



(12) 发明专利

(10) 授权公告号 CN 103164416 B

(45) 授权公告日 2016. 08. 03

(21) 申请号 201110412484. 7

审查员 马鑫

(22) 申请日 2011. 12. 12

(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岛资本大厦一座
四层 847 号邮箱

(72) 发明人 张正威 张立柱 杨廷辉

(74) 专利代理机构 北京同达信恒知识产权代理
有限公司 11291

代理人 郭润湘

(51) Int. Cl.

G06F 17/30(2006. 01)

(56) 对比文件

CN 102044009 A, 2011. 05. 04,

CN 101510856 A, 2009. 08. 19,

CN 101499101 A, 2009. 08. 05,

CN 102035891 A, 2011. 04. 27,

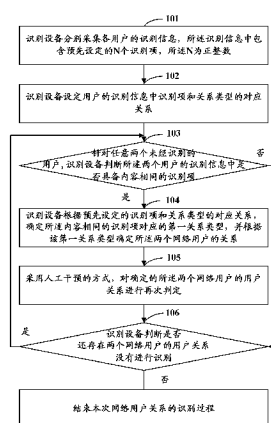
权利要求书2页 说明书12页 附图4页

(54) 发明名称

一种用户关系的识别方法及设备

(57) 摘要

本申请公开了一种用户关系的识别方法及设备,识别设备从不同渠道采集各用户的识别信息,所述识别信息中包含预先设定的若干个识别项,针对任意两个用户,所述识别设备判断该两个用户的识别信息中是否具备内容相同的识别项,若是,则识别设备根据确定的所述两个用户的识别信息中内容相同的识别项所对应的关系类型,确定所述两个用户的关系,使得所述两个用户关系的建立更为准确、真实,解决了现有网络中用户关系无法准确、真实地反映用户实际关系,从而造成网络信息安全存在较高风险的问题。



1. 一种用户关系的识别方法,其特征在于,包括:

识别设备分别采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数;

针对任意两个用户,所述识别设备判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则识别设备根据预先设定的识别项和关系类型的对应关系,确定所述内容相同的识别项对应的第一关系类型;以及,

识别设备根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值,并根据采集到的各识别项的渠道来源,按照渠道来源的真实度,调整所述N个识别项的权重值;并根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值,以及根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型;

所述识别设备根据所述第一关系类型和所述第二关系类型确定所述两个用户的关系。

2. 如权利要求1所述的用户关系的识别方法,其特征在于,所述识别设备通过以下方式设定识别项和关系类型的对应关系:

识别设备将所述N个识别项划分为多组,并为划分后的每一组设定对应的关系类型,其中,每一组内包含至少一个识别项,且任意两组内的识别项至少不完全相同;

确定所述内容相同的识别项对应的第一关系类型,具体包括:

识别设备从划分的多组中选择出至少一组,其中,选择出的组中包含的全部识别项是所述内容相同的识别项中的部分项或全部项;

识别设备根据为每一组设定的对应关系类型,确定选择出的组对应的关系类型,作为所述内容相同的识别项对应的第一关系类型。

3. 一种用户关系的识别方法,其特征在于,包括:

识别设备分别采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,且根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值,并根据采集到的各识别项的渠道来源,按照渠道来源的真实度,调整所述N个识别项的权重值;所述N为正整数;

针对任意两个用户,所述识别设备判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则识别设备根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值;

识别设备根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,并根据该第二关系类型确定所述两个用户的关系。

4. 一种用户关系的识别设备,其特征在于,包括:

数据采集模块,用于采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数;

关系识别模块,用于针对任意两个用户,判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则根据预先设定的识别项和关系类型的对应关系,确定所述内容相同的识别项对应的第一关系类型;

权重分配模块,用于根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值;并根据采集到的各识别项的渠道来源,按照渠道来源的真实度,调整所述N个识别项的权重值;

所述关系识别模块,还用于针对任意两个用户,在判断所述两个用户的识别信息中具备内容相同的识别项时,根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值,根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,并根据所述第一关系类型和所述第二关系类型确定所述两个用户的关系。

5.如权利要求4所述的用户关系的识别设备,其特征在于,还包括:

关系对应模块,用于将所述N个识别项划分为多组,并为划分后的每一组设定对应的关系类型,其中,每一组内包含至少一个识别项,且任意两组内的识别项至少不完全相同;

所述关系识别模块,具体用于从划分的多组识别项中选择出至少一组,其中,选择出的组中包含的全部识别项是所述内容相同的识别项中的部分项或全部项,并根据为每一组设定的对应关系类型,确定选择出的组对应的关系类型,作为所述内容相同的识别项对应的第一关系类型,并根据该第一关系类型确定所述两个用户的关系。

6.如权利要求4或5所述的用户关系的识别设备,其特征在于,还包括:

人工识别模块,用于提供操作界面来接收人工干预信息,利用接收到的人工干预信息对关系识别模块所得到的识别结果再次进行判定。

7.一种用户关系的识别设备,其特征在于,包括:

数据采集模块,用于采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,并根据采集到的各识别项的渠道来源,按照渠道来源的真实度,调整所述N个识别项的权重值;所述N为正整数;

权重分配模块,用于根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值;

关系识别模块,用于针对任意两个用户,判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值,根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,并根据该第二关系类型确定所述两个用户的关系。

一种用户关系的识别方法及设备

技术领域

[0001] 本申请涉及互联网技术领域,尤其涉及一种用户关系的识别方法及设备。

背景技术

[0002] 随着互联网技术的不断发展,互联网业务也得到飞速发展,用户可以利用各种互联网网站与同处互联网的另一用户进行商品的买卖以及交友活动等,从而与该用户在互联网网站上建立相应的用户关系,如交易关系、朋友关系等,将现实生活中人与人之间的关系扩展到虚拟网络领域。

[0003] 如用户利用互联网网站进行交友活动时,可以根据自身的爱好、教育水平、所处地区等选择另一用户并与之建立朋友关系,或者是具有真实朋友关系的用户使用在交友网站中的用户账号,相互呼叫建立互联网中的朋友关系。

[0004] 当用户之间建立用户关系后,可通过该用户关系扩展互联网业务。例如:用户A与用户B网络上朋友关系,且用户A注册在某一网站服务器中,当用户B也希望注册在该网站服务器时,网站服务器可根据用户A和用户B之间的朋友关系,确定用户B的安全性,允许用户B的注册。

