面向能源受限的物联网设备的联邦学习：系统映射研究

Rachid EL Mokadem、Yann Ben Maissa 和 Zineb El Akkaoui

{elmokadem.rachid, benmaissa, elakkaoui}@inpt.ac.ma

电信系统、网络与服务实验室，国家邮政与电信研究所，拉巴特，10587，摩洛哥。

摘要

联邦机器学习（Fed ML）是一种新型分布式机器学习技术，通过协作使用客户端本地数据训练全局模型，而无需传输这些数据。节点只发送参数更新（例如神经网络中的权重更新），服务器将其融合以构建全局模型。由于不泄露节点数据，Fed ML 可确保数据的机密性；这是网络安全的关键方面，使其能够应用于对数据敏感的物联网（IoT）和移动应用场景，例如智能地理定位和智能电网。然而，大多数物联网设备都受到能源限制，因此有必要优化 Fed ML 流程，以实现高效的训练任务和优化的功耗。本文据我们所知，首次针对能源受限的物联网设备上的 FedML 优化技术开展系统映射研究（SMS）。在总计 800 多篇论文中，我们筛选出 67 篇符合标准的论文，并通过一组精心设计的研究问题，对该领域进行结构化概述。最后，我们尝试分析能源受限场景下 Fed ML 的研究现状，并为研究界提出一些潜在建议。

关键词： 联邦　　机器学习、能耗优化、物联网、边缘计算与移动计算、端侧智能

1 引言

背景。机器学习（ML）已成为不同应用领域中一种重要且日益普及的范式。在过去十年中，物联网计算机系统及其潜在应用（例如智慧城市、智能电网）取得了长足发展，使其能够在如此庞大而复杂的环境中受益于机器学习的能力。此外，物联网在工业界和学术界的广泛应用（例如通过树莓派和 Arduino 等快速原型开发平台）也使人们对广泛关键应用中的数据隐私保护和高效资源利用产生了更高期待。因此，鉴于机器学习在分布式系统和敏感数据方面的局限性，McMahan 等人提出了联邦机器学习（联邦机器学习，简称 Fed ML）以应对这些限制。该方法将模型训练任务委派给客户端设备，由它们协作构建全局共享模型，在整合各自的本地数据学习成果的同时，避免任何私人数据离开其原始设备 [32]。自这篇开创性论文发表以来，联邦机器学习已成为机器学习领域的“热门话题”之一。

问题。物联网和移动设备面临与能源来源相关的重大约束，因此必须针对分配的任何任务优化这些设备的功耗。特别是，机器学习算法被认为是一种高度耗能的多任务过程 [17]。在分布式机器学习环境中，节点必须

持续与主节点交换数据，这可能会增加系统的开销成本。联邦机器学习试图通过限制交换数据仅包含本地模型权重[46]，（由节点训练），而不是交换大量原始数据，来解决这一问题。同时，为应对与物联网和移动设备特征相关的其他关键挑战，联邦机器学习仍需进一步改进，即资源受限和能源约束[54]。因此，过去几年中，科学界提出了越来越多针对这些方面的联邦机器学习研究。鉴于这一不断发展的文献体系，亟需开展一项综合性研究，以全面概述能耗优化方法，并为研究界提出新的研究方向。

贡献。一些研究实际上已经着手解决最初联邦机器学习方案（称为 FedAvg）的局限性，并提出了许多优化方法，主要涉及通信负载、数据交换和其他方面，这些方法能够直接或间接地应对有限的能量约束。据我们所知，本文旨在针对能量受限设备上的联邦机器学习优化开展首个系统映射研究（SMS）。本系统映射研究的任务包括：i）基于五个研究问题，对该主题已发表的相关主要研究进行统计和分类；ii）分析并讨论研究结果，为研究界清晰了解近期改进提供依据；以及 iii）协助工程师为物联网和移动设备开发创新的联邦机器学习解决方案。

目录。本文其余部分安排如下：首先，我们在第2节介绍相关研究和综述。然后，在第3节中，我们通过原始 FedAvg 算法介绍一些理论基础，并阐述能源优化问题的表述。在第4节中，我们介绍本研究采用的方法，包括论文筛选与筛选过程，以及研究问题（RQs）。

表1 相关研究

Year	标题	Type	重点
2020	系统性文献综述 on 联邦 机器 Lea习：F rom A 软件 工程 视角。[41]	SLR	软件 engineering 方面
2020	关于联邦学习的系统性文献综述 学习：从模型质量的角 度。 [40]	SLR	模型 质量
2020	移动边缘网络中的联邦学习：一项 全面综述。[38]	综述	通用
2020	联邦学习综述 ：赋能技术、协议与应用。[1]	综述	应用
2020	隐私保护综述 联邦学习 物联网。[6]	调查	隐私

在第5节中，我们回答这些问题，并分析从所研究论文中得出的结果。随后，我们在第6节中展开讨论，并针对研究方向提出一些建议。第7节列出了本研究有效性面临的一些威胁。最后，我们在第8节总结全文，并概述一些可能的未来研究工作。

2 相关研究

本节介绍了三篇调查研究和两篇与本研究相关的文献综述（见表1）。[41]从软件工程的角度对联邦学习进行了系统性文献综述，概述了联邦学习系统的整体情况，重点关注软件开发方面以及实际应用中的一般挑战。另一方面，[40]，从模型质量的角度对联邦学习进行了系统性文献综述，研究了提升联邦机器学习模型和数据质量的方法。此外，作者还比较了在相同数据上采用联邦学习和非联邦学习所得到的模型。此外，[38]针对移动边缘网络中的联邦学习开展了一项调查研究，探讨了良好性能的特征和局限性、资源分配、通信成本以及数据隐私问题。此外，[1]对联邦机器学习的赋能技术、协议和应用进行了调查研究。他们

提供了最相关的协议、平台，以及联邦学习在现实生活中的应用案例，帮助数据科学家为各行业构建更优的隐私保护解决方案；他们还探讨了联邦机器学习在实际应用中的挑战与优势。最后，[5] 从隐私保护的角度对联邦学习进行了调查研究。

尽管这些综述和系统性文献综述都很出色，但我们认为，本研究探讨了一些尚未被它们直接涉及的方面。除了 [38]，之外，它们并未关注联邦学习优化中的能量因素，而 [38]，对这一问题的探讨也不够深入。我们尝试阐明 FedML 在物联网中的功耗问题。思科在 [9]，中报告称，到 2022 年，物联网连接将占全球所有联网设备和连接总数（285 亿）的半数以上（146 亿），这表明物联网在人们生活中的普及程度日益提高 [14]。我们日常使用智能手机和智能手表时，需要频繁充电，有时甚至会觉得颇为不便，这也是这一问题在现实中的体现。最后，无线传感器网络也有其特殊情况：它们可能部署在恶劣环境中，完全无法补充能量。

3 背景

本节将介绍联邦学习的整体流程、FedAvg 算法、能耗问题以及其他一些背景信息。

3.1 联邦学习

联邦机器学习是通过结合小型设备的计算能力和本地数据[46]，在由这些设备组成的大规模分布式系统上开发准确模型的过程。其目标是解决单台中央计算机无法解决的一类问题，例如涉及用户个人数据、实时计算和端侧人工智能的问题[32]。

FedML 基于分布式架构，由多个节点在各自的本地数据上执行训练任务，并交换其模型的参数，并由中央服务器进行处理。服务器随后根据本地模型构建全局聚合模型，该模型等同于使用所有节点汇总后的数据训练得到的模型。对于 FedAvg，全局模型W_g是通过对本地模型W_i进行加权平均构建的（见公式1）。

$$\mathrm{Wg} = \sum_i \frac{n_i}{n} \mathrm{W_i} \quad (1)$$

全局目标函数f的优化可以表示为对所有参与节点$i = 1,\ldots,n_i$的局部目标函数f_i的平均值进行优化，如方程2 [54]所示。

$$\min_w \mathrm{f(w)} = \min_w \frac{1}{n} \sum_i \mathrm{f_i(w)} \quad (2)$$

其中：

$$f_i(w) \coloneqq \frac{1}{k} \sum_{\xi \in D_i} l(w, \xi) \quad (3)$$

f_i定义为每个节点i在其本地样本点上局部损失函数l的平均值，$D_i = \xi_{i1},\cdots,\xi_{im}$适用于$i \in [n],$，其中$D_i$是节点$i$的本地数据集，由$m$个数据点组成；$\xi_i$而$w$是模型参数。

$$min_w f(w) \coloneqq min_w \frac{1}{nk} \sum_{\xi \in D} l(w, \xi) \quad (4)$$

最后，为求解方程4，每个节点都使用梯度下降法，在其本地训练数据D_i上最小化损失l_i，最终聚合后的模型W_g将使全局目标函数达到最小值。

3.1.1 联邦学习伪算法

算法1 展示了联邦平均（FedAvg），由[46] 为联邦机器学习提出。

/* 在服务器上运行*/

初始化 w_0;

对 每一轮次 $t = 1,2, \ldots$ **执行**

$m \leftarrow$

$S_t \leftarrow$ 随机选取m个客户端); (

对 每个客户端 $k \in S_t$ 并行 **执行**

$w_{t+1} \leftarrow$ 客户端更新 (k, w_t);

$w_{t+1} \leftarrow \mathrm{P}\ _{K\,k=1}\ n_k \underline{n}\ w_{k\,t+1}$;

end

max(C.K,1);

end

/* 在客户端 k 上运行*/

函数 客户端更新(k,w): $B \leftarrow$(分割 P_k 分成大小为**B**的批次）；**对于**每个本地第 i 个 *epoch*，从 *1* 到 E **执行**

对于批次 $b \in B$ **执行**

$w \leftarrow w - \mu \nabla l(w,b)$;

end

end

返回w 至服务器;

end

算法 1： FedAvg 伪算法

该算法中使用的符号说明如下。

- **C** 每轮选取的客户端比例 • **K** 客户端总数 • **m** 每轮随机选取的客户端数量 • S_t 每轮的客户端集合 • w_t 第 t 轮的全局模型参数

- w_t^k 第 t 轮从客户端 k 接收的模型参数 • n_k 客户端 k 的数据点数量 • n 所有客户端的数据点总数 • P_k 客户端 k 的本地数据集 • **B** 客户端训练所用的本地数据集小批量大小 • B 用于本地训练的数据集小批量集合

- **E** 每个客户端在向服务器发送更新前执行的训练轮次数。
- μ 学习率 • l 损失函数 • w 本地模型参数

如上述算法所示，服务器初始化模型参数 w_0，然后在每个轮次中，确定要选取的数量 **m** ，即从总共 **K** 个客户端中按比例 **C** 选取参与训练的客户端。设备子集 S_t由随机方式确定，然后每个客户端设备 **k** 从服务器接收模型参数 w_t，以便在各自的

