MarS-FL: Enabling Competitors to Collaborate in Federated Learning

考虑FL中的竞争关系，并设计对应的模型，如配合率。

Xiaohu Wu and Han Yu

Abstract—Federated learning (FL) is rapidly gaining popularity and enables multiple data owners (*a.k.a.* FL participants) to collaboratively train machine learning models in a privacy-preserving way. A key unaddressed scenario is that these FL participants are in a competitive market, where market shares represent their competitiveness. Although they are interested to enhance the performance of their respective models through FL, market leaders (who are often data owners who can contribute significantly to building high performance FL models) want to avoid losing their market shares by enhancing their competitors' models. Currently, there is no modeling tool to analyze such scenarios and support informed decision-making. In this paper, we bridge this gap by proposing the market share-based decision support framework for participation in FL (MarS-FL). We introduce *two notions of δ-stable market* and *friendliness* to measure the viability of FL and the market acceptability of FL. The FL participants' behaviours can then be predicted using game theoretic tools (i.e., their optimal strategies concerning participation in FL). If the market δ-stability is achievable, the final model performance improvement of each FL-PT shall be bounded, which relates to the market conditions of FL applications. We provide tight bounds and quantify the friendliness, κ, of given market conditions to FL. Experimental results show the viability of FL in a wide range of market conditions. Our results are useful for identifying the market conditions under which collaborative FL model training is viable among competitors, and the requirements that have to be imposed while applying FL under these conditions.

Index Terms—Federated learning, competitive market, game theory, performance allocation

1 Introduction

Federated learning (FL) is an emerging privacy-preserving collaborative machine learning (ML) paradigm and has gained widespread attention [1]. An individual data owner (*a.k.a.* FL participants (FL-PTs)) often has insufficient data to train its learning model. Multiple FL-PTs can collaboratively build better ML models with their local data while no local data of an individual are exposed and transferred outside, thereby preserving data privacy by design [2], [3]. For example, in a centralized FL architecture, the FL-PTs periodically upload their respective local ML model updates to a central server that will aggregate these updates for collaboratively training a global ML model, instead of uploading local data. This process is illustrated in Figure 1(a) and iterated until the global ML model converges.

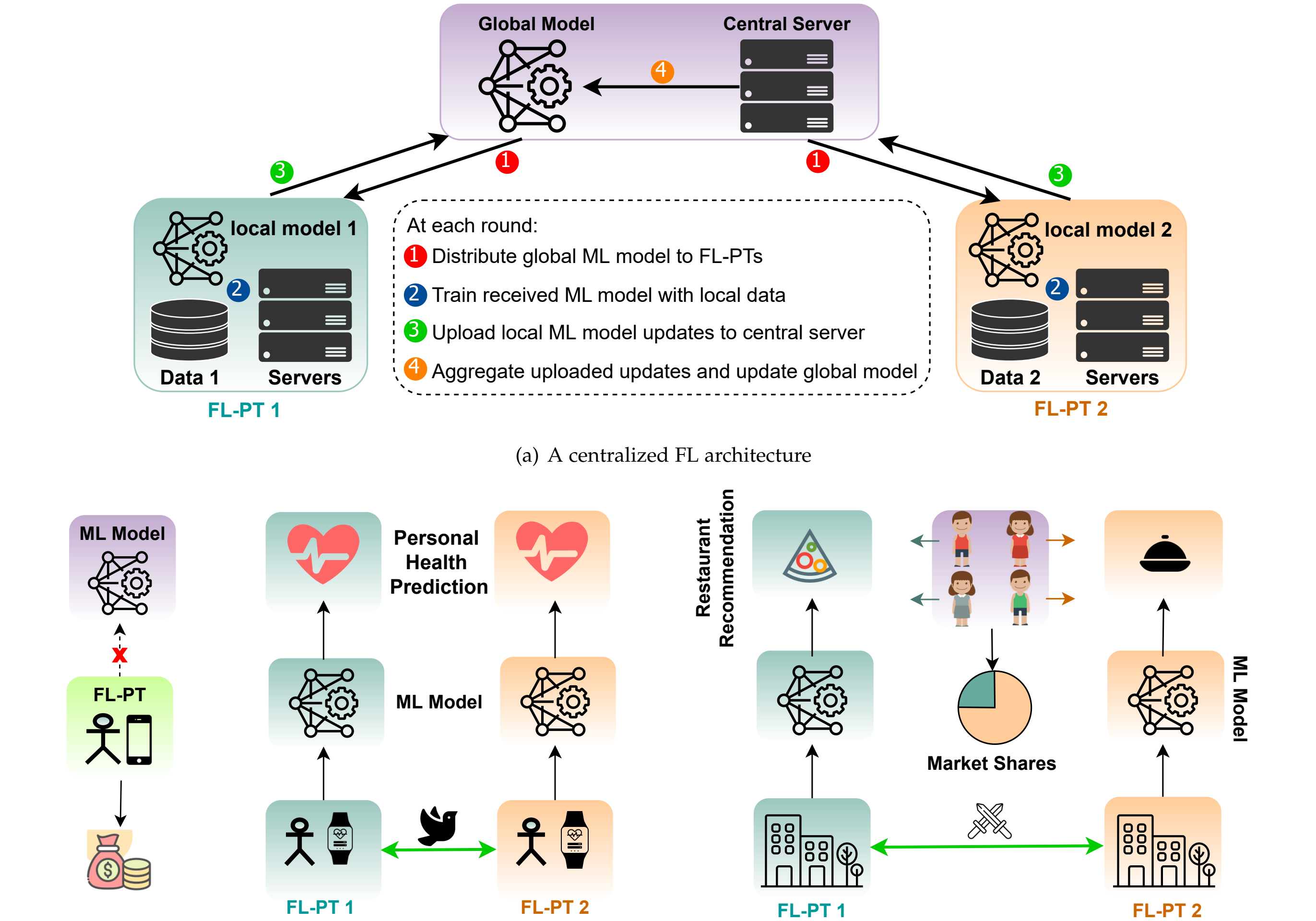
As data privacy has become a growing concern in modern societies and world-wide governments are enacting related laws [3], FL has become especially important and it has many promising applications such as digital banking [4], [5], recommender systems for online services [6], [7], safety monitoring [8], healthcare [9], [10] and mobile applications [11]. With the development of 5G/6G networks, the applications of FL are also expanding rapidly [12], [13]. In such context, proper incentive mechanisms are urgently needed to incentivize data owners to join the FL training process [1], [14], [15]. Based on the relationship of the eventual model users and FL-PTs, the prevailing use cases of FL can be categorized as follows:

• *Xiaohu Wu is with the WeBank-NTU Joint Research Centre on Fintech, Nanyang Technological University, Singapore. Han Yu is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. E-mail: {xiaohu.wu, han.yu}@ntu.edu.sg*

Manuscript received April 19, 2005; revised August 26, 2015.

1) The FL-PTs are not the model users. The model users benefit from the model performance improvement, which has no direct impact on FL-PTs' well-being. In this case, since FL-PTs are not interested in the model performance, monetary rewards are often used to incentivize FL-PTs to contribute more local resources to the FL training process [16], [17], [18], [19], [20], [21], [22]. This is illustrated in Figure 1(b).
2) The FL-PTs are the eventual model users, and an FL-PT's utility is not affected by the model performance improvements obtained by other FL-PTs. This is illustrated in Figure 1(c) where each person only cares about its own health and ML model accuracy. In this case, different FL-PTs achieve different levels of model performance improvements and monetary transfer is taken to let the FL-PTs of higher payoff compensate the FL-PTs of lower payoff [23].
3) The FL-PTs are the eventual model users. However, an FL-PT's utility is dependent on the model performance improvements of all FL-PTs. This is illustrated in Figure 1(d) where each customer compares different FL-PTs (e.g., apps) and weighs the ML model performances of all FL-PTs before choosing an app for use.

