

“© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Dynamic Federated Learning-Based Economic Framework for Internet-of-Vehicles

Yuris Mulya Saputra, Dinh Thai Hoang, Diep N. Nguyen, Le-Nam Tran, Shimin Gong, and Eryk Dutkiewicz

Abstract—Federated learning (FL) can empower Internet-of-Vehicles (IoV) networks by leveraging smart vehicles (SVs) to participate in the learning process with minimum data exchanges and privacy disclosure. The collected data and learned knowledge can help the vehicular service provider (VSP) improve the global model accuracy, e.g., for road safety as well as better profits for both VSP and participating SVs. Nonetheless, there exist major challenges when implementing the FL in IoV networks, such as dynamic activities and diverse quality-of-information (QoI) from a large number of SVs, VSP's limited payment budget, and profit competition among SVs. In this paper, we propose a novel dynamic FL-based economic framework for an IoV network to address these challenges. Specifically, the VSP first implements an SV selection method to determine a set of the best SVs for the FL process according to the significance of their current locations and information history at each learning round. Then, each selected SV can collect on-road information and offer a payment contract to the VSP based on its collected QoI. For that, we develop a multi-principal one-agent contract-based policy to maximize the profits of the VSP and learning SVs under the VSP's limited payment budget and asymmetric information between the VSP and SVs. Through experimental results using real-world on-road datasets, we show that our framework can converge 57% faster (even with only 10% of active SVs in the network) and obtain much higher social welfare of the network (up to 27.2 times) compared with those of other baseline FL methods.

Index Terms—Federated learning, IoV, quality-of-information, contract theory, profit optimization, vehicular networks.

1 INTRODUCTION

According to the recent Allied Market Research's outlook [1], the global market value of the Internet-of-Vehicles (IoV) is expected to grow by more than 215% in 2024 due to significant demands of road safety from smart vehicles (SVs), e.g., autonomous cars and electric vehicles. Through the IoV, the vehicular service provider (VSP) can build on-road services to enhance the driving safety for the SV users. In particular, the VSP can first obtain on-road data from active SVs e.g., road conditions, location, and driving activities. Using the centralized learning process, the VSP can then generate the meaningful on-road information with high quality [2], and share the updated on-road information to requesting customers (e.g., SVs, public transportations, local governments, and mobile users) in its coverage areas. In fact, the aforementioned mechanism has been widely adopted by several commercial on-road applications, e.g., the Placemeter (www.placemeter.com), Google Maps (www.google.com/maps), and Waze (www.waze.com). However, sharing raw on-road data to the VSP in a centralized manner may face the huge computation and storage

costs for the VSP, privacy disclosure for the SV users, and severe network congestion in the IoV (when a huge amount of data is transferred over the network). Therefore, federated learning (FL) as a highly-effective distributed machine learning approach has a great potential to address these issues [3]. Specifically, instead of sharing raw on-road data, each SV in the FL-based IoV network can train its local on-road data to generate the local trained model. This local model is then collected by the VSP to update the global on-road model, aiming at enhancing the on-road model accuracy. By doing so, we can mitigate the inherent problems that occur in the centralized learning systems.

Although the applications of FL-based IoV networks have been studied in a few works, e.g., [4], [5], [6], the conventional FL model is inefficient or even impractical to be implemented in real-world IoV networks due to several reasons. First, the behaviors of SVs in the networks are very dynamic in practice, e.g., some SVs move frequently, while some other SVs are occasionally disconnected from the network (e.g., due to unreliable communication channels, multiple handovers, and Internet unavailability at particular periods). Furthermore, training local datasets and collecting local trained models from all the SVs in the networks for the learning process are costly and impractical due to huge communication overheads, especially with a large number of moving SVs. Second, the quality-of-information (QoI), that includes on-road data in various locations at different times [7], obtained by the SVs is remarkably diverse. That makes the concurrent learning process from all SVs even worse when some SVs have low QoI. Third, due to the VSP's limited payment budget and the inherent profit competition among SVs, how to optimally incentivize the "right" SVs for their contributions to the FL process while maximizing the profit of VSP still remains as an open issue, especially under

- Y. M. Saputra is with the School of Electrical and Data Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia (e-mail: yurismulya.saputra@student.uts.edu.au) and Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.
- D. T. Hoang, D. N. Nguyen, and E. Dutkiewicz are with the School of Electrical and Data Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia (e-mail: {hoang.dinh, diep.nguyen, [eryk.dutkiewicz](mailto:eryk.dutkiewicz@uts.edu.au)}@uts.edu.au).
- L. N. Tran is with the School of Electrical and Electronic Engineering, University College Dublin, Belfield, Dublin 4, Ireland (e-mail: nam.tran@ucd.ie).
- S. Gong is with the Sun Yat-sen University, China (e-mail: gong-shm5@mail.sysu.edu.cn).

the information asymmetry between the VSP and SVs.

To address the first challenge due to dynamic activities of a large number of SVs, solutions to select/schedule only a set of FL local learners, e.g., smartphones, IoT devices, and vehicles, in the wireless networks can be adopted. For example, in [8], the authors introduce random scheduling, round-robin, and proportional fair methods to choose FL learners. Other random selection methods using asynchronous FL and heterogeneous networks are also discussed in [9] and [10], respectively. However, these works assume that the selected FL learners can implement local learning process equally without considering their QoI (the second aforementioned challenge). In fact, the QoI obtained by SVs in the IoV networks varies dramatically [7], e.g., some SVs are not moving for a certain period, and thus they may provide low QoI. Consequently, the low QoI can degrade the prediction accuracy or cause FL performance instability (when the low-quality trained local models are aggregated together with other local models during the learning process) [11], leading to a lower profit for the VSP. Furthermore, all the above works do not consider any incentive mechanism to compensate the selected FL learners for the learning process, and thus discourage the FL learners to participate in and contribute high QoI for the next FL processes [3]. Note that a “right” incentive mechanism can also in turn help to “recruit” the “right” SVs. Designing such an incentive model is challenging, especially under the limited budget of the VSP, the competition among the SVs, as well as the information asymmetry between the VSP and selected SVs.

In this paper, we aim to address the aforementioned challenges through proposing a novel dynamic FL-based economic framework that selectively engages a subset of best SVs for the FL process at each learning round. Specifically, the VSP can first select active SVs whose current locations are within significant areas. These significant areas are determined based on the Average Annual Daily Flow (AADF), i.e., the average number of vehicles passing through predefined roads in a particular area on one day [12]. From the set of active SVs based on the location significance, the VSP can further choose the best SVs based on their information significance (or QoI) history, e.g., due to the VSP’s limited payment budget for the learning SVs [13]. Through combining both location and information significance, the VSP can obtain useful on-road information [14] and reliable trained model updates [3] from the selected SVs, aiming at enhancing the learning quality of the FL. From this SV selection, the VSP can provide an incentive mechanism for the best SVs which share their current information significance for the FL process. Nonetheless, due to the the VSP’s limited payment budget and the existence of multiple learning SVs at each round, each selected SV may compete with other selected SVs in obtaining the corresponding payment as the incentive of its local training process. To address this issue, the authors in [15], [16], [17] propose a non-collaborative Stackelberg game model where the VSP and mobile devices act as a leader and followers, respectively. However, these methods are only applicable when the mobile devices know the full information of the VSP’s payment budget (referred to as *information symmetry*). In practice, the VSP usually keeps its payment budget confidential (referred to as *information asymmetry*) and the best SVs may not want to

follow controls by the VSP completely (due to the conflict of economic interests between them) [18]. Hence, the above approaches are ineffective to implement in our problem.

Given the above, we develop a *multi-principal one-agent (MPOA)* contract-based economic model [19], where the best SVs (as *principals*) non-collaboratively offer contract agreements (containing information significance and offered payment) to the VSP. The VSP (as *the agent*) is then in charge of optimizing the offered contracts. In this case, we formulate the contract model as a non-collaborative learning contract problem to maximize the profits of the VSP and the best SVs at each round under the VSP’s information-asymmetry and common constraints, i.e., the individual rationality (IR) and incentive compatibility (IC). The IR constraints ensure that the VSP always joins the FL process. Meanwhile, the IC constraints guarantee that the VSP always obtains the maximum profit when the best contract determined for the VSP is applied. To find the optimal contracts for the SVs, we first transform the contract optimization problem into an equivalent low-complexity problem. We then develop an iterative algorithm which can quickly find the equilibrium solution, i.e., optimal contracts, for the transformed problem. After that, the VSP and selected SVs can execute the FL algorithm to improve the global on-road model accuracy at each learning round. Moreover, at the end of each round, the VSP can update its net profit considering the current round’s global model accuracy and freshness (i.e., how up-to-date the global model in the FL process is). We then conduct experiments using real-world on-road datasets from all major roads of 190 areas/local districts and 1.5M traffic accidents in the United Kingdom (UK) between 2000 and 2016. The experiment results demonstrate that our framework can improve the social welfare of the network up to 27.2 times and speed up the learning convergence up to 57% compared with those of the baseline FL methods [8].

Although a few recent works, e.g., [20], [21], have studied the contract theory in FL, their proposed approaches have some limitations and are inapplicable in IoV networks. First, they only select the FL learners based on the QoI without considering location significance which is one of the most important factors in IoV networks. Second, both works assume that the FL learners’ data with certain quality have been collected prior to the FL learner selection. This is impractical in IoV networks where the SVs are dynamically moving, and thus the QoI may vary over time. Third, they are limited to the *one-principal multi-agent*-based contract policy, i.e., not accounting for the competition among FL learners. Moreover, the work in [20] implements the contract optimization only once and utilizes the same FL learners for the entire FL process until reaching a convergence. Meanwhile, the FL learner selection in [21] is determined by the vehicular users (not the central server). As such, there is a possibility that some candidate FL learners with high QoI may not be able to join in the FL process, and thus the overall prediction model accuracy can be degraded. To the end, the major contributions of our work are as follows:

- Propose a novel dynamic FL-based economic framework for the IoV network, aiming at selecting the best SVs and maximizing profits for the VSP as well as learning SVs to quickly achieve the converged

global prediction model for road safety.

- Design the dynamic SV selection method taking both location and QoI of SVs into account. As such, we can select the best quality SVs for the FL process, thereby improving the overall learning process and yielding more profits for the VSP.
- Develop an MPOA-based contract problem which can incorporate the learning payment competition among the selected SVs under the VSP's limited and unknown payment budget.
- Propose a transformation method to reduce the complexity of the original optimization problem, and then develop a light-weight iterative algorithm which can quickly find the optimal contracts for SVs.
- Theoretically analyze the convergence of the proposed dynamic FL and then provide detailed analysis to show the impacts of the global model accuracy and its freshness on the net profit of the VSP.
- Conduct extensive experiments to evaluate the proposed framework using two real-world on-road datasets in the UK. These results provide useful information to help the VSP in designing the effective FL with SV selection method in the IoV network.

The rest of this paper is organized as follows. Section 2 describes the proposed economic framework for the IoV. Section 3 introduces the proposed SV selection method, and Section 4 presents the proposed MPOA contract-based problem and solution. Section 5 describes the FL algorithm with the selected SVs. Moreover, the VSP's profit analysis based on the global model accuracy and freshness is provided in Section 6. The performance evaluation is given in Section 7, and then the conclusion is drawn in Section 8.

2 DYNAMIC FEDERATED LEARNING-BASED ECONOMIC FRAMEWORK

Consider an IoV network including one VSP and multiple active SVs who can participate in the dynamic FL process. Let $\mathcal{I} = \{1, \dots, i, \dots, I\}$ denote the set of active SVs in the IoV network. Periodically, each SV- i can monitor its current on-road status through its embedded sensor devices, e.g., the camera, weather, and global positioning system (GPS) sensors. From this status, the SV can pre-process and extract meaningful on-road information and store it in a log file at the SV's local storage. In particular, the information may include the SV's visited location history, date, time, light condition (e.g., dark or light), weather condition (e.g., windy, snowy, or rainy), and road surface condition (e.g., dry or wet). To collaborate with the VSP in the FL process, the interested SV users can first deploy an on-road service application developed by the VSP at their SVs' built-in Android or iOS platforms, e.g., Live Traffic NSW (www.livetraffic.com) and Live Traffic Info (www.highwaysengland.co.uk). Then, the SV users can enable the application's access permission to allow the VSP accessing basic information including the SVs' current location and information significance in a secure manner [24]. In this way, the VSP can observe useful information to select the best SVs (which act as the FL learners at each learning round, e.g., every hour or day) within its coverage (e.g., via

road side units [22]) with minimum privacy disclosure of the SVs' collected on-road data. Note that the accessed current location information is only used to help the VSP select the best SVs for each learning round (as similarly implemented in location-based services such as Google Maps, Placemeter, and Waze), and is totally different from on-road sensing data that the selected SVs collect when they run on the road. Specifically, let $\mathcal{T} = \{0, 1, \dots, t, \dots, T\}$ denote the set of rounds of an FL process and $\mathcal{M}(t) = \{1, \dots, m, \dots, M\}$, where $\mathcal{M}(t) \subset \mathcal{I}$, be the set of selected SVs based on their location significance at round t . The selection of M active SVs guarantees that they will provide useful on-road information due to their locations in the significant areas [14]. $\mathcal{N}(t) = \{1, \dots, n, \dots, N\}$, where $\mathcal{N}(t) \subset \mathcal{M}(t)$, denotes the set of selected best quality SVs based on their information significance at round t . The selection of N best SVs can ensure that the local trained model updates are reliable to provide high learning accuracy with faster convergence [20]. Alternatively, we denote the collected data size, i.e., number of samples, at SV- n for each learning round t as $\eta_n(t)$.

To compensate the selected SVs in $\mathcal{N}(t)$, the VSP allocates a maximum payment budget of $B_{max}(t)$ at each round t . Due to its economic benefit, the VSP practically keeps its willingness to pay (in regards to the received information significance from the FL learners) as a private information for the SVs, and it can be expressed as the type of the VSP [18]. This type is influenced by the VSP's current payment budget: a higher type indicates the willingness to pay more for the SVs in $\mathcal{N}(t)$ due to its higher payment budget. For that, we can define a finite set of the VSP's types as $\Theta = \{\theta_1, \dots, \theta_j, \dots, \theta_J\}$, where $j, j \in \mathcal{J} = \{1, \dots, j, \dots, J\}$, indicates the type index and we have $\theta_J > \theta_{J-1} > \dots > \theta_j > \dots > \theta_2 > \theta_1$ without loss of generality. Correspondingly, we denote $B_j(t)$ to be the payment budget for the VSP with type θ_j , where $B_j(t) = \frac{\theta_j}{\theta_J} B_{max}(t)$, $j \in \mathcal{J}$. As the type of the VSP is private, the SVs only can observe the VSP's type distribution, i.e., $\rho_j(t)$, where $\sum_{j=1}^J \rho_j(t) = 1, \forall j \in \mathcal{J}, \forall t \in \mathcal{T}$ [23]. According to $B_j(t)$, the VSP requires to determine the payment proportion for each SV- n . Particularly, given the VSP with type θ_j , we define payment proportion vector of all N SVs as $\boldsymbol{\rho}(t) = [\rho_1(t), \dots, \rho_j(t), \dots, \rho_J(t)]$, where $\rho_j(t) = [\rho_j^1(t), \dots, \rho_j^n(t), \dots, \rho_j^N(t)]$, and $0 \leq \rho_j^n(t) \leq 1, \forall j \in \mathcal{J}, \forall n \in \mathcal{N}(t)$. Moreover, for each learning round t , we denote $\boldsymbol{\zeta}(t) = [\zeta_1(t), \dots, \zeta_j(t), \dots, \zeta_J(t)]$ and $\boldsymbol{\varphi}(t) = [\varphi_1(t), \dots, \varphi_j(t), \dots, \varphi_J(t)]$ to be the vectors of information significance and corresponding payments of all participating SVs for all types, respectively. In this case, $\zeta_j(t) = [\zeta_j^1(t), \dots, \zeta_j^n(t), \dots, \zeta_j^N(t)]$, and $\varphi_j(t) = [\varphi_j^1(t), \dots, \varphi_j^n(t), \dots, \varphi_j^N(t)]$. As the VSP's type increases, each SV- n can offer higher information significance to the VSP since it can obtain higher payment from the VSP. Overall, the proposed dynamic FL-based economic framework for the IoV can be illustrated in Fig. 1. For each learning round t , the VSP performs the following processes:

- Step 1: The VSP observes active SVs in the network.
- Step 2: The VSP selects M active SVs based on their location significance using the K-means algorithm.
- Step 3: From M SVs, the VSP selects N best SVs which have the highest QoI (defined in Section 3).

