Proceedings of the 4th International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/2026.33761

Federated Learning for Intelligent Transportation Systems: A Comprehensive Investigation

Haozhen Wen
School of Computer Science, Nanjing University, Nanjing, China
231840160@smail.nju.edu.cn

Abstract. Because of the fast development of ITS technologies, the amount of data that could be collected from connected cars, roadside infrastructures, and even mobile phones has significantly increased. Centralized approaches to collecting and processing data bring up a number of problems associated with data security, distribution difficulties, and compliance issues. One of the promising decentralized frameworks that have recently gained popularity is federated learning (FL). FL is the process in which no raw data but rather the updates to the model are transmitted from one client to another. Thus, using FL preserves data privacy while lowering the costs related to communication. In this paper, the potential applications of federated learning in transportation are considered, paying special attention to its ability to solve specific tasks associated with traffic forecasting, autonomous driving, and V2X. The paper first describes the concept of federated learning and its implementation within the sphere of transportation, where the characteristics of data heterogeneity and mobility add new complexity to the approach. Then, literature is systematically divided into three primary spheres of application: (i) traffic flow and mobility prediction, (ii) autonomous driving coordination, and (iii) intelligent mobility services. For each of the spheres, architectural design choices and model aggregations as well as the effect of non-independent and identically distributed (non-IID) data are analyzed.

Keywords: Federated Learning, Intelligent Transportation Systems, Data Heterogeneity, Privacy Preservation

1. Introduction

The ITS attempts to improve the transportation systems in terms of efficiency, safety, and sustainability using modern information and communication technologies.

Connected and autonomous vehicles, CAVs along with various IoT sensors have also generated a huge amount of data on transportation [1, 2]. In traditional data-processing methods, all the acquired original data will be transmitted to a central server. This results in several issues: it is necessary to shield people's and confidential stuff like whereabouts so that it does require steps taken to protect the users' information, processing these huge volumes of info is quite costly, and various safety and management procedures obstruct inter-departmental data exchanges [3, 4]. Though there have been many attempts by the scientific world to overcome this obstacle, both old data handling methods as

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.33761

well as newer forms of FL are unable to solve all of these issues at once and do so even less when it comes to the constantly changing environment of vehicular networks.

Hence federated learning is a good tech to store the data locally but send the model. It can make sure that protecting the own privacy as well as save some core-network bandwidth [5]. To be more precise regarding Federated Learning: it is feasible for the edge node to transmit only a change on the model like gradient rather than original data which could guarantee that all kinds of information would not go through core network. But because there is dynamic change with traffic environment as well as the mobility and heterogeneous nature of data, implementing this on Federated Learning in those scenarios would have its own set of problems and therefore it is understandable why reviewing the current literature regarding FL application for Transportation.

2. Applications of federated learning in transportation

These applications can be systematically categorized based on functional roles in ITS, including (i) traffic-level prediction, (ii) vehicle-level intelligence, and (iii) network-level coordination.

2.1. Traffic flow and mobility prediction

2.1.1. Real-time traffic flow forecasting

Real-time traffic flow prediction is probably the best opportunity this paper has to mention my FL before ITS. In terms of FL, traffic flow prediction models can be developed collaboratively by traffic control centers located in different regions around the world without exposing any private information about the place [1]. Different from the traditional method which needs to gather raw traffic data at various places to refine the model so that it could be exposed and unstable due to its leakage [3]. FL frameworks allows traffic management centers and edge device to create their own model individually on raw traffic information like traffic volume, speed, occupancy by not sharing this information with each other's.

2.1.2. Cross-regional transit demand prediction

Xiang Jiacheng's transit demand prediction make use FL for there is big non-IID one can see across different cities transport data [6]. In the urban setting there are clearly different kinds of traffic, due to different demographics and different types of roads [3]. Fl-based forecasting structure considers such kind of variability using the idea of personal federations; thus, each city has its very own version and they share a common spatiotemporal representation as in parameter averages. Some more recent examples are the MoE models that dynamically routes its prediction query to various sub model that is specialized on different part of the urban area. So, they do beat the purely local train and a simple Fed-Avg baseline with the multi city - transit data sets [1].

