2023 设备智能，计算和通信技术国际会议，(DICCT) | 978-1-6654-7491-7/23/$31.00 ©2023 IEEE | DOI: 10.1109 / DICCT56244.2023.10110075

联合学习: 应用、安全隐患和防御措施

Sonam Tyagi
计算图形时代希尔大学学院
Haldwani，印度
sonamtyagi@gehu.ac.in

Ishwari Singh Rajput 计算机
图形学院，*印度 Haldwani*
Hill 大学
ishwarirajput@gehu.ac.in

理查德 Pandey
计算图形时代希尔大学学院
Haldwani，印度
richapandey@gehu.ac.in

***摘要*:联邦学习(FL)是一种前沿的分布式学习方法，它使多个用户能够在保持个人数据隐私的同时共享训练结果。随着数据安全变得越来越重要，从不同的数据所有者那里收集数据以进行机器学习预测变得越来越具有挑战性。联邦学习在克服机器学习和深度学习模型面临的挑战的同时，还可以增加训练数据，保护用户隐私。由于数据隐私和安全是一个世界性的问题，联邦学习的概念从理论到实践的层面日益增加。本文综述了联邦学习框架的概况、类型、不同的应用、几种类型的攻击和防御机制。**

***索引术语*:联邦学习，水平 FL，垂直 FL，联邦 TL，中毒攻击**

我的介绍。

在这个几乎所有数据都通过网络传输的大数据时代，数据安全已成为重中之重。尽管网络数据安全保护技术取得了惊人的进步，但机密性和关键信息泄露问题仍然存在。随着计算能力的提高，机器学习已经成为分析和处理大量数据的首选方法，开辟了新的应用领域[1]。在机器学习取得进步的同时，出现了两个必须解决的问题。数据安全是难以保证的，隐私保护变得越来越重要。然而，在大数据时代，数据共享并不常见;由于共享数据的新趋势以及组织对数据保护的日益重视，以防止数据泄露[2]。不幸的是，由于难以促进用于机器学习训练目的的数据交换，数据孤岛已经出现。Google 引入了联邦学习的概念，其明确目标是打破数据孤岛现象[3]。其设计以信息安全、终端数据隐私和个人数据隐私为基础，优先考虑在众多用户之间有效交换大数据和执行机器学习[4]。除了神经网络，联邦学习还可以利用其他流行的机器学习技术，比如随机森林。除了加强企业与企业之间数据交换的安全性之外这种交流的结果是训练模型的精度，该模型在跨组织共享数据的同时也成功地解决了隐私保护问题[5]。随着联邦学习的发展，该模型的实用性已经扩展到各种新的环境中自 2020 年底以来，互联网上发生了重大数据泄露事件，促使监管机构对大多数大型数据公司进行调查。2020 年 3 月发生了影响 5.38 亿微博用户的数据泄露事件。泄露的数据包括用户 id 和电话号码，用户名和头像等基本账户信息，以及地理位置数据和每个用户的关注者数量。2020 年 12 月 16 日，来自世界各地的 4500 万张医学图片被泄露到网上。大数据公司的数据泄露使整个行业成为国内外关注的焦点。对这些公司的审查表明，各国的立法和监管正在加强，但它也揭示了全球范围内数据泄露和客户隐私泄露的规模。显然，这表明，各企业都应重视数据隐私保护的质量。百度指数显示，在 2019 年 9 月至 2021 年 4 月期间，联邦学习关键字的搜索指数大幅增加，表明越来越多的人开始意识到联邦学习。如何获取数据、如何处理数据以及数据安全管理都是现代数据隐私保护视角下的重要问题[6]。随着联邦学习的兴起，数据的真正潜力终于得以实现[7]。它解决了数据隐私和共享问题，这些问题在很大程度上阻碍了传统机器学习和深度学习[8]。由于没有收集到足够的信息，因此得出的模型是不准确的。这对可用于保护敏感数据的最先进的人工智能算法提出了有用的挑战。

我们研究的主要目标如下:

1) 介绍当前最先进的联邦学习概念的概述。
2) 深入了解联邦学习的分类。
3) 确定联邦学习在不同领域中的应用。

4) 分析联邦学习中的安全隐患及防御措施

本文的组织结构如下。第二节介绍了联邦学习的概念。第一部分将介绍联邦学习的各种类别
III. 第四节将讨论联邦学习在各个领域的应用。第五节介绍了各种安全隐患及其防范措施。最后的讨论载于第六节。

2 什么是联邦学习?

