



(12) 发明专利

(10) 授权公告号 CN 111666346 B

(45) 授权公告日 2024. 08. 16

(21) 申请号 201910167233.3

G06Q 40/04 (2012.01)

(22) 申请日 2019.03.06

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 103294818 A, 2013.09.11

申请公布号 CN 111666346 A

CN 107798541 A, 2018.03.13

CN 108681936 A, 2018.10.19

(43) 申请公布日 2020.09.15

审查员 郭坚

(73) 专利权人 京东科技控股股份有限公司

地址 100176 北京市大兴区北京经济技术

开发区科创十一街18号C座2层221室

(72) 发明人 周石磊

(74) 专利代理机构 北京品源专利代理有限公司

11332

专利代理师 孟金喆

(51) Int. Cl.

G06F 16/28 (2019.01)

G06Q 20/40 (2012.01)

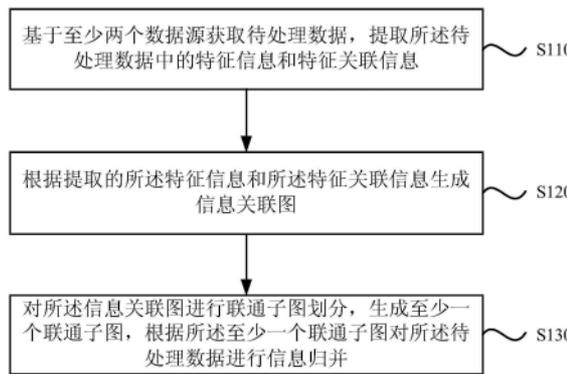
权利要求书3页 说明书13页 附图4页

(54) 发明名称

信息归并方法、交易查询方法、装置、计算机及存储介质

(57) 摘要

本发明实施例公开了信息归并方法、交易查询方法、装置、计算机及存储介质。其中信息归并方法包括基于至少两个数据源获取待处理数据，提取待处理数据中的特征信息和特征关联信息；根据提取的特征信息和特征关联信息生成信息关联图；对信息关联图进行联通子图划分，生成至少一个联通子图，根据至少一个联通子图对待处理数据进行信息归并。通过将待处理数据中的特征信息通过关联关系形成信息关联图，并基于信息关联图中特征节点的联通性对信息关联图进行划分，得到多个彼此独立的联通子图，基于联通子图将特征信息进行归并，通过图形方式简化了信息归并，方便直观，解决了数据库中海量数据无法清晰判断数据关联关系的问题，提高了图像归并效率。



1. 一种信息归并方法,其特征在于,包括:

基于至少两个数据源获取待处理数据,提取所述待处理数据中的特征信息和特征关联信息;

根据提取的所述特征信息和所述特征关联信息生成信息关联图,其中,所述信息关联图由特征节点和连接特征节点的关联边组成,每一个所述特征节点上具有一个特征信息,具有关联关系的两个所述特征信息对应的两个所述特征节点之间设置有关联边,所述关联边用于连接上述两个特征节点;

对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至少一个联通子图对所述待处理数据进行信息归并,其中,所述联通子图中的任意两个特征节点通过一个或多个关联边联通,且任意两个联通子图中的任意特征节点之间不存在关联边;

在对所述信息关联图进行联通子图划分,生成至少一个联通子图之后,还包括:

通过生成的联通子图中的任意特征节点的特征信息与历史联通子图进行匹配,通过所述生成的联通子图更新匹配成功的历史联通子图,其中,更新方式是遍历所述生成的联通子图中与第一特征节点具有关联边的第二特征节点,确定所述历史联通子图中是否具有所述第二特征节点,若否,则设置所述第二特征节点,并设置所述第一特征节点和第二特征节点的关联边,若是,则确定所述历史联通子图中第一特征节点和第二特征节点是否设置关联边,如果有,则无需进行更新,如果没有,则设置第一特征节点和第二特征节点的关联边;

相应的,根据所述至少一个联通子图对所述待处理数据进行信息归并,包括:

根据所述至少一个更新联通子图对所述待处理数据进行信息归并;

其中,所述特征信息包括账号、会员信息、银行卡信息、证件信息、手机号信息、设备信息和WIFI信息中的至少一项;

所述特征节点包括账号节点、会员信息节点、银行卡信息节点、证件信息节点、手机号信息节点、设备信息节点和WIFI信息节点中的至少一项,并通过所述特征信息的关联关系,将所述特征节点之间设置关联边。

2. 根据权利要求1所述的方法,其特征在于,提取所述待处理数据中的节点信息和节点关联信息,包括:

根据预设特征标识在所述待处理数据中进行匹配,确定所述预设特征标识对应的特征信息;

遍历所述待处理数据,确定任意两个特征信息之间的关联关系。

3. 根据权利要求1所述的方法,其特征在于,根据提取的所述特征信息和所述特征关联信息生成信息关联图,包括:

根据所述特征信息设置特征节点;

根据所述特征关联信息将存在关联关系的两个特征节点间设置关联边,生成所述信息关联图。

4. 根据权利要求3所述的方法,其特征在于,对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至少一个联通子图对所述待处理数据进行信息归并,包括:

遍历所述信息关联图中的特征节点,将基于关联边连接的特征节点划分至同一联通子图,其中,任意两个联通子图中的任意特征节点之间不存在关联边;

将所述同一联通子图中特征节点对应的特征信息归并为同一群组信息。

5. 根据权利要求1-4任一所述的方法,其特征在于,所述数据源包括电商平台的用户行为数据源,相应的,所述信息关联图为电商平台的用户信息关联图,所述联通子图为具有关联关系的用户信息集合。

6. 一种交易查询方法,其特征在于,包括:

获取已知风险用户信息,根据所述风险用户信息在至少一个联通子图中进行匹配,确定与所述已知风险用户信息相匹配的目标联通子图,其中,所述至少一个联通子图为根据权利要求1-5任一所述信息归并方法确定的;

提取所述目标联通子图中关联用户信息;

确定所述关联用户信息的当前交易,将所述关联用户信息的当前交易确定为风险交易。

7. 根据权利要求6所述的方法,其特征在于,根据所述风险用户信息在至少一个联通子图中进行匹配,确定与所述已知风险用户信息相匹配的目标联通子图,包括:

根据预设特征标识提取所述已知风险用户信息中的风险特征信息;

将所述风险特征信息与至少一个联通子图中的特征节点的特征信息进行匹配,当匹配成功时,将匹配成功的特征节点所属的联通子图确定为目标联通子图。

8. 根据权利要求6所述的方法,其特征在于,所述联通子图的每一个特征节点包括至少一个行为属性信息;相应的,当根据所述已知风险用户信息匹配失败时,获取所述已知风险用户的行为特征信息,将所述行为特征信息与所述联通子图中的行为属性信息进行匹配,根据匹配结果确定目标联通子图。

9. 一种信息归并装置,其特征在于,包括:

信息提取模块,用于基于至少两个数据源获取待处理数据,提取所述待处理数据中的特征信息和特征关联信息;

