applications of blockchain in different scenarios have been presented. More specifically, [21] focused on the information systems management scenario and its security; [22] discussed the benefits, challenges, and functionalities when applying blockchain to government, financial, manufacturing, and healthcare sectors; [23] explored the current status, applications, and directions of blockchain in supply chain management; [24] reviewed the trends and challenges in blockchain-based Internet of things (IoT) and blockchain-based data management; [25] examined the applications of blockchain technology in food supply chains, agricultural insurance, smart farming, and transactions of agricultural products; [26] presented a detailed survey on current blockchain applications for artificial intelligence (AI) and open challenges. Bernabe *et al.* [27] and Zhang *et al.* [28] investigated the security and privacy of blockchain. Bernabe *et al.* [27] focused on the security and privacy of diverse blockchain scenarios encompassing, eGovernment, eHealth, cryptocurrencies, smart cities, and cooperative intelligent transport systems, whereas Zhang *et al.* [28] focused on the security and privacy of blockchain with respect to concept, attributes, techniques, and systems.

For surveys on covert communication, Makhdoom *et al.* [29] highlighted the techniques, limitations, and challenges of covert communication. Gopalan *et al.* [30] and GGCT *et al.* [31] focused on audio steganography methods for realizing covert communication. Qiao *et al.* [32] and Shen *et al.* [33] conducted underwater covert communication, where Shen *et al.* [33] reviewed the development status of underwater bionic covert communication and Qiao *et al.* [32] particularly focused on bionic covert underwater acoustic communication using cetacean vocals. Jing *et al.* [34] discussed key technologies of constructing covert channels for new emerging networks. Dai *et al.* [35] reviewed wireless covert communication. Li *et al.* [36] analyzed the covert channel in Android systems and the detection schemes. Caviglione *et al.* [37] investigated challenges in the development of countermeasures against network covert channels.

For surveys on communications in blockchain, related reviews mainly focused on the communication in vehicle-to-everything (V2X) [38], IoT [39], [40], 5th generation mobile communication technology (5G) [41], [42], and aerial communications [43].

For surveys on combining blockchain and covert communication, Makhdoom *et al.* [29] displayed a few typical blockchain-based covert communication schemes. However, they only compare the pros and cons of blockchain-based approaches and classic steganography methods at a high level. Tian *et al.* [34] presented five covert channels in the context of the blockchain. Nevertheless, they focused on introducing new network environments that can be used to build covert channels, including streaming media, blockchain, and Internet protocol version 6 (IPv6). Bhutta *et al.* [44] and Singh *et al.* [45] presented that covert communication in blockchain enhanced the reliability, confidentiality, and integrity of data transmission, while they focused on the security and attacks of blockchain systems. Gadekallu *et al.* [46] presented an implementation of covert communication in blockchain for data privacy of edge of things, while they focused on applications, opportunities, and challenges of blockchain for edge of things. Kim *et al.* [47] presented that the covert channel in blockchain can perform decentralized identifiers (DID) wallet system attacks, while they focused on the security of blockchain-based DID services.

B. Scope of This Survey and Contributions

A systematic survey of covert communication in blockchain is still missing, which motivates us to launch a comprehensive and deep investigation and provide a roadmap for follow-up researchers. Compared with related work, we provide the first survey of blockchain-based covert communication, including channel construction, channel evaluation, privacy aspects, challenges, and future directions. The main contributions of this survey are listed as follows:

This article has been accepted for publication in IEEE Communications Surveys & Tutorials. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/COMST.2022.3204281

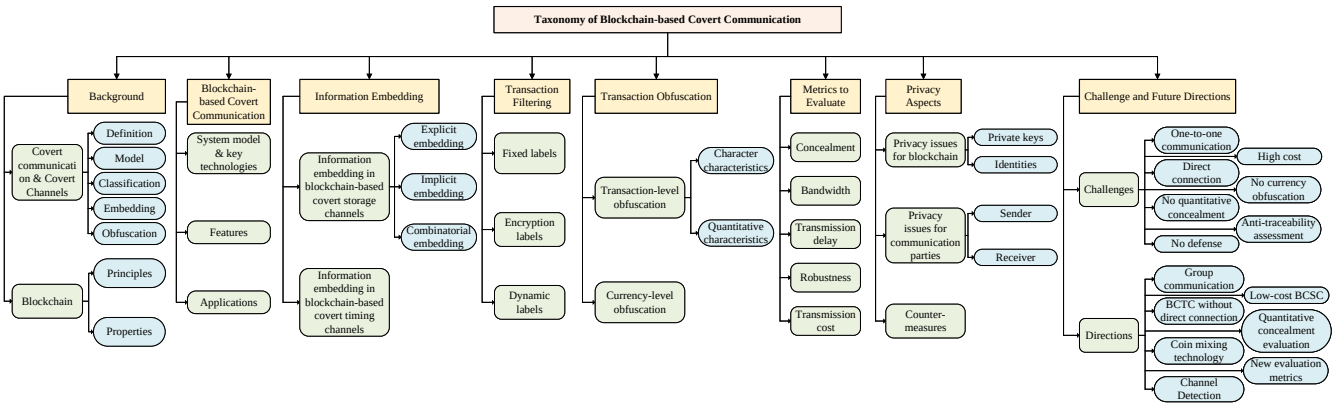


Fig. 1: The taxonomy of this survey.

- *Summarize blockchain-based covert communication:* To our best knowledge, we are the first to study blockchain-based covert communication systematically, and the survey covers almost all achievements related to blockchain-based covert communication up to November 2021.
- *Provide an overview of blockchain-based covert communication:* We provide a comprehensive overview of blockchain-based covert communication, including the definition, the model, and features. We also discuss potential application scenarios.
- *Detail existing blockchain-based covert channels:* We summarize three key technologies on channel building: information embedding, transaction filtering, and transaction obfuscation. These three technologies are then comprehensively introduced and classified. Before each technology is introduced, we briefly describe how it works in traditional covert channels. Meanwhile, we discuss their pros and cons.
- *Describe metrics to evaluate blockchain-based covert channels:* We outline essential metrics of evaluating blockchain-based covert channels and discuss their evaluation methods, which sets a foundation for unified and quantitative evaluation.
- *Discuss privacy aspects of the covert communication in blockchain:* We discuss privacy issues that may arise from the covert communication in blockchain.
- *Present challenges and future directions:* We present several existing challenges and future directions of blockchain-based covert communication, which guide the direction for follow-up researchers.

The illustration of the scope and taxonomy of this survey is shown in Fig. 1. Section II introduces covert communication, covert channels, and blockchain. Section III presents the model, key technologies, features, and applications of blockchain-based covert communication. Section IV - VI reviews existing schemes and details the key technologies on channel building. Section VII analyzes evaluation metrics and methods of blockchain-based covert channels. Section VIII discusses the privacy aspects of the covert communication in the blockchain. Section IX presents challenges and future directions. Finally, Section X concludes the survey.

II. Background

We provide background on covert communication, covert channels, and blockchain, and describe how the blockchain addresses drawbacks of traditional covert communication in this section.

A. Covert Communication and Covert Channels

We introduce covert communication and covert channel from definition, model, classification, information embedding, and obfuscation.

1) Definition: Covert communication [48] means that two parties transmit data via steganography and hide the existence of the transmission from a third party. It is a special kind of communication in which the communication behaviors are undetectable [3]. To implement covert communication, Lampson *et al.* [4] first gave the formal definition of covert channels in 1973. In their original definition, covert channels allow a process to transmit information to other unauthorized processes within a host [49]. The rapid development of computer networks facilitates covert channels [50] because of the diversity of network protocols and the large volume of network traffic. Currently, covert channels are the media to implement covert communication [51], and they are deployed [52], [53], [54] and evaluated [55], [56], [57] based on traditional centralized networks such as the Internet.

2) Model: Simmons *et al.* [58] first proposed the prisoner problem in 1983. Fig. 2 depicts the prisoner problem. Alice and Bob, who are prisoners, are separated in prison, and they need to negotiate a breakout way. The only communication way is to send messages through Wendy, who guards them. All the communications with Wendy seem like innocuous conversations, while the breakout way is hidden through the conversations and keeps secret against Wendy.

The prisoner problem has also become the model of covert communication. Alice and Bob are two parties that need to transmit covert information through the intermediate routing Wendy. The two parties hide covert information through public channels to cover their communication behaviors. Everyone,

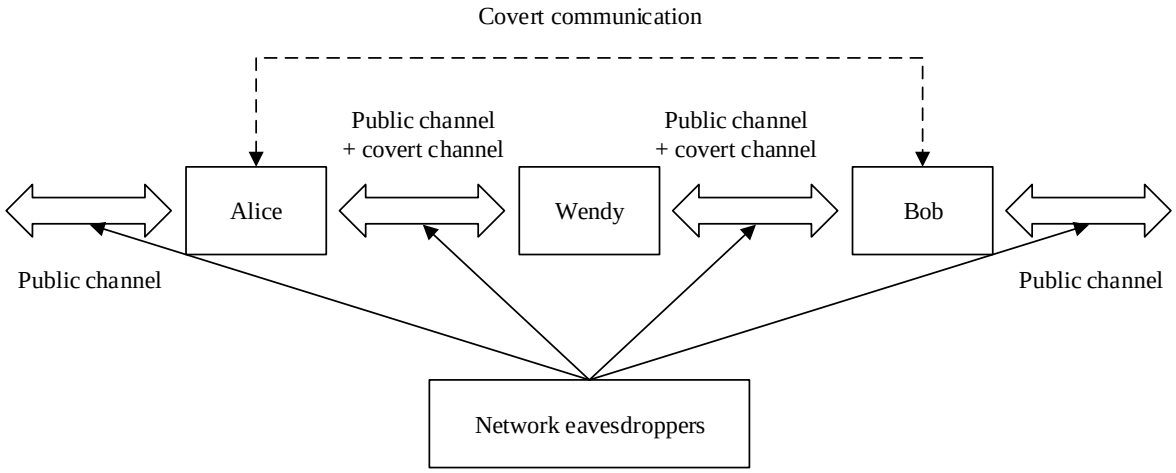


Fig. 2: The model of covert communication.

including network eavesdroppers, has access to the open channel. Network eavesdroppers aim to detect and destroy covert communication, further tracing Alice and Bob's identities. Communication channels constructed by Alice and Bob are called covert channels.

3) Classification: According to types of carriers, covert channels are mostly divided into the covert storage channel (CSC) and the covert timing channel (CTC) [59], [34]. CSC generally employs PDU's storage fields such as redundant and reserved fields of HTTP packages to embed covert information [60] and has a high capacity due to the diversity of protocols, while the regularity of these fields leads to low concealment. CTC utilizes time-correlated information between PDUs as carriers [61], e.g., IPD between TCP packages [62]. It needs more than one PDU to encode covert information, which leads to lower bandwidth and higher concealment. The network delay may add noise to channels so that the receiver may receive inaccurate information. There also exist other classifications. According to the seven layers of the open system interconnect (OSI) model, covert channels can be categorized into seven types, i.e., covert channels in physical, data-link, network, transport, session, presentation, and application layers [63]. According to the noise setting, covert channels can be divided into noisy channels and noiseless channels [64].

4) Information Embedding: Covert information is mostly embedded in storage fields of network protocols. These storage fields exist in each layer of the OSI model. For example, *the data link layer's* medium access control [65] and 802.11 protocols [66] can transmit covert information. *In the network layer*, the identification field [67], don't fragment flag [68], timestamp header extension bits [63], time to live field [69] of the IPv4 protocol, protocol header fields [70], and hop limit field [71] of the IPv6 protocol can be manipulated to embed covert information. The sequence number, confirmation sequence number [72], control flags [73], [74], [75], and timestamp [76] of the TCP protocol *in the transport layer* are employed to achieve covert channels. Storage fields *in the application layer's* protocols, such as HTTP [77], [78], [79], [80] and FTP [81], [82], are also used to embed information.

CTCs primarily embed covert information into IPD, such as the interval of network data packets and the number of network data packets per unit time. Various network protocols such as TCP [83], [84], RTP [85], and HTTP [86] can be utilized to implement IPD-based covert channels. In addition, covert information can also be embedded into the data rate of communication data streams [63] and sequences of PDUs [87], [88], [89], [90].

Traditional covert channels are easily eliminated or restricted. The elimination directly blocks or normalizes fields of protocols and destroys covert channels [91], [92]. For example, setting the identification field to 0 and rewriting the time to live field eliminate corresponding covert channels. The restriction modifies IPD and significantly increases the bit error rate of CTCs [93]. Hence the restriction reduces the channel capacity to a minimal extent, furthermore restricts corresponding covert channels ultimately [94], [95], [96]. Traditional covert channels are unreliable, while decentralization of blockchain ensures the reliability of blockchain-based covert channels.

5) Obfuscation: The purpose of obfuscation is to enhance concealment of covert channels. Storage fields containing covert information are more sensitive to content-based detection methods [97], [98], [99]. CTCs have higher concealment [100], partly because their development undergoes more obfuscation. We take their development as an example to briefly introduce the idea of obfuscation.

Cabuk *et al.* [101] proposed the first IP-based covert timing channel (IPCTC). They encoded the number of data packets within a period of time as 0 and 1. For example, the number of data packages exceeds ten transmits bit 1, and that the number of data packages is less than ten transmits bit 0. However, this kind of encoding intervenes in the IPD distribution between data packets and results in a notable difference between the IPD distribution of normal and encoded data packets. Then Cabuk *et al.* [64] proposed a time-replay covert timing channel. It records the IPD of normal data packets and divides them into two groups. One group represents bit 0, and the other

This article has been accepted for publication in IEEE Communications Surveys & Tutorials. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/COMST.2022.3204281

represents bit 1. The time-replay covert timing channel has stronger concealment than IPCTC since its IPD distribution conforms to that of normal data packets.

Gianvechio *et al.* [102] also proposed the model-based covert timing channel to improve IPCTC. They constructed exponential, logarithmic, and other distribution functions to fit the normal IPD distribution. Then they selected the model with the smallest standard deviation as their final model. However, the model can only be established according to the law of normal IPD in a certain period. The distribution fitted by their model only conforms to the normal distribution in the current period and cannot synchronize the normal distribution in real-time. Given the new problem, Zhang *et al.* [103] and Xin *et al.* [104] proposed covert channels that spontaneously adapt to normal network traffic.

The improvement process of the above IP-based covert channels essentially fits the IPD distribution between special and normal data packets, i.e., obfuscation. However, the difficulty of obtaining whole traffic in a traditional network and the huge computing resources required to analyze traffic make it challenging to obfuscate traditional covert channels. The openness and transparency of blockchain allow everyone to obtain the whole ledger easily. It is hence easier to obfuscate special transactions. In addition, multiple cryptography fields do not need to be obfuscated due to the randomness of cryptography primitives.

B. Blockchain

We present the working principle of the blockchain and its properties, and show advantages of covert communication in blockchain.

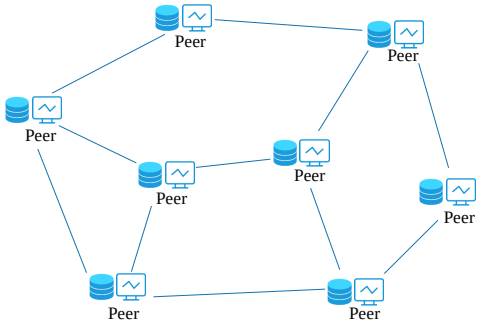


Fig. 3: The network structure of blockchain.

1) Working Principles: Cryptographic primitives, consensus mechanism, transaction broadcast, etc. are the basic components of a blockchain [105]. We introduce how the blockchain works by these basic components.

- *Cryptographic primitive.* Hash functions and digital signature algorithms are two main cryptographic primitives used in blockchain. The *hash functions* are also known as computable one-way compression functions with one input and one output. Computable means that it is efficient to compute the output given the input; One-way means it is difficult to compute the input given the output; Compression means that the length of the output string is fixed and usually less than that of the input string. Hash functions are usually used in blockchain to calculate the unique identification of a long text. The *digital signature algorithm* outputs a proof of the authenticity of a message. Such proof guarantees that the message was indeed sent by a certain entity and has not been tampered with. Digital signature algorithms are commonly utilized for transaction creation in blockchains to indicate that the transaction creator holds the inputs used to create the transaction and to guarantee that the transaction has not been tampered with.
- *Consensus mechanism.* In the blockchain, the consensus mechanism can be understood as a protocol that makes the system state-maintained locally by each node in the distributed network consistent. We briefly present the Proof of Work (PoW) algorithm of Bitcoin as an example. All nodes in a distributed network compete to generate new blocks and update the system state for a reward. These nodes decide who generates the new block by solving a computational problem. The problem is difficult to solve while easy to verify. The first solver publishes the solution and other nodes verify the solution. If the solution is valid, they agree that the node is the first to generate a new block and update the system state. Other consensus mechanisms like Proof of Stake can also promise the consistency of a distributed system through different principles.
- *Broadcasting.* The broadcasting refers to the way that data such as a new block and a new transaction is synchronized to the entire blockchain network. When a node wants to send data, it will broadcast the data to its neighbor nodes. Then the neighbor node will verify the validity of the received data, and thereafter transfer it to the child neighbor nodes except the data source node. Finally, the data is forwarded to each node.

2) Properties: Blockchain can be seen as a ledger with flooding propagation, decentralization, and anonymity [106]. These properties contribute to addressing the drawbacks of existing traditional covert channels and building state-of-the-art covert channels.

- *Flooding propagation.* Transactions are broadcasted through flooding propagation in the blockchain network and transmitted to all nodes. When a node receives transactions, it stores these transactions locally and forwards them to other nodes. All nodes perform the above same behavior and store the same transactions. In a nutshell, transactions are transmitted non-directionally during flooding propagation. Every node could be the sender or the receiver of covert communication. The sender does not communicate with the sender via a fixed routing.
- *Decentralization.* No centralized hardware or organization exists in the blockchain network, and all nodes maintain the blockchain system together. Fig. 3 shows the network structure of blockchain, i.e., a peer-to-peer network. Each node maintains a global ledger himself/herself. No one can control the whole system as a central authority.

TABLE I: Comparison of existing covert methods and blockchain-based covert communication

Existing drawbacks	Advantages of blockchain-based covert communication
Simply detected	Avoid direct connection and communication between the sender and the receiver
Easily restricted and eliminated	Guarantee the immutability of the blockchain ledger and transaction data
Simply traceable	Replace the users real identity with random anonymous addresses

Decentralization provides reliability for blockchain-based covert channels. Transactions cannot be modified, and blockchain-based covert channels cannot be destroyed.

