Blockchain-Empowered Federated Learning: Benefits, Challenges, and Solutions

Zeju Cai, Jianguo Chen, *Member, IEEE*, Yuting Fan, Zibin Zheng, *Fellow, IEEE*, and Keqin Li, *Fellow, IEEE*

(Survey-Tutorial Paper)

***Abstract*—Federated learning (FL) is a distributed machine learning approach that protects user data privacy by training models locally on clients and aggregating them on a parameter server. While effective at preserving privacy, FL systems face limitations such as single points of failure, lack of incentives, and inadequate security. To address these challenges, blockchain technology is integrated into FL systems to provide stronger security, fairness, and scalability. However, blockchain-empowered FL (BC-FL) systems introduce additional demands on network, computing, and storage resources. This survey provides a comprehensive review of recent research on BC-FL systems, analyzing the benefits and challenges associated with blockchain integration. We explore why blockchain is applicable to FL, how it can be implemented, and the challenges and existing solutions for its integration. Additionally, we offer insights on future research directions for the BC-FL system.**

***Index Terms*—Blockchain-empowered federated learning (FL), distributed artificial intelligence, security and privacy.**

I. Introduction

ARTIFICIAL Intelligence (AI) technologies drive the Fourth Industrial Revolution, with user data being essential for training diverse Machine Learning (ML) models [1]. Training high-quality ML models often involves a centralized approach, necessitating internal storage of user data. This raises privacy concerns [2], [3], [4] and highlights the need for stringent privacy protections [5]. In recent years, regions such as the European Union [6], [7], the United States [8], and Singapore [5] have enacted relevant laws and regulations to govern the use of personal data, enhancing privacy protection but potentially hindering the utilization of high-quality data.

Federated Learning (FL) is a privacy-preserving distributed machine learning paradigm that balances user data protection and effective utilization [9], [10], [11]. FL involves training local models on user devices and aggregating these local models into a global model on a server without requiring users to upload their data, addressing the aforementioned privacy concerns. Initially applied to training Gboard [12], FL has proven successful. Its potential extends beyond this, as it can also address the issue of data silos. Data silos refer to the isolated or dispersed nature of data, making access to this data extremely challenging [13]. One cause of data silos is the reluctance of organizations to share data due to privacy or competitive concerns. For instance, due to privacy protection, hospitals may be unwilling to share patient data [14]. In summary, the judicious use of FL can break down data barriers, leading to its widespread application in healthcare [15], [16], finance [17], [18], industry [19], [20] and so on.

While privacy protection and data utilization benefits have popularized FL across industry and academia, they also introduce specific challenges. First, there is a lack of trust among nodes within the FL system [21], [22]. Nodes may worry that their training contributions will be intentionally tampered with or miscalculated, damaging their reputation and deserved rewards. Second, FL systems are vulnerable to attacks from malicious nodes [23], [24]. Malicious users may intentionally provide incorrect information to prevent model convergence and disrupt model training, while malicious servers can recover users' training data from the uploaded models. Third, FL is prone to single point of failure issues [25]. In traditional FL architectures, the central server is responsible for aggregating and updating global model parameters. If the central server is attacked or fails, the entire system's operation is severely affected, leading to interruptions in the training process, data loss, and irrecoverable model states.

Blockchain is essentially a distributed ledger, and its successful application in cryptocurrencies demonstrates its potential to build trust, security, and transparency [26], [27], [28]. Consequently, numerous studies have integrated blockchain with FL systems to enhance functionality, creating blockchain-empowered FL (BC-FL) systems. Analyzing existing BC-FL literature, we find that blockchain's enhancement of different aspects of FL originates from its distinct properties. First, blockchain's transparency and immutability can alleviate the lack of trust among nodes within the FL system. By recording data requiring consensus in the FL system on the blockchain, these data cannot be tampered with by malicious nodes,

Received 5 July 2024; revised 9 December 2024; accepted 28 January 2025. Date of publication 13 February 2025; date of current version 4 September 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62372486 and in part by the Natural Science Foundation of Guangdong Province under Grant 2023A1515011179. Recommended for acceptance by M. Huang. *(Corresponding author: Jianguo Chen.)*

Zeju Cai and Yuting Fan are with the School of Software Engineering, Sun Yat-sen University, Guangzhou 510275, China (e-mail: caizj9@mail2.sysu.edu.cn; fanyt6@mail2.sysu.edu.cn).

Jianguo Chen and Zibin Zheng are with the School of Software Engineering, Sun Yat-sen University, Guangzhou 510275, China, and also with the Guangdong Engineering Technology Research Center of Blockchain, Guangzhou 511493, China (e-mail: chenjg33@mail.sysu.edu.cn; zhzibin@mail.sysu.edu.cn).

Keqin Li is with the Department of Computer Science, State University of New York, New Paltz, NY 12246 USA (e-mail: lik@newpaltz.edu).

Digital Object Identifier 10.1109/TBDATA.2025.3541560

2332-7790 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

enhancing trust relationships. Second, through cross-validation of blockchain nodes and other mechanisms, the resistance of the FL system to malicious nodes is improved. Finally, blockchain can replace the centralized server to avoid single point of failure issues. By designing a reasonable consensus mechanism, suitable clients can be selected to undertake model aggregation tasks in each communication round. With the advent of blockchain 2.0, users can develop smart contracts running on the blockchain, endowing BC-FL with greater scalability for automatically running various algorithms [29].

The introduction of blockchain has further driven the development of FL, but blockchain is not a panacea for FL. Our research indicates that blockchain integration poses challenges related to runtime efficiency and storage capacity. First, the consensus mechanism of blockchain adds communication and computation overhead to the BC-FL system. Second, due to the distributed storage nature of blockchain, full nodes need to back up the entire blockchain data. Additionally, the introduction of blockchain can also bring additional security issues, such as Sybil attacks [30].

Currently, several surveys on BC-FL systems have been published. Some focus on the integration of BC-FL with other fields, such as the Internet of Things [31], [32], drones [33], and healthcare [34], [35]. These studies emphasize the specific applications of BC-FL systems rather than their commonalities. Other surveys investigate BC-FL systems in general. Qu et al. conducted a detailed study on the performance of decentralization, attack resistance, and incentive mechanisms in BC-FL systems, and surveyed the system architecture forms of BC-FL [36]. However, they did not investigate transparent reputation mechanisms in BC-FL and thoroughly analyze why blockchain can enhance FL systems, merely classifying the functions of BC-FL. Zhu et al. divided BC-FL system models into three categories and surveyed real-world applications of BC-FL [33]. However, they lacked a comprehensive investigation of single point of failure, reputation mechanisms, security, and privacy issues. Additionally, the aforementioned surveys neglected to conduct a detailed investigation into the negative effects blockchain can bring to BC-FL systems. Sameera et al. summarized the general architecture of BC-FL and detailed how BC-FL addresses security and privacy threats [37]. However, they lacked in-depth research on BC-FL's reputation mechanisms, incentive mechanisms, and system efficiency and storage issues. We believe that the various enhancements and potential challenges blockchain brings to FL systems stem from certain properties or functionalities of blockchain. Meanwhile, some properties of blockchain can play different roles in FL systems depending on their application. The contributions of this work are as follows:

- Starting from the characteristics and functionalities of blockchain, we introduce how blockchain enhances FL systems in terms of decentralization, reputation mechanisms, incentive mechanisms, and security.
- We comprehensively investigate the additional challenges that arise from using blockchain in FL systems, the reasons behind these challenges, and existing solutions.
- We summarize future research directions for blockchain-based FL systems based on existing research.

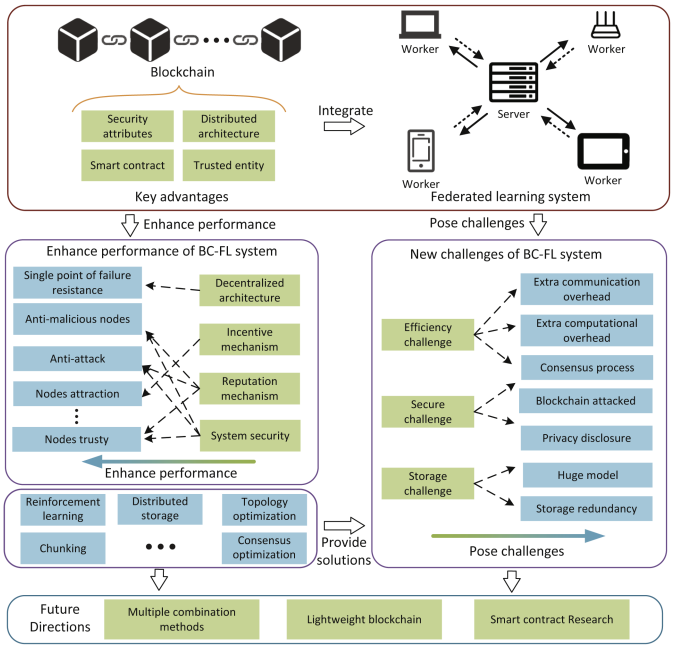


Fig. 1. Main scope of this survey. We begin by exploring the characteristics of blockchain and investigate its enhancement of Federated Learning systems. Next, we discuss the additional challenges introduced by using blockchain in FL systems and review existing solutions. Finally, we outline future research directions for Blockchain-empowered Federated Learning systems.

Investigating how blockchain enhances FL systems from the perspective of blockchain's characteristics is an unaddressed area in existing surveys. This survey can complement similar recent surveys, filling a gap in the research on BC-FL. This survey.'s main scope is illustrated in Fig. 1. Section II provides relevant background knowledge on FL, blockchain. Section III details general BC-FL architecture and how blockchain enhances FL systems. Section IV elaborates on the challenges of using blockchain in FL systems and existing solutions. Section V-A we point out future research directions for BC-FL systems. Finally, Section VI concludes the paper.

II. Background

A. Federated Learning

FL is a privacy-preserving distributed machine learning paradigm proposed by Google [9]. This paradigm involves a network of multiple clients (users) alongside a central server. The clients are tasked with developing local models, which are then consolidated by the server into a unified, global model [38], [39], [40]. This structure allows participants considerable autonomy, enabling them to contribute to the FL framework without disclosing their data to any FL node. Participation in FL training remains at the discretion of the clients. Fig. 3 visually represents the standard FL training methodology. The task initiator selects an FL server to publish the training task. Subsequently, clients related to the training task join the FL training, and the server initializes the global model. In each communication round, the server selects clients to participate in that round of training and distributes the global model to these clients. The clients then use

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

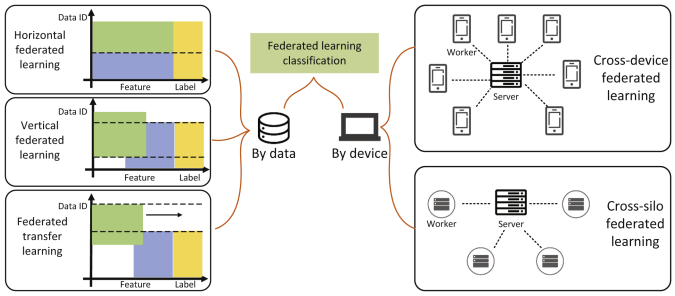


Fig. 2. Categories of FL systems.

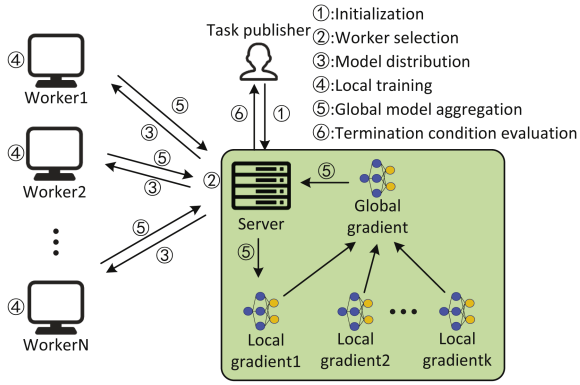


Fig. 3. Standard federated learning (FL) training methodology.

their local datasets to train the global model, resulting in local models, which they send back to the server. The server aggregates these local models into a new global model according to certain rules. The FL process stops when the training termination condition is met; otherwise, training continues.

FL generally has two classification methods, as depicted in Fig. 2. Based on the sample ID and feature distribution of local datasets, FL systems can be divided into Horizontal FL (HFL), Vertical FL (VFL), and Federated Transfer Learning (FTL). In HFL, clients have different sample IDs, but the features for each sample ID are the same. In VFL, clients share the same sample IDs, but the sample features differ, with the features distributed across different clients. In FTL, both the sample IDs and features of the clients differ. Generally, current research on BC-FL is predominantly based on HFL, with only a few studies on VFL [41], [42] and FTL [43]. Despite the datasets' different characteristics, blockchain's role does not fundamentally differ.

Moreover, based on the different types of user devices involved in training, FL can be divided into cross-device FL and cross-silo FL. Cross-device FL involves an extensive array of mobile and IoT devices, potentially numbering up to 10^{10}, often with poor device performance and network conditions [44], [45]. In cross-silo FL systems, the participating clients are typically professional computing nodes maintained by specialized institutions, and the number of clients is generally fewer than 100 [46], [47].

At present, a considerable part of the research on FL is focused on improving performance. However, with the development of FL, many researchers have found that simply improving the performance of FL is not enough. FL is a decentralized system, and its security and ecological construction are very important for the promotion of FL [48]. The introduction of blockchain provides ideas for the security and ecological construction of FL.

B. Blockchain

Blockchain is a distributed ledger system designed to record data in a decentralized manner that ensures security, transparency, and immutability [49], [50]. It consists of a series of interconnected blocks arranged in chronological order. Each block consists of a block header and a block body. The block header contains metadata, such as the block number, timestamp, and the hash of the previous block, ensuring immutability and integrity. The block body stores specific information, such as the training data for FL.

In a blockchain network, there is no centralized authority node, and its autonomous operation relies on consensus mechanisms. In essence, a consensus mechanism is a method by which multiple nodes reach an agreement, ensuring that all nodes have a consistent recognition of the data in the blockchain. Mainstream consensus mechanisms can be categorized into proof-based and committee-based approaches [26]. Proof-based consensus mechanisms prioritize nodes with certain resources, granting them a higher likelihood of adding blocks. For example, Proof of Work (PoW) requires competing nodes to solve a puzzle, with the winner being able to add the block. This method favors nodes with high computational resources. PoS selects a leader to add a new block using a cryptographic random algorithm, with the probability of selection being proportional to the number of tokens held by the node. Therefore, nodes with more tokens have a higher priority [51]. Committee-based consensus mechanisms use a voting process, where a predefined number of votes are required to validate a new block. For instance, in Practical Byzantine Fault Tolerance (PBFT) [52], nodes are divided into primary and backup nodes. The primary node is responsible for proposing new blocks and reaching consensus through a three-phase protocol (pre-prepare, prepare, and commit). In each phase, more than two-thirds of the nodes must vote in favor for the process to continue. Raft [53] is a simpler consensus mechanism that streamlines data synchronization by electing a leader. The leader is responsible for receiving requests and replicating operation logs to other nodes in the cluster, with other nodes maintaining consistency by following the leader's log.