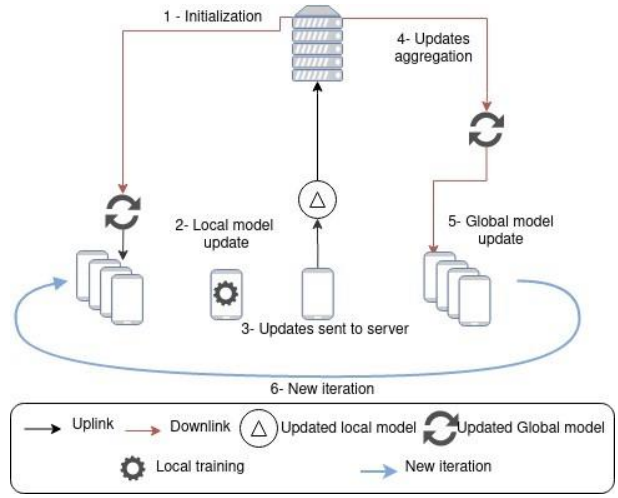


图 1 联邦学习全局示意图

本地数据集 P_k。此训练过程会将本地数据划分为大小为 **B** 的小批量，并执行若干轮 **E** 本地训练轮次来训练本地模型。最后，所有选中的客户端都会计算参数 w 的更新，并将其发送回服务器，由服务器对这些更新取平均值，从而得到新的全局模型参数 w_{t+1}。服务器会决定重复此轮次多次，直至达到目标性能。

3.1.2 联邦学习流程

联邦机器学习架构由客户端节点和中央服务器组成。服务器接收来自客户端设备的计算得到的更新，并执行聚合操作以构建全局模型。随后，通过在节点上运行额外的迭代来训练其本地模型，使其持续改进，直至获得理想结果。

图1 总体展示了 联邦机器学习 架构中涉及的组件，以及 FedAvg 算法的执行阶段。在每一轮训练中，都会执行以下操作：

1. **定义模型结构、随机初始化参数并选择参与设备**：中央服务器必须在训练开始前定义参数 **E**、**C** 和 **B**，并且必须选择一部分客户端参与每一轮训练。

2. **基于本地数据的模型更新**：每个选定的客户端都会按照中央服务器设定的次数执行本地训练迭代，以计算全局模型的更新。

3. **向服务器传输本地模型更新**：每个参与设备都会发送计算出的模型更新。4. **聚合所有接收的模型更新**：服务器聚合接收的更新，从而构建全局模型。5. **与设备共享更新后的全局模型**。

3.1.3 异质性

在实际应用中，联邦学习的参与节点通常拥有不均衡的资源和训练数据，我们分别将其称为系统异构性 和统计异质性[35]。

系统异构性

在全局模型的协同训练过程中，不同节点的能力各不相同（例如 CPU、电池、内存、带宽）。因此，如果忽略这一事实，收敛速度会非常慢，而且能力较弱的客户端会在训练结束前耗尽资源，导致模型性能不佳。

统计异质性

FedAvg 最初由 [46], 提出时，基于以下假设：各节点的数据 独立同分布（iid），这通过使用梯度下降优化方法，为关于均衡本地数据集 4 的方程 D_i提供了理论解。然而，对于物联网和用户设备上的大多数分布式数据而言，这一假设并不成立；这是联邦机器学习**在现实场景** 中部署 [76] 所面临的一大限制因素。事实上，该主题下发表的大多数研究表明，算法在独立同分布数据上表现良好，但在非独立同分布设置下表现较差，这会导致对全局模型性能以及训练所需时间和能量产生不利影响 [76]。这一重大问题促使多个团队开发技术，以使原始联邦学习算法能够适应这两种异质性[11, 34, 70]。

3.2 能耗建模

面向能量受限设备的 FedML 优化，主要目标是在构建优质全局模型的同时，尽可能降低节点的功能性能耗。通常，无线设备的总能耗 ET 可分为三个主要部分：Enet、Ec 和 Esys（公式 5）。

$$E_T = E_{net} + E_c + E_{sys} \quad (5)$$

E_{net}是设备与其他设备或服务器进行更新交换时通信所消耗的能量。Ec 是设备的本地处理单元和内存完成训练计算所消耗的能量。E_{sys}是设备执行常规系统操作所消耗的能量，与其参与联邦训练无关。

请注意，与物联网中的总能耗相比，E_{sys}通常很小，可以忽略不计。此外，它并非本研究所考虑问题的特定部分，因此我们在此建模中将其省略。

此外，在操作量相同的情况下，通信通常比处理消耗更多能量（这也说明了在数据传输前采用多种聚合方法的合理性）。公式 6 给出了网络通信所消耗的能量，其表示形式由一组与本文背景相关的参数构成。

$$E_{net} \simeq \sum_{i=1}^{N_T} N_{Tbit}^{i} \frac{P_T}{R_T} + \sum_{i=1}^{N_R} N_{Rbit}^{i} \frac{P_R}{R_R} + c \quad (6)$$

N_T和 N_R分别表示设备发送和接收的更新数量。P_T和 P_R分别表示发送和接收时的收发器功率。R_T和 R_R分别表示发送和接收的比特率。和分别表示给定更新 i 中发送和接收的比特数，c 表示信道噪声、传输错误等无关因素消耗的能量。N_{Tbit} N_{Rbit}

如果 $P_R = P_T = P$ 和 $R_R = R_T = R$，则方程 6 可简化为方程 7：

$$E_{net} \simeq \left(\sum_{i=1}^{N_T} N_{Tbit}^{i} + \sum_{i=1}^{N_R} N_{Rbit}^{i}\right)\frac{P}{R} + c_1 \quad (7)$$

此外，每个客户端设备本地计算的能耗可由方程8近似表示。

$$E_c \simeq \sum_{i=1}^{N_{round}} T_{training}^{i} \times P_c^{i} \quad (8)$$

其中，P_c^i代表第i轮次的单位训练时间功耗，$T_{training}$表示计算操作的持续时间，N_{round}表示给定设备需要执行的操作数量。

如果$T_{training}$和P_c在给定设备的所有轮次中均相同，则方程8可简化为：

$$E_c \simeq N_{round} T_{training} P_c \quad (9)$$

总之，每个客户端设备的近似总能耗（方程5）可用方程10表示。

$$E_T \simeq N_{round} T_{training} P_c + \left(\sum_{i=1}^{N_T} N_{Tbit}^{i} + \sum_{i=1}^{N_R} N_{Rbit}^{i}\right)\frac{P}{R} \quad (10)$$

根据上述能耗公式，我们可以确定一系列会影响联邦学习中参与客户端设备能耗的参数：交换更新的数量N_T和N_R，每次交换更新所含的比特数N_{Tbit}^{i}和N_{Rbit}^{i}，传输功率P、传输比特率R、本地训练的持续时间$T_{training}$，以及本地训练轮数N_{round}。

3.3 联邦机器学习优化参数

基于上一节建立的方程以及所研究的入选论文，我们确定了若干能耗优化方面。

因此，为了最小化公式9中的总能耗，加速模型收敛的本地训练任务优化应减少联邦轮次的数量N_{round}。此外，若降低训练模型的复杂度，训练时长$T_{training}$也会缩短，而模型复杂度会影响能量效率。通过减少与中央服务器交换的数据量，以最低成本聚合更新（即，减少N_{Tbit}和N_{Rbit}，见公式7），有助于延长电池寿命。此外，模型更新交换的频率会影响更新总数N_T和N_R（公式7），从而进一步优化通信能耗。通过充分利用异构节点的计算资源参与训练，也能实现进一步优化；此外，优化客户端选择，以平衡参与节点的负载并选出最适合加速收敛的节点。最后，降低传输功率P并提高比特率R（公式7）也有助于减少总能耗。

这项分析将帮助我们在之后对文献中提出的不同方法和技术进行分类，正如我们将在第5.3小节中看到的那样。

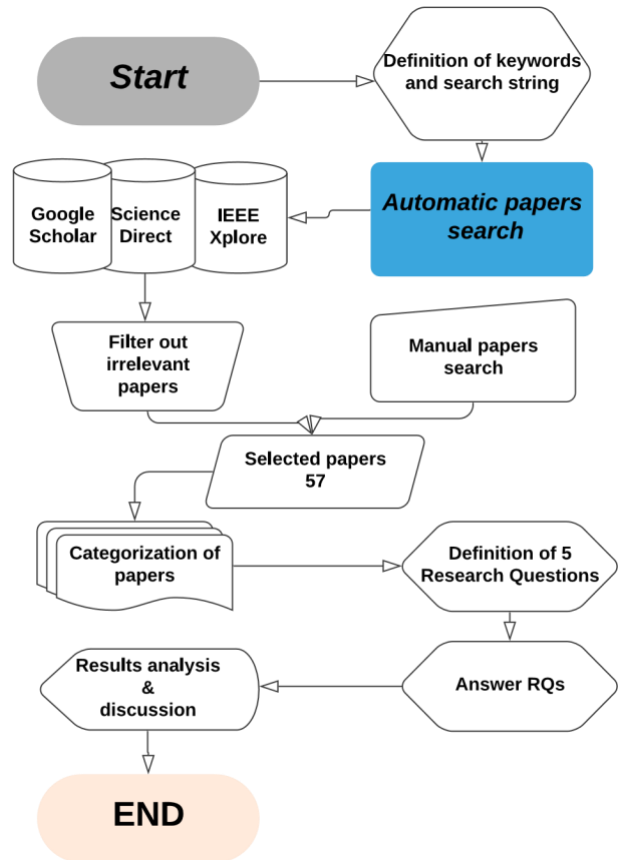


图2 我们的检索流程

4 系统映射研究流程

本节介绍了本系统映射研究全过程所遵循的流程。

更多资料可在为此创建的在线存储库中获取[1]。

图2展示了所采取的步骤。首先，根据预先定义的关键词和检索字符串，在三个常用数据库中进行自动检索；随后，根据标题筛选相关论文。接着，我们根据摘要进一步筛选，并通过阅读全文进一步缩小范围。最后，我们又进行了手动检索，以发现自动检索未能找到的文章。下文将详细介绍工作流程的各个步骤。

4.1 论文筛选

为了获取与研究相关的所有论文，我们使用清单1中的检索字符串查询了三个主要数据库（Google Scholar、IEEE Explore 和 ScienceDirect）。该检索字符串主要由以下关键词构成：联邦机器学习、边缘计算、端侧智能、能量和优化。

清单1 检索查询

"(" 联邦机器学习" OR "联邦学习") AND ("边缘计算" OR "端侧智能 ") AND (能量 OR 功率) AND (优化 OR 最优 OR 高效 OR 效率)"

筛选论文。我们对初始搜索结果进行筛选，仅保留符合以下所有纳入和排除标准的论文。

i. 纳入标准：

- 2016年至2021年7月发表的论文
- 英语论文
- 提出联邦学习能耗优化方案的论文，所采用的技术包括降低通信成本或缩短训练时间
- 总体上针对物联网或移动设备的论文

ii. 排除标准：

- 有关分布式机器学习的研究，但未明确应用于资源受限设备上的联邦学习
- 同一作者的相关研究

手动检索。为尽可能全面地覆盖文献，我们增加了一个步骤，以查找此前可能遗漏的潜在论文：通过查阅来进行反向滚雪球法

表2 研究问题

研究问题编号	问题
RQ1	发表趋势如何?
RQ2	Wh在网络 rk 架构 r被提出 ?
RQ3	如何实现能耗优化?
RQ4	如何 h优 化结果已验证 ?
RQ5	报告了哪些优化结果?

