An Incentive Mechanism for Federated Learning in Wireless Cellular Networks: An Auction Approach

Tra Huong Thi Le, *Member, IEEE*, Nguyen H. Tran, *Senior Member, IEEE*, Yan Kyaw Tun,
Minh N. H. Nguyen, *Member, IEEE*, Shashi Raj Pandey, Zhu Han, *Fellow, IEEE*,
and Choong Seon Hong, *Senior Member, IEEE*

Abstract—Federated Learning (FL) is a distributed learning framework that can deal with the distributed issue in machine learning and still guarantee high learning performance. However, it is impractical that all users will sacrifice their resources to join the FL algorithm. This motivates us to study the incentive mechanism design for FL. In this paper, we consider a FL system that involves one base station (BS) and multiple mobile users. The mobile users use their own data to train the local machine learning model, and then send the trained models to the BS, which generates the initial model, collects local models and constructs the global model. Then, we formulate the incentive mechanism between the BS and mobile users as an auction game where the BS is an auctioneer and the mobile users are the sellers. In the proposed game, each mobile user submits its bids according to the minimal energy cost that the mobile users experiences in participating in FL. To decide winners in the auction and maximize social welfare, we propose the primal-dual greedy auction mechanism. The proposed mechanism can guarantee three economic properties, namely, truthfulness, individual rationality and efficiency. Finally, numerical results are shown to demonstrate the performance effectiveness of our proposed mechanism.

Index Terms—Federated learning, auction game, resource allocation, wireless network, incentive mechanism.

I. INTRODUCTION

CURRENTLY, according to the report of International Data Corporation, there are nearly 3 billions smart-phones on the world [1], [2], which generate a

Tra Huong Thi Le, Yan Kyaw Tun, Minh N. H. Nguyen, Shashi Raj Pandey, and Choong Seon Hong are with the Department of Computer Science and Engineering, Kyung Hee University, Yongin-si 17104, South Korea (e-mail: huong_tra25@khu.ac.kr; ykyawtun7@khu.ac.kr; minhnhn@khu.ac.kr; shashiraj@khu.ac.kr; cshong@khu.ac.kr).

Nguyen H. Tran is with the School of Computer Science, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: nguyen.tran@sydney.edu.au).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Yongin-si 17104, Korea.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2021.3062708.

Digital Object Identifier 10.1109/TWC.2021.3062708

huge amount of personal data. Nowadays, mobile devices equipped with specialized hardware architectures and computing engines can handle the machine learning problem effectively. In addition, the application of machine learning techniques in mobile devices has grown rapidly. Furthermore, due to the limitation of wireless communication resources and privacy protection problem, the conventional central machine learning techniques, which upload all data of mobile devices to the central sever, are becoming less attractive. For this reason, federated learning (FL) is promoted, which is implemented distributively at the edge of the network [3], [4]. In FL, mobile users can collaboratively train a global model using their own local data. Mobile users compute the updates of the current global model, and then send back the updates to the central server for aggregation and build a new global model. This process is repeated until an accuracy level of the global learning model is achieved. By this way, FL can preserve the personal information and data of mobile users. In addition, FL will significantly promote services that have unparalleled versatile data collection and model training on a large scale. Take the app Waze as an example. This application can help the users in avoiding heavy traffic roads, but users have to share their own locations to the server. If FL is applied to this app, users only need to send the intermediate gradient values to the server rather than the raw data [5]. Last but not least, the development of mobile edge computing provides an immense exposure to extract the benefits of FL [6], [7].

In spite of the above mentioned benefits of FL, there are remaining two key challenges of having an efficient FL framework. The first challenge is the economic challenge. Data samples per mobile device are small to train a high-quality learning model so a large number of mobile users are needed to ensure cooperation. In addition, the mobile users who join the learning process are independent and uncontrollable. Here, mobile users may not be willing to participate in the learning due to the energy cost incurred by model training. In other words, the base station (BS), which generates the global model, has to stimulate the mobile users for participation. The second challenge is the technical challenge. On the one hand, we need users to collectively provide a large number of data samples to enable FL without sharing their private data. On the other hand, we need to protect model from imperfect updates. The global loss minimization problem should enable in (a) proper assessment of the quality of the local solution to

improve personalization and fairness amongst the participating clients while training a global model, (b) effective decoupling of the local solvers, thereby balancing communication and computation in the distributed setting. Moreover, we need to consider wireless resource limitations (such as time, antenna number and bandwidth) affecting the performance of FL. Besides, the limited energy of wireless devices is a crucial challenge for deploying FL. Indeed, it is necessary to optimize the energy efficiency for FL implementation because of these resource constraints.

To deal with the above challenges, in this paper, we model the FL service between the BS and mobile users as an auction game in which the BS is buyer and mobile users are sellers. In particular, the BS first initiates and announces a FL task. When each mobile user receives the FL task information, they decide the amount of resources required to participate in the model training. After that, each mobile user submits a bid, which includes the required amount of resource, local accuracy, and the corresponding energy cost, to the BS. Moreover, the BS plays the role of auctioneer to decide the winners among mobile users as well as clear payment for the winning mobile users. In addition, the auction used in this paper is a type of combinational auction [8], [9] since each mobile user can bid for combinations of resources. However, the proposed auction mechanism allows mobile users sharing the resources at the BS, which is different from the conventional combinatorial auction. The proposed mechanism directly determines the trading rules between the buyer (BS) and sellers (mobile users) and motivates the mobile users to participate in the model training. Compared with other incentive mechanism approaches (e.g., contract theory [10]) in which the service market is a monopoly market, where mobile users can only decide whether or not to accept the contracts, the proposed auction enables mobile users to bids any combinations of resources. Moreover, the proposed auction mechanism can simultaneously provide truthfulness and individual rationality. An auction mechanism is truthful if a bidder's utility does not increase when that bidder makes other bidding strategies, rather than the true value. Revealing the true value is a dominant strategy for each participating user regardless of what strategies other users use [11]. An absent-truthfulness auction mechanism could leave the door to possible market manipulation and produce inferior results [12]. Additionally, if the value of any bidder is non-negative, an auction process will ensure individual rationality. The difference between our proposed auction scheme and other work, e.g., [11] is that we included user cost calculations in the submitted bid. The cost is applied by users to evaluate the value of learning participation. Users not only consider the payment they can receive from learning but also the impact of their bidding result on global performance (time constraints, global accuracy). In other words, users try to obtain high payment and help the system to achieve high social welfare. Moreover, user independently evaluates the value of learning without leaking their privacy parameters and information.

The contributions of this paper are summarized as follows:

- We propose an auction framework for the wireless FL services market. *Then, we present the bidding cost in every user's bid submitted to the BS.* From the mobile users' perspective, each mobile user makes optimal decisions on the amount of resources and local accuracy to minimize weighted sum of completion time and energy costs while the delay requirement for FL is satisfied. To solve the cost decision problem, a low-complexity iterative algorithm is proposed.

- From the perspective of the BS, we formulate the winner selection problem in the auction game as the social welfare maximization problem, which is an NP-hard problem considering the limitation of the wireless resource. We propose a primal-dual greedy algorithm to deal with the NP-hard problem of selecting the winning users and critical value - based payment. We also proved that the proposed auction mechanism is truthful, individual rational and computationally efficient.

- Finally, we carry out the numerical study to show that a proposed auction mechanism can guarantee the approximation factor of the integrality to the maximal welfare that is derived by the optimal solution and outperforms compared with baseline.

The rest of this paper is organized as follows. Section II summarizes the related work. The system model is introduced in Section III. We describe the problem formulation in Section IV. We present the auction-based resource purchasing mechanism in Section V. Simulation results are given in Section VI. Finally, Section VII concludes the paper.

II. RELATED WORKS

Due to the resource constraints and the heterogeneity of mobile users, some focus issues are resource allocation, client selection and incentive mechanism to improve the efficiency of FL. The authors in [5] posed the joint learning and transmission energy minimization problem for FL. In this paper, all users upload their learning model to the BS in a synchronous manner. The work in [13] also considered the latency and energy consumption minimization problem for the case of asynchronous transmission. The work in [14] explored the problem of reducing the learning loss function by considering packet errors over wireless links, but this research ignored the computation delay of the local learning model. In [15], the authors suggested energy-efficient strategies for allocating bandwidth and scheduling while at the same time, guaranteeing learning efficiency. The derived optimal policies allocate more bandwidth to those scheduled devices with weaker channels or lower computing capacities, which are the bottlenecks of synchronized model updates in FL.

However, the works in [5], [13]–[16] overlooked the problem of client selection to build a high-quality machine learning model. The authors in [17] designed a protocol called FedCS. The FedCS protocol has a resource request phase to gather information such as computing power and wireless channel states from a subset of randomly selected clients, i.e., FL workers, that are able to finish the local training punctually. A q-FedAvg training algorithm for selecting the client by the computational power was proposed in [18]. This proposed algorithm in [18] can improve the training efficiency and solve the fairness issue. The study in [19]

recommended combining deep reinforcement learning (DRL) and FL frameworks with mobile edge systems to optimize computing, caching and communication. The study in [20] jointly considered the device selection and beamforming fast global model aggregation. They used the principle of over-the-air computation to exploit signal superposition multiple access channels. In [16], the authors proposed a reliable worker selection scheme for federated learning tasks. Workers are selected based on their reputation values, which are calculated according to local reputation opinions generated from direct interaction histories, and recommended reputation opinions of other task publishers stored on an open-access consortium blockchain named reputation blockchain. However, the performance of the proposed work in [16] depends on the reputation threshold, which remains to be an open issue.

