ICL: An Incentivized Collaborative Learning Framework

Xinran Wang[1], Qi Le[1], Ahmad Faraz Khan[2], Jie Ding[3], and Ali Anwar[1]
[1]Computer Science and Engineering, University of Minnesota-Twin Cities, Minneapolis, MN 55455, USA
[2]Computer Science, Virginia Tech, Blacksburg, VA 24060, USA
[3]School of Statistics, University of Minnesota-Twin Cities, Minneapolis, MN 55455, USA
Email: {wang8740, le000288, dingj, aanwar}@umn.edu, ahmadfk@vt.edu

2024 IEEE International Conference on Big Data (BigData) | 979-8-3503-6248-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/BigData62323.2024.10825643

***Abstract*—Collaborations among various entities, such as companies, research labs, AI agents, and edge devices, have become increasingly crucial for achieving machine learning tasks that cannot be accomplished by a single entity alone. This is likely due to factors such as security constraints, privacy concerns, and limitations in computation resources. As a result, Collaborative Learning has been gaining momentum. However, a significant challenge in practical applications of Collaborative Learning is how to effectively incentivize multiple entities to collaborate before any collaboration occurs. In this study, we propose ICL, an architectural framework for Incentivized Collaborative Learning, and provide insights into the critical issue of when and why incentives can improve collaboration performance. We showcase the concepts of ICL to specific use cases in federated learning, assisted learning, and multi-armed bandit, corroborating with both theoretical and experimental results.**

I. Introduction

Motivation. Over the past decade, Artificial Intelligence (AI) has achieved significant success in engineering and scientific domains, e.g., robotic control [1], natural language processing [2], and computer vision [3]. With this trend, a growing number of entities, e.g., governments, hospitals, companies, and edge devices, are integrating AI models into their workflows to facilitate data analysis and enhance decision-making. While a variety of standardized Machine Learning (ML) models are readily available for entities to implement AI tasks, model performance heavily depends on the quality and availability of local training data, models, and computation resources [4]. For example, a local bank's financial model may be constrained by the small size of its subjects and the number of feature variables. However, this bank could possibly improve its model by integrating additional observations and feature variables from other banks or industry sectors. Therefore, there is a strong need for collaborative learning that allows entities to enhance their model performance while respecting the proprietary nature of local resources. This has motivated recent research on learning frameworks, such as Federated Learning (FL) [5, 6] and Assisted Learning (AL) [7], which can improve learning performance from distributed data. ML entities, similar to humans, can collaborate to accomplish tasks that benefit each participant. However, these entities possess local ML resources that can be highly heterogeneous in terms of training procedures, computation cost, sample size, and data quality. A key challenge in facilitating such collaborations is understanding the motivations and incentives that drive entities to participate in the first place. An effective incentive mechanism is crucial for facilitating a "benign" collaboration in which high-quality entities are suitably motivated to maximize the overall benefit [8].

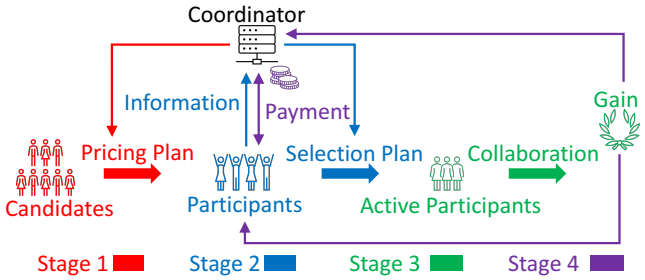

Fig. 1: Overview of ICL

Limitation of state-of-the-art approaches. The need to deploy collaborative learning systems in the real world has motivated the development of incentive mechanisms in FL. However, these incentive mechanisms are designed only to fulfill specific needs for a narrow set of applications as each application has unique requirements and processes [9, 10]. This requires spending development time and cost as well as research hours on creating an incentive mechanism for every new application [11–14]. This is particularly due to the lack of any generic incentive framework that can be used for a wider spectrum of applications. Furthermore, the existing literature has studied different aspects of incentives in particular application cases, e.g., using contract theory to set the price for participating in FL, or evaluating contributions for reward or penalty allocation, which we will review in Subsection II. However, understanding when and why incentive mechanism design can enhance collaboration performance is under-explored.

Key insights. In this work, we aim to address these challenges by developing a generic incentivized framework for Collaborative Learning (ICL) to abstract common application scenarios. For this task, we first observe the common roles of all actors in a broad spectrum of collaborative learning applications. As illustrated in Fig. 1, a set of learners play

three roles as the game proceeds: 1) *candidates*, who decide whether to participate in the game based on a pricing plan, 2) *participants*, whose final payment, which can be negative if interpreted as a reward, depends on the pricing plan and actual outcomes of collaboration, and 3) *active participants*, who jointly realize a collaboration gain to be enjoyed by all participants. The system-level goal is to promote high-quality collaborations for an objective. Examples of the collaboration gain include improved models, predictability, and rewards.

Solution. Driven by the above observations, we propose the following design principles of incentives to benefit collaboration: (1) Each participant can simultaneously play the roles of contributor and beneficiary of the collaboration gain. (2) Each participant will pay to participate in return for a collaboration gain if the improvement over its local gain outweighs its participation cost. (3) The pricing plan determines each entity's participation cost and can be positive or non-positive, tailored to reward those who contribute positively and charge those who hinder collaboration or disproportionately benefit from it. (4) The system for collaboration may incur a net zero cost, while still engaging entities to achieve the maximum possible collaborative gains. Our framework provides a unified understanding for the modular design of incentives from a system design perspective. We will show how incentives can be used to reduce exploration complexity and create win-win situations for participants from collaboration.

Contributions. We make three main contributions through this work. First, we propose a generic framework for ICL along with design principles. These collectively formalize the role of incentives in learning, ensuring that eligible entities are motivated to actively foster collaboration and benefit all the participants. Second, we showcase the adaptability of ICL by integrating it into diverse collaborative learning scenarios, including Federated Learning (FL), Assisted Learning (AL), and Multi-Armed Bandits (MAB). Our approach emphasizes the framework's versatility and modularity, illustrating its potential to enhance a broad spectrum of applications, which in turn could result in significant savings in time, cost, and resources. Lastly, through a series of experimental studies, we validate our theoretical constructs and provide practical insights. Specifically, our results highlight the pivotal role of well-designed pricing and selection strategies in minimizing exploration costs in learning environments, ultimately fostering mutually beneficial outcomes for all participants.

II. Related work

To address collaborative learning challenges, existing studies have focused on various aspects, such as security [15], privacy [16], fairness [17], personalization [18], model heterogeneity [19], and lack-of-labels [20]. However, a fundamental question remains: why would participants want to join collaborative learning in the first place? This has motivated an active line of research to use incentive schemes to enhance collaboration. We briefly review them below.

Promoter of incentives. Who want to design mechanisms to incentivize participants and initiate a collaboration? From this angle, existing work can be roughly categorized in two classes: server-centric, meaning that a collaboration is initiated by a server who owns the model and aims to incentivize edge devices to join model training [13, 21], and participant-centric where the incentives are designed at the participants' interest [22].

Different goals of incentives. What is the objective of an incentive mechanism design? Most existing work on incentivized collaborative learning, in particular FL, have adopted some common rules for incentive mechanism design, e.g. incentive compatibility and individual rationality [13]. The eventual objective for incentivized collaboration is often maximizing profit from the perspective of the incentive mechanism designer, which is either the coordinator (also called "server", "platform") [13] or the participants (also called "clients" in FL) [22]. Another commonly studied objective is maximizing global model performance in FL, which can be commercialized and turned into profit [23]. Other objectives being studied include budget balance [14], computational efficiency [12], fairness [24], and Pareto efficiency [23].

Overall, the role of incentives in collaborative learning has inspired many recent studies on bringing economic concepts to design learning platforms. Most existing work has focused on FL, especially mobile edge computing scenarios. Nonetheless, the need for collaboration extends beyond FL, as shown in [24] which studied synthetic data generation based on collaborative data sharing, and [8] which developed an AL framework where an entity being assisted is bound to assist others based on implicit mechanism design. Moreover, incentives in collaborative learning is under-studied in two critical aspects. Firstly, how to design incentives under a unified architecture, considering the existing work often focuses on specific application scenarios? Secondly, when and why do incentives improve collaboration performance? Prior work has often focused on designing an incentive as a separate problem based on an existing collaboration scheme, instead of treating incentive as part of the learning itself. These gaps motivated this work on ICL.

