河海大学
HOHAI UNIVERSITY

Edge-Cloud Enabled Smart Sensing Applications with Personalized Federated Learning in IoT

Yingchi Mao[a,b], Yi Rong, Jiajun Wang[b], Zibo Wang[b], Xiaoming He[a,b], and Jie Wu[c]

[a] College of Computer Science and Software Engineering, Hohai University, Nanjing, China

[b] College of Internet of Things, Nanjing University of Posts and Telecommunication, Nanjing, China

[c] Center for Networked Computing, Temple University, Philadelphia, USA

IEEE GLOBECOM

2024 IEEE Global Communications Conference

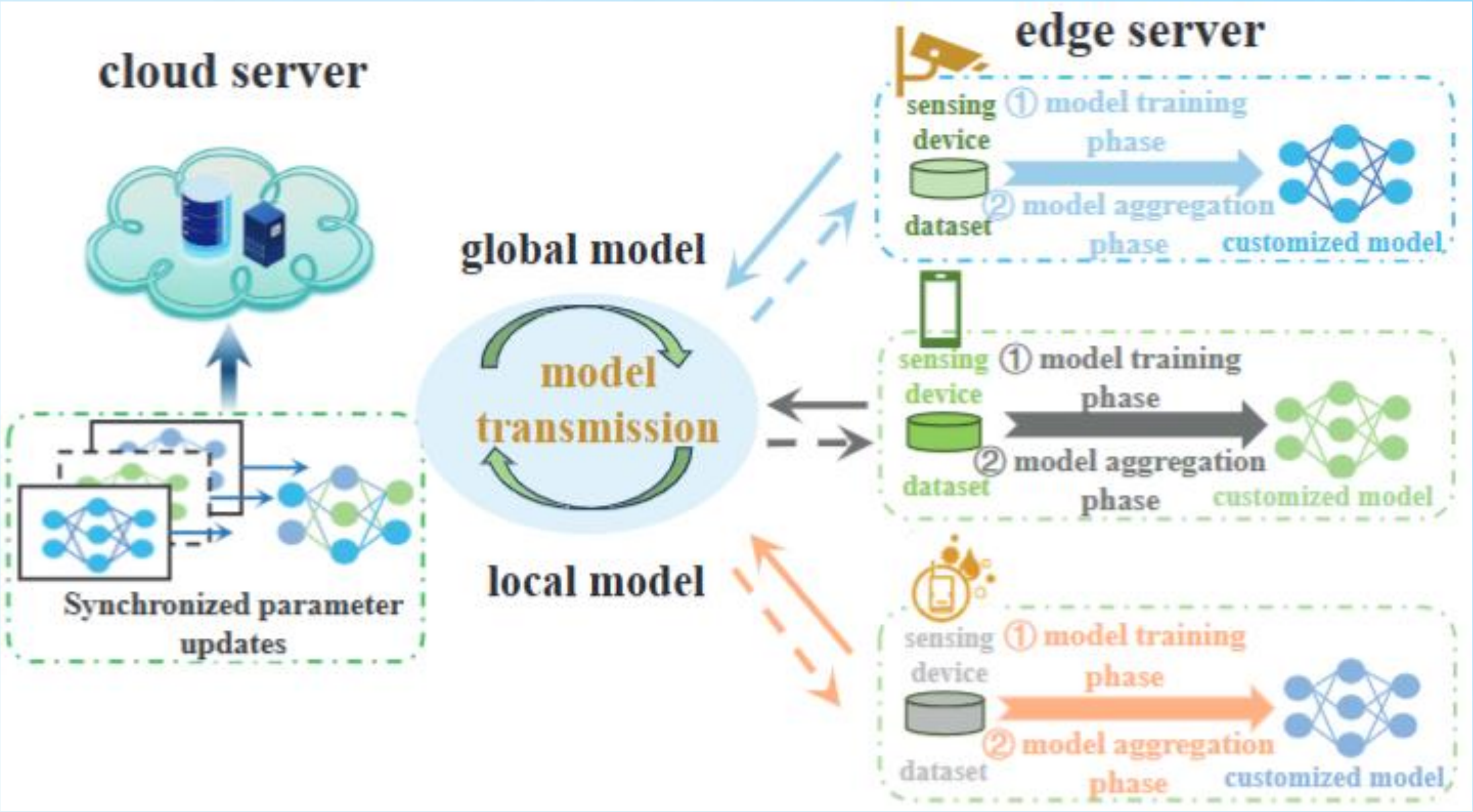
Dec. 8-12, 2024

Content

Introduction

■ Background

The rapid development of deep learning technologies and the widespread deployment of sensing devices have brought considerable attention to Internet of Things. The smart sensing application is one of the popular applications in IoT. Personalized Federated Learning (pFL) is a replacement to traditional Federated Learning to tackle the statistical heterogeneity of clients' private datasets.



