

成绩\_\_\_\_\_

苏州大学

课 程      文献阅读和科技写作

学院(部)    计算机科学与技术学院

类 型      专题综述

学 号      1927405160

姓 名      张昊

日 期      2022 年 6 月 19 日

# 边缘计算的联邦学习算法

1927405160 张昊

(苏州大学计算机科学与技术学院, 江苏 苏州 215006)

**摘 要** 联邦学习是一种分布式机器学习方法,可以在保证数据隐私的前提下利用边缘设备的计算资源训练机器学习模型.但是边缘计算场景下的资源往往受限,数据往往不平衡,在边缘设备上应用联邦学习来分布式地训练机器学习模型带来了新的挑战.为此,我们探讨了边缘计算中的联邦学习的现有研究.我们首先从联邦学习的定义出发,指出现有的问题和挑战,并对各类联邦学习算法的优缺点进行了总结,最后探讨了边缘计算中联邦学习算法的进一步研究方向.

**关键词** 联邦学习; 边缘计算; 设备端训练; 资源受限设备

## Federated learning algorithms in edge computing

Hao Zhang

(School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu, 215006, China)

**Abstract:** Federated learning (FL) is a distributed machine learning strategy that trains a machine learning model by using the computational resources of edge devices while ensuring data privacy. However, there are often limited resources and unbalanced data in edge computing. Training a machine learning model in a distributed way using FL on edge devices brings new challenges. To this end, we explore the existing studies on FL in edge computing. We first give the definition of FL, and point out the existing problems and challenges. We then summarize the advantages and disadvantages of various FL algorithms. Finally, we discuss the future research directions of FL algorithms in edge computing.

**Key Words:** Federated learning; edge computing; on-device training; resource-constrained devices

### 1 引言

联邦学习 (Federated Learning, FL) 是一种分布式机器学习方法,广泛应用于边缘计算场景中,利用众多离散的边缘设备的计算能力训练机器学习模型.在传统的机器学习方法中,客户端需要将数据共享给云端的数据中心,这将面临着用户隐私暴露等一系列问题,尤其是

用户需要共享敏感数据的时候。

联邦学习就是这样一种可以在保证数据隐私安全及合法合规的基础上，充分利用边缘设备上本地数据进行模型训练的方法，从而用户无需与其他实体分享数据。2016 年，谷歌公司提出了联邦学习这一概念<sup>[1]</sup>，用于解决安卓手机终端用户在本地上更新模型的问题。近年来，得益于物联网技术的迅速发展和联邦学习带来的隐私保护等优点，联邦学习技术在边缘计算领域获得了广泛应用。然而，边缘设备的计算和通信等资源是受限的，且不同设备上的数据分布也不尽相同。因此，在这些边缘设备上应用联邦学习来分布式地训练机器学习模型带来了新的挑战，如训练难以在预期时间内达到收敛、消耗较高的能耗等等。而要在边缘计算场景中实现高效的联邦学习，就需要对其算法流程进行一系列的优化和改进。

本文首先为联邦学习进行了定义，分析了其应用于边缘计算中的挑战，并总结了各类联邦学习算法的优点和缺点。最后，探讨了边缘计算中联邦学习算法的进一步研究方向。

## 2 研究现状

### 2.1 联邦学习定义

联邦学习可以定义为一种分布式机器学习方法，由多个客户端（也称为计算节点，worker）和一个或多个服务器（server）共同参与相同的模型训练任务。其中服务器维护全局共享的模型，客户端不会向服务器共享参与训练的数据，而是在本地训练模型，服务器通过定期聚合各客户端的模型信息来更新全局共享模型。