信息关联图生成模块,用于根据提取的所述特征信息和所述特征关联信息生成信息关联图,其中,所述信息关联图由特征节点和连接特征节点的关联边组成,每一个所述特征节点上具有一个特征信息,具有关联关系的两个所述特征信息对应的两个所述特征节点之间设置有关联边,所述关联边用于连接上述两个特征节点;

信息归并模块,用于对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至少一个联通子图对所述待处理数据进行信息归并,其中,所述联通子图中的任意两个特征节点通过一个或多个关联边联通,且任意两个联通子图中的任意特征节点之间不存在关联边;

更新联通子图确定模块,用于在对所述信息关联图进行联通子图划分,生成至少一个联通子图之后,

通过生成的联通子图中的任意特征节点的特征信息与历史联通子图进行匹配,通过所述生成的联通子图更新匹配成功的历史联通子图,其中,更新方式是遍历所述生成的联通子图中与第一特征节点具有关联边的第二特征节点,确定所述历史联通子图中是否具有所述第二特征节点,若否,则设置所述第二特征节点,并设置所述第一特征节点和第二特征节点的关联边,若是,则确定所述历史联通子图中第一特征节点和第二特征节点是否设置关联边,如果有,则无需进行更新,如果没有,则设置第一特征节点和第二特征节点的关联边;

相应的,所述信息归并模块用于:根据所述至少一个更新联通子图对所述待处理数据

进行信息归并；

其中,所述特征信息包括账号、会员信息、银行卡信息、证件信息、手机号信息、设备信息和WIFI信息中的至少一项；

所述特征节点包括账号节点、会员信息节点、银行卡信息节点、证件信息节点、手机号信息节点、设备信息节点和WIFI信息节点中的至少一项,并通过所述特征信息的关联关系,将所述特征节点之间设置关联边。

10.一种交易查询装置,其特征在于,包括:

第一目标联通子图确定模块,用于获取已知风险用户信息,根据所述风险用户信息在至少一个联通子图中进行匹配,确定与所述已知风险用户信息相匹配的目标联通子图,其中,所述至少一个联通子图为根据权利要求1-5任一所述信息归并方法确定的;

关联用户信息确定模块,用于提取所述目标联通子图中关联用户信息;

风险交易确定模块,用于确定所述关联用户信息的当前交易,将所述关联用户信息的当前交易确定为风险交易。

11.一种计算机设备,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,其特征在于,所述处理器执行所述程序时实现如权利要求1-5中任一所述的信息归并方法。

12.一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1-5中任一所述的信息归并方法。

13.一种计算机设备,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,其特征在于,所述处理器执行所述程序时实现如权利要求6-8中任一所述的交易查询方法。

14.一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求6-8中任一所述的交易查询方法。

信息归并方法、交易查询方法、装置、计算机及存储介质

技术领域

[0001] 本发明实施例涉及数据处理技术,尤其涉及一种信息归并方法、交易查询方法、装置、计算机及存储介质。

背景技术

[0002] 随着互联网技术的不断发展以及电商平台的兴起,在电商平台上的团伙欺诈越来越多,黑产规模也越来越大。

[0003] 在电商平台上,同一用户具有多个账号的现象十分常见,一种情况是用户的正常网络活动,通过多个账号满足用户的活动需求;另一种情况是不法分子通过大量站好进行非法牟利,例如刷单、黑产或者洗钱等。为了提高电商平台的安全性判断多个交易个体是否为同一用户、多个欺诈行为是否为同一用户操作以及多个账号是否属于同一欺诈团体,在风控反欺诈中越来越重要。

[0004] 针对上述问题,目前常采用如下方式进行判断,其一,基于业务数据,根据固定的判断规则进行匹配,根据匹配结果判断不同账号是否属于同一用户,例如可以是将身份证号和注册手机号相同的账号确定为同一用户的账号;其二,基于用户的基本数据,确定账号对应的特征向量,通过无监督聚类方式,对账号的特征向量进行聚类,将聚类得到的账号确定为相似账号。

[0005] 在实现本发明的过程中,发明人发现现有技术中至少存在以下技术问题:对于第一种判断方式,存在数据缺失导致无法判断的问题,例如在申请账号时,身份证信息不属于必填字段,存在大量账号的身份证字段缺失。进一步的,黑产用户通过实名认证时使用的身份证、手机号、银行卡等大部分由黑市购买得到,无法保证信息的准确性。对于第二种判断方式,通过无监督聚类算法,可以是将用户信息归并到一个特定的群组,但是如果存在较大的群组(包含大量账号)时,对于非数值型的属性,无法量化两个账号之间的形似程度,有效判断的准确性差。

发明内容

[0006] 本发明提供信息归并方法、交易查询方法、装置、计算机及存储介质,以提高信息归并的准确性。

[0007] 第一方面,本发明实施例提供了一种信息归并方法,包括:

[0008] 基于至少两个数据源获取待处理数据,提取所述待处理数据中的特征信息和特征关联信息;

[0009] 根据提取的所述特征信息和所述特征关联信息生成信息关联图;

[0010] 对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至少一个联通子图对所述待处理数据进行信息归并。

[0011] 第二方面,本发明实施例还提供了一种交易查询方法,包括:

[0012] 获取已知风险用户信息,根据所述风险用户信息在至少一个联通子图中进行匹

配,确定与所述已知风险用户信息相匹配的目标联通子图,其中,所述至少一个联通子图为根据本申请任意实施例提供的信息归并方法确定的;

[0013] 提取所述目标联通子图中关联用户信息;

[0014] 确定所述关联用户信息的当前交易,将所述关联用户信息的当前交易确定为风险交易。

[0015] 第三方面,本发明实施例还提供了一种信息归并装置,包括:

[0016] 信息提取模块,用于基于至少两个数据源获取待处理数据,提取所述待处理数据中的特征信息和特征关联信息;

[0017] 信息关联图生成模块,用于根据提取的所述特征信息和所述特征关联信息生成信息关联图;

[0018] 信息归并模块,用于对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至少一个联通子图对所述待处理数据进行信息归并。

[0019] 第四方面,本发明实施例还提供了一种交易查询装置,其特征在于,包括:

[0020] 第一目标联通子图确定模块,用于获取已知风险用户信息,根据所述风险用户信息在至少一个联通子图中进行匹配,确定与所述已知风险用户信息相匹配的目标联通子图,其中,所述至少一个联通子图为根据本申请任意实施例提供的信息归并方法确定的;

[0021] 关联用户信息确定模块,用于提取所述目标联通子图中关联用户信息;

[0022] 风险交易确定模块,用于确定所述关联用户信息的当前交易,将所述关联用户信息的当前交易确定为风险交易。

[0023] 第五方面,本发明实施例还提供了一种计算机设备,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,所述处理器执行所述程序时实现如本申请任意实施例提供的信息归并方法。

[0024] 第六方面,本发明实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如本申请任意实施例提供的信息归并方法。

[0025] 第七方面,本发明实施例还提供了一种计算机设备,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,所述处理器执行所述程序时实现如本申请任意实施例提供的交易查询方法。

