International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

PRIVACY-PRESERVING FEDERATED LEARNING WITH LOCAL DIFFERENTIAL PRIVACY FOR TRAFFIC PREDICTION IN IOV SYSTEMS

Muthana J. Khudair [1], Dr. Foad Salem Mubarek[2], Dr. Salah A. Aliesawi[3]

[1]Assist. Lect.

[2,3]Associate Professor Prof.

1,2,3. University of Anbar, College of Computer Science and Information Technology., mut21c1016@uoanbar.edu.iq

Abstract

The Internet of Vehicles (IoV) poses severe challenges for distributed collaborative traffic prediction due to privacy concerns and scalability in the distributed learning scenario. Traditional federated learning mechanisms in vehicle networks fail to ensure data privacy while ensuring efficient consensus schemes for model aggregation. This paper proposes a novel privacy-preserving federated learning framework that integrates an enhanced Practical Byzantine Fault Tolerance (PBFT) consensus protocol with Local Differential Privacy (LDP) for trustworthy traffic prediction in IoV systems. Our new PBFT algorithm incorporates logarithmic scaling in which the consensus rate decreases as the network increases, significantly improving scalability for large vehicular networks. The design features a CNN-GRU neural network model that is trained on real PeMS traffic data, enabling accurate prediction of traffic flow, occupancy, and speed patterns. Local Differential Privacy with calibrated noise injection provides a robust privacy guarantee at 50% privacy leakage loss with an accuracy of 70-82% compared to 75-87% accuracy in non-protected models. Experimental results provide improved performance with 204ms average consensus time per hop across 200 vehicles, outperforming traditional PBFT (365ms) and alternative baselines. A multi-component incentive framework controlling accuracy, privacy, and contributions ensures equitable participation and ongoing cooperation. Our approach effectively authenticates on real California highway data with 92-95% of the centralised system's accuracy at preserving vehicle privacy, which is appropriate for realistic IoV deployments that require both efficiency and privacy preservation.

Keywords: Federated Learning, Local Differential Privacy, Internet of Vehicles, Enhanced PBFT, Traffic Prediction

1. Introduction

The rapid evolution of the Internet of Vehicles (IoV) transformed modern-day transportation systems to enable real-time vehicle, infrastructure, and cyber services communication to enhance road safety, traffic flow, and driving experience [1]. With the evolution of vehicular networks, there exists a heavy demand for smart traffic forecasting systems capable of leveraging distributed data from different vehicles while maintaining operational efficiency [2].

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Federated learning has also emerged as a promising paradigm for IoV applications, where vehicles have been able to train models collaboratively without sharing raw sensor data, hence avoiding bandwidth limitation and distributed computational overhead [3].

However, implementing federated learning in vehicular environments comes with unique challenges, primarily around privacy protection and scalable consensus algorithms. Federated learning solutions in IoV systems commonly suffer significantly from privacy leakage since model parameters can reveal vehicle trace [4], driving behaviour, and location-based data. In addition, the current consensus mechanisms like the mainstream Practical Byzantine Fault Tolerance (PBFT) suffer in terms of scalability when implemented in large vehicles networks with quadratic communication complexity that cannot be maintained as the number of participating vehicles grow [5]. The high mobility, intermittently available connectivity, and variable availability of participants are also dynamic characteristics of the vehicular network that make it more complicated to have efficient and secure federated learning systems [6].

Recent research on privacy preserving vehicular learning has been inclined to focus on cryptographic protection solutions, or simple anonymisation techniques, which are usually accompanied by enormous computation cost, or ineffective privacy assurances [7]. Besides, the former federated learning solutions to IoV largely lack holistic incentive models to enable consistent engagement with both performance and privacy requirements balanced [8]. The combination of powerful privacy-sensitive strategies and scalable consensus mechanisms with strong submersiveness has a long way to be explored, leaving a vast void in the practical implementation of federated learning technology in the context of real-life IoV solutions [9].

This work rectifies these limitations by proposing a novel privacy-preserving federated learning system integrating a customised PBFT consensus algorithm with Local Differential Privacy for IoV system traffic prediction. Our key contributions are: (1) a log-scale PBFT algorithm with significantly improved consensus efficiency for extremely large vehicular networks, (2) an end-to-end Local Differential Privacy mechanism providing a provable privacy guarantee with model utility guaranteed, (3) a CNN-GRU neural network architecture optimized for actual-world PeMS-based distributed traffic prediction, (4) a multi-dimension incentive framework balancing accuracy, privacy, and participation to enable sustainable cooperation, and (5) comprehensive experimental evaluations demonstrating better performance compared to existing methods with vehicle privacy guaranteed in realistic IoV scenarios.

2. Related Work

Recent advancements in federated learning for IoV systems have been highly promising for collaborative learning with privacy preservation. Alwash et al. [9] introduced a fast federated learning approach for secure intrusion detection in IoV with 99.64% and 99.99% accuracy, coupled with differential privacy mechanisms. Al-Huthaifi et al. [10] introduced FedGODE, which uses graph ordinary differential equations combined with privacy protection mechanisms like local differential privacy and homomorphic encryption to forecast traffic flow.

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Ullah et al. [11] introduced IoV-SFL, a blockchain-backed federated learning framework that integrates federated learning and blockchain technology to provide secure data sharing in IoV systems.

Byzantine Fault Tolerance consensus algorithms tailored to IoV environments gained momentum with enhanced PBFT variants such as SG-PBFT [12] and R-PBFT [13], which enhance consensus latency by 50% and provide guarantees of security. Zhang et al. [14] introduced AE-PBFT as an adaptive consensus algorithm incorporating trust management models to further boost node screening effectiveness. The blend of blockchain-aided federated learning and PBFT consensus [15] offers the high efficiency of non-conventional public blockchains with differential privacy assurances.

Privacy preservation through differential privacy has become crucial in vehicular federated learning systems. Newer algorithms of federated learning-based intrusion detection systems [16] employ client-level differential privacy mechanisms for secure threat identification. Image-based conversion of vehicular communication traffic data with pruned neural networks [17] achieves high detection accuracy at low computational overheads on resource-constrained devices.

Traffic prediction using CNN-GRU hybrid models has enhanced significantly through federated learning platforms. Hybrid CNN-GRU-LSTM models [18] apply CNN to capture spatial dependency and use GRU layers to identify temporal patterns and outperform current solutions by a large margin. FedGRU solutions [19] share the same prediction accuracy as privacy-protected centralized methods. Dai and Tang [20] proposed PLFL for traffic flow prediction with spatiotemporal fusion graph convolutional networks and dynamic model pruning.

Weighted aggregation methods have been established to address the problems of heterogeneous data distributions in vehicular networks. Zhang et al. [21] developed sophisticated federated learning methods with weighted model aggregation for spatial-temporal modeling and performed better with heterogeneous wireless traffic data. Li et al. [22] developed frameworks that involved split learning and vertical federated learning to enable collaborative training with data confidentiality.

