

学校编码: 10384

学 号: 200215001

厦門大學

博 士 学 位 论 文

资产重组：一个资不抵债国有企业的最佳选择

**Capital Reorganization: The Best Choice for  
A State-Owned Enterprise with Financial Crisis**

作者姓名

指导教师姓名: XXXX 教授

专 业 名 称:

论文提交日期: 年 月

论文答辩日期: 年 月

学位授予日期: 年 月

20 年 月



## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为( )课题(组)的研究成果,获得( )课题(组)经费或实验室的资助,在( )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

本人声明该学位论文不存在剽窃、抄袭等学术不端行为,并愿意承担因学术不端行为所带来的一切后果和法律责任。

声明人 (签名):

指导教师(签名):

年 月 日



## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（        ） 1. 经厦门大学保密委员会审查核定的涉密学位论文，  
于        年    月    日解密，解密后适用上述授权。

（        ） 2. 不涉密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。涉密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年    月    日



## 学位论文答辩委员会名单

主席	赵 XX	厦门大学	教授	博士生导师
委员	刘 XX	厦门大学	教授	博士生导师
	杨 XX	中国 XXXX 科学院 XXXXXXX 研究所	研究员	博士生导师
	黄 XX	XXXX 大学	教授	博士生导师
	周 XX	XXXX 大学	副教授	硕士生导师
秘书	吴 XX	厦门大学	助理教授	-

答辩时间：2020 年 5 月 10 日上午

答辩地点：厦门大学海韵园行政 C-505 会议室





## 摘 要

论文的摘要是对论文研究内容和成果的高度概括。摘要应对论文所研究的问题及其研究目的进行描述，对研究方法和过程进行简单介绍，对研究成果和所得结论进行概括。摘要应具有独立性和自明性，其内容应包含与论文全文同等量的主要信息。使读者即使不阅读全文，通过摘要就能了解论文的总体内容和主要成果。论文摘要的书写应力求精确、简明。切忌写成对论文书写内容进行提要的形式，尤其要避免“第 1 章……；第 2 章……；……”这种或类似的陈述方式

**关键词：**形状记忆；聚氨酯；织物；合成；应用（一般选 3~8 个单词或专业术语，且中英文关键词必须对应。）



## **Abstract**

An abstract of a dissertation is a summary and extraction of research work and contributions. Included in an abstract should be description of research topic and research objective, brief introduction to methodology and research process, and summarization of conclusion and contributions of the research. An abstract should be characterized by independence and clarity and carry identical information with the dissertation. It should be such that the general idea and major contributions of the dissertation are conveyed without reading the dissertation.

**Keywords:** shape memory properties; polyurethane; textile; synthesis; application



## 目 录

摘 要 .....	I
Abstract .....	II
目 录 .....	III
主要缩略词表 .....	VII
主要符号表 .....	VIII
图索引 .....	IX
表索引 .....	X
第 1 章 绪论 .....	1
1.1 研究意义及问题 .....	1
1.2 国内外研究现状 .....	2
1.3 研究内容和主要贡献 .....	4
1.4 本文主要内容和章节安排 .....	5
1.4.1 形状记忆聚氨酯的研究进展 .....	5
1.4.2 水系聚氨酯及聚氨酯整理剂 .....	6
第 2 章 基于服务端中心计算的联邦学习框架 .....	7
2.1 问题描述 .....	7
2.2 联邦学习框架 .....	10
2.2.1 两阶段模型分割训练策略 .....	10
2.2.2 模型分割方法 .....	10
2.2.3 模型参数更新策略 .....	11
2.2.4 动态批次组建 .....	11
2.2.5 数据异构解决方案 .....	12
2.2.6 历史特征交换采样 .....	13
2.3 实验设置 .....	13
第 3 章 总结与展望 .....	14

参考文献 .....	15
附录 A 关于 XXX 的证明 .....	19
附录 B Maxwell Equations .....	20
致 谢 .....	21
在学期间完成的相关学术成果 .....	22

## Contents

<b>Abstract-Chinese</b>	I
<b>Abstract-English</b>	II
<b>Contents</b>	III
<b>List of Acronyms</b>	VII
<b>List of Symbols</b>	VII
<b>List of Figures</b>	IX
<b>List of Tables</b>	X
<b>Chapter 1 Introduction</b>	1
<b>1.1 Research Significance and Problems</b>	1
<b>1.2 Research Progress Overview in Home and Abroad</b>	2
<b>1.3 Research Content and Main Contributions</b>	4
<b>1.4 Major Contents and Chapter Arrangement</b>	5
1.4.1 Research Progress of Shape Memory Polyurethane	5
<b>2.1 Problem Description</b>	7
<b>2.2 Federated Learning Framework</b>	10
2.2.1 Model Scalability Solution	10
2.2.2 Model Splitting Method	10
2.2.3 Model Parameter Update Strategy	11
2.2.4 Dynamic Batching	11
2.2.5 Data Heterogeneity Solution	12
2.2.6 Stable Feature Fusion	13
<b>2.3 Experimental Settings</b>	13
<b>References</b>	15
<b>Appendix A Proof of XXX</b>	19
<b>Acknowledgements</b>	21

<b>Relevant Academic Achievements Completed During the Academic Period .....</b>	<b>22</b>
--	-----------



## 主要缩略词表

缩写	英文全称	中文含义
FFT	Fast Fourier Transform	快速傅里叶变换



## 主要符号表

符号	中文含义
$c$	光速



## 图索引

图 1.1	热塑性形状记忆聚氨酯的形状记忆机理示意图 .....	5
图 2.1	MedFed 联邦学习框架模型分割图 .....	11
图 2.2	MedFed 联邦学习框架效率图 .....	12
图 2.3	MedFed 联邦学习框架结构图 .....	13



## 表索引

表 1.1 水系聚氨酯分类 .....	6
---------------------	---





## 第 1 章 绪论

### 1.1 研究意义及问题

医学图像获取依赖于多种成像技术，包括计算机断层扫描（CT）<sup>[1-2]</sup>、磁共振成像（MRI）、放射性核素扫描和超声波等。这些技术各自揭示了生物组织的独特特征和信息，但也增加了图像的复杂性和多样性。CT 成像基于 X 射线，通过捕捉不同组织对 X 射线的吸收差异来生成图像。与此相比，MRI 利用磁场和无损耗的非离子辐射，对组织中的水分和脂肪进行成像，展现不同的组织特性。这些多样的成像方法提供了不同视角的生物组织图像，有助于更全面地了解人体内部结构和状况。医学图像分割是一项关键任务，旨在从整体医学图像中精确提取和标记出特定的结构或区域。这包括对器官、病变、血管等目标的识别和界定。这一过程对于深入理解和分析医学图像至关重要，它为医学研究、疾病诊断和治疗规划提供了关键支持。医学图像分割在定量分析和计算机辅助诊断系统中扮演着核心角色，使得对复杂医学图像的解读更加精确和高效。然而医学图像分割的数据标注过程通常需要耗费大量的人力和时间，且标注结果的质量和准确性往往受到标注者的经验和主观因素的影响。此外，医学图像的数据通常涉及隐私和安全问题，这也使得数据的共享和使用受到了严格的限制。因此，如何利用有限的标注数据并且保护数据隐私的情况下，实现医学图像分割的高效和精确成为了一个重要的研究问题。