III. ICL Design

A. System Overview

In this section, we provide an overview of the ICL formulation. As illustrated in Fig. 1, a collaboration consists of four stages. In Stage 1, the coordinator sets a pricing plan based on prior knowledge of the candidates' potential gains (e.g., from previous rounds), and each candidate decides whether to be a participant by committing a payment at the end of this round. In Stage 2, the coordinator collects participants' information (e.g., their estimated gains) and uses a selection plan to choose the active participants. In Stage 3, the active participants collaborate to produce an outcome, which is enjoyed by all participants (including non-active ones). In Stage 4, the coordinator charges according to the pricing plan, the *realized* collaboration gain, and individual gains of active participants. Here, a gain (e.g., decrease in test loss) is assumed to be a function of the realized outcome (e.g., trained model).

B. ICL components

The ICL system includes two parties: candidate entities and coordinator. For notational convenience, we will first introduce a single-round game and extend it in Section IV.

Candidates. Consider M candidates indexed by $[M] \triangleq \{1, \ldots, M\}$. In an ICL game, each candidate m can *potentially* produce an outcome $x_m \in \mathcal{X}$, such as a model parameter. Any element in $\mathcal{X}$ can be mapped to a gain $Z \in \mathbb{R}$, e.g., reduced prediction loss. But such a gain will not necessarily be realized unless the candidate becomes an active collaborator of the game. At the beginning of a game, a candidate will receive a pricing plan from the coordinator specifying the cost of participating in the game and use that to decide whether to become a *participant* of the game. If a candidate participates, it has the opportunity to be selected as an *active participant*. All active participants will then collaborate to produce an outcome (e.g., model or prediction protocol), which also generates a collaboration gain. This outcome is distributed among all participants to benefit them. At the end of the game, all participants must pay according to the pre-specified pricing plan, with the actual price depending on the realized collaboration gain.

We let $\mathbb{I}_{\text{P}}$ and $\mathbb{I}_{\text{A}}$ denote the set of participants and active participants, respectively (so $\mathbb{I}_{\text{A}} \subseteq \mathbb{I}_{\text{P}} \subseteq [M]$). Given the above, an entity has a *consumer-provider bi-profile*, meaning that it can serve as a consumer who wishes to benefit from and also a provider who contributes to the collaboration.

Coordinator. A coordinator, e.g., company, government agency, or platform, orchestrates the game by performing the following actions in order: determine a pricing plan of the participation costs based on initial information collected from candidates, select active participants from those candidates that have chosen to become participants, realize the collaboration gain, and charge the participants according to the gain. The coordinator can be a virtual entity rather than a physical one.

Collaboration gain. Given active participants represented by $\mathbb{I}_{\text{A}}$, the collaborative gain is a function of their individual outcomes, denoted by $\mathcal{G}: (x_m, m \in \mathbb{I}_{\text{A}}) \mapsto z_{\mathbb{I}_{\text{A}}} \in \mathbb{R}$. This gain will be enjoyed by all participants and the coordinator, e.g., in the form of an improved model distributed by the coordinator. We also use $\mathcal{G}: x_m \mapsto z_m \in \mathbb{R}$, $m \in [M]$, to denote the gain of an individual outcome.

Pricing plan. The pricing plan is a function from $\mathbb{R}^{|\mathbb{I}_{\text{A}}|+1}$ to $\mathbb{R}^{|\mathbb{I}_{\text{P}}|}$ that maps the collaboration gain and individual gains of active participants to a cost needed to participate in the game:

$$\mathcal{P}: (z, z_m, m \in \mathbb{I}_{\text{A}}) \mapsto (c_j, j \in \mathbb{I}_{\text{P}}), \tag{1}$$

where z denotes the realized collaboration gain. In practice, we may parameterize $\mathcal{P}$ so that it is low for active and good-performing entities, medium for non-active entities, and high for active and laggard/disruptive entities, a point we will demonstrate in the experiments. We assume that the active participants will share their individual gains, namely $z_m, m \in \mathbb{I}_{\text{A}}$, so that all other participants' cost can be evaluated. The $\mathcal{P}$ will provide incentives to each candidate to decide to participate or not. As such, we denote the set of participants by $\mathbb{I}_{\text{P}} = \text{Incent}(\mathcal{P})$.

Profit. For each party, the profit will consist of two components: monetary profit from participation fees and gain-converted profit from collaboration gains. More specifically, let c_m denote the final participation cost for entity m. Let the Utility-Income (UI) function $z \mapsto \mathcal{U}(z)$ determine the amount of income uniquely associated with any particular gain Z. We suppose the UI function is the same for participants and the system. Then, the profit of client m is

$$\text{PROFIT}_m \triangleq \mathbb{1}_{m \in \mathbb{I}_{\text{P}}} \cdot (-c_m + \mathcal{U}(z_{\mathbb{I}_{\text{A}}}) - \mathcal{U}(z_m)), \tag{2}$$

where $z_{\mathbb{I}_{\text{A}}} \triangleq \mathcal{G}(x_m, m \in \mathbb{I}_{\text{A}})$, and the last term contrasts with its standalone learning. We define the system-level profit as the overall income from participation, $\sum_{m \in \mathbb{I}_{\text{P}}} c_m$, weighted plus the amount converted from collaboration gain, $\mathcal{U}(z_{\mathbb{I}_{\text{A}}})$, namely

$$\text{PROFIT}_{\text{sys}} \triangleq \lambda \sum_{m \in \mathbb{I}_{\text{P}}} c_m + \mathcal{U}(z_{\mathbb{I}_{\text{A}}}), \tag{3}$$

where $\lambda \geq 0$ is a pre-specified control variable that balances the monetary income and collaboration gain. We can regard the system-level profit as the coordinator's profit.

Coordinator's profit. One may put additional constraints on the coordinator's monetary income $\sum_{m \in \mathbb{I}_{\text{P}}} c_m$. A particular case is to restrict that $\sum_{m \in \mathbb{I}_{\text{P}}} c_m = 0$, which may be interpreted that the system does not need actual monetary income but rather uses the mechanism for model improvement. This is typical in coordinator-free decentralized learning (to revisit in Section IV-B).

Selection plan. The coordinator will select the active participants $\mathbb{I}_{\text{A}}$ from $\mathbb{I}_{\text{P}}$ based on a set of available information, denoted by $\mathcal{I}$. We assume that the $\mathcal{I}$ consists of the coordinator's belief of the distributions of x_m (namely the realizable gain) for each client m in $\mathbb{I}_{\text{P}}$. The selection plan is a function that maps from $\mathcal{I}$ and $\mathbb{I}_{\text{P}}$ to a set $\mathbb{I}_{\text{A}} \subseteq \mathbb{I}_{\text{P}}$, denoted by

$$\mathcal{S}: (\mathcal{I}, \mathbb{I}_{\text{P}}) \mapsto \mathbb{I}_{\text{A}}. \tag{4}$$

This can be a randomized map, e.g., when each participant is selected with a certain probability (Section III-D2). In practice, $\mathcal{I}$ may refer to the coordinator's estimates of the underlying distribution of x_m, $m \in \mathbb{I}_{\text{P}}$, based on historical performance on the participant side.

Objective of mechanism design. Our objective in designing a collaboration mechanism is to maximize the system-level profit under constraints tied to candidates' individual incentives, which will be revisited in Section III-D. The maximization is over the pricing plan $\mathcal{P}$ and selection plan $\mathcal{S}$. With the earlier discussions, the objective is

$$\max_{\mathcal{P}, \mathcal{S}} \mathbb{E}\left\{\lambda \sum_{m \in \mathbb{I}_{\text{P}}} c_m + \mathcal{U}(z_{\mathbb{I}_{\text{A}}})\right\}, \text{ where} \tag{5}$$

$$c_m \text{ is specified by } \mathcal{P} \text{ in (1)}, \; z_{\mathbb{I}_{\text{A}}} = \mathcal{G}(x_m, m \in \mathbb{I}_{\text{A}}), \tag{6}$$

$$\mathbb{I}_{\text{A}} = \mathcal{S}(\mathcal{I}, \mathbb{I}_{\text{P}}), \text{ s.t. } \mathbb{I}_{\text{P}} = \text{Incent}(\mathcal{P}). \tag{7}$$

We will elaborate on (7) in Section III-C.