Based on node participation constraints, blockchain can be classified into two main types: permissionless blockchain and permissioned blockchain. Permissionless blockchains (e.g., Bitcoin and Ethereum) allow unrestricted participation of nodes. Permissioned blockchains, on the other hand, are managed by specific organizations, with node access restricted to authorized participants, thereby ensuring enhanced data privacy and security—critical for applications in industries such as finance and government.

The blockchain's replica is stored on every node, which provides the system with high transparency. Combined with an appropriate consensus mechanism and the chain structure, blockchain effectively prevents data tampering, ensuring data

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

security. Furthermore, the combination of transparency and immutability makes blockchain auditable. These features give blockchain significant application value in FL.

III. Blockchain-Empowered Federated Learning

This section introduces how blockchain technology enhances the FL system. These enhancements can be categorized into four aspects: decentralization, reputation evaluation mechanism, incentive evaluation mechanism, and security. Each of these four aspects will be discussed separately.

A. Decentralization

FL traditionally relies on a central parameter server, where clients must continuously communicate with a single FL server. This centralized structure poses significant risks, such as single points of failure and potential malicious server behaviour. Furthermore, node reputation information is solely managed by the server, which is not ideal for developing an open FL ecosystem. Blockchain's decentralization is a core feature that fundamentally addresses these issues and provides a new architectural approach to FL systems. Decentralization can enhance the security, transparency, and reliability of FL systems, with these benefits manifesting in various facets of blockchain's advantages for FL.

Our review of existing literature identifies several factors influencing the decentralization of BC-FL, including system architecture, consensus mechanisms, and smart contracts. The system architecture determines which nodes maintain the blockchain, while the consensus mechanism dictates which nodes have the right to manage the system. Smart contracts can automate various algorithms within the BC-FL system, offering greater scalability. Table I compiles representative BC-FL systems, detailing their use of smart contracts, architecture, consensus mechanisms, and experimental platforms.

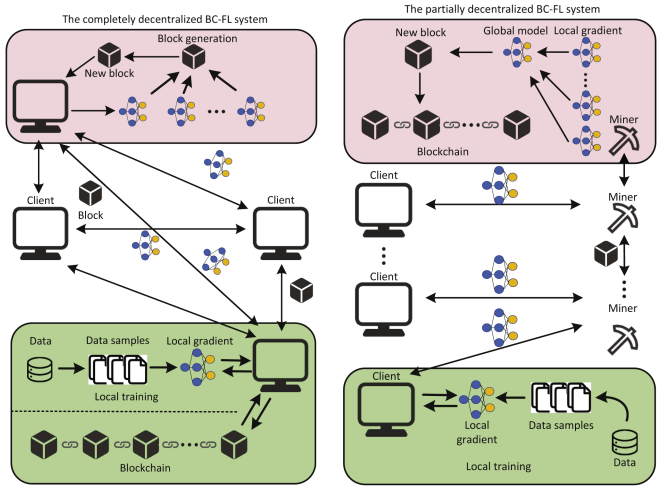
Architecture: BC-FL systems can be categorized by their degree of decentralization: complete and partial. In the completely decentralized BC-FL, all nodes are eligible to participate in the consensus process of the blockchain. Fig. 4 shows its general system architecture. This approach demands high computational and storage capacities from all nodes. Conversely, partially decentralized BC-FL involves only a subset of nodes running the blockchain, while others focus solely on FL training. Fig. 4 shows its general system architecture. The selected nodes that operate the blockchain system are known as super nodes and typically have stronger computing power and better communication conditions. This approach sacrifices some transparency for increased efficiency.

Consensus Mechanism: A considerable portion of the work adopts common blockchain consensus mechanisms such as PoW and PBFT. PoW involves blockchain nodes competing to solve a mathematical problem, with the first solver aggregating models and training information into a new block. Other nodes then verify the block's correctness, and upon majority approval, it is added to the blockchain. In PBFT, a set of consensus nodes is chosen within the BC-FL system, from which a leader node aggregates the model and generates a new block. Other nodes in the set verify the leader's block. Some work has specifically developed consensus mechanisms for BC-FL, such as Proof of Reputation [61]. These custom consensus mechanisms are usually designed to enhance FL functionality or mitigate the disadvantages of blockchain, which will be elaborated on in the subsequent sections.

Smart Contracts: Smart contracts significantly enhance the scalability of BC-FL systems. For instance, model aggregation can be executed via smart contracts, increasing transparency. Additionally, smart contracts can deploy algorithms for detecting and handling malicious nodes, thereby improving system efficiency. They can also manage node reputation

TABLE I
BC-FL Systems Based on Blockchain and Federated Learning

Ref.	Smart contract	Architecture	Consensus algorithm	Platform
Abdel [54]	✓	Complete	Algorand	Other
Fang [55]	×	Partial	Algorand	Other
Feng [56]	✓	Partial	PBFT/Raft	Hyperledger Fabric
Guo [57]	✓	Partial	PBFT	Hyperledger Fabric
Jiang [58]	×	Complete	DPoS	Other
Liu [59]	×	Partial	PoW+PoA	Other
Lu [60]	×	Complete	DPoS	Other
Nguyen [61]	×	Partial	PoR	Other
Nguyen [62]	×	Partial	PoW	Other
Qi [63]	✓	Partial	-	Hyperledger Fabric
Qi [64]	✓	Complete	Modified PBFT	Ethereum
Qu [65]	×	Complete	PoW	Other
Rehman [66]	✓	Complete	-	Ethereum
Wu [67]	×	Complete	PoW	Other
Xu [68]	×	Complete	-	Other
Xu [69]	✓	Complete	-	Other
Zhang [70]	×	Partial	PoW	Other
Zhao [71]	×	Partial	PBFT	Other
Wang [72]	×	Complete	-	Other
Huang [72]	✓	Partial	Raft, HotStuff	FISCO
Ouyang [73]	✓	Partial	PoS	Ethereum
Yuan [74]	✓	Partial	Raft, Modified DAG	Hyperledger, DAG
aloqaily [75]	×	Partial	-	Other
Mu [76]	×	Complete	-	Other
Wahrstatter [77]	×	Complete	PoS	Ethereum



(a) Complete decentralization. (b) Partial decentralization.

Fig. 4. Two decentralized architectures of the Blockchain-Empowered Federated Learning system.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

evaluations and incentive algorithms, further enhancing system transparency.

Next, we will examine some representative architectures of completely decentralized BC-FL systems.

In [69], Xu et al. proposed a BC-FL framework named Blockchain Empowered Secure and Incentive Federated Learning (BESIFL). BESIFL enables any node in the network to initiate FL training requirements. Upon receipt of a requirement, BESIFL selects computing nodes with high computation reputation scores to form a computing pool and assigns them the task of model training. Meanwhile, BESIFL chooses verification nodes with high verification reputation scores to form a verification pool and assigns them the task of model aggregation and verification using pre-defined procedures specified by the smart contract. BESIFL can combat malicious nodes and improve the performance and security of FL, but the node selection scheme is relatively complex and will increase time overhead. Li et al. also proposed a completely decentralized BF-FL system, where each client acts as both a FL trainer and a blockchain miner [78]. After training their local models, clients initiate blockchain transaction requests and broadcast their models by attaching them to the transaction information. Each client aggregates the global model locally after receiving local models from all other clients and starts mining. The winning miner broadcasts a block containing global model information, which is verified by other clients and then written into the blockchain. This method effectively resists threats such as single point failures and malicious attacks, ensures reliable model updates, and provides theoretical analysis. However, this system assumes that all clients possess equal computational power, which may not be realistic in practice. In addition to the two aforementioned decentralized methods, Qu et al. designed a novel approach that utilizes a rotation mechanism with randomness to select committee members for participating in blockchain consensus [79]. This proposed blockchain consensus mechanism greatly reduces additional consumption generated by the blockchain consensus process compared to the PoW mechanism. Committee members are only responsible for aggregating and validating the global model and do not participate in training. The global model is generated by committee members and stored in the blockchain after verification. While the rotation mechanism ensures the mobility of committee members, it can ensure some level of system security. However, this consensus mechanism is only applicable in situations where the number of malicious nodes is small.

The BC-FL systems described below follow the partial decentralization architecture. Feng et al. proposed a BC-FL system for UAVs that maintains the blockchain system only in entities with high computing and storage capabilities, such as base stations and roadside nodes [56]. The approach implements model aggregation operations through smart contracts, replacing traditional parameter servers, and achieving a balance between efficiency and transparency. To address the challenge of online and offline state changes among BC-FL participants, the authors set the maximum waiting time and the required number of local models for each learning round. If any of these conditions are met, the model update contract is triggered, ensuring timely updates while accommodating BC-FL participant availability.

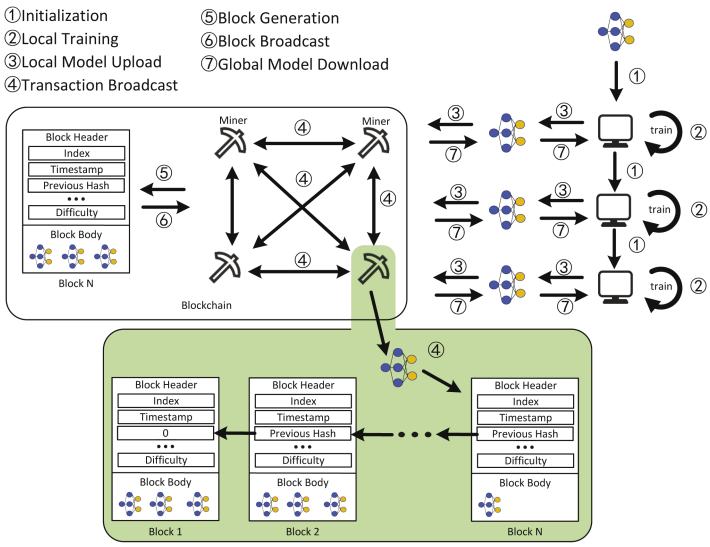


Fig. 5. Overall structure and workflow of the blockchain-empowered federated learning system.

However, the authors did not provide solutions such as model compression based on the weak communication capabilities of drones. In [59], Liu et al. proposed a framework for training vehicle intrusion model. The blockchain is maintained by roadside units and stores and shares the global models for the BC-FL system. After receiving the global model, the vehicle uses the data collected by itself to train the model and upload it to the connected roadside unit nodes. The consensus mechanism in place combines PoW and PoA, with the roadside node that has achieved the highest accuracy being written into the block to encourage the training of high-precision models. PoA can significantly shorten block generation time and improve system efficiency. However, the experiment was ideal and did not consider issues such as signal stability.

Workflow of BC-FL Systems: The overall workflow of the BC-FL system is illustrated in Fig. 5. Different BC-FL systems may be adjusted according to specific circumstances. The steps are explained as follows:

1. Initialization: Each client initializes the environment based on prior negotiation, including model parameter, and training parameter. The blockchain can assist clients in negotiation by storing initialization parameters on the chain and using smart contracts. 2.Local model training: Each client trains the global model using their local dataset. 3.Local model upload: Clients upload training-related data and local models to the blockchain system. To alleviate storage pressure on the blockchain, clients may upload only model-related information rather than the entire model, as detailed in Section IV-C. 4.Transaction broadcast: Upon receiving the transaction, blockchain nodes broadcast it within the system for cross-validation. The nodes inspect the transaction content (e.g., model) based on pre-defined rules. If no issues are found, the transaction is added to the local transaction pool. 5.Block generation: Blockchain nodes select the node with the right to generate blocks for the current round based on the consensus protocol. This node aggregates the local models to generate the global model, compiles relevant model

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

and training information, and creates a block. 6.Block Broadcast: The blockchain system broadcasts the newly generated block. Upon receiving it, validation nodes verify the block according to specific rules. If the majority of nodes validate the block, it is added to their locally maintained blockchain, achieving consensus across the network. 7.Global Model Download: Clients download the latest global model from the blockchain system. 8.End condition judgment: Based on pre-negotiated rules, the FL process evaluates whether it has reached the end condition. If not, the process returns to Step 2 to continue training.

B. Reputation Evalutation Mechanism

FL is a collaborative approach to training a shared model that requires the participation of multiple clients with local data. However, clients may have varying motivations and behaviors, such as seeking rewards for their assistance, hoping to obtain a trained model, or attempting to benefit from the global model without contributing to the training process. In some cases, clients may even have malicious intentions, seeking to undermine the effectiveness of FL due to conflicts of interest in reality or other factors. Compared to traditional distributed learning methods, FL prioritizes user data privacy, which means that the parameter server has limited access to information about the local environment of each client. Therefore, it is essential for the FL task publisher to implement a reputation management mechanism that can assist in managing, rewarding, or punishing FL clients based on their contributions and behavior.

Several studies have proposed the use of some reputation management mechanisms in a centralized way on the parameter server [57], [80]. While this approach can serve as a foundation for client management, reward and punishment schemes, its lack of transparency remains a concern. Data owners who contribute to the training process may worry about potential inaccuracies in the parameter server's reputation calculations, while those seeking to obtain a trained model may be concerned that the parameter server could intentionally manipulate reputations to undermine FL models. Given the importance of attracting high-quality data owners to ensure optimal FL model performance, the transparent reputation management mechanism is particularly well-suited for FL systems. Additionally, a trustworthy parameter server aims to calculate reputation in a transparent manner to discourage malicious nodes. To address these concerns, the BC-FL system leverages blockchain technology to ensure the transparency and credibility of the reputation management mechanism.

After conducting our analysis, we have identified two crucial functions that blockchain can perform within the reputation management mechanism.

1. The blockchain acts as a reliable third-party ledger in the BC-FL system to document crucial information regarding each node's reputation, including but not limited to its reputation value [57], [81], [82] and various calculation bases [59], [63], [83].
2. In the BC-FL system, the reputation computation process can be deployed on the blockchain through a specialized reputation calculation smart contract [63], [83], [84]. This approach serves to ensure both transparency and automation throughout the entire computation process, thereby guaranteeing dependable and consistent outcomes.

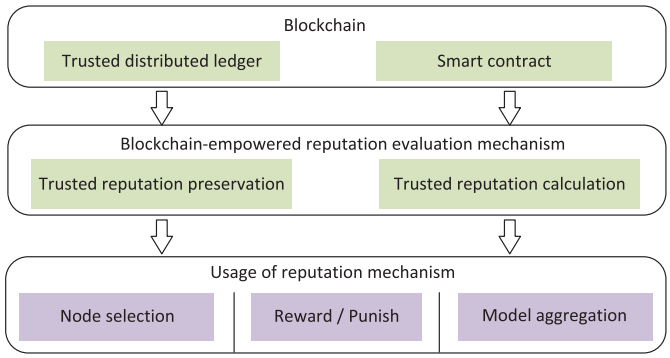


Fig. 6. Reputation management mechanisms based on blockchain. Blockchain is commonly utilized as a reliable distributed ledger or transparent smart contract platform for reputation management mechanisms. This allows the system to store clients' reputation value and the reputation calculation basis on the blockchain, or use smart contracts to compute the reputation in a transparent way. The primary function of reputation management mechanisms is to facilitate node selection, model aggregation, and incentivization.

