

# federated learning

## Personalized Federated Learning with Parameter Propagation 参数传播的个性化联邦学习

- 利用从不同客户端收集的分散数据，提出了一种个性化的联邦学习范式，用于训练机器学习模型，而无需交换来自本地客户端的原始数据。本文从隐私保护迁移学习的角度深入探讨了个性化联邦学习，并指出了以往个性化联邦学习算法的局限性。首先，以往的工作在关注所有客户的整体绩效时，对一些客户的知识可转移性为负。其次，明确学习客户端之间的统计任务相关性需要较高的沟通成本。第三，将从有经验的客户那里学到的知识推广到新客户那里，计算成本很高。为了解决这些问题，本文提出了一种新的用于个性化联邦学习的联邦参数传播(FEDORA)框架。具体来说，我们将标准的个性化联邦学习重新表述为一个保护隐私的迁移学习问题，目标是提高每个客户端的泛化性能。FEDORA背后的关键思想是学习如何传输以及是否同时传输，包括(1)自适应参数传播:强制一个客户端根据其任务相关性(例如，通过分布相似性显式度量)自适应地将其参数传播给其他客户端，以及(2)选择性正则化:只有当这些参数与其局部模型的泛化性能呈正相关时，每个客户端才会使用接收到的参数对其局部个性化模型进行正则化。在各种联邦学习基准上的实验证明了所提出的FEDORA框架在最先进的个性化联邦学习基线上的有效性。
- <https://dl.acm.org/doi/10.1145/3580305.3599464>
- August 2023

**Proposed Algorithm: FEDORA**

Jun Wu

□ **Federated Parameter Propagation (FEDORA)**


$$\min_{\theta_k, \hat{\theta}_k} \underbrace{\sum_{k=1}^K \frac{1}{\lambda_k n_k} \sum_{i=1}^{n_k} \ell(x_i^k, y_i^k; \theta_k)}_{\text{Local Training}} + \underbrace{\sum_{k=1}^K \|\theta_k - \hat{\theta}_k\|_2^2}_{\text{Consistency}} + \underbrace{\frac{\alpha}{2} \sum_{k=1}^K \sum_{k'=1}^K \frac{w_{kk'}}{D_{kk}} \|\hat{\theta}_k - \hat{\theta}_{k'}\|_2^2}_{\text{Smoothness}} \quad \Leftarrow \text{Objective}$$

**Client update:**  $\min_{\theta_k} \frac{1}{n_k} \sum_{i=1}^{n_k} \ell(x_i^k, y_i^k; \theta_k) + \lambda_k \|\theta_k - \hat{\theta}_k\|_2^2$  (Fix  $\hat{\theta}_k$ , update  $\theta_k$ )

**Server update:**  $\min_{\hat{\theta}_k} \sum_{k=1}^K \|\theta_k - \hat{\theta}_k\|_2^2 + \frac{\alpha}{2} \sum_{k=1}^K \sum_{k'=1}^K \frac{w_{kk'}}{D_{kk}} \|\hat{\theta}_k - \hat{\theta}_{k'}\|_2^2$  (Fix  $\theta_k$ , update  $\hat{\theta}_k$ )

$\Leftarrow \text{Solution}$

- 5 -



## Negative Transfer Mitigation



### Accuracy

$$ACC(\theta_k^*) = \frac{1}{n_{test}} \sum_{i=1}^{n_{test}} [y_i = y_i^{pred}]$$

### Relative Accuracy

$$R-ACC(\theta_k^*) = \frac{ACC(\theta_k^*) - ACC(\theta_k^{LOCAL})}{ACC(\theta_k^{LOCAL})}$$

### Positive Transferability Ratio

$$PTR = \frac{1}{K} \sum_{k=1}^K \mathbb{I}[ACC(\theta_k^*) - ACC(\theta_k^{LOCAL})]$$

Model	Rotated MNIST			Rotated Fashion-MNIST			CIFAR-10		
	Acc ↑	R-Acc ↑	PTR ↑	Acc ↑	R-Acc ↑	PTR ↑	Acc ↑	R-Acc ↑	PTR ↑
LOCAL	0.7642	-	-	0.7057	-	-	0.7617	-	-
FedAvg [25]	0.6889	-0.0976	0	0.6441	-0.0847	0.1250	0.6531	-0.1382	0.3000
FedAvg+FT	0.7411	-0.0293	0.3056	0.6848	-0.0283	0.3472	0.7992	0.0513	0.9000
FedProx [21]	0.5375	-0.2962	0	0.5968	-0.1521	0	0.6984	-0.0799	0.2000
FedProx+FT	0.6893	-0.0973	0.0278	0.6788	-0.0358	0.3056	0.7953	0.0460	0.9000
LG-FedAvg [23]	0.7804	0.0214	0.9444	0.7137	0.0115	0.7361	0.7656	0.0054	0.8000
FedPer [1]	0.7741	0.0135	0.6389	0.6725	-0.0457	0.1389	0.8352	0.0990	<b>1.0000</b>
pFedHN [33]	0.8004	0.0486	0.8611	0.7215	0.0249	0.6944	0.7766	0.0221	0.6000
APFL [6]	0.7871	0.0303	0.8889	0.7134	0.0112	0.7639	0.8258	0.0866	0.9000
Ditto [20]	0.7806	0.0220	0.7222	0.7212	0.0232	0.7361	0.8078	0.0630	0.9000
IFCA [9]	0.7915	0.0365	0.6944	0.7305	0.0370	0.7639	0.8227	0.0828	0.9000
FeSEM [46]	0.7720	0.0110	0.6111	0.7074	0.0051	0.5278	0.8547	0.1255	<b>1.0000</b>
FedFOMO [47]	0.7749	0.0140	0.9167	0.7110	0.0076	0.7639	0.8242	0.0797	<b>1.0000</b>
FedU [7]	0.7837	0.0260	0.8889	0.7208	0.0225	0.8056	0.7836	0.0295	0.9000
FedAMP [13]	0.7869	0.0298	<b>1.0000</b>	0.7203	0.0213	0.8056	0.7953	0.0457	0.8000
<b>FEDORA</b>	<b>0.8251</b>	<b>0.0806</b>	<b>1.0000</b>	<b>0.7433</b>	<b>0.0548</b>	<b>0.9028</b>	<b>0.8570</b>	<b>0.1288</b>	<b>1.0000</b>