Interpretation of the objective. We discuss three interesting cases of the objective. First, when $\lambda = 1$, the objective is equivalent to maximizing the overall profit. Second, when $\lambda = 0$, the objective is to improve the modeling through a collaboration mechanism. In this case, the system has no interest in the participation income but only provides a platform to incentivize the non-active to pay for the gain obtained by the active. Since the system need not pay for any participants, it is natural to assume the "zero-balance" constraint $\sum_{m\in\mathbb{I}_\text{P}} c_m = 0$. Thus, we have

$$\text{Objective: } \max_{\mathcal{P},\mathcal{S}} \mathbb{E}\{\mathcal{U}(z_{\mathbb{I}_\text{A}})\}. \tag{8}$$

Third, as $\lambda \to \infty$, the objective reduces to maximizing the system profit, $\max_{\mathcal{P},\mathcal{S}} \mathbb{E}\{\sum_{m\in\mathbb{I}_\text{P}} c_m\}$. Intuitively, the collaboration gain should still be reasonable to attract sufficiently many participants. Lastly, the following proposition shows that by properly replacing the λ, the system's objective can be interpreted as an alternative objective that combines the system's and participants' gains.

Proposition 1. Let $\lambda' \triangleq \frac{\lambda-1}{|\mathbb{I}_\text{P}|+1}$. The Objective (5) where λ' is replaced with λ is equivalent to maximizing the average social welfare defined by $(\text{PROFIT}_\text{sys} + \sum_{m\in\mathbb{I}_\text{P}} \text{PROFIT}_m)/(|\mathbb{I}_\text{P}|+1)$.

C. Incentives of participation

We study the incentives of collaboration from the candidates' perspectives. First, we will elaborate on (7) here. For each candidate, the incentive to become a participant is the larger profit of receiving the collaboration gain compared with realizing a gain on its own. Then, candidate m has the *incentive* to participate in the game if and only if

$$\text{Incent}_m: \ \mathbb{E}\big\{-c_m + \mathcal{U}(z_{\mathbb{I}_\text{A}}) - \mathcal{U}(z_m)\big\} \geq 0, \tag{9}$$

where $z_{\mathbb{I}_\text{A}}$ and c_m were introduced in (6) and (7). Here, $\mathbb{E}$ denotes the expectation regarding the random quantities, including the active participant set and the realized gains.

Inaccurate candidate. A candidate may have its own expectation $\mathbb{E}_m$ in place of $\mathbb{E}$ in (9) when making its decision. In this case, if the candidate is overly confident about the collaboration gain – its expected z tends to be larger than the actual, either intentionally or not – it will participate in the game. The system can have a further screening of it: 1) if this participant is selected as an active participant, it will likely suffer from a penalty since its realized gain will be seen by the coordinator, which will implicitly give feedback as an incentive to that candidate; 2) if not selected, it will become an inactive participant, which will contribute to the system's profit but not harm the collaboration. In this way, a candidate will have a limited extent to harm the system.

D. Mechanism design for the ICL game

The idea of mechanism design in economic theory is to devise mechanisms to jointly regulate the decisions of multiple parties in a game to eventually attain a system's desired goal (see, e.g., [25]). In our ICL game, the system's desired goal is to maximize (5), and the mechanisms to design include $\mathcal{P}$ and $\mathcal{S}$. Section III-C discussed the incentives from the candidates' view. This subsection studies the mechanism designs from the system's perspective.

1) Pricing plan, from candidates to participants: From the system's view, we can cast the M candidates and the coordinator as the parties in a game. Consider the following strategy choices of each party. Each candidate m has two choices: whether to participate or not, represented by $b_m \triangleq \mathbb{1}_{m\in\mathbb{I}_\text{P}} \in \{0,1\}$; the coordinator has a choice of the pricing and selection plans, denoted by $(\mathcal{P},\mathcal{S})$. Following the notation in (4), for a set of participants that exclude m, denoted by $\mathbb{I}_\text{P}^{(-m)}$, we let $\mathbb{I}_\text{A}^{(-m)} \triangleq \mathcal{S}(\mathcal{I}, \mathbb{I}_\text{P}^{(-m)})$ and $\mathbb{I}_\text{A} \triangleq \mathcal{S}(\mathcal{I}, \mathbb{I}_\text{P}^{(-m)} \cup \{m\})$. We have the following condition under a Nash equilibrium.

Theorem 1 (Equilibrium condition)**.** The condition to attain Nash equilibrium is

$$\text{Incent}_m: \mathbb{E}\big\{-c_m + \mathcal{U}(z_{\mathbb{I}_\text{A}}) - \mathcal{U}(z_m)\big\} \geq 0, \text{ iff. } m \in \mathbb{I}_\text{P}, \tag{10}$$

$$\text{Incent}_\text{sys}: \mathbb{E}\big\{\lambda c_m + \mathcal{U}(z_{\mathbb{I}_\text{A}}) - \mathcal{U}(z_{\mathbb{I}_\text{A}^{(-m)}})\big\} \geq 0, \text{ iff. } m \in \mathbb{I}_\text{P}. \tag{11}$$

Pricing as a part of the collaborative learning. A critical reader may wonder why not price participants directly based on the realized gains, which we refer to as post-collaboration pricing, e.g., using the Shapley value [18, 26]. The main distinction is that our studied pricing plan can not only generate profit or reallocate resources on the system side but also influence collaboration gains. Specifically, the pricing plan can screen higher-quality candidates to allow the coordinator to improve model performance in the subsequent collaboration. For instance, consider the case where the sole purpose is to maximize collaboration gain, namely $\lambda = 0$. In this situation, an entity m violating the condition in (11) is treated as a laggard, and c_m can be designed accordingly to ensure this client will not participate, as per violating (10).

2) Selection plan, from participants to active participants: We introduce a general probabilistic selection plan. Assume the information transmitted from participant m is a distribution of x_m, denoted by $\mathcal{P}_m$ for all $m \in \mathbb{I}_\text{P}$. Suppose the system expects to select $\rho \in (0,1]$ proportion of the participants. Consider a probabilistic selection plan that will select each client m in $\mathbb{I}_\text{P}$ with probability $q_m \in [0,1]$. Let $q = [q_m]_{m\in\mathbb{I}_\text{P}}$. We thus have the constraint

$$q \in Q(\rho, \mathbb{I}_\text{P}) \triangleq \bigg\{q : \sum_{m\in\mathbb{I}_\text{P}} q_m = \rho|\mathbb{I}_\text{P}|\bigg\}. \tag{12}$$

Let b_m denote an independent Bernoulli random variable with $\mathbb{P}(b_m = 1) = q_m$, or $b_m \sim \text{B}(q_m)$. Then, conditional on the existing participants $\mathbb{I}_\text{P}$, maximizing any system objective, e.g., (5) and (8), will lead to a particular law of client selection represented by q. For example, for the objective (8), we may solve the following problem. $q^* = \arg\max_{q\in Q(\rho,\mathbb{I}_\text{P})} U(q) \triangleq \mathbb{E}\big\{\mathcal{U}(z_{\mathbb{I}_\text{A}}) = \mathcal{U}(\mathcal{G}(x_m, m \in \mathbb{I}_\text{A}))\big\}$, where the expectation is over $b_m \sim \text{B}(q_m), x_m \in \mathcal{P}_m$, and $\mathbb{I}_\text{A} = \{m \in \mathbb{I}_\text{P} : b_m = 1\}$. We will show specific examples in Section IV. The existing works have examined client sampling from perspectives other than incentives, such as minimizing gradient variance [27].

