纵向联邦线性模型在线推理过程中成员推断攻击的隐私保护研究

尹虹舒，周旭华，周文君
中国电信股份有限公司研究院安全技术研究所，上海 201315

摘要
随着大数据的发展以及数据安全相关法规的出台，人们的隐私保护意识逐渐加强，“数据孤岛”现象愈发严重。联邦学习技术作为解决该问题的有效方法之一，已成为当下备受关注的热点。在纵向联邦学习在线推理过程中，当前的主流方法并未考虑对数据标识的保护。针对此问题，提出一种适用于纵向联邦线性模型在线推理过程中的成员推断攻击的隐私保护方法，通过构造具有假阳率的过滤器来避免对数据标识的精确定位，从而保证数据的安全性；使用同态加密实现在线推理过程的全密态，保护中间计算结果；根据同态加密的密文倍乘性质，使用随机数乘法盲化操作，保证最终推理结果的安全性。该方案进一步提高了纵向联邦学习在线推理过程中用户隐私的安全性，且具有更低的计算开销和通信开销。

关键词
联邦学习；纵向联邦线性模型；在线推理；部分同态加密；数据盲化

中图分类号：TP309.2 文献标志码：A doi: 10.11959/j.issn.2096-0271.2022056

Research on privacy preservation of member inference attacks in online inference process for vertical federated learning linear model

YIN Hongshu, ZHOU Xuhua, ZHOU Wenjun
Security Technology Research Division, China Telecom Research Institute, Shanghai 201315, China

Abstract
With the development of big data and the introduction of data security regulations, the awareness of privacy protection has gradually increased, and the phenomenon of data isolation has become more and more serious. Federated learning technology as one of the effective methods to solve this problem has become a hot spot of concern. In the online inference process of vertical federated learning, the current mainstream methods do not consider the protection of data identity, which is easy to leak user privacy. A privacy protection method for member inference attacks was proposed in the online inference process of the vertical federated linear model. A filter with a false positive rate was constructed to avoid the accurate positioning of data identity to ensure the security of data. Homomorphic encryption was used to realize the full encrypted state of the online inference process and protect the intermediate calculation results. According to the ciphertext multiplication property of homomorphic encryption, the random number multiplication method was

used to mask data, which ensured the security of the final inference result. This scheme further improved the security of user privacy in the online inference process of vertical federated learning and had lower computation overhead and communication costs.

Key words
federated learning, vertical federated learning linear model, online inference, partial homomorphic encryption, data masking

0 引言

随着大数据的迅速发展以及数据安全相关法规的出台，人们对数据安全与隐私保护的意识逐渐加强，企业之间的数据共享变得愈加困难，“数据孤岛”现象愈发严重。联邦学习是目前在保护数据隐私前提下解决“数据孤岛”问题的有效方式[1]。联邦学习能在将各方数据保存在本地的同时进行模型训练，降低了隐私泄露的风险。在实际应用中，联邦学习分为3种，分别是横向联邦学习、纵向联邦学习、联邦迁移学习。其中，纵向联邦学习在数据赋能、数据变现等场景中的应用较为普遍，受到了越来越多的关注，它表现为各方数据集的用户重叠部分较大，用户特征重叠部分较小[2]，例如在金融领域中，银行与电商之间的联合建模能更准确地识别信贷风险[3]。

纵向联邦学习在应用过程中可分为两个阶段：联邦模型训练和联邦在线推理。参与联邦建模的机构（即数据拥有方，后文均称之为参与方）先进行加密样本对齐与加密模型训练，此过程被称为联邦模型训练；在完成模型训练并建立预测模型后，后续的预测由参与方在各自的数据上使用模型参数计算结果，这一预测过程被称为联邦在线推理。

以两个参与方的场景为例，预测发起方（以下简称发起方）开展在线推理时，需要将包含数据标识的请求体发送给另一个参与方（后文均称之为响应方），响应方根据请求体中的数据标识查找己方对应的数据，并使用模型参数计算部分预测结果，然后将该部分预测结果返回给发起方；与此同时，发起方计算己方的部分预测结果，并与响应方的部分预测结果进行合并，从而完成整个在线推理过程。该合并结果即完整的预测结果。