The reputation management mechanism based on blockchain in BC-FL is illustrated in Fig. 6.

Various papers adopt distinct approaches in calculating the reputation of BC-FL clients. Some calculate reputation values solely on the basis of local model test accuracy, while others take into account evaluations from other clients or factor in the interaction effect between clients and the blockchain system. Moreover, researchers have leveraged clients' reputations in various ways. For instance, some deploy reputation as a criterion for selecting participating clients, whereas others utilize it to ascertain the weight assigned to global model aggregation. Additionally, there are those who offer incentives and penalties to clients based on their respective reputations. We present a comprehensive analysis of BC-FL systems that utilize blockchain technology to establish transparent reputation management mechanisms. Table II summarizes the key attributes of these systems.

Kang et al. proposed a reputation value calculation method based on multi-weight subjective logic, allowing each evaluator to form a unique reputation assessment [81], [83]. Specifically, an evaluator i directly evaluates a target a, resulting in a direct reputation value $V_{i,a}$. At the same time, i also considers the evaluations of a made by other evaluators j, denoted as $V_{j,a}$ (indirect reputation). The similarity $S_{i,j,a}$ between $V_{i,a}$ and $V_{j,a}$ is calculated using a modified cosine similarity measure. This similarity $S_{i,j,a}$ is then used as a weight to combine the direct and indirect reputation values into an integrated reputation value according to a specific formula. This method incorporates evaluations from multiple entities within the system, theoretically providing accurate and comprehensive reputation assessments. However, the introduction of multiple factors makes the calculation process more complex, potentially increasing system latency and resource consumption. Additionally, the method involves numerous hyperparameters, which may affect its feasibility and efficiency in large-scale deployments.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

TABLE II
BLOCKCHAIN-BASED REPUTATION MECHANISM IN BC-FL SYSTEMS

Ref.	Reputation Source			Blockchain Usage		Reputation Usage		
	Global Model	Other Workers	Blockchain	Usage 1	Usage 2	Model Aggregation	Node Selection	Reward or Punishment
Chen [85]	✓	✓	-	-	-	-	✓	-
Gao [86]	✓	-	-	✓	✓	-	✓	✓
Guo [57]	✓	-	-	-	-	-	✓	-
Haddaji[82]	✓	-	✓	✓	-	-	✓	-
He [80]	✓	-	-	-	-	-	✓	-
Kang [83]	✓	-	-	✓	-	-	✓	-
Liu [59]	✓	-	-	✓	-	-	✓	✓
Qi [63]	✓	✓	-	✓	✓	✓	-	✓
Qiu [84]	✓	-	✓	✓	✓	✓	-	-
Rahman [87]	✓	-	-	✓	-	-	✓	-
Xu [68]	✓	-	-	✓	✓	✓	✓	✓
Zhao [88]	✓	-	-	✓	-	-	✓	✓
Wang [72]	✓	-	✓	✓	-	✓	-	-
Lin [89]	✓	-	✓	✓	-	-	✓	-
Fu [90]	✓	-	-	✓	-	-	-	✓
Wahrstatter [77]	✓	-	-	✓	✓	-	-	✓

In [63], Qi et al. proposed a novel reputation evaluation mechanism for multi-model aggregators in FL. Each model aggregator has its test dataset, and the reputation of each participating client is calculated separately by each aggregator. The winning aggregator is selected based on a set of rules, and the winning aggregator updates the client's reputation value to the blockchain. The model aggregators calculate the client's reputation in two steps. In the first step, each model aggregator uses a fair-value game [91] to test the quality of the local model with its test dataset. When the result of a formula containing model test accuracy reaches a certain threshold, the corresponding reputation update is activated. In the second step, the model aggregator synthesizes the results given by other aggregators on the network to obtain the indirect reputation value of the node. Finally, the reputation evaluation value of the modified model aggregator for the node in this round is obtained from the results of the first and second steps. This approach ensures fairness in reputation evaluation across different aggregators and improves the accuracy of the final reputation value. Additionally, authors introduced a grouping mechanism to handle the complexity of large participant numbers, which improves upon Kang et al.'s method [81]. However, the grouping basis is limited, potentially leading to issues like data homogeneity within groups and reduced learning effectiveness.

In [86], Gao et al. designed a time-decaying subjective logic model (SLM) algorithm to measure the client's reputation and a lightweight approach based on gradient similarity to measure client contribution. The final task publisher determines the client's reward share by multiplying the contribution and reputation metrics. They used reputation metrics to measure client reliability and select clients with high reputations to ensure high system stability, which enables BC-FL to work stably in unreliable environments. However, this method does not fully consider the impact of data heterogeneity and may harm honest clients.

C. Incentive Evaluation Mechanism

In FL systems, clients not only need to contribute local data but also consume significant amounts of computing resources and network bandwidth [92], [93]. Without tangible incentives, it may be difficult to attract enough clients to participate in the FL systems. Therefore, introducing an incentive mechanism in FL systems is critical. The introduction of incentives can help incentivize clients to join the FL systems and contribute their valuable data. Adequate participation is crucial for FL to train accurate models with good generalization. Additionally, incorporating incentives can increase clients' engagement and motivation, leading to contributing better data and participation in more training epochs [94]. Furthermore, the incentive mechanism can help achieve fairness in FL systems by rewarding clients based on their data quality and computing power.

A transparent and open incentive mechanism is crucial for attracting clients to participate in federated learning. As it involves vital interests, each client hopes to supervise the calculation of rewards. The BC-FL system utilizes the blockchain to provide a transparent and open incentive mechanism. The blockchain is a decentralized ledger that is maintained on each participating node, requiring the joint efforts of blockchain nodes instead of a centralized organization. This architecture ensures transparency and openness and facilitates tracking and auditing of data necessary for calculating incentives, thereby establishing clients' trust in the incentive results. Furthermore, the incentive algorithm can be written as a smart contract and deployed on the blockchain for automatic incentive calculation and distribution, further strengthening clients' trust in the incentive results. The transparent and open incentive mechanism provided by the blockchain can help to attract more clients to participate in the FL process, contributing high-quality data and computing resources. Consequently, it promotes the accuracy and generalization of the trained model and enhances the efficiency of the BC-FL system.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

TABLE III
STATISTICS OF BLOCKCHAIN BASED INCENTIVE MECHANISM IN BC-FL SYSTEMS

Ref.	Incentives	Incentive Basis	Smart Contract
[54]	Manufacturer's discount	Reputation	✓
[95]	Token	Model accuracy	✓
[86]	Token	Model accuracy, reputation	✓
[96]	Token	Model accuracy	-
[59]	Reputation	Model accuracy	-
[11]	Ethereum	Model accuracy	✓
[97]	Financial incentive	Model accuracy	✓
[98]	Not mentioned	Data size	-
[99]	Not mentioned	Computing power, local data	-
[66]	Token	Reputation	✓
[100]	Reputation, revenue	Reputation, shaply values, and model aggregration	✓
[101]	Token	Model accuracy, block mining	-
[68]	Token	Model accuracy, training time	✓
[102]	Not mentioned	Model accuracy, consensus Participation	-
[103]	Token	Training speed, computing power, and feature extractors sharing	✓
[104]	Ethereum	Model accuracy, data size	✓
[72]	Token	Model accuracy, block mining	-
[105]	Token	Model accuracy, node behavior	✓
[77]	Ethereum	Model accuracy, reputation	✓

We focus on BC-FL systems that provide transparent and open incentive mechanisms based on the blockchain. We believe that understanding this incentive mechanism requires consideration of three aspects: incentive basis, incentives, and incentive algorithms. The settings of these aspects should be tailored to the specific FL tasks. Table III outlines several prominent BC-FL systems developed in recent years.

The incentive basis refers to the criteria that the system uses to reward clients, which may include factors such as node reputation, data quality and quantity, and learning behavior. For instance, Qu et al. rewarded the clients based on the amount of data they contributed [98], but this approach may not accurately reflect the overall contribution of a client to the global model. Factors such as data quality and participation frequency can also significantly impact the effectiveness of the training process. In contrast, Li et al. focused solely on model accuracy as the basis for awarding nodes, as it is verifiable and reflects their contribution [96]. This method does not fully consider the enthusiasm of the participants. Meanwhile, Gao et al. argued that rewards should be based on both model accuracy and node reputation, as this incentivizes continued contributions to the global model [86]. In addition, to compensate the data owner, Zhang et al. considered the energy consumption of the data owner during training and incorporated this factor into the calculation of rewards [103].

Incentives refer to the rewards that clients receive in a system, and they can take various forms such as economic items, tokens, and reputation. Economic items provide monetary benefits to data owners, such as cryptocurrencies like Bitcoin or Ethereum. Tokens, on the other hand, are generated by the BC-FL system and can be used to purchase services within the system, including trained models or tasks for model training. The circulation of tokens promotes a self-sustaining ecosystem within the system that encourages participants to contribute and collaborate. In [86], [95], [101], researchers have utilized tokens within their proposed BC-FL systems as rewards Liu et al. used Ethereum as a reward for training, providing real-world economic incentives [11]. In addition to cryptocurrency rewards, Abdel et al. proposed a BC-FL system for the Industrial Internet of Things that offers clients maintenance services or discounts on products from manufacturers as incentives [54].

The incentive algorithm determines the specific implementation method of the incentive mechanism. Generally, the algorithm involves quantifying each incentive basis and inputting it as a variable into the reward function, which yields the corresponding reward value. For instance, Xu et al. proposed a reward formula balancing model accuracy and training time [68]. The reward for each client is calculated based on their accuracy and training time during training. A tunable parameter, α, allows prioritization between accuracy and time depending on the task requirements. This flexibility enables adaptation to tasks with varying sensitivity to these factors.

D. Security Enhancement

The BC-FL system achieves the establishment of a trustworthy relationship in the system through blockchain technology. As a distributed database, blockchain aligns with the distributed nature of FL. With certain consensus mechanisms, the blockchain can still maintain the consistency and correctness of the system even in the presence of malicious clients. Therefore, the robustness of blockchain against malicious nodes makes it well-suited for an environment where malicious nodes could exist in the FL system. Furthermore, due to the robustness of the blockchain, the BC-FL system allows for the storage of vulnerable data in the blockchain, enhancing the security of the entire system. The security issues in the BC-FL system is illustrated in Fig. 7.

To explicate the specific security properties of blockchain necessary for implementation in a BC-FL system, we conducted an extensive study of representative BC-FL systems from recent years. The results of this research are presented in Table IV.

Transparency: Transparency is one of the key features of the blockchain. All the information stored on the blockchain is accessible to full nodes, while light nodes can query certain information by sending requests to the full nodes. In the BC-FL system, transparency refers to the transparent operation of algorithms and the disclosure of data. This includes but is not limited to, the parameter aggregation operation, the reputation of each node, and the reward operation of the system. The transparent nature of blockchain is derived from the distributed maintenance of the blockchain across all nodes in the network, with each node maintaining a local copy of the blockchain ledger.

Auditability: Auditability is a significant feature of blockchain systems, enabling the tracing and analysis of data using specialized algorithms. In the BC-FL system, auditability becomes particularly valuable when specific circumstances arise, such as ineffective model training or the need to review client operations. The recorded data on the blockchain - including local gradients

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
STATISTICS ON THE SECURITY PURPOSE OF INTRODUCING BLOCKCHAIN IN BC-FL SYSTEMS

Ref.	Transparency	Auditability	Anti-malicious nodes	Traceability	Immutability	Anti-single point of failure
[106]	-	✓	✓	-	✓	-
[107]	-	-	✓	-	-	-
[108]	-	-	✓	-	✓	-
[109]	-	✓	-	-	✓	-
[96]	-	✓	✓	✓	-	✓
[110]	-	-	✓	-	-	-
[111]	-	-	-	-	-	✓
[112]	✓	-	✓	✓	-	✓
[113]	-	✓	-	-	✓	-
[114]	-	✓	✓	-	✓	-
[68]	✓	-	-	-	✓	✓
[103]	✓	-	-	-	✓	-
[70]	-	-	-	✓	✓	✓
[115]	-	-	-	-	✓	-
[73]	-	✓	✓	-	✓	✓
[89]	-	-	✓	-	-	✓
[75]	-	-	✓	-	-	-
[76]	-	-	✓	-	-	✓
[90]	✓	-	-	-	✓	✓
[105]	-	✓	✓	-	✓	✓
[116]	-	-	✓	✓	✓	-

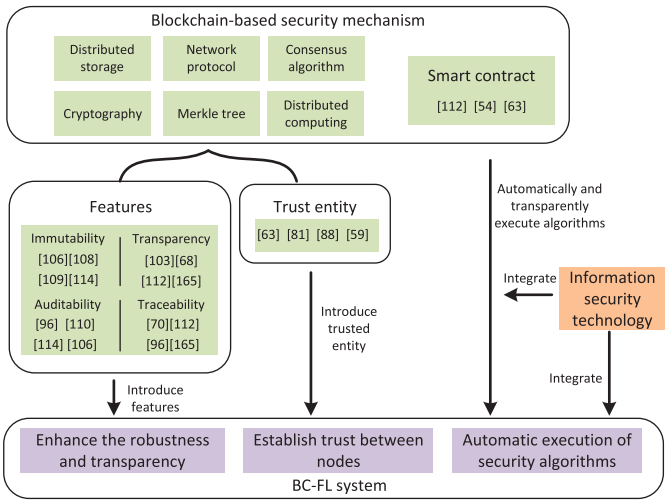


Fig. 7. Security provided by blockchain for the BC-FL system. By employing appropriate techniques, blockchain can impart its security features (e.g. immutability and traceability) to the FL system. Moreover, in a partially trusted FL environment, blockchain can act as a reliable entity to foster trust relationships. Furthermore, deploying security-enhancing algorithms on the blockchain via smart contracts can further enhance the security of the BC-FL system.

- can be extracted for detailed analysis. By analyzing previously recorded information on the blockchain, such as local gradients, nodes can be penalized for producing undesirable outcomes.

Anti-malicious nodes: In blockchain systems, malicious nodes can take on various forms, including those that propagate false blocks or launch attacks against the system. Byzantine robust consensus algorithms can be used to mitigate these types of malicious behavior. In the BC-FL system, malicious nodes are those that can undermine the effectiveness of the system, such as through poisoning attacks or privacy violations. To address these issues, specific consensus algorithms can be designed to thwart malicious activity, or security techniques can be incorporated into the system via smart contracts. Anti-malicious nodes and auditability both play a role in dealing with malicious nodes, but the former aims to prevent the impact of malicious nodes in real time, while the latter focuses on identifying the source of the attack after the fact.