Free-rider and adversarial participants. A free-rider is an entity m with a low local gain z_m but hopes to participate to enjoy the collaboration gain realized by other more capable participants with a relatively small participation cost. To that end, the entity may deliberately inform the system of a poor $\mathcal{P}_m$ so that the system, if following the above-optimized selection plan, will not select it as active. Consequently, the free-rider's actual local gain will not be revealed and may not suffer a high participation cost. This case motivates us to adopt a random selection to a certain extent in selecting the active participants. More specifically, suppose every participant $m \in \mathbb{I}_{\text{P}}$ will have at least a $\bar{\rho} > 0$ probability of being selected to be active. Then, it expects to pay at least $\bar{\rho} \cdot \mathbb{E}\{c_m\} = \bar{\rho} \cdot \mathcal{P}(z, z_i, i \in \mathbb{I}_{\text{A}})$ in return for an additional model gain of $\mathbb{E}\{\mathcal{U}(z_{\mathbb{I}_{\text{A}}}) - \mathcal{U}(z_m)\}$, where $\mathbb{I}_{\text{A}}$ contains m. Thus, it is not worth the entity m's participation should the system design a cost function that meets the following: $\bar{\rho} \cdot \mathbb{E}\{\mathcal{P}(z, z_m, m \in \mathbb{I}_{\text{A}})\} \geq \mathbb{E}\{\mathcal{U}(z_{\mathbb{I}_{\text{A}}}) - \mathcal{U}(z_m)\}$ for all z_m overly small. For example, the coordinator may impose a high cost whenever the realized local gain z_m revealed after the collaboration exceeds a pre-specified threshold. On the other hand, there may be an adversarial participant with a poor local gain but informs the system of an excellent $\mathcal{P}_m$ so that the system will select it to be active. In such cases, the same argument regarding the choice of the pricing plan applies, so no adversarial entity would dare to risk paying an excessively high cost after participation.

Random sampling for noninformative scenarios. We show that if the system is noninformative, random sampling can be close to the optimal selection. Suppose the information from participant m is the mean and variance of $x_m \in \mathbb{R}$, denoted by μ_m, σ^2 for $m \in \mathbb{I}_{\text{P}}$. A large σ^2 means less information. The result below shows random sampling is close to the optimal for large σ.

Proposition 2. Assume the gain is defined by $\mathcal{G}(x) \triangleq -\mathbb{E}(x-\mu)^2$, where μ represents the underlying parameter of interest, and the participants' weights ζ_m's are the same. Assume that $\sigma^2/(|\mathbb{I}_{\text{P}}| \cdot \max_{m \in \mathbb{I}_{\text{P}}}(\mu_m - \mu)^2) \to \infty$ as $|\mathbb{I}_{\text{P}}| \to \infty$.
Then, we have $U(q)/U(q^*) \to_p 1$ as $|\mathbb{I}_{\text{P}}| \to \infty$.

IV. Use cases of ICL

In this section, we present example use cases of ICL. Additional use cases, detailed theory, and experimental evaluations are available in the arXiv version [28].

A. ICL for Federated Learning

Federated learning (FL) [5, 6] is a popular distributed learning framework where the main idea is to learn a joint model using the averaging of locally learned model parameters. Its original goal is to exploit the resources of massive edge devices (also called "clients") to achieve a global objective orchestrated by a central coordinator ("server") in a way that the training data do not need to be transmitted. In line with FL, we suppose that at any particular round, the outcome of client m, x_m, represents a model. The collaboration will generate an outcome in an additive form: $z_{\mathbb{I}_{\text{A}}} \triangleq \mathcal{G}(x_m, m \in \mathbb{I}_{\text{A}}) = \mathcal{G}(\sum_{i \in \mathbb{I}_{\text{A}}} \zeta_i x_i / \sum_{i \in \mathbb{I}_{\text{A}}} \zeta_i)$, where ζ_i's are the pre-determined unnormalized positive weights associated with all the candidates, e.g., according to the sample size [6] or uncertainty quantification [29]. Let $\mathbb{I}_{\text{P}} \triangleq [K]$, where $K \leq M$ is the number of participants, and M is the number of candidates.

Algorithm 1 Incentivized Federated Learning

Input: Datasets $D_{1:M}$ distributed on M local clients, active rate $\rho \in (0, 1]$, number of rounds T, objective parameter λ, clients' unnormalized weights $\zeta_{1:M}$, test loss L, pricing plan $\mathcal{P}_\theta : (z, z_m, m \in \mathbb{I}_{\text{A}}) \mapsto (c_j, j \in \mathbb{I}_{\text{P}})$, where $\theta = [\theta_1, \theta_2]$ is the unknown parameter, $c_j \triangleq A_1(z; \theta_1) + \mathbb{1}_{j \in \mathbb{I}_{\text{A}}} A_2(z - z_m; \theta_2)$, and A_1, A_2 are pre-specified functions, server's test dataset D_{test}, sigmoid function $\sigma_s : v \mapsto (1 + e^{-v/s})^{-1}$ where $s > 0$ is a hyperparameter, learning rate $\eta > 0$ for optimizing θ. Recall that $\tau_{m,t} \leq t$ denotes the last round when the client m was active before round t.

Output: Server's updated model $\bar{x}_t$, overall realized profit $\lambda \sum_{m \in \mathbb{I}_{\text{P},t}} c^*_{m,t} + \mathcal{U}(\bar{x}_t)$, $t = 1, \ldots, T$

Initialization: θ_0, $\bar{x}_0$, $\bar{z}_0 \triangleq \mathcal{G}(\bar{x}_0)$

System executes:

for each round $t = 1, \ldots, T$ **do**

$\theta_t \triangleq (\theta_{1,t}, \theta_{2,t}) \leftarrow$ **ServerStrategy**$(\bar{z}_{\tau_{m,t}}, z_{m,\tau_{m,t}}, \bar{x}_{\tau_{m,t}}, x_{\tau_{m,t}}, \zeta_m, m \in [M], \theta_{t-1})$

Distribute $\mathcal{P}_t \equiv \mathcal{P}_{\theta_t}$ to all M clients

$b_{m,t} \leftarrow$ **ClientStrategy**$(\mathcal{P}_t, z_{m,\tau_{m,t}}, \bar{z}_{t-1}), m \in [M]$

$\mathbb{I}_{\text{P},t} \leftarrow \{m \in [M] : b_{m,t} = 1\}$

$\mathbb{I}_{\text{A},t} \leftarrow$ randomly sample $\max(\lfloor \rho \cdot |\mathbb{I}_{\text{P},t}| \rfloor, 1)$ active clients from $\mathbb{I}_{\text{P},t}$

for each client $m \in \mathbb{I}_{\text{A},t}$ in parallel **do**

Distribute the server's model parameter $\bar{x}_{t-1}$ to local client m

$x_{m,t} \leftarrow$ ClientUpdate$(D_m, \bar{x}_{t-1})$ // use any standard local update

Send $x_{m,t}$ to the server

end

Receive model parameters from active clients, and calculate $\bar{x}_t = (\sum_{m \in \mathbb{I}_{\text{A},t}} \zeta_m)^{-1} \sum_{m \in \mathbb{I}_{\text{A},t}} \zeta_m x_{m,t}$

Calculate the collaboration gain $\bar{z}_t = \mathcal{G}(x_{m,t}, m \in \mathbb{I}_{\text{A}})$ and broadcast to all candidate clients

Calculate the individual gain of each active participant $z_{m,t} = \mathcal{G}(x_{m,t})$ and return it to the associated client m

Participant m pays $c^*_{m,t} \triangleq A_1(\bar{z}_t; \theta_{1,t}) + \mathbb{1}_{m \in \mathbb{I}_{\text{A},t}} \cdot A_2(\bar{z}_t - z_{m,t}; \theta_{2,t})$, for all $m \in \mathbb{I}_{\text{P},t}$

end

ServerStrategy $(\bar{z}_{\tau_{m,t}}, z_{m,\tau_{m,t}}, \bar{x}_{\tau_{m,t}}, x_{\tau_{m,t}}, \zeta_m, m \in [M], \theta_{t-1})$**:**

Define objective function of θ to maximize:

$$O_t(\theta) = \sum_{m=1}^{M} \sigma_s(\delta_{m,t}) \cdot \Big\{ \lambda \cdot c_{m,t}(\theta) - \frac{\zeta_m}{\sum_{j \in \mathbb{I}_{\text{A},\tau_{m,t}}} \zeta_j} \cdot f'(\bar{x}_{\tau_{m,t}}) \cdot (\bar{x}_{\tau_{m,t}} - x_{m,\tau_{m,t}}) \Big\},$$

$$c_{m,t}(\theta) \triangleq A_1(\bar{z}_{\tau_{m,t}}; \theta_1) + \rho \cdot A_2(\bar{z}_{\tau_{m,t}} - z_{m,\tau_{m,t}}; \theta_2) \quad (13)$$

$$\delta_{m,t} \triangleq \mathcal{U}(\bar{z}_{t-1}) - \mathcal{U}(z_{m,\tau_{m,t}}) - c_{m,t}(\theta) \quad (14)$$

where f is defined by

$$f(x) \triangleq \mathcal{U}\big(-L(x, D_{\text{test}})\big). \quad (15)$$

Return $\theta_t \leftarrow \theta_{t-1} + \eta \cdot \nabla_\theta O_t(\theta_{t-1})$

