Journal of Electrical Engineering and Information Technologies, Vol. **10**, No. 1 pp. 41-46 (2025)
Article 231
In print: ISSN 2545–4250
Received: 16.12.2024
On line: ISSN 2545–4269
Accepted: 14.04.2025
UDC: 621.311.1:621.371/.372]:004.85.032.24

Original scientific paper

INTEROPERABILITY BETWEEN FEDERATED LEARNING AND WIRELESS POWERED COMMUNICATIONS: MINIMIZING TRAINING DELAYS

Slavche Pejoski [1)] Marija Poposka Zoran Hadzi-Velkov

Ss Cyril and Methodius University in Skopje, Faculty of Electrical Engineering and Information Technologies, Republic of North Macedonia

[1)] slavchep@feit.ukim.edu.mk

Abstract: This paper explores an interoperability scenario between wireless powered transfer and federated learning (FL) technologies. In this scenario, the base station does not only coordinate training of the global FL model, but it also charges energy harvesting (EH) clients, which are responsible for training the local models. These EH clients are equipped with rechargeable batteries, enabling them to perform local processing and energy harvesting concurrently. We propose an efficient resource allocation scheme that optimizes both computing and communication parameters to minimize training latency. Simulation results show a significant latency reduction compared to state-of-the-art FL system that operates with non-overlapping local processing and energy harvesting phases.

Keywords: wireless powered communication networks, resource allocation, federated learning

ИНТЕРОПЕРАБИЛНОСТ ПОМЕЃУ ФЕДЕРАТИВНО УЧЕЊЕ И БЕЗЖИЧНО НАПОЈУВАНИ КОМУНИКАЦИИ: МИНИМИЗИРАЊЕ НА ВРЕМЕТО НА ТРЕНИРАЊЕ

Апстракт: Овој труд истражува сценарио на интероперабилност помеѓу технологиите за безжичен пренос на енергија и за федеративно учење. Во ова сценарио, базната станица не го координира само процесот на тренирање на глобалниот модел добиен со федеративно учење, туку истовремено ги полни со енергија и клиентите што жнеат енергија, кои пак се одговорни за тренирање на нивните локални модели. Овие клиенти што жнеат енергија се опремени со батерии кои може да се надополнуваат што им дозволува истовремено локално да процесираат и да жнеат енергија. Во трудот предлагаме ефикасно доделување на ресурси што ги оптимизира процесирачките и комуникациските параметри за да се минимизира времето потребно за тренирање на моделот. Симулациските резултати покажуваат значајно намалување на времето потребно за тренирање во споредба со врвните системи за федеративно учење кои работат во режим без преклопување помеѓу фазите на локално процесирање и жетва на енергија.

Клучни зборови: мрежи со безжичен пренос на енергија, доделување ресурси, федеративно учење

I. INTRODUCTION

NEXT-generation communication systems are expected to be the primary vehicles for distributed machine learning, such as federated learning (FL). In turn, as stipulated in the IEEE 3652.1 standard [1], FL will contribute to solving data privacy and information security concerns relevant to these communication systems. The FL concept is particularly useful for the Internet of Things (IoT) ecosystem, whose number of devices and traffic volumes have grown exponentially in recent years, where FL offers opportunities to improve their limited computing capabilities and privacy guarantees [2]. Additionally, IoT devices have limited power supply despite the requirements for their energy sustainability over extended periods of time. In this regard, feasible technologies to facilitate energy sustainability of IoT are wireless power transfer (WPT) and radio frequency (RF) energy harvesting (EH) [3]. Therefore, the fusion of FL and wireless power communications appears as the natural mixture of technologies for the design of energy self-sustainable intelligent systems and services for the next-generation communications.

To present, several papers available in the literature have studied wireless powered FL systems [4]-[6]. Paper [4] assumes that EH clients (EHCs) first harvest energy and then perform local computation and data offloading to the FL parameter server. Papers [5] and [6] propose resource allocation and scheduling schemes for wireless powered FL systems by allowing simultaneous energy harvesting and local model processing at EHCs. However, they use separate dedicated channels for communications and energy harvesting and do not offer closed-form solutions for proposed resource allocations, which makes their practical real-time implementation difficult. Furthermore, the algorithms proposed in [5] and [6] rely on alternating optimization of resource allocation subproblems, which are both suboptimal and computationally intensive.

