An Incentive Mechanism for Federated Learning: A Continuous Zero-Determinant Strategy Approach

Changbing Tang, *Senior Member, IEEE*, Baosen Yang, Xiaodong Xie, Guanrong Chen, *Life Fellow, IEEE*, Mohammed A. A. Al-qaness, and Yang Liu, *Senior Member, IEEE*

Abstract—As a representative emerging machine learning technique, federated learning (FL) has gained considerable popularity for its special feature of "making data available but not visible". However, potential problems remain, including privacy breaches, imbalances in payment, and inequitable distribution. These shortcomings let devices reluctantly contribute relevant data to, or even refuse to participate in FL. Therefore, in the application of FL, an important but also challenging issue is to motivate as many participants as possible to provide high-quality data to FL. In this paper, we propose an incentive mechanism for FL based on the continuous zero-determinant (CZD) strategies from the perspective of game theory. We first model the interaction between the server and the devices during the FL process as a continuous iterative game. We then apply the CZD strategies for two players and then multiple players to optimize the social welfare of FL, for which we prove that the server can keep social welfare at a high and stable level. Subsequently, we design an incentive mechanism based on the CZD strategies to attract devices to contribute all of their high-accuracy data to FL. Finally, we perform simulations to demonstrate that our proposed CZD-based incentive mechanism can indeed generate high and stable social welfare in FL.

Index Terms—Federated learning (FL), game theory, incentive mechanism, machine learning, zero-determinant strategy.

I. INTRODUCTION

THE emergence of artificial intelligence and the rapid development of 5G technologies have made the Internet of Things gradually evolve into the Internet of Intelligences [1]. Many studies on the machine learning (ML) and the iterative learning control of multi-agents once attracted everyone's attention [2]−[4]. However, issues such as data silos and privacy security directly hindered their development [5], [6]. As a representative distributed ML technique, federated learning (FL) has attracted considerable attention for its characteristic of "making data available but not visible" [7]−[9]. In FL, all participants work in a collaborative manner [10], [11]. There is no need to upload private personal data during training [12], and the devices only need to use the shared model for local iterative learning, global aggregation [13] and updating [14].

However, when applying FL in real applications, devices are often reluctant to contribute relevant data, or even refuse to participate in FL, as they are faced with sensitive issues such as privacy risks and break-even results. There are some studies considering these issues. For example, in [15] an incentive scheme was designed to maximize the collective utility while minimizing the inequality among data owners. However, some studies were based on unreasonable assumptions, such as mobile devices voluntarily participating in FL without receiving any rewards [15], [16], which is unrealistic in real-life scenarios. Furthermore, risks such as resource consumption and privacy leakage make the devices more reluctant to participate in FL. Therefore, an important and also challenging issue is: *Determining how to design an effective incentive mechanism that not only optimizes the social welfare of FL, but also encourages devices to contribute all of their high-quality data to FL.*

To address the aforementioned problem, game theory offers an ideal approach since it is an effective tool for competition and cooperation of individuals participating in an interacting process [17]−[19], such as FL, or more generally, ML. There are already some works using game theory to design effective incentive mechanisms, for example to optimize social welfare [20]−[24]. To select reliable devices participating in FL, in [20] a reputation-based incentive mechanism was proposed based on contract theory. Some studies combine effective computational algorithms with game theory to enhance FL efficiency [25]−[29]. Nevertheless, further research is still needed to balance schematic efficiency and social welfare.

As a new methodology for studying various strategies in game theory, zero-determinant (ZD) strategies [30] have found many applications. By using ZD strategies, a player can unilaterally control the opponent's expected utility, regardless of the opponent's strategy used. On one hand, some work aims to enhance the applicability of ZD strategies through the-

C. Tang, X. Xie, and M. Al-qaness are with the College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua 321004, China (e-mail: tangcb@zjnu.edu.cn; xiaodong@zjnu.edu.cn; alqaness@zjnu.edu.cn).

B. Yang is with the School of Mathematical Sciences, Zhejiang Normal University, Jinhua 321004, China (e-mail: bsyang@zjnu.edu.cn).

G. Chen is with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China (e-mail: eegchen@cityu.edu.hk).

Y. Liu is with Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, and also with the School of Mathematical Sciences, Zhejiang Normal University, Jinhua 321004, China (e-mail: liuyang@zjnu.zinu.cn).

oretical analysis [31]−[34]. For example, in [31] the ZD strategies were classified into three types (fairness, extortion, and generosity) to study cooperation and control in social dilemmas. On the other hand, some work incorporates practical consideration into modeling and their performances [35]−[38]. For instance, in [36] a collective extortion strategy was developed to accelerate the learning process. Notably, very few work focuses on continuous strategy games, especially in multi-player continuous strategy game scenarios of FL. This motivates us to design an incentive mechanism for FL based on continuous ZD (CZD) strategies.

In the practical FL scenario, there are some advantages of considering CZD strategies. First of all, the incomplete cooperation of contributing a portion of high-quality data is widely present in FL. That is, the continuous action space is more practical than the discrete action space in FL. Actually, the ZD strategies are special cases of CZD strategies. The CZD strategies maintain the control effect of the ZD strategies and extend the action spaces from discrete to continuous. Secondly, as continuous strategies contain more complete information compared to discrete strategies, the CZD strategies have broader incentive ranges than the ZD strategies. Therefore, compared with the ZD strategies, the incentive mechanism designed by the CZD strategies can more effectively motivate the participating FL devices to contribute high-quality data. At last, the fair distribution of rewards can be achieved by adjusting the parameters in the CZD strategies. That is, the incentive mechanism designed by applying the CZD strategies can effectively achieve "no pain no gain", in which devices contributing high accuracy data can receive high rewards and the server gains a high-quality global model. Therefore, the incentive mechanism designed by the CZD strategies can lead to a win-win situation in FL.

In this paper, we consider the continuous strategy game scenario in FL and propose an incentive mechanism based on the CZD strategies. We first model the interaction between devices and servers during FL as an iterative game. We then design an incentive mechanism to optimize social welfare. We further apply CZD with two players, and then multiple players, to optimize social welfare in FL. Based on theoretical analysis, we demonstrate that servers can control social welfare at a high and stable level by applying the CZD strategies with refinement.

The main contributions of this study are summarized as follows:

1) We model the interaction between the server and the devices in the FL process as a continuous iterative game, based on which the ZD strategies are extended to the CZD strategies. Practically, the CZD strategies cover more complete information than the ZD strategies, which results in problem where the CZD strategies are more suitable to demonstrate the incomplete cooperation scenario of FL. Our work further develops ZD theory and enriches the application of ZD strategies in continuous strategy scenarios.

2) Based on the CZD strategies with two-player and multi-player scenarios, we design an incentive mechanism for FL. This mechanism can not only motivate as many devices as possible to contribute high-quality data in FL, but also enhance the efficiency of FL. Compared with the existing ZD-based studies, our proposed approach can achieve broader incentive range, more efficient motivation, more effective incentive compatibility, and more accurate fair distribution of rewards.

3) We combine numerical and real experiments to validate the advantages of the proposed method in simulations. On one hand, the strengths of the CZD strategies are proved by numerical experiments. On the other hand, the effectiveness of the CZD-based incentive algorithm is verified by FL experiments with real datasets. To our knowledge, among the existing FL incentive mechanisms based on the ZD strategies, we are the first to perform both numerical and real experiments.

The remainder of the paper is organized as follows. In Section II, a literature review is presented. In Section III, the system model based on the CZD strategies in FL is established. In Section IV, by applying the CZD strategies to the continuous scenarios with two players and then multiple players, an incentive mechanism algorithm is designed for resolving the dilemma in FL. Finally, in Section V, simulation results are demonstrated, and in Section VI the investigation is concluded.

II. RELATED WORK

Several works [20]−[24] designed effective incentive mechanisms based on game-theoretic analyses. To optimize the system payoff, a hierarchical incentive framework based on coalitional and contractual games was proposed in [21]. In [22], the stackelberg game was applied from the perspective of communication efficiency to optimize FL. In [23], an incentive mechanism was designed for FL based on the auction theory. Moreover, some works combined computational algorithms with game theory [25]−[29], aiming to optimize social welfare and solve the incentive problem in FL. However, the aforementioned works are either from a theoretical perspective to improve FL efficiency or from an algorithmic perspective to optimize the social welfare, which did not or were unable to balance the efficiency and effectiveness through the FL process.

As a new approach in game theory, the ZD strategies [30] led to extensive research in recent years. In [32], the ZD strategies were applied to both discrete and continuous strategy scenarios for wireless resource sharing, where the number of players was extended from two to *N*. Also, in [33] the fairness of the ZD strategies was studied and in [34] the ZD strategies were refined to a ruling strategy in a repeated game with a limited number of players and action sets. Furthermore, to encourage collaborative mining in PoW-based blockchain networks, an incentive mechanism based on the ZD strategies was proposed in [35]. In [37], the ZD strategies were applied to solve the reward allocation problem in mobile crowdsourcing. The multi-player multi-action ZD (MMZD) strategy was applied in [38] to address social dilemmas in cross-silo FL. However, it was noticed that some factors such as high incentive costs and the random nature of device' strategies could still reduce FL efficiency and even harm social welfare.

Inspired by the above works, we consider the continuous strategy game scenarios in FL and propose an incentive mech-

anism based on the CZD strategies. Compared with the existing works, our proposed approach has several differences. Firstly, continuous strategies cover more complete information than discrete strategies, which is caused by the dynamic interaction between FL participants. Therefore, we apply CZD strategies to solve the incentive problem in FL. This differs from existing studies [30]−[38] who focused on discrete ZD strategies. Secondly, the CZD strategies can promote full cooperation of FL participants, which is confirmed by theoretical derivations in two-player and multi-player scenarios. Based on this, the designed CZD-based incentive algorithm can optimize social welfare and also improve FL efficiency. This is different from the existing studies. In [20]−[24], game methods were adopted to optimize social welfare; in [25]−[29], several efficient computational algorithms were applied to improve FL efficiency. Thirdly, the effectiveness of the CZD-based incentive algorithm is verified by FL experiments with real datasets. This is different from existing studies [36]−[38] which were verified by numerical experiments without real experiments.

III. PROBLEM FORMULATION

A. System Model

Fig. 1 shows an FL scenario consisting of one server and N devices. First, the server issues the FL task (including the global model, deadline, and expected reward), and the devices choose whether to join based on whether their local data matches the FL task sufficiently. Secondly, the server selects candidate devices by pre-training and sends the global model. The candidate devices do not need to upload any sensitive data, but only need to receive the global model for local training. Third, if the model parameters and gradient values meet the thresholds set by the server, which were obtained after several rounds of local iterative computation, the devices can apply to upload the local models. Finally, after local model aggregation and global model updating, the server will distribute rewards to the devices participating in FL.

However, due to the lack of trust and the unavoidable information gap, selfish devices may resort to "cheating" (e.g., contributing their incomplete or irrelevant data during FL) to save computational resources and demand rewards that do not match their contributions. Obviously, it is more effective to prevent and control problems by taking preventive measures than to solve them during the working process. Therefore, to avoid selfish behavior of a device, the server can dynamically modulate rewards based on the contribution of the device. Conveniently, we summarize the notations in Table I.