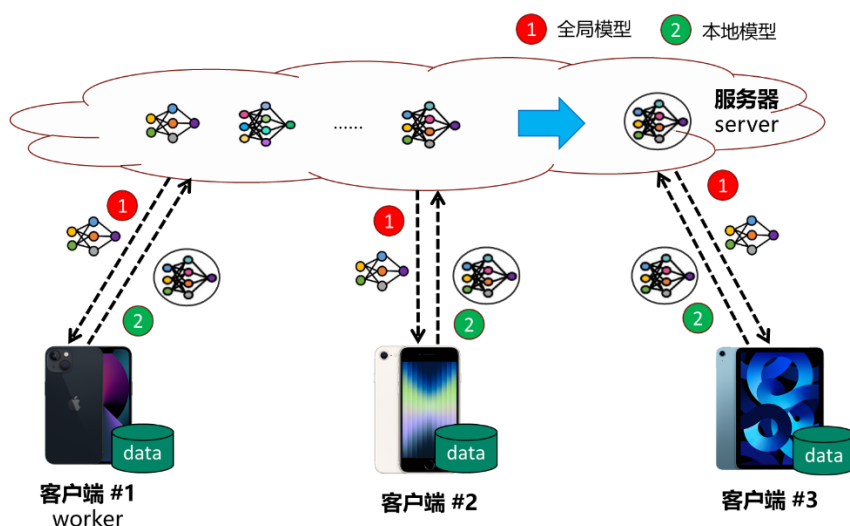


图 1 遵循客户端—服务器架构的联邦学习过程示意图。

如图 1 所示，联邦学习的过程通常包括三个步骤<sup>[2]</sup>：

- (1) 在初始阶段, 服务器设定学习任务, 初始化全局模型且向客户端广播该全局模型;
- (2) 每个参与者收到全局模型后, 利用其本地数据进行模型训练, 新的模型与服务器共享;
- (3) 服务器从参与者处接收到本地模型后执行聚合并生成更新的全局模型, 再次与参与者共享;

重复执行步骤 (2) ~ (3), 直至全局模型收敛.

一般地, 假设所有参与联邦学习的客户端总量为  $K$ , 服务器每次选择比例为  $C$  的客户端进行训练. 每个客户端  $k \in [1, K]$  保存有参与训练的数据集  $D_k$ , 记  $n_k = |D_k|$  为数据集  $D_k$  中数据样本的数量. 对于机器学习问题, 函数  $l(\mathbf{w}; x_i, y_i)$  表示数据样本  $(x_i, y_i)$  的损失, 其中  $\mathbf{w}$  表示神经网络的模型参数. 相应地, 客户端  $k$  在本地数据集  $D_k$  上的损失函数  $f_k(\mathbf{w})$  可表示为各数据样本的损失之和

$$f_k(\mathbf{w}) = \frac{1}{n_k} \sum_{i \in D_k} l(\mathbf{w}; x_i, y_i). \quad (1)$$

记  $n = \sum_{k=1}^K n_k$  表示参与联邦学习的数据样本总量, 从而我们可以简单地定义客户端  $k$  对联邦学习的影响为  $p_k = \frac{n_k}{n}$ , 则联邦学习的全局损失函数  $F(\mathbf{w})$  相应表示为

$$F(\mathbf{w}) = \sum_{k=1}^K p_k f_k(\mathbf{w}). \quad (2)$$

联邦学习算法的优化目标为最小化全局损失函数  $F(\mathbf{w})$ .

## 2.2 在边缘设备上进行联邦学习

在边缘计算中应用联邦学习不可避免地需要边缘设备的参与. 边缘设备可以概括性地定义为计算能力、通信能力、存储容量和功耗等资源受限的实体<sup>[3]</sup>, 这类实体拥有不会与其他实体共享的本地数据. 相较于传统的机器学习方法, 在边缘设备上进行联邦学习面临着更多的挑战:

### (1) 非独立同分布的数据

联邦学习中, 各客户端上参与训练的数据通常是与用户偏好以及使用情况密切相关的, 因此任一个客户端上的本地数据往往具有不同的规模、特征和类别分布, 不能代表所有参与联邦学习训练的数据集的总体统计分布, 从而导致了非独立同分布 (Non Independent and Identically Distributed, 非 IID) 数据的出现<sup>[1]</sup>. 联邦学习中可能会遇到类别不平衡、分布不平衡和数量不平衡三种非 IID 数据的情况<sup>[4]</sup>.

