上海创智学院

Shanghai Innovation Institute

CVPR

JUNE 3-7, 2026

DENVER

COLORADO

FedRAC: Rolling Submodel Allocation for Collaborative Fairness in Federated Learning

Zihui Wang[1,2], Yuhang Fu[1], Mengmeng Du[5], Zhimin Yuan[6],

Yachen Liu[4], Weisheng Liao[2], Kaiyu Wang[1,*] , Zheng Wang[3,4,*]

[1]Harbin Engineering University, [2]Pengcheng Laboratory, [3]Shanghai Innovation Institution,

[4]Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, China

[5]Fuyang Normal University, [6]Nanyang Normal University

wangzh08@pcl.ac.cn, yuhangfu@hrbeu.edu.cn, zwang@stu.xmu.edu.cn,wang kaiyu@hrbeu.edu.cn

Paper ID: 40217

Code: https://github.com/ZiHuiWangpcl1/FedRAC

Background (1)：Collaborative Fairness (CF)

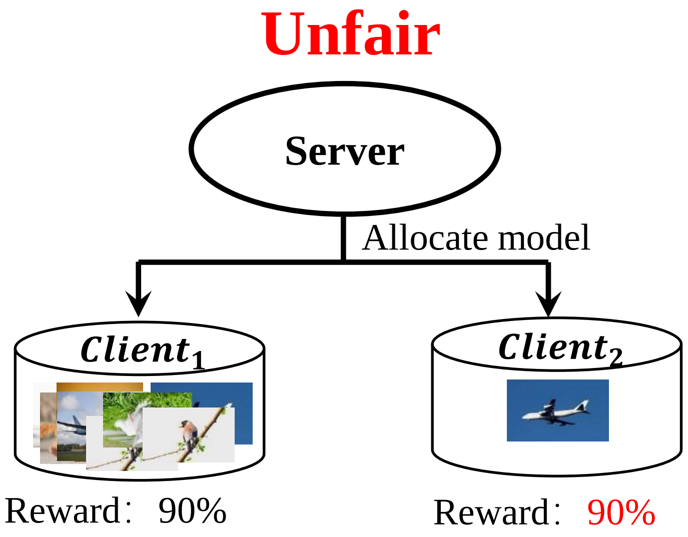


Figure 1: Unfair example in FL

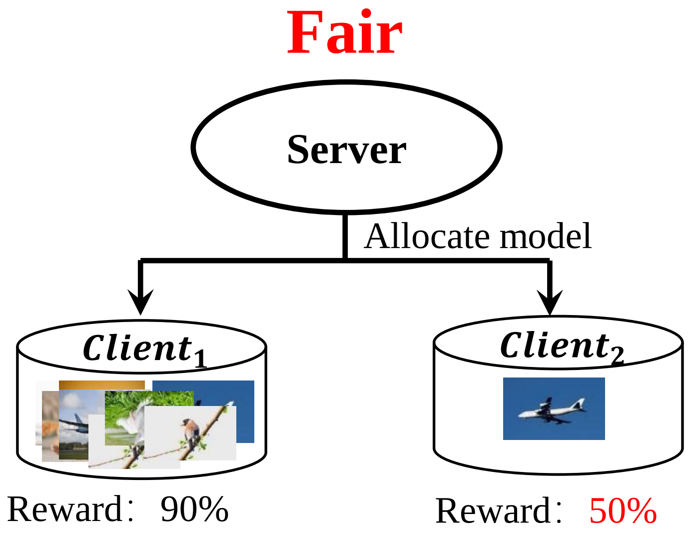


Figure 2: Fair example in FL

- Unfair: Early FL frameworks [1] usually distributed the same model to all clients without considering their distinct contributions to the model performance, resulting in unfairness to high-contributing clients

- Fair: CF [2] stands as an essential element in federated learning to encourage client participation by equitably distributing rewards based on individual contributions

---

[1] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Artificial intelligence and statistics. PMLR, 2017: 1273-1282.
[2] Xu X, Lyu L, Ma X, et al. Gradient driven rewards to guarantee fairness in collaborative machine learning[J]. Advances in Neural Information Processing Systems, 2021, 34: 16104-16117.

Background (2)：Existing methods

- **CFFL** [IJCAI workshop-20]
  - **CFFL** allocates more gradients to higher reputation client, and the reputation is calculated by the local accuracy and data sizes (or label diversity)
- **CGSV** [NeurIPS-21]
  - **CGSV** assigns more gradients to clients whose local model gradients is more similar to the global gradients
- **FedAVE** [Neurocomputing-24]
  - **FedAVE** assigns more gradients to clients whose data distribution information is more similar to the ideal dataset
- **IAFL** [ICLR-24]
  - **IAFL** assigns differentiated training-time model rewards by allowing higher-contribution clients to aggregate a larger proportion of local model updates
- **FedSAC** [KDD-24]
  - **FedSAC** allocates high-performance submodels containing more key neurons to higher-contribution clients based on dynamic reputation.

