Auction based Incentive Design for Efficient Federated Learning in Cellular Wireless Networks

Tra Huong Thi Le[1], Nguyen H. Tran[2], Yan Kyaw Tun[1], Zhu Han[3], Choong Seon Hong[1]
[1] Dept. of Computer Science and Engineering, Kyung Hee University, Korea
[2] School of Computer Science, The University of Sydney, Sydney, NSW 2006, Australia
[3] Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004-4005 USA
E-mail:{huong_tra25, ykyawtun7, cshong}@khu.ac.kr, nguyen.tran@sydney.edu.au, hanzhu22@gmail.com.

***Abstract*—Federated learning is an prominent machine learning technique that model is trained distributively by using local data of mobile users, which can preserve the privacy of users and still guarantee high learning performance. In this paper, we deal with the problem of incentive mechanism design for motivating users to participate in training. In this paper, we employ the randomized auction framework for incentive mechanism design in which the base station is a seller and mobile users are buyers. Concerning the energy cost incurred due to join the training, the users need to decide how many uplink subchannels, transmission power and CPU cycle frequency and then claim them in submitted bids to the base station. After receiving the submitted bids, the base station needs algorithms to select winners and determine the corresponding rewards so that the social cost is minimized. The proposed mechanism can guarantee three economic properties, i.e., truthfulness, individual rationality and efficiency. Finally, numerical results are provided to demonstrate the effectiveness, and efficiency of our scheme.**

I. Introduction

With the development of machine learning technology, mobile users have a great service experience during using the mobile applications. However, conventional machine learning techniques require the training data to be aggregated on one central server. On the other hand, the privacy protection of mobile users is a rising problem. Therefore, there is a demand in development of a distributed learning framework that mobile users can use the local collected data to train the learning model distributively. One of the most popular of distributed learning framework is the federated learning algorithm [1]. Federated learning first introduced by Google and is used to design a virtual keyboard application for smart phones named Gboard. Nowadays, federated learning can be a great solution for many services. Take the app Ware as an example. This app can support the users to avoid the heavy traffic roads but users have to share their own locations to the server. If federated learning is applied to this app, users need sending only the intermediate gradient values to the server instead of the raw data [2]

In federated learning, mobile users can collaboratively train a global model while all the training data are kept on their own devices. In particular, a mobile user computes the model updates based on its local data, which then sends back to the central server and aggregated there. That process is performed again until reaching an accuracy level of the learning model. By this way, federated learning can protect the personal information of mobile users as well as reduce the computation and storage cost of the central server. With the widely implementation of mobile edge computing, federated learning is easily deployed in real. Moreover, federated learning is also beneficial to such services like caching, transportation, shopping, and hospital which requires an unprecedented large-scale flexible data collection and model training. Different from conventional machine learning approaches, in federated learning algorithm, the user must transmit the training local model to the BS over a wireless links. Recently, several existing works studied the implementation problem of FL over wireless link [1], [3], [4].

However, there are still critical challenges for both base station and users in implementing the federated learning over wireless link due to the limit of wireless resource (bandwidth) and the limited energy cost of mobile users' devices. On the one hand, the bandwidth is limited and the base station (BS) do not have fully personal knowledge of CPU cycle or transmission power of mobile users' devices. On the other hand, not all mobile users contribute their resources unconditionally due to energy costs suffered model training. Therefore, it is necessary to design an incentive mechanism to motivate the self-interest mobile users to take part in model training and reveal the true information to the BS. Otherwise, some unreliable mobile users may perform undesirable behaviors, which affect to the performance of global training of a federated learning task. The work in [5] designs a incentive mechanism using the contract theory to motivate the high reputation users to participate in model training for preventing the poisoning attacks in federated learning. The contract is designed for different types of users, which categorized based on data quality of users and the considering resource is CPU cycle frequency. However, in this paper, we consider using the randomized auction [6] to stimulate the mobile users to join model training and share the resource. In the submitted bid, the mobile user clarifies the sub channel bundle, CPU cycle frequency and transmission power and the corresponding cost.

