Online Federated Learning on Distributed Unknown Data Using UAVs

Xichong Zhang[1], Haotian Xu[1], Yin Xu[1], Mingjun Xiao[1], Jie Wu[2], and Jinrui Zhou[1]

[1]University of Science and Technology of China
[2]Temple University

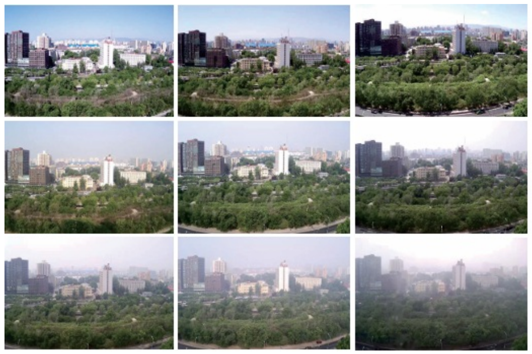
Outline

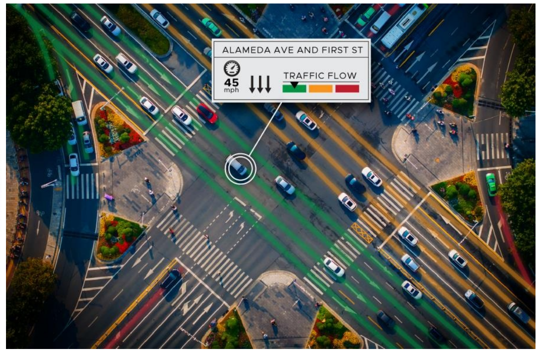
- Background & Challenges
- Related Works & Problem Formulation
- Basic Idea & Solution
- Evaluation & Conclusion

Background

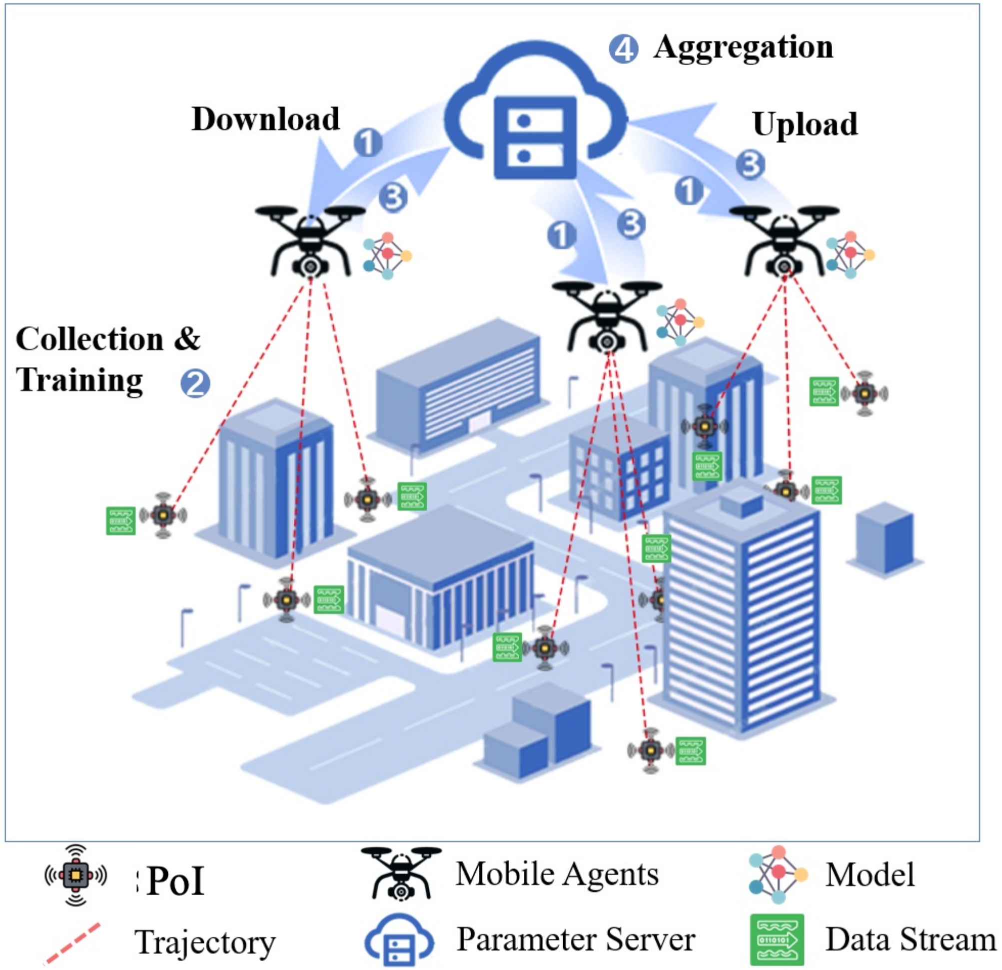
- Low-altitude Economy (LAE)
  - Services conducted at altitudes below 3000 meters
  - Logistics, Transportation, Traffic Offloading…
  - Will reach $127 billion by 2030

- Integration of UAV and Federated Learning
  - UAV: High Mobility & Flexibility
  - FL: Collaborative training without sharing data
  - Intelligent UAV applications, such as infering the air quality from images





Motivation



Traditional FL

- **Two-layer architecture:** Client-Server;
- **Inseparable client-data association:** The client maintains static data sets.

