Journal of King Saud University Computer and Information Sciences (2025) 37:120
https://doi.org/10.1007/s44443-025-00134-y

ORIGINAL PAPER

Blockchain-based federated learning framework for malicious node detection in internet of vehicles (IoV) networks using fog and cloud computing

Srinivas Reddy Bandarapu[1] · Muhammad Bilal[2] · Pushpalika Chatterjee[3] · Adnan Mustafa Cheema[2] · Junaid Rashid[4] · Jungeun Kim[5]

Received: 29 January 2025 / Accepted: 22 June 2025

Abstract

Due to the continuous digitalization, IoV networks are vulnerable to various communication attacks by malicious network nodes. In these attacks, the malicious entities disseminate faulty information in the network, which affects quick and intelligent decision-making in the network. Many deep learning and machine learning techniques are proposed for the classification of legitimate and malicious vehicular entities. These techniques have a centralized model training structure, which has low classification accuracy and is vulnerable to privacy leakage. To address these issues, we propose a blockchain-based federated learning framework for distributed classification of malicious and legitimate vehicles. The proposed model uses the capabilities of Long short-term memory (LSTM) and Naive Bayes (NB) for efficient and reliable malicious node detection. In our proposed model, the distributed models are trained on each locally installed virtual machine with a federated learning mechanism and then a unified model is generated at the centralized cloud server. The proposed model not only enhances the accuracy and privacy preservation but also solves the issues of centralized Internet of Vehicles (IoV) networks such as single point of failure and performance bottlenecks by utilizing the capabilities of blockchain. We used the Vehicular Reference Misbehavior (VeReMi) dataset for evaluation of our proposed model. The results show that our proposed LSTM and NB-based model outperforms centralized benchmark classification methods in malicious node detection. With an accuracy of 95%, the LSTM-based model demonstrates superior performance in identifying both malicious and legitimate vehicles, achieving a precision of 0.96 and a recall of 0.97. The high value of precision and recall shows that our model can efficiently discriminate between malicious and legitimate vehicles in the IoV network.

Keywords Blockchain · Cloud computing · Federated learning · Fog computing · Internet of vehicles

1 Introduction

In recent decades, nearly all domains, including finance, healthcare, education, and sensor networks, have undergone significant transformation due to advancements in tracking technologies using AI models (Nguyen et al. 2024) and the evolution of secure, trust-aware systems in vehicular networks (Alalwany and Mahgoub 2024). Similarly, in the last ten years, the industry of automobile has changed to a large extent due to the high demand for smart cities (Hemmati et al. 2024). The automobile industry tries to ensure the safety of users while simultaneously minimizing their overheads (Sun et al. 2024a). In this regard, the Internet of Vehicles (IoVs) plays an important role, it utilizes the data of various network entities and ensures quick and intelligent decision-making

✉ Junaid Rashid
junaid.rashid@sejong.ac.kr

✉ Jungeun Kim
jekim@inha.ac.kr

1 DigiTech Labs, 15915 NE 117th Way, Redmond, WA 98052, USA

2 Department of Information Technology, Rawalpindi Women University, Rawalpindi, Pakistan

3 The Huntington National Bank, Columbus, Georgia 43074, USA

4 Department of Artificial Intelligence and Data Science, Sejong University, Seoul, Republic of Korea

5 Department of Computer Science and Engineering, Inha University, Incheon, Republic of Korea

Published online: 21 July 2025

(Guo et al. 2024). In this way, the IoV network not only provides a secure and reliable network for users but also ensures the optimization of resources (Zou et al. 2024). The IoV network can organize itself on its own and ensures that the data is transferable to each network entity in real-time, which ultimately helps in the optimization of resources of the network and the safety of users (Shinde et al. 2024). However, the IoV network faces many issues such as privacy leaks during continuous data transfer processes and the presence of malicious nodes (Khezri et al. 2025). Malicious nodes can cause severe attacks and attempt to dominate the vehicular network (Miao et al. 2024). These malicious nodes disseminate faulty information in the network, which ultimately affects the decision-making process in the network (Wu et al. 2024). Hence, it is the reason that the internal vehicular entities are not comfortable sharing their sensitive information in the network (Yan et al. 2024). On the other hand, external users are also reluctant to join such networks, as their critical information and sensitive records may be exposed to risk due to faulty message broadcasting caused by malicious nodes in the network (Yang et al. 2024).

Many techniques are proposed for identifying and removing internal malicious nodes from the IoV network to solve the issues of privacy leakage and network data loss. A context-aware protocol is proposed to collect various parameters from the vehicular network (Rehman et al. 2022). After this, the artificial intelligence-enabled context awareness technique is used to create an environment for the event. This context-based event is then analyzed to evaluate the true value of network entities, while simultaneously ensuring multi-level uncertainty handling. Beside this, a multi-layer trust management scheme is proposed, where the trust of each vehicular entity is evaluated at the transport layer (Ahmad et al. 2021). However, these techniques have a centralized structure and all their operations are dependent on a centralized third party. Therefore, these malicious node detection mechanisms are vulnerable to different issues such as single point of failure, high network overhead, large monetary cost, and performance bottlenecks.

Deep learning (Ahmed et al. 2021) and machine learning techniques (Wang et al. 2022) are also proposed for the identification and removal of malicious nodes in the IoV network. These techniques are efficient in the classification of malicious and legitimate nodes and can secure the IoV network against various attacks possessed by malicious nodes. However, these machine learning and deep learning techniques are involved with centralized model training. The actual data of vehicular network entities is shared with remote network units, which causes the issues of privacy leakage. Therefore, the centralized model training techniques not only compromise the accuracy of the classification process but also cause privacy leakage of network entities.

Blockchain is a decentralized protocol (Nakamoto 2008), and all the entities either they are seller or buyer, act as the foundation of the network. Furthermore, blockchain is immutable, and all entries recorded on the blockchain network cannot be forged or tampered (Mulligan et al. 2024). Therefore, any malicious entity in the network cannot manipulate the data, and network vehicles can be easily authenticated without large network overhead, which ultimately enhances the trust and security of the overall network. Furthermore, the blockchain is benefited by the large size of the network, which ultimately enhances the overall security of the network through decentralization (Popoola et al. 2024). When the number of nodes in the blockchain network is high then it is very difficult for any malicious node to change the history of the network. In this way, the blockchain network can ensure transparency in a network while simultaneously enhancing the trust between unknown and dispersed entities (Liang et al. 2024). The adoption of blockchain by a firm can play a pivotal role in shaping its innovation strategy and enhancing its environmental performance, particularly in the context of sustainability-oriented technological advancement (Zhu et al. 2024). Nowadays, there is an increasing threat of malicious node activity in IoV networks and the limitations of centralized intrusion detection systems. Furthermore, traditional systems are vulnerable to scalability issues, privacy violations, and high latency. To solve these aforementioned issues, we utilize the capabilities of federated learning over fog nodes, which enables distributed training close to data sources, preserving user privacy, and ensuring faster detection of anomalies. Furthermore, the LSTM classifier is combined with Naive Baye, which uses temporal patterns in the data and maintains computational feasibility on resource-constrained vehicular nodes. This hybrid model enhances detection accuracy by leveraging both deep learning and probabilistic reasoning. The fog-based architecture minimizes response time by processing data locally before cloud aggregation. Blockchain integration ensures tamper-proof model updates and secures inter-node communication. Overall, the system achieves a balance between accuracy, privacy, and efficiency suitable for dynamic IoV environments. Therefore, in this paper, blockchain-based federated learning framework is proposed to solve the issues caused by third-party involvement. Its utilizes the capabilities of Long short-term memory (LSTM) and Naive Bayes (NB) classification algorithms for efficient and reliable identification of malicious nodes while simultaneously solving the issues associated with centralized model training. We also used fog and cloud computing in our proposed malicious

node detection model to enhance its efficiency and scalability. Fog computing enables real-time data processing close to the source, such as at roadside units or sink nodes, reducing latency and ensuring quicker decision-making. The contributions of our paper are given as follows.

- A blockchain-based federated learning mechanism is proposed for classification of malicious and legitimate nodes while utilizing the capabilities of LSTM and NB algorithm and solving issues associated with centralized model training.
- A blockchain-based architecture is introduced for the IoV network to mitigate issues such as large network overhead, single points of failure, and latency caused by third-party involvement.
- The proposed model incorporates fog and cloud computing to enable real-time data processing and scalable network data storage, thereby ensuring low-latency detection of malicious nodes in IoV networks.

2 Related work

The Internet of Vehicles (IoV) shares a vast amount of data within the IoV ecosystem to support decision-making and data trading (Jabbar et al. 2022; Bhattacharya et al. 2022). The functionalities of the IoV network are optimized with their integration with cloud servers and edge computing. However, autonomous vehicles in IoV networks are vulnerable to various security attacks, as malicious nodes can easily gain access to network data (Kebande et al. 2021). Therefore, there is a need for an efficient authentication mechanism that ensures access control so that only authorized and legitimate nodes can access and process the data. They propose a blockchain-based decentralized authentication process to solve the issue of unauthorized access. The digital signatures are used by the proposed model to authenticate cloud servers and IoVs in the network. The authentication process is based on various factors, which ultimately help in hardening the authentication process. Different aspects like single sign-on and markup language are used in the proposed model to further enhance the authentication process of IoVs. To evaluate the performance of the proposed authentication scheme, the authors conducted a security strength analysis. The results show that all the principles of data integrity and confidentiality are satisfied by the proposed scheme. Furthermore, a probability-based time algorithm is also used in the proposed model to eliminate the weaknesses of traditional authentication models. In the proposed multi-factor authentication, the users need to present two sets of credentials for logging into the system. First of all, the first set of credentials is matched with the already stored credential, if the match is successful then the user is allowed to enter the PIN or password. In this authentication, the context and behavior of transmitted data are also considered for providing a secure and reliable access control system. First of all, the vehicles sense the data from the network and then send the ID of the cloud server to which they are associated. This cloud server receives the data from vehicles and aggregates that data. Furthermore, this cloud server is also responsible for establishing the between vehicles and clouds. When a vehicle joins the network, then this vehicle is asked to get registration in the network by generating a hash chain. After that, a one-time password is used by vehicles to update the hash chain to establish a secure connection with the cloud server. Lastly, the vehicles and clouds can easily share their data through this secure channel. All data trading transactions are validated by miner nodes and then added to the immutable digital ledger. The proposed model is evaluated regarding confidentiality, data integrity, percentage of distributed attacks, and number of authenticated nodes. In the future, the authors are interested in studying different use cases to design a multifactor efficient authentication process, which can use attribution processes and behavioral models.

