A Robust Game-theoretical Federated Learning Framework with Joint Differential Privacy
一种基于联合差分隐私的博弈联邦学习框架

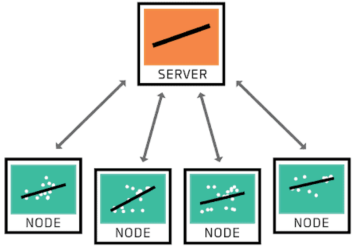
汇报人：黄其涵　　导师：章静教授

IEEE Transactions on Knowledge and Data Engineering, 2022

2023 年 3 月 2 日

目录

1.1 Problem



现有联邦学习框架中的问题：

- **对于客户端的激励通常被大多数联邦学习框架所忽视。**激励是指增加客户参加训练后的奖励或收益，使得他们能够更加积极地贡献数据参与训练；
- **恶意参与者试图破坏联邦学习的全局模型。**联邦学习通常容易受到敌对客户的操纵，尤其是关于模型和数据投毒攻击，即恶意参与者在训练集中引入大量带有修改标签的数据样本，导致全局模型的错误分类。
- **私人数据仍然可以通过从参数更新中推断信息来泄露。**因此，针对中毒攻击的最佳安全策略还应该通过推理提供针对隐私泄露的保护。

1.2 Contributions

- 提出了一种基于联合差分隐私的联邦学习框架，可限制攻击者的影响并提供严格的隐私保证。
- 提出了两种基于真实客户端数据的博弈论机制来选择参与的客户端进行训练。
- 对于提出机制进行理论证明和分析。客户端选择机制满足博弈论属性，并且对客户端的策略操纵具有鲁棒性。
- 进行了关于所设计机制的基本原理实验。使用真实世界数据集进行实验，证明了所提出的联邦学习方案的实际有效性。

目录

2 Preliminaries

2.1 Game Theory

博弈论，又称为对策论（Game Theory）、赛局理论等，既是现代数学的一个新分支，也是运筹学的一个重要学科。博弈论主要研究公式化了的激励结构间的相互作用，是研究具有斗争或竞争性质现象的数学理论和方法。博弈论考虑游戏中的个体的预测行为和实际行为，并研究它们的优化策略。**在联邦学习中，服务器需要招募一些客户端来执行训练任务。但是，由于每个客户端的数据集和计算能力不同，他们参与训练的精力和时间成本等也各不相同。**

在文本中：
我们的框架将客户选择过程视为拍卖游戏，客户提交他们的成本 c_i 作为出价，服务器根据其预算 B 决定赢家和付款 p_i。如果服务器选择了客户，则其效用为 $u_i = p_i$-c_i，**否则** $u_i = 0$。
假定客户是理性利己的，每个客户只关心自己的效用并试图最大化它（尽管有可能得到虚假信息）。服务器被激励以其有限的预算 B 吸引尽可能多的参与客户。

2.2 Definitions

具体来说，客户选择机制应该满足以下博弈论定义:

Definition 1 (Dominant-strategy truthful)

A client selection mechanism $\mathcal{M}$ is dominant-strategy truthful if, for every client, truthfully reporting his cost c_i is a dominant strategy. That is, given c_i and all possible $c_i' \in \mathbb{R}$:

$$u_i\left(c_i, c_{-i}\right) \geq u_i\left(c_i', c_{-i}\right),$$

where c_{-i} represents the strategy profile of clients except for the i th component, $c = (c_i, c_{-i})$.

真实意味着没有客户可以通过歪曲数据来提高他们的效用，这是开发博弈论解决方案的一个重要属性。

2.2 Definitions

Definition 2 (Individual rationality)

A mechanism $\mathcal{M}$ is individual rational for each client if the clients gain nonnegative utility by participating in $\mathcal{M}$.

$$u_i = p_i - c_i \geq 0.$$

个人理性保证了每个客户的基本效用。如果参与一种机制可能导致负效用，客户可能会完全拒绝参与。

Definition 3 (Budget balance)

A client selection mechanism $\mathcal{M}$ is budget-balanced if the total amount paid by the server does not exceed its budget. Formally, given m selected clients, each receives a payment p_i, then

