(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 106295349 B
(45)授权公告日 2020.06.05

(21)申请号 201510289825.4

(22)申请日 2015.05.29

(65)同一申请的已公布的文献号
申请公布号 CN 106295349 A

(43)申请公布日 2017.01.04

(73)专利权人 阿里巴巴集团控股有限公司
地址 英属开曼群岛大开曼资本大厦一座四层847号邮箱

(72)发明人 谭纯平

(74)专利代理机构 北京清源汇知识产权代理事务所(特殊普通合伙) 11644
代理人 冯德魁

(51)Int.Cl.
G06F 21/57(2013.01)

(56)对比文件
CN 103024744 A,2013.04.03,权利要求1-6,说明书第0019-0047段.

审查员 张曼

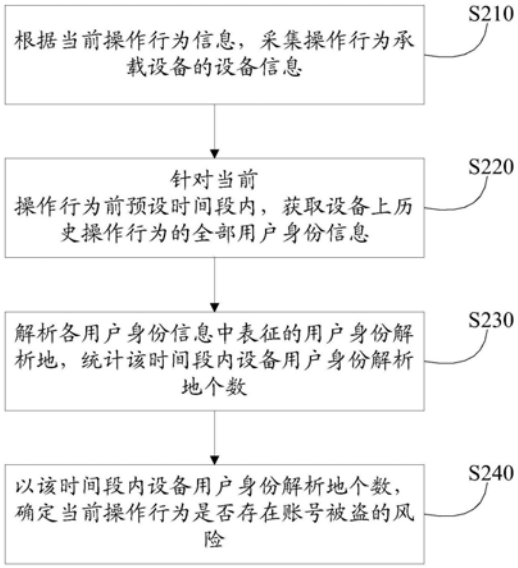
权利要求书3页 说明书11页 附图2页

(54)发明名称

账号被盗的风险识别方法、识别装置及防控系统

(57)摘要

本申请提供一种账号被盗的风险识别方法、风险识别装置及防控系统。该方法采集操作行为登陆设备的设备信息,获取操作行为前预设时间段内设备上历史操作行为的全部用户身份信息,解析各用户身份解析地并统计该时间段内设备用户身份解析地个数,以此来确定当前操作行为是否存在账号被盗的风险。该装置具有设备信息采集模块、用户信息获取模块、用户身份解析模块及盗账风险评定模块,可以根据当前操作行为前预设时间段内设备用户身份解析地个数来确定当前操作行为是否存在账号被盗的风险。该系统采用通过用户身份解析地个数来评定风险等级的装置,可以更有效地识别、处理盗账风险,可以为用户提供安全的网络环境。



CN 106295349 B

1.一种账号被盗的风险识别方法,其特征在于,包括:

根据当前操作行为信息,采集操作行为承载设备的设备信息;

针对当前操作行为前预设时间段内,获取设备上历史操作行为的全部用户身份信息;

解析各用户身份信息中表征的用户身份解析地,统计该时间段内设备用户身份解析地个数;以及

根据该时间段内设备用户身份解析地个数,确定当前操作行为是否存在账号被盗的风险,包括:根据该时间段内设备用户身份解析地个数,结合用户身份解析地的区间操作行为数、账盗操作数,确定设备的盗账风险等级;结合当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数,确定当前操作行为用户的盗账风险等级;结合设备的盗账风险等级和当前操作行为用户的盗账风险等级来计算账号被盗风险值,在风险值大于预设阈值时识别为账号被盗。

2.如权利要求1所述的风险识别方法,其特征在于,所述用户身份信息包括用户注册信息中的证件信息;

所述解析各用户身份信息中表征的用户身份解析地,统计该时间段内设备用户身份解析地个数步骤,具体包括:根据各个用户注册信息中的证件类型和证件号码,获取用户身份解析地并统计设备用户身份解析地个数。

3.如权利要求2所述的风险识别方法,其特征在于,所述根据各个用户注册信息中的证件类型和证件号码,获取用户身份解析地并统计设备用户身份解析地个数步骤,包括:

根据证件类型的种类,确定用户身份解析地的解析方式;

在证件类型为中国境内居民身份证时,解析每一证件号码的前六位来获取用户身份解析地,并据此统计设备用户身份解析地个数;

在证件类型为中国境内非居民身份证或中国境外证件时,推定每一证件类型或每一证件号码对应一个用户身份解析地,并据此统计设备用户身份解析地个数。

4.如权利要求1所述的风险识别方法,其特征在于,所述根据当前操作行为信息,采集操作行为承载设备的设备信息步骤,包括:通过采集设备的设备标识码,来获取设备相应的设备信息。

5.如权利要求4所述的风险识别方法,其特征在于,所述通过采集设备的设备标识码,来获取设备相应的设备信息步骤,包括:

根据设备的种类,确定设备信息的采集内容;

在设备为PC时,采集的设备信息包括MAC、IP和/或UMID;

在设备为移动终端时,采集的设备信息包括MAC、IMEI、TID和/或手机号。

6.如权利要求4所述的风险识别方法,其特征在于,包括:

根据由设备标识码识别出的设备数量,确定设备用户身份解析地个数的统计方式;

在识别出唯一设备时,解析统计该设备用户身份解析地个数;

在识别出多个设备时,解析统计各设备用户身份解析地个数;

在未识别出设备时,将设备用户身份解析地个数计为0;

得到的设备用户身份解析地个数作为预设评分模型的输入变量,用来评定设备的盗账风险等级。

7.如权利要求1～6任一项所述的风险识别方法,其特征在于,所述结合当前操作行为

用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数，确定当前操作行为用户的盗账风险等级，包括：结合当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数，并结合设备上全部用户个数、当前用户当次操作行为信息与历史操作行为信息的差异和/或当次操作行为路由特征信息与历史操作行为路由特征信息是否相同，来评定当前操作行为用户的盗账风险等级。

8.一种账号被盗的风险识别装置，其特征在于，包括：

设备信息采集模块，根据当前操作行为信息，采集操作行为承载设备的设备信息；

用户信息获取模块，针对当前操作行为前预设时间段内，获取设备上历史操作行为的全部用户身份信息；

用户身份解析模块，解析各用户身份信息中表征的用户身份解析地，统计该时间段内设备用户身份解析地个数；以及

盗账风险评定模块，根据该时间段内设备用户身份解析地个数，确定当前操作行为是否存在账号被盗的风险，包括：根据该时间段内设备用户身份解析地个数，结合用户身份解析地的区间操作行为数、账盗操作数，确定设备的盗账风险等级；结合当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数，确定当前操作行为用户的盗账风险等级；结合设备的盗账风险等级和当前操作行为用户的盗账风险等级来计算账号被盗风险值，在风险值大于预设阈值时识别为账号被盗。