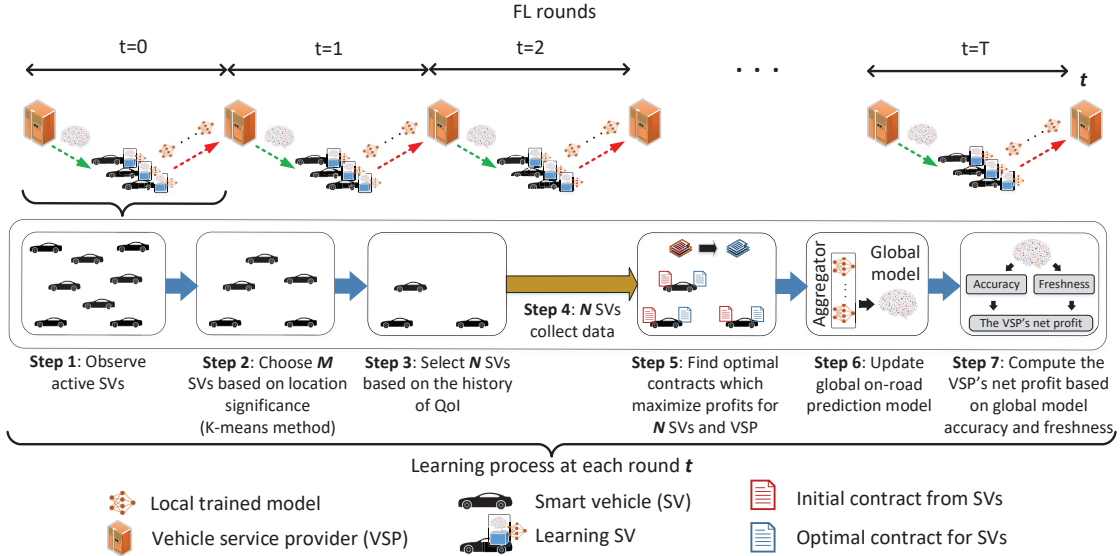


Fig. 1: The proposed dynamic FL-based economic framework for the IoV.

- Step 4: The N selected SVs collect on-road information data for the current learning round.
- Step 5: In the end of each learning round, these N SVs send initial contracts based on the collected on-road data to the VSP. Then, the VSP will determine and send the optimal contracts to the SVs.
- Step 6: Upon receiving the optimal contracts, all N SVs train their local datasets to obtain local trained models. These local models are then collected by the VSP to update the current global model.
- Step 7: The VSP updates its net profit based on the current round's global model accuracy and freshness.

At each learning round, the VSP may choose a different set of best SVs because of their diverse mobility behaviors. The above processes are repeated until the global on-road prediction model converges or after a pre-defined number of learning rounds is achieved. More details can be found in the following sections.

3 LOCATION AND INFORMATION SIGNIFICANCE-BASED SV SELECTION

In this section, we discuss on how the VSP can choose a set of best active SVs based on the location and information significance prior to the FL process at each learning round t .

3.1 Location Significance-Based SV Selection

In this method, the VSP can monitor the current location of the active SVs in the IoV using the GPS information stored in their on-road service applications locally. This current location information sharing can be securely protected, e.g., using anonymization or obfuscation-based approaches [24], [25]. From the active SVs observed in each learning round, the VSP can select M SVs, where $M \leq I$, whose locations are in the significant areas (i.e., the areas are of high interests by the VSP and/or the areas with high vehicle traffic volume based on their total AADFs for specific periods), e.g., big cities, central business districts, or tourist attractions. Let $\mathcal{D} = \{1, \dots, d, \dots, D\}$ denote the set of considered areas,

the fixed total AADF for each area- d prior to the learning process can be expressed by [12]

$$V_d = \sum_{\varpi_1=1}^{\varpi_1^{max}} \sum_{\varpi_2=1}^{\varpi_2^{max}} \frac{v_d^{\varpi_1, \varpi_2}}{365}, \quad (1)$$

where $\varpi_1 \in \{1, 2, \dots, \varpi_1^{max}\}$ and $\varpi_2 \in \{1, 2, \dots, \varpi_2^{max}\}$ specify the indices of considered year periods and days for a year (i.e., $\varpi_2^{max} = 365$), respectively. In addition, $v_d^{\varpi_1, \varpi_2}$ represents the vehicle traffic volume in area- d for day ϖ_2 of year ϖ_1 . The area with higher V_d value implies the area with higher significance. The reason is that many vehicles tend to visit the centralized areas where profitable employment market, desirable lifestyle, better educational occasion, and tourist attractions exist. In this case, the vehicles in those areas typically accommodate more useful and meaningful on-road information, e.g., traffic jams and/or accidents, due to their high vehicle traffic volume [14].

Using V_d from all areas in \mathcal{D} , the VSP can categorize those areas into insignificant and significant areas. For that, we implement a binary classification scheme using the K-means algorithm [26] in which groups '0' and '1' represent insignificant and significant areas, respectively. Consider that the VSP has a road traffic dataset containing area IDs, their major road locations, and their total AADFs of the vehicles $V_d, \forall d \in \mathcal{D}$. To attain the sets of insignificant and significant areas, the total AADF centroids \hat{V}^0 and \hat{V}^1 that minimize the overall squared distance between each V_d and the centroids can be determined using the following:

$$\min_{\{\sigma, \hat{V}\}} \sum_{d \in \mathcal{D}} \sigma_d^0 (V_d - \hat{V}^0)^2 + \sigma_d^1 (V_d - \hat{V}^1)^2, \quad (2)$$

$$\text{s.t. } \sigma_d^0 + \sigma_d^1 = 1, \forall d \in \mathcal{D}, \quad (3)$$

$$\sigma_d^0, \sigma_d^1 \in \{0, 1\}, \forall d \in \mathcal{D}, \quad (4)$$

where σ_d^0 and σ_d^1 are binary variables which indicate the total AADF closeness of area- d to the centroids \hat{V}^0 and \hat{V}^1 , respectively. The constraints (3) imply that the total AADF of each area- d can be classified into one group only. To find the optimal sets of insignificant and significant areas, the centroids \hat{V}^0 and \hat{V}^1 are first updated at each K-means

iteration l , i.e., $\hat{V}_{(l+1)}^0 = \frac{\sum_{d \in \mathcal{D}} \sigma_{d,(l)}^0 V_d}{\sum_{d \in \mathcal{D}} \sigma_{d,(l)}^0}$ and $\hat{V}_{(l+1)}^1 = \frac{\sum_{d \in \mathcal{D}} \sigma_{d,(l)}^1 V_d}{\sum_{d \in \mathcal{D}} \sigma_{d,(l)}^1}$, respectively, until $\hat{V}_{(l+1)}^0 = \hat{V}_{(l)}^0$ and $\hat{V}_{(l+1)}^1 = \hat{V}_{(l)}^1$. Hence, the VSP can obtain the optimal sets of insignificant areas, i.e., $\mathcal{D}^0 \subset \mathcal{D}$, and significant areas, i.e., $\mathcal{D}^1 \subset \mathcal{D}$, based on the value '1' of both σ_d^0 and σ_d^1 , $\forall d \in \mathcal{D}$, for all learning rounds. Then, the VSP can dynamically select M SVs whose current locations are within the set of significant areas \mathcal{D}^1 at each learning round t .

3.2 Information Significance-Based SV Selection

From the selected M SVs, the VSP can further choose N SVs as the FL learners, where $N \leq M$, to efficiently execute the FL algorithm for each round t . To this end, we introduce the information significance-based SV selection method to obtain N SVs which have the highest QoI. In particular, adopted from [7], suppose that $\mathcal{K} = \{1, \dots, k, \dots, K\}$ and $\mathcal{L} = \{1, \dots, l, \dots, L\}$ to be the set of captured on-road information timespans, e.g. between day 1 and day 7 for a week, and location IDs, e.g., between location 1 and location 100, from the SVs in $\mathcal{M}(t)$, respectively. Utilizing the timespan- k and location- l , each SV- m can calculate a spatio-temporal variability $\eta_m^{k,l}(t)$, i.e., the data size in location- l for timespan- k , from its total data size, i.e., $\eta_m(t)$, where $\eta_m(t) = \sum_{k=1}^K \sum_{l=1}^L \eta_m^{k,l}(t)$. This calculation can be conducted by the on-road service application provided by the VSP and stored at the SVs.

To implement the FL process at each round t , the VSP can determine the required spatio-temporal variability at timespan- k and location- l from SV- m , i.e., $y_m^{k,l}(t)$. However, due to the different mobile activities of the SV- m , $\forall m \in \mathcal{M}(t)$, on the roads in diverse locations for different times, each SV- m may have various actual spatio-temporal variability for each timespan- k and location- l , i.e., [7]

$$x_m^{k,l}(t) = \begin{cases} \eta_m^{k,l}(t), & \text{if } \eta_m^{k,l}(t) \leq y_m^{k,l}(t), \\ y_m^{k,l}(t), & \text{otherwise.} \end{cases} \quad (5)$$

Equation (5) implies that the actual spatio-temporal variabilities cannot exceed the required ones. From $x_m^{k,l}(t)$ and $y_m^{k,l}(t)$, we can define two matrices of actual and required spatio-temporal variabilities for each SV- m as follows:

$$\mathbf{X}_m(t) = \begin{pmatrix} x_m^{1,1}(t) & x_m^{1,2}(t) & \dots & x_m^{1,L}(t) \\ x_m^{2,1}(t) & x_m^{2,2}(t) & \dots & x_m^{2,L}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_m^{K,1}(t) & x_m^{K,2}(t) & \dots & x_m^{K,L}(t) \end{pmatrix}, \text{ and} \quad (6)$$

$$\mathbf{Y}_m(t) = \begin{pmatrix} y_m^{1,1}(t) & y_m^{1,2}(t) & \dots & y_m^{1,L}(t) \\ y_m^{2,1}(t) & y_m^{2,2}(t) & \dots & y_m^{2,L}(t) \\ \vdots & \vdots & \ddots & \vdots \\ y_m^{K,1}(t) & y_m^{K,2}(t) & \dots & y_m^{K,L}(t) \end{pmatrix}.$$

Then, we can derive the information significance metric as the normalized spatial length of a matrix [27] in Frobenius norm using (6), which is

$$\zeta^m(t) = 1 - \frac{\|\mathbf{Y}_m(t) - \mathbf{X}_m(t)\|_F}{\|\mathbf{Y}_m(t)\|_F} \quad (7)$$

$$= 1 - \frac{\sqrt{\sum_{k=1}^K \sum_{l=1}^L (y_m^{k,l}(t) - x_m^{k,l}(t))^2}}{\sqrt{\sum_{k=1}^K \sum_{l=1}^L (y_m^{k,l}(t))^2}}.$$

From (7), we can observe that the information significance for SV- m , $\forall m \in \mathcal{M}(t)$, is between 0 and 1, i.e., $0 \leq \zeta^m(t) \leq 1$. Specifically, $\zeta^m(t) = 0$ implies that no actual spatio-temporal variability is collected by the SV- m . Meanwhile, $\zeta^m(t) = 1$ specifies that the required spatio-temporal variabilities for all timespans and location IDs of the SV- m are fully satisfied. Using this information significance metric, the VSP can determine the dataset quality of the SVs in $\mathcal{M}(t)$ without disclosing any sensitive information of the SVs' actual datasets. Then, the VSP can select N SVs that will implement the FL process based on their largest $\zeta^m(t)$ values for each round t , i.e.,

$$\mathcal{N}(t) = \max_{[N]} \{\zeta^1(t), \dots, \zeta^M(t)\}. \quad (8)$$

Equation (8) specifies that only N SVs will exchange their local on-road models to update the global on-road prediction model at each round t in the FL process.

4 MPOA-BASED LEARNING CONTRACT OPTIMIZATION PROBLEM AND SOLUTION

Based on the list of selected SVs, the VSP can inform the selected SVs in $\mathcal{N}(t)$ to first collect on-road data for a particular period. Then, the SVs can send initial contracts at each learning round t including the information significance $\zeta_j^n(t) = \frac{\theta_j}{\theta_j} \zeta^n(t)$, $\forall n \in \mathcal{N}(t)$, $\forall j \in \mathcal{J}$, based on the collected data, and payment $\varphi_j^n(t) = v \zeta_j^n(t)$, $\forall n \in \mathcal{N}(t)$, $\forall j \in \mathcal{J}$ [6], where v is the information significance price unit. These initial contracts will be used to optimize the non-collaborative contract problem at the VSP. In this section, we first describe the MPOA-based contract optimization problem for the VSP (as the agent) and learning SVs in $\mathcal{N}(t)$ (as the principals). After that, a transformation method is introduced to reduce the problem complexity and an iterative algorithm is developed to find the equilibrium contract for the selected SVs.

4.1 Profit Optimization for the VSP

Considering the payment proportion vector for the VSP with type θ_j , i.e., $\boldsymbol{\rho}_j(t)$, information significance vector for the VSP with type θ_j , i.e., $\boldsymbol{\zeta}_j(t)$, and payment vector for the VSP with type θ_j , i.e., $\boldsymbol{\varphi}_j(t)$, the profit that the VSP with type θ_j can obtain from performing the learning process together with SVs in $\mathcal{N}(t)$ at round t can be formulated by

$$\mu_{VSP}^j(t) = \theta_j S(\boldsymbol{\rho}_j(t), \boldsymbol{\zeta}_j(t)) - C(\boldsymbol{\rho}_j(t), \boldsymbol{\varphi}_j(t)). \quad (9)$$

Specifically, $S(\boldsymbol{\rho}_j(t), \boldsymbol{\zeta}_j(t))$ and $C(\boldsymbol{\rho}_j(t), \boldsymbol{\varphi}_j(t))$ indicate the satisfaction and cost functions of the VSP, respectively. Moreover, the θ_j describes the weight of $S(\boldsymbol{\rho}_j(t), \boldsymbol{\zeta}_j(t))$ for the VSP with type θ_j . As such, the VSP with a higher type has a higher weight due to its willingness to pay more towards the FL process. For the satisfaction function, we use a squared-root function as similarly implemented in [23]. The reason is that the satisfaction increases when the QoI, i.e., information significance, gets better. However, the VSP may have less interest to further increase the satisfaction when the higher data quality triggers less on-road model accuracy improvement [28]. Thus, the satisfaction function can be expressed by

$$S(\boldsymbol{\rho}_j(t), \boldsymbol{\zeta}_j(t)) = \lambda \sqrt{\sum_{n \in \mathcal{N}(t)} \rho_j^n(t) \zeta_j^n(t)}, \quad (10)$$

where $\lambda > 0$ is a conversion parameter corresponding to the monetary unit of obtaining a certain information significance based on the existing data trading market [23]. For the cost function of the VSP, we can formulate it as the total payment transfer to the participating SVs in $\mathcal{N}(t)$ with respect to their information significance values used in the FL process, which is

$$C(\boldsymbol{\varrho}_j(t), \boldsymbol{\varphi}_j(t)) = \sum_{n \in \mathcal{N}(t)} \varrho_j^n(t) \varphi_j^n(t). \quad (11)$$

From (9)-(11), we can devise an optimization problem to maximize the profit of VSP with type θ_j by

$$(\mathbf{P}_1) \quad \max_{\boldsymbol{\varrho}_j(t)} \theta_j S(\boldsymbol{\varrho}_j(t), \boldsymbol{\zeta}_j(t)) - C(\boldsymbol{\varrho}_j(t), \boldsymbol{\varphi}_j(t)). \quad (12)$$

$$\text{s.t.} \quad C(\boldsymbol{\varrho}_j(t), \boldsymbol{\varphi}_j(t)) \leq B_j(t), \quad (13)$$

$$0 \leq \varrho_j^n(t) \leq 1, \forall n \in \mathcal{N}(t), \quad (14)$$

where the constraint (13) implies that the total payment for SVs in the FL process cannot exceed the maximum payment budget of the VSP with type θ_j , i.e., $B_j(t)$. Due to the convexity of objective function (12), i.e., the satisfaction and cost functions are concave and linear functions, respectively, and the linearity of the constraints (13)-(14) in (\mathbf{P}_1) , it can be guaranteed that we can find the optimal $\hat{\boldsymbol{\varrho}}_j(t) = [\hat{\varrho}_j^1(t), \dots, \hat{\varrho}_j^n(t), \dots, \hat{\varrho}_j^N(t)]$, $\forall j \in \mathcal{J}$, straightforwardly.