Data Collection

Data-driven FL

- **Three-layer architecture:** Data-Client-Server;
- **Separable client-data association :** data are separated from model trainers;
- **Data-driven**: FL is embedded with a dynamic data collection process.

Challenges

➢ **How to make the decision of PoI selection for UAVs to benefit model training mostly?**

- FL is embedded with an issue of **matching** between UAVs and PoIs.

  → **Quantify** the impact of matching result on the model training of FL?

- **Unknown data distribution**: UAVs do not know the data distribution of PoIs in advance.

  → UAVs have to learn data distributions while making online decisions, which involves a **trade-off** between exploration and exploitation.

Outline

- Background & Challenges
- Related Work & Game Formulation
- Basic Idea & Solution
- Evaluation & Conclusion

Related Works

- **UAV FL:** deploy FL on UAVs to combine their advantages

  e.g., Y. Wang et al. "Learning in the air: Secure federated learning for uav-assisted crowdsensing," IEEE TNSE, vol. 8, no. 2, pp. 1055–1069, 2020.

- **Data Selection:** selecting informative and valuable samples for training

  e.g., A. Li, L. Zhang, J. Tan, et al. "Sample-level data selection for federated learning," in IEEE INFOCOM, 2021, pp. 1–10.

- **Restless MAB:** all bandits might evolve stochastically

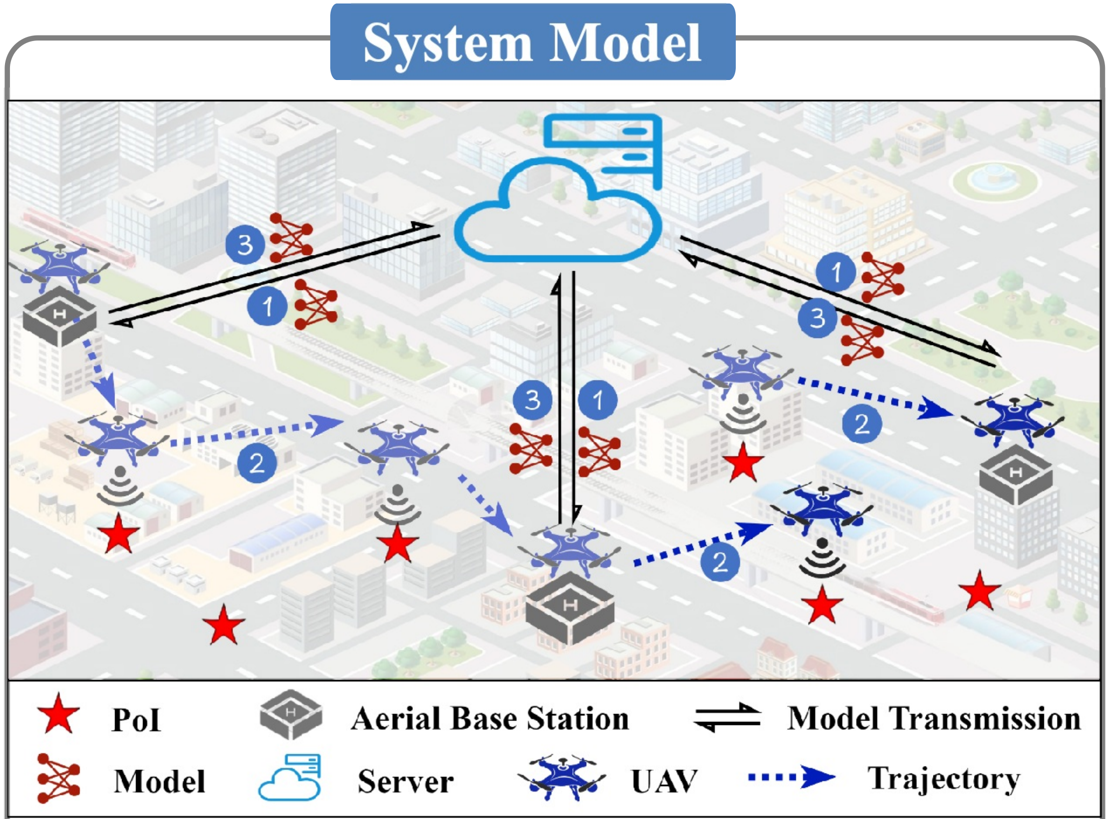
  e.g., Whittle P. "Restless bandits: Activity allocation in a changing world", in Journal of applied probability, 1988, 25(A): 287-298.

Ignore the impact of PoI selection on convergence

Ignore the dynamics of training data

We aim to design a scheduling mechanism for UAV FL so that the collected data contribute to the model training mostly.

