Localizing Partial Model for Personalized Federated Learning

Heewon Park, **Miru Kim** and **Minhae Kwon**
{heewon012,48rlaalfm}@soongsil.ac.kr, minhae@ssu.ac.kr
Soongsil University, Seoul, Republic of Korea

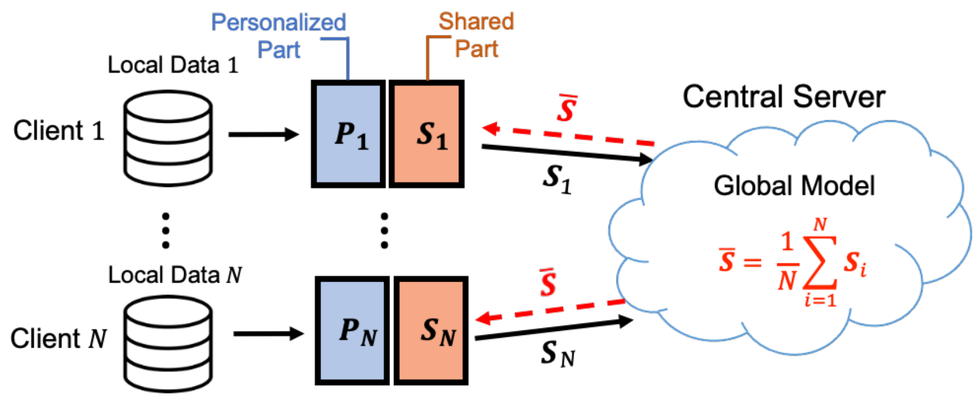
Motivation & Objectives

Federated learning offers a privacy-preserving solution by training models across multiple devices without exchanging raw data [1]

- Limitation1: **The heterogeneous client data distribution** hinders the generation of local models optimized for individual data
- Limitation2: Vanilla FL is **vulnerable to attacks** because it operates under the assumption that all clients are trustworthy

→ We propose an algorithm that **localizes the partial model** and updates local models in **two steps**

Partial Sharing Federated Learning



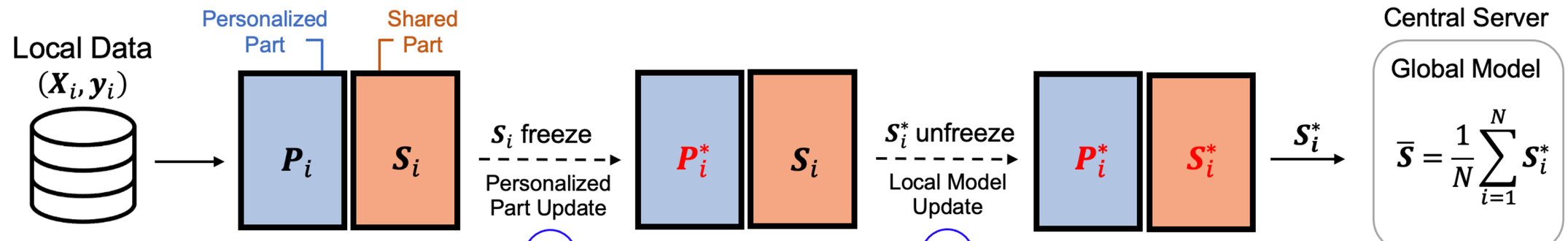
Devide local model into two parts: **personalized-shared part** [2]

Goal: Increase Performance of Personalization (PoP)

$$PoP = \frac{1}{N}\frac{1}{T_{total}}\sum_{i=1}^{N}\sum_{k=1}^{K} T_{k,i}q_{k,i}$$

$\mathbf{q}_{k,i}$: The proportion of k-th label data of i-th client
$\mathbf{T}_{total}$: # total test set
$\mathbf{T}_{k,i}$: # correct prediction of k-th label predicted by i-th client

Proposed Partial Sharing Algorithm: pFedFrz



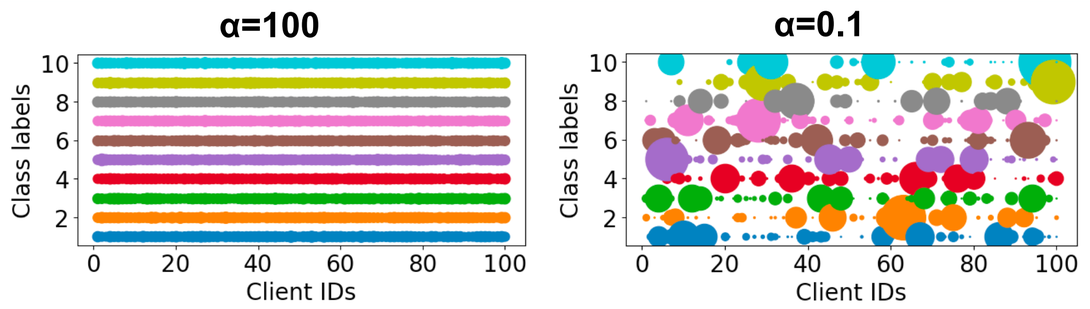
η: Learning rate L_i : Loss function $\mathbf{P}_i^*$: Updated personalized part $\mathbf{S}_i^*$: Updated shared part

