CS402 Final Report: Intersections of Federated Learning, Economics, and Game Theory

Aicha Slaitane*

Abstract

With the increase in the application of machine learning in all aspects of life, concerns about the privacy of the training and testing data have emerged. In response to these concerns, a new technology was introduced: federated learning (FL). Federated learning is an application of machine learning in distributed data environments. In FL, instead of sending the available datasets to a central server for training, the model parameters are sent to data providers. Each provider (called client) trains the model locally with its local dataset and only sends the updated model parameters to the server, thus preserving data privacy. The server then aggregates all the model updates. Despite its advantages, federated learning faces significant challenges particularly when handling non-IID (non-Independent and Identically Distributed) data which makes the global model difficult to converge. Some clients may have limited computation resources and may be concerned about data privacy. Motivating them to participate in the training process is also another challenge of FL. This survey explores existing research aimed at addressing these issues, with a focus on how concepts from economics and game theory can be leveraged to overcome them.

1 Introduction

Artificial Intelligence (AI) has an undeniable significance. Its application in various domains has revolutionized the world, rendering intelligent systems more accessible and cost-effective. Machine learning is also extensively used in domains dealing with sensitive data. For example, hospitals train machine learning models using clients' medical data to predict potential illnesses for individuals. Another example is banks that use clients' data to train models assessing loan eligibility. In such applications of machine learning, where the training data is sensitive, there is a greater emphasis on the privacy of individuals' and companies' data. Recently, regulations and laws on data sharing have increased after data privacy concerns have multiplied, especially with prominent big tech companies like Meta [2]. These regulations include the Data Privacy Act of 2023 in the United States, the General Data Protection Regulation (GDPR) in the European Union, and the Personal Protection Law (PIPL) in China.

As a response to the increase in data privacy concerns, researchers have proposed different solutions. One prominent solution is federated learning. Unlike traditional machine learning which trains the model centrally using the whole training dataset (which contains local datasets from different entities) in a central server, federated learning trains the model locally on each data contributor's local computer(s). This technique enhances data privacy by not sharing individual data sets with a central server, rather each contributor trains the model locally and only sends the updated model's parameters to the server for aggregation to create a global model. Notably, Google's mobile keyboard prediction system represents one of the first pioneering successes of federated learning [7].

Recent years have seen an increase in federated learning research given its importance and high potential. Researchers have investigated optimal ways to aggregate the locally trained models' parameters in the central server, to personalize the global model to each client, to deal with challenges stemming from the nature of the non-IID (non-Independent and Identically Distributed) distribution of the clients' data, to motivate clients to participate in federated learning training and more.

Game theory, an important field in economics, has been leveraged by researchers to address some of the critical challenges of federated learning. It is a mathematical framework used for analyzing situations where multiple decision-makers (clients in FL) interact, each trying to maximize their personal goals called utilities or payoffs. In the context of FL, these goals include minimizing the loss value or computation costs of each client's local model or minimizing the training budget for the central server. In this survey, we present the different applications of contract theory and game theory in solving two of feder-

*Duke Kunshan University, Email: aicha.slaitane@dukekunshan.edu.cn

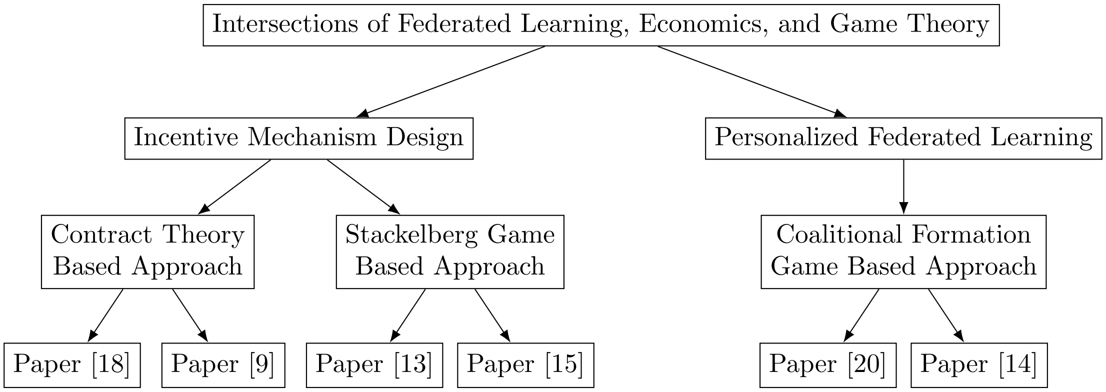

Figure 1: Structure of the survey

ated learning's major challenges: model personalization and incentive mechanism designs. Discussions of these applications are within each corresponding section.