## Model elasticity for hardware heterogeneity in federated learning systems 联邦学习系统中硬件异构的模型弹性

- 迄今为止提出的大多数联邦学习(FL)算法通过聚合通常共享相同架构的多个本地模型来获得全局模型，从而忽略了对边缘设备硬件异构性的影响。为了解决这个问题，我们提出了一个基于模型弹性新概念的模型-体系结构协同设计框架。更准确地说，我们使本地设备能够训练属于同一架构家族的不同模型，选择以匹配各种边缘设备的资源预算(例如，延迟，内存，功率)。我们在EMNIST和CIFAR-10上对IID和非IID情况的研究结果显示，每轮通信传输的数据减少了2.44倍，通信轮数减少了100倍，同时提供与现有方法相同或更好的准确性。
- <https://dl.acm.org/doi/10.1145/3556557.3557954>
- October 2022

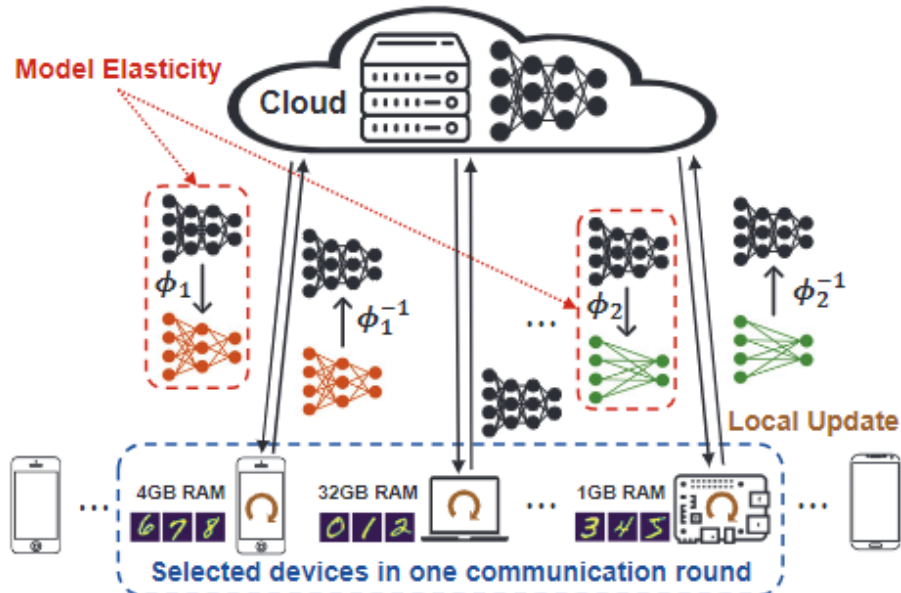


Figure 1: We propose *model elasticity* to allow the cloud to distribute elastic models that satisfy the heterogeneous hardware constraints of edge devices.

## Federated Learning with Label-Masking Distillation 基于标签掩蔽蒸馏的联邦学习

- 联邦学习提供了一种保护隐私的方式，通过全局服务器的协调，在分布在多个本地客户机上的数据上协作训练模型。在本文中，我们重点研究了联邦学习中的标签分布偏差，其中由于客户端用户行为的不同，不同客户端之间的标签分布存在显著差异。面对这种情况，现有的大多数方法由于客户端对标签分布信息的利用不足，会导致次优优化。受此启发，我们提出了一种称为FedLMD的标签屏蔽蒸馏方法，通过感知每个客户端的各种标签分布来促进联邦学习。我们根据训练过程中每个类的样本数量将标签分为多数标签和少数标签。客户端模型从本地数据中学习多数标签的知识。蒸馏过程掩盖了来自全局模型的多数标签的预测，因此它可以更专注于保留客户端的少数标签知识。一系列实验表明，该方法在各种情况下都能达到最先进的性能。此外，考虑到客户端有限的资源，我们提出了一种不需要额外教师的FedLMD-Tf变体，它在不增加计算成本的情况下优于以前的轻量级方法。我们的代码可在<https://github.com/wnma3mz/FedLMD>上获得。
- <https://dl.acm.org/doi/10.1145/3581783.3611984>
- 27 October 2023

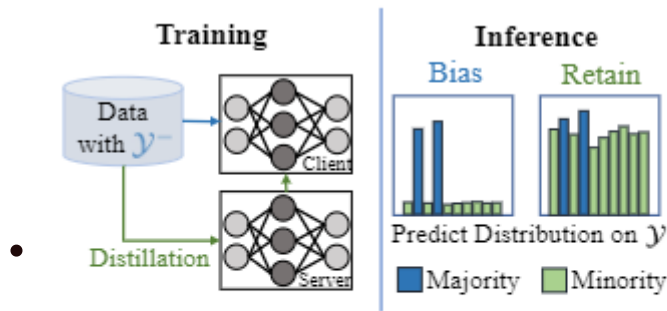


Figure 1: The model trained on the private dataset of a client with partial class labels  $\mathcal{Y}^-$  is generally biased to  $\mathcal{Y}^-$  due to knowledge missing over complete class labels  $\mathcal{Y}$ . Our FedLMD method proposes to alleviate it by utilizing the global model from the server to retain the knowledge of minority labels  $\mathcal{Y} \setminus \mathcal{Y}^-$ .

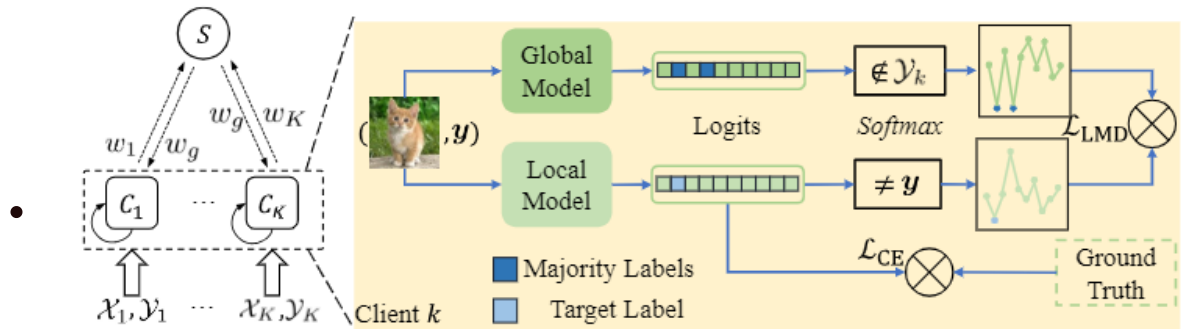


Figure 3: The framework of our approach. For the aggregation process, for the uploaded weight  $w_1, \dots, w_K$  of the model are calculated as weighted averages to obtain  $w_g$ . For each client, the training loss is the combination of the cross-entropy loss  $\mathcal{L}_{CE}$  for learning from local data and the label-masking distillation loss  $\mathcal{L}_{LMD}$  for distilling from the global model.