- *Anonymity.* Anonymity means that all users anonymously participate in the blockchain network. Specifically, each user appears with addresses that are indeed a random string. It is hard to link users to their addresses. The blockchain ledger only stores addresses and enables users' anonymity. Anonymity provides anti-traceability and conceals communication parties' identities.

Components such as cryptographic primitives and consensus mechanisms provide blockchains with decentralization and anonymity that address drawbacks of traditional covert channels. Table I shows drawbacks of existing covet methods and advantages of blockchain-based covert communication. The existing traditional covert methods is simply detected, restricted, eliminated, and traceable. Achieving covert communication in blockchain avoids the direct connection between the sender and the receiver, and thus provides stronger concealment against detection techniques. Blockchain guarantees the reliability of covert channels through the immutability of the distributed ledger. Blockchain replaces the users real identity with random anonymous addresses and thus makes covert channels difficult to trace.

III. Blockchain-based Covert Communication

We present the system model, key technologies, features, and applications of blockchain-based covert communication.

A. System Model and Key Technologies

We provide an overview of blockchain covert communication and summarize three key techniques for constructing covert channels.

Blockchain-based covert communication refers to launching covert communication via blockchain networks. Fig. 4 describes the system model of blockchain-based covert communication. The model comprises three types of entities: the sender, the receiver, and the blockchain network. The sender and the receiver conduct a secret negotiation before communications start.

The sender holds covert information and addresses to create covert transactions, i.e., special transactions. In the following, we call them covert transactions for better understanding. The sender first embeds covert information into transactions. Second, the sender generates labels that can only be identified by the receiver according to the pre-negotiated transaction filtering method. The above transactions are then obfuscated to obtain covert transactions. Finally, the sender sends these covert transactions to the blockchain network.

The receiver is the recipient of covert information. The only operation the receiver executes, related to the Internet, is requesting data from the blockchain. Then the receiver filters out covert transactions according to the pre-negotiated transaction filtering method and performs de-obfuscation. Afterward, the receiver extracts covert information from covert transactions via corresponding covert information extraction algorithms. The request is identical to other blockchain users, and the filtering and extraction are performed locally. Moreover, addresses involved in covert transactions have nothing to do with the receiver. The identical behaviors, local operations, and irrelevant addresses promise the anonymity of the receiver.

The blockchain network is the public channel to implement blockchain-based covert communication. Covert transactions are hidden in the blockchain network. The sender broadcasts covert transactions to the blockchain network, and the receiver requests the blockchain ledger. The blockchain network is the medium to transmit covert information.

We summarize three key technologies in the blockchain-based covert communication model: information embedding, transaction filtering, and transaction obfuscation. Section IV - VI comprehensively details these technologies by analyzing existing blockchain-based covert channels.

B. Features

Due to inherent properties of blockchain, blockchain-based covert communication shows advances in strong concealment, high reliability, and anti-traceability. These features are described in this section as below:

Strong Concealment is mainly attributed to flooding propagation, the large number of active users, massive transactions, and complicated cryptographic primitives. A large number of active users and their same behaviors as those of communication parties make behaviors of the sender and receiver have no distinctive features. Massive normal transactions make it possible to hide covert transactions in normal transactions. Taking Bitcoin and Ethereum as examples, Bitcoin has about 300,000 daily transactions and Ethereum has more than 1,700,000 daily transactions [107]. Complicated cryptographic primitives provide storage fields [108], [109] that can cover original features of covert information. Cryptographic storage fields containing covert information cannot be distinguished from those of normal transactions due to the indistinguishability of cryptographic primitives.

High Reliability is guaranteed by decentralized blockchain and the immutability of transactions. Decentralized blockchain makes it difficult to jeopardize the blockchain system by

This article has been accepted for publication in IEEE Communications Surveys & Tutorials. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/COMST.2022.3204281

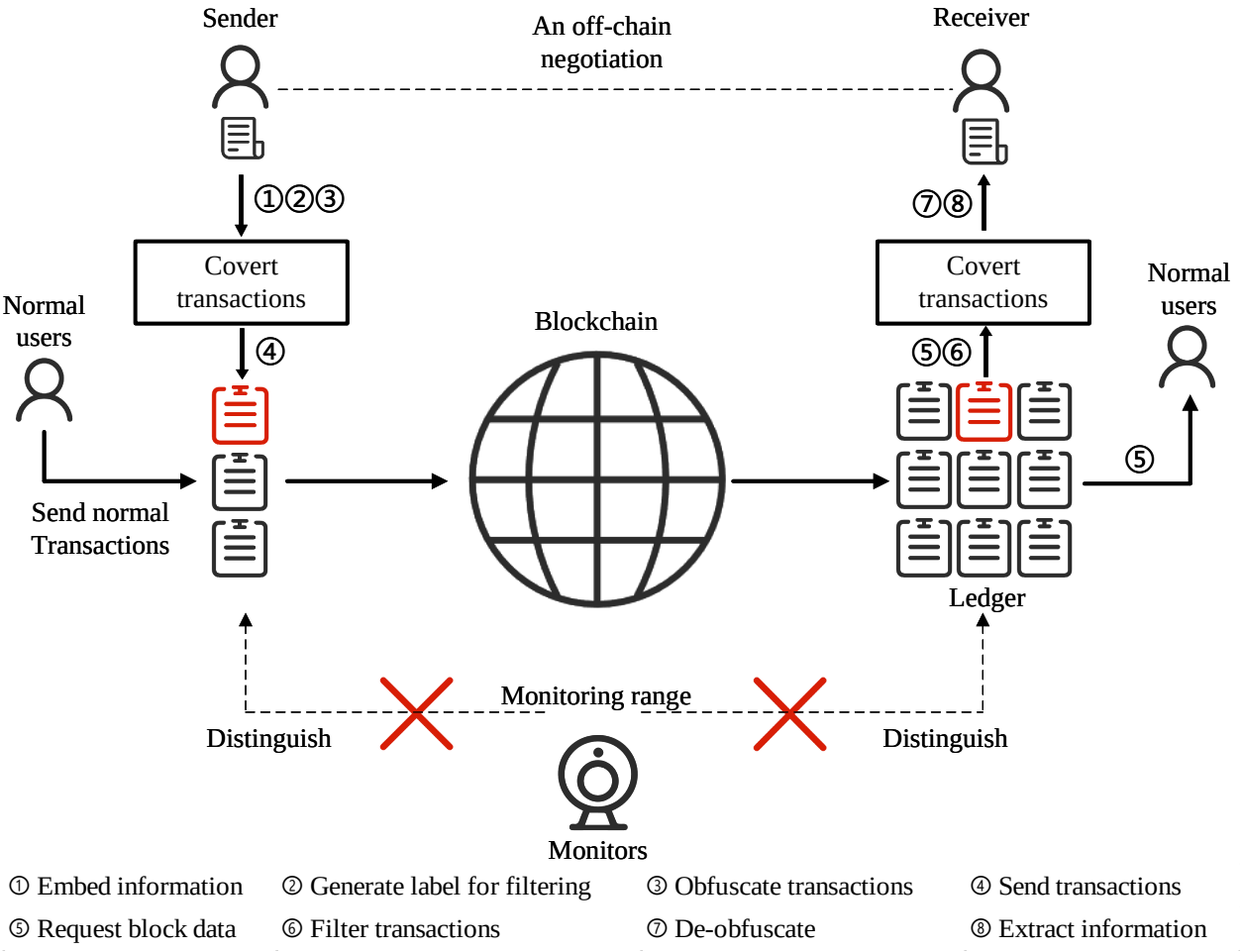


Fig. 4: The system model of blockchain-based covert communication.

traditionally compromising the central server in the centralized network. Due to flooding propagation, covert transactions can bypass compromised nodes and be broadcast to the whole blockchain network. Techniques for eliminating or restricting traditional covert channels cannot work well in blockchain-based covert channels, so channels' reliability can be significantly improved.

Anti-traceability is enabled due to the anonymity property of blockchain. The anonymity of the receiver is completely ensured since covert transactions do not contain information about the receiver. The only information related to the sender's identity is its addresses. Even though covert channels are detected, adversaries only discover anonymous addresses rather than communication parties' identities.

C. Applications

We investigate the existing applications based on blockchain covert communication, which can be divided into sensitive data transmission, botnet building, de-anonymization attack implementation, digital evidence preservation, identity verification, and covert broadcast channels.

Blockchain-based covert communication pays more attention to hiding the existence of transmission behaviors than transmission contents itself. Discoveries of communication behaviors often lead to exposures of identities, further exposing the entire organization [110], [111]. Communication parties, such as the government and enterprises, whose identities and groups are sensitive, thus need to transmit data via blockchain-based covert communication.

Blockchain-based covert communication provides a natural implementation environment for the botnet's command and control (C&C) system [112], [113], [114], [115], [116], and it also brings optimization for the maintenance cost of the C&C system [117], botmaster's anonymity [118], and botnet's reliability [119], [120]. Sincerely, there have been botnets whose C&C system is constructed by blockchain, and part of botnets have been detected [121], [122].

De-anonymization attacks could be implemented through blockchain-based covert communication. Blockchain networks such as Zcash [123] and Monero [124] are designed to protect the privacy of blockchain users further. Similar to the fact that covert communication could de-anonymize anonymous networks [125], [126], [127], Biryukov *et al.* [128] found that covert channels in Zcash might destroy the privacy protection function provided by Zcash, thus tracing Zcash users' identities.

Blockchain-based covert channels contribute to digital ev-

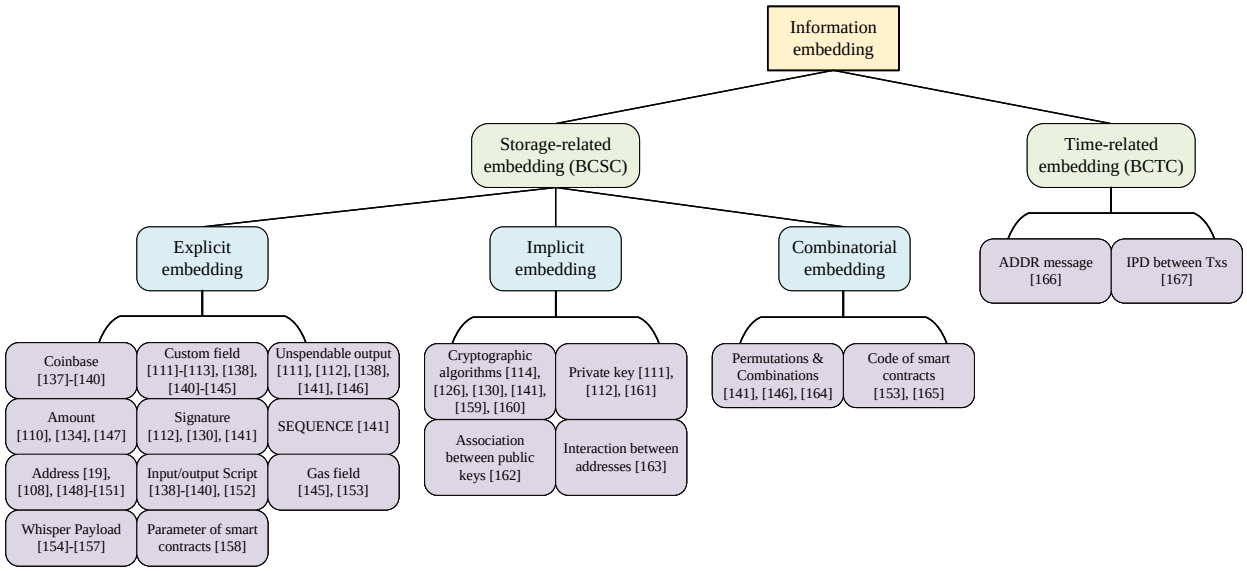


Fig. 5: Classifications of information embedding.

idence preservation. Benefitted from blockchain's anonymity and transparency features, Wang *et al.* [129] proposed a digital evidence preservation system with anonymity, transparency, scalability, and lightweight.

Blockchain-based covert communication can also be used for identity verification. Chen *et al.* [130] applied blockchain-based covert channels to remote authentication between IoT servers and IoT devices to avoid centralized attacks and the high cost of maintaining certificate authority. Coincidentally, Gai *et al.* [131] combined covert channels and edge computing to realize identity authorization in smart grids, which made their proposed model be able to resist all kinds of attacks.

Openness and transparency make it possible for attackers to report malicious and illegal activities through blockchain-based covert communication. Alsalami *et al.* [132] pointed out that cryptographic features of public blockchains enabled wallet version attacks aiming at mining users' privacy. In addition, they also argued that the peer-to-peer (P2P) network architecture of blockchain contributed to realizing covert broadcast channels [133].

On the one hand, blockchain-based covert communication facilitates the development of a wide range of applications. On the other hand, blockchain-based covert communication can also lead to attacks that disrupt the original functionality of the blockchain system.

D. Summary and Insights

In this section, we give an overview of blockchain-based covert communication. We also model blockchain-based covert communication and discuss its features. Additionally, based on these features, we present applications from blockchain-based covert communication. From the discussion above, blockchain-based covert communication can provide privacy preservation for digital evidence and identity verification. On the other hand, new privacy issues like de-anonymization may occur, which will be discussed in Section VIII.

IV. Information Embedding

We discuss the existing information embedding techniques in this section. Specifically, we classify these techniques into storage-related embedding and time-related embedding. The storage-correlated embedding employs transaction storage fields to embed covert information. Further, storage-related embeddings are divided into explicit embedding, implicit embedding, and combinatorial embedding according to whether the sender directly encodes covert information as a transaction field and the embedded field's property (private or public), and channels that implement information embedding with storage-related embedding methods are called blockchain-based covert storage channels (BCSC); The time-correlated embedding utilizes time-correlated carriers to embed covert information, and corresponding channels are called blockchain-based covert timing channels (BCTC). Fig. 5 shows the overview of our classification with references. Given the significant differences between these two information embedding methods, we introduce these two types of information embedding methods successively below.

A. Information Embedding in BCSC

Existing information embedding methods in BCSC are shown in this section. In fact, Bitcoin is able to store data [134], [135] immutably, transparently, and permanently, boosting many blockchain-based applications [136], [137], [138]. These applications employ the data storage function of blockchain to record and display data, which has a particular gap between covert channels. However, the idea of data storage inspires BCSCs. We investigate storage fields for embedding covert information according to our classification:

1) Explicit embedding: *Explicit embedding* refers to the embedding approach that directly encodes covert information as the public transaction fields.

- **The coinbase transaction of Bitcoin.** A coinbase transaction is created when a new block is generated. The miner, who commits a new block, can embed data in this coinbase transaction. Satoshi Nakamoto left a message in the Bitcoin genesis block [139]. BCSCs are built by directly embedding covert information into coinbase transactions [140], [141], [142]. However, only the miner is able to build such BCSCs, which makes it challenging to achieve this embedding method in practice. Even if such BCSCs can be implemented, such coinbase transactions containing covert information are easily exposed.
- **Custom storage fields.** Existing blockchain systems generally have a custom storage field, such as OP_RETURN of Bitcoin and INPUT of Ethereum. Transaction creators entirely control these fields. Covert information can be embedded into OP_RETURN [143], [144], [115], [114], [140], [113], [142]. However, OP_RETURN has obvious formats and characteristics [145], which leads to a significant difference between normal OP_RETURN and special OP_RETURN. Yin *et al.* [146] considered cleaning the trace of covert transactions to overcome this drawback. The sender first sends covert transactions with low fees. When the receiver identifies covert transactions in the transaction pool, the sender creates other normal transactions with high fees whose inputs are the same as covert transactions. The covert transactions will not be on-chain in the end. Basuki *et al.* [147] utilized INPUT to embed covert information. INPUT containing information also has a difference from normal INPUT. Using custom storage fields to embed covert information lacks concealment due to differences between normal fields and special fields. When embedding with custom fields, the sender needs to study how to make special fields indistinguishable from normal fields.
- **Unspendable outputs.** Unspendable outputs are utilized to embed covert information [143], [114], [140], [113], [148]. Fig. 6 shows the embedding principle. The sender directly encodes covert information as transaction outputs or addresses. However, these generated outputs cannot be spent since corresponding private keys are needed to spend them. Above private keys cannot be inferred through generated addresses or public keys since the elliptic curve public key generation algorithm and hash functions are irreversible.

  Unspendable outputs have high concealment since addresses, public keys, and encrypted covert information are all random and indistinguishable from each other. The shortcoming is that unspendable outputs incur more cost. The amount the sender transfers to special addresses should be as small as possible When embedding information in this way.
- **Amount fields.** Covert information can be embedded into transaction's amount fields of Bitcoin lightning network [112], Bitcoin network [136], and Ethereum network [149]. The cost of cryptocurrencies and the precision of amount fields limit channel capacity. The sender also needs to simulate the special amount based on the distribution of the normal amount.

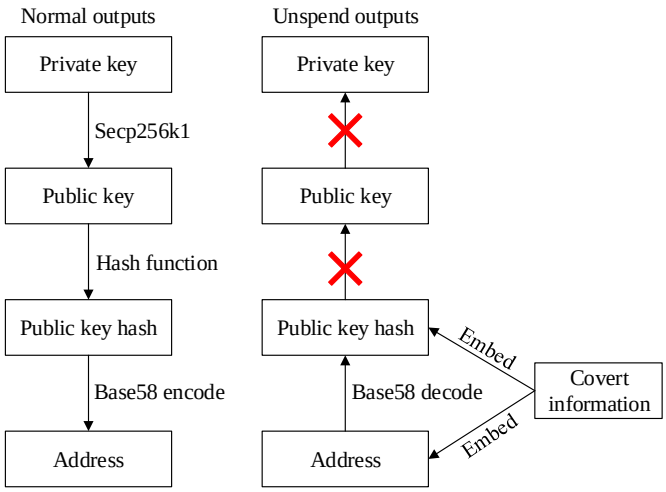


Fig. 6: Unspendable output.

- **Signatures.** The signatures of transactions are obtained through calculation and cannot be completely controlled. However, the sender can control a few bits of signatures by brute force computing. The sender keeps modifying the unsigned transaction or the signing process's random parameters until a few bits of signatures meet expected demands [132], [143], [114]. For example, Fig. 7 demonstrates the process of controlling the last hexadecimal digit of signature to be 6. Channels embedding covert information into signatures have high concealment and low channel capacity, and the embedding process needs huge computing resources. Fields with more controllable bits are better for information embedding.