9.如权利要求8所述的风险识别装置，其特征在于，所述用户身份信息包括用户注册信息中的证件信息；

所述用户身份解析模块根据各个的用户注册信息中的证件类型和证件号码，获取用户身份解析地并统计设备用户身份解析地个数。

10.如权利要求9所述的风险识别装置，其特征在于，所述用户身份解析模块根据证件类型的种类，确定用户身份解析地的解析方式；在证件类型为中国境内居民身份证时，解析每一证件号码的前六位来获取用户身份解析地，并据此统计设备用户身份解析地个数；在证件类型为中国境内非居民身份证或中国境外证件时，推定每一证件类型或每一证件号码对应一个用户身份解析地，并据此统计设备用户身份解析地个数。

11.如权利要求8所述的风险识别装置，其特征在于，所述设备信息采集模块通过采集设备的设备标识码，来获取设备相应的设备信息。

12.如权利要求8所述的风险识别装置，其特征在于，所述设备信息采集模块根据设备的种类，确定设备信息的采集内容；在设备为PC时，采集的设备信息包括MAC、IP和/或UMID；在设备为移动终端时，采集的设备信息包括MAC、IMEI、TID和/或手机号。

13.如权利要求8所述的风险识别装置，其特征在于，所述用户身份解析模块根据由所述设备信息采集模块通过设备标识码识别出的设备数量，来确定设备用户身份解析地个数的统计方式：识别出唯一设备时，解析统计该设备用户身份解析地个数；识别出多个设备时，解析统计各设备用户身份解析地个数；未识别出设备时，将设备用户身份解析地个数计为0；得到的设备用户身份解析地个数作为所述盗账风险评定模块预设评分模型的输入变量，用来评定设备的盗账风险等级。

14.如权利要求8～12任一项所述的风险识别装置，其特征在于，所述盗账风险评定模

块结合当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数，确定当前操作行为用户的盗账风险等级，包括：结合设备上全部用户个数、当前用户当次操作行为信息与历史操作行为信息的差异和/或当次操作行为路由特征信息与历史操作行为路由特征信息是否相同，来评定当前操作行为用户的盗账风险等级。

15.一种账号被盗的风险防控系统，其特征在于，包括：具有如权利要求8～14任一项所述的风险识别装置、盗账上报装置及风险处理装置，其中：

所述风险识别装置用于计算操作行为平台中账号被盗风险值，在风险值大于预设阈值时识别账号被盗；

盗账上报装置，用于在所述风险识别装置识别账号被盗时，上报账号被盗消息于所述风险处理装置及用户接收设备；

所述风险处理装置，用于收到账号被盗消息时，冻结用户被盗账号并拦截与被盗账号关联的风险数据。

16.如权利要求15所述的风险防控系统，其特征在于，所述系统包括：具有案件数据库，用于存放所述风险处理装置拦截的所述风险数据，以供所述风险处理装置对所述风险数据进行核查，以及供所述风险识别装置对预设评分模型进行校验。

账号被盗的风险识别方法、识别装置及防控系统

技术领域

[0001] 本申请涉及网络安全技术，尤其涉及一种账号被盗的风险识别方法、识别装置及防控系统。

背景技术

[0002] 网上交易、移动支付以及其它应用在给用户带来便利的同时，也存在突出的安全隐患。若账号被盗，用户在蒙受财产损失之外，还可能承担盗账者利用被盗账号从事不法行为的风险。如何及时有效地识别账号是否被盗，以便为用户提供尽可能安全的网络环境，这是网络应用服务提供商无法回避也必须解决的一个重要问题。现有技术提出了较多的盗账风险识别方案，以下举例简要说明。

[0003] 一类方案是通过监控交易用户的交易请求是否异常来识别盗账风险。例如，检测用户是否为异地登陆，在发生异地登陆时要求用户进行验证；如验证不成功，将冻结该用户账号。异地登录是常见的盗账表现形式，因此通过监控异地登陆请求有助于及时识别出盗账风险。但是，由于网络运营商可能改变其拥有的IP地址池，特别是城市之间的IP地址调配时，将会导致正常用户被识别成风险用户，这导致盗账识别的误差率较高。

[0004] 另一类方案是通过监控重点设备来识别盗账风险。例如，统计交易登陆设备上的交易用户数，作为识别盗账风险评分模型的输入变量，由此来评价该设备上盗账的风险等级。若一个设备上交易的用户较少，发生盗账风险的概率相对较低；否则，若设备上交易的用户很多，则发生盗账的风险概率大为增加。因此，重点监控这类交易用户较多的设备，在一定程度上能识别出盗账事件。但是，设备上的交易用户数这个变量的区分能力和稳定性较差，对于正常的单设备多用户交易的情形而言，该方案容易导致识别错误。

[0005] 其它网络操作行为的盗账风险识别方案也经常出现误判、漏判的问题，其盗账风险区分能力不够强，导致这些方案的总体效果不够理想。有鉴于此，有必要设计一种新的盗账风险识别方案。

发明内容

[0006] 针对现有技术存在的缺陷，本申请的目的在于提供一种账号被盗的风险识别方法、识别装置及防控系统，以便有效地提高盗账风险区分能力。

[0007] 为解决以上技术问题，本申请提供一种账号被盗的风险识别方法，包括：

[0008] 根据当前操作行为信息，采集操作行为承载设备的设备信息；

[0009] 针对操作行为前预设时间段内，获取设备上历史操作行为的全部用户身份信息；

[0010] 解析各用户身份信息中表征的用户身份解析地，统计该时间段内设备用户身份解析地个数；以及

[0011] 根据该时间段内设备用户身份解析地个数，确定当前操作行为是否存在账号被盗的风险。

[0012] 较优地，所述用户身份信息包括用户注册信息中的证件信息；所述解析各用户身

份信息中表征的用户身份解析地，统计该时间段内设备用户身份解析地个数步骤，具体包括：根据各个的用户注册信息中的证件类型和证件号码，获取用户身份解析地并统计设备用户身份解析地个数。

