(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 110457601 B
(45)授权公告日 2023.10.24

(21)申请号 201910755376.6

(22)申请日 2019.08.15

(65)同一申请的已公布的文献号
申请公布号 CN 110457601 A

(43)申请公布日 2019.11.15

(73)专利权人 腾讯科技(武汉)有限公司
地址 430000 湖北省武汉市江夏经济开发区庙山阳光五路特1号

(72)发明人 范小龙

(74)专利代理机构 北京康信知识产权代理有限责任公司 11240
专利代理师 周婷婷

(51)Int.Cl.
G06F 16/9536(2019.01)
G06Q 50/00(2012.01)

(56)对比文件
CN 107992738 A,2018.05.04
CN 109213857 A,2019.01.15
CN 109949172 A,2019.06.28
US 2016277424 A1,2016.09.22
US 8745698 B1,2014.06.03
WO 2018121113 A1,2018.07.05
CN 103853841 A,2014.06.11
曲强等.社交网络异常用户检测技术研究进展.《网络与信息安全学报》.2018,第13-23页.

审查员 何华

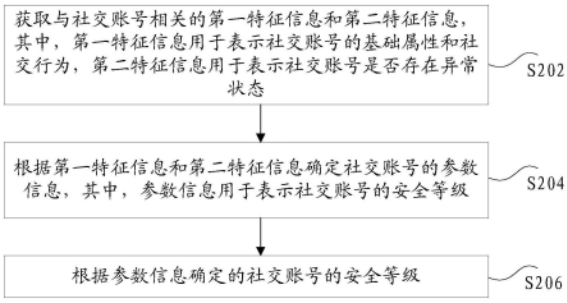
权利要求书2页 说明书12页 附图5页

(54)发明名称

社交账号的识别方法和装置、存储介质及电子装置

(57)摘要

本发明公开了一种社交账号的识别方法和装置、存储介质及电子装置,其中,该方法包括:获取与所述社交账号相关的第一特征信息和第二特征信息,其中,所述第一特征信息用于表示所述社交账号的基础属性和社交行为,所述第二特征信息用于表示所述社交账号是否存在异常状态;根据所述第一特征信息和第二特征信息确定所述社交账号的参数信息,其中,所述参数信息用于表示所述社交账号的安全等级;根据所述参数信息确定的所述社交账号的安全等。通过本发明解决了由于基于RTMP的连麦方案的抗网络抖动能力较差会导致音视频传输不流畅的技术问题。



CN 110457601 B

1.一种社交账号的识别方法,其特征在于,包括:

获取与所述社交账号相关的第一特征信息和第二特征信息,其中,所述第一特征信息用于表示所述社交账号的基础属性和社交行为,所述第二特征信息用于表示所述社交账号是否存在异常状态;

根据所述第一特征信息和第二特征信息确定所述社交账号的参数信息,其中,所述参数信息用于表示所述社交账号的安全等级,所述参数信息为基于所述第一特征信息对应的基础参数数据和所述第二特征信息对应的附加参数信息之间的相关性关系确定得到的,所述相关性关系包括:正相关关系和负相关关系,所述负相关关系用于表示第一附加参数信息能减少所述基础参数信息的安全等级,所述正相关关系用于表示第二附加参数信息能增加所述基础参数信息的安全等级,所述附加参数信息包括:所述第一附加参数信息和所述第二附加参数信息,所述第一附加参数信息为根据所述第二特征信息中用于表示所述社交账号存在异常状态的信息确定得到的,所述第二附加参数信息为根据所述第二特征信息中用于表示所述社交账号未存在异常状态的信息确定得到的;

根据所述参数信息确定的所述社交账号的安全等级。

2.根据权利要求1所述的方法,其特征在于,所述根据所述第一特征信息和第二特征信息确定所述社交账号的参数信息,包括:

根据所述第一特征信息中所述社交账号的基础属性和社交行为确定出与所述社交账号相关的所述基础参数信息;

根据所述第二特征信息中所述社交账号是否存在异常行为确定出与所述社交账号相关的所述附加参数信息;

根据所述基础参数数据和所述附加参数信息的相关性关系确定所述参数信息。

3.根据权利要求2所述的方法,其特征在于,根据所述第一特征信息中所述社交账号的基础属性和社交行为确定出与所述社交账号相关的所述基础参数信息,包括:

根据所述第一特征信息中与所述社交账号相关的基础属性确定出所述社交账号的活跃度;

根据所述第一特征信息中的与所述社交账号的社交行为相关的社交行为确定出所述社交账号的安全系数;

对所述活跃度和所述安全系数进行加权处理得到所述基础参数信息。

4.根据权利要求2所述的方法,其特征在于,所述根据所述第一特征信息和第二特征信息确定所述社交账号参数信息,包括:

根据所述基础参数信息与所述第一附加参数信息的负相关关系,以及所述基础参数与所述第二附加参数信息的正相关关系确定出所述参数信息。

5.一种社交账号的识别装置,其特征在于,包括:

获取模块,用于获取与所述社交账号相关的第一特征信息和第二特征信息,其中,所述第一特征信息用于表示所述社交账号的基础属性和社交行为,所述第二特征信息用于表示所述社交账号是否存在异常状态;

第一确定模块,用于根据所述第一特征信息和第二特征信息确定所述社交账号的参数信息,其中,所述参数信息用于表示所述社交账号的安全等级,所述参数信息为基于所述第一特征信息对应的基础参数数据和所述第二特征信息对应的附加参数信息之间的相关性

关系确定得到的，所述相关性关系包括：正相关关系和负相关关系，所述负相关关系用于表示第一附加参数信息能减少所述基础参数信息的安全等级，所述正相关关系用于表示第二附加参数信息能增加所述基础参数信息的安全等级，所述附加参数信息包括：所述第一附加参数信息和所述第二附加参数信息，所述第一附加参数信息为根据所述第二特征信息中用于表示所述社交账号存在异常状态的信息确定得到的，所述第二附加参数信息为根据所述第二特征信息中用于表示所述社交账号未存在异常状态的信息确定得到的；

第二确定模块，用于根据所述参数信息确定的所述社交账号的安全等级。

6.根据权利要求5所述的装置，其特征在于，所述第一确定模块包括：

第一确定单元，用于根据所述第一特征信息中所述社交账号的基础属性和社交行为确定出与所述社交账号相关的所述基础参数信息；

第二确定单元，用于根据所述第二特征信息中所述社交账号是否存在异常行为确定出与所述社交账号相关的所述附加参数信息；

第三确定单元，用于根据所述基础参数数据和所述附加参数信息的相关性关系确定所述参数信息。

7.一种存储介质，其特征在于，所述存储介质中存储有计算机程序，其中，所述计算机程序被设置为运行时执行所述权利要求1至4任一项中所述的方法。

8.一种电子装置，包括存储器和处理器，其特征在于，所述存储器中存储有计算机程序，所述处理器被设置为通过所述计算机程序执行所述权利要求1至4任一项中所述的方法。

社交账号的识别方法和装置、存储介质及电子装置

技术领域

[0001] 本发明涉及计算机领域，具体而言，涉及一种社交账号的识别方法和装置、存储介质及电子装置。

背景技术

[0002] 现有的社交账号安全等级测评上，大部分是采用特定的安全系数来计算，比如检测是否有身份认证，是否有绑定手机号，是否有木马等等，但是该方式在所有安全指数都有的情况下安全分数会比较高，但依然很难有效区分出黑产和正常社交账号。

