Federated Learning with Privacy-Preserving Incentives for Aerial Computing Networks

Peng Wang, *Student Member, IEEE,* Yi Yang, *Studen Member, IEEE,* Wen Sun, *Senior Member, IEEE,* Qubeijian Wang, *Member, IEEE,* Bin Guo, *Senior Member, IEEE,* Jianhua He, *Senior Member, IEEE,* Yuanguo Bi, *Member, IEEE*

Abstract—With the help of artificial intelligence (AI) model, aerial computing can help analyze and predict the network dynamics and support intelligent decision-making to improve the performance of 6G space-air-ground integrated networks. Federated learning has been proposed to tackle the challenges of limited energy and data shortage for the application of AI models in aerial computing networks. A critical problem of FL for aerial computing is the lack of incentives due to privacy concerns. On the one hand, the information needed to measure users' learning quality may be eavesdropped. On the other hand, users' real costs for determining payments may also undertake inference attacks. In this paper, we design a privacy-preserving and learning quality-aware incentive mechanism for federated learning in aerial computing networks. We propose differential privacy based scheme to protect the privacy of the real cost. In addition, utilize Combinatorial Multi-Armed Bandit (CMAB) algorithm to evaluate the user learning quality without any participant information. Simulation results demonstrate that our scheme can significantly motivate high-quality participants with guaranteed privacy preservation and achieve effective federated learning under the constraint of the limited budget.

Index Terms—federated learning, aerial computing network, privacy-preserving, incentive mechanism

1 Introduction

With the large-scale commercialization of 5G, the number of users and service demands in wireless networks is constantly increasing. To meet the high requirements for coverage, service quality, and communication speed, researchers have started exploring emerging technologies and applications for the future 6G space-air-ground integrated network [1]. Integrated with cloud computing, edge computing, big data, artificial intelligence (AI), transmission rate, end-to-end delay, reliability, spectrum efficiency and energy consumption of network will be significantly improved. The 6G space-air-ground integrated network shows great promise to achieve global coverage, ubiquitous intelligence, and reliable services in a real sense [2].

Aerial computing is considered a key enabler for achieving seamless global coverage, thanks to its convenient mobile access, distributed computing, high flexibility, and scalability [3]. Aerial network consisting of interconnected unmanned aerial vehicles (UAVs) as air domain infrastructure and the cloud with powerful computing power and sufficient resources. The aerial network can provide enhanced communication and computing services for ground networks, particularly in extreme fields such as wartime communication and post-disaster rescue [4]. In addition, with the help of AI model, the network dynamics can be predicted, and intelligent decision-making can be achieved. For example, when the ground base station is destroyed after the disaster, UAVs equipped with AI models can cooperate to provide communication relay services. [5]. However, building AI models in dynamic and complex aerial computing networks faces challenges. On the one hand, UAVs have limited energy and computing resources and cannot complete model training independently. On the other hand, building the AI model on the UAV requires massive data from the network, which faces the risk of privacy leakage during transmission.

Federated learning is believed to construct AI models in an effective and privacy-preserving manner. Yang *et al.* [6] developed an asynchronous federated learning framework for UAVs-enabled networks, and conducted distributed training locally without transmitting sensitive data to UAV servers. Alferaid *et al.* [7] proposed a resource management approach with federated learning in MEC, where clients are assigned with different subnetworks according to the status of their local resources to reach elastic and efficient utilization of energy. Yeom *et al.* [8] proposed an energy-

- *This work was supported in part by the National Natural Science Foundation of China under Grant 62272391, in part by the Key Industry Innovation Chain of Shaanxi under Grant 2021ZDLGY05-08, in part by the Defense Industrial Technology Development Program under Grants JCKY2020204B027, JCKY2021110B143, and JCKY2021608B001, in part by EPSRC with RC Grant reference EP/Y027787/1 and UKRI under grant number EP/Y028317/1, in part by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 824019 and No 101022280, and in part by Horizon Europe MSCA programme under grant agreement No 101086228.*
- *Peng Wang is with the School of Cyber Engineering, Xidian University, Xi'an 710071, China (e-mail: pengw@stu.xidian.edu.cn)*
- *Wen Sun, Qubeijian Wang and Yi Yang are with the School of Cybersecurity, Northwestern Polytechnical University, Xi'an, China (e-mail: yiyang@nwpu.edu.cn, sunwen@nwpu.edu.cn, qubeijian.wang@nwpu.edu.cn).*
- *Bin Guo is with the Department of Computer Science, Northwestern Polytechnical University, Xi'an, China (e-mail: guob@nwpu.edu.cn).*
- *Jianhua He is with the School of Computer Science and Electronic Engineering, University of Essex, UK (e-mail: j.he@essex.ac.uk).*
- *Yuanguo Bi is with the School of Computer Science and Engineering, Northeastern University, Shenyang, China (e-mail: biyuanguo @mail.neu.edu.cn).*
- *Yi Yang is the corresponding author.*

efficient SAGIN based on federated learning, in which IoT devices choose appropriate satellites or UAVs for task offloading. Wei *et al.* [9] proposed a lightweight privacy-preserving federated learning scheme for large-scale IoT devices. The authors presented a reusable masking with the secret-sharing protocol to protect the privacy of individual local data while reducing the computing and communication overhead. Despite the potential benefits, achieving efficient federated learning is hindered by low participation enthusiasm and significant learning quality variety from users.

Efficient incentive mechanisms are indispensable to attract high-quality users to participate in federated learning and build high-performance AI models. Chen *et al.* [10] integrated the reputation-based and blockchain-based incentives, which suggests the utilization of cryptocurrency to promote honest participation in federated learning-based data-sharing. Lee *et al.* [11] proposed a hierarchical federated learning incentive mechanism based on multi-leader Stackelberg games to improve learning efficiency. Kim *et al.* [12] formulated the incentive between the aggregator and clients as an auction game, utilizing primal-dual greedy algorithms to solve NP-hard problems of selecting high-quality users.

Nevertheless, the information needed to measure users' learning quality, such as data set distribution, location, computing power, etc., is privacy-sensitive, and the information can be eavesdropped during transmission. In addition, users' real costs, which are indispensable for determining payments in incentives, may undertake inference attacks, affecting the truthfulness of the incentive mechanism [13]. In this paper, to protect user privacy and improve the federated learning model accuracy, we study the privacy-protecting and quality-ware incentives for federated learning in aerial computing networks. The specific contributions of this paper are summarized as follows.

- We propose a privacy-preserving incentive mechanism with learning quality-awareness called Combinatorial Multi-Armed Bandit-based and differential privacy-protected auction (CamPRA) for federated learning in aerial computing networks. In each round, the UAV as the aggregator adaptively selects high-quality clients for learning task only with users' learning cost information.
- We utilize Combinatorial Multi-Armed Bandit (CMAB) to evaluate the learning quality without any participant information and achieve high performance of the AI model. Moreover, we use differential privacy to protect the privacy of the clients' real costs.
- We theoretically prove the proposed CamPRA satisfies truthfulness, individual rationality, budget balance, and the convergence of CMAB regret. Numerical results show that our CamPRA is superior to others in terms of privacy protection, learning accuracy, and learning cost.

The rest of the paper is organized as follows: Section 2 reviews related work, Section 3 presents the system model, Section 4 introduces our privacy-preserving incentive mechanism for federated learning, Section 5 discusses the numerical results, and Section 6 concludes the paper.

2 Related Work

The emergence of federated learning has enabled users to conduct distributed training and collaborate on building AI models without disclosing raw data [14]. However, users often lose their motivation to be involved in model training due to challenges such as resource consumption, computational overhead, and data privacy. By introducing incentive mechanisms, users can be motivated to participate in federated learning and contribute data and computing resources actively. In recent years, there has been extensive research on incentive mechanisms in federated learning [15] [16] [17]. Heiss *et al.* [18] developed a real-time contribution measurement method which obtained the contribution rate of each participant based on attention aggregation, to provide reasonable rewards to participants with significant contributions. Oktian *et al.* [19] proposed an enhanced Shapey value method for incentive mechanisms in federated learning with multiple influencing factors as weights to measure income distribution. Han *et al.* [20] designed a tokenized incentive mechanism in which Tagging is a means of paying for client. This scheme uses new metrics such as tag reduction rate to measure the contribution rate of clients. However, these incentive mechanisms improve the effectiveness of federated learning by collecting user private information, resulting in privacy disclosure of participants. It is necessary to encourage users to participate in federated learning while protecting privacy.

Researchers have begun to pay attention to privacy protection issues of incentive mechanisms in untrusted environments. Liu *et al.* [21] formulated the computing resource problem in federated learning as a tackelberg game. Gonccalves *et al.* [22] proposed an incentive framework for federated learning based on differentially private and 3-D contract approach. Fantacci *et al.* [23] established a noncooperative-game-enabled incentive mechanism, and Xiong *et al.* [24] utilize blockchain to desigen a value-driven incentive mechanism to guarantee data privacy and provide auditability for the whole training process. However, these works focus on protecting users' private information in incentives, while ignoring the measurement of users' learning quality to improve the learning performance in federated learning.

