

LeFi: Learn to Incentivize Federated Learning in Automotive Edge Computing

Ming Zhao, Yuru Zhang, Qiang Liu
 University of Nebraska-Lincoln
 qiang.liu@unl.edu

Tao Han
 New Jersey Institute of Technology
 tao.han@njit.edu

Abstract—Federated learning (FL) is the promising privacy-preserve approach to continually update the central machine learning (ML) model (e.g., object detectors in edge servers) by aggregating the gradients obtained from local observation data in distributed connected and automated vehicles (CAVs). The incentive mechanism is to incentivize individual selfish CAVs to participate in FL towards the improvement of overall model accuracy. It is, however, challenging to design the incentive mechanism, due to the complex correlation between the overall model accuracy and unknown incentive sensitivity of CAVs, especially under the non-independent and identically distributed (Non-IID) data of individual CAVs. In this paper, we propose a new learn-to-incentivize algorithm to adaptively allocate rewards to individual CAVs under unknown sensitivity functions. First, we gradually learn the unknown sensitivity function of individual CAVs with accumulative observations, by using compute-efficient Gaussian process regression (GPR). Second, we iteratively update the reward allocation to individual CAVs with new sampled gradients, derived from GPR. Third, we project the updated reward allocations to comply with the total budget. We evaluate the performance of extensive simulations, where the simulation parameters are obtained from realistic profiling of the CIFAR-10 dataset and NVIDIA RTX 3080 GPU. The results show that our proposed algorithm substantially outperforms existing solutions, in terms of accuracy, scalability, and adaptability.

Index Terms—Federated Learning, Incentive Mechanism, Vehicular Networks, Edge Computing

I. INTRODUCTION

Federated learning is the promising paradigm to achieve large-scale privacy-preserve distributed machine learning. To collaboratively train a global model, a generic federated learning system involves massive participants (which compute the model update with their local data) and a central parameter server, which aggregates distributed updates in the global model, such as FedAvg [1]. Instead of sharing the raw training data, only computed model updates (e.g., gradients) will be transmitted to the central server, which substantially alleviates the data privacy concerns and improves communication efficiency. Hence, there are ever-increasing research efforts in exploring and exploiting the advantages of federated learning in a wide range of application domains, such as healthcare, finance, and retail [2].

The incentive mechanism (i.e., incentivizing massive participants to contribute their data) plays a key role in achieving the goal of federated learning [3]. This is based on the observation that individual participants are generally reluctant to participate in federated learning without monetary incentives. On the one hand, the local training consumes non-trivial computation resources, which would squeeze the limited computation capacity of participants (usually edge devices). On the other hand, the transmission of the computed

model updates requires non-negligible mobile traffic, which may increase the data usage of participants (usually mobile devices). As a result, the sensitivity function of participants is different concerning monetary incentives. In addition, recent findings show that not all participants are equally important, in terms of improving the overall accuracy of the global model, because of their heterogeneous data quantity and quality [4], such as independently and identically distributed (IID).

There are extensive existing works in designing incentive mechanisms in the domain of federated learning [3], [5]. However, these works generally assume that the information of local data (e.g., quality and quantity) in participants is known by the central server, while this assumption may not always hold. For example, participants can be selfish, and optimize their contribution (e.g., data and computation) dynamically in each federated learning round. Recent works focus on using model-free approaches (e.g., deep reinforcement learning [6], [7]) to learn and incentivize participants without the need for prior knowledge. However, they typically use deep neural networks (DNNs) to parameterize their policies, which tend to be low sample efficiency and can hardly scale under dynamic participants over time. Therefore, it is imperative to design a sample-efficient and model-free approach to incentivize diverse participants in federated learning.