2.1.3. Heterogeneous traffic state estimation

Heterogeneous traffic state estimation deals with the fact that real world traffic sensor deployments have quite some variations in terms of data quality, time coverage and measurement type. As for the operational ITS infrastructure, there are a variety of different sensors such as inductive loops and crowdsourced GPS traces which all have different sample rates and levels of accuracy [2]. More evolved versions of these types of FL frameworks could also be used to tackle multi variable heterogeneity by giving them different weights, those would then be updated periodically according

DOI: 10.54254/2755-2721/2026.33761

to how "good" they did in regards for that specific client's dataset [6]. Dual-driven is like global anomaly detection of the corrupted stream + some local noising. So, such a heterogeneous Fl Structure is more Robust to prediction in real life since when it encounters unbalanced dataset it does not perform badly.

2.2. Autonomous driving coordination

2.2.1. Collaborative perception across vehicles

Collaborative perception with FL, this will help CAVs improve in terms of detecting objects and the environment but they don't have to give out their raw sensor stream to 3rd parties Different vehicles gather various experience from different road conditions, accumulate a lot of perceptual knowledge about observation by themselves. FL can make the vehicle have access, by means of secure gradient sharing among themselves to other people as well, both what it has experienced on its own and what others have experienced, so that their samples are greater. As of lately it is possible to see more applications regarding FL on CAVs perception that use domain adaptation methods, so that they can cope with heterogeneous problems that result out of using different sets up as for each manufacturer [7]. It's better at generalizing towards the rare, more important safety case of a pedestrian that is occluded.

2.2.2. Distributed Motion Planning and trajectory prediction

In Distributed Motion Planning and trajectory prediction using FL it handles the computational issue caused by vehicle's behavior being processed in one place and privacy as well. For those motion-planning algos that use social compliance, they need lots of driving logs to learn their policies. And if collecting these logs, having private data that shows sensitive things [4]. With FL based frameworks allows cars improve on their own lanes changing knowledge as well as how they handle intersections with other cars however it is kept secret what any given car has driven before [7]. And meta-learn and be adaptable to all types of these very different drive cultures on any part of the planet Earth. And hence, a federated way of achieving better generalization in all sorts of different driving conditions.

2.2.3. Predictive maintenance and vehicle diagnostics

And another is called predictive maintenance when like a whole bunch of cars all make some forecast about which part will start to break down, separate for each car and don't collect or share proprietary auto data [7]. In modern vehicles, a lot of sensors monitor the engines and batteries as they degrade [2]. And then with Fl, fleet owners makers get to make use of this spread-out watch without getting on nobody's consumer privacy and competitors feel good/bad about power train. Federated prediction models would have the ability to detect very soon-to-come failures, like detecting out-of-the-ordinary shaking before bearings get worn down so that all the distributed populations are improving the model in an anonymous way. Techniques like this can provide statistical significance for an anomaly detection threshold because there is already some variance in the operating environment.

DOI: 10.54254/2755-2721/2026.33761

2.3. V2X communication and edge intelligence

2.3.1. Resource Allocation Optimization in V2X networks

Resource Allocation Optimization is another important application of FL in V2X Systems where the allocation of resources like spectrum, power & computational power must be done as per rapidly changing network topology [2]. A lot of movement when there's a traditional central algorithm that doesn't respond in real-time because the channel state info has become old by the time it optimizes with the cloud [8]. With FL this paper can now learn good, almost optimal schedule policies locally, i.e., at the edge on an edge node as opposed to centrally where it will be much slower. In the better version is to use a federation system such that it can allow many agents' reinforcement learning but no on costs their own location or even where are they currently moving. And better performance is found from them too, mostly in crowded towns.