To select appropriate users for federated learning, incentive mechanisms based on learning quality measurement have been extensively studied. Hu *et al.* [25] deduced the optimal strategy for servers and users by solving Stackelberg equilibrium, and Lee *et al.* [11] proposed a hierarchical federated learning incentive mechanism based on multi-leader Stackelberg games. Additional technologies, including auctionmechanisms, reputation mechanisms, and contract theory, have also been used in quality-aware incentive mechanisms [26] [27] [28]. For example, Le *et al.* [29] formulated the incentive between the aggregator and clients as an auction game, utilizing primal dual greedy algorithms to solve NP-hard problems of selecting winners. Moudoud *et al.* [30] designed a reverse auction-based incentive method, perceiving user quality and selecting users through comprehensive reputation and bidding price selection. However, these methods did not consider the privacy protection and quality issues of model updates. Challenges still exist in

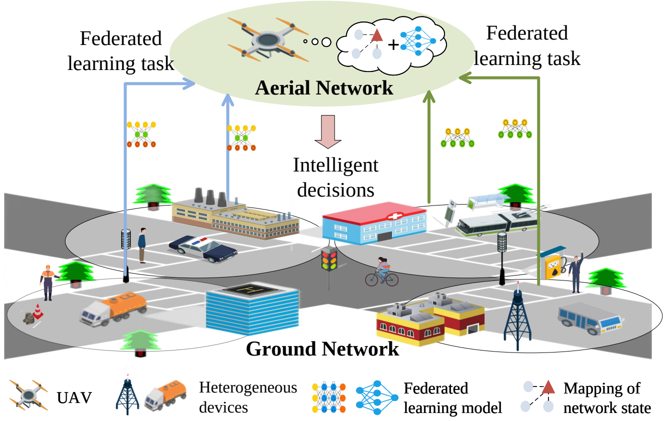


Fig. 1: Federated learning enabled aerial computing networks.

measuring user quality through privacy-preserving methods while designing incentive mechanisms that are compatible with federated learning.

It is noted that most of the existing works have not focused on balancing privacy protection and learning performance in incentives. In this work, we propose a privacy-preserving incentive mechanism with quality awareness for federated learning in the dynamic and complex aerial computing networks, which protects users' incentive-related information and improves the learning efficiency.

3 System Model

Fig. 1 depicts the federated learning-enable aerial computing network architecture. The aerial computing network consists of the terrestrial network with massive heterogeneous devices and the aerial network with interconnected UAVs. The UAV in aerial network serve as aerial infrastructure, and provide supplementary connectivity and reliable data transmission for the terrestrial network extreme situations or service congestion. Moreover, with the assistance of the AI model, UAVs can monitor and analyze the network status within their coverage in real-time [31], and support intelligent decision making to improve network performance, such as channel resource allocation, cooperative communication relay, etc. To alleviate the computing burden and privacy risk, the UAV uses federated learning to train the AI models. Ground devices with sufficient data execute the federated learning task in parallel without exposing their private data. The trained local model finally aggregated on the UAV. In this way, the AI model can integrate the global knowledge of the network in an energy-efficient and privacy-preserving way, and its generality is greatly improved.

3.1 Federated Learning Process

We use $\mathcal{N} = \{1, 2, ..., N\}$ to denote the set of ground devices. In the t-th round of global iteration in federated learning, assuming $\Phi(t)$ is the client set which are chosen from $\mathcal{N}$ for federated learning task. First, the UAV as aggregator publishes its initialized model $\omega(t-1)$ and distribute it to each client n in $\Phi(t)$. Then each client receives the initialized global model $\omega(t-1)$ and updates the parameters of its local model $\omega_n(t)$ using its private data set D_n. During this process, each client trains its local model (which is called local update) to find the optimal parameters that minimize the loss function $L(\omega_n(t))$.

The loss function for each client n can be defined as follows:

$$L(\omega_n(t)) = \frac{1}{|D_n|} \sum_{d_i \in D_n} l_i(\omega_n(t), d_i), \tag{1}$$

where the $l_i(\omega_n(t), d_i)$ is the loss function of data sample $d_i \in D_n$. Each client performs local model training to minimize the loss function of the AI model, and then updates its local model parameters $\omega_n(t)$. The update rule for client n is as follows:

$$\omega_n(t) = \omega(t-1) - \eta \bigtriangledown L(\omega_n(t-1)), \tag{2}$$

where $\eta > 0$ is the learning step and $\bigtriangledown$ is the gradient of the loss function. After executing the local update, all clients upload their local models ω_n^t to the aggregator. The aggregator performs aggregation of the weighted model parameters and obtains the global model $\omega(t)$, which is called global update. The global update complies with the following rule:

$$\omega(t) = \frac{1}{\sum_{n \in \Phi(t)} |D_n|} \sum_{n \in \mathcal{N}(t)} |D_n| \omega_n(t). \tag{3}$$

The above processes will iterate until the global loss function $L(\omega)$ converges.

$$L(\omega) = \frac{1}{\sum_{n \in \Phi(t)} |D_n|} \sum_{n \in \Phi(t)} |D_n| L(\omega_n). \tag{4}$$

TABLE 1: Notation Setting

Notation	Meaning
$\omega_n(t)$	The local model of client n at time slot t
$c_n(t)$	The learning cost of client n at time slot t
$c_n^{'}(t)$	The perturbed cost by differential privacy
$\rho_n(t)$	The update significance of client n at time slot t
$\tau_n(t)$	The update latency of client n at time slot t
$q_n(t)$	The learning quality of client n at time slot t
$\hat{q}_n(t)$	The estimated learning quality using UCB
$\Phi_n(t)$	The client selection result at time slot t
$p_n(t)$	The payment for client n at time slot t
$B(t)$	The budget used to recruit clients
$\mu_n(t)$	UCB index
$\beta_n(t)$	The unit cost of learning quality
K	The number of selected clients in each round
ϵ	The differential privacy budget

3.2 Learning Cost of Clients

To execute the federated learning task, each client needs to collect data, train the local model iteratively with their computing resource, and upload the local model to the aggregator consuming communication resource, which inevitably incurs data cost, computing cost, and communication cost.

- **Data cost.** Data cost involves the cost of collection, storage, maintenance, and preprocessing for model training. Assuming client n has a unit data cost ζ_n^{data} for task π, the total data cost of n for task π is expressed as

$$c_n^{data} = \zeta_n^{data}|D_n|. \tag{5}$$

- **Computation cost.** The computational cost mainly comes from the consumption of local model updating per iteration. We use f_n to represent the computation capacity of the client (i.e.,the CPU frequency), The CPU cycles required to train one dataset sample locally is denoted by I_n^{local}. The time consumption of local model update is denoted by

$$T_n^{comp} = \frac{c_n|\mathcal{D}_n|}{f_n}. \tag{6}$$

The client n's computation cost of executing one round of local model training can be expressed as

$$c_n^{comp} = \zeta_n^{comp}\delta_n T_n^{comp} f_n^3 = \delta_n \zeta_n^{comp}|D_n| f_n^2, \tag{7}$$

where ζ_n^{comp} is the client n's computation cost of unit CPU frequency, and the δ_n is client n's local iteration rounds.

- **Communication cost.** In this paper, we mainly consider the communication cost of local model uploading. The transmission time of local model updating is

$$T_n^{com} = \frac{|\boldsymbol{\omega}_n|}{r_n}, \tag{8}$$

where r_n is the model upload rate and $|\omega_n|$ is the size of the local model. The communication cost can be expressed as

$$c_n^{com} = W_n T_n^{com} \zeta_n^{com}, \tag{9}$$

where $|\omega_n|$ is the local model size, and ζ_n^{com} is the unit cost for communication.

Hence the total learning cost c_n of client n can be calculated by

$$c_n = c_n^{data} + c_n^{comp} + c_n^{com}. \tag{10}$$

3.3 Learning Quality Measurement

To motivate the client with high learning quality for effective federated learning, a learning quality measurement mechanism is indispensable. However, comprehensively measuring learning quality requires multi-dimensional information such as data set size, data distribution, computation and communication ability, etc. Additionally, privacy concerns make it difficult to obtain these relevant information. In this paper, we use two hindsight observation-based metrics to measure the learning quality, i.e., update significance and learning latency.

- **Update significance.** Update significance has been widely studied to measure the client's contribution to the global model in federated learning. In this paper, the update significance is calculated by the contribution of local models to global model loss reduction. Supposing that for each iteration t, its start time is t^{start} and end time is t^{end}, then we can explain the update significance of client n for as

$$\rho_n(t) = L(t^{start}) - L_n(t^{end}), \tag{11}$$

where $L(t^{start})$ is the loss of global model ωat t^{start}, and $L_n(t^{end})$ is the client n's local model loss at $L_n(t^{end})$.

- **learning latency.** Due to the strict requirements of federated learning on update delay, the learning latency cannot be ignored. The time consumption of each iteration mainly consists of three components: initial model distribution time, local model training time, and local model uploading time, denoted by $\tau_n^{dis}(t)$, $\tau_n^{upd}(t)$ and $\tau_n^{upl}(t)$, respectively. Then, the total time consumed by client n for global learning in round t can be given as

$$\tau_n(t) = \min\{\tau_n^{dis}(t) + \tau_n^{upd}(t) + \tau_n^{upl}(t), \tau_{\max}\}, \tag{12}$$

where $\tau_{\max}$ is the maximum waiting time that the UAV can tolerate. Although the local update time can be calculated with Equation (8), the other two components are still difficult to obtain due to the heterogeneity of devices and dynamics of the network environment. In this paper, the UAV as aggregator observes $\tau_n(t)$ directly thus reducing difficulties and computation overhead.

Therefore, qualities of clients can be defined as the weighted average of learning quality $\rho_n(t)$ and training latency $\tau_n(t)$:

$$q_n(t) = \lambda \cdot \rho_n(t) + \frac{\sigma}{\tau_n(t)} \tag{13}$$

where λ and σ are weighted parameters and $\lambda + \sigma = 1$.

4 Privacy-Preserving Incentive for Federated Learning

4.1 Problem Formulation

In this paper, we aim to design an incentive scheme with privacy protection and quality awareness to maximize the federated learning performance in aerial computing networks. An appropriate clients set adaptive to the time-varying network environment is dynamically selected before each round of learning t. The participant selection problem can be defined as:

$$\begin{array}{lll}
\textbf{P1:} & \underset{\boldsymbol{\Phi}}{maximize} & \sum_{t=1}^{T} \mathcal{G}(\Phi(t), q_n(t)), \\
\text{s.t.} & x_n(t) \in \{0,1\}, \forall n \in \mathcal{N} \quad (C1) & \\
& p_n(t) \geq c_n(t), \forall n \in \Phi(t), \quad (C2) & (14) \\
& \sum_{n \in \Phi(t)} p_n(t) \leq B(t) \quad (C3), & (15)
\end{array}$$

where $\Phi = \{\Phi(1), \Phi(2), ..., \Phi(t), ..., \Phi(T)\}$ is the client selection results, T is the termination of the iteration rounds, $\mathcal{G}(\cdot)$ is the model aggregation function. The constraint $C1$ is used to indicate whether client n is selected by task $\pi(t)$. $C2$ limits the payment to clients should be higher than its learning cost. $C3$ is the budget constraint that the total payments should not exceed the budget $B(t)$.

