

图结构模型训练和垃圾账号识别

相关申请的交叉引用

- [01]本专利申请要求于 2018 年 3 月 14 日提交的、申请号为 201810209270.1、发明名称为“图结构模型训练、垃圾账号识别方法、装置以及设备”的中国专利申请的优先权，
- 5 该申请的全文以引用的方式并入本文中。

技术领域

[02]本说明书涉及计算机软件技术领域，尤其涉及图结构模型训练、垃圾账户识别方法、装置以及设备。

背景技术

- 10 [03]随着计算机和互联网技术的迅速发展，很多业务可以在网上进行，用户要使用这些业务，往往需要注册相应的账户，比如电商平台账户、第三方支付平台账户、论坛平台账户等。

- [04]一些用户或者组织出于不良目的，会注册大量账户，并利用这些账户进行一些异常操作，比如传播留言、推销虚假广告、刷单等，这些账户可能给平台带来风险，而且对于平台价值也较低，被视为垃圾账户。
- 15

[05]在现有技术中，一般通过用户举报的方式，判定垃圾账户并进行相应的处理，比如冻结、注销等。

[06]基于现有技术，需要有效的垃圾账户识别方案。

发明内容

- 20 [07]本说明书实施例提供图结构模型训练、垃圾账户识别方法、装置以及设备，用以解决如下技术问题：需要有效的垃圾账户识别方案。

[08]为解决上述技术问题，本说明书实施例是这样实现的：

- [09]本说明书实施例提供一种图结构模型训练方法，包括：获取账户媒介网络图，所述账户媒介网络图中的节点表示账户和媒介，至少部分边表示其连接的节点间具有登录
- 25 行为关系；获取所述节点的特征数据和风险标注数据，所述特征数据反映对应节点在时

间序列上的登录行为；根据所述账户媒介网络图、所述特征数据和所述风险标注数据，训练预定义的图结构模型，用以识别垃圾账户。

[10]本说明书实施例提供的一种垃圾账户识别方法，包括：获取待识别账户的特征数据，以及获取所述待识别账户所属的账户媒介网络图；将所述待识别账户的特征数据，以及该账户媒介网络图中对应于所述待识别账户的拓扑结构，输入利用上述图结构模型训练方法训练后的图结构模型进行计算；根据所述训练后的图结构模型输出的预测数据，判定所述待识别账户是否为垃圾账户。

[11]本说明书实施例提供的一种图结构模型训练装置，包括：第一获取模块，获取账户媒介网络图，所述账户媒介网络图中的节点表示账户和媒介，至少部分边表示其连接的节点间具有登录行为关系；第二获取模块，获取所述节点的特征数据和风险标注数据，所述特征数据反映对应节点在时间序列上的登录行为；训练识别模块，根据所述账户媒介网络图、所述特征数据和所述风险标注数据，训练预定义的图结构模型，用以识别垃圾账户。

[12]本说明书实施例提供的一种垃圾账户识别装置，包括：获取模块，获取待识别账户的特征数据，以及获取所述待识别账户所属的账户媒介网络图；输入模块，将所述待识别账户的特征数据，以及该账户媒介网络图中对应于所述待识别账户的拓扑结构，输入利用上述图结构模型训练方法训练后的图结构模型进行计算；判定模块，根据所述训练后的图结构模型输出的预测数据，判定所述待识别账户是否为垃圾账户。

[13]本说明书实施例提供的一种图结构模型训练设备，包括：至少一个处理器；以及，与所述至少一个处理器通信连接的存储器。其中，所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够：获取账户媒介网络图，所述账户媒介网络图中的节点表示账户和媒介，至少部分边表示其连接的节点间具有登录行为关系；获取所述节点的特征数据和风险标注数据，所述特征数据反映对应节点在时间序列上的登录行为；根据所述账户媒介网络图、所述特征数据和所述风险标注数据，训练预定义的图结构模型，用以识别垃圾账户。

[14]本说明书实施例采用的上述至少一个技术方案能够达到以下有益效果：通过上述基于账户媒介网络图的图嵌入方案，能够利用垃圾账户的媒介聚集性和时间聚集性，有效地识别垃圾账户。