为了在医学图像分割任务中提高效果并保护隐私，联邦学习提供了一种创新的解决方案。这种分布式学习方法的核心优势在于它允许多个数据持有者在不共享数据的前提下共同训练模型。虽然联邦学习一定程度上缓解了数据不足的问题，但在面对标注数据不足的挑战时，还可以采用自监督学习、半监督学习和迁移学习等策略。结合联邦学习和这些学习策略的应用，不仅能有效利用有限的标注数据提升模型性能，这还与 OpenAI 提出的模型 *scaling laws* 相契合，因为根据这些定律，大规模预训练模型往往需要大量的数据，而为大量数据进行标注在实际操作中并不可行。因此，这种结合联邦学习和以上各种学习策略的方法已成为实现医学图像分割的高效、精确且成本效益高的重要途径。

然而，该方向上仍存在以下亟待研究的关键问题：\* 在模型规模日趋庞大的情况下，之前的联邦学习框架都不能实现大规模模型的高效训练。这是因为联邦学

习的核心特点是在不共享数据的前提下进行模型训练，这就意味着在联邦学习框架下，模型的训练数据是分布在多个数据持有者的设备上的，而这些设备的计算能力和存储能力往往是有限的。因此，如何在这种情况下实现高效的模型训练是一个重要的研究问题。\* 由于未标注数据和标注数据，不同的联邦学习参与者的数据分布往往是不同的，对模型性能有着不同程度的损害。如何消除这些非独立同分布的数据对模型训练的影响也是一个重要的研究问题。\* **scaling laws** 在大量的其他领域被验证是正确的，如何利用这一规律来指导医学图像分割模型的训练是一个重要的研究问题。因为对于医学图像分割这一领域而言，对于精度的要求往往是非常高的，因此扩大模型的规模来获得更好的性能受益是非常有意义的。

## 1.2 国内外研究现状

传统的集中式学习要求将在本地设备（如手机）上收集的所有数据集中存储在数据中心或云服务器上。这一要求不仅提出了隐私风险和数据泄露的担忧，而且当数据量巨大时，对服务器的存储和计算能力提出了高要求。尽管分布式数据并行<sup>[3]</sup>，它允许多台机器并行训练具有不同数据组的模型副本，可能作为解决存储和计算能力问题的潜在解决方案，但它仍需要访问整个训练数据以将其分割成均匀分布的片段，导致可能对数据的安全和隐私问题。

联邦学习（FL）旨在训练一个全局模型，该模型可以在不同设备上分布的数据上进行训练，同时保护数据隐私。2016年，McMahan 等人<sup>[4]</sup>首次介绍了基于数据并行性的 FL 概念，并提出了一种联邦平均（FedAvg）算法。作为一种去中心化的机器学习方法，FedAvg 允许多个设备合作训练机器学习模型，同时保持用户数据本地存储。FedAvg 免除了将用户的敏感数据上传到集中式服务器的需要，并使得边缘设备能够在自己的本地数据集内部本地训练共享模型。通过聚合本地模型的更新（梯度），FedAvg 满足了隐私保护和数据安全的基本要求。

虽然 FL 提供了一种保护隐私的有希望的方法，但与集中式学习相比，当 FL 应用于现实世界时，许多挑战随之而来<sup>[5]</sup>。这些挑战包括传输参数所需的通信成本、本地设备所需的计算能力和能源消耗，以及学习过程中可能大量本地设备的异质性和随机性。大量的研究致力于解决上述挑战，包括减少通信成本<sup>[4,6-8]</sup>，考虑硬件约束的 FL<sup>[9-10]</sup>，以及针对对抗性攻击的额外保护<sup>[11-13]</sup>。

尽管<sup>[4]</sup>中的作者声称 FedAvg 能够在一定程度上应对非独立同分布（Non-IID）数据，但许多研究已经指出，在 Non-IID 或异构数据上 FL 的准确性几乎不可避免

地会恶化<sup>[14]</sup>。性能下降主要可以归因于由于 Non-IID 导致的本地模型的权重分歧。也就是说，具有相同初始参数的本地模型会由于本地数据分布的异质性而收敛到不同的模型。在 FL 过程中，通过

平均上传的本地模型获得的共享全局模型与理想模型（当本地设备上的数据是 IID 时获得的模型）之间的差异持续增加，减缓了收敛速度并恶化了学习性能。

由于对 FL 的研究兴趣迅速增加，文献中已经发表了几篇有价值的 FL 综述论文。在<sup>[5,15]</sup>中给出了 FL 及其应用的一般介绍，<sup>[11,16]</sup>中详细讨论了进展和挑战。在<sup>[17-20]</sup>中介绍了 FL 中的威胁分析和额外的隐私保护技术。也报告了 FL 应用于物联网和边缘设备<sup>[21-23]</sup>、无线网络<sup>[24]</sup>、移动设备<sup>[10]</sup>和医疗保健<sup>[25]</sup>的概述。

尽管 Kulkarni 等人<sup>[26]</sup>提供了对处理 Non-IID 数据的个性化方法的简要介绍，但现有工作没有详细探讨 Non-IID 数据对 FL 的影响。为了填补这一空白，本文致力于对 Non-IID 数据上的 FL 进行全面调查，包括对各种数据分布的深入分析、它们对模型聚合的影响、现有处理偏斜数据分布技术的分类及其优缺点的讨论，以及 FL 面临的 Non-IID 数据的剩余挑战和未来研究的概述。

FL 是一个快速发展的主题，我们在这里只讨论与之密切相关的方法。关于这一领域的综合性研究已经在文献<sup>[11,27]</sup>中出现。通常的 FL 设置包括两种类型的更新，即服务器和设备更新，每种更新都与最小化某些局部损失函数相关，这个函数本身可能会在不同的轮次中动态更新。在任何轮次中，都有方法尝试进行完全优化，或者提出不精确优化。我们特别关注那些解决四种 FL 场景（大规模分布、异质性、不可靠链接和不平衡数据）的相关工作。

一系列工作提出基于局部 SGD<sup>[28]</sup>的更新，其中每个参与设备执行单个局部 SGD 步骤。然后服务器平均收到的模型。与局部 SGD 相比，我们的方法提出最小化一个局部惩罚的经验损失。

shortfedAvg<sup>[4]</sup>是局部 SGD 的一个泛化，它提出了每轮更多的局部 SGD 步骤。尽管如此，shortfedAvg 在设备端优化的解决方案是不精确的。确定何时停止最小化，以便获得良好的准确性-通信权衡，是基于调整轮次数和学习率<sup>[4,29]</sup>。尽管 shortfedAvg 在 IID 设置中具有强大的实证性能，但在非 IID 场景中性能下降<sup>[14]</sup>。