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2019-0-01287, Evolvable Deep Learning Model Generation Platform for Edge Computing). *Dr. CS Hong is the corresponding author.

978-1-7281-3106-1/20/$31.00 ©2020 IEEE

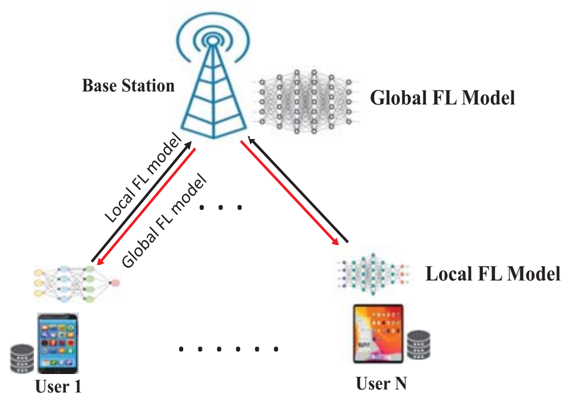

Fig. 1: System model.

With a given subchannel bundle and local accuracy level, the mobile user determines the CPU cycle frequency and transmission power in order to minimize the total cost. Based on the set of submitted bids of mobile users, the BS selects the different set of mobile users in different global iteration. The proposed mechanism can provide the guarantee of efficiency while approximate truthfulness and individual rationality are ensured simultaneously. The contributions of this paper are summarized as follows:

- First, we present a communication and computation model for federated learning in a cellular-connected wireless system. Due to the limitation of the uplink bandwidth, the BS needs to select proper users to execute the federated learning algorithm to minimize the social cost.
- Then, we present the bidding cost in each bid that the user submits to the BS. With the given the pre-defined subchannel bundle and local accuracy level, the user must optimize the transmission power and CPU cycle frequency s minimize the energy consumption while delay requirement of federated learning is satisfied.
- We adopt the a randomized auction mechanism to deal with the NP-hard social cost minimization problem in selecting the winning users and corresponding reward.
- Finally, the numerical studies show that a randomized auction mechanism can guarantee the approximation factor of the integrality to the optimal minimum cost.

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

II. System Model

Consider a cellular network consisting of one BS and a set $\mathcal{N}$ of N users. The BS and users collaboratively carry out a federated learning algorithm. The BS generates the initial global model and broadcast to users. Using its collected training data, users train a federated learning model. The federated learning model that is trained at each user is called the local model. Then, the BS collects all local models and generate a shared federated learning model, which is called global model. As shown in Fig. 1, the uplink from the users to the BS is used to transmit the parameters related to the local model while the downlink is used to transmit the parameters related to the global model. In this paper, we just consider the uplink bandwidth allocation due to the relation of the uplink bandwidth and the total cost of one user in learning a global model. Because the wireless resource is limited and users may be reluctant to participate in model training, we consider the interaction between the BS and users as the auction game. In this game, the BS and users have roles as one buyer and multiple sellers, respectively. The BS pays the reward for the users participating in the training model. The users join the training model and get the reward. User n submits a set $\mathcal{B}_n$ of bids to the BS. The ith bid submitted by the user n is denoted by b_n^i. Bid b_n^i consists of the resource (subchannel bundle C_n^i, power transmission p_n^i, CPU cycle frequency f_n^i and local accuracy level ϵ_n^i) and the claimed cost v_n^i due to the model training. Each user n determines individually its true cost V_n^i, which will be presented in Section III. Let $x(b_n^i)$ be a binary variable indicating the bid b_n^i wins.

After receiving all the bids from users, the BS computes the set $\mathcal{Q}$ of all winner profile with correlated probabilistic weight β. At the beginning of each global iteration, an winner profile $x^q = \{x(b_n^i)\}_{b_n^i \in \mathcal{B}}$ is selected with probability β^q by the BS. Let $r(b_n^i)$ be the amount of reward paid to a winning bid b_n^i.

In each global iteration, a selected winning mobile user joins model training and sends back the model update to the BS and gets the rewards. We assume that each subchannel can be allocated to more than one users.

