



(12)发明专利申请

(10)申请公布号 CN 110647522 A

(43)申请公布日 2020.01.03

(21)申请号 201910842703.1

(22)申请日 2019.09.06

(71)申请人 中国建设银行股份有限公司  
地址 100033 北京市西城区金融大街25号  
申请人 建信金融科技有限责任公司

(72)发明人 梁学甲 覃鹏 龚苇 梁永健  
 禰栋雄

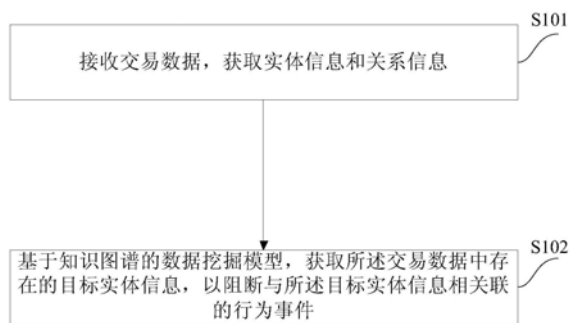
(74)专利代理机构 中原信达知识产权代理有限  
 责任公司 11219  
代理人 张一军 李阳

(51)Int.Cl.  
G06F 16/215(2019.01)  
G06F 16/2458(2019.01)  
G06F 16/26(2019.01)

权利要求书1页 说明书8页 附图4页

(54)发明名称  
一种数据挖掘方法、装置及其系统

(57)摘要  
本发明公开了数据挖掘方法、装置及其系统,涉及计算机技术领域。该方法的一具体实施方式包括:接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息;基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。从而,本发明的实施方式能够解决挖掘黑产数据效率低、精度差的问题。



1. 一种数据挖掘方法,其特征在于,包括:

接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息;

基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

2. 根据权利要求1所述的方法,其特征在于,获取所述交易数据中存在的目标实体信息之后,包括:

根据知识图谱的数据挖掘模型的关系网络距离,对获取的目标实体信息分级添加至预设的目标实体名单中。

3. 根据权利要求1所述的方法,其特征在于,还包括:

根据预设间隔时间段,获取包括有所述目标实体信息的交易数据;

利用知识图谱的数据挖掘模型,计算间隔一段时间所获得的知识图谱的相似度,以获得所述目标实体信息的行为变化数据。

4. 根据权利要求1-3任一所述的方法,其特征在于,还包括:

以图形方式展示目标实体信息的资金流信息和行为流信息,以列表形式展示基于知识图谱的数据挖掘模型获得的知识图谱中的所有节点及其关系。

5. 一种数据挖掘装置,其特征在于,包括:

数据处理模块,用于接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息;

知识图谱侦测引擎,用于基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

6. 根据权利要求5所述的装置,其特征在于,所述知识图谱侦测引擎获取所述交易数据中存在的目标实体信息之后,包括:

根据知识图谱的数据挖掘模型的关系网络距离,对获取的目标实体信息分级添加至预设的目标实体名单中。

7. 根据权利要求5所述的装置,其特征在于,所述知识图谱侦测引擎,还用于:

根据预设间隔时间段,获取包括有所述目标实体信息的交易数据;

利用知识图谱的数据挖掘模型,计算间隔一段时间所获得的知识图谱的相似度,以获得所述目标实体信息的行为变化数据。

8. 根据权利要求5-7任一所述的装置,其特征在于,还包括:

可视化管理模块,用于以图形方式展示目标实体信息的资金流信息和行为流信息,以列表形式展示基于知识图谱的数据挖掘模型获得的知识图谱中的所有节点及其关系。

9. 一种电子设备,其特征在于,包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-4中任一所述的方法。

10. 一种计算机可读介质,其上存储有计算机程序,其特征在于,所述程序被处理器执行时实现如权利要求1-4中任一所述的方法。

## 一种数据挖掘方法、装置及其系统

### 技术领域

[0001] 本发明涉及计算机技术领域,尤其涉及一种数据挖掘方法、装置及其系统。

### 背景技术

[0002] 金融黑产所涉的交易行为一般具有不公平的属性,通过欺诈套取他人资金的交易动机以及为逃避司法打击而隐匿资金的交易行为,导致涉及金融黑产账户呈现出异于一般客户和账户的种种异常特征,包括交易行为异常、交易对象异常、资金走向异常等。

