Cross-silo Federated Learning with Record-level Personalized Differential Privacy

Junxu Liu
Renmin University of China
Beijing, China
geminiljx@gmail.com

Jian Lou
Zhejiang University
Hangzhou, China
jian.lou@zju.edu.cn

Li Xiong
Emory University
Atlanta, USA
lxiong@emory.edu

Jinfei Liu
Zhejiang University
Hangzhou, China
jinfeiliu@zju.edu.cn

Xiaofeng Meng*
Renmin University of China
Beijing, China
xfmeng@ruc.edu.cn

Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security (CCS '24), October 14–18, 2024, Salt Lake City, UT, USA. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3658644.3670351

差分隐私背景

多个医院合作改进疾病诊断模型，但不能共享患者的原始数据。
采用联邦学习，用户数据不上传到服务器。

差分隐私的作用

1. 保护模型更新：在模型更新过程中**添加随机噪声**。
2. 隐私预算 (ε)：ε 值越小，噪声越大，隐私保护越强，但模型性能可能下降。

如何在保证有效隐私保护的同时，最大限度地提高模型性能，是一个关键的问题。

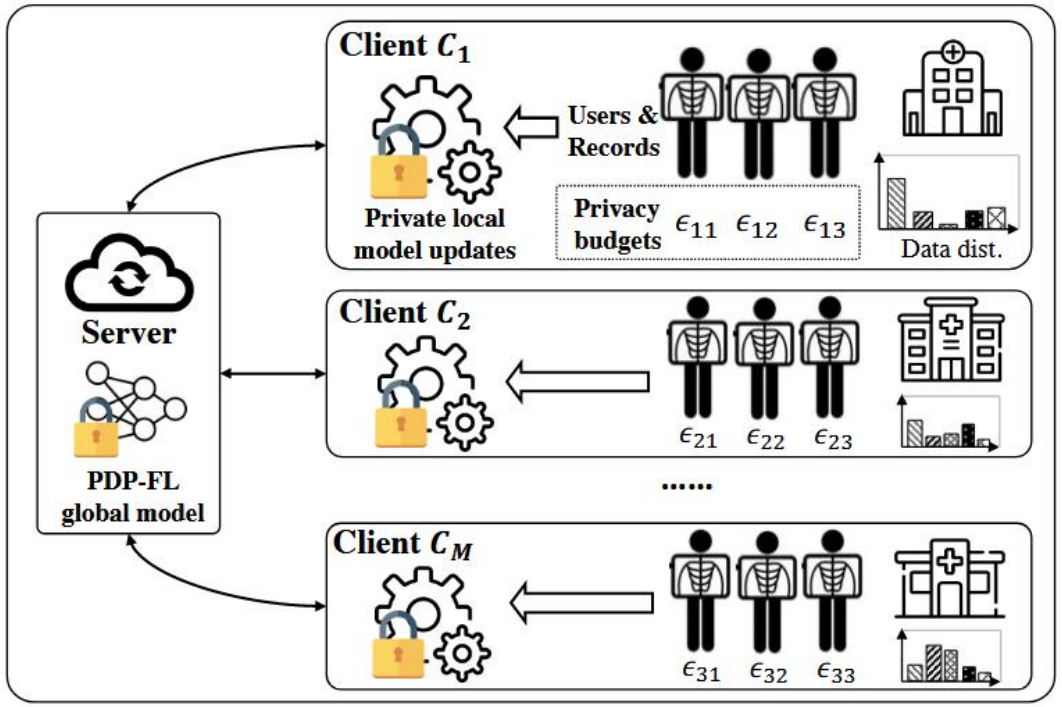
Personalized Differential Privacy

标准DP通过隐私预算 ε 量化隐私保护的程度。
它对每个参与者都施加了**相同的隐私保护**。

这种一致性不能反映人们之间隐私期望差异的现实情况，并且会导致显著的效用成本。

客户端级PDP - FL。

记录级rPDP - FL（本文）。



rPDP的现实背景

假设一个医疗研究项目收集患者的医疗记录，使用差分隐私（DP）技术来保护隐私。
项目中有三个参与者：

1. 参与者A：非常担心隐私泄露，希望得到最大保护，即使研究结果的准确性会受影响。
2. 参与者B：更关心研究成果，愿意接受较低的隐私保护以提高数据效用。
3. 参与者C：医学专业人士，希望保护特定的个人信息，但对总体统计数据的披露不太在意。