### (2) 通信开销

联邦学习中，客户端与服务器之间需要频繁地交换模型信息。由于联邦学习的过程中传递的数据量很大，设备与服务器的通信代价往往比设备端的一次计算代价更大，与服务器的频繁交互会消耗更多资源。同时，客户端的可用状态也是时刻变化的。这一系列特点加大了在有限的通信开销下应用联邦学习的难度。

### (3) 客户端的选择和调度问题

在联邦学习的过程中，服务器不可避免地需要选择使用哪些客户端参与训练，并且需要考虑如何公平地组织这些客户端设备。考虑到资源受限的边缘设备与服务器的频繁交互可能消耗更多资源，客户端有时还会出现掉队等现象<sup>[5]</sup>，调度问题显得尤为重要。客户端的选择和调度问题在减少能耗和利用更少带宽方面可以发挥至关重要的作用<sup>[2]</sup>。

综上所述，在边缘设备上进行联邦学习有许多亟待解决的问题。而要解决这些问题，着手点主要在于对联邦学习算法的改进和优化。

## 2.3 边缘计算的联邦学习算法

### 2.3.1 联邦平均算法

McMahan 等人提出了联邦学习的概念后介绍了联邦平均（Federated Averaging, FedAvg）算法<sup>[1]</sup>。

**算法 1.** 联邦平均（Federated Averaging, FedAvg）算法。

输入：全局迭代次数  $T$ ，本地训练轮次  $E$ ，客户端数量  $K$ ，客户端比例系数  $C$ ，客户端本地批大小  $B$ ，学习率  $\eta$ 。  
输出：最终模型参数  $\mathbf{w}$ 。

#### 服务器的进程

初始化全局模型参数  $\mathbf{w}_0$ 。

**对于** 每轮迭代  $t \in \{0, 1, \dots, T-1\}$  执行：

$c \leftarrow \max(\lfloor C \cdot K \rfloor, 1)$ 。

随机选择  $c$  个客户端作为本轮客户端集合  $R_t$ 。

向客户端集合  $R_t$  广播全局模型参数  $\mathbf{w}_t$ 。

**对于** 客户端集合  $R_t$  中的每个客户端  $k \in [1, c]$  并行执行：

接收客户端  $k$  的本地更新的模型参数  $\mathbf{w}_t^k$ 。

$\mathbf{w}_{t+1} \leftarrow \sum_{k=1}^c p_k \mathbf{w}_t^k$ 。

**返回** 最终模型参数  $\mathbf{w}_T$ 。

#### 客户端 $k$ 的进程

接收全局模型参数  $\mathbf{w}$ 。

将本地数据集  $D_k$  划分为大小为  $B$  的小批次  $\mathcal{B}$ 。

**对于** 每轮本地训练  $i \in \{1, 2, \dots, E\}$  执行：

**对于** 数据样本  $b \in \mathcal{B}$  进行训练：

$\mathbf{w} \leftarrow \mathbf{w} - \eta \nabla l(\mathbf{w}; b)$ 。

**返回** 本地更新的模型参数  $\mathbf{w}$ 。

算法 1 描述了 FedAvg 算法的逻辑。服务器首先初始化训练模型的参数  $\mathbf{w}_0$ ，随机选择一定比例的客户端参与训练，并在每轮迭代  $t$  中向客户端广播全局模型参数  $\mathbf{w}_t$ 。随后，每个客户端使用随机梯度下降（Stochastic Gradient Descent, SGD）算法，利用从本地数据集中选择的小批次进行多轮本地训练，最小化公式 1 给出的损失函数。服务器一旦收到来自各客户端的更新的模型，就聚合并更新全局模型。算法重复这个过程，不断最小化公式 2 给出的损失，直到达到设定的迭代次数，最后返回最终的模型参数。

文献[6]中分析了 FedAvg 算法的复杂度，从理论上证明了 FedAvg 对于强凸和光滑问题以  $O(1/T)$  的速度（ $T$  为随机梯度下降迭代次数）收敛于全局最优。

面对 2.2 节总结的一系列挑战，有关研究提出了许多算法，这些算法大多是基于 FedAvg 算法的。接下来将对其中具有代表性的算法做简单介绍，并阐明它们的主要贡献。

