FedCompetitors: Harmonious Collaboration in Federated Learning with Competing Participants

Shanli Tan[1*], Hao Cheng[2*], Xiaohu Wu[1*†], Han Yu[3*], Tiantian He[4†], Yew-Soon Ong[3,4], Chongjun Wang[2], Xiaofeng Tao[1]

[1]National Engineering Research Center of Mobile Network Technologies, Beijing University of Posts and Telecommunications, China
[2]State Key Laboratory for Novel Software Technology, Nanjing University, China
[3]School of Computer Science and Engineering, Nanyang Technological University, Singapore
[4]Agency for Science, Technology and Research, Singapore
{xiaohu.wu, tanshanli2019, taoxf}@bupt.edu.cn, chengh@smail.nju.edu.cn, {han.yu, asysong}@ntu.edu.sg, he_tiantian@cfar.a-star.edu.sg, chjwang@nju.edu.cn

Abstract

Federated learning (FL) provides a privacy-preserving approach for collaborative training of machine learning models. Given the potential data heterogeneity, it is crucial to select appropriate collaborators for each FL participant (FL-PT) based on data complementarity. Recent studies have addressed this challenge. Similarly, it is imperative to consider the inter-individual relationships among FL-PTs where some FL-PTs engage in competition. Although FL literature has acknowledged the significance of this scenario, practical methods for establishing FL ecosystems remain largely unexplored. In this paper, we extend a principle from the balance theory, namely "the friend of my enemy is my enemy", to ensure the absence of conflicting interests within an FL ecosystem. The extended principle and the resulting problem are formulated via graph theory and integer linear programming. A polynomial-time algorithm is proposed to determine the collaborators of each FL-PT. The solution guarantees high scalability, allowing even competing FL-PTs to smoothly join the ecosystem without conflict of interest. The proposed framework jointly considers competition and data heterogeneity. Extensive experiments on real-world and synthetic data demonstrate its efficacy compared to five alternative approaches, and its ability to establish efficient collaboration networks among FL-PTs.

Introduction

Federated Learning (FL) represents a paradigm within distributed machine learning (ML) that facilitates the collaborative training of ML models by leveraging data from multiple parties while upholding privacy considerations (Yang et al. 2019). Each participant in FL (referred to as FL-PT) acts as a custodian of data and directly employs its dataset to locally train a model. In the well-established Federated Averaging (FedAvg) framework (McMahan et al. 2017), a central server (CS) periodically gathers model updates from individual FL-PTs, which are then aggregated to refine a global model. Similarly, each FL-PT regularly acquires the latest global model from the CS and further enhances it through local training. This iterative interplay between the CS and FL-PTs persists until the global model achieves convergence. FL has demonstrated significant promise across diverse domains, including healthcare, digital banking, ridesharing, recommender systems, and drug discovery (Sheller et al. 2020; Long et al. 2020; Yang et al. 2020; Wang et al. 2022; Oldenhof et al. 2023; Sun et al. 2023).

For example, consider a clinical research network of multiple hospitals (Fleurence et al. 2014). These hospitals possess the capacity to collaboratively construct ML models. In an optimal setting, the global model derived from FL should outperform models crafted by individual FL-PTs. However, a potential complication arises from the non-independent and non-identically distributed (Non-IID) nature of data across these FL-PTs (Zhu et al. 2021). Each FL-PT undertakes local model training, which might lead it to a distinct local optima, diverging from the global optima. Consequently, the model performance of an FL-PT might experience degradation due to the FL process (Wang et al. 2019). The diversity in data characteristics among FL-PTs can be graphically portrayed using a directed benefit graph denoted as $\mathcal{G}_b$ (Cui et al. 2022). In this graphical representation, an edge from FL-PT v_i to v_j signifies that the data from v_i can potentially enhance the learning outcomes of v_j through the FL process.

Besides data heterogeneity, another important factor is the relationships among FL-PTs. For instance, in the context of hospitals located in different cities, they serve distinct populations. As depicted in Figure 1, the hospital in city C solely focuses on improving its own ML model, and its utility is independent of any FL-PT in other cities. Such two FL-PTs are considered "independent", where the shared global model in FL functions as a public good, similar to a radio signal where each individual only values the received signal quality (Tang and Wong 2021). In contrast, hospitals within the

*These authors contributed equally.
†Corresponding authors.

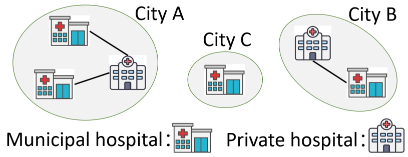


Figure 1: Illustration of hospital relationships: the black line denotes the competing relationship between two hospitals.

same city (e.g., city B) serve the same population, which can include both public and private hospitals. Then, competition arises where the utility of an FL-PT also depends on the model performance of its competitor (Brekke, Siciliani, and Straume 2011). Such FL-PTs are considered "competitive". The inter-individual relationship between any two FL-PTs can be represented by an undirected graph $\mathcal{G}_c$.

In the presence of both data heterogeneity and competition, selecting suitable collaborators for each FL-PT is a crucial challenge. Recently, Cui et al. (2022) consider the data heterogeneity case (i.e., the edge set of $\mathcal{G}_b$ is non-empty and the edge set of $\mathcal{G}_c$ is empty) and leverages the concept of core-stable coalition from cooperative games to effectively address this. All FL-PTs are partitioned into disjoint groups/coalitions. Let $\pi(i)$ denote the coalition to which v_i belongs where π is called a coalition structure, and v_i's utility depends on the FL-PTs in $\pi(i)$. For a core-stable coalition structure π, there is no other coalition $\mathcal{C}$ such that every FL-PT v_i in $\mathcal{C}$ prefers $\mathcal{C}$ over $\pi(i)$ (Aziz and Savani 2016). Nevertheless, there is no existing work addressing the issue of competition among a part of FL-PTs when establishing collaborations in FL ecosystems.

In this paper, we propose the `FedCompetitors` approach to bridge this gap. It is general in the sense that (i) the edge set of $\mathcal{G}_c$ is empty or non-empty except the complete graph case and (ii) the edge set of $\mathcal{G}_b$ is non-empty. The presence of competing FL-PTs has been recognized as an important aspect in the FL literature (Kairouz et al. 2021; Zhan et al. 2022; Shi, Yu, and Leung 2023). In balance theory, a principle, namely "the friend of my enemy is my enemy", can avoid conflict of interest (Leskovec, Huttenlocher, and Kleinberg 2010; Cartwright and Harary 1956). We apply its extended version to establish collaboration among FL-PTs. Specifically, suppose v_i and v_k compete, and v_j is the friend of v_i (i.e., v_i benefits from the data of v_j in FL training). The FL-PT v_i, its friend v_j, and other FL-PTs who benefit v_i and v_j are in an alliance. Then, the CS regulates that v_k will not make a contribution to any FL-PT in the alliance, which ensures that no FL-PTs directly or indirectly assist their competitors. If two FL-PTs can collaborate together, they are independent of each other. In a group of independent FL-PTs, an FL-PT can freely collaborate with other FL-PTs in the group, thereby maximizing the social welfare of the entire FL ecosystem.