Similarly, Hîrţan et al. (2020) identified that some malicious entities in vehicular networks share falsified data to achieve their goals. It is very crucial to rely on this compromised data to make any decision in real-time. Therefore the authors propose a mechanism for calculating the reputation of nodes in the vehicular network. Furthermore, the proposed model also ensures the confidentiality of data by using public and private keys. A pair of public and private keys is generated by public key infrastructure. The vehicle that wants to share its data over the network encrypts the data with the public key of the receiver vehicle or roadside unit. After encryption, the plain readable text is converted into ciphertext, which is then transmitted over the network. This ciphertext is not readable by any intermediate node, as these nodes do not possess the decryption key. This solves the issue of man in the middle attack. The cipher text can only be decrypted by the private key of the receiver. It is the reason that only the receiver can decrypt and use the sent data. In this way, the proposed mechanism helps achieve data confidentiality. The authors state that there are many existing schemes that can establish trust in the network without considering the privacy of users. However, the model proposed by the authors not only ensures trust within the vehicular network but also preserves the confidentiality of users identities and data. The reputation system is entirely dependent on blockchain technology, as users can retrieve data from network vehicles and similarly transmit their sensed data over the network. Each data trading is tracked and validated by blockchain high-resourced miner nodes. After validation of each transaction, the block is added to the blockchain, and reputation is given to the respective sender node. Due to the integration of blockchain and reputation system, the network data

is reliable and it is very convenient to make a final decision after a traffic incident while depending upon that trustworthy data. The proposed model is evaluated by simulating the proposed model of real-time data of Beijing, San Francisco, and Rome. The proposed model is evaluated while considering the parameters of the percentage of blocked intersections, blockchain storage size, total number of cluster messages, number of cars, and network overhead. In the future, the system model can be extended to accurately identify the GPS of users having more impact over the network. Furthermore, the proposed model can be evaluated by conducting the experiments while considering more parameters.

It is very challenging for vehicular networks to handle traffic events and ensure non-repudiated service provisioning (Ning et al. 2021). Furthermore, existing schemes fail to guarantee user data security and introduce significant network delays during the computation process. The authors mention that it is very important to ensure network data security and efficient and reliable transmission of data for the progress of the IoVs ecosystem. Therefore, the authors propose a blockchain-based decentralized framework to reduce system latency while simultaneously ensuring network data security. Spilt and integration optimization objectives are considered to divide the problem into two parts to provide an efficient solution. In vehicular networks, there is a tradeoff between network latency due to the integration of blockchain and the security of network data. The first part of the problem is used to minimize this tradeoff to provide a suitable and reliable solution for both data security and network latency. Furthermore, reinforcement learning with the integration of deep learning technology is also used in the network to solve identified issues by selecting reliable and trustworthy miner nodes from a set of roadside units. In the second part of the problem, the authors are further interested in optimizing user comfort by maximizing the overall user comfort of all users in the network. Moreover, a distributed algorithm is used for the selection of the most reliable computation modes to perform blockchain high computation tasks in a distributed manner. The proposed model is evaluated by conducting several experiments and it is observed from the results that the proposed model is capable of selecting the miner nodes transaction validation. Furthermore, it is also observed that the proposed model is efficiently securing the data while simultaneously solving the issue of latency in the network. Besides this, the proposed model is evaluated in terms of reputation value, grade value, iteration index, task ratio, CPU frequency, number of reliable miner nodes, social value, network overhead, and total computing power. The proposed model solves the issues of lack of data security and network latency simultaneously. However, the proposed model uses a crowdsourcing framework for reliable communication in the IoV ecosystem, which is not suitable to perform all types of tasks in the network. There is a need for domain experts to complete complex and specialized tasks. It is the reason that the problem-solving scope of this technology is limited to a large extent.

Besides this, identify that all the existing studies use different human-centric visual data for the detection of human-object interaction (Ozaki et al. 2024). All these studies are mainly focused on improving the accuracy of the method while utilizing widely used open datasets. These methods are not adaptable to different fields of life and do not have any robustness. Furthermore, the existing models fail to provide efficient services because these models do not have high extensibility and are not speedy and efficient, which ultimately results in large computational overhead. To solve these aforementioned issues, a model is proposed that combines the object detection mechanism with the skeletal method for predicting the human and different objects from the visual data while simultaneously achieving high extensibility and robustness (Ozaki et al. 2024). The proposed model is adaptable to various fields of life as it is lightweight and does not require a huge volume of resources to be operated. Initially, the YOLOv5 object detection method and the MediaPipe estimation algorithm are used to extract object features and human postures, respectively. The proposed model uses logistic regression for the extraction of the most relevant features instead of all feature extraction performed by all other deep learning models. The proposed model does not require a large visual image or video dataset for learning the process of human interaction because it uses coordinates of objects and human position landmarks. This is the reason the proposed model is efficient and lightweight in extracting the features of human postures and distinct objects. Furthermore, the robustness of the model is increased by the prediction of object and human features with coordinates. After that, the prediction probability of each class is calculated by the model, and the final prediction is performed based on high probability. The model includes two convolutional layers: convolution layer A is used to estimate the coordinates of the object bounding box, while convolution layer B is responsible for estimating human posture landmarks. The proposed model is evaluated in terms of accuracy, robustness, ablation study, recognition speed, and computational overhead. The proposed model is compared with different state-of-the-art deep learning mechanisms in terms of robustness and computational overhead. The experimental results show that the proposed model outperforms all deep learning models with high robustness and low computational overhead. In the future, the authors will extend the proposed system model with an application that can learn the spatial utilization and all characteristics of users.

Recent advances show the significance of federated learning (FL) in ensuring data privacy and decentralization in intelligent transportation systems (ITS). Several studies such as Ullah et al. (2025); Xie et al. (2024); Korba et al. (2024)

have employed FL for vehicular data sharing to minimize the risk of centralized points of failure. Similarly, the advent of fog computing has demonstrated substantial advantages in enabling real-time data processing near the data source (Lingamallu et al. 2024; Liu et al. 2024; Gu et al. 2024), thereby reducing latency and enhancing system scalability. To address security challenges on the Internet of Vehicles (IoV), Ullah et al. (2025) proposed a blockchain-based federated learning framework for intrusion detection, which facilitates decentralized model training while preserving data privacy. Their integration of blockchain ensures integrity and traceability in model updates. However, the reliance on a central aggregator introduces a potential single point of failure, and the blockchain component incurs notable computational overhead, particularly under non-independent and identically distributed (non-IID) data conditions. In a related study, Xie et al. (2024) introduced IoV-BCFL, a hybrid architecture combining FL for local intrusion detection and blockchain for secure parameter exchange. This framework strengthens both model security and data confidentiality in decentralized vehicular networks. Nonetheless, the approach faces practical challenges due to limited computational resources on vehicles and the latency introduced by blockchain consensus mechanisms, which may impair real-time responsiveness. Furthermore, system scalability also becomes an issue as the number of participating vehicles increases. Korba et al. (2024) presented Zero-X, a blockchain-enabled open-set federated learning system for detecting zero-day attacks in IoV environments. By integrating deep neural networks with a novel Proof-of-Accuracy consensus mechanism, the framework supports secure and decentralized model training. While the design ensures robustness against novel threats, the blockchain consensus process introduces computational and communication overhead at MEC nodes, potentially affecting scalability. Although CAVs operate with unlabeled benign data, SOCs require labeled attack samples, which may limit adaptability in rapidly evolving threat landscapes without adequate threat intelligence.

Table 1 presents a critical analysis of the literature, highlighting the strengths, limitations, and key features of existing approaches. It shows that most models rely heavily on centralized architectures and often lack privacy-preserving mechanisms. The proposed model leverages the capabilities of Long Short-Term Memory (LSTM) to capture sequential malicious behaviors, combined with the computational simplicity of Naive Bayes, all within a federated learning architecture operates over fog nodes. This design enables effective handling of heterogeneous and distributed data while preserving data privacy and minimizing dependence on centralized servers. Accordingly, proposed model constitutes

Table 1 Comprehensive analysis of existing approaches: challenges, innovations, evaluations, and future directions

Addressed Limitations	Proposed Contributions	Performance Parameters	Future Work/ Limitation
The IoVs networks are vulnerable to various security attacks; the malicious nodes can easily access the networks' data (Kebande et al. 2021)	Blockchain-based multi-factor authentication for access control and probabilistic algorithm to remove traditional mechanism weaknesses	Confidentiality, data integrity, percentage of distributed attacks, and number of authenticated nodes	Different use cases will be studied to design a multifactor efficient authentication process, which can use attribution processes and behavioral models
Malicious entities in vehicular networks share falsified data. It is very dangerous to rely on this compromised data to make any decision in real-time (Hîrţan et al. 2020)	Blockchain-based reputation mechanism not only guarantees trust in the vehicular network but also ensures the confidentiality of users' identity and data	Percentage of blocked intersections, blockchain storage size, total number of cluster messages, number of cars, and network overhead	System model can be extended to accurately identify the GPS and model can be evaluated by conducting experiments
Existing schemes are unable to ensure data security of users and cause large network delays in computation processes (Ning et al. 2021)	Multi-objective optimization is proposed for solving the issues of network latency and data security, and a deep learning algorithm is proposed to select the most authentic nodes for task handling	Reputation value, grade value, iteration index, task ratio, CPU frequency, number of reliable miner nodes, social value, network overhead, and total computing power	The problem-solving scope of this crowdsourcing is limited as domain experts are not suitable for this technology
Existing deep learning methods for human-object interaction are not adaptable, robust, efficient, and extensible (Ozaki et al. 2024)	Hybrid model of object detection mechanism with the skeletal method for predicting the human and different objects, YOLOv5 object detection method and eMediaPi estimation algorithms are used for feature extraction	Accuracy, robustness, ablation study, recognition speed, and computational overhead	An application will be developed for learning spatial utilization and all characteristics of users

a novel integration of low-latency, privacy-aware learning for the IoV, where both detection performance and data confidentiality are critical. The model introduces a hybrid federated learning technique using LSTM and Naive Bayes for malicious node detection in IoV, uniquely combining deep temporal learning with lightweight probabilistic classification. Unlike previous works, it adopts a layered fog–cloud architecture that reduces latency and supports real-time model updates. Blockchain with Proof of Authority and Keccak-256 hashing ensures secure, low-overhead validation and decentralized trust. By sharing only trained models, our model preserves privacy and avoids data leakage common in centralized methods. This integration addresses scalability, computational efficiency, and privacy, challenges that previous approaches could not fully overcome.

3 Proposed model

3.1 Network architecture

The IoV network is composed of various kinds of entities such as ordinary sensing vehicles, roadside units, and sink nodes. These entities are deployed in the network based on their computational capabilities and defined roles. The ordinary sensing vehicles are mobile in the network and responsible for sensing and sharing data with their precise coordinates to roadside units. The roadside unit units receive this data from ordinary sensing vehicles and process it. All the duplicate and redundant values are removed from the data. After that, this data is stored on roadside units for quick and intelligent decision-making in the network. Since roadside units lack substantial processing and storage capabilities, their data is transmitted to a centralized sink node. Each sink node is responsible for managing the data of a specific vehicular region. The sink node collects the data from various roadside units and further processes it. Both the roadside units and sink nodes ensure on-demand safe and reliable delivery of data to each entity in the network, which is then used for quick and intelligent decision-making. The IoV network is self-organizing and ensures real-time data transfer among network entities, ultimately contributing to resource optimization and enhanced user safety. However, the IoV networks are vulnerable to the issue of faulty message dissemination by malicious network entities, which ultimately affects the decision-making process.