Although the third use case has been well recognized [5], [7], [14], [24], its formal study is missing in existing literature, which hinders FL-based collaboration in practical situations. This is probably due to the unclearness of the metrics and additional constraints for judging and maintaining the viability of FL in a competitive market. In contrast, in the first and second use cases, such questions do not exist since the FL viability is obvious. In this paper, we study the third use case, which is commonly found in horizontal



(a) A centralized FL architecture

(b) Case 1 (c) Use case 2 (d) Use case 3: the use case of this paper

存在竞争关系

Fig. 1: FL overview and its three use cases

federated learning (HFL) scenarios in which the FL model is used for profit-making activities (*e.g.*, personalized loan interest determination in digital banking) [2]. In such scenarios, FL-PTs provide the same services and compete for a same group of customers. Market share is a key indicator of FL-PTs' market competitiveness (*i.e.*, how well a firm is doing against its competitors) [14], [25]. For example [14], [24], when multiple digital banks collaboratively train a FL model to predict the creditworthiness of customers, larger banks with more high quality data may be reluctant to join FL for fear of benefiting its smaller competitors and eroding its market share. The underlying challenge has been framed as the "free-rider problem" in FL [1] for which monetary incentives are not effective.

This challenge has inspired trust-based FL ecosystems to emerge, which are built on the proposition that the FL model performance improvement obtained by each FL-PT shall be proportional to its contribution [1], [15]. However, there is no analytical tool to help FL-PTs determine how it affects their market shares under different market conditions. Thus, the data owners still face uncertainty regarding potential market share erosion as a result of joining FL, which hinders the adoption of FL in HFL scenarios involving competitors. As FL can improve FL-PTs' model performance which, in turn, results in better products or services to enhance the collective utility of the businesses and customers involved [1], [2], a clear understanding of the impact of joining FL on each FL-PT's market share to facilitate wider adoption of FL is crucial.

Suppose that the original market size is P and there are n firms (*i.e.*, FL-PTs) $\mathcal{C} = \{1, 2, \cdots, n\}$ in a given market which can join FL. The original market share of FL-PT $i \in \mathcal{C}$ is $MS_i \in (0, 1)$, where $\sum_{i=1}^{n} MS_i = 1$. After joining FL, the market size becomes P' and the market share of FL-PT $i \in \mathcal{C}$ becomes $MS'_i \in (0, 1)$, where $\sum_{i=1}^{n} MS'_i = 1$. We now define a notion of market stability to measure the viability of FL. Intuitively, no FL-PT is willing to accept a significant reduction in its market share as a result of joining FL (i.e., no FL-PT's interest shall be sacrificed), even if all FL-PTs benefit from FL with improved model performance and quality of service. Formally, we have:

Definition 1 (δ-Stable Market). In a given FL ecosystem, the market is δ-stable if the market share of every FL-PT $i \in \mathcal{C}$ satisfies the following condition: $V_i = MS_i - MS'_i \leq \delta$, where $\delta \in (-1, 1)$ is an upper bound of the difference of the market shares for any given FL-PT before and after joining FL.

We refer to V_i as *the market variance of FL-PT i* as a result of joining FL. Ideally, the market share of each FL-PT i does not vary (i.e., $\delta = 0$). However, this is hard to achieve in practice. Thus, we relax it to δ not exceeding a small positive real number within $(0, 1)$ and that each FL-PT i experiences an acceptable reduction in its market share after joining FL in the worst case. Then, the FL ecosystem is said to be viable only if the market is δ-stable. The value of δ is negotiated among the n FL-PTs.

In this paper, we propose an analytical framework for market share-based decision support about participation in FL (MarS-FL) by leveraging game theory and marketing models [26], [27]. Given the original FL-PTs' market shares before joining FL, the market dynamics are driven by exogenous factors such as the customer loyalty, and the leaving and switching rates of customers for each FL-PT; also see Figure 2. It can be embodied by a general mathematical model in economic literature [27]. We also take an additional factor into account: the possible market growth rate. In the FL context, the switching customer of a FL-PT will reconsider choosing one of the n FL-PTs. Due to market growth, a customer who newly enters the market will choose one of the n FL-PTs. In both cases, the possibility of choosing an FL-PT is positively correlated to the service quality of the FL-PT which, in turn, depends on the model performance of this FL-PT. Intuitively, the model performance improvement of each FL-PT is bounded, resulting in bounded variations in market shares.

For the decision-maker of an FL ecosystem, its decision variables are the relative ML model performance improvements $\{Q_i\}_{i=1}^n$ of the n FL-PTs where $Q_i \in [0, 1]$ and $\sum_{i=1}^n Q_i = 1$. MarS-FL provides a tight lower bound $Q_i^{min} \in [0, 1]$ such that the market δ-stability is maintained when $Q_i \geq Q_i^{min}$. Based on these bounds, we define a notion of *friendliness* κ to measure how conducive a market is for FL where $\kappa = 1 - \sum_{i=1}^n Q_i^{min}$. FL is viable when $\kappa \in [0, 1]$; then there exists a feasible allocation of $\{Q_i\}_{i=1}^n$. The value of κ implies the space size of decision variables; the larger the value of κ, the higher the friendliness. Further, MarS-FL also provides a sufficient and necessary condition for the viability of FL in a competitive market. Through numerical experiments, we demonstrate the capability of MarS-FL to analyze the viability of FL under a wide range of market conditions. Our results are useful for identifying the market conditions under which collaborative FL model training is viable among competitors, and the requirements that have to be imposed while applying FL under these conditions.

The rest of this paper is organized as follows. In Section 2, we introduce the related work. In Section 3, we define the FL market model and our problem formulation. The analytical framework and results of MarS-FL are presented in Section 4. In Section 5, extensive experimental results illustrate the FL viability in different market environments. Finally, we conclude the paper in Section 6.

2 Related Work

Our work in this paper is broadly related to incentive mechanism design for motivating participation in FL. Existing research in this domain can be broadly divided into two categories [1]: 1) monetary FL incentive schemes, and 2) non-monetary FL incentive schemes.

Monetary incentive schemes are generally designed for the first FL use case. These schemes that provide monetary rewards to FL-PTs in order to motivate them to contribute more local resources to FL model training. FL-PTs in this case are often resource-constrained devices such as mobile devices. The types of resources include not only local data but also computational and communication resources. Sarikaya *et al.* [17] model the interaction between the FL server (*i.e.,* the model user) and FL-PTs as a Stackelberg game with the goal of improving the FL model performance by jointly optimizing the commitment of local computational resources and the allocation of the incentive budget. Zhan *et al.* [19] leverage the Stackelberg game to model the interaction to incentivize FL-PTs to contribute more data. Zeng *et al.* [22] consider motivating FL-PTs to contribute multiple types of resources by applying multi-dimensional procurement auction theory and propose a scheme that selects and rewards a fixed number of FL-PTs for FL training. Song *et al.* [16] and Zhang *et al.* [18] propose schemes to select and pay FL-PTs based on reputation and reverse auction. The FLI approach [20], [21] has been proposed which supports fair allocation of incentive payout to FL-PTs using future revenues generated by the FL model. These approaches leverage incentive mechanism research results in economics and game theory. As they generally assume that the FL-PTs care only about monetary rewards, the intermediate FL models and the final FL model are generally freely shared during the FL training process.

