Optimizing Data-Driven Federated Learning in UAV Networks

Datian Li*, Mingjun Xiao*, Yin Xu*, Jie Wu†
*School of Computer Science and Technology & Suzhou Institute for Advanced Research,
University of Science and Technology of China, China
†Department of Computer and Information Sciences, Temple University, USA
Corresponding Email: xiaomj@ustc.edu.cn

***Abstract*—Federated Learning (FL) is an emerging privacy-preserving distributed machine learning paradigm that enables numerous clients to collaboratively train a global model without transmitting private datasets to the FL server. Unlike most existing research, this paper introduces a Data-Driven FL system in Unmanned Aerial Vehicle (UAV) networks, named DDFL, which features an innovative three-layer architecture. In this architecture, UAVs, inherently lacking data, serve as mobile clients and are tasked with periodically collecting desired data from a group of Points of Interest (PoIs) for FL training. Meanwhile, a Base Station (BS) coordinates these UAVs to efficiently train a high-quality global model. Our objective is to determine a data collection strategy for each UAV to minimize the time required to meet the global model's loss requirement within the constraint of energy consumption. Through theoretical analysis, we establish a bound for the convergence speed of DDFL and quantify the impact of collecting data from different PoIs on the global model's loss function. Leveraging these analyses, we formulate the PoI selection problem as a novel two-stage Combinatorial Multi-Armed Bandit (CMAB) problem with multiple constraints. We then propose an Adaptive Two-stage CMAB-based algorithm, named FedATC, to jointly optimize the data collection route and UAV velocity. Extensive simulations demonstrate that FedATC significantly reduces the time required to achieve desired model quality compared to state-of-the-art algorithms.**

***Index Terms*—Data Collection, Data-Driven Federated Learning, Multi-Armed Bandit, UAVs.**

I. Introduction

Federated Learning (FL) [1] has emerged as a promising distributed machine learning paradigm, enabling a group of clients to collaboratively train a global model under the coordination of a central server without the need to transmit private datasets to the server. This key feature effectively addresses the challenge of data silos and preserves data privacy, leading to its broad applications across various industries, including security [2]–[4], finance [5], [6], and healthcare [7], [8].

In this paper, we introduce and explore a novel Data-Driven FL system in Unmanned Aerial Vehicle (UAV) networks, named DDFL. As shown in Fig. 1, the DDFL system features an innovative three-layer architecture, comprising a server layer represented by a Base Station (BS), a client layer represented by UAVs, and a data layer represented by Points of Interest (PoIs). Consider a practical scenario where UAVs from different service providers inherently lack data and are required to periodically collect traffic data from a group of PoIs to monitor traffic conditions. Service providers are usually unwilling to share their private data, thus it is necessary to adopt an FL framework for UAVs to collect data and update models under the coordination of the BS. Since traffic data is time-sensitive, it is crucial to minimize the time taken by UAVs for data collection and to coordinate the training of a high-quality global model as quickly as possible. Meanwhile, we need to further take the following two constraints into consideration. On one hand, in order to ensure the model's high quality, the BS needs to keep the global loss function within a specified limit. On the other hand, UAVs face a significant constraint on energy consumption during data collection and model training due to their limited energy resources [9], [10]. Therefore, a critical issue is how to minimize the time consumed by UAVs, taking into account both the global loss function and the UAVs' energy consumption constraints.

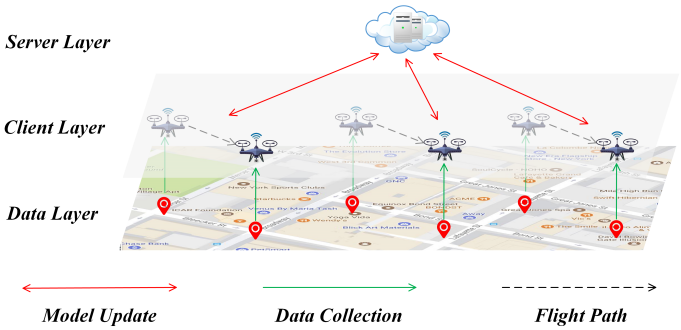


Fig. 1: The Data-Driven FL system in UAV networks.

There are two major challenges in the above-mentioned DDFL system issue. **The first challenge arises from the unknown and dynamic data quality of different PoIs**. Since the performance of the global model significantly depends on the quality of data [11], the BS aims to collect high-quality data as efficiently as possible. This requires us not only to quantify the data quality of PoIs but also to continuously

This research was supported in part by the National Natural Science Foundation of China (NSFC) (Grant No. 62172386, 62436010, 61872330, 61572457), the Natural Science Foundation of Jiangsu Province in China (Grant No. BK20231212, BK20191194), and NSF Grants CNS 2128378, CNS 2107014, CNS 1824440, CNS 1828363, CNS 1757533, CNS 1629746, and CNS 1651947.

explore and update it. Several works have been devoted to the selection of sample data within the client in recent years [12], [13]. However, most of the existing work assumes that the data quality of PoIs is known and invariable in advance, which is not true in practice. **The second challenge is the energy consumption and speed limitations of UAVs, which make accelerating the DDFL training process through effective PoI selection more difficult.** Given the limited energy consumption of UAVs, only a few PoIs can be selected for data collection at a time, limiting the convergence speed of FL. Additionally, the selection of diverse PoIs, coupled with the varying speeds of different UAVs, lead to differing levels of energy consumption and time delays. For instance, some PoIs may offer high-quality data but are located far from the UAVs, leading to higher energy use and potentially rendering them less ideal choices for the UAVs' data collection [14]. Therefore, the decision-making process for selecting the optimal PoIs under multiple constraints is highly challenging.

To address the above challenges, we devise an Adaptive Two-stage CMAB-based FL algorithm, named FedATC. Through rigorous theoretical analysis, we first establish an upper bound for the convergence speed of the global model in relation to the data from PoIs, guiding the quantification of data quality for different PoIs. We then formalize the PoI selection problem and apply transformations to recast it as a Combinatorial Multi-Armed Bandit (CMAB) problem. We employ a two-stage CMAB-based algorithm to obtain an approximately optimal PoI selection with multiple constraints and optimize the route using the 2-Opt-TSP heuristic algorithm [15]. We also propose an iteration-based algorithm to jointly optimize the data collection route and UAV velocity.

Our major contributions are summarized as follows:

- We introduce a DDFL system in UAV networks, where UAVs are required to periodically collect desired data from a group of PoIs so as to minimize time delays within the constraints of energy consumption and the global loss function. To the best of our knowledge, this is the first FL system that abstracts a data layer in addition to the conventional server layer and client layer.
- We derive an upper bound for the convergence speed of the novel DDFL system, whereby we analyze the relationship between the training loss of the global model and the data from PoIs. Based on this analysis, we quantify the data quality of PoIs and model the PoI selection problem as a novel two-stage CMAB problem.
- We propose the FedATC algorithm. First, we divide the data collection route problem into two steps: PoI selection and route planning. We employ a two-stage CMAB-based algorithm to solve the PoI selection problem and prove the approximate regret of this algorithm. Subsequently, we propose an iteration-based algorithm to jointly optimize the data collection route and UAV velocity.
- We conduct extensive simulations based on multiple datasets. The results demonstrate that the performance of the FedATC algorithm is superior to other algorithms.