标准DP为所有人设置相同的隐私保护强度 ε，无法满足这三者不同的隐私期望。
A觉得保护不够，B觉得限制太多，C则需要更细粒度的保护。

隐私预算分配策略

实现 rPDP 的本质在于确保每个记录在整个训练过程中的累积隐私成本与其预定的隐私预算保持一致。这就强调需要一个**有效的隐私预算分配策略**。

现有方案：

①当某个记录的隐私预算耗尽，该记录将不再参与后续的训练。

②保证最严格的统一(ε, δ) - DP。（全部取最严格的隐私预算要求。）

③Dropout策略，将隐私预算低于阈值的个体丢弃。（隐私预算过于严格的个体不参与训练。）

隐私预算分配策略

现有方案的问题：

①造成灾难性遗忘。（模型在后续的训练中遗忘这些记录中的关键知识，导致模型性能下降。）

②噪声过多，模型性能较差。

③一些隐私预算严格（保守）的个体，不参与训练，造成模型性能下降。

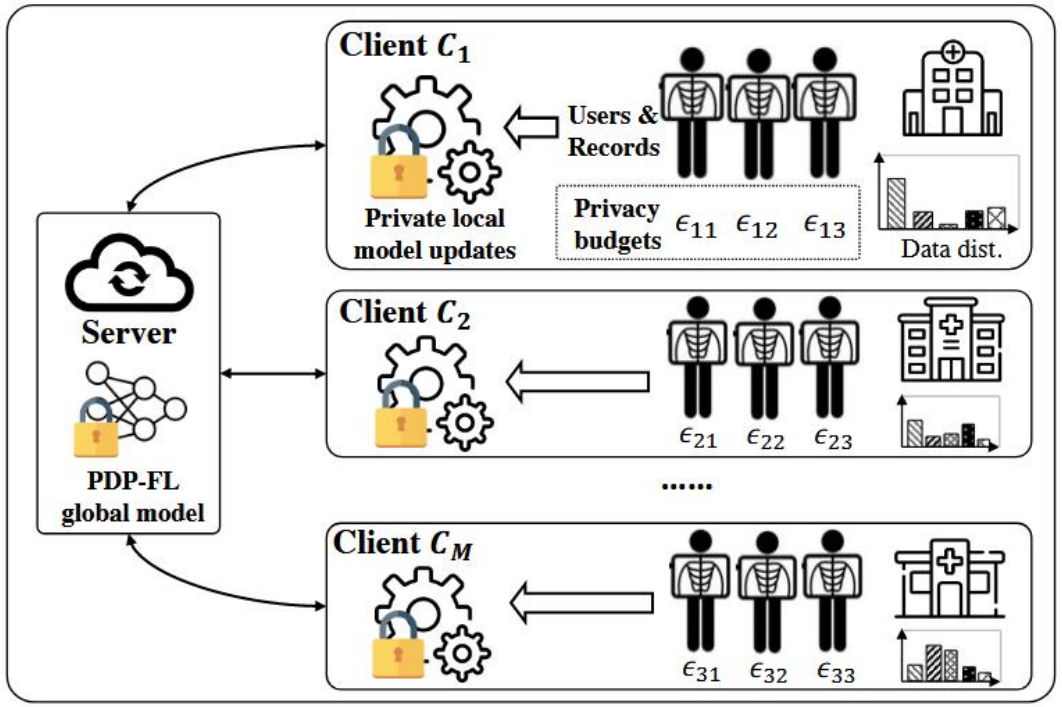
隐私预算分配策略

上述方案，对隐私预算分配不够细致。

理想的分配方案：

每个个体（记录）分配一个合适的隐私预算，使得在训练过程中总隐私预算刚好耗尽。

可以在确保隐私保护的前提下最优化模型性能。



两阶段混合采样方案。

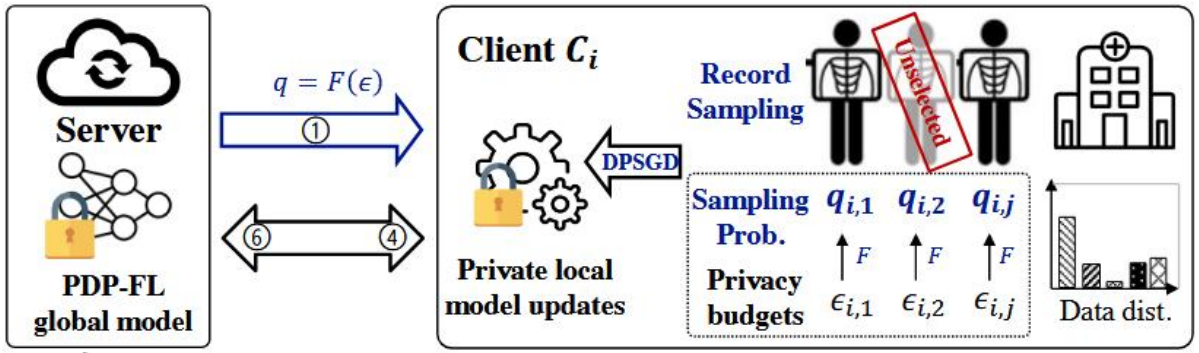
本文开发了一种**非均匀采样机制**，实现所有记录的**隐私预算同时耗尽。**

- 阶段1：客户端泊松采样(均匀)：服务器以均匀的每个客户端采样概率 λ，通过泊松采样选择一个随机的客户端子集 C 。