联邦学习的目标是在许多分散的边缘设备上训练机器学习算法，这些设备保存本地数据样本而不交换它们。由于在联邦学习过程中训练数据与成员保存在本地，因此该方法既可以实现每个成员训练数据的共享，又可以保证每个成员隐私的保护[9]。联邦学习的基本过程如图 1 所示，包括以下几个步骤:

1) 参与者从服务器下载全局*模*型 αt-1
2) 参与者 n 训练局部数据得到局部模型 αt,n
3) 参与者将本地模型更新上传到中央服务器
4) 服务器接收到所有参与者的数据后，进行加权聚合操作，得到全局模型 αt
   其中，表述如下:
   *αt,n=*第 n 个参与者的第 th 轮通信的局部模型更新
   *αt=*第 th 轮通信的全局模型更新。

联邦学习技术的特点如下:

1) 联邦学习中使用的数据永远不会以纯文本形式传输并且始终存储在客户端本地，只有模型更新信息需要与中央服务器进行交互。
2) 作为联邦学习的一部分，每个为模型的 α 训练提*供*数据的人都可以访问生成的模型。
3) 最终，联邦学习模型的准确性可与集中的机器学习模型相媲美。
4) 全局模型的准确性随着联邦学习用户使用的训练数据标准的提高而提高。

III. 联邦学习的类型

联邦学习已被证明在隐私敏感的环境中具有巨大的潜力，例如银行业、制造业和其他数据感知领域。根据数据的分布，联邦学习可以分为(图 2):

- 水平联合学习
- 垂直联合学习
- 联邦迁移学习

A. 水平联合学习

数据集的集成是水平联邦学习的基本组成部分。参与者数据和用户数据之间的高度相似性表明这是感兴趣的领域可以用于协作模型训练的数据是数据的子集，其中数据的属性由两个参与者共享[10]，尽管消费者是不同的(图 3)。就数据而言，水平联邦学习应用程序的可能情况范围更大。

B. 垂直联合学习

在垂直联邦学习中，拥有同一组用户的多个属性的不同方可以一起训练机器学习模型，而不需要共享原始数据或模型参数[11]。当用户之间的重叠较多，数据特征之间的重叠较少时，可以用于联合建模训练的数据是双方相同用户在数据属性上下文中不相同的数据子集(图 4)。

C. 联邦迁移学习

在大多数情况下，数据不像水平联邦学习和垂直联邦学习那样占用相同的样本区域或特征空间。因此，这里的主要问题是数据标签的稀缺性和低质量[12]。当参与者的特征和样本之间几乎没有相似之处时，联邦迁移学习是有用的(图 5)。联邦迁移学习的基本概念是不同的参与者具有独特的属性。

*由于联邦学习的诸多*优点，它正在迅速取代传统的机器学习。以下是一些优势:

1) 用户隐私保护: 为了保护用户的隐私，联邦学习中的数据保留在本地，个人用户的数据不被共享[13]。
2) 在海量数据集上训练模型: 通过访问大量的训练数据，可以提高训练模型的质量。与传统的机器学习相比用联合学习训练模型提供了更高的准确性，同时还需要更少的硬件，并允许在大量数据负载下进行更快的训练。

联邦学习模型是如何工作的?

1) 联邦学习允许许多个人远程共享数据，以便共同训练单个深度学习模型并逐步增强它。
2) 每一方都从云数据中心获得模型，这通常是一个已经经过训练的基础模型。

图 1 所示。 联邦学

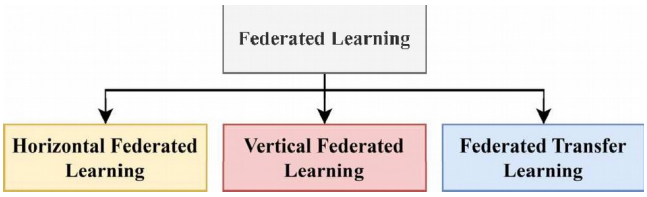

图 2 所示。联邦学习的分类

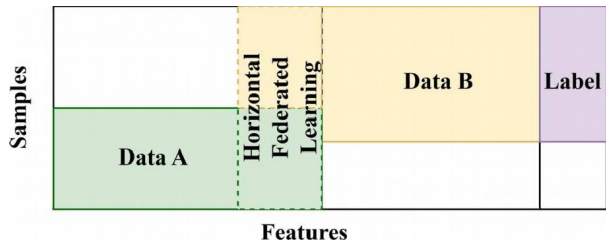

图 3 所示。水平联合学习

3) 一旦他们使用个人数据进行训练，模型的新配置就会被总结和加密。
4) 模型的变化被传送回云，在那里它们被解密、平均，并添加到主模型中。
5) 协同训练进行到模型

