**涉密论文** □ **公开论文** □



**本科生毕业论文（设计）**

**题目 基于联邦学习的光刻热区检测**

**姓名与学号**  王唯一 3160102407

**指导教师**  卓成

**年级与专业**  2016级信息工程

**所在学院**  信息与电子工程学院

**提交日期**  2020年7月8日

**浙江大学本科生毕业论文（设计）承诺书**

1. 本人郑重地承诺所呈交的毕业论文（设计），是在指导教师的指导下严格按照学校和学院有关规定完成的。

2. 本人在毕业论文（设计）中除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 **浙江大学** 或其他教育机构的学位或证书而使用过的材料。

3. 与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

4. 本人承诺在毕业论文（设计）工作过程中没有伪造数据等行为。

5. 若在本毕业论文（设计）中有侵犯任何方面知识产权的行为，由本人承担相应的法律责任。

6. 本人完全了解**浙江大学**有权保留并向有关部门或机构送交本论文（设计）的复印件和磁盘，允许本论文（设计）被查阅和借阅。本人授权**浙江大学**可以将本论文（设计）的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编本论文（设计）。

作者签名： 导师签名：

签字日期： 年 月 日 签字日期： 年 月 日

**致谢**

浙大四年，白驹过隙，在这里我曾彷徨，曾挣扎，也曾坚定，曾解脱。在浙大，我遇见的每一个人，经历的每一件事都对我的人生带来或大或小的涟漪。我心中充满着对许许多多的人的感激。

首先感谢我的毕业设计导师卓成老师，他为我指引了毕业设计的方向，为我提供了毕业设计所需的实验环境和实验平台，经常与我就课题的内容进行讨论，提出了许多新的改进与思路，毕业设计能够有最终的效果，与他对我的指导息息相关。

感谢香港中文大学的Bei Yu老师和陈然同学，他们为我提供了工业界的数据集，让我能长期使用他们的实验平台，使得我的毕业设计能够顺利进行下去。

感谢徐金明老师在百忙之中抽出时间与我讨论，在分布式优化方面给予许多我指点。

感谢高迪师姐，她提供了许多有价值的资料，多次抽出时间与我进行讨论，对我的困惑作出解答。

感谢我的父母，是他们将我养育成人，在我本科求学和做毕业设计时，一直支持我，给予我鼓励和肯定。

感谢我身边所有的同学和好友，感谢他们本科四年的陪伴，感谢他们为我带来的欢乐。

寥寥数语难以言尽我的心情，希望未来还能和陪伴我的人携手并进。

**摘要**

机器学习已经广泛地应用于光刻热区检测领域，并在最近几年涌现出了许多不俗的成果。但是仍有一个矛盾点一直没能解决：芯片设计公司和研究机构出于保护商业利益的目的，往往不愿意共享数据，研究人员能获得的数据体量极其有限。这导致过往的许多光刻热区检测模型或多或少存在过拟合，普适性不足等问题。联邦学习的问世给这个问题的解决提供了思路，本文提出了一种全新的系统框架，将本地光刻热区检测算法与联邦学习FederatedAveraging算法结合起来，既保护了各个芯片设计单位的数据隐私，又使得光刻热区检测模型鲁棒且具有普适性。在客户端，本文采用已经十分成熟的BL算法，以保证本地模型在本地数据集上表现最佳；在服务器端，本文采用FederatedAveraging算法，对客户上传的模型进行加权平均。而后本文以ADMM算法优化了整个系统并在系统中添加了差分隐私以进一步保护客户的数据隐私。实验结果表明，结合BL算法与FederatedAveraging算法得到的光刻热区检测模型可以与以往高性能模型相媲美，验证了用联邦学习来解决光刻热区检测中的痛点是可行的。但是以ADMM算法优化系统并没有带来性能上的提升，这需要更深层次的探讨。

**关键词：**光刻热区检测；联邦学习；隐私保护；BL算法；ADMM算法

**Abstract**

Machine learning has been widely used in the field of lithography hotspot detection, and many advanced results have emerged in recent years.But there is still a contradiction that has not been resolved: chip design companies and research institutions are often reluctant to share data for the purpose of protecting commercial interests, so the amount of data available to researchers is extremely limited. The creation of federated learning provides an idea for solving this problem. This thesis proposes a brand new system framework that combines the local lithography hotspot detection algorithm with the federated learning FederatedAveraging algorithm to protect the data privacy of each chip design company, it also makes the lithography hotspot detection model robust and scalable. On the client side, this thesis uses the advanced BL algorithm to ensure that the local model performs best on the local data set; on the server side, this thesis uses the FederatedAveraging algorithm to weight the average of the model uploaded by the client. Then this thesis optimizes the entire system with ADMM algorithm and adds differential privacy to the system to further protect the customer's data privacy. The experimental results show that the lithography hotspot detection model combined with the BL algorithm and FederatedAveraging algorithm can be comparable to the previous high-performance models, which verifies that it is feasible to solve the contradiction in lithography hotspot detection using federated learning. However, optimizing the system with the ADMM algorithm has not brought performance improvements, which requires a further discussion.

**Keywords:** Hotspot Detection; Federated Learning; Privacy Detection; BL Algorithm; ADMM

**目录**

**第一部分 毕业论文（设计）**

[1 绪论 1](#_Toc43291709)

[1.1 背景介绍 1](#_Toc43291710)

[1.1.1 光刻热区检测 1](#_Toc43291711)

[1.1.2 联邦学习 1](#_Toc43291712)

[1.2 国内外研究现状 2](#_Toc43291713)

[1.2.1 光刻热区检测研究现状 2](#_Toc43291714)

[1.2.1 联邦学习与隐私保护 3](#_Toc43291715)

[1.3 主要研究内容 4](#_Toc43291716)

[2 客户端光刻热区检测基本算法 6](#_Toc43291717)

[2.1 特征张量提取 6](#_Toc43291718)

[2.2 客户端神经网络 7](#_Toc43291719)

[2.2.1 神经网络基本架构 7](#_Toc43291720)

[2.2.2 参数更新策略 8](#_Toc43291721)

[2.2.3 Biased Learning 9](#_Toc43291722)

[3 服务器端模型聚合算法及系统框架 13](#_Toc43291723)

[3.1 FederatedAveraging算法 13](#_Toc43291724)

[3.2 系统框架 15](#_Toc43291725)

[3.3 系统优化 16](#_Toc43291726)

[3.3.1 ADMM优化系统 16](#_Toc43291727)

[3.3.2 差分隐私的应用 17](#_Toc43291728)

[4 实验结果 20](#_Toc43291729)

[4.1 实验环境与基本实验参数 20](#_Toc43291730)

[4.2 模型训练 20](#_Toc43291731)

[4.3 模型测试 20](#_Toc43291732)

[5 结论及未来展望 26](#_Toc43291733)

[参考文献 27](#_Toc43291734)

[附录 29](#_Toc43291735)

