# A Robust Game-Theoretical Federated Learning Framework With Joint Differential Privacy

Huo Mingda

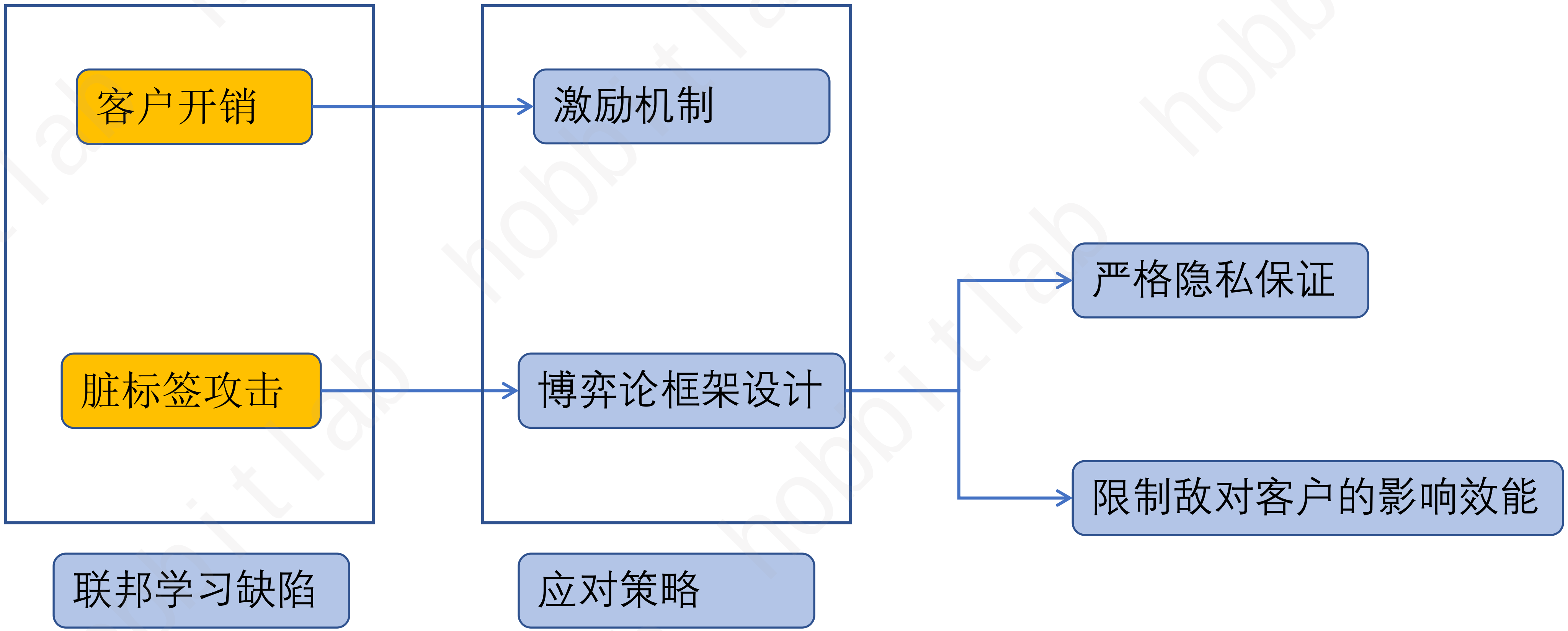
Jinan University, Guangzhou

May 11, 2023

# 研究内容：博弈论和差分隐私的联邦学习框架

Huo Mingda, Jinan University

用户选择机制：两套博弈论机制
用户策略操作（理性攻击）：受制于鲁棒性机制



# 研究内容：差分隐私



## 联邦学习差分隐私流程

1. **本地计算**

   客户端 $i$ 根据本地数据库 $\mathcal{D}_i$ 和接受的服务器的全局模型$w_G^t$作为本地的参数，即$w_i^t = w_G^t$，进行梯度下降策略进行本地模型训练得到$w_i^{t+1}$ （$t$ 表示当前round）。

