

An Incentive Design To Perform Federated Learning

Shashi Raj Pandey, Sabah Suhail, Yan Kyaw Tun, Madyan Alsenwi, Choong Seon Hong*

Department of Computer Engineering,

Kyung Hee University,

Yongin, 17104 Korea

Email: {shashiraj, sabah, ykyawtun7, malsenwi, cshong}@khu.ac.kr

Abstract

Federated Learning is a distributed model training approach having privacy preserving benefits. Under this machine learning approach, a number of devices will collaborate to train a learning model of interest without sharing their raw dataset. Instead, they share their local model parameters, and an aggregator such as a central multi-access edge computing (MEC) server will do the local parameters aggregation to build a global model. The global model is then broadcast back to the devices for the next round of local training, and the iterative process continues henceforth. However, in doing so, a key challenge is to ensure a number of active devices joining in the training process. In this work, we propose an incentive-based methodology to ensure active participation for federated learning, namely *Crowd-FL*. In particular, we characterize the impact of participation for a level of local accuracy and evaluate the learning performance against the random device selection approach as the baseline. To that end, we derive a utility maximization problem for incorporating participation and model training paradigm. Simulation results show that our proposed methodology will not only improve the average model accuracy, but also minimize the number of global communication rounds.

Keywords – federated learning, incentive mechanism, crowdsourcing, utility maximization

I. INTRODUCTION

Privacy concerns have been an important aspect to consider in designing training paradigms for various machine learning models. Recently, a distributed model training approach with privacy preserving benefits was proposed, namely Federated Learning [1]. In the general setting of federated learning, a number of devices collaborate to train a high-quality machine learning model of interest. In doing so, they communicate with a central server which acts as an aggregator to facilitate the global model building process. Within each round of communication with the MEC server, the devices do multiple rounds of local iterations over their raw dataset. Basically, the devices will adopt algorithms such as stochastic gradient descent (SGD) to train a local model. Then, the local model parameters are sent to the MEC server following various channel access techniques [13], [9], where they are aggregated to build a global model. Once done, the global model is broadcast back to the devices for running another phase of local iterations. Hence, the local model parameters are updated, again sent to the MEC server for the next round of aggregation. This iterative process continues henceforth.

In this regard, it has been evidently known that an increase in the number of devices will significantly improve the global model performance. Thus, first, a key challenge is to ensure a number of active devices joining in the training process. Second, numerous methodologies to efficiently handle the communication cost involved during model parameters exchanges has been studied [3], [2], [5]. However, delicate consideration in incorporating participation [8] and model training paradigm has been largely overlooked so far. Thus, to address this issue, we propose an incentive-based methodology to ensure

active participation for federated learning, namely *Crowd-FL*. In particular, a crowd of appropriate clients are influenced to participate in the model training process with the offered reward rate. The incentive for participating would cover the incurred local computation cost and communication cost for the devices in the federated learning framework [7], [10]. Furthermore, we derive a utility maximization problem for incorporating participation and model training paradigm. The utility maximization problem characterizes the performance of federated learning model for a given global accuracy level. Subsequently, we relate it with available devices to train the model, and the impact of incentive mechanism on them. To validate the efficacy of proposed approach, we rely on a naive random device selection method.

The rest of this paper is organized as follows. Section II discusses the system model and Section III explains the proposed Crowd-FL solution approach. Section IV shows simulation results. Finally, Section V concludes the paper with future work.

II. SYSTEM MODEL

We consider a set of devices $\mathcal{K} \in \{1, 2, \dots, K\}$ willing to solve a learning problem in the Crowd-FL. The devices are associated with a central server that performs local model parameters aggregation during the training of a high-quality global model in the federated learning setting. To that end, with the sample data $\{x_i, y_i\}$, a typical learning problem for an input sample vector x_i (e.g., the pixels of an image) is to find the *model parameter vector* $w \in \mathbb{R}^d$ that characterizes the output y_i (e.g., the labeled output of the image, such as the corresponding product names in a store) with a loss function $f_i(w)$. In this regard, some examples of loss functions

include $f_i(w) = \frac{1}{2}(x_i^T w - y_i)^2$, $y_i \in \mathbb{R}$ for a linear regression problem and $f_i(w) = \max\{0, 1 - y_i x_i^T w\}$, $y_i \in \{-1, 1\}$ for support vector machines. The term $x_i^T w$ is often called a *linear mapping function*.

For each client k , the empirical risk with respect to w on the local data set \mathcal{D}_k is

$$J_k(w) := \frac{1}{D_k} \sum_{i=1}^{D_k} f_i(w). \quad (1)$$

Then, we define the empirical risk with respect to w on all distributed data samples as the finite-sum objective of the form

$$\min_{w \in \mathbb{R}^d} J(w) \quad \text{where} \quad J(w) := \sum_{k=1}^K \frac{D_k}{D} \cdot J_k(w). \quad (2)$$

The objective for the set of devices in the federated learning framework is to find the optimal model parameter w^* that solves the learning problem (2) without sharing their available raw dataset. In the next section, we propose of method to enable a set of devices for solving this finite-sum objective problem.