附图说明

- [15]为了更清楚地说明本说明书实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本说明书中记载的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。
- [16]图 1 为本说明书的方案在一种实际应用场景下涉及的一种整体架构示意图；
- [17]图 2 为本说明书实施例提供的一种图结构模型训练方法的流程示意图；
- [18]图 3 为本说明书实施例提供的一种垃圾账户识别方法的流程示意图；
- [19]图 4 为本说明书实施例提供的上述各方法的一种实施方案示意图；
- 10 [20]图 5 为本说明书实施例提供的对应于图 2 的一种图结构模型训练装置的结构示意图；
- [21]图 6 为本说明书实施例提供的对应于图 3 的一种垃圾账户识别装置的结构示意图；
- [22]图 7 为本说明书实施例提供的对应于图 2 的一种图结构模型训练设备的结构示意图；
- [23]图 8 为本说明书实施例提供的对应于图 3 的一种垃圾账户识别设备的结构示意图。

具体实施方式

- 15 [24]本说明书实施例提供图结构模型训练、垃圾账户识别方法、装置以及设备。
- [25]为了使本技术领域的人员更好地理解本说明书中的技术方案，下面将结合本说明书实施例中的附图，对本说明书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申请一部分实施例，而不是全部的实施例。基于本说明书实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都应当属于本申请保护的范围。
- 20 [26]一般地，用于非正常行为的账户均可以视为垃圾账户，比如通过机器自动大量注册的账户等。在大多数业务平台的风控体系中，识别垃圾账户在风控安全上具有重要意义，难点体现为新注册的账户并没有足够的账户画像信息判定其是否为垃圾账户。本说明书考虑到了垃圾账户往往具有的两种特性，媒介聚集性和时间聚集性，进而根据这两种特性提出了一种有监督图嵌入的垃圾账户识别方案，能够有效地识别垃圾账户。这里，图嵌入可以指将图中节点的一些原始数据在指定的特征空间（本说明书称为隐特征空间）
- 25

进行映射，得到相应的嵌入向量，用于表示节点。

[27]媒介聚集性可以指：同一个恶意用户注册的多个垃圾账户往往是通过同一个或者少数几个媒介注册的。导致媒介聚集性的原因在于：恶意用户往往寻求利益上的追求，他们并没有足够的资源通过大量媒介注册大量账户。

- 5 [28]时间聚集性可以指：同一个恶意用户控制的垃圾账户往往在某一个短时间段内形成大量非正常行为。导致时间聚集性的原因在于：恶意用户往往追求短期的利益目标，造成在他们控制下的账户必须在短时间内产生大量非正常行为。

- [29]图 1 为本说明书的方案在一种实际应用场景下涉及的一种整体架构示意图。该整体架构中，主要涉及预定义的图结构模型所在有监督学习服务器，以及训练图结构模型能够使用的三类数据：反映指定行为关系的账户媒介网络图、账户媒介网络图中节点反映时间序列上的指定行为的特征数据、节点的风险标注数据。指定行为比如是登录行为、注册行为、交易行为等。图结构模型训练后，能够用于识别垃圾账户。
- 10

[30]这些训练用的数据可以由有监督学习服务器或者其他设备生成，或者也可以人工编写。

- 15 [31]下面基于图 1 中示例性的架构，对本说明书的方案进行详细说明。

[32]图 2 为本说明书实施例提供的一种图结构模型训练方法的流程示意图。图 2 中的流程包括以下步骤：

[33]S202：获取账户媒介网络图，所述账户媒介网络图中的节点表示账户和媒介，至少部分边表示其连接的节点间具有登录行为关系。

- 20 [34]在本说明书实施例中，账户媒介网络图是一种异质网络图，异质指节点性质差异，比如，图中某些节点可能表示账户，某些节点可能表示账户相关的媒介。账户通过媒介注册或者实现后续行为，媒介比如是设备、IP 网络、物理地址等。

- [35]在本说明书实施例中，账户媒介网络图可以根据账户一定时间范围内的历史数据生成。历史数据可以包括账户的注册行为数据，比如，账户是通过怎样的媒介注册的、账户注册时填写的注册信息等；历史数据也可以包括账户注册后的行为数据，比如账户的登录行为数据、交易行为数据等。对于一定时间范围，这里不做具体限定，可以预先设定，比如最近若干天等。
- 25