2.3.2. Federated learning over Vehicular edge computing

In FL over Vehicular edge computing (VEC) multi access edge is coming in the vehicular network which solve the problem of computations and timing due to vehicle' s limited bandwidth [2]. The VEC architecture would employ something closer to the mobile vehicle(s), like an RSU, so it can do those FL Agg calculations with far less delay [8]. But vehicles are mobile too: there can be temporary outages at the time of handoff and clients may not participate for all times. With regards to FLs which considers FL that takes the mobility factor into consideration, it represents this dynamic. It dynamically determines how edges are formed or which participants should be selected for training.As Data Offloading schemes get added to such structure, it lets cars move workloads about so as when it is best (when closer-in spots), channel conditions there are clearer.

2.3.3. Secure and adaptive model aggregation for V2V collaboration

It's safe and it is possible to change a combination model to deal with the question about whether it will be continued Automatic to Car's V2V group: Traditional FL algos make steady popls & friendly behavior, but that's often broken in vehicular places where people are untrustworthy & there can be bad guys. In modern federated learning architectures within V2V systems, an Aggregator System is responsible for managing the overall process and assessing whether a particular vehicle is legitimate based on contextual information. Privacy mechanisms like diff private or homomorphic encryption give good ways of sharing what the car knows between themselves. Concept drift detection and mitigation are also incorporated to ensure model stability over time.

3. Challenges & future directions

3.1. Existing challenges

3.1.1. Heterogeneity and non-IID data in traffic prediction

Federate Traffic forecast scene can clearly show the great problem for real convergence because of the data heterogeneity and spread convergence is blocked [3]. There is an awful lot of variances around when-space and accuracy for one sensor on the traffic. So, what ultimately left with are very non-IID datasets which aren't amenable to many optimization [6]; Simple federated averaging performs poorly because low-quality or limited data from some clients can distort the global model,

DOI: 10.54254/2755-2721/2026.33761

and distribution differences between urban and rural groups make knowledge sharing more difficult. Moreover, that different authorities have their own schedules for collecting data makes things worse because what has been updated might be referring to very differently constituted traffic conditions.

3.1.2. Communication overhead and client mobility in autonomous driving

Fl's also being used for autonomous driving. It runs into big troubles about talking plus moving around [2]. In training round, it requires both direction transmitting the params so there is heavy use of bandwidth for uplink band which causes problems in vehicle spectrum allotment [9]. The overhead scales linearly for every single model as well as for any number of fleets which might participate in that fleet, thus it does not make sense to employ the use of a neural net. At the same time, large vehicle mobility causes participation intimacy as they traverse coverage boundaries or experience signal attenuation. These types of fleeting trends are also no good for any kind of convergence's promises; they only make those non-IIDs that show up all over.

3.1.3. Security vulnerabilities and privacy leakage in V2X systems

FL- based V2X framework still have some security issues because people can see what other cars are doing when exchanging their gradients. The adverse parties will insert purposeful model updates to ruin the main general model and cause an adverse effect on the collision avoidance results. Model poisoning attacks have special worries when the bad cars take part in making things better, needing hardly any check. Meanwhile, gradient inversion and membership inference attacks intend to obtain the secret private info from what should be private data updates. Vanilla FL implementations on vehicles make drivers easy for people to see their location when they do not send too much information.

3.2. Future research directions

3.2.1. Heterogeneity-aware personalized federated learning

To address the non-IID data problem in traffic forecasting, more sophisticated heterogeneity-aware personalized federated learning structures are required [3]. Construction of Federation Hierarchies: If it looks like nearby, they can aggregate near each other into intermediate aggregations. Use these aggregations to figure out what things look like regionally before they're aggregated back up to the full picture [6]. Personalized is like the FL but it does share a little bit and has more personal freedom because while every client gets to play with their very own mini version of the model, each one still gets to keep its own private sub-model. Mixture-of-experts type extensions which dynamically route based on current traffic conditions seem like they could be a good way to have both global knowledge and local flexibility.