The incentive mechanism is required to select optimal clients effectively while ensuring the following desirabilities.

- **Privacy-preserving.** The incentive mechanism must comprehensively protect users' private information, including learning quality-related information and real learning costs.
- **Individual rationality.** The incentive mechanism must guarantee that the payment received by each user is not less than its real cost. In other words, each user n should earn a non-negative profit when participating in the inventive mechanism.
- **Budget balance.** For the UAV, the total payments for to the selected clients should not exceed its budget per round.
- **Truthfulness.** The incentive scheme must guarantee that no user can improve its profit by strategically submitting a cost to the UAV.

4.2 CMAB Modeling

The formulated incentive problem **P1** is difficult to mathematically externalize in terms of model aggregation functions due to the lack of appropriate metrics to quantify the quality of individual model updates and global aggregated models. In the following, we introduce CMAB to deal with **P1**.

A CMAB problem is an online reinforcement learning-based sequential decision problem. The player should make decisions on which arms of the bandit to pull in each round t, where pulling an arm is called an action. It is noted that taking any action will incur a cost and the player has limited budget for arm selection. To solve the CMAB problem, the player has to continuously learn which set of arms can bring high reward according to the results of each round, and make sequential decisions to maximize the total expected reward, i.e., $\mathbb{E}(Q(x)) = \sum_{t=1}^{T} x_n(t)q_n(t)$ under the budget constraint, where $q_n(t) \in [0, 1]$ is the observed reward. **P1** can be modeled as a CMAB problem, where the UAV and the clients are regarded as the player and the arms, respectively. We formalize the federated learning task participants selection problem by maximizing the total expected rewards $\mathbb{E}(Q(x))$ obtained by the UAV:

$$
\begin{aligned}
\textbf{P2:} \quad & maximize \quad \sum_{t=1}^{T} x_n(t)q_n(t) \\
\text{s.t.} \quad & x_n(t) \in \{0, 1\}, \forall n \in \mathcal{N}, \quad (C1) \\
& p_n(t) \geq c_n(t), \forall n \in \Phi(t), \quad (C2) \qquad (16) \\
& \sum_{n \in \Phi(t)} p_n(t) \leq B(t) \quad (C3), \qquad (17)
\end{aligned}
$$

Noting that **P2** can be considered as a knapsack problem, which has been proven to be NP-hard. It is impossible to find a deterministic algorithm to address the problem in polynomial time. Auctions are widely recognized as an effective method to find approximate-optimal solutions for the knapsack problem.

Definition 1. *((α, β)-approximation solution) For $\alpha, \beta \leq 1$, (α, β)-approximation solution takes an expectation vector $q_1, 1_2, ...q_N$ as input, and outputs a set of users(or arms) $A \in \mathcal{N}$. such that $\mathbb{P}[Q(A) \geq \alpha \cdot Q^{opt}] \geq \beta$, where Q^{opt} is the optimal reward, and β is the success probability of the solution.*

When using an (α, β)-approximation algorithm, it is inappropriate to compare the performance of a CMAB solution against the optimal reward as the regret. Instead, we should compare the algorithm with the (α, β) fraction of the optimal reward.

Definition 2. *((α, β)-approximation regret) (α, β)-approximation regret of a CMAB algorithm after T rounds of running that outputs the (α, β)-approximation solution is defined as*

$$
Reg_{\alpha,\beta}(T) = T \cdot \alpha \cdot \beta \cdot Q^{opt} - \mathbb{E}[\sum_{t=1}^{T} Q(t)]. \qquad (18)
$$

4.3 CMAB-based Privacy-Preserving and Quality-Aware Incentive Framework

To protect clients' private information while motivating high-quality clients for the federated learning, we propose a privacy-preserving and quality-aware incentive mechanism called Combinatorial Multi-Armed Bandit-based and differential privacy-protected auction (CamPRA). Figure. 2 illustrates the CamPRA framework where the UAV acts as the buyer and the clients act as sellers. clients do not need to expose any information besides their real cost as the bid. Specifically, the aggregator learns the qualities of all its capable clients by letting them execute the task and observe the obtained reward so that the total expected reward is as close to the optimal reward as possible.

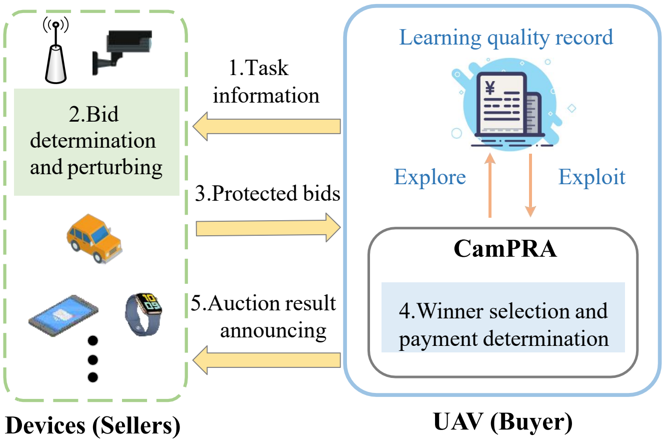

Learning quality record
1.Task information
2.Bid determination and perturbing
3.Protected bids
Explore
Exploit
CamPRA
5.Auction result announcing
4.Winner selection and payment determination
Devices (Sellers)
UAV (Buyer)

Fig. 2: CamPRA framework.

4.3.1 UCB-based Adaptive Learning Quality Measurement

In the CMAB model, as the available information about arms (which means the quality-related information) is limited, the player has to handle the trade-off between exploitation and exploration. The exploitation refers to selecting arms that performed well in the past rounds of federated learning. The exploration means selecting arms that might bring higher rewards in the future. Moreover, the available budget for selecting arms is also limited, posing extra economic challenges to balancing exploitation and exploration. In most CMAB methods [32], the exploration phase and the exploitation phase are separated. During the exploration phase, the player takes turns selecting a set of arms and observes the rewards until reaching the preset exploration times or exhausting the exploration budget. Then in the exploitation phase, the player selects the optimal set of arms based on the exploration results.

However, if low-quality clients are frequently selected during the exploration phase, the performance of the global model will be affected. In the proposed CamPRA, we do not separate the exploitation and exploration phases, so as to avoid the slow improvement of the global model when exploring client quality. Specifically, we design a UCB-based adaptive learning quality estimation scheme. Assume client n has been selected $s_n(t)$ times to perform learning task π in the first t rounds. The update rule of $s_n(t)$ can be defined as

$$s_n(t) = \begin{cases} s_n(t-1)+1, & x_n(t)=1; \\ s_n(t-1), & x_n(t)=0. \end{cases} \tag{19}$$

We use $\overline{q}_n(t) \in [0,1]$ to denote the sample average learning quality of client n until round t. The update rule of $\overline{q}_n(t)$ can be expressed by

$$\overline{q}_n(t) = \begin{cases} \dfrac{\overline{q}_n(t-1)s_n(t-1)+q_n(t)}{s_n(t-1)+1}, & x_n(t)=1; \\ \overline{q}_n(t-1), & x_n(t)=0, \end{cases} \tag{20}$$

where the $q_n(t)$ is the quality learned in the current round and is calculated according to (5). Hence the client n's learning quality $\hat{q}_n(t)$ can be estimated based on UCB interval,

$$\hat{q}_n(t) = min\{\overline{q}_n(t) + \mu_n(t), 1\}, \tag{21}$$

$$\mu_n(t) = \sqrt{\frac{3\text{ln}t}{2s_n(t)}}, \tag{22}$$

where $\mu_n(t)$ is the UCB index to address the dynamics of learning quality.

4.3.2 Real Cost Protection

To achieve a truthful auction mechanism, each data owner should submit its true cost as the bid. However, the true cost information may be subject to inference attacks from malicious users. In this paper, aiming at safeguarding the privacy of bids, we employ the exponential mechanism based on differential privacy (DP) to perturb the bid. We suppose that for the bid distribution space C, the UAV possesses its prior knowledge. The exponential mechanism probabilistically maps $c_n(t)$ to $c_n^{'}(t) \in C$ for any true bid $c_n(t) \in C$, with a probability of

$$\mathbb{P}(c_n^{'}(t)|c_n(t)) \propto \exp(\frac{\epsilon g(c_n(t), c_n^{'}(t))}{2\Delta g}), \tag{23}$$

Here, ϵ represents the privacy budget allocated for perturbing bids, and $g(c_n(t), c_n^{'}(t))$ is a function used to measure the distance between $c_n(t)$ and $c_n^{'}(t)$. The sensitivity of the function g is denoted by Δq. More importantly, the UAV determines the bid perturbation function, which is then employed by every client with the same budget ϵ to promote the subsequent process of winner selection. As $c_n(t)$ and $c_n^{'}(t)$ become closer, the probability $\mathbb{P}(c_n^{'}(t)|c_n(t))$ increases. This implies that $g(c_n^{'}(t)|c_n(t))$ is a monotonically decreasing function about $|c_n(t) - c_n^{'}(t)|$. As a result, we define $g(c_n^{'}(t)|c_n(t))$ as

$$g(c_n(t), c_n^{'}(t)) = -|c_n(t) - c_n^{'}(t)|^{1/2}. \tag{24}$$

Therefore, the sensitivity is represented by $\Delta g = \max|g - g'| = (c_n(t)_{max} - c_n(t)_{min})^{1/2} = \Delta c_n^{1/2}(t)$, where $\Delta c_n(t) = c_n(t)_{max} - c_n(t)_{min}$ is the range of C. For any $c_n^{'}(t) \in C$, the aggregator can calculate the mapping probability $\mathbb{P}(c_n^{'}(t)|c_n(t))$ with the following equation

$$\mathbb{P}(c_n^{'}(t)|c_n(t)) = \frac{\exp(\epsilon\frac{-|c_n(t)-c_n^{'}(t)|^{1/2}}{2\Delta c_n^{1/2}(t)})}{\sum_{c_n^*(t)\in C}\exp\left(\epsilon\frac{-|c_n(t)-c_n^*(t)|^{1/2}}{2\Delta c_n^{1/2}(t)}\right)}. \tag{25}$$

Then a random bid $c_n^{'}(t) \in C$ is chose as the perturbed bid with probability $\mathbb{P}(c_n^{'}(t)|c_n(t))$ and is submitted to the aggregator.