查看了入选论文中的引用文献。因此，额外纳入了一些论文，总计67篇论文。在本研究的其余部分，我们将根据论文的发表时间顺序，为入选论文分配编号：P1、P2，直至P67。所有论文及其分类列表见附录A中的表5。

4.2 研究问题

为系统地分析文献并比较提出的技术，我们定义了一组研究问题，以指导分析（表 2）。研究问题1指出论文的时间线和来源；研究问题2介绍各篇论文所考虑的网络拓扑；研究问题3考察各篇论文所涉及的 FedML 能耗优化方面；研究问题4介绍用于验证相关方法的实验设置；研究问题5则衡量实验中的优化改进。

[1] https://gitlab.com/rachid-el-mokadem/fedmlsysrev

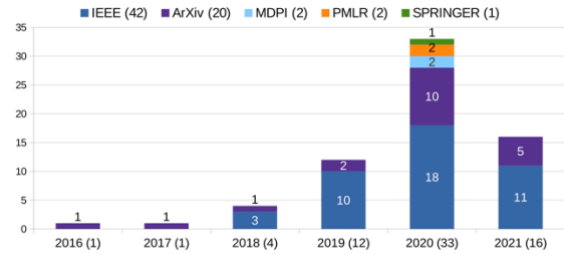


图 3 联邦机器学习论文的发表趋势随时间的变化

5 问题解答

本节将分析入选论文的研究结果，并按照第 4.2 小节中定义的研究问题逐一作答。

5.1 研究问题1 - 发表趋势如何

回答这一研究问题将有助于了解论文发表数量的变化、论文在各出版渠道中的分布情况以及论文的性质，并考察它们对FedML领域的影响。

图 3 展示了论文发表数量随时间的变化趋势。过去 3 年论文数量显著增长，仅 2020 年就发表了 33 篇。鉴于首篇论文[32] 发表于 2016年，我们可以清楚地看到

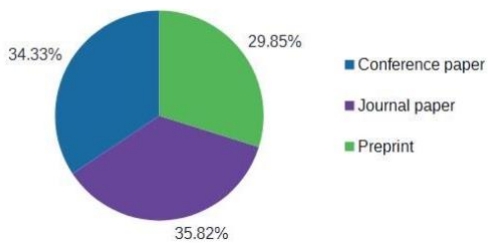


图 4 论文类型分布

这一主题正受到世界各地多个研究团队的广泛关注。

如图 4所示，大多数论文发表于期刊（≈36%）和会议（≈34%）。这表明该主题日趋成熟，也体现了科学界为此投入的积极努力。67 篇论文中，还有 20 篇（(≈29%）以预印本形式发表在-预印本服务器上 arXiv 数据库中，其中 10 篇发表于 2020 年。鉴于该主题发展迅速、反馈及时，这一情况也就不足为奇。我们还纳入了 Konecnˊ、Jakub 等人[32, 33]，未经同行评审的论文，因为它们被认为是该领域影响力最大的论文，引用次数分别为 747 次和 1733 次。同一团队还提出了具有开创性的 FedML 方案[46]。

此外，我们还统计了每篇论文的引用次数，如图 5 所示，以衡量它们对该主题的影响。显然，较早发表的论文往往比新论文获得更多引用。不过，这确实能大致反映出 论文的学术关注度在学术界的情况。从图中可以看到，有几篇论文的引用次数明显较高。对于 P1 至 P8 等较早发表的论文，这在一定程度上是合理的。然而，P15（[57]）获得了 345 次引用，P19（[8]）获得了 110 次引用，P35（[54]）获得了 145 次引用，这无疑表明这些论文影响力很大。关于这些论文所用技术的更多详情，请参见小节 5.3

5.2 研究问题2——提出了哪些网络架构

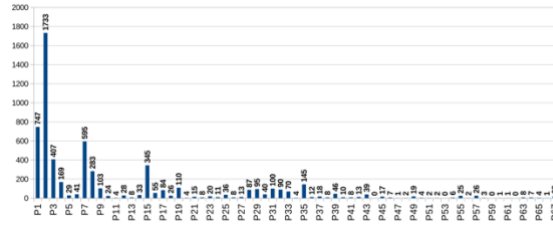


图 5 每篇论文的引用次数

在这个问题中，我们考察所研究论文提出的网络架构。这一点很重要，因为网络拓扑会影响通信成本，进而影响能耗。

架构如下。

- **集中式**：基于中央服务器，确保参与设备之间的通信和模型参数交换。由于设备与服务器之间需要远距离通信，因此这种方案会消耗较多能量，并且需要较高的传输功率 P （公式 7）。此外，该方案还存在单点故障问题。

• **去中心化**：基于节点间通信，无需中央服务器。在这种设置下，设备只需在节点之间进行短距离通信，便能节省大量能量 [15]。

• **混合式**：该架构至少由设备的三层组成，其中中间层设备部署在中央服务器与终端设备之间。

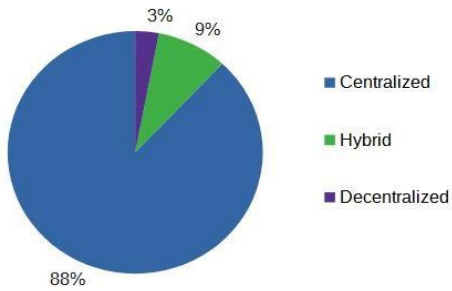


图 6 按网络拓扑划分的论文数量

混合式架构通过在主服务器与 终端设备之间添加边缘服务器来实现。这些中间设备可以发挥多种作用，例如管理其控制下的直接客户端，从而减轻中央服务器的负担，并缩短汇总多个接收的更新所需的等待时间。在某些情况下，也可以利用该边缘服务器减轻终端设备进行本地更新计算的负担：它可以定期从选定的客户端查询训练数据，以更高的计算容量执行更新，并代表终端节点与服务器通信。因此，该架构能够显著优化设备能耗，但也会对数据隐私构成一定威胁，尤其是在这些边缘服务器不可信且数据高度敏感的情况下。

图 6 显示，大多数论文（67 篇中有 59 篇）基于集中式设置，另有 2 篇采用完全去中心化方案，6 篇提出了混合式架构。集中式方案占主导地位，可能是因为原始论文中的架构[46], 来自 Google。此外，完全去中心化方案在没有中央设备的情况下聚合模型时，面临一些算法和实际挑战 [30]。

5.3 研究问题3 - 如何实现能耗优化

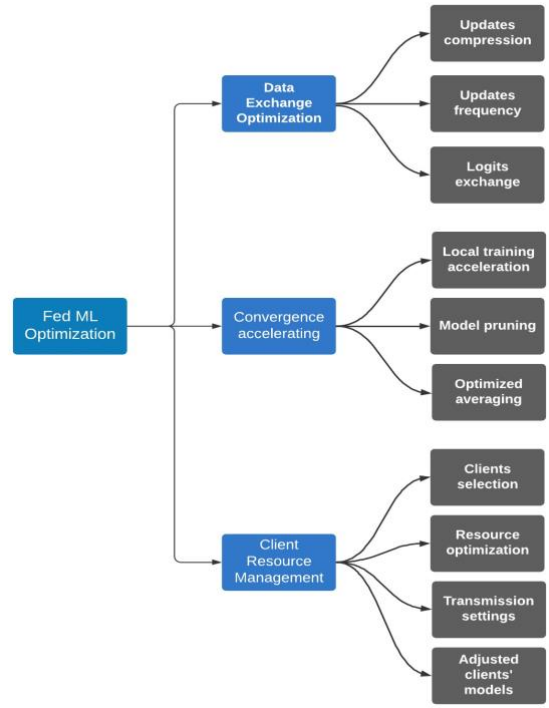


图 7 联邦学习优化技术（回顾）

在这个问题中，我们分析论文采用的优化**联邦机器学习**的技术。我们的研究侧重于功耗降低，因此，如小节 3.2所述，所有研究的优化方面都通过公式 10与能量相关联。我们将这些技术分为以下几类：（1）收敛加速；（2）数据交换优化；（3）客户端资源管理。图 7 概述了不同的技术。表 3列出了每篇受研究论文所涉及的优化方面。

5.3.1 收敛加速

在联邦机器学习中，客户端节点执行训练任务，以构建一个全局

在中央服务器的编排下训练模型，经过尽可能多的轮次以达到良好的性能。为了延长客户端设备的电池寿命，达到

全局模型收敛时间可通过多种方法缩短。

本地训练加速

许多研究采用了不同的优化方法来加速本地训练，例如自适应学习率 [42], 和 Adam优化方法[47]。

公式 11 用于联邦学习的原始版本。 w 表示模型权重，b 表示模型偏置，∇ 表示损失函数l 的梯度，μ 表示学习率。

在这一版本中，服务器会在开始时设定一个学习率参数，用于计算本地训练中的梯度下降步长。与此相对的是前述方法，它们会确定最佳步长，以便快速实现全局模型的收敛。

$$w \leftarrow w - \mu \nabla l(w, b) \quad (11)$$

这些技术的优点是减少模型收敛所需的轮次数 N_{round}（公式 9），从而降低参与设备的能耗。

因此，[47] 提出了 CE-FedAvg，该算法通过采用以学习率性能更优而著称的 Adam方法（而非原始 FedAvg 算法中使用的随机梯度下降），改进了节点计算本地更新的方式。该算法的权重更新方法由每个客户端执行，如公式12所示。

$$w_k, m_k, v_k \leftarrow AdamSGD(w_k, m_k, v_k) \quad (12)$$

其中，w_k 表示模型权重，m_k 表示 Adam 的一阶矩，v_k 表示 Adam 的二阶矩。这些参数用于通过对所有接收的更新或梯度取平均值来计算 Adam 步长，并在下一轮训练中将其发送回客户端。

特征增强是一种机器学习技术，可提升类别分布不平衡情况下的训练性能（[75]）。类似地，在联邦学习领域，FedFusion 是 [72] 提出的一种算法，它采用名为特征融合的技术，通过结合全局模型的特征空间与本地模型的特征空间来训练本地模型，从而加快全局模型的训练。该算法将全局模型用作特征提取器，然后采用多种类型的特征融合来高效聚合所有特征。此外，[71] 提出了一种结合最大均值差异（MMD）的双流模型学习方法：节点并行训练两个模型，这两个模型都以全局模型参数初始化，但其中一个（全局模型）在训练期间保持不变。该方法计算两个模型输出之间的 MMD 损失，并以此优化本地模型。这项技术常用于标准机器学习中的迁移学习和知识蒸馏；将其应用于联邦学习有助于加快训练并降低通信成本。其核心是利用全局模型参数约束本地模型训练，以避免本地模型过拟合本地数据，从而在更少的训练轮次内构建出良好的全局模型 N_{round}。 [4] 采用自适应丢弃方案，通过降低本地模型的复杂度和可训练参数数量来缩短收敛时间。实际应用中，每轮都会向每个参与客户端发送一个随机的全局模型子网络 w_c，然后利用激活分数图 M 跟踪最佳子模型的索引 A，以便在后续轮次中重复使用。