3.2.2. Gradient compression and adaptive client selection

And so, it is necessary to improve on the gradient compression, as well as the choice of what clients are adaptive, in order to avoid these connection and mobility issues with an autonomous vehicle: [9]. By applying gradient compression methods, the payload goes down significantly per round but it's still got good enough for convergence [2]. And take a complementary approach to select the clients by picking them both because their channel is good and the amount of dwell time they have left inside the area of coverage will result in being even better off. And when it comes to hierarchical FL

DOI: 10.54254/2755-2721/2026.33761

structure, assume that the roadside unit does intermediate gathering for those updates brought forward by the nearby car and truck; then in face of such fleeting actions on behalf of client.

3.2.3. Integrated Security-by-Design and blockchain-augmented trust

In order to fill in any security gaps with the FL centric V2X system, it is required to think of a Security by Design [10]. The differential privacy framework that adds carefully tuned up noise also gives out some formal private assurances even though I've been taking care making sure it has adequate safe enough to be used application [2]. And with that it is possible to use some other secure Aggregation Protocols via homomorphic Encryptions not to worry about any changes on it being transferred to the curios Server. When building a blockchain based trust management system, it can help learn what kind of reputation is worth verifying and eliminating those perpetually bad guys. When putting all these systems into the future digital twin and 6G environment it would help to increase its scalability and speed.

4. Conclusion

This review provides an in-depth analysis of the application of FL in Intelligent Transportation Systems, stressing its potential benefits in fostering collaborative learning while preserving privacy and reducing communication overheads. While centralized learning is less efficient for tasks like traffic forecasting and autonomous vehicle coordination as well as V2X communication, FL can offer a feasible approach without revealing the data. However, it faces many difficulties caused by the fact that there is non-IID data distribution in transportation systems and they are mobile with very strict time constraints. Though it might appear that current studies dealing with the issue of heterogeneity, mobility, security has shown some promising results but still more such kind of study are needed where they should be conducted in real life situation. As for future research, it is required to study personalized and hierarchical FL, more efficient methods of communication, etc. Combining FL with next technology like digital twinning, 6G network etc. will further drive the development of ITS.

References

[1] Chong, Y.-W., et al. (2025). Federated Learning for Intelligent Transportation Systems: Use Cases, Open Challenges, and Opportunities. IEEE Intelligent Transportation Systems Magazine.
[2] Patil, R., et al. (2025). Federated Learning in Vehicular Networks for Resource Management, Privacy, and Security: Challenges and Future Directions. IEEE Access, 13, 190766–190785.
[3] Li, T., et al. (2020). Federated Learning: Challenges, Methods, and Future Directions. IEEE Signal Processing Magazine, 37(3), 50–60.
[4] Kairouz, P., et al. (2021). Advances and Open Problems in Federated Learning. Foundations and Trends in Machine Learning, 14(1–2), 1–210.
[5] McMahan, B., et al. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS).
[6] Sakhipov, A., et al. (2026). Deep Heterogeneity Learning for Cross-City Transit Forecasting: A Differentially Private Federated Framework with Mixture-of-Experts and Seasonal Decomposition. Frontiers in Future Transportation.
[7] Chellapandi, V. P., et al. (2024). Federated Learning for Connected and Automated Vehicles: A Survey of Existing Approaches and Challenges. IEEE Transactions on Intelligent Vehicles, 9(1), 119–137.
[8] Shi, Y., et al. (2025). Mobility-Aware Dynamic Resource Allocation for FL in VEC. Proc. of 2025 International Conference on MEC.
[9] Xing, L., et al. (2025). A survey of federated learning-based gradient compression for internet of vehicles. Engineering Applications of Artificial Intelligence, 159, 111662.

DOI: 10.54254/2755-2721/2026.33761

[10] Selvaraj, A. K., et al. (2025). Federated Learning and Digital Twin-Enabled Distributed Intelligence Framework for 6G Autonomous Transport Systems. IEEE Journals & Magazine.