[36]对于新注册的账户，比如，可以根据账户注册行为数据，和/或注册后一定时间范围

(通常是某个短时间范围)内的指定行为数据,生成账户媒介网络图,以便于尽量提前识别垃圾账户。

[37]为了便于描述,下面一些实施例主要以媒介为设备,指定行为为登录行为为例进行说明,则账户媒介网络图具体为账户设备网络图。

- 5 [38]在生成账户设备网络图时,可以先确定所要表示的各账户和各设备,将要表示的每个账户分别用一个节点表示,每个设备也可以分别用一个节点表示,任意一个节点要么表示账户,要么表示设备。进一步地,若两个节点间具有登录关系,则在这两个节点间建立一条表示该登录行为关系的边,从而生成账户设备网络图。

- 10 [39]这里,登录行为关系主要指账户与设备间的关系,若某账户一定时间范围内在某设备上登录过,则可以称该账户与该设备间具有登录行为关系。需要说明的是,在实际应用中,若有需求,登录行为关系的具体含义也可以拓宽,比如,登录行为关系也可以包括账户与账户间的关系,若某账户与另一账户一定时间范围内曾在同一设备上的登录过,则可以称该账户与该另一账户间具有登录行为关系。

- 15 [40]在本说明书实施例中,账户媒介网络图可以是无向图,也可以是有向图,这里不做具体限定。一般地,若只反映登录行为关系,则采用无向图即可;而若还反映诸如交易行为关系等更多的关系,也可以采用有向图,在有向图中,边的指向表明业务关系方向,比如,若A节点表示买家账户,B节点表示卖家账户,则表示A节点与B节点间交易行为关系的边可以从A节点指定B节点,该指向也能够反映资金流动方向。

- 20 [41]在本说明书实施例中,为了便于图计算,账户媒介网络图可以用矩阵进行表示。可以使矩阵的不同的单行、单列分别表示账户媒介网络图中不同节点,矩阵中的不同元素分别表示,其所在行与列表示的节点间的登录行为关系。

- [42]例如,对于表示账户设备网络图的矩阵,比如将矩阵记作 $G \in \mathbb{R}^{n \times n}$,矩阵为 n 行 n 列, n 表示所要表示的账户数加设备数。假定账户设备网络图为一个二部图,只有表示账户的节点与表示设备的节点间才可能有边,若有边,则对应的元素为1,否则为0,比如,
- 25 若表示账户 i 与设备 j 的节点间有边,则 G 的第 i 行 j 列的元素 $g_{i,j}=1$ 。

[43]S204: 获取所述节点的特征数据和风险标注数据,所述特征数据反映对应节点在时间序列上的登录行为。

[44]在本说明书实施例中,步骤S204中的节点可以是账户媒介网络图中的部分节点,而未必是全部节点。比如,可以是表示账户的至少部分节点,当然,还可以获取表示媒

介的至少部分节点的特征数据,表示媒介的节点的特征数据未必要反映对应节点在时间序列上的登录行为,这里不做具体限定,若媒介是设备,其特征数据比如可以反映诸如设备类型、设备厂商等设备信息。