Traceability: The blockchain system inherently preserves all state changes since its genesis block. When tracing back to a previous state, the system can be readily restored to a specific point in history. In the BC-FL system, traceability refers to the ability to restore a previously trained model or parameters saved by the current work in case of severe damage or loss due to central server failure.

Immutability: The immutable nature of the blockchain can be attributed to the sound design that underlies its consensus algorithm. Each full node in a blockchain network maintains a local copy of the ledger, which ensures that malicious nodes are unable to dictate terms to other nodes unless they comply with the consensus algorithm. Any attempts to tamper with the local copy by modifying incorrect blockchains will result in the creation of new blocks that cannot be recognized by other honest nodes. Therefore, as long as the majority of computing power is held by honest nodes, the blockchain remains immutable. In the BC-FL system, critical information such as client reputation and model hash values can be securely stored on the blockchain to ensure the accuracy of this data.

Anti-single point of failure: The term "single point of failure" refers to a scenario where a sole parameter server becomes the bottleneck for FL security, rendering the entire system inoperable if it fails due to an attack or power outage, among other reasons. To tackle this problem, the BC-FL system replaces the role of the parameter server with blockchain technology. As discussed in Section III, the issue of single point of a failure is elaborated upon.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

Transparency, Auditability, Traceability, and Immutability are inherently interconnected, they enhance the security and robustness of the BC-FL system. Among these attributes, immutability and transparency can be considered fundamental for the following reasons. Immutability is foundational as it underpins trust in the blockchain system. Without immutability, the reliability of auditability and traceability would be compromised, as historical data could be altered, leading to inaccurate analyses or the inability to restore accurate prior states. Transparency is essential for enabling auditability and traceability. Without access to the recorded data (provided by transparency), audit processes cannot inspect historical operations, and the system cannot reconstruct past states.

To provide a comprehensive understanding of the utilization of blockchain technology in enhancing the security of the BC-FL system, we will discuss prominent literature in this field.

In [106], Sana et al. regarded the blockchain as an immutable, decentralized, and reliable entity, which they incorporate into their proposed BC-FL framework called blockchain-based privacy-preserving FL (BC-based PPFL). The utilization of blockchain provides auditability, thereby enhancing the resilience of BC-based PPFL against malicious clients. Specifically, the assumption of semi-honest clients in the universal FL system is further elevated to the assumption of malicious clients. However, the authors' experiment is relatively simple and fails to fully illustrate the tolerance of malicious nodes. Qi et al. introduced the adoption of smart contracts to handle FL tasks [63]. These smart contracts encompass various functions such as task initiation, member selection, FL execution, reputation evaluation, reward distribution, and query processing. This approach can greatly improve the credibility and reliability of FL, but the extensive use of smart contracts consumes a lot of computing resources, so it may be limited in practical applications. In [88], Zhao et al. combined Multi-Krum with reputation mechanisms as well as aggregation mechanisms to rule out malicious gradients and penalize malicious clients. However, the MNIST dataset used in the experiment is small, and its effect needs to be verified on more datasets. In [54], Qi et al. proposed a smart contract called Hunter Contract (HC) to prevent malicious clients. HC acts as a hunter by randomly selecting a client and verifying whether the gradient uploaded by that client causes a decline in the global model accuracy. If the reduction surpasses a predefined threshold, the client is classified as malicious. This method requires careful consideration of the threshold, otherwise it may be misjudged.

In a blockchain system, individual nodes follow the consensus mechanism to ensure the consistency, validity, and accuracy of the data. In a BC-FL system, the data or training results of the FL process are stored on the blockchain, and the blockchain's consensus mechanism can be used to verify the content of the FL. Consequently, some researchers have improved the security of FL by adjusting the blockchain's consensus mechanism.

In [96], Li et al. proposed a Byzantine-resistant consensus mechanism named Proof of Accuracy, which serves to identify models of poor quality. This consensus algorithm takes into consideration not only the exclusion of local models that are deemed too poor for aggregation into the global model but also the potential for a local model with a high loss value to aid the global model in escaping local optimal solutions. To fulfil this requirement, the consensus algorithm employs two critical thresholds: the accuracy oscillation threshold (AOT) and the accuracy deviation threshold (ADT). The AOT determines the maximum acceptable accuracy reduction permitted by the accepted model, while the ADT determines the maximum absolute difference in accuracy among different client models. These two thresholds are subject to dynamic adjustments as the algorithm progresses. The AOT and ADT methods are highly dependent on the choice of test sets, which need to be extensive and representative when constructing test datasets. In [84], Qiu et al. increased the security of the BC-FL system through the introduction of a novel consensus protocol called Proof of Learning (PoL). In contrast to PoW, PoL requires nodes to compete for the privilege of accounting rights through calculation by training a FL model, where the node with the smallest loss value adds a new block as the winner. Other clients aggregate the winner's local model based on the reputation value against the winning node after verifying the authenticity of the newly added block. Ouyang et al. utilized smart contracts to authenticate participating nodes and prevent malicious nodes from participating [73]. However, the authors mainly validate the effectiveness of the method based on theoretical models and simulation experiments, without validating it on multiple datasets.

Furthermore, the security provided by blockchain plays a critical role in constructing a robust ecosystem for FL, wherein transparency enables participants to conduct real-time supervision, thereby establishing trust and attracting diverse stakeholders to engage in the ecosystem. Smart contracts autonomously execute incentive algorithms and reputation management mechanisms, ensuring fairness and further enhancing client trust and participation enthusiasm. In [117], blockchain precisely records the contributions of each model provider, equitably distributing rewards based on predictive accuracy while simultaneously improving system security. This approach incentivizes entities with high-quality model resources to actively participate, consequently expanding the diversity of the model repository.

As mentioned earlier, certain security technologies from the field of information security have been considered for use in the BC-FL system to enhance their security. While not directly related to the security of the BC-FL system, smart contracts can serve as a platform for running certain algorithms. Hence, we will provide a brief overview of this topic. To safeguard client privacy, the utilization of homomorphic encryption and differential privacy algorithms [5] is common, and researchers have developed advanced algorithms building upon these fundamental techniques. We have organized this material in Section IV-B.

IV. Challenges and Solutions in BC-FL Systems

While blockchain can indeed enhance the capabilities of the FL systems and mitigate certain limitations, it is imperative to duly recognize and confront the accompanying drawbacks. In this section, we will delve into the principal challenges entailed in the integration of blockchain into FL and the corresponding

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

TABLE V
COMPARISON BETWEEN DIFFERENT SOLUTIONS FOR EACH EFFICIENCY CHALLENGE

Ref.	Solutions	Detailed methods	Dataset	Evaluation indicators
Cao [121]	Blockchain topology	DAG blockchain	MNIST	Accuracy, loss, iteration delay
Cheng [107]	Consensus algorithm, blockchain topology	Two-layer blockchain, Raft, PBFT	-	Latency reduction
Feng [108]	Consensus algorithm, blockchain topology	Two-layer blockchain, sharding	MNIST	Accuracy, time cost
Hieu [122]	RL	DRL	-	Energy consumption, latency, total payment
Li [123]	Consensus algorithm	committee consensus mechanism	FEMNIST	accuracy, communication overhead
Lu [60]	RL, blockchain	DAG blockchain, DRL	-	Accuracy, time cost, agent reward, cumulative cost
Nguyen [61]	Consensus algorithm	Proof of reputation	DarkCOVID, ChestCOVID	Running latency, block verification latency, Accuracy, Loss
Nguyen [62]	RL	DRL, A2C	SVHN, Fashion-MNIST	Accuracy, agent reward, latency, Loss
Qi [64]	Consensus algorithm	Modified PBFT	Diabetes Breast Cancer	Accuracy, time cost, gas cost
Qu [99]	Consensus algorithm	Proof of federalism	CIFAR-10	Accuracy
Xu [102]	Consensus algorithm, blockchain topology	Two-layer blockchain, proof of credit, efficient BFT	MNIST	Latency, communication overhead, data throughput
Zhao [71]	RL	Federated DDQL	-	Agent reward, latency
Wang [115]	Blockchain topology	Two-layer blockchain	TSP, FMNIST	Accuracy, energy consumption, learning utility
Yuan [74]	Consensus algorithm, blockchain topology	Blockchain sharding, DAG-based mainchain	MNIST, Penn Treebank	Accuracy, training Latency, testing perplexity
Lin [89]	Blockchain topology, RL	Blockchain sharding, DRL-based sharding	MNIST, KMNIST, FMNIST, CIFAR-10	Accuracy, agent reward, reputation of nodes

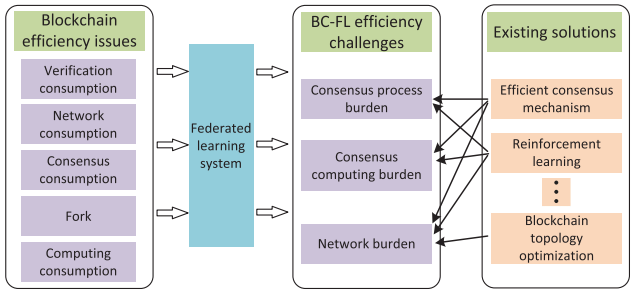


Fig. 8. Efficiency challenges and related solutions in the BC-FL systems. The efficiency of blockchain is susceptible to factors such as network and computing overhead. Consequently, BC-FL systems potentially lead to a decrease in overall efficiency. In response to thus challenges, multiple strategies are contributed to mitigate the reduction in system efficiency. These methods include but not limited to, efficient consensus mechanisms, reinforcement learning, and optimized blockchain topologies.

solutions, which can be broadly classified into three key aspects: efficiency, security, and storage.

A. *Efficiency Challenges and Solutions*

The processing capacity of blockchain systems is inherently limited. For instance, Bitcoin can only handle seven transactions per second [118]. In contrast, modern centralized payment systems can process thousands of transactions per second [119]. Fig. 8 illustrates the efficiency challenges faced by blockchain in BC-FL systems. Unlike centralized systems, blockchain systems necessitate additional steps such as verification, communication, and network-wide consensus to maintain normal operations, which reduces the efficiency when integrating blockchain with FL systems. For instance, with 200 nodes, Biscotti takes approximately 13 times longer than FL to achieve the same accuracy [120]. Biscotti's consensus protocol is based on PoS. According to experiments by Wang et al. BC-FL systems based on PoW require roughly 10 times more runtime than those based on PoS [25]. This implies that, without optimization, using PoW-based BC-FL systems results in a runtime overhead approximately 100 times that of traditional FL systems. Current BC-FL systems address these efficiency issues through various methods, including efficient consensus mechanisms, reinforcement learning, and optimized blockchain topologies. A summary of the pertinent literature is provided in Table V.

1) Efficient Consensus Algorithms: The PoW consensus protocol provides robust resistance against Sybil attacks on the public chain, ensuring a strong defense against malicious nodes. However, a primary drawback of the PoW mechanism lies in its requirement for a block generation rate that is slower than the rate of block propagation across the network, aimed at minimizing the risk of a blockchain fork. Current research reveals relatively modest transactions per second (TPS) for both PoW and PoS consensus protocols [124]. Typically, the PoW protocol achieves TPS figures below 100, while the PoS protocol reaches less than 1000 TPS. In actuality, the TPS tends to be even lower; for instance, Bitcoin operates at a mere 7 TPS [125]. Additionally, the competitive nature among miners vying for block mining rewards escalates energy consumption. Moreover, suboptimal network conditions of edge devices heighten the likelihood of forks. In the BC-FL systems, underpinned by a partially decentralized architecture, the scenario improves to some extent. Nonetheless, achieving consensus across the entire network still requires considerable time, impeding the speed of model aggregation. Consequently, numerous researchers are dedicating their efforts to the development of efficient consensus protocols that can enhance the overall operational velocity of the BC-FL system, all the while aligning with the requirements of the federated learning process.

In [61], Nguyen et al. contended that certain established consensus algorithms introduce substantial communication overhead while striving for consensus. For example, DPoS necessitates that each blockchain node communicates with

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

a minimum of half the nodes within the BC-FL system for confirmation, leading to redundant validations among these nodes. To tackle this challenge, they designed a streamlined consensus mechanism known as Proof of Reputation (PoR). Within the POR algorithm, every blockchain node is permitted to validate with just a single other node during the consensus process, resulting in a significant reduction in validation delays. However, an environment with highly heterogeneous data can easily lead to deviations in PoR's reputation evaluation. In [99], Qu et al. introduces a Proof-of-Federalism (PoF) consensus algorithm, which builds upon the foundation of PoW. PoF leverages the training of FL tasks as a viable alternative to the challenge of discovering a fitting nonce in PoW, effectively sidestepping the computational resources typically expended during the consensus calculation process. Before each training round commences, intelligent contracts sift through unfavorable local model parameters and cherry-pick local models that lend themselves well to global aggregation. During cross validation, each node singles out the most optimal set of global models. Upon reaching a predetermined time threshold, the participant who boasts the highest number of selected global models emerges as the victorious contender. However, the authors only uses CIFAR-10 to evaluate the method, which cannot fully illustrate the superiority of the method.

In [102], Xu et al. proposed a lightweight blockchain network for FL systems called micro-chain to address the issues of low transaction throughput and poor scalability. Participants in FL are divided into multiple small-scale micro chains, each of which is unified through an advanced inter-chain network using Byzantine fault-tolerant consensus protocols. Within each micro-chain, block consistency is achieved using the Proof of Credit (PoC) algorithm, where committee members are responsible for generating new blocks. Then, a new committee is randomly selected at the end of each dynasty round. Ledger consensus is achieved using the Vote-based Chain Finality (VCF) protocol, where committee member nodes vote to select the preferred branch in case of network forks. In [123], Li et al. introduced an innovative committee consensus mechanism aimed at significantly reducing the required consensus computation. The proposed mechanism selects multiple clients as committee nodes in each training round, utilizing the data on these committee nodes as the validation set. The final scores for each trained client are then determined by taking the median of the scores of these clients. These scores are subsequently used to perform global model aggregation by selecting a specific number of clients with the highest scores. However, authors store the model on the blockchain, resulting in a lot of storage overhead.

2) Reinforcement Learning: RL is a machine learning algorithm that enables an agent to interact with the environment, learn from its experiences, and take action accordingly. The ultimate objective is to maximize the cumulative reward obtained by the agent over time. The traditional optimization methods are ineffective in the BC-FL system because of the system's complexity, a large number of participants, and their limited computing and communication resources. To address these challenges and achieve better results, RL can be utilized to optimize resource allocation and schedule the resources of each client based on signals received from them. This can potentially reduce system delays and lead to improved performance.