System Model



① Server determines the scheduling of all UAVs for data collection and training

② Each UAV *n* sequentially visits the designated PoIs perform data collection and local training.

③ UAVs return to the nearest ABS for charging and global model aggregation.

System Model

- **Core Variables:** $\{1, 2, \ldots, N\}$ and $\{1, 2, \ldots, M\}$ denotes UAVs and PoIs;
  The decision variabl $\{A_1(r), \cdots, A_n(r), \cdots, A_N(r)\}$, where $A_n(r)$ represents the order in which UAV n visits PoIs at time r. And $\mathcal{P}_r = \cup_{n \in \mathcal{N}} A_n(r)$
- **Data Model:** Considering the impact of higher information entropy and more data on model generalization, we define the utility function $U_m(r)$ of PoI m in round r as $U_m(r) \triangleq |D_r^m| \cdot H_m$

$$H_m = -\sum_{b=1}^{B} q_m^b \log q_m^b$$

- **Energy Model:** when a UAV is flying at a fixed speed v, the power can be given by the expression $P_{fly} = p_1(1 + \frac{3v^2}{U_p^2}) + p_2\sqrt{(1 + \frac{v^4}{4v_0^2})^{1/2} - \frac{v^2}{2v_0^2}} + \frac{1}{2}p_3 v^3$;

  Travel Energy:

$$E_{total}(r) = E_{travel} + E_{train}$$

Problem Formulation

$$\textbf{P1:} \quad \max_{\mathcal{A}} \sum_{r=1}^{R} \sum_{m \in \mathcal{P}_r} |D_r^m| \cdot H_m,$$

$$\text{s.t.} \quad \mathbb{E}[F(\mathbf{w}_R)] - F(\mathbf{w}^*) \leqslant \epsilon,$$

$$\sum_{n \in \mathcal{N}} E_n(r) \leqslant E_0, \quad \forall r \in \mathcal{R}.$$

Optimization goal:

Find a scheduling strategy ***A*** that maximizes the generalizability of model training.

- **Constraint 1:** the gap between the **expected global loss** after R rounds and the **optimal loss**
- **Constraint 2:** the total energy constraint of UAVs, where $E_n(r) = E_{travel} + E_{train}$

Difficult to quantify the impact of PoI data on convergence.

The utility of PoI data is dynamic and unknown.

Outline

- Background & Challenges
- Related Work & Game Formulation
- Basic Idea & Solution
- Evaluation & Conclusion

Convergence Analysis

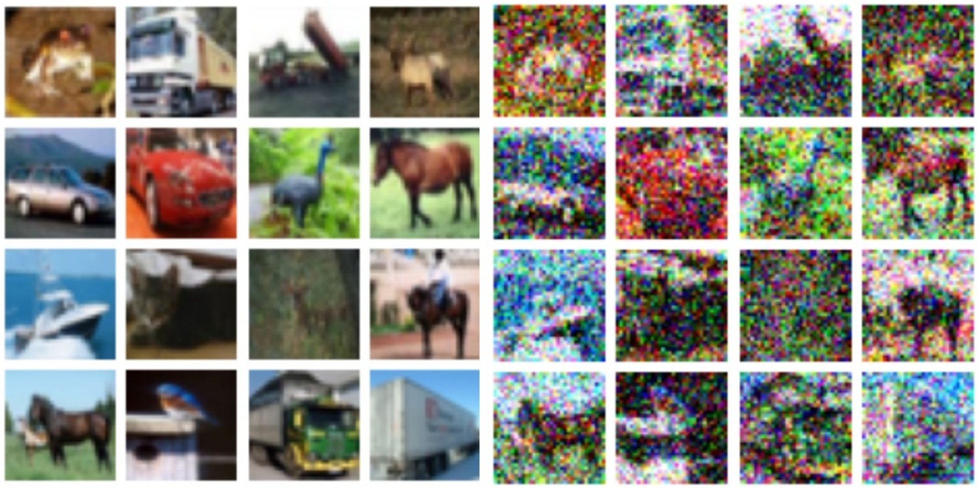
Theorem 1（Convergence bound）

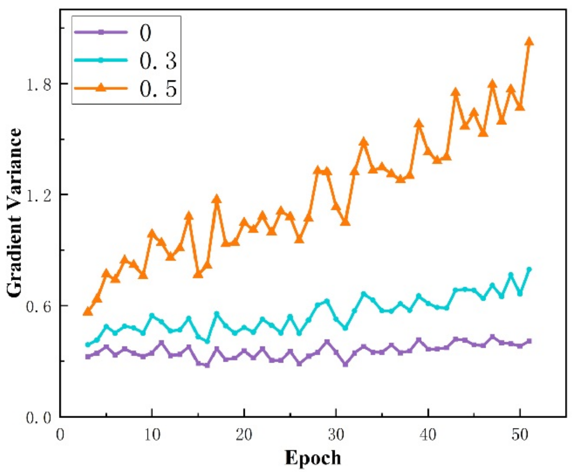
Given the PoI selection strategy $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_R\}$, after R rounds training, the difference between the expected global training loss $\mathbb{E}[F(\mathbf{w}_R)]$ and the optimal value $F(\mathbf{w}^*)$ satisfies