Many deep learning and machine learning models are also proposed for malicious node detection in which centralized model training is being performed (Ahmed et al. 2021; Wang et al. 2022). The actual data of vehicular network entities is shared with a central network server, which is responsible for the training and testing of classification models based on provided features. Due to this, the privacy of network entities is compromised. Furthermore, the centralized servers are not able to ensure high classification accuracy as these servers are not able to efficiently process large-scale, diverse, and dynamic data. Moreover, these centralized training servers are vulnerable to issues such as single points of failure and model overfitting. Therefore, our proposed blockchain-based federated learning model utilizes the capabilities of LSTM and Naive Bayes NB classification algorithm for efficient and reliable identification of malicious nodes while simultaneously preserving the privacy of network entities and enhancing classification accuracy. Following are some assumptions on which our proposed proposed model is dependent.

- The ordinary sensing vehicles continuously enter and leave the sub-networks.
- All the roadside units and sink nodes are stationary and considered to be legitimate.
- The sink nodes generates cryptographic keys of ordinary vehicles while utilizing the capabilities of public key infrastructure.
- The legitimate ordinary sensing vehicles, sink nodes, and roadside units have their unique Ethereum addresses.
- The sink nodes are responsible for permanent data storage and data aggregation.

3.2 Proposed blockchain-based federated learning framework for IoV network

In our proposed blockchain-based federated learning model, the vehicular network is divided into various sub-networks. Each sub-network contains stationary roadside units and sink nodes. Further, ordinary sensing vehicles are mobile and they continuously enter and leave the network. The malicious ordinary vehicles try to overcome the network by disseminating faulty messages (Kharche et al. 2024). Three heterogeneous sub-networks are considered in our proposed blockchain-based federated learning framework for the detection and removal of malicious vehicular entities, as shown in Fig. 1. The data of each sub-network is collected by the respective roadside unit and processed. After that, this data is sent to the sink nodes for further refinement of data. Each sink node is associated with a particular virtual machine. The sink node processes the data and sends it to the virtual machine associated with it for local model training. The virtual machine initially collects the data and then trains a local model based on features provided by the sink nodes. In last, the data is sent to the cloud server to generate a unified model. The cloud server is the highly resource-enriched node and is responsible for the collection of data from the locally trained model and their fusion for the generation of the unified central model. The cloud server generates a fused model and sends it back to each virtual machine associated with a respective sink node.

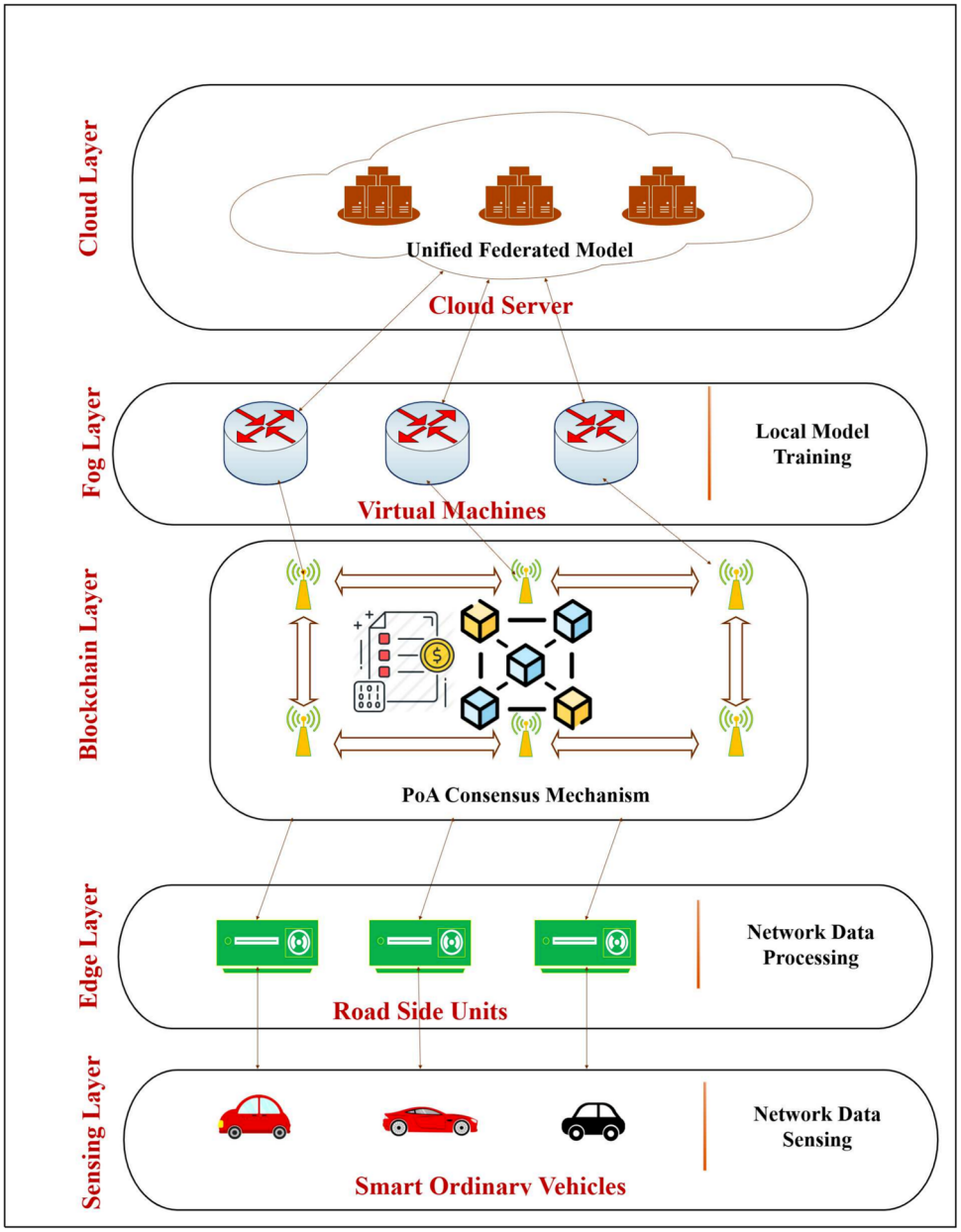

Fig. 1 Proposed blockchain-based federated learning framework for internet of vehicle networks

The virtual machines collect the unified model from the cloud server and classify the legitimate and malicious vehicles of their respective sink nodes. In our proposed model, the locally trained models are shared with the central cloud server instead of the actual data of each sink node. In this way, the significant features that can reveal sensitive and important information about network vehicular entities are not shared outside the respective sub-network, which ultimately helps preserve the privacy of network vehicular entities (Jain and Khare 2024). Furthermore, the accuracy of the malicious node classification process is also enhanced. The reason is that various local models trained by rich and significant features of dispersed vehicular entities are shared with cloud servers. As a result, a robust and accurate model is generated, enabling efficient and reliable classification of malicious and legitimate vehicular entities. Figure 1 shows the layered architecture of our proposed federated learning model. The initial layer of our model is the sensing layer in which different ordinary sensing vehicles are moving and responsible for sensing various network parameters such as traffic congestion, road conditions, temperature, humidity, etc. The second layer is the edge layer in which stationary roadside units are placed that collect the data of sensing vehicles and process it (Xu et al. 2024). Then there is a blockchain having sink nodes that are responsible for the generation of public and private keys and transaction validation by utilizing the

capabilities of the proof of authority (PoA) consensus algorithm and blockchain. The fourth fog layer has various virtual machines are installed for local model training from various features of associated sink nodes. The last layer cloud layer in which various models are trained locally and their fusion. Instead of actual data of each sink node, the locally trained models are shared with the central cloud server.

Algorithm 1 shows the step-by-step implementation process of the proposed blockchain-based federated learning framework model designed for IoV networks.

- The input to the proposed algorithm includes several sets: $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ represents the vehicles in the network, $\mathcal{R} = \{r_1, r_2, \ldots, r_k\}$ refers to roadside units (RSUs), $\mathcal{S} = \{s_1, s_2, \ldots, s_m\}$ denotes sink nodes, $\mathcal{F} = \{f_1, f_2, \ldots, f_m\}$ represents virtual machines, and $\mathcal{C}$ is the cloud server. The goal is to classify vehicles as malicious ($\mathcal{M}$) or legitimate ($\mathcal{L}$) and generate a unified global model $\mathcal{M}_{global}$.
- The process begins in the sensing layer, where each vehicle $v_i \in \mathcal{V}$ senses the network parameters $\mathcal{P} = \{p_1, p_2, \ldots, p_l\}$. These parameters may include traffic congestion, road conditions, temperature, and humidity. The data sensed by each vehicle, represented as $\mathcal{D}_i$, is transmitted to the nearest roadside unit $r_j \in \mathcal{R}$ for initial processing.
- In the edge layer, each roadside unit r_j aggregates data from vehicles within its range, forming a dataset $\mathcal{D}_j = \bigcup_i \mathcal{D}_i$. This refined data, $\mathcal{D}'_j$, is then transmitted to its associated sink node $s_j \in \mathcal{S}$ for further processing and validation (Matei and Cocoşatu 2024).
- At the blockchain layer, each sink node s_j generates public and private keys, $\mathcal{K}_{pub}$ and $\mathcal{K}_{priv}$, respectively. Using the proof of authority (PoA) consensus algorithm, the sink node validates transactions to ensure data integrity. Once validated, the refined dataset $\mathcal{D}''_j$ is sent to its corresponding virtual machine $f_j \in \mathcal{F}$ for local model training.
- At the fog layer, virtual machines perform feature extraction, $\mathcal{F}_j = \text{FeatureExtraction}(\mathcal{D}''_j)$, to identify significant attributes. Each virtual machine f_j then trains a local model $\mathcal{M}^j_{local}$ using LSTM and NB algorithms. The locally trained models $\mathcal{M}^j_{local}$ are transmitted to the cloud server $\mathcal{C}$ instead of the raw data to preserve privacy (Gautam et al. 2024).
- The cloud layer aggregates all locally trained models, $\mathcal{M}_{local} = \{\mathcal{M}^1_{local}, \mathcal{M}^2_{local}, \ldots, \mathcal{M}^m_{local}\}$, and fuses them to generate a unified global model $\mathcal{M}_{global} = \text{Fusion}(\mathcal{M}_{local})$. The fused model $\mathcal{M}_{global}$ is then broadcast back to each virtual machine f_j.
- Each virtual machine f_j uses the unified model $\mathcal{M}_{global}$ to classify vehicles as legitimate ($\mathcal{L}_j = \{v_i | \mathcal{M}_{global}(v_i) = 1\}$) or malicious ($\mathcal{M}_j = \{v_i | \mathcal{M}_{global}(v_i) = 0\}$). This classification is performed locally for each sub-network.
- The algorithm ensures privacy by only sharing locally trained models with the cloud server, not the actual datasets. This approach maintains the confidentiality of vehicular network entities sensitive information. Additionally, sharing models instead of raw data enhances the accuracy of malicious node classification, as diverse and rich features from multiple sub-networks contribute to the global model.
- The hierarchical layered architecture, including sensing, edge, blockchain, fog, and cloud layers, enables efficient detection and removal of malicious vehicular entities while maintaining data privacy.