The structure of the survey is as follows:

- First, in the first section, we present the fundamentals of federated learning, contract theory, and game theory. This section is important to better understand the papers discussed later.

- In the second section, we present and discuss two papers that leverage Contract Theory in FL to motivate clients to participate in training.

- In the third section, we present and discuss two papers on the application of Stackelberg Games in FL.

- In the fourth section, we also present and discuss two papers that address the challenge of model personalization in FL using Coalitional Game Theory.

- Finally, we finish we a conclusion and future research directions.

2 Fundamentals of Federated Learning and Game Theory

2.1 Federated Learning

In federated learning, clients aim to train the same machine-learning model for a specific task. For example, if the clients are hospitals, their goal may be to train a model to recognize potential illnesses based on their patients' data. Since each hospital has a limited number of data samples, it collaborates with other hospitals to train the same model under a federated learning framework to preserve the privacy of patients' data. In particular, each client/hospital aims to minimize its local model loss:

$$\min_w \mathcal{L}_i(w) = \frac{1}{n_i} \sum_{j=1}^{n_i} \ell(w; x_j, y_j)$$

where:

$\min_w \mathcal{L}_i(w)$: is the minimization of the local loss function $\mathcal{L}_i(w)$ with respect to the model parameters w.

$\frac{1}{n_i}$: is the normalization term, where n_i is the number of data points on client i.

$\ell(w; x_j, y_j)$: is the loss function for a single data point (x_j, y_j) using model parameters w, which could represent various losses like mean squared error, cross-entropy, etc.

After each client trains the model locally, it sends its updated model parameters w to the server via secure communication. The server then aggregates all the model's parameters updates for all clients K. The local training and aggregation is considered one round of federated learning training. The server can aggregate the model parameters in different ways depending on the distribution of data or the specifics of the task at hand. Here are some common aggregation strategies:

- Average Aggregation:

$$w = \frac{1}{K} \sum_{i=1}^{K} w_i$$

where K is the number of clients and w_i represents the model parameters from client i.

- Clipped Aggregation:

$$w = \frac{1}{K} \sum_{i=1}^{K} \text{clip}(w_i, \tau)$$

where $\text{clip}(w_i, \tau)$ clips the model parameters w_i to a threshold τ.

- Weighted Aggregation:

$$w = \sum_{i=1}^{K} \frac{n_i}{N} w_i$$

where n_i is the number of data points on client i and $N = \sum_{i=1}^{K} n_i$ is the total number of data points across all clients.

As mentioned in the introduction, federated learning faces some critical challenges related to the non-independent and identical distribution of clients' data which affects the convergence of the global/aggregated model and the personalization of clients' local models. On the other hand, clients may be reluctant to share their data and train the global model if they don't have enough incentives to do so. In particular, these clients may have concerns about data privacy leakages during the transmission of model parameters, so it's important to design effective mechanisms to encourage them to participate in training. To tackle these issues, researchers have sought help from important principles in economics. To understand how these principles address the aforementioned challenges, we first give an overview of the important principles discussed in this survey with respect to their application in federated learning.

2.2 Principles in Economics and Game Theory

Several economics and game theory approaches are leveraged to tackle some of the important challenges in Federated Learning. In this survey, we focus on Contract Theory and Game Theory based approaches. We give a general overview of these principles in this section.

2.2.1 Contract Theory

Contract theory is a crucial area within economics that examines how contractual arrangements are structured to align the interests of parties with potentially divergent goals. It addresses the design of formal and informal agreements to ensure that all parties involved have the right incentives to fulfill their obligations effectively. In federated learning, the parties are the participating clients which may have different incentives for the training and this can affect the convergence of the global model.