$$\sum_{i=1}^{m} p_i \leq B$$

最后，从服务器的角度来看，其总支出必须保持在其预定义的预算范围内。

目录

3 Methodology

3.1 Overview of the mechanism

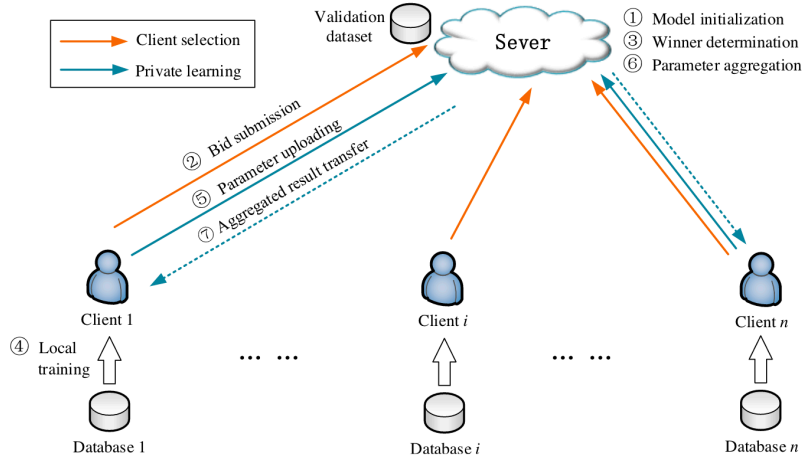


Fig. 1. Our jointly differentially private federated learning framework

3.2 Alogrithm 1 - Step③ Winner Determination

Algorithm 1 Truthful client selection mechanism $\mathcal{M}_1$

Input: cost c_i from each client, the server's total budget B .
Output: winning clients for FL training, the payment p_i for each client.

1: **for** $i \leq n$ **do**
2: $q_i \leftarrow \frac{c_i}{d_i}$.
3: **end for**
4: **sort** q_i in an increasing order, $q_1 \leq q_2 \leq \cdots \leq q_n$, breaking ties arbitrarily.
5: **find** the largest $m \in [n]$ that satisfies $q_1 \leq q_2 \leq \cdots \leq q_m$ and $q_m \leq \frac{B}{\sum_{i \in S} d_i}$, breaking ties arbitrarily.
6: **for** $1 \leq i \leq m$ **do**
7: $p_i \leftarrow \min\{\frac{B}{\sum_{i \in S} d_i}, q_{m+1}\} \cdot d_i$.
8: **end for**
9: **for** $m < i \leq n$ **do**
10: $p_i \leftarrow 0$.
11: **end for**

c_i：训练成本（能源消耗、通信成本和训练数据集的估值来评估成本）
d_i：数据量
q_i：单位数据消耗
$\frac{B}{\sum_{i \in S} d_i}$：平均单位数据消耗
p_i：选择该模型需要的支出

3.2 Alogrithm 2 - Step③ Winner Determination

Algorithm 2 Truthful client selection mechanism $\mathcal{M}_2$

Input: cost c_i from each client, the server's total budget B .
Output: winning clients for FL training, the payment p_i for each client.

for $i \leq n$ **do**
  $r_i \leftarrow \frac{d_i}{c_i}$.
end for
sort r_i in an decreasing order, $r_1 \geq r_2 \geq \cdots \geq r_n$, breaking ties arbitrarily.
$paid \leftarrow 0$, $selected \leftarrow \emptyset$, $i = 1$.
while $paid + c_i \leq B$ **do**
  $selected \leftarrow selected \cup \{i\}$.
  $paid \leftarrow paid + c_i$.
  $i \leftarrow i + 1$.
end while
if $\sum_{j \in selected} d_j \geq d_{j+1}$ **then**
  Output $selected$.
else
  Output $\{i^*\}$ that satisfies $i^* = \arg\max_i d_i$.
end if
for $i \in selected$ **do**
  $p_i \leftarrow \int_0^{c_i} z \cdot \frac{\mathrm{d}}{\mathrm{d}z} \mathsf{alloc}_i(z, c_{-i}) \mathrm{d}z$.
end for

Knapsack Problems:

$$\max \sum_{i=1}^{n} x_i \cdot d_i,$$

$$\text{s.t.} \sum_{i=1}^{n} x_i \cdot c_i \leq B, \forall i, x_i \in \{0, 1\}.$$

是典型的 0/1 背包问题, 其中:
d_i 代表每个客户的个性化权重值
c_i 是相应的消耗
x_i 是带不带它，取值为 0 或 1
r_i：ranking，其实就是按效率的排序
p_i：选择该模型需要的支出

