electronics

MDPI

Article

Blockchain-Empowered AI for 6G-Enabled Internet of Vehicles

Ferheen Ayaz [1,*], Zhengguo Sheng [1,*], Daxin Tian [2], Maziar Nekovee [1] and Nagham Saeed [3]

[1] Department of Engineering, University of Sussex, Falmer, Brighton BN1 9RH, UK
[2] School of Transportation Science and Engineering, Beihang University, Beijing 100191, China
[3] School of Computing and Engineering, University of West London, 92-98 Warwick Road, London W5 5RF, UK
* Correspondence: f.ayaz@sussex.ac.uk (F.A.); z.sheng@sussex.ac.uk (Z.S.)

Abstract: The 6G communication technologies are expected to provide fast data rates and incessant connectivity to heterogeneous networks, such as the Internet of Vehicles (IoV). However, the resulting unprecedented surge in data traffic, massive increase in the number of nodes with high mobility, and low-latency requirements give rise to serious security, privacy, and trust challenges. The blockchain could potentially ensure trust and security in IoV due to its features, including consensus for credibility and immutability for tamper proofing. In parallel, federated learning (FL) is a privacy-preserving artificial-intelligence paradigm that does not require to share data for model training in machine learning. It can reduce data traffic and resolve privacy challenges of intelligent IoV networks. The blockchain can also complement FL by ensuring the decentralization and securing distribution of incentives. This article reviews the trends and challenges of the blockchain and FL in 6G IoV networks. Then, the impact of their combination, challenges in implementation, and future research directions are highlighted. We also evaluate our proposal of blockchain-based FL to protect IoV security and privacy that utilizes smart contract and secure transactions of incentives via the blockchain to protect FL. Compared with other solutions, the failure rate of the proposed solution was at least 5% lower with 30% malicious nodes in the network.

Keywords: blockchain; federated learning; Internet of Vehicles; security; privacy; AI

Citation: Ayaz, F.; Sheng, Z.; Tian, D.; Nekovee, M.; Saeed, N. Blockchain-Empowered AI for 6G-Enabled Internet of Vehicles. *Electronics* **2022**, *11*, 3339. https://doi.org/10.3390/electronics11203339

Academic Editors: Nurul I. Sarkar and Juan-Carlos Cano

Received: 20 September 2022
Accepted: 11 October 2022
Published: 17 October 2022

1. Introduction

The vision of ubiquitous intelligence and connectivity in the 6G era requires vehicles to fully leverage their computing resources for producing distributed artificial intelligence (AI) solutions and exploit various communication models and approaches, for example, vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) and vehicle-to-everything (V2X) in the heterogeneous environment of the Internet of Vehicles (IoV). The 6G-enabled IoV can effectively offer multiple services, including resource sharing, task offloading, and infotainment. However, one of the primary objectives of IoV is the real-time sharing of information and emergency messages to achieve safe driving conditions and healthy traffic flow [1]. For example, the timely exchange of messages about an incident or traffic jam can result in incoming vehicles planning a better journey. Nevertheless, due to the massive growth in the number of connected nodes, the IoV comes across several security challenges that are yet to be addressed.

Maintaining trust among nodes in the IoV is one of the challenges for secure and credible message dissemination. An unknown message is inherently considered to be untrusted because a malicious node may generate a false message about an incident that did not actually occur. The blockchain can be used to evaluate message credibility. It is originally a distributed ledger for securely recording a history of transactions after verifying them through a consensus [1]. The consensus is a mutual agreement between nodes to ensure a transaction's credibility. The credibility of a message is typically assessed with three approaches. The first is pre-event, where each node has its trust rating based on

whether its originated message is true or false. If its trust rating falls below a certain threshold, then the message generated by the node is marked as false. This approach is either centralized, where a central authority (CA) stores trust ratings [2], or decentralized, where a blockchain is used to record trust or reputation ratings [3]. The second is postevent, in which a message is validated on the basis of endorsements or votes from the neighboring nodes. In a postevent approach, a blockchain-based solution runs voting-based consensus to receive endorsements for a message [4]. The third is the hybrid approach that combines the advantages of the first two, i.e., only trusted nodes take part in postevent validation, which is offered by a blockchain consensus, namely, delegated proof of stake (DPoS) [5]. Furthermore, as the 6G-assisted IoV encompasses several communication protocols, a multitude of applications, and a large number of diverse and untrusted nodes, it is not feasible for a CA to be in charge of the entire network. Alternatively, decentralization offers better flexibility in such cases. The decentralized architecture, the process of verifying blocks in the presence of malicious nodes, and the cryptographic nature of the blockchain render it a suitable candidate for implementing security and privacy in the IoV.

Another pressing concern is the increasing number of sensors in autonomous vehicles (AVs) generating a large volume of data, commonly termed as big data. Big data are an essential component to the artificial-intelligence (AI) mechanisms of both AVs and intelligent transportation systems (ITSs) for perceiving the surrounding environment, and predicting traffic conditions and hazards. It may contain private information about an AV and its passengers. Sharing big data over the network poses serious security threats [6]. This is why on-device machine learning is a secure and private solution, where an inference machine-learning model is trained on a node with its own data. However, it is usually a noncooperative approach and results in the inefficient utilization of computational resources because no knowledge of an on-device trained model is shared with other nodes; therefore, one node cannot utilize the benefits of a machine-learning model produced by another. On the other hand, in a centralized solution or off-device learning, data from all connected nodes are gathered on a cloud or a central node that produces an inference model for all nodes [7]. In off-device learning, it is essential that the central node employs secure and privacy-preserving techniques for managing and storing big data. In addition, transmitting big data to a central node consumes a large amount of time and also requires a stable channel, which is impractical in high-mobility vehicle networks. Thus, to manage the diverse nature of big data and their privacy preserving requirements, federated learning (FL) is considered to be a suitable AI technique for various cooperative applications in the IoV. In FL, nodes train local models individually on their own private data. Instead of whole data, they share only local models with a central node or aggregator that combines all local models to form a global model [7].