One of the primary concerns in contract theory is dealing with issues such as moral hazard, adverse selection, and information asymmetry. These issues arise when one client has more or better information than the others, such as more knowledge about the data distribution of the participating clients, potentially leading to suboptimal outcomes. For example, moral hazard occurs when a client is more likely to take risks because the negative consequences are borne by other clients.

Contract theory provides a framework for understanding how contracts can be designed to mitigate these problems. It includes various models and approaches, such as principal-agent models, which analyze the relationship between a principal (e.g., a server) and an agent (e.g., clients) and how to structure incentives to ensure the clients act in the server's best interest [6].

2.2.2 Game Theory

Game theory is a mathematical framework used to analyze strategic interactions among rational decision-makers. In the context of FL, the decision-makers are the server and the clients. Game theory provides tools to model and predict the behavior of individuals or groups in situations where their decisions affect each other [3]. There are different approaches in game theory, notably Stackelberg games, Coalitional games, and non-cooperative games. We present each of these approaches below, but before that, we need to define some important terminology in game theory [8]:

- **Players**: A decision-maker within the game that can take different actions, potentially affecting the situations of decision-makers in the game. In the context of FL, a decision-maker is a client.

- **Strategy**: A strategy of a player within a game is the plan he follows to maximize his best interests.

- **Payoff**: The reward of the player after effectively completing a specific task. In FL, the reward can be a minimal local model loss or a monetary reward for participating in the training.

- **Utility Function**: A utility function computes a numerical value to each possible outcome of a game. This value represents the preference of the client for that outcome [16].

- **Equilibrium**: A situation in the game in which all players can't make any other decisions or take actions.

- **Stackelberg Games**:

  Stackelberg games are a type of strategic game where players take turns making their moves, rather than acting simultaneously. These games involve a leader (the first mover) who commits to a strategy and followers who observe this strategy and make their decisions accordingly [11].

This sequential decision-making process is particularly useful for federated learning when coordinating the actions of multiple clients.

The Stackelberg equilibrium is the solution concept used to determine the outcomes of these games. It is defined by the leader's ability to anticipate the follower's optimal response and choose a strategy that maximizes their own benefit, given this response.

In federated learning, a central server (leader) can set a global model update strategy, which the clients (followers) then observe and use to update their local models. This hierarchical approach helps in efficiently managing the learning process, ensuring that the clients' updates are aligned with the overall objective set by the central server. By anticipating the clients' responses, the central server can optimize the global model updates to achieve better performance and faster convergence.

This method is beneficial in scenarios where the central server needs to guide the learning process while considering the individual contributions and constraints of each client, leading to a more coordinated and effective federated learning system.

- **Coalitional Games**:

  Coalitional games, also known as cooperative games, focus on the formation of coalitions among players. In these games, players can benefit by cooperating and forming groups to achieve a better outcome than they would individually. The main concepts in coalitional games include:

  - **The Core:** A set of payoff distributions where no subset of players can achieve a better outcome by forming a coalition and acting independently.

  - **The Shapley Value:** A method of distributing the total gains to players based on their marginal contributions to the coalition.

  - **The Bargaining Set:** A set of payoff distributions where players negotiate and agree on a fair division of the total gains.

  Coalitional games are particularly relevant in federated learning because they provide a framework for understanding how multiple clients (or devices) can collaborate to achieve a common goal. In federated learning, clients often need to form coalitions to share their local models and data insights without compromising privacy. By forming these coalitions, clients can improve the overall performance of the global model. The concepts of the core and the Shapley value in coalitional games help in determining fair and efficient ways to distribute the benefits of collaboration among the clients. This ensures that all participants are incentivized to contribute their data and computational resources, leading to a more robust and effective federated learning system.

3 Federated Learning Meets Contract Theory

Contract theory, which examines how contractual arrangements are structured to align the interests of clients, is used in federated learning to design effective incentive mechanisms that encourage clients to contribute their data and participate in training. There is a great benefit to incentive mechanism designs as they bring clients with diverse datasets of various distributions. Training the global model with these datasets improves its performance.

