Joint Accuracy and Resource Allocation for Green Federated Learning Networks

Xu Chu[1], Xiaoyang Liu[1], Qimei Chen[1](✉), Yunfei Xiong[2], Juanjuan Wang[3], Han Yu[2], and Xiang Hu[2]

[1] School of Electronic Information, Wuhan University, Wuhan 430072, China
{chu_xu,liuxiaoyang,chenqimei}@whu.edu.cn
[2] Wuhan Fiberhome Technical Services Co. Ltd., Wuhan 433072, China
{yfxiong,yuhan,huxiang}@fiberhome.com
[3] School of Business Administration, Zhongnan University of Economics and Law, Wuhan 430072, China
Wangjj@zuel.edu.cn

Abstract. This paper studies the energy and time resource optimization of federated learning (FL) in wireless communication networks. In the considered network model, each client uses local data for model training, and then sends the trained FL model to the central server. However, the energy budget of the local computing and transmission process is limited. Therefore, reducing energy consumption should be given priority when we consider the FL efficiency and accuracy. We invest a green communication joint learning issue and expressed as an optimization problem. To minimize the energy consumption under the condition that the overall FL time is constrained, we propose an iterative algorithm based on Lyapunov optimization. Our algorithm selects the clients participating in each round and allocates the different bandwidth to each client. At the same time, the connection between local training and communication process is considered so that we can get the optimal client local calculation force.

Keywords: Client selection · Resource allocation · Green communication

1 Introduction

In the era of big data, AI, and deep learning [1–3] continue to develop. Related technologies use massive amounts of data for training, so as to achieve the purpose of intelligence. The increasing requirements for storage space, computer capabilities, and privacy requirements have become urgent issues. The goal of federated learning is to enable mobile devices to collaboratively learn a shared machine learning model. Instead of sharing the entire training data to the central server, Federated learning only uploads the partial learning model to it [4]. Due to limited wireless resources, this process may affect the performance of FL.

M. Qiu et al. (Eds.): SmartCom 2021, LNCS 13202, pp. 154–163, 2022.
https://doi.org/10.1007/978-3-030-97774-0_14

In recent years, many studies have done a lot of work to optimize FL. [5–8] introduced several new designs of the FL algorithm and various problems as well as solutions to improve the FL effect. However, they mainly focused on the delay or accuracy of FL. As the focus on energy issues gradually increasing, the topic of Green Communication on how to reduce FL energy consumption continued to attract people's attention. In [8], the author proposed a new FL algorithm that can minimize the communication cost. [10–12] studied the scheme of reducing FL energy consumption or time consuming. However, the optimization problem of [9–12] is formed by treating each learning round equally, so equal network resources are allocated between learning rounds. In [13], the author found a phenomenon called "later is better" in the representative machine learning tasks such as image classification and text generation. They explicitly considered the different significance of each FL round for the final result, and studied the bandwidth allocation as well as customer selection under the conditions of long-term energy constraints and uncertain wireless channel information. However, paper [13] did not consider the efficiency and accuracy of FL. It also ignored the energy distribution between local and communication.

Our paper aims to describe the problem of client selection and bandwidth allocation in FL, while considering the optimization of the client's local energy consumption and transmission energy consumption in the presence of total time constraints. We have made better use of the phenomenon in [13] to reduce FL energy consumption and formulated a problem of minimizing FL energy under a time consuming limit. This work aims to minimizes the total energy consumption of local computing and wireless transmission without decreasing the FL efficiency and accuracy. We propose an algorithm based on Lyapunov optimization to solve the above optimization problem. The simulation results show that compared with the traditional FL method, the proposed scheme can achieve about 56.5% energy saving.

2 System Model

We now build a wireless joint learning network, which consists of K clients and a central server. Each client k has a local data set D_k. For each data set $D_k = \{x_{ki}, y_{ki}\}_{i=1}^{D_k}$, $x_{ki} \in R^d$ is the input vector of client k, y_{ki} is its output. For conveniently describing the relationship, the FL model trained by the data set of each client is called the local FL model. The FL model generated by the central server using the input of the local FL model of all clients is called the global FL model.