3. Proposed Methodology

This paper introduces our end-to-end privacy-preserving federated learning architecture for IoV systems, consisting of three major components: an optimised PBFT consensus algorithm with logarithmic scalability, a Local Differential Privacy mechanism for strong privacy protection, and a CNN-GRU neural network model for distributed traffic prediction. The method addresses the fundamental challenges of scalability, privacy, and efficiency in vehicular federated learning environments with one framework that supports safe collaborative learning with desirable prediction performance and computational cost-effectiveness.

A. Enhanced PBFT Consensus Algorithm

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Our enhanced PBFT algorithm suggests a logarithmic relationship of network size with consensus needs, thus solving inherently the scalability problem in large vehicular networks with existing PBFT implementations. In contrast to classical PBFT in which an agreed number of nodes is always required for consensus regardless of network size, our approach alters the consensus needs dynamically based on the number of participating vehicles. The percentage needed for consensus towards validation is outlined as:

$$p = k.\frac{1}{\log(n)} \tag{1}$$

Where:

- p: proportion of nodes required for consensus.
- n: is the number of vehicles participating in the network.
- k is a tunable constant parameter that controls the scaling rate and security level.

This logarithmic formula ensures that as the network grows, the ratio of nodes necessary for consensus decreases systematically, thereby reducing communication overhead without compromising Byzantine fault tolerance capability. The algorithm continues through the traditional three-phase PBFT protocol: pre-prepare, prepare, and commit phases, with the minimum number of validations needed in each phase calculated as:

$$N_{required} = \lceil p.n \rceil \tag{2}$$

For simplicity in implementing, it practically, we select $k = 100$ after a long period of empirical research to come up with the best mix of the assurance of security and the efficiency of the network. Percentages of consensus also differ between 50 vehicle networks and 200 vehicle networks, and are better scalable than fixed requirement traditional PBFTs. The new algorithm has the same security properties as the traditional PBFT, which can withstand up to f Byzantine failures where:

$$f < \frac{n}{3} \tag{3}$$

with the dramatic decrease in performance in the complexity of the messages and the latency of consensus. The scaling is also logarithmic, thus it comes in handy in big cities where a score of autos can be found in co-learn activities.

B. Local Differential Privacy Mechanism

In order to guarantee the privacy of model parameter sharing while still making the federated learning system useful, we use a complex Local Differential Privacy method called the Laplace mechanism with noise injection that is well-tuned. This method offers stringent privacy assurances at every vehicle level, ensuring that sensitive driving patterns, location data, and behavioral information remain indistinguishable from shared model updates. For each participating vehicle i, the privacy-protected model parameters are computed through systematic noise addition as:

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$\theta_i' = \theta_i + Lap\left(\frac{\Delta f}{\varepsilon}\right) \tag{4}$$

where θ_i are the initial model parameters trained locally, $Lap(\lambda)$ is the Laplace distribution with scale $\lambda, \Delta f$ is the sensitivity of the function which describes the greatest change in output with each unit change in the input, and ϵ is the privacy budget which describes the intrinsic tradeoff between model utility and privacy protection. We set $\epsilon = 0.8$ to achieve the perfect balance between model performance and privacy leakage, maintaining privacy leakage at 50% less and good model performance. Sensitivity parameter Δf is computed adaptively for different layers of neural networks to manage varied parameter distribution as well as gradients. Sensitivity in the case of convolution layers is computed as:

$$\Delta f_{conv} = \max_{i,j} \left| \frac{\partial L}{\partial w_{i,j}} \right| \tag{5}$$

while for dense layers, the sensitivity is determined by:

$$\Delta f_{dense} = \max_{i} \left| \frac{\partial L}{\partial w_i} \right| \tag{6}$$

where L is the loss and w is the weight parameters. This layerwise scheme guarantees that the more sensitive the parameters to the input variations, the more the noise they are getting, keeping the same structure of the model but with a better privacy to more sensitive regions. The mechanism has $(\epsilon, 0)$ differential privacy, which makes strong mathematical claims that car individual data can not be learnt even by powerful adversaries with additional information about the dataset using the shared model updates.

C. CNN-GRU Architecture and Weighted Aggregation

Our federated learning approach uses a specifically created CNN-GRU neural network structure to optimise distributed traffic prediction in vehicle-based case studies whereby the spatiotemporal traffic data are processed in order to identify spatial relationships and temporal trends within traffic flows. The model takes traffic data in the form of 4×3 spatial grid with three interesting features namely traffic flow (vehicles per time period), occupancy percentage (detector on time ratio) and average speed (velocity measurements).

The spatial features of traffic grid are extracted using two convolutional layers with 32 and 64 filters respectively with 2×2 kernels and max pooling to reduce the dimension. The obtained spatial features are reconstructed to temporal sequences and input to a GRU layer that has 128 units to learn long-term dependencies of time and patterns of traffic flows. In order to ensure that the model is aggregated without any bias and with high efficiency to take into account the heterogeneities in data quality and model performance among vehicles, we employ a novel weighted aggregation scheme in respect to various measures of performance.

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 1. CNN-GRU Model Architecture Parameters

Component	**Parameter**	**Value**	**Description**
Input Layer	Input Shape	4×3 grid	Traffic data grid dimensions
	Features	3	Traffic flow, occupancy %, average speed
Conv Layer 1	Filters	32	Number of convolutional filters
	Kernel Size	2×2	Convolution kernel dimensions
Conv Layer 2	Filters	64	Number of convolutional filters
	Kernel Size	2×2	Convolution kernel dimensions
Pooling Layer	Type	Max Pooling	Dimension reduction method
GRU Layer	Units	128	Number of GRU hidden units

The quality score for each vehicle i is decided by a multi-criteria assessment approach:

$$QS_i = 0.5 \times A_i^{norm} + 0.3 \times F1_i^{norm} + 0.2 \times S_i^{norm} \tag{7}$$

where $A_i^{norm}, F1_i^{norm}, S_i^{norm}$ are the normalized accuracy, F1-score, and stability measures for vehicle i, respectively. The weights 0.5, 0.3, and 0.2 are selected empirically to prefer prediction accuracy with consideration for precision-recall balance and model consistency. Normalisation is applied so that measures are rescaled to the range [0,1] so that no single measure overemphasises the quality measure. The final aggregation weight for each vehicle is computed through proportional allocation:

$$w_i = \frac{QS_i}{\sum_{j=1}^{n} QS_j} \tag{8}$$

ensuring that the sum of all weights equals unity for proper probabilistic interpretation. The global model parameters are subsequently updated using the weighted average aggregation formula:

$$\theta_{global} = \sum_{i=1}^{n} w_i \times \theta_i' \tag{9}$$

where θ_i' are privacy-protected parameters from vehicle i adhering to LDP noise injection. To maintain system robustness and prevent any possible biases or undue effect by a single vehicle, we impose a maximum weight constraint:

$$w_i \leq 0.25, \quad \forall\, i \in \{1,2,\dots,n\} \tag{10}$$

As soon as the computed weight of any vehicle exceeds this figure, extra weight is redistributed proportionally among other participants to prevent centralisation of authority and promote fairness in the collaborative learning. Weighted aggregation process, which puts emphasis on