Several papers have studied federated learning in the context of contract theory. Tian et al. introduce an incentive mechanism for federated learning (FL) using contract theory to tackle the challenge of creating appropriate incentives for FL clients [18]. The goal of the paper is to minimize the incentive budget while ensuring that clients remain individually rational (IR) and incentive-compatible (IC) during each training round. In particular, the authors design a contract model that considers two private types of clients: data quality and computation effort. The data quality measures the coverage quality of a dataset. The coverage qualities are noted by $\Theta = \{\theta_1, \ldots, \theta_I\}$ and sorted in ascending order

$$\theta_1 \leq \ldots \leq \theta_i \leq \ldots \leq \theta_I$$

The set Φ defines the contract:

$$\Phi = \{\varphi_i = (f_i, R_i(f_i)) \mid i \in \{1, \ldots, I\}\}$$

There are I contracts corresponding to the I quality types. Each contract has a fixed registration fee f_i that clients of type i need to pay in order to participate in the training and a reward $R_i(f_i)$ awarded to clients whose model passes the test by reaching the corresponding generalization accuracy M_i. The authors define the utility functions of the clients and the server. The utility function for player i is given by:

$$U_i = \theta_i e_i R_i - f_i - \frac{c}{2} e_i^2$$

where θ_i is the quality type and e_i^2 the effort of willingness of the client.

The utility function of the server is given by:

$$U_s = \sum_{i=1}^{I} \beta_i U_i = \sum_{i=1}^{I} \beta_i (f_i + \theta_i e_i (G(M_i) - R_i))$$

where β_i is the type distribution of clients and $G(M_i)$ is the revenue generated by the model of client i.

The contract optimization problem this paper is thus:

$$\max \sum_{i=1}^{I} \beta_i (f_i + \theta_i e_i (G(M_i) - R_i))$$

subject to:

$$\text{(IR)} \quad \theta_i e_i R_i - f_i - \frac{c}{2} e_i^2 \geq 0$$

$$\text{(IC)} \quad \theta_i e_i R_i - f_i - \frac{c}{2} e_i^2 \geq \theta_i e_{ji} R_j - f_j - \frac{c}{2} (e_{ji})^2$$
$$\forall j \neq i, \ i, j \in \{1, \ldots, n\}$$

The IR (Individual Rationality) constraint ensures that the utility of each client is positive, and the IC (Individual Compatibility) constraint ensures that the clients receive the highest utility when they choose the contract corresponding to their type.

The authors solve this optimization after simplifying the constraints by leveraging the monotonicity between f_i, θ_i, and R_i.

There is a big limitation in this paper with regard to the optimization problem. The main objective function mainly considers the budget, the rewards, and the fees of the game, and does not emphasize the importance of the performance of the models. It's true that they refer to the performance with the generalization threshold M_i but it is still ambiguous. A better approach would consider a convergence bound of the loss function of the global model and aim to minimize this bound. This approach would be more mathematically rigorous.

A second paper that integrates contract theory in the incentive mechanism design of federated is that of Kang et al [9]. The authors of this paper solve a similar problem as that of the previous one but introduce different metrics in the utility functions of the clients and the server and use a different definition of the quality type. They define the quality type in terms of the computation resources of the clients. This paper introduces a reputation metric to evaluate and select reliable players. This metric is managed using blockchain technology to ensure security and tamper-resistance. We will not go into details on how the reputation score is measured. This paper also introduces a measure for the computational model for federated learning that relies on the CPU cycle frequency from each client. In particular, they calculate the energy consumption of a client i in a single iteration as

$$E_{\text{cmp},i}(f_i) = \zeta c_i s_i f_i^2$$

where ζ is a computation coefficient, s_i is the number of data points of client i, c_i the CPU cycle for training one data point, and f_i CPU frequency of type i client.

The utility function of the server (called task publisher in the paper) is given as:

$$\max_{(R_n, f_n)} U_{TP} = \omega \ln (T_{\max} - T_{t,i}) - lR$$

where ω is the satisfaction degree parameter of the task publisher. $T_{\max}$ is the task publisher's maximum tolerance time of federated learning, and l is the unit cost of the rewards for the workers. $T_{t,i}$ is the time it takes a client to train the local model in one iteration and a function of f_i, c_i, s_i, and θ_i.