A combination of the blockchain and FL ensures an intelligent, trusted, secure and private IoV network in 6G. The motivation behind using blockchain-enabled FL is threefold. First, the blockchain is used to manage the incentive distribution mechanism among nodes participating in FL. Blockchain-based smart contracts can securely automate the transactions of incentives in the form of cryptocurrency. Second, the blockchain introduces decentralization in FL. Instead of submitting local models to a central aggregator, nodes can add their local models as blocks into a blockchain. Third, the blockchain provides security against malicious nodes. A smart blockchain contract can detect malicious or inaccurate local models before they are added into the blockchain. Figure 1 illustrates the proposed integration of the blockchain and FL in IoV, where mobile nodes (vehicles) act as local model trainers of FL, and a road-side unit (RSU) performs aggregation.

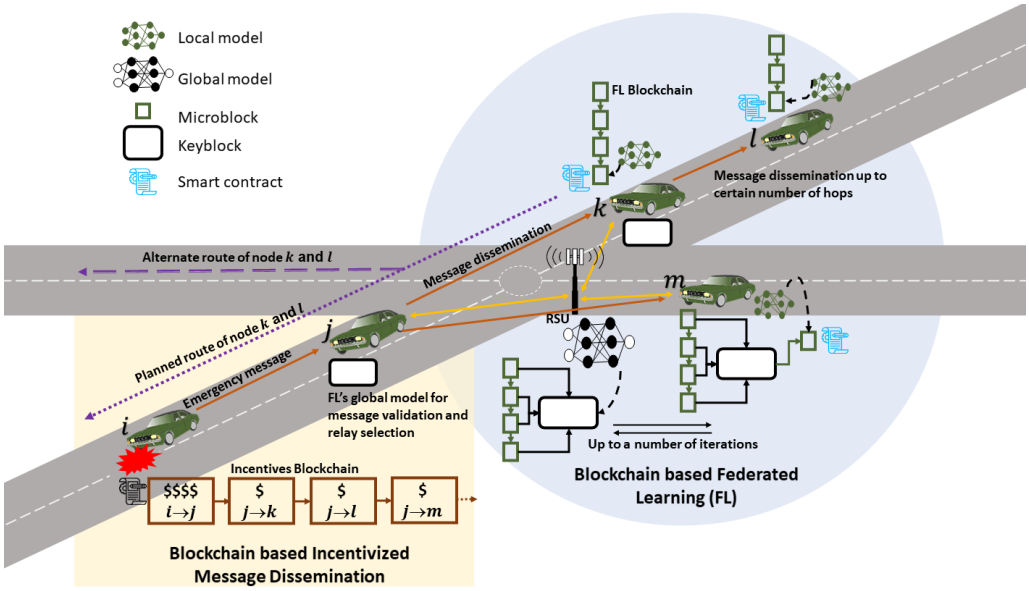


Figure 1. Blockchain-enabled FL and message dissemination in the IoV.

In this paper, we discuss the challenges and potential solutions of the blockchain and FL implementation in IoV. Then, we highlight a proposal of an integrated blockchain and FL approach for message dissemination in IoV networks and discuss its performance. Future research directions and the conclusion are provided at the end. The contributions of this paper are as follows.

- We discuss the challenges of the blockchain and FL in IoV, and highlight future research directions.
- We present an integrated solution of blockchain-empowered FL for security and privacy in the IoV. The proposed solution utilizes smart contracts and incentive transactional features of blockchain to provide security to FL.
- We computed the failure rate of the proposed solution and compared it with that of other blockchain solutions. The proposed solution resulted in a 5% reduction in failure rate as compared to other FL-integrated blockchain solutions with a high percentage of malicious nodes.

The rest of the paper is organized as follows. Section 2 individually discusses the blockchain and FL in the IoV and related works of blockchain-enabled FL. Section 3 describes the proposed solution. Results and a discussion are presented in Section 4. Section 5 concludes the paper.

2. Blockchain and Federated Learning in the IoV

2.1. Blockchain

A blockchain is typically defined as a peer-to-peer electronic cash system for recording transactions in the form of blocks linked with each other through cryptographic hash [8]. It has recently attracted attention as a potential solution to security issues in IoT. Its decentralized framework is well-suited to large-scale networks. Many blockchains incorporate cryptocurrency or virtual credits that can be utilized to manage incentives among cooperative nodes. Figure 1 shows an incentivized message dissemination solution managed by the blockchain in which the source node of a message compensates cooperative nodes that are affected by incidents with virtual credits.

Another application of the blockchain in trusted IoV networks is the evaluation of message credibility and dissemination in a decentralized manner. The blockchain consensus for validating transactions can be employed to authenticate messages from untrusted nodes. Proof-of-stake (PoS) consensus where trust ratings are stakes [9] or voting-based blockchain consensus [1] is used to measure message credibility. In addition, similar to

the computation of a mathematical puzzle in proof-of-work (PoW) consensus, the nodes perform computations to find the most appropriate relay node among themselves. An example of such computation is the quality factor (QF_i) of node i, which calculates the suitability of becoming a relay node, i.e.,

$$QF_i = DF_i \cdot F(SINR_i), \quad (1)$$

where DF_i denotes the distance from node i to its previous sender, and $F(SINR_i)$ denotes the estimated signal-to-interference-and-noise ratio (SINR) at the time when node i transmits [4].