[0005] 但由于目前互联网网站并没有提供用户身份验证的机制,无法判断各用户所提供的关于自身信息的真假,因此,上述任意两个用户之间朋友关系的建立是没有准确的依据的,可能存在网络上两个用户为朋友关系但现实生活中两者并无关系的情况,也就是说,用户在网络上的关系信息并不能反映实际中用户的关系,当网站服务器或第三方机构利用网络上用户之间的关系进行相应的互联网业务时,可能会存在由于用户关系的真实度欠缺所造成的网络信息安全性存在风险的问题。

[0006] 因此,如果能找到一种能够准确反映用户实际关系的用户关系识别方案,识别出具有真实关系的用户之间在互联网上的用户关系,则可以将所述识别出的互联网中的用户关系应用在需要较高网络信息安全性的互联网业务中,如第三方机构利用所述识别出的互联网中的用户关系,有效地分析、跟踪用户行为,得到对该用户较为准确的评价,从而进行相应的互联网业务,如贷款调查等。

发明内容

[0007] 本申请的目的在于:提供一种用户关系的识别方法及设备,用以解决现有网络中用户关系无法准确、真实地反映用户实际关系,从而造成网络信息安全性存在风险的问题。

[0008] 一种用户关系的识别方法,包括:

[0009] 识别设备分别采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数;

[0010] 针对任意两个用户,所述识别设备判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则识别设备根据预先设定的识别项和关系类型的对应关系,确定所述内容相同的识别项对应的第一关系类型,并根据该第一关系类型确定所述两个用户的关

系。

[0011] 一种用户关系的识别方法,包括:

[0012] 识别设备分别采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,且根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值,所述N为正整数;

[0013] 针对任意两个用户,所述识别设备判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则识别设备根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值;

[0014] 识别设备根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,并根据该第二关系类型确定所述两个用户的关系。

[0015] 一种用户关系的识别设备,包括:

[0016] 数据采集模块,用于采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数;

[0017] 关系识别模块,用于针对任意两个用户,判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则根据预先设定的识别项和关系类型的对应关系,确定所述内容相同的识别项对应的第一关系类型,并根据该第一关系类型确定所述两个用户的关系。

[0018] 一种用户关系的识别设备,包括:

[0019] 数据采集模块,用于采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数;

[0020] 权重分配模块,用于根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值;

[0021] 关系识别模块,用于针对任意两个用户,判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值,根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,并根据该第二关系类型确定所述两个用户的关系。

[0022] 本申请的有益效果如下:

[0023] 本申请实施例提供了一种用户关系的识别方法及设备,识别设备采集各用户的识别信息,所述识别信息中包含预先设定的若干个识别项,针对任意两个用户,所述识别设备根据所述两个用户的识别信息中内容相同的识别项所对应的关系类型,确定所述两个用户的关系,使得所述两个用户关系的建立更为准确、真实,解决了现有网络中用户关系无法准确、真实地反映用户实际关系,从而造成网络信息安全性存在较高风险的问题。

附图说明

[0024] 图1所示为本申请实施例一中用户关系的识别方法流程示意图;

[0025] 图2所示为用户A的多维网状关系示意图;

[0026] 图3所示为本申请实施例二中用户关系的识别方法流程示意图;

[0027] 图4所示为本申请实施例三中用户关系的识别设备结构示意图;

[0028] 图5所示为本申请实施例四中用户关系的识别设备结构示意图。

具体实施方式

[0029] 为实现本申请目的,本申请实施例提出了一种用户关系的识别方法及设备,识别设备采集各用户的识别信息,由于针对每个用户的识别信息中包含预先设定的N个识别项,因此,所述识别设备可以针对任意两个用户,判断这两个用户的识别信息中是否具备内容相同的识别项,若是,则根据这两个用户相同的识别项,最终确定这两个用户的关系,通过本申请的用户关系识别方案,使得任意两个用户之间的关系建立得更为准确、真实,避免了网络信息安全的高风险问题。

[0030] 下面结合说明书附图对本申请实施例的方案进行详细描述。

[0031] 实施例一

[0032] 如图1所示,为本申请实施例一中用户关系的识别方法流程示意图,所述识别方法包括以下步骤:

[0033] 步骤101:识别设备分别采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数。

[0034] 所述用户可以为在互联网网站上进行用户账号注册、从而进行相应互联网业务行为的个人、企业或组织等。

[0035] 具体地,所述识别设备至少可以通过以下两种方式来采集各用户的识别信息:

[0036] (1)采用主动挖掘的方式。

[0037] 识别设备通过用户注册或登录互联网网站时,在网站服务器中存储的字段,主动搜索并采集与各个字段匹配、相似的信息。

[0038] 例如,可以在用户登录相应网站时,采集登录时所使用的IP地址或者MAC信息等,再例如,还可以采集该用户在该网站内的用户账号信息,如登录ID,采集的所述IP地址、MAC信息、登录ID与该用户的用户账号信息相关联,成为该用户的识别信息。

[0039] (2)采用定时更新的方式。

[0040] 识别设备通过与各个识别信息渠道建立更新机制,实时监测记录了多种渠道产生的识别信息的数据库,当某一数据库中保存的识别信息发生变化时,则可通过API接口、文件包等形式,同步获取更新后的识别信息。

[0041] 例如,实时监测交易平台的数据库,当某一用户在该交易平台上进行了一次交易,则在该数据库中将更新该用户的交易信息,此时,可通过数据库的接口,同步获取数据库中该用户更新后的交易信息。所述交易信息包括:买家用户标识、卖家用户标识、交易标的物、交易金额、交易时间等与一次交易过程相关的明细。

[0042] 所述识别信息渠道包括但不限于:可以实现用户间业务交易的电子商务平台;银行网银系统、企业征信系统、个人征信系统等第三方网站或平台;即时聊天工具、在线交易软件;工商单位、民事单位、司法单位等各种政府、民间机构等。

[0043] 所述识别信息是用于识别用户的信息,包含一个或多个识别项。针对个人用户,所述识别信息可以为身份证号码、姓名、登录ID、聊天ID、电子邮箱、家庭电话、交易信息、担保信息、借贷信息等,其中身份证号码、登录账号等即为识别信息中的识别项。

