

# 面向多层无线边缘环境下的联邦学习通信优化的研究

赵罗成 屈志昊 谢在鹏

河海大学计算机与信息学院 南京 211100

(cheng114yang@hhu.edu.cn)

**摘要** 现有的联邦学习模型同步方法大多基于单层的参数服务器架构,难以适应当前异构无线网络场景,同时存在单点通信负载过重、系统延展性差等问题。针对这些问题,文中提出了一种面向边缘混合无线网络的联邦学习高效模型同步方法。在混合无线网络环境中,边缘移动终端将本地模型传输给附近的小型基站,小型基站收到边缘移动终端模型后执行聚合算法,并将聚合后的模型发送给宏基站完成全局模型更新。考虑到信道性能的异构性和数据传输在无线信道上的竞争关系,文中提出了一种新型的分组异步模型同步方法,并设计了基于传输速率感知的信道分配算法。在真实的数据集上进行了实验,实验结果表明,与传统的模型更新算法相比,所提基于分组异步模型更新的信道分配方法可以缩短 25%~60% 的训练通信时间,大幅度提升了联邦学习的训练效率。

**关键词**: 联邦学习; 异步更新; 信道分配; 异构无线网络; 模型聚合

**中图分类号** TP393.1

## Study on Communication Optimization of Federated Learning in Multi-layer Wireless Edge Environment

ZHAO Luo-cheng, QU Zhi-hao and XIE Zai-peng

School of Computer and Information, Hohai University, Nanjing 211100, China

**Abstract** Existing model synchronization mechanisms of federated learning (FL) are mostly based on single-layer parameter server architecture, which are difficult to adapt to current heterogeneous wireless network scenarios. There are some problems such as excessive communication load on single-point and poor scalability of FL. In response to these problems, this paper proposes an efficient model synchronization scheme for FL in hybrid wireless edge networks. In a hybrid edge wireless network, edge devices transmit local models to nearby small base stations. After receiving local models from edge devices, small base stations execute the aggregation algorithm and send the aggregated models to the macro base station to update the global model. Considering the heterogeneity of channel performance and the competitive relationship of data transmission on the wireless channel, this paper proposes a new type of grouping asynchronous model synchronization scheme and designs a transmission rate aware channel allocation algorithm. Experiments are carried out on real data sets. Experimental results show that the proposed transmission rate aware channel allocation algorithm in grouping asynchronous model synchronization scheme can reduce communication time by 25%~60% and greatly improve the training efficiency of FL.

**Keywords** Federated learning, Asynchronous update, Channel allocation, Heterogeneous wireless network, Model aggregation

### 1 引言

近年来,随着人工智能、无线通信和物联网的快速发展,语音识别、自动驾驶、智慧医疗等智能应用在日常生活和工业场景中都得到了广泛应用<sup>[1-3]</sup>。传统集中式机器学习方法将用户数据传输到云服务器上进行集中式训练,会造成无法容忍的通信成本和时延。为此,Google 于 2016 年提出了联邦学习<sup>[4]</sup>,将边缘计算、人工智能、大数据进行有机融合,旨在充

分利用边缘环境中丰富的计算通信资源,通过多个边缘设备协同训练高精度机器学习模型。联邦学习由于具有易于分布式协作、边缘数据就近处理等优点,被认为是实现边缘智能的关键使能技术,成为了学术界和工业界的热点研究课题。

在联邦学习中,移动终端基于本地数据进行训练,从而生成局部模型,由参数服务器对局部模型进行聚合并对全局模型进行更新,通过多轮迭代将全局模型收敛到预期精度。但是,随着各种智能应用日益复杂,联邦学习模型的参数不断

到稿日期:2021-08-05 返修日期:2021-09-06

基金项目:中央高校业务费(B200202176, B210202079);中国博士后基金面上项目(2019M661709)

This work was supported by the Fundamental Research Funds for the Central Universities(B200202176, B210202079) and Project Funded by China Postdoctoral Science Foundation(2019M661709).

通信作者:屈志昊(quzhihao@hhu.edu.cn)

增多,例如自动驾驶等智能应用模型的参数量达到数百万甚至上亿,在训练过程中会产生大量的计算开销和通信开销。此外,移动终端的硬件资源受限,边缘网络带宽不足。因此,如何在保证模型精度的同时降低计算开销与通信开销,并提升联邦学习的收敛速度<sup>[5-8]</sup>是边缘环境中联邦学习的关键问题。联邦学习的训练速度受两个因素影响:1)模型精度(或者损失)收敛到某一精度时需要的全局迭代轮数;2)每一轮的通信开销和计算开销<sup>[9-11]</sup>。大多数联邦学习的无线通信优化方法,如压缩<sup>[12]</sup>、资源分配<sup>[13]</sup>和节点调度<sup>[14-16]</sup>等,都是建立在参数服务器与移动终端直接相连的单层结构上。联邦学习在单层结构下具有以下缺点:1)参数服务器要与所有参与训练的移动终端通信,存在单点通信负载过重的问题;2)单层结构不适用于现实中复杂的网络架构,延展性不强,无法支撑大规模节点;3)参数服务器与移动终端间的无线链路有高度动态性和异构性<sup>[16]</sup>,而不同移动终端的计算性能有差异,在严格模型同步的方法下,快节点需要等待慢节点完成本地训练和模型传输,才能进行下一轮训练,这严重拖慢了联邦学习的训练进度,造成大量的资源浪费。