3.2 Review

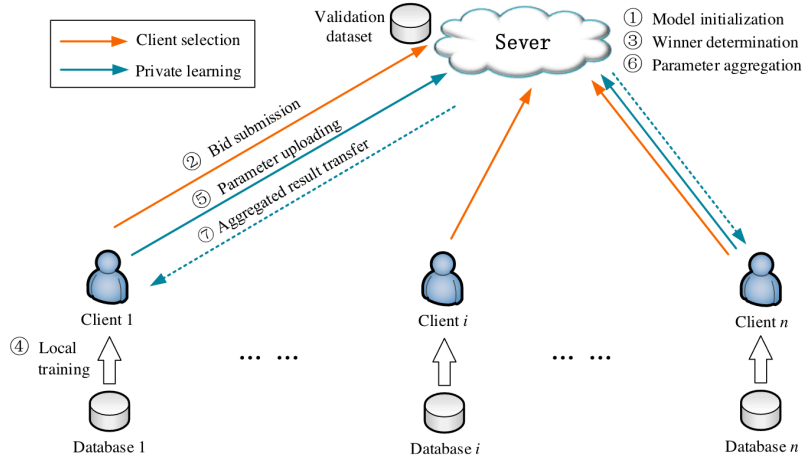


Fig. 1. Our jointly differentially private federated learning framework

3.3 Alogrithm 3 - Step⑤ Parameter Uploading

Algorithm 3 Local perturbation mechanism $\mathcal{M}_3$

Input: the private data θ_i of client i, privacy budget ϵ_L, sensitivity $\Delta_2 f$, clipping threshold C.

Output: perturbed data $\hat{\theta}_i$.

1: sample $Z_i \sim \mathcal{N}(0, \sigma)$, where $\sigma = \sqrt{2 \ln \frac{1.25}{\delta}} \cdot \frac{\Delta_2 f}{\epsilon_L}$

2: $\theta_i \leftarrow \theta_i / \max\{1, \frac{\theta_i}{C}\}$.

3: $\hat{\theta}_i \leftarrow \theta_i + Z_i$.

4: **output** $\hat{\theta}_i$.

DP-SGD

θ_i：需要扰动的参数

Z_i：高斯噪声

C：梯度范数阈值

p_i：选择该模型需要的支出

ϵ：隐私预算

δ：松弛系数

3.4 Alogrithm 4 - Step⑥ Parameter Aggregation

Algorithm 4 Parameter aggregation mechanism $\mathcal{M}_4$

Input: the noisy data $\hat{\theta}_i$ of each client, privacy budget ϵ_E
Output: aggregated data (agg, sum) for each client

construct m client groups $g_i = \{Cj\}_{j \neq i}$.
for each group g_i **do**
  $agg_i \leftarrow \sum_{j \in g_i} d_j \hat{\theta}_j$, $sum_i \leftarrow \sum_{j \in g_i} d_j$.
  $\hat{\theta}_i = agg_i / sum_i$.
end for
for $1 \leq i \leq m$ **do**
  compute $y \leftarrow \mathsf{Acc}_D(\hat{\theta}_i)$.
end for
$\Delta y \leftarrow \frac{1}{m-1}$.
pick up a g_i^* with probability $\propto \frac{\epsilon_E \cdot y(D,g)}{2\Delta y}$.
for $1 \leq i \leq m$ **do**
  send client C_i the aggregated data pair (agg_{i^*}, sum_{i^*}) computed from g_i^*.
end for

$$\theta = \sum_{i=1}^{m} d_i \hat{\theta}_i / \sum_{i=1}^{m} d_i$$

g_i：客户端模型组合，总数其实是 $C_m^{m-1} = m$
agg_i：每组合模型个性化权重值与参数乘积的累和
sum_i：每组合模型个性化权重值的累和
$\hat{\theta}_i$：每组合模型数据量的累和
ACC_D：得分函数 (来自每个客户端组的聚合参数的优秀程度)
Δy：评分函数的灵敏度
g_i^*：以概率选出的客户端模型组合

3.5 Alogrithm 5 - Step⑦ Parameter Result Transfer

Algorithm 5 Jointly differentially private local update $\mathcal{M}_5$

Input: the data pair (agg, sum), C_i's true parameters θ_i
Output: jointly differentially private parameter for client i

1: $\theta_i \leftarrow \frac{agg+d_i\theta_i}{sum+d_i}$.
2: **test** the new parameters θ_i using client i's local dataset.
3: **generate** the parameters for next round of submission.

agg_i：每组合模型个性化权重值与参数乘积的累和
sum_i：每组合模型个性化权重值的累和
d_i 代表每个客户的个性化权重值
θ_i：本地参数（没扰动）

目录

4 Conclusions

Conclusions

4 Conclusions

- 在本文中，提出了一个强大的联邦学习框架，解决了激励客户参与联邦学习的实际问题。
- 其次，提出了两种新颖的博弈论机制，将客户选择制定为拍卖博弈。客户将他们的成本报告为出价，服务器使用不同的支付策略来最大化其目标。整体框架是联合差分隐私的，这限制了敌对客户的影响。
- 此外，我们还对真实世界的数据集进行了模拟，以验证框架的性能。
- 结果表明，在我们的方案下，敌对客户对全局模型的准确性几乎没有影响。

Q&A

感谢您的聆听和反馈