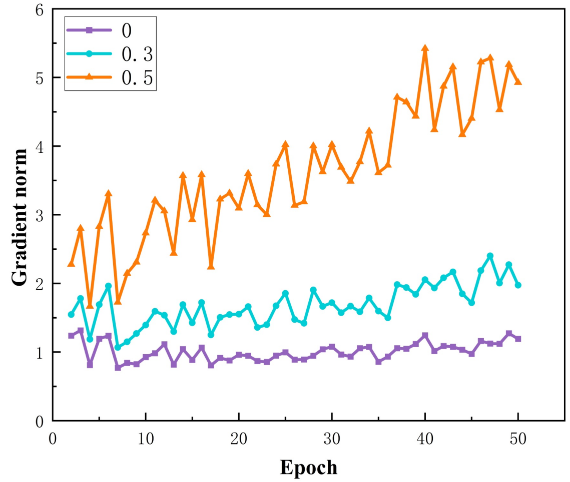
$$\mathbb{E}[F(\mathbf{w}_R)] - F(\mathbf{w}^*) \leq \frac{L}{2}(1 - \frac{\nu\bar{\eta}}{2})^R \|\mathbf{w}_0 - \mathbf{w}^*\|^2 + \boxed{\frac{L\beta_1}{2M_0\nu} \sum_{r=1}^{R} \sum_{m\in\mathcal{P}_r} [G_m^2 + \sigma_m^2]},$$

where $\beta_1 = \tilde{\eta}[\nu\tilde{\eta}K^2 + 2(K-1)^2 + 1]$ is a function of the training process with respect to the hyperparameters.

The gradient variance σ_m and the gradient upper bound G_m can be used as the value measure of PoI data from the perspective of convergence.







Convert Problem

P1:

$$\max_{\mathcal{A}} \sum_{r=1}^{R} \sum_{m \in \mathcal{P}_r} |D_r^m| \cdot H_m,$$

$$\text{s.t.} \quad \mathbb{E}[F(\mathbf{w}_R)] - F(\mathbf{w}^*) \leqslant \epsilon,$$

$$\sum_{n \in \mathcal{N}} E_n(r) \leqslant E_0, \quad \forall r \in \mathcal{R}.$$

P2:

$$\max_{\mathcal{A}} \sum_{r}^{R} \sum_{m \in \mathcal{P}_r} |D_r^m| \cdot H_m,$$

$$\text{s.t.} \quad \sum_{r=1}^{R} \sum_{m \in \mathcal{P}_r} [G_m^2 + \sigma_m^2] \leqslant \epsilon_0,$$

$$\sum_{n \in \mathcal{N}} E_n(r) \leqslant E_0, \quad \forall r \in \mathcal{R}.$$

Stage 1

In the first stage, the selected PoI set is determined. The optimization objective and constraint 1 are combined using the Lagrange dual method to transform it into the minimum-maximum problem.

Stage 2

After finding the optimal PoI set, the second restriction challenge is to consider the matching problem between drones and PoIs as an open path multi-traveling salesman problem, and use existing algorithms to solve it.

Basic Idea

- Transform the PoI selection problem into the corresponding **dual problem P3**

P3 :
$$\min_{\lambda}\max_{\mathcal{P}} \mathcal{L}(\mathcal{P},\lambda)=\sum_{r=1}^{R}\sum_{m\in\mathcal{P}_r}(|D_r^m|\cdot H_m-\lambda(\sigma_m^2+G_m^2))+\lambda\epsilon_0,$$
s.t.
$$\lambda\geq 0.$$

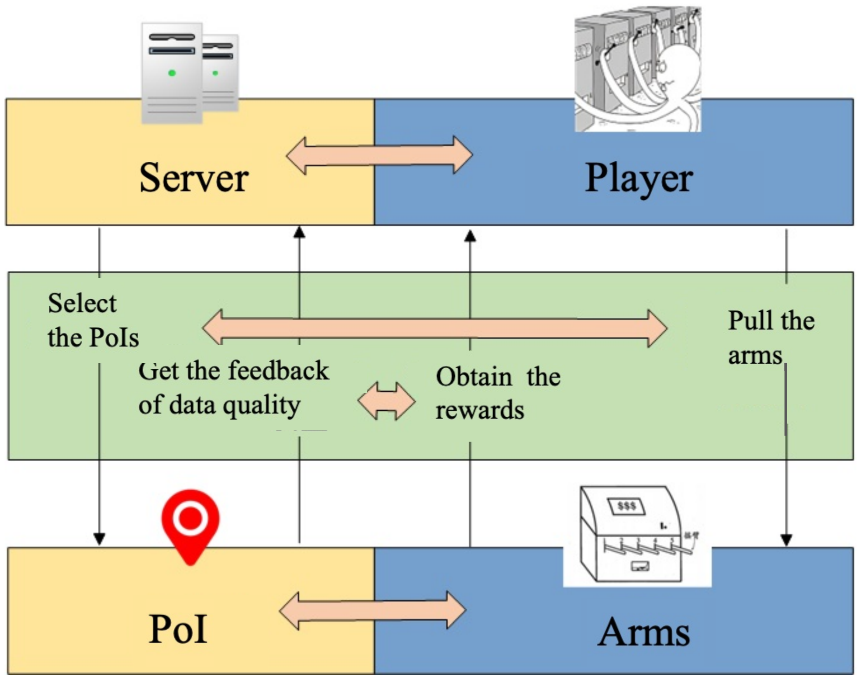
- Solve $\max_{\mathcal{P}}\mathcal{L}(\mathcal{P},\lambda)$: finding the optimal selection $\mathcal{P}$ for any given λ.