混合无线边缘环境下的联邦学习架构一般是由一个宏基站与多个小型基站相连,小型基站再与移动终端相连组成<sup>[17-20]</sup>。混合无线网络有以下优点<sup>[17]</sup>:1)小型基站可以实现蜂窝频谱的高频重用或高密度空间重用,在模型分发上具有高数据率、低丢失率与低时延的优点<sup>[21-22]</sup>;2)小型基站与移动终端的距离小,混合无线网络中同步方案在降低能耗上很有优势;3)小型基站提高了联邦学习系统的拓展性,丰富了参与训练的数据,在提高全局模型精度方面很有潜力。小型基站与移动终端之间的距离短,传输时间相差小,可以采用同步更新策略。不同小型基站的覆盖范围不同,完成局部同步更新的时间差距大,小型基站与宏基站之间可以采用异步更新策略<sup>[23-26]</sup>。由于信道资源的有限性,移动终端和小型基站之间会出现信道竞争问题,需要对信道进行分配,可以减少小型基站的等待时间,提高混合无线边缘环境下联邦学习的收敛速度。

本文研究的是边缘混合无线网络的联邦学习高效模型同步方法,设计了基于传输速率感知的信道分配算法,提高了联邦学习的收敛速度。本文的主要贡献如下:

(1)针对联邦学习在单层结构中的不足,研究了边缘混合无线网络的联邦学习通信优化问题,对信道调度问题进行了形式化描述。

(2)提出了一种新型的分组异步模型同步方法,在小型基站间采用异步更新,并设计了基于传输速率感知的信道分配算法,可以有效地缩短联邦学习的收敛时间。

(3)在真实数据集上进行了实验,实验结果表明,与完全同步的方法相比,所提出的分组异步模型同步方法与信道分配算法可以缩短25%~60%的收敛时间,随着信道数量的增加,缩短的比例提升了10%左右。

## 2 相关工作

无线传输是联邦学习中重要的一部分,移动终端与云

服务器之间频繁的无线通信会产生巨大的通信量,给无线通信环境带来了巨大的负担<sup>[5]</sup>。文献[15]建立了一个分析模型来描述无线网络中联邦学习的性能,并比较了随机调度、循环调度和比例公平3种不同调度策略在不同场景下联邦学习的收敛速度。文献[11]提出了在无线联邦学习中优化收敛时间需要考虑的两个权衡:1)最佳训练精度参数与计算和通信学习的时间;2)训练时间与通信和计算所消耗的本地设备的能量。压缩可以有效地降低传输量,文献[12]研究了压缩方法量化和稀疏化是如何影响训练的联邦学习的收敛速度的。并非所有的设备都需要参与每一轮的联邦学习,文献[6]考虑到无线因素和能量约束,提出了无线资源分配和客户选择联合优化问题。为了找到最优的无线资源分配、用户选择和发射功率,提出了基于匈牙利算法的算法来寻找最优分配策略。文献[13]提出了一种针对不完备信道状态信息的新的客户端调度和资源联合分配策略,并详细分析了模型参数、可用的计算资源与通信资源以及非独立同分布数据对联邦学习性能的影响。异步更新缓解了同步联邦学习算法在同步问题上的开销<sup>[24]</sup>,文献[26]提出了一种以不同频率聚合和更新DNN浅层和深层参数的异步更新策略,减少了服务器和客户端之间需要通信的参数数量,提高了每轮的通信效率。

在联邦学习中,移动终端与云服务器直接通信会造成通信负载过大以及时延过长,而边缘服务器能够与客户端进行更高效的通信,分层联邦学习可以通过高效的边缘更新来减少本地迭代的运行时间与次数<sup>[19]</sup>。文献[19]通过预实验验证了FedAvg扩展到分层之后能够继续收敛,研究了移动终端与边缘服务器以及边缘服务器与云服务器之间两个级别的模型聚合应以怎样的频率进行才能更好地平衡时延与能耗的问题。文献[17]利用小型基站的广播特性,设计了一种部分同步并行方案,小型基站接收到全局参数与梯度之后执行局部模型平均,将结果广播给宏基站和相邻的移动终端,打破了传统的传输顺序,有效地减少了通信负载与时延。文献[18]在分层的基础上考虑了移动终端间的异构性,提出了一种跨异构蜂窝网络的分层联邦学习。为了缩短端到端的延迟,应用了稀疏化和周期平均等分布式高效无线通信技术。文献[20]提出了一种利用局域网的带宽资源的分层联邦学习范式,分析了同一区域内和不同区域间的设备应如何协同训练以及设备之间的拓扑结构对局域网带宽利用率的影响。该框架还可以根据局域网的带宽动态地选择参与训练的设备,以提高训练速度。

