Proceedings of the 5th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/99/20251774

# A Survey of Incentive Mechanisms for Federated Learning

**Yuzhe Zhou**[1,a,*]

[1]*Queen Mary School Hainan, Beijing University of Posts and Telecommunications, Beijing, China*
*a. shuaimike@bupt.edu.cn*
**corresponding author*

***Abstract:*** Federated learning is a cutting-edge distributed machine learning technique that lowers the possibility of data leaking from centralized uploads by enabling users to cooperatively train models. It not only protects client privacy but also better utilizes decentralized data resources, demonstrating high learning performance. To employ their models in the training process, however, a significant number of clients must be involved, which inexorably uses up the device's resources—computational power, communication bandwidth, and energy. Therefore, incentivizing more clients to participate in federated learning is crucial. Typical methods can provide numerical or model-based compensation to ensure they contribute their valuable model resources. This paper provides a brief introduction to federated learning and incentive mechanisms, and surveys related research on federated learning incentives. Specifically, this paper categorizes existing federated learning incentive mechanisms into four technical approaches: Shapley values, Stackelberg games, auctions, and contracts. Finally, the paper discusses some future directions for incentivizing clients in federated learning.

***Keywords:*** Federated learning, Incentive Mechanism, Survey.

## 1. Introduction

As a novel distributed learning paradigm, federated learning effectively addresses data privacy issues. During the federated learning training process, the central server first initializes a global model, including its initial parameters and structure, and distributes the global model and its parameters to the local devices or clients participating in federated learning. Each participant then uses their local dataset for training. After completing a certain number of local training iterations, they send the updated model parameters back to the central server. Upon receiving all local device model updates, the central server aggregates all parameters to generate new global parameters and begins the next round of training. This process iterates until predefined stopping conditions are met. From the above training process, it can be seen that federated learning significantly protects participants' data privacy since there is no need to upload raw data.

Ever since its conception, federated learning has attracted a lot of interest from the academic and business communities. To address the issue of non-independent and identically distributed data, McMahan et al. proposed the FedAvg method based on iterative model averaging [1]. The core idea is to upload the parameters of local models to the server, assign a weight to the model parameters uploaded by each device, perform a weighted average, and then broadcast this average back to all local devices. This process can be iterated multiple times until convergence. Federated learning primarily uses gradient descent to optimize the loss function. In their work, Zheng et al. introduced the concept of update age to calculate how long it has been since the server last received a client's most recent update [2]. As a result, the global model and local data are not excessively disconnected from one another, which speeds up the convergence of the model. The server is able to receive federated learning (FL) training updates from clients in a timely manner.

Currently, federated learning systems are very dependent on the caliber of local client updates. However, without appropriate compensation, clients may not continue to participate in the learning process. Additionally, federated learning is susceptible to attacks during training. In paper [3], Cao et al. found that the attack success rate is linearly related to the number of poisoned training samples and attackers. Malicious agents have the ability to take control of the local models of several clients, and malicious clients have the ability to introduce malicious data into the local model processing, which can eventually be used to manipulate the global model. Owing to these hazards, unless they obtain enough incentives, clients could be less inclined to engage in federated learning activities.

## 2. Background

### 2.1. Federated Learning

Over the past few years, artificial intelligence (AI) has developed rapidly, from facial recognition and AlphaGo defeating human Go player Lee Sedol to autonomous driving. AI has permeated various aspects of our lives. However, researchers often overlook the fundamental principle of AI: it is trained on vast amounts of data, especially high-quality data. In reality, apart from a few giant companies, most enterprises face the problem of insufficient data. Therefore, the secure flow of data is an inevitable trend. However, data exchange also involves interest exchange, and data often appears in isolated silos between companies or even departments within a company. To address data silos and privacy protection issues, federated learning has emerged.

In recent years, AI has advanced rapidly, from facial recognition and AlphaGo defeating human Go player Lee Sedol to autonomous driving. AI has permeated various aspects of our lives. However, researchers often overlook the fundamental principle of AI: it is trained on vast amounts of data, especially high-quality data. In reality, apart from a few giant companies, most enterprises face the problem of insufficient data. Therefore, the secure flow of data is an inevitable trend. However, data exchange also involves interest exchange, and data often appears in isolated silos between companies or even departments within a company. To address data silos and privacy protection issues, federated learning has emerged.

The following describes the usual federated learning architecture and training procedure. Initialization is the first step. In accordance with the training job, the parameter server selects the global training model, pre-trains it on a public dataset, and then distributes the first model parameters to the chosen clients. Local Training is Step 2. Every client uses the starting settings to train on its local dataset. Step 3: The parameter server receives the modified model parameters. Step 4: Aggregation of Global Model Updates. The parameter server updates the global model with the averaged model after averaging the uploaded local update parameters for each round. Until a few predetermined requirements are satisfied, this process is repeated.