[0026] 第八方面,本发明实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如本申请任意实施例提供的交易查询方法。

[0027] 本发明实施例提供的技术方案,通过将待处理数据中的特征信息通过特征信息间的关联关系形成信息关联图,并基于信息关联图中特征节点的联通性对信息关联图进行划分,得到多个彼此独立的联通子图,基于联通子图将特征信息进行归并,通过图形方式简化了信息归并,方便直观,解决了数据库中海量数据无法清晰判断数据关联关系的问题,提高了图像归并效率。

附图说明

[0028] 图1为本发明实施例一提供的一种信息归并方法的方法流程图;

[0029] 图2为本发明实施例一提供的一种信息关联图的示意图;

[0030] 图3为本发明实施例二提供的一种交易查询方法的方法流程图;

- [0031] 图4是本发明实施例三提供的一种信息归并装置的结构示意图；
 [0032] 图5是本发明实施例四提供的一种交易查询装置的结构示意图；
 [0033] 图6为本发明实施例五提供的一种计算机设备的结构示意图；
 [0034] 图7为本发明实施例七提供的一种计算机设备的结构示意图。

具体实施方式

[0035] 下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是,此处所描述的具体实施例仅仅用于解释本发明,而非对本发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与本发明相关的部分而非全部结构。

[0036] 实施例一

[0037] 图1为本发明实施例一提供的一种信息归并方法的流程图,本实施例可适用于对大量数据进行信息归并的情况,该方法可以由本申请实施例提供的信息归并装置来执行,具体包括如下步骤:

[0038] S110、基于至少两个数据源获取待处理数据,提取所述待处理数据中的特征信息和特征关联信息。

[0039] 其中,数据源用于提供不同的待处理数据,待处理数据可以是数据源传输的实时数据,还可以是数据源存储的预设时间段的待处理数据。根据待归并数据确定特征标识,根据特征标识在待处理数据中提取对应的特征信息,以及特征信息之间的关联关系。其中,特征标识可以是特征信息的名称或者用于表征特征信息的字符或者字符串等,该特征标识可以是预先确定的,还可以是根据待归并数据从待处理数据中筛选得到的。示例性的,以电商平台的数据源为例,若待归并的数据为账号,则确定特征标识可以是与账号相关的信息标识,例如账号名称、账号注册用户、账号注册手机号等。

[0040] 可选的,提取所述待处理数据中的节点信息和节点关联信息,包括:根据预设特征标识在所述待处理数据中进行匹配,确定所述预设特征标识对应的特征信息;遍历所述待处理数据,确定任意两个特征信息之间的关联关系。在本实施例中,预先设置特征标识,将预设特征标识在待处理数据中进行逐一匹配,得到每一个预设特征标识对应的特征信息,以及任意两个特征信息之间的关联关系。以电商平台的数据源为例,所述数据源包括电商平台的用户行为数据源,电商平台的用户行为数据源可以包括但不限于注册信息表、会员信息表、绑卡信息表、支付信息表、实名认证信息表、订单信息表和支付信息表,相应的,待归并数据可以是账号,从待处理数据中提取的特征信息可以包括但不限于账号、会员信息、银行卡信息、证件信息、手机号信息、设备信息和WIFI信息。基于上述预设特征标识在上述电商平台的用户行为数据源进行匹配,以得到预设特征标识对应的特征信息和特征信息间的关联关系。可选的,可预先确定匹配各个预设特征标识和关联关系的数据源,提高特征信息提取的针对性,避免无效的数据匹配过程。示例性的,参见表1和表2,表1为特征信息与数据源的对应关系,表2为特征信息间关联关系与数据源的对应关系。

[0041] 表1

[0042]

数据类型	特征名称	数据源
node_type_pin	账号	注册信息表
node_type_pay	会员信息	会员信息表

node_type_card	银行卡信息	绑卡信息表、支付信息表
node_type_idcard	证件信息	实名认证信息表
node_type_phone	手机号信息	注册信息表、订单信息表
node_type_eid	设备信息	注册信息表、支付信息表

[0043] 表2

关系类型	特征信息的关联关系	数据源
register_with_pin_phone	账号与注册手机号	注册信息表
register_on_pin_eid	账号与注册设备	注册信息表

[0044]

bind_to_pin_pay	账号与会员 ID	会员信息表
login_on_pin_eid	账号与登录设备	登录信息表
auth_by_pin_idcard	账号与证件信息	实名认证信息表
reserve_with_card_phone	银行卡与持卡人手机号	用户绑卡信息表
owns_id_idcard_card	银行卡与持卡人证件	用户绑卡信息表
bind_to_pin_card	账号与绑定银行卡	用户绑卡信息表
consignee_with_pin_phone	账号与收货手机号	订单信息表
trade_with_pin_card	账号与支付银行卡	支付信息表

[0045]

[0046] S120、根据提取的所述特征信息和所述特征关联信息生成信息关联图。

[0047] 其中,信息关联图通过图形的形式展示特征信息和所述特征关联信息,示例性的,可以是将具有关联关系的特征信息进行图形连接,以得到全部特征信息之间的关联图形。

[0048] 可选的,根据提取的所述特征信息和所述特征关联信息生成信息关联图,包括:根据所述特征信息设置特征节点;根据所述特征关联信息将存在关联关系的两个特征节点间设置关联边,生成所述信息关联图。其中,信息关联图有特征节点和连接特征节点的关联边组成,每一个特征节点上具有一个特征信息,具有关联关系的两个特征信息对应的两个特征节点之间设置有关联边,关联边用于连接上述两个特征节点。示例性的,参加图2,图2是本发明实施例提供的一种信息关联图的示意图。在图2中,包括账号A、账号B、设备C,其中账号A、账号B均通过设备C进行注册,即账号A、账号B分别与设备C存在关联关系,根据账号A、账号B和设备C分别设置特征节点,并将账号A与设备C的特征节点间设置关联边,将账号B与设备C的特征节点间设置关联边,生成图2中的信息关联图。通过基于特征信息和特征关联信息生成信息关联图,将特征信息通过连接节点的图形形式进行展示,相对于数据表中的文字形式,可直观的确定存在关联关系的特征信息,提高了特征信息的查阅的便利性和直观性。

[0049] S130、对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至

少一个联通子图对所述待处理数据进行信息归并。

[0050] 其中,在确定信息关联图之后,基于信息关联图中特征节点的连通性对信息关联图进行联通子图划分,通过特征信息的关联关系对特征信息进行划分,得到多个联通子图,其中,联通子图中的任意两个特征节点可通过一个或多个关联边联通,且任意两个联通子图中的任意特征节点之间不存在关联边。

[0051] 可选的,对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至少一个联通子图对所述待处理数据进行信息归并,包括:遍历所述信息关联图中的特征节点,将基于关联边连接的特征节点划分至同一联通子图,其中,任意两个联通子图中的任意特征节点之间不存在关联边;将所述同一联通子图中特征节点对应的特征信息归并为同一群组信息。具体的,对于信息关联图中的第一特征节点,根据该第一特征节点的关联边,可确定与第一特征节点存在关联关系的一个或多个第二特征节点,其中,第二特征节点的数量与第一特征节点的关联边的数量相同;进一步的,根据每一个第二特征节点的关联边,可确定与第二特征节点存在关联关系的一个或多个第三特征节点,其中,第三特征节点与第一特征节点不重复,以此类推,可得到第一特征节点所属的联通子图。