The FL process between the client and the central server includes three steps in each iteration: First, each client participates in the FL round is updated locally. Afterwards, the client transmits the local FL model to the central server. Finally, the results are aggregated and broadcast globally on the central server. Each iteration is called a learning round. The wireless joint learning network needs to select the appropriate client to upload the updated local model in each round. It depends on the channel state of each client in the current round, the previous accumulated time consuming and bandwidth allocation. The method will optimize the learning performance at the same time minimize the energy consumption. Therefore, we use $a_k^t = \{0, 1\}$ to indicate whether client k is selected in round t and $a_k^t = 1$ indicates that the client k is selected.

2.1 Resource Consumption Model

Local Training Process. The local training time can be defined by the following formula:

$$\tau_k^{t,l} = \frac{I_k C_k D_k}{f_k^t}, \tag{1}$$

where f_k^t represents the computing power of client k, which is determined by the number of CPU cycles per second. I_k is the number of local iterations, and C_k is the number of CPU cycles required to process a piece of data. According to [10], the total local energy consumption of client k is

$$E_k^{t,l} = I_k \kappa C_k D_k (f_k^t)^2, \tag{2}$$

where κ is the effective switched capacitance that depends on the chip architecture.

Transmission Process. Use b_k^t as the bandwidth allocated by client k in round t. Since the bandwidth of the system has an upper limit, we have $\sum_{k=1}^{K} b_k^t \leq B^t$, where B^t is the upper limit of the total bandwidth in round t. According to Shannon's formula, for the achievable transmission rate of client k participating in round t, we have

$$r_k^{t,c} = b_k^t \log_2(1 + \frac{g_k^t p_k^t}{N_0 b_k^t}), \tag{3}$$

where p_k^t is the average transmit power of client k, g_k^t is the channel gain between client k and the central server, and N_0 is the power spectral density of Gaussian noise. According to (3) we can define the time consuming of upload process as $\tau_k^{t,c}$. It should be noted that the transmission efficiency defined by Shannon's formula is the upper limit of the transmission rate. In order to transmit the data completely, the inequation $r_k^t \tau_k^{t,c} \geq L$ should be satisfied, where L represents the data size of the machine learning model used. In order to reduce the time consuming, we order $\tau_k^{t,c} = \frac{L}{r_k^t}$. Therefore, we derive the average transmission power according to Eq. (3) and $\tau_k^{t,c}$ as

$$p_k^t = \frac{b_k^t N_0}{g_k^t} (2^{\frac{L}{b_k^t \tau_k^{t,c}}} - 1). \tag{4}$$

Then, the transmission energy consumption of client k in round t can be obtained as

$$E_k^{t,c} = \frac{b_k^t N_0 \tau_k^{t,c}}{g_k^t} (2^{\frac{L}{b_k^t \tau_k^{t,c}}} - 1). \tag{5}$$

Therefore, in round t, the total time consuming of client k is

$$\tau_k^t = \tau_k^{t,l} + \tau_k^{t,c}. \tag{6}$$

The total energy consumption of client k in round t is

$$E_k^t = E^{t,l} + E^{t,c}. \tag{7}$$

2.2 Packet Error Rates

In the process of transmission, there may be errors in the data. These errors will affect the accuracy of FL. To simplify the problem, we assume that each local FL model will be transmitted as a single packet in the uplink. The packet error rate when the client k transmits the local model to the central server is given by [14]:

$$w_k^t = 1 - e^{-\frac{m b_k^t N_0}{p_k^t g_k^t}}, \tag{8}$$

where m being a waterfall threshold [15].