### 2.3.2 适用于非 IID 数据的算法

在 FedAvg 算法中，各客户端的训练过程中使用了相同的学习率和本地训练轮次。但是，本地训练轮次设置对收敛速度的影响至关重要，而数据分布不一致会恶化这一问题。Li 等人引入了 FedProx 算法<sup>[7]</sup>，使用边缘设备上的非 IID 数据进行联邦学习，实现了全局模型的快速收敛。FedProx 算法的改进之处在于客户端  $k$  不是最小化在本地数据集  $D_k$  上的损失函数  $f_k(\cdot)$ ，而是增加近端项惩罚，近似最小化  $h_k$ ：

$$\min_{\mathbf{w}} h_k(\mathbf{w}; \mathbf{w}^g) = f_k(\mathbf{w}) + \frac{\mu}{2} \|\mathbf{w} - \mathbf{w}^g\|^2. \quad (3)$$

其中  $\mu$  称为近端项惩罚系数， $\mathbf{w}$  为本地更新的模型参数， $\mathbf{w}^g$  为服务器下发的全局模型参数。当  $\mu = 0$  时，FedProx 退化为 FedAvg。利用这种方法，FedProx 算法可以在不需要手动设置本地训练轮次的前提下，考虑了数据的分布不平衡性，使本地更新模型更接近全局模型，提高了模型收敛的稳健性和稳定性。但与 FedAvg 算法类似的是，在 FedProx 算法中，所有客户端在聚合阶段的权重是相等的，因此缺少了对客户端功能差异（如硬件差异等）的考虑。

此外，Wang 等人发现在设备上训练数据的分布与基于这些数据训练的模型参数之间存在隐含的联系，可以根据这一点利用客户端其上传的模型参数推测该设备上的数据分布，并据此提出了一种基于经验的联邦学习框架 FAVOR<sup>[8]</sup>。FAVOR 采用深度 Q 学习机制，智能地选择联邦学习每轮迭代中参与的客户端，以抵消不平衡数据的影响，加快了收敛的速度。

### 2.3.3 通信高效的算法

减少通信的数据量和通信次数是这类算法的核心思想。一个直观的想法是对模型进行压缩，从而减少通信的数据量。Mills 等人提出了通信高效的联邦平均（Communication-Efficient

FedAvg, CE-FedAvg) 算法, 利用分布式 Adam 优化代替传统的 SGD 算法, 并压缩上传的模型, 减少了达到目标精度所需的通信轮数和每轮上传的总数据<sup>[5]</sup>, 从而使得联邦学习的过程更为通信高效. 在算法所需的通信次数方面, Reisizadeh 等人提出了周期平均量化的联邦学习算法 (Federated learning algorithm with Periodic Averaging and Quantization, FedPAQ) 算法<sup>[9]</sup>, 其主要思想是执行定期平均, 从而减少了通信频率. 具体而言, FedPAQ 算法不是在每次迭代时将更新的模型与服务器同步, 而是使客户端能够在与服务器共享更新的模型之前对模型执行多轮本地更新.

另外注意到在 FedAvg 算法中, 服务器对客户端发来的更新的模型参数  $w_t^k$  直接进行了加权平均, 这可能会使得某些参数失效, 对模型产生不利影响, 并显著增加了客户端的通信负担. 匹配平均的联邦学习 (Federated Learning with Matched Averaging, FedMA) 算法<sup>[10]</sup>考虑了神经网络中参数的置换不变性, 改进了服务器的聚合策略. FedMA 算法要求通信轮数等于网络的层数, 通过匹配和平均具有相似特征的隐藏元素, 逐层地聚合全局模型. 算法的大致流程如下: 服务器首先只从客户端收集模型的第一层参数, 并执行单层匹配以获得全局模型的第一层参数; 然后服务器将这些参数广播给客户端, 客户端继续训练后续的层, 同时冻结已经匹配的层. 不断重复此过程, 直至最后一层. 在此过程中服务器根据每个客户端数据的类别比例对其进行加权平均. 可以看到, FedMA 算法的优点之一是它比 FedAvg 算法更有效地利用了通信轮次, 减少了通信量. 然而, FedMA 可能容易受到攻击, 使模型中毒<sup>[4]</sup>.

除此之外, 还有很多方案来使联邦学习的过程更为通信高效, 包括下一节将提到的异步联邦学习方案. Liu 等人基于客户端—边缘侧—云端层次架构提出了层次联邦平均 (Hierarchical Federated Averaging, HierFAVG) 算法<sup>[11]</sup>. 该算法在边缘服务器聚合固定数量客户端的更新的模型, 同时在云服务器聚合各边缘服务器的模型, 得到全局模型. 与现有客户端—服务器架构相比, 这样的多级结构交换模型参数时在通信次数和数据量上更加高效. 但是, HierFAVG 算法并不能很好的解决客户端掉队的问题.