A. Computation Model for Federated Learning

We denote s_n as the local data samples that mobile user $n \in \mathcal{N}$ uses to participate in the federated learning task. The computation resource that mobile user n contributes for local model training is denoted as f_n. The required number of CPU cycles for mobile user n to computing one sample of data in local model training is denoted by c_n. Therefore, the time duration of one local iteration is $\frac{c_n s_n}{f_n}$, and the energy consumption for one local iteration is presented as

$$E_n^{com}(f_n) = \zeta c_n s_n f_n^2, \tag{1}$$

where ζ is the effective capacitance parameter of computing chipset for mobile user n.

B. Communication Model for Fedarated Learning

In federated learning, the process of training the shared global model among mobile users is iterative until a global accuracy level of learning is achieved. The communication round between the mobile users and the BS is called global iterations. During a global iteration, the mobile users send the update of their own local model to the BS through wireless communications. Each local model update has a local accuracy which is denoted as ε_n. A larger local ε_n accuracy results in fewer local and global iterations. When the global accuracy is fixed, the number of local iterations is $\log(\frac{1}{\varepsilon_n})$. The duration time of a global iteration is the sum of the computation time of a local iteration and the uplink communication time of a local model update. The time duration of a local iteration of the mobile user n is denoted by

$$T_n^{comp} = \frac{c_n s_n}{f_n} \tag{2}$$

We assume that the BS allocate a set of $\mathcal{C} = \{1, 2, ..., C\}$ subchannels for uplink communication of local model update. Each subchannel has bandwidth of W. The number of antennas is assumed to be large to achieve the massive MIMO effect.

The achievable uplink data rate for user n in subchannel k is expressed as [7]

$$r_n^k = W \log(1 + \frac{(A_n - 1)p_n h_n^k}{W N_0}), \tag{3}$$

where p_n is the transmission power of mobile user n, h_n^k is the channel gain of peer to peer link between mobile user and the BS. N_0 is the background noise. A_n is the number of antennas the BS assigns to the mobile user n. The uplink data rate of user n is presented as

$$r_n = \sum_{k \in \mathcal{C}_n} r_n^k, \tag{4}$$

where C_n is the set of subchannels 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 mobile users. Therefore, the transmission time of a local model update is

$$T_n^{com} = \frac{\sigma}{r_n}. \tag{5}$$

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

$$T_{n,k}^t = \log(\frac{1}{\varepsilon_n}) T_n^{comp} + T_{n,k}^{com}. \tag{6}$$

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

$$E^{com} = T^{com} p_n = \frac{\sigma p_n}{r_n}. \tag{7}$$

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

$$E_{n,k}^t = \log(\frac{1}{\varepsilon_n}) E_n^{comp} + E_{n,k}^{com}. \tag{8}$$

III. Deciding users's bid

To transmit the local model update to the BS, users need bandwidth resource. However, there is a dependency between bandwidth and the corresponding cost considering the maximum tolerance time of federated learning. In this section, we present the way to determine the transmission power and the computation resource with a given subchannel bundle so that the total energy consumption in one global iteration is minimized. In other words, for bid b_n^i, user n calculates the transmission power p_n^i, the computation resource f_n^i and the cost v_n^i corresponding to a given subchannel bundle C_n^i. However, the process to decide users' bid is the same for every submitted bids. Thus, we remove the bid index i in this section. Thus, the energy cost of mobile user n is defined as

$$\begin{aligned} \textbf{P1:} \quad & \min_{f_n, p_n} \quad E_n^t && (9a)\\ & s.t. \quad T_n^t \le T_{max}, && (9b)\\ & \qquad f_n \in [f_n^{\min}, f_n^{\max}], && (9c)\\ & \qquad p_n \in (0, p_n^{\max}], && (9d)\\ & \qquad \varepsilon_n, \text{ and } C_n \text{ are given.} && (9e) \end{aligned}$$

In **P1**, the allocation of the uplink transmission power p and computation resources f are coupled from each other in the constraint (9b). **P1** can be solved by optimizing communication and computation resources iteratively.