[0003] 此外，当前的账号异常检测大部分只采用了恶意分类及群体挖掘等方法实现，只是一个单向的恶意账号识别，不能有效对账号的安全信用做出不同等级的划分。而当前的信用评分主要是针对用户身份的信用评分，其实现路径为：根据客户的信用历史及综合资料，主要是交易及历史消费等特征，利用一定的信用评分模型，得到不同等级的信用分数。

[0004] 但是，社交账号的安全指数计算容易被黑产模拟突破，即身份资料信息等很容易被黑产获取到，并且进行模拟出高分数的安全分数。另外，当前恶意账号检测结果过于单一，无法比较准确的测评出是否真的是恶意账号或正常账号。

[0005] 针对现有技术中的上述问题，目前尚未提出有效的解决方案。

发明内容

[0006] 本发明实施例提供一种社交账号的识别方法和装置、存储介质及电子装置，以至少解决由于基于RTMP的连麦方案的抗网络抖动能力较差会导致音视频传输不流畅的技术问题。

[0007] 根据本发明实施例的一个方面，提供了一种社交账号的识别方法，包括：获取与所述社交账号相关的第一特征信息和第二特征信息，其中，所述第一特征信息用于表示所述社交账号的基础属性和社交行为，所述第二特征信息用于表示所述社交账号是否存在异常状态；根据所述第一特征信息和第二特征信息确定所述社交账号的参数信息，其中，所述参数信息用于表示所述社交账号的安全等级；根据所述参数信息确定的所述社交账号的安全等级。

[0008] 根据本发明实施例的另一方面，还提供了一种社交账号的识别装置，包括：获取模块，用于获取与所述社交账号相关的第一特征信息和第二特征信息，其中，所述第一特征信息用于表示所述社交账号的基础属性和社交行为，所述第二特征信息用于表示所述社交账号是否存在异常状态；第一确定模块，用于根据所述第一特征信息和第二特征信息确定所述社交账号的参数信息，其中，所述参数信息用于表示所述社交账号的安全等级；第二确定模块，用于根据所述参数信息确定的所述社交账号的安全等级。

[0009] 根据本发明实施例的又一方面，还提供了一种存储介质，该存储介质中存储有计算机程序，其中，该计算机程序被设置为运行时执行上述方法。

[0010] 根据本发明实施例的又一方面，还提供了一种电子装置，包括存储器、处理器及存

储在存储器上并可在处理器上运行的计算机程序，其中，上述处理器通过计算机程序执行上述的方法。

[0011] 在本发明实施例中，采用根据第一特征信息和第二特征信息确定社交账号的用于表示社交账号安全等级的参数信息，而该第一特征信息用于表示社交账号的基础属性和社交行为，第二特征信息用于表示社交账号是否存在异常状态；也即在本实施例中采用了多个维度对社交账号的安全等级进行识别，而不只是单一的从某几个方面对社交账号进行识别，从而解决了相关技术对于社交账号的安全等级的测评维度比较单一导致测评结果并不能反应真实社交账号状态的技术问题，使得对于社交账号的安全等级的识别更加精准。

附图说明

[0012] 此处所说明的附图用来提供对本发明的进一步理解，构成本申请的一部分，本发明的示意性实施例及其说明用于解释本发明，并不构成对本发明的不当限定。在附图中：

[0013] 图1是根据本发明实施例的一种社交账号的识别方法的应用场景的示意图；

[0014] 图2是根据本发明实施例的一种可选的社交账号的识别方法的流程示意图；

[0015] 图3是根据本发明实施例的另一种可选的社交账号的识别方法的流程示意图；

[0016] 图4是根据本发明实施例的一种可选的社交账号风险的识别的示意图；

[0017] 图5是根据本发明实施例的社交账号信用评分结果的示意图；

[0018] 图6是根据本发明实施例的一种可选的社交账号的识别装置的结构示意图；

[0019] 图7是根据本发明实施例的又一种可选的社交账号的识别装置的结构示意图；

[0020] 图8是根据本发明实施例的一种可选的电子装置的结构示意图。

具体实施方式

[0021] 为了使本技术领域的人员更好地理解本发明方案，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分的实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本发明保护的范围。

[0022] 需要说明的是，本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0023] 根据本发明实施例的一个方面，提供了一种社交账号的识别方法。可选地，上述社交账号的识别方法可以但不限于应用于如图1所示的应用环境中。如图1所示，在终端102内下载有一款用于社交的应用，用户可以通过登录自己在该社交应用中的账号，并将社交数据通过网络104发送到服务器106，服务器106将社交数据发送到其他终端108中的社交应用。

[0024] 可选地，在本实施例中，上述终端可以包括但不限于以下至少之一：手机、平板电脑等。上述网络可以包括但不限于：有线网络，无线网络，其中，该有线网络包括：局域网、城域网和广域网，该无线网络包括：蓝牙、WIFI及其他实现无线通信的网络。上述只是一种示例，本实施例对此不做任何限定。

[0025] 可选地，在本实施例中，作为一种可选的实施方式，如图2所示，上述社交账号的识别方法可以包括：

[0026] 步骤S202，获取与社交账号相关的第一特征信息和第二特征信息，其中，第一特征信息用于表示社交账号的基础属性和社交行为，第二特征信息用于表示社交账号是否存在异常状态；

[0027] 步骤S204，根据第一特征信息和第二特征信息确定社交账号的参数信息，其中，参数信息用于表示社交账号的安全等级；

[0028] 步骤S206，根据参数信息确定的社交账号的安全等级。

[0029] 可选地，在本实施例中，社交账号的基础属性至少包括：社交账号的用户信息、登录社交账号的频率以及每次登录社交账号的时长，登录社交账号的地点及其时间，是否有真实身份认证等等，与社交账号日常使用相关的信息均属于该社交账号的基础属性；而社交账号的社交行为至少包括：加好友的行为，发送消息/群发消息行为，社交账号发消息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为等等。

[0030] 可选地，本实施例中的社交账号的异常状态主要通过社交账号的以下信息来判别：看社交账号是否是自动机操作的虚假账号(养号，通过团伙关联分析及恶意分类模型实现)，是否有作恶风险(通过风险预测模型实现，主要采用监督的二分类模型实现)，是否有违规操作记录(是否被多个场景打击并封号)，关联群体识别(主要看好友团伙之间是否有恶意群体存在)等等。

[0031] 通过上述步骤S202至步骤S206，采用根据第一特征信息和第二特征信息确定社交账号的用于表示社交账号安全等级的参数信息，而该第一特征信息用于表示社交账号的基础属性和社交行为，第二特征信息用于表示社交账号是否存在异常状态；也即在本实施例中采用了多个维度对社交账号的安全等级进行识别，而不只是单一的从某几个方面对社交账号进行识别，从而解决了相关技术对于社交账号的安全等级的测评维度比较单一导致测评结果并不能反应真实社交账号状态的技术问题，使得对于社交账号的安全等级的识别更加精准。

[0032] 可选地，在本实施例中，对于本实施例中步骤S204中涉及到的根据第一特征信息和第二特征信息确定社交账号的参数信息的方式，进一步可以包括：

[0033] 步骤S204-11，根据第一特征信息中社交账号的基础属性和社交行为确定出与社交账号相关的基础参数信息；

[0034] 可选地，在实施例中，该步骤S204-11进一步可以包括：

[0035] S1，根据第一特征信息中与社交账号相关的基础属性确定出社交账号的活跃度；

[0036] S2，根据第一特征信息中的与社交账号的社交行为相关的社交行为确定出社交账号的安全系数；

[0037] S3，对活跃度和安全系数进行加权处理得到基础参数信息。

[0038] 需要说明的是，本实施例中的社交账号的基础属性至少包括：社交账号的用户信息、登录社交账号的频率以及每次登录社交账号的时长，登录社交账号的地点及其时间，是否有真实身份认证等等，与社交账号日常使用相关的信息均属于该社交账号的基础属性；因此，社交账号的活跃度是结合了上述基础属性中具体属性来综合进行计算的。