目前联邦模型训练阶段的安全性已被广泛研究，而联邦在线推理阶段的安全性研究相对较少。在上述纵向联邦在线推理过程中，预测请求的请求体内包含的数据标识可能会让响应方直接定位到具体的用户，再结合联合建模的业务特点，响应方很容易推测出业务背后隐含的用户需求，导致用户隐私泄露，如用户是否有贷款需求。

可以发现，纵向联邦在线推理过程中有用户隐私泄露的可能性，针对此问题以及现有研究的不足，本文提出一种面向纵向联邦线性模型在线推理过程中的成员推断攻击的隐私保护方法，对传输过程中的数据标识进行处理，并优化现有方案，进而避免用户隐私泄露。

1 相关工作

在联邦学习技术中，当前的安全性研究主要集中在训练阶段，作为联邦学习的最后一环，在线推理阶段在实际应用场景中的使用频率最高，然而该阶段中的数据安全性问题仍然存在。

1.1 联邦学习用户隐私保护的国内外研究现状

1.1.1 模型训练阶段

目前针对联邦学习的隐私安全问题，诸多研究专注于联邦学习的训练过程。Nasr M等人[4]基于模型训练的参数泄露，利用随机梯度下降的隐私漏洞提出了一种推理攻击算法。罗丹等人[5]提出了一种应用差分隐私技术保护模型训练过程中的参数的方法，通过合理分配隐私预算实现用户隐私保护。对于隐私攻击方式，Barreno M等人[6]提到了规避/探索攻击，此类攻击方式会导致输出一个错误的结果，或者通过收集关于模型特征的信息进行攻击。Bouacida N等人[7]讨论了联邦学习工作流程中涉及的多种隐私泄露情况，包括通信状态中模型被恶意替换、梯度泄露、通过模型参数或训练数据来破坏训练过程、篡改聚合模型更新、聚合算法配置错误等，因此依旧需要相关安全策略来降低隐私泄露的风险。

1.1.2 在线推理阶段

Luo X J等人[8]研究了纵向联邦学习在线推理阶段的隐私泄露问题，提出了基于模型预测的特征推理的攻击方案。针对联邦学习在线推理阶段，Lyu L J等人[9]提到推理阶段的攻击可以分为白盒攻击（可以完全访问联邦学习模型）和黑盒攻击（只能够查询联邦学习模型），其中模型的传输步骤使得任何恶意客户端都可以访问该模型，因此需要采取一些措施来预防白盒攻击。

目前联邦学习的安全性问题涉及的大多是数据训练以及推理过程中的特征推理攻击，而本文研究的问题是在线推理过程中的数据标识泄露问题。这些攻击很大程度上依赖于训练过程中的训练数据样本或破坏模型更新过程、篡改数据特征、泄露交换的模型梯度。为了解决这些隐私安全问题，目前安全多方计算、差分隐私和同态加密等方法均在联邦学习中得到了广泛应用[10-12]。

1.2 纵向联邦在线推理介绍与安全风险

纵向联邦学习是联邦学习中比较常见的一类场景，它适用于几个数据集共享相同的数据样本但特征空间不同的情况[2]。例如同一城市的两家不同的机构，一方为银行，另一方为电商。两方的用户集可能为该地区的大多数居民，因此数据样本重合较多，然而银行与电商的特征空间有很大不同，银行具有用户收支行为和信用评级的特征，电商的特征则为用户的浏览和购买历史。在金融领域中，以数字银行贷款业务为例，一家银行的目标是建立一个机器学习模型，通过联合电商公司的数据来评估是否可以为某些用户提供贷款。只有银行拥有训练数据集和测试数据集中的标签信息，即是否应批准贷款，于是银行与电商两方联合建模共同完成了针对金融风控的模型训练。为了训练纵向联邦学习模型，各参与方以安全的方式迭代交换某些中间结果，直到获得完成联邦训练的模型，最后将训练好的模型发布给各方。获得训练好的模型后，各方再利用该模型协同进行“预测数据集”中新样本的模型预测。