2. **模型扰动**

   每个客户端产生一个随机噪音$n$，$n$ 是符合高斯分布的，使用$\overline{w_i}^{t+1} = w_i^{t+1} + n$ 扰动本地模型（这里注意$w$是一个矩阵，那么$n$就对矩阵的每一个元素产生噪音）。

3. **模型聚合**

   服务器使用FedAVG算法聚合从客户端收到的$\overline{w_i}^{t+1}$得到新的全局模型参数$w_G^{t+1}$，也就是扰动过的模型参数。

4. **模型广播**

   服务器将新的模型参数广播给每个客户端。

5. **本地模型更新**

   每个客户端接受新的模型参数，重新进行本地计算。

# 研究内容：联邦学习框架设计



1. 定义损失函数

$$F_i(\theta) = \frac{1}{d_i} \sum_{j \in D_i} f_j(\theta),$$

参与培训过程的客户必须找到使给定损失函数最小化的参数 θ

2. 聚合客户参数

$$\theta = \sum_{i=1}^{n} \frac{d_i}{d} \theta_i, \quad d = \sum_{i=1}^{n} d_i.$$

服务器通过计算每个客户端的加权平均值来聚合来自每个客户端的模型参数 θ j

3. 寻找优化最值

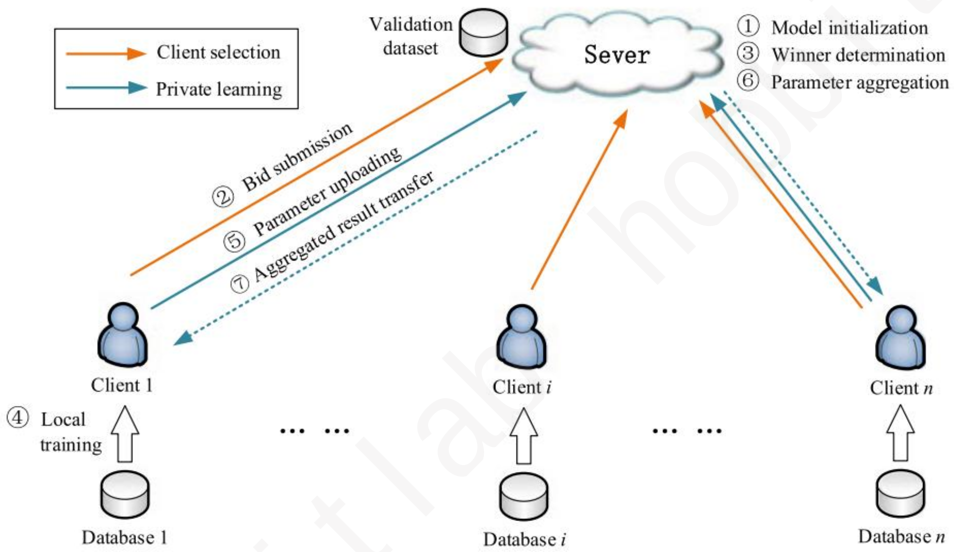
$$\theta^* = \arg\min_{\theta} \frac{d_i}{d} F_i(\theta).$$

