上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

SJTU

Federated Learning Extension

施宏建

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

目录

1 Preliminary

2 Directions

3 Examples

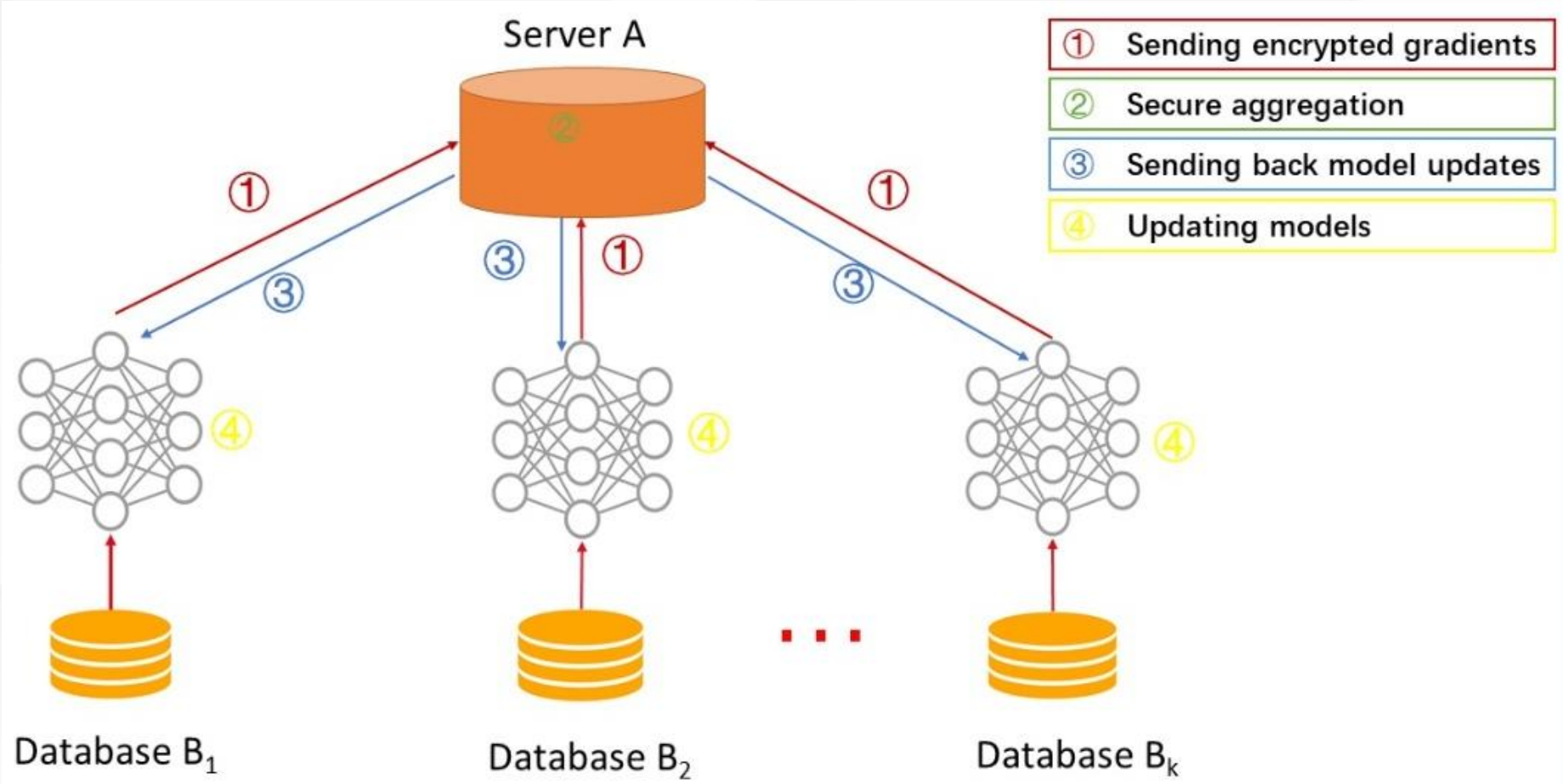
01

Preliminary

- Introduction
- Structure
- Procedure

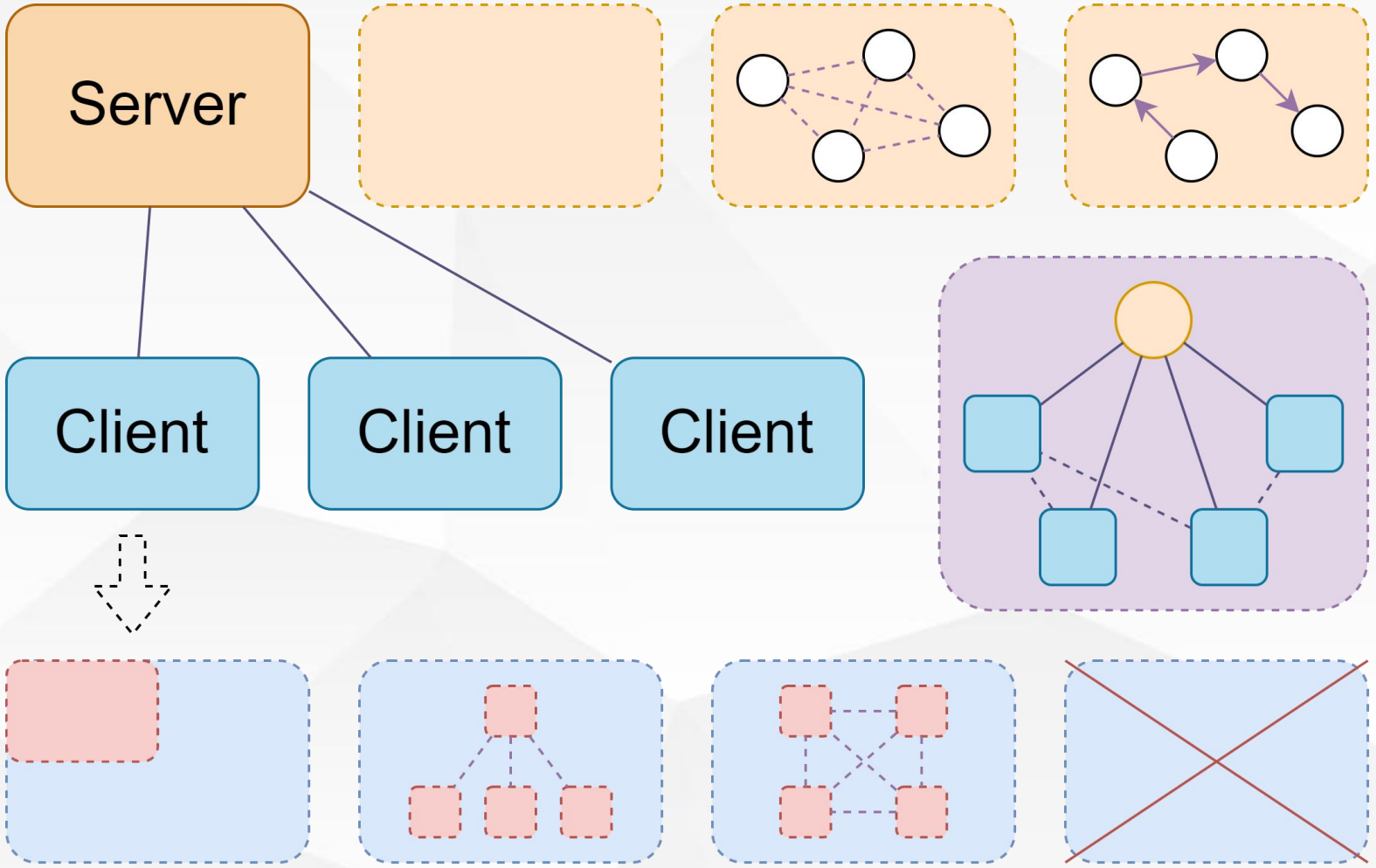
联邦学习介绍

- 联邦学习是一种分布式机器学习技术，或机器学习框架。
- 在保证**数据隐私安全**及**合法合规**的基础上，实现**共同建模**，提升AI模型的效果。
- 本质上是通过多个用户设备共同训练一个代表所有用户设备的全局模型
- 训练过程不需要用户数据的交换，更强调隐私性。



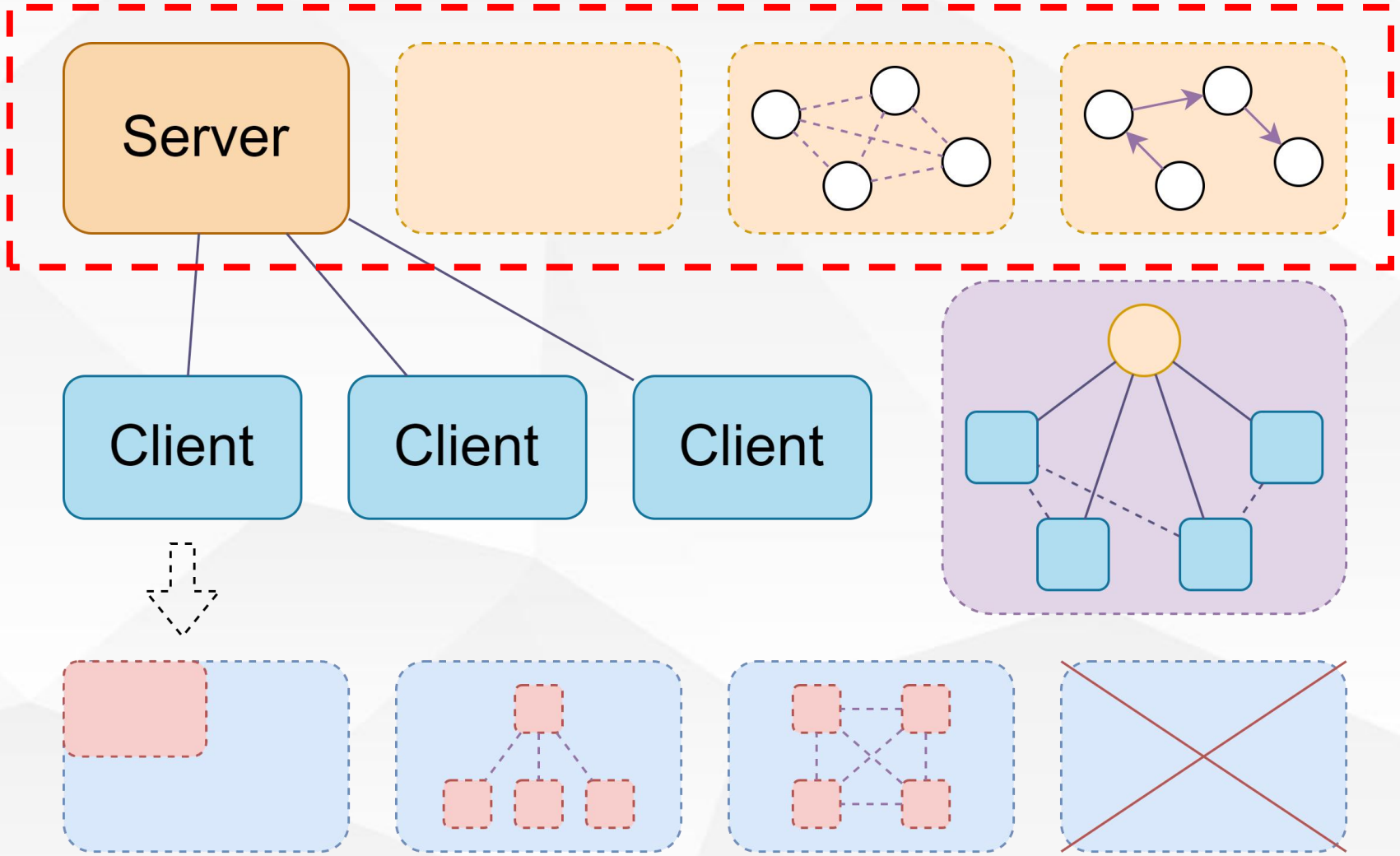
总体架构

- 服务器端
  - 有、无 - 控制方式
  - 单点、多点 - 一致性
  - 区块链 - 可信
- 网络连接
  - 去中心、中心、半中心
- 客户端
  - 资源占用 - 全部、部分
  - 拓扑结构 - 层次化
  - 计算能力 - 有无



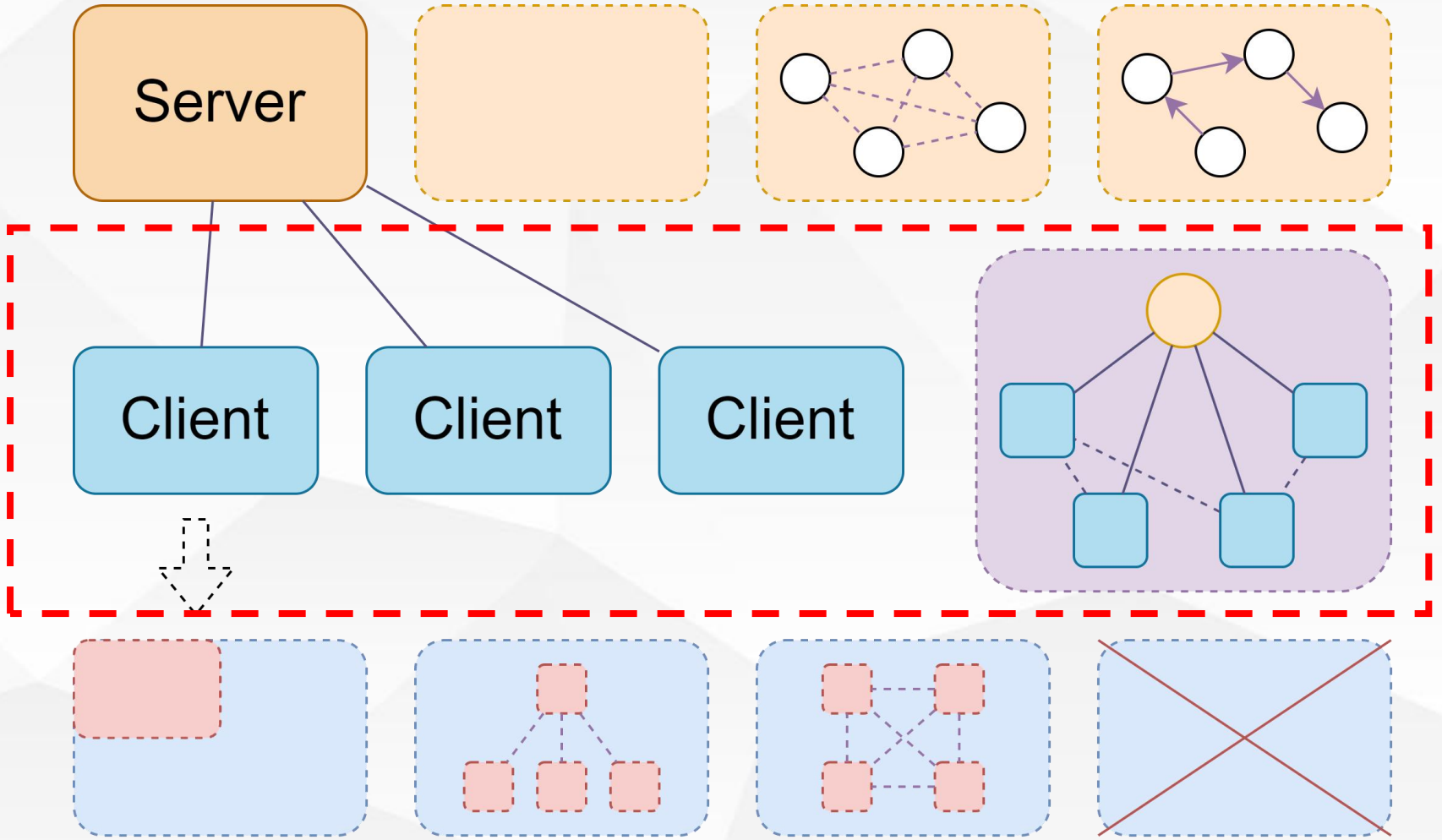
联邦学习框架-服务器端

- 控制方式
  - 中心化服务器
  - 去中心化控制
- 一致性
  - 单点服务器
  - 多点服务器集群
- 可信第三方
  - 区块链服务器



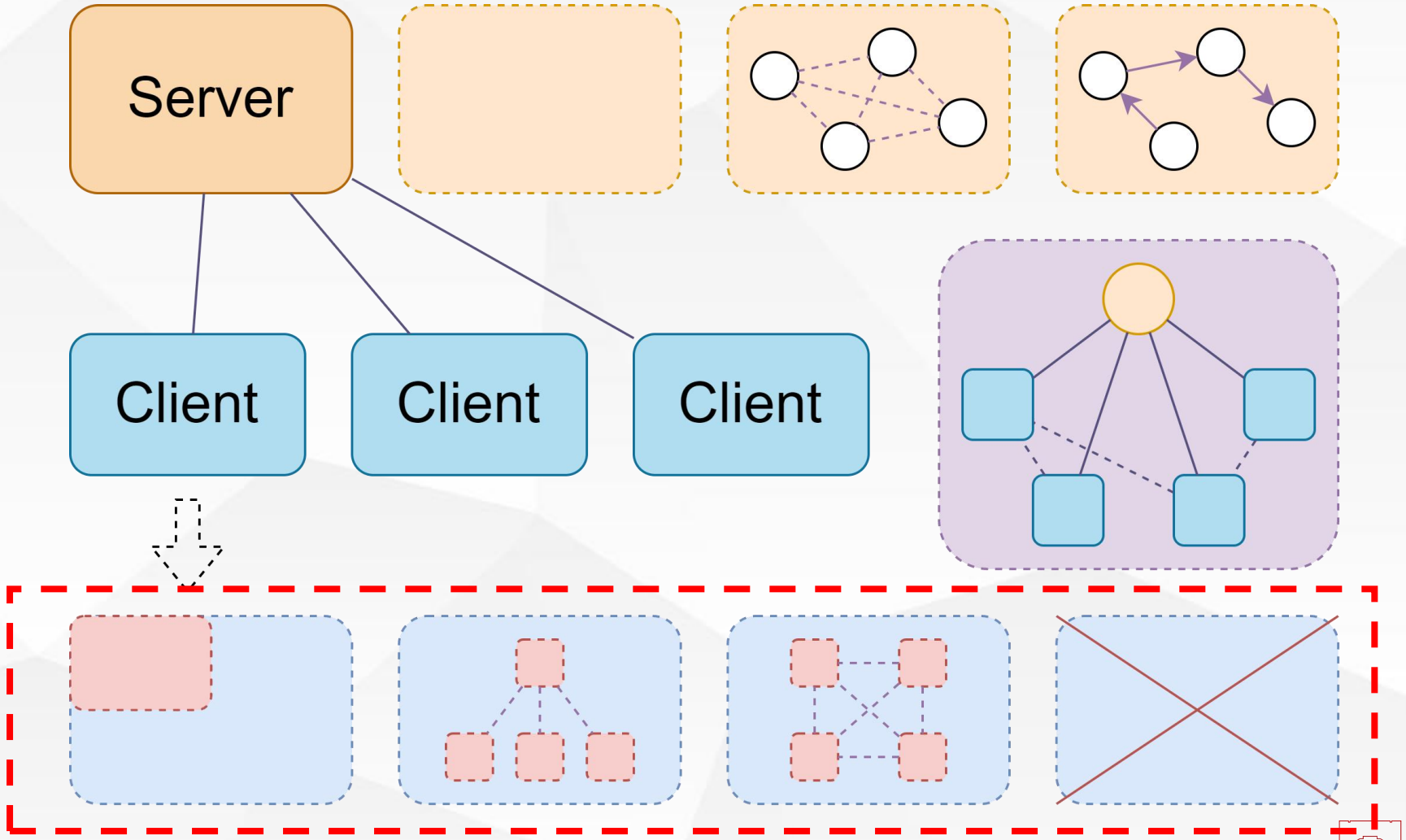
联邦学习框架-网络连接

- 中心化框架
  - 服务器端-客户端
- 去中心化框架
  - 客户端-客户端
- 半中心化框架
  - 服务器端-客户端
  - 客户端-客户端



联邦学习框架-客户端

- 资源占用
  - 虚拟机/容器
  - 资源隔离性
- 拓扑结构
  - 多层中心化
  - 集群分布式
- 设备能力
  - 低计算速度
  - 少存储空间



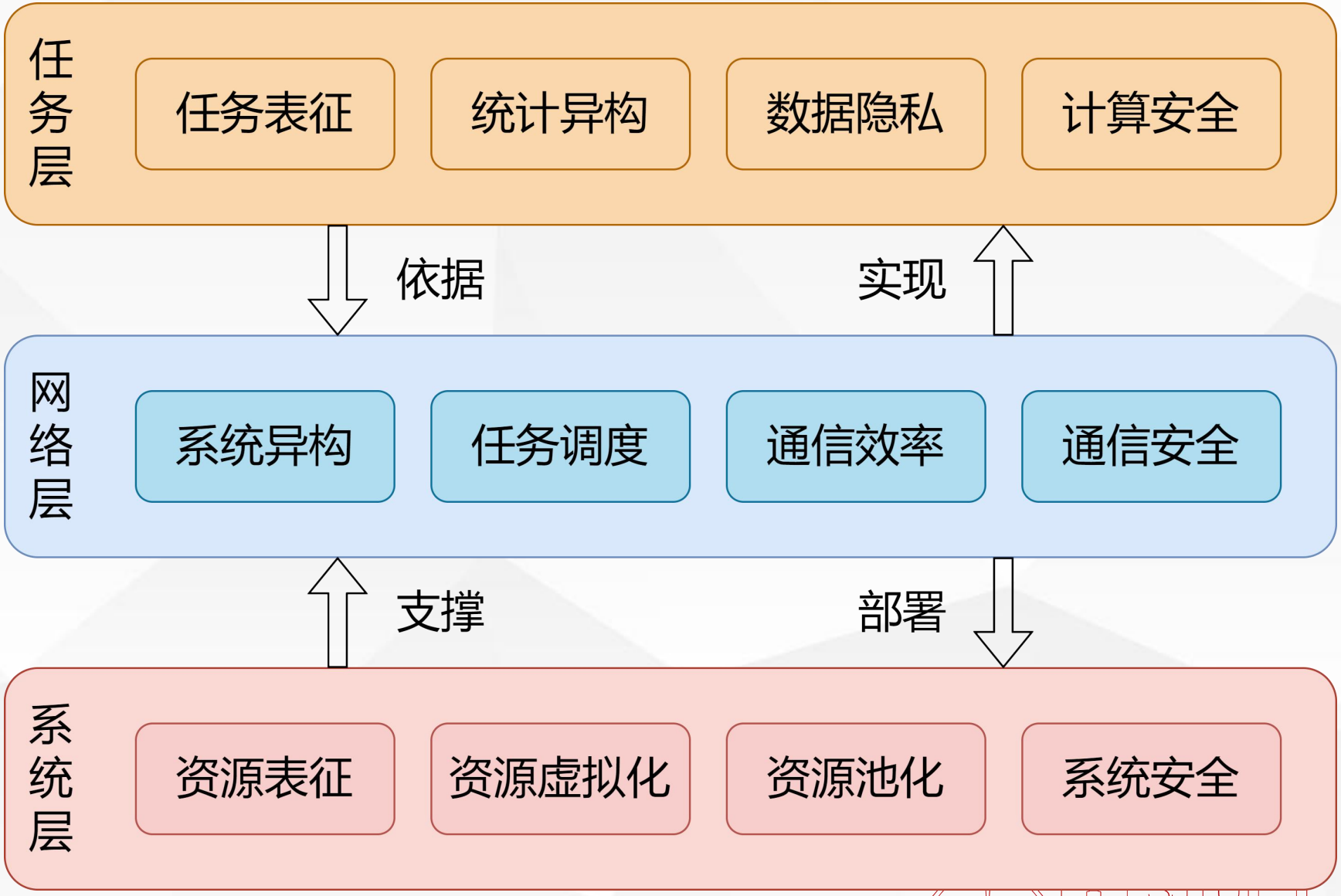
02

Directions

- Server
- Client
- Network
- Others

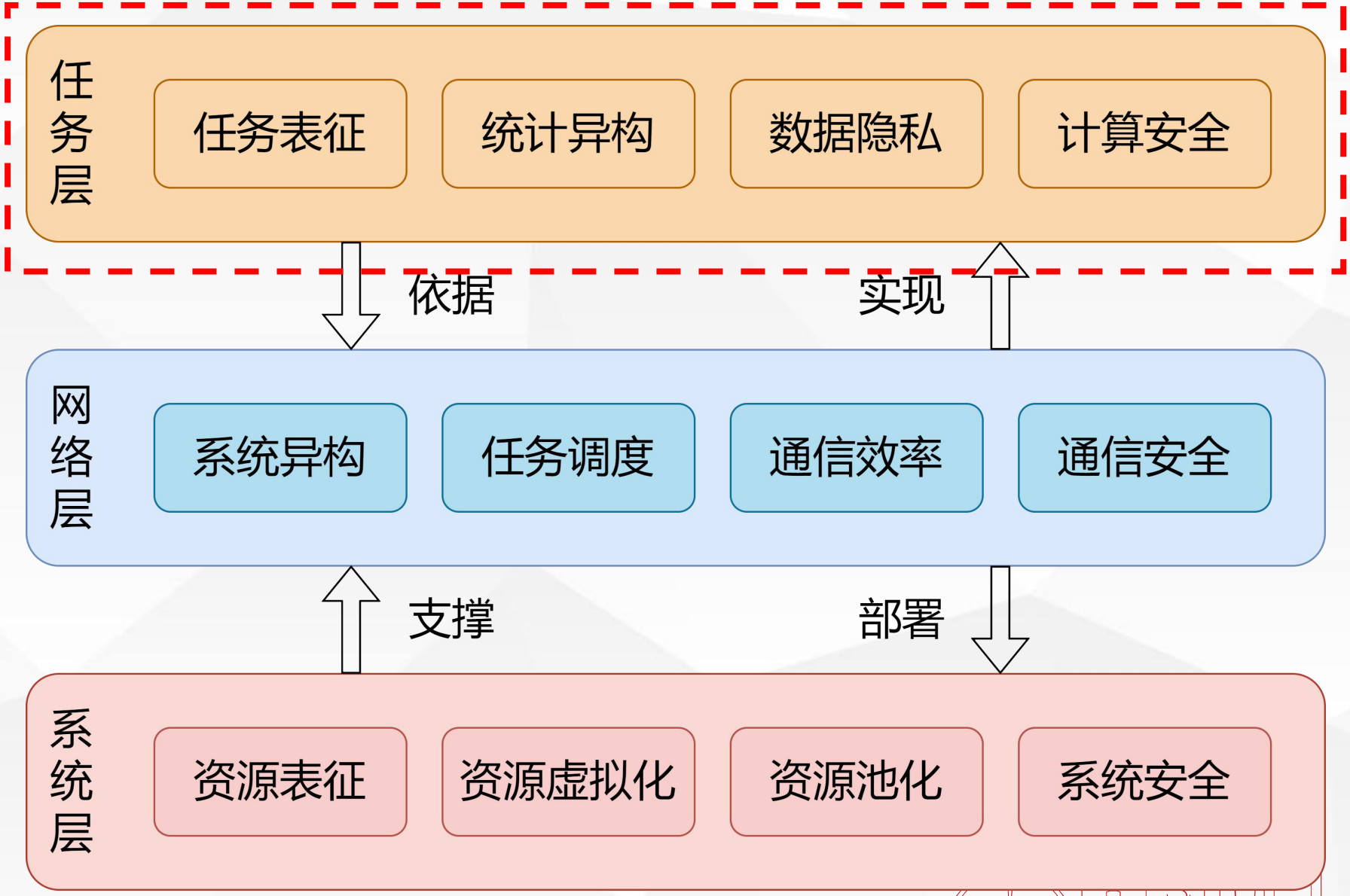
总体架构

- 任务层：上层应用角度
  - 明确用户需求
  - 应用场景建模
- 网络层：中层传输角度
  - 明确环境限制
  - 网络状态建模
- 系统层：下层资源角度
  - 明确硬件条件
  - 硬件单元建模