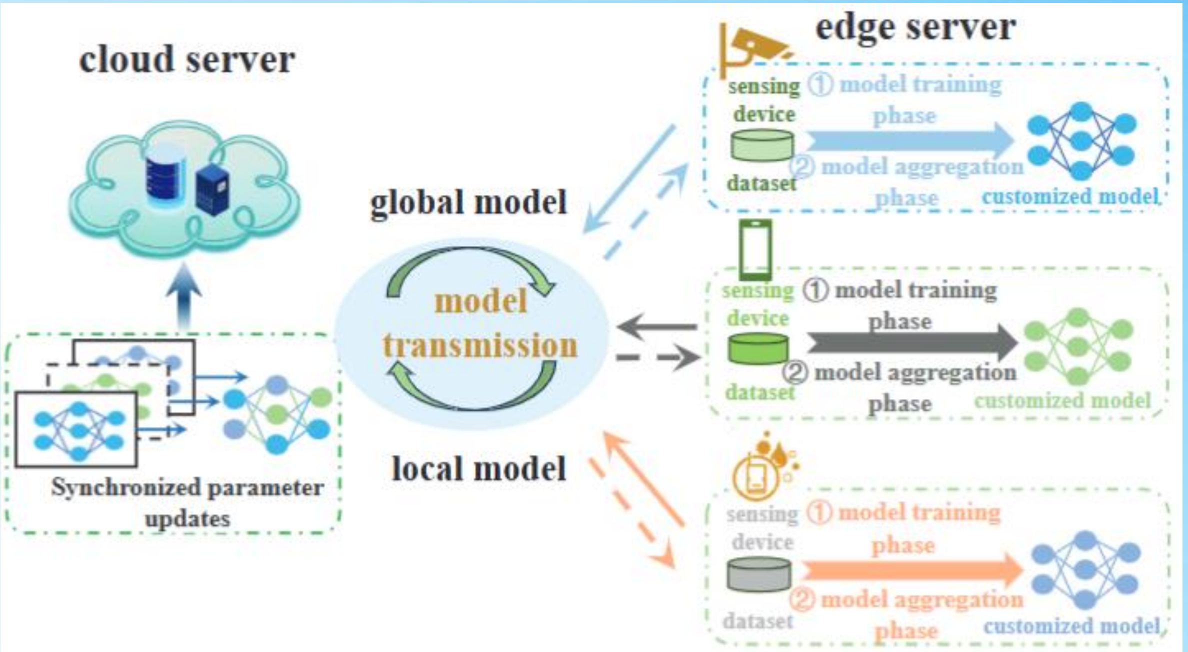
Introduction

- ■ **Motivations**: What are challenges that exist in Smart Sensing Applications?

However, existing pFL methods encounter two challenges in smart sensing applications:

- ➢ **The effect of the global model preference**: Existing pFLs based on weighted model aggregation have not succeeded in reducing the adverse effect of the global model preference.

- ➢ **Dynamic role differences of model hierarchy**: Few studies have investigated hierarchical aggregation, but there is a lack of adaptive aggregation weights to capture the dynamic changes of hierarchical roles, leading to inaccurate personalization.



Approach

■ **Proposal**: How to solve these challenges?

a novel edge-cloud enabled weighted model aggregation-based pFL framework named pFL-Sensing for smart sensing applications

01 Problem formulation

We first formally define the pFL-based smart sensing application. The optimization objective is determinated.