Algorithm 1 Blockchain-based federated learning model for IoV network.

Input: $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ (vehicular entities), $\mathcal{R} = \{r_1, r_2, \ldots, r_k\}$ (roadside units), $\mathcal{S} = \{s_1, s_2, \ldots, s_m\}$ (sink nodes), $\mathcal{F} = \{f_1, f_2, \ldots, f_m\}$ (virtual machines), $\mathcal{C}$ (cloud server).

Output: $\mathcal{M}_{global}$ (unified model), classification of malicious ($\mathcal{M}$) and legitimate ($\mathcal{L}$) vehicles.

Initialize $\mathcal{M}^i_{local}$ for each $f_i \in \mathcal{F}$.
Layer 1: Sensing Layer
for each $v_i \in \mathcal{V}$ **do**
  v_i senses network parameters $\mathcal{P} = \{p_1, p_2, \ldots, p_l\}$.
  Data $\mathcal{D}_i$ is transmitted to nearest roadside unit r_j.
end for
Layer 2: Edge Layer
for each $r_j \in \mathcal{R}$ **do**
  Aggregate data $\mathcal{D}_j = \bigcup_i \mathcal{D}_i$ from vehicles.
  Transmit refined data $\mathcal{D}'_j$ to corresponding sink node s_j.
end for
Layer 3: Blockchain Layer
for each $s_j \in \mathcal{S}$ **do**
  Generate public and private keys $\mathcal{K}_{pub}, \mathcal{K}_{priv}$.
  Validate transactions using PoA consensus algorithm.
  Transmit validated data $\mathcal{D}''_j$ to corresponding f_j.
end for
Layer 4: Fog Layer
for each $f_j \in \mathcal{F}$ **do**
  Extract features $\mathcal{F}_j = \text{FeatureExtraction}(\mathcal{D}''_j)$.
  Train local model $\mathcal{M}^j_{local}$ using LSTM and NB.
  Transmit $\mathcal{M}^j_{local}$ to $\mathcal{C}$.
end for
Layer 5: Cloud Layer
Aggregate models $\mathcal{M}_{local} = \{\mathcal{M}^1_{local}, \mathcal{M}^2_{local}, \ldots, \mathcal{M}^m_{local}\}$.
Fuse models to generate global model $\mathcal{M}_{global} = \text{Fusion}(\mathcal{M}_{local})$.
Broadcast $\mathcal{M}_{global}$ to all $f_j \in \mathcal{F}$.
Malicious Vehicle Detection
for each $f_j \in \mathcal{F}$ **do**
  Classify vehicles: $\mathcal{L}_j = \{v_i | \mathcal{M}_{global}(v_i) = 1\}$.
  $\mathcal{M}_j = \{v_i | \mathcal{M}_{global}(v_i) = 0\}$.
end for
Preservation of Privacy
Ensure no actual data $\mathcal{D}_j$ is shared with $\mathcal{C}$, only $\mathcal{M}_{local}$.
End of Algorithm

3.3 LSTM and NB classification techniques

In our proposed model, the LSTM and NB classification techniques are used in virtual machines associated with sink nodes for the classification of legitimate and malicious vehicular entities.

LSTM classification technique is efefctive for sequential data such as end-to-end delay and time series data of vehicular networks and can capture temporal dependencies in the communication among various network entities. Multiple memory cells and gating mechanisms are used for capturing the temporal patterns of vehicular communications. At each time step t, the LSTM operations are governed by a series of equations. The forget gate determines the discared information from the previous cell state C_{t-1}, using the activation function σ to regulate this process, as given in (1).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad \text{(forget gate)} \tag{1}$$

The input gate identifies the new information to be added to the cell state, using the input (x_t) and the hidden state from the previous time step (h_{t-1}), as given in (2).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad \text{(input gate)} \tag{2}$$

Simultaneously, the candidate cell state is computed, representing a potential update to the cell state, as given in (3).

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \quad \text{(candidate cell state)} \tag{3}$$

The updated cell state is then calculated by combining the results of the forget gate and input gate, allowing the model to retain relevant information, as given in (4).

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \quad \text{(cell state update)} \tag{4}$$

The output gate determines the final output of the LSTM cell by regulating the hidden state, as given in (5).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad \text{(output gate)} \tag{5}$$

Finally, the hidden state output is computed by combining the output gate with the updated cell state, and this information is passed to the next time step, as given in (6).

$$h_t = o_t \odot \tanh(C_t), \quad \text{(hidden state output)} \tag{6}$$

Where, f_t, i_t, o_t are forget, input, and output gates, respectively, at time step t. C_t shows the cell state at time t, h_t represents the hidden state output.The x_t is input vector at time t, W_f, W_i, W_C, W_o are the weight matrices, and b_f, b_i, b_C, b_o shows various bias terms. Lastly, σ and tanh are the sigmoid activation function and hyperbolic tangent activation function, respectively. The process of malicious vehicular node detection using the LSTM-based approach in the proposed model is detailed in Algorithm 2. This series of operations enables the LSTM model to effectively process and analyze temporal data in vehicular networks, making it a robust solution for sequential data classification tasks.

Algorithm 2 shows that the input dataset $\mathcal{D}$ consists of features X_t and their corresponding labels Y_t, collected over T time steps. The dataset is then divided into training $(\mathcal{D}_{train})$ and testing $(\mathcal{D}_{test})$ subsets, with a predefined training percentage P_{train}. The LSTM network, initialized with parameters $\{W_f, W_i, W_C, W_o, b_f, b_i, b_C, b_o\}$, captures temporal dependencies by maintaining cell states C_t and hidden states h_t at each time step. The forget gate $(f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f))$ selectively retains or discards past information, while the input gate $(i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i))$ determines how much new information to incorporate.

Algorithm 2 LSTM-based malicious node detection in IoV network.

Input: $\mathcal{D} = \{(X_t, Y_t)\}_{t=1}^{T}$ (sequential dataset), T (total time steps), E (epochs), η (learning rate), P_{train} (training data percentage).
Output: Classification of vehicular nodes into legitimate or malicious.
Initialization:
Initialize LSTM parameters: $\{W_f, W_i, W_C, W_o, b_f, b_i, b_C, b_o\}$.
Set cell state $C_0 = 0$ and hidden state $h_0 = 0$.
Split dataset: $\mathcal{D}_{train} \leftarrow P_{train}\%$ of $\mathcal{D}$, $\mathcal{D}_{test} \leftarrow (1 - P_{train})\%$ of $\mathcal{D}$.
Step 1: Preprocessing
for each $(X_t, Y_t) \in \mathcal{D}$ **do**
  Extract features: end-to-end delay (D_{e2e}) and honesty (H).
  Normalize features: $\tilde{X}_t \leftarrow \frac{X_t - \mu}{\sigma}$, where μ and σ are the mean and standard deviation of X_t.
end for
Step 2: LSTM Training
for epoch $e \in \{1, \ldots, E\}$ **do**
  for each $(X_t, Y_t) \in \mathcal{D}_{train}$ **do**
    Compute forget gate: $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$.
    Compute input gate: $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$.
    Compute candidate cell state: $\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$.
    Update cell state: $C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$.
    Compute output gate: $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$.
    Compute hidden state: $h_t = o_t \odot \tanh(C_t)$.
  end for
  Update weights using backpropagation through time (BPTT) and gradient descent with learning rate η.
end for
Step 3: Prediction and Classification
for each $X_t \in \mathcal{D}_{test}$ **do**
  Compute sequence of hidden states $\{h_1, h_2, \ldots, h_T\}$ using the trained LSTM model.
  Calculate classification score: $\hat{Y}_t = \sigma(W_{out} \cdot h_t + b_{out})$.
  if $\hat{Y}_t > 0.5$ **then**
    Assign $Y_t = 1$ (Malicious).
    Remove node from network.
  else
    Assign $Y_t = 0$ (Legitimate).
  end if
end for
End of Algorithm

The candidate cell state ($\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$) represents potential updates, which are combined with the forget and input gates to update the cell state: $C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$. The output gate ($o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$) and the updated cell state determine the hidden state: $h_t = o_t \odot \tanh(C_t)$. The training process optimizes these parameters using backpropagation through time (BPTT) and gradient descent, with a learning rate η over multiple epochs E. After training, the model predicts the class of nodes in the test dataset $\mathcal{D}_{test}$. For each input sequence X_t, the trained LSTM computes hidden states $\{h_1, h_2, \ldots, h_T\}$, which are used to calculate a classification score $\hat{Y}_t = \sigma(W_{out} \cdot h_t + b_{out})$. A threshold of 0.5 is applied to $\hat{Y}_t$, assigning $Y_t = 1$ for malicious nodes and $Y_t = 0$ for legitimate ones.

Meanwhile, the NB technique provides a mechanism in which probabilistic classification is performed for all vehicular entities. The probabilistic classification does not require large computational overhead, which is suitable for resource-constrained IoV networks. NB computes the probability of a vehicular node N_i being malicious ($C = 1$) or legitimate ($C = 0$) given the feature set $X = \{x_1, x_2, \ldots, x_n\}$. The probability is calculated by (7).

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \tag{7}$$

By assuming conditional independence of the features, the posterior probability can be expressed, as shown in (8).

$$P(C|X) \propto P(C) \prod_{i=1}^{n} P(x_i|C) \tag{8}$$

Where $P(C)$ is the prior probability of a malicious or legitimate vehicular entity. $P(x_i|C)$ shows the likelihood of the feature x_i given the class C and $P(C|X)$ represents the posterior probability of the class C given the feature set X. In the classification process, NB calculates the probabilities for legitimate and malicious classes and assigns the node to the class with the highest posterior probability. The process of malicious vehicular node detection in our proposed model by utilizing the capabilities of NB is given in Algorithm 3.

Algorithm 3 Naive Bayes-based malicious node detection.

