# Distributed Mechanism Design in Continuous Space for Federated Learning over Vehicular Networks

Mina Montazeri, Hamed Kebriaei*, *Senior Member, IEEE*, Babak N. Araabi, Jiawen Kang, Dusit Niyato, *Fellow, IEEE*

***Abstract*—With the rapid development of the 5G network and the promising of 6G network, vehicles can report a large amount of real-time traffic information to Road-Side Units (RSUs). However, due to the large communication cost, limited computing resources, and privacy leakage risks, centralized data processing by RSUs is not efficient and not secure. Thereby, federated learning is introduced to enable vehicles to train local models and send the model parameters to the RSUs, without the need for revealing their personal data. Nevertheless, without an efficient incentive mechanism, the vehicles, as data owners, may be unwilling to join the federated learning task. In this paper, we adopt contract theory to design an incentive mechanism with asymmetric information and continuum of types for interaction between the RSU and vehicles. The designed mechanism satisfies the incentive compatibility, individual rationality and also maximizes the RSU's profit. The technical challenge is to solve a non-quadratic functional optimization in continuous space with coupling constraints among the vehicles. To address this challenge, we propose two iterative algorithms to find the RSU's optimal strategies as the functions of vehicles' types . In the first algorithm, a combination of the calculus of variation and dual decomposition methods is utilized to achieve an analytical solution for the optimal strategy of RSU, while in the second one, a combination of approximate dynamic programming and neural networks is used to estimate the optimal strategy of RSU.**

***Index Terms*—Federated learning, contract theory, asymmetric information, vehicular network, optimal control, dual decomposition, approximate dynamic programming.**

## I. Introduction

Nowadays, 5G network is being deployed in many countries. Meanwhile, the artificial intelligence-enabled 6G network is being conceived by researchers for the future evolution of network intelligentization [1]. These networks have brought fundamental changes into a smart city and caused the possibilities for advanced vehicular services and applications such as autonomous driving and traffic prediction which can yield an improved driving experience [2]. Most of the traditional machine learning techniques need to gather vehicle data into a central server to perform model training. However, such techniques face some limitations in the computational and wireless networks resources, and also the vehicles may suffer from violating their information privacy by sending personal raw data to a central system [3]. To address these concerns, Federated Learning (FL) as a distributed machine learning technique has recently been introduced by Google in which, the model training can be performed in a distributed fashion [4]. Specifically, federated learning assigns the training task to distributed data owners, and instead of sending the raw data, they only sent the intermediate information like gradient values to the central server (task publisher) to train the model [5]. When FL is adopted in the vehicular network, the RSU is considered as a central server for aggregating trained parameters from different vehicles and transferring the aggregated trained parameter to vehicles back [6].

Due to computational resource costs for model training and privacy leakage risks, vehicles may be reluctant to participate in the FL task. Therefore, incentive mechanisms have been proposed to motivate vehicles to participate in FL tasks in return for certain rewards [7]. The goal of this paper is to study optimal contracts in the federated learning task when the RSU as a task publisher has incomplete information from vehicles' private information (i.e. data quality). This contract is proposed such that the RSU provides incentive rewards and computational resources to vehicles in order to achieve the optimal FL task allocation. The incentive reward and computational resource allocation which are modeled as the functions of the vehicle's private information are designed to achieve three objectives simultaneously: motivate vehicles to participate in the FL task, ensure truthful reporting of the private information of vehicles, and maximize the RSU's utility. In the designing process, we consider that the private information of each vehicle is drawn from a specific continuous distribution. This assumption makes the problem more realistic and challenging since we face a double continuum of incentive constraints. Considering different distributions over private information of vehicles poses some technical challenges in dealing with the coupling constraints among them.

In this paper, we propose two algorithms to solve the

This research was supported in part by Key-Area and Development Program of Guangdong Province (Grant No. 2020B0101090004), NSFC under grant No. 62102099. This research was also supported in part by the National Research Foundation (NRF) and Infocomm Media Development Authority under the Future Communications Research and Development Programme (FCP), under the AI Singapore Programme (AISG) (AISG2-RP-2020-019), under Energy Research Test-Bed and Industry Partnership Funding Initiative, part of the Energy Grid (EG) 2.0 programme and under DesCartes and the Campus for Research Excellence and Technological Enterprise (CREATE) programme. The work was also in part supported by Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT2022B12). This work was also in part supported by a grant from the Institute for Research in Fundamental Sciences (IPM) under grant number: CS 1400-4-451. (Corresponding Author: Hamed Kebriaei).

Mina Montazeri, Hamed Kebriaei and Babak N. Araabi are with the School of ECE, College of Engineering, University of Tehran, Tehran, Iran. H. Kebriaei is also with the School of Computer Science, Institute for Research in Fundamental Sciences (IPM), P.O. Box 19395-5746, Tehran, Iran. Emails: {mina.montazeri@ut.ac.ir, kebriaei@ut.ac.ir, araabi@ut.ac.ir}. Jiawen Kang is with Automation of School, Guangdong University of Technology, China. Email: {kavinkang@gdut.edu.cn}. Dusit Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. Email: {dniyato@ntu.edu.sg}.

functional optimization of the contract design problem. In the first algorithm, the optimal computation resource allocation and incentive reward functions are determined analytically by using calculus of variation and dual decomposition methods. In the second algorithm, the optimal strategy is approximated using the Approximate Dynamic Programming (ADP) method via actor and critic neural networks. The first algorithm converges in a less number of iterations than that of the second algorithm. However, each iteration of the first algorithm takes more time than that of the second algorithm. Thus, the total computation time of the second algorithm becomes less than that of the first algorithm. The main contributions of the paper are as follows:

- We propose a new contract theory based incentive mechanism for federated learning tasks in vehicular networks. The novel optimal contract design is the first mechanism for federated learning that is based on continuous private information of vehicles. More importantly, the private information is drawn from a continuous distribution and is not necessarily identical for different data owners.
- To solve the specific contract design problem with functional optimization, we transformed the contract model as an optimal control problem with coupling constraints and derive an explicit relation between two decision functions of the RSU which simplifies the solution of the proposed constrained functional optimization problem.
- We propose two algorithms to solve functional optimization with coupling constraints. In the first algorithm, we introduce an analytical method to determine the optimal strategy using a combination of the calculus of variation and the dual decomposition method. In the second algorithm, we propose an ADP algorithm with neural networks to approximate the optimal strategy of RSU.

The rest of this paper is organized as follows: Section II reviews the literature about using federated learning in mobile edge computing and vehicular networks. Section II describes the system model and problem formulation. In Section IV, we discuss the contract theory requirements and model the contract problem as an optimal control problem, In Section V, two iterative algorithms are proposed to solve contract optimization. In Section VI, simulations are conducted to show the method's features. Finally, Section VII concludes the paper.

## II. Literature Review

As a privacy-preserving learning tool, FL has attracted more studies recently. The authors in [8] provided a comprehensive study on FL and its applications in terms of FL's security and privacy. They presented a detailed review of security and privacy concerns that needed to be considered in the FL method. The authors in [9] formulated an FL over a wireless network as an optimization problem that captured the trade-off between the FL time and data owner energy consumption. They also showed how this non-convex optimization can be solved. [10] applied FL for classification under the condition of imbalance and noise varying. The proposed FL method can achieve high accuracy with a low risk of data leakage.

Most of the existing work made some unrealistic assumptions about assigning FL tasks to data owners. Firstly, they assume that all the data owners voluntarily participate in FL [11], which is not practical in the real world due to resource costs incurred by model training [12]. Secondly, if the central system wants to allocate the FL task to the data owners optimally, it needs to know all their private information i.e. data quality. Most of the studies assume that all data owners report such information truthfully to the task publisher [13]. However, this may not be applicable in the real world as agents have the incentive to announce their private information incorrectly if it can make more profit for them [12]. In this case, contract theory offers a framework to design mechanisms under asymmetric information. In this framework, the task publisher as a designer provides incentive rewards to data owners to persuade them to join the FL task and report their information truthfully [14]. The author in [15] used the VCG mechanism to encourage the selfish agents to play truthfully by paying appropriate incentives. In [16], the authors proposed an auction as an incentive mechanism for the wireless FL market. In the proposed auction, each mobile user submits its bids, and the base station decides the winner such that maximizes social welfare. Both [15] and [16] try to maximize social welfare. However, most of the time base station as a selfish participant would like to maximize its profit instead of maximizing social welfare. In [17], [18], the authors model the interaction between the task publisher and the data owners as a contract with asymmetric information. In [17], the authors considered private information with one dimension, while in [18] the authors consider multi-dimensional private information. In both of these papers, the authors assumed that the data owners are divided into a discrete and finite set of classes. In each of these classes, the data owners have specific private information. Although the task publisher does not know the exact true class of a data owner, it has the knowledge of the probability that a data owner belongs to a certain class. The assumption of this paper is not realistic for two reasons. Firstly, a discrete and finite set of classes is not realistic in real word applications and essentially simplifies the solution since the double continuum of incentive compatibility constraints is reduced to a finite number of constraints. Nevertheless, the continuum of constraint is more realistic and technically rigorous [19]. Secondly, in real-world applications, each data owner's private information is drawn from a different distribution. However, in [17], it was assumed that all data owners' private information has the same distribution.