Concerning the incentive mechanism design, the authors in [21] proposed a Stackelberg game model to investigate the interactions between the server and the mobile devices in a cooperative relay communication network. The mobile devices determine the price per unit of data for individual profit maximization, while the server chooses the size of training data to optimize its own profit. The authors in [10] studied the block-chained FL architecture and proposed the contract theory based payment mechanism to incentivize the mobile devices to take part in the FL. However, [10] largely provided a latency analysis for the related applications. The work in [22] proposed a hierarchical incentive mechanism design for FL that considers multiple model owners and the formation of numerous federations for mobile crowdsensing in the system model. The authors in [22] leveraged on the self-revealing mechanism in contract theory under information asymmetry for the incentive mismatch challenge between workers and model owners. For the incentive challenge among model owners, to ensure the stability of a federation through preventing free-riding attacks, we use the coalitional game theory approach that rewards model owners based on their marginal contributions. The model owner will choose to join the federation to maximize marginal payoffs while minimizing contractual costs. However, the wireless nature of the communication medium is not taken into account in [22]. The work in [23] designed and analyzed a novel crowdsourcing framework to enable FL. In [23], a two-stage Stackelberg game model was adopted to jointly study the utility maximization of the participating clients and multi-access edge computing (MEC) server interacting via an application interface to construct a high-quality learning model. In [24], a Stackelberg game for FL in IoT was proposed to tackle the challenge of incentivizing people to join the FL by contributing their computational power and data. For the cases where the knowledge of participants' decisions and accurate contribution evaluation are accessible, the Nash Equilibrium was derived, and an algorithm based on DRL was built to unknown the knowledge of participants' decisions and accurate contribution evaluation for the cases. However, both [23] and [24] studied only the uniform pricing scheme for participants.

Different from the Stackelberg game and contract theory, the auction mechanism allows mobile users to actively report its cost. Therefore, the BS is capable of understanding their

Description	Notation
Global loss function	$L(\omega)$
Computing energy consumption of user n	E_n^{comp}
Computing time of one local iteration of user n	T_n^{comp}
Transmission time of user n	T_n^{com}
Transmission energy of user n	E_n^{com}
Total transmission time of user n	T_n^{tol}
Total energy consumption of user n	E_n^{tol}
Local accuracy of user n	ε_n
Computing resource of user n	f_n
Transmission power of user n	p_n
Antenna number desired by user n	A_n
The real cost of the ith bid of user n	V_{ni}
The claimed cost of the ith bid of user n	v_{ni}
The satisfaction level of the ith bid of user n	χ_{ni}
The winner indicator of the ith bid of user n	x_{ni}

status and requests adequately. Reference [25] adopted the multi-dimensional procurement auction to motivate nodes to participate in FL. However, there exist some differences between [25] and our work: (a) [25] the bids submitted by edge node clarifies the combination of resources and the expected payment, which is based on the private cost parameters while in our work, the bid declares the combination of resource, local accuracy and the cost, which is determined based on latency and energy cost models; (b) the winner selection in [25] based on the scoring function announced by the aggregator while in our work, winners are selected in order to optimize social welfare and ensure resource efficiency.

III. SYSTEM MODEL: FEDERATED LEARNING SERVICES MARKET

Table I presents the key notations used in this paper.

A. Preliminary of Federated Learning

Consider a cellular network in which one BS and a set $\mathcal{N}$ of N users cooperatively perform a FL algorithm for model learning, as shown in Fig. 1. Each user n has s_n local data samples. Each data set $s_n = \{a_{nk}, b_{nk,1 \leq k \leq s_n}\}$ where a_{nk} is an input and b_{nk} is its corresponding output. The FL model trained by the dataset of each user is called the local FL model, while the FL model at the BS aggregates the local model from all users as the global FL model. We define a vector ω as the model parameter. We also introduce the loss function $l_n(\omega, a_{nk}, b_{nk})$ that captures the FL performance over input vector a_{nk} and output b_{nk}. The loss function may be different, depending on the different learning tasks. The total loss function of user n will be

$$L_n(\omega) = \frac{1}{s_n} \sum_{k=1}^{s_n} l_n(\omega, a_{nk}, b_{nk}). \tag{1}$$

Then, the learning model is the minimizer of the following global loss function minimization problem

$$\min_{\omega} \quad L(\omega) = \frac{1}{S} \sum_{n=1}^{N} \sum_{k=1}^{s_n} l_n(\omega, a_{nk}, b_{nk}), \tag{2}$$

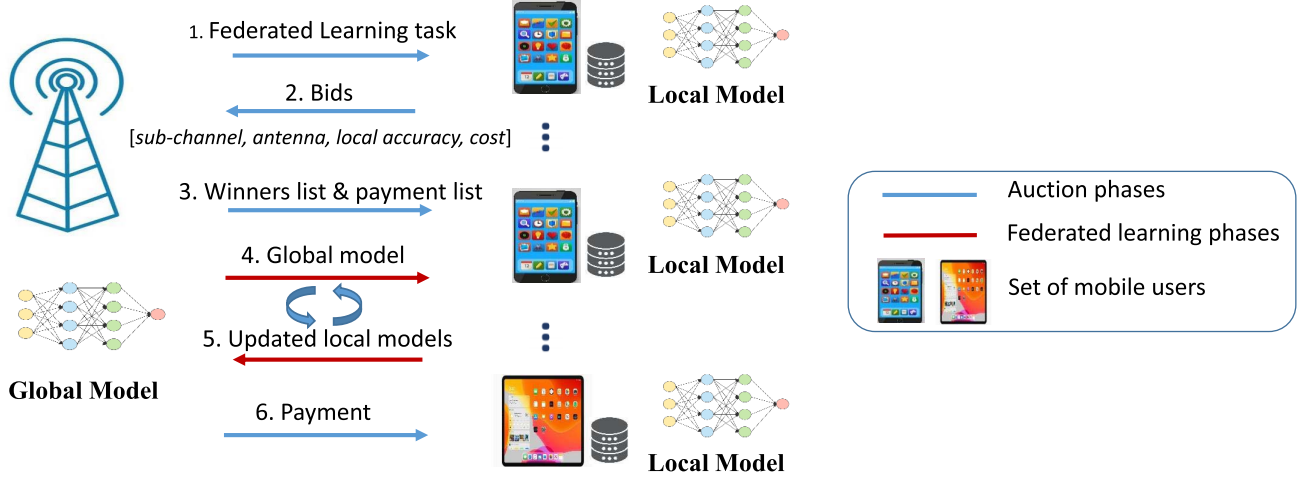where $S = \sum_{n=1}^{N} s_n$ is the total data samples of all users.

Fig. 1.   System model.

To solve the problem in (2), we adopt the FL algorithm of [26]. The algorithm uses an iterative approach that requires a number of global iterations (i.e., communication rounds) to achieve a global accuracy level. In each global iteration, there are interactions between the users and BS. Specifically, at a given global iteration t, users receive the global parameter ω^t, users computes $\nabla L_n(\omega^t), \forall n$ and send it to the BS. The BS computes [13]

$$\nabla L(\omega^t) = \frac{1}{N} \sum_{n=1}^{N} \nabla L_n(\omega^t), \qquad (3)$$

and then broadcasts the value of $\nabla L(\omega^t)$ to all participating users. Each participating user n will use local training data s_n to solve the local FL problem is defined as

$$\min_{\phi_n} \quad \mathcal{G}_n(\omega^t, \phi_n)$$
$$= L_n(\omega^t + \phi_n) - \left(\nabla L_n(\omega^t) - \varpi \nabla L(\omega^t)\right)^T \phi_n, \quad (4)$$

where ϕ_n represents the difference between global FL parameter and local FL parameter for user n. Each participating user n uses the gradient method to solve (4) with local accuracy ε_n that characterizes the quality of the local solution, and produces the output ϕ_n that satisfies

$$\mathcal{G}_n(\omega^t, \phi_n) - \mathcal{G}_n(\omega^t, \phi_n^*) < \varepsilon_n(\mathcal{G}_n(\omega^t, 0) - \mathcal{G}_n(\omega^t, \phi_n^*)). \quad (5)$$

Solving (4) also takes multiple local iterations to achieve a particular local accuracy. Then each user n sends the local parameter ϕ_n to the BS. Next, the BS aggregates the local parameters from the users and computes

$$\omega^{t+1} = \omega^t + \frac{1}{N} \sum_{n=1}^{N} \phi_n^t, \qquad (6)$$

and broadcasts the value to all users, which is used for next iteration $t+1$. This process is repeated until the global accuracy γ of (2) is obtained.

Assume that $L_n(\omega)$ is H-Lipschitz continuous and π-strongly convex, i.e.,

$$\pi \boldsymbol{I} \preceq \nabla^2 L_n(\omega) \preceq H\boldsymbol{I}, \quad \forall n \in \mathcal{N},$$

the general lower bound on the number of global iterations is depends on local accuracy ε and the global accuracy γ as [13]:

$$I^g(\gamma, \varepsilon) = \frac{C_1 \log(1/\gamma)}{1 - \varepsilon}, \qquad (7)$$

where the local accuracy measures the quality of the local solution as described in the preceding paragraphs.

In (7), we observe that a very high local accuracy (small ε) can significantly boost the global accuracy γ for a fixed number of global iterations I^g at the BS to solve the global problem. However, each user n has to spend excessive resources in terms of local iterations, I_n^l to attain a small value of ε_n. The lower bound on the number of local iterations needed to achieve local accuracy ε_n is derived as [13]

$$I_n^l(\varepsilon_n) = \vartheta_n \log\left(\frac{1}{\varepsilon_n}\right), \qquad (8)$$

where $\vartheta_n > 0$ is a parameter choice of user n that depends on parameters of $L_n(\omega)$ [13]. In this paper, we normalize $\vartheta_n = 1$. Therefore, to address this trade-off, the BS can setup an economic interaction environment to motivate the participating users to enhance local accuracy ε_n. Correspondingly, with the increased payment, the participating users are motivated to attain better local accuracy ε_n (i.e., smaller values), which as noted in (7) can improve the global accuracy γ for a fixed number of iterations I^g of the BS to solve the global problem. In this case, the corresponding performance bound in (7) for the heterogeneous responses ϵ_n can be updated to catch the statistical and system-level heterogeneity regarding the worst case of the participating users' responses as:

$$I^g(\gamma, \varepsilon_n) = \frac{\varpi \log(1/\gamma)}{1 - \max_n \varepsilon_n}, \quad \forall n. \qquad (9)$$

B. Computation and Communication Models for Federated Learning

The contributed computation resource that user n contributes for local model training is denoted as f_n. Then, c_n denotes the number of CPU cycles needed for the user n to perform one sample of data in local training. Thus, energy consumption of the user for one local iteration is presented as

$$E_n^{com}(f_n) = \zeta c_n s_n f_n^2, \tag{10}$$

where ζ is the effective capacitance parameter of computing chipset for user n. The computing time of a local iteration at the user n is denoted by

$$T_n^{comp} = \frac{c_n s_n}{f_n}. \tag{11}$$

It is noted that the uplink from the users to the BS is used to transmit the parameters of the local FL model while the downlink is used for transmitting the parameters of the global FL model. In this paper, we just consider the uplink bandwidth allocation due to the relation of the uplink bandwidth and the cost that user experiences during learning a global model. We consider the uplink transmission of an OFDMA-based cellular system. A set of $\mathcal{B} = \{1, 2, \ldots, B\}$ subchannels each with bandwidth W. Moreover, the BS is equipped with A antennas and each user equipment has a single antenna (i.e., multi-user MIMO). We assume A to be large (e.g., several hundreds) to achieve massive MIMO effect which scales up traditional MIMO by orders of magnitude. Massive MIMO uses spatial-division multiplexing. In this paper, we assume that the BS has perfect channel state information (CSI) and the channel gain is perfectly estimated, similar to [27], [28]. Then, the achievable uplink data rate of mobile user n is expressed as [28], [29]:

$$r_n = b_n W \log_2 \left(1 + \frac{(A_n - 1)p_n h_n}{b_n W N_0}\right), \tag{12}$$

where p_n is the transmission power of user n, h_n is the channel gain of peer to peer link between user and the BS, N_0 is the background noise, A_n is the number of antennas the BS assigns to user n, and b_n is the number of sub-channels that user n uses to transmit the local model update to the BS.

