(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 111400357 A
(43)申请公布日 2020.07.10

(21)申请号 202010107808.5

(22)申请日 2020.02.21

(71)申请人 中国建设银行股份有限公司
地址 100033 北京市西城区金融大街25号
申请人 建信金融科技有限责任公司

(72)发明人 黄鸿铿 黄建德

(74)专利代理机构 中原信达知识产权代理有限责任公司 11219
代理人 张一军 王安娜

(51)Int.Cl.
G06F 16/2457(2019.01)
H04L 29/06(2006.01)
G06N 20/00(2019.01)

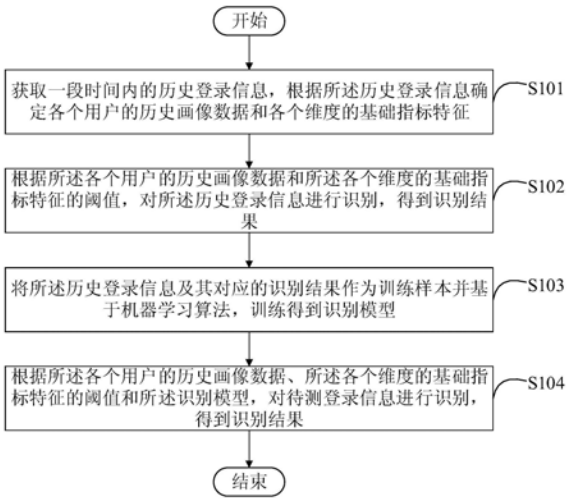
权利要求书3页 说明书13页 附图5页

(54)发明名称

一种识别异常登录的方法和装置

(57)摘要

本发明公开了一种识别异常登录的方法和装置，涉及计算机技术领域。该方法的一具体实施方式包括：获取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和登录维度的基础指标特征；根据所述各个用户的历史画像数据和所述各个基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果；将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型；根据所述各个用户的历史画像数据、所述各个基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。该实施方式能够解决异常登录识别结果不准确的技术问题。



CN 111400357 A

1.一种识别异常登录的方法,其特征在于,包括:

获取一段时间内的历史登录信息,根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征;

根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值,对所述历史登录信息进行识别,得到识别结果;

将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法,训练得到识别模型;

根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型,对待测登录信息进行识别,得到识别结果。

2.根据权利要求1所述的方法,其特征在于,所述历史登录信息包括登录时间、登录渠道、设备信息、IP地址、账户信息、浏览器信息和登录返回码;

所述历史画像数据包括常用IP、所述常用IP的归属地、常用设备信息和常用浏览器信息;和/或,

所述各个维度的基础指标特征至少包括IP维度的指标特征、设备维度的指标特征和账户维度的指标特征。

3.根据权利要求1所述的方法,其特征在于,根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值,对所述历史登录信息进行识别,得到识别结果,包括:

对所述各个维度的基础指标特征进行组合,并根据所述各个维度的基础指标特征的阈值,得到组合模型;

基于所述组合模型对各条所述历史登录信息进行筛选,得到异常列表;其中,所述异常列表中包括至少一条历史登录信息;

基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别,从而得到识别结果。

4.根据权利要求3所述的方法,其特征在于,基于所述组合模型对各条所述历史登录信息进行筛选,得到异常列表,包括:

对于每个组合模型,基于所述组合模型中的基础指标特征及其对应的阈值,从所述各个维度的基础指标特征中筛选出异常的登录信息;

根据各个所述异常的登录信息,从各条所述历史登录信息中筛选出异常列表。

5.根据权利要求3所述的方法,其特征在于,基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别,从而得到识别结果,包括:

对于所述异常列表中的每条历史登录信息,判断所述历史登录信息与所述历史登录信息对应的用户的历史画像数据是否一致;

若是,则将所述历史登录信息从所述异常列表中剔除;

若否,则识别所述历史登录信息为异常。

6.根据权利要求1所述的方法,其特征在于,将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法,训练得到识别模型,包括:

将所述历史登录信息对应的各个用户的历史画像数据和各个维度的基础指标特征以及所述历史登录信息对应的识别结果作为训练样本,并采用梯度提升决策数模型进行有监督学习,从而训练得到识别模型。

7.根据权利要求1所述的方法，其特征在于，根据所述各个用户的历史画像数据、所述各个基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果，包括：

根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对待测登录信息进行识别，得到第一异常识别结果；

基于所述识别模型对所述待测登录信息进行识别，得到第二异常识别结果；

将所述第一异常识别结果和所述第二异常识别结果取并集，作为所述待测登录信息的识别结果。

8.一种识别异常登录的装置，其特征在于，包括：

计算模块，用于获取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征；

第一识别模块，用于根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果；

训练模块，用于将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型；

第二识别模块，用于根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。

9.根据权利要求8所述的装置，其特征在于，所述历史登录信息包括登录时间、登录渠道、设备信息、IP地址、账户信息、浏览器信息和登录返回码；

所述历史画像数据包括常用IP、所述常用IP的归属地、常用设备信息和常用浏览器信息；和/或，

所述各个维度的基础指标特征至少包括IP维度的指标特征、设备维度的指标特征和账户维度的指标特征。

10.根据权利要求8所述的装置，其特征在于，所述第一识别模块还用于：

对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型；

基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表；其中，所述异常列表中包括至少一条历史登录信息；

基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果。

11.根据权利要求10所述的装置，其特征在于，所述第一识别模块还用于：

对于每个组合模型，基于所述组合模型中的基础指标特征及其对应的阈值，从所述各个维度的基础指标特征中筛选出异常的登录信息；

根据各个所述异常的登录信息，从各条所述历史登录信息中筛选出异常列表。

12.根据权利要求8所述的装置，其特征在于，所述第一识别模块还用于：

对于所述异常列表中的每条历史登录信息，判断所述历史登录信息与所述历史登录信息对应的用户的历史画像数据是否一致；

若是，则将所述历史登录信息从所述异常列表中剔除；

若否，则识别所述历史登录信息为异常。

13.根据权利要求8所述的装置,其特征在于,所述训练模块还用于:

将所述历史登录信息对应的各个用户的历史画像数据和各个维度的基础指标特征以及所述历史登录信息对应的识别结果作为训练样本,并采用梯度提升决策数模型进行有监督学习,从而训练得到识别模型。

14.根据权利要求8所述的装置,其特征在于,所述第二识别模块还用于:

根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值,对待测登录信息进行识别,得到第一异常识别结果;

基于所述识别模型对所述待测登录信息进行识别,得到第二异常识别结果;

将所述第一异常识别结果和所述第二异常识别结果取并集,作为所述待测登录信息的识别结果。

15.一种电子设备,其特征在于,包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-7中任一所述的方法。

16.一种计算机可读介质,其上存储有计算机程序,其特征在于,所述程序被处理器执行时实现如权利要求1-7中任一所述的方法。

一种识别异常登录的方法和装置

技术领域

[0001] 本发明涉及计算机技术领域，尤其涉及一种识别异常登录的方法和装置。

背景技术

[0002] 近年来安全威胁发生了很大变化，定向威胁攻击事件、新型威胁越来越多，依据现有规则、关联分析等技术进行安全风险识别已越来越难以满足未来的发展趋势。

[0003] 目前，各大网站登录的方式主要以手机APP登录为主，攻击者无法篡改登录的请求消息内容，可以通过各种风险介质例如设备指纹，登录IP等来对用户的登录进行异常检测，整体的检测维度可以比较单一明确，检测精准率较高。但是，以浏览器登录方式登录网站，由于风险介质可人为篡改，比如伪造设备信息等，从单一的维度无法准确地定位异常行为，很容易造成误报。