## III. System Model of Federated Learning Task in Vehicular Networks

We consider the federated learning framework in vehicular networks as a monopoly market with an RSU as a task publisher and a set of vehicles as data owners $\mathcal{N} = \{1, 2, \dots, N\}$ [9], [17], and take traffic sign recognition as a task example of applications of federated learning in vehicular networks [20].

Each vehicle is equipped with computing and caching resources [20] and the RSU would like to use this computational potential for traffic sign recognition. Thus, the RSU designs contract items for incentivizing vehicles to join federated learning. For the federated learning task, the vehicles are the nodes to perform computation on their local data in order to update a global model. During a global iteration, the vehicles iteratively train a shared global model with local model updates generated using their private local data. Then, the vehicles upload their local model updates to the RSU for updating the global model. Note that the data size of global model is small enough to be fastly transmitted to the moving vehicles. Thus, without loss of generality, we consider that the average transmission rate is stable and the effects of moving vehicles could be ignored. Figure 1 shows the model of federated learning tasks for traffic sign recognition.

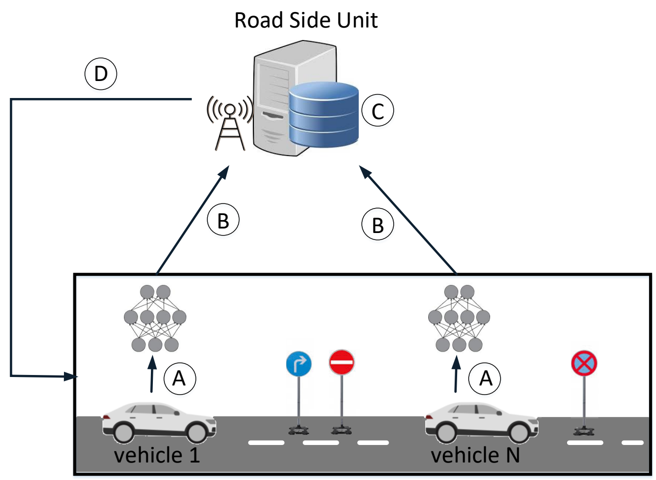



Fig. 1. The system model of federated learning task for traffic sign recognition in vehicular networks [21]. The training process contains four steps: A. update the local model: all the vehicles train the machine learning model with their own data locally. B. send local model updates: each vehicle uploads some parameters of the trained model to RSU. C. aggregate local models and update global model: The RSU performs secure aggregation over the uploaded parameters from vehicles to obtain the global model. D. send model updates: The RSU distributes the parameters of the global model to the vehicles.

The total energy consumption of the vehicles in federated learning includes two parts: i) energy consumption of CPU computation to generate local model updates, and ii) energy consumption of the model update transmission [9]. The computational time and energy consumption of the vehicle $i$ for a local iteration are expressed in (1) [9].

$$T_i^{cmp} = \frac{c_i \times s_i}{x_i}, \;\; E_i^{cmp} = \zeta_i \times c_i \times s_i \times x_i^2, \tag{1}$$

where $\zeta_i$ is the effective capacitance parameter of computing chip set for vehicle $i$, $s_i$ is the size of local data sample for vehicle $i$, $c_i$ is the number of CPU cycles for a vehicle $i$ to perform local model training using one data sample, and $x_i$ is the computational resource of this vehicle.

Each vehicle has local data quality (i.e. local data accuracy or local data reliability) which is denoted by $\epsilon_i$ [17]. A vehicle with high-quality data leads to fewer local and global iterations [22]. Given a fixed global accuracy, the number of local iterations needed is $\left(\phi \times log(\frac{1}{\epsilon_i})\right)$ where $\phi$ is the predefined coefficient [9]. Therefore, for a global iteration, the computational energy consumption of the vehicle $i$ is $\left(\phi \times log(\frac{1}{\epsilon_i}) \times E_i^{cmp}\right)$. We define $\frac{1}{\phi \times log(\frac{1}{\epsilon_i})}$ as a type of the vehicle $i$ which is denoted by $\theta_i \in [\underline{\theta}, \bar{\theta}]$. Each vehicle's type is its private information (i.e. only known to the vehicle itself and neither the RSU nor other vehicles know the type). However, its cumulative distribution $F(\theta_i)$ is common knowledge [23]. This knowledge can be obtained, for example, through analyzing historical data [17]. We assume that $F(\theta_i)$ is continuously differentiable. The total time of participating in a global iteration for the vehicle $i$ is as follows:

$$T_i^t = \left(\frac{1}{\theta_i} \times T_i^{cmp}\right) + T_i^{com}, \tag{2}$$

where $T_i^{com}$ is the average transmission time to transmit local model updates in a global iteration for vehicle $i$. Also the total energy consumption of the vehicle $i$ is as follows:

$$E_i^t = \left(\frac{1}{\theta_i} \times E_i^{cmp}\right) + E_i^{com}, \tag{3}$$

where $E_i^{com}$ is the energy consumption by vehicle $i$ to transmit local model updates in a global iteration. Due to similar wireless communication environments, we consider the average transmission time and transmission energy consumption to be the same statistically for all vehicles. These statistically identical vehicles are practical in many situations such as when the mobile units of vehicles are of the same type.

Let $\hat{\theta}_i$ be the announced type by a vehicle $i$ to RSU which is not necessarily its true type, i.e. $\theta_i$. Then, the computation resource and the corresponding reward offered by the RSU for the vehicle $i$ are denoted by $x_i(\hat{\theta}_i)$ and $R_i(\hat{\theta}_i)$, respectively. Thus the utility function of the RSU is defined as follows:

$$U_{TP}(x_i(\hat{\theta}_i), R_i(\hat{\theta}_i)) = \sum_{i=1}^{N} \boldsymbol{E_{\theta_i}}[S_i(\hat{\theta}_i, x_i(\hat{\theta}_i)) - R_i(\hat{\theta}_i)], \tag{4}$$

where $S_i(\theta_i, x_i(\theta_i)) \equiv w \ln(T_{max} - T_i^t)$ is the satisfaction function of the task publisher for the total time of a global iteration [17]. Here $w > 0$ is the satisfaction degree parameter of RSU and $T_{max}$ is the RSU's maximum tolerance time for the FL task. Note that both the higher quality (higher type) and larger computational resource can improve the profit for the RSU, i.e., $\frac{\partial U_{TP}}{\partial \theta} > 0$, $\frac{\partial U_{TP}}{\partial x} > 0$ and $\frac{\partial U_{TP}}{\partial \epsilon} > 0$. From the RSU's point of view, a constraint that is related to the limitation of the total reward budget to be satisfied while maximizing its utility. Therefore, the RSU needs to solve the following constrained maximization problem:

$$\max_{R_i(\hat{\theta}_i), x_i(\hat{\theta}_i)} \quad U_{TP}(x_i(\hat{\theta}_i), R_i(\hat{\theta}_i)), \tag{5}$$
$$s.t. \quad \sum_{i=1}^{N} \boldsymbol{E_{\theta_i}}[R_i(\theta_i)] \leq R_{max},$$

where $R_{max}$ is the total reward budget of the RSU. The utility function of a vehicle $i$ is defined as follows:

$$U_i^D(\theta_i, \hat{\theta}_i) = R_i(\hat{\theta}_i) - \mu E_i^t(\theta_i, x_i(\hat{\theta}_i)), \tag{6}$$

where $\mu$ is a pre-defined parameter for energy consumption.

