2021 IEEE International Conference on Big Data and Smart Computing (BigComp) | 978-1-7281-8924-6/20/$31.00 ©2021 IEEE | DOI: 10.1109/BigComp51126.2021.00033

A Decentralized Game Theoretic Approach for Energy-aware Resource Management in Federated Learning

Chit Wutyee Zaw, Choong Seon Hong
Department of Computer Science and Engineering, Kyung Hee University,
Yongin-si, Republic of Korea
Email: {cwyzaw, cshong}@khu.ac.kr

***Abstract*—The resource management in Federated Learning (FL) system has been a challenging issue since mobile users can save the energy consumption by limiting their computing resources and dataset in training their local models in which users have the energy limitation. We analyze the performance of the global model on the size of dataset and computing resources used for the local training. The performance of the final model is significantly influenced by the resource management of users. Moreover, the decisions of the users on the communication, computing resources and size of dataset can affect the time taken for one computing round. Since a large number of mobile users participate in the FL, a centralized resource management is not practical. Thus, we formulate an energy-aware resource management problem for FL in which users are interested in minimizing the time taken for one computing round with the constraints of energy consumption, communication resources and performance of the training model. Due to the coupling in the communication resource allocation, we formulate the resource management problem as a Generalized Nash Equilibrium Problem (GNEP) and propose a decentralized algorithm. In addition, we analyze the performance of the proposed algorithm on the resource management, energy and time consumption.**

***Index Terms*—Decentralized approach, Energy-aware resource management, Federated learning, Generalized nash equilibrium problem**

I. Introduction

In the world of big data generated by a large number of Internet of Things, it is important to learn the behavior of mobile users or assist them with helpful recommendations. Due to the privacy concerns of mobile users, Federated Learning (FL) approach has emerged where users train a statistical model on the mobile devices with their local dataset. A global model can be learned by the contributions of the local models of participating mobile users. FL brings the multiple advantages to the learning at the edge network such as the privacy of the users and the communication efficiency without sending the local datasets to the centralized server. However, there are various challenges and open problems to address in FL [1].

Due to the energy limitation at the mobile users, the resource management in FL has become challenging. The performance of the local model is highly influenced by the local datasets and computing resources used for the training. Thus, there is a tradeoff between the model training loss and the energy consumption of the mobile users. To improve the performance of the local model, a large amount of datasets and computing resources are needed for the training. However, the high usage of datasets and computing resources will impose a high energy consumption on the mobile devices. In addition, a centralized solution approach for the resource management is not efficient due to a large number of participators in FL. The motivation and challenges for FL in wireless networks are studied in [2].

The resource management in FL has been explored in the existing works. Authors in [3] discussed the effect of wireless channel conditions on the learning model. They proposed a joint learning, channel allocation and user selection where the convergence of the FL algorithm is derived. A joint optimization of communication time and computing resource allocation of mobiles users for FL is proposed in [4] where a general upper bound on global iterations for FL is considered. Authors in [5] proposed a joint dataset and user selection for FL where the model loss and cost for the training are jointly minimized. The energy and time consumption of users is jointly minimized by the resource allocation for FL for wireless networks in [6] where a centralized solution for the formulated resource allocation problem is proposed. A centralized solution for optimizing the convergence time of the learning algorithm for FL is proposed in [7]. Authors in [8] analyzed the channel uncertainty in FL where the client scheduling and wireless resource allocation are jointly optimized. The communication latency for decentralized learning in wireless network is analyzed in [9]. In addition, the distributed game for the resource management using the Generalized Nash Equilibrium Game is proposed in [10] and [11]. This type of game is useful in addressing the resource coupling among players which is difficult to solve by optimization techniques.

In this paper, the tradeoff between the model training loss and the energy consumption is addressed with respect to the training data samples and computing resources used for FL. An energy-aware resource management problem where the uplink bandwidth allocation, the fraction of dataset and computing resources used in the training are managed to minimize the time consumption of FL while improving the training model within the limited energy of the users. Due

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.

