NEURAL INFORMATION PROCESSING SYSTEMS

MARS

Multimedia Analysis & Reasoning Lab

Dynamic Personalized Federated Learning with Adaptive Differential Privacy

Xiyuan Yang[1], Wenke Huang[1], Mang Ye[1]

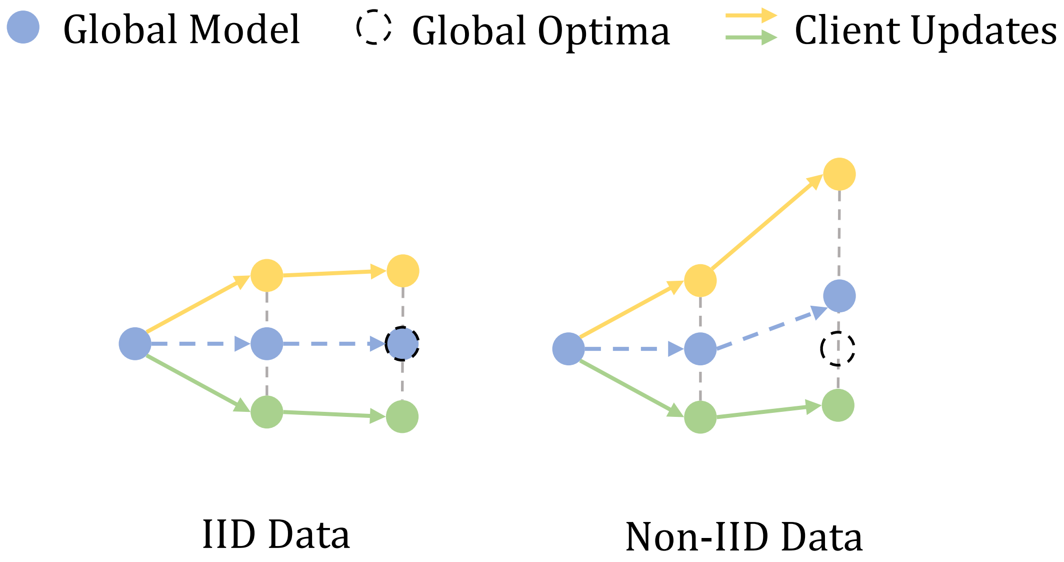
[1]School of Computer Science, Wuhan University, Wuhan, China

Paper: https://nips.cc/virtual/2023/poster/71639

Project Page: https://github.com/xiyuanyang45/DynamicPFL

Introduction

Personalized Federated Learning (PFL)

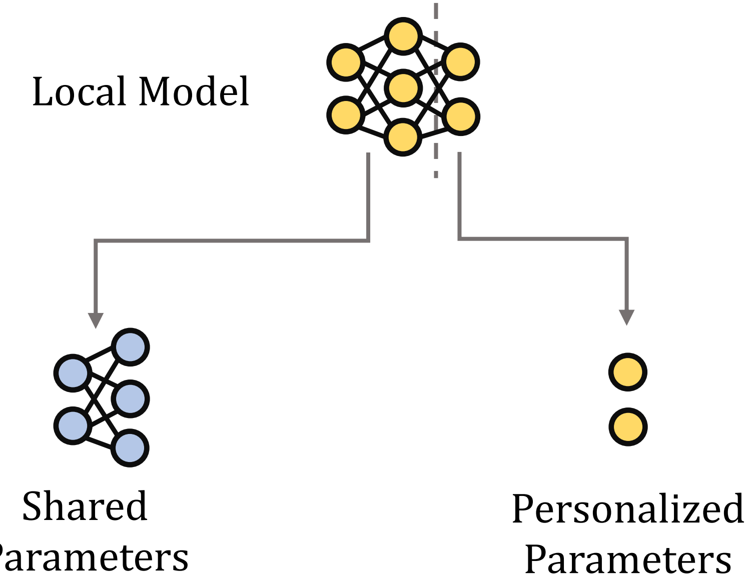


Uniformed Updates

Divergent Updates

Poor Generalization

PFL:



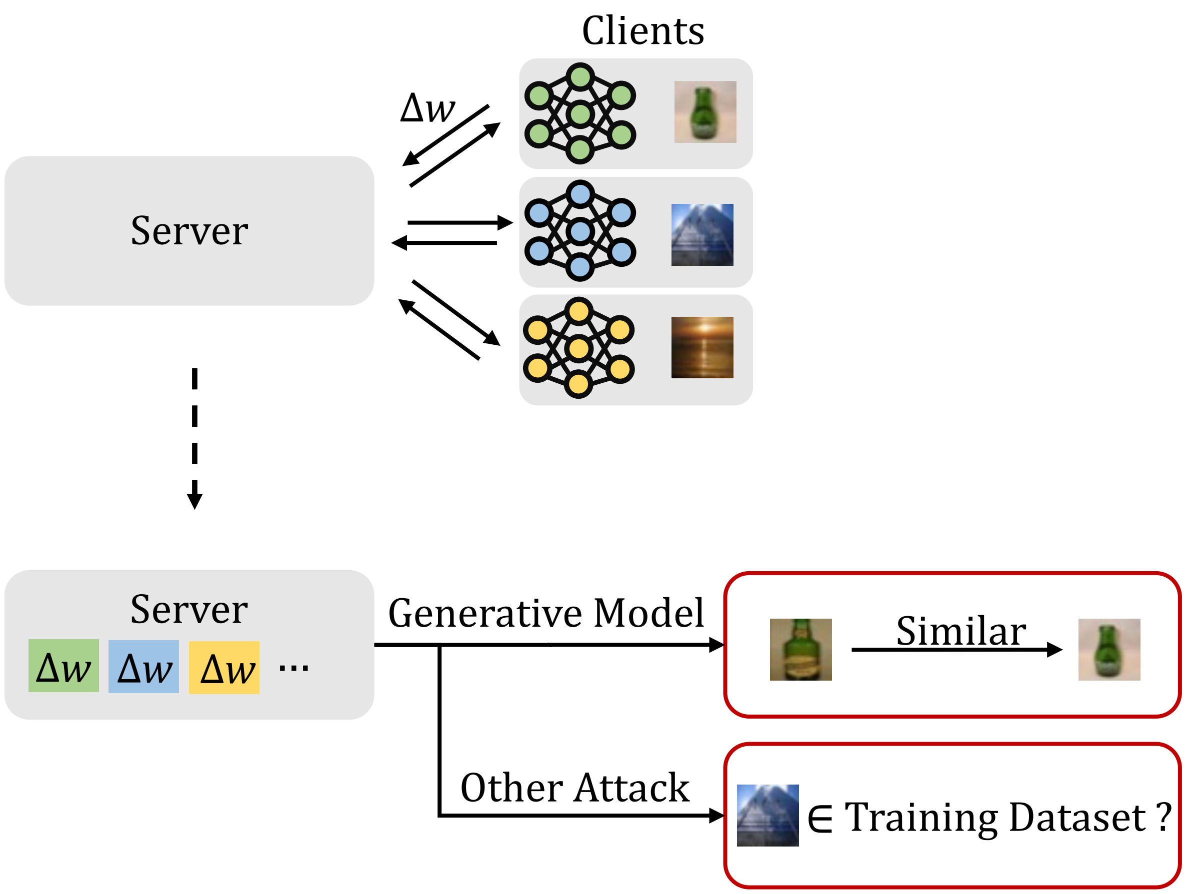
- **shared** by multiple tasks
- learn from **other clients**

- **Personalized** for client
- learn **local** knowledge

Introduction

Differential Privacy in PFL

Privacy Leakage Issue



User-Level Differential Privacy:

Clipping:

$$\Delta w \leftarrow \frac{\Delta w}{\max(1, \frac{\|\Delta w\|_2}{C})}$$

Introducing randomness into the training algorithm to protect privacy

Adding Noise:

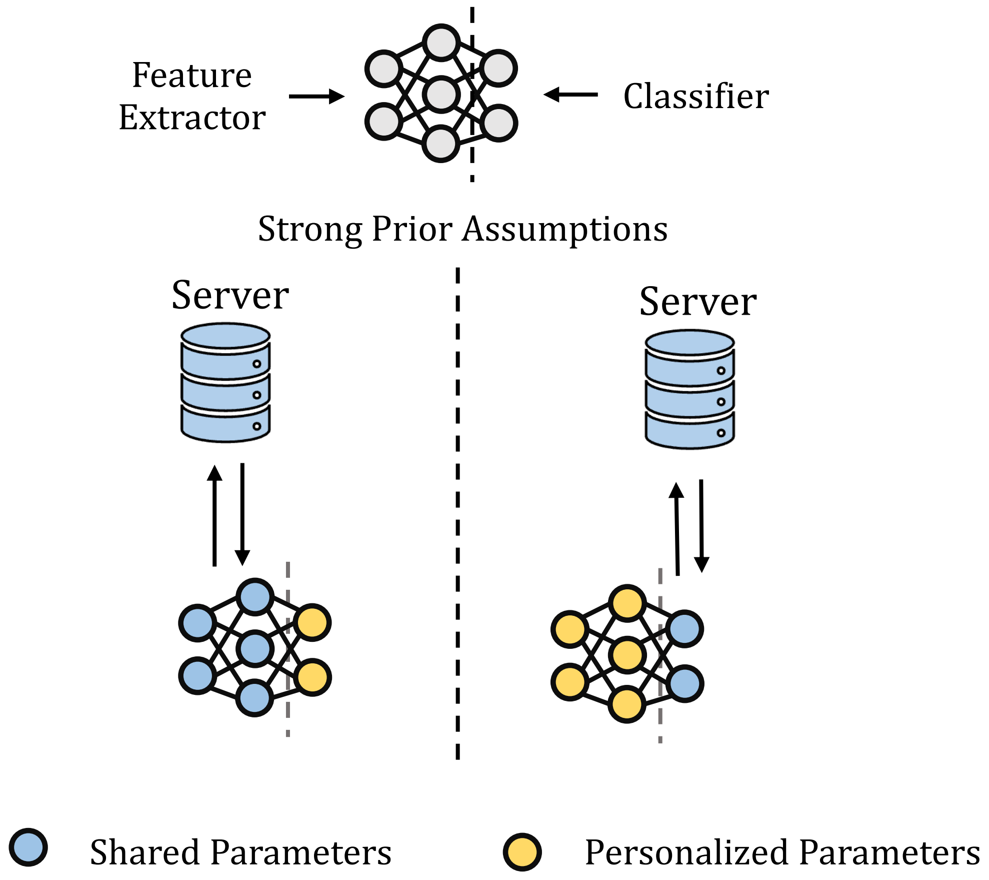
$$\Delta w \leftarrow \Delta w + \mathcal{N}(0, C^2\sigma^2/|m^t|)$$