B. Game Formulation

In the FL process, all participants have their own preferences for the choice of strategies and make decisions independently by combining their costs and benefits. In fact, the devices and the server involved in FL have no way to obtain strategy information of others in interaction, and there are multiple decision participants. Therefore, the interaction between devices and the server can be modeled as a multi-player synchronous game.
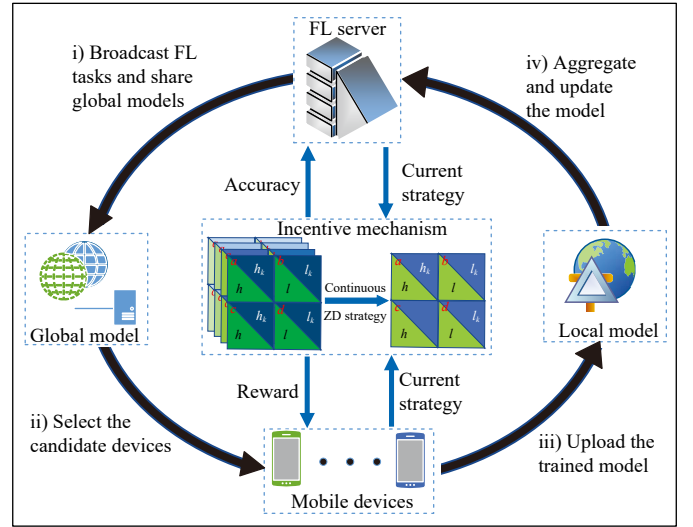

Fig. 1. System model. The process of the FL task can be divided into four stages: i) Broadcasting of the FL task and initialization of the global model; ii) Selection of candidate devices and transmission of the global model; iii) Training and uploading of local models; iv) Model aggregation and updating, and reward distribution. During the FL process, devices may resort to cheating to save their computational resources. To enhance the efficiency of FL, it is necessary to design an effective incentive mechanism that can encourage devices to contribute their high-quality data to FL.

TABLE I
NOTATION

Notation	Meaning
N	The total number of devices
Δ_k	The earth mover's distance (EMD) of the k-th device
$x_k \in [l_k, h_k]$	The strategy of the k-th device
S_k	The FL-related data size of the k-th device
θ	The accuracy of global model
$\overline{\Delta}$	The average EMD value of all devices
$y \in [l, h]$	The strategy of the server
r_k	The reward determined by the server to the k-th device
$p_k(x_k)$	The conditional probability that the k-th device chooses x_k
$q(y)$	The conditional probability that the server chooses y
u_k	The utility equation of the k-th device
C_k	The total cost of the k-th device
u_s	The utility equation of the server
R	The transmission and aggregation cost of the server
U_k	The payoff matrix of the k-th device
U_s	The payoff matrix of the server
E_k	The expected utility of the k-th device
E_s	The expected utility of the server
χ	A non-zero scalar in the CZD strategies

In addition, in terms of the participants' strategies, they will have more than just two discrete outcomes. Specifically, for the server, it will not foolishly choose to pay in full in any situation, nor will it recklessly choose a zero-payment strategy. For the devices, they can contribute partial data instead of par-

ticipating or not participating during FL. Therefore, in this paper, we focus on the continuous strategy scenario in FL, which is more comprehensive and realistic in practice.

The strategy of the k-th device ($k = 1, 2, \ldots, N$) in the current round can be defined as $x_k \in [l_k, h_k]$, that is, local data accuracy is controlled by the device and chosen to invest in the FL task for model training. Using incomplete information among FL participants, l_k and h_k can be obtained by pre-training. The strategy of the server in the current round can be defined as $y \in [l, h]$; that is, the reward is given to all devices under the control of the server, which is effective on the accuracy of the trained global model.

Based on the experimental validation reported in [27], the strategy of the k-th device can be modeled by the following function via curve fitting:

$$x_k = x_k(S_k, \Delta_k) = \pi(\Delta_k) - c_1 e^{-c_2(c_3 S_k)^{\pi(\Delta_k)}} \tag{1}$$

where $\pi(\Delta_k) = c_4 e^{-\left(c_5 + \frac{\Delta_k}{c_6}\right)^2}$, and Δ_k is the earth mover's distance (EMD) metric [26] of the k-th device for FL, which is leveraged to quantify the dissimilarity of its data distributions. A large EMD indicates a high dissimilarity that adversely influences the accuracy of the local model. Also, S_k represents the FL-related data size of the k-th device. Moreover, c_i ($i \in \{1, 2, 3, 4, 5, 6\}$) are the positive curve fitting parameters. On the right-hand side, the first term $\pi(\Delta_k)$ captures the degradation of the model performance when Δ_k increases. The exponential term $c_1 e^{-c_2(c_3 S_k)^{\pi(\Delta_k)}}$ reflects that the higher the data quality and the data size, the better the model accuracy.

Similarly, by curve fitting [27], the accuracy of the global model can be expressed as

$$\theta = \pi(\overline{\Delta}) - c_1 e^{-c_2\left(\frac{c_3}{N} \sum_{k=1}^{N} x_k S_k\right)^{\pi(\overline{\Delta})}} \tag{2}$$

where $\overline{\Delta} = \frac{1}{N} \sum_{k=1}^{N} \Delta_k$ is the average EMD value of all devices in the FL task.

Under the conditions that the k-th device takes x'_k and that the server takes y' in the previous round, the conditional probabilities of the k-th device and the server taking x_k and y in the current round are defined by $p_k(x_k) = p_k(x'_1, \ldots, x'_N, y'; x_k)$ and $q(y) = q(x'_1, \ldots, x'_N, y'; y)$, respectively. Based on this definition, the mixed strategy of the k-th device has a constraint where

$$\int_{l_k}^{h_k} p_k(x_k) dx_k = 1. \tag{3}$$

Similarly, the mixed strategy of the server has a constraint where

$$\int_{l}^{h} q(y) dy = 1. \tag{4}$$

To compensate for the devices' costs in FL, a payment strategy $r_k = \varsigma_k y \left(\frac{x_k}{h_k} + \tau\right)$ is determined by the server to each k-th device involved in the FL task, as a reward to its contribution, where ς_k and τ are adjustment scalars controlled by the server. Then, the utility equation of the k-th device can be expressed as

$$u_k = \sigma_k r_k - \delta_k C_k(x_k) \tag{5}$$

where $\sigma_k, \delta_k > 0$ are the weight parameters for balancing revenue and cost, and C_k is the total cost of the k-th device, which is monotonically increasing in x_k.

Correspondingly, the utility equation of the server can be expressed as

$$u_s = \sigma_s v\theta - \delta_s \left[\sum_{k=1}^{N} (r_k) + \rho R\right] \tag{6}$$

where σ_s and $\delta_s > 0$ are the weight parameters to balance the model accuracy and the total payment, v is the expected return of the server corresponding to the unit global model accuracy, ρ is a positive adjustment parameter, and R is the transmission and aggregation expense of the server in FL.

Based on the interaction between the server and the devices, the FL game can be precisely defined as follows.

Definition 1 (FL Game): In the FL scenario consisting of one server and N devices, the devices can determine the quality and quantity of data, and the server can control the rewards to the devices. The interaction between the devices and the server is equivalent to a multiplayer synchronous repetitive game, called the FL Game, $\mathcal{G} = (N \cup \{s\}, [l_k, h_k] \cup [l, h], \{u_k\} \cup \{u_s\})$ with $k \in \{1, 2, \ldots, N\}$.

C. Dilemma in the FL Game

Optimal actions of rational players correspond to actions that maximize their own profits. However, this is a dilemma in the FL game, as further discussed below.

Proposition 1: In the FL game described in Definition 1, the best actions of the k-th device and server are l_k and l, respectively.

Proof: For the k-th device, according to (5), if the server's strategy y is fixed, the utility u_k is a monotonically decreasing function about x_k. That is, whatever the server's action is, the best action of the k-th device is always l_k.

Without loss of generality, two extreme cases are considered here to verify the conclusion. If $y = h$, then r_k in (5) is taken to the maximum value. Considering the total cost $C_k(l_k) < C_k(h_k)$, one has $u_k(l_k) > u_k(h_k)$, which leads to the best action of $x_k = l_k$. If $y = l$, so it will not gain more by choosing $x_k = h_k$.

Similarly, for the server, if the k-th device's strategy x_k is fixed, the utility u_s is a monotonically decreasing function about y. Since $r_k(l) < r_k(h)$, one has $u_s(l) > u_s(h)$. According to (6), u_s can be maximized only if r_k becomes as small as possible, which corresponds to $y = l$.

In other words, regardless of the action taken by the device, the best action for the server is $y = l$. ∎

It follows from Proposition 1 that the best action of the device is to contribute incomplete high-accuracy data, and the best action of the server is to set a lower level of reward.

This dilemma will seriously harm the social welfare or even hinder the implementation of FL. As the initiator and manager of FL, the server has the motivation and ability to take precautions (e.g., an incentive mechanism) to attract contribution from the device so as to avoid the dilemma.

However, on one hand, the utility of the server is affected

by the strategies of all participants, which makes it difficult to unilaterally adopt optimal strategies. On the other hand, the additional costs of the preventive measures are critical challenges. Inspired by [30], the ZD strategies can help the adopter unilaterally control the proportional relationship between the expected utilities of two participants, which is expected to resolve the FL dilemma. With this consideration, in [36] the ZD strategies were extended to multiplayer discrete strategy scenarios for accelerating the learning process.

Notice that the traditional ZD strategies are applicable to discrete strategy games with two players, while the strategies of participants in FL are usually continuous. Therefore, this paper studies continuous strategy scenarios with two players and then multiple players, and designs an incentive algorithm for FL based on the CZD strategies to attract contributions from devices.

IV. INCENTIVE MECHANISM IN FL BASED ON THE CZD STRATEGIES

A. CZD Strategies With Two Players

When $N = 1$, the strategy functions of the devices and the server are x_k and y, respectively, and the utility of the devices and the server are $u_k(x_k, y)$, and $u_s(x_k, y)$, respectively. The probability of the k-th device and the server to choose strategies x'_k and y' in the previous round and the current round are $z(x'_k, y')$ and $z(x_k, y)$, respectively.

If the steady-state vector satisfies $z = z(x'_k, y') = z(x_k, y)$, the expected utility of the k-th device within each round can be expressed as

$$E_k = \int_l^h \int_{l_k}^{h_k} [u_k(x_k, y) \cdot z] dx_k dy \tag{7}$$

and the expected utility of the server within each round can be expressed as

$$E_s = \int_l^h \int_{l_k}^{h_k} [u_s(x_k, y) \cdot z] dx_k dy. \tag{8}$$

Based on the theory of the ZD strategies, the following lemma is established.