The utility function of the clients is given as:

$$\max_{(R_n, f_n)} U_D = R_n - \mu \left[\frac{\psi}{\theta_n} \zeta c_n s_n f_n^2 + \sigma \frac{\rho_n}{B} \ln \left(1 + \frac{\rho_n h_n}{N_0} \right) \right]$$

This function contains some of the same parameters as those in the utility function of the server. The other parameters are not that relevant and will not be explained on this survey.

Given these utility functions, the author define the optimization problem, which includes individual rationality and incentive compatibility constraints. Similar [18], the authors of this paper leverage the monotonicity priorities of some of the parameters in the utility functions to simplify the constraint and solve the optimization problems.

The limitation of this paper is similar to that of the previous one. They also focus on maximizing profits rather than improving the performance of the models.

The last paper discussed in this survey that integrates contract theory with federated learning is [21]. This paper builds on the approach used by [9], but introduces the additional complexity of multiple task publishers, each with its own set of contracts. This scenario presents a significant challenge as clients must navigate and select the optimal contract from the various options provided by different servers. The authors address this by developing a mechanism that ensures clients can efficiently choose the best contract, thereby optimizing their participation in the federated learning process. This approach not only enhances the flexibility and scalability of federated learning systems but also improves the overall efficiency and effectiveness of the learning process.

4 Federated Learning Meets Stackelberg Games

Under the same umbrella of incentive mechanism designs in federated learning, in this section, we discuss some of the mechanisms that leverage principles from Stackelberg games.

First, in [13], Luo et al. address two important issues in FL at the same time: incentive mechanism design and randomized client participation. In general, in federated learning, it is impractical for all clients to participate in the training in each round because of the limited computation resources they may have or their unavailability (phones as clients may be out of battery). That's why, it is often the case that either all clients participate in the training (full clients participation) or a deterministic subset of clients is sampled based on their data quantity, computation, or communication resources and participates in all training rounds [12]. However, these practices can lead to a biased global model because of the non-iid distribution of clients' datasets. To address the problem of client sampling, Luo et al. propose an innovative approach that allows the server to sample clients based on their participation levels (joining probabilities) q. They prove that this approach results in an unbiased global model. In addition to the randomized client sampling, the authors design an incentive mechanism that leverages the participation levels to optimize the utility functions of the clients and the server. They solve the problem using a two-stage Stackelberg game approach. Two other important contributions of this paper are the intrinsic value and the bi-directional payment. The intrinsic value denotes the preference level of model improvements. It is formulated as:

$$V_i := v_i \left(F(w_i^*) - E[F(w_{R(q)})] \right)$$

where $F(w_n^*)$ is the optimal local model loss of client i, $E[F(w_{R(q)})]$ is the expected loss of the global model after R rounds, and v_n is the preference level for improvement for client i. Thanks to the intrinsic value, the authors introduce the concept of bi-directional payment where not only does the server reward the clients for training, but the clients pay the server to join the training depending on their intrinsic value. If a client's intrinsic value is high, he has a greater motive to train the model and benefit from it, thus he pays the server to join the training.

The authors in [13] formulate the utility functions of the server as:

$$\min_P \; U_s(P, q) := E[F(w_{R(q)})],$$

subject to:

$$\sum_{i=1}^{N} P_i q_i \leq B.$$

where P_i is the payment from the server to client i, q_i is the participation level of client i, and B is the budget of training set by the server. The budget constraints the payment amounts to the clients.

The utility function of the clients is given as:

$$\max_{q_i} \; U_i(q_i, P_i) := P_i q_i - C_i + V_i$$

subject to:

$$0 \leq q_i \leq q_{i,\max}.$$

where C_i is the cost of training of client i and V_i is its intrinsic value.

The authors solve this problem using a two-stage Stackelberg game approach where the server is the leader and the clients are the followers. In stage I, the server determines the budget B and the payments $\mathbf{P} = (P_1, \ldots, P_N)$. In stage II, the clients determine their participation levels q_i given the payment P_i.