To apply RL in the BC-FL system, there are several fundamental steps to follow. First, the system designer must define the environment based on specific circumstances, such as the parameters of the client and network conditions. This environment can be modelled as a Markov decision process. Second, the agent's action space should be defined, which includes factors such as the energy consumed by the device during training and the block generation difficulty. Third, defining the reward is essential. In general, the reward in the system can be based on overall training delay that encourages the agent to find ways to reduce the system delay effectively. Finally, RL training is performed using a specific algorithm. The agent learns how to optimize resource allocation within the BC-FL system under different environmental scenarios through continuous interaction with the environment.

In [122], Hieu et al. used the deep reinforcement learning method [126] to control the data and energy used for training and block generation in the device. By judiciously allocating resources, they were able to mitigate the system delay and enhance overall system efficiency. In [60], Lu et al. used the Deep Q-learning (DQL) [126] method to facilitate client selection for the FL process. They formulated a joint optimization plan by considering the client's available wireless transmission rate, client computing power (CPU frequency), and the current selection status of clients as the state of the DQL method. The reward function is designed as a weighted sum of the loss function of each node, the computation time, and the communication time. This approach leads to a high level of model accuracy while maintaining a low global system cost. The proposed algorithm design shows promising results in performance evaluation, indicating its potential in real-world applications.

In [71], Zhao et al. proposed a BC-FL system for vehicle networks. The proposed system allows autonomous vehicles (AVs) to offload part of their computing tasks to edge servers (ESs), effectively reducing local computation latency, communication latency, and blockchain consensus latency. To achieve this, the authors employed a federated duel deep Q-learning (DDQL) algorithm [127] and deployed it to each AV to enable them to take action according to the changing external environment. The state space of the proposed DDQL includes wireless channel conditions, data set quality, and packet error rate, where AVs select offload strategy, wireless channel, and CPU-cycle frequency based on the DDQL algorithm.

In [62], Nguyen et al. applied the DRL method based on a parameterized advantage actor-critic (A2C) algorithm [128] to a multi-server edge computing scenario to reduce the overall system latency. Their proposed hybrid discrete-continuous action DRL algorithm takes into account various factors such as data size, channel state, broadband state, computation state, and hash power to determine whether an edge node should perform computation offloading. In case of offloading, the agent needs to decide on the corresponding channel selection, power allocation and other transmission necessary parameters. In case of non-offloading, the agent needs to decide on the necessary parameters for training such as the hash power allocation for local computation. Unlike existing purely discrete or purely continuous action

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

DRL algorithms, the authors proposed a hybrid model where resource allocation is continuous, while the offloading decision is discrete, leading to improved training performance.

3) Optimized Blockchain Topology: The topological structure of a blockchain system is a crucial factor that impacts information transmission and significantly influences the system's efficiency and scalability. Modifying the topology of the blockchain can potentially improve its efficiency, which has been demonstrated in some papers in the BC-FL systems [129], [130]. The topology of a blockchain system includes the physical and logical topology, both of which can affect the system's efficiency.

Improving the physical topology involves considering the node layout, physical location, and network topology. For instance, positioning relevant nodes near the data source can reduce the network delay, altering node connections' topology can enhance network transmission efficacy, and using edge computing can reduce the computing burden of clients.

The logical topology of the blockchain refers to how transactions and blocks are verified and added, and it can impact the processing speed and scalability of the system. The initially proposed blockchain has a linear chain structure, and data records are processed serially. The Directed Acyclic Graph (DAG) blockchain adopts the organization method of a directed acyclic graph, where multiple preceding blocks point to one block. Compared with linear chain blockchains, it features high concurrency and weak synchronization. Cao et al. utilized DAG-based blockchains to reduce resource consumption and address the issues of device asynchrony and anomaly detection [121]. However, there are problems such as decreased verifiability and difficulty in model convergence. To solve these problems, Zhang et al. proposed the TGFL, a BC-FL system based on the tree-graph blockchain, which supports verifiable and semi-asynchronous training [131]. TGFL can improve the efficiency of the system while ensuring model convergence.

In addition, there are some improvements that involve both the physical and logical topology of the blockchain. One such improvement is the deployment of a two-layered blockchain architecture, which comprises two relatively autonomous blockchains – the main-chain and the sub-chain. The sub-chain is responsible for interfacing with peripheral devices and executing swift consensus algorithms. Meanwhile, a subset of nodes within the sub-chain are nominated to constitute the main-chain. Typically, the main-chain utilizes Byzantine fault-tolerant consensus algorithms to ensure the security of the system.

In [107], Cheng et al. proposed a BC-FL system based on a two-layer blockchain architecture. The lower-layer blockchain is responsible for connecting devices to achieve strong consistency and a high consensus rate. Within a short period of time, the lower-layer blockchain needs to reach a consensus while only considering the problems of equipment failure and omission. To this end, the Raft protocol is employed, which is more efficient despite lacking Byzantine fault tolerance. The upper-layer blockchain connects various lower-layer blockchains and is designed to prevent malicious nodes and resolve Byzantine faults. Thus, the PBFT algorithm is employed, which can effectively resist Byzantine attacks but requires a longer time frame for consensus. The upper-layer blockchain's nodes are super nodes with robust computing power selected from the lower-layer blockchain. The upper-layer blockchain has fewer nodes, reducing the traffic required for consensus. At the same time, these nodes have stronger computing power, and these two reasons together increase the consensus speed of the system. However, the assumptions in the author's experiment are relatively ideal, and problems such as actual network fluctuations are not considered.

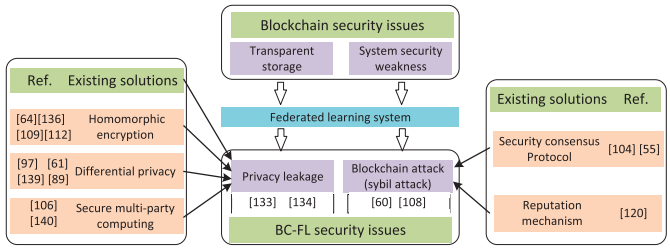


Fig. 9. Security challenges and related solutions in the BC-FL systems. Malicious nodes pose a threat to the blockchain within the BC-FL system through two distinct avenues: privacy leakage and consensus mechanisms. The former capitalizes on the blockchain's data transparency to breach access to model information stored within it, whereas the latter employs attacks via the straightforward consensus mechanism inherent in the BC-FL system. In response to these challenges, contemporary solutions are predominantly centered around the development of diverse privacy protection algorithms and the implementation of exceptionally secure consensus mechanisms. .

B. Secure Challenges and Solutions

Integrating blockchain into FL systems holds the potential to significantly bolster system security. However, the successful execution of such integration in BC-FL systems hinges greatly upon the scrupulous deliberation of system designers and the implementation of effective combination strategies. Inadequate integration of blockchain may give rise to supplementary predicaments. The security challenges and related solutions in the BC-FL systems are evidenced in Fig. 9.

As shown in Fig. 9, the transparent nature of blockchain data raises concerns about storing sensitive information, potentially leading to violations of privacy. Additionally, extant attack methods targeting blockchain systems, such as Sybil attacks [132], have the capability to compromise the security of the BC-FL system. An examination of recent BC-FL systems has unveiled several instances wherein Sybil attacks and breaches of privacy remain plausible.

1) Privacy Leakage: The immutability and transparency inherent in blockchain play a pivotal role in safeguarding the integrity of a system. Blockchain data can be validated by all clients, and it remains impervious to unauthorized tampering by malicious entities. However, this approach also brings forth a potential vulnerability, as malevolent nodes can effortlessly access sensitive data stored on the blockchain. In BC-FL systems, multiple research endeavors permit clients to store local models or gradients on the blockchain, along with their retrieval methods [133], [134]. Regrettably, this allowance opens the door for malicious clients to potentially deduce sensitive worker data. To tackle this issue, several scholars suggested the implementation of diverse cryptographic techniques [56], [135]. These techniques serve to fortify the system's privacy protection capabilities while mitigating the potential privacy hazards.

Homomorphic Encryption (HE) represents an encryption technology that facilitates direct computations on encrypted

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

data, empowering aggregators to execute model aggregation operations without necessitating the decryption of local models [64], [109], [112], [136]. In [136], Sun et al. designed a Bresson-Catalano-Pointcheval based homomorphic noise mechanism to secure gradient values and pinpoint malevolent data owners. However, the encryption time of this method is proportional to the number of parameters, which limits the size of the model used. Meanwhile, in [109], Jia et al. seamlessly incorporated the homomorphic encryption scheme Paillier [137] into K-means clustering, distributed random forest, and distributed AdaBoost components in the BC-FL systems. The scheme offers a privacy-preserving solution for client data when employing these machine learning algorithms. In another study [112], Miao et al. harnessed Fully Homomorphic Encryption (FHE) to facilitate secure model aggregation. However, FHE will incur huge time overhead. Concurrently, they leveraged blockchain to ensure the transparency of the aggregation process. In [114], Sun et al. introduced a verification procedure in [104] before local update aggregation to fend off poisoning attacks. They introduced differential privacy noise during the verification process to obfuscate local updates, thereby enhancing privacy. Additionally, in [55], Fang et al. outlined a secure and verifiable local update aggregation scheme, replacing differential privacy technology with the Shamir Secret Sharing technique [138] to ensure the correctness of confidential sharing.

Multiple studies also employed differential privacy to protect the privacy of FL clients [61], [80], [97]. In [97], Ma et al. delved into a differential privacy solution for the BC-FL system, where noise is added to the local data features to uphold local privacy and pseudo-noise sequences are adopted to identify inactive clients. Similarly, in [139], Abadi et al. incorporated tailored noise into the data prior to sharing, effectively obscuring the actual data values while maintaining usability even after noise integration. Within BC-FL systems integrating differential privacy, it is customary for clients to introduce noise to the model prior to uploading the local model, thereby ensuring privacy protection. In [88], Zhao et al. employed differential privacy to safeguard the privacy of individual clients by applying it to the extracted data features of each client. Additionally, Qu's work [99] presented an enhanced differential privacy algorithm built upon generative adversarial networks, offering a means of preserving the privacy of local models.

Secure Multi-Party Computing (SMPC) stands out as another promising avenue for ensuring privacy of BC-FL systems [106], [140]. SMPC represents a versatile cryptographic tool that empowers distributed parties to collaboratively compute diverse functions while withholding their confidential inputs and outputs [141]. Within the BC-FL system incorporating SMPC, every client employs the SMPC protocol to join forces and aggregate the global model. SMPC can be instantiated as a smart contract on the blockchain, with these contracts delineating computation rules and guaranteeing proper protocol execution. In [106], Awan et al. designed a meticulously algorithm that leverages homomorphic encryption and proxy re-encryption grounded in the Paillier encryption algorithm. This technique involves encrypting each local model, thereby preventing the model aggregator from accessing individual models. Nevertheless, upon aggregating the encrypted local models, the aggregator can obtain an unencrypted global model, thus preserving the confidentiality of each client's data.

Several studies explored alternative approaches to address the privacy concerns within the BC-FL system [55], [142], [143]. For instance, in [142], Wei et al. introduced a chameleon hash scheme with a modifiable trapdoor (CHCT) as a countermeasure to potential privacy leaks on the blockchain, effectively creating an adaptable blockchain structure. The CHCT employs trapdoors to generate hash collisions, resulting in identical hash values. When sensitive or erroneous data is identified on the blockchain, clients can utilize CHCT to amend the relevant data. However, strict adherence to a well-defined set of procedures is imperative when modifying the blockchain to safeguard its reputation as a trusted third-party entity. In [55], Fang et al. employed a privacy-preserving strategy to store the gradient's commitment on the blockchain and mapped it to an elliptic curve point. Simultaneously, the gradient is obscured using a Pseudorandom generator-based mask, which can subsequently be removed to restore the accurate global gradient once all local gradients are incorporated. Similarly, [143], Guo et al. presented a blockchain-based obfuscation transmission mechanism, shielding the local models of FL edge nodes from external scrutiny by potential attack devices. The blockchain is initially divided into distinct branches starting from the genesis block, each corresponding to a training device. A hash key block on each branch stores the hash key function published by the server. Qin et al. applied model compression techniques to protect model privacy [144]. However, their work lacks ablation experiments, leaving the impact of model compression techniques on performance unclear.

2) Sybil Attacks: Sybil attacks have garnered extensive attention within the blockchain field, owing to their potential to compromise the integrity and security of blockchains [132]. Thus attacks involve an assailant generating numerous false identities or nodes within the network, affording them the means to manipulate the system's dynamics [145]. Established methods like PoW and PoS have demonstrated some degree of resilience against Sybil attacks [146], [147]. Within the context of the BC-FL system, certain endeavors have adopted lightweight consensus protocols or rapid information transmission methods to bolster system speed, inadvertently rendering them susceptible to Sybil attacks [60], [108], [120]. For instance, in [60], the Raft protocol is harnessed to expedite consensus within the underlying blockchain. However, this approach exposes a vulnerability where an attacker could subvert the leader election process through the creation of fabricated identities. This disruption might impede the proper selection of legitimate leaders or lead the system astray from its intended behavior. In another instance, Feng et al. employed a localized model update chain facilitated by inter-device communication for efficient blockchain information transfer [108]. While inter-device communication offers improved network performance and reduced communication costs, it also presents a vulnerability to Sybil attacks [120]. In the realm of inter-device communication, attackers exploit the creation of multiple spurious identities or devices to gain a foothold in the network, inundating it with counterfeit traffic or acquiring sensitive information.

Another group of research tried to employ various consensus mechanisms to counter Sybil attacks [55], [104], [120],

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

TABLE VI
REAL WORLD APPLICATIONS OF BC-FL SYSTEMS

Ref.	Domain	Model	Dataset	Blockchain's Role	Potential Disadvantage
[153]	Medical imaging	ResNet-34	Fashion-MNIST, CIFAR10	Reputation mechanism provision, security improvement, decentralization	Not validated on medical datasets
[154]	Covid-19 detection	VGG16, DenseNet, etc.	COVID-19 patient dataset	Security improvement, trust building	Lacking ablation experiment
[155]	Medical image analysis	2D U-net	Prostate, Camelyon17	Security improvement, incentive mechanism provision	PoW brings large computational overhead
[156]	Disease prediction	DNN	PBMC transcriptome dataset, X-ray dataset, etc.	Security improvement, Decentralization	-
[157]	Digital twin	DRL model	CIFAR10	Security improsvement, permission control	Using a single dataset, generalization cannot be verified
[56]	Drones	CNN	EMNIST	Decentralization, security improvement, identity authentication	No evaluation on real-world datasets
[158]	Internet of vehicles	-	Simulating dataset	Data integrity and immutability ensurance	Experiment fails to considering communication capability differences
[159]	Drones for disaster response	CNN, MobileNet	EMNIST, Real disaster dataset	Reputation mechanism provision, security improvement	The mobility and energy limitations of drones are not fully considered
[160]	Autonomous vehicles	CNN	Traffic-Light data sets	Decentralization, Security improvement, incentive mechanism provision	Using a single dataset
[161]	Traffic prediction	LSTM	DelDOT traffic flow dataset, PeMS Bay dataset	Data integrity and security ensurance	Not consider dynamic traffic environment
[162]	Intelligent transportation	-	UNSW-NB15	reputation mechanism provision, security improvement	Using a single dataset
[163]	Mobile crowdsourcing	CNN	Fashion-MNIST	incentive mechanism provision, security improvement	Not considering poisoning attack
[82]	Internet of vehicles	ANN	NSL-KDD	reputation mechanism provision, security improvement	The impact of frequent node entry and exit in a dynamic network environment is not considered
[66]	Industrial internet of things	NN	Turbofan engine degradation simulation dataset	reputation mechanism provision, fairness and credibility ensurance	Not considering device heterogeneity, single type dataset
[164]	Device failure detection	LR, NN	Air-conditioning systems dataset	reputation mechanism provision, data integrity and auditability ensurance	The experiment was small-scale
[165]	Vessel collision avoidance	ConvLSTM	Generated dataset representing ships mobility	Security, transparency and traceability ensurance	Still has problems of missed detection and inaccurate position

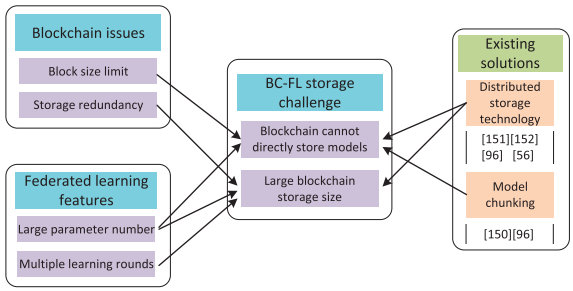


Fig. 10. Storage challenges and related solutions in BC-FL systems.