Input: $\mathcal{N} = \{N_1, N_2, \ldots, N_m\}$ (vehicular nodes), $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$ (feature set), $P(C)$ (prior probabilities), $\mathcal{C} = \{C_0, C_1\}$ (legitimate or malicious classes).
Output: $\mathcal{M} \subseteq \mathcal{N}$ (detected malicious nodes).
Step 1: Initialization
Extract feature set $\mathcal{X}$ from vehicular communication data for all nodes $\mathcal{N}$.
Define classes $\mathcal{C}$: C_0 for legitimate, C_1 for malicious nodes.
Assign prior probabilities $P(C_0)$ and $P(C_1)$ based on historical data.
Step 2: Likelihood Calculation
for each node $N_i \in \mathcal{N}$ **do**
  for each feature $x_j \in \mathcal{X}$ **do**
    Compute likelihoods for both classes:
    $P(x_j|C_k) = \frac{\text{Frequency of } x_j \text{ in class } C_k}{\text{Total instances in class } C_k}, \quad k \in \{0, 1\}$
  end for
end for
Step 3: Posterior Probability Calculation
for each node $N_i \in \mathcal{N}$ **do**
  Calculate posterior probabilities for both classes:
  $P(C_k|X_i) \propto P(C_k) \prod_{j=1}^{n} P(x_j|C_k), \quad k \in \{0, 1\}$
  Normalize:
  $P(C_k|X_i) = \frac{P(C_k) \prod_{j=1}^{n} P(x_j|C_k)}{\sum_{k \in \{0,1\}} P(C_k) \prod_{j=1}^{n} P(x_j|C_k)}$
end for
Step 4: Classification
for each node $N_i \in \mathcal{N}$ **do**
  if $P(C_1|X_i) > P(C_0|X_i)$ **then**
    Classify N_i as malicious (C_1).
  else
    Classify N_i as legitimate (C_0).
  end if
end for
Step 5: Malicious Node Removal
Identify malicious nodes: $\mathcal{M} = \{N_i \in \mathcal{N} : P(C_1|X_i) > P(C_0|X_i)\}$.
Remove malicious nodes $\mathcal{M}$ from the IoV network to enhance security.
End of Algorithm

In Algorithm 3, the process begins by defining the input, which includes the set of vehicular nodes $\mathcal{N} = \{N_1, N_2, \ldots, N_m\}$, a feature set $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$ extracted from communication data, prior probabilities $P(C_0)$ and $P(C_1)$ representing legitimate and malicious classes, respectively, and the class labels $\mathcal{C} = \{C_0, C_1\}$ (Chen et al. 2024). The algorithm aims to classify each vehicular node as either legitimate (C_0) or malicious (C_1) based on the posterior probabilities calculated using the NB method. It initializes by extracting relevant features from vehicular communication data, such as packet transmission behavior, end-to-end delay, or integrity indicators. Prior probabilities $P(C_0)$ and $P(C_1)$ are assigned based on historical or training data to reflect the likelihood of nodes being in each class.

For each node $N_i \in \mathcal{N}$, the algorithm calculates the likelihood $P(x_j|C_k)$ for each feature $x_j \in \mathcal{X}$ under each class $C_k \in \{C_0, C_1\}$. The likelihood is computed as the frequency of feature x_j in the class C_k, normalized by the total instances in that class. Using the calculated likelihoods, the posterior probabilities $P(C_0|X_i)$ and $P(C_1|X_i)$ for each node are determined as the product of the prior probability and the likelihoods of the features, given by $P(C_k|X_i) \propto P(C_k) \prod_{j=1}^{n} P(x_j|C_k)$. The posterior probabilities are normalized to ensure their values fall between 0 and 1. The node is classified as malicious if $P(C_1|X_i) > P(C_0|X_i)$; otherwise, it is classified as legitimate. The set of malicious nodes $\mathcal{M} = \{N_i \in \mathcal{N} : P(C_1|X_i) > P(C_0|X_i)\}$ is then removed from the vehicular network.

3.4 Blockchain-based federated learning mechanism for IoV networks

We propose a blockchain-based federated learning mechanism for malicious vehicular entities detection (Singh et al. 2024; Almazroi et al. 2024). We use the Keccak-256 hashing algorithm to enhance the integrity of data while simultaneously preserving the privacy of users. Furthermore, all the entities of the IoV network are registered with the blockchain network. Therefore, malicious entities can not change or modify data because unauthorized access can be easily detected in the blockchain network with the help of the Merkle tree structure. In our proposed model, the new vehicles can only join the network using the smart contract that is written and published on the Ethereum Remix.

This is the contract code attached to the Metamask wallet. When an unauthorized vehicle tries to access the data of the blockchain network, an alert message is sent to all registered cars via a web-based system (Meng et al. 2024). In this way, our proposed model ensures the reliability of data transferred between self-driving cars through the use of blockchain-based systems. After this, the data is validated by validator nodes with the help of the proof of authority (PoA) consensus algorithm. After that, each entity in the IoV network can verify that the transaction made is valid and authentic (Kaushik et al. 2024). Moreover, the security of the blockchain in our proposed model is made more secure by adding a unique code by adding nonce into it. This nonce is a very a unique part of PoA consensus techniques that adds another layer of security to the cryptographic hash generated for every transaction. In this way, our blockchain network not only secures the sensitive and important information of IoV network entities but also establishes a strong foundation of the highest trust and transparency in the interaction between the vehicular entities of the IoV network. The blockchain-based federated learning mechanism integrated with the blockchain network ensures the effective and reliable detection of malicious vehicular entities (Abou El Houda et al. 2024). LSTM uses sequential data such as end-to-end delay and time series data of vehicular networks and captures temporal dependencies in the communication among various network entities (Kamble et al. 2024).

While the NB performs probabilistic classification for all vehicular entities in the IoV network. The probabilistic classification does not require large computational overhead, which is suitable for resource-constrained IoV networks (Kaur et al. 2024). These trained models are deployed within a blockchain-secured IoV network, which ultimately enhances resilience and trustworthiness, empowering vehicles to securely share critical data for quick and intelligent decision-making for traffic management (Sun et al. 2024b).

The whole process of blockchain transaction validation with the federated learning mechanism for the IoV network is described in Algorithm 4. The algorithm outlines a blockchain-based federated learning mechanism for ensuring the secure classification of legitimate and malicious vehicular

Algorithm 4 Blockchain-based federated learning mechanism for IoV network.

Input: $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ (vehicular entities), $\mathcal{B} = \{b_1, b_2, \ldots, b_k\}$ (blocks in blockchain), $\mathcal{T} = \{t_1, t_2, \ldots, t_m\}$ (transactions), $\mathcal{SC}$ (smart contract), $\mathcal{H}$ (Keccak-256 hash function), $\mathcal{PIA}$ (Proof of Authority consensus algorithm), $\mathcal{F}$ (federated learning model).
Output: Validated and secured transactions in the blockchain network with malicious nodes detected.
Step 1: Registration and Initialization
for each $v_i \in \mathcal{V}$ **do**
  Register v_i in the blockchain network using $\mathcal{SC}$.
  Assign unique credentials to v_i for secure communication.
end for
Step 2: Data Collection and Hashing
for each $v_i \in \mathcal{V}$ **do**
  Collect communication data $\mathcal{X}_i$ (e.g., end-to-end delay, time series data).
  Compute hash $H_i = \mathcal{H}(\mathcal{X}_i)$ for integrity verification.
end for
Step 3: Validation of Transactions
for each $t_j \in \mathcal{T}$ **do**
  Broadcast t_j to validator nodes.
  Verify t_j using $\mathcal{H}$ and $\mathcal{PIA}$ consensus algorithm.
  if t_j is valid **then**
    Add t_j to a new block b_{k+1}.
    Append b_{k+1} to the blockchain $\mathcal{B}$.
  else
    Reject t_j and alert all registered $v_i \in \mathcal{V}$.
  end if
end for
Step 4: Federated Learning-Based Classification
for each $v_i \in \mathcal{V}$ **do**
  Train federated model $\mathcal{F}$ using $\mathcal{X}_i$.
  Apply LSTM for sequential data analysis.
  Update cell state C_t and hidden state h_t (Eq. 4–6).
  Perform NB-based probabilistic classification.
  Calculate $P(C|X) \propto P(C) \prod_{i=1}^{n} P(x_i|C)$.
  Classify v_i as legitimate or malicious.
end for
Step 5: Alert Mechanism for Malicious Entities
for each $v_i \in \mathcal{V}$ classified as malicious **do**
  Broadcast alert to all $v_j \in \mathcal{V}$ via web-based system.
  Prevent unauthorized access to $\mathcal{B}$ by v_i.
end for
Step 6: Security Enhancement with Nonce
for each $b_k \in \mathcal{B}$ **do**
  Add a unique nonce η_k to ensure secure hash computation.
  Recompute block hash $H_k = \mathcal{H}(b_k \parallel \eta_k)$.
end for
Step 7: Verification and Transparency
for each $b_k \in \mathcal{B}$ **do**
  Allow all $v_i \in \mathcal{V}$ to verify the integrity of b_k.
end for
End of Algorithm

entities in the Internet of Vehicles (IoV) network. The process begins with the registration and initialization of vehicular entities ($\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$) in the blockchain network using a smart contract ($\mathcal{SC}$). Each vehicular entity is assigned unique credentials for secure communication. The communication data ($\mathcal{X}_i$) generated by each vehicular entity is hashed using the Keccak-256 hashing function ($\mathcal{H}$), ensuring data integrity and preventing unauthorized modification. These transactions ($\mathcal{T} = \{t_1, t_2, \ldots, t_m\}$) are then broadcast to validator nodes for verification through the Proof of Authority (PoA) consensus algorithm ($\mathcal{PA}$). Transactions that pass validation are added to a new block (b_{k+1}) and appended to the blockchain ($\mathcal{B} = \{b_1, b_2, \ldots, b_k\}$), while invalid transactions are rejected, and alerts are sent to all registered entities.

The algorithm incorporates a federated learning model ($\mathcal{F}$) for efficient classification. Each vehicular entity trains the model locally using its collected data ($\mathcal{X}_i$). Sequential data such as end-to-end delay and time series information are analyzed using LSTM to capture temporal dependencies, updating the cell state (C_t) and hidden state (h_t) at each time step. Probabilistic classification is performed using the NB technique, calculating the posterior probability $P(C|X) \propto P(C) \prod_{i=1}^{n} P(x_i|C)$, where $P(C)$ is the prior probability of a class, $P(x_i|C)$ is the likelihood of feature x_i given the class C, and $P(C|X)$ is the posterior probability of the class C given the feature set X.

The classification process identifies legitimate and malicious vehicular entities, and an alert mechanism broadcasts warnings about malicious nodes to all other vehicles in the network. To enhance blockchain security, each block (b_k) is augmented with a unique nonce (η_k), ensuring secure cryptographic hash generation ($H_k = \mathcal{H}(b_k||\eta_k)$) for every transaction. This method promotes transparency by allowing all vehicular entities to verify the integrity of blockchain transactions, establishing trust and resilience in the IoV network.

We face several challenges while developing our proposed model. First, we need to carefully coordinate model updates across different fog nodes that have varying computing power and network conditions to make sure the model converges. Second, LSTM is efficient in the classification of malicious and legitimate vehicles due to learning temporal dependencies. However, it has a large computational overhead, which makes real-time use of edge devices difficult and is not suitable for resource-constrained IoV networks. To solve this issue, our model distributes LSTM training to fog nodes with sufficient computational capacity rather than directly on in-vehicle devices. Moreover, we use truncated backpropagation through time, reduced sequence lengths, gradient compression, and adaptive learning rates to optimize training time and memory usage. Third, we need to carefully combine the results from the LSTM and Naive Bayes models to make sure the predictions stay accurate and use the strengths of both models. Lastly, we must protect the model updates from harmful attacks and make sure the model works well even in tough or tricky situations.