02 Problem modeling

1. model training phase
2. model aggregation phase

Approach

01 Problem formulation

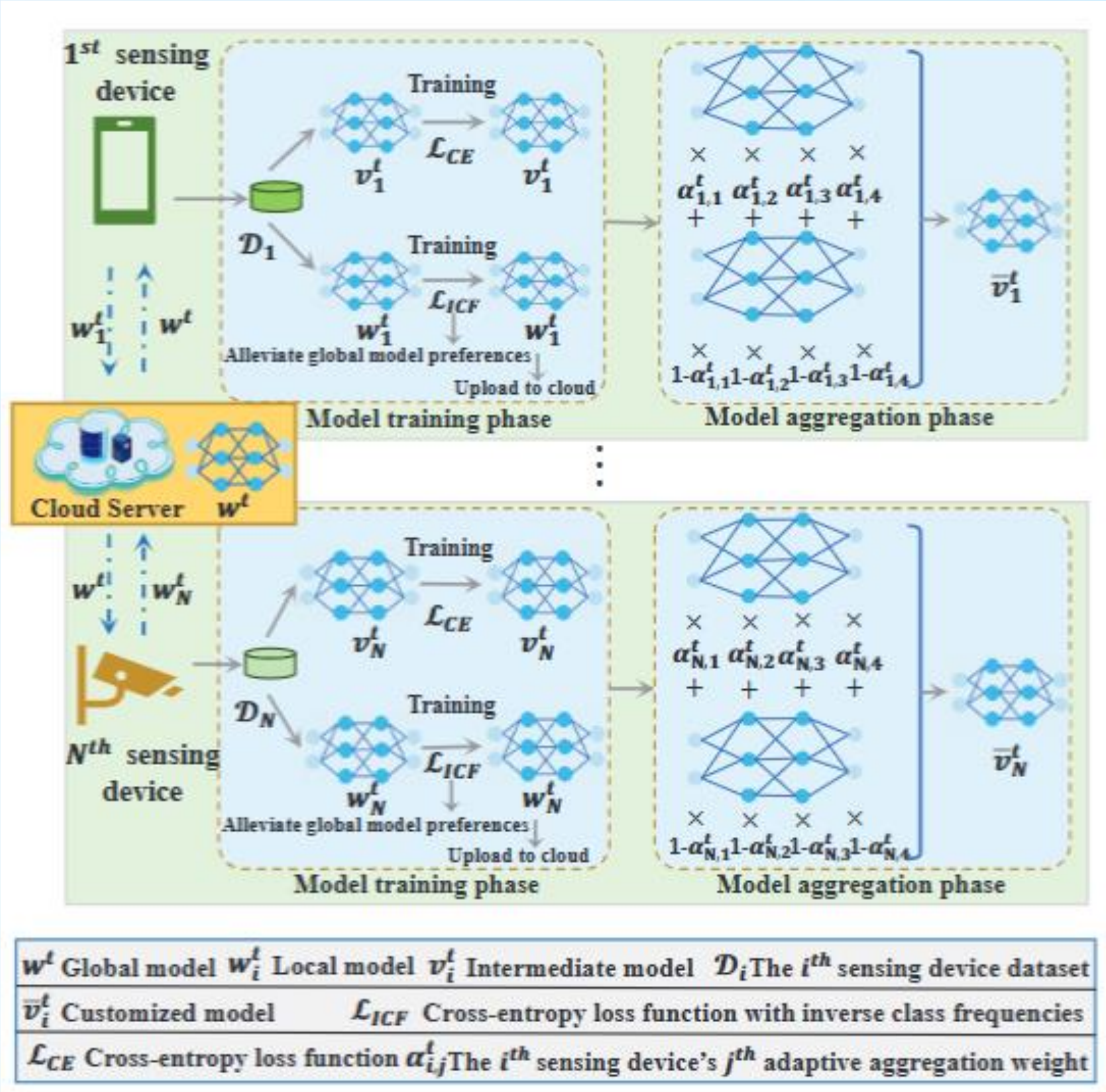
We collaboratively train customized models $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_N$ through $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_N$, where $\bar{v}_i$ is a customized model for the i^{th} sensing device. The optimal customized models are acquired through minimizing the global loss:

$$(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_N) = \arg\min_{\bar{v}} \sum_{i=1}^{N} \frac{m_i}{M} \mathcal{L}_i$$
$$\mathcal{L}_i = \mathcal{L}(\mathcal{D}_i; \bar{v}_i; w), \forall i \in [N]$$

where $\mathcal{L}(\cdot)$ stands for the global loss function. $\mathcal{L}_i$ represents as the loss function of the i^{th} sensing device related to dataset $\mathcal{D}_i$, measuring the discrepancies between the predicted value and the real label of m_i data samples. w stands for the global model, which provides useful information to train the i^{th} sensing device's customized model.