[0013] 较优地，所述根据各个用户注册信息中的证件类型和证件号码，获取用户身份解析地并统计设备用户身份解析地个数步骤，包括：

[0014] 根据证件类型的种类，确定用户身份解析地的解析方式；

[0015] 在证件类型为中国境内居民身份证时，解析每一证件号码的前六位来获取用户身份解析地，并据此统计设备用户身份解析地个数；

[0016] 在证件类型为中国境内非居民身份证或中国境外证件时，推定每一证件类型或每一证件号码对应一个用户身份解析地，并据此统计设备用户身份解析地个数。

[0017] 较优地，所述根据当前操作行为信息，采集操作行为承载设备的设备信息步骤，包括：通过采集设备的设备标识码，来获取设备相应的设备信息。

[0018] 较优地，所述通过采集设备的设备标识码，来获取设备相应的设备信息步骤，包括：

[0019] 根据设备的种类，确定设备信息的采集内容；

[0020] 在设备为PC时，采集的设备信息包括MAC、IP和/或UMID；

[0021] 在设备为移动终端时，采集的设备信息包括MAC、IMEI、TID和/或手机号。

[0022] 较优地，所述通过采集设备的设备标识码，来获取设备相应的设备信息步骤，包括：

[0023] 根据由设备标识码识别出的设备数量，确定设备用户身份解析地个数的统计方式；

[0024] 在识别出唯一设备时，解析统计该设备用户身份解析地个数；

[0025] 在识别出多个设备时，解析统计各设备用户身份解析地个数；

[0026] 在未识别出设备时，将设备用户身份解析地个数计为0；

[0027] 得到的设备用户身份解析地个数作为预设评分模型的输入变量，用来评定设备的盗账风险等级。

[0028] 较优地，所述根据该时间段内设备用户身份解析地个数，确定当前操作行为是否存在账号被盗的风险步骤，包括：结合设备上全部用户个数、当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP 地址个数、当前用户的当次操作行为用信息与历史操作行为信息的差异和/或当次操作行为路由特征信息与历史操作行为路由特征信息是否相同，来评定当前操作行为用户的盗账风险等级。

[0029] 较优地，所述根据该时间段内设备用户身份解析地个数，确定当前操作行为是否存在账号被盗的风险步骤，进一步包括：结合设备的盗账风险等级和当前操作行为用户的盗账风险等级来计算账号被盗风险值，在风险值大于预设阈值时识别为账号被盗。

[0030] 在此基础上，本申请还提供一种账号被盗的风险识别装置，包括：

[0031] 设备信息采集模块，根据当前操作行为信息，采集操作行为登陆设备的设备信息；

[0032] 用户信息获取模块，针对操作行为前预设时间段内，获取设备上历史操作行为的全部用户身份信息；

[0033] 用户身份解析模块，解析各用户身份信息中表征的用户身份解析地，统计该时间

段内设备用户身份解析地个数;以及

[0034] 盗账风险评定模块,以该时间段内设备用户身份解析地个数,确定当前操作行为是否存在账号被盗的风险。

[0035] 较优地,所述用户身份信息包括用户注册信息中的证件信息;所述用户身份解析模块根据各个用户注册信息中的证件类型和证件号码,获取用户身份解析地并统计设备用户身份解析地个数。

[0036] 较优地,所述用户身份解析模块根据证件类型的种类,确定用户身份解析地的解析方式;在证件类型为中国境内居民身份证时,解析每一证件号码的前六位来获取用户身份解析地,并据此统计设备用户身份解析地个数;在证件类型为中国境内非居民身份证或中国境外证件时,推定每一证件类型或每一证件号码对应一个用户身份解析地,并据此统计设备用户身份解析地个数。

[0037] 较优地,所述设备信息采集模块通过采集设备的设备标识码,来获取设备相应的设备信息。

[0038] 较优地,所述设备信息采集模块根据设备的种类,确定设备信息的采集内容;在设备为PC时,采集的设备信息包括MAC、IP和/或UMID;在设备为移动终端时,采集的设备信息包括MAC、IMEI、TID和/或手机号。

[0039] 较优地,所述用户身份解析模块根据由所述设备信息采集模块通过设备标识码识别出的设备数量,来确定设备用户身份解析地个数的统计方式:识别出唯一设备时,解析统计该设备用户身份解析地个数;识别出多个设备时,解析统计各设备用户身份解析地个数;未识别出设备时,将设备用户身份解析地个数计为0;得到的设备用户身份解析地个数作为所述盗账风险评定模块预设评分模型的输入变量,用来评定设备的盗账风险等级。

[0040] 较优地,所述盗账风险评定模块结合设备上全部用户个数、当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数、当前用户的当次操作行为用信息与历史操作行为信息的差异和/或当次操作行为路由特征信息与历史操作行为路由特征信息是否相同,来评定当前操作行为用户的盗账风险等级。

[0041] 较优地,所述盗账风险评定模块结合设备的盗账风险等级和当前操作行为用户的盗账风险等级来计算账号被盗风险值,在风险值大于预设阈值时识别为账号被盗。

[0042] 在此基础上,本申请还相应提供一种账号被盗的风险防控系统,包括:具有上述风险识别装置、盗账上报装置及风险处理装置,其中:

[0043] 所述风险识别装置用于计算操作行为平台中账号被盗风险值,在风险值大于预设阈值时识别账号被盗;

[0044] 盗账上报装置,用于在所述风险识别装置识别账号被盗时,上报账号被盗消息于所述风险处理装置及用户接收设备;

[0045] 所述风险处理装置,用于收到账号被盗消息时,冻结用户被盗账号并拦截与被盗账号关联的风险数据。

[0046] 较优地,所述系统包括:具有案件数据库,用于存放所述风险处理装置拦截的所述风险数据,以供所述风险处理装置对所述风险数据进行核查,以及供所述风险识别装置对所述评分模型进行校验。

[0047] 与现有技术相比,本申请提出了一种基于设备用户身份解析地个数来识别盗账风

险的方案，其通过采集操作行为前一段时间内登陆设备上全部用户身份信息，来解析统计该设备用户身份解析地个数，并将该统计量作为风险评分模型的输入变量来评定盗账风险等级，可以有效地提高盗账风险区分能力。这是因为：最近一段时间内登陆设备上全部操作行为用户的不同身份解析地是一类更加有效和稳定的变量，如果操作行为发生时设备在最近一段时间内有多个不同的身份解析地，操作行为账号发生盗账的风险很高；同时，一个设备上相同地方不同用户操作行为的情况远比不同地方不同用户操作行为的情况更为普遍，因而该变量能够有效地剔除部分多用户操作行为但风险较低的情况，这就有助于提升风险评分模型变量区分风险的能力和稳定性。

附图说明

[0048] 通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本申请的限制。而且在整个附图中，用相同的参考符号来表示相同的部件。在附图中：

[0049] 图1示出某网络平台MAC设备在当前操作行为前7天内的设备用户身份解析地个数与盗账风险的关系；

[0050] 图2示出了根据本申请一个实施例账号被盗的风险识别方法的流程图；

[0051] 图3示出了根据本申请一个实施例账号被盗的风险识别装置的方框图；

[0052] 图4示出了根据本申请一个实施例账号被盗的风险防控系统的方框图。

具体实施方式

[0053] 本申请的以下实施例在风险评分模型中引入当前操作行为前一段时间内设备用户身份解析地个数这类输入变量，来提高模型变量的区分风险能力。这一方案需要基于数据挖掘技术来建立风险评分模型，主要的建模步骤包括确定研究目标、确定数据源及抽取样本、数据探索、模型开发和模型验证等等，本申请的重点在于为评分模型构建合适的输入变量。因为评分模型本身不是本申请的主要关注点，故建模的其它细节不再展开，具体请参考习知的现有技术。