《浙江大学本科生毕业论文（设计）任务书》

《浙江大学本科生毕业论文（设计）考核表》

**第二部分 文献综述和开题报告**

指导教师对文献综述和开题报告具体内容要求

一、文献综述 1

二、开题报告 7

三、外文翻译 13

四、外文原文 29

《浙江大学本科生文献综述和开题报告考核表》

第一部分

**毕业论文（设计）**

# 绪论

## 背景介绍

### 1.1.1 光刻热区检测

随着半导体工艺节点的不断推进，晶体管特征尺寸进入了纳米领域。传统光刻系统的光源波长局限着光刻工艺，这已经成为集成电路产量、良品率的瓶颈。即使采用各种先进的工艺诸如design rule check(DRC), optical proximity correction(OPC)等，某些特殊的版图区域仍会导致制造上的缺陷，这些版图区域被称为光刻热区。而光刻热区检测即为在一定时间范围内，准确地定位出版图中存在的光刻热区。因此，光刻热区检测是集成电路版图设计物理验证阶段的一个重要环节。

传统的光刻热区检测很大程度上依赖于光刻仿真，该方法能以极高的准确率检测出版图中的热区，但同时计算复杂度高，耗费时间长，一般用于制造最后的验证阶段。

一般用三个指标来评价一个光刻热区检测方法：

1. Accuracy：被正确检测的热区片段与所有热区片段的比值。
2. False Alarm：实际为非热区却被检测为热区的片段总数。
3. ODST(Overall Detection and Simulation Time)：模型测试的时间与被检测为光刻热区片段的光刻仿真时间的总和.

为了在测试阶段快速而准确地检测出光刻热区，现阶段有两种主流的方法，分别位模式匹配和机器学习，两种方法从不同的思路尝试去解决光刻热区检测中的问题，各有优劣。

### 1.1.2 联邦学习

在信息爆炸与机器学习火热的时代，模型的训练者们总是期望着能获得更多的数据，使训练模型更具有普适性、鲁棒性，而现在诸多的移动终端给开发者们提供了这种可能。例如，开发者借助手机的录音功能，获取海量用户的语音，训练语音识别模型，从而提供位用户提供更好的交互体验。然而对用户而言，这许多的信息、数据具有极强的隐私性，用户并不希望厂商、开发者们获得与个人隐私相关的数据。数据的隐私性与开发者对数据的期望之间的矛盾越来越成为社会、学界关注的焦点。联邦学习正是在这种背景下应运而生。

联邦学习这一概念最早由McMahan H B等人提出[1]，简而言之即为将训练数据分布在各个移动终端，各自进行训练，将所得模型上传至中心服务器进行聚合，而后分享给各个终端使用。联邦学习使得用户不需将数据上传至中心服务器便能享受到经过充分训练的模型带来的良好用户体验。

在一个基本的联邦学习框架中，有一个中心服务器(server)和若干个客户(client)。每个客户存储着自己的不共享的本地数据，各自用本地数据训练模型，将模型上传至中心服务器。中心服务器负责组织各客户的本地训练，聚合所得模型，将模型分享回客户，如此为一轮。客户再在聚合后的模型上进行训练，周而复始，直至最终模型收敛。

联邦学习有几点特性是其他学习框架所不具备的：

1. 隐私性。不仅仅是数据来源的隐私性，本地训练的模型也是隐私的，对于服务器而言，所得模型的来源也是隐私的。
2. 各客户的数据集可能是非独立同分布，也就意味着某一个客户的数据并不一定具有整体的特性。
3. 通信成本需要被考虑在整体性能之内。一个联邦学习框架中，会出现某个或某几个客户可能会中途退出训练，或者各客户与服务器之间的通信成本不同等等情况。

如何开发一套快速计算的本地算法、如何在服务器端对模型进行聚合、如何减少通信成本都是未来联邦学习研究与发展的方向。

## 国内外研究现状

### 1.2.1 光刻热区检测研究现状

现阶段光刻热区检测主要有三种方法：光刻仿真、模式匹配、机器学习。如前文所说，光刻仿真虽然准确率极高，但是计算复杂度高，耗时长，不便于在测试阶段快速准确地检测光刻热区。因此，模式匹配和机器学习是当前的主流手段。

在采用模式匹配方法的诸多光刻热区检测框架中，Wen W Y等人[2]将基于密度的版图编码和主成分分析法、改良的曼哈顿距离法相结合；在[3]中，Yu Y T等人用一些重要的设计规则来表征光刻热区的拓扑特征；在[4]中，提出了一种改进的切线空间距离度量来进行热点模式分析和分类。尽管模式匹配能准确快速地检测出光刻热区，但是对于一些未知的光刻热区模式，其准确率并不尽人意。另一方面，机器学习往往能提取到版图的深度特征，能显著提升光刻热区检测的准确率。Drmanac等人[5]提出了一种基于直方图的布局表示和一种无监督支持向量机(SVM)模型来预测光刻过程的变化；Ding等人[6]采用了分层人工神经网络和SVM来降低false alarm；Zhang H等人[7]提出了一种将在线学习与信息论相结合的方法来进行热区检测，他们的方法在ICCAD 2012 Contest数据集上的表现优于许多前人的工作；Yang H Y等人[8]在卷积神经网络的基础上提出了BL(Biased Learning)算法及其改良版BBL(Batch Biased learning)算法；Chen R等人[9]提出了一种能直接在一个大的版图区域中检测多个光刻热区的框架，效果显著。

自从机器学习成为学界的热点，越来越多的关注被投放到用机器学习的方法来进行光刻热区检测。但是无论是模式匹配还是说机器学习，它们都没有解决一个问题：False Alarm过高的问题，也即有很多非热区被检测为了热区。加之各芯片设计厂商之间数据不互通，研究人员很难获得大量数据进行模型训练，这很容易导致模型的过拟合问题，如何获得各厂商的芯片设计版图而不侵犯他们的商业权益也是一个亟待解决的问题。

### 联邦学习与隐私保护

McDonald等人[21]和Povey等人[22]研究了通过迭代平均局部训练模型进行分布式训练的感知器和语音识别DNNs；Zhang等人[23]研究了一种“软”平均的异步方法；Neverova等[24]还讨论了在设备上保存敏感用户数据的优点；Shokri和Shmatikov的工作[25]在某种程度上和本文是相关的：他们专注于训练深度网络，强调隐私的重要性，通过在每一轮通信中只共享参数的子集来解决通信成本问题这些工作并没有考虑不平衡和非独立同分布的数据集，而这对于联邦学习而言是非常重要的。

联邦学习可分为三类：横向联邦学习（Horizontal Federated Learning），纵向联邦学习（Vertical Federated Learning），联邦迁移学习（Federated Transfer Learning）[11]。Yang Q等人在[11]中阐述了这三种联邦学习框架。

与那些把数据存储在中心服务器的训练框架不同，联邦学习把数据存储在客户本地。即使把数据集匿名化，一般的训练框架仍然会使数据的来源——客户处于隐私泄露的危险中，因为各客户通过中心服务器互联在一起。对比而言，联邦学习所需的数据通信是完全匿名的，因为模型的更新本身并不包含客户的个体信息，而模型更新的来源对中心服务器来说也不重要，因此，联邦学习完全尊重用户数据隐私。