4.2 Profit Optimization for Learning SVs

In this section, we formulate the FL contract optimization problem to maximize the expected profit of each SV- n , $n \in \mathcal{N}(t)$, individually under the constraints of the VSP. Specifically, from the optimal payment proportion $\hat{\boldsymbol{\varrho}}_j(t)$, $\forall j \in \mathcal{J}$, each SV- n can calculate its expected profit through taking all possible types of the VSP into account such that

$$\mu_n(\boldsymbol{\zeta}(t), \boldsymbol{\varphi}(t)) = \sum_{j=1}^J \left(\hat{\varrho}_j^n(t) \varphi_j^n(t) - \hat{\varrho}_j^n(t) \zeta_j^n(t) \xi_n(t) \right) \rho_j(t), \quad (15)$$

where $\xi_n(t)$ indicates the computation and memory costs of SV- n in training the local dataset with certain information significance at round t . Additionally, the utilization of $\sum(\cdot)$ specifies that the expected profit of each SV- n at round t depends on the VSP's type distribution $\rho_j(t)$, $\forall j \in \mathcal{J}$.

In this optimization problem, we can obtain the optimal contracts which satisfy the individual rationality (IR) and incentive compatibility (IC) constraints from the VSP. The IR constraints guarantee that the VSP with a certain type will always obtain a positive profit as described in Definition 1.

Definition 1. *IR constraints: The VSP must achieve a non-negative profit, i.e.,*

$$\theta_j S(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\zeta}_j(t)) - C(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\varphi}_j(t)) \geq 0, \forall j \in \mathcal{J}, \quad (16)$$

for each $t \in \mathcal{T}$ to participate in the FL contract optimization.

Due to the information asymmetry between the VSP and SVs, in addition to the IR constraints, the optimal contracts must also satisfy the IC constraints which ensure the feasibility of the contracts. In particular, the IC constraints guarantee that the VSP can always obtain the maximum profit if an appropriate contract, i.e., the contract designed for that current type of the VSP, is utilized. This condition can be defined in the following Definition 2.

Definition 2. *IC constraints: The VSP with current type θ_j will likely to choose a contract designed for its current type θ_j rather than with another type θ_{j^*} , i.e.,*

$$\theta_j S(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\zeta}_j(t)) - C(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\varphi}_j(t)) \geq \quad (17)$$

$$\theta_{j^*} S(\hat{\boldsymbol{\varrho}}_{j^*}(t), \boldsymbol{\zeta}_{j^*}(t)) - C(\hat{\boldsymbol{\varrho}}_{j^*}(t), \boldsymbol{\varphi}_{j^*}(t)), j \neq j^*, \forall j, j^* \in \mathcal{J}.$$

Given the expected profit of each SV- n in (15) as well as IR and IC constraints of the VSP in (16) and (17), respectively, we can formulate a non-collaborative learning contract optimization due to the selfish nature of participating SVs. As such, the FL contract optimization problem (\mathbf{P}_2) to maximize the expected profit for SV- n independently at the VSP for each t can be expressed by

$$(\mathbf{P}_2) \quad \max_{\boldsymbol{\zeta}(t), \boldsymbol{\varphi}(t)} \mu_n(\boldsymbol{\zeta}(t), \boldsymbol{\varphi}(t)), \forall n \in \mathcal{N}(t), \quad (18)$$

$$\text{s.t.} \quad C(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\varphi}_j(t)) \leq B_j(t), \forall j \in \mathcal{J}, \quad (19)$$

$$\theta_j S(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\zeta}_j(t)) - C(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\varphi}_j(t)) \geq 0, \forall j \in \mathcal{J}, \quad (20)$$

$$\theta_j S(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\zeta}_j(t)) - C(\hat{\boldsymbol{\varrho}}_j(t), \boldsymbol{\varphi}_j(t)) \geq \quad (21)$$

$$\theta_{j^*} S(\hat{\boldsymbol{\varrho}}_{j^*}(t), \boldsymbol{\zeta}_{j^*}(t)) - C(\hat{\boldsymbol{\varrho}}_{j^*}(t), \boldsymbol{\varphi}_{j^*}(t)), j \neq j^*, \forall j, j^* \in \mathcal{J}.$$

Using the optimization problem (\mathbf{P}_2) , each SV- n via the VSP can maximize its profit under the competition from other SVs in $\mathcal{N}(t)$ and the limited payment budget of the VSP.

4.3 Social Welfare of the Internet-of-Vehicles

To further derive the social welfare of the IoV, i.e., the total actual profits of the VSP and all learning SVs in $\mathcal{N}(t)$, we can find the actual profit of each SV- n when the VSP has the type θ_j as $\hat{\varrho}_j^n(t) \varphi_j^n(t) - \hat{\varrho}_j^n(t) \zeta_j^n(t) \xi_n(t)$. Then, given the VSP has type θ_j , we can calculate the total actual profit of all SVs in $\mathcal{N}(t)$ by

$$\sum_{n \in \mathcal{N}(t)} \left(\hat{\varrho}_j^n(t) \varphi_j^n(t) - \hat{\varrho}_j^n(t) \zeta_j^n(t) \xi_n(t) \right). \quad (22)$$

As a result, we can calculate the social welfare at round t for type θ_j according to (9)-(11) and (22) by

$$\begin{aligned} \mu_{SW}^j(t) &= \theta_j \lambda \sqrt{\sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \zeta_j^n(t) - \sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \varphi_j^n(t)} \\ &+ \sum_{n \in \mathcal{N}(t)} \left(\hat{\varrho}_j^n(t) \varphi_j^n(t) - \hat{\varrho}_j^n(t) \zeta_j^n(t) \xi_n(t) \right) \\ &= \theta_j \lambda \sqrt{\sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \zeta_j^n(t) - \sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \zeta_j^n(t) \xi_n(t)}, \forall j \in \mathcal{J}. \end{aligned} \quad (23)$$

4.4 Contract Problem Transformation

In Proposition 1, we can observe that solving (\mathbf{P}_2) requires computational complexity $O(J^2)$. As a result, this will incur a large of computational resources, especially when the number of VSP's types and the number of learning rounds are high.

Proposition 1. *The computational complexity of solving (\mathbf{P}_2) is $O(J^2)$.*

Proof. From (19)-(21), we can observe that there are J payment budget constraints, J IR constraints, and $J(J-1)$ IC constraints. Thus, the computational complexity of solving problem (\mathbf{P}_2) is $O(J^2)$. \square

Then, using the following transformation method, i.e., transforming the IR and IC constraints, we can reduce the

computational complexity of solving (\mathbf{P}_2) from $O(J^2)$ into $O(J)$. Specifically, we first demonstrate that when the VSP's type is higher (i.e., the willingness to pay more due to higher payment budget), the learning SV- $n, \forall n \in \mathcal{N}(t)$, will offer higher information significance values to the VSP. This aims to obtain higher payments from the VSP. This condition can be formally written in Lemma 1.

Lemma 1. *Let $(\zeta(t), \varphi(t))$ denote any feasible contracts from the SVs to the VSP such that $\zeta_j(t) \geq \zeta_{j^*}(t)$ if and only if $\theta_j \geq \theta_{j^*}$, where $j \neq j^*, j, j^* \in \mathcal{J}$.*

Proof. See Appendix A. \square

Since $\zeta_j(t) \geq \zeta_{j^*}(t)$, in the following Proposition 2, we show that the SVs will request higher payment to the VSP correspondingly when the VSP has type θ_j compared with when it occupies type θ_{j^*} .

Proposition 2. *If $\zeta_j(t) \geq \zeta_{j^*}(t)$, then $\varphi_j(t) \geq \varphi_{j^*}(t)$, where $j \neq j^*, j, j^* \in \mathcal{J}$.*

Proof. See Appendix B. \square

According to Lemma 1 and Proposition 2, we can observe in the following Proposition 3 that the profit of the VSP is a monotonic increasing function of its type.

Proposition 3. *For any feasible contract $(\zeta(t), \varphi(t))$, the profit of the VSP must satisfy*

$$\begin{aligned} \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) &\geq \\ \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)), \end{aligned} \quad (24)$$

where $\theta_j \geq \theta_{j^*}, j \neq j^*, j, j^* \in \mathcal{J}$.

Proof. See Appendix C. \square

From Proposition 3, we can reduce the number of IR constraints by using the VSP's minimum type, i.e., θ_1 . As such, we can derive that

$$\begin{aligned} \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) &\geq \\ \theta_j S(\hat{\rho}_1(t), \zeta_1(t)) - C(\hat{\rho}_1(t), \varphi_1(t)) & \\ \geq \theta_1 S(\hat{\rho}_1(t), \zeta_1(t)) - C(\hat{\rho}_1(t), \varphi_1(t)) &\geq 0. \end{aligned} \quad (25)$$

In other words, the IR constraints for other types $\theta_j, j > 1$, will automatically hold if and only if we can satisfy the IR constraint for θ_1 . To this end, we can modify the IR constraints in (20) into

$$\theta_1 S(\hat{\rho}_1(t), \zeta_1(t)) - C(\hat{\rho}_1(t), \varphi_1(t)) \geq 0. \quad (26)$$

For the IC constraints in (21), we can also reduce the number of constraints using the following transformation stated in Lemma 2.

Lemma 2. *The IC constraints in (21) of (\mathbf{P}_2) can be modified into the following local downward incentive constraints (LDIC), i.e.,*

$$\begin{aligned} \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) &\geq \\ \theta_j S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t)) - C(\hat{\rho}_{j-1}(t), \varphi_{j-1}(t)), & \\ \forall j \in \{2, \dots, J\}, & \end{aligned} \quad (27)$$

where $\zeta_j(t) \geq \zeta_{j-1}(t), \forall j \in \{2, \dots, J\}$.

Proof. See Appendix D. \square

The conditions in (27) imply that if the IC constraint regarding the type that is one type lower than type θ_j holds, i.e., θ_{j-1} , then all other IC constraints are also satisfied as

long as the conditions in Lemma 1 hold. From (26)-(27), we can update the optimization problem (\mathbf{P}_2) into the problem (\mathbf{P}_3) as follows:

$$(\mathbf{P}_3) \quad \max_{\zeta(t), \varphi(t)} \mu_n(\zeta(t), \varphi(t)), \forall n \in \mathcal{N}(t), \quad (28)$$

s.t. (19), and,

$$\theta_1 S(\hat{\rho}_1(t), \zeta_1(t)) - C(\hat{\rho}_1(t), \varphi_1(t)) \geq 0, \quad (29)$$

$$\theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) \geq$$

$$\theta_j S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t)) - C(\hat{\rho}_{j-1}(t), \varphi_{j-1}(t)),$$

$$\forall j \in \{2, \dots, J\}, \quad (30)$$

$$\zeta_j(t) \geq \zeta_{j-1}(t), \forall j \in \{2, \dots, J\}. \quad (31)$$

From problem (\mathbf{P}_3) , we demonstrate that the problem now can be solved with computational complexity $O(J)$ as stated in the following Proposition 4.

Proposition 4. *The computational complexity of solving (\mathbf{P}_3) is $O(J)$.*

Proof. From the constraints (19), (29)-(31), there are $J + 1 + (J - 1) + (J - 1) = 3J - 1$ number of the constraints. Hence, the computational complexity of solving problem (\mathbf{P}_3) becomes $O(J)$. \square

4.5 Learning Contract Iterative Algorithm

Based on the problem (\mathbf{P}_3) , we can find the optimal contracts $(\hat{\zeta}(t), \hat{\varphi}(t))$ by implementing an iterative algorithm as described in Algorithm 1. Specifically, the SVs in $\mathcal{N}(t)$ first offer the initial contracts at iteration $\varsigma = 0$, i.e., $(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t)), \forall n \in \mathcal{N}(t)$, where $\zeta_{(\varsigma)}^n(t) = [\zeta_{1,(\varsigma)}^n(t), \dots, \zeta_{j,(\varsigma)}^n(t), \dots, \zeta_{J,(\varsigma)}^n(t)]$ and $\varphi_{(\varsigma)}^n(t) = [\varphi_{1,(\varsigma)}^n(t), \dots, \varphi_{j,(\varsigma)}^n(t), \dots, \varphi_{J,(\varsigma)}^n(t)]$. Given $(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t)), \forall n \in \mathcal{N}(t)$, the VSP can find the optimal payment proportion values $\hat{\rho}_{(\varsigma)}(t)$ which maximize the objective function of (\mathbf{P}_1) . Then, using the obtained $\hat{\rho}_{(\varsigma)}(t)$ and other SVs' current contracts $(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t))$ [29], the VSP can iteratively find a new contract for the SV- $n, n \in \mathcal{N}(t)$, to maximize the objective function of (\mathbf{P}_3) at iteration $\varsigma + 1$.

The algorithm continues until the gaps between the expected profits at iteration ς and $\varsigma + 1$ of all the SVs is equal or less than the optimality tolerance γ , and thus the algorithm converges and the equilibrium contract solution can be found (as proven in Theorems 1 and 2). In this case, the expected profit of the SV- n at the equilibrium contract solution $(\hat{\zeta}^n(t), \hat{\varphi}^n(t))$ will achieve the highest value compared with all other contract solutions it takes, given the equilibrium contract of other SVs $(\hat{\zeta}^{-n}(t), \hat{\varphi}^{-n}(t))$. Alternatively, no SV- n can improve its expected profit by unilaterally deviating from its equilibrium strategy $(\hat{\zeta}^n(t), \hat{\varphi}^n(t))$, as described in Definition 3.

Definition 3. *Equilibrium solution for non-collaborative learning contract optimization problem: The optimal contracts $(\hat{\zeta}(t), \hat{\varphi}(t)) = (\hat{\zeta}^n(t), \hat{\varphi}^n(t), \hat{\zeta}^{-n}(t), \hat{\varphi}^{-n}(t))$ are the equilibrium solution of (\mathbf{P}_3) at learning round t if and only if*

$$\begin{aligned} \mu_n(\hat{\zeta}^n(t), \hat{\varphi}^n(t), \hat{\zeta}^{-n}(t), \hat{\varphi}^{-n}(t)) &\geq \\ \mu_n(\zeta^n(t), \varphi^n(t), \hat{\zeta}^{-n}(t), \hat{\varphi}^{-n}(t)), \forall n \in \mathcal{N}(t), & \end{aligned} \quad (32)$$

hold and the optimal contracts $(\hat{\zeta}(t), \hat{\varphi}(t))$ still satisfy the constraints (29)-(31).