Based on network topology, data availability, and data partitioning, federated learning can be categorized as follows:

1. Network Topology-Based Classification: Federated learning has two common network topology types: centralized configuration based on a parameter server and decentralized configuration. The former divides the computing machines into two roles: parameter server and client. The parameter server provides model aggregation and distribution functions, while clients hold training data and provide the necessary computational power. The latter does not distinguish between

DOI: 10.54254/2755-2721/99/20251774

parameter servers and clients; each node has the same functionality, undertaking computational tasks while interacting with data from other nodes.

2. Data Availability-Based Classification: Federated learning can be divided into cross-silo federated learning and cross-device federated learning based on data availability. The former refers to training on isolated datasets with fewer clients, where data is partitioned by samples or features. The latter typically involves a larger number of clients with lower reliability.

Data Partitioning-Based Classification: Based on data partitioning, federated learning may be divided into three categories: vertical federated learning, horizontal federated learning, and federated transfer learning. Similar to horizontally dividing data in a table view, where each participant has separate user data but the same data characteristics, horizontal federated learning is appropriate for situations when each participant's datasets have the same feature space but distinct sample spaces. Vertical federated learning is applicable when two datasets have a high overlap of users but low overlap of user features. The datasets are vertically partitioned (i.e., by feature dimension), and the overlapping user data with different features is used for training. Federated transfer learning combines the concept of transfer learning to handle situations where there are only a few overlapping samples and features between participants' datasets, and the datasets differ significantly in scale and distribution.

## 2.2. Incentive mechanism

Using certain techniques and management systems to maximize workers' dedication to the company and their jobs is known as an incentive mechanism. It includes the structure, procedures, connections, and patterns of development of interactions and limitations between the incentive object and the incentive subject system, which utilizes different incentive ways and standardizes them. The incentive mechanism serves as a bridge in order for businesses to turn ideas into tangible outcomes, aiming to attain a win-win outcome between both corporate and individual objectives by stimulating employees' enthusiasm and creativity. However, due to the heterogeneity of devices and data in federated learning systems, it is challenging to effectively evaluate user contributions and dynamically adjust the system. This prevents the direct application of incentive mechanism designs from other fields to federated learning. The main issue currently faced in this field is how to design incentive mechanisms tailored to the unique characteristics of federated learning.

# 3. Incentive mechanism in federated learning

This section primarily classifies current incentive mechanisms from a technical perspective, including Shapley values, Stackelberg games, auctions, and contracts. These four incentive mechanisms all have significant advantages in terms of fairness. However, the applicability of federated learning incentive mechanisms based on Shapley values may be limited by the federated learning scenario. For example, in vertical federated learning, the direct application of Shapley values may be challenging due to the vertical partitioning of data features. Federated learning incentive mechanisms based on Stackelberg games allow leaders and followers to dynamically adjust their strategies according to each other's actions, achieving optimal resource allocation. However, there may be cooperation risks in practical applications. For instance, if there are trust issues or conflicts of interest between mobile devices or between model owners and mobile devices, the incentive mechanism may fail or perform poorly. Auction-based federated learning incentive mechanisms promote the efficient allocation of resources (such as computational resources and data resources) by allowing them to flow to the highest bidder or the most significant contributor through market competition. Contract-based federated learning incentive mechanisms clearly define the rights and obligations of participants, helping to establish long-term stable cooperative relationships. When formulating and executing contracts, it is necessary to ensure compliance with relevant laws and regulations. In federated learning, this may require additional legal compliance costs and may face legal risks and disputes.

### 3.1. Shapley value

The Shapley value is a classic concept in cooperative game theory used for the fair distribution of cooperative gains. It defines a unique payoff scheme and is often used in machine learning to interpret model predictions and measure the contribution of features to the prediction results. To accurately assess participants' contributions to federated learning, Song et al.[4]proposed a contribution index based on the Shapley value, aiming to evaluate each data provider's contribution to federated learning. However, directly calculating the contribution index requires significant time and effort. Therefore, the authors approximate the reconstruction of models from different dataset combinations using intermediate results from the training process, avoiding additional training.

Moreover, calculating the Shapley value in federated learning often faces two issues: first, it requires detailed evaluation of each data source subset's model performance, incurring high communication costs; second, given that the relevance of a data source may vary depending on when it is trained, it goes against the sequential structure of federated learning by ignoring the order in which the data sources are used during training. Federated Shapley Value (FSV), a variation of the Shapley value appropriate for federated learning, was developed by Wang et al. [5]. While capturing the effect of the participation order on data value and enabling computation without extra communication costs, FSV maintains the beneficial qualities of the standard Shapley value.