Lemma 1: For $x_k \in [l_k, h_k]$, $y \in [l, h]$, if the server's strategy $\bar{q}(x'_k, y'; h)$ meets

$$\bar{q}(x'_k, y'; h) = \begin{cases} q(x'_k, y'; h) & y \in [l, h) \\ q(x'_k, y'; h) + 1 & y = h \end{cases} \tag{9}$$

where

$$\bar{q}(x'_k, y'; h) = \chi \left[\alpha u_k(x'_k, y') + \beta u_s(x'_k, y') - \gamma \right] \tag{10}$$

then the expected utility of all FL participants will satisfy

$$\alpha E_k + \beta E_s - \gamma = 0 \tag{11}$$

where the factor χ is a non-zero scalar, α and β are weight factors, and γ is a parameter.

Proof: The proof is given in Appendix A. ∎

When the server takes the CZD strategies as in (9) and (10), social welfare can always be kept at the desired level, whatever the strategies of the devices are. The objective can be transformed to the following optimization problem:

$$\boldsymbol{P1} : \max_{q(y)} \quad \alpha E_k + \beta E_s \quad \forall p_k(x_k)$$

$$\text{s.t.} \begin{cases} 0 \le q(x'_k, y'; y) \le 1 \\ \alpha E_k + \beta E_s - \gamma = 0. \end{cases} \tag{12}$$

By solving the problem $\boldsymbol{P1}$, the following theorem can be established.

Theorem 1: When the strategy of the server satisfies

$$\bar{q}(x'_k, y'; h) = \begin{cases} \chi \left(\alpha u_k^0 + \beta u_s^0 - \gamma_{\max} \right) & \text{if } y' < h \\ \chi \left(\alpha u_k^0 + \beta u_s^0 - \gamma_{\max} \right) + 1 & \text{if } y' = h \end{cases} \tag{13}$$

the social welfare γ will reach a maximum value, as

$$\gamma_{\max} = \begin{cases} \max \left(\alpha u_k^1 + \beta u_s^1 \right) & \text{if } \chi < 0 \\ \max \left[\alpha u_k \left(x_k^0, h \right) + \beta u_s \left(x_k^0, h \right) \right] & \text{if } \chi > 0. \end{cases} \tag{14}$$

where $u_k^i = u_k(x_k^i, y^i)$, $u_s^i = u_s(x_k^i, y^i)$, $x_k^i \in [l_k, h_k], y^i \in [l, h]$, $i = 0, 1$.

Proof: The proof is given in Appendix B. ∎

At this point, the social welfare of the FL can be unilaterally controlled at $\gamma_{\max}$. That is, by applying the CZD strategies, the server can unilaterally control the social welfare of the FL at the desired level whatever the devices' strategies are.

B. CZD Strategies With Multiple Players

In this subsection, consider a scenario where there is one server and N devices playing the game with their continuous strategies in the FL iterative task, i.e., $\forall x_k \in [l_k, h_k]$ ($k = 1, 2, \ldots, N$), $\forall y \in [l, h]$.

The corresponding formula in the continuous strategy setting with two players is extended to $N + 1$ dimensions.

According to Definition 1, when the strategy functions of the k-th device and the server are x_k and y, respectively, the utility of the k-th device can be defined as $u_k = u_k(x_1, \ldots, x_N, y)$ ($k = 1, 2, \ldots, N$). Additionally, the utility of the server can be defined as $u_s = u_s(x_1, \ldots, x_N, y)$.

To reveal the relationship between the variables more intuitively, set

$$\Phi(x_1, \ldots, x_N, y)$$

$$= \left[\sum_{k=1}^N \alpha_k u_k(x_1, \ldots, x_N, y) \right] + \beta u_s(x_1, \ldots, x_N, y). \tag{15}$$

The probability that the devices and the server select the corresponding strategies x'_k and y' in the previous round can be defined as $z(x'_1, \ldots, x'_N, y')$. In the current round, the probability can be defined as $z(x_1, \ldots, x_N, y)$. Thus, the corresponding transformation equation can be written as

$$M(x'_1, \ldots, x'_N, y'; x_1, \ldots, x_N, y)$$

$$= q(x'_1, \ldots, x'_N, y'; y) \prod_{k=1}^N p_k(x'_1, \ldots, x'_N, y'; x_k).$$

Consequently, one has

$$z(x'_1, \ldots, x'_N, y') M(x'_1, \ldots, x'_N, y'; x_1, \ldots, x_N, y)$$

$$= z(x_1, \ldots, x_N, y).$$

If the steady-state vector satisfies $z = z(x'_1,\ldots,x'_N,y') = z(x_1,\ldots,x_N,y)$, the expected utility of the k-th device can be expressed as

$$E_k = \int_l^h \int_{l_1}^{h_1} \cdots \int_{l_N}^{h_N} (u_k \cdot z) \quad dx_N \cdots dx_1 dy \qquad (16)$$

and the expected utility of the server can be expressed as

$$E_s = \int_l^h \int_{l_1}^{h_1} \cdots \int_{l_N}^{h_N} (u_k \cdot z) \quad dx_N \cdots dx_1 dy. \qquad (17)$$

The total utilities of all the players with weight factors α_k ($k = 1, 2, \ldots, N$) and β can be defined as the social welfare, i.e.,

$$\bar{E} = \left(\sum_{k=1}^N \alpha_k E_k \right) + \beta E_s. \qquad (18)$$

Similar to the previous section, the following lemma can be established based on the theory of the ZD strategies.

Lemma 2: For $x_k \in [l_k, h_k]$, $k = 1, 2, \ldots, N$, $y \in [l, h]$, if the server's strategy $\tilde{q}(h) = \tilde{q}(x'_1, \ldots, x'_N, y'; h)$ meets

$$\tilde{q}(h) = \begin{cases} q(x'_1,\ldots,x'_N,y';h) & y \in [l,h) \\ q(x'_1,\ldots,x'_N,y';h) + 1 & y = h \end{cases} \qquad (19)$$

where

$$q(x'_1,\ldots,x'_N,y';h) = \chi \left[\Phi\left(x'_1,\ldots,x'_N,y'\right) - \gamma \right]$$
$$= \chi \left[\left(\sum_{k=1}^N \alpha_k u_k \right) + \beta u_s - \gamma \right] \qquad (20)$$

then the expected utility of all FL participants will satisfy

$$\left(\sum_{k=1}^N \alpha_k E_k \right) + \beta E_s - \gamma = 0 \qquad (21)$$

where the factor χ is a non-zero scalar, α_k ($k = 1, 2, \ldots, N$) and β are weight factors, and γ is a parameter.

Proof: The proof is given in Appendix C. ∎

Algorithm 1 Incentive Algorithm

Input: Threshold value Θ_k, deadline $\mathbb{T}$;
Input: The CZD strategies $q(y)$;
Output: Reward vector $\mathbf{r} = \{r_1, r_2, \ldots, r_n\}$.
1: Initialize parameters, post FL tasks and recruit devices
2: **for all** devices that have sent requests **do**
3:   **if** the sample data accuracy of the k-th device $\Xi_k \geq \Theta_k$ **then**
4:     Evaluate the k-th device's accuracy upper bounds h_k, lower bounds l_k, and weight ϖ_k, and then update Θ_k
5:   **end if**
6: **end for**
7: **for** $k = 1$ to N **do**
8:   Compute and update local models
9:   **if** $x_k \geq \Theta_k$ **then**
10:     Upload the local model to the server
11:   **end if**
12: **end for**
13: **while** $T_k \leq \mathbb{T}$ **do**
14:   Model aggregation $\theta \leftarrow \theta = \frac{1}{N} \sum_{k=1}^N (\varpi_k \cdot x_k)$
15:   **if** $\theta \geq \Theta$ **then**
16:     Calculate $\mathbf{r} = \{r_1, r_2, \ldots, r_n\}$ by the CZD strategies

---

17:   **end if**
18: **end while**
19: **return** Reward vector $\mathbf{r} = \{r_1, r_2, \ldots, r_n\}$.

When the server takes the CZD strategies as in (19) and (20), the social welfare could always be kept at the desired level, whatever the strategies of the devices are. Accordingly, the optimization problem to be solved can be expressed as

$$\boldsymbol{P2} : \max_{q(y)} \quad \bar{E} = \left(\sum_{k=1}^N \alpha_k E_k \right) + \beta E_s \quad \forall p_k(x_k)$$

$$\text{s.t.} \begin{cases} 0 \leq q(y) = q\left(x'_1,\ldots,x'_N,y';y\right) \leq 1 \\ \left(\sum_{k=1}^N \alpha_k E_k \right) + \beta E_s - \gamma = 0. \end{cases} \qquad (22)$$

By solving the problem $\boldsymbol{P2}$, the following theorem can be established.

Theorem 2: When the strategy of the server satisfies

$$\tilde{q}(h) = \begin{cases} \chi \left[\Phi\left(x'_1,\ldots,x'_N,y'\right) - \gamma \right] & \text{if} \quad y' \in [l,h) \\ \chi \left[\Phi\left(x'_1,\ldots,x'_N,y'\right) - \gamma \right] + 1 & \text{if} \quad y' = h \end{cases} \qquad (23)$$

the social welfare γ can reach a maximum value as

$$\gamma_{\max} = \begin{cases} \max\left(\Phi\left(x'_1,\ldots,x'_N,y'\right)\right) & \text{if} \quad \chi < 0 \\ \max\left(\Phi\left(x'_1,\ldots,x'_N,h'\right)\right) & \text{if} \quad \chi > 0. \end{cases} \qquad (24)$$

where $x'_k \in [l_k, h_k]$, $k = 1, 2, \ldots, N$, $y' \in [l, h]$.

Proof: The proof is given in Appendix D. ∎

At this point, social welfare can be controlled unilaterally to the value of $\gamma_{\max}$, regardless of the strategies adopted by the devices.

In summary, when the server applies the CZD strategies, it can unilaterally control the expected utility of all participants, which makes the social welfare of FL stable at a high level.

C. Incentive Algorithm

Considering the selfishness of the FL game players and the continuity of the adopted strategies, we propose a CZD-based incentive mechanism as follows with the illustration in Fig. 2 and the pseudo-code in Algorithm 1.

1) Initialization: The server initializes model parameters, publishes the FL task and recruits devices (Line 1).

2) Selection: In pre-training, if the accuracy of the k-th device's sample data satisfies the threshold value Θ_k, the server can evaluate its accuracy upper bound h_k, the lower bound l_k, and weight ϖ_k, and then update Θ_k (Lines 2−6).

3) Model Training: The selected devices calculate and update the local model by applying their high-quality data. If the trained local model meets the requirements, it can be sent to the server (Lines 7−12).

4) Model Aggregation and Reward Distribution: During local model aggregation, the k-th device's contribution x_k is recorded. And the server assigns the reward r_k by applying the CZD strategies (Lines 13−19).

