



# 南京大学

## 研究生毕业论文 (申请硕士学位)

论文题目 南京大学毕业论文 L<sup>A</sup>T<sub>E</sub>X 模板

作者姓名 陈晨曦

学科、专业名称 计算机技术

研究方向 联邦学习

指导教师 叶保留教授

2022 年 2 月 15 日

学 号：MF1933004

论文答辩日期：xxxx 年 xx 月 xx 日

指 导 教 师： (签字)

# **L<sup>A</sup>T<sub>E</sub>X NJU thesis template**

by

**Chenxi Chen**

Supervised by

**Professor Baoliu Ye**

A dissertation submitted to  
the graduate school of Nanjing University  
in partial fulfilment of the requirements for the degree of  
MASTER  
in  
Computer Science and Technology



Department of Computer Science and Technology  
Nanjing University

February 15, 2022



# 南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目：\_\_\_\_\_ 标题第一行  
\_\_\_\_\_ 标题第二行用于长标题换行  
计算机技术 专业 2019 级硕士生姓名： 陈晨曦  
指导教师（姓名、职称）： \_\_\_\_\_ 叶保留教授

## 摘 要

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

关键词： 关键词 1    关键词 2

## 南京大学研究生毕业论文英文摘要首页用纸

THESIS: englishabstractitlea  
nenglishabstractitleb  
SPECIALIZATION: Computer Science and Technology  
POSTGRADUATE: Chenxi Chen  
MENTOR: Professor Baoliu Ye

### **Abstract**

As a new machine learning paradigm, federated learning uses distributed edge devices to train the model without exposing privacy data, providing thorough protection for user privacy. However, federated learning faces severe communication challenges due to frequent and massive model parameter updates. Moreover, its training delay suffers from the heterogeneity of edge devices. Federated learning places the majority of the deep learning (DL) training tasks on the edge devices with heterogeneous computation and communication abilities, leading to tremendous training delay. And the intense exchanges of heavy model parameters severely exacerbate this problem. DL models' layered structure allows a device to partition and offload part of the model training to the server, which reduces the size of transmitted data and computation amount of devices. This paper proposes a method based on the model partition, which speeds up training by reducing the high communication overhead and mitigating the heterogeneous problem of federated learning in the edge environment. We model the training process of a partitioned model in the wireless network, construct an optimization problem that minimizes the training through adjustment of partition and resource allocation strategies. And we design an polynomial algorithm that finds a sub-optimal to the problem by dividing it into two sub-problems and solving them using polynomial algorithms. Extensive simulation results verify that our algorithm can effectively reduce the training delay compared to existing solutions.

**keywords:** federated learning    model partition

# 前言

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

作者

20xx 年夏于南京大学





# 目 录

中文摘要 .....	i
英文摘要 .....	ii
前 言 .....	iii
目 录 .....	v
插图清单 .....	vii
附表清单 .....	ix
第一章 绪论 .....	1
1.1 研究背景 .....	1
1.2 章节 .....	2
1.3 章节 .....	3
1.4 论文结构 .....	3
第二章 算法 .....	5
第三章 实验验证 .....	7
第四章 总结与展望 .....	9
致 谢 .....	11
参考文献 .....	13
A 附录代码 .....	15
A.1 main 函数 .....	15
简历与科研成果 .....	17
《学位论文出版授权书》 .....	19



# 插图清单

3-1	单图示例 .....	7
3-2	实验硬件设备总览 .....	7
3-3	实验测量示意图 .....	7
3-4	子图样例 .....	8



## 附表清单



# 第一章 绪论

## 1.1 研究背景

基于神经网络的深度学习在人工智能的各种领域，如图像识别、语音识别和自然语言处理，中取得了前所未有的精度。深度学习技术也因此被广泛应用于现实生活场景，从根本上改变了人们的生活方式。主流的深度学习模型训练框架都是基于云的，这类框架需要首先将用户产生的海量原始数据收集到云中心，随后集群使用收集来的数据训练深度学习模型。然而，这样的训练方式暴露了一些缺点。一是隐私问题。基于云的训练矿建允许服务器直接访问用户的数据，这威胁到用户的隐私安全。随着网络的普及，越来越多的人意识到保护网络隐私的重要性，关于网络隐私保护的立法也在不断完善，例如：欧盟通用数据保护条例 (GDPR)[?]、《中华人民共和国网络安全法》[?]、《中华人民共和国民法通则》`citegeneralprinciple` 限制了对用户敏感数据的使用。

其次，随着物联网的发展，越来越多的联网设备——智能手机、摄像机和智能传感器等被生产和使用，人们平均每天产生的数据也在急剧增长。根据思科的 2020 年度互联网报告 [?]，网络设备数量将从 2018 年的 184 亿台增加到 2023 年的 293 亿台。存储如此规模的数据势必会给云中心带来巨大的存储压力。

为了解决上述问题，提出了一种称为联邦学习 [?] 的新 DL 训练框架。在联邦学习中，参与者（参与训练的最终用户）使用他们的数据合作训练联邦学习服务器所需的深度学习模型，然后将他们的模型更新而不是原始数据上传到服务器进行聚合。这些步骤重复多次，直到达到所需的精度。这种训练方式可以防止服务器直接访问用户的敏感数据；因此，用户隐私受到保护。并且服务器不需要存储所有的训练数据，可以节省相当多的存储空间。此外，由于计算能力以惊人的速度增长的无数终端设备的杠杆作用，联邦学习在大规模并行计算中具有巨大的潜力。