Yang et al.[6] designed a supervised fuzzy Shapley value incentive mechanism, which successfully achieves fair and Pareto-efficient optimal payoff distribution. Besides computational issues, directly using the Shapley value can expose feature values or data sample distributions. Wang et al. [7]were the first to identify this threat in vertical federated learning and introduced a variant called the Shapley Group Value to evaluate the utility of feature subsets while avoiding the disclosure of any private feature information in vertical federated learning. Specifically, they integrate several private features into a joint feature and calculate the Shapley Group Value of this joint feature in a two-participant scenario. As the number of participants increases, the computation process becomes considerably complex.

### 3.2. Stackelberg game

Game theory can be divided into two types based on the mode of play: static games and dynamic games. Static games assume that all participants make decisions simultaneously, while dynamic games consider participants making decisions in a sequential order. In federated learning, the relationship between the federation and participants is typically led by the federation, with participants following suit. Therefore, this relationship can be modeled as a Stackelberg game. Guo et al. [8]utilized the Stackelberg game to design an incentive mechanism for federated learning (IMD). The clients follow the server, who serves as their leader. They played a game to maximize their profits and suggested an enhanced NSGA-II. They did this by combining NSGA-II with GA to determine the Stackelberg game's Nash equilibrium. The Nash equilibrium states that it can offer methods for the clients as well as the server. In order to encourage edge nodes to participate in model training, Qin et al. [9]developed an incentive mechanism for federated learning. balancing costs and training under a limited budget. Specifically, they modeled the utility between the server and edge nodes and used the Stackelberg game to obtain the optimal solution for the utility model. Considering data leakage issues, Yi et al.[10] preserved the differential privacy system in wireless federated learning and maximized server utility based on the Stackelberg game.

DOI: 10.54254/2755-2721/99/20251774

### 3.3. Auction

Auction-based mechanisms, due to their simplicity, enable federations to easily obtain private information from participants. The auction framework known as AFL was created by Pang et al.[11]. It works by breaking down the social cost minimization problem into a set of Winner Determination Problems (WDP) that are determined by the number of global iterations. AFL uses a payment algorithm to compute the winners' prizes and a greedy algorithm to identify the winners in order to solve each WDP. Ultimately, AFL provides the best answer among all WDPs. Combinatorial Auction and Bargaining (CAB) is a two-stage federated learning incentive system that Xu et al.[12] devised to improve the overall utility of the federated learning platform and mobile users. The "Combinatorial Auction" stage and the "Bargaining" stage are the two stages of the mechanism. This mechanism is superior to other baseline mechanisms and yields higher platform profits than other mechanisms. Its effectiveness is supported by theoretical and numerical analysis, which also demonstrates that it is individually rational and incentive-compatible. The FAIR federated learning system was proposed by Deng et al.[13]. Reverse auctions are used as a model for the FAIR system's incentive mechanism, which motivates top-tier users to take part in learning.

### 3.4. Contract

Contract theory is a branch of economics that posits moral norms and political order stem from rational agreements or contracts. The theory asserts that individuals, in a hypothetical original state, voluntarily enter into contracts to constrain their behavior to achieve common interests and avoid conflicts. An incentive program with information asymmetry was put out by Cao et al. [14]and was based on the principle of a two-period dynamic contract. By balancing the model owner's weighted preferences for service latency and Age of Information (AoI), this strategy aims to increase the utility of the model owner by incentivizing additional data owners to take part in model training. Compared to traditional contracts and uniform pricing strategies, the model owner gains more profit from the proposed contract. However, existing research on federated learning incentive mechanisms only considers scenarios with a single task publisher and multiple worker nodes. In scenarios with multiple task publishers and multiple worker nodes, competition among different task publishers complicates the entire research process, making it impossible to directly apply contract designs from single task publisher scenarios. An incentive mechanism based on contract theory for multi-task publisher situations was presented by Xuan et al. [15]in order to overcome this problem. This mechanism can incentivize worker nodes to join and increase the effectiveness of federated learning.

## 4. Future studies of incentive in federated learning

Future incentive schemes should focus on enhancing the overall performance of FL at a low cost by attracting more participants. Firstly, the ultimate goal of any incentive mechanism is undoubtedly to improve FL performance. Otherwise, even if the scheme attracts more participants, it would be meaningless. Another crucial aspect is that the designed incentive scheme should remain lightweight, as resource-constrained nodes are often hesitant to engage in high-cost incentive methods.

Researchers should place greater emphasis on cross-silo FL. The decision-making processes of large enterprises or organizations differ significantly from those of end-users and mobile devices, necessitating a novel cross-silo FL incentive mechanism. Furthermore, the widespread application of FL in cross-silo environments makes incentive schemes increasingly important and indispensable.