Approach

02 Problem modeling

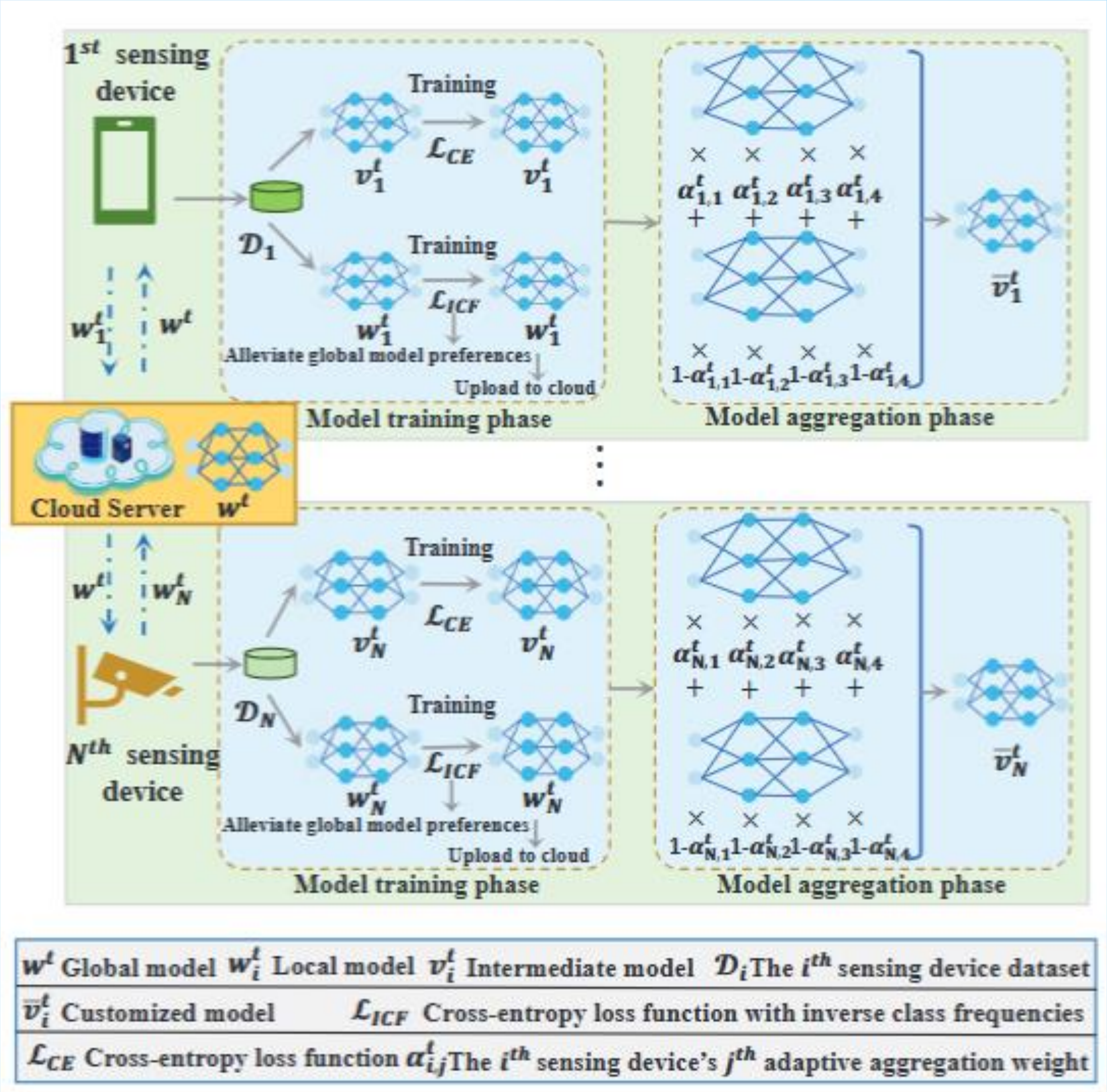


1 model training phase

In the model training phase, to alleviate the low performance of the local model on the minority class in sensing device data due to the global model preference, we introduce ICF into the cross-entropy loss function for updating the local model.