[0052] 本实施例中,根据联通子图对特征信息进行归并,例如可以是将联通子图中的同一类型的特征信息归并为同一群组,示例性的,参数图2,账号A、账号B、设备C属于同一联通子图,进一步的可以是将账号A和账号B确定为同一用户的账号,或者同一组织的账号。

[0053] 本实施例的技术方案,通过将待处理数据中的特征信息通过特征信息间的关联关系形成信息关联图,并基于信息关联图中特征节点的联通性对信息关联图进行划分,得到多个彼此独立的联通子图,基于联通子图将特征信息进行归并,通过图形方式简化了信息归并,方便直观,解决了数据库中海量数据无法清晰判断数据关联关系的问题,提高了图像归并效率。

[0054] 在一些实施例中,在对所述信息关联图进行联通子图划分,生成至少一个联通子图之后,还包括:若存在至少一个历史联通子图,则将生成的至少一个联通子图与所述至少一个历史联通子图进行合并,生成至少一个更新联通子图。在本实施例中,可以通过新建的联通子图对历史联通子图进行更新,具体的,可以通过新建的联通子图中的任意特征节点的特征信息与历史联通子图进行匹配,通过新建的联通子图更新匹配成功历史联通子图,其中,更新方式可以是遍历新建的联通子图中第一特征节点具有关联边的第二特征节点,确定历史联通子图中是否具有该第二特征节点,若否,则设置第二特征节点,并设置第一特征节点和第二特征节点的关联边,若是,则确定历史联通子图中第一特征节点和第二特征节点是否设置关联边,如果有,则无需进行更新,如果没有,则设置第一特征节点和第二特征节点的关联边。

[0055] 相应的,根据所述至少一个联通子图对所述待处理数据进行信息归并,包括:根据所述至少一个更新联通子图对所述待处理数据进行信息归并。在本实施例中,通过新建的联通子图不断更新历史联通子图,提高联通子图的特征信息的全面性,以及信息归并的准确性。

[0056] 以电商平台的用户行为数据源为例,对电商平台的用户数据进行归并,可以是:基于商平台的用户行为数据源获取待处理数据,提取所述待处理数据中的特征信息和特征关联信息;其中,数据源包括注册信息表、会员信息表、绑卡信息表、支付信息表、实名认证信

息表、订单信息表和支付信息表,特征信息可以包括账号、会员信息、银行卡信息、证件信息、手机号信息、设备信息和WIFI信息。

[0057] 根据提取的上述特征信息和上述特征关联信息生成电商平台的用户信息关联图,该电商平台的用户信息关联图中包括账号节点、会员信息节点、银行卡信息节点、证件信息节点、手机号信息节点、设备信息节点和WIFI信息节点,并通过上述特征信息的关联关系,将上述节点之间设置关联边。

[0058] 对所述信息关联图进行联通子图划分,生成至少一个联通子图,所述联通子图是具有关联关系的用户信息集合,根据所述至少一个联通子图对所述待处理数据进行信息归并。其中,可以是将同一联通子图中的账号确定为同一用户的账号,或者同一组织的账号。通过联通子图对信息进行归并,可通过任一特征信息快速确定与该特征信息关联其他信息,便于后续通过关联账号进行刷单、黑产、欺诈团伙等犯罪行为的查处,提高信息查询效率,便于对特征信息的统一管理。

[0059] 实施例二

[0060] 图3是本发明实施例提供的一种交易查询方法的方法流程图,本实施例可适用于对风险交易进行查询的情况,该方法可以由本申请实施例提供的交易查询装置来执行,具体包括如下步骤:

[0061] S310、获取已知风险用户信息,根据所述风险用户信息在至少一个联通子图上进行匹配,确定与所述已知风险用户信息相匹配的目标联通子图,其中,所述至少一个联通子图为根据上述实施例提供的信息归并方法确定的。

[0062] S320、提取所述目标联通子图中关联用户信息。

[0063] S330、确定所述关联用户信息的当前交易,将所述关联用户信息的当前交易确定为风险交易。

[0064] 其中,已知风险用户信息可以是包括多个特征信息,已知风险用户信息可以是进行违规操作的用户信息,例如进行欺诈、刷单、洗钱等违规操作的用户信息。例如用户实名认证信息、账号信息、身份证信息等。通过已知风险用户信息中的一个或多个信息在至少一个联通子图上进行匹配,确定上述已知风险用户信息所属的联通子图,将该联通子图中其他的关联用户确定为风险用户,将关联用户正在进行的当前交易确定为风险交易。

[0065] 可选的,根据所述风险用户信息在至少一个联通子图上进行匹配,确定与所述已知风险用户信息相匹配的目标联通子图,包括:根据预设特征标识提取所述已知风险用户信息中的风险特征信息;将所述风险特征信息与至少一个联通子图中的特征节点的特征信息进行匹配,当匹配成功时,将匹配成功的特征节点所属的连通子图确定为目标联通子图。

[0066] 在本实施例中,可以是将建立的联通子图存储至图数据库中,其中,图数据库可以是HBase数据库,基于ElasticSearch作为图数据库的索引工具。ElasticSearch索引表主要保存用户根据节点或边的属性进行查询节点和边的信息,示例性的,以订单支付数据为例:将订单的支付详细信息保存到ElasticSearch表中,当根据订单支付方式为条件进行查询时,则在ElasticSearch表中查询符合这些条件的特征节点与关联边,然后再在图中查询这些特征节点和关联边的关联信息,根据查询结果得到目标联通子图。示例性的,参见表3,表3是本发明实施例提供的支付数据表的示意图。

[0067] 表3

	表名称	rnetwork.trade_base_info	
	基本说明	支付数据表	
	字段名称	字段意义	数据类型
	ORDERID	订单 id	STRING
	PIN	京东账号	STRING
	CARD_NO	银行卡	STRING
	EN_CARD_NO	银行卡脱敏信息	STRING
	CARD_BANK	发卡银行	STRING
[0068]	CARD_TYPE	银行卡类型	INTEGER
	NEW_MEMBER_ID	会员 ID	STRING
	CREATETIME	支付时间	STRING
	PAY_TYPE	支付类型	INTEGER
	PRO_NAME	商品名称	STRING
	PRO_PRICE	商品价格	DOUBLE
	PRO_NUM	商品数量	INTEGER
	ADD_NAME	收货人	STRNG
	ADD_MOBILE	收货手机号	STRING
	ADD_WHERE	收货地址	STRING
[0069]	EID	设备	STRING
	IP	IP 地址	STRING

[0070] 可选的,所述联通子图的每一个特征节点包括至少一个行为属性信息;相应的,当根据所述已知风险用户信息匹配失败时,获取所述已知风险用户的行为特征信息,将所述