[0039] 例如，以社交账号A一周内的使用情况为例，首先获取该社交账号A的用户信息，用户信息就包括性别，年龄，账户昵称，个性签名等等，再统计一周内登陆的时长和频率，以及登录的地点；如在周一至周五该社交账号A的用户在地点1出差，出差期间工作时间只是将社交账号登陆，并未使用，在周末后，用户回到常住地点2，并在周末通过社交账号A与好友进行交流，通过社交账号A浏览相关实时新闻和文章，并对其社交账号中的好友状态进行评论，在线使用时长为4个小时；上述社交账号A在该一周内的基础属性均是要被统计的。当然，上述仅仅是举例说明，也可以统计社交账号A一个月内的使用情况，而且基础属性还可以包括是否存在真实身份认证等其他基础属性，其中，真实身份认证的方式包括：是否有身份证实名认证，认证的时长，是否有绑定手机等。

[0040] 此外，社交账号的社交行为至少包括：加好友的行为，发送消息/群发消息行为，社交账号发消息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为等等。

[0041] 对于社交账号的社交行为还是以社交账号A在一周内的使用情况为例，则需要对该社交账号A在该一周内的加好友的行为，发送消息/群发消息行为，社交账号发消息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为进行统计，例如在这一周内社交账号A加了5个好友，并删除了一个长期不联系的好友，此外接收到了10个账号发送的垃圾消息，对20个账号好友的状态进行了评论等等。

[0042] 在对社交账号A进行统计之后，由于统计数值数据，范围不可定，所以需要进行特征映射归一到0-1之间，比如一个账号登录次数为1000次，通过映射函数直接转换为0.99左右。其中，该映射函数在实施例中优选为sigmd，当然也可以是其他应用函数。

[0043] 需要说明的是，本可选实施方式中的基础参数信息的表现形式优选为分值，即通过分值可以直观的确定该基础参数信息所对应的安全等级，当然分值越高所对应的安全等级也就越高。因此，在根据相关映射函数计算出统一的基础数值，再通过融合函数综合计算出基础信用分。

[0044] 在本实施例中的映射函数可以是：

[0045] $$\mathcal{F}(x) = sigmd(x) \quad ---①$$

[0046] $$\mathcal{F}(x) = (a + x)/(b + x) \quad ---②$$

[0047] 需要说明的是，两种函数在不同的数值类型中可以做不同选择。

[0048] 可选地，本实施例中的加权融合函数为：

[0049] $$S\ (x\ ) = \sum_{k\ =0}^{n} a\ _{k}\ F\ (x\ _{k}\ )$$

[0050] 步骤S204-12，根据第二特征信息中社交账号是否存在异常行为确定出与社交账

号相关的附加参数信息；

[0051] 可选地，在本实施例中步骤S204-12进一步可以包括：

[0052] 步骤S1，根据第二特征信息中用于表示社交账号存在异常状态的信息确定出与社交账号的信第一附加参数信息；

[0053] 步骤S2，根据第二特征信息中用于表示社交账号未存在异常状态的信息确定出与社交账号相关的第二附加参数信息，其中，附加参数信息包括：第一附加参数信息和第二附加参数信息。

[0054] 基于上述本实施例中的社交账号的异常状态主要通过社交账号的以下信息来判别：看社交账号是否是自动机操作的虚假账号(养号，通过团伙关联分析及恶意分类模型实现)，是否有作恶风险(通过风险预测模型实现，主要采用监督的二分类模型实现)，是否有违规操作记录(是否被多个场景打击并封号)，关联群体识别(主要看好友团伙之间是否有恶意群体存在)等等。

[0055] 对于社交账号出现了上述异常状态后是需要对该社交账号进行惩罚的，如果在实施例的具体应用场景中依然也是以分值来表现该第一附加参数信息和第二附加参数信息的情况下，则出现一次上述异常状态则该第一附加参数信息的分值将会根据不同的异常类型的权重加上相应的分值，反之则在第二附加参数的分值上加上相应的分值。

[0056] 还是以上述社交账号A在一周内的使用情况为例，除了统计基础属性中的该社交账号A的用户信息，用户信息就包括性别，年龄，账户昵称，个性签名等等，再统计一周内登陆的时长和频率，以及登录的地点；如在周一至周五该社交账号A的用户在地点1出差，出差期间工作时间只是将社交账号登陆，并未使用，在周末后，用户回到常住地点2，并在周末通过社交账号A与好友进行交流，通过社交账号A浏览相关实时新闻和文章，并对其社交账号中的好友状态进行评论，在线使用时长为4个小时；上述社交账号A在该一周内的基础属性均是要被统计的；以及社交账号的社交行为还是以社交账号A在一周内的使用情况为例，则需要对该社交账号A在该一周内的加好友的行为，发送消息/群发消息行为，社交账号发消息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为进行统计，例如在这一周内社交账号A加了5个好友，并删除了一个长期不联系的好友，此外接收到了10个账号发送的垃圾消息，对20个账号好友的状态进行了评论等等外，对于该一周内出现的异常状态进行统计，例如，是否被盗号存在有作恶风险，或者与好友交流过程存在一些有损国家形象或与国家政治导向相悖的言论，或者发表一些低俗的言论或图片等违规操作记录等等。

[0057] 步骤S204-13，根据基础参数数据和附加参数信息的相关性关系确定参数信息。

[0058] 需要说明的是，本实施例中的相关性关系包括：正相关关系和负相关关系，其中，负相关关系用于表示第一附加参数信息能减少基础参数信息的安全等级，正相关关系用于表示第二附加参数信息能增加基础参数信息的安全等级。

[0059] 基于此，本实施例步骤S206中根据第一特征信息和第二特征信息确定社交账号的参数信息的方式，可以是：根据基础参数信息与第一附加参数信息的负相关关系，以及基础参数与第二附加参数信息的正相关关系确定出参数信息。

[0060] 下面结合一具体实施方式对本实施例进行举例说明；

[0061] 在本具体实施方式中提供了一种社交账号安全信用分计算的方法，其主要整合海

量的社交、安全、账号属性、交易流水、关联网络等海量大数据，结合先进的智能分析技术，建设一体化的账号风控信用体系。其中，该方法采用基于基础账号社交活跃属性评分卡及恶意判别的机器学习分类方法等多模型融合，综合实现对每个社交账号进行安全信用分计算。

[0062] 需要说明的是，安全信用分为上述本实施例中参数信息的一种表现形式，当然在其他应用场景中也可以由其他形式来来表现该参数信息。

[0063] 在本具体实施方式总，该社交账号安全信用分计算方法包括：基础数据输入，基础评分，双向判别模型融合，综合校验，以及输出分析五个部分组成，总体流程图，如图3所示，

[0064] (1)基础数据输入：基础账号属性/安全/行为/反馈等数据等收集分析及特征挖掘提取；

[0065] 其中，基础数据输入整合了注册，登录，社交行为，账号属性，用户反馈等海量数据，深度提取多维度的有效特征，例如包括：活跃天数，作恶次数及时间，常用操作习惯，社交活跃度，使用天数，登录天数等多个基础特征维度，特征数量根据场景还可以不断扩充。