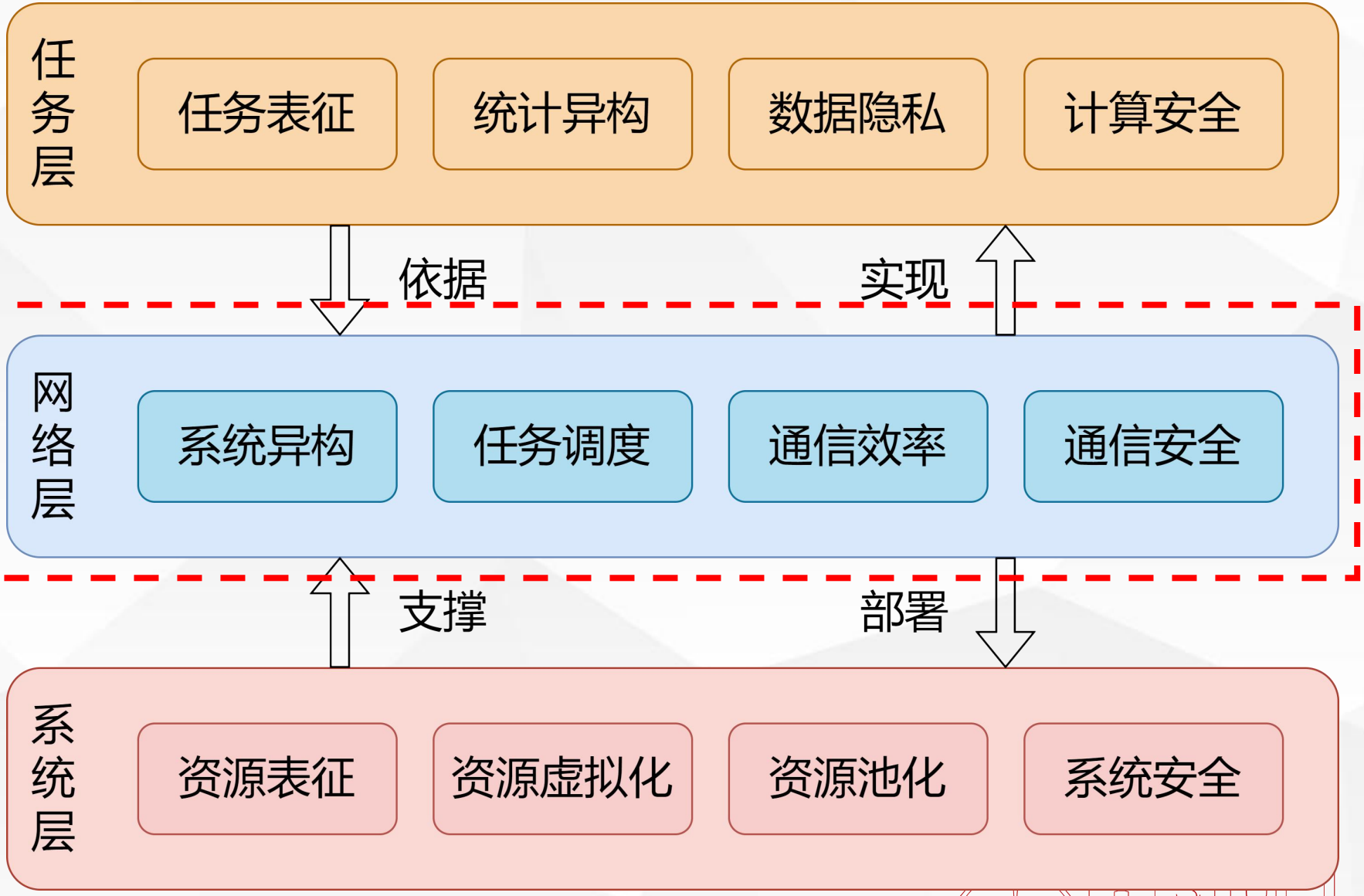
联邦学习挑战-任务层

- 上层应用带来的挑战
- **任务表征**：如何让系统理解任务需求
- **统计异构**：如何满足不同任务的需求
- **数据隐私**：如何保障任务数据不泄露
- **计算安全**：如何保障计算过程正确性



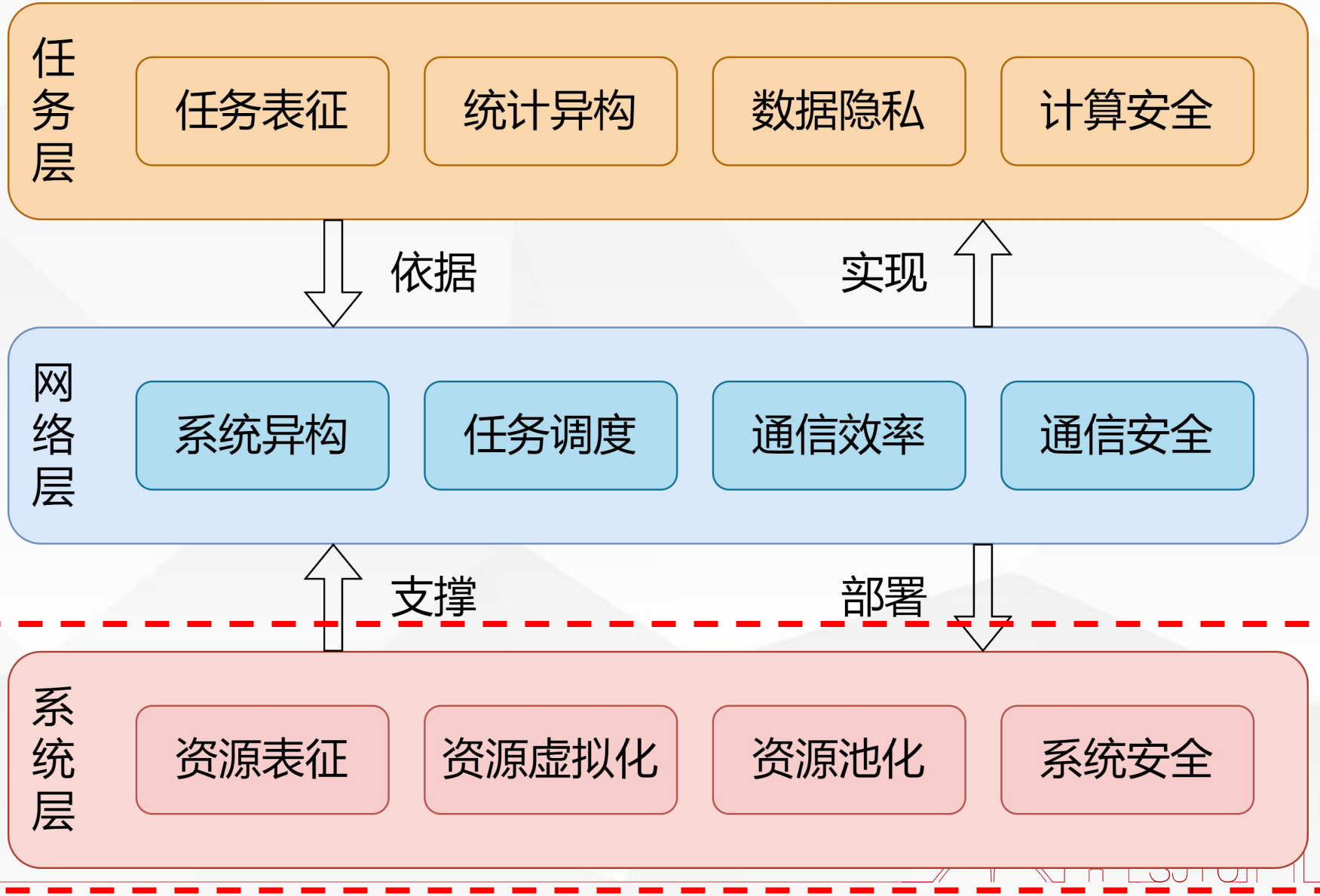
联邦学习挑战-网络层

- 网络传输带来的挑战
- **系统异构**：如何合理控制不同的设备
- **任务调度**：如何高效匹配任务和设备
- **通信效率**：如何减少通信带来的延迟
- **通信安全**：如何防止通信带来的错误

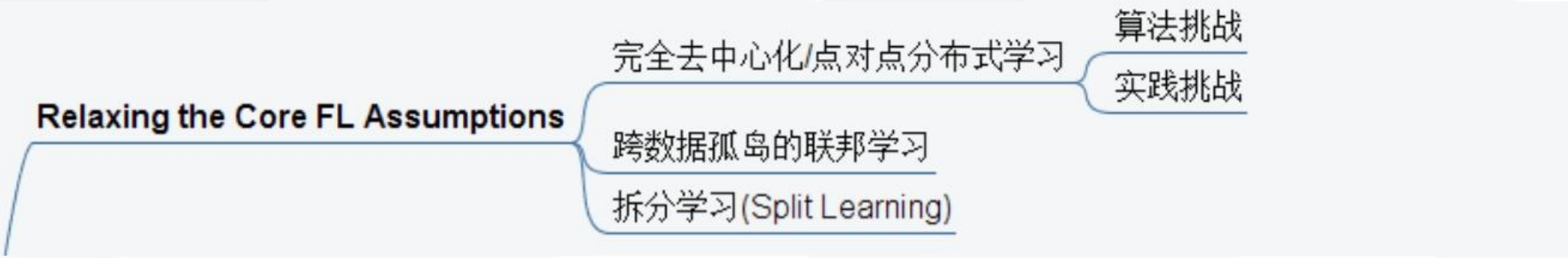


联邦学习挑战-系统层

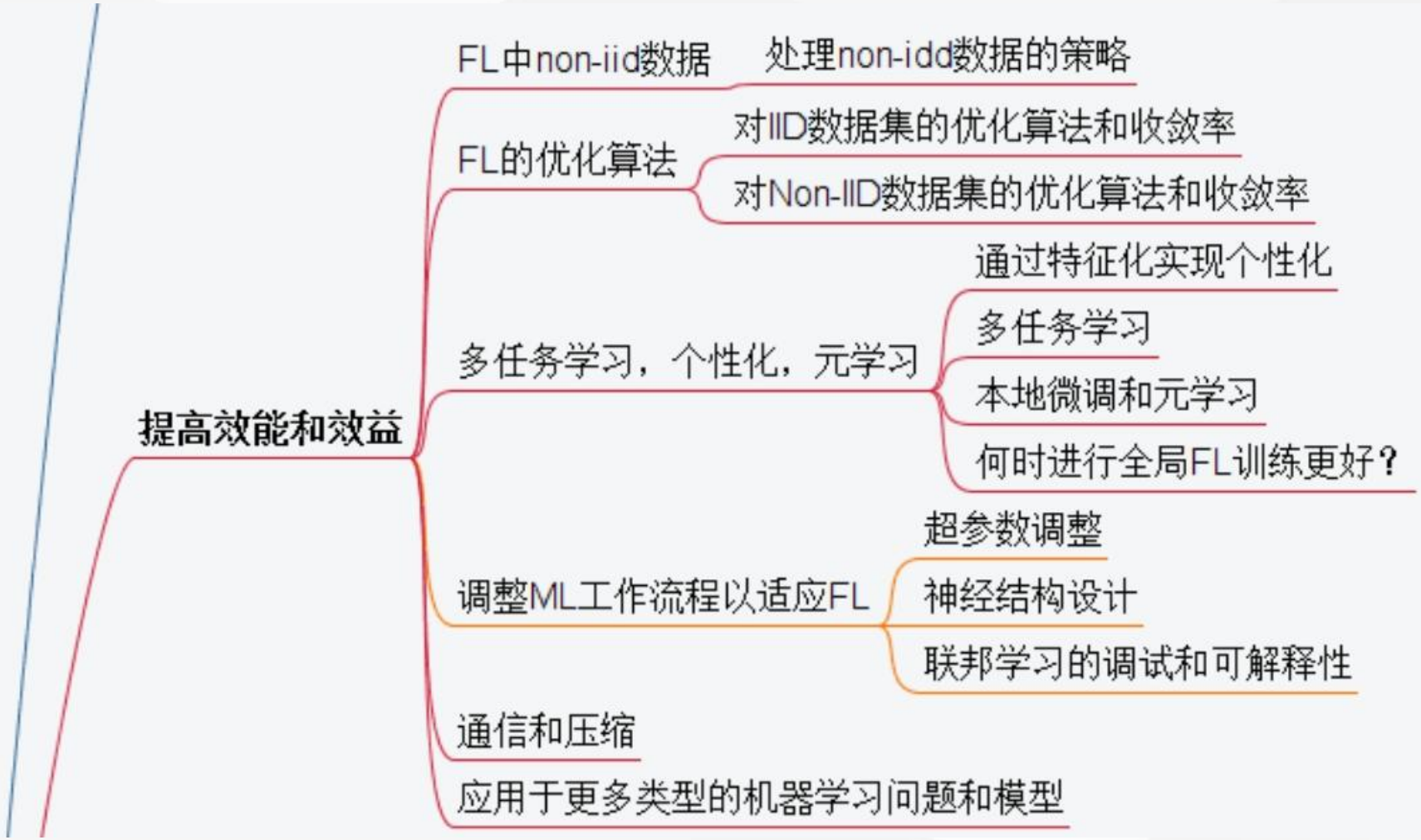
- 资源角度带来的挑战
- **资源表征**：如何让系统理解资源能力
- **资源虚拟化**：如何让系统调用资源
- **资源池化**：如何让系统统一管理资源
- **系统安全**：如何防止系统受入侵攻击



Surveys


Relaxing the Core FL Assumptions
完全去中心化/点对点分布式学习
算法挑战
实践挑战
跨数据孤岛的联邦学习
拆分学习(Split Learning)

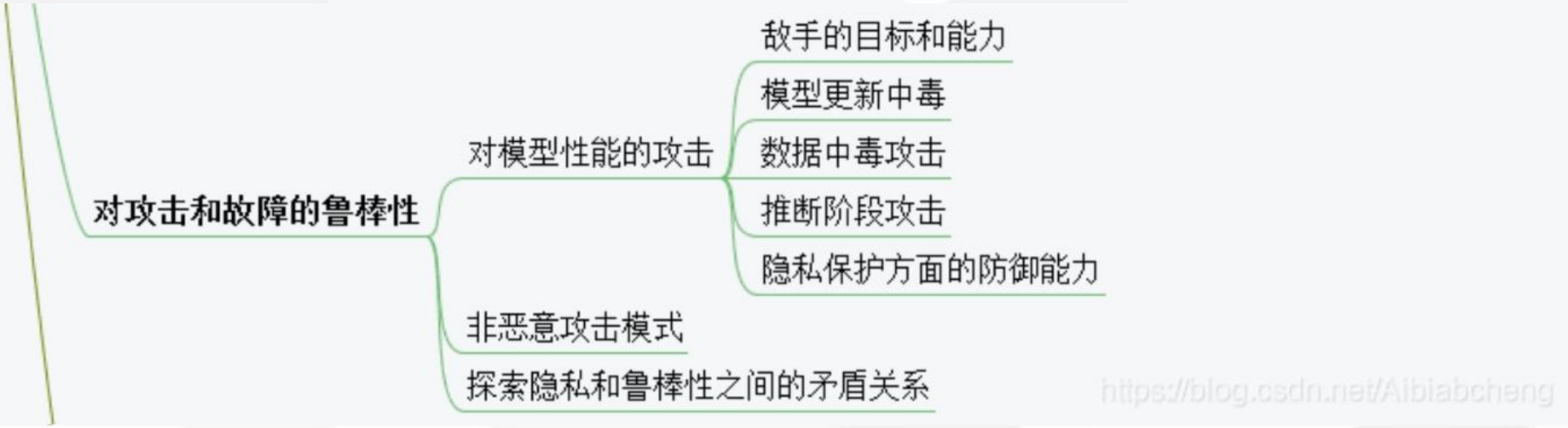
Surveys


提高效能和效益
FL中non-iid数据
处理non-idd数据的策略
FL的优化算法
对IID数据集的优化算法和收敛率
对Non-IID数据集的优化算法和收敛率
多任务学习，个性化，元学习
通过特征化实现个性化
多任务学习
本地微调和元学习
何时进行全局FL训练更好？
调整ML工作流程以适应FL
超参数调整
神经结构设计
联邦学习的调试和可解释性
通信和压缩
应用于更多类型的机器学习问题和模型

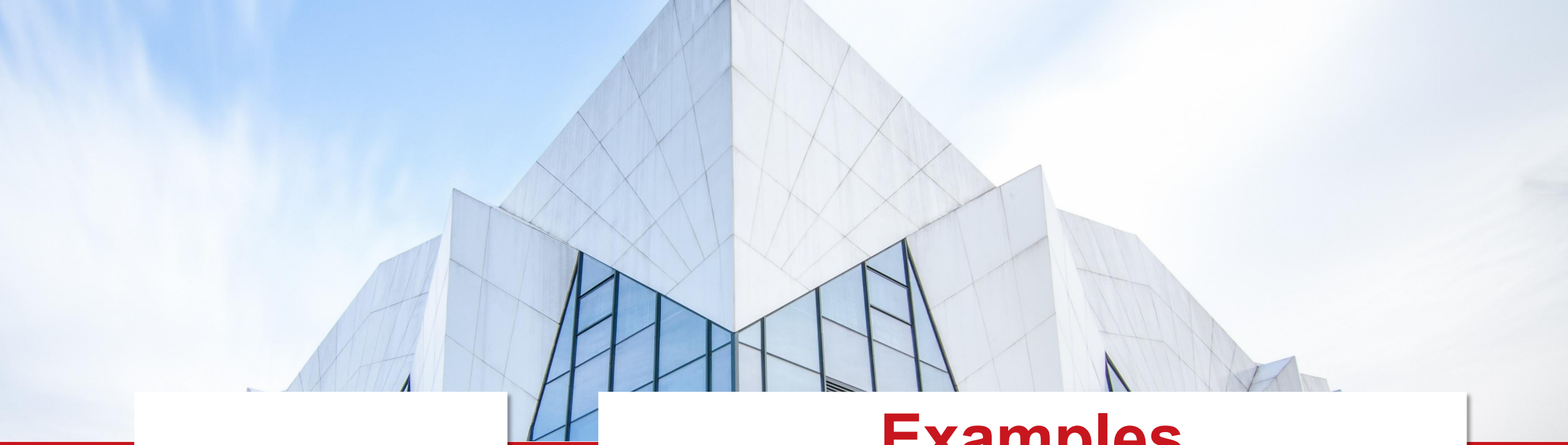
Surveys


联邦学习综述
对用户数据的隐私保护
参与者，模型威胁，与深层隐私
工具和技术
安全计算
隐私保护披露
可验证性
防范外部恶意参与者
评估迭代轮次和最终模型
考虑中心差分隐私的模型训练
迭代隐蔽
对演变数据的重复分析
防止模型被盗和误用
针对恶意服务器的保护
挑战（通信通道，Sybil Attacks(女巫攻击)）与选择
现有解决方案的缺陷
分布式差分隐私训练
子模型训练时的隐私保护
用户感知
了解特定分析任务的隐私需求
引发隐私偏好的行为研究

Surveys



Surveys



03

Examples

- Our published papers

团队代表性论文

1. Hongjian Shi, Weichu Zheng, Zifei Liu, Ruhui Ma, Haibing Guan, Automatic Pipeline Parallelism: A Parallel Inference Framework for Deep Learning Applications in 6G Mobile Communication Systems, in IEEE Journal on Selected Areas in Communications (JSAC)，2023.
2. Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Zhangcheng Lv, Xiulang Jin, Zhengui Xue, Ruhui Ma, Haibing Guan, Siren: Byzantine-robust Federated Learning via Proactive Alarming, Proceedings of ACM Symposium on Cloud Computing (SoCC)，2021.
3. Jianqing Zhang, Yang Hua, Hao Wang, Tao Song, Zhengui Xue, Ruhui Ma, Haibing Guan, FedALA: Adaptive Local Aggregation for Personalized Federated Learning, Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI), 2022.
4. Hongjian Shi, Ruhui Ma, Dongmei Li, Haibing Guan, Hierarchical Adaptive Collaborative Learning: A Distributed Learning Framework for Customized Cloud Services in 6G Mobile Systems, in IEEE Network，2023.
5. Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Ruhui Ma, Xiulang Jin, Zhengui Xue, Haibing Guan, SIREN+: Robust Federated Learning with Proactive Alarming and Differential Privacy, IEEE Transactions on Dependable and Secure Computing (TDSC), 2023.

团队代表性论文

6. 施宏建，马汝辉，管海兵，基于区块链辅助的半中心化联邦学习框架，计算机研究与发展，2023
7. Hongjian Shi, Ilyas Bayanbayev, Ruhui Ma, Haibing Guan, Cloud-based Collaborative Agricultural Learning with Flexible Model Size and Adaptive Batch Number, ACM Transactions on Sensor Networks (TOSN), 2023.
8. Hanxi Guo, Qing Yang, Hao Wang, Yang Hua, Tao Song, Ruhui Ma, and Haibing Guan. SpaceDML: Enabling Distributed Machine Learning in Space Information Networks. IEEE Network，2021.
9. Ruhui Ma, Hongjian Shi, Honghao Gao, Haibing Guan, Muddesar Iqbal, Shahid Mumtaz, cFedDT: Cross-domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products, IEEE Transactions on Consumer Electronics (TCE), 2023.

团队代表性论文

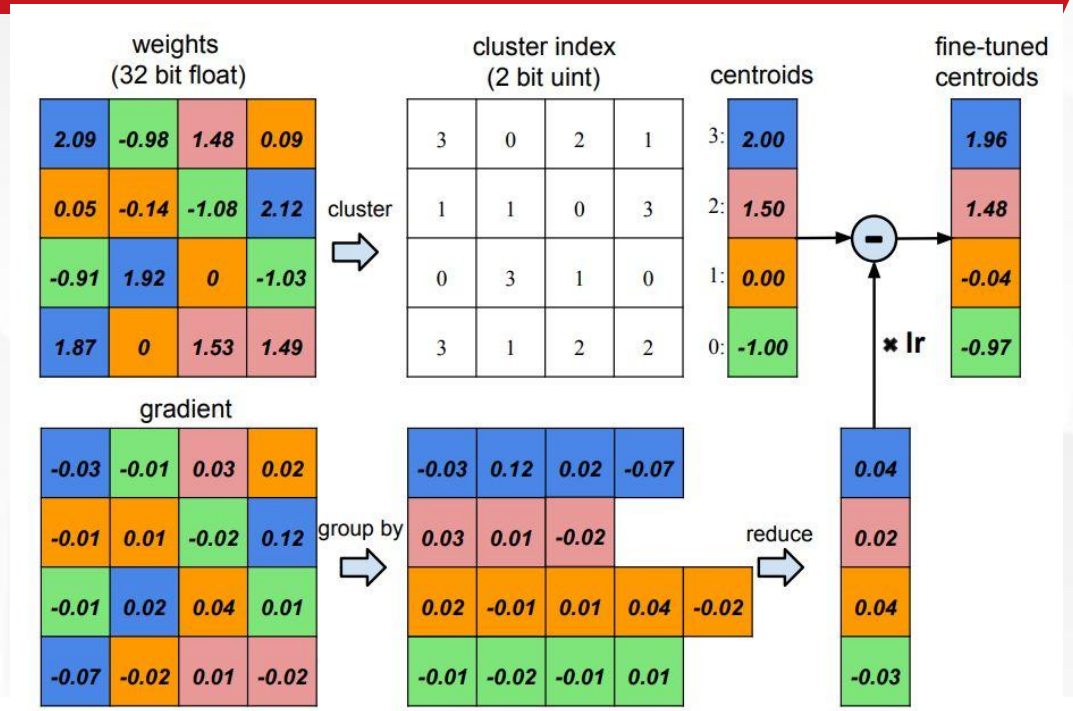
11. Yunhui Wang, Weichu Zheng, Zifei Liu, Jinyan Wang, Hongjian Shi, Minyu Gu, Ruhui Ma, and Haibing Guan, A Federated Network Intrusion Detection System with Multi-branch Network and Vertical blocking Aggregation, Electronics, 2023.
12. Hongjian Shi, Jianqing Zhang, Shuming Fan, Ruhui Ma, Haibing Guan, pFedEff: An Efficient and Personalized Federated Cognitive Learning Framework in Multi-agent Systems, IEEE Transactions on Cognitive and Development Systems (TCDS), 2023.

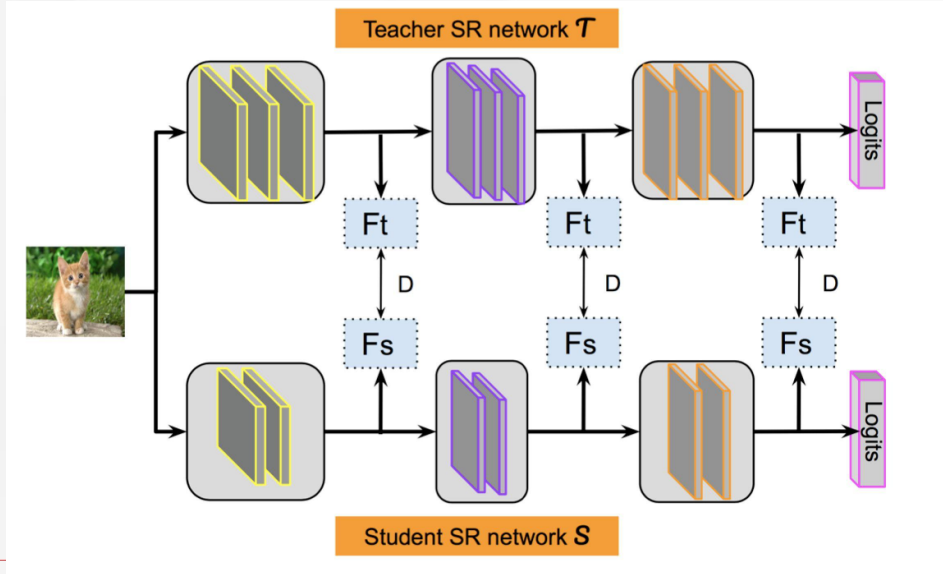
联邦学习效率

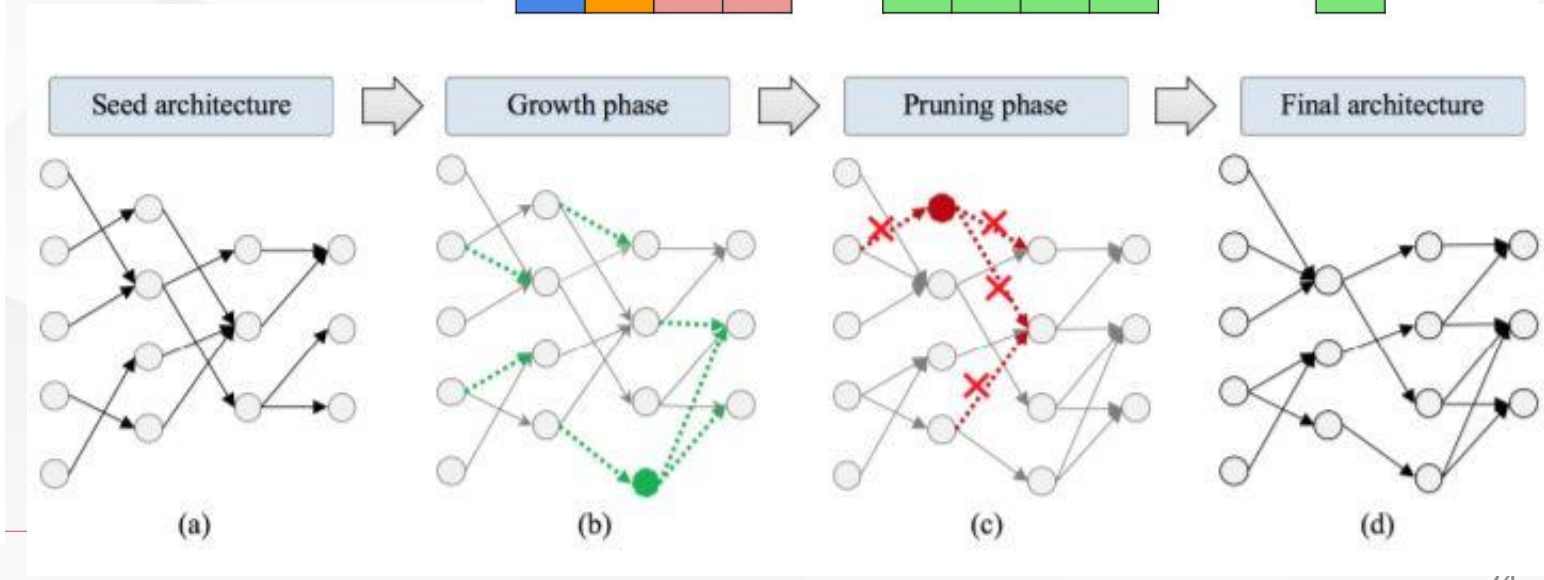
- 数据传输量大 - 模型压缩
- 网络环境复杂 - 传输调度

联邦学习效率-模型压缩

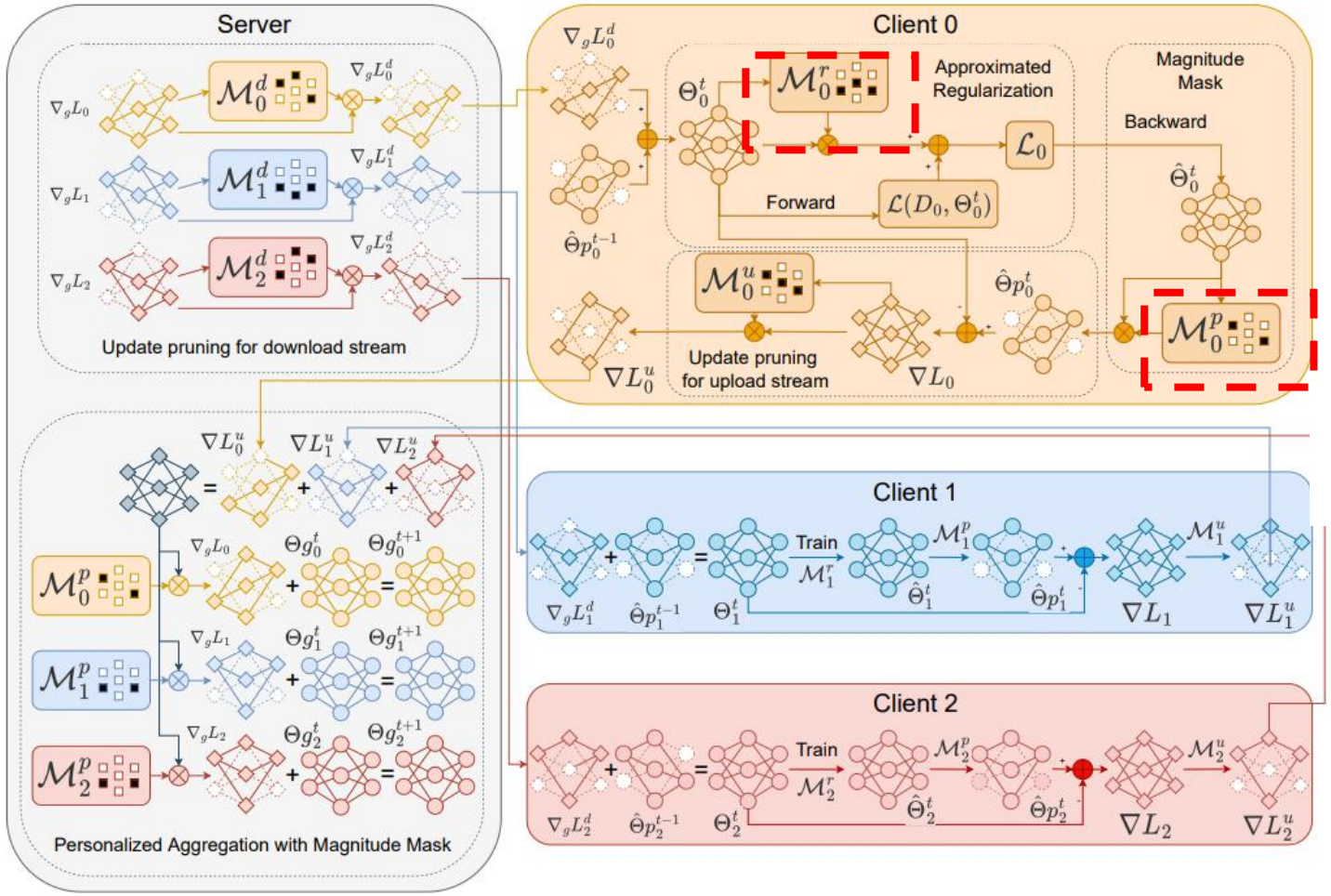
- 通信效率
  - 智能算法本身通信量大，通信开销大
- 模型压缩：**模型剪枝，模型量化，知识蒸馏**
  - 减少数据通信、计算量
  - 构造准确但小的神经网络模型







联邦学习效率



- **绝对值剪枝**

$$\mathcal{H}(w) = |w|.$$

- **损失函数正则化**
- 定位对推理结果影响较低的参数

$$\mathcal{L}_i = l(D_i, \Theta_i^t) + \lambda_1 \cdot ||\Theta_i^t \odot \mathcal{M}_i^r||_1.$$

- **训练后模型剪枝**
- 过滤对推理结果影响较低的参数

$$\hat{\Theta}p_i^t \leftarrow \hat{\Theta}_i^t \odot \mathcal{M}_i^p.$$

Hongjian Shi, Jianqing Zhang, Shuming Fan, Ruhui Ma, Haibing Guan, pFedEff: An Efficient and Personalized Federated Cognitive Learning Framework in Multi-agent Systems, IEEE Transactions on Cognitive and Development Systems, 2023.

