联邦学习研究综述

周传鑫，孙奕，汪德刚，葛桦玮
（信息工程大学，河南 郑州 450001）

摘　要：联邦学习由于能够在多方数据源聚合的场景下协同训练全局最优模型，近年来迅速成为安全机器学习领域的研究热点。首先，归纳了联邦学习定义、算法原理和分类；接着，深入分析了其面临的主要威胁与挑战；然后，重点对通信效率、隐私安全、信任与激励机制 3 个方向的典型研究方案对比分析，指出其优缺点；最后，结合边缘计算、区块链、5G 等新兴技术对联邦学习的应用前景及研究热点进行展望。

关键词：联邦学习；隐私保护；区块链；边缘计算
中图分类号：TP309.2
文献标识码：A
DOI: 10.11959/j.issn.2096−109x.2021056

Survey of federated learning research

ZHOU Chuanxin, SUN Yi, WANG Degang, GE Huawei
Information Engineering University, Zhenghzou 450001, China

Abstract: Federated learning has rapidly become a research hotspot in the field of security machine learning in recent years because it can train the global optimal model collaboratively without the need for multiple data source aggregation. Firstly, the federated learning framework, algorithm principle and classification were summarized. Then, the main threats and challenges it faced, were analysed indepth the comparative analysis of typical research programs in the three directions of communication efficiency, privacy and security, trust and incentive mechanism was focused on, and their advantages and disadvantages were pointed out. Finally, Combined with application of edge computing, blockchain, 5G and other emerging technologies to federated learning, its future development prospects and research hotspots was prospected.

Keywords: federated learning, privacy protection, blockchain, edge of computing

1　引言

随着计算机算力的提升，机器学习作为海量数据的分析处理技术，已经广泛服务于人类社会。然而，机器学习技术的发展过程中面临两大挑战：一是数据安全难以得到保障，隐私数据泄露问题亟待解决；二是网络安全隔离和行业隐私，不同行业、部门之间存在数据壁垒，导致

收稿日期：2020−06−23；**修回日期**：2020−10−10
通信作者：孙奕，sunyi-1001@163.com
基金项目：国家自然科学基金（61702550）
Foundation Item: The National Natural Science Foundation of China（61702550）

论文引用格式：周传鑫，孙奕，汪德刚，等. 联邦学习研究综述[J]. 网络与信息安全学报, 2021, 7(5): 77-92.
ZHOU C X, SUN Y, WANG D G, et al. Survey of federated learning research[J]. Chinese Journal of Network and Information Security, 2021, 7(5): 77-92.

数据形成“孤岛”无法安全共享[1]，而仅凭各部门独立数据训练的机器学习模型性能无法达到全局最优化。

为了解决以上问题，谷歌提出联邦学习[2]（FL，federated learning）技术，其通过将机器学习的数据存储和模型训练阶段转移至本地用户，仅与中心服务器交互模型更新的方式有效保障了用户的隐私安全。作为网络安全领域新的研究热点，联邦学习吸引了大量关注与研究。为了更加深入地展开研究，本文主要对现有研究成果做初步的梳理和总结，对典型方案进行详细分析与比较，指出它们的优势与不足，并结合边缘计算、区块链、5G 等新兴技术对联邦学习的应用前景和研究热点进行展望[3]。

2 联邦学习基本概念

2.1 联邦学习

传统的机器学习算法需要用户将源数据上传到高算力的云服务器上集中训练，这种方式导致了数据流向的不可控和敏感数据泄露问题。Mcmahan 等在 2016 年提出联邦学习技术[2]，允许用户在机器学习过程中既可以保护用户隐私，又能够无须源数据聚合形成训练数据共享。

联邦学习本质上是一种分布式的机器学习技术，其流程如图 1 所示。

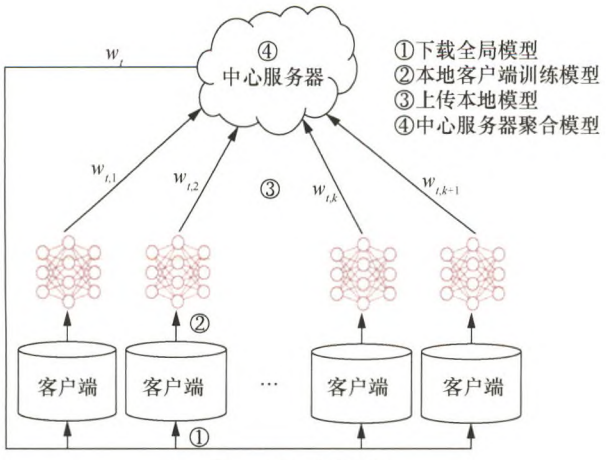


图 1 联邦学习流程

Figure 1 Process of federated learning

客户端（如平板电脑、手机、物联网设备）在中心服务器（如服务提供商）的协调下共同训练模型，其中客户端负责训练本地数据得到本地模型（local model）。中心服务器负责加权聚合本地模型，得到全局模型（global model），经过多轮迭代后最终得到一个趋近于集中式机器学习结果的模型 w，有效地降低了传统机器学习源数据聚合带来的许多隐私风险。

联邦学习的一次迭代过程如下。

（1）客户端从服务器下载全局模型 w_{t-1}。

（2）客户端 k 训练本地数据得到本地模型 $w_{t,k}$（第 k 个客户端第 t 轮通信的本地模型更新）。

（3）各方客户端上传本地模型更新到中心服务器。

（4）中心服务器接收各方数据后进行加权聚合操作，得到全局模型 w_t（第 t 轮通信的全局模型更新）。

综上，联邦学习技术具有以下几个特点。① 参与联邦学习的原始数据都保留在本地客户端，与中心服务器交互的只是模型更新信息。② 联邦学习的参与方联合训练出的模型 w 将被各方共享。③ 联邦学习最终的模型精度与集中式机器学习相似。④ 联邦学习参与方的训练数据质量越高，全局模型精度越高。

2.2 联邦学习算法原理

典型的联邦学习场景是在本地客户端设备负责存储和处理数据的约束下，只上传模型更新的梯度信息，在数千万到数百万个客户端设备上训练单个全局模型 w。中心服务器的目标函数 $F(w)$ 通常表现为

$$\min_w F(w), F(w)=\sum_{k=1}^{m}\frac{n_k}{n}F_k(w) \tag{1}$$

其中，m 是参与训练的客户端设备总数，n 是所有客户端数据量总和，n_k 是第 k 个客户端的数据量，$F_k(w)$ 是第 k 个设备的本地目标函数。

$$F_k(w)=\frac{1}{n_k}\sum_{i\in d_k} f_i(w) \tag{2}$$

其中，d_k 是第 k 个客户端的本地数据集，$f_i(w)=\alpha(x_i,y_i,w)$ 是具有参数 w 的模型对数据集 d_k 中的实例 (x_i,y_i) 产生的损失函数。d_k 中所有实例产生的损失函数之和除以客户端 k 的总数据量就是本地客户端的平均损失函数，损失函数与模型精度成反比，因此，机器学习的目标函数优化

通常是让损失函数达到最小值[4]。

联邦学习的目标函数优化算法中，通常采用大批量随机梯度下降（SGD）算法，即通过本地客户端模型训练的损失函数，乘以固定的学习率 η，计算出新一轮的权重更新。因此，本地客户端的模型权重更新如下：

$$w_{t,k} = w_{t-1,k} - \eta \nabla F_k(w) \tag{3}$$

第 t 轮通信中心服务器的模型聚合更新如下：

$$w_t = \sum_{k=1}^{K} \frac{n_k}{n} w_{t,k} \tag{4}$$

2.3 联邦学习分类

联邦学习的应用场景不同，客户端之间持有的数据集特征各不相同。假设 D_m 代表客户端 m 持有的数据，I 表示样本 ID，Y 表示数据集的标签信息，X 表示数据集的特征信息，因此，一个完整的训练数据集 D 应由(I,Y,X)构成。根据参与训练客户端的数据集特征信息 X 的不同，联邦学习被分为横向联邦学习、纵向联邦学习和联邦迁移学习[5]。

2.3.1 横向联邦学习

横向联邦学习的特点是数据集特征 X 和标签信息 Y 相同，但样本 ID 不同，其公式表达如下。

$$X_m = X_n, Y_m = Y_n, I_m \neq I_n, \forall D_m, \forall D_n, m \neq n \tag{5}$$

横向联邦学习如图 2 所示，$u_1 \sim u_6$ 表示数据集实例。

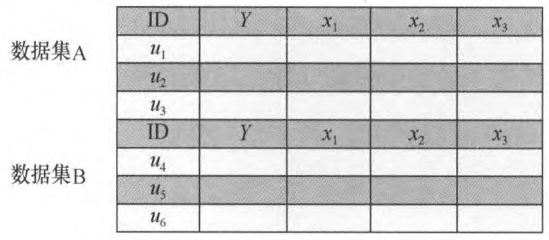


图 2　横向联邦学习
Figure 2　Horizontal federated learning

在用户输入法数据上训练的下一词预测模型[6]是典型的横向联邦学习应用。不同的手机用户具有相同的数据特征，数百万个安卓手机在云服务器的协调下训练共享的全局模型，其本质是将多方对不同目标的相同特征描述进行训练提取。