4.3.3 Winner Selection and Payment Determination

After receiving the perturbed bids from the participants $\mathcal{N}(t)$, the aggregator needs to select a set of clients with high learning qualities in a cost-effective manner due to the budget limitation. In the traditional auction mechanism, the aggregator can make optimal selection decisions based on the client's real bid. However, the aggregator cannot know the real cost $c_n(t)$ from $c_n^{'}(t)$. For this, for any client n, we use the expected bid $\mathbb{E}(c_n(t)|c_n^{'}(t))$ to approximately estimate the true bid $c_n(t)$.

$$\begin{aligned} \mathbb{E}[c_n(t)|c_n^{'}(t)] &= \sum_{c_n(t)\in C}\mathbb{P}(c_n(t)|c_n^{'}(t))c_n(t) \\ &= \frac{\sum_{c_n(t)\in C}\mathbb{P}(c_n^{'}(t)|c_n(t))P(c_n(t))c_n(t)}{\sum_{c_n(t)\in C}\mathbb{P}(c_n^{'}(t)|c_n(t))P(c_n(t))}, \end{aligned} \tag{26}$$

where $\mathbb{P}(c_n(t)) = f(c_n(t))/|C|$ is the proportion of bid $c_n(t)$ in C, and $f(c_n(t))$ is the frequency of bid $c_n(t)$.

Then the aggregator selects K clients as the winners according to the expected bid $\mathbb{E}[c_n(t)|c_n^{'}(t)]$ and the current learned learning quality $\hat{q}_n(t)$. The value of K is determined by the budget $B(t)$ and the $\beta_n(t) = \hat{q}_n(t-1)/\mathbb{E}[c_n(t)|c_n^{'}(t)]$ (which means the unit cost of learning quality). Specifically, the aggregator sorts the received candidate set $\mathcal{N}(t)$ in descending order according to the value of $\beta_n(t)$ and gets a sorted candidate set:

$$\mathbb{N}(t) = \{n_{l_1}(t), n_{l_2}(t), ..., n_{l_{|\mathbb{N}|}}(t)\}, \tag{27}$$

where

$$\beta_{l_1}(t) \geq \beta_{l_2}(t) \geq ... \geq \beta(t)_{l_{|\mathbb{N}|}}. \tag{28}$$

And then the client chooses the first K clients in $\mathbb{N}(t)$ to the winner set $\Phi(t)$ under the budget constraint $B(t)$, i.e.,

$$K = \begin{cases} \arg\max\limits_{K}\{K|\dfrac{B(t)}{\sum_{l_i=1}^{K}\beta_{l_i}(t)}, \forall n(t)_{l_i} \in \mathbb{N}(t)\} & \text{if } \mathbb{N}(t) \neq \emptyset; \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

To ensure the truthfulness of CamPRA, according to Myerson's Theorem, we design a critical value-based payment determination scheme. Each winner n should be paid the critical value p_n^*. Any bid must win the auction if its value $c_n \leq p_n^*$, otherwise it won't win.

Then the payment for each winner in $\Phi(t)$ is determined by the following rules:

$$p_n(t) = \min\left\{\frac{B(t)}{\sum_{i\in\Phi(t)}\beta_i(t)}\beta_n(t), \frac{q_n(t)}{\beta_{l_{k+1}}(t)}\right\}. \tag{30}$$

4.3.4 The Detailed Algorithm

In this part, the details of CamPRA algorithm are proposed, as shown in Algorithm 1.

- **Task publication.** To carry out the $t-$th round of federated learning global iteration, the client publishes a task π^t. For task $\pi(t)$, its task information is $I(t) = \{l(t), B(t)\}$, where $l(t)$ is types of data set required by $\pi(t)$, and $B(t)$ is the budget of $\pi(t)$.
- **Bid determination and perturbing.** After receiving the task information, each client n can calculate its learning cost in the current round $c_n(t)$ according to the task information and its own energy state. Then it perturb its real bid using Equation (25), then submits its perturbed bid $c_n^{'}(t)$ to the aggregator.
- **learning quality initialization.** In the initial phase, i.e., $t = 1$, the aggregator set the initial quality $\hat{q}_n(0)$ of each client n to be 1.
- **Winner selection.** In the $t-$th round of CamPRA, after receiving bids from clients, the aggregator needs to select clients for effective federated learning according to the qualities learned by the UCB model and the bids. The aggregator first calculate the expected bids $\mathbb{E}[c_n(t)|c_n^{'}(t)]$ for each $c_n^{'}(t)$ according to Equation (26). Then the aggregator selects K clients for the task according to Equation (29).
- **Payment determination.** The aggregator calculates the critical payment using Equation (30).
- **learning quality updating.** At last, $\hat{q}_n^t$ is updated according to the global aggregation result.

Those processes continue until the model is successfully constructed and the federated learning terminates.

4.4 Theoretical Analysis

In this section, we theoretically analyze the desired properties of CamPRA, including γ-truthfulness, budget balance, individual rationality, privacy preserving, and regret upper bound of CMAB.

4.4.1 Truthfulness

We first define utility functions to quantify the revenues of clients and the aggregator in the auction. In the t-th round, the utility of each client n can be denoted by

$$U_n^{client}(t) = x_n(t)(p_n(t) - c_n(t)). \tag{31}$$

The utility of the aggregator can be denoted by

$$U^{aggregator}(t) = \sum_{n\in\Phi(t)} Q_n(t) - p_n(t), \tag{32}$$

where $Q_n(t)$ is the revenue earned in the current round of federated learning, and is function of the learning quality $q_n(t)$.

```
Algorithm 1: CamPRA
  Input: Sellers set $\mathcal{N}$, task $\pi$, budget of the task $B$,
         differential privacy budget $\epsilon$.
  Output: Winner set $\Phi$, Payment $P$
  $t = 0$;
  learning quality initialization: $\forall n \in \mathcal{N},\ \overline{q}_n(t) = 0$;
  while true do
      $t = t + 1$;
      client:
      foreach $n \in \mathcal{N}^t$ do
          Calculate bid: $b_n(t)$;
          Perturb bid $b_n^{'}(t)$ with Equation (25);
          Submit $b_n^{'}(t)$to the aggregator;
      end
      aggregator:
      if $|\mathcal{N}(t)| = 0$ then
          No seller will be selected;
          $\Phi(t) = \emptyset, P(t) = \emptyset$;
          Break;
      end
      foreach $n \in \mathcal{N}(t)$ do
          Calculate the expected bid $\mathbb{E}(c_n(t)|c_n^{'}(t))$ by
            Equation (26)
          $\beta_n(t) = \hat{q}_n(t-1)/c_n(t)$;
          $\mathbb{N}(t) = \{n_{l_1}(t), n_{l_2}(t), ..., n_{l_{|\mathbb{N}(t)|}}(t)\}$, where
            $\beta_{l_1}(t) \geq \beta_{l_2}(t) \geq ... \geq \beta(t)_{l_{|\mathbb{N}(t)|}}$;
          Calculate $K$ according to Equation (29);
          $\Phi(t) = \Phi(t) \cup \{n_{l_1}(t), n_{l_2}(t), ..., n_{l_K}(t)\}$;
          Calculate each payment $p_n(t)$ according to
            Equation (30);
          $P(t) = P(t) \cup p_n(t)$
          Update $s_n(t)$, $\overline{q}_n(t)$, $\hat{q}_n(t)$ according to
            Equation (19), (20), (21);
      end
  end
```

Definition 3. *(γ-truthfulness) An incentive mechanism is γ-truthful if for a bid $c_n(t) \neq c_n^{'}(t)$, and any bid of other clients $c_{-n}(t)$, there is:*

$$\mathbb{E}[U^{client}(c_n^{'}(t), c_{-n}(t))] \geq \mathbb{E}[U^{client}(c_n(t), c_{-n}(t))] - \gamma \tag{33}$$

Theorem 1. *The CamPRA is γ-truthful.*

Proof. We use $c_1(t)$ and $c_2(t)$ to denote the two different bids from client 1 and client 2. According to the obfuscation function, we can get $\mathbb{P}(c^{'}(t)|c_1(t)) \leq exp(\epsilon)\mathbb{P}(c^{'}(t)|c_2(t))$. Thus, the expectation of utility of the client satisfies:

$$\begin{aligned}\mathbb{E}[U^{client}(c_2(t))] &= \sum_{c^{'}\in C} U^{client}(c^{'})\mathbb{P}(c^{'}|c_2(t)) \\ &\geq \sum_{c^{'}\in C} U^{client}(c^{'})exp(-\epsilon)\mathbb{P}(c^{'}|c_1(t)) \\ &= exp(-\epsilon)\mathbb{E}[U^{client}(c_1(t))] \\ &\geq (1-\epsilon)\mathbb{E}[U^{client}(c_1(t))] \\ &= \mathbb{E}[U^{client}(c_1(t))] - \epsilon\mathbb{E}[U^{client}(c_1(t))].\end{aligned} \tag{34}$$

With the payment determination strategy of CamPRA, we have

$$\begin{aligned} U^{client}(t) &= p_n(t) - c_n(t) \\ &= \min\{\frac{B^t}{\sum_{i\in\Phi(t)}\beta_i(t)}\beta_n(t)\frac{q_n(t)}{\beta_{l_{k+1}}(t)} \\ &\leq \beta_{max}(t) - \beta_{min}(t) + \beta_{l_{k+1}}(t) \\ &\leq \Delta c. \end{aligned} \tag{35}$$

We have $\mathbb{E}[U^{client}(t)] \leq \Delta c$, Hence $\mathbb{E}[U^{client}(c_2(t))] \geq \mathbb{E}[U^{client}(c_1(t))] - \epsilon\Delta c$. According to the definition of γ-truthfulness, we can prove that the CamPRA satisfies $\epsilon\Delta$-truthfulness.