Additionally, the authors study the Stackelberg equilibrium of the game, the impact of the server's budget, and that of the client's parameters. They summarize the results in theorems and present several experiments.

One of the first papers on motivating clients in federated learning under a Stackerlberg game approach is [15]. In this paper, Sarikaya and Ercetin address the issue of the limited computation resources of clients, which may hinder their participation in FL training. Usually, clients divide their local resources into different local tasks, and they need to decide how much resources they can allocate to FL. Specifically, the authors try to solve the following question: "How much Central Processor Unit (CPU) resource of heterogeneous workers should be allocated to the training task of the model owner?" [15]. Similar to [13], they derive the solution with a two-stage Stackelberg game, where the server or task publisher is the leader and the clients are the followers. The main parameter in their optimization problem is the time elapsed $T_{i,t}$ for a client i to finish local training in round t and update the gradients. $T_{i,t}$ for each client is assumed to be random, independently distributed, and is exponentially distributed with mean $\frac{P_i}{c_i}$ where P_i is the computation power (represented in the CPU power) and c_i is the number of CPU cycles necessary to complete the training task. The utility function of the clients is given as

$$U_i(P_i, q_i) = q_i P_i - \kappa c_i(P_i)^2$$

κ is a parameter depending on the hardware design of the clients and q_i is the price of one unit of the client's CPU power. The utility function of the server:

$$\min_q \Delta = VE\left(\max_i T_{i,t}\right) + \sum_{i=1}^{K} q_i P_i$$

subject to

$$\sum_{i=1}^{K} q_i P_i \leq B$$

where V is a positive constant optimization parameter.

In the first stage of the game, the authors solve for the optimal solution of the CPU power P_i for each client using the first derivative of the utility function. The optimal P_i value is a function of [?]. Then, in the second stage, the authors implement an efficient update algorithm to reach the equilibrium point and solve for the optimal values of q_i.

This paper presents an interesting approach to designing an incentive mechanism with Stackelberg games, however, it lacks a deeper analysis of the convergence of the local loss functions for each client which makes the implementation of this approach impractical in real work scenarios.

5 Federated Learning Meets Coalitional Games

As mentioned in the introduction, federated learning faces a significant challenge stemming from heterogeneous data distributions across clients, often referred to as non-IID (non-Independent and Identically Distributed) data. Personalized Federated Learning (PFL) addresses this challenge by tailoring the global model to better fit the specific data distribution of each client. The goal is to combine the generalization capabilities of the global model with the specificity of local models. Some common approaches of PFL involve fine-tuning [17], meta-learning [5], clustered models [17], and coalition-based models [4].

In this section, we discuss two papers on the application of Coalitional Game Theory (also known as Cooperative Games) in federated learning.

In [20], Wu et al. address the challenge of statistical heterogeneity in clients' local data. The authors model the problem of client collaboration in federated learning as a coalition formation game. Each client decides whether to join a coalition based on the potential benefits in terms of improved model performance. The Shapley value is used to fairly distribute the gains from collaboration among the clients. This value quantifies each client's contribution to the coalition's overall performance. In the paper, a client's i contribution to all possible coalitions is measured by

$$\phi_i(v) = \frac{1}{|N|!} \sum_{\pi \in \Pi} \left[v(C_{\pi(i)} \cup \{i\}) - v(C_{\pi(i)}) \right]$$

where π refers to a coalition, Π represents the set of all coalitions in the game, and the function $v(\pi) : 2^n \to R$ assigns to every coalition $\pi \subseteq \Pi$ a real number representing the gain obtained by the coalition.

To determine the coalitions' structure, the authors implement an iterative algorithm that forms and adjusts coalitions by maximizing the Shapley value for each client. Clients continuously update their models based on the aggregated results from their respective coalitions. Within each coalition, the algorithm performs personalized model aggregation. This means that the global model is adapted to better fit the specific data distribution of each client in the coalition. The algorithm keeps running until convergence, i.e. the game has reached an equilibrium state and no client has an incentive to diverge or leave the game.