2375-9356/21/$31.00 ©2021 IEEE
DOI 10.1109/BigComp51126.2021.00033

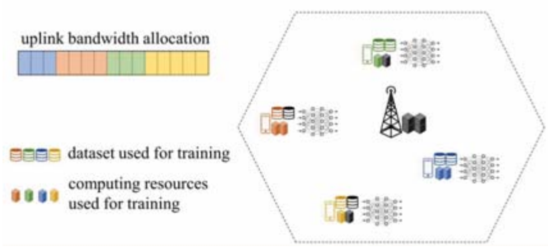

Fig. 1. Resource management in federated learning system.

to the coupling in the resource management decisions, we decouple the formulated problem and derive the closed form solution for the fraction of the local dataset. The decentralized generalized nash game is formulated for the uplink bandwidth and computing resource management. In addition, we propose an energy-aware decentralized game algorithm by penalizing the coupling constraints of the formulated game and taking the training model loss into account. The convergence of the algorithm is analyzed by performing the simulation.

II. System Model

We consider a FL model where a set of users $\mathcal{U} = \{1, 2, \cdots, U\}$ performs the local model training by using the local dataset, $\mathcal{K}_u$, and available computing resources, P_u. The model parameters are transferred to the edge sever, S, which is deployed at a base station where a global model is built by aggregating the local models. Since the mobile users are limited in terms of battery life, it is not efficient to use all the local dataset and computing resources. Moreover, the performance of the training model is influenced by the selected datasamples and computing resources. The resource management for FL is shown in Fig. 1.

A. Learning Model

User u participates in the FL by training the local model based on the local dataset, $\mathcal{K}_u$. Thus, user u minimize the loss function on the local dataset, $\mathcal{K}_u$, as follows.

$$\underset{\mathbf{w}_u}{\text{minimize}} \sum_{k\in\mathcal{K}_u} l(\mathbf{w}_u, x_k, y_k),$$

where $\mathbf{w}_u$ is the weight vector which describes the local model obtained by user u, x_k and y_k are the features and truth value of sample $k \in \mathcal{K}_u$. The Mean Squared Error is used as a loss function in performing the linear regression in our paper. Once the local training is done by all participating users, the weight vector, $\mathbf{w}_u$, is sent to the edge node where the model aggregation is done to build a global model as follows.

$$\bar{\mathbf{w}} = \frac{\sum_{u\in\mathcal{U}} |\mathcal{K}_u| \mathbf{w}_u}{\sum_{u\in\mathcal{U}} |\mathcal{K}_u|},$$

where the local weight, $\mathbf{w}_u$, is balanced by the contribution of user u in FL.

In our proposed model, we consider that users are allowed to train the local model by deciding a fraction of dataset denoted by κ_u depending on the energy level and time constraint. In order to save the energy consumption of the mobile devices, the performance of the training model has to be compromised. Let $\tilde{\mathcal{K}}_u$ be the dataset in which the data samples are selected randomly where $|\tilde{\mathcal{K}}_u| = \kappa_u |\mathcal{K}_u|$. Thus, the performance of the training model and the resulting local model, $\mathbf{w}_u$ of user u will be changed. The local training of user u is as follows.

$$\underset{\mathbf{w}_u}{\text{minimize}} \sum_{k\in\hat{\mathcal{K}}_u} l(\mathbf{w}_u, x_k, y_k). \tag{1}$$

The contribution of users might change according to their selected datasets in the model aggregation as follows.

$$\bar{\mathbf{w}} = \frac{\sum_{u\in\mathcal{U}} |\hat{\mathcal{K}}_u| \mathbf{w}_u}{\sum_{u\in\mathcal{U}} |\hat{\mathcal{K}}_u|}. \tag{2}$$

B. Computation of the Learning Model

Since a statistical model is built by updating the weights for multiple iterations, the training model is improved after multiple iterations but this will consume a significant amount of energy of mobile devices. Thus, users are allowed to stop the model training at any iteration which is converted to the fraction of computing resources, ρ_u, which influences the performance of the training model.

