西安交通大学电信学部

博士研究生学科发展基础文献集阅读总结报告

报告题目 面向数据异构的联邦学习安全机制与优化聚合研究

姓　　名 岳高峰

学　　科 电子信息

指导教师 言涅

日　　期 2023 年 6 月 20 日

“学科发展基础文献集”是本学科研究生教育和学术研究所需的基础的专著、论文、报告等的目录汇总，学科文献阅读是引领研究生特别是博士生夯实学术根基、迅速进入学科专业研究领域和培养研究生科研素养的重要手段，是课程知识扩展和补充的重要方式。“学科发展基础文献集”的建立旨在帮助研究生更高效地、全面地了解学科历史脉络及发展前沿，掌握核心知识及研究方法，为研究生拓展学术视野打下基础。

研究生自入学开始在本学科发展基础文献集平台进行文献阅读学习。通用文献适用于该学科所有学生；拓展文献适用于对应方向学生，其他学科或方向学生选读。通用及拓展文献的学习篇目和数量由学科或学院选定。理工医类学科文献学习时间一般不少于 60 小时。文献集阅读以学生自学为主，学科可辅以导读和讲座等形式。

内容包括：前言，主题内容，总结，参考文献列表

字数要求：不低于 5000 字（含参考文献列表）

文献集阅读篇数由各学院自行确定最低篇数标准（列在参考文献列表中的文献篇数）

摘要

随着物联网和数据成为研究热点，物联网连接了传感器、设备、网络和云计算，数据成为其核心。在这一背景下，机器学习被广泛应用于处理大规模数据集，但隐私保护法的出台限制了标准分布式机器学习。为解决隐私保护问题，联邦学习应运而生。联邦学习是一种特殊的分布式机器学习方法，允许在分布式环境中进行模型训练，本地计算减少了数据传输和存储开销。联邦学习的提出为保护个人隐私和数据安全提供了解决方案，吸引了广泛的关注和研究。然而，联邦学习面临着隐私和安全性、数据异质性和可扩展性等挑战。通过克服这些挑战，联邦学习有望在数据驱动的机器学习中发挥重要作用，为实现隐私保护和高性能模型训练提供有效的解决方案。

。

前言

随着科技的迅猛发展，物联网和数据成为当今社会中备受关注的研究领域。物联网是一种将传感器、设备、网络和云计算等技术相结合，实现物体之间相互连接和通信的网络系统[1]。数据则是物联网的核心[2]，通过收集、存储和分析大量的数据，我们可以获得有关物体状态、环境信息和人类行为等方面的有价值的洞察和决策支持。在过去的几年里，物联网的应用范围不断扩大，涵盖了各个领域，如智能家居、智能城市、智能交通、智能医疗等[3]，这也导致大量本地数据的产生。

为了分析大量的数据和获取有价值的信息，人工智能被应用于边缘计算，因为它在预测未来事件方面具有非凡的能力。作为最具代表性的人工智能技术，机器学习被应用于众多领域（例如，物体检测、语音识别和自然语言处理）。处理大量数据的方法中，机器学习被认为是最可行的[4]。标准的分布式机器学习是一种将机器学习算法和技术与分布式计算相结合的方法，旨在处理大规模数据集和复杂模型训练的问题[5]。分布式机器学习的核心思想是将大规模数据集划分为多个子集，并将这些子集分布在不同的计算节点上。每个计算节点都可以独立地训练一个模型，然后通过合并和整合各个节点上的模型参数来生成最终的全局模型。这种分布式训练的方法可以大大减少训练时间，特别是在处理大规模数据集和复杂模型时更为明显，但是节点会按照一定的规律和其他工作节点进行通信（通信的内容主要是子模型参数或者参数更新）。

然而随着一些国家和地区出台了专门的隐私保护法，如欧洲的《通用数据保护条例》（GDPR）[6]、美国的《加州消费者隐私法》（CCPA）[7]等，加大了对个人隐私权利的重视和保护的加强。严格的隐私保护法限制了对数据的直接使用[8]，这些法律法规的目的是确保在信息社会中，个人的隐私得到充分的尊重和保护，防止个人信息被滥用、泄露或不当收集。同时隐私保护法要求数据处理者采取合理的安全措施，保护个人信息的安全性和机密性，防止数据的泄露、损毁或被未经授权的人访问。在这种情况下，数据裸露下的标准分布式机器学习变得不可用[9-10]，因为其没有隐私保护的能力，会导致个人敏感数据泄露。