为了处理非 IID 设置，提出了 shortfedAvg 的几种修改版本。这些变体包括使用递减的学习率<sup>[29]</sup>；动态修改设备经验损失<sup>[30]</sup>；或修改服务器端更新<sup>[31-32]</sup>。使用递减学习率或自定义服务器端更新的方法仍然依赖于设备内的局部 SGD 更新。虽然这些工作确实认识到局部和全局静止点的不兼容性，但他们提出的解决方案是

基于不精确的最小化。此外，为了在非 IID 情况下建立收敛性，这些工作施加了额外的“有界非 IID”条件。

$\text{shortfedProx}^{[30]}$  与我们的方法相关。像我们一样，他们提出了一个动态调节器，这个调节器基于服务器提供的模型进行修改。这个调节器惩罚与服务器模型相距较远的更新。尽管如此，结果中的调节器并没有使全局和局部静止点对齐，因此需要不精确的最小化，并且他们通过仔细选择学习率和轮次来实现。此外，调整需要一些关于统计异质性的知识。

在类似的方向上，有工作通过添加额外的设备变量来增强更新，这些变量也与模型一起传输<sup>[33-34]</sup>。这些工作通过添加设备依赖的调节器来证明收敛保证。尽管如此，他们遭受额外的通信成本，并且他们没有在深度神经网络上进行广泛的实验。其中， $\text{scaf}^{[33]}$  是一个与我们工

作密切相关的工作，尽管它传输了额外的变量，并且在第??节中给出了更详细的比较。

另一系列分布式优化方法<sup>[34-41]</sup>可以在这种设置中考虑。此外，有工作将 SGD 类型方法的分析扩展到 FL 设置<sup>[42-44]</sup>。然而，这些算法是为全设备参与案例提出的，这不符合 FL 的一个重要方面。 $\text{FedSVRG}^{[35]}$  和  $\text{DANE}^{[34]}$  需要在每一轮从所有设备获取梯度信息，并且它们不能直接应用于部分 FL 设置。例如， $\text{FedDANE}^{[45]}$  是 DANE 的一个版本，适用于部分参与。然而， $\text{FedDANE}$  在部分参与情况下的实证表现不如  $\text{FedAvg}^{[45]}$ 。与这些工作类似， $\text{FedPD}^{[46]}$  方法是在不同参与概念下提出的分布式优化方法。 $\text{FedPD}$  在每一轮要么激活所有设备要么不激活任何设备，这再次不满足 FL 中的部分参与。

最后，另一组工作旨在通过压缩传输模型来减少通信成本<sup>[47-49]</sup>。他们通过降低传输的比特率来节省通信成本。这些想法是我们工作的补充，它们可以集成到我们提出的解决方案中。

### 1.3 研究内容和主要贡献

因此本文的主要贡献可以总结如下：

- (1) 提出了一种新的联邦学习框架，该框架可以在不共享数据的前提下实现大规模模型的高效训练。
- (2) 设计了一种数据分布自适应的学习方法，该方法可以消除非独立同分布数据对模型训练的影响。

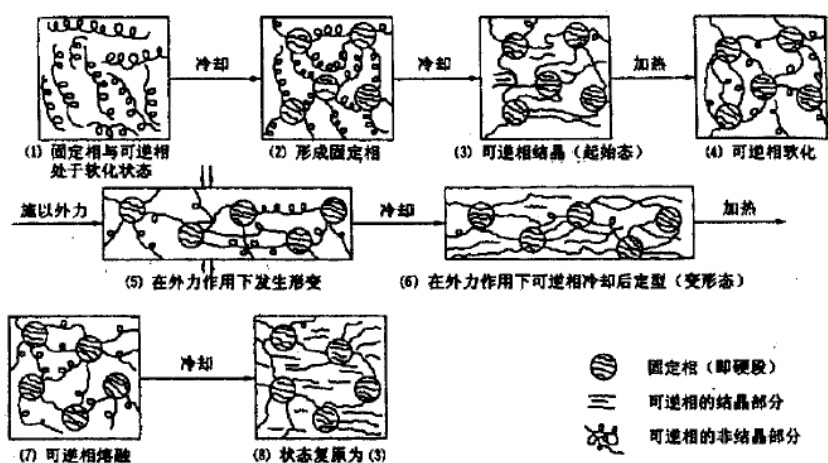


图 1.1 热塑性形状记忆聚氨酯的形状记忆机理示意图

数据来源：交代图形的出处或者来源，表明“作者、来源名称、时间”，用小五宋体，置图左下方。交代图形的出处或者来源，表明“作者、来源名称、时间”，用小五宋体，置图左下方。

(3) 提出了 SimLoss，一种新的损失函数，该损失函数可以在大规模模型训练中提高模型的性能。

(4) 设计了一种新的 router 路由方式，大大减少了 moe 中 drop token 的数量。

## 1.4 本文主要内容和章节安排

形状记忆聚合物（SMP）是继形状记忆合金后在 80 年代发展起来的一种新型形状记忆材料 [50]。形状记忆高分子材料在常温范围内具有塑料的性质，即刚性、形状稳定恢复性；同时在一定温度下（所谓记忆温度下）具有橡胶的特性，主要表现为材料的可变形性和形变恢复性。即“记忆初始态—固定变形—恢复起始态”的循环。

固定相只有物理交联结构的聚氨酯称为热塑性 SMPU，而有化学交联结构称为热固性 SMPU。热塑性和热固性形状记忆聚氨酯的形状记忆原理示意图如图??所示

### 1.4.1 形状记忆聚氨酯的研究进展

首例 SMPU 是日本 Mitsubishi 公司开发成功的……。

表 1.1 水系聚氨酯分类

类别	水溶型	胶体分散型	乳液型
状态	溶解 ~ 胶束	分散	白浊
外观	水溶型	胶体分散型	乳液型
粒径 / $\mu\text{m}$	$< 0.001$	$0.001 - 0.1$	$> 0.1$
重均分子量	$1000 \sim 10000$	数千 ~ 20万	$> 5000$

#### 1.4.2 水系聚氨酯及聚氨酯整理剂

水系聚氨酯的形态对其流动性，成膜性及加工织物的性能有重要影响，一般分为三种类型 [50]，如表 1.1所示。

由于它们对纤维织物的浸透性和亲和性不同，因此在纺织品染整加工中的用途也有差别，其中以水溶型和乳液型产品较为常用。另外，水系聚氨酯又有反应性和非反应性之分。虽然它们的共同特点是分子结构中不含异氰酸酯基，但前者是用封闭剂将异氰酸酯基暂时封闭，在纺织品整理时复出。相互交联反应形成三维网状结构而固着在织物表面。……