## DFL: High-Performance Blockchain-Based Federated Learning

### DFL:基于区块链的高性能联邦学习

- 任何研究人员都建议用区块链系统取代联邦学习中的聚合服务器，以提高隐私性、鲁棒性和可扩展性。在这种方法中，客户将他们更新的模型上传到区块链分类账，并使用智能合约执行模型平均。然而，区块链系统的显著延迟和有限的计算能力使得支持区块链上的机器学习应用程序效率低下。在本文中，我们提出了一种名为DFL的新的公共区块链架构，该架构专门针对分布式联邦机器学习进行了优化。我们的架构继承了传统区块链系统的优点，同时通过放弃全球共识实现低延迟和低资源消耗。为了评估我们架构的性能和鲁棒性，我们实现了一个原型并在物理四节点网络上进行了测试，并开发了一个模拟器来模拟更大的网络和更复杂的情况。我们的实验表明，DFL架构对非i.i.d的准确率可以达到90%以上。数据集，即使存在模型中毒攻击，同时确保区块链部分消耗的硬件资源少于5%
- <https://dl.acm.org/doi/epdf/10.1145/3600225>
- 18 September 2023

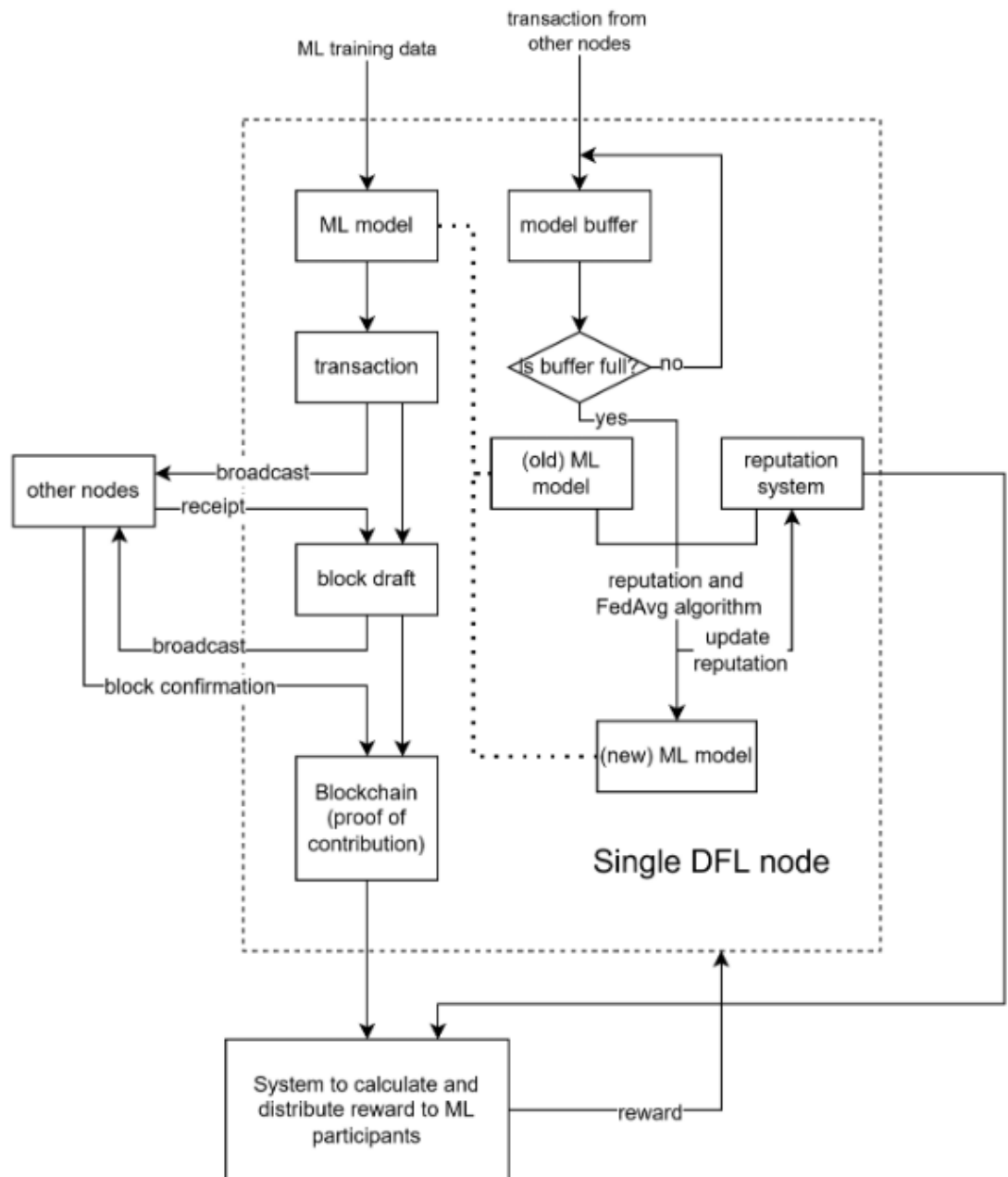


Fig. 1. DFL overview.

## Inferring Class-Label Distribution in Federated Learning 联邦学习中类标签分布的推断

- 联邦学习(FL)已经成为一种流行的分布式学习方法，通过使用个人客户端的私有数据来训练分类器。客户的数据通常被认为是保密的，但它们的异质性和潜在的类不平衡会对训练模型的准确性产生不利影响。阶级不平衡可能不是常识，甚至可能本身就是机密信息。因此，从性能和隐私的角度来看，训练数据的类标签分布的推断都很重要。在本文中，我们从对抗的角度研究了基于发送到参数服务器的模型参数更新的类标签分布推断问题。首先，我们给出了可能进行精确推理的条件。然后，我们介绍了四种新的方法来估计一般FL设置下的类标签分布。我们在四个不同的数据集上评估了提议的差分方法，我们的结果表明它们明显优于最先进的方法。
- <https://dl.acm.org/doi/epdf/10.1145/3560830.3563725>