#### 2.3.4 调度高效的算法

在边缘计算中有两种主要的联邦学习方案, 分别是同步联邦学习和异步联邦学习. FedAvg 算法属于典型的同步联邦学习方案, 其中服务器只有接收到所有指定的客户端更新的模型参数之后, 才会进行聚合并广播新的全局模型参数, 这样就导致每轮迭代的训练时间主要取决于所有边缘节点的最大训练时间. 文献[12]中, 作者采用了异步的联邦学习方案, 并考虑了边缘设备资源受限的问题, 提出了通信高效的异步联邦学习 (Communication-Efficient

Asynchronous Federated Learning, CE-AFL) 算法, 服务器根据每轮迭代中所有边缘节点的到达顺序, 利用所有边缘设备中占比 $\alpha$ 的部分来对更新后的模型进行聚合. 这种调度方案在理论上显著增加了客户端的计算时间占比, 使客户端可以少通信、多计算, 从而提高了联邦学习的效率.

在公平地组织边缘设备方面, FedAvg 算法的优化过程会引入不同客户端之间的不公平性. 在文献[13]中, 作者提出了  $q$  公平联邦学习 (q-Fair Federated Learning, q-FFL) 问题来解决联邦学习中公平问题. q-FFL 问题在公式 2 给出的联邦学习的优化目标中引入了参数  $q$ :

$$\min_{\mathbf{w}} F_q(\mathbf{w}) = \sum_{k=1}^K \frac{p_k}{q+1} f_k^{q+1}(\mathbf{w}). \quad (4)$$

这样实现了对不同设备损失的重新加权计算, 使得损失较高的设备具有较高的相对权重, 并且设计了一种通信高效的 q-FedAvg 算法来解决这一问题. 这种方法减小了准确度分布方差, 实现了设备间更公平的准确度分布. 然而, 这种方法在降低方差的同时降低了部分客户端的性能.

### 3 总结与展望

随着物联网等领域的迅速发展, 联邦学习技术在边缘计算领域获得了广泛应用. 本文中我们讨论了在边缘计算场景中应用联邦学习的众多挑战与解决方案, 现有的算法和解决方案大多是从 FedAvg 算法出发, 针对细分的挑战进行改进和优化. 这些方法会在一定程度上加快训练进程, 但往往是以失去某些预测或分类的准确性为代价的. 因此, 未来的工作需要在准确性与通信高效之间进行权衡, 提出准确且高效的联邦学习算法. 我们注意到, 目前的工作更多的假设是凸优化问题, 为保证在具有受限的资源的情况下利用非 IID 数据进行的异步联邦学习仍然收敛, 未来的工作中需要将非凸优化问题的损失函数纳入考虑, 并提出支持算法. 此外, 异步的联邦学习会有更优的效率, 在通信效率等方面有一定的优势. 我们认为, 异步联邦学习本质上是如何平衡同步频率和收敛效果的问题, 更高的同步频率会带来更小的损失, 但这也对算法的收敛性提出了更高的要求.

综上所述, 对边缘计算的联邦学习算法进行设计时要考虑边缘多方面的受限因素, 结合上述算法的优势, 在保持收敛性的前提下使算法更为准确高效.



## 4 写作心得

在写作本文前，我查阅了许多有关联邦学习的文献，对联邦学习的一些经典算法有了较为深入的了解，对联邦学习这一种方法有了较为全面的认识。阅读这些外文文献的过程不仅是对前沿科技的学习和探索，更是对我外文阅读能力、数理基础的全面提升，特别是读到一些数学上的证明的时候。这也为我完成这篇综述打下了一定的基础。在本文的写作过程中，我从边缘计算的联邦学习算法中的有关研究中选取了三个广泛研究的问题进行探讨，基于FedAvg算法存在的不足，对有关算法进行了简要的整理和分析。在写作的过程中我也认识到，文献综述应不像“文献报告”一样，单纯把所阅读的文献一一陈述，还应该带有作者的评述。我自认为有关评述的内容写得还不够好，部分内容逻辑性并不强。最后，完成这篇综述也让我对有关算法有了一点自己的想法，不过仍需后续更广泛的文献阅读来验证。