模型剪枝

模型剪枝是深度学习中广泛使用的另一种技术，它基于训练数据减少模型参数的数量，从而加快训练速度。参数数量的减少简化了模型，进而缩短了计算时间

($T_{training}$ 在公式 9），本地训练能耗 E_c，同时保持良好的模型性能。 [29] 实现了一种名为 PruneFL 的算法：首先由选定的客户端使用其本地数据进行剪枝。随后，服务器在每一轮中根据训练效率，结合所有客户端的更新，通过移除或重新添加一些参数，迭代调整得到的较小模型。为使参数的添加和删除可逆，作者使用由 0 和 1 组成的掩码，分别表示被移除和保留的权重。同样，[69] 提出了一种结合权重量化和选择性更新的结构化模型剪枝方法，以加速训练并降低设备上的计算成本。具体而言，

作者采用 l_1 – 范数 对模型权重进行剪枝，剪枝比例可在 0 到 90% 之间变化。

优化平均

传统联邦学习通过收集本地模型更新并直接对其求平均来运作。多篇论文提出采用更先进的平均方法，以实现更快的训练收敛。因此，[23] 提出了联邦动量（FedMom），这是一种采用有偏梯度的技术，根据方程13 和14，利用动量法更新全局模型：

$$v_{t+1} = w_t - \eta \sum_{k=1}^{K} \frac{n_k}{n} (w_t - w_{t+1}^k) \quad (13)$$

$$w_{t+1} = v_{t+1} + \beta(v_{t+1} - v_t) \quad (14)$$

其中，v_t是上一轮更新的平均值，而 beta β（通常等于 0.9）是用于计算更新随时间变化的移动平均值的参数。另一方面，[39]采用了分层架构，在中央服务器与客户端节点之间引入了L台边缘服务器。每台边缘服务器都有一个客户端子集s，并在将更新转发至主服务器之前对其进行聚合。作者表示，该方法能够缩短训练时间并降低节点能耗。

5.3.2 数据交换优化

全局模型由中央服务器收集并聚合参与节点的更新而构建。为实现通信高效联邦学习，多项研究对计算得到的更新的交换频率及其数据大小进行了优化，从而大幅延长参与节点的电池寿命，同时不影响全局模型的性能。在 FedAvg 中，本地模型的聚合按照以下方程15进行：

$$w_{t+1} \leftarrow \sum_{k}^{K} \frac{n_k}{n} w_k^{t+1} \quad (15)$$

其中，w_k表示每个节点学习得到的权重，n_k表示每个节点的数据点数量，且 n 表示所有K个参与客户端的数据点总数。

更新压缩

FedAvg 的一个已知局限是，参与客户端必须在每轮训练中上传完整的计算得到的更新，这会影响这些设备的功耗。所提出的优化方法可以减少节点与服务器之间交换的数据，同时保持全局模型的质量。

除了使用较少比特对最终更新进行编码的常规数据压缩算法（例如 霍夫曼编码，由[7, 43], 采用的数据大小缩减还通过其他几种方法实现，例如更新量化、稀疏化和草图化。所有这些技术的目标都是减少每轮传输的比特数 N_{Tbit}，通过无线接口发送，从而显著降低交换模型更新（公式 7）所消耗的能量。

机器学习模型的量化是通过使用较低的浮点精度来表示模型权重，从而缩减比特大小（[2]）。[68]提出了一种名为联邦训练三值量化（FTTQ）的方法，可减少上行和下行流量。该方法在训练过程中使用可调阈值进行逐层权重量化，此外还可降低训练任务的能耗预算。同样，[27, 43, 47, 54] 使用量化来缩减数据大小，在大多数情况下会与其他技术结合使用。此外，[27, 44] 提出了一种用于更新量化的自适应方案，以实现通信高效训练。

此外，还通过对全局模型和本地模型进行稀疏化，压缩交换数据：将计算出的更新中低于给定阈值的梯度值置为零。此操作会生成稀疏模型更新，可用少量比特进行编码，从而优化通信成本和能耗预算。因此，[57] 提出了稀疏三元压缩（STC），这是一种专为资源受限设备上的联邦学习需求而设计的新型压缩框架。STC 扩展了现有的压缩方法（

特别是 top-k 稀疏化 [60]），以支持下游压缩；此外，作者将稀疏化与量化和 Golomb 编码相结合，以获得更好的优化结果。 [20] 开发了一种自适应梯度稀疏化方法，该方法基于双向 top-k 梯度稀疏化，以降低服务器与节点之间双向通信的通信成本。其稀疏化参数k 由服务器确定，在通信与全局模型准确率之间进行权衡。此外，[62] 采用了带梯度校正的梯度稀疏化，以累积被消除的微小梯度，并在之后将其加回，从而加快模型训练的收敛。

为此采用的其他技术还包括梯度草图化 [33, 55] 和子采样。第一种方法基于使用名为 Count Sketch [61]的数据结构对更新进行压缩。第二种方法 [33] 要求客户端仅发送从计算所得更新中派生出的较小更新，即通过随机采样其数值生成更新。随后，服务器对所有接收到的子采样更新进行平均，以获得全局模型参数的估计值。此外， [52] 采用了双侧低秩压缩，以减小服务器与节点之间双向传输的模型大小。最后，[36] 采用了基于层的参数选择，以便仅传输每个模型层中的重要参数。

更新频率

在原始的 **联邦机器学习** 算法中，客户端在模型训练的每次迭代中都会发送更新；随着与服务器交换的更新次数N_T和N_R 增加，这会导致通信成本和能耗较高。为了在给定的资源预算下实现这一点，[65]提出了一种控制算法，用于确定本地更新与全局参数聚合之间的最佳权衡。该算法在分布式训练过程中学习数据分布和系统特征，然后根据资源约束动态确定全局模型聚合的频率。另一个方案是，[8] 提出了一种基于模型逐层频率的不同方法。这意味着，重要层的参数比不太重要层的参数交换得更频繁。原因在于，深度模型的前几层往往学习不同数据集的通用特征，而更深的层则学习更具体的特征。因此，每个节点将其模型拆分为浅层权重w_g 和深层权重w_s，并在服务器的控制下分别异步地与服务器交换这些权重。服务器会为每个客户端确定要考虑的权重类型，并执行时间加权聚合，以提高最新接收模型的权重。 [64] 提出了另一种方法：客户端降低从服务器拉取全局模型的频率（以减少下行能耗），并通过本地更新来弥补差距。

Logits 交换

有些研究选择不与服务器交换更新：仅将训练后的本地模型的输出（称为*logits*）发送到服务器，从而使联邦学习的通信成本降低多个数量级 [58]；然而，所有客户端和服务器都必须共享公共数据样本，才能计算并共享各自的输出。

为了利用这些经过精简的数据构建全局模型，[24, 26, 50, 58] 的作者采用了一种名为知识蒸馏 [59], 的迁移学习技术：多个教师模型（本地模型）将其所学知识传递给单个学生模型（全局模型）[21]；在所有情况下，蒸馏任务均由服务器执行，但[24] 除外：服务器会对发送来的 logits 求平均，然后将其发回客户端，由客户端自行执行蒸馏。在这种情况下，双向通信成本都会降低，因为服务器无需将整个模型发送给节点。

5.3.3 客户端资源管理

许多方法都能让客户端设备以最优能耗参与模型训练。客户端资源通常指 CPU 时间、内存和无线带宽，这些因素往往与能耗有关。客户端资源管理中最常用的两种方法会影响：（1）客户端参与；（2）传输设置。

客户端选择

在最初的 FedML 中，每轮都会从可用节点中随机选择参与方。后续分析表明，这种方法会导致模型收敛效果不佳，并造成节点资源浪费（[49]）。因此，许多研究通过自适应 且最优的客户端选择来解决这一问题，并根据每轮可用的资源和数据进行选择。[49] 提出了 FedCS，这是一种采用优化客户端选择的联邦学习算法：服务器首先随机选择一组客户端，然后根据客户端资源以及计算更新所需的时间，进行更有针对性的选择。此外，[3] 提出了一种服务器端强化学习 方案，该方案以能量单位e_n、CPU 周期数c_n以及每轮中各客户端用于训练的数据点数量为依据。

随后，服务器根据这些数值计算奖励，以帮助其找到最优策略和动作，从而在资源使用最优的情况下实现高效训练。同样地，[53] 提出了一种名为 FedMCCS 的多准则客户端选择模型，该模型根据 CPU、能量、内存和时间，对客户端设备进行区分式选择。服务器通过与节点进行请求/响应辅助数据交换，在训练过程中跟踪这些数值。系统使用这些属性训练线性回归模型，以预测客户端是否有足够的资源参与训练任务。此外，[67]提出根据客户端的参与历史来选择客户端，因为参与历史会影响全局模型的性能。另外，[56] 提出在每一轮中采用数据不平衡感知的参与方选择方式，确保所有数据类别至少被覆盖一次。为此，系统向每个节点请求一个位掩码 η，其中包含与可用数据类别相对应的 C 比特。随后，服务器按置位数量从多到少对这些位掩码进行排序，并通过平均更新尽量减少覆盖所有类别所需的客户端数量。

混合方案

其他论文提出了一种混合方案，基于我们的第二个研究问题中介绍的架构，以优化客户端设备的资源。[12] 开发了一种自平衡系统，该系统基于中介边缘服务器，收集客户端近似均匀的数据分布子集，并在将训练后的模型发送至中央服务器以构建全局模型之前，对这些模型进行聚合；同样，[39] 通过平衡与L 边缘服务器之间的参数交换、训练时间和通信预算，实现了能耗降低，其中每个边缘服务器包含少量客户端；[66]采用了模型更新的分层聚合，以克服节点与服务器之间的通信过载。此外，[25]和 [74] 提出了一种云-边缘-客户端方案，其中客户端将部分或全部训练任务卸载至边缘服务器，而边缘服务器获取部分客户端数据用于训练。该方法在客户端数据的通信开销和隐私问题方面存在一些缺陷，但在某些特定应用场景中可能具有相关性。

传输设置

无线传输通常会给移动设备和物联网设备带来较高的能耗。正如我们在公式 7 中所见，它与网络接口的传输功率 P 和比特率 R 有关。许多论文都针对无线传输和带宽设置的优化展开了研究。[48]提出了一种传输功率和速率优化技术，用于节点与中央服务器之间的节能通信。他们将最小化总能耗视为一个关于比特率、传输功率和本地更新所需 CPU 频率的联合优化问题。另一方面，[28] 提出了一种基于分段 Gossip 协议的节点间完全去中心化通信方式。在这种通信方案中，工作节点将其更新分割成较小的片段，分别与部分参与方共享；随后，这些参与方将片段逐步转发给其他工作节点，直到所有工作节点都获得全部片段。系统会选择带宽更快的对等工作节点来拉取更新。此外，[63] 提出了一种能量感知工作节点调度算法：节点在每一轮监控自身能耗，并将其与可调阈值进行比较，以决定是否参与下一训练轮次。随后，每个工作节点将其更新划分为 M 个片段，并通过 M 个经过调整传输功率的子信道分别传输。

自适应本地模型
最近，[73] 和[11] 提出根据节点的能力和数据对本地模型进行适配，这与当前最先进的 Fed Avg 算法相反，后者为所有客户端提供相同的模型架构。在这两项研究中，每个客户端都会从全局模型中导出一个适配模型，作为其子网络，并且具有更少的层和参数。随后，使用不同的平均方法来聚合更新。这些新方法能够感知异质性和数据不平衡，并可自适应地节省训练和通信所消耗的能量。