We denote σ as the data size of a local model update and it is the same for all users. Therefore, the transmission time of a local model update is

$$T_n^{com}(p_n, A_n, b_n) = \frac{\sigma}{r_n}. \tag{13}$$

To transmit local model updates in a global iteration, the user n uses the amount of energy given as

$$E^{com}(p_n, f_n, A_n, b_n) = T^{com} p_n = \frac{\sigma p_n}{r_n}. \tag{14}$$

Hence, the total time of one global iteration for user n is denoted as

$$T_n^{tol}(p_n, f_n, A_n, b_n, \varepsilon_n)$$
$$= \log\left(\frac{1}{\varepsilon_n}\right) T_n^{comp}(f_n) + T_n^{com}(p_n, A_n, b_n). \tag{15}$$

Therefore, the total energy consumption of a user n in one global iteration is denoted as follows

$$E_n^{tol}(p_n, f_n, A_n, b_n, \varepsilon_n)$$
$$= \log\left(\frac{1}{\varepsilon_n}\right) E_n^{comp}(f_n) + E_n^{com}(p_n, A_n, b_n). \tag{16}$$

C. Auction Model

As described in Fig. 1, the BS first initializes the global network model. Then, the BS announces the auction rule and advertises the FL task to the mobile users. The mobile users then report their bids. Here, mobile user n submits a set of I_n of bids to the BS. A bid Δ_{ni} denotes the ith bid submitted by the mobile user n. Bid b_{ni} consists of the resource (sub-channel number b_{ni}, antenna number A_{ni}, local accuracy level ε_{ni}) and the claimed cost v_{ni} for the model training. Each mobile user n has its own discretion to determine its true cost V_{ni}, which will be presented in Section IV. Let x_{ni} be a binary variable indicating the bid Δ_{ni} wins or not. After receiving all the bids from mobile users, the BS decides winners and then allocates the resource to the winning mobile users. The winning mobile users join the FL and receive the payment after finishing the training model.

Remark: In each bid, the bidder declares the requested resources, the local accuracy, and the corresponding cost. And the cost is calculated before submitting bids. Therefore, the cost corresponding to the requesting resources can be included in the bid during the bidding process.

Following we discuss one practical usage of our proposed auction scheme in FL. Let's consider a concrete example of a mobile phone keyboard such as Gboard (Google Keyboard). A large amount of local data will be generated when users interact with the keyboard app on their mobile devices. Suppose that Google server wants to train a next-word prediction model based on users' data. The server can announce the learning project to users through the app and encourage their participation. If a user wants to know more about this project, the app will display an interface to submit the bids and calculate the expected cost. If the user is interested in learning, he/she will download apps, calculate cost and submit the bids through the interface. Once the BS receive all bids in certain time, the BS will start the training process by broadcasting an initial global model to all the winning users. On behalf of the user, the app will download this global model and upload the model updates generated by the training on the user's local data. After finishing the model training project, the BS will give users rewards (e.g., money) based on the bid it wins.

IV. DECIDING MOBILE USERS'S BID

To transmit the local model update to the BS, mobile users need sub-channels and antennas resources. However, given the maximum tolerable time of FL, there is a correlation between resource and corresponding energy cost. In this section, we present the way mobile users decide bids. Specially, for bid Δ_{ni}, mobile user n calculates transmission power p_{ni}, computation resource f_{ni} and cost v_{ni} corresponding to a given sub-channel number b_{ni} and antenna number A_{ni}. However, for simplicity, the process to decide mobile users'

bid is the same for every submitted bids. Thus, we remove the bid index i in this section. The energy cost of mobile user n is defined after user n solve the weighted sum of completion time and total energy consumption in the submitted bid, which is given as

$$\textbf{P1}: \min_{f_n, p_n, A_n, b_n, \varepsilon_n} I_0^n \left(E_n^{tol}(p_n, f_n, \varepsilon_n) + \rho T_n^{tol}(p_n, f_n, \varepsilon_n) \right) \tag{17a}$$

$$s.t.\ I_0^n T_n^{tol}(p_n, f_n, A_n, b_n, \varepsilon_n) \le T_{max}, \tag{17b}$$

$$f_n \in [f_n^{\min}, f_n^{\max}], \tag{17c}$$

$$p_n \in (0, p_n^{\max}], \tag{17d}$$

$$\varepsilon_n \le (0, 1], \tag{17e}$$

$$A_n \in (0, A_n^{\max}], \tag{17f}$$

$$b_n \in (0, b_n^{\max}], \tag{17g}$$

where $f_n^{\max}$ and $p_n^{\max}$ are the maximum local computation capacity and maximum transmit power of mobile user n, respectively. $A_n^{\max}$ and $b_n^{\max}$ are the maximum antenna and maximum sub-channel that mobile user n can request in each bid, respectively. $A_n^{\max}$ and $b_n^{\max}$ are chosen by mobile user n. $I_0^n = \frac{C_1 \log(1/\gamma)}{1 - \varepsilon_n}$ is the lower bound of the number global iterations corresponding to local accuracy ε_n. Note that the cost to the mobile user cannot be the same over iterations. ρ is the weight. However, to make the problem more tractable, we consider minimizing the approximated cost rather than the actual cost, similar to approach in [23], [30]. Constraint (17b) indicates delay requirement of FL task.

According to **P1**, the maximum number of antennas and sub-channels are always energy efficient, i.e., the optimal antenna is $A_n = A_n^{max}, b_n = b_n^{max}$ and $\varepsilon_n^*, p_n^*, f_n^*$ are the optimal solution to:

$$\textbf{P2}: \min_{f_n, p_n, \varepsilon_n} I_0^n \left(E_n^{tol}(p_n, f_n, \varepsilon_n) + \rho T_n^{tol}(p_n, f_n, \varepsilon_n) \right)$$

$$s.t.\ I_0^n T_n^{tol}(p_n, f_n, \varepsilon_n) \le T_{max},$$

$$f_n \in [f_n^{\min}, f_n^{\max}],$$

$$\varepsilon_n \in (0, 1],$$

$$p_n \in (0, p_n^{\max}]. \tag{18}$$

Because of the non convexity of **P2**, it is challenging to obtain the global optimal solution. To overcome the challenge, an iterative algorithm with low complexity is proposed in the following subsection.

A. Iterative Algorithm

The proposed iterative algorithm basically involves two steps in each iteration. To obtain the optimal, we first solve (**P2**) with fixed ε_n, and then ε_n is updated based on the obtained f_n, p_n in the previous step. In the first step, we consider the first case when ε_n is fixed, and **P2** becomes

$$\textbf{P3}: \min_{f_n, p_n, \varepsilon_n} I_0^n \left(E_n^{tol}(p_n, f_n, \varepsilon_n) + \rho T_n^{tol}(p_n, f_n, \varepsilon_n) \right)$$

$$s.t.\ I_0^n T_n^{tol}(p_n, f_n, \varepsilon_n) \le T_{max},$$

$$f_n \in [f_n^{\min}, f_n^{\max}],$$

$$p_n \in (0, p_n^{\max}]. \tag{19}$$

P3 can be decomposed into two sub-problems as follows.

Algorithm 1 Optimal Uplink Power Transmission

1 Calculate $\phi(p_n^{max})$
2 Calculate p_n^{min} so that $I_0^n T_n^{tol}(p_n^{min}) = T_{max}$
3 **if** $\phi(p_n^{max} < 0)$ **then**
4 | $p_n^* = p_n^{max}$
5 **else**
6 | $p_1 = \max(0, p_n^{min})$ and $p_2 = p_n^{max}$
7 | **while** $(p_2 - p_1 \le \epsilon)$ **do**
8 | | $p_u = (p_1 + p_1)/2$
9 | | **if** $\phi(p_u) \le 0$ **then**
11 | | | $p_1 = p_u$
12 | | **else**
14 | | | $p_2 = p_u$
15 | | **end**
16 | **end**
17 | $p_n^* = (p_1 + p_2)/2$
18 **end**

1) Optimization of Uplink Transmission Power: Each mobile user assigns its transmission power by solving the following problem:

$$\textbf{P3a}: \min_{p_n} f(p_n)$$

$$s.t.\ I_0^n T_n^{tol}(p_n, f_n, \varepsilon_n) \le T_{max},$$

$$p_n \in (0, p_n^{\max}],$$

$$f_n, \varepsilon_n\ \text{are given}. \tag{20}$$

where $f(p_n) = \frac{\sigma(1+\rho)p_n}{b_n W \log_2(1 + \frac{(A_n-1)p_n h_n}{b_n W N_0})}$.

Note that $f(p_n)$ is quasiconvex in the domain [31]. A general approach to the quasiconvex optimization problem is the bisection method, which solves a convex feasibility problem each time [32]. However, solving convex feasibility problems by an interior cutting-plane method requires $O(\kappa^2/\alpha^2)$ iterations, where κ is the dimension of the problem [31]. On the other hand, we have

$$f'(p_n) = \frac{\sigma \log_2(1 + \theta_n p_n h_n) + \frac{\sigma(1+\rho)p_n \theta_n h_n}{\ln 2(1 + \theta_n p_n h_n)}}{b_n W (\log(1 + \theta_n p_n h_n))^2}, \tag{21}$$

where $\theta_n = \frac{(A_n - 1)}{W N_0}$. Then, we have

$$\phi(p_n) = \sigma \log_2(1 + \theta_n p_n h_n) + \frac{\sigma(1 + \rho)p_n \theta_n h_n}{\ln 2(1 + \theta_n p_n h_n)} \tag{22}$$

is a monotonically increasing transcendental function and negative at the starting point $p_n = 0$ [31]. Therefore, in order to obtain the optimal power allocation p_n as shown in Algorithm 1, we follow a low-complexity bisection method by calculating $\phi(p_n)$ rather than solving a convex feasibility problem each time.