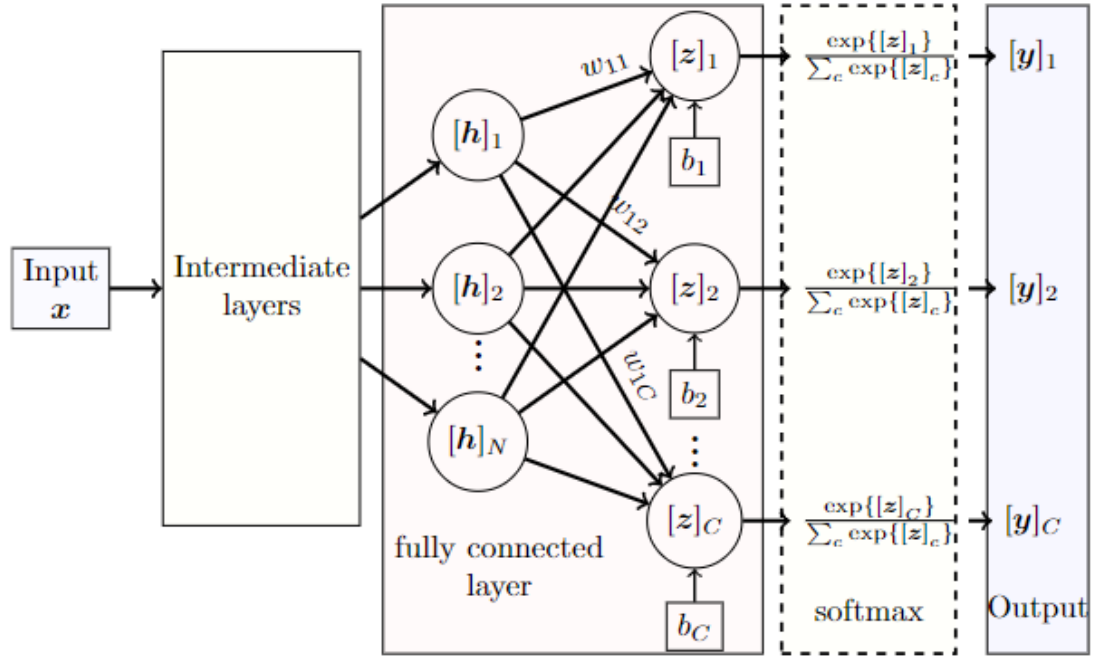


Figure 1: A generic neural network classifier with fully-connected output layer and softmax layer before the classification output. The weights and biases from the fully-connected output layer are used to estimate the class-label distribution in this paper.

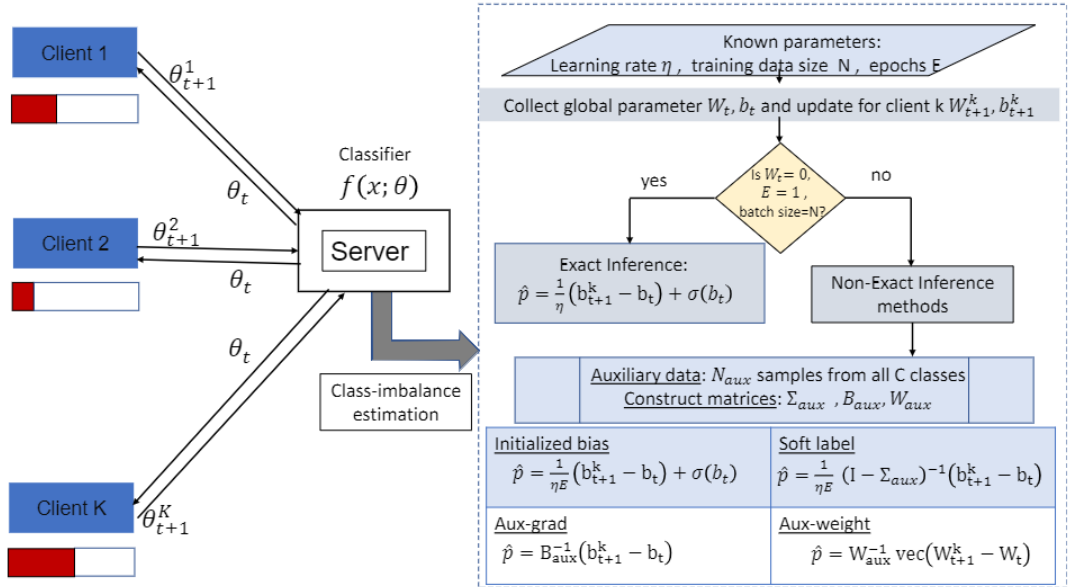


Figure 2: Block diagram of the proposed class-label distribution inference attacks.

# FedDefender: Client-Side Attack-Tolerant Federated Learning (FedDefender:客户端容忍攻击的联邦学习)

- 联邦学习可以在不损害隐私的情况下从分散的数据源中学习，这使得它成为一项至关重要的技术。然而，它很容易受到模型中毒攻击，其中恶意客户端会干扰训练过程。以前的防御机制通过谨慎地使用模型聚合来关注服务器端，但是当数据分布不相同或攻击者可以访问良性客户端的信息时，这可能不有效。在本文中，我们提出了一种专注于客户端的新的防御机制，称为feddefender，以帮助良性客户端训练健壮的本地模型，并避免来自攻击者的恶意模型更新的不利影响，即使在服务器端防御无法识别或移除对手时也是如此。该方法由两个主要部分组成:(1)容错局部元更新和(2)容错全局知识蒸馏。这些组件用于寻找抗噪声模型参数，同时从可能损坏的全局模型中准确地提取知识。我们的客户端防御策略具有灵活的结构，可以与任何现有的服务器端策略结合使用。跨多个数据集的真实场景评估表明，所提出的方法增强了联邦学习对模型中毒攻击的鲁棒性
- <https://dl.acm.org/doi/epdf/10.1145/3580305.3599346>
- 04 August 2023