While the RSU aims to maximize its utility over $x_i(\hat{\theta}_i)$ and $R_i(\hat{\theta}_i)$, it faces two main problems: the RSU must ensure that vehicles are sufficiently incentivized to participate and also to declare their types truthfully. Hence, due to information asymmetry, the RSU should design specific contracts for different types of vehicles in order to maximize its profits. In the next section, these criteria are taken into account, and by utilizing contract theory, the optimal incentive strategy of the RSU is achieved.

## IV. Contract Between RSU and vehicles

The major steps of the contracting game are as follows: The RSU designs strategy functions $(R_i(\hat{\theta}_i), x_i(\hat{\theta}_i))$ for each vehicle and announces the strategy function of each vehicle. Then, Each vehicle determines its optimal announcement type that maximizes its utility and finally, based on the RSU's strategy function and vehicles' announcement types, the payoff of all vehicles and the task RSU are determined. An optimal contract must attract the vehicles to participate in the FL and ensure that each vehicle announces its type truthfully. The following is a definition of these constraints.

**Definition 1.** *A mechanism is Individually Rational (IR) if the vehicle's utility is non-negative by announcing its type truthfully:*

$$U_i^D(\theta_i, \theta_i) \geq 0, \qquad \forall \theta_i \in [\underline{\theta}, \bar{\theta}], i \in \mathcal{N}. \tag{7}$$

**Definition 2.** *A mechanism is Incentive Compatible (IC) if the vehicle cannot be better off by misreporting its type:*

$$U_i^D(\theta_i, \theta_i) \geq U_i^D(\theta_i, \hat{\theta}_i), \quad \forall \hat{\theta}_i, \theta_i \in [\underline{\theta}, \bar{\theta}], i \in \mathcal{N}. \tag{8}$$

From the contract theory point of view, the RSU needs to maximize its utility, subject to IR and IC constraints. Therefore, optimization (5), can be reformulated as follows.

$$\max_{R_i(\hat{\theta}_i), x_i(\hat{\theta}_i)} \sum_{i=1}^{N} \boldsymbol{E}_{\boldsymbol{\theta_i}}[S_i(\hat{\theta}_i, x_i(\hat{\theta}_i)) - R_i(\hat{\theta}_i)], \tag{9a}$$

$s.t.$

$$R_i(\theta_i) - \mu E_i^t(\theta_i, x_i(\theta_i)) \geq R_i(\hat{\theta}_i) - \mu E_i^t(\theta_i, x_i(\hat{\theta}_i)), \tag{9b}$$
$$\forall \hat{\theta}_i, \theta_i \in [\underline{\theta}, \bar{\theta}], i \in \mathcal{N}(IC_{\theta_i, \hat{\theta}_i}),$$

$$R_i(\theta_i) - \mu E_i^t(\theta_i, x_i(\theta_i)) \geq 0, \quad \forall \theta \in [\underline{\theta}, \bar{\theta}], i \in \mathcal{N}(IR_{\theta_i}), \tag{9c}$$

$$\sum_{i=1}^{N} \boldsymbol{E}_{\boldsymbol{\theta_i}}[R_i(\theta_i)] \leq R_{max}. \tag{9d}$$

The $IR_{\theta_i}$ inequality reflects the fact that the vehicle $i$ is incentivized to participate in this contract. The $IC_{\theta_i, \hat{\theta}_i}$ inequality reflects the fact that for the vehicle $i$ with type $\theta_i$, truthful reporting of type is the best choice, and hence, there is no incentive to misrepresent itself as a vehicle with unreal type.

### A. Reformulation as Optimal Control Problem

The optimization problem (9) is hard to be solved. Thus, we reformulate it as an optimal control problem.

**Lemma 1.** *In equation* (9)*, the $IR_{\theta_i}$ constraint is satisfied for all $\theta_i$ if $IR_{\underline{\theta}}$ is binding (or active) which means $U_i^D(\underline{\theta}, \underline{\theta}) = 0$.*

*Proof.* According to [14], if $\frac{\partial U_i^D}{\partial \theta_i}(x_i(\hat{\theta}_i), \theta_i) > 0$, then $IR_{\underline{\theta}}$ and $IC_{\theta_i, \hat{\theta}_i}$ imply $IR_{\theta_i}$. Furthermore, it proves that $IR_{\underline{\theta}}$ is binding. Otherwise, we could decrease $R_i(\theta_i)$ for all $\theta_i \in [\underline{\theta}, \bar{\theta}]$ by $\epsilon > 0$, which would preserve all constraints of optimization (9) and also increase the utility of RSU. □

**Lemma 2.** *The solution of $(R_i(\theta_i), x_i(\theta_i))$ in equation* (9) *is IC if and only if both of the following conditions hold:*

$$x_i'(\theta_i) \geq 0, \tag{10}$$

$$R_i(\theta_i) = \mu E_i^t(x_i(\theta_i), \theta_i) - \mu \int_{\underline{\theta}}^{\theta_i} \frac{\partial E_i^t}{\partial \theta_i}(x_i(y), y) dy. \tag{11}$$

*Proof.* We divide the proof of this lemma into two parts, the forward direction "if", and the backward direction, "only if".

To show "if" part, by replacing $E_i^t(x_i(\theta_i), \theta_i)$ from equation (3), for $\theta_i > \hat{\theta}_i$, constraint $IC_{\theta_i, \hat{\theta}_i}$ can be re-written as:

$$\mu \int_{\theta_i}^{\hat{\theta}_i} -\frac{1}{y^2} \zeta_i c_i s_i x_i^2(y) dy \geq \mu \zeta_i c_i s_i x_i^2(\hat{\theta}_i)(\frac{1}{\hat{\theta}_i} - \frac{1}{\theta_i}). \tag{12}$$

For $\theta_i < \hat{\theta}_i$, we have:

$$\mu \zeta_i c_i s_i x_i^2(\hat{\theta}_i)(\frac{1}{\theta_i} - \frac{1}{\hat{\theta}_i}) \geq \mu \int_{\hat{\theta}_i}^{\theta_i} -\frac{1}{y^2} \zeta_i c_i s_i x_i^2(y) dy, \tag{13}$$

which both equations (12) and (13) hold true due to the monotonicity of $x_i(\theta_i)$. To show "only if" part, we prove that truthfulness implies monotonicity of $x_i(\theta_i)$. According to the IC constraint in equation (9), we can obtain:

$$R_i(\theta_i) - \mu E_i^t(\theta_i, x_i(\theta_i)) \geq R_i(\hat{\theta}_i) - \mu E_i^t(\theta_i, x(\hat{\theta}_i))(IC_{\theta_i, \hat{\theta}_i}), \tag{14}$$

$$R_i(\hat{\theta}_i) - \mu E_i^t(\hat{\theta}_i, x_i(\hat{\theta}_i)) \geq R_i(\theta_i) - \mu E_i^t(\hat{\theta}_i, x_i(\theta_i))(IC_{\hat{\theta}_i, \theta_i}). \tag{15}$$

Given $E_i^t(x_i(\theta_i), \theta_i)$ in (3), and by summation of equations (14) and (15), we get:

$$x_i^2(\hat{\theta}_i) \times (\frac{1}{\theta_i} - \frac{1}{\hat{\theta}_i}) \geq x_i^2(\theta_i) \times (\frac{1}{\theta_i} - \frac{1}{\hat{\theta}_i}). \tag{16}$$

which implies monotonicity of $x_i(\theta_i)$. To derive (11), we can rearrange equations (14) and (15) as follows:

$$\mu E_i^t(\theta_i, x_i(\theta_i)) - \mu E_i^t(\theta_i, x_i(\hat{\theta}_i)) \leq R_i(\theta_i) - R_i(\hat{\theta}_i) \leq$$
$$\mu E_i^t(\hat{\theta}_i, x_i(\theta_i)) - \mu E_i^t(\hat{\theta}_i, x_i(\hat{\theta}_i)). \tag{17}$$

Given $E_i^t(x_i(\theta_i), \theta_i)$ in (3), and by considering $\hat{\theta}_i = \theta_i + \epsilon$ and dividing throughout (17) by $\epsilon$, and letting $\epsilon \to 0$, we have:

$$\mu c_i s_i \frac{1}{\theta_i} \frac{d}{d\theta_i} x_i^2(\theta_i) \leq \frac{d}{d\theta_i} R_i(\theta_i) \leq \mu c_i s_i \frac{1}{\theta_i} \frac{d}{d\theta_i} x_i^2(\theta_i). \tag{18}$$