[0003] 在实现本发明过程中,发明人发现现有技术中至少存在如下问题:

[0004] 目前,在黑产数据(黑产数据是指欺诈团伙所使用的账号、设备、手机号、位置等数据,同一个团伙一般存在一定的关联共性。)挖掘应用方面,主要使用关系数据库、简单规则来实现。比如在Greenplum(Greenplum是一家为用户提供新型企业级数据仓库公司)上使用MADlib(MADlib为一个开源机器学习库)算法库中的简单图算法进行关联关系规则统计。可以看出,现有的黑产数据挖掘方式,主要针对资金交易行为总结专家规则,只能进行简单的部分转入转出,挖掘黑产的深度较低,且效率低,精度差。

### 发明内容

[0005] 有鉴于此,本发明实施例提供一种数据挖掘方法、装置及其系统,能够解决挖掘黑产数据效率低、精度差的问题。

[0006] 为实现上述目的,根据本发明实施例的一个方面,提供了一种数据挖掘方法,包括接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息;基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

[0007] 可选地,获取所述交易数据中存在的目标实体信息之后,包括:

[0008] 根据知识图谱的数据挖掘模型的关系网络距离,对获取的目标实体信息分级添加至预设的目标实体名单中。

[0009] 可选地,还包括:

[0010] 根据预设间隔时间段,获取包括有所述目标实体信息的交易数据;

[0011] 利用知识图谱的数据挖掘模型,计算间隔一段时间所获得的知识图谱的相似度,以获得所述目标实体信息的行为变化数据。

[0012] 可选地,还包括:

[0013] 以图形方式展示目标实体信息的资金流信息和行为流信息,以列表形式展示基于知识图谱的数据挖掘模型获得的知识图谱中的所有节点及其关系。

[0014] 另外,根据本发明实施例的一个方面,提供了一种数据挖掘装置,包括数据处理模块,用于接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息;知识图谱侦测引擎,用于基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相

关联的行为事件。

[0015] 可选地,所述处理模块获取所述交易数据中存在的目标实体信息之后,包括:

[0016] 根据知识图谱的数据挖掘模型的关系网络距离,对获取的目标实体信息分级添加至预设的目标实体名单中。

[0017] 可选地,还用于:

[0018] 根据预设间隔时间段,获取包括有所述目标实体信息的交易数据;

[0019] 利用知识图谱的数据挖掘模型,计算间隔一段时间所获得的知识图谱的相似度,以获得所述目标实体信息的行为变化数据。

[0020] 可选地,还包括:

[0021] 可视化管理模块,用于以图形方式展示目标实体信息的资金流信息和行为流信息,以列表形式展示基于知识图谱的数据挖掘模型获得的知识图谱中的所有节点及其关系。

[0022] 根据本发明实施例的另一个方面,还提供了一种电子设备,包括:

[0023] 一个或多个处理器;

[0024] 存储装置,用于存储一个或多个程序,

[0025] 当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现上述任一数据挖掘实施例所述的方法。

[0026] 根据本发明实施例的另一个方面,还提供了一种计算机可读介质,其上存储有计算机程序,所述程序被处理器执行时实现上述任一基于数据挖掘实施例所述的方法。

[0027] 上述发明中的一个实施例具有如下优点或有益效果:本发明通过接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息;基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。因此,本发明能够实现一套基于知识图谱的黑产数据挖掘方法、装置及其系统,简化黑产数据挖掘操作流程,进而实现高效率、高精度的黑产数据挖掘。

[0028] 上述的非惯用的可选方式所具有的进一步效果将在下文中结合具体实施方式加以说明。

## 附图说明

[0029] 附图用于更好地理解本发明,不构成对本发明的不当限定。其中:

[0030] 图1是根据本发明第一实施例的数据挖掘方法的主要流程的示意图;

[0031] 图2是根据本发明第二实施例的数据挖掘方法的主要流程的示意图;