然而索引扩散并不总是有效率的，它也会带来带宽开销。一方面，扩散更多的索引可以使搜索更快地返回，减少了搜索带宽开销；另一方面，由于 P2P 中结点和数据处于不断动态变化之中，当数据失效或更新时（如结点离线、删除或更新数据），数据的索引也相应失效，必须加以更新维护。因此，扩散更多的索引意味着维护开销的增加。于是在带宽开销方面，搜索开销与索引维护开销之间存在着折衷关系（trade-off）。与以往工作中仅考虑搜索开销不同，本章的模型中我们同时考虑搜索和维护两方面，给出了索引扩散方法对搜索整体性能的影响和数学关系。通过模型我们发现索引数量是决定宽松约束一般性搜索性能的至关重要的因素，采用最优索引分布可以很大程度上提高性能，降低系统开销。与一般认为的 P2P 无偏向性搜索难于扩展（non-scalable）恰恰相反，模型显示在最优的索引扩散策略下，基于无偏向性搜索具备很好的可扩展性，其结点负载和带宽开销随系统规模  $N$ （结点数）增长具有  $O$  的增长关系。这种平方根关系保证了对大规模 P2P 系统很好的适应性。

快速傅里叶变换（Fast Fourier Transform, FFT）是离散傅立叶变换的一种快速算法。光速的符号定义为  $c$ 。

## 第 2 章 基于服务端中心计算的联邦学习框架

### 2.1 问题描述

所有的联邦学习框架的核心思想都是在不共享数据的前提下进行模型训练。这就意味着在联邦学习框架下，模型的训练数据是分布在多个数据持有者的设备上的，而这些设备的计算能力和存储能力往往是有限的。存在严重的算力不均衡的问题，这就导致了在联邦学习框架下，模型的训练效率往往是非常低的。这就是联邦学习中的计算异构问题。为了解决这一问题，我们提出了一种新的联邦学习框架（MedFed），该框架可以在不共享数据的前提下实现大规模模型的高效训练。

同时联邦学习还有一个重要的问题是非独立同分布数据的问题。由于未标注数据和标注数据，不同的联邦学习参与者的数据分布往往是不同的，对模型性能有着不同程度的损害。为了解决联邦学习中的数据异构问题，之前已经有了一些工作，比如 FedAvg<sup>[51]</sup>、FedProx<sup>[52]</sup>、FedOpt<sup>[?] ]</sup>、FedNova<sup>[?] ]</sup>、FedMA<sup>[?] ]</sup>、FedPD<sup>[?] ]</sup>、FedMGDA<sup>[?] ]</sup>等。

相关工作。联邦学习是一个快速发展的主题，我们只在这里描述与之密切相关的方法。关于联邦学习的综合研究已经出现在 Kairouz 等人（2019）；Li 等人（2020）的文献中。一般的联邦学习设置涉及两种类型的更新，即服务器端和设备端，每种更新都与最小化某种局部损失函数相关联，该损失函数本身可以在不同轮次动态更新。在任何一轮中，有些方法试图完全优化，而其他方法则提出了近似优化。我们专注于解决四种联邦学习场景（大规模分发、异质性、不可靠链接和数据不平衡）的相关工作。一种工作提出了基于本地 SGD（Stich, 2019）的更新，其中每个参与设备执行单个本地 SGD 步骤。然后服务器平均接收到的模型。与本地 SGD 相比，我们的方法提出了最小化本地惩罚经验损失。FedAvg（McMahan 等人，2017）是本地 SGD 的一般化，提出了每轮更多的本地 SGD 步骤。尽管如此，FedAvg 仍然不完全解决了设备端的优化问题。确定何时停止最小化以获得良好的精度-通信折衷是基于调整时期数和学习率的（McMahan 等人，2017；Li 等人，2020b）。尽管 FedAvg 在 IID 设置中具有较强的经验性能，但在非 IID 场景中性能下降（Zhao 等人，2018）。已经提出了几种修改 FedAvg 以处理非 IID 设置的方法。这些变体包括使用递减的学习率（Li 等人，2020b）；动态修改设备经验损失（Li 等人，2020a）；或修改服务器端更新（Hsu 等人，2019；Reddi 等人，2020）。使用递减学习率或

定制服务器端更新的方法仍然依赖于设备内的本地 SGD 更新。虽然这些工作确实承认了本地和全局稳定点的不兼容性，但其提出的修复方法基于近似最小化。此外，为了在非 IID 情况下建立收敛性，这些工作还施加了额外的“有界非 IID”条件。FedProx (Li 等人, 2020a) 与我们的方法相关。与我们一样，他们提出了一种动态正则化器，该正则化器基于服务器提供的模型进行修改。然而，由此产生的正则化器并不能使全局和本地稳定点对齐，因此需要进行近似最小化，他们通过精心选择学习率和时期来实现这一点。此外，调整需要对统计异质性有一定的了解。类似的，还有一些工作将更新与额外的设备变量相结合，这些变量也与模型一起传输 (Karimireddy 等人, 2019; Shamir 等人, 2014)。这些工作通过添加设备相关的正则化器来证明收敛性保证。然而，它们会增加额外的通信成本，并且在深度神经网络方面没有进行广泛的实验。其中，SCAFFOLD (Karimireddy 等人, 2019) 是一项与本文密切相关的工作，尽管它传输了额外的变量，但在第 2 节中给出了更详细的比较。另一类分布式优化方法 (Konecny 等人, 2016; Makhdoumi & Ozdaglar, 2017; Shamir 等人, 2014; Yuan & Ma, 2020; Pathak & Wainwright, 2020; Liang 等人, 2019; Li 等人, 2020c; Condat 等人, 2020) 可以在这种设置中考虑。此外，还有一些工作将 SGD 类型方法的分析扩展到 FL 设置中 (Gorbunov 等人, 2020; Khaled 等人, 2020b; Li & Richtarik, 2020)。然而，这些算法是为全设备参与的情况提出的，未能满足 FL 的一个重要方面。FedSVRG (Konecny 等人, 2016) 和 DANE (Shamir 等人, 2014) 需要每轮从所有设备获取梯度信息，因此不直接适用于部分 FL 设置。例如，FedDANE (Li 等人, 2019) 是 DANE 的一个版本，可在部分参与中运行。然而，FedDANE 在部分参与的情况下实验性能比 FedAvg 差 (Li 等人, 2019)。与这些工作类似，FedPD (Zhang 等人, 2020) 方法是在不同参与方式的分布式优化中提出的。FedPD 每轮激活所有设备或不激活任何设备，这也不能满足 FL 的部分参与要求。最后，另一类工作旨在通过压缩传输模型来减少通信成本 (Dutta 等人, 2019; Mishchenko 等人, 2019; Alistarh 等人, 2017)。通过降低传输的比特率来节省通信成本。这些想法与我们的工作互补的，它们可以集成到我们提出的解决方案中。

与此同时，联邦学习还有一个问题是说，不同的客户端的通信带宽和延迟是不同的。这就导致了在联邦学习框架下，模型的训练效率往往是非常低的。这就是联邦学习中的通信异构问题。为了解决这一问题，我们提出了一种新的联邦学习框架 (MedFed)，该框架可以在不共享数据的前提下实现大规模模型的高效训练。