Approach

02 Problem modeling

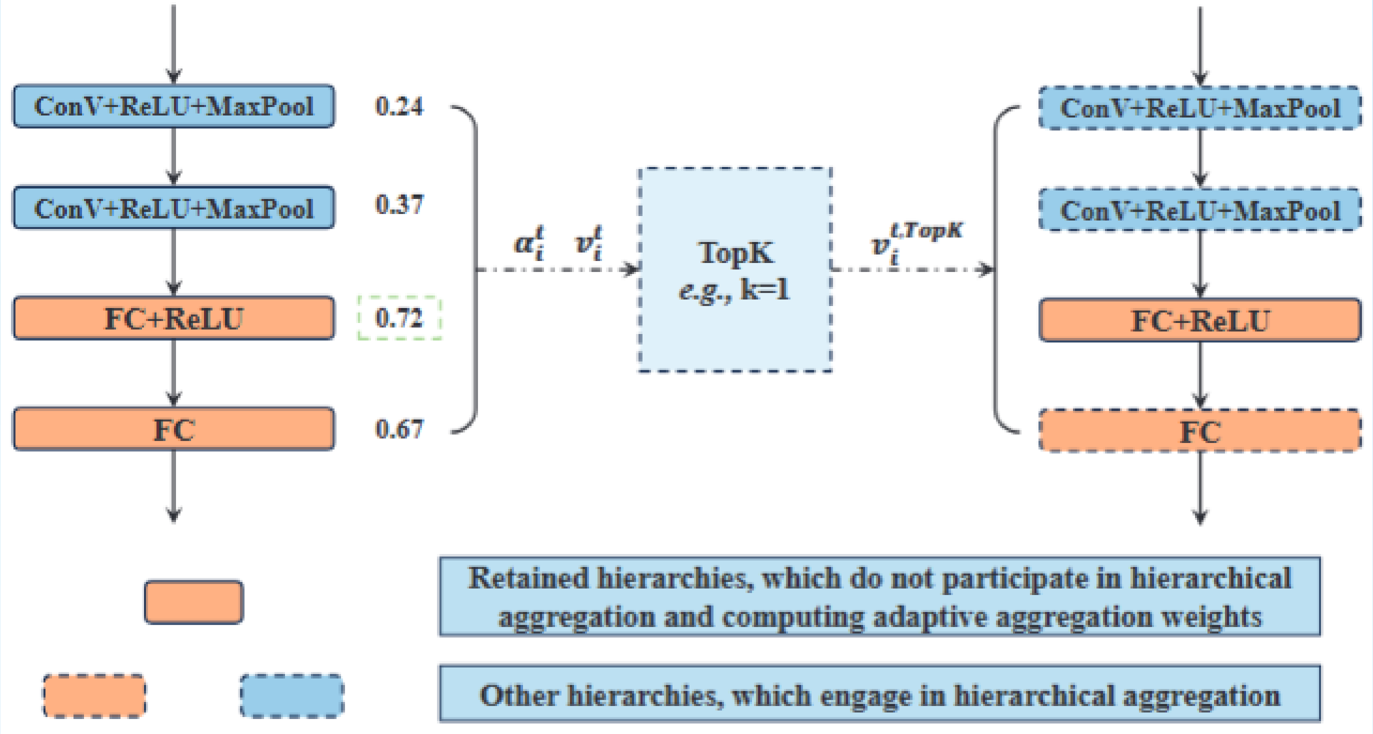


2 model aggregation phase

In the model aggregation phase, to capture the dynamic role differences of model hierarchy, the hierarchical aggregation and an Adaptive Weight Calculation (AWC) mechanism are introduced to implement the aggregation of the local model and the intermediate model for each sensing device.

Approach

- Adaptive Weight Calculation (AWC) mechanism



a TopK strategy is introduced into the AWC. The key idea of the TopK is to heuristically retain the partial hierarchies of the intermediate model based on the top k aggregation weights. These preserved hierarchies are directly included as a part of the customized model without participating in hierarchical aggregation, thus reducing the computational cost of aggregation weights for k hierarchies.

Experiment

■ Experiment Setup

- We conduct experiments to evaluate the performance of pFL-Sensing.
- Specifically, we choose the CIFAR-100, Tiny-ImageNet, and AG News as experiment datasets.
- We employ the Dirichlet distribution simulate the data heterogeneity.

Dataset →

Dataset
CIFAR-100
Tiny-ImageNet
AG News

Experiment Setup

- The parameter settings are shown in the table.

Parameter	Value
Communication round E	500
Batch size	10
The number of clients N	20
Initial aggregation weights	0.5
Learning rate η	0.1
The defaultparameter of the Dirichlet function β	0.1

- And the indicators include classification **accuracy**.

Experiment

■ Analysis of Classification Performance Comparison

Pathological Heterogeneous Setting: Compared with these baselines, pFL-Sensing achieves the best performance, demonstrating how considering the global model preference with the ICA and dynamic role differences of model hierarchy with the AWC can improve classification precision.