[0032] 图3是根据本发明第三实施例的数据挖掘方法的主要流程的示意图;

[0033] 图4是根据本发明实施例的数据挖掘装置的主要模块的示意图;

[0034] 图5是根据本发明另一实施例的数据挖掘装置的主要模块的示意图;

[0035] 图6是根据本发明实施例的数据挖掘系统的示意图;

[0036] 图7是本发明实施例可以应用于其中的示例性系统架构图;

[0037] 图8是适于用来实现本发明实施例的终端设备或服务器的计算机系统的结构示意图。

## 具体实施方式

[0038] 以下结合附图对本发明的示范性实施例做出说明,其中包括本发明实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本发明的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0039] 图1是根据本发明第一实施例的数据挖掘方法的主要流程的示意图,所述数据挖掘方法可以包括:

[0040] 步骤S101,接收交易数据,获取实体信息和关系信息。

[0041] 在实施例中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息。基于知识图谱的数据挖掘模型主要考虑实体信息的选取及关系信息的处理。例如:以用户的证件号、账号、设备号、手机号、住址等各种基本信息作为实体,对实体之间包括签约、登录、转账等的不同行为和各类属性特征作为关系,构建“实体—关系—实体”的数据集合,送入图形数据库,将不同用户及其属性关系可视化。

[0042] 步骤S102,基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

[0043] 较佳地,在获取所述交易数据中存在的目标实体信息之后,可以根据知识图谱的数据挖掘模型的关系网络距离,对获取的目标实体信息分级添加至预设的目标实体名单中。

[0044] 在进一步地实施例中,按基于知识图谱的数据挖掘模型获得的知识图谱中关系网络距离筛选上下游交易实体,再对上游实体结合其资金往来情况及登录情况导出的可视化图谱自动分析并将其归类,例如黑产和高危客户。同时,对数据库中相同特征信息按图谱中关系网络距离分级添加黑名单,并导出可视化图谱,展示清晰的关联路径,变冗杂的数据表为直观的图形展示。

[0045] 其中,分级添加黑名单为了后续进行交易控制,根据级别不同,对有些目标实体直接拦截交易,有些则可以进行事后进一步分析等。

[0046] 在一个具体的实施例中,在对获取的目标实体信息分级添加至预设的目标实体名单中时,将基于知识图谱的数据挖掘模型初始构造的大图,通过社区发现算法划分为各个子图,针对每个子图,会重新计算其基本特征(出入度、中心性、中介性、PageRank等),然后筛选出包括目标实体信息的子图,并计算出每个子图中目标实体信息的数量,根据基本特征的排名以及目标实体信息节点在图中的位置,分别标记出图中每个节点的风险等级,将排名比目标实体信息节点高以及和目标实体信息节点直接相连的节点标记为高风险,其余的节点标记为低风险。

[0047] 作为一个可参考的实施例,步骤S102中的基于知识图谱的数据挖掘模型可以通过如下过程获得:

[0048] 首先构建基于知识图谱的数据挖掘模型所需的实体,包括用户实体、设备实体等,并根据资金流、行为流等数据信息抽取多维度关联关系,存入图数据库中以构建多维度的知识图谱,同时计算图中各实体的出入度(Degree)、中心性(Closeness)、中介性(Betweenness)、PageRank等基本属性特征。然后,从已知的黑产数据出发,可直接查询出与其关联的可疑账户图谱,同时以标记的黑产数据为样本,使用相关社区发现算法训练知识

图谱黑产挖掘模型,并不断调优。

[0049] 其中,出入度(Degree)是图中与节点相关联的边的数量,对于有向图来说,节点的出边条数称为该节点的出度,节点的入边条数称为该节点的入度。中心性(Closeness)是节点与图中所有其他可达节点之间的最短路径的均值,这个值越小就说明这个节点到其他所有节点的路径越短,也就说明这个节点距离其他所有节点越近。中介性(Betweenness)是计算经过一个节点的最短路径的数量,数量越多,就说明它的中介性越高。PageRank是一种链接分析算法,它将数字权重(介于0和1之间)分配给有向图中的每个节点,从而“测量”它们在图中的相对重要性。