ClientStrategy $(\mathcal{P}_t, z_{m,\tau_{m,t}}, \bar{z}_{t-1})$**:**

Return $b_{m,t} \triangleq \mathbb{1}_{\delta_{m,t} > 0}$, where $\delta_{m,t}$ is the same as in (14) but with $\theta = \theta_t$

1) Large-participation approximation: Suppose the selection plan is based on a random sampling of $\mathbb{I}_{\mathrm{P}}$ with a given probability, say $\rho \in (0,1)$, for each participant to be active. Assume b_m, for $m \in [K]$, are IID Bernoulli random variables with $\mathbb{P}(b_m = 1) = \rho$. Let $\mathcal{U} \circ \mathcal{G}$ denote the composition of $\mathcal{U}$ and $\mathcal{G}$, and $(\mathcal{U} \circ \mathcal{G})'$ its derivative. Then, we have the following result to approximate the equilibrium conditions in Section III-D under a large number of participating clients. Let $\bar{\zeta}_{\mathbb{I}_{\mathrm{P}}}$ and $\bar{x}_{\mathbb{I}_{\mathrm{P}}}$ denote all the participants' average of ζ and weighted average of X, namely $\bar{\zeta}_{\mathbb{I}_{\mathrm{P}}} \triangleq (\sum_{i\in\mathbb{I}_{\mathrm{P}}} \zeta_i)/|\mathbb{I}_{\mathrm{P}}|$, $\bar{x}_{\mathbb{I}_{\mathrm{P}}} \triangleq (\sum_{i\in\mathbb{I}_{\mathrm{P}}} \zeta_i x_i)/(\sum_{i\in\mathbb{I}_{\mathrm{P}}} \zeta_i)$.

2) Iterative mechanism design: To optimize the mechanism in practice, the server cannot evaluate the collaboration gain $\mathcal{G}(x_m, m \in \mathbb{I}_{\mathrm{A}})$ due to its complex dependency on the individual outcomes x_m. Alternatively, the server can optimize its mechanism by maximizing the surrogate of the collaboration gain. The determination of the pricing plan will involve multiple candidates' choices. Since a practical FL system involves multiple rounds, we suppose each client will use the results from the previous rounds to approximate the current strategy set and make decisions.

In Algorithm 1, we give a more specific incentivized FL setup and operational algorithm based on the development in Section IV-A. The general Objective (5) could be regarded as a collaboration game in any particular round of FL. As FL consists of multiple rounds, we use the previous rounds as a basis for each candidate client to decide whether to participate in the current round. We use a subscript t to highlight the dependence of a quantity on the FL round. For example, the outcome x_m will be replaced with $x_{m,t}$ for round t. Derivations and further discussions of Algorithm 1 are included in [28].

B. ICL for Assisted Learning

Assisted learning (AL) [7, 8, 30, 31] is a decentralized learning framework that allows organizations to autonomously improve their learning quality within only a few assistance rounds and without sharing local models, objectives, or data. AL has primarily focused on vertically partitioned data, where entities possess data with distinct feature variables collected from the same cohort of subjects. Recently, AL has been extended to support organizations with horizontally partitioned data based on the idea of transmitting model training trajectories, within applications to reinforcement learning [32], in which the assisting agent has diverse environments, and unsupervised domain adaptation [33], where the assisting agent possesses supplementary data domains.

Unlike FL schemes, AL is 1) decentralized in that there is no global model to be shared or synchronized among all the entities in the training and 2) assistance-persistent in the sense that an entity still needs the output from other entities in the prediction stage. From the perspective of incentivized collaboration, the above naturally leads to two considerations with complementary insights into the pricing and selection plans compared with FL in Section IV-A.

• *Consideration 1: Autonomous incentive design without a coordinator.* Since each entity can initiate and terminate assistance, it is natural to consider a coordinator-free scenario, where entities can autonomously reach a consensus on collaboration partners based on their pricing plans.

• *Consideration 2: Limited information for incentive design.* In AL, an entity aims to seek assistance to enhance prediction performance without sharing proprietary local models. Thus, we suppose the communicated information for collaboration is limited to gains (z) rather than outcomes (x).

To put it into perspective, we will study a three-entity setting in Section IV-B1 to develop insights into the incentive that allows for a consensus on collaboration in Stage 1 (Fig. 1). In Section IV-B2, we will further study a multi-entity setting where multiple less-competitive participants are allowed to enter Stage 2 to enjoy the collaboration gain, but they will not compete for being active participants.

1) Consensus of competing candidates: In this section, we study three candidate entities, Alice, Bob, and Carol, and suppose each candidate aims to maximize its profit. Suppose a collaboration round can only consist of two entities. Then, the collaboration will only occur when two out of the three, say Alice and Bob, can maximally assist each other. From Alice's perspective, Carol is less competitive than Bob, and meanwhile, from Bob's perspective, Carol is less competitive than Alice. We will provide conditions to reach a consensus. Suppose each entity has its own payment plan: entity i will pay a price $p_i(z - z_i)$ for any given collaboration gain z and its local gain z_i, for all $i \in [M]$. So, if entities i and j collaborate, the actual price i will pay is $p_i(z-z_i)-p_j(z-z_j)$. The goal of each entity is to maximize the expected gain-converted profit minus the participation cost, namely the quantity in (2). For simplicity, suppose $p_i(\Delta z) = c_i \cdot \Delta z$ and $\mathcal{U}(z) = u \cdot z$. Let $\mu_{i,j} \triangleq \mathbb{E}(\mathcal{U}(z_{i,j}))$ denote the expected income of the collaboration gain formed by entities i and j, and $\mu_{j\leftarrow i} \triangleq \mathbb{E}(\mathcal{U}(z_{i,j}) - \mathcal{U}(Z_j))$ the additional gain brought by i to j.

2) Consensus of non-competing candidates: We will further study a multi-entity setting where multiple less-competitive participants are allowed to enjoy the collaboration gain, but they will not compete for being active participants. The objective is to develop an incentive to maximize the collaboration gain, namely Objective (8), that will eventually benefit all the participants. For ease of presentation, we will consider only one active participant (namely $|\mathbb{I}_{\mathrm{A}}| = 1$). In general, we may regard a set of participants as one "mega" participant. Following the above two considerations of AL, we will first study the following setup. Suppose K entities decide to participate in a collaboration, where one of them will be selected to realize the collaboration gain. For example, if participant m is active, it will realize a model gain $z_m \sim \mathcal{G}(x_m)$, where $x_m \sim \mathcal{P}_m$ is the potential outcome of participant m. Let $\mathcal{P}^*_m$ denote the distribution of z_m induced by $\mathcal{P}_m$ for $m \in [K]$ and suppose they are the *shared* information among participants. In line with Consideration 1, the system is set to have a zero balance,

namely $\sum_{m\in[K]} c_m = 0$. Following the notation in (1), we consider the following particular pricing plan:

$$\mathcal{P} : z \mapsto C(z)\mathbb{1}_{j\notin\mathbb{I}_{\text{A}}} - (K-1)C(z)\mathbb{1}_{j\in\mathbb{I}_{\text{A}}},\ j \in \mathbb{I}_{\text{P}}, \qquad (16)$$

where C is non-decreasing so that a larger gain is associated with a larger cost. In other words, each of the $K-1$ non-active participants will pay a cost of $C(z)$, which depends on the realized z, to the active participant. Then, the condition to reach a consensus among participants is the existence of a participant, say participant 1, such that when it is active, (i) the collaboration gain is maximized, and (ii) every participant sees that its individual profit is maximized, as formalized below.