## 参 考 文 献

- [1] McMahan H B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data//Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Fort Lauderdale, USA, 2017: 1273-1282.
- [2] Imteaj A, Thakker U, Wang S, et al. A survey on federated learning for resource-constrained IoT devices. IEEE Internet of Things Journal, 2022, 9(1): 1-24.
- [3] Dhar S, Guo J, Liu J, et al. A survey of on-device machine learning: An algorithms and learning theory perspective. ACM Transactions on Internet of Things, 2021, 2(3): 1-49.
- [4] Wahab O A, Mourad A, Otrouk H, et al. Federated machine learning: Survey, multi-level classification, desirable criteria and future directions in communication and networking systems. IEEE Communications Surveys & Tutorials, 2021, 23(2): 1342-1397.
- [5] Mills J, Hu J, Min G. Communication-efficient federated learning for wireless edge intelligence in IoT. IEEE Internet of Things Journal, 2020, 7(7): 5986-5994.
- [6] Li X, Huang K, Yang W, et al. On the convergence of FedAvg on non-IID data//Proceedings of the 8th International Conference on Learning Representations. Addis Ababa, Ethiopia, 2020: 1-26.
- [7] Li T, Sahu A K, Zaheer M, et al. Federated optimization in heterogeneous networks//Proceedings of Machine Learning and Systems, Austin, USA, 2020: 429-450.
- [8] Wang H, Kaplan Z, Niu D, et al. Optimizing federated learning on non-IID data with reinforcement learning//Proceedings of the 39th IEEE Conference on Computer Communications, Toronto, Canada, 2020: 1698-1707.
- [9] Reisizadeh A, Mokhtari A, Hassani H, et al. FedPAQ: A communication-efficient federated learning method with periodic averaging and quantization//Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics, Palermo, Italy, 2020: 2021-2031.

- [10] Wang H, Yurochkin M, Sun Y, et al. Federated learning with matched averaging//Proceedings of the 8th International Conference on Learning Representations, Addis Ababa, Ethiopia, 2020: 1-16.
- [11] Liu L, Zhang J, Song S, et al. Client-edge-cloud hierarchical federated learning//Proceedings of 2020 IEEE International Conference on Communications, Dublin, Ireland, 2020: 1-6.
- [12] Liu J, Xu H, Xu Y, et al. Communication-efficient asynchronous federated learning in resource-constrained edge computing. Computer Networks, 2021, 199: 108429.
- [13] Li T, Sanjabi M, Beirami A, et al. Fair resource allocation in federated learning//Proceedings of the 8th International Conference on Learning Representations, Addis Ababa, Ethiopia, 2020: 1-27.

知网个人查重服务报告单(全文对照)

报告编号:BC202206190931015563839995 检测时间:2022-06-19 09:31:01

篇名: 边缘计算的联邦学习算法  
作者: 张昊  
检测类型: 学术研究  
时间范围: 1900-01-01至2022-06-19

检测结果

去除本人文献复制比: 4.1% 去除引用文献复制比: 4.1% 总文字复制比: 4.1%  
单篇最大文字复制比: 1.9% (基于分布式优化思想的配电网用电负荷多层协同预测方法)  
重复字符数: [294] 单篇最大重复字符数: [134] 总字符数: [7119]



1. 边缘计算的联邦学习算法 总字符数: 7119

相似文献列表

去除本人文献复制比: 4.1%(294)	文字复制比: 4.1%(294)
1 基于分布式优化思想的配电网用电负荷多层协同预测方法 谭嘉;李知艺;杨欢;赵荣祥;鞠平; - 《上海交通大学学报》- 2021-12-28	1.9% (134) 是否引证: 否
2 基于卷积神经网络的联邦学习算法研究 尹祥(导师: 文泉) - 《电子科技大学硕士论文》- 2021-03-24	0.9% (62) 是否引证: 否
3 电容型设备缺陷预测模型研究 张亚萌(导师: 郑泽忠) - 《电子科技大学硕士论文》- 2019-05-17	0.7% (48) 是否引证: 否
4 物联网应用中的隐私保护机器学习模型研究 冯传凯(导师: 许艳) - 《安徽大学硕士论文》- 2021-06-30	0.5% (39) 是否引证: 否