The packet error rate helps us filter the client. If the central server determines that there is an error in the received local FL model, the local model will be refused to participate in this round of global aggregation process, and the client will be rejected to participate in the next FL process. We define $c_k^t = \{0, 1\}$, and $c_k^t = 0$ means that client k does not participate in the subsequent FL process.

2.3 Problem Formulation

For general high-accuracy, high-robust machine learning problems, fewer clients are selected in the early learning rounds, while more clients are selected in the later learning rounds. This distribution method will behave better than the average distribution. We call this phenomenon "later is better". Therefore, time weight λ^t is introduced to reflect the impact of choosing different numbers of clients in different learning stages on FL energy and performance.

To decrease the energy consumption of FL while avoiding the side effect, we propose an optimization problem. Its goal is to minimize the energy consumption of the client in the presence of a time consuming limit:

$$\text{P1:} \quad \min_{\boldsymbol{a}^0, \boldsymbol{b}^0, \boldsymbol{f}^0, \ldots, \boldsymbol{a}^{T-1}, \boldsymbol{b}^{T-1}, \boldsymbol{f}^{T-1}} \sum_{t=0}^{T-1} E^t(\boldsymbol{a}^t, \boldsymbol{b}^t, \boldsymbol{f}^t), \tag{9}$$

$$\text{s.t.} \sum_{t=0}^{T-1} \left(\frac{I_k C_k D_k}{f_k} + \frac{L}{b_k^t \log_2 \left(1 + \frac{p_k^t g_k^t}{N_0 b_k^t}\right)} \right) \leq T_k, \forall k, \tag{9a}$$

$$b_{\min} \leq b_k^t \leq b_{\max}, \forall k, \forall t, \sum_{k=1}^{K} b_k^t \leq B^t, \forall t, \tag{9b}$$

$$f_{\min} \leq f_k^t \leq f_{\max}, \forall k, \forall t, \tag{9c}$$

$$a_k^t = \{0, 1\}, c_k^t = \{0, 1\}, \forall k, \forall t, \tag{9d}$$

where $E^t\left(\boldsymbol{a}^t, \boldsymbol{b}^t, \boldsymbol{f}^t\right) = \lambda^t \sum_{k=1}^{K} E_k^t a_k^t c_k^{t-1}$. It should be noted that c_k^{t-1} represents the impact of the previous round of packet error rate on the current round of client selection decisions. Constraint (9a) requires that the total time consuming for the client k in T rounds to complete the FL task does not exceed the time consuming limit T_k. Constraint (9b) is the range of bandwidth allocation. Constraint (9c) limits the user's CPU computing power. Constraint (9d) jointly determine the decision of client selection.

3 Resource Allocation For Energy Minimization

In this section, we propose a low-complexity algorithm to solve P1 on a basis of the Lyapunov optimization. We assume that in terms of measuring energy and time consuming, the client's local update and upload processes are independent of each other. Therefore, we consider the optimization problems of these two processes separately.

3.1 Local Iterative Optimization Algorithm

For the selected client, the optimization problem of the local iteration is

$$\text{P2:} \quad \min_{\boldsymbol{f}^0,\ldots,\boldsymbol{f}^{T-1}} \sum_{t=0}^{T-1} I_k \kappa C_k D_k \left(f_k^t\right)^2, \tag{10}$$

According to (10), using minimum f_k^t is always effective. If each task has a time-consuming threshold $\tilde{\tau}$, then we have $\frac{I_k C_k D_k}{f_k} \leq \tilde{\tau}$. So in order to minimize (10), the optimal solution of f_k^t is $f^* = \frac{I_k C_k D_k}{\tilde{\tau}}$. Therefore, the local time consuming satisfies $T_k^l = \sum_{t=0}^{T-1} \tilde{\tau}$. According to the total time consuming limit, it can be deduced that the time consuming limit of the upload process is $T_k^c = T_k - T_k^l$.