[0050] 进一步地,在构建基于知识图谱的数据挖掘模型的实体和关联关系时,可以对历史黑产数据进行维护、标记、分类。其中,标记是为了支持不同的数据挖掘模型,可以建立多个细分模型,通过标记来控制不同的数据,跑不同的模型,同时也会根据运行结果调整参数等。分类是将黑产数据划分为实体信息和关系信息,实体信息包括用户信息(例如账号)和设备信息,关系信息包括资金流信息和行为流信息。

[0051] 另外,设置不同的数据源及需要分析挖掘的数据维度。其中,数据维度包含:客户(姓名、住址、联系人)、账号(卡bin、归属地、开户网点)、设备(设备号、设备类型、设备标识)、IP(IP段、归属地、代理等标识)、手机号(号段、归属地、运营商、高频等标识)、位置(GPS、区域、位置标识)、交易(金额、频率、时间段)、登录(时间段、频率)。

[0052] 作为具体的实施例,节点(实体)和边(关系)构成一个图,实体分为用户实体、设备实体,每个实体有一些属性值,用户实体有姓名、账号、客户号等属性,设备实体有设备类型、设备号等属性。关系分为资金关系、登录关系,资金关系主要是资金的流入流出关系,对应的节点为客户实体,包含流入节点、流出节点、总金额、总笔数、时间范围等,登录关系是客户实体与设备实体之间的关系,是单向的关系,包含登录次数、时间段等。

[0053] 图中实体以及关系的构造,通过分析用户交易痕迹、登录痕迹构造,抽取一段时间的交易、登录数据,进行统计分析,分别按照实体及关系的结构构造出实体及关系,从而绘制出一个初始的大图。

[0054] 图2是根据本发明第二实施例的数据挖掘方法的主要流程的示意图,所述数据挖掘方法可以包括:

[0055] 步骤S201,接收交易数据,获取实体信息和关系信息。

[0056] 其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息。

[0057] 步骤S202,基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

[0058] 步骤S203,根据预设间隔时间段,获取包括有所述目标实体信息的交易数据。

[0059] 步骤S204,利用知识图谱的数据挖掘模型,计算间隔一段时间所获得的知识图谱的相似度。

[0060] 步骤S205,获得所述目标实体信息的行为变化数据。

[0061] 在图2所示的的实施例中,利用知识图谱的相似性计算对一段时间内图像的变化进行统计,挖掘异常值并纳入后续监测名单。也就是说,针对同一批黑产数据,每间隔预设的一段时间,会将新的交易数据添加进来重复计算其知识图谱,然后根据相似性算法比对

其不同时间的行为变化数据,特别是新增了某些关联节点等。若存在较大变化时,会调整监控范围等。

[0062] 图3是根据本发明第三实施例的数据挖掘方法的主要流程的示意图,所述数据挖掘方法可以包括:

[0063] 步骤S301,接收交易数据,获取实体信息和关系信息。

[0064] 其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息。

[0065] 步骤S302,基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

[0066] 步骤S303,以图形方式展示目标实体信息的资金流信息和行为流信息,以列表形式展示基于知识图谱的数据挖掘模型获得的知识图谱中的所有节点及其关系。

[0067] 另外,还值得说明的是,在图3中所示的实施例可以支持用户制定数据挖掘任务,用户通过拖拽方式展开制定数据挖掘任务,通过自定义方式叠加定义选取符合特定条件或者特定挖掘算法的用户信息,根据需要动态的调整分析数据的广度和分析挖掘的维度,对挖掘结果进行精确的控制,进而获得最优的分析结果。

[0068] 也就是说,提供了可视化页面并选择模型运行,模型分为两种:一个是图的展示形式,一个是列表的展示形式。用户可以筛选出对应的节点,选择条件包含节点、关系、出入度、中心性、中介性、PageRank、日期、节点数等,同时筛选需要运行的模型,可以进行后续的模式分析。

[0069] 综上所述,本发明的数据挖掘方法创造性地基于知识图谱进行黑产挖掘,即通过多维度的分析发现客户、账户之间隐藏的关联关系,构建包含多个数据源的知识图谱,提高挖掘黑产数据的效率。同时,本发明的数据挖掘方法动态调整分析数据的广度和分析挖掘的维度,实现黑产数据可视化的深入挖掘。