4 Experimental results

5 Dataset description

We use the Vehicular Reference Misbehavior (VeReMi) dataset (VeReMi Project Contributors 2024) for testing and training of our proposed model. The VeReMi dataset provides several unique features and carefully structured instances for effective and reliable testing. The dataset consists of various instances that simulate vehicular communication under both legitimate and malicious scenarios. It features a wide variety of malicious behaviors such as false position attacks, where nodes broadcast their incorrect credentials and message fabrication, which ultimately reflects diverse real-world security threats. This dataset provides a a wide range of scenarios such as diverse attack patterns and node densities, which makes it ideal for evaluating the reliability and effectiveness of malicious node detection models. The results from the VeReMi dataset validate that our LSTM and NB-enabled malicious node detection model is not only able to enhance the privacy of network vehicular entities but also the accuracy of the detection mechanism. The simulation parameters of our model are given in Table 2.

Table 2 Simulation parameters

Parameter Name	Values
Sensing Area	1000 X 1000 m^2
Total Ordinary Sensing Vehicles	200
Number of Virtual Machines	8
Number of RSU	8
Number of SNs	3
Communication Range	300 meters
Speed of SVs	30-100 km/h
Simulation Duration	600 seconds
RSU Communication Delay	3 ms
Blockchain Consensus Mechanism	PoA
Federated Learning Model	LSTM and NB
Probabilistic Classifier	NB
Malicious Node Injection Rate	5% of total nodes
Data Transmission Frequency	2 packets/s
Packet Loss Probability	1%
Encryption Protocol	AES-256
Power Consumption (Avg)	3W per node

5.1 Evaluation metrics

To ensure a comprehensive performance assessment, all models are evaluated using standard classification metrics: accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the model's predictions as represented in (9).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

Besides this, precision and recall provide insights into the model's ability to correctly identify malicious nodes without generating false alarms. Both precision and recall are calculated according to (10) and (11), respectively.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{11}$$

The F1-score harmonizes precision and recall into a single metric to balance detection performance, represented in (12).

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

It is important to emphasize that centralized models, despite being effective in classification, inherently suffer from privacy risks since raw vehicular data must be shared with a central entity, potentially exposing sensitive information. Moreover, centralized training struggles with scalability and efficiency when processing large-scale, heterogeneous, and dynamic IoV data, often leading to performance bottlenecks and increased latency. These models are also vulnerable to single points of failure, where any disruption at the central server compromises the entire detection system. Our proposed model efficiently addresses these critical limitations by enabling distributed training across fog nodes, preserving privacy by sharing only model updates, and improving robustness and scalability in IoV networks.

5.2 Centralized benchmark models

Various well-established centralized machine learning and deep learning models are implemented and evaluated on the same preprocessed IoV dataset. The deep learning model consists of two convolutional layers followed by max-pooling layers and a fully connected layer activated by ReLU functions, designed to capture spatial features effectively. The machine learning classifiers include SVM configured with a radial basis function (RBF) kernel to handle non-linear data separability, and Decision Tree classifiers using the Gini impurity criterion for feature splitting and node purity. These models are trained and tested in a centralized manner, meaning all raw vehicular data is aggregated at a central server for model development. The parameters for each model, such as the number of neurons in fully connected layers, kernel parameters for SVM, and maximum tree depth for Decision Trees, are tuned based on validation sets to optimize performance. This benchmark setup serves as a baseline for evaluating the improvements introduced by our proposed model.

In addition to the above models, we consider the Long Short-Term Memory (LSTM) network, Naive Bayes (NB), Decision Tree (DT), and Support Vector Machine (SVM) to establish robust baselines. The LSTM model is employed due to its capability to capture temporal dependencies within sequential vehicular data. It is configured with an input layer matching the feature dimension, followed by one or more LSTM layers with 64 hidden units, and a dropout layer to prevent overfitting. The model is trained using the Adam optimizer with a learning rate of 0.001 and a binary cross-entropy loss for classification.

The Naive Bayes (NB) classifier assumes conditional independence among features and is particularly lightweight, making it a fast and interpretable baseline. We use the Gaussian variant of NB, which models the likelihood of continuous features using Gaussian distributions. The model updates are based on closed-form probability estimates derived from the training data without iterative backpropagation.

The Decision Tree (DT) classifier is parameterized with a maximum depth of 10, a minimum sample split of 2, and uses the Gini impurity as the splitting criterion. These parameters are optimized using grid search on validation sets to prevent overfitting and ensure generalization. The tree structure is recursively built by selecting features that yield the highest information gain, and pruning strategies are optionally applied to reduce complexity.

The Support Vector Machine (SVM) is implemented with an RBF kernel to map input features into higher-dimensional space for better classification of non-linear patterns. The hyperparameters include the regularization parameter $C = 1.0$ and kernel coefficient $\gamma = 0.01$, both tuned via cross-validation. The SVM model is trained using the Sequential Minimal Optimization (SMO) algorithm, and support vectors are identified based on the margins between classes.

These centralized models are updated based on traditional training regimes involving gradient descent (in deep learning models) or rule-based statistical updates (in NB and DT) and serve as a strong baseline to contrast the distributed learning advantages of our proposed model framework.

5.3 Hybrid classification strategy with LSTM and Naive Bayes in federated IoV networks

In our proposed model, LSTM and Naive Bayes are trained separately on the same dataset distributed across fog nodes.

LSTM captures temporal patterns, while Naive Bayes provides fast probabilistic classification based on feature distributions. After this, the final prediction is obtained by combining outputs through a weighted voting mechanism, where higher confidence scores are given precedence. This hybrid integration enhances model robustness, which ultimately allows our model to benefit from both deep temporal learning and lightweight statistical inference. In this way, our model ensures independent model maintenance and ensures effective and reliable updates, and modular deployment in diverse vehicular environments. In our proposed model for malicious node detection, the data of each sub-network is trained on a local virtual machine. We use the PoA consensus algorithm for the validation of transactions among legitimate vehicular entities in contrast to the benchmark scheme that uses the proof of work (PoW) consensus algorithm. Figure 2 shows the consumption of gas for various numbers of transactions for both PoA and PoW consensus algorithms. The figure shows that PoW consistently consumes around 2,300,000 Gwei gas units for transactions from 100 to 300. While PoA maintains a lower and consistent gas consumption of 2,200,000 Gwei for transactions 100 to 400, also converges to 2,100,000 Gwei at 500 transactions. It is clear that the consumption of gas for validation of transactions for PoW is comparatively higher than the PoA consensus algorithm, as shown in Table 3. The reason is that the PoW consensus algorithms need to find a nonce such that the resulting hash is below a specified target value. First of all, the performed transaction is broadcasted in the network then all the transactions are collected by the miner nodes in the block. Then miner nodes solve a puzzle to find a nonce, it is a special magical number that is added to the hash to calculate a particular type of hash. The nonce is a random number,

Table 3 Gas consumption (Gwei) for PoW and PoA across various transactions

No. of Transactions	PoW (Gwei)	PoA (Gwei)
100	2,310,000	2,202,000
200	2,302,000	2,220,000
300	2,302,000	2,201,000
400	2,211,000	2,208,000
500	2,114,000	2,112,000

which adds a dash of unpredictability and makes it tough to change the hash. Nonce varies significantly in the PoW consensus mechanism, which changes the difficulty and rewards to find that perfect hash meeting a special target value of the hash. The whole process of finding the magical nonce is called mining and it requires a large amount of computational resources because the miner nodes have to perform different hash operations for the calculation of the correct nonce. Moreover, some validator nodes are selected in PoA based on their stakes and reputations. As PoA consensus algorithm relies on the capabilities of pre-approved validators other than the mining process. Therefore, the overall computational overhead is very small in PoA as compared to the PoW consensus algorithm.

Furthermore, we use the Keccak-256 hashing algorithm in our proposed model to ensure the integrity of data in contrast to the benchmark scheme that uses SHA-256. Figure 3 shows the comparison of transaction latencies of SHA-256 and Keccack-256 hashing algorithms. It can be depicted that the transaction latency for the Keccak-256 algorithm is lower for all the transactions as compared to the SHA-256 hashing technique. For 1,000 transactions, Keccak-256 takes 0.8 seconds, whereas SHA-256 requires 1.2 seconds, shows that

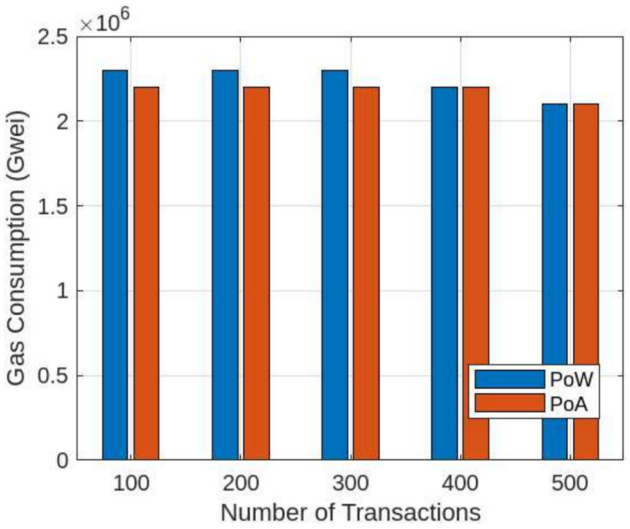

Fig. 2 The gas consumption of PoW and PoA for different transactions

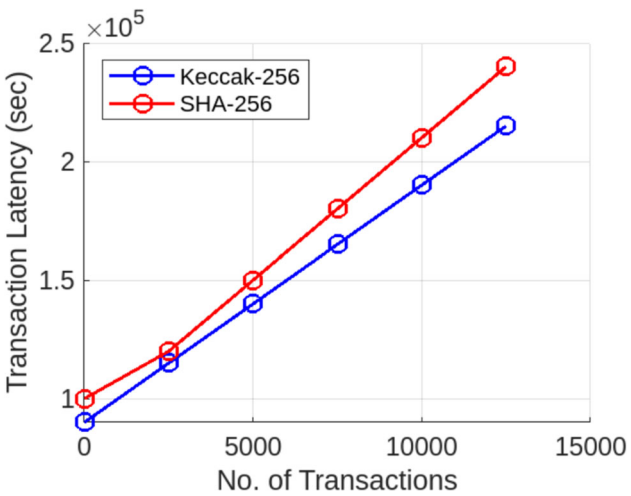

Fig. 3 Transaction latency for different blockchain transactions

Table 4 Transaction latency comparison of Keccak-256 and SHA-256

No. of Transactions	Keccak-256 (sec)	SHA-256 (sec)
1000	0.8	1.2
2000	1.4	1.8
3000	2.0	2.5
4000	2.6	3.2
5000	3.0	3.8

SHA-256 is comparatively slower. This performance difference is illustrated in Table 4. The reason for the lower transaction latency of the Keccack-256 hashing algorithm is that it uses sponge construction, which takes a very small amount of time to hash as compared to the SHA-256 Markel tree technique.