Since all x_k and y are continuous bounded variables, the water-filling algorithm [39] in complexity of $O(N)$ is com-
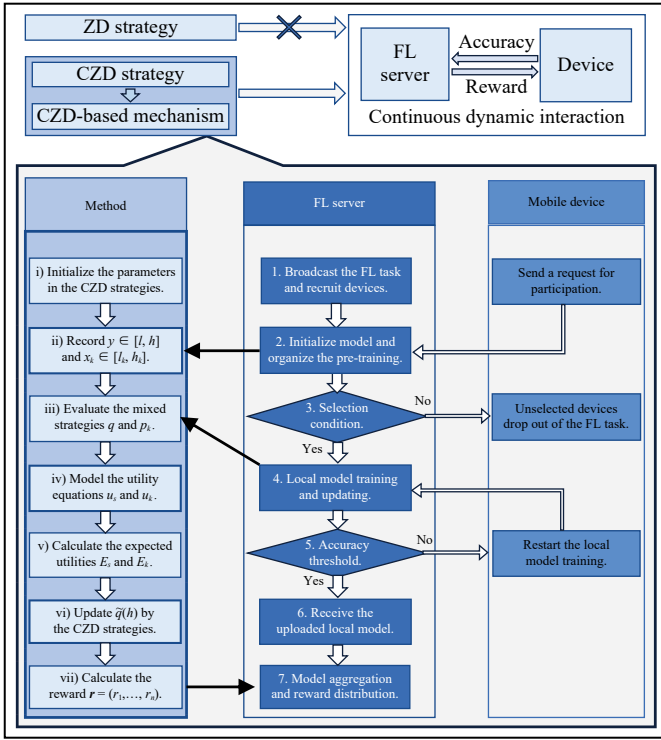
Fig. 2.   Flowchart of the CZD-based incentive mechanism.

monly used to approach the optimal solution. Since this algorithm is equivalent to a traversal of all participating devices, the complexity of Algorithm 1 is also $O(N)$.

V. SIMULATIONS

In this section, to evaluate the performance of the proposed incentive algorithm, the datasets used for the experimental tests are MNIST [40] and CIFAR-10 [41]. MNIST contains 10 000 test samples and 60 000 training samples. CIFAR-10 contains 10 000 test samples and 50 000 training samples.

The deep learning models to be trained are the multilayer-perceptron (MLP) [42] and ResNet-18 [43] models. The mini-batch stochastic gradient descent (SGD) [44] optimizer is used. The codes of the incentive algorithm are edited in Python 3.7.3 with PyTorch 1.2.0. The machine used for simulation is a desktop PC with an Intel Core i5 running at 2.7 GHz. To enhance the confidence of the numerical simulation and eliminate errors, the corresponding data in the correlation diagram is obtained by repeating the experiment 20 times and then taking the average values.

In the FL scenario with one server and 10 devices, the server applies the CZD strategies with $\tilde{q}(h) = \tilde{q}(x_1', \ldots, x_N', y'; h)$ satisfying (20) and (21). The scalars independent of the FL are set as $c_1 = 0.361$, $c_2 = 4.348$, $c_3 = 0.001$, $c_4 = 0.993$, $c_5 = 0.178$ and $c_6 = 1.743$. Other parameters in the simulation are listed in Table II.

A. Effectiveness of the CZD Strategies

To show the effectiveness of the CZD strategies, as displayed in Table III, the characteristics of the strategies CE [36] and MMZD [38] are compared with CZD. Note that both CE and MMZD are based on the discrete strategy scenarios,

TABLE II
PARAMETER SETTINGS

Parameter	Value	Parameter	Value
N	10	τ	8
ς_k	3	r_k	$[0, 10]$
σ_k, δ_k	0.5, 0.5	v	28
σ_s, δ_s	0.6, 0.4	S_k	0.5 MB
R	7	$\mathbb{T}$	5

TABLE III
COMPARISON OF CHARACTERISTICS

Strategy	Strategy scenario
CE [36]	Discrete with two players and with multiple players
MMZD [38]	Discrete with multiple players
Our CZD	Continuous with two players and with multiple players

while the CZD strategies are based on the continuous strategy scenarios with two players and multiple players, respectively.

Observing that MMZD has a similar setup to this paper, so we take MMZD as an example for performance comparison to show the effectiveness of CZD. In Table IV, we record the average accuracy of FL models based on different incentive mechanisms with different models and different datasets. To represent the performance of the different methods more concretely, we create real FL experiments by taking the MLP model based on the datasets MNIST and CIFAR-10, the results are represented in Tables V and VI, respectively. The data in Tables V and VI are the average accuracy of FL models, and running time.

TABLE IV
AVERAGE ACCURACY OF FL MODELS BASED ON DIFFERENT INCENTIVES

Model	Dataset	Epoch	Incentive method for FL	
			MMZD [38]	CZD
MLP	MNIST	10	75.83%	84.12%
		100	86.67%	93.57%
	CIFAR-10	10	53.23%	52.07%
		100	77.08%	84.49%
ResNet-18	MNIST	10	76.40%	86.63%
		100	89.70%	97.14%
	CIFAR-10	10	54.05%	51.91%
		100	79.82%	85.35%

In Table IV, it can be seen that CZD outperforms MMZD regardless of the dataset or training model applied. For example, when we choose model ResNet-18 with dataset CIFAR-10, CZD improves model accuracy by 5.53% over MMZD. The results in Tables V and VI are consistent with those in Table IV. In Table V, it can be seen that the proposed method improves model accuracy to about 90% after 20 epochs, and the difference in model accuracy between the two methods after 10 epochs of training is 8.29%. Based on our proposed

TABLE V
PERFORMANCE COMPARISON WITH MMZD BASED ON MNIST

Epoch	MMZD	CZD
5	(65.33%, 109 s)	(73.94%, 94 s)
10	(75.83%, 205 s)	(84.12%, 179 s)
20	(85.00%, 418 s)	(89.60%, 355 s)
50	(90.37%, 938 s)	(92.76%, 819 s)
100	(86.67%, 2222 s)	(93.57%, 1751 s)
150	(91.92%, 3187 s)	(95.02%, 2854 s)
200	(95.94%, 3993 s)	(98.33%, 3604 s)

TABLE VI
PERFORMANCE COMPARISON WITH MMZD BASED ON CIFAR-10

Epoch	MMZD	CZD
5	(46.21%, 752 s)	(44.17%, 647 s)
10	(53.23%, 1413 s)	(52.07%, 1233 s)
20	(57.76%, 2882 s)	(58.64%, 2447 s)
40	(60.47%, 5653 s)	(67.62%, 4968 s)
60	(65.21%, 8789 s)	(73.57%, 7775 s)
80	(74.31%, 12253 s)	(76.83%, 10494 s)
100	(77.08%, 15317 s)	(84.49%, 12069 s)

method, about 400 s of training time is saved in 200 epochs. Compared to MNIST, CIFAR-10 is more complex, which directly hinders the performance of FL during model training. In Table VI, CZD improves accuracy by 7% compared to MMZD in 100 epochs of FL and saves 3000 s of training time. In other words, when the player's strategy is extended from discrete to continuous, devices not only can choose whether to participate or not but also can control their contribution specifically in conjunction with some conditions such as revenues and costs. Therefore, compared with strategies in discrete scenarios, CZD strategies that cover more complete information will be more valuable in practical applications.

To confirm that the server can control any evolutionary device to fully cooperate after adopting the CZD strategies, it is now compared with the classical strategies, such as TFT (tit-for-tat), WSLS (win-stay-lose-shift) [45] and GTFT (generous TFT, TFT with forgiveness) [46].

In TFT, the player acts according to the opponent's previous actions. In WSLS, the player holds an action if it would lead to a high benefit; otherwise, its makes a switch. Similar to TFT, players who take GTFT will repeat their opponent's choice from the previous round, but have the probability to choose to forgive when betrayed. Here, we set this probability as 1/3.

The strategies of the devices and the server in the FL game can be divided into two cases: cooperation (C) and defection (D). For the k-th device, C corresponds to $x_k \in [\lambda_k, h_k]$, where $\lambda_k = \int_l^h \int_{l_k}^{h_k} (x_k \cdot z) \, dx_k dy \in [l_k, h_k]$. Similarly, for the server, C corresponds to $y \in [\mu, h]$, where $\mu = \int_l^h \int_{l_k}^{h_k} (y \cdot z) \, dx_k dy \in [l, h]$. Otherwise, the action corresponds to D.

Figs. 3 and 4 show the variation of social welfare with the global model epochs in the cases of $N = 1$ and $N = 10$, respec-

tively, where the server adopts the CZD strategies, TFT, WSLS and GTFT, and the devices apply TFT, WSLS and GTFT. For example, CZD vs TFT corresponds to the case where the server adopts the CZD strategies while the devices adopt TFT. As GTFT was shown to be effective in promoting cooperation in heterogeneous groups [46], we specifically compare GTFT with other strategies in Figs. 3(a) and 3(b).

In the two-player continuous strategy scenario, Fig. 3(a) shows that the social welfare is unstable when the devices adopt the TFT strategy while the server takes WSLS or TFT. If the devices take WSLS while the server adopts WSLS or TFT, the social welfare will be stabilized at a low level. When the server applies CZD strategies, social welfare will converge to a high and stable value, regardless of the strategies adopted by the devices. In Fig. 3(b), it can be seen that the forgiveness of GTFT does increase the social welfare of FL. However, social welfare converges to the optimal value only in the case of CZD vs GTFT and GTFT vs GTFT, and CZD is 70 epochs ahead of GTFT.

In Fig. 4, the number of devices participating in FL is expanded from 1 to 10. Compared with Fig. 3, social welfare converges to a higher stable value in all cases. In particular, when the server adopts the CZD strategies, social welfare increases from 10 to 16 as N changes. GTFT can be effective in promoting cooperation, but forgiveness may hurt adopters' gains. Similar to Fig. 3(b), in Fig. 4(b), the performance of GTFT is better than that of TFT and WSLS. Social welfare can reach an optimal value in the case of CZD vs GTFT and GTFT vs GTFT. However, CZD shows a stronger control effect than other strategies.

B. Fairness of the CZD Strategies

To prove the fairness of CZD strategies, Fig. 5 describes the stable values of FL participants' relative utilities when the server and devices takes different strategies. The value of the relative utility is equal to the quotient of the actual utility divided by the utility in ALLC (all cooperative).

Figs. 5(a) and 5(b) show that only the CZD strategies can make the relative utility of the devices and the server steadily go towards 1. As for GTFT, the relative utility of FL participants converge to 1 only when the GTFT player faces opponents who adopt the same strategy. The strategies TFT and WSLS push the relative utility below 1 and can even reduce the server's revenue.

Fig. 6 shows the variation of the relative utility when all devices take the WSLS strategy and the server adopts different strategies. Similar to Fig. 5, the relative utilities of FL participants steadily converge to 1 only when the server adopts the CZD strategies. For GTFT, the relative utility can converge to 1 only if it encounters an opponent who adopts the same strategy, otherwise it will be below 1. The strategies TFT and WSLS make the relative utility tend to stabilize at a value lower than 1.