III. CROWD-FL SOLUTION APPROACH

In our proposed Crowd-FL solution approach, the MEC offers reward to the participating devices in the model training process. Then, the participating devices will solve the learning problem (2) in its dual form, i.e., exploiting the primal-dual structure to distributively solve the local subproblems [5]. In such distributed optimization framework, the increase in the number of devices to train the learning model will improve the model accuracy within smaller communication rounds. Note that, in each communication round (global iteration), the MEC will aggregate the local model parameters in the form of dual variables, just like the vanilla federated learning algorithm FedAvg [1]. In this regard, for any strongly convex objective function, we have [6]:

$$I_{\text{global}} = \frac{\nu \cdot \log(\frac{1}{\epsilon})}{1 - \Theta}, \quad (3)$$

where $\nu > 0$ is a constant, ϵ is the global model accuracy, and Θ is the relative accuracy at each device's local subproblem characterized by the duality gap [7]. To that end, for a fixed global model accuracy ϵ , the key performance indicator, i.e., the number of global iterations I_{global} will depend on the local relative accuracy.

We characterize the improvement in the number of devices to train the global model with the offered reward rate in terms of the requirement for a value of Θ at each device. Here, the MEC evokes a number of devices to ensure a minimum level of Θ to improve the global model at an accuracy level. However, Θ and the possible influenced devices with the incentive plan are unknown. We then derive the optimal value of Θ for a number of training devices to analyze the improvement in the learning problem against random selection, and no guaranteed local accuracy. Consider a case where the devices solution to the local problem result is uniformly distributed $\Theta \in [\underline{\Theta}, \bar{\Theta}]$. We consider a maximum number of global iterations is $t \in [1, T_{\text{max}}]$. Then, with K devices available to join the training

process with the offered reward rate, we have only a subset of subset of devices joining in based on their responses. This means, devices performing worse than $\bar{\Theta}$ will be removed.

The number of devices joining the process for a given time t is $K(t) = K \cdot \left[\frac{\Theta(t) - \underline{\Theta}}{\bar{\Theta} - \underline{\Theta}} \right]$. Then, the MEC maximizes its utility with the known number of devices joining the training process as $\mathcal{U}(\cdot) + (1 - \Theta) \cdot K(t)$ such that $\mathcal{U}(\cdot) \equiv \gamma (1 - 10^{-(\eta(1-\Theta)+\zeta)})$, where $0 < \gamma \leq 1$ is a coefficient, and $\eta \geq 0, \zeta \leq 0$ are defined parameters that characterizes the model performance in terms of global accuracy, as in [4]. Therefore, for a number of known $K(t)$, the utility maximization problem in a Crowd-FL framework is

$$\begin{aligned} \max_{\Theta(t)} \quad & \gamma \left(1 - 10^{-(\eta(1-\Theta(t))+\zeta)} \right) + (1 - \Theta(t)) \cdot K(t) \\ \text{s.t.} \quad & \Theta \in [\underline{\Theta}, \bar{\Theta}]. \end{aligned} \quad (4)$$

The problem in (4) is a convex problem whose solution can be obtained using the Karush-Khun-Tucker (KKT) conditions [11]. Following the KKT condition (*details omitted for brevity*), the optimal value of Θ should satisfy

$$K = \frac{\ln(10) \cdot (\gamma\eta) \cdot 10^{-(\eta(1-\Theta^*(t))+\zeta)} \cdot (\bar{\Theta} - \underline{\Theta})}{1 - 2\Theta^*(t) + \underline{\Theta}}. \quad (5)$$

Next, we adopt an iterative method to solve the problem in (5) [12], starting with an initial average guess on Θ .

IV. SIMULATION RESULTS

We set up the network for $K = 20$ devices where each device has a portion of the total MNIST dataset by the order of $1/K$, i.e., we consider for an i.i.d setting. Based upon the offered reward, an appropriate number of devices will participate, instead of randomly selected subset of devices, to give a baseline comparison for the model training process. Furthermore, the general learning settings is as follows: each participating client will run local iterations for a fixed number of epochs in one round of training. The number of local epoch is 5, batch size is 20, and the learning rate is 0.01 at each device. We then compare the average accuracy of the trained model with and without the proposed incentive mechanism. As in Fig. 1, we observe the improvement in the model accuracy with the proposed Crowd-FL solution approach. In particular, the proposed method attains 80% accuracy in nearly 40 global iterations while the conventional method takes it over 50 for the same. Subsequently, the proposed incentive mechanism based participation methodology will reduce the number of global iterations to maintain a certain accuracy level. In doing so, the communication cost to train a high quality model is considerably lower, proving the efficacy of the proposed mechanism.