Thus $\mu c_i s_i \frac{1}{\theta_i} \frac{d}{d\theta_i} x_i^2(\theta_i) = \frac{d}{d\theta_i} R_i(\theta_i)$. Integrating equation (18) with respect to $\theta_i$ from $\underline{\theta}$ to $\theta_i$ and applying integration by parts we have:

$$R_i(\theta_i) = \mu c_i s_i \frac{1}{\theta_i} x_i^2(\theta_i) - \mu c_i s_i \int_{\underline{\theta}}^{\theta_i} -\frac{1}{y^2} x_i^2(y) dy. \tag{19}$$

Considering definition of $E_i^t(x_i(\theta_i), \theta_i)$ in equation (3), equation (19) can be written as (11) which completes the proof. □

**Definition 3.** *$h(t) \equiv \frac{f(t)}{1-F(t)}$ is known as the hazard rate of $t$ in the statistics literature [14].*

**Proposition 1.** *The optimal solution to the optimization problem* (9) *is the same as the following optimization solution.*

$$\max_{x_i(\theta_i)} \sum_{i=1}^{N} \int_{\underline{\theta}}^{\bar{\theta}} U_i^{TP-PW}(x_i(\theta_i),\theta_i) d\theta_i, \tag{20a}$$

$$s.t.\ x_i'(\theta_i) = u_i(\theta), \tag{20b}$$

$$u_i(\theta_i) \geq 0, \tag{20c}$$

$$\sum_{i=1}^{N} \int_{\underline{\theta}}^{\bar{\theta}} R_i^{PW} d\theta_i \leq R_{max}, \tag{20d}$$

*where* $U_i^{TP-PW} \equiv [S_i(x_i(\theta_i),\theta_i) + \mu \frac{\partial E_i^t(x_i(\theta_i),\theta_i)}{\partial \theta_i} \frac{1}{h(\theta_i)} - \mu E_i^t(x_i(\theta),\theta_i)] f(\theta_i)$ *and* $R_i^{PW} \equiv [\mu E_i^t(x_i(\theta_i),\theta_i) - \mu \frac{\partial E_i^t(x_i(\theta_i),\theta_i)}{\partial \theta_i} \frac{1}{h(\theta_i)}] f(\theta_i)$.

*Proof.* Replacing $R_i(\theta_i)$ from equation (11) in the utility function of RSU, we have:

$$U_{TP}(x_i(\theta_i)) = \sum_{i=1}^{N} \big[ \int_{\underline{\theta}}^{\bar{\theta}} [S_i(x_i(\theta_i)) - \mu E_i^t(x_i(\theta_i),\theta_i) \tag{21}$$
$$+ \mu \int_{\underline{\theta}}^{\theta_i} \frac{\partial E_i^t}{\partial \theta_i}(x_i(s),s) ds] f(\theta_i) d\theta_i \big].$$

The integration by parts of the second term gives:

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[ \int_{\underline{\theta}}^{\theta_i} \frac{\partial E_i^t}{\partial \theta_i}(x_i(s),s) ds \right] f(\theta_i) d\theta_i = \tag{22}$$
$$\int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial E_i^t(x_i(\theta_i),\theta_i)}{\partial \theta_i} \frac{1-F(\theta_i)}{f(\theta_i)} f(\theta_i) d\theta_i.$$

According to (21) and (22), we can rewrite $U_{TP}$ as follows:

$$U_{TP}(x_i(\theta_i)) = \sum_{i=1}^{N} \big[ \boldsymbol{E}_{\boldsymbol{\theta_i}} [S_i(x_i(\theta)) \tag{23}$$
$$+ \mu \frac{\partial E_i^t(x_i(\theta_i),\theta_i)}{\partial \theta_i} \frac{1-F(\theta_i)}{f(\theta_i)} - \mu E_i^t(x_i(\theta_i),\theta_i)] \big].$$

In a similar way $\boldsymbol{E}_{\boldsymbol{\theta_i}}[R_i(\theta_i)]$, in (9d) can be rewritten as:

$$\boldsymbol{E}_{\boldsymbol{\theta_i}}[\mu E_i^t(x_i(\theta_i),\theta_i) - \mu \frac{\partial E_i^t(x_i(\theta_i),\theta_i)}{\partial \theta_i} \frac{1-F(\theta_i)}{f(\theta_i)}]. \tag{24}$$

Thus, the optimization (9) can be written as (20)

$$\max_{x_i(\theta_i)} \sum_{i=1}^{N} \big[ \int_{\underline{\theta}}^{\bar{\theta}} [S_i(x_i(\theta_i)) + \mu \frac{\partial E_i^t(x_i(\theta_i),\theta_i)}{\partial \theta_i} \frac{1}{h(\theta_i)} \tag{25a}$$
$$- \mu E_i^t(x_i(\theta_i),\theta_i)] f(\theta_i) d\theta_i \big],$$

$$s.t.\quad x_i'(\theta_i) \geq 0, \tag{25b}$$

$$\sum_{i=1}^{N} \int_{\underline{\theta}}^{\bar{\theta}} [\mu E_i^t(x_i(\theta_i),\theta_i) - \mu \frac{\partial E_i^t(x_i(\theta_i),\theta_i)}{\partial \theta_i} \frac{1}{h(\theta_i)}] f(\theta_i) d\theta_i \tag{25c}$$
$$\leq R_{max}.$$

□

The optimization problem (20) appears in the form of an optimal control problem in dynamic optimization, with state variables $x_i(\theta_i)$ and control variables $u_i(\theta_i)$. Note that constraint (20d) is the coupling constraint. In order to handle the coupling constraint and also, to avoid the high computational costs imposed by a centralized method for large-scale systems, the dual decomposition method is used to decompose the problem into multiple individual sub-problems [24]. Also, since we face a functional optimization, where the decision variables are functions rather than vectors or scalars, the optimal control theory is employed to solve optimization (25). Nevertheless, since the closed-form solution of each sub-problem does not exist, two approaches are utilized to solve optimization (20): (i) the algorithm based on the calculus of variation and optimal control theory, (ii) approximate dynamic programming algorithm that approximates the optimal strategy by using neural networks.

## V. Distributed Solution of Contract Model

In this section, we investigate the solution for the functional optimization problem (20). In order to handle the coupling constraint (20d), the dual decomposition method is used to decompose the main optimization problem into multiple individual sub-problems. Thus, optimization (20) can be solved in an iterative fashion, where the RSU updates the dual variables using a sub-gradient method and broadcasts the dual variables to the vehicles. Then, $N$ sub-problems are solved by vehicles using the value of dual variables in the last iteration [24]. In what follows, we explain two algorithms that can solve optimization (20).

### A. Calculus of variation with dual decomposition method

The Lagrangian function of the optimization problem (20) by considering the constraints is defined as follows:

$$L(u_i,\lambda_i,w_i,\nu) = \sum_{i=1}^{N} U_{TP-PW} + \sum_{i=1}^{N} \lambda_i u_i + \tag{26}$$
$$\sum_{i=1}^{N} w_i u_i + \nu(\sum_{i=1}^{N} R_i^{PW} - R_{max}),$$

where $\lambda_i$, $w_i$, and $\nu$ are multipliers associated with constraints (20b), (20c), and (20d), respectively. Since $\frac{\partial^2}{\partial \theta_i^2} U_i^{TP-PW}(x_i(\theta_i),\theta_i) < 0$, the objective function of optimization (20a) is strictly concave, and also the constraints (20b), (20c) and (20d) are convex. Therefore, this optimization is strictly convex and hence, the strong duality holds [25]. Sub-problem $i$ is an optimal control problem that is defined as follows:

$$\max_{u_i(\theta_i),\lambda_i(\theta_i),w_i(\theta_i)} U_i^{TP-PW} + \lambda_i u_i + \ w_i u_i + \nu R_i^{PW}. \tag{27}$$

The necessary conditions for optimality are:

$$\frac{\partial L}{\partial u_i} = \lambda_i + w_i = 0, \qquad w_i \geq 0, \quad w_i \times u_i = 0, \tag{28}$$

$$\frac{d\lambda_i}{d\theta_i} = -\frac{\partial L}{\partial x_i} = -\frac{\partial U_i^{TP-PW}}{\partial x_i} - \nu \frac{\partial R_i^{PW}}{\partial x_i}, \quad \lambda_i(\bar{\theta}) = 0. \tag{29}$$

