Appendix: Contract Design for Adaptive Federated Learning

Ni YANG, Yue CUI, Man Hon Cheung

April 11, 2024

1 Proof to Lemma 1

Proof. For a contract $\phi_{i,j}$, if the server chooses a smaller payment $\underline{p_{i,j}} < p_i^*$, the IR constraint is not satisfied. If the server chooses a larger payment $\overline{p_{i,j}} > p_{i,j}^*$, the cost will be higher as $\frac{\partial C(\boldsymbol{r},\boldsymbol{p})}{\partial p_{i,j}} > 0$. Thus, $p_{i,j}^* = \theta_i r_{i,j}$. □

2 Proof to Lemma 2

Proof.

1. The non-negative weighted sum of convex (concave) functions is convex (concave). [1] $g_1(\boldsymbol{r}) = \sum_{i=1}^{I}\sum_{j=1}^{J} N_{i,j} r_{i,j}\gamma_j$ and $g_2(\boldsymbol{r}) = \sum_{i=1}^{I}\sum_{j=1}^{J} N_{i,j}\theta_i r_{i,j}$ are weighted sum of variables $r_{i,j}$, where $N_{i,j}$, γ_j and θ_i are weights. Thus, $g_1(\boldsymbol{r})$ and $g_2(\boldsymbol{r})$ are both convex and concave.
2. The scalar composition $f(x) = h(g(x))$ is convex if $g(\cdot)$ is concave, $h(\cdot)$ is convex and $\hat{h}$ is non-increasing, where $\hat{h}$ is the extended-value extension of function h, which assigns the value ∞ to points not in **dom** h. Suppose $h(x) = \frac{1}{\sqrt{x}}$, which is a convex and $\hat{h}$ is non-increasing. Thus, $f(\boldsymbol{r}) = h(g_1(\boldsymbol{r}))$ is convex.
3. $C_{com}(\boldsymbol{r})$ is non-negative weighted sum of convex functions $f(\boldsymbol{r})$ and $g_2(\boldsymbol{r})$. Thus, $C_{com}(\boldsymbol{r})$ is convex.

□

3 Proof to Lemma 3

Proof.

1. Condition 1 is derived from IR constraints. As $\theta_I > \theta_i$, $p_{I,j} - \theta_i r_{I,j} > p_{I,j} - \theta_I r_{I,j} \geq 0$ for $i \in \mathcal{I}\backslash\{I\}$. This condition ensures type-$\{i,j\}, \forall i \in \mathcal{I}$ clients at least receive a non-negative payoff by accepting contract $\phi_{I,j}$.
2. Condition 2 is derived from IC constraints. It reveals the order of required resources and payment. The proof consists of three steps:
   (a) First, we prove that $r_{i,j} \geq r_{i',j}$ if and only if $p_{i,j} \geq p_{i',j}$, $i, i' \in \mathcal{I}$ and $i \neq i'$.
      i. Forward direction: $r_{i,j} \geq r_{i',j} \Rightarrow p_{i,j} \geq p_{i',j}$. IC constraints for type-$\{i,j\}$ clients:
      $p_{i,j} - \theta_i r_{i,j} \geq p_{i',j} - \theta_i r_{i',j}$
      $\Leftrightarrow \theta_i(r_{i',j} - r_{i,j}) \geq p_{i',j} - p_{i,j}$.
      Thus, if we have $r_{i,j} \geq r_{i',j}$, then $0 \geq \theta_i(r_{i,j} \geq r_{i',j}) \geq p_{i',j} - p_{i,j}$, which means that $p_{i,j} \geq p_{i',j}$.
      ii. Backward direction: $p_{i,j} \geq p_{i',j} \Rightarrow r_{i,j} \geq r_{i',j}$. IC constraints for type-$\{i',j\}$ clients:
      $p_{i',j} - \theta_{i'} r_{i',j} \geq p_{i,j} - \theta_{i'} r_{i,j}$
      $\Leftrightarrow \theta_{i'}(r_{i,j} - r_{i',j}) \geq p_{i,j} - p_{i',j}$.
      Thus, if we have $p_{i,j} \geq p_{i',j}$, then $\theta_{i'}(r_{i,j} - r_{i',j}) \geq p_{i,j} - p_{i',j} \geq 0$, which tells that $r_{i,j} \geq r_{i',j}$.
   (b) Then, for type-$\{i',j\}$ and type-$\{i,j\}$, we will prove that when $\theta_i < \theta_{i'}$, $r_{i,j} \geq r_{i',j}$. We prove it through contradiction. Suppose that there exists the case where $\theta_i < \theta_{i'}$ and $r_{i,j} < r_{i',j}$. Then $(\theta_{i'} - \theta_i)(r_{i',j} - r_{i,j}) > 0$ holds. According to the IC constraints, we have

$$\begin{cases} p_{i,j} - \theta_i r_{i,j} \geq p_{i',j} - \theta_i r_{i',j}, \\ p_{i',j} - \theta_{i'} r_{i',j} \geq p_{i,j} - \theta_{i'} r_{i,j}. \end{cases}$$

   After adding them up, we have $(\theta_{i'} - \theta_i)(r_{i',j} - r_{i,j}) \leq 0$, which is in contradiction with the previous assumption. Thus, when $\theta_i < \theta_{i'}$, $r_{i,j} \geq r_{i',j}$

(c) Combining the two steps together, we can derive condition 2.