[0054] 本申请发明人为了构建合适的变量，收集了用户在某网络平台操作行为的海量数据进行数据分析与数据挖掘。例如，对于上述平台而言，用户的每笔操作行为发生时会采集到设备信息，且还会统计这些设备上最近一段时间内的操作行为上述平台的用户数，通过设备上的上述平台的用户数可以识别盗账风险，但这经常会发生误判的问题。本申请发明人基于数据分析发现：对于上述平台操作行为而言，相同设备上相同地方人操作行为的情况远比不同地方人操作行为的情况更为普遍。因此，更加有效和稳定的变量是统计最近一段时间内该设备上操作行为上述平台用户数的不同的身份解析地，且该身份解析地最好解析到县(市)这一粒度；如果操作行为发生时设备在最近一段时间内有多个不同的身份解析地，这样的操作行为风险是非常高的。

[0055] 上述规律在上述平台操作行为中是非常稳定的，可以理解的是，这一规律同样适应于其它网络应用中的操作行为。有鉴于此，本申请所指“操作行为”是一个广义的概念，其并不仅仅限于上述平台等存在资金和货物转移的商务行为，各种网络应用中用户与服务平台之间、不同用户之间进行的数据交换也都属于本申请“操作行为”的范畴，如社交网站的

登陆事件就属于本申请所指的“操作行为”。

[0056] 由于当前操作行为前一段时间内设备用户身份解析地个数与盗账风险之间存在很强的关联，故本申请发明人将操作行为前预设时间段内设备用户身份解析地个数作为风险评分模型的输入变量。这类变量能够剔除部分多用户操作行为但风险低的情况，从而提升了变量区分风险的能力和稳定性。

[0057] 据此，本申请发明人提出以下基本构思：以当前操作行为前一段时间内设备用户身份解析地个数作为输入变量，通过预设评分模型来评定设备的盗账风险等级，以便更及时有效地识别用户账号被盗的风险。这一技术构思主要涉及三方面内容：变量构建、识别过程、模型校验，以下进一步详细说明如下。

[0058] 1、变量构建

[0059] 在构建有效变量时，需要考虑变量主体、变量对象、时间间隔、统计指标等要素，具体说明如下。

[0060] (1)变量主体：操作行为(如某支付平台操作行为)登陆设备的设备信息。设备信息可通过采集设备的设备标识码，如MAC(Medium Access Control，物理地址)、UMID(Unique Material Identifier，唯一素材识别码)、IP(Internet Protocol，网络互连协议)地址，IMEI(International Mobile Equipment Identity，移动设备国身份码)、TID(THREAD Identifier，线程控制符)、手机号等来识别。通常而言，PC (Personal Computer，个人计算机)可采集设备MAC、IP和/或UMID，移动终端可采集MAC、IMEI、TID和/或手机号。具体的采集、识别方案请参考习知的现有技术即可，不再赘述。

[0061] (2)变量对象：操作行为登陆设备上的用户身份解析地。用户身份解析地通常根据证件类型和证件号码来确定，例如中国居民身份证的前六位就可以表示为县(市)，识别这前六位就可以知道用户来自哪个行政区域，由此也就获得用户身份解析地。

[0062] (3)时间间隔：操作行为前多久时间内的间隔(如30分钟、2小时、12小时、1天、3天、7天等)。各种操作行为平台存在较大差异，具体可根据操作行为请求、操作行为性质等因素确定即可，不再赘述。

[0063] (4)统计指标：统计设备上用户身份解析地个数。正常操作行为环境下，设备上用户身份解析地个数通常较少；若设备用户身份解析地个数较多，表明盗账风险较高，这是根据大量数据分析得出的一个十分可靠的结论。

[0064] 综合上述要素，本申请将具体变量确定为当前操作行为前预设某一时间间隔内操作行为设备的全部用户的不同身份解析地个数，这个统计变量可以提高模型变量的区分风险能力。该变量与盗账风险存在很强的关联性，如果操作行为前最近一段时间内设备上有多个不同的身份解析地，则该设备上操作行为发生账号盗账的风险较高。

[0065] 2、识别过程

[0066] 将上述统计量作为风险评分模型的输入变量之后，可基于不同设备的用户身份解析地来对盗账风险进行识别，其优点在于可提高变量区分风险的能力和稳定性。具体地，盗账风险识别时需遵循以下流程。

[0067] (1)设备身份解析地个数的获取，包括：

[0068] a.获取当前操作行为的设备信息；获取该设备操作行为前的特定时间段内 (比如3天)的全部操作行为用户的身份信息；

[0069] b.解析这些身份信息，获取身份信息中对应的用户地域，并统计设备中不同用户身份解析地个数。此处的用户身份解析地通常为标识为城市，此处的“城市”指的是行政区域，不能狭义地理解为与农村相对的概念。

[0070] c.特殊情况的处理：若存在多个设备，则各自分别统计用户身份解析地个数；若无法获取设备信息，则用户身份解析地个数记为0。

[0071] (2)评定风险等级

[0072] 设备是以用户身份解析地个数来评定风险等级的，具体而言：获得当前操作行为前预设时间段内设备用户身份解析地个数后，将其作为变量输入到风险评分模型之中打分，在综合考虑模型中各变量的权重后，就可以获得设备的盗账风险等级。若得分高，表明盗账风险较高，此时需要重点监控此设备。

[0073] 3、模型校验

[0074] 输入变量“当前操作行为前预设时间段内设备用户身份解析地个数”对于风险评分模型的预测效果的影响如何，应当进行校验。若该变量有效，则可以根据前述第2步的流程来自动识别用户的盗账风险；否则，需要重新调整评分模型及有关输入变量。

[0075] 对于本申请而言，模型校验时需考虑以下因素：

[0076] (1)历史操作行为标识是否为案件。本申请中引入当前操作行为前预设时间段内设备用户身份解析地个数这一统计量时，需要关联设备上的历史操作行为数据，以便证明引入的这类变量是有效的。也就是说，变量是否有效需要通过历史操作行为数据来衡量；换而言之，历史操作行为数据可以对是否盗账进行区分。

[0077] 假定历史操作行为为盗账，标识为“坏”；否则标志位“好”。如果通过第2步所识别盗账风险结果为“坏”，而历史操作行为标识也为“坏”；或者，通过第2步所识别盗账风险结果为“好”，而历史操作行为标识也为“好”；则认为校验通过，否则校验不通过。若校验通过的概率较高，就表明当前操作行为前预设时间段内设备用户身份解析地个数作为输入变量引入到评分模型中是有效的，即该变量具有较高的风险区分能力。