C. Communication Model

The uplink transmission is required to transfer the local model, $\mathbf{w}_u$, to the edge node for the model aggregation. The Orthogonal Frequency Multiple Access for the uplink transmission is considered where the uplink bandwidth is orthogonality divided among the users. The instantaneous rate of user u for the uplink data transmission is as follows.

$$R_u(\omega_u) = \omega_u \log_2\left(1 + \frac{p_u h_u}{n_0}\right),$$

where ω_u is the uplink bandwidth allocated to user u, p_u is the uplink transmit power, h_u is the channel gain and n_0 is additive white Gaussian noise.

D. Energy Consumption Model

The energy consumption of user, u, is defined as follows.

$$e_u(\kappa_u, \rho_u, \omega_u) = p_u\left(\frac{f(|\mathbf{w}_u|)}{R_u(\omega_u)}\right) + \psi\kappa_u f(|\mathcal{K}_u|)\tau^2\rho_u^2 P_u^2, \tag{3}$$

where ψ is the chip capacitance, τ is the CPU cycles required to process one byte data, and P_u is the total computing resources which is the CPU frequency of the mobile device of user u. $f(x)$ is a linear function which calculate the size of x in bytes. According to (3), $e_u(\kappa_u, \rho_u, \omega_u)$ is increasing with respect to κ_u and ρ_u, but it is decreasing with ω_u.

E. Time Consumption Model

The time required for one computing round in FL is calculated as follows.

$$t_u(\kappa_u, \rho_u, \omega_u) = \frac{\kappa_u f(|\mathcal{K}_u|)\tau}{\rho_u P_u} + \frac{f(|\mathbf{w}_u|)}{R_u(\omega_u)}. \tag{4}$$

According to (4), the time consumption of user u is decreasing with ρ_u and ω_u, but increasing with respect to κ_u.

III. Problem Formulation

Since the model training loss is compromised by saving the energy consumption, the edge server imposes a minimum threshold for the training loss which is the loss at the previous iteration. Moreover, time taken for one computing round taken is affected by the slowest user due to the synchronous model update. Thus, the objective of the training manager is to minimize the total time taken within the energy limit of mobile users while achieving the required training loss.

$$\underset{[\boldsymbol{\kappa},\boldsymbol{\rho}]\in\mathcal{B}^{|\mathcal{U}|},\boldsymbol{\omega}\in\mathbb{R}_+^{|\mathcal{U}|}}{\text{minimize}} \quad \max_u \{t_u(\kappa_u,\rho_u,\omega_u)\} \tag{5a}$$

$$\text{subject to} \quad \sum_{k\in\hat{\mathcal{K}}_u} l(\mathbf{w}_u,x_k,y_k) \le L(\bar{\mathbf{w}}^{t-1}), \forall u\in\mathcal{U}, \tag{5b}$$

$$e_u(\kappa_u,\rho_u,\omega_u) \le \bar{e}_u, \forall u\in\mathcal{U}, \tag{5c}$$

$$\sum_{u\in\mathcal{U}} \omega_u \le \bar{\omega}, \tag{5d}$$

where $\mathcal{B} := [0,1]$, $L(\bar{\mathbf{w}}^{t-1})$ is the global model loss at the previous iteration, $\bar{\omega}$ is the total available uplink bandwidth, and $\boldsymbol{\kappa} = [\kappa_u]^T_{u\in\mathcal{U}}$, $\boldsymbol{\rho} = [\rho_u]^T_{u\in\mathcal{U}}$, and $\boldsymbol{\omega} = [\omega_u]^T_{u\in\mathcal{U}}$.

Centralized solutions are not efficient for a large number of users. Thus, we present a decentralized resource management problem for each user $u, \forall u\in\mathcal{U}$ as follows:

$$\begin{aligned} \underset{[\kappa_u,\rho_u]\in\mathcal{B},\omega_u\in\mathbb{R}_+}{\text{minimize}} \quad & t_u(\kappa_u,\rho_u,\omega_u) \\ \text{subject to} \quad & \text{(5b), (5c), (5d)}, \end{aligned} \tag{6}$$

Since there is a coupling in the resource management of users, a decentralized game formulation is proposed.