行为特征信息与所述联通子图中的行为属性信息进行匹配,根据匹配结果确定目标联通子图。示例性的,账号的行为属性信息可以是注册账号的ip归属地、注册时间、注册来源等,下单信息的行为属性信息包括收货人、收货地址、商品类型等。行为特征信息可以包括但不限于注册账号的ip归属地、交易时间、支付方式、收货ip归属地。通过提取已知风险用户的交易行为特征信息,在各个联通子图中进行匹配,确定目标联通子图,通过多维度数据对目标联通子图进行搜索,提高目标联通子图的确定精度和速度。

[0071] 在本实施例中,在确定风险交易之后,可以是对风险交易进行验证,当验证成功时,拦截该风险交易。其中,对风险交易进行验证可以是人工审核,还可以是根据预设条件进行验证,其中,预设条件可以是交易时间、交易类型等,当风险交易符合预设条件时,可进行拦截,提高交易安全性。

[0072] 实施例三

[0073] 图4是本发明实施例提供的一种信息归并装置的结构示意图,如图4所示,该装置包括:信息提取模块410、信息关联图生成模块420和信息归并模块430。

[0074] 信息提取模块410,用于基于至少两个数据源获取待处理数据,提取所述待处理数据中的特征信息和特征关联信息;

[0075] 信息关联图生成模块420,用于根据提取的所述特征信息和所述特征关联信息生成信息关联图;

[0076] 信息归并模块430,用于对所述信息关联图进行联通子图划分,生成至少一个联通子图,根据所述至少一个联通子图对所述待处理数据进行信息归并。

[0077] 可选的,信息提取模块410用于:

[0078] 根据预设特征标识在所述待处理数据中进行匹配,确定所述预设特征标识对应的特征信息;

[0079] 遍历所述待处理数据,确定任意两个特征信息之间的关联关系。

[0080] 可选的,信息关联图生成模块420用于:

[0081] 根据所述特征信息设置特征节点;

[0082] 根据所述特征关联信息将存在关联关系的两个特征节点间设置关联边,生成所述信息关联图。

[0083] 可选的,信息归并模块430用于:

[0084] 遍历所述信息关联图中的特征节点,将基于关联边连接的特征节点划分至同一联通子图,其中,任意两个联通子图中的任意特征节点之间不存在关联边;

[0085] 将所述同一联通子图中特征节点对应的特征信息归并为同一群组信息。

[0086] 可选的,所述装置还包括:

[0087] 更新联通子图确定模块,用于在对所述信息关联图进行联通子图划分,生成至少一个联通子图之后,若存在至少一个历史联通子图,则将生成的至少一个联通子图与所述至少一个历史联通子图进行合并,生成至少一个更新联通子图;

[0088] 相应的,信息归并模块430用于:根据所述至少一个更新联通子图对所述待处理数据进行信息归并。

[0089] 可选的,所述数据源包括电商平台的用户行为数据源,相应的,所述信息关联图为电商平台的用户信息关联图,所述联通子图为具有关联关系的用户信息集合。

[0090] 上述产品可执行本发明任意实施例所提供的信息归并方法,具备执行信息归并方法相应的功能模块和有益效果。

[0091] 实施例四

[0092] 图5是本发明实施例提供的交易查询装置的结构示意图,该交易查询装置包括第一目标联通子图确定模块510、关联用户信息确定模块520和风险交易确定模块530。

[0093] 第一目标联通子图确定模块510,用于获取已知风险用户信息,根据所述风险用户信息在至少一个联通子图中进行匹配,确定与所述已知风险用户信息相匹配的目标联通子图,其中,所述至少一个联通子图为根据本申请任意实施例提供的信息归并方法确定的;

[0094] 关联用户信息确定模块520,用于提取所述目标联通子图中关联用户信息;

[0095] 风险交易确定模块530,用于确定所述关联用户信息的当前交易,将所述关联用户信息的当前交易确定为风险交易。

[0096] 可选的,第一目标联通子图确定模块510用于:

[0097] 根据预设特征标识提取所述已知风险用户信息中的风险特征信息;

[0098] 将所述风险特征信息与至少一个联通子图中的特征节点的特征信息进行匹配,当匹配成功时,将匹配成功的特征节点所属的联通子图确定为目标联通子图。

[0099] 可选的,所述联通子图的每一个特征节点包括至少一个行为属性信息;

[0100] 相应的,所述装置还包括:

[0101] 第二目标联通子图确定模块,用于当根据所述已知风险用户信息匹配失败时,获取所述已知风险用户的行为特征信息,将所述行为特征信息与所述联通子图中的行为属性信息进行匹配,根据匹配结果确定目标联通子图。

[0102] 上述产品可执行本发明任意实施例所提供的交易查询方法,具备执行交易查询方法相应的功能模块和有益效果。

[0103] 实施例五

[0104] 图6为本发明实施例五提供的一种计算机设备的结构示意图。图6示出了适于用来实现本发明实施方式的计算机设备612的框图。图6显示的计算机设备612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。设备612典型的是承担信息归并功能的计算设备。

[0105] 如图6所示,计算机设备612以通用计算设备的形式表现。计算机设备612的组件可以包括但不限于:一个或者多个处理器616,存储装置628,连接不同系统组件(包括存储装置628和处理器616)的总线618。

[0106] 总线618表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构(Industry Standard Architecture,ISA)总线,微通道体系结构(Micro Channel Architecture,MCA)总线,增强型ISA总线、视频电子标准协会(Video Electronics Standards Association,VESA)局域总线以及外围组件互连(Peripheral Component Interconnect,PCI)总线。

[0107] 计算机设备612典型地包括多种计算机系统可读介质。这些介质可以是任何能够被计算机设备612访问的可用介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0108] 存储装置628可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器(Random Access Memory, RAM) 630和/或高速缓存存储器632。计算机设备612可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统634可以用于读写不可移动的、非易失性磁介质(图6未显示,通常称为“硬盘驱动器”)。尽管图6中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如只读光盘(Compact Disc-Read Only Memory, CD-ROM)、数字视盘(Digital Video Disc-Read Only Memory, DVD-ROM) 或者其它光介质)读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线618相连。存储装置628可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模块,这些程序模块被配置以执行本发明各实施例的功能。

[0109] 具有一组(至少一个)程序模块626的程序636,可以存储在例如存储装置628中,这样的程序模块626包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块626通常执行本发明所描述的实施例中的功能和/或方法。

[0110] 计算机设备612也可以与一个或多个外部设备614(例如键盘、指向设备、摄像头、显示器624等)通信,还可与一个或者多个使得用户能与该计算机设备612交互的设备通信,和/或与使得该计算机设备612能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口622进行。并且,计算机设备612还可以通过网络适配器620与一个或者多个网络(例如局域网(Local Area Network, LAN),广域网Wide Area Network, WAN)和/或公共网络,例如因特网)通信。如图所示,网络适配器620通过总线618与计算机设备612的其它模块通信。应当明白,尽管图中未示出,可以结合计算机设备612使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、磁盘阵列(Redundant Arrays of Independent Disks, RAID)系统、磁带驱动器以及数据备份存储系统等。