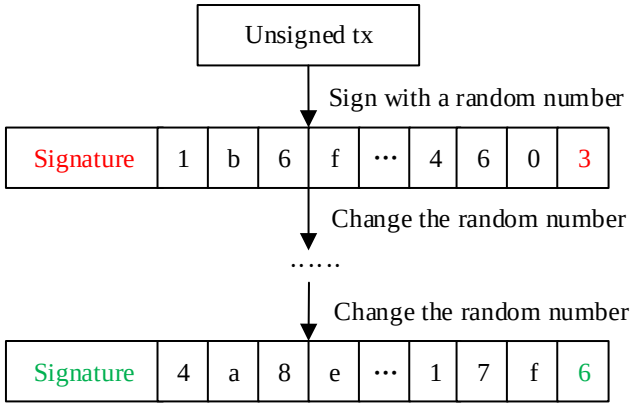


Fig. 7: Control the signature by brute force computing.

- **The SEQUENCE field of Bitcoin.** SEQUENCE of Bitcoin transactions controls whether a transaction needs a delay to take effect. Transaction creators completely control this field so that the sender can embed covert information into it [143]. However, the value of most trans-

actions' SEQUENCE is $0XFFFFFFFF$, which means that the transaction takes effect immediately. Transactions whose SEQUENCE is not $0XFFFFFFFF$ are distinct from other transactions. Embedding covert information into SEQUENCE lacks concealment.

- **Addessses.** Unlike unspendable outputs, addresses here are controllable. The sender possesses private keys of the above addresses and controls a few bits of these addresses via brute force computing. Partala *et al.* [19] proposed a blockchain covert channel (BLOCCE) for the first time and proved its security and feasibility. BLOCCE employs the least significant bit (LSB) of addresses to store 1-bit covert information. Later work [150], [151], [110], [152] improved BLOCCE for increasing channel capacity. Most studies [150], [151], [110] increased channel capacity by increasing the number of bits carried in each address. In addition, Wang *et al.* [152] used Monero's ring signature mechanism to increase the number of addresses that carried covert information to increase channel capacity. Qin *et al.* [153] calculated the hash of output addresses and divided the result into two sets: one represented bit 1, and the other represented bit 0. Covert channels that embed covert information into addresses have low capacity and high concealment. The sender should look for ways to embed more information in the address without increasing the amount of computation.
- **Input/output scripts of Bitcoin.** Covert information can be embedded into input/output scripts of Bitcoin transactions [140], [154], [141], [142]. Transactions whose scripts can be embedded into covert information are divided into two types: standard transactions and non-standard transactions. Standard transactions mainly refer to M-N multi-signature transactions. These transactions have an M-N multi-signature address. As long as N signatures are offered, the unspent transaction output (UTXO) of this M-N address can be spent. Other $M-N$ input/output scripts can be used for information embedding. Non-standard transactions mean transactions whose lock scripts or unlock scripts are customized. Transaction creators add redundant data as covert information into lock scripts and generate redundant scripts. Although these redundant scripts are non-standard, the miner thinks them valid and involves them in new blocks. Such an embedding method generally achieves high-capacity channels. Nevertheless, the number of multi-signature transactions and non-standard transactions is small. Whether this embedding method has high concealment or not lacks theoretical and experimental support.
- **Gas-correlated fileds of Ethereum.** Basuki *et al.* [147] and Liu *et al.* [155] embedded information into gas-correlated fields of Ethereum, such as gasLimit and gasPrice. These two fields affect the fees of transactions. Accordingly, the capacity of covert channels based on the gas-correlated fields is also low.
- **The Whisper protocol of Ethereum.** Whisper is an identity-based communication system based on the Ethereum P2P framework [156]. Whisper allows users to broadcast custom encrypted data in the payload field. The receiver filters target envelopes through the topic field. Fig. 8 demonstrates a diagram of the data unit in Whisper, i.e., the envelope.

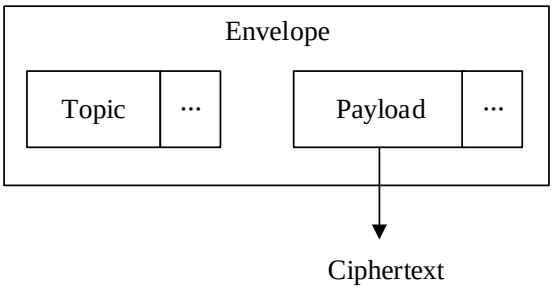


Fig. 8: Diagram of the envelope.

Covert channels can be implemented based on Whisper [157], [158], [159]. All data broadcast around Whisper is encrypted provides concealment. The capacity of the encrypted data is related to a process of mining. The harder the mining process, the higher the capacity.

- **Parameters of smart contracts.** In addition to transferring digital currencies, transactions can also execute a script that achieves specific functions. The smart contract of Ethereum is designed to implement such functions. Transactions containing smart contracts are called contract transactions. Calling contract Transactions generally requires parameters. Fig. 9 displays that these parameters can be used to transmit covert information. Zhang *et al.* [160] handled the parameters of the voting contract and the bidding contract to embed covert information and realized covert channels based on smart contracts.

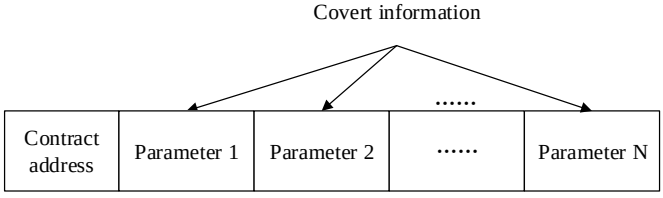


Fig. 9: Information embedding in parameters of Ethereum smart contracts.

2) Implicit embedding: Implicit embedding is the embedding approach that embeds covert information into intermediate variables in the transaction generation process, and these variables are usually private and invisible.

- **Random parameters of cryptography algorithms.** The digital signature algorithm guarantees the integrity of transactions. Random parameters are needed to perform the signing process. The sender can control these random parameters to transmit covert information [143], [132], [161], [162], [116]. In addition, blockchain systems are integrated with cryptography algorithms such as zero-knowledge proofs and ring signatures to enhance anonymity. These cryptography algorithms also provide

multiple random parameters to embed covert information. Biryukov *et al.* [128] constructed two kinds of covert channels based on Zcash. They embedded covert information into random parameters of the zero-knowledge proofs generation process.

Channels employing random parameters to embed covert information are with high concealment. Random parameters are not directly displayed in transactions. Only signatures and proofs are visible. Cryptography primitives provide randomness for these signatures and proofs. The randomness makes them indistinguishable from other normal signatures or proofs. However, the receiver has difficulty recovering these random parameters. Specifically, take the elliptic curve digital signature algorithm (ECDSA) for an example:

$$ks = h + rd \tag{1}$$

Equation (1) shows the core step of ECDSA. r and s are two parts of the signature. h denotes the hash of an unsigned transaction. r, s, and h can be obtained from the public blockchain ledger. k is the random parameter used for signing, and d represents the private key.

In addition to public r, s, and h, the receiver needs d in advance to extract random parameter k. Thus the receiver can spend the sender's digital currencies or forge covert transactions. The above consequence is fatal in botnets since honeypots in botnets may destroy the whole botnets by forging covert transactions. The sender needs to design a method to recover random parameters without sharing the private key.

- **Private keys.** A private key is essentially a random number that can be manipulated to embed covert information. The private key is not displayed in public transaction fields, either. Only the address, the hash of the private key, is visible, and the address is also with randomness provided by cryptography primitives. The sender needs to share the random parameter with the receiver that helps to recover covert information. Tian *et al.* [163] and Ali *et al.* [113] related two covert transactions with the same random parameter to recover the private key. However, this kind of sharing is unreasonable since the same random parameter makes it easy for everyone to identify these two transactions [114]. Covert channels embedding covert information into private keys have high concealment while the sender needs to share random parameters reasonably and avoid reusing random parameters.
- **The association between public keys.** Cao *et al.* [164] embedded covert information into the association between public keys. The association is essentially a hard mathematical problem. The sender shares all possible solutions with the receiver in advance so that only the receiver can extract covert information by traversing these solutions. Their proposed channel's capacity can be dynamically adjusted and is inversely proportional to needed computing resources of recovering covert information. As the channel capacity increases, the amount of computation required is also increased. Information embedding methods should avoid this waste of computing resources.
- **The interaction between addresses.** Xiang *et al.* [165] proposed a novel scheme to embed covert information. The sender represents bit 1 and 0 through whether there is a transaction between two addresses, i.e., the interaction between addresses. Fig. 10 shows the information embedding scheme. The interaction is represented by a matrix I. Element $I_{01} = 1$ means that address a_0 sends a transaction to address a_1. Similarly, $I_{02} = 0$ means that there is no transaction between a_0 and a_2. The embedding process consists of multiple rounds. Only the upper-right or lower-left part of the diagonal of the matrix works in each round. The example in Fig. 10 transmits "101" when the upper-right part of the diagonal of the matrix works. Each transaction can only carry 1-bit information through this embedding method. The sender should increase the amount of information each transaction can embed.

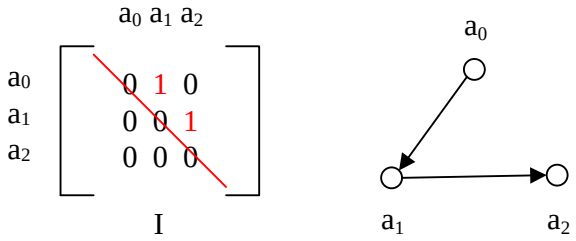


Fig. 10: Embedding in interaction between addresses.

3) Combinatorial embedding: *Combinatorial embedding* is the embedding approach that indirectly utilizes the storage properties of transaction fields like the permutation of outputs to embed covert information.

- **Permutations and combinations of addresses, amounts, and transactions.** Torki *et al.* [148] suggested applying the permutation of transaction input/output addresses to information embedding. Fionov *et al.* [143] indicated that the permutation and combination of amounts could also embed covert information. Fig. 11 shows an example that embeds covert information into the permutation of output addresses. The sender and the receiver negotiate encoding and decoding modes, i.e., the permutation-code table. The sender transmits code d by permutating output addresses into $(1, 2, 0)$. The permutation means that the address with index 1 is the largest, the address with index 2 is the second largest, and the address with index 0 is the third largest.

  The capacity of the above channel depends on the number of inputs/output addresses of a transaction. Transactions with large input/output addresses account for a tiny percentage in reality. This embedding method cannot take both channel capacity and concealment into consideration. Xu *et al.* [166] permutated transactions of a block to overcome the incompatibility problem. However, their scheme can only be implemented by a miner who commits new blocks.

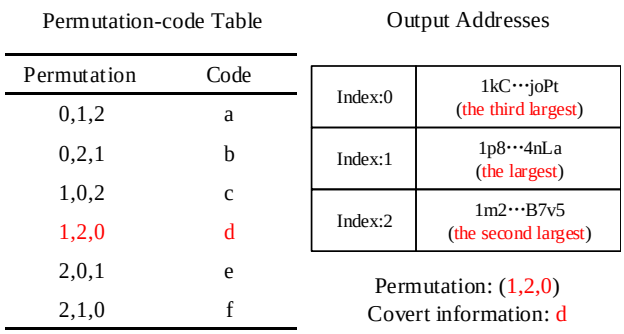


Fig. 11: Permutations of output addresses.

- **Codes of Ethereum smart contract.** Li *et al.* [167] and Liu *et al.* [155] embedded covert information into codes of smart contracts. Among them, Li *et al.* [167] deployed smart contracts to control the botnet's C&C system. In comparison, Liu *et al.* [155] introduced rearrangement and replacement algorithms based on the source code and the byte code to hide covert information. Rearrangement refers to reordering objects such as functions and variables. Replacement means replacing sentences that play the same role, such as $a++$ and $a = a + 1$. These rearrangements and replacements change the hash of smart contracts. Different hashes represent different covert information.

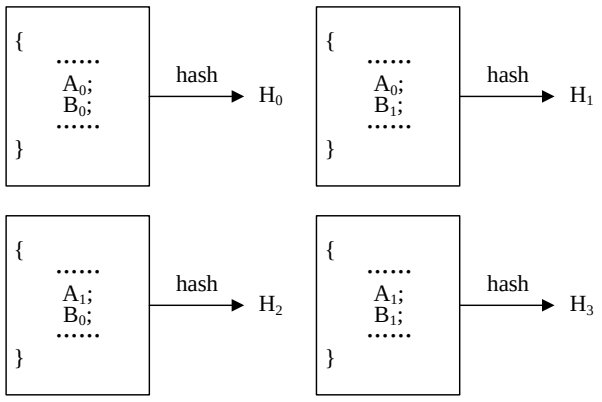


Fig. 12: Embedding of replacement in smart contract source code.

Fig. 12 depicts the principle of replacements. Suppose that $A = (A_0, A_1)$, $B = (B_0, B_1)$. A and B represent pre-negotiated sets whose elements can be used for replacement. Each element in A has the same effect, and each element in B has the same effect. A and B appear twice totally in the source code. Accordingly, The replacement algorithm can generate four codes performing the same function. Suppose their hashes are H_0, H_1, H_2, and H_3. The sender and the receiver negotiate that the order of these hash values determines the encoding result. For example, the hash value H_0 is the second smallest, and the code with (A_0, B_0), whose hash is H_0, transmits bits 01.

Lessons learned: Information embedding in BCSC is essentially a process of finding controllable transaction storage fields. These fields can be an inherent part of the transaction (OP_RETURN, INPUT, and amount, etc.) or self-constructed by the sender (the association between public keys and the interaction between addresses, etc.). When embedding information in these fields, the sender needs to guarantee the indistinguishability between embedded fields and normal fields. If using private keys and random parameters of cryptographic algorithms for embedding, the sender also needs to pay special attention to avoiding sharing private keys with the receiver and reusing random parameters. Sharing private keys with the receiver makes the receiver able to forge covert transactions and steal the sender's cryptocurrencies. Reusing random parameters makes the first half part of two signatures identical, thus exposing the covert transactions.

B. Information Embedding in BCTC

We list the existing information embedding methods in BCTC in this section. Only two proposals build CTCs via blockchain networks. Li *et al.* [168] theoretically proposed a BCTC based on transactions' on-chain time interval for the first time. However, the sender cannot control transactions' on-chain time, thus unfeasibility of their scheme. Lv *et al.* [169] constructed a BCTC based on the IP address (ADDR) message. The ADDR message exchanges neighboring nodes' information and consists of IP addresses. Both the sender and the receiver maintain a Bitcoin node. These two nodes connect to each other and communicate via ADDR messages. Covert information is hidden in these ADDR messages. Nevertheless, nodes push ADDR messages once a day limits channel capacity. In addition, the sender and the receiver directly connect to each other, so the channel may be easier to detect compared to BCSCs.

Lessons learned: It is difficult to implement BCTC by treating transactions as PDUs of traditional CTCs since the transaction confirmation time cannot be guaranteed. Therefore, the communication traffic among blockchain nodes is more suitable for building BCTC. However, existing BCTC still require a direct connection between communicating parties, which might incur channel exposure. We will discuss building BCTC without such a connection in future directions.

C. Summary and Insights

In this section, we discuss the first key technology for building blockchain-based covert channels, i.e., information embedding. Its essence is to find carriers to embed information. Blockchain transactions and traffic can provide such carriers for embedding covert information. The amount of covert information that these carriers can carry varies

from each other. Table II summarizes and compares existing information embedding methods of BCSC from applicable blockchain, controllable capacity, and characteristics (capacity and concealment). We do not include embedding methods of BCTC in the comparison since there are few samples. As can be seen, the explicit embedding methods generally have a higher capacity, and the implicit embedding methods possess higher concealment. In addition, several embedding methods may raise security issues like private key leakage. The above findings motivate us to build high-performance channels, which is discussed in Section IX.

V. Transaction Filtering

Blockchain-based covert communication has strong concealment partly because the sender does not directly communicate with the receiver. However, this also leads to the difficulty of identifying covert transactions. Filtering mechanisms need to be designed so that the receiver can quickly identify covert transactions from loads of blockchain transactions. Existing feasible BCTC [169] does not need filtering mechanisms since it allows a direct connection between the sender and the receiver. We only investigate the filtering mechanisms of BCSCs and classify them into fixed labels, decryption labels, and dynamic labels. In this section, the existing filtering methods are introduced one by one according to this classification.

A. Fixed Labels

We first describe the existing fixed label methods. Filtering by fixed labels means that communication parties always use a transaction field with a fixed value to filter transactions.

Fixed addresses are utilized as labels to filter covert transactions [147], [19], [115], [112], [114], [150], [151], [113]. The sender and the receiver agree on a fixed address. The sender uses this address as the input/output address to create covert transactions, and the receiver only filters transactions related to this address. As the number of communications increases, more and more transactions related to the same address make this address notable. Besides, the sender's identity is easier to trace as the number of times the address is used increases. Although this transaction filtering method is efficient, it leads to low concealment.

B. Decryption Labels

Filtering by encryption labels means that the receiver performs the information recovery process for all transactions, decrypts the obtained data, and filters the transactions based on whether a meaningful plaintext can be successfully decrypted.

Recabarren *et al.* [154] filtered covert transactions through encryption and decryption. The sender and the receiver negotiate a key in advance to encrypt and decrypt covert information. The sender encrypts covert information before embedding it. The receiver extracts the ciphertext and then decrypts it with the negotiated key. If the receiver gets a meaningful plaintext, covert transactions are successfully filtered. This method is more concealed than fixed labels while having lower efficiency. The receiver needs to traverse all on-chain transactions, and encryption and decryption processes require many computing resources. In addition, if the capacity of each covert transaction is less than the shortest length of ciphertext, the filtering method is useless since the receiver cannot decrypt a segment of ciphertext.

C. Dynamic Labels

We show existing dynamic label methods in this section. Filtering by dynamic labels means communication parties utilize transaction fields with dynamic values to filter transactions.

Tian *et al.* [163] proposed a dynamic label generation algorithm based on the probability distribution of OP_RETURN. The sender utilizes the domain generation algorithm (DGA) to generate a certain sequence of dynamic domains and obtain dynamic labels. The sender and the receiver negotiate a seed to generate the same domains, further acquiring the same labels. They embedded labels into OP_RETURN and proved that these labels were resistant to K-S and conditional entropy tests. Their dynamic labels achieve a compromise of concealment and efficiency. However, the receiver needs to know the same OP_RETURN distribution as when the sender generates labels. It is tough to synchronize the OP_RETURN distribution when creating and filtering covert transactions in practice.