- 阶段2：记录级泊松采样(非均匀)：每个被选择的客户端在本地进行迭代。

每次迭代，通过泊松采样，以**非均匀的每记录采样概率q，**从局部数据集中抽取小批量数据。

预算较高的记录，采样概率q更高。预算较低的个体，采样概率q更低。



Simulation-CurveFitting策略

问题就变成：

已知单个记录的隐私预算ε，如何确定其采样概率q？

即：如何建立 采样概率q 和 隐私预算ε 之间的对应关系？

Simulation-CurveFitting策略

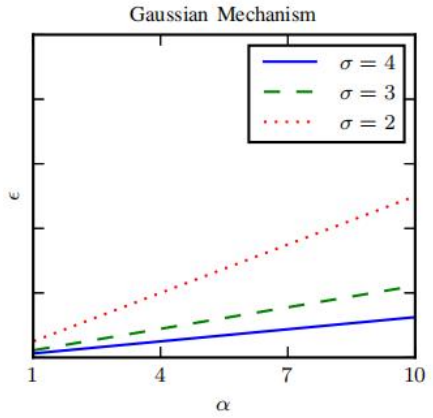
Rényi差分隐私：(α, ρ)-RDP （更加精细的隐私预算度量工具。）

与传统的 (ε, δ) -DP 不同，RDP利用Rényi散度来量化隐私保护。

优点：RDP允许一个连续谱的隐私度量，这一特性使得RDP能够提供更清晰的隐私量化。

Rényi散度公式：

$$D_\alpha(P \parallel Q) \triangleq \frac{1}{\alpha - 1} \log \mathbb{E}_{o \sim Q} \left[\left(\frac{P(o)}{Q(o)} \right)^\alpha \right]$$



从 RDP 转换到 DP：

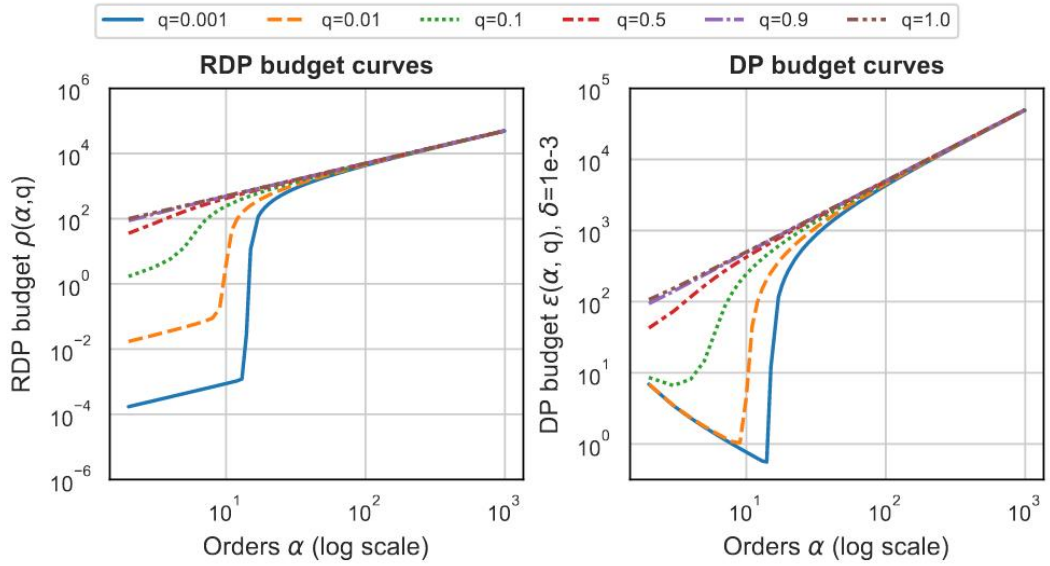
$$\varepsilon = \rho + \frac{\log(1/\delta)}{\alpha - 1}$$