For the second use case, Tang *et al.* [23] assume that each FL-PT will commit all its data for local model training, and characterize the interaction among FL-PTs as a non-cooperative game. They propose, under mild assumptions, an incentive scheme that achieves social welfare maximization, individual rationality and budget balance. The scheme features a distributed algorithm and incurs additional costs and transmission for organizations. To address this, Chen *et al.* [28] propose to use the multi-player multi-action zero-determinant strategy to maximize the social welfare. Xu *et al.* [29] focus on non-monetary schemes and design a gradient reward scheme with a fairness guarantee. At training time, they use a sparsifying gradient trick to control the quality of the aggregated gradient downloaded from the server as reward to each FL-PT such that its quality is commensurate to the quality of the FL-PT's uploaded/contributed gradient. Zhang *et al.* [30] model FL-PTs' long-term selfish participation behaviors as an infinitely repeated game and derive the optimal subgame perfect Nash equilibrium which minimizes the number of free riders while maximizing the amount of local data for model training.

In the third FL use case in which the FL-PTs are also the end users of the final FL models and competing in the same market, which is the focus of our study, sharing the same FL model among all participants without regards to their contributions has been shown to cause breakdown of collab-

This article has been accepted for publication in IEEE Transactions on Big Data. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2022.3186991

oration [24]. Non-monetary incentive schemes which assign each FL-PT a different model in each training iteration with performance reflecting its contribution are starting to emerge [15], [24], [31]. However, these existing approaches do not take FL-PTs' market shares into account when allocating different versions of FL models to them. MarS-FL bridges this gap by identifying the optimal FL participation strategies, the viable operational space of an FL system, and the market conditions under which FL can be beneficial for a given FL-PT. We refer readers to [32], [33] for two more comprehensive surveys on FL incentive schemes.

3 FL MARKET DYNAMICS

In this section, we provide a detailed analysis of the market dynamics surrounding the third FL use case scenario. We first introduce two typical FL system architectures.

3.1 FL System Architectures

In FL, each FL-PT $i \in \{1, 2, \cdots, n\}$ has a local model with a common structure. The training process iterates for multiple rounds. We use w_i^t to denote the parameters of the local model of FL-PT i at round $t \in \{1, 2, \cdots, T\}$. In the *centralized FL architecture*, there is a central FL server to mediate the training process [23]. In the beginning of round t, let w^t denote the parameters of the global FL model owned by the FL server and $\hat{w}_i^t$ denote the model parameters sent from the FL server to FL-PT i. In the current design, $\hat{w}_i^t = w^t$, which is not applicable to the use case of this paper; however, the related questions can be addressed in future based on the results of this paper. Each FL-PT i downloads $\hat{w}_i^t$. i uses a batch of its local data to train model $\hat{w}_i^t$ and computes the gradient ∇w_i^t. The updated local model is denoted as $w_i^t = \hat{w}_i^t - \eta \nabla w_i^t$ where η is the learning rate. i then uploads w_i^t (or ∇w_i^t) to the FL server. The FL server produces the global model w^{t+1} by aggregating the received local model updates from all FL-PTs. Let x_i denote the amount of data that i decides to use for its local model training. In the classic Federated Averaging strategy, the aggregating process is defined as [34], [35]:

$$w^{t+1} = \frac{x_i}{\sum_{j=1}^{n} x_j} w_i^t.$$

In the *decentralized FL architecture*, the training process iterates without a central FL server [1]. In the beginning of round t, each FL-PT i uses a batch of its local data to train its local model w_i^{t-1} obtained at the last round and computes the gradient ∇w_i^t. The updated model is denoted as $w_i^t = w_i^{t-1} - \eta \nabla w_i^t$. i then shares the updates ∇w_i^t with other FL-PTs which have agreed to establish collaborative training partnership with i (based on considerations such as trust [24]). In the meantime, i also downloads local model updates from other partners. Then, i further updates its model w_i^t by aggregating the received local model updates from its partners.

3.2 FL Market Model

Let us consider n firms (*i.e.*, FL-PTs) $\mathcal{C} = \{1, 2, \cdots, n\}$ in a given market which can join FL. The market size of FL-PT i is P_i (*e.g.*, in terms of the number of customers) and their aggregated market size is denoted as P. The (relative) market share of i is $MS_i \in (0, 1]$ where

$$P_i = MS_i \times P. \tag{1}$$

Thus, we have $\sum_{i=1}^{n} MS_i = 1$. Firm i owns a local dataset $\mathcal{D}^i$ with $D^i = |\mathcal{D}^i|$ samples. The local model performance can be measured by the model loss value, which in turn, affects the quality of service offered by the firm. Smaller loss values imply better quality of service.The variable $x_i \in [0, D^i]$ denotes the amount of local data that FL-PT i decides to use for FL training. If i trains its model solely using its local data $\mathcal{D}^i$ without joining FL, the final loss function L_i of the resulting model can be expressed as:

$$L_i = L_i\left(D^i\right).$$

If i joins FL, it can leverage local models from other FL-PTs. After the FL model training process ends, the model loss function L_i' of FL-PT i can be expressed as:

$$L_i' = L_i'\left(\left\{D^j\right\}_{j=1}^{n}, \{x_j\}_{j=1}^{n}, Trad\right).$$

$Trad$ denotes a given model exchange scheme among FL-PTs (*e.g.*, under the centralized FL architecture [34], or the decentralized FL architecture [24]). It determines how much information i obtains during the FL process and thus the value of L_i'.

In this paper, we consider FL ecosystems with non-monetary incentive schemes in which an FL-PT's reward is reflected by the final loss function value of the model it obtains. We make a natural assumption that for an FL-PT $i \in \mathcal{C}$, its contribution to the FL system increases as it commits more of its local data for FL training, and the loss function value of i decreases as its contribution to FL increases. Formally, the assumption is expressed as follows.

Assumption 1. Suppose $D^i \geq x_i^{(1)} > x_i^{(2)} \geq 0$. Then, we have

$$L_i'\left(x_i^{(1)}\right) < L_i'\left(x_i^{(2)}\right).$$

After the FL model training, the model performance improvement of FL-PT $i \in \mathcal{C}$ is measured by:

$$d_i = L_i - L_i' \geq 0. \tag{2}$$

We focus on studying the effect of the final ML model performance improvements $\{d_i\}_{i=1}^{n}$ on the market shares. $\sum_{j=1}^{n} d_j$ denotes the aggregate model performance improvement for all FL-PTs. The relative model performance improvement Q_i for FL-PT i is defined as:

$$Q_i = \frac{d_i}{\sum_{j=1}^{n} d_j}, \tag{3}$$

where we have:

$$Q_i \in [0, 1] \text{ and } \sum\nolimits_{i=1}^{n} Q_i = 1. \tag{4}$$

Given the value of $\sum_{j=1}^{n} d_j$, $\{Q_i\}_{i=1}^{n}$ uniquely corresponds to $\{d_i\}_{i=1}^{n}$.

The service quality improvement S_i' of FL-PT i is proportional to the model performance improvement d_i and is defined as:

$$S_i' = \alpha \cdot d_i \tag{5}$$

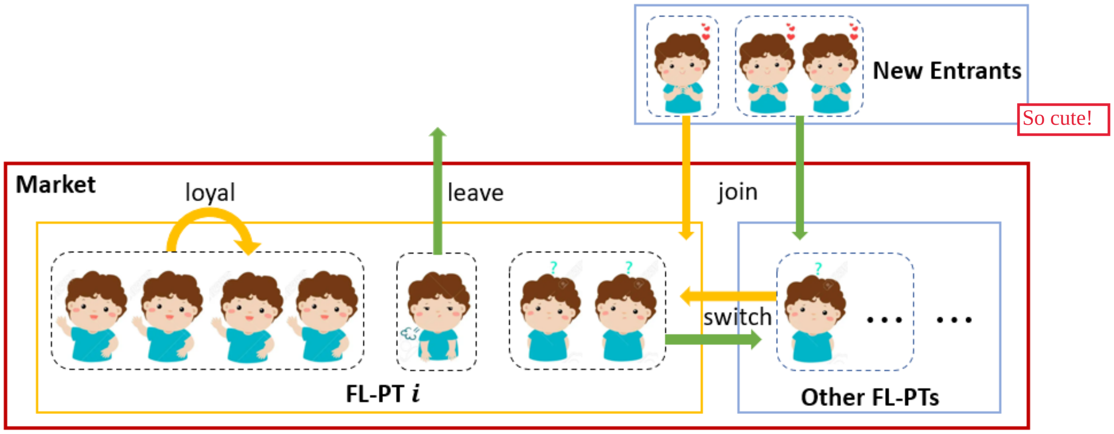


Fig. 2: The possible movements of customers among FL-PTs, and into and out of the market.