4.4.2 Individual Rationality

Theorem 2. *The CamPRA satisfies individual rationality.*

Proof. ***Case 1.*** In CamPRA, if client n is not selected by the client, its utility $U_n^{aer} = 0$.

Case 2. If client n wins the auction, it will receive a payment $p_n(t)$.

1) If $p_n(t) = \frac{B(t)}{\sum_{i\in\Phi(t)}\beta_i(t)}\beta_n(t)$, the utility of client n is

$$\begin{aligned} U_n^{aer}(t) &= p_n(t) - c_n(t) \\ &= \frac{B(t)}{\sum_{i\in\Phi(t)}\beta_i(t)}\beta_n(t) - \frac{q_n(t)}{\beta_n(t)} \\ &= q_n(t)\left(\frac{B(t)}{\sum_{i\in\Phi(t)}\beta_i(t)} - \frac{1}{\beta_n(t)}\right) \\ &\geq q_n(t)\left(\frac{B(t)}{\sum_{i\in\Phi(t)}\beta_i(t)} - \frac{1}{\beta_{l_k}(t)}\right) \geq 0. \end{aligned} \tag{36}$$

2) If $p_n(t) = \frac{q_n(t)}{\beta_{l_{k+1}}(t)}$, the utility of client n is

$$U_n^{aer}(t) = p_n(t) - c_n(t) = \frac{q_n(t)}{\beta_{l_{k+1}}(t)} - \frac{q_n(t)}{\beta_n(t)} \geq 0. \tag{37}$$

In any case, the client's utility $U_n^{aer}(t) \geq 0$. Thus the individual rationality of each client is proved and **Theorem 3** holds.

4.4.3 Budget Balance

Theorem 3. *The CamPRA satisfies budget balance.*

Proof. In each round t, the client will finally pay $P(t)$ to the winners totally, as our CamPRA strictly follows the budget constraint, we have

$$\sum_{n\in\Phi(t)} p_n(t) \leq B(t) \tag{38}$$

Hence the budget balance of each aggregator is proved and the **Theorem 4** holds.

4.4.4 Privacy Preserving

Theorem 4. *Our proposed CamPRA satisfies satisfies ϵ-differential privacy for any $\epsilon > 0$.*

Proof. Assuming c_1 andc_2 are two different bids from client 1 and client 2, respectively (for simplicity, we omit the time index t). The probabilities of obfuscating c_1 andc_2 to $c^{'}$ are denoted by $\mathbb{P}(c^{'}|c_1)$ and $\mathbb{P}(c^{'}|c_2)$. The relative probability of CamPRA for given b1 and b2 can be defined as follows:

$$\begin{aligned} \frac{\mathbb{P}(c^{'}|c_1)}{\mathbb{P}(c^{'}|c_2)} &= \frac{\dfrac{\exp(-\epsilon\frac{|c_1 - c^{'}|^{1/2}}{2\Delta c^{1/2}})}{\sum_{c^*\in C}\exp(-\epsilon\frac{|c_1 - c^*|^{1/2}}{2\Delta c^{1/2}})}}{\dfrac{\exp(-\epsilon\frac{|c_2 - c^{'}|^{1/2}}{2\Delta c^{1/2}})}{\sum_{c^*\in C}\exp(-\epsilon\frac{|c_2 - c^*|^{1/2}}{2\Delta c^{1/2}})}} \\ &= \frac{\exp(-\epsilon\frac{|c_1 - c^{'}|^{1/2}}{2\Delta c^{1/2}})}{\sum_{c^*\in C}\exp(-\epsilon\frac{|c_2 - c^{'}|^{1/2}}{2\Delta c^{1/2}})} \cdot \frac{\exp(-\epsilon\frac{|c_2 - c^*|^{1/2}}{2\Delta c^{1/2}})}{\sum_{c^*\in C}\exp(-\epsilon\frac{|c_1 - c^*|^{1/2}}{2\Delta c^{1/2}})} \\ &\leq \exp(\frac{\epsilon\Delta c^{1/2}}{2\Delta c^{1/2}}) \cdot \frac{\sum_{c^*\in C}\exp(-\epsilon\frac{|c_1 - c^{'}|^{1/2} + \Delta c^{1/2}}{2\Delta c^{1/2}})}{\sum_{c^*\in C}\exp(-\epsilon\frac{|c_1 - c^*|^{1/2}}{2\Delta c^{1/2}})} \\ &= (\exp(\frac{\epsilon}{2}))^2 \end{aligned} \tag{39}$$

Thus we have

$$\frac{\mathbb{P}(c^{'}|c_1)}{\mathbb{P}(c^{'}|c_2)} = X \cdot Y \leq \exp(\epsilon) \tag{40}$$

Till now, we have proved our CamPRA satisfies ϵ-Differential Privacy.

4.4.5 Regret Analysis

In any round t, a winner set $\Phi(t)$is "bad" if $Q(\Phi(t)) < \alpha \cdot Q^{opt}$. $\Phi^{bad} = \{\Phi|Q(\Phi) < \alpha \cdot Q^{opt}\}$ denotes the set of bad winner set. For a client $n \in \mathcal{N}$, we have

$$\Delta_{min}^n = \alpha \cdot Q^{opt} - max\{Q(\Phi)|\Phi \in \Phi^{bad}, n \in \Phi\}, \tag{41}$$

$$\Delta_{max}^n = \alpha \cdot Q^{opt} - min\{Q(\Phi)|\Phi \in \Phi^{bad}, n \in \Phi\}. \tag{42}$$

We define $\Delta_{max} = max_{n\in\mathcal{N}}\Delta_{max}^n$ and $\Delta_{min} = min_{n\in\mathcal{N}}\Delta_{min}^n$. Hence we provide the regret bound of CMABA as follow.

Theorem 5. *The (α, β)-approximation regret of the CamPRA is at most*

$$\left(\frac{6lnT}{f^{-1}(\Delta_{min})^2} + \frac{\pi^2}{3} + 1\right) \cdot N \cdot \Delta_{max}. \tag{43}$$

Where $f(\cdot)$ is a bounded smoothness function.

Proof. In any round t, we use $E(t)$ to denote the event that our CamPRA can not conduct an α-approximate solution, i.e., $R(\Phi(t)) < R^{opt}$, $Pr[E(t)] = \mathbb{E}[\mathbb{I}\{E(t)\}] \leq 1 - \beta$. $\mathbb{I}\{H\} = 1$ if the H is true, and $\mathbb{I}\{H\} = 0$ if H is false.

Moreover, we define $\varepsilon_n(t)$ as the counter for client n after the initialization round and $\varepsilon_n(t)$ is updated as follow:

$$n = argmin_{i\in\Phi(t)}\varepsilon_i(t), \varepsilon_i(t) = \varepsilon_i(t) + 1. \tag{44}$$

If multiple clients satisfy this condition, we select an arbitrary client. Hence, when the winner set selected in round t is not the optimal set, one element of $\varepsilon_i(t)$ will increase by 1. In other words, the bad rounds number (i.e., the optimal client set is not selected) in the first t rounds is less than or equal to $\sum_{n\in\mathbb{N}}\varepsilon_n(t)$.

We define $u(t) = 6lnt/f^{-1}(\Delta_{min})^2$ as the sampling threshold for round t. If $\Phi(t) \in \Phi^{bad}$ is selected and ε_n is updated (i.e., round t is a bad round). Then we have

$$\begin{aligned}
&\sum_{n\in\mathbb{N}}\varepsilon_n(T) - N\cdot(u(T)+1) = \sum_{t=N+1}^{T}\mathbb{I}\{\Phi(t)\in\Phi^{bad}\} - N\cdot u(T)\\
&\le \sum_{t=N+1}^{T}\sum_{n\in\mathbb{N}}\mathbb{I}\{\Phi(t)\in\Phi^{bad}, \varepsilon_n(t) > \varepsilon_n(t-1), \varepsilon_n(t-1) > u(t)\}\\
&= \sum_{t=N+1}^{T}\mathbb{I}\{\Phi(t)\in\Phi^{bad}, \forall n\in\mathbb{N}, \varepsilon_n(t-1) > u(t)\}\\
&\le \sum_{t=N+1}^{T}(\mathbb{I}\{E(t)\} + \mathbb{I}\{\neg E(t), \Phi(t)\in\Phi^{bad}, \forall n\in\Phi(t), \varepsilon_n(t-1)\\
&\quad > u(t)\})\\
&\le \sum_{t=N+1}^{T}(\mathbb{I}\{E(t)\} + \mathbb{I}\{\neg E(t), \Phi(t)\in\Phi^{bad}, \forall n\in\Phi(t), s_n(t-1)\\
&\quad > u(t)\})
\end{aligned} \tag{45}$$

For the counter $s_n(t)$, we define its standard difference as $\mu_n(t) = min\{\sqrt{3lnt/2s_n^t}, 1\}$, the maximum standard difference is $\mu(t) = max\{\mu_n(t)|n \in \Phi(t)\}$. Let $H(t) = \{\forall n\in\mathbb{N}, |\hat{q}_n(s_n(t)) - q_n| \le \mu_n(t)\}$ to denote the event "the difference between the sampling average quality $\hat{q}_n(s_n(t))$ and the actual quality q_n is below the standard difference". SInce for any $n\in\mathbb{N}$,

$$\begin{aligned}
&\mathbb{P}\{|\hat{q}_n(s_n(t-1)) - q_n| \ge \mu_n(t-1)\}\\
&= \sum_{i=1}^{t-1}\mathbb{P}\{\hat{q}_n(s_n(t-1)) - q_n| \ge \mu_n(t-1), s_n(t-1) = ti\}\\
&\le \sum_{i=1}^{t-1}\mathbb{P}\{\hat{q}_n(t-1) - q_n| \ge \sqrt{\frac{3lnt}{2s_n(t-1)}}, s_n(t-1) = i\}.
\end{aligned} \tag{46}$$