- [45]在本说明书实施例中,特征数据可以根据账户一定时间范围内的历史数据生成。针对前面提到的时间聚集性,在生成特征数据时,不光考虑账户的登录行为本身,还考虑账户的登录行为与时间之间的关系,比如,可以将时间序列化(如划分多个时间区间、或者采样离散的时间点等),确定在账户的登录行为在时间序列上的分布情况,比如,登录行为发生的具体时刻、持续时间、单位时间内登录行为发生的次数等。特征数据一般可以表示为向量或者矩阵,下面一些实施例主要以特征数据表示为向量为例进行说明。
- 10 [46]在本说明书实施例中,以账户设备网络图中的各节点为例。某些节点可能表示垃圾账户,某些节点可能表示垃圾账户登录过的设备,这些情况尚且未明确,需要通过特定手段明确其中的至少部分情况,才能够得到有训练标签的训练样本,进而才能够用于后续的有监督学习。特定手段这里不做具体限定,比如,可以基于抽样精确追踪分析,也可以基于用户举报等手段。
- 15 [47]通过明确的上述至少部分情况,能够预先或者实时地为部分节点标注风险标注数据,风险标注数据能够表明节点所存在的风险,比如,是否表示垃圾账户,是否表示登录过垃圾账户的设备等。在实际应用中,这里的风险可以不局限于垃圾账户相关内容,比如也可以表示正常账户存在的容易受到攻击的风险等。上述的训练标签可以根据风险标注数据得到,一般地,风险标注数据可以直接作为训练标签。
- 20 [48]风险标注数据的表示形式是多样的,这里不做具体限定。比如,若确定某节点与垃圾账户无关,该节点的风险标注数据可以记作 1,若确定某节点表示垃圾账户或者表示登录过垃圾账户的设备,该节点的风险标注数据可以记作 0;等等。
- [49]另外,在实际应用中,也可以只对表示账户的节点标注风险标注数据,而不对表示媒介的节点标注风险标注数据。
- 25 [50]S206: 根据所述账户媒介网络图、所述特征数据和所述风险标注数据,训练预定义的图结构模型,用以识别垃圾账户。
- [51]在本说明书实施例中,图结构模型的至少部分参数是基于图结构的,这部分参数可以用根据至少部分账户媒介网络图和/或特征数据进行赋值。图结构模型还有一部分参数需要通过训练优化求解。

[52]例如，在一种实际应用场景下，图结构模型用于根据节点的特征数据，以及账户媒介网络图中对应于所述节点的拓扑结构，计算隐特征空间中所述节点多次迭代后的嵌入向量；进一步地，图结构模型还用于根据所述嵌入向量，计算所述节点的预测数据，预测数据表示所述节点对应于垃圾账户的可能性。

- 5 [53]预测数据的形式是多样的，这里不做具体限定，比如是概率值、非概率值的分值、或者分类类别标识等形式。

[54]在实际应用中，图结构模型也未必要计算预测数据，可以在计算出嵌入向量后输出给别的模型使用，本说明书不详细分析这种情况，下面一些实施例主要还是基于上例进行说明。

- 10 [55]在本说明书实施例中，图结构模型训练后，即可以用于分类或者回归，以预测输入数据的风险性质。

[56]例如，输入数据可以是待识别账户对应的特征数据，以及待识别账户所属账户媒介网络图（未必是步骤 S202 中的账户媒介网络图）中对应的拓扑结构，通过训练后的图结构模型的计算，输出预测数据，从而能够判定待识别账户是否为垃圾账户。其中，待识别账户可以是步骤 S202 中的账户媒介网络图中节点所表示的，也可以是该账户媒介网络图之外的；对于前一种情况，输入数据已经确定，因此可以直接进行识别，而对于后一种情况，输入数据可能尚未确定，则可以采用本说明书的方案，先确定输入数据，再进行识别。

- 20 [57]当然，根据风险标注数据的具体内容，图结构模型除了用于识别垃圾账户以外，还可能用于预测其他方面的风险，原理都是相同的，这里不再赘述。

[58]通过图 2 的方法，通过上述基于账户媒介网络图的图嵌入方案，能够利用垃圾账户的媒介聚集性和时间聚集性，有效地识别垃圾账户。

[59]基于图 2 的方法，本说明书实施例还提供了该方法的一些具体实施方案，以及扩展方案，下面进行说明。

- 25 [60]在本说明书实施例中，根据上面的例子可知，对于步骤 S206，所述识别垃圾账户，具体可以包括：获取待识别账户的特征数据，以及获取所述待识别账户所属的账户媒介网络图；将所述待识别账户的特征数据，以及该账户媒介网络图中对应于所述待识别账户的拓扑结构输入训练后的所述图结构模型进行计算；获取训练后的所述图结构模型计算后输出的预测数据，以判定所述待识别账户是否为垃圾账户。

[61]在本说明书实施例中，前面已经提到，时间序列可以通过划分时间范围得到，在这种该情况下，对于步骤 S204，所述获取所述节点的特征数据，具体可以包括：获取所述节点一定时间范围内的登录行为数据；将所述一定时间范围进行划分，得到时间序列；根据所述时间序列中所述登录行为数据的分布情况，生成特征向量，作为所述节点的特征数据。