联邦学习的在线建模结束后，各个用户只能得到与自己相关的模型参数，因此纵向联邦学习的在线推理阶段需要所有用户协作完成。两个用户场景下的纵向联邦学习在线推理流程如**图1**所示。

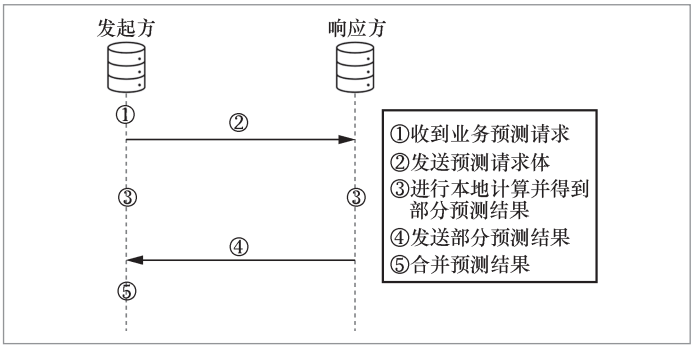


图1 纵向联邦学习在线推理流程

①收到业务预测请求：发起方收到业务预测请求。

②发送预测请求体：发起方将包含数据及模型标识等信息的请求体发送给响应方。

③进行本地计算并得到部分预测结果：发起方和响应方分别基于本地部分模型参数进行预测计算，得到部分预测结果。

④发送部分预测结果：响应方将己方的部分预测结果发送给发起方，便于发起方聚合双方结果。

⑤合并预测结果：发起方聚合双方的部分预测结果，得到最终的预测结果。

由②可以发现，数据标识传输会造成信息泄露的风险，形成成员推断攻击，例如电商（响应方）可以通过银行（发起方）发送的数据标识推断出哪些用户存在贷款需求。

针对上述问题以及现有研究的不足，本文提出一种面向纵向联邦线性模型在线推理过程中的成员推断攻击的隐私保护方法，旨在完善纵向联邦线性模型的在线推理过程，保障用户隐私安全。

2 方法设计与实现

在纵向联邦学习的在线推理阶段，预测发起方将数据标识发送给响应方的过程中，在部分场景下，安全的求交方式使得双方知道共有用户的情况，这可能会间接泄露该信息附带的隐私信息，比如共有用户正在寻求贷款。该数据标识的传输使得响应方能够间接获取用户的隐私，造成信息泄露。

针对此问题以及现有研究的不足，本文提出一种基于具有假阳率的过滤器、支持密文倍乘计算的加法同态加密算法和随机数乘法盲化的面向纵向联邦线性模型在线推理过程中的成员推断攻击的隐私保护方法，对传输过程中的数据标识进行处理，并优化现有方案，进而避免用户隐私泄露。

2.1 符号说明

为了方便理解，对本文用到的符号做如下说明，具体见**表1**。其中，加密使用的是部分同态加密算法，提供加法同态计算，具有密文倍乘性质。

2.2 总体框架

在纵向联邦学习的在线推理过程中，本文方法的总体流程如**图2**所示，发起方收到在线推理请求后，使用数据标识id构造出具有假阳率的过滤器F_{id}（如布隆过滤器（Bloom filter），它是一种用于数据过滤的概率数据结构，可以返回假阳性结果[13-15]），并将该过滤器提供给响应方；响应方使用该过滤器进行全库筛选，获得数据标识集I_{id}，满足$\mathrm{id}\in I_{\mathrm{id}}$且$|I_{\mathrm{id}}|>1$，因此响应方需返回针对多条数据的部分预测结果，而不是仅返回原始id对应的单条数据，从而使发起方无法精确定位到原数据标识对应的用户。在后续过程中，发起方聚合双方的预测结果，并对该结果进行盲化处理，进一步保护用户隐私安全。