The inequality $|\hat{q}_n(s_n(t-1)) - q_n| \ge \mu_n(t-1)\}$ has two cases: $\hat{q}_n(s_n(t-1)) - q_n \ge \mu_n(t-1)\}$ and $\hat{q}_n(s_n(t-1)) - q_n \le -\mu_n(t-1)\}$. Apply Chernoff-Hoeffding bound[XXX], we have

$$\begin{aligned}
&\sum_{i=1}^{t-1}\mathbb{P}\{\hat{q}_n(t-1) - q_n| \ge \sqrt{3lnt/2s_n(t-1)}, s_n(t-1) = i\}\\
&\le 2te^{-3lnt} = 2t^{-2}.
\end{aligned} \tag{47}$$

Thus $Pr\{\neg H(t)\} \le 2Nt^{-2}$. In other words, if the number of times n was selected $s_n(t)$ is large enough, the CamPRA can get a nice estimation of q_n and is unlikely ti select a bad winner set. As $Pr[H(t), \neg E(t), \Phi(t) \in \Phi^{bad}, s_n(t-1) > u(t)] = 0$, we have $Pr[\neg E(t), \Phi(t) \in \Phi^{bad}] \le Pr[\neg H(t) \le 2Nt^{-2}$. We have

$$\begin{aligned}
&\mathbb{E}[\sum_{i=1}^{N}s_n(t)] \le N\cdot(u(t)+1) + (1-\beta)(T-N) + \sum_{t=1}^{T}\frac{2N}{t^2}\\
&\le \frac{6Tlnt}{f^{-1}(\Delta_{min})^2} + (\frac{\pi^2}{3}+1)\cdot T + (1-\beta)(T-N).
\end{aligned} \tag{48}$$

It is noticed that in each round CamPRA outputs a bad winner set, the regret is at most $\Delta_{max} \ge \alpha\cdot R^{opt} - R(\Phi(t))$. Then we can get the regret upper bound:

$$\begin{aligned}
&Reg_{\alpha,\beta}(T)\\
&\le T\cdot\alpha\cdot\beta\cdot Q^{opt} - (T\cdot\alpha\cdot Q^{opt} - \mathbb{E}[\sum_{t=1}^{T}s_n(t)]\cdot\Delta_{max})\\
&\le \frac{6Tlnt}{f^{-1}(\Delta_{min})^2} + (\frac{\pi^2}{3}+1)\cdot T + (1-\beta)(T-N)\cdot\Delta_{max}\\
&- (1-\beta)T\cdot\alpha\\
&\le (\frac{6Tlnt}{f^{-1}(\Delta_{min})^2} + \frac{\pi^2}{3} + 1).
\end{aligned} \tag{49}$$

5 Numerical Results

5.1 Simulation Settings

In this section, we utilize federated learning to construct the AI model in Aerial Computing Networks with Pytorch 1.13.1 software and evaluate the performance of CamPRA using the MNIST dataset. MNIST is commonly used in federated learning with 60,000 training samples and 10,000 test samples. The global iteration round is 20. We differentiate the learning quality of clients by varying their dataset size and local iteration rounds. The number of clients $N = 20$ and budget per round $B(t) = 10$ in default. The differential privacy budget ϵ is between $(0.1, 1.1)$, and $\epsilon = 1.1$ in default.

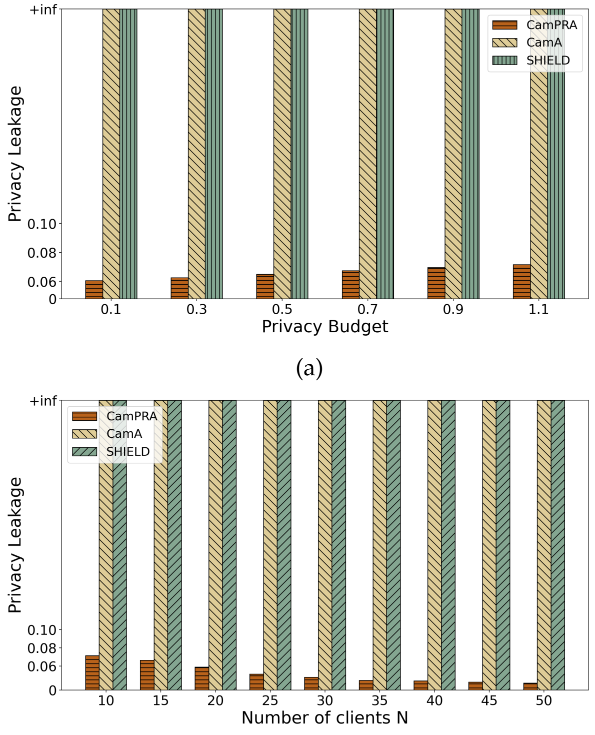
To evaluate the performance of our proposed CamPRA, we compared it with three benchmark methods, which are described below. 1)CamA. It is the proposed UCB-based learning quality-aware scheme without bid protection. We use CamA to analyze the effectiveness of our bid perturbation method. 2)SHIELD [13]. It is another privacy-preserving incentive scheme in which bids are perturbed on the trusted platform. The platform selects winners with the goal of minimizing social costs. We compare SHIELD with CamPRA to analyze the effect of our scheme on learning quality awareness. 3)Optimal. It assumes we know the real learning quality of each client in advance. Then the algorithm uses our designed auction to select winners and determine payments in each round of incentives. We compare optimal with our CamPRA to analyze its effectiveness on learning performance improvement.

To evaluate and verify the performance of our CamPRA, we adopt various metrics, including privacy leakage, learning accuracy, total reward, total payment, the number of clients, and the budget for incentives per round. Among them, privacy leakage refers to the degree of the disclosure of true bids in the incentive process. According to the definition in [33], the privacy leakage can be expressed as

$$PL = \frac{1}{\sum_{c\in C} \upsilon(c) \ln \frac{1}{\mathbb{P}[c^* = c]}}, \tag{50}$$

where c is the true submitted bid, c^* is a perturbed bid received by the aggregator. $\upsilon(b)$ is the probability of a true bid c in the set of received bids.

5.2 Simulation Results



(a)

(b)

Fig. 3: Comparison of privacy leakage in different schemes.

Fig. 3 compares the privacy risks of different schemes. Figure 3(a) shows the effect of privacy budget ϵ on privacy leakage. It can be observed that the privacy leakage risk of CamPRA increases with ϵ. This is because a lower value of ϵ results in a higher probability of confusing a bid c with other bids C^* , which provides more protection to users' bidding privacy. In contrast, the privacy leakage values of both CamA and SHIELD are infinitely large, because on the one hand, in CamA, clients submit true bids, and on the other hand, in SHIELD, users' bids are confused after being uploaded to the trusted platform, and in the event of an untrustworthy network environment or platform, SHIELD is unable to protect users' bid privacy. Fig. 3(b) shows the influence of the N on privacy leakage among different schemes with a fixed privacy budget $\epsilon = 1$. We can see CamPRA's privacy leakage decreases with increasing N since a higher probability of confusing a bid c with other bids C^* provides greater protection to users' bidding privacy. Similarly, the privacy leakage of CamA and SHIELD remains infinitely large.

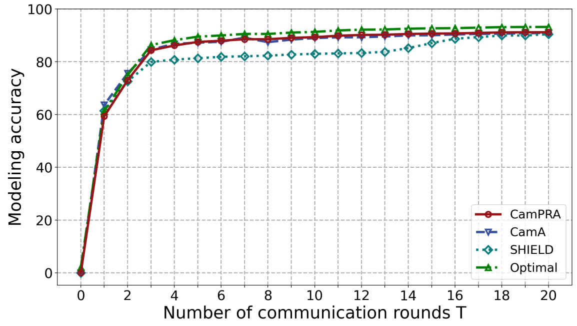
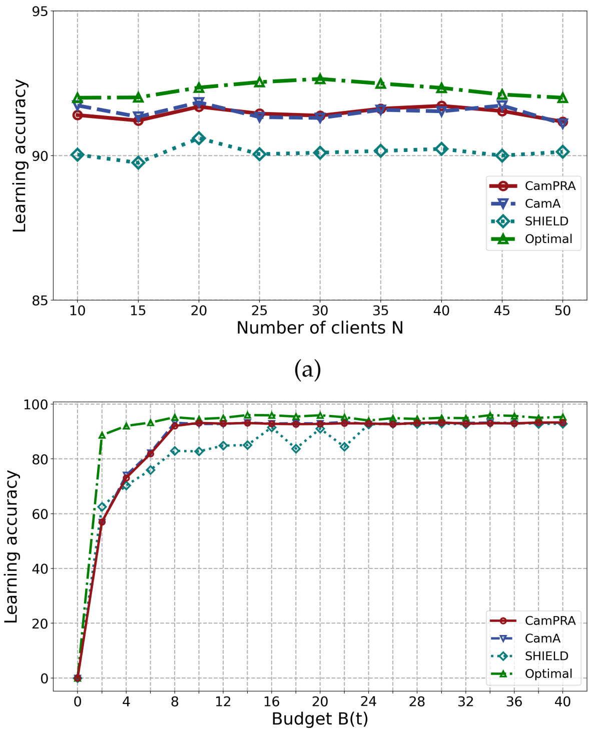


Fig. 4: Comparison of learning accuracy in different schemes.