现有许多隐私保护的框架和技术可以用于联邦学习。同态加密技术由Rivest R L等人最早提出[14]，同态加密除了能实现基本的加密操作之外，还能实现密文间的多种计算功能，即先计算后解密可等价于先解密后计算。安全多方计算（Secure Muti-party Computation）由Yao A C等人在1982年正式提出[12]。安全多方计算容许多个数据拥有者在互不信任的情况下进行协同计算，输出计算结果，并保证任何一方均无法获取除应得的计算结果之外的其他任何信息。差分隐私是另一项隐私保护的技术。差分隐私（Differential Privacy）是[密码学](https://baike.baidu.com/item/%E5%AF%86%E7%A0%81%E5%AD%A6)中的一种手段，旨在提供一种当从[统计数据库](https://baike.baidu.com/item/%E7%BB%9F%E8%AE%A1%E6%95%B0%E6%8D%AE%E5%BA%93)[查询](https://baike.baidu.com/item/%E6%9F%A5%E8%AF%A2)时，最大化数据查询的准确性，同时最大限度减少识别其[记录](https://baike.baidu.com/item/%E8%AE%B0%E5%BD%95)的机会[13]。差分隐私已经有许多知名的应用，例如苹果公司已在其IOS操作系统中使用差分隐私，改进其虚拟助手和建议技术[10]。

## 主要研究内容

如前文所说，出于芯片设计厂商对商业权益的保护而导致的样本稀缺，很容易使光刻热区检测模型出现过拟合的现象，因此，本文将通过联邦学习的方法来尝试解决这一问题。

本文的主要贡献如下：

1. 采取一种先进的BL(Biased Learning)及其改良版BBL(Batch Biased Learning)算法[8]作为客户端光刻热区检测基本算法，该算法使光刻热区检测在各客户端上达到性能最佳。
2. 将BL算法与FederatedAveraging算法[1]结合，建立基于联邦学习的光刻热区检测基本框架。
3. 使用ADMM算法优化模型聚合，减少因联邦学习而带来的性能损失。
4. 在系统中加入差分隐私，最大限度保护客户端数据隐私。

下文将有以下几章内容。

第二章介绍在客户端运行的光刻热区检测基本算法。客户端在本地对芯片版图提取特征张量，以此作为神经网络输入。在神经网络中，采取MGD方法作为参数更新策略，并对non-hotspot样本真实值作出一定偏置，这一算法被称为BL算法。该算法使得本地模型在本地数据集上的表现达到最佳。

第三章阐述在服务器端运行的模型聚合算法及ADMM算法[16]、差分隐私优化系统整体，阐述整个系统框架。基本系统框架在客户端采用第二章所述的BL算法，在服务器端采用FederatedAveraging算法对上传的模型进行聚合。ADMM算法适用于凸优化分布式问题，在大数据的分析学习中往往有较好的表现。在理想情况下，在本系统中客户是完全匿名的，但是其模型还有可能受到攻击，在系统中加入差分隐私能更好地保护模型不受外来攻击。

第四章罗列了实验结果。结果表明采用联邦学习的思想来进行光刻热区检测是可行的，能够解决当前光刻热区检测中存在的痛点。ADMM优化系统并没有可见的性能提升，差分隐私保护模型，同时牺牲了一定的系统吸能。第五章对理论和实验作出总结并探讨未来方向。

# 客户端光刻热区检测基本算法

Yang H Y等人[8]针对光刻热区检测提出特征张量提取算法，以此算法提取版图频域特征，后续以BL(Biased Learning)算法做分类，这种方法在ICCAD 2012 Contest数据集上达到了很高的准确率，并且False Alarm也很低。本文把该方法作为客户端运行的光刻热区检测算法，本章将对此作出具体介绍。

## 特征张量提取

本文所描述的这个框架中，第一步是要找到一个合适的版图特征表示。值得注意的是，版图中的热区与光的衍射相关，因此，一块版图是否包含光刻热区不仅仅取决于这块版图本身，还会受到周边版图的影响。鉴于此，在提取特征时，还要注意保留版图的空间相关性。

本文采取了一种基于离散余弦变换(DCT)的特征张量提取方法，该方法提供原始版图片段的低维表示同时保留了版图片段的空间相关性。在提取特征张量后，每幅版图图像I被转换成一幅具有以下性质的超图像（通道数量自定义的图像）F：1）每个通道的大小远小于I；2）从F中可以近似地恢复出I。

特征张量的提取包括以下几个步骤。

步骤1*:* 把每个版图片段分为个子区域。

步骤2*:* 把每个版图片段的每个子区域用如下离散余弦变换转换到频域：

 (1)

式中为子区域的大小，和为版图及其频域表示的各自坐标。显而易见的，频谱中的低频分量包含了绝大部分的信息与能量。

步骤3*:* 把按照锯齿形遍历展平成如下一维向量，其中大的索引意味着更高频的频域系数：

 (2)

步骤4*:* 选取的前个元素

 (3)

然后把所有的结合起来并保留他们的空间相关性。最终，特征张量如下所示：

， (4)

回溯上述步骤，即可从特征张量复原出原始的版图片段。

DCT的性质保证了高频系数趋近于0，较大的响应只出现在索引较小处，即低频区域。因此即使有大量的元素被丢弃了，信息量却没有很大的损失。

该特征张量在应用于神经网络时还具有以下优点：1）与卷积神经网络中的数据包传输高度兼容；2）与使用原始版图图像作为输入相比，由于神经网络的规模随着输入规模的减小而减小，前向传播时间显著缩短，节省了训练和测试时间。

## 客户端神经网络

### 2.2.1 神经网络基本架构

卷积神经网络(CNN)在图像分类领域有卓越的表现，BL算法采用卷积神经网络作为客户端神经网络的基本模型。卷积神经网络由几层执行特征提取的卷积层和几层生成样本分类概率的全连接层构成。

在本文中，卷积神经网络由两个卷积部分和两层全连接层构成，每个卷积部分包括三层：1）一个卷积层；2）一个ReLU层；3）一个最大池化层(max-pooling layer)。在每一次的卷积过程中，一系列的卷积核在特征张量F上执行如下的卷积操作：

 (5)

式中,。本文所有卷积核的大小都为3×3，两个卷积部分的输出通道数分别为16和32。ReLU是在卷积层之后，对每个元素进行操作的一种激活函数，其表达式如下所示，它保证了整个神经网络是非线性且稀疏的。

 (6)

最大池化层对上一层的输出作2×2的降采样，并作为当前卷积部分的输出层。卷积部分之后是两个全连接层，其输出节点分别为250和2。在训练时，以50%的概率对第一个全连接层执行dropout，以减轻过拟合第二个全连接层是整个神经网络的输出层，两个节点输出输入样本为光刻热区和非热区的预测概率。网络模型配置具体见表1。

**表1 神经网络模型配置**

|  |  |  |  |
| --- | --- | --- | --- |
| 层 | 卷积核大小 | 步长 | 输出 |
| conv1-1  conv1-2  maxpooling1 | 3  3  2 | 1  1  2 | 12×12×16  12×12×16  6×6×16 |
| Conv2-1  Conv2-2  Maxpooling2 | 3  3  2 | 1  1  2 | 6×6×32  6×6×32  3×3×32 |
| fc1  fc2 | -  - | -  - | 250  2 |

