Federated Learning Framework
Coping with Hierarchical Heterogeneity in Cooperative ITS

Rui Song[1,2] , Liguo Zhou[2] , Venkatnarayanan Lakshminarasimhan[2] , Andreas Festag[1,3] and Alois Knoll[2]

Abstract— Deep learning is a key approach for the environment perception function of Cooperative Intelligent Transportation Systems (C-ITS) with autonomous vehicles and smart traffic infrastructure. The performance of the object recognition and detection strongly depends on the data volume in the model training, which is usually not sufficient due to the limited data collected by a typically small fleet of test vehicles. In today's C-ITS, smart traffic participants are capable of timely generating and transmitting a large amount of data. However, these data can not be used for model training directly due to privacy constraints. In this paper, we introduce a federated learning framework coping with Hierarchical Heterogeneity (H^2-Fed), which can notably enhance the conventional pretrained deep learning model. The framework exploits data from connected public traffic agents in vehicular networks without affecting user data privacy. By coordinating existing traffic infrastructure, including roadside units and road traffic clouds, the model parameters are efficiently disseminated by vehicular communications and hierarchically aggregated. Considering the individual heterogeneity of data distribution, computational and communication capabilities across traffic agents and roadside units, we employ a novel method that addresses the heterogeneity in different aggregation layers of the framework architecture, i.e., aggregation in layers of roadside units and cloud. The experimental results indicate that our method can well balance the learning accuracy and stability according to the knowledge of heterogeneity in current communication networks. Compared to other baseline approaches, the evaluation on a Non-IID MNIST dataset shows that our framework is more general and capable especially in application scenarios with low communication quality. Even when 90% of the agents are timely disconnected, the pre-trained deep learning model can still be forced to converge stably, and its accuracy can be enhanced from 68% to over 90% after convergence.

I. INTRODUCTION

Artificial intelligence, especially deep learning, is one of the most important technologies in today's *Cooperative Intelligent Transportation System* (C-ITS) development. Its application requires a huge amount of data collected on traffic roads. However, the deep learning models, i.e., *Deep Neural Network* (DNN) models, are usually trained by an insufficient amount of data because automotive manufacturers (Original Equipment Manufacturers, OEMs) train their DNN models for autonomous driving only with data from a very limited number of test vehicles. The resulting bias in the data distribution can lead to errors in autonomous driving functions and threaten traffic safety.

In most regions, the acquisition of raw data from public road traffic is challenged due to privacy constraints. Fortunately, federated learning [1] has been proposed as a distributed learning approach that does not affect the privacy of user data. It provides C-ITS a great opportunity to gain data from public road traffic and enhance the centralized pre-trained DNN models. *Vehicle-to-Everything* (V2X) communication [2] enables the data exchange among vehicles, roadside infrastructure, backends and cloud systems, and facilitates federated learning in C-ITS. Based on abundant *Roadside Units* (RSUs), digital twins [3] can provide accurate labels for the learning process.

As many other federated learning application scenarios, C-ITS faces heterogeneity problems. These problems exist due to the lack of full access and control to traffic agents. In C-ITS, heterogeneity derives mainly from *(i)* various available computational capabilities, *(ii)* timely varied communication quality, and *(iii)* individual local training datasets, known as not-independent and identically distributed (Non-IID) data. It can lead to inconsistency in the training process across the traffic agents and cause low convergence speed and instability in federated learning in C-ITS.

To guarantee the success of federated learning and enhance the pre-trained DNN models in C-ITS, we develop a novel federated learning framework as shown in Fig. 1. The framework allows all connected vehicles to collaboratively train a common DNN model via hybrid vehicular communications, i.e., *Vehicle-to-Infrastructure* (V2I) and *Vehicle-to-Network* (V2N). Through sharing both global and roadside federated learning models in our framework, the effect of hierarchical heterogeneity on the learning process can be beneficially reduced, and hence the performance of pre-trained DNN models is powered by data acquired from public traffic. The key contributions of the paper are:

- Formulate the problem with Non-IID data across RSUs, individual communication and computation capabilities at the connected traffic agents, and analyze the impact on the federated learning performance in vehicular networks,
- Address individual heterogeneity in different layers of the hierarchical federated learning system and develop

[1]Rui Song and Andreas Festag are with Fraunhofer Institute for Transportation and Infrastructure Systems IVI, Ingolstadt, Germany, e-mail: `{rui.song, andreas.festag}@ivi.fraunhofer.de.`

[2]Rui Song, Liguo Zhou, Venkatnarayanan Lakshminarasimhan and Alois Knoll are with Technical University of Munich, Robotics, Artificial Intelligence and Real-Time Systems, Garching, Germany, e-mail: `{rui.song, liguo.zhou, venkat.lakshmi}@tum.de, knoll@in.tum.de.`