where, in a given market, the weight $\alpha > 0$ is the degree of improvement to service quality brought by a unit of model performance improvement. The relative service quality improvement for i, S_i, is defined as:

$$S_i = \frac{S'_i}{\sum_{j=1}^{n} S'_j} \overset{(a)}{=} Q_i \in [0, 1] \tag{6}$$

where the equality (a) is a result of Eq. (5). We focus on a competitive market in which each firm serves a group of customers with no overlap. The attractiveness of firms to customers changes based on their relative service quality. The customers of a firm may switch to another firm. If the model performance of the firms improves in general, the entire market becomes more attractive, resulting in increases in the market size. Such market dynamics are detailed in the next subsection. We use MS'_i to denote the market share of i after the FL process.

3.3 FL-PT Market Share Dynamics

In this subsection, we adapt the classic model of Rust and Zahorik [27] to characterize the FL-PTs' market shares $\{MS'_i\}_{i=1}^n$ after FL training concludes. The market dynamics are mainly driven by (i) the loyalty and the leaving or switching actions of the existing customers of different FL-PTs and (ii) the action of the customers who newly enter the market, which is illustrated in Figure 2. Initially, the customers for $i \in \mathcal{C}$ (whose number totals to $P_i = MS_i \times P$) can be classified into three types:

- Let $r_i \in [0, 1]$ denote the proportion of customers loyal to i. After FL model training, $(1 - r_i)P_i$ customers will leave i. They will either exit the market or switch to other FL-PTs.
- Let $\nu_i \in [0, 1 - r_i]$ denote the proportion of customers who leave the market where we have

$$r_i + \nu_i \in [0, 1]. \tag{7}$$

  After FL model training, $\nu_i P_i$ customers of i will exit the market.
- The remaining $(1 - r_i - \nu_i)P_i$ customers are still active in the market, but will consider the possibility of switching to other FL-PTs. These customers are referred to as "free customers" from i.

Up to θP new customers may join the market and be served by the n firms, where $\theta \geq 0$. Generally, the market share of FL-PT $i \in \mathcal{C}$ consists of the following:

- The number of loyal customers of i is:

$$P_{l,i} = r_i P_i. \tag{8}$$

- We use S_i defined in Eq. (6) to denote the attractiveness of i to customers. The fraction of free customers from $j \in \mathcal{C}$ joining i is linearly proportional to S_i. Specifically, the number of customers leaving j for i equals to $(1 - r_j - \nu_j)P_j S_i$. The total number of free customers from itself and the other $(n - 1)$ FL-PTs joining i is:

$$P_{f,i} = \sum_{j \in \mathcal{C}} (1 - r_j - \nu_j)P_j S_i. \tag{9}$$

- θP new customers will be divided among the n FL-PTs. The number of new customers joining i is proportional to its attractiveness:

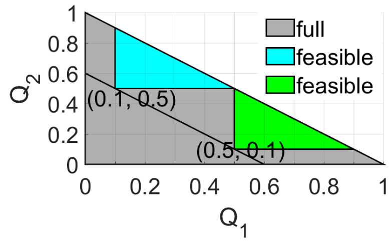
$$P_{e,i} = S_i \theta P. \tag{10}$$

After FL training, $\sum_{i=1}^n \nu_i P_i$ customers will exit the market; we assume that this amount is small such that the whole market size is still positive and the new market size becomes:

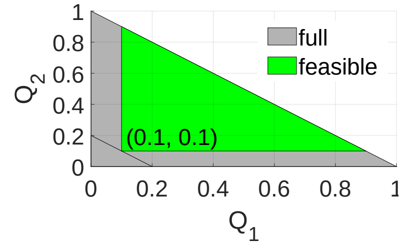
$$P' = (1 + \theta)P - \sum_{i=1}^{n} \nu_i P_i > 0. \tag{11}$$

Based on Eq. (1) and Eq. (8)–(11), the market share of FL-PT i becomes:

$$\begin{aligned} MS'_i &= \frac{P_{l,i} + P_{f,i} + P_{e,i}}{P'} \\ &= \frac{r_i MS_i + S_i \sum\limits_{j \in \mathcal{C}} (1 - r_j - \nu_j) MS_j + S_i \theta}{(1 + \theta) - \sum\limits_{j=1}^{n} \nu_j MS_j}. \end{aligned} \tag{12}$$



(a) $\kappa = 1 - \sum_{i=1}^{2} Q_i^{\min} = 0.4$



(b) $\kappa = 1 - \sum_{i=1}^{2} Q_i^{\min} = 0.8$

Fig. 3: Friendliness of market environments to FL ($n = 2$).

3.4 Decision Support Tasks

Below, we formulate the decision support tasks to be addressed while understanding the role of FL in a competitive market.

3.4.1 *A Non-Cooperative Game*

The model of non-cooperative game provides insight into the interaction of FL-PTs. If a game has a dominant strategy equilibrium, such an equilibrium will be an ideal notion to predict the best course of action by any given player [26].

The n FL-PTs are assumed to be selfish but not malicious. All FL-PTs compete against each other. Like [30], an FL-PT can choose its participation level (i.e., the amount of data that it commits to FL training) to maximize its own interest and to avoid that other FL-PTs become free-riders. However, it has no intention to damage the FL ecosystem. Each FL-PT i has a decision variable x_i which determine how many local data samples are used for FL training. x_i, in turn, determines the loss function value L_i' for i after the FL model training concludes. This corresponds to a non-cooperative game where the n FL-PTs are players. i's strategy space is $\mathcal{X}_i = \left[0, D^i\right]$ and a single strategy is $x_i \in \mathcal{X}_i$. The strategy space of the other $(n-1)$ FL-PTs is denoted as

$$\mathcal{X}_{-i} = \{(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \,|\, x_{i'} \in \mathcal{X}_{i'}, i' \neq i\}.$$

i aims to enhance its market status in a competitive market, and its payoff is defined as its market share MS_i' (or equivalently $MS_i'P'$). An ideal situation is that each FL-PT $i \in \mathcal{C}$ can decide its best strategy $x_i^* \in \mathcal{X}_i$ to maximize its market share, even if it does not have any knowledge of the strategies of other FL-PTs. Formally, $x^* = \{x_i^*\}_{i=1}^n$ is a dominant strategy equilibrium of the game if we have that, for each FL-PT $i \in \mathcal{C}$ and any alternate strategy $x_i' \in \mathcal{X}_i$ different than x_i^*,

$$MS_i'(x_i^*, x_{-i}) \geq MS_i'(x_i', x_{-i}), \forall x_{-i} \in \mathcal{X}_{-i}. \quad (13)$$

3.4.2 *FL Viability and Market Friendliness*

Another decision support task of importance to the manager of an FL ecosystem is to understand the viability of FL in a competitive market. As discussed above, the market dynamics are parameterized by θ, $\{r_j\}_{j=1}^n$ and $\{\nu_j\}_{j=1}^n$. The effect of FL on model performance is parameterized by $\{Q_j\}_{j=1}^n = \{S_j\}_{j=1}^n$. The resulting market shares $\{MS_i'\}_{i=1}^n$ after the FL model training are functions of these parameters.