IV. Decentralized Generalized Nash Game

In this section, we present the generalized Nash game formulation to address the resource coupling among the users. Due to the unpredictable behavior of the training loss function with respect to the resource management, we will omit the training loss constraint in this section. However, we will take the loss constraint into account in the proposed Algorithm.

The formulated resource management problem is decoupled to solve efficiently where the local dataset management problem is formulated as follows.

$$\begin{aligned} \underset{\kappa_u\in\mathcal{B}}{\text{minimize}} \quad & t_u(\kappa_u,\rho_u,\omega_u) \\ \text{subject to} \quad & \text{(5c)}, \end{aligned} \tag{7}$$

where the objective function and (5c) are increasing with κ_u. The closed form solution of κ_u is derived as follows.

$$\kappa_u = \min\left\{ \left(\bar{e}_u - p_u \frac{f(|\mathbf{w}_u|)}{\omega_u R_u(\omega_u)}\right) \frac{1}{\psi f(|\mathcal{K}_u|)\tau^2 \rho_u^2 P_u^2}, 1 \right\}. \tag{8}$$

The Generalized Nash Equilibrium (GNE) game $\mathrm{G}_u(\rho_u,\boldsymbol{\omega})$ of user u is formulated follows.

$$\mathrm{G}_u(\rho_u,\boldsymbol{\omega}): \underset{[\rho_u,\omega_u]\in\Gamma_u}{\text{minimize}} \quad t_u(\kappa_u,\rho_u,\omega_u). \tag{9}$$

where the strategy set of user u, Γ_u, is defined as $\Gamma_u = \Big\{ e_u(\kappa_u,\rho_u,\omega_u) \le \bar{e}_u, \omega_u \le \sum_{j\in\mathcal{U},j\neq u}\omega_j, \omega_u \ge 0, 0\le\rho_u\le 1 \Big\}$.

Lemma 1. *The objective function of user u, $t_u(\kappa_u,\rho_u,\omega_u)$, is convex with respect to ρ_u and ω_u. The strategy set of user u, Γ_u, is jointly convex with the other users' strategies.*

Based on Lemma 1, the existence of GNE is as follows.

Theorem 1. *There exist at least one Generalized Nash Equilibrium (GNE) for the formulated game $G_u(\rho_u,\boldsymbol{\omega})$.*

The formulated GNE game has the coupling in the resource management among users. To solve this problem efficiently, an energy-aware decentralized game algorithm is proposed.

V. Energy-aware Decentralized Game Algorithm

The formulated GNE game is hard to solve due to the coupling in the resource management strategies among the users. Thus, we propose a penalty approach for the formulated GNE game where the coupling constraints are penalized for the violation. The resource management problem of user u, $\bar{\mathrm{G}}_u(\rho_u,\boldsymbol{\omega}), \forall u\in\mathcal{U}$, with the penalty parameter for the coupling constraints is defined as follows.

$$\underset{[\rho_u,\omega_u]\in\bar{\Gamma}_u}{\text{minimize}} \quad t_u(\kappa_u,\rho_u,\omega_u) + \hat{\eta}\hat{e}_u(\kappa_u,\rho_u,\omega_u) + \tilde{\eta}\tilde{\omega}_u(\boldsymbol{\omega}), \tag{10}$$

where $\bar{\Gamma}_u = \{\omega_u \ge 0, 0\le\rho_u\le 1\}$, $\hat{e}_u(\kappa_u,\rho_u,\omega_u) = e_u(\kappa_u,\rho_u,\omega_u) - \bar{e}_u$ and $\tilde{\omega}_u(\boldsymbol{\omega}) = \omega_u - \sum_{j\in\mathcal{U},j\neq u}\omega_j$. $\hat{\eta}_u$ and $\tilde{\eta}_u$ are the penalty parameters which are updated as follows.