然而，在模型训练过程中，客户端数量较多，往往容易发生客户端恶意连接企图窃取信息，Li 等[7]对此提出基于检测的方法拒绝恶意客户端的接入。同时，该方式需要考虑服务器对客户端模型信息的可见性，针对不可信服务器，Bonawitz 等[8-9]引入安全多方计算来保护客户端更新的隐私性。

2.3.2 纵向联邦学习

纵向联邦学习的特点是各数据集特征 X 和标签信息 Y 不同，但样本 ID 信息相同，其公式表达如下。

$$X_m \neq X_n, Y_m \neq Y_n, I_m = I_n, \forall D_m, \forall D_n, m \neq n \tag{6}$$

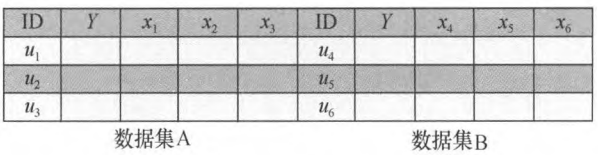


图 3　纵向联邦学习
Figure 3　Vertical federated learning

纵向联邦学习中一方掌握训练的标签信息 Y，各方通过输入特征信息 X，得到纵向全局模型。其典型应用场景如同一地区的银行和电商平台：银行拥有当地用户的收支记录 x_1，电商平台拥有用户的消费记录和浏览记录 x_2，双方想通过数据联合对客户信用 Y 进行评级，从而提供更个性化的服务，其本质是将多方对相同目标的不同特征描述进行训练提取。为防止纵向联邦学习中恶意用户推测出他方私有用户数据，Cheng 等[10]通过 RSA 和哈希函数确保参与训练的各方只能获得基于各方共有用户特征训练得到的模型。

2.3.3 联邦迁移学习

联邦迁移学习的特点是数据集特征 X、标签信息 Y 和样本 ID 信息都不同，其公式表达如下：

$$X_m \neq X_n, Y_m \neq Y_n, I_m \neq I_n, \forall D_m, \forall D_n, m \neq n \tag{7}$$

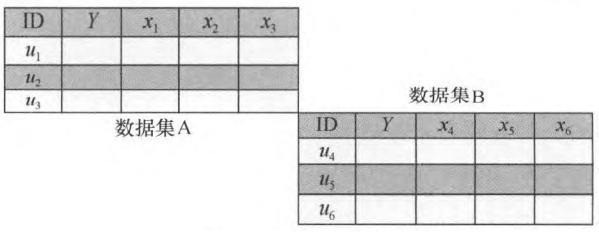


图 4　联邦迁移学习
Figure 4　Federated transfer learning

联邦迁移学习被用于解决标签样本少和数据

集不足的问题[5]，如中国的电商平台与其他国家银行之间的数据迁移，由于跨部门跨国的数据交流很难实现，通过联邦迁移学习可以很好地解决这类痛点问题。

3 联邦学习存在的威胁与挑战

自联邦学习的概念提出后，其迅速得到了学术界广泛的关注与研究，但是目前这一研究方向仍有许多威胁与挑战亟待解决，其中，最核心的问题包括通信效率短板明显、隐私安全仍有缺陷、缺乏信任与激励机制，这些问题极大地限制了联邦学习的进一步发展与应用。

3.1 通信效率短板明显

在联邦学习网络中，服务器与远程客户端之间往往需要进行不断的通信来交互模型更新信息，动辄万计的客户端很容易对通信网络造成巨大的带宽负担。通常，全局模型训练时间分为数据处理时间和通信传输时间两部分，而随着计算机设备算力的提升，数据处理时间不断降低，联邦学习的通信传输效率变成限制其训练速度的主要因素[11]。

联邦学习与分布式计算的区别是联邦学习的数据集来自各个终端用户，这些用户产生的数据特征往往呈非独立同分布（Non-IID）。Non-IID指的是在概率统计理论中，各数据集中的随机变量不服从同一分布，即对于不同的客户端 i 和 j，它们的数据集概率分布 $P_i \neq P_j$。而传统的分布式框架算法只有在处理独立同分布（IID）数据时表现良好，而在处理 Non-IID 数据时会造成训练过程难以收敛、通信轮数过多等问题[2,11]。另外，在互联网环境中，大量本地模型的更新、上传会导致中心服务器通信开销过大，无法满足正常的应用要求，同时相邻的模型更新中可能包含许多重复更新或者与全局模型不相关的更新[12]。

综上，联邦学习的通信效率优化具有重要的研究意义。通常改进方案有两个目标：减少每轮通信传输的数据大小；减少模型训练的总轮数。目前，改进通信效率方案主要是通过优化联邦学习框架算法、压缩模型更新和采用分层分级的训练架构。这些方案一定限度上提升了联邦学习模型训练速度、减小了数据通信量，对联邦学习技术的完善具有重大意义，但现阶段仍然存在许多难以解决的问题。例如，优化算法在处理Non-IID数据时相对于处理IID数据的时间开销成倍增长[2]；压缩算法虽然能够显著降低通信数据大小，但同时会严重影响模型精度，在通信效率和模型精度之间的平衡成为挑战[11,13]；分层分级的训练架构也不适合于所有的联邦学习场景，有时这种物理结构并不存在。

3.2 隐私安全仍有缺陷

联邦学习通过源数据不出本地而仅交互模型更新（如梯度信息）的方式来保护用户的敏感数据，开创了数据安全的新范式。理想情况下，联邦学习中客户端通过训练源数据上传本地模型，服务器仅负责聚合和分发每轮迭代形成的全局模型。然而，在真实的网络环境中，模型反演攻击、成员推理攻击、模型推理攻击层出不穷，参与训练的客户端动机难以判断，中心服务器的可信程度难以保证，仅通过模型更新来保护用户隐私的方式显然是不够的。

研究表明，梯度信息会泄露用户的隐私数据[14-20]，攻击者可以通过客户端上传的梯度信息间接推出标签信息和数据集的成员信息。Carlini等[15]从训练用户语言数据的递归神经网络中提取出了用户的敏感数据，如特定的银行卡号。Fredrikson 等[16] 研究了如何从模型信息中窃取数据隐私，并通过药量预测实验实现了对线性回归模型的反演攻击，获得了患者的敏感信息。Hitaj等[18]用生成对抗网络（GAN）对模型聚合发起攻击，实验结果表明，恶意客户端能够通过产生相似的本地模型更新来窃取用户数据隐私。Gei 等[19]证明了从梯度信息重建输入数据的可行性与深度网络架构无关，并将一批输入图像用余弦相似度和对抗攻击的方法恢复出来。

如图 5 所示，联邦学习主要存在 3 种威胁：恶意客户端修改模型更新，破坏全局模型聚合；恶意分析者通过对模型更新信息的分析推测源数据隐私信息；恶意服务器企图获得客户端的源数据。针对以上威胁，增强联邦学习隐私安全性的主流方案与经典机器学习隐私保护技术结合，包括差分隐私（DP，differential privacy）、安全多方

计算（MPC，secure multi-party computation）、同态加密（HE，homomorphic encryption）等技术[20]。大量的研究表明，联邦学习与这些隐私保护技术的结合能够提供足够强的安全性，但仍然存在一些问题需要解决。例如，与差分隐私的结合在较少客户端参与的联邦学习中，模型精度受到较大的影响，虽然在大量客户端参与时能够通过模型加权平均抵消噪声误差，但算法中包含的大量超参数仍然限制了进一步的应用[21-22]；与安全多方计算和同态加密技术的结合能够提供无损全局模型的构建，但同时会造成较大的通信开销[21-23]，如何平衡通信负担和模型安全是一个相当大的挑战。

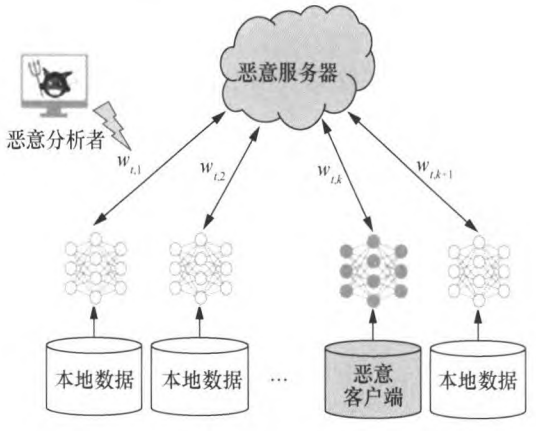


图 5　联邦学习中的安全威胁

Figure 5　Security threats in federated learning

3.3　缺乏信任与激励机制

联邦学习为现代社会建立了一个数据安全共享的架构，在未来万物互联的场景中，不同的机构、部门之间的数据联合会形成一个巨大的联邦学习联盟，旨在构建基于大数据和多特征融合的智能分析决策模型。但是，数据联盟需要吸引大量客户端参与到训练过程中，没有高效的激励机制很难吸引足够的训练数据，无法保证最终的智能模型质量；另外，联邦学习并没有针对客户端的信任机制，对于客户端的信誉没有统一的分数评价，这严重影响了对优质客户端的选择，从而导致全局模型精度降低。