Background (3)： Challenges

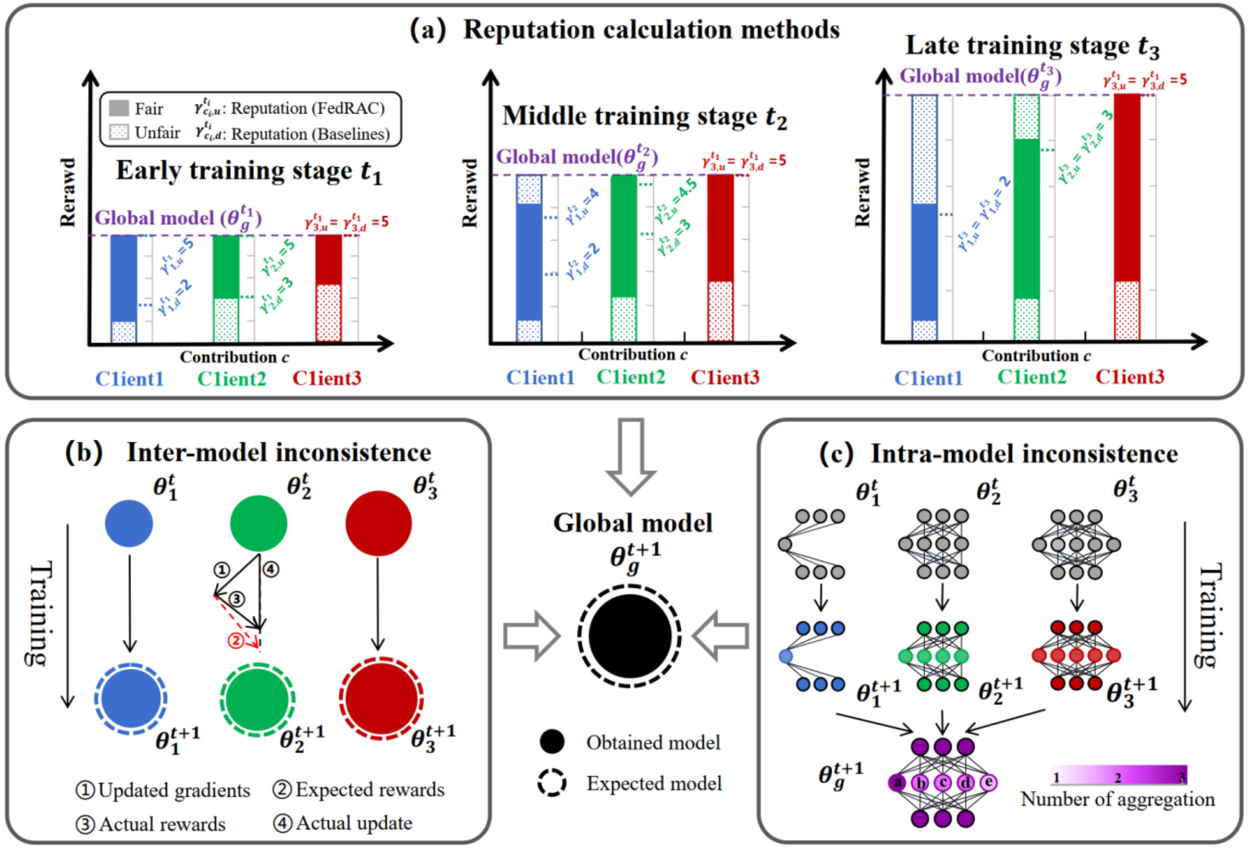
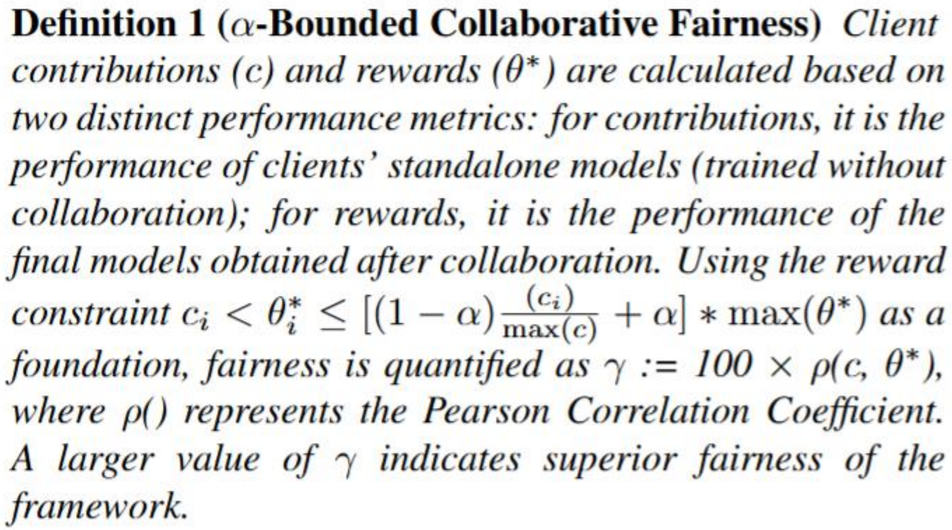


Figure 3: Problem illustration of collaborative fairness in FL

Existing methods to enhance CF still face two kinds of challenges:

1) In the reputation calculation phase, existing approaches overlook the fact that the performance of the global model improves progressively. By assigning fixed reputation proportions to all clients throughout training, they under-reward low-contribution clients in the early stages, which ultimately leads to a decline in the performance of the aggregated model.
2) In the reward allocation phase, conventional methods often lead to inter-model (i.e., gradient-based) or intra-model (i.e., submodel-based) inconsistencies, thereby degrading the performance of the aggregated model.