[3]Andreas Festag is with Technische Hochschule Ingolstadt, CARISSMA Institute for Electric, COnnected, and Secure Mobility (C-ECOS), Ingolstadt, Germany, e-mail: `andreas.festag@carissma.eu.`
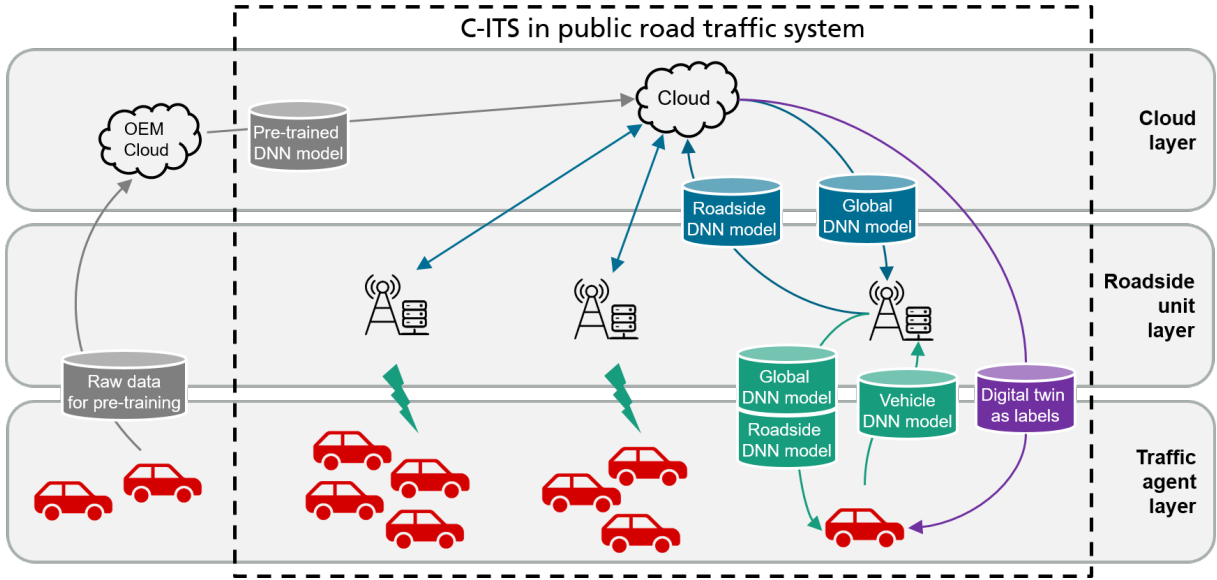
Fig. 1: Proposed framework for federated learning in a Cooperative Intelligent Transportation System (C-ITS): Blue, green and purple arrows indicate I2N, V2I and N2V communication, respectively. The gray line represents OEM-internal communication.

multiple proximal terms in the learning algorithm,

- Propose a general and empirical federated forward training framework for enhancing deep learning tasks in C-ITS by means of V2X networks without affecting the data privacy,
- Evaluate our methods by federated training a DNN model on benchmark datasets considering the Non-IID data properties and heterogeneous communication qualities in C-ITS,
- Provide comprehensive experiment results and in-depth analysis of parameters in our framework with heterogeneously connected traffic agents in public road traffic,
- Compare our framework with other federated learning approaches and address the trade-off between stability and accuracy in federated learning application scenarios. The software implementation of our framework is publicly available as open source at GitHub.[1]

The remainder of this paper is structured as follows. Section II summarizes related work in federated learning and its application in C-ITS. The problem with individual hierarchical heterogeneity in vehicular networks and federated learning is formulated in Section III. The federated learning methods are introduced in Section IV and the proposed H^2-Fed framework is presented in Section V. The system is comprehensively evaluated in Section VI considering heterogeneity in individual layers of the framework architecture with respect to C-ITS scenarios. Section VII provides a summary and concludes.

II. RELATED WORK

Federated learning is a particular distributed learning framework, which was proposed in [1] as *FedAvg* and can

be employed for training models without accessing the user's data in massively distributed systems. To achieve joint learning in a federated fashion, especially wireless communication networks for sharing the locally training results across infrastructure are required. Specifically, deployment of C-V2X networks [4] according to corresponding standards [5] facilities federated learning in C-ITS, where many traffic services and applications need to be developed based on DNN models, e.g., ResNet [6], V2X-ViT [7], etc., for traffic object recognition and detection functions. In this context, [8]–[10] indicate the great potential benefits of federated learning in C-ITS.

As with many other federated learning application scenarios, heterogeneity problems exist while implementing federated as well. Research in [10]–[12] and others address heterogeneity problems and provide general federated learning algorithms (e.g., *FedProx*) by adjusting the training objective to reduce the influence of straggling devices. However, depending on specific federated learning applications and associated communication infrastructure, the heterogeneity can be caused by various reasons and needs to be optimized individually.