A. Optimization of Uplink Transmission Power:

Each mobile user assigns its transmission power by solving the following problem:

$$\begin{aligned} \textbf{P2:} \quad & \min_{p_n} \quad f(p_n)\\ & s.t. \quad T_n^t \le T_{max},\\ & \qquad p_n \in (0, p_n^{\max}],\\ & \qquad f_n \text{ are given.} \end{aligned} \tag{10}$$

where $f(p_n) = \frac{\sigma p_n}{\sum_{k \in C_n} W \log(1 + \frac{(A_n - 1)p_n h_k}{W No})}$.

Note that $f(p_n)$ is quasiconvex in the domain. A general solution to deal with the quasiconvex optimization problem is the bisection method which a convex feasibility problem is solved each time. However, this solution requires $O(m^2/\alpha^2)$ iterations, where m is the dimension of the problem [8]. On the other hand, we have

$$f'(p_n) = \frac{\sigma \sum_{k \in C_n} \log(1 + \theta_n p_n h_n^k) + \sum_{k \in C_n} \frac{\sigma p_n \theta_n h_n^k}{\ln 2(1 + \theta_n p_n h_n^k)}}{(\sum_{k \in C_n} \log(1 + \theta_n p_n h_n^k))^2}, \tag{11}$$

where $\theta_n = \frac{(A_n - 1)}{W No}$. Then, we have

$$\phi(p_n) = \sigma \sum_{k \in C_n} \log(1 + \theta_n p_n h_n^k) + \sum_{k \in C_n} \frac{\sigma p_n \theta_n h_n^k}{\ln 2(1 + \theta_n p_n h_n^k)} \tag{12}$$

is a monotonically increasing transcendental function and negative at the starting point $p_n = 0$. Therefore, in this paper, a low-complexity bisection method is adopted. Specially, we calculate $\phi(p_n)$ each time, to get the optimal power allocation p_n as shown in Algorithm 1.

B. Optimization of CPU cycle frequency:

$$\begin{aligned} \textbf{P3:} \quad & \min_{f_n} \quad \log(\frac{1}{\varepsilon_n}) E_n^{comp}\\ & s.t. \quad T_n^t \le T_{max},\\ & \qquad f_n \in [f_n^{\min}, f_n^{\max}],\\ & \qquad p_n \text{ are given.} \end{aligned} \tag{13}$$

Algorithm 1: Optimal Uplink Power Transmission

Calculate $\phi(p_n^{max}) = \sigma \sum_{k \in C_n} \log(1 + \theta_n p_n^{max} h_k) + \sum_{k \in C_n} \frac{\sigma p_n^{max} \theta_n h_k}{\ln 2(1+\theta_n p_n^{max} h_k)}$
Calculate p_n^{min} so that $T_n^t(p_n^{min}) = T_{max}$
if $\phi(p_n^{max} < 0)$ **then**
  $p_n^* = p_n^{max}$
else
  $p_s = \max(0, p_n^{min})$ and $p_t = p_n^{max}$
  while $(p_t - p_s \leq \epsilon)$ **do**
    $p_l = (p_t + p_s)/2$
    if $\phi(p_l) \leq 0$ **then**
      $p_s = p_l$
    else
      $p_t = p_l$
    end
  end
  $p_n^* = (p_t + p_s)/2$
end

P3 is the convex problem, so we can solve it by any convex optimization tool. After deciding the bids, mobile users submit bids to the BS. The following section present the auction mechanism between the BS and users to select winners, allocate subchannels and decide the payment.