C. Effect of System Parameters

As seen in Section IV, system parameters play an essential role in regulating the utility relationship between the devices and the server. To reveal the effect of the system parameters
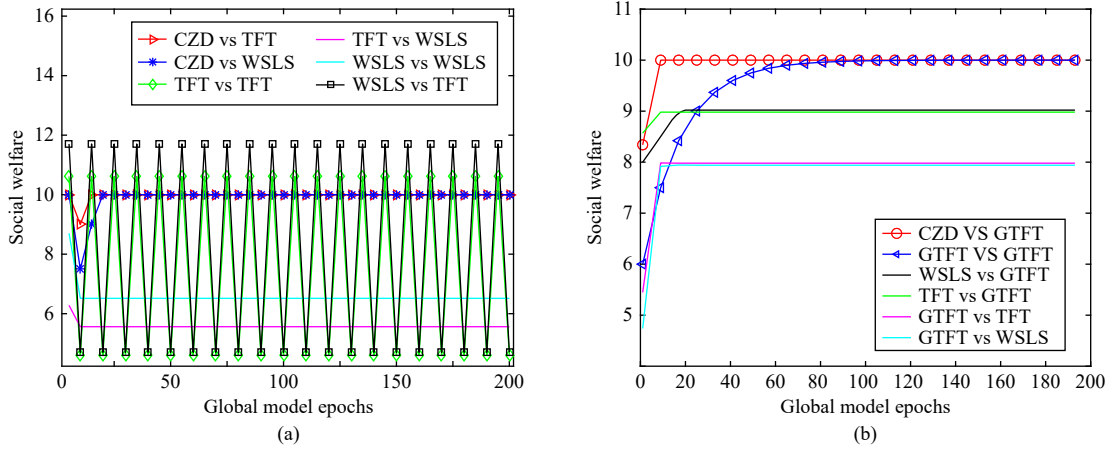
Fig. 3. Social welfare with global model epochs in the continuous strategy setting when $N = 1$. (a) Comparison of CZD, TFT, and WSLS; (b) Comparison of CZD, TFT, WSLS, and GTFT.
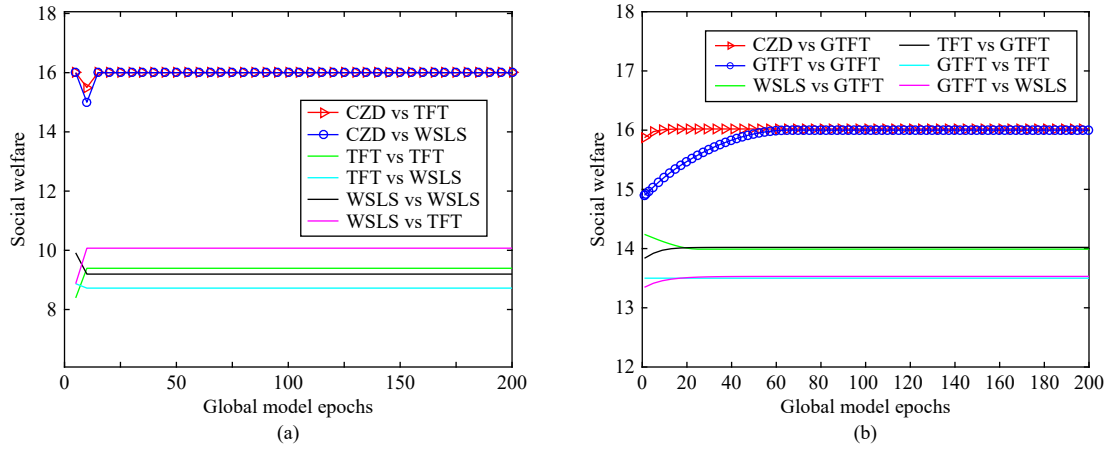

Fig. 4. Social welfare with global model epochs in the continuous strategy setting when $N = 10$. (a) Comparison of CZD, TFT, and WSLS; (b) Comparison of CZD, TFT, WSLS, and GTFT.
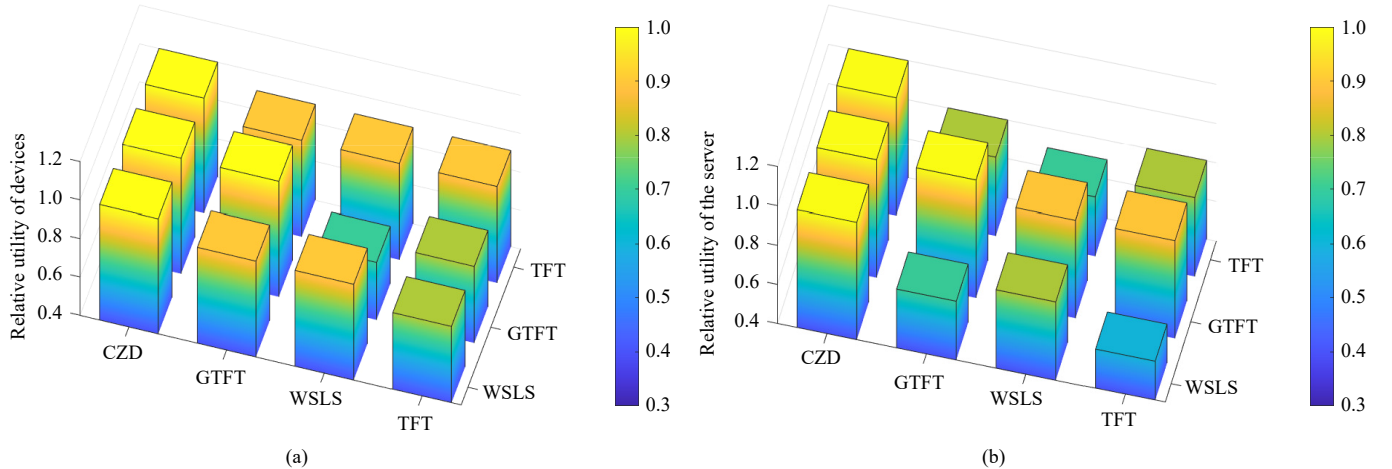

Fig. 5. Stable values of FL participants' relative utilities when the server and devices adopt different strategies, where the strategies of the server are CZD, GTFT, WSLS and TFT, and the strategies of devices are WSLS, GTFT and TFT. (a) Relative utility of devices; (b) Relative utility of the server.

on the FL game, Fig. 7 shows the variation of social welfare with the global model epochs when the system parameters take different values. It can be observed that the change of the χ value will lead to fluctuation of social welfare, while the change of the β value will affect the convergence rate of social welfare. However, the social welfare of FL will eventually stabilize at an identical level. This is because the CZD strategies can regulate the actions of the devices and make them contribute their high-accuracy data, which in turn guarantees that the expected utility of all participants tends to be stable,
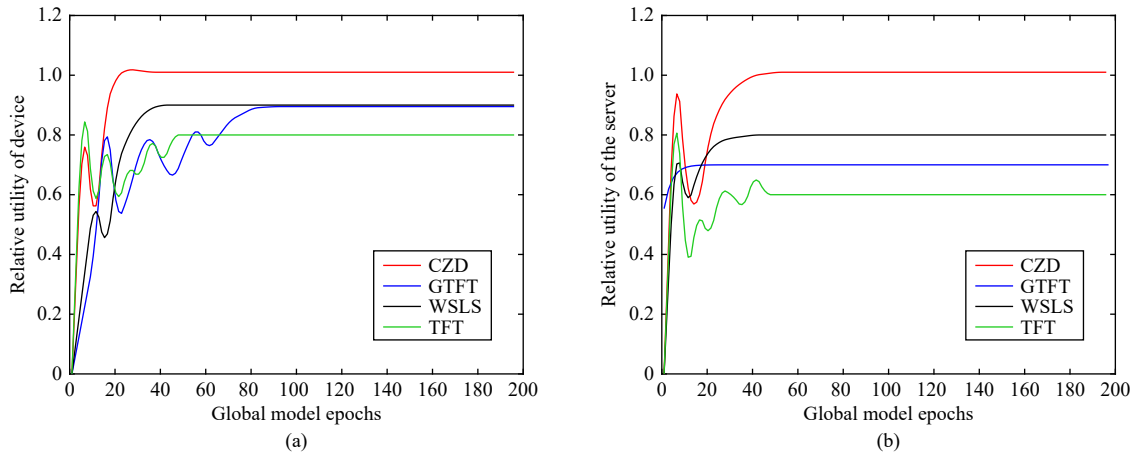
Fig. 6.   Variation of the FL participants' relative utility when all devices take the WSLS strategy and the server adopts different strategies. (a) Variation of the devices' relative utility; (b) Variation of the server's relative utility.
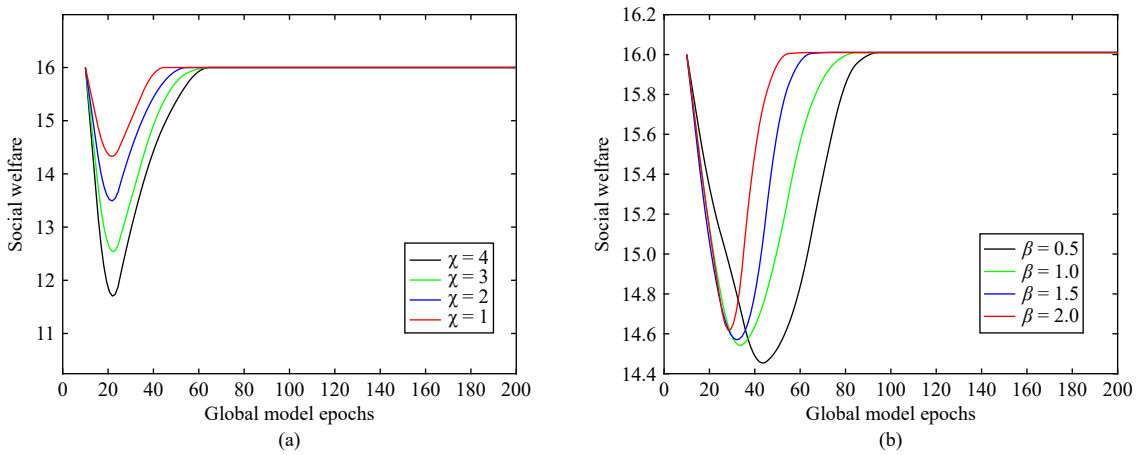

Fig. 7.   Variation of the social welfare with global model epochs. (a) When the factor χ takes different values; (b) When β takes different values.

resulting in the social welfare converging to the same level.

In summary, the designed CZD-based incentive algorithm has several advantages. Firstly, our proposed method can be effective in avoiding anyone to increase their own interests by harming the collective interests. Secondly, the CZD strategies extend the action spaces of the ZD strategies from discrete to continuous. Compared with the discrete action space, the continuous action space covers all incomplete cooperative behaviors including full cooperation and full defection. Thirdly, in FL, the continuous strategy contains more complete information than the discrete strategy. As the objects considered by the CZD strategies are more comprehensive, the CZD-based incentive algorithm can more effectively motivate the full contribution of FL participants. Fourthly, the fair distribution of rewards can be achieved by adjusting the system parameters.

VI.   CONCLUSION AND DISCUSSION

In this paper, to investigate the issue of incentive mechanisms in FL, we model the interaction between the devices and the server as a continuous multi-player iterative game. We apply the CZD strategies to the continuous strategy scenarios with two players and multiple players. Then, we demonstrate

that the server can unilaterally control social welfare regardless of the strategies adopted by the devices. Based on this, we develop a CZD-based incentive algorithm which can motivate as many devices as possible to participate in FL and contribute high-quality data. Furthermore, by comparing strategies of TFT, GTFT and WSLS in game theory and other methods in existing works, we present some simulations to verify the advantages of the CZD-based incentive algorithm. Our proposed approach can attract devices to contribute all their high-accuracy data in FL, and then effectively optimize the social welfare and improve the efficiency of FL.