Moreover, to utilize the hierarchical infrastructure for federated learning in C-ITS, multi-layer aggregation can be used. [13] provides a federated learning framework in hierarchical clustering steps and [14] proposes hierarchical federated learning (*HierFAVG*) in a client-edge-cloud system, which can be regarded as similar to C-ITS considered in this paper. However, heterogeneity is not considered. [15] focuses on federated learning in cellular networks but do not take the possible direct communication among vehicles and with RSUs via sidelink into account. Sidelink is regarded as a typical communication method in V2X networks and

---

[1] https://github.com/rruisong/H2-Fed

TABLE I: Overview of heterogeneity metrics while hierarchically applying federated learning system in C-ITS

Heterogeneity metric	Notation	Main cause	Effect on Federated Learning
Local Aggregation Round	LAR	Diverse Quality of Service (QoS)	The pre-aggregation [2] round will be done in each RSU
Connection Success Ratio	CSR	Varied communication quality	The number of agents connected for federated learning
Stable Connection Duration	SCD	Varied communication quality	Stable participating time of one agent, once it is connected
Full-task Success Ratio [3]	FSR	Individual available computation resource	Finished training epochs in each agent

[1] Aggregation in an RSU before the global aggregation in the cloud.
[2] FSR is not the focus in this paper and has a similar effect as CSR.

is considerably faster than up- and downlink [16], [17]. Meanwhile, due to the various V2X messaging services, especially for traffic safety, as evaluated in [18], the priority of data transmission for federated learning in V2X networks can be low compared to safety-related messages, which causes additional time-variant heterogeneity problems across traffic agents.

In this paper, we develop a hierarchical federated learning framework specifically for hybrid V2X networks in C-ITS with consideration of heterogeneity problems. Differentiating from other existing research work, we analyze the heterogeneity caused by V2X communication and available computational capabilities in traffic agents. By employing multiple proximal terms in the federated learning algorithm, the individual heterogeneity in different layers of the framework can be coped for the successes of model training.

III. Heterogeneity Problems

Heterogeneity in datasets. We assume that RSUs are equipped with homogeneous computational capabilities and benefit from a stable communication quality; in another word, the traffic cloud reliably receives the aggregated models from all RSUs. However, the models generated in each RSU are based on a dataset that is notably different from other RSUs due to the individual position of each RSU. For instance, the model in an RSU equipped at a highway is trained with data in highway scenarios, where the involved dynamic objects in the datasets are typically passenger cars or trucks that move fast in simple road typologies. In contrast, there are various types of dynamic objects with low speed in complex road networks in the datasets, when an RSU is located in cities. In addition, even urban scenarios in various locations can be marked distinct, e.g., at crossings, roundabouts, etc. Therefore, we believe that the models generated across RSUs are based on Non-IID datasets. Besides, the diverse average traffic flows cause the unbalanced agent number at RSUs, which can finally impact the learning results while aggregating the models at the traffic cloud.

Heterogeneity in communication. The direct communication between traffic agents and RSUs achieves local aggregation (or pre-aggregation) in federated learning at each RSU, where traffic agents and the RSU represent the clients and server in non-hierarchical federated learning. Then, before one global aggregation, each RSU can implement several *Local Aggregation Rounds* (*LAR*), which accelerates the learning process by means of pre-aggregation between two

global aggregation steps. Through sidelink communication in Cellular-V2X, the models at the RSU can be pre-aggregated up to 50 times, if we aggregate models from RSUs every 1 s globally in the cloud. In fact, due to the various other traffic services in V2X communication, we need to set the *LAR* in a proper range with respect to the current *Quality of Service* (QoS) at each RSU. Additionally, according to the current data traffic in the communication channel, not all traffic agents can continuously upload their models to the RSUs. Some traffic agents might need to allocate the communication resource for other tasks with higher priority. The metric *Connection Success Ratio* (*CSR*) indicates the current connection situation specifically for federated learning at each RSU j:

$$CSR_j = \frac{N_{j,connected\,agent}}{N_{j,participant}}, (CSR \in [0,1]), \quad (1)$$

where $N_{j,participant}$ is the number of all connected agents at RSU j and $N_{j,connected\,agent}$ is the number of successfully connected agents, which stably upload the models in a predefined duration, specified as *Stable Connection Duration* (SCD) in seconds.

Heterogeneity in computation. Apart from communication, the available computation resources can also vary in each traffic agent. This leads to the fact that not all traffic agents finish the expected local training epochs E. Normally, federated learning requires each agent to finish local training by the same E. However, the different computational capabilities and currently occupied computational resources cannot always satisfy the request on E within a specific aggregation frequency, which causes further heterogeneity. *Full-task Success Ratio* (FSR) indicates the ratio of traffic agents that can finish the requested E in time. Similarly, FSR can also be individual in each FSR. Note that if an agent cannot even finish one epoch, its results will be discarded, which has the same effect as a failed connection. Thus, in this paper, we will address heterogeneity especially considering varied *CSR*. All heterogeneity metrics are summarized in Tab. I

IV. Federated Learning Methods

Federated learning can enhance the pre-trained DNN model for intelligent traffic applications since it can further improve the model with broad data collected by public traffic without affecting user data privacy. Considering the V2X communication networks for federated learning, the DNN models from agents can be uploaded via direct communication. Thanks to the high-defined sensors at intelligent

infrastructure, the highly accurate perception are achieved and can be set for 3D digital twin services, which can be deduced into sensor raw data from a global coordinate system and taken as labels for supervised federated learning implementation.