As formally analyzed in the "Market Stability" subsection, for every FL-PT $i \in \mathcal{C}$, there is a tight lower bound of the relative ML model performance improvement, which is denoted by $\hat{Q}_i^{\min}$, to maintain the market δ-stability. Formally, the market is δ-stable if and only if $Q_i \geq \hat{Q}_i^{\min}, \forall i \in \mathcal{C}$. Mathematically, it is possible that $\hat{Q}_i^{\min} < 0$. This means that the market share reduction of FL-PT i will not exceed δ even if i does not obtain model performance improvement in the FL model training process. However, in a real FL ecosystem, each FL-PT i can achieve some level of model performance improvement, *i.e.*, $Q_i \geq 0$. Thus, we let

$$Q_i^{\min} = \max\left\{\hat{Q}_i^{\min}, 0\right\} \geq 0, \forall i \in \mathcal{C} \quad (14)$$

and the market δ-stability is achievable if and only if the following relation is satisfied:

$$Q_i \geq Q_i^{\min}, \forall i \in \mathcal{C}. \quad (15)$$

We now define an index to measure the friendliness of market environments towards FL. Let

$$Q_i = Q_i^{\min} + y_i, \forall i \in \mathcal{C}. \quad (16)$$

Definition 2. Suppose that we are given the bounds $\{Q_i^{\min}\}_{i=1}^n$ and the decision variables are $\{Q_i\}_{i=1}^n$. The level of market friendliness towards FL, κ, is defined as:

$$\kappa \triangleq 1 - \sum_{i=1}^{n} Q_i^{\min} \stackrel{(a)}{=} \sum_{i=1}^{n} y_i \quad (17)$$

where the equality (a) is due to Eq. (4) and Eq. (16).

The value of κ indicates the viability of FL. If $\kappa \in [0, 1]$, then the market δ-stability can be maintained and FL is viable in a competitive market. If $\kappa \notin [0, 1]$, then FL is unviable. The explanation is as follows. If $\kappa \in [0, 1]$, then there exists a feasible allocation of $\{y_i\}_{i=1}^n$ among FL-PTs such that $y_i \geq 0, \forall i \in \mathcal{C}$; further, the relation (15) is satisfied by Eq. (16) and the market δ-stability is maintained. Thus, FL is viable. The value of κ determines the decision space size of an FL designer; the larger the value of κ, the more friendly a market is towards FL:

- When $\kappa = 1$, then $\sum_{i=1}^n Q_i^{\min} = 0$. Further, by (14), we have $Q_i^{\min} = 0$, $\forall i \in \mathcal{C}$. This indicates that any FL framework can satisfy the market δ-stability, since Eq. (15) holds naturally.

- The case of $\kappa \in (0,1)$ is illustrated in Figures 3(a) and 3(b) where the grey triangle denotes the full region in which each point represents a possible pair (Q_1, Q_2). Each colored triangle (blue or green) denotes the feasible region that is the decision space of an FL designer and in which each point corresponds to a feasible pair (Q_1, Q_2) that can achieve the market δ-stability. In Figure 3(a), we illustrate the cases with $\kappa = 0.4$, but different $(Q_1^{\min}, Q_2^{\min})$ values. In Figure 3(b), we illustrate the case with a larger $\kappa = 0.8$.
- When $\kappa = 0$, then $\sum_{i=1}^n Q_i^{\min} = 1$. By Eq. (4), we have $\sum_{i=1}^n Q_i = \sum_{i=1}^n Q_i^{\min}$; the only feasible solution that can satisfy the relation (15) is $Q_i = Q_i^{\min}$, $\forall i \in \mathcal{C}$. Then, the FL framework has to strictly satisfy this solution in order to achieve the market δ-stability.

In contrast, if $\kappa < 0$, then $\sum_{i=1}^n y_i < 0$: there exists an FL-PT $i \in \mathcal{C}$ such that $y_i < 0$. This indicates that the relation (15) cannot be satisfied. If $\kappa > 1$, then $\sum_{i=1}^n Q_i^{\min} < 0$, which contradicts Eq. (14).

In the next section, we will analyze the dominant strategy equilibrium of the game above, the bound $Q_i^{\min}$ for all $i \in \mathcal{C}$ and the friendliness κ.

4 FL-PT Behaviour and Achievability of Market's δ-Stability

4.1 A Dominant Strategy Equilibrium

Proposition 1. The dominant strategy equilibrium of the non-cooperative game is $x^* = (D^1, D^2, \cdots, D^n)$.

Proof 1. It suffices to show that Eq. (13) holds when $x_i^* = D^i$. With abuse of notation, we let Q_1, Q_2, $\cdots$, Q_n (*resp.* Q_1', Q_2', $\cdots$, Q_n') denote the relative model performance improvements of the n FL-PTs under the strategy profile (x_i^*, x_{-i}) (*resp.* (x_i', x_{-i})). Based on Assumption 1, we have $L_i'(x_i^*) < L_i'(x_i')$. The loss function values of other FL-PTs remain unchanged. Thus, from Eq. (2) and Eq. (3), we have:

$$Q_i > Q_i' \text{ and } Q_j < Q_j', \forall j \in \mathcal{C} - \{i\} \tag{18}$$

Let

$$\varepsilon_i = Q_i - Q_i' \overset{(a)}{>} 0. \tag{19}$$

where the inequality (a) is due to Eq. (18). Finally,

$$\begin{aligned} & MS_i'(x_i^*, x_{-i}) - MS_i'(x_i', x_{-i}) \\ \overset{(c)}{=} & \frac{\sum\limits_{j\in\mathcal{C}} (1 - r_j - \nu_j) MS_j \varepsilon_i + \varepsilon_i \theta}{1 + \theta - \sum\limits_{j=1}^n \nu_j MS_j} \overset{(d)}{\geq} 0 \end{aligned}$$

The equality (c) is due to Eq. (6) and Eq. (12); the inequality (d) is due to Eq. (7), Eq. (11) and Eq. (19).

In a competitive market, each FL-PT aims to maximize its market share. By Proposition 1, an FL-PT's market share is maximized when it uses all of its local data for FL training, regardless of others' strategies. Ideally, under the dominant strategy solution, the FL ecosystem will also produce the best global ML model.

4.2 Market Stability

We define the following factors to represent the overall market dynamics and show that the lower bound $Q_i^{\min}$ in (15) is governed by some of these factors under a simple mathematical structure. Let

$$\begin{aligned} v_o &= \sum_{i=1}^n v_i MS_i \\ e &= 1 + \theta - v_o \\ f_o &= \frac{\theta + \sum\limits_{i=1}^n (1 - r_j - \nu_i) MS_i}{e} \\ \hat{r}_i &= \frac{r_i MS_i}{e} \end{aligned} \tag{20}$$

After the FL training process, we have

- v_o represents the rate at which the customers leave the market (i.e., the total number of such customers is $v_o P$).
- $(e \times P)$ represents the total number of customers in the market by Eq. (11). We refer to $e \in (0, +\infty)$ as **the expanded scale of the market** in relation to the original population size P.
- $r_i MS_i P$ is the number of loyal customers of FL-PT i in the original market. These customers will continue to be served by FL-PT i after the FL training process. $\hat{r}_i$ is the ratio of $r_i MS_i P$ to eP. We refer to $\hat{r}_i$ as **the proportion of old customers of FL-PT** i in relation to the new population size eP.
- $(1 - r_j - \nu_i) MS_i P$ is the number of free customers of FL-PT i who will reconsider which FL-PT to join after FL model training. θP is the number of new customers who consider joining one of the n FL-PTs. Thus, $\theta + \sum_{i=1}^n (1 - r_j - \nu_i) MS_i \times P$ represents the total number of customers who eventually joins one FL-PT. We refer to f_o as **the proportion of vacillating customers** in relation to the new population size eP.