[0044] 针对企业用户,所述识别信息可以为机构代码、企业名称、企业类型、客户资源、企业邮箱、企业地址等。

[0045] 在进行识别信息的采集之前,识别设备可预先设定一个或多个用于进行用户分析的识别项,这样做的目的为:由于表示用户的识别信息可以有多种,若针对所有的识别项来进行用户分析的话,可能会使得分析的数据量过大,同时也不一定利于用户关系的识别,因此,可以从大量的识别项中选择部分比较有代表性的识别项,一方面可以减少分析的数据量,另一方面,这种精准的分析也能够提高用户关系确定的准确性。

[0046] 例如:识别设备可按照表1所示,预先设定多个识别项的列表,并将采集的识别信息与所述列表中的各识别项进行匹配,得到可用于进行关系分析的标准格式的识别信息。

[0047]

用户标识 识别信息		
识别项	身份证号码	
	姓名	
	登录 ID	
	MAC 地址	
	电子邮箱	
	交易信息	
	借贷信息	
	家庭电话	
	担保信息	
	IP 地址	
	...	

[0048] 表1

[0049] 假设采集的个人用户A的识别信息为:姓名、家庭电话、交易信息、登录ID、MAC地址,可得到该用户A的识别信息如表2所示:

[0050]

用户标识 识别信息		用户 A
	身份证号码	
	姓名	张三
	登录 ID	zhangsan01

[0051]

识别项	MAC 地址	00-11
	电子邮箱	
	交易信息	买家 A、卖家 B、交易金额 C 元等
	借贷信息	
	家庭电话	020-1234
	担保信息	
	IP 地址	
	...	...

[0052] 表2

[0053] 识别设备采集的识别信息中并不一定包含预先设定的所有N个识别项,也就是说,若将预设的N个识别项以表1的形式设置,针对每次采集的用户的识别信息,将识别信息中的内容填充至表1所示的数据表中时,可能存在某些识别项中有内容,某些识别项中内容为空的情况,如表2所示。

[0054] 当针对某次采集的用户的识别信息,不包含表1所示的某些识别项的内容,即出现表2所示的某些识别项为空的情况时,识别设备可以采取主动获取的方式或定时更新的方式,重复多次采集该用户的识别信息,以期使采集的识别信息中尽可能地包含所述N个识别项的内容,使得表2所示的识别信息达到相对完整的状态。

[0055] 需要说明的是,由于识别信息是从不同的识别信息渠道采集获得,可能出现针对同一用户的相同名称的识别项出现不同内容的情况,例如:通过第一渠道的数据库采集到用户A的家庭电话与通过第二渠道的数据库采集到的用户A的家庭电话不同的情况,本实施例包括但不限于以下两种做法:

[0056] 第一种做法:识别设备根据识别信息采集的来源不同,为不同来源的识别信息分配权重,当针对同一用户的相同名称的识别项出现不同内容的情况时,确定不同内容所属来源的权重,保留权重较高的来源提供的内容,同时标明所述识别项内容的渠道来源。

[0057] 第二种做法:识别设备将不同来源的针对同一识别项的不同的内容,都作为该识别项的内容保留,同时标明所述识别项内容的渠道来源。

[0058] 在采集到各用户的识别信息后,所述识别设备可将该识别信息以表2所示的数据表的形式存储至数据库中,较优地,可为存储的识别信息分配标识,通过该分配的标识不仅可以从数据库中查询出各用户的识别信息,还可以识别出该用户是个人用户还是企业用户等。例如:针对表2所示的个人用户A的识别信息分配的标识可为A_001,其中,A表示对应的识别信息是个人用户的识别信息,001为个人用户中的识别信息的序列号。再例如,针对某一企业用户的识别信息分配的标识可为B_001,其中,B表示对应的识别信息是企业用户的识别信息,001为企业用户中的识别信息的序列号。

[0059] 另外,为了使本申请中采集的识别信息能够方便地进行数据库存储并进行分析,采集的识别信息中各识别项的指代需一致。例如:针对表1中名称为姓名的识别项,不论从

何处采集的识别信息,针对姓名这一识别项,其内容都应当表示用户的姓名。

[0060] 需要说明的是,在本步骤101中,由于各用户在不同互联网网站或平台已注册的用户账号可能有多个,且所述同一用户的多个用户账号可能并不相同,因此,识别设备在进行用户的识别信息的采集时,可以先确定同一用户所注册的各个用户账号,具体可以采用以下步骤来确定各个用户账号是否为同一用户所注册:

[0061] 步骤一:识别设备采集各用户账号的识别信息,所述识别信息中包含设定的关键字段和模型字段。

[0062] 所述关键字段表示能够唯一或是接近唯一地识别用户账号信息注册者的身份的识别信息,例如,针对个人用户,关键字段的名称可以为身份证号码、护照号码等,针对身份证号码这一关键字段,其内容可以为用户在注册用户账号时填写的身份证号码。

[0063] 所述模型字段为识别信息中的非关键字段,例如,针对个人用户,模型字段的名称可以为性别、出生日期、国籍、电子邮箱、教育水平,针对出生日期这一模型字段,其内容可以为用户在注册用户账号时填写的出生日期。

[0064] 步骤二:针对任意两个用户账号的识别信息,所述识别设备判断所述两个用户账号的识别信息中是否有至少一个关键字段的内容相同;若是,则确定该两个用户账号是由同一用户注册的;否则,确定识别信息中的模型字段的内容相似度,判断所述两个用户账号的识别信息中的模型字段的内容相似度是否达到设定条件,若是,则确定该两个用户账号是由同一用户注册的,否则,确定该两个用户账号是由不同用户注册的。

[0065] 具体地,所述识别设备确定识别信息中的模型字段的内容相似度包括以下方法:

[0066] 方法一:

[0067] 识别设备确定各模型字段的权重值,分别将两个用户账号的识别信息中名称相同的模型字段的内容进行相似度运算,并将各运算结果与对应的模型字段的权重值进行加权求和,得到模型字段的内容相似度。

[0068] 方法二:

[0069] 识别设备确定识别信息中的注册信息、从第三方获取的信息和用户账号运行时产生的信息的权重值,分别将两个用户账号的识别信息中名称相同的模型字段的内容进行相似度运算,并确定进行相似度运算的内容所属信息的权重值;将各运算结果与进行相似度运算的内容所属信息的权重值进行加权求和,得到模型字段的内容相似度。