发明内容

[0004] 有鉴于此，本发明实施例提供一种识别异常登录的方法和装置，以解决异常登录识别结果不准确的技术问题。

[0005] 为实现上述目的，根据本发明实施例的一个方面，提供了一种识别异常登录的方法，包括：

[0006] 获取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征；

[0007] 根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果；

[0008] 将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型；

[0009] 根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。

[0010] 可选地，所述历史登录信息包括登录时间、登录渠道、设备信息、IP地址、账户信息、浏览器信息和登录返回码；

[0011] 所述历史画像数据包括常用IP、所述常用IP的归属地、常用设备信息和常用浏览器信息；和/或，

[0012] 所述各个维度的基础指标特征至少包括IP维度的指标特征、设备维度的指标特征和账户维度的指标特征。

[0013] 可选地，根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果，包括：

[0014] 对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型；

[0015] 基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表；其中，所述

异常列表中包括至少一条历史登录信息；

[0016] 基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果。

[0017] 可选地，基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表，包括：

[0018] 对于每个组合模型，基于所述组合模型中的基础指标特征及其对应的阈值，从所述各个维度的基础指标特征中筛选出异常的登录信息；

[0019] 根据各个所述异常的登录信息，从各条所述历史登录信息中筛选出异常列表。

[0020] 可选地，基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果，包括：

[0021] 对于所述异常列表中的每条历史登录信息，判断所述历史登录信息与所述历史登录信息对应的用户的历史画像数据是否一致；

[0022] 若是，则将所述历史登录信息从所述异常列表中剔除；

[0023] 若否，则识别所述历史登录信息为异常。

[0024] 可选地，将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型，包括：

[0025] 将所述历史登录信息对应的各个用户的历史画像数据和各个维度的基础指标特征以及所述历史登录信息对应的识别结果作为训练样本，并采用梯度提升决策数模型进行有监督学习，从而训练得到识别模型。

[0026] 可选地，根据所述各个用户的历史画像数据、所述各个基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果，包括：

[0027] 根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对待测登录信息进行识别，得到第一异常识别结果；

[0028] 基于所述识别模型对所述待测登录信息进行识别，得到第二异常识别结果；

[0029] 将所述第一异常识别结果和所述第二异常识别结果取并集，作为所述待测登录信息的识别结果。

[0030] 另外，根据本发明实施例的另一个方面，提供了一种识别异常登录的装置，包括：

[0031] 计算模块，用于获取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征；

[0032] 第一识别模块，用于根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果；

[0033] 训练模块，用于将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型；

[0034] 第二识别模块，用于根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。

[0035] 可选地，所述历史登录信息包括登录时间、登录渠道、设备信息、IP地址、账户信息、浏览器信息和登录返回码；

[0036] 所述历史画像数据包括常用IP、所述常用IP的归属地、常用设备信息和常用浏览器信息；和/或，

[0037] 所述各个维度的基础指标特征至少包括IP维度的指标特征、设备维度的指标特征和账户维度的指标特征。

[0038] 可选地，所述第一识别模块还用于：

[0039] 对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型；

[0040] 基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表；其中，所述异常列表中包括至少一条历史登录信息；

[0041] 基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果。

[0042] 可选地，所述第一识别模块还用于：

[0043] 对于每个组合模型，基于所述组合模型中的基础指标特征及其对应的阈值，从所述各个维度的基础指标特征中筛选出异常的登录信息；

[0044] 根据各个所述异常的登录信息，从各条所述历史登录信息中筛选出异常列表。

[0045] 可选地，所述第一识别模块还用于：

[0046] 对于所述异常列表中的每条历史登录信息，判断所述历史登录信息与所述历史登录信息对应的用户的历史画像数据是否一致；

[0047] 若是，则将所述历史登录信息从所述异常列表中剔除；

[0048] 若否，则识别所述历史登录信息为异常。

[0049] 可选地，所述训练模块还用于：

[0050] 将所述历史登录信息对应的各个用户的历史画像数据和各个维度的基础指标特征以及所述历史登录信息对应的识别结果作为训练样本，并采用梯度提升决策数模型进行有监督学习，从而训练得到识别模型。

[0051] 可选地，所述第二识别模块还用于：

[0052] 根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对待测登录信息进行识别，得到第一异常识别结果；

[0053] 基于所述识别模型对所述待测登录信息进行识别，得到第二异常识别结果；

[0054] 将所述第一异常识别结果和所述第二异常识别结果取并集，作为所述待测登录信息的识别结果。

[0055] 根据本发明实施例的另一个方面，还提供了一种电子设备，包括：

[0056] 一个或多个处理器；

[0057] 存储装置，用于存储一个或多个程序，

[0058] 当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现上述任一实施例所述的方法。

[0059] 根据本发明实施例的另一个方面，还提供了一种计算机可读介质，其上存储有计算机程序，所述程序被处理器执行时实现上述任一实施例所述的方法。

[0060] 上述发明中的一个实施例具有如下优点或有益效果：因为采用根据各个用户的历史画像数据和各个维度的基础指标特征的阈值对历史登录信息进行识别，将历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法训练得到识别模型，从而结合各个用户的历史画像数据、各个维度的基础指标特征的阈值和识别模型对待测登录信息进行识

别的技术手段，所以克服了现有技术中异常登录识别结果不准确的技术问题。本发明实施例通过规则模型和有监督机器学习模型相结合的方式识别网站异常登录，一方面，通过规则模型的识别结果将无监督学习转化为有监督学习，并通过机器学习的方式解决了复杂规则以及规则阈值无法衡量的问题；另一方面，保证了模型的稳定性和更大范围的召回，从而能够准确识别攻击威胁与潜在未知风险，降低人工风险排查成本，提升安全纵深防御能力。

[0061] 上述的非惯用的可选方式所具有的进一步效果将在下文中结合具体实施方式加以说明。

附图说明

[0062] 附图用于更好地理解本发明，不构成对本发明的不当限定。其中：

[0063] 图1是根据本发明实施例的识别异常登录的方法的主要流程的示意图；

[0064] 图2是根据本发明一个可参考实施例的识别异常登录的方法的主要流程的示意图；

[0065] 图3是根据本发明另一个可参考实施例的识别异常登录的方法的主要流程的示意图；

[0066] 图4是根据本发明实施例的识别异常登录的装置的主要模块的示意图；

[0067] 图5是本发明实施例可以应用于其中的示例性系统架构图；

[0068] 图6是适于用来实现本发明实施例的终端设备或服务器的计算机系统的结构示意图。

具体实施方式

[0069] 以下结合附图对本发明的示范性实施例做出说明，其中包括本发明实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本发明的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

[0070] 在实现本发明过程中，发明人发现现有技术中至少存在如下问题：

[0071] 1)基于规则模型进行检测：当有很多业务属性并且强属性较少时，人工总结规则较困难；规则识别对于不同登录风险的识别覆盖率不够高，很难同时保证精确率和召回率。