Simulation-CurveFitting策略

取两组候选值：
一组是各种采样概率 q 。
另一组是 (α, ρ)-RDP 的参数 α 。

根据泊松采样高斯机制，可以计算出(α, ρ)-RDP参数 ρ 。
然后从 (α, ρ)-RDP 转换到 (ε, δ) -DP。

最终得到固定采样概率q 下，参数α ——参数ε 之间的对应关系。



$$\rho_{PoiSG}(\alpha, q) \leq \frac{1}{\alpha-1} \log \left\{ (1-q)^{\alpha-1}(\alpha q - q + 1) + \sum_{\ell=2}^{\alpha} \binom{\alpha}{\ell} (1-q)^{\alpha-\ell} q^{\ell} e^{(\ell-1)\rho(\ell)} \right\}. \quad (2)$$

$$\varepsilon = \rho + \frac{\log(1/\delta)}{\alpha - 1}$$

Simulation-CurveFitting策略

固定采样概率q，分配的隐私预算要尽量贴合所需隐私预算的上界。

（尽量做到完成训练时，隐私预算同步耗尽。）

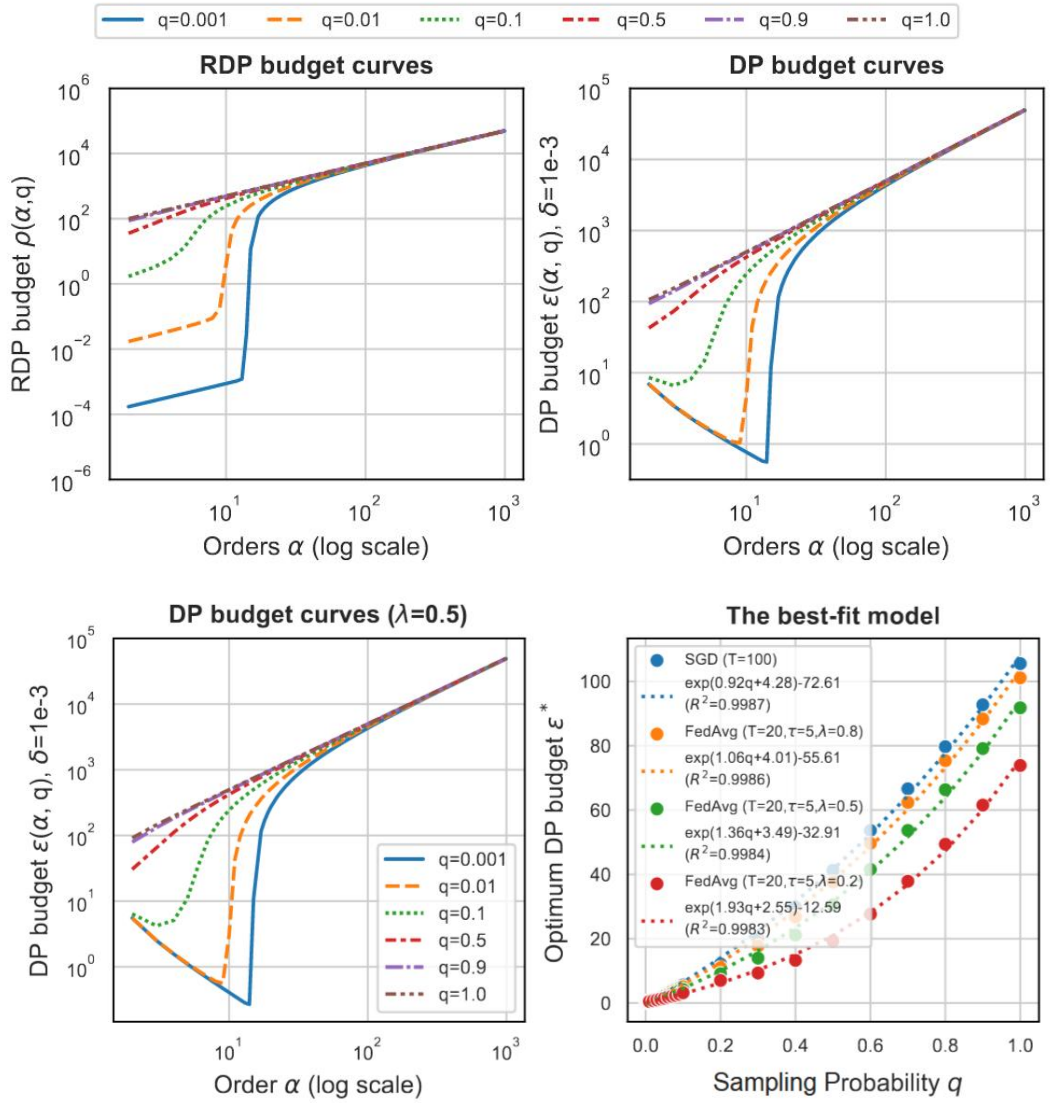
最优隐私预算 ε* 即为图中曲线的最小值。

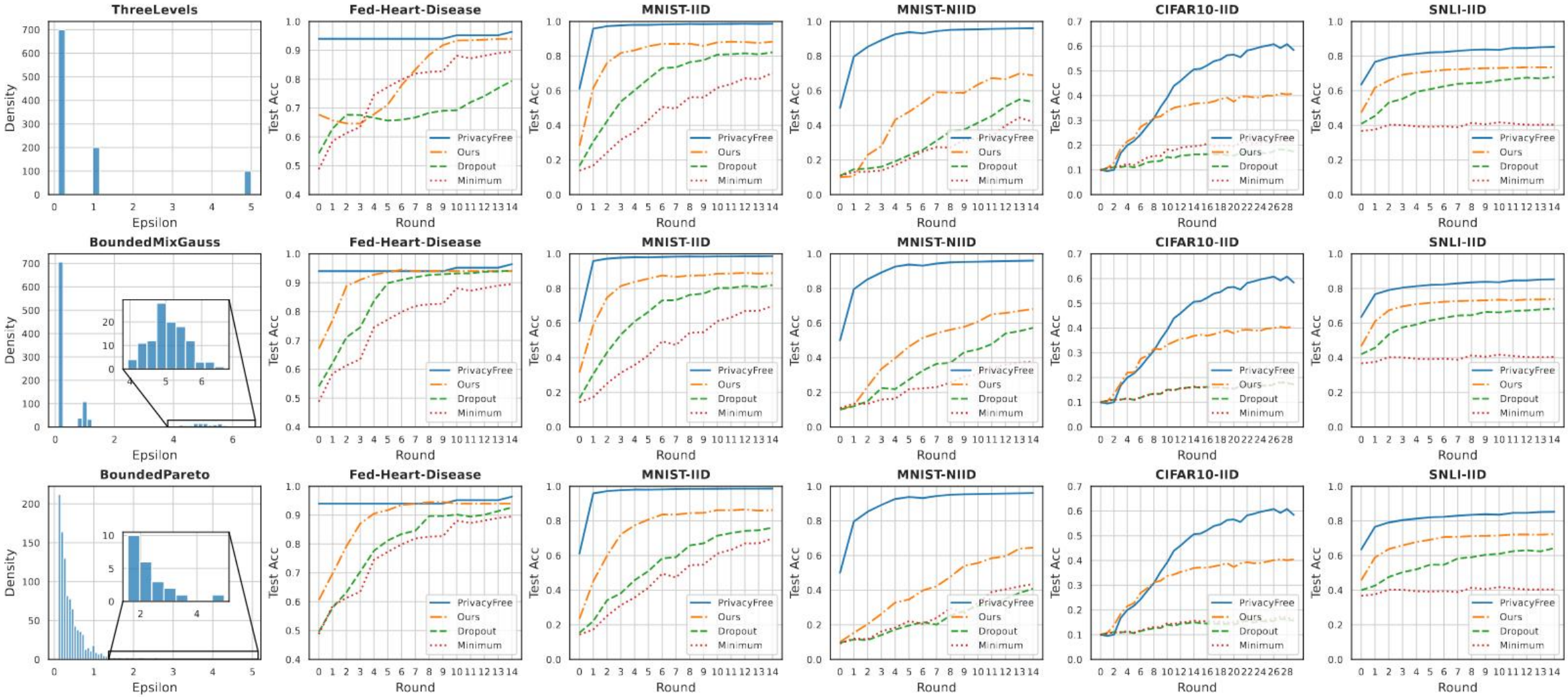
随后可以得到不同 采样概率q 和 最优隐私预算 ε* 的对应曲线。

用函数 f(q) — ε* 去拟合该曲线。

$$\hat{\varepsilon}^* \approx f(q) \triangleq e^{a \cdot q + b} + c,$$

$$q = f^{-1}(\varepsilon)$$




ThreeLevels
Fed-Heart-Disease
MNIST-IID
MNIST-NIID
CIFAR10-IID
SNLI-IID
BoundedMixGauss
BoundedPareto
Density
Epsilon
Test Acc