V. CONCLUSION

In this work, we have proposed an incentive-based methodology to ensure active participation for federated learning, namely *Crowd-FL*. In doing so, we have characterized the impact of participation for a level of local accuracy and evaluated the learning performance against the random device

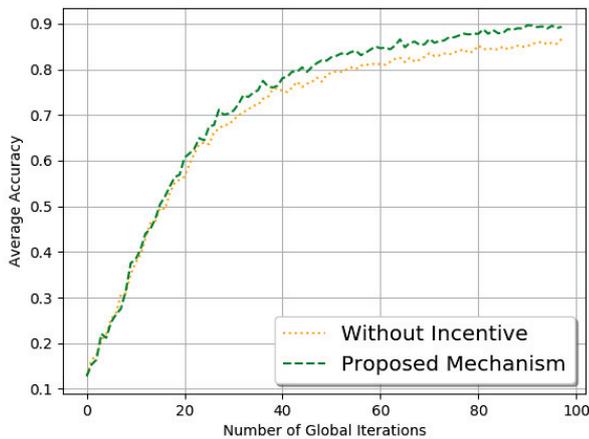


Fig. 1. Average accuracy versus the number of global iterations.

selection approach as the baseline. We have derived a utility maximization problem for incorporating participation and model training paradigm. In the simulation results, we have shown that our proposed methodology not only improves the average model accuracy, but also minimizes the number of global communication rounds.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2019-0-01287, Evolvable Deep Learning Model Generation Platform for Edge Computing). *Dr. CS Hong is the corresponding author.

REFERENCES

- [1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson et al., "Communication-efficient learning of deep networks from decentralized data," arXiv preprint arXiv:1602.05629, Feb 2016.
- [2] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, "When edge meets learning: Adaptive control for resource-constrained distributed machine learning," arXiv preprint arXiv:1804.05271, Aug 2018.
- [3] N. H. Tran, W. Bao, A. Zomaya, N. M. NH, and C. S. Hong, "Federated learning over wireless networks: Optimization model design and analysis," in IEEE Conference on Computer Communications (INFOCOM), Paris, France, April 29–May 2, 2019, pp. 1387–1395.
- [4] Pandey, S.R., Tran, N.H., Bennis, M., Tun, Y.K., Han, Z. and Hong, C.S., 2019, December. Incentivize to build: A crowdsourcing framework for federated learning. In 2019 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE.
- [5] Ma, Chenxin, Jakub Konečný, Martin Jaggi, Virginia Smith, Michael I. Jordan, Peter Richtárik, and Martin Takáč. "Distributed optimization with arbitrary local solvers." *Optimization Methods and Software* 32, no. 4 (2017): 813-848.
- [6] Konečný, Jakub, H. Brendan McMahan, Daniel Ramage, and Peter Richtárik. "Federated optimization: Distributed machine learning for on-device intelligence." arXiv preprint arXiv:1610.02527 (2016).
- [7] Pandey, S.R., Tran, N.H., Bennis, M., Tun, Y.K., Manzoor, A. and Hong, C.S., 2020. A crowdsourcing framework for on-device federated learning. *IEEE Transactions on Wireless Communications*.
- [8] Pandey, Shashi Raj, and Choong Seon Hong. "Response driven efficient task load assignment in mobile crowdsourcing." In 2018 International Conference on Information Networking (ICOIN), pp. 442-446. IEEE, 2018.
- [9] Alsenwi, Madyan, Ibrar Yaqoob, Shashi Raj Pandey, Yan Kyaw Tun, Anupam Kumar Bairagi, Lok-won Kim, and Choong Seon Hong. "Towards coexistence of cellular and wifi networks in unlicensed spectrum: A neural networks based approach." *IEEE Access* 7 (2019): 110023-110034.
- [10] Le, Tra, Tran, N.H., Tun, Y.K., N. M. NH, Han, Zhu, and Hong, C.S., 2020. "An Incentive Mechanism for Federated Learning in Wireless Cellular network: An Auction Approach." *IEEE Transactions on Wireless Communications* (submitted).
- [11] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [12] S. D. Conte and C. De Boor, *Elementary numerical analysis: an algorithmic approach*. SIAM, Dec 2017, vol. 78.
- [13] Alsenwi, M., Tran, N.H., Bennis, M., Pandey, S.R., Bairagi, A.K. and Hong, C.S., 2020. Intelligent Resource Slicing for eMBB and URLLC Coexistence in 5G and Beyond: A Deep Reinforcement Learning Based Approach. arXiv preprint arXiv:2003.07651.