[0072] 2)基于有监督机器学习模型进行检测：有监督机器学习需要一定的样例，在风险登录识别场景下，可用的样例可能很少；有监督模型对于已知登录风险具有较好的识别效果，而对于未知的登录风险，只基于有监督机器学习模型无法带来更多的召回率。

[0073] 3)基于无监督机器学习模型(如iForest等异常检测方法)进行检测：iForest在孤立点发现方面没有考虑业务知识，学习获得结果的可解释性不强；在有一定标注样例的情况下，无监督机器学习没有充分利用已有的标注样例。

[0074] 为了解决现有技术中存在的技术问题，本发明实时采用规则模型和有监督机器学习模型相结合的识别方法，基于规则模型可以发现更多的异常登录风险，同时基于有监督机器学习模型可以有效地识别已知的登录风险，从而提高异常登录识别结果的准确性。

[0075] 图1是根据本发明实施例的识别异常登录的方法的主要流程的示意图。作为本发明的一个实施例，如图1所示，所述识别异常登录的方法可以包括：

[0076] 步骤101,获取一段时间内的历史登录信息,根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征。

[0077] 首先获取过去一段时间内(比如过去1个月内、2个月天内或者3个月内等)的历史登录信息,其中,每条历史登录信息可以包括登录时间、登录渠道(APP登录、手机网页登录、电脑登录等)、设备信息(设备标识、mac地址、bios系列号,硬盘系列号等)、IP地址、账户信息、浏览器信息(浏览器名称、版本号等)和登录返回码(成功,失败的类型:用户不存在、密码错误等)。

[0078] 然后,可以通过大数据组件(比如hdfs、hive等组件)对获取的各条历史登录信息进行预处理,预处理主要包括数据过滤和数据异常处理。可选地,数据过滤主要是根据登录渠道过滤出非APP登录的用户登录信息,比如过滤出手机网页登录、电脑登录等非APP登录方式的用户登录信息。可选地,数据异常处理主要是根据预先制定的规则,将异常的信息置为空。例如,异常IP:根据IP的规则识别出异常IP,如果非*.*.*.*的IP地址,则判断为异常IP,置为空;异常设备:通用设备的标识长度小于等于5,根据设备的标识判断是否为异常设备,如果是,则置为空。Mac地址异常:如果Mac地址长度不是12位,则异常,置为空。

[0079] 最后,根据预处理后的历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征。可选地,所述历史画像数据包括常用IP、所述常用IP的归属地、常用设备信息和常用浏览器信息。具体地,可以通过大数据组件加工各个用户的历史画像数据,主要包括各个用户常用的IP地址以及常用IP的归属地、常用的设备信息、常用的浏览器信息等,从而将这些维度的信息泛化为各个用户的画像信息。

[0080] 比如,如果通过统计历史登录信息,发现某用户长期采用同一个IP进行登录,则认为该IP地址属于本人IP,该IP作为该用户的历史画像数据。IP的归属地也是重要画像,如果用户长期在某个地区登录,则可以认为该地区为该用户的历史画像数据。设备信息和浏览器信息同理,不再赘述。

[0081] 可选地,所述各个维度的基础指标特征至少包括IP维度的指标特征、设备维度的指标特征和账户维度的指标特征。具体地,可以通过大数据组件(比如hive的sql)加工各个维度的基础指标特征。例如:

[0082] IP1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录返回码为:密码错误,次数、账号数;

[0083] IP1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录返回码为:用户不存在,次数、账号数;

[0084] IP1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录的账号数、失败账号数、敏感时间(凌晨)登录账号数、失败账号数;

[0085] IP1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录的IP登录前后间隔时间的最大值,均值、方差;

[0086] 设备1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录的次数、失败次数、敏感时间(凌晨)登录次数、失败次数;

[0087] 设备1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录前后间隔时间的最大值,均值、方差;

[0088] 账号1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录的次数、失败次数、敏感时

间(凌晨)登录次数、失败次数；

[0089] 账号1分钟、10分钟、30分钟、60分钟、1天、3天、7天登录的账号数、失败账号数、敏感时间(凌晨)登录账号数、失败账号数。

[0090] 在本发明的实施例中，可以根据撞库扫号模型、暴力破解模型、机器人登录模型、IP跨地域异动模型等规则模型来构建各个维度的基础指标特征，从而根据历史登录信息计算出各个维度的基础指标特征。

[0091] 比如，撞库扫号模型的典型表现为：在一定时间段内，某些客户端/IP集合使用x个不同的密码访问了m个以上的不同账户并且登录成功率低于一定阈值n。那么根据撞库扫号模型构建的基础指标特征可以包括IP的登录次数、IP的登录账户数、设备的登录次数、设备的登录账户数、IP/设备对应账户的登录次数汇总、IP/设备对应账户登录返回码(不同返回码的个数分别统计，撞库扫号主要对应的返回码为账号不存在)等；时间窗口取10分钟、60分钟、1天、3天、7天。

[0092] 暴力破解模型的典型表现为：在一定时间段内，某个设备/IP/账户登录过多次、失败次数较多或者失败的账户数量较多，比如1天内设备/账户登录次数>＝80；10分钟内设备/IP/账户登录失败次数>＝80；10分钟内设备/IP登录失败的账户数>＝16。那么根据暴力破解模型构建的基础指标特征可以包括IP一天的登录次数、一天的登录账户数，登录失败的账户数；时间窗口取10分钟、60分钟、1天、3天、7天。

[0093] 机器人登录模型的典型表现为：在时间序列上可能存在一定规律(比如登录时间间隔分布有规律)，同时操作行为和访问路径相似并且登录失败率高。那么根据机器人登录模型构建的基础指标特征可以前后两次登录的时间差值的均值、登录间隔时间的方差。

[0094] IP跨地域异动模型的典型表现为：账户/设备短时间内的位置切换距离过大且可能出现失败。那么根据IP跨地域异动模型构建的基础特征指标可以包括账户对应的IP单位时间内登录省份的数量、账户对应的IP单位时间内登录的跨地区的距离。时间窗口取10分钟、60分钟、1天、3天、7天。

[0095] 本发明实施例利用大数据组件从海量的数据中提取数据并对其进行处理和计算，解决大数据量性能的问题，而且从价值密度低的大量数据里，有效提取出关键信息，数据的可信度高，有助于提高识别准确性。

[0096] 步骤102，根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果。

[0097] 根据步骤101得到的所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对各条历史登录信息进行识别，识别出异常的历史登录信息。

[0098] 可选地，步骤102可以包括：对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型；基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表；其中，所述异常列表中包括至少一条历史登录信息；基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果。在本发明的实施例中，基于业务规则对基础指标特征进行组合，各种组合模型的维度覆盖全面，有助于提高筛选准确性。可选地，组合模型中的基础指标特征及其对应的阈值可以根据经验确定。

[0099] 可选地，可以根据撞库扫号模型、暴力破解模型、机器人登录模型、IP跨地域异动

模型、非本人登录模型等对应的基础指标特征构建组合模型。比如将撞库扫号模型对应的基础指标特征和非本人登录模型对应的基础指标特征进行组合，得到组合模型，如果命中该组合模型，则将命中的历史登录信息都写入异常列表中。比如，组合模型可以是：一天登录次数大于x，一天失败比例大于x，当前登录非本人常用IP概率小于5％。又比如，组合模型可以是：设备一天登录次数大于x，设备一天失败比例大于x，当前登录非本人常用设备概率小于5％。