因此，Google 的研究团队在 2016 年提出联邦学习来解决机器学习中隐私保护的问题[11]。联邦学习是一种新型的特殊的分布式机器学习的方法，旨在解决数据隐私和安全性的问题[12]。联邦学习允许在分布式环境中进行模型训练，每个设备或数据持有者可以在本地进行计算，减少了数据传输和存储的开销。同时联邦学习通过在不共享数据的情况下进行模型训练，使得不同数据持有者可以协作学习，共同提高模型性能。

联邦学习迅速引起了学术界和工业界的广泛关注，并在之后的几年中得到了持续的研究和发展。联邦学习的原理和应用也在不同领域的研究和实践中得到了广泛应用和探索。随着对数据隐私和安全性的需求增加，联邦学习在数据驱动的机器学习中扮演着重要的角色，并成为了研究的热点之一。联邦学习作为一种分布式机器学习方法，也面临着一些挑战和困难[13]，包括以下几个方面：

1. 隐私和安全性：联邦学习的一个关键目标是保护参与者的数据隐私和安全性。然而，在实践中，仍然存在一些隐私攻击的风险，例如模型反推攻击和信息泄露攻击。确保在联邦学习过程中的数据安全和隐私保护，是一个重要而复杂的挑战。
2. 数据异质性：在联邦学习中，参与者通常拥有不同的数据分布和特征。这种数据的异质性可能导致模型的性能下降，因为模型在一个参与者的数据上训练的效果可能无法很好地泛化到其他参与者的数据上。解决这一问题的挑战在于如何有效地融合不同数据来源的信息，以提高全局模型的性能。

3. 联邦学习的可扩展性：随着参与者数量的增加和数据规模的扩大，联邦学习的可扩展性成为一个挑战。在大规模联邦学习中，如何高效地管理和协调大量参与者的计算和通信，以及保持模型的一致性和准确性，是一个需要解决的问题。

主题一：一种保护隐私和高度安全的联邦学习方案

联邦学习（FL），允许多个参与者共同训练机器学习模型，通过避免暴露本地数据来加强隐私保护。然而，在联邦学习过程中，攻击可能针对上传的模型更新，以窃取客户的敏感信息[14]。同时，客户可能有不诚实的行为或上传恶意的数据来阻碍全局模型的训练[15]。

近年来，在联邦学习中保护隐私的问题得到了极大的关注。缓解客户端隐私泄露的最先进的解决方案是通过在联邦学习中使用差分隐私。Shi 等人提出了一种新的隐私保护方案，称为 HFL-LDP[16]，它建立在云-边缘-客户的分层结构上。这种方法引入了时刻计算的概念，可以对累积的隐私损失进行监控。文章表明，所提出的算法在适当的扰动水平下满足了 LDP 的要求，并在隐私和效用之间取得了平衡，在提高性能的同时提供了更高的隐私水平。

Hu 等人于 2022 年提出了联邦学习与稀疏化模型扰动（Fed-SMP）的隐私破坏范式[17]，该范式在客户端层面采用 DP 来保持模型的准确性。Fed-SMP 策略有效地结合了随机稀疏化和局部模型的 Top-k 稀疏化。同时，本文还加强了隐私与精度的权衡，展示了 Fed-SMP 的有效性，并降低了通信成本。

Jiang 等人于 2022 年介绍了 FLASHE 方案[18]，该方案通过使用 HE 和模块化加法运算来掩盖局部梯度更新。作者放弃了非对称密钥的设计以优化计算效率，但 FLASHE 适应于跨域的联邦学习，而不是各种类型的联邦学习。Mou 等人提出了一种基于 SMC 和 DP 的特殊隐私保护联邦学习方案[19]，SMC 的加入减少了 DP 引入的噪声，与基于 DP 的联邦学习相比，模型的准确性得到了提高，同时在联邦学习过程中仍有能力保护客户的隐私。然而，由于使用了加密，基于 SMC 的联邦学习增加了计算成本和通信开销。