从图 8可以看出，数据交换优化类别的论文数量最多（32 篇），其次是客户端资源管理（25 篇），最后是收敛加速（10 篇）。图9更详细地展示了各类别使用的优化技术

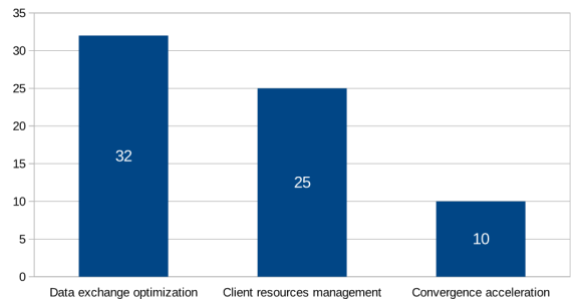


图 8 各优化类别的论文数量

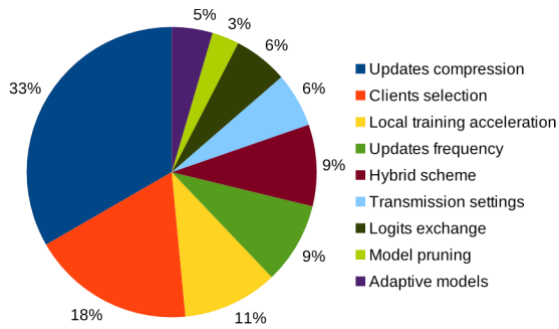


图 9 各优化技术的论文数量

论文。更新压缩和客户端选择占据了大多数论文，这并不令人意外，因为原始 FedAvg 算法在这些方面存在明显的局限性。此外，这对设备的能量节约影响最大，因为无线通信消耗了运行能耗中的大部分。其余论文则采用了其他技术，其中本地训练加速略占优势。

5.4 研究问题4 - 如何验证优化

该研究问题关注不同论文用于验证其提出的方法的实验设置（模型、数据集和测试平台）。

表 4 中的分类数据表明，大多数论文在研究的实验部分仅考虑了使用 MNIST 和 CIFAR-10 数据集的神经网络（具体而言是卷积神经网络）。虽然这似乎限制了验证范围，但这可以用以下原因解释：这些知名数据集易于获取，也易于基于 PyTorch 和 TensorFlow 等流行的机器学习框架实现神经网络。一些论文还对逻辑回归、逻辑回归和支持向量机等其他机器学习模型进行了额外验证。然而，针对非神经网络模型的验证结果仍严重不足，而这类模型的复杂度可能较低，因此资源消耗也较低。

关于实验平台，不同论文考虑了不同数量的参与节点，从 2 个到 50000 个不等。大多数论文（57 篇中的 52 篇）使用多GPU计算机上的仿真节点，而只有五篇论文（P7、P14、P17、P34、P58）在树莓派TM和智能手机等真实设备上进行了实验。仿真设备有助于深入了解验证情况，但仍需在真实设备上、现实场景中开展更多实验，尤其要关注无线通信、能源约束和计算能力。工业界和学术界快速原型开发平台的普及（例如，ArduinoTM、树莓派TM & ESP31TM）也是推动这一工作的另一项动因。

5.5 研究问题5 - 报告了哪些优化结果

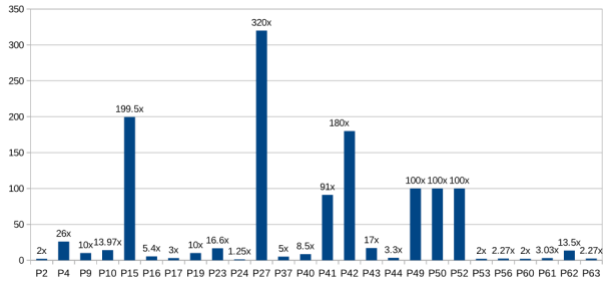


图 10 通信成本改进（按论文统计）

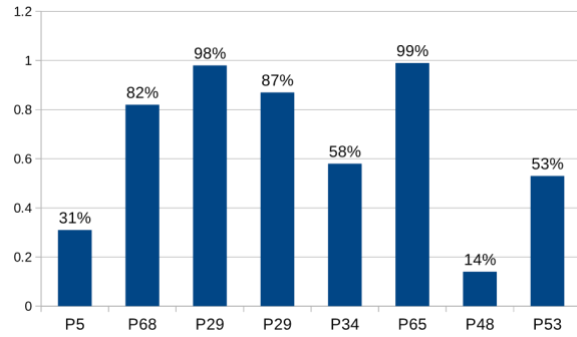


图 11 能耗改进（按论文统计）

表 3 按优化技术分类的论文列表

类别	优化技术	论文
数据交换优化	更新压缩	P2 P15 P17 P19 P23 P45 P24 P26 P27 P35 P37 P39 P40 P42 P64 P52
	更新频率	P1 P7 P8 P20 P56
	Logits 交换	P4 P10 P46 P49
客户端资源管理	客户端选择	P3 P5 P18 P25 P28 P31 P32 P38 P44 P47 P48
	混合方案	P16 P29 P34 P36 P51
	传输设置	P67 P22 P30 P57
	自适应模型	P54 P55
收敛加速	本地训练加速 模型剪枝	P6 P11 P12 P13 P33 P63
		P14 P58
	优化模型平均	P21

在这个问题中，我们列出并比较了调查论文中获得的优化结果。数值结果分为三类：（1）通信成本；（2）收敛时间；（3）能耗。每篇论文都量化了相较于标准 FedAvg [46] 算法所取得的优化幅度，各类别的数值结果列于图表中。从图 10 可以看出，全局通信成本降低的幅度范围很广，从 2 倍到 320 倍不等。从图 12 可以看出，收敛时间的改善幅度从 3% 到 98% 不等。至于能耗，图 11 显示改善幅度介于 14% 和 99% 之间。

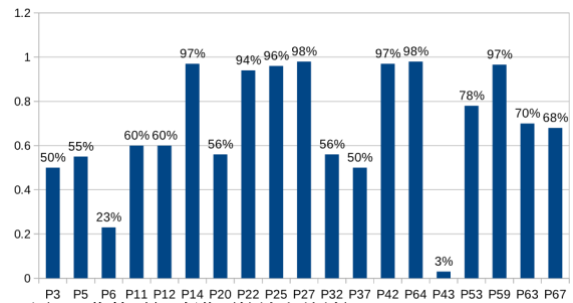


图 12 收敛时间改进（按论文统计）

收敛时间和通信成本优化的结果非常令人鼓舞，这与社区对这两个方面的浓厚兴趣一致（图 8）。尽管它们会直接影响能耗，但请注意，在我们的阅读中，评估过优化

的效益对能耗的直接影响的论文相对较少（67 篇中仅有 8 篇），而这对我们的研究至关重要。

6 讨论

本节结合前文分析，依据第5节中的研究问题，讨论系统映射研究的结果。对于提出的每个研究问题（RI），我们将介绍：（1）与能源受限联邦机器学习相关的一些显著局限性；（2）面向研究界的一些改进方向和建议。

6.1 RI1：完全去中心化方案

在基于联邦机器学习的集中式方案中，客户端节点在训练期间交换数据，并与中央服务器通信，该服务器通常位于云端。因此，节点必须使用远距离无线通信连接服务器，这意味着功耗较高 [15]。为克服这一问题，我们必须利用节点之间功耗低得多的短距离通信，通过点对点通信交换更新。

实现完全去中心化 FedML 的所提方法会给资源受限设备带来过载，这是因为节点需要执行额外操作来弥补中央服务器的作用。在[28], 中，节点还必须扮演聚合服务器的角色，而[13] 提出了一种技术：所有节点以链式方案（采用某种多跳

表4 论文验证设置

论文	ML model	数据集	数量 f nodes
P1	RNN	Bl帖子数据集	1000
P2	CNN - RNN	CIFAR1 0 - Reddit 公共帖子	100 - 1024
P3	CNN	CIFAR1 0 Fas hionMNIST	1000
P4	CNN	MNIST	10
P5	RNN	-	3
P6	CNN : AlexNet	CIFAR10 - M NIST	2 – 1 00
P7	CNN	CIFAR10 - MNIST	5（Raspberry Pi）– 500
P8	-	-	50
P9	CNN - 长短期记忆网络	MNIST	30
P10	ANN	MNIST	25
P11	CNN	MNIST - CIFAR1 0	-
P12	CNN	MNIST - CIFAR1 0	-
P13	CNN	CIFAR-10	20
P14	卷积神经网络：VGG 11 - L eNet	FeMNIS T	5 - 10 (RaspberryPi)
P15	卷积神经网络：VGG11 - 长短期记忆网络 - 逻辑 回归	CIFAR-10 - 关键词识别数据集 - FashionMNIST- NIST - MNIST	100
P16	CNN	EMNIS T - CIFAR-10 - CINIC-10	500
P17	CNN	MNIST CIFAR1 0	10 - 4 0（Raspberry Pi）
P18	Linea r model	Rando m in整数数据	20
P19	卷积神经网络 循环神经网络	MNIST 人类活动识别	20
P20	ANN	MNIST	15
P21	逻辑回归 - 卷积神经网络（Res 1 Net 8）	CIFAR-10	20
P22	CNN	CIFAR10 - MNIST - SVHN	20
P23	CNN	FEMNIST	50
P25	回归模型 - 卷积神经网络 - 支持向量机	波士顿房价数据集 - MNIST - KDD Cup'99 数据集	5 - 100 - 500
P26	SNN	MNIST-D VS 数据集	2
P27	卷积神经网络 (LeNet-5 – CifarNet – DenseNet-121)	MNIST - CIFAR-10 - ImageNet	64
P28	CNN	MNIST	50
P29	CNN	MNIST - CIFAR1 0	50
P30	ANN	MNIST	50
P31	CNN	-	80
P32	CNN	Fashion/MNIST - CIFAR1 0	100
P33	CNN	MNIST	3
P35	线路 r R出口 ion - CNN	MNIST - CIFAR1 0	50

P36	CNN	MNIST	500个客户端/10台边缘服务器r
P37	CNN - Lin ear R出口流量ion	MINST - Californi一个房屋数据集	10
P39	卷积神经网络（AlexNet）- Transformer （GPT2-small）	CIFAR-10 - PersonaChat	10000 至 50000
P40	卷积神经网络 - 逻辑回归	CIFAR-10 - FashionMNIST - Senti-ment140	100
P41	CNN	CiFar10 - MINST	10
P42	卷积神经网络 MNIST	CIFAR-10	1 到 64
P43	卷积神经网络（残差网络）	CiFar10 - CiFar100	16
P44	ANN	MNIST - FEMNIST	100
P46	CNN	MNIST - CIFAR-10	10 至 20
P47	CNN	FeMNIS - CiFar10 - CiFar100	100
P49	卷积神经网络 - 长短期记忆网络	MNIST - IMDb	NA
P50	卷积神经网络（残差网络）	CIFAR-10	10
P52	CNN	MNIST - EMNIST	20
P53	人工神经网络 - 卷积神经网络	EMNIST	50 - 1000
P55	卷积神经网络（残差网络）- Transformer	MNIST - CIFAR-10 - WikiText-2	100
P56	CNN	MNIST	30
P57	CNN	MNIST - CIFAR-10 和 SVHN	20
P58	CNN	MNIST - CIFAR-10	3（Core i5 个人电脑）
P59	CNN	MNIST	10-18
P60	CNN	CIFAR-10 - FEMNIST - IMDB	10
P61	CNN	CIFAR-10	10
P63	CNN	FEMNIST - Shakespeare - Se ntiment140	-
P64	CNN	MNIS T – CIFAR-10	12
P65	CNN	CIFAR-10 - F ashionMNIST	8
P66	CNN	MNIST - F ashionMNIST - CIFAR1 0	-
P67	CNN - 长短期记忆网络	FEMNIST - 莎士比亚数据集	-