[0078] (2)风险区分能力的量化

[0079] 上述变量的风险区分能力到底如何，可以进一步进行量化。具体可通过分段计算“当前操作行为前预设时间段内设备身份解析地个数”的盗账区分能力指标来实现，这些量化指标主要包括两类：提升度和区间IV值(Infofmation Value，信息值)。

[0080] 盗账区分能力指标的计算公式如下：

[0081] 提升度＝区间盗账户交易浓度/平均盗账户交易浓度

[0082] $$WOE=\ln\left(\frac{\text{区间非盗账户交易占全部非盗账户交易比}}{\text{区间盗账户交易占全部盗账户交易比}}\right)\times 100$$

[0083] 区间IV＝WOE×(区间非盗账户交易占全部非盗账户交易比-区间盗账户交易占全部盗账户交易比)

[0084] IV＝各区间IV之和

[0085] 上述公式中，为了分析的方便，将WOE(Weight Of Eividence，证据权重) 乘以系数100，其内涵与数据挖掘中的指标woe并无本质区别。可以理解的是，上述公式中的“交易”应理解为广义的网络操作行为，而不仅仅局限为资金支付、商品转移等实际的商务活动。

[0086] 按照上述公式计算“当前操作行为前预设时间段内设备身份解析地个数”这一变

量的风险区分能力结果，可以有效验证引入评分模型的有效性。下面以“MAC 设备在当前操作行为前7天内的设备用户身份解析地个数”的盗账区分能力指标为例进行说明，计算结果如表1所示：

[0087] 表1:MAC设备操作行为前7天设备用户身份解析地个数盗账区分能力

[0088]

区间	区间操作行为数	盗账操作行为数	区间盗账操作行为浓度	平均盗账操作行为浓度	提升度	区间IV值	IV值
0	578,007	1,934	0.3%	1.0%	0.33	32.07	171.74
[1,2]	704,478	4,602	0.7%	1.0%	0.65	7.9	171.74
(2,327]	48,887	6,756	13.8%	1.0%	13.82	131.77	171.74

[0089] 表1可用图形化的方式来进行展示。参见图1，其示出某支付平台MAC设备在当前操作行为前7天内设备用户身份解析地个数与盗账风险的关系。由表1和图 1可知，在设备用户身份解析地个数大于2时的提升度为13.82，即通过MAC设备7 天内操作行为用户身份解析地个数来识别盗账风险的能力较提升了13.82倍，由此表明这一变量的盗账区分能力是十分有效的。

[0090] 类似地，对于其它具体实例的校验，各量化指标也是比较理想的。这就表明，本申请在风险评分模型中引入当前操作行为前一段时间内设备用户身份解析地个数这类输入变量来评定盗账风险等级，可以提高模型变量的区分风险能力，由此识别的盗账风险识别效果较为理想。

[0091] 值得注意的是，表1和图中的IV值较大，在某些情况下存在“过预测”可能，为了消除这一现象，本申请还结合设备上全部用户个数、当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数、当前用户的当次操作行为用信息与历史操作行为信息的差异和/或当次操作行为路由特征信息与历史操作行为路由特征信息是否相同等指标，来综合评定用户账号是否被盗的风险，以消除单一变量“过预测”带来的不利影响。

[0092] 以上对本申请利用当前操作行为前预设时间段内设备用户身份解析地个数这一统计量来识别盗账风险的技术构思进行了系统性、原理性的阐述，以下进一步对该技术构思的具体实现方案进行说明。基于前文分析，在确定风险评分模型及输入变量并校验成功之后，只要按照前述第2步在服务器段部署应用程序来具体实施即可，并不需要重复建模及校验。

[0093] 参见图2，出了根据本申请一个实施例账号被盗的风险识别方法的流程图。如图2所示，该风险识别方法包括以下步骤210～步骤240等主要步骤，以下详细进行描述。

[0094] S210、根据当前操作行为信息，采集操作行为登陆设备的设备信息。

[0095] 此步骤响应于当前操作行为信息，由服务器端相应设备来采集操作行为登陆设备的设备信息，一般是通过采集设备的设备标识码来得到的。

[0096] 一般地，客户端的登陆设备会有多种类型，如PC设备经常会有MAC、IP和/或 UMID，移动终端经常会有MAC、IMEI、TID和/或手机号等等，因此需要根据设备的种类来确定设备信息的采集内容。通常而言，PC采集MAC、IP和/或UMID，移动终端采集MAC、、TID和/或手机号即可，具体的信息采集及识别方法请参照现有技术，不再赘述。

[0097] 值得注意的是，此S210步骤中所指的当前操作行为可以是针对用户账号的登录请求，也可以是针对用户账号的预设数据操作请求等。其中，针对用户账号的预设数据操作请求可以包括：针对用户账号的密码修改请求、针对用户账号的余额转账请求、针对用户账号的物品买卖请求等。可以理解的是，预设的数据操作请求可以由服务器预先设置，也可以由用户通过客户端预先设置，在此不进行限定。

[0098] 在客户端发起针对用户账号的登录请求时，用户的登录信息通常会包括用户标识，以及用户发起登录请求的客户端的信息和接收登录请求的服务器的信息。因此，根据用户的登录信息获取该用户登录的路由路径，并从用户登录的路由路径中提取当次路由特征信息，并通过比较当次操作行为路由特征信息与历史操作行为路由特征信息是否相同，也可以来评定当前操作行为用户的盗账风险等级。

[0099] S220、针对操作行为前预设时间段内，获取设备上历史操作行为的全部用户身份信息。

[0100] 可以理解的是，对于单个账号单次操作行为的风险识别非常复杂且也难以实施的，而借助挖掘多个账号操作行为之间的联系是一种非常有效的方法。如前文所述，本申请通过当次操作行为前预设时间段内设备用户身份解析地个数来评定设备的盗账风险，这就需要提取该时间段内设备上历史操作行为的全部用户身份信息，特别是提取用户身份解析地尤为重要。

[0101] 在实际应用中，一般可以根据操作行为平台、操作行为请求、操作行为性质等因素确定设备上用户上述时间间隔(如30分钟、2小时、12小时、1天、3天、7天等)。获得这段时间内历史操作行为的用户信息之后，就可以进一步解析各用户的身份地域，在统计该时间段内设备用户身份解析地个数之后，就可以将其作为风险评分模型的变量来进行评分。

[0102] S230、解析各用户身份信息中表征的用户身份解析地，统计该时间段内设备用户身份解析地个数。

[0103] 如前文所述，当前操作行为前最近一段时间内该设备上操作行为某支付平台用户数的不同的身份解析地是一类非常有效和稳定的变量，因此可将该时间段内的设备用户身份解析地个数的统计量作为变量输入到风险评分模型中评分，最终达到识别用户账号是否被盗的目的。