α −Bounded Collaborative Fairness (α − BCF)



Definition 1 (α**-Bounded Collaborative Fairness)** *Client contributions (c) and rewards (θ^*) are calculated based on two distinct performance metrics: for contributions, it is the performance of clients' standalone models (trained without collaboration); for rewards, it is the performance of the final models obtained after collaboration. Using the reward constraint* $c_i < \theta_i^* \leq [(1-\alpha)\frac{(c_i)}{\max(c)} + \alpha] * \max(\theta^*)$ *as a foundation, fairness is quantified as* $\gamma := 100 \times \rho(c, \theta^*)$*, where* $\rho()$ *represents the Pearson Correlation Coefficient. A larger value of* γ *indicates superior fairness of the framework.*

We propose **α −Bounded Collaborative Fairness (α −BCF)** to tackle the issue of insufficient incentives for the high-contributing client. $\boldsymbol{\alpha}$− BCF could ensure $c_i < \theta_i^* < \left[(1-\alpha)\frac{(c_i)}{max(c)} + \alpha\right] * \max(\theta^*)$ and then quantitative fairness with the Pearson Correlation Coefficient ρ(c;θ^*). The rationale behind the formula in ***Definition* 1** aims to amplify significant distinctions in rewards.

- $c_i < \theta_i^*$ ensures that clients' rewards exceed their contributions.
- $\theta_i^* < \left[(1-\alpha)\frac{(c_i)}{max(c)} + \alpha\right] * \max(\theta^*)$prevents excessive rewards for clients with low contributions.

FedRAC overall architecture

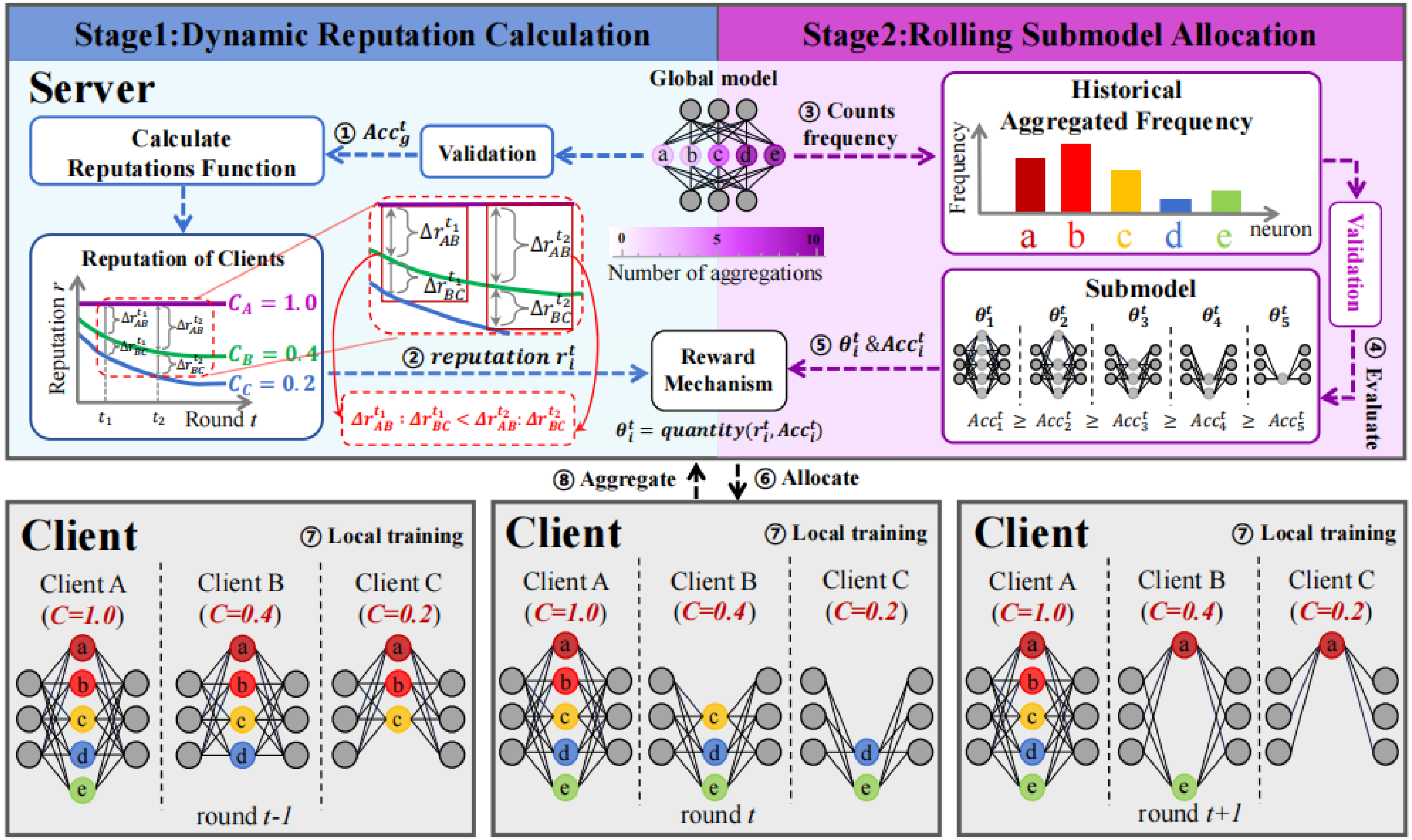
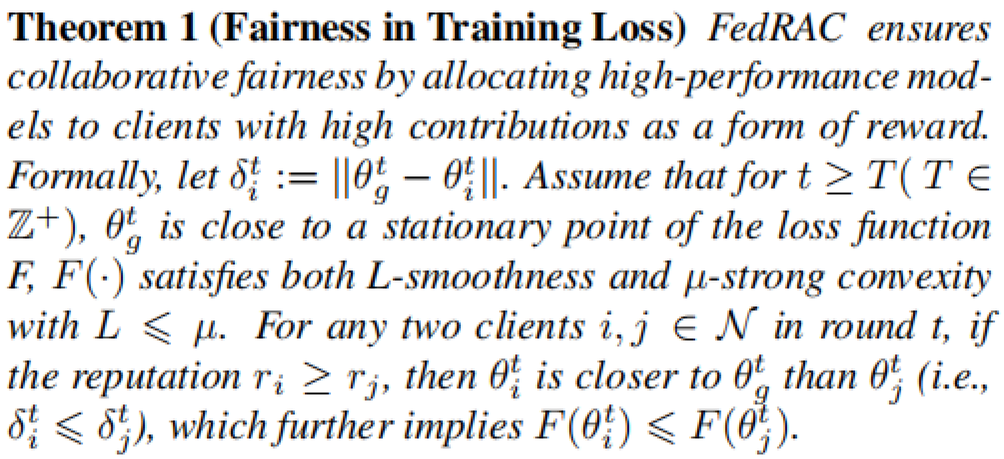