[62]例如，假定一定时间范围被设定为过去的 m 天，按小时划分，则能够划分得到 $m*24$ 个时间分段构成的时间序列，可以根据账户 i 在各时间分段内的登录次数，生成 d 维特征向量 x_i 。这里并不限定 x_i 的具体构建方式，比如， d 可以等于 $m*24$ ， x_i 的每个元素可以分别表示账户 i 在其中一个时间分段内的登录次数， x_i 的元素可以是经过归一化处理。

[63]在本说明书实施例中，隐特征空间中节点在第 t 次迭代后的嵌入向量可以是根据所述节点的特征数据、账户媒介网络图中对应于所述节点的拓扑结构，以及隐特征空间中所述节点在第 $t-1$ 次迭代后的嵌入向量计算得到的。更直观地，结合上面的一些例子，一种示例性的图结构模型的定义及训练过程如下所示：

15 “初始化图结构模型待优化求解的参数： $W_1 \in \mathbb{R}^{d,k}$ 、 $W_2 \in \mathbb{R}^{k,k}$ 、 $w \in \mathbb{R}^k$ ，比如，采用标准高斯分布初始化等；

迭代训练设定次数或者直至训练收敛：

初始化 $\Phi^{(1)}$ ；

{for $t=1$ to N : // N 次迭代执行，以计算嵌入向量

20 $\Phi^{(t+1)} = \sigma(XW_1 + G\Phi^{(t)}W_2)$; } // (公式一)

$pred_i = w^T \phi_i$; // (公式二) 根据嵌入向量，计算预测数据

optimize $\arg \min_{W_1, W_2, w} \sum_i L(y_i, pred_i)$; // (公式三) 优化参数

[64]其中， $\Phi^{(t+1)}$ 表示隐特征空间中至少一个所述节点在第 $t+1$ 次迭代后的嵌入向量， σ 表示非线性变换函数（比如，Relu、Sigmoid、Tanh 等函数）， W_1 、 W_2 表示权重矩阵， X 表示所述至少一个所述节点的特征数据， G 表示所述账户媒介网络图中对应于所述至少一个所述节点的拓扑结构； $pred_i$ 表示第 i 个所述节点经过迭代后的预测数据， ϕ_i 表示隐

特征空间中第 i 个所述节点所述多次迭代后的嵌入向量, w^T 表示用于将 ϕ_i 分值化的参数向量, T 表示转置运算; y_i 表示第 i 个所述节点的风险标注数据, L 表示用于度量所述预测数据与其对应的风险标注数据的一致性差距的损失函数, 这里不做具体限定, 比如, 可以采用 logistic loss、hinge loss、cross_entropy 等损失函数。

- 5 [65] 在前面的一个例子中, $G \in \mathbb{R}^{n,n}$, 此时 G 表示所述账户媒介网络图完整的拓扑结构, 在这种情况下, 可以使 X 表示所述账户媒介网络图中全部节点的特征数据, 以及可以使 Φ 表示所述账户媒介网络图中全部节点的嵌入向量, 比如, $X \in \mathbb{R}^{n,d}$, X 的每行分别表示一个节点的特征数据, $\Phi \in \mathbb{R}^{n,k}$, k 表示嵌入的隐特征空间的维度, Φ 的每行分别表示一个节点的嵌入向量。

- 10 [66] 当然, 也可以使 G 只表示账户媒介网络图完整的拓扑结构的一部分, 相应地, X 、 Φ 也可以只包含账户媒介网络图中一部分节点的数据。

[67] 在本说明书实施例中, 可以以预测数据与其对应的风险标注数据的一致性最大化为训练目标, 训练图结构模型。则在上例的场景下, 对于步骤 S206, 所述训练预定义的图结构模型, 具体可以包括: 利用反向传播算法和所述风险标注数据, 对

- 15 $\arg \min_{w_1, w_2, w} \sum_i L(y_i, \text{pred}_i)$ 进行优化, 求得最优的 w_1 、 w_2 、 w 。

[68] 上面的公式一、公式二、公式三是示例性的, 并非唯一方案。比如, 公式一中 X 、 G 分别的所在项可以通过乘法、指数或者对数等运算进行变形, 或者还可以合并这两项, 或者还可以删除其中一项; 再比如, 公式二中也可以利用 *softmax* 函数对 ϕ_i 进行分值化; 再比如, 若公式三的损失函数表示预测数据与对应的风险预测数据的一致化程度, 则公