在联邦学习中，有多种类型的不诚实客户端攻击可能发生，包括物理离线攻击、中毒客户的攻击和来自模型集中器的攻击等等。Hei 等人提出了一个基于区块链-联邦学习的云入侵检测系统（BFL-CIDS）来解决分布式环境中的恶意攻击联邦学习[20]，该方案引入了云计算机中心的使用，利用区块链技术存储训练过程参数和行为信息，这种方法减少了错误警报的可能性，提高了联邦学习的准确性。

考虑到防御有针对性的模型中毒攻击，Gu 等人于 2021 年提出了一个名为 Fedcvae 的策略[21]。在这种方法中，中央服务器采用条件变异自动编码器（VAE）来检测和消除无监督联邦学习系统中的恶意模型，Fedcvae 克服了变异自动编码器模型的一些弱点，它只采用特定的恶意攻击，例如同值攻击。在同一年，Cao 等人开发了一种安全的联邦学习方法[22]，通过利用随机选择的客户端子集来抵御恶意客户端，这种方法在某些情况下对数量有限的恶意客户端是稳健的；作者还介绍了一种蒙特卡洛算法来计算认证的安全级别，并在 MNIST 数据集

上实现了对恶意客户端的防御性能的提高。

在实际应用中，Abubaker 等人于 2022 年提出了一种基于超越第五代（Beyond fifth-Generation，B5G）的区块链技术检测和删除物联网（IoT）中恶意节点的解决方案[23]。这个解决方案对于在所有实体之间建立信任非常重要，本文的创新之处在于实现了带有级联加密的组合数字签名，以确保全球服务器和本地客户端的不可抵赖性。

大多数现有的隐私保护方法很少处理来自两方面的威胁。因此，提出一种新型的联邦学习方案迫在眉睫，与现有方法相比，所提出的技术必须可以保证客户的隐私保护，并在联邦学习处理过程中防止一些威胁。

主题二：大规模异构客户端的优化聚合

在联邦学习训练中，SGD 方向是由所有参与的客户协同决定的，他们拥有 Non-IID 数据和异质系统的能力，一些客户的本地模型更新方向可能会偏离全局模型的最陡峭的 SGD 方向（即客户漂移[24]），从而导致收敛缓慢或模型精度低。

Xu 等人利用获得的见解开发了一种新的算法[25]，只利用当前可用的无线信道信息，但可以实现长期的性能保证。Yoshida 等人[26]提出了当客户的计算和通信资源无法估计时，联合学习（FL）的客户选择方法；该方法利用移动客户的丰富数据和计算资源来训练机器学习（ML）模型，而无需在中央系统中收集他们的数据。Huang 等人将保证公平性的客户端选择建模为 Lyapunov 优化问题[27]，然后提出了

一种基于 C2MAB 的方法来估计每个客户端和服务器之间的模型交换时间，在此基础上设计了一种公平性保证算法，称为 RBCS-F，用于解决问题。Abdulrahman 等人提出 FedMCCS，一种基于多标准的 FL 中的客户端选择方法[28]。Yu 等人提出通过选择合适的客户端并在 CPU 频率和传输功率方面分配适当的资源，动态调整和优化所选客户端数量最大化和客户端总能耗最小化之间的权衡[29]。Zhang 等人提出一种经验驱动的控制算法[30]，自适应地选择客户端设备参与每一轮 FL。

上述方法都是在小规模且各个客户端异构性较弱的情况下使用，对于大规模异构客户端的选择与聚合显得无能为力，这个方面也是亟待研究的。

主题三：智慧城市中联邦学习的应用及测试

随伴随着城市公共交通的电气化，供电系统正面临着越来越严峻的压力。在各种公共交通车辆中，电动汽车的快速扩张给许多电动汽车的先驱国家，如中国，印度和美国的电力供应方面带来了前所未有的压力。准确的电动汽车能源需求预测是有效协调充电服务提供商、充电站和电动汽车之间能源供应的基础，这被视为缓解城市电力供应压力的关键。然而，由于电动汽车充电行为的不确定性和复杂性，这个问题的解决是不容易的。近年来，许多工作被提出来预测电动车的能源需求[31-33]。

一般来说，他们用统计模型来描述电动车的充电行为，并利用建立的模型来预测电动车的能量需求变化。为了捕捉各种特征之间更复

杂的非线性关系，一些基于机器学习的方法已经被引入到电动汽车的能源需求预测中。各种基于机器学习的模型，如深度神经网络（DNN）、卷积神经网络（CNN）和长短时记忆（LSTM）网络，被用来提取与电动汽车或充电站能耗相关的复杂模式和高维特征，并取得了比传统统计模型高得多的能源需求预测性能。