3.2 Optimization Algorithm for Uploading Process

Under the time consuming limit, the client's reasonable selection and assignment of tasks will reduce the overall time-consuming of FL. The time consumed in each round will cause the client to gradually approach the time-consuming limit T_k^c, which will be a dynamic problem. To solve this problem, we construct a time deviation queue $q_k(t)$ based on Lyapunov optimization, which is used to represent the deviation between the total time consuming of the current transmission process of client k and the time limit T_k^c. Client selection and bandwidth allocation decisions should be guided under the conditions of satisfying time constraints. We define the update of the queue as

$$q_k(t+1) = \left[\tau_k^{t,c} - \frac{T_k^c}{T} + q_k(t)\right]^+, \tag{11}$$

where $[\cdot]^+ = \max\{\cdot, 0\}$.

With $q_k(t)$, the constraint (9a) in P1 can be connected with the minimization problem (9). In this way the infinite time domain minimum average cost problem that is difficult to solve is transformed into the minimum value problem in each round. Now we are about to solve the following problem:

$$\text{P3:} \quad \min_{\boldsymbol{a}^0, \boldsymbol{b}^0, \ldots, \boldsymbol{a}^{T-1}, \boldsymbol{b}^{T-1}} \left(V\lambda^t E^{t,c} + \sum_{k=1}^{K} q_k(t)\, \tau_k^{t,c} \right) a^t c^t, \tag{12}$$

$$\text{s.t. } (11), (13), (14). \tag{12a}$$

Among them, V is a set of penalty factors. The purpose is to make a trade-off between minimizing energy loss and satisfying time consuming limit. It is used to adjust the weight relationship between Lyapunov drift and penalty in each round t. The additional term $\sum_{k=1}^{K} q_k(t) \cdot \tau_k^{t,c}$ in Eq. (13) helps us to transform constraint (9a) into each round of problems, which appears in the form of penalty terms. P3 takes into account the deviation between the previous accumulated time of the client in the current round and the upper limit of the evenly allocated time, which can be called a time deficit. If $q_k(t)$ is larger, the client will be more inclined to invest in the round that suits it (always in the later rounds) to ensure that the total time consuming of communication process does not exceed the time consuming limit T_k^c. In this way, we have avoided meeting the time constraints while predicting future channel conditions, at the same time reduced energy consumption as much as possible.

Note that P3 still contains a_k^t and c_k^t. It is still a mixed integer problem. In order to solve P3, we need to strip a_k^t and c_k^t from the equation solving. Therefore, metric $\rho_k^t = \frac{V\lambda^t \tilde{p} + q_k(t)}{g_k^t}$ is introduced, where $\tilde{p}$ is the upper limit of the client's target transmit power. Based on the metric ρ_k^t, the client is gradually added to the selection set S in ascending order of ρ_k^t. The smaller ρ_k^t is, the higher the priority. This selection set will assist us in bandwidth allocation and client selection. For every selection set, bandwidth allocation is transformed into the following optimization problem:

$$\text{P4:} \quad \min_{\{b_k^t\}_{k\in S}} \sum_{k\in S} \left[V\lambda^t \tilde{p} + q_k(t) \right] \cdot \frac{L}{b_k^t \log_2\left(1 + \frac{\tilde{p}\cdot g_k^t}{N_0 b_k^t}\right)}, \tag{13}$$

$$\text{s.t. } b_{\min} \le b_k^t \le b_{\max}, \forall k \in S, \tag{13a}$$

$$\sum_{k\in S} b_k^t \le B^t, \forall t. \tag{13b}$$

For each selection set S, the corresponding client optimal bandwidth allocation result is $b^*(S)$. For the initial data set (i.e., a data set with only one client), we make the optimal bandwidth allocation of the first client to be $b^*(S_0) = b_{\max}$. As the client is continuously added to S, its termination condition is $B^t - \sum_{m=0}^{M-1} b^*(S_m) < b^*(S_M)$, which can control the choice of the client.