[148]. For instance, in [104], Zhang et al. utilize a validator committee selection scheme akin to the Algorand consensus algorithm [148], utilizing verifiable random numbers to thwart Sybil attacks. In [55], Fang et al. designed a secure aggregation protocol that directly applies the Algorand consensus algorithm to fend off Sybil and tampering attacks. The protocol uses pairwise random masks to impede Sybil attacks. Shayan et al. [120] introduced a fully decentralized system to effectively mitigate Sybil attacks by judiciously defining reputation levels. They used blockchain and cryptographic primitives to defends against known attacks.

C. Storage Challenges and Solutions

The storage requirements for blockchain systems are inherently cumbersome, as each full node is required to maintain a complete backup of the entire system. This leads to a linear increase in the total storage size with the number of full nodes. In FL, clients transmit their local models to a central server and download the global model. The server is responsible for storing both the local and global models and various FL-related data and parameters, thereby becoming the node with the highest storage demand in the FL system. When enhancing the FL system with blockchain, the BC-FL system must inevitably store diverse information on the blockchain, resulting in significant storage overhead. Furthermore, most blockchain platforms currently impose limitations on transaction or block size. For instance, Bitcoin has a block size limit of 1 MB, and while Ethereum does not have a theoretical block size limit, its gas limit effectively restricts the size of transactions [149]. If the BC-FL system requires direct storage of large volumes of data within blocks, such as model parameters, this could surpass the blockchain system's storage capabilities.

As illustrated in Fig. 10, the storage challenges of the BC-FL system are primarily twofold:

Constrained storage capacity: the limited block size makes storing some data that takes up storage space difficult.

Redundant storage demands: a large amount of training-related data is stored in the blockchain, which brings unnecessary information redundancy and terrible storage challenges to the entire BC-FL system.

As depicted in Fig. 10, the current landscape presents two prevailing strategies to tackle the storage challenges in BC-FL systems. The first approach entails chunking the FL models or data into distinct segments, which are then stored on the blockchain with constrained block size [150]. This methodology necessitates prior negotiation of a serialization plan among nodes. Subsequently, each split data's size is logged as supplementary information within the transaction block. Gradual storage of the data on the BC-FL system is accomplished through the initiation of transactions. This approach incurs significant

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

communication overhead. Fang et al. addressed this issue by employing the Chinese Remainder Theorem to divide gradients into k parts and compress them, reducing the size of each gradient to $\frac{1}{k}$ of its original size and thus lowering communication overhead within the system [55]. Despite this improvement, the communication overhead remains higher than that of the second solution. Therefore, we consider that such techniques possess restricted applicability and are suitable only for systems characterized by a few supernodes, each endowed with robust storage capabilities capable of managing storage redundancy.

The second solution involves utilizing distributed storage technology to house the model, while retaining only the acquisition method on the blockchain [66], [88], [135], [151], [152]. For example, the InterPlanetary File System (IPFS) employs content addressing for file storage and retrieval, allowing users to access files using the hash value associated with the file [56]. In this methodology, solely the hash of the respective model finds its place on the blockchain. Additionally, Xu et al. incorporated a model producer within the system to provide download links to other nodes [96]. The blockchain then retains the model hashes and corresponding download links solely as part of this innovative approach. These approaches address the intricate interplay between blockchain and FL requirements, paving the way for more efficient and effective storage management within BC-FL systems.

Generally, the storage limitations of blockchain do not necessarily affect the choice of models. This is because many works do not store model-related information on the blockchain. Even if storage is required, some distributed storage solutions can be used. The models that BC-FL used can choose are consistent with those of FL.

V. Discussion

A. Future Research Directions

1) Combination Blockchain With VFL and FTL: In the main sections of this survey, we did not differentiate between HFL, VFL, and FTL, as most existing BC-FL systems are based on HFL. In those BC-FL systems based on VFL and FTL, the objectives and methods involving blockchain are similar to those in HFL-based systems. However, VFL and FTL involve many additional steps. Taking VFL as an example, clients do not have complete data features, so tasks such as encrypted entity alignment and model splitting are required, making it more complex than HFL. We hope that future researchers will explore how blockchain can play a role in the unique steps of VFL and FTL, further enhancing their security, efficiency, and credibility, and driving the advancement of this field.

2) Lightweight Blockchain Solutions: In FL systems, particularly in cross-device FL, clients typically exhibit constrained communication and computational capacities. Introducing blockchain on each client might further burden the communication and computational resources of edge devices. The majority of blockchains in BC-FL systems maintain a rather general-purpose nature, with only a handful being meticulously customized for these systems. The forthcoming challenge lies in the advancement of consensus algorithms, topology structures, communication methodologies, and other enhancements aimed at enhancing the compatibility of blockchain systems with the FL framework.

3) Commercial Applications: Despite the emergence of numerous BC-FL papers, most of them only conduct local simulations or small-scale experiments, making it difficult to determine whether their proposed methods are effective in large-scale applications. Additionally, there is currently a lack of commercial applications for large-scale BC-FL systems. Therefore, future research needs to focus on applying BC-FL systems to large-scale commercial applications to validate their effectiveness and commercial value.

B. Combination of Blockchain and Federated Learning

The integration of blockchain into FL necessitates addressing two critical questions: which nodes operate the blockchain and what data is stored on the blockchain. They significantly impact the system's performance, security, and efficiency, and are highly dependent on the specific application scenario.

Blockchain Node Selection: A fully decentralized approach is well-suited for scenarios where transparency, security, or trust are paramount, and where computational and storage resources are evenly distributed. Examples include inter-hospital collaborations [24], [156] and drone networks [56]. This approach fosters a highly democratic and transparent system, avoiding excessive control by any single entity. However, it requires each node to possess substantial computational and storage capabilities. In contrast, a partially decentralized approach is better suited for heterogeneous environments with nodes performing different functions or in applications with hierarchical structures. For instance, supernodes run the blockchain in industrial scenario, while other nodes focus on FL tasks and interact with the blockchain through these supernodes [80], [166].

Data Storage on the Blockchain: Storing training-related data, such as model parameters, on the blockchain enhances security and ensures data integrity, making it suitable for scenarios with numerous malicious nodes or stringent security requirements [106]. Storing incentive-related data supports a fair and transparent FL environment, fostering a healthy FL marketplace, such as FL crowdsourcing [88]. Reputation data stored on the blockchain can improve system security [167] or serve as a basis for incentives [81]. Additionally, the use of smart contracts offers flexibility and scalability for various applications [63], [110].

The integration of blockchain and FL is highly adaptable and context-dependent. For example, in a drone-based BC-FL deployment, the selection of nodes to run the blockchain would depend on factors such as the number of drones, their computational and communication capabilities, and the specific security requirements of the mission. Simultaneously, what data is stored on the blockchain is selected based on the execution mission and runtime environment. The design of a BC-FL system should be based on a meticulous evaluation of the specific requirements and characteristics of the application at hand.

C. Real World Applications

In Table VI, we investigate and enumerate the real-world applications of BC-FL systems for readers' reference. Currently, BC-FL systems find application in a wide array of domains, including healthcare, intelligent transportation and industry.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

VI. Conclusion

Blockchain-empowered Federated Learning (BC-FL) has emerged as a promising research area. This survey explored how blockchain enhances FL by improving security, preventing single points of failure, and enabling reputation and incentive mechanisms. We also discussed key challenges, including efficiency, storage, and security issues, along with existing solutions. Finally, we discussed real-world applications, integration strategies, and future research directions. We hope this work provides valuable insights and accelerates further exploration in BC-FL.

References

[1] C. Meurisch et al., "Data protection in AI services: A survey," *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–38, 2021.

[2] M. Bashir et al., "Online privacy and informed consent: The dilemma of information asymmetry," in *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–10, 2015, .

[3] A. K. R. Nadikattu, "IoT and the issue of data privacy," *Int. J. Innov. Eng. Res. Technol.*, vol. 5, no. 10, pp. 23–26, 2018.

[4] J. Isaak et al., "User data privacy: Facebook, Cambridge analytica, and privacy protection," *Computer*, vol. 51, no. 8, pp. 56–59, 2018.

[5] X. Yin et al., "A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions," *ACM Comput. Surv.*, vol. 54, no. 6, pp. 1–36, 2021.

[6] C. Tikkinen-Piri et al., "EU general data protection regulation: Changes and implications for personal data collecting companies," *Comput. Law Secur. Rev.*, vol. 34, no. 1, pp. 134–153, 2018.

[7] C. F. Mondschein et al., "The EU's general data protection regulation (GDPR) in a research context," in *Fundamentals of Clinical Data Science*. Cham, Switzerland: Springer, 2019, pp. 55–71.

[8] S. M. Boyne, "Data protection in the United States," *Amer. J. Comp. Law*, vol. 66, no. suppl_1, pp. 299–343, 2018.

[9] B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," in *Proc. Artif. Intell. Statist.*, PMLR, 2017, pp. 1273–1282.

[10] X. Wang, C. Wang, X. Li, V. C. M. Leung, and T. Taleb, "Federated deep reinforcement learning for Internet of Things with decentralized cooperative edge caching," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9441–9455, Oct. 2020.

[11] Y. Liu, J. Peng, J. Kang, A. M. Iliyasu, D. Niyato, and A. A. A. El-Latif, "A secure federated learning framework for 5G networks," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 24–31, Aug. 2020.

[12] N. Ding, Z. Fang, and J. Huang, "Incentive mechanism design for federated learning with multi-dimensional private information," in *Proc. IEEE Int. Symp. Model. Optim. Mobile, Ad Hoc, Wireless Netw.*, 2020, pp. 1–8.

[13] Q. Li, Y. Diao, Q. Chen, and B. He, "Federated learning on non-IID data silos: An experimental study," in *Proc. IEEE Int. Conf. Data Eng.*, 2022, pp. 965–978.

[14] D. C. Nguyen et al., "Federated learning for smart healthcare: A survey," *ACM Comput. Surv.*, vol. 55, no. 3, pp. 1–37, 2022.

[15] H. Li et al., "Review on security of federated learning and its application in healthcare," *Future Gener. Comput. Syst.*, vol. 144, pp. 271–290, 2023.

[16] S. Rani et al., "Federated learning for secure IOMT-applications in smart healthcare systems: A comprehensive review," *Knowl.-Based Syst.*, vol. 274, 2023, Art. no. 110658.

[17] W. Zheng et al., "Federated meta-learning for fraudulent credit card detection," in *Proc. Int. Conf. Int. Joint Conf. Artif. Intell.*, 2021, pp. 4654–4660.

[18] X. Zhang, A. Mavromatis, A. Vafeas, R. Nejabati, and D. Simeonidou, "Federated feature selection for horizontal federated learning in IoT networks," *IEEE Internet Things J.*, vol. 10, no. 11, pp. 10095–10112, Jun. 2023.

[19] A. Arunan, Y. Qin, X. Li, and C. Yuen, "A federated learning-based industrial health prognostics for heterogeneous edge devices using matched feature extraction," *IEEE Trans. Automat. Sci. Eng.*, vol. 21, no. 3, pp. 3065–3079, Jul. 2024.

[20] S. Zeng et al., "HFedMS: Heterogeneous federated learning with memorable data semantics in industrial metaverse," *IEEE Trans. Cloud Comput.*, vol. 11, no. 3, pp. 3055–3069, Third Quarter 2023.

[21] S. Ji, J. Zhang, Y. Zhang, Z. Han, and C. Ma, "LAFED: A lightweight authentication mechanism for blockchain-enabled federated learning system," *Future Gener. Comput. Syst.*, vol. 145, pp. 56–67, 2023.

[22] F. Yang et al., "An explainable federated learning and blockchain-based secure credit modeling method," *Eur. J. Oper. Res.*, vol. 317, no. 2, pp. 449–467, 2024.

[23] Z. A. E. Houda, A. S. Hafid, and L. Khoukhi, "MiTFed: A privacy preserving collaborative network attack mitigation framework based on federated learning using sdn and blockchain," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 4, pp. 1985–2001, Jul./Aug. 2023.

[24] A. P. Kalapaaking, I. Khalil, and X. Yi, "Blockchain-based federated learning with SMPC model verification against poisoning attack for healthcare systems," *IEEE Trans. Emerg. Topics Comput.*, vol. 12, no. 1, pp. 269–280, First Quarter, 2024.

[25] Y. Wang, J. Zhou, G. Feng, X. Niu, and S. Qin, "Blockchain assisted federated learning for enabling network edge intelligence," *IEEE Netw.*, vol. 37, no. 1, pp. 96–102, Jan./Feb. 2023.

[26] J. Xu et al., "A survey of blockchain consensus protocols," *ACM Comput. Surv.*, vol. 55, no. 13s, pp. 1–35, 2023.

[27] 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. 117134–117151, 2019.

[28] H.-N. Dai, Z. Zheng, and Y. Zhang, "Blockchain for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8076–8094, Oct. 2019.

[29] S. Aggarwal et al., "Blockchain 2.0: Smart contracts," in *Advances in Computers*, vol. 121. Amsterdam, The Netherlands: Elsevier, 2021, pp. 301–322.

[30] A. Hafid, A. S. Hafid, and M. Samih, "A tractable probabilistic approach to analyze sybil attacks in sharding-based blockchain protocols," *IEEE Trans. Emerg. Topics Comput.*, vol. 11, no. 1, pp. 126–136, First Quarter, 2023.