[0100] 非本人登录模型的典型表现为：一个账户常用的登录IP段、地点、时间段等，非本人登录可能使用非本人常用环境同时失败率高。构建的基础指标特征：用户常用的登录IP、常用的登录地区、常用的设备、当前登录的地区、当前登录的设备。

[0101] 可选地，基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表，包括：对于每个组合模型，基于所述组合模型中的基础指标特征及其对应的阈值，从所述各个维度的基础指标特征中筛选出异常的登录信息(比如异常的IP、异常的设备、异常的用户账户)；根据各个所述异常的登录信息，从各条所述历史登录信息中筛选出异常列表。

[0102] 可选地，基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果，包括：对于所述异常列表中的每条历史登录信息，判断所述历史登录信息与所述历史登录信息对应的用户的历史画像数据是否一致；若是，则将所述历史登录信息从所述异常列表中剔除；若否，则识别所述历史登录信息为异常。根据IP、设备和用户账户等异常的风险介质筛选出异常列表后，还需要进一步结合用户的历史画像数据来判断异常列表中的各条历史登录信息(因为有可能有些属于正常的用户行为)，从而准确地识别出异常登录。

[0103] 本发明实施例通过集成异常风险登录常用的几个规则模型构建特征，并建立多维度多时间窗口的全方面特征，可以提高识别准确性，进而有助于提高训练得到的识别模型的可靠性。

[0104] 步骤103，将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型。

[0105] 由于在步骤102中对过去一段时间内的历史登录信息进行了识别，因此可以将所述历史登录信息及其对应的识别结果作为训练样本，并基于计算学习模型进行训练，从而训练得到识别模型。可选地，步骤103可以包括：将所述历史登录信息对应的各个用户的历史画像数据和各个维度的基础指标特征以及所述历史登录信息对应的识别结果作为训练样本，并采用梯度提升决策数模型进行有监督学习，从而训练得到识别模型。在本发明的实施例中，可以将步骤101中得到各个用户的历史画像数据和各个维度的基础指标特征作为训练样本，还可以增加用户账户的注册时间、用户年龄等用户基础信息也作为训练样本，从而增加模型的整体泛化能力，提高识别准确性。

[0106] 步骤104，根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。

[0107] 可选地，待测登录信息可以是近期的登录信息，比如前一天的登录信息、前两天的登录信息、前三天的登录信息等，结合步骤102的识别过程和步骤103训练得到的识别模型对近期的登录信息进行准确识别。可选地，待测登录信息也可以当前的某一条登录信息，本发明实施例对此不作限制。

[0108] 可选地，步骤104可以包括：根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对待测登录信息进行识别，得到第一异常识别结果；基于所述识别模型对所述待测登录信息进行识别，得到第二异常识别结果；将所述第一异常识别结果和所述第二异常识别结果取并集，作为所述待测登录信息的识别结果。本发明实施例将两个异常识别结果取并集，并集中的登录信息均为异常登录，从而得到了异常登录的用户账户，然后对这些异常登录的用户账户下触发的线上操作(比如转账、下单、支付等操作)进行召回。

[0109] 本发明实施例将规则模型和有监督机器学习模型的识别结果取并集，因此不需要再针对每个基础指标特征设定准确的阈值。如果采用规则模型识别异常登录，往往阈值比较严苛，虽然准确率较高，但是召回率比较低，本发明实施例在规则模型的基础上，进一步结合识别模型，并将两者的识别结果取并集，既保证了模型的识别准确率，同时又增加了召回率。

[0110] 根据上面所述的各种实施例，可以看出本发明通过根据各个用户的历史画像数据和各个维度的基础指标特征的阈值对历史登录信息进行识别，将历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法训练得到识别模型，从而结合各个用户的历史画像数据、各个维度的基础指标特征的阈值和识别模型对待测登录信息进行识别的技术手段，解决了现有技术中异常登录识别结果不准确的技术问题。本发明实施例通过规则模型和有监督机器学习模型相结合的方式识别网站异常登录，一方面，通过规则模型的识别结果将无监督学习转化为有监督学习，并通过机器学习的方式解决了复杂规则以及规则阈值无法衡量的问题；另一方面，保证了模型的稳定性和更大范围的召回，从而能够准确识别攻击威胁与潜在未知风险，降低人工风险排查成本，提升安全纵深防御能力。

[0111] 图2是根据本发明一个可参考实施例的识别异常登录的方法的主要流程的示意图。作为本发明的又一个实施例，如图2所示，所述识别异常登录的方法可以包括：

[0112] 步骤201，获取一段时间内的历史登录信息。

[0113] 可选地，可以从数据库中获取距离当前最近的1个月、2个月或者3个月内的历史登录信息，每条历史登录信息可以包括登录时间、登录渠道(APP登录、手机网页登录、电脑登录等)、设备信息(设备标识、mac地址、bios系列号，硬盘系列号等)、IP地址、账户信息、浏览器信息(浏览器名称、版本号等)和登录返回码(成功，失败的类型：用户不存在、密码错误等)。

[0114] 步骤202，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征。

[0115] 可选地，所述历史画像数据包括常用IP、所述常用IP的归属地、常用设备信息和常用浏览器信息。具体地，可以通过大数据组件加工各个用户的历史画像数据，主要包括各个用户常用的IP地址以及常用IP的归属地、常用的设备信息、常用的浏览器信息等，从而将这些维度的信息泛化为各个用户的画像信息。可选地，所述各个维度的基础指标特征至少包括IP维度的指标特征、设备维度的指标特征和账户维度的指标特征。具体地，可以通过大数据组件加工各个维度的基础指标特征。可以根据撞库扫号模型、暴力破解模型、机器人登录模型、IP跨地域异动模型等规则模型来构建各个维度的基础指标特征，从而根据历史登录信息计算出各个维度的基础指标特征。

[0116] 步骤203，对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础

指标特征的阈值，得到组合模型。

[0117] 可选地，可以根据撞库扫号模型、暴力破解模型、机器人登录模型、IP跨地域异动模型、非本人登录模型等对应的基础指标特征构建组合模型，并根据经验设定基础指标特征的阈值。

[0118] 步骤204，基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表。

[0119] 具体地，如果某条历史登录信息命中组合模型，则将命中的历史登录信息写入异常列表中。

[0120] 步骤205，基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果。

[0121] 对于异常列表中的每条历史登录信息，结合用户的历史画像数据进一步判断是否异常，从而提高识别准确性。

[0122] 步骤206，将所述历史登录信息对应的各个用户的历史画像数据和各个维度的基础指标特征以及所述历史登录信息对应的识别结果作为训练样本，并采用梯度提升决策数模型进行有监督学习，从而训练得到识别模型。

[0123] 在训练模型时，还可以增加用户账户的注册时间、用户年龄等用户基础信息也作为训练样本，从而增加模型的整体泛化能力，提高识别准确性。

[0124] 步骤207，根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对待测登录信息进行识别，得到第一异常识别结果。