针对以上问题，学术界通过结合区块链技术做出了大量研究。区块链是比特币的底层技术，它作为一种安全可靠、不可篡改和支持查询验证的分布式分类账，被应用于解决各类数据安全存储和信任问题[26-27]。联邦学习通过集成区块链能够以一种安全、高度抗中断和可审计的方式记录其模型更新，为系统框架提供可问责性和不可否认性。同时，区块链的激励机制作为一种经济回报能够根据构建模型时客户端的贡献给予相应的奖励。

4　联邦学习技术研究进展

针对联邦学习中存在的威胁与挑战，目前已经存在许多解决方案，本节对大量文献进行总结，分别就联邦学习的通信效率、隐私安全、信任与激励机制 3 方面展开研究。

4.1　通信效率

目前的研究中针对通信效率的改进主要有以下 3 种方法。

（1）算法优化：开发适合处理 Non-IID 和非平衡分布数据的模型训练算法，减少用于传输的模型数据大小，加快模型训练的收敛速度。

（2）压缩：压缩能够有效降低通信数据大小，但对数据的压缩会导致部分信息的丢失，此类方法需要在模型精度和通信效率之间寻找最佳的平衡。

（3）分散训练：将联邦学习框架分层分级，降低中心服务器的通信负担。

在大多数情况下，这几种方法是相辅相成的，通过特定的方法把这几种方案结合是研究的热点方向[28-29]。表 1 给出现有通信效率算法的性能比较。

4.1.1　算法优化

算法优化是对分布式机器学习框架的改进，使该框架更适用于海量客户端、高频率、低容量、数据特征不均的联邦学习环境，实现通信轮数和模型更新数据的减少。

在分布式计算框架中，客户端每运行一次 SGD 算法训练，机器学习模型就会向中心服务器上传本轮产生的本地模型更新。但是，频繁的通信交互会对参与训练各方造成不必要的通信负担。McMahan 等[2]针对联邦学习的低带宽环境提出 FedAvg 算法，要求客户端在本地多次执行 SGD 算法，然后与中心服务器交互模型更新，实现用更少的通信轮数训练出相同精度的模型。相比于基准算法 FedSGD[30]，其在训练不同神经网络的

通信轮数上减少了 1%~10%，但该算法对于非凸问题没有收敛保证，在非 IID 数据集上难以收敛[31]。

自 FedAvg 算法被提出，后续大量研究在此基础上做进一步的拓展，但 FedAvg 算法本身有一定的缺陷[32]。首先，服务器端聚合时根据客户端数据量大小来分配相应的权重，这导致拥有大量重复数据的客户端能够轻易影响全局模型；其次，客户端仅执行 SGD 算法和执行固定次数的 SGD 算法一定限度上限制了模型训练的速度。对此，Li 等[33]提出 FedProx 算法，根据客户端设备可用的系统资源执行可变次数的 SGD 算法，缩短收敛时间的同时将模型更新数据压缩了 1/2~1/3，更加适用于客户端数据质量、计算资源等联邦学习场景。同样是针对联邦学习框架的改进，Liu 等[34]认为传统的 FL 仅利用一阶梯度下降（GD），忽略了对梯度更新的先前迭代，提出了 MFL 方案，在联邦学习的本地模型更新阶段使用动量梯度下降（MGD），实验证明，在一定条件下该方案显著提升了模型训练的收敛速度。Huang 等[35]提出迭代自适应的 LoAdaBoost 算法，通过分析客户端更新的交叉熵损失，调整本地客户端 epoch 次数，相对于传统 FedAvg 算法固定 epoch，准确度与收敛速度均有显著提升。

除了对最初的 FedAvg 算法的各种改进以外，在客户端或者服务器上增加筛选算法也是研究方向之一。Wang 等[12] 认为客户端上传的本地模型更新中含有大量的冗余和不相关信息，严重占用通信带宽，因此提出 CMFL 算法，该算法要求客户端筛选本地模型更新与上一轮全局模型的相关度，通过模型梯度正负符号相同的百分比来避免上传达不到阈值要求的本地模型更新，实现通信开销的降低，但该算法建立在客户端按照协议执行的基础上，系统的鲁棒性较弱。Jiang 等[36]提出了 BACombo 算法，利用 gossip 协议和 epsilon-greedy 算法检查客户端之间随时间变化的平均带宽，最大限度地利用带宽容量，进而加快收敛速度。

4.1.2 压缩

压缩方案通常分为两种：梯度压缩和全局模型压缩。通常情况下，梯度压缩相比于全局模型压缩对通信效率的影响更大，因为互联网环境中上行链路速度比下载链路速度慢得多，交互通信的时间主要集中在梯度数据上传阶段。

横向联邦学习中往往有大量的本地客户端，很难保证每个客户端都拥有稳定可靠的网络连接，低质量的通信会严重降低通信速度。Konečný 等[11] 提出针对本地模型的结构化更新和草图更新算法，客户端被要求在一个低秩或随机掩码后的有限空间中进行模型学习，然后草图更新算法对模型更新进行量化、随机旋转和子采样等压缩操作，该方案被证明在 SGD 迭代方面显著减慢了收敛速度。在上述基础上，Caldas 等[13] 将该方法应用于对全局模型更新的压缩中，同时提出 Federated Dropout 思想优化模型更新，中心服务器随机选择全局模型的更小子集并采用量化、随

表 1 通信效率算法的性能比较

Table 1 Performance comparison of communication efficiency algorithms

文献	压缩	本地优化	算法性能	算法特点
[30]			弱	分布式计算基准算法
[2]	√	√	弱	FedAvg 算法与其优化
[34-35]		√	强	针对 FedAvg 算法的优化
[12]		√	强	优化筛选机制、过滤无关更新
[11]	√		弱	传统压缩方法
[13]	√	√	强	传统压缩方法+算法优化
[38]	√		强	适应性修改压缩阈值
[31]	√		强	对 non-IID 数据表现较好

注：以 FedAvg 为基准，算法性能大于 3 倍为强压缩，小于 3 倍为弱压缩，“√”表示通信效率算法的类别。

机旋转和子采样等压缩操作，客户端接收到全局模型后解压缩并进行本地模型训练，从而减少了联邦学习对客户端设备资源的影响，允许培训更高容量的模型，并接触到更多样化的用户。Reisizadeh 等[37]选择将算法优化与压缩的思路结合起来，其提出的 FedPAQ 算法要求服务器只选择一小部分客户端参与训练，同时客户端减少上传本地模型次数并在上传之前进行量化更新操作减小通信量。

但是，上述算法采取的都是固定阈值的压缩通信，这种方式在客户端之间模型更新差异较大时显得并不合理。对此，Lu 等[38]提出自适应阈值梯度压缩算法，客户端通过判断梯度变化，计算得到适当的阈值用于压缩通信，同时保证模型的性能损失较小。

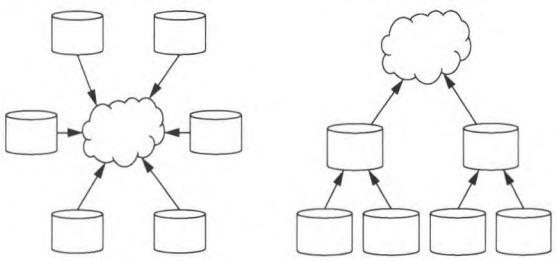
另外，现有的大部分压缩方法只在呈 IID 分布的客户端数据下表现良好，这些方法并不适合联邦学习场景。对此，Sattler 等[31]提出一种新的稀疏三元压缩（STC）框架，STC 扩展了现有的 top-k 梯度稀疏化压缩技术，通过 Golomb 无损编码压缩联邦框架交互的模型更新，使算法更适用于高频率低容量的联邦学习环境，同时保证了在大量客户端参与下的鲁棒性。

4.1.3　分散训练

在联邦学习中，通信拓扑通常是星形拓扑，但这往往会造成中心服务器的通信成本太大，分散拓扑（客户端只与它们的邻居通信）可以作为一种替代方案，如图 6 所示。在低带宽或高时延网络上运行时，分散拓扑被证明比星形拓扑训练速度更快[32-40]。联邦学习的分散拓扑[41-44]先设定边缘服务器聚合来自客户端设备的本地更新，然后边缘服务器充当客户端的角色与中心服务器交互。例如，Sharma 等[43]构建了一个多层分布式计算防御框架，通过数据层、边缘层、雾层和云层的协同决策，解决海量数据集中传输的问题。通过这种分层通信的方法可以有效降低中央服务器的通信负担，但它并不适用于所有的场景，因为这种物理层次可能不存在，也不可能预先知道。