Algorithm 1. GFL (Green Federated Learning) Algorithm

Require: $q_k(0) = 0, \forall k$

Rank the clients according to ρ. Therefore we have $\rho_1 \le \rho_2 \le \dots \rho_K$

Let $S_0 = \{k : \rho_k = \rho_1\}, S = S_0, b^*(S_0) = b_{\max}$

for $k = \mid S_0 \mid +1, \dots, K-1$ **do**

  Set $S = S + \{k\}$

  Solve **P5** and gain $\boldsymbol{\eta}^*(S)$ and $\boldsymbol{b}^*(S)$

  if $w > \tilde{w}$ **then**

    Refuse client k

  end if

  if $B - \sum_{m=1}^{k} b_m^* < b_k^*$ **then**

    $S = S - \{k\}$

    Let $S = S^*$

    Stop the iteration

  end if

end for

Obtain $\boldsymbol{a}^*$ where $a_k^* = \mathbf{1}\{k \in S^*\}, \forall k$ and $\boldsymbol{b}^* = \boldsymbol{b}^*(S^*)$

To find the optimal solution of P4 more intuitively, the variable $\eta_k^t = \frac{b_k^t}{g_k^t}$ is introduced. It represents the bandwidth under unit channel gain. Then P4 is transformed into P5. The pseudo code of this process is given in Algorithm 1.

$$\text{P5:} \quad \min_{\{b_k^t\}_{k\in S}} \sum_{k\in S} \frac{V\lambda^t \tilde{p} + q_k(t)}{g_k^t} \cdot \frac{L}{\eta_k^t \log_2\left(1 + \frac{\tilde{p}}{N_0 \eta_k^t}\right)}. \tag{14}$$

$$\text{s.t. } b_{\min} \le b_k^t \le b_{\max}, \forall k \in S, \tag{14a}$$

$$\sum_{k\in S} b_k^t \le B^t, \forall t. \tag{14b}$$

With regards to the complexity of GFL algorithm, we divide it into two parts. The first part is to sort the client with bubble sorting. The complexity in the worst case is $\mathcal{O}(k^2)$ and the best case is $\mathcal{O}(k)$. In the second part, we solve the minimum problem in cyclic iteration. We only consider the worst case, and the time complexity is $\mathcal{O}(k)$.

3.3 Overall Energy Optimization

In this subsection, we summarize the energy in the local and transmission process to describe and analyze the rationality of Algorithm 1. How the selection set assists us in selecting clients and allocating the bandwidth is also outlined. We make the following analysis to prove the optimality of Algorithm 1.

Proposition 1: In any FL round t, the allocated bandwidth of the client added to the S set will not be greater than the allocated bandwidth of the current existing client in the S set.

Proposition 1 shows that, according to the definition of the metric ρ_k^t, with smaller ρ_k^t, the client has better channel condition and less time consuming. Therefore, according to Proposition 1, we will allocate more bandwidth to clients with smaller ρ_k^t. According to constraint (13a), there are upper and lower limits for the bandwidth allocation of client k. The lower limit $b_{\min}$ is the minimum bandwidth that maintains the client's normal transmission of the local model, and is limited by the packet error rate. The relationship can be expressed as $b_{\min} = \frac{\tilde{p} \cdot g_k^t}{m N_0} \ln \left(\frac{1}{1-\tilde{w}} \right)$, where $\tilde{w}$ is the error rate threshold. As bandwidth is limited and the termination condition $B^t - \sum_{m=0}^{M-1} b^* \left(S_m \right) < b^* \left(S_M \right)$ is met, in round t, we select the first M clients by bandwidth.

Corollary 1: Assume that the size of the local data of all clients is the same, i.e., $D_{k_1} = D_{k_2}, \forall k_1 \neq k_2$. The number of local iterations of all clients and the number of CPU cycles required to process one data are also the same. In this way f^* only depends on the time-consuming threshold $\tilde{\tau}$. Therefore, it can be proved that the local energy minimization and the transmission energy minimization cooperate to solve the overall energy minimization problem.