An important interoperability issue between WPT and FL is whether the EH process runs concurrently with the FL process to enhance efficiency [7] [8]. While [4] does not assume concurrent processes, papers [5] and [6] consider parallel execution, with WPT occurring over a dedicated frequency channel separate from the communication channel. In scenarios where the local computation load at the clients is substantial, such as during FL training, enabling simultaneous energy harvesting and local computation can result in significant improvements in overall system performance. In this paper, we extend [4] by assuming that the clients are equipped with rechargeable batteries and thus can simultaneously harvest energy and carry out local processing, which changes the energy management model and the optimal resource allocation. The existence of rechargeable batteries at EHCs is already implemented in some of the off-the-shelf energy harvesting devices [9], making the analyzed system closer to practical implementation. Additionally, contrary to [5] and [6], we assume that communication and energy transmission are carried out over the same channel, using the time division principle that allows for a more compact design with a single base station and single communication channel. Thus, we design and optimize a wireless powered system for FL training where local processing and energy harvesting occur simultaneously over a common communication bandwidth. We formulate a resource allocation problem to minimize delay in a time-division multiple access (TDMA)-based wireless powered communication network (WPCN) used for FL training. Our system design is optimized using an exact analytical solution for resource allocation parameters, ensuring lower computational complexity and reduced latency compared to state-of-the-art schemes.

TABLE I

LIST OF NOTATIONS

Notation	Description
K	Number of EHC
τ_{LC}	Duration of joint local training/harvesting phase and communication phase
t_k	Duration of kth communications subphase
$\mathbf{w}_k$, $\mathbf{w}$	Local model parameters/Global model parameters
D_k	Number of samples in local dataset at kth EHC
L_k	CPU cycles/sample to process single data sample at kth EHC
f_k	CPU cycles per second at kth EHC
I_k	Number of iterations to reach some local accuracy within each TR for kth EHC
a_k	Total number of CPU cycles needed by kth EHC to process single local model update
a_0	Total number of CPU cycles needed by any EHC to process single local model update when all users have the same I_k, L_k and D_k
E_k^{EH}	Amount of energy stored by kth EHC during the joint phase
P_0	Transmit power for BS RF energy broadcasts
ξ	Energy conversion efficiency of the EH circuit
h_k	Gain of the wireless channel between BS and kth EHC
E_k^{LC}	Energy required for kth EHC to train its local model during a single TR
α	Energy efficiency coefficient
b_0	Size in bits of local model parameters transmitted from kth EHC to BS
p_k	Transmit power of kth EHC
B	Communications bandwidth
N_0	Thermal noise power spectral density
E_k^{WC}	Energy consumed by kth EHC for transmission to BS
f_{max}	Maximum CPU frequency allowed at each EHC

II. SYSTEM MODEL

We consider FL training process realized over WPCN, consisting of K EHCs, each executing its local FL model, and single base station (BS), acting as the FL parameter server. We assume each node is equipped with single antenna and operate in half-duplex mode. In addition, each EHC is equipped with EH circuit comprised of rectifying antenna and rechargeable battery.

As per Figure 1, the FL training time is divided into training rounds (TRs), where each TR consists of two phases: joint local training/harvesting phase and communication phase. During the joint phase, with duration τ_{LC}, BS emits RF energy, which is collected by EHC in channel-dependent amount. The communications phase is subdivided into K subphases based upon TDMA, where each subphase is dedicated to information transmission from single EHC. Specifically, during kth communications subphase, with duration t_k, kth EHC ($1 \le k \le K$) offloads to BS its local model parameters, $\mathbf{w}_k$. At the end of the communication phase, BS updates the global model using local model parameters from all EHCs, and then sends these parameters back to EHCs to update their local FL models. The global model update in each TR is realized according to specific FL algorithm. When the FedAvg algorithm [10] is used, assuming equal size of local datasets from all EHCs, global model updates at the end of given TR are calculated according to the following equation:

$$\mathbf{w} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{w}_k. \tag{1}$$

The time needed by BS to update the global model and to broadcast the global model parameters to EHC is neglected due to substantial bandwidth and power available to BS. This procedure is repeated over multiple TR iterations until global loss function convergences at its minimum with some prescribed accuracy.