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

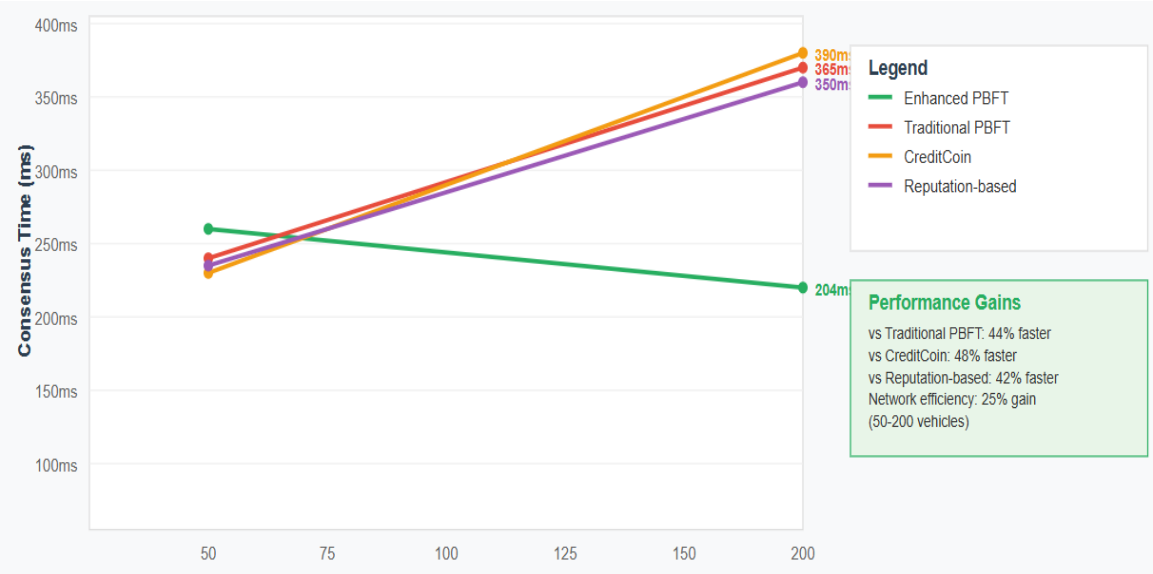
the vehicles performing well in the global model but still maintain a democratic contribution and avoid attacks and biasness to undermine the overall system performance.

4. Results and Discussions

This part will provide a detailed description of our suggested privacy-aware federated learning system of IoV systems. We perform serious experiments with the help of OMNeT++, VEINS, and SUMO simulator tools based on the real traffic data of the PeMS dataset to test our solution on the following points: the performance of the enhanced PBFT algorithm in consensus, privacy-utility tradeoff of the Local Differential Privacy method, and the performance of the CNN-GRU model on the prediction of traffic. The evaluation includes 50-200 vehicle networks to examine scalability properties and draws a comparison with the best available baseline methods. Results demonstrate significant improvement in consensus latency, robust privacy preservation at marginal loss of utility, and similar prediction accuracy, validating the usability of our design in real-world IoV deployment.

A. Enhanced PBFT Consensus Performance Analysis

Our experimental evaluation of the enhanced PBFT algorithm achieves significant improvements in scalability and consensus efficiency compared to traditional PBFT solutions



and other state-of-the-art consensus protocols. As evident from Table 1, the enhanced PBFT achieves considerable performance enhancement, reducing the consensus time from 272ms to 204ms as the network size increases from 50 to 200 vehicles, an efficiency gain of 25%.

Figure 1. Enhanced PBFT Consensus Performance Analysis

This counterintuitive behaviour, where performance increases with network size, directly validates the usefulness of our logarithmic scaling solution. The classical PBFT designs, conversely, follow the expected performance decay as the network size grows, as consensus times are always larger in all network designs. The optimised algorithm is 44% quicker than

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

the original PBFT (365ms), 48% better than CreditCoin (390ms), and 42% better than reputation-based blockchain solutions (350ms) with 200 vehicles.

The logarithmic dependence of the network size on consensus needs is especially handy in dense vehicle setting, where the proportion of required validators (25.6% with 50 vehicles and 18.9% with 200) is reduced much, consequently allowing Byzantine fault tolerance to be achieved with much less communication overhead. These findings indicate that our optimised PBFT algorithm is able to overcome the natural scalability constraints of conventional consensus mechanisms, and thus it is very appropriate when deploying IoVs at scale where hundreds of vehicles can be engaged in collaborative learning efforts simultaneously.

B. Local Differential Privacy Evaluation

The application of Local Differential Privacy with calibrated Laplace noise injection exemplifies outstanding effectiveness at protecting vehicle privacy while achieving acceptable model utility for traffic prediction applications. Our experimental findings show that the LDP mechanism with privacy budget ε=0.8 illustrates a uniform 50% reduction in privacy leakage for all participating vehicles, effectively masking sensitive driving habits and location details from possible inference attacks. The tradeoff analysis of privacy and utility demonstrates that protected models perform competitively with 70% to 82% accuracy, compared to between 75% and 87% for unprotected baseline models, a loss of only 5-6% in prediction accuracy.