## 3 两层无线联邦学习模型

本文考虑一个两层的异构无线网络中的联邦学习场景,其是包括一个宏基站与多个小型基站相连以及小型基站与多个移动终端相连的两层结构。在多层异构无线网络中,多层小型基站的作用是对数据进行中继转发,用一层小型基站代替,保留小型基站与移动终端、小型基站与宏基站的连接即可,因此采用两层的异构无线网络。构建的两层异构无线网络联邦学习的模型如图1所示,包括1个宏基站(MBS),S个小型基站(SBS),N个移动终端,C个信道( $C < N$ )。

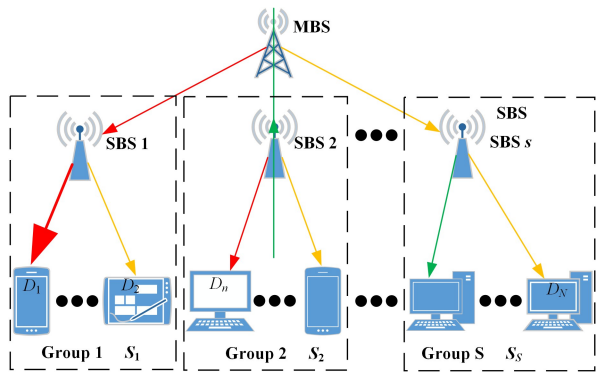


图1 两层无线联邦学习模型

Fig.1 Two-tier FL model in wireless environment

$N$  个移动终端根据地理位置划分为  $S$  组, 每个移动终端  $n$  由相邻的小型基站  $s$  负责, 小型基站  $s$  负责的移动终端集合为  $S_s$ , 移动终端  $n$  的数据集用  $D_n$  表示, 小型基站  $s$  所覆盖的移动终端的总训练数据为  $D_s$ 。

$$D_s = \sum_{n=1}^{S_s} D_n \quad (1)$$

### 3.1 两层无线联邦学习训练过程

在两层无线联邦学习模型中, 宏基站、小型基站与移动终端之间相互协作, 共同完成模型训练。宏基站首先将全局模型广播给所有的小型基站, 小型基站  $s$  再将收到的全局模型转发给移动终端。当移动终端  $n$  本地更新完毕之后, 将生成的本地模型上传给所关联的小型基站  $s$ 。由于小型基站  $s$  与移动终端之间距离近, 因此小型基站与所覆盖移动终端之间采用同步更新策略。小型基站  $s$  收到全部覆盖的移动终端更新后的本地模型后, 对本地模型进行加权平均, 再将结果发送给宏基站。由于组间移动终端的数量、计算能力和通信能力不同, 小型基站一轮同步更新的时间相差较大, 因此宏基站与小型基站之间的更新采用异步更新策略。当小型基站  $s$  执行完模型更新后, 立即与宏基站进行通信, 进行全局模型更新。具体的训练过程如下:

- (1) 时刻  $t$ , 宏基站向小型基站广播全局模型  $\omega_k$ 。
- (2) 小型基站  $s$  向移动终端推送全局模型  $\omega_k$ 。
- (3) 移动终端  $n$  进行本地训练, 生成局部模型  $\omega_k^n$ 。
- (4) 移动终端  $n$  向所关联的小型基站  $s$  发送局部模型  $\omega_k^n$ 。
- (5) 小型基站  $s$  收到移动终端的局部模型后, 根据式(2)更新局部模型为  $\omega_{k,s}$ 。

$$\omega_{k,s} = \frac{1}{D_s} \sum_{n=1}^{S_s} \omega_k^n \quad (2)$$

(6) 小型基站  $s$  将生成的局部平均模型  $\omega_{k,s}$  上传到宏基站。

(7) 宏基站收到小型基站  $s$  的局部平均模型  $\omega_{k,s}$  后, 当前全局模型为  $\omega_{k'}$ , 根据式(3)更新全局模型为  $\omega_{k'+1}$ , 其中  $\alpha$  是混合超参数, 与陈旧性  $k' - k$  有关。

$$\omega_{k'+1} = (1 - \alpha)\omega_{k'} + \alpha\omega_{k,s} \quad (3)$$

- (8) 重复步骤(1)~步骤(7), 直至全局模型收敛。

上述 8 个步骤是两层无线联邦学习的一次迭代过程, 全局模型在多次迭代之后逐渐收敛。本文涉及的符号及其含义如表 1 所列。

表 1 本文中的符号及其说明

Table 1 Notations and their description in this paper

符号	说明	符号	说明
$N$	移动终端总数量	$\omega_{k,s}$	$s$ 在第 $k$ 轮迭代时的平均局部模型
$C$	信道总数量	$\alpha$	混合超参数
$S$	小型基站的数量	$t_n^{cmp}$	$n$ 的训练时间
$n$	移动终端 $n$	$t_n^{com}$	$n$ 的通信时间
$s$	小型基站 $s$	$t_s^{ul}$	$s$ 的通信时间
$S_s$	小型基站 $s$ 负责的移动终端集合	$T_k$	第 $k$ 轮迭代所需的总时延
$D_n$	$n$ 的数据集	$X$	$s$ 与移动终端间的信道分配策略
$D_s$	小型基站 $s$ 覆盖 $n$ 的总数据集	$x_{n,c,t}$	$n$ 在时刻 $t$ 的信道分配结果
$k$	第 $k$ 轮迭代	$Y$	宏基站与 $s$ 的信道分配策略
$\omega_k$	第 $k$ 轮全局模型	$y_{s,c',t}$	$s$ 在时刻 $t$ 的信道分配结果
$\omega_k^n$	移动终端 $n$ 在第 $k$ 轮的局部模型	$T$	收敛所需的时间