学习过程表述为一个优化问题旨在寻找全局损失函数的最小值

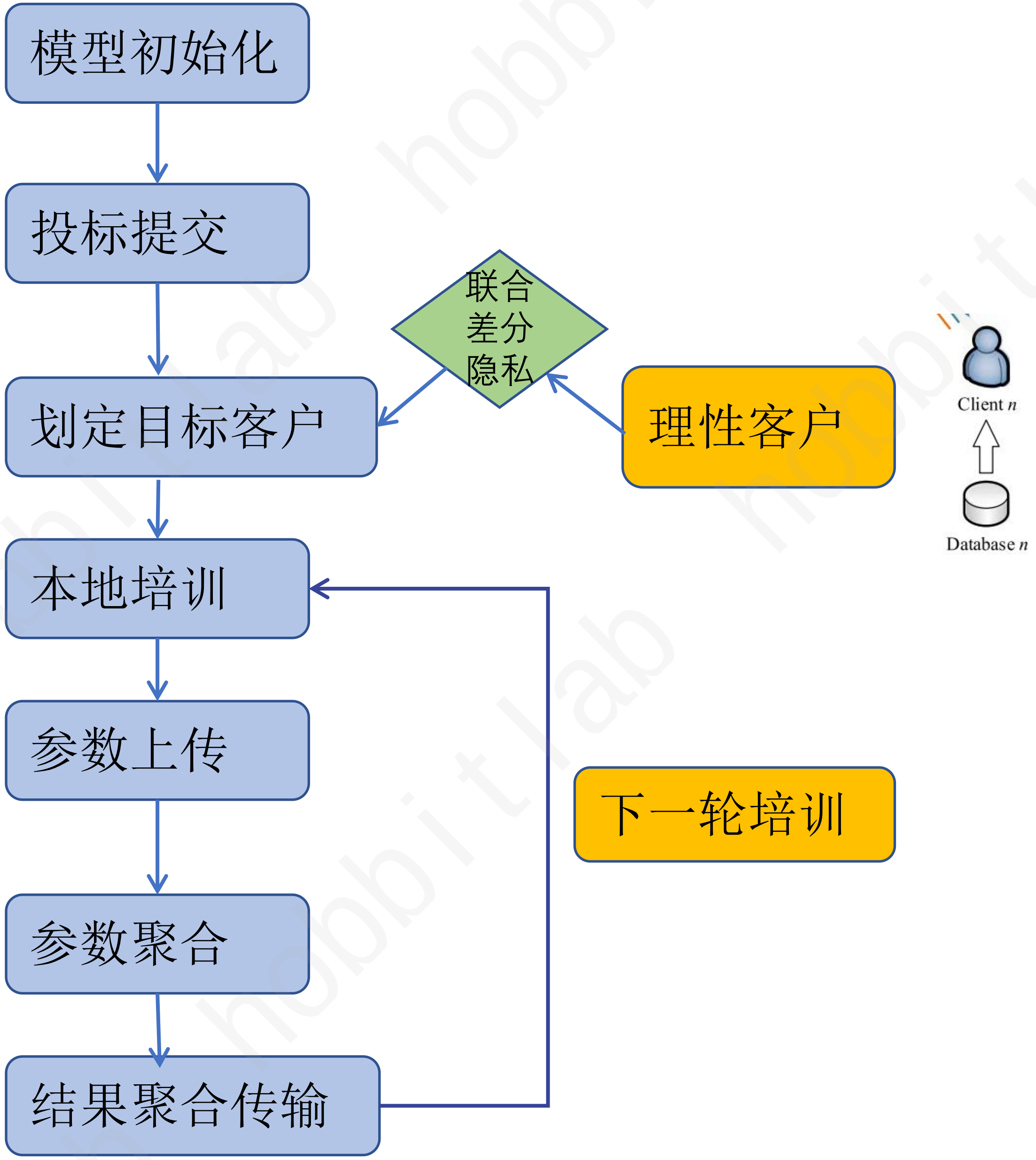
经过足够多轮的训练和参数聚合，优化问题 θ *的解全局收敛到最优值

# 研究设计：联邦学习框架设计

Huo Mingda, Jinan University




拍卖过程



# 研究设计：算法设计-客户选择1



**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.

**for** $i \leq n$ **do**

  $q_i \leftarrow \frac{c_i}{d_i}$.

**end for**

**sort** $q_i$ in an increasing order, $q_1 \leq q_2 \leq \cdots \leq q_n$, breaking ties arbitrarily.

**find** the largest $m \in [n]$ that satisfies $q_1 \leq q_2 \leq \cdots \leq q_m$ and $\boxed{q_m \leq \frac{B}{\sum_{i \in S} d_i}}$ breaking ties arbitrarily.