通信）计算并交换各自的更新。第一种技术的局限在于，节点需要承担额外开销才能扮演服务器角色；第二种技术则要求节点在整个训练过程中始终运行：这两种情况下，节点层面都需要更多的能量和资源

混合方案似乎克服了其中一些问题，因为它设有一个基于云的中央服务器，仅用于管理客户端参与和选择，节点的数据查询也非常有限，同时模型聚合仍在客户端设备之间进行。在

与此同时，为了倡导完全去中心化方案，需要建立一种新的理论框架，以支持完全去中心化的模型聚合，并实现合理的能量和资源消耗。

6.2 RI2：大型模型缩减

一些规模较大且包含大量可训练参数的**联邦机器学习**模型（例如，深度神经网络）需要耗费大量计算资源进行训练，[17] 这对能源受限的物联网设备而言是一项挑战。

如果我们将联邦机器学习视为一种自助聚合 [22]，在不同分布式节点的数据集上构建全局模型，我们就可以缩减本地模型的规模和复杂度，从而让设备更轻松地完成训练任务。通过反复聚合分布式模型，我们仍然可以构建高性能的全局模型（例如，通过多数投票）。

另一种可能的解决方案是 Frankle 和 Carbin 提出的彩票中奖票假说 [16]，。该假说指出，稠密神经网络中包含一个稀疏子网络，其训练后可达到与初始网络相当的性能。将这一技术应用于联邦学习场景，可以大幅缩减全局模型的规模和复杂度，从而加快训练速度，并降低节点上的计算负载。目标稀疏模型可通过自适应迭代剪枝（AIP）[19] 或神经架构搜索（NAS）[51]，从全局模型中获得；服务器可根据模型性能和可用客户端资源自适应地执行这些方法。

6.3 RI3：能量感知数据压缩

许多论文提出了不同的技术，以减少向服务器发送或从服务器接收的数据量 [33, 54, 57]。所用的压缩技术包括：草图化、稀疏化、量化和数据编码。这些技术显著降低了客户端设备的通信成本。然而，它们也给节点增加了额外任务，导致压缩、编码和解码传输数据需要占用内存和 CPU。

许多研究 [37, 37] 揭示了误差受控有损压缩方案在压缩率和计算效率方面的优势。我们建议研究一种适用于联邦学习更新压缩的等效技术，以进一步降低通信能耗。

6.4 RI4：异质性感知优化

节点异质性是联邦学习在现实世界应用 [35]中面临的一个关键问题。因此，这一主题引起了研究界的关注，已有多项研究 [49,70]，提出了基于判别式参与节点选择的不同方法。这些方法只选择同时具备所需 资源和训练数据的设备。虽然这似乎解决了异质性问题，但剔除缺少其中任一条件的设备，可能会影响模型性能和收敛时间。

在这种情况下，分别利用设备的数据和计算能力来管理设备会更有利。有些节点可以利用自身的数据参与，其他节点可以贡献计算容量，其余节点则两者兼顾。为实现这种灵活性，一些设备可以交换原始数据、提取特征或数据标签。为在这些通信过程中保护数据隐私（这是联邦机器学习 的依据之一），我们可以采用轻量级加密技术（例如，基于椭圆曲线的技术）进行节点间通信，或针对不可信的节点间关系使用同态加密方案 [18] 。

我们还可以提出一种基于将客户端节点分为两组的异质性管理方法。第一组由性能强大、资源充足、专用于训练任务的设备组成；另一组则由资源有限、仅使用自身数据进行验证的节点组成。验证分数可用作反馈，以自适应地调整模型聚合参数。

6.5 RI5：结果验证

大多数受研究的论文都使用功能强大的计算机上的仿真节点来验证其方法。此外，研究重点还集中在使用卷积神经网络（CNNs）执行图像识别任务，以及使用 MNIST 和 CIFAR-10 等常见数据集。选择图像数据进行验证可能是

由于其敏感性和较大的规模，从而导致较高的通信成本，因此为 **联邦机器学习** 的使用提供了理由。然而，在物联网以及移动设备中，还存在其他类型的潜在应用，它们具有不同形式的数据和学习任务（例如，温度或湿度等环境量）。

为了以透明且可复现的方式验证所提出的技术和取得的成果，我们强调，应在真实世界条件下，使用真实的物联网或移动设备以及多样化的学习任务，开展先进的测试平台验证。此外，我们建议建立标准化的联邦学习性能分析基准测试，使世界各地的研究人员能够使用真实且多样化的数据和现实场景验证其研究成果。

6.6 RI6：联邦推理

大多数关于协同机器学习的文献都集中于训练阶段。尽管推理任务在能量和资源方面的开销较低，但在能量受限的物联网设备的 FedML 场景中，我们可能仍需考虑如何协同计算预测结果或对事件进行分类。这对于涉及大量数据和模型的音频及视频处理也有益处。此外，当需要利用多个节点之间的相关性来分类或预测某个值时，这一主题的重要性尤为明显。

因此，我们建议为 FedML 开发协同推理框架，使各节点能够相互协助，平衡预测或分类负载，而不是依赖服务器完成这项任务。同样，必须采用适当的加密机制来保障数据隐私。

7 效度威胁

任何调查或系统映射研究（包括我们的研究）都可能存在一些常见局限性 [10], , 涉及文献覆盖范围以及处理所研究条目时产生的偏差。为了尽可能减少这些威胁，我们尝试遵循一个定义明确的流程 [31]。首先，我们在不同数据库中全面检索相关论文，并利用检索词同义词尽可能获取更多有效结果。随后，我们分多个阶段手动筛选论文：先查看标题和关键词，再阅读摘要，最后研读全文。我们至少重复了这一流程两次：一次在研究开始时，另一次在几个月后。

然而，由于我们使用的是当时可用的资源，因此可能存在与检索字符串选择、数据收集过程、研究问题选择和时间跨度相关的问题。

关于检索字符串的选择，尽管我们使用了明确的关键词，但由于关键词索引不完善，仍可能遗漏一些相关文献。我们对先前经过验证的论文进行了人工滚雪球检索，这帮助我们发现了自动检索流程遗漏的一些文章。不过，这种方法也未必总是足够。

关于数据收集过程，每篇文章（标题、摘要和全文）均由一名研究人员审阅，这可能导致一些错误。我们通过彼此讨论，部分解决了这一问题。

关于研究问题的选择，尽管我们进行了广泛讨论，力求尽可能全面、清晰，但仍可能有一些方面未被涵盖。

关于时间跨度，我们涵盖了从奠基性论文于2016年发表到2021年7月这一时期。此后可能还发表了一些有趣的论文。

最后，我们希望未来能有更多资源来解决前文提到的潜在不足，以及同行研究人员指出的其他问题。

8 结论与未来工作

摘要。据我们所知，本文首次对**面向能量受限物联网的联邦机器学习** 开展了系统映射研究

设备。我们通过可复现的研究流程，筛选出自奠基论文由 [46] 发布以来与该主题相关的67篇论文，并尝试通过滚雪球检索和手动检索来弥补可能存在的偏差。

结果分析围绕与总体发表趋势、**联邦机器学习** 网络架构以及能耗优化方案（报告结果与验证）相关的 5 个研究问题展开。结果表明，文献对更新压缩和客户端选择的关注度最高，并在降低通信成本（最高达 320 倍）、收敛时间（最高达 98%）和能耗（最高达 99%）方面取得了有意义的成果。

通过分析，我们确定了 6 个研究问题及相关建议：完全去中心化方案、大型模型缩减、能量感知数据压缩、异构性利用、真实场景结果验证以及联邦推理。相关建议包括采用全局模型尺寸缩减和高效数据压缩方案等方法，以帮助降低节点的通信成本和计算成本。为高效应对系统异构性，我们提出对资源受限设备实施自适应且灵活的管理，并将其纳入训练。最后，我们强调需要建立一个专用的标准基准，以便在真实世界条件和真实测试平台上，对结果进行透明且严格的验证。

未来工作。我们计划针对完全去中心化的**联邦机器学习**这一特定主题开展一项系统性文献综述（系统性文献综述），这似乎非常有趣。事实上，这种方式消除了单点故障，也带来了在没有任何中心节点的情况下进行更新聚合所面临的挑战。系统性文献综述旨在深入探讨某个具体问题，而系统性映射研究则是对整个领域进行广泛梳理。因此，我们认为，这项工作是本研究顺理成章的延伸。

参考文献

[1] M. Aledhari、R. Razzak、R. M. Parizi 和 F. Saeed。《联邦学习：赋能技术、协议与应用综述》。*IEEE Access*，8:140699–140725，2020。

[2] D. Alistarh、D. Grubic、J. Li、R. Tomioka 和 M. Vojnovic。Qsgd：通过梯度量化与编码实现高效通信的随机梯度下降。神经信息处理系统进展，30：1709–1720，2017。

[3] T. T. Anh、N. C. Luong、D. Niyato、D. I. Kim 和 L.-C. Wang。《移动群智机器学习的高效训练管理：一种深度强化学习方法》。*IEEE* 无线通信快报，8 (5):1345– 1348，2019。[4] N. Bouacida、J.Hou、H. Zang 和 X. Liu。《自适应联邦丢弃：提升联邦学习的通信效率和泛化能力》。*arXiv* 预印本 *arXiv:2011.04050*，2020。[5] C. Briggs、Z. Fan 和 P. Andr´as。《面向物联网的隐私保护联邦学习综述》。*arXiv*：学习，2020。[6] C. Briggs、Z. Fan 和 P. Andras。《面向物联网的隐私保护联邦学习综述》。*arXiv* 电子预印本，页码 arXiv– 2004，2020。[7] Z. Chai、Y. Chen、L. Zhao、Y. Cheng 和 H. Rangwala。《FedAT：一种在非独立同分布数据下采用异步层级的通信高效联邦学习方法》。*arXiv* 预印本 *arXiv:2010.05958*，2020。[8] Y. Chen、X. Sun 和 Y. Jin。《通过逐层异步模型更新和时间加权聚合实现通信高效的联邦深度学习》。*IEEE*神经网络与学习系统汇刊，31(10):4229–4238，2019。[9] V. Cisco。思科可视化网络指数：2017–2022 年预测与趋势。白皮书，1:1，2018。[10] F. Q. Da Silva、M. Suassuna、A. C. C. Franca、A. M. Grubb、T. B. Gouveia、C. V. Monteiro 和 I.