### 2.2.2 参数更新策略

确定梯度的计算方式和制定参数更新策略是训练神经网络之前非常关键的两项环节。后向传播(Back-propagation)[15]在计算大型神经网络的梯度中已经被验证是有效和快速的，且已经被广泛应用。每一个训练样本F都有对应的梯度集，集合里的每一个元素都是某一层的梯度矩阵，*v*是总层数。神经网络的所有参数都用来更新。

随机梯度下降(SGD)参数更新策略已经被无数次地证明能够有效地训练大数据集。但是当数据集的规模达到百万这一数量级时，SGD就很难有效地利用计算计算资源，在训练中需要很长的时间才能遍历到每一个训练样本。可以退一步采用一种小批量梯度下降(Mini-Batch Gradient Descent，MGD)的方法，该方法随机选取一组样本进行梯度下降。

对于大型非线性神经网络而言，后向传播和MGD并没有一个严格的收敛标准。通常会把25%的训练集拿出来作为验证集，验证集从不参与参数更新，而是每过几轮用它来测试模型。当测试结果没有太大变化或者开始逐渐变差的时候，就可以认为模型已经收敛了。

算法1中详细说明了MGD的细节，其中W为神经元权重，为学习率，为学习率衰减系数，*k*为学习率衰减步数，和分别为光刻热区和非热区的真实向量值。指标*j*对迭代进行计数(第4行)，在每一个迭代周期随机选择*m*个训练样本(第5行)。训练样本的梯度由后向传播计算得出(第6-8行)。然后神经元权重通过减去所选样本的平均梯度来进行更新(第14行)。当*j*是*k*的整数倍时，学习率衰减为，此后每*k*个迭代周期学习率都照此衰减(第11-13行)。在MGD算法的最后，返回一个在验证集上满足要求的模型(第15行)。

|  |
| --- |
| **算法1** MGD |
| 1: **function** MGD(**W**, , , *k*, , ) |
| 2: 初始化参数 *j* 0, **W** > 0 |
| 3: **while not** stop condition **do** |
| 4: *j j* + 1; |
| 5: 选取*m*个样本; |
| 6: **for** *i* 1, 2, …, *m* **do** |
| 7: backprop(); |
| 8: **end for** |
| 9: 计算梯度 |
| 10: 更新参数 **W** **W** |
| 11: **if** *j* mod *k* = 0 **then** |
| 12: , *j* 0; |
| 13: **end if** |
| 14: **end while** |
| 15: **return** Trained model *f*; |
| 16: **end function** |

### 2.2.3 Biased Learning

Softmax cross entropy能加速反向传播，同时具有和均方误差(mean square error)相当的性能。在一个n分类的项目中，一个属于c类的样本有一个实数空间内的n维真实值(ground truth)，其真实值满足且。的每一项数值可以看作是样本属于对应类的概率，模型输出的预测标签向量*y*与其定义相同。

在光刻热区检测这一任务中，非热区和热区的真实值分别为。为了输出与真实值相关的损失函数，神经网络输出向量通过如下所示的softmax函数缩放到（0，1）区间内：

 (7)

然后，交叉熵(cross entropy)如下计算：

 (8)

另外，需要特殊定义：

 (9)

因为经过softmax计算后的标签的每一项都是给定样本为非热区和热区的概率，所以可以定义

(10)

(11)

为了提高热区的预测准确率，改动决策边界是一种非常直接的方法，如下所示：

(12)

式中是边界位移常数。然而，这种方法会导致false alarm的大量增加。

如上文所述，非热区和热区的真实值分别为。对于非热区样本而言，整个神经网络将朝着输出为的方向训练。当模型收敛时，对于绝大部分的非热区版图片段，模型将以很高的概率（接近于1）预测其为非热区。但是可以从式(10)可以看到，只要预测概率超过0.5，那样本就会被分类为非热区，因此从某种程度上来说，这个分类器是过于“自信”的。直观地看，一个过于自信的分类器并不一定能提供良好的预测性能，反而会降低预测准确率甚至是模型过拟合。因此，可以做出这样一个假设：光刻热区的预测准确率

定义一个已经收敛的卷积神经网络，其样本真实值向量分别为，其在一个测试集上对于光刻热区的预测准确率为*a*。微调非热区的真实值，能在新的模型上得到新的测试集热区预测准确率和新的false alarm 。在原始模型上改动决策边界使其在测试集上热区预测准确率达到，其对应false alarm为，那么将有

对于任意非热区样本，其有无biased常数的损失函数分别由(13)式和(14)式给出。

(13)

(14)

网络训练的速度取决于损失函数对于神经元权重的梯度计算。根据后向传播，预测向量*y*和模型权重的关系是确定的，因此对于非热区样本，有无biased常数情况下的训练速度取决于损失函数对于预测向量的梯度，由下二式给出：

(15)

(16)

注意到当时，网络会作出错误的预测

(17)

这说明了如果biased常数直接应用于一个随机初始化的神经网络上时，这个网络将更新得比没有biased常数要慢很多。因此本文将biased常数应用于预先训练过的模型上。

Biased常数并不是不受限制的，因为在训练的某个节点上，大多数非热区样本输出将会越过中间线，也就是50%，从而导致false alarm显著增加。这种方法通过牺牲非热区样本的置信度来提升热区检测的准确率，因此称之为BL(Biased Learning)算法。总的来说，BL算法就是迭代地执行变化着的非热区真实值的MGD算法，如算法2所示。在算法 2中，即biased常数，其初始值为0(第1行)，为biased常数递增步数，*t*为BL算法最大迭代次数，光刻热区的真实值被固定在[0, 1]区间内，非热区的真实值为[1 — , ]。经过预训练，模型达到初步收敛后，以的步长递增。继续迭代执行BL算法，直到迭代次数达到*t*(第2-6行)。

|  |
| --- |
| **算法2** BL |
| **Require:** *, , t,* **W**, , , *k*, , |
| 1: *i* 0, 0, [0, 1]; |
| 2: **while** *i < t* **do** |
| 3: ; |
| 4: MGD(**W**, , , *k*, , ); **1** |
| 5: *i* *i + 1*,  *+* |
| 6: **end while** |

相比于前人所做的研究，BL算法利用CNN中ReLU激活函数的优点，在对hotspot的预测准确率和false alarm之间找到了一个绝佳的平衡点[8]。因此，本文将BL算法作为在客户端执行的光刻热区检测算法。

# 服务器端模型聚合算法及系统框架

如前文所述，各个芯片设计厂商出于对各自商业利益的考虑，他们的数据集是不公开、不互通的，这给光刻热区检测研究人员带来极大的困难。研究人员很难获得大量数据进行模型训练，从而极易导致模型过拟合，得到的模型不具有普适性，难以推广。

而联邦学习的问世给以上问题的解决提供了全新的思路。联邦学习客户端于服务器之间传输的信息是训练模型所必需的最小更新。这个更新本身的生存周期是很短暂的，它也不会包含比原始训练数据更多的信息量，通常只会远少于。而且，服务器也并不需要了解更新具体来自于哪一个客户。联邦学习的隐私特性完美地契合了光刻热区检测现存的痛点，故从直觉上来讲，把联邦学习的框架套在光刻热区检测上是可行的。