[0070] 图4是根据本发明实施例的数据挖掘装置,如图4所示,所述数据挖掘装置400包括数据处理模块401和知识图谱侦测引擎402。其中,数据处理模块401用于接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息。知识图谱侦测引擎402用于基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

[0071] 较佳地,所述知识图谱侦测引擎402获取所述交易数据中存在的目标实体信息之后,包括:根据知识图谱的数据挖掘模型的关系网络距离,对获取的目标实体信息分级添加至预设的目标实体名单中。

[0072] 作为另一个实施例,所述知识图谱侦测引擎402,还用于:

[0073] 根据预设间隔时间段,获取包括有所述目标实体信息的交易数据;

[0074] 利用知识图谱的数据挖掘模型,计算间隔一段时间所获得的知识图谱的相似度,以获得所述目标实体信息的行为变化数据。

[0075] 作为又一个实施例,如图5所示,所述数据挖掘装置500包括数据处理模块501、知识图谱侦测引擎502和可视化管理模块503。其中,数据处理模块501用于接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息。知识图谱侦测引擎502用于基于知识图谱的数据挖掘模型,获取所述

交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。所述可视化管理模块503用于以图形方式展示目标实体信息的资金流信息和行为流信息,以列表形式展示基于知识图谱的数据挖掘模型获得的知识图谱中的所有节点及其关系。

[0076] 还值得说明的是,所述数据挖掘装置还可以包括黑产数据管理模块和知识图谱黑产挖掘模型训练模块。其中,黑产数据管理模块可以提供对黑产数据进行维护、标记、分类等的功能。进一步地,黑产数据管理模块可以将黑产数据划分为不同类型,并设置不同的数据源及需要分析挖掘的数据维度,后续传入知识图谱黑产挖掘模型训练模块。

[0077] 知识图谱黑产挖掘模型训练模块首先构建知识图谱所需的实体,包括客户实体、设备实体等,并根据资金流、行为流等数据抽取出多维度的关联关系,存入图数据库中构建多维度知识图谱的数据挖掘模型,同时计算图中各实体的出入度、中心性、中介性、PageRank等基本属性特征。然后从已知的黑产数据出发,可直接查询出与其关联的可疑账户图谱。同时以标记的黑产数据为样本,使用社区发现算法训练知识图谱的数据挖掘模型,并不断调优。最后,将训练的知识图谱数据挖掘模型发布到知识图谱侦测引擎进行后续挖掘。另外,知识图谱黑产挖掘模型训练模块支持多种维度的不同算法模型训练。

[0078] 需要说明的是,在本发明所述数据挖掘方法和所述数据挖掘装置在具体实施内容上具有相应关系,故重复内容不再说明。

[0079] 图6是根据本发明实施例的数据挖掘系统的示意图,所述数据挖掘系统可以包括:数据准备层、侦测处理层和应用服务层。其中,数据处理模块设置在数据准备层,知识图谱侦测引擎和知识图谱黑产挖掘模型训练设置在侦测处理层,黑产数据管理模块、可视化管理模块和交易控制模块设置在应用服务层。

[0080] 进一步地,黑产数据管理模块将标记、分类后的黑产数据传输至知识图谱黑产挖掘模型训练模块,构建基于知识图谱的数据挖掘模型(图6中所示的1.a)。知识图谱黑产挖掘模型训练模块将训练完成的知识图谱的数据挖掘模型固化到知识图谱侦测引擎(图6中所示的2)。数据处理模块对用户资金流、行为流数据进行清洗,输出给知识图谱侦测引擎(图6中所示的1.b)。知识图谱侦测引擎将侦测挖掘出的黑产数据及其关联关系图传输到可视化管理模块进行展示(图6中所示的3)。可视化管理模块将核实后确认欺诈的黑产数据同步至黑产数据管理模块,用于黑产挖掘模型的调优及后续管理(图6中所示的4.a)。可视化管理模块将黑产数据同步至交易控制模块,阻断黑产客户的后续交易等行为(图6中所示的4.b)。