Gao *et al.* [144] utilized signatures modified by Kleptography as labels. Kleptography is a cryptography trapdoor technology that generates indistinguishable signatures via adding a controllable backdoor to signature algorithms, such as ECDSA and RSA. Fig. 13 demonstrates how Kleptography recognizes modified signatures and restores the private key. The sender uses Kleptography to reform ECDSA and creates modified signatures with a pre-negotiated Kleptography key. The receiver filters covert transactions via recovering private keys corresponding to input addresses of covert transactions. Both modified and normal signatures are random and indistinguishable, so the labels have high concealment. The receiver needs to traverse all on-chain transactions, which leads to low filtering efficiency. In addition, two transactions with the same input address are required to restore covert information, which may be an obvious feature of covert transactions.

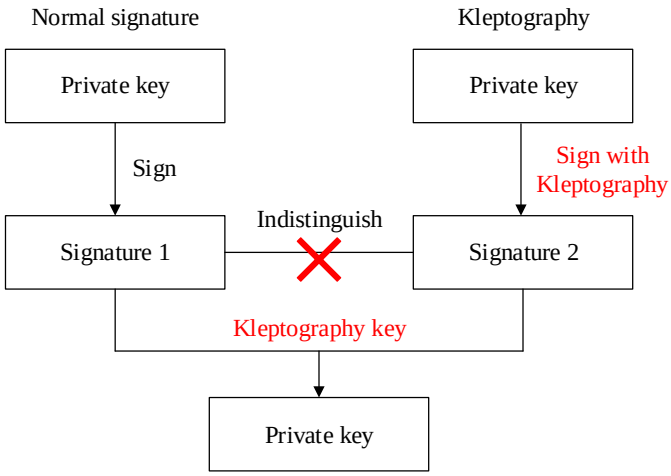


Fig. 13: Kleptography.

TABLE II: Summary of existing feasible information embedding schemes in BCSCs

Literature	Embedded filed	Applicable blockchain	Classification	Controllable capacity/bits	Characteristics
[19]	LSB of Addresses	Bitcoin	Explicit embedding	1	Low capacity; high concealment
[143], [144], [115], etc.	OP_RETURN	Bitcoin	Explicit embedding	640	High capacity; low concealment
[147]	Input, Address, and gasLimit	Ethereum	Explicit embedding	29	Low capacity
[114]	Unspendable outputs	Bitcoin	Explicit embedding	160	High cost
[116]	Random factors of ECDSA	Blockchain containing ECDSA	Implicit embedding	256	High concealment; share private keys; avoid reusing random number
[161]	Random factors of ring sigunature	Monero	Implicit embedding	2560	High capacity; high concealment; only apply to Monero
[128]	Random factors of zk-SNARK	Zcash	Implicit embedding	~70	High concealment; the capacity is limited by computing resources
[149]	Amount	Ethereum	Explicit embedding	~28	High concealment; low capacity
[143]	Sigunature/public key	All blockchain	Explicit embedding	~16	High concealment; the capacity is limited by computing resources
[163]	Private key	All blockchain	Implicit embedding	256	high capacity; need to design how to recovert the private key
[142]	1-3 muti-signature transactions	Bitcoin	Explicit embedding	1040	High capacity; low concealment due to the small number of multi-signature transactions
[143]	Permutations of inputs/outputs	Bitcoin	Combinatorial embedding	$\log_2 m!$, where m is the number of inputs/outputs	Can't take both the capacity and concealment into consideration
[164]	Connection between public keys	All blockchain	Implicit embedding	Variable	Capacity is related to the computing resources consumed
[160]	Parameter of smart contract	Ethereum	Explicit embedding	Variable	Capacity is related to the specific contract
[157], [158], [159]	Payload field of whisper	Ethereum	Explicit embedding	Variable	Capacity is related to the computing resources consumed
[165]	Address interaction relationship	Bitcoin	Implicit embedding	1	Low capacity; large interactions between a fixed address set
	Value	Bitcoin	Explicit embedding	~23	Low capacity; high concealment
[155]	Source code of Ethereun smart contract	Ethereum	Explicit embedding	202	High capacity; low concealment
	Byte code of Ethereun smart contract	Ethereum	Explicit embedding	387.36	
	Transaction fields including fromAddress, toAddress, gasLimit, and gasPrice, etc.	Ethereum	Explicit embedding	48-64	Low capacity; high concealment

Si *et al.* [170] filtered covert transactions according to dynamic addresses. The sender and the receiver negotiate the generation rule of private keys and calculate identical addresses. These addresses are indistinguishable from other normal addresses due to the randomness of the elliptic curve algorithm and hash functions. This filtering method is efficient since the application programming interface (API) often provides services for inquiries by addresses. The receiver filters covert transactions through addresses without traversing all transactions. However, the sender and the receiver are required to share all private keys in advance, so threats such as stealing cryptocurrencies and forging covert transactions exist.

Lan *et al.* [161] and Guo *et al.* [110] implemented BCSC in Monero that utilizes one-time addresses to filter covert transactions. The one-time address mechanism is designed to protect the receiver's identity. A Monero user possesses a public key and a private key. The public key is used for other users to generate one-time Monero addresses belonging to the Monero user. The private key verifies whether a Monero address belongs to the Monero user. Monero addresses generated by the different private keys are indistinguishable. Fig. 14 shows the one-time address mechanism of Monero. However, the

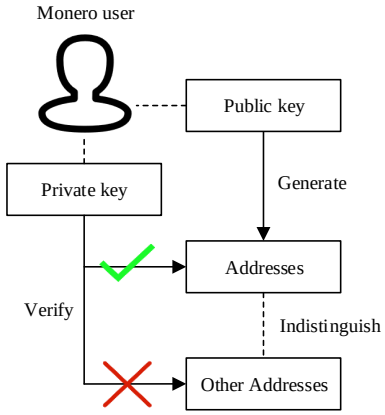


Fig. 14: The one-time address mechanism of Monero.

filtering method has flaws. The output of covert transactions is the receiver's one-time Monero addresses during the covert communication process. The sender transfers a certain number of Moneroes to the receiver for each communication. The sender cannot hold his/her currencies, and the receiver can forge covert transactions.

Torki *et al.* [148] applied hierarchical deterministic wallets (HDW) to transaction filtering. The sender and the receiver negotiate the private key generation rule, i.e., HDW, for constructing labels. Digital currency wallets generally manage users' private keys through HDW.

$$x_i = y + H(k||i) \tag{2}$$

Equation (2) describes HDW, where y and k are secrets negotiated by the sender and the receiver, H denotes a secure hash function, and x_i is the $i-th$ private key. The receiver calculates private keys and addresses according to HDW to filter covert transactions. Threats of sharing private keys also exist in this filtering scheme.

Biryukov *et al.* [128] pointed out the essence of transaction filtering, i.e., the sender and the receiver negotiate secrets of generating labels that both parties can only identify. They then proposed a transaction filtering scheme employing zkSNARK proofs as labels. The labels also have high concealment due to cryptography primitives. Nevertheless, generating labels needs huge computing resources because of the high computational complexity of zero-knowledge proof. The proposal combines transaction filtering and information embedding processes, causing a high error-filtering rate or a low channel capacity.

D. Summary and Insights

The broadcasting of blockchain makes two parties to communicate without a direct connection. However, this needs a filtering mechanism to help the receiver to identify the covert information-carrying carrier. We outline existing transaction filtering methods involved in BCSCs, and Table III illustrates their principles and characteristics. Existing transaction filtering schemes cannot simultaneously meet concealment, efficiency, security, and feasibility, which motivates us to design dynamic label technologies with the following properties:

- The sender avoids sharing private keys with the receiver.
- Nobody can identify labels other than the receiver.
- The receiver filters covert transactions efficiently.

VI. Transaction Obfucation

We briefly explain why transactions need to be obfuscated and describe the research state of transaction confusion. Transaction fields used for embedding and filtering have unique statistical characteristics since they are formed by covert information encoding. These characteristics may be different from those of normal transaction fields. Transactions are stored on-chain permanently and openly, so monitors have sufficient ability and adequate time to analyze all transactions. Numerous organizations have developed mature tools and technologies for blockchain data statistics [171], [172] and steganographic analysis [173], [174]. If adversaries detect covert transactions through the above techniques, they can trace communication parties' identities via address tracing technologies [175], [176]. Then adversaries destroy the anti-traceability of blockchain-based covert communication. It is consequently imperative to conceal covert transactions and increase the difficulty for adversaries to distinguish covert transactions and normal transactions. Specifically, the sender obtains characteristics of the normal transactions and obfuscates covert transactions to increase the concealment of covert transactions. Characteristics that the sender needs to obfuscate can be divided into transaction-level and currency-level characteristics. We then present transaction-level obfuscation and currency-level obfuscation in this section.

A. Transaction-level Obfuscation

We first introduce obfuscation of transaction-level characteristics. Transaction-level characteristics include character characteristics and quantitative characteristics. Existing obfuscation mainly focuses on character characteristics of transaction fields, such as amount and OP_RETURN.

TABLE III: Summary of transaction filtering in BCSCs

Literature	Classification	Principle	Filtering field	Characteristics
[147], [19], [115] etc.	Fixed labels	Fixed output/input addresses of covert transactions	Addresses	Efficient; low concealment
[161], [110]	Fixed labels	One-time addresses of Monero	Addresses	High concealment; unreasonable currency transferring; threats of forging covert transactions
[154]	Decryption labels	Decrypt successfully or not	The same as embedding fields	Inefficient; the channel capacity must be longer than the length of the cipher text
[163]	Dynamic labels	Domain Generation Algorithm (DGA)	The OP_RETURN	High concealment; difficult to implement in practice
[144]	Dynamic labels	Kleptography	Signatures	High concealment; inefficient
[170]	Dynamic labels	HMAC	Addresses	Efficient; threats of sharing private keys
[148]	Dynamic labels	HDW	Addresses	Efficient; threats of sharing private keys
[128]	Dynamic labels	Brute force computing	Proofs of Zk-SNARK	High concealment; high computational complexity

1) Character characteristics: Partala *et al.* [19] obfuscated amount fields based on historical transaction amounts of a fixed address. Xiang *et al.* [165] obfuscated the length of transactions' amounts according to its distribution. They analyzed normal transactions' amounts and simulated the length of transactions' amounts to avoid amounts with abnormal length. Liu *et al.* [149] obfuscated Ethereum transactions' amounts. They counted orders of magnitudes of normal transactions' amounts to determine that of covert transactions' amounts. In addition, they calculated the information entropy of normal transactions' amounts and utilized several bits of covert transactions' amounts to adapt these amounts to normal transactions' amounts. Although these obfuscations reduce channel capacity, it is imperative to conceal covert transactions through obfuscation.

Tian *et al.* [163] obfuscated OP_RETURN based on the character distribution of normal OP_RETURN. The sender ensures that the probability of each character in special OP_RETURN is equal to that of normal OP_RETURN. However, their scheme actually cannot hide special OP_RETURN since they only adapt the probability of a single character. Consecutive characters of obfuscated OP_RETURN may expose covert channels since many protocols use OP_RETURN to extend applications beyond transfer [145]. These protocols lead to a certain format of OP_RETURN. Obfuscating OP_RETURN only according to the probability of a single character cannot guarantee the format of OP_RETURN. For example, Omni [177] is one of the most widely used protocols in Bitcoin, and most OP_RETURN are utilized to execute Omni. Fig. 15 displays the structure of Omni. The protocol consists of a 4-byte Omni protocol flag, 2-byte transaction version, 2-byte transaction type, 4-byte currency identifier, and the 8-byte transfer amount. However, extra bytes except the transfer amount have a fixed format. Characters of OP_RETURN obfuscated by the single character's distribution [163] cannot match consecutive characters of OP_RETURN performing Omni.

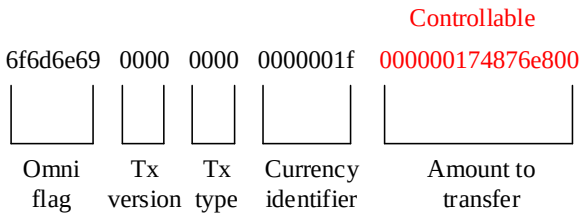


Fig. 15: The structure of the Omni protocol.

Liu *et al.* [155] innovatively applied deep neural networks to transaction obfuscation. The sender designs an RNN generator that outputs covert transactions and inputs normal transactions and covert information. For covert transactions generated by the RNN, a CNN estimates whether they are normal transactions or not via a loss function. The original CNN is trained in advance. The whole model, including RNN and CNN, is then trained to obfuscate covert transactions. The CNN optimizes the loss function, and RNN generates covert transactions with more concealment. RNN and CNN make up a generative adversarial network. RNN embeds covert information into Ethereum transaction fields, including fromAddress and toAddress. However, the sender may not possess corresponding private keys of the generated fromAddress or toAddress, so covert transactions may be invalid or cause unspendable outputs.

2) Quantitative characteristics: Obfuscation of quantitative characteristics includes the number of transaction input and output addresses. The capacity of combinatorial channels [143] that encode data through permutations and combinations of transaction input or output addresses is limited by transactions' quantitative characteristics. More than half of transactions have

only one input and two outputs [178]. Transactions with thousands of input or output addresses are notable. High-capacity combinatorial channels often generate transactions with many addresses and hence lack concealment in reality. Whether covert information is embedded in input or output addresses or not, it is necessary to obfuscate the number of input and output addresses and other quantitative characteristics to enhance concealment.

B. Currency-level Obfuscation

We show the currency-level obfuscation in this section. Currency-level obfuscation refers to obfuscating the currency relation between transactions. Multiple transactions are usually required to transmit covert information once. Cryptocurrencies of these transactions belong to the sender. The adversary can infer links between the input of a transaction and the output of the last transaction from cryptocurrencies. For example, the UTXO model of Bitcoin links a sequence of transactions. The adversary can trace the original covert transaction when detecting covert transactions. Then cryptocurrencies of the original covert transaction are traced. These cryptocurrencies are usually obtained from digital currency exchanges, which provide cash-currency exchanging and record exchangers' identities. Therefore, tracing the original covert transaction may expose covert channels and the sender's identity. However, little research considers obfuscating currency-level characteristics of transactions. Fortunately, several coin mixing technologies have been developed to blur associations between transactions [179], [180]. These coin mixing technologies, initially used for privacy protection, are not very popular due to neglect of privacy. Therefore, mixed transactions with many input and output addresses are rare, and obvious features of mixed transactions make them unsuitable for covert communication. Currency-level obfuscation still needs development, and we will introduce it in Section IX.

C. Summary and Insights

Transaction obfuscation is one of the key technologies for achieving blockchain-based covert communication. In this section, we explain the necessity of obfuscation for concealing covert transactions, although transaction obfuscation decreases channel capacity. Table IV summarizes the existing research. Cryptographically related fields do not require obfuscation. Almost half of the proposals that should have done obfuscation do not perform the step, and most of the obfuscation-performed schemes employ statistical methods for obfuscation. It is worth mentioning that a proposal [155] applies neural networks to obfuscation. This provides novel ideas for obfuscation, and we will discuss them in Section IX.

VII. Metrics to Evaluate Blockchain-based Covert Channels

In this section, we present the evaluation of blockchain-based covert channels. We summarize evaluation metrics of evaluating blockchain-based covert channels, including concealment, bandwidth, transmission delay, robustness, and transmission cost. Each metric and its evaluation methods are detailed in the following context. The "one-time covert transmission behavior" involved later may consist of a certain number of covert transactions. The specific number depends on the number of transactions required to transmit sufficiently small amounts of covert information. For example, DLchain [163] employs two covert transactions to achieve one-time covert transmission behavior.

A. Concealment

Concealment refers to the ability of covert channels not to be detected. It is also called antidetection or security. Existing work mainly evaluates concealment through indistinguishability between covert transactions and normal transactions. Most studies only qualitatively analyze the concealment of their proposed schemes [166], [115], [161], [151], [170]. Other methods of evaluating concealment can be divided into theoretical proof [19], [164] and experimental tests [163], [168], [149].

1) Theoretical Proof: Theoretical proof methods prove indistinguishability between covert and normal transactions based on provable security. Partala *et al.* [19] conducted a rigorous theoretical proof of their proposed channel's concealment. They utilized LSB of addresses to transmit covert information and obfuscated the amount field according to historical amounts. They then proved that addresses and amount fields of covert transactions are indistinguishable from those of normal transactions. Cao *et al.* [164] embedded covert information into the association between public keys. They theoretically proved that addresses containing covert information are indistinguishable from normal addresses.

2) Experimental tests: The experimental test method quantificationally compares embedding or filtering fields between covert and normal transactions via the K-S test, KLD, and so on. Li *et al.* [168] utilized the entropy rate (ER) to evaluate concealment, and Tian *et al.* [163] also demonstrated concealment of their proposed dynamic labels through ER. The closer ER of covert transactions is to that of normal transactions, the more concealment the channel has. Equation (3) gives the calculation of ER:

$$ER = \min_{i=1,2,\ldots,m} CCE(X_i|X_{i-1}) \tag{3}$$

where CCE refers to the corrected conditional entropy in Equation (4):

$$CCE(X_m|X_{m-1}) = H(X_m|X_{m-1}) + p(X_m)H(X_1) \tag{4}$$

Among them, $H(X_m|X_{m-1})$ is conditional entropy. $p(X_m)$ denotes the probability of X_m. $H(X_1)$ is the entropy of X_1. X_i represents the feature of embedding or filtering fields such as the length of OP_RETURN and magnitudes of the amount field.