原文内容	相似内容来源
此处有 40 字相似 ng; on-device training; resource-constrained devices 1 引言 联邦学习(Federated Learning, FL)是一种分布式机器学习方法, 广泛应用于边缘计算场景中,利用众多离散的边缘设备的计算能力训练机器学习模型。在传统的机器学习方法中,客户端需要将数据共享	物联网应用中的隐私保护机器学习模型研究 冯传凯 - 《安徽大学硕士论文》- 2021-06-30 (是否引证: 否) 1. 需求,保证图像数据保留在拥有者的本地设备上,研究者们致力于在分布式设备之间进行协作机器学习。文献[6]提出了一种称为联邦学习(Federated Learning, 简称 FL)的分散机器学习方法。在 FL 中,终端设备使用它们的本地图像数据来训练服务器所需的机器学习模型。然后终端设备将模型更新而不是图像数据
此处有 30 字相似 就需要对其算法流程进行一系列的优化和改进。 2 本文首先为联邦学习进行了定义,分析了其应用于边缘计算中的挑战,并总结了各类 联邦学习算法的优点和缺点。最后,探讨了边缘计算中联邦学习算法	基于卷积神经网络的联邦学习算法研究 尹祥 - 《电子科技大学博士论文》- 2021-03-24 (是否引证: 否) 1., E 代表局部的训练的轮次, T 代表整个代表的轮次。β 代表学习的速率。可以看出算法上分为服务器和客户端执行步骤。在联邦平均学习算法中,已经体现了联邦学习算法框架中一些关键步骤,这些步骤不管是在联合全局还是个性化模型类型中都是极其重要的。1.

	<p>的进一步研究方向.</p> <p>2 研究现状</p> <p>2.1 联邦学习定义</p> <p>联邦学习可以定义为一种分布式机器学习方法, 由多个客户端 (也</p>	<p>客户端上局部更新步骤。当客</p>
3	<p>此处有 45 字相似</p> <p>个或多个服务器 (server) 共同参与相同的模型训练任务. 其中服务器维护全局共享的模型, 客户端不会向服务器共享参与训练的</p> <p>数据, 而是在本地训练模型, 服务器通过定期聚合各客户端的模型信息来更新全局共享模型.</p> <p>图</p> <p>1 遵循客户端-服务器架构的联邦学习过程示意图.</p> <p>如图 1所示, 联邦学习的过程通常包括三个步骤[2]:</p> <p>(1) 在初</p>	<p>基于分布式优化思想的配电网用电负荷多层协同预测方法 谭嘉;李知艺;杨欢;赵荣祥;鞠平; - 《上海交通大学学报》- 2021-12-28 (是否引证: 否)</p> <p>1. 学习根据是否需要第三方服务器分为两种, 分别是客户端-服务器架构[26]和对等网络架构[27]. 前者是让多个客户端利用本地数据训练本地模型, 通过中央服务器聚合各个客户端的模型参数并更新全局模型, 然后将全局模型参数下发给各个客户端, 继续用本地数据更新模型, 直到全局模型收敛为止. 后者在训练过程中, 不需要第三方服务器,</p>
4	<p>此处有 50 字相似</p> <p>edAvg算法的逻辑. 服务器首先初始化训练模型的参数, 随机选择一定比例的客户端参与训练, 并在每轮迭代t中向客户端广播全局</p> <p>模型参数. 随后, 每个客户端使用随机梯度下降 (Stochastic Gradient Descent, SGD) 算法, 利用从本地数据集中选择的小批次进行多轮本地训练, 最小化公式1给出的损失函数. 服务器一旦收到来自各客户端的更</p>	<p>电容型设备缺陷预测模型研究 张亚萌 - 《电子科技大学博士学位论文》- 2019-05-17 (是否引证: 否)</p> <p>1. 含层的权重矩阵和偏置向量, (7)2(8)W 和(7)2(8)b 表示的是隐含层到输出层的权重矩阵和偏置向量. 关于模型参数的最优化, 随机梯度下降 (Stochastic Gradient Descent) 是最为电子科技大学硕士学位论文10常用的一种方法. 它的主要原理就是迭代训练, 对所有参数随机初始化, 计算梯度,</p>
5	<p>此处有 30 字相似</p> <p>radient Descent, SGD) 算法, 利用从本地数据集中选择的小批次进行多轮本地训练, 最小化公式1给出的损失函数.</p> <p>服务器一旦收到来自各客户端的更新的模型, 就聚合并更新全局模型</p> <p>. 算法重复这个过程, 不断最小化公式2给出的损失, 直到达到设定的迭代次数, 最后返回最终的模型参数.</p> <p>文献[6]中分析了F</p>	<p>基于分布式优化思想的配电网用电负荷多层协同预测方法 谭嘉;李知艺;杨欢;赵荣祥;鞠平; - 《上海交通大学学报》- 2021-12-28 (是否引证: 否)</p> <p>1. 器分为两种, 分别是客户端-服务器架构[26]和对等网络架构[27]. 前者是让多个客户端利用本地数据训练本地模型, 通过中央服务器聚合各个客户端的模型参数并更新全局模型, 然后将全局模型参数下发给各个客户端, 继续用本地数据更新模型, 直到全局模型收敛为止. 后者在训练过程中, 不需要第三方服务器,</p>
6	<p>此处有 64 字相似</p> <p>FedMA算法要求通信轮数等于网络的层数, 通过匹配和平均具有相似特征的隐藏元素, 逐层地聚合全局模型. 算法的大致流程如下:</p> <p>服务器首先只从客户端收集模型的第一层参数, 并执行单层匹配以获得全局模型的第一层参数; 然后服务器将这些参数广播给客户端, 客户端继续</p> <p>训练后续的层, 同时冻结已经匹配的层. 不断重复此过程, 直至最后一层. 在此过程中服务器根据每个客户端数据的类别比例对其进行加</p>	<p>基于分布式优化思想的配电网用电负荷多层协同预测方法 谭嘉;李知艺;杨欢;赵荣祥;鞠平; - 《上海交通大学学报》- 2021-12-28 (是否引证: 否)</p> <p>1. 器分为两种, 分别是客户端-服务器架构[26]和对等网络架构[27]. 前者是让多个客户端利用本地数据训练本地模型, 通过中央服务器聚合各个客户端的模型参数并更新全局模型, 然后将全局模型参数下发给各个客户端, 继续用本地数据更新模型, 直到全局模型收敛为止. 后者在训练过程中, 不需要第三方服务器, 客户端之间可以直接通信. 为了降低通信开销</p>