2) Optimization of CPU Cycle Frequency and Number of Antennas:

$$\textbf{P3b}: \min_{f_n} I_0^n \log\left(\frac{1}{\varepsilon_n}\right) c_n s_n \left(\zeta f_n^2 + \rho/f_n\right)$$

$$s.t.\ I_0^n \left(\log\left(\frac{1}{\varepsilon_n}\right)\frac{c_n s_n}{f_n} + T_n^{com}\right) \le T_{max},$$

$$f_n \in [f_n^{\min}, f_n^{\max}],$$

$$p_n, \varepsilon_n\ \text{are given}. \tag{23}$$

Algorithm 2 Optimal Local Accuracy

---

1 Initialize $\varepsilon_n = \varepsilon_n^{(0)}$, set $j = 0$
2 **repeat**
3   Calculate $\varepsilon_n^* = \frac{\alpha_1}{(\ln 2)\xi^j}$
4   Update $\xi^{(j+1)} = \frac{\gamma_1 \log_2(1/\varepsilon_n) + \gamma_2}{1 - \varepsilon_n}$
5   Set $j = j + 1$
6 **until** $|H(\xi^{(n+1)})|/|H(\xi^{(n)})| < \epsilon_2$;

Algorithm 3 Iterative Algorithm

---

1 Initialize a feasible solution p_n, f_n, ε_n and set $j = 0$.
2 **repeat**
3   With $\varepsilon_n^{(j)}$ obtain the optimal $p_n^{(j+1)}, f_n^{(j+1)}$ of problem **P2** :
4   With $p_n^{(j+1)}, f_n^{(j+1)}$ obtain the optimal $\varepsilon_n^{(j+1)}$ of problem **P2** :
5   Set $j = j + 1$
6 **until** *Objective value of P2 converges*;

P3b is the convex problem, so we can solve it by any convex optimization tool.

In the second step, **P2** can be simplified by using f_n and p_n calculated in the first step as:

$$\textbf{P4}: \quad \min_{\varepsilon_n} \frac{\gamma_1 \log_2(1/\varepsilon_n) + \gamma_2}{1 - \varepsilon_n} \tag{24a}$$

$$s.t. \quad T_n^{tol} \leq T_{max}, \tag{24b}$$

where $\gamma_1 = a\left(E_n^{comp} + T_n^{comp}\right)$ and $\gamma_2 = a\left(E_n^{com} + T_n^{com}\right)$. The constraint (24b) is equivalent to $T_n^{com} \leq \vartheta(\varepsilon_n)$, where $\vartheta(\varepsilon_n) = \frac{1-\varepsilon_n}{m} T_{max} + \frac{c_n s_n \log_2 \varepsilon_n}{f_n}$. We have $\vartheta(\varepsilon_n)'' < 0$, and therefore, $\vartheta(\varepsilon_n)$ is a concave function. Thus, constraint (24b) can be equivalent transformed to $\varepsilon_n^{min} \leq \varepsilon_n \leq \varepsilon_n^{max}$, where $\vartheta(\varepsilon_n^{min}) = \vartheta(\varepsilon_n^{max}) = T_n^{com}$. Therefore, ε_n is the optimal solution to

$$\textbf{P5}: \quad \min_{\varepsilon_n} \frac{\gamma_1 \log_2(1/\varepsilon_n) + \gamma_2}{1 - \varepsilon_n}$$
$$s.t. \quad \varepsilon_n^{min} \leq \varepsilon_n \leq \varepsilon_n^{max}. \tag{25}$$

Obviously, the objective function of **P5** has a fractional in nature, which is generally difficult to solve. According to [13], [33], solving **P5** is equivalent to finding the root of the nonlinear function $H(\xi)$ defined as follows

$$H(\xi) = \min_{\varepsilon_n^{min} \leq \varepsilon_n \leq \varepsilon_n^{max}} \gamma_1 \log_2(1/\varepsilon_n) + \gamma_2 - \xi(1 - \varepsilon_n) \tag{26}$$

Function $H(\xi)$ with fixed ξ is convex. Therefore, the optimal solution ε_n can be obtained by setting the first-order derivative of $H(\xi)$ to zero, which leads to the optimal solution is $\varepsilon_n^* = \frac{\gamma_1}{(\ln 2\xi)}$. Thus, similar to [13], problem **P5** can be solved by using the Dinkelbach method in [33] (shown as Algorithm 2).

B. Convergence Analysis

The algorithm that solves problems **P2** is given in Algorithm 4, which iteratively solves problems **P3** and **P4**. Since the optimal solution of problem **P3** and **P4** is obtained in each step, the objective value of problem **P2** is non-increasing in each step. Moreover, the objective value of problem **P2** is

lower bounded by zero. Thus, Algorithm 4 always converges to a local optimal solution.

C. Complexity Analysis

Because of the non-convexity of **P2**, it is challenging to obtain the global optimal solution. To overcome the challenge, an iterative algorithm with low-complexity is proposed in the following subsection. In particular, to solve the general energy-efficient resource allocation problem **P2** using Algorithm 3, the major complexity in each step lies in solving problems **P3** and **P4**. To solve problem **P3**, the complexity is $\mathcal{O}(L_e \log_2(1/\epsilon_1))$, where ϵ_1 is the accuracy of solving **P3** with the bisection method and L_e is the number of iterations for optimizing f_n and p_n. To solve problem **P4**, the complexity is $\mathcal{O}(\log_2(1/\epsilon_2))$ with accuracy ϵ_2 by using the Dinkelbach method [33]. As a result, the total complexity of the proposed Algorithm 4 is $H_e S$, where H_e is the number of iterations for problems **P3** and **P4** and S is equal to $\mathcal{O}(L_e \log_2(1/\epsilon_1)) + \mathcal{O}(\log_2(1/\epsilon_2))$.

After deciding the bids, the mobile users submit bids to the BS. The following section describes the auction mechanism between the BS and mobile users for selecting winners, allocating bandwidth and deciding on payment.

V. AUCTION MECHANISM BETWEEN BS AND MOBILE USERS

After receiving all bids submitted by mobile users, the BS decides a set of winners by solving the problem (**P6**), aiming to maximize social welfare. The BS's aim is to achieve social welfare because the BS needs to incentive mobile users to participate in learning. Here, the BS's freedom in designing the incentive mechanism is the payment determination, which can force participant mobile users to be truthful. Moreover, if the BS wants to select winners to maximize its utility, the BS needs to know the distribution of mobile users' private information in advance [8], which is assumed to be unavailable in our work. In case the prior distribution of mobile users' private information is not available, worst-case analysis can be applied, but that method could lead to overly pessimistic results [8].

A. Problem Formulation

In bid Δ_{ni} that mobile user n submits to the BS includes the number of subchannels b_{ni}, the number of antennas A_{ni}, local accuracy ϵ_{ni}, and claimed cost v_{ni}. The utility of one bid is the difference between the payment g_{ni} and the real cost V_{ni}.

$$U_{ni} = \begin{cases} g_{ni} - V_{ni}, & \text{if bid } \Delta_{ni} \text{ wins,} \\ 0, & \text{otherwise.} \end{cases} \tag{27}$$

The payment that the BS pays for winning bids is $\sum_{n,i} g_{ni}$. As we described in Section III-A, high local accuracy will significantly improve the global accuracy for a fixed number of global iterations. The utility of the BS is the difference between the BS's satisfaction level and the payment for mobile users. The satisfaction level of the BS to bid Δ_{ni} is measured

based on the local accuracy that mobile user n can provide in the ith bid and is defined as follows

$$\chi_{ni} = \frac{\tau}{\varepsilon_{ni}}. \tag{28}$$

Thus, the total utilities of the system or the social welfare is

$$\sum_{n,i} (\chi_{ni} - v_{ni}) x_{ni}. \tag{29}$$

If mobile users truthfully submit their cost, $V_{ni} = v_{ni}$, we have the social welfare maximization problem defined as follows:

$$\mathbf{P6} : \max_x \sum_{n,i} (\chi_{ni} - v_{ni}) x_{ni} \tag{30a}$$

$$s.t. \sum_n x_{ni} b_{ni} \leq B_{max}, \tag{30b}$$

$$\sum_n x_{ni} A_{ni} \leq A_{max}, \tag{30c}$$

$$\sum_i x_{ni} \leq 1, \forall n, \tag{30d}$$

$$x_{ni} = \{0, 1\}, \tag{30e}$$

where (30b) and (30c) indicate the bandwidth resource (i.e., sub-channels) and the antennas limitation constraints of the BS, respectively. Then, (30d) shows that a mobile user can win at most one bid and (30e) is the binary constraint that presents whether bid Δ_{ni} wins or not.

Problem **P6** is a minimization knapsack problem, which is known to be NP-hard. This implies that no algorithm is able to find out the optimal solution of **P6** in polynomial time. It is also known that a mechanism with Vickrey-Clarke-Groves (VCG) payment rule is truthful only when the resource allocation is optimal. Hence, using VCG payment directly is unsuitable due to the problem **P6** is computationally intractable. To deal with the NP-hard problem, we proposed the primal-dual based greedy algorithm. The following economic properties are desired.

Truthfulness: An auction mechanism is truthful if and only if for every bidder n can get the highest utility when it reports true value.

Individual Rational: If each mobile user reports its true information (i.e., cost and local accuracy), the utility for each bid is nonnegative, i.e., $U_{ni} \geq 0$.

Computation Efficiency: The problem can be solved in polynomial time.

Among these three properties, truthfulness is the most challenging one to achieve. In order to design a truthful auction mechanism, we introduce the following definitions.

Definition 1: (Monotonicity): If mobile user n wins with the bid $\Delta_{ni} = \{v_{ni}, 1/\varepsilon_{ni}, b_{ni}, A_{ni}\}$, then mobile user n can win the bid with $\Delta_{nj} = \{v_{nj}, 1/\varepsilon_{nj}, b_{nj}, A_{nj}\} \succ \Delta_{ni} = \{v_{ni}, 1/\varepsilon_{ni}, b_{ni}, A_{ni}\}$.