As a baseline method, we first employ FedAvg [1] without dropping any clients in the hierarchical federated learning framework. For an RSU k, the aim is to minimize:

$$\min_w f_k(w) = \sum_{i \in \mathscr{P}_k^t} \frac{n_{i,k}}{n_k} F_{i,k}(w), \qquad (2)$$

where $\mathscr{P}_k^t$ is a partition of traffic agents selected at time point t at RSU k and n_k is the total number of data points in all traffic agents selected by the RSU federated learning orchestrator. The agent i trains with $n_{i,k}$ data points and $F_{i,k}(w)$ is averaged optimized objective in the agent i with respect to non-convex DNN optimization. If K RSU train a DNN model jointly, we can extend the objective form in (2) as:

$$\min_w h(w) = \sum_{k=1}^{K} \sum_{i \in \mathscr{P}_k^t} \frac{n_{i,k}}{n_k} F_{i,k}(w) \qquad (3)$$

To address the individual heterogeneity in each layer of the federated learning system, as analyzed in Sec. III, we employ the γ_k^t-inexact solution [11] with multiple proximal terms in the loss function of the hierarchical federated learning. The aim of the traffic agent i connecting RSU k is then to minimize:

$$\min_{w_{i,k}} h_k(\cdot) = F_k(w_{i,k}) + \sum_{l=1}^{L} \frac{\mu_{k,l}}{2} ||w_{i,k} - w_l||^2 \qquad (4)$$

where L is the number of aggregation layers. We note $\mathscr{M} = \{\mu_{k,l} | k = 1, 2, ..., K, l = 1, 2, ..., L\}$ as a set of main parameters in 4. In vehicular networks, $L = 2$. When $l = 1$, μ_1 indicates the heterogeneity of data from traffic agents and w_1 is the current weight of DNN model in the connected RSU. For $l = 2$, μ_2 indicates the heterogeneity of models across RSUs and w_2 is the current model weight in the traffic cloud. To simplify the notation, we define

$$\begin{aligned} & \textit{for an RSU } k\text{: } w_k \overset{\text{def}}{=\!=} w_{l=1}, \\ & \textit{for the traffic cloud: } w \overset{\text{def}}{=\!=} w_{l=2}. \end{aligned} \qquad (5)$$

We can rewrite (4) as a double-layer aggregation system in accordance with the application of hierarchical federated learning in vehicular networks as

$$\min_{w_{i,k}} h_k(\cdot) = F_k(w_{i,k}) + \frac{\mu_{k,1}}{2} ||w_{i,k} - w_k||^2 + \frac{\mu_{k,2}}{2} ||w_{i,k} - w||^2, \qquad (6)$$

where for each RSU k, the set of main parameters is $\mathscr{M}_k = \{\mu_{k,l} | l = 1, 2\}$.

The method defined in (6) reduces the effect of individual heterogeneity of data in different layers by setting the parameters $\mu_{k,1}$ and $\mu_{k,2}$ in $\mathscr{M}_k$, which can stabilize federated optimization through adding the additional trailers in the loss.

Though the multiple proximal terms compared with a single term might slow down the convergence of training [11], it provides more tunable parameters to separately consider the heterogeneity among RSUs and across agents in one RSU, which will be evaluated in Section VI.

V. SYSTEM FRAMEWORK

We divide the C-ITS into three layers: traffic agent layer, roadside unit layer, and cloud layer (Fig. 1). We first consider a pre-trained process, where raw data are collected by automotive OEMs internal test vehicles the in traffic agent layer. A DNN model is trained in a centralized fashion based on the data managed in the cloud layer. Since the amount of the collected data is not sufficient for training, the pre-trained DNN model is sent to a public traffic cloud, which can orchestrate federated learning for model enhancements in public C-ITS.

The cloud shares the pre-trained DNN model with all connected RSUs in the roadside unit layer. This pre-trained DNN model is taken as the initial global and roadside FL model and forwarded to the traffic agents by the RSUs. Once the traffic agents in the traffic agent layer receive the duplicated model via V2I communication, the training process is run based on the cached datasets. The data labeling can be achieved by digital twin services via V2N communication [3].