In this paper, we propose the *LeFi* algorithm to learn and incentivize participants, i.e., connected and automated vehicles (CAVs), in automotive edge computing. The optimization objective is to maximize the overall accuracy of the global model under the total monetary budget, by optimizing the reward weight of individual participating CAVs. First, we develop a sample-efficient surrogate model to approximate the unknown sensitivity function of participants, according to the accumulative Server-to-CAV interactions over time. Second, we design a sample-based gradient descent method to update the reward weight of individual CAVs and project the reward weights under the total budget if applicable. Third, we use Karush-Kuhn-Tucker (KKT) conditions to derive the data selection in each CAV independently. Then, the central server will receive the actual data size of local training in each CAV, which will be used to update the surrogate model continuously. We evaluate the proposed algorithm via extensive simulations, where the simulation parameters are obtained by profiling a Convolutional neural network (CNN)-based model under the CIFAR-10 dataset and Nvidia RTX 3080 GPU. The results show that our proposed algorithm substantially outperforms existing solutions, in terms of convergence, accuracy, and scalability.

II. SYSTEM MODEL

We consider a generic federated learning system, including a central server and a massive number of CAVs denoted as $\mathcal{N} = \{1, 2, \dots, N\}$. To train the global DNN model, e.g., object detection and segmentation, CAVs perceive surrounding environments with diverse onboard sensors (e.g., LiDAR and cameras) and are responsible for the local training process before the calculated gradients are sent to the central server. The central server will aggregate the gradients received from CAVs, e.g., averaging, to update the global DNN model. The global model will be then disseminated to the participating CAVs, e.g., changed parameters, which creates the next round of federated learning.

Non-IID Data. Recent findings [4] show that the distribution of training data of individual participants (i.e., CAVs) would generate a non-trivial impact on the overall accuracy of the global model. The training data is IID, only if its samples are drawn from the same probability distribution, and are independent of one another statistically. However, the diversified trajectories of individual CAVs will naturally create non-IID training data. Specifically, Non-IID data can be characterized by the allocation of a proportion of samples for each label to each party according to a Dirichlet distribution [8]. Without loss of generality, we introduce the variable $\pi \in [0, 1]$ to quantify the degree of Non-IID in the training data. A smaller π value indicates a higher degree of Non-IID, signifying greater disparity in the proportion of samples for each label across parties. In contrast, as π approaches 1, it suggests that each party possesses an identical proportion of the samples for each label.

Accuracy Model. The overall accuracy of the global model is related to not only the non-IID severity but also the size of training data in individual participating CAVs. In Fig. 2, we show the correlations by training a CNN model over the CIFAR-10 dataset (in total 50K training data) with 5 CAVs (different data sizes and same non-IID severities). It can be seen that higher accuracy can be generally achieved by increasing the data size under lower non-IID severity. Without loss of generality, the accuracy model can be built by fitting these data points with various regression models, such as linear, polynomial and exponential. However, the non-IID severity of training data in CAVs is mostly different, which generates intractable complexity in regressing the accuracy model, especially under thousands of CAVs, if not more. For the sake of tractability, we approximately formulate the overall accuracy of the global model as the combination of the individual CAVs' accuracy

$$A_{mean} = \frac{w}{N} \sum_{n \in \mathcal{N}} A(D_n, \pi_n), \quad (1)$$

where w is the weight, which relates to not only the distribution of non-IID severity among all CAVs but also the number of participating CAVs. Here, $A(D_n, \pi_n)$ is the accuracy function of the n -th CAV, which may be regressed with tractable profiling data points.

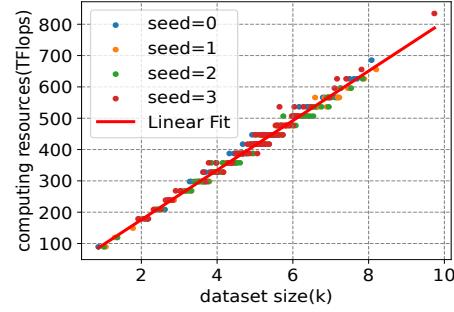


Fig. 1: Computing resources needs in different dataset sizes

Incentive Mechanism. To improve the overall accuracy of the global model, it is essential to incentivize the contribution of selfish CAVs, in terms of computation resources and unique training data. As we quantified the aforementioned accuracy model, we designed the incentive mechanism by offering monetary rewards to individual CAVs. In particular, we define the reward function for the n -th CAV as

$$f(D_n) = \alpha_n (1 - e^{-\beta_n D_n}), \quad (2)$$

where α_n denotes the reward weight, D_n is the data size of the CAV, and β_n is a pre-defined non-negative hyperparameter. The rationale behind this is that, by optimizing and allocating the reward cap (i.e., α_n), the reward function $f(D_n)$ will incentivize the n -th CAV to contribute more training data. In addition, using the exponential-related functions (rather than linear) can represent a wide range of correlations between the size of training data and the reward.