Figure 6: FedRAC overall architecture

The overall framework of FedRAC that achieves α-BCF by maintaining consistency across local models. FedRAC consists of two module:

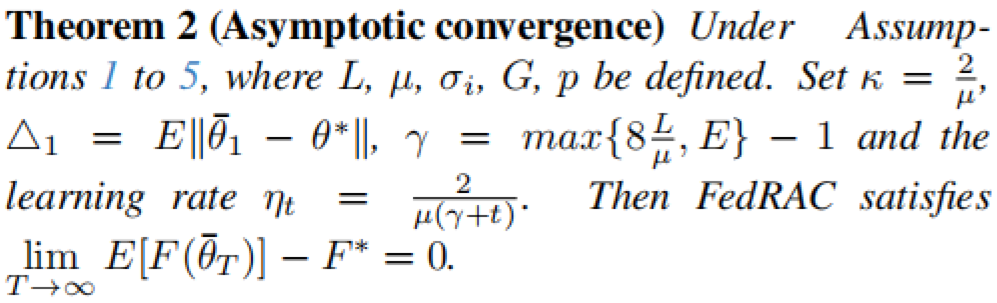
1) **dynamic Reputation Calculation module** computes real-time client reputations by integrating local performance and global model progress, enabling accurate contribution-aware incentives;
2) **Rolling Submodel Allocation module** maintains a historical neuron aggregation frequency table and prioritizes low-frequency neurons when constructing submodels. It then assigns high-performance submodels to high-reputation clients, ensuring both inter-model consistency (aligned update directions) and intra-model consistency (balanced neuron training), while significantly improving overall model performance.

Theoretical guarantee



Theorem 1 (Fairness in Training Loss) *FedRAC ensures collaborative fairness by allocating high-performance models to clients with high contributions as a form of reward. Formally, let* $\delta_i^t := \|\theta_g^t - \theta_i^t\|$. *Assume that for* $t \geq T(T \in \mathbb{Z}^+)$, θ_g^t *is close to a stationary point of the loss function F,* $F(\cdot)$ *satisfies both L-smoothness and* μ*-strong convexity with* $L \leqslant \mu$. *For any two clients* $i, j \in \mathcal{N}$ *in round t, if the reputation* $r_i \geq r_j$*, then* θ_i^t *is closer to* θ_g^t *than* θ_j^t *(i.e.,* $\delta_i^t \leqslant \delta_j^t$*), which further implies* $F(\theta_i^t) \leqslant F(\theta_j^t)$.

If client *i* holds a higher reputation than client *j* ($r_i \geq r_j$), and the submodel θ_i^t obtained by client *i* encompasses the submodel θ_j^t obtained by client *j* ($\theta_i^t \in \theta_j^t \in \theta_g^t$). Then, the submodel θ_j^t obtained by client *i* will exhibit closer alignment with the aggregated model θ_g^t in round *t*



Theorem 2 (Asymptotic convergence) *Under Assumptions 1 to 5, where* L, μ, σ_i, G, p *be defined. Set* $\kappa = \frac{2}{\mu}$, $\triangle_1 = E\|\bar{\theta}_1 - \theta^*\|$, $\gamma = max\{8\frac{L}{\mu}, E\} - 1$ *and the learning rate* $\eta_t = \frac{2}{\mu(\gamma+t)}$. *Then FedRAC satisfies* $\lim_{T \to \infty} E[F(\bar{\theta}_T)] - F^* = 0$.