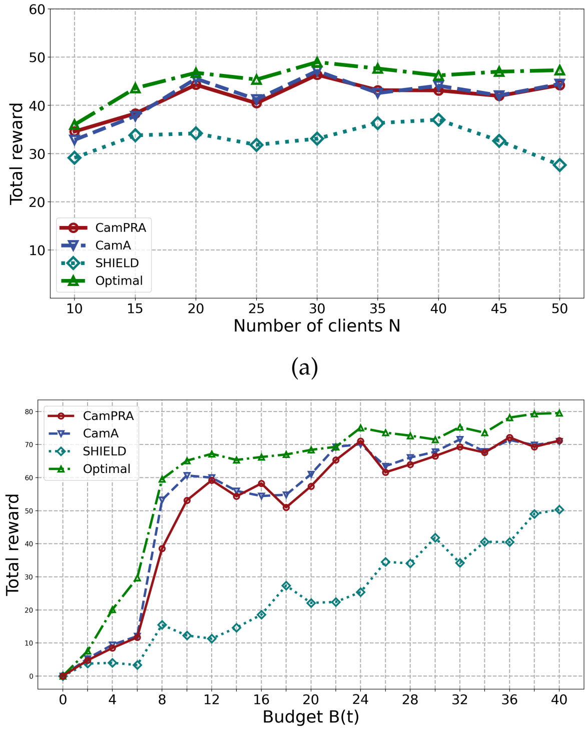
(a)

(b)

Fig. 5: The impact of the number of clients N and incentive budget per round $B(t)$ on learning accuracy.

Figure 4 and Figure 5 evaluate the effect of different schemes on the accuracy improvement of federated learning. Fig. 4 shows the model accuracy of the four schemes under different global communication rounds. It is observed that the accuracy of CamPRA is very close to optimal, indicating the effectiveness of its model quality perception method. CamPRA and CamA have the same linear height, indicating that CamPRA's bidding confusion mechanism does not affect the learning performance while protecting privacy. In contrast, the SHIELD scheme has the lowest accuracy due to its focus on reducing social costs, which ignores the improvement of learning accuracy. Fig. 5 shows the influence of client number N and the budget $B(t)$ per round on learning accuracy. A higher number of clients N makes it more difficult to select high-quality participants under a fixed limited budget $B(t)$. It can be seen from

Fig. 5(a) that as N increases, CamPRA can always optimize participants based on accurate learning quality perception and achieve close to optimal learning accuracy. The budget $B(t)$ per round limits the exploration efficiency of the algorithm for learning quality and affects learning accuracy. It can be seen from Fig. 5(b) that as $B(t)$ increases, CamPRA can quickly achieve close to optimal model accuracy, while SHIELD's focus on reducing learning costs leads to slow and unstable improvement in learning accuracy.



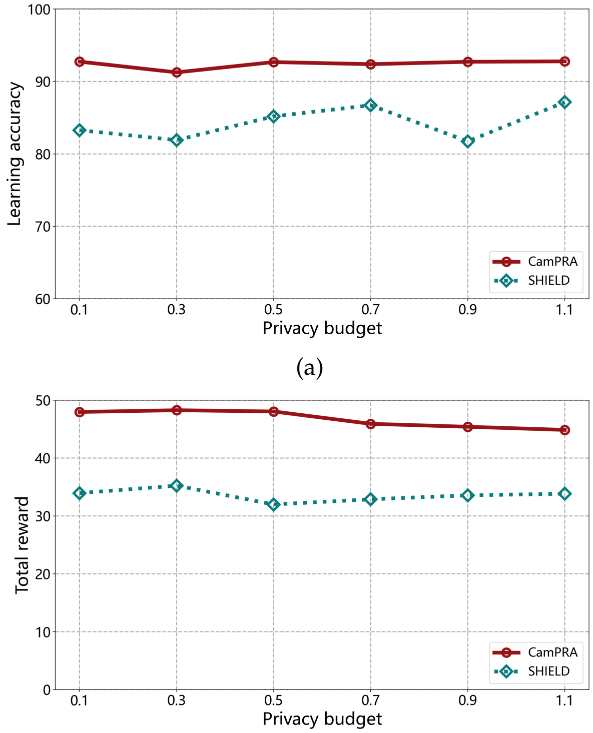
(a)

(b)

Fig. 6: The impact of the number of clients N and incentive budget per round $B(t)$ on total reward.

Fig. 6 evaluates the total rewards achieved by different schemes. In this paper, we define the total reward as the sum of the weighted sum of learning quality and learning time of the selected clients. We evaluate the effects of our proposed auction scheme with privacy-preserving and quality-aware by comparing the total rewards under different values of N and $B(t)$. It can be observed from Fig. 5(a) that our scheme can always achieve close to optimal reward as N increases. However, the total reward of CamPRA decreases when $N = 25$ and $N = 35$. It is because using multi-armed bandit to perceive and determine learning quality requires exploration of the unknown, which inevitably leads to the selection of clients with lower learning quality. It can be seen from Fig. 5(b) that as $B(t)$ increases, CamPRA consistently achieves close to optimal total reward and significantly outperforms that of the SHIELD scheme. The fluctuations in the reward are also due to the quality exploration of the unknown which cannot be avoided.

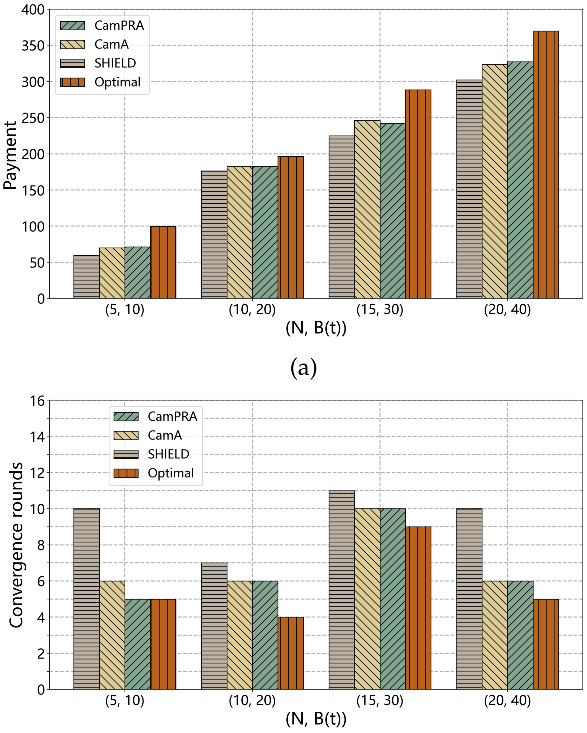
Figure 7 compares the performance of the two differential privacy-based schemes. As shown in Figure 7, As the differential privacy budget increases, the learning accuracy and achieved total reward of CamPRA do not change sig-



(a)

(b)

Fig. 7: The impact of differential privacy budget ϵ on learning accuracy and total reward.



(a)

(b)

Fig. 8: Comparison of time and economic cost of different schemes.

nificantly and are both higher than that of SHIELD. Figure 8 shows the cost of different schemes when the federated learning model converges under different B(t) and N. From Figure 8(a), we can see that the payment of our CamPRA is higher than that of SHIELD, but the improvement of the learning quality is also much higher than that of SHIELD. In Figure 8(b), when the federated learning model converges, the number of communication rounds of CamPRA is lower than that of SHIED, and close to that of the optimal scheme. Our CamPRA enables energy-efficient and accurate federated learning.

6 Conclusions

In this paper, we proposed an auction-based incentive scheme with privacy-protection and learning quality-awareness for effective federated learning in aerial computing networks. Combinatorial Multi-Armed Bandit was introduced to evaluate the user learning quality without any private information. Furthermore, we used differential privacy to protect the real client cost from inferring attacks. We further theoretically proved the proposed privacy-preserving incentive mechanism satisfies truthfulness, individual rationality, budget balance, and the convergence of CMAB regret. Simulation results demonstrated that our scheme can well balance the trade-off between privacy preservation and learning accuracy improvement. However, this work only focuses on exploring and optimizing learning quality in scenarios with a fixed number of clients. In the future, we will consider scenarios where clients can freely join and exit. The main challenge lies in adapting quickly to the dynamics of clients and effectively learning their true quality, which undoubtedly poses a significant challenge in the context of CMAB.

References

[1] F. A. Dicandia, N. J. Fonseca, M. Bacco, S. Mugnaini, and S. Genovesi, "Space-air-ground integrated 6g wireless communication networks: A review of antenna technologies and application scenarios," *Sensors*, vol. 22, no. 9, p. 3136, 2022.

[2] H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, "A survey on space-air-ground-sea integrated network security in 6g," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 53–87, 2022.

[3] Q.-V. Pham, R. Ruby, F. Fang, D. C. Nguyen, Z. Yang, M. Le, Z. Ding, and W.-J. Hwang, "Aerial computing: A new computing paradigm, applications, and challenges," *IEEE Internet of Things Journal*, vol. 9, no. 11, pp. 8339–8363, 2022.

[4] Z. Jia, Q. Wu, C. Dong, C. Yuen, and Z. Han, "Hierarchical aerial computing for internet of things via cooperation of haps and uavs," *IEEE Internet of Things Journal*, 2023.

[5] R. Kaewpuang, M. Xu, D. Niyato, H. Yu, and Z. Xiong, "Resource allocation in quantum key distribution (qkd) for space-air-ground integrated networks," in *2022 IEEE 27th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. IEEE, 2022, pp. 71–76.

[6] H. Yang, J. Zhao, Z. Xiong, K.-Y. Lam, S. Sun, and L. Xiao, "Privacy-preserving federated learning for uav-enabled networks: Learning-based joint scheduling and resource management," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 10, pp. 3144–3159, 2021.

[7] A. Alferaidi, K. Yadav, Y. Alharbi, W. Viriyasitavat, S. Kautish, and G. Dhiman, "Federated learning algorithms to optimize the client and cost selections," *Mathematical Problems in Engineering*, vol. 2022, 2022.

[8] S. Yeom, S. S. Kolekar, and K. Kim, "Effective edge server placement for efficient federated clustering," in *2022 23rd Asia-Pacific Network Operations and Management Symposium (APNOMS)*. IEEE, 2022, pp. 1–4.