II. System Model and Problem Formulation

A. System Overview

We consider a DDFL system in UAV networks, comprising a central BS, a set of UAVs denoted as $\mathcal{U} \triangleq \{1, 2, \ldots, u, \ldots, U\}$, and a group of PoIs denoted as $\mathcal{P} \triangleq \{1, 2, \ldots, p, \ldots, P\}$, as illustrated in Fig. 1. The BS updates the global model and coordinates the flight paths and velocities of the UAVs. Each UAV u is assigned to collect data and train the local model along a flight path $\mathcal{R}_u^t$. Hereinafter, the unordered set of PoIs corresponding to $\mathcal{R}_u^t$ is denoted as $\mathcal{P}_u^t$. Each PoI can be selected by at most one UAV at a time. Specifically, the whole FL system works as follows.

1) The BS initiates the process by transmitting the initial model and flight route to the UAVs. All UAVs simultaneously commence their flights and begin data collection.
2) Once UAV u has collected data sets $\mathcal{D}_p$ from PoI p, it commences a major training round on $\mathcal{D}_p$, encompassing m local rounds. The loss function for each local round is articulated as:
$$F_u^t(w_u^{t-1,i}; D_p) = \frac{1}{|D_p|} \sum_{(x,y)\in D_p} f\left(w_u^{t-1,i}, x, y\right), \quad (1)$$
where $w_u^{t-1,i}$ denotes the model parameters of UAV u at the i-th local round within the t-th major round, and $f(\cdot)$ is a loss function defined by the server, e.g., cross-entropy loss. Then, all local updates along the path $\mathcal{P}_u^t$ can be cumulatively represented as:
$$w_u^t = w^{t-1} - \eta_t \sum_{p\in\mathcal{P}_u^t} \sum_{i=0}^{m-1} \nabla F_u^t(w_u^{t-1,i}; D_p), \quad (2)$$
where η_t is the learning rate for the t-th major round. Here, w^{t-1} is the global model from the previous major round, which also serves as the initial model for the current round. After collecting data and training the model along path $\mathcal{P}_u^t$, the local model of UAV u is updated to w_u^t. Finally, UAV u uploads w_u^t to the server.
3) In an effort to minimize the waiting time for straggling devices, the BS employs an asynchronous update mechanism. Upon receiving the local model uploaded by UAV u, the BS updates the global model as follows:
$$w^t = \left(1 - \frac{1}{U}\right) w^{t-1} + \frac{1}{U} w_u^t. \quad (3)$$
After aggregation, the BS then sends w^t along with the updated flight route and flight velocity to UAV u. UAV u subsequently initiates a new round of data collection and model training.

Overall, the global loss function is defined as follows:
$$F(w^t) \triangleq \frac{1}{U} \sum_{u=1}^{U} F_u^t(w^t; D_u^t). \quad (4)$$
where D_u^t is the local training dataset of UAV u. The goal of the model training process is to obtain the optimal model parameter vector $(w^t)^*$ so as to minimize $F(w^t)$, i.e.,
$$(w^t)^* = \arg\min_{w^t} F(w^t). \quad (5)$$

B. Energy and Time-Delay Model

In this section, we will conduct a detailed analysis of the energy consumption and time delay incurred during the data collection and model training process of the UAVs.

Fristly, when the UAV is flying at a velocity v_u, the power consumption $(\rho_u)_{\text{fly}}$ is given by [9]:

$$(\rho_u)_{\text{fly}} = \rho_1 \left(1 + \frac{3(v_u)^2}{v_{\text{tip}}^2}\right) + \rho_2 \left(1 + \frac{(v_u)^4}{4v_0^4} - \frac{(v_u)^2}{2v_0^2}\right)^{\frac{1}{2}} + \frac{1}{2}\rho_3 (v_u)^3, \tag{6}$$

where ρ_1, ρ_2, and ρ_3 are constants corresponding to blade profile power, induced power, and air resistance, respectively. The term v_{tip} signifies the tip speed of the rotor blade, while v_0 denotes the mean rotor-induced velocity during hovering. Additionally, when the UAV collects data, it hovers (i.e., $v_u = 0$). The power consumption, $(\rho_u)_{\text{hover}}$, is given by:

$$(\rho_u)_{\text{hover}} = \rho_1 + \rho_2. \tag{7}$$

The total data collection time T_{collect} for n PoIs is $T_{\text{collect}} = \tau_{\text{collect}} \cdot n$, assuming uniform collection time τ_{collect} for all PoIs. Moreover, the total UAV motion time T_{move} can be calculated as $T_{\text{move}} = \frac{L}{v_u}$, where L is the flying distance and v_u is the UAV's speed.

The total energy consumption of the UAV, accounting for both flying and hovering phases, can be calculated as:

$$E_{\text{flight-hover}}(L, n, v_u) = (\rho_u)_{\text{fly}} \cdot T_{\text{move}} + (\rho_u)_{\text{hover}} \cdot T_{\text{collect}}. \tag{8}$$

The energy consumption for local training when collecting data from n PoIs is given by [16]:

$$E_{\text{local-train}}(n) = n \cdot m \cdot \varsigma \cdot \omega \cdot \vartheta^2 \cdot Z, \tag{9}$$

where ϑ represents the CPU clock frequency of each UAV u, ω is the number of CPU cycles required per bit of data, ς is the energy consumption coefficient of the UAV's chip, and Z signifies the number of bits required for one local round of model training. Based on the above analysis, the overall energy consumption is calculated as:

$$E_{\text{total}}(L, n, v_u) = E_{\text{flight-hover}}(L, n, v_u) + E_{\text{local-train}}(n). \tag{10}$$

For simplicity, we assume that backup energy reserves are sufficient to cover the energy consumption for data transmission and the return flight to the charging station. Consequently, these energy costs are not included in E_{total} and γ_{E_u}.

To minimize the overall time delay, UAVs conduct local training synchronously during UAV flight. Therefore, the total time elapsed is:

$$T_{\text{total}}(L, n, v_u) = T_{\text{move}} + T_{\text{collect}}. \tag{11}$$

C. Problem Formulation

To jointly design the UAV data collection strategy and the FL algorithm, we formulate an optimization problem that aims to minimize the time delay for the requested UAV u, under the constraints of the loss function and energy consumption. Additionally, practical considerations require us to account for the upper limit on the UAV's velocity. Consequently, the minimization problem is formalized as follows:

TABLE I: Description of Major Notations

Variable	Description
$\mathcal{U}, u$	The set of UAVs and a specific UAV.
$\mathcal{P}, p$	The group of PoIs and a specific PoI.
$\mathcal{R}_1^*, \mathcal{R}_2^*$	The optimal paths for the first and second stages.
γ_{E_u}	The energy consumption limit of UAV u.
γ_{L}	The limit of the global model loss function.
F_u^t	The local loss function of UAV u in the t-th major round.
$w_u^{t,i}, w_u^t$	Model parameters of UAV u in the i-th local round of the t-th and t-th major rounds.
η_t	The learning rate in the t-th major round.
$D_u^t, \mathcal{D}_p$	The local dataset of UAV u in the t-th major round, and the dataset collected from PoI p.
v_u	The velocity of UAV.
L, l_p	The total flying distance, and the distance between the UAV's location and PoI p.
T_{total}	The total time in the t-th major round.
E_{total}	The total energy consumption in the t-th major round.
$\lambda, \mathcal{L}$	The vector of Lagrange multipliers and the Lagrange dual function in the t-th major round.
q_1^t, q_2^t	The quality values for PoIs at different stages in the t-th major round.
c_1, c_2	The costs for selecting a PoI at a distance l at different stages in the t-th major round.
B_1, B_2	Total energy consumption bounds for different stages in the t-th major round.

$$\textbf{P1:} \quad \min_{\mathcal{R}_u^t, v_u} \quad T_{\text{total}}(L, n, v_u), \tag{12a}$$

$$\text{s. t.} \quad F(w^t) - F^*(w^t) \leq \gamma_{\text{L}}, \tag{12b}$$

$$E_{\text{total}}(L, n, v_u) \leq \gamma_{\text{E}_u}, \tag{12c}$$

$$0 \leq v_u \leq v_{\max}, \tag{12d}$$

where γ_{L} and γ_{E_u} are predetermined positive constants. $F^*(w^t)$ denotes the minimized global loss function under the constraints of UAV speed and energy consumption. Our optimization variables are the UAV's route $\mathcal{R}_u^t$ and velocity v_u, as L, n, and $F(w^t)$ are all controlled by $\mathcal{R}_u^t$. For ease of reference, we list the major notations in Table I.