Suppose we are given an arbitrary FL training framework in which each FL-PT $i \in \mathcal{C}$ uses x_i local data samples for FL model training. This process determines the values of $\{Q_1, \ldots, Q_n\}$ via Eq. (3). The proposition below shows that the δ-stability of the market is achieved when $\{Q_1, \ldots, Q_n\}$ satisfy a particular relation with the original market status and dynamics.

Proposition 2. Define

$$\hat{Q}_i^{\min} = \frac{(MS_i - \delta) - \hat{r}_i}{f_o}. \tag{21}$$

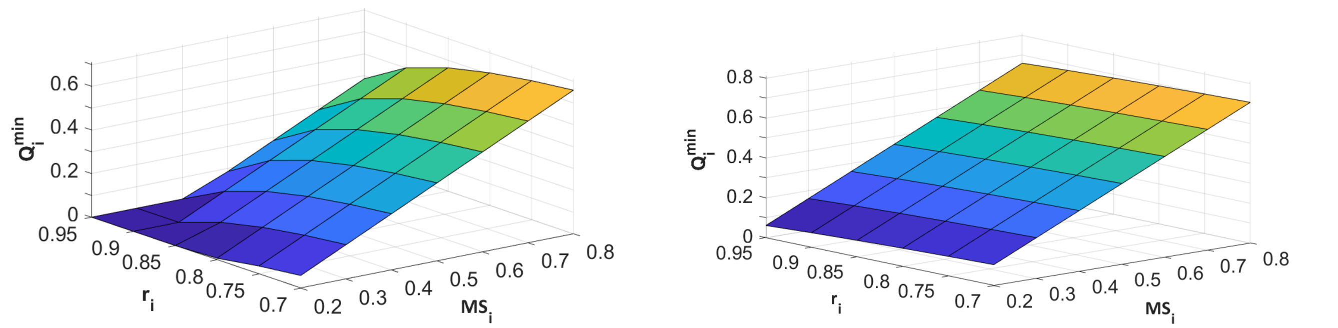
The market stability is achieved if and only if

$$Q_i \geq Q_i^{\min}, \forall i \in \mathcal{C},$$

where $Q_i^{\min} = \max\left\{\hat{Q}_i^{\min}, 0\right\}$.

Proof 2. To achieve market stability, based on Eq. (12) and Definition 1, we have for each FL-PT $i \in \mathcal{C}$ that:

$$\begin{aligned} V_i &= MS_i - MS_i' \\ &= MS_i - \frac{r_i MS_i + S_i \sum\limits_{j\in\mathcal{C}} (1 - r_j - \nu_j) MS_j + S_i \theta}{(1+\theta) - \sum\limits_{j=1}^n \nu_j MS_j} \\ &\leq \delta. \end{aligned}$$



(a) Mature market

(b) Fast-growing market

Fig. 4: The minimum relative improvement, $Q_i^{\min}$, required to maintain the market δ-stability

Further, we have

$$S_i \geq \frac{(MS_i - \delta)\left(1 + \theta - \sum_{j=1}^{n} v_j MS_j\right) - r_i MS_i}{\theta + \sum_{j=1}^{n}(1 - r_j - v_j)MS_j}$$

By Eq. (6), we have $S_i = Q_i$. Further, based on Eq. (15) and Eq. (20), Proposition 2 holds.

Proposition 2 provides the minimum relative ML model performance improvements $\{Q_i^{\min}\}_{i=1}^{n}$ required to maintain the market δ-stability. The lower bound $Q_i^{\min}$ of FL-PT i is governed by its own market features and the overall market dynamics, including the original market share MS_i, the proportion $\hat{r}_i$ of old customers of FL-PT i, and the proportion f_o of the vacillating customers of all FL-PTs. For FL-PTs with $\hat{Q}_i^{\min} \leq 0$, their δ-stability can be maintained even if they do not benefit from the FL training process and their ML model performance is not improved. Conversely, we refer to an FL-PT $i \in \mathcal{C}$ with $\hat{Q}_i^{\min} > 0$ as *a sensitive FL-PT*, and its ML model performance has to be improved to maintain the δ-stability. The set of sensitive FL-PTs is denoted by $\mathcal{C}'$ and we let $n' = |\mathcal{C}'|$. Next, by Definition 2, we can obtain the friendliness towards FL under any given market environment.

Proposition 3. The friendliness of market environment to FL is:

$$\kappa = 1 - \sum_{i=1}^{n} \max\left\{0, \frac{(MS_i - \delta) - \hat{r}_i}{f_o}\right\}$$
$$= 1 - \frac{\sum_{i \in \mathcal{C}'}(MS_i - \hat{r}_i) - n'\delta}{f_o}$$

Proof 3. It follows from Definition 2 and Proposition 2.

By Proposition 3, κ is governed by the number n' of sensitive FL-PTs, the total market share $\sum_{i \in \mathcal{C}'} MS_i$ of these sensitive FL-PTs $\mathcal{C}'$, the total proportion $\sum_{i \in \mathcal{C}'} \hat{r}_i$ of the old customers of $\mathcal{C}'$, and the proportion f_o of vacillating customers. It is observed that the friendliness of market environments to FL simply by the four factors that reflect the overall market dynamics, instead of the individual's market features.

Proposition 4. FL is viable in a competitive market if and only if

$$\frac{\left(\sum_{i \in \mathcal{C}'} MS_i - \hat{r}_i\right) - f_o}{|\mathcal{C}'|} \leq \delta.$$

Proof 4. FL is viable in a competitive market if and only if $\kappa > 0$. The proposition follows from Proposition 3.

5 Numerical Studies and Analysis

In this section, we illustrate the minimum requirement of the relative ML model performance improvement to maintain the market δ-stability, and the friendliness to FL in typical market environments.

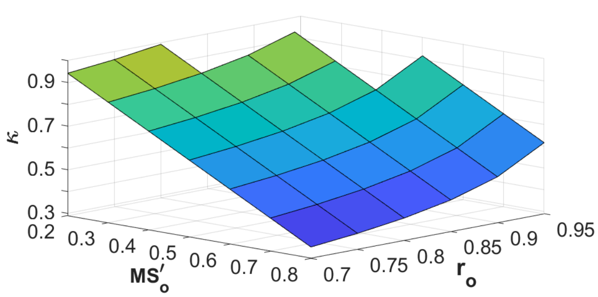
5.1 General Settings

The number n of FL-PTs can be arbitrary. Different types of markets have different growth rates. We set θ to 0.1 and 0.5 to simulate a mature market and a fast-growing market, respectively. The average proportion of loyal customers, r_o, is set to the range of $[0.7, 0.95]$ (i.e., we set $r_1 = \cdots = r_n = r_o$ and for $i \in \mathcal{C}$, r_i ranges from 0.7 to 0.95 with a stepsize 0.05). The rates at which customers leave an FL-PT and all the n FL-PTs are generally small and set to 0.02 (i.e., $v_1 = \cdots = v_n = v_o = 0.02$).