Algorithm 1 Learning Contract Iterative Algorithm

- 1: Set $\varsigma = 0$ and $\gamma > 0$
 - 2: The VSP informs the selected SVs in $\mathcal{N}(t)$ to send initial contracts
 - 3: **for** $\forall n \in \mathcal{N}(t)$ **do**
 - 4: SV- n offers $(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t))$ to the VSP
 - 5: **end for**
 - 6: **repeat**
 - 7: Obtain $\hat{\varrho}_{(\varsigma)}(t)$ that maximize (\mathbf{P}_1) given $(\zeta_{(\varsigma)}(t), \varphi_{(\varsigma)}(t))$
 - 8: **for** $\forall n \in \mathcal{N}(t)$ **do**
 - 9: Find the new contract $(\zeta_{new}^n(t), \varphi_{new}^n(t))$, which maximizes (\mathbf{P}_3) using $\hat{\varrho}_{(\varsigma)}(t)$ and $(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$
 - 10: **if** $\left[\mu_n(\zeta_{new}^n(t), \varphi_{new}^n(t), \zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) - \mu_n(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t), \zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) \right] > \gamma$ **then**
 - 11: Set $(\zeta_{(\varsigma+1)}^n(t), \varphi_{(\varsigma+1)}^n(t)) = (\zeta_{new}^n(t), \varphi_{new}^n(t))$
 - 12: **else**
 - 13: Set $(\zeta_{(\varsigma+1)}^n(t), \varphi_{(\varsigma+1)}^n(t)) = (\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t))$
 - 14: **end if**
 - 15: **end for**
 - 16: $\varsigma = \varsigma + 1$
 - 17: **until** $\mu_n(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t), \zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)), \forall n \in \mathcal{N}(t)$, do not change anymore
 - 18: Generate optimal contracts $(\hat{\zeta}(t), \hat{\varphi}(t))$
-

4.6 Convergence, Equilibrium, and Complexity Analysis of the Learning Contract Iterative Algorithm

To analyze the convergence and equilibrium contract for Algorithm 1, we consider two steps for each iteration ς . First, each SV- n , $n \in \mathcal{N}(t)$, generates contract $(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t))$ and sends it to the VSP. Then, the VSP helps the SVs to find their optimal payment proportions $\hat{\varrho}_{(\varsigma)}(t)$ locally due to the unknown contract information among the SVs. In this case, we can show that Algorithm 1 converges to the equilibrium contract solution by using the best response method [29]. In particular, the best response of SV- n at iteration $\varsigma + 1$ given $(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$ can be expressed by

$$\Phi_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) = \quad (33)$$

$$\arg \max_{\{\zeta_{new}^n(t), \varphi_{new}^n(t)\} \in \mathbb{C}_n} \mu_n(\zeta_{new}^n(t), \varphi_{new}^n(t), \zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)),$$

where \mathbb{C}_n is the non-empty contract space [36] for SV- n and $\mathbb{C} = \prod_{n \in \mathcal{N}(t)} \mathbb{C}_n$. As the monotonicity is not ensured for each SV- n , we can use $(\zeta_{new}^n(t), \varphi_{new}^n(t)) \in \Phi_{(\varsigma+1)}^n$ to be $(\zeta_{(\varsigma+1)}^n(t), \varphi_{(\varsigma+1)}^n(t))$ if the following condition satisfies

$$\left[\mu_n(\zeta_{new}^n(t), \varphi_{new}^n(t), \zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) - \mu_n(\zeta_{(\varsigma)}^n(t), \varphi_{(\varsigma)}^n(t), \zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) \right] > \gamma. \quad (34)$$

The above iteration stops when the algorithm converges for all SVs in $\mathcal{N}(t)$ as stated in Theorem 1.

THEOREM 1. The best response iterative Algorithm 1 converges under the optimality tolerance γ .

Proof. See Appendix E. \square

Next, we need to ensure that the Algorithm 1 also converges to the equilibrium solution $(\hat{\zeta}(t), \hat{\varphi}(t))$. Specifically, we first investigate that the equilibrium solution exists by finding a fixed point in a set-valued function $\Phi, \Phi: \mathbb{C} \rightarrow 2^{\mathbb{C}}$, which is $\Phi = [\Phi^n(\zeta^{-n}(t), \varphi^{-n}(t)), \Phi^{-n}(\zeta^n(t), \varphi^n(t))]$. The revelation of this fixed point is equivalent to the equilibrium solution [19], [36], in which the Algorithm 1 converges to the equilibrium contract solution, i.e., all the SVs in $\mathcal{N}(t)$ obtain the maximum expected profits where $(\hat{\zeta}(t), \hat{\varphi}(t))$ is found. This is formally stated in Theorem 2.

THEOREM 2. If $(\hat{\zeta}(t), \hat{\varphi}(t))$ is a fixed point in Φ , then the best response iterative process in Algorithm 1 converges to its equilibrium contract solution $(\hat{\zeta}(t), \hat{\varphi}(t))$.

Proof. See Appendix F. \square

Finally, the complexity of the Algorithm 1 with N SVs is formally stated in the following Theorem 3.

THEOREM 3. Algorithm 1 has polynomial complexity $\log(N) + e^z + O(1/N)$, where z is the Euler constant.

Proof. See Appendix G. \square

5 FEDERATED LEARNING WITH SV SELECTION

5.1 Learning Process

To improve the global on-road prediction model accuracy, we implement the FL process with SV selection at each round [11] using the latest information significance (based on the above optimal contracts $(\hat{\zeta}(t), \hat{\varphi}(t))$ of SVs in $\mathcal{N}(t)$). For the FL algorithm, we adopt a deep learning approach leveraging deep neural networks (DNN) for a classification prediction model, i.e., when the output layer of DNN produces discrete prediction values. Particularly, we specify $\mathbf{A}_n(t)$ and $\mathbf{G}_n(t)$ to be the training feature data (using features as columns) and ground-truth label data within $\eta_n(t)$ of SV- n , $n \in \mathcal{N}(t)$, respectively. Thus, the total number of samples for all SVs in $\mathcal{N}(t)$ can be derived by $\eta(t) = \sum_{n \in \mathcal{N}(t)} \eta_n(t)$. When multiple layers of the DNN are considered, we have $\mathbf{A}_n^\ell(t)$ such that $\mathbf{A}_n^1(t) = \mathbf{A}_n(t)$, where ℓ is the training layer, $\ell \in [1, 2, \dots, \ell_{max}]$. For each layer- ℓ , we can compute the training output data $\hat{\mathbf{G}}_n^\ell(t)$ as described by

$$\hat{\mathbf{G}}_n^\ell(t) = \alpha_n^\ell(\mathbf{A}_n^\ell(t)\mathbf{W}(t)), \quad (35)$$

where $\mathbf{W}(t)$ is the global model matrix containing weights. $\alpha_n^\ell(\cdot)$ indicates the activation function of SV- n for nonlinear transformation. For $\ell \in [1, 2, \dots, \ell_{max} - 1]$, we adopt a *tanh* activation function [30], i.e.,

$$\alpha_n^\ell(\mathbf{A}_n^\ell(t)\mathbf{W}(t)) = \frac{e^{\mathbf{A}_n^\ell(t)\mathbf{W}(t)} - e^{-\mathbf{A}_n^\ell(t)\mathbf{W}(t)}}{e^{\mathbf{A}_n^\ell(t)\mathbf{W}(t)} + e^{-\mathbf{A}_n^\ell(t)\mathbf{W}(t)}}. \quad (36)$$

For the final layer, i.e., $\ell = \ell_{max}$, we utilize a *softmax* activation function which is typically used to interpret a probability distribution for a classification model [30], i.e.,

$$\alpha_n^{\ell_{max}} \left(\mathbf{A}_n^{\ell_{max}}(t) \mathbf{W}(t) \right) = \frac{e^{\mathbf{A}_n^{\ell_{max}}(t) \mathbf{W}(t)}}{\sum e^{\mathbf{A}_n^{\ell_{max}}(t) \mathbf{W}(t)}}. \quad (37)$$

To extract more meaningful features from the training input, we utilize several hidden layers ℓ , where $1 < \ell < \ell_{max}$, such that $\mathbf{A}_n^{\ell+1}(t) = \hat{\mathbf{G}}_n^\ell(t)$. At the final layer ℓ_{max} , we can obtain the predicted label data $\hat{\mathbf{G}}_n^{\ell_{max}}(t)$ to calculate the local loss function of SV- n at each t using a squared Frobenius norm as follows:

$$\begin{aligned} \epsilon_n \left(\mathbf{W}(t) \right) &= \frac{1}{\eta_n(t)} \left\| \mathbf{G}_n(t) - \hat{\mathbf{G}}_n^{\ell_{max}}(t) \right\|_F^2 \\ &= \frac{1}{\eta_n(t)} \sum_{s=1}^{\eta_n(t)} \left(g_n^s(t) - \hat{g}_n^s(t) \right)^2, \end{aligned} \quad (38)$$

where $g_n^s(t)$ and $\hat{g}_n^s(t)$ are the elements of ground truth label data $\mathbf{G}_n(t)$ and predicted label data $\hat{\mathbf{G}}_n^{\ell_{max}}(t)$ at row- s for round t , respectively. From $\epsilon_n \left(\mathbf{W}(t) \right)$, we can compute the local loss gradient at SV- n by

$$\nabla \mathbf{W}_n(t) = \frac{\partial \epsilon_n \left(\mathbf{W}(t) \right)}{\partial \mathbf{W}(t)}. \quad (39)$$

Using this gradient $\nabla \mathbf{W}_n(t)$, each SV- n can update the local on-road model $\mathbf{W}_n(t)$ through utilizing the *Adam* optimizer [31] which minimizes $\epsilon_n \left(\mathbf{W}(t) \right)$. This optimizer operates as the adaptive learning rate to obtain high robustness and fast convergence to the $\mathbf{W}_n(t)$. In this case, we can derive the local update rules of the exponential moving average of the $\nabla \mathbf{W}_n(t)$, i.e., $p_n(t)$, and the squared $\nabla \mathbf{W}_n(t)$ to obtain the variance, i.e., $q_n(t)$, as follows:

$$\begin{aligned} p_n^{(\tau+1)}(t) &= \beta_{p_n}^{(\tau)}(t) p_n^{(\tau)}(t) + \left(1 - \beta_{p_n}^{(\tau)}(t) \right) \nabla \mathbf{W}_n(t), \\ q_n^{(\tau+1)}(t) &= \beta_{q_n}^{(\tau)}(t) q_n^{(\tau)}(t) + \left(1 - \beta_{q_n}^{(\tau)}(t) \right) \left(\nabla \mathbf{W}_n(t) \right)^2, \end{aligned} \quad (40)$$

where $\beta_{p_n}^{(\tau)}(t) \in [0, 1)$ and $\beta_{q_n}^{(\tau)}(t) \in [0, 1)$ represent $p_n^{(\tau)}(t)$'s and $q_n^{(\tau)}(t)$'s steps of the exponential decays at local iteration τ , respectively. Using the learning step κ_n , we can determine how often the $\mathbf{W}_n(t)$ is updated for the next local iteration $\tau + 1$, which can be calculated by

$$\kappa_n^{(\tau+1)}(t) = \kappa_n \frac{\sqrt{1 - \beta_{q_n}^{(\tau+1)}(t)}}{1 - \beta_{p_n}^{(\tau+1)}(t)}. \quad (41)$$

Then, we can update the $\mathbf{W}_n^{(\tau+1)}(t)$ for the next $\tau + 1$ by

$$\mathbf{W}_n^{(\tau+1)}(t) = \mathbf{W}_n^{(\tau)}(t) - \kappa_n^{(\tau+1)}(t) \frac{p_n^{(\tau+1)}(t)}{\sqrt{q_n^{(\tau+1)}(t) + \varepsilon}}, \quad (42)$$

where ε specifies a constant to avoid zero division when $\sqrt{q_n^{(\tau+1)}(t)}$ reaches zero. This local process continues until a pre-defined local iteration threshold τ_{th} is reached, and thus $\kappa_n^{(\tau_{th})}(t)$ and $\mathbf{W}_n^{(\tau_{th})}(t)$ are obtained. To this end, the SV- n can send its $\mathbf{W}_n^{(\tau_{th})}(t)$ to the VSP for the global on-road model update $\mathbf{W}(t + 1)$ by aggregating all $\mathbf{W}_n^{(\tau_{th})}(t)$, $\forall n \in \mathcal{N}(t)$, as follows:

$$\mathbf{W}(t + 1) = \frac{1}{\eta(t)} \sum_{n \in \mathcal{N}(t)} \eta_n(t) \mathbf{W}_n^{(\tau_{th})}(t). \quad (43)$$

To this end, we can compute the global loss function at

Algorithm 2 FL Process with SV Selection Algorithm

- 1: Set initial τ_{th} , t_{th} , $\mathbf{W}(0)$, and $t = 0$
 - 2: **while** $t \leq t_{th}$ **and** $\Psi(\mathbf{W}(t))$ does not converge **do**
 - 3: The VSP determines SVs in $\mathcal{N}(t) \subset \mathcal{M}(t) \subset \mathcal{I}$ using (8)
 - 4: All SVs in $\mathcal{N}(t)$ collect on-road data for a certain time
 - 5: Execute Algorithm 1 for all SVs in $\mathcal{N}(t)$ to obtain optimal contracts $\left(\hat{\zeta}(t), \hat{\varphi}(t) \right)$
 - 6: **for** $\forall n \in \mathcal{N}(t)$ **do**
 - 7: Set $\mathbf{A}_n(t)$, $\mathbf{G}_n(t)$ from $\eta_n(t)$ based on the optimal $\hat{\zeta}^n(t)$
 - 8: Calculate $\hat{\mathbf{G}}_n^{\ell_{max}}(t)$ using $\mathbf{A}_n(t)$ and $\mathbf{W}(t)$
 - 9: Derive $\epsilon_n \left(\mathbf{W}(t) \right)$ and $\mathbf{W}_n^{(\tau_{th})}(t)$
 - 10: Send $\mathbf{W}_n^{(\tau_{th})}(t)$ to the VSP
 - 11: **end for**
 - 12: Update the global on-road model $\mathbf{W}(t+1)$ using $\mathbf{W}_n^{(\tau_{th})}(t)$ and $\eta_n(t)$, $\forall n \in \mathcal{N}(t)$
 - 13: Determine the global loss $\Psi(\mathbf{W}(t + 1))$ using $\epsilon_n \left(\mathbf{W}_n^{(\tau_{th})}(t) \right)$, $\forall n \in \mathcal{N}(t)$
 - 14: $t = t + 1$
 - 15: **end while**
 - 16: Obtain the final global on-road model \mathbf{W}^* and global loss $\Psi^*(\mathbf{W}^*)$
-

round $t + 1$ which is expressed by

$$\Psi(\mathbf{W}(t + 1)) = \frac{1}{N} \sum_{n \in \mathcal{N}(t)} \epsilon_n \left(\mathbf{W}_n^{(\tau_{th})}(t) \right). \quad (44)$$

This global process repeats until the global loss converges or the number of rounds reaches a given threshold t_{th} , and thus the final global on-road model \mathbf{W}^* and the final global loss $\Psi^*(\mathbf{W}^*)$ are obtained. The proposed FL algorithm with SV selection and MPOA contract-based model is summarized in Algorithm 2.

5.2 Convergence Analysis

In this section, we investigate the convergence of the proposed FL algorithm with SV selection using the gap between the expected global loss after t^\diamond rounds, i.e., $\mathbb{E}[\Psi(\mathbf{W}(t^\diamond))]$ and the final global loss, i.e., $\Psi^*(\mathbf{W}^*)$. Consider the probability that an SV will be chosen by the VSP for the FL execution at any round is $\frac{N}{T}$. Then, we can demonstrate that the global loss gap is upper bounded by the expected squared L2-norm global model gap, i.e., $\mathbb{E}[\|\mathbf{W}(t^\diamond) - \mathbf{W}^*\|_2^2]$, and it eventually reaches zero as stated in Theorem 4.