To guarantee convergence to the global optimum, we make the assumption that each neuron in the aggregated model is equally allocated over *T* rounds

Experiment(1): Datasets & Baselines

1. Metric

- Fairness
- Test accuracy
- Rate

2. Dataset

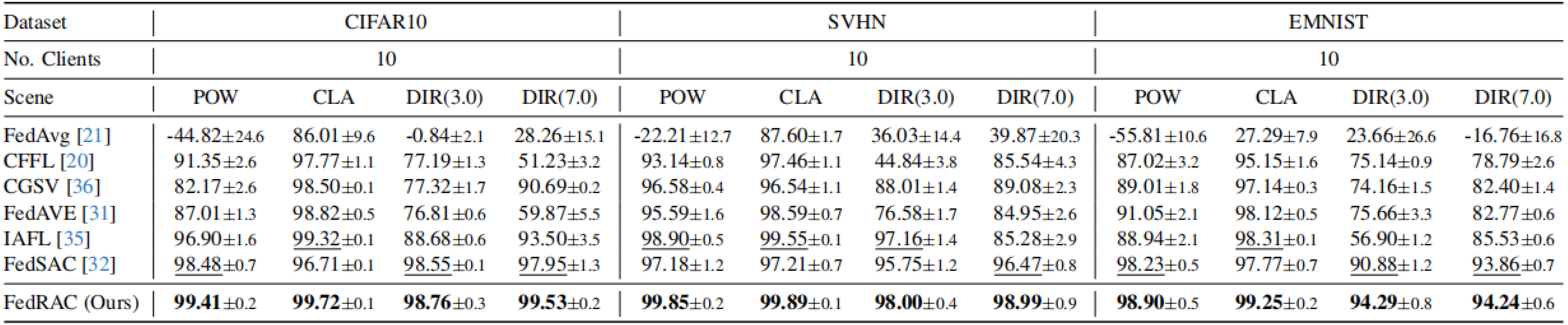
- CIFAR-10 [3], SVHN [4], EMNIST [5], and Tiny-ImageNet[6]
- The non-I.I.D. constructions follows [7-9]

3. Baseline

- FedAvg [10], CFFL [11], CGSV [12], FedAVE [13], IAFL[14], and FedSAC[15]

[3] Alex Krizhevsky, et al. 2009. Learning multiple layers of features from tiny images. Master's thesis, Department of Computer Science, University of Toronto, 2009. (2009)
[4] Yuval Netzer, et al. 2011. Reading digits in natural images with unsupervised feature learning. In NIPS Workshop on Deep Learning and Unsupervised Feature Learning 2011 (2011)
[5] Gregory Cohen, Saeed Afshar, Jonathan Tapson, and Andre van Schaik. Emnist: an extension of mnist to handwritten letters. arXiv preprint arXiv:1702.05373, 2017.
[6] Yann Le, Xuan Yang, et al. Tiny imagenet visual recognition challenge. CS 231N, 7(7):3, 2015.
[7] Ang Li, et al. 2021. Fedmask: Joint computation and communication-efficient personalized federated learning via heterogeneous masking. In Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems. 42–55.
[8] Tao Lin, et al. 2020. Ensemble distillation for robust model fusion in federated learning. Advances in Neural Information Processing Systems 33 (2020), 2351–2363.
[9] Md Palash Uddin, et al. 2020. Mutual information driven federated learning. IEEE Transactions on Parallel and Distributed Systems 32, 7 (2020), 1526–1538.
[10] Brendan McMahan, et al. 2017. Communication-efficient learning of deep networks from decentralized data. In Artificial Intelligence and Statistics. PMLR, 1273–1282.
[11] Lingjuan Lyu, et al. 2020. Collaborative fairness in federated learning. In Federated Learning. Springer, 189–204.
[12] Xinyi Xu, et al. 2021. Gradient driven rewards to guarantee fairness in collaborative machine learning. Advances in Neural Information Processing Systems 34 (2021), 16104–16117.
[13] Zihui Wang, et al. 2024. FedAVE: Adaptive data value evaluation framework for collaborative fairness in federated learning. Neurocomputing (2024), 127227.
[14] Zhaoxuan Wu, Mohammad Mohammadi Amiri, Ramesh Raskar, and Bryan Kian Hsiang Low. Incentive-aware federated learning with training-time model rewards. In The Twelfth International Conference on Learning Representations, 2024.
[15] Zihui Wang, Zheng Wang, Lingjuan Lyu, Zhaopeng Peng, Zhicheng Yang, Chenglu Wen, Rongshan Yu, Cheng Wang, and Xiaoliang Fan. Fedsac: Dynamic submodel allocation for collaborative fairness in federated learning. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pages 3299–3310, 2024.