4.2　隐私安全

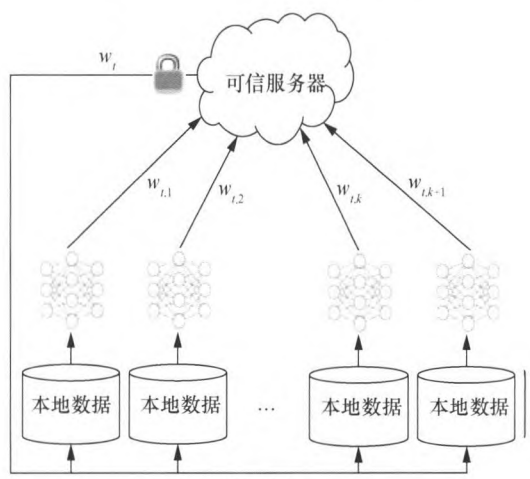
为解决联邦学习中暴露的隐私泄露问题，学术界做了大量研究来增强隐私安全性。根据隐私保护细粒度的不同，联邦学习的隐私安全被分为全局隐私（global privacy）和本地隐私（local privacy），如图 7 所示。全局隐私假定中心服务器是安全可信任的，即每轮通信的模型更新中心服务器可见。本地隐私假定中心服务器同样可能存在恶意行为，因此本地模型更新在上传到中心服务器之前需要进行加密处理。表 2 为改进联邦学习隐私安全性的算法对比。



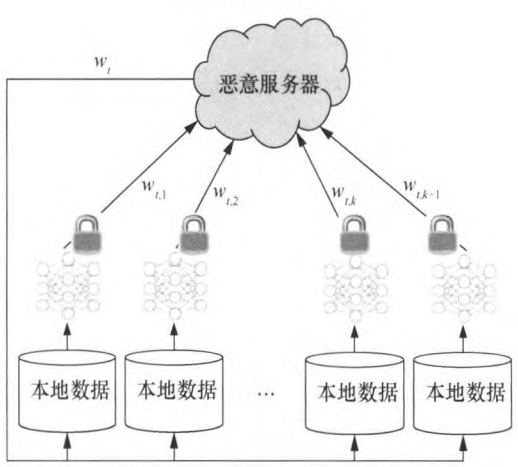
(a) 星形拓扑　　(b) 分散拓扑

图 6　两种不同的网络拓扑结构

Figure 6　Two different network topologies



(a) 全局隐私



(b) 本地隐私

图 7　两种不同的隐私保护方案

Figure 7　Two different privacy protection schemes

表 2 改进联邦学习隐私安全性的算法对比

Table 2 Comparison of algorithms for improving the privacy and security of federated learning

文献	技术结合	隐私保护类型	特点
[14, 21-22]	差分隐私	全局隐私	隐藏客户端的贡献
[51]	差分隐私	全局隐私	减少超参数数目
[23-24]	安全多方计算	本地隐私	模型精度无损、开销大
[29, 54-56]	差分隐私	本地隐私	精度略有损失
[45, 57]	同态加密	本地隐私	模型精度无损
[58]	差分隐私	本地隐私	应用于 MEC 与 IoT
[7]	恶意检测	模型更新检测	深度学习检测恶意模型
[63]	恶意检测	模型更新检测	算力资源消耗较小

4.2.1 典型隐私保护技术

现有的方案主要通过结合典型隐私保护技术来提供进一步的隐私增强，如差分隐私、安全多方计算、同态加密等技术，这些技术在之前的研究中已经被广泛应用于传统机器学习的隐私保护[45]。

定义 1 差分隐私。设随机化算法 A，对于两个至多有一条数据不同的数据集D和D'以及任意可能的输出 S，若算法 A 满足

$$\mathrm{pr}[\mathrm{A}(D)\in S]\leqslant e^{\varepsilon}\mathrm{pr}[\mathrm{A}(D')\in S]+\delta \quad (8)$$

则称随机化算法 A 满足(ε,δ)差分隐私保护。其中，ε代表隐私保护预算，δ是算法允许的误差，通常为较小的常数。

Dwork 等[46]于 2006 年提出差分隐私概念，并使用严格的数学推导给出了安全性证明。通常差分隐私算法的噪声机制分为指数噪声、Laplace 噪声和高斯噪声，其中，指数噪声主要用于处理离散数据集，Laplace 噪声和高斯噪声主要用于处理连续数据集。

定义 2 安全多方计算[48]。假设有 n 个参与方 $P_1, P_2,\cdots,P_n$ 分别拥有自己的敏感数据 $m_1, m_2,\cdots, m_n$，这 n 个参与者在不泄露各自输入数据的前提下共同执行一个协议函数 $f(m_1,m_2,\cdots,m_n)$。

安全多方计算的研究焦点是在没有可信第三方的条件下，参与训练各方安全计算的一个共同的约束函数。姚期智[49]于 1983 年提出安全多方计算的概念，通过混淆电路、不经意传输、秘密分享等技术实现多方共同运算，并确保各方数据的安全性。

定义 3 同态加密。设有明文数据 $d_1,d_2,\cdots,d_n$，这 n 个数据对应的加密数据为 $m_1,m_2,\cdots,m_n$，若加密算法满足

$$\begin{aligned}&\mathrm{Enc}(f(m_1,m_2,\cdots,m_n))=\\&f(\mathrm{Enc}(m_1),\mathrm{Enc}(m_2),\cdots,\mathrm{Enc}(m_n))\end{aligned} \quad (9)$$

则称该加密算法满足同态加密。同态加密能够直接对密文数据进行密码学运算，最终运算结果经解密后与在明文上直接运算结果一致。Rivest 等[50]于 1978 年提出同态加密概念。同态加密分为全同态加密和部分同态加密，其中部分同态加密分为乘法同态和加法同态，若一个算法既满足乘法同态又满足加法同态，则称为全同态加密算法。

4.2.2 全局隐私

在全局隐私中，假设存在一个受信任的服务器，外部敌手可能是恶意客户端、分析师、使用学习模型的设备或它们的任何组合。恶意客户端可以从中心服务器接收到它们参与轮的所有模型迭代信息，分析师可以在不同的训练轮中使用不同的超参数来研究模型迭代信息。因此，对中间迭代过程和最终模型进行严格的加密保护十分重要。

在联邦学习进程中，恶意客户端能够通过对分布式模型的分析，获得客户端在训练过程中的贡献及数据集信息。Geyer 等[21] 提出一种针对客户端的差分隐私保护联邦优化算法，实现了对模型训练期间客户端贡献的隐藏，在有足够多客户端参与的情况下，能够以较小的模型性能成本来达到用户级差分隐私。McMahan 等[22]同样使用差分隐私加密全局模型更新，证明了如果参与联邦

学习的客户端数量足够多，对模型更新信息的加密就会以增加计算量为代价而不会降低模型精度。Bhowmick 等[14] 利用差分隐私技术，通过限制潜在对手的能力，提供同等隐私保护程度的同时保证了更好的模型性能。

但是，上述方案中都存在许多影响通信效率和精度的超参数，用户必须谨慎选择才能达到预期效果。Thakkar 等[51]针对这个缺点提出自适应梯度裁剪策略，对特定层添加不同的噪声，同时对迭代差分隐私机制应用自适应分数剪裁，有效缓解了差分隐私算法中超参数过多的问题。

4.2.3 本地隐私

针对不可信服务器和恶意敌手反演攻击的问题，结合传统的安全多方计算和同态加密等技术，能实现模型信息的无损加解密，但却大大增加了通信成本与计算开销。Bonawitz 等[23]提出 Secure Aggregation 模型，结合秘密分享等技术使服务器无法解密单一客户端的梯度信息，仅能执行聚合操作得到全局模型，从而实现对恶意服务器的信息隐藏。Mandal 等[24]在此工作基础上做了通信效率的改进，引入它非交互式成对密钥交互计算（NIKE）技术，在离线阶段计算主密钥的同时限定用户最多与 L 个邻居进行掩码操作，从而有效减少了秘密分享的时间开销。Dong 等[28]将秘密分享与同态加密应用于通信效率算法（TernGrad），解决了隐私泄露的同时大幅提升了框架的通信和计算开销。Hao 等[52]通过改进 BGV 同态加密算法，消除了密钥交换操作并增加了纯文本空间，提供后量子安全性的同时避免了交互密钥导致的通信负担。在纵向联邦学习场景中，各部门进行训练数据对齐时可能造成标签信息和私有数据的泄露。Cheng 等[25]通过改进 XGBoost 树模型提出 SecureBoost 算法，其利用 RSA 和哈希函数实现各方数据的共有样本 ID 对齐，同时使用加法同态加密保护各方交互的标签信息和梯度直方图信息，最终实现了与不添加隐私保护的联邦学习相同的模型精度。Aono 等[53] 对深度神经网络模型进行同态加密的思想为联邦学习提供了新方向。