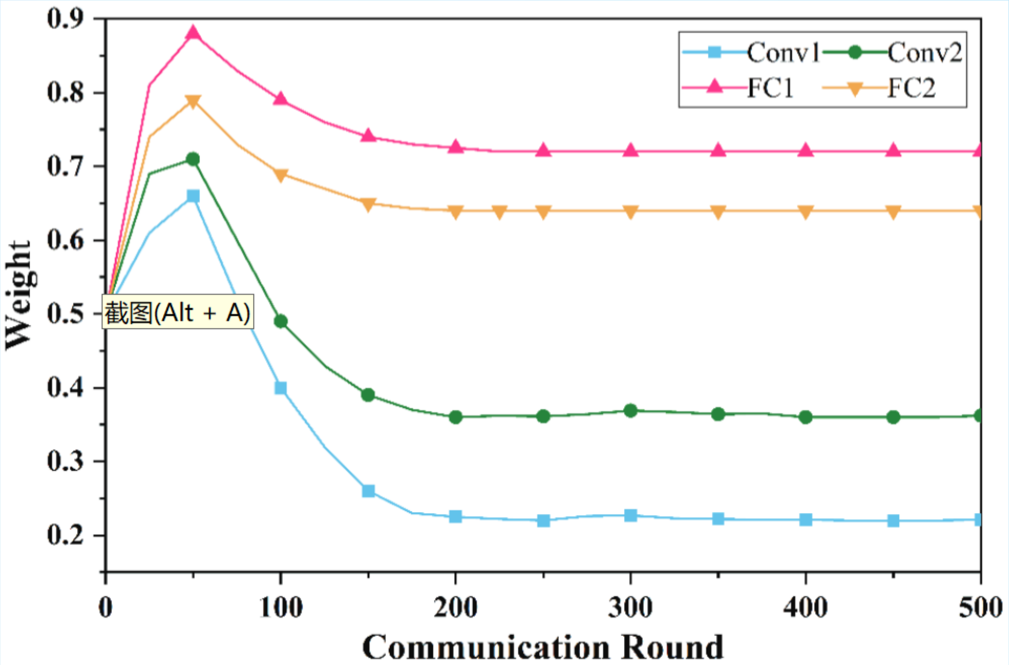
Practical Heterogeneous Setting: Compared with these baselines, pFL-Sensing achieves the highest classification accuracy because of the inclusion of the global model preference and dynamic role differences of model hierarchy. In contrast to the 4-hierarchy CNN, ResNet-18 is regarded as a large backbone with more hierarchies.

Setting	pathological heter setting		practical heter setting			
Methods	CIFAR-100	TINY	CIFAR-100	TINY	TINY*	AG News
FedAvg	25.98	14.20	31.98	19.46	19.45	79.57
Per-FedAvg	56.80	28.06	44.28	25.07	21.81	93.27
Ditto	67.23	40.23	52.87	32.15	35.92	95.45
pFedMe	58.20	27.71	47.34	26.93	33.44	91.41
FedAMP	64.34	37.15	47.69	27.99	29.11	94.18
FedPHP	63.09	37.88	50.52	35.69	29.90	94.38
FedFomo	62.49	35.87	45.39	30.33	32.84	95.84
PartialFed	65.35	37.76	51.37	32.78	36.91	94.87
pFL-Sensing	68.76	42.16	54.19	39.38	40.42	95.93

■ Analysis of Aggregation Weight Evolution

In the early stages, since the customized model needs to learn personalized knowledges from individual clients, the intermediate model is given larger weights.

As the communication rounds iterate, generalized information requires to be considered for the customized model. Meanwhile, the customized model gradually converges. As a result, the aggregation weights of the intermediate model progressively decrease and tend to stabilize.



Conclusion

Conclusion

- In this paper, we have proposed an edge-cloud enabled weighted model aggregation-based pFL framework named pFL-Sensing for smart sensing applications. Each edge server has generated a customized model through two phases-a model training phase and a model aggregation phase.

➢ In the former phase, to alleviate the global model preference, we have introduced the ICF into the loss function for local model training.

➢ In the latter phase, we have integrated hierarchical aggregation and the AWC. We have proposed hierarchical aggregation to aggregate each hierarchy of the local model and the intermediate model with aggregation weights to produce an individual customized model for each client. We also have designed the AWC to adaptively update aggregation weights based on dynamic role differences of model hierarchy.

Future Work

- Experiments on large-scale datasets
- Further improve the accuracy

At Last

Thanks for your attention!