It is noticed that we have considered the case of one server in this paper, but the situation is more complex in real applications. As the number of servers increases, the interaction among FL participants will become more complex, and the competition among FL participants will also become more competitive. Thereby, the malicious competitive behavior among servers will also become more prominent, and even pose a serious threat to the sustainability of the FL system. Furthermore, the proposed CZD-based incentive mechanism can motivate participants to invest in FL, but has limited effect on the optimization of the learning algorithm. Thus, it is meaningful to consider the balance between the incentive

effect and learning efficiency.

In the future, we intend to expand the number of servers from one to multiple, and consider dynamic interactions between multiple servers and multiple devices in FL, which corresponds to a more complex network structure. Under this scene, we will consider applying the ZD alliance theory to overcome the effect of increased servers and avoid malicious competitive behaviors. In addition, we will consider using the reinforcement learning method to optimize the strategies selection in the FL process of decision making, where the balance between the incentive effect and the learning efficiency can be achieved.

APPENDIX

A. Proof of Lemma 1

Proof: Firstly, a sufficiently small number ε_k is introduced to divide the k-th device's strategy x_k into l_k, $l_k + \varepsilon_k$, …, $l_k + n\varepsilon_k = h_k$. This is equivalent to dividing the data used by the device for the FL training task into n parts, where the size of each part is $\varepsilon_k = \frac{h_k - l_k}{n}$.

Similarly, a sufficiently small number ε is introduced to divide the server's strategy y into l, $l + \varepsilon, \ldots, l + n\varepsilon = h$. This is equivalent to dividing the rewards that can be set by the server into n parts, and the size of each part is $\varepsilon = \frac{h-l}{n}$. When n is sufficiently large, or ε_k and ε approach zero, one can consider the strategy equations of devices and the server as continuous versions.

The conditional probability of the devices can be defined as g_{ij}^k, and the conditional probability of the server can be defined as h_{ij}^k, $i, j, k \in \{1, 2, \ldots, n\}$, where in the previous round the strategy function of the devices and the server are $(l_k + i\varepsilon_k)$ and $(l + j\varepsilon)$, respectively. In the current round, the strategy function of the server is $(l + k\varepsilon)$.

According to the strategy functions of the devices and the server defined with various initial conditions, the process of the two-player iterative game can be regarded as a stochastic process with the characteristics of a Markov chain.

By introducing a stationary vector $\boldsymbol{v}$, the state transition matrix of the Markov chain can be defined as $\bar{\boldsymbol{M}}$, so that

$$\bar{\boldsymbol{M}} = [\boldsymbol{H}_{11}, \ldots, \boldsymbol{H}_{1n}, \boldsymbol{H}_{21}, \ldots, \boldsymbol{H}_{2n}, \ldots, \boldsymbol{H}_{n1}, \ldots, \boldsymbol{H}_{nn}]$$

and

$$\boldsymbol{v}^T \cdot \bar{\boldsymbol{M}} = \boldsymbol{v}^T$$

where $\boldsymbol{H}_{ij}$, $i, j \in \{1, 2, \ldots, n\}$, is the conditional probability that the devices take strategy $(l_k + i\varepsilon_k)$ and the server takes strategy $(l + j\varepsilon)$ in the current round, given all the outcomes of the previous round. Thus, for $\forall i, j \in \{1, 2, \ldots, n\}$, one has

$$\boldsymbol{H}_{ij} = \left[g_{11}^i h_{11}^j, \ldots, g_{1n}^i h_{n1}^j, \ldots, g_{n1}^i h_{1n}^j, \ldots, g_{nn}^i h_{nn}^j \right]^T.$$

The payoff matrix of the device can be defined as

$$\boldsymbol{U_k} = \left[u_k^{11}, \ldots, u_k^{n1}, u_k^{12}, \ldots, u_k^{n2}, \ldots, u_k^{1n}, \ldots, u_k^{nn} \right]^T$$

and the payoff matrix of the server can be defined as

$$\boldsymbol{U_s} = \left[u_s^{11}, \ldots, u_s^{n1}, u_s^{12}, \ldots, u_s^{n2}, \ldots, u_s^{1n}, \ldots, u_s^{nn} \right]^T.$$

Consequently, the expected utility of the devices and the server is $E_k' = \boldsymbol{v}^T \cdot \boldsymbol{U_k}$ and $E_s' = \boldsymbol{v}^T \cdot \boldsymbol{U_s}$, respectively.

Suppose that $\bar{\boldsymbol{M}}' = \bar{\boldsymbol{M}} - \boldsymbol{I}$, where $\boldsymbol{I}$ is the unit diagonal matrix. Then, $\boldsymbol{v}^T \bar{\boldsymbol{M}}' = \boldsymbol{0}$.

According to Cramer's rule, one has

$$adj(\bar{\boldsymbol{M}}') \bar{\boldsymbol{M}}' = \det(\bar{\boldsymbol{M}}') \boldsymbol{I} = \boldsymbol{0} \tag{A1}$$

where $adj(\bar{\boldsymbol{M}}')$ is the adjugate matrix of $\bar{\boldsymbol{M}}'$.

Accordingly, every row of $adj(\bar{\boldsymbol{M}}')$ is proportional to $\boldsymbol{v}$. Therefore, the first row of $adj(\bar{\boldsymbol{M}}')$ can be replaced by $\boldsymbol{v}$. Furthermore, $\sum_{k=1}^{n} g_{ij}^k = 1$.

Here, a determinant is used to express the dot product of $\boldsymbol{f} = [f_1, f_2, \ldots, f_{n^N}]^T$ with the stationary vector $\boldsymbol{v}$. By using an elementary transformation and introducing a vector

$$\boldsymbol{h_{nn}} = \left[h_{11}^n, \ldots h_{(n-1)n}^n, h_{n1}^n - 1, \ldots, h_{nn}^n - 1 \right]^T$$

it can be seen that the result depends on the strategy of only one participant, i.e.,

$$\boldsymbol{v}^T \cdot \boldsymbol{f} = D(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}) = \det \begin{pmatrix} f_1 & \cdots & h_{11}^n \\ \vdots & \vdots & \vdots \\ f_{(n-1)n} & \cdots & h_{(n-1)n}^n \\ f_{(n-1)n+1} & \cdots & h_{n1}^n - 1 \\ \vdots & \vdots & \vdots \\ f_{nn} & \cdots & h_{nn}^n - 1 \end{pmatrix}.$$

When $\boldsymbol{f} = \alpha \boldsymbol{U_k} + \beta \boldsymbol{U_s} - \gamma \boldsymbol{1}$, with α and β as weight factors, and γ is a nonzero scalar, one has

$$\boldsymbol{v} \cdot \boldsymbol{f} = \boldsymbol{v} (\alpha \boldsymbol{U_k} + \beta \boldsymbol{U_s} - \gamma \boldsymbol{1}) = \alpha E_k' + \beta E_s' - \gamma. \tag{A2}$$

When

$$\boldsymbol{h_{nn}} = \chi (\alpha \boldsymbol{U_k} + \beta \boldsymbol{U_s} - \gamma \boldsymbol{1}) \tag{A3}$$

using properties of the determinant, one obtains

$$\alpha E_k' + \beta E_s' - \gamma = 0. \tag{A4}$$

When ε_k and ε approach zero, since the strategy functions of the devices and the server are continuous, the expected utility of the device E_k' approaches E_k, and the expected utility of the server E_s' approaches E_s, i.e., $\alpha E_k + \beta E_s - \gamma = 0$. ∎

B. Proof of Theorem 1

Proof: Based on the existing conditions and the strategy of the server, the problem $\boldsymbol{P1}$ can be solved by solving the following problem:

$$\max_{q(y)} \gamma$$

$$\text{s.t.} \begin{cases} 0 \leq q\left(x_k', y'; y\right) \leq 1 \\ \bar{q}(x_k', y'; h) = \chi \left[\alpha u_k\left(x_k', y'\right) + \beta u_s\left(x_k', y'\right) - \gamma \right] \\ \chi \neq 0. \end{cases} \tag{B1}$$

If $\chi < 0$, noting the constraints $q(x_k', y'; y) \leq 1$, and setting $u_k^0 = u_k(x_k^0, y^0)$, $u_s^0 = u_s(x_k^0, y^0)$, one has

$$\gamma_{\max} = \min(w_1), \quad \forall x_k^0 \in [l_k, h_k], \forall y^0 \in [l, h]$$

where

$$w_1 = \begin{cases} \dfrac{1}{\chi} + \alpha u_k^0 + \beta u_s^0 & y^0 \in [l, h) \\ \alpha u_k^0 + \beta u_s^0 & y^0 = h. \end{cases}$$

Noting the constraint $q(x_k', y'; y) \geq 0$, and setting $u_k^1 = u_k(x_k^1, y^1)$, $u_s^1 = u_s(x_k^1, y^1)$, one has

$$\gamma_{\min} = \max(w_2), \quad \forall x_k^1 \in [l_k, h_k], \forall y^1 \in [l, h]$$

where

$$w_2 = \begin{cases} \alpha u_k^1 + \beta u_s^1 & y^1 \in [l, h) \\ -\dfrac{1}{\chi} + \alpha u_k^1 + \beta u_s^1 & y^1 = h. \end{cases}$$

Obviously, $\gamma_{\min} \leq \gamma_{\max}$; that is,

$$\max(w_2) \leq \min(w_1), \forall x_k^0, x_k^1 \in [l_k, h_k], \forall y^0, y^1 \in [l, h].$$

Since χ can be any negative number with a small absolute value, one obtains

$$\gamma_{\max} = \max(\alpha u_k^1 + \beta u_s^1), \forall x_k^1 \in [l_k, h_k], \forall y^1 \in [l, h].$$

Similarly, if $\chi > 0$, then

$$\gamma_{\max} = \max(\alpha u_k(x_k^0, h) + \beta u_s(x_k^0, h)), \forall x_k^0 \in [l_k, h_k].$$

In summary, the maximum value of γ is

$$\gamma_{\max} = \begin{cases} \max(\alpha u_k^1 + \beta u_s^1) & \text{if } \chi < 0 \\ \max(\alpha u_k(x_k^0, h) + \beta u_s(x_k^0, h)) & \text{if } \chi > 0. \end{cases} \quad \text{(B2)}$$

Therefore, to make social welfare stable at a high value, the CZD strategies adopted by the server need to satisfy

$$\bar{q}(x_k', y'; h) = \begin{cases} \chi(\alpha u_k^0 + \beta u_s^0 - \gamma_{\max}) & \text{if } y' \in [l, h] \\ \chi(\alpha u_k^0 + \beta u_s^0 - \gamma_{\max}) + 1 & \text{if } y' = h. \end{cases} \quad \text{(B3)}$$

∎

C. Proof of Lemma 2

Proof: If there are $(N-1)$ devices and one server in the FL game, according to the proof of Lemma 1, the conditional probability that the strategy of the k-th device is $l_k + k\varepsilon_k$ in the current round can be defined as

$$g_k^i = [g_1^{k,i}, g_2^{k,i}, \ldots, g_{n^N}^{k,i}]^T$$

for all $i \in \{1, 2, \ldots, n\}$, $k \in \{1, 2, \ldots, N-1\}$, and the conditional probability that the strategy function of the server is $l + k\varepsilon$ in the current round can be defined as

$$h^j = [h_1^j, h_2^j, \ldots, h_{n^N}^j]^T, \forall j \in \{1, 2, \ldots, n\}.$$

Therefore, the transition matrix M_N of the Markov chain is

$$M_N = [G_1, \ldots, G_n, G_{n+1}, \ldots, G_{n^2}, \ldots, G_{n^N}]$$

and the stationary vector v of the Markov chain is

$$v^T \cdot M_N = v^T$$

where

$$G_{t_0 t_1 \cdots t_{N-1}} = h^{t_0} \circ g_1^{t_1} \circ \cdots \circ g_{N-1}^{t_{N-1}}$$

for all $t_k \in \{1, 2, \ldots, n\}$, $\forall k \in \{1, 2, \ldots, N-1\}$, where $\circ$ is the Hadamard product in matrix multiplication.