Challenges and Potential Solutions

The metrics used to evaluate the performance of the blockchain include adversary control, which is defined as the maximal percentage of malicious nodes that it can tolerate in the network; throughput, which is defined as the number of blocks generated per unit time; and latency, which is defined as the time required to validate a transaction. If a blockchain is employed for emergency message dissemination in IoV, it must support high throughput and low latency. Famous blockchain bitcoin uses a PoW consensus, which has a latency of around 10 min [4]. Due to the short-lived connectivity among high-speed mobile nodes, PoW is not appropriate for IoV. In addition, the high computational complexity and power requirements of PoW are not suitable for on-board units (OBUs) on vehicles. Alternatives to PoW were proposed for achieving high throughput in which the PoS [9] and voting-based practical byzantine fault-tolerant (PBFT) consensus [1,8] are the most recommended for IoV. Some consensus algorithms were designed specifically for message validation and dissemination in vehicular networks, for example, joint PoS and PoW [3], and proof-of-quality factor (PoQF) [4]. Edge computing nodes that are specifically dedicated to offload complex computations can also be utilized to complete a PoW consensus.

Another challenge in the blockchain is the possible creation of forks. A fork is a block added in parallel to another block. According to the longest chain rule [4], all parallel blocks except the one connected with the longest chain are discarded. This is why forks are sometimes intentionally created by malicious nodes as an attempt to take control of the blockchain by adding further blocks connected with an invalid fork. However, in V2V communications, forks may be generated unintentionally by mobile nodes, leading to the probable deletion of valid blocks by the longest chain rule. This is because, in V2V communications, a new block announcement from a node i can only be received by other nodes that are in its transmission range. At the same time, an honest node j that is out of the transmission range of node i can create a fork by announcing another block. This problem of fork occurrence can be resolved if parallel microblocks are allowed to be generated by moving nodes with limited connectivity, and keyblocks are only added by an authoritative node, for example, RSU, as shown in Figure 1. Mobile nodes must be rendered responsible to update their copy of a blockchain. In addition, 5G- or beyond-5G-assisted IoV could potentially resolve the fork issue by utilizing a base station for the generation and announcement of blocks [2].

2.2. Federated Learning (FL)

FL has emerged as a new machine-learning approach to reach an optimized level of security and privacy with acceptable latency, communication, and computational costs. Similar to on-device learning, local models (usually deep neural network (DNN) models) are trained separately on each node with their own local data. Instead of transferring a large amount of raw data, only a trained local model is sent to a central node, called aggregator. All local models are consolidated by the aggregator, and a global model is sent back to all nodes for retraining and updating local models. This process is repeated up to

several iterations until a minimal possible loss function is attained by the global model [7]. The loss function of a global model at kth iteration is defined as follows.

$$L(w_G^k) = \frac{1}{N}\sum_{i=1}^{N} L(w_i^k), \tag{2}$$

where w_G^k are weights of the global model, and w_G^k are weights of the local model produced by node i at the kth iteration. One of the commonly used loss functions is mean squared error (MSE), which is defined as

$$MSE = \frac{1}{M}\sum_{i=1}^{M} L(y_i - y_i')^2, \tag{3}$$

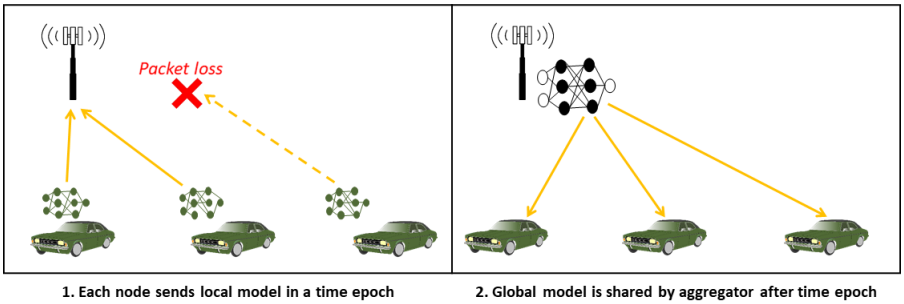
where M is the size of the test data, and y_i and y_i' are the expected and predicted outputs of a model, respectively.

FL was suggested to be a suitable technique to the fulfil privacy requirements of the IoV [10]. It is not only suitable for protecting data privacy, but also efficient in utilizing most varied computational resources and data across different nodes in 6G networks by integrating all local models. Since nodes in the IoV travel in different paths and environments, they are able to train models using diverse data, for example, varied speed limits, different propagation paths, and multiple network densities. This is why local models designed on a variety of data could potentially lead to an accurate all-encompassing global model. In [11], FL was indicated to be a solution of constrained energy operations in vehicular offloading. Apart from vehicular nodes, unmanned aerial vehicles (UAVs) were recommended to perform FL in air–ground intelligent vehicular networks in [12].