[9] Z. Wei, Q. Pei, N. Zhang, X. Liu, C. Wu, and A. Taherkordi, "Lightweight federated learning for large-scale iot devices with privacy guarantee," *IEEE Internet of Things Journal*, vol. 10, no. 4, pp. 3179–3191, 2023.

[10] Y. Chen, Y. Zhang, S. Wang, F. Wang, Y. Li, Y. Jiang, L. Chen, and B. Guo, "Dim-ds: Dynamic incentive model for data sharing in federated learning based on smart contracts and evolutionary game theory," *IEEE Internet of Things Journal*, vol. 9, no. 23, pp. 24 572–24 584, 2022.

[11] J. Lee, D. Kim, and D. Niyato, "A novel joint dataset and incentive management mechanism for federated learning over mec," *IEEE Access*, vol. 10, pp. 30 026–30 038, 2022.

[12] J. Y. Kim, "A unique and robust social contract: An application to negotiations with probabilistic conflicts," *Global Economic Review*, vol. 51, no. 1, pp. 61–74, 2022.

[13] C. Ying, H. Jin, X. Wang, and Y. Luo, "Double insurance: Incentivized federated learning with differential privacy in mobile crowdsensing," in *2020 International Symposium on Reliable Distributed Systems (SRDS)*, 2020, pp. 81–90.

[14] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021.

[15] Y. Zhan, J. Zhang, Z. Hong, L. Wu, P. Li, and S. Guo, "A survey of incentive mechanism design for federated learning," *IEEE Transactions on Emerging Topics in Computing*, vol. 10, no. 2, pp. 1035–1044, 2021.

[16] Y. Zhan, P. Li, S. Guo, and Z. Qu, "Incentive mechanism design for federated learning: Challenges and opportunities," *IEEE Network*, vol. 35, no. 4, pp. 310–317, 2021.

[17] L. Witt, M. Heyer, K. Toyoda, W. Samek, and D. Li, "Decentral and incentivized federated learning frameworks: A systematic literature review," *IEEE Internet of Things Journal*, 2022.

[18] J. Heiss, E. Grünewald, S. Tai, N. Haimerl, and S. Schulte, "Advancing blockchain-based federated learning through verifiable off-chain computations," in *2022 IEEE International Conference on Blockchain (Blockchain)*. IEEE, 2022, pp. 194–201.

[19] Y. E. Oktian, B. Stanley, and S.-G. Lee, "Building trusted federated learning on blockchain," *Symmetry*, vol. 14, no. 7, p. 1407, 2022.

[20] J. Han, A. F. Khan, S. Zawad, A. Anwar, N. B. Angel, Y. Zhou, F. Yan, and A. R. Butt, "Tokenized incentive for federated learning," in *Proceedings of the Federated Learning Workshop at the Association for the Advancement of Artificial Intelligence (AAAI) Conference*, 2022.

[21] T. Liu, B. Di, S. Wang, and L. Song, "A privacy-preserving incentive mechanism for federated cloud-edge learning," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 1–6.

[22] J. P. D. B. Gonçalves and R. D. S. Villaça, "A blockchained incentive architecture for federated learning," in *2022 IEEE International Conference on Blockchain (Blockchain)*. IEEE, 2022, pp. 482–487.

[23] R. Fantacci and B. Picano, "A d2d-aided federated learning scheme with incentive mechanism in 6g networks," *IEEE Access*, 2022.

[24] A. Xiong, Y. Chen, H. Chen, J. Chen, S. Yang, J. Huang, Z. Li, and S. Guo, "A truthful and reliable incentive mechanism for federated learning based on reputation mechanism and reverse auction," *Electronics*, vol. 12, no. 3, p. 517, 2023.

[25] R. Hu and Y. Gong, "Trading data for learning: Incentive mechanism for on-device federated learning," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.

[26] J. Weng, J. Weng, J. Zhang, M. Li, Y. Zhang, and W. Luo, "Deepchain: Auditable and privacy-preserving deep learning with blockchain-based incentive," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 5, pp. 2438–2455, 2019.

[27] J. Zhang, Y. Wu, and R. Pan, "Auction-based ex-post-payment incentive mechanism design for horizontal federated learning with reputation and contribution measurement," *arXiv preprint arXiv:2201.02410*, 2022.

[28] L. Li, X. Yu, X. Cai, X. He, and Y. Liu, "Contract theory based incentive mechanism for federated learning in health crowdsensing," *IEEE Internet of Things Journal*, 2023.

[29] T. H. T. Le, N. H. Tran, Y. K. Tun, M. N. Nguyen, S. R. Pandey, Z. Han, and C. S. Hong, "An incentive mechanism for federated learning in wireless cellular networks: An auction approach," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4874–4887, 2021.

[30] H. Moudoud, S. Cherkaoui, and L. Khoukhi, "Towards a secure and reliable federated learning using blockchain," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 01–06.

[31] W. Sun, S. Lian, H. Zhang, and Y. Zhang, "Lightweight digital twin and federated learning with distributed incentive in air-ground 6g networks," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 3, pp. 1214–1227, 2023.

[32] M. Xiao, B. An, J. Wang, G. Gao, S. Zhang, and J. Wu, "Cmab-based reverse auction for unknown worker recruitment in mobile crowdsensing," *IEEE Transactions on Mobile Computing*, vol. 21, no. 10, pp. 3502–3518, 2022.

[33] J. Lin, D. Yang, M. Li, J. Xu, and G. Xue, "Frameworks for privacy-preserving mobile crowdsensing incentive mechanisms," *IEEE Transactions on Mobile Computing*, vol. 17, no. 8, pp. 1851–1864, 2018.

Peng Wang (Student Member,IEEE) received the B.Eng. degree software engineering from Harbin Institute of Technology, Harbin, China in 2017. He is currently working toward the Ph.D. degree in cyber engineering with Xidian University, Xi'an, China. His research interests include blockchain, access control, Internet of things.

Yi Yang (Student Member,IEEE) received the B.Eng. degree software engineering from Northwestern Polytechnical University, Xi'an, China in 2017. He is currently working toward the Ph.D. degree in Cybersecurity with Northwestern Polytechnical University, Xi'an, China. His research interests include incentive mechanism, wireless mobile communications, digital twins, Internet of things, federated learning.

Wen Sun (Senior Member, IEEE) received the B.E. degree from the Harbin Institute of Technology, Harbin, China, in 2009, and the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2014. She is currently a Full Professor with the School of Cybersecurity, Northwestern Polytechnical University, Xi'an, China. She has authored or coauthored more than 50 peer-reviewed papers in various prestigious IEEE journals and conferences, including IEEE Transactions on Industrial Informatics, IEEE Transactions on Wireless Communications, IEEE Network, and IEEE Wireless Communications. Her research interests include wireless mobile communications, IoT, 5G, and blockchain. She was the recipient of the Best Paper Award of GlobeCom2019. She is the publicity Chair of WiMob2019 and CNS2020, and a TPC Member of ICC and GlobeCom in 2018 and 2019.

Qubeijian Wang (Member, IEEE) received the B.E. degree in electrical engineering from Xi'an Jiaotong-Liverpool University, Suzhou, China, and the University of Liverpool, Liverpool, U.K., in 2015, the M.E. degree in telecommunications from the University of Melbourne, Parkville, VIC, Australia, in 2017, and the Ph.D. degree in electronic information technology from Macau University of Science and Technology, Macau, China, in 2020. He is currently an Assistant Professor with the School of Cybersecurity, Northwestern Polytechnical University, Xi'an, China. His research interests include UAV-aided communications, physical-layer security, and large-scale network performance analysis. Dr. Wang serves as a TPC Member for conferences, including GlobeCom 2021 and 2022, and VTC2021-Fall and a reviewer for various prestigious IEEE journals and conferences.

Bin Guo (Senior Member, IEEE) received the Ph.D. degree in computer science from Keio University, Japan, in 2009. He was a Post-Doctoral Researcher with the Institute Telecom SudParis, France. He is currently a Professor with Northwestern Polytechnical University, China. He has published more than 90 papers in refereed journals, conference proceedings, and book chapters. His research interests include ubiquitous computing, mobile crowd sensing, and HCI. He is the General Co-Chair of the 12th IEEE International Conference on Ubiquitous Intelligence and Computing (IEEE UIC'15) and the Program Chair of IEEE CPSCom'16, ANT'14, and UIC'13. He has served as an Associate Editor for the IEEE Communications Magazine and the IEEE TRANSACTIONS ON HUMAN-MACHINE-SYSTEMS and a Guest Editor for the ACM Transactions on Intelligent Systems and Technology and the IEEE INTERNET OF THINGS.

Jianhua He (Senior Member, IEEE) received the Ph.D. degree from Nanyang Technological University, Singapore, in 2002. He was with the University of Bristol, Swansea University, and Aston University. He is currently a Reader with the University of Essex, U.K. He has published more than 150 research papers in refereed international journals and conferences. He is the Coordinator of EU Horizon2020 projects COSAFE and VESAFE on cooperative connected autonomous vehicles. His main research interests include 5G/6G wireless communications and networks, connected vehicles, autonomous driving, the Internet of Things, mobile edge computing, intelligent transport systems, data analytics, AI, and machine learning. He was the Workshop Chair of MobiArch'20 and ICAV'21 and a Steering Committee Member of MobiArch'21 and MobiArch'22. He is a member of the editorial board of IEEE Wireless Communications Letters and The Computer Journal.

Yuanguo Bi (Member, IEEE) received the Ph.D. degree in computer science from Northeastern University, Shenyang, China, in 2010.,He was a visiting Ph.D. student with the BroadBand Communications Research Lab, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, from 2007 to 2009. He is currently a Professor with the School of Computer Science and Engineering, Northeastern University.His research interests include medium access control, QoS routing, multihop broadcast, and mobility management in vehicular networks, software-defined networking, and mobile edge computing.