E. dos Santos。软件工程研究中实证研究的复现：一项系统映射研究。《实证软件工程》，19(3):501–557，2014。

[11] E. Diao、J. Ding 和 V. Tarokh。Heterofl：面向异构客户端的计算与通信高效联邦学习。*arXiv* 预印本 *arXiv:2010.01264*，2020。

[12] M. Duan、D. Liu、X. Chen、Y. Tan、J. Ren、L. Qiao 和 L. Liang。Astraea：一种用于提高移动深度学习应用分类准确率的自平衡联邦学习方法。载于*2019 IEEE* 第*37*届国际计算机设计会议（*ICCD*），第246–254页。IEEE，2019。

[13] A. Elgabli、J. Park、A. S. Bedi、C. B. Issaid、M. Bennis 和 V. Aggarwal。Qgadmm：用于通信高效的去中心化机器学习的量化群组交替方向乘子法。《*IEEE*通信汇刊》，2020。

[14] G. Fettweis 和 E. Zimmermann。信息通信技术能耗的趋势与挑战。载于第十一届无线个人多媒体通信国际研讨会论文集，第 2 卷，第 6 页。Citeseer，2008 年。

[15] X. Foukas、K. Kontovasilis 和 M. K. Marina。移动设备的短距离协作可实现节能垂直切换。《无线通信与移动计算》，2018，2018。

[16] J. Frankle 和 M. Carbin。彩票假设：寻找稀疏可训练神经网络。*arXiv* 预印本 *arXiv:1803.03635*，2018 年。

[17] E. Garc´ıa-Mart´ın、C. F. Rodrigues、G. Riley 和 H. Grahn。《机器学习中的能耗估算》。《并行与分布式计算杂志》，134：75– 88，2019。

[18] C. Gentry 等人。一种全同态加密方案，第 20 卷。斯坦福大学，斯坦福，2009。

[19] Y. Gordienko、Y. Kochura、V. Taran、N. Gordienko、A. Bugaiov 和 S. Stirenko。用于加速深度神经网络的自适应迭代剪枝。载于 *2019* 年第十一届国际电子与信息技术科学实践会议（*ELIT*），第 173–178 页。IEEE，2019 年。

[20] P. Han、S. Wang 和 K. K. Leung。面向高效联邦学习的自适应梯度稀疏化：一种在线学习方法。*arXiv* 预印本 *arXiv:2001.04756*，2020。

[21] G. Hinton、O. Vinyals 和 J. Dean。神经网络中的知识蒸馏。*arXiv* 预印本 *arXiv:1503.02531*，2015。

[22] T. Hothorn 和 B. Lausen。双重装袋：通过自助聚合组合分类器。《模式识别》，36(6):1303–1309，2003。

[23] Z. Huo、Q. Yang、B. Gu、L. C. Huang 等。使用新型联邦动量算法实现更快的设备端训练。*arXiv* 预印本 *arXiv:2002.02090*，2020。

[24] S. Itahara、T. Nishio、Y. Koda、M. Morikura 和 K. Yamamoto。基于蒸馏的半监督联邦学习：利用非独立同分布的私有数据实现通信高效的协同训练。*arXiv* 预印本 *arXiv:2008.06180*，2020。

[25] J. Jeon、S. Park、M. Choi、J. Kim、Y.-B. Kwon 和 S. Cho。《高性能且稳定的能量高效联邦学习平台的最优用户选择》。《电子学》，9(9):1359，2020。

[26] E. Jeong、S. Oh、H. Kim、J. Park、M. Bennis 和 S.-L. Kim。《面向设备端机器学习的通信高效方法：非独立同分布私有数据下的联邦蒸馏与数据增强》。*arXiv* 预印本 *arXiv:1811.11479*，2018。

[27] D. Jhunjhunwala、A. Gadhikar、G. Joshi 和 Y. C. Eldar。面向通信高效联邦学习的模型更新自适应量化。载于 *ICASSP 2021—2021 IEEE*声学、语音与信号处理国际会议*(ICASSP)*，第 3110–3114 页。IEEE，2021。

[28] J. Jiang、L. Hu、C. Hu、J. Liu 和 Z. Wang。BACombo——带宽感知去中心化联邦学习。《电子学》，9(3):440，2020。

[29] Y. Jiang、S. Wang、B. J. Ko、W.-H. Lee 和 L. Tassiulas。模型剪枝支持在边缘设备上进行高效的联邦学习。*arXiv* 预印本 *arXiv:1909.12326*，2019。

[30] P. Kairouz、H. B. McMahan、B. Avent、A. Bellet、M. Bennis、A. N. Bhagoji、K. Bonawitz、Z. Charles、G. Cormode、R. Cummings 等。联邦学习的进展与开放问题。*arXiv* 预印本 *arXiv:1912.04977*，2019。

[31] B. Kitchenham 和 S. Charters。软件工程中开展系统性文献综述的指南。2007 年。

[32] J. Koneˇcny`、H. B. McMahan、D. Ramage 和 P. Richt´arik。联邦优化：面向端侧智能的分布式机器学习。*arXiv* 预印本 *arXiv:1610.02527*，2016年。

[33] J. Koneˇcny`、H. B. McMahan、F. X. Yu、P. Richt´arik、A. T. Suresh 和 D. Bacon。联邦学习：提升通信效率的策略。*arXiv*预印本 *arXiv:1610.05492*，2016年。

[34] L. Li、D. Shi、R. Hou、H. Li、M. Pan 和 Z. Han。To talk or to work：面向异构移动边缘设备上节能联邦学习的灵活通信压缩。*arXiv* 预印本 *arXiv:2012.11804*，2020 年。[35] T. Li、A. K. Sahu、A. Talwalkar 和 V. Smith。联邦学习：挑战、方法与未来方向。《*IEEE*信号处理杂志》，37(3)：50–60，2020 年。[36] Z. Lian、W. Wang 和 C. Su。COFEL：采用本地差分隐私的通信高效优化联邦学习。载于 *ICC 2021-IEEE*国际通信会议，第 1–6 页。IEEE，2021 年。

[37] X. Liang、S. Di、D. Tao、S. Li、S. Li、H. Guo、Z. Chen 和 F. Cappello。针对科学数据集高压缩比优化的误差可控有损压缩。载于*2018*年*IEEE*国际大数据会议（*Big Data*），第438–447页。IEEE，2018年。

[38] W. Y. B. Lim、N. C. Luong、D. T. Hoang、Y. Jiao、Y.-C. Liang、Q. Yang、D. Niyato 和 C. Miao。移动边缘网络中的联邦学习：全面综述。*IEEE*通信调查与教程，22(3)：2031–2063，2020。

[39] L. Liu、J. Zhang、S. Song 和 K. B. Letaief。客户端-边缘-云分层联邦学习。载于 *ICC 2020-2020 IEEE*国际通信会议（*ICC*），第 1–6 页。IEEE，2020。

[40] Y. Liu、L. Zhang、N. Ge 和 G. hao Li。联邦学习系统性文献综述：模型质量视角。*ArXiv*，abs/2012.01973，2020。

[41] S. K. Lo、Q. Lu、C. Wang、H. Paik 和 L. Zhu。联邦机器学习系统性文献综述：软件工程视角。*arXiv*预印本 *arXiv:2007.11354*，2020。

[42] Z. Ma、Y. Xu、H. Xu、Z. Meng、L. Huang 和 Y. Xue。《面向资源受限边缘计算中联邦学习的自适应批量大小》。*IEEE*移动计算汇刊，2021。[43] A. Malekijoo、M. J. Fadaeieslam、H. Malekijou、M. Homayounfar、F. Alizadeh-Shabdiz 和 R. Rawassizadeh。《Fedzip：一种压缩

用于通信高效联邦学习的框架。*arXiv* 预印本 *arXiv:2102.01593*，2021。

[44] Y. Mao, Z. Zhao, G. Yan, Y. Liu, T. Lan, L. Song, 和 W. Ding。带自适应量化的通信高效联邦学习。*arXiv* 预印本 *arXiv:2104.06023*，2021年。

[45] B. Martinez、M. Monton、I. Vilajosana 和 J. D. Prades。模型的力量：物联网设备的功耗建模。*IEEESensors Journal*，15(10):5777–5789，2015年。

[46] H. B. McMahan、E. Moore、D. Ramage 和 B. A. y Arcas。使用模型平均的深度网络联邦学习。*arXiv* 预印本 *arXiv:1602.05629*，2016年。

[47] J. Mills、J. Hu 和 G. Min。面向物联网无线边缘智能的通信高效联邦学习。*IEEE*物联网期刊，7(7):5986–5994，2019。[48] X. Mo 和 J. Xu。采用通信与计算联合设计的能效联邦边缘学习。*arXiv* 预印本 *arXiv:2003.00199*，2020。[49] T. Nishio 和 R. Yonetani。移动边缘中面向异构资源的联邦学习客户端选择。载于*ICC 2019-2019 IEEE*国际通信会议（*ICC*），第 1–7 页。IEEE，2019。[50] J. Park、S. Wang、A. Elgabli、S. Oh、E. Jeong、H. Cha、H. Kim、S.-L. Kim 和 M. Bennis。在网络边缘蒸馏端侧智能。*arXiv* 预印本 *arXiv:1908.05895*，2019。[51] H. Pham、M. Guan、B. Zoph、Q. Le 和 J. Dean。通过参数共享实现高效的神经架构搜索。载于国际机器学习会议，第 4095–4104 页。PMLR，2018。[52] Z. Qiao、X. Yu、J. Zhang 和 K. B. Letaief。通信高效联邦学习采用双侧低秩压缩。*arXiv* 预印本 *arXiv:2104.12416*，2021。

[53] S. A. Rahman、H. Tout、A. Mourad 和 C. Talhi。FedMCCS：面向最优物联网联邦学习的多标准客户端选择模型。

*IEEE*物联网期刊，2020。

[54] A. Reisizadeh、A. Mokhtari、H. Hassani、A. Jadbabaie 和 R. Pedarsani。Fedpaq：一种采用周期性平均和量化的通信高效联邦学习方法。发表于 国际人工智能与统计会议，第 2021–2031 页。PMLR，2020。

[55] D. Rothchild、A. Panda、E. Ullah、N. Ivkin、I. Stoica、V. Braverman、J. Gonzalez 和 R. Arora。Fetchsgd：采用草图化的通信高效联邦学习。载于国际机器学习会议，第 8253–8265 页。PMLR，2020 年。[56] D. Sarkar、S. Rai 和 A. Narang。Catfedavg：优化联邦学习中的通信效率和分类准确率。*arXiv* 预印本 *arXiv:2011.07229*，2020 年。[57] F. Sattler、S. Wiedemann、K.-R. Mu¨ller 和 W. Samek。基于非独立同分布数据的鲁棒且通信高效的联邦学习。*IEEE* 神经网络与学习系统汇刊，31(9)：3400–3413，2019 年。[58] F. Sattler、A. Marban、R. Rischke 和 W. Samek。通信高效的联邦蒸馏。*arXiv* 预印本 *arXiv:2012.00632*，2020 年。[59] H. Seo、J. Park、S. Oh、M. Bennis 和 S.-L. Kim。联邦知识蒸馏。*arXiv* 预印本 *arXiv:2011.02367*，2020 年。[60] S. Shi、X. Chu、K. C. Cheung 和 S. See。理解分布式深度学习中的 top-k 稀疏化。*arXiv* 预印本 *arXiv:1911.08772*，2019 年。