本章将分三小节具体阐述服务器端运行的模型聚合算法以及客户端和服务器端是如何协作解决问题。第一小节介绍基本的FederatedAveraging模型聚合算法[1]；第二小节将第二章所述的BL算法与FederatedAveraging算法结合，组成一个完整的系统；第三小节简要介绍如何利用ADMM框架[16]对系统整体进行优化。

## FederatedAveraging算法

联邦学习，简单地来说，可以看作是如下的最小化有限和目标：

 (18)

式中*n*为总样本数。对于一个机器学习问题，我们通常会有，即样本在模型上的损失函数。假定把所有样本随机分给*K*个客户，是客户*k*持有的样本集合，有。因此，式(18)还可以写成：

 (19)

如果训练集随机均匀地分给每个客户，那么有，即客户损失函数的期望等于总体损失函数。这是典型的分布式优化算法的独立同分布(IID)假设；而在联邦学习中这种假设往往是不成立的，可以是*f*的任意差的近似，也即非独立同分布(non-IID)。

联邦学习中，在每一个大的轮次里，以比例*C*选取一部分客户，并计算这些客户所持有的所有样本的损失函数的梯度。*C*可以看作是全局的batch size，当*C* = 1时就对应着全局非随机(full-batch)梯度下降。不妨把这种最基本的算法叫做FederatedSGD(FedSGD)。

对于一个典型的FedSGD的实现，有*C* = 1以及一个固定的学习率，对每一个客户计算，即在当前模型下本地数据的平均梯度，然后中央服务器对梯度进行聚合并执行参数更新，因为式(19)，参数更新过程可以写成如下：

(20)

式(20)有一种等价写法。因为

(21)

所以有

(22)

也就是说，每一个客户在本地用本地数据在全局模型上执行一轮梯度下降，服务器对所有客户的模型进行加权平均。接下来就可以对算法添加更多细节。在上传模型至服务器前，客户可以对本地的模型更新（梯度下降）进行如下多次迭代。

(23)

这种算法就是FederatedAveraging(FedAvg)算法。整体计算量由三个参数控制：*C*，每一轮被选取参与模型更新的客户比例；*E*，每一个客户本地迭代次数；*B*，本地迭代的batch size。不妨定义当时，整个本地训练集被当成是一个batch。对于一个样本数量为的一个客户，其本地每一大轮次更新总次数为

(24)

算法完整伪代码见算法3。在算法一开始，先初始化一个全局模型(第2行)，而后选取*m*个客户参与本轮训练(第4-5行)。这*m*个客户各自在本地进行模型训练、参数更新，将得到的模型返回给服务器(第11-18行)，由服务器对模型进行聚合(第9行)，并将聚合后的模型分享给所有客户作为下一轮训练的基础(第10行)。

对于一般的非凸目标，参数空间中的平均模型可以产生一个任意坏的模型。

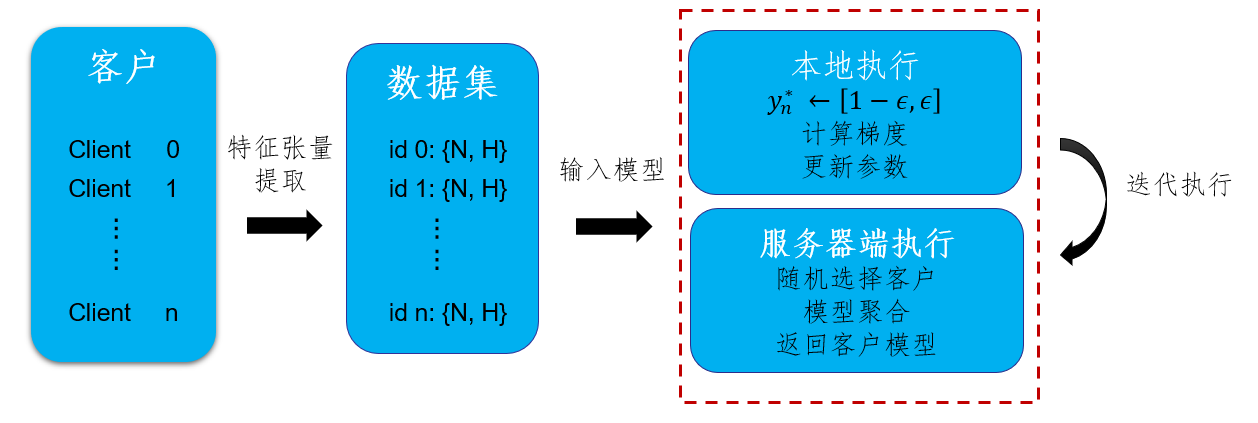
|  |
| --- |
| **算法3** FederatedAveraging. |
| 1: 服务器端执行**:** |
| 2: 初始化 |
| 3: **for** each round *t* = 1, 2, … **do** |
| 4: *m* max*(,* 1*)*; |
| 5: (任意选取*m*个客户); |
| 6: **for** each client **in parallel** **do** |
| 7: ClientUpdate(*k*, ); |
| 8: **end for** |
| 9: ;  10: 发送至每一个客户 |
| 11: **ClientUpdate**(*k*, ): //在客户编号为*k*上执行 |
| 12: (把按照batch size为*B*的大小划分) |
| 13: **for** each local epoch *i* from 1 to *E* **do** |
| 14: **for** batch *b* **do** |
| 15: ; |
| 16: **end for**  17: **end for** |
| 18:  **return**  to server |

前人实验表明，FedAveraging算法对于不平衡和非独立同分布的大数据集是非常鲁棒的[1]，因此本文将其作为服务器端模型聚合算法。

## 系统框架

到目前为止，客户端光刻热区检测算法和服务器端模型聚合算法都已经完整地阐述过了。把二者结合起来也不是一件特别困难的事，只需将算法3中的ClientUpdate改成算法2 BL算法即可。

如图1所示，有*n*个芯片厂商作为系统的*n*个客户，对每个客户的芯片版图进行特征张量提取，作为算法的数据集。选择部分客户在本地各自迭代地执行BL算法而后同步上传模型至中央服务器。中央服务器对模型进行聚合，将聚合后的模型同步返回给所有客户，在新的一轮中，客户在新的模型上继续迭代执行BL算法。



**图1 基于联邦学习的光刻热区检测基本系统框架**

算法伪代码见算法4。与算法3相比，算法4将ClientUpdate完全替换成算法2的BL算法。

|  |
| --- |
| **算法4** 基于联邦学习的光刻热区检测基本框架 |
| 1: 服务器端执行**:** |
| 2: 初始化 |
| 3: **for** each round *t* = 1, 2, … **do** |
| 4: *m* max*(,* 1*)*; |
| 5: (任意选取*m*个客户); |
| 6: **for** each client **in parallel** **do** |
| 7: ClientUpdate(*k*, ); |
| 8: **end for** |
| 9: ;  10: 发送至每一个客户 |
| 11: **ClientUpdate**(*k*, ): //在客户编号为*k*上执行 |
| 12: BL(*k*, ); **2** |
| 13:  **return**  to server |