[0066] (2)基础评分：进账号在业务场景的活跃程度，包括登录/消息操作/社交属性等综合来计算的，还有融合账号的安全属性数据，来进行整合加权计算基础的账号安全信用分；

[0067] 其中，基础评分部分主要针对多种基础特征数据，总结分析后，建立依据账号活跃度+安全系数的多维度的基础评分模型，主要提取多种活跃度和安全等级；

[0068] 活跃度计算：结合账号社交数据(发消息/加好友/社群等其他活跃行为)及登录操作(登录频次/在线时长)等综合来计算。

[0069] 安全等级：是指在个业务场景中的违规操作记录统计以及身份安全认证数据，业务违规包括：发垃圾消息次数，骚扰添加好友次数，被好友删除的次数等等，身份安全认证数据主要包括：是否有身份证实名认证，认证的时长，是否有绑定手机等。

[0070] 需要说明的是，这些都是统计数值数据，范围不可定，所以需要进行特征映射归一到0-1之间，比如一个账号登录次数为1000次，通过sigmd函数直接转换为0.99左右，这就是映射方法。再根据相关映射函数计算出统一的基础数值，再通过融合函数综合计算出基础信用分。

[0071] (3)双向判别模型：融合了核心的账号风险模型，以及多维度的安全大数据，综合对分数进行修正。

[0072] 其中，双向判别模型融合了核心的账号风险模型，以及多维度的安全大数据，综合对分数进行修正，其中，分为信用惩罚网络和信用奖励网络。

[0073] 信用惩罚：主要看号码是否是自动机操作的虚假账号(养号，通过团伙关联分析及恶意分类模型实现)，是否有作恶风险(通过风险预测模型实现，主要采用监督的二分类模型实现)，是否有违规操作记录(是否被多个场景打击并封号)，关联群体识别(主要看好友团伙之间是否有恶意群体存在)。通过这些子纬度不同的风险评测机械能恶意加权处理后，做信用惩罚，降低信用分。

[0074] 信用奖励：确定在没有上述异常风险指标后，结合身份安全认证实数进行相关优质等级判别，在原始基础分数上进行加分处理。优质定义主要看好友之间是否都没有违规操作记录，边会对整个小团体定义为优质团伙。

[0075] 可选地，信用惩罚中最核心的是账号风险的识别，如图4所示，主要根据上报的这

些数据(登录流水/注册流水/发消息流水/操作设备及网络环境流水),提取出相关的账号特征,可以简述为建立出多维度的特征图谱,然后将特征图谱输入到恶意的二分类模型中,进行相关恶意分类预测。恶意二分类模型不限定于具体单个类别,可以是神经网络/XGBOOST/RF等。

[0076] 其中,包括:账号风险概率预测和养号群体的挖掘,异常账号的挖掘模型见图2所示,其是一个独立的子模型,根据账号画像数据,建设数据特征挖掘处理,经过多种分析模型,包括账号风险恶意分类(GBDT/RF/XGBoost),及图关联分析群体养号团伙挖掘,最终输出账号风险分(0-100)以及养号等级。其中,输出的账号风险分范围不一定限制在0-100,可以跟进场景进行调节到指定范围。

[0077] 账号风险挖掘模型最终会输出,账号的恶意分数,以及养号的风险等级;最后跟进账号的高风险概率以及高养号等级融合对信用分进行惩罚处理。对于安全属性并且预测的风险指数较低的账号,会做不同程度的信用奖励,提升账号安全信用分。

[0078] (4)综合校验,对优质和恶意分数段的号码进行确判,进一步提升准确率和可解释性,还有结合不同的场景实现对不同基本的号码做信用分的加权处理。

[0079] 其中,综合校验是对优质和恶意分数段的号码进行确判,寻找关键特征点做校验确判,比如优质账号中:无养号/低风险/高安全属性,高危号码中:低活跃/多次作恶记录等,进一步提升准确率和可解释性;还会结合不同的场景实现对不同基本的号码做信用分的加权处理,比如企业/公共号码进行一定的奖励和惩罚,以便在相关场景下更精准。

[0080] (5)输出分析:输出不同层级的账号安全信用分,可供给不同场景应用。

[0081] 其中,最终输出不同层级的账号安全信用分,可供给不同场景应用。

[0082] 如图5所示,是设备应用分是在真实场景下的K-S值的评测效果数据,可以看到整体对恶意/正常设备的综合差值大于60%。能有效的区分出恶意和正常设备。可选地,输出的账号信用分及等级范围不一定限制在0-100/5个等级,可以跟进场景进行调节到指定范围及等级需求。

[0083] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明并不受所描述的动作顺序的限制,因为依据本发明,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本发明所必须的。

[0084] 根据本发明实施例的另一个方面,还提供了一种社交账号的识别装置,如图6所示,该装置包括:

[0085] (1)获取模块602,用于获取与社交账号相关的第一特征信息和第二特征信息,其中,第一特征信息用于表示社交账号的基础属性和社交行为,第二特征信息用于表示社交账号是否存在异常状态;

[0086] (2)第一确定模块604,用于根据第一特征信息和第二特征信息确定社交账号的参数信息,其中,参数信息用于表示社交账号的安全等级;

[0087] (3)第二确定模块606,用于根据参数信息确定的社交账号的安全等级。

[0088] 可选地,在本实施例中,社交账号的基础属性至少包括:社交账号的用户信息、登录社交账号的频率以及每次登录社交账号的时长,登录社交账号的地点及其时间,是否有

真实身份认证等等，与社交账号日常使用相关的信息均属于该社交账号的基础属性；而社交账号的社交行为至少包括：加好友的行为，发送消息/群发消息行为，社交账号发消息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为等等。

[0089] 可选地，本实施例中的社交账号的异常状态主要通过社交账号的以下信息来判别：看社交账号是否是自动机操作的虚假账号(养号，通过团伙关联分析及恶意分类模型实现)，是否有作恶风险(通过风险预测模型实现，主要采用监督的二分类模型实现)，是否有违规操作记录(是否被多个场景打击并封号)，关联群体识别(主要看好友团伙之间是否有恶意群体存在)等等。

[0090] 可选地，在本实施例中，如图7所示，第一确定模块604包括：

[0091] (1)第一确定单元702，用于根据第一特征信息中社交账号的基础属性和社交行为确定出与社交账号相关的基础参数信息；

[0092] 其中，该第一确定单元702进一步可以包括：第一确定子单元，用于根据第一特征信息中与社交账号相关的基础属性确定出社交账号的活跃度；第二确定子单元，用于根据第一特征信息中的与社交账号的社交行为相关的社交行为确定出社交账号的安全系数；处理子单元，用于对活跃度和安全系数进行加权处理得到基础参数信息。

[0093] 需要说明的是，本实施例中的社交账号的基础属性至少包括：社交账号的用户信息、登录社交账号的频率以及每次登录社交账号的时长，登录社交账号的地点及其时间，是否有真实身份认证等等，与社交账号日常使用相关的信息均属于该社交账号的基础属性；因此，社交账号的活跃度是结合了上述基础属性中具体属性来综合进行计算的。