[0070] 识别设备通过上述步骤识别出同一用户所注册的多个用户账号后,可以按照该用户所对应的多个用户账号来进行所述用户识别信息的采集,确保针对每一用户所采集到的识别信息更加全面。

[0071] 另外,所述识别设备也可以将不同互联网网站或平台上已注册的各用户账号看成是与用户为一一对应关系,即一个用户账号代表一个用户。

[0072] 步骤102:识别设备设定用户的识别信息中识别项和关系类型的对应关系。

[0073] 具体地,所述识别设备可以通过以下方式设定所述识别项和关系类型的对应关系:

[0074] 识别设备将所述N个识别项划分为多组,并为划分后的每一组设定对应的关系类型,其中,每一组内包含至少一个识别项,且任意两组内的识别项至少不完全相同。

[0075] 具体地,假设预先设定3个识别项,分别为识别项A、识别项B和识别项C,则可将所

述识别项划分为7组:组1为识别项A;组2为识别项B;组3为识别项C;组4为识别项A、B;组5为识别项A、C;组6为识别项B、C;组7为识别项A、B、C。

[0076] 识别设备可为上述7组分别设定对应的关系类型,也可只为其中部分组分别设定对应的关系类型。针对设定了关系类型的组,一组对应一个关系类型,多组可对应同一关系类型。

[0077] 例如:可将组7设定对应第一维关系,组5设定对应第二维关系,组2设定对应第三维关系。由于组7内的识别项包含组5内的识别项且多于组5内的识别项,因此,第一维关系所表示的用户之间的相关性高于第二维关系所表示的用户之间的相关性;同理,第二维关系所表示的用户之间的相关性高于第三维关系所表示的用户之间的相关性。例如:第一维关系表示两个用户之间的相关性最高,如用户自身;第二维关系表示两个用户之间的相关性较高,但低于第一维关系表示的用户之间的相关性,如家庭成员、朋友等;第三维关系表示两个用户之间的相关性较低,如朋友的朋友、交易伙伴等。

[0078] 优选地,识别设备可根据识别项所能够反映的用户之间具有的相应关系,进一步设定包含识别项的组对应的关系类型,如家庭成员类型、朋友类型、交易者类型等。

[0079] 例如:若识别项A、B、C分别是反映用户之间交易的信息,如:交易平台ID、交易时间、交易标的物、交易价格等,则若两个用户的识别信息中,针对反映用户之间交易信息的识别项A、B、C都相同,则可确定这两个用户的关系为交易者,此时,组7(识别项A、B、C)对应的关系类型为交易伙伴类型。

[0080] 步骤103:针对任意两个未经识别的用户,识别设备判断所述两个用户的识别信息中是否具备内容相同的识别项,若是则进行步骤104,否则进行步骤106。

[0081] 具体地,可以采用精确匹配或者模糊匹配的方式来逐个比较采集到的所述两个用户的识别信息中名称相同的各识别项的内容是否相同,并以数据表的形式存储内容相同的所述识别项。

[0082] 步骤104:识别设备根据预先设定的识别项和关系类型的对应关系,确定所述内容相同的识别项对应的第一关系类型,并根据该第一关系类型确定所述两个用户的关系。

[0083] 具体地,识别设备可以通过以下步骤确定内容相同的识别项对应的第一关系类型:

[0084] 第一步:识别设备根据步骤103中确定的内容相同的识别项,从步骤102中划分的多个组中选择出至少一组,其中,选择出的组中包含的全部识别项是所述内容相同的识别项中的部分项或全部项。

[0085] 具体地,假设针对用户A和用户B,确定的所述内容相同的识别项为识别项A和识别项C,仍以步骤102中划分的7个组为例,可选择组5、组1、组3中任意一组或多组。

[0086] 第二步:识别设备根据为每一组设定的对应关系类型,确定选择出的组对应的关系类型,作为所述内容相同的识别项对应的第一关系类型。

[0087] 假设在第一步中选择了多组,且多组对应多个关系类型,则可将任一关系类型作为所述第一关系类型;或者,将对应的多个关系类型中,表示关系相关性最高的关系类型作为第一关系类型。

[0088] 识别设备在确定所述内容相同的识别项对应的第一关系类型后,则可以根据该第一关系类型确定所述两个用户的关系,如上述第二步中,若确定第一关系类型为家庭成员

类型,则可确定两个用户的关系为家庭成员。

[0089] 进一步地,为了更加准确地识别所述两个用户的关系,在本申请实施例一的方案中,识别设备还可以通过以下方式确定两个用户之间的第二关系类型,进而结合第一关系类型确定两个用户之间的最终关系。

[0090] 具体地,识别设备可以通过以下步骤确定两个用户之间的第二关系类型:

[0091] 步骤一:识别设备根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值。

[0092] 具体地,对于表示用户之间的相关性越高的识别项,识别设备为其分配的权重也就越高,较优地,本实施例中可以采用模型算法来确定各识别项的权重值,具体可以采用层次分析模型的方式,通过建立层次结构模型、构造成对比较阵、计算权向量并做一致性检验的步骤来得到各识别项的权重值。

[0093] 进一步地,识别设备还可以根据采集到的各识别项的渠道来源,按照渠道来源的真实度,调整所述N个识别项的权重值,对于渠道来源真实度越高的识别项,为其分配的权重值也越高。

[0094] 步骤二:识别设备根据步骤103确定的内容相同的识别项,利用所述识别项的权重值,确定所述两个用户之间的关系值。

[0095] 具体地,可以将所述内容相同的识别项的权重值进行相加求和,按照权重值之和的大小来确定所述两个用户之间的关系值,例如:权重值之和越大,两个用户之间的关系值越大。

[0096] 步骤三:识别设备根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型。

[0097] 具体地,识别设备可以按照关系值的高低来设定各关系值所对应的关系类型,如关系值越高,则其对应的关系类型所表示的用户的相关性也越高。

[0098] 步骤四:识别设备根据确定的所述第一关系类型和所述第二关系类型确定所述两个用户的关系。

[0099] 具体地,可以将所述第一关系类型和所述第二关系类型均作为确定的所述两个用户的关系,也可以选择其中相关性较低的关系类型作为所述两个用户的最终关系。

[0100] 步骤105:采用人工干预的方式,对确定的所述两个用户的用户关系进行再次判定。

[0101] 较优地,为了保证判断结果的可靠性,还可以对步骤104的结果进一步采用人工干预的方式再次进行判定,本实施例的方案也不限于:若人工干预的方式再次判定的结果与步骤104的确定结果不同时,以人工干预的结果为准。