常见的联邦学习系统通常都是以下三种架构：1) 服务端-客户端架构；2) 点对点架构；3) 序列式架构；其中后两者在当下模型数据规模日趋庞大的情况下，都无法突破模型训练中的内存墙。因为这两种方式，都要求每个模型存储完整的模型，并在本地进行模型训练，这就意味着，每个模型都需要存储完整的模型参数和中间的激活值。具体来说，随着模型大小的增加，存放单个模型所需要的空间也越大。而每个医院（客户端）的存储空间是有限的。简单地说，想要训练一个模型，我们需要存储模型的参数和中间的激活值。在机器数量固定的情况下，从极端的情况来说，中间的激活值可以同过反向过程中的重计算<sup>[53]</sup>来减少存储空间，但是模型的参数是更难被减少的，这是因为，在机器数量不增加的情况下，当下的减少模型存储空间的办法主要有 offload<sup>[54-56]</sup>，但是受限于 pcie 的带宽限制，这种方法往往以牺牲训练效率为代价。下面，我们简单计算一下存储一个模型需要的空间大小。通常每一个模型参数对应着三个部分的存储，分别是模型权重，模型梯度和优化器状态。我们用  $P_{\text{model}}$  表示模型参数的大小， $P_{\text{grad}}$  表示模型梯度的大小， $P_{\text{opt}}$  表示优化器状态的大小。那么，存储一个模型需要的空间大小为  $P_{\text{model}} + P_{\text{grad}} + P_{\text{opt}}$ 。下面，我们将分别计算模型在 float32 训练和 float16 训练下（bfloat16 同），存储一个模型需要的空间大小。设模型参数数量为  $N$ ，float32 下一个参数占用 4 个字节，梯度占用 4 个字节，优化器状态占用 8 个字节（ADAM 优化器需要存储动量和方差），因此，float32 下存储一个模型需要的空间大小为  $N \times 4 + N \times 4 + N \times 8 = 16N$ 。而在 float16 下，一个参数占用 2 个字节，梯度占用 2 个字节，优化器状态则需要占用 16 个字节（包括高精度的参数、梯度的副本和动量、方差），因此，float16 下存储一个模型需要的空间大小为  $N \times 2 + N \times 2 + N \times 16 = 20N$ 。假设我们要训练一个 100B 的分割模型，那么在 float32 下，存储一个模型需要的空间大小为 1600G，而在 float16 下，存储一个模型需要的空间大小为 2000G。这是绝大多数医院都无法承受的存储空间。

因此，我们需要一种新的联邦学习框架，这个联邦学习框架需要满足以下几个条件：1) 拥有良好的扩展性，能够适应不同大小的模型；2) 尽可能消除非独立同分布数据和长尾数据对模型训练的影响；3) 尽可能消除不同医院之间的计算异构和通信异构，实现模型的高效训练，满足当下模型快速迭代的需求；4) 尽可能保护患者隐私，保证数据的安全性。

## 2.2 联邦学习框架

如图 3.1 所示，我们提出了一种新的联邦学习框架（MedFed），该框架主要包括 4 个模块，分别是：初始化模块、特征交换模块、模型训练更新模块和弹性计算模块。

我们首先着重介绍一下模型训练更新模块。模型训练更新模块是联邦学习框架的核心模块，和以往的联邦学习架构不同的是，本文所提出的 MedFed 框架旨在保护患者隐私的情况下，尽可能实现模型的高效训练，而要解决模型的高效训练，首先要解决的就是计算异构和通信异构的问题。计算异构指的是不同客户端的计算能力和存储能力是不同的，通信异构指的是不同客户端的通信带宽和延迟是不同的。前者导致了不同联邦学习客户端的模型训练是存在着巨大的差距的，后者则会导致计算资源的巨大浪费，因为为了保证计算的正确性，模型在更新的同时是不能继续训练的。根据上文提出的四个条件，我们将逐一介绍我们是如何在 MedFed 框架下解决这些问题的。

### 2.2.1 两阶段模型分割训练策略

在服务端-客户端的联邦学习框架下，如何设计模型在服务端和客户端上的分布，和客户端与服务端需要实现的功能，是框架设计的核心要素，本节旨在阐述模型的分割方法，与模型的参数更新的策略。

### 2.2.2 模型分割方法

正如前文所提到的，现有的联邦学习框架，即使采用服务端-客户端的实现架构，也仅仅只是把参数的更新这一步放到服务端进行。而没有考虑每一个客户端的实际存储实际上是非常受限的，这个和数据中心动辄数千卡的存储空间是不可同日而语的。在训练一个大规模模型的时候，我们需要存储模型的参数、梯度和优化器状态，如果要在每一个客户端都完整地存储以上内容，这无疑将绝大多数的医院排除在外，也不符合人工智能应当普惠的初衷。为了缩减每一个客户端的存储压力，在我们的 medfed 框架下，我们将模型切分成两个部分，一部分模型较小，存储在每一个医院客户端，一部分模型较大，存储在服务端。也就是说，我们将模型的大部分参数存储在数据中心，而将模型的一小部分参数存储在每一个医院客户端。这种分布式的存储方式类似于流水线并行<sup>[57-58]</sup>，这种方式的灵活性在于，我们可以轻易地调整分配到医院客户端的模型参数的大小，以适应不同大小

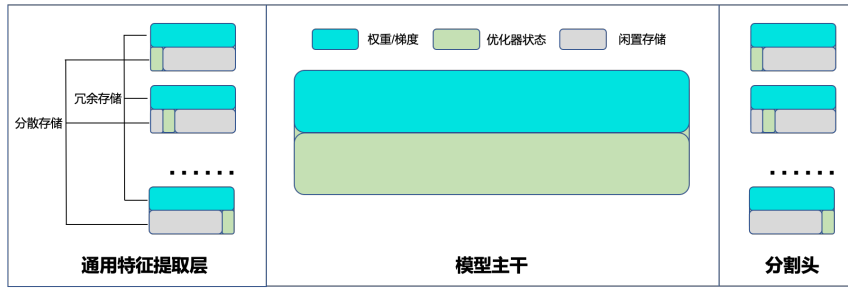


图 2.1 MedFed 联邦学习框架模型分割图

的模型。thapad 等人提出的 `splitfed`<sup>[59]</sup> 也采用了类似的架构，我们的联邦学习框架也借鉴了他们的设计，但是 `splitfed` 的局限性在于：1) `splitfed` 存在隐私泄露的问题，因为 `splitfed` 会在数据中心存储完整的标签，这在一定程度上会泄露患者的隐私；2) 该方法只是在单机上验证了可行性，在实际的生产环境中，该方法的性能是无法保证的。3) `splitfed` 完全没有考虑通信异构和计算异构的问题，这就意味着，`splitfed` 的训练效率总是受限于医院客户端中最慢的那一个，这在实际的生产环境中会导致训练效率的极大下降，对于宝贵的计算资源是一种极大的浪费。4) `splitfed` 虽然讲模型分割成了两个部分，但是对于客户端上存储的模型，仍然需要存储完整的优化器状态，正如前文所提到的，优化器状态往往是要比模型参数更大的，这进一步增加了客户端的存储压力。