需要注意的是，因为本文旨在解决现有光刻热区检测研究中的痛点并验证联邦学习应用于光刻热区检测的可行性，故本文所阐述的系统是不考虑通信成本的。也就是说所有客户和服务器之间的通信是稳定实时的，没有信息量损耗的，不会出现某个或某几个客户会中途退出训练，各客户与服务器之间的通信完全相同。

## 系统优化

### 3.3.1 ADMM优化系统

交替方向乘子法(Alternating Direction Method of Multipliers, ADMM)是一种求解优化问题的计算框架，适用于求解分布式凸优化问题。ADMM通过分解协调(Decomposition-Coordination)过程，将大的全局问题分解为多个较小、较容易求解的局部子问题，并通过协调子问题的解而得到大的全局问题的解。ADMM的适用环境与本文所研究的问题相当契合，可以尝试将ADMM应用于本系统，优化整体算法。

ADMM最早分别由Glowinski & Marrocco及Gabay & Mercier于1975年和1976年提出，并被Boyd等人于2011年重新综述并证明其适用于大规模分布式优化问题[16]。

本文将就ADMM如何应用于本系统作简要介绍，不作具体论证。应用ADMM优化系统算法如算法5所示。与前文所述算法相比，算法5在客户端的参数更新过程添加了一个对偶变量*y*以约束模型参数的变化。

|  |
| --- |
| **算法5** ADMM优化系统 |
| 1: 服务器端执行**:** |
| 2: 初始化 |
| 3: **for** each round *t* = 1, 2, … **do** |
| 4: *m* max*(,* 1*)*; |
| 5: (任意选取*m*个客户); |
| 6: **for** each client **in parallel** **do** |
| 7: ClientUpdate(*k*, ); |
| 8: **end for** |
| 9: ;  10: 发送至每一个客户 |
| 11: **ClientUpdate**(*k*, ): //在客户编号为*k*上执行 |
| 12: BL(*k*, ); |
| 13: ; |
| 14:  **return**  to server |

### 3.3.2 差分隐私的应用

最近对成员攻击和模型反转攻击的研究从多个维度揭示了潜在的隐私风险。深度神经网络有大量的隐藏层，从而产生了足够大的有效容量能够将个体数据的某些细节编码为模型参数，甚至可以存储整个数据集[17]。有许多研究已经表明，个体信息是可以有效地从神经网络中提取出来[18][19]。因此随着深度学习应用和深度学习服务的广泛部署，隐私问题的严重性越来越被凸显出来。为了保护本系统中客户的模型不受攻击，还需更加强力的隐私保护机制。差异隐私（Differential privacy, DP）是保证聚合数据库上算法隐私的一个强有力的标准。DP描述了两个输入数据集之间的输出差异，这两个数据集最多有一个元素的差异。通俗地说，它要求对数据集中的单个数据点的任何更改只能对算法的输出造成统计上不显著的更改。

**定义1** (差分隐私)：一个随机机制提供了-差分隐私如果对任意两个只有一个元素差异的相邻数据集，映射满足

(25)

实现差分隐私的标准方法是灵敏度方法，该方法向输出添加一些与查询函数的灵敏度成比例的噪声。灵敏度为对由于单个数据集样本的更改而导致的输出的最大更改的度量。

**定义2** (灵敏度)：一个查找函数的灵敏度定义为：

 (26)

式中为任意两个只有一个元素差异的相邻数据集，表示L1或L2范数。

算法6描述了如何将差分隐私应用在本系统上。在本地训练的每一个迭代周期，计算一个batch里每个样本的梯度，以梯度的L2范数作适当放缩并计算平均值，而后加上高斯噪声以保护隐私，再以此执行一步参数更新。

|  |
| --- |
| **算法6** 差分隐私优化系统 |
| 1: 服务器端执行**:** |
| 2: 初始化 |
| 3: **for** each round *t* = 1, 2, … **do** |
| 4: *m* max*(,* 1*)*; |
| 5: (任意选取*m*个客户); |
| 6: **for** each client **in parallel** **do** |
| 7: ClientUpdate(*k*, ); |
| 8: **end for** |
| 9: ;  10: 发送至每一个客户 |
| 11: **ClientUpdate**(*k*, ): //在客户编号为*k*上执行  12: |
| 13: 选取*m*个样本; |
| 14: **for** *i* 1, 2, …, *m* **do** |
| 15: backprop(); |
| 16: **end for** |
| 17: ;  18: ;  19: **;** |
| 19:  **return**  to server |

# 实验结果

## 实验环境与基本实验参数

本文使用Tensorflow库在Python中实现了系统框架，并在配有Xeon E5-2650 CPU和Nvidia 1080Ti GPU的平台上进行训练和测试。为了能充分地对整个系统进行评估，本文使用了5个数据集，其中一个为ICCAD 2012 Contest，其余四个为工业界的数据集Industry1-Industry4。这5个数据集的基本信息见表2。表格四列分别列出训练集和测试集中光刻热区(hotspot)和非光刻热区(non-hotspot)的总数。

**表2 数据集基本信息**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集 | 训练集 | | 测试集 | |
| HotSpot# | Non-HotSpt# | HotSpot# | Non-HotSpt# |
| ICCAD  Industry1  Industry2  Industry3  Industry4 | 1204  34281  15197  24776  3629 | 17096  15635  48758  49315  80299 | 2524  17157  7520  12228  942 | 138848  7801  24457  24817  20412 |

## 模型训练

表3列出详细列出了模型训练的基本参数。“Adam”是一种高性能的神经网络优化器，它能使模型更快地收敛[20]。参数和*k*表明学习率每过*k*个迭代周期衰减至。特征张量通道表示在特征张量提取过程中除去高频分量后剩余的元素数。

表3 模型训练基本参数

|  |  |
| --- | --- |
| 参数配置 | 值 |
| Optimizer  学习率初值  学习率衰减系数  学习率衰减步数*k*  Batch Size  特征张量通道 | Adam  0.001  0.75  6400  64  32 |

## 模型测试

首先，对ICCAD，Industry1-4五个数据集（数据集基本信息见表2）分别执行BL算法(算法2)，其结果作为对照组。实验结果在表4中列出。“Accu(%)”和“FA#”分别表示对hotspot的预测准确率和false alarm总数（定义见1.1.1）。“FA(%)”为false alarm总数与测试集中non-hotspot个数之比。

**表4 BL算法在各个数据集上的实验结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 数据集 | ICCAD | Industry1 | Industry2 | Industry3 | Industry4 |
| Accu (%) | 97.4 | 98.5196 | 94.9601 | 91.8629 | 97.8769 |
| FA# | 2312 | 113 | 1464 | 3750 | 76 |
| FA(%) | 1.6651 | 1.4485 | 5.986 | 15.1106 | 0.3723 |

从表中不难看出，各个数据集差异是比较大的，尤其是Industry2，Industry3，同样的训练参数，由Industry2，Industry3数据集训练出来的模型在其测试集上表现出来的性能稍逊于其他数据集。由于Industry数据是商业机密，无从得知版图之间的具体差异，只能初步猜测Industry2，Industry3版图相比于其他数据集更为复杂难以预测，需要进一步的参数调整。