THEOREM 4. *The proposed FL with SV selection in Algorithm 2 will converge to the minimum global loss $\Psi^*(\mathbf{W}^*)$ according to the global loss gap condition $[\mathbb{E}[\Psi(\mathbf{W}(t^\diamond))] - \Psi^*(\mathbf{W}^*)] \leq \frac{\delta}{2} \mathbb{E}[\|\mathbf{W}(t^\diamond) - \mathbf{W}^*\|_2^2]$, where δ is a positive constant.*

Proof. Given that the global loss function after t^\diamond rounds, i.e., $\Psi(\mathbf{W}(t^\diamond))$, is δ -smooth, we can show that $\mathbb{E}[\Psi(\mathbf{W}(t^\diamond))] - \Psi^*(\mathbf{W}^*) \leq \frac{\delta}{2} \mathbb{E}[\|\mathbf{W}(t^\diamond) - \mathbf{W}^*\|_2^2]$. More details are provided in Appendix H. \square

To this end, it is intractable to provide the exact convergence rate theoretically due to the dynamic set of learning SVs with various QoI at each learning round. Nonetheless, in the simulation results of Section 7, i.e., Fig. 8 and Fig. 9, we can show that the proposed FL algorithm with SV selection can speed up the convergence to the $\Psi^*(\mathbf{W}^*)$, where $[\mathbb{E}[\Psi(\mathbf{W}(t^\diamond))] - \Psi^*(\mathbf{W}^*)] = 0$, at a certain learning round t^\diamond .

6 VSP'S PROFIT ANALYSIS BASED ON THE GLOBAL MODEL ACCURACY AND FRESHNESS

When more active SVs with high information significance are recruited for the FL execution, the converged global on-road model with high accuracy can be achieved faster. In this way, the VSP can obtain more up-to-date global on-road model, and thus achieve a higher freshness value when selling this model to the potential customers, such as SVs, public transportations, and mobile users. The adoption of information freshness in FL has been recently investigated in the literature. For example, the authors in [37] propose an FL scheduling strategy considering the freshness of FL learners' local trained models, aiming at computing the staleness level of the local models for the global model update. In [38], the authors present an incentive mechanism for the FL process accounting for the freshness of collected data at each FL learner towards the completion of FL process. Different from the existing works, we consider the freshness of produced global model at each learning round and analyze its impact on the VSP's profit.

In practice, the profit of the VSP is influenced by not only the decision of contract optimization, but also the obtained global model accuracy and freshness. To the best of our knowledge, this is the first analysis in the literature accounting for the influence of global model accuracy and freshness towards the economic aspect of the typical FL process. Specifically, we introduce a dynamic global model economic value considering the product of global model accuracy and freshness at each learning round t , which is expressed by

$$\omega(t) = a\chi(t)e^{-bt}, \quad (45)$$

where a, b are control parameters which can be defined in advance based on the accuracy and freshness of global model, $\chi(t)$ is the global prediction model accuracy at learning round t , and e^{-bt} is the global model freshness function with respect to the learning round t [32], [33]. The relationship between the global model accuracy and freshness in (45) implies that there is a trade-off between them in the FL process. Specifically, to obtain a global model with high accuracy, the FL algorithm usually requires to be executed for multiple learning rounds, that leads to a lower freshness value, and vice versa. Furthermore, the proposed dynamic global model economic value can be used to control the VSP's net profit effectively. Specifically, if the $\omega(t)$ is higher, the satisfaction in (10) will increase and vice versa. Thus, the net profit of the VSP and social welfare for type θ_j after the current learning round is completed can be respectively defined as follows:

$$\hat{\mu}_{VSP}^j(t) = \theta_j \omega(t) \lambda \sqrt{\sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \zeta_j^n(t)} - \sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \varphi_j^n(t), \quad (46)$$

and

$$\begin{aligned} \hat{\mu}_{SW}^j(t) &= \theta_j \omega(t) \lambda \sqrt{\sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \zeta_j^n(t)} \\ &\quad - \sum_{n \in \mathcal{N}(t)} \hat{\varrho}_j^n(t) \zeta_j^n(t) \xi_n(t), \forall j \in \mathcal{J}. \end{aligned} \quad (47)$$

From (46), we can observe the relationship between the net profit of the VSP and the global model economic

value (through the global model accuracy and freshness values). In particular, the earlier the VSP can obtain the high-accurate global prediction model, the more freshness and thus the higher economic value the global model can achieve (as illustrated in Fig. 2). However, in practice, at the early learning rounds, the accuracy of the global prediction model is not really high due to insufficient number of trained samples in the FL process (as illustrated in Fig. 2(c)) [3]. Moreover, as shown in Fig. 2(d), the net profit of the VSP will first increase up to the first 15 learning rounds due to high freshness value and increasing accuracy of the global model. Although the global model accuracy gets improved for the rest of learning rounds, the existence of lower freshness value will trigger lower global model economic value, and thus the net profit of the VSP also decreases gradually. Therefore, this observation provides an insightful information for the VSP to determine the best global model economic value which improves its profit in the IoV accounting for the trade-off between the global model accuracy and freshness for each learning round.

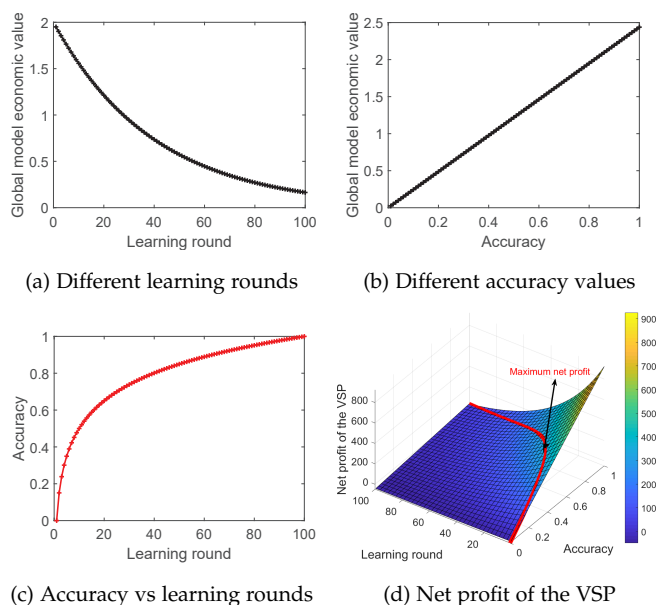


Fig. 2: The impact of global model accuracy and freshness (represented by learning rounds) on the global model economic value and the VSP's net profit.

7 PERFORMANCE EVALUATION

7.1 Dataset Preparation

We carry our experiments using the actual two road traffic datasets in the UK between 2000 and 2016 [34] to evaluate the performance of the proposed economic framework. For the first dataset, there are 275K traffic AADFs from all major roads of 190 areas/local districts in the UK that will be applied to obtain the set of significant areas. From this dataset, we divide AADFs from all major roads based on the area IDs such that each area is represented by the total AADF V_d as defined in (1). For the second dataset, it includes 1.5M traffic accidents of all major roads in the UK along with 34 features. Based on these features, we extract 7 useful

features including accident location with 416 location IDs, accident date, accident time, accident additional conditions, i.e., light, weather, and road surface conditions (as *training features*) and accident severity (as *training label*). We also convert the accident date and time into categorical features, i.e., 7-day (i.e., 1, 2, ..., 7) and 24-hour (i.e., 0, 2, ..., 23) categories, respectively.

7.2 Experiment Setup

We use *TensorFlow CPU 2.2* containing *TensorFlow Federated* in a shared cluster. We also consider 100 active SVs. i.e., $I = 100$, to compare the performance of our proposed method, i.e., proposed-LIS, with other baseline FL methods, i.e., random scheduling and round-robin scheduling [8]. For the random scheduling, the VSP selects N SVs from I active SVs randomly. For the round-robin scheduling, the VSP chooses N SVs from I active SVs sequentially in a round-robin manner. Both methods do not consider the SVs' location and information significance at each round. We also present a location-significance scheduling (referred to as proposed-LS), i.e., the VSP selects N SVs from M active SVs randomly without considering the SVs' information significance at each round, to show the impact of selected SVs' current location areas on the performance results. To this end, we do not consider an information-significance scheduling because this method may provide inaccurate and useless on-road information regardless its QoI. We then divide the active SVs into 3 categories, i.e., high-QoI SVs, i.e., h -SVs, medium-QoI SVs, i.e., m -SVs, and low-QoI SVs, i.e., l -SVs. In particular, we generate their data sizes and spatio-temporal variabilities in the following order: h -SVs > m -SVs > l -SVs. At each learning round, all SVs are randomly scattered in both significant and insignificant (which are determined by the K-means algorithm in Section. 3.1) as illustrated in Fig. 3. We also account for various values of N , i.e., 5 and 10 SVs, for each round to evaluate the proposed framework efficiency with respect to various payment budgets.

In the contract optimization, we consider one agent, i.e., the VSP, and N principals corresponding to N learning SVs. We set λ at 1.2 monetary unit (MU) per $\zeta^n(t) = 0.1, \forall n \in \mathcal{N}(t), \forall t \in \mathcal{T}$ [23]. We also set v and $\xi_n(t)$ at 2.1MU and 0.5MU per $\zeta^n(t) = 0.1, \forall n \in \mathcal{N}(t), \forall t \in \mathcal{T}$, respectively. To demonstrate various performances of the VSP, we consider 10 types, i.e., $J = 10$, with the same distribution of the types, i.e., $\rho_j(t) = 0.1$ at each round. Then, we set $B_{max}(t)$ at 125MU and 250MU [35] at each round for the cases of 5 and 10 learning SVs, respectively.

Next, we split the accident dataset into 0.8 training set and 0.2 testing set. Additionally, we consider two data distribution scenarios, i.e., i.i.d and non-i.i.d scenarios. For i.i.d scenario, it occurs when most of the learning SVs in the IoV network have visited many locations at different times frequently, e.g., taxi or long-trip public transport, and thus they can collect on-road data with various accident severity levels. In this case, we distribute the training set randomly to the h -SVs, m -SVs, and l -SVs based on their information significance. Meanwhile, for non-i.i.d scenario, it takes place when most of the learning SVs have only passed through specific limited locations most of the times,



Fig. 3: An illustration of significant (red) and insignificant (blue) areas in the UK from the dataset in [34].

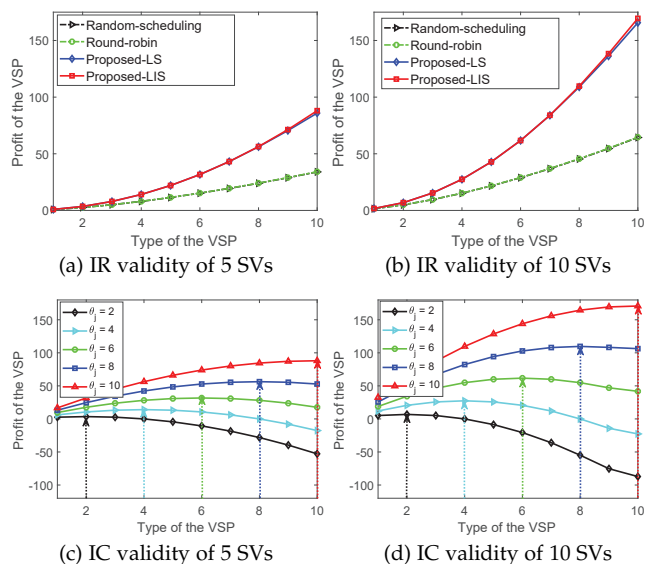


Fig. 4: The validity of IR and IC constraints for i.i.d scenario.

e.g., the school bus or city public transport, such that each learning SV has very-limited and fixed accident severity information. Specifically, we first divide the training set randomly into 3 training subsets for h -SVs, m -SVs, and l -SVs. Then, we sort each training subset according to the training label. Each sorted training subset is distributed to the corresponding number of learning SVs based on their information significance. To implement the DNN, we use two hidden layers with 128 and 64 neurons per layer and one final layer with 3 neurons (which corresponding to 3 labels). We also use the Adam optimizer with initial step size 0.01.

7.3 Learning Contract Performance

7.3.1 Common constraint validity and profit of the VSP

We first evaluate the validity of the VSP's common constraints, i.e., IR and IC constraints, considered in the contract optimization at a specific learning round, e.g., the first round in this case. Specifically, we can observe the IR constraints of the VSP for all methods in Fig. 4(a)-(b) and Fig. 5(a)-(b) for i.i.d and non-i.i.d scenarios, respectively, with various numbers of learning SVs. In this case, the profits of proposed methods coincide with each other due to the VSP's limited budget constraint for the VSP's each type. From

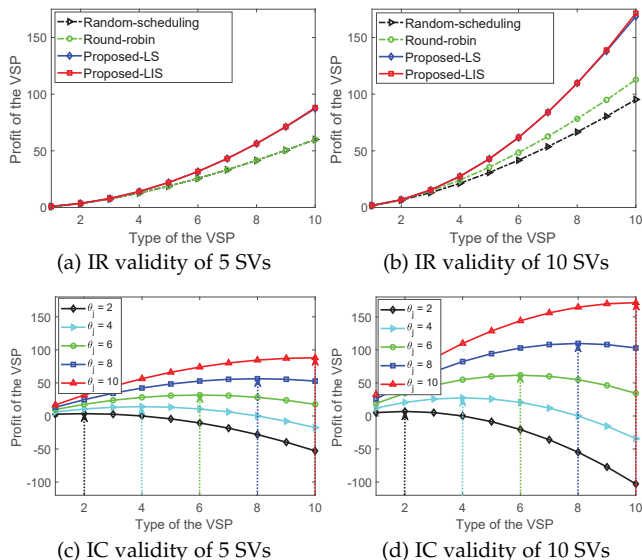


Fig. 5: The validity of IR and IC constraints for non-i.i.d scenario.

these figures, the VSP can always obtain positive profits for all types of the VSP, thereby ensuring the feasibility of the IR constraints as shown in (16). These profits monotonically increase with respect to the VSP's type such that the VSP with a higher type will obtain a higher profit. The reason is that the VSP is willing to pay more money to the learning SVs due to higher payment budget. Additionally, we can observe that more learning SVs, i.e., 10 learning SVs, will produce higher profits for the VSP up to 1.95 times, respectively, compared with that of 5 learning SVs. This is because more learning SVs can provide larger training data sizes and better information significance values for the FL process. Meanwhile, compared with other methods, the proposed-LIS can achieve higher VSP's profit up to 2.63 and 1.8 times for i.i.d and non-i.i.d scenarios, respectively. For Fig. 4(c)-(d) and Fig. 5(c)-(d), we observe that the VSP can always obtain the highest profit when it utilizes the right contract for its true type. In this way, the IC constraints for all types of the VSP are also satisfied. Particularly, the VSP with types 2, 4, 6, 8, and 10 will produce the highest profit when the suitable contracts for those types are used. As both IR and IC constraints hold, we can find the feasible contracts for all learning SVs to maximize their profits at each round.

7.3.2 The profits of SVs and social welfare of the IoV

We then analyze the total profit of learning SVs and social welfare of the IoV network at the first round as shown in Fig. 6 and Fig. 7. Similar to the VSP's profit performance, the proposed-LIS can achieve higher learning SVs' total profits up to 21 and 18.2 times for 5 and 10 learning SVs, respectively, in the i.i.d scenario as well as 5.27 and 8.93 times for 5 and 10 learning SVs, respectively, in the non-i.i.d scenario, compared with those of random and round-robin scheduling methods. The reason is that both methods are likely to train the dataset with very low data size and quality without considering the location-information significance. We also observe that although the proposed-LS can achieve the total profit of learning SVs close to that of proposed-LIS, the proposed-LIS can increase the performance gap by 1.38

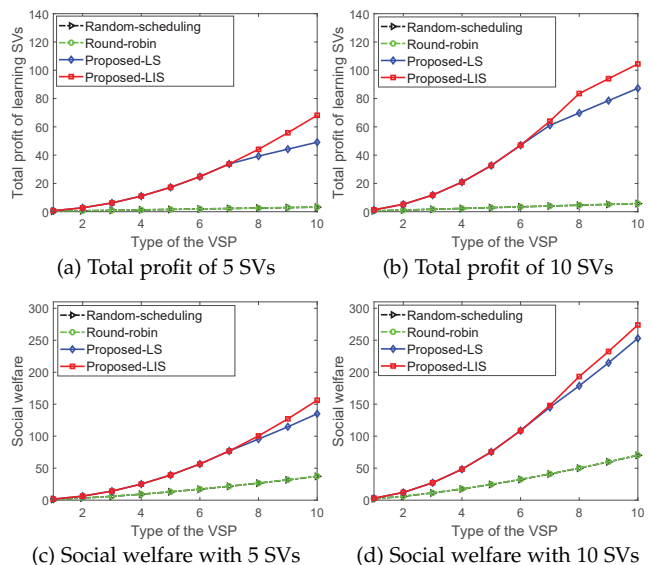


Fig. 6: The total profit of learning SVs and social welfare for i.i.d scenario.