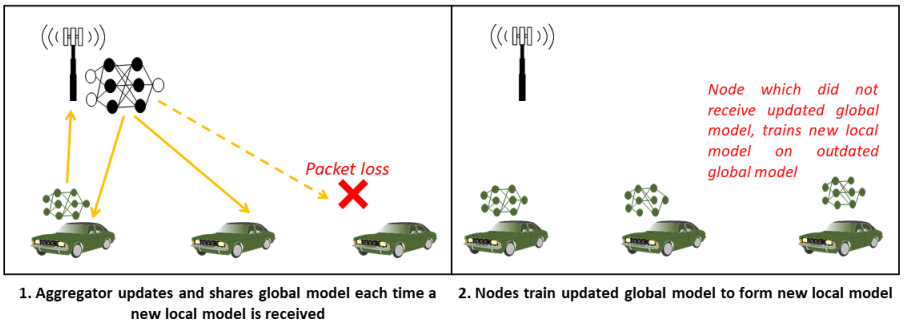
Challenges and Potential Solutions

FL can be carried out in two modes: synchronous and asynchronous [5]. In synchronous FL, nodes are given a time limit in which they submit their local models. A global model is formed after a time epoch. In asynchronous FL, nodes can send their models at their own convenience. A global model is updated each time a new model is received, thereby resulting in high communication efficiency [13]. However, a node may not be able to upload its local model in a fixed time epoch due to high mobility in a vehicular network, as shown in Figure 2a. On the other hand, in asynchronous FL in the IoV, it is possible that a node loses the connection with the aggregator and is unable to receive every update in the global model. It may continue training its local model on the basis of an outdated global model, thereby leading to wasting its resources, as shown in Figure 2b. To increase accuracy and reduce packet loss, a modified synchronous FL can be used in the IoV in which the aggregator waits for a certain number of local models to be received instead of a time epoch. A semisynchronous solution leveraging the advantages of both approaches was presented in [13].

Furthermore, since FL relies on a central aggregator, and vehicular networks tend to be decentralized, it is challenging to implement conventional FL in the IoV. Instead of one central node, multiple RSUs or base stations can perform the aggregator task. An incentive mechanism is also required to motivate nodes to contribute towards FL [7]. In an incentive distribution mechanism, nodes are more motivated to participate if their expected or predicted utility is high. Probabilistic modeling can be used to determine the expected utility of nodes prior to task initiation. It is also possible that nodes may behave maliciously to gain unfair incentives. A malicious node can use false data or deliberately produce an inaccurate local model. FL needs some means to detect malicious behavior.



(**a**) Synchronous.



(**b**) Asynchronous.

Figure 2. Modes of FL in IoV.

2.3. Related Works of Blockchain-Enabled FL

Blockchain-enabled FL is a potential solution to ensure decentralization, privacy, and security. As discussed, FL requires decentralization in IoV, an incentive distribution mechanism, and security against malicious behavior, all of which can be enabled by a blockchain. The existing literature proposed various solutions of blockchain-enabled FL in the IoV for different applications [14–16]. The communication latency challenge in IoV, consensus delays, and the unsuitability of PoW in vehicular networks were discussed in [10]. Therefore, several delay-sensitive consensuses were presented for FL in the IoV. For example, in [14], a PBFT-based blockchain-supported FL was used to ensure trust in the network. In [15], delegated PBFT (DPBFT) was proposed as the consensus, whereas FL was used for traffic flow prediction. A hierarchical blockchain for knowledge sharing was incorporated in [16], where the bottom chain was managed by mobile nodes and an RSU control top chain. All nodes, including vehicles and RSUs, participate in local model training. An AI-based proof-of-knowledge (PoK) consensus was proposed in [16].

Several related works utilized the blockchain to protect against security threats in FL. In [14], the blockchain was used to ensure decentralization and avoid single points of failure. The local models trained by individual nodes were transferred through a cloud server via edge nodes. Only local models of specific nodes (e.g., police vehicles and ambulances) were passed through PBFT consensus to guarantee their submission to cloud server, even if any edge node was faulty. In [15], a blockchain consensus provided security against poisoning attacks. In [17], a blockchain eliminated any kind of spoofing, forging, and/or reverse engineering attacks during FL in the IoV. The threat of untrusted leaders and byzantine faults during FL were resolved by the blockchain in [18].

A blockchain for securing FL is recommended not only for the IoV, but generally for all distributed frameworks incorporating mobile edge computing (MEC) as a potential machine-learning technique in 6G [19,20]. The blockchain was also recommended to improve the scalability of FL in MEC [19]. In UAVs, it was highlighted as a potential solution to secure computing in beyond-5G networks [21].

3. Proposed Methodology of Blockchain-Enabled FL

In this paper, we analyzed a smart blockchain contract to ensure automation, decentralization, and security, and evaluated the performance of the proposed solution for the application of message dissemination and relay node selection via proof-of-federated-earning (PoFL) consensus.

3.1. Smart-Contract-Based Blockchain for Incentivized FL

For smart-contract-based blockchain-enabled FL, we highlight three main steps in Figure 3. Detailed analysis is presented in [7], and we describe the steps below.

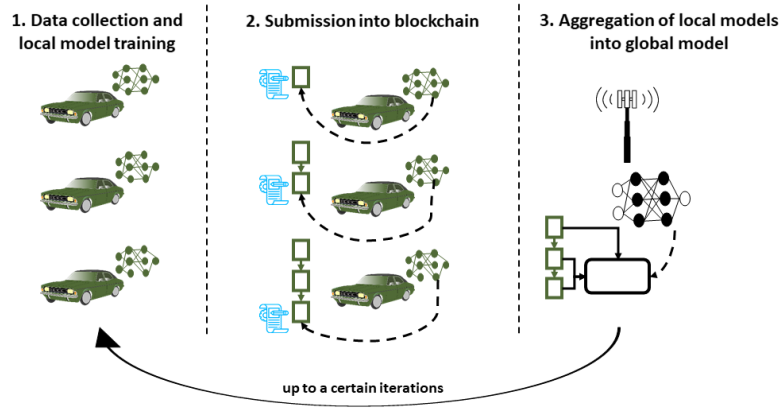


Figure 3. The proposed solution of blockchain-enabled FL in IoV.