然而，所有这些方法都严重依赖直接从电动车上收集的能源使用数据或移动轨迹，这导致了电动车司机的重大隐私泄漏风险。因此，上述方法不适用于现在的实际情况。

为了克服上述问题，基于联邦学习的隐私保护的能源需求预测框架急需建立。

总结

联邦学习是一种分布式机器学习方法，通过在本地设备上进行模型训练，保护数据隐私的同时，实现模型的全局性能提升。它面临着数据异质性、不平衡数据和标签、通信和计算开销、隐私和安全性以及可扩展性等挑战。联邦学习的意义在于解决了数据隐私保护和大规模分布式数据处理的问题，促进了数据共享与协作学习。联邦学习有望进一步推动数据驱动的机器学习发展，为实现安全、高效、个性化的模型训练提供解决方案，并且在智慧城市、医疗保健、物联网等领域具有广泛的应用前景。

参考文献

[1] Joel Höglund. Public key infrastructure and its applications for

resource-constrained IoT [D]. Uppsala University, 2023, 2-8.

[2] Jitendra Bhatia, Kiran Italiya, Kuldeepsinh Jadeja, et al. An overview of fog data analytics for IoT applications[J]. Sensors, 2023, 23(1): 199.

[3] bdullah Sevin, Abdu Ahmed Osman Mohammed. A survey on software implementation of lightweight block ciphers for IoT devices[J]. Journal of Ambient Intelligence and Humanized Computing, 2021, 14(3):1801-1815.

[4] He Li, Kaoru Ota, Mianxiong Dong. Learning IoT in edge: deep learning for the internet of things with edge computing[J]. IEEE Network, 2018, 32(1):96-101

[5] Qingshan Liu, Zhigang Zeng, Yaochu Jin. Distributed machine learning, optimization and applications[J]. Neurocomputing, 2022, 489: 486-487.

[6] Majid Mollaeefar, Silvio Ranise. Identifying and quantifying trade-offs in multi-stakeholder risk evaluation with applications to the data protection impact assessment of the GDPR[J]. Computers & Security, 2023, 129:103206.

[7] Richmond Y. Wong, Andrew Chong, and R. Cooper Aspegren. Privacy Legislation as Business Risks: How GDPR and CCPA are Represented in Technology Companies' Investment Risk Disclosures[C]. Proceedings of the ACM on Human-Computer Interaction, 2023.

[8] Alberto Bietti, Chen-Yu Wei, Miroslav Dudik, et al. Personalization

Improves Privacy-Accuracy Tradeoffs in Federated Learning[C]. Proceedings of the 39th International Conference on Machine Learning, 2022:1945-1962.

[9] Philip C. Treleaven, Malgorzata Smietanka, Hirsh Pithadia. Federated Learning: The Pioneering Distributed Machine Learning and Privacy-Preserving Data Technology[J]. Computer, 2022, 55(4): 20-29.

[10] Sheng Shen, Tianqing Zhu, Di Wu, et al. From distributed machine learning to federated learning: In the view of data privacy and security[J]. Concurrency Computation Practice Experience, 2022, 34(16),1:4.

[11] Cholakoska Ana, Pfitzner Bjarne, Gjoreski Hristijan, et al. Differentially Private Federated Learning for Anomaly Detection in EHealth Networks[C]. ACM International Conference on Ubiquitous Computing, 2021.

[12] Sarhad Arisdakessian, Omar Abdel Wahab , Azzam Mourad , et al. A survey on IoT intrusion detection: federated learning, game theory, social psychology, and explainable AI as future directions[J]. IEEE Internet Things, 2023, 10(5): 4059-4092.

[13] Nuria Rodríguez-Barroso, Daniel Jiménez-López, M. Victoria Luzón, et al. Survey on federated learning threats: Concepts, taxonomy on attacks and defences, experimental study and challenges[J]. Information Fusion, 2023, 90: 148-173.

[14] Geiping J , Bauermeister H , Drge H , et al. Inverting Gradients -- How easy is it to break privacy in federated learning?[C]. Conference on Neural Information Processing Systems, 2020.