20 式三中可以调整为求最大值而不是求最小值; 等等。

[69] 进一步地, 本说明书实施例还提供了基于上述图结构模型的一种垃圾账户识别方法的流程示意图, 如图 3 所示。图 3 中的流程包括以下步骤:

[70] S302: 获取待识别账户的特征数据, 以及获取所述待识别账户所属的账户媒介网络图。

- 25 [71] S304: 将所述待识别账户的特征数据, 以及该账户媒介网络图中对应于所述待识别账户的拓扑结构, 输入利用上述图结构模型训练方法训练后的图结构模型进行计算。

[72] S306: 根据所述训练后的图结构模型输出的预测数据, 判定所述待识别账户是否为

垃圾账户。

[73]根据上面的说明，本说明书实施例还提供了上述各方法的一种实施方案示意图，如图4所示。

- [74]图4的方案可以包括以下步骤：获取过去的 m 天内的账户设备网络图、每个账户的
5 登录行为数据和风险标注数据；通过有监督学习，训练预定义的图结构模型，得到训练后的图结构模型；对待预测数据（如一个或者多个账户），获取对应的账户设备网络图、每个账户的登录行为数据；利用训练后的图结构模型进行预测，得到预测结果。

[75]上面对本说明书实施例提供的方法进行了说明，基于同样的思路，本说明书实施例还提供了对应的装置和设备，如图5~图8所示。

- 10 [76]图5为本说明书实施例提供的对应于图2的一种图结构模型训练装置的结构示意图，该装置可以位于图2中流程的执行主体，包括：第一获取模块501，获取账户媒介网络图，所述账户媒介网络图中的节点表示账户和媒介，至少部分边表示其连接的节点间具有登录行为关系；第二获取模块502，获取所述节点的特征数据和风险标注数据，所述特征数据反映对应节点在时间序列上的登录行为；训练识别模块503，根据所述账户媒介
15 网络图、所述特征数据和所述风险标注数据，训练预定义的图结构模型，用以识别垃圾账户。

[77]可选地，所述媒介包括设备。

[78]可选地，所述图结构模型用于根据所述节点的特征数据，以及所述账户媒介网络图中对应于所述节点的拓扑结构，计算隐特征空间中所述节点多次迭代后的嵌入向量。

- 20 [79]可选地，所述图结构模型还用于根据所述嵌入向量，计算所述节点的预测数据，所述预测数据表示所述节点对应于垃圾账户的可能性。

- [80]可选地，所述训练识别模块503识别垃圾账户，具体包括：训练识别模块503获取待识别账户的特征数据，以及获取所述待识别账户所属的账户媒介网络图；将所述待识别账户的特征数据，以及该账户媒介网络图中对应于所述待识别账户的拓扑结构输入训练后的所述图结构模型进行计算；获取训练后的所述图结构模型计算后输出的预测数据，
25 以判定所述待识别账户是否为垃圾账户。

[81]可选地，所述第二获取模块502获取所述节点的特征数据，具体包括：所述第二获取模块502获取所述节点一定时间范围内的登录行为数据；将所述一定时间范围进行划分，得到时间序列；根据所述时间序列中所述登录行为数据的分布情况，生成特征向量，

作为所述节点的特征数据。

[82]可选地, 隐特征空间中所述节点在第 t 次迭代后的嵌入向量是根据所述节点的特征数据、所述账户媒介网络图中对应于所述节点的拓扑结构, 以及隐特征空间中所述节点在第 $t-1$ 次迭代后的嵌入向量计算得到的。

- 5 [83]可选地, 所述根据所述节点的特征数据, 以及所述账户媒介网络图中对应于所述节点的拓扑结构, 计算隐特征空间中所述节点多次迭代后的嵌入向量, 具体包括按照如下公式, 计算隐特征空间中所述节点多次迭代后的嵌入向量:

$$\Phi^{(t+1)} = \sigma(XW_1 + G\Phi^{(t)}W_2);$$

- [84]其中, $\Phi^{(t+1)}$ 表示隐特征空间中至少一个所述节点在第 $t+1$ 次迭代后的嵌入向量, σ 表示非线性变换函数, W_1 、 W_2 表示权重矩阵, X 表示所述至少一个所述节点的特征数据, G 表示所述账户媒介网络图中对应于所述至少一个所述节点的拓扑结构。
- 10