Selfish Participating CAVs. The local training process and sharing of data require non-negligible costs in each CAV, e.g., computing resources and mobile traffic. Hence, it is reasonable to assume that each CAV is selfish in participating the federated learning, e.g., monetary rewards to compensate for local resource consumption. In particular, we profile the computational complexity of the local training process under different training data sizes in Fig. 1, by using the metric of TFLOPS (Tera Floating-point operations per second). By fitting the profiled data point with linear regression, we obtain the computation complexity $C(D_n) = pD_n + q$, where $p = 79.1259$ and $q = 17.6219$.

Client-Side Problem. On the client side, the objective is to maximize the revenue (i.e., the received reward minus its cost), by optimizing the data size for local training. Therefore, we formulate the client-side optimization problem of data selection in the n -th CAV as follows

$$\begin{aligned} \mathbb{P}_0 : \quad & \max_{D_n} \alpha_n (1 - e^{-\beta_n D_n}) - \theta \cdot C(D_n) \\ \text{s.t.} \quad & \begin{cases} C(D_n)/F_n \leq T_{max}, \\ D_n \leq D_{max}, \end{cases} \end{aligned} \quad (3)$$

where we introduce the latency requirement by dividing the computation complexity of local training and the computation capacity of the n -th CAV (i.e., F_n). Here, T_{max} is the maximum latency allowance and D_{max} is the maximum data size. In addition, θ is the rate per computation capability, e.g.,

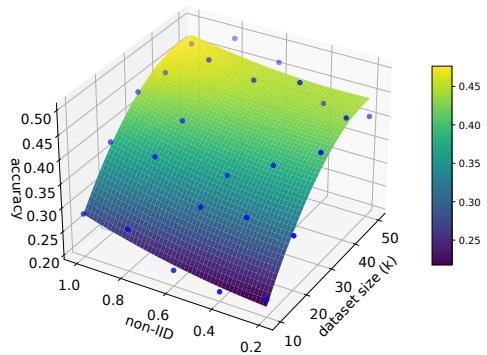


Fig. 2: Model accuracy varies in different dataset sizes and Non-IID severity

per TFLOPS. Note that, the reward weight α is given by the server side, when solving this client-side problem.

Server-Side Problem. On the server side, the objective is to maximize the overall accuracy under the given budget, by optimizing the reward weight of individual CAVs. Therefore, we formulate the server-side optimization problem of incentive design as follows

$$\begin{aligned} \mathbb{P}_1 : \max_{\{\alpha_n, n \in \mathcal{N}\}} \quad & \sum_{n \in \mathcal{N}} A(D_n(\alpha_n), \pi_n) \\ \text{s.t.} \quad & \sum_{n \in \mathcal{N}} \alpha_n (1 - e^{-\beta_n D_n}) \leq M_{max}, \end{aligned} \quad (4)$$

where we exclude the w, N in Eq. 1 without affecting the overall optimization problem, and M_{max} is the total budget. Here, we consider the non-IID severity π_n is known by the n -th CAV and fixed in each federated learning round. Note that, we denote **sensitivity function** as $D_n(\alpha_n)$, which represents the size of training data in the n -th CAV under a given reward weight α_n . This is because, selfish CAVs will determine their contributed size of training data independently, where the underlying calculations are dependent on various factors, e.g., user sensitivity of rewards and computation capacities. From the perspective of the server side, the sensitivity function of CAVs is unknown.

Challenges. The technical challenge of solving the aforementioned problems lies in the unknown sensitivity function of CAVs from the perspective of the server side. As a result, the server-side problem falls into the black-box optimization domain, where conventional linear, nonlinear, and convex optimization methods cannot work directly. With the massive number of CAVs in federated learning, it is intractable to employ exhaustive search techniques in practice. Although Bayesian optimization provides a generic framework to solve blackbox problems, its scalability remains uncertain under thousands of CAVs (if not more) in federated learning.