2.3 详细流程

在纵向联邦学习在线推理阶段，本文方法的详细流程如**图3**所示。

（1）响应方预先生成同态加密算法的密钥对(PK,SK)，并将公钥PK发送给发起方。

（2）当发起方收到业务系统的在线推理请求后，根据预定规则，针对数据标识id生成具有假阳率的过滤器F_{id}，并发送给响应方。过滤器构造规则根据业务需求的不同而不同，可以选择布隆过滤器、n-前/后缀过滤器等。

（3）响应方和发起方同步进行如下步骤。

响应方根据F_{id}生成满足过滤器规则的数据标识集I_{id}，并查找与之对应的特征数据集$X_{I_{\mathrm{id}}}^{\mathrm{H}}$；针对$X_{I_{\mathrm{id}}}^{\mathrm{H}}$中的每条数据，响应方使用部分模型$w^{\mathrm{H}}$计算得到对应的部分预测结果$y_i^{\mathrm{H}}\left(i \in I_{\mathrm{id}}\right)$后，使用公钥PK对该结果进行加密，得到密文$\left[\!\left[y_{\mathrm{id}}^{\mathrm{H}} \right]\!\right]_{\mathrm{PK}}$，形成部分预测结果集密文$\left[\!\left[Y_{I_{\mathrm{id}}}^{\mathrm{H}} \right]\!\right]_{\mathrm{PK}}$；最后将$\left[\!\left[Y_{I_{\mathrm{id}}}^{\mathrm{H}} \right]\!\right]_{\mathrm{PK}}$以及$I_{\mathrm{id}}$发送给发起方。

发起方查找特征数据$x_{\mathrm{id}}^{\mathrm{G}}$，并使用部分模型$w^{\mathrm{G}}$计算对应的部分预测结果$y_{\mathrm{id}}^{\mathrm{G}}$；使用公钥PK对$y_{\mathrm{id}}^{\mathrm{G}}$进行加密，得到密文$\left[\!\left[y_{\mathrm{id}}^{\mathrm{G}} \right]\!\right]_{\mathrm{PK}}$。

（4）发起方根据I_{id}从集合$\left[\!\left[Y_{I_{\mathrm{id}}}^{\mathrm{H}} \right]\!\right]_{\mathrm{PK}}$中提取出与id对应的响应方的部分预测结果密文$\left[\!\left[y_{\mathrm{id}}^{\mathrm{H}} \right]\!\right]_{\mathrm{PK}}$，并进行聚合计算，得到$\left[\!\left[y_{\mathrm{id}} \right]\!\right]_{\mathrm{PK}} = \left[\!\left[y_{\mathrm{id}}^{\mathrm{H}} \right]\!\right]_{\mathrm{PK}} + \left[\!\left[y_{\mathrm{id}}^{\mathrm{G}} \right]\!\right]_{\mathrm{PK}}$。

（5）发起方选择随机数$\mathrm{mask}_{\mathrm{id}}^{\mathrm{G}}$，并将其与$\left[\!\left[y_{\mathrm{id}} \right]\!\right]_{\mathrm{PK}}$相乘，基于同态加密的密文倍乘性质，得到盲化密文$\left[\!\left[\mathrm{mask}_{\mathrm{id}}^{\mathrm{G}} \cdot y_{\mathrm{id}} \right]\!\right]_{\mathrm{PK}}$，并将盲化密文发送给响应方。

（6）响应方使用私钥SK进行解密，得到结果$\mathrm{mask}_{\mathrm{id}}^{\mathrm{G}} \cdot y_{\mathrm{id}}$，并将该结果发送给发起方。

表1 符号说明

符号	说明
PK,SK	同态加密算法的密钥对，PK为公钥，SK为私钥
G	发起方
H	响应方
w^{G}	发起方的部分模型
w^{H}	响应方的部分模型
id	待预测数据的唯一标识，如身份证号的哈希值等
F_{id}	针对数据标识id创建的具有假阳率的过滤器
I_{id}	满足过滤器规则的数据标识集
$\lvert S \rvert$	集合S中的元素个数
$X_{I_{\mathrm{id}}}^{\mathrm{H}}$	针对过滤器F_{id}的响应方H的特征数据集
$y_{\mathrm{id}}^{\mathrm{H}}$	响应方H的部分预测结果
$Y_{I_{\mathrm{id}}}^{\mathrm{H}}$	针对过滤器F_{id}的响应方H的部分预测结果集
$x_{\mathrm{id}}^{\mathrm{G}}$	发起方G的特征数据
$x_{\mathrm{id}}^{\mathrm{H}}$	响应方H的特征数据
y_{id}^{G}	发起方G的部分预测结果
y_{id}	模型的最终预测结果
$\mathrm{mask}_{\mathrm{id}}^{\mathrm{G}}$	发起方G的用于数据盲化的随机数
$\left[\!\left[R \right]\!\right]_{\mathrm{PK}}$	使用公钥PK对某数值R进行同态加密计算得到的密文