The Lagrangian is strictly concave in $x_i$ (since $\frac{\partial^2 U_i^{TP-PW}}{\partial {x_i}^2} \geq 0$), and hence there is no jump in $x_i$. Thus $u_i$ is bounded. Integrating equation (29) gives us:

$$\lambda_i = \int_{\theta_i}^{\bar{\theta}} -\frac{\partial}{\partial x_i} U_i^{TP-PW}(x_i(y), y) - \nu \frac{\partial}{\partial x_i} R_i^{PW}(x_i(y), y) dy. \tag{30}$$

According to (28), either $\lambda_i < 0$ and $u_i = 0$, or else $\lambda_i = 0$ and $u_i > 0$. Thus, substituting $\lambda_i$ from (30) we have:

$$\int_{\theta_i}^{\bar{\theta}} -\frac{\partial}{\partial x_i} U_i^{TP-PW}(x_i(y), y) - \nu \frac{\partial}{\partial x_i} R_i^{PW}(x_i(y), y) dy \leq 0. \tag{31}$$

On an interval which $u_i > 0$, equation (31) holds with equality that gives $\frac{\partial}{\partial x_i} U_i^{TP-PW} + \nu \frac{\partial}{\partial x_i} R_i^{PW} = 0$. Thus, in this case by definitions of $U_i^{TP-PW}$ and $R_i^{PW}$, we have:

$$\begin{aligned} &\frac{w}{T_{max} - T^{com} - \frac{c_i s_i}{\theta_i x_i(\theta_i)}} \times \frac{c_i s_i}{\theta_i x_i^2(\theta_i)} \\ &- 2\mu\zeta_i c_i s_i x_i(\theta_i) \times (\frac{1}{\theta_i^2 h(\theta_i)} + \frac{1}{\theta_i}) \times (1-\nu) = 0. \end{aligned} \tag{32}$$

There can be two kinds of intervals for $u_i = 0$. First, in the initial interval, i.e. $0 < \theta_i < \theta_i^0$. Second, the interior intervals i.e. $\theta_i^{j1}, < \theta_i < \theta_i^{j2}$ , where $j \in \{1, 2, \ldots, M\}$ and $M$ is the number of interior intervals in which $u_i = 0$. Since $x_i'(\theta_i) = 0$ in these ranges, the vehicles with the types in these ranges are assigned the same computational resource (and same reward). Hence $x_i(\theta_i) = x_i(0)$ is assigned to the vehicles with types in the range $0 < \theta_i < \theta_i^0$, where $\theta_i^0$ is implicitly defined by:

$$\begin{aligned} &\frac{w}{T_{max} - T^{com} - \frac{c_i s_i}{\theta_i^0 x_i(0)}} \times \frac{c_i s_i}{\theta_i^0 x_i^2(0)} - \\ &2\mu\zeta_i c_i s_i x_i(0) \times \left(\frac{1}{(\theta_i^0)^2 h(\theta_i^0)} + \frac{1}{\theta_i^0}\right) \times (1-\nu) = 0. \end{aligned} \tag{33}$$

On an interval which $u_i > 0$, we have:

$$\begin{aligned} &\frac{d}{d\theta_i}[\frac{\partial}{\partial x_i} U_i^{TP-PW}(x_i(\theta_i), \theta_i) + \nu \frac{\partial}{\partial x_i} R_i^{PW}(x_i(\theta_i), \theta_i)] = \\ &\frac{\partial^2}{\partial x_i^2} U_i^{TP-PW}(x_i(\theta_i), \theta_i) \frac{\partial x_i}{\partial \theta_i} + \nu \frac{\partial^2}{\partial x_i^2} R_i^{PW}(x_i(\theta_i), \theta_i) \frac{\partial x_i}{\partial \theta_i} + \\ &\frac{\partial}{\partial \theta_i}[\frac{\partial}{\partial x_i} U_i^{TP-PW}(x_i(\theta_i), \theta_i) + \nu \frac{\partial}{\partial x_i} R_i^{PW}(x_i(\theta_i), \theta_i)] = 0. \end{aligned} \tag{34}$$

Since $\frac{\partial x_i}{\partial \theta_i} = 0$ for the vehicle with types in range $\theta_i^{j1} < \theta_i < \theta_i^{j2}$, we can rewrite equation (34) for vehicles with type $\theta_i^{j1}$ as follows:

$$\frac{\partial}{\partial \theta_i}[\frac{\partial}{\partial x_i} U_i^{TP-PW}(x_i(\theta_i), \theta_i) + \nu \frac{\partial}{\partial x_i} R_i^{PW}(x_i(\theta_i), \theta_i)] = 0. \tag{35}$$

Given the definitions of $U_i^{TP-PW}$, $S_i(x_i(\theta_i), \theta_i)$, and $E_i(x_i(\theta_i), \theta_i)$, equation (35) can be re-written as:

$$\begin{aligned} &\frac{w c_i s_i}{(T_{max} - T^{com} - \frac{c_i s_i}{\theta_i x_i(\theta_i)})^2 \times \theta_i^3} - \\ &\frac{w x_i(\theta_i)}{(T_{max} - T^{com} - \frac{c_i s_i}{\theta_i x_i(\theta_i)}) \times \theta_i^2} - \\ &2\mu\zeta_i x_i^4(\theta_i)(1-\nu) \frac{\partial}{\partial \theta_i}(\frac{1}{\theta_i^2 h(\theta_i)} + \frac{1}{\theta_i}) = 0. \end{aligned} \tag{36}$$

A set of pairs $(\theta_i^{j1}, x(\theta_i^{j1}))$ can be found by solving equations (36) and (32), simultaneously. Equation (31) holds with equality for vehicles with both types $\theta_i^{j1}$ and $\theta_i^{j2}$. Thus,

$$\begin{aligned} \int_{\theta_i^{j1}}^{\theta_i^{j2}} &[\frac{w}{T_{max} - T^{com} - \frac{c_i s_i}{y x_i(y)}} \times \frac{c_i s_i}{y x_i^2(y)} \\ &- 2\mu\zeta_i c_i s_i x_i(y) \times (\frac{1}{y^2 h(y)} + \frac{1}{y}) \times (1-\nu)] dy = 0. \end{aligned} \tag{37}$$

In this interval, since $u_i = 0$, we have $x_i(y) = x(\theta_i^{j1})$ which is determined in the previous steps. Thus $\theta_i^{j2}$ can be calculated by equation (37).

The solution procedure of sub-problem optimization is summarized in Algorithm 1. This algorithm first determines $\theta_i^0$ as the endpoints of the initial interval in which $u_i = 0$. Also, it determines interior intervals $[\theta_i^{j1}, \theta_i^{j2}]$ in which $u_i = 0$. As shown in Algorithm 1, $x_i(\theta_i)$ can be calculated by this procedure. After all $N$ sub-problems are solved, the central system updates Lagrange multiplier $\nu$ at iteration $(k+1)$ using subgradient method as follows

$$\nu_{k+1} = \nu_k + \alpha_{k+1}[\sum_{i=1}^{N} R_i^{PW} - R_{max}]^+, \tag{38}$$

**Algorithm 1:** The proposed calculus of variation solution procedure of sub-problems that the RSU is faced

**Result:** $x_i(\theta_i)$
define sampling rate $\theta_n$;
define discrete set $\tilde{\Theta}_i$ such that $\tilde{\theta}_i = [0 : \theta_n : \bar{\theta}_i] \in \tilde{\Theta}_i$;
solve problem (33) and calculate $\theta_0^i$ using Newton method [26] ;
solve problems (36) and (32) simultaneously using Newton method and calculate set of pairs$(\theta_i^{j1}, x_i(\theta_i^{j1}))$ ;
$G \leftarrow \{(\theta_i^{j1}, x(\theta_i^{j1}))\}$;
$M \leftarrow |G|$;
$x_i(\tilde{\theta}_i) \leftarrow x_i(0), \quad \forall \tilde{\theta}_i < \theta_i^0$;
**for** $j = 1 : M$ **do**
  solve problem (37) using Newton method and calculate $\theta_i^{j2}$ ;
  $x_i(\tilde{\theta}_i) \leftarrow x(\theta_i^{j1}), \quad \forall \ \theta_i^{j1} < \tilde{\theta}_i < \theta_i^{j2}$;
  **if** $j == 1$ **then**
    solve problem (32) using Newton method and calculate $x_i(\tilde{\theta}_i)$ for each $\tilde{\theta}_i$ such that $\theta_i^0 < \tilde{\theta}_i < \theta_i^{j1}$ ;
  **else**
    solve problem (32) using Newton method and calculate $x_i(\tilde{\theta}_i)$ for each $\tilde{\theta}_i$ such that $\theta_i^{(j-1)2} < \tilde{\theta}_i < \theta_i^{j1}$ ;
  **end**
  **if** $j == M$ **then**
    solve problem (32) using Newton method and calculate $x_i(\tilde{\theta}_i)$ for each $\tilde{\theta}_i$ such that $\theta_i^{(j-1)2} < \tilde{\theta}_i < \bar{\theta}$ ;
  **end**
**end**

where $[\cdot]^+ = max(0, \cdot)$ and $\alpha_k$ is the step size at iteration $k$.