**for** $1 \leq i \leq m$ **do**

  $p_i \leftarrow \boxed{\min\{\frac{B}{\sum_{i \in S} d_i}, q_{m+1}\} \cdot d_i.}$

**end for**

**for** $m < i \leq n$ **do**

  $p_i \leftarrow 0.$

**end for**

服务器S根据其商品的销售单价从卖方（客户）处购买商品（数据）

找到m最大值

确定中标客户支付额度

放弃为投标失败客户支付

**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**
  $p_i \leftarrow \int_0^{c_i} z \cdot \frac{d}{dz} alloc_i(z, c_{-i}) dz$.
**end for**

符合服务器需求的客户端选择

使用近似方法解决可能存在的np问题

- 步骤1：按照正常的贪婪法求解，得到一个解，设其价值为$C_r$；
- 步骤2：挑选价值最大的物品装入背包，设其价值为$C_s$；
- 步骤3：选择$C_r$、$C_s$两者最大的作为算法的输出。

分配函数描述客户端分配状态

# 研究设计：算法设计-隐私学习



**Algorithm 3.** Local Perturbation Mechanism $\mathcal{M}_3$

**Input:** the private data $\theta_i$ of client $i$, privacy budget $\epsilon_L$, sensitivity $\Delta_2 f$, clipping threshold $C$.
**Output:** perturbed data $\hat{\theta}_i$.

sample $Z_i \sim \mathcal{N}(0, \sigma)$, where $\boxed{\sigma = \sqrt{2 \ln \frac{1.25}{\delta}} \cdot \frac{\Delta_2 f}{\epsilon_L}}$
$\theta_i \leftarrow \theta_i / \max\{1, \frac{\theta_i}{C}\}$.
$\hat{\theta}_i \leftarrow \theta_i + Z_i$.
**output** $\hat{\theta}_i$.

**Algorithm 4.** Parameter Aggregation Mechanism $\mathcal{M}_4$

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

**construct** $m$ client groups $g_i = \{Cj\}_{j \neq i}$.
**for** each group $g_i$ **do**
  $agg_i \leftarrow \sum_{j \in g_i} d_j \theta_j$, $sum_i \leftarrow \sum_{j \in g_i} d_j$.
  $\hat{\theta}_i = agg_i / sum_i$
**end for**
**for** $1 \leq i \leq m$ **do**
  **compute** $y \leftarrow Acc_D(\hat{\theta}_i)$.
**end for**
$\Delta y \leftarrow \frac{1}{m-1}$.
**pick up** a $g_i^*$ with probability $\propto \frac{\epsilon_E y(D,g)}{2\Delta y}$.
**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**