Key Point	Updates local model in **two steps** ①: Update **only the personalized part** → Increase compatibility between personalized-shared parts ②: Update both personalized and shared parts → **Generate a local model optimized** to local data

Simulation Results

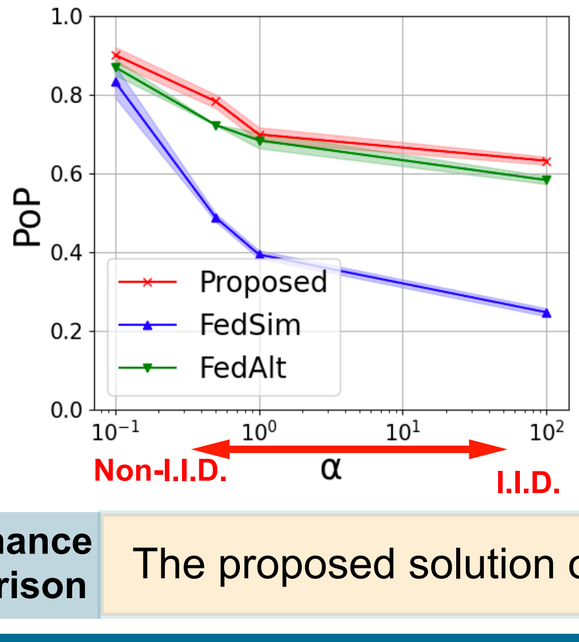
Simulation Setup

- Dataset: MNIST, EMNIST
- Attacker: **Gradient Sign Flip Attack**
- **Heterogeneous** setting with **Dirichlet Distribution**



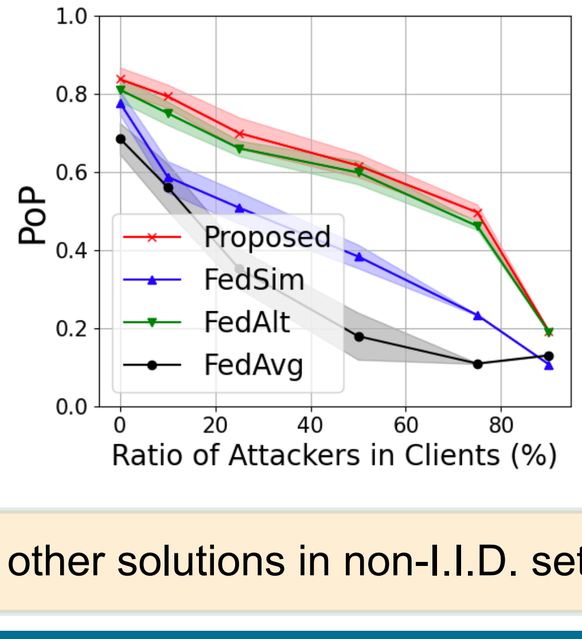
Performance of **Personalization**

EMNIST

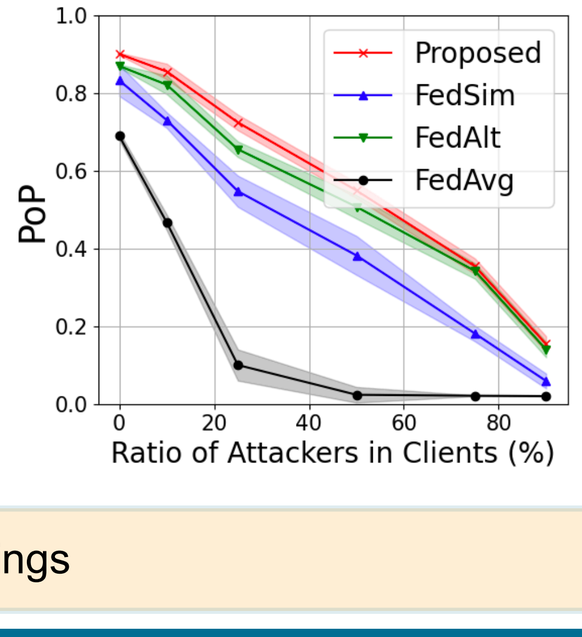


Robustness against Gradient Sign Flip Attack

non-I.I.D. MNIST (α=0.1)



non-I.I.D. EMNIST (α=0.1)



Performance Comparison	The proposed solution outperforms other solutions in non-I.I.D. settings

Conclusion

- We propose pFedFrz for local update in partial sharing federated learning that trains local model in two steps
- The proposed solution generates local models optimized to individual datasets in non-I.I.D. setting
- The proposed solution builds robust local models against attackers

[1]B. McMahan, et al., "Communication-efficient learning of deep networks from decentralized data," AISTATS 2017.
[2]K. Pillutla, et al., "Federated Learning with Partial Model Personalization," ICML 2023.