$$\hat{\eta}_u^t = \begin{cases} \hat{\eta}_u^{t-1}, & \text{if } \hat{e}_u(\kappa_u^t,\rho_u^t,\omega_u^t) \le 0 \\ \hat{\eta}_u^{t-1} + \hat{\Delta}, & \text{otherwise} \end{cases} \tag{11}$$

$$\tilde{\eta}_u^t = \begin{cases} \tilde{\eta}_u^{t-1}, & \text{if } \tilde{\omega}_u(\boldsymbol{\omega}^t) \le 0 \\ \tilde{\eta}_u^{t-1} + \tilde{\Delta}, & \text{otherwise} \end{cases}, \tag{12}$$

where $\hat{\Delta}$ and $\tilde{\Delta}$ are the increments to for the violated constraints to converge to a steady feasible solution.

Algorithm 1 Energy-aware Decentralized Game Algorithm

Output: $\bar{\mathbf{w}}$, $\boldsymbol{\kappa}$, $\boldsymbol{\rho}$, $\boldsymbol{\omega}$

```
t ← 0
each user u, ∀u ∈ U initializes κ_u^0, ρ_u^0, ω_u^0, η̂_u, η̃_u.
repeat
    t ← t + 1
    repeat
        update ρ_u^t, ω_u^t by solving (10)
        update κ_u^t according to (8)
        update η̂_u^t and η̃_u^t as in (11) and (12).
    until ê_u(κ_u^t, ρ_u^t, ω_u^t) ≤ 0, ω̃_u(ω^t) ≤ 0, ∀u ∈ U.
    update w_u^t by training the local model with K̂_u^t, ρ_u^t.
until Σ_{k∈K̂_u^t} l(w_u^t, x_k, y_k) ≤ L(w̄^{t−1}), ∀u ∈ U.
```

VI. Simulation Results

We consider a single cell macro base stations with an edge server where the linear regression model is trained on the Life Expectancy dataset from WHO by 100 mobile devices. Fig.

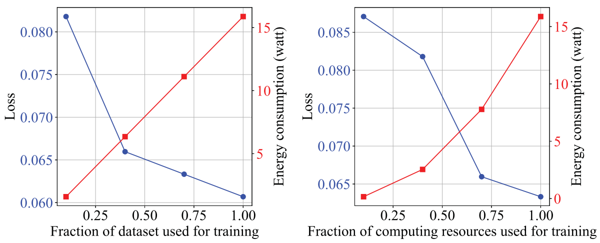

Fig. 2. Loss and energy consumption on dataset size and computing resources.

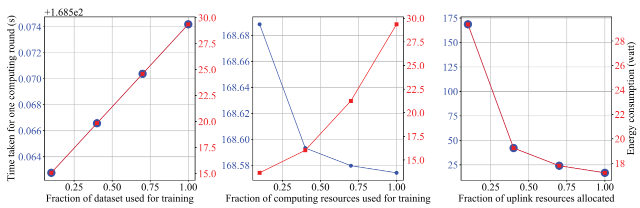

Fig. 3. Time and energy consumption on the allocation of resources.

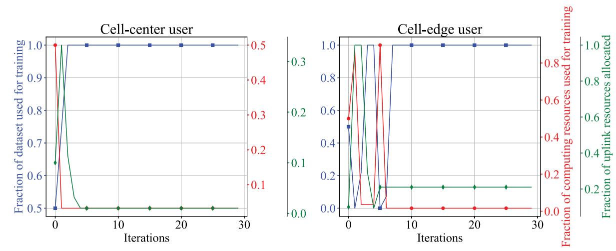

Fig. 4. Performance of the algorithm on the resources allocation.

