DOI:10.13718/j.cnki.xsxb.2021.07.011

基于暹罗网络的云计算隐私保护算法①

朱利华[1], 朱玲玲[2]

1. 常州信息职业技术学院 软件与大数据学院, 江苏 常州 213164;
2. 南通大学 信息科学技术学院, 江苏 南通 226200

摘要: 针对云计算难以在准确性和隐私性之间取得可接受的折衷问题, 提出一种基于暹罗网络的云计算隐私保护移动分析算法. 该算法通过暹罗网络对特征提取模块进行微调来选择适用于主要任务, 但不适用于其他辅助任务的专有特征, 采用主成分分析(Principle Component Analysis, PCA)法对专有特征进行降维, 增加其私密性, 降低通信开销, 通过在特征向量里嵌入多维噪声来进一步提高隐私性. 最后, 基于对任意敏感变量统计分析的隐私评估方法评估隐私和验证专有特征提取的质量. 实验结果表明: 本文提出的框架可以在实用性、隐私性和准确性之间实现理想的平衡.

关 键 词: 云计算; 隐私保护; 暹罗网络结构; 深度学习

中图分类号: TP393 **文献标志码:** A **文章编号:** 1000-5471(2021)07-0084-06

近年来随着互联网的发展, 云计算为用户服务增加了一种新的模式, 允许用户随时随地以按次付费的方式访问信息技术服务[1-2]. 智能手机和智能家电等设备大多数收集各种形式的数据并将其传输到云端, 以便从基于云的数据挖掘服务中获益, 如推荐系统、定向广告、安全监控、健康监测和城市规划等[3-4]. 隐私问题是基于云的物联网应用程序构成的重要威胁之一, 用户通过共享敏感数据并允许服务提供商获取、分析或将其数据货币化, 从而存在暴露这些数据的风险[5-6]. 将数据卸载给云服务可能会带来立即或未来潜在的可扩展性和隐私风险, 此外依赖于在用户端执行完整分析的技术也有其自身的资源限制(如存储和带宽限制、能源限制)和用户体验惩罚.

隐私保护是为了防止个人信息和存储在云端的敏感信息暴露[7-8]. 通常采用访问控制、加密、信任技术或上述方法的多种组合进行用户信息和个人隐私保护[9-10]. 云计算的动态计算和存储服务推动了信息技术领域的重大变革, 同时也给安全和隐私带来了巨大的影响, 虚拟化技术和相关的多租户模型可能会导致同一物理设备中的数据丢失. 由于用户与云平台之间缺乏信任, 因此用户不相信数据是在云平台上使用的. 当信息高度集中时, 安全手段必须满足云端处理的要求[11]. 如何实现对数据资源安全和高效的访问控制已成为云计算中的关键问题.

文献[12]提出一种基于边缘云计算的人脸识别隐私保护深度学习算法, 该算法利用差分私有机制实现了基于隐私保护的深度卷积神经网络人脸识别模型的边缘训练, 解决了将深度神经网络(DNN)训练任务外包给不可信服务器时个人的隐私问题. 文献[13]提出一种基于稀疏去噪自编码器的深度神经网络, 用于隐私保护大数据分析. 该算法采用改进的稀疏去噪自动编码器对数据进行变换以提高学习特征的鲁棒性, 使用卷积神经网络(CNN)对变换后的数据进行分类, 提高了深层神经网络对隐私保护数据分类的性能. 文献[14]提出一种基于信息论约束的深度私有特征提取算法, 在适度的资源利用率下, 该算法在保证敏感信

① 收稿日期: 2020-07-14
基金项目: 国家自然科学基金项目(61673384); 全国高等院校计算机基础教育研究会重点课题(GZYZD2018014).
作者简介: 朱利华, 副教授, 主要从事计算机应用研究.

息隐私的前提下可以提高主要任务的精度，但是该算法主要关注删除与单个用户定义的敏感变量相关的信息，最终用户可能无法完全推断出哪些定义为敏感变量.