Liu *et al.* [149] evaluated the concealment of their embedding scheme through information entropy, and Equation (5) shows the formula:

$$H(X) = -\Sigma_{i=1}^{n} P(X_i) \log_2 P(X_i) \tag{5}$$

TABLE IV: Summary of transaction obfuscation

Field		Need to be obfuscated	Literature	Embedding or filtering	Obfuscated or not	Description
Input addresses		Do not need	[148], [170]	Filtering	Do not need	Calculated though cryptography primitives; are with natural randomness and indistinguishability, do not need to be obfuscated.
Output scripts/addresses		Do not need	[148], [132], [19], [114], [113], [151], [150], [110], [152]	Embedding	Do not need	
		Do not need	[161]	Filtering	Do not need	
Signatures		Do not need	[114]	Embedding & filtering	Do not need	
Zero-knowledge proofs		Do not need	[128]		Do not need	
Association between public keys		Do not need	[164]	Embedding & filtering	Do not need	
Smart contracts' parameters		Need	[160]	Embedding	×	Obfuscation depends on the contracts.
Custom storage fields	OP_RETURN	Need	[114], [113], [144], [115], [150]	Embedding	×	Need to obfuscate from the length, the format, and the character distribution, etc.
		Need	[163]	Filtering	✓	
	Input	Need	[147]	Embedding	×	
Random parameters of cryptography algorithms	Single-signature transactions	Do not need	[114], [113], [116], [163], [161], [132]	Embedding	Do not need	Do not need to obfuscate; need to avoid the reuse of random parameters.
	Multi-signature transactions	Need	[181], [154]	Embedding	✓	The proportion of multi-signature transactions which affects concealment needs to be considered.
The gasLimit		Need	[147]	Embedding	×	Need to obfuscate from the distribution.
Permutations of addresses/amounts		Need	[181]	Embedding	×	Need to obfuscate from the number of input/output addresses and amounts.
The amount		Need	[112]	Embedding	×	Need to obfuscate from the length and precision.
		Need	[149]	Embedding	✓	
		Need	[165]	Embedding	✓	
The payload of Whisper		Need	[158]	Embedding	✓	The length of the payload needs to be obfuscated; the content is all encrypted and does not need to be obfuscated.
The SEQUENCE of Bitcoin		Need	[181]	Embedding	×	Need to obfuscate from the distribution.
Transaction fields in Ethereum		Need	[155]	Embedding	✓	Applied deep neural networks to obfuscation, while covert transactions may be invalid or cause unspendable outputs.

They embedded covert information into the amount field and calculated the information entropy $H(X)$ of binary transaction amounts. X represents the transaction amount, and $P(X_i)$ denotes the probability of the binary bit X_i (0 or 1) appearing in X. Concealment of the embedding scheme is assessed by comparing normal transaction amounts' information entropy to that of covert transaction amounts.

In addition, Tian *et al.* [163] and Liu *et al.* [149] applied the K-S test to evaluate the amount field. The K-S test evaluates whether two distributions, i.e.,$F(x)$ and $G(x)$, can be considered the same distribution. The greater the result D_{ks} is, the lower probability that these two distributions belong to the same distribution. Equation (6) displays the formula:

$$D_{ks} = \sup_x |F(x) - G(x)| \tag{6}$$

Liu *et al.* [155] designed a machine learning steganalyzer, which converts smart contracts, i.e., carriers carrying covert information, into a binary sequence and counts frequencies of 0 and 1. The closer frequencies are to frequencies of normal smart contracts, the more concealed the channel is.

B. Bandwidth

Bandwidth refers to the maximum amount of covert information transmitted through covert channels within limited carriers. The bandwidth of traditional covert channels is generally represented by the amount of covert information transmitted per unit time (for CTC) [182] or per unit data packet (for CSC) [183]. Similarly, the maximum number of bits that each covert transaction could carry is used to describe the bandwidth of BCSC [147], [143], [144], [161], [149], [164], [128]. Equation (7) displays the formula of BCSC's bandwidth:

$$cap = \frac{bits}{Num_{tx}} \tag{7}$$

where $bits$ is the number of bits that can be reliably transmitted within one-time covert transmission behavior, and Num_{tx} represents the number of transactions required to realize this one-time covert transmission behavior. cap denotes the bandwidth of covert channels, and its unit is bit/transaction. This evaluation method provides a uniform standard to compare different BCSCs. Therefore, this way of evaluating bandwidth is used the most.

Xu *et al.* [166] denotes the bandwidth of their proposed BCSC by the maximum number of bits of covert information that each block can carry because this channel utilizes a whole block of transactions to embed covert information. The unit is bit/block, and the formula is:

$$cap = \frac{bits}{Num_{block}} \tag{8}$$

where Num_{block} is the number of blocks required to realize one-time covert transmission behavior.

Li *et al.* [168] used the maximum bits of covert information that can be correctly transmitted per unit time to represent the bandwidth of BCTC. The unit is bit/s, and the formula is:

$$cap = \frac{bits}{t} \tag{9}$$

where t represents the unit of time.

C. Transmission Delay

Transmission delay refers to the time from sending covert transactions to extracting covert information within one-time covert transmission behavior [112], [114], [170], [116], [146]. Transmission delay generally depends on:

- **Receiver's access method.** If the receiver accesses the blockchain network via a full node, he/she can request transaction data through the full nodes RPC function. Since the transaction data is stored locally when the receiver maintains a full node, the response of RPC is fast with almost no delay. If the receiver accesses the blockchain network via external APIs, there will be a 2-3 second delay when the external API responds, which is set by the API provider to resist DoS attacks. Moreover, API providers often set limits on the number of API accesses, e.g., 1,000 accesses per IP per hour, to resist DoS attacks. If the receiver requests data too frequently, it needs to wait for a while to gain access to the API.
- **The scheme itself.** For blockchain-based covert communication schemes, the confirmed order of covert transactions may determine whether covert information can be correctly extracted. In this case, the next covert transaction should be sent after confirming the previous covert transaction. If one-time covert transmission behavior requires multiple transactions, transmission delay needs to take transaction confirmation time into account.

The calculation formula of transmission delay is:

$$delay = T_2 - T_1 \tag{10}$$

where T_1 represents the time when the sender sends the first covert transaction within one-time transmission behavior, and T_2 is the time when the receiver extracts covert information.

D. Robustness

Robustness is the ability of the receiver to extract covert information correctly. BCSC cannot be eliminated due to the immutability of the blockchain. BCSC is generally considered robust and reliable [163], [116]. Nevertheless, network delay may affect BCTC. The sender cannot promise that the receiver extracts covert information anticipatively. Robustness is generally presented by the bit error rate (BER).

$$BER = \frac{S_{error}}{S_{all}} \tag{11}$$

Equation (11) describes BER, where S_{error} denotes the number of error symbols transmitted in one-time covert transmission behavior, and S_{all} is the total number of symbols in this one-time covert transmission behavior.

Equation (12) shows the other definition of BER, i.e., the ratio of error bits:

$$BER = \frac{\sum_{i=1}^{k} e(m(i), m'(i))}{S_{all}} \tag{12}$$

where $m(i)$ and $m'(i)$ represent the original and decoded i-th bit of covert information. $e(*,*)$ is a binary function, and if the two parameters of the function are identical, it returns 1; otherwise, it returns 0. The higher BER, the weaker robustness of covert channels.

E. Transmission Cost

Creating transactions in blockchain requires fees, and implementing blockchain-based covert channels needs a cost. The transmission cost refers to a cost required to transmit a certain amount of data. The cost of transmitting a unit byte is usually defined as the transmission cost of covert channels [144], [115], [114], [150], [113], [154]. Equation (13) shows the formula:

$$cost = \frac{\sum fee}{bytes} \tag{13}$$

where $bytes$ denotes the number of bytes of covert information that covert channels can reliably transmit in one-time covert transmission behavior. fee refers to the total fees required to realize this transmission. The unit of transmission cost is related to dependent cryptocurrency $coin$, i.e., $coin$/Byte.

In addition, studies applying blockchain-based covert communication to botnets' C&C systems define transmission cost as the cost of transmitting an instruction [112], [135] or maintaining a botnet [157].

F. Summary and Insights

We present evaluation metrics and methods to evaluate blockchain-based covert channels in this section. We also compare existing blockchain-based covert channels from blockchain platforms, embedding, filtering, concealment, bandwidth, and feasibility in Table V. We only summarize channels whose evaluation metrics contain concealment. Covert channels whose evaluation metrics do not contain concealment are meaningless since concealment is imperative to covert channels. Concealment evaluation in the table depends on the evaluation methods of the proposed scheme. Schemes with formal proof are considered high concealment; schemes

TABLE V: Summary of Blockchain-based Covert Channels

Literature	Blockchain	Embedding	Filtering	Concealment	Bandwidth(bits/tx)	Feasibility
[132]	Bytecoin	Random number in ring signature	\	high	128	✓
	Monero	Random number in ring signature	\	high	256	✓
[163]	Bitcoin	Random number in ECDSA	OP_RETURN	medium	128	×
[147]	Ethereum	Input, gas limit, address	Fixed input address	low	29	✓
[19]	All blockchains	Output address	Fixed input address	high	1	✓
[143]	Bitcoin	Permutations of addresses/amounts	\	low	$\log_2 m!$, where m is the number of inputs/outputs	✓
		Amount distribution	\	low	$log_2\binom{V-V_0+m-1}{m-1}$, where m is the number of output, V represents the total amount, and V_0 denotes the minimum output amount amount	✓
[144]	Bitcoin	\	Signature	low	\	✓
[166]	All blockchains	Transaction ordering in block	\	low	$log_2 \frac{m!}{\prod_{e \in U_c} f(e)!}$, where m is the number of transactions in the block, U_c denotes all transactions with different addresses, and $f(e)$ represents the number of times that the transaction e repeats	×
[115]	Bitcoin testnet	OP_RETURN	Fixed input address	low	640	✓
[112]	Bitcoin lighting network	Amount	Fixed input address	low	\	✓
[113], [114]	Bitcoin	OP_RETURN	Fixed input address	low	640	✓
		Unspendable outputs	Fixed input address	low	160	✓
		Private key	Reusing random number in signature	low	256	✓
		Signature scripts	Fixed input address	low	~14	✓

Continued on next page

Continued from last page

Literature	Blockchain	Embedding	Filtering	Concealment	Bandwidth(bits/tx)	Feasibility
[161]	Monero	Random number in ring signature	One-time address in Monero	high	$2560m$, where m is the number of inputs	✓
[150]	Bitcoin	Output address	Fixed input address	low	~24	✓
[151]	All blockchains	Output address	Fixed input address	low	\	✓
[168]	All blockchains	IPD of transactions	\	medium	\	×
[170]	All blockchains	\	Dynamic address	low	\	✓
[116]	All blockchains	Random number in signing	Fixed address	low	256	✓
[154]	Bitcoin	Multi-signature transaction input script	\	low	448	✓
[149]	Ethereum	Amount	\	medium	~28	✓
[164]	All blockchains	Association between public keys	Association between public keys	high	Variable	✓
[128]	Zcash	Zero-knowledge proof	Zero-knowledge proof	high	~70	✓
[158], [159], [157]	Whisper in Ethereum	Payload	Topic	low	$256n$, where n is an integer and affects the computational cost required to send covert information	✓
[110]	Monero	Input address	One-time address in Monero	high	11	✓
[148]	Bitcoin	Unspendable outputs	Input address of HDW	low	160	✓
		Output address and its ordering	Output address of HDW	low	$mn + logn!$, where n is the number of outputs, and m represents the number of bits of covert information carried in each address	✓
[152]	Bitcoin	Output address	Off-chain	low	34	✓
[160]	Ethereum	Smart contracts parameters	Fixed contract address	low	Variable	✓
[146]	Bitcoin	OP_RETURN, transaction fileds	Off-chain with the third party	low	640	✓

Continued on next page

Continued from last page

Literature	Blockchain	Embedding	Filtering	Concealment	Bandwidth(bits/tx)	Feasibility
[153]	All blockchains	Output address	Fixed input address	low	1	✓
[165]	Bitcoin	Amount and address interaction relationship	Fixed input address set	medium	14-24	✓
[169]	Bitcoin	ADDR message	Point-to-point connection	low	240bits / day	✓
[155]	Ethereum	Source code of Ethereun smart contract	Fixed input address	medium	202	✓
	Ethereum	Byte code of Ethereun smart contract	Fixed input address	medium	387.36	✓
	Ethereum	Transaction fields	Fixed input address	medium	48-64	✓

with experiment evaluation are considered medium concealment; schemes with only qualitative description are considered low concealment. If covert channels are implemented in Bitcoin, their bandwidths are calculated according to covert transactions with one input and two outputs by default. This kind of transaction has the largest proportion of Bitcoin transactions [178]. Feasibility is determined by whether covert channels could be achieved in practice or not. If only a block proposer can implement the scheme, it is considered unfeasible. We only evaluate concealment, bandwidth, and feasibility of covert channels since most papers only evaluate these three metrics. The table shows that: 1) there is no quantitative evaluation method for concealment; 2) the feasibility of schemes needs to be considered; 3) there is no unified assessment dimension for blockchain-based covert communication schemes. We discuss standardization and quantification of evaluation metrics in Section IX.

VIII. Privacy Aspects of Covert Communication in Blockchain

Covert channels are originally used by attackers for system data leakage. Does the covert communication in blockchain raise privacy concerns like data leakage? There possibly exist privacy leakage of blockchain users and communicating parties. We discuss these privacy issues and countermeasures in this section.

A. Privacy Issues for Blockchain Users

Private keys and anonymous identities are important privacy for blockchain users. However, the covert communication in blockchains may make them be leaked.

1) Leakage of private keys: Gadekallu *et al.* [46] argued that blockchain-based services like DID may face private key leakage attacks due to the existence of covert channels, and Davenport *et al.* [184] deployed such a covert channel to attack the DID wallet system and expose users' private keys. Alsalami *et al.* [132] also pointed out that the covert communication in blockchain may expose the private keys of blockchain users, especially the covert communication in close-source hardware wallets. The above private key exposure attacks are indeed possible, and we discuss the implementation principles of these attacks with examples.

Example 1: The DLchain scheme [163] implements twice ECDSA with the same random parameter for information embedding. This idea can be utilized to construct a backdoor. We denote the twice ECDSAs as $s_1 = k^{-1}(h_1 + r_1 d)$ and $s_2 = k^{-1}(h_2 + r_2 d)$. Then the private key d can be calculated by $d = \frac{h_1 s_1^{-1} s_2 - h_2}{r_2 - r_1 s_1^{-1} s_2}$. Suppose a backdoored Bitcoin hardware wallet performs ECDSA in the following way. Whenever a new ECDSA is executed, the wallet will first determine whether the private key used for this ECDSA has been recorded in its memory. If the private key has not been recorded in memory, the wallet performs a normal ECDSA and records the private key and the corresponding random parameter in memory; otherwise, the wallet performs this ECDSA with the last recorded random parameter. Once a private key is used twice, the wallet generates two transactions with the same r since r is only determined by k ($r = kG$). The attacker observes transactions with the same r and recovers corresponding private keys. In this way, the attacker can build a backdoored hardware wallet that utilizes DLchain for private key leakage.

Example 2: If the sender embeds covert information into private keys, he/she can share the random parameter in advan-

ce to help the receiver recover the private key. However, this may raise privacy leakage issues. Suppose that a backdoored Bitcoin hardware wallet generates the parameter through $k = G(\cdot)$ rather than generates the random parameter k randomly, where $G(\cdot)$ is a deterministic pseudorandom number generator (PRNG). The attacker keeps the seed of the PRNG to learn the sequence of generated pseudorandom numbers. For all transactions created by such a wallet, the attacker can calculate the corresponding private key d through $d = r^{-1}(G(\cdot)s - h)$.

2) Leakage of identities: The leakage of identities refers to linking a transaction to the IP address [185], [186], [187] of the transaction creator. Biryukov *et al.* [128] believed that subliminal channels in Zash [188] may destroy the anonymity of Zcash users. Specifically, Zcash allows a third party to generate proofs for transactions in a lightweight wallet scenario due to the high computational complexity of the proof generation process. The third party can implement a covert channel in the process of proof generation which transmits the IP address of the transaction creator. The third party is then able to extract the IP address and link it to corresponding transactions.

Similar leakage can also occur in hardware wallets. Hardware wallet sellers can design covert channels in their wallets to link transactions to the real identities of transaction creators. Hardware wallet sellers record each wallet device ID and the identity of its buyer. Hardware wallet sellers also deploy a covert channel for each wallet. The covert channel transmits the wallet ID. As long as transactions are generated with such hardware wallets, sellers can recover the ID and associate the transaction with the buyer's real identity based on their ID-buyer records.

B. Privacy Issues for Communication Parties

Communicating parties' identities might be still leaked since they can be inferred through transaction tracing technologies. Besides, if communication parties access the blockchain network through the third-party API, the API providers can also record communication parties' identities.

1) For the sender: Input addresses of covert transactions belong to the sender. Although the input addresses are anonymous, organizations like Chainalysis [189] can still infer identities behind these addresses [190], [191] through currency-level traceability and real-name authentication information recorded by the exchange. The fact also emphasizes the necessity of currency-level obfuscation. Besides, one flaw of BCSC is that covert transactions are permanently and publicly stored on-chain and cannot be changed. The flaw may also incur privacy issues since there is always a possibility of covert transactions being identified. Once covert transactions are identified, the sender's identity may be traced through transaction tracing technologies. Yin *et al.* [146] proposed a cleaning scheme to prevent covert transactions from being permanently stored on-chain. They first created a covert transaction with a low fee. They then created another transaction (called a conflicting transaction) with the same UTXO as the first covert transaction, while the fee of the conflicting transaction is set higher. They broadcast the conflicting transaction after the receiver has successfully read covert information and before the covert transaction is on-chain. However, they just came up with the concept without experimenting with it. They also pointed out that the way to set the fees for conflicting transactions and covert transactions, and the time to send conflicting transactions are critical problems to be considered. Therefore, whether the clearing method is feasible still needs to be verified.

The way the sender broadcasts covert transactions may also expose the sender's identity. The sender can broadcast covert transactions in two ways: third-party APIs and self-maintained full nodes. Broadcasting transactions through third-party APIs will enable third parties to link the broadcasted transactions to the sender's IP address; broadcasting transactions through self-maintained full nodes may also lead to identity traceability with a low probability [190].

2) For the receiver: the way for the receiver to access the blockchain network may expose its real identity. The receiver can also request transaction data through third-party APIs or a self-maintained full node. Long-term requests for transactional data from third-party APIs may raise suspicion and expose the fact that the receiver is communicating. It is relatively more secure to request transaction data through a self-maintaining full node.

C. Countermeasures

To prevent the privacy leakage of blockchain users and the identity leakage of communication parties, we list countermeasures as below:

1) For blockchain users:

- **Avoid private key reuse to defend against private key exposure attacks.** Private keys can be leaked when they are used for the second time. Therefore, as long as the private key is not reused, it cannot be leaked. Blockchain users can avoid reusing private keys by always creating a new address for change. In UTXO model-based blockchains like Bitcoin, this new address is called the change address.
- **Use secure open-source wallets for management.** In account model-based blockchains such as Ethereum, there is no above change address because there is only one output address of the transaction. Using a hardware wallet in an account-based blockchain is risky because the reuse of private keys cannot be avoided. Users of such blockchains are advised to create transactions with secure open-source wallets and do not use closed-source hardware wallets.
- **Create transactions locally to resist identity exposure attacks.** Regardless of how much computing resource the process of creating a transaction consumes, blockchain users are suggested to create transactions locally without relying on third parties.