5.4 Performance comparison with centralized models

Figure 4 shows a comparative evaluation of the proposed blockchain-based federated learning model against several benchmark distributed classification methods, including Naive Bayes (NB), Long Short-Term Memory (LSTM), Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Decision Tree (DT). The models are assessed using four key performance metrics: accuracy, precision, recall, and F1-score. The results show that SVM achieves an accuracy of 86%, which is comparatively low due to its reliance on static features, limiting its effectiveness in the dynamic IoV environments. Similarly, CNN achieves a higher accuracy of 91%. However, it lacks the capacity

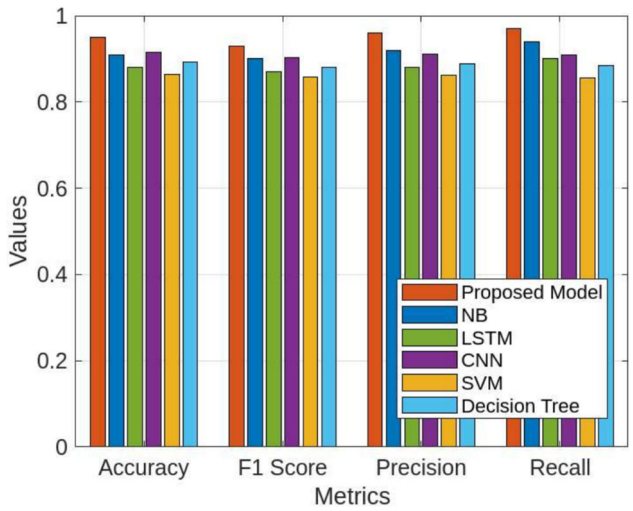

Fig. 4 Performance comparison of the proposed model with centralized models

Table 5 Comparison between the proposed model and centralized benchmark models

Metric	Proposed Model	NB	LSTM	SVM	CNN	Decision Tree
Accuracy	0.95	0.91	0.88	0.86	0.91	0.89
F1 Score	0.93	0.90	0.87	0.85	0.90	0.88
Precision	0.96	0.92	0.88	0.86	0.91	0.88
Recall	0.97	0.94	0.90	0.85	0.90	0.88

to model temporal dependencies, which are essential for accurate malicious node detection in vehicular networks. The decision tree classifier attains an accuracy of 89% but exhibits vulnerability to overfitting and demonstrates low generalization, particularly when handling imbalanced and sequential data. Overall, the proposed model consistently outperforms all benchmark centralized methods across all evaluation metrics. This improvement is attributed to the fact that our model trains a unified model using dispersed features from multiple sub-networks. Therefore, the model is enriched with meaningful features and subsequently shared with each sub-network for classification, which ultimately enhances the accuracy of our proposed model for malicious node detection. Notably, the proposed model achieves the highest accuracy of 0.95%, as shown in Table 5. This performance gain is due to its effective learning of temporal patterns and dependencies in sequential data. The temporal correlations among vehicular entities are captured efficiently by the model's LSTM component, which contributes to robust classification performance. In terms of F1-score, the proposed model achieved a value of 0.93, outperforming both Naive Bayes and other centralized benchmark schemes. This improvement indicates a stronger ability to accurately distinguish between malicious and legitimate vehicles by leveraging temporal correlations. Furthermore, the model achieves a precision of 0.96 and a recall of 0.97, which are the highest values among all compared methods. These results show that our proposed model can effectively capture temporal dependencies and complex patterns in sequential data of network vehicular entities, which ultimately helps in the accurate identification of malicious nodes with minimal false positives and false negatives.

5.5 Comparison with federated learning models

Figure 5 presents a comparative analysis of the proposed blockchain-based federated learning model against other federated learning models, including FedAvg-LSTM (Li et al. 2024), FedProx (Yu et al. 2024), and FedAvg-CNN (Si et al. 2024), evaluated across four key performance metrics: accuracy, precision, recall, and F1-score. The proposed model

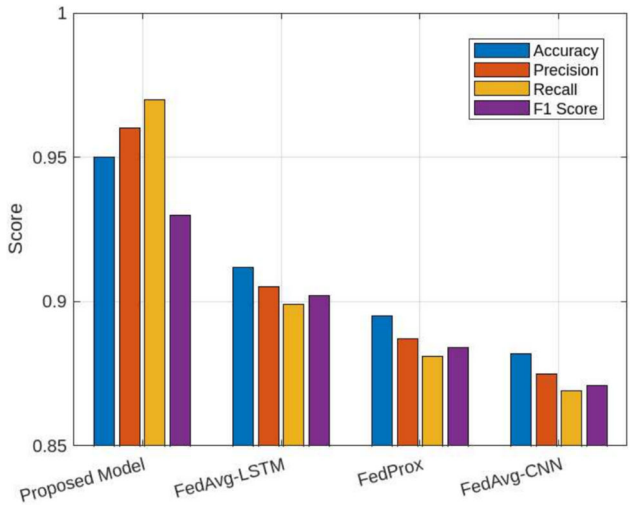

Fig. 5 Performance comparison of the proposed model with other federated learning models

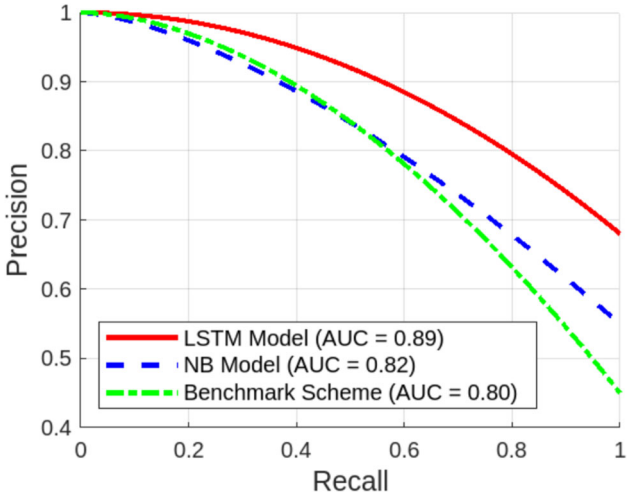

Fig. 6 Performance comparison based on the PR curve for LSTM, Naive Bayes (NB), and the benchmark scheme

achieves the highest classification accuracy of 95.0%, outperforming FedAvg-LSTM (91.2%), FedProx (89.5%), and FedAvg-CNN (88.2%). In addition, the proposed model demonstrates superior performance in precision (96.0%), recall (97.0%), and F1-score (93.0%), indicate its robustness and reliability to detect malicious nodes. The enhanced performance is attributed to the hybrid architecture that integrates LSTM for modeling temporal dependencies in sequential vehicular data with Naive Bayes for computationally efficient probabilistic inference. However, FedAvg-LSTM and FedProx exhibit suboptimal performance due to the non-IID data distribution in IoV environments, while FedAvg-CNN lacks temporal modeling capabilities. Furthermore, the proposed model ensures data privacy and enables scalable deployment across distributed fog nodes, making it well-suited for real-time and privacy-sensitive vehicular networks.

5.6 Performance evaluation using PR curve and end-to-end delay

A Precision-Recall (PR) curve is presented in Fig. 6 that compares the performance of LSTM and NB and a benchmark centralized classification technique for malicious node detection in our deployed IoV network. LSTM model outperforms NB and benchmark models and achieves the highest Area Under the Curve (AUC) of 0.89. It indicates that LSTM is superior in balancing precision and recall as the LSTM model effectively identifies malicious nodes while simultaneously minimizing false positives. However, the NB model has an AUC of 0.82 and performs effectively. However, it struggles as it overlooks complex relationships in the IoV network. Lastly, the benchmark scheme has the lowest AUC of 0.80% among all, which shows that the benchmark scheme is not reliable and effective in the classification of legitimate and malicious vehicular entities.

We consider various features such as sequential data, end-to-end delay, and honesty for evaluating the behavior of various vehicular entities. Figure 7 shows the end-to-end delay of various randomly selected IoV vehicular entities. Its shows that the end-to-end delay value of various vehicles is different for different vehicular entities. The reason is that the response time of different vehicles is different due to the types and sizes of shared data. Furthermore, the vehicles having higher end-to-end delay have maximum chances to be identified as malicious nodes by the classifiers because the abnormal value of any parameters is easily detected by our proposed blockchain-based federated learning model.

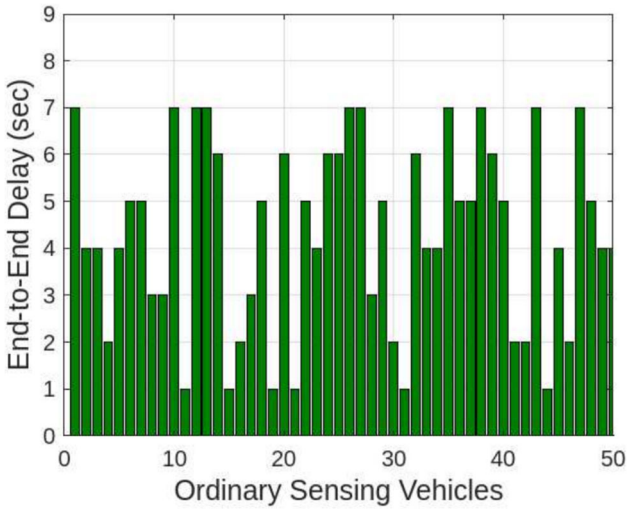

Fig. 7 End-to-end delay of ordinary sensing vehicles

6 Conclusion

In this paper, a blockchain-based federated learning framework for malicious node detection and distributed classification of legitimate and malicious vehicles is proposed. The capabilities of LSTM and NB are utilized in the blockchain-based federated learning model for efficient and reliable malicious node detection. The distributed models are trained on each locally installed virtual machine using a federated learning mechanism. After this, a unified model is generated at the centralized cloud server. We use VeReMi to evaluate the performance of our proposed model. The simulation results indicate that the LSTM and NB techniques outperforms centralized benchmark classification methods in malicious node detection. An accuracy of 95% is achieved by the LSTM model, which shows the efficiency of proposed model in identification of both malicious and legitimate vehicles. In the future, we will further use a novel deep learning algorithm to optimize the federated learning model for real-time malicious node detection in dynamic IoV networks. Furthermore, we will use InterPlanetary File System (IPFS) to address the issue of costly storage in blockchain-enabled IoV network systems.

Author Contributions Srinivas Reddy Bandarapu: Writing – original draft, Visualization, Methodology, Conceptualization. Muhammad Bilal: Resources, Project administration, Writing – review & editing. Pushpalika Chatterjee: Writing – review & editing, Resources, Validation. Adnan Mustafa Cheema: Conceptualization, Writing – review & editing, Validation. Junaid Rashid: Conceptualization, Visualization, Formal analysis, Writing – review & editing. Jungeun Kim: Formal analysis, Validation, Writing – review & editing.