The extended principle and the resulting problem above can be formulated via graph theory and integer linear programming. We further propose a polynomial-time algorithm that is to determine the collaborators of each FL-PT. Using the proposed solution, even competing FL-PTs can seamlessly join without conflict of interest and the FL ecosystem thus exhibits a high level of scalability and is trusted by FL-PTs with conflicting interests (Tariq et al. 2023; Yu et al. 2014). Extensive experiments on both synthetic and real-world datasets demonstrate the effectiveness of `FedCompetitors` over the state of the art.

Related Work

We focus on the context of cross-silo FL, where FL-PTs are typically companies or organizations and they both contribute their data and utilize the trained ML models. In the existing research, two scenarios have been extensively investigated: (i) any two FL-PTs in the FL ecosystem are independent of each other and an FL-PT solely focuses on improving its own model performance, and (ii) any two FL-PTs in the FL ecosystem compete against each other where $\mathcal{G}_c$ is a complete graph. In this paper, we mainly consider the scenario where there exists competition among a part of FL-PTs and an FL-PT will not collaborate with its competitors and other FL-PTs with potential conflict of interest.

Firstly, in the independent scenario, prior studies focus on alleviating the side effect of data heterogeneity. While applying Hedonic games that are a type of cooperative games (Aziz and Savani 2016), stable coalition structures are sought to establish collaboration among FL-PTs. Donahue and Kleinberg (2021) provide an analytical understanding of what partition of FL-PTs leads to a stable coalition structure for mean estimation and linear regression. Chaudhury et al. (2022) treat all FL-PTs as a grand coalition and optimizes a common model for all FL-PTs, which is considered core-stable if there is no other coalition $\mathcal{S}$ of FL-PTs that could significantly benefit by training a model with only their data. Another way that learns personalized models for FL-PTs works as follows (Tan et al. 2022): (i) use the CS to train a global model, and (ii) adapt the model to the local data of FL-PTs. Several approaches, such as meta-learning, and multi-task learning, have been employed for personalization (Fallah, Mokhtari, and Ozdaglar 2020; Smith et al. 2017). Ding and Wang (2022) study the case when the FL ecosystem expands to have numerous independent FL-PTs. A group of FL-PTs that has similar contributors is a group of collaboration partners. The authors propose to partition all FL-PTs into K groups and adaptively learn a small number K of models for n FL-PTs, where $1 \ll K \ll n$.

Secondly, in the competition scenario, all FL-PTs offer the same service in a given market. Wu and Yu (2022) aim to achieve the objective of maintaining a negligible change in market share after FL-PTs join the FL ecosystem (Wu, De Pellegrini, and Casale 2023), and analyze the achievability of this objective. Afterwards, two other works study the profitablity of FL-PTs in the given market after FL-PTs join the FL ecosystem, but are taken under different assumptions on the source of extra profit brought by FL. Specifically, Tsoy and Konstantinov (2023) use the following assumption: (i) each consumer has a fixed budget that is allocated to multiple services from different markets, and (ii) if an FL-PT has a higher model quality, its service quality is higher and the consumer will allocate more of its budget to consume the service. Huang, Ke, and Liu (2023) consider duopoly business competition between two FL-PTs and assume that, if the model-related service can be improved by FL, customers will have willingness to pay more and FL-PTs thus have opportunities to increase their profits.

Model and Assumptions

We use graph theory to describe our model of interest and mathematically formulate the extended principle. Specifically, let us consider a set of n FL-PTs denoted by $\mathcal{V} = \{v_1, v_2, \cdots, v_n\}$. Each FL-PT v_i possesses a local dataset $\mathcal{D}_i$. The FL-PTs contemplate joining a collaborative FL network, facilitated by the CS. However, challenges such as data heterogeneity and competition arise among the FL-PTs. To characterize the various relationships among the FL-PTs, three graphs are employed.

Competing graph $\mathcal{G}_c$. An undirected graph $\mathcal{G}_c = (\mathcal{V}, E_c)$ is used to represent the competing relations between any two FL-PTs, where $\mathcal{V}$ is the set of nodes/FL-PTs and E_c is the set of edges. An edge $(v_i, v_j) \in E_c$ signifies a competitive relationship between FL-PTs v_i and v_j. The adjacency matrix of $\mathcal{G}_c$ is denoted as $S_{n\times n}$: its main diagonal elements are set to zero, i.e., $s_{i,i} = 0$; when $i \neq j$, $s_{i,j} = 1$ if v_i competes with v_j, and $s_{i,j} = 0$ if v_i is independent of v_j. Each FL-PT v_i will report its competitors to CS, as it hopes that CS will correctly utilize this information to prevent its competitors from benefiting from its data. Thus, CS has the knowledge of $\mathcal{G}_c$.

Benefit graph $\mathcal{G}_b$. A benefit graph is employed to depict the impact of sample distribution discrepancies among the n FL-PTs. For any two FL-PTs v_i and v_j, if $w_{j,i} = 0$, it indicates that v_i cannot benefit from the data of v_j. Conversely, if $w_{j,i} > 0$, it implies that v_i can benefit from v_j's data, with larger values of $w_{j,i}$ signifying greater benefit to v_i. These values $w_{j,i}$ define a directed graph denoted as $\mathcal{G}_b = (\mathcal{V}, E_b)$, referred to as the benefit graph: $(v_j, v_i) \in E_b$ if and only if $i \neq j$ and $w_{j,i} > 0$. The adjacency matrix of $\mathcal{G}_b$ is denoted as $W_{n\times n}$, where the i-th column comprises the weights $w_{1,i}, w_{2,i}, \cdots, w_{n,i}$, representing the importance of the n FL-PTs to v_i. The level of potential (LoP) of an FL-PT v_i contributing to the other FL-PTs $\mathcal{V} - \{v_i\}$ is defined as

$$w_i = \sum\nolimits_{j\neq i} w_{i,j}, \tag{1}$$

which measures the importance of v_i to the FL ecosystem. The graph $\mathcal{G}_b$ can be obtained by the hypernetwork technique in (Cui et al. 2022; Navon et al. 2021).

Data usage graph $\mathcal{G}_u$. Although v_i may benefit from v_j's data ($w_{j,i} > 0$), CS has the authority to determine whether v_i can actually utilize v_j's local model update information (i.e., indirectly use v_j's data) in the FL training process or not. Let $X = (x_{j,i})$ be a $n \times n$ matrix where

$$x_{j,i} \in \{0, 1\} \tag{2}$$

is a decision variable: for two different FL-PTs v_i and v_j, $x_{j,i}$ is set to one if v_j will contribute to v_i (i.e., v_i will utilize v_j's local model update information) in the FL training process and $x_{j,i}$ is set to zero otherwise. X defines a directed graph $\mathcal{G}_u = (\mathcal{V}, E_u)$, called the data usage graph: $(v_j, v_i) \in E_u$ if and only if $j \neq i$ and $x_{j,i} = 1$; then, v_j is said to be a collaborator or friend of v_i. Consider any pair of FL-PTs v_i and v_j. If v_j's data cannot benefit v_i ($w_{j,i} = 0$), we set $x_{j,i} = 0$. Only when v_j's data can benefit v_i, there is a possibility that $x_{j,i} = 1$. Consequently, E_u is a subset of E_b, leading directly to the following conclusion.