Round
PrivacyFree
Ours
Dropout
Minimum

AVE	0.05	0.1	0.2	0.4	0.8	1.4	2.8	5.6	11.2	22.4
1	**1.008682**	0.809368	0.725618	0.785445	0.736802	0.783802	0.733693	0.735382	0.735386	0.749253
2	1.042804	**0.59935**	0.597186	0.540947	0.510446	0.508259	0.531975	0.52349	0.507271	0.513549
3	1.491051	0.599983	**0.481246**	0.542943	0.465396	0.456353	0.460885	0.472369	0.491436	0.506605
4	1.982205	0.738332	0.503775	0.492367	0.453479	0.448563	0.447345	**0.433352**	0.43841	0.453412
5	2.271299	0.798641	0.556186	**0.476036**	0.460125	0.475915	0.458337	0.445076	0.463469	0.452184
7.5	2.300377	2.176939	0.586889	0.515522	**0.447086**	0.503011	0.443103	0.457394	0.448825	**0.422854**
10	2.335641	2.294126	0.780405	0.521683	0.482541	**0.449558**	**0.437227**	0.458747	0.468188	0.459458
15	2.570177	2.299228	1.819052	0.551724	0.473429	0.469611	0.472104	0.480327	0.474193	0.469296
20	3.092263	2.313276	2.300976	0.761321	0.512418	0.467892	0.457985	0.468698	**0.431905**	0.444558