Result (1)：Fairness



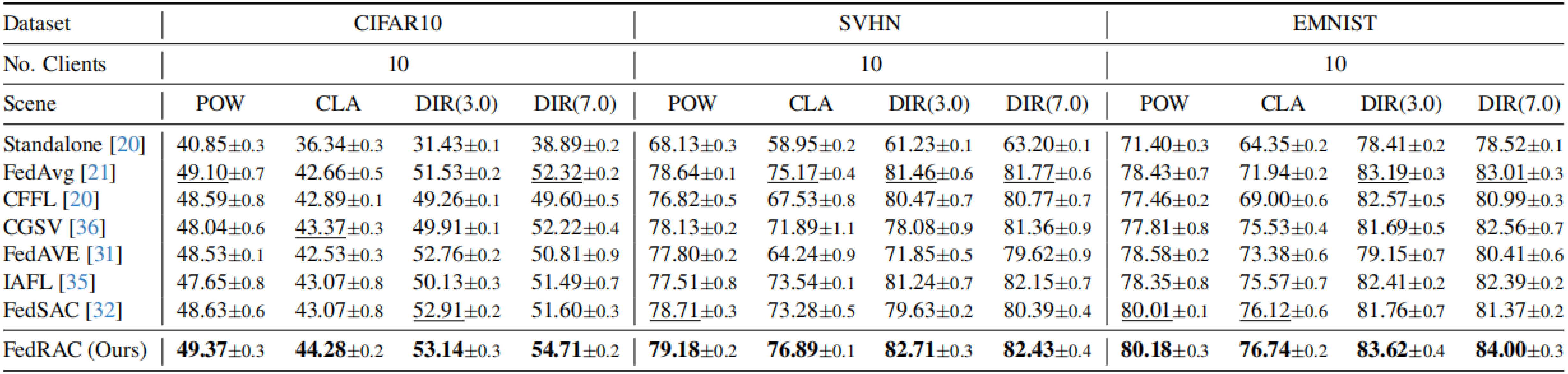
Dataset	CIFAR10				SVHN				EMNIST			
No. Clients	10				10				10			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)	POW	CLA	DIR(3.0)	DIR(7.0)	POW	CLA	DIR(3.0)	DIR(7.0)
FedAvg [21]	-44.82±24.6	86.01±9.6	-0.84±2.1	28.26±15.1	-22.21±12.7	87.60±1.7	36.03±14.4	39.87±20.3	-55.81±10.6	27.29±7.9	23.66±26.6	-16.76±16.8
CFFL [20]	91.35±2.6	97.77±1.1	77.19±1.3	51.23±3.2	93.14±0.8	97.46±1.1	44.84±3.8	85.54±4.3	87.02±3.2	95.15±1.6	75.14±0.9	78.79±2.6
CGSV [36]	82.17±2.6	98.50±0.1	77.32±1.7	90.69±0.2	96.58±0.4	96.54±1.1	88.01±1.4	89.08±2.3	89.01±1.8	97.14±0.3	74.16±1.5	82.40±1.4
FedAVE [31]	87.01±1.3	98.82±0.5	76.81±0.6	59.87±5.5	95.59±1.6	98.59±0.7	76.58±1.7	84.95±2.6	91.05±2.1	98.12±0.5	75.66±3.3	82.77±0.6
IAFL [35]	96.90±1.6	99.32±0.1	88.68±0.6	93.50±3.5	98.90±0.5	99.55±0.1	97.16±1.4	85.28±2.9	88.94±2.1	98.31±0.1	56.90±1.2	85.53±0.6
FedSAC [32]	98.48±0.7	96.71±0.1	98.55±0.1	97.95±1.3	97.18±1.2	97.21±0.7	95.75±1.2	96.47±0.8	98.23±0.5	97.77±0.7	90.88±1.2	93.86±0.7
FedRAC (Ours)	**99.41**±0.2	**99.72**±0.1	**98.76**±0.3	**99.53**±0.2	**99.85**±0.2	**99.89**±0.1	**98.00**±0.4	**98.99**±0.9	**98.90**±0.5	**99.25**±0.2	**94.29**±0.8	**94.24**±0.6

Table 1. **Comparison results of fairness** $\rho \in [-100, 100]$ with different methods on three datasets.

Question: How does the fairness of our FedRAC compare to various state-of-the-art (SOTA) methods?

Answer: Table 1 demonstrates that the proposed FedRAC **outperforms** the state-of-the-art approaches in fairness, and validated the effectiveness of our method: high-contributing clients obtain high-performance models.