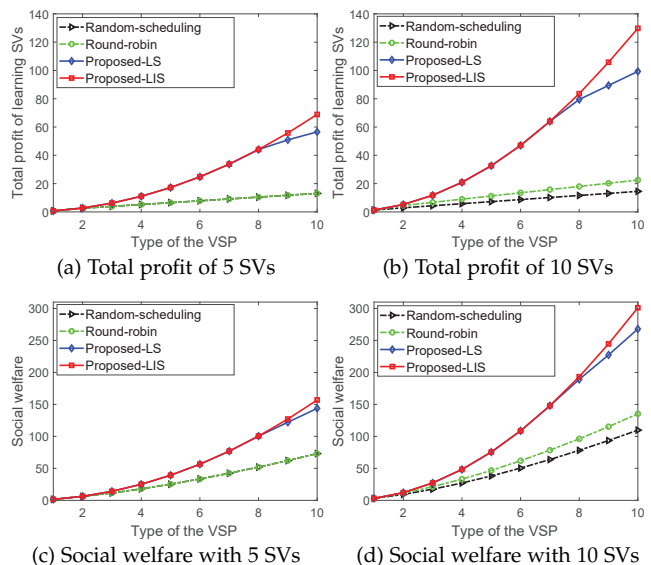


Fig. 7: The total profit of learning SVs and social welfare for non-i.i.d scenario.

and 1.31 times for i.i.d and non-i.i.d scenarios, respectively, when a higher type is employed by the VSP. The above performance aligns with the social welfare of the network in Fig. 6(c)-(d) and Fig. 7(c)-(d). Specifically, the proposed-LS and proposed-LIS can improve the social welfare up to 3.63 and 4.2 times (i.i.d scenario) as well as 2.44 and 2.74 times (non-i.i.d scenario), respectively, compared with those of random and round-robin scheduling methods.

7.4 Dynamic Federated Learning Accuracy Performance

7.4.1 i.i.d scenario

We first investigate the global model accuracy for the i.i.d scenario when the number of learning SVs, i.e., N , increases and the VSP has type 10. In particular, when 5 SVs are used to perform the FL algorithm for each round as shown in

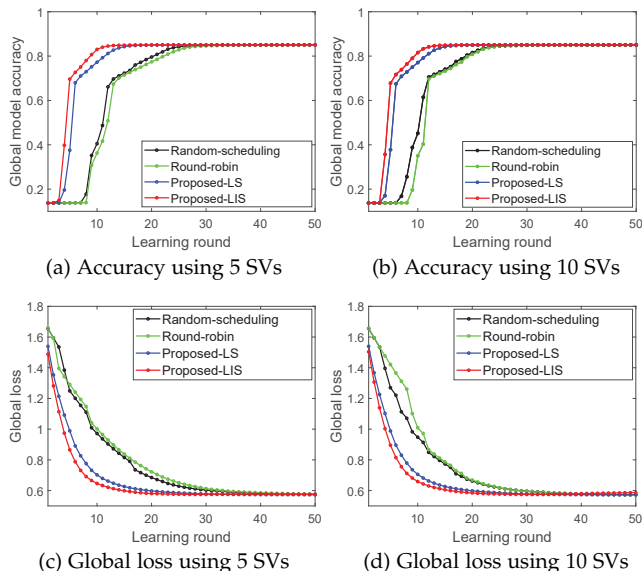


Fig. 8: The performance of proposed FL with SV selection using various learning SVs with i.i.d datasets.

Fig. 8(a), the proposed-LIS can converge 57% (or 17 rounds) earlier than those of the random scheduling and round-robin with an accuracy of 85%. Furthermore, the proposed-LIS can achieve the convergence 28% (or 5 rounds) faster than that of the proposed-LS at the same accuracy level. This is because the proposed-LIS only selects an SV with high information significance within the significant areas. Meanwhile, the baseline FL methods select the learning SVs from the total active SVs without considering the location and information significance-based SV selection. In this case, there exists high probability that they only use l -SVs most of the learning rounds, leading to a slow convergence. For the proposed-LS, it only selects SVs in the significant areas without information significance consideration. As a result, its convergence accuracy will be slightly delayed. These accuracy performances align with the global loss ones in Fig. 8(c) where the proposed-LIS can reach the fastest minimum convergence compared with other learning methods.

The interesting point of accuracy performance can be observed when 10 learning SVs are considered in Fig. 8(b). Specifically, the learning methods other than the proposed-LIS can improve the accuracy convergence speed by 18% (or 6 rounds) because of the existence of more accurate/meaningful local models from more learning SVs. Nevertheless, this observation does not apply for the proposed-LIS. The reason is that when each learning SV contains i.i.d local dataset with high information significance, the aggregation of all local models at the VSP will generate the same model accuracy compared with the condition when one h -SV is used. As such, the additional datasets with the same high information significance from other SVs do not further improve the convergence speed [11].

7.4.2 Non-i.i.d scenario

In contrast to the i.i.d scenario, the convergence speed for the proposed-LIS gets faster when more learning SVs are selected at each round for the non-i.i.d scenario. As shown in Fig. 9(a), all FL methods suffer from the fluctuated/unstable

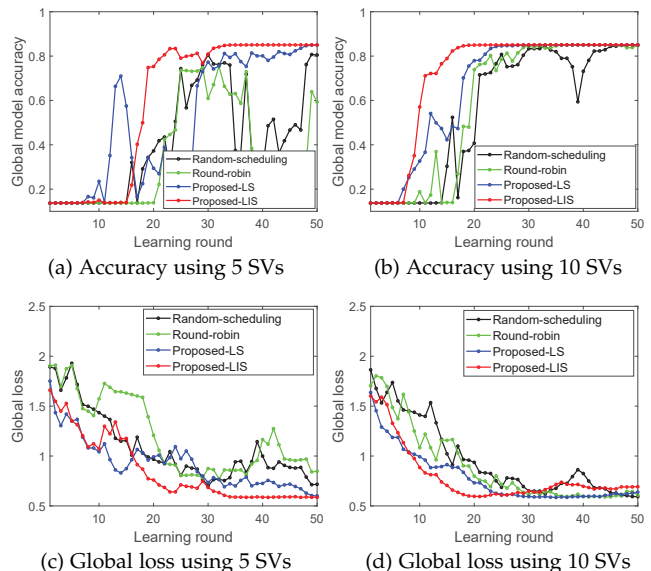


Fig. 9: The performance of proposed FL with SV selection using various learning SVs with non-i.i.d datasets.

learning performance when the VSP utilizes 5 learning SVs (due to biased local dataset from each active SV [11]). Nevertheless, the proposed-LIS eventually can achieve the convergence within 30 learning rounds with accuracy level 85%. Specifically, the proposed-LIS still outperforms the proposed-LS up to 17.5% (or 7 rounds) in terms of convergence speed due to the selection guarantee of h -SVs. Moreover, other baseline methods cannot even reach the convergence because they likely to select l -SVs with very biased local datasets. Similar to the i.i.d scenario, the global loss performance in Fig. 9(c) aligns with the accuracy one where the proposed-LS and proposed-LIS can reach the fastest minimum convergence compared with those of the random scheduling and round-robin methods.

When more learning SVs are selected by the VSP, the accuracy performances for all FL methods can be improved. As observed in Fig. 9(b), the proposed-LIS can speed up the accuracy convergence by 45% (or 15 rounds) when the VSP schedules 10 learning SVs for each round. The reason is that the proposed-LIS can train more h -SVs and m -SVs with less biased local datasets, thereby reducing the unfairness among different SVs. Furthermore, the proposed-LIS can achieve the convergence speed up to 62% (or 29 rounds) faster than those of other methods when 10 learning SVs are used, respectively. These results can provide useful information for the VSP in practice to determine the best SV selection method for the FL process in terms of stability, robustness, and flexibility.

7.5 Relationship Between Contract and Federated Learning Performance

This section discusses how the global model accuracy and the global model freshness obtained from the learning process at each round can influence the profits of the VSP (with type 10) and learning SVs as well as the social welfare of the network obtained from the contract optimization. As observed in Fig. 10(a)-(b) when i.i.d datasets are used, the net profit of the VSP (when using the proposed-LS and

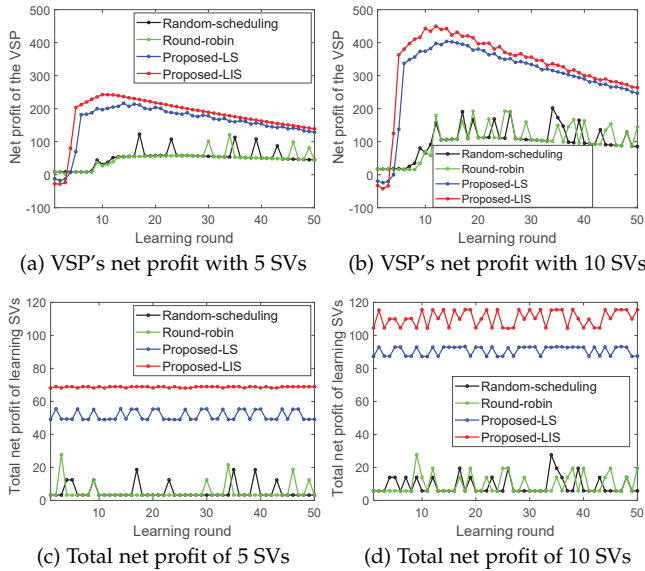


Fig. 10: Net profits of the VSP and learning SVs for all learning rounds when i.i.d datasets are trained.

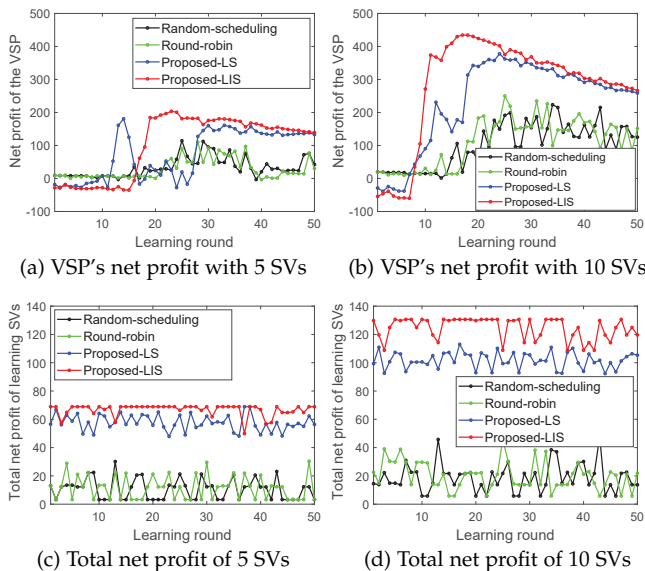


Fig. 11: Net profits of the VSP and learning SVs for all learning rounds when non-i.i.d datasets are trained.

proposed-LIS) first suffers from negative profit during the first three learning rounds for all cases. The reason is that although the VSP can obtain high information significance values from the learning SVs at these rounds, the obtained global model accuracy cannot compensate the satisfaction function due to the higher cost, i.e., high payment for the learning SVs with high QoI. As the learning round increases, the global model accuracy start to improve significantly by 60%, thereby enhancing the net profit of the VSP (with the maximum profit between round 10 and 15). Due to the global model freshness impact and less accuracy improvement at further learning rounds, the profit of the VSP gradually decreases. Nonetheless, our proposed-LS and proposed-LIS can still outperform other methods up to 7.25 and 8.92 times in terms of the VSP's net profit, respectively.

For the total net profit of learning SVs, the performance

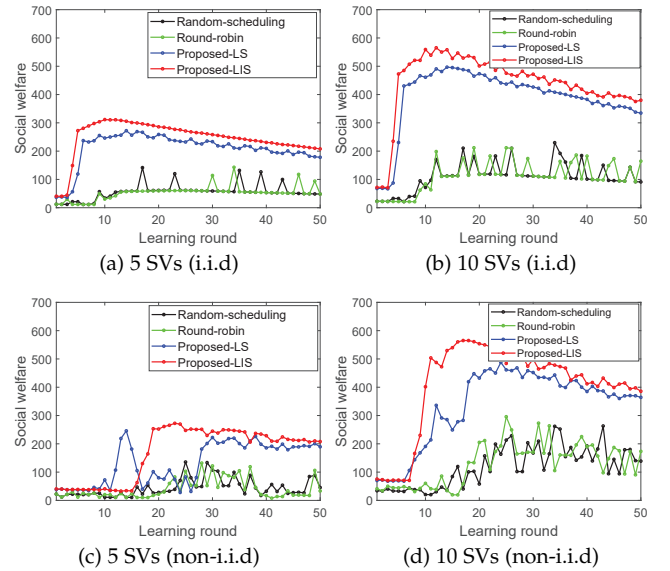


Fig. 12: Final social welfare for all learning rounds when i.i.d and non-i.i.d datasets are trained.

is not affected by the global model accuracy and freshness. Alternatively, it remains the same as the one obtained from the contract optimization. The reason is that the profits of learning SVs depend only on their current round's trained datasets with certain information significance values, and computation as well as memory costs at the learning SVs. As can be seen in Fig. 10(c)-(d), our proposed-LIS can achieve the highest total profit performance of learning SVs for most rounds due to the use of more h -SVs. Similar trends with more fluctuated profit performance can be observed in Fig. 11 when non-i.i.d datasets are used for the cases of 5 and 10 learning SVs. However, the use of 5 learning SVs in Fig. 11(a) and Fig. 11(c) triggers low net profits for both VSP and learning SVs of the proposed-LS and proposed-LIS during the first 16 and 27 learning rounds, respectively (due to the unstable accuracy performance in Fig. 9(a)). To support the performances in Fig. 10 and Fig. 11, we show the final social welfare for both i.i.d and non-i.i.d scenarios in Fig. 12. Specifically, the trend of social welfare follows the one of the VSP's net profit with increased profit values due to addition of learning SVs' total profit. In this case, our proposed-LIS can achieve the social welfare up to 24.5 and 27.2 times for i.i.d and non-i.i.d scenarios, respectively, compared with those of other methods. To this end, we can summarize that the i.i.d scenario can provide more stable economic benefit performance throughout the learning rounds compared with that of the non-i.i.d scenario. However, the non-i.i.d scenario will provide more practical economic benefit performance because of the unique and biased dataset (generated from diverse mobile activities) at each learning SV in the practical IoV network.

8 CONCLUSION

In this paper, we have proposed the novel economic framework for the IoV network to maximize the profits for the VSP and learning SVs in the dynamic FL process. Specifically, we have introduced the dynamic SV selection method to determine a set of active SVs based on the significance of

their locations and information at each learning round. From the set of selected SVs, each SV can first collect on-road data and then offer a payment contract to the VSP with respect to its QoI. In this case, we have developed the MPOA-based contract optimization problem under the competition of learning SVs and the VSP's common constraints (i.e., the IR and IC constraints) as well as incomplete information of payment budget. To find the optimal contracts, we have transformed the problem into the equivalent low-complexity problem and then implemented the iterative contract algorithm which can achieve the equilibrium solution for all the learning SVs at each learning round. We have also analyzed the convergence of proposed FL and investigated on how the accuracy and freshness of the global prediction model obtained by the FL process at each round affect the net profits of the VSP and learning SVs as well as the social welfare of the network. Through the experimental results, we have shown that our proposed framework can significantly improve the FL convergence speed, profits of the VSP and learning SVs, and the social welfare of the network compared with other baseline FL methods.