III. PROPOSED SOLUTION

In this section, we introduce the Learn-to-incentivize Federated learning (LeFi) algorithm, which allows the server to iteratively learn to allocate reward weights to individual CAVs without prior knowledge. Different from vanilla Bayesian optimization, the *LeFi* algorithm works by following

three steps. First, we use a sample-efficient surrogate model to approximate the size of training data under different reward weights for each CAV. Second, we design a sample-based gradient descent method to update the reward weight of individual CAVs. Third, we project the reward weights under the total budget. The server will receive the actual data size of local training in each CAV, which will be used to update the surrogate model continuously. These steps iterate until the convergence of the algorithm.

A. Unknown Function Approximation

Here, the goal is to find a sample-efficient surrogate model to learn and fit the unknown sensitivity function. The surrogate model can be updated via accumulative query-and-observation. In particular, we focus on learning the sensitivity function of individual CAVs, rather than the overall accuracy model. Although linear and polynomial regression are generally both compute- and sample- efficient, their approximation capabilities are limited to represent complex and unknown functions in selfish CAVs. In contrast, deep learning (DL) utilizes multiple neurons and multi-layer neural networks to effectively fit nonlinear and high-dim complex functions. Nonetheless, DL practically requires a substantial amount of training datasets, with a low sample efficiency.

Alternatively, GPR offers several advantages in terms of sample-efficiency and approximation capability. Unlike linear regression, GPR does not assume a predefined form of the relationship between inputs and outputs, allowing for the capture of complex, non-linear patterns. This non-parametric approach leverages a covariance function, or kernel, which encapsulates assumptions about the function to be learned. In addition, GPR can provide a probabilistic prediction along with the uncertainty of the estimation, which is crucial for decision-making processes.

The Gaussian process is defined by its mean function $m(x)$ and covariance function $k(x, x')$, which is expressed as:

$$m(x) = \mathbb{E}[f(x)], \quad (5)$$

$$k(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))], \quad (6)$$

where $f(x)$ represents the unknown function to learn. The choice of the kernel function k is pivotal as it determines the smoothness and other properties of the functions drawn from the process. The resultant predictive distribution for a new input x^* is then given by a normal distribution characterized by a mean $\mu(x^*)$ and variance $\sigma^2(x^*)$, making GPR a robust tool that inherently quantifies the uncertainty of its predictions.

With the GPR, we define its input as the reward weight and the output as the resulting data size per CAV. In each iteration, the GPR of CAVs is updated by using all the accumulated reward-to-size pairs in the previous iterations.

B. Reward Allocation Updating

Here, the goal is to conduct a one-step update on the reward weight of all CAVs given the current GPRs, towards the optimal solution.

First, we deviate the last reward weight α_n with Δ , and predict its mean and std by inferring the GPR as

$$\mu, \sigma = GPR(\alpha_n + \Delta), \quad (7)$$

where μ and σ are the mean and std, respectively.

Second, we estimate the gradient at the last reward weight α_n as

$$\frac{\partial D_n}{\partial \alpha_n} = \frac{Normal(\mu, \sigma) - GPR(\alpha_n)}{\delta}, \quad (8)$$

where we sample from the Normal distribution $Normal(\mu, \sigma)$. Here, we use the sampled value (rather than the mean) to incorporate its variance in the following gradient descent.

Third, we calculate the gradient by following

$$\nabla \bar{g} = \frac{\partial A_n}{\partial \alpha_n} = \frac{\partial A_n}{\partial D_n} \cdot \frac{\partial D_n}{\partial \alpha_n}, \quad (9)$$

where,

$$\frac{\partial A_n}{\partial D_n} = 2aD_n + c\pi_n + d, \quad (10)$$

where a , c , and d are constants obtained from polynomial regression of accuracy function in CAVs.

Fourth, we update the last reward weight as follows

$$\alpha_n^{t+1} = \alpha_n^t + \eta \cdot \nabla \bar{g}, \quad (11)$$

Fifth, we project the generated α_n^{t+1} under the total budget (if applicable) as follow

$$\alpha_n^p = M_{max} / \sum_{n \in N} \alpha_n (1 - e^{-\beta_n D_n}), \quad (12)$$

which happens only if $\sum_{n \in N} \alpha_n (1 - e^{-\beta_n D_n}) \geq M_{max}$.