[0104] 可以理解的是，用户身份解析度的区分粒度对于评分模型的输出结果存在较大关联性。较优地，本申请以城市为用户身份解析地的区分粒度，使得盗账风险识别效果达到较为满意的效果。

[0105] 对于很多网络操作行为平台而言，用户账户注册需要进行实名认证，这对于提高网络安全性是较为有利的。为此，本步骤S230根据各个用户注册信息中的证件类型和证件号码，获取用户身份解析地并统计设备用户身份解析地个数，具体是根据证件类型的种类，确定用户身份解析地的解析方式：

[0106] 若证件类型为中国境内居民身份证，其前六位为县(市)级行政区域，因此可以简单地解析每一证件号码的前六位来获取用户身份解析地，并据此统计设备用户身份解析地个数；

[0107] 若证件类型为中国境内非居民身份证(如军官证)或中国境外证件(如护照) 时，无法直接识别用户身份所在行政区域。但这种情况是比较少的，因此可以简单地考虑推定

每一证件类型或每一证件号码对应一个用户身份解析地，并据此统计设备用户身份解析地个数。当然，若建模时获得这些证件类型的编号方式，则可以根据具体证件号码来获得用户身份解析地，不再赘述。

[0108] 值得注意的是，步骤S210中获取的设备信息可能存在不同的情形：在大多数情况下，可以同时采集到多种设备信息，如MAC、IMEI等；但由于技术原因，某些场景或系统限制下无法采集操作行为时的设备信息；或者，操作行为时采集到的设备信息是一个明显的热点，需要予以排除；等等。针对这些情况，需要相应地调整设备用户身份解析地个数的解析及统计方式。

[0109] 为此，步骤S220～S230中需要根据设备标识码识别出的设备数量，来确定设备用户身份解析地个数的统计方式，具体是：

[0110] 若识别出唯一设备，解析该唯一设备中每一用户身份解析地，并统计该设备用户身份解析地个数；

[0111] 若识别出多个设备，分别解析各个设备中每一用户身份解析地，并统计各个设备用户身份解析地个数；

[0112] 若未识别出设备，将设备用户身份解析地个数计为0；

[0113] 按照上述方式得到的设备用户身份解析地个数，作为预设评分模型的输入变量，引入到风险评分模型来评定设备的盗账风险等级，以便实现通过历史操作行为数据来衡量评分模型中变量对盗账风险的区分能力。

[0114] S240、以该时间段内设备用户身份解析地个数，确定当前操作行为是否存在账号被盗的风险。

[0115] 通常地，以该时间段内设备用户身份解析地个数为预设评分模型的输入变量，评定设备的盗账风险等级。该盗账风险等级表征账号被盗的风险，若盗账风险等级超过设定的阈值，认定账号被盗；否则，认定账号未被盗。

[0116] 根据前述步骤S210～S230获得操作行为前预设时间段内设备用户身份解析地个数的统计量后，就可以将其引入到风险评分模型中评分而得到设备的盗账风险等级，由此实现对账号被盗的风险识别，以便及时采取处理措施来消除风险。

[0117] 通过以上过程实现了对设备的盗账风险等级评定之后，为了保证进一步提高盗账风险识别能力，本申请还进一步结合设备上全部用户个数、当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数、当前用户的当次操作行为用信息与历史操作行为信息的差异和/或当次操作行为路由特征信息与历史操作行为路由特征信息是否相同等要素，来评定当前操作行为用户的盗账风险等级。这样，在结合多种因素之后，本申请的盗账风险识别的能力大大提高。

[0118] 本申请结合设备的盗账风险等级和当前操作行为用户的盗账风险等级来计算账号被盗风险值，在风险值大于预设阈值时识别为账号被盗，并且在识别账号被盗时输出相应的盗账提示信息，以便由操作行为平台和用户及时进行处理，从而消除盗账的安全隐患，避免造成财产损失或其它问题。

[0119] 以上对账号被盗的风险识别方法(以下简称方法)进行了详细描述。在此基础上，本申请还相应地提供账号被盗的风险识别装置(以下简称装置)，以下进行详细的描述。

[0120] 顺便指出的是，本实施例装置中如有描述不尽之处，请参见前文方法部分的描述

内容;同样地,前述方法部分中如涉及到装置的结构,也可以引见以下的描述内容。

[0121] 参见图3,示出了根据本申请一个实施例账号被盗的风险识别装置。该装置300 由设备信息采集模块310、用户信息获取模块320、用户身份解析模块330及盗账风险评定模块340等部分组成,以下对各部分进行说明。

[0122] 设备信息采集模块310,可以根据当前操作行为信息,采集操作行为登陆设备的设备信息。此处,该设备信息采集模块310通过采集设备的设备标识码,来获取设备相应的设备信息,具体根据设备的种类,确定设备信息的采集内容,即:对于 PC,采集MAC、IP和/或UMID;对于移动终端,采集MAC、IMEI、TID和/或手机号。

[0123] 用户信息获取模块320,可以针对操作行为前预设时间段内,获取设备上历史操作行为的全部用户身份信息。该用户信息获取模块320获得相应时间段内历史操作行为的用户信息之后,提供给用户身份解析模块330来解析各用户的身份地域,并统计该时间段内设备用户身份解析地个数,之后就可以将其作为风险评分模型的变量来进行评分。

[0124] 用户身份解析模块330,可以解析各用户身份信息中表征的用户身份解析地,统计该时间段内设备用户身份解析地个数。特别地,用户身份解析模块330以城市为用户身份解析地的区分粒度,其用户身份信息包括用户注册信息中的证件信息,并根据各个用户注册信息中的证件类型和证件号码,获取用户身份解析地并统计设备用户身份解析地个数,具体是根据证件类型的种类来确定用户身份解析地的解析方式,具体而言:若证件类型为中国境内居民身份证,解析每一证件号码的前六位来获取用户身份解析地,并据此统计设备用户身份解析地个数;若证件类型为中国境内非居民身份证或中国境外证件时,推定每一证件类型或每一证件号码对应一个用户身份解析地,并据此统计设备用户身份解析地个数。

[0125] 此外,该用户身份解析模块330可以根据设备信息采集模块310通过设备标识码识别出的设备数量,来确定设备用户身份解析地个数的统计方式,即:若识别出唯一设备,解析统计该设备用户身份解析地个数;若识别出多个设备,解析统计各设备用户身份解析地个数;若未识别出设备,将设备用户身份解析地个数计为0;由此得到的设备用户身份解析地个数,作为盗账风险评定模块340预设评分模型的输入变量,来评定设备的盗账风险等级。