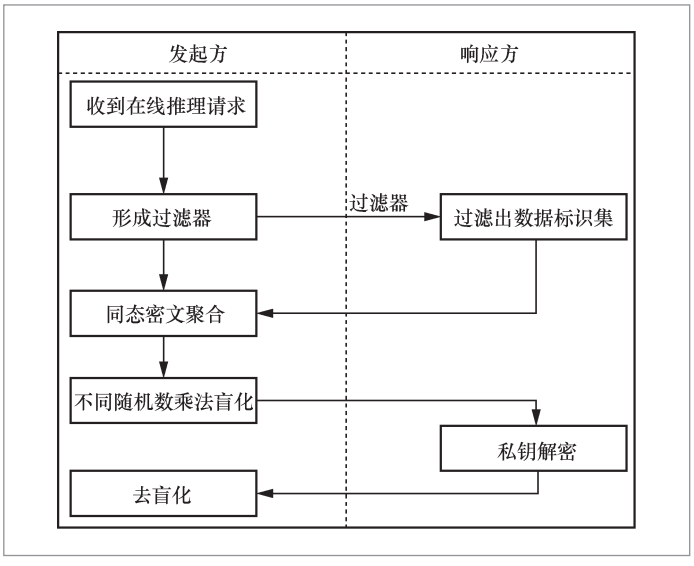


图 2 总体流程

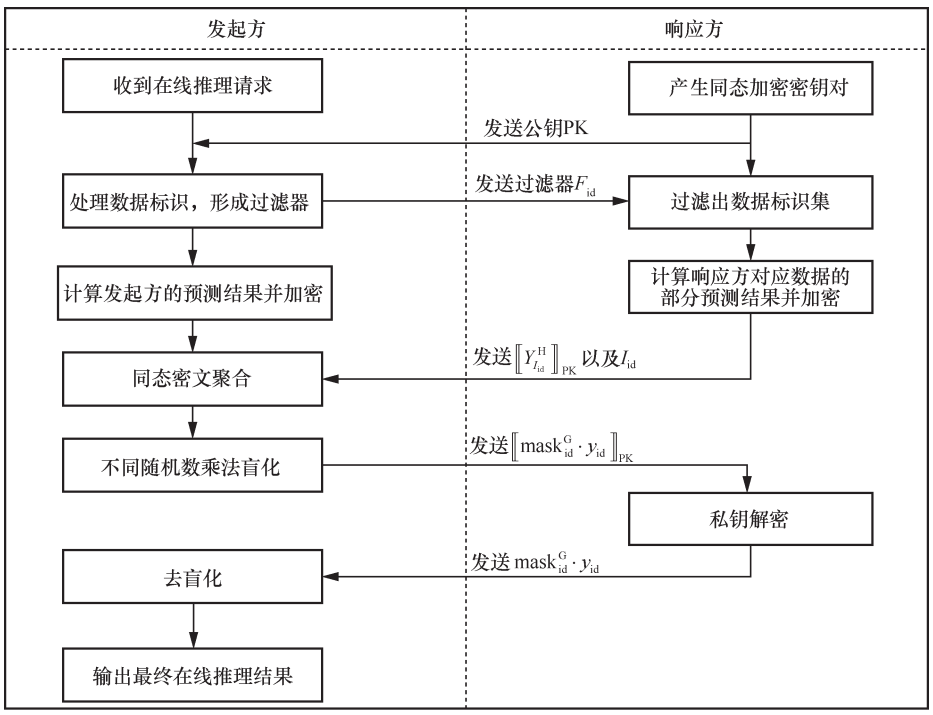


图 3 详细流程

（7）发起方进行去盲化操作，得到最终的预测结果 $y_{id}=\text{mask}_{id}^{G}\cdot y_{id}\cdot\frac{1}{\text{mask}_{id}^{G}}$。