**Assumption 1.** *Sub-gradient step sizes satisfy* $\alpha_k \geq 0$, $\sum_{k=0}^{\infty} \alpha_k = \infty$, $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$, *and* $\alpha_\infty = 0$.

**Lemma 3.** *Lagrange multipliers* $\nu$ *converge to their optimal values.*

*Proof.* According to Assumption 1, the Lagrange multiplier converges to its optimal value and the proof is provided in Proposition 6.3.1 in [27]. □

### B. Approximate Dynamic Programming with Neural Network

In the second algorithm, we use the approximate dynamic programming algorithm with neural network approximators to solve optimization (20) [28]. To utilize this method, we first discretize the dynamic system (20). Thus, the Lagrangian function of problem (20) by considering the coupling constraints (20d) becomes as follows:

$$L^d = (\sum_{i=1}^{N} \sum_{n \in \Theta_i} L_i(x_i(n), u_i(n), \nu)) - \nu R_{max}, \quad (39)$$

where $L_i = U_i^{TP-PW} + \nu R_i^{PW}$ and $\Theta_i = \{\underline{\theta_i}, \dots, \bar{\theta}_i\}$. In the ADP algorithm, vehicle $i$ aims to find $u_i$ which maximizes the following cost-to-go function:

$$V_i(x_i(\theta_i)) = \sum_{n=\theta_i}^{\bar{\theta}_i} L_i(x_i(n), u_i(n), \nu)) - \nu R_{max}. \quad (40)$$

The optimal cost-to-go function must satisfy the following Bellman's equation:

$$V_i^*(x_i(\theta_i)) = \quad (41)$$
$$\max_{u_i(\theta_i)} \left[ L_i(x_i(\theta_i), u_i(\theta_i), \nu) - \nu R_{max} + V_i(x_i(\theta_i) + u_i(\theta_i)) \right].$$

Since equation (41) does not have a closed-form solution, the ADP solves this problem iteratively. In an iterative procedure, the estimation of optimal control input and optimal cost-to-go function of agent $i$ are updated, respectively, as follows [29]:

$$u_i^{k+1}(\theta_i) = arg \max_{u_i(\theta_i)} \quad (42)$$
$$[L_i(x_i(\theta_i), u_i^k(\theta_i), \nu) - \nu R_{max} + V_i^k(x_i(\theta_i) + u_i^k(\theta_i))],$$
$$V_i^{k+1}(\theta_i) = (1 - c^k) V_i^k(\theta_i) + (c^k)^2 \quad (43)$$
$$[L_i(x_i(\theta_i), u_i^{k+1}(\theta_i), \nu) - \nu R_{max} + V_i^k(x_i(\theta_i) + u_i^{k+1}(\theta_i))],$$

where $c^k$ is the learning rate at iteration $k$. The control input (42) and cost-to-go function (43) are approximated by two neural networks named actor and critic, respectively. These two neural networks are as follows [30]:

$$\hat{u_i}^k(\theta_i) = \mu_{i,a}^k(\theta_i) \sigma(x_i(\theta_i)), \quad (44)$$

$$\hat{V_i}^k(\theta_i) = \mu_{i,c}^k(\theta_i) \phi(x_i(\theta_i)), \quad (45)$$

where $\mu_{i,a}^k(\theta_i)$ and $\mu_{i,c}^k(\theta_i)$ are the weight matrices of the actor and critic networks of agent $i$ with type $\theta_i$ at iteration $k$, respectively. $\sigma(x_i(\theta_i))$ and $\phi(x_i(\theta_i))$ are two radial basis functions as general functions approximation. According to the Bellman equation [31], the loss functions of the critic and actor networks are as follows:

$$E_{i,c}^{k+1}(\theta_i) = \frac{1}{2} \Big( (1 - c^k) \hat{V_i}^k(\theta_i) + (c^k)^2 [L_i(x_i(\theta_i), \hat{u_i}^{k+1}(\theta_i), \nu)$$
$$- \nu R_{max} + \hat{V_i}^k(x_i(\theta_i) + u_i^{k+1}(\theta_i))] - \hat{V_i}^{k+1}(\theta_i) \Big), \quad (46)$$

$$E_{i,a}^{k+1}(\theta_i) = \frac{1}{2} \Big( arg \max_{u_i(\theta_i)} [L_i(x_i(\theta_i), u_i^k(\theta_i), \nu) - \nu R_{max}$$
$$+ \hat{V_i}^k(x_i(\theta_i) + u_i^k(\theta_i))] - \hat{u_i}^{k+1}(\theta_i) \Big). \quad (47)$$

The weights of the actor and critic networks are updated by the gradient-based algorithm as follows:

$$\mu_{i,a}^{k+1,j+1}(\theta_i) = \mu_{i,a}^{k+1,j}(\theta_i) - \beta_a \frac{\partial E_{i,c}^{k+1}(\theta_i)}{\partial \hat{u_i}^{k+1}(\theta_i)} \frac{\partial \hat{u_i}^{k+1}(\theta_i)}{\partial \mu_{i,c}^{k+1,j}(\theta_i)}, \quad (48)$$

$$\mu_{i,c}^{k+1,j+1}(\theta_i) = \mu_{i,c}^{k+1,j}(\theta_i) - \beta_c \frac{\partial E_{i,c}^{k+1}(\theta_i)}{\partial \hat{V_i}^{k+1}(\theta_i)} \frac{\partial \hat{V_i}^{k+1}(\theta_i)}{\partial \mu_{i,a}^{k+1,j}(\theta_i)}, \quad (49)$$

where $j$ is the iteration step for updating the neural networks' weights. $\beta_a$ and $\beta_c$ are the learning rate parameters of actor and critic networks, respectively. Then, using the updated control variable, $R_i^{PW}$ is calculated and the Lagrange multiplier $\nu$ at iteration $(k+1)$ is updated by using equation (38). Algorithm 2 summarizes the iterative solution to solve optimization (20) using ADP and dual decomposition methods.

Based on Theorems 3 and 4 in [29], for a sufficiently large number of neurons of neural networks, the weights of actor and critic neural networks converge to their optimal values if $\beta_a \leq \frac{2}{||\sigma(x_i(\theta_i))||^2}$ and $\beta_c \leq \frac{2}{||\phi(x_i(\theta_i))||^2}$. Also according to this assumption and assumption 1, the control input converges to its optimal value and the algorithm converges to the optimal solution of optimization (20).

## VI. Performance Evaluation

To evaluate the performance of our proposed approach, we carry out a traffic sign classification task through federated learning on a dataset named "Belgium Traffic Sign Dataset" [21] which contains 4591 training samples and 2534 test samples of 62 different sign types. There exist 100 moving vehicles with cameras as data owners in the federated learning task. The simulations are provided by MATLAB R2021a and performed on a Desktop PC with 12 GB of RAM and a 1.80 GHz processor. We consider uniform distribution on the interval $[2, 10]$ for types of all vehicles. We set $\epsilon_i \in [0.5, 0.91]$, $c_i = 5$, $s_i = 46$, $T_i^{com} = 10$ and $E_i^{com} = 20$ for $i = \{1, \dots, 100\}$. The maximum time of task and total amount of given reward are 10, and 300, respectively. Also, we consider 20 neurons in neural network [32], [17].