III. Theoretical Analysis

In this section, we derive two theorems that lay the theoretical foundation for the algorithm design in Section IV. Theorem 1 establishes an upper bound on the convergence rate of the DDFL system, which is beneficial for solving $F^*(w^t)$. Theorem 2 derives an upper bound for the left-hand side of Eq. (12b), aiding in the resolution of problem.

To facilitate our theoretical analysis, we adhere to a classical assumption widely employed in the FL literature [17]–[20].

Assumption 1 (Lipschitz Gradient)**.** *For each UAV $u \in \mathcal{U}$, the loss function $F_u^t(w)$ is $\mathcal{K}_u$-Lipschitz gradient, i.e., $\|\nabla F_u^t(w_1) - \nabla F_u^t(w_2)\|_2 \leq \mathcal{K}_u \|w_1 - w_2\|_2$, which implies that the global loss function $F(w)$ is $\mathcal{K}$-Lipschitz gradient with $\mathcal{K} = \frac{1}{U}\sum_{u \in \mathcal{U}} \mathcal{K}_u$.*

Theorem 1 (Global Loss Reduction)**.** *Given Assumption 1, when a UAV collects data along path $\mathcal{R}_u^t$ for training local*

models, the reduction of the aggregated global loss $F(w^t)$ is bounded as follows:

$$\begin{aligned} &F\left(w^t\right)-F\left(w^{t-1}\right) \\ &\leq \sum_{p\in\mathcal{P}_u^t}\sum_{i=0}^{m-1}\sum_{(x,y)\in D_p}\Big(\alpha_p\left\|\nabla f\left(w_u^{t,i},x,y\right)\right\|^2 \\ &\quad -\beta_p\left\langle\nabla F\left(w^{t-1}\right),\nabla f\left(w_u^{t,i},x,y\right)\right\rangle\Big), \end{aligned} \tag{13}$$

where $\alpha_p=\frac{\mathcal{K}}{2U^2}\left(\frac{\eta}{|D_p|}\right)^2$ *and* $\beta_p=\frac{1}{U}\left(\frac{\eta}{|D_p|}\right)$.

Theorem 2. *Given Assumption 1, we can derive an upper bound for the discrepancy between the flight path $\mathcal{R}_u^t$ utilized by UAV u in the t-th major round and the optimal path $\mathcal{R}_1^*$ under the constraints of UAV speed and energy consumption as denoted in Eq. (12b). The bound is formulated as:*

$$F(w^t)-F^*(w^t)\leq\sum_{p\in\mathcal{P}_1^*}\sum_{i=0}^{m-1}\boldsymbol{V}_p-\sum_{p\in\mathcal{P}_u^t}\sum_{i=0}^{m-1}\boldsymbol{V}_p, \tag{14}$$

where the term $\boldsymbol{V}_p$ is defined as:

$$\begin{aligned}\boldsymbol{V}_p \;&=\tfrac{\eta}{U}\left\langle\textstyle\sum_{(x,y)\in D_p}\frac{1}{|D_p|}\nabla f\left(w^{t-1},x,y\right),\nabla F\left((w^t)^*\right)\right\rangle \\ &\quad+\tfrac{\mathcal{K}\eta^2}{2U^2}\left\|\textstyle\sum_{(x,y)\in D_p}\frac{1}{|D_p|}\nabla f\left(w^{t-1},x,y\right)\right\|^2,\end{aligned} \tag{15}$$

and can be interpreted as the data quality for PoI p.

IV. Algorithm Design

A. Using the Convergence Bound to Convert Problem

According to Theorem 2, we have obtained an upper bound for the left-hand side of Eq. (12b), through which we can control the satisfaction of constraint 12b. Hence, Problem P1 can be reformulated as follows:

$$\textbf{P2:}\quad \min_{\mathcal{R}_u^t,v_u}\quad T_{\text{total}}(L,n,v_u), \tag{16a}$$

$$\text{s. t.}\quad \sum_{p\in\mathcal{P}_1^*}\sum_{i=0}^{m-1}\boldsymbol{V}_p-\sum_{p\in\mathcal{P}_u^t}\sum_{i=0}^{m-1}\boldsymbol{V}_p\leq\gamma_{\text{L}}, \tag{16b}$$

$$E_{\text{total}}(L,n,v_u)\leq\gamma_{\text{E}_u},\quad\forall u\in\mathcal{U}, \tag{16c}$$

$$0\leq v_u\leq v_{\max}. \tag{16d}$$

According to Problem P2, we need to determine the UAV's velocity v_u and its flight path $\mathcal{R}_u^t$, because L, n, and $\mathcal{P}_u^t$ are determined by $\mathcal{R}_u^t$. However, path planning is an NP-hard problem, and finding even an approximate solution in this scenario is challenging. Furthermore, both $\mathcal{P}_u^t$ and v_u simultaneously affect energy consumption and time delay, making the problem more complex.

To address this challenge, we decouple v_u from $\mathcal{R}_u^t$ and divide the path planning problem into two steps: selecting a subset of PoIs and solving for the optimal path based on the selected PoIs. In Section IV-B, we assume that the UAV's velocity is fixed and formulate the PoI selection problem as a novel two-stage CMAB problem for resolution. In Section IV-C, we solve for the optimal path based on the selected PoIs and adaptively adjust the UAV's velocity using an iterative method.

B. CMAB Modeling and Solution

To solve the PoI selection problem, we formulate it as a CMAB problem. Unlike the typical CMAB problem, $\mathcal{P}_1^*$ is an unknown set that requires resolution. To tackle this, we adopt a two-stage CMAB-based approach. In the first stage, we seek $\mathcal{P}_1^*$ that satisfies only the energy consumption constraint. In the second stage, based on $\mathcal{P}_1^*$, we find the optimal set of PoIs $\mathcal{P}_2^*$ corresponding to Problem P2. However, this problem is a special case of the 0-1 knapsack problem, which is NP-hard [21], indicating that there is no polynomial-time optimal algorithm. To address this problem, we extend the UWR algorithm [22].

Firstly, we utilize the Lagrangian Dual approach to transform Problem P2 into a max-min problem, denoted as Problem P3, which is formulated as follows:

$$\begin{aligned}\textbf{P3:}\quad &\max_{\lambda}\min_{\mathcal{P}_u^t}\quad \mathcal{L}(\mathcal{P}_u^t,\lambda)\\ &\text{s.t.}\quad \lambda\geq 0.\end{aligned} \tag{17}$$

Here, the Lagrange dual function $\mathcal{L}(\mathcal{P}_u^t,\lambda)$ is defined as:

$$\mathcal{L}(\mathcal{P}_u^t,\lambda)=T_{\text{total}}-\lambda_1 g_1-\lambda_2 g_2, \tag{18}$$

where:

- $g_1=\gamma_{\text{L}}-\sum_{p\in\mathcal{P}_1^*}\sum_{i=0}^{m-1}\boldsymbol{V}_p+\sum_{p\in\mathcal{P}_u^t}\sum_{i=0}^{m-1}\boldsymbol{V}_p$,
- $g_2=\gamma_{\text{E}_u}-E_{\text{total}}$, for all $u\in\mathcal{U}$.