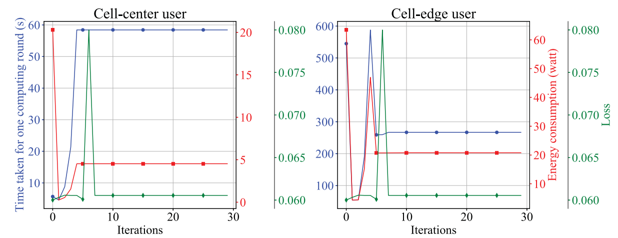

Fig. 5. Performance of the algorithm on loss, time and energy consumption.

2 shows the model training loss and energy consumption on the fraction of local dataset and computing resources used for training. The performance of the model is improved but the energy consumption is increased as the dataset size and computing resource increase. Fig. 3 shows that both the time and energy consumption of user increases with the fraction of dataset usage while the time and energy consumption with respect to the computing resources are contradict to each other. The time and energy consumption decreases as the allocated bandwidth increases which is limited and shared among the participating users. Fig. 4 shows that the algorithm converges to a steady point after a few iterations. Both the cell-center and cell-edge user use the whole dataset while the only a small fraction of computing resource is allocated for the model training. This is because the energy consumption is increasing linearly with the size of dataset while it is a quadratic function of the allocated computing resource. To minimize the time consumption, the higher fraction of uplink bandwidth resource is allocated to the cell-edge user than the cell-center user due to the channel conditions. Fig. 5 shows that there is a difference in the time consumption among two users although the higher bandwidth is allocated to the cell-edge user than the cell-center user. Moreover, the energy consumption of the cell-center user is less than the cell-edge user due to the transmit power used for the uplink transmission. The model training loss increases at iteration 7 because the resource management algorithm has not converged to a steady point.

VII. Conclusion

An energy-aware resource management problem for FL is proposed to save the energy consumption of the mobile devices with the cost of the training model performance. Due to the coupling in the limited uplink bandwidth resources among users, a decentralized generalized Nash Game is formulated. We then proposed an energy-aware decentralized game algorithm to solve the formulated problem efficiently. The simulation is performed to analyze the tradeoff between the model training loss and the energy consumption.

References

[1] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings *et al.*, "Advances and open problems in federated learning," *arXiv preprint arXiv:1912.04977*, Dec 2019.

[2] S. Niknam, H. S. Dhillon, and J. H. Reed, "Federated learning for wireless communications: Motivation, opportunities, and challenges," *IEEE Communications Magazine*, vol. 58, no. 6, pp. 46–51, Jul 2020.

[3] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," *arXiv preprint arXiv:1909.07972*, Sep 2019.

[4] 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 INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, Paris, Apr 2019, pp. 1387–1395.

[5] C. Feng, Y. Wang, Z. Zhao, T. Q. Quek, and M. Peng, "Joint optimization of data sampling and user selection for federated learning in the mobile edge computing systems," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, Dublin, Jun 2020, pp. 1–6.

[6] Z. Yang, M. Chen, W. Saad, C. S. Hong, and M. Shikh-Bahaei, "Energy efficient federated learning over wireless communication networks," *arXiv preprint arXiv:1911.02417*, Nov 2019.

[7] M. Chen, H. V. Poor, W. Saad, and S. Cui, "Convergence time optimization for federated learning over wireless networks," *arXiv preprint arXiv:2001.07845*, Jan 2020.

[8] M. M. Wadu, S. Samarakoon, and M. Bennis, "Federated learning under channel uncertainty: Joint client scheduling and resource allocation," *arXiv preprint arXiv:2002.00802*, Feb 2020.

[9] N. Naderializadeh, "On the communication latency of wireless decentralized learning," *arXiv preprint arXiv:2002.04069*, Feb 2020.

[10] C. W. Zaw, N. N. Ei, H. Y. R. Im, Y. K. Tun, and C. S. Hong, "Cost and latency tradeoff in mobile edge computing: A distributed game approach," in *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, Kyoto, Feb 2019, pp. 1–7.

[11] C. W. Zaw, N. H. Tran, W. Saad, Z. Han, and C. S. Hong, "Generalized nash equilibrium game for radio and computing resource allocation in co-located mec," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, Dublin, Jun 2020, pp. 1–6.