According to the specific federated learning orchestration, the traffic agents train the DNN model several epochs and share it back to their connected RSUs. Then RSUs aggregate the models from the selected partition of agents and update the roadside FL model. Before uploading the model to the cloud, the local aggregation in the roadside unit layer can be implemented a few times, as the model exchange between the roadside unit and traffic agent layers using sidelink communication is considered to be faster than up- and downlink. The roadside unit layer can be regarded as the first aggregation layer. After a necessary pre-aggregation, the roadside FL models from all RSUs are sent to the cloud via I2N communication and aggregated globally. The cloud layer is then the second aggregation layer. The new global FL model is sent back to all RSUs, where the old roadside and global FL models in the roadside unit layer are replaced.

To address the heterogeneity problems in C-ITS as analyzed in the previous Section III, we employ the method defined in (6) for federated optimization in hierarchical systems and propose a *Federated Learning Framework Coping with Hierarchical Heterogeneity* (H^2-*Fed*). The pseudo-code for traffic agents, RSUs and the traffic cloud are given in Algorithm 1, 2 and 3, respectively. By designing parameters in the traffic agent and RSU layers, our H^2-*Fed* system can be deployed as a general framework for other federated learning methods with respect to (4) : *(i)* $\mu_{k,l} = 0$ and $L = 1$ correspond to the idea of *FedAvg* [1]; *(ii)* $\mu_{k,l} > 0$ and $L = 1$ are equivalent to *FedProx* [11]; *(iii)* $\mu_{k,l} = 0$ and $L > 1$ match *HierFAVG* [14]. Furthermore, the trade-off between stability and accuracy can be tuned by designing all parameters for

Algorithm 1 Traffic agent i connecting RSU k in H^2-Fed

Input: η, $\mathcal{M}_{i,k}$, E, digital twin data
Output: $w_{i,k}$
  1: receive w_k and w from RSU k
  2: generate labels from digital twin data
  3: **for** each epoch $\tau = 1,...,E$ **do**
  4:     $w_{i,k} \leftarrow \eta \nabla \mathcal{L}^*_{\mathcal{M}_{i,k}}(w_{i,k}; w_k; w)$ based on Equation (6)
  5: **end for**
  6: send $w_{i,k}$ to RSU k

Algorithm 2 RSU k in H^2-Fed

Input: *LAR*
Output: $w_{i,k}$
  1: **for** each local round **do**
  2:     download w from cloud
  3:     generate $\mathscr{P}^t_k$ based on orchestration
  4:     **for** each clients $i \in \mathscr{P}^t_k$ **do**
  5:         send w and w_k to client i
  6:         receive $w_{i,k}$ from client i
  7:     **end for**
  8:     $w_k \leftarrow \sum_{i \in \mathscr{P}^t_k} \frac{n_{i,k}}{n_k} w_k$
  9: **end for**
10: upload w_k to cloud

Algorithm 3 cloud in H^2-Fed

Input: initial w from a private cloud
Output: w
  1: **for** each global round **do**
  2:     **for** each RSU k **do**
  3:         send w to RSU k
  4:         receive w_k from RSU k
  5:     **end for**
  6:     $w \leftarrow \sum_k \frac{n_k}{n} w_k$
  7: **end for**

---

traffic agents at different RSUs in H^2-*Fed*. This will be evaluated in Sec. VI.

VI. Experimental Results

We set up an experiment with a DNN model with a size of $130\,kB$ trained on the MNIST [19] dataset, which has 10 labels as road traffic scenarios and is processed as Non-IID data partitions. Similar to the methods in [20], those partitions are divided into 110 traffic agents, where the first 10 agents exclude a few labels and are used for pre-training as the addressed application scenarios. The pre-trained DNN model with 68% testing accuracy (ACC) is set as the initial model in all three layers as described in Section V. Each of the other 100 agents is able to train the DNN model locally and communicate to one RSU if the connection can successfully be established. As formulated in Section III, we vary *CSR* and *SCD*, and consider the heterogeneous communication quality in V2X networks.

To evaluate the enhancement of the pre-trained model specifically by various μ_1, we note the testing accuracy changes at time t compared to the pre-trained DNN model as $\Delta ACC = ACC_t - ACC_{pre}$, and define *ACC Enhancement Degree* (AED) as

$$AED = (\Delta ACC^{\mu_1>0} - \Delta ACC^{\mu_1=0})/\Delta ACC^{\mu_1=0}. \quad (7)$$

The metric AED indicates the testing accuracy enhancement by increasing μ_1 ($\mu_1 > 0$) in the set of parameters $\mathcal{M}$ (Algorithm 1) for agents connecting to the RSUs, which can only provide the data with very limited label types. As the first column of Fig. 2 shows, when we set $\mu_1 = 0.001$ the

communication quality satisfied is (CSR=100%), the AED is overall positive (under 0.5%) with $\mu_2 = 0$ or $\mu_2 = 0.001$ after convergence (from around 10 seconds). By increasing μ_1, the AED is clearly raised. Each row in Fig. 2 shows the AED during federated learning with various μ_1 and fixed μ_2 with different qualities of communication. The results indicate that the AED can be obviously increased when the communication quality is rather bad (low CSR), which can oft happen in V2X networks in C-ITS due to the low priority of federated learning-related communication messages. When only 20% traffic agents are successfully connected to the RSUs, $\mu_1 = 0.001$ can enhance the testing ACC of the federated trained DNN model up to 20% compared to the model trained without setting μ_1.