The notation $\succ$ denotes the preference over bid pairs. Specifically, $\Delta_{nj} = \{v_{nj}, 1/\varepsilon_{nj}, b_{nj}, A_{nj}\} \succ \Delta_{ni} = \{v_{ni}, 1/\varepsilon_{ni}, b_{ni}, A_{ni}\}$ if $\varepsilon_{nj} < \varepsilon_{ni}$ for $v_{nj} = v_{ni}, b_{nj} = b_{ni}, A_{nj} = A_{ni}$ or $v_{nj} < v_{ni}, b_{nj} < b_{ni}, A_{nj} < A_{ni}$ for $\varepsilon_{nj} = \varepsilon_{ni}$. The monotonicity implies that the chance to obtain a required bundle of resources can only be enhanced by either increasing the local accuracy or decreasing the amount of resources required or decreasing the cost.

Definition 2: (Critical Value): For a given monotone allocation scheme, there exists a critical value c_{ni} of each bid Δ_{ni} such that $\forall n, i(\chi_{ni} - v_{ni}) \geq c_{ni}$ will be a winning bid, while $\forall n, i(\chi_{ni} - v_{ni}) < c_{ni}$ is a losing bid.

In our proposed mechanism, the difference between the satisfaction based on local accuracy and cost of one bid can be considered as the value of that bid. Therefore, the critical value can be seen as the minimum value that one bidder has to bid to obtain the requested bundle of resources. With the concepts of monotonicity and critical value, we have the following lemma.

Lemma 1: An auction mechanism is truthful if the allocation scheme is monotone and each winning mobile user is paid the amount that equals to the difference between the satisfaction based on the local accuracy and the critical value.

Proof: Similar Lemma 1 and Theorem 1 in [11]. ∎

In the next subsection, we propose a primal-dual greedy approximation algorithm for solving problem **P6**. The algorithm iteratively updates both primal and dual variables and the approximation analysis is based on duality property. As the result, we firstly relax $1 \geq x_{ni} \geq 0$ of **P6** to have the linear programming relaxation (LPR) of **P6**. Then, we introduce the dual variable vectors $\boldsymbol{y}$, $\boldsymbol{z}$ and $\boldsymbol{t}$ corresponding to constraints (30b), (30c) and (30d) and we have the dual of problem LPR of **P6** can be written as

$$\mathbf{P7} : \max_{\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}} \sum_{n \in \mathcal{N}} y_n + z B_{max} + t A_{max} \tag{31a}$$

$$s.t. \ y_n + z A_{ni} + t B_{ni} \geq q_{ni}, \quad \forall n, i, \tag{31b}$$

$$y_n \geq 0, \quad \forall n, \tag{31c}$$

$$z, t \geq 0. \tag{31d}$$

In Section V-B, we devise an greedy approximation algorithm and Section V-C, a theoretical bound is achieved for the approximation ratio of the proposed algorithm.

B. Approximation Algorithm Design

In this section, we use a greedy algorithm to solve problem **P6** The main idea of the greedy algorithm is to allocate the resource to bidders with the larger normalized value. The winner selection process is described in the Algorithm 4. The process consists 3 steps:

Step 1: Based on the bid's value and the weighted sum of requested resources, each bid Δ_{ni} calculates the normalized value. The bid's value is defined as the difference between the satisfaction level of the BS and the cost declared in this bid, $q_{ni} = \chi_{ni} - v_{ni}$. The weighted sum of different types of resources declared in this bid is defined as $s_n^i = \eta_b B_n + \eta_a A_n$, where η_b, η_a are the weights. The normalized value of the bid is defined as the ration between the value of this bid and the weighted sum of requested resources, and is denoted as

$$\bar{q}_{ni} = \frac{q_{ni}}{s_{ni}}.$$

Step 2: The bid with maximum $\bar{q}_{ni}$ wins the bidding.

Step 3: Delete user n from the list of bidders. Then go back to Step 2 until either one of the following termination conditions is satisfied:

Algorithm 4 The Greedy Approximation Algorithm

---

1 **Input:** $(B, A, \chi, v, B_{max}, A_{max})$
2 **Output:** solution **x**
3 $\mathcal{U} = \varnothing$, $\mathbf{x} = \mathbf{0}$
4 $\forall n : y_n = 0, \psi = 0$;
5 $\varphi = 0, B = 0, A = 0$;
6 $s_n^i = \eta_b B_{ni} + \eta_a A_{ni}$;
7 $q_{kj} = \chi_{ni} - v_{ni}$;
8 **for** $n \in \mathcal{N}$ **do**
9 $\quad$ $i_n = \arg\max_i\{q_{ni}\}$;
10 **end**
11 $\kappa = \max \frac{s_{ni}}{s_{ni'}}$;
12 **while** $\mathcal{N} \neq \varnothing$ **do**
13 $\quad$ $\mu = \arg\max_{n\in\mathcal{N}} \frac{q_{ni}}{s_{ni_n}}$;
14 $\quad$ **if** $B + b_{\mu i_\mu} <= B_{max}$ and $A + a_{\mu i_\mu} <= A_{max}$ **then**
15 $\quad\quad$ $x_{\mu i_\mu} = 1$; $\quad y_\mu = q_{\mu i_\mu}$;
16 $\quad\quad$ $\varphi = \varphi + q_{\mu i_\mu}$;
17 $\quad\quad$ $\psi = \frac{\sum_{n\in\mathcal{U}} q_{ni_n}}{\sum_{n\in\mathcal{U}} s_{ni_n}}$;
18 $\quad\quad$ $\mathcal{U} = \mathcal{U} \cup \{\mu\}$ and $\mathcal{N} = \mathcal{N} \setminus \{\mu\}$
19 $\quad$ **else**
20 $\quad\quad$ break;
21 $\quad$ **end**
22 **end**
23 $\bar{\psi} = \kappa\psi$;
24 $z = \eta_b\bar{\psi}$, $\quad t = \eta_a\bar{\psi}$

---

i) The BS has not enough resource to satisfy the demand;
ii) All the mobile users win one bid.

C. Approximation Ratio Analysis

In this subsection, we analyze approximation ratio of Algorithm 4. Our approach is to use the duality property to derive a bound for approximation algorithm. We denote the optimal solution and the optimal value of LPR of **P6** as x_{ni}^* and OP_f. Furthermore, let OP and φ as the optimal value of **P6** and the primal value of **P6** obtained by Algorithm 4. Our analysis consists of two steps. First, Theorem 1 shows that Algorithm 4 generates a feasible solution to **P7**, and Proposition 1 provides approximation factor.

Theorem 1: Algorithm 4 provides a feasible solution to **P7**.
$\quad$ *Proof:* We discuss the following three cases:

- Case 1: mobile user μ wins, i.e., $\mu \in \mathcal{U}$ and $b_{\mu i_\mu} = \max_{i'\in\mathcal{I}_\mu}\{q_{\mu i'}\}$. Then we have $y_\mu = q_{\mu i_\mu} \geq q_{\mu i'}, \forall i' \in \mathcal{I}_\mu$. Thus, constraint (31b) is satisfied for all mobile users in $\mathcal{U}$.

- Case 2: mobile user μ loses the auction, i.e., $\mu \in \mathcal{N} \setminus \mathcal{U}$. According to the while loop, it is evident that

$$\frac{q_{ni_n}}{s_{ni_n}} > \frac{q_{\mu i_\mu}}{s_{\mu i_\mu}}, \quad \forall n \in \mathcal{U}.$$

Therefore, $\psi > \frac{q_{\mu i_\mu}}{s_{\mu i_\mu}}$. Thus,

$$\bar{\psi} \geq \kappa \frac{q_{\mu i_\mu}}{s_{\mu i_\mu}} \geq \frac{q_{\mu i_\mu}}{s_{\mu i_\mu}}.$$

In addition, we have

$$q_{\mu i_\mu} \geq q_{\mu i'} \quad \text{and} \quad \kappa > \frac{s_{\mu i_\mu}}{s_{\mu i'}}, \quad \forall i' \neq i_\mu.$$

Therefore,

$$\bar{\psi} \geq \frac{q_{\mu i'}}{n_{\mu i'}}, \quad \forall i' \neq i_\mu.$$

Therefore, we have

$$\eta_b\bar{\psi}B_{in} + \eta_a\bar{\psi}A_{in} \geq q_{in}, \quad \forall i' \neq i_\mu.$$

or

$$zC_{in} + tA_{in} \geq q_{in}, \quad \forall i' \neq i_\mu.$$

Therefore, constraint (31b) is also satisfied for all mobile users in $\mathcal{N} \setminus \mathcal{U}$.
$\quad\blacksquare$

Proposition 1: The upper bound of integrality gap α between **P6** and its relaxation and the approximation ratio of Algorithm 4 are $1 + \frac{\kappa\Upsilon}{\Upsilon - S}$, where $\Upsilon = \eta_b B_{max} + \eta_a A_{max}$, $S = \max_{n,i} s_{ni}$.