下一步的实验是另外一个对照组实验。将ICCAD，Industry1-4五个数据集合并成一个数据集，执行BL算法(算法2)。实验结果在表5列出。为了更直观地显示将数据集合并带来的影响，图2简单对比了这两次实验。不难看出，合并数据集会对模型在测试集上的表现造成负面影响：模型在所有测试集上对hotspot的预测准确率有所降低，false alarm也有不同程度的增加。原因不难猜测，因为不同芯片设计公司设计的版图可能设计风格，设计工艺不同，设计复杂程度不同。用一套训练参数来同时训练这么一个大的合并数据集，是可以预见到性能的下降的。

**表5 合并数据集执行BL算法实验结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 以ICCAD&Industry1-4训练模型 | | | | | |
| 测试集 | ICCAD | Industry1 | Industry2 | Industry3 | Industry4 |
| Accu (%) | 78.0507 | 96.6944 | 93.391 | 91.4699 | 75.5829 |
| FA# | 10431 | 864 | 3823 | 5237 | 347 |
| FA(%) | 7.5125 | 11.0755 | 15.6315 | 21.1025 | 1.6999 |

**图2 数据集合并前后效果对比**

为了验证3.2中所描述的框架是合理有效的，下一步需要进行联邦学习实验(算法4)。把每一个数据集各自当成一个客户，用如表3所示的相同的训练参数各自进行本地训练，服务器对5个上传的模型进行加权平均（即每轮5个客户都参与训练），执行多轮后得到实验结果。而后将每个数据集分为5个客户，10个客户再次进行联邦学习实验以观察客户数量对实验结果的影响（每轮选取50%客户参与训练）。实验结果见表6。

**表6 联邦学习不同客户数在各个数据集上的实验结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 以ICCAD&Industry1-4训练联邦学习模型 | | | | | | |
| 训练集 | ICCAD | | | Industry1 | | |
| 客户数 | 1 | 5 | 10 | 1 | 5 | 10 |
| Accu (%) | 94.5721 | 88.9065 | 86.2916 | 79.7674 | 75.9515 | 73.0005 |
| FA# | 15107 | 12366 | 11498 | 591 | 474 | 460 |
| FA(%) | 10.8802 | 8.9061 | 8.281 | 7.576 | 6.0761 | 5.8967 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Industry2 | | | Industry3 | | | Industry4 | | |
| 1 | 5 | 10 | 1 | 5 | 10 | 1 | 5 | 10 |
| 92.2473 | 93.9229 | 91.1170 | 79.6215 | 78.6147 | 78.2450 | 88.2166 | 94.3737 | 89.1720 |
| 3903 | 3909 | 3560 | 4471 | 4128 | 4344 | 690 | 1096 | 887 |
| 15.9586 | 15.9832 | 14.5562 | 18.0159 | 16.6338 | 17.5041 | 3.3804 | 5.3694 | 4.3455 |

图3更直观地表现了客户数量对模型性能的影响。从整体趋势上来看，客户数量增加会同时降低对hotspot的预测准确率和false alarm，这是否带来性能上的改变需要作ODST（定义见1.1.1）计算，本文不作记述。同时可以观察到，false alarm与对hotspot的预测准确率存在一定正相关关系，这也是符合直观感受的。

1. (b) (c)

(d) (e)

**图3 客户数量对模型性能的影响。训练集分别为(a)ICCAD，(b)Industry1，(c)Industry2，(d)Industry3，(e)Industry4。**

取表6每一测试集项的中间一列和表5的实验数据制成对比图，如图4。可见联邦学习模型虽然并没有带来模型性能上的提升，但对hotspot的预测准确率和false alarm大体上与非联邦学习模型相近，因此很容易得出结论：将BL算法(算法2)与FederatedAveraging(算法3)结合，用联邦学习的思想来做光刻热区检测是切实可行的。

**图4 联邦学习与非联邦学习实验结果对比**

后续以ADMM算法对系统进行优化(算法5)，每个数据集当成一个客户，用如表3所示的训练参数进行训练。实验结果见表7。取表6每一测试项的第一列和表7的实验数据制成对比图，见图5。从图5看来，ADMM算法并没有给系统整体性能带来明显地提升。

**表7 ADMM优化联邦学习框架实验结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ADMM优化系统 | | | | | |
| 测试集 | ICCAD | Industry1 | Industry2 | Industry3 | Industry4 |
| Accu (%) | 92.3876 | 78.1703 | 94.7557 | 77.8168 | 84.3518 |
| FA# | 14004 | 573 | 3969 | 4193 | 622 |
| FA(%) | 10.0858 | 7.3452 | 16.2285 | 16.8957 | 3.0472 |

**图5 ADMM优化前后实验结果对比**

对于在系统中添加差分隐私，本文只是做了一次简单的试验。在梯度计算中添加了*N*(0, 4)分布的高斯噪声以观察差分隐私对系统的影响, 每个数据集当成一个客户，用如表3所示的训练参数进行训练。实验结果在表8中列出。取表6每一测试集项的第一列和表8的实验数据制成对比图，见图6。可见，在梯度计算中添加了*N*(0, 4)分布的高斯噪声会对系统整体性能造成一定负面影响，这是和隐私保护之间的权衡。

**表8 联邦学习框架加入差分隐私实验结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 差分隐私优化系统 | | | | | |
| 测试集 | ICCAD | Industry1 | Industry2 | Industry3 | Industry4 |
| Accu (%) | 90.3312 | 84.79 | 87.6683 | 73.146 | 84.5571 |
| FA# | 16289 | 611 | 4144 | 4206 | 731 |
| FA(%) | 11.7315 | 7.8323 | 16.944 | 16.9481 | 3.5812 |

m

**图6 差分隐私优化前后实验结果对比**

# 结论及未来展望

为了解决如今光刻热区检测领域中公共数据集稀少与提高模型预测准确率的矛盾，本文提出了一种基于联邦学习的光刻热区检测系统框架。在执行该框架之前，对数据集采用一种特征张量提取的算法提取其频域特征，作为后续框架输入。该框架采用一种高性能的BL算法作为客户在本地运行的光刻热区检测算法，该算法对非光刻热区样本的真实值作出偏置，一定程度上牺牲了非光刻热区样本的损失函数，但该算法保证本地模型在本地数据集上表现最佳；采用FederatedAveraging算法作为服务器运行的模型聚合算法，该算法证明了全局模型参数更新等价于所有客户模型的加权平均，其对于不平衡、非独立同分布的数据集有很鲁棒的良好表现。本文进一步以ADMM算法尝试优化了整个系统框架，并探索性地在系统中添加了差分隐私。

实验结果表明，在本地运行光刻热区检测算法，在服务器端对模型进行聚合的思想是可行的，虽然没有性能上的提升，但是能够解决当前数据集稀少和隐私的问题。ADMM算法优化系统相比于优化之前，模型在各数据集上并没有更好的表现，原因可能是多方面的。这需要对ADMM算法作更深入的学习与研究，思考其应当如何应用于本问题，更好地优化系统。在系统中添加差分隐私能够进一步地保护客户的数据隐私，但是要付出模型性能下降的代价。相信在模型性能和隐私之间存在一个完美的平衡，在未来将继续学习差分隐私的相关概念和计算，尝试添加不同类型、不同参数的噪声，对在本地训练添加噪声和在模型上传添加噪声进行实验对比，期望在隐私保护的同时尽量少地损失模型性能。