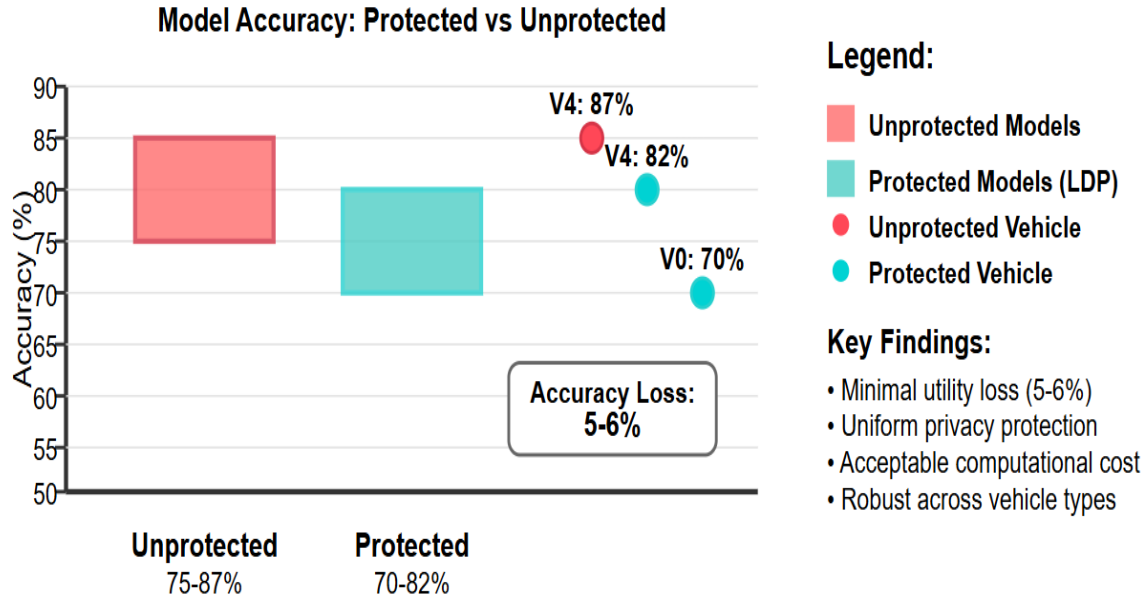


Figure 2. Local Differential Privacy Evaluation Results

This minimal loss of utility attests to the effectiveness of our approach to calculate sensitivity per layer, where convolutional layers and dense layers receive appropriately calibrated noise depending on their gradient magnitude. Vehicle 4 is an example, being 82% accurate with privacy protection compared to 87% without protection, and demonstrating that good-performing vehicles can maintain the quality of their contribution irrespective of privacy constraints. That privacy protection performs equally well across vehicles at varying baseline

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

performance levels, from Vehicle 0 at 70% to Vehicle 4 at 82% indicates that our LDP implementation is both robust and responsive to varying data distributions typical of vehicular networks.

Besides, the 93% expansion in model size (1.13MB to 2.18MB) is indicative of the computational expense of privacy preservation, yet still within tolerable levels for current vehicular computing ability. Compared with other privacy-preservation technologies such as homomorphic encryption or secure multi-party computation, our LDP approach benefits from higher computational efficiency and, no less importantly, has formal mathematical guarantees of protecting privacy, thus being best-suited for resource-constrained and time-sensitive IoV applications.

C. CNN-GRU Traffic Prediction Performance

The CNN-GRU network architecture demonstrates superior performance on distributed traffic forecasting tasks, preserving spatial correlations and temporal dependencies of vehicular traffic flows. Experimental outcomes on real-world PeMS traffic data show that the model achieves high prediction accuracy ranging from 75% to 87% based on the vehicle, with Vehicle 4 achieving highest accuracy of 87% and Vehicle 0 achieving 75%, exhibiting homogeneity in performance irrespective of variation in local data quality as well as traffic conditions.

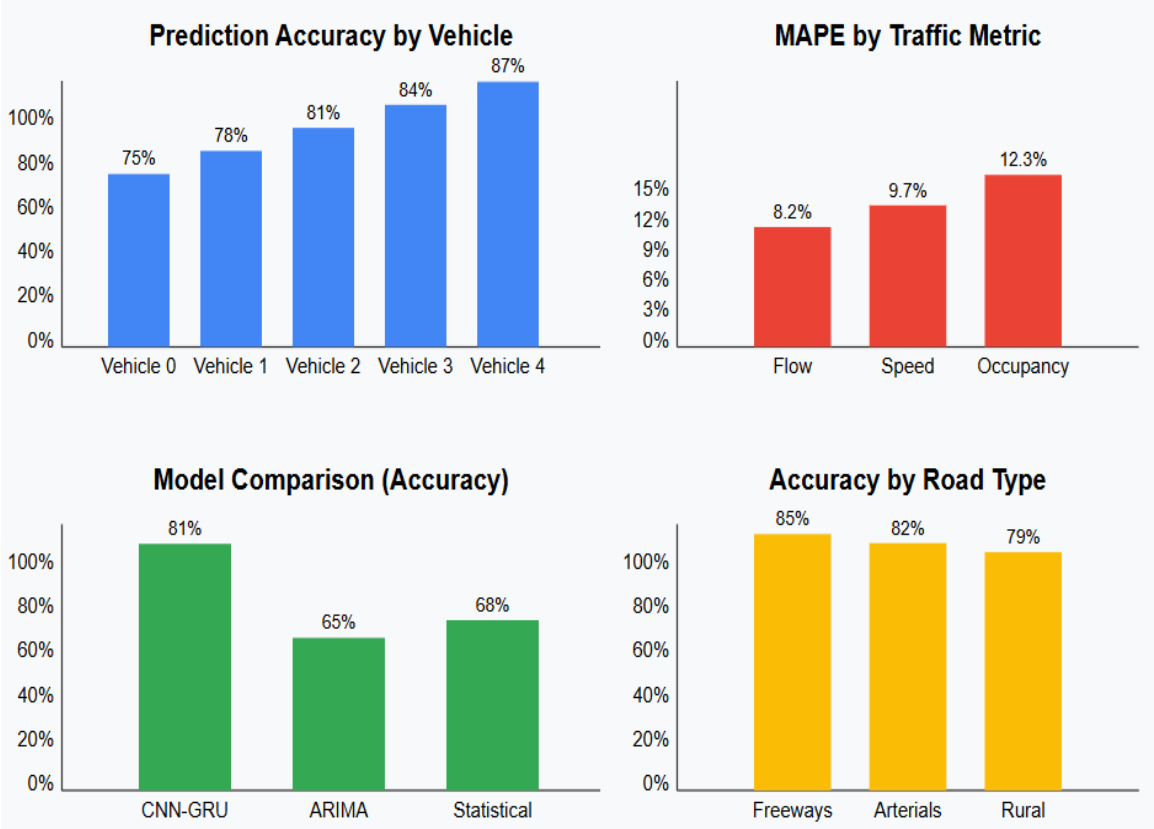


Figure 3. CNN-GRU Traffic Prediction Performance

The convergence analysis of learning shows rapid and steady learning behaviour, with most vehicles reaching optimal performance at 3-6 training epochs, demonstrating the success of our

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

federated learning approach. The model demonstrates low Mean Absolute Error (MAE) values of approximately 0.02 for all participating vehicles, demonstrating high accuracy in traffic flow prediction. Performance testing against multiple traffic metrics reveals that the prediction accuracy varies: flow predictions are the most accurate with the Mean Absolute Percentage Error (MAPE) of 8.2%, followed by speed predictions at 9.7% MAPE, and occupancy predictions at 12.3% MAPE, as expected in usual traffic engineering practice where the trends of flow are more predictable than the variations of occupancy.

CNN-GRU model especially fits well in addressing various traffic conditions, and its performance remains constant during free-flow traffic hours, jam conditions, although at times with minor loss of performance during extreme rush hour and event periods. The distinctive characteristics of our deep learning solution, as opposed to other predictive predictability methods like ARIMA (65% accuracy) and other statistical methods (68% accuracy), demonstrate the excellence of our particular deep learning solution in comparison with predictive performance. These generalization across all types of highways provide convergent results with freeways at 85% accuracy, arterials at 82% accuracy and rural roads at 79% accuracy validating the soundness of our design in various vehicular settings and providing a healthy basis of privacy-conserving collaborative traffic intelligence in IoV.