接下来，我们将介绍我们是如何解决这些问题的。首先，我们将模型切割成了两个部分，但是这种切割方式并不是简单的像 `splitfed` 那样将模型的前几层存储在客户端，将模型的后几层存储在服务端。这种切割方式带来的坏处是，我们需要将每一个样本的标签仍然存储在服务端，这在一定程度上会泄露患者的隐私。如图 3.2 所示

### 2.2.3 模型参数更新策略

### 2.2.4 动态批次组建

为了提升模型的训练速度，和减少数据中心模型的闲置时间，我们引入了动态批次组建的方法。具体来说，我们引入两个参数，全局批次大小  $GBS$  和局部批次大小  $MBS$ 。不同的医院客户端的计算能力、与数据中心的通信带宽和延迟都是不同的。因此，不同的医院客户端在给数据中心发送激活值的时候，发送的量和

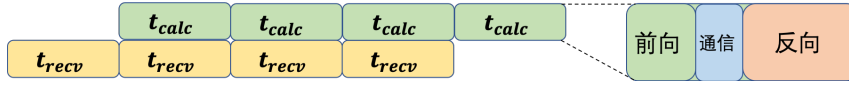


图 2.2 MedFed 联邦学习框架效率图

频率都是不同的。在数据中心收到  $MBS$  个激活值之后，我们就会将这  $MBS$  个激活值组成一个批次，然后进行一次前向和反向传播。这一步需要的时间与数据中心分配给这个任务的算力成正比，设这个时间为  $t_{calc}$ 。同时，我们设置数据中心收到  $MBS$  个激活值的时间为  $t_{recv}$ 。

我们进一步考虑了数据中心在处理这些批次时的效率问题。我们定义了一个效率比率  $\eta$ ，它是基于计算时间  $t_{calc}$  和接收时间  $t_{recv}$  的函数，用来衡量数据中心处理批次的效率。公式如下：

$$\eta = \frac{t_{calc}}{t_{recv}} \quad (2-1)$$

理想情况下，我们希望  $\eta$  接近于 1，这意味着数据中心的计算能力与接收激活值的速度保持平衡。这是因为，一个批次的处理时间由接收激活值的时间和前向反向传播中较长的一个决定。当我们通过调节数据中心的算力，使得  $\eta$  接近于 1 时，我们的训练过程就如图 3.2 所示，可以看到，数据中心的计算能力和接收激活值的速度是保持平衡的。

在这种情况下，我们可以计算的数据中心对于算力的浪费，因为数据中心的优化器更新不需要通信，因此在耗时上总是更短的，设数据中心的优化器更新时间为  $t_{opt\_center}$ ，医院客户端的优化器更新时间为  $t_{opt\_client}$ 。则数据中心对的有效率计算时间为：

$$t_{efficiency} = \frac{\frac{GBS}{MBS} * t_{calc} + t_{opt\_center}}{\frac{GBS}{MBS} * t_{calc} + 1 * t_{calc} + t_{opt\_client}} \quad (2-2)$$

### 2.2.5 数据异构解决方案

第二个问题是尽可能消除非独立同分布数据对模型训练的影响。数据的异构性一直是联邦学习中的一个重要问题<sup>[60]</sup>，因为不同的医院客户端的数据分布是不同的，如果不做任何处理，这就会导致模型的训练效果很差。首先我们确保来自不同的数据分布的数据在一次训练中是均匀分布的（需要引用）。。。。。

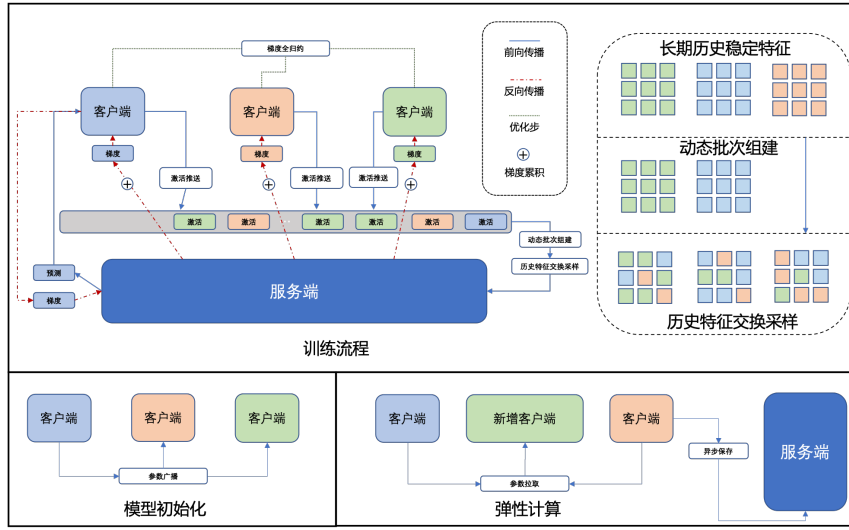


图 2.3 MedFed 联邦学习框架结构图

### 2.2.6 历史特征交换采样

其次根据我们所采用的实际的分割模型（存在一个教师模型，和一个学生模型，学生模型通过指数平均算法更新教师模型<sup>[61]</sup>）也就是教师模型是更新的相对缓慢的，因此对同一个样本而言，输入到相邻训练步数的教师模型中，得到的特征是相似的，因此我们可以定义一个历史窗口，这个历史窗口内的模型我们认为他们是有效的模型，但是距离当前时间点越远的模型，他的权重会越小，公式如下：

$$\omega_t = \frac{1}{1 + \alpha(t_{\text{cur}} - t)} \quad (2-3)$$

其中  $\omega_t$  表示模型的权重， $t_{\text{cur}}$  表示当前时间点， $t$  表示模型的时间点， $\alpha$  表示一个超参数，用来调整模型的权重。

## 2.3 实验设置



## 第3章 总结与展望

本文采用……。 (结论作为学位论文正文的最后部分单独排写，但不加章号。结论是对整个论文主要结果的总结。在结论中应明确指出本研究的创新点，对其应用前景和社会、经济价值等加以预测和评价，并指出今后进一步在本研究方向进行研究工作的展望与设想。结论部分的撰写应简明扼要，突出创新性。)