[61] W. Siblini、F. Meyer 和 P. Kuntz。利用计数草图降低使用梯度下降训练模型时的内存消耗。载于*2019*年国际神经网络联合会议 *(IJCNN)*，第 1–8 页。IEEE，2019 年。[62] H. Sun、S. Li、F. R. Yu、Q. Qi、J. Wang 和 J. Liao。面向结合边缘计算的物联网通信高效联邦学习。*IEEE*物联网期刊，7(11)：11053–11067，2020 年。[63] Y. Sun、S. Zhou 和 D. Gu¨ndu¨z。面向冗余数据联邦学习的节能模拟聚合。载于*ICC 2020-2020 IEEE*国际通信会议 *(ICC)*，第 1–7 页。IEEE，2020 年。[64] H. Wang、Z. Qu、S. Guo、X. Gao、R. Li 和 B. Ye。采用本地补偿的间歇式拉取，实现通信高效的分布式学习。*IEEE*新兴计算主题汇刊，2020 年。[65] S. Wang、T. Tuor、T. Salonidis、K. K. Leung、C. Makaya、T. He 和 K. Chan。资源受限边缘计算系统中的自适应联邦学习。*IEEE*通信领域精选期刊，37(6)：1205–1221，2019 年。[66] Z. Wang、H. Xu、J. Liu、H. Huang、C. Qiao 和 Y. Zhao。边缘计算中采用分层聚合的资源高效联邦学习。载于 *IEEE INFOCOM 2021-IEEE*计算机通信会议，第 1–10 页。IEEE，2021 年。[67] W. Wu、L. He、W. Lin、R. Mao、C. Maple 和 S. A. Jarvis。SAFA：一种低开销的快速联邦学习半异步协议。*IEEE*计算机汇刊，2020 年。[68] J. Xu、W. Du、Y. Jin、W. He 和 R. Cheng。用于通信高效联邦学习的三元压缩。*IEEE*神经网络与学习系统汇刊，2020 年。

[69] W. Xu, W. Fang, Y. Ding, M. Zou, 和 N. Xiong。通过剪枝、量化和选择性更新，加速大数据分析中的物联网联邦学习。*IEEEAccess*，9：38457–38466，2021。[70] C. Yang, Q. Wang, M. Xu, S. Wang, K. Bian, 和 X. Liu。异质性感知联邦学习。*arXiv preprint arXiv:2006.06983*，2020。[71] X. Yao, C. Huang, 和 L. Sun。双流联邦学习：降低通信成本。载于*2018*年*IEEE*视觉通信与图像处理会议 *(VCIP)*，第1–4页。IEEE，2018。[72] X. Yao, T. Huang, C. Wu, R. Zhang, 和 L. Sun。迈向更快、更优的联邦学习：一种特征融合方法。载于*2019*年*IEEE*国际图像处理会议 *(ICIP)*，第175–179页。IEEE，2019。[73] R. Yu 和 P. Li。面向移动边缘计算中的资源高效联邦学习。*IEEE Network*，35(1)：148–155，2021。[74] C. W. Zaw, S. R. Pandey, K. Kim, 和 C. S. Hong。多接入边缘计算系统中联邦学习的能量感知资源管理。*IEEE Access*，9：34938–34950，2021。[75] Y. Zhang, B. Sun, Y. Xiao, R. Xiao, 和 Y. Wei。利用条件混合 WGAN 进行不平衡分类的数据增强。信号处理：图像通信，75：89–99，2019。[76] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, 和 V. Chandra。非独立同分布数据下的联邦学习。*arXiv preprint arXiv:1806.00582*，2018。

A. 附录

本附录列出了本研究纳入的所有论文，并按时间顺序标记为 P1至P67。

表5 论文列表

编号	论文标题	优化技术
P1	F联邦优化 优化 n:Distri分布式 Machine Learning for On-D设备 In智能 度	数据交换优化 h化 优化 n
P2	F联邦学习 rning: Str策略 r Im提升通信 mmunicatio效率 ciency	数据交换 h优化 优化 n
P3	移动边缘中异构资源下的联邦学习客户端选择	客户端资源 Man管理
P4	通信高效的设备端机器学习：非独立同分布的私人数据下的联邦蒸馏与 数据增强	数据交换优化
P5	移动群智机器学习的高效训练管理：一种深度 强化学习方法	客户端资源管理
P6	T双流 m F联邦学习 习：R 减少 Co mm信成本	收敛加速 n
P7	Adapti布式联邦 Lea习在资源 受限边缘计 m算系统	数据交 h换优 化
P8	无线网络上的联邦学习：优化模型设计与分析	数据交 h换优 优化
P9	CMFL：移ti门控 Co mmuni cation 开销 销 for F联邦学 rning	数据交换优化 h换优化 化
P10	知识蒸馏g On-D设备 Intelli网络上的智能 rk Edge	数据交换优化 h换优化 化 n
P11	通过客户端附加机制降低通信成本的联邦学习	收敛加速 n
P12	T迈向更快的 Better F联邦学 习：一种特 征 Fusion A方 h	收敛加速 n
P13	通过空中计算进行二阶优化的端侧联邦学习	收敛加速 n
P14	Mode剪枝使 能 Effi效的联邦学习 在边缘设 备上	收敛加速 n
P15	R鲁棒且 mmunicatio高效的联邦学习 来自 No非 IID 数据 ata	数据交换优化 h数据交换优化 数据交换优化
P16	Astraea：用于提升移动深度学习应用分类准确率的自平衡联邦学习	客户端资源管理
P17	Co通信 低效 联邦学习 用于 Wirel边缘 In智能 nce 物联网	数据与xchan能优化 化 ion
P18	性能 优化 imiz化 F 联邦学习 rnin网络上的irel无线网络	Clien资源 M管理 ment
P19	通信高效的联邦深度学习：采用逐层异步 模型更新和时间加权聚合	数据交换优化
P20	面向通信高效联邦学习的本地补偿间歇式拉取	数据交 h换优 化
P21	更快的 O n-D设备 训练 使用新的联邦 M动量 m Algorithm	收敛加速 n
P22	BACombo—Ban dwidt h-Aware 去中心化联邦学 习	客户端资源 Man管理
P23	Ternar 与压缩 n for 通信 catio高效 cient F联邦学习 rning	数据交换 h优化 优化
P24	面向通信高效联邦学习的动态采样与选择性掩蔽	数据交 h换优 化
P25	SAFA：一种低开销的快速联邦学习半异步协议	客户端资源 Man理
P26	面向低功耗边缘智能的脉冲神经网络联邦神经形态学习	数据交换优化
P27	面向物联网边缘计算的通信高效联邦学习	数据交换优化
P28	Energy-高效cient R无线资源 Allocatio分配，用于联 邦边缘学习 rning	客户端资源 M管理
P29	客户端-边缘-云 分层 F 联合学习 rning	客户端资源 M管理
P30	Energy-Aware Analog A聚合 io的联邦学习 习 结合冗余 数据	客户端资源 Man管理
P31	收敛 时间最小化 联邦学习 无线 rnin网络 ir无线网络	Clients 资源 M管理 ment

P32	Opt化 g F联邦学习 学习于 No非独立同分布数据 数据与it强化 强制 L学习 g	客户端资源 Man管理
P33	Accelerating F联邦学习 通过 M 学习 动量 m 梯度下降	收敛加速 n
P34	面向高性能、稳定节能联邦学习平台的最优用户选择	客户端资源管理
P35	FedPAQ：一种采用周期性平均和量化的通信高效联邦学习方法	数据交换优化
P36	在移动边缘计算系统中加速与可靠性无关的客户端上的联邦学习	客户端资源管理
P37	Q-GADMM：面向通信高效的去中心化机器学习的量化群组交替方向乘子法	数据交换优化
P38	FedMCCS：面向最优物联网联邦学习的多准则客户端选择模型	客户端资源 Man管理
P39	F取 D: 通信 catio能效 高效联邦学习 rning 与草图化	数据交换 h优化 优化
P40	FedAT：一种在非独立同分布数据下采用异步层级的通信高效联邦学习方法	数据交换优化
P41	惰性 A ggr聚合 Qua ntized Gradient Innovation 用于 Co通信 n-高效 F 联邦式学习	数据交 h换优 化 n
P42	面向物联网边缘计算的通信高效联邦学习	数据交换优化
P43	组 Kn知识 迁移：联 邦学习 大型卷积神经网络在边缘侧	客户端资源 M管理
P44	CatFedAvg：优化联邦学习中的通信效率与分类准确率	客户端资源管理
P45	面向高效联邦学习的自适应梯度稀疏化：一种在线学习方法	数据交换 h范围优化 化
P46	Communicatio高效联邦 蒸馏 蒸馏	数据交换 h优化 优化 n
P47	一种面向物联网设备联邦学习的可信且能量感知的双重深度强化学习调度策略	客户端资源管理
P48	基于时分多址模式的无线网络节能联邦学习设备调度	客户端资源管理
P49	蒸馏-B ased Se mi-Supe rvised Federated Lea训练 for Co mmunicatio n-使用非独立同分布的私人数据进行高效协同训练	数据交换优化
P50	Time-Correl联邦空间 rsifi分配 n f或通信 catio能效 高效联邦学习 rning	数据交换 h能优化 化 n
P51	多接入边缘计算系统中联邦学习的能耗感知资源管理	客户端资源管理
P52	FEDZIP：一种用于通信高效联邦学习的压缩框架	数据交换 hange opt imization
P53	FedProf: Opt imizin g Federated Lea rning with Dynami c Data 资源配置 g	Clien资源 M管理 ment
P54	T迈向资源 高效联邦学习 在移动边缘计算中的学习 移动边缘计算	客户端资源 M管理 ment
P55	HeteroFL：面向异构客户端的计算与通信高效联邦学习	客户端资源管理
P56	无线供能联邦边缘学习：收敛与功率传输之间的最优权衡	数据交换优化
P57	衰落信道中空中联邦学习的梯度统计感知功率控制	Clients 资源 M管理
P58	通过剪枝、 量化和选择性更新加速物联网大数据分析中的联邦学习	收敛加速
P59	Communicatio n 高效联邦学习 效联邦学习 习与 A dapti ve Qua ntization	数据 exc h范围优化 优化 n
P60	资源受限边缘计算中的联邦学习自适应批量大小	Conver能加速 n
P61	采用双侧低秩压缩的通信高效联邦学习	数据交换 h化优化 化
P62	边缘计算中基于分层聚合的资源高效联邦学习	客户端资源 Management
P63	自适应联邦丢弃：提升联邦学习的通信效率和泛化能力	收敛加速
P64	交流还是工作：面向异构移动边缘设备上节能 联邦学习的灵活通信压缩	数据交换优化

P65	面向通信高效联邦学习的模型更新自适应量化	数据交 h换优	化
P66	COFEL：采用本地差分隐私的通信高效优化联邦学习	数据交换优化	
P67	采用通信与计算联合设计的能效联邦边缘学习	客户端资源	Management