### 3.2 无线传输模型

在两层无线联邦学习的一次迭代过程中, 每个步骤都会产生一定的时间延迟。首先, 小型基站与移动终端的距离近, 步骤(2)中的时延可忽略不计。其次, 宏基站和小型基站的计算能力强于移动终端的计算能力, 且仅负责聚合模型以及进行模型平均, 步骤(5)与步骤(7)中的时延可忽略不计。由于移动终端的计算能力与通信能力有限, 因此考虑步骤(3)中的移动终端  $n$  进行本地训练的时延  $t_n^{cmp}$  与步骤(4)中上传局部模型的时延  $t_n^{com}$ , 则:

$$t_n^{com} = \frac{D_n}{A_n} * L_n \quad (4)$$

其中,  $A_n$  表示移动终端  $n$  的计算能力, 假设为常数,  $L_n$  表示进行  $L_n$  轮本地训练。

$$t_n^{com} = \frac{|\omega_k^n|}{r_n} \quad (5)$$

其中,  $|\omega_k^n|$  表示移动终端  $n$  生成的局部模型  $\omega_k^n$  的大小,  $r_n$  表示在香农定理下的传输速率, 则:

$$r_n = W_c \log_2 \left( 1 + \frac{P_n}{N_s} \right) \quad (6)$$

其中,  $W_c$  表示信道  $c$  的带宽,  $P_n$  表示移动终端  $n$  的传输功率,  $N_s$  表示高斯噪声, 假设其为常数。

由于移动终端之间并行做本地训练, 小型基站  $s$  与移动终端之间采用同步更新策略, 因此小型基站  $s$  组内同步更新时间  $t_s^{syn}$  取决于计算与通信耗时最长的移动终端的时间, 则:

$$t_s^{syn} = \max_{n \in S_s} (t_n^{cmp} + t_n^{com}) \quad (7)$$

由于在无线环境中, 上传速率一般小于下载速率, 因此考虑步骤(6)中小型基站将聚合后的局部模型上传到宏基站的时延  $t_s^{ul}$ , 则:

$$t_s^{ul} = \frac{|\omega_{k,s}|}{r_s} \quad (8)$$

其中,  $|\omega_{k,s}|$  为小型基站  $s$  进行局部模型平均后的平均局部模型  $\omega_{k,s}$  的大小,  $r_s$  表示在香农定理下的传输速率。

在两层无线联邦学习的一轮迭代过程中, 仅考虑步骤(4)~步骤(6)所产生的总时延  $T_k$ , 则:

$$T_k = t_s^{syn} + t_s^{ul} \quad (9)$$

## 4 问题定义

在两层无线联邦学习模型中,组间的信道可以复用,每个组内的信道相互干扰。信道的数量  $C$  小于各个组内移动终端的数量  $|S_s|$ ,存在多个移动终端同时竞争同一信道的问题。不同移动终端的性能异构,根据式(6)可知,不同信道的传输速率与时延不同,如果不对信道进行合理地分配,则存在等待时延,时延越长,收敛速度就越慢。设计合理的信道分配策略可以提高收敛速度。对于小型基站  $s$ ,用  $X$  表示小型基站  $s$  与移动终端间的信道分配策略, $X$  中的元素  $x_{n,c,t}$  表示信道分配结果,则:

$$x_{n,c,t} = \begin{cases} 0, & t \text{ 时刻,信道 } c \text{ 未分配给节点 } n \\ 1, & t \text{ 时刻,信道 } c \text{ 分配给节点 } n \end{cases}$$

移动终端与小型基站间采用同步更新策略,则信道分配策略定义为最小化联邦学习一轮迭代过程中局部模型聚合所需时间  $t_s^{\max}$  的问题:

$$\min t_s^{\max} \quad (10)$$

$$\text{s. t. } \begin{cases} \sum_{c=1}^C x_{n,c,t} \leq 1, \forall n, t & (10a) \\ \sum_{n=1}^S x_{n,c,t} \leq C, \forall c, t & (10b) \\ \sum_{t=0}^{t_s^{\max}} x_{n,c,t} r_n \geq |\omega_k^n|, \forall n, c & (10c) \\ t_n^{\text{comp}} \leq t_n^{\text{com}}, \forall n & (10d) \end{cases}$$