Next, we evaluate the effect of positive μ_2 on federated learning. We notice in Fig. 2 that a positive value of μ_2 can reduce the AED to some extent. However, for most federated learning application scenarios, a failure in the learning process is caused by unreliable connections between clients and the server. Thus, learning stability can be more significant for federated learning. The first row of Fig. 3 shows that the concussion in the learning process due to the low communication quality can be well stabilized through a large μ_2. We also present the *Mean Squared Error* (MSE) of the testing accuracy to centralized learning results in the second row of Fig. 3, where the raw data from all traffic agents are collected in the cloud layer and used for centralized model training. When $\mu_2 = 0.005$, the jitters of the testing accuracy curve are extraordinarily coped, where the performance of federated learning is almost the same as learning with $CSR = 90\%$.

Finally, we compare our H^2-*Fed* with other two baseline methods – *FedProx* [11], *HierFAVG* [14] – considering Non-IID datasets and heterogeneous communication quality with $CSR = 10\%$ and $SCD = 1\,s$. We demonstrate all those federated DNN model enhancement approaches in two different empirical C-ITS application scenarios with 100 traffic agents and 10 RSUs. In *Scenario I*, the datasets across RSUs are Non-IID, while the datasets in all traffic agents under one RSU have the same distribution. In contrast, the datasets across traffic agents at one RSU are Non-IID in *Scenario II*, where each RSUs has the same data distribution.

Fig. 4a shows our framework H^2-*Fed* can enhance the pre-trained model stably from beginning to convergence, while
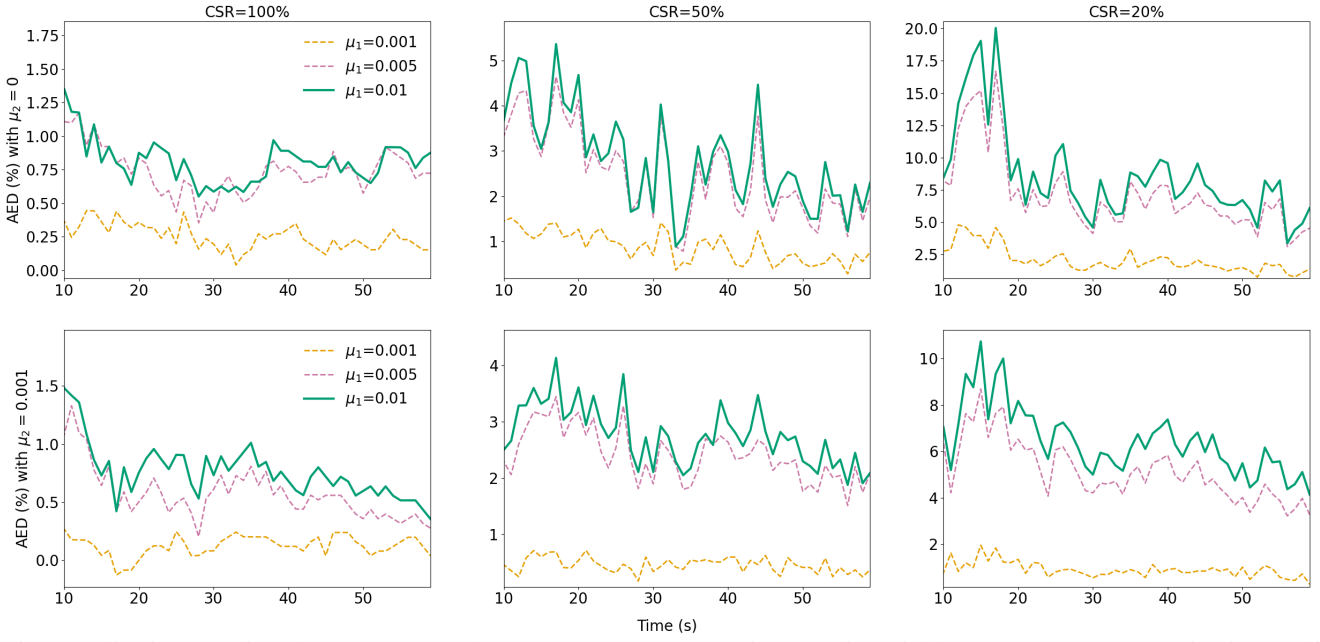
Fig. 2: Adjusting μ_1 raises the *Accuracy Enhancement Degree* (AED) in scenarios with heterogeneous communication quality.
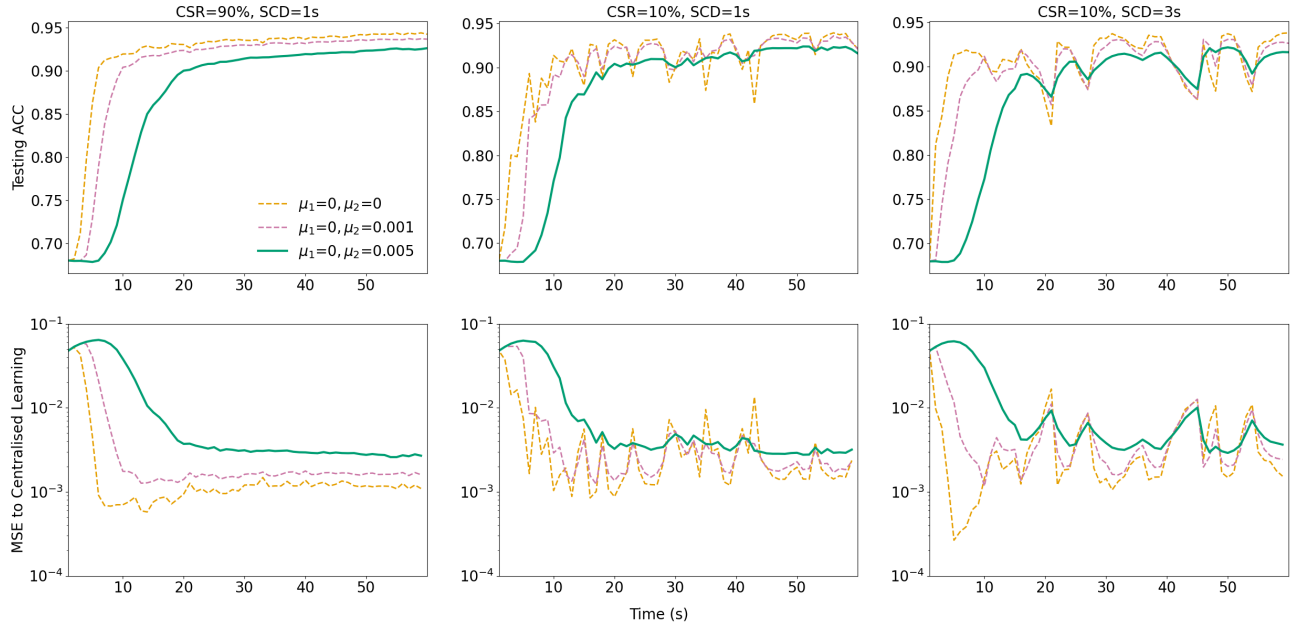