在参数聚合过程中保证联合差异隐私

高斯分布的随机噪声扰动

假设受培训客户户为敌对状态

构造m个客户端组

计算每一组内参数的加权平均值

模型在服务器的验证数据集上评估的准确性

指数概率选择聚合结果输出

受提交参数质量评分限制，对抗性客户端只能以一定的概率影响参数聚合的输出

# 研究设计：算法设计-差分隐私输出



**Algorithm 5.** Jointly Differentially Private Local Update $\mathcal{M}_5$

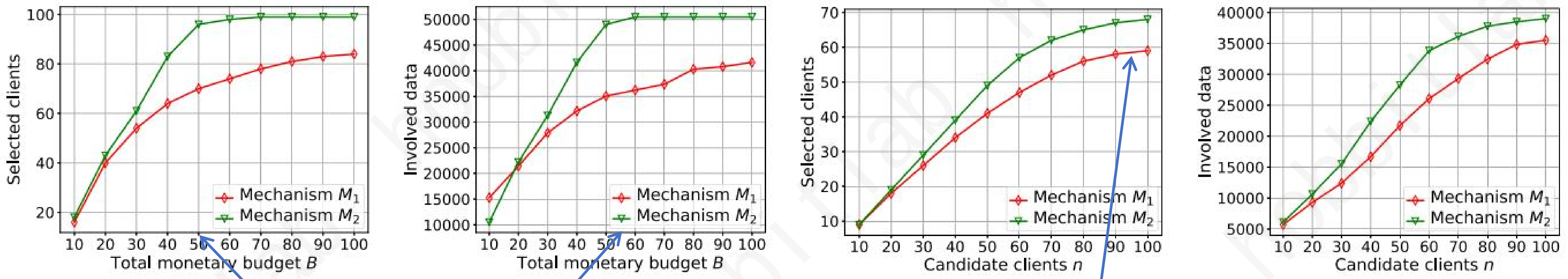
**Input:** the data $\boxed{\text{pair } (agg, sum)}$, $C_i$'s true parameters $\theta_i$
**Output:** jointly differentially private parameter for client $i$
1: $\theta_i \leftarrow \frac{agg + d_i \theta_i}{sum + d_i}$.
2: **test** the new parameters $\theta_i$ using client $i$'s local dataset.
3: **generate** the parameters for next round of submission.

本地数据来测试新的参数

为下一轮培训过程进行参数更新

# 实验对比：MNIST、CIFAR数据集测试

Huo Mingda, Jinan University




(a) The number of selected clients vs. the server's monetary budget

(b) The number of involved data vs. the server's monetary budget

(c) The number of selected clients vs. the number of candidate clients

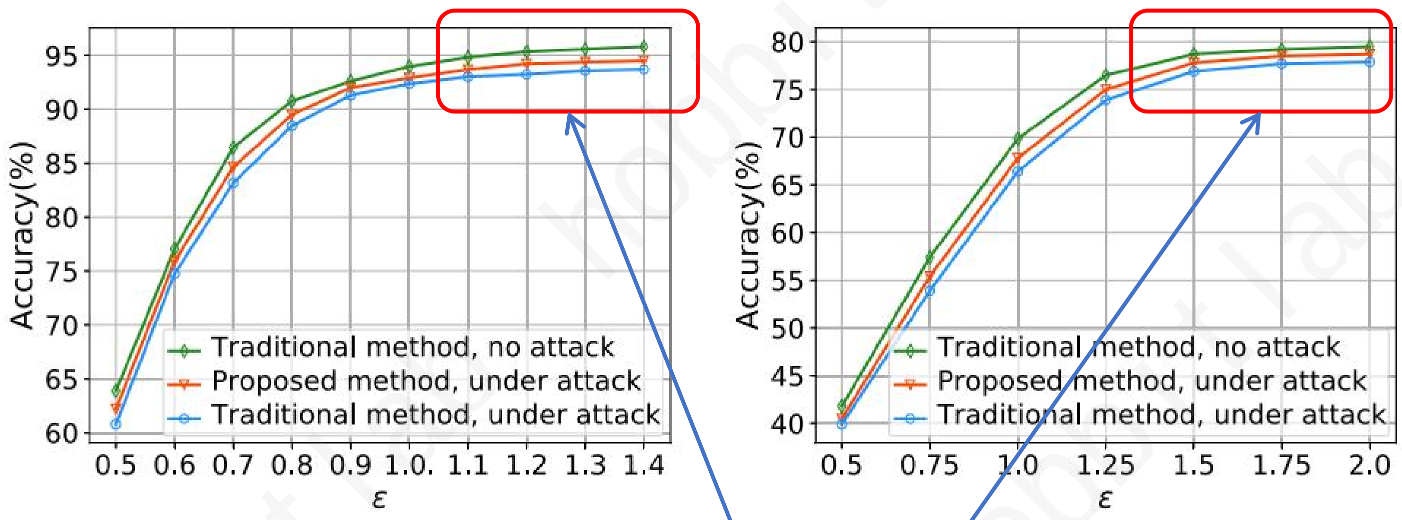
(d) The number of involved data vs. the number of candidate clients