[0081] 图7示出了可以应用本发明实施例的数据挖掘方法或数据挖掘装置的示例性系统架构700。

[0082] 如图7所示,系统架构700可以包括终端设备701、702、703,网络704和服务器705。网络704用以在终端设备701、702、703和服务器705之间提供通信链路的介质。网络704可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0083] 用户可以使用终端设备701、702、703通过网络704与服务器705交互,以接收或发送消息等。终端设备701、702、703上可以安装有各种通讯客户端应用,例如购物类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等(仅为示例)。

[0084] 终端设备701、702、703可以是具有显示屏并且支持网页浏览的各种电子设备,包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。



[0085] 服务器705可以是提供各种服务的服务器,例如对用户利用终端设备701、702、703所浏览的购物类网站提供支持的后台管理服务器(仅为示例)。后台管理服务器可以对接收到的产品信息查询请求等数据进行分析等处理,并将处理结果(例如目标推送信息、产品信息—仅为示例)反馈给终端设备。

[0086] 需要说明的是,本发明实施例所提供的数据挖掘方法一般由服务器705执行,相应地,数据挖掘装置一般设置于服务器705中。

[0087] 应该理解,图7中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的。

[0088] 下面参考图8,其示出了适于用来实现本发明实施例的终端设备的计算机系统800的结构示意图。图8示出的终端设备仅仅是一个示例,不应对本发明实施例的功能和使用范围带来任何限制。

[0089] 如图8所示,计算机系统800包括中央处理单元(CPU)801,其可以根据存储在只读存储器(ROM)802中的程序或者从存储部分808加载到随机访问存储器(RAM)803中的程序而执行各种适当的动作和处理。在RAM803中,还存储有系统800操作所需的各种程序和数据。CPU801、ROM802以及RAM803通过总线804彼此相连。输入/输出(I/O)接口805也连接至总线804。

[0090] 以下部件连接至I/O接口805:包括键盘、鼠标等的输入部分806;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分807;包括硬盘等的存储部分808;以及包括诸如LAN卡、调制解调器等网络接口卡的通信部分809。通信部分809经由诸如因特网的网络执行通信处理。驱动器810也根据需要连接至I/O接口805。可拆卸介质811,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器810上,以便于从其上读出的计算机程序根据需要被安装入存储部分808。

[0091] 特别地,根据本发明公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本发明公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分809从网络上被下载和安装,和/或从可拆卸介质811被安装。在该计算机程序被中央处理单元(CPU)801执行时,执行本发明的系统中限定的上述功能。

[0092] 需要说明的是,本发明所示的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或者是上述两者的任意组合。计算机可读存储介质例如可以是——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可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(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。在本发明中,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。而在本发明中,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。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。

[0093] 附图中的流程图和框图,图示了按照本发明各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图或流程图中的每个方框、以及框图或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0094] 描述于本发明实施例中所涉及到的模块可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的模块也可以设置在处理器中,例如,可以描述为:一种处理器包括数据处理模块和知识图谱侦测引擎。其中,这些模块的名称在某种情况下并不构成对该模块本身的限定。

[0095] 作为另一方面,本发明还提供了一种计算机可读介质,该计算机可读介质可以是上述实施例中描述的设备中所包含的;也可以是单独存在,而未装配入该设备中。上述计算机可读介质承载有一个或者多个程序,当上述一个或者多个程序被一个该设备执行时,使得该设备包括:接收交易数据,获取实体信息和关系信息;其中,所述实体信息包括用户信息和设备信息,关系信息包括资金流信息和行为流信息;基于知识图谱的数据挖掘模型,获取所述交易数据中存在的目标实体信息,以阻断与所述目标实体信息相关联的行为事件。

[0096] 根据本发明实施例的技术方案,能够解决挖掘黑产数据效率低、精度差的问题。

[0097] 上述具体实施方式,并不构成对本发明保护范围的限制。本领域技术人员应该明白的是,取决于设计要求和因素,可以发生各种各样的修改、组合、子组合和替代。任何在本发明的精神和原则之内所作的修改、等同替换和改进等,均应包含在本发明保护范围之内。