25	4.490566	2.321458	2.298353	1.162863	0.551798	0.476411	0.456821	0.453023	0.47047	0.464432

AVG	0.05	0.1	0.2	0.4	0.8	1.4	2.8	5.6	11.2	22.4
1	**0.7487**	0.8248	0.8493	0.8287	0.8466	0.8326	0.8475	0.833	0.8412	0.8429
2	0.7187	**0.8559**	0.8612	0.8671	0.8885	0.8908	0.8744	0.8835	0.8853	0.883
3	0.5384	0.853	**0.8883**	0.8636	0.8927	0.8901	0.8844	0.8816	0.8828	0.8772
4	0.37725	0.7994	0.8727	0.8775	0.8902	0.8861	0.8974	0.8897	**0.8955**	0.891
5	0.157884	0.779	0.8638	**0.8902**	0.8907	0.8837	0.8898	**0.8944**	0.8936	0.8898
7.5	0.1089	0.2416	0.8526	0.8673	**0.893**	0.8758	0.8971	0.8885	0.885	**0.8966**
10	0.1208	0.121675	0.7842	0.8645	0.8854	**0.896**	**0.8983**	0.8903	0.8836	0.8906
15	0.1193	0.111375	0.4057	0.8661	0.8779	0.8825	0.8813	0.8782	0.8927	0.8873
20	0.1176	0.096175	0.098258	0.8027	0.8755	0.8843	0.8917	0.8827	0.8926	0.8913
25	0.1172	0.1043	0.11155	0.6753	0.8639	0.8819	0.8966	0.8913	0.8854	0.8924