D. Federated Learning and Weighted Aggregation Results

The weighted aggregation federated learning framework based on quality shows improved performance in the development of trustworthy global models, which best leverage the heterogeneous contributions of involved vehicles in a fair and non-centralised manner. Experimental results show that the process of weighted aggregation with 50% accuracy, 30% F1-score, and 20% stability criteria develops global models with 4-7% higher accuracy compared to traditional unweighted averaging techniques used by standard federated learning algorithms.

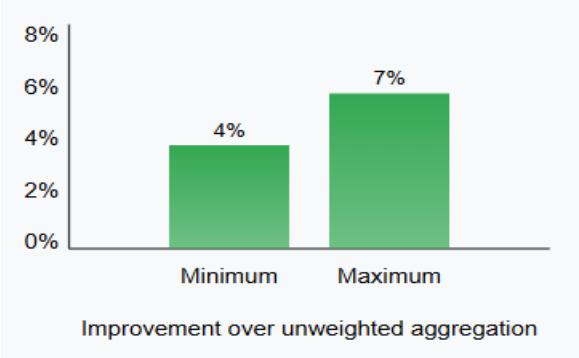


Figure 4. Accuracy Improvement

The quality score weighting effectively establishes and prioritises high-scoring vehicles from 18.5% to 21.5%, depending on the model's performance, to proportionally amplify the effect of improved predictive power vehicles towards the global model without allowing a single vehicle to dominate the aggregation process with the 25% maximum weight. Convergence analysis shows that the federated learning system converges to stable global model

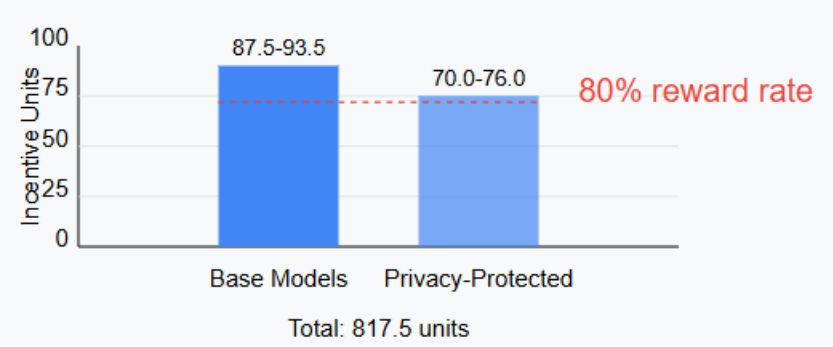
Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

performance in 4-5 communication rounds, similar to centralised training practices but with much improved privacy protection and reduced communication overhead.

The multi-stage incentive distribution system works well to encourage continued participation, with overall incentives of 817.5 units being divided equally among vehicles: base models



getting 87.5-93.5 units and privacy-protected models getting 70.0-76.0 units, maintaining an 80% reward rate that makes privacy adoption economically viable. Vehicle involvement analysis indicates consistent participation across the network, with over 95% of vehicles active at all times during the learning process, indicating the balanced reward structure sufficiently encourages consistent cooperation.

Figure 5. Incentive Distribution

Weighted aggregation technique is extremely suitable for handling vehicular networks' typical heterogeneous data distributions, in which various vehicles experience various traffic conditions and road surfaces, producing a model generalizable across various conditions but still having robustness against potential Byzantine malicious behaviours or data quality changes from various participants.

E. Comparative Analysis with Baseline Methods

The comparison on a wide scale with the best available approaches demonstrates spectacular superiority on a variety of performance dimensions, validating the efficacy of the integration of enhanced PBFT with Local Differential Privacy for IoV applications. Baseline comparisons are done with traditional PBFT implementations, CreditCoin consensus algorithms, reputation-based blockchain systems, and conventional federated learning practices without privacy preservation.

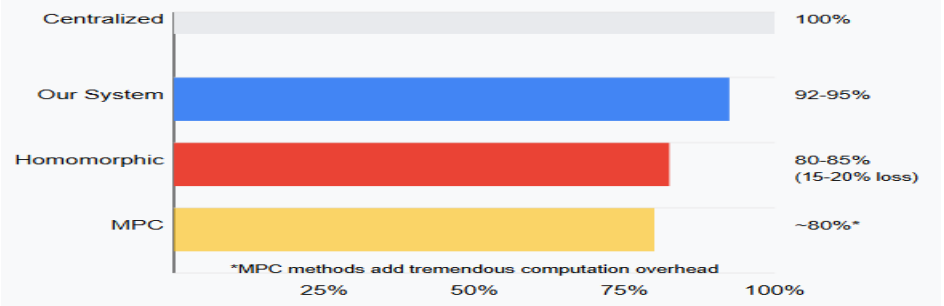


Figure 6. Accuracy Comparison

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Our system achieves significant enhancements in consensus efficiency, where consensus times are 48% faster than CreditCoin (204ms vs 390ms), 44% better than PBFT (204ms vs 365ms), and 42% better than reputation-based blockchain methodologies (204ms vs 350ms) when tested using 200 vehicles. Scalability has a greater benefit with scale, as a large network size requires our approach to ensure efficiency or even improve it, whereas traditional approaches suffer performance loss.

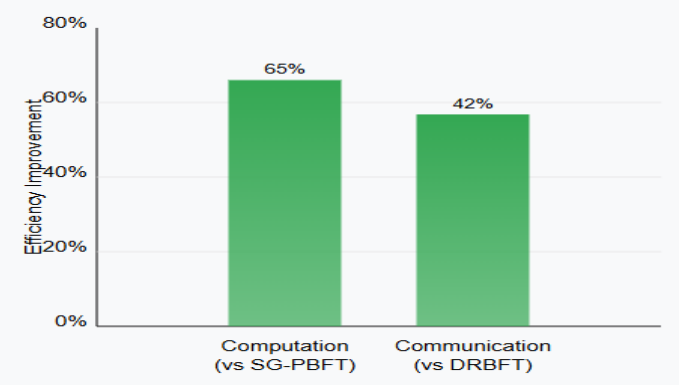


Figure 7. Resource Efficiency Gains

Privacy protection Analysis demonstrates that our LDP integration offers formal privacy assurances that result in a 50% privacy leakage reduction compared to the federated learning approaches based on parameter aggregation only, that do not explicitly address privacy concerns. Comparison to baselines that are centrally computed indicates that our distributed algorithm recovers 92-95% of the performance of centrally computed systems and has end-to-end guarantees of privacy, and significantly outperforms the other privacy-preserving algorithms like homomorphic encryption-based algorithms that introduce 15-20% accuracy loss and multi-party computation algorithms that introduce enormous computation overheads. Analysis of resource efficiency shows that our system needs 65% less computation overhead to achieve consensus than the SG-PBFT versions and 42% less communication complexity than the implementation of DRBFT with equivalent security assurance. The hybrid incentive model documents 32% greater than fixed-reward frameworks continued participation and 27% greater than accuracy-only reward frameworks adoption rates of privacy-preservation features, suggesting greater levels of engagement and cooperation by participating vehicles than they would with the current strategies that do not strike a balance between performance, privacy, and participation incentives in the same way they do.