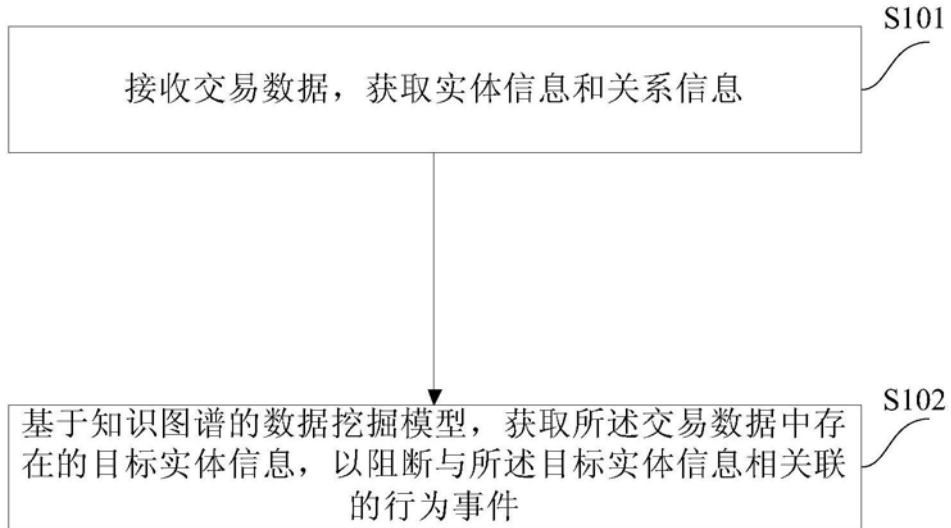


图1

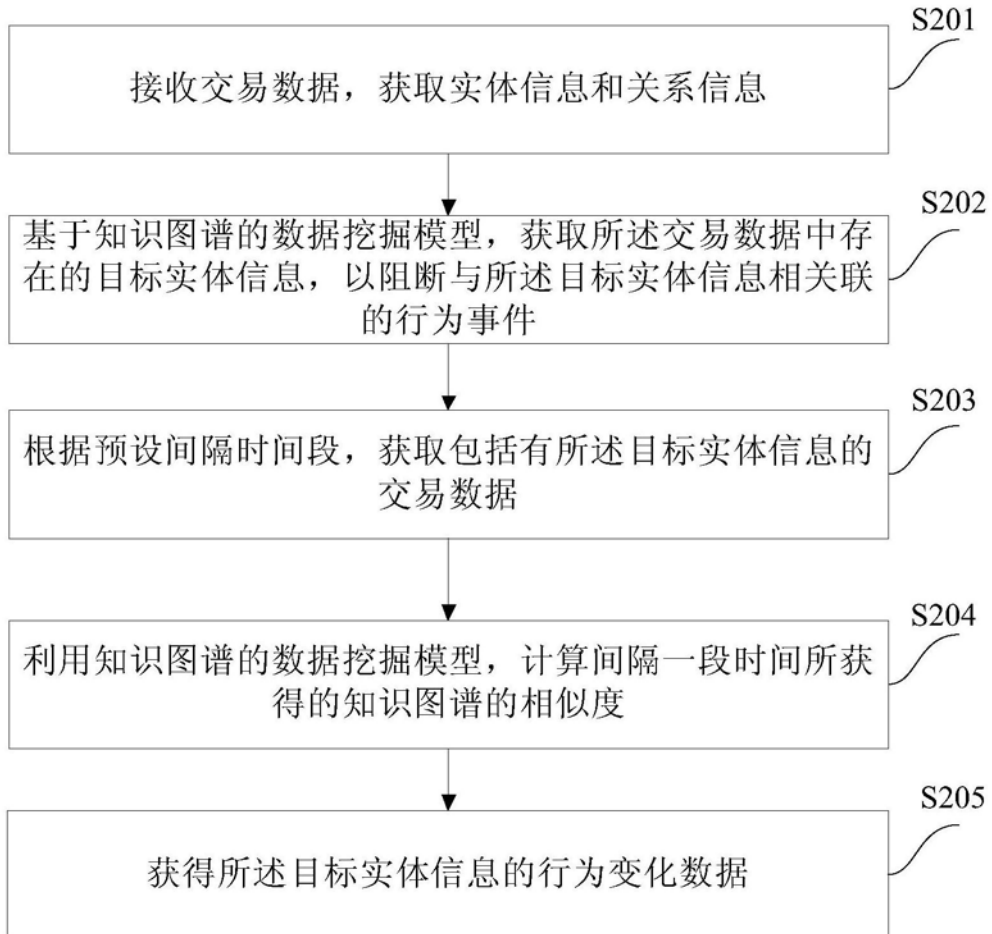


图2

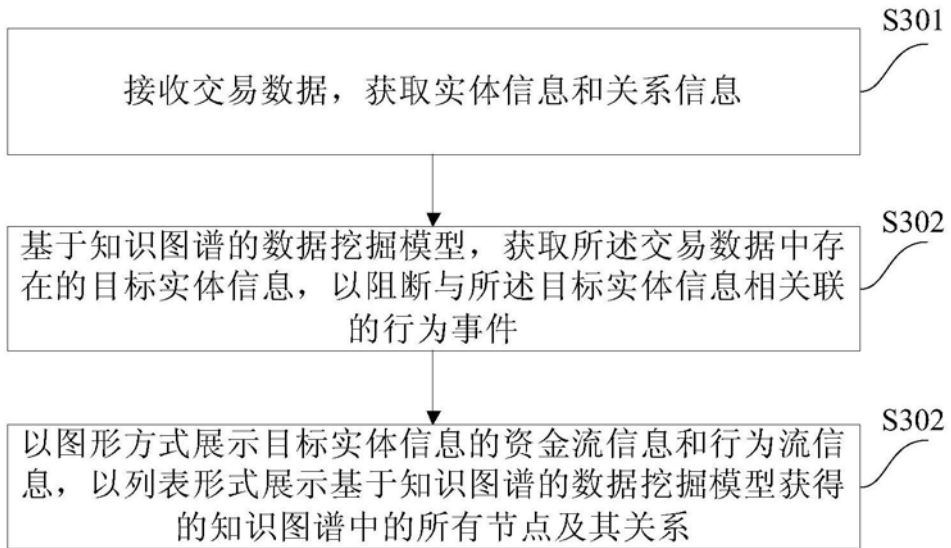