另一个研究热点是联邦学习与差分隐私的融合，由于差分隐私不增加客户端通信成本，被广泛应用于模型更新的隐私保护。学术界的研究主要致力于在保护隐私信息的前提下，尽可能地减少噪声对模型训练的影响，进而提升模型性能。Liu 等[54]提出一种自适应隐私保护的 APFL 方案，通过分析数据集的特征向量 x_i 对输出模型的影响，为不同贡献的特征向量分配不同的隐私预算 ε，同时减少贡献较少的数据集的噪声，实现严格差分隐私的同时高效保证了全局模型精度与性能。Huang 等[55]针对客户端之间的不平衡数据提出 DP-FL 框架，其根据每个用户的数据量设置不同的差分隐私预算 ε，设计具有自适应梯度下降算法的差分隐私专用卷积神经网络，来更新每个用户的训练参数，结果证明相较于传统的 FL 框架，该方案在不平衡数据集中表现较好。Wei 等[56]将差分隐私与 FL 的结合做了深入的分析，证明存在最优的 K 值（$1 \leqslant K \leqslant$ 总客户端数 N），可以在固定的隐私保护级别上实现最佳的收敛性能。Cao 等[29]则从通信效率和隐私保护的结合出发，结合本地差分隐私，为物联网终端低算力设备提供了资源消耗的隐私保护框架。

但是，上述方案主要致力解决服务器不可信的问题，没有考虑服务器是否正确执行指定聚合操作，恶意服务器很有可能会回传虚假全局模型，蓄意破坏特定客户端对全局模型的使用。针对这类信任问题，Xu 等[57]提出具有隐私保护和模型可验证的联邦学习框架 VerifyNet，通过双掩码协议保证客户端本地梯度的保密性，同时将中心服务器欺骗客户端的困难性转移到解 NP-hard 数学难题上，保证了全局模型的完整性和正确性。

随着联邦学习在移动边缘计算（MEC）和物联网（IoT）中的广泛应用，其存在的安全与隐私问题开始受到关注。Lu 等[58]提出了一种差分隐私异步联邦学习（DPAFL）方案，通过将本地差分隐私引入联邦学习中，在本地模型的 SGD 更新中加入高斯噪声以保护隐私性，同时开发了一个新的异步联邦学习架构，它利用分布式的点对点更新方案，而不是集中式更新，以减轻集中式服务器带来的单点安全威胁，更适用于 MEC 环境。后来，Lu 等[59]将这种方案应用于车载网络物理系统，解决车辆物联网环境下敏感数据泄露的问题。

Hu 等[60]在异构物联网环境中使用联邦学习结合差分隐私保障用户隐私，提出一种对用户设备异质性具有鲁棒性的 FL 算法。

4.2.4 模型更新检测

对于模型更新的异常检测同样是确保训练过程安全的重要方式，Fang 等[61]通过客户端的本地模型发起中毒攻击使全局模型具有较大的测试错误率，并对 4 种拜占庭鲁棒性联邦学习框架进行了攻击研究，证明了联邦学习对局部模型中毒防御的必要性。

在联邦学习环境中，通常有数以万计的设备参与训练，服务器如果无法及时检测恶意客户端，很容易造成全局模型被污染甚至隐私泄露问题。Li 等[7] 提出基于检测的算法，通过一个预先训练的自动编码器神经网络来检测异常的客户行为，并消除其负面影响，给出各客户端信用评分并拒绝恶意客户端的连接。Zhao[62]等通过在服务器端部署 GAN，通过客户端模型参数生成审计数据集，并利用该数据集检查参与者模型的准确性，确定是否存在中毒攻击。实验证明，该方法相比传统的模型反演方法，生成的审计数据集质量更高。

但是，上述提出的检测算法需要消耗服务器大量的算力审核客户端本地模型，这导致在全诚实客户端参与的联邦学习中，资源遭到极大的浪费。对此，为减少算力消耗，Kang 等[63]通过经典的 RONI 中毒攻击检测算法，比较数据库中有没有相似的本地模型更新效果来判断是否中毒，然后对客户端给出信誉分以供任务发布者选择信誉值高的客户端参与训练，进而排除恶意客户端攻击的可能。Fung 等[64]将这种比较放在本地模型与上一轮全局模型上，通过比较本地模型更新与全局模型更新向量方向的相似性，判断客户端是否存在恶意。Chen 等[65]基于受信任的执行环境，设计了训练完整性协议用于检测不诚实的行为，如篡改本地训练模型和延迟本地训练进程，实验证明该方案具有训练完整性与实用性。

4.3 信任与激励机制

联邦学习中，一方面，由于服务器的中心协调地位，往往存在单点故障、执行环境不可信等信任问题；另一方面，如何建立激励机制使参与方自愿消耗算力参与到数据联邦中是一项重大的挑战。鉴于此，学术界主要通过结合区块链技术为联邦学习提供信任与激励机制。区块链具有的数据库不可篡改、安全可验证的特性解决了联邦学习在发展过程中的痛点问题，表 3 为基于区块链的联邦学习方案对比。

表 3 基于区块链的联邦学习方案对比

Table 3 Comparison of federated learning schemes based on blockchain

文献	关注问题	区块链作用
[45]	单点故障、安全可信	分布式架构+安全可验证
[46-48,50,58,83]	激励机制	可信存储+建立共识
[45,49,64,80]	可审计、鲁棒性	矿工检验+智能合约

联邦学习中，由于服务器的中心聚合作用，其设备一旦受到敌手单点攻击就会对整个学习框架造成很大的安全隐患。为提升框架的安全性、可信性、可靠性，Majeed 等[67]提出基于区块链的 FLchain 架构，以提升联邦学习的安全性，在 FLchain 中，对于每个全局模型，框架都会创建一个新的通道来存储特定通道分类账，同时创建“全局模型状态树”来跟踪全局模型的权重更新，FLchain 以一种不可篡改的方式保证了 FL 模型的起源性和可审计性。Sharma 等[43]在分布式多层计算框架的基础上使用离线区块链和在线区块链实时存储大量节点的临时训练数据，利用多层和多链结构有效减少网络故障、恶意攻击对联邦学习的影响，但没有将区块链的激励机制作为提升模型性能的辅助措施。Arachchige 等[68]通过融合差分隐私、联邦学习、以太坊区块链和智能合约构建了名为 PriModChain 的框架，为联邦学习在工业物联网中的应用提供了隐私性、安全性与可靠性，但该框架的运行效率限制了其进一步的发展。Lu 等[69]提出了一种新的混合区块链架构，其由许可区块链和本地有向无环图（DAG）组成，以实现车联网中的有效数据共享，提升学习模型的可靠性。Pokhrel 等[70]通过私有区块链提出一种多级信任框架，以实现本地模型更新从观察到学习和验证的端到端可信性。

联邦学习框架不仅存在单点故障问题，在没有良好激励机制的情况下参与训练的客户端可能

会不上传或上传虚假的模型更新。针对激励机制的设计，Kim 等[71]提出了 BlockFL 架构，其中每个设备将本地模型更新上传给区块链网络中的关联矿工，矿工负责对模型更新进行交换和验证后记录到区块链中，并提供相应的奖励。Kang 等[72]引入声誉概念作为客户端信任度的衡量指标，利用多权重的主观逻辑模型设计了基于声誉的可信客户端选择方案，同时通过区块链的不可篡改性实现分布式的信誉管理，并使用契约理论分析参与构建模型的客户端的算力投入、模型质量等因素给予相对应的回报。Weng[73]等提出 DeepChain 方案，通过区分客户端在训练过程中表现的活性和兼容性，促使客户端发送正确的、高质量的模型更新，同时使用区块链技术保证模型安全和训练过程的可审核性，实现保密性、可审核和公平公正的目标。Kim 等[74] 通过区块链技术对所有的模型更新进行完整的记录，并给予丰厚的奖励来激励用户参与联邦学习，提出了基于权重的客户端子集选择方案，通过每个客户端局部模型的精度和参与训练的频率来选择用于训练的客户端，实现了较高的稳定性和较快的收敛速度。Zhan 等[75]设计了一种基于深度强化学习的（DRL）激励机制，将传统的资源分配策略应用于 FL 分布式特殊场景，以达到边缘节点的最佳训练策略和定价策略。