IV. Base station-User Auction Mechanism

A. Problem Formulation

In bid b_n^i that the user n submits to the BS, the subchannel bundle C_n^i and claimed cost v_n^i. The utility of one bid is the difference between the reward r_n^i and the true cost V_n^i.

$$U(b_n^i) = \begin{cases} r(b_n^i) - V_n^i, & \text{if bid } b_n^i \text{ wins,} \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

The payment that the BS for winning bids is $\sum_{b_n^i \in \mathcal{B}} r_n^i$. If users truthfully submit their cost, $V_n^i = v_n^i$, we have the social cost minimization problem defined as follows:

$$\textbf{P4:} \quad \min_{x} \quad \sum_{b_n^i \in \mathcal{B}} v_n^i x(b_n^i) \tag{15a}$$

$$s.t. \quad \sum_{b_n^i : c_k \in C_n^i} x(b_n^i) \geq 1, \forall c_k \in \mathcal{C}, \tag{15b}$$

$$\sum_{i} x(b_n^i) \leq 1, \forall i \in \mathcal{B}_n, \tag{15c}$$

$$x(b_n^i) = \{0, 1\}, \tag{15d}$$

where c_k denotes for one subchannel of set $\mathcal{C}$. The objective function of **P4** is to minimize the social cost which is the sum of the true cost of bids selected by the BS. The first constraint (15b) requires that each subchannel should be allocated to at least one user, the second constraint (15c) indicates at most one bid submitted by one user can win and the constraint (15d) indicates whether bid b_n^i wins.

The problem **P4** is a minimization knapsack problem which is proven to be NP-hard. It is infeasible to solve **P4** optimally in polynomial time. In addition, a VCG mechanism based payment rule is truthful only when the solution is optimal. Hence, we cannot apply the VCG payment directly.

To deal with the NP-hard problem, we adopt the framework of randomized auction [6], [9]. The following economic properties are desired.

Approximate-truthfulness in Expectation The dominant strategy of each user that aims to maximize the lower bound of expected utility is truthful submitting. This means $L(E(U(V_n^i, -\boldsymbol{b}_n^i))) \geq L(E(U((v_n^i, -\boldsymbol{b}_n^i)))$, where $L(E(U(V_n^i, -\boldsymbol{b}_n^i)))$ is the lower bound of the expected utility of b_n^i.

Individual Rational If each user reports its true cost, expected utility for each bid is nonnegative, i.e., $E(U(b_n^i)) \geq 0$.

$\ln(C)$**-Approximation Efficiency** The social cost is at most $\ln(C)$ times of minimum social cost denoted by S, i.e., $\sum_{b_n^i \in \mathcal{B}} v_n^i x(b_n^i) \leq \ln(C) S$.

B. Design of Randomized Auction

The randomized auction framework has three key steps: 1) an greedy algorithm that achieves $\ln(C)$-approximation efficiency, 2) convex decomposition, 3) VCG-based payment rule.

1) Approximation Algorithm: In this section, we use a greedy algorithm to solve the problem **P4** and the dual fitting method is applied to achieve an upper bound α.

Let y and z be the dual variable vector corresponding to constraints (15b) and (15c). Then the dual of problem LPR can be written as

$$\textbf{P5:} \quad \max_{\mathbf{y,z}} \quad \sum_{c_k \in \mathcal{C}} y_{c_k} - \sum_{n \in \mathcal{N}} z_n \tag{16a}$$

$$s.t. \quad \sum_{c_k : c_k \in \mathcal{C}^{\rangle}_{\langle}} y_{c_k} \leq v_n^i + z_n, \forall b_n^i \in \mathcal{B}, \tag{16b}$$

$$y_{c_k} \geq 0, \forall c_k \in \mathcal{C}, \tag{16c}$$

$$z_n \geq 0, \forall n \in \mathcal{N}. \tag{16d}$$

In the following, we introduce the Algorithm 2 to solve **P5** and compute α via dual fitting.

Proposition 1. *The upper bound of the integrality gap α between* ***P4*** *and its relaxation, and the approximation ratio of Algorithm 2 are* $\ln(C)$.