参考文献

1. McMahan H B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[J]. arXiv preprint arXiv:1602.05629, 2016.
2. Wen W Y, Li J C, Lin S Y, et al. A fuzzy-matching model with grid reduction for lithography hotspot detection[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2014, 33(11): 1671-1680.
3. Yu Y T, Chan Y C, Sinha S, et al. Accurate process-hotspot detection using critical design rule extraction[C]//Proceedings of the 49th Annual Design Automation Conference. ACM, 2012: 1167-1172.
4. Guo J , Yang F , Sinha S , et al. Improved tangent space based distance metric for accurate lithographic hotspot classification[J]. 2012.
5. Drmanac, Dragoljub Gagi, Liu, Frank, Wang, LiC. Predicting variability in nanoscale lithography processes.[J]. 2009.
6. Ding, Duo, Torres, Andres J, Pikus, Fedor G, et al. High performance lithographic hotspot detection using hierarchically refined machine learning[C]// Asia & South Pacific Design Automation Conference. IEEE, 2011.
7. Zhang H, Yu B, Young E F Y. Enabling online learning in lithography hotspot detection with information-theoretic feature optimization[C]//Proceedings of the 35th International Conference on Computer-Aided Design. ACM, 2016: 47.
8. Yang H, Su J, Zou Y, et al. Layout hotspot detection with feature tensor generation and deep biased learning[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2018, 38(6): 1175-1187.
9. Chen R, Zhong W, Yang H, et al. Faster Region-based Hotspot Detection[C]//Proceedings of the 56th Annual Design Automation Conference 2019. ACM, 2019: 146.
10. McMahan H B, Ramage D, Talwar K, et al. Learning differentially private recurrent language models[J]. arXiv preprint arXiv:1710.06963, 2017.
11. 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): 12.
12. Yao A C. Protocols for secure computations[C]//23rd annual symposium on foundations of computer science (sfcs 1982). IEEE, 1982: 160-164.
13. Dwork C. Differential privacy[J]. Encyclopedia of Cryptography and Security, 2011: 338-340.
14. Rivest R L, Adleman L, Dertouzos M L. On data banks and privacy homomorphisms[J]. Foundations of secure computation, 1978, 4(11): 169-180.
15. Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. nature, 1986, 323(6088): 533-536.
16. Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foundations and Trends® in Machine learning, 2011, 3(1): 1-122.
17. Zhang C, Bengio S, Hardt M, et al. Understanding deep learning requires rethinking generalization[J]. arXiv preprint arXiv:1611.03530, 2016.
18. Fredrikson M, Jha S, Ristenpart T. Model inversion attacks that exploit confidence information and basic countermeasures[C]//Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. 2015: 1322-1333.
19. Shokri R, Stronati M, Song C, et al. Membership inference attacks against machine learning models[C]//2017 IEEE Symposium on Security and Privacy (SP). IEEE, 2017: 3-18.
20. Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
21. Mcdonald R T , Hall K , Mann G . Distributed Training Strategies for the Structured Perceptron[C]// Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 2-4, 2010, Los Angeles, California, USA. Association for Computational Linguistics, 2010.
22. Povey D, Zhang X, Khudanpur S. Parallel training of deep neural networks with natural gradient and parameter averaging[J]. arXiv preprint arXiv:1410.7455, 2014.
23. Zhang S, Choromanska A E, LeCun Y. Deep learning with elastic averaging SGD[C]//Advances in neural information processing systems. 2015: 685-693.
24. Neverova N, Wolf C, Lacey G, et al. Learning human identity from motion patterns[J]. IEEE Access, 2016, 4: 1810-1820.
25. Shokri R, Shmatikov V. Privacy-preserving deep learning[C]//Proceedings of the 22nd ACM SIGSAC conference on computer and communications security. 2015: 1310-1321.

附录

无

**本科生毕业论文（设计）任务书**

**一、题目：基于联邦学习的光刻热区检测**

**二、指导教师对毕业论文（设计）的进度安排及任务要求**

**任务要求：**

**（1） 熟悉联邦学习和光刻热区检测的基本概念等。**

**（2） 学习联邦学习和隐私保护的框架并熟悉基于Tensor Flow的编程环境。**

**（3） 设计应用于热区检测的联邦学习框架，根据隐私保护需求进行优化。**

**（4） 根据实际数据设计实验，进行数据仿真与功能验证分析。**

**（5） 完成毕业设计的总结工作和毕业论文的撰写。**

**进度安排（可以按一周或两周进行安排）：**

**2019-11-11至2019-11-24：查阅相关资料**

**2019-11-25至2019-12-08：调研国内相关方面的最新研究进展情况。**

**2019-12-09至2019-12-22：完成资料翻译。**

**2019-12-23至2020-01-05：撰写文献综述、文献翻译和开题报告，且准备开题答辩。**

**2020-01-06至2020-01-16：开题答辩。**

**2020-01-18至2020-02-20：寒假。**

**2020-02-24至2020-03-08：完成联邦学习和光刻热区检测的研究并进行仿真分析。**

**2020-03-09至2020-04-19：设计应用于热区检测的联邦学习框架，根据隐私保护需求进行优化，进行仿真分析。**

**2020-04-20至2020-05-03：构建样本测试实验，完成针对实际数据的指标测量，对结果进行分析比较验证。**

**2020-05-04至2020-05-17：撰写、修改毕业论文。**

**2020-05-18至2019-06-07：查重、论文答辩。**

**起讫日期：2019年11月11日-2020年7月11日**

**指导教师**（**签名）**  **职称**

**三、系或研究所审核意见**

**任务要求适中，进度安排合理。**

**负责人**（**签名）**

**年 月 日**

**毕 业 论 文（设计） 考 核**

**一、指导教师对毕业论文（设计）的评语**

**该论文探讨了一种利用联邦学习实现光刻热区检验的方法。总体而言，论文调研充分，推导可信，框架完善，但是方法本身可以更深入讨论以提高论文全面性。**

**指导教师(签名）**

**2020年 06月 15日**

**二、答辩小组对毕业论文（设计）的答辩评语及总评成绩**

**论文研究了基于联邦学习的光刻热区检测。**

**论文主要的研究成果是：验证应用光刻热区检测的思想进行光刻热区检测是可行的，在系统框架中添加差分隐私能够更有效地保护客户隐私。**

**论文工作表明，作者理论基础夯实，具有较强的科研能力。论文书写规范，引用资料详实，理论分析充分，仿真实验结果充分验证了假设，论文以达到本科生学士学位论文要求。**

**论文答辩表达清楚，回答问题正确。答辩委员会经过讨论，一致同意王唯一同学本科毕业答辩通过，建议授予学士学位。**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **成绩**  **比例** | **文献综述**  **占（10%）** | **开题报告**  **占（15%）** | **外文翻译**  **占（5%）** | **毕业论文（设计）质量及答辩**  **占（70%）** | **总评成绩** |
| **分值** | 8 | 13 | 4 | 63 | 88 |

**答辩小组负责人（签名）**

**2020年 07月 06日**

第二部分

**文献综述和开题报告**