$\lambda=(\lambda_1,\lambda_2)$ represents the vector of Lagrange multipliers.

To address Problem P3, we first aim to solve the problem of $\min_{\mathcal{P}_u^t}\mathcal{L}(\mathcal{P}_u^t,\lambda)$, which means we need to derive the corresponding optimal set of PoIs $\mathcal{P}_2^*$ to minimize $\mathcal{L}(\mathcal{P}_u^t,\lambda)$ for any given λ. However, since $\mathcal{P}_1^*$ is unknown, we need to derive the optimal solution for the following problem P4, corresponding to $\mathcal{P}_1^*$, in the first stage:

$$\begin{aligned}\textbf{P4}\quad &\text{minimize}\quad F^*(w^t)\\ &\text{s.t.}\quad E_{\text{total}}\leq\gamma_{\text{E}_u},\quad\forall u\in\mathcal{U}.\end{aligned} \tag{19}$$

1) The First Stage: Inspired by the UWR algorithm [22], we formulate Problem P4 as a CMAB problem. Each available PoI is treated as an arm, and its contribution to the reduction of the global model loss function is considered as the reward.

We introduce N_p^t and $\bar{\nabla f_p^t}$ to record the number of times that PoI p has been selected up to round t and the average gradient of PoI p, respectively:

$$N_p^t=\begin{cases}N_p^{t-1}+1; & p\in\mathcal{P}_u^t,\\ N_p^{t-1}; & p\notin\mathcal{P}_u^t,\end{cases} \tag{22}$$

$$\bar{\nabla f_p^t}=\begin{cases}\frac{\bar{\nabla f_p^{t-1}}N_p^{t-1}+\nabla f_p^t}{N_p^{t-1}+1}; & p\in\mathcal{P}_u^t,\\ \bar{\nabla f_p^t}; & p\notin\mathcal{P}_u^t.\end{cases} \tag{23}$$

Then, we substitute $\bar{\nabla f_p^t}$ for $\nabla f\left(w_u^{t,i},x,y\right)$. According to Theorem 1, we define the quality value for each PoI as follows:

$$q_1^t(p)=\sum_{i=0}^{m-1}\sum_{x,y\in D_p}\left(\alpha_p\left\|\bar{\nabla f_p^t}\right\|^2-\beta_p\left\langle\nabla F\left(w^{t-1}\right),\bar{\nabla f_p^t}\right\rangle\right). \tag{24}$$

Algorithm 1 The Two-Stage CMAB-Based Algorithm

Require: Initial speed v_1, estimated quality $\widehat{q}_i^t(p)$
Ensure: Optimal set of PoIs $\mathcal{P}_2^t$

Initialize *total_cost* = 0
while *total_cost* $< B_1$ **do**
  For all p such that *total_cost* $+ c_1(l_p) < B_1$, select:

$$p^* = \underset{p \in \mathcal{P} \setminus \mathcal{P}_u^t}{\operatorname{argmax}} \frac{\widehat{q}_1^t(p)}{c_1(l_p)} \tag{20}$$

  if no p satisfies the constraint **then**
    Break
  Add p^* to $\mathcal{P}_1^t$
  total_cost += $c_1(l_{p^*})$
Estimate $F((w^t)^*)$ for $\mathcal{P}_1^t$ using Theorem 1 and $\bar{\nabla f_p^t}$
Compute the sum of V_p for $\mathcal{P}_1^t$ using Equation (17) and $\bar{\nabla f_p^t}$, and determine Budget B_2
Set *total_cost* = 0
while *total_cost* $< B_2$ **do**
  For all p such that *total_cost* $+ c_2(l_p) < B_2$, select:

$$p^* = \underset{p \in \mathcal{P} \setminus \mathcal{P}_u^t}{\operatorname{argmax}} \frac{\widehat{q}_2^t(p)}{c_2(l_p)} \tag{21}$$

  if no p satisfies the condition **then**
    Break
  Add p^* to $\mathcal{P}_2^t$
  total_cost += $c_2(l_{p^*})$

In order to balance the relationship between exploitation and exploration, we use $\widehat{q}_i^t(p)$ to denote the UCB-based quality value. Here, $i = 1$ is applicable for the first stage, and $i = 2$ for the second stage. That is,

$$\widehat{q}_i^t(p) = \bar{q}_i^t(p) + Q_{t,p}; \quad Q_{t,p} = \sqrt{\frac{2\ln\left(\sum_{p' \in \mathcal{P}} N_{p'}^t\right)}{N_p^t}}. \tag{25}$$

The corresponding cost for each PoI is estimated as $c_1(l_p) = T_{\text{total}}(l_p, 1, v_1)$, where l_p represents the distance between the UAV's location and the PoI, with a total energy consumption bound of $B_1 = \gamma_{E_u}$.

Based on the equations presented earlier, we adopt a greedy approach to determine the set of PoIs. In the initial stages, $\mathcal{P}_u^t$ is initialized as empty. Subsequently, we search for the element in $\mathcal{P} \setminus \mathcal{P}_u^t$ that can maximally boost the UCB-based quality function $\widehat{q}_i^t(p)$ per unit cost. In other words, our selection criterion is the ratio between the marginal value of the function $\widehat{q}_i^t(p)$ and its cost, as described below:

$$p^* = \underset{p \in \mathcal{P} \setminus \mathcal{P}_u^t}{\operatorname{argmax}} \frac{\widehat{q}_i^t(p)}{c_i(l_p)}. \tag{26}$$

Following this criterion, we ensure that the algorithm prioritizes the exploration of PoIs that have not yet been selected because $\widehat{q}_i^t(p)$ tends to infinity. We continuously select p^* to be added to $\mathcal{P}_u^t$ until no PoI meets the criterion such that the cumulative cost c_1 remains within the bound B_1. At that point, we have obtained P_1^*.

Algorithm 2 The FedATC Algorithm

Require: Local model w_u^t, collected gradient information ∇f_p^t, threshold for convergence v_{step}
Ensure: Global model w^t, UAV data collection path $\mathcal{R}_2^*$, UAV flight speed v_u

Initialize $v_1 \leftarrow 0$, $v_u \leftarrow v_{\max}$
$w^t \leftarrow \frac{U-1}{U} w^{t-1} + \frac{1}{U} w_i^t$
Update $\bar{\nabla f_p^t}$ using ∇f_p^t and Eq. (23)
Update $\widehat{q}_i^t(p)$ using $\bar{\nabla f_p^t}$ and Eq. (25)
while $|v_u - v_1| < v_{\text{step}}$ **do**
  $v_1 \leftarrow v_u$
  $\mathcal{P}_2^t \leftarrow$ Alg. 1$(v_1, \widehat{q}_i^t(p))$
  Apply gradient ascent to obtain optimal λ^* and corresponding $\mathcal{P}_2^*$
  $\mathcal{R}_2^* \leftarrow$ 2-OPT-TSP$(\mathcal{P}_2^*, l_u)$
  Compute optimal speed v_u based on $\mathcal{R}_2^*$