[0125] 步骤208，基于所述识别模型对所述待测登录信息进行识别，得到第二异常识别结果。

[0126] 步骤209，将所述第一异常识别结果和所述第二异常识别结果取并集，作为所述待测登录信息的识别结果。

[0127] 另外，在本发明一个可参考实施例中识别异常登录的方法的具体实施内容，在上面所述识别异常登录的方法中已经详细说明了，故在此重复内容不再说明。

[0128] 图3是根据本发明另一个可参考实施例的识别异常登录的方法的主要流程的示意图。作为本发明的另一个实施例，如图3所示，所述识别异常登录的方法可以包括：

[0129] 步骤301，取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征。

[0130] 步骤302，对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型。

[0131] 步骤303，对于每个组合模型，基于所述组合模型中的基础指标特征及其对应的阈值，从所述各个维度的基础指标特征中筛选出异常的登录信息(比如异常的IP、异常的设备、异常的用户账户)。

[0132] 步骤304，根据各个所述异常的登录信息，从各条所述历史登录信息中筛选出异常列表。

[0133] 步骤305，对于所述异常列表中的每条历史登录信息，判断所述历史登录信息与所述历史登录信息对应的用户的历史画像数据是否一致；若是，则执行步骤306；若否，则执行步骤307。

[0134] 步骤306，将所述历史登录信息从所述异常列表中剔除。

[0135] 步骤307，识别所述历史登录信息为异常。

[0136] 步骤308，将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型。

[0137] 步骤309，根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。

[0138] 另外，在本发明一个可参考实施例中识别异常登录的方法的具体实施内容，在上面所述识别异常登录的方法中已经详细说明了，故在此重复内容不再说明。

[0139] 图4是根据本发明实施例的识别异常登录的装置的主要模块的示意图。如图4所示，所述识别异常登录的装置400包括计算模块401、第一识别模块402、训练模块403和第二识别模块404。计算模块401用于获取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征；第一识别模块402用于根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果；训练模块403用于将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型；第二识别模块404用于根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。

[0140] 可选地，所述历史登录信息包括登录时间、登录渠道、设备信息、IP地址、账户信息、浏览器信息和登录返回码；

[0141] 所述历史画像数据包括常用IP、所述常用IP的归属地、常用设备信息和常用浏览器信息；和/或，

[0142] 所述各个维度的基础指标特征至少包括IP维度的指标特征、设备维度的指标特征和账户维度的指标特征。

[0143] 可选地，所述第一识别模块402还用于：

[0144] 对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型；

[0145] 基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表；其中，所述异常列表中包括至少一条历史登录信息；

[0146] 基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果。

[0147] 可选地，所述第一识别模块402还用于：

[0148] 对于每个组合模型，基于所述组合模型中的基础指标特征及其对应的阈值，从所述各个维度的基础指标特征中筛选出异常的登录信息；

[0149] 根据各个所述异常的登录信息，从各条所述历史登录信息中筛选出异常列表。

[0150] 可选地，所述第一识别模块402还用于：

[0151] 对于所述异常列表中的每条历史登录信息，判断所述历史登录信息与所述历史登录信息对应的用户的历史画像数据是否一致；

[0152] 若是，则将所述历史登录信息从所述异常列表中剔除；

[0153] 若否，则识别所述历史登录信息为异常。

[0154] 可选地，所述训练模块403还用于：

[0155] 将所述历史登录信息对应的各个用户的历史画像数据和各个维度的基础指标特征以及所述历史登录信息对应的识别结果作为训练样本，并采用梯度提升决策数模型进行有监督学习，从而训练得到识别模型。

[0156] 可选地，所述第二识别模块404还用于：

[0157] 根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对待测登录信息进行识别，得到第一异常识别结果；

[0158] 基于所述识别模型对所述待测登录信息进行识别，得到第二异常识别结果；

[0159] 将所述第一异常识别结果和所述第二异常识别结果取并集，作为所述待测登录信息的识别结果。

[0160] 根据上面所述的各种实施例，可以看出本发明通过根据各个用户的历史画像数据和各个维度的基础指标特征的阈值对历史登录信息进行识别，将历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法训练得到识别模型，从而结合各个用户的历史画像数据、各个维度的基础指标特征的阈值和识别模型对待测登录信息进行识别的技术手段，解决了现有技术中异常登录识别结果不准确的技术问题。本发明实施例通过规则模型和有监督机器学习模型相结合的方式识别网站异常登录，一方面，通过规则模型的识别结果将无监督学习转化为有监督学习，并通过机器学习的方式解决了复杂规则以及规则阈值无法衡量的问题；另一方面，保证了模型的稳定性和更大范围的召回，从而能够准确识别攻击威胁与潜在未知风险，降低人工风险排查成本，提升安全纵深防御能力。

[0161] 需要说明的是，在本发明所述识别异常登录的装置的具体实施内容，在上面所述识别异常登录的方法中已经详细说明了，故在此重复内容不再说明。

[0162] 图5示出了可以应用本发明实施例的识别异常登录的方法或识别异常登录的装置的示例性系统架构500。

[0163] 如图5所示，系统架构500可以包括终端设备501、502、503，网络504和服务器505。网络504用以在终端设备501、502、503和服务器505之间提供通信链路的介质。网络504可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

[0164] 用户可以使用终端设备501、502、503通过网络504与服务器505交互，以接收或发送消息等。终端设备501、502、503上可以安装有各种通讯客户端应用，例如购物类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等(仅为示例)。

[0165] 终端设备501、502、503可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。

[0166] 服务器505可以是提供各种服务的服务器，例如对用户利用终端设备501、502、503所浏览的购物类网站提供支持的后台管理服务器(仅为示例)。后台管理服务器可以对接收到的物品信息查询请求等数据进行分析等处理，并将处理结果(例如目标推送信息、物品信息——仅为示例)反馈给终端设备。

[0167] 需要说明的是，本发明实施例所提供的识别异常登录的方法一般由服务器505执行，相应地，所述识别异常登录的装置一般设置在服务器505中。本发明实施例所提供的识别异常登录的方法也可以由终端设备501、502、503执行，相应地，所述识别异常登录的装置可以设置在终端设备501、502、503中。

[0168] 应该理解，图5中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需

要,可以具有任意数目的终端设备、网络和服务器。

[0169] 下面参考图6,其示出了适于用来实现本发明实施例的终端设备的计算机系统600的结构示意图。图6示出的终端设备仅仅是一个示例,不应对本发明实施例的功能和使用范围带来任何限制。

[0170] 如图6所示,计算机系统600包括中央处理单元(CPU)601,其可以根据存储在只读存储器(ROM)602中的程序或者从存储部分608加载到随机访问存储器(RAM)603中的程序而执行各种适当的动作和处理。在RAM 603中,还存储有系统600操作所需的各种程序和数据。CPU 601、ROM 602以及RAM603通过总线604彼此相连。输入/输出(I/O)接口605也连接至总线604。

[0171] 以下部件连接至I/O接口605:包括键盘、鼠标等的输入部分606;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分607;包括硬盘等的存储部分608;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分609。通信部分609经由诸如因特网的网络执行通信处理。驱动器610也根据需要连接至I/O接口605。可拆卸介质611,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器610上,以便于从其上读出的计算机程序根据需要被安装入存储部分608。

[0172] 特别地,根据本发明公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本发明公开的实施例包括一种计算机程序,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分609从网络上被下载和安装,和/或从可拆卸介质611被安装。在该计算机程序被中央处理单元(CPU)601执行时,执行本发明的系统中限定的上述功能。