联邦学习效率

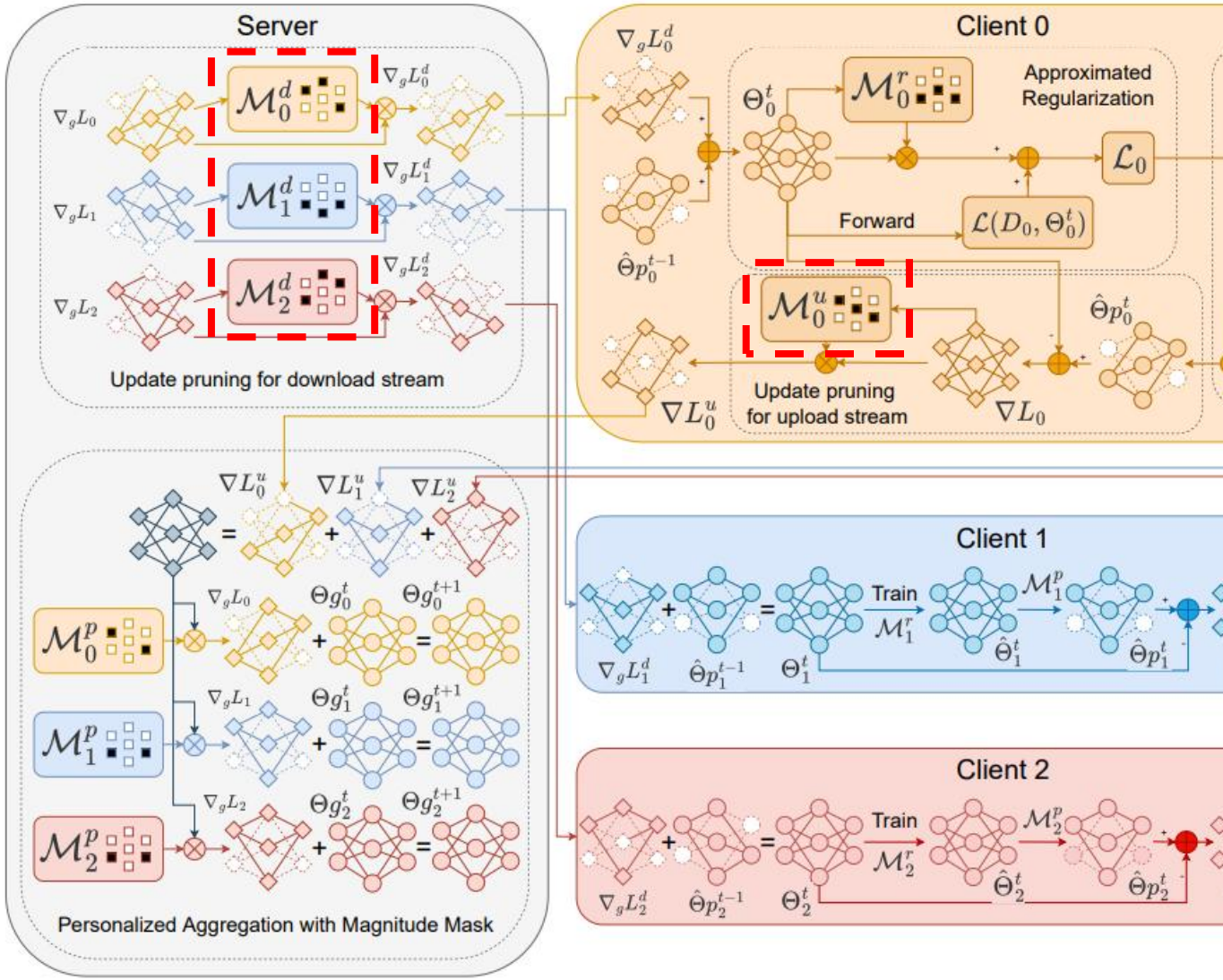
- 重要性剪枝

$$\begin{aligned}\mathcal{I}(w) &= |\mathcal{L}(D, \Theta_{w\to 0}) - \mathcal{L}(D, \Theta)| \\ &= |\frac{\partial \mathcal{L}(D,\Theta)}{\partial w} \cdot (0 - w) + o(w^2)| \\ &= |\frac{\partial \mathcal{L}(D,\Theta)}{\partial w} \cdot w|.\end{aligned}$$

- 上传/下载传输剪枝
- 过滤更新量较小的上传/下载模型更新

$$\nabla_g \mathcal{L}_i^u \leftarrow \nabla \mathcal{L}_i \odot \mathcal{M}_i^u.$$

$$\nabla \mathcal{L}_i^d = (\sum_{j=1}^{M} k_j \cdot \nabla_g \mathcal{L}_i) \odot \mathcal{M}_i^d$$



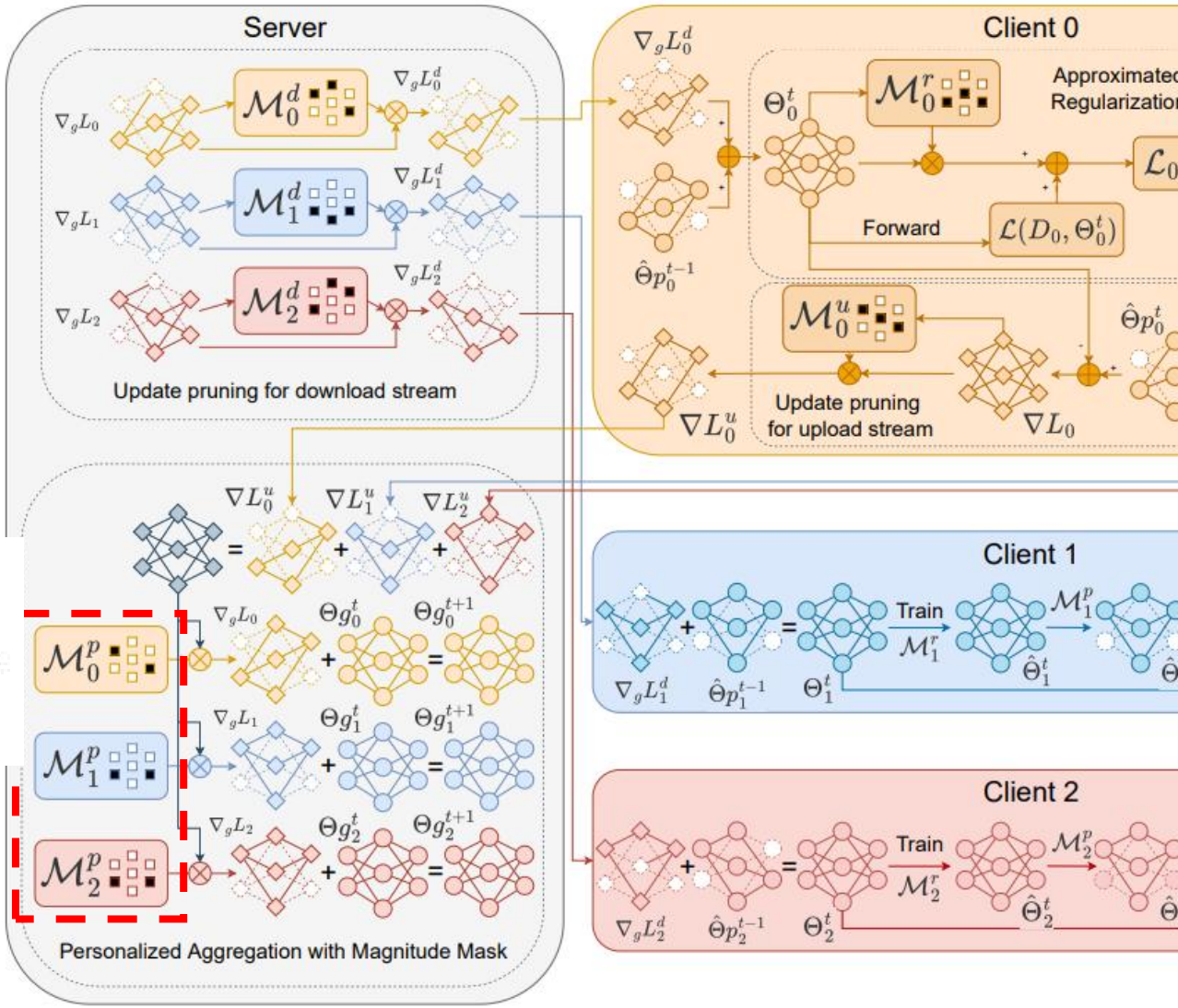
Hongjian Shi, Jianqing Zhang, Shuming Fan, Ruhui Ma, Haibing Guan, pFedEff: An Efficient and Personalized Federated Cognitive Learning Framework in Multi-agent Systems, IEEE Transactions on Cognitive and Development Systems, 2023.

联邦学习效率

- 个性化模型聚合
- 只对有效更新进行聚合

$$\Theta g_i^{t+1} \leftarrow \lambda_2 \cdot \Theta g_i^t +$$

$$(1-\lambda_2)\cdot \mathcal{M}_i^p \odot \sum_{j=1}^{M} k_j \cdot (\Theta g_j^t - \eta \cdot \nabla_g \mathcal{L}_j^u),$$



Hongjian Shi, Jianqing Zhang, Shuming Fan, Ruhui Ma, Haibing Guan, pFedEff: An Efficient and Personalized Federated Cognitive Learning Framework in Multi-agent Systems, IEEE Transactions on Cognitive and Development Systems, 2023.

联邦学习效率-模型压缩

- 在大幅减少通信量（~94%）的情况下
- 尽可能保障了训练准确率（<6%）

Metrics & Baselines			Dataset&Data distributions									
			MNIST					FMNIST				
			IID	DIR75	DIR50	DIR25	PAT	IID	DIR75	DIR50	DIR25	PAT
	Acc.	FedAvg	**0.9913**	0.9907	0.9881	0.9833	0.9576	0.9142	0.9021	0.9029	0.8923	0.7795
		FedPer	0.9687	0.9569	0.9512	0.9291	0.8778	0.8490	0.8373	0.8439	0.8230	0.7334
		ADP-50	0.9900	0.9899	0.9882	0.9848	0.9593	0.9016	0.9022	0.9047	0.8920	0.8006
		Hermes-50	0.9902	0.9899	0.9883	0.9852	0.9596	**0.9152**	0.9030	0.9031	0.8930	0.7803
		SplitFed	0.9911	**0.9909**	0.9882	0.9834	0.9545	0.9134	0.9022	0.9050	0.8919	0.7680
		FCCL	0.9820	0.9808	0.9821	**0.9866**	**0.9981**	0.8831	0.9037	0.9122	**0.9254**	**0.9949**
		LotteryFL-75	0.9700	0.9700	0.9750	0.9800	0.9775	**0.9152**	**0.9275**	**0.9125**	0.9250	0.9825
		LotteryFL-50	0.9700	0.9725	0.9700	0.9775	0.9875	0.8625	0.9075	0.9075	0.9225	0.9875
		LotteryFL-25	0.9350	0.9400	0.9350	0.9525	0.9600	0.7625	0.7675	0.7075	0.7850	0.8600
		pFedEff-75	0.9908	0.9873	**0.9888**	0.9814	0.9729	0.9108	0.8887	0.8961	0.8989	0.9293
		pFedEff-50	0.9897	0.9843	0.9862	0.9786	0.9715	0.9086	0.8774	0.8867	0.8894	0.9585
		pFedEff-25	0.9863	0.9723	0.9745	0.9601	0.9745	0.9018	0.8427	0.8544	0.8557	0.9500
	Conv.	FedAvg	198	198	192	197	192	196	200	200	198	200
		FedPer	196	197	199	200	197	196	196	199	197	200
		ADP-50	148	187	**179**	162	197	174	197	198	200	192
		Hermes-50	**90**	160	193	199	198	178	184	198	177	198
		SplitFed	192	198	198	189	199	196	189	194	190	193
		FCCL	79	187	167	132	184	151	127	191	189	51
		LotteryFL-75	131	178	198	**125**	98	**70**	109	93	152	41
		LotteryFL-50	197	**153**	183	168	149	159	182	157	165	81
		LotteryFL-25	187	189	181	159	172	199	**0**	**4**	**8**	97
		pFedEff-75	197	193	196	194	**2**	200	187	195	198	**4**
		pFedEff-50	199	199	195	193	4	190	191	199	196	6
		pFedEff-25	195	200	200	200	19	184	200	199	199	16
	Comm. Vol.	FedAvg	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		FedPer	0.0895	0.0895	0.0895	0.0895	0.0895	0.0895	0.0895	0.0895	0.0895	0.0895
		ADP-50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		Hermes-50	1.0037	1.0037	1.0037	1.0037	1.0037	1.0037	1.0037	1.0037	1.0037	1.0037
		SplitFed	0.1071	0.1071	0.1071	0.1071	0.1071	0.1071	0.1071	0.1071	0.1071	0.1071
		FCCL	0.1031	0.1031	0.1031	0.1031	0.1031	0.1031	0.1031	0.1031	0.1031	0.1031
		LotteryFL-75	0.7350	0.7351	0.7351	0.7351	0.7350	0.7350	0.7351	0.7351	0.7351	0.7350
		LotteryFL-50	0.4876	0.4877	0.4877	0.4877	0.4876	0.4876	0.4877	0.4877	0.4877	0.4876
		LotteryFL-25	0.2377	0.2378	0.2378	0.2378	0.2377	0.3362	0.2430	0.2575	0.2523	0.2377
		pFedEff-75	0.5277	0.5625	0.5625	0.5625	0.5149	0.5548	0.5625	0.5625	0.5625	0.5623
		pFedEff-50	0.2398	0.2500	0.2500	0.2500	0.2407	0.2500	0.2500	0.2500	0.2500	0.2500
		pFedEff-25	**0.0596**	**0.0625**	**0.0625**	**0.0625**	**0.0625**	**0.0625**	**0.0625**	**0.0625**	**0.0625**	**0.0625**

Hongjian Shi, Jianqing Zhang, Shuming Fan, Ruhui Ma, Haibing Guan, pFedEff: An Efficient and Personalized Federated Cognitive Learning Framework in Multi-agent Systems, IEEE Transactions on Cognitive and Development Systems, 2023.

联邦学习效率-模型压缩

- 不同绝对值mask下不同压缩率结果
- pre-training mask和post-training mask能够有效减少通信量，但是对准确率影响较大

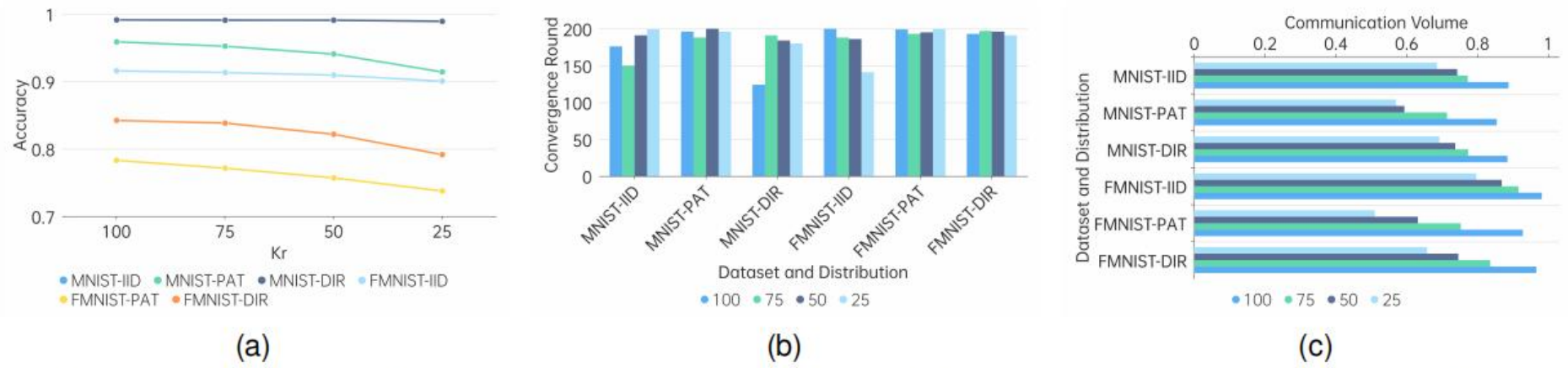


Fig. 5. The metrics when using different K_i^r for different distributions on different datasets. (a) Accuracy; (b) Convergence rate; (c) Communication Volume. We can observe that K_i^r can decrease the communication volume but introduce minor accuracy loss.

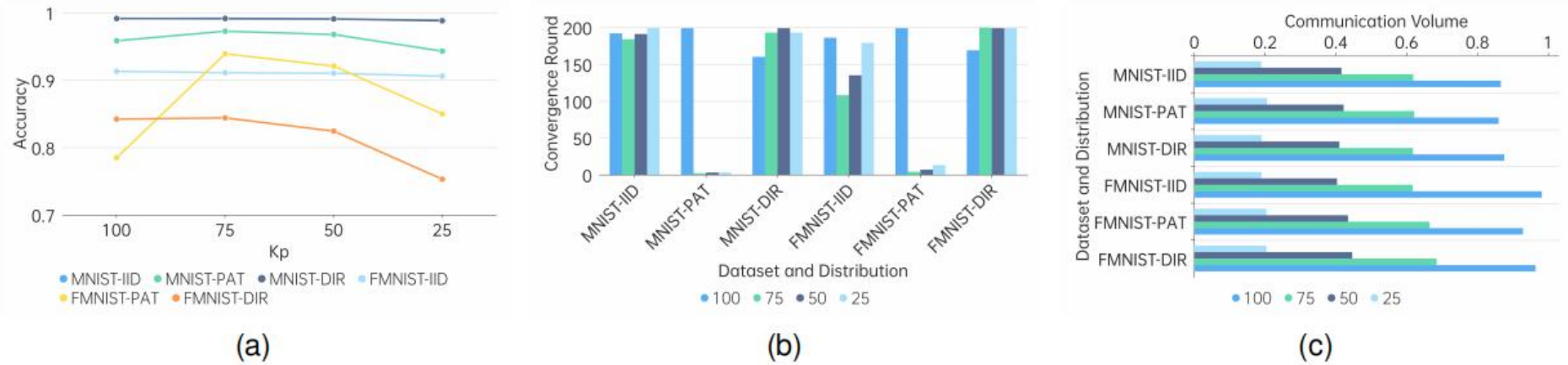
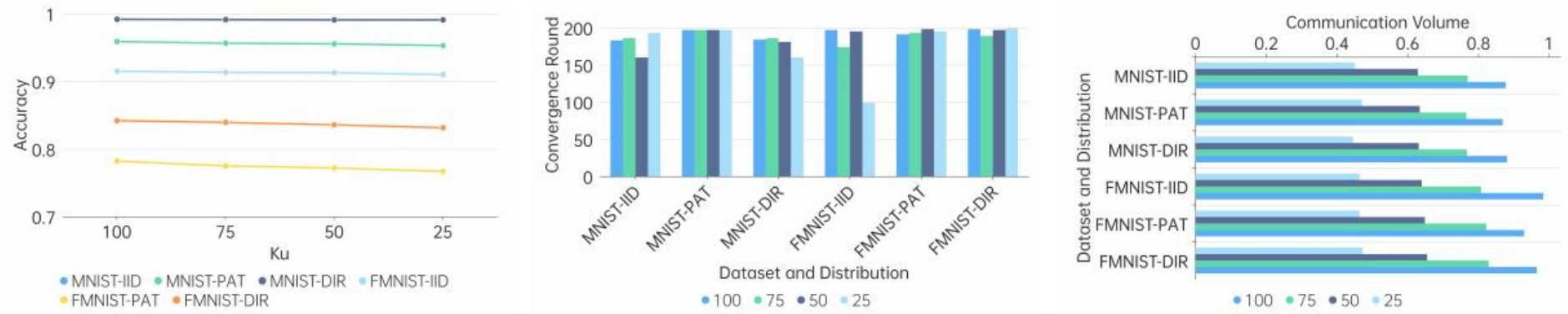


Fig. 6. The metrics when using different K_i^p for different distributions on different datasets. (a) Accuracy; (b) Convergence rate; (c) Communication Volume. We can observe that K_i^p significantly reduces the communication volume but introduce relatively large accuracy loss.

Hongjian Shi, Jianqing Zhang, Shuming Fan, Ruhui Ma, Haibing Guan, pFedEff: An Efficient and Personalized Federated Cognitive Learning Framework in Multi-agent Systems, IEEE Transactions on Cognitive and Development Systems, 2023.

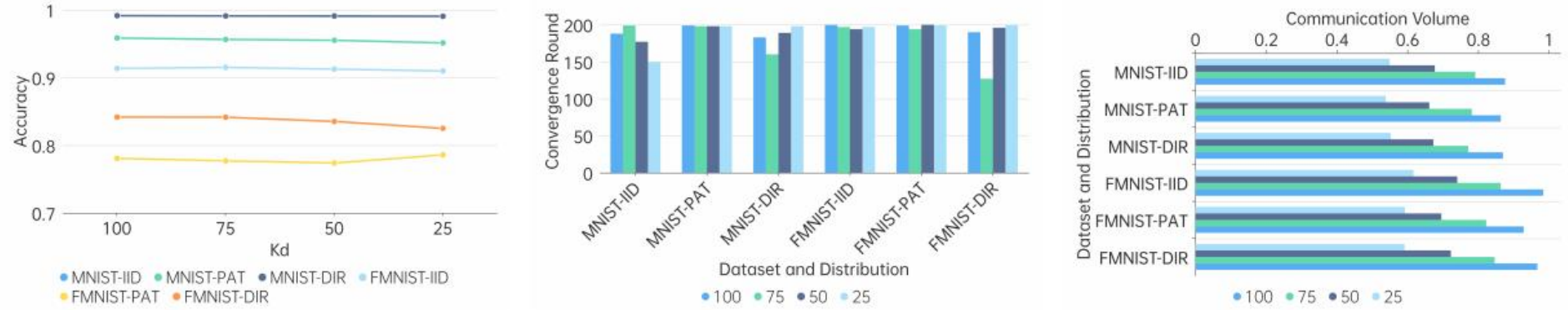
联邦学习效率-模型压缩

- 不同重要性mask下不同压缩率结果
- upload mask和download mask虽然通信量减少有限，但是能够有效维持准确率



(a) (b) (c)

Fig. 7. The metrics when using different K_i^u for different distributions on different datasets. (a) Accuracy; (b) Convergence rate; (c) Communication Volume. We can observe that K_i^u can keep the training accuracy while decreasing the communication volume.



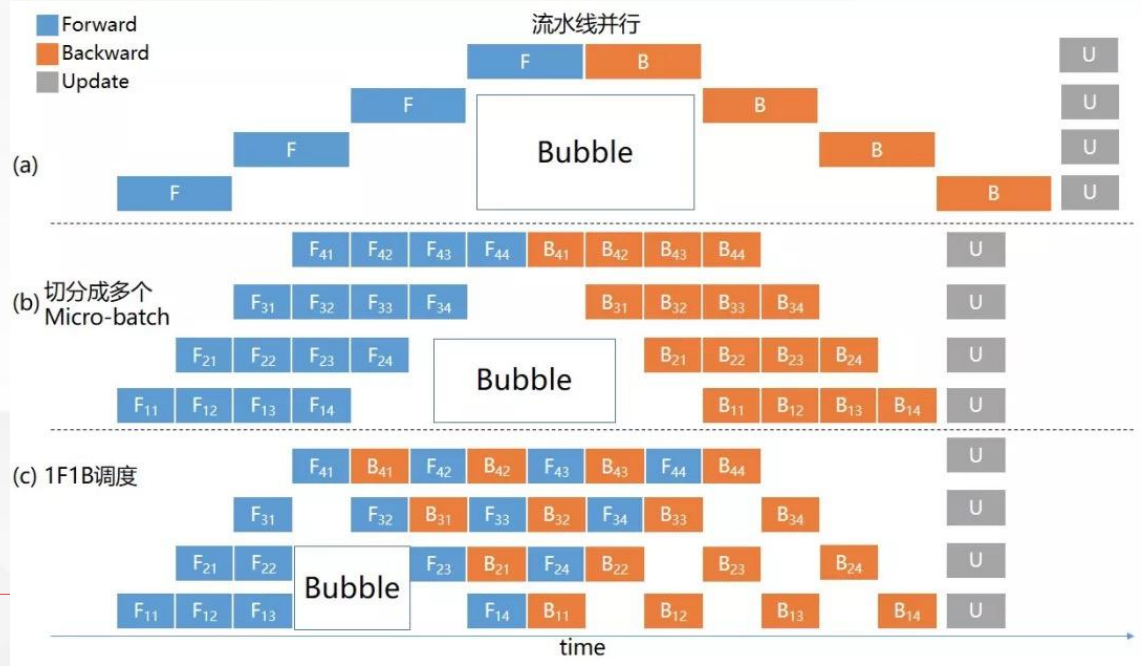
(a) (b) (c)

Fig. 8. The metrics when using different K_i^d for different distributions on different datasets. (a) Accuracy; (b) Convergence rate; (c) Communication Volume. We can observe that K_i^d can keep the training accuracy while decreasing the communication volume.

Hongjian Shi, Jianqing Zhang, Shuming Fan, Ruhui Ma, Haibing Guan, pFedEff: An Efficient and Personalized Federated Cognitive Learning Framework in Multi-agent Systems, IEEE Transactions on Cognitive and Development Systems, 2023.