Let $M_N' = M_N - I$, thus $v M_N' = 0$. Based on the Cramer's rule, one obtains

$$adj(M_N') M_N' = \det(M_N') = \mathbf{0} \quad \text{(C1)}$$

where every row of $adj(M_N')$ is proportional to v. Therefore, the first row of $adj(M_N')$ can be replaced by v. Here, the determinant is used to express the dot product of $f = [f_1, f_2, \ldots, f_{n^N}]^T$ with the stationary vector v.

By using an elementary transformation and introducing the vector

$$\bar{h}^n = [h_1^n, \ldots, h_{n^{(N-1)}}^n, h_{n^{(N-1)}+1}^n - 1, \ldots, h_{n^N}^n - 1]^T$$

it can be seen that the result depends on the strategy of only one participant, i.e.,

$$v^T \cdot f = \det(f \quad \cdots \quad \bar{h}^n). \quad \text{(C2)}$$

The payoff matrix of the k-th device can be defined as

$$U_k = [u_k^1, u_k^2, \ldots, u_k^{n^N}]^T$$

and the payoff matrix of the server can be defined as

$$U_s = [u_s^1, u_s^2, \ldots, u_s^{n^N}]^T.$$

Therefore, the utility of the k-th device is

$$E_k' = v^T \cdot U_k.$$

$k \in \{1, 2, \ldots, N-1\}$, and the utility of the server is

$$E_s' = v^T \cdot U_s.$$

When the server takes strategy $\bar{h}^n$, and $f = \alpha_1 U_1 + \cdots + \alpha_{N-1} U_{N-1} + \beta U_s - \gamma \mathbf{1}$, with $\alpha_1, \ldots, \alpha_{N-1}$, β and γ as scalars, one has

$$v^T \cdot f = v^T(\alpha_1 U_1 + \cdots + \alpha_N U_{N-1} + \beta U_s - \gamma \mathbf{1})$$
$$= \alpha_1 E_1' + \cdots + \alpha_N E_{N-1}' + \beta E_s' - \gamma = 0. \quad \text{(C3)}$$

When ε_k and ε approach zero, since the strategy functions of the devices and the server are continuous, E_k' approaches E_k, $\forall k = 1, 2, \ldots, N-1$, and E_s' approaches E_s, i.e.,

$$\alpha_1 E_1 + \cdots + \alpha_N E_{N-1} + \beta E_s - \gamma = 0. \quad \text{(C4)}$$

In addition, when there are N devices with one server, the conditional probability that the strategy functions of the k-th devices and the server are $(l_k + i\varepsilon_k)$ and $(l + j\varepsilon)$ in the current round can be defined as

$$g_k^i = [g_1^{k,i}, g_2^{k,i}, \ldots, g_{n^{N+1}}^{k,i}]^T$$

and

$$h^j = [h_1^j, h_2^j, \ldots, h_{n^{N+1}}^j]^T$$

where $i, j \in \{1, 2, \ldots, n\}$, $k \in \{1, 2, \ldots, N\}$.

Based on the definition, when the number of players in the FL game is $(N+1)$, one has

$$M_{N+1} = M_N \otimes g_N$$

and

$$v^T \cdot M_{N+1} = v^T$$

where M_{N+1} is the Markov chain transition matrix, v is a stationary vector, $g_N = [g_1, g_2, \ldots, g_n]^T$, and $\otimes$ is the Kronecker product.

Let $M'_{N+1} = M_{N+1} - I$. Then, $v \cdot M'_{N+1} = 0$. Based on the Cramer's rule, one obtains

$$adj\left(M'_{N+1}\right) M'_{N+1} = \det\left(M'_{N+1}\right) = \mathbf{0} \tag{C5}$$

where every row of $adj(M'_{N+1})$ is proportional to v. Therefore, the first row of $adj(M'_{N+1})$ can be used to replace v.

In addition, define a vector by

$$f' = [f_1, f_2, \ldots, f_{n^{N+1}}]^T = \left[f'_1, f'_2, \ldots, f'_n\right]^T$$

where

$$f'_k = \left[f_{(k-1)n^N+1}, f_{(k-1)n^N+2}, \ldots, f_{(k-1)n^N+n^N}\right].$$

As $\sum_{t=1}^{n} g_t = 1$, with some elementary transformations, the dot product of any vector f' and stationary vector v can be expressed as

$$v^T \cdot f' = \det(M_N \otimes g_N - I)$$

$$= \det\left(\begin{bmatrix} f'_1 & G_2 & \cdots & G_{n^N} \\ \vdots & \vdots & \ddots & \vdots \\ f'_n & G_2 & \cdots & G_{n^N} \end{bmatrix} - \begin{bmatrix} \mathbf{1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{1} \end{bmatrix}\right)$$

where the $\mathbf{1}$ is a $1 \times n^{N-1}$-dimensional vector and all its elements are 1.

Therefore, in the FL game with N players, one can use a determinant to express the dot product of f' and stationary vector v, where one can see that the column depends only on the strategy of the server, i.e.,

$$v^T \cdot f' = \det(f \quad \cdots \quad \tilde{h}^n) \tag{C6}$$

where $\tilde{h}^n = [h_1^n, h_2^n, \ldots, h_{n^N}^n, h_{n^N+1}^n - 1, \ldots, h_{n^{N+1}}^n - 1]^T$.

The payoff matrix of the k-th device is

$$U_k = \left[u_k^1, u_k^2, \ldots, u_k^{n^{N+1}}\right]^T$$

and the payoff matrix of the server is

$$U_s = \left[u_s^1, u_s^2, \ldots, u_s^{n^{N+1}}\right]^T.$$

Therefore, the utility of the k-th device is $E'_k = v^T \cdot U_k$, $k \in \{1, 2, \ldots, N\}$, and the utility of the server is $E'_s = v^T \cdot U_s$.

When the server takes strategy $\bar{h}^n$, and $f' = \alpha_1 U_1 + \cdots + \alpha_N U_N + \beta U_s - \gamma \mathbf{1}$, with $\alpha_1, \ldots, \alpha_N, \beta$ and γ as scalars, one has

$$v^T \cdot f' = v^T (\alpha_1 U_1 + \cdots + \alpha_N U_N + \beta U_s - \gamma \mathbf{1})$$

$$= \alpha_1 E'_1 + \cdots + \alpha_N E'_N + \beta E'_s - \gamma = 0. \tag{C7}$$

When ε and ε_k approach 0, since the strategy functions of the devices and the server are continuous, E'_k approaches E_k, $\forall k = 1, 2, \ldots, N$ and E'_s approaches E_s, i.e.,

$$\alpha_1 E_1 + \cdots + \alpha_N E_N + \beta E_s - \gamma = 0. \tag{C8}$$

Based on mathematical induction, it follows that in N-player FL games ($N \geq 2$), when the server adopts the CZD strategies, it can control the expected utility of all participants unilaterally, which can keep the social welfare at a high and stable level.   ∎

D. Proof of Theorem 2

The problem **P2** can be solved by solving the following problem:

$$\max_{q(y)} \gamma$$

$$\text{s.t.} \begin{cases} 0 \leq q\left(x'_1, \ldots, x'_N, y'; y\right) \leq 1 \\ \tilde{q}(x'_1, \ldots, x'_N, y'; h) = \chi\left[\Phi\left(x'_1, \ldots, x'_N, y'\right) - \gamma\right] \\ \chi \neq 0. \end{cases} \tag{D1}$$

Similar to the Proof of **Theorem 1**, one has

$$\gamma_{\max} = \begin{cases} \max\left(\Phi\left(x'_1, \ldots, x'_N, y'\right)\right) & \text{if} \quad \chi < 0 \\ \max\left(\Phi\left(x'_1, \ldots, x'_N, h'\right)\right) & \text{if} \quad \chi > 0. \end{cases} \tag{D2}$$

Therefore, to optimize the social welfare of FL, the CZD strategies adopted by the server need to satisfy

$$\tilde{q}\left(x'_1, \ldots, x'_N, y'; h\right)$$

$$= \begin{cases} \chi\left(\Phi\left(x'_1, \ldots, x'_N, y'\right) - \gamma\right) & \text{if} \quad y' \in [l, h) \\ \chi\left(\Phi\left(x'_1, \ldots, x'_N, y'\right) - \gamma\right) + 1 & \text{if} \quad y' = h. \end{cases} \tag{D3}$$

where $x'_k \in [l_k, h_k]$, $k = 1, 2, \ldots, N$, $y' \in [l, h]$.   ∎

REFERENCES

[1] K. Letaief, W. Chen, Y. Shi, J. Zhang, and Y. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[2] C. Zhou, H. Tao, Y. Chen, V. Stojanovic, and W. Paszke, "Robust point-to-point iterative learning control for constrained systems: A minimum energy approach," *Int. J. Robust Nonlinear Control*, vol. 32, no. 18, pp. 10139–10161, Sept. 2022.

[3] H. Tao, J. Qiu, Y. Chen, V. Stojanovic, and L. Cheng, "Unsupervised cross-domain rolling bearing fault diagnosis based on time-frequency information fusion," *J. Frankl. Inst.*, vol. 360, no. 2, pp. 1454–1477, Jan. 2023.

[4] Z. Zhuang, H. Tao, Y. Chen, V. Stojanovic, and W. Paszke, "An optimal iterative learning control approach for linear systems with nonuniform trial lengths under input constraints," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 53, no. 6, pp. 3461–3473, Jun. 2023.

[5] N. Aitzhan and D. Svetinovic, "Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams," *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 5, pp. 840–852, Sep.–Oct. 2018.

[6] L. Liu, J. Zhang, S. Song, and K. Letaief, "Client-edge-cloud hierarchical federated learning," in *Proc. IEEE Int. Conf. Commun.*, Dublin, Ireland, 2020, pp. 1–6.

[7] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, "Federated learning," *Synth. Lect. Artif. Intell. Mach. Learn.*, vol. 13, no. 3, pp. 1–207, Dec. 2019.

[8] A. Hard, K. Rao, R. Mathews, S. Ramaswamy, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage, "Federated learning for mobile keyboard prediction," 2018, [Online]. Available: https https://arxiv.org/abs/1811.03604.