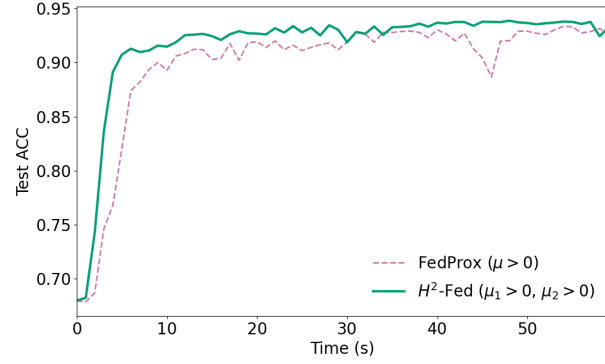Fig. 3: Increasing μ_2 stabilizes federated learning in scenarios with heterogeneous communication quality.

the jitters of *HierFAVG* curve are more visibly affected by the bad communication situation. Furthermore, as shown in Fig. 4a, our framework H^2-*Fed* in *Scenario II* outperforms *FedProx* remarkably, as all models at each RSU are pre-aggregated multiple times, which accelerates the convergence of federated learning. The comparison results indicate that our framework can address hierarchical heterogeneity by adapting more tunable parameters which can result in the success of federated learning in C-ITS.

VII. CONCLUSION

In this work, we highlight one of the most important bottlenecks in AI model training for C-ITS – insufficient data, which can be solved by federated learning in V2X communication networks. We address the individual hetero-geneity problems in different layers of C-ITS and proposed a novel method, as well as an associated general framework named H^2-*Fed* for enhancing the pre-trained DNN model. Our empirical evaluation demonstrates that the framework can guarantee the success of federated learning even with large-scale disconnected traffic agents in C-ITS. The com-parison with other baseline methods indicates the flexibility
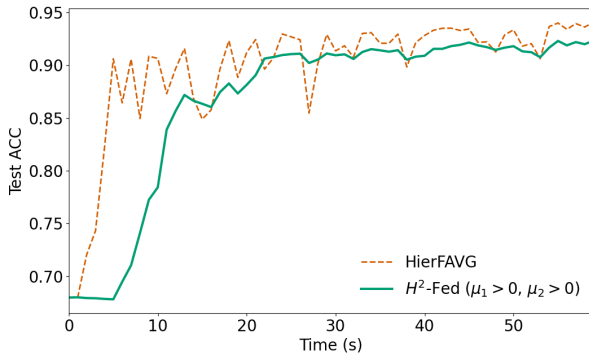
(a) Scenario I: Non-IID datasets only in RSU layer   (b) Scenario II: Non-IID datasets only in traffic agent layer