2) For communication parties:

- **Implement currency-level obfuscation.** The sender should utilize the coin mixing technology to obfuscate the association between addresses, thereby reducing the possibility of identity being traced.

- **Employ proxy technologies.** The sender and the receiver may protect identity privacy by changing IP through proxy technologies [192] and requesting data from different APIs.

D. Summary and Insights

While enabling more imperceptible channels, covert communication in blockchain may also cause privacy issues. Covert communication in blockchain may undermine the security of the blockchain system by leaking the privacy of blockchain users. Blockchain users can avoid such privacy leakage via privacy-conscious countermeasures such as avoiding address reuse. The effectiveness of these countermeasures confirms the necessity of creating a new change address for each transaction and that the UTXO model is more secure than the account model. On the other hand, communicating parties' identities may still be traced in the blockchain environment. However, covert communication in blockchain is still more secure than traditional covert communication – something is better than nothing.

IX. Challenges and Future Directions

Based on the above systematic studies on blockchain-based covert communication, we describe challenges and potential research directions as follows.

A. Challenges

1) Existing research is mostly the implementation of one-to-one communication and lacks blockchain-based covert group communication schemes: Blockchain is suitable for covert group communication since each blockchain node can be the receiver of group communication. Moreover, a defect of traditional covert group communication is that leakage incurred by malicious receivers may expose other receivers' identities. Whereas the receivers' addresses do not appear in blockchain-based covert group communication, and any receiver can neither learn other receivers' addresses nor identify their identities.

2) Existing BCSCs are not suitable for megabyte-level data transmission due to their low bandwidth and high cost: On the one hand, the bandwidth of existing BCSCs requires an increase. Although BCSCs increase their bandwidth at the expense of concealment, these low-concealment BCSCs only allow 80-byte data embedding per transaction [115], [114], [113]. On the other hand, implementing BCSC needs to create transactions with a high cost. Often a Bitcoin transaction carrying less than 80 bytes of covert information costs 2.36 USD (for Nov.18, 2021) [193]. We do not consider creating covert transactions through free testnets due to their small number of normal transactions. Therefore, transmitting MB-level pictures or files may require a huge cost of more than 30,932 USD. The huge cost of transmitting large files greatly limits the application of blockchain-based covert channels.

3) Existing blockchain-based covert timing channels still require a direct connection between communicating parties, which might incur channel exposure: Lv *et al.* [169] constructed a BCTC by embedding covert information into messages broadcasted between Bitcoin nodes. The advantage is that huge traffic generated by Bitcoin nodes can be embeded covert information without cost, and the BCTC is inexpensive. However, in their model, both the sender and the receiver deploy a node, and both nodes connect each other to communicate. The direct connection is the same as traditional CTC and does not take advantage of the non-directional sending of blockchain-based covert communication. Besides, covert information only exists in a certain period. BCTC with a direct connection lacks concealment since the sender directionally sends traffic to the receiver.

4) Existing currency-level obfuscation techniques are not suitable for blockchain-based covert communication: Existing anonymity enhancement technologies further protect blockchain participants' identities via mixing relations between addresses. Bonneau *et al.* [194] proposed Mixcoin, a centralized currency mixing service with signature accountability. A centralized organization mixes multiple uncorrelated transactions into one or several transactions with many inputs and outputs. However, Mixcoin cannot prevent the centralized organization from privately backing up original transaction records. A dishonest centralized organization still causes disclosure of transaction information before it is mixed. To overcome the centralized limitation of Mixcoin, Maxwell [195] proposed CoinJoin, a decentralized currency mixing service that merges multiple transactions into one mixed transaction. Inputs of original transactions constitute the input of the mixed transaction, and so does the output. Relations between inputs and outputs of original transactions are obfuscated so that adversaries cannot infer relations between input addresses and output addresses of the mixed transaction. However, CoinJoin is threatened with denial of service attacks, and further improvements are proposed [196], [197], [198]. However, all transactions generated by these techniques have a large number of inputs and outputs, making these transactions obvious.

5) There are no unified evaluation methods to compare the concealment of different schemes: Existing research evaluates concealment through qualitative evaluations such as theoretical proof and experimental tests. The qualitative evaluation incurs a large evaluation gap of concealment between different schemes.

6) Existing research ignores the evaluation of anti-traceability: The receiver passively reads blockchain without creating transactions, so its identity can be fully anonymized. However, the sender's identity might be traced through transaction tracing technologies. The sender's identity is not completely anonymous. The anonymity of the sender represents the anti-traceability of covert channels.

7) There are no countermeasures to avoid malicious attacks implemented through the blockchain-based covert channel: Blockchain-based covert channels may apply to deanonymization attacks and malicious activities reports. Detecting covert channels in advance helps defend against these attacks and activities. There have been studies on blockchain steganography analysis [173], [174]. However, these studies do not detect any covert channel in the blockchain. The reason may be huge differences between carriers of blockchain-based covert channels and traditional covert channels. Carriers of blockchain-based covert channels are transaction fields, while

carriers of traditional covert channels are PDUs. In addition, carriers such as addresses and signatures have extremely high concealment. Consequently, detection methods of traditional covert channels do not apply to blockchain-based covert channels. Test methods such as the K-S test and conditional entropy, which perform well in traditional covert channel detection, cannot detect blockchain-based covert channels. These test methods designed for detection even become tools for transaction obfuscation [163], [168], [149].

B. Future directions

1) Blockchain-based covert group communication: Constructing covert group communication in blockchain is a valuable research direction. In addition to building such broadcast channels, another key point is to design group key update methods. The key updating is necessary since malicious receivers know the transaction filtering method and are able to generate fake spam transactions to interfere with covert group communication. Besides, traditional covert group communication needs a one-to-one connection between the sender and each receiver when updating the group key. However, blockchain fails to provide such a one-to-one connection. It is, therefore, necessary to study group key update methods and realize blockchain-based covert group communication.

2) BCSC with high bandwidth and low cost: The interplanetary file system (IPFS) [199] is a blockchain-based peer-to-peer hypermedia protocol designed to preserve and grow humanity's knowledge. It also possesses properties of blockchain, such as immutability and decentralization. Besides, IPFS is an inexpensive and efficient system that contributes to building high-bandwidth and low-cost BCSCs [200].

3) BCTC without direct connection: The sender and the receiver do not need to build such a directly-attached network route since traffic in blockchain networks is also broadcasted via flooding propagation. Intermediate nodes help transmit traffic from the sender to the receiver. However, constructing CTC without direct connections in a blockchain network needs to overcome new challenges. For example, the intermediate node's neighbor nodes may interfere with the covert communication process since they also send traffic to the intermediate node. The neighbor nodes' traffic may reorder the intermediate node's traffic, which may decrease the robustness of covert channels. Robust embedding and filtering methods must be designed to construct BCTCs without direct connections.

4) Coin mixing technology without a special transaction structure: Transactions generated by existing mixing techniques have a large number of inputs and outputs, and transactions with this structure are rare. Therefore, creating covert transactions with a mixed address makes them easier for eavesdroppers to identify the covert transactions. It is necessary to study the coin mixing method without a special transaction structure to enhance the anti-traceability of blockchain-based covert communication and protect the identity privacy of communicating parties.

5) Quantitative concealment evaluation methods for blockchain-based covert channels: Concealment is one of the most important properties of covert channels, and a quantitative standard is required to quantify concealment and compare different channels [201].

6) New evaluation metrics for blockchain-based covert channels: Anti-traceability should also be one of the metrics to evaluate blockchain-based covert channels. Anti-traceability of traditional covert channels can be evaluated according to continuous intervals [202], anonymity set size [203], and information entropy [204], which may help evaluate the anti-traceability of blockchain-based covert channels.

7) Blockchain-based covert channel detection: Transactions are mostly obfuscated via character distributions. The openness and transparency of blockchain make the entire ledger easy to obtain and analyze, such that obfuscation of character distributions conceals covert transactions well. In addition to character distributions, transaction structures (such as input/output number) and time-correlated characteristics of transactions (such as UTXO in the same block of Bitcoin) are also crucial to identify covert transactions. Based on the above characteristics, covert transactions may be easier to detect. Utilizing these features to research blockchain-based covert channel detection may work. Besides, combining machine learning [205], [206] and the above characteristics may contribute to detecting blockchain-based covert channels [207], [208], [209].

X. Conclusion

This paper presents a comprehensive survey of blockchain-based covert communication. First, we introduce the model, features, and applications. Then, we extract three key technologies of building blockchain-based covert channels based on existing research. We classify and detail the key technologies and sum up metrics to evaluate blockchain-based covert channels. Privacy issues caused by covert communication in blockchain are also discussed. Finally, challenges and future directions are presented. Our survey provides a roadmap for follow-up researchers. We are dedicated to forwarding the research of covert communication. We believe that our work inspires readers who focus on covert channels.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant U1836212 and Grant 61872041, and in part by the Shandong Provincial Key Research and Development Program under Grant 2021CXGC010106.

References

[1] E. Nakashima, "Researchers identify sophisticated chinese cyberespionage group," *Washington Post*, vol. 28, 2014.
[2] R. J. Deibert, R. Rohozinski, A. Manchanda, N. Villeneuve, and G. Walton, "Tracking ghostnet: Investigating a cyber espionage network," 2009.
[3] R. Rosenthal, "Covert communication in the psychological experiment." *Psychological Bulletin*, vol. 67, no. 5, p. 356, 1967.
[4] B. W. Lampson, "A note on the confinement problem," *Communications of the ACM*, vol. 16, no. 10, pp. 613–615, 1973.
[5] C. Sanders, *Practical packet analysis: Using Wireshark to solve real-world network problems*. No Starch Press, 2017.

[6] P. Yang, Y. Li, and Y. Zang, "Detecting dns covert channels using stacking model," *China Communications*, vol. 17, no. 10, pp. 183–194, 2020.

[7] S. Zander, "Detecting covert channels in fps online games," in *2017 IEEE 42nd Conference on Local Computer Networks (LCN)*. IEEE, 2017, pp. 555–558.

[8] G. Venkataramani, J. Chen, and M. Doroslovacki, "Detecting hardware covert timing channels," *IEEE Micro*, vol. 36, no. 5, pp. 17–27, 2016.

[9] M. Chourib, "Detecting selected network covert channels using machine learning," in *2019 International Conference on High Performance Computing & Simulation (HPCS)*. IEEE, 2019, pp. 582–588.

[10] F. Rezaei, M. Hempel, and H. Sharif, "Towards a reliable detection of covert timing channels over real-time network traffic," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 3, pp. 249–264, 2017.

[11] X. Liu, H. Xue, and Y. Dai, "A self adaptive jamming strategy to restrict covert timing channel," in *2011 2nd International Symposium on Intelligence Information Processing and Trusted Computing*. IEEE, 2011, pp. 1–4.

[12] X. Liu, H. Xue, X. Feng, and Y. Dai, "Design of the multi-level security network switch system which restricts covert channel," in *2011 IEEE 3rd International Conference on Communication Software and Networks*. IEEE, 2011, pp. 233–237.

[13] A. Belozubova, K. Kogos, and A. Epishkina, "On/off covert channel capacity limitation by adding extra delays," in *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*. IEEE, 2021, pp. 2318–2322.

[14] A. Belozubova, A. Epishkina, and K. Kogos, "Dummy traffic generation to limit timing covert channels," in *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. IEEE, 2018, pp. 1472–1476.

[15] A. Belozubova, A. Epishkina, and K. Kogos, "Random delays to limit timing covert channel," in *2016 European Intelligence and Security Informatics Conference (EISIC)*. IEEE, 2016, pp. 188–191.

[16] D. Frolova, K. Kogos, and A. Epishkina, "Traffic normalization for covert channel protecting," in *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*. IEEE, 2021, pp. 2330–2333.

[17] M. A. Elsadig and Y. A. Fadlalla, "A balanced approach to eliminate packet length-based covert channels," in *2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS)*. IEEE, 2017, pp. 1–7.

[18] A. A. Monrat, O. Schelén, and K. Andersson, "A survey of blockchain from the perspectives of applications, challenges, and opportunities," *IEEE Access*, vol. 7, pp. 117 134–117 151, 2019.

[19] J. Partala, "Provably secure covert communication on blockchain," *Cryptography*, vol. 2, no. 3, p. 18, 2018.

[20] U. Bodkhe, S. Tanwar, K. Parekh, P. Khanpara, S. Tyagi, N. Kumar, and M. Alazab, "Blockchain for industry 4.0: A comprehensive review," *IEEE Access*, vol. 8, pp. 79 764–79 800, 2020.

[21] D. Berdik, S. Otoum, N. Schmidt, D. Porter, and Y. Jararweh, "A survey on blockchain for information systems management and security," *Information Processing & Management*, vol. 58, no. 1, p. 102397, 2021.

[22] O. Ali, A. Jaradat, A. Kulakli, and A. Abuhalimeh, "A comparative study: Blockchain technology utilization benefits, challenges and functionalities," *IEEE Access*, vol. 9, pp. 12 730–12 749, 2021.

[23] S. E. Chang and Y. Chen, "When blockchain meets supply chain: A systematic literature review on current development and potential applications," *IEEE Access*, vol. 8, pp. 62 478–62 494, 2020.

[24] Y. Lu, "The blockchain: State-of-the-art and research challenges," *Journal of Industrial Information Integration*, vol. 15, pp. 80–90, 2019.

[25] H. Xiong, T. Dalhaus, P. Wang, and J. Huang, "Blockchain technology for agriculture: applications and rationale," *frontiers in Blockchain*, vol. 3, p. 7, 2020.

[26] K. Salah, M. H. U. Rehman, N. Nizamuddin, and A. Al-Fuqaha, "Blockchain for ai: Review and open research challenges," *IEEE Access*, vol. 7, pp. 10 127–10 149, 2019.

[27] J. B. Bernabe, J. L. Canovas, J. L. Hernandez-Ramos, R. T. Moreno, and A. Skarmeta, "Privacy-preserving solutions for blockchain: Review and challenges," *IEEE Access*, vol. 7, pp. 164 908–164 940, 2019.

[28] R. Zhang, R. Xue, and L. Liu, "Security and privacy on blockchain," *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, pp. 1–34, 2019.

[29] I. Makhdoom, M. Abolhasan, and J. Lipman, "A comprehensive survey of covert communication techniques, limitations and future challenges," *Computers & Security*, p. 102784, 2022.

[30] K. Gopalan, "Audio steganography for information hiding and covert communication–a tutorial," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*. IEEE, 2018, pp. 0242–0243.

[31] J. GGCT, "Audio steganography for secure data communication: A review."

[32] G. Qiao, M. Bilal, S. Liu, Z. Babar, and T. Ma, "Biologically inspired covert underwater acoustic communication-a review," *Physical Communication*, vol. 30, pp. 107–114, 2018.

[33] J. Shen, K. Lian, and Q. Yang, "Research on underwater bionic covert communication," in *International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy*. Springer, 2020, pp. 223–228.

[34] J. Tian, G. Xiong, Z. Li, and G. Gou, "A survey of key technologies for constructing network covert channel," *Security and Communication Networks*, vol. 2020, 2020.

[35] Y. Dai, G. Liu, P. Cao, W. Liu, and J. Zhai, "A survey of wireless covert communications," *Nanjing Xinxi Gongcheng Daxue Xuebao*, vol. 12, no. 1, pp. 45–56, 2020.

[36] S. Li, L. Zhang, and C. Zhao, "An overview of android covert channel," in *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. IEEE, 2018, pp. 482–485.

[37] L. Caviglione, "Trends and challenges in network covert channels countermeasures," *Applied Sciences*, vol. 11, no. 4, p. 1641, 2021.

[38] R. Shrestha, S. Y. Nam, R. Bajracharya, and S. Kim, "Evolution of v2x communication and integration of blockchain for security enhancements," *Electronics*, vol. 9, no. 9, p. 1338, 2020.

[39] M. Wazid, A. K. Das, S. Shetty, and M. Jo, "A tutorial and future research for building a blockchain-based secure communication scheme for internet of intelligent things," *IEEE Access*, vol. 8, pp. 88 700–88 716, 2020.

[40] J. Ali and S. Sofi, "Ensuring security and transparency in distributed communication in iot ecosystems using blockchain technology: Protocols, applications and challenges," *International Journal of Computing and Digital System*, 2021.

[41] M. Tahir, M. H. Habaebi, M. Dabbagh, A. Mughees, A. Ahad, and K. I. Ahmed, "A review on application of blockchain in 5g and beyond networks: Taxonomy, field-trials, challenges and opportunities," *IEEE Access*, vol. 8, pp. 115 876–115 904, 2020.

[42] R. Gupta, A. Kumari, and S. Tanwar, "Fusion of blockchain and artificial intelligence for secure drone networking underlying 5g communications," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4176, 2021.

[43] R. L. Kumar, Q.-V. Pham, F. Khan, M. J. Piran, and K. Dev, "Blockchain for securing aerial communications: Potentials, solutions, and research directions," *Physical Communication*, vol. 47, p. 101390, 2021.

[44] M. N. M. Bhutta, A. A. Khwaja, A. Nadeem, H. F. Ahmad, M. K. Khan, M. A. Hanif, H. Song, M. Alshamari, and Y. Cao, "A survey on blockchain technology: evolution, architecture and security," *IEEE Access*, vol. 9, pp. 61 048–61 073, 2021.

[45] S. Singh, A. S. Hosen, and B. Yoon, "Blockchain security attacks, challenges, and solutions for the future distributed iot network," *IEEE Access*, vol. 9, pp. 13 938–13 959, 2021.

[46] T. R. Gadekallu, Q.-V. Pham, D. C. Nguyen, P. K. R. Maddikunta, N. Deepa, B. Prabadevi, P. N. Pathirana, J. Zhao, and W.-J. Hwang, "Blockchain for edge of things: applications, opportunities, and challenges," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 964–988, 2021.

[47] B. G. Kim, Y.-S. Cho, S.-H. Kim, H. Kim, and S. S. Woo, "A security analysis of blockchain-based did services," *IEEE Access*, vol. 9, pp. 22 894–22 913, 2021.

[48] Y.-A. Xie, J. Kang, D. Niyato, N. T. T. Van, N. C. Luong, Z. Liu, and H. Yu, "Securing federated learning: A covert communication-based approach," *IEEE Network*, 2022.