[0102] 在步骤105中若进行过人工干预操作,则识别设备可以通过日志信息的方式记录人工干预的操作过程以及操作结果。

[0103] 步骤106:识别设备判断是否还存在两个用户的用户关系没有进行识别,若是,则跳转至步骤103,否则,结束本次用户关系的识别过程。

[0104] 当识别设备识别出网络中任意两个用户的关系后,可以建立一个针对任一用户的关系网,图2所示即为针对用户A的多维网状关系展示图,对于用户A来说,按照确定的其他用户与用户A的关系相关性的,建立了一个针对用户A的三维关系网,处在第一维关系

的可以为用户A自身,或者经过用户关系识别确定的与用户A为同一用户的用户,处在第二维关系的可以为与用户A为家庭成员、朋友等关系的用户,处在第三维关系的可以为与用户A的朋友为朋友关系等的用户、与用户A为交易伙伴关系的用户等。

[0105] 通过本申请实施例一的方案,识别设备可以采用较低的成本、快速采集到各识别信息渠道中更新后的用户的识别信息,降低了大量用户识别信息的获取成本;针对任意两个用户,识别设备根据确定的所述两个用户的识别信息中内容相同的识别项所对应的关系类型,可以识别出任意两个用户之间的关系;利用识别出的所述用户的关系,识别设备可以建立一个针对任一用户的更为准确、真实的用户关系网,并且可以将所述用户的关系网应用在需要较高网络信息安全性的互联网业务中,如第三方机构利用所述用户的关系网,有效地分析、跟踪所述用户的行为,得到对该用户较为准确的评价,从而进行相应的互联网业务,如贷款调查等。

[0106] 实施例二

[0107] 在实施例一的方案中,识别设备是根据预先设定的识别项和关系类型的对应关系,来确定两个用户的第一关系类型的,在本实施例二的方案中,提出了根据各识别项的权重值来确定两个用户的第二关系类型的方案,如图3所示,为本实施例二中用户关系的识别方法流程示意图,包括以下步骤:

[0108] 步骤201:识别设备分别采集各用户的识别信息,所述识别信息中包含N个识别项。

[0109] 本步骤201与实施例一中的步骤101的方案相同,此处不再赘述。

[0110] 步骤202:识别设备根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值。

[0111] 步骤203:针对任意两个未经识别的用户,所述识别设备判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则执行步骤203,否则,执行步骤207。

[0112] 本步骤203与实施例一中的步骤103的方案相同,此处不再赘述。

[0113] 步骤204:识别设备根据确定的内容相同的识别项,利用所述识别项的权重值,确定所述两个用户之间的关系值。

[0114] 具体地,可以将所述内容相同的识别项的权重值进行相加求和,按照权重值之和的大小来确定所述两个用户之间的关系值,例如:权重值之和越大,两个用户之间的关系值越大。

[0115] 步骤205:识别设备根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,根据确定的所述第二关系类型确定所述两个用户的关系。

[0116] 具体地,识别设备可以按照关系值的高低来设定各关系值所对应的关系类型,如关系值越高,则为其设定的对应的关系类型所表示的用户的相关性也越高。

[0117] 步骤206:采用人工干预的方式,对确定的所述两个用户的用户关系进行再次判定。

[0118] 本步骤206与实施例一中的步骤105的方案相同,此处不再赘述。

[0119] 步骤207:识别设备判断是否还存在两个用户的用户关系没有进行识别,若是,则跳转至步骤203,否则,结束本次用户关系的识别过程。

[0120] 实施例三

[0121] 如图4所示,为与本申请实施例一属于同一发明构思下的用户关系的识别设备结构示意图,所述用户关系的识别设备包括数据采集模块11和关系识别模块12。

[0122] 所述数据采集模块11用于采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数。

[0123] 具体地,所述数据采集模块11至少可以通过以下两种方式来采集各用户的识别信息:

[0124] (1)采用主动挖掘的方式。

[0125] 通过用户注册或登录互联网网站时,在网站服务器中存储的字段,主动搜索并采集与各个字段匹配、相似的信息。

[0126] 例如,可以在用户登录相应网站时,采集登录时所使用的IP地址或者MAC信息等,再例如,还可以采集该用户在该网站内的用户账号信息,如登录ID,采集的所述IP地址、MAC信息、登录ID与该用户的用户账号信息相关联,成为该用户的识别信息。

[0127] (2)采用定时更新的方式。

[0128] 通过与各个识别信息渠道建立更新机制,实时监测记录了多种渠道产生的识别信息的数据库,当某一数据库中保存的识别信息发生变化时,则可通过API接口、文件包等形式,同步获取更新后的识别信息。

[0129] 例如,实时监测交易平台的数据库,当某一用户在该交易平台上进行了一次交易,则在该数据库中将更新该用户的交易信息,此时,可通过数据库的接口,同步获取数据库中该用户更新后的交易信息。所述交易信息包括:买家用户标识、卖家用户标识、交易标的物、交易金额、交易时间等与一次交易过程相关的明细。

[0130] 所述识别信息渠道包括但不限于:可以实现用户间业务交易的电子商务平台;银行网银系统、企业征信系统、个人征信系统等第三方网站或平台;即时聊天工具、在线交易软件;工商单位、民事单位、司法单位等各种政府、民间机构等。

[0131] 所述识别信息是用于识别用户的信息,包含一个或多个识别项。针对个人用户,所述识别信息可以为身份证号码、姓名、登录ID、聊天ID、电子邮箱、家庭电话、交易信息、担保信息、借贷信息等,其中身份证号码、登录账号等即为识别信息中的识别项。

[0132] 针对企业用户,所述识别信息可以为机构代码、企业名称、企业类型、客户资源、企业邮箱、企业地址等。

[0133] 所述关系识别模块12用于针对任意两个用户,判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则根据预先设定的识别项和关系类型的对应关系,确定所述内容相同的识别项对应的第一关系类型,并根据该第一关系类型确定所述两个用户的关系。

[0134] 具体地,所述用户的识别设备还包括关系对应模块13,所述关系对应模块13用于将所述N个识别项划分为多组,并为划分后的每一组设定对应的关系类型,其中,每一组内包含至少一个识别项,且任意两组内的识别项至少不完全相同。