Problem **P3** can be decoupled to Problem **P4**

P4 :
$$\max_{\mathcal{P}} \mathcal{L}(\mathcal{P},\lambda)=\sum_{r=1}^{R}\sum_{m\in\mathcal{P}_r}(|D_r^m|\cdot H_m-\lambda(\sigma_m^2+G_m^2)).$$

Combinatorial Multi-Armed Bandit

- **Formulation:** The online learning and decision-making process can be modeled as a **Combinatorial Multi-Armed Bandit (CMAB)**



Servers are considered multi-armed bandit players.

The process of selecting PoIs is regarded as pulling the arms.

The data quality of the selected PoIs is regarded as the reward of the arms.

PoI is like arm the bandit.

Gaussian Process

- **Gaussian Process(GP):** Modeling the distribution of functions in function space using a Gaussian distribution → Capturing the dynamics of the reward functions flexibly.

- **Basic process：**

  a. Let $h_m(j)$ denote a specific reward function for each PoI m, where j is the index associated with the round r.

  b. Given several observations $\{(j_i, h_i)\}_{i=1}^{n}$ on the function $h_m(j)$, denoted as $\mathbf{Y}_{m,C} = (Y_{m,1}, \ldots, Y_{m,C})^T$.

  c. Then the posterior mean and variance based on C observations are:

$$\mu_r(m) = \mathbf{g}_C(j)^T(\mathbf{G}_C + \sigma^2 I)^{-1}\mathbf{Y}_{m,C},$$
$$\delta_r(m) = g_m(j, j') - \mathbf{g}_C(j)^T(\mathbf{G}_C + \sigma^2 I)^{-1}\mathbf{g}_C(j').$$

where $\mathbf{g}_C(j)$ is the prediction point correlation vector, and $\mathbf{G}_C$ is the semi-positive definite covariance matrix.

GP-CMAB-based PoI Selection

Algorithm 1: GP-CMAB-based PoI Selection

Require: $\mathcal{M}$, α_r, λ, $J_{m,r-1}$, $\mathbf{Y}_{m,C}$ for all $m \in \mathcal{M}$;
Ensure : the representative subset of PoIs, $\mathcal{P}$;

for *PoI* $m = 1, ..., M$ **do**
    Update the covariates $J_{m,r}$ according to Eq. (17);
    Calculate kernel matrix $\mathbf{G}_C = [g_m(j_i, j_l)]_{i,l=1}^{C}$;
    Train GP-CMAB model with $\mathbf{G}_C$ and observations;
    Update $\mu_r(m)$ and $\delta_r^2(m)$ by Eqs. (18) and (19);
    Calculate $\Psi_r(m)$ by Eq. (20);
Sort the PoIs according to the UCB value Ψ;
Select the top M_0 PoIs as $\mathcal{P}_r$;
Observe the feedback for each selected PoI m;
Return the set of selected PoIs in round r, $\mathcal{P}_r$.

✓ **Upper Confidence Bound (UCB):** The classic algorithm tackling the exploitation-exploration dilemma.

✓ UCB index: $\Psi_r(m) = \mu_r(m) + \alpha_r \delta_r(m)$

The classic UCB algorithm process. After sorting according to the UCB index, greedily select several rocker arms with the largest UCB index.

➢ Solving $\boldsymbol{min_{\lambda}\, \mathcal{L}(\mathcal{P}, \lambda)}$: finding the optimal $\boldsymbol{\lambda}$ through grid-search tuning.

Framework

Algorithm 1: Training process of OFL-UD^2

Input: $\mathcal{M}, \mathcal{N}, \mathcal{S}, \mathcal{R}, M_0, \mathbf{w}_0, \eta_r, \epsilon, E_0$, and λ;

Output: The global model $\mathbf{w}_R$;

1 **for** *each round* $r = 1, 2, ..., R$ **do**

2 The service provider selects a set of PoIs for UAVs according to Algorithm 1;

3 $\mathcal{A}_r \leftarrow \text{SModel}(\mathcal{A}_{r-1}, \mathcal{P}_r)$;

4 The service provider sends the current global model $\mathbf{w}_{r-1}$ and decision variables $\mathcal{A}_r$ to each UAV;