[31] W. Issa et al., "Blockchain-based federated learning for securing Internet of Things: A comprehensive survey," *ACM Comput. Surv.*, vol. 55, no. 9, pp. 1–43, 2023.

[32] M. Ali et al., "Integration of blockchain and federated learning for Internet of Things: Recent advances and future challenges," *Comput. Secur.*, vol. 108, 2021, Art. no. 102355.

[33] J. Zhu et al., "Blockchain-empowered federated learning: Challenges, solutions, and future directions," *ACM Comput. Surv.*, vol. 55, pp. 1–31, 2022.

[34] R. Myrzashova, S. H. Alsamhi, A. V. Shvetsov, A. Hawbani, and X. Wei, "Blockchain meets federated learning in healthcare: A systematic review with challenges and opportunities," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14418–14437, Aug. 2023.

[35] A. M. Eldin et al., "Federated blockchain system (FBS) for the healthcare industry," *Sci. Rep.*, vol. 13, no. 1, 2023, Art. no. 2569.

[36] Y. Qu et al., "Blockchain-enabled federated learning: A survey," *ACM Comput. Surv.*, vol. 55, no. 4, pp. 1–35, 2022.

[37] K. Sameera et al., "Privacy-preserving in blockchain-based federated learning systems," *Comput. Commun.*, vol. 222, pp. 38–67, 2024.

[38] S. Vargaftik et al., "EDEN: Communication-efficient and robust distributed mean estimation for federated learning," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2022, pp. 21984–22014.

[39] Y. Wang et al., "Communication-efficient adaptive federated learning," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2022, pp. 22802–22838.

[40] C. Zhang et al., "A survey on federated learning," *Knowl.-Based Syst.*, vol. 216, 2021, Art. no. 106775.

[41] Z. Teimoori, A. Yassine, and M. S. Hossain, "A secure cloudlet-based charging station recommendation for electric vehicles empowered by federated learning," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6464–6473, Sep. 2022.

[42] S. Xu et al., "A method of federated learning based on blockchain," in *Proc. Int. Conf. Comput. Sci. Appl. Eng.*, 2021, pp. 1–8.

[43] P. Zhang, H. Sun, J. Situ, C. Jiang, and D. Xie, "Federated transfer learning for IIoT devices with low computing power based on blockchain and edge computing," *IEEE Access*, vol. 9, pp. 98630–98638, 2021.

[44] D. Chen et al., "FS-REAL: Towards real-world cross-device federated learning," in *Proc. ACM SIGKDD Conf. Knowl. Discov. Data Mining*, 2023, pp. 3829–3841.

[45] R. Dorfman et al., "DoCoFL: Downlink compression for cross-device federated learning," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2023, pp. 8356–8388.

[46] C. Huang, M. Tang, Q. Ma, J. Huang, and X. Liu, "Promoting collaborations in cross-silo federated learning: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 62, no. 4, pp. 82–88, Apr. 2024.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

[47] S. Yuan et al., "Adaptive incentivize for cross-silo federated learning in IIoT: A multi-agent reinforcement learning approach," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 15048–15058, May 2024.
[48] Y. Jiao, P. Wang, D. Niyato, B. Lin, and D. I. Kim, "Toward an automated auction framework for wireless federated learning services market," *IEEE Trans. Mobile Comput.*, vol. 20, no. 10, pp. 3034–3048, Oct. 2021.
[49] Z. Zheng et al., "Blockchain challenges and opportunities: A survey," *Int. J. Web Grid Serv.*, vol. 14, no. 4, pp. 352–375, 2018.
[50] T. Ye, M. Luo, Y. Yang, K.-K. R. Choo, and D. He, "A survey on redactable blockchain: Challenges and opportunities," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 3, pp. 1669–1683, May/Jun. 2023.
[51] C. Lepore et al., "A survey on blockchain consensus with a performance comparison of PoW, PoS and pure PoS," *Mathematics*, vol. 8, no. 10, 2020, Art. no. 1782.
[52] W. Li, C. Feng, L. Zhang, H. Xu, B. Cao, and M. A. Imran, "A scalable multi-layer PBFT consensus for blockchain," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 5, pp. 1146–1160, May 2021.
[53] D. Huang, X. Ma, and S. Zhang, "Performance analysis of the raft consensus algorithm for private blockchains," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 1, pp. 172–181, Jan. 2020.
[54] M. Abdel-Basset, N. Moustafa, and H. Hawash, "Privacy-preserved cyberattack detection in industrial edge of things (IEoT): A blockchain-orchestrated federated learning approach," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 7920–7934, Nov. 2022.
[55] C. Fang et al., "A privacy-preserving and verifiable federated learning method based on blockchain," *Comput. Commun.*, vol. 186, pp. 1–11, 2022.
[56] C. Feng, B. Liu, K. Yu, S. K. Goudos, and S. Wan, "Blockchain-empowered decentralized horizontal federated learning for 5G-enabled UAVs," *IEEE Trans. Ind. Informat.*, vol. 18, no. 5, pp. 3582–3592, May 2022.
[57] S. Guo et al., "Sandbox computing: A data privacy trusted sharing paradigm via blockchain and federated learning," *IEEE Trans. Comput.*, vol. 72, no. 3, pp. 800–810, Mar. 2023.
[58] L. Jiang, H. Zheng, H. Tian, S. Xie, and Y. Zhang, "Cooperative federated learning and model update verification in blockchain-empowered digital twin edge networks," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11154–11167, Jul. 2022.
[59] H. Liu et al., "Blockchain and federated learning for collaborative intrusion detection in vehicular edge computing," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6073–6084, Jun. 2021.
[60] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, "Blockchain empowered asynchronous federated learning for secure data sharing in Internet of Vehicles," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4298–4311, Apr. 2020.
[61] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, and A. Y. Zomaya, "Federated learning for COVID-19 detection with generative adversarial networks in edge cloud computing," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10257–10271, Jun. 2022.
[62] D. C. Nguyen, S. Hosseinalipour, D. J. Love, P. N. Pathirana, and C. G. Brinton, "Latency optimization for blockchain-empowered federated learning in multi-server edge computing," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 12, pp. 3373–3390, Dec. 2022.
[63] J. Qi, F. Lin, Z. Chen, C. Tang, R. Jia, and M. Li, "High-quality model aggregation for blockchain-based federated learning via reputation-motivated task participation," *IEEE Internet Things J.*, vol. 9, no. 19, pp. 18378–18391, Oct. 2022.
[64] M. Qi et al., "A blockchain-enabled federated learning model for privacy preservation: System design," in *Proc. Australas. Conf. Inf. Secur. Privacy*, Springer, 2021, pp. 473–489.
[65] Y. Qu, S. R. Pokhrel, S. Garg, L. Gao, and Y. Xiang, "A blockchained federated learning framework for cognitive computing in industry 4.0 networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 4, pp. 2964–2973, Apr. 2021.
[66] M. H. U. Rehman, A. M. Dirir, K. Salah, E. Damiani, and D. Svetinovic, "TrustFed: A framework for fair and trustworthy cross-device federated learning in IIoT," *IEEE Trans. Ind. Informat.*, vol. 17, no. 12, pp. 8485–8494, Dec. 2021.
[67] D. Wu et al., "A blockchain-based multi-layer decentralized framework for robust federated learning," in *Proc. Int. Joint Conf. Neural Netw.*, 2022, pp. 1–8.
[68] C. Xu, Y. Qu, P. W. Eklund, Y. Xiang, and L. Gao, "BAFL: An efficient blockchain-based asynchronous federated learning framework," in *Proc. IEEE Symp. Comput. Commun.*, 2021, pp. 1–6.
[69] Y. Xu et al., "BESIFL: Blockchain empowered secure and incentive federated learning paradigm in IoT," *IEEE Internet Things J.*, vol. 10, no. 8, pp. 6561–6573, Apr. 2023.
[70] P. Zhang, Y. Hong, N. Kumar, M. Alazab, M. D. Alshehri, and C. Jiang, "BC-edgeFL: A defensive transmission model based on blockchain-assisted reinforced federated learning in IIoT environment," *IEEE Trans. Ind. Informat.*, vol. 18, no. 5, pp. 3551–3561, May 2022.
[71] N. Zhao, H. Wu, F. R. Yu, L. Wang, W. Zhang, and V. C. M. Leung, "Deep-reinforcement-learning-based latency minimization in edge intelligence over vehicular networks," *IEEE Internet Things J.*, vol. 9, no. 2, pp. 1300–1312, Jan. 2022.
[72] Z. Wang, Q. Hu, R. Li, M. Xu, and Z. Xiong, "Incentive mechanism design for joint resource allocation in blockchain-based federated learning," *IEEE Trans. Parallel Distrib. Syst.*, vol. 34, no. 5, pp. 1536–1547, May 2023.
[73] L. Ouyang, F. -Y. Wang, Y. Tian, X. Jia, H. Qi, and G. Wang, "Artificial identification: A novel privacy framework for federated learning based on blockchain," *IEEE Trans. Computat. Social Syst.*, vol. 10, no. 6, pp. 3576–3585, Dec. 2023.
[74] S. Yuan, B. Cao, Y. Sun, Z. Wan, and M. Peng, "Secure and efficient federated learning through layering and sharding blockchain," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 3, pp. 3120–3134, May/Jun. 2024.
[75] M. Aloqaily, I. A. Ridhawi, and S. Kanhere, "Reinforcing industry 4.0 with digital twins and blockchain-assisted federated learning," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 11, pp. 3504–3516, Nov. 2023.
[76] J. Mu, W. Ouyang, T. Hong, W. Yuan, Y. Cui, and Z. Jing, "Digital twin-enabled federated learning in mobile networks: From the perspective of communication-assisted sensing," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 10, pp. 3230–3241, Oct. 2023.
[77] A. Wahrstatter et al., "OpenFL: A scalable and secure decentralized federated learning system on the ethereum blockchain," *Internet Things*, vol. 26, 2024, Art. no. 101174.
[78] J. Li et al., "Blockchain assisted decentralized federated learning (BLADE-FL): Performance analysis and resource allocation," *IEEE Trans. Parallel Distrib. Syst.*, vol. 33, no. 10, pp. 2401–2415, Oct. 2022.
[79] Y. Qu, C. Xu, L. Gao, Y. Xiang, and S. Yu, "FL-SEC: Privacy-preserving decentralized federated learning using signSGD for the internet of artificially intelligent things," *IEEE Internet Things Mag.*, vol. 5, no. 1, pp. 85–90, Mar. 2022.
[80] X. He, Q. Chen, L. Tang, W. Wang, and T. Liu, "CGAN-based collaborative intrusion detection for UAV networks: A blockchain-empowered distributed federated learning approach," *IEEE Internet Things J.*, vol. 10, no. 1, pp. 120–132, Jan. 2023.
[81] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10700–10714, Dec. 2019.
[82] A. Haddaji et al., "Federated learning with blockchain approach for trust management in IoV," in *Proc. Int. Conf. Adv. Inf. Netw. Appl.*, Springer, 2022, pp. 411–423.
[83] J. Kang et al., "Optimizing task assignment for reliable blockchain-empowered federated edge learning," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1910–1923, Feb. 2021.
[84] C. Qiu, G. S. Aujla, J. Jiang, W. Wen, and P. Zhang, "Rendering secure and trustworthy edge intelligence in 5G-enabled IIoT using proof of learning consensus protocol," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 900–909, Jan. 2023.
[85] H. Chen et al., "RepBFL: Reputation based blockchain-enabled federated learning framework for data sharing in internet of vehicles," in *Proc. Int. Conf. Parallel Distrib. Comput. Appl. Technol.*, Springer, 2022, pp. 536–547.
[86] L. Gao et al., "FGFL: A blockchain-based fair incentive governor for federated learning," *J. Parallel Distrib. Comput.*, vol. 163, pp. 283–299, 2022.
[87] M. A. Rahman, M. S. Hossain, M. S. Islam, N. A. Alrajeh, and G. Muhammad, "Secure and provenance enhanced internet of health things framework: A blockchain managed federated learning approach," *IEEE Access*, vol. 8, pp. 205071–205087, 2020.
[88] Y. Zhao et al., "Privacy-preserving blockchain-based federated learning for IoT devices," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1817–1829, Feb. 2021.
[89] Y. Lin et al., "DRL-based adaptive sharding for blockchain-based federated learning," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 5992–6004, Oct. 2023.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

[90] Y. Fu, C. Li, F. R. Yu, T. H. Luan, and P. Zhao, "An incentive mechanism of incorporating supervision game for federated learning in autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 12, pp. 14800–14812, Dec. 2023.

[91] S. Gollapudi et al., "Profit sharing and efficiency in utility games," in *Proc. Annu. Eur. Symp. Algorithms*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017, pp. 43:1–43:14.

[92] H. Yu et al., "A fairness-aware incentive scheme for federated learning," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc.*, 2020, pp. 393–399.

[93] Y. Zhan, J. Zhang, Z. Hong, L. Wu, P. Li, and S. Guo, "A survey of incentive mechanism design for federated learning," *IEEE Trans. Emerg. Topics Comput.*, vol. 10, no. 2, pp. 1035–1044, Second Quarter, 2022.

[94] Y. Zhan, P. Li, Z. Qu, D. Zeng, and S. Guo, "A learning-based incentive mechanism for federated learning," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6360–6368, Jul. 2020.

[95] Y. Chen et al., "DIM-DS: Dynamic incentive model for data sharing in federated learning based on smart contracts and evolutionary game theory," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 24572–24584, Dec. 2022.

[96] Z. Li et al., "Byzantine resistant secure blockchained federated learning at the edge," *IEEE Netw.*, vol. 35, no. 4, pp. 295–301, Jul./Aug. 2021.

[97] C. Ma et al., "When federated learning meets blockchain: A new distributed learning paradigm," *IEEE Comput. Intell. Mag.*, vol. 17, no. 3, pp. 26–33, Aug. 2022.

[98] Y. Qu et al., "Decentralized privacy using blockchain-enabled federated learning in fog computing," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5171–5183, Jun. 2020.

[99] Y. Qu, L. Gao, Y. Xiang, S. Shen, and S. Yu, "FedTwin: Blockchain-enabled adaptive asynchronous federated learning for digital twin networks," *IEEE Netw.*, vol. 36, no. 6, pp. 183–190, Nov./Dec. 2022.

[100] Z. Wang et al., "Blockchain empowered federated learning for data sharing incentive mechanism," *Procedia Comput. Sci.*, vol. 202, pp. 348–353, 2022.

[101] J. Weng, J. Weng, J. Zhang, M. Li, Y. Zhang, and W. Luo, "DeepChain: Auditable and privacy-preserving deep learning with blockchain-based incentive," *IEEE Trans. Dependable Secure Comput.*, vol. 18, no. 5, pp. 2438–2455, Sep./Oct. 2021.