Clip client updates in L2-norm, limiting the contribution of single participant

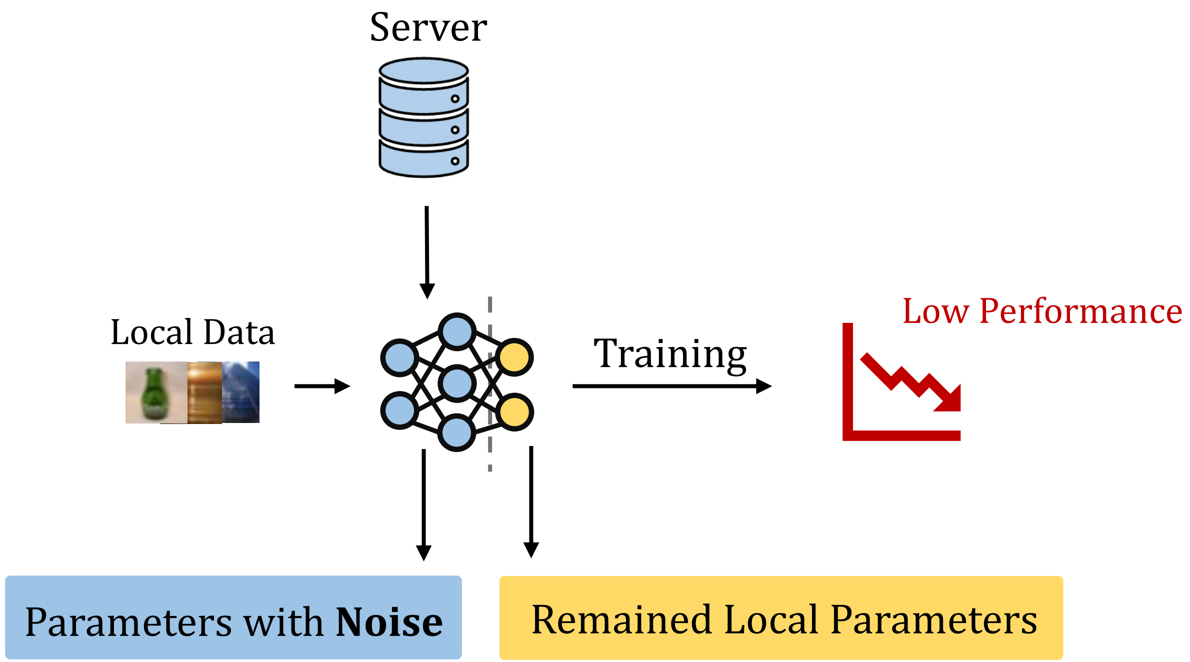
C: Clipping Bound σ:noise scale m: number of clients