The benefits of this paper rest in improved personalization, fair collaboration, and scalability. By forming optimal coalitions and using personalized model aggregation, the algorithm significantly improves the accuracy of the local models for individual clients which is proved in the experiments. pFedSV outperforms all other personalization algorithms (implemented prior to the publication of this paper) significantly. Additionally, the use of Shapley value ensures that the benefits of collaboration are fairly distributed among clients, encouraging participation and cooperation. Moreover, the algorithm is robust and can handle a large number of clients with different data distributions.

In [14], the authors also leverage coalitional game theory, or more precisely hedonic coalition formation game, in federated learning. In this paper, the authors propose a serverless hierarchical FL framework where clients are organized into coalitions. Each coalition has a leader who coordinates the learning process within the coalition and communicates with other coalition leaders. This hierarchical structure aims to reduce communication overhead and improve scalability. The paper introduces a reputation-aware mechanism to form coalitions. Clients with higher reputations are more likely to be selected as coalition leaders. Reputation is based on factors such as historical performance, reliability, and contribution to the learning process. The serverless environment leverages edge computing resources, which reduces the dependency on centralized servers and enhances the robustness of the FL process.

The authors use a special type of coalition formation process called hedonic coalition formation. It is a concept in cooperative game theory that focuses on the formation of coalitions (groups) based on the preferences of individual players [19]. The term "hedonic" refers to the idea that each player's satisfaction or utility depends solely on the members of their own coalition, rather than on the overall structure of the game or the payoffs distributed among players.

The serverless hierarchical federated learning framework adopted in this paper is composed of a two-layer architecture. In the top layer, the cluster heads exchange parameters, while in the bottom layer, the clients are grouped into coalition where they train their local models under the coalition heads.

The cluster heads are elected based on their reputation scores. The reputation score is calculated using a multiweight subjective logic scheme [10]. A cluster head with higher reputation score attracts more clients as it is viewed as a non-malicious and reliable in performing the intermediate model aggregation and relay.

The authors conduct extensive experiments to evaluate the performance of the proposed approach. The results demonstrate that the reputation-aware hedonic coalition formation significantly improves the efficiency and effectiveness of the hierarchical FL process. The approach outperforms traditional FL methods in terms of communication overhead, convergence speed, and model accuracy.

There are nonetheless some limitations to [14]. The serverless architecture, while reducing the risk of a single point of failure, incurs significant communication overheads. This can be a bottleneck, especially in large-scale deployments where frequent communication between cluster heads and workers is required. Additionally, the effectiveness of the reputation-aware mechanism heavily depends on the accuracy and reliability of the reputation scores. If the reputation system is compromised or manipulated, it could lead to suboptimal coalition formations and reduced overall performance.

6 Conclusion and Future Work

In summary, federated learning is one of the prominent technologies in machine learning that aims to preserve data privacy in distributed data environments. It was applied in many important applications like Google's GBoard prediction feature [7]. However, it still faces some challenges, notably because of the non-IID nature of the local datasets of data providers. This results in model personalization challenges for the clients. Additionally, because of potential privacy leakage during model parameters communication or global model aggregation, fairness issues, and computation cost limitations, some clients may be reluctant to participate in FL. To encourage the client to participate, a direction of research in FL has been created that specifically designs incentive mechanisms. Some researchers have leveraged important fields in economics to address these challenges. Notably, contract theory, stackelberg games, and coalitional formation games have been used to design incentive mechanism designs and to implement more accurate and robust personalized models. These techniques have proved to be successful in research. However, data sharing between clients and the server is still considered extremely challenging even in academia. That is one of the main reasons that federated learning is not widely used in industry. To overcome data sharing challenges, some researchers have proposed to construct AI and data markets [1]. In these markets, researchers can exchange AI models and datasets while being governed by strict sharing rules. This is another interesting research direction that is gaining more attention.

References

[1] Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar. Too much data: Prices and inefficiencies in data markets. *American Economic Journal: Microeconomics*, 14(4):218–256, 2022.

[2] Diane Bartz. Meta, us government spar in court over toughened privacy order. *Reuters*, October 2023. https://www.reuters.com/legal/meta-us-government-spar-court-over-toughened-privacy-order-2023-10-17/.