[15] Eugene Bagdasaryan, Andreas Veit, Yiqing Hua, Deborah Estrin, Vitaly Shmatikov. How To Backdoor Federated Learning[C]. AISTATS, 2020: 2938-2948.

[16] Lu Shi, Jiangang Shu, Weizhe Zhang, et al. HFL-DP: hierarchical federated learning with differential privacy[C]. Global Communications Conference, 2021:1-7.

[17] Rui Hu, Yanmin Gong, Yuanxiong Guo. Federated learning with sparsified model perturbation: Improving accuracy under client-level differential privacy[J]. CoRR, 2022, abs/2202.07178.

[18] Zhifeng Jiang, Wei Wang, Yang Liu. FLASHE: additively symmetric homomorphic encryption for cross-silo federated learning[J]. CoRR, 2021, abs/2109.00675.

[19] Wenhao Mou, Chunlei Fu, Yan Lei, et al. A verifiable federated learning scheme based on secure multi-party computation[C]. Wireless Algorithms, Systems, and Applications, 2021:198-209.

[20] Xinhong Hei, Xinyue Yin, Yichuan Wang, Ju Ren, and Lei Zhu. A trusted feature aggregator federated learning for distributed malicious attack detection. Comput. Secur., 99, 2020.

[21] Zhipin Gu, Yuexiang Yang. Detecting malicious model updates

from federated learning on conditional variational autoencoder[C]. IEEE International Parallel and Distributed Processing Symposium, 2021:671-680.

[22] Xiaoyu Cao, Jinyuan Jia, Neil Zhenqiang Gong. Provably secure federated learning against malicious clients[C]. Association for the Advancement of Artificial Intelligence, 2021:6885-6893.

[23] Zain Abubaker, Nadeem Javaid, Ahmad Almogren, et al. Blockchained service provisioning and malicious node detection via federated learning in scalable internet of sensor things networks[J]. Computer Networks, 2022, 204.

[24] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, et al. SCAFFOLD:stochastic controlled averaging for federated learning[C]. International Conference on Machine Learning, 2020:5132-5143.

[25] Jie Xu, Heqiang Wang. Client Selection And Bandwidth Allocation In Wireless Federated Learning Networks: A Long-Term Perspective[J]. IEEE Transactions on Wireless Communications, 2021, 20(2):1188-1200.

[26] Naoya Yoshida, Takayuki Nishio, Masahiro Morikura, et al. MAB-based Client Selection for Federated Learning with Uncertain Resources in Mobile Networks[C]. IEEE Global Communications Conference, 2020:1-6.

[27] Tiansheng Huang, Weiwei Lin, Wentai Wu, et al. An Efficiency-boosting Client Selection Scheme for Federated Learning with Fairness Guarantee[J]. IEEE Transactions on Parallel and Distributed Systems, 2021, 32(7):1552-1564.

[28] Sawsan Abdulrahman, Hanine Tout, Azzam Mourad, et al. FedMCCS: Multicriteria Client Selection Model for Optimal IoT Federated Learning[J]. IEEE Internet of Things Journal, 2021, 8(6):4723-4735.

[29] Liangkun Yu, Rana Albelaihi, Xiang Sun, et al. Jointly Optimizing Client Selection and Resource Management in Wireless Federated Learning for Internet of Things[J]. IEEE Internet of Things Journal, 2022, 9(6):4385-4395.

[30] Hangjia Zhang; Zhijun Xie; Roozbeh Zarei, et al. Adaptive Client Selection in Resource Constrained Federated Learning Systems: A Deep Reinforcement Learning Approach[J]. IEEE Access, 2021, 9:98423-98432.

[31] Mostafa Majidpour, Charlie Qiu, Peter Chu, et al. Fast prediction for sparse time series: Demand forecast of ev charging stations for cell phone applications[J]. IEEE TII, 2015, 11(1):242-250.

[32] Farshad Rassaei, Wee-Seng Soh, Kee-Chaing Chua. A statistical modelling and analysis of residential electric vehicles' charging demand in smart grids[C]. In Proc. of IEEE ISGT, 2015:1-5.

[33] Feng Tianheng, Yang Lin, Gu Qing, et al. A supervisory control strategy for plug-in hybrid electric vehicles based on energy demand prediction and route preview[J]. IEEE TVT, 2015, 64(5):1691-1700.