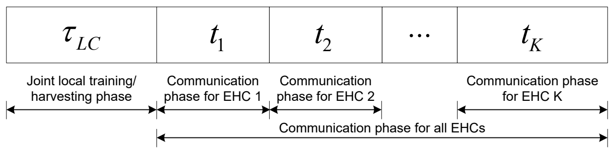

Fig. 1. Structure of single training round

A. Joint local training and energy harvesting phase

We assume that each EHC has a battery with certain initial energy stored so that each EHC can perform local processing while simultaneously harvesting energy. Let us assume that kth EHC has i.i.d. local dataset $\mathcal{D}_k$ with D_k number of samples. The computational load to process single data sample is equal to L_k CPU cycles/sample, and CPU frequency of kth EHC is equal to f_k CPU cycles/second. If the number of iterations to reach some local accuracy within each TR is equal to I_k, then the total number of CPU cycles needed by kth EHC to process single local model update is given by $a_k = L_k D_k I_k$. Without loss of generality, the values of L_k, D_k and I_k are assumed equal for all EHCs ($L_k = L_0$, $D_k = D_0$, and $I_k = I_0$), which leads to $a_k = a_0 = L_0 D_0 I_0, \forall k$. For analytical tractability, we assume the local training phases of all EHCs are completed at the same time instant. In this case, the duration of the joint phase is set as equal to:

$$\tau_{LC} = \max_{1\le k\le K}\left\{\frac{a_0}{f_k}\right\}. \tag{2}$$

Now, let us denote the gain of the wireless channel between BS and kth EHC ($1 \le k \le K$) by h_k. This channel is assumed to be reciprocal. While the local model update process is taking place at EHCs, BS broadcasts RF energy at fixed transmit power P_0 that is harvested by all EHCs. Assuming that EH circuit at EHC complies with linear EH model, the amount of energy stored by kth EHC during the joint phase is determined by:

$$E_k^{EH} = \xi P_0 \tau_{LC} h_k, \tag{3}$$

where $\xi (0 < \xi \le 1)$ denotes the energy conversion efficiency of the EH circuit. On the other hand, the energy required for kth EHC to train its local model during single TR is determined by:

$$E_k^{LC} = \alpha a_0 f_k^2, \tag{4}$$

where α is an energy efficiency coefficient that depends on EHC's CPU architecture [11] [12].

B. Communication phase

During kth subphase of the communication phase, kth EHC transmits to BS its local model parameters. Let us denote the size in bits of the local model parameters transmitted from kth EHC to BS by b_0. The capacity of the wireless channel between kth EHC and BS should be high enough to sustain the reliable transmission of b_0 bits, which yields the following condition:

$$t_k B \log_2\left(1+\frac{p_k h_k}{BN_0}\right) \ge b_0, \forall k, \tag{5}$$

where p_k is the transmit power of kth EHC, B is the communications bandwidth, and N_0 is the power spectral density of the thermal noise at the receiver. In this case, energy consumed by kth EHC for transmission to BS is equal to $E_k^{WC} = p_k t_k$. Therefore, the total amount of energy consumed during TR should be no higher than the total amount of energy harvested, i.e., $E_k^{WC} + E_k^{LC} \le E_k^{EH}$.

III. Latency Minimization

We aim to minimize FL training duration in the considered WPCN by optimally allocating communication parameters (τ_{LC}, t_k, p_k) and computation parameters (f_k). Given that FL training duration is defined as:

$$TR = \tau_{LC} + \sum_{k=1}^{K} t_k, \tag{6}$$

we address the following resource allocation problem:

$$\begin{aligned}
&\underset{\tau_{LC}, f_k, t_k, p_k}{\text{Minimize}}\ \tau_{LC} + \sum_{k=1}^{K} t_k \\
&\text{subject to:} \\
&C1: p_k t_k + \alpha a_0 f_k^2 \le \xi P_0 \tau_{LC} h_k, \forall k \\
&C2: t_k B \log_2\left(1+\frac{p_k h_k}{BN_0}\right) \ge b_0, \forall k \\
&C3: f_k \le f_{max}, \forall k \\
&C4: \frac{a_0}{f_k} \le \tau_{LC}, \forall k
\end{aligned} \tag{7}$$

In (7), constraint $C1$ implies that energy harvested by EHC during TR is completely spent for local training and data transmission in the same TR. Constraint $C2$ refers to the capacity of the uplink channel, whereas $C3$ applies maximum CPU frequency constraint on each EHC, f_{max}. Constraint $C4$ imposes an upper bound on duration of local training phases of all EHCs, thus satisfying (2).