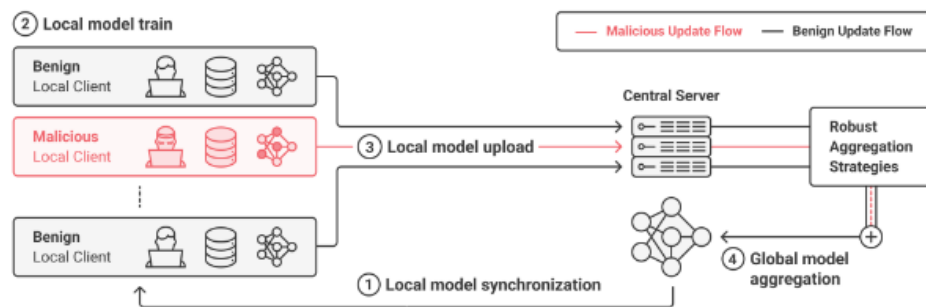
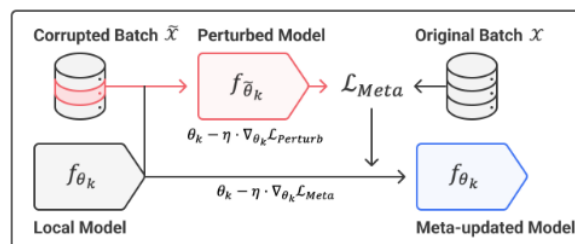


Figure 1: Illustration of federated learning under poisoning attacks. Attackers attempt to send malicious updates to the central server to corrupt the aggregated model. Unlike existing works that primarily focus on robust aggregation on the server side (stage ④), FedDefender focuses on training robust local models by benign clients (stage ②) to protect against malicious updates from adversaries.

Step.1 Attack-Tolerant Local Meta Update (Sec. 4.1)



Step.2 Attack-Tolerant Global Knowledge Distillation (Sec. 4.2)

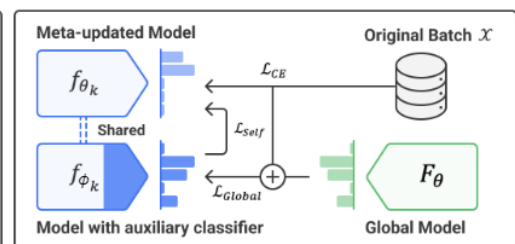


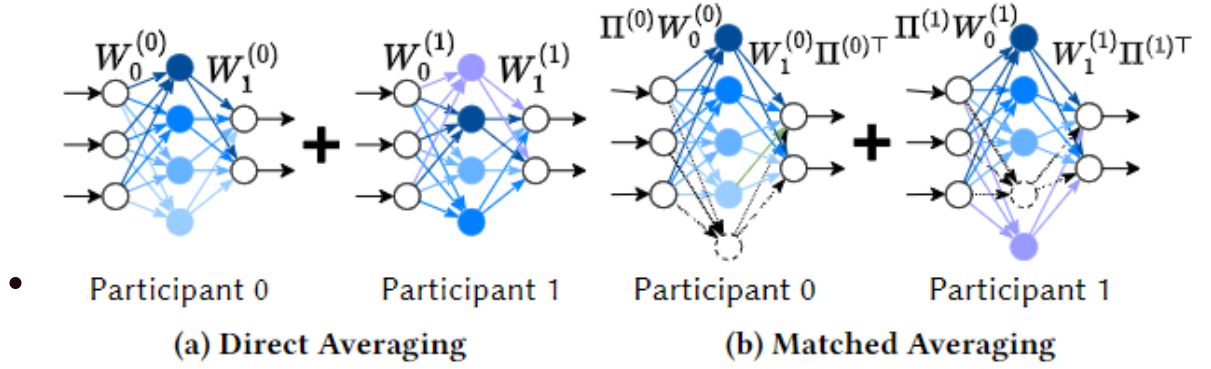
Figure 2: The overall design of FedDefender framework. FedDefender comprises two steps: (1) attack-tolerant local meta-update, which finds local model parameters that are less prone to overfitting to noise, and (2) attack-tolerant global knowledge distillation, which aims to convey correct global knowledge from a potentially contaminated global model, regularizing the local model to mitigate data bias.

## Communication-Efficient Generalized Neuron Matching for Federated Learning 面向联邦学习的通信高效广义神经元匹配

- 联邦学习(FL)是一种流行的分布式机器学习范式，它允许多个参与者在共享原始数据的情况下协作训练模型。在每一轮FL中，所有参与者并行地训练本地模型，并且服务器将本地模型聚合以创建全局模型。神经元匹配是一种很有前途的聚合方法，它利用排列不变性来提高全局模型的质量并降低通信成本。然而，现有的神经

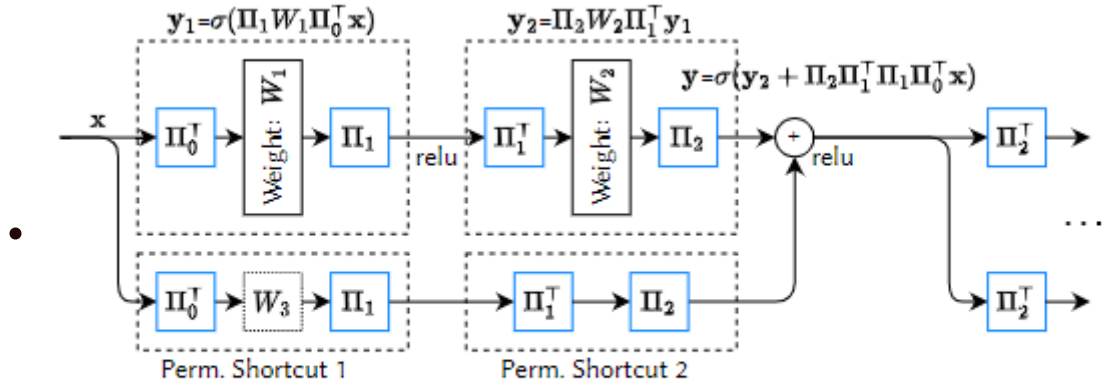
匹配方法有两个严格的限制:1)它们只适用于具有简单顺序网络的模型，而不能支持像resnet这样的高级模型。2)他们的方法显著增加了匹配后的模型尺寸，随着训练的进行，这导致了大量的内存和通信成本。在本文中，我们提出了一种新的神经元匹配算法，称为联邦广义匹配平均(FedGMA)来缓解这些限制。实验表明，该方法适用于复杂网络结构，如ResNet和InceptionNet，并且达到了与直接平均整个模型的算法相当的精度。它还降低了通信成本，并在数据异构场景中展示了鲁棒性。

- <https://dl.acm.org/doi/epdf/10.1145/3605573.3605726>
- 13 September 2023



**Figure 1: Neuron Matching Concept Illustration**

( $W_i^{(j)}$  represents the  $i$ -th layer of the  $j$ -th participant,  $\Pi^{(j)}$  are permutation matrices.)