3.1.1. Data Collection and Local Model Training

To train a local model, the nodes collect their private data while traveling on roads. The type of collected data depends on the task requirements. For example, if an FL task is aimed to estimate traffic density on roads, collected data may include the number of nodes at different routes and varied times. On the other hand, if a task requires vehicle detection and classification, images of vehicles are collected as data. When the collected data reach a required size, the nodes start to train the local model.

3.1.2. Submission into the Blockchain

During model training, a malicious node can deliberately change the training data or weights of the local model to produce inaccurate results. To protect local models from malicious attacks, they are required to be submitted into the blockchain after passing through a consensus algorithm. The consensus algorithms discussed earlier can be applied to validate a local model. For example, a local model may be considered to be credible if the trust rating of the trainer exceeds a certain threshold or the trusted neighbor nodes may check the local model with their own private data and compare them with the expected results. Alternatively, a smart contract can automate checking the local model with prerecorded data and the expected results. We propose an AI-based technique embedded in a smart contract to detect inaccuracies or anomalies in a local model. As shown in Figure 3, a customized blockchain structure consisting of parallel blocks for the submission of local models is proposed.

3.1.3. Aggregation of Local Models into a Global Model

Instead of a central node, some authorized nodes such as RSU or nodes with high trust ratings can act as aggregator. They are responsible for regularly updating their copy of blockchain until a desired number of local models are received. Among many aggregators, the one that first receives a required number of local models combines them to form a global model and add it into the blockchain as a keyblock for retraining. The creation of a global model marks the completion of a single iteration of the FL process. The proposed

steps are repeated until a specified number of iterations or the desired loss function defined in (2) are satisfied by the global model.

4. Results and Discussion

4.1. Performance Evaluation

We evaluated the performance of the proposed blockchain-enabled FL using the Tensorflow library of machine learning in Python. Data collection and blockchain updates were simulated in OMNeT++ and integrated with Simulation of Urban Mobility (SUMO). The nodes moved on a bidirectional road following the Krauss model [7] at a maximal speed of 110 km/h for 200 s in a simulated area of 10 km $\times$ 10 km. One RSU was used to aggregate the local models, and the number of mobile nodes was in the range of [10, 300]. Results were evaluated as the average of 100 simulation runs. The local models were trained on the parameters of node speed, its distance from the previous sender, moving direction, and traffic density to select a relay node in a distributed manner through the PoFL consensus.

Figure 4 shows the MSE of the global model after 100 iterations of FL calculated according to (3). Malicious nodes use false data to produce inaccurate local models. We used a machine-learning algorithm (isolation forest) in a blockchain-based smart contract to detect and reject malicious local models prior to their addition into the blockchain [7]. This led to an average reduction of 8.3% in MSE as compared to the same approach used without smart-contract-based security.

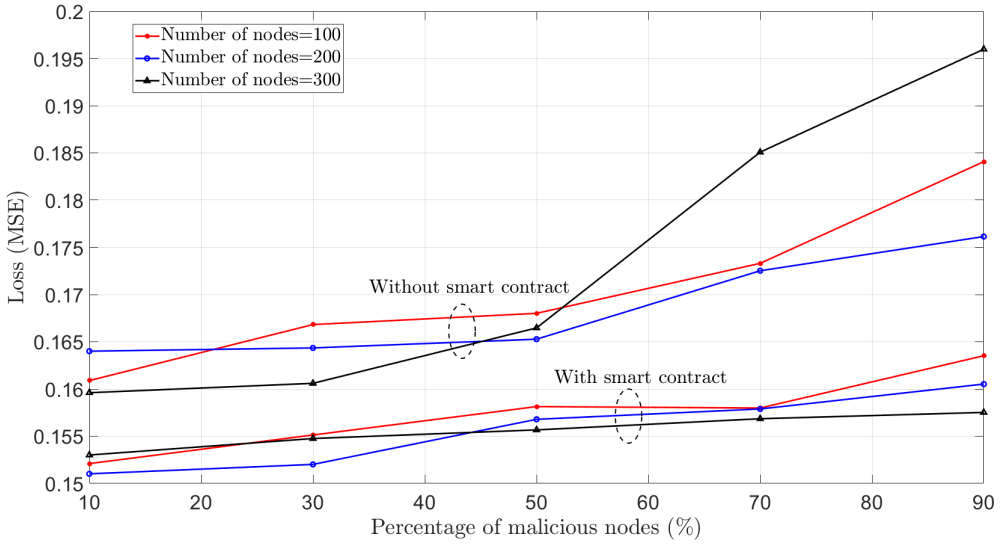


Figure 4. Loss (MSE) of global model.

Table 1 lists a comparison of various solutions for blockchain-enabled FL that shows that the smart-contract-based PoFL solution covered incentive distribution, the impact of the increase in the number of nodes, and a comparison of the solution without blockchain [7]. In [15,16] and the proposed solution, the blockchain is utilized to provide security against malicious nodes. Therefore, all local models are verified through consensus to ensure that they are not deliberately trained to produce inaccurate results. The failure rate of [15,16] and the proposed solution in the presence of malicious nodes is shown in Figure 5. Failure rate refers to the degradation in performance of a solution, i.e., inaccurate traffic flow estimation in [15], inaccurate block addition in [16], and the percentage of nodes not receiving the message in our solution. Figure 5 shows that the failure rate increased significantly with the malicious-node percentage in [16]. PoFL resulted in the least failure rate as the percentage of malicious nodes increased in the network. At 30% malicious nodes in the network, the failure rate was at least 5% lower than that of other solutions.