The solution of (7) is given by the following theorem.

Theorem 1: The optimal transmit power of kth EHC is given by:

$$p_k^* = \frac{N_0 B}{h_k}\left[-1 - g_{1k}(f_0^*) W_{-1}\left(-\frac{1}{g_{1k}(f_0^*)} e^{-\frac{1}{g_{1k}(f_0^*)}}\right)\right] \quad (8)$$

where $W_{-1}(z)$ is second branch of the Lambert-W function, which is real-valued for $-1/e \le z < 0$. The auxiliary function $g_{1k}(f_0^*)$ is defined by:

$$g_{1k}(f_0^*) = \frac{a_0 h_k \left(\xi P_0 h_k - \alpha (f_0^*)^3\right)}{b_0 N_0 f_0^* \log(2)}, \quad (9)$$

where f_0^* is equal for all EHCs and given by:

$$f_k^* = f_0^* = \begin{cases} f_c^*, & f_c^* < f_{max} \\ f_{max}, & f_c^* \ge f_{max} \end{cases}, \forall k. \quad (10)$$

In (10), the value of f_c^* is found as the solution of the following transcendental equation:

$$\xi P_0 \sum_{k=1}^{K} \frac{h_k}{g_{2k}(f_c^*)} + \frac{1}{2\alpha (f_c^*)^3} \sum_{k=1}^{K} \frac{1}{g_{2k}(f_c^*)} = 1, \quad (11)$$

where the auxiliary function $g_{2k}(f_c^*)$ is defined by:

$$g_{2k}(f_c^*) = \left(\frac{N_0 B}{h_k} + p_k^*\right) \log\left(1 + \frac{h_k p_k^*}{N_0 B}\right) - p_k^*, \quad (12)$$

where p_k^* is a function of f_c^* according to (8). The optimal duration of the joint local training/harvesting phase is given by:

$$\tau_{LC}^* = \frac{a_0}{f_0^*}, \quad (13)$$

and the optimal transmission duration of kth EHC equals:

$$t_k^* = \frac{b_0}{B \log_2\left(1 + \frac{p_k^* h_k}{N_0 B}\right)}, \quad \forall k. \quad (14)$$

Proof: Please refer to Appendix A.

IV. NUMERICAL RESULTS

To study the performance of the proposed resource allocation scheme, we consider WPCN with $K = 10$ EHCs used for FL training. EHCs are distributed at distances $d_k = 2 \times k$ meters around BS, where k is the EHC's index. The channel between BS and kth EHC is exposed only to deterministic fading with a path loss exponent equal to 3, i.e., $h_k = 10^{-3} d_k^{-3}$. The thermal noise power spectral density equals $N_0 = -160$ dBm/Hz. The EHC's computational load equals $L_0 = 1000$ CPU cycles/sample, the computation efficiency $\alpha = 10^{-28}$, and the number of local iterations is set to $I_0 = 3$, yielding $a_0 = 3 \cdot 10^{-3}$. We also set $B = 1$ MHz and $b_0 = 1$ Mbit. The size of the EHC's local dataset equals $D_0 = 1000$ samples. On the following figures, the proposed scheme is denoted as "Parallel local processing and harvesting". As benchmark, we consider a comparable resource allocation scheme developed in [4], where local processing and EH phases do not overlap. The same system parameter settings as described above are also applied to the benchmark scheme, which is denoted as "Separate phases".