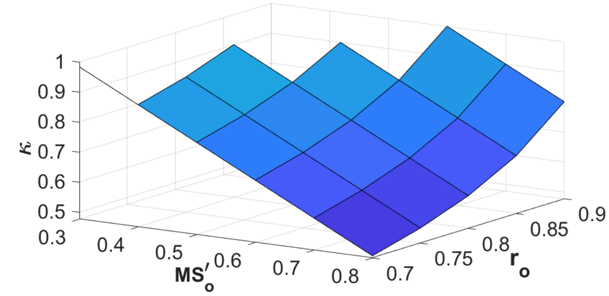
5.2 The Minimum Requirement $\{Q_i^{\min}\}_{i=1}^{n}$

The market share represents the competitiveness of a firm $i \in \mathcal{C}$. We set MS_i to the range of $[0.2, 0.8]$ with a stepsize 0.1. The proportion $\hat{r}_i$ of old customers and the proportion f_o of vacillating customers are computed by Eq. (20). The minimum ML model performance improvement requirement of an individual FL-PT i, $Q_i^{\min}$, is illustrated in Figure 4. Overall, different FL-PTs have different requirements $Q_i^{\min}$ to maintain the market δ-stability based on their market shares, their customer loyalties and the market dynamics such as the market growth rate:

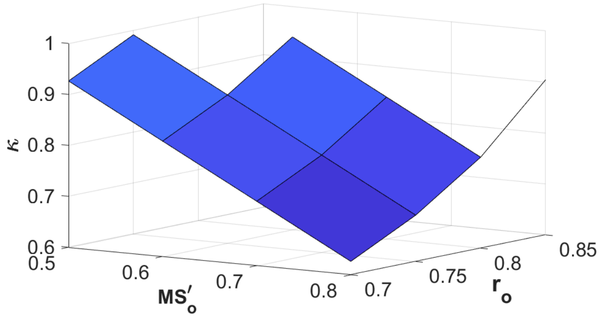
- In spite of the market growth rate, we can see from each plot of Figure 4 that (i) the higher the market share MS_i of an FL-PT i, the larger the required



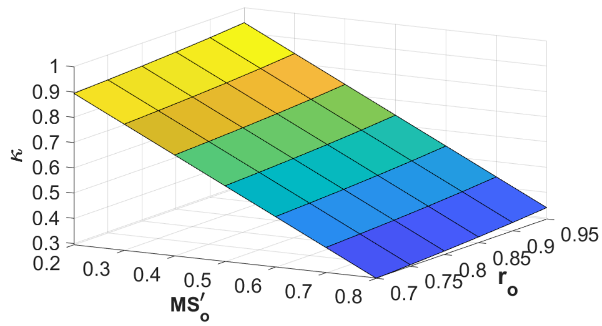
(a) Mature market, $n' = 1$



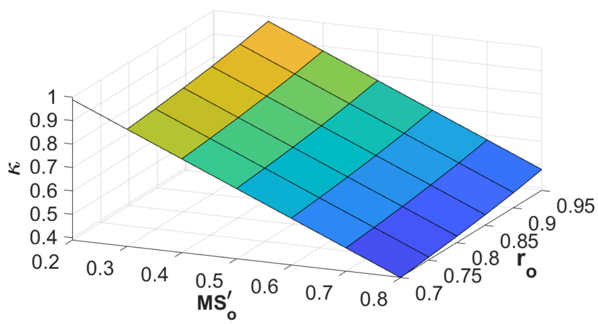
(b) Mature market, $n' = 2$



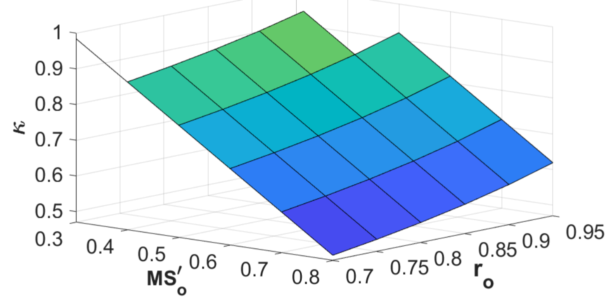
(c) Mature market, $n' = 3$



(d) Fast-growing market, $n' = 1$



(e) Fast-growing market, $n' = 2$



(f) Fast-growing market, $n' = 3$

Fig. 5: The Friendliness κ

value of $Q_i^{\min}$, and (ii) the higher the customer loyalty r_i of an FL-PT, the smaller the required value of $Q_i^{\min}$.

- Under the same market share and customer loyalty (MS_i, r_i), the minimum requirement $Q_i^{\min}$ in a mature market is lower than its counterpart in a fast-growing market.

While designing an FL scheme for FL-PTs in a competitive market, it is important to ensure that model performance improvement achievable for each FL-PT satisfies the minimum requirement $\{Q_i^{\min}\}_{i=1}^n$.

5.3 The Friendliness κ

Overall, the friendliness, κ, in Proposition 3 reflects how accommodating a competitive market is to FL and its meaning is explained when we define it in Definition 2. Let $MS'_o = \sum_{i \in \mathcal{C}'} MS_i$. We set MS'_o to range from 0.2 to 0.8 with a stepsize 0.1. r_o represents the proportion of the loyal customers of all FL-PTs in the original market. Figure 5 illustrates the numerical results in a mature and fast-growing market where (MS'_o, r_o) satisfies $\sum_{i \in \mathcal{C}'} (MS_i - \hat{r}_i) - n'\delta = MS'_o \left(1 - \frac{r_o}{e}\right) > 0$. Given the overall customer leaving rate v_o, a high market growth rate θ leads to a high expanded scale e of the market by Eq. (20). By Proposition 3, the friendliness κ is increasing in the overall customer loyalty r_o and the number n' of sensitive FL-PTs, and decreasing in the total market share MS'_o of sensitive FL-PTs and the market growth rate θ, which are also reflected in the six plots of the figure. Encouragingly, $\kappa \in (0, 1)$ in a wide range of market environments; so is the FL viability:

- In spite of the market growth rate θ and the number n' of sensitive FL-PTs, it is observed from each plot of Figure 5 that $\kappa \in (0, 1)$ under all the feasible pairs (MS'_o, r_o).
- The friendliness κ in a mature market is higher, when the number n' of sensitive FL-PTs, the total market share MS'_o of sensitive FL-PTs, and the overall customer loyalty r_o are the same.
- From each of Figures 5(a)-5(c), it is observed in a mature market that the friendliness κ is especially

high when either the total market share MS'_o of sensitive FL-PTs is small or the overall customer loyalty r_o is high.

- From each of Figures 5(d)-5(f), it is observed in a fast-growing market that (i) if the number n' of sensitive FL-PTs is small (e.g., $n' = 1, 2$), the friendliness κ is especially high when the total market share MS'_o of sensitive FL-PTs is small, in spite of the overall customer loyalty r_o, and (ii) if n' is large (e.g, $n' = 3$), we have the same observation as the mature market above.

6 Conclusions and Future Work

In this paper, we propose an analytical framework to understand the impact of FL on firms' market shares under various market settings. For each FL-PT, we characterize the process by which it joins FL as a non-cooperative game and derive its dominant strategy. For the decision-makers of an FL ecosystem, MarS-FL provides a tight lower bound of the minimum relative ML model performance improvement required by each FL-PT in order to maintain the market δ-stability. MarS-FL also provides a notion of *friendliness* to measure how conducive a market is for FL, and a sufficient and necessary condition for the viability of FL in a competitive market. The results of this paper can guide non-monetary FL incentive mechanisms to allocate model performance improvements among FL-PTs in order to encourage larger data owners to overcome their fear of smaller FL-PTs free-riding on them and join FL.

In future, we will consider designing under the centralized architecture a generic parametric algorithm to determine the allocation of ML model performance while maintaining the market δ-stability. Furthermore, in practice, the related market parameters such as customer loyalty and switching are sometimes estimated with uncertainty [36]. We will leverage the theory of robust optimization [37] to address this. Overall, this paper provides a conceptual framework that serves as a starting point and allows for the further development of FL in the scenarios whenever FL-PTs are in a competitive market [5], [7], [14].

Acknowledgements

This research is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG2-RP-2020-019); the Future Communications Research & Development Programme (FCP-NTU-RG-2021-014); the Joint NTU-WeBank Research Centre on Fintech (Award No: NWJ-2020-008); the Nanyang Assistant Professorship (NAP); and the RIE 2020 Advanced Manufacturing and Engineering (AME) Programmatic Fund (No. A20G8b0102), Singapore; and Future Communications Research & Development Programme (FCP-NTU-RG-2021-014). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

References

[1] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings *et al.*, "Advances and open problems in federated learning," *Foundations and Trends® in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021.

[2] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, pp. 12:1–12:19, 2019.