[0111] 处理器616通过运行存储在存储装置628中的程序,从而执行各种功能应用以及数据处理,例如实现本发明上述实施例所提供的信息归并方法。

[0112] 实施例六

[0113] 本发明实施例六还提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如本发明实施例所提供的信息归并方法。

[0114] 当然,本发明实施例所提供的一种计算机可读存储介质,其上存储的计算机程序不限于如上所述的方法操作,还可以执行本发明任意实施例所提供的信息归并方法。

[0115] 本发明实施例的计算机存储介质,可以采用一个或多个计算机可读的介质的任意组合。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。计算机可读存储介质例如可以是一—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(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPR0M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。在本文件中,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。

[0116] 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。

[0117] 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。

[0118] 可以以一种或多种程序设计语言或其组合来编写用于执行本发明操作的计算机程序代码,所述程序设计语言包括面向对象的程序设计语言——诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言——诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0119] 实施例七

[0120] 图7为本发明实施例七提供的一种计算机设备的结构示意图。图7示出了适于用来实现本发明实施方式的计算机设备712的框图。图7显示的计算机设备712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。设备712典型的是承担交易查询功能的计算设备。

[0121] 如图7所示,计算机设备712以通用计算设备的形式表现。计算机设备712的组件可以包括但不限于:一个或者多个处理器716,存储装置728,连接不同系统组件(包括存储装置728和处理器716)的总线718。

[0122] 总线718表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构(Industry Standard Architecture,ISA)总线,微通道体系结构(Micro Channel Architecture,MCA)总线,增强型ISA总线、视频电子标准协会(Video Electronics Standards Association,VESA)局域总线以及外围组件互连(Peripheral Component Interconnect,PCI)总线。

[0123] 计算机设备712典型地包括多种计算机系统可读介质。这些介质可以是任何能够被计算机设备712访问的可用介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0124] 存储装置728可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器(Random Access Memory,RAM)730和/或高速缓存存储器732。计算机设备712可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统734可以用于读写不可移动的、非易失性磁介质(图7未显示,通常称为“硬盘驱动器”)。尽管图7中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如只读光盘(Compact Disc-Read Only Memory,CD-

ROM)、数字视盘(Digital Video Disc-Read Only Memory,DVD-ROM) 或者其它光介质) 读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线718 相连。存储装置728可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模块,这些程序模块被配置以执行本发明各实施例的功能。

[0125] 具有一组(至少一个)程序模块726的程序736,可以存储在例如存储装置728中,这样的程序模块726包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块726通常执行本发明所描述的实施例中的功能和/或方法。

[0126] 计算机设备712也可以与一个或多个外部设备714(例如键盘、指向设备、摄像头、显示器724等)通信,还可与一个或者多个使得用户能与该计算机设备712交互的设备通信,和/或与使得该计算机设备712能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口722进行。并且,计算机设备712还可以通过网络适配器720与一个或者多个网络(例如局域网(Local Area Network,LAN),广域网Wide Area Network,WAN)和/或公共网络,例如因特网)通信。如图所示,网络适配器720通过总线718与计算机设备712的其它模块通信。应当明白,尽管图中未示出,可以结合计算机设备712使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、磁盘阵列(Redundant Arrays of Independent Disks,RAID)系统、磁带驱动器以及数据备份存储系统等。

[0127] 处理器716通过运行存储在存储装置728中的程序,从而执行各种功能应用以及数据处理,例如实现本发明上述实施例所提供的交易查询方法。

[0128] 实施例八

[0129] 本发明实施例八还提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如本发明实施例所提供的交易查询方法。

[0130] 当然,本发明实施例所提供的一种计算机可读存储介质,其上存储的计算机程序不限于如上所述的方法操作,还可以执行本发明任意实施例所提供的交易查询方法。

[0131] 本发明实施例的计算机存储介质,可以采用一个或多个计算机可读的介质的任意组合。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。计算机可读存储介质例如可以是一—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(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(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。在本文件中,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。

[0132] 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。

[0133] 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。

[0134] 可以以一种或多种程序设计语言或其组合来编写用于执行本发明操作的计算机程序代码,所述程序设计语言包括面向对象的设计语言——诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言——诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0135] 注意,上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解,本发明不限于这里所述的特定实施例,对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本发明的保护范围。因此,虽然通过以上实施例对本发明进行了较为详细的说明,但是本发明不仅仅限于以上实施例,在不脱离本发明构思的情况下,还可以包括更多其他等效实施例,而本发明的范围由所附的权利要求范围决定。