经过了充分的训练，迭代又迭代。

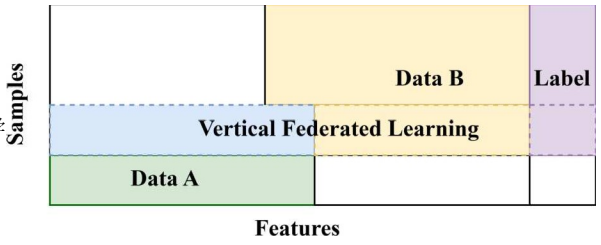

图 4 所示。 垂直联合学习

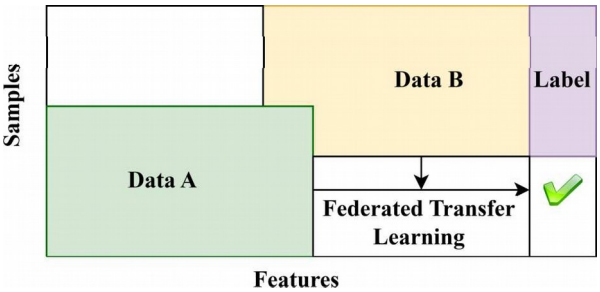

图 5 所示。 联邦迁移学习

IV. 联邦学习的应用

自成立以来，联邦学习已被用于各种用途，尽管其中大多数仍处于实验阶段。公众对联邦学习的认识已经从理论和概念发展到实践[14]。联邦学习在很多领域都非常有用(图 6)。其中一些如下:

A. 入侵检测

与传统的深度学习方法相比，当联合学习与这些训练过程结合使用时，入侵检测的准确性得到了提高。数据隐私得到了保障，避免了数据丢失、被盗等诸多安全问题[15]，[16]。

B. 电力工业

在具有泛在能力的物联网背景下使用联邦学习大大增强了处理能力问题。同时满足了联邦学习的数据分布需求，进一步增强了数据的安全性[17]。

C. 金融行业

数据孤岛是银行业的一个主要问题。所有企业都有兴趣保护客户的个人信息，并从充分利用他们的数据中获得回报。由于监管力度加大，数据安全问题将得到更多关注。如果将联邦学习适当地应用于金融行业，实现普遍适用性那么不同金融机构的数据效果将更加显著[18]。一个优点是，它保护了客户最私人的财务信息的保密性。然而，联邦学习框架的分布式计算提高了计算能力。

D. 通信行业

由于通信行业的特殊性和大量的用户数据，隐私保护已经演变成一个复杂的问题。因此，在通信领域实施联邦学习，不仅可以有效利用运营商持有的大量用户数据，还可以保证所有用户信息的保密性[19]。

E. 医疗保健

由于其作为一种改变游戏规则的方法来保护患者数据的潜力，联邦学习在医疗保健行业越来越受欢迎[20]。每个医疗机构可能都有大量的患者数据，但这可能不足以训练独立的预测模型。联邦学习和疾病预测的整合是减少跨机构分析障碍的良好选择之一[21]。

V. 安全隐患与化解˙联邦学习中的防御措施

联邦学习已经显示出巨大的前景，并在各种环境中产生了令人印象深刻的结果;但是，没有一种模式是没有缺陷的。当联邦学习与实际应用程序集成时，它不可避免地容易受到不道德用户和罪犯的攻击(图 7)。虽然馈入式学习的好处是显而易见的，其发展也符合当前的趋势，但在应用于实践之前，仍需要进行大量的测试以确保其可靠性[22]。在联邦学习中存在以下几个威胁:

A. 中毒攻击

当恶意用户通过攻击训练数据集来操纵机器学习模型的预测时，这被称为中毒攻击。在联邦学习中，中毒攻击可以通过两种方式进行:数据中毒和模型中毒[23]。 术语“数据中毒”指的是攻击者通过引入不正确的标签或有偏差的数据来破坏训练集的场景，从而降低用于训练模型的样本的完整性，并可能损害其可信度或有效性。术语“模型中毒”指的是攻击者在不更改用于创建模型的底层数据的情况下更改模型的场景。在全局聚合过程中，攻击者通过传递错误参数或损坏模型来干扰学习过程[24]。