**Figure 6: A Residual Block after Matching**



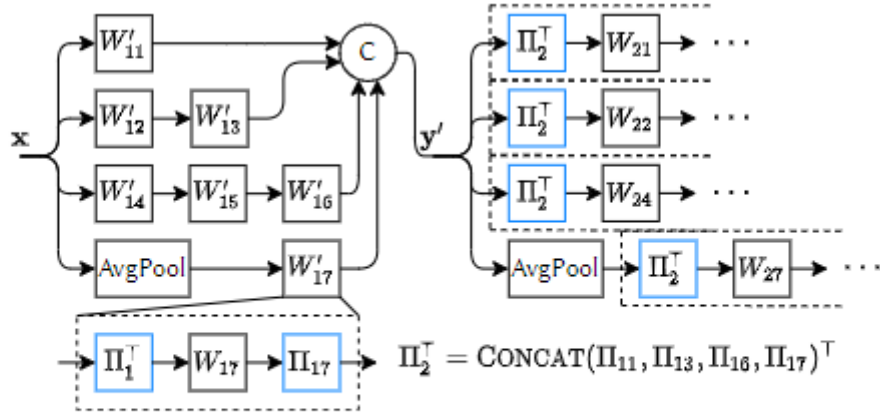


Figure 7: Two Inception Blocks after Matching

## ID-Based Multireceiver Homomorphic Proxy Re-Encryption in Federated Learning 联邦学习中基于id的多接收者同态代理重加密

- 随着机器学习的进步，数据隐私问题日益受到关注。联邦学习(FL)是谷歌在2016年发明的一种机器学习。在FL中，主要目的是通过聚合参与者上传的局部模型来训练一个高精度的全局模型，并且过程中的所有数据都保存在本地。但是，云服务器中的安全性或参与者之间的安全性折衷使得此过程不够安全。为了解决这个问题，本文提出了一种基于身份的多接收者同态代理重加密(IMHPRE)方案，该方案利用同态操作和重加密来提供改进的加密数据处理和访问控制。采用该方案时，参与者可以直接使用公开身份进行加密。IMHPRE方案对于选择明文攻击也是安全的。对比结果表明，由于IMHPRE允许云服务器对多个接收方的重新加密模型进行模型聚合，因此其性能优于同类算法。
- <https://dl.acm.org/doi/epdf/10.1145/3540199>
- 29 November 2022

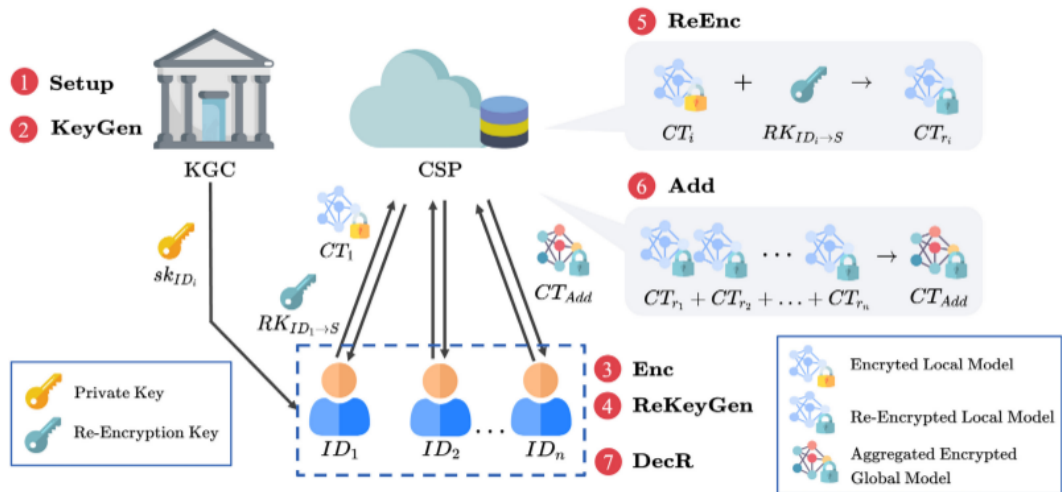


Fig. 3. Algorithms applied in federated learning.