[3] Femke Bekius and Sharlene L. Gomes. A framework to design game theory-based interventions for strategic analysis of real-world problems with stakeholders. *European Journal of Operational Research*, 309(2):925–938, 2023. Accessed October 12, 2024.

[4] Kate Donahue and Jon Kleinberg. Model-sharing games: Analyzing federated learning under volun-

tary participation. *arXiv preprint arXiv:2010.00753*, 2020.

[5] Alireza Fallah, Aryan Mokhtari, and Asuman Ozdaglar. Personalized federated learning: A meta-learning approach. In *Advances in Neural Information Processing Systems*, pages 6965–6975. Springer, 2020.

[6] James Gordley and Hao Jiang. The maze of contemporary contract theory and a way out. *The American Journal of Jurisprudence*, 68(1):1–32, 2023.

[7] Andrew Hard, Kanishka Rao, Rajiv Mathews, Swaroop Ramaswamy, Françoise Beaufays, Sean Augenstein, Hubert Eichner, Chloé Kiddon, and Daniel Ramage. Federated learning for mobile keyboard prediction. *ArXiv*, 2018. Accessed February 20, 2024.

[8] Investopedia. Game theory, 2024. Accessed: 2024-10-12.

[9] Jiawen Kang, Zehui Xiong, Dusit Niyato, Shengli Xie, and Junshan Zhang. Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory. *IEEE Internet of Things Journal*, 6(6):10700–10714, 2019.

[10] Jiawen Kang, Zehui Xiong, Dusit Niyato, Shengli Xie, and Junshan Zhang. Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory. *IEEE Internet of Things Journal*, 6(6):10700–10714, Dec 2019.

[11] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3):50–60, 2020.

[12] Wee Y. B. Lim, Zehui Xiong, Chunyan Miao, Dusit Niyato, Qiang Yang, Cyril Leung, and H. Vincent Poor. Hierarchical incentive mechanism design for federated machine learning in mobile networks. *IEEE Internet of Things Journal*, 7(10):9575–9588, 2020.

[13] Bing Luo, Yutong Feng, Shiqiang Wang, Jianwei Huang, and Leandros Tassiulas. Incentive mechanism design for unbiased federated learning with randomized client participation. *Information Sciences*, 664:120330, 2024. Accessed October 12, 2024.

[14] Jer Shyuan Ng, Wei Yang Bryan Lim, Zehui Xiong, Xianbin Cao, Jiangming Jin, Dusit Niyato, Cyril Leung, and Chunyan Miao. Reputation-aware hedonic coalition formation for efficient serverless hierarchical federated learning. *IEEE Transactions on Parallel and Distributed Systems*, 33(11):2675–2686, 2022.

[15] Yunus Sarikaya and Ozgur Ercetin. Motivating workers in federated learning: A stackelberg game perspective. *arXiv preprint arXiv:1908.03092*, 2019.

[16] Branislav L. Slantchev. Preferences and expected utility, 2012.

[17] Alysa Ziying Tan, Han Yu, Lizhen Cui, and Qiang Yang. Towards personalized federated learning. *arXiv preprint arXiv:2103.00710*, 2021.

[18] Mengmeng Tian, Yuxin Chen, Yuan Liu, Zehui Xiong, Cyril Leung, and Chunyan Miao. A contract theory based incentive mechanism for federated learning. *arXiv preprint arXiv:2108.05568*, 2021.

[19] Lexing Wang, Tenghai Qiu, Zhiqiang Pu, Jianqiang Yi, Jinying Zhu, and Wanmai Yuan. Hedonic coalition formation for distributed task allocation in heterogeneous multi-agent system. *International Journal of Control, Automation and Systems*, 22:1212–1224, 2024.

[20] Leijie Wu, Song Guo, Yaohong Ding, Yufeng Zhan, and Jie Zhang. A coalition formation game approach for personalized federated learning. *arXiv preprint arXiv:2202.02502*, 2022.

[21] Shichang Xuan, Mengda Wang, Jingyi Zhang, Wei Wang, Dapeng Man, and Wu Yang. An incentive mechanism design for federated learning with multiple task publishers by contract theory approach. *Information Sciences*, 664:120330, 2024. Accessed October 12, 2024.