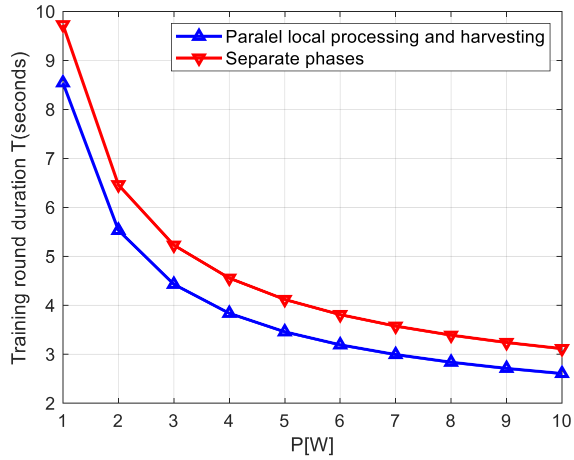

Fig. 2. Training round duration vs. BS transmit power for $f_{max} = 1$ GHz.

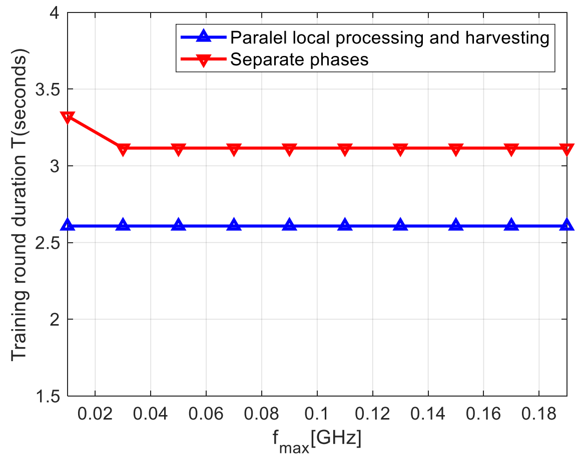

Fig. 3. Training round duration vs. f_{max} for $P_0 = 10$ W.

Figure 2 depicts the training round duration versus BS transmit power. The latency of the proposed scheme is consistently below that of the benchmark scheme, because the benchmark scheme spends extra time for local processing. Figure 3 shows the relationship between the training round duration and the maximum frequency of energy harvesting clients (EHCs). When f_{max} is low, the benchmark scheme experiences increased latency - a trend not observed with the proposed algorithm. In the benchmark scheme, the maximum frequency in this range becomes a bottleneck, leading to the optimal frequency being kept at its upper limit. Due to separation of local training and EH phases, the low optimal frequency significantly prolongs the local training phase, increasing TR duration. In contrast, the proposed algorithm allows EH and local processing to occur concurrently, enabling the local processing to run at frequency lower than the maximum without extending TR duration.

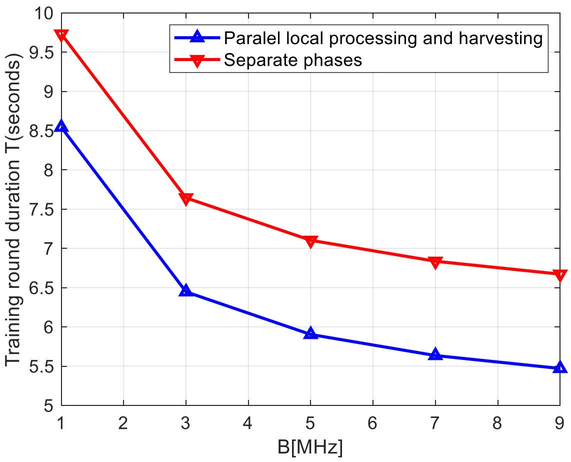

Fig. 4. Training round duration vs. B for $P_0 = 10$ W and $f_{max} = 1$ GHz.

Figure 4 illustrates the relationship between training round duration and available bandwidth. In the proposed scheme, as bandwidth increases, transmission time decreases, leading to reduced training duration. In contrast, the local processing period in the benchmark scheme is only slightly impacted by the increased bandwidth. The performance gap between the proposed and benchmark schemes increases with increasing bandwidth.