[85]可选地, 所述根据所述嵌入向量, 计算所述节点的预测数据, 具体包括按照如下公式, 计算所述节点的预测数据:

$$pred_i = w^T \phi_i;$$

- 15 [86]其中, $pred_i$ 表示第 i 个所述节点经过迭代后的预测数据, ϕ_i 表示隐特征空间中第 i 个所述节点所述多次迭代后的嵌入向量, w^T 表示用于将 ϕ_i 分值化的参数向量, T 表示转置运算。

- [87]可选地, 所述训练识别模块 503 训练预定义的图结构模型, 具体包括: 所述训练识别模块 503 以所述预测数据与其对应的风险标注数据的一致性最大化为训练目标, 训练预定义的图结构模型。
- 20

[88]可选地, 所述训练识别模块 503 训练预定义的图结构模型, 具体包括: 所述训练识别模块 503 利用反向传播算法和所述风险标注数据, 对 $\arg \min_{W_1, W_2, w} \sum_i L(y_i, pred_i)$ 进行优化, 求得最优的 W_1 、 W_2 、 w 。其中, y_i 表示第 i 个所述节点的风险标注数据, L 表示用于度量所述预测数据与其对应的风险标注数据的一致性差距的损失函数。

- 25 [89]图 6 为本说明书实施例提供的对应于图 3 的一种垃圾账户识别装置的结构示意图, 该装置可以位于图 3 中流程的执行主体, 包括: 获取模块 601, 获取待识别账户的特征数据, 以及获取所述待识别账户所属的账户媒介网络图; 输入模块 602, 将所述待识别

账户的特征数据,以及该账户媒介网络图中对应于所述待识别账户的拓扑结构,输入利用上述图结构模型训练方法训练后的图结构模型进行计算;判定模块 603,根据所述训练后的图结构模型输出的预测数据,判定所述待识别账户是否为垃圾账户。

[90]图 7 为本说明书实施例提供的对应于图 2 的一种图结构模型训练设备的结构示意图,所述设备包括:至少一个处理器;以及,与所述至少一个处理器通信连接的存储器。其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够:获取账户媒介网络图,所述账户媒介网络图中的节点表示账户和媒介,至少部分边表示其连接的节点间具有登录行为关系;获取所述节点的特征数据和风险标注数据,所述特征数据反映对应节点在时间序列上的登录行为;根据所述账户媒介网络图、所述特征数据和所述风险标注数据,训练预定义的图结构模型,用以识别垃圾账户。

[91]图 8 为本说明书实施例提供的对应于图 3 的一种垃圾账户识别设备的结构示意图,所述设备包括:至少一个处理器;以及,与所述至少一个处理器通信连接的存储器。其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够:获取待识别账户的特征数据,以及获取所述待识别账户所属的账户媒介网络图;将所述待识别账户的特征数据,以及该账户媒介网络图中对应于所述待识别账户的拓扑结构,输入利用上述图结构模型训练方法训练后的图结构模型进行计算;根据所述训练后的图结构模型输出的预测数据,判定所述待识别账户是否为垃圾账户。

[92]基于同样的思路,本说明书实施例还提供了对应于图 2 的一种非易失性计算机存储介质,存储有计算机可执行指令,所述计算机可执行指令设置为:获取账户媒介网络图,所述账户媒介网络图中的节点表示账户和媒介,至少部分边表示其连接的节点间具有登录行为关系;获取所述节点的特征数据和风险标注数据,所述特征数据反映对应节点在时间序列上的登录行为;根据所述账户媒介网络图、所述特征数据和所述风险标注数据,训练预定义的图结构模型,用以识别垃圾账户。

[93]基于同样的思路,本说明书实施例还提供了对应于图 3 的一种非易失性计算机存储介质,存储有计算机可执行指令,所述计算机可执行指令设置为:获取待识别账户的特征数据,以及获取所述待识别账户所属的账户媒介网络图;将所述待识别账户的特征数据,以及该账户媒介网络图中对应于所述待识别账户的拓扑结构,输入利用上述图结构模型训练方法训练后的图结构模型进行计算;根据所述训练后的图结构模型输出的预测