2) The Second Stage: To derive P_2^*, we first note that since λ_1 is a positive number, the problem $\min_{\mathcal{P}_u^t} \mathcal{L}(\mathcal{P}_u^t, \lambda)$ is equivalent to $\min_{\mathcal{P}_u^t} \frac{\mathcal{L}(\mathcal{P}_u^t, \lambda)}{\lambda_1}$. After rearranging, the Lagrange dual function can be expressed as:

$$\begin{aligned}\frac{\mathcal{L}(\mathcal{P}_u^t, \lambda)}{\lambda_1} &= -\sum_{p \in \mathcal{P}_u^t} \sum_{i=0}^{m-1} \boldsymbol{V}_p - \left[-\left(\tfrac{1}{\lambda_1} T_{\text{total}} + \tfrac{\lambda_2}{\lambda_1} E_{\text{total}} \right) \right. \\ &\quad \left. + \left(\gamma_L - \sum_{p \in \mathcal{P}_1^*} \sum_{i=0}^{m-1} \boldsymbol{V}_p + \tfrac{\lambda_2}{\lambda_1} \gamma_{\mathrm{E}_u} \right) \right]. \end{aligned} \tag{27}$$

Therefore, we replace $\nabla f\left(w_u^{t,i}, x, y\right)$ with $\bar{\nabla f_p^t}$, to obtain the new quality value for each PoI as follows:

$$q_2^t(p) = \sum_{i=0}^{m-1} \boldsymbol{V}_p. \tag{28}$$

Then, the estimated cost for an individual PoI and the total resource bound are updated as:

$$c_2(l_p) = \frac{1}{\lambda_1} T_{\text{total}}(l_p, 1, v_1) + \frac{\lambda_2}{\lambda_1} E_{\text{total}}(l_p, 1, v_1), \tag{29}$$

$$B_2 = \gamma_L - \sum_{p \in \boldsymbol{\mathcal{P}_1^*}} \sum_{i=0}^{m-1} \boldsymbol{V}_p + \frac{\lambda_2}{\lambda_1} \gamma_{E_u}. \tag{30}$$

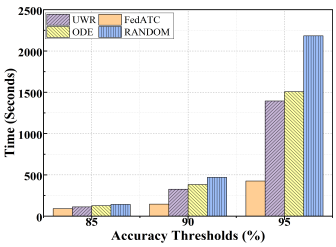
Finally, similarly to the first stage, we employ a greedy strategy to select PoIs within the total cost constraint of B_2, ultimately deriving the approximate optimal solution for $\min_{\mathcal{P}_u^t} \mathcal{L}(\mathcal{P}_u^t, \lambda)$ for any given λ.

Theorem 3. *Let Alg. 1 be implemented in the DDFL system and let the number of options for the UAVs be denoted as P_f. Then, the worst α-approximate regret of Alg. 1, symbolized by $R_2(B_2)$, can be expressed as:*

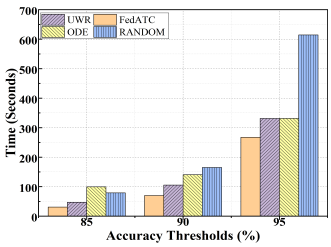
$$R_2(B_2) = O\left(P_f \ln(B_2 + P_f \ln B_1)\right). \tag{31}$$

C. The FedATC Algorithm

Since Alg. 1 has derived the solution of $\min_{\mathcal{P}_u^t} \mathcal{L}(\mathcal{P}_u^t, \lambda)$ for any given λ. Then, to solve the Lagrange dual function for Problem P3, i.e., $\max_\lambda \min_{\mathcal{P}_u^t} \mathcal{L}(\mathcal{P}_u^t, \lambda)$, we only need

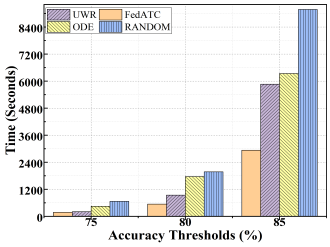


(a) 4 UAVs performance

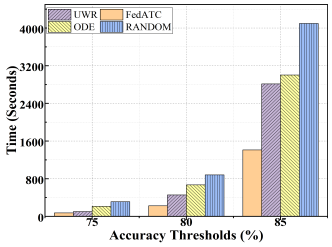


(b) 8 UAVs performance

Fig. 2: Performance across UAVs on MNIST

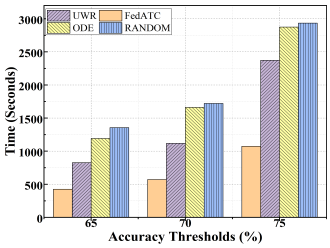


(a) 4 UAVs performance

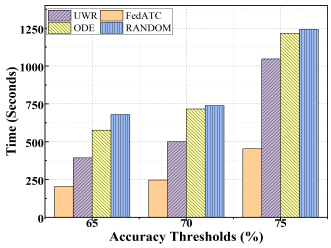


(b) 8 UAVs performance

Fig. 3: Performance across UAVs on FMNIST

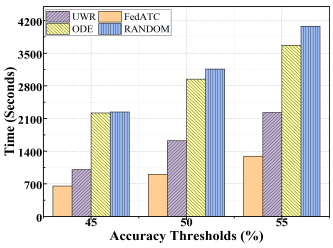


(a) 4 UAVs performance

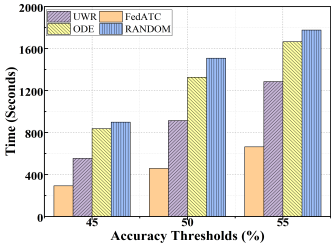


(b) 8 UAVs performance

Fig. 4: Performance across UAVs on SVHN



(a) 4 UAVs performance



(b) 8 UAVs performance

Fig. 5: Performance across UAVs on CIFAR10

to focus on the problem of $\max_\lambda L(\mathcal{P}^*, \lambda)$, i.e., finding an optimal λ to maximize $L(\mathcal{P}^*, \lambda)$. However, it is hard to find the optimal solution to solve this problem. Therefore, we will find the optimal λ (or locally optimal λ) by computing subgradients of the dual function and using the gradient ascent method [23]. At this point, we have successfully solved the PoI selection problem with the UAV speed set to v_1.

Now, having completed the PoI selection problem and obtained $\mathcal{P}_2^*$, we next calculate the optimal path. The computation of this optimal path is a generalization of the Traveling Salesman Problem (TSP). In this paper, we utilize the 2-Opt-TSP heuristic algorithm, which serves as a example.

Addressing the UAV's speed and flight path concurrently is particularly challenging. So we adopt an iterative approach to handle these components separately. Specifically, we initiate the UAV's speed as v_1 and then obtain an approximate solution to Problem P2 based on the aforementioned method. After obtaining the optimal path $\mathcal{R}_2^*$ for P2, we then re-estimate the optimal speed. We can approach the optimal speed solution using gradient descent, but considering that the UAV speed is often a positive integer, we can adopt a more efficient method. We start from $v_{\max}$ and decrement until the first speed that meets the energy consumption limit, which is considered the optimal speed. This process is repeated until the difference in speeds between consecutive iterations is less than a predetermined threshold.

Finally, we analyze the performance of the FedATC algorithm. The computational complexity of the 2-Opt-TSP algorithm is $O(N_T P^2)$, where N_T represents the number of iterations and has been proven to converge [15]. Assuming Step 5 of FedATC requires N_L iterations, N_L is also bounded as $v_{\max}/v_{\text{step}}$. In practical applications, N_T and N_L can be manually capped to provide a balance between computational resources, time efficiency, and model performance. Consequently, the computational complexity of FedATC, dominated by the main loop (Steps 5-10), is $O(P^2 v_{\max}(N_L + N_T)/v_{\text{step}})$.