REFERENCES

- [1] Internet of Vehicle Market Outlook: 2024, May 2018. Available Online: <https://www.alliedmarketresearch.com/internet-of-vehicles-market>. Last Accessed on: September 2020.
- [2] W. Nie, *et al.*, "A quality-oriented data collection scheme in vehicular sensor networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5570-5584, Jul. 2018.
- [3] W. Y. B. Lim, *et al.*, "Federated learning in mobile edge networks: a comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031-2063, Apr. 2020.
- [4] S. Samarakoon, M. Bennis, W. Saad, and M. Debbah, "Federated learning for ultra-reliable low-latency V2V communications," in *IEEE GLOBECOM*, Dec. 2018, pp. 1-7.
- [5] S. R. Pokhrel and J. Choi, "Improving TCP performance over WiFi for internet of vehicles: a federated learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6798-6802, Jun. 2020.
- [6] Y. M. Saputra, D. Nguyen, H. T. Dinh, T. X. Vu, E. Dutkiewicz, and S. Chatzinotas, "Federated learning meets contract theory: economic-efficiency framework for electric vehicle networks," *IEEE Trans. Mobile Comput.*, Early Access, Dec. 2020.
- [7] Z. Song, *et al.*, "QoI-aware multitask-oriented dynamic participant selection with budget constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4618-4632, Nov. 2014.
- [8] H. H. Yang, Z. Liu, T. Q. S. Quek, and H. V. Poor, "Scheduling policies for federated learning in wireless networks," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 317-333, Jan. 2020.
- [9] C. Xie, S. Koyejo, and I. Gupta, "Asynchronous federated optimization," arXiv:1903.03934 [cs.DC], Sept. 2019.
- [10] T. Li, *et al.*, "Federated optimization in heterogeneous networks," arXiv:1812.06127 [cs.LG], Apr. 2020.
- [11] M. Amiri, D. Gunduz, S. R. Kulkarni, and H. V. Poor, "Convergence of update aware device scheduling for federated learning at the wireless edge," arXiv:2001.10402 [cs.IT], May 2020.
- [12] Road Traffic Estimates. Available Online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524848/annual-methodology-note.pdf. Last Accessed: Sept. 2020.
- [13] Z. Zheng, *et al.*, "A budget feasible incentive mechanism for weighted coverage maximization in mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 16, no. 9, pp. 2392-2407, Sept. 2017.
- [14] R. Hussain and S. Zeadally, "Autonomous cars: research results, issues, and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1275-1313, Secondquarter 2019.
- [15] Y. Zhan, P. Li, Z. Qu, D. Zeng, and S. Guo, "A learning-based incentive mechanism for federated Learning," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6360-6368, Jul. 2020.
- [16] L. U. Khan, *et al.*, "Federated learning for edge networks: resource optimization and incentive mechanism," *IEEE Commun. Mag.*, vol. 58, no. 10, pp. 88-93, Oct. 2020.
- [17] Y. Sarikaya and O. Ercetin, "Motivating workers in federated learning: a stackelberg game perspective," *IEEE Networking Letters*, vol. 2, no. 1, pp. 23-27, Mar. 2020.
- [18] P. Bolton and M. Dewatripont, *Contract Theory*. Cambridge, MA, USA: MIT Press, 2005.
- [19] B. D. Bernheim and M. D. Whinston, "Common agency," *Econometrica*, vol. 54, no. 4, pp. 923-942, Jul. 1986.
- [20] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10700-10714, Dec. 2019.
- [21] 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. 23920-23935, Feb. 2020.
- [22] Service requirements for V2X services, Mar. 2017. Available Online: https://www.etsi.org/deliver/etsi_ts/122100_122199/122185/14.03.00_60/ts_122185v140300p.pdf. Last Accessed on: December 2020.
- [23] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu, "Privacy or utility in data collection? a contract theoretic approach" *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 7, pp. 1256-1269, Oct. 2015.
- [24] G. P. Corser, H. Fu, and A. Banihani, "Evaluating location privacy in vehicular communications and applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 9, pp. 2658-2667, Sep. 2016.
- [25] J. Lim, *et al.*, "Preserving location privacy of connected vehicles with highly accurate location updates," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 540-543, Mar. 2017.
- [26] P. S. Bradley, K. P. Bennett, and A. Demiriz, "Constrained K-means clustering," *Microsoft Research MSR-TR-2000-65*, May 2000.
- [27] A. L. Custodio, H. Rocha, and L. N. Vicente, "Incorporating minimum Frobenius norm models in direct search," *Comput. Optim. Appl.*, vol. 46, no. 2, pp. 265-278, Jun. 2010.
- [28] P. A. Samuelson and W. D. Nordhaus, *Microeconomics*, 18th ed. Boston, MA, USA: McGraw-Hill, 2005.
- [29] Z. Zhu, S. Lambotharan, W. H. Chin, and Z. Fan, "A game theoretic optimization framework for home demand management incorporating local energy resources," *IEEE Trans. Ind. Informat.*, vol. 11, no. 2, pp. 353-362, Apr. 2015.
- [30] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: a survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224-2287, Thirdquarter 2019.
- [31] D. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *ICLR 2015*, May 2015, pp. 1-15.
- [32] X. Chen, *et al.*, "Age of information aware radio resource management in vehicular networks: a proactive deep reinforcement learning perspective," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2268-2281, Apr. 2020.
- [33] A. A. Abdellatif, C. F. Chiasserini, and F. Malandrino, "Active learning-based classification in automated connected vehicles," in *IEEE INFOCOM WKSHPs*, Jul. 2020, pp. 598-603.
- [34] UK Traffic Accidents Between 2005 and 2014, Sept. 2017. Available Online: <https://www.kaggle.com/daveianhickey/2000-16-traffic-flow-england-scotland-wales>. Last Accessed on September 2020.
- [35] D. Zhao, X. Li, and H. Ma, "Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 647-661, Apr. 2016.
- [36] J. Fraysse, "Common agency: existence of an equilibrium in the case of two outcomes," *Econometrica*, vol. 61, no. 5, pp. 1225-1229, Sept. 1993.
- [37] H. H. Yang, A. Arafat, T. Q. S. Quek and H. Vincent Poor, "Age-based scheduling policy for federated learning in mobile edge networks," in *IEEE ICASSP*, May 2020, pp. 8743-8747.
- [38] W. Y. B. Lim, *et al.*, "When information freshness meets service latency in federated learning: a task-aware incentive scheme for smart industries," *IEEE Trans. Ind. Informat.*, Early Access, Dec. 2020.
- [39] S. Durand and B. Gaujal, "Complexity and optimality of the best response algorithm in random potential games," in *SAGT*, Sept. 2016, pp. 40-51.

APPENDIX A PROOF OF LEMMA 1

We first prove that $\zeta_j(t) \geq \zeta_{j^*}(t)$ if and only if $\theta_j \geq \theta_{j^*}$ using the IC constraint in (17), which is

$$\begin{aligned} & \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) \geq \\ & \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)). \end{aligned} \quad (48)$$

If the VSP has type j^* , then we obtain

$$\begin{aligned} & \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)) \geq \\ & \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)), \end{aligned} \quad (49)$$

where $j \neq j^*$ and $j, j^* \in \mathcal{J}$. From (48) and (49), we can obtain

$$\begin{aligned} & \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) + \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) \geq \\ & \theta_j S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) + \theta_{j^*} S(\hat{\rho}_j(t), \zeta_j(t)), \\ & \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - \theta_{j^*} S(\hat{\rho}_j(t), \zeta_j(t)) \geq \\ & \theta_j S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)), \\ & S(\hat{\rho}_j(t), \zeta_j(t))(\theta_j - \theta_{j^*}) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))(\theta_j - \theta_{j^*}). \end{aligned} \quad (50)$$

By removing $(\theta_j - \theta_{j^*})$, we have $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$. Since the satisfaction function in (10) is monotonically increasing in $\zeta_j(t)$, thus if $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$, then $\zeta_j(t) \geq \zeta_{j^*}(t)$ given $\hat{\rho}_j = \hat{\rho}_{j^*}$.

Then, we prove that $\theta_j \geq \theta_{j^*}$ if and only if $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$. From (48)-(50), we have

$$\begin{aligned} & \theta_j (S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))) \geq \\ & \theta_{j^*} (S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))). \end{aligned} \quad (51)$$

By eliminating both $S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$, we obtain $\theta_j \geq \theta_{j^*}$ considering $S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) \geq 0$ since $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$. As a result, we prove that if $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$, then $\theta_j \geq \theta_{j^*}$. This concludes the proof.

APPENDIX B PROOF OF PROPOSITION 2

Based on (48) and (49), we can derive the following expressions:

$$C(\hat{\rho}_j(t), \varphi_j(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)) \leq \theta_j (S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))), \quad (52)$$

$$\begin{aligned} & C(\hat{\rho}_j(t), \varphi_j(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)) \geq \\ & \theta_{j^*} (S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))). \end{aligned} \quad (53)$$

Given $\hat{\rho}_j = \hat{\rho}_{j^*}$, if $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$, then we can obtain $C(\hat{\rho}_j(t), \varphi_j(t)) \geq C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t))$ from (53), and thus $\varphi_j(t) \geq \varphi_{j^*}(t)$ since the cost function (11) is monotonically increasing in $\varphi_j(t)$. Moreover, if $C(\hat{\rho}_j(t), \varphi_j(t)) \geq C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t))$, then we can obtain from (52) that $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$, and thus $\zeta_j(t) \geq \zeta_{j^*}(t)$.

APPENDIX C PROOF OF PROPOSITION 3

From Lemma 1 and Proposition 2, we have $\zeta_j(t) \geq \zeta_{j^*}(t)$ and $\varphi_j(t) \geq \varphi_{j^*}(t)$. If $\theta_j \geq \theta_{j^*}$, then

$$\begin{aligned} & \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) \geq \\ & \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)) \geq \\ & \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)) \geq 0. \end{aligned} \quad (54)$$

Thus, $\theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) \geq \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t))$ when $\theta_j \geq \theta_{j^*}$.

APPENDIX D PROOF OF LEMMA 2

We first define downward ICs (DICs) and upward ICs (UICs) as the IC constraints between the VSP with type indices j and j^* , $\forall j^* \in \{1, \dots, j-1\}$, and between the VSP with type indices j and j^* , $\forall j^* \in \{j+1, \dots, J\}$, respectively, in which

$$\begin{aligned} & \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) \geq \\ & \theta_{j^*} S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t)) - C(\hat{\rho}_{j^*}(t), \varphi_{j^*}(t)). \end{aligned} \quad (55)$$

Then, we prove that DICs can be simplified into two consecutive types called local DICs (LDICs), which are the IC constraints between the VSP with type indices j and $j-1$. Consider that $\theta_{j-1} < \theta_j < \theta_{j+1}$, $j \in \{2, \dots, J-1\}$, we can derive

$$\begin{aligned} & \theta_{j+1} S(\hat{\rho}_{j+1}(t), \zeta_{j+1}(t)) - C(\hat{\rho}_{j+1}(t), \varphi_{j+1}(t)) \geq \\ & \theta_{j+1} S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)), \end{aligned} \quad (56)$$

$$\begin{aligned} & \theta_j S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) \geq \\ & \theta_j S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t)) - C(\hat{\rho}_{j-1}(t), \varphi_{j-1}(t)). \end{aligned} \quad (57)$$

From Lemma 1, we have proven that $S(\hat{\rho}_j(t), \zeta_j(t)) \geq S(\hat{\rho}_{j^*}(t), \zeta_{j^*}(t))$ if $\theta_j \geq \theta_{j^*}$. Then, we can obtain

$$\begin{aligned} & \theta_{j+1} (S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t))) \geq \\ & \theta_j (S(\hat{\rho}_j(t), \zeta_j(t)) - S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t))) \geq \\ & C(\hat{\rho}_j(t), \varphi_j(t)) - C(\hat{\rho}_{j-1}(t), \varphi_{j-1}(t)), \end{aligned} \quad (58)$$

and

$$\begin{aligned} & \theta_{j+1} S(\hat{\rho}_{j+1}(t), \zeta_{j+1}(t)) - C(\hat{\rho}_{j+1}(t), \varphi_{j+1}(t)) \geq \\ & \theta_{j+1} S(\hat{\rho}_j(t), \zeta_j(t)) - C(\hat{\rho}_j(t), \varphi_j(t)) \geq \\ & \theta_{j+1} S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t)) - C(\hat{\rho}_{j-1}(t), \varphi_{j-1}(t)). \end{aligned} \quad (59)$$

As $\theta_{j+1} S(\hat{\rho}_{j+1}(t), \zeta_{j+1}(t)) - C(\hat{\rho}_{j+1}(t), \varphi_{j+1}(t)) \geq \theta_{j+1} S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t)) - C(\hat{\rho}_{j-1}(t), \varphi_{j-1}(t))$, it can be expanded from the VSP with type index $j-1$ to type index 1 to show that all the DICs satisfy, i.e.,

$$\begin{aligned} & \theta_{j+1} S(\hat{\rho}_{j+1}(t), \zeta_{j+1}(t)) - C(\hat{\rho}_{j+1}(t), \varphi_{j+1}(t)) \geq \\ & \theta_{j+1} S(\hat{\rho}_{j-1}(t), \zeta_{j-1}(t)) - C(\hat{\rho}_{j-1}(t), \varphi_{j-1}(t)) \geq \dots \geq \\ & \theta_{j+1} S(\hat{\rho}_1(t), \zeta_1(t)) - C(\hat{\rho}_1(t), \varphi_1(t)), \forall j \in \{1, \dots, J-1\}. \end{aligned} \quad (60)$$

As a result, we can conclude that by using the LDICs, all the DICs satisfy, and then the DICs can be reduced. Likewise, we can prove that all the UICs hold considering the local UICs (LUICs) and Lemma 1.

APPENDIX E

PROOF OF THEOREM 1

Consider \mathbb{C} has $(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t))$ and $(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$ such that

$$(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)) > (\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)). \quad (61)$$

Given that $\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) = \max \Phi_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$, we can show that the choice $\hat{\Phi}^n$ in Φ^n at SV- n is increasing, which is

$$\hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)) \geq \hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)). \quad (62)$$

Using $(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)) > (\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$ and $\hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)) < \hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$ to derive a contradiction, we can obtain

$$\begin{aligned} & \mu_n \left(\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)), \zeta_{(\varsigma+2)}^{-n}(t), \varphi_{(\varsigma+2)}^{-n}(t) \right) + \\ & \mu_n \left(\hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)), \zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t) \right) \geq \\ & \mu_n \left(\hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)), \zeta_{(\varsigma+2)}^{-n}(t), \varphi_{(\varsigma+2)}^{-n}(t) \right) + \\ & \mu_n \left(\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)), \zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t) \right). \quad (63) \end{aligned}$$

Based on the definition of $\Phi_{(\varsigma+1)}^n$, $\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) \in \Phi_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$ describes that

$$\begin{aligned} & \mu_n \left(\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)), \zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t) \right) \geq \\ & \mu_n \left(\hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)), \zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t) \right). \quad (64) \end{aligned}$$

Then, from (63) and (64), we have

$$\begin{aligned} & \mu_n \left(\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)), \zeta_{(\varsigma+2)}^{-n}(t), \varphi_{(\varsigma+2)}^{-n}(t) \right) \geq \\ & \mu_n \left(\hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)), \zeta_{(\varsigma+2)}^{-n}(t), \varphi_{(\varsigma+2)}^{-n}(t) \right), \quad (65) \end{aligned}$$

and thus $\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) \in \Phi_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t))$. In other words, it can be noticed that $\hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t)) > \hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t))$. Based on the definition of $\hat{\Phi}_{(\varsigma+1)}^n$, it can be observed that $\hat{\Phi}_{(\varsigma+2)}^n(\zeta_{(\varsigma+1)}^{-n}(t), \varphi_{(\varsigma+1)}^{-n}(t)) \geq \hat{\Phi}_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$ which contradicts with the above condition. As a result, $\hat{\Phi}^n$ is an increasing function.