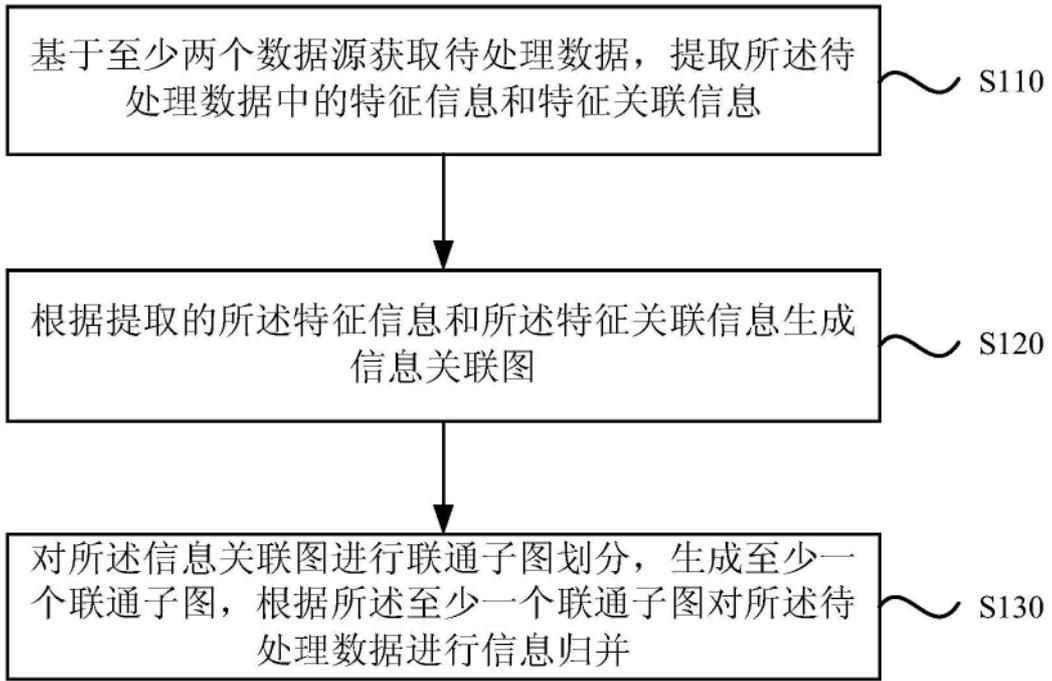


图1

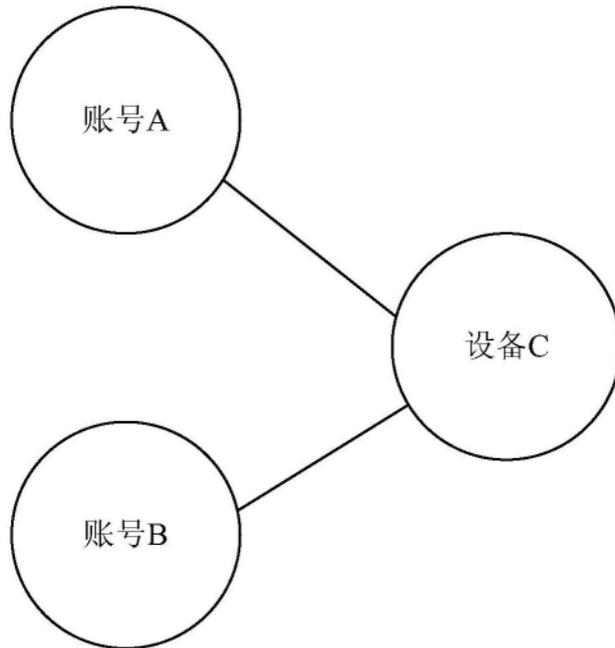


图2

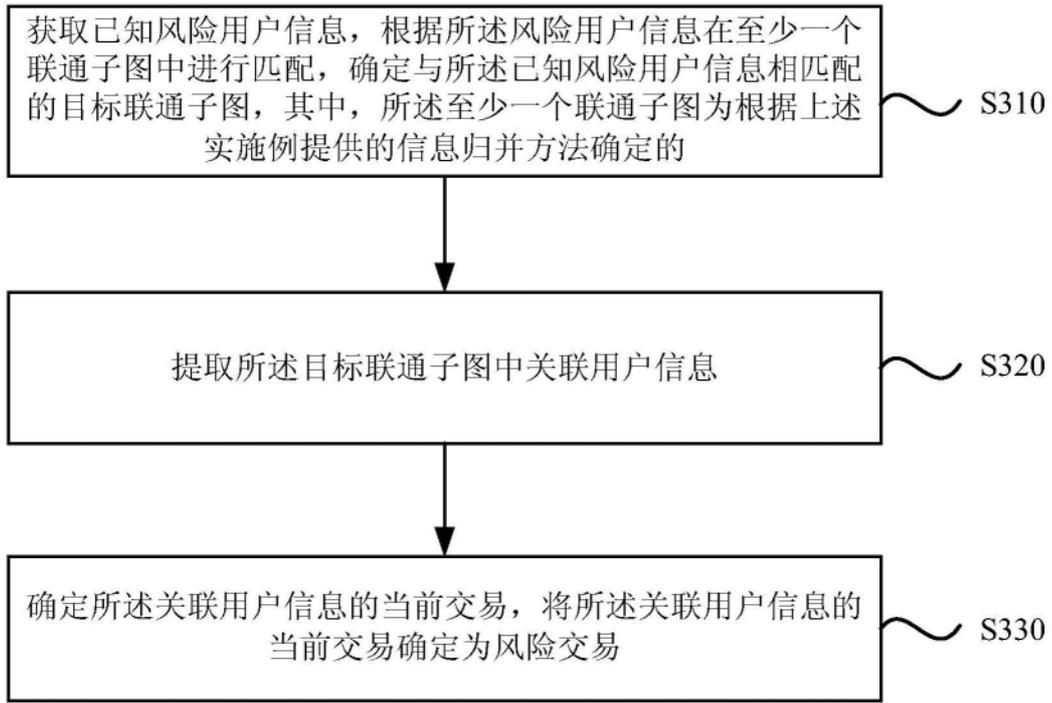


图3

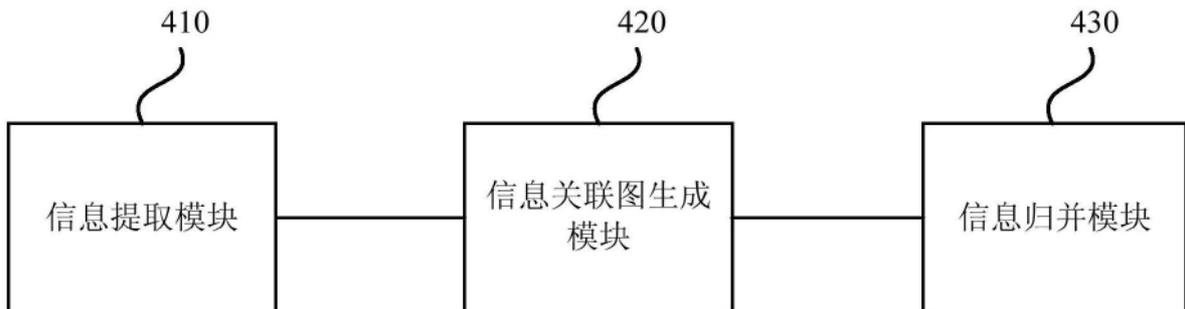