4 Simulation Results And Analyze

We use the CIFAR10 data set to simulate the FL process. The total number of clients is set to 20, each client is allocated 500 training samples. The number of local iterations $I_k = 50$ and the number of CPU cycles required to process one data is set as $C_k = 2 \times 10^4$ cycles/sample. The total bandwidth of the wireless network during transmission is $B = 20$ MHz, and the power spectral density of Gaussian white noise is set as $N_0 = -174$ dBm/Hz. To limit the energy consumption of the client, we set the upload power as $\tilde{p} = 30$ dBm. The maximum allocated bandwidth $b_{\max} = 2$ MHz. For each client, FL runs a total of 150 rounds and the time consuming limit $T_k = 100$ s. We named our algorithm GFL (Green Federated Learning) and compared the performance with the Select-All FL model. We make all 20 clients selected by Select-All in each round. GFL dynamically selects clients based on "later is better" throughout the FL process.

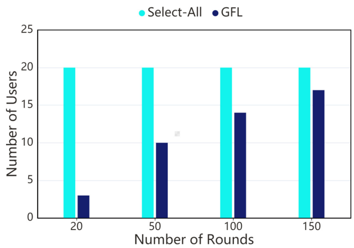

Fig. 1. Clients selection in each round

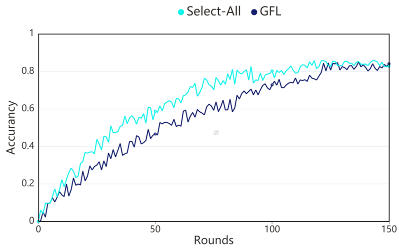

Fig. 2. Comparison of accuracy

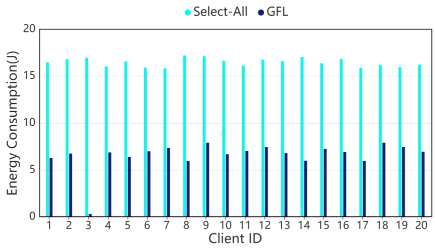

Fig. 3. Comparison of energy consumption

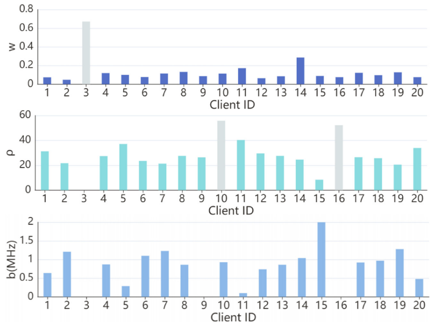

Fig. 4. Client selection and bandwidth allocation strategy

Figure 1 shows the number of clients selected by GFL and Select-All in each round. Select-All uses all 20 clients in each round, while GFL selects fewer clients in the early rounds and more clients in the later rounds. The distribution of client choices with rounds can be flexibly adjusted by time weights λ^t.

Figure 2 shows the accuracy of GFL and Select-All. Considering that Select-All has used all clients, it has good convergence efficiency and accuracy. Although the accuracy of GFL falls behind Select-All in the previous rounds, at the end of the FL round, its accuracy is close to Select-All.

Compared with GFL, Select-All generates a considerable amount of energy consumption. Figure 3 compares the energy consumption of the two methods for a single client. Since Select-All used all clients, its energy consumption was significantly higher than GFL at all stages. GFL saves a lot of energy consumption while ensuring the overall accuracy. It can be seen that the client 3 consumes very little energy. But it is not because it takes an advantage over other clients. We will introduce the specific reasons below.