联邦学习效率-传输调度

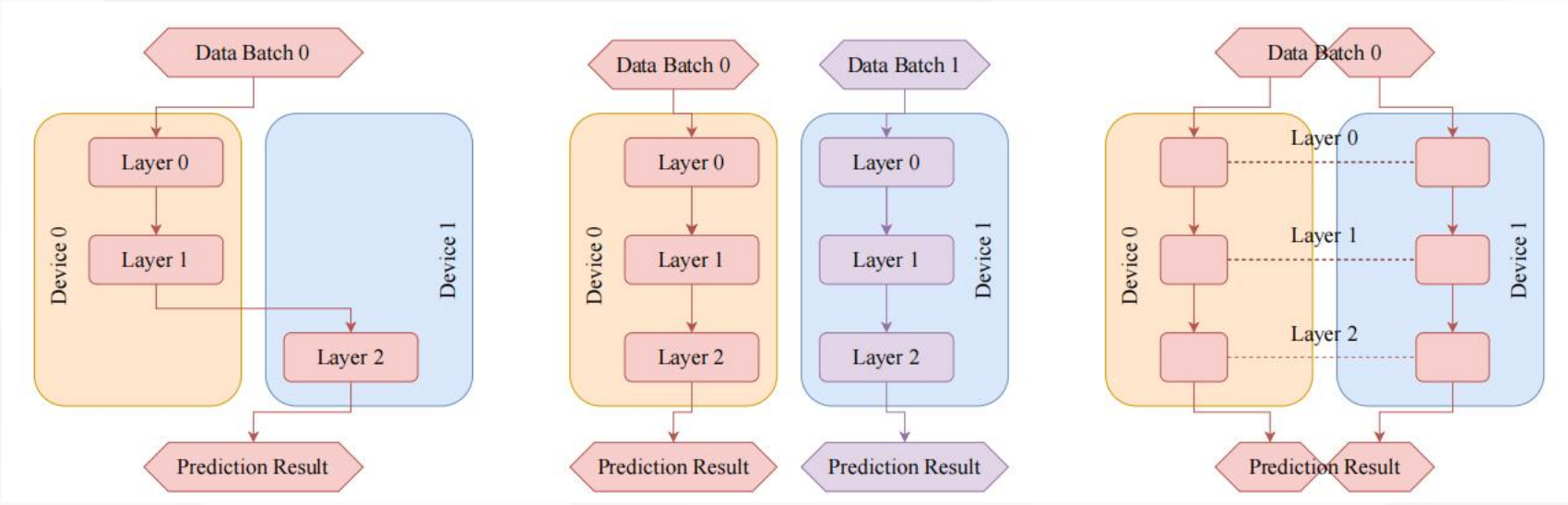
- 任务调度
  - 任务复杂，资源丰富，需要调度计算子任务
- 并行算法：
  - 计算and计算、计算and通信、通信and通信
  - 数据并行、模型并行、算子并行、优化器并行

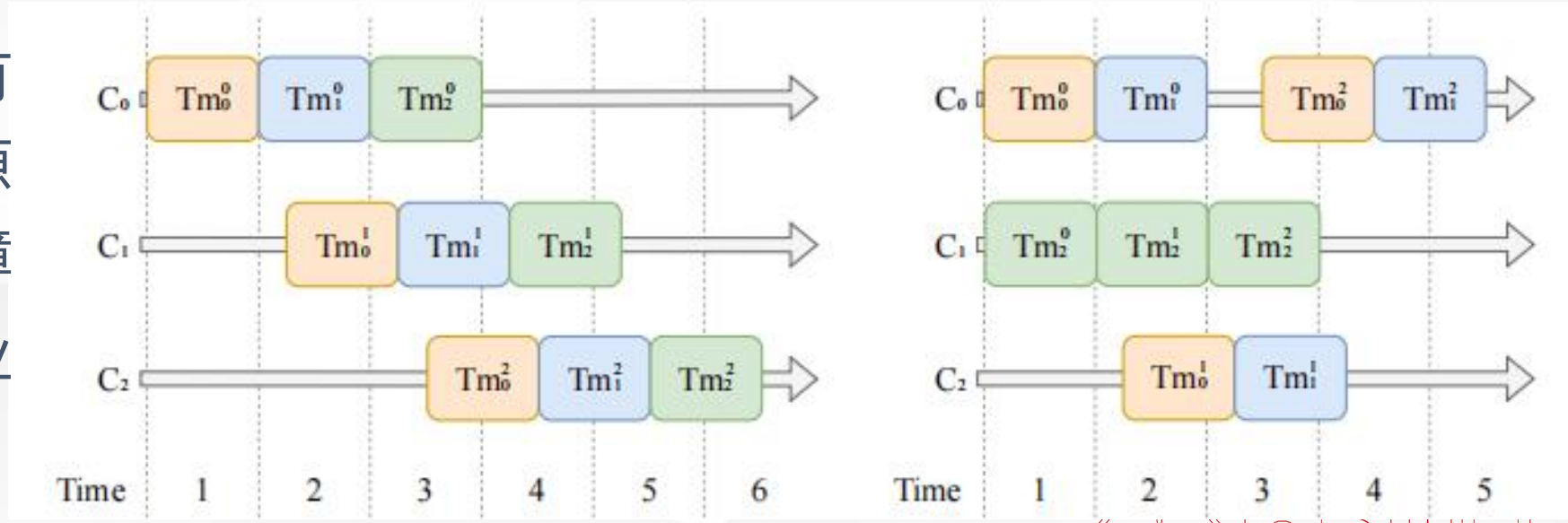


Parallel Strategy	Affiliation
Pangu[2]	Huawei
Pipedream[3]	Microsoft
Gpipe[4]	Google
Fairscale[5]	Facebook
Megatron-LM[6]	NVIDIA
DeepSpeed[7]	Microsoft
Colossal-AI[8]	HPC-AI
FlexFlow[9]	Stanford
OneFlow[10]	OneFlow

联邦学习效率-传输调度

- 并行算法
  - 模型：减少空间占用
  - 数据：减少处理延迟
  - 算子：提高处理速度
- 存在的问题
  - 单一并行算法效率有限，存在大量的资源浪费，可用性无保障
  - 多并行算法需要专业设计

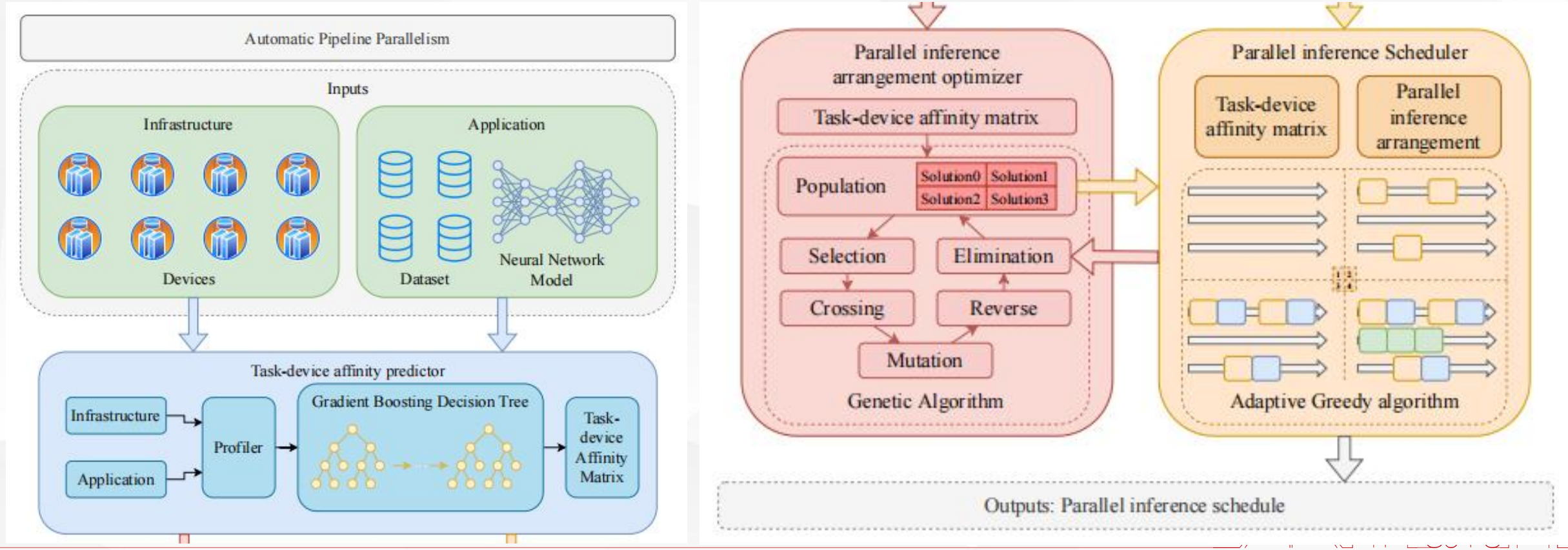




Hongjian Shi, Weichu Zheng, Zifei Liu, Ruhui Ma, Haibing Guan, Automatic Pipeline Parallelism: A Parallel Inference Framework for Deep Learning Applications in 6G Mobile Communication Systems, in IEEE Journal on Selected Areas in Communications，2023.

联邦学习效率-传输调度

- 先采用GBDT预测各任务在各设备上的运行时间，来减少profiling的用时
- 而后采用遗传算法决定任务位置，确认各个任务所在设备，以减少任务处理延迟
- 遗传算法中会用贪心算法排列各设备上的任务，得到适应度函数，并得到实际可运行的任务时间表



Hongjian Shi, Weichu Zheng, Zifei Liu, Ruhui Ma, Haibing Guan, Automatic Pipeline Parallelism: A Parallel Inference Framework for Deep Learning Applications in 6G Mobile Communication Systems, in IEEE Journal on Selected Areas in Communications，2023.

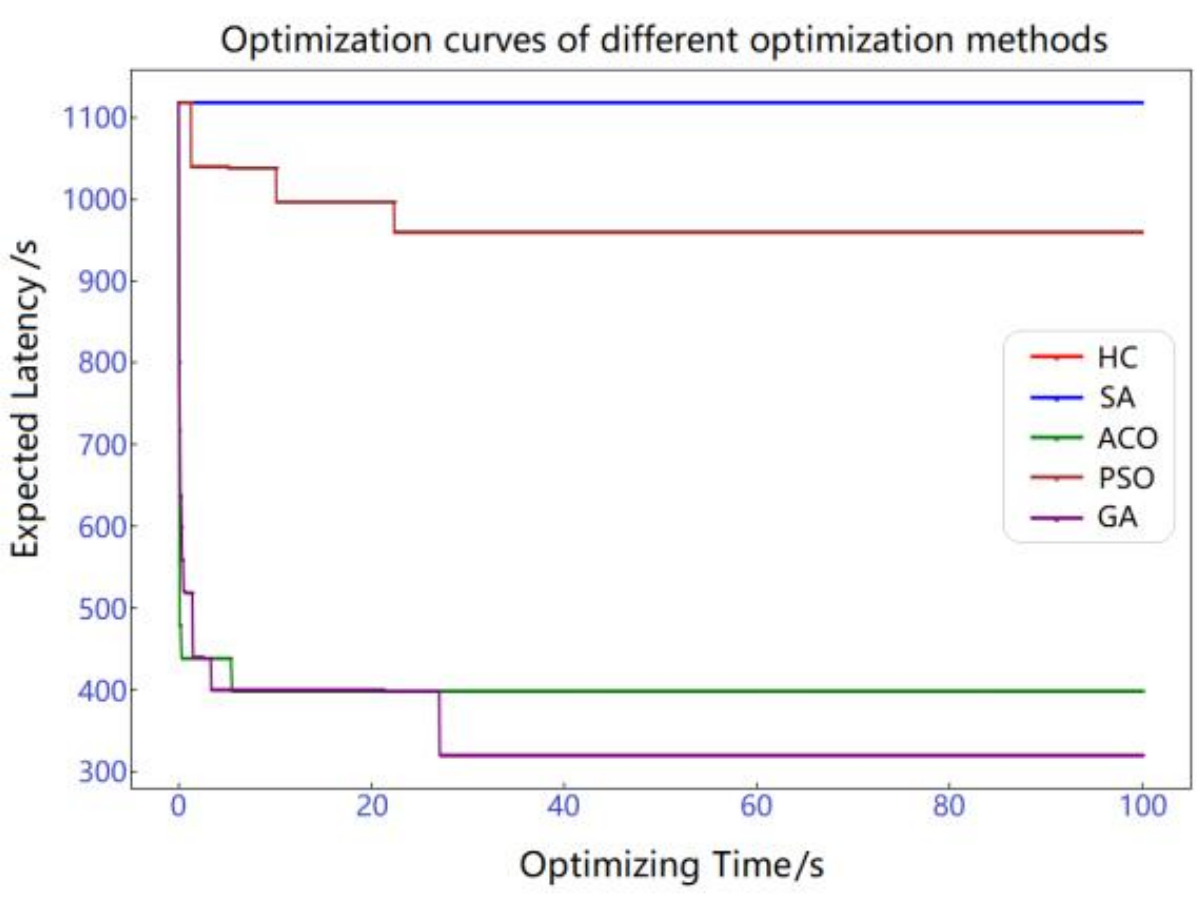
联邦学习效率

- 整体算法能够提高设备利用率
- 减少延迟提高吞吐量
- 同时能提高可用性

- 采用遗传算法的效果和收敛性较好

Method	Latency(s)	Throu.(/s)	Dev. Util.(%)	Rel.(%)
Standalone	6.1053	1310.34	12.2831	0
Model#8	9.6717	827.15	8.3009	0
Data	0.9468	8449.12	83.9462	100
Pipeline#8	2.9603	2702.43	27.0113	0
AP^2#5	1.3739	5822.89	49.9417	100

Method	Latency(s)	Throu.(/s)	Dev. Util.(%)	Rel.(%)
Standalone	6.9814	1432.37	9.8450	0
Model#5	11.9720	835.29	6.2150	0
Data	0.9734	10273.53	82.3580	100
Pipeline#5	1.7692	5652.16	39.2073	89
AP^2#5	1.7178	5821.57	45.4050	100



Hongjian Shi, Weichu Zheng, Zifei Liu, Ruhui Ma, Haibing Guan, Automatic Pipeline Parallelism: A Parallel Inference Framework for Deep Learning Applications in 6G Mobile Communication Systems, in IEEE Journal on Selected Areas in Communications，2023.

联邦学习效率

- GBDT预测方法准确率明显由于其他预测方法
- 并且由于是线下训练，训练时间不会对线上推理产生负面影响
- 主要是由于GBDT对于表格数据和feature较少的数据有着十分优越的表现

Method	Test Accuracy	Train Accuracy	Training time(ms)
GBDT	**0.7176±0.0001**	**0.8058±0.0000**	107.48±0.64
FCNN-leakyrelu	0.6583±0.0003	0.7670±0.0000	6620.46±108.63
FCNN-relu [78]	0.6719±0.0002	0.7664±0.0000	6699.71±114.25
FCNN-tanh	0.6860±0.0000	0.7674±0.0000	6674.88±95.96
FCNN-sigmoid	0.7039±0.0000	0.7624±0.0000	6574.73±106.43
ABDT [79]	0.7114±0.0000	0.8001±0.0000	70.20±1.06
RF [80]	0.7096±0.0011	0.7738±0.0004	9.56±0.57
SVR [81]	0.7166±0.0000	0.7399±0.0000	**7.97±0.52**

n_estimators	Train Accuracy	Test Accuracy	Training Time(ms)
100	0.7410±0.0000	0.7172±0.0000	**48.39±0.72**
120	0.7634±0.0000	0.7251±0.0000	58.17±0.57
140	0.7783±0.0000	**0.7266±0.0000**	67.46±0.79
160	0.7882±0.0000	0.7238±0.0000	76.44±0.68
180	0.7978±0.0000	0.7209±0.0000	86.48±1.20
200	0.8058±0.0000	0.7176±0.0001	95.27±0.77
220	0.8120±0.0000	0.7132±0.0001	106.06±4.01
240	0.8219±0.0025	0.7064±0.0029	115.31±3.25
260	0.8273±0.0020	0.7020±0.0025	125.07±4.59
280	**0.8309±0.0000**	0.6992±0.0036	132.34±1.07

Hongjian Shi, Weichu Zheng, Zifei Liu, Ruhui Ma, Haibing Guan, Automatic Pipeline Parallelism: A Parallel Inference Framework for Deep Learning Applications in 6G Mobile Communication Systems, in IEEE Journal on Selected Areas in Communications，2023.

联邦学习安全

- 拜占庭攻击 - 双端协同预警
- 不可信第三方 - 区块链辅助聚合

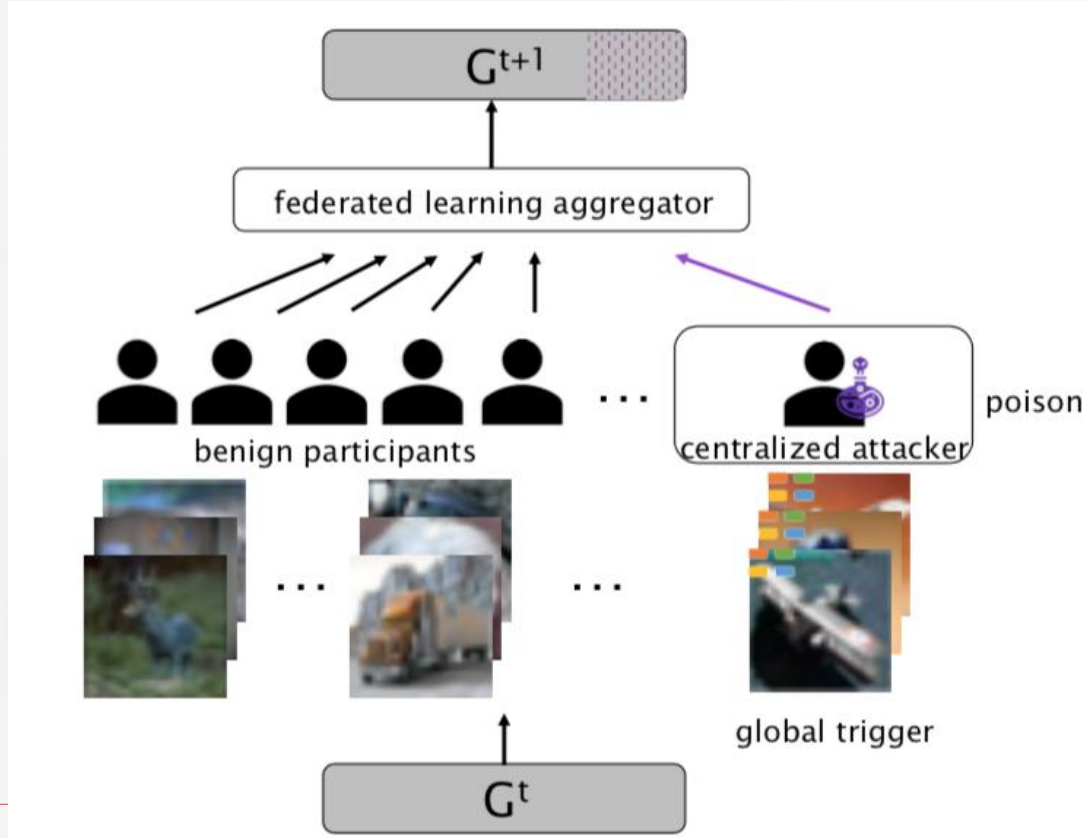
联邦学习安全---拜占庭攻击

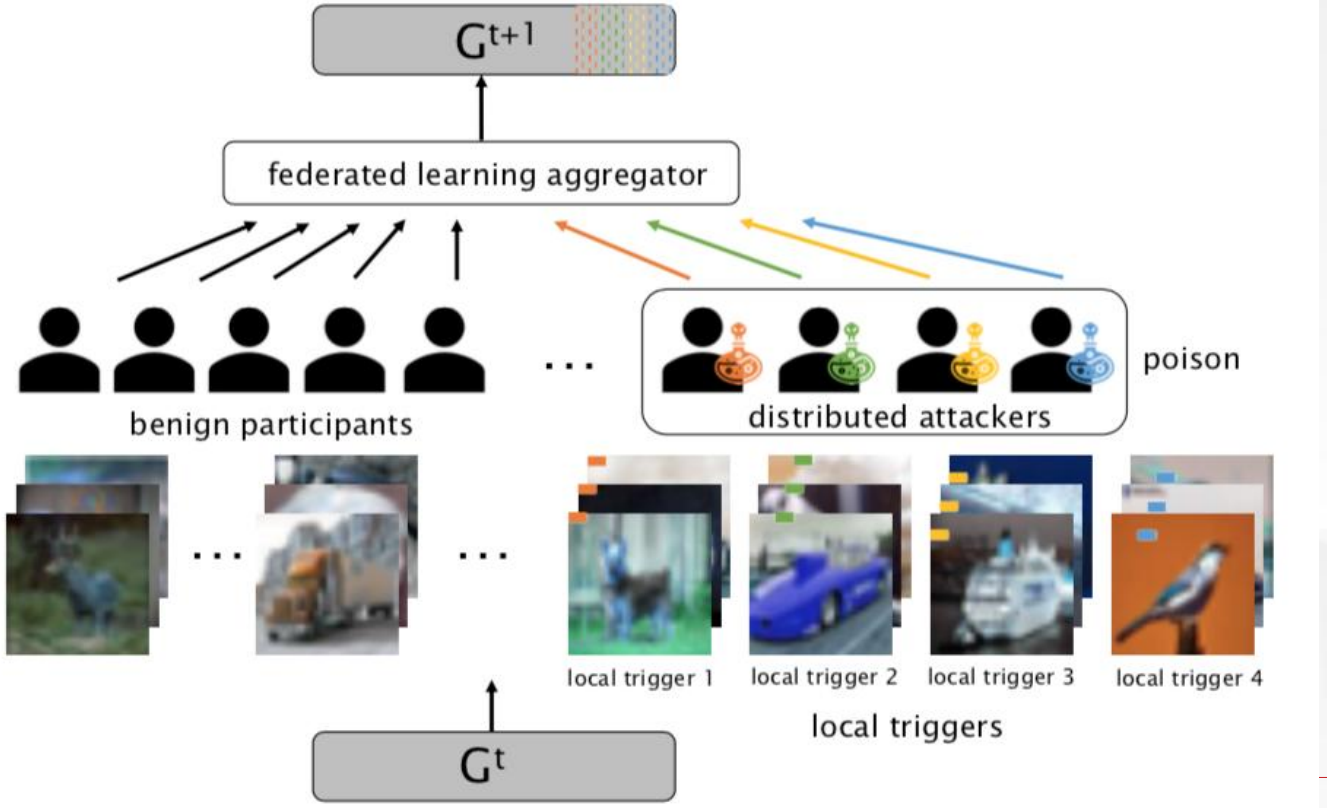
计算安全

- 拜占庭攻击，系统中存在恶意节点通过传输错误数据破坏训练模型

异常更新判断：

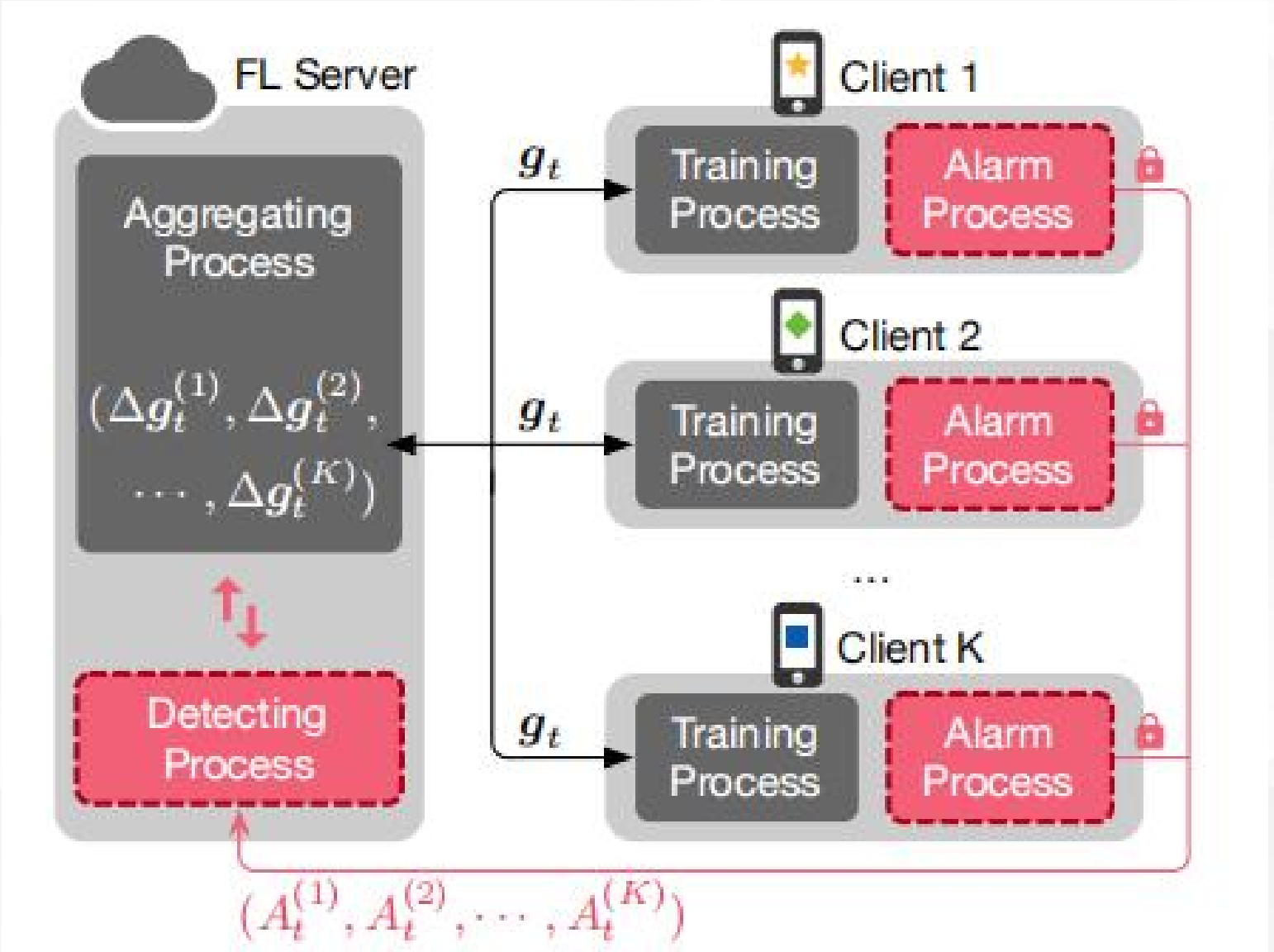
- 判断模型更新中的离群点
- 剔除离群点或是抑制离群点贡献





联邦学习安全---拜占庭攻击

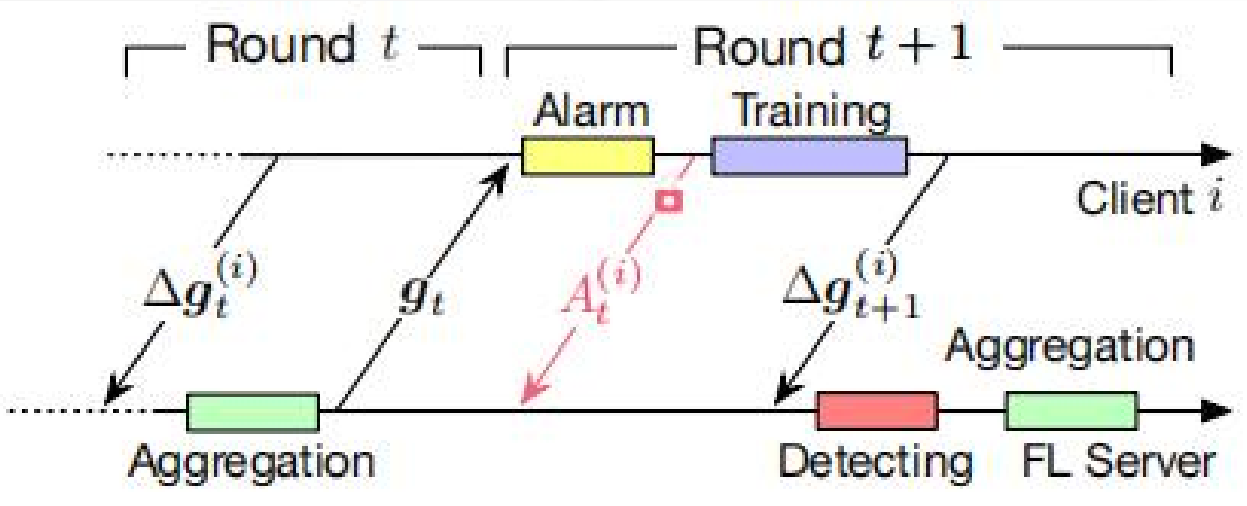
- 利用客户端信息辅助排除恶意节点
- 客户端侧拥有**示警机制**，可以和服务器侧的决策流程互相配合
- 服务器侧使用基于准确度筛查和权重分析的**决策流程**
- 系统提供两种辅助机制，分别为**惩罚机制**和**奖励机制**

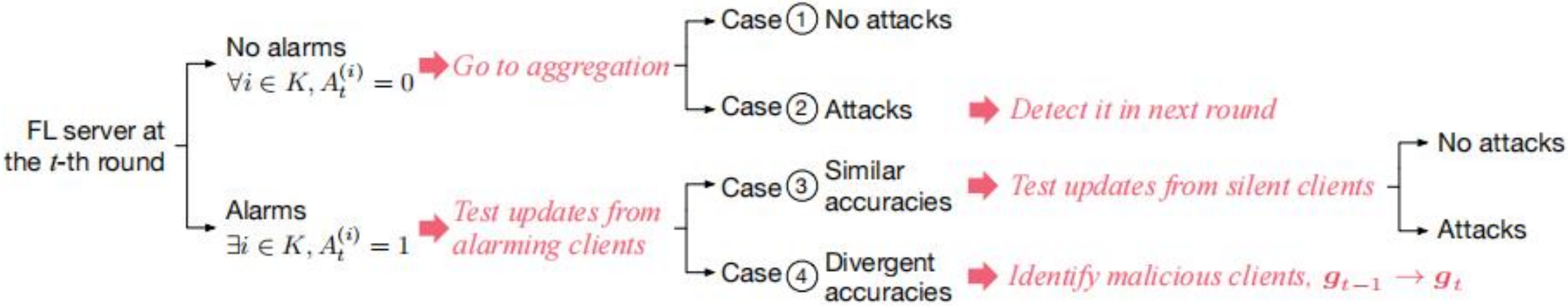


Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Zhangcheng Lv, Xiulang Jin, Zhengui Xue, Ruhui Ma, Haibing Guan, Siren: Byzantine-robust Federated Learning via Proactive Alarming, Proceedings of ACM Symposium on Cloud Computing, pp. 47--60, 2021.

联邦学习安全---拜占庭攻击

- 客户端侧结构与方法流程
  - 根据全局模型和本地模型**区别**进行判断
- 服务器侧结构与方法流程
  - 根据客户端**告警**判断是否存在恶意攻击





Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Zhangcheng Lv, Xiulang Jin, Zhengui Xue, Ruhui Ma, Haibing Guan, Siren: Byzantine-robust Federated Learning via Proactive Alarming, Proceedings of ACM Symposium on Cloud Computing, pp. 47--60, 2021.