[0094] 例如，以社交账号A一周内的使用情况为例，首先获取该社交账号A的用户信息，用户信息就包括性别，年龄，账户昵称，个性签名等等，再统计一周内登陆的时长和频率，以及登录的地点；如在周一至周五该社交账号A的用户在地点1出差，出差期间工作时间只是将社交账号登陆，并未使用，在周末后，用户回到常住地点2，并在周末通过社交账号A与好友进行交流，通过社交账号A浏览相关实时新闻和文章，并对其社交账号中的好友状态进行评论，在线使用时长为4个小时；上述社交账号A在该一周内的基础属性均是要被统计的。当然，上述仅仅是举例说明，也可以统计社交账号A一个月内的使用情况，而且基础属性还可以包括是否存在真实身份认证等其他基础属性，其中，真实身份认证的方式包括：是否有身份证实名认证，认证的时长，是否有绑定手机等。

[0095] 此外，社交账号的社交行为至少包括：加好友的行为，发送消息/群发消息行为，社交账号发消息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为等等。

[0096] 对于社交账号的社交行为还是以社交账号A在一周内的使用情况为例，则需要对该社交账号A在该一周内的加好友的行为，发送消息/群发消息行为，社交账号发消息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为进行统计，例如在这一周内社交账号A加了5个好友，并删除了一个长期不联系的好友，此外接收到了10个账号发送的垃圾消息，对20个账号好友的状态进行了评论等等。

[0097] 在对社交账号A进行统计之后，由于统计数值数据，范围不可定，所以需要进行特

征映射归一到0-1之间，比如一个账号登录次数为1000次，通过映射函数直接转换为0.99左右。其中，该映射函数在实施例中优选为sigmd，当然也可以是其他应用函数。

[0098] 需要说明的是，本可选实施方式中的基础参数信息的表现形式优选为分值，即通过分值可以直观的确定该基础参数信息所对应的安全等级，当然分值越高所对应的安全等级也就越高。因此，在根据相关映射函数计算出统一的基础数值，再通过融合函数综合计算出基础信用分。

[0099] 在本实施例中的映射函数可以是：

[0100] $$\mathcal{F}(x) = sigmd(x) \quad ---①$$

[0101] $$\mathcal{F}(x) = (a + x)/(b + x) \quad ---②$$

[0102] 需要说明的是，两种函数在不同的数值类型中可以做不同选择。

[0103] 可选地，本实施例中的加权融合函数为：

[0104] $$S\ (x\ ) = \sum_{k\ =0}^{n} a\ _{k}\ F\ (x\ _{k}\ )$$

[0105] (2)第二确定单元704，用于根据第二特征信息中社交账号是否存在异常行为确定出与社交账号相关的附加参数信息；

[0106] 其中，该第二确定单元包括：第三确定子单元，用于根据第二特征信息中用于表示社交账号存在异常状态的信息确定出与社交账号的信第一附加参数信息；第四确定子单元，用于根据第二特征信息中用于表示社交账号未存在异常状态的信息确定出与社交账号相关的第二附加参数信息，其中，附加参数信息包括：第一附加参数信息和第二附加参数信息。

[0107] 基于上述本实施例中的社交账号的异常状态主要通过社交账号的以下信息来判别：看社交账号是否是自动机操作的虚假账号(养号，通过团伙关联分析及恶意分类模型实现)，是否有作恶风险(通过风险预测模型实现，主要采用监督的二分类模型实现)，是否有违规操作记录(是否被多个场景打击并封号)，关联群体识别(主要看好友团伙之间是否有恶意群体存在)等等。

[0108] 对于社交账号出现了上述异常状态后是需要对该社交账号进行惩罚的，如果在实施例的具体应用场景中依然也是以分值来表现该第一附加参数信息和第二附加参数信息的情况下，则出现一次上述异常状态则该第一附加参数信息的分值将会根据不同的异常类型的权重加上相应的分值，反之则在第二附加参数的分值上加上相应的分值。

[0109] 还是以上述社交账号A在一周内的使用情况为例，除了统计基础属性中的该社交账号A的用户信息，用户信息就包括性别，年龄，账户昵称，个性签名等等，再统计一周内登陆的时长和频率，以及登录的地点；如在周一至周五该社交账号A的用户在地点1出差，出差期间工作时间只是将社交账号登陆，并未使用，在周末后，用户回到常住地点2，并在周末通过社交账号A与好友进行交流，通过社交账号A浏览相关实时新闻和文章，并对其社交账号中的好友状态进行评论，在线使用时长为4个小时；上述社交账号A在该一周内的基础属性均是要被统计的；以及社交账号的社交行为还是以社交账号A在一周内的使用情况为例，则需要对该社交账号A在该一周内的加好友的行为，发送消息/群发消息行为，社交账号发消

息被评论行为，评论好友账号的行为，发垃圾消息的行为，接收垃圾消息的行为，骚扰添加好友的行为，被好友删除的行为进行统计，例如在这一周内社交账号A加了5个好友，并删除了一个长期不联系的好友，此外接收到了10个账号发送的垃圾消息，对20个账号好友的状态进行了评论等等外，对于该一周内出现的异常状态进行统计，例如，是否被盗号存在有作恶风险，或者与好友交流过程存在一些有损国家形象或与国家政治导向相悖的言论，或者发表一些低俗的言论或图片等违规操作记录等等。

[0110] (3)第三确定单元706，用于根据基础参数数据和附加参数信息的相关性关系确定参数信息。

[0111] 需要说明的是，本实施例中的相关性关系包括：正相关关系和负相关关系，其中，负相关关系用于表示第一附加参数信息能减少基础参数信息的安全等级，正相关关系用于表示第二附加参数信息能增加基础参数信息的安全等级。

[0112] 基于上述相关性关系，本实施例中的第二确定模块606，还用于根据基础参数信息与第一附加参数信息的负相关关系，以及基础参数与第二附加参数信息的正相关关系确定出参数信息。

[0113] 通过上述本实施例中的社交账号的识别装置，采用根据第一特征信息和第二特征信息确定社交账号的用于表示社交账号安全等级的参数信息，而该第一特征信息用于表示社交账号的基础属性和社交行为，第二特征信息用于表示社交账号是否存在异常状态；也即在本实施例中采用了多个维度对社交账号的安全等级进行识别，而不只是单一的从某几个方面对社交账号进行识别，从而解决了相关技术对于社交账号的安全等级的测评维度比较单一导致测评结果并不能反应真实社交账号状态的技术问题，使得对于社交账号的安全等级的识别更加精准。

[0114] 根据本发明的实施例的又一方面，还提供了一种存储介质，该存储介质中存储有计算机程序，其中，该计算机程序被设置为运行时执行上述任一项方法实施例中的步骤。

[0115] 可选地，在本实施例中，上述存储介质可以被设置为存储用于执行以下步骤的计算机程序：

[0116] S1，获取与社交账号相关的第一特征信息和第二特征信息，其中，第一特征信息用于表示社交账号的基础属性和社交行为，第二特征信息用于表示社交账号是否存在异常状态；