Proof: Let OP and OP_f be the optimal solution for **P6** and LPR of **P6**. We can obtain the following:

$$OP \leq OP_f \leq \sum_{n=1}^{N} y_n + zB_{max} + tA_{max}$$

$$\leq \sum_{n=1}^{N} y_n + \bar{\psi}(\eta_b B_{max} + \eta_a A_{max})$$

$$\leq \sum_{n\in\mathcal{N}} q_{ni_n} + \bar{\psi}(\eta_b B_{max} + \eta_a A_{max})$$

$$\leq \left(\sum_{n\in\mathcal{N}} q_{ni_n}\right)\left(1 + \frac{(\eta_b B_{max} + \eta_a A_{max})\kappa}{\eta_b B_{max} + \eta_a A_{max} - S}\right)$$

$$\leq \varphi\left(1 + \frac{\Upsilon\kappa}{\Upsilon - S}\right),$$

Therefore, the integrality α is given as

$$OP_f/OP \leq OP_f/\varphi$$
$$\leq \left(1 + \frac{\kappa\Upsilon}{\Upsilon - S}\right).$$

The approximation ratio is

$$OP/\varphi \leq OP_f/\varphi \leq \left(1 + \frac{\kappa\Upsilon}{\Upsilon - S}\right).$$

$\quad\blacksquare$

D. Payment

Then we will find the critical value which is the minimum value a bidder has to bid to win the requested bundle of resources. In this paper, we consider the bid combinations submitted by mobile user n as the combinations of bid submit by virtual bidders, in which each virtual bidder can submit one bid. Therefore, the number of virtual bidders corresponding to mobile user n is equal to the number of bids I_n that mobile user n submits. Denote by m the losing mobile user with the highest normalized value if mobile user n is not participating in the auction. Accordingly, the minimum value mobile user n needs to place is $\frac{q_{mi_m}}{s_{mi_m}}s_{ni_n}$, where i_m and i_n are the indexes of highest normalized value bids of mobile user m and n, respectively. Thus, the payment of winning mobile user n in the pricing scheme is $g_{ni_n} = \chi_{ni_n} - \frac{q_{mi_m}}{s_{mi_m}}s_{ni_n}$.
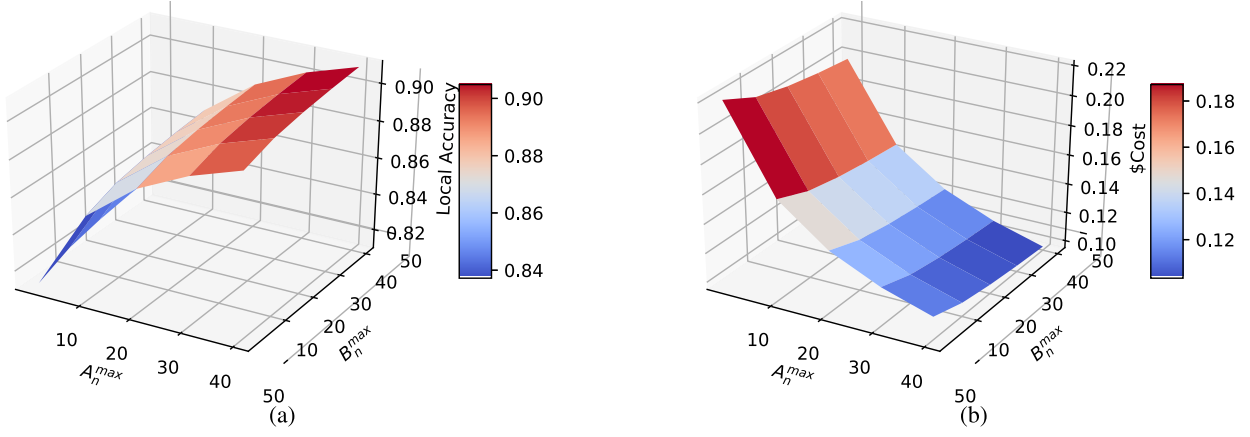
Fig. 2. Numerical results a) The changing of the local accuracy when the maximum number sub-channels and antennas in one bid vary, b) The changing of the energy cost when the maximum number sub-channels and antennas in one bid vary.

E. Properties

Now, we show that the winner determination algorithm is monotone and the payment determined for a winner mobile user is the difference between the local accuracy based satisfaction and the critical value of its bid. From line 13 of the Algorithm 4, it is clear that a mobile user can increase its chance of winning by increasing its bid. Also, a mobile user can increase its chance to win by decreasing the weighted sum of the resources. Therefore, the winner determination algorithm is monotone with respect to mobile user's bids. Moreover, the value of a winning bidder is equals to the minimum value it has to bid to win its bundle, i.e., its critical value. This is done by finding the losing bidder m who would win if bidder n would not participate in the auction. Thus, the proposed mechanism has a monotone allocation algorithm and payment for the winning bidder equals to the difference between the local accuracy based satisfaction and the critical value of its bid. We conclude that proposed mechanism is a truthful mechanism according to Lemma 1.

Next, we prove that the proposed auction mechanism is individually rational. For any mobile user n bidding its true value, we consider two possible cases:

- If mobile user n is a winner with its bid ith, its payment is

$$
\begin{aligned}
U_{ni} &= g_{ni} - v_{ni} \\
&= (\chi_{ni} - \frac{q_{mi_m}}{s_{mi_m}} s_{ni} - v_{ni} \\
&= \left(\frac{\chi_{ni} - v_{ni}}{s_{ni}} - \frac{q_{mi_m}}{s_{mi_m}} \right) s_{ni} \\
&= \left(\frac{q_{ni}}{s_{ni}} - \frac{q_{mi_m}}{s_{mi_m}} \right) s_{ni} \geq 0
\end{aligned}
$$

where m the losing bidder with the highest normalized valuation if n does not participate in the auction and the last inequality follows from Algorithm 4.

- If mobile user n is not a winner. Its utility is 0.

Therefore, the proposed auction mechanism is individually rational.

Finally, we show that the proposed auction mechanism is computationally efficient. We can see that in Algorithm 4,

the while-loop (lines 12-22) takes at most N times, linear to input. Calculating the payment takes at most $N(N-1)$ times. Therefore, the proposed auction mechanism is computationally efficient. Therefore, the time complexity of Algorithm 4 is $\mathcal{O}(N^2)$.

VI. SIMULATION RESULTS

In this section, we provide some simulation results to evaluate the proposed mechanism. The parameters for the simulation are set the following. The required CPU cycles for performing a data sample c_n is uniformly distributed between $[10, 50]$ cycles/bit [13]. The size of data samples of each mobile user is $s_n = 80 \times 10^6$. The effective switched capacitance in local computation is $\xi = 10^{-28}$ [13]. We assume that the noise power spectral density level N_0 is -174dBm/Hz, the sub-channel bandwidth is $W = 15$ kHz and the channel gain is uniformly distributed between $[-90, -95]$ dB 27. In addition, the maximum and minimum transmit power of each mobile user is uniformly distributed between $[6, 10]$ mW and between $[0, 2]$ mW, respectively. The maximum and minimum computation capacity is uniformly distributed between $[3, 5]$ GHz and between $[10, 20]$ Hz, respectively. We also assume that the total number of sub-channels and antennas of the BS are 100 and 100, respectively. Firstly, we use the iterative Algorithm 3 to perform the characteristic of evaluating bids when $\rho = 1$. The maximum number of sub-channels B_n^{max} and antennas A_n^{max} for mobile user n to request in each bid vary from 10 to 50. Fig. 2a shows the accuracy level that mobile user n requires to provide increases when the maximum number of sub-channels B_n^{max} and antennas A_n^{max} increase. In particular, when the sub-channels and antennas are both 50, the local accuracy 0.92 while when the sub-channels and antennas are both 10, the local accuracy 0.81. This is because the transmission time and transmission cost decrease when wireless resources increase. It requires less global round to satisfy the learning task performance. Therefore, the local accuracy increases or. As shown in Fig. 2b, the energy cost decreases when the number of sub-channels and antennas increases. This is because mobile user n can keep low contributing CPU cycle frequency and transmission rate while guaranteeing the delay constraint.
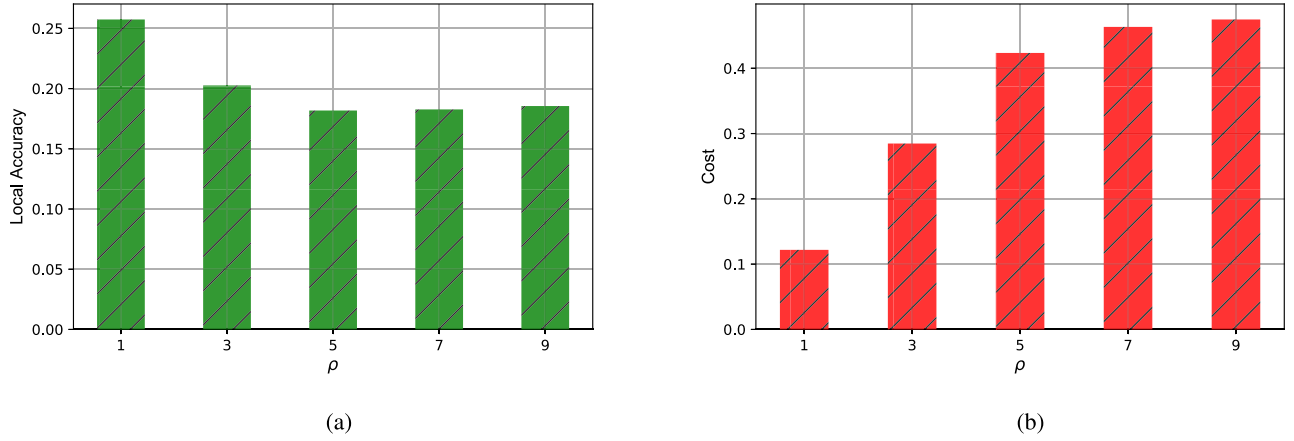
Fig. 3. Numerical results a) Local accuracy v.s. ρ b) Energy cost v.s. ρ.
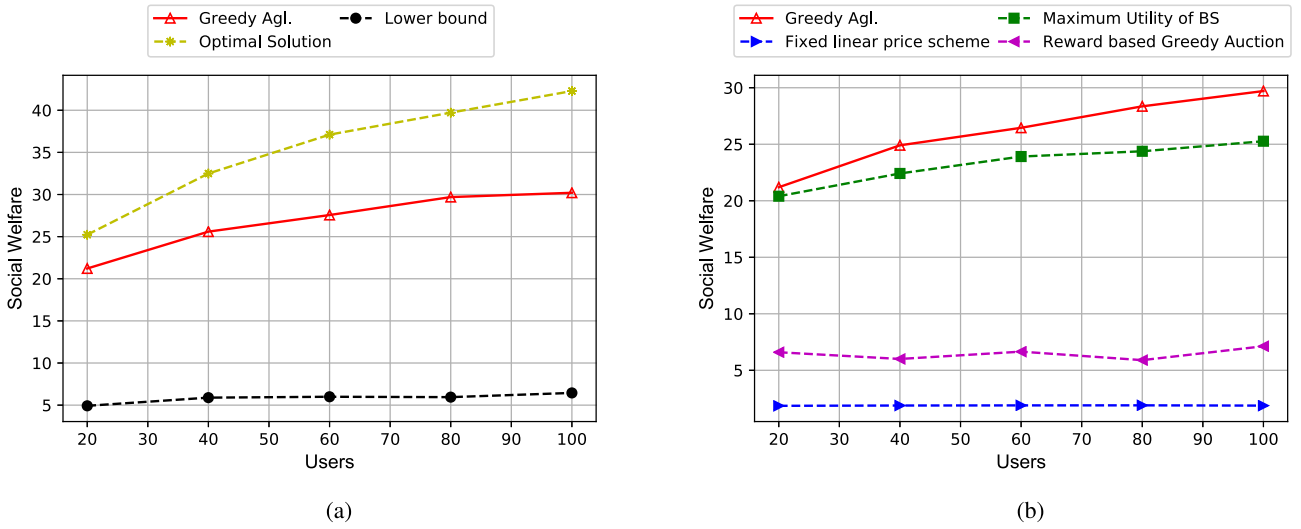

Fig. 4. Numerical results for social welfare a) three schemes: Optimal solution, proposed greedy algorithm and lower bound b) four schemes: proposed greedy algorithm, Reward-based greedy auction, Maximum utility of the BS and fixed price scheme.