Fig. 4: Comparison between our framework H^2-*Fed* and other baseline methods for heterogeneous communication quality. Note that the baseline methods can also be seen as our framework with dedicated parameter combinations.

and generality of H^2-*Fed* that provides opportunities to well balance the trade-off between accuracy and stability in different application scenarios. Besides, we believe that in our framework the robustness of federated learning can be further improved by dynamic parameter settings, which will be validated in future simulations. Moreover, time-variant latency due to road traffic topology and efficient encoding for dissemination of large DNN models remain.

REFERENCES

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, vol. 54, 2017, pp. 1273–1282, arXiv:1602.05629 [cs.LG].

[2] A. Festag, "Standards for vehicular communication – from IEEE 802.11p to 5G," *Springer Elektrotech. Inftech.*, no. 132, p. 409–416, 2016, doi:10.1007/s00502-015-0343-0.

[3] C. Creß and A. C. Knoll, "Intelligent Transportation Systems using external infrastructure: A literature survey," 2021. [Online]. Available: https://arxiv.org/abs/2112.05615

[4] V. Lakshminarasimhan and A. Knoll, "C-V2X resource deployment architecture based on moving network convoys," in *IEEE VTC2020-Spring*, 2020, p. 6, doi:10.1109/VTC2020-Spring48590.2020.9128410.

[5] A. Festag, "Cooperative Intelligent Transport Systems standards in Europe," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 166–172, 2014, doi:10.1109/MCOM.2014.6979970.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778, doi:10.1109/CVPR.2016.90.

[7] R. Xu *et al.*, "V2X-ViT: Vehicle-to-everything cooperative perception with vision transformer," 2022, arXiv:2203.10638 [cs.CV].

[8] J. Posner, L. Tseng, M. Aloqaily, and Y. Jararweh, "Federated learning in vehicular networks: Opportunities and solutions," *IEEE Network*, vol. 35, no. 2, pp. 152–159, 2021, doi:10.1109/MNET.011.2000430.

[9] D. Ye, R. Yu, M. Pan, and Z. Han, "Federated learning in vehicular edge computing: A selective model aggregation approach," *IEEE Access*, vol. 8, pp. 23 920–23 935, 2020, doi:10.1109/ACCESS.2020.2968399.

[10] A. M. Elbir, B. Soner, and S. Coleri, "Federated learning in vehicular networks," 2020. [Online]. Available: https://arxiv.org/abs/2006.01412

[11] T. Li *et al.*, "Federated optimization in heterogeneous networks," Apr. 2020, arXiv1812.06127 [cs.LG].

[12] J. Wang, Q. Liu, H. Liang, G. Joshi, and H. V. Poor, "A novel framework for the analysis and design of heterogeneous federated learning," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5234–5249, 2021, doi:10.1109/TSP.2021.3106104.

[13] C. Briggs, Z. Fan, and P. Andras, "Federated learning with hierarchical clustering of local updates to improve training on non-IID data," in *International Joint Conference on Neural Networks (IJCNN)*, 2020, p. 9, doi:10.1109/IJCNN48605.2020.9207469.

[14] L. Liu, J. Zhang, S. Song, and K. B. Letaief, "Client-edge-cloud hierarchical federated learning," in *IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6, doi:10.1109/ICC40277.2020.9148862.

[15] M. S. H. Abad, E. Ozfatura, D. Gunduz, and O. Ercetin, "Hierarchical federated learning across heterogeneous cellular networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 8866–8870, doi:10.1109/ICASSP40776.2020.9054634.

[16] A. Hegde and A. Festag, "Mode switching performance in Cellular-V2X," in *IEEE Vehicular Networking Conference (VNC)*, 2020, p. 8, doi:10.1109/VNC51378.2020.9318394.

[17] M. H. C. Garcia *et al.*, "A tutorial on 5G NR V2X communications," *IEEE Communications Surveys Tutorials*, vol. 23, no. 3, pp. 1972–2026, 2021, doi:10.1109/COMST.2021.3057017.

[18] S. Kühlmorgen *et al.*, "Evaluation of congestion-enabled forwarding with mixed data traffic in vehicular communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 233–247, 2020, doi:10.1109/TITS.2018.2890619.

[19] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi:10.1109/5.726791.

[20] S. Caldas *et al.*, "Leaf: A benchmark for federated settings," Dec. 2019, arXiv:1812.01097 [cs.LG].