5 **for** *each UAV* $n = 1, 2, ..., N$ *in parallel* **do**

6 **for** *each PoI* $m \in \mathcal{A}_n(r)$ **do**

7 Travel and collect data at PoI m;

8 Get the observation values of $|D_r^m|$, H_m;

9 Estimate $G_m^2 = \mathbb{E}[\|\nabla F_{r,m}(\mathbf{w}_{r-1}, \xi)\|^2]$ and $\sigma_m^2 = \mathbb{E}[\|\nabla F_{r,m}(\mathbf{w}_{r-1}, \xi_m) - \nabla F_{r,m}(\mathbf{w}_{r-1})\|^2]$;

10 Calculate the observed reward based on Eq. (16): $|D_r^m| \cdot H_m - \lambda(\sigma_m^2 + G_m^2)$;

11 Initialize $\mathbf{w}_{r,n}^0 = \mathbf{w}_{r-1}$;

12 **for** *each local iteration* $k = 0, 1, ..., K-1$ **do**

13 $\mathbf{w}_{r,n}^{k+1} = \mathbf{w}_{r,n}^k - \eta_r \nabla F_{r,n}(\mathbf{w}_{r,n}^k, \xi)$;

14 Travel back to the ABS according to $\mathcal{A}_r$;

15 Upload the final model parameters $\mathbf{w}_{r,n}^K$ and reward observations to the service provider;

16 The service provider aggregates the local models to form the global model: $\mathbf{w}_r = \frac{1}{N}\sum_{i \in \mathcal{N}} p_i \mathbf{w}_{r,i}^K$;

17 The service provider learns the dynamic reward function for each PoI in the GP-CMAB model;

Line 5-11:
Each time a PoI is visited, its reward observation (data volume, information entropy, and gradient information) is obtained for subsequent fitting.

Line 12-17:
Classic federated learning process. The UAVs need to periodically return to the ABSs to communicate and charge.

Algorithm Analysis

Lemma

- ✓ (Deviation estimation) For any PoI m and constant $a \leq 0$ at episode r, it holds that

$$\int_a^{\infty} \mathbb{P}(h_m(j) - \mu_r(m) \geq x \mid \mathcal{F}_{r-1}) dx \leq \sqrt{2\pi}\delta_r(m) \exp\left\{-\frac{a^2}{2\delta_r^2(m)}\right\}$$

- ✓ (Instantaneous regret) Assume we select $\mathcal{P}_r$ according to the UCB index Ψ at time r, then instantaneous regret satisfies $\mathbb{E}[\sum_{i \in \mathcal{P}_r^*} h_i(j) - \sum_{i \in \mathcal{P}_r} h_i(j) \mid \mathcal{F}_{r-1}] \leqslant \frac{\sqrt{2\pi}}{r^2} + \alpha_r \sum_{i \in \mathcal{P}_r} \delta_r(i).$

- ✓ (Bounding GP variance) $\sum_{r=1}^{R} \sum_{i \in \mathcal{P}_r} \delta_r^2(i) \leqslant C_1 M \gamma_R,$ where $C_1 = 1/(\log(1 + \delta^{-2}))$.

Theorem

- ✓ **Theorem 2 (Regret)** The expected cumulative regret of the GP-CMAB-based PoI selection strategy satisfies

$$\mathbb{E}[Reg_R] \leqslant O(\sqrt{M M_0 R \gamma_R \log(M R J_0)}),$$

where γ_R is the maximum information gain after R rounds.

Outline

- ➢Background & Challenges
- ➢Related Work & Game Formulation
- ➢Basic Idea & Solution
- ➢Evaluation & Conclusion

Evaluation

Experimental Settings

Dataset

- Outdoor-images → a real world dataset collected via crowdsensing
- CIFAR-10, SVHN, HAM and Animals

Parameter settings

- Hardware settings: refering to the specifications of commercial drones
- Statistical settings: Non-IID level {0,1,2,4,5,10}

Compared Algorithms

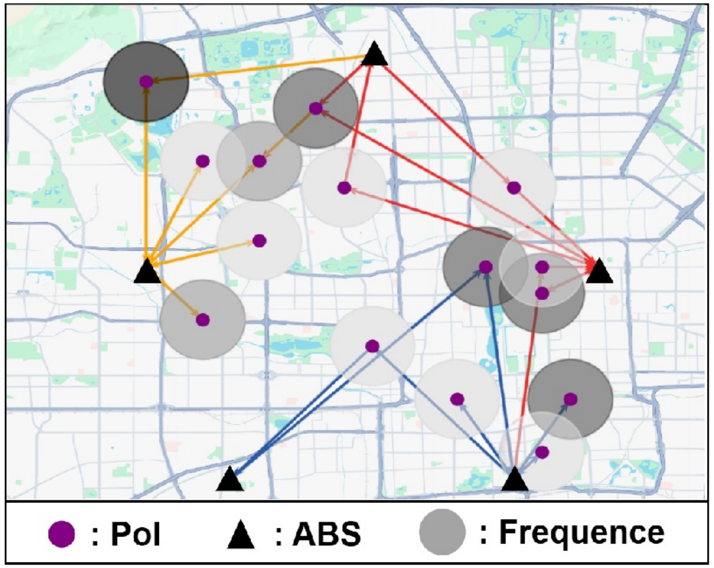
- Our method (OFL-UD^2)
- Random Sampling (RS)
- Importance Sampling (IS)
- Online Data Selection (ODE)