Result (2)：Test accuracy



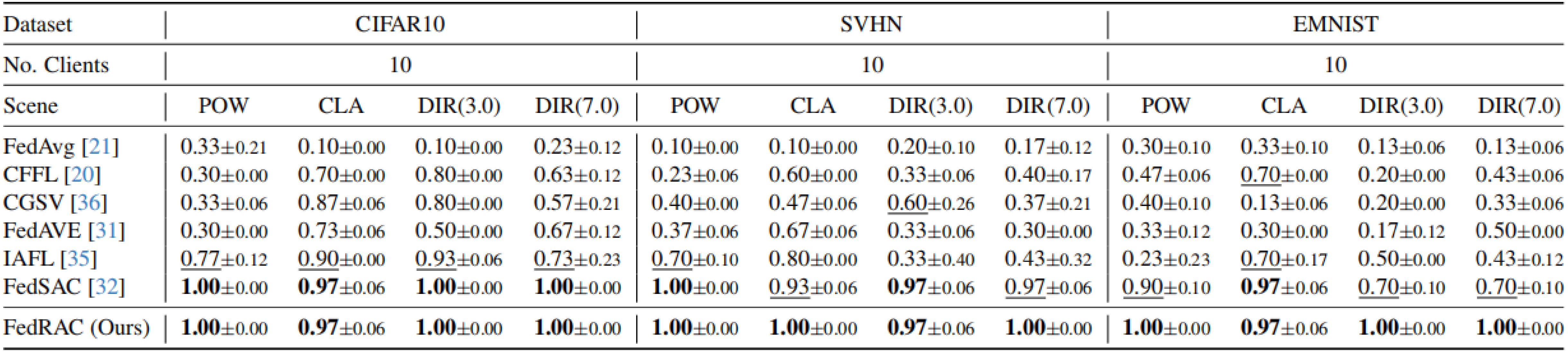
Dataset	CIFAR10				SVHN				EMNIST			
No. Clients	10				10				10			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)	POW	CLA	DIR(3.0)	DIR(7.0)	POW	CLA	DIR(3.0)	DIR(7.0)
Standalone [20]	40.85±0.3	36.34±0.3	31.43±0.1	38.89±0.2	68.13±0.3	58.95±0.2	61.23±0.1	63.20±0.1	71.40±0.3	64.35±0.2	78.41±0.2	78.52±0.1
FedAvg [21]	49.10±0.7	42.66±0.5	51.53±0.2	52.32±0.2	78.64±0.1	75.17±0.4	81.46±0.6	81.77±0.6	78.43±0.7	71.94±0.2	83.19±0.3	83.01±0.3
CFFL [20]	48.59±0.8	42.89±0.1	49.26±0.1	49.60±0.5	76.82±0.5	67.53±0.8	80.47±0.7	80.77±0.7	77.46±0.2	69.00±0.6	82.57±0.5	80.99±0.3
CGSV [36]	48.04±0.6	43.37±0.3	49.91±0.1	52.22±0.4	78.13±0.2	71.89±1.1	78.08±0.9	81.36±0.9	77.81±0.8	75.53±0.4	81.69±0.5	82.56±0.7
FedAVE [31]	48.53±0.1	42.53±0.3	52.76±0.2	50.81±0.9	77.80±0.2	64.24±0.9	71.85±0.5	79.62±0.9	78.58±0.2	73.38±0.6	79.15±0.7	80.41±0.6
IAFL [35]	47.65±0.8	43.07±0.8	50.13±0.3	51.49±0.7	77.51±0.8	73.54±0.1	81.24±0.7	82.15±0.7	78.35±0.8	75.57±0.7	82.41±0.2	82.39±0.2
FedSAC [32]	48.63±0.6	43.07±0.8	52.91±0.2	51.60±0.3	78.71±0.3	73.28±0.5	79.63±0.2	80.39±0.4	80.01±0.1	76.12±0.6	81.76±0.7	81.37±0.2
FedRAC (Ours)	**49.37**±0.3	**44.28**±0.2	**53.14**±0.3	**54.71**±0.2	**79.18**±0.2	**76.89**±0.1	**82.71**±0.3	**82.43**±0.4	**80.18**±0.3	**76.74**±0.2	**83.62**±0.4	**84.00**±0.3

Table 2. **Comparison results of the maximum test accuracy(%)** with different methods on three datasets.

Question: How does the predictive accuracy of FedRAC differ from SOTA methods on multiple datasets?

Answer: Table 2 shows that the proposed FedRAC outperforms all baselines in terms of accuracy

Result (3)： Rate



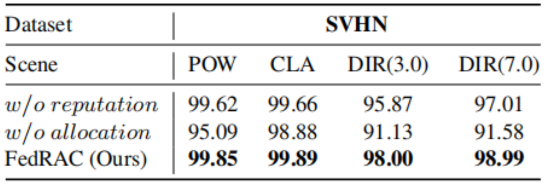
Dataset	CIFAR10				SVHN				EMNIST			
No. Clients	10				10				10			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)	POW	CLA	DIR(3.0)	DIR(7.0)	POW	CLA	DIR(3.0)	DIR(7.0)
FedAvg [21]	0.33±0.21	0.10±0.00	0.10±0.00	0.23±0.12	0.10±0.00	0.10±0.00	0.20±0.10	0.17±0.12	0.30±0.10	0.33±0.10	0.13±0.06	0.13±0.06
CFFL [20]	0.30±0.00	0.70±0.00	0.80±0.00	0.63±0.12	0.23±0.06	0.60±0.00	0.33±0.06	0.40±0.17	0.47±0.06	0.70±0.00	0.20±0.00	0.43±0.06
CGSV [36]	0.33±0.06	0.87±0.06	0.80±0.00	0.57±0.21	0.40±0.00	0.47±0.06	0.60±0.26	0.37±0.21	0.40±0.10	0.13±0.06	0.20±0.00	0.33±0.06
FedAVE [31]	0.30±0.00	0.73±0.06	0.50±0.00	0.67±0.12	0.37±0.06	0.67±0.06	0.33±0.06	0.30±0.00	0.33±0.12	0.30±0.00	0.17±0.12	0.50±0.00
IAFL [35]	0.77±0.12	0.90±0.00	0.93±0.06	0.73±0.23	0.70±0.10	0.80±0.00	0.33±0.40	0.43±0.32	0.23±0.23	0.70±0.17	0.50±0.00	0.43±0.12
FedSAC [32]	**1.00**±0.00	**0.97**±0.06	**1.00**±0.00	**1.00**±0.00	**1.00**±0.00	0.93±0.06	**0.97**±0.06	0.97±0.06	0.90±0.10	**0.97**±0.06	0.70±0.10	0.70±0.10
FedRAC (Ours)	**1.00**±0.00	**0.97**±0.06	**1.00**±0.00	**1.00**±0.00	**1.00**±0.00	**1.00**±0.00	**0.97**±0.06	**1.00**±0.00	**1.00**±0.00	**0.97**±0.06	**1.00**±0.00	**1.00**±0.00

Table 3. **Comparison results of rate** with different methods on three datasets.