联邦学习安全---拜占庭攻击

- 在sign-flipping, label-flipping和targeted model poisoning攻击下可以有效收敛到一个较优准确率
- 也可以看出针对targeted model poisoning时，Siren稳定性较好

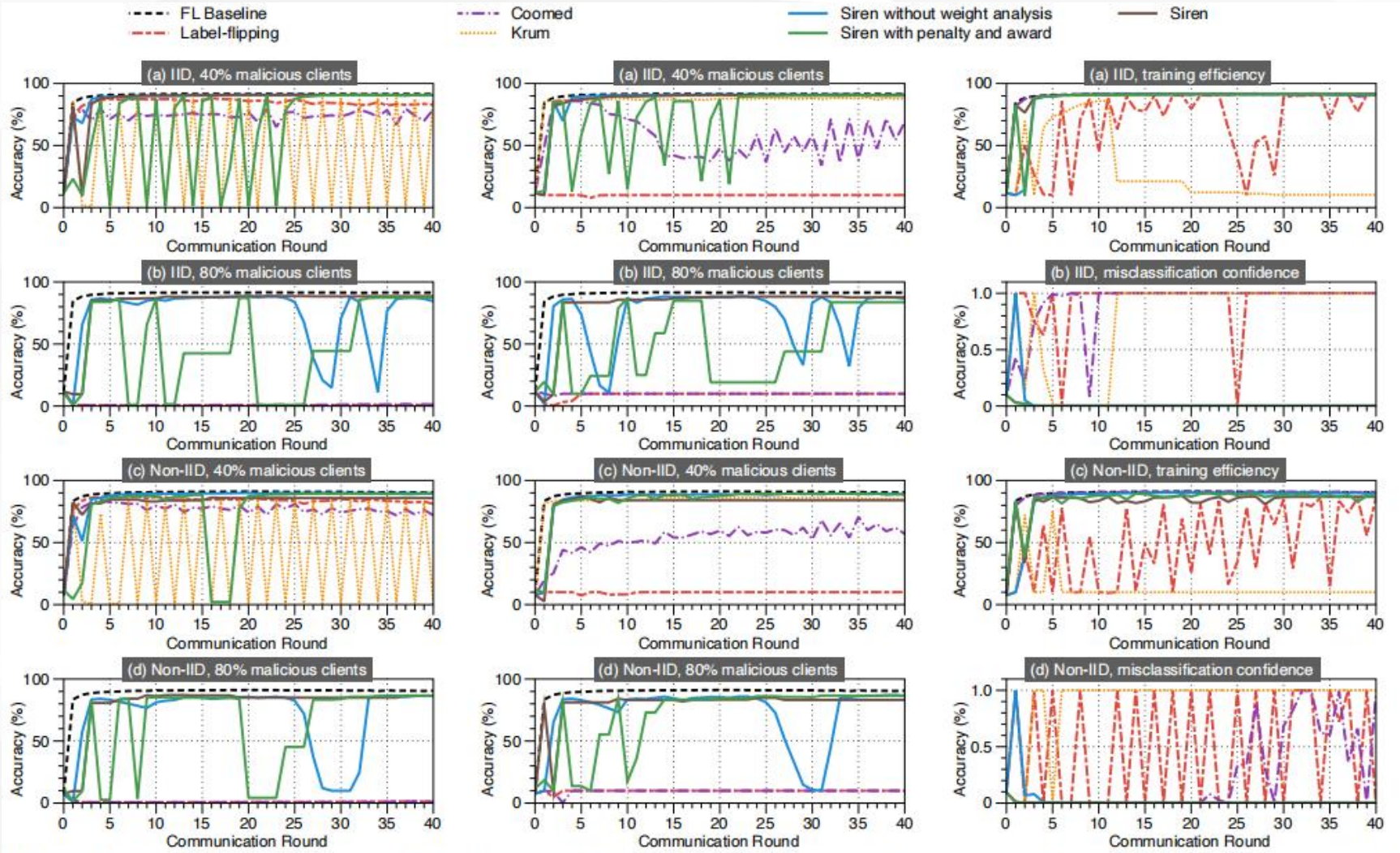


Figure 5: Training efficiency under label-flipping attack when $|K| = 10$.

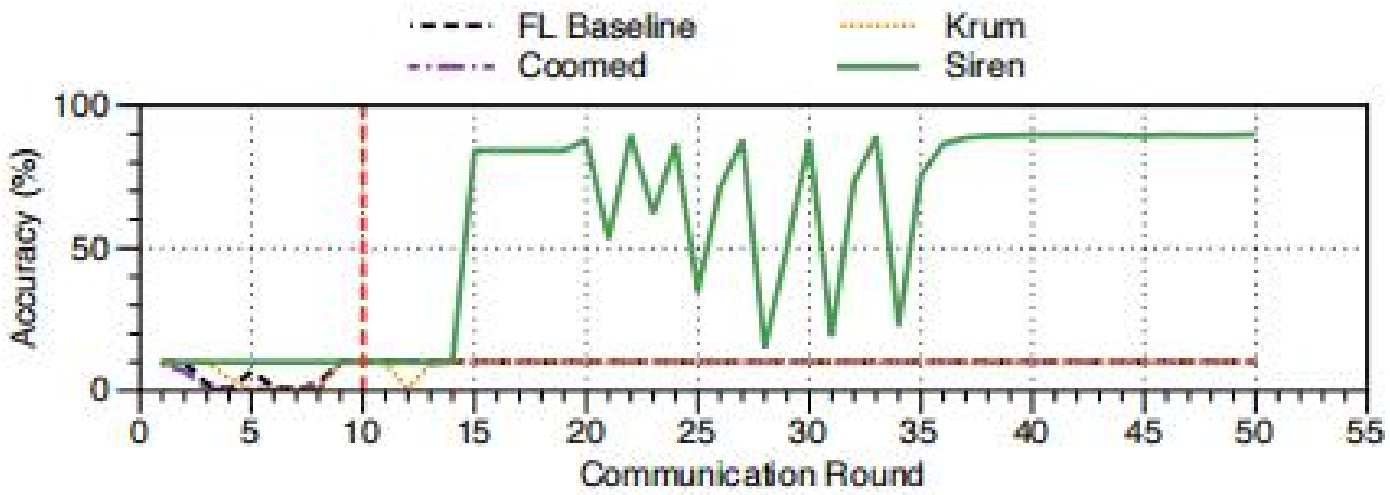
Figure 6: Training efficiency under sign-flipping attack when $|K| = 10$.

Figure 7: Training efficiency and misclassification confidence under targeted model poisoning, $|K| = 10$.

Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Zhangcheng Lv, Xiulang Jin, Zhengui Xue, Ruhui Ma, Haibing Guan, Siren: Byzantine-robust Federated Learning via Proactive Alarming, Proceedings of ACM Symposium on Cloud Computing, pp. 47--60, 2021.

联邦学习安全---拜占庭攻击

- Siren可以在大比例恶意客户端的情况下（80%）拥有较强的防御能力
- Siren可以在系统已经遭受攻击的情况下恢复原有系统的训练准确率



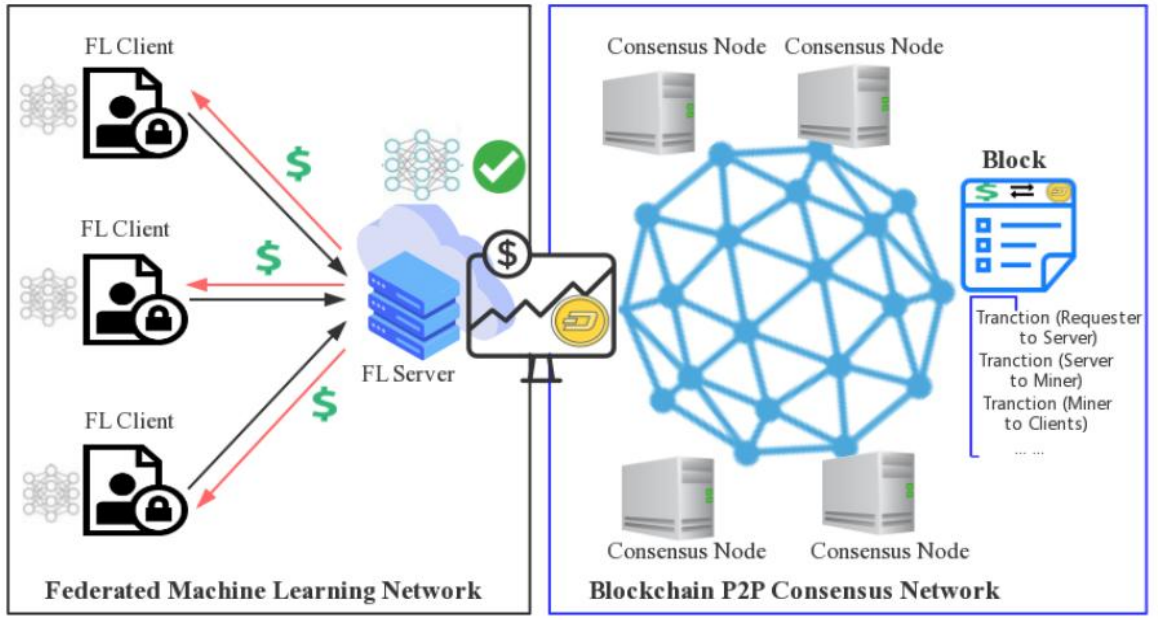
Attack Type	Defense Methods	Malicious Proportion	Accuracy
None	None	0%	55.33%
Sign-flipping	None	40%	10.01%
		80%	10.01%
	Krum[1]	40%	41.53%
	Coomed	40%	34.25%
		80%	9.99%
	SIREN	40%	**51.58%**
		80%	**45.50%**
Label-flipping	None	40%	34.70%
		80%	11.68%
	Krum[1]	40%	44.72%
	Coomed	40%	37.21%
		80%	9.60%
	SIREN	40%	**49.82%**
		80%	**43.52%**

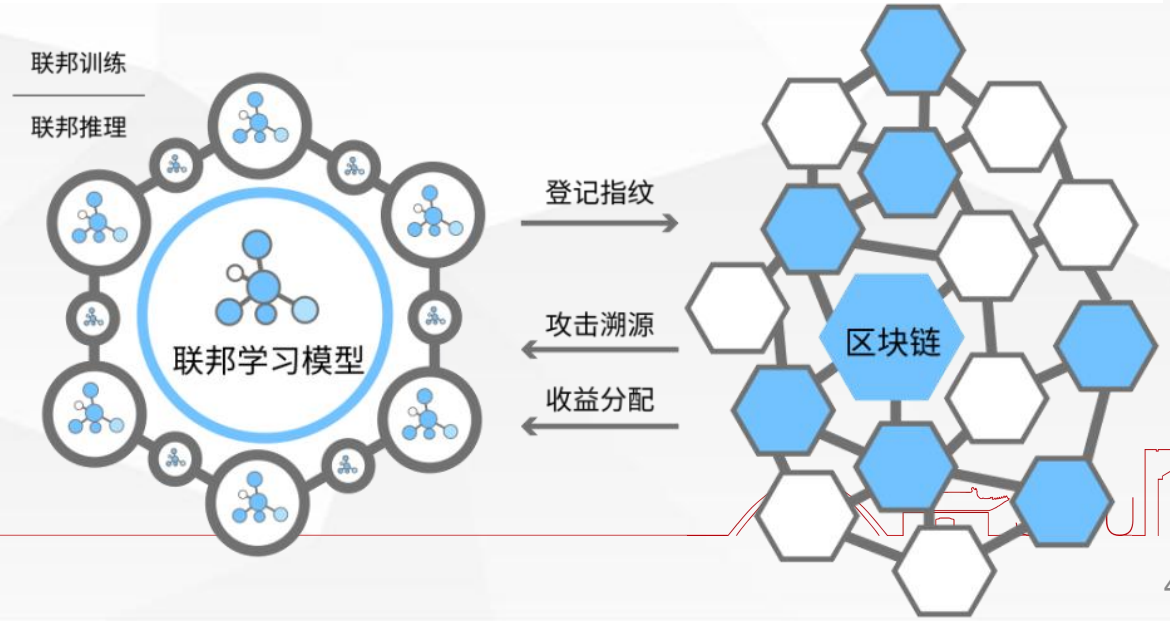
[1] Krum cannot work properly when malicious clients' proportion reaches 80%.

Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Zhangcheng Lv, Xiulang Jin, Zhengui Xue, Ruhui Ma, Haibing Guan, Siren: Byzantine-robust Federated Learning via Proactive Alarming, Proceedings of ACM Symposium on Cloud Computing, pp. 47--60, 2021.

联邦学习安全---不可信第三方

- 通信安全
  - 无可信第三方服务器，易受攻击
- 联邦学习+区块链
  - 用区块链替换原有server保存全局模型，可以提高全局模型的防篡改能力
  - 区块链本身的共识机制和联邦学习的聚合权重有很多相似之处



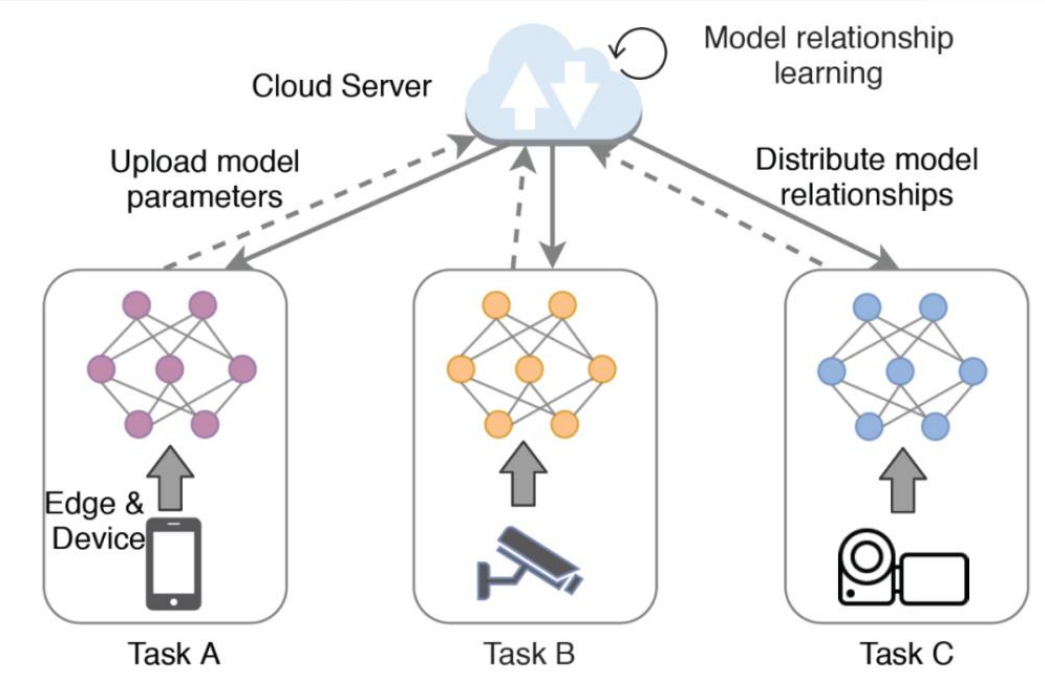


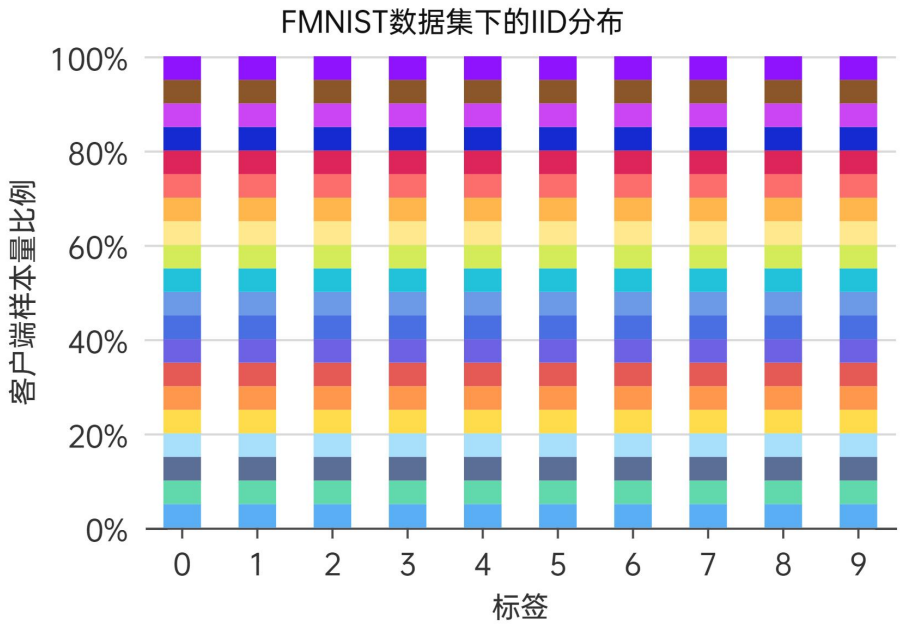
联邦学习异构性

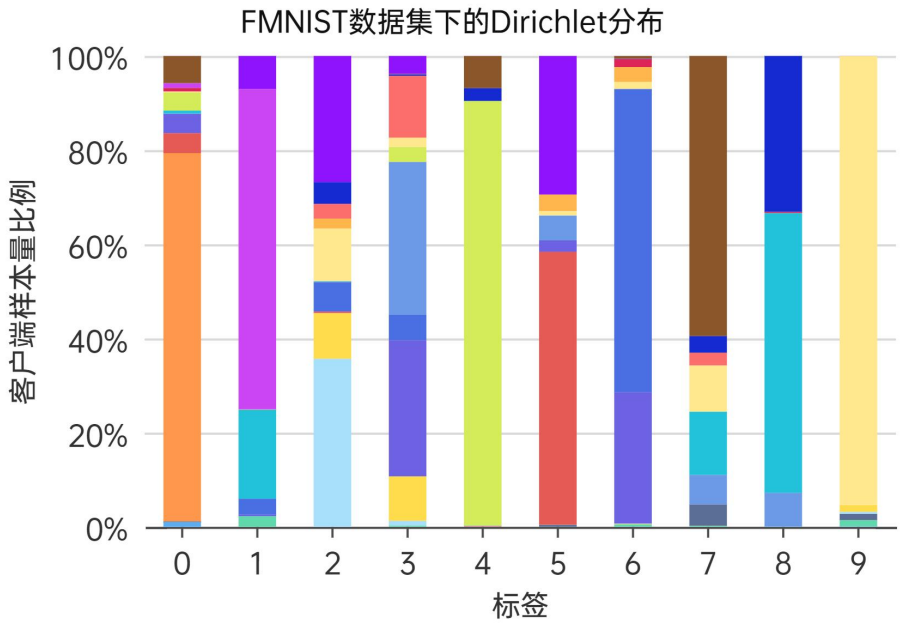
- 统计异构性 - 非聚合联邦学习
- 系统异构性 - 适应性训练调整

联邦学习异构性---统计异构性

- 统计异构
  - 各客户端数据分布不同，模型不准确
- 跨域联邦学习：
  - 模型之间交换知识，不单纯进行聚合
  - 模型在本地测试集上的效果会更好







联邦学习异构性---统计异构性

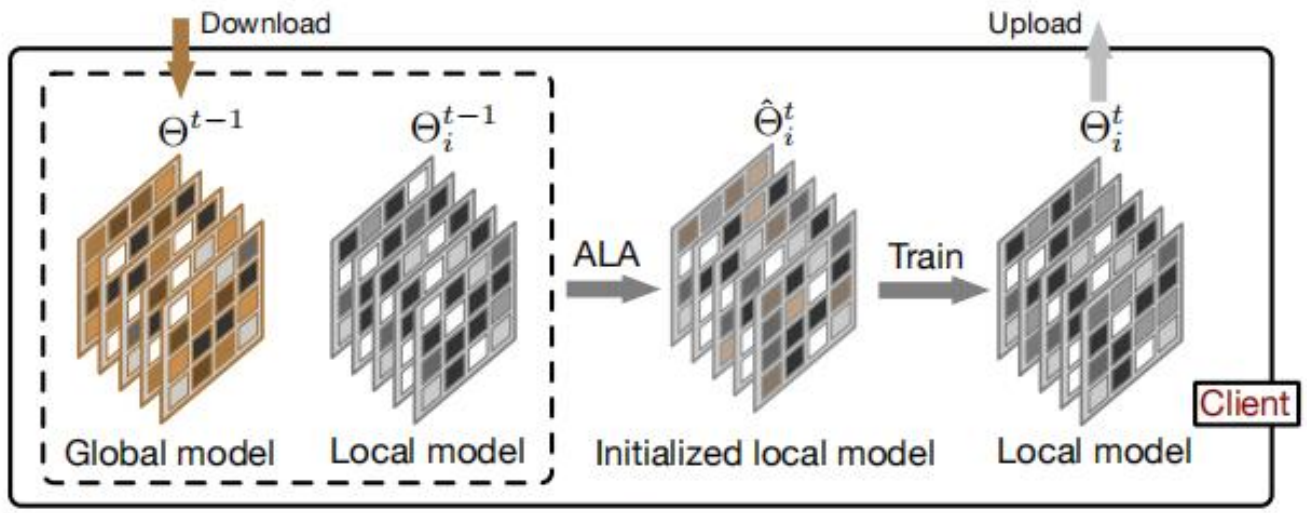
- 引入个性化mask做element-wise的更新

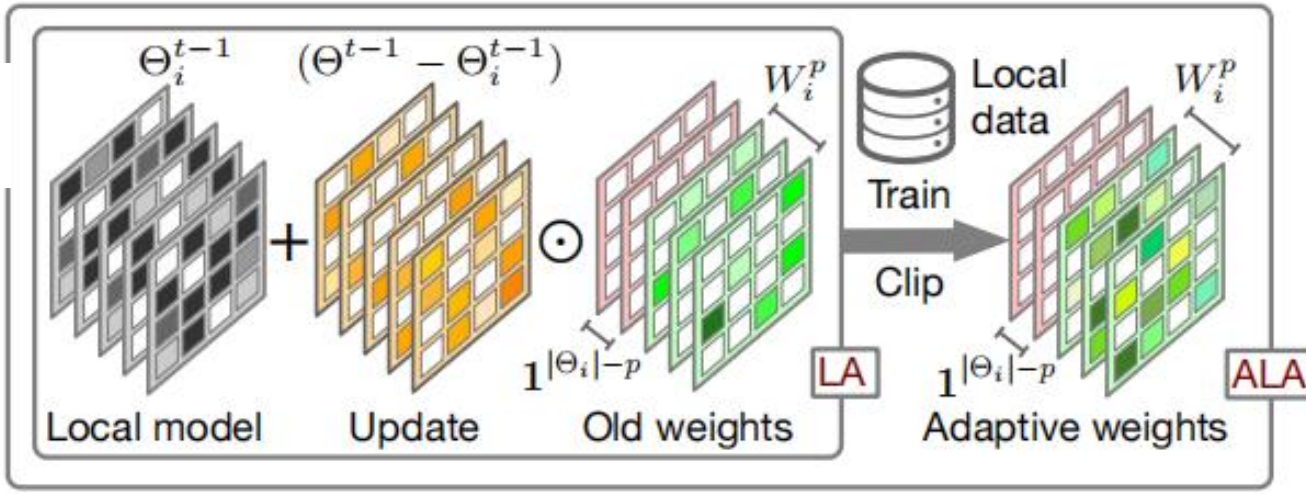
$$\hat{\Theta}_i^t := \Theta_i^{t-1} \odot W_{i,1} + \Theta^{t-1} \odot W_{i,2},$$
$$s.t. \quad w_1^q + w_2^q = 1, \forall \text{ valid } q$$

$$\hat{\Theta}_i^t := \Theta_i^{t-1} + (\Theta^{t-1} - \Theta_i^{t-1}) \odot W_i,$$

$$\hat{\Theta}_i^t := \Theta_i^{t-1} + (\Theta^{t-1} - \Theta_i^{t-1}) \odot [\mathbf{1}^{|\Theta_i|-p}; W_i^p],$$

$$W_i^p \leftarrow W_i^p - \eta \nabla_{W_i^p} \mathcal{L}(\hat{\Theta}_i^t, D_i^{s,t}; \Theta^{t-1}),$$





Jianqing Zhang, Yang Hua, Hao Wang, Tao Song, Zhengui Xue, Ruhui Ma, Haibing Guan, FedALA: Adaptive Local Aggregation for Personalized Federated Learning, Proceedings of the 37th AAAI Conference on Artificial Intelligence, 2022.

联邦学习异构性---统计异构性

- 在多个数据集多种设定上均取得了最高的准确率

Table 2: The test accuracy (%) in the pathological heterogeneous setting and practical heterogeneous setting.

Settings	Pathological heterogeneous setting			Practical heterogeneous setting				
Methods	MNIST	Cifar10	Cifar100	Cifar10	Cifar100	TINY	TINY*	AG News
FedAvg	97.93±0.05	55.09±0.83	25.98±0.13	59.16±0.47	31.89±0.47	19.46±0.20	19.45±0.13	79.57±0.17
FedProx	98.01±0.09	55.06±0.75	25.94±0.16	59.21±0.40	31.99±0.41	19.37±0.22	19.27±0.23	79.35±0.23
FedAvg-C	99.79±0.00	92.13±0.03	66.17±0.03	90.34±0.01	51.80±0.02	30.67±0.08	36.94±0.10	95.89±0.25
FedProx-C	99.80±0.04	92.12±0.03	66.07±0.08	90.33±0.01	51.84±0.07	30.77±0.13	38.78±0.52	96.10±0.22
Per-FedAvg	99.63±0.02	89.63±0.23	56.80±0.26	87.74±0.19	44.28±0.33	25.07±0.07	21.81±0.54	93.27±0.25
FedRep	99.77±0.03	91.93±0.14	67.56±0.31	90.40±0.24	52.39±0.35	37.27±0.20	39.95±0.61	96.28±0.14
pFedMe	99.75±0.02	90.11±0.10	58.20±0.14	88.09±0.32	47.34±0.46	26.93±0.19	33.44±0.33	91.41±0.22
Ditto	99.81±0.00	92.39±0.06	67.23±0.07	90.59±0.01	52.87±0.64	32.15±0.04	35.92±0.43	95.45±0.17
FedAMP	99.76±0.02	90.79±0.16	64.34±0.37	88.70±0.18	47.69±0.49	27.99±0.11	29.11±0.15	94.18±0.09
FedPHP	99.73±0.00	90.01±0.00	63.09±0.04	88.92±0.02	50.52±0.16	35.69±3.26	29.90±0.51	94.38±0.12
FedFomo	99.83±0.00	91.85±0.02	62.49±0.22	88.06±0.02	45.39±0.45	26.33±0.22	26.84±0.11	95.84±0.15
APPLE	99.75±0.01	90.97±0.05	65.80±0.08	89.37±0.11	53.22±0.20	35.04±0.47	39.93±0.52	95.63±0.21
PartialFed	99.86±0.01	89.60±0.13	61.39±0.12	87.38±0.08	48.81±0.20	35.26±0.18	37.50±0.16	85.20±0.16
FedALA	**99.88±0.01**	**92.44±0.02**	**67.83±0.06**	**90.67±0.03**	**55.92±0.03**	**40.54±0.02**	**41.94±0.05**	**96.52±0.08**

Jianqing Zhang, Yang Hua, Hao Wang, Tao Song, Zhengui Xue, Ruhui Ma, Haibing Guan, FedALA: Adaptive Local Aggregation for Personalized Federated Learning, Proceedings of the 37th AAAI Conference on Artificial Intelligence, 2022.

联邦学习异构性---统计异构性

FedALA对在不同统计异构性和扩展性的情况都展开了测试，在多种设定下其效果均比传统方法要好。

	Heterogeneity			Scalability		Applicability of ALA			
Datasets	Tiny-ImageNet		AG News	Cifar100		Tiny-ImageNet		Cifar100	
Methods	$Dir(0.01)$	$Dir(0.5)$	$Dir(1)$	50 clients	100 clients	Acc.	Imps.	Acc.	Imps.
FedAvg	15.70±0.46	21.14±0.47	87.12±0.19	31.90±0.27	31.95±0.37	40.54±0.17	21.08	55.92±0.15	24.03
FedProx	15.66±0.36	21.22±0.47	87.21±0.13	31.94±0.30	31.97±0.24	40.53±0.26	21.16	56.18±0.65	24.19
FedAvg-C	49.88±0.11	16.21±0.05	91.38±0.21	49.82±0.11	47.90±0.12	—	—	—	—
FedProx-C	49.84±0.02	16.36±0.19	92.03±0.19	49.79±0.14	48.02±0.02	—	—	—	—
Per-FedAvg	39.39±0.30	16.36±0.13	87.08±0.26	44.31±0.20	36.07±0.24	30.90±0.28	5.83	48.68±0.36	4.40
FedRep	55.43±0.15	16.74±0.09	92.25±0.20	47.41±0.18	44.61±0.20	37.89±0.31	0.62	53.02±0.11	0.63
pFedMe	41.45±0.14	17.48±0.61	87.08±0.18	48.36±0.64	46.45±0.18	27.30±0.24	0.37	47.91±0.21	0.57
Ditto	50.62±0.02	18.98±0.05	91.89±0.17	54.22±0.04	52.89±0.22	40.75±0.06	8.60	56.33±0.07	3.46
FedAMP	48.42±0.06	12.48±0.21	83.35±0.05	44.39±0.35	40.43±0.17	28.18±0.20	0.19	48.03±0.23	0.34
FedPHP	48.63±0.02	21.09±0.07	90.52±0.19	52.44±0.16	49.70±0.31	40.16±0.24	4.47	54.28±0.21	3.76
FedFomo	46.36±0.54	11.59±0.11	91.20±0.18	42.56±0.33	38.91±0.08	—	—	—	—
APPLE	48.04±0.10	24.28±0.21	84.10±0.18	55.06±0.20	52.81±0.29	—	—	—	—
PartialFed	49.38±0.02	24.20±0.10	91.01±0.28	48.95±0.07	39.31±0.01	35.40±0.02	0.14	48.99±0.05	0.18
FedALA	**55.75±0.02**	**27.85±0.06**	**92.45±0.10**	**55.61±0.02**	**54.68±0.57**	—	—	—	—

Jianqing Zhang, Yang Hua, Hao Wang, Tao Song, Zhengui Xue, Ruhui Ma, Haibing Guan, FedALA: Adaptive Local Aggregation for Personalized Federated Learning, Proceedings of the 37th AAAI Conference on Artificial Intelligence, 2022.