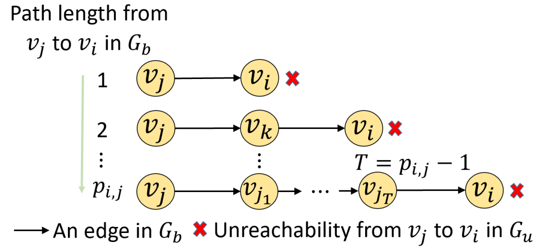


Figure 2: Illustration of Assumption 1: v_j is reachable to v_i in $\mathcal{G}_b$, while v_i and v_j compete against each other.

Lemma 1. *For any two nodes v_j and v_i, if there is no path from v_j to v_i in the benefit graph $\mathcal{G}_b$, then this also holds in the data usage graph $\mathcal{G}_u$.*

Principle for Avoiding Conflict of Interest

Below, we extend the principle that "the friend of my enemy is my enemy".

Assumption 1. *For any two competing FL-PTs v_i and v_j (i.e., $(v_i, v_j) \in E_c$), v_j is unreachable to v_i in the data usage graph $\mathcal{G}_u$.*

Assumption 1 is implemented while establishing the collaboration relationships among FL-PTs. Suppose there is a path from v_j to v_i in the benefit graph $\mathcal{G}_b$ whose length is $p_{i,j}$. We use Figure 2 to explain the implication of Assumption 1. If $p_{i,j} = 1$, it posits that one FL-PT refuses to contribute to its competitor. If $p_{i,j} = 2$, we use v_k to denote the intermediate node. If v_i benefits from v_k, v_k is v_i's friend; v_j is not willing to see the enhancement of v_i's model. Assumption 1 posits that, if $(v_k, v_i) \in E_u$, then $(v_j, v_k) \notin E_u$, i.e., v_j doesn't help the friend v_k of its enemy v_i. Generally, for any $p_{i,j}$, the path from v_j to v_i in $\mathcal{G}_b$ is denoted as

$$P_j^i = (v_{j_0}, v_{j_1}, \cdots, v_{j_{p_{i,j}}}), \tag{3}$$

where $j_0 = j$ and $j_{p_{i,j}} = i$. If any, let t be the minimum integer in $[1, p_{i,j} - 1]$ such that $(v_{j_l}, v_{j_{l+1}}) \in E_u$ for every $l \in [t, p_{i,j} - 1]$ where v_{j_l} helps $v_{j_{l+1}}$. Then, FL-PTs v_{j_t}, $v_{j_{t+1}}, \cdots, v_{j_{p_{i,j}}}$ are said to be in an alliance, and v_j will not help any member in this alliance. Assumption 1 follows a common logic in reality that nobody wants to see others help its enemy and its enemy's friends. By applying Assumption 1, it is strictly guaranteed that each FL-PT will not make a contribution to its competitors directly or indirectly.

For any competing FL-PTs v_i and v_j, let $\mathcal{P}_{j,i}$ denote the set of all reachable paths from v_j to v_i in the graph $\mathcal{G}_b$. Assumption 1 can be characterized by $\mathcal{G}_c$, $\mathcal{G}_b$, and $\mathcal{G}_u$.

Proposition 1. *Assumption 1 holds if and only if the following condition is satisfied:*

$$x_{j,j_1} + x_{j_1,j_2} + \cdots + x_{j_{p_{i,j}},i} \leqslant p_{i,j} - 1, \qquad \forall (v_i, v_j) \in E_c,\ \forall P_j^i \in \mathcal{P}_{j,i}. \tag{4}$$

(a) v_i is reachable to the red node in the oval in the graph $\mathcal{G}_u$, which is the competitor of the blue nodes.

(b) v_j is reachable from the red node in the oval in the graph $\mathcal{G}_u$, which is the competitor of the golden nodes.

Figure 3: Effect on Assumption 1 after adding an edge (v_j, v_i) in the data usage graph $\mathcal{G}_u$.

Proof. Firstly, we prove the reverse direction. By Lemma 1, to satisfy Assumption 1, we only need to focus on such v_j and v_i that are reachable in $\mathcal{G}_b$. If Eq. (4) holds, then, for any $P_j^i \in \mathcal{P}_{j,i}$ in Eq. (3), there exist two adjacent nodes v_{j_l} and $v_{j_{l+1}}$ in P_j^i, where $l \in [0, p_{i,j} - 1]$, such that $x_{j_l,j_{l+1}} = 0$ and $(v_{j_l}, v_{j_{l+1}}) \notin E_u$. Thus, there are no reachable paths from v_j to v_i in $\mathcal{G}_u$ and Assumption 1 is satisfied. Secondly, we prove the forward direction by contradiction. If Eq. (4) doesn't hold, then, for any $l \in [0, p_{i,j} - 1]$, $x_{j_l,j_{l+1}} = 1$ and there is an edge from v_{j_l} to $v_{j_{l+1}}$ in $\mathcal{G}_u$, which contradicts Assumption 1 where v_j is not reachable to v_i in $\mathcal{G}_u$. □

In this paper, we aim to construct an FL ecosystem without conflict of interest. Mathematically, our problem is to determine the matrix $X_{n\times n}$ of decision variables that satisfy Eq. (2) and (4), which determines the collaborators of FL-PTs. The absence of conflicting interests among FL-PTs is guaranteed by Eq. (4).

Polynomial-Time Algorithm

We propose a polynomial-time algorithm to determine the matrix $X_{n\times n}$ subject to Eq. (2) and (4). We begin by describing the algorithm's initial states. The LoP w_i in Eq. (1) measures the importance of v_i to the FL ecosystem. We sort the LoPs of all FL-PTs in non-increasing order, and without loss of generality, we assume:

$$w_1 \geqslant w_2 \geqslant \cdots \geqslant w_n. \tag{5}$$

The initial values of $X_{n\times n}$ are set as follows:

$$x_{j,i} = 1 \text{ if } i = j, \text{ and } x_{j,i} = 0 \text{ if } i \neq j. \tag{6}$$

This defines the initial $\mathcal{G}_u$, which will be updated as the algorithm runs. We also define a connectivity matrix $C_{n\times n}$ of $\mathcal{G}_u$: when $i \neq j$, $c_{j,i} = 1$ if there is a path from v_j to v_i and $c_{j,i} = 0$ otherwise; $c_{i,i}$ is always set to one trivially. Initially, $C_{n\times n}$ is set as an identity matrix.

The proposed algorithm is presented as Algorithm 1. The n FL-PTs are considered sequentially from v_1 to v_n (line 3). At the step for v_i (line 4), the decision variables to be determined are $\{x_{j,i}\}_{j\neq i}$ and we maximize the benefit of v_i:

$$\text{maximize} \sum_{j\neq i} w_{j,i} \cdot x_{j,i} \tag{7}$$

subject to Eq. (2) and (4). Afterwards, $X_{n\times n}$ is updated and the collaborators of v_i are determined. Next, we solve the integer linear programming (ILP) problem (7). Let $\mathcal{B}_i$ denote all FL-PTs that can benefit v_i but are independent of v_i, which can be defined by the adjacency matrix $W_{n\times n}$ of $\mathcal{G}_b$ and the adjacency matrix $S_{n\times n}$ of $\mathcal{G}_c$:

$$\mathcal{B}_i = \{v_j \in \mathcal{V} \,|\, j \neq i, w_{j,i} > 0, s_{j,i} = 0\}. \tag{8}$$

Algorithm 1: Collaborator Selection

Data: $S_{n\times n}$, and $W_{n\times n}$
Result: $X_{n\times n}$
1 Initialize $X_{n\times n}$ by Eq. (6) and $C_{n\times n}$ to be an identity matrix;
2 Generate the sorted sequence (i.e., Eq. (5));
3 **for** *v_i in the sorted sequence* **do**
4 Solve the ILP problem (7) by Algorithm 2;

$\mathcal{B}_i$ includes all possible collaborators of v_i.