Question: How does the proportion of clients lying within the $\boldsymbol{\alpha}-$BCF interval (i.e., the rate) compare with that obtained by SOTA methods on different datasets?

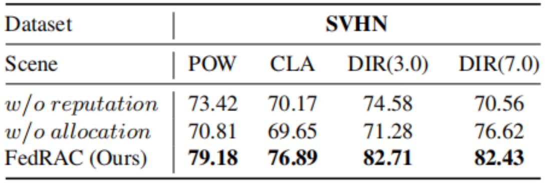
Answer: Table 3 shows that the proposed FedRAC outperforms all baseline methods in terms of Rate

Result (4)：Ablation studies



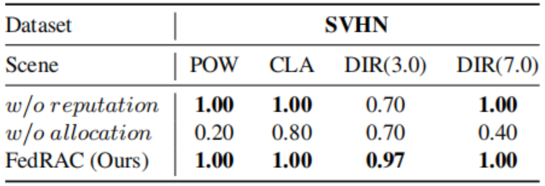
Dataset	SVHN			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)
w/o reputation	99.62	99.66	95.87	97.01
w/o allocation	95.09	98.88	91.13	91.58
FedRAC (Ours)	**99.85**	**99.89**	**98.00**	**98.99**

Table 4. Ablation studies on FedRAC for the fairness on SVHN.



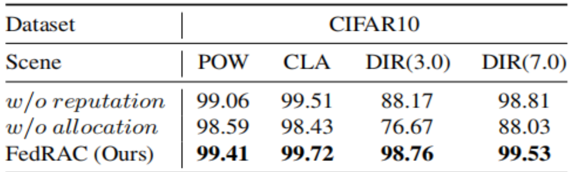
Dataset	SVHN			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)
w/o reputation	73.42	70.17	74.58	70.56
w/o allocation	70.81	69.65	71.28	76.62
FedRAC (Ours)	**79.18**	**76.89**	**82.71**	**82.43**

Table 5. Ablation studies on FedRAC for the maximum test accuracy on SVHN.



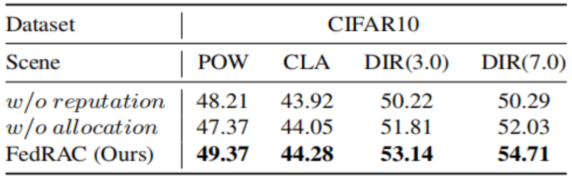
Dataset	SVHN			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)
w/o reputation	**1.00**	**1.00**	0.70	**1.00**
w/o allocation	0.20	0.80	0.70	0.40
FedRAC (Ours)	**1.00**	**1.00**	**0.97**	**1.00**

Table 6. Ablation studies on FedRAC for the rate on SVHN.



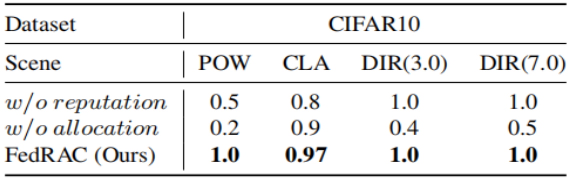
Dataset	CIFAR10			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)
w/o reputation	99.06	99.51	88.17	98.81
w/o allocation	98.59	98.43	76.67	88.03
FedRAC (Ours)	**99.41**	**99.72**	**98.76**	**99.53**

Table 7. Ablation studies on FedRAC for the fairness on CIFAR10.



Dataset	CIFAR10			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)
w/o reputation	48.21	43.92	50.22	50.29
w/o allocation	47.37	44.05	51.81	52.03
FedRAC (Ours)	**49.37**	**44.28**	**53.14**	**54.71**

Table 8. Ablation studies on FedRAC for the maximum test accuracy on CIFAR10.



Dataset	CIFAR10			
Scene	POW	CLA	DIR(3.0)	DIR(7.0)
w/o reputation	0.5	0.8	1.0	1.0
w/o allocation	0.2	0.9	0.4	0.5
FedRAC (Ours)	**1.0**	**0.97**	**1.0**	**1.0**

Table 9. Ablation studies on FedRAC for the rate on CIFAR10.

Question: How do different components (i.e., dynamic reputation calculation module and rolling submodel allocation module) affect the fairness, the predictive model performance, and the rate?

Answer: Tables demonstrate that the two designed modules in FedRAC are crucial and significant in enhancing bounded collaborative fairness.

FedRAC: Rolling Submodel Allocation for Collaborative Fairness in Federated Learning

Zihui Wang[1,2], Yuhang Fu[1], Mengmeng Du[5], Zhimin Yuan[6],

Yachen Liu[4], Weisheng Liao[2], Kaiyu Wang[1,*] , Zheng Wang[3,4,*]

[1]Harbin Engineering University, [2]Pengcheng Laboratory, [3]Shanghai Innovation Institution,

[4]Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, China

[5]Fuyang Normal University, [6]Nanyang Normal University

wangzh08@pcl.ac.cn, yuhangfu@hrbeu.edu.cn, zwang@stu.xmu.edu.cn,wang kaiyu@hrbeu.edu.cn

Paper ID: 40217

Code: https://github.com/ZiHuiWangpcl1/FedRAC