数据，判定所述待识别账户是否为垃圾账户。

[94]上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下，在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

[95]本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于装置、设备、非易失性计算机存储介质实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

[96]本说明书实施例提供的装置、设备、非易失性计算机存储介质与方法是对应的，因此，装置、设备、非易失性计算机存储介质也具有与对应方法类似的有益技术效果，由于上面已经对方法的有益技术效果进行了详细说明，因此，这里不再赘述对应装置、设备、非易失性计算机存储介质的有益技术效果。

[97]在 20 世纪 90 年代，对于一个技术的改进可以很明显地区分是硬件上的改进（例如，对二极管、晶体管、开关等电路结构的改进）还是软件上的改进（对于方法流程的改进）。然而，随着技术的发展，当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此，不能说一个方法流程的改进就不能用硬件实体模块来实现。例如，可编程逻辑器件（Programmable Logic Device, PLD）（例如现场可编程门阵列（Field Programmable Gate Array, FPGA））就是这样一种集成电路，其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统“集成”在一片 PLD 上，而不需要请芯片制造厂商来设计和制作专用的集成电路芯片。而且，如今，取代手工地制作集成电路芯片，这种编程也多半改用“逻辑编译器（logic compiler）”软件来实现，它与程序开发撰写时所用的软件编译器相类似，而要编译之前的原始代码也得用特定的编程语言来撰写，此称之为硬件描述语言（Hardware Description Language, HDL），而 HDL 也并非仅有一种，而是有许多种，如 ABEL（Advanced Boolean Expression Language）、AHDL（Altera Hardware Description Language）、Confluence、CUPL（Cornell University Programming Language）、HDCal、JHDL（Java Hardware Description Language）、Lava、Lola、MyHDL、PALASM、RHDL（Ruby Hardware Description Language）等，目前最

普遍使用的是 VHDL(Very-High-Speed Integrated Circuit Hardware Description Language) 与 Verilog。本领域技术人员也应该清楚, 只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中, 就可以很容易得到实现该逻辑方法流程的硬件电路。

[98] 控制器可以按任何适当的方式实现, 例如, 控制器可以采取例如微处理器或处理器以及存储可由该 (微) 处理器执行的计算机可读程序代码 (例如软件或固件) 的计算机可读介质、逻辑门、开关、专用集成电路 (Application Specific Integrated Circuit, ASIC)、可编程逻辑控制器和嵌入微控制器的形式, 控制器的例子包括但不限于以下微控制器: ARC 625D、Atmel AT91SAM、Microchip PIC18F26K20 以及 Silicone Labs C8051F320, 存储器控制器还可以被实现为存储器的控制逻辑的一部分。本领域技术人员也知道, 除了以纯计算机可读程序代码方式实现控制器以外, 完全可以通过将方法步骤进行逻辑编程来使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件, 而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至, 可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

[99] 上述实施例阐明的系统、装置、模块或单元, 具体可以由计算机芯片或实体实现, 或者由具有某种功能的产品来实现。一种典型的实现设备为计算机。具体的, 计算机例如可以为个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任何设备的组合。

[100] 为了描述的方便, 描述以上装置时以功能分为各种单元分别描述。当然, 在实施本说明书时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

[101] 本领域内的技术人员应明白, 本说明书实施例可提供为方法、系统、或计算机程序产品。因此, 本说明书实施例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, 本说明书实施例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器、CD-ROM、光学存储器等) 上实施的计算机程序产品的形式。

[102] 本说明书是参照根据本说明书实施例的方法、设备 (系统)、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据

处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

[103] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

[104] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

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

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

[107] 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方法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)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(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暂存电脑可读媒体(transitory media)，如调制的数据信号和载波。

[108] 还需要说明的是，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、商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、商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“包括一个……”限定的要素，并不排除在

包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[109] 本领域技术人员应明白，本说明书实施例可提供为方法、系统或计算机程序产品。因此，本说明书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，本说明书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器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

[110] 本说明书可以在由计算机执行的计算机可执行指令的一般上下文中描述，例如程序模块。一般地，程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本说明书，在这些分布式计算环境中，由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中，程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

[111] 本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于系统实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

[112] 以上所述仅为本说明书实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理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权利要求范围之内。