[102] R. Xu and Y. Chen, "μDFL: A secure microchained decentralized federated learning fabric atop IoT networks," *IEEE Trans. Netw. Service Manag.*, vol. 19, no. 3, pp. 2677–2688, Sep. 2022.

[103] C. Zhang et al., "A blockchain-based model migration approach for secure and sustainable federated learning in IoT systems," *IEEE Internet Things J.*, vol. 10, no. 8, pp. 6574–6585, Apr. 2023.

[104] Z. Zhang et al., "Refiner: A reliable incentive-driven federated learning system powered by blockchain," in *Proc. VLDB Endowment*, vol. 14, no. 12, pp. 2659–2662, 2021.

[105] Y. He et al., "A game theory-based incentive mechanism for collaborative security of federated learning in energy blockchain environment," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 21294–21308, Dec. 2023.

[106] S. Awan et al., "Poster: A reliable and accountable privacy-preserving federated learning framework using the blockchain," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2019, pp. 2561–2563.

[107] R. Cheng, Y. Sun, Y. Liu, L. Xia, D. Feng, and M. A. Imran, "Blockchain-empowered federated learning approach for an intelligent and reliable D2D caching scheme," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 7879–7890, Jun. 2021.

[108] L. Feng, Z. Yang, S. Guo, X. Qiu, W. Li, and P. Yu, "Two-layered blockchain architecture for federated learning over the mobile edge network," *IEEE Netw.*, vol. 36, no. 1, pp. 45–51, Jan./Feb. 2022.

[109] B. Jia, X. Zhang, J. Liu, Y. Zhang, K. Huang, and Y. Liang, "Blockchain-enabled federated learning data protection aggregation scheme with differential privacy and homomorphic encryption in IIoT," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 4049–4058, Jun. 2022.

[110] Y. Liu, Y. Kang, C. Xing, T. Chen, and Q. Yang, "A secure federated transfer learning framework," *IEEE Intell. Syst.*, vol. 35, no. 4, pp. 70–82, Jul./Aug. 2020.

[111] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for 5G beyond," *IEEE Netw.*, vol. 35, no. 1, pp. 219–225, Jan./Feb. 2021.

[112] Y. Miao, Z. Liu, H. Li, K. -K. R. Choo, and R. H. Deng, "Privacy-preserving byzantine-robust federated learning via blockchain systems," *IEEE Trans. Inf. Forensics Secur.*, vol. 17, pp. 2848–2861, 2022.

[113] V. Mothukuri, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and K. -K. R. Choo, "FabricFL: Blockchain-in-the-loop federated learning for trusted decentralized systems," *IEEE Syst. J.*, vol. 16, no. 3, pp. 3711–3722, Sep. 2022.

[114] J. Sun, Y. Wu, S. Wang, Y. Fu, and X. Chang, "Permissioned blockchain frame for secure federated learning," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 13–17, Jan. 2022.

[115] X. Huang, Y. Wu, C. Liang, Q. Chen, and J. Zhang, "Distance-aware hierarchical federated learning in blockchain-enabled edge computing network," *IEEE Internet Things J.*, vol. 10, no. 21, pp. 19163–19176, Nov. 2023.

[116] G. Xu, Z. Zhou, J. Dong, L. Zhang, and X. Song, "A blockchain-based federated learning scheme for data sharing in industrial Internet of Things," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 21467–21478, Dec. 2023.

[117] J. Weng, J. Weng, H. Huang, C. Cai, and C. Wang, "FedServing: A federated prediction serving framework based on incentive mechanism," in *Proc. IEEE Conf. Comput. Commun.*, 2021, pp. 1–10.

[118] S. M. H. Bamakan et al., "A survey of blockchain consensus algorithms performance evaluation criteria," *Expert Syst. Appl.*, vol. 154, 2020, Art. no. 113385.

[119] C. Fan, S. Ghaemi, H. Khazaei, and P. Musilek, "Performance evaluation of blockchain systems: A systematic survey," *IEEE Access*, vol. 8, pp. 126927–126950, 2020.

[120] M. Shayan, C. Fung, C. J. M. Yoon, and I. Beschastnikh, "Biscotti: A blockchain system for private and secure federated learning," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 7, pp. 1513–1525, Jul. 2021.

[121] M. Cao, L. Zhang, and B. Cao, "Toward on-device federated learning: A direct acyclic graph-based blockchain approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 4, pp. 2028–2042, Apr. 2023.

[122] N. Q. Hieu, T. A. Tran, C. L. Nguyen, D. Niyato, D. I. Kim, and E. Elmroth, "Deep reinforcement learning for resource management in blockchain-enabled federated learning network," *IEEE Netw. Lett.*, vol. 4, no. 3, pp. 137–141, Sep. 2022.

[123] Y. Li, C. Chen, N. Liu, H. Huang, Z. Zheng, and Q. Yan, "A blockchain-based decentralized federated learning framework with committee consensus," *IEEE Netw.*, vol. 35, no. 1, pp. 234–241, Jan./Feb. 2021.

[124] D. Mingxiao, M. Xiaofeng, Z. Zhe, W. Xiangwei, and C. Qijun, "A review on consensus algorithm of blockchain," in *Proc. IEEE Int. Conf. Syst. Man, Cybern.*, 2017, pp. 2567–2572.

[125] J. Göbel and A. E. Krzesinski, "Increased block size and bitcoin blockchain dynamics," in *Proc. IEEE Int. Telecommun. Netw. Appl. Conf.*, 2017, pp. 1–6.

[126] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, Nov. 2017.

[127] Z. Wang et al., "Dueling network architectures for deep reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2016, pp. 1995–2003.

[128] T. Zahavy et al., "A self-tuning actor-critic algorithm," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 20913–20924.

[129] W. Hao et al., "Towards a trust-enhanced blockchain P2P topology for enabling fast and reliable broadcast," *IEEE Trans. Netw. Service Manag.*, vol. 17, no. 2, pp. 904–917, Jun. 2020.

[130] J. Xie, F. R. Yu, T. Huang, R. Xie, J. Liu, and Y. Liu, "A survey on the scalability of blockchain systems," *IEEE Netw.*, vol. 33, no. 5, pp. 166–173, Sep./Oct. 2019.

[131] C. Zhang, Y. Xu, X. Wu, E. Wang, H. Jiang, and Y. Zhang, "A semi-asynchronous decentralized federated learning framework via tree-graph blockchain," in *Proc. IEEE Int. Conf. Comput. Commun.*, 2024, pp. 1121–1130.

[132] S. Zhang and J. -H. Lee, "Double-spending with a sybil attack in the Bitcoin decentralized network," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5715–5722, Oct. 2019.

[133] J. Zhang, Y. Chen, and H. Li, "Privacy leakage of adversarial training models in federated learning systems," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 107–113.

[134] Z. Yuan et al., "SecretGen: Privacy recovery on pre-trained models via distribution discrimination," in *Proc. Eur. Conf. Comput. Vis.*, Springer, 2022, pp. 139–155.

[135] U. Majeed, L. U. Khan, A. Yousafzai, Z. Han, B. J. Park, and C. S. Hong, "ST-BFL: A structured transparency empowered cross-silo federated learning on the blockchain framework," *IEEE Access*, vol. 9, pp. 155634–155650, 2021.

[136] Z. Sun et al., "A blockchain-based audit approach for encrypted data in federated learning," *Digit. Commun. Netw.*, vol. 8, no. 5, pp. 614–624, 2022.

[137] A. Acar et al., "A survey on homomorphic encryption schemes: Theory and implementation," *ACM Comput. Surveys*, vol. 51, no. 4, pp. 1–35, 2018.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.

[138] Q. Li and M. G. Christensen, "A privacy-preserving asynchronous averaging algorithm based on Shamir's secret sharing," in *Proc. IEEE Eur. Signal Process. Conf.*, 2019, pp. 1–5.
[139] M. Abadi et al., "Deep learning with differential privacy," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2016, pp. 308–318.
[140] J. Passerat-Palmbach et al., "Blockchain-orchestrated machine learning for privacy preserving federated learning in electronic health data," in *Proc. IEEE Int. Conf. Blockchain*, 2020, pp. 550–555.
[141] C. Zhao et al., "Secure multi-party computation: Theory, practice and applications," *Inf. Sci.*, vol. 476, pp. 357–372, 2019.
[142] J. Wei et al., "A redactable blockchain framework for secure federated learning in industrial Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 18, pp. 17901–17911, Sep. 2022.
[143] J. Guo, J. Wu, A. Liu, and N. N. Xiong, "LightFed: An efficient and secure federated edge learning system on model splitting," *IEEE Trans. Parallel Distrib. Syst.*, vol. 33, no. 11, pp. 2701–2713, Nov. 2022.
[144] Z. Qin, X. Yan, M. Zhou, and S. Deng, "BlockDFL: A blockchain-based fully decentralized peer-to-peer federated learning framework," in *Proc. ACM Web Conf.*, 2024, pp. 2914–2925.
[145] P. Otte et al., "TrustChain: A sybil-resistant scalable blockchain," *Future Gener. Comput. Syst.*, pp. 770–780, 2020.
[146] M. Baza et al., "Detecting sybil attacks using proofs of work and location in VANETs," *IEEE Trans. Dependable Secure Comput.*, vol. 19, no. 1, pp. 39–53, Jan./Feb. 2022.
[147] R. Matzutt, J. Pennekamp, E. Buchholz, and K. Wehrle, "Utilizing public blockchains for the sybil-resistant bootstrapping of distributed anonymity services," in *Proc. ACM Asia Conf. Comput. Commun. Secur.*, 2020, pp. 531–542.
[148] Y. Gilad et al., "Algorand: Scaling byzantine agreements for cryptocurrencies," in *Proc. Symp. operating Syst. Princ.*, 2017, pp. 51–68.
[149] R. Nagayama, R. Banno, and K. Shudo, "Identifying impacts of protocol and internet development on the Bitcoin network," in *Proc. IEEE Symp. Comput. Commun.*, 2020, pp. 1–6.
[150] P. Ramanan and K. Nakayama, "BAFFLE: Blockchain based aggregator free federated learning," in *Proc. IEEE Int. Conf. Blockchain*, 2020, pp. 72–81.
[151] M. M. Salim and J. H. Park, "Federated learning-based secure electronic health record sharing scheme in medical informatics," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 2, pp. 617–624, Feb. 2023.
[152] L. Ouyang, Y. Yuan, and F. -Y. Wang, "Learning markets: An AI collaboration framework based on blockchain and smart contracts," *IEEE Internet Things J.*, vol. 9, no. 16, pp. 14273–14286, Aug. 2022.
[153] Z. Lian, W. Wang, Z. Han, and C. Su, "Blockchain-based personalized federated learning for internet of medical things," *IEEE Trans. Sustain. Comput.*, vol. 8, no. 4, pp. 694–702, Fourth Quarter 2023.
[154] R. Kumar et al., "Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging," *IEEE Sensors J.*, vol. 21, no. 14, pp. 16301–16314, Jul., 2021.
[155] J. Mu, M. Kadoch, T. Yuan, W. Lv, Q. Liu, and B. Li, "Explainable federated medical image analysis through causal learning and blockchain," *IEEE J. Biomed. Health Informat.*, vol. 28, no. 6, pp. 3206–3218, Jun. 2024.
[156] S. Warnat-Herresthal et al., "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, no. 7862, pp. 265–270, 2021.
[157] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, "Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 5098–5107, Jul. 2021.
[158] B. Ghimire, D. B. Rawat, and A. Rahman, "Efficient information dissemination in blockchain-enabled federated learning for IoV," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 15310–15319, May 2024.
[159] K. T. Pauu, J. Wu, Y. Fan, Q. Pan, and M.-I.-V. Maka, "Differential privacy and blockchain-empowered decentralized graph federated learning enabled UAVs for disaster response," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 20930–20947, Jun. 2024.
[160] Y. He, K. Huang, G. Zhang, F. R. Yu, J. Chen, and J. Li, "Bift: A blockchain-based federated learning system for connected and autonomous vehicles," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12311–12322, Jul. 2022.
[161] C. Meese et al., "Adaptive traffic prediction at the ITS edge with online models and blockchain-based federated learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 9, pp. 10725–10740, Sep. 2024.
[162] Z. A. El Houda, H. Moudoud, B. Brik, and L. Khoukhi, "Blockchain-enabled federated learning for enhanced collaborative intrusion detection in vehicular edge computing," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 7, pp. 7661–7672, Jul. 2024.
[163] H. Ma et al., "Blockchain-based privacy-preserving federated learning for mobile crowdsourcing," *IEEE Internet Things J.*, vol. 11, no. 8, pp. 13884–13899, Apr. 2024.
[164] W. Zhang et al., "Blockchain-based federated learning for device failure detection in industrial IoT," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5926–5937, Apr. 2021.
[165] W. Hammedi, B. Brik, and S. M. Senouci, "Toward optimal MEC-based collision avoidance system for cooperative inland vessels: A federated deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2525–2537, Feb. 2023.
[166] Y. Li, X. Tao, X. Zhang, J. Liu, and J. Xu, "Privacy-preserved federated learning for autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8423–8434, Jul. 2022.
[167] J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang, and M. Guizani, "Reliable federated learning for mobile networks," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 72–80, Apr. 2020.

Zeju Cai received the BE degree from Jinan University, China, in 2022. He is currently working toward the master's degree with the School of Software Engineering of Sun Yat-sen University, China. His research interests include federated learning and Blockchain.

Jianguo Chen (Member, IEEE) is an associate professor with Sun Yat-sen University, China. His research focuses on high-performance AI, federated learning, distributed computing, and their applications in intelligent transportation and medicine. He has published more than 60 papers in top conferences and journals including *IEEE Transactions on Industrial Informatics*, and *IEEE Transactions on Parallel and Distributed Systems*. He serves as an associate editor for the *International Journal of Embedded Systems and Journal of Current Scientific Research*.

Yuting Fan received the BE degree from Sun Yat-sen University, China, in 2020. He is currently working toward the master's degree with the School of Software Engineering of Sun Yat-sen University, China. His research interests include federated learning and drone.

Zibin Zheng (Fellow, IEEE) is a professor with Sun Yat-sen University. His research focuses on service computing, data mining, software reliability, and Blockchain. He is an IET fellow, ACM outstanding scientist, and Global highly cited scientist. He has published more than 200 papers with 35000+ citations and an H-index of 87. He has received multiple awards including the Ministry of Education Natural Science Second Prize and ACM SIGSOFT Distinguished Paper Award.

Keqin Li (Fellow, IEEE) is a SUNY distinguished professor of computer science with the State University of New York and a national distinguished professor with Hunan University, China. His research focuses on cloud computing, fog computing, mobile edge computing, and energy efficient computing. He has published more than 850 papers, holds 70+ patents, and is among the world's top 5 most influential scientists in parallel and distributed computing.

Authorized licensed use limited to: National Huaqiao University. Downloaded on November 19,2025 at 09:54:39 UTC from IEEE Xplore. Restrictions apply.