[0135] 具体地,所述关系识别模块12用于从划分的多组识别项中选择出至少一组,其中,选择出的组中包含的全部识别项是所述内容相同的识别项中的部分项或全部项,并根据为每一组设定的对应关系类型,确定选择出的组对应的关系类型,作为所述内容相同的识别项对应的第一关系类型,并根据该第一关系类型确定所述两个用户的关系。

[0136] 优选地,所述用户的识别设备还可以包括权重分配模块14,所述权重分配模块14用于根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值;具体地,对于表示用户之间的相关性越高的识别项,其分配的权重也就越高,较优地,还可以采用模型算法的方式来确定各识别项的权重值,进一步地,还可以根据采集到的各识别项的渠道来源,按照渠道来源的真实度,调整所述N个识别项的权重值,对于渠道来源真实度越高的识别项,为其分配的权重值也越高。

[0137] 具体地,所述关系识别模块12还用于在针对任意两个用户,判断所述两个用户的识别信息中具备内容相同的识别项后,根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值,根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,并根据所述第一关系类型和所述第二关系类型确定所述两个用户的关系。

[0138] 较优地,所述用户关系的识别设备还可以包括数据存储模块15和人工识别模块16。

[0139] 所述数据存储模块15用于存储数据采集模块11所采集到的各用户的识别信息;较优地,可为存储的识别信息分配标识,通过该分配的标识不仅可以从数据存储模块15中查询出各用户的识别信息,还可以识别出该用户是个人用户还是企业用户等。

[0140] 所述人工识别模块16用于提供操作界面来接收人工干预信息,利用接收到的人工干预信息对关系识别模块12所得到的识别结果再次进行判定,若人工干预的方式再次判定的结果与关系识别模块12的确定结果不一致时,以人工干预的结果为准。

[0141] 较优地,当采用人工干预的方式对识别结果再次进行判定时,可以设定多名网络管理员并为其分配唯一编号和相应的权限,同时需要记录各网络管理员操作过程中的MAC地址和IP地址,在本IP段内允许所述网络管理员行使人工干预权限,非本IP段则提示出错,以保证识别信息的安全。

[0142] 较优地,所述数据存储模块15还用于将确定的任意两个用户的关系识别结果进行存储,如可以采用数据表或关系网等方式进行存储,同时采用日志信息的方式记录人工识别模块16对本地存储的任意两个用户的关系识别结果进行的相应操作结果和操作过程。

[0143] 较优地,所述数据存储模块15还可以提供用户关系识别结果查询接口和导出接口给其它系统;通过所述查询接口,可以查询识别出的任意两个用户的关系,并可以通过多维关系网展示图的方式来显示。

[0144] 较优地,所述用户关系的识别设备还可以具有识别结果自学习功能,如:当采用人工干预的方式,将用户关系识别设备识别出的100个A关系均更改为B关系时,则识别设备可以对遇到的下一个类似的关系直接定为B关系。

[0145] 实施例四

[0146] 如图5所示,为与本申请实施例二属于同一发明构思下的用户关系的识别设备结构示意图,所述用户关系的识别设备包括数据采集模块11、关系识别模块12和权重分配模块14。

[0147] 所述数据采集模块11用于采集各用户的识别信息,所述识别信息中包含预先设定的N个识别项,所述N为正整数;所述权重分配模块14用于根据识别项所表示的用户之间的相关性,为所述N个识别项分别设定权重值;所述关系识别模块12用于针对任意两个用户,

判断所述两个用户的识别信息中是否具备内容相同的识别项,若是,则根据所述内容相同的识别项的权重值,确定所述两个用户之间的关系值,根据关系值区间与关系类型的对应关系,确定该关系值所在的关系值区间对应的第二关系类型,并根据该第二关系类型确定所述两个用户的关系。

[0148] 较优地,所述用户关系的识别设备还可以包括数据存储模块15和人工识别模块16。

[0149] 所述数据存储模块15用于存储数据采集模块11所采集到的各用户的识别信息;较优地,可为存储的识别信息分配标识,通过该分配的标识不仅可以从数据存储模块15中查询出各用户的识别信息,还可以识别出该用户是个人用户还是企业用户等。

[0150] 所述人工识别模块16用于提供操作界面来接收人工干预信息,利用接收到的人工干预信息对关系识别模块12所得到的识别结果再次进行判定,若人工干预的方式再次判定的结果与关系识别模块12的确定结果不一致时,以人工干预的结果为准。

[0151] 较优地,所述数据存储模块15还用于将确定的任意两个用户的关系识别结果进行存储,如可以采用数据表或关系网等方式进行存储,同时采用日志信息的方式记录人工识别模块16对本地存储的任意两个用户的关系识别结果进行的相应操作结果和操作过程。

[0152] 较优地,所述数据存储模块15还可以提供用户关系识别结果查询接口和导出接口给其它系统;通过所述查询接口,可以查询识别出的任意两个用户的关系,并可以通过多维关系网展示图的方式来显示。

[0153] 显然,本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样,倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内,则本申请也意图包含这些改动和变型在内。