C. Selfish Dataset Selection

From the perspective of individual CAVs, optimizing the data size of local training is a key approach to maximizing its own revenue. Selecting a larger data size can generally improve the received reward from the server side, but this also introduces additional costs, which could potentially decrease the revenue. Here, we employ the KKT conditions to identify the optimal solution for training data size.

Initially, we establish the Lagrangian function with the introduction of Lagrange multipliers to account for the imposed constraints. Let λ_1 be the Lagrange multiplier for the first constraint, and λ_2, λ_3 be the multipliers for the inequality constraints $D_{min} \leq D_n$ and $D_n \leq D_{max}$, respectively. Moreover, the Lagrange multipliers must fulfill the non-negativity conditions, $\lambda_1, \lambda_2, \lambda_3 \geq 0$.

The Lagrangian, \mathcal{L} , can be articulated as follows

$$\begin{aligned} \mathcal{L}(D_n, \lambda_1, \lambda_2, \lambda_3) = & \alpha_n (1 - e^{-\beta_n D_n}) - \theta \cdot (aD_n + b) \\ & + \lambda_1 \left(T_{max} - \frac{aD_n + b}{F_n} \right) \\ & + \lambda_2 (D_n - D_{min}) \\ & - \lambda_3 (D_n - D_{max}), \end{aligned} \quad (13)$$

Then, we differentiate \mathcal{L} with respect to D_n and set the derivative equal to zero

$$\frac{\partial \mathcal{L}}{\partial D_n} = -\alpha_n \beta_n e^{-\beta_n D_n} - \theta \cdot a - \lambda_1 \frac{a}{F_n} + \lambda_2 - \lambda_3, \quad (14)$$

Algorithm 1: The *LeFi* algorithm

Input: $M_{max}, F_n, \theta_n, \pi_n, D_{min}, D_{max}, \eta$
Output: D_n, α_n

- 1 Initialize the reward weight α_n and warm start the GPR model; **for** $t = 0, 1, \dots, Rounds$ **do**
- 2 $\text{// } * * \text{ GPR Prediction } * * /$
- 3 Build training dataset with $[\alpha, D]$ and train GPR model;
- 4 Get mean $\mu(D_n)$ and std σ of deviated dataset size $\alpha + \Delta$ from GPR;
- 5 Sample dataset size from normal distribution $Normal(\mu, \sigma)$;
- 6 $\text{// } * * \text{ Sample-based Gradient Descent } * * /$
- 7 Update reward weight α_n based on Eq. 11;
- 8 $\text{// } * * \text{ Project Reward Weight } * * /$;
- 9 **if** $\sum_{n \in N} \alpha_n (1 - e^{-\beta_n D_n}) \geq M_{max}$ **then**
- 10 Project reward weights based on Eq. 12;
- 11 **return** optimal D_n and α_n for all CAVs;

For the inequality constraints, the following conditions must be satisfied

$$\begin{aligned} \lambda_1 \left(T_{max} - \frac{aD_n + b}{F_n} \right) &= 0, \\ \lambda_2 (D_n - D_{min}) &= 0, \\ \lambda_3 (D_n - D_{max}) &= 0. \end{aligned} \quad (15)$$

Note that, the data selection problem in individual CAVs is not known by the server side. Our proposed algorithm makes no assumption on data selection methods and can be easily extended to accommodate other methods in selfish CAVs.

D. The *LeFi* Algorithm

Based on the above analysis, we introduce the *Learn-to-incentivize Federated learning (LeFi)* algorithm, whose pseudocode is provided in Alg. 1.