其中,约束条件(10a)表示任意时刻  $t$ ,任意移动终端  $n$  最多可以分配到 1 个信道;约束条件(10b)表示任意时刻  $t$ ,所有移动终端至多分配到  $C$  个信道;约束条件(10c)表示任意移动终端  $n$  在信道的可用时间内把局部模型  $\omega_k^n$  传输完毕;约束条件(10d)表示任意的移动终端  $n$  在本地训练完成后再上传。

小型基站与宏基站之间也采用无线传输,其中信道数量  $C'$  小于小型基站的数量  $S$ ,也存在多个小型基站同时竞争同一信道的问题,用  $Y$  表示宏基站与小型基站之间的信道分配策略, $Y$  中的元素  $y_{s,c',t}$  表示信道分配结果,则:

$$y_{s,c',t} = \begin{cases} 0, & t \text{ 时刻,信道 } c' \text{ 未分配给小型基站 } s \\ 1, & t \text{ 时刻,信道 } c' \text{ 分配给小型基站 } s \end{cases}$$

小型基站与宏基站之间采用异步更新策略,则小型基站之间的信道分配策略定义为最小化联邦学习收敛所需时间  $T$  的问题,则:

$$\min T \quad (11)$$

$$\text{s. t. } \begin{cases} \sum_{c'=1}^{C'} y_{s,c',t} \leq 1, \forall s, t & (11a) \\ \sum_{s=1}^S y_{s,c',t} \leq C', \forall c', t & (11b) \\ \sum_{t=0}^{t_{s,k}^{\max}} y_{s,c',t} r_s \geq |\omega_{k,s}|, \forall s, c' & (11c) \\ t_{s,k}^{\text{com}} \geq t_{s,k}^{\text{comp}}, \forall s, k & (11d) \\ t_{s,k+1}^{\text{comp}} \geq t_{s,k}^{\text{com}}, \forall s, k & (11e) \\ 0 \leq k \leq K, & (11f) \end{cases}$$

其中,约束条件(11a)表示任意时刻  $t$ ,任意小型基站  $s$  最多可以分配到 1 个信道;约束条件(11b)表示任意时刻  $t$ ,所有小型基站至多分配到  $C'$  个信道;约束条件(11c)表示任意的小型

基站  $s$  在信道的可用时间内将  $\omega_{k,s}$  传输完;约束条件(11d)表示任意的小型基站  $s$  在第  $k$  轮迭代中在局部模型聚合完成后再上传;约束条件(11e)表示任意的小型基站  $s$  在第  $k+1$  轮迭代中的局部模型聚合在第  $k$  轮迭代过程中的传输完成之后进行;约束条件(11f)表示全部的迭代轮数为  $K$ 。

## 5 两层无线联邦学习信道分配算法

### 5.1 整体系统流程

图 2 为两层无线环境下联邦学习的整体系统流程图,实验与结果分析也将根据该系统流程进行仿真。

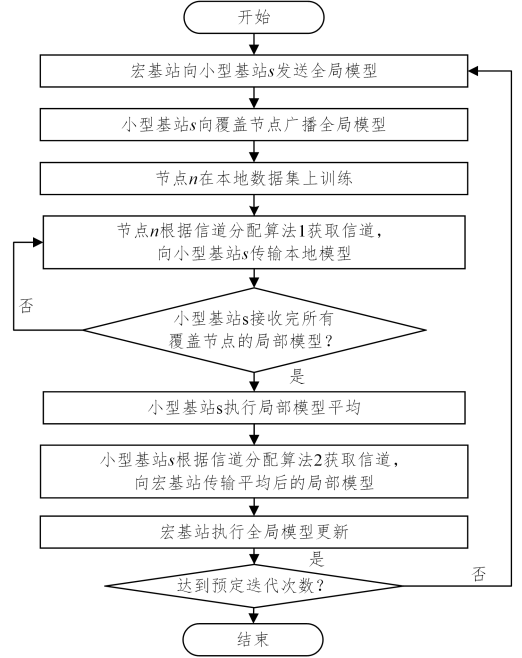


图 2 两层无线联邦学习流程图

Fig. 2 Flow chart of two-tier FL in wireless environments

### 5.2 信道分配算法

在无线边缘环境中,移动终端和小型基站的传输功率以及信道带宽不同。根据式(5)可知,传输速率与传输功率和信道带宽成正比,因此考虑基于传输速率感知的信道分配算法,将信道带宽大的空闲信道优先分配给传输功率大的移动终端  $n$  与小型基站  $s$ ,使得较慢的移动终端与小型基站能够更快地获得信道使用权,从而最小化联邦学习一轮迭代过程中局部模型聚合所需的时间  $t_s^{\max}$  以及联邦学习收敛所需的时间  $T$ ,加快联邦学习的收敛速度。

移动终端与小型基站间执行同步更新策略,信道分配算法 1 的具体流程如下:

- (1)时刻  $t$ ,小型基站  $s$  获取完成本地训练的移动终端集合,按传输功率从大到小排序。
- (2)小型基站  $s$  获取空闲信道集合,按带宽从大到小排序。
- (3)小型基站  $s$  依次将带宽最大的信道  $c$  分配给传输功率最大的移动终端  $n$ ,将信道  $c$  的状态设为繁忙。
- (4)移动终端  $n$  获得信道  $c$  后进行传输。
- (5)移动终端  $n$  传输完毕后,等待小型基站广播新一轮的全局模型,将信道  $c$  的状态设为空闲。

(6)未分配到信道的移动终端等待下一时刻 $t'$ ,重复步骤(1)。

小型基站与宏基站之间执行异步更新策略,信道分配算法2的具体流程如下:

(1)时刻 $t$ ,宏基站获取完成局部模型聚合的小型基站集合,按小型基站的传输功率从大到小排序。

(2)宏基站获取空闲信道的集合,按带宽从大到小排序。

(3)宏基站依次将带宽最大的信道 $c'$ 分配给传输功率最大的小型基站 $s$ ,并将该信道状态设置为繁忙。

(4)获得信道 $c'$ 的小型基站 $s$ 进行传输。

(5)小型基站 $s$ 传输完毕后,将信道 $c'$ 的状态设为空闲,宏基站进行全局模型更新,并将更新后的全局模型回传给 $s$ 。

(6)小型基站 $s$ 进行下一轮迭代。

(7)未分配到信道的小型基站等待下一时刻 $t'$ ,重复步骤(1)。

## 6 实验与结果分析

### 6.1 实验设置

在本文的仿真实验中,基于python设计并实现了两层的无线联邦学习模型。考虑的场景包括一个宏基站、多个小型基站与若干个移动终端。小型基站以同步更新的方式进行局部模型平均,宏基站以异步更新的方式进行全局模型更新。采用的模型为卷积神经网络(CNN),由3个卷积层与2个全连接层组成。移动终端的数据集为手写数字数据集MNIST与普通物体识别的数据集CIFAR。

为了验证本文提出的联邦学习模型同步方法在两层无线网络中的效率,我们针对无线网络传输进行仿真。考虑到基于功率控制的无线传输优化超出了本文的研究范围,本文采用一种简化的无线信道仿真方法,以发送速率和丢包率刻画无线信道传输性能。由于丢包恢复的代价偏大,且部分丢包对联邦学习的收敛影响不大,因此采用丢包后不重传的传输协议。移动终端的数据集独立同分布,本地训练轮数为5,一次训练所选取的样本数为10,学习率为0.01。信道的数量为2或5,信道的传输速率为64 Kbps~1 Mbps不等,丢包率在20%~80%中随机选择。本文对两层无线联邦学习的收敛时间进行了分析,对比实验是移动终端与小型基站以及小型基站与宏基站之间都采用同步更新策略的两层同步无线联邦学习。

### 6.2 实验结果分析

图3—图5给出了两层同步无线联邦学习与本文提出的两层无线联邦学习的损失值随时间的变化情况,横坐标是时间,单位为s,纵坐标是损失值,本文提出的两层无线联邦学习的模型损失值随时间变化用实线表示,两层同步无线联邦学习用虚线表示。其中,图3—图5的数据集与信道数量分别为MNIST与2个信道、CIFAR与2个信道、MNIST与5个信道。图3给出了4组实验的结果,图3(a)—图3(d)给出了有2个信道时,小型基站数量分别为3,5,10和20的情况。图4和图5给出了有5个信道时,小型基站数量分别为10和20的情况。表2—表4所列的结果分别与图3—图5对应,给出了相应图中的模型收敛时间、达到的准确率以及所提方法减少收敛时间的比例。