For any $v_j \in \mathcal{B}_i$, let $\mathcal{V}_j^-$ denote a set consisting of all nodes that are reachable to v_j in $\mathcal{G}_u$, as well as v_j itself, which can be defined by the connectivity matrix $C_{n\times n}$:

$$\mathcal{V}_j^- = \{v_k \in \mathcal{V} \,|\, c_{k,j} = 1\}. \tag{9}$$

Let $\mathcal{S}_j^-$ denote all competitors of the nodes in $\mathcal{V}_j^-$, and $\mathcal{S}_{i,j}^-$ denote the nodes of $\mathcal{S}_j^-$ that are reachable from v_i in $\mathcal{G}_u$:

$$\mathcal{S}_j^- = \left\{v_k \in \mathcal{V} \,|\, \exists v_p \in \mathcal{V}_j^- : s_{k,p} = 1\right\}, \tag{10}$$

$$\mathcal{S}_{i,j}^- = \left\{v_k \in \mathcal{S}_j^- \,|\, c_{i,k} = 1\right\} \subseteq \mathcal{S}_j^-. \tag{11}$$

As illustrated in Figure 3(a), if $\mathcal{S}_{i,j}^- \neq \emptyset$, we have $x_{j,i} = 0$; otherwise, some nodes in $\mathcal{V}_j^-$ will be reachable to its competitor (e.g., the node in the oval) in $\mathcal{G}_u$, which violates Eq. (4). Let $\mathcal{V}_i^+$ denote a set consisting of all nodes that are reachable from v_i in $\mathcal{G}_u$, as well as v_i itself:

$$\mathcal{V}_i^+ = \{v_k \in \mathcal{V} \,|\, c_{i,k} = 1\}. \tag{12}$$

Let $\mathcal{S}_i^+$ denote all competitors of the nodes in $\mathcal{V}_i^+$, and $\mathcal{S}_{i,j}^+$ denote the nodes of $\mathcal{S}_i^+$ that are reachable to v_j in $\mathcal{G}_u$:

$$\mathcal{S}_i^+ = \{v_k \in \mathcal{V} \,|\, \exists v_p \in \mathcal{V}_i^+ : s_{p,k} = 1\}, \tag{13}$$

$$\mathcal{S}_{i,j}^+ = \{v_k \in \mathcal{S}_i^+ \,|\, c_{k,j} = 1\} \subseteq \mathcal{S}_i^+. \tag{14}$$

Here, by Eq. (9), (11), (12), and (14), we have

$$\mathcal{S}_{i,j}^- = \mathcal{V}_i^+ \cap \mathcal{S}_j^- \subseteq \mathcal{V}_i^+ \text{ and } \mathcal{S}_{i,j}^+ = \mathcal{V}_j^- \cap \mathcal{S}_i^+ \subseteq \mathcal{V}_j^-. \tag{15}$$

As illustrated in Figure 3(b), if $\mathcal{S}_{i,j}^+ \neq \emptyset$, then $x_{j,i} = 0$; otherwise, some nodes in $\mathcal{V}_i^+$ will be reachable from its competitor (e.g., the node in the oval) in $\mathcal{G}_u$, violating Eq. (4).

Based on the above understanding, we propose Algorithm 2 to solve the problem (7). For a node $v_j \in \mathcal{B}_i$, $w_{j,i}$ represents the importance of v_j to v_i. We sort the nodes of $\mathcal{B}_i$ in the non-increasing order of their values $w_{j,i}$ (line 1). The nodes of $\mathcal{B}_i$ are considered sequentially (line 2). For each

Algorithm 2: ILP Solver

```
  Data: W_{n×n}, S_{n×n}, and C_{n×n}
  Result: the updated X_{n×n}, and C_{n×n}
1 Sort the nodes of B_i in non-increasing order of their
    values w_{j,i}, generating a sorted sequence;
2 for v_j in the sorted sequence do
3   if S^+_{i,j} = ∅ ∧ S^-_{i,j} = ∅ then
4     x_{j,i} ← 1, c_{j,i} ← 1;
5     for any two integers p ∈ [1,n] and q ∈ [1,n]
        with p ≠ q and (p,q) ≠ (j,i) do
6       if c_{p,q} = 0 ∧ c_{p,j} = 1 ∧ c_{i,q} = 1 then
7         c_{p,q} ← 1;
```

node $v_j \in \mathcal{B}_i$, if $\mathcal{S}^+_{i,j} = \emptyset$ and $\mathcal{S}^-_{i,j} = \emptyset$, the algorithm sets v_j as the collaborator of v_i (i.e., $x_{j,i} = 1$), with the connectivity from v_j to v_i is updated (lines 3-4). Finally, we consider the effect of setting $x_{j,i} = 1$ on the connectivity between any two nodes v_p and v_q in the graph $\mathcal{G}_u$, except (v_j, v_i) (line 5). In the graph $\mathcal{G}_u$, if we have before executing line 4 that v_p is not reachable to v_q, v_p is reachable to v_j, and v_i is reachable to v_q, then v_p becomes reachable to v_q (lines 6-7).

Lemma 2. *Given $W_{n\times n}$, $S_{n\times n}$ and $C_{n\times n}$, the time complexity of finding $\mathcal{B}_i$ is $\mathcal{O}(n)$ while the time complexity of finding $\mathcal{S}^-_{i,j}$ or $\mathcal{S}^+_{i,j}$ is $\mathcal{O}(n^2)$.*

Proof. By Eq. (8), the (time) complexity of finding $\mathcal{B}_i$ is $\mathcal{O}(n)$ where $|\mathcal{B}_i| \leqslant n$. By Eq. (9), the complexity of finding $\mathcal{V}^-_j$ is $\mathcal{O}(n)$ where $|\mathcal{V}^-_j| \leqslant n$. By Eq. (10), $\mathcal{S}^-_j$ can be found by (i) checking every $v_k \in \mathcal{V}$ and (ii) judging whether there exists a node $v_p \in \mathcal{V}^-_j$ such that $s_{k,p} = 1$; the resulting complexity is $\mathcal{O}(n^2)$; here, $|\mathcal{S}^-_j| \leqslant n$. Given $\mathcal{S}^-_j$, by Eq. (11), the complexity of finding $\mathcal{S}^-_{i,j}$ is $\mathcal{O}(n)$. Finally, the complexity of finding $\mathcal{S}^-_{i,j}$ is $\mathcal{O}(n^2)$. Similarly to $\mathcal{S}^-_{i,j}$, the complexity of finding $\mathcal{S}^+_{i,j}$ is also $\mathcal{O}(n^2)$. □