联邦学习异构性---统计异构性

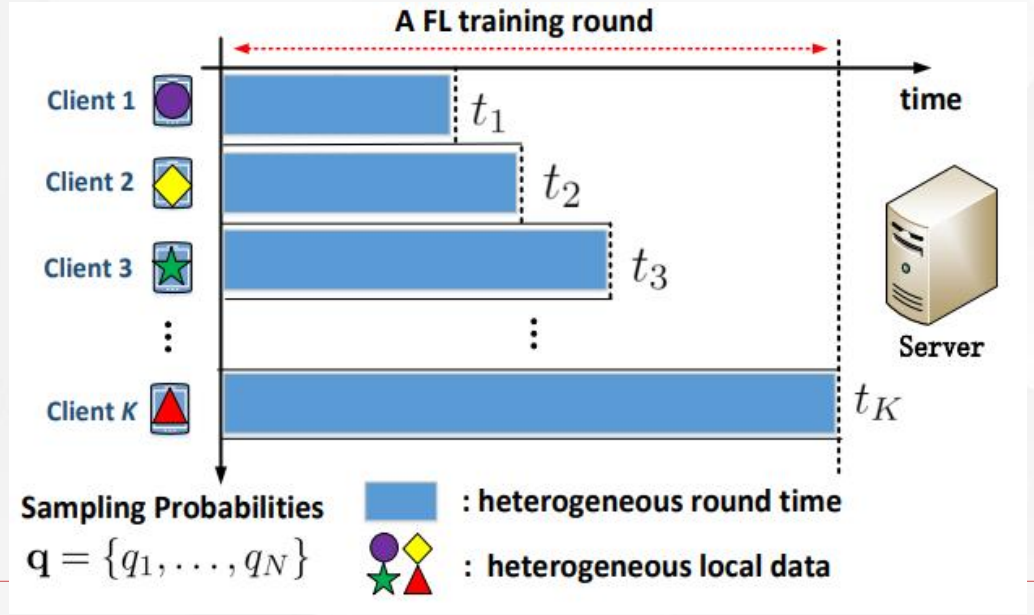
- 同时算法相比其他baseline，减少了一定的时间，提高了整体的计算效率和训练效率
- 其通信量和FedAvg相当，并没有引入额外的传输

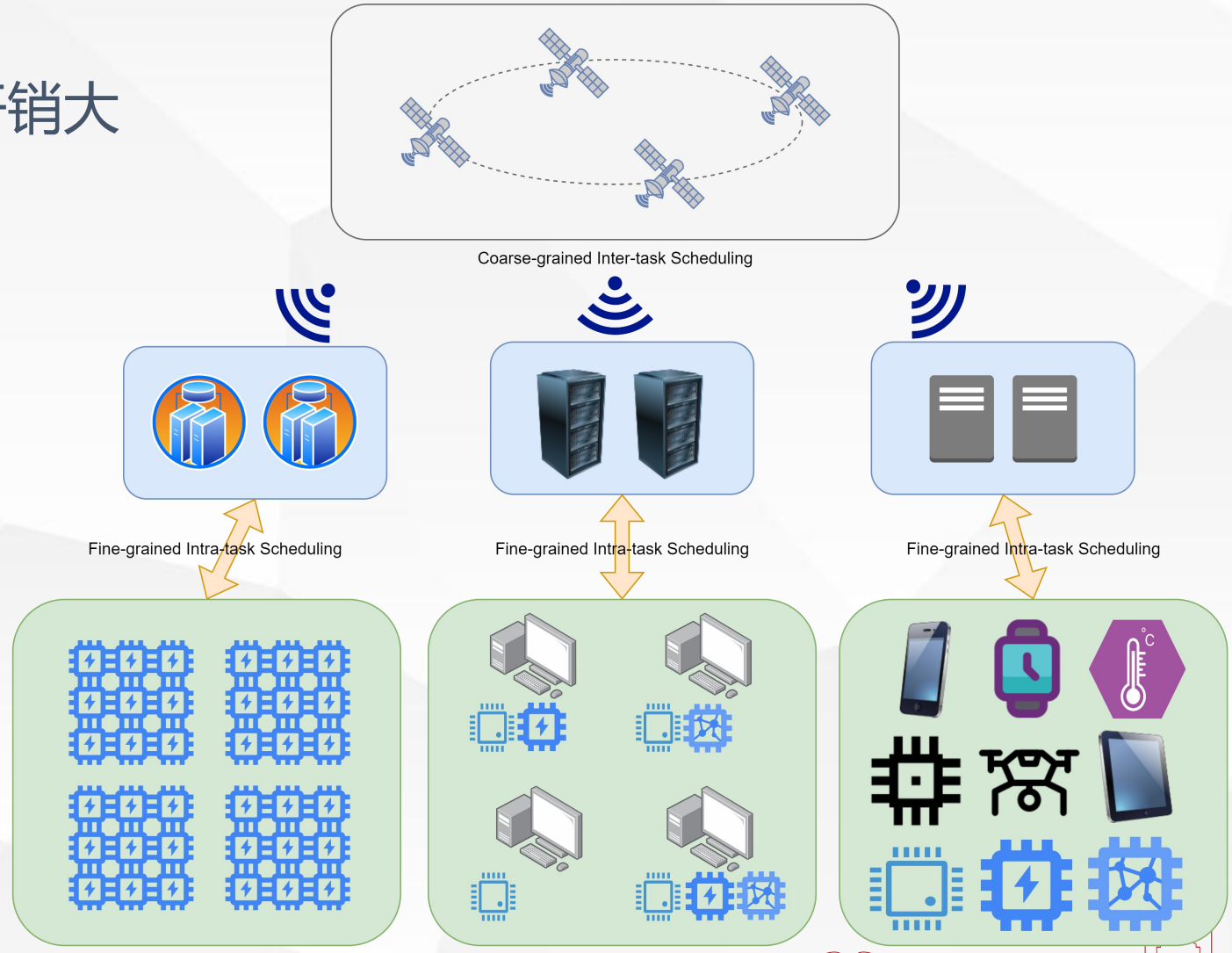
	Computation		Communication
	Total time	Time/iter.	Param./iter.
FedAvg	365 min	1.59 min	$2 * \Sigma$
FedProx	325 min	1.99 min	$2 * \Sigma$
FedAvg-C	607 min	24.28 min	$2 * \Sigma$
FedProx-C	711 min	28.44 min	$2 * \Sigma$
Per-FedAvg	121 min	3.56 min	$2 * \Sigma$
FedRep	471 min	4.09 min	$2 * \alpha_f * \Sigma$
pFedMe	1157 min	10.24 min	$2 * \Sigma$
Ditto	318 min	11.78 min	$2 * \Sigma$
FedAMP	92 min	1.53 min	$2 * \Sigma$
FedPHP	264 min	4.06 min	$2 * \Sigma$
FedFomo	193 min	2.72 min	$(1 + M) * \Sigma$
APPLE	132 min	2.93 min	$(1 + M) * \Sigma$
PartialFed	693 min	2.13 min	$2 * \Sigma$
FedALA	7+116 min	1.93 min	$2 * \Sigma$

Jianqing Zhang, Yang Hua, Hao Wang, Tao Song, Zhengui Xue, Ruhui Ma, Haibing Guan, FedALA: Adaptive Local Aggregation for Personalized Federated Learning, Proceedings of the 37th AAAI Conference on Artificial Intelligence, 2022.

联邦学习异构性---系统异构性

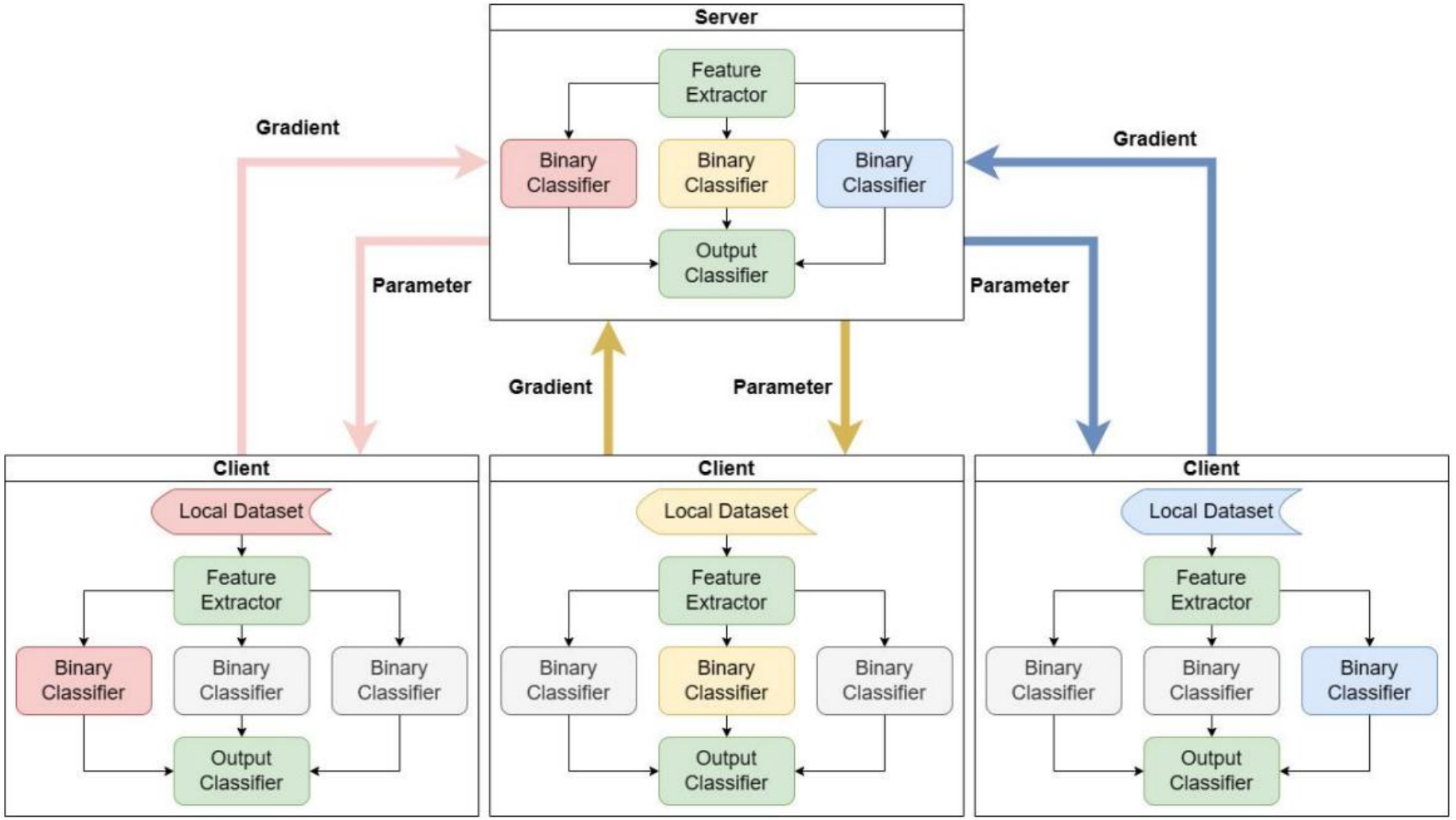
- 系统异构
  - 各客户端计算能力不同，通信开销大
- 适应性训练调整
  - 采用不同模型不同训练参数
  - 采用异步或者回合制聚合





联邦学习异构性---系统异构性

- 将多分类模型分解成多个二分类模型
- 训练时只对**本地标签对应的分支**进行训练
- 传输时只传输**经过训练的模型分支**
- 聚合时只聚合**传输上来的对应分支**

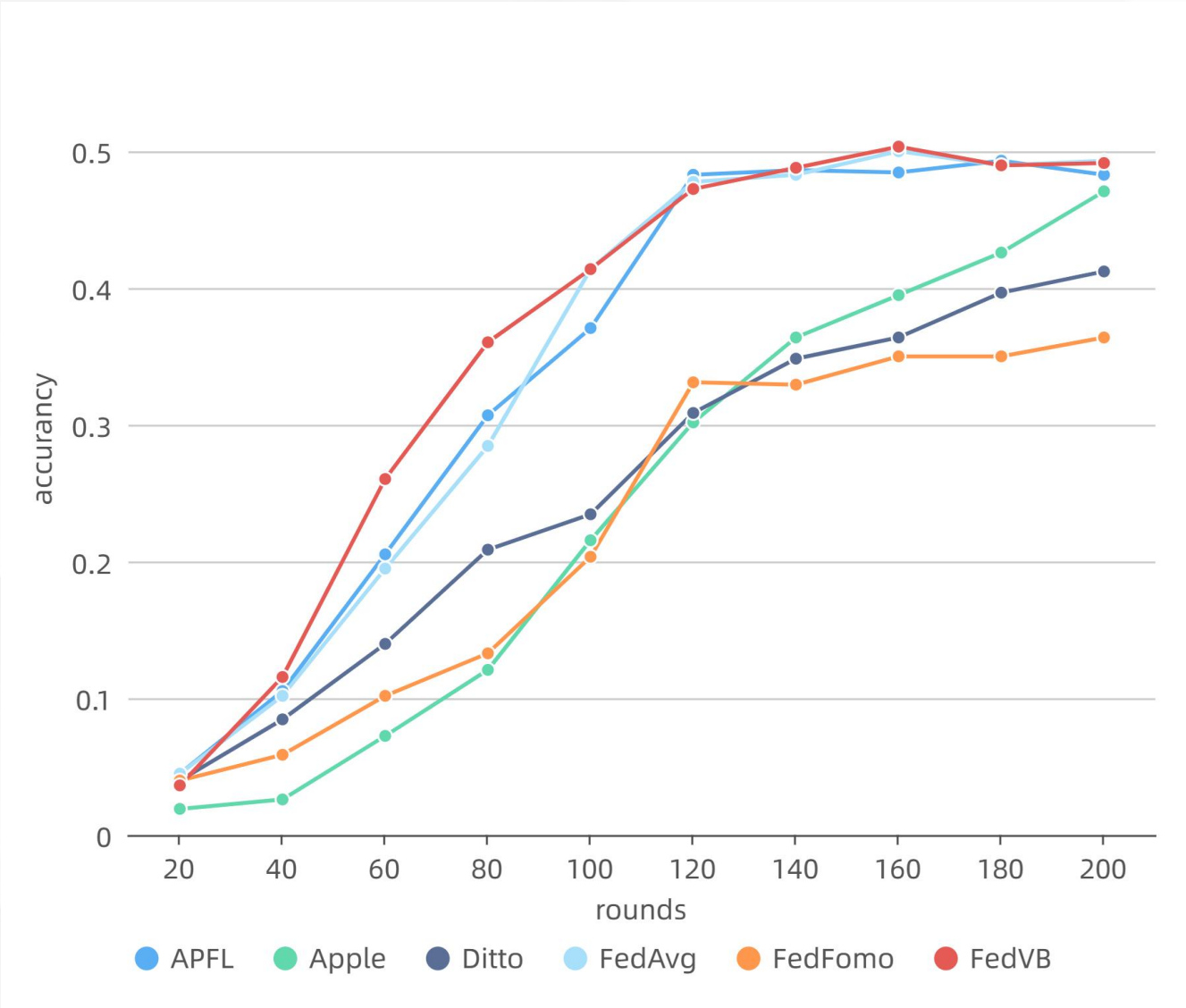


Yunhui Wang, Weichu Zheng, Zifei Liu, Jinyan Wang, Hongjian Shi, Minyu Gu, Ruhui Ma, and Haibing Guan, A Federated Network Intrusion Detection System with Multi-branch Network and Vertical blocking Aggregation, 2023.

联邦学习异构性---系统异构性

- FedVB相比其他baseline收敛更快，准确率更高
- 多分支模型相比传统模型能够在准确率上有所提升

Algorithm	Simple classifier	Multi-branch Network
FedAvg	0.3983	0.5000
APFL	0.4276	0.5000
Apple	0.3983	0.4707
Ditto	0.3137	0.4259
FedBN	0.3983	0.5017
FedMTL	0.3224	0.5017
FedProx	0.4000	0.5017
pFedMe	0.0259	0.0379
FedFomo	0.8897	0.8500
LOCAL	0.8638	0.8621
PerAvg	0.6000	0.7483



Yunhui Wang, Weichu Zheng, Zifei Liu, Jinyan Wang, Hongjian Shi, Minyu Gu, Ruhui Ma, and Haibing Guan, A Federated Network Intrusion Detection System with Multi-branch Network and Vertical blocking Aggregation, 2023.

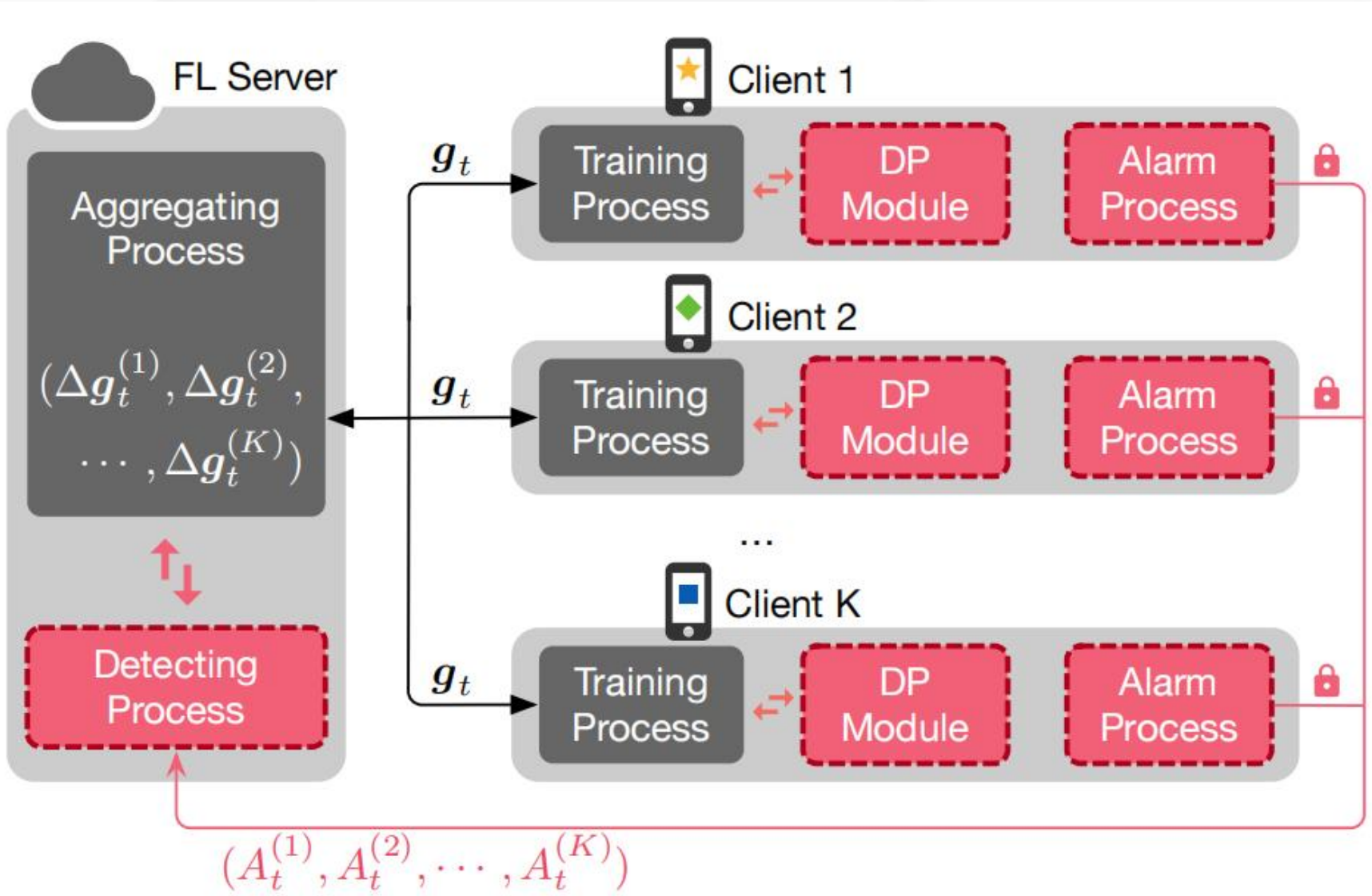
03

多技术融合成果

- 效率 + 安全
- 效率 + 异构性
- 安全 + 异构性
- 效率 + 安全 + 异构性

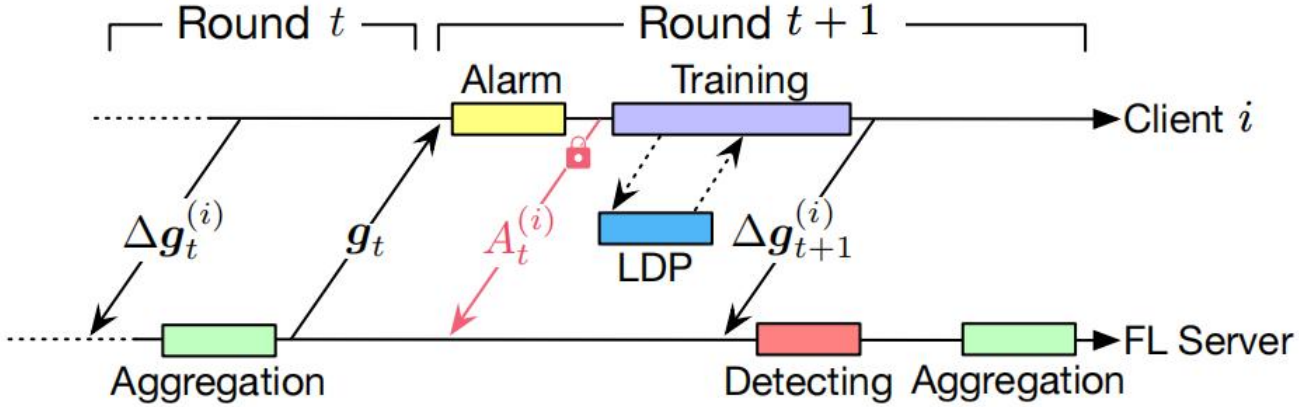
联邦学习安全

- 在原本Siren的基础上融入和差分隐私以进一步保障数据隐私



Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Ruhui Ma, Xiulang Jin, Zhengui Xue, Haibing Guan, SIREN+: Robust Federated Learning with Proactive Alarming and Differential Privacy, IEEE Transactions on Dependable and Secure Computing, 2023.

- 在原本Siren基础上将客户端部分加入了差分隐私




```
// training process
function ClientUpdate(i, g_t)
    A_t^(i) ← Alarm(g_t, g_t^(i));
    if A_t^(i) is 0 then
        g^(i) ← g_t;
    else
        g^(i) ← g_t^(i);
    for each epoch e = 1, ··· , E do
        if not use DP then
            // normal local training
            train the model g^(i) with optimizer on the
             local dataset D^(i), and obtain g_{t+1}^(i);
        else
            // local training with DP
            train the model g^(i) with LDP on the local
             dataset D^(i), and obtain g_{t+1}^(i);
    // calculate model updates
    Δg_{t+1}^(i) ← g_{t+1}^(i) − g_t;
    return Δg_{t+1}^(i);
```

```
// alarming process
function Alarm(g_t, g_t^(i))
    ω_t ← testing g_t on the local test dataset D_0^(i);
    ω_t^(i) ← testing g_t^(i) on the local test dataset D_0^(i);
    if ω_t ≥ ω_t^(i) · (1 − C_c) then
        // the global model is normal
        A_t^(i) ← 0;
    else
        // the global model is abnormal
        A_t^(i) ← 1;
    send A_t^(i) in a secure tunnel to the FL server;
    return A_t^(i);
```

Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Ruhui Ma, Xiulang Jin, Zhengui Xue, Haibing Guan, SIREN+: Robust Federated Learning with Proactive Alarming and Differential Privacy, IEEE Transactions on Dependable and Secure Computing, 2023.

- 与原本Siren性能相当

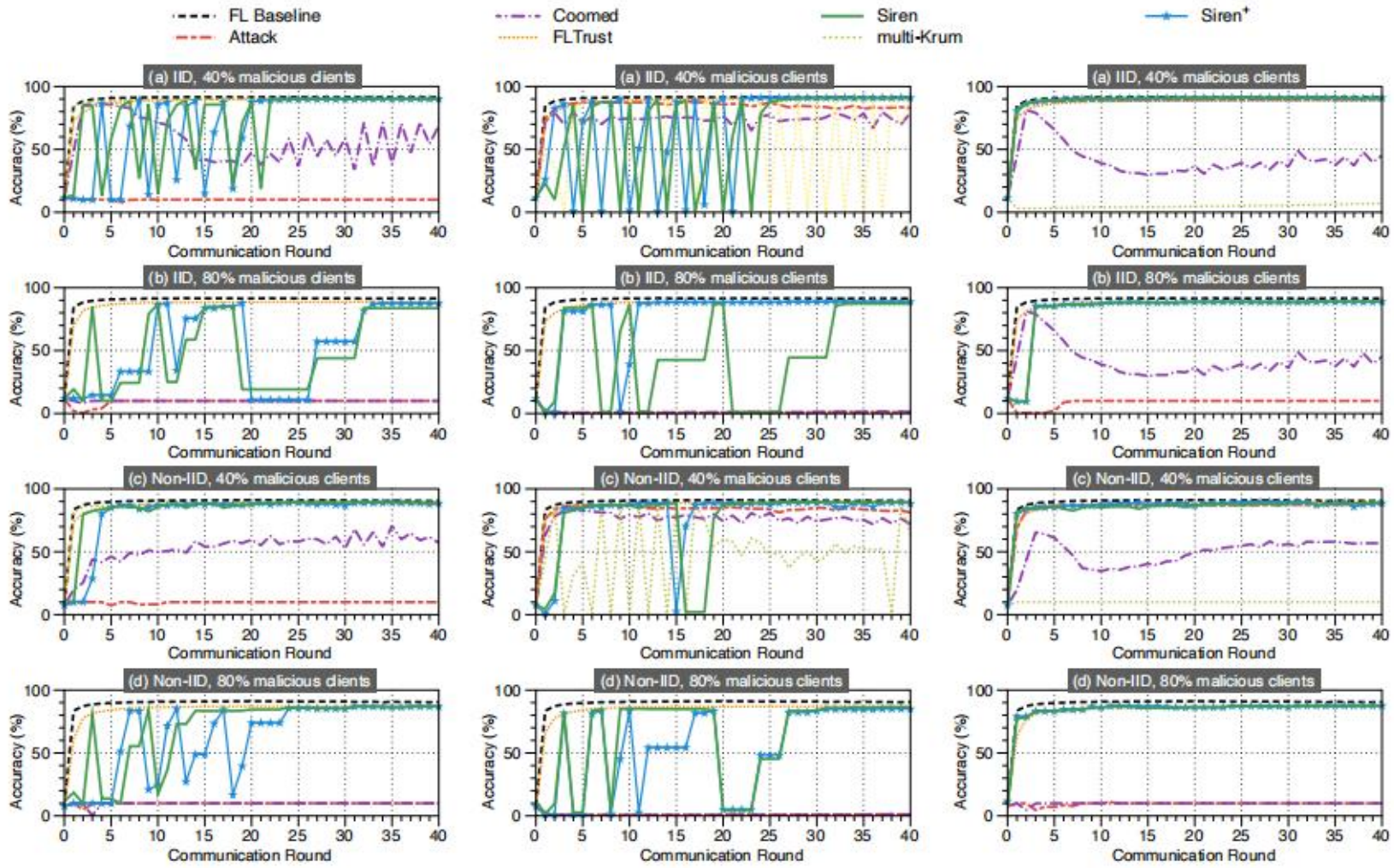


Fig. 5: Training efficiency under sign-flipping attack when $|K| = 10$.

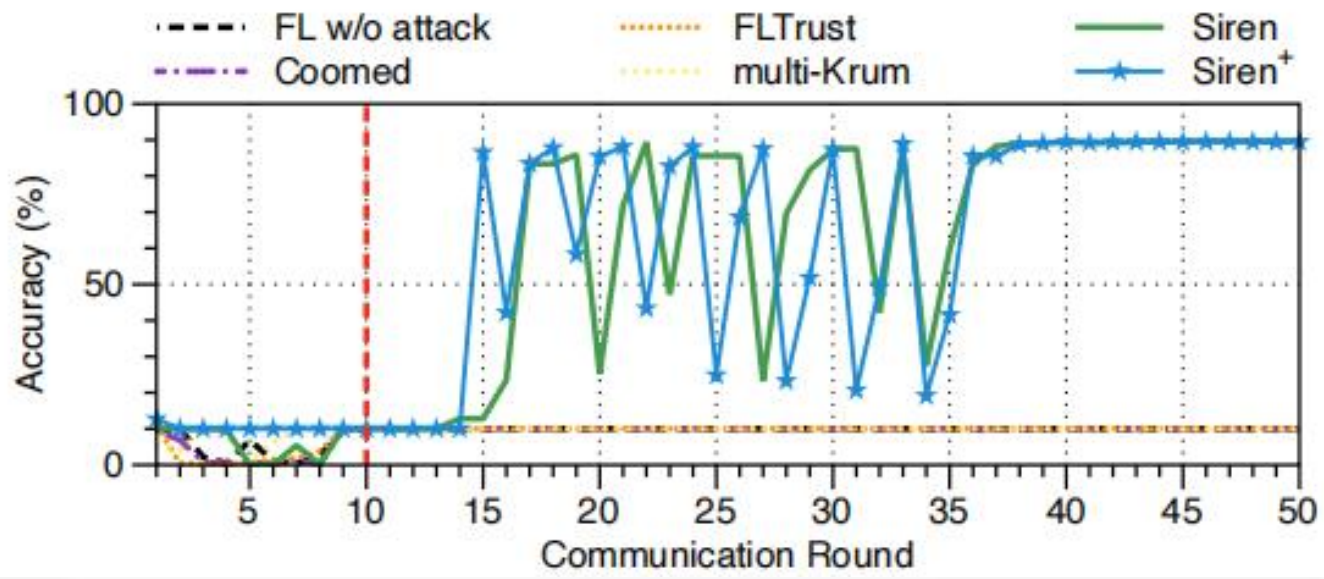
Fig. 6: Training efficiency under label-flipping attack when $|K| = 10$.

Fig. 7: Training efficiency under adaptive attack when $|K| = 10$.

Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Ruhui Ma, Xiulang Jin, Zhengui Xue, Haibing Guan, SIREN+: Robust Federated Learning with Proactive Alarming and Differential Privacy, IEEE Transactions on Dependable and Secure Computing, 2023.