[3] Y. Cheng, Y. Liu, T. Chen, and Q. Yang, "Federated learning for privacy-preserving ai," *Communications of the ACM*, vol. 63, no. 12, pp. 33–36, 2020.

[4] G. Long, Y. Tan, J. Jiang, and C. Zhang, "Federated learning for open banking," in *Federated Learning: Privacy and Incentive*, Q. Yang, L. Fan, and H. Yu, Eds. Springer, 2020, pp. 240–254.

[5] Q. Yang, "Toward responsible ai: An overview of federated learning for user-centered privacy-preserving computing," *ACM Transactions on Interactive Intelligent Systems*, vol. 11, no. 3-4, pp. 1–22, 2021.

[6] B. Tan, B. Liu, V. Zheng, and Q. Yang, "A federated recommender system for online services," in *Fourteenth ACM Conference on Recommender Systems (RecSys'20)*, 2020, pp. 579–581.

[7] S. Kalloori and S. Klingler, "Horizontal cross-silo federated recommender systems," in *Fifteenth ACM Conference on Recommender Systems*, 2021, pp. 680–684.

[8] Y. Liu, A. Huang, Y. Luo, H. Huang, Y. Liu, Y. Chen, L. Feng, T. Chen, H. Yu, and Q. Yang, "Fedvision: An online visual object detection platform powered by federated learning," in *Proceedings of the 32nd Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-20)*, 2020, pp. 13 172–13 179.

[9] M. J. Sheller, B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, D. M. Weilin Xu, R. R. Colen, and S. Bakas, "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data," *Scientific Reports*, vol. 10, no. 12598, pp. doi:10.1038/s41 598–020–69 250–1, 2020.

[10] G. A. Kaissis, M. R. Makowski, D. Rackert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, pp. 305–311, 2020.

[11] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konečný, S. Mazzocchi, B. McMahan *et al.*, "Towards federated learning at scale: System design," *Proceedings of Machine Learning and Systems*, vol. 1, pp. 374–388, 2019.

[12] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6g: Ai empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, 2019.

[13] Z. Yang, M. Chen, K.-K. Wong, H. V. Poor, and S. Cui, "Federated learning for 6g: Applications, challenges, and opportunities," *arXiv preprint arXiv:2101.01338*, 2021.

[14] Y. Zhan, J. Zhang, Z. Hong, L. Wu, P. Li, and S. Guo, "A survey of incentive mechanism design for federated learning," *IEEE Transactions on Emerging Topics in Computing*, p. doi:10.1109/TETC.2021.3063517, 2021.

[15] L. Lyu, X. Xu, Q. Wang, and H. Yu, "Collaborative fairness in federated learning," in *Federated Learning*, Q. Yang, L. Fan, and H. Yu, Eds. Springer, 2020, pp. 189–204.

[16] T. Song, Y. Tong, and S. Wei, "Profit allocation for federated learning," in *Proceedings of the 2019 IEEE International Conference on Big Data (BigData'19)*, 2019, pp. 2577–2586.

[17] Y. Sarikaya and O. Ercetin, "Motivating workers in federated learning: A stackelberg game perspective," *IEEE Networking Letters*, vol. 2, no. 1, pp. 23–27, 2019.

[18] J. Zhang, Y. Wu, and R. Pan, "Incentive mechanism for horizontal federated learning based on reputation and reverse auction," in *Proceedings of the 2021 ACM Web Conference (WWW'21)*, 2021, pp. 947–956.

[19] Y. Zhan, P. Li, Z. Qu, D. Zeng, and S. Guo, "A learning-based incentive mechanism for federated learning," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6360–6368, 2020.

[20] H. Yu, Z. Liu, Y. Liu, T. Chen, M. Cong, X. Weng, D. Niyato, and Q. Yang, "A fairness-aware incentive scheme for federated learning," in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES'20)*, 2020, pp. 393–399.

[21] ——, "A sustainable incentive scheme for federated learning," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 58–69, 2020.

[22] R. Zeng, S. Zhang, J. Wang, and X. Chu, "Fmore: An incentive scheme of multi-dimensional auction for federated learning in

mec," in *Proceedings of the IEEE 40th International Conference on Distributed Computing Systems (ICDCS'20)*, 2020, pp. 278–288.

[23] M. Tang and V. W. Wong, "An incentive mechanism for cross-silo federated learning: A public goods perspective," in *Proceedings of the IEEE 40th Conference on Computer Communications (INFOCOM'21)*, 2021, pp. 1–10.

[24] L. Lyu, J. Yu, K. Nandakumar, Y. Li, X. Ma, J. Jin, H. Yu, and K. S. Ng, "Towards fair and privacy-preserving federated deep models," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 11, pp. 2524–2541, 2020.

[25] P. W. Farris, N. Bendle, P. E. Pfeifer, and D. Reibstein, *Marketing metrics: The definitive guide to measuring marketing performance*. Pearson Education, 2010.

[26] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge University Press, 2007.

[27] R. T. Rust and A. J. Zahorik, "Customer satisfaction, customer retention, and market share," *Journal of retailing*, vol. 69, no. 2, pp. 193–215, 1993.

[28] J. Chen, Q. Hu, and H. Jiang, "Social welfare maximization in cross-silo federated learning," in *Proceedings of the 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'22)*. IEEE, 2022, pp. 4258–4262.

[29] X. Xu, L. Lyu, X. Ma, C. Miao, C. S. Foo, and B. K. H. Low, "Gradient driven rewards to guarantee fairness in collaborative machine learning," *Advances in Neural Information Processing Systems*, vol. 34, pp. 16 104–16 117, 2021.

[30] N. Zhang, Q. Ma, and X. Chen, "Enabling long-term cooperation in cross-silo federated learning: A repeated game perspective," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2022.

[31] X. Xu and L. Lyu, "Towards building a robust and fair federated learning system," *arXiv preprint arXiv:2011.10464*, 2020.

[32] Y. Shi, H. Yu, and C. Leung, "Towards fairness-aware federated learning," *arXiv preprint arXiv:2111.01872v2*, 2022.

[33] X. Tu, K. Zhu, N. C. Luong, D. Niyato, Y. Zhang, and J. Li, "Incentive mechanisms for federated learning: From economic and game theoretic perspective," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2022.

[34] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS'17)*, 2017, pp. 1273–1282.

[35] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, "On the convergence of fedavg on non-iid data," in *International Conference on Learning Representations (ICLR'20)*, 2020, pp. 1–16.

[36] W. Verbeke, D. Martens, C. Mues, and B. Baesens, "Building comprehensible customer churn prediction models with advanced rule induction techniques," *Expert systems with applications*, vol. 38, no. 3, pp. 2354–2364, 2011.

[37] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM review*, vol. 53, no. 3, pp. 464–501, 2011.

Xiaohu Wu received his PhD from Telecom Paris/Institut Polytechnique de Paris, France in 2016, in computer science. Recently, he has been a research fellow in Nanyang Technological University, Singapore, and will be joining Beijing University of Posts and Telecommunications as an associate professor. Before, he had held research or visiting positions at Fondazione Bruno Kessler (Trento, Italy), Aalto University (Espoo, Finland), University of California at Berkeley, and Singapore University of Technology and Design. His research interests include federated learning, network economics, modeling and performance analysis, information security and so on.

Han Yu is a Nanyang Assistant Professor (NAP) in the School of Computer Science and Engineering (SCSE), Nanyang Technological University (NTU), Singapore. He held the prestigious Lee Kuan Yew Post-Doctoral Fellowship (LKY PDF) from 2015 to 2018. He obtained his PhD from the School of Computer Science and Engineering, NTU. His research focuses on federated learning and algorithmic fairness. He has published over 150 research papers and book chapters in leading international conferences and journals. He is a co-author of the book *Federated Learning* - the first monograph on the topic of federated learning. His research works have won multiple awards from conferences and journals.