Theorem 2. Assume $\mathcal{U}$ is a pre-specified non-decreasing function. Consider the pricing plan in (16) where C can be any non-negative and non-decreasing function. Let $\mu_m \triangleq \mathbb{E}_{\mathcal{P}^*_m}\{\mathcal{U}(z_m)\}$ denote the expected gain of participant m when it is active, $m \in [K]$. The necessary and sufficient condition for reaching a pricing consensus is the existence of a participant, say participant 1, that satisfies $\mu_1 - u_j \geq \mathbb{E}_{\mathcal{P}^*_1}\{C(z_1)\} + \mathbb{E}_{\mathcal{P}^*_j}\{(K-1)C(z_j)\}$ for all $j \neq 1$. If we further assume the linearity $\mathcal{U}(z) \triangleq u\cdot z$ and $C(z) = c\cdot z$, this inequality becomes $c \leq \min_{j\neq 1} u\cdot(\mu_1-\mu_j)/(\mu_1+(K-1)\mu_j)$.

V. Experimental Studies

A. Experimental Setup for Federated Learning

In this study, we used the following experimental setup.

Data. We evaluate the FashionMNIST [34], CIFAR10 [35], and CINIC10 [36] datasets. In Table I, we present the key statistics of each dataset.

TABLE I: Statistics of datasets in our experiments

Datasets	FashionMNIST	CIFAR10	CINIC10
Training instances	60,000	50,000	20,000
Test instances	10,000	10,000	10,000
Features	784	1,024	1,024
Classes	10	10	10

Setting. Standard FL can be vulnerable to poor-quality or adversarial clients. It is envisioned that a reasonable incentive mechanism can enhance the robustness of FL training against participation of these unwanted clients. To demonstrate this intuition, we consider an extreme case where there exist Byzantine attacks carried out by malicious or faulty clients [37]. We then demonstrate the robustness of the incentivized FL (labeled as "ICL" in the results) against two types of Byzantine attacks [38, 39]: random modification, which is a training data-based attack [37, 40], and label flipping, which is a parameter-based attack [41, 42]. Specifically, the random modification is based on generating local model parameters from a uniform distribution between -0.25 and 0.25, and the label flipping uses cyclic alterations, such as changing 'dog' to 'cat', and vice versa. Byzantine client ratios are adjusted to two levels: $\{0.2, 0.3\}$. In the reported results below, *Byzantine-0.2* refers to a scenario where 20% of the total clients are adversarial. We generate 100 clients using IID partitioning of the training part of each dataset. Among participating clients, the server will select 10% of clients as active clients per round. We adopt the model architecture and hyperparameter settings similar to [43].

Pricing. We consider the pricing plan with

$$\begin{aligned} c_j &= A_1(z;\theta_1) + \mathbb{1}_{j\in\mathbb{I}_{\text{A}}} A_2(z - z_m;\theta_2) \\ &= \theta_1\cdot z + \mathbb{1}_{j\in\mathbb{I}_{\text{A}}}\theta_1\cdot z\cdot(-1+\gamma\sigma_s(z-z_m-\theta_2)). \end{aligned}$$

Accordingly, we replace the calculation of cost in (13) with

$$c_{m,t}(\theta) \triangleq \theta_1\cdot\bar{z}_{\tau_{m,t}}\big(1+\rho\cdot(-1+\gamma\sigma_s(\bar{z}_{\tau_{m,t}} - z_{m,\tau_{m,t}} - \theta_2))\big).$$

In our study, we conduct an ablation test with different hyperparameters γ: 11, 101, and 2001. These are referred to as ICL plans 1, 2, and 3 in the following results. A larger value of γ implies a more substantial penalty for the under-performing clients. The θ_2 is initialized in each global communication round using the Jenks natural breaks technique [44] and optimized based on the objective function.

Profit. Recall that the profit of a client or the server consists of monetary profit from participation fees and gain-converted profit from collaboration gains. In FL, we define the collaboration gain Z_t as the negative test loss of the updated server model at round t, so the larger, the better. More specifically,

$$\begin{aligned} \bar{z}_t &= \mathcal{G}(x_{m,t}, m\in\mathbb{I}_{\text{A}}) = -L(\bar{x}_t, D_{\text{test}}) \\ &\triangleq -\sum_{(\text{input},\text{output})\in D_{\text{test}}} \ell(\text{input},\text{output};\bar{x}_t), \end{aligned}$$

where ℓ is the cross-entropy loss under the model parameterized by $\bar{x}_t$. Likewise, the local gain of a client m is

$$\begin{aligned} z_{m,t} &= \mathcal{G}(x_{m,t}) \triangleq -L(x_{m,t}, D_{\text{test}}) \\ &\triangleq -\sum_{(\text{input},\text{output})\in D_{\text{test}}} \ell(\text{input},\text{output};x_{m,t}). \end{aligned}$$

Notably, the test set only needs to be stored and operated by the server, and the clients only need to access their own historical gains and the server's gains. We use $\lambda = 0.1$ and $\mathcal{U} : z \mapsto z$ (the identity map) in the experiment.

Results. We visualize the learning curves on each dataset in Figures 2, 3, and 4. The model performance is assessed using the top-1 accuracy on the test dataset. We also summarize the best model performance in Tables II, III, and IV, respectively.

The following key points can be drawn from the results. Firstly, in comparison with non-incentivized Federated Learning (FL) based on FedAvg, our proposed incentivized FL (denoted as "ICL") algorithm can deliver a higher and more rapidly increasing model performance (representing Collaboration Gain in our context), as defined in (8). Across all settings–including two types of adversarial attacks, two ratios of adversarial clients, and three datasets–ICL with Pricing Plan 3 (ICL PP-3) consistently outperforms FedAvg by a significant margin. On the other hand, ICL with pricing plan 1 (ICL PP-1) underperforms, which is expected as it only imposes a mild penalty on laggard/adversarial active clients. The results

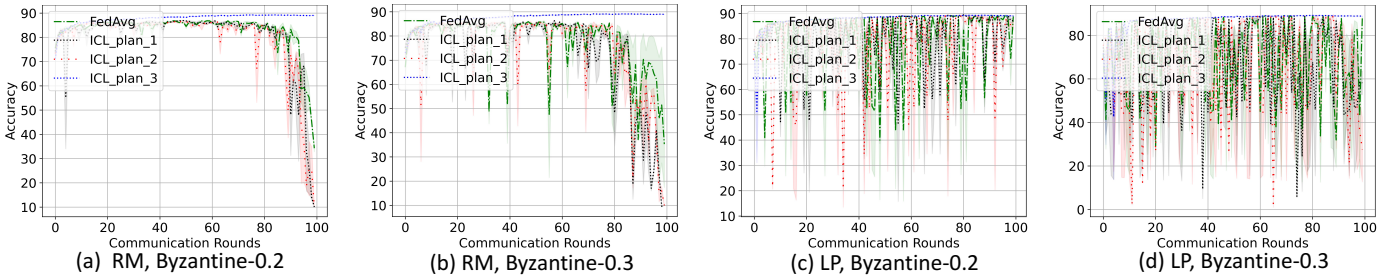

Fig. 2: Learning curves of ICL (incorporating three pricing plans) and FedAvg measured by Accuracy, assessed for random modification (RM) label flipping (LP) attacks and two malicious client ratios (0.2 and 0.3), applied to the **FashionMNIST**.

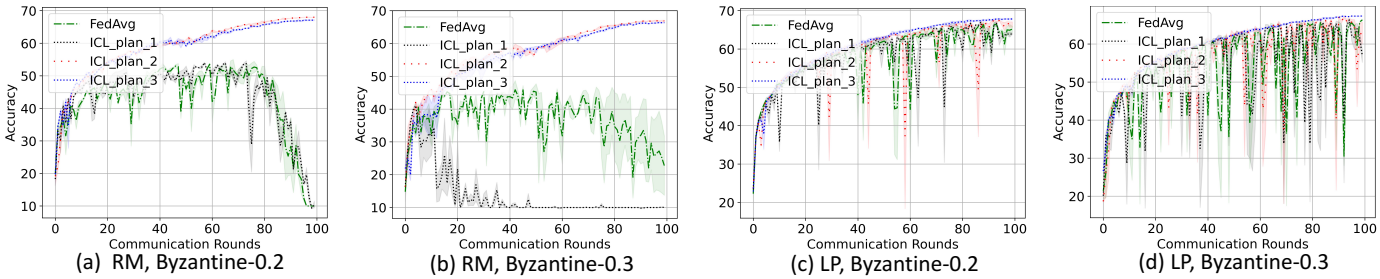

Fig. 3: Learning curves of ICL (incorporating three pricing plans) and FedAvg measured by Accuracy, assessed for random modification (RM) label flipping (LP) attacks and two malicious client ratios (0.2 and 0.3), applied to the **CIFAR10**.