## 参考文献

- [1] Ludwig W, Schmidt S, Lauridsen E M, et al. X-ray computed tomography[J]. Nature Reviews Methods Primers, 2008. <https://www.nature.com/articles/nrmeth.2015.16>.
- [2] of Nuclear Medicine Technology J. Principles of ct and ct technology[J]. Journal of Nuclear Medicine Technology, 2023. <https://tech.snmjournals.org/content/early/2023/02/15/jnmt.122.263632>.
- [3] Dean J, Corrado G S, Monga R, et al. Large scale distributed deep networks[J]. 2012.
- [4] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Artificial Intelligence and Statistics. 2017: 1273-1282.
- [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] 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.
- [7] Mills J, Hu J, Min G. Communication-efficient federated learning for wireless edge intelligence in iot[J]. IEEE Internet of Things Journal, 2019, 7(7): 5986-5994.
- [8] Xu J, Du W, Jin Y, et al. Ternary compression for communication-efficient federated learning [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020.
- [9] Duan M, Liu D, Chen X, et al. Astraea: Self-balancing federated learning for improving classification accuracy of mobile deep learning applications[C]//2019 IEEE 37th International Conference on Computer Design (ICCD). IEEE, 2019: 246-254.
- [10] Lim W Y B, Luong N C, Hoang D T, et al. Federated learning in mobile edge networks: A comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2020, 22(3): 2031-2063.
- [11] Kairouz P, McMahan H B, Avenet B, et al. Advances and open problems in federated learning[J]. arXiv preprint arXiv:1912.04977, 2019.
- [12] Zhang C, Li S, Xia J, et al. Batchcrypt: Efficient homomorphic encryption for cross-silo federated learning[C]//2020 USENIX Annual Technical Conference (USENIX ATC 20). USENIX Association, 2020: 493-506. <https://www.usenix.org/conference/atc20/presentation/zhang-chengliang>.
- [13] Zhu H, Wang R, Jin Y, et al. Distributed additive encryption and quantization for privacy preserving federated deep learning[J]. arXiv preprint arXiv:2011.12623, 2020.
- [14] Zhao Y, Li M, Lai L, et al. Federated learning with non-iid data[J]. arXiv preprint arXiv:1806.00582, 2018.
- [15] Zhang C, Xie Y, Bai H, et al. A survey on federated learning[J]. Knowledge-Based Systems, 2021, 216: 106775. <https://www.sciencedirect.com/science/article/pii/S0950705121000381>.
- [16] Li T, Sahu A K, Talwalkar A, et al. Federated learning: Challenges, methods, and future directions[J]. IEEE Signal Processing Magazine, 2020, 37(3): 50-60.

- [17] Vepakomma P, Swedish T, Raskar R, et al. No peek: A survey of private distributed deep learning [J]. arXiv preprint arXiv:1812.03288, 2018.
- [18] Li Q, Wen Z, Wu Z, et al. A survey on federated learning systems: vision, hype and reality for data privacy and protection[J]. arXiv preprint arXiv:1907.09693, 2019.
- [19] Briggs C, Fan Z, Andras P. A review of privacy preserving federated learning for private iot analytics[J]. arXiv preprint arXiv:2004.11794, 2020.
- [20] Lyu L, Yu H, Yang Q. Threats to federated learning: A survey[J]. arXiv preprint arXiv:2003.02133, 2020.
- [21] Lan Q, Zhang Z, Du Y, et al. An introduction to communication efficient edge machine learning [J]. arXiv preprint arXiv:1912.01554, 2019.
- [22] Imteaj A, Thakker U, Wang S, et al. Federated learning for resource-constrained iot devices: Panoramas and state-of-the-art[J]. arXiv preprint arXiv:2002.10610, 2020.
- [23] Shi Y, Yang K, Jiang T, et al. Communication-efficient edge ai: Algorithms and systems[J]. IEEE Communications Surveys & Tutorials, 2020, 22(4): 2167-2191.
- [24] Hosseinalipour S, Brinton C G, Aggarwal V, et al. From federated learning to fog learning: Towards large-scale distributed machine learning in heterogeneous wireless networks[J]. arXiv preprint arXiv:2006.03594, 2020.
- [25] Xu J, Glicksberg B S, Su C, et al. Federated learning for healthcare informatics[J]. Journal of Healthcare Informatics Research, 2021, 5(1): 1-19.
- [26] Kulkarni V, Kulkarni M, Pant A. Survey of personalization techniques for federated learning [C]//2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4). IEEE, 2020: 794-797.
- [27] Li T, Sahu A K, Talwalkar A, et al. Federated learning: Challenges, methods, and future directions[J]. IEEE Signal Processing Magazine, 2020, 37(3): 50-60.
- [28] Stich S U. Local SGD converges fast and communicates little[J]. International Conference on Learning Representations (ICLR), 2019: arXiv:1805.09767. <https://arxiv.org/abs/1805.09767>.
- [29] Li X, Huang K, Yang W, et al. On the convergence of fedavg on non-iid data[C]//International Conference on Learning Representations. 2020. <https://openreview.net/forum?id=HJxNANvTDS>.
- [30] Li T, Sahu A K, Zaheer M, et al. Federated optimization in heterogeneous networks[C]//Proceedings of Machine Learning and Systems 2020. 2020: 429-450.
- [31] Hsu T H, Qi H, Brown M. Measuring the effects of non-identical data distribution for federated visual classification[J]. CoRR, 2019, abs/1909.06335. <http://arxiv.org/abs/1909.06335>.
- [32] Reddi S, Charles Z, Zaheer M, et al. Adaptive federated optimization[J]. arXiv preprint arXiv:2003.00295, 2020.
- [33] Karimireddy S P, Kale S, Mohri M, et al. SCAFFOLD: stochastic controlled averaging for on-device federated learning[J]. CoRR, 2019, abs/1910.06378. <http://arxiv.org/abs/1910.06378>.
- [34] Shamir O, Srebro N, Zhang T. Communication-efficient distributed optimization using an approximate newton-type method[C]//International conference on machine learning. 2014: 1000-