7	<p>此处有 35 字相似</p> <p>调度高效的算法</p> <p>在边缘计算中有两种主要的联邦学习方案，分别是同步联邦学习和异步联邦学习。FedAvg算法属于典型的同步联邦学习方案，其中服务器只有接收到所有指定的客户端更新的模型参数之后，才会进行聚合并广播新的全局模型参数，这样就导致每轮迭代的训练时间主要取决于所有边缘节点的最大训练时间。文献[12]中，作</p>	<p>基于卷积神经网络的联邦学习算法研究 尹祥 - 《电子科技大学博士论文》 - 2021-03-24 (是否引证: 否)</p> <p>1. I. I. D. 37.06 37.84 36.68N. I. I. D. 30.94 30.63 28.99当个性化的联邦学习算法完成，服务器下发每个客户端最终的个性化模型参数。类似于元学习算法，每个客户端接收到专有的个性化模型之后，依然可以基于该模型在自己的数据集上进行微调，以期取得更好的效</p>

- 说明：
- 1. 总文字复制比:被检测文献总重复字符数在总字符数中所占的比例
  - 2. 去除引用文献复制比:去除系统识别为引用的文献后, 计算出来的重合字符数在总字符数中所占的比例
  - 3. 去除本人文献复制比:去除系统识别为作者本人其他文献后, 计算出来的重合字符数在总字符数中所占的比例
  - 4. 单篇最大文字复制比:被检测文献与所有相似文献比对后, 重合字符数占总字符数比例最大的那一篇文献的文字复制比
  - 5. 复制比按照“四舍五入”规则, 保留1位小数
  - 6. 红色文字表示文字复制部分;绿色文字表示引用部分;棕灰色文字表示系统依据作者姓名识别的本人其他文献部分
  - 7. 系统依据您选择的检测类型、比对截止日期生成本报告