图3

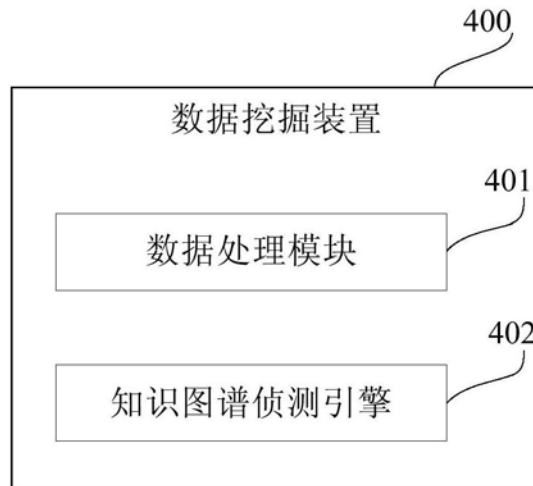


图4

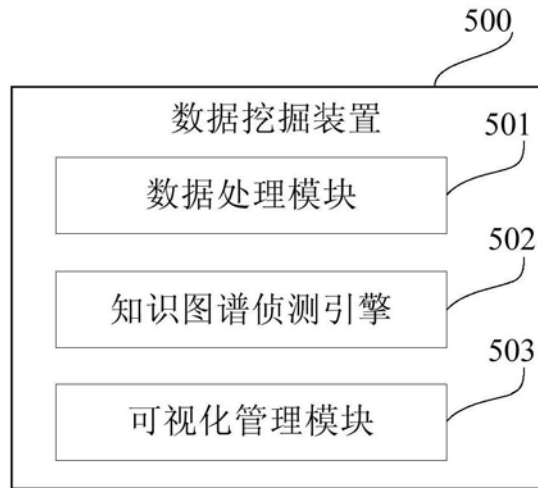


图5