[0126] 盗账风险评定模块340,可以该时间段内设备用户身份解析地个数为预设评分模型的输入变量,评定设备的盗账风险等级。该盗账风险评定模块340进一步结合设备上全部用户个数、当前操作行为用户绑定手机号码个数、当前用户历史操作行为设备个数、当前用户历史操作行为IP地址个数、当前用户的当次操作行为用信息与历史操作行为信息的差异和/或当次操作行为路由特征信息与历史操作行为路由特征信息是否相同,来评定当前操作行为用户的盗账风险等级。在此基础上,该盗账风险评定模块340通过结合设备的盗账风险等级和当前操作行为用户的盗账风险等级来计算账号被盗风险值,在风险值大于预设阈值时识别为账号被盗。

[0127] 以上对本申请账号被盗的风险识别装置进行了描述,其具有较好的盗账风险识别区分能力,且风险识别稳定性较好。在此基础上,本申请相应构建账号被盗的风险防控系统,以下进行简要进行描述。

[0128] 参见图4,出了根据本申请一个实施例账号被盗的风险防控系统。该风险防控系统

适用于用户(图未示出)与操作行为平台(图未示出)的操作行为风险防控,其具有风险识别装置300、盗账上报装置200、风险处理装置100及案件数据库400。

[0129] 风险防控系统各部分的连接关系如图4所示,相应的功能实现过程为:风险识别装置300计算操作行为平台中账号被盗风险值,在风险值大于预设阈值时识别账号被盗;盗账上报装置200在风险识别装置100识别账号被盗时,上报账号被盗消息于风险处理装置400及用户接收设备(如手机)500;风险处理装置100收到账号被盗消息时,冻结用户被盗账号并拦截与被盗账号关联的风险数据;案件数据库400 存放风险处理装置100拦截的风险数据,以供风险处理装置300对风险数据进行核查,以及风险识别装置300对评分模型进行校验。

[0130] 上述风险防控系统中,风险识别装置300请参照图3所示结构,其它装置可以选择习知的设备或应用程序。这种风险防控系统可以及时识别用户账号被盗的风险,一旦确认账号被盗则可以及时进行处理,由此可以更好地为提供安全的网络操作行为环境,因而具有较好的应用价值。

[0131] 在下面的描述中阐述了很多具体细节以便于充分理解本申请。但是本申请能够以很多不同于在此描述的其它方式来实施,本领域技术人员可以在不违背本申请内涵的情况下做类似推广,因此本申请不受下面公开的具体实施例的限制。

[0132] 本申请虽然以较佳实施例公开如上,但其并不是用来限定本申请,任何本领域技术人员在不脱离本申请的精神和范围内,都可以做出可能的变动和修改,因此本申请的保护范围应当以本申请权利要求所界定的范围为准。

[0133] 在一个典型的配置中,计算设备包括一个或多个处理模块(CPU)、输入/输出接口、网络接口和内存。

[0134] 内存可能包括计算机可读介质中的非永久性存储模块,随机存取存储模块 (RAM)和/或非易失性内存等形式,如只读存储模块(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0135] 1、计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何系统或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储模块(SRAM)、动态随机存取存储模块(DRAM)、其他类型的随机存取存储模块(RAM)、只读存储模块(ROM)、电可擦除可编程只读存储模块(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储模块 (CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括非暂存电脑可读媒体 (transitory media),如调制的数据信号和载波。

[0136] 2、本领域技术人员应明白,本申请的实施例可提供为系统、系统或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储模块、CD-ROM、光学存储模块等)上实施的计算机程序产品的形式。

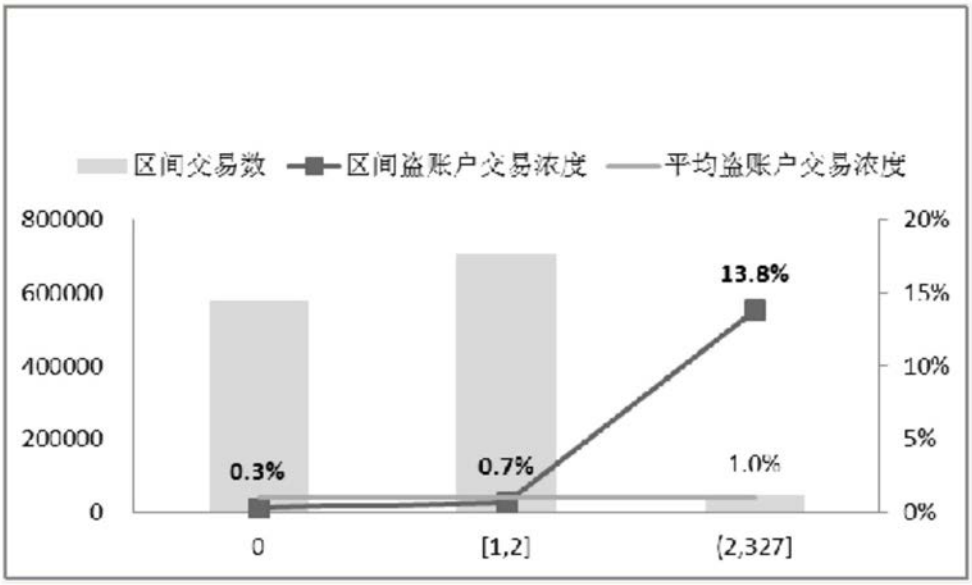

区间交易数
区间盗账户交易浓度
平均盗账户交易浓度
800000
600000
400000
200000
0
20%
15%
10%
5%
0%
0.3%
0.7%
1.0%
13.8%
0
[1,2]
[2,327]