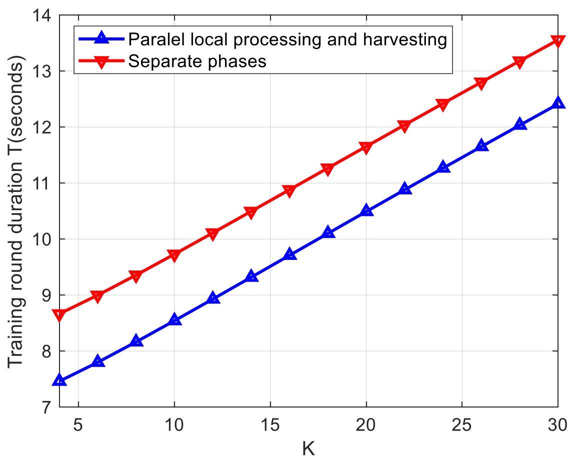

Fig. 5. Training round duration vs. K for $P_0 = 10$ W and $f_{max} = 1$ GHz.

To illustrate how the proposed method scales up with increasing number of clients and more complex networks, particularly in large-scale IoT deployments, Figure 5 shows the training round duration versus the number of EHCs. For this figure, EHCs were placed at equal distance between 2 and 20 m. The gain obtained using parallel local processing and energy harvesting is almost constant for the whole range of K, meaning that it is most important when the number of EHCs is small. Namely, as K increases, the duration of the transmission phase increases and becomes dominant, hence the reduction in training round duration due to parallel processing becomes less important.

V. Conclusion

This paper proposes a resource allocation scheme aimed at minimizing FL training duration over wireless powered communication system. To reduce latency, local model training in each round is performed concurrently with the network's energy broadcasting and harvesting process. The resource allocation problem is analytically addressed, enabling efficient online implementation of the proposed scheme. The simulation results show that this approach significantly reduces training latency compared to systems with non-overlapping energy harvesting and local processing phases. The performance gains are particularly notable when client maximum CPU frequencies are lower, emphasizing the need for rechargeable batteries in energy-harvesting devices. In future work, we will extend our study to account for the impact of random fading in the wireless channel.

Appendix A
Proof of Theorem

The proof follows similar derivation steps as in [4]. We first use the substitution $e_k = p_t t_k$ to transform (7) into a convex optimization problem, where the solution of the Lagrangian dual problem is the desired optimal solution. The Lagrangian of the transformed problem is expressed as:

$$\mathcal{L} = \left(\tau_{LC} + \sum_{k=1}^{K} t_k\right) + \sum_{i=1}^{K} \lambda_k \left[e_k + \alpha a_0 f_k^2 - \xi P_0 \tau_{LC} h_k \right] + \sum_{k=1}^{K} \mu_k \left[c_0 - t_k \log\left(1 + \frac{e_k h_k}{t_k N_0 B}\right)\right] + \sum_{k=1}^{K} \gamma_k (f_k - f_{max}) + \sum_{k=1}^{K} \beta_k \left(\frac{a_0}{f_k} - \tau_{LC}\right), \tag{15}$$

where λ_k, μ_k, γ_k and β_k are non-negative Lagrangian multipliers associated with $C1, C2, C3$ and $C4$ in (7), respectively. Note, $c_0 = b_0 \log(2)/B$. Next, we set the derivatives of $\mathcal{L}$ with respect to τ_{LC}, f_k, t_k and e_k to zero, yielding:

$$\frac{\partial \mathcal{L}}{\partial \tau_{LC}} = \sum_{k=1}^{K} \beta_k + \sum_{k=1}^{K} \xi P_0 h_k \lambda_k - 1 = 0 \tag{16}$$

$$\frac{\partial \mathcal{L}}{\partial f_k} = 2\alpha a_0 f_k \lambda_k + \gamma_k - \beta_k \frac{a_0}{f_k^2} = 0 \tag{17}$$

$$\frac{\partial \mathcal{L}}{\partial t_k} = \mu_k \left[\log\left(1 + \frac{e_k h_k}{t_k N_0 B}\right) - \frac{e_k h_k}{t_k N_0 B + e_k h_k} \right] = 0 \tag{18}$$

$$\frac{\partial \mathcal{L}}{\partial e_k} = \mu_k - \lambda_k \frac{t_k N_k + e_k h_k}{t_k h_k} = 0. \quad (19)$$

The multipliers λ_k, μ_k and β_k are strictly positive $\forall k$, and thus, each of the constraints $C1$, $C2$ and $C4$ in (7) must be satisfied with strict equality.