[49] S. Zander, G. Armitage, and P. Branch, "A survey of covert channels and countermeasures in computer network protocols," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 3, pp. 44–57, 2007.

[50] J. Ullrich, T. Zseby, J. Fabini, and E. Weippl, "Network-based secret communication in clouds: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1112–1144, 2017.

[51] J. Kang, D. Niyato, J. Zhang, D. I. Kim *et al.*, "Reconfigurable intelligent surface-aided joint radar and covert communications: Fundamentals, optimization, and challenges," *arXiv preprint arXiv:2203.02704*, 2022.

[52] K. Shahzad and X. Zhou, "Covert wireless communications under quasi-static fading with channel uncertainty," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1104–1116, 2020.

[53] S.-H. Lee, L. Wang, A. Khisti, and G. W. Wornell, "Covert communication with channel-state information at the transmitter," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2310–2319, 2018.

[54] L. Zhang, T. Huang, W. Rasheed, X. Hu, and C. Zhao, "An enlarging-the-capacity packet sorting covert channel," *IEEE Access*, vol. 7, pp. 145 634–145 640, 2019.

[55] A. Epishkina and K. Kogos, "Covert channels parameters evaluation using the information theory statements," in *2015 5th International Conference on IT Convergence and Security (ICITCS)*. IEEE, 2015, pp. 1–5.

[56] F. Rezaei, M. Hempel, P. L. Shrestha, and H. Sharif, "Evaluation and verification of automated covert channel modeling using a real network platform," in *2014 IEEE Military Communications Conference*. IEEE, 2014, pp. 12–17.

[57] H. Hovhannisyan, K. Lu, and J. Wang, "A novel high-speed ip-timing covert channel: Design and evaluation," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 7198–7203.

[58] G. J. Simmons, "The prisoners problem and the subliminal channel," in *Advances in Cryptology*. Springer, 1984, pp. 51–67.

[59] D. Llamas, C. Allison, and A. Miller, "Covert channels in internet protocols: A survey," in *Proceedings of the 6th Annual Postgraduate Symposium about the Convergence of Telecommunications, Networking and Broadcasting, PGNET*, vol. 2005. Citeseer, 2005.

[60] M. A. Elsadig and Y. A. Fadlalla, "Survey on covert storage channel in computer network protocols: detection and mitigation techniques," *International Journal of Advances in Computer Networks and Its Security*, vol. 6, no. 3, pp. 11–17, 2016.

[61] F. Chen, Y. Wang, H. Song, and X. Li, "A statistical study of covert timing channels using network packet frequency," in *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)*. IEEE, 2015, pp. 166–168.

[62] R. Archibald and D. Ghosal, "Design and performance evaluation of a covert timing channel," *Security and Communication Networks*, vol. 9, no. 8, pp. 755–770, 2016.

[63] T. G. Handel and M. T. Sandford, "Hiding data in the osi network model," in *International Workshop on Information Hiding*. Springer, 1996, pp. 23–38.

[64] S. Cabuk, "Network covert channels: Design, analysis, detection, and elimination," Ph.D. dissertation, Purdue University, 2006.

[65] G. Xu, W. Yang, and L. Huang, "Hybrid covert channel in lte-a: modeling and analysis," *Journal of Network and Computer Applications*, vol. 111, pp. 117–126, 2018.

[66] K. Sawicki, G. Bieszczad, and Z. Piotrowski, "Stegoframeordermac layer covert network channel for wireless ieee 802.11 networks," *Sensors*, vol. 21, no. 18, p. 6268, 2021.

[67] V. K. Vishnoi, "An offline and efficient storage covert channel detection mechanism," 2019.

[68] L. Huang, L. Zhou, and Y. Guo, "Detecting technology of network storage covert channel based on optics algorithm," in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2018, pp. 891–896.

[69] H. Gunadi and S. Zander, "Bro covert channel detection (broccade) framework: scope and background," Technical report, Tech. Rep., 2017.

[70] M. N. Islam, V. C. Patil, and S. Kundu, "Determining proximal geolocation of iot edge devices via covert channel," in *2017 18th International Symposium on Quality Electronic Design (ISQED)*. IEEE, 2017, pp. 196–202.

[71] M. Zhang and X. Tang, "Hop-limit path mapping algorithm for virtual network embedding," *Wireless Personal Communications*, vol. 95, no. 3, pp. 2033–2048, 2017.

[72] M. A. Elsadig and Y. A. Fadlalla, "Network protocol covert channels: Countermeasures techniques," in *2017 9th IEEE-GCC Conference and Exhibition (GCCCE)*. IEEE, 2017, pp. 1–9.

[73] V. K. Vishnoi, "An offline and efficient storage covert channel detection mechanism," 2019.

[74] J. Oakley, L. Yu, X. Zhong, G. K. Venayagamoorthy, and R. Brooks, "Protocol proxy: An fte-based covert channel," *Computers & Security*, vol. 92, p. 101777, 2020.

[75] P. Nowakowski, P. Żórawski, K. Cabaj, and W. Mazurczyk, "Study of the error detection and correction scheme for distributed network covert channels," in *The 16th International Conference on Availability, Reliability and Security*, 2021, pp. 1–8.

[76] A. Ameri and D. Johnson, "Covert channel over network time protocol," in *Proceedings of the 2017 International Conference on Cryptography, Security and Privacy*, 2017, pp. 62–65.

[77] K. G. Kogos, E. I. Seliverstova, and A. V. Epishkina, "Review of covert channels over HTTP: Communication and countermeasures," in *2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. IEEE, 2017, pp. 459–462.

[78] W. Graniszewski, J. Krupski, and K. Szczypiorski, "Somsteg-framework for covert channel, and its detection, within http." *J. Univers. Comput. Sci.*, vol. 24, no. 7, pp. 864–891, 2018.

[79] G. Daneault and D. Johnson, "Client-initiated HTTP covert channels using relays," in *2016 4th International Symposium on Digital Forensic and Security (ISDFS)*. IEEE, 2016, pp. 32–37.

[80] R. Bortolameotti, T. van Ede, M. Caselli, M. H. Everts, P. Hartel, R. Hofstede, W. Jonker, and A. Peter, "Decanter: Detection of anomalous outbound HTTP traffic by passive application fingerprinting," in *Proceedings of the 33rd Annual computer security applications Conference*, 2017, pp. 373–386.

[81] B. Wu and H. Liu, "A behavior-based covert channel based on gps deception for smart mobile devices," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–6.

[82] J. Xing, Q. Kang, and A. Chen, "Netwarden: Mitigating network covert channels while preserving performance," in *29th {USENIX} Security Symposium ({USENIX} Security 20)*, 2020, pp. 2039–2056.

[83] A. Liguori, F. Benedetto, G. Giunta, N. Kopal, and A. Wacker, "Analysis and monitoring of hidden tcp traffic based on an open-source covert timing channel," in *2015 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2015, pp. 667–674.

[84] P. Wang, S. Lan, J. Zhang, and G. Liu, "Covert timing channel method based on tcp timestamp option," *Journal of PLA University of Science and Technology(Natural Science Edition)*, vol. 16, no. 2, pp. 120–125, 2015.

[85] M. Azadmanesh, M. Mahdavi, and B. S. Ghahfarokhi, "A reliable and efficient micro-protocol for data transmission over an rtp-based covert channel," *Multimedia Systems*, vol. 26, no. 2, pp. 173–190, 2020.

[86] Y.-a. Tan, X. Xu, C. Liang, X. Zhang, Q. Zhang, and Y. Li, "An end-to-end covert channel via packet dropout for mobile networks," *International Journal of Distributed Sensor Networks*, vol. 14, no. 5, p. 1550147718779568, 2018.

[87] J. Zheng, S. Li, S. Hao, Y. Li, and Y. Zhang, "Zm-ctc: Covert timing channel construction method based on zigzag matrix," *Computer Communications*, 2021.

[88] X. Zhang, L. Pang, L. Guo, and Y. Li, "Building undetectable covert channels over mobile networks with machine learning," in *International Conference on Machine Learning for Cyber Security*. Springer, 2020, pp. 331–339.

[89] X. Zhang, C. Liang, Q. Zhang, Y. Li, J. Zheng, and Y.-a. Tan, "Building covert timing channels by packet rearrangement over mobile networks," *Information Sciences*, vol. 445, pp. 66–78, 2018.

[90] X. Zhang, L. Zhu, X. Wang, C. Zhang, H. Zhu, and Y.-a. Tan, "A packet-reordering covert channel over volte voice and video traffics," *Journal of Network and Computer Applications*, vol. 126, pp. 29–38, 2019.

[91] M. Byrenheid, M. Rossberg, G. Schaefer, and R. Dorn, "Covert-channel-resistant congestion control for traffic normalization in uncontrolled networks," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–7.

[92] M. A. Elsadig and Y. A. Fadlalla, "A balanced approach to eliminate packet length-based covert channels," in *2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS)*. IEEE, 2017, pp. 1–7.

[93] A. Sokolov and K. Kogos, "Inter-packet delays normalization to limit ip covert timing channels," *Procedia Computer Science*, vol. 169, pp. 400–406, 2020.

[94] B. B. Yilmaz, N. Sehatbakhsh, A. Zajić, and M. Prvulovic, "Communication model and capacity limits of covert channels created by software activities," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1891–1904, 2019.

[95] R. Soltani, D. Goeckel, D. Towsley, and A. Houmansadr, "Fundamental limits of covert packet insertion," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3401–3414, 2020.

[96] A. Belozubova and K. Kogos, "How to limit capacity of timing covert channel by adding extra delays," *Procedia Computer Science*, vol. 190, pp. 64–70, 2021.

[97] J. Chow, X. Li, and X. Mountrouidou, "Raising flags: Detecting covert storage channels using relative entropy," in *2017 IEEE International*

Conference on Intelligence and Security Informatics (ISI). IEEE, 2017, pp. 25–30.

[98] Z. Wang, L. Huang, W. Yang, and Z. He, "A classifier method for detection of covert channels over lte," in *2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*. IEEE, 2017, pp. 454–460.

[99] J. Xing, A. Morrison, and A. Chen, "Netwarden: Mitigating network covert channels without performance loss," in *11th {USENIX} Workshop on Hot Topics in Cloud Computing (HotCloud 19)*, 2019.

[100] Y. Li, L. Ding, J. Wu, Q. Cui, X. Liu, B. Guan, and Y. Wang, "Survey on key issues in networks covert channel," *Journal of Software*, vol. 8, pp. 2470–2490, 2019.

[101] S. Cabuk, C. E. Brodley, and C. Shields, "Ip covert timing channels: design and detection," in *Proceedings of the 11th ACM conference on Computer and communications security*, 2004, pp. 178–187.

[102] S. Gianvecchio, H. Wang, D. Wijesekera, and S. Jajodia, "Model-based covert timing channels: Automated modeling and evasion," in *International Workshop on Recent Advances in Intrusion Detection*. Springer, 2008, pp. 211–230.

[103] X. Zhang, L. Guo, Y. Xue, and Q. Zhang, "A two-way volte covert channel with feedback adaptive to mobile network environment," *IEEE Access*, vol. 7, pp. 122 214–122 223, 2019.

[104] G. Xin, Y. Liu, T. Yang, and Y. Cao, "An adaptive audio steganography for covert wireless communication," *Security and Communication Networks*, vol. 2018, 2018.

[105] M. Belotti, N. Božić, G. Pujolle, and S. Secci, "A vademecum on blockchain technologies: When, which, and how," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3796–3838, 2019.

[106] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, and M. H. Rehmani, "Applications of blockchains in the internet of things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1676–1717, 2018.

[107] BitinfoCharts. Bitcoin, ethereum transactions historical chart. Accessed: August 9, 2021. [Online]. Available: https://bitinfocharts.com/comparison/transactions-btc-eth.html/

[108] T. Tiemann, S. Berndt, T. Eisenbarth, and M. Liskiewicz, ""
act natural!": Having a private chat on a public blockchain," *Cryptology ePrint Archive*, 2021.

[109] A. Hartl, R. Annessi, and T. Zseby, "A subliminal channel in eddsa: Information leakage with high-speed signatures," in *Proceedings of the 2017 International Workshop on Managing Insider Security Threats*, 2017, pp. 67–78.

[110] Z. Guo, L. Shi, M. Xu, and H. Yin, "Mrcc: A practical covert channel over monero with provable security," *IEEE Access*, vol. 9, pp. 31 816–31 825, 2021.

[111] S. Adair, R. Deibert, R. Rohozinski, N. Villeneuve, and G. Walton, "Shadows in the cloud: Investigating cyber espionage 2.0," *A joint report of the Information Warfare Monitor and Shadowserver Foundation, Toronto*, 2010.

[112] A. Kurt, E. Erdin, M. Cebe, K. Akkaya, and A. S. Uluagac, "Lnbot: a covert hybrid botnet on bitcoin lightning network for fun and profit," in *European Symposium on Research in Computer Security*. Springer, 2020, pp. 734–755.

[113] S. T. Ali, P. McCorry, P. H.-J. Lee, and F. Hao, "Zombiecoin: Powering next-generation botnets with bitcoin," in *International Conference on Financial Cryptography and Data Security*. Springer, 2015, pp. 34–48.

[114] S. T. Ali, P. McCorry, P. H.-J. Lee, and F. Hao, "Zombiecoin 2.0 : managing next-generation botnets using bitcoin," *International Journal of Information Security*, vol. 17, no. 4, pp. 411–422, 2018.

[115] F. Franzoni, I. Abellan, and V. Daza, "Leveraging bitcoin testnet for bidirectional botnet command and control systems," in *International Conference on Financial Cryptography and Data Security*. Springer, 2020, pp. 3–19.

[116] D. Frkat, R. Annessi, and T. Zseby, "Chainchannels: Private botnet communication over public blockchains," in *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. IEEE, 2018, pp. 1244–1252.

[117] D. Frkat, "Subliminal channels in blockchain applications for hidden botnet communication," Ph.D. dissertation, Wien, 2019.

[118] H. Zhao, H. Shu, F. Kang, and Y. Xing, "High resistance botnet based on smart contract," *Chinese Journal of Network and Information Security*, vol. 7, no. 4, 2021.

[119] D. Kamenski, A. Shaghaghi, M. Warren, and S. S. Kanhere, "Attacking with bitcoin: Using bitcoin to build resilient botnet armies," in *Computational Intelligence in Security for Information Systems Conference*. Springer, 2019, pp. 3–12.

[120] F. Schär, "Cryptocurrencies: Miner heterogeneity, botnets, and proof-of-work efficiency," *Frontiers in Blockchain*, vol. 3, p. 16, 2020.

[121] Analyzing the pony c&c server hidden in bitcoin. Accessed: August 24, 2021. [Online]. Available: https://www.freebuf.com/articles/blockchain-articles/221822.html/

[122] K. Eisenkraft and A. Olshtein, "Pony's C&C servers hidden inside the bitcoin blockchain," Technical Report. Check Point. https://research.checkpoint.com/2019/ponys , Tech. Rep., 2019.

[123] E. B. Sasson, A. Chiesa, C. Garman, M. Green, I. Miers, E. Tromer, and M. Virza, "Zerocash: Decentralized anonymous payments from bitcoin," in *2014 IEEE Symposium on Security and Privacy*. IEEE, 2014, pp. 459–474.

[124] S. Noether, "Ring signature confidential transactions for monero." *IACR Cryptol. ePrint Arch.*, vol. 2015, p. 1098, 2015.

[125] T. Lu, P. Yao, L. Zhao, Y. Li, F. Xie, and Y. Xia, "An analysis of attacks against anonymous communication networks," in *2014 7th International Conference on Security Technology*. IEEE, 2014, pp. 38–40.

[126] I. S. Moskowitz, R. E. Newman, and P. F. Syverson, "Quasi-anonymous channels," NAVAL RESEARCH LAB WASHINGTON DC CENTER FOR HIGH ASSURANCE COMPUTING SYSTEMS , Tech. Rep., 2003.

[127] J. Fan, J. Xu, M. H. Ammar, and S. B. Moon, "Prefix-preserving ip address anonymization: Measurement-based security evaluation and a new cryptography-based scheme," *Computer Networks*, vol. 46, no. 2, pp. 253–272, 2004.

[128] A. Biryukov, D. Feher, and G. Vitto, "Privacy aspects and subliminal channels in zcash," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019, pp. 1813–1830.

[129] M. Wang, Q. Wu, B. Qin, Q. Wang, J. Liu, and Z. Guan, "Lightweight and manageable digital evidence preservation system on bitcoin," *Journal of Computer Science and Technology*, vol. 33, no. 3, pp. 568–586, 2018.

[130] T.-H. Chen, W.-B. Lee, H.-B. Chen, and C.-L. Wang, "Revisitedthe subliminal channel in blockchain and its application to iot security," *Symmetry*, vol. 13, no. 5, p. 855, 2021.

[131] K. Gai, Y. Wu, L. Zhu, L. Xu, and Y. Zhang, "Permissioned blockchain and edge computing empowered privacy-preserving smart grid networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7992–8004, 2019.

[132] N. Alsalami and B. Zhang, "Uncontrolled randomness in blockchains: Covert bulletin board for illicit activity," in *2020 IEEE/ACM 28th International Symposium on Quality of Service (IWQoS)*. IEEE, 2020, pp. 1–10.

[133] N. Alsalami and B. Zhang, "Utilizing public blockchains for censorship-circumvention and iot communication," in *2019 IEEE Conference on Dependable and Secure Computing (DSC)*. IEEE, 2019, pp. 1–7.

[134] M. Conti, E. S. Kumar, C. Lal, and S. Ruj, "A survey on security and privacy issues of bitcoin," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3416–3452, 2018.

[135] A. Zhou, Y. Zhong, Z. Zuo, and L. Zhang, "D-bitbot: a p2p duplex botnet model in bitcoin network," *Journal of Harbin Institute of Technology*, vol. 52, no. 5, pp. 66–74, 2020.

[136] M. D. Sleiman, A. P. Lauf, and R. Yampolskiy, "Bitcoin message: Data insertion on a proof-of-work cryptocurrency system," in *2015 International Conference on Cyberworlds (CW)*. IEEE, 2015, pp. 332–336.

[137] A. Tomescu and S. Devadas, "Catena: Efficient non-equivocation via bitcoin," in *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2017, pp. 393–409.

[138] M. Ali, J. Nelson, R. Shea, and M. J. Freedman, "Blockstack: A global naming and storage system secured by blockchains," in *2016 {USENIX} annual technical conference ({USENIX}{ATC} 16)*, 2016, pp. 181–194.