联邦学习安全

- Siren+延续了Siren的训练恢复能力
- Siren+特别对Adaptive Attack进行了测试，依旧有着较强的防御效果

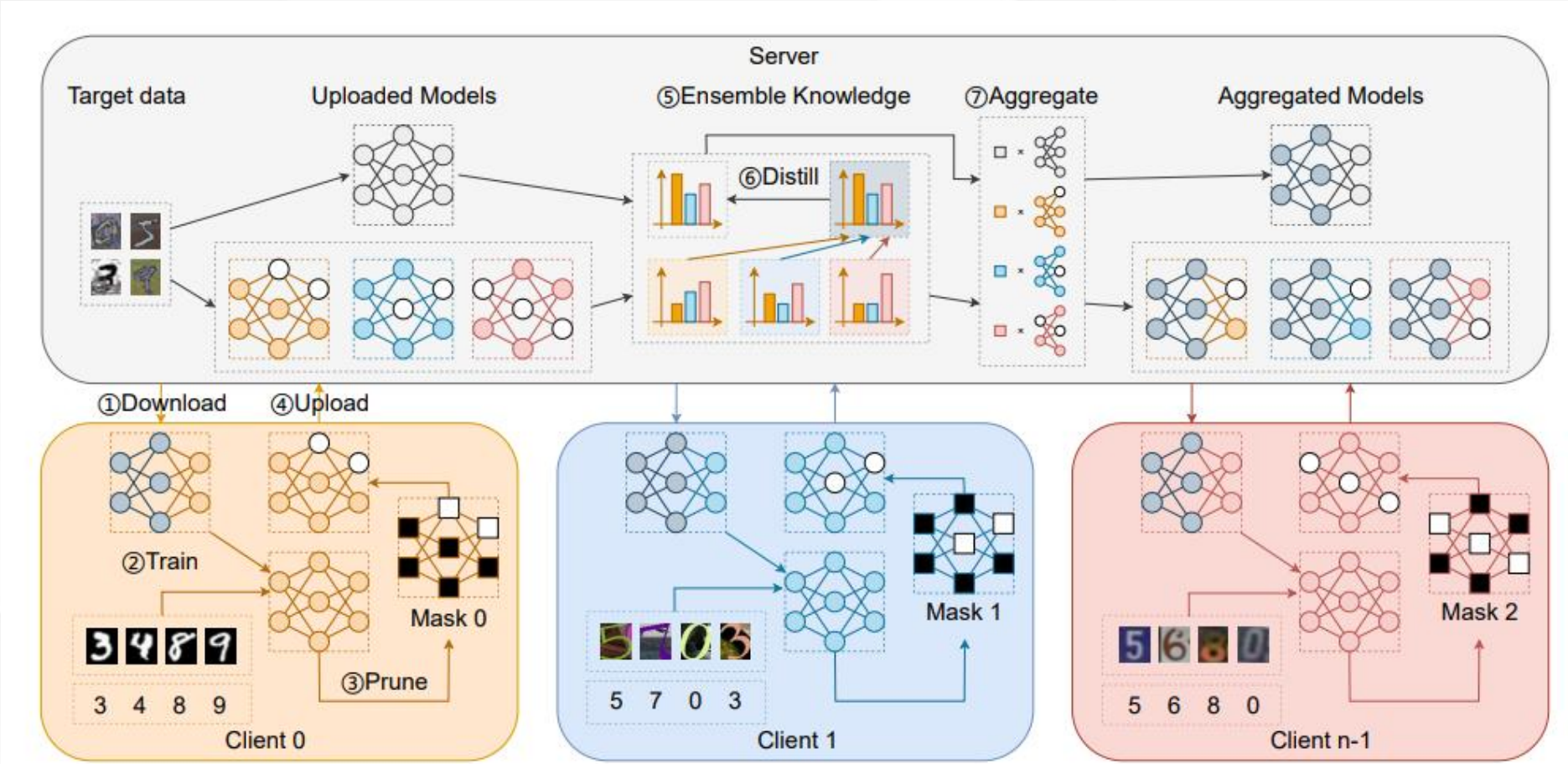


Attack Type	Malicious Client Proportion	Accuracy
None	0%	80.23%
Sign-flipping	20%	76.32%
	40%	74.4%
	80%	63.81%
Label-flipping	20%	76.94%
	40%	75.42%
	80%	64.53%
Adaptive Attack	20%	78.41%
	40%	75.79%
	80%	61.57%

Hanxi Guo, Hao Wang, Tao Song, Yang Hua, Ruhui Ma, Xiulang Jin, Zhengui Xue, Haibing Guan, SIREN+: Robust Federated Learning with Proactive Alarming and Differential Privacy, IEEE Transactions on Dependable and Secure Computing, 2023.

联邦学习效率 + 异构性

- 跨域联邦学习中，源域模型不一定适用于目标域
- 标签蒸馏：聚合源域标签，用于训练目标域模型



Ruhui Ma, Hongjian Shi, Honghao Gao, Haibing Guan, Muddesar Iqbal, Shahid Mumtaz, cFedDT: Cross-domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products, IEEE Transactions on Consumer Electronics, 2023.

联邦学习效率 + 异构性

适应性训练批次调整

- 解决系统异构性

减少客户端训练时间差异减少等待时间

$$B_i^{t+1} = B_i^t * \frac{\tau^{t+1}}{\tau_i^t} \qquad \tau^{t+1} = \frac{1}{N}\sum_{i=0}^{n-1}\tau_i^t$$

减少更新步长提高收敛效果

$$B^{t+1} = B^t * \frac{t}{t+1}$$

$$B_i^{t+1} = B_i^t * \frac{\frac{1}{N}\sum_{i=0}^{n-1}\tau_i^t}{\tau_i^t} * \frac{t_{max}-t}{t_{max}}$$

模型剪枝

- 减少网络传输量

$$m_{ik}^t = I(|\hat{\Theta}_{ik}^t| \geq h_i^t) \rightarrow \Theta_i'^t = \hat{\Theta}_i^t \odot M_i^t$$

基于比例的剪枝阈值设计

$$h_i^t = \min_k(|\hat{\Theta}_{ik}^t|) + p * [\max_k(|\hat{\Theta}_{ik}^t|) - \min_k(|\hat{\Theta}_{ik}^t|)]$$

基于分位数的剪枝阈值设计

$$h_i^t = quantile(|\hat{\Theta}_{ik}^t|)$$

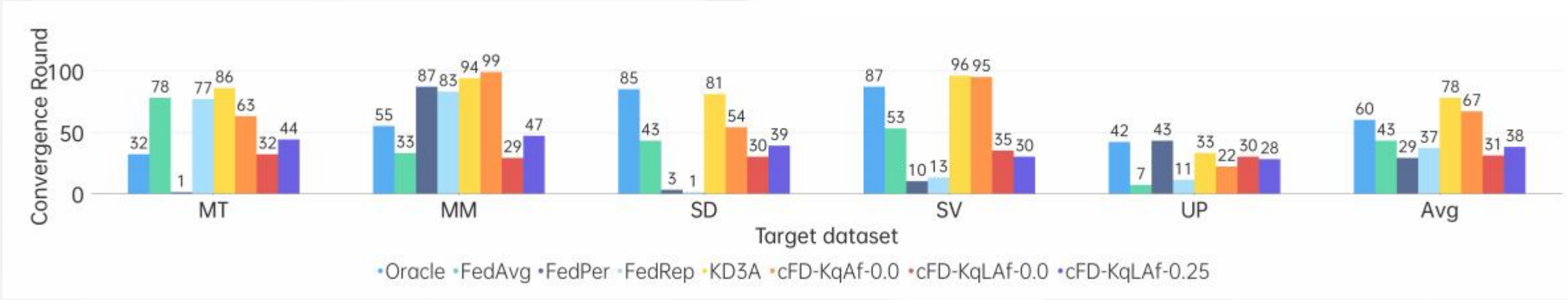
Ruhui Ma, Hongjian Shi, Honghao Gao, Haibing Guan, Muddesar Iqbal, Shahid Mumtaz, cFedDT: Cross-domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products, IEEE Transactions on Consumer Electronics, 2023.

联邦学习效率 + 异构性

- cFedDT能够在跨域联邦学习的场景下从源域学习到更准确的知识，为目标域构造更加准确的pesudo-label
- 同时其收敛效率在大部分情况下比其他SOTA算法高

TABLE II
THE TARGET ACCURACIES FOR SOTA ALGORITHMS.

Algorithm	Target Dataset					Avg.
	MT	MM	SD	SV	UP	
Oracle	0.996	0.980	0.990	0.927	0.994	0.977
FedAvg	0.993	0.739	0.902	0.822	0.969	0.885
FedPer	0.119	0.116	0.095	0.098	0.061	0.098
FedRep	0.074	0.122	0.095	0.104	0.201	0.119
KD3A	0.993	0.880	0.913	0.885	0.978	0.932
cFD-KqAf-0.0	**0.994**	0.886	0.915	**0.888**	**0.977**	0.932
cFD-KqLAf-0.0	0.991	**0.978**	0.911	0.884	0.974	**0.948**
cFD-KqLAf-0.25	0.993	0.958	**0.916**	0.876	0.974	0.943



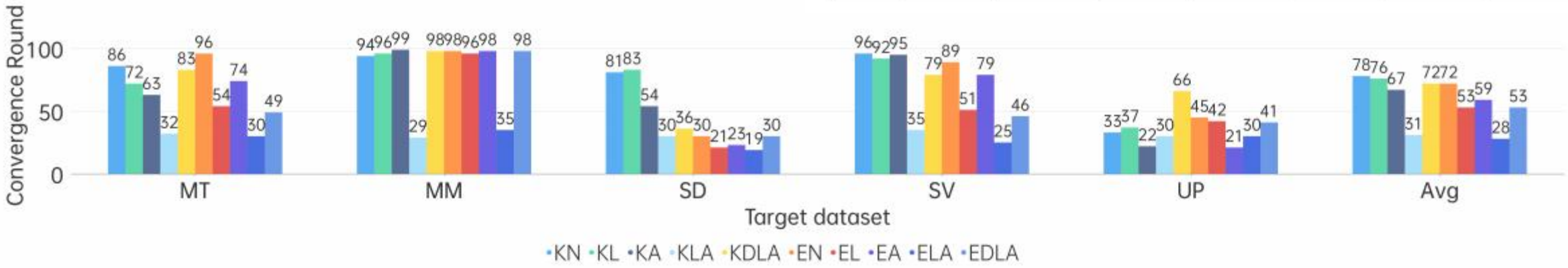
Ruhui Ma, Hongjian Shi, Honghao Gao, Haibing Guan, Muddesar Iqbal, Shahid Mumtaz, cFedDT: Cross-domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products, IEEE Transactions on Consumer Electronics, 2023.

联邦学习效率 + 异构性

- 批次量调整时，减少客户端用时差距和批次量递减均对准确率有正面效果
- 合并使用两种批次调整方案能够有效的提高收敛效率，单一任意一种都没有办法做到

TABLE III
THE TARGET ACCURACIES FOR DIFFERENT BATCH ADJUSTMENT METHODS.

KE	BA	Target Dataset					Avg.
		MT	MM	SD	SV	UP	
K	N	0.993	0.880	0.913	0.885	0.978	0.930
	L	**0.994**	0.914	0.912	0.883	0.980	0.937
	A	**0.994**	0.886	**0.915**	**0.888**	0.977	0.932
	LA	0.991	**0.978**	0.911	0.884	0.974	**0.948**
	DLA	0.993	0.914	0.913	0.880	0.980	0.936
E	N	0.992	0.895	0.884	0.864	0.978	0.923
	L	0.989	0.913	0.880	0.860	**0.981**	0.925
	A	0.992	0.901	0.883	0.860	0.979	0.923
	LA	0.989	0.944	0.870	0.844	0.973	0.924
	DLA	0.992	0.914	0.876	0.857	0.977	0.923



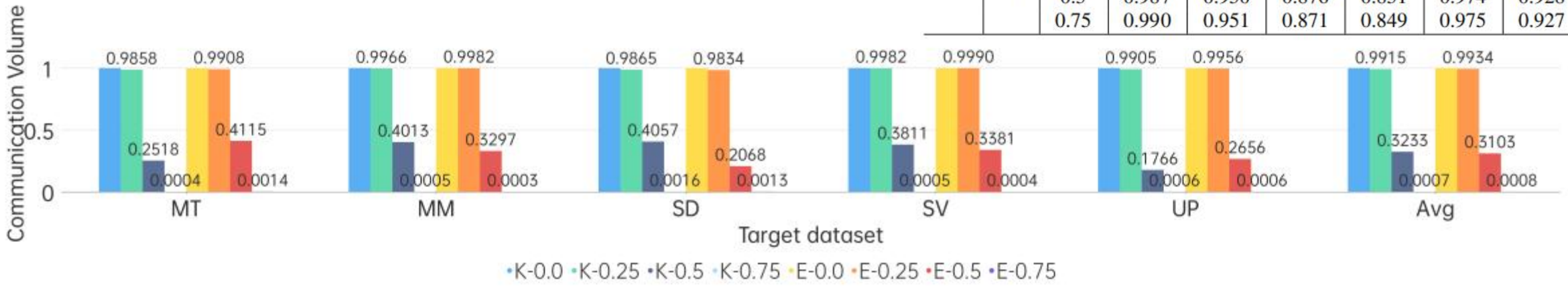
Ruhui Ma, Hongjian Shi, Honghao Gao, Haibing Guan, Muddesar Iqbal, Shahid Mumtaz, cFedDT: Cross-domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products, IEEE Transactions on Consumer Electronics, 2023.

联邦学习效率 + 异构性

- 采用基于分位数的剪枝阈值设计方案相对而言对准确率的负面影响较小
- 压缩率不同对收敛速度基本没有影响
- 但能够将原本模型参数量压缩至0.05%

TABLE IV
THE TARGET ACCURACIES FOR DIFFERENT PRUNING TYPES AND PRUNING RATIOS.

KE	PT	p	Target Dataset					Avg.
			MT	MM	SD	SV	UP	
K	r	0.0	0.991	0.959	0.909	0.878	0.972	0.942
		0.25	0.992	0.961	0.914	0.878	**0.977**	0.944
		0.5	0.991	0.949	0.912	0.879	**0.977**	0.942
		0.75	**0.993**	0.954	0.912	0.881	0.975	0.943
	q	0.0	0.991	**0.978**	0.884	**0.911**	0.974	**0.948**
		0.25	**0.993**	0.958	**0.916**	0.876	0.974	0.943
		0.5	0.992	0.956	0.912	0.875	0.973	0.942
		0.75	0.991	0.951	0.905	0.876	0.973	0.939
E	r	0.0	0.988	0.941	0.875	0.852	**0.977**	0.927
		0.25	0.990	0.946	0.875	0.850	0.974	0.927
		0.5	0.986	0.933	0.866	0.850	0.974	0.922
		0.75	0.990	0.951	0.873	0.846	0.971	0.926
	q	0.0	0.989	0.944	0.870	0.844	0.973	0.924
		0.25	0.988	0.946	0.875	0.839	0.970	0.924
		0.5	0.987	0.950	0.876	0.851	0.974	0.928
		0.75	0.990	0.951	0.871	0.849	0.975	0.927



Ruhui Ma, Hongjian Shi, Honghao Gao, Haibing Guan, Muddesar Iqbal, Shahid Mumtaz, cFedDT: Cross-domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products, IEEE Transactions on Consumer Electronics, 2023.

联邦学习效率 + 异构性

- 聚合部分参数（特征提取or分类器）并不能提高收敛后的准确率
- 并且对整体的收敛速度并没有太多影响

TABLE V

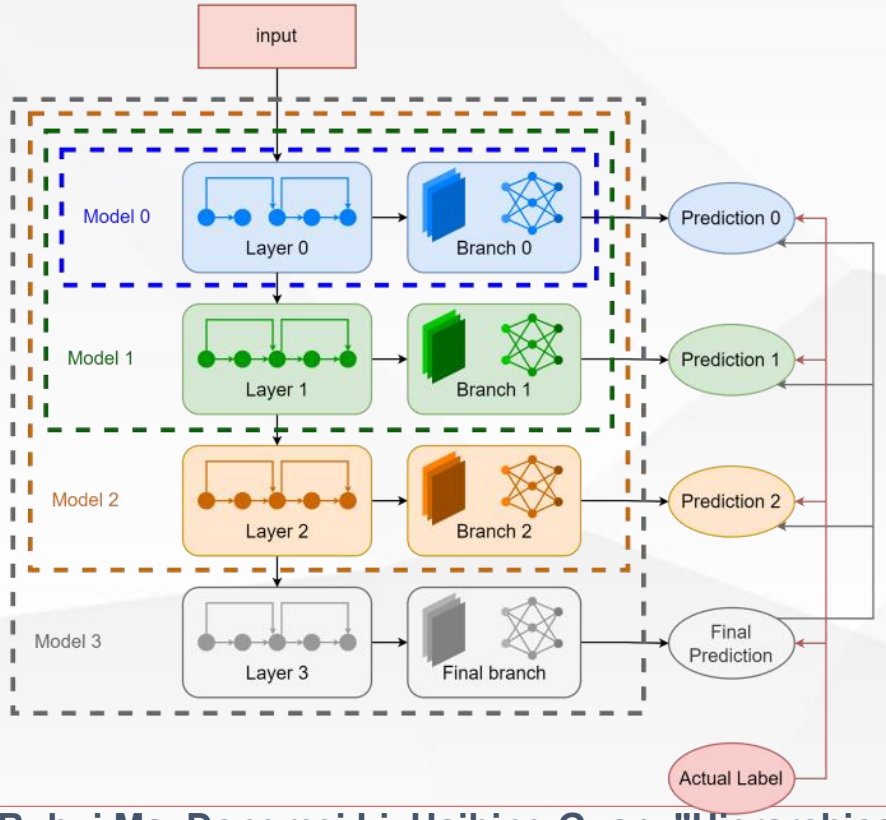
THE TARGET ACCURACIES FOR DIFFERENT AGGREGATION TYPES.

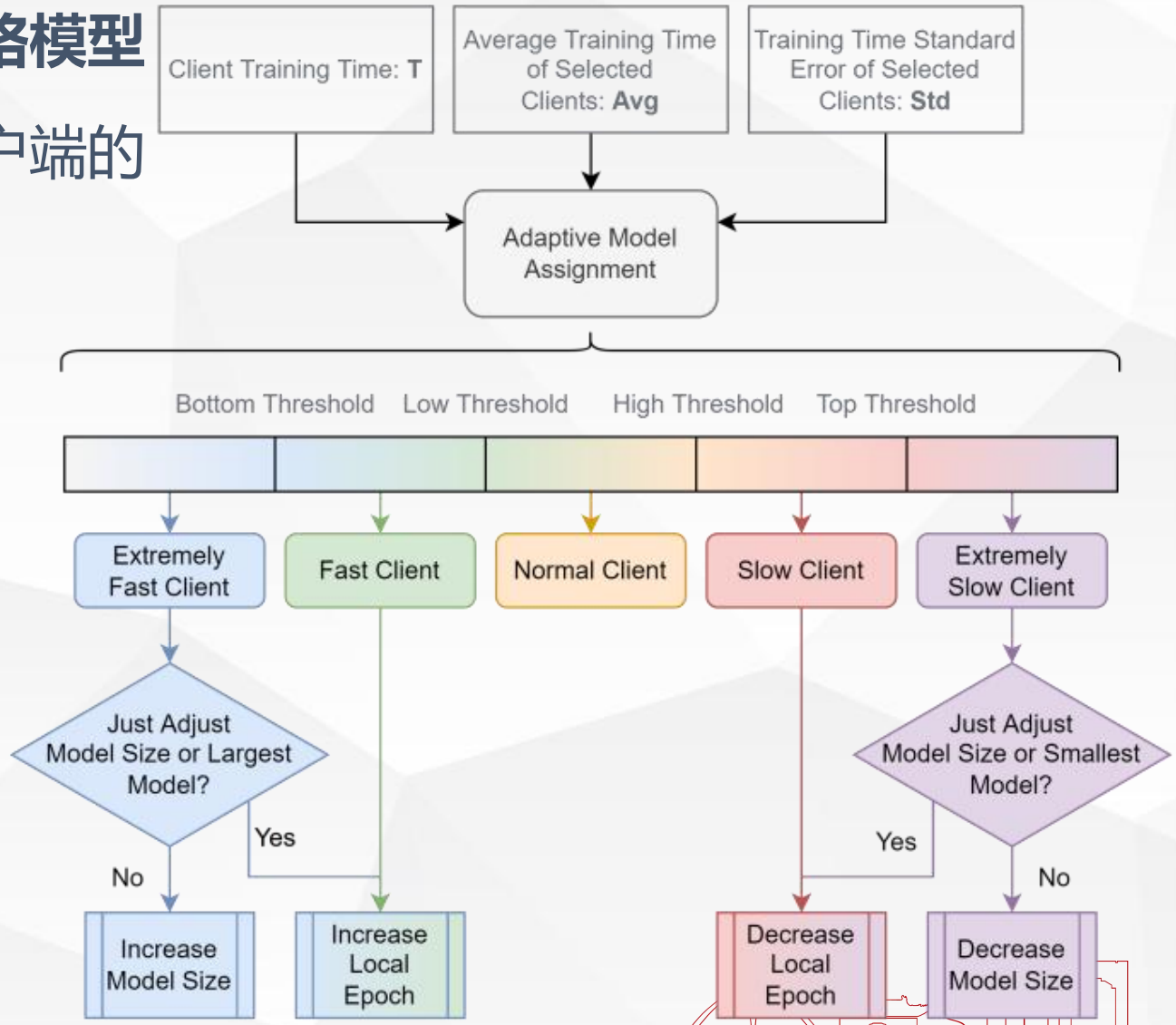
KE	AT	Target Dataset					Avg.
		MT	MM	SD	SV	UP	
K	f	0.991	**0.959**	0.909	0.878	0.972	0.942
	c	**0.993**	0.950	0.910	**0.884**	0.973	0.942
	fc	0.992	0.950	**0.914**	0.882	0.976	**0.943**
E	f	0.988	0.941	0.875	0.852	**0.977**	0.927
	c	0.988	0.951	0.871	0.849	0.966	0.925
	fc	0.986	0.942	0.877	0.851	0.969	0.925

Ruhui Ma, Hongjian Shi, Honghao Gao, Haibing Guan, Muddesar Iqbal, Shahid Mumtaz, cFedDT: Cross-domain Federated Learning in Digital Twins for Metaverse Consumer Electronic Products, IEEE Transactions on Consumer Electronics, 2023.

联邦学习效率 + 异构性

- 采用自蒸馏的方式，构造**多出口神经网络模型**
- 根据各客户端训练时间区别，调整各客户端的**模型大小，本地训练轮次**

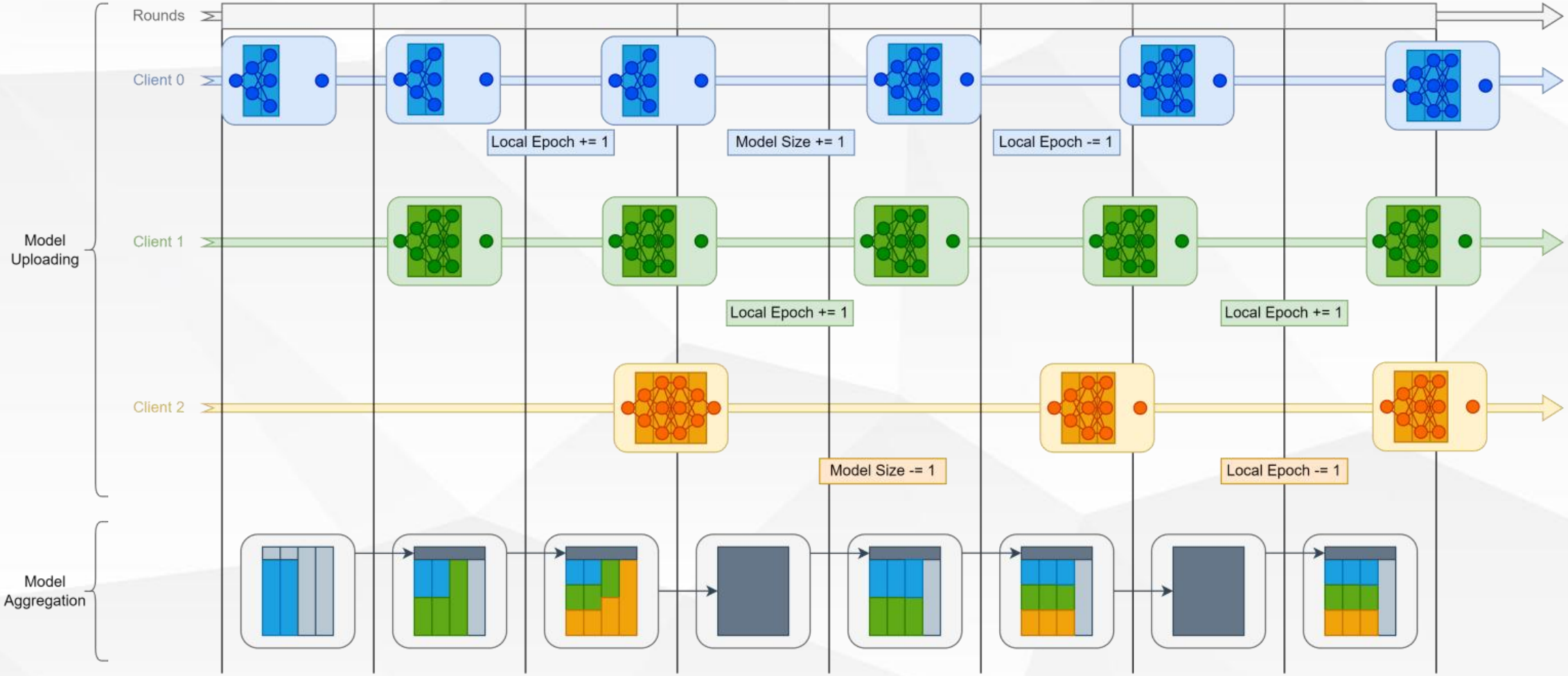




Hongjian Shi, Ruhui Ma, Dongmei Li, Haibing Guan, "Hierarchical Adaptive Collaborative Learning: A Distributed Learning Framework for Customized Cloud Services in 6G Mobile Systems," in IEEE Network，2023.

联邦学习效率 + 异构性

- 采用**回合制聚合**的方式进一步减少通信开销，处理极端设备



Hongjian Shi, Ruhui Ma, Dongmei Li, Haibing Guan, "Hierarchical Adaptive Collaborative Learning: A Distributed Learning Framework for Customized Cloud Services in 6G Mobile Systems," in IEEE Network，2023.