Fig. 3a and Fig. 3b present the cost of one bid of the mobile user and local accuracy, respectively, when the weight ρ varies from 1 to 9. As shown in Fig. 3a and Fig. 3b, when ρ increases, the local accuracy decreases and the energy cost increases. This is because when ρ increases, the objective focuses more on minimizing the time completion of one global round. It requires more computation resource as well as better quality of data (low local accuracy).

In the following, we evaluate the performance of the proposed auction algorithm. To compare with the proposed algorithm, we use four baselines:

- Optimal Solution: **P6** is solved optimally.
- Fixed Price Scheme [34]: In this scheme, price vector $f = \{f_b, f_a\}$ is the price mobile users need to pay for the resource. In this scheme, the mobile users are served in a first-come, first-served basic until the resources are exhausted. The mobile user can get the resource when the valuation of mobile user's bid is at least $F_{ni} = B_{ni}f_b + A_{ni}f_a$ which is the sum of the fixed price of each resource in their bid. We consider three kinds of price vector: linear

price ($f_i = f_o \times \eta_i, i = a, b$), sub-linear price vector ($f_i = f_o \times \eta_i^{0.85}, i = a, b$) [34], and a super-linear price vector ($f_i = f_o \times \eta_i^{1.15}, i = a, b$) [34]. Here, we call f_o as the basic price. Unless specified otherwise, we choose $f_o = 0.01$.

- Reward-based greedy auction [35].
- Maximum utility of the BS.

Fig. 4a reports the performance of the optimal solution, the lower bound, and the proposed greedy scheme. The lower bound is determined by the fractional optimal solution divided by gap when the number of mobile users varies from 20 to 100 with the step size of 20. We note that with the number of mobile users increasing, all schemes produce higher social welfare. This is because there is more chances to choose winning bids with the higher value. Although the social welfare obtained through the proposed greedy scheme is lower than through optimal solution and much higher than the lower bound.

Fig. 4b shows the social cost achieved by the proposed greedy scheme, maximum utility of the BS, reward based
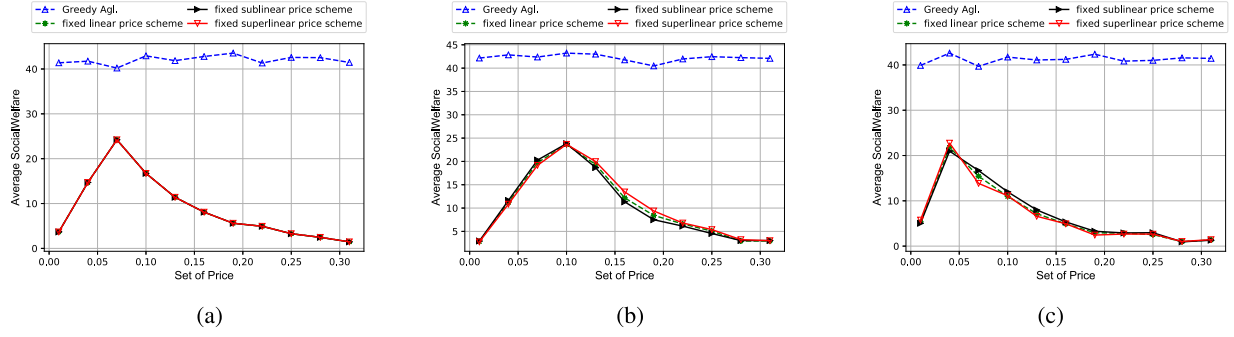
Fig. 5. Numerical results for social welfare when a) $\eta_a = 1, \eta_b = 0.5$, b) $\eta_a = 1, \eta_b = 1$, c) $\eta_a = 1, \eta_b = 2$.
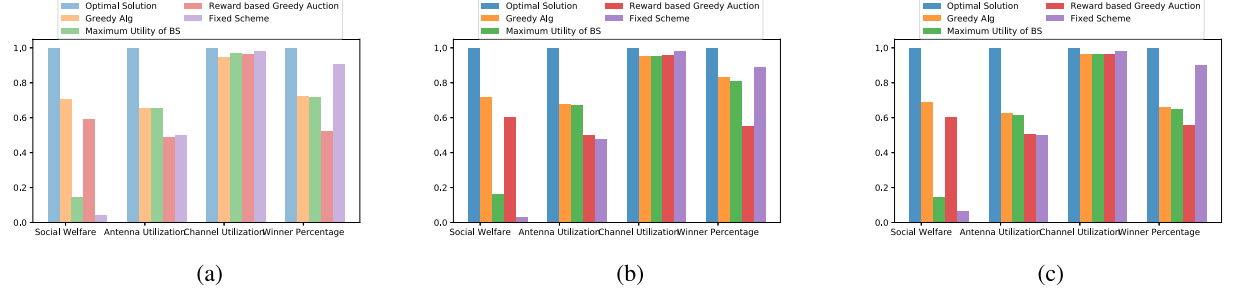

Fig. 6. Numerical results for normalized ratio when a) $\eta_a = 1, \eta_b = 0.5$, b) $\eta_a = 1, \eta_b = 1$, c) $\eta_a = 1, \eta_b = 2$.

greedy auction and fixed linear scheme when the number of mobile users varies 20 to 100 with the step size of 20. We can see that the proposed greedy scheme can provide the much higher social welfare than the baseline. When the number of users is 100, the social welfare obtained by proposed greedy algorithm is approximately 16% higher than the one obtained by the maximum utility of BS. The result is that our proposed algorithm focus on maximizing the social welfare. In addition, when the number of users is 100, the social welfare obtained by the proposed greedy algorithm is approximately 4 times and 15 times higher than the one obtained by the reward based Greedy Auction and fixed linear price scheme, which ignore the wireless resource limitation when deciding the winning bids.

Since the fixed price scheme heavily depends on the prices of resources, the next experiment helps us to decide whether the fixed-price vector or the performance of the proposed mechanisms is better when we change the basic price f_o between $[0.01, 0.31]$ with the step is 0.03. Fig. 5a, Fig. 5b and Fig. 5c show that the social welfare of fixed price firstly increases and then decreases and equal to 0 when the initial price increases. This is because when the basic price becomes too high, the sum of the price is higher than the valuation of the resources claimed in a bid. Moreover, the social welfare achieved by linear, sublinear and superlinear price schemes are lower than by the proposed greedy scheme. This proves our proposed auction scheme outperforms the fixed price scheme.

In Fig. 6a, Fig. 6b and Fig. 6c, we observe the metrics: social welfare, resource utilization and percentage of five schemes: greedy proposed scheme and fixed price schemes with other baselines. We perform in terms of the ratio with proposed greedy scheme. Among these schemes, the optimal solution is the highest in terms of all metrics. Compared with the proposed scheme, the fixed price can utilize more resources

and more mobile users but provides less social welfare. This is due to the fact that the fixed price mechanism heavily depends on the prices of the resources. In addition, the resource utilization of our proposed scheme is competitive to the one of maximum utility of the BS scheme and the reward based greedy auction scheme. Furthermore, the proposed scheme provides more social welfare.

VII. CONCLUSION

This paper focus on the incentive mechanism design to stimulate mobile users to participate in FL. We formulated the incentive problem between the BS and mobile users in the FL service market as the auction game with the objective of maximizing social welfare. Then, we presented the method for mobile users to decide the bids submitted to the BS so that mobile users can minimize the energy cost. We also proposed the iterative algorithm with low complexity. In addition, we proposed a primal-dual greedy algorithm to tackle the NP-hard winner selection problem. Finally, we showed that the proposed auction mechanism guarantee truthfulness, individual rationality and computation efficiency. Simulation results demonstrated the effectiveness of the proposed mechanism where social welfare obtained by our proposed mechanism is 400% larger than by the fixed price scheme. The model in our work can be extended to multi BS when users are one the large area. One BS can not cover the whole area. In that case, one BS perform edge aggregations of local models which are transmitted from devices in proximity. When each of BS achieves a given learning accuracy, updated models at the edge are transmitted to the cloud or MBS for global aggregation. Intuitively, this hierarchical model can help to reduce significant communication overhead between device users and the cloud via edge model aggregations and reduce the latency. In addition, through the coordination by the

edge servers in proximity, more efficient communication and computation resource allocation among device users can be achieved. Moreover, we can consider the hierarchical auction mechanism consisting of two hierarchical auction models. i.e. a single-seller multiple-buyer model where the lower stage is between BS and mobile users and the higher stage is between the cloud and BSs. Another direction is the case which there are many BSs from different organizers who are interested in using the data from the set of users to train similar types of machine learning models. In that situation, there may be competition of BSs. This will make BSs' decision making different from our work. Therefore, we can also consider it as the future work.

REFERENCES

[1] *Smartphone Shipments Expected to Further Decline in 2018 Before Returning to Growth in 2019*. Accessed: Dec. 2019. [Online]. Available: https://www.idc.com/getdoc.jsp?containerId=prUS43773018

[2] S. R. Pandey, N. H. Tran, M. Bennis, Y. K. Tun, Z. Han, and C. S. Hong, "Incentivize to build: A crowdsourcing framework for federated learning," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[3] K. Bonawitz *et al.*, "Towards federated learning at scale: System design," 2019, *arXiv:1902.01046*. [Online]. Available: http://arxiv.org/abs/1902.01046

[4] T. H. T. Le, N. H. Tran, Y. K. Tun, Z. Han, and C. S. Hong, "Auction based incentive design for efficient federated learning in cellular wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[5] N. H. Tran, W. Bao, A. Zomaya, M. N. H. Nguyen, and C. S. Hong, "Federated learning over wireless networks: Optimization model design and analysis," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Paris, France, Apr. 2019, pp. 1387–1395.