Proposition 2. *Suppose $X_{n\times n}$ satisfies Eq. (2) and (4) before v_i is considered. Algorithm 2 gives a feasible solution to the ILP problem* (7) *with a time complexity $\mathcal{O}(n^3)$ when v_i is considered.*

Proof. By Proposition 1, Eq. (4) is equivalent to Assumption 1. Firstly, we prove by contradiction that Algorithm 2 gives a feasible solution. Before v_i is considered, no two competitors in $\mathcal{V}$ are reachable in $\mathcal{G}_u$ by Assumption 1. Setting $x_{j,i} = 1$ is equivalent to adding an edge (v_j, v_i) in $\mathcal{G}_u$. By the definition of $\mathcal{V}^-_j$ and $\mathcal{V}^+_i$, the addition of (v_j, v_i) can only affect the reachability from the nodes of $\mathcal{V}^-_j$ to the nodes of $\mathcal{V}^+_i$ in $\mathcal{G}_u$. Suppose there exists a node $v_j \in \mathcal{B}_i$ satisfying $\mathcal{S}^+_{i,j} = \emptyset$ and $\mathcal{S}^-_{i,j} = \emptyset$, such that, Assumption 1 is violated after setting $x_{j,i} = 1$. Thus, the addition of (v_j, v_i) leads to that some node of $\mathcal{V}^-_j$ is reachable to and competes with some node of $\mathcal{V}^+_i$ in $\mathcal{G}_u$. Then, there exists a node v_k such that either $v_k \in \mathcal{V}^-_j$ and v_k is a competitor of some node in $\mathcal{V}^+_i$ (i.e., $v_k \in \mathcal{S}^+_{i,j}$ by Eq. (13) and (15)), or $v_k \in \mathcal{V}^+_j$ and v_k is a competitor of the nodes of $\mathcal{V}^-_i$ (i.e., $v_k \in \mathcal{S}^-_{i,j}$ by Eq. (10) and (15)). $\mathcal{S}^-_{i,j}$ and $\mathcal{S}^+_{i,j}$ are non-empty, which contradicts the condition in line 3 that leads to $x_{j,i} = 1$. Secondly, we show the complexity of Algorithm 2. Given $\mathcal{B}_i$, the time complexity of sorting the nodes of $\mathcal{B}_i$ is $\mathcal{O}(n \log n)$, e.g., using the mergesort algorithm. Thus, by Lemma 2, the time complexity in line 1 is $\mathcal{O}(n \log n)$. For the for-loop in line 2, its time complexity is $\mathcal{O}(n)$ where $|\mathcal{B}_i| \leqslant n$; by Lemma 2, the time complexity in line 3 is $\mathcal{O}(n^2)$. For the for-loop in line 5, the time complexity is $\mathcal{O}(n^2)$. The total time complexity in lines 2–7 is $\mathcal{O}(n^3)$. Finally, Algorithm 2 has a time complexity $\mathcal{O}(n^3)$. □

We show the correctness of Algorithm 1. At the beginning of Algorithm 1, $X_{n\times n}$ satisfies Eq. (2) and (4) by Eq. (6). After each step for v_i in line 4, $X_{n\times n}$ still satisfies these constraints by Proposition 2. When Algorithm 1 ends, the final collaborating relationships among FL-PTs are determined by $X_{n\times n}$. By Eq. (1), the (time) complexity of computing w_i for each FL-PT v_i is $\mathcal{O}(n)$; thus, the complexity of computing $w_1, w_2, \cdots, w_n$ is $\mathcal{O}(n^2)$. The complexity of sorting $w_1, w_2, \cdots, w_n$ is $\mathcal{O}(n \log n)$. Thus, the complexity in line 2 of Algorithm 1 is $\mathcal{O}(n^2)$. By Proposition 2, the complexity in lines 3-4 is $\mathcal{O}(n^4)$. Thus, the complexity of Algorithm 1 is $\mathcal{O}(n^4)$. We note that, the overall optimization objective of this paper can be social welfare maximization (e.g., $\max \sum_{i=1}^{n} \sum_{j=1}^{n} w_{j,i} x_{j,i}$), subject to Eq. (2) and (4). This is a 0-1 multidimensional knapsack problem with binary weights where greedy heuristics are typically used to find a solution (Gorski, Paquete, and Pedrosa 2012; Akçay, Li, and Xu 2007; Wu et al. 2021).

Experimental Evaluation

We conduct experiments on synthetic data and the CIFAR-10 dataset. To investigate the practicality of `FedCompetitors`, we also adopt the electronic health record (EHR) dataset eICU (Pollard et al. 2018) to illustrate the collaboration relationships of FL-PTs on a real-world network of multiple hospitals.

Comparison Baselines

Compared with the proposed approach in the last section, we now give a more intuitive procedure to address the competing relationships among FL-PTs. This procedure makes the previous FL approaches (e.g., FedAvg) applicable to the scenario of this paper. At a high level, we will find a partition of all FL-PTs into several disjoint groups such that the FL-PTs in each group are independent of each other, without conflict of interest. Then, baselines can be generated by directly applying the previous FL approaches to each group of FL-PTs. Specifically, the competing graph $\mathcal{G}_c$ describes the competing relationship among FL-PTs. Let $\mathcal{G}^-_c$ denote the complement of $\mathcal{G}_c$: the nonexistence of an edge between v_i and v_j in $\mathcal{G}_c$ leads to the existence of an edge (v_i, v_j) in $\mathcal{G}^-_c$, and vice versa. Each edge in the graph $\mathcal{G}^-_c$ indicates that the two FL-PTs connected by this edge are independent. A clique is a subset of nodes of $\mathcal{G}^-_c$ such that every two nodes

Weakly Non-IID setting (MSE)								
	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
Local	0.23±0.08	0.23±0.09	0.87±0.41	0.82±0.26	0.23±0.10	0.23±0.07	0.82±0.24	0.78±0.30
FedAvg	0.20±0.06	0.20±0.06	0.20±0.10	0.19±0.07	0.19±0.06	0.19±0.06	0.19±0.08	0.19±0.10
FedProx	0.16±0.06	0.17±0.07	0.15±0.09	0.17±0.08	0.17±0.06	0.17±0.06	0.16±0.09	0.18±0.07
SCAFFOLD	0.17±0.07	0.17±0.07	0.16±0.09	0.16±0.07	0.18±0.06	0.18±0.07	0.18±0.08	0.18±0.08
CE	**0.14±0.10**	0.14±0.11	1.14±0.67	1.20±0.88	**0.15±0.08**	0.16±0.09	1.23±0.37	1.22±0.81
`FedCompetitors`	0.14±0.12	**0.14±0.07**	**0.13±0.06**	**0.15±0.06**	**0.15±0.08**	**0.14±0.06**	**0.14±0.07**	**0.14±0.07**