Proof. Based on the Algorithm 2, we have the integrality gap α is

$$\begin{aligned} O/O_f &\leq w/O_f \\ &= \ln(C)(\sum_{c_k \in \mathcal{C}} y_{c_k} - \sum_{n \in \mathcal{N}} z_n)/O_f \\ &\leq \ln(C) \end{aligned}$$

The approximation ratio is $w/O \leq w/O_f \leq \ln(C)$, where O is the optimal objective value of **P4**, O_f is the optimal objective of relaxation of **P4**, i.e., $1 \geq x(b_n^i) \geq 0$. □

Algorithm 2: Approximate Algorithm

for *each bid* b_n^i *in* $\mathcal{B}$ **do**
  $x(b_n^i) = 0$
end
subchannelnum = C
usernum=N
while $(\mathcal{C} \neq \varnothing)$ *AND* $(\mathcal{N} \neq \varnothing)$ **do**
  for *each bid* b_n^i *in* $\mathcal{B}$ **do**
    if $\mathcal{C} \bigcap C_n^i = \varnothing$ **then**
      remove b_n^i from $\mathcal{B}$
    else
      $C_n'^i = \mathcal{C} \bigcap C_n^i$
    end
  end
  [n,i] = $\arg\min_{b_n^i \in \mathcal{B}} \{v_n^i / |C_n'^i|\}$
  price = $\min_{b_n^i \in \mathcal{B}} \{v_n^i / |C_n'^i|\}$
  $x(b_n^i) = 1; w = w + v_n^i; \mathcal{C} = \mathcal{C} - C_n^i; \mathcal{N} = \mathcal{N} - \{i\}$
  $z_n = v_n^i$; $\mathcal{B} = \mathcal{B} - \mathcal{B}_n$;
  for *each subchannel* c_k *in winning bid* b_n^i **do**
    y_{c_k}= price/ln(subchannelnum)+$z_n / |C_n'^i|$
  end
end

2) Convex Decomposition: By relaxing $1 \geq x(b_n^i) \geq 0$ of **P4**, we have the LPR of **P4**. Solving LPR of **P4** gives us an optimal fractional x_n^{i*} and the corresponding OPT_f

$$\sum_{c_k \in \mathcal{C}} y_{c_k} \leq \alpha OPT_f, \tag{17}$$

In this step, we aim to find the combination weighting coefficient β^q, such that

$$\sum_{q \in \mathcal{Q}} \beta^q = 1, \text{and} \sum_{q \in \mathcal{Q}} \beta^q \mathbf{x}^q \leq \alpha \mathbf{x}^*. \tag{18}$$

The weight β^q is the probability that feasible integer solution $\mathbf{x}^q$ is selected. In order to obtain the selecting probability β^q for each integer solution $\mathbf{x}^q$, we have to solve the following problem:

$$\begin{aligned} \textbf{Primal:} \quad & \max_{\boldsymbol{\beta}} \quad \sum_{q \in \mathcal{Q}} \beta^q \\ & s.t. \quad \sum_{q \in \mathcal{Q}} \beta^q \mathbf{x}^q \leq \alpha \mathbf{x}^*, \\ & \qquad \sum_{q \in \mathcal{Q}} \beta^q \leq 1, \\ & \qquad \beta^q \geq 0, \ \forall q \in \mathcal{Q}. \end{aligned} \tag{19}$$

This problem is hard to solve because there is an exponential number of variables. Thus, we transform (19) into the dual form with an exponential number of constraints. The corresponding dual problem of (19) is given as follows:

$$\begin{aligned} \textbf{Dual:} \quad & \min_{\boldsymbol{\omega}, \iota} \quad \sum_{b_n^i \in \mathcal{B}} \alpha x^*(b_n^i) \omega_{b_n^i} + \iota \\ & s.t. \quad \sum_{b_n^i \in \mathcal{B}} x^q(b_n^i) \omega_{b_n^i} + \iota \geq 1, \ \forall q \in \mathcal{Q}, \\ & \qquad \iota \geq 0, \quad \omega_{b_n^i} \geq 0 \quad \forall b_n^i \in \mathcal{B}. \end{aligned} \tag{20}$$

To solve the dual problem in polynomial time, the ellipsoid algorithm is applied. Specially, a polynomial number of hyperplanes is used to obtain the optimal solution. Our proposed approximation algorithm is used to create hyperplane, which correlates with a feasible integral solution. Each feasible integral solution further corresponds to a variable in the primal problem. This allows us to solve the primal problem in polynomial time. Hence, we can solve the primal problem in polynomial time.