[6] L. U. Khan, M. Alsenwi, Z. Han, and C. S. Hong, "Self organizing federated learning over wireless networks: A socially aware clustering approach," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Barcelona, Spain, Jan. 2020, pp. 453–458.

[7] D. Chen *et al.*, "Federated learning based mobile edge computing for augmented reality applications," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Big Island, HI, USA, Feb. 2020, pp. 767–773.

[8] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[9] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[10] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10700–10714, Dec. 2019.

[11] J. Wang, D. Yang, J. Tang, and M. C. Gursoy, "Enabling radio-as-a-service with truthful auction mechanisms," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2340–2349, Apr. 2017.

[12] P. Klemperer, "What really matters in auction design," *J. Econ. Perspect.*, vol. 16, no. 1, pp. 169–189, Feb. 2002.

[13] Z. Yang, M. Chen, W. Saad, C. Seon Hong, and M. Shikh-Bahaei, "Energy efficient federated learning over wireless communication networks," 2019, *arXiv:1911.02417*. [Online]. Available: http://arxiv.org/abs/1911.02417

[14] M. Chen, Z. Yang, W. Saad, C. Yin, H. Vincent Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," 2019, *arXiv:1909.07972*. [Online]. Available: http://arxiv.org/abs/1909.07972

[15] Q. Zeng, Y. Du, K. K. Leung, and K. Huang, "Energy-efficient radio resource allocation for federated edge learning," 2019, *arXiv:1907.06040*. [Online]. Available: http://arxiv.org/abs/1907.06040

[16] J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang, and M. Guizani, "Reliable federated learning for mobile networks," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 72–80, Apr. 2020.

[17] T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019, pp. 1–7.

[18] T. Li, M. Sanjabi, A. Beirami, and V. Smith, "Fair resource allocation in federated learning," 2019, *arXiv:1905.10497*. [Online]. Available: http://arxiv.org/abs/1905.10497

[19] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning," *IEEE Netw.*, vol. 33, no. 5, pp. 156–165, Sep. 2019.

[20] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2022–2035, Mar. 2020.

[21] S. Feng, D. Niyato, P. Wang, D. I. Kim, and Y.-C. Liang, "Joint service pricing and cooperative relay communication for federated learning," in *Proc. Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Phys. Social Comput. (CPSCom) IEEE Smart Data (SmartData)*, Atlanta, GA, USA, Jul. 2019, pp. 815–820.

[22] W. Y. B. Lim *et al.*, "Hierarchical incentive mechanism design for federated machine learning in mobile networks," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9575–9588, Oct. 2020.

[23] S. R. Pandey, N. H. Tran, M. Bennis, Y. K. Tun, A. Manzoor, and C. S. Hong, "A crowdsourcing framework for on-device federated learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3241–3256, May 2020.

[24] Y. Zhan, P. Li, Z. Qu, D. Zeng, and S. Guo, "A learning-based incentive mechanism for federated learning," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6360–6368, Jul. 2020.

[25] R. Zeng, S. Zhang, J. Wang, and X. Chu, "FMore: An incentive scheme of multi-dimensional auction for federated learning in MEC," 2020, *arXiv:2002.09699*. [Online]. Available: http://arxiv.org/abs/2002.09699

[26] J. Konečný, H. Brendan McMahan, D. Ramage, and P. Richtárik, "Federated optimization: Distributed machine learning for on-device intelligence," 2016, *arXiv:1610.02527*. [Online]. Available: http://arxiv.org/abs/1610.02527

[27] Y. Jiao, P. Wang, D. Niyato, B. Lin, and D. I. Kim, "Toward an automated auction framework for wireless federated learning services market," *IEEE Trans. Mobile Comput.*, early access, May 14, 2020, doi: 10.1109/TMC.2020.2994639.

[28] Y. Hao, Q. Ni, H. Li, and S. Hou, "Energy-efficient multi-user mobile-edge computation offloading in massive MIMO enabled HetNets," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019, pp. 1–6.

[29] H. Quoc Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[30] S. Wang *et al.*, "When edge meets learning: Adaptive control for resource-constrained distributed machine learning," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Honolulu, HI, USA, Apr. 2018, pp. 63–71.

[31] X. Lyu, H. Tian, C. Sengul, and P. Zhang, "Multiuser joint task offloading and resource optimization in proximate clouds," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3435–3447, Apr. 2017.

[32] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[33] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[34] S. Zaman and D. Grosu, "Combinatorial auction-based allocation of virtual machine instances in clouds," *J. Parallel Distrib. Comput.*, vol. 73, no. 4, pp. 495–508, Apr. 2013.

[35] A.-L. Jin, W. Song, P. Wang, D. Niyato, and P. Ju, "Auction mechanisms toward efficient resource sharing for cloudlets in mobile cloud computing," *IEEE Trans. Services Comput.*, vol. 9, no. 6, pp. 895–909, Nov. 2016.
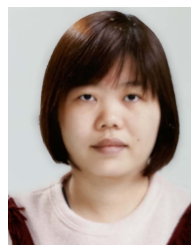
Tra Huong Thi Le (Member, IEEE) received the B.S. and M.S. degrees in electric and electronics engineering from the Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam, in 2010 and 2012, respectively, and the Ph.D. degree from the Department of Computer Science and Engineering, Kyung Hee University, in 2021. Her research interests include resource management, game theory, and machine learning.

Nguyen H. Tran (Senior Member, IEEE) received the B.S. degree from the HCMC University of Technology and the Ph.D. degree from Kyung Hee University in 2005 and 2011, respectively, all in electrical and computer engineering. He was an Assistant Professor with the Department of Computer Science and Engineering, Kyung Hee University, from 2012 to 2017. Since 2018, he has been with the School of Computer Science, The University of Sydney, where he is currently a Senior Lecturer. His research interests include distributed computing, machine learning, and networking. He received the Best KHU Thesis Award in engineering in 2011 and several best paper awards, including IEEE ICC 2016 and ACM MSWiM 2019. He receives the Korea NRF Funding for Basic Science and Research 2016–2023 and ARC Discovery Project 2020–2023. He was the Editor of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING from 2016 to 2020, and the Associate Editor of IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS 2020/2021 in the area of distributed machine learning/federated learning.

Yan Kyaw Tun received the B.E. degree in marine electrical systems and electronics engineering from Myanmar Maritime University, Thanlyin, Myanmar, in 2014, and the Ph.D. degree in computer science and engineering from Kyung Hee University, South Korea, in 2021. He is currently working as a Post-Doctoral Researcher with Kyung Hee University. His research interests include network economics, game theory, network optimization, wireless communication, wireless network virtualization, mobile edge computing, and wireless resource slicing for 5G.

Minh N. H. Nguyen (Member, IEEE) received the B.E. degree in computer science and engineering from the Ho Chi Minh City University of Technology, Vietnam, in 2013, and the Ph.D. degree in computer science and engineering from Kyung Hee University, Seoul, South Korea, in 2020. He is currently working as a Lecturer with The University of Da Nang, Vietnam, and a Post-Doctoral Researcher with the Networking Intelligence Laboratory, Kyung Hee University. His research interests include wireless communications, mobile edge computing, federated learning, and distributed machine learning. He received the Best KHU Ph.D. Thesis Award in engineering in 2020.

Shashi Raj Pandey received the B.E. degree in electrical and electronics with a specialization in communication from Kathmandu University, Nepal, in 2013. He is currently pursuing the Ph.D. degree in computer science and engineering with Kyung Hee University, South Korea. After graduation, he has served as a Network Engineer for Huawei Technologies Nepal Company Pvt. Ltd., Nepal, from 2013 to 2016. His research interests include network economics, game theory, wireless communications, edge computing, and distributed machine learning.

Zhu Han (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University in 1997 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, MD, USA, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer with JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, Boise, ID, USA. He is currently a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department, University of Houston, Houston, TX, USA, and the Computer Science Department, University of Houston. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the Best Paper Award for the *EURASIP Journal on Advances in Signal Processing* in 2015, IEEE Leonard G. Abraham Prize in the field of communications systems [Best Paper Award in IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC)] in 2016, and several best paper awards in IEEE conferences. He was/has been an IEEE Communications Society Distinguished Lecturer from 2015 to 2018, an AAAS Fellow since 2019, and an ACM Distinguished Member since 2019. He has been 1% highly cited researcher since 2017 according to Web of Science. He is also the winner of 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks."

Choong Seon Hong (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Kyung Hee University, Seoul, South Korea, in 1983 and 1985, respectively, and the Ph.D. degree from Keio University, Tokyo, Japan, in 1997.

In 1988, he joined KT, South Korea, where he was involved in broadband networks as a Member of the Technical Staff. Since 1993, he has been with Keio University. He was with the Telecommunications Network Laboratory, KT, as a Senior Member of Technical Staff and as the Director of the Networking Research Team until 1999. Since 1999, he has been a Professor with the Department of Computer Science and Engineering, Kyung Hee University. His research interests include future Internet, intelligent edge computing, network management, and network security. He is also a member of the Association for Computing Machinery (ACM), the Institute of Electronics, Information and Communication Engineers (IEICE), the Information Processing Society of Japan (IPSJ), the Korean Institute of Information Scientists and Engineers (KIISE), the Korean Institute of Communications and Information Sciences (KICS), the Korean Information Processing Society (KIPS), and the Open Standards and ICT Association (OSIA). He has served as the general chair, the TPC chair/member, or an organizing committee member for international conferences, such as the Network Operations and Management Symposium (NOMS), International Symposium on Integrated Network Management (IM), Asia-Pacific Network Operations and Management Symposium (APNOMS), End-to-End Monitoring Techniques and Services (E2EMON), IEEE Consumer Communications and Networking Conference (CCNC), Assurance in Distributed Systems and Networks (ADSN), International Conference on Parallel Processing (ICPP), Data Integration and Mining (DIM), World Conference on Information Security Applications (WISA), Broadband Convergence Network (BcN), Telecommunication Information Networking Architecture (TINA), International Symposium on Applications and the Internet (SAINT), and International Conference on Information Networking (ICOIN). He was an Associate Editor of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, the IEEE JOURNAL OF COMMUNICATIONS AND NETWORKS, and the *International Journal of Network Management*, and an Associate Technical Editor of the *IEEE Communications Magazine*. He also serves as an Associate Editor for the *International Journal of Network Management* and *Future Internet Journal*.