A. Case 1: $\gamma_k = 0, \forall k$

In this case, $f_k = f_0 < f_{max}, \forall k$. Combining $C1$, $C2$ and $C4$, and using the substitution $e_k = p_k t_k$ we obtain:

$$\frac{c_0 p_k}{\log\left(1 + \frac{p_k h_k}{N_k}\right)} = \xi P_0 \frac{a_0}{f_0} h_k - \alpha a_0 f_0^2. \quad (20)$$

Note, (20) has a general form $p_k = b \cdot \log\left(1 + a p_k\right)$, where $ab > 1$, and can be solved in closed form as $p_k = -a^{-1} - b \cdot W_{-1}\left(-(ab)^{-1} \cdot \exp((ab)^{-1})\right)$, where $W_{-1}(\cdot)$ denotes the second branch of the Lambert-W function. Therefore, the closed form solution of (20) is given by (8). The optimal value of τ_{LC} is obtained from $C4$, yielding (13), whereas the optimal value of t_k is obtained from $C2$, yielding (14). Next, we insert (19) into (18), which yields:

$$\lambda_k = \frac{1}{g_{2k}(f_c^*)}, \quad (21)$$

where $g_{2k}(f_c^*)$ is given in (12). We obtain (11) by inserting (17) and (21) in (16).

B. Case 1: $\gamma_k > 0, \forall k$

In this case, $f_k = f_0 = f_{max}, \forall k$

REFERENCES

[1] IEEE Guide for Architectural Framework and Application of Federated Machine Learning, IEEE Std 3652.1-2020, March 2021.

[2] S. Niknam, H. S. Dhillon and J. H. Reed, “Federated Learning for Wireless Communications: Motivation, Opportunities, and Challenges”, IEEE Commun. Mag., vol.58, no.6, pp. 46-51, June 2020.

[3] Z. Chu, F. Zhou, Z. Zhu, R. Q. Hu and P. Xiao, “Wireless Powered Sensor Networks for Internet of Things: Maximum Throughput and Optimal Power Allocation”, IEEE Internet of Things Journal, vol.5, no.1, pp. 310-321, February 2018.

[4] M. Poposka, S. Pejoski, V. Rakovic, D. Denkovski, H. Gjoreski and Z. Hadzi-Velkov, “Delay Minimization of Federated Learning Over Wireless Powered Communication Networks”, IEEE Communications Letters, vol.28, no.1, pp. 108-112, January 2024.

[5] R. Hamdi, M. Chen, A. B. Said, M. Qaraqe and H. V. Poor. “Federated Learning Over Energy Harvesting Wireless Networks”, IEEE Internet of Things Journal, vol.9, no.1, pp. 92-103, January 2022.

[6] Q. Zeng, Y. Du and K. Huang. “Wirelessly Powered Federated Edge Learning: Optimal Tradeoffs Between Convergence and Power Transfer”, IEEE Trans. Wir. Comm., vol 21, pp. 680-695, January 2022.

[7] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas. “Communication-Efficient Learning of Deep Networks from Decentralized Data,” Proc. AISTATS, vol.54, 2017, pp. 1273–1282.

[8] J. Yao et al. “GoMORE: Global Model Reuse for Resource-Constrained Wireless Federated Learning”, IEEE Wir. Commun. Lett., vol.12, no.9, pp. 1543-1547, September 2023.

[9] S. Shen, J. Kim and B. Clerckx. “Closed-Loop Wireless Power Transfer With Adaptive Waveform and Beamforming: Design, Prototype, and Experiment”. IEEE Journal of Microwaves, vol.3, no.1, pp. 29-42, January 2023.

[10] H. B. McMahan et al. “Communication-Efficient Learning of Deep Networks from Decentralized Data”. Proc. 20th Int. Conf. Artif. Intell. Statist., Lauderdale, FL, USA, 2017, pp. 1273-1282.

[11] Q. Zeng, Y. Du and K. Huang, “Wirelessly Powered Federated Edge Learning: Optimal Tradeoffs Between Convergence and Power Transfer”, IEEE Trans. Wir. Commun., vol.21, no.1, pp. 680-695, January 2022.

[12] Y. Wu et al. “Non-Orthogonal Multiple Access Assisted Federated Learning via Wireless Power Transfer: A Cost-Efficient Approach”, IEEE Trans. Commun., vol.70, no.4, pp. 2853-2869, April 2022.