To compare the performance of ADP and calculus of variation algorithms for solving optimization (20), Figure 2 and Figure 3 are drawn. These figures show the convergence of the RSU's total reward paid to the vehicles and expected computational resource of vehicle $i$. Based on these figures, both total reward (i.e. $\sum_{i=1}^{N} \boldsymbol{E_{\theta_i}}[R_i(\theta_i)]$) and expected computational

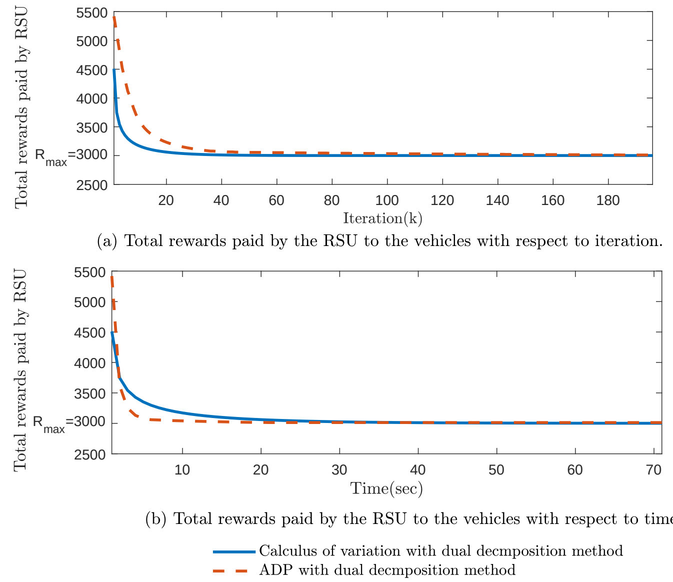
**Algorithm 2:** Combination of ADP and dual decomposition methods for solving the functional optimization

**Result:** $x(\theta)$ , $\nu$
**while** $||\nu_k - \nu_{k-1}|| > \epsilon$ **do**
  **for** $i = 1 : N$ **do**
    **for** $\theta_i = \underline{\theta_i} : \bar{\theta}_i$ **do**
      **while** $||\mu_{i,c}^{k,j} - \mu_{i,c}^{k,j-1}|| > \epsilon$ *or* $||\mu_{i,a}^{k,j} - \mu_{i,a}^{k,j-1}|| > \epsilon$ **do**
        Update $\mu_{i,c}^{k,j+1}$ and $\mu_{i,a}^{k,j+1}$ using (48) and (49), respectively.;
        $\mu_{i,c}^{k,j} \leftarrow \mu_{i,c}^{k,j+1}$;
        $\mu_{i,c}^{k,j} \leftarrow \mu_{i,a}^{k,j+1}$;
        $j = j + 1$
      **end**
      Compute $\hat{u_i}^k(\theta_i)$ using (44);
      Compute $\hat{v_i}^k(\theta_i)$ using (45);
      Apply $\hat{u_i}^k(\theta_i)$ to the system and obtain the next state.
    **end**
  **end**
  Compute $\sum_{i=1}^{N} \sum_{\theta_i=\underline{\theta_i}}^{\bar{\theta}_i} R_i^{PW}(\theta_i)$ and update $\nu_{k+1}$ using (38);
  $\nu_k \leftarrow \nu_{k+1}$;
  $k = k + 1$
**end**

resource (i.e. $E_{\theta_i}[x_i(\theta_i)]$) converge to the same value in both of the calculus of variation and the ADP methods. As it was expected, after some iterations of dual decomposition, the RSU's total reward becomes less than the total budget of the RSU in both algorithms, which means that constraint in (9d) is satisfied. According to Figure 2(a) and Figure 3(a), the number of iterations needed for convergence in the calculus of variation method is less than that of the ADP method. Specifically, the calculus of variation and the ADP methods reach the optimal solution in about 72 and 196 iterations, respectively. However, since each iteration of the calculus of variation takes more CPU time than that of the ADP method, based on Figures 2(b) and 3(b), convergence time in the ADP method is less than the calculus of variation method.

To validate the IR and IC constraints in the proposed scheme, the utilities of five specific types of vehicle $i$ when it announces different types are depicted in Figure 4. The black stars marked on the curves indicate the maximum points of utility function for each type of vehicle. Figure 4 shows that when a vehicle as a vehicle announces its own type truthfully, it gains a positive and maximal utility. However, deceiving the RSU by announcing an unreal type causes a loss for the vehicle. Such results verify that the optimal strategies of the proposed mechanism satisfy IR and IC constraints.

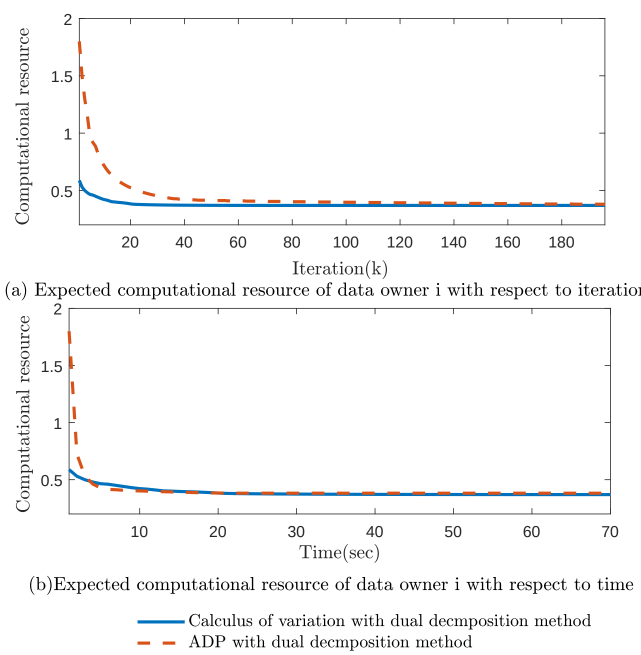
We evaluate the utility of vehicles and RSU by comparing the proposed contract scheme with the discriminatory contract scheme and the contract scheme without type verification. The discriminatory contract scheme is considered to be a baseline where the RSU has full knowledge about the vehicles [33]. In




(a) Total rewards paid by the RSU to the vehicles with respect to iteration.

(b) Total rewards paid by the RSU to the vehicles with respect to time

Fig. 2. Total rewards paid by the RSU to the vehicles




(a) Expected computational resource of data owner i with respect to iteration

(b)Expected computational resource of data owner i with respect to time

Fig. 3. Expected computational resource of vehicle $i$

the contract scheme without type verification, the RSU does not know the types of vehicles and mistakenly assumes that vehicles announce their types truthfully. However, the vehicles as selfish agents may announce their type falsely to profit more [33]. Note that as shown in Figures 2 and 3, the result of ADP and calculus of variation algorithms are very close to each other. Thus, we only plot the ADP algorithm results as the proposed method in the following figures.

The performance of these three schemes is illustrated in Table I, Figure 5, and Figure 6. In the discriminatory scheme, the mechanism is designed under complete information about the vehicles such that the task publisher maximizes its utility by minimizing the utility of the vehicles. As expected, in

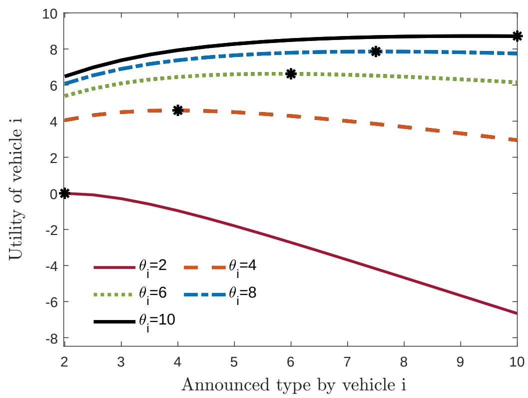



Fig. 4. Utilities of 5 types of vehicle $i$ ($\theta_i = 2, 4, 6, 8, 10$) when announcing different types as their types.

this case, the highest profit for the RSU and zero utility for vehicles are observed. However, the utility of RSU in the proposed scheme is very close to the baseline scheme which shows the high performance of the proposed scheme. As shown in Table I, the utility of the RSU in the proposed scheme is higher than that in the contract scheme without type verification. The reason is that untruthful reporting the types by vehicles makes RSU to be deceived and achieves less utility than those in the other schemes. Figure 5 also shows that $U_i^D(\underline{\theta}, \underline{\theta}) = 0$ which is consistent with Lemma1. Figure 6 shows both the optimal computational resource and the optimal incentive reward increase with the vehicles' type in the proposed scheme. Therefore, the higher the local data quality of the vehicle, the more the RSU allocates computational resource and incentive reward. Figure 6(a) also presents that the optimal computational resource increases with the vehicle type, which is consistent with $x_i'(\theta_i) \geq 0$ in Lemma 2.