图4

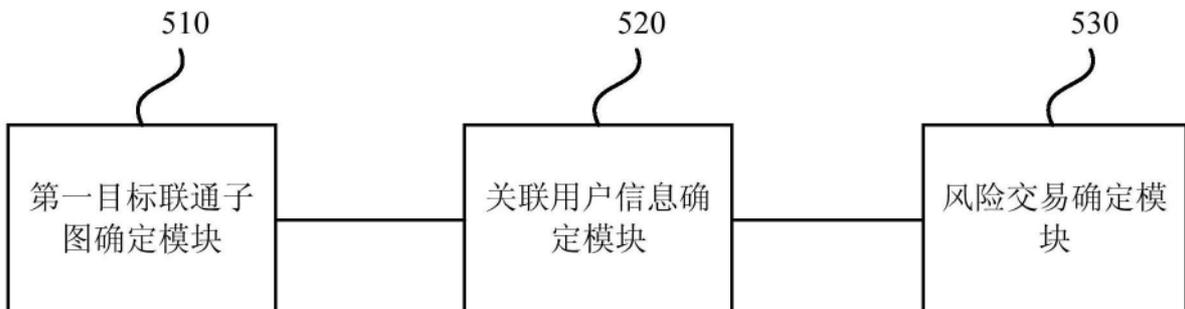


图5

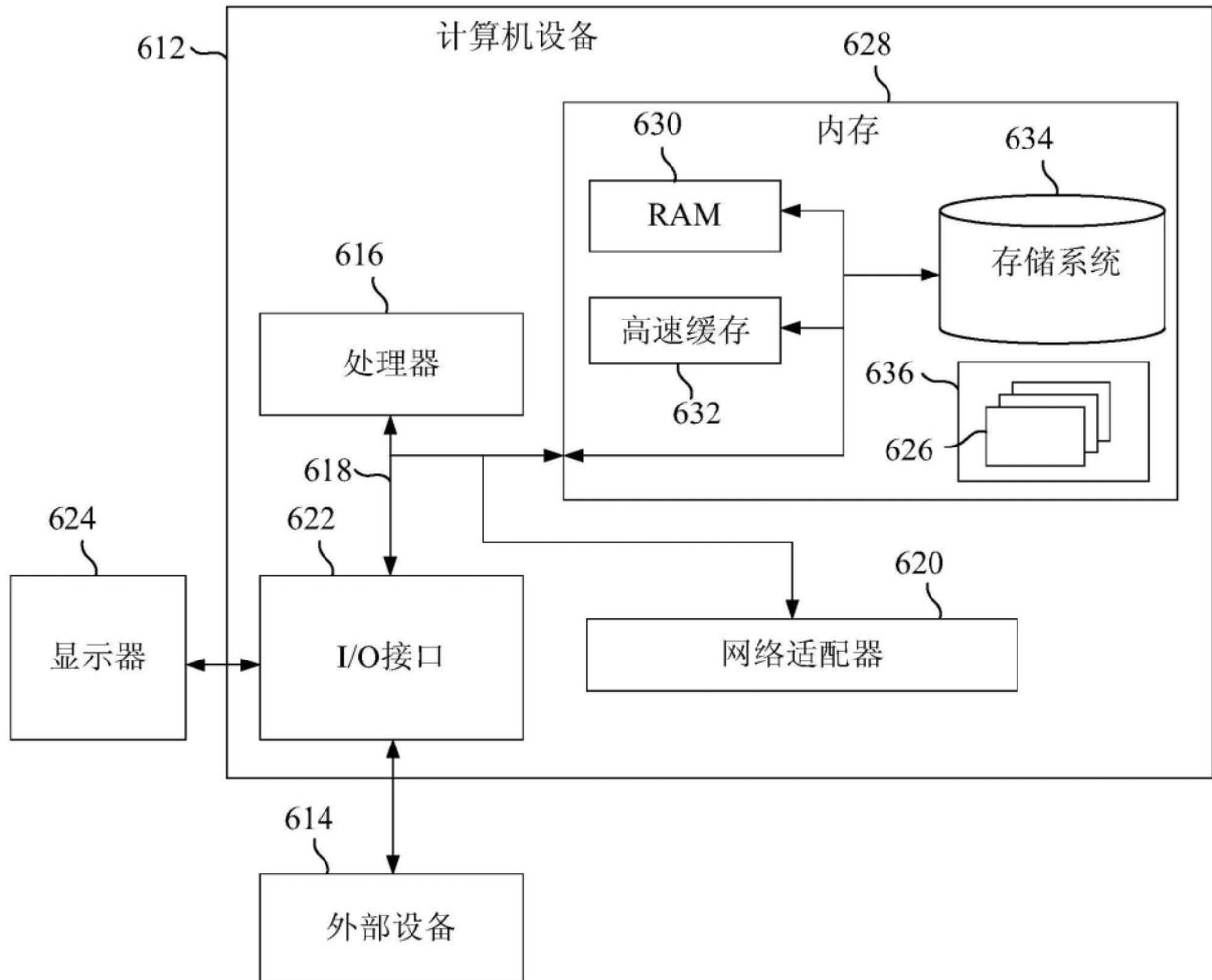


图6

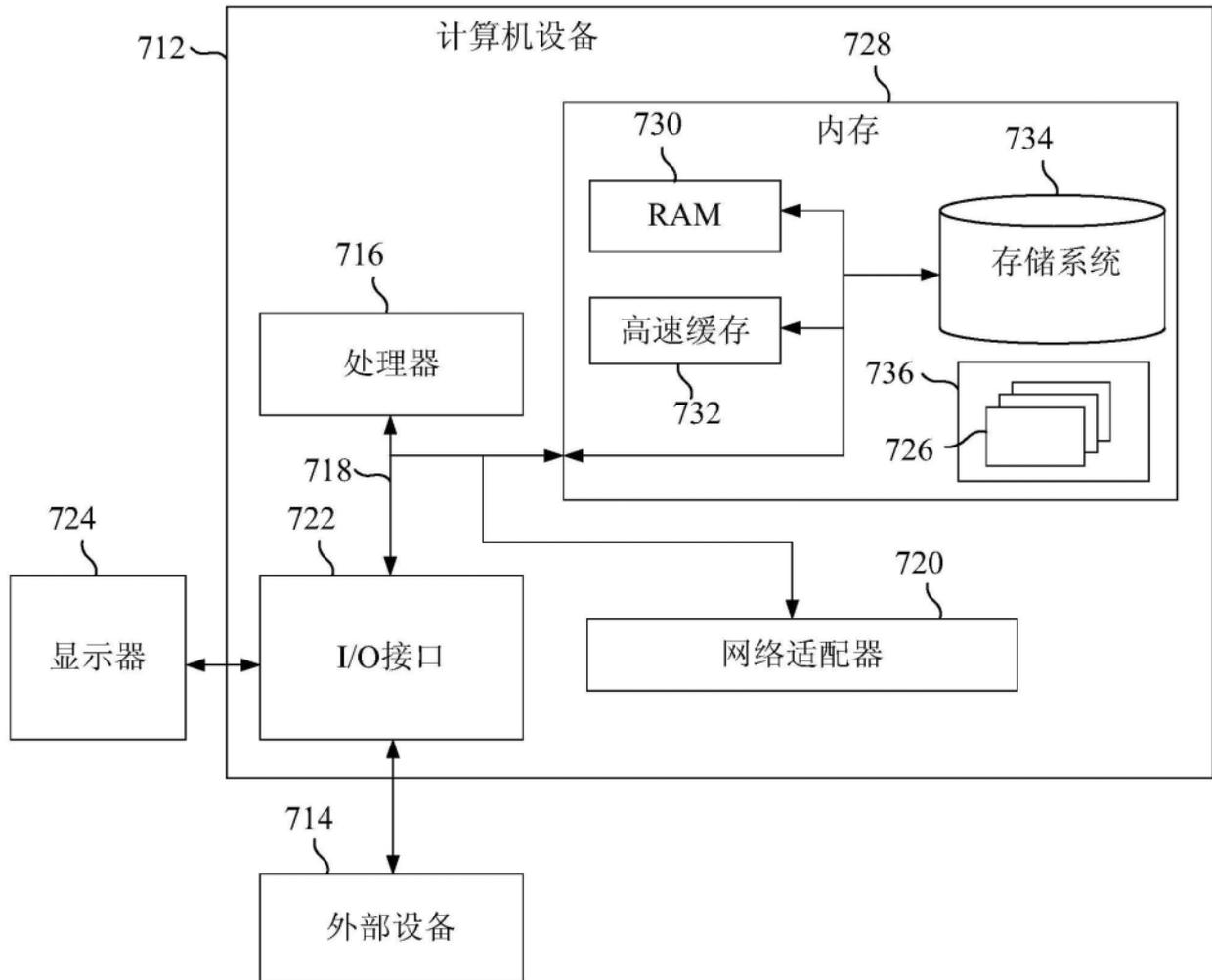


图7