图1

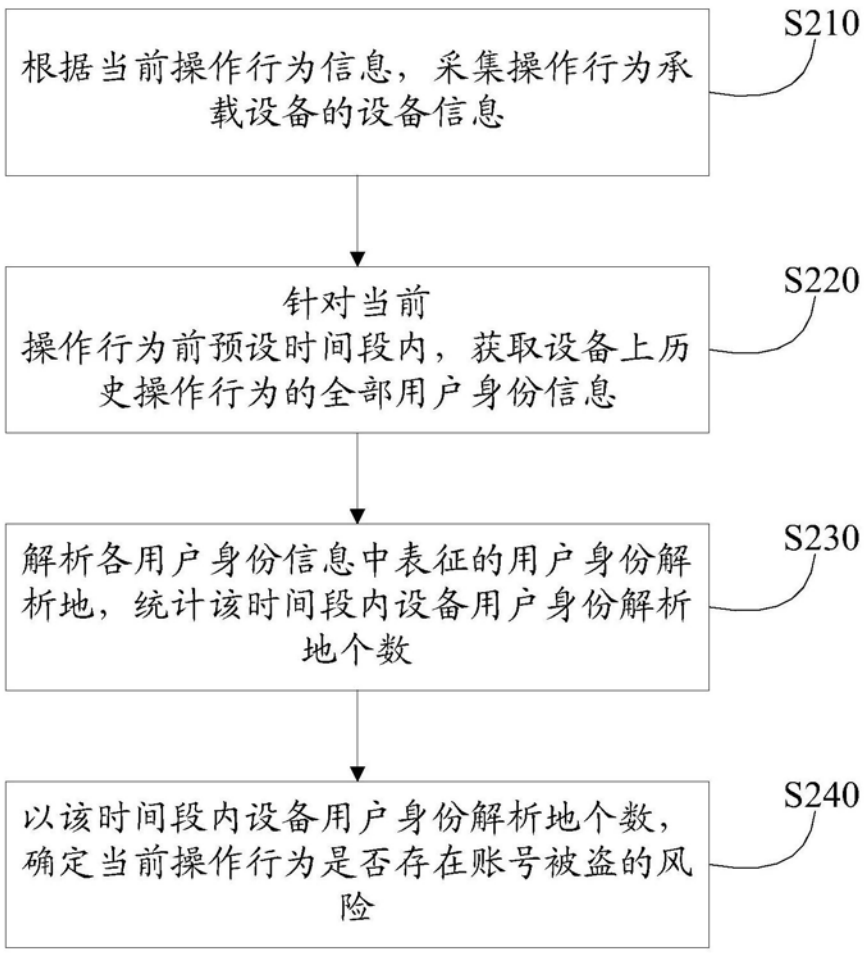

根据当前操作行为信息，采集操作行为承载设备的设备信息
S210
针对当前
操作行为前预设时间段内，获取设备上历史操作行为的全部用户身份信息
S220
解析各用户身份信息中表征的用户身份解析地，统计该时间段内设备用户身份解析地个数
S230
以该时间段内设备用户身份解析地个数，确定当前操作行为是否存在账号被盗的风险
S240

图2

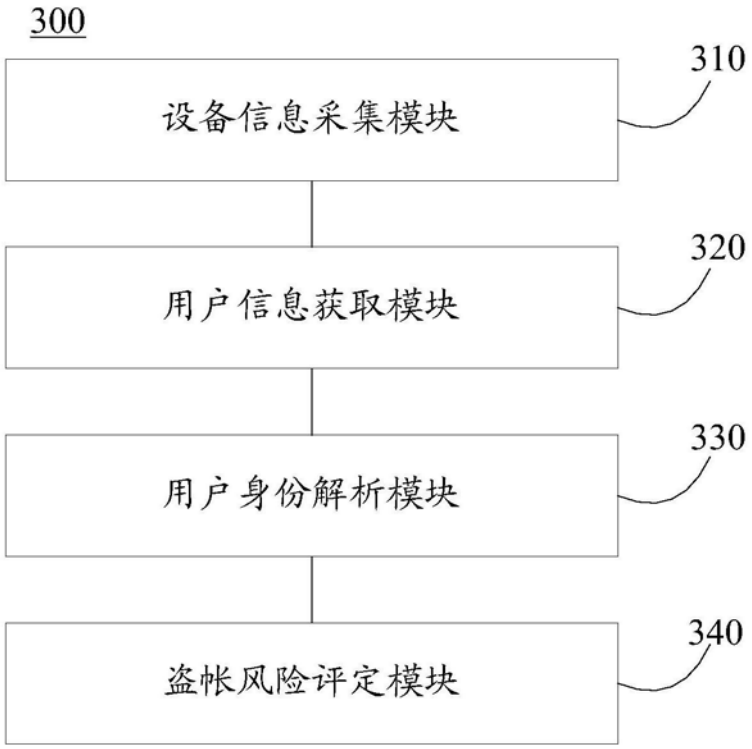

300
310
设备信息采集模块
320
用户信息获取模块
330
用户身份解析模块
340
盗帐风险评定模块

图3

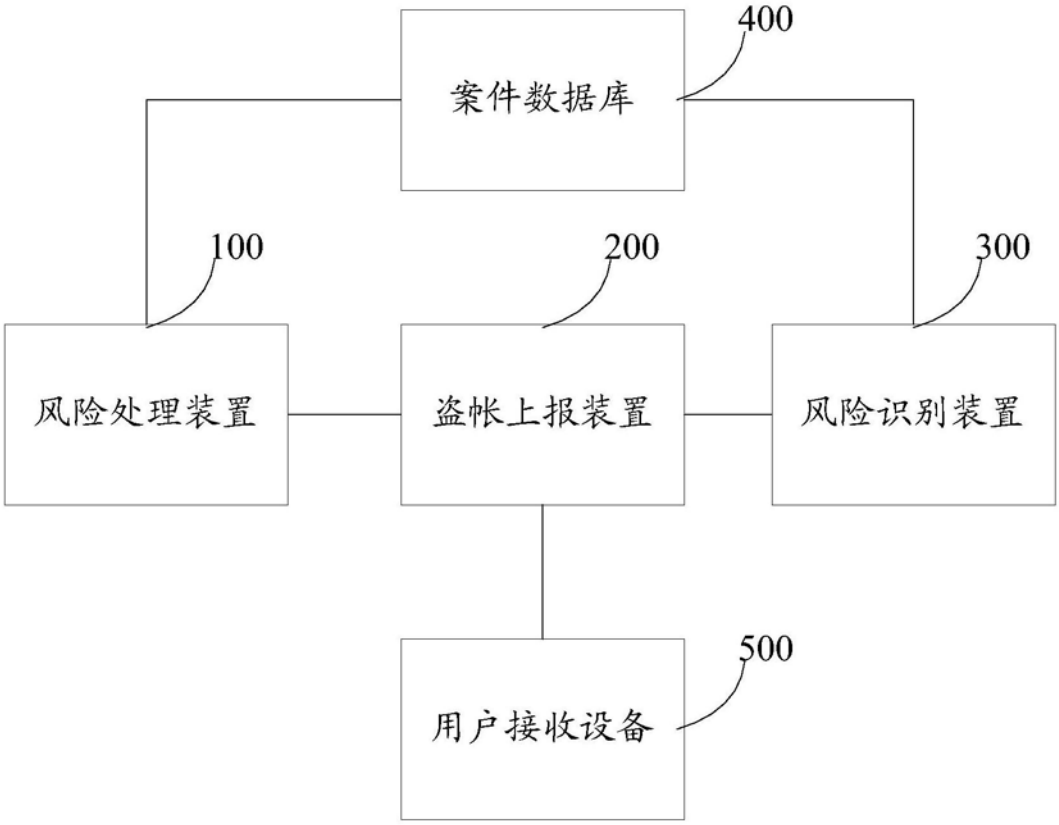

400
案件数据库
100
200
300
风险处理装置
盗帐上报装置
风险识别装置
500
用户接收设备

图4