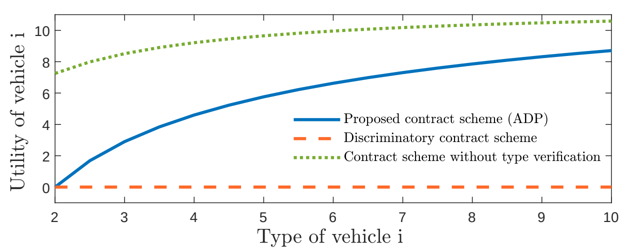



Fig. 5. Utility of vehicle $i$ under different schemes.

TABLE I
UTILITY OF RSU UNDER DIFFERENT SCHEMES.

| Scheme | RSU's utility |
|---|---|
| Proposed contract scheme (ADP) | 498890 |
| Discriminatory contract scheme | 501783 |
| Contract scheme without type verification | 458095 |

## VII. CONCLUSION

We studied optimal mechanism strategies for the federated learning task over the wireless vehicular network with information asymmetry and captured the interactions among the RSU and vehicles. We considered vehicles may not reveal their private information truthfully to achieve more payoff. Moreover, we assumed the distribution of each agent's private information can be different from the distribution of other agents' private information. The RSU as a task publisher designs a mechanism based on contract theory that contains its strategies including the computation resource and the corresponding incentive reward toward all private information of vehicles. To solve the constrained functional optimization problem induced by the proposed incentive mechanism, two algorithms were proposed. In the first algorithm, we used calculus of variation theory and dual decomposition while in the second algorithm ADP and dual decomposition were used.

Numerical results verified theoretical results and showed that both of these algorithms converge to the same optimal solution. The calculus of variation algorithm outperforms in terms of the number of iterations. However, due to the large time required for each iteration in this algorithm, it requires more execution time than that of the proposed ADP algorithm.

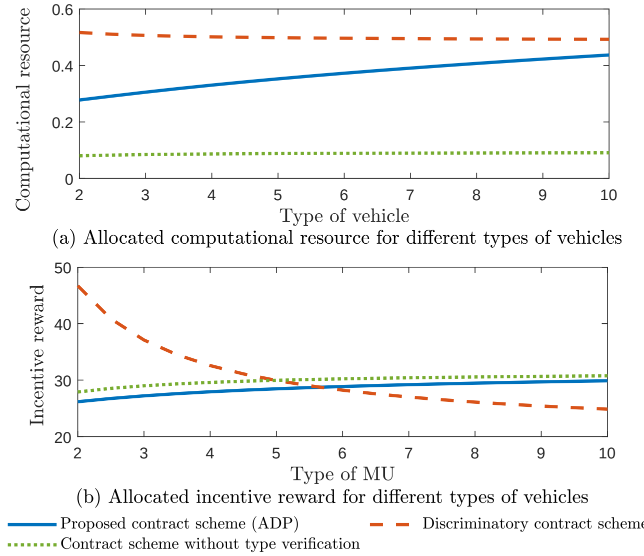



Fig. 6. Utility of vehicle $i$ under different schemes.

## REFERENCES


[1] H. Cao, J. Du, H. Zhao, D. X. Luo, N. Kumar, L. Yang, and F. R. Yu, "Resource-ability assisted service function chain embedding and scheduling for 6g networks with virtualization," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3846–3859, 2021.

[2] H. Shu, T. Liu, X. Mu, and D. Cao, "Driving tasks transfer using deep reinforcement learning for decision-making of autonomous vehicles in unsignalized intersection," *IEEE Transactions on Vehicular Technology*, 2021.

[3] Y. Li, X. Tao, X. Zhang, J. Liu, and J. Xu, "Privacy-preserved federated learning for autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[4] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, 2017, pp. 1273–1282.

[5] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 2031–2063, 2020.

[6] Federated machine learning for ai self-driving cars. [Online]. Available: https://www.aitrends.com/ai-insider/federated-machine-learning-for-ai-self-driving-cars/

[7] A. Ali, I. Ilahi, A. Qayyum, I. Mohammed, A. Al-Fuqaha, and J. Qadir, "Incentive-driven federated learning and associated security challenges: A systematic review," 2021.

[8] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, 2021.

[9] N. H. Tran, W. Bao, A. Zomaya, M. N. Nguyen, and C. S. Hong, "Federated learning over wireless networks: Optimization model design and analysis," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 1387–1395.

[10] Y. Wang, G. Gui, H. Gacanin, B. Adebisi, H. Sari, and F. Adachi, "Federated learning for automatic modulation classification under class imbalance and varying noise condition," *IEEE Transactions on Cognitive Communications and Networking*, 2021.

[11] H. Kim, J. Park, M. Bennis, and S.-L. Kim, "Blockchained on-device federated learning," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1279–1283, 2019.

[12] Z. Zhou, P. Liu, J. Feng, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Computation resource allocation and task assignment optimization in vehicular fog computing: A contract-matching approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3113–3125, 2019.

[13] Y. Zhan, P. Li, Z. Qu, D. Zeng, and S. Guo, "A learning-based incentive mechanism for federated learning," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6360–6368, 2020.

[14] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, "Algorithmic game theory, 2007," *Google Scholar Digital Library Digital Library*.

[15] S. Kim, "Incentive design and differential privacy based federated learning: A mechanism design perspective," *IEEE Access*, vol. 8, 2020.

[16] T. H. T. Le, N. H. Tran, Y. K. Tun, M. N. Nguyen, S. R. Pandey, Z. Han, and C. S. Hong, "An incentive mechanism for federated learning in wireless cellular networks: An auction approach," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4874–4887, 2021.

[17] 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 of Things Journal*, vol. 6, no. 6, pp. 10 700–10 714, 2019.

[18] N. Ding, Z. Fang, and J. Huang, "Optimal contract design for efficient federated learning with multi-dimensional private information," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, 2020.

[19] J. Mirrlees, "Information and incentives: The economics of carrots and sticks," *The Economic Journal*, vol. 107, no. 444, pp. 1311–1329, 1997.

[20] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Differentially private asynchronous federated learning for mobile edge computing in urban informatics," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 2134–2143, 2019.

[21] R. Timofte, K. Zimmermann, and L. Van Gool, "Multi-view traffic sign detection, recognition, and 3d localisation," *Machine vision and applications*, vol. 25, no. 3, pp. 633–647, 2014.

[22] A. Richardson, A. Filos-Ratsikas, and B. Faltings, "Rewarding high-quality data via influence functions," *arXiv:1908.11598*, 2019.

[23] Y. Zhang, L. Liu, Y. Gu, D. Niyato, M. Pan, and Z. Han, "Offloading in software defined network at edge with information asymmetry: A contract theoretical approach," *Journal of Signal Processing Systems*, vol. 83, no. 2, pp. 241–253, 2016.

[24] A. Falsone, K. Margellos, S. Garatti, and M. Prandini, "Dual decomposition for multi-agent distributed optimization with coupling constraints," *Automatica*, vol. 84, pp. 149–158, 2017.

[25] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[26] K. E. Atkinson, *An introduction to numerical analysis*. John wiley & sons, 2008.

[27] D. P. Bertsekas, W. Hager, and O. Mangasarian, "Nonlinear programming. athena scientific belmont," *Massachusets, USA*, 1999.

[28] T. Bian, Y. Jiang, and Z.-P. Jiang, "Adaptive dynamic programming and optimal control of nonlinear nonaffine systems," *Automatica*, vol. 50, no. 10, pp. 2624–2632, 2014.

[29] P. Rokhforoz, H. Kebriaei, and M. N. Ahmadabadi, "Large-scale dynamic system optimization using dual decomposition method with approximate dynamic programming," *Systems & Control Letters*, vol. 150, 2021.

[30] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, "Discrete-time nonlinear hjb solution using approximate dynamic programming: Convergence proof," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 4, pp. 943–949, 2008.

[31] S. Peng, "A generalized dynamic programming principle and hamilton-jacobi-bellman equation," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 38, no. 2, pp. 119–134, 1992.

[32] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, and J. Zhao, "Toward secure blockchain-enabled internet of vehicles: Optimizing consensus management using reputation and contract theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2906–2920, 2019.

[33] M. Montazeri, H. Kebriaei, B. N. Araabi, and D. Niyato, "Optimal mechanism design in the sponsored content service market," *IEEE Communications Letters*, vol. 25, no. 9, pp. 3051–3054, 2021.