3) Randomized User Selection and Reward Determination: In this step, each possible integer solution $\mathbf{x}^q$ is chosen with a probability β^q that we can obtain in step convex decomposition. If i-th bid of user n is selected, then we use the VCG-based payment to decide the reward to user n as follows

$$r_n^i = r_n^{i*} \frac{\sum_i v_n^i x^q(b_n^i)}{\sum_i v_n^i x^*(b_n^i)}. \tag{21}$$

where $r_n^{i*} = V_{\mathcal{B}|\{b_n^i\}} - \sum_{b_n^{i'} \in \mathcal{B}|\{b_n^i\}} v_n^i x*(b_n^{i'})$, $V_{\mathcal{B}|\{b_n^i\}}$ is the minimum social cost of LPR of **P4** when b_n^i is deleted from $\mathcal{B}$, and $\sum_{b_n^{i'} \in \mathcal{B}|\{b_n^i\}} v_n^i x*(b_n^{i'})$ is the social cost of LPR of **P4** minus the cost of b_n^i.

Proposition 1 ensures the property of $\ln(C)$-approximation efficiency. Following the property of VCG-based payment rule, the randomized auction mechanism can provide the approximate truthfulness in expectation and individual rationality.

V. Simulation Result

In this section, we provide some simulation results to evaluate the proposed mechanism. For simulation parameters, the CPU cycles for performing a data sample is $c_n = 20$, the size of data samples is $s_n = 800 \times 10^3$, the maximum tolerance time of a federated learning task is $T_{max} = [25, 75]$. Some other predefined parameters $p_{max} = 5$, $\zeta = 2$, $No = -174 dBm/Hz$, $W = 15kHz$.

Fig. 2 shows the energy cost of one bid of a user when the upper limit of local data accuracy varies from 35% to 75%. When the local data accuracy increases, the energy cost decreases. This is because that it needs fewer iterations.

Fig. 3 presents the energy cost of one bid of a user when the maximum tolerance time varies from 35 to 75. When the maximum tolerance time increases, the energy cost decreases. Because user can keep low transmission power and CPU cycle frequency so that it can satisfy the delay constraint. Fig. 4 reports approximation ratio as the number of subchannels C and the number of users increases. When the subchannel number increases, the theoretical approximation ratio, which is equal to $\ln(C)$, increases. We can see that the actual ratio is much smaller than the theoretical ratio.

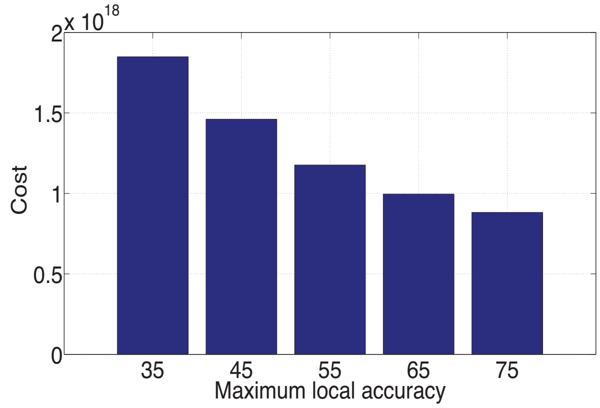

Fig. 2: The social cost under different levels of local training data.

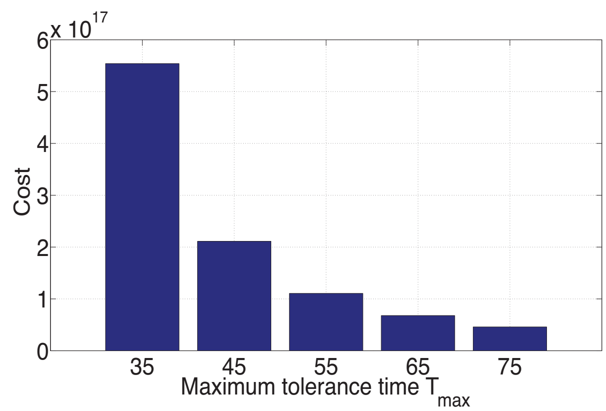

Fig. 3: The social cost under different levels of local training data.