[9] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: conceptand applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, Jan. 2019.

[10] L. Khan, S. Pandey, N. Tran, W. Saad, Z. Han, M. Nguyen, and C. Hong, "Federated learning for edge networks: Resource optimization and incentive mechanism," *IEEE Commun. Mag.*, vol. 58, no. 10,

pp. 88–93, Oct. 2020.

[11] W. Lim, J. Huang, Z. Xiong, J. Kang, D. Niyato, X. Hua, C. Leung, and C. Miao, "Towards federated learning in uavenabled internet of vehicles: A multi-dimensional contractmatching approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5140–5154, Aug. 2021.

[12] N. Ding, Z. Fang, and J. Huang, "Optimal contract design for efficient federated learning with multi-dimensional private information," *IEEE J. Sel. Areas. Commun.*, vol. 39, no. 1, pp. 186–200, Jan. 2021.

[13] N. Tran, W. Bao, A. Zomaya, M. Nguyen, and C. Hong, "Federated learning over wireless networks: Optimization model design and analysis," in *Proc. IEEE INFOCOM*, Paris, France, 2019, pp. 1387–1395.

[14] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, and D. Niyato, "Federated learning for 6G communications: Challenges, methods, and future directions," *China Commun.*, vol. 17, no. 9, pp. 105–118, Sept. 2020.

[15] H. Yu, Z. Liu, Y. Liu, T. Chen, M. Cong, X. Weng, D. Niyato, and Q. Yang, "A fairness-aware incentive scheme for federated learning," in *Proc. AAAI/ACM Conf. AI, Ethics, and Society*, pp. 393–399, Feb. 2020.

[16] X. Wang, C. Wang, X. Li, V. Leung, and T. Taleb, "Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9441–9455, Oct. 2020.

[17] L. Xue, C. Sun, D. Wunsch, Y. Zhou, and F. Yu, "An adaptive strategy via reinforcement learning for the prisoner's dilemma game," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 1, pp. 301–310, Jan. 2018.

[18] J. Wang, Y. Hong, J. Wang, J. Xu, Y. Tang, Q. Han, and J. Kurths, "Cooperative and competitive multi-agent systems: From optimization to games," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 5, pp. 763–783, May 2022.

[19] P. Y. Zhang, M. C. Zhou, C. X. Li, and A. Abusorrah, "Dynamic evolutionary game-based modeling, analysis and performance enhancement of blockchain channels," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 1, pp. 188–202, Jan. 2023.

[20] 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.

[21] W. Lim, Z. Xiong, C. Miao, D. Niyato, Q. Yang, C. Leung, and H. Poor, "Hierarchical incentive mechanism design for federated machine learning in mobile networks," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9575–9588, Oct. 2020.

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

[23] T. Thi Le, N. Tran, Y. Tun, M. Nguyen, S. Pandey, Z. Han, and C. Hong, "An incentive mechanism for federated learning in wireless cellular networks: An auction approach," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 8, pp. 4874–4887, Aug. 2021.

[24] S. Fan, H. Zhang, Z. Wang, and W. Cai, "Mobile devices strategies in blockchain-based federated learning: A dynamic game perspective," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 3, pp. 1376–1388, May-Jun. 2023.

[25] Q. Zhang, J. Zhu, S. Gao, Z. Xiong, Q. Ding, and G. Piao, "Incentive mechanism for federated learning based on blockchain and Bayesian game," *Sci. Sin. Inform.*, vol. 52, no. 6, pp. 971–991, Jun. 2022.

[26] Z. Su, Y. Wang, T. Luan, N. Zhang, F. Li, T. Chen, and H. Cao, "Secure and efficient federated learning for smart grid with edge-cloud collaboration," *IEEE Trans. Industr. Inform.*, vol. 18, no. 2, pp. 1333–1344, Feb. 2022.

[27] Y. Jiao, P. Wang, D. Niyato, B. Lin, and D. Kim, "Toward an automated auction framework for wireless federated learning services market," *IEEE Trans. Mob. Comput.*, vol. 20, no. 10, pp. 3034–3048, Oct. 2021.

[28] 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.

[29] Y. Zhan and J. Zhang, "An incentive mechanism design for efficient edge learning by deep reinforcement learning approach," in *Proc. IEEE INFOCOM*, Toronto, Canada, 2020, pp. 2489–2498.

[30] W. Press and F. Dyson, "Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 26, pp. 10409–10413, May 2012.

[31] C. Hilbe, A. Wu, A. Traulsen, and M. Nowak, "Cooperation and control in multiplayer social dilemmas," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 46, pp. 16425–16430, Oct. 2014.

[32] H. Zhang, D. Niyato, L. Song, T. Jiang, and Z. Han, "Zero-determinant strategy for resource sharing in wireless cooperations," *IEEE Trans. Wirel. Commun.*, vol. 15, no. 3, pp. 2179–2192, Mar. 2016.

[33] M. Wu and C. Tang, "Price of fairness for zero-determinant strategies in iterated prisoner's dilemma," in *Proc. IEEE Chinese Control Conf.*, Shanghai, China, 2021, pp. 770–775.

[34] R. Tan, Q. Su, B. Wu, and L. Wang, "Payoff control in repeated games," in *Proc. IEEE Chin. Control Decis. Conf.*, Kunming, China, 2021, pp. 997–1005.

[35] C. Tang, C. Li, X. Yu, Z. Zheng, and Z. Chen, "Cooperative mining in blockchain networks with zero-determinant strategies," *IEEE Trans. Cybern.*, vol. 50, no. 10, pp. 4544–4549, Oct. 2020.

[36] Q. Hu, S. Wang, Z. Xiong, and X. Cheng, "Nothing wasted: Full contribution enforcement in federated edge learning," *IEEE Trans. Mob. Comput.*, vol. 22, no. 5, pp. 2850–2861, May 2023.

[37] Q. Hu, Z. Wang, M. Xu, and X. Cheng, "Blockchain and federated edge learning for privacy-preserving mobile crowdsensing," *IEEE Internet Things J.*, vol. 10, no. 14, pp. 12000–12011, Jul. 2023.

[38] J. Chen, Q. Hu, and H. Jiang, "Social welfare maximization in cross-silo federated learning," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Singapore, 2022, pp. 4258–4262.

[39] W. Yu, G. Ginis, and J. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas. Commun.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[40] L. Deng, "The MNIST database of handwritten digit images for machine learning research," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, Nov. 2012.

[41] Y. Abouelnaga, O. Ali, H. Rady, and M. Moustafa, "CIFAR-10: KNN-based ensemble of classifiers," in *Proc. IEEE Int. Conf. Comput. Sci. Comput. Intell.*, Las Vegas, USA, 2016, pp. 1192–1195.

[42] S. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification," *IEEE Trans. Neural Networks*, vol. 3, no. 5, pp. 683–697, Sept. 1992.

[43] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, Las Vegas, NV, USA, 2016, pp. 770–778.

[44] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Arcas, "Communication-efficient learning of deep networks from decentralized data," *Proc. Int. Conf. Artif. Intell. Stat.*, vol. 54, pp. 1273–1282, Apr. 2017.

[45] C. Hilbe, M. Nowak, and K. Sigmund, "Evolution of extortion in iterated prisoner's dilemma games," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 17, pp. 6913–6918, Aug. 2013.

[46] M. Nowak and K. Sigmund, "Tit for tat in heterogeneous populations," *Nature*, vol. 355, pp. 250–253, Jan. 1992.

Changbing Tang (Senior Member, IEEE) received the Ph.D. degree from the Department of Electronic Engineering, Fudan University, in 2014. He received the B.S. and M.S. degrees in mathematics and applied mathematics from Zhejiang Normal University, in 2004 and 2007, respectively.

Dr. Tang was the recipient of the Academic New Artist Doctoral Post Graduate from the Ministry of Education of China in 2012 and the recipient of the Academician Pairing Training Program for Young Talents of Zhejiang Province in 2019. He is currently an Associate Professor with the College of Physics and Electronic Information Engineering, Zhe-

jiang Normal University. His research interests include game theory and its applications, networks and distributed optimization.

Baosen Yang received the B.S. degree in mathematics and applied mathematics from Qilu Normal University in 2021. He is currently pursuing for the M.S. degree in the School of Mathematical Sciences of Zhejiang Normal University.

His research interests include game theory and its applications, zero-determinant strategy, federated learning and networks.

Xiaodong Xie received the B.S. degree in electronic science and technology, and the B.S. degree in communication engineering from Nanjing University of Posts and Telecommunications in 2021. He is currently pursuing for the M.S. degree in the College of Physics and Electronic Information Engineering, Zhejiang Normal University.

His research interests include smart grid, game theory and blockchain.

Guanrong Chen (Life Fellow, IEEE) received the M.Sc. degree in computer science from Sun Yat-sen University, in 1981, and the Ph.D. degree in applied mathematics from Texas A&M University, USA, in 1987. He was a tenured Full Professor at the University of Houston, USA, before he joined City University of Hong Kong, Hong Kong, China, as a Chair Professor in 2000, where he has been the Founding Director of the Centre for Complexity and Complex Networks, and is now the Hong Kong Shun Hing Education and Charity Fund Chair Professor in engineering.

Prof. Chen was awarded the 2011 Euler Gold Medal from Russia, and conferred Honorary Doctor Degrees by the Saint Petersburg State University, Russia in 2011 and by the University of Normandy, France in 2014. He has been a member of the Academia Europaea since 2014 and a Fellow of The World Academy of Sciences since 2015.

Prof. Chen's research interests are in the fields of complex networks, nonlinear dynamics and control systems. He has been a Highly Cited Researcher in Engineering continuously for some ten years according to Clarivate Web of Science.

Mohammed A. A. Al-qaness received the B.S., M.S., and Ph.D. degrees from Wuhan University of Technology, in 2010, 2014, and 2017, respectively, all in information and communication engineering.

He was an Assistant Professor with the School of Computer Science, Wuhan University. He was also a Postdoc Researcher at the State Key Laboratory for Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. Now, he is a Professor at College of Physics and Electronic Information Engineering, Zhejiang Normal University. His current research interests include wireless sensing, mobile computing, machine learning, signal and image processing, and natural language processing.
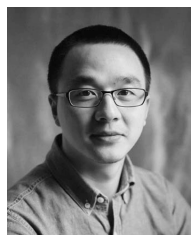
Yang Liu (Senior Member, IEEE) received the B.S. degree in mathematics from Zhejiang Normal University, in 2003, and the Ph.D. degree from Tongji University, in 2008.

He is currently a Vice Dean of Jinhua Intelligent Manufacturing Research Institute. He is also a Professor with the School of Mathematical Sciences and with Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University. He has authored over 100 publications and three books. He is an Associate Editor of *Neural Processing Letters* (Springer). He was recognized by Elsevier as a Most Cited Chinese Researcher in 2020–2022, and by Clarivate Analytics as a Highly Cited Researcher in 2019–2022. He was a Supervisor of ICCM Best Thesis Award in 2016 and 2022. His research interests include logical systems, hybrid systems and distributed optimization.