联邦学习中的客户端可能遭受恶意攻击而上传恶意模型更新，破坏全局模型聚合过程，而区块链的审计性与可靠性结合联邦学习具有广阔的研究价值。Preuveneers 等[76]提出了一个基于区块链的联邦学习模型审计方案，客户端上传的模型更新需要进行异常检测并链接到分布式分类账上，检测结果大于预定义阈值的客户端将被问责，同时为避免中心服务器的单点故障问题，区块链被用于替换中心服务器与客户端的直接交互，因此联邦学习中的每个节点都拥有一个完整的分类账副本，并且可以计算汇总的权重更新。Zhu 等[77]引入区块链技术管理联邦学习的安全问题，建立安全的协同训练机制来检测客户端的可靠性。实验结果表明，当拜占庭故障设备是客户端成员的一部分时，该方案具有明显的优势。Qu 等[64]提出 FL-Block 方案，利用区块链的性质与其提出的增强协议能够有效抵御中毒攻击，其要求矿工在记录模型更新到区块链之前先验证其正确性，然后将模型更新存储在与其关联的候选块中。Liu 等[80]在区块链中基于智能合约交互模型更新，以自动验证模型更新防御恶意和不可靠的参与者，同时引入本地差分隐私技术，防止成员推理攻击，实现了 5G 网络中的隐私安全 FL。

联邦学习与区块链的结合使系统成为一个完善的闭环学习机制。一方面，联邦学习技术能够为具有隐私数据的参与方提供跨域安全共享方案；另一方面，区块链技术作为核心数据库为参与方提供了安全存储、信任管理、细粒度区分和激励回报等应用需求，促使拥有数据的用户积极参与到数据联邦中。

5　研究热点和前景展望

5.1　研究热点

不同于传统的分布式机器学习技术，海量客户端与 Non-IID 数据集对联邦学习提出了新的挑战。目前，学术界对于联邦学习的研究十分活跃，研究者可能不仅需要掌握机器学习技术，还需要掌握分布式算法优化、密码学、压缩量化、信息论、统计等技术[80]。本文介绍了联邦学习在通信效率、隐私安全、信任与激励机制等方向上的研究进展，但仍有一些其他研究方向值得探索。

（1）系统异构。在联邦学习环境中，由于参与训练的客户端之间硬件配置、网络带宽、电池容量等不同，各终端设备的计算能力、通信速度和存储能力各不相同[81]。除此之外，联邦学习架构通常会限制终端设备参与训练的数量，尤其是在数百万设备参与的训练中，处于活跃状态的往往只有数百个客户端。每个客户端并不一定可靠，随时可能因为网络故障、算力限制等问题退出现有训练，这些系统级别的异构会给模型整体效能造成极大的挑战。因此，适用于系统异构的联邦学习算法必须满足 3 点要求：客户端的低参与率；兼容不同的硬件结构；能够容忍训练设备的中途退出。

（2）统计异构。不同的终端设备通常使用各式各样的方式生成、存储和传输数据，因此各设备之间数据的特征和体量可能有很大的不同，导

致数据呈 Non-IID 分布和非平衡分布。尽管这类分布的数据集可以通过通信效率优化的方式处理，但仍然存在一些针对统计异构的解决方法，如通过多任务学习框架学习不同的局部模型[82]。类似于元学习，多任务学习由于对个性化和特定于设备建模的支持，已经成为解决数据统计异构性的主流方法。

（3）无线通信。在 5G 技术日益普及的今天，联邦学习开始被逐渐应用于无线网络领域。由于无线信道的带宽容量有限，因此在发送信息之前，需要对模型更新进行量化压缩，在这种模式下，一个重要的考虑因素是存在量化误差时模型更新的鲁棒性。除了通信带宽外，无线通信中复杂的噪声和干扰也是加剧信道瓶颈的因素[83]。因此，开发适用于无线通信的联邦学习算法具有突出的研究意义[84]。

除了对联邦学习本身技术的改进，最新的研究进展包括结合边缘计算在物联网领域的应用[58,85-87]，如图 8 所示。由于部分终端设备并没有足够的计算资源，同时为了满足智能决策的低时延响应，边缘计算在云中心和边缘设备之间添加了边缘服务器作为中介层，联邦学习作为其“操作系统”满足了智能边缘设备实时决策、多点协同、自主可控的要求。充分利用智能边缘服务器计算、存储、传输能力，改变传统集中上传数据进行决策的方式，破解了传统集中式机器学习数据难以聚合、隐私难以保护、云中心的单点故障等问题，为未来多功能集群、跨多智能设备的实时安全决策提供了可靠的技术保障。

5.2 前景展望

在大数据时代，如何在保障数据安全和隐私的前提下，实现数据共享，促进多源数据的碰撞、融合，最大限度地释放数据价值，成为学术界和产业界面临的挑战之一。而联邦学习作为应对该挑战的一项新兴技术，在诸多领域具有广阔的应用前景。

（1）边缘计算和物联网。随着智能手机和移动互联网的普及应用，大量数据产生在设备的边缘端，移动边缘计算使计算发生在本地设备，而不需要将隐私数据发送到云端。而联邦学习作为边缘计算的操作系统，提供了一种各方协作与共享的协议规范，它能够让边缘设备在不向云端设备发送源数据的情况下，合作训练出一个最优的全局机器学习模型。未来，随着物联网的进一步发展，人工智能和边缘计算将朝着一体化的方向大步向前。

（2）智慧医疗。为了降低人工成本和减少人为操作失误的可能，机器学习技术开始越来越多地应用在医疗领域，用于提升医疗诊治的效率和准确率。但是，由于医疗机构的数据对于隐私和安全的敏感性，医疗数据中心很难收集到足够数量的、特征丰富的、可以全面描述患者症状的数据，而性能良好的机器学习模型往往需要来自多个数据源，包括医疗报告、病例特征、生理指标、

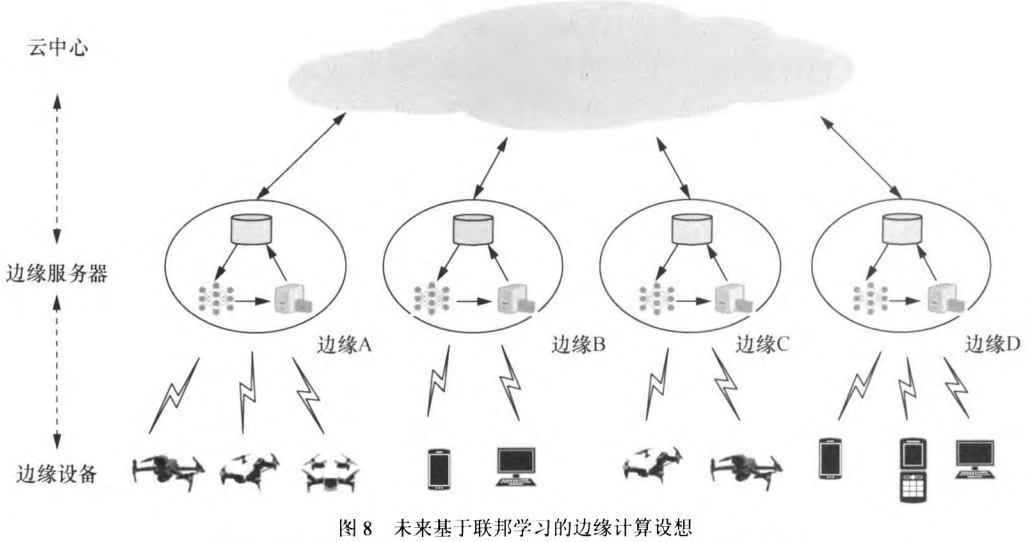


图 8 未来基于联邦学习的边缘计算设想

Figure 8 Future vision of edge computing based on federated learning

基因序列等。联邦迁移学习是解决这类问题的有效方法，无须交换各医疗机构的私有数据，协同所有的训练参与方训练一个共享模型，同时迁移学习技术可以扩展训练数据的样本空间和特征空间，有效降低各医疗机构之间样本分布的差异性。

（3）金融风控。为了维持金融行业稳定、风险控制和防止金融诈骗，银行和金融企业都希望利用人工智能技术为客户提供有效且安全的金融服务。在实际应用中，对客户“肖像”特征的描述通常包括资质信息、购买能力、购买偏好及商品特征等，而这些信息分别分布在银行、电子商务平台和用户的私人社交网络中。出于隐私安全的考虑，将三方数据聚合并不现实，而联邦学习为构建跨企业、跨数据平台以及跨领域的大数据和 AI 系统提供了良好的技术支持。

（4）智慧城市。随着人工智能、物联网和 5G 技术的发展，智慧城市的概念已经跃然纸上。然而，在城市的不同信息部门中，如后勤、应急、维稳、安保等，会产生大量的异构数据，形成多个数据孤岛，无法整合利用。联邦学习的异构数据处理能力能够帮助人们创造迅速响应市民需求的智慧城市，解决数据“孤岛”问题，同时基于智慧城市构建的机器学习模型为企业提供个性化服务带来了更多的机遇[79]。

（5）涉密数据的安全共享。大数据环境背景下，数据的安全交换显得尤为敏感。常规共享交换使多部门数据汇集的方法，极有可能导致权限难以控制、责任划分不清、问题难以追责，甚至造成失泄密等重大安全事故。如何解决涉密数据的安全共享难题，联邦学习技术的跨域共享特性使各部门之间无须汇集数据即可实现敏感数据的跨域安全共享。

6　结束语

本文介绍了联邦学习技术概念、算法原理与分类，并对目前联邦学习中的 3 个痛点问题的研究进展做出归纳总结，最后展望了联邦学习在各领域的发展前景。随着社会对于隐私安全的日益重视，政府正在逐步加强对私人数据的管控，传统的机器学习模式可能不再符合安全法规。联邦学习作为下一代人工智能大规模协作的基础理论，为目前发展人工智能面临的小数据和隐私等关键问题提供了有效的解决思路。同时，对于联邦学习的国际标准在积极制定中，未来建立在统一标准下的联邦学习必然能够更好地应用于各行各业，发挥更大的效能，进一步推动网络安全的发展[3]。

参考文献：

[1] 微众银行 AI 项目组. 联邦学习白皮书 V1.0[R]. 2018.
WeBank AI Project Team. Federated learning white paper V1.0 [R]. 2018.