尽管联邦学习相对于基于云的训练方法具有优势，但在实践中的部署方面面临着一些挑战。主要挑战之一是通信开销。在联邦学习中训练 DL 模型涉及

数百甚至数千轮模型交换，而如今的 DL 模型往往具有巨大的参数大小——数百兆字节或更多。因此，使用联邦学习来训练 DL 模型需要大量的数据传输。计算设备的异构性带来了另一个挑战。联邦学习采用批量同步并行（BSP）聚合，即在每一轮训练中，服务器在所有设备完成计算和上传结果后聚合更新。因为联邦学习中设备的计算和通信能力有很大差异，但是这些设备的计算量和通信量是相同的。所以在每一轮中，完成任务的最快设备必须等待最慢的设备。上述两个挑战都严重减缓了联邦学习的训练。

这些年来已经提出了应对这些挑战的研究。处理高通信开销的主流方法是压缩 [1] [2] [3] [4] [5]，它使用量化、稀疏化或修剪来减小模型或更新的大小。然而，由于压缩过程中的信息丢失，这种压缩方法会降低模型的精度。通常，模型在压缩后需要进行更多轮训练才能达到与压缩前模型相同的精度，这最终会增加整体训练时间。异步并行 (ASP) 聚合 [6]，其中服务器会在每次更新时立即聚合，以及过时的同步并行 (SSP) [7]，BSP 和 ASP 的折衷方案，都可以用来解决异构问题。然而，使用 ASP 或 SSP 模型收敛会恶化，这会导致收敛速度慢和模型精度低。此外，压缩和 ASP 聚合都无法应对上述挑战。

## 1.2 章节

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.



## 1.3 章节

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

## 1.4 论文结构

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

- 一级 item
  - 二级 item
    - ◆ 三级 item
- 一级 item



## 第二章 算法

---

算法 2.1 算法名字

---

输入: 这是输入

输出: 这是输出

1: **while** flag **do**

2:   这是语句

3: **end while**

---



## 第三章 实验验证



图 3-1: 单图示例

实验硬件设备如图 3-2 所示。

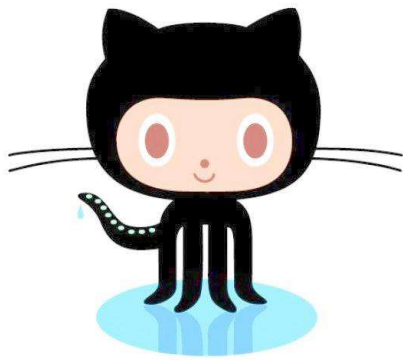


图 3-2: 实验硬件设备总览

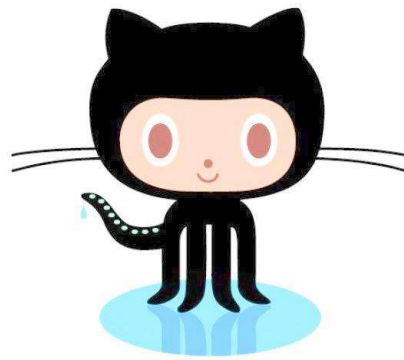


图 3-3: 实验测量示意图

图 3-4 所示子图 3-4(a) 和子图 3-4(b)。



(a) 子图



(b) 子图

图 3-4: 子图样例

## 第四章 总结与展望

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.





# 致 谢

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.



## 参考文献



# 附录 A 附录代码

## A.1 main 函数

```
int main()  
{  
    return 0;  
}
```



# 简历与科研成果

## 基本信息

韦小宝，男，汉族，1985 年 11 月出生，江苏省扬州人。

## 教育背景

2007 年 9 月 — 2010 年 6 月	南京大学计算机科学与技术系	硕士
2003 年 9 月 — 2007 年 6 月	南京大学计算机科学与技术系	本科

## 攻读硕士学位期间完成的学术成果

1. Xiaobao Wei, Jinnan Chen, “Voting-on-Grid Clustering for Secure Localization in Wireless Sensor Networks,” in Proc. IEEE International Conference on Communications (ICC) 2010, May. 2010.
2. Xiaobao Wei, Shiba Mao, Jinnan Chen, “Protecting Source Location Privacy in Wireless Sensor Networks with Data Aggregation,” in Proc. 6th International Conference on Ubiquitous Intelligence and Computing (UIC) 2009, Oct. 2009.

## 攻读硕士学位期间参与的科研课题

1. 国家自然科学基金面上项目“问题研究”（课题年限 2010 年 1 月 — 2012 年 12 月），负责相关问题的研究。





# 《学位论文出版授权书》

本人完全同意《中国优秀博硕士学位论文全文数据库出版章程》（以下简称“章程”），愿意将本人的学位论文提交“中国学术期刊（光盘版）电子杂志社”在《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》中全文发表。《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》可以以电子、网络及其他数字媒体形式公开出版，并同意编入《中国知识资源总库》，在《中国博硕士学位论文评价数据库》中使用和在互联网上传播，同意按“章程”规定享受相关权益。

作者签名：\_\_\_\_\_

\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

论文题名	南京大学毕业论文 L <sup>A</sup> T <sub>E</sub> X 模板				
研究生学号	MF1933004	所在院系	计算机科学与技术系	学位年度	2022
论文级别	<div><input type="checkbox"/> 硕士<div><input type="checkbox"/> 硕士专业学位</div></div> <div><input type="checkbox"/> 博士<div><input type="checkbox"/> 博士专业学位</div></div> <div>(请在方框内画勾)</div>				
作者 Email	sample@smail.nju.edu.cn				
导师姓名	叶保留教授				

论文涉密情况：

☐ 不保密

☐ 保密，保密期(\_\_\_\_\_年\_\_\_\_月\_\_\_\_日至\_\_\_\_\_年\_\_\_\_月\_\_\_\_日)

注：请将该授权书填写后装订在学位论文最后一页（南大封面）。