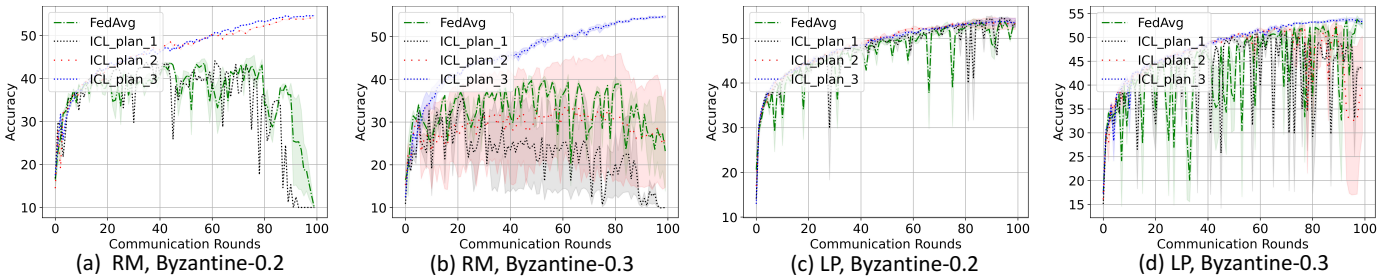

Fig. 4: Learning curves of ICL (incorporating three pricing plans) and FedAvg measured by Accuracy, assessed for random modification (RM) label flipping (LP) attacks and two malicious client ratios (0.2 and 0.3), applied to the **CINIC10**.

suggest that it is possible to significantly mitigate the influence of malicious clients by precluding them from participating in an FL round, given that the pricing penalty is sufficiently large. Furthermore, the figures also indicate that the random modification attack poses a more significant threat compared to the label flipping attack, making it particularly difficult for non-incentivized FedAvg to converge.

TABLE II: Best model prediction accuracy of ICL (incorporating three pricing plans) and FedAvg measured by Accuracy, assessed for random modification (RM) and label flipping (LP) attacks, with two malicious client ratios (0.2 and 0.3), applied to the **FashionMNIST** dataset. "PP-1" uses the smallest penalty for the underperforming clients.

Byzantine Method	Random Modification		Label Flipping	
	0.2	*0.3*	*0.2*	*0.3*
FedAvg	87.1 ± 0.1	86.5 ± 0.1	89.6 ± 0.0	89.3 ± 0.0
ICL PP-1	87.2 ± 0.0	86.4 ± 0.0	89.5 ± 0.1	89.0 ± 0.1
ICL PP-2	87.0 ± 0.0	86.4 ± 0.3	89.3 ± 0.1	89.4 ± 0.2
ICL PP-3	89.4 ± 0.0	89.2 ± 0.1	89.4 ± 0.1	89.2 ± 0.1

TABLE III: Best model prediction accuracy of ICL (incorporating three pricing plans) and FedAvg measured by Accuracy, assessed for random modification (RM) and label flipping (LP) attacks, with two malicious client ratios (0.2 and 0.3), applied to the **CIFAR10** dataset. "PP-1" uses the smallest penalty for the underperforming clients.

Byzantine Method	Random Modification		Label Flipping	
	0.2	*0.3*	*0.2*	*0.3*
FedAvg	55.7 ± 0.4	50.1 ± 0.3	67.3 ± 0.2	66.9 ± 0.2
ICL PP-1	54.8 ± 0.2	42.0 ± 0.9	67.0 ± 0.4	66.9 ± 0.3
ICL PP-2	68.1 ± 0.0	67.0 ± 0.2	67.6 ± 0.1	67.1 ± 0.1
ICL PP-3	67.1 ± 0.2	66.4 ± 0.1	67.9 ± 0.2	67.4 ± 0.1

B. ICL for Assisted Learning

We provide an experimental study to corroborate the insights in Section IV-B. The Inequality can be interpreted that the total gain of entity 1 received from entity 2, which consists of the collaboration-generated gain $(u - c_1)\mu_{1,2}$ and the pricing-based gain $c_2\mu_{2\leftarrow 1}$, is no larger than that from entity 3.

TABLE IV: Best model prediction accuracy of ICL (incorporating three pricing plans) and FedAvg measured by Accuracy, assessed for random modification (RM) and label flipping (LP) attacks, with two malicious client ratios (0.2 and 0.3), applied to the **CINIC10** dataset. "PP-1" uses the smallest penalty for the underperforming clients.

Byzantine Method	Random Modification		Label Flipping	
	0.2	*0.3*	*0.2*	*0.3*
FedAvg	45.9 ± 0.8	41.0 ± 0.6	54.3 ± 0.4	53.8 ± 0.8
ICL PP-1	44.1 ± 0.0	38.6 ± 2.1	54.4 ± 0.3	53.9 ± 0.1
ICL PP-2	54.1 ± 0.0	41.2 ± 7.8	55.4 ± 0.8	54.7 ± 0.6
ICL PP-3	54.6 ± 0.0	54.7 ± 0.3	55.1 ± 0.6	54.2 ± 0.3

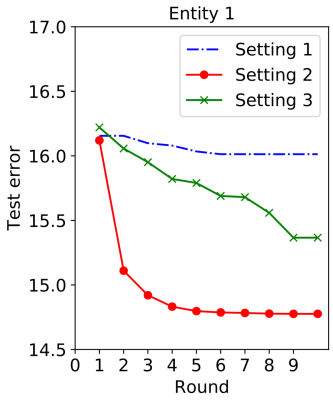

Fig. 5: Test error of Entity 1 in three settings. In Setting i, entity i pays none ($c_i = 0$) while entity $j \neq i$ pays with $c_j \neq 0$.

To show our pricing plan can promote mutually beneficial collaboration, we apply ICL to the parallel assisted learning (PAL) framework [8] to develop an incentivized PAL. More detailed algorithms and experimental studies are included in [28]. We show an experiment using real-world clinical data MIMIC [45]. Suppose three divisions (Entity 1, 2, 3) collect heterogeneous features from the same patients for different tasks: predict the heart rate, systolic blood pressure, and length of stay. In Setting i, Entity i does not provide an incentive while the other two entities do. The results in Fig. 5 show that in Setting 1, entity 1 does not gain much as it does not provide incentives; in other two settings, it gains from collaborating with the entity with mutual benefits.

C. Practical Guide on Implementing ICL

We synthesize empirical findings from three case studies to offer pragmatic insights for ICL deployment [28].

- ICL's versatility in diverse collaborative learning environments necessitates case-specific adaptations. For instance, applying the ICL equilibrium condition (Theorem 1) across the three use cases yields distinct formulas and implications.
- Synergizing the pricing and selection plans can effectively minimize learning exploration complexity, fostering mutually beneficial outcomes in collaboration. In the absence of a deliberate selection plan, such as the random selection in FL, the pricing plan becomes pivotal in filtering suitable collaborators.
- Collaboration consensus among participants can be achieved independently of a central coordinator or financial motivations within the system. Utilizing tokens, for instance, facilitates decentralized collaboration while maintaining zero balance at the system level.
- Collaboration gains differ based on whether gains are evaluated cumulatively over multiple rounds (as in multi-armed bandit scenarios in [28]) or at a single point in time (as in federated learning). Consequently, incentive plan optimization must consider these contextual variances in striking tradeoffs between exploration costs and gains.

VI. Conclusion

Collaboration among entities becomes increasingly important for enhancing their performance. We proposed an incentive framework to study how entities can be properly incentivized to create common benefits. There are several limitations worth further investigation. For instance, future studies could examine the functional forms of pricing plans, use cases to promote model security [46, 47] and privacy [48], and trade-offs between collaboration and competition.

VII. Acknowledgement

This paper is based upon work supported by the 3M Science and Technology Graduate Fellowship, Samsung Global Research Outreach award, National Science Foundation under CAREER Grant No. 2338506.

References

[1] M. P. Deisenroth, G. Neumann, and J. Peters, *A Survey on Policy Search for Robotics*, 2013.

[2] J. Li, W. Monroe, A. Ritter, D. Jurafsky, M. Galley, and J. Gao, "Deep reinforcement learning for dialogue generation," in *Proc. Conference on Empirical Methods in Natural Language Processing*, 2016, pp. 1192–1202.