3. Condition 3 is derived by the IC constraint of type-$\{i, j\}$ and $\{i', j\}$ client, which ensures that the adjacent type of clients in group-γ_j will not choose another's contract. This property can be applied to all types in group-γ_j

$$\begin{cases} p_{i,j} - \theta_i r_{i,j} \geq p_{i',j} - \theta_i r_{i',j} \\ p_{i',j} - \theta_{i'} r_{i',j} \geq p_{i,j} - \theta_{i'} r_{i,j} \end{cases} \Leftrightarrow \text{Condition 3}$$

□

4 Proof to Lemma 4

Proof. As $\frac{\partial C(\boldsymbol{r},\boldsymbol{p})}{\partial p_{i,j}} > 0$, the server will make the payment as small as possible.

1. For type-$\{I, j\}$ clients, to satisfy condition 1 in Lemma 3, it is optimal to make $p^*_{I,j} = \theta_I r_{I,j}$.

2. For type-$\{i, j\}$ clients, $i \in \{1, \ldots, I-1\}$

   (a) Condition 3 in Lemma 3 gives the bounds of $p_{i,j}$. Given $p^*_{I,j}$, the server will choose the lower bound as $p^*_{I-1,j} = \theta_{I-1} r_{I-1} + r_I(\theta_I - \theta_{I-1})$.

   (b) Applying the same operation to type-$\{i, j\}$, $i \in \{1, \ldots, I-2\}$ we can conclude that $p^*_{i,j} = \theta_i r_{i,j} + \sum_{l=i+1}^{I} r_{l,j}(\theta_l - \theta_{l-1})$ for $i \in \{1, \ldots, I-1\}$.

□

5 Proof to Lemma 5

Proof. The proof is similar to the proof to Lemma 2. The only difference is that the second dominating term is $g_3(\boldsymbol{r}) = \sum_{i=1}^{I} \sum_{j=1}^{J} N_{i,j}(\theta_i r_{i,j} + \sum_{l=i+1}^{I} r_{l,j}(\theta_l - \theta_{l-1}))$, rather than $g_2(\boldsymbol{r})$. However, $g_3(\boldsymbol{r})$ is still a non-negative weighted sum of variables $r_{i,j}$. Thus, $C_{in}(\boldsymbol{r})$ is convex in $\boldsymbol{r}$. □

6 Proof to Theorem 2

Proof. The KKT conditions under the incomplete information and the complete information only differ in the stationarity condition, while the other three conditions are the same.
The stationarity condition for Problem 2 is $\lambda^*_{i,j} = \alpha_2[N_{i,j}\theta_i + \sum_{l=1}^{i-1} N_{l,j}(\theta_i - \theta_{i-1})] - \frac{\alpha_1 N_{i,j}\gamma_j}{2(\sum_{i=1}^{I}\sum_{j=1}^{J}\gamma_j N_{i,j} r^*_{i,j})^{\frac{3}{2}}}$.
Suppose

$$A_{i,j} = \begin{cases} \frac{\gamma_j}{\theta_i}, & \text{when } i = 1; \\ \frac{N_{i,j}\gamma_j}{N_{i,j}\theta_i + \sum_{l=1}^{i-1} N_{l,j}(\theta_i - \theta_{i-1})}, & \text{when } i \in \{2, \ldots, I\}. \end{cases}$$

$\Rightarrow$ DF: $(\sum_{i=1}^{I}\sum_{j=1}^{J}\gamma_j N_{i,j} r^*_{i,j})^{\frac{3}{2}} \geq \frac{\alpha_1}{2\alpha_2} \cdot A_{i,j}$
We will prove that $A_{1,1} = \arg\max_{i \in \mathcal{I}, j \in \mathcal{J}}\{A_{i,j}\}$.

1. For any specific $j \in \mathcal{J}$ and $i \in \{2, \ldots, I\}$, as $\theta_1 < \theta_i$, $\sum_{l=1}^{i-1} N_{l,j}(\theta_i - \theta_{i-1}) > 0$. Thus $A_{1,j} = \frac{\gamma_j}{\theta_1} = \frac{N_{i,j}\gamma_j}{N_{i,j}\theta_1} > A_{i,j} = \frac{N_{i,j}\gamma_j}{N_{i,j}\theta_i + \sum_{l=1}^{i-1} N_{l,j}(\theta_i - \theta_{i-1})}$.

2. For any $j \in \{2, \ldots, J\}$, as $\gamma_1 > \gamma_j$, $A_{1,1} = \frac{\gamma_1}{\theta_1} > A_{1,j} = \frac{\gamma_j}{\theta_1}$

3. Thus, $A_{1,1} = \arg\max_{i \in \mathcal{I}, j \in \mathcal{J}}\{A_{i,j}\}$

Then, following the same deduction in proof of Theorem 1, we can get the same result. □

References

[1] S. P. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.