防御中毒攻击: 由于在联邦学习中，中毒攻击可以通过两种方式进行:数据中毒和模型中毒;因此，对于这两种攻击也可以考虑防御[25]。防止数据中毒最有效的策略是保证模型训练前的数据是安全可信的。其次，如果不确定所训练的数据是否安全可靠，则必须确保模型在训练过程中足够稳定，以保护数据的完整性[26]。由于投毒方以正常用户的身份参与投毒，如果不被发现，其所有行为都将被误认为是正常用户的行为，从而无法检测出投毒者的数据样本。此外，毒药的使用者可以通过自适应地调整模型的数据和参数来优化模型的中毒攻击，以适应自己独特的中毒情况。

B. 对手的攻击

对抗性攻击是故意通过操纵输入数据来欺骗模型给出假阳性结果。对抗性样本是指在原始样本中引入噪声而产生的输入样本[27]。

防御:防御对抗性攻击: 对抗训练是对付对抗性攻击最常用的策略。这意味着将真实数据集和对抗样本结合起来，进行训练，然后对结果模型进行分析和改进。虽然这种类型的训练可以提高

图 6 所示。 联邦学习的应用

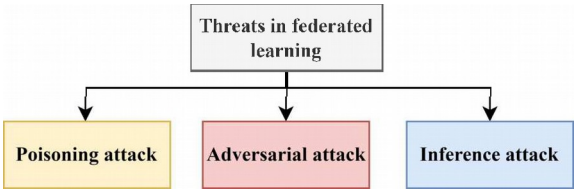

图 7 所示。联邦学习中的威胁

模型的鲁棒性和稳定性，它也有一个明显的缺点，即它容易受到训练的对抗性样本模型的影响，但不受任何未知攻击的影响。数据增强是对抗对抗性攻击的另一种技术[27]。在这种情况下，原始数据被随机改变，以提高模型的泛化能力，这可以用来对抗像图像裁剪、缩放、隐写等攻击。

C. 推理攻击

作为联合学习方法的一部分，个人可以在本地训练他们的数据。每个人都独立工作。参与者对当地信息的间接访问受到限制。隐私和安全可以在一定程度上得到维护。不过，一定程度的安全措施仍然存在，那就是隐私泄露的可能性。例如，恶意用户可能会利用共享参数推断其他用户的隐私信息，然后使用这些信息创建恶意数据，从而破坏模型[28]。前向推理攻击的目标是通过推导模型的参数来利用模型的保密性。反向推理攻击的目标是通过从用于训练模型的训练集中推断出敏感用户数据来获取敏感用户数据。

防御推理攻击: 差分隐私、秘密共享机制、同态加密和混合防御机制都是可以用来实现安全的方法

1) 由于不同的地方包含不同数量的数据，数据不会以一致的方式分散;因此，训练模型的准确性和有效性存在差异。

抵抗推理攻击[29]。差异隐私提供了统计保护，防止对手可能从随机过程的结果中收集到的知识。通过在算法的计算中引入随机性，它建立了单个用户对最终结果影响程度的绝对最大值。秘密共享机制是指将机密信息细分并分发给所有相关方进行安全存储的方法。如果没有收集到一定比例的秘密碎片，则无法获取信息，但收集到这些碎片后可以更改隐藏的信息。另一种流行且安全的加密类型是同态加密。它可以在解密后对密文执行与对明文相同的操作而无需直接访问明文。使用同态加密，服务器不需要知道用户的隐私设置;它只需要对密文进行训练，并在聚合过程中进行聚合。

VI. 讨论

在快速发展的大数据、数据安全和人工智能领域中，使用联邦学习使我们能够利用多个数据源和模型，提高系统效率。本文简要介绍了联邦学习，探讨了潜在的漏洞，并在实际应用程序的上下文中提供了这些安全问题的解决方案。

操作和维护通信网络涉及大量的智能设备，并导致大量的数据输出。直接对内部数据进行训练会导致数据分布不平衡，通用性有限。如果联邦学习在通信部门得到广泛应用，它就有机会解决目前影响该部门的一些问题，如下所述:

2) 由于隐私问题，不能可靠地共享存储的本地数据。它使典型的机器学习训练无法生成高质量的模型，并且使跨学科协作更具挑战性。

联邦学习有可能与更广泛的机器学习算法一起工作，尽管这最终取决于每种算法的通信、加密、编码等细节。

联邦学习的实际应用通常涉及异构客户端设备的分布式网络，例如位于网络外围的移动电话和计算机。测试和调整在移动应用程序(如手机)上运行的机器学习算法更具挑战性和复杂性。现在的问题是如何获得既有效又实用的联邦算法。在这方面所做的工作还需要学者们进行更多的探索，进行深入的分析和研究，使其能够适用于现实场景。本文简单概述了联邦学习、它的类型、威胁和各种应用。

未来，它将被广泛应用于解决通信领域存在的问题。许多人将在一个安全的框架下一起工作和训练，以达到一个共同的目标。

参考文献

[1] W. Y. B. Lim 等人，“移动边缘网络中的联邦学习：《综合调查》，IEEE common。测量员教程，vol. 22, no。3, pp. 2031-2063, july 2020, doi: 10.1109 / COMST.2020.2986024。

[2] M. Khan, F. G. Glavin 和 M. Nickles，“作为隐私解决方案的联邦学习-概述”，《计算机程序》。科学。，第 217 卷，第 316-325 页，2023 年 1 月，doi: 10.1016 / J.PROCS.2022.12.227。

[3] 阮德昌，丁明，P. N. Pathirana, A. Seneviratne，李俊，等 H. Vincent Poor，“物联网的联邦学习：《综合调查》，IEEE 公共版。测量员教程，第 23 卷，第 2 期。3, pp. 1622-1658, july 2021, doi: 10.1109 / COMST.2021.3075439。

[4] M. Abdel-Basset, H. Hawash, N. Moustafa, I. Razzak 和 M. Abd Elfattah，“雾辅助物联网中非 id 数据的隐私保护学习：一种联合学习方法，“数字。Commun。网络，2022 年 12 月，doi: 10.1016 / J.DCAN.2022.12.013。

[5] 顾晓霞，李俊，张涛，任伟，周坤荣，“联邦学习中隐私、准确性和模型公平性的权衡”，计算机学报。安全内核。， vol. 122, p. 102907, Nov. 2022, doi: 10.1016 / J.COSE.2022.102907。

[6] M. Aledhari, R. Razzak, R. M. Parizi 和 F. Saeed，“联邦学习：《使能技术、协议和应用调查》，IEEE Access, vol. 8, pp. 140699-140725, 2020, doi: 10.1109 / AC - CESS.2020.3013541。

[7] 郭勇，刘峰，周涛，蔡志强，肖宁，“眼见为实：迈向联邦学习中数据隐私的交互式可视化探索，" infm。的过程。等内容。，第 60 卷，第 6 期。2, p. 103162, 2023 年 3 月，doi: 10.1016 / J.IPM.2022.103162。

[8] 张涛，李莉，吴明辉，余伟，王晓明，徐振哲，“paggroup：高性能联邦学习的隐私感知分组框架，“J.并行分布。第一版。，第 175 卷，第 37-50 页，2023 年 5 月，doi: 10.1016 / J.JPDC.2022.12.011。

[9] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik 和 Y. Jararweh，“联邦学习回顾：基础、使能技术和未来应用。的过程。等内容。，第 59 卷，第 5 期。6, p. 103061, Nov. 2022, doi: 10.1016 / J.IPM.2022.103061。

[10] 黄伟，李涛，王东，杜树生，张俊，黄涛，“水平联合学习的公平性和准确性”，第 2 期。科学。(纽约)。， vol. 589, pp. 170-185, april . 2022, doi: 10.1016 / J.INS.2021.12.102。

[11] 冯思，“基于垂直联合学习的非重叠样本特征选择”，《专家系统》。达成。， vol. 208, p. 118097, december 2022, doi: 10.1016 / J.ESWA.2022.118097。

[12] 姚志和赵晨，“FedTMI：知识辅助联邦迁移学习用于工业缺失数据的输入，“J.过程控制，vol. 117, pp. 206-215, sept . 2022, doi: 10.1016 / J.JPROCONT.2022.08.004。

[13] 陈建军，薛军，王勇，黄丽丽，周志强，“基于隐私保护和可追溯的联邦学习在工业物联网应用中的应用”，专家系统。达成。，第 213 卷，第 119036 页，2023 年 3 月，doi: 10.1016 / J.ESWA.2022.119036。