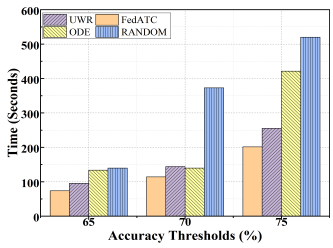
V. PERFORMANCE EVALUATION

A. Evaluation Methodology

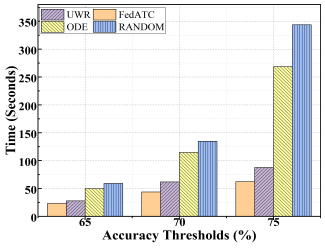
1) Simulation Setup: Our research involves extensive simulations utilizing four major datasets: MNIST [24], Fashion-MNIST (FMNIST) [25], SVHN [26], and CIFAR-10 [27].

In our default simulation setup, we deploy a network of 20 PoIs with an energy bound set to 30×10^4. We consider data distribution across different PoIs to be Independent and Identically Distributed (IID) by default, and achieve non-IID [28] effects by controlling the dataset size differences among PoIs and the ordered splitting of datasets. The default model for different datasets is Convolutional Neural Network (CNN), equipped with dataset-specific hyperparameters and learning rate strategies. We ensure consistency in model initialization, data distribution, the location of PoIs and the BS, as well as the initial flight paths across simulations. To note, to mimic the data quality of different PoIs, we introduce varying degrees of feature noise and label noise to different PoIs.

2) Algorithms for Comparison: In the context of UAV-based FL system, our FedATC algorithm addresses data-driven challenges, for which no existing algorithms are directly applicable. To the best of our knowledge, the UWR algorithm, as proposed in [22] and based on the CMAB framework, is the most relevant existing approach. However, we need to modify it to adapt to the DDFL system, specifically to minimize the global model's loss function under the constraints of UAV speed and energy consumption.

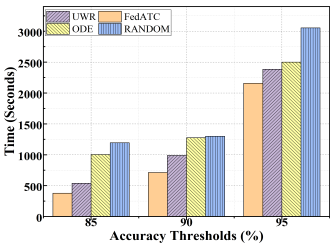


(a) 4 UAVs performance

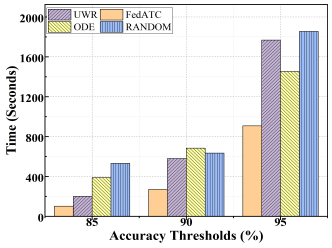


(b) 8 UAVs performance

Fig. 6: Performance of LR on MNIST



(a) 4 UAVs performance



(b) 8 UAVs performance

Fig. 7: Performance on MNIST with non-IID Data

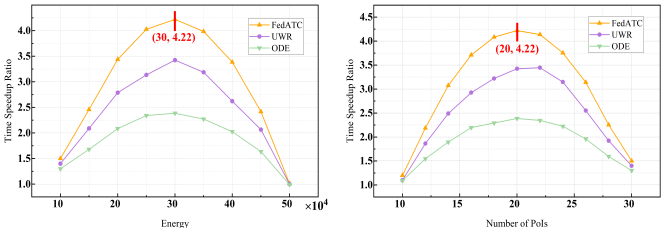


Fig. 8: Time Speedup Ratio

Additionally, we implement the ODE algorithm [11] and a Random algorithm for a more comprehensive comparison. Similar to our quantification of data quality for PoI p, denoted as V_p, the ODE algorithm employs a greedy methodology, selecting the most valuable sample at each step while ensuring that energy consumption remains within the defined limits. In subsequent simulations, we modify the concept of value to align with the convergence rate derived in Theorem 1, thereby fitting the unique requirements of our DDFL system.

B. Evaluation Results

In this section, we present a comprehensive evaluation of the FedATC algorithm's performance under various conditions. Our analysis spans several facets, including algorithmic efficiency across different datasets, the effect of model complexity by transitioning from CNN to Logistic Regression (LR) models, the robustness of FedATC in non-IID data scenarios, and the impact of energy bounds and the number of PoIs on the time speed-up ratio. The results highlight FedATC's consistent outperformance of other algorithms and underscore its adaptability to diverse scenarios and configurations.

First, we compare the cumulative time taken by four algorithms to achieve the required accuracy on different datasets and with varying numbers of UAVs, as shown in Fig. 2, Fig. 3, Fig. 4, and Fig. 5. Accuracy measures the number of correct predictions. For each of the four datasets, we set three typical accuracy thresholds, specifically the nearest accuracy divisible by 5 that approximates model convergence. In Fig. 2, it is observed that, regardless of whether there are 4 or 8 UAVs, the time taken by FedATC is significantly less than the three other algorithms. Moreover, the higher the required accuracy, the more pronounced the time superiority of FedATC is. The results for FMNIST, SVHN, and CIFAR10, as shown in Fig. 3, Fig. 4, and Fig. 5 respectively, exhibit similar patterns as the MNIST dataset.

Upon transitioning from a CNN to a LR model, we conduct a comparative analysis of the cumulative time taken by the four algorithms to reach the desired accuracy levels. Due to limited space, we only illustrate the result of MNIST. As depicted in Fig. 6, the FedATC algorithm consistently outperforms the other three algorithms. Additionally, deploying 8 UAVs significantly reduces the cumulative time compared to 4 UAVs, although the relative performance gap between the different algorithms narrows. This reduction could be attributed to the diminishing number of PoIs available for selection by each UAV, which in turn slightly diminishes FedATC's time efficiency advantage. These results demonstrate the robustness of the FedATC algorithm across different models.

Furthermore, we extend our analysis to non-IID data scenarios to evaluate the robustness of FedATC. As depicted in Fig. 7, FedATC demonstrates better time efficiency than the comparative algorithms, particularly at the accuracy threshold of 95%. Nevertheless, the performance gain of FedATC over other algorithms is less significant with non-IID data, highlighting the influence of data distribution on algorithm efficiency. Despite this, the results confirm FedATC's adaptability and consistent performance advantage, even under the challenging conditions presented by non-IID data distributions.

Finally, we examine the impact of varying energy bounds and varying numbers of PoIs on the time speed-up ratio across different algorithms. The time speed-up ratio is computed by $\Delta = \frac{T_{\text{random}}}{T}$. As depicted in Fig. 8, the analysis on the MNIST dataset illustrates a non-linear relationship between the energy bounds and the time speed-up ratio. It is observed that all considered algorithms shows an initial increase in the time speed-up ratio with an increase in the energy bound, reaching a peak at 30×10^4 units. Beyond this energy level, the performance advantage begins to diminish, suggesting an optimal energy threshold for maximizing the time speed-up ratio. Concurrently, a scarcity of PoIs leads to only a marginal performance improvement over the RANDOM algorithm, indicating limited choices for FedATC. Conversely, an abundance of PoIs results in a prolonged learning phase for data quality, thereby increasing the convergence time and reducing the time speed-up ratio. These findings underscore the importance of balancing the number of PoIs and the energy bound.

VI. RELATED WORK

Federated Learning: Many studies in FL have focused on numerous optimizations and efficiency improvements [29]–[31]. Although some research has considered hierarchical FL architectures [32], [33], they primarily focus on the edge server layer and overlook the data layer. Current research in FL on the data layer mainly concentrates on the selection of sample data within clients [11]–[13], neglecting scenarios where data needs to be actively collected. Especially when the data quality of different PoIs in the data layer is unknown and varies, this issue becomes more challenging.