上述步骤为纵向联邦学习在线推理阶段隐私保护方法的详细流程，其中使用了具有假阳率的过滤器，使得响应方无法定位到单条数据，只需计算过滤出的数据标识所对应的数据预测结果，降低了计算开销。在同态密文聚合后，加入不同随机数进行乘法盲化，进一步提高了隐私保护的安全性。

3 性能与安全性分析

本节从性能与安全性两个方面分析本文提出的面向纵向联邦线性模型在线推理过程中的成员推断攻击的隐私保护方法。

3.1 性能分析

为了更好地表示性能分析效果，设发起方发送给响应方的请求体中的数据量为 n，过滤器的假阳率为P。其中下标1表示原方法，下标2表示本文方法。仅考虑比较耗时的操作，不考虑算术运算。

3.1.1 计算量分析

本节的计算量分析针对基于本地部分模型参数的预测计算、过滤器运算、加解密运算、聚合运算以及盲化与去盲化运算。其中，α表示基于本地部分模型参数的预测计算量，β表示过滤器处理运算的计算量，γ表示加解密运算的计算量，δ表示聚合运算的计算量，ε、ϵ 分别表示盲化、去盲化运算的计算量。

在纵向联邦线性模型在线推理原过程中，由第1.2节介绍的流程可知，整个计算过程中的计算量可以表示为：

$$Q_1=\alpha_1+\delta_1 \tag{1}$$

在本文方法中，根据第2.3节的(3)可知，响应方需要使用过滤器F_{id}处理原数据标识集，获得的新的数据标识集大小为$|I_{id}|=nP$。在发起方进行聚合计算后，与原方法相比，本文方法后续增加了盲化与去盲化过程，整个计算过程中的计算量可以表示为：

$$Q_2=\alpha_2+\beta_2+\gamma_2+\delta_2+\varepsilon_2+\epsilon_2=nP\times(\alpha_1+\delta_1)+\beta_2+\gamma_2+\varepsilon_2+\epsilon_2 \tag{2}$$

对比式(1)与式(2)可以发现，与原方法相比，本文方法增加了数据标识处理运算、加解密以及盲化与去盲化的计算量。其中由于过滤器对原数据标识进行了处理，响应方需要处理的数据量增加，因此$Q_2>Q_1$。

由上述分析可知，本文方法的计算量增加了，但是可以使用技术手段降低本文方法带来的计算性能损耗，例如可以使用GPU硬件加速、同态加密打包技术等提高计算效率。

3.1.2 通信量分析

针对纵向联邦线性模型在线推理过程中的通信量，本节分析通信轮次以及与传输数据量相关的通信量。本文使用Paillier同态加密算法，密文长度是1 024 bit，原数据标识长度为m bit。

在纵向联邦线性模型在线推理的原方法中，由第1.2节介绍的流程可知，整个过程包含1轮次数据传输通信。由第2.3节的(5)~(7)可知，相比原方法，本文方法增加了1轮次数据传输通信。原方法的通信量可以表示为：

$$T_1=m \tag{3}$$

在本文方法中，由于使用了过滤器，响应方需要计算nP条数据的预测结果，将该结果加密后传输给发起方的通信量可以表示为：

$$T_2=np\times(1\ 024+m) \tag{4}$$

由式(3)和式(4)可知，本文方法的通信量增加了。在后续盲化过程中，同态加密后的结果与盲化因子相乘后长度不变，对通信量开销无影响。

与原方法相比，本文方法增加了1轮次数据传输通信。常数次通信开销的增加对系统的影响可忽略不计。由过滤器的使用造成的额外通信开销可以使用控制过滤器的假阳率不能过大的方法寻找平衡点；根据密钥的生命周期较短、使用后即可丢弃的特性，可通过降低密钥长度来降低通信开销；可使用同态加密打包技术，减少传输过程中的数据量，从而降低额外的密文通信开销。