Strongly Non-IID Setting (MSE)								
	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
Local	0.23±0.08	0.23±0.08	0.22±0.07	0.23±0.08	0.23±0.06	0.22±0.06	0.22±0.08	0.23±0.07
FedAvg	24.47±4.98	24.85±4.82	24.85±5.03	24.73±5.67	24.15±3.00	24.47±2.78	24.17±4.40	24.97±3.81
FedProx	17.80±7.54	17.82±6.42	17.88±7.68	17.86±7.64	17.69±7.14	17.76±6.23	17.68±5.94	17.73±7.04
SCAFFOLD	17.22±2.85	17.44±2.17	17.39±4.02	17.20±3.58	16.87±2.75	17.13±2.79	17.00±2.41	17.33±2.59
CE	0.15±0.12	0.14±0.11	0.14±0.07	**0.14±0.07**	0.14±0.06	**0.14±0.06**	0.12±0.05	**0.12±0.05**
`FedCompetitors`	**0.14±0.07**	**0.13±0.06**	**0.13±0.06**	0.14±0.09	**0.13±0.07**	**0.14±0.06**	**0.11±0.04**	0.13±0.07

Table 1: Experiments with synthetic data under fixed competing graphs

in the clique are adjacent, that is, a clique is a subgraph that is complete. A clique cover of $\mathcal{G}_c^-$ is a partition of all nodes into cliques within which every two nodes in the clique are adjacent and independent of each other (Tomita, Tanaka, and Takahashi 2006). A minimum clique cover is a clique cover that uses as few cliques as possible.

The FL-PTs in each clique are grouped together to take FL training, without involving the FL-PTs from other cliques. We apply four typical FL approaches directly to the nodes of each clique for FL training: **FedAvg**, **CE**, **FedProx** (Li et al. 2020) and **SCAFFOLD** (Karimireddy et al. 2020), which generates four baselines. The collaboration equilibrium (CE) approach is proposed in (Cui et al. 2022) where each coalition is defined as a strongly connected component of $\mathcal{G}_b$; its effectiveness has been validated against several other approaches. FedProx and SCAFFOLD represent two typical approaches that make the aggregated model at the CS close to the global optima and are two benchmarks in (Li et al. 2022) for showing the FL performance under Non-IID data settings. The fifth baseline is **Local** where each FL-PT takes local ML training without collaboration.

General experimental setting. Like (Cui et al. 2022), the hypernetwork technique in (Navon et al. 2021) is used to compute the benefit graph $\mathcal{G}_b$ and a hypernetwork is constructed by a multilayer perceptron (MLP). When it comes to a specific dataset, all approaches have the same network structure for each FL-PT to execute the learning tasks.

Synthetic Experiments

We show the experimental results on synthetic data with fixed competing graphs. Specifically, let us consider 8 FL-PTs $\{v_1, v_2, \cdots, v_8\}$. The synthetic features are generated by $x \sim \mathcal{U}[-1.0, 1.0]$. Given the FL-PT v_i, the grand truth weights $u_{i,l} = v_l + r_{i,l}$ are sampled as $v \sim \mathcal{U}[0.0, 1.0]$ and $r_{i,l} \sim \mathcal{N}(0.0, \rho^2)$ where $l \in \{1, 2, 3\}$; the noise $\epsilon \sim \mathcal{N}(0.0, 0.1^2)$ is added to each label.

Weakly Non-IID setting. ρ^2 measures the data distribution discrepancy among FL-PTs. We set $\rho = 0.01$, which means that the generated data are weakly non-iid in terms of sample features and labels. The same type of polynomial regression tasks is learned by all FL-PTs and the synthetic labels are defined as: $y = \sum_{l=1}^{3} u_{i,l}^T x^l + \epsilon$. The network used for predicting the label at each FL-PT is an MLP with one hidden layer. FL-PTs v_1, v_2, v_5 and v_6 have 2000 samples, while the other FL-PTs have 100 samples. Thus, there exists quantity skew, i.e., a significant difference in the sample quantities of FL-PTs. Two large FL-PTs v_1 and v_2 are independent and compete with the other two large FL-PTs v_5 and v_6 that are independent. Each small FL-PT competes one large FL-PT: (v_1, v_7), (v_2, v_8), (v_3, v_5), and (v_4, v_6) are edges in the competing graph $\mathcal{G}_c$. Such $\mathcal{G}_c$ leads to a unique clique cover. Under this setting, the minimum clique cover of $\mathcal{G}_c^-$ is $\{v_i\}_{i=1}^4$ and $\{v_i\}_{i=5}^8$, and small FL-PTs benefit large FL-PTs little. The experimental results (measured by mean squared error (MSE)) are given in Table 1. On average, CE has the worst performance since small FL-PTs v_3, v_4, v_7 and v_8 cannot benefit from large FL-PTs. Particularly, `FedCompetitors` has the best performance compared with the five baselines.

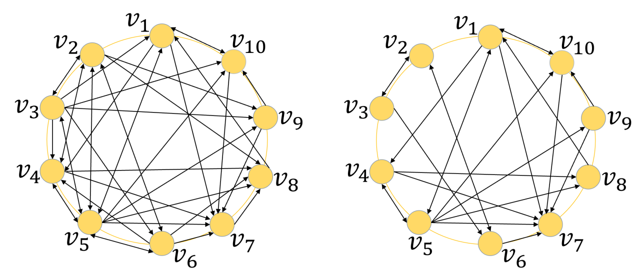
Strongly Non-IID setting. This setting is the same as the setting above expect three aspects. Firstly, each FL-PT has 2000 samples and there is no quantity skew. Secondly, we generate conflicting learning tasks by flipping over the labels of some FL-PTs: $y = -\sum_{l=1}^{3} u_{i,l}^T x^l + \epsilon$ for $i \in \{5, 6, 7, 8\}$, which leads to strongly Non-IID among the eight FL-PTs in terms of the labels. Thirdly, we test on a different competing graph where there are two independent groups of FL-PTs $\{v_i\}_{i=1}^4$ and $\{v_i\}_{i=5}^8$: for $i \in \{1, 5\}$, the FL-PTs v_i and v_{i+1} are independent of each other and compete with v_{i+2} and v_{i+3} that are also independent of each other. Under this setting, all FL-PTs in the same group can benefit each other; the minimum clique cover of $\mathcal{G}_c^-$ is $\{v_1, v_2, v_5, v_6\}$ and $\{v_3, v_4, v_7, v_8\}$. The experimental results are given in Table 1. FedAvg, FedProx, and SCAFFOLD perform the worst since training a global model cannot simultaneously satisfy the FL-PTs in the same clique with conflicting learning tasks. It is observed that `FedCompetitors` has the best performance compared with the five baselines.

	AUC									
	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}
Local	76.12	69.46	68.94	68.04	76.46	40.00	69.30	60.53	56.94	49.12
FedAvg	75.26	72.09	68.87	74.13	83.72	41.67	79.37	54.41	66.67	38.10
CE	**83.53**	75.64	**74.38**	74.46	80.89	82.61	71.43	66.67	66.67	80.00
FedCompetitors	81.50	**78.23**	69.18	**83.52**	**85.91**	**89.58**	**80.70**	**68.89**	**90.48**	**95.24**

Table 2: Experiments with eICU under a fixed competing graph

	MTA
Local	86.46 ± 4.12
FedAvg	52.99 ± 4.38
FedProx	51.13 ± 7.10
SCAFFOLD	51.20 ± 7.09
CE	87.80 ± 7.18
FedCompetitors	$\mathbf{91.33 \pm 4.14}$

Table 3: Experiments with CIFAR-10 under randomly generated competing graphs



(a) $\mathcal{G}_b$. (b) $\mathcal{G}_b$.