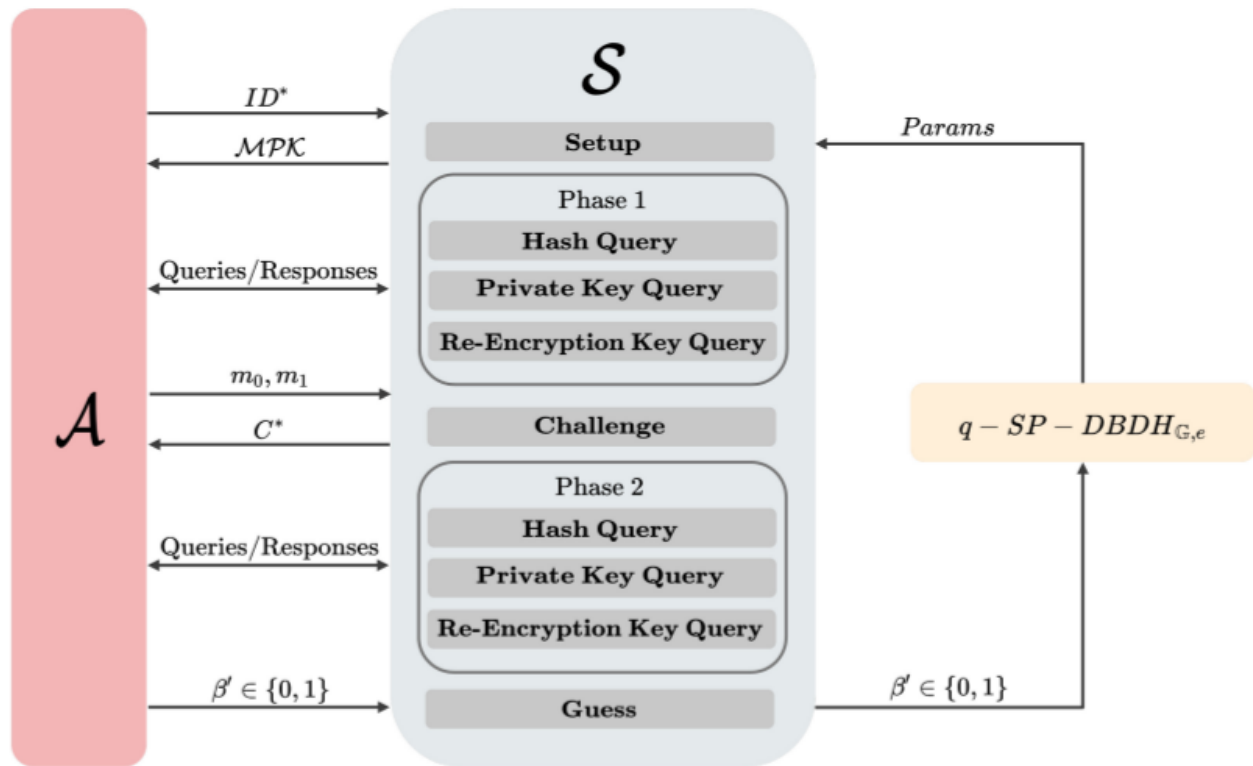
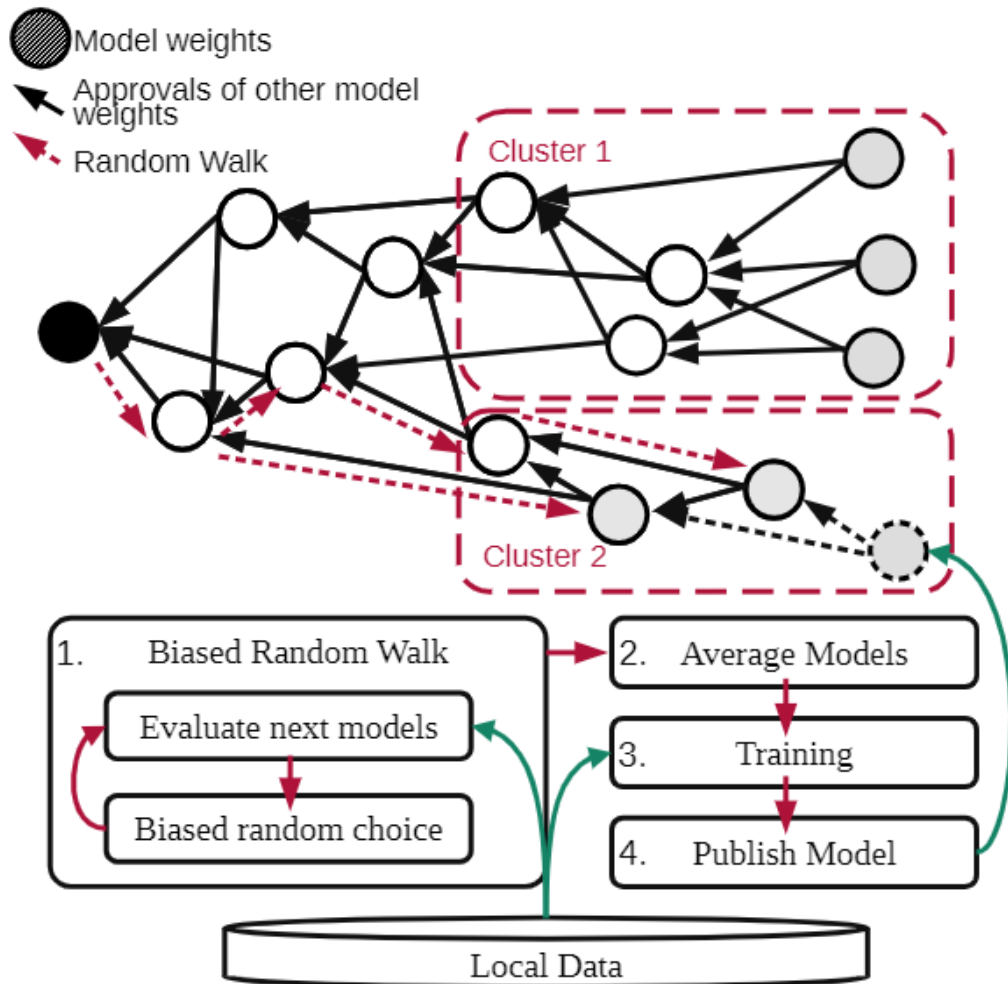


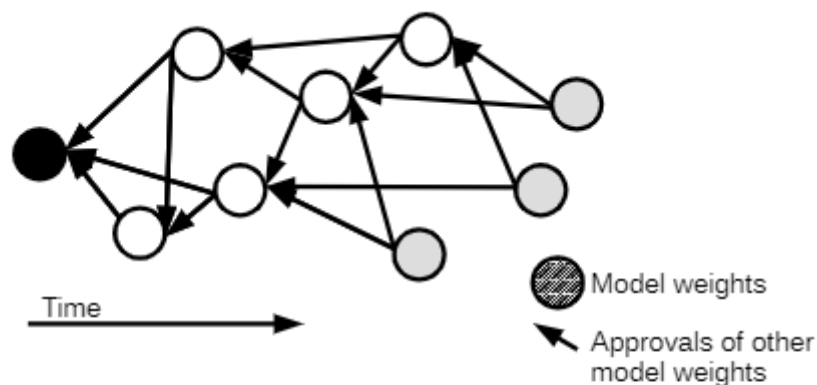
Fig. 4. The IND-sID-CPA game.

## Implicit model specialization through dag-based decentralized federated learning通过基于dag的分散联邦学习实现隐式模型专门化

- 联邦学习允许一组分布式客户端在私有数据上训练通用机器学习模型。模型更新的交换由中央实体或以分散的方式管理，例如通过区块链。然而，所有客户机之间的强泛化使得这些方法不适合非独立和同分布(non-IID)数据。我们提出了一种基于模型更新的有向无环图(DAG)的联邦学习中去中心化和个性化的统一方法。客户不再训练单一的全局模型，而是专注于本地数据，同时根据各自数据的相似性使用来自其他客户的模型更新。这种专门化隐式地出现在基于dag的通信和模型更新的选择中。因此，我们支持专门模型的发展，这些模型专注于数据的一个子集，因此比集中式或基于区块链的设置中的联邦学习更好地覆盖非iid数据。据我们所知，所提出的解决方案是第一个在完全分散的联邦学习中统一个性化和中毒鲁棒性的解决方案。我们的评估表明，模型的专门化直接来自基于dag的三个不同数据集的模型更新通信。此外，与联邦平均相比，我们展示了稳定的模型准确性和更小的客户端方差。
- <https://dl.acm.org/doi/10.1145/3464298.3493403>
- December 2021



**Figure 1:** We use a biased random walk through a DAG of model updates to find models that perform well on local data, resulting in clusters emerging in the DAG.