联邦学习效率 + 异构性

- 收敛速度更快
- 每回合训练用时更少
- 各客户端训练用时差距更小
- 总参数量变小
- 总计算量变小
- 内存读写量变小

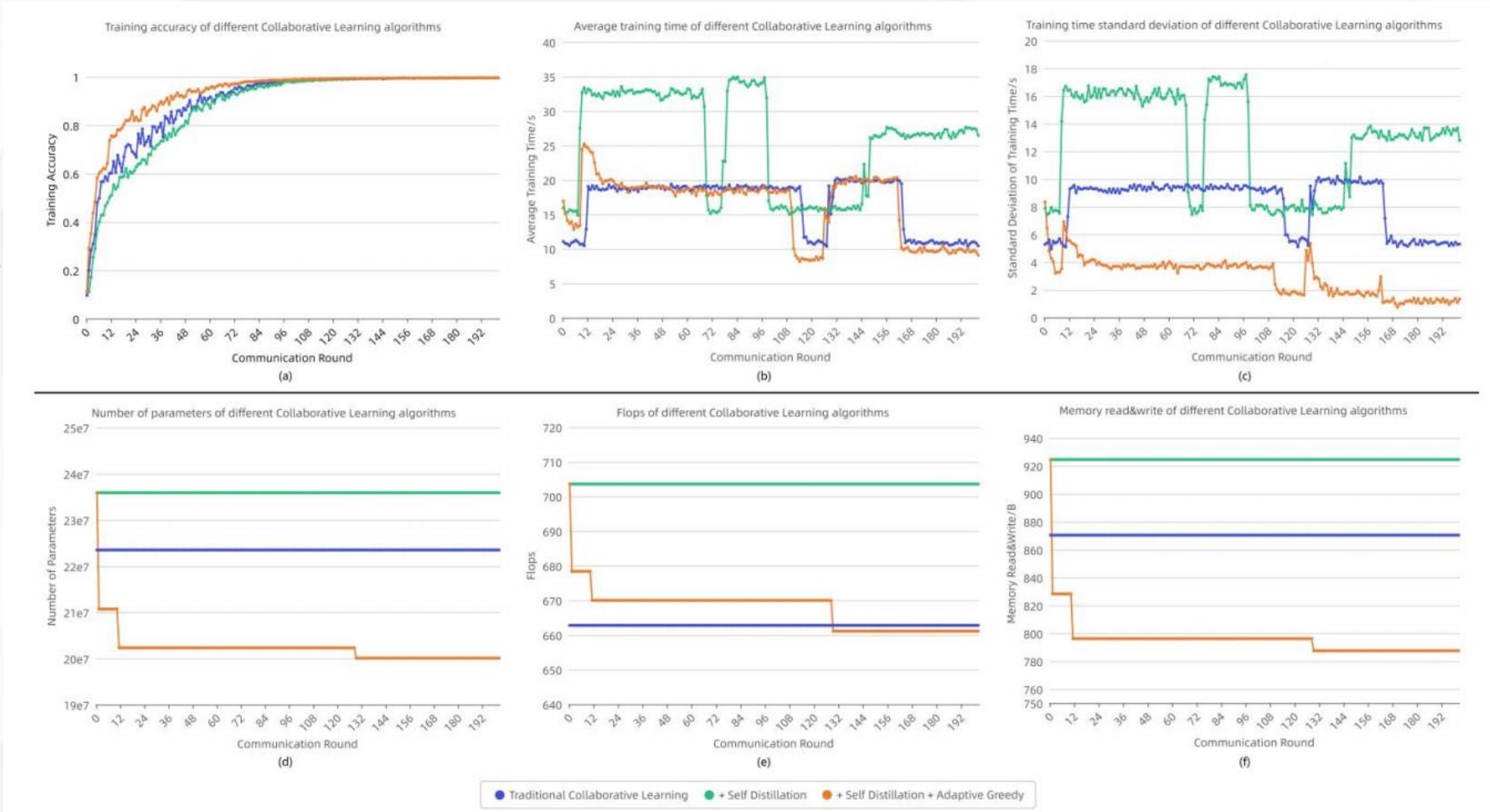
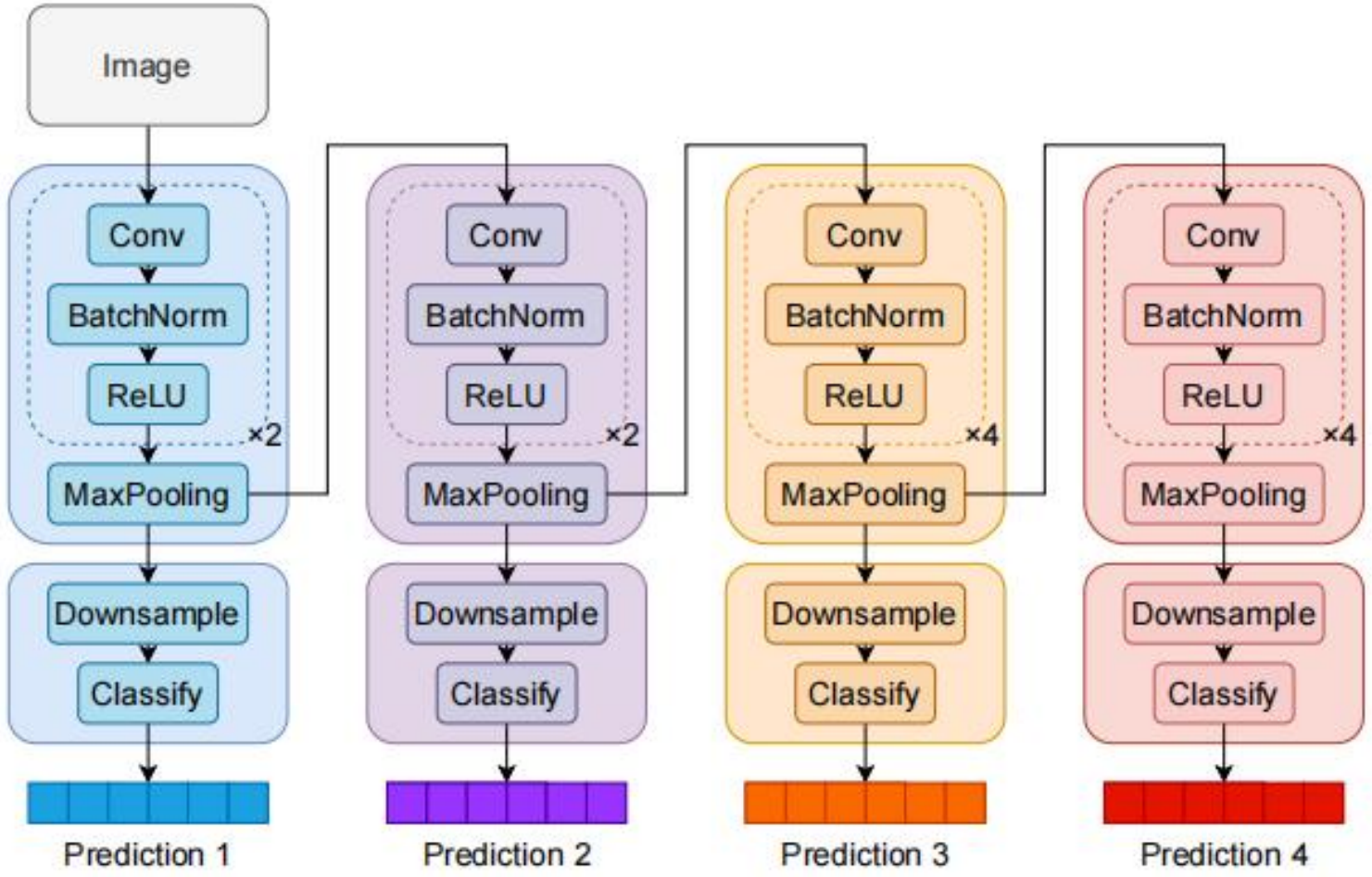


Fig.5 The training curve of different metrics in different communication rounds for the three algorithms. (a) Training accuracy. (b) Average training time. (c) Training time standard deviations. (d) The Number of parameters. (e) Training FLOPS. (f) Memory read&write volume.

Hongjian Shi, Ruhui Ma, Dongmei Li, Haibing Guan, "Hierarchical Adaptive Collaborative Learning: A Distributed Learning Framework for Customized Cloud Services in 6G Mobile Systems," in IEEE Network，2023.

联邦学习效率 + 异构性

- 考虑到不同客户端存储计算能力不同，不一定能够放下所有模型
- 采用了**多出口的神经网络模型**
- 能够给出多种大小的模型供不同计算存储能力的设备使用
- 不同大小的模型对不同大小的目标物体检测能力不同

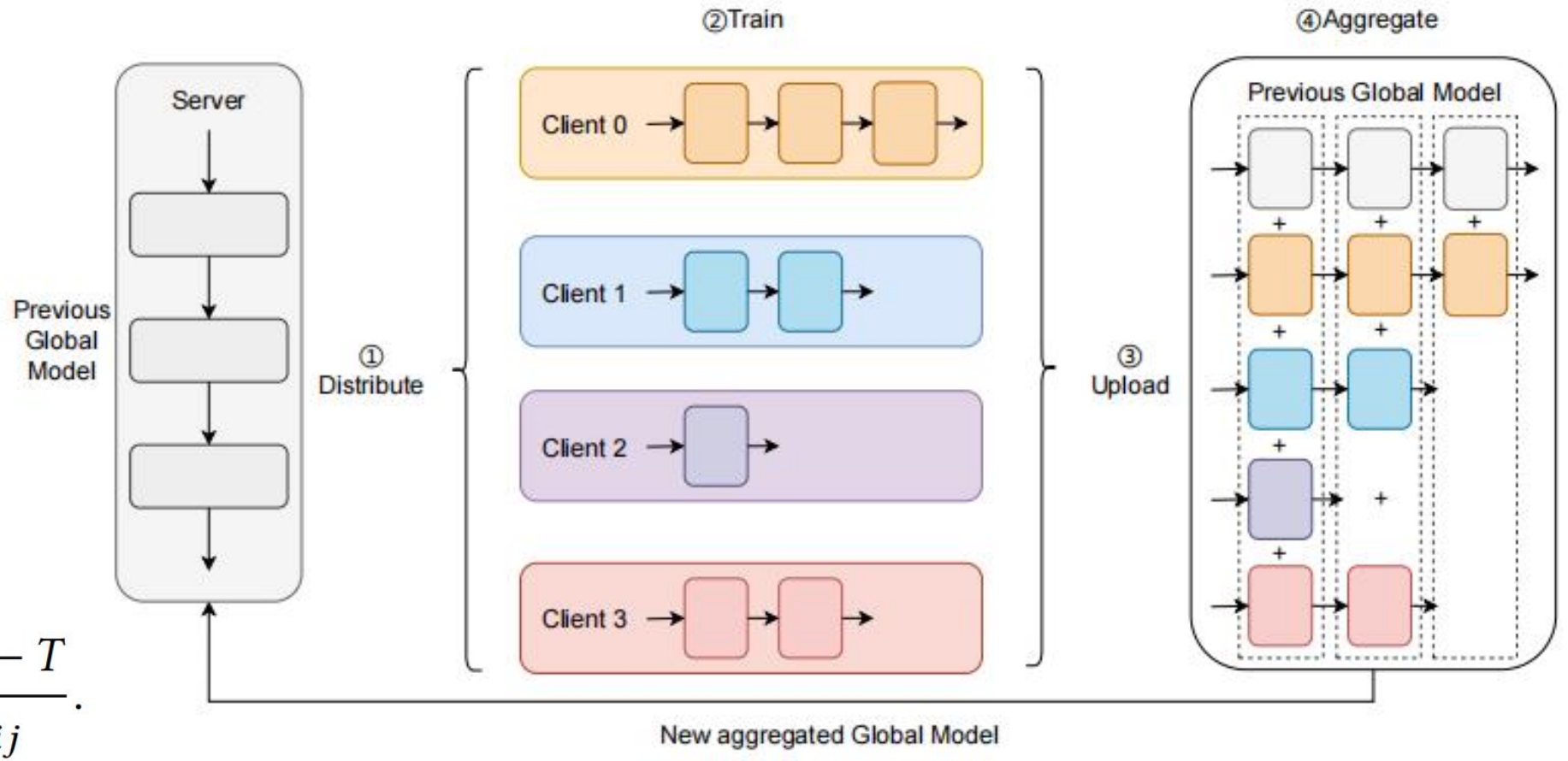


Hongjian Shi, Ilyas Bayanbayev, Ruhui Ma, Haibing Guan, Cloud-based Collaborative Agricultural Learning with Flexible Model Size and Adaptive Batch Number, ACM Transactions on Sensor Networks, 2023.

联邦学习效率 + 异构性

- 考虑到模型不同，需要设计不同部分的**聚合权重**保障收敛性
- 考虑到系统异构性，采用逼近式的训练**批次量调整**

$$B_{ij}^{t+1} = B_{ij}^{t} - \gamma * B_{ij}^{t} * \frac{\tau_{ij} - T}{\tau_{ij}}.$$



Hongjian Shi, Ilyas Bayanbayev, Ruhui Ma, Haibing Guan, Cloud-based Collaborative Agricultural Learning with Flexible Model Size and Adaptive Batch Number, ACM Transactions on Sensor Networks, 2023.

联邦学习效率 + 异构性

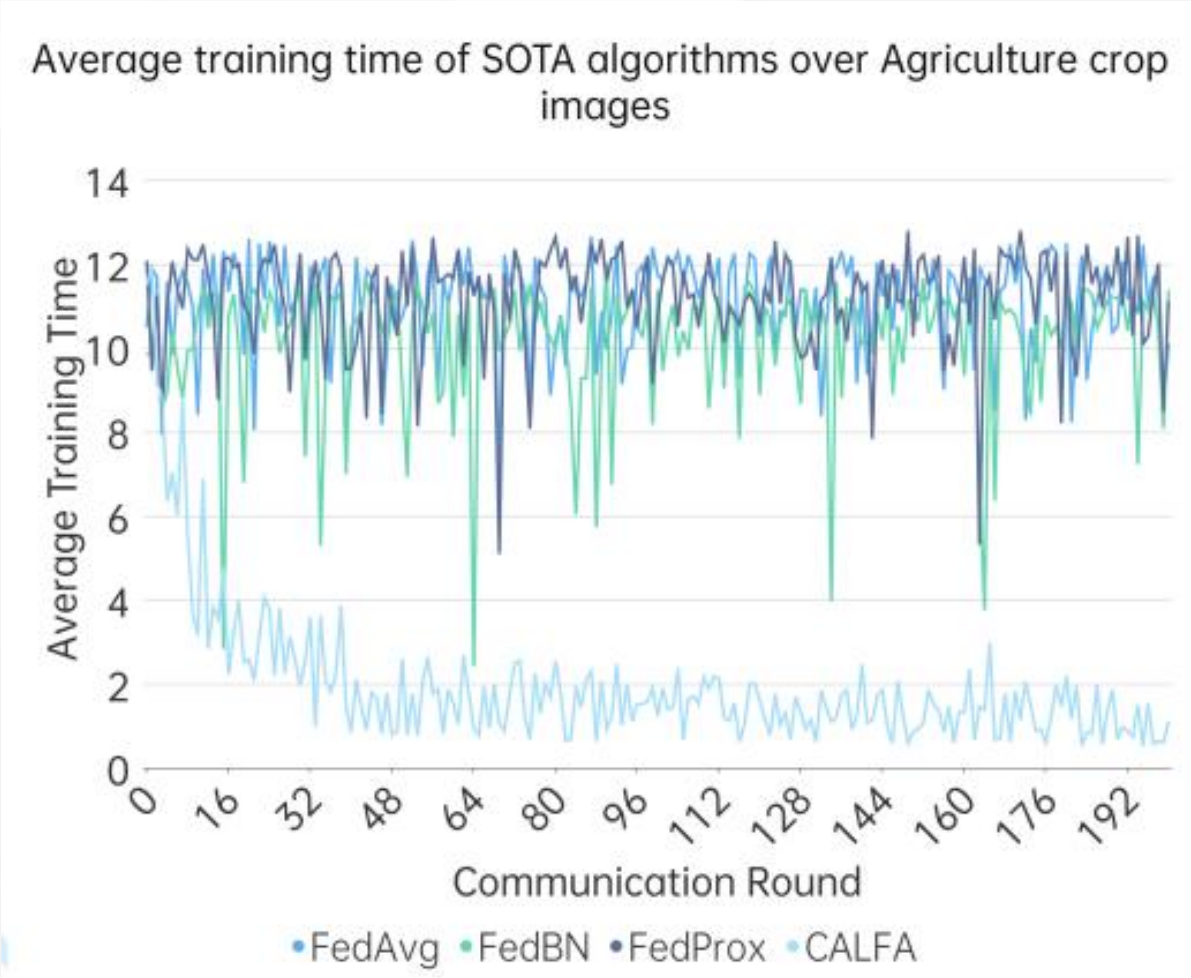
- 能够在拥有和原本FedAvg等算法相当的准确率

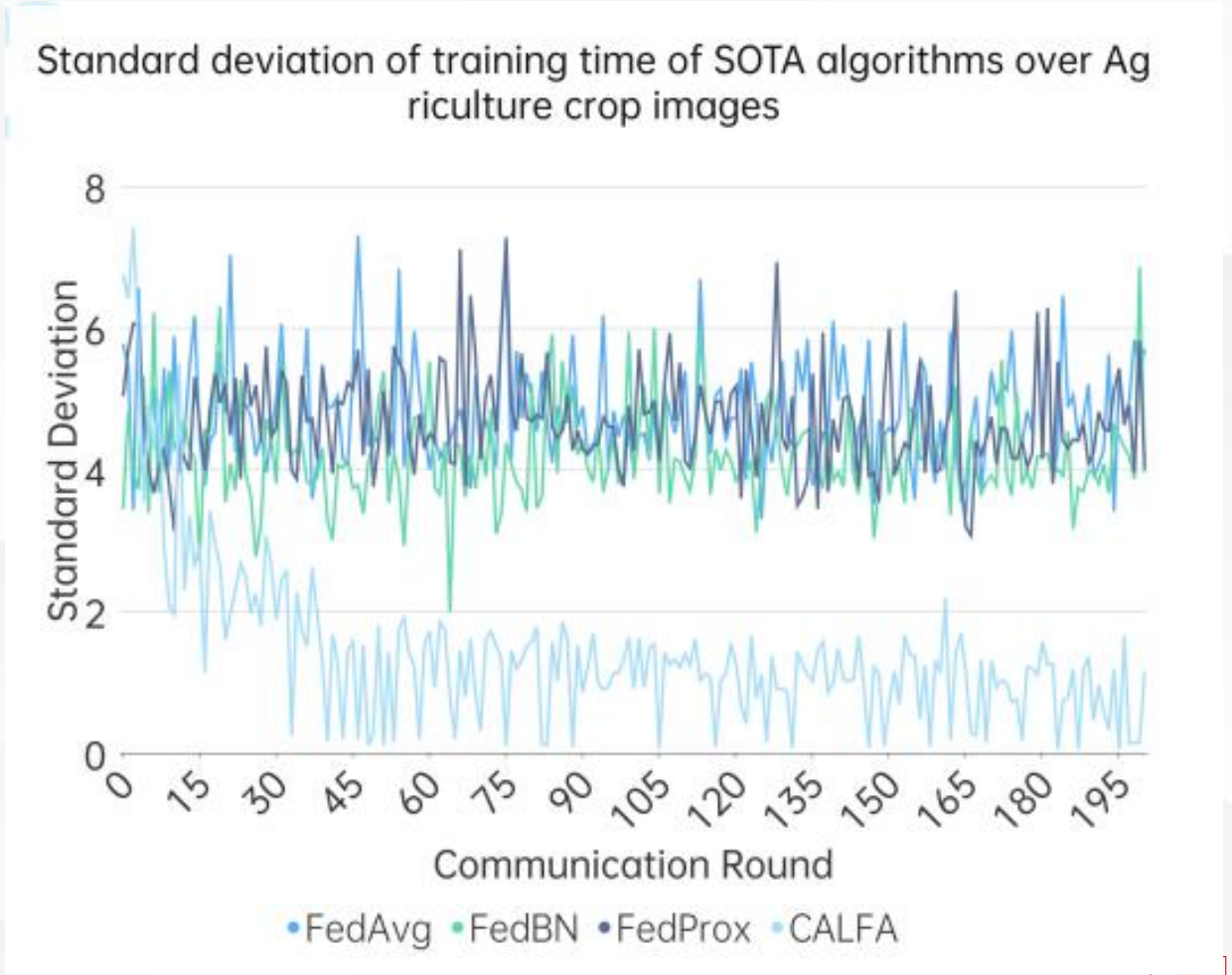
Algorithm	Accuracy	AUC	Conv.	Avg. Time/s	Std. Time/s	App.
Local	0.2913±0.0073	0.5807±0.0053	126±36	11.2562±4.8366	4.5228±1.4563	False
FedAvg[27]	**0.8027**±0.0219	0.9436±0.0104	101±52	12.5468±2.0883	5.0903±0.6760	False
FedProx[18]	**0.8027**±0.0201	**0.9477**±0.0057	121±41	12.1191±4.2455	4.5073±1.1736	False
FedBN[19]	0.7900±0.0180	0.9397±0.0088	**75**±20	11.9506±2.5389	5.0324±0.3404	False
FedPer[1]	0.3288±0.0000	0.6221±0.0192	36±27	12.3796±2.2522	4.8215±0.3423	False
FedRep[5]	0.3288±0.0000	0.6209±0.0189	28±5	15.6082±4.8483	6.0099±1.2850	False
APFL[6]	0.2320±0.0162	0.5086±0.0280	29±58	21.8305±10.0183	7.3776±3.4201	False
FedPHP[20]	0.2265±0.0275	0.5074±0.0333	9±16	11.7139±4.1097	4.8673±1.4517	False
FedMTL[34]	0.2941±0.0074	0.5594±0.0113	28±21	12.4020±2.9406	4.7836±0.5254	False
Ditto[17]	0.2904±0.0062	0.5552±0.0111	43±37	9.4509±3.3148	4.3855±0.6869	False
FedFomo[41]	0.3251±0.0018	0.6339±0.0015	179±31	8.7981±1.4074	4.0164±0.7786	False
FedProto[36]	0.3142±0.0124	0.6123±0.0034	164±106	11.5173±1.2865	4.6279±0.6379	False
CALFA w/o A_1	0.7945±0.0238	0.9472±0.0099	103±41	**3.4714**±1.2558	1.6892±0.0920	False
CALFA w/o A_2	0.8000±0.0137	0.9456±0.0068	94±42	7.5063±3.1591	3.2267±2.5267	True
CALFA	0.8018±0.0249	0.9416±0.0057	131±49	5.7313±3.4893	**0.9158**±0.8337	True

Hongjian Shi, Ilyas Bayanbayev, Ruhui Ma, Haibing Guan, Cloud-based Collaborative Agricultural Learning with Flexible Model Size and Adaptive Batch Number, ACM Transactions on Sensor Networks, 2023.

联邦学习效率 + 异构性

- 能够在保障准确率的情况下，减少训练用时和系统异构性

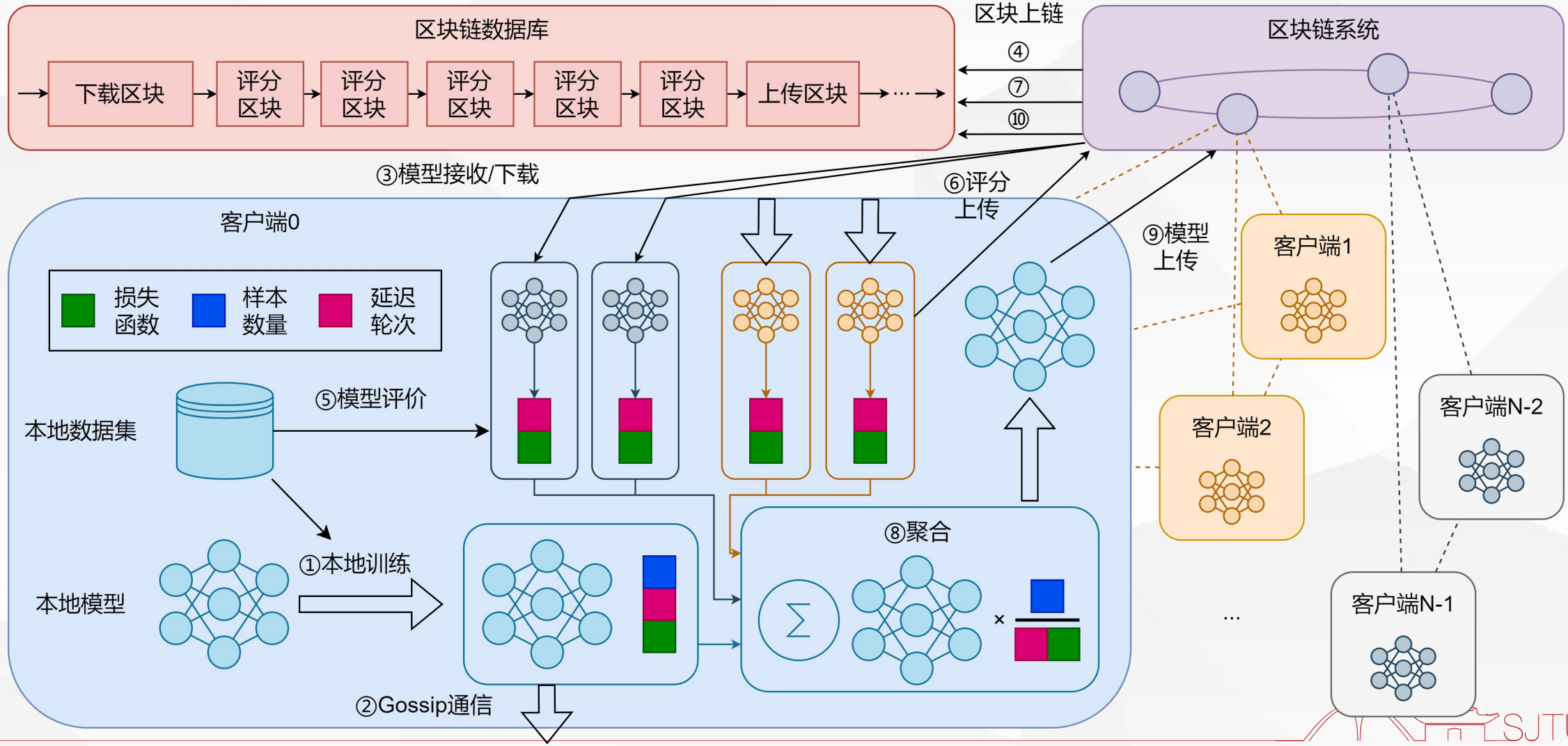




Hongjian Shi, Ilyas Bayanbayev, Ruhui Ma, Haibing Guan, Cloud-based Collaborative Agricultural Learning with Flexible Model Size and Adaptive Batch Number, ACM Transactions on Sensor Networks, 2023.

联邦学习效率 + 安全 + 异构性

半中心化：可信客户端-可信客户端（去中心） + 不可信客户端-区块链系统（中心）



施宏建，马汝辉，管海兵，基于区块链辅助的半中心化联邦学习框架，计算机研究与发展，2023

联邦学习效率 + 安全 + 异构性

基于损失函数的聚合权重调整方案

$$w_{1,j}^{t}=\frac{1}{L(\theta_{j}^{t_j}\mid D_i)},w_{1,k}^{t}=\frac{1}{L(\theta_{k}^{t_k}\mid D_i)}$$

基于延迟轮次的过时模型补偿方案

$$w_{2,j}^{t}=\begin{cases}e^{(t_j-t)}, t_j<t\\ 1, \quad t_j\geq t\end{cases},w_{2,k}^{t}=\begin{cases}e^{(t_k-t)}, t_k<t\\ 1, \quad t_k\geq t\end{cases}$$

原始权重

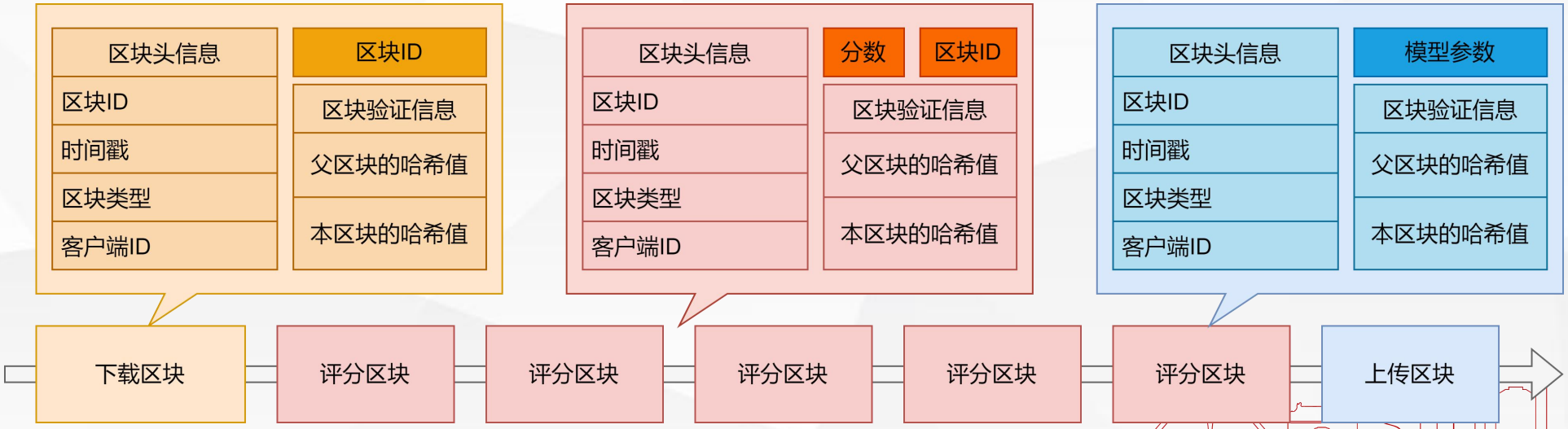
$$w_i^t=\frac{N_i}{\sum_{i=0}^{N-1}N_i}$$

最终的聚合权重

$$\theta_i^t=\frac{w_i^t w_{1,i}^t}{w^t}\tilde{\theta}_i^{t-1}+\sum_{C_j\in C_i^{neigh}}\frac{w_j^t w_{1,j}^t w_{2,j}^t}{w^t}\tilde{\theta}_j^{t_j}+\sum_{C_k\notin C_i^{neigh}}\frac{w_k^t w_{1,k}^t w_{2,k}^t}{w^t}\tilde{\theta}_k^{t_k}$$

区块设计

- 下载区块
- 评分区块
- 上传区块



联邦学习效率 + 安全 + 异构性

表 2 本文框架与其他联邦学习框架在 FMNIST-DIR 下的实验结果

算法框架	测试准确率	测试 AUC	平均训练时间(s)	设备计算时间比例(%)
FedAvg	0.7995 ± 0.0024	0.9666 ± 0.0245	1.3887 ± 0.0600	35.64 ± 3.77
FedProx	0.7993 ± 0.0022	0.9666 ± 0.0246	1.5989 ± 0.0752	36.27 ± 3.39
FedBN	0.7994 ± 0.0023	0.9666 ± 0.0245	1.6163 ± 0.0395	33.21 ± 1.74
FedPer	0.9735 ± 0.0001	0.9967 ± 0.0057	1.4584 ± 0.0864	26.43 ± 1.84
FedRep	0.9743 ± 0.0005	0.9974 ± 0.0061	2.2325 ± 0.1026	38.73 ± 4.17
FedBABU	0.7685 ± 0.0070	0.9950 ± 0.0090	1.4313 ± 0.0733	35.48 ± 3.38
APFL	0.9720 ± 0.0002	0.9976 ± 0.0077	3.6855 ± 0.1621	29.11 ± 1.21
FedPHP	0.0919 ± 0.0185	0.5072 ± 0.2313	3.9221 ± 0.3462	36.44 ± 1.55
Ditto	0.9714 ± 0.0005	0.9986 ± 0.0103	3.5450 ± 0.2870	30.81 ± 1.29
FedFomo	0.9719 ± 0.0004	0.9971 ± 0.0179	2.0139 ± 0.2788	38.14 ± 1.94
FedAMP	0.9720 ± 0.0006	0.9971 ± 0.0114	1.6826 ± 0.0546	29.34 ± 2.52
APPLE	0.9638 ± 0.0006	0.9907 ± 0.0206	30.1721 ± 13.1385	36.25 ± 4.49
FedAsync	0.8352 ± 0.0372	0.9482 ± 0.1870	1.4260 ± 0.0749	100.00 ± 0.00
Ours	0.8901 ± 0.0092	0.9760 ± 0.1303	1.8396 ± 0.0213	100.00 ± 0.00
Ours - delay	0.8749 ± 0.0070	0.9714 ± 0.1291	1.6460 ± 0.0517	100.00 ± 0.00
Ours - delay -loss	0.6595 ± 0.0772	0.8994 ± 0.2153	1.4766 ± 0.0417	100.00 ± 0.00

施宏建，马汝辉，管海兵，基于区块链辅助的半中心化联邦学习框架，计算机研究与发展，2023

Q&A

饮 水 思 源　爱 国 荣 校