[0117] S2，根据第一特征信息和第二特征信息确定社交账号的参数信息，其中，参数信息用于表示社交账号的安全等级；

[0118] S3，根据参数信息确定的社交账号的安全等级。

[0119] 可选地，在本实施例中，上述存储介质可以被设置为存储用于执行以下步骤的计算机程序：

[0120] S1，根据第一特征信息中社交账号的基础属性和社交行为确定出与社交账号相关的基础参数信息；

[0121] S2，根据第二特征信息中社交账号是否存在异常行为确定出与社交账号相关的附加参数信息；

[0122] S3，根据基础参数数据和附加参数信息的相关性关系确定参数信息。

[0123] 可选地，在本实施例中，上述存储介质可以被设置为存储用于执行以下步骤的计

算机程序：

[0124] S1,根据第一特征信息中与社交账号相关的基础属性确定出社交账号的活跃度；

[0125] S2,根据第一特征信息中的与社交账号的社交行为相关的社交行为确定出社交账号的安全系数；

[0126] S3,对活跃度和安全系数进行加权处理得到基础参数信息。

[0127] 可选地,在本实施例中,上述存储介质可以被设置为存储用于执行以下步骤的计算机程序：

[0128] S1,根据第二特征信息中用于表示社交账号存在异常状态的信息确定出与社交账号的信第一附加参数信息；

[0129] S2,根据第二特征信息中用于表示社交账号未存在异常状态的信息确定出与社交账号相关的第二附加参数信息,其中,附加参数信息包括:第一附加参数信息和第二附加参数信息。

[0130] 可选地,在本实施例中,上述存储介质可以被设置为存储用于执行以下步骤的计算机程序：

[0131] S1,根据基础参数信息与第一附加参数信息的负相关关系,以及基础参数与第二附加参数信息的正相关关系确定出参数信息,其中,相关性关系包括:正相关关系和负相关关系,其中,负相关关系用于表示第一附加参数信息能减少基础参数信息的安全等级,正相关关系用于表示第二附加参数信息能增加基础参数信息的安全等级。

[0132] 可选地,在本实施例中,本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令终端设备相关的硬件来完成,该程序可以存储于一计算机可读存储介质中,存储介质可以包括:闪存盘、只读存储器(Read-Only Memory,ROM)、随机存取器(Random Access Memory,RAM)、磁盘或光盘等。

[0133] 根据本发明实施例的又一个方面,还提供了一种用于实施上述社交账号的识别方法的电子装置,如图8所示,该电子装置包括:处理器802、存储器804、显示器806、用户接口808、传输装置810等。该存储器中存储有计算机程序,该处理器被设置为通过计算机程序执行上述任一项方法实施例中的步骤。

[0134] 可选地,在本实施例中,上述电子装置可以位于计算机网络的多个网络设备中的至少一个网络设备。

[0135] 可选地,在本实施例中,上述处理器可以被设置为通过计算机程序执行以下步骤：

[0136] S1,获取与社交账号相关的第一特征信息和第二特征信息,其中,第一特征信息用于表示社交账号的基础属性和社交行为,第二特征信息用于表示社交账号是否存在异常状态；

[0137] S2,根据第一特征信息和第二特征信息确定社交账号的参数信息,其中,参数信息用于表示社交账号的安全等级；

[0138] S3,根据参数信息确定的社交账号的安全等级。

[0139] 可选地,本领域普通技术人员可以理解,图8所示的结构仅为示意,电子装置也可以是智能手机(如Android手机、iOS手机等)、平板电脑、掌上电脑以及移动互联网设备(Mobile Internet Devices,MID)、PAD等终端设备。图8其并不对上述电子装置的结构造成限定。例如,电子装置还可包括比图8中所示更多或者更少的组件(如网络接口等),或者具

有与图8所示不同的配置。

[0140] 其中，存储器804可用于存储软件程序以及模块，如本发明实施例中的天社交账号的识别方法和装置对应的程序指令/模块，处理器802通过运行存储在存储器804内的软件程序以及模块，从而执行各种功能应用以及数据处理，即实现上述社交账号的识别方法。存储器804可包括高速随机存储器，还可以包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器804可进一步包括相对于处理器802远程设置的存储器，这些远程存储器可以通过网络连接至终端。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0141] 上述的传输装置810用于经由一个网络接收或者发送数据。上述的网络具体实例可包括有线网络及无线网络。在一个实例中，传输装置810包括一个网络适配器(Network Interface Controller,NIC)，其可通过网线与其他网络设备与路由器相连从而可与互联网或局域网进行通讯。在一个实例中，传输装置810为射频(Radio Frequency,RF)模块，其用于通过无线方式与互联网进行通讯。

[0142] 上述本发明实施例序号仅仅为了描述，不代表实施例的优劣。

[0143] 上述实施例中的集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在上述计算机可读取的存储介质中。基于这样的理解，本发明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全部或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存储介质中，包括若干指令用以使得一台或多台计算机设备(可为个人计算机、服务器或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。

[0144] 在本发明的上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其他实施例的相关描述。

[0145] 在本申请所提供的几个实施例中，应该理解到，所揭露的客户端，可通过其它的方式实现。其中，以上所描述的装置实施例仅仅是示意性的，例如所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，单元或模块的间接耦合或通信连接，可以是电性或其它的形式。

[0146] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0147] 另外，在本发明各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