Problem 1

Inflexible Personalization



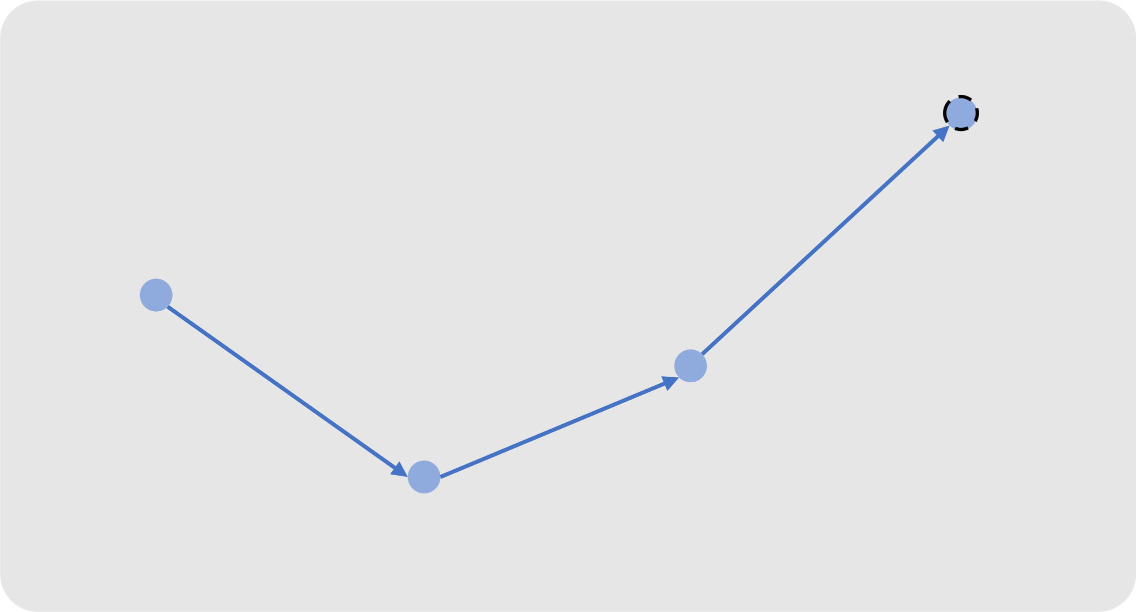
Noise Effects



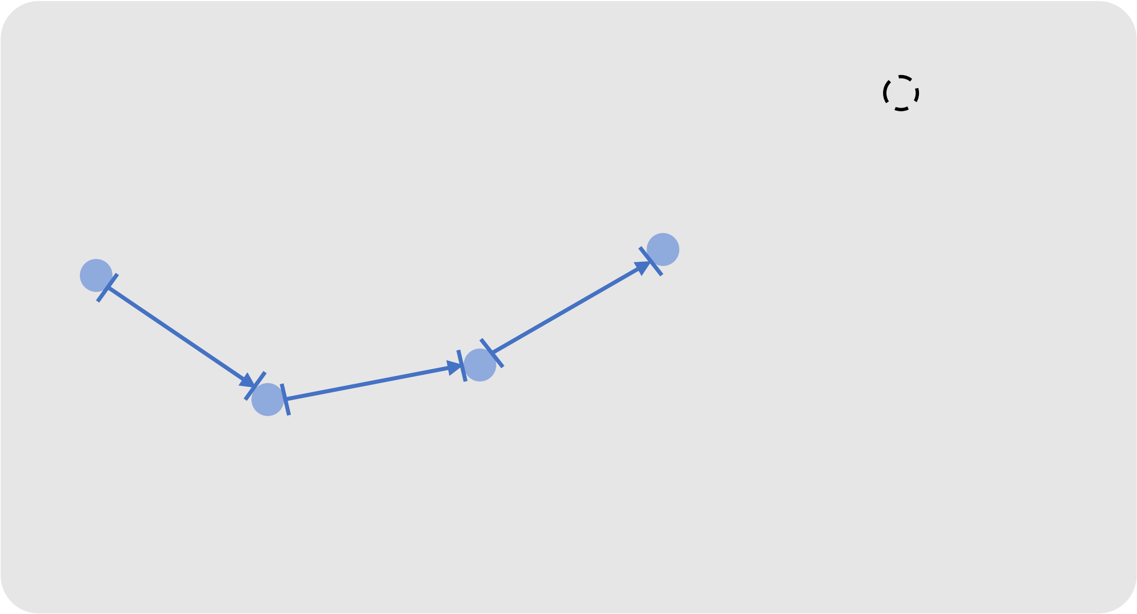
Problem 2

Model Convergence Difficulty

Conventional Gradient Calculation:



Clipped Gradient:



Global Model Global Optima Clipped Gradient

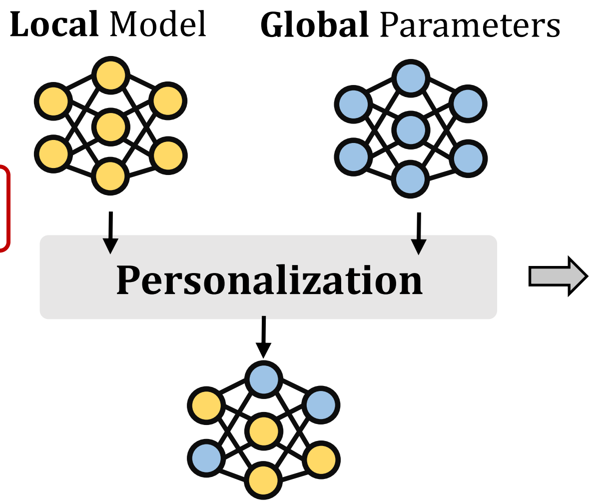
Methodology

Dynamic Personalization:

We introduce **Fisher Information** value to dynamically select out personalized parameters.

large Fisher value → high information content → personalized

When server distributing parameters, client's personalized part is remained, thus parameters with more information are **not affected by noise**



Personalization:

Local Model → Calculate Fisher Value

informative parameters → Personalized: **u**

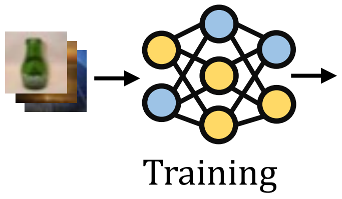
less informative parameters substitute by Shared: **v**

Adaptive Differential Privacy:

We design an **adaptive** gradient adjustment term

For **u**, we limit the update norm to preserve informative parameters

For **v**, we make it closer to C, thus mitigating the impact of clipping



$\mathcal{L}oss = \mathcal{L}_{CE} + \mathcal{L}_1 + \mathcal{L}_2$

Constraint:

penalty $\mathcal{L}_1$: $||\Delta \mathbf{u}||$

penalty $\mathcal{L}_2$: $|| \; || \Delta \mathbf{v} || - C \; ||$

v: Shared Parameters C: Clipping Bound

u: Personalized Parameters

Experiments

Comparison with Stat-Of-The-Art Methods

Methods	FEMNIST $\epsilon=2$	FEMNIST $\epsilon=4$	FEMNIST $\epsilon=8$	FEMNIST $\epsilon=16$	SVHN $\epsilon=2$	SVHN $\epsilon=4$	SVHN $\epsilon=6$	SVHN $\epsilon=8$
DP-FedAvg [12]	71.92	72.80	73.26	75.02	53.91	54.76	56.33	57.65
BLUR+LUS [8]	72.80	74.23	74.58	75.16	53.76	57.52	56.90	58.57
PPSGD [3]	68.20	69.60	71.03	71.93	55.89	56.15	56.74	59.37
DP-FedSAM [38]	73.13	73.94	74.54	74.66	53.04	52.84	54.03	56.08
FedDPA(Ours)	**74.46** ↑1.33	**77.27** ↑3.04	**77.42** ↑2.84	**76.99** ↑1.83	**58.78** ↑2.89	**62.63** ↑5.11	**63.66** ↑6.76	**64.57** ↑5.29

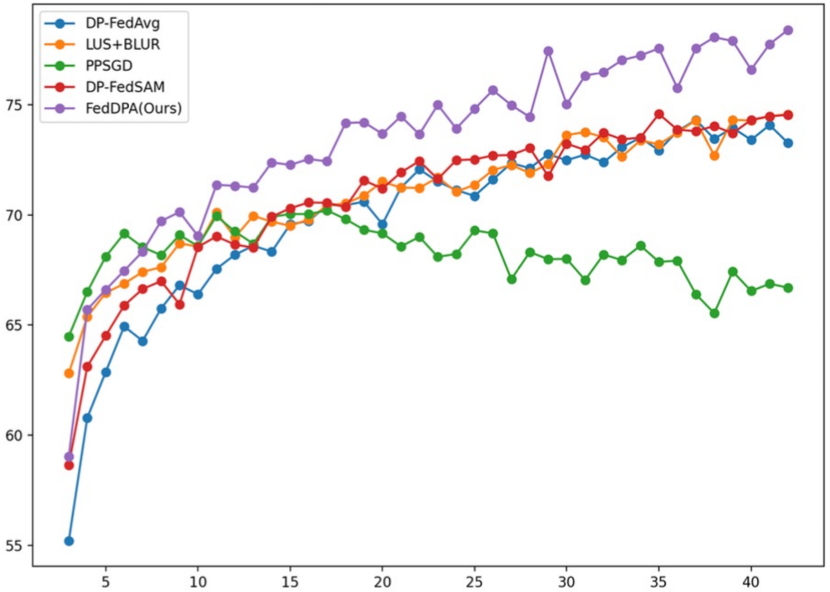
Smaller ε sepresents Stronger Privacy Protection

Ablation Study

Components DFP	Components AC	Accuracy on clients dataset $\epsilon = 2$	$\epsilon = 4$	$\epsilon = 8$	$\epsilon = 16$
		71.92	72.80	73.26	75.02
✓		72.18	74.56	74.38	75.33
✓	✓	**74.46** ↑2.28	**77.27** ↑2.71	**77.42** ↑2.93	**76.99** ↑1.66

Accuracy Curve

Ours



Conclusion

1. In this paper, we delve into the challenges of inflexible personalization and convergence difficulties under differential privacy in personalized federated learning

2. We introduce FedDPA with two innovative components, Dynamic Fisher Personalization (DFP) and Adaptive Constraint (AC), that utilize layer-wise Fisher information to dynamically personalize and adaptively constrain parameters, effectively addressing the challenges

3. The efficacy of our proposed methods has been thoroughly validated has been validated against many state-of-the-art methods on various datasets

Thanks for watching!

Xiyuan Yang