Table 1. Blockchain-enabled FL for vehicular networks.

Solution	Incorporated Incentive Distribution	Blockchain Structure	Adversarial Threat	Comparison without Blockchain	Application
PBFT [14]	No	Linear	None	No	Trustworthy AI
DPBFT [15]	No	Linear	Poisoning	Yes	Traffic flow prediction
PoK [16]	Yes	Hierarchical	Integrity, double spending and dishonest behavior	No	Image classification
PoFL (Proposed)	Yes	Parallel blocks for local models	Poisoning, selfish and dishonest behavior	Yes	Message dissemination

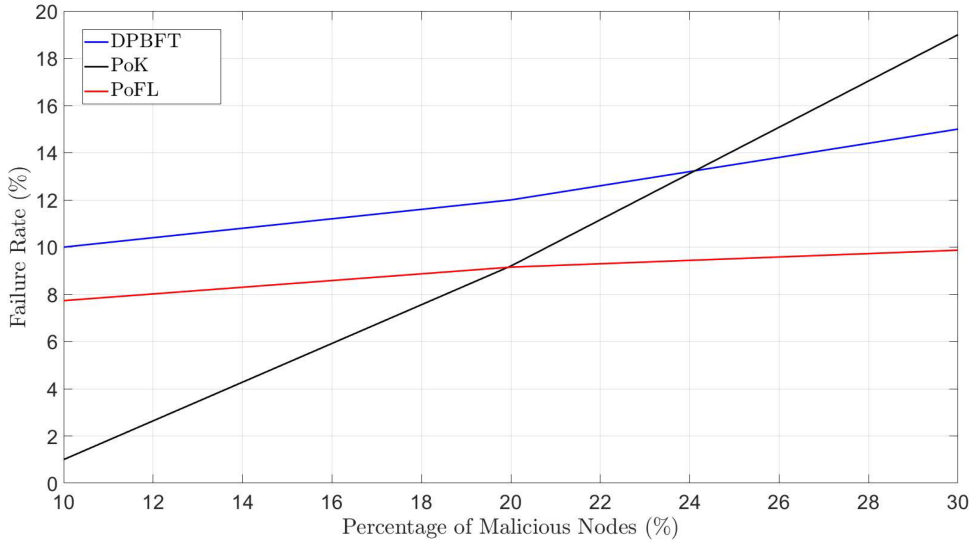


Figure 5. Failure rate in the presence of malicious nodes.

4.2. Challenges and Potential Solutions

4.2.1. Effective and Efficient Consensus

The selection of an appropriate consensus is one of the challenges of blockchain-enabled FL in the IoV. In [10], the drawback of additional delays due to blockchain management in the FL process of the IoV was discussed. A possible reason of the delay was PoW being used by nodes to record and validate local models. To reduce the delay, DPoS was used as the consensus of blockchain-supported FL in [5]. The proposed approach was faster, but decentralization was compromised because only RSUs could generate blocks. In addition, a consensus algorithm in blockchain-enabled FL requires both time efficiency and security against malicious nodes. As a solution, a machine-learning-based smart contract [7] and PoK consensus [16] were introduced to detect malicious local models. While employing AI techniques, it is important to consider that the consensus is lightweight and could meet the on-device resource and latency requirements in an IoV application.

4.2.2. Synchronized and Customized Ledger

Due to the distributed nature of the IoV, synchronization issues and forks may arise if blockchain updates are not adequately scheduled by nodes. The nodes must responsibly and regularly update their blockchain and ensure that they are training their local models on the latest global model. In addition, the regular exchange of updates related to the blockchain increases the communication cost. To reach an optimal solution, hierarchical or customized structures allowing for parallel block addition could potentially replace linear

ledgers. Examples of pioneering works include parallel off-chain blocks [7], the directed acyclic graph (DAG)-based ledger [5], and the multiple blockchain layer approach [16].

4.2.3. Data Quality and Size

The quality and quantity of collected data are crucial in determining the accuracy of a global model. Data must contain all dependent variables that affect the outcome, such as the position and speed of a node, channel quality parameters, collision probability, transmitting power, and SINR. If data collection requires cooperation from other nodes via beacon messages or acknowledgments, malicious or selfish nodes may inject false samples in the collected data, thereby leading to an inaccurate local model. In this case, it is necessary to design a fair mechanism to punish malicious nodes and incentivize honest nodes contributing to FL. Example solutions such as AI-based smart contract can provide extra robustness to ensure security without third-party dependence.

The size of the data also affects the efficiency of a global model [7]. Large data lead to increased accuracy, but take more time and energy to collect. Additionally, the accuracy of a global model is proportional to the number of nodes participating in FL [7]. However, in an incentivized FL, large data and a greater number of participating nodes may decrease their incentives. Therefore, an optimal number of nodes and data sizes for a particular FL task are required in order to attain reasonable efficiency, accuracy, and incentives. Incentive mechanisms should be analyzed with respect to all dependent parameters. A potential solution is machine-learning methods that could develop dynamic incentive mechanisms adapting to varying data sizes and numbers of nodes.

4.3. Future Directions

4.3.1. Quantum-Enhanced Blockchain

Although security is one of the prominent features of the blockchain due to its cryptographic nature, existing blockchain frameworks rely on digital signatures that are vulnerable to attack by quantum computers [22]. Therefore, robust cryptographic schemes are required to maintain blockchain security. However, the suitability of a cryptographic scheme with respect to the processing power of mobile nodes and latency requirements of the IoV must also be considered. Therefore, a future direction is to explore efficient but computationally simple quantum-resistant cryptographic schemes, such as lattice-based schemes [23] and chameleon hashes [24]. Quantum networks provide secure communications, as quantum states cannot be copied or measured without being altered, thereby preventing impersonation attacks. Hence, quantum-enhanced blockchains address the issue of security and also offer faster processing [22].