Evaluation Metrics

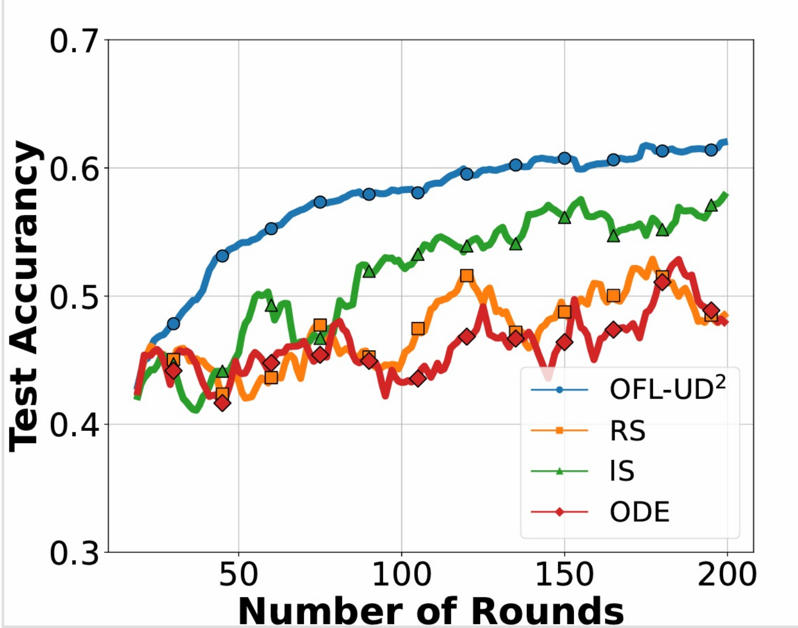
- Test Accuracy
- Cumulative Regret

Evaluation

- Overall Performance on the real-world distributed Outdoor-images dataset:



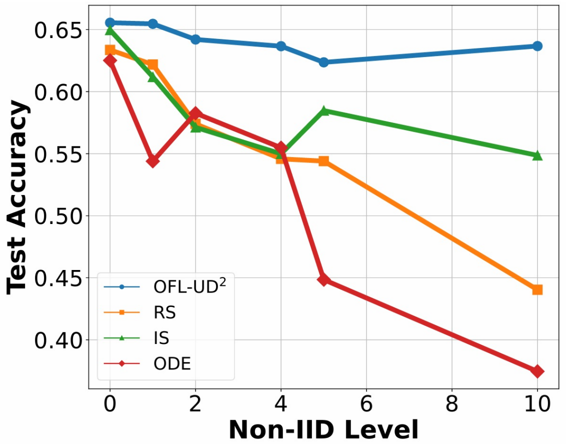
Trajectory

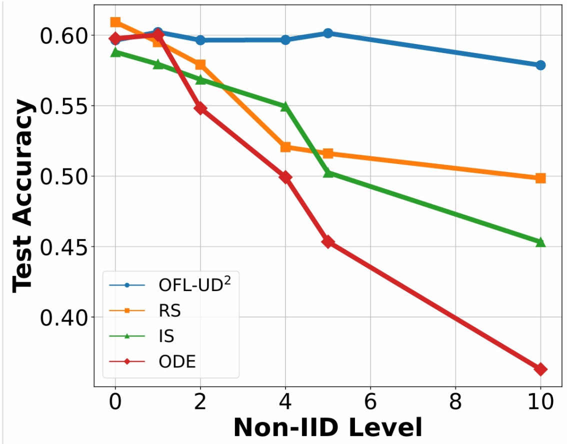


Test Accuracy

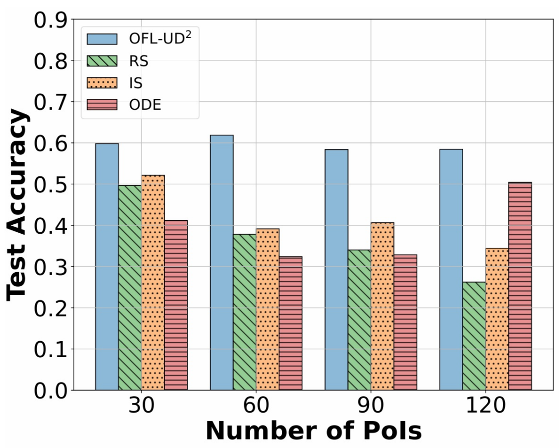
The test accuracy of OFL-UD^2 on the real-world dataset Outdoor-images is significantly higher than other methods.