UAV Data Collection: Many works have studied UAV-based data collection in communication networks. For example, Zhan et al. [34] optimize the wake-up schedule and trajectory of UAVs to minimize the energy consumption of sensor nodes, while Gong et al. [35] minimize the UAV flight time while ensuring reliable data upload from sensors. Additionally, recent studies have explored mobile crowdsensing using UAVs to minimize the age of information (AoI) and energy consumption [9], as well as the energy-efficient and cooperative data collection from low-level sensors while recharging from multiple charging stations [10]. However, the integration of UAV-based data collection with FL is a novel area that, to the best of our knowledge, remains unexplored.

Multi-Armed Bandit: MAB is a well-established approach for making decisions that require a balance between exploration and exploitation [36]–[38]. CMAB algorithms provide solutions for scenarios where decisions are combinatorial in nature [22], [39]. Recent applications of MAB in FL have aimed at optimizing resource allocation among clients [40], [41]. Nonetheless, the application of a Two-stage CMAB-based algorithm for PoI selection in FL presents a unique and unexplored challenge that this work seeks to address.

VII. CONCLUSION

In this paper, we introduce a novel DDFL system in UAV networks, which for the first time abstracts a data layer beyond the conventional server and client layers. Utilizing UAVs for data collection, we conduct a theoretical analysis of the relationship between the convergence speed of the DDFL system and the data of PoIs. Building on this foundation, we model the PoI selection problem as a CMAB issue and employ a two-stage CMAB-based algorithm to achieve an approximate optimal PoI selection. We then propose the FedATC algorithm to jointly optimize the data collection route and UAV velocity, minimizing time latency under the constraints of energy consumption and the global loss function. Extensive simulations on four real datasets have substantiated the efficacy of our approach. Future research will focus on exploring the heterogeneous data distribution among PoIs.

APPENDIX

A. Proof of Theorem 1

Proof. First, we analyze the convergence rate of the global loss function in the DDFL system. Based on Assumption 1, we have

$$F(w^t) - F(w^{t-1}) \leq \langle \nabla F(w^{t-1}), w^t - w^{t-1} \rangle + \frac{\mathcal{K}}{2} \|w^t - w^{t-1}\|^2. \tag{32}$$

Next, by utilizing the asynchronous aggregation formula, Equation (3), we can further expand the above equation as:

$$\begin{aligned}(32) &= \langle \nabla F(w^{t-1}), \frac{1}{U} w_u^t + \frac{U-1}{U} w^{t-1} - w^{t-1} \rangle \\ &\quad + \frac{\mathcal{K}}{2} \left\| \frac{1}{U} w_u^t + \frac{U-1}{U} w^{t-1} - w^{t-1} \right\|^2 \\ &\leq \frac{1}{U} \langle \nabla F(w^{t-1}), w_u^t - w^{t-1} \rangle + \frac{\mathcal{K}}{2U^2} \|w_u^t - w^{t-1}\|^2.\end{aligned} \tag{33}$$

Integrating Equation (2), we obtain the following equation:

$$\begin{aligned}(33) \leq \frac{1}{U} \Bigg[&\frac{\mathcal{K}}{2U} \left\| \sum_{p \in \mathcal{P}_u^t} \sum_{i=0}^{m-1} \eta \nabla F_u^t(w_u^{t-1,i}; D_p) \right\|^2 \\ &- \eta \sum_{p \in \mathcal{P}_u^t} \sum_{i=0}^{m-1} \langle \nabla F(w^{t-1}), \nabla F_u^t(w_u^{t-1,i}; D_p) \rangle \Bigg].\end{aligned} \tag{34}$$

Finally, by incorporating the local training loss function, Equation (1), we can consolidate the following theorem:

$$\begin{aligned}(34) \leq \sum_{p \in \mathcal{P}_u^t} \sum_{i=0}^{m-1} \sum_{(x,y) \in D_p} \Big(&\alpha_p p \|\nabla f(w_u^{t,i}, x, y)\|^2 \\ &- \beta_p \langle \nabla F(w^{t-1}), \nabla f(w_u^{t,i}, x, y) \rangle \Big),\end{aligned} \tag{35}$$

where $\alpha_p = \frac{\mathcal{K}}{2U^2} \left(\frac{\eta}{|D_p|} \right)^2$ and $\beta_p = \frac{1}{U} \left(\frac{\eta}{|D_p|} \right)$. □

B. Proof of Theorem 2

Proof. Combining Equation (1) and Equation (2), we have

$$\begin{aligned}w^t - (w^t)^* = \frac{\eta}{U} \Bigg[&\sum_{p \in \boldsymbol{\mathcal{P}_1^*}} \sum_{i=0}^{m-1} \sum_{(x,y) \in D_p} \frac{1}{|D_p|} \nabla f(w^{t-1}, x, y) \\ &- \sum_{p \in \boldsymbol{P}_t} \sum_{i=0}^{m-1} \sum_{(x,y) \in D_p} \frac{1}{|D_p|} \nabla f(w^{t-1}, x, y) \Bigg].\end{aligned} \tag{36}$$

According to Assumption 1, we have

$$F(w^t) - F^*(w^t) \leq \langle \nabla F^*(w^t), w^t - (w^t)^* \rangle + \frac{\mathcal{K}}{2} \|w^t - (w^t)^*\|^2. \tag{37}$$

Substituting Equation (36) in, we get

$$(37) = \frac{\eta}{U} \langle \nabla F^*(w^t), g_{\boldsymbol{\mathcal{P}_1^*}} - g_{\boldsymbol{P}} \rangle + \frac{\mathcal{K}\eta^2}{2U^2} \|g_{\boldsymbol{\mathcal{P}_1^*}} - g_{\boldsymbol{P}}\|^2, \tag{38}$$

where

$$g_{\boldsymbol{P}} = \sum_{d \in \boldsymbol{P}} \sum_{i=0}^{m-1} \sum_{(x,y) \in D_p} \frac{1}{|D_p|} \nabla f(w^{t-1}, x, y). \tag{39}$$

Then, we organize $g_{\boldsymbol{P}}$ and obtain the following conclusion:

$$F(w^t) - F^*(w^t) \leq \sum_{p \in \mathcal{P}_1^*} \sum_{i=0}^{m-1} \boldsymbol{V}_p - \sum_{p \in \mathcal{P}_u^t} \sum_{i=0}^{m-1} \boldsymbol{V}_p. \tag{40}$$

□

C. Proof of Theorem 3

Proof. Due to limited space, we borrow the basic idea and some symbol notation in [22] to present our proof sketch. We can view the first stage of Alg. 1 as a special case of UWR. Consequently, we obtain the following result:

$$R_1(B_1) \quad \leq \alpha \sum_{p=1}^{\tau^\star(B_1)} q_1(p) - \mathbb{E}\left[\sum_{p=1}^{\tau^*(B_1)} q_1(p)\right] \tag{41}$$

$$= O\left(P_f \ln B_1\right). \tag{42}$$

Given that in Alg. 1 we actually use B_2' estimated by the first stage, the equation is formulated as follows:

$$B_2' \leq \gamma_L - \sum_{p \in \boldsymbol{\mathcal{P}}_1^\star} \sum_{i=0}^{m-1} \boldsymbol{V}_p + O\left(P_f \ln B_1\right) + \frac{\lambda_2}{\lambda_1}\gamma_{E_u}. \tag{43}$$

We can derive the worst α-approximate regret of Alg. 1:

$$R_2(B_2') = O\left(P_f \ln(B_2 + P_f \ln B_1)\right). \tag{44}$$

□

REFERENCES

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *AIStats*, 2017, pp. 1273–1282.

[2] C. Ma, J. Li, M. Ding, H. H. Yang, F. Shu, T. Q. Quek, and H. V. Poor, "On safeguarding privacy and security in the framework of federated learning," *IEEE Network*, vol. 34, no. 4, pp. 242–248, 2020.