The process begins with the initialization of reward weight α_n . Using the α and the optimal dataset size D obtained from the KKT conditions, we construct the training set and train the GPR of each CAV. Next, the GPR model predicts with the last deviated $\alpha_n + \Delta$ and uses the mean and std as parameters to randomly generate $D_n(\alpha_n + \Delta)$ followed by the Normal distribution. Then, the gradient is calculated and the reward weight of all CAVs is updated accordingly. If the reward allocated exceeds the given budget, the reward weights are projected back into the range of the budget. In each iteration, all previously observed values of D_n and α_n are used to train the GPR, which would gradually improve its approximation accuracy towards the convergence of the algorithm.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the proposed algorithm with extensive simulations. To obtain the simulation parameters, we profile a CNN-based model under the CIFAR-10 dataset over a desktop with NVIDIA 3080 GPU whose computation capacity

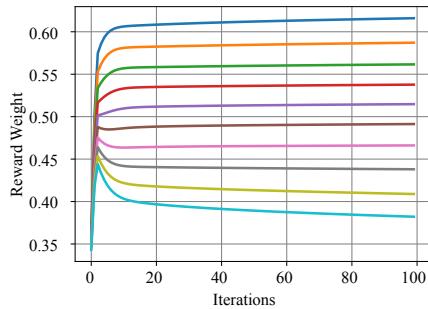


Fig. 3: Convergence of rewards weights.

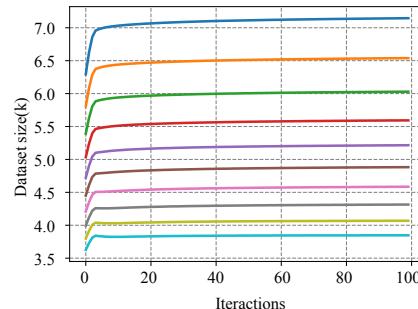


Fig. 4: Convergence of data selection.

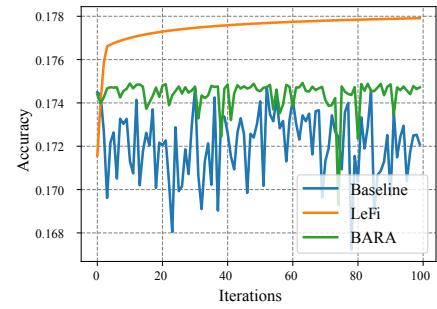


Fig. 5: Convergence of accuracy.

is 28.9 FLOPS. All the regression models are obtained based on the profiled measurement, including the overall accuracy and the computation complexity. Based on the pricing of GPU rental services in the market [9], we set the cost weight θ by generating variations by varying from 0.5 to 2 times 0.9 USD. Besides, the non-IID severity is uniformly generated between 0.1 and 1.0 for all CAVs. The maximum training time of each round is 200s by default. The maximum size of training data in each CAV is 10K entries. We conduct 2-order polynomial regression to obtain the accuracy model $A(D_n, \pi_n) = aD_n^2 + b\pi_n^2 + cD_n\pi_n + dD_n + e\pi_n + f$ with root mean square error (RMSE) 0.03, whose parameters are $a = -0.000152, b = 0.071, c = -0.00117, d = 0.0151, e = 0.011, f = 0.073$. The default number of CAVs is 10, the budget is 5 USD, and the latency requirement is 1 minute.

We compared with the following algorithms¹:

- **Baseline:** the *Baseline* method uses random search to find the combination of reward weight of all CAVs, which serves as the baseline performance.
- **BARA:** the *BARA* method [7] uses Bayesian optimization to optimize the reward weight of all CAVs, where a GPR is created to approximate the correlation between all the weights of CAVs and the overall accuracy of the global model. The expected improvement (EI) is used for the acquisition function.

Convergence. We show the convergence of the *LeFi* algorithm, in terms of reward weights and data sizes in Fig. 3 and Fig. 4, respectively. It can be seen that the reward weights (server-side) and data sizes (client-side) converge in about only 20 iterations. Notably, the reward weight of all CAVs increases in the very beginning (e.g., 5 iterations) and then diverges accordingly, which is because the total budget is reached, and the projection is invoked to ensure the budget for all CAVs. As the *LeFi* algorithm converges, we can see that the optimal reward weights are different for all CAVs, which is mainly attributed to their heterogeneous sensitivity functions. In addition, we show the convergence of overall accuracy under all the solutions in Fig. 5. As compared to *Baseline*, the *BARA* can generally achieve better accuracy performance, because it re-uses the previous observations for future searching. As compared to other solutions, the *LeFi*

¹We omit DRL-based solutions because they can hardly learn in very limited Server-to-CAV interactions (i.e., maximum 100).

algorithm shows a very smooth convergence and eventually higher accuracy performance throughout the iterations. Note that, although the absolute numeric difference is relatively small, such improvement in the accuracy of the global model is non-trivial. These results verify the superior convergence of the *LeFi* algorithm.