Using the increasing function of $\hat{\Phi}^n$, the VSP can update the contract of SV- n at iteration $\varsigma + 1$ if the condition in (34) is satisfied. To this end, the use of γ can be denoted as the optimality tolerance to stop the iterative algorithm. In particular, when all the SVs in $\mathcal{N}(t)$ do not hold the condition in (34) at iteration $\varsigma = \varkappa$, we have $(\zeta_{(\varkappa+1)}^n(t), \varphi_{(\varkappa+1)}^n(t)) = (\zeta_{(\varkappa)}^n(t), \varphi_{(\varkappa)}^n(t))$, $\forall n \in \mathcal{N}(t)$. Hence, for the rest of ς

values starting from \varkappa , the iterative algorithm provides the same $\mu_n(\zeta_{(\varsigma+1)}^n(t), \varphi_{(\varsigma+1)}^n(t), \zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$, $\forall n \in \mathcal{N}(t)$. It means that the iterative algorithm converges under γ .

APPENDIX F

PROOF OF THEOREM 2

To prove that the iterative algorithm can converge to the equilibrium solution, we first show that an equilibrium exists by finding a fixed point of Φ . Suppose that \mathbb{C}^* contains $(\zeta(t), \varphi(t)) \in \mathbb{C}$. This \mathbb{C}^* is a non-empty contract space because $\hat{\Phi}(\min \mathbb{C}) \geq \min \mathbb{C}$, $\hat{\Phi} \in \Phi$, and thus $\hat{\Phi}(\max \mathbb{C}^*) \geq \max \mathbb{C}^*$. Since $\hat{\Phi}$ is an increasing function, we obtain $\hat{\Phi}(\hat{\Phi}(\max \mathbb{C}^*)) \geq \hat{\Phi}(\max \mathbb{C}^*)$, and thus $\hat{\Phi}(\max \mathbb{C}^*) \in \mathbb{C}^*$. Then, we have $\hat{\Phi}(\max \mathbb{C}^*) \leq \max \mathbb{C}^*$ and obtain $\hat{\Phi}(\max \mathbb{C}^*) = \max \mathbb{C}^*$. Consequently, $\max \mathbb{C}^*$ is a fixed point of Φ which contains $(\hat{\zeta}(t), \hat{\varphi}(t))$ and indicates that the equilibrium exists. Since the profits of all SV- n , $\forall n \in \mathcal{N}(t)$, follow the increasing function until they converge, no SVs can further improve the profits when their optimal contracts are obtained. In this case, it implies that the iterative algorithm must converge to the equilibrium contract solution $(\hat{\zeta}(t), \hat{\varphi}(t))$, to guarantee that there exists no such SV- n can improve its profit, i.e., $\mu_n(\hat{\zeta}^n(t), \hat{\varphi}^n(t), \hat{\zeta}^{-n}(t), \hat{\varphi}^{-n}(t))$, unilaterally [19], [36].

APPENDIX G

PROOF OF THEOREM 3

Let $n^\dagger \subset N$ denote the number of SVs in $\mathcal{N}(t)$ which complete the best response optimizations at iteration $\varsigma + 1$, i.e., $\Phi_{(\varsigma+1)}^n(\zeta_{(\varsigma)}^{-n}(t), \varphi_{(\varsigma)}^{-n}(t))$, without modifying their current contracts. We also define $U \in [0, 1]$ as the current normalized candidate profit. Given that R is the total number of contract policies, n^\dagger -th SV obtains $R - 1$ new contract policies in \mathbb{C}_{n^\dagger} (whose normalized candidate profits need to be compared with U) when the SV computes its $\Phi_{(\varsigma+1)}^{n^\dagger}(\zeta_{(\varsigma)}^{-n^\dagger}(t), \varphi_{(\varsigma)}^{-n^\dagger}(t))$. In this case, the probability that all the $R - 1$ policies cannot drive the next normalized candidate profit higher than U is U^{R-1} . Thus, the $(n^\dagger + 1)$ -th SV can try to find its $\Phi_{(\varsigma+1)}^{n^\dagger+1}(\zeta_{(\varsigma)}^{-(n^\dagger+1)}(t), \varphi_{(\varsigma)}^{-(n^\dagger+1)}(t))$. In contrast, the probability that one of the $R - 1$ policies is the best response can be defined as $1 - U^{R-1}$. As such, the normalized candidate profit changes from U to a value larger than $\hat{\mu}_{n^\dagger}(\zeta_{(\varsigma)}^{n^\dagger}(t), \varphi_{(\varsigma)}^{n^\dagger}(t), \zeta_{(\varsigma)}^{-n^\dagger}(t), \varphi_{(\varsigma)}^{-n^\dagger}(t)) + \gamma$ with the probability $1 - \tilde{\mu}_{n^\dagger}^{R-1}$, where $\tilde{\mu}_{n^\dagger} = \hat{\mu}_{n^\dagger} + \gamma$, and $\hat{\mu}_{n^\dagger} \in [0, 1]$ is the normalized current profit of SV- n^\dagger (we simplify the form of $\hat{\mu}_{n^\dagger}(\cdot)$ into $\hat{\mu}_{n^\dagger}$). At this state, we set the number of SVs in $\mathcal{N}(t)$ that have finished the best response back to 1.

Utilizing the Markov chain, let $\vartheta(U, n^\dagger)$ to be the number of steps of the Algorithm 1, when n^\dagger SVs have finished the best response (without changing the contracts with normal-

ized candidate profit U) before reaching the convergence. Then, we obtain

$$\begin{aligned} \vartheta(U, n^\dagger) &= U^{R-1}\vartheta(U, n^\dagger + 1) \\ &+ \int_U^1 (R-1)\tilde{\mu}_{n^\dagger}^{R-2}(\vartheta(\tilde{\mu}_{n^\dagger}, 1) + 1)d\tilde{\mu}_{n^\dagger}, \end{aligned} \quad (66)$$

where the first component specifies that the $(n^\dagger + 1)$ -th SV does not modify its contract, i.e., $\vartheta(U, n^\dagger) = \vartheta(U, n^\dagger + 1)$, with probability U^{R-1} . Additionally, the second component represents that the $(n^\dagger + 1)$ -th SV finds a new best response with normalized candidate profit $\tilde{\mu}_{n^\dagger}$ and probability density $(R-1)\tilde{\mu}_{n^\dagger}^{R-2}$, and thus one additional step is produced, i.e., $\vartheta(U, n^\dagger) = \vartheta(\tilde{\mu}_{n^\dagger}, 1) + 1$. Given the boundary conditions $\vartheta(U, N) = 0$ and $\vartheta(1, n^\dagger) = 0$ for all U and n^\dagger , respectively, and $\Xi(U) = \int_U^1 (R-1)\tilde{\mu}_{n^\dagger}^{R-2}(\vartheta(\tilde{\mu}_{n^\dagger}, 1) + 1)d\tilde{\mu}_{n^\dagger}$, we have

$$\begin{cases} \vartheta(U, 1) &= U^{R-1}\vartheta(U, 2) + \Xi(U), \\ \vdots &= \vdots \\ \vartheta(U, N-2) &= U^{R-1}\vartheta(U, N-1) + \Xi(U), \\ \vartheta(U, N-1) &= \Xi(U). \end{cases} \quad (67)$$

From (67), we can obtain that $\vartheta(U, 1) = \Xi(U)\Pi(U)$, where $\Pi(U) = 1 + U^{R-1} + \dots + U^{(R-1)(N-2)}$. (68)

As such, the differential equation of $\vartheta(U, 1)$ with respect to U is
$$\frac{d\vartheta(U, 1)}{dU} + \left((R-1)U^{R-2}\Pi(U) - \frac{1}{\Pi(U)} \frac{d\Pi(U)}{dU} \right) \vartheta(U, 1) = - (R-1)U^{R-2}\Pi(U). \quad (69)$$

Given that the boundary condition $\vartheta(1, 1) = 0$ and

$$\begin{aligned} \Omega(U) &= \int_0^U \left((R-1)\tilde{\mu}_1^{R-2}\Pi(\tilde{\mu}_1) - \frac{1}{\Pi(\tilde{\mu}_1)} \frac{d\Pi(\tilde{\mu}_1)}{d\tilde{\mu}_1} \right) d\tilde{\mu}_1 \\ &= \int_0^U (R-1)\tilde{\mu}_1^{R-2}\Pi(\tilde{\mu}_1)d\tilde{\mu}_1 - \log(\Pi(U)), \end{aligned} \quad (70)$$

the function $\vartheta(U, 1)$ becomes $\vartheta(U, 1) = e^{-\Omega(U)} \int_U^1 (R-1)\tilde{\mu}_1^{R-2}\Pi(\tilde{\mu}_1)e^{\Omega(\tilde{\mu}_1)}d\tilde{\mu}_1$. As $\vartheta(U, 1)$ is decreasing in U , the number of steps is upper-bounded by $\vartheta(0, 1)$. Specifically, considering $\Omega(0) = 0$ from (70) and $\Pi(0) = 1$ from (68), then

$$\begin{aligned} \vartheta(0, 1) &= \int_0^1 (R-1)\tilde{\mu}_1^{R-2} e^{\sum_{n=0}^{N-2} \frac{\tilde{\mu}_1^{(R-1)(n+1)}}{n+1}} d\tilde{\mu}_1 \\ &= \int_0^1 e^{\sum_{n=0}^{N-2} \frac{\tilde{\mu}_1^{(R-1)(n+1)}}{n+1}} d\tilde{\mu}_1^{(R-1)} \\ &= \int_0^{1-\frac{1}{N}} e^{\sum_{n=1}^{N-1} \frac{\tilde{\mu}_1^{(R-1)n}}{n}} d\tilde{\mu}_1^{(R-1)} + \\ &\int_{1-\frac{1}{N}}^1 e^{\sum_{n=1}^{N-1} \frac{\tilde{\mu}_1^{(R-1)n}}{n}} d\tilde{\mu}_1^{(R-1)} \quad (71) \\ &\leq \int_0^{1-\frac{1}{N}} e^{\sum_{n=1}^{\infty} \frac{\tilde{\mu}_1^{(R-1)n}}{n}} d\tilde{\mu}_1^{(R-1)} + \frac{1}{N} e^{\sum_{n=1}^{N-1} \frac{1}{n}} \\ &= \int_0^{1-\frac{1}{N}} \frac{d\tilde{\mu}_1^{(R-1)}}{1 - \tilde{\mu}_1^{(R-1)}} + e^z + O(1/N) \\ &= \log(N) + e^z + O(1/N), \end{aligned}$$

where $z \approx 0.5772$ is the Euler constant. From (71), we can see that the number of steps of the Algorithm 1 is bounded above to the polynomial complexity $\log(N) + e^z + O(1/N)$ [39].

APPENDIX H PROOF OF THEOREM 4

We first consider that the local loss function for all learning SVs, i.e., $\epsilon_n, \forall n \in \mathcal{N}(t)$, are all δ_1 -smooth and δ_2 -strongly convex [11], i.e.,

$$\begin{aligned} \epsilon_n(\mathbf{W}(t)) - \epsilon_n(\hat{\mathbf{W}}(t)) &\leq \langle \mathbf{W}(t) - \hat{\mathbf{W}}(t), \nabla \epsilon_n(\hat{\mathbf{W}}(t)) \rangle \\ &+ \frac{\delta_1}{2} \|\mathbf{W}(t) - \hat{\mathbf{W}}(t)\|_2^2, \forall n \in \mathcal{N}(t), \end{aligned} \quad (72)$$

and

$$\begin{aligned} \epsilon_n(\mathbf{W}(t)) - \epsilon_n(\hat{\mathbf{W}}(t)) &\geq \langle \mathbf{W}(t) - \hat{\mathbf{W}}(t), \nabla \epsilon_n(\hat{\mathbf{W}}(t)) \rangle \\ &+ \frac{\delta_2}{2} \|\mathbf{W}(t) - \hat{\mathbf{W}}(t)\|_2^2, \forall n \in \mathcal{N}(t), \end{aligned} \quad (73)$$

respectively. Then, given that $0 < \kappa(t) \leq \min(1, \frac{1}{\delta_2 \tau_{th}}), \forall t$, where δ_2 is a positive constant and $\kappa(t) = \kappa_1^{(\tau_{th})}(t) = \dots = \kappa_N^{(\tau_{th})}(t)$ from (41) with fixed $\beta_{p_n}^{(\tau_{th})}(t)$ and $\beta_{q_n}^{(\tau_{th})}(t)$, $\forall n \in \mathcal{N}(t)$, we can obtain the expected squared L2-norm global model gap [11] as follows:

$$\begin{aligned} \mathbb{E}[\|\mathbf{W}(t) - \mathbf{W}^*\|_2^2] &\leq \left(\prod_{t^*=0}^{t-1} f_1(t^*) \right) \|\mathbf{W}(0) - \mathbf{W}^*\|_2^2 \\ &+ \sum_{t^\dagger=0}^{t-1} f_2(t^\dagger) \prod_{t^*=t^\dagger+1}^{t-1} f_1(t^*), \end{aligned} \quad (74)$$

where

$$\begin{aligned} f_1(t^*) &\triangleq 1 - \delta_2 \kappa(t^*) (\tau_{th} - \kappa(t^*) (\tau_{th} - 1)), \quad (75) \\ f_2(t^*) &\triangleq \frac{(I-N)\kappa^2(t^*)\tau_{th}^2\Upsilon^2}{N(I-1)} + 2\kappa(t^*)(\tau_{th}-1)\Lambda + \\ &(1 + \delta_2(1 - \kappa(t^*)))\kappa^2(t^*)\Upsilon^2 \frac{\tau_{th}(\tau_{th}-1)(2\tau_{th}-1)}{6} + \\ &\kappa^2(t^*)(\tau_{th}^2 + \tau_{th} - 1)\Upsilon^2, \end{aligned} \quad (76)$$

$$\Lambda \triangleq \Psi^*(\mathbf{W}^*) - \frac{1}{I} \sum_{i=1}^I \epsilon_i(\mathbf{W}^*) \geq 0, \quad (77)$$

and $\Upsilon = \max \|\mathbf{W}(t)\|$ is the maximum global model for all learning rounds. In this case, the first component $\left(\prod_{t^*=0}^{t-1} f_1(t^*) \right) \|\mathbf{W}(0) - \mathbf{W}^*\|_2^2$ in (74) indicates the global model gap in the initial round. Meanwhile, the second component $\sum_{t^\dagger=0}^{t-1} f_2(t^\dagger) \prod_{t^*=t^\dagger+1}^{t-1} f_1(t^*)$ in (74) implies the cumulate impact of the proposed SV selection method on the global model convergence.

Using the δ -smoothness function in (72) for the global loss $\Psi(\cdot)$, we have the below expression after t^\diamond rounds, i.e.,

$$\begin{aligned} \mathbb{E}[\Psi(\mathbf{W}(t^\diamond))] - \Psi^*(\mathbf{W}^*) &\leq \frac{\delta}{2} \mathbb{E}[\|\mathbf{W}(t^\diamond) - \mathbf{W}^*\|_2^2] \\ &\leq \frac{\delta}{2} \left(\prod_{t^*=0}^{t^\diamond-1} f_1(t^*) \right) \|\mathbf{W}(0) - \mathbf{W}^*\|_2^2 + \\ &\frac{\delta}{2} \sum_{t^\dagger=0}^{t^\diamond-1} f_2(t^\dagger) \prod_{t^*=t^\dagger+1}^{t^\diamond-1} f_1(t^*). \end{aligned} \quad (78)$$

Equation (78) shows that the global loss gap is upper bounded by $\frac{\delta}{2} \mathbb{E}[\|\mathbf{W}(t^\diamond) - \mathbf{W}^*\|_2^2]$. Then, due to the decreasing learning rate $\lim_{t \rightarrow \infty} \kappa(t) = 0$, we can observe that $\lim_{t^\diamond \rightarrow \infty} [\mathbb{E}[\Psi(\mathbf{W}(t^\diamond))] - \Psi^*(\mathbf{W}^*)] = 0$, which implies that the global loss gap eventually will reach zero. This concludes the proof.