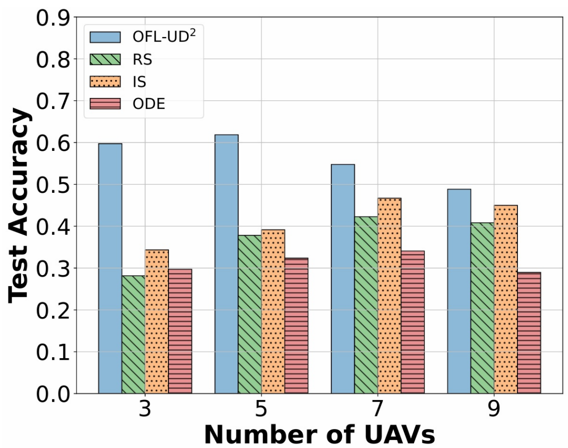
Evaluation





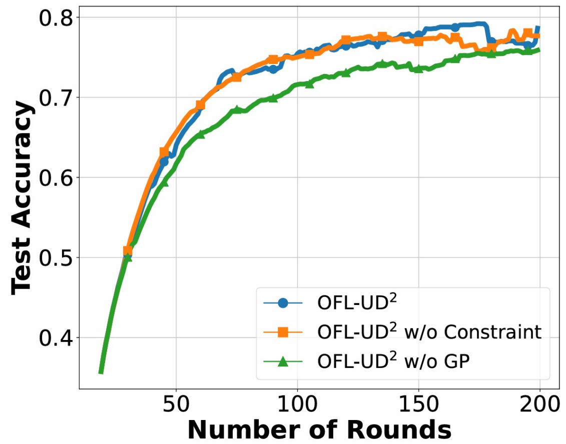
OFL-UD^2 shows good robustness on datasets with different heterogeneity levels.

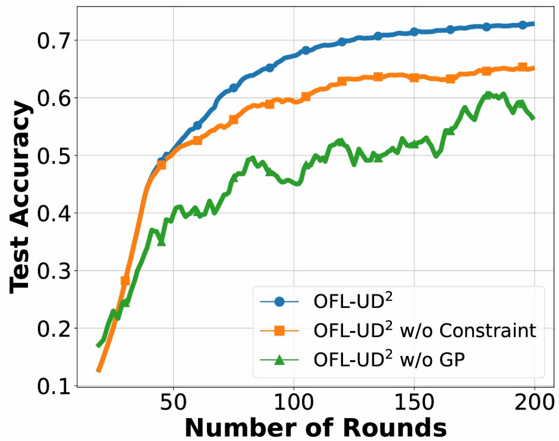




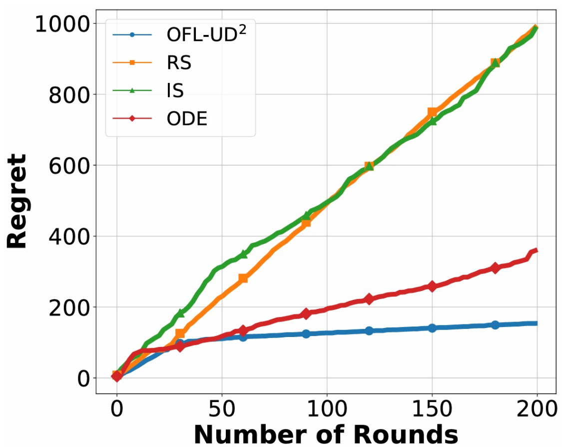
OFL-UD^2 has different performance impacts under different PoI numbers and UAV scales.

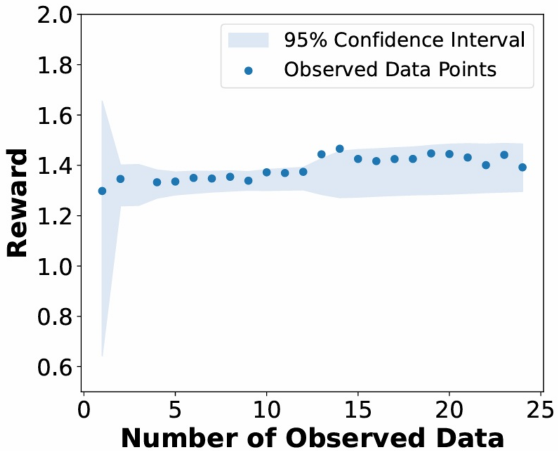
Evaluation





Ablation experiments show that both strategies based on gradient information and Gaussian processes are effective.





OFL-UD^2 can quickly and accurately estimate the parameters.

Conclusion

- Introduce a **novel online data-driven FL framework OFL-UD^2** , in which each round of FL is embedded with an issue of matching between UAVs and PoIs.

- Design a GP-CMAB-based PoI selection strategy, enabling UAVs to be scheduled to optimal PoIs for near-optimal model training performance.

- Prove the **approximate optimality** of **OFL-UD^2** and evaluate the algorithm performance via simulations.

Future work:

- Explore other decisions in UAV ML tasks, such as data selective collection, collection time, etc.
- Investigate on fine-grained integration of **Sequential FL** and **UAV.**

中国科学技术大学
University of Science and Technology of China

41st IEEE International Conference on Data Engineering

—— HONG KONG SAR, CHINA | MAY 19 – 23, 2025 ——

Question?

Xichong Zhang yyu18@mail.ustc.edu.cn