两种算法提供的客户数和数据数对比

由于总货币预算是固定的，因出现了饱和点，由于客户的基本成本，客户数量无法提高

# 实验对比





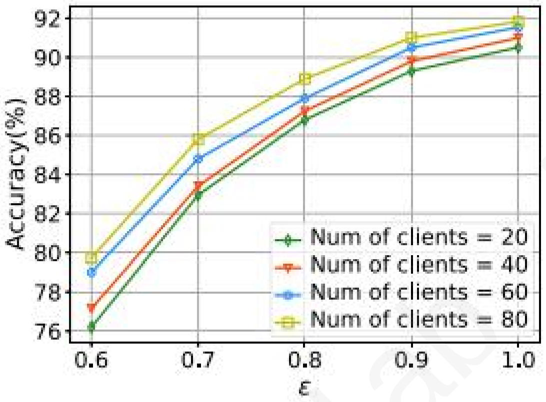

(a) Accuracy of the global model vs. privacy budget, MNIST vs. dataset.

(b) Accuracy of the global model privacy budget, CIFAR-10 dataset.

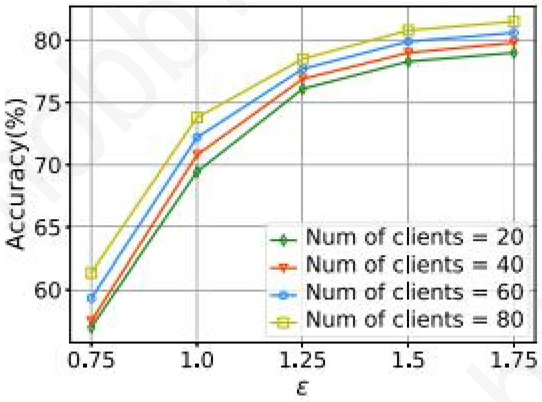
隐私预算上升到一定水平噪声精度影响削弱

# 实验对比： 可解释公平性分析

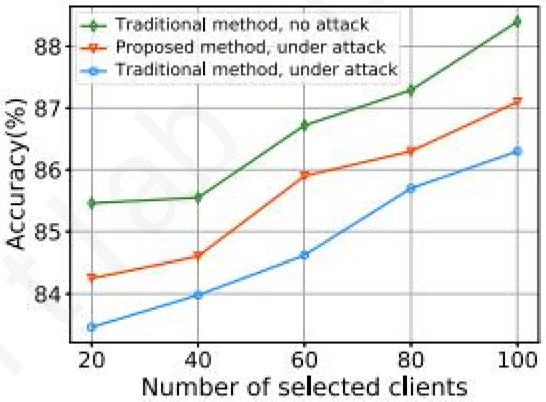




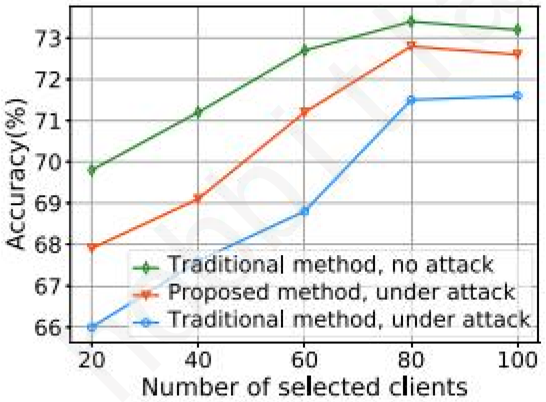

(a) The accuracy of the global model vs. the privacy budget, MNIST dataset.




(b) The accuracy of the global model vs. the privacy budget, CIFAR-10 dataset.




(c) The accuracy of the global model vs. the number of client, MNIST dataset.




(d) The accuracy of the global model vs. the number of client, CIFAR-10 dataset.

Tradeoff：鲁棒性vs.性能

隐私预算低：客户个体相对影响弱

隐私预算高：相对影响强

# Thanks!