[14] 李莉，范勇，谢明，林坤英，“联合学习的应用综述”，计算机科学。印第安纳州,Eng。， vol. 149, p. 106854, Nov. 2020, doi: 10.1016 / J.CIE.2020.106854。

[15] 李斌，吴勇，宋建军，吕仁，李涛，赵磊，“深度人工智能：工业网络物理系统中入侵检测的联合深度学习，" IEEE Trans。《工业信息学》，第 17 卷，第 17 期。8, pp. 5615-5624, Aug. 2021, doi: 10.1109 / TII.2020.3023430。

[16] S. Agrawal 等人，“入侵检测系统的联邦学习：概念、挑战和未来方向。Commun。， vol. 195, pp. 346-361, Nov. 2022, doi: 10.1016 / J.COMCOM.2022.09.012。

[17] W. Zhang et al.，“短期光伏发电预测的半异步个性化联合学习”，《数字》。Commun。网络，2022 年 3 月，doi: 10.1016 / J.DCAN.2022.03.022。

[18] I. Ullah, U. U. Hassan 和 M. I. Ali，“面向工业 4.0 的多级联合学习——一种众包方法”，《计算机学报》。科学。， vol. 217, pp. 423-435, january 2023, doi: 10.1016 / J.PROCS.2022.12.238。

[19] 陈明明，杨志强，杨志强，尹志强，“一种基于无线网络的联合学习和通信框架”，IEEE，Wirel。Commun。，第 20 卷，no。1, pp. 269-283, january 2021, doi: 10.1109 / TWC.2020.3024629。

[20] W. Oh 和 G. N. Nadkarni，“使用结构化医疗数据的医疗保健联邦学习”，广告。肾脏说。愈合。，第 30 卷，第 3 期。1，第 4 - 16 页，2023 年 1 月，doi: 10.1053 / J.AKDH.2022.11.007。

[21] 郝明，李红红，罗晓霞，徐国光，杨红红，刘士生，“基于高效隐私增强的工业人工智能联邦学习”，《计算机工程与应用》(ei)。《工业信息学》，第 16 卷第 1 期。10, pp. 6532-6542, october 2020, doi: 10.1109 / TII.2019.2945367。

[22] A. Yang 等，“联邦学习模型与安全隐患防护应用进展综述”，数字。Commun。网络，2022 年 11 月，doi: 10.1016 / J.DCAN.2022.11.006。

[23] 赵毅，张杰，曹毅，“操纵脆弱性：面向图像分类的联邦云边缘客户端学习中的中毒攻击与对策，“基于知识的系统”。， vol. 259, p. 110072, january 2023, doi: 10.1016 / J.KNOSYS.2022.110072。

[24] 陈志明，田平，廖文文，“针对联邦学习系统的多目标模型投毒攻击”，《高置信度计算》。，第 1 卷，第 1 期。1, p. 100002, 2021 年 6 月，doi: 10.1016 / J.HCC.2021.100002。

[25] 张振华，张勇，郭德华，姚丽丽，李振华，“secfedids：基于联邦学习的网络入侵检测系统中毒攻击的鲁棒防御[j]。总的来说。第一版。系统。， vol. 134, pp. 154-169, september 2022, doi: 10.1016 / J.FUTURE.2022.04.010。

[26] n . rodriguez ´ıguez-Barroso, D. Jime´nez-Lo´pez, M. V. Luzo´n, F. Herrera 和 E. Mart´ınez-Ca´mara，“联邦学习威胁的调查：概念，攻击和防御的分类，实验研究和挑战，”Inf。 《聚变》，第 90 卷，第 148-173 页，2023 年 2 月，doi: 10.1016 / J.INFFUS.2022.09.011。

[27] A. K. Nair, E. D. Raj 和 J. Sahoo，“联邦学习环境中对抗性攻击的稳健分析”，computer。的立场。《接口》，第 86 卷，第 103723 页，2023 年 8 月，doi: 10.1016 / J.CSI.2023.103723。

[28] 顾毅，白毅，徐思，“CS-MIA：联合学习中基于预测置信度序列的隶属推理攻击[j]。安全内核。达成。， vol. 67, p. 103201, june . 2022, doi: 10.1016 / J.JISA.2022.103201。

[29] K. Wei 等人，“差分隐私的联邦学习：算法与性能分析”，IEEE 译。正无穷。取证安全内核。，第 15 卷，第 3454-3469 页，2020,doi: 10.1109 / TIFS.2020.2988575。