Carefully managing the number of local epochs to regulate the number of communication rounds through incentive mechanisms is not only a daunting task but also a challenging one. Moreover, the computation and communication costs for each client in each training round may vary, further complicating the design of incentive mechanisms that drive federated learning. To our knowledge, there are currently no research results on this topic.

## 5. Conclusion

A key component of the architecture of new federated learning systems is incentive mechanisms. The paper provides an in-depth analysis of current approaches and suggest avenues for further research. First, the paper presents the basic ideas of federated learning and incentive systems. Then, the paper conducts a detailed review, analysis, and comparison of solutions to emerging implementation challenges when creating incentive structures for federated learning. These challenges include Shapley values, Stackelberg games, auctions, and contracts. Finally, the paper proposes several future research directions, such as how to apply incentive mechanisms to cross-silo FL environments and how to regulate the number of communication rounds through incentive mechanisms. In summary, incentive mechanisms play a vital role in federated learning systems as they provide a sufficient number of clients for the system to function in practice. With the rapid proliferation of machine learning applications, developing effective and efficient incentive mechanisms for federated learning is a dynamic new field. The paper hopes this survey will encourage more researchers to work in this area.

## References

[1] *McMahan, H. B. et al. (2017). Communication-efficient learning of deep networks from decentralized data. in Proc. Artif. Intell. Statist. Conf., pp. 1273–1282*
[2] *Zheng, Z., Hong, Y., Li, K., Chen, Q. (2024). An Incentive Mechanism Based on AoU, Data Quality, and Data Quantity for Federated Learning. 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Tianjin, China, pp. 1832-1837.*
[3] *Cao, D., Chang, S., Lin, Z., Liu, G. and Sun, D. (2019). Understanding Distributed Poisoning Attack in Federated Learning. 2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS), Tianjin, China, pp. 233-239.*
[4] *Song, T., Tong, Y. and Wei, S. (2019). Profit Allocation for Federated Learning. 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, pp. 2577-2586.*
[5] *Wang, T., Rausch, J., Zhang, C., Jia, R., Song, D. (2020). A Principled Approach to Data Valuation for Federated Learning. In: Yang, Q., Fan, L., Yu, H. (eds) Federated Learning. Lecture Notes in Computer Science, vol 12500. Springer, Cham.*
[6] *Yang, X., Xiang, S., Peng, C., Tan, W., Yue, W., Liu. H., and Ding, H. (2024). Federated Learning Incentive Mechanism with Supervised Fuzzy Shapley Value. Axioms, 13, no. 4: 254.*
[7] *Wang, G., Dang, C. X. and Zhou, Z. (2019). Measure Contribution of Participants in Federated Learning. 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, pp. 2597-2604.*
[8] *Guo, W., Wang, Y., Jiang, P. (2023). Incentive mechanism design for Federated Learning with Stackelberg game perspective in the industrial scenario. Computers & Industrial Engineering, 184: 0360-8352.*
[9] *Qin, S., Ye, J., Tang, X. and Li, X. (2023). Stackelberg Game-Base Incentive Scheme for Federated Learning in Artificial Intelligence of Things. 2023 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Danzhou, China, pp. 159-165.*
[10] *Yi, Z., Jiao, Y., Dai, W., Li, G., Wang, H. and Xu, Y. (2022). A Stackelberg Incentive Mechanism for Wireless Federated Learning With Differential Privacy. in IEEE Wireless Communications Letters, vol. 11, no. 9, pp. 1805-1809.*
[11] *Pang, J., Yu, J., Zhou, R. and Lui, J. C. S. (2023). An Incentive Auction for Heterogeneous Client Selection in Federated Learning. in IEEE Transactions on Mobile Computing, vol. 22, no. 10, pp. 5733-5750.*
[12] *Xu, B., Zuo, L., Jin, J., et al. (2023). CAB: a combinatorial-auction-and-bargaining-based federated learning incentive mechanism. World Wide b, 26, 2351–2372.*
[13] *Deng, Y. et al. (2021). FAIR: Quality-Aware Federated Learning with Precise User Incentive and Model Aggregation. IEEE INFOCOM 2021 - IEEE Conference on Computer Communications, Vancouver, BC, Canada, pp. 1-10.*

DOI: 10.54254/2755-2721/99/20251774

*[14] Cao, M., Wang, Q., Wang, Q. (2024). Federated learning in smart home: A dynamic contract based approach with task preferences. Computer Networks, 249: 1389-1286.*
*[15] Xuan, S., Wang, M., Zhang, J., Wang, W., Man, D., Yang, W. (2024). An incentive mechanism design for federated learning with multiple task publishers by contract theory approach. Information Sciences, 664: 0020-0255.*