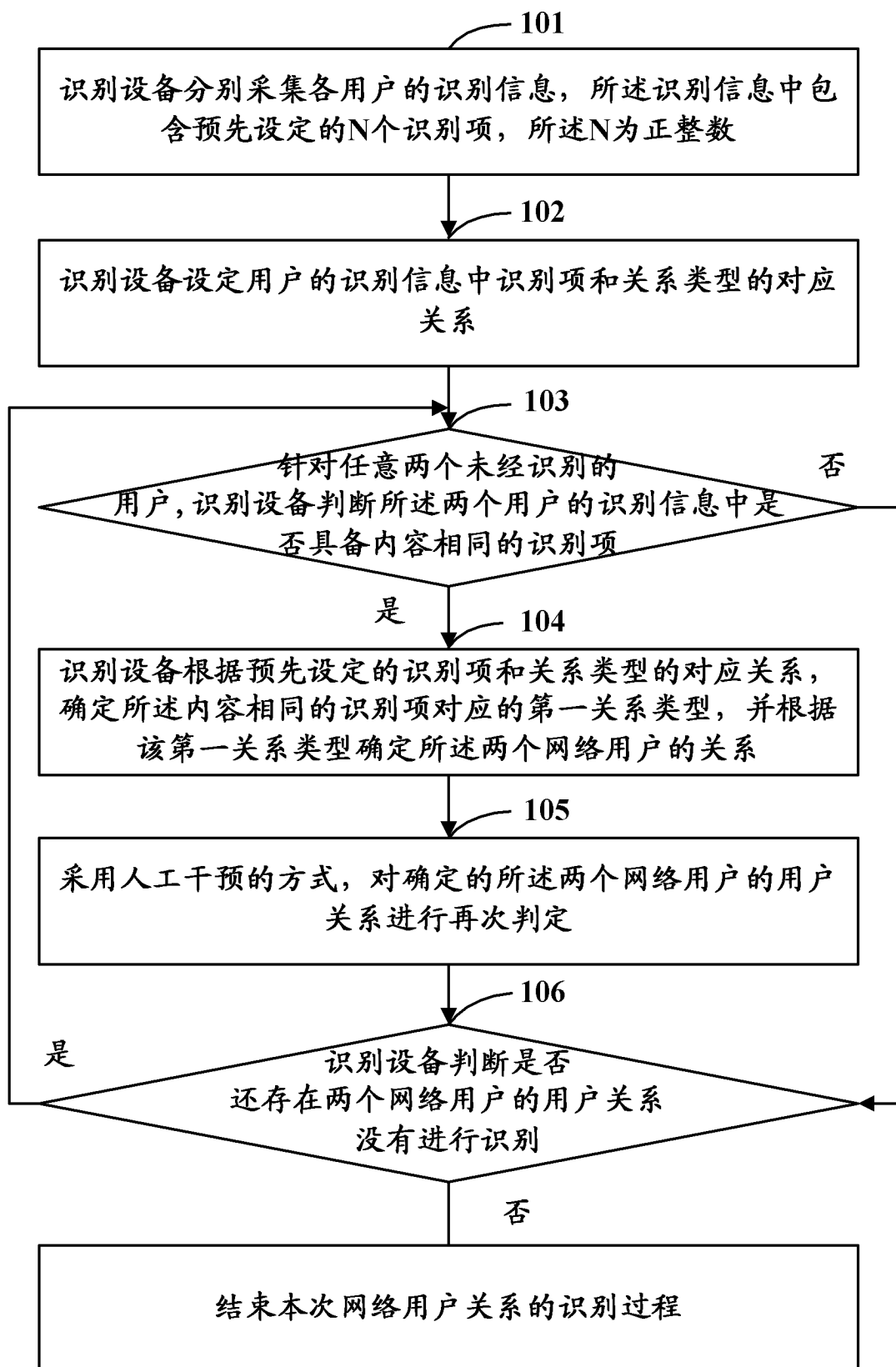


图1

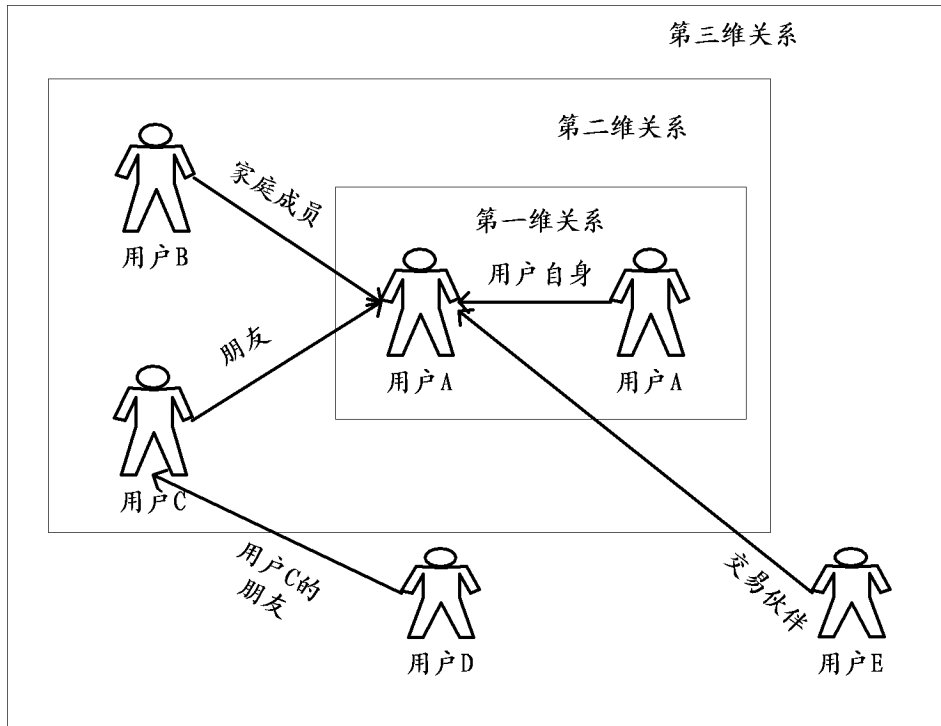


图2

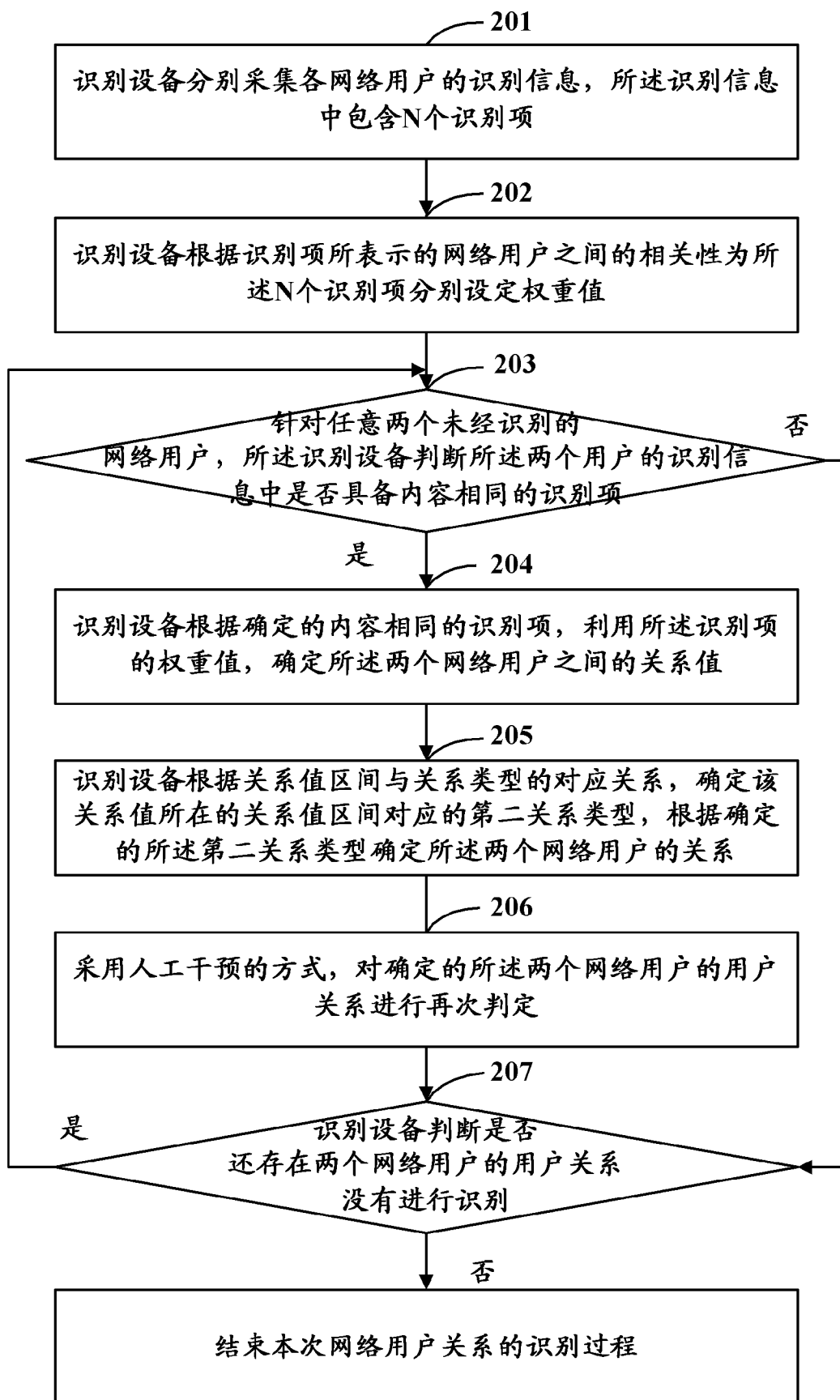