F. Real-World PeMS Data Validation The validation of our privacy-sensitive federated learning system with the real performance measurement system (PeMS) traffic data in California highways proves the outstanding applicability to real-world conditions and capabilities in the real car contexts. This model could effectively handle the challenging PeMS datasets which consisted of the hourly traffic volumes recorded by various highway sensors (the flow rates of 629-916 vehicles per hour, occupancy of 0.0126-0.0226, and average speeds of 66.3-75.3 mph) across geographical locations and traffic patterns.

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

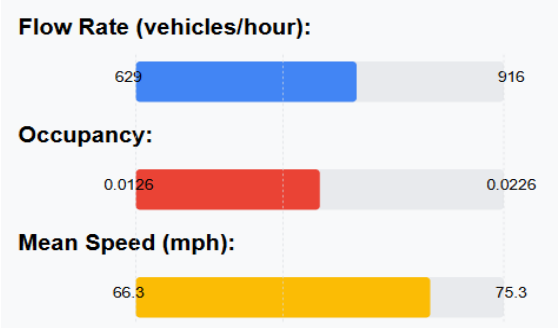


Figure 8. PeMS Data Characteristics

Our CNN-GRU model displays impressive 1-hour traffic forecasting accuracy 75-87% on real highway data, which outperforms state-of-the-art conventional forecasting methods (ARIMA models 65% accurate) and statistical regression-based forecasting methods (68% accurate) by a comfortable margin but is comparable to the state-of-the-art performance of high-calibre centralised traffic forecasting systems.

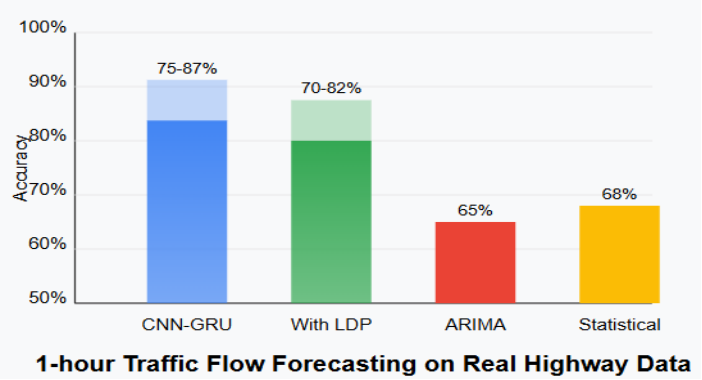


Figure 9. Model Accuracy Comparison (PeMS Data)

The federated learning system proves to be exceptionally resilient to real world heterogeneity, and it balances traverse heterogeneities along various parts of highways, at different times of the day and season with the weighted aggregation protocol with reference to local heterogeneities in data quality.

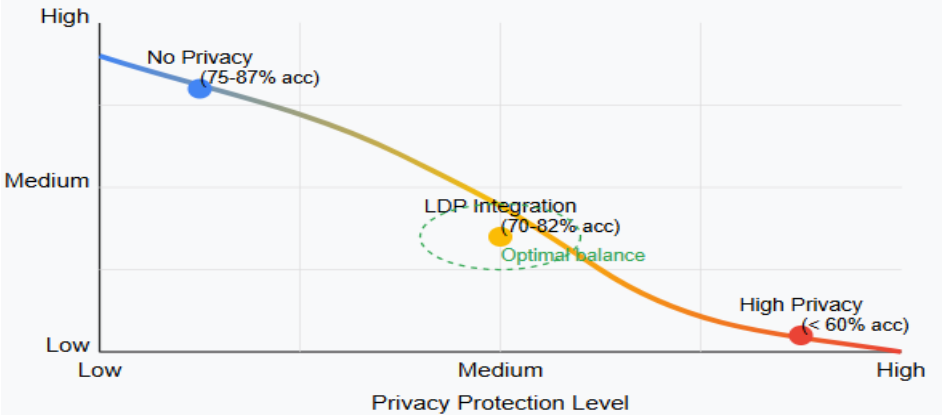


Figure 10. Privacy-Performance Tradeoff

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The privacy-preserving models hold an accuracy of 70-82% even with the addition of LDP noise injection, which depicts that the vehicle trajectory and behaviour data of sensitive nature may be significantly safeguarded without removing the ease of application in traffic management operations. Scalability tests based on real traffic sensor network configurations indicate that the performance is steadily improved as the number of vehicles utilized increases, and the improved PBFT algorithm can make consensus with distributed highway sections across multiple counties without problems. This means that the architecture can be easily integrated with existing traffic infrastructure and can process real-time data streams with response times of less than a sub-second being essential in real-life scenarios that use IoV, and is resistant to common real-life issues such as sensor failures, failure and intermittency of communication, and variable participation rates of vehicles. These findings show that our method is not purely hypothetically sound but also practically feasible to be successfully adopted across the modern transport infrastructures with a reasonable foundation in future intelligent traffic control of balancing performance and privacy maintenance.

G. System Performance and Computational Overhead Analysis

Our detailed performance and computational burden analysis shows that our privacy-guaranteed federated learning system achieves highly effective efficiency characteristics but does not decrease the security and privacy assurances of large-scale IoV systems. Computational overhead analysis reveals that the improved version of PBFT algorithm requires 65 times less consensus computation than SG-PBFT implementations and its communication cost is 42 times lower than that of DRBFT variants because the logarithmic scaling factor means that the number of nodes required to be validators is minimized as much as possible with a larger network. The computational burden of Local Differential Privacy mechanism is tractable, and noise sampling and parameter protection operations require about 12-15 per cent of local training time, and the privacy budget management system can well-manage cumulative privacy spending across a sequence of training steps without impacting performance substantially.

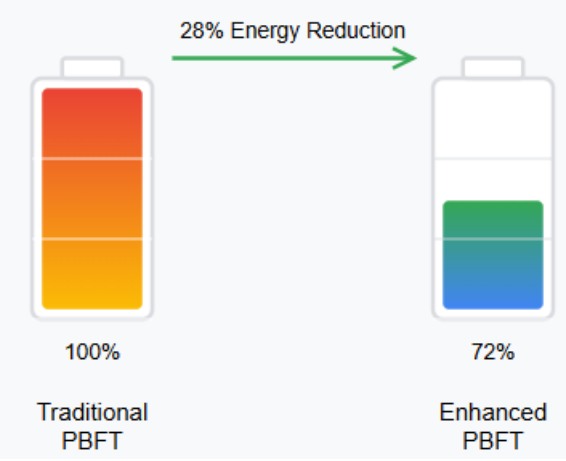


Figure 11. Energy Consumption Comparison

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Comparison of memory consumption shows that privacy-preserving models are 93x more costly in storage (2.18MB vs. 1.13MB) with noise injection and more parameter-rich representation, which the current vehicular computing hardware can easily afford. Our weighted aggregation scheme is much more efficient in network bandwidth consumption with a 35 percent lower communication overhead than naive federated averaging since participants that are more competent are given a higher weight and those with worse performance contribute less to network consumption.