Figure 4 shows our client selection and bandwidth allocation strategy. The top subplot shows that if the packet error rate w^t is higher than the threshold, the client will be refused to participate in the subsequent FL process. The packet error rate of client 3 exceeds the threshold, so it will not be selected. Therefore, although it consumes very little energy, it makes no contribution to the result of FL. The subplot in the middle shows the selection metric ρ_k^t. Gray clients indicate that they are not selected. The bottom subplot shows the bandwidth allocation of the selected client. According to our algorithm, clients with higher priority have lower channel conditions and time consuming, and they will be allocated more bandwidth.

5 Conclusion

Federated learning has attracted much attention as a popular artificial intelligence technology that can protect data privacy and solve the problem of data islands. The allocation of wireless network resources is one of the important and commonplace topics. This paper studied the client selection and bandwidth allocation issues in wireless communication networks. Based on the time dependence of different learning rounds, we formulated a joint accuracy and resource allocation problem to minimize the energy consumed in the training and transmission process with a total time consuming limit. We used Lyapunov optimization to find the solution of the problem. The results showed that our scheme is superior to the classic algorithm in terms of total energy consumption. It can achieve about 56.5% energy saving.

References

1. Qiu, H., Qiu, M., Lu, Z.: Selective encryption on ECG data in body sensor network based on supervised machine learning. Inf. Fusion **55**, 59–67 (2020)
2. Qiu, L., Gai, K., Qiu, M.: Optimal big data sharing approach for tele-health in cloud computing. In: IEEE SmartCloud, pp. 184–189 (2016)
3. Lu, R., Jin, X., Zhang, S., Qiu, M., Wu, X.: A study on big knowledge and its engineering issues. IEEE Trans. Knowl. Data Eng. **31**(9), 1630–1644 (2018)
4. Zhu, G., Liu, D., Du, Y., You, C., Zhang, J., Huang, K.: Towards an intelligent edge: wireless communication meets machine learning (2018). arXiv:1809.00343
5. Bonawitz, K., et al.: Towards federated learning at scale: system design. In: Proceedings of Systems and Machine Learning Conference, Stanford, CA, USA, pp. 1–15 (2019)
6. McMahan, H.B., Moore, E., et al.: Communication-efficient learning of deep networks from decentralized data. arXiv preprint arXiv:1602.05629 (2016)
7. Zhao, X., Yang, K., Chen, Q., et al.: Deep learning based mobile data offloading in mobile edge computing systems. Future Gener. Comput. Syst. **99**, 346–355 (2019)
8. Yang, Y., Chen, Q.: Distributed resource allocation under mobile edge computing networks: invited paper. In: 17th ISWCS, pp. 1–6 (2021)
9. Konečný, J., McMahan, B., Ramage, D.: Federated optimization: distributed optimization beyond the datacenter, November 2015. arXiv:1511.03575
10. Yang, Z., Chen, M., Saad, W., Hong, C.S., Shikh-Bahaei, M.: Energy efficient federated learning over wireless communication networks. IEEE Trans. Wirel. Comm. **20**(3), 1935–1949 (2021)
11. Hu, Y., Huang, H., Yu, N.: Device scheduling for energy-efficient federated learning over wireless network based on TDMA mode. In: International Conference on WCSP, pp. 286–291 (2020)
12. Tran, N.H., Bao, W., Zomaya, A., Nguyen, M., Hong, C.S.: Federated learning over wireless networks: optimization model design and analysis. In: IEEE INFOCOM 2019, pp. 1387–1395 (2019)
13. Xu, J., Wang, H.: Client selection and bandwidth allocation in wireless federated learning networks: a long-term perspective. IEEE Trans. Wirel. Comm. **20**(2), 1188–1200 (2021)
14. Zeng, Q., Du, Y., Huang, K., Leung, K.K.: Energy-efficient radio resource allocation for federated edge learning. In: Proceedings of IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–6, June 2020
15. Xi, Y., Burr, A., Wei, J., Grace, D.: A general upper bound to evaluate packet error rate over quasi-static fading channels. IEEE Trans. Wirel. Commun. **10**(5), 1373–1377 (2011)