图3

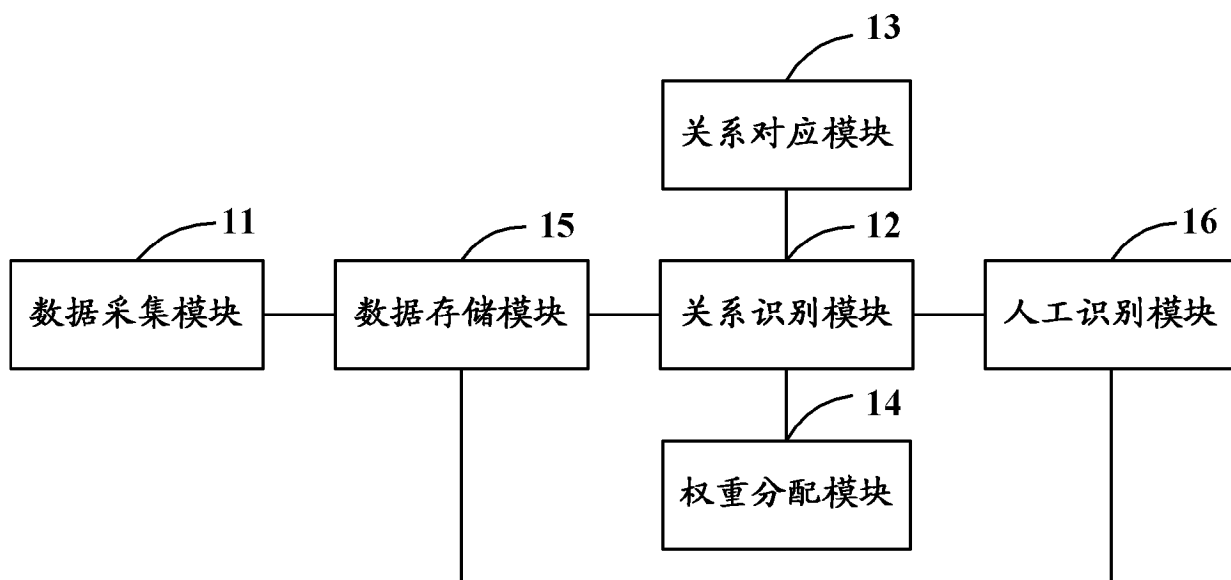


图4

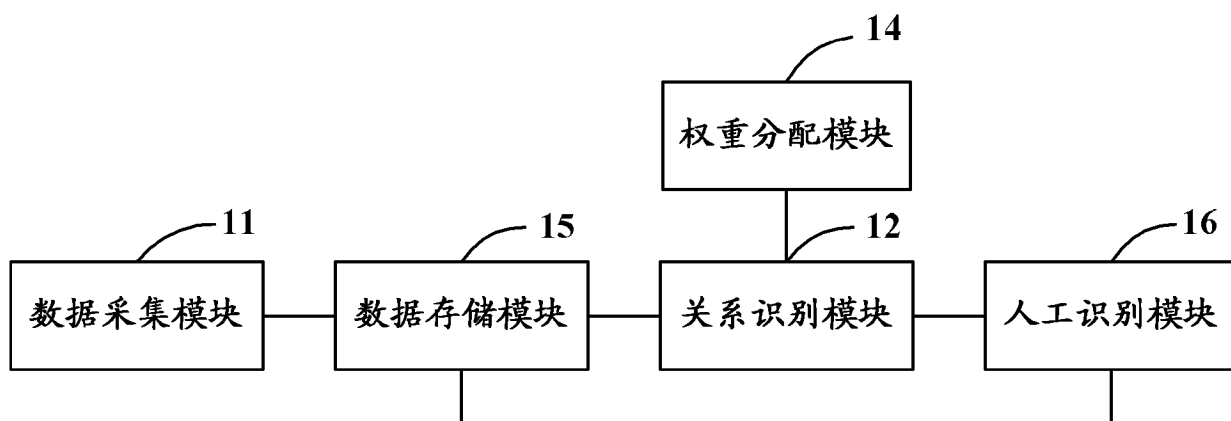


图5