[0173] 需要说明的是,本发明所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本发明中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本发明中,计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:无线、电线、光缆、RF等等,或者上述的任意合适的组合。

[0174] 附图中的流程图和框图,图示了按照本发明各种实施例的系统、方法和计算机程序的可能实现的体系架构、功能和操作。在这点上,流程图或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,上述模块、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，两个接连地表示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，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操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。

[0175] 描述于本发明实施例中所涉及到的模块可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的模块也可以设置在处理器中，例如，可以描述为：一种处理器包括计算模块、第一识别模块、训练模块和第二识别模块，其中，这些模块的名称在某种情况下并不构成对该模块本身的限定。

[0176] 作为另一方面，本发明还提供了一种计算机可读介质，该计算机可读介质可以是上述实施例中描述的设备中所包含的；也可以是单独存在，而未装配入该设备中。上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被一个该设备执行时，使得该设备包括：获取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征；根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果；将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型；根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果。

[0177] 根据本发明实施例的技术方案，因为采用根据各个用户的历史画像数据和各个维度的基础指标特征的阈值对历史登录信息进行识别，将历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法训练得到识别模型，从而结合各个用户的历史画像数据、各个维度的基础指标特征的阈值和识别模型对待测登录信息进行识别的技术手段，所以克服了现有技术中异常登录识别结果不准确的技术问题。本发明实施例通过规则模型和有监督机器学习模型相结合的方式识别网站异常登录，一方面，通过规则模型的识别结果将无监督学习转化为有监督学习，并通过机器学习的方式解决了复杂规则以及规则阈值无法衡量的问题；另一方面，保证了模型的稳定性和更大范围的召回，从而能够准确识别攻击威胁与潜在未知风险，降低人工风险排查成本，提升安全纵深防御能。

[0178] 上述具体实施方式，并不构成对本发明保护范围的限制。本领域技术人员应该明白的是，取决于设计要求和其他因素，可以发生各种各样的修改、组合、子组合和替代。任何在本发明的精神和原则之内所作的修改、等同替换和改进等，均应包含在本发明保护范围之内。

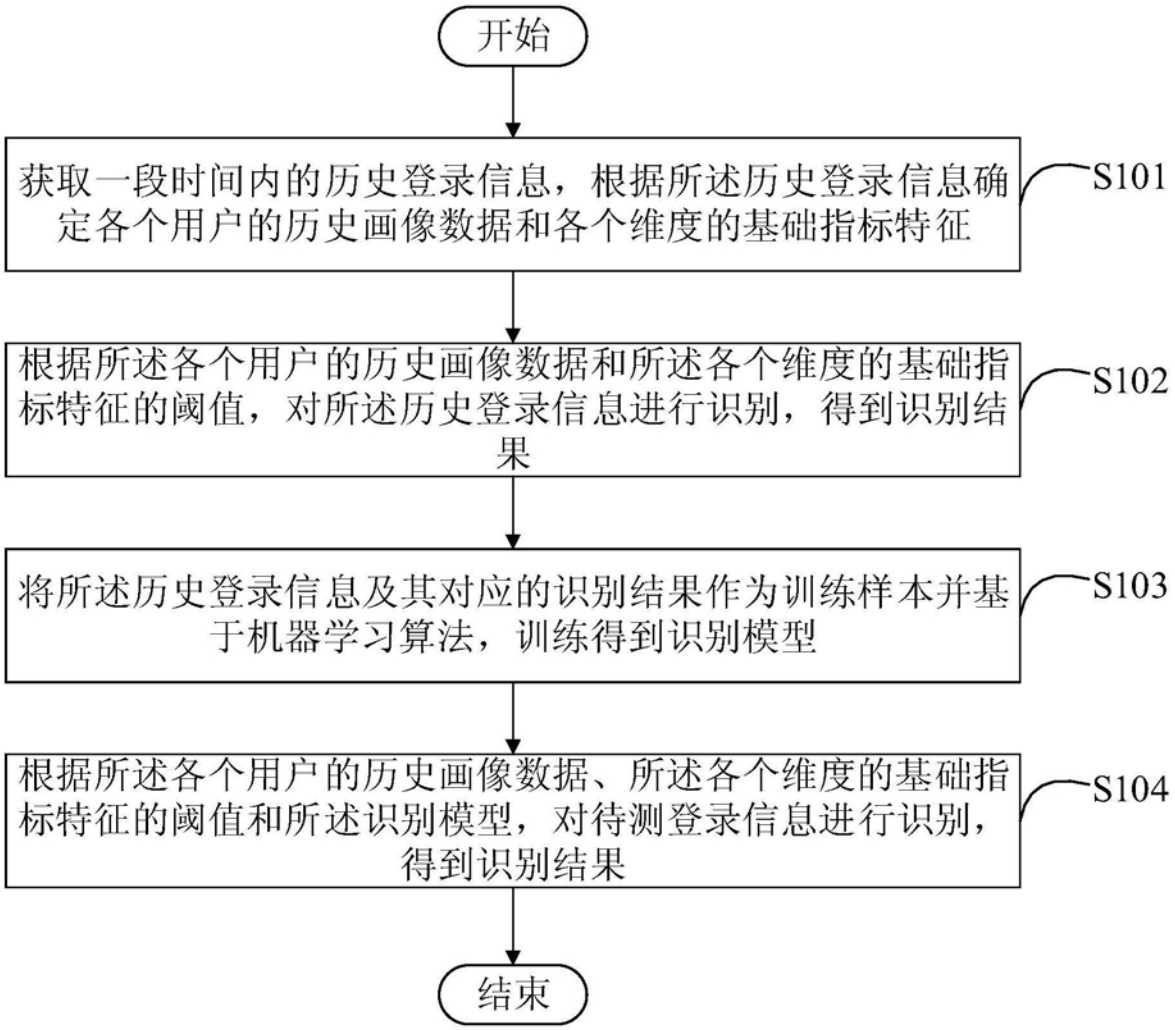

开始
获取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征
S101
根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对所述历史登录信息进行识别，得到识别结果
S102
将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型
S103
根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果
S104
结束