Figure 4: Illustration of hospital collaboration.

Benchmark Experiments

We conduct experiments on CIFAR-10 with competing graphs that are generated randomly. CIFAR-10 is an image classification dataset and has 10 classes, each with 6000 images. We follow the setting in (Cui et al. 2022) for CIFAR-10 to construct Non-IID data and network structures, and to measure performance. There are 10 FL-PTs, and each FL-PT randomly obtains 2 of the 10 classes to simulate the Non-IID setting. The model performance is measured by the mean test accuracy (MTA). To simulate competition, we set the probability of two FL-PTs competing against each other to 0.2, thus generating a random competing graph $\mathcal{G}_c$, which constrains the collaboration between some FL-PTs. Table 3 shows the experimental results. It is observed that FedCompetitors has the best performance. FedAvg, FedProx, and SCAFFOLD perform worst since training a global model cannot simultaneously satisfy the FL-PTs in the same clique with data heterogeneity. FedCompetitors performs better than CE by 3.53%.

Hospital Collaboration Example

eICU is a dataset collecting EHRs from many hospitals across the United States admitted to the intensive care unit (ICU). The task is to predict mortality during hospitalization. We use this dataset to illustrate a benefit graph $\mathcal{G}_b$ and a data usage graph $\mathcal{G}_u$ in the real world. The setting here is the same as the setting in (Cui et al. 2022) for eICU, including the data pre-processing procedure, the way of choosing hospitals, the network structures, and the performance metric. There are 10 hospitals: each of the first 5 hospitals $\{v_i\}_{i=1}^5$ has about 1000 patients and each of the remaining hospitals $\{v_i\}_{i=6}^{10}$ has about 100 patients. Label imbalance occurs since more than 90% samples have negative labels; thus, AUC is used to measure the utility of each FL-PT. The generated benefit graph $\mathcal{G}_b$ is illustrated in Figure 4(a).

Let us consider the case where more than one large hospital may be located in the same city while small hospitals are dispersed in rural areas with lower population densities; competition mainly occurs among large hospitals. We assume that v_2 competes with v_5, while v_3 competes with v_4 and v_5, respectively. For the baselines except the local approach, the way of generating the clique cover is independent of $\mathcal{G}_b$ where FL-PTs in each clique collaborate together; the generated clique cover is $\{v_4, v_5\}$ and $\{v_i\}_{i=1}^3 \cup \{v_i\}_{i=6}^{10}$. For FedCompetitors, the generated data usage graph $\mathcal{G}_u$ is illustrated in Figure 4(b), which fully utilizes the information on $\mathcal{G}_b$ by Algorithm 1. Compared with the baselines, it is observed from Figure 4(b) that the local model update information of v_4 and v_5 can also be utilized by other FL-PTs $\{v_1, v_7, v_8, v_9, v_{10}\}$ while v_4 and v_5 can similarly benefit from v_1 in the FL training process. This is an advantage of FedCompetitors and is reflected in the experimental results, which are given in Table 2. Overall, FedCompetitors achieves the best performance.

Conclusions

We consider in this paper an open research problem in which a subset of FL-PTs in the FL ecosystem engage in competition. We extend a principle from balance theory that "the friend of my enemy is my enemy" to guarantee that no conflict of interest occurs among FL-PTs. The resulting FL ecosystem thus exhibits a high level of scalability since FL-PTs that even compete can join smoothly. We formulate the problem and show that it is mathematically solvable in polynomial time. Thus, an efficient algorithm is proposed to determine the collaboration relationships of FL-PTs. The framework of this paper is also general since it considers both competition and data heterogeneity, which is another important aspect in FL. Extensive experiments demonstrate the effectiveness of the proposed framework.

Acknowledgments

This research was supported in part by the National Key R&D Program of China (No. 2022YFB2902900). This research/project is also supported, in part, by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-RP-2020-019); the RIE 2020 Advanced Manufacturing and Engineering (AME) Programmatic Fund (No. A20G8b0102), Singapore; the A*STAR RIE2025 Manufacturing, Trade and Connectivity (MTC) Industry Alignment Fund- Pre-Positioning (IAF-PP) (No. M23L4a0001), Singapore; and the Center for Frontier AI Research (CFAR), Agency for Science, Technology and Research (A*STAR), Singapore. The work of Hao Cheng and Chongjun Wang was supported by the National Natural Science Foundation of China (Grant No. 62192783, 62376117). The work of Shanli Tan was done when he was a research intern with Xiaohu Wu at the National Engineering Research Center of Mobile Network Technologies, Beijing University of Posts and Telecommunications, China.

References

Akçay, Y.; Li, H.; and Xu, S. H. 2007. Greedy algorithm for the general multidimensional knapsack problem. *Annals of operations research*, 150: 17–29.

Aziz, H.; and Savani, R. 2016. Hedonic Games. In Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D., eds., *Handbook of Computational Social Choice*, 356–376. Cambridge University Press.

Brekke, K. R.; Siciliani, L.; and Straume, O. R. 2011. Hospital competition and quality with regulated prices. *Scandinavian Journal of Economics*, 113(2): 444–469.

Cartwright, D.; and Harary, F. 1956. Structural balance: A generalization of Heider's theory. *Psychological Review*, 63(5): 277.

Chaudhury, B. R.; Li, L.; Kang, M.; Li, B.; and Mehta, R. 2022. Fairness in federated learning via core-stability. In *Advances in Neural Information Processing Systems (*NeurIPS'22*)*, volume 35, 5738–5750.

Cui, S.; Liang, J.; Pan, W.; Chen, K.; Zhang, C.; and Wang, F. 2022. Collaboration equilibrium in federated learning. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (*KDD'22*)*, 241–251.

Ding, S.; and Wang, W. 2022. Collaborative learning by detecting collaboration partners. In *Advances in Neural Information Processing Systems (*NeurIPS'22*)*, volume 35, 15629–15641.

Donahue, K.; and Kleinberg, J. 2021. Model-sharing games: Analyzing federated learning under voluntary participation. *Proceedings of the 35th AAAI Conference on Artificial Intelligence (*AAAI'21*)*, 35(6): 5303–5311.

Fallah, A.; Mokhtari, A.; and Ozdaglar, A. 2020. Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach. In *Advances in Neural Information Processing Systems (*NeurIPS'20*)*, volume 33, 3557–3568.

Fleurence, R. L.; Curtis, L. H.; Califf, R. M.; Platt, R.; Selby, J. V.; and Brown, J. S. 2014. Launching PCORnet, a national patient-centered clinical research network. *Journal of the American Medical Informatics Association*, 21(4): 578–582.

Gorski, J.; Paquete, L.; and Pedrosa, F. 2012. Greedy algorithms for a class of knapsack problems with binary weights. *Computers & Operations Research*, 39(3): 498–511.

Huang, C.; Ke, S.; and Liu, X. 2023. Duopoly business competition in cross-silo federated learning. *IEEE Transactions on Network Science and Engineering*, 1–13.