[2] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[J]. Artificial Intelligence and Statistics, 2017: 1273-1282.

[3] 杨强. AI 与数据隐私保护：联邦学习的破解之道[J]. 信息安全研究, 2019, 5(11): 961-965.
YANG Q. AI and data privacy protection: the cracking method of federated learning[J]. Information Security Research. 2019, 5(11): 961-965.

[4] 潘碧莹, 丘海华, 张家伦. 不同数据分布的联邦机器学习技术研究[M]. 5G 网络创新研讨会 (2019) 论文集, 2019.
PAN B Y, QING H H, ZHANG J L. Research on federal machine learning technology with different data distribution[M]. 5G Network Innovation Seminar (2019) Proceedings, 2019.

[5] YANG Q, LIU Y, CHEN T, et al. Federated machine learning: Concept and applications[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2019, 10(2): 1-19.

[6] HARD A, RAO K, MATHEWS R, et al. Federated learning for mobile keyboard prediction[J]. arXiv preprint arXiv:1811.03604, 2018.

[7] LI S, CHENG Y, LIU Y, et al. Abnormal client behavior detection in federated learning[J]. arXiv preprint arXiv:1910.09933, 2019.

[8] BONAWITZ K, IVANOV V, KREUTER B, et al. Practical secure aggregation for privacy-preserving machine learning[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017: 1175-1191.

[9] MANDAL K, GONG G, LIU C. NIKE-based fast privacy-preserving high-dimensional data aggregation for mobile devices[R]. CACR Technical Report, 2018.

[10] CHENG K, FAN T, JIN Y, et al. Secureboost: a lossless federated learning framework[J]. arXiv preprint arXiv:1901.08755, 2019.

[11] KONEČNÝ J, MCMAHAN H B, YU F X, et al. Federated learning: strategies for improving communication efficiency[J]. arXiv preprint arXiv:1610.05492, 2016.

[12] WANG L P, WANG W, LI B. CMFL: mitigating communication overhead for federated learning[C]//2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2019: 954-964.

[13] CALDAS S, KONEČNY J, MC-MAHAN H B, et al. Expanding the reach of federated learning by reducing client resource requirements[J]. arXiv preprint arXiv:1812.07210, 2018.

[14] BHOWMICK A, DUCHI J, FREUDIGER J, et al. Protection

against reconstruction and its applications in private federated learning[J]. arXiv preprint arXiv:1812.00984, 2018.

[15] CARLINI N, LIU C, KOS J, et al. The secret sharer: Measuring unintended neural network memorization & extracting secrets[J]. arXiv preprint arXiv:1802.08232, 2018.

[16] FREDRIKSON M, LANTZ E, JHA S, et al. Privacy in pharmacogenetics: an end-to-end case study of personalized warfarin dosing[C]//23rd {USENIX} Security Symposium ({USENIX} Security 14). 2014: 17-32.

[17] MELIS L, SONG C, DE-CRISTOFARO E, et al. Exploiting unintended feature leakage in collaborative learning[C]//2019 IEEE Symposium on Security and Privacy (SP). 2019: 691-706.

[18] HITAJ B, ATENIESE G, PEREZ-CRUZ F. Deep models under the GAN: information leakage from collaborative deep learning[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017: 603-618.

[19] GEI P J, BAUERMEISTER H, DRGE H, et al. Inverting gradients - how easy is it to break privacy in federated learning[R]. 2020.

[20] SONG M, WANG Z, ZHANG Z, et al. Analyzing user-level privacy attack against federated learning[J]. IEEE Journal on Selected Areas in Communications, 2020.

[21] GEYER R C, KLEIN T, NABI M. Differentially private federated learning: a client level perspective[J]. arXiv preprint arXiv:1712.07557, 2017.

[22] MC MAHAN H B, RAMAGE D, TALWAR K, et al. Learning differentially private recurrent language models[J]. arXiv preprint arXiv:1710.06963, 2017.

[23] BONAWITZ K, IVANOV V, KREUTER B, et al. Practical secure aggregation for privacy-preserving machine learning[C]//Proceedings of the 2017 ACM Sigsac Conference on Computer and Communications Security. 2017: 1175-1191.

[24] MANDAL K, GONG G, LIU C. NIKE-based fast privacy-preserving high-dimensional data aggregation for mobile devices[R]. CACR Technical Report, 2018.

[25] CHENG K, FAN T, JIN Y, et al. Secureboost: a lossless federated learning framework[J]. arXiv preprint arXiv:1901.08755, 2019.

[26] 袁勇, 王飞跃. 区块链技术发展现状与展望[J]. 电子学报, 2016, 42(4): 481-494.
YUAN Y, WANG F Y. Blockchain: the state of the art and future trends[J]. Acta Automatica Sinica, 2016, 42(4): 481-494.

[27] YANG R, YU F R, SI P, et al. Integrated blockchain and edge computing systems: a survey, some research issues and challenges[J]. IEEE Communications Surveys & Tutorials, 2019, 21(2): 1508-1532.

[28] DONG Y, CHEN X, SHEN L, et al. EaSTFLy: efficient and secure ternary federated learning[J]. Computers & Security, 2020, 94: 1-15.

[29] CAO H, LIU S, ZHAO R, et al. IFed: a novel federated learning framework for local differential privacy in power internet of things[J]. International Journal of Distributed Sensor Networks, 2020, 16(5): 1-3.

[30] CHEN J, PAN X, MONGA R, et al. Revisiting distributed synchronous SGD[J]. arXiv preprint arXiv:1604.00981, 2016.

[31] SATTLER F, WIEDEMANN S, MÜLLER K R, et al. Robust and communication-efficient federated learning from Non-IID data[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019.

[32] XIAO P, CHENG S, STANKOVIC V, et al. Averaging is probably not the optimum way of aggregating parameters in federated learning[J]. Entropy, 2020, 22(3): 314.

[33] LI T, SAHU A K, ZAHEER M, et al. Federated optimization in heterogeneous networks[J]. arXiv preprint arXiv:1812.06127, 2018.

[34] LIU W, CHEN L, CHEN Y, et al. Accelerating federated learning via momentum gradient descent[J]. IEEE Transactions on Parallel and Distributed Systems, 2020, 31(8): 1754-1766.

[35] HUANG L, YIN Y, FU Z, et al. LoAdaBoost: loss-based AdaBoost federated machine learning with reduced computational complexity on IID and non-IID intensive care data[J]. PLoS ONE, 2020 15(4): 1-6.

[36] JIANG J, HU L, HU C, et al. BACombo—bandwidth-aware decentralized federated learning[J]. Electronics, 2020, 9(3): 440.

[37] REISIZADEH A, MOKHTARI A, HASSANI H, et al. Fedpaq: a communication-efficient federated learning method with periodic averaging and quantization[C]//International Conference on Artificial Intelligence and Statistics. 2020: 2021-2031.

[38] LU X, LIAO Y, LIO P, et al. Privacy-preserving asynchronous federated learning mechanism for edge network computing[J]. IEEE Access, 2020, 8: 48970-48981.

[39] HE L, BIAN A, JAGGI M. Cola: decentralized linear learning[C]//Advances in Neural Information Processing Systems. 2018: 4536-4546.

[40] LALITHA A, WANG X, KILINC O, et al. Decentralized Bayesian learning over graphs[J]. arXiv preprint arXiv:1905.10466, 2019.

[41] LIN T, STICH S U, PATEL K K, et al. Don't use large mini-batches, use local SGD[J]. arXiv preprint arXiv:1808.07217, 2018.

[42] LIU L, ZHANG J, SONG S H, et al. Edge-assisted hierarchical federated learning with non-iid data[J]. arXiv preprint arXiv:1905.06641, 2019.

[43] SHARMA P K, PARK J H, CHO K. Blockchain and federated learning-based distributed computing defence framework for sustainable society[J]. Sustainable Cities and Society, 2020: 102220.