**Figure 2:** Communicating model updates in federated learning through a DAG: The nodes in the graph are model weight updates and the edges connect a weight update to the two other weight updates that were used as a basis for its training. Tips of the DAG (gray) are updates that didn't receive any approvals yet.

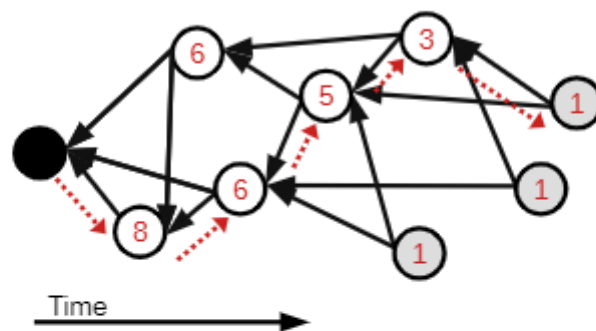


Figure 3: In DAG-based consensus schemes, traditionally weights of transactions are calculated by counting the number of approving transactions (also considering all transactions as self-approving). Thus, the weights are a global property of the DAG itself. The dashed red arrows show a random walk that always chooses the highest weight.

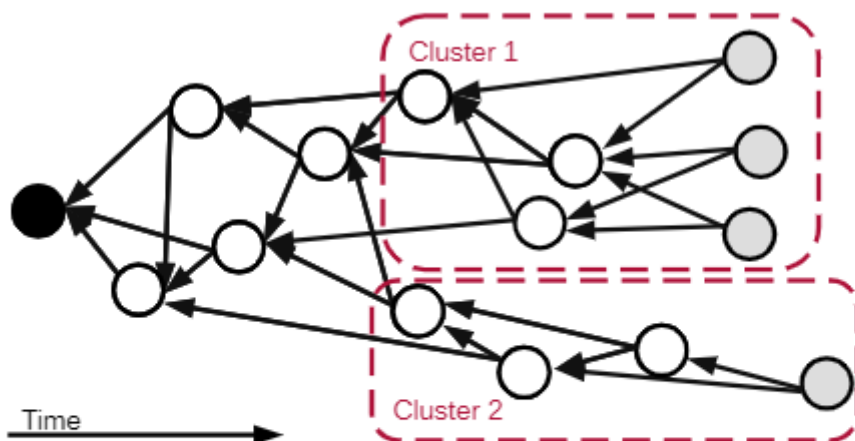


Figure 4: Tip selection using random walks biased by the model performance on local data leads to specialization and clustering in the directed acyclic graph.

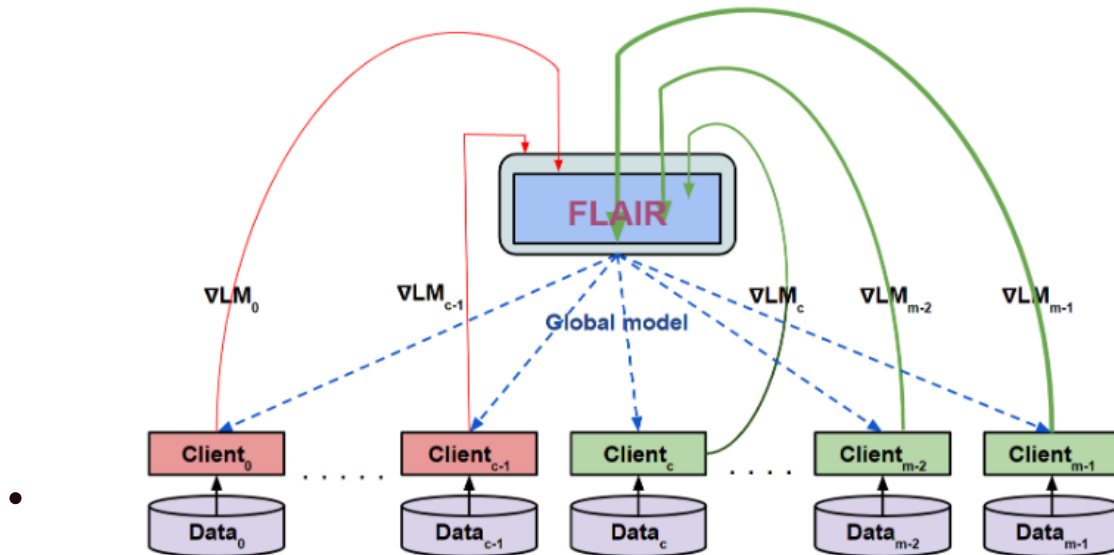
## FLAIR: Defense against Model Poisoning Attack in Federated Learning

FLAIR:联邦学习中的模型中毒攻击防御

联邦学习——分散环境中的多方分布式学习——比集中式学习更容易受到模型中毒攻击。这是因为恶意客户端可以串通并发送精心定制模型更新，以使全局模型不准确。这推动了拜占庭弹性联邦学习算法的发展，如Krum、Bulyan、FABA和FoolsGold。然而，最近开发的一种非目标模型中毒攻击表明，所有先前的防御都可以绕过。这种攻击利用的直觉是，对于一组恶意客户端，只需改变优化器正在计算的梯度更新的符号，就可以从优化器中转移模型，从而增加测试错误率。在这项工作中，我们开发了flair -一种针对这种定向偏差攻击

(DDA)的防御，这是一种最先进的模型中毒攻击。FLAIR基于我们的直觉，即在联邦学习中，梯度翻转的某些模式表示攻击。这种直觉在不同的学习算法、模型和数据集上都非常稳定。FLAIR根据客户在培训阶段的行为为参与的客户分配声誉分数，然后对客户的贡献进行加权。我们表明，当20-30%的客户端是恶意客户端时，现有的FABA [IJCAI ' 19]，FoolsGold [Usenix ' 20]和FLTrust [NDSS ' 21]的防御基线失效，FLAIR提供了高达45%的恶意客户端百分比的拜占庭鲁棒性。我们还表明，FLAIR甚至可以针对DDA的白盒版本提供健壮性。

- <https://dl.acm.org/doi/epdf/10.1145/3579856.3582836>
- 10 July 2023



**Figure 1: FLAIR’s architecture where  $c$  out of  $m$  clients are malicious and send carefully crafted values of their local models to throw the global model off convergence. FLAIR weighs the gradients, received from the clients, by their reputation scores before aggregation represented by varying thicknesses of arrows from the clients.**