- 1008.
- [35] Konečný J, McMahan H B, Yu F X, et al. Federated learning: Strategies for improving communication efficiency[J]. arXiv preprint arXiv:1610.05492, 2016.
  - [36] Makhdoumi A, Ozdaglar A. Convergence rate of distributed admm over networks[J]. IEEE Transactions on Automatic Control, 2017, 62(10): 5082-5095.
  - [37] Yuan H, Ma T. Federated accelerated stochastic gradient descent[J]. arXiv preprint arXiv:2006.08950, 2020.
  - [38] Pathak R, Wainwright M J. Fedsplit: An algorithmic framework for fast federated optimization [J]. arXiv preprint arXiv:2005.05238, 2020.
  - [39] Liang X, Shen S, Liu J, et al. Variance reduced local sgd with lower communication complexity [J]. arXiv preprint arXiv:1912.12844, 2019.
  - [40] Li Z, Kovalev D, Qian X, et al. Acceleration for compressed gradient descent in distributed and federated optimization[J]. arXiv preprint arXiv:2002.11364, 2020.
  - [41] Condat L, Malinovsky G, Richtárik P. Distributed proximal splitting algorithms with rates and acceleration[J]. arXiv preprint arXiv:2010.00952, 2020.
  - [42] Gorbunov E, Hanzely F, Richtárik P. A unified theory of sgd: Variance reduction, sampling, quantization and coordinate descent[C]//International Conference on Artificial Intelligence and Statistics. PMLR, 2020: 680-690.
  - [43] Khaled A, Sebbouh O, Loizou N, et al. Unified analysis of stochastic gradient methods for composite convex and smooth optimization[J]. arXiv preprint arXiv:2006.11573, 2020.
  - [44] Li Z, Richtárik P. A unified analysis of stochastic gradient methods for nonconvex federated optimization[J]. arXiv preprint arXiv:2006.07013, 2020.
  - [45] Li T, Sahu A K, Zaheer M, et al. Feddane: A federated newton-type method[C]//2019 53rd Asilomar Conference on Signals, Systems, and Computers. IEEE, 2019: 1227-1231.
  - [46] Zhang X, Hong M, Dhople S, et al. Fedpd: A federated learning framework with optimal rates and adaptivity to non-iid data[J]. arXiv preprint arXiv:2005.11418, 2020.
  - [47] Dutta A, Bergou E H, Abdelmoniem A M, et al. On the discrepancy between the theoretical analysis and practical implementations of compressed communication for distributed deep learning [J]. arXiv preprint arXiv:1911.08250, 2019.
  - [48] Mishchenko K, Gorbunov E, Takáč M, et al. Distributed learning with compressed gradient differences[J]. arXiv preprint arXiv:1901.09269, 2019.
  - [49] Alistarh D, Grubic D, Li J, et al. Qsgd: Communication-efficient sgd via gradient quantization and encoding[C]//Advances in Neural Information Processing Systems. 2017: 1709-1720.
  - [50] 姜敏, 彭少贤, 酆华兴. 形状记忆聚合物研究现状与发展[J]. 现代塑料加工应用, 2005, 17 (2): 53-56.
  - [51] McMahan H B, Moore E, Ramage D, et al. Federated learning of deep networks using model averaging[J]. CoRR, 2016, abs/1602.05629. <http://arxiv.org/abs/1602.05629>.
  - [52] Sahu A K, Li T, Sanjabi M, et al. On the convergence of federated optimization in heterogeneous

- networks[J]. CoRR, 2018, abs/1812.06127. <http://arxiv.org/abs/1812.06127>.
- [53] Beaumont O, Eyraud-Dubois L, Herrmann J, et al. Optimal checkpointing for heterogeneous chains: how to train deep neural networks with limited memory[J]. CoRR, 2019, abs/1911.13214. <http://arxiv.org/abs/1911.13214>.
- [54] S.B. S, Garg A, Kulkarni P. Dynamic memory management for gpu-based training of deep neural networks[C]//2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). 2019: 200-209.
- [55] Ren J, Rajbhandari S, Aminabadi R Y, et al. Zero-offload: Democratizing billion-scale model training[J]. CoRR, 2021, abs/2101.06840. <https://arxiv.org/abs/2101.06840>.
- [56] Rajbhandari S, Ruwase O, Rasley J, et al. Zero-infinity: Breaking the GPU memory wall for extreme scale deep learning[J]. CoRR, 2021, abs/2104.07857. <https://arxiv.org/abs/2104.07857>.
- [57] Narayanan D, Shoeybi M, Casper J, et al. Efficient large-scale language model training on GPU clusters[J]. CoRR, 2021, abs/2104.04473. <https://arxiv.org/abs/2104.04473>.
- [58] Narayanan D, Phanishayee A, Shi K, et al. Memory-efficient pipeline-parallel DNN training[J]. CoRR, 2020, abs/2006.09503. <https://arxiv.org/abs/2006.09503>.
- [59] Thapa C, Chamikara M A P, Camtepe S, et al. Splitfed: When federated learning meets split learning[Z]. 2022.
- [60] Zhu H, Xu J, Liu S, et al. Federated learning on non-iid data: A survey[J]. CoRR, 2021, abs/2106.06843. <https://arxiv.org/abs/2106.06843>.
- [61] Tarvainen A, Valpola H. Weight-averaged consistency targets improve semi-supervised deep learning results[J]. CoRR, 2017, abs/1703.01780. <http://arxiv.org/abs/1703.01780>.

## 附录 A 关于 XXX 的证明

附录相关内容...



## 附录 B Maxwell Equations

因为在柱坐标系下， $\bar{\mu}$  是对角的，所以 Maxwell 方程组中电场  $\mathbf{E}$  的旋度所以  $\mathbf{H}$  的各个分量可以写为：

$$H_r = \frac{1}{\mathbf{i}\omega\mu_r} \frac{1}{r} \frac{\partial E_z}{\partial \theta} \quad (\text{B-1a})$$

$$H_\theta = -\frac{1}{\mathbf{i}\omega\mu_\theta} \frac{\partial E_z}{\partial r} \quad (\text{B-1b})$$

同样地，在柱坐标系下， $\bar{\epsilon}$  是对角的，所以 Maxwell 方程组中磁场  $\mathbf{H}$  的旋度

$$\nabla \times \mathbf{H} = -\mathbf{i}\omega\mathbf{D} \quad (\text{B-2a})$$

$$\left[ \frac{1}{r} \frac{\partial}{\partial r} (r H_\theta) - \frac{1}{r} \frac{\partial H_r}{\partial \theta} \right] \hat{\mathbf{z}} = -\mathbf{i}\omega\bar{\epsilon}\mathbf{E} = -\mathbf{i}\omega\epsilon_z E_z \hat{\mathbf{z}} \quad (\text{B-2b})$$

$$\frac{1}{r} \frac{\partial}{\partial r} (r H_\theta) - \frac{1}{r} \frac{\partial H_r}{\partial \theta} = -\mathbf{i}\omega\epsilon_z E_z \quad (\text{B-2c})$$

由此我们可以得到关于  $E_z$  的波函数方程：

$$\frac{1}{\mu_\theta\epsilon_z} \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial E_z}{\partial r} \right) + \frac{1}{\mu_r\epsilon_z} \frac{1}{r^2} \frac{\partial^2 E_z}{\partial \theta^2} + \omega^2 E_z = 0 \quad (\text{B-3})$$





## 致 谢

本论文的工作是在导师……。



## 在学期间完成的相关学术成果

### 学术论文：

- [1] 高凌. 交联型与线形水性聚氨酯的形状记忆性能比较 [J]. 化工进展, 2006, 532 — 535. (核心期刊)
- [2] 杨轶, 张宁欣, 任天令, 等. 硅基铁电微声学器件中薄膜残余应力的研究 [J]. 中国机械工程, 2005, 16(14):1289-1291.

### 专利：

- [3] 任天令, 杨轶, 朱一平, 等. 硅基铁电微声学传感器畴极化区域控制和电极连接的方法: 中国, CN1602118A[P]. 2005-03-30.
- [4] Ren T L, Yang Y, Zhu Y P, et al. Piezoelectric micro acoustic sensor based on ferroelectric materials: USA, No.11/215, 102[P]. (美国发明专利申请号.)

