(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111988327 A
(43)申请公布日 2020.11.24

(21)申请号 202010867322.1

(22)申请日 2020.08.25

(71)申请人 北京天融信网络安全技术有限公司
地址 100000 北京市海淀区上地东路1号院3号楼四层
申请人 北京天融信科技有限公司
北京天融信软件有限公司

(72)发明人 黄娜 李建国 余小军

(74)专利代理机构 北京超凡宏宇专利代理事务所(特殊普通合伙) 11463
代理人 蒋姗

(51)Int.Cl.
H04L 29/06(2006.01)
H04L 12/24(2006.01)
H04L 12/26(2006.01)

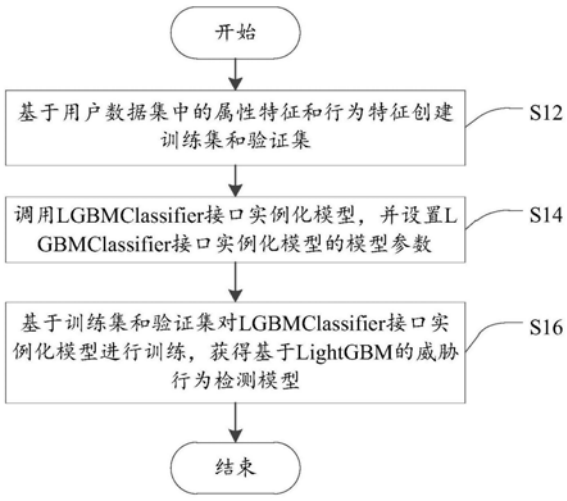
权利要求书2页 说明书9页 附图3页

(54)发明名称

威胁行为检测和模型建立方法、装置、电子设备及存储介质

(57)摘要

本申请提供一种威胁行为检测和模型建立方法、装置、电子设备及存储介质，涉及网络安全技术领域。威胁行为检测模型建立方法包括：基于用户数据集中的属性特征和行为特征创建训练集和验证集；调用LGBMClassifier接口实例化模型，并设置该接口实例化模型的模型参数；基于训练集和验证集对接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型，该检测模型用于基于输入的检测特征输出不合法的概率，检测特征包括待检测用户的属性特征和行为特征。通过威胁行为检测模型进行威胁行为检测不需要对每个用户设置单独检测模型，且LightGBM算法具有并行计算的特性，提高检测的效率，降低了其对计算资源的消耗。



CN 111988327 A

1.一种威胁行为检测模型建立方法,其特征在于,所述方法包括:

基于用户数据集中的属性特征和行为特征创建训练集和验证集,所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种,所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种;

调用LGBMClassifier接口实例化模型,并设置所述LGBMClassifier接口实例化模型的模型参数;

基于所述训练集和所述验证集对所述LGBMClassifier接口实例化模型进行训练,获得基于LightGBM的威胁行为检测模型,所述威胁行为检测模型用于基于输入的检测特征输出所述检测特征的标签为不合法的概率,所述检测特征包括待检测用户的属性特征和行为特征。

2.根据权利要求1所述的方法,其特征在于,所述基于用户数据的属性特征和行为特征创建训练集和验证集,包括:

基于所述用户标识获得所述用户数据集中的所述属性特征和所述行为特征创建训练集和验证集的合并数据;

基于所述用户标识对所述合并数据分别添加标签,获得标签数据,所述标签包括用于表示所述用户标识对应的行为特征合法的标签,以及用于表示所述用户标识对应的行为特征不合法的标签;

对所述标签数据进行数据预处理,获得预处理数据;

将所述预处理数据按照预设比例划分为所述训练集和所述验证集。

3.根据权利要求1所述的方法,其特征在于,所述设置所述LGBMClassifier接口实例化模型的模型参数,包括:

将二分类对数损失函数设置为目标函数;

将所述属性特征中的职能特征设置为类别型特征,所述职能特征包括所述用户标识、所述岗位信息。

4.根据权利要求3所述的方法,其特征在于,所述模型参数包括叶子数、最大深度、叶子节点最小样本数以及学习率和L2正则化系数,所述设置所述LGBMClassifier接口实例化模型的模型参数,还包括:

根据所述训练集的数据量规模,设置所述LGBMClassifier接口实例化模型的所述叶子数、所述最大深度、所述叶子节点最小样本数以及所述学习率和所述L2正则化系数。

5.一种威胁行为检测方法,其特征在于,所述方法包括:

获取检测特征,所述检测特征包括待检测用户的属性特征以及所述待检测用户的行为特征中各项行为的次数,所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种,所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种;

将所述检测特征输入如权利要求1-4中任一项所述的威胁行为检测模型,获得所述威胁行为检测模型的输出结果,所述输出结果表示所述检测特征的标签为不合法的概率;

基于所述输出结果和预设模型阈值的数值对比结果确定所述待检测用户是否具有威胁行为。

6.根据权利要求5所述的方法,其特征在于,在所述基于所述输出结果和预设模型阈值

的数值对比结果确定所述待检测用户是否具有威胁行为之前，所述方法还包括：

在不同的模型阈值下计算验证集对应的所述威胁行为检测模型的精确率和召回率；

将使所述精确率和所述召回率满足预设精确阈值的模型阈值设置为所述预设模型阈值，所述输出结果大于所述预设模型阈值时表示所述待检测用户具有威胁行为，所述输出结果小于或等于所述预设模型阈值时表示所述待检测用户不具有威胁行为。

7.一种威胁行为检测模型建立装置，其特征在于，所述装置包括：

数据集创建模块，用于基于用户数据集中的属性特征和行为特征创建训练集和验证集，所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；

调用模块，用于调用LGBMClassifier接口实例化模型，并设置所述LGBMClassifier接口实例化模型的模型参数；

训练模块，用于基于所述训练集和所述验证集对所述LGBMClassifier接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型，所述威胁行为检测模型用于基于输入的检测特征输出所述检测特征的标签为不合法的概率，所述检测特征包括待检测用户的属性特征和行为特征。

8.一种威胁行为检测装置，其特征在于，所述装置包括：

检测特征获取模块，用于获取检测特征，所述检测特征包括待检测用户的属性特征以及所述待检测用户的行为特征中各项行为的次数，所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；

模型检测模块，用于将所述检测特征输入如权利要求1-4中任一项所述的威胁行为检测模型，获得所述威胁行为检测模型的输出结果，所述输出结果表示所述检测特征的标签为不合法的概率；

威胁判定模块，用于基于所述输出结果和预设模型阈值的数值对比结果确定所述待检测用户是否具有威胁行为。

9.一种电子设备，其特征在于，所述电子设备包括存储器和处理器，所述存储器中存储有程序指令，所述处理器运行所述程序指令时，执行权利要求1-6中任一项所述方法中的步骤。

10.一种存储介质，其特征在于，所述存储介质中存储有计算机程序指令，所述计算机程序指令被一处理器运行时，执行权利要求1-6任一项所述方法中的步骤。

威胁行为检测和模型建立方法、装置、电子设备及存储介质

技术领域

[0001] 本申请涉及网络安全技术领域，具体而言，涉及一种威胁行为检测和模型建立方法、装置、电子设备及存储介质。

背景技术

[0002] 内部威胁是指，内部人员或外部攻击者伪装成的内部人员，利用合法身份及权限，破坏系统或数据、窃取信息、滥用资源等，对企业或组织构成安全危害。随着防火墙等安全防护技术的逐渐强大，内部威胁逐渐成为攻击者的一种常用手段，通过密码窃取、利益诱使等实施内部威胁行为。

[0003] 内部威胁检测通常基于用户在内部网络中的日志数据，相关研究中普遍应用了K-Means、孤立森林、长短期记忆网络、贝叶斯网络等机器学习算法，刻画用户的行为模式，达到检测异常的目的。具有不同角色、职责的用户，往往行为模式也有所不同。但是现有技术采用上述算法进行用户行为模式刻画时，为了区分个体差异，采取对每个用户单独建立检测模型的方法，这种方法会引起模型不统一、数量较多的问题。内部人员的日志数量庞大，会造成内部威胁行为检测效率低、占用计算资源多的问题。

发明内容

[0004] 有鉴于此，本申请实施例的目的在于提供一种威胁行为检测和模型建立方法、装置、电子设备及存储介质，以改善现有技术中存在的内部威胁行为检测效率低、占用计算资源多问题。

[0005] 本申请实施例提供了一种威胁行为检测模型建立方法，所述方法包括：基于用户数据集中的属性特征和行为特征创建训练集和验证集，所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；调用LGBMClassifier接口实例化模型，并设置所述LGBMClassifier接口实例化模型的模型参数；基于所述训练集和所述验证集对所述LGBMClassifier接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型，所述威胁行为检测模型用于基于输入的检测特征输出所述检测特征的标签为不合法的概率，所述检测特征包括待检测用户的属性特征和行为特征。

[0006] 在上述实现方式中，通过结合用户的属性特征及行为特征，使模型根据节点分支自动识别出不同的用户行为模式，避免对每个用户设置单独的检测模型，提高了建模效率，同时用户日志的数量庞大，比较消耗计算资源和时间，利用LightGBM算法并行计算的优势，降低了资源和时间消耗，且能够保证较高的准确性。

[0007] 可选地，所述基于用户数据的属性特征和行为特征创建训练集和验证集，包括：基于所述用户标识获得所述用户数据集中的所述属性特征和所述行为特征创建训练集和验证集的合并数据；基于所述用户标识对所述合并数据分别添加标签，获得标签数据，所述标签包括用于表示所述用户标识对应的行为特征合法的标签，以及用于表示所述用户标识对

应的行为特征不合法的标签;对所述标签数据进行数据预处理,获得预处理数据;将所述预处理数据按照预设比例划分为所述训练集和所述验证集。

[0008] 在上述实现方式中,对用户的合并数据进行行为特征合法与否的标签添加,获得训练集和验证集,能够使根据该训练集和验证集的模型能够判定用户标识对应的用户行为是否合法。

[0009] 可选地,所述设置所述LGBMClassifier接口实例化模型的模型参数,包括:将二分类对数损失函数设置为目标函数;将所述属性特征中的职能特征设置为类别型特征,所述职能特征包括所述用户标识、所述岗位信息。

[0010] 在上述实现方式中,将属性特征中的职能特征设置为类别型特征,使模型能够针对不同职能特征对应的行为特征进行威胁判定,提高了威胁行为判定的准确性。

[0011] 可选地,所述模型参数包括叶子数、最大深度、叶子节点最小样本数以及学习率和L2正则化系数,所述设置所述LGBMClassifier接口实例化模型的模型参数,还包括:根据所述训练集的数据量规模,设置所述LGBMClassifier接口实例化模型的所述叶子数、所述最大深度、所述叶子节点最小样本数以及所述学习率和所述L2正则化系数。

[0012] 在上述实现方式中,基于训练集的规模对LGBMClassifier接口实例化模型进行参数设定,可以在保证训练获得的威胁行为检测模型的准确性,同时提高模型训练与数据规模的匹配性和效率。

[0013] 本申请实施例提供了一种威胁行为检测方法,所述方法包括:获取检测特征,所述检测特征包括待检测用户的属性特征以及所述待检测用户的行为特征中各项行为的次数,所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种,所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种;将所述检测特征输入任一项上述的威胁行为检测模型,获得所述威胁行为检测模型的输出结果,所述输出结果表示所述检测特征的标签为不合法的概率;基于所述输出结果和预设模型阈值的数值对比结果确定所述待检测用户是否具有威胁行为。

[0014] 在上述实现方式中,采用威胁行为检测模型进行内部威胁检测,通过结合用户的属性特征及行为特征,使模型根据节点分支自动识别出不同的用户行为模式,避免对每个用户设置单独的检测模型,提高了建模效率,同时用户日志的数量庞大,比较消耗计算资源和时间,利用LightGBM算法并行计算的优势,降低了资源和时间消耗,且能够保证较高的准确性。

[0015] 可选地,在所述基于所述输出结果和预设模型阈值的数值对比结果确定所述待检测用户是否具有威胁行为之前,所述方法还包括:在不同的模型阈值下计算验证集对应的所述威胁行为检测模型的精确率和召回率;将使所述精确率和所述召回率满足预设精确阈值的模型阈值设置为所述预设模型阈值,所述输出结果大于所述预设模型阈值时表示所述待检测用户具有威胁行为,所述输出结果小于或等于所述预设模型阈值时表示所述待检测用户不具有威胁行为。

[0016] 在上述实现方式中,基于威胁行为检测模型的精确率和召回率作为预设模型阈值设定条件,在威胁行为检测模型的输出结果满足该预设模型阈值时确定用户行为为内部威胁行为,从而提高了内部威胁检测准确率。

[0017] 本申请实施例还提供了一种威胁行为检测模型建立装置,所述装置包括:数据集

创建模块，用于基于用户数据集中的属性特征和行为特征创建训练集和验证集，所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；调用模块，用于调用LGBMClassifier接口实例化模型，并设置所述LGBMClassifier接口实例化模型的模型参数；训练模块，用于基于所述训练集和所述验证集对所述LGBMClassifier接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型，所述威胁行为检测模型用于基于输入的检测特征输出所述检测特征的标签为不合法的概率，所述检测特征包括待检测用户的属性特征和行为特征。

[0018] 在上述实现方式中，通过结合用户的属性特征及行为特征，使模型根据节点分支自动识别出不同的用户行为模式，避免对每个用户设置单独的检测模型，提高了建模效率，同时用户日志的数量庞大，比较消耗计算资源和时间，利用LightGBM算法并行计算的优势，降低了资源和时间消耗，且能够保证较高的准确性。

[0019] 可选地，所述数据集创建模块具体用于：基于所述用户标识获得所述用户数据集中的所述属性特征和所述行为特征创建训练集和验证集的合并数据；基于所述用户标识对所述合并数据分别添加标签，获得标签数据，所述标签包括用于表示所述用户标识对应的行为特征合法的标签，以及用于表示所述用户标识对应的行为特征不合法的标签；对所述标签数据进行数据预处理，获得预处理数据；将所述预处理数据按照预设比例划分为所述训练集和所述验证集。

[0020] 在上述实现方式中，对用户的合并数据进行行为特征合法与否的标签添加，获得训练集和验证集，能够使根据该训练集和验证集的模型能够判定用户标识对应的用户行为是否合法。

[0021] 可选地，所述调用模块具体用于：将二分类对数损失函数设置为目标函数；将所述属性特征中的职能特征设置为类别型特征，所述职能特征包括所述用户标识、所述岗位信息。

[0022] 在上述实现方式中，将属性特征中的职能特征设置为类别型特征，使模型能够针对不同职能特征对应的行为特征进行威胁判定，提高了威胁行为判定的准确性。

[0023] 可选地，所述模型参数包括叶子数、最大深度、叶子节点最小样本数以及学习率和L2正则化系数，所述调用模块具体用于：根据所述训练集的数据量规模，设置所述LGBMClassifier接口实例化模型的所述叶子数、所述最大深度、所述叶子节点最小样本数以及所述学习率和所述L2正则化系数。

[0024] 在上述实现方式中，基于训练集的规模对LGBMClassifier接口实例化模型进行参数设定，可以在保证训练获得的威胁行为检测模型的准确性，同时提高模型训练与数据规模的匹配性和效率。

[0025] 本申请实施例还提供了一种威胁行为检测装置，所述装置包括：检测特征获取模块，用于获取检测特征，所述检测特征包括待检测用户的属性特征以及所述待检测用户的行为特征中各项行为的次数，所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；模型检测模块，用于将所述检测特征输入如任一项上述的威胁行为检测模型，获得所述威胁行为检测模型的输出结果，所述输出结果表示所述检测特征的标签为不

合法的概率;威胁判定模块,用于基于所述输出结果和预设模型阈值的数值对比结果确定所述待检测用户是否具有威胁行为。

[0026] 在上述实现方式中,采用威胁行为检测模型进行内部威胁检测,通过结合用户的属性特征及行为特征,使模型根据节点分支自动识别出不同的用户行为模式,避免对每个用户设置单独的检测模型,提高了建模效率,同时用户日志的数量庞大,比较消耗计算资源和时间,利用LightGBM算法并行计算的优势,降低了资源和时间消耗,且能够保证较高的准确性。

[0027] 可选地,所述威胁判定模块具体用于:在不同的模型阈值下计算验证集对应的所述威胁行为检测模型的精确率和召回率;将使所述精确率和所述召回率满足预设精确阈值的模型阈值设置为所述预设模型阈值,所述输出结果大于所述预设模型阈值时表示所述待检测用户具有威胁行为,所述输出结果小于或等于所述预设模型阈值时表示所述待检测用户不具有威胁行为。

[0028] 在上述实现方式中,基于威胁行为检测模型的精确率和召回率作为预设模型阈值设定条件,在威胁行为检测模型的输出结果满足该预设模型阈值时确定用户行为为内部威胁行为,从而提高了内部威胁检测准确率。

[0029] 本申请实施例还提供了一种电子设备,所述电子设备包括存储器和处理器,所述存储器中存储有程序指令,所述处理器读取并运行所述程序指令时,执行上述任一实现方式中的步骤。

[0030] 本申请实施例还提供了一种可读取存储介质,所述可读取存储介质中存储有计算机程序指令,所述计算机程序指令被一处理器读取并运行时,执行上述任一实现方式中的步骤。

附图说明

[0031] 为了更清楚地说明本申请实施例的技术方案,下面将对本申请实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,因此不应被看作是对范围的限定,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他相关的附图。

[0032] 图1为本申请实施例提供的一种威胁行为检测模型建立方法的流程示意图。

[0033] 图2为本申请实施例提供的一种训练集和验证集创建步骤的流程示意图。

[0034] 图3为本申请实施例提供的一种威胁行为检测方法的流程示意图。

[0035] 图4为本申请实施例提供的一种威胁行为检测模型建立装置的模块示意图。

[0036] 图5为本申请实施例提供的一种威胁行为检测装置的模块示意图。

[0037] 图标:30-威胁行为检测模型建立装置;31-数据集创建模块;32-调用模块;33-训练模块;40-威胁行为检测装置;41-检测特征获取模块;42-模型检测模块;43-威胁判定模块。

具体实施方式

[0038] 下面将结合本申请实施例中附图,对本申请实施例中的技术方案进行描述。

[0039] 现有的网络安全防护技术中,对内部威胁进行检测的方式包括:方式一,收集用户

属性信息，包括姓名、年龄、性别、心理评定、人际交往情况、工作完成质量、工作满意度等；对数据进行清洗和预处理后，通过用户画像计算用户之间的相似度；采用K-Means算法对用户属性画像聚类，得到行为模式相近的用户群；方式二，获取用户行为信息和标识信息；根据用户的行为样本集，训练得到长短期记忆网络分类模型；根据用户标识信息判断是否正确分类；方式三，首先检测用户的多域行为，然后利用熵权法融合多域行为的检测结果。

[0040] 上述方式一、方式二和方式三通常对每个用户设置单独的检测模型，且方式一使用了用户属性特征，没有使用行为特征，仅检测动机较高的危险用户，没有直接检测内部威胁行为。方式二将用户行为信息作为训练集特征，用户标识信息作为判断标准，而没有将两类特征结合作为训练数据。方式三仅使用用户行为信息，采用决策级融合检测多域行为，容易忽略数据中的原始信息。

[0041] 因此现有技术中对内部威胁的检测存在建模及检测效率较低，占用计算资源较多，且准确率不足的问题。

[0042] 为了解决上述问题，本申请实施例提供了一种威胁行为检测模型建立方法，请参考图1，图1为本申请实施例提供的一种威胁行为检测模型建立方法的流程示意图，该威胁行为检测模型建立方法的具体步骤可以如下：

[0043] 步骤S12：基于用户数据集中的属性特征和行为特征创建训练集和验证集。

[0044] 本实施例中的用户数据集为收集的内部员工用户的身份标识和历史日志数据，身份标识包括用户标识、用户属性、岗位信息和工作满意度，岗位信息包括部门和职务等，用户属性包括姓名、年龄、性别、心理评定、人际交往情况、工作完成质量、工作满意度等，将身份标识作为属性特征，历史日志数据包括系统访问(登陆和注销)、文件访问(读取、写入、复制和删除)、外设连接(连接和断开)、网络访问(访问、上传和下载)、邮件收发(接收和发送)等用户行为，统计用户在一天时间内做出每一类用户行为的历史数量作为行为特征。

[0045] 具体地，请参考图2，图2为本申请实施例提供的一种训练集和验证集创建步骤的流程示意图，如图所示步骤S12具体可以包括如下子步骤：

[0046] 步骤S122：基于用户标识获得用户数据集中的属性特征和行为特征创建训练集和验证集的合并数据。

[0047] 将属性特征用X表示，行为特征用Y表示，并将属性特征X和行为特征Y按照用户标识对应合并获得合并数据。

[0048] 其中，用户标识可以是用于表示用户的唯一身份标识，其可以为数字、英文、汉字及其组合等任意字符串。

[0049] 步骤S124：基于用户标识对合并数据分别添加标签，获得标签数据，标签包括用于表示用户标识对应的行为特征合法的标签，以及用于表示用户标识对应的行为特征不合法的标签。

[0050] 可选地，本实施例中的标签可以用l_i表示，不同取值的i对应不同的用户标识，其取值为0或1，0表示合法，1表示不合法。

[0051] 应当理解的是，在其他实施例中，标签可以有其他适用于行为特征合法性的任意表示方式，合法与不合法的取值也可以进行灵活选取。

[0052] 步骤S126：对标签数据进行数据预处理，获得预处理数据。

[0053] 可选地，本实施例中的数据预处理可以包括清洗和空值填充。

[0054] 清洗是剔除错误的数据或具有特殊性的数据，如去除某个日期的行为统计。

[0055] 空值填充是因为属性特征中可能存在信息不全，行为统计时也会存在员工没有做出的行为类型，将空值用统一的数值(例如0)填充。

[0056] 此外，步骤S124的标签标记步骤也可以视为数据预处理的一部分，因此步骤S124和步骤S126之间的顺序关系不受限定。

[0057] 步骤S128：将预处理数据按照预设比例划分为训练集和验证集。

[0058] 可选地，本实施例中的训练集和验证集的划分比例可以根据模型训练的具体需求进行调整，例如验证集的比例为5%～20%。

[0059] 应当理解的是，划分训练集和验证集的步骤也可以在步骤S122～步骤S126的任意步骤之前或之后进行，因此步骤S122的顺序关系不受限定。

[0060] 步骤S14：调用LGBMClassifier接口实例化模型，并设置LGBMClassifier接口实例化模型的模型参数。

[0061] 上述LGBMClassifier接口实例化模型可以是机器学习库Sklearn中所提供，LGBM为LightGBM(Light Gradient Boosting Machine)，其为实现GBDT(Gradient Boosting Decision Tree)算法的框架，支持高效率的并行训练。GBDT是机器学习中一个长盛不衰的模型，其主要思想是利用弱分类器(决策树)迭代训练以得到最优模型，该模型具有训练效果好、不易过拟合等优点。GBDT在工业界应用广泛，通常被用于点击率预测，搜索排序等任务。LGBM提出的主要原因是为了解决GBDT在海量数据遇到的问题，让GBDT可以更好更快地用于工业实践。LGBM具有以下优点：更快的训练速度、更低的内存消耗、更好的准确率和分布式支持，可以快速处理海量数据。

[0062] 可选地，本实施例中可以将二分类对数损失函数设置为基于LightGBM的威胁行为检测模型的目标函数，其具体计算公式如下：

[0063] $$\mathrm{Loss} = -\frac{1}{N}\sum_{i=1}^{N}(l_i \log P_{l_i} + (1-l_i)\log(1-P_{l_i}))$$

[0064] 其中，N表示用户数量，P_{l_i}表示模型预测用户标签为l_i概率，Loss为模型的损失值。

[0065] 进一步地，将属性特征X中的职能特征(包括用户标识和岗位信息)设置为类别型特征，使模型能够区分具有不同职能属性的用户。

[0066] 此外，在进行模型训练前还需要根据训练集的数据量规模，设置LGBMClassifier接口实例化模型的叶子数(num_leaves)、最大深度(max_depth)、叶子节点最小样本数(min_data_in_leaf)以及学习率(feature_fraction)和L2正则化系数。例如，在训练集的数据量规模为百万级时，可以设为num_leaves＝2048，max_depth＝15，min_data_in_leaf＝50，feature_fraction＝0.7，L2正则化系数＝0.1。

[0067] 步骤S16：基于训练集和验证集对LGBMClassifier接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型。

[0068] 经过训练获得的基于LightGBM的威胁行为检测模型的输出结果为$P(l_x=1)$，其表示特征x标签为1(即不合法)的概率。

[0069] 在完成威胁行为检测模型的建立后，可以使用该威胁行为检测模型对内部用户的行为特征进行检验，判定其是否属于内部威胁行为，因此本实施例提供了一种威胁行为检

测方法，请参考图3，图3为本申请实施例提供的一种威胁行为检测方法的流程示意图，该威胁行为检测方法的具体步骤可以如下：

[0070] 步骤S22：获取检测特征。

[0071] 检测特征包括待检测用户的属性特征以及待检测用户的行为特征中各项行为的次数，属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种。

[0072] 步骤S24：将检测特征输入威胁行为检测模型，获得威胁行为检测模型的输出结果。

[0073] 输出结果表示检测特征的标签为不合法的概率，即$P(l_x=1)$，其表示特征x标签为1(即不合法)的概率。

[0074] 步骤S26：基于输出结果和预设模型阈值的数值对比结果确定待检测用户是否具有威胁行为。

[0075] 在不同阈值下，计算模型测试验证集的精确率(Precision)和召回率(Recall)，当精确率和召回率能够同时满足需求时，将相应的阈值设置为预设模型阈值T，若$P(l_x=1)>T$，判定该检测数据不合法，即该用户做出了内部威胁行为；反之，判定该检测数据为合法。

[0076] 为了配合上述威胁行为检测模型建立方法，本申请实施例还提供了一种威胁行为检测模型建立装置30，请参考图4，图4为本申请实施例提供的一种威胁行为检测模型建立装置的模块示意图。

[0077] 威胁行为检测模型建立装置30包括：

[0078] 数据集创建模块31，用于基于用户数据集中的属性特征和行为特征创建训练集和验证集，属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；

[0079] 调用模块32，用于调用LGBMClassifier接口实例化模型，并设置LGBMClassifier接口实例化模型的模型参数；

[0080] 训练模块33，用于基于训练集和验证集对LGBMClassifier接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型，威胁行为检测模型用于基于输入的检测特征输出检测特征的标签为不合法的概率，检测特征包括待检测用户的属性特征和行为特征。

[0081] 可选地，数据集创建模块31具体用于：基于用户标识获得用户数据集中的属性特征和行为特征创建训练集和验证集的合并数据；基于用户标识对合并数据分别添加标签，获得标签数据，标签包括用于表示用户标识对应的行为特征合法的标签，以及用于表示用户标识对应的行为特征不合法的标签；对标签数据进行数据预处理，获得预处理数据；将预处理数据按照预设比例划分为训练集和验证集。

[0082] 可选地，调用模块32具体用于：将二分类对数损失函数设置为目标函数；将属性特征中的职能特征设置为类别型特征，职能特征包括用户标识、岗位信息。

[0083] 可选地，模型参数包括叶子数、最大深度、叶子节点最小样本数以及学习率和L2正则化系数，调用模块32具体用于：根据训练集的数据量规模，设置LGBMClassifier接口实例化模型的叶子数、最大深度、叶子节点最小样本数以及学习率和L2正则化系数。

[0084] 为了配合上述威胁行为检测方法，本申请实施例还提供了一种威胁行为检测装置40，请参考图5，图5为本申请实施例提供的一种威胁行为检测装置的模块示意图。

[0085] 威胁行为检测装置40包括：

[0086] 检测特征获取模块41，用于获取检测特征，检测特征包括待检测用户的属性特征以及待检测用户的行为特征中各项行为的次数，属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；

[0087] 模型检测模块42，用于将检测特征输入威胁行为检测模型，获得威胁行为检测模型的输出结果，输出结果表示检测特征的标签为不合法的概率；

[0088] 威胁判定模块43，用于基于输出结果和预设模型阈值的数值对比结果确定待检测用户是否具有威胁行为。

[0089] 可选地，威胁判定模块43具体用于：在不同的模型阈值下计算验证集对应的威胁行为检测模型的精确率和召回率；将使精确率和召回率满足预设精确阈值的模型阈值设置为预设模型阈值，输出结果大于预设模型阈值时表示待检测用户具有威胁行为，输出结果小于或等于预设模型阈值时表示待检测用户不具有威胁行为。

[0090] 本申请实施例还提供了一种电子设备，该电子设备包括存储器和处理器，所述存储器中存储有程序指令，所述处理器读取并运行所述程序指令时，执行本实施例提供的威胁行为检测模型建立方法或威胁行为检测方法中任一项所述方法中的步骤。

[0091] 应当理解是，该电子设备可以是个人电脑(Personal Computer，PC)、平板电脑、智能手机、个人数字助理(Personal Digital Assistant，PDA)等具有逻辑计算功能的电子设备。

[0092] 本申请实施例还提供了一种可读取存储介质，所述可读取存储介质中存储有计算机程序指令，所述计算机程序指令被一处理器读取并运行时，执行威胁行为检测模型建立方法或威胁行为检测方法中的步骤。

[0093] 综上所述，本申请实施例提供了一种威胁行为检测和模型建立方法、装置、电子设备及存储介质，所述方法包括：基于用户数据集中的属性特征和行为特征创建训练集和验证集，所述属性特征包括用户标识、用户属性、岗位信息和工作满意度中的至少一种，所述行为特征包括系统访问、文件访问、外设连接、网络访问和邮件收发中的至少一种；调用LGBMClassifier接口实例化模型，并设置所述LGBMClassifier接口实例化模型的模型参数；基于所述训练集和所述验证集对所述LGBMClassifier接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型，所述威胁行为检测模型用于基于输入的检测特征输出所述检测特征的标签为不合法的概率，所述检测特征包括待检测用户的属性特征和行为特征。

[0094] 在上述实现方式中，通过结合用户的属性特征及行为特征，使模型根据节点分支自动识别出不同的用户行为模式，避免对每个用户设置单独的检测模型，提高了建模效率，同时用户日志的数量庞大，比较消耗计算资源和时间，利用LightGBM算法并行计算的优势，降低了资源和时间消耗，且能够保证较高的准确性。

[0095] 在本申请所提供的几个实施例中，应该理解到，所揭露的设备，也可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的，例如，附图中的框图显示了根据本申请的多个实施例的设备的可能实现的体系架构、功能和操作。在这点上，框图中的每个方框可以代表一个模块、程序段或代码的一部分，所述模块、程序段或代码的一部分包含一个或

多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现方式中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图中的每个方框、以及框图的组合，可以用执行规定的功能或动作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

[0096] 另外，在本申请各个实施例中的各功能模块可以集成在一起形成一个独立的部分，也可以是各个模块单独存在，也可以两个或两个以上模块集成形成一个独立的部分。

[0097] 所述功能如果以软件功能模块的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。因此本实施例还提供了一种可读取存储介质中存储有计算机程序指令，所述计算机程序指令被一处理器读取并运行时，执行区块数据存储方法中任一项所述方法中的步骤。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备(可以是个人计算机，服务器，或者网络设备等)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器(ROM，Read-Only Memory)、随机存取存储器(RAM，RanDom Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0098] 以上所述仅为本申请的实施例而已，并不用于限制本申请的保护范围，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步定义和解释。

[0099] 以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。

[0100] 需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

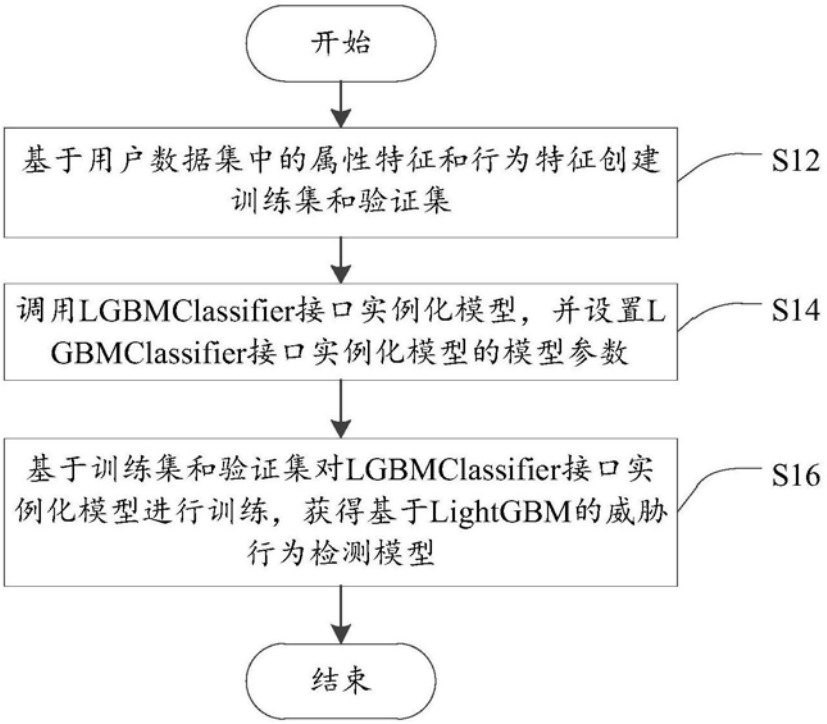

开始
基于用户数据集中的属性特征和行为特征创建训练集和验证集
S12
调用LGBMClassifier接口实例化模型，并设置LGBMClassifier接口实例化模型的模型参数
S14
基于训练集和验证集对LGBMClassifier接口实例化模型进行训练，获得基于LightGBM的威胁行为检测模型
S16
结束

图1

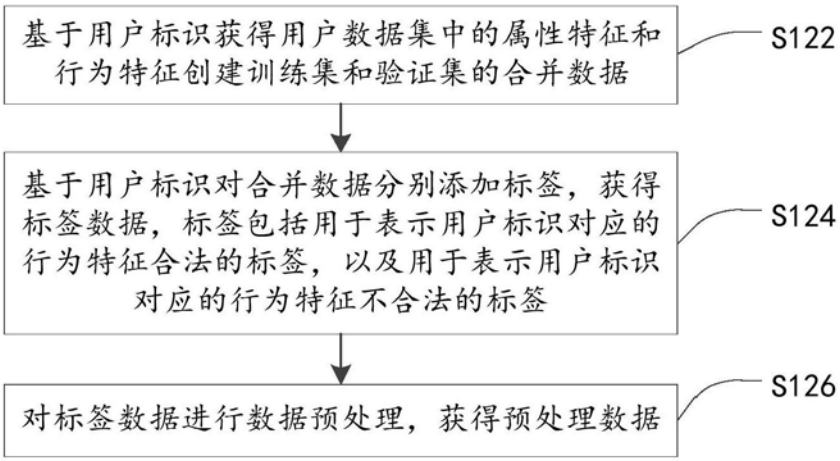

基于用户标识获得用户数据集中的属性特征和行为特征创建训练集和验证集的合并数据
S122
基于用户标识对合并数据分别添加标签，获得标签数据，标签包括用于表示用户标识对应的行为特征合法的标签，以及用于表示用户标识对应的行为特征不合法的标签
S124
对标签数据进行数据预处理，获得预处理数据
S126

图2

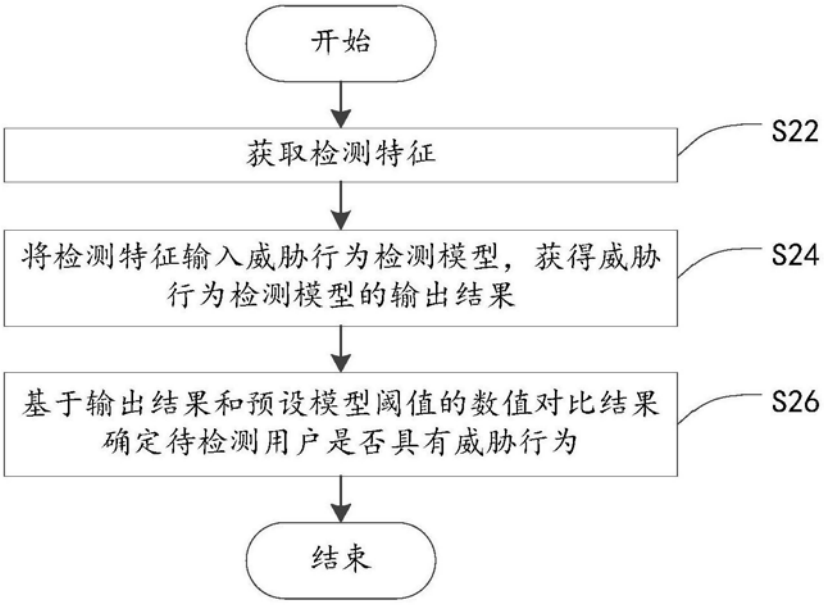

开始
获取检测特征
S22
将检测特征输入威胁行为检测模型，获得威胁行为检测模型的输出结果
S24
基于输出结果和预设模型阈值的数值对比结果确定待检测用户是否具有威胁行为
S26
结束

图3

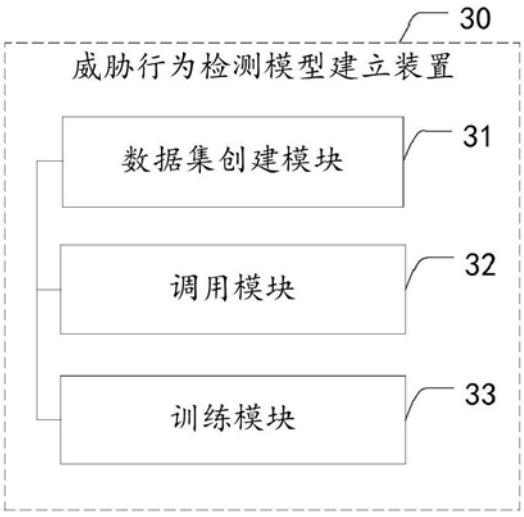

30
威胁行为检测模型建立装置
数据集创建模块
31
调用模块
32
训练模块
33

图4

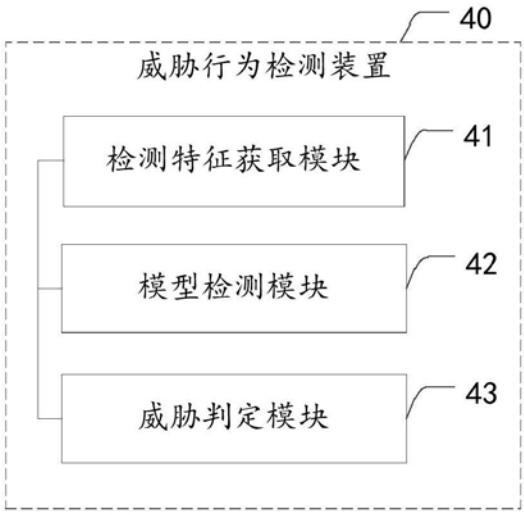


图5