Kairouz, P.; McMahan, H. B.; Avent, B.; Bellet, A.; Bennis, M.; Nitin Bhagoji, A.; Bonawitz, K.; Charles, Z.; Cormode, G.; Cummings, R.; D'Oliveira, R. G. L.; Eichner, H.; El Rouayheb, S.; Evans, D.; Gardner, J.; Garrett, Z.; Gascón, A.; Ghazi, B.; Gibbons, P. B.; Gruteser, M.; Harchaoui, Z.; He, C.; He, L.; Huo, Z.; Hutchinson, B.; Hsu, J.; Jaggi, M.; Javidi, T.; Joshi, G.; Khodak, M.; Konecný, J.; Korolova, A.; Koushanfar, F.; Koyejo, S.; Lepoint, T.; Liu, Y.; Mittal, P.; Mohri, M.; Nock, R.; Özgür, A.; Pagh, R.; Qi, H.; Ramage, D.; Raskar, R.; Raykova, M.; Song, D.; Song, W.; Stich, S. U.; Sun, Z.; Suresh, A. T.; Tramèr, F.; Vepakomma, P.; Wang, J.; Xiong, L.; Xu, Z.; Yang, Q.; Yu, F. X.; Yu, H.; and Zhao, S. 2021. Advances and Open Problems in Federated Learning. *Foundations and Trends in Machine Learning*, 14(1–2): 1–210.

Karimireddy, S. P.; Kale, S.; Mohri, M.; Reddi, S.; Stich, S.; and Suresh, A. T. 2020. Scaffold: Stochastic controlled averaging for federated learning. In *Proceedings of the 37th International Conference on Machine Learning (*ICML'20*)*, volume 119, 5132–5143.

Leskovec, J.; Huttenlocher, D.; and Kleinberg, J. 2010. Predicting positive and negative links in online social networks. In *Proceedings of the 19th International Conference on World Wide Web (*WWW'10*)*, 641–650.

Li, Q.; Diao, Y.; Chen, Q.; and He, B. 2022. Federated learning on non-iid data silos: An experimental study. In *Proceedings of the IEEE 38th International Conference on Data Engineering (*ICDE'22*)*, 965–978.

Li, T.; Sahu, A. K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; and Smith, V. 2020. Federated optimization in heterogeneous networks. In *Proceedings of Machine Learning and Systems*, volume 2, 429–450.

Long, G.; Tan, Y.; Jiang, J.; and Zhang, C. 2020. Federated learning for open banking. In *Federated Learning*, 240–254. Springer.

McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; and Arcas, B. A. 2017. Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (*AISTATS'17*)*, 1273–1282.

Navon, A.; Shamsian, A.; Fetaya, E.; and Chechik, G. 2021. Learning the Pareto Front with Hypernetworks. In *International Conference on Learning Representations (*ICLR'21*)*.

Oldenhof, M.; Ács, G.; Pejó, B.; Schuffenhauer, A.; Holway, N.; Sturm, N.; Dieckmann, A.; Fortmeier, O.; Boniface, E.; Mayer, C.; et al. 2023. Industry-scale orchestrated federated learning for drug discovery. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 15576–15584.

Pollard, T. J.; Johnson, A. E.; Raffa, J. D.; Celi, L. A.; Mark, R. G.; and Badawi, O. 2018. The eICU collaborative research database, a freely available multi-center database for critical care research. *Scientific data*, 5(1): 1–13.

Sheller, M. J.; Edwards, B.; Reina, G. A.; Martin, J.; Pati, S.; Kotrotsou, A.; Milchenko, M.; Xu, W.; Marcus, D.; Colen, R. R.; et al. 2020. Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, 10(1): 1–12.

Shi, Y.; Yu, H.; and Leung, C. 2023. Towards fairness-aware federated learning. *IEEE Transactions on Neural Networks and Learning Systems*, 1–17.

Smith, V.; Chiang, C.-K.; Sanjabi, M.; and Talwalkar, A. S. 2017. Federated multi-task learning. In *Advances in neural information processing systems (*NIPS'17*)*, volume 30.

Sun, C.; Huang, C.; Shou, B.; and Huang, J. 2023. Federated Learning in Competitive EV Charging Market. *arXiv preprint arXiv:2310.08794*.

Tan, A. Z.; Yu, H.; Cui, L.; and Yang, Q. 2022. Towards personalized federated learning. *IEEE Transactions on Neural Networks and Learning Systems*, 1–17.

Tang, M.; and Wong, V. W. 2021. An incentive mechanism for cross-silo federated learning: A public goods perspective. In *Proceedings of the 2022 IEEE Conference on Computer Communications (*INFOCOM'22*)*, 1–10. IEEE.

Tariq, A.; Serhani, M. A.; Sallabi, F.; Qayyum, T.; Barka, E. S.; and Shuaib, K. A. 2023. Trustworthy Federated Learning: A Survey. *arXiv preprint arXiv:2305.11537*.

Tomita, E.; Tanaka, A.; and Takahashi, H. 2006. The worst-case time complexity for generating all maximal cliques and computational experiments. *Theoretical computer science*, 363(1): 28–42.

Tsoy, N.; and Konstantinov, N. 2023. Strategic data sharing between competitors. *Thirty-seventh Conference on Neural Information Processing Systems*.

Wang, Y.; Tong, Y.; Zhou, Z.; Ren, Z.; Xu, Y.; Wu, G.; and Lv, W. 2022. Fed-LTD: Towards Cross-Platform Ride Hailing via Federated Learning to Dispatch. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (*KDD'22*)*, 4079–4089.

Wang, Z.; Dai, Z.; Póczos, B.; and Carbonell, J. 2019. Characterizing and avoiding negative transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (*CVPR'19*)*, 11293–11302.

Wu, X.; De Pellegrini, F.; and Casale, G. 2023. Delay and price differentiation in cloud computing: A service model, supporting architectures, and performance. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, 8(3): 1–40.

Wu, X.; Liu, Y.; Tang, X.; Cai, W.; Bai, F.; Khonstantine, G.; and Zhao, G. 2021. Multi-Agent Pickup and Delivery with Task Deadlines. In *Proceedings of the International Symposium on Combinatorial Search*, volume 12, 206–208.

Wu, X.; and Yu, H. 2022. MarS-FL: Enabling competitors to collaborate in federated learning. *IEEE Transactions on Big Data*, 1–11.

Yang, L.; Tan, B.; Zheng, V. W.; Chen, K.; and Yang, Q. 2020. Federated recommendation systems. In *Federated Learning: Privacy and Incentive*, 225–239. Springer.

Yang, Q.; Liu, Y.; Chen, T.; and Tong, Y. 2019. Federated machinelearning: concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2): 12:1–12:19.

Yu, H.; Miao, C.; An, B.; Shen, Z.; and Leung, C. 2014. Reputation-aware task allocation for human trustees. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, 357–364.

Zhan, Y.; Zhang, J.; Hong, Z.; Wu, L.; Li, P.; and Guo, S. 2022. A survey of incentive mechanism design for federated learning. *IEEE Transactions on Emerging Topics in Computing*, 10(2): 1035–1044.

Zhu, H.; Xu, J.; Liu, S.; and Jin, Y. 2021. Federated learning on non-IID data: A survey. *Neurocomputing*, 465: 371–390.