Scalability. We show the scalability of the *LeFi* algorithm in Fig. 6 by evaluating its accuracy performance under different numbers of CAVs. Here, the metric of relative accuracy improvement (RAI) is defined as the ratio between the accuracy improvement of a solution and the *LeFi* algorithm. The accuracy improvement denotes the achieved accuracy increase of the *LeFi* algorithm as compared to arbitrary reward weights. In general, a higher RAI indicates that the solution can achieve better accuracy. As we can see, the RAI of both *Baseline* and *BARA* are below 1.0, which means they are less comparative with the *LeFi* algorithm. In addition, their RAI difference becomes larger, as the increment of the number of CAVs, which suggests they are less scalable as compared to *LeFi* algorithm. When there are a massive number of CAVs in federated learning, the searching space for both *Baseline* and *BARA* increases exponentially, which decreases their effectiveness in finding the optimal reward weights of all CAVs.

Adaptability. We evaluate the adaptability of the *LeFi* algorithm under varying budget and latency requirements in Fig. 7 and Fig. 8, respectively. As more budget allows higher data sizes in all CAVs, the accuracy performance of all solutions is increased accordingly, as shown in Fig. 7. When the budget is no longer a significant constraint, all CAVs would like to contribute all of their training data, which is the main reason why all solutions share similar accuracy when the budget is 25 USD. However, we see that the *LeFi* algorithm can achieve higher improvement than other solutions, especially when the budget is smaller (e.g., 1). As a more strict latency requirement is enforced, we see the accuracy performance of all solutions generally reduce in Fig. 8, because CAVs cannot contribute more training data to match the allocated reward weights. In the meantime, the *LeFi* algorithm typically achieves higher accuracy gain than other solutions, especially when the latency requirement is loosened. These results justify the better adaptability of the *LeFi* algorithm under different conditions in federated learning.

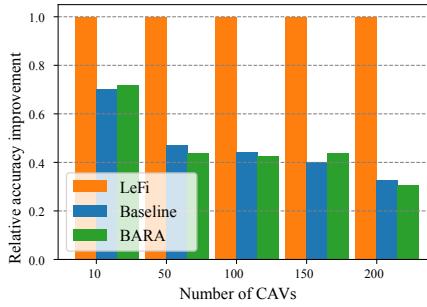


Fig. 6: Relative accuracy improvement under different number of CAVs.

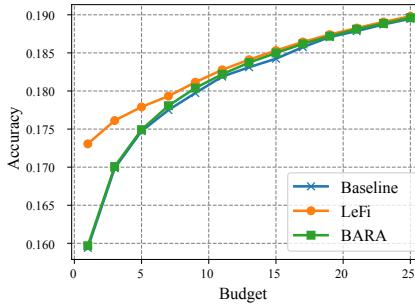


Fig. 7: Accuracy under different budget.

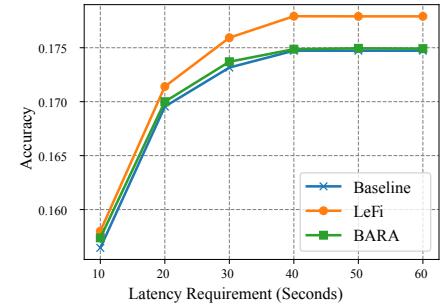


Fig. 8: Accuracy under different latency requirement.