4.3.2. Modified FL Approach

Due to skewed and heterogeneous data in vehicular networks, the performance of a global model in FL may vary significantly across different nodes, environments, and traffic situations [25]. One global model may not be accurate for all traffic conditions or road maps. A selective model aggregation approach to incorporate data asymmetry was proposed to improve accuracy in vehicular networks [26]. On the other hand, the communication efficiency of FL can be improved with over-the-air (OTA) learning, where nodes send only local gradients to the aggregator instead of the large number of weights of a deep neural network. The aggregator computes both the average gradient and the new corresponding global weights for the next iteration [27].

4.3.3. Integrating Blockchain and FL with the Latest 6G Trends

The blockchain and FL frameworks are easily applicable with advanced 6G technologies. Latest trends include the integration of digital twin networks with the blockchain and FL for ITS [28], physical-layer security with blockchain for full duplex nonorthogonal multiple access (FD-NOMA)-based V2X systems [29], and reconfigurable intelligent surfaces

(RISs) with FL for a NOMA-based UAV network [30]. Therefore, both the blockchain and FL, and their combination could potentially find applications in future 6G IoV networks.

5. Conclusions

This article provided a brief overview of the blockchain, FL, and their integration to achieve security and privacy in the IoV. Taking each approach in turn, we presented the technical challenges and potential solutions, reviewing the advantages, limitations, and possible research directions in the blockchain and FL. The proposed integration of the blockchain and FL could possibly overcome limitations of current real-time vehicular systems, given that the OBUs of vehicles are capable of training and executing FL models. The suggested approaches could likely offer a way out of various security and privacy problems, but many open issues remain. On the basis of the current research progress, better solutions can be planned ahead to practically and feasibly implement the discussed approaches, and meet the growing security and privacy requirements of the IoV.

Author Contributions: All authors contributed extensively to the work presented in this paper, i.e., to the paper conceptualization, to the resources analysis, as well as to the writing, review, and editing processes. All authors have read and agreed to the published version of the manuscript.

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101006411 and Royal Society Kan Tong Po International Fellowship (KTP\R1\201007).

Data Availability Statement: Source code is available at https://zenodo.org/record/5575863#.Y0f-ZNfMKUk.

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

References

1. Ayaz, F.; Sheng, Z.; Tian, D.; Guan, Y.L.; Leung, V. A Voting Blockchain based Message Dissemination in Vehicular Ad-Hoc Networks (VANETs). In Proceedings of the International Communications Conference, Dublin, Ireland, 7–11 June 2020.
2. Cui, J.; Zhang, X.; Zhong, H.; Ying, Z.; Liu, L. RSMA: Reputation System-Based Lightweight Message Authentication Framework and Protocol for 5G-Enabled Vehicular Networks. *IEEE Internet Things J.* **2019**, *6*, 6417–6428. [CrossRef]
3. Yang, Z.; Yang, K.; Lei, L.; Zheng, K.; Leung, V.C.M. Blockchain based Decentralized Trust Management in Vehicular Networks. *IEEE Internet Things J.* **2019**, *6*, 1495–1505. [CrossRef]
4. Ayaz, F.; Sheng, Z.; Tian, D.; Guan, Y.L. A Proof-of-Quality-Factor (PoQF) based Blockchain and Edge Computing for Vehicular Message Dissemination. *IEEE Internet Things J.* **2021**, *8*, 2468–2482. [CrossRef]
5. Lu, Y.; Huang, X.; Zhang, K.; Maharjan, S.; Zhang, Y. Blockchain Empowered Asynchronous Federated Learning for Secure Data Sharing in Internet of Vehicles. *IEEE Trans. Veh. Technol.* **2020**, *69*, 4298–4311. [CrossRef]
6. Husmaldin, S.; Saeed, N. Big Data Analytics Correlation Taxonomy. *Information* **2019**, *11*, 17. [CrossRef]
7. Ayaz, F.; Sheng, Z.; Tian, D.; Guan, Y.L. A Blockchain Based Federated Learning for Message Dissemination in Vehicular Networks. *IEEE Trans. Veh. Technol.* **2022**, *71*, 1927–1940. [CrossRef]
8. Ayaz, F.; Sheng, Z.; Tian, D.; Leung, V. Blockchain-Enabled Security and Privacy for Internet-of-Vehicles. In *Internet of Vehicles and its Applications in Autonomous Driving*; Gupta, N., Prakash, A., Tripathi, R., Eds.; Springer: Cham, Switzerland, 2020; pp. 123–148.
9. Han, Q.; Yang, Z.; Ma, Z.; Li, J.; Shi, Y.; Zhang, J.; Yang, S. CMBIoV: Consensus Mechanism for Blockchain on Internet of Vehicles. In *International Conference on Blockchain and Trustworthy Systems, Proceedings of the International Conference on Blockchain and Trustworthy Systems, Dali, China, 6–7 August 2020*; Springer: Singapore, 2020.
10. Pokhrel, S.R.; Choi, J. Federated Learning with Blockchain for Autonomous Vehicles: Analysis and Design Challenges. *IEEE Trans. Commun.* **2020**, *68*, 4734–4746. [CrossRef]
11. Shinde, S.S.; Bozorgchenani, A.; Tarchi, D.; Ni, Q. On the Design of Federated Learning in Latency and Energy Constrained Computation Offloading Operations in Vehicular Edge Computing Systems. *IEEE Trans. Veh. Technol.* **2022**, *71*, 2041–2057. [CrossRef]
12. Shinde, S.S.; Tarchi, D. Towards a Novel Air–Ground Intelligent Platform for Vehicular Networks: Technologies, Scenarios, and Challenges. *Smart Cities* **2021**, *4*, 1469–1495. [CrossRef]
13. Liang, F.; Yang, Q.; Liu, R.; Wang, J.; Sato, K.; Guo, J. Semi-Synchronous Federated Learning Protocol with Dynamic Aggregation in Internet of Vehicles. *IEEE Trans. Veh. Technol.* **2022**, *71*, 4677–4691. [CrossRef]
14. Otoum, S.; Al Ridhawi, I.; Mouftah, H.T. Blockchain-Supported Federated Learning for Trustworthy Vehicular Networks. In Proceedings of the IEEE GLOBECOM, Taipei, Taiwan, 7–11 December 2020.