[0148] 以上所述仅是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本发明的保护范围。

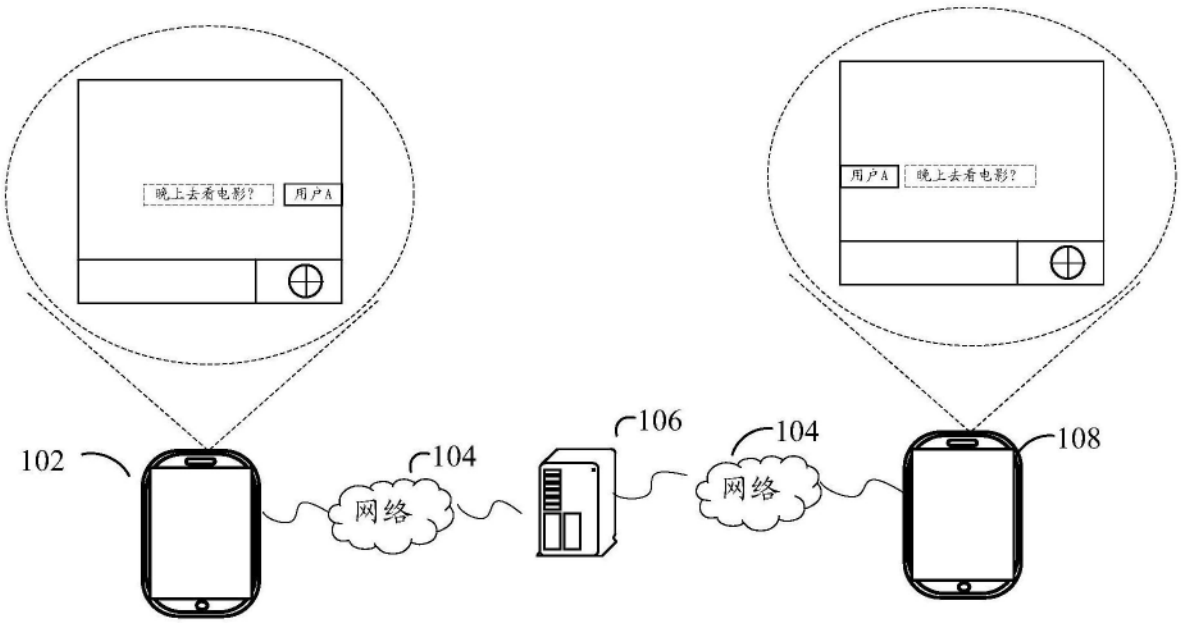

晚上去看电影?
用户A
用户A
晚上去看电影?
102
104
网络
106
104
网络
108

图1

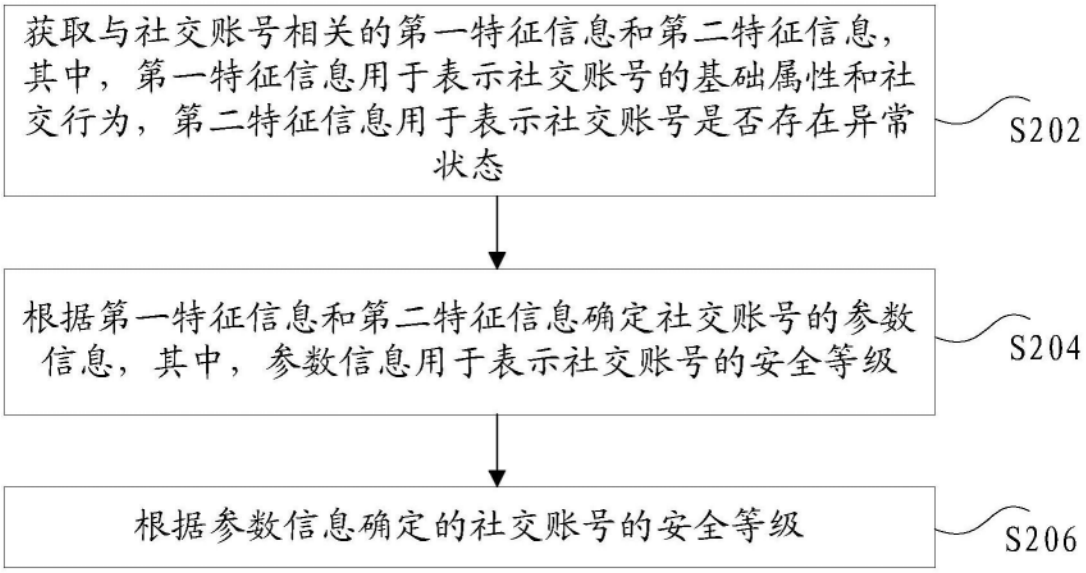

获取与社交账号相关的第一特征信息和第二特征信息，其中，第一特征信息用于表示社交账号的基础属性和社交行为，第二特征信息用于表示社交账号是否存在异常状态
S202
根据第一特征信息和第二特征信息确定社交账号的参数信息，其中，参数信息用于表示社交账号的安全等级
S204
根据参数信息确定的社交账号的安全等级
S206

图2

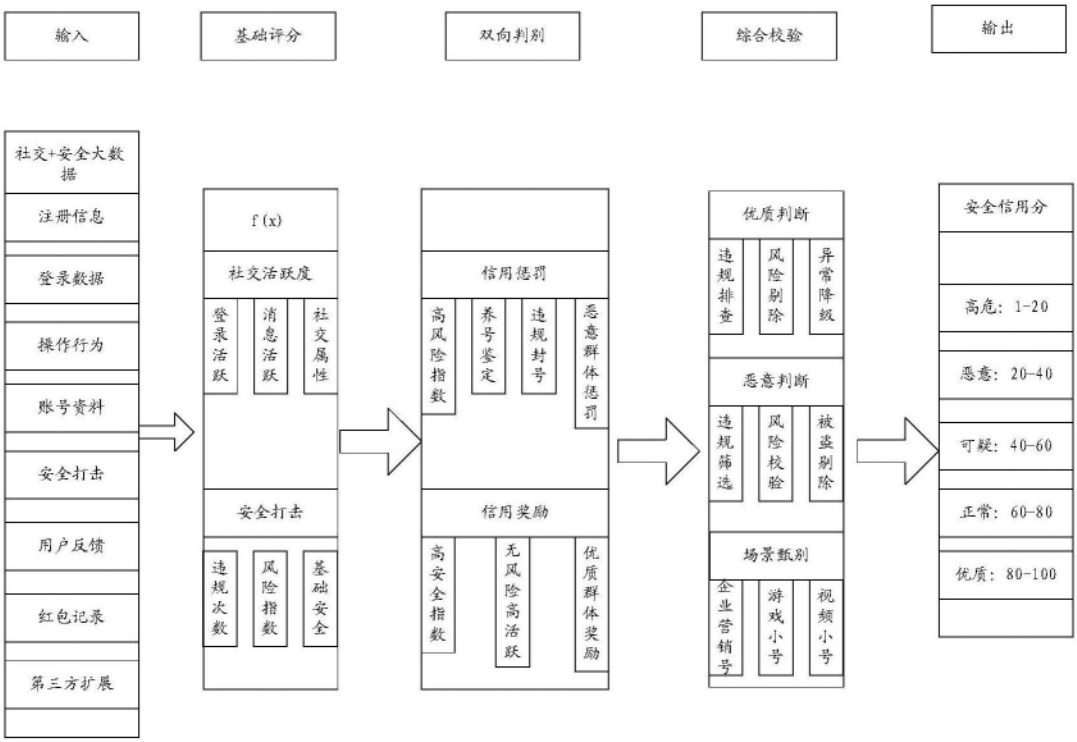

输入
基础评分
双向判别
综合校验
输出
社交+安全大数据
注册信息
登录数据
操作行为
账号资料
安全打击
用户反馈
红包记录
第三方扩展
f(x)
社交活跃度
登录活跃
消息活跃
社交属性
安全打击
违规次数
风险指数
基础安全
信用惩罚
高风险指数
养号鉴定
违规封号
恶意群体惩罚
信用奖励
高安全指数
无风险高活跃
优质群体奖励
优质判断
违规排查
风险剔除
异常降级
恶意判断
违规筛选
风险校验
被盗刷除
场景甄别
企业营销号
游戏小号
视频小号
安全信用分
高危：1-20
恶意：20-40
可疑：40-60
正常：60-80
优质：80-100

图3

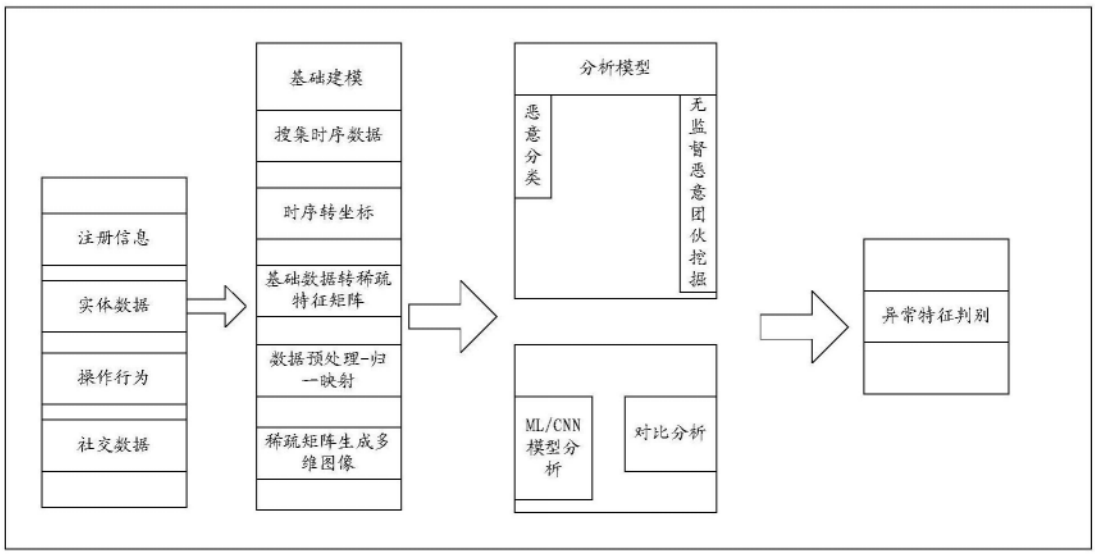

注册信息
实体数据
操作行为
社交数据
基础建模
搜集时序数据
时序转坐标
基础数据转稀疏特征矩阵
数据预处理-归一映射
稀疏矩阵生成多维图像
分析模型
恶意分类
无监督恶意团伙挖掘
ML/CNN模型分析
对比分析
异常特征判别