[3] X. Zhang, F. Li, Z. Zhang, Q. Li, C. Wang, and J. Wu, "Enabling execution assurance of federated learning at untrusted participants," in *IEEE INFOCOM*, 2020, pp. 1877–1886.

[4] X. Yuan, X. Ma, L. Zhang, Y. Fang, and D. Wu, "Beyond class-level privacy leakage: Breaking record-level privacy in federated learning," *IEEE Internet of Things Journal*, vol. 9, no. 4, pp. 2555–2565, 2021.

[5] D. Byrd and A. Polychroniadou, "Differentially private secure multi-party computation for federated learning in financial applications," in *ACM AI in Finance*, 2020, pp. 1–9.

[6] H. Yu, Z. Liu, Y. Liu, T. Chen, M. Cong, X. Weng, D. Niyato, and Q. Yang, "A sustainable incentive scheme for federated learning," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 58–69, 2020.

[7] D. C. Nguyen, Q.-V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, O. Dobre, and W.-J. Hwang, "Federated learning for smart healthcare: A survey," *ACM CSUR*, vol. 55, no. 3, pp. 1–37, 2022.

[8] J. Li, Y. Meng, L. Ma, S. Du, H. Zhu, Q. Pei, and X. Shen, "A federated learning based privacy-preserving smart healthcare system," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, 2021.

[9] Z. Dai, H. Wang, C. H. Liu, R. Han, J. Tang, and G. Wang, "Mobile crowdsensing for data freshness: A deep reinforcement learning approach," in *IEEE INFOCOM*, 2021, pp. 1–10.

[10] C. H. Liu, C. Piao, and J. Tang, "Energy-efficient uav crowdsensing with multiple charging stations by deep learning," in *IEEE INFOCOM*, 2020, pp. 199–208.

[11] C. Gong, Z. Zheng, F. Wu, Y. Shao, B. Li, and G. Chen, "To store or not? online data selection for federated learning with limited storage," in *Proceedings of the ACM Web Conference 2023*, 2023, pp. 3044–3055.

[12] A. Li, L. Zhang, J. Tan, Y. Qin, J. Wang, and X.-Y. Li, "Sample-level data selection for federated learning," in *IEEE INFOCOM*, 2021, pp. 1–10.

[13] L. Nagalapatti, R. S. Mittal, and R. Narayanam, "Is your data relevant?: Dynamic selection of relevant data for federated learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 7, 2022, pp. 7859–7867.

[14] A. Trotta, M. Di Felice, L. Bononi, E. Natalizio, L. Perilli, E. F. Scarselli, T. S. Cinotti, and R. Canegallo, "Bee-drones: Energy-efficient data collection on wake-up radio-based wireless sensor networks," in *IEEE INFOCOM WKSHPS*, 2019, pp. 547–553.

[15] J. K. Lenstra and A. R. Kan, "Complexity of vehicle routing and scheduling problems," *Networks*, vol. 11, no. 2, pp. 221–227, 1981.

[16] 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," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 269–283, 2020.

[17] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, "Adaptive federated learning in resource constrained edge computing systems," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1205–1221, 2019.

[18] Y. Jin, L. Jiao, Z. Qian, S. Zhang, and S. Lu, "Learning for learning: predictive online control of federated learning with edge provisioning," in *IEEE INFOCOM*, 2021, pp. 1–10.

[19] P. Sun, X. Chen, G. Liao, and J. Huang, "A profit-maximizing model marketplace with differentially private federated learning," in *IEEE INFOCOM*, 2022, pp. 1439–1448.

[20] Y. Liu, L. Xu, X. Yuan, C. Wang, and B. Li, "The right to be forgotten in federated learning: An efficient realization with rapid retraining," in *IEEE INFOCOM*, 2022, pp. 1749–1758.

[21] E. L. Lawler, "Fast approximation algorithms for knapsack problems," in *IEEE 18th SFCS*, 1977, pp. 206–213.

[22] G. Gao, J. Wu, M. Xiao, and G. Chen, "Combinatorial multi-armed bandit based unknown worker recruitment in heterogeneous crowdsensing," in *IEEE INFOCOM*, 2020, pp. 179–188.

[23] D. W. Hearn and S. Lawphongpanich, "Lagrangian dual ascent by generalized linear programming," *Operations research letters*, vol. 8, no. 4, pp. 189–196, 1989.

[24] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[25] H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms," *arXiv preprint arXiv:1708.07747*, 2017.

[26] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, "Reading digits in natural images with unsupervised feature learning," in *NIPS workshop*, vol. 2011, 2011, p. 5.

[27] A. Krizhevsky, "Learning multiple layers of features from tiny images," Citeseer, Tech. Rep., 2009.

[28] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-iid data," *arXiv preprint arXiv:1806.00582*, 2018.

[29] X. Zhou, J. Zhao, H. Han, and C. Guet, "Joint optimization of energy consumption and completion time in federated learning," in *IEEE ICDCS*, 2022, pp. 1005–1017.

[30] J. Perazzone, S. Wang, M. Ji, and K. S. Chan, "Communication-efficient device scheduling for federated learning using stochastic optimization," in *IEEE INFOCOM*, 2022, pp. 1449–1458.

[31] Y. Jin, L. Jiao, Z. Qian, S. Zhang, S. Lu, and X. Wang, "Resource-efficient and convergence-preserving online participant selection in federated learning," in *IEEE ICDCS*, 2020, pp. 606–616.

[32] Z. Wang, H. Xu, J. Liu, H. Huang, C. Qiao, and Y. Zhao, "Resource-efficient federated learning with hierarchical aggregation in edge computing," in *IEEE INFOCOM*, 2021, pp. 1–10.

[33] Y. Deng, J. Ren, C. Tang, F. Lyu, Y. Liu, and Y. Zhang, "A hierarchical knowledge transfer framework for heterogeneous federated learning," in *IEEE INFOCOM*, 2023, pp. 1–10.

[34] C. Zhan, Y. Zeng, and R. Zhang, "Energy-efficient data collection in uav enabled wireless sensor network," *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 328–331, 2018.

[35] J. Gong, T.-H. Chang, C. Shen, and X. Chen, "Flight time minimization of uav for data collection over wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 1942–1954, 2018.

[36] J. Ye, D. Lin, K. Cai, C. Zhou, J. He, and J. C. Lui, "Data-driven rate control for rdma networks: A lightweight online learning approach," in *IEEE ICDCS*, 2023, pp. 1–11.

[37] Z. S. Karnin and O. Anava, "Multi-armed bandits: Competing with optimal sequences," *NIPS*, vol. 29, 2016.

[38] S. Zhang, M. Xiao, G. Gao, Y. Xu, and H. Sun, "Offloading tasks to unknown edge servers: A contextual multi-armed bandit approach," in *IEEE INFOCOM WKSHPS*, 2023, pp. 1–6.

[39] M. Xiao, J. Wang, H. Zhao, and G. Gao, "Unknown worker recruitment in mobile crowdsensing using cmab and auction," in *IEEE ICDCS*, 2020, pp. 1145–1155.

[40] H. Huang, R. Li, J. Liu, S. Zhou, K. Lin, and Z. Zheng, "Contextfl: Context-aware federated learning by estimating the training and reporting phases of mobile clients," in *IEEE ICDCS*, 2022, pp. 570–580.

[41] P. Han, S. Wang, and K. K. Leung, "Adaptive gradient sparsification for efficient federated learning: An online learning approach," in *IEEE ICDCS*, 2020, pp. 300–310.