基于上述文献的思路，本文提出一种基于暹罗网络的云计算隐私保护分析算法. 为了保护数据隐私不受未经授权的任务影响，基于暹罗网络对特征提取模块进行微调以选择适合主任务的专有特征. 为了提高数据的隐私性，对专有特征进行主成分分析(Principal Component Analysis，PCA)降维和多维噪声嵌入. 为了度量专有特征提取器的质量，本文设计了一种评估隐私和验证特征提取模块的新方法. 实验表明，该算法可以通过防止不相关的信息暴露到云端来有效地保护用户隐私.

1　问题定义

假设用户希望通过云服务执行一项主要任务(如语音识别或图像分析)，云服务采用分类模型从用户数据中推断出主要的特征作为主要任务. 将传感器数据发送给基于云的服务提供商进行活动识别(主要任务)，并对其活动做出不同的反应(如在开会时使手机静音). 但是，用户不希望服务提供商能够从数据中推断出其他潜在的敏感信息(如身份或性别)，用户希望以安全、隐私保护的方式使用提供的基于云的预训练服务.

对原始数据进行边缘化预处理可以防止暴露数据不需要的特征，但是由于客户端的各种限制，这样的任务需要有最小的负担. 为了实现这一点，本文提出一种用于云计算隐私保护分析的基于暹罗网络的特征提取模型. 客户端对用户原始数据进行边缘化预处理提取其专有特征，并将该特征而不是原始数据上传到云端进行进一步处理，从而防止用户的敏感信息暴露. 在这种情况下，专有特征应具有以下属性：必须保留尽可能多与主要特征相关的信息以防止主任务性能下降；必须隐藏或丢弃所有其他不必要的信息以防止任何敏感措施的推断. 由于必要信息和敏感信息相互关联，并且特征提取需要在客户端具有最小的开销，因此其主要挑战是如何正确设计特征提取模块，使其在保护敏感信息的同时保留有关主任务的必要信息.

2　基于暹罗网络的云计算隐私保护

为了在保护敏感信息的同时保留有关主任务的必要信息，本文提出了一种基于暹罗网络的特征提取模型，用于云计算隐私保护分析. 云服务向用户的设备提供特征提取程序，设备通过该程序提取数据的专有特征，并将其上传到云中进行其余的过程.

2.1　基于预训练神经网络的分层

深度神经网络(Deep Neural Network，DNN)在机器学习和数据挖掘应用中非常流行，它们提供从原始数据中提取高级信息的高精度分类器. 通过预训练将卷积神经网络分解为一个部署在用户设备上的特征提取模块和一个在云端运行的分类器模块. 在 DNN 模型中，高层主要任务变得越来越具体，同时丢失了包含敏感信息以外的其他无关信息. 选择 DNN 的中间层作为分离点，将 DNN 的输入层到中间层(包括中间层)作为特征提取器存储在边缘，中间层之后的层存储在云中构成分类器模块，将中间层的输出称为中间特性. 利用层分离机制同时实现两个基本目标：获得所需的特征提取器；高层输出对主要任务更具针对性，包含较少的无关信息.

2.2　基于暹罗网络的主任务模型微调

在选择中间层时会有一个折衷，从高层选择它导致敏感信息更高的隐私，但也会增加客户端的计算成本. 解决方案是在向云服务显示中间特性时，使用暹罗网络对现有主要任务的深度模型进行微调，以便所有其他敏感度量变得不可预测. 为了使同一类的特征位于彼此非常小的邻域内，可以更好地保护输入数据的隐私，在应用层分离之前使用暹罗网络对云上预先训练的模型进行微调.

暹罗网络是训练学习模型的一种常用方法，它提供一个特征空间，数据点的相似性由欧氏距离定义，其主要思想是使语义相似点的表示尽可能接近，而不同点的表示则彼此相差较远. 例如，在确定两幅图像是否属于同一个人为目标的人脸验证问题中，暹罗微调能够使属于同一人的任意两幅图像落在特征空间的局部邻域内，并且具有不匹配人脸的任意两幅图像的特征变得彼此远离.