图4

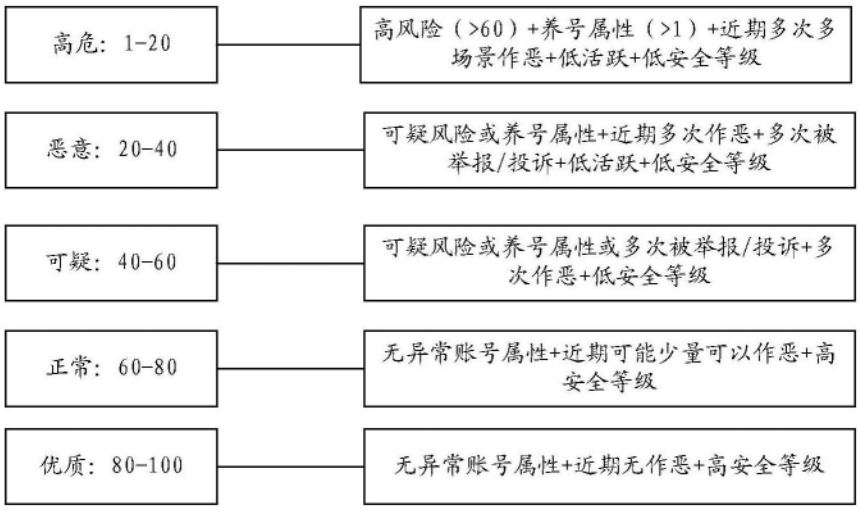

高危：1-20
高风险（>60）+养号属性（>1）+近期多次多场景作恶+低活跃+低安全等级
恶意：20-40
可疑风险或养号属性+近期多次作恶+多次被举报/投诉+低活跃+低安全等级
可疑：40-60
可疑风险或养号属性或多次被举报/投诉+多次作恶+低安全等级
正常：60-80
无异常账号属性+近期可能少量可以作恶+高安全等级
优质：80-100
无异常账号属性+近期无作恶+高安全等级

图5

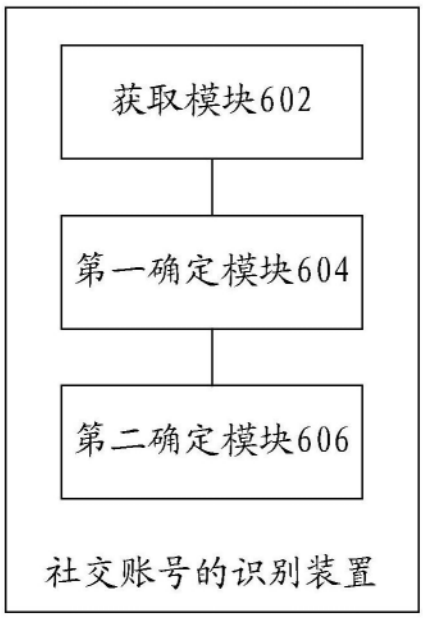

获取模块602
第一确定模块604
第二确定模块606
社交账号的识别装置

图6

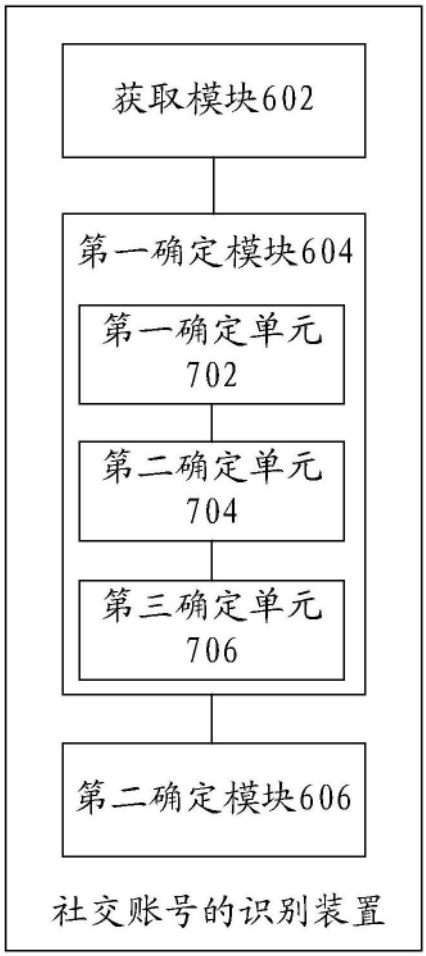

获取模块602
第一确定模块604
第一确定单元702
第二确定单元704
第三确定单元706
第二确定模块606
社交账号的识别装置

图7

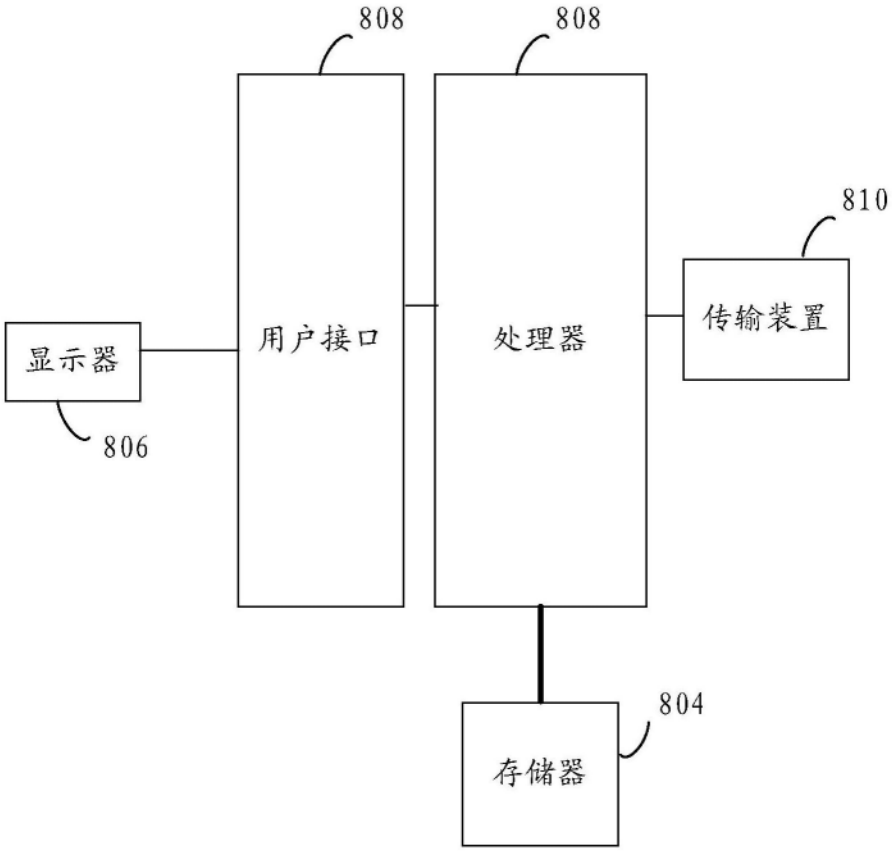


图8