3.2 安全性分析

当参与方数量为两个时，本文方法是安全的。由于在计算机网络安全方面，很难做到万无一失，任何破坏与攻击都有可能产生，因此本节做出如下假设：①本文方法应用于受控环境中，受控环境可通过传统的安全加固措施实现；②所有参与方均为半诚实的，即参与方都会执行预设的方案步骤，但也会尝试推断方案之外的信息；③本文仅考虑两个参与方的情形。

3.2.1 数据标识的安全分析

在纵向联邦线性模型在线推理过程中，当预测发起方发起推理请求后，请求体中需要包含的数据标识为发起方用于匹配样本的标识，其可能为设备号或手机号的哈希值等。响应方收到该数据标识后，可以反向推断出该数据标识对应用户的其他业务功能标签，造成隐私泄露。因此本文

方法考虑对该数据标识进行处理，将具有假阳率的过滤器发送给响应方。发起方根据己方设定的规则将该过滤器发送至响应方后，响应方获得被扰动的数据集[13]，该扰动数据集中包含原数据标识，数据范围的扩大使得响应方无法精确定位到原数据标识的对应用户，从而较好地保证发起方的隐私安全。

3.2.2 响应方计算结果的安全分析

在纵向联邦线性模型在线推理过程中，所有计算均在同态加密密文上进行，因此发起方无法获知计算的中间结果明文，其中包括响应方的部分预测结果明文。

3.2.3 最终推理结果的安全分析

发起方使用随机数乘法盲化对聚合结果进行处理，每次选择不同的随机数进行盲化，再将盲化后的预测结果发送给响应方。发起方利用同态加密的密文倍乘性质，实现了对最终推理结果的盲化处理；即使响应方使用私钥对盲化的最终推理结果进行解密，也无法得到去盲化的最终推理结果；只有发起方使用盲化随机数进行去盲化处理后，才能恢复出最终推理结果。由此可以看出，最终推理结果只能由发起方获得，保证了最终推理结果的安全性。

综上分析，在纵向联邦线性模型的在线推理阶段，本文方法的贡献可归纳为如下3点：①通过构造具有假阳率的过滤器来避免响应方对数据标识的精确定位，从而保证数据标识的安全性；②使用同态加密实现在线推理过程的全密态，保护响应方的中间计算结果；③根据同态加密的密文倍乘性质，使用随机数乘法盲化操作，使得最终推理结果只能由发起方获得，保证了最终推理结果的安全性。上述3点保证了纵向联邦线性模型在线推理阶段中参与方交互信息的安全性。

4 结束语

针对纵向联邦线性模型在线推理阶段中的成员推断攻击问题，本文提出了一种面向纵向联邦线性模型在线推理过程中的成员推断攻击的隐私保护方法，确保经过处理后的数据标识的传输不会造成用户隐私泄露，并保证了整个过程交互信息的安全性。在性能方面，与原方法相比，本文方法的计算量与通信量虽有所增加，但可以使用相关技术来降低额外开销；在安全性方面，在本文假设条件下，本文方法能够有效避免用户隐私泄露。

参考文献:

[1] 杨强. AI与数据隐私保护：联邦学习的破解之道[J]. 信息安全研究, 2019, 5(11): 961-965.
YANG Q. AI and data privacy protection: the way to federated learning[J]. Journal of Information Security Research, 2019, 5(11): 961-965.

[2] YANG Q, LIU Y, CHEN T J, et al. Federated machine learning: concept and applications[J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2): 1-19.

[3] 杨强, 童咏昕, 王晏晟, 等. 群体智能中的联邦学习算法综述[J]. 智能科学与技术学报, 2022, 4(1): 29-44.
YANG Q, TONG Y X, WANG Y S, et al. A survey on federated learning in crowd intelligence[J]. Chinese Journal of Intelligent Science and Technology, 2022, 4(1): 29-44.