15. Qi, Y.; Hossain, M.S.; Nie, J.; Li, X. Privacy-preserving Blockchain-based Federated Learning for Traffic Flow Prediction. *Future Gener. Comput. Syst.* **2021**, *117*, 328–337. [CrossRef]
16. Chai, H.; Leng, S.; Chen, Y.; Zhang, K. A Hierarchical Blockchain-Enabled Federated Learning Algorithm for Knowledge Sharing in Internet of Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 3975–3986. [CrossRef]
17. Chen, J.; Chen, M.; Zeng, G.; Weng, J. BDFL: A Byzantine-Fault-Tolerance Decentralized Federated Learning Method for Autonomous Vehicle *IEEE Trans. Veh. Technol.* **2021**, *70*, 8639–8652. [CrossRef]
18. Ghimire, B.; Rawat, D.B. Secure, Privacy Preserving and Verifiable Federating Learning using Blockchain for Internet of Vehicles *IEEE Consum. Electron. Mag.* **2022**, *11*, 67–74. [CrossRef]
19. Nguyen, D.C.; Ding, M.; Pham, Q.V.; Pathirana, P.N.; Le, L.B.; Seneviratne, A.; Li, J.; Niyato, D.; Poor, H.V. Federated Learning Meets Blockchain in Edge Computing: Opportunities and Challenges. *IEEE Internet Things J.* **2021**, *8*, 12806–12825. [CrossRef]
20. Muscinelli, E.; Shinde, S.S.; Tarchi, D. Overview of Distributed Machine Learning Techniques for 6G Networks. *Algorithms* **2022**, *15*, 210. [CrossRef]
21. Saraswat, D.; Verma, A.; Bhattacharya, P.; Tanwar, S.; Sharma, G.; Bokoro, P.N.; Sharma, R. Blockchain-Based Federated Learning in UAVs Beyond 5G Networks: A Solution Taxonomy and Future Directions. *IEEE Access* **2022**, *10*, 33154–33182. [CrossRef]
22. Fedorov, A.K.; Kiktenko, E.O.; Lvovsky, A.I. Quantum Computers put Blockchain Security at Risk. Available online: https://ora.ox.ac.uk/objects/uuid:a2fdf02c-7b8a-4b38-b6b4-68b34cb636db/download_file?file_format=pdf&safe_filename=Federov%2Bedit2-AF-2-AL.pdf&type_of_work=Journal+article (accessed on 13 September 2022).
23. Alkim, E.; Ducas, L.; Pöppelmann, T. Post-Quantum Key Exchange—A New Hope. In Proceedings of the 25th USENIX Security Symposium, Austin, TX, USA, 10–12 August 2016.
24. Wu, C.; Ke, L.; Du, Y. Quantum Resistant Key-Exposure Free Chameleon Hash and Applications in Redactable Blockchain. *Inf. Sci.* **2020**, *548*, 438–449. [CrossRef]
25. Li, T.; Hu, S.; Beirami, A.; Smi, V. Ditto: Fair and Robust Federated Learning Through Personalization. In Proceedings of the 38th International Conference on Machine Learning, Virtual Event, 18–24 July 2021.
26. Ye, D.; Yu, R.; Pan, M.; Han, Z. Federated Learning in Vehicular Edge Computing: A Selective Model Aggregation Approach. *IEEE Access* **2020**, *8*, 23920–23935. [CrossRef]
27. Xue, Y.; Su, L.; Lau, V. FedOComp: Two-Timescale Online Gradient Compression for Over-the-Air Federated Learning. *IEEE Internet Things J.* **2022**, *9*, 19330–19345. [CrossRef]
28. Liu, J.; Zhang, L.; Li, C.; Bai, J.; Lv, H.; Lv, Z. Blockchain-Based Secure Communication of Intelligent Transportation Digital Twins System. *IEEE Trans. Intell. Transp. Syst.* **2022**, 1–11. [CrossRef]
29. Ayaz, F.; Sheng, Z.; Ho, I.; Tian, D.; Ding, Z. Blockchain-enabled FD-NOMA based Vehicular Network with Physical Layer Security. In Proceedings of the IEEE 95th VTC-Spring, Helsinki, Finland, 19–22 June 2022.
30. Wang, H.-F.; Huang, C.-S.; Wang, L.-C. RIS-assisted UAV Networks: Deployment Optimization with Reinforcement-Learning-Based Federated Learning. In Proceedings of the IEEE WOCC, Taipei, Taiwan, 7–8 October 2021.