[44] ZHANG J, ZHAO Y, WANG J, et al. FedMEC: improving efficiency of differentially private federated learning via mobile edge computing[J]. Mobile Networks and Applications, 2020: 1-13.

[45] 刘俊旭, 孟小峰. 机器学习的隐私保护研究综述[J]. 计算机研究与发展, 2020, 57(2): 346.
LIU J X, MENG X F. A survey of research on privacy protection in machine learning[J]. Computer Research and Development, 2020, 57(2): 346.

[46] DWORK C, MC-SHERRY F, NISSIM K, et al. Calibrating noise to sensitivity in private data analysis[C]//Theory of Cryptography Conference. 2006: 265-284.

[47] 叶青青, 孟小峰, 朱敏杰, 等. 本地化差分隐私研究综述[J]. 软件学报, 2018, 29(7): 159-183.
YE Q Q, MENG X F, ZHU M J, et al. Survey of localized differential privacy research[J]. Journal of Software, 2018, 29(7): 159-183.

[48] 苏冠通, 徐茂桐. 安全多方计算技术与应用综述[J]. 信息通信技术与政策, 2019 (5): 19-22.
SU G T, XU M T. Survey of secure multiparty computing technology and application[J]. Information and Communication Technol-

ogy and Policy, 2019 (5): 19-22.

[49] DOLEV D, YAO A. On the security of public key protocols[J]. IEEE Transactions on Information Theory, 1983, 29(2): 198-208.

[50] RIVEST R L, SHAMIR A, ADLEMAN L. A method for obtaining digital signatures and public-key cryptosystems[J]. Communications of the ACM, 1978, 21(2): 120-126.

[51] THAKKAR O, ANDREW G, MC-MAHAN H B. Differentially private learning with adaptive clipping[J]. arXiv preprint arXiv:1905.03871, 2019.

[52] HAO M, LI H, LUO X, et al. Efficient and privacy-enhanced federated learning for industrial artificial intelligence[J]. IEEE Transactions on Industrial Informatics, 2019, 16(10): 6532-6542.

[53] AONO Y, HAYASHI T, WANG L, et al. Privacy-preserving deep learning via additively homomorphic encryption[J]. IEEE Transactions on Information Forensics and Security, 2017, 13(5): 1333-1345.

[54] LIU X, LI H, XU G, et al. Adaptive privacy-preserving federated learning[J]. Peer to Peer Networking and Applications, 2020, 13: 2356-2366.

[55] HUANG X, DING Y, JIANG Z L, et al. DP-FL: a novel differentially private federated learning framework for the unbalanced data[J]. World Wide Web, 2020: 1-17.

[56] WEI K, LI J, DING M, et al. Federated learning with differential privacy: algorithms and performance analysis[J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 3454-3469.

[57] XU G, LI H, LIU S, et al. VerifyNet: secure and verifiable federated learning[J]. IEEE Transactions on Information Forensics and Security, 2019, 15: 911-926.

[58] LU Y, HUANG X, DAI Y, et al. Differentially private asynchronous federated learning for mobile edge computing in urban informatics[J]. IEEE Transactions on Industrial Informatics, 2019.

[59] LU Y, HUANG X, DAI Y, et al. Federated learning for data privacy preservation in vehicular cyber-physical systems[J]. IEEE Network, 2020, 34(3): 50-56.

[60] HU R, GUO Y, LI H, et al. Personalized federated learning with differential privacy[J]. IEEE Internet of Things Journal, 2020, 7(10): 9530-9539.

[61] FANG M, CAO X, JIA J, et al. Local model poisoning attacks to Byzantine-robust federated learning[J]. arXiv preprint arXiv:1911.11815, 2019.

[62] ZHAO Y, CHEN J, ZHANG J, et al. Detecting and mitigating poisoning attacks in federated learning using generative adversarial networks[J]. Concurrency and Computation: Practice and Experience, 2020: 1-2.

[63] KANG J, XIONG Z, NIYATO D, et al. Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory[J]. IEEE Internet of Things Journal, 2019, 6(6): 10700-10714.

[64] FUNG C, YOON C J M, BESCHASTNIKH I. Mitigating sybils in federated learning poisoning[J]. arXiv preprint arXiv:1808.04866, 2018.

[65] CHEN Y, LUO F, LI T, et al. A training-integrity privacy-preserving federated learning scheme with trusted execution environment[J]. Information Sciences, 2020, 522: 69-79.

[66] LYU L, YU J, NANDAKUMAR K, et al. Towards fair and privacy-preserving federated deep models[J]. IEEE Transactions on Parallel and Distributed Systems, 2020, 31(11): 2524-2541.

[67] MAJEED U, HONG C S. FLchain: federated learning via MEC-enabled blockchain network[C]//2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS). 2019: 1-4.

[68] ARACHCHIGE P C M, BERTOK P, KHALIL I, et al. A trustworthy privacy preserving framework for machine learning in industrial iot systems[J]. IEEE Transactions on Industrial Informatics, 2020, 16(9): 6092-6102.

[69] LU Y, HUANG X, ZHANG K, et al. Blockchain empowered asynchronous federated learning for secure data sharing in internet of vehicles[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4): 4298-4311.

[70] POKHREL S R. Towards efficient and reliable federated learning using blockchain for autonomous vehicles[J]. Computer Networks, 2020: 107431.

[71] KIM H, PARK J, BENNIS M, et al. On-device federated learning via blockchain and its latency analysis[J]. arXiv preprint arXiv:1808.03949, 2018.

[72] KANG J, XIONG Z, NIYATO D, et al. Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory[J]. IEEE Internet of Things Journal, 2019, 6(6): 10700-10714.

[73] WENG J, WENG J, ZHANG J, et al. Deepchain: auditable and privacy-preserving deep learning with blockchain-based incentive[J]. IEEE Transactions on Dependable and Secure Computing, 2019.

[74] KIM Y J, HONG C S. Blockchain-based node-aware dynamic weighting methods for improving federated learning performance[C]//20th Asia-Pacific Network Operations and Management Symposium (APNOMS). 2019: 1-4.

[75] ZHAN Y, LI P, QU Z, et al. A learning-based incentive mechanism for federated learning[J]. IEEE Internet of Things Journal, 2020, 7(7): 6360-6368.

[76] PREUVENEERS D, RIMMER V, TSINGENOPOULOS I, et al. Chained anomaly detection models for federated learning: an intrusion detection case study[J]. Applied Sciences, 2018, 8(12): 2663.

[77] ZHU X, LI H, YU Y. Blockchain-based privacy preserving deep learning[C]//International Conference on Information Security and Cryptology. 2018: 370-383.

[78] QU Y, GAO L, LUAN T H, et al. Decentralized privacy using blockchain-enabled federated learning in fog computing[J]. IEEE Internet of Things Journal, 2020, 7(6): 5171-5183.

[79] LIU Y, PENG J, KANG J, et al. A secure federated learning framework for 5G networks[J]. arXiv preprint arXiv:2005.05752, 2020.

[80] KAIROUZ P, MC-MAHAN H B, AVENT B, et al. Advances and open problems in federated learning[J]. arXiv preprint arXiv:1912.04977, 2019.

[81] LI T, SAHU A K, TALWALKAR A, et al. Federated learning: challenges, methods, and future directions[J]. arXiv preprint arXiv:1908.07873, 2019.

[82] SMITH V, CHIANG C K, SANJABI M, et al. Federated multi-task learning[C]//Advances in Neural Information Processing Systems. 2017: 4424-4434.

[83] ANG F, CHEN L, ZHAO N, et al. Robust federated learning with noisy communication[J]. IEEE Transactions on Communications, 2020.

[84] NIKNAM S, DHILLON H S, REED J H. Federated learning for wireless communications: motivation, opportunities, and challenges[J]. IEEE Communications Magazine, 2020, 58(6): 46-51.

[85] REN J, WANG H, HOU T, et al. Federated learning-based computation offloading optimization in edge computing-supported internet of things[J]. IEEE Access, 2019, 7: 69194-69201.

[86] WANG X, HAN Y, WANG C, et al. In-edge AI: intelligentizing mobile edge computing, caching and communication by federated learning[J]. IEEE Network, 2019, 33(5): 156-165.

[87] WANG S, TUOR T, SALONIDIS T, et al. Adaptive federated learning in resource constrained edge computing systems[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(6): 1205-1221.

[作者简介]

周传鑫（1997– ），男，安徽蚌埠人，信息工程大学硕士生，主要研究方向为数据安全交换、机器学习和隐私保护。

孙奕（1979– ），女，河南郑州人，博士，信息工程大学副教授，主要研究方向为网络与信息安全、数据安全交换。

汪德刚（1996– ），男，陕西安康人，信息工程大学硕士生，主要研究方向为数据安全交换、恶意流量检测。

葛桦玮（1998– ），男，浙江临海人，主要研究方向为数据安全交换。