为了微调暹罗网络，训练数据集必须包含标记为相似或不相似的成对样本. 成对样本使用损失函数计

算两个输出的距离. 对于损失函数，两个不同点的距离最大，两个相似点的距离最小. 这种方法使特征提取程序更加私有化，保护用户免受云上的推断攻击，损失函数为

$$L(f_1, f_2) = \begin{cases} \| f_1 - f_2 \|_2^2, & \text{相似} \\ \max(0, margin - \| f_1 - f_2 \|_2^2), & \text{不相似} \end{cases} \tag{1}$$

式(1)中 f_1 和 f_2 是数据点的映射，$margin$ 是控制特征空间方差的超参数.

利用暹罗网络微调来增加特征提取器的隐私性，通过定义一个新的相似性准则将输入数据映射到特征空间，使得具有相同类标签的样本彼此靠近，而具有不同类标签的样本变得遥远，然后使用对比损失函数对模型进行微调. 例如，如果主要任务是性别识别，通过暹罗微调，所有“男性”类的图像将映射到特征空间中的一个小局部区域，所有“女性”图像都将映射到远离男性的另一个区域.

通过定义中间层的对比损失，得到一个多目标优化问题. 它试图通过最小化分类损失来提高主变量预测的准确性，同时通过最小化对比损失来增加排他特征的隐私性. 本文在主分类损失(加权和)中加入对比损失作为正则化项，并尝试用梯度下降算法优化新的损失函数.

本文将这种嵌入方法称为暹罗嵌入. 暹罗微调的整个过程仅由服务提供商在云中完成一次，然后应用层分离并将特征提取模块交付给最终用户. 通过适当地定义相似性准则，可以将这种思想应用到任何程序中. 实验表明，使用暹罗嵌入可以在保持原始任务准确性的同时保留隐私.

2.3 基于主成分分析对中间特征降维

所有基于云服务的一个重要问题是通信成本过高，本文使用主成分分析(Principle Component Analysis，PCA)通过减少中间特征的维数来解决这个问题. 主成分分析(PCA)是一种多变量统计方法，它是最常用的降维方法之一. PCA 采用线性变换，通过矩阵相乘的方法来实现降维和重构. PCA 尽量保留输入信号的主要结构，并删除所有其他不必要的细节，该方法减少了用户和云之间的通信开销.

暹罗微调使特征空间更加健壮，这样在微调空间上应用 PCA 不会显著降低主要任务的精度. 在中间特征空间上进行降维，在不显著降低主要任务精度的情况下，带来了两个优势：①极大地降低了边缘到云的通信成本；②基于降维重构过程的性质，极大地提高了隐私性.

对中间特征应用 PCA 的过程如下：服务提供者分别在特征提取模块的最后一层添加 PCA 映射，在分类模块的第一层之前附加 PCA 重构矩阵. 所提取的中间特征是一个低维向量，可以很容易地传输到云端，且通信成本低. 该方法通过对中间特征进行降维处理，得到排他特征.

2.4 噪声嵌入

除了暹罗微调和降维外，特征提取模块还可以在特征向量中添加多维噪声进一步提高隐私性，本文称这种技术为噪声嵌入. 虽然暹罗微调尝试将具有不同敏感类的数据点映射到单个点，但这些点之间可能仍有很小的距离. 通过在特征中加入随机噪声，可以大大增加敏感变量的不确定性，而敏感变量中的不确定性随着嵌入噪声方差的增加而增加，从而更好地保护隐私. 但是，高方差噪声也会降低主变量的预测精度，因为它会导致数据点脱离正确的类区域. 因此，在增加噪声量时需要在隐私和准确性之间进行权衡. 暹罗微调的一个显著好处是允许嵌入具有更高方差的噪声，因为经过微调后主变量的类内方差减小，而类间方差增加. 因此，通过暹罗微调本文提出的框架可以容忍更高的方差噪声，而主任务的性能不会出现明显下降.

本文将结合降维和加噪的暹罗微调称为暹罗高级嵌入方法. 在将特征提取程序推送到最终用户的设备之前，先在云端执行暹罗微调. 当用户尝试使用服务时，设备上的特征提取模块从中间层输入数据中提取特征向量，通过应用 PCA 或自动编码器对所获得的特征进行降维处理. 在上传到云之前，特征提取器在降维特性中嵌入多维噪声，以获得专有特性. 驻留在云中的分类器模块接收独占特性并执行解压缩操作，最后将重构后的特征输入神经网络，得到预期的结果.

3 专有特征的隐私验证和评估

本节定义了一种基于对任意敏感变量统计分析的隐私评估方法，用于验证与离散敏感变量(如人的身份)相关的排他特性的隐私性. 由于该方法需要在不同上下文中处理隐私问题，本文使用似然等级作为隐

私度量来评估框架的隐私性，考虑似然等级作为隐私度量具有理论研究依据，因为它可以被看作是 k- 匿名性和 top-k 准确性的扩展，也是猜测熵的估计. 事实上，平均等级相当于期望的 k- 匿名性，低等级导致 top-k 准确性降低，从而增加隐私性. 此外，平均等级是熵的经验估计，这是一种众所周知的不确定性度量，也可以用作隐私度量.

设数据集为 $D=\{(x_i, s_i)\}_{i=1:N}$，其中 x_i 为输入数据，s_i 是一个离散敏感类，N 为离散敏感类的数量，可以从集合{1，2，…，K} 获取值. 本文在 D 上应用特征抽取器得到特征集 $F=\{(f_i, s_i)\}_{i=1:N}$，然后通过对 F 加噪建立噪声特征 $Z=\{(z_i, s_i)\}_{i=1:N}$. 为了度量某个噪声特征 z_i 的隐私性，计算所有敏感类的条件似然性$\{P(s \mid z_i) \mid 1 \leqslant s \leqslant K\}$. 因此，首先估计任意敏感类 s 的 $P(z_i \mid s)$.

$$P(z_i \mid s)=\int_f P(z_i, f \mid s)\mathrm{d}f=\int_f P(z_i \mid f)P(f \mid s)\mathrm{d}f \tag{2}$$

如果以 f 为条件，s 变得独立于 z_i，并且有

$$P(z_i \mid s)=\int_f P(z_i \mid f)P(f \mid s)\mathrm{d}f=E_{f\sim P(f\mid s)}[P(z_i \mid f)] \tag{3}$$

设 $F_s=\{f_1, f_2, \cdots, f_{N_s}\}$ 为数据集中敏感类 s 的提取特征集，可以用 F_s 上计算的样本平均值来估算上述期望值

$$\hat{P}(z_i \mid s)=\frac{1}{N_s}\sum_{f_j\in F_s}P(z_i \mid f_j) \tag{4}$$

通过使用贝叶斯规则可以得到

$$\hat{P}(s \mid z_i) \propto \hat{P}(z_i \mid s)P(s) \tag{5}$$

然后，可以计算出所有敏感类在给定噪声特征 z_i 下的相对似然性. 当知道 z_i 的正确敏感类是 s_i 时，在集合$\{P(s \mid z_i) \mid 1 \leqslant s \leqslant N_s\}$ 中找到了类 $P(s_i \mid z_i)$ 的似然秩，按降序排列(可能性越低得到的秩越高)，作为 z_i 的隐私

$$Privacy(z_i)=\frac{Rank(s_i)}{K} \tag{6}$$

本文将秩除以 K(类的总数)，以归一化秩使其的值介于 0 和 1 之间的值. 对于 Z 的所有成员，可以通过平均个人隐私值来估计传输数据的总隐私

$$Privacy_{total}=\frac{1}{N}\sum_{i=1}^{N}Privacy(z_i) \tag{7}$$

深度可视化可以对输入可视化任务(如图像)专有特征的隐私性有一个深入的了解. 为了了解专有特征中敏感信息的数量，本文使用自动编码器可视化技术，该技术在处理图像数据集时特别有用. 该技术在每一层的数据表示上设计了一个译码器，通过译码器来重建原始输入图像. 因此，可以通过将重建图像与原始输入进行比较来分析每层中保留的敏感信息.

使用迁移学习来衡量提取的特征对主要任务的特殊性程度. 假设本文已经为主要变量 N_1 使用 DNN 分类训练构建训练网络 DNNN_1，为辅助变量 N_2 使用 DNN 分类训练构建训练网络，根据以下过程推断任意敏感变量：首先将开始层的权值从 N_1 的前 i 层复制到 N_2 的前 i 层，然后用随机权值初始化 N_2 的剩余层，并冻结 N_2 的前 i 层(不更新其权重). 最后使用迁移训练 N_2 进行敏感变量推理，学习剩余的敏感变量参数. 经过训练后，敏感变量预测精度直接关系到从第 i 层提取的特征的通用程度. 敏感变量预测精度越低，意味着主要任务的特征越具体.

在实践中，最终用户或第三方可以使用这些方法来验证服务提供商特征提取器模块的隐私，以决定是否应该信任服务提供商.

4　实验结果与分析

所有实验均是在配置为 iOS12.2，Hexa-core 64bit CPU，4GB LPDDR4X RAM 和 256 GB 内存的平台上进行. 基于 VGG-16 架构的具有 94%精确度的预训练模型进行实验，在 Conv5-1，Conv5-2 和 Conv5-3 上将其分解为不同的中间层，并在每个中间层上评估该模型. 对中间特征应用 PCA，将 Conv5-3，Conv5-2 和

Conv5-1的维数分别降低到4,6和8. 选取MotionSense数据集[15],该数据集包含加速度计和陀螺仪传感器(姿态、重力、用户加速度和旋转速度)产生的时间序列数据,这些数据是由24名参与者前口袋的iphone7s收集的,在15次试验中进行以下活动:下楼、上楼、散步和慢跑.

为了评估不同嵌入方法如何影响性别分类和活动识别这两个不同主要任务的准确性,表1和表2给出了性别分析和活动识别在不同嵌入方法的精度. 可以看出暹罗嵌入和简单嵌入的预测精度接近,这意味着暹罗微调并不一定会降低主要任务的精确度. 应用降维简单嵌入的预测精度明显降低,应用降维和加噪的暹罗高级嵌入精度变化不大,即暹罗嵌入比简单嵌入更健壮.

表1 性别分类的预测准确性

方 法	Conv5-1	Conv5-2	Conv5-3
简单嵌入	94.0	94.0	94.0
降维后的简单嵌入	89.7	87.0	94.0
暹罗嵌入	92.5	92.7	93.5
暹罗高级嵌入	92.6	92.9	93.8

表2 活动识别的预测准确性

方 法	Conv-4	FC-1	FC-2
简单嵌入	93.0	92.7	93.2
降维后的简单嵌入	85.3	92.5	93.1
暹罗嵌入	92.8	92.6	94.2
暹罗高级嵌入	92.4	92.3	94.2

为了评估性别分类任务的身份隐私性,本文向Alexnet解码器提供了不同嵌入性别分类模型的唯一特征,然后对解码器进行微调以尽可能多地重建输入图像. 与其他算法相比,本文算法可以最大程度地移除个体的身份特征(在保留性别特征的基础上尽可能模糊图像),证明本文提出的高级嵌入算法具有最好的隐私保护性能.

5 结 语

为了在保护敏感信息的同时保留有关主任务的必要信息,本文提出了基于暹罗网络的云计算隐私保护移动分析算法. 该算法将特征提取程序推送到最终用户的设备之前先在云端执行暹罗微调,以选择一个非常适合主任务但不适合任何其他次要任务的专有特性. 当用户使用服务时,设备上的特征提取模块从中间层输入数据中提取特征向量,通过应用PCA特征进行降维处理. 在上传到云之前,在降维特性中嵌入多维噪声以获得专有特性. 驻留在云中的分类器模块接收专有特性并执行解压缩操作,最后将重构后的特征输入神经网络,得到预期的结果. 实验结果表明,该算法在准确性和隐私性之间取得可接受的折衷. 未来的工作是将该算法扩展到递归神经网络,以处理时间序列输入数据,如语音或视频.

参考文献:

[1] SUN P J. Privacy Protection and Data Security in Cloud Computing: a Survey, Challenges, and Solutions [J]. IEEE Access, 2019, 7: 147420-147452.

[2] SUN J F, HU S N, NIE X Y, et al. Efficient Ranked Multi-Keyword Retrieval with Privacy Protection for Multiple Data Owners in Cloud Computing [J]. IEEE Systems Journal, 2020, 14(2): 1728-1739.

[3] OCANSEY S K, AMETEPE W, LI X W, et al. Dynamic Searchable Encryption with Privacy Protection for Cloud Computing [J]. International Journal of Communication Systems, 2018, 31(7): 3403-3411.

[4] ZAHRAH A A, JHANJHI N. Comprehensive Review: Privacy Protection of User in Location-Aware Services of Mobile Cloud Computing [J]. Wireless Personal Communications, 2020, 111(1): 541-564.

[5] SINGH N, SINGH A K. Data Privacy Protection Mechanisms in Cloud [J]. Data Science and Engineering, 2018, 3(1): 24-39.

[6] ALSMADI D, PRYBUTOK V. Sharing and Storage Behavior via Cloud Computing: Security and Privacy in Research and Practice [J]. Computers in Human Behavior, 2018, 85: 218-226.

[7] 江芝蒙，侯翔，李杰. 核子空间投影和广义特征值分解的云数据隐私保护 [J]. 计算机应用与软件，2019，36(4)：268-272，280.

[8] CUI J, ZHANG X Y, ZHONG H, et al. Extensible Conditional Privacy Protection Authentication Scheme for Secure Vehicular Networks in a Multi-Cloud Environment [J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 1654-1667.

[9] WANG T, MEI Y X, JIA W J, et al. Edge-Based Differential Privacy Computing for Sensor-Cloud Systems [J]. Journal of Parallel and Distributed Computing, 2020, 136: 75-85.

[10] WEN Y P, LIU J X, DOU W C, et al. Scheduling Workflows with Privacy Protection Constraints for Big Data Applications on Cloud [J]. Future Generation Computer Systems, 2020, 108: 1084-1091.

[11] DOU Y, CHAN H C B, AU M H. A Distributed Trust Evaluation Protocol with Privacy Protection for Intercloud [J]. IEEE Transactions on Parallel and Distributed Systems, 2019, 30(6): 1208-1221.

[12] Mao Yunlong, Yi Shanhe, Li Qun, et al. A privacy-preserving deep learning approach for face recognition with edge computing [C]//USENIX Workshop Hot Topics Edge Comput. Boston: USENIX, 2018.

[13] ALGULIYEV R M, ALIGULIYEV R M, ABDULLAYEVA F J. Privacy-Preserving Deep Learning Algorithm for Big Personal Data Analysis [J]. Journal of Industrial Information Integration, 2019, 15: 1-14.

[14] OSIA S A, TAHERI A, SHAMSABADI A S, et al. Deep Private-Feature Extraction [J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 32(1): 54-66.

[15] MALEKZADEH M, CLEGG R G, CAVALLARO A, et al. Protecting Sensory Data Against Sensitive Inferences [C]// Proceedings of the 1st Workshop on Privacy by Design in Distributed Systems. New York: ACM, 2018.

Cloud Computing Privacy Protection Algorithm Based on Siamese Network

ZHU Li-hua[1], ZHU Ling-ling[2]

1. *School of Software and Big Data, Changzhou College of Information Technology, Changzhou Jiangsu 213164, China;*
2. *School of Information Science and Technology, Nantong University, Nantong Jiangsu 226200, China*

Abstract: Aiming at the problem that cloud computing is difficult to achieve an acceptable compromise between accuracy and privacy, a mobile analysis algorithm for cloud computing privacy protection based on Siam Network has been proposed. The algorithm uses the Siamese network to fine-tune the feature extraction module to select proprietary features that are suitable for the main task but not suitable for other auxiliary tasks. Principal component analysis (PCA) is used to reduce the dimension of the special features to increase its privacy and reduce the communication overhead. The privacy is further improved by embedding multi-dimensional noise into the feature vector. Finally, a privacy evaluation method based on statistical analysis of any sensitive variable is used to evaluate privacy and verify the quality of proprietary feature extraction. Experimental results show that the proposed framework can achieve an ideal balance between practicality, privacy and accuracy.

Key words: cloud computing; privacy protection; Siamese network structure; deep learning

责任编辑　夏　娟