图1

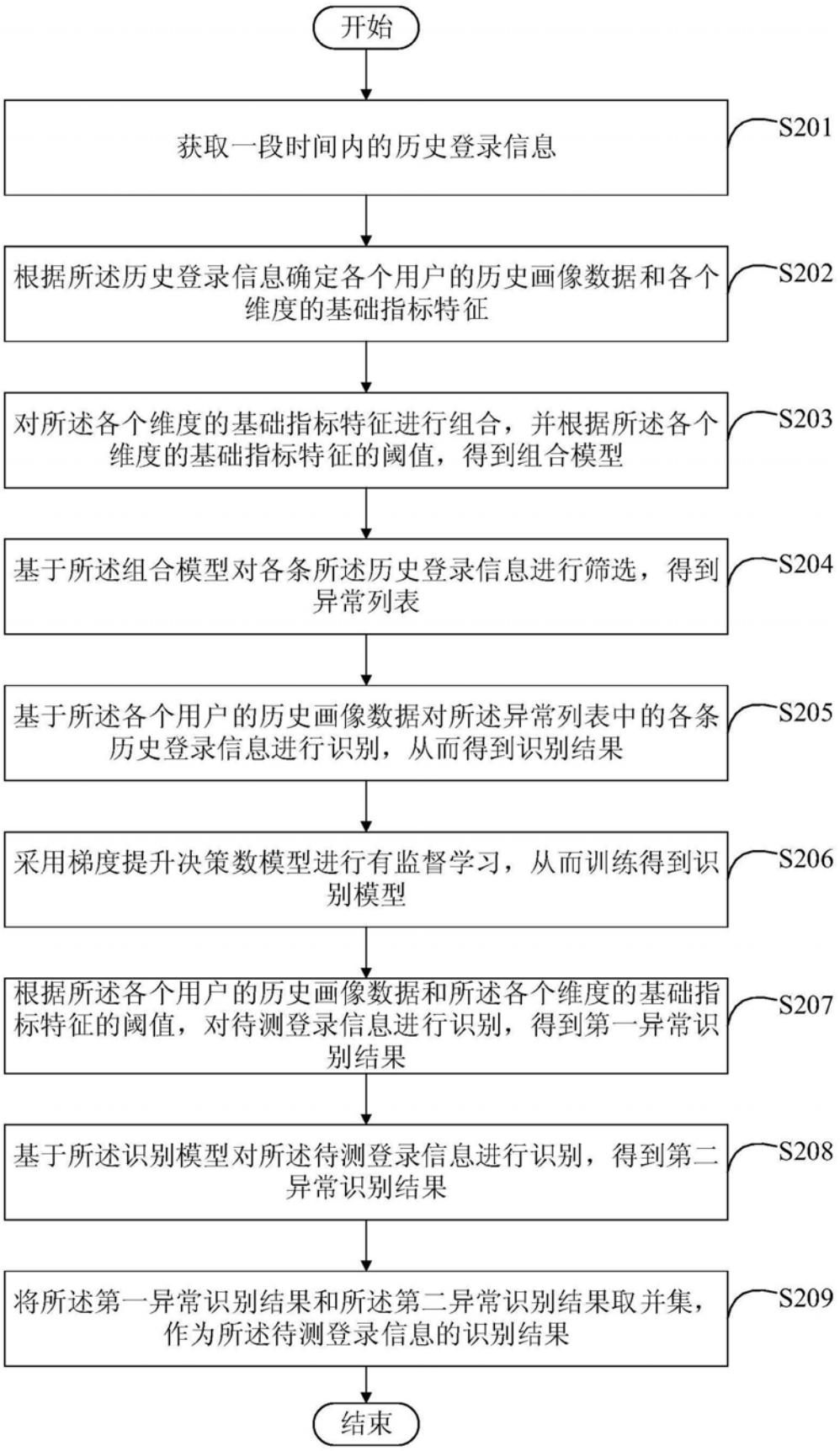

开始
获取一段时间内的历史登录信息
S201
根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征
S202
对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型
S203
基于所述组合模型对各条所述历史登录信息进行筛选，得到异常列表
S204
基于所述各个用户的历史画像数据对所述异常列表中的各条历史登录信息进行识别，从而得到识别结果
S205
采用梯度提升决策数模型进行有监督学习，从而训练得到识别模型
S206
根据所述各个用户的历史画像数据和所述各个维度的基础指标特征的阈值，对待测登录信息进行识别，得到第一异常识别结果
S207
基于所述识别模型对所述待测登录信息进行识别，得到第二异常识别结果
S208
将所述第一异常识别结果和所述第二异常识别结果取并集，作为所述待测登录信息的识别结果
S209
结束

图2

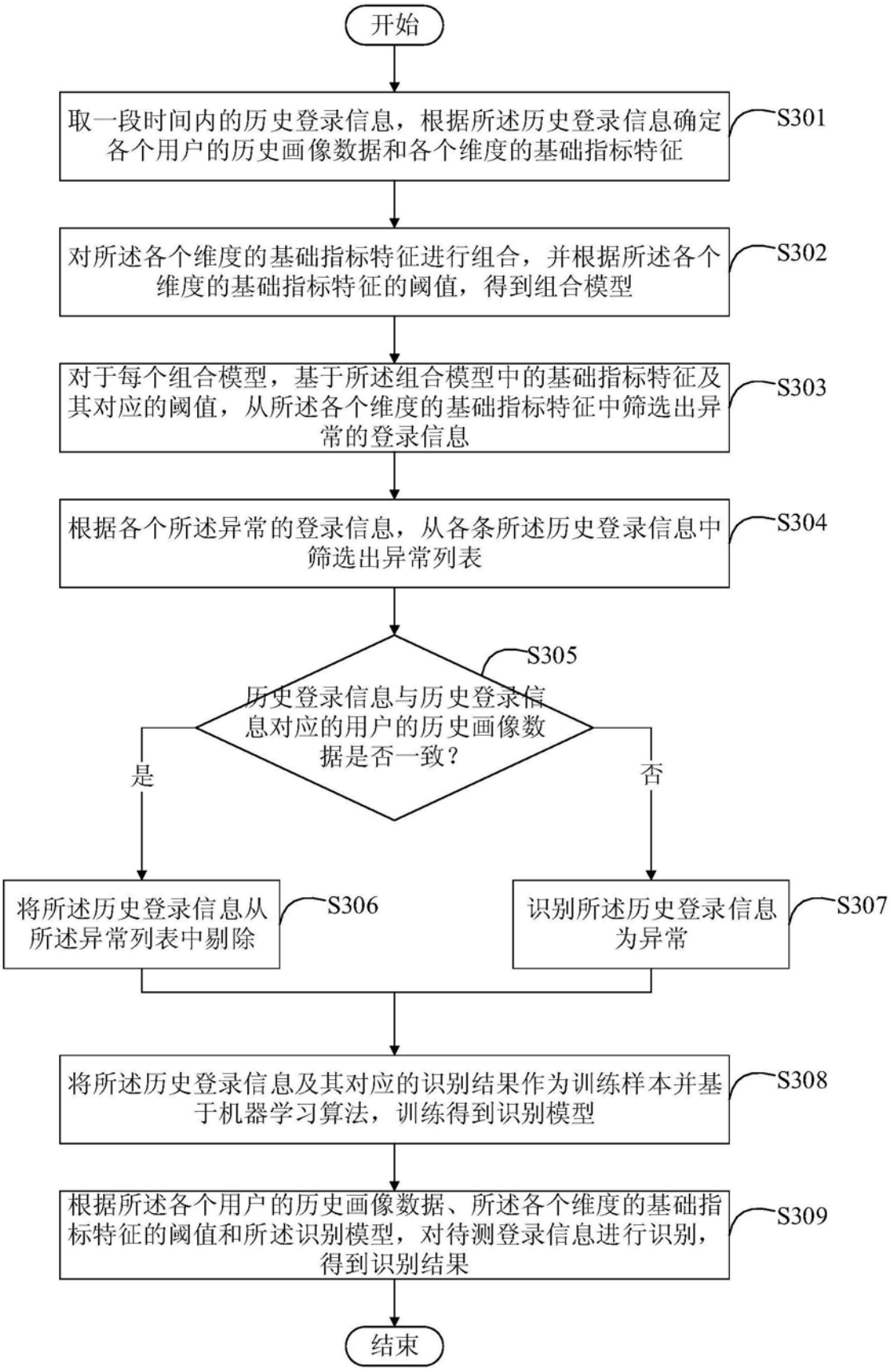

开始
取一段时间内的历史登录信息，根据所述历史登录信息确定各个用户的历史画像数据和各个维度的基础指标特征
S301
对所述各个维度的基础指标特征进行组合，并根据所述各个维度的基础指标特征的阈值，得到组合模型
S302
对于每个组合模型，基于所述组合模型中的基础指标特征及其对应的阈值，从所述各个维度的基础指标特征中筛选出异常的登录信息
S303
根据各个所述异常的登录信息，从各条所述历史登录信息中筛选出异常列表
S304
S305
历史登录信息与历史登录信息对应的用户的历史画像数据是否一致？
是
否
将所述历史登录信息从所述异常列表中剔除
S306
识别所述历史登录信息为异常
S307
将所述历史登录信息及其对应的识别结果作为训练样本并基于机器学习算法，训练得到识别模型
S308
根据所述各个用户的历史画像数据、所述各个维度的基础指标特征的阈值和所述识别模型，对待测登录信息进行识别，得到识别结果
S309
结束