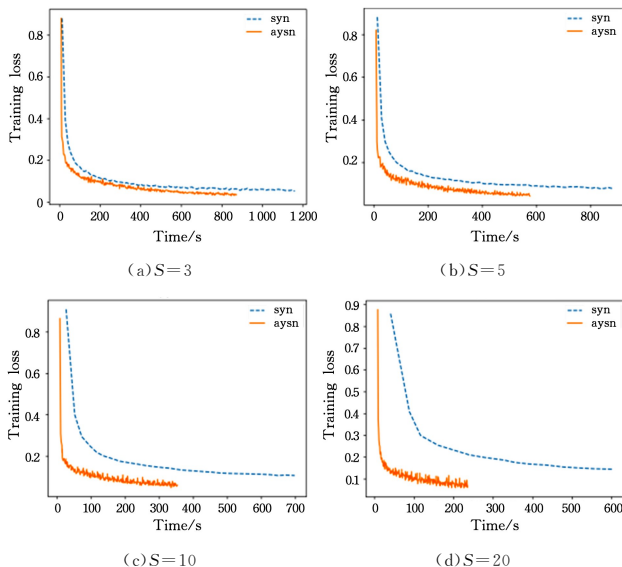


图3 CNN-MNIST-2 模型损失值随时间的变化

Fig. 3 Variation of loss of CNN-MNIST-2 with time

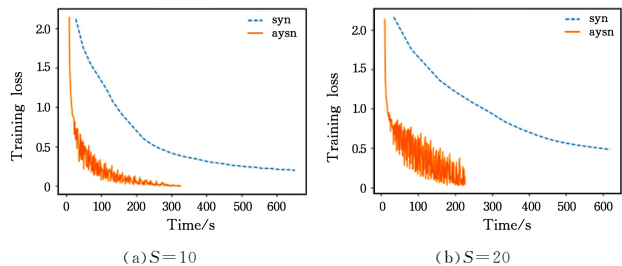


图4 CNN-CIFAR-2 模型损失值随时间的变化

Fig. 4 Variation of loss of CNN-CIFAR-2 with time

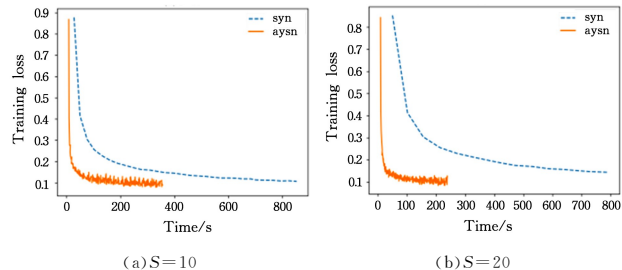


图5 CNN-MNIST-5 模型损失值随时间的变化

Fig. 5 Variation of loss of CNN-MNIST-5 with time

表2 CNN-MNIST-2 的收敛时间与准确率

Table 2 Convergence time and accuracy of CNN-MNIST-2

小型基站 数量/个	同步		异步		收敛时间 缩短/%
	时间/s	准确率/%	时间/s	准确率/%	
3	1158.58	98.8	871.40	98.6	24.8
5	878.28	98.7	575.10	98.9	34.5
10	697.74	98.8	353.25	98.9	49.4
20	601.20	98.6	235.22	99.1	60.9

表3 CNN-CIFAR-2 的收敛时间与准确率

Table 3 Convergence time and accuracy of CNN-CIFAR-2

小型基站 数量/个	同步		异步		收敛时间 缩短/%
	时间/s	准确率/%	时间/s	准确率/%	
10	651.41	55.1	325.93	53.76	50.0
20	618.94	58.3	225.75	58.0	63.5

表4 CNN- MNIST-5的收敛时间与准确率

Table 4 Convergence time and accuracy of CNN-MNIST-5

小型基站 数量/个	同步		异步		收敛时间 缩短/%
	时间/s	准确率/%	时间/s	准确率/%	
10	852.30	98.9	355.10	98.8	58.3
20	791.15	98.7	236.92	99.0	70.1

从图3—图5可以看出,采用异步更新的收敛速度整体上快于同步更新的收敛速度,同样由表2—表4可知,异步更新和同步更新在达到十分相近的准确率时,异步更新的收敛时间要短于同步更新的收敛时间,而且随着小型基站数量的增加,收敛时间越短。

由图3与表2可知,当有3个小型基站与2个信道时,小型基站与信道数量相差不大,异步更新相比同步更新缩短了24.8%的收敛时间。当有5,10与20个小型基站时,小型基站数量与信道数量的差距开始逐渐增大,异步更新可以缩短30%~60%左右的收敛时间。这是因为当信道不能满足多个小型基站同时传输时,同步更新时多个小型基站必须等待其他小型基站占用信道结束后才能使用信道,更新完的小型基站也只能等待其他小型基站更新完毕才能进行下一轮更新,增加了小型基站的等待时间。异步更新时采用基于传输速率感知的信道分配算法,小型基站传输完毕后宏基站立即更新,小型基站不必等待其他小型基站更新就可以进行下一轮更新,缩短了一轮迭代所需的时间。由图4与表3、图3(c)和图3(d)与表2的对比可知,所提两层异步联邦学习与信道分配算法适用于不同的数据集,所达到的缩短联邦学习收敛时间的效果相近。

由图5与表4、图3(c)和图3(d)与表2的对比可知,当信道数量从2增加到5时,两层异步联邦学习提出的两层异步联邦学习与信道分配算法能够进一步减少收敛时间,收敛时间减少比例相比同等数量的小型基站提高了10%左右,进一步提高了收敛速度。因为随着信道数量的增加,两层异步联邦学习能够更加充分地利用信道资源,更多的小型基站可以利用信道与基站进行全局模型更新,缩短了小型基站等待信道的的时间,从而加快了联邦学习的收敛速度。

**结束语** 本文提出了一种两层的无线联邦学习模型与分组异步模型同步方法,小型基站与移动终端之间进行同步局部模型聚合,宏基站与小型基站进行异步全局模型更新。为了减少联邦学习的收敛时间,提出了基于传输速率感知的信道分配算法,该算法可以减少一轮迭代过程中移动终端与小型基站的上传模型以及等待分配信道的的时间,从而加快联邦学习的收敛速度。仿真实验结果表明,在达到相同的模型准确率时,所提信道分配算法使得两层异步无线联邦学习模型在收敛时间上明显少于两层同步的无线联邦学习,且随着小型基站与信道数量的差距变大,收敛时间缩短比例逐渐增大,收敛速度也越快。