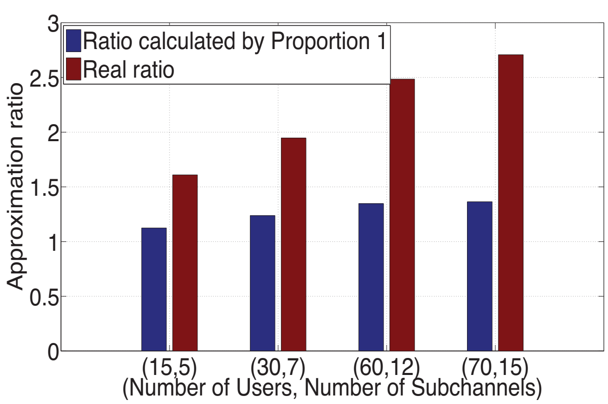

Fig. 4: Approximation versus (number of users, number of subchannels).

Fig. 5 shows the social cost achieve by the proposed scheme and by solving the relaxation of **P4** when the number of users varies. The upper bound of expected cost is equal to $\ln(C)$ times the social cost by solving the relaxation of **P4**. When

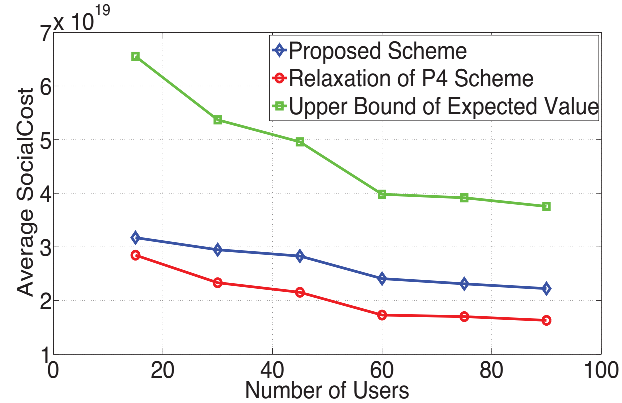

Fig. 5: Social Cost versus users under different schemes.

the user number increases, the social cost decreases. This is because the BS can choose the bids with lower cost as the winners.

VI. Conclusion

This paper focuses on incentive mechanism design to stimulate users to participate in federated learning. We present the method for users to decide the bids submitted to the BS. We also employ the randomized auction framework in mechanism design to tackle the NP-hard user selection problem and to guarantee approximate truthfulness in expectation, individual rationality and computation efficiency. Simulation results show the effectiveness of the proposed mechanism.

References

[1] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konecny, S. Mazzocchi, H. B. McMahan *et al.*, "Towards federated learning at scale: System design," *arXiv preprint arXiv:1902.01046*, 2019.

[2] N. H. Tran, W. Bao, A. Zomaya, and C. S. Hong, "Federated learning over wireless networks: Optimization model design and analysis," *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, 2019.

[3] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," *arXiv preprint arXiv:1909.07972*, 2019.

[4] 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," *https://arxiv.org/abs/1911.01046v1*, 2019.

[5] 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 of Things Journal*, 2019.

[6] J. Li, Y. Zhu, Y. Hua, and J. Yu, "Crowdsourcing sensing to smartphones: A randomized auction approach," *IEEE Transactions on Mobile Computing*, vol. 16, no. 10, pp. 2764–2777, 2017.

[7] Y. Hao, Q. Ni, H. Li, and S. Hou, "Energy-efficient multi-user mobile-edge computation offloading in massive mimo enabled hetnets," *2019 IEEE International Conference on Communications (ICC)*, 2019.

[8] X. Lyu, H. Tian, C. Sengul, and P. Zhang, "Multiuser joint task offloading and resource optimization in proximate clouds," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3435–3447, 2016.

[9] W. Zhong, K. Xie, Y. Liu, C. Yang, and S. Xie, "Efficient auction mechanisms for two-layer vehicle-to-grid energy trading in smart grid," *2017 IEEE International Conference on Communications (ICC)*, 2017.