[139] Bitcoin transaction 4a5e1e4baab89f3a32518a88c31bc87f618f76673e2cc77ab2127b7afdeda33b. Accessed: August 9, 2021. [Online]. Available: https://btc.bitaps.com/4a5e1e4baab89f3a32518a88c31bc87f618f76673e2cc77ab2127b7afdeda33b/

[140] R. Matzutt, J. Hiller, M. Henze, J. H. Ziegeldorf, D. Müllmann, O. Hohlfeld, and K. Wehrle, "A quantitative analysis of the impact of arbitrary blockchain content on bitcoin," in *International Conference*

on Financial Cryptography and Data Security. Springer, 2018, pp. 420–438.

[141] R. Matzutt, O. Hohlfeld, M. Henze, R. Rawiel, J. H. Ziegeldorf, and K. Wehrle, "Poster: I don't want that content! on the risks of exploiting bitcoin's blockchain as a content store," in *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 2016, pp. 1769–1771.

[142] A. Sward, I. Vecna, and F. Stonedahl, "Data insertion in bitcoin's blockchain," *Ledger*, vol. 3, 2018.

[143] A. Fionov, "Exploring covert channels in bitcoin transactions," in *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*. IEEE, 2019, pp. 0059–0064.

[144] F. Gao, L. Zhu, K. Gai, C. Zhang, and S. Liu, "Achieving a covert channel over an open blockchain network," *IEEE Network*, vol. 34, no. 2, pp. 6–13, 2020.

[145] M. Bartoletti and L. Pompianu, "An analysis of bitcoin op_return metadata," in *International Conference on Financial Cryptography and Data Security*. Springer, 2017, pp. 218–230.

[146] J. Yin, X. Cui, C. Liu, Q. Liu, T. Cui, and Z. Wang, "Coinbot: A covert botnet in the cryptocurrency network," in *International Conference on Information and Communications Security*. Springer, 2020, pp. 107–125.

[147] A. I. Basuki and D. Rosiyadi, "Joint transaction-image steganography for high capacity covert communication," in *2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*. IEEE, 2019, pp. 41–46.

[148] O. Torki, M. Ashouri-Talouki, and M. Mahdavi, "Blockchain for steganography: advantages, new algorithms and open challenges," *arXiv preprint arXiv:2101.03103*, 2021.

[149] S. Liu, Z. Fang, F. Gao, B. Koussainov, Z. Zhang, J. Liu, and L. Zhu, "Whispers on ethereum: Blockchain-based covert data embedding schemes," in *Proceedings of the 2nd ACM International Symposium on Blockchain and Secure Critical Infrastructure*, 2020, pp. 171–179.

[150] L. Zhang, Z. Zhang, W. Wang, R. Waqas, C. Zhao, S. Kim, and H. Chen, "A covert communication method using special bitcoin addresses generated by vanitygen," *Comput., Mater. Continua*, vol. 65, no. 1, pp. 597–616, 2020.

[151] S. Song and W. Peng, "BLOCCE + : An improved blockchain-based covert communication approach," Master's thesis, 2020.

[152] W. Wang and C. Su, "Ccbrsn: A system with high embedding capacity for covert communication in bitcoin," in *IFIP International Conference on ICT Systems Security and Privacy Protection*. Springer, 2020, pp. 324–337.

[153] J. Qin, Y. Luo, X. Xiang, and Y. Tan, "A novel network covert channel model based on blockchain transaction parity," in *International Conference on Artificial Intelligence and Security*. Springer, 2021, pp. 54–63.

[154] R. Recabarren and B. Carbunar, "Tithonus: A bitcoin based censorship resilient system," *arXiv preprint arXiv:1810.00279*, 2018.

[155] J. Liu, "Research on information hiding methods based on blockchain technology," 2020.

[156] V. Gluhovsky. Eip-627: Whisper specification. Accessed: August 9, 2021. [Online]. Available: https://eips.ethereum.org/EIPS/eip-627/

[157] M. Baden, C. F. Torres, B. B. F. Pontiveros, and R. State, "Whispering botnet command and control instructions," in *2019 Crypto Valley Conference on Blockchain Technology (CVCBT)*. IEEE, 2019, pp. 77–81.

[158] L. Zhang, Z. Zhang, Z. Jin, Y. Su, and Z. Wang, "An approach of covert communication based on the ethereum whisper protocol in blockchain," *International Journal of Intelligent Systems*, vol. 36, no. 2, pp. 962–996, 2021.

[159] Z. Zhang, L. Zhang, W. Rasheed, Z. Jin, T. Ma, H. Chen, and G. Xu, "The research on covert communication model based on blockchain: A case study of ethereums whisper protocol," in *International Conference on Frontiers in Cyber Security*. Springer, 2020, pp. 215–230.

[160] L. Zhang, Z. Zhang, W. Wang, Z. Jin, Y. Su, and H. Chen, "Research on a covert communication model realized by using smart contracts in blockchain environment," *IEEE Systems Journal*, 2021.

[161] Y. Lan, F. Zhang, and H. Tian, "Using monero to realize covert communication," *Journal of Xidian University*, vol. v.47, no. 05, pp. 23–31, 2020.

[162] A. Hartl, R. Annessi, and T. Zseby, "A subliminal channel in eddsa: Information leakage with high-speed signatures," in *Proceedings of the 2017 International Workshop on Managing Insider Security Threats*, 2017, pp. 67–78.

[163] J. Tian, G. Gou, C. Liu, Y. Chen, G. Xiong, and Z. Li, "DLchain: A covert channel over blockchain based on dynamic labels," in *International Conference on Information and Communications Security*. Springer, 2019, pp. 814–830.

[164] H. Cao, H. Yin, F. Gao, Z. Zhang, B. Khoussainov, S. Xu, and L. Zhu, "Chain-based covert data embedding schemes in blockchain," *IEEE Internet of Things Journal*, 2020.

[165] L. Xiangyang, Z. Pei, Z. Mingliang, L. Hao, and Q. Cheng, "A novel covert communication method based on bitcoin transaction," *IEEE Transactions on Industrial Informatics*, 2021.

[166] M. Xu, H. Wu, G. Feng, X. Zhang, and F. Ding, "Broadcasting steganography in the blockchain," in *International Workshop on Digital Watermarking*. Springer, 2019, pp. 256–267.

[167] D. Li, D. Hu, Y. Liu, and Y. Jiang, "Research on command control channel of botnet based on blockchain technology," *Modern Computer*, 2020.

[168] Y. Li, L. Ding, J. Wu, Q. Cui, X. Liu, and B. Guan, "Research on a new network covert channel model in blockchain environment," *Journal on Communications*, vol. 40, no. 5, pp. 67–78, 2019.

[169] J. Lv and X. Cao, "Covert communication technology based on bitcoin," *Journal of Cyber Security*, vol. 6, no. 2, pp. 143–152, 2021.

[170] C. Si, F. Gao, L. Zhu, G. Gong, C. Zhang, Z. Chen, and R. Li, "Covert data transmission mechanism based on dynamic label in blockchain," *Journal of Xidian University*, vol. 47, no. 5, p. 9, 2020.

[171] D. W. Hook and S. J. Porter, "Scaling scientometrics: Dimensions on google bigquery as an infrastructure for large-scale analysis," *Frontiers in Research Metrics and Analytics*, vol. 6, p. 5, 2021.

[172] G. Lopez, D. T. Seaton, A. Ang, D. Tingley, and I. Chuang, "Google bigquery for education: Framework for parsing and analyzing edx mooc data," in *Proceedings of the fourth (2017) ACM conference on learning@ scale*, 2017, pp. 181–184.

[173] A. A. Giron, J. E. Martina, and R. Custódio, "Steganographic analysis of blockchains," *Sensors*, vol. 21, no. 12, p. 4078, 2021.

[174] A. A. Giron, J. E. Martina, and R. Custódio, "Bitcoin blockchain steganographic analysis," in *International Conference on Applied Cryptography and Network Security*. Springer, 2020, pp. 41–57.

[175] K. Gagneja, "Traceability of cryptocurrency transactions using blockchain analytics," *International Journal of Computing and Digital Systems*, vol. 9, no. 2, pp. 159–165, 2020.

[176] G. Kappos, H. Yousaf, M. Maller, and S. Meiklejohn, "An empirical analysis of anonymity in zcash," in *27th {USENIX} Security Symposium ({USENIX} Security 18)*, 2018, pp. 463–477.

[177] J. Willett, M. Hidskes, D. Johnston, R. Gross, and M. Schneider, "Omni protocol specification (formerly mastercoin)," *white paper), accessed January*, vol. 28, 2016.

[178] One-input and two-output transactions. Accessed: August 9, 2021. [Online]. Available: https://transactionfee.info/charts/transactions-1in-2out/

[179] R. Xiao, W. Ren, T. Zhu, and K.-K. R. Choo, "A mixing scheme using a decentralized signature protocol for privacy protection in bitcoin blockchain," *IEEE Transactions on Dependable and Secure Computing*, 2019.

[180] T. Ruffing, P. Moreno-Sanchez, and A. Kate, "P2p mixing and unlinkable bitcoin transactions." in *NDSS*, 2017, pp. 1–15.

[181] A. Fionov, "Exploring covert channels in bitcoin transactions," in *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*. IEEE, 2019, pp. 0059–0064.

[182] P. L. Shrestha, M. Hempel, H. Sharif, and H.-H. Chen, "An event-based unified system model to characterize and evaluate timing covert channels," *IEEE Systems Journal*, vol. 10, no. 1, pp. 271–280, 2014.

[183] J.-C. Wang, H.-M. Lee, C.-W. Chen, and A. B. Jeng, "Estimating intent-based covert channel bandwidth by time series decomposition analysis in android platform," in *2017 IEEE Conference on Application, Information and Network Security (AINS)*. IEEE, 2017, pp. 31–36.

[184] A. Davenport and S. Shetty, "Air gapped wallet schemes and private key leakage in permissioned blockchain platforms," in *2019 IEEE International Conference on Blockchain (Blockchain)*. IEEE, 2019, pp. 541–545.

[185] A. Biryukov, D. Khovratovich, and I. Pustogarov, "Deanonymisation of clients in bitcoin p2p network," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014, pp. 15–29.

[186] A. Biryukov and S. Tikhomirov, "Deanonymization and linkability of cryptocurrency transactions based on network analysis," in *2019 IEEE European symposium on security and privacy (EuroS&P)*. IEEE, 2019, pp. 172–184.

[187] A. Biryukov and S. Tikhomirov, "Transaction clustering using network traffic analysis for bitcoin and derived blockchains," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2019, pp. 204–209.

[188] D. Hopwood, S. Bowe, T. Hornby, and N. Wilcox, "Zcash protocol specification," *GitHub: San Francisco, CA, USA*, p. 1, 2016.

[189] Chainalysis, "The 2020 state of crypto crime," 2020.

[190] F. Gao, H. Mao, Z. Wu, M. Shen, L. Zhu, and Y. Li, "Lightweight transaction tracing technology for bitcoin," *Chinese Journal of Computers*, vol. 41, no. 5, pp. 989–1004, 2018.

[191] L. Zhu, F. Gao, M. Shen, Y. Li, B. Zheng, H. Mao, and Z. Wu, "Survey on privacy preserving techniques for blockchain technology," *Journal of Computer Research and Development*, vol. 54, no. 10, p. 2170, 2017.

[192] R. Tourani, S. Misra, T. Mick, and G. Panwar, "Security, privacy, and access control in information-centric networking: A survey," *IEEE communications surveys & tutorials*, vol. 20, no. 1, pp. 566–600, 2017.

[193] Bitcoin average transaction fee. Accessed: November 19, 2021. [Online]. Available: https://ycharts.com/indicators/bitcoin_average_transaction_fee/

[194] J. Bonneau, A. Narayanan, A. Miller, J. Clark, J. A. Kroll, and E. W. Felten, "Mixcoin: Anonymity for bitcoin with accountable mixes," in *International Conference on Financial Cryptography and Data Security*. Springer, 2014, pp. 486–504.

[195] G. Maxwell, "Coinjoin: Bitcoin privacy for the real world," in *Post on Bitcoin forum*, 2013.

[196] R. Xiao, W. Ren, T. Zhu, and K.-K. R. Choo, "A mixing scheme using a decentralized signature protocol for privacy protection in bitcoin blockchain," *IEEE Transactions on Dependable and Secure Computing*, 2019.

[197] Y. Liu, R. Li, X. Liu, J. Wang, C. Tang, and H. Kang, "Enhancing anonymity of bitcoin based on ring signature algorithm," in *2017 13th International conference on computational intelligence and security (CIS)*. IEEE, 2017, pp. 317–321.

[198] T. Ruffing and P. Moreno-Sanchez, "Valueshuffle: Mixing confidential transactions for comprehensive transaction privacy in bitcoin," in *International Conference on Financial Cryptography and Data Security*. Springer, 2017, pp. 133–154.

[199] J. Benet, "Ipfs-content addressed, versioned, p2p file system," *arXiv preprint arXiv:1407.3561*, 2014.

[200] W. She, L. Huo, Z. Tian, Y. Zhuang, C. Niu, and W. Liu, "A double steganography model combining blockchain and interplanetary file system," *Peer-to-Peer Networking and Applications*, pp. 1–14, 2021.

[201] H. Du, D. Niyato, Y.-a. Xie, Y. Cheng, J. Kang, and D. I. Kim, "Performance analysis and optimization for jammer-aided multi-antenna uav covert communication," *arXiv preprint arXiv:2202.00973*, 2022.

[202] A. Castiglione, M. Nappi, and C. Pero, "Towards the design of a covert channel by using web tracking technologies," in *International Conference on Dependability in Sensor, Cloud, and Big Data Systems and Applications*. Springer, 2019, pp. 231–246.

[203] L. Zhang, T. Huang, X. Hu, Z. Zhang, W. Wang, D. Guan, C. Zhao, and S. Kim, "A distributed covert channel of the packet ordering enhancement model based on data compression," *CMC-COMPUTERS MATERIALS & CONTINUA*, vol. 64, no. 3, pp. 2013–2030, 2020.

[204] C. Diaz, S. Seys, J. Claessens, and B. Preneel, "Towards measuring anonymity," in *International Workshop on Privacy Enhancing Technologies*. Springer, 2002, pp. 54–68.

[205] J. Kang, Z. Xiong, X. Li, Y. Zhang, D. Niyato, C. Leung, and C. Miao, "Optimizing task assignment for reliable blockchain-empowered federated edge learning," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1910–1923, 2021.

[206] J. Kang, X. Li, J. Nie, Y. Liu, M. Xu, Z. Xiong, D. Niyato, and Q. Yan, "Communication-efficient and cross-chain empowered federated learning for artificial intelligence of things," *IEEE Transactions on Network Science and Engineering*, 2022.

[207] S. Saeli, F. Bisio, P. Lombardo, and D. Massa, "DNS covert channel detection via behavioral analysis: a machine learning approach," *arXiv preprint arXiv:2010.01582*, 2020.

[208] Y. Sun, L. Zhang, and C. Zhao, "A study of network covert channel detection based on deep learning," in *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. IEEE, 2018, pp. 637–641.

[209] İ. G. Çavuşo, H. Alemdar, E. Onur *et al.*, "Covert channel detection using machine learning," in *2020 28th Signal Processing and Communications Applications Conference (SIU)*. IEEE, 2020, pp. 1–4.

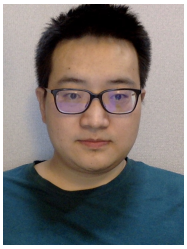
Zhuo Chen received the B.E. degree in information security from the North China Electric Power University, Beijing, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Cyberspace Science and Technology, Beijing Institute of Technology. His current research interests include blockchain technology and covert communication.

Liehuang Zhu (Senior Member, IEEE) is a Full Professor with the School of Cyberspace Science and Technology, Beijing Institute of Technology. He is selected into the Program for New Century Excellent Talents in University from Ministry of Education, China. He has published over 60 SCI-indexed research papers in these areas, as well as a book published by Springer. His research interests include Internet of Things, cloud computing security, Internet, and mobile security. He serves on the editorial boards of three international journals, including IEEE Internet of Things Journal, IEEE Network, and IEEE Transactions on Vehicular Technology. He won the Best Paper Award at IEEE/ACM IWQoS 2017 and IEEE TrustCom 2018.

Peng Jiang (Member, IEEE) received her Ph.D. degree from Beijing University of Posts and Telecommunications in 2017. She is currently an Associate Professor of the School of Cyberspace Science and Technology, Beijing Institute of Technology. She has published more than 40 papers in the area of cybersecurity and has served as a program committee member in many international conferences. Her research interests include cryptography, information security, and blockchain.

Can Zhang received the B.E. degree in computer science and technology from the Beijing Institute of Technology, Beijing, China, in 2017. He is currently working toward the Ph.D. degree with the School of Cyberspace Science and Technology, Beijing Institute of Technology. His current research interests include security & privacy in VANET, cloud computing security, and blockchain technology.

This article has been accepted for publication in IEEE Communications Surveys & Tutorials. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/COMST.2022.3204281



Feng Gao received a Ph.D. degree in computer science and technology from Beijing Institute of Technology, China, in 2018. His research interests include blockchain and cybersecurity.

Dawei Xu is a PhD student with the School of Computer Science and Technology, Beijing Institute of Technology. He engages in science research and education work in College of Cybersecurity, Changchun University. His current research interests include blockchain technology, anonymous communication, big data privacy protection, and machine learning.

Jialing He received the M.S. and Ph.D. degrees from the Beijing Institute of Technology, Beijing, China, in 2018 and 2022, respectively, where she is currently an assist research fellow in college of computer science, Chongqing University, Chongqing, China. Her current research interests include differential privacy and user behavior mining.

Yan Zhang (Fellow, IEEE) received the B.S. degree from the Nanjing University of Post and Telecommunications, Nanjing, China, the M.S. degree from Beihang University, Beihang, China, and the Ph.D. degree from the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore. He is currently a Full Professor with the Department of Informatics, University of Oslo, Oslo, Norway. His research interests include next-generation wireless networks leading to 6G and green and secure cyber-physical systems (e.g., smart grid and transport). Dr. Zhang is a Fellow of IET, an Elected Member of Academia Europaea, the Royal Norwegian Society of Sciences and Letters (DKNVS), the Norwegian Academy of Technological Sciences. Since 2018, he has been a recipient of the global Highly Cited Researcher Award (Web of Science top one).