未来的研究方向是面向实际应用中更加复杂的场景:1)小型基站可以和宏基站共用无线信道,存在信道干扰的问题;2)每一个移动终端可能同时连接到多个小型基站。在这种情况下,研究如何通过一个精确的模型刻画无线网络的

特征,再对联邦学习的通信进行优化。

## 参考文献

- [1] MOHANTA B K, JENA D, SATAPATHY U, et al. Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology [J]. *Internet of Things*, 2020, 11: 100227-100244.
- [2] LI S, DA X L, ZHAO S. 5G Internet of Things: A survey [J]. *Journal of Industrial Information Integration*, 2018, 10: 1-9.
- [3] QADRI Y A, NAUMAN A, ZIKRIA Y B, et al. The future of healthcare internet of things: a survey of emerging technologies [J]. *IEEE Communications Surveys & Tutorials*, 2020, 22(2): 1121-1167.
- [4] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C] // *Artificial intelligence and statistics*. PMLR, 2017: 1273-1282.
- [5] 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.
- [6] CHEN M, YANG Z, SAAD W, et al. A joint learning and communications framework for federated learning over wireless networks [J]. *IEEE Transactions on Wireless Communications*, 2020, 20(1): 269-283.
- [7] KONČENÝ J, MCMAHAN H B, YU F X, et al. Federated learning: Strategies for improving communication efficiency [J]. *arXiv*: 1610.05492, 2016.
- [8] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C] // *Artificial Intelligence and Statistics*. PMLR, 2017: 1273-1282.
- [9] CHEN M, POOR H V, SAAD W, et al. Convergence time optimization for federated learning over wireless networks [J]. *IEEE Transactions on Wireless Communications*, 2020, 20(4): 2457-2471.
- [10] TRAN N H, BAO W, ZOMAYA A, et al. Federated learning over wireless networks: Optimization model design and analysis [C] // *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019: 1387-1395.
- [11] PILLA L L. Optimal task assignment for heterogeneous federated learning devices [C] // *2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE, 2021: 661-670.
- [12] JIANG P, AGRAWAL G. A linear speedup analysis of distributed deep learning with sparse and quantized communication [C] // *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. 2018: 2530-2541.
- [13] WADU M M, SAMARAKOON S, BENNIS M. Joint client scheduling and resource allocation under channel uncertainty in federated learning [J]. *IEEE Transactions on Communications*, 2021, 69(9): 5962-5974.
- [14] YANG H H, LIU Z, QUEK T Q S, et al. Scheduling policies for

- federated learning in wireless networks[J]. *IEEE Transactions on Communications*, 2019, 68(1): 317-333.
- [15] XIA W, QUEK T Q S, GUO K, et al. Multi-armed bandit-based client scheduling for federated learning[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(11): 7108-7123.
- [16] NISHIO T, YONETANI R. Client selection for federated learning with heterogeneous resources in mobile edge[C]// 2019 IEEE International Conference on Communications (ICC 2019). IEEE, 2019: 1-7.
- [17] QU Z, GUO S, WANG H, et al. Partial synchronization to accelerate federated learning over relay-assisted edge networks[J/OL]. *IEEE Transactions on Mobile Computing*. <https://ieeexplore.ieee.org/abstract/document/9439928>.
- [18] ABAD M S H, OZFATURA E, GUNDUZ D, et al. Hierarchical federated learning across heterogeneous cellular networks[C]// 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 8866-8870.
- [19] LIU L, ZHANG J, SONG S H, et al. Client-edge-cloud hierarchical federated learning[C]// 2020 IEEE International Conference on Communications (ICC 2020). IEEE, 2020: 1-6.
- [20] YUAN J, XU M, MA X, et al. Hierarchical federated learning through lan-wan orchestration[J]. *arXiv*: 2010. 11612, 2020.
- [21] LIM W Y B, NG J S, XIONG Z, et al. Decentralized edge intelligence: a dynamic resource allocation framework for hierarchical federated learning[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2021, 33(3): 536-550.
- [22] LUO S, CHEN X, WU Q, et al. Hfel: Joint edge association and resource allocation for cost-efficient hierarchical federated edge learning[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(10): 6535-6548.
- [23] WU W, HE L, LIN W, et al. SAFA: A semi-asynchronous protocol for fast federated learning with low overhead[J]. *IEEE Transactions on Computers*, 2020, 70(5): 655-668.
- [24] XIE C, KOYEJO S, GUPTA I. Asynchronous federated optimization[J]. *arXiv*: 1903. 03934, 2019.
- [25] CHEN Y, NING Y, SLAWSKI M, et al. Asynchronous online federated learning for edge devices with non-iid data[C]// 2020 IEEE International Conference on Big Data. IEEE, 2020: 15-24.
- [26] CHEN Y, SUN X, JIN Y. Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 31(10): 4229-4238.



**ZHAO Luo-cheng**, born in 1998, post-graduate. His main research interests include distributed computing and federated learning.



**QU Zhi-hao**, born in 1989, assistant researcher. His main research interests include federated learning, cloud edge collaboration and distributed machine learning.

(责任编辑:喻藜)