[4] NASR M, SHOKRI R, HOUMANSADR A. Comprehensive privacy analysis of deep

learning: passive and active white-box inference attacks against centralized and federated learning[C]//Proceedings of 2019 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2019: 739-753.

[5] 罗丹, 徐茹枝, 关志涛. 物联网环境中基于深度学习的差分隐私预算优化方法[J]. 物联网学报, 2022, 6(2): 65-76.
LUO D, XU R Z, GUAN Z T. Differential privacy budget optimization based on deep learning in IoT[J]. Chinese Journal on Internet of Things, 2022, 6(2): 65-76.

[6] BARRENO M, NELSON B, SEARS R, et al. Can machine learning be secure?[C]//Proceedings of 2006 ACM Symposium on Information, Computer and Communications Security. New York: ACM Press, 2006: 16-25.

[7] BOUACIDA N, MOHAPATRA P. Vulnerabilities in federated learning[J]. IEEE Access, 2021, 9: 63229-63249.

[8] LUO X J, WU Y C, XIAO X K, et al. Feature inference attack on model predictions in vertical federated learning[C]//Proceedings of IEEE 37th International Conference on Data Engineering. Piscataway: IEEE Press, 2021: 181-192.

[9] LYU L J, YU H, YANG Q. Threats to federated learning: a survey[J]. arXiv preprint, 2020, arXiv:2003.02133.

[10] 李宗育, 桂小林, 顾迎捷, 等. 同态加密技术及其在云计算隐私保护中的应用[J]. 软件学报, 2018, 29(7): 1830-1851.
LI Z Y, GUI X L, GU Y J, et al. Survey on homomorphic encryption algorithm and its application in the privacy-preserving for cloud computing[J]. Journal of Software, 2018, 29(7): 1830-1851.

[11] 李顺东, 窦家维, 王道顺. 同态加密算法及其在云安全中的应用[J]. 计算机研究与发展, 2015, 52(6): 1378-1388.
LI S D, DOU J W, WANG D S. Survey on homomorphic encryption and its applications to cloud security[J]. Journal of Computer Research and Development, 2015, 52(6): 1378-1388.

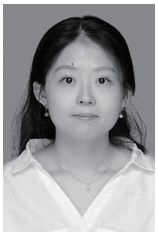
[12] 陈前昕, 毕仁万, 林劼, 等. 支持多数不规则用户的隐私保护联邦学习框架[J]. 网络与信息安全学报, 2022, 8(1): 139-150.
CHEN Q X, BI R W, LIN J, et al. Privacy-preserving federated learning framework with irregular-majority users[J]. Chinese Journal of Network and Information Security, 2022, 8(1): 139-150.

[13] PATGIRI R, NAYAK S, MUPPALANENI N B. Is Bloom filter a bad choice for security and privacy?[C]//Proceedings of 2021 International Conference on Information Networking. Piscataway: IEEE Press, 2021: 648-653.

[14] BRODER A, MITZENMACHER M. Network applications of Bloom filters: a survey[J]. Internet Mathematics, 2004, 1(4): 485-509.

[15] SELVARAJ S, SADASIVAM G S, GOUTHAM D T, et al. Privacy preserving Bloom recommender system[C]//Proceedings of 2021 International Conference on Computer Communication and Informatics. Piscataway: IEEE Press, 2021: 1-6.

作者简介



尹虹舒(1993-),女,中国电信股份有限公司研究院安全技术研究所中级工程师、安全技术研究员,主要研究方向为数据安全、信息安全等。

周旭华（1983– ），男，博士，中国电信股份有限公司研究院安全技术研究所研究员，主要研究方向为隐私保护计算、密码学、数据安全等。

周文君（1980– ），女，中国电信股份有限公司研究院安全技术研究所研究员，主要研究方向为数据安全、系统与应用安全等。

收稿日期： 2022-04-01
通信作者： 周旭华，zhouxh5@chinatelecom.cn
基金项目： 国家重点研发计划资助项目（No.2021YFB3101300）
Foundation Item: The National Key Research and Development Program of China (No.2021YFB3101300)