Funding This work was partly supported by the Institute of Information & Communications Technology Planning & Evaluation(IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government(MSIT)(IITP-2025-RS-2023-00259678) and by the IITP(Institute of Information & Communications Technology Planning & Evaluation)-ITRC(Information Technology Research Center) grant funded by the Korea government(Ministry of Science and ICT)(IITP-2025-RS-2024-00438335).

Availability of data and material The datasets are publicly available.

Declarations

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abou El Houda Z, Moudoud H, Brik B, Khoukhi L (2024) Blockchain-enabled federated learning for enhanced collaborative intrusion detection in vehicular edge computing. IEEE Trans Intell Transport Syst 25:7661–7672

Ahmad F, Kurugollu F, Kerrache CA, Sezer S, Liu L (2021) Notrino: A novel hybrid trust management scheme for internet-of-vehicles. IEEE Trans Vehic Technol 70:9244–9257

Ahmed I, Jeon G, Ahmad A (2021) Deep learning-based intrusion detection system for internet of vehicles. IEEE Consumer Electron Mag 12:117–123

Alalwany E, Mahgoub I (2024) Security and trust management in the internet of vehicles (iov): Challenges and machine learning solutions. Sensors 24:368

Almazroi AA, Alqarni MA, Al-Shareeda MA, Alkinani MH, Almazroey AA, Gaber T (2024) Fca-vbn: Fog computing-based authentication scheme for 5g-assisted vehicular blockchain network. Int Things 25:101096

Bhattacharya S, Victor N, Chengoden R, Ramalingam M, Selvi GC, Maddikunta PKR, Donta PK, Dustdar S, Jhaveri RH, Gadekallu TR (2022) Blockchain for internet of underwater things: State-of-the-art, applications, challenges, and future directions. Sustainability 14:15659

Chen CM, Miao Q, Kumari S, Khan MK, Rodrigues JJ (2024) A privacy-preserving authentication protocol for electric vehicle battery swapping based on intelligent blockchain. IEEE Int Things J 11:17538–17551

Gautam D, Thakur G, Kumar P, Das AK, Park Y (2024) Blockchain assisted intra-twin and inter-twin authentication scheme for vehicular digital twin system. IEEE Trans Intellt Transport Syst 25(10):15002–15015

Gu K, Ouyang X, Wang Y (2024) Malicious vehicle detection scheme based on spatio-temporal features of traffic flow under cloud-fog computing-based iovs. IEEE Trans Intell Transport Syst 25(9):11534–11551

Guo J, Bilal M, Qiu Y, Qian C, Xu X, Choo KKR (2024) Survey on digital twins for internet of vehicles: Fundamentals, challenges, and opportunities. Digital Commun Netw 10:237–247

Hemmati A, Zarei M, Rahmani AM (2024) Big data challenges and opportunities in internet of vehicles: a systematic review. Int J Pervasive Comput Commun 20:308–342

Hîrţan LA, Dobre C, González-Vélez H (2020) Blockchain-based reputation for intelligent transportation systems. Sensors 20:791

Jabbar R, Dhib E, Said AB, Krichen M, Fetais N, Zaidan E, Barkaoui K (2022) Blockchain technology for intelligent transportation systems: A systematic literature review. IEEE Access 10:20995–21031

Jain A, Khare V (2024) Blockchain-based financial and economic analysis of green vehicles: Path towards intelligent transportation. Towards a Green Economy Sustain Development, Interconnected Modern Multi-Energy Netw Intell Transport Syst, pp 344–378

Kamble S, Kounte M et al (2024) Blockchain technology: A study applied on intelligent transport system. In: AI and blockchain applications in industrial robotics. IGI Global Scientific Publishing, pp 253–275

Kaur G, Shobana M, Kavin F, Sellakumar S, Meenakshi D, Bharathiraja N (2024) Secured intelligent transportation with privacy retention through blockchain framework. J Intell Fuzzy Syst 46:10507–10521

Kaushik P, Rathore SPS, Sachdeva L, Poonia M, Singh D, Bir L (2024) Intelligent transportation systems trusted user's security and privacy. In: 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), IEEE, pp 1–6

Kebande VR, Awaysheh FM, Ikuesan RA, Alawadi SA, Alshehri MD (2021) A blockchain-based multi-factor authentication model for a cloud-enabled internet of vehicles. Sensors 21:6018

Kharche A, Badholia S, Upadhyay RK (2024) Implementation of blockchain technology in integrated iot networks for constructing scalable its systems in india. Blockchain: Res Appl 5:100188

Khezri E, Hassanzadeh H, Yahya RO, Mir M (2025) Security challenges in internet of vehicles (iov) for its: A survey. Tsinghua Sci Technol 30:1700–1723

Korba AA, Boualouache A, Ghamri-Doudane Y (2024) Zero-x: A blockchain-enabled open-set federated learning framework for zero-day attack detection in iov. IEEE Trans Vehicular Technol 73:12399–12414

Li Y, Zhang R, Zhao P, Wei Y (2024) Feature-attended federated lstm for anomaly detection in the financial internet of things. Appl Sci 14:5555

Liang W, Liu Y, Yang C, Xie S, Li K, Susilo W (2024) On identity, transaction, and smart contract privacy on permissioned and permissionless blockchain: A comprehensive survey. ACM Comput Surv 56:1–35

Lingamallu RK, Balasubramani P, Arvind S, Rao PS, Ammisetty V, Gupta KG, Sharath M, Kumar YN, Mittal V (2024) Securing iot networks: A fog-based framework for malicious device detection. In: MATEC Web of Conferences, EDP Sciences, pp 01103

Liu X, Tan Z, Liang L, Li G (2024) A multidimensional trust evaluation mechanism for improving network security in fog computing. IEEE Trans Indust Inf

Matei A, Cocoşommaatu M (2024) Artificial internet of things, sensor-based digital twin urban computing vision algorithms, and blockchain cloud networks in sustainable smart city administration. Sustainability 16:6749

Meng X, Liu B, Meng X, Liang Y, Deng H (2024) A lightweight group authentication protocol for blockchain-based vehicular edge computing networks. IEEE Trans Intell Transport Syst

Miao J, Wang Z, Ning X, Shankar A, Maple C, Rodrigues JJ (2024) A uav-assisted authentication protocol for internet of vehicles. IEEE Trans Intell Transport Syst 25(8):10286–10297

Mulligan C, Morsfield S, Cheikosman E (2024) Blockchain for sustainability: A systematic literature review for policy impact. Telecommun Policy 48:102676

Nakamoto S (2008) Bitcoin: A peer-to-peer electronic cash system. Accessed November 15, 2024

Nguyen TT, Nguyen HH, Sartipi M, Fisichella M (2024) Lammon: language model combined graph neural network for multi-target multi-camera tracking in online scenarios. Mach Learn 113:6811–6837

Ning Z, Sun S, Wang X, Guo L, Guo S, Hu X, Hu B, Kwok RY (2021) Blockchain-enabled intelligent transportation systems: A distributed crowdsensing framework. IEEE Trans Mobile Comput 21:4201–4217

Ozaki H, Tran DT, Lee JH (2024) Effective human-object interaction recognition for edge devices in intelligent space. SICE J Control, Measurement, Syst Integration 17:1–9

Popoola O, Rodrigues M, Marchang J, Shenfield A, Ikpehai A, Popoola J (2024) A critical literature review of security and privacy in smart home healthcare schemes adopting iot & blockchain: problems, challenges and solutions. Blockchain: Res Appl 5:100178

Rehman A, Hassan MF, Hooi YK, Qureshi MA, Shukla S, Susanto E, Rubab S, Abdel-Aty AH (2022) Ctmf: Context-aware trust management framework for internet of vehicles. IEEE Access 10:73685–73701

Shinde V, Dhanawat V, Almogren A, Biswas A, Bilal M, Naqvi RA, Rehman AU (2024) Copy-move forgery detection technique using graph convolutional networks feature extraction. IEEE Access 12:121675–121687

Si C, Wang H, Chen L, Zhao J, Min Y, Xu F (2024) Robust co-modeling for privacy-preserving short-term load forecasting with incongruent load data distributions. IEEE Trans Smart Grid 15(3):2985–2999

Singh A, Rani P, Ramesh JVN, Athawale SV, Alkhayyat AH, Aledaily AN, Prola TA, Sharma R (2024) Blockchain-based lightweight authentication protocol for next-generation trustworthy internet of vehicles communication. IEEE Trans Consumer Electron 70:4898–4907

Sun G, Wang Z, Su H, Yu H, Lei B, Guizani M (2024) Profit maximization of independent task offloading in mec-enabled 5g internet of vehicles. IEEE Trans Intell Transport Syst 25(11):16449–16461

Sun X, Dou H, Chen S, Zhao H (2024) A novel block-chain based secure cross-domain interaction approach for intelligent transportation systems. Phys Commun 63:102223

Ullah I, Deng X, Pei X, Mushtaq H, Khan Z (2025) Securing internet of vehicles: a blockchain-based federated learning approach for enhanced intrusion detection. Cluster Comput 28:256

VeReMi Project Contributors (2024) Vehicular reference misbehavior dataset (veremi). https://veremi-dataset.github.io/. Accessed November 17, 2024

Wang S, Hu Y, Qi G (2022) Blockchain and deep learning based trust management for internet of vehicles. Simulation Modell Pract Theory 120:102627

Wu C, Fan H, Wang K, Zhang P (2024) Enhancing federated learning in heterogeneous internet of vehicles: A collaborative training approach. Electronics 13:3999

Xie N, Zhang C, Yuan Q, Kong J, Di X (2024) Iov-bcfl: An intrusion detection method for iov based on blockchain and federated learning. Ad Hoc Networks 163:103590

Xu S, Wang T, Sun A, Tong Y, Ren Z, Zhu R, Song HH (2024) Post-quantum anonymous, traceable and linkable authentication scheme based on blockchain for intelligent vehicular transportation systems. IEEE Trans Intell Transport Syst 25(9):12108–12119

Yan G, Liu K, Liu C, Zhang J (2024) Edge intelligence for internet of vehicles: A survey. IEEE Trans Consumer Electron 70(2):4858–4877

Yang M, Zhang B, Wang T, Cai J, Weng X, Feng H, Fang K (2024) Vehicle interactive dynamic graph neural network-based trajectory prediction for internet of vehicles. IEEE Int Things J 11:35777–35790

Yu T, Wang X, Hu J, Yang J (2024) Multi-agent proximal policy optimization-based dynamic client selection for federated ai in 6g-oriented internet of vehicles. IEEE Trans Vehic Technol 73(9):13611–13624

Zhu J, Feng T, Lu Y, Jiang W (2024) Using blockchain or not? a focal firm's blockchain strategy in the context of carbon emission reduction technology innovation. Business Strategy Environ 33:3505–3531

Zou Y, Zhang Z, Zhang C, Zheng Y, Yu D, Yu J (2024) A distributed abstract mac layer for cooperative learning on internet of vehicles. IEEE Trans Intell Transport Syst 25(8):8972–8983