[3] F. Liu, S. Li, L. Zhang, C. Zhou, R. Ye, Y. Wang, and J. Lu, "3dcnn-dqn-rnn: A deep reinforcement learning framework for semantic parsing of large-scale 3d point clouds," in *Proc. International Conference on Computer Vision (ICCV)*, 2017, pp. 5679–5688.

[4] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[5] J. Konecny, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," *arXiv preprint arXiv:1610.05492*, 2016.

[6] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. AISTATS*, 2017, pp. 1273–1282.

[7] X. Xian, X. Wang, J. Ding, and R. Ghanadan, "Assisted learning: A framework for multi-organization learning," in *Advances in Neural Information Processing Systems*, vol. 33, 2020.

[8] X. Wang, J. Zhang, M. Hong, Y. Yang, and J. Ding, "Parallel assisted learning," *IEEE Transactions on Signal Processing*, vol. 70, pp. 5848–5858, 2022.

[9] L. Lyu, X. Xu, Q. Wang, and H. Yu, "Collaborative fairness in federated learning," *Federated Learning: Privacy and Incentive*, pp. 189–204, 2020.

[10] R. H. L. Sim, Y. Zhang, M. C. Chan, and B. K. H. Low, "Collaborative machine learning with incentive-aware model rewards," in *International conference on machine learning*. PMLR, 2020, pp. 8927–8936.

[11] X. Tu, K. Zhu, N. C. Luong, D. Niyato, Y. Zhang, and J. Li, "Incentive mechanisms for federated learning: From economic and game theoretic perspective," *IEEE Transactions on Cognitive Communications and Networking*, 2022.

[12] J. Xu, J. Xiang, and D. Yang, "Incentive mechanisms for time window dependent tasks in mobile crowdsensing," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6353–6364, 2015.

[13] J. Kang, Z. Xiong, D. Niyato, H. Yu, Y.-C. Liang, and D. I. Kim, "Incentive design for efficient federated learning in mobile networks: A contract theory approach," in *2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*. IEEE, 2019, pp. 1–5.

[14] M. Tang and V. W. Wong, "An incentive mechanism for cross-silo federated learning: A public goods perspective," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.

[15] A. N. Bhagoji, S. Chakraborty, P. Mittal, and S. Calo, "Analyzing federated learning through an adversarial lens," in *International Conference on Machine Learning*. PMLR, 2019, pp. 634–643.

[16] S. Truex, N. Baracaldo, A. Anwar, T. Steinke, H. Ludwig, R. Zhang, and Y. Zhou, "A hybrid approach to privacy-preserving federated learning," in *Proceedings of the 12th ACM workshop on artificial intelligence and security*, 2019, pp. 1–11.

[17] H. Yu, Z. Liu, Y. Liu, T. Chen, M. Cong, X. Weng, D. Niyato, and Q. Yang, "A sustainable incentive scheme for federated learning," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 58–69, 2020.

[18] A. F. Khan, X. Wang, Q. Le, A. A. Khan, H. Ali, J. Ding, A. Butt, and A. Anwar, "Pi-fl: Personalized and incentivized federated learning," *arXiv preprint arXiv:2304.07514*, 2023.

[19] C. He, E. Mushtaq, J. Ding, and S. Avestimehr, "Fednas: Federated deep learning via neural architecture search," 2020.

[20] E. Diao, J. Ding, and V. Tarokh, "SemiFL: Communication efficient semi-supervised federated learning with unlabeled clients," in *Advances in Neural Information Processing Systems*, 2022.

[21] S. R. Pandey, N. H. Tran, M. Bennis, Y. K. Tun, Z. Han, and C. S. Hong, "Incentivize to build: A crowdsourcing framework for federated learning," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[22] G. Huang, X. Chen, T. Ouyang, Q. Ma, L. Chen, and J. Zhang, "Collaboration in participant-centric federated learning: A game-theoretical perspective," *IEEE Transactions on Mobile Computing*, 2022.

[23] R. Zeng, S. Zhang, J. Wang, and X. Chu, "Fmore: An incentive scheme of multi-dimensional auction for federated learning in MEC," in *2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2020, pp. 278–288.

[24] S. S. Tay, X. Xu, C. S. Foo, and B. K. H. Low, "Incentivizing collaboration in machine learning via synthetic data rewards," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 9, 2022, pp. 9448–9456.

[25] E. S. Maskin, "Mechanism design: How to implement social goals," *American Economic Review*, vol. 98, no. 3, pp. 567–76, 2008.

[26] A. E. Roth, *The Shapley value: essays in honor of Lloyd S. Shapley*. Cambridge University Press, 1988.

[27] W. Chen, S. Horvath, and P. Richtarik, "Optimal client sampling for federated learning," *arXiv preprint arXiv:2010.13723*, 2020.

[28] X. Wang, Q. Le, A. F. Khan, J. Ding, and A. Anwar, "A framework for incentivized collaborative learning," *arXiv preprint arXiv:2305.17052*, 2023.

[29] H. Chen, J. Ding, E. Tramel, S. Wu, A. K. Sahu, S. Avestimehr, and T. Zhang, "Self-aware personalized federated learning," *Conference on Neural Information Processing Systems*, 2022.

[30] E. Diao, J. Ding, and V. Tarokh, "GAL: Gradient assisted learning for decentralized multi-organization collaborations," in *Advances in Neural Information Processing Systems*, 2022.

[31] E. Diao, V. Tarokh, and J. Ding, "Privacy-preserving multi-target multi-domain recommender systems with assisted autoencoders," *arXiv preprint arXiv:2110.13340*, 2022.

[32] C. Chen, J. Zhou, J. Ding, and Y. Zhou, "Assisted learning for organizations with limited data," *Transactions on Machine Learning Research*, 2023.

[33] C. Chen, J. Zhang, J. Ding, and Y. Zhou, "Assisted unsupervised domain adaptation," *Proc. International Symposium on Information Theory*, 2023.

[34] H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms," *arXiv preprint arXiv:1708.07747*, 2017.

[35] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, "Reading digits in natural images with unsupervised feature learning," 2011.

[36] L. N. Darlow, E. J. Crowley, A. Antoniou, and A. J. Storkey, "Cinic-10 is not imagenet or cifar-10," *arXiv preprint arXiv:1810.03505*, 2018.

[37] J. Shi, W. Wan, S. Hu, J. Lu, and L. Y. Zhang, "Challenges and approaches for mitigating byzantine attacks in federated learning," in *2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. IEEE, 2022, pp. 139–146.

[38] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of byzantine attacks," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 16–29, 2008.

[39] M. Fang, X. Cao, J. Jia, and N. Z. Gong, "Local model poisoning attacks to byzantine-robust federated learning," in *Proceedings of the 29th USENIX Conference on Security Symposium*, 2020, pp. 1623–1640.

[40] Q. Xia, Z. Tao, Z. Hao, and Q. Li, "Faba: an algorithm for fast aggregation against byzantine attacks in distributed neural networks," in *IJCAI*, 2019.

[41] N. M. Jebreel, J. Domingo-Ferrer, D. Sánchez, and A. Blanco-Justicia, "Defending against the label-flipping attack in federated learning," *arXiv preprint arXiv:2207.01982*, 2022.

[42] W. Gill, A. Anwar, and M. A. Gulzar, "Feddebug: Systematic debugging for federated learning applications," *arXiv preprint arXiv:2301.03553*, 2023.

[43] Q. Li, Y. Diao, Q. Chen, and B. He, "Federated learning on non-iid data silos: An experimental study," in *2022 IEEE 38th International Conference on Data Engineering (ICDE)*. IEEE, 2022, pp. 965–978.

[44] G. F. Jenks and F. C. Caspall, "Error on choroplethic maps: definition, measurement, reduction," *Annals of the Association of American Geographers*, vol. 61, no. 2, pp. 217–244, 1971.

[45] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "Mimic-iii, a freely accessible critical care database," *Sci. Data*, vol. 3, p. 160035, 2016.

[46] X. Wang, Y. Xiang, J. Gao, and J. Ding, "Information laundering for model privacy," *Proc. ICLR*, 2021.

[47] X. Xian, G. Wang, J. Srinivasa, A. Kundu, X. Bi, M. Hong, and J. Ding, "Understanding backdoor attacks through the adaptability hypothesis," *Proc. International Conference on Machine Learning*, 2023.

[48] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Proc. EUROCRYPT*. Springer, 2006, pp. 486–503.