V. RELATED WORK

The incentive mechanism in federated learning has been extensively investigated in recent years. Chai *et al.* [10] employed game theory and blockchain to encourage secure knowledge sharing by designing a transactional market-based approach. Saputra *et al.* [11] proposed employing Smart CAVs (SVs) as federal members to enhance network functionality with minimal data exposure. They devise a contract-based policy focusing on the optimal selection of SVs and payment contracts, dependent on the data quality. Yang *et al.* [7] introduced BARA, a Bayesian optimization-based algorithm for dynamic reward budget allocation in federated learning, aimed at optimizing model utility and accuracy. This novel approach effectively handles the complex relationship between reward allocation and model performance in federated learning environments. Weng *et al.* [12] presented FedServing, a federated prediction serving framework that incentivizes truthful machine learning predictions through a Bayesian game theory-based mechanism and enhances accuracy using truth discovery algorithms. This framework also ensures model privacy and exchange fairness by employing blockchain and Trusted Execution Environments (TEEs). Toyoda *et al.* [13] introduced a mathematical framework using contest theory to analyze and optimize reward distribution in a model updating environment. Emphasizing Bayesian-Nash incentive compatibility ensures that workers maximize their payoffs with no better alternative strategies, and it establishes an optimal reward distribution policy. However, existing works either assume the known sensitivity function of all CAVs or cannot scale to the massive number of CAVs (thousands if not more). In this work, we proposed the high-scalable *LeFi* algorithm to learn and incentivize heterogeneous CAVs without the need for prior knowledge of their sensitivity functions.

VI. CONCLUSION

In this work, we proposed a new *LeFi* algorithm to incentivize the participation of heterogeneous CAVs in federated learning. The key observation is that the sensitivity function of CAVs is generally unknown by the server side, due to a variety of practical factors, such as computation capacity and monetary sensitivity. Our key idea to solve this challenge is to develop a sample-efficient surrogate model to learn and approximate the unknown sensitivity function, from

accumulated Server-to-CAV interactions. Then, we design a sample-based gradient descent method to iteratively update the reward allocations and apply projection if needed to meet the budget. The simulation results show that our proposed algorithm can significantly outperform existing solutions, in terms of accuracy, scalability, and adaptability.

ACKNOWLEDGEMENT

This work is supported by the US National Science Foundation under Grant No. 2321699.

REFERENCES

- [1] B. McMahan *et al.*, “Communication-efficient learning of deep networks from decentralized data,” in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.
- [2] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated machine learning: Concept and applications,” *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–19, 2019.
- [3] 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.
- [4] Y. Zhan, J. Zhang, Z. Hong, L. Wu, P. Li, and S. Guo, “A survey of incentive mechanism design for federated learning,” *IEEE Transactions on Emerging Topics in Computing*, vol. 10, no. 2, pp. 1035–1044, 2021.
- [5] 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. 10700–10714, 2019.
- [6] Y. Zhan, P. Li, S. Guo, and Z. Qu, “Incentive mechanism design for federated learning: Challenges and opportunities,” *IEEE Network*, vol. 35, no. 4, pp. 310–317, 2021.
- [7] Y. Yang, Y. Zhou, M. Hu, D. Wu, and Q. Z. Sheng, “Bara: Efficient incentive mechanism with online reward budget allocation in cross-silo federated learning,” *arXiv preprint arXiv:2305.05221*, 2023.
- [8] Q. Li, Y. Diao, Q. Chen, and B. He, “Federated learning on non-iid data silos: An experimental study,” in *2022 IEEE 38th International Conference on Data Engineering (ICDE)*. IEEE, 2022, pp. 965–978.
- [9] Vast AI. (2023) Homepage. [Online]. Available: <https://vast.ai/>
- [10] H. Chai, S. Leng, Y. Chen, and K. Zhang, “A hierarchical blockchain-enabled federated learning algorithm for knowledge sharing in internet of vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 3975–3986, 2020.
- [11] Y. M. Saputra, H. T. Dinh, D. Nguyen, L.-N. Tran, S. Gong, and E. Dutkiewicz, “Dynamic federated learning-based economic framework for internet-of-vehicles,” *IEEE Transactions on Mobile Computing*, 2021.
- [12] J. Weng, J. Huang, C. Cai, and C. Wang, “Fedserving: A federated prediction serving framework based on incentive mechanism,” in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.
- [13] K. Toyoda, J. Zhao, A. N. S. Zhang, and P. T. Mathiopoulos, “Blockchain-enabled federated learning with mechanism design,” *IEEE Access*, vol. 8, pp. 219744–219756, 2020.