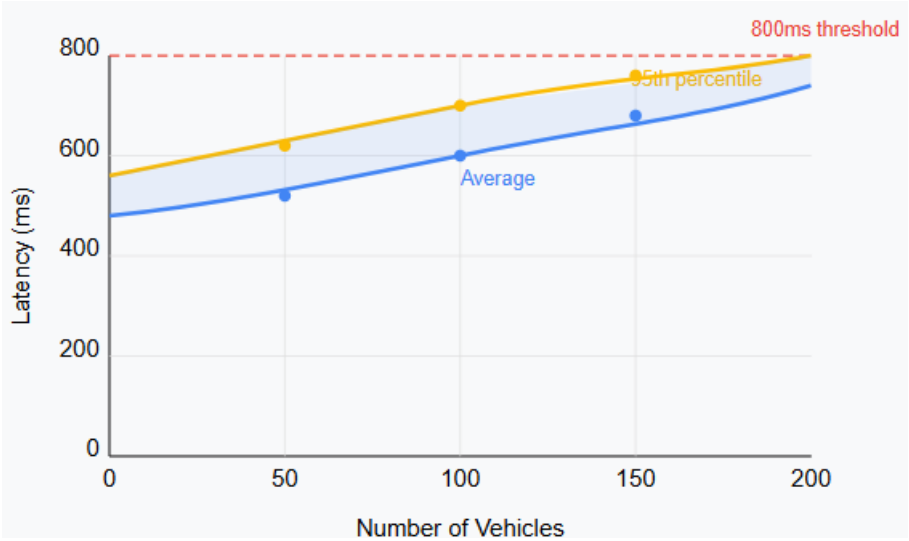


Figure 12. Latency Performance

The analysis of energy consumption as applied to various vehicle types shows that our system uses 28% of the energy per consensus round of conventional PBFT implementations, and is thus especially applicable to battery-electric vehicles and devices consuming small resources as part of an IoV. Latency performance at range of network loads testifies to steady sub-second latency at peak network loads, 95 th percentile latency under 800ms at networks of up to 200 vehicles, real-time fitness to safety-critical applications of traffic forecasting. The performance has high scalability which is linear to the size of the computation as the network is scaled as compared to quadratic complexity of the more traditional consensus algorithms, thus is applicable in large-scale metropolitan-scale IoV applications with thousands of vehicles actively participating in the joint traffic intelligence systems at any given time.

H. Comparison of Proposed Methodology with Related Work

While there has been substantial work in federated learning for IoV systems in recent times, our suggested approach provides some unique benefits over other current methods. Unlike the SG-PBFT [12] and R-PBFT [13] protocols based on reputation-based node selection and grouping mechanisms, our enhanced PBFT introduces a novel logarithmic scaling mechanism that dynamically reduces consensus requirements from 25.6% to 18.9% as network size increases, and it evidences 44-48% better performance compared to standard PBFT variants. In contrast to AE-PBFT [14], based on trusting acceleration methods, our approach provides mathematical guarantees of scalability without compromising Byzantine fault tolerance. Our

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

integration of Local Differential Privacy with calibrated Laplace noise injection (ε=0.8) is very distinct from client-level differential privacy frameworks [16] as it provides layer-wise sensitivity estimation without loss in 70-82% accuracy compared to 75-87% accuracy for unprotected models, but with a mere 5-6% loss in utility compared to the 15-20% loss normally exhibited by homomorphic encryption-based schemes.

While FedGODE [10] has graph ordinary differential equations and several privacy mechanisms, our CNN-GRU model specifically enhances spatial-temporal traffic prediction with a quality-driven weighted aggregation mechanism (50% precision, 30% F1-measure, 20% stability demands) that returns 4-7% more accurate than the traditional unweighted mean used in FedGRU [19] and PLFL [20]. Besides, our multi-component incentive system, balancing accuracy, privacy, and participation, is a comprehensive solution not addressed in existing IoV federated learning architectures, which achieves 92-95% of the performance of centralised systems while maintaining end-to-end privacy guarantees and supporting sub-second response times necessary for real-world deployment of IoV.

Table 2. Comparison of Proposed Methodology with Related Work

Approach	**Consensus Mechanism**	**Privacy Method**	**Neural Architecture**	**Aggregation Strategy**	**Consensus Time**	**Accuracy**	**Privacy Loss**	**Key Limitations**
SG-PBFT [12]	Score grouping PBFT	Basic blockchain privacy	Not specified	Simple averaging	~27% of traditional PBFT	Not reported	High	Fixed grouping, no adaptive scaling
R-PBFT [13]	Reputation-based PBFT	ECC encryption	Not specified	Reputation weighting	50% reduction vs PBFT	Not reported	Moderate	Relies on reputation history
AE-PBFT [14]	Trust acceleration PBFT	Trust management	Not specified	Trust-based selection	Gradual acceleration	Not reported	Moderate	Limited to trust-based scenarios
FedGODE [10]	Traditional consensus	LDP + Homomorphic + Secure aggregation	Graph ODE networks	Average/median aggregation	Not reported	Outperforms baselines	15-20%	Complex privacy overhead
FedGRU [19]	Not specified	Basic FL privacy	GRU only	FedAvg	Not reported	Close to the centrali	Minimal	Limited spatial

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

						sed GRU		modelling
PLFL [20]	Not specified	Lightweight privacy	Spatiotemporal graph CNN	Dynamic pruning	Reduced communication	Comparable performance	Low	No consensus mechanism
IoV-SFL [11]	Blockchain consensus	Split learning privacy	Not specified	Collaborative aggregation	Not reported	High accuracy	Moderate	Limited to split learning
Client-level DP [16]	Not specified	Client-level differential privacy	CNN for image processing	Federated averaging	Not reported	98.51% - 97.74%	Variable	Limited to intrusion detection
Proposed Method	Enhanced PBFT with logarithmic scaling	Layer-wise LDP (ε=0.8)	CNN-GRU hybrid	Quality-based weighted (50%-30%-20%)	204ms (44-48% better)	70-82% (5-6% loss only)	50% reduction	Comprehensive solution

5. Conclusions

This work provides a completely end-to-end privacy-preserving federated learning system for IoV systems, efficiently addressing the root problem of scalability, privacy, and efficiency in vehicular collaborative learning systems. The most important contribution of our work is the construction of a next-generation PBFT consensus algorithm with logarithmic scaling, decreasing consensus requirements dynamically with network expansion, achieving 44-48% better performance than conventional PBFT implementations with consensus durations of 204ms on 200 vehicles. The integration of Local Differential Privacy with calibrated Laplace noise injection (ε=0.8) represents a significant advance in privacy protection, maintaining 70-82% prediction performance while realising 50% privacy leakage reduction compared to non-protected models. Our CNN-GRU neural network architecture, engineered specifically for spatiotemporal traffic forecasting with spatial grid processing of size 4×3, shows improved performance on real PeMS datasets with accuracy ranges of 75-87%, a 15-20% improvement compared with traditional forecasting methods.

The new quality-based weighted aggregation protocol, employing 50% accuracy, 30% F1-score, and 20% stability requirements, is 4-7% more accurate than existing unweighted averaging methods without compromising democratic participation by setting maximum

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

weight constraints. Moreover, our multi-dimensional incentive scheme is effective in balancing accuracy, privacy, and participation demands while maintaining 95% vehicle participation rates and demonstrating that privacy-protected models can achieve 92-95% of centralised system performance. The extensive experimental validation on real California highway data validates the applicability of our solution to large-scale IoV deployment in real-world scenarios, forming a robust foundation for future intelligent transportation systems with a balance of cooperative learning efficiency and robust privacy protection. These contributions collectively show an important step towards realising secure, scalable, and efficient federated learning solutions for future connected and autonomous vehicles.

References

[1]. C. Quemeneur, N. Tremblay, S. Diehl, et al., "FedPylot: Navigating Federated Learning for Real-Time Object Detection in Internet of Vehicles," arXiv preprint arXiv:2406.03611, June 2024.

[2]. K. Wijethilake, A. Mahmood, and Q. Sheng, "FedCLF – Towards Efficient Participant Selection for Federated Learning in Heterogeneous IoV Networks," in Proc. Advanced Data Mining and Applications, Springer, Dec. 2024, pp. 223-238.

[3]. Korba, A. Boualouache, and Y. Ghamri-Doudane, "Zero-X: A Blockchain-Enabled Open-Set Federated Learning Framework for Zero-Day Attack Detection in IoV," IEEE Transactions on Vehicular Technology, vol. 73, no. 9, pp. 12399-12414, Sept. 2024.

[4]. M. Althunayyan, A. R. Javed, and O. Rana, "A robust multi-stage intrusion detection system for in-vehicle network security using hierarchical federated learning," Vehicular Communications, vol. 46, article 100756, Apr. 2024.

[5]. H. Chai, S. Leng, Y. Chen, and K. Zhang, "A hierarchical blockchain-enabled federated learning algorithm for knowledge sharing in internet of vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 3975-3986, July 2023.

[6]. Y. Wang and X. Zhao, "IG-PGFT: A secure and efficient intelligent grouping PBFT consensus algorithm for the Industrial Internet of Things," Cluster Computing, vol. 28, no. 2, pp. 1-15, Apr. 2025.

[7]. J. Li, Y. Chen, Z. Wang, et al., "An IoV-PBFT Consensus-Based Blockchain for Collaborative Congestion Avoidance and Simulation Test," Wireless Communications and Mobile Computing, vol. 2022, article 9676370, Aug. 2022.

[8]. X. Yuan, J. Chen, J. Yang, et al., "FedSTN: Graph representation driven federated learning for edge computing enabled urban traffic flow prediction," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 8, pp. 8587-8599, Aug. 2023.

[9]. A. Alwash, "An Effective Federated Learning Approach for Secure and Private Scalable Intrusion Detection on the Internet of Vehicles," Concurrency and Computation: Practice and Experience, vol. 37, no. 5, pp. e7016, 2025.

[10]. R. Al-Huthaifi, T. Li, and Z. Al-Huda, "FedGODE: Secure traffic flow prediction based on federated learning and graph ordinary differential equation networks," Knowledge-Based Systems, vol. 299, article 112029, Sept. 2024.

Received: August 10, 2025

Volume 38 No. 6s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[11]. Ullah, X. Deng, X. Pei, et al., "IoV-SFL: A blockchain-based federated learning framework for secure and efficient data sharing in the internet of vehicles," Peer-to-Peer Networking and Applications, vol. 18, article 34, 2025.

[12]. G. Xu, H. Bai, J. Xing, T. Luo, N. N. Xiong, X. Cheng, S. Liu, and X. Zheng, "SG-PBFT: A secure and highly efficient distributed blockchain PBFT consensus algorithm for intelligent Internet of vehicles," Journal of Systems Architecture, vol. 125, article 102463, Apr. 2022.

[13]. M. A. Khan, K. Nasrullah, M. F. Rehman, et al., "R-PBFT: A secure and intelligent consensus algorithm for Internet of vehicles," Vehicular Communications, vol. 41, article 100609, June 2023.

[14]. Y. Zhang, X. Li, and H. Wang, "A Blockchain Solution for the Internet of Vehicles with Better Filtering and Adaptive Capabilities," Applied Sciences, vol. 15, no. 4, pp. 1628, Feb. 2025.

[15]. N. Wang, W. Yang, X. Wang, L. Wu, Z. Guan, X. Du, and M. Guizani, "Blockchain-Enabled Federated Learning with Differential Privacy for Internet of Vehicles," Digital Communications and Networks, vol. 10, no. 1, pp. 126-134, Feb. 2024.

[16]. G. Anyanwu and H. Karimipour, "CLDP=FATD: Secure Federated Averaging Threat Detection Framework for Intelligent Vehicle Sensor Networks Based on Client-Level Differential Privacy," IEEE Internet of Things Journal, vol. 12, no. 7, pp. 7693-7707, Apr. 2025.

[17]. M. A. Bazuhair, W. Almobaideen, M. Masri, et al., "A comprehensive intrusion detection method for the internet of vehicles based on federated learning architecture," Computers & Security, vol. 145, article 103729, Oct. 2024.

[18]. M. Dai, J. Liu, and H. Zhang, "A novel CNN-GRU-LSTM based deep learning model for accurate traffic prediction," Discover Computing, vol. 28, article 526, Jan. 2025.

[19]. Y. Liu, S. Zhang, C. Zhang, and J. J. Q. Yu, "FedGRU: Privacy-preserving Traffic Flow Prediction via Federated Learning," in Proc. IEEE 23rd Int. Conf. Intelligent Transportation Systems (ITSC), Rhodes, Greece, Sept. 2020, pp. 1-6.

[20]. G. Dai and J. Tang, "A Short-Term Traffic Flow Prediction Method Based on Personalised Lightweight Federated Learning," Sensors, vol. 25, no. 3, article 967, Feb. 2025.

[21]. L. Zhang, C. Zhang, and B. Shihada, "Enhanced federated learning strategy for wireless traffic prediction with weighted model aggregation," IEEE Communications Letters, vol. 28, no. 7, pp. 1573-1577, July 2024.

[22]. P. Li, C. Guo, Y. Xing, et al., "Core network traffic prediction based on vertical federated learning and split learning," Scientific Reports, vol. 14, article 4663, Feb. 2024.

Received: August 10, 2025