图3

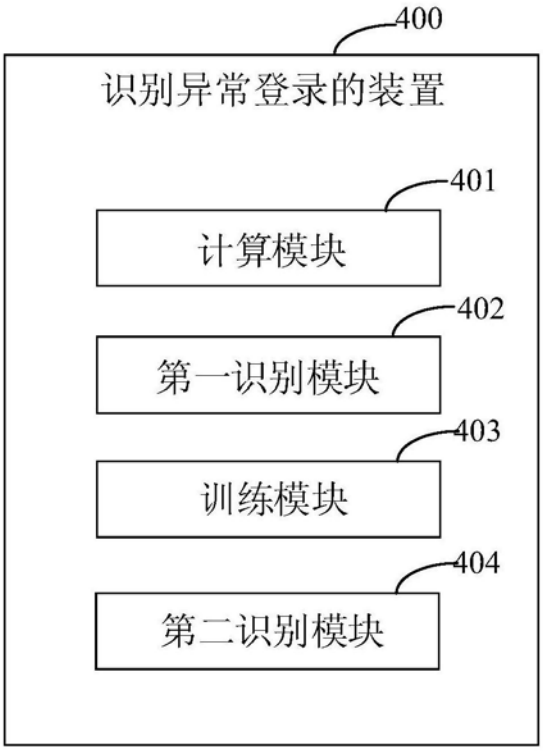

400
识别异常登录的装置
401
计算模块
402
第一识别模块
403
训练模块
404
第二识别模块

图4

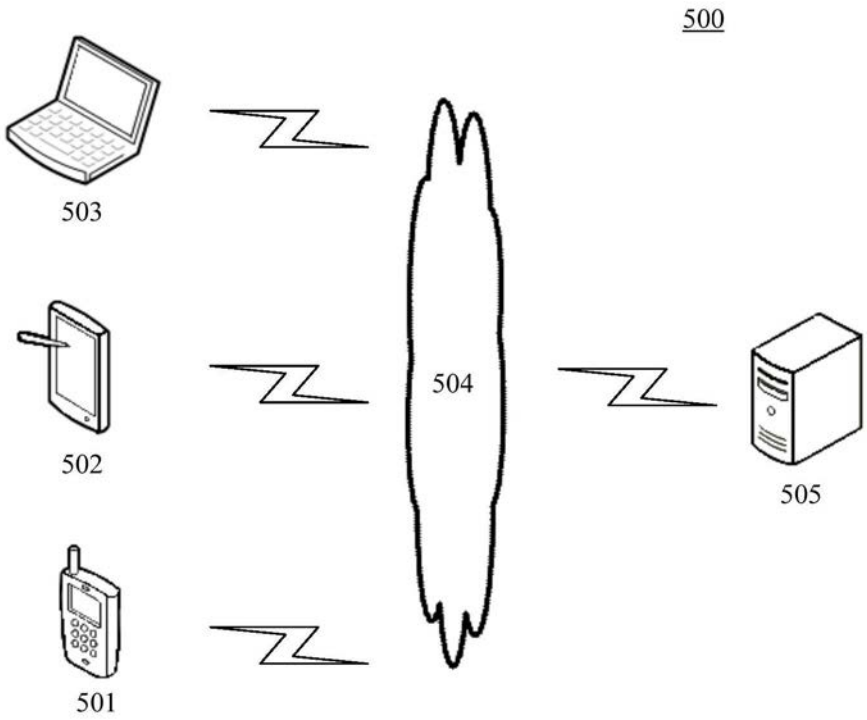

500
503
502
501
504
505

图5

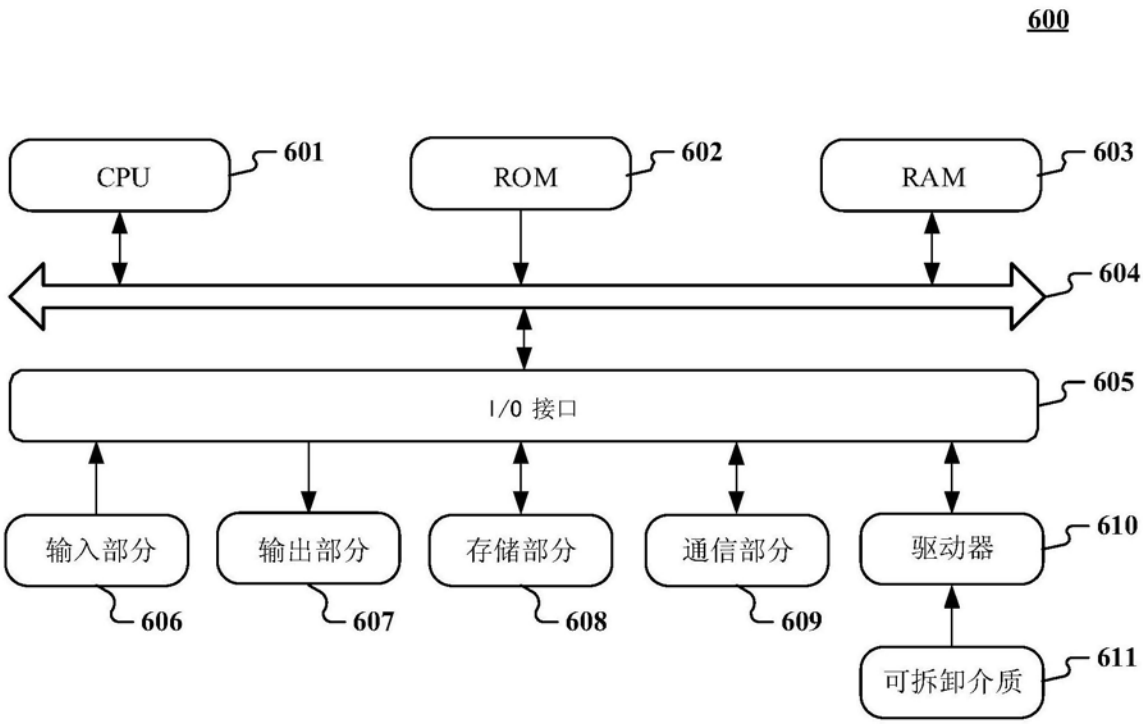

600
CPU
601
ROM
602
RAM
603
604
I/O 接口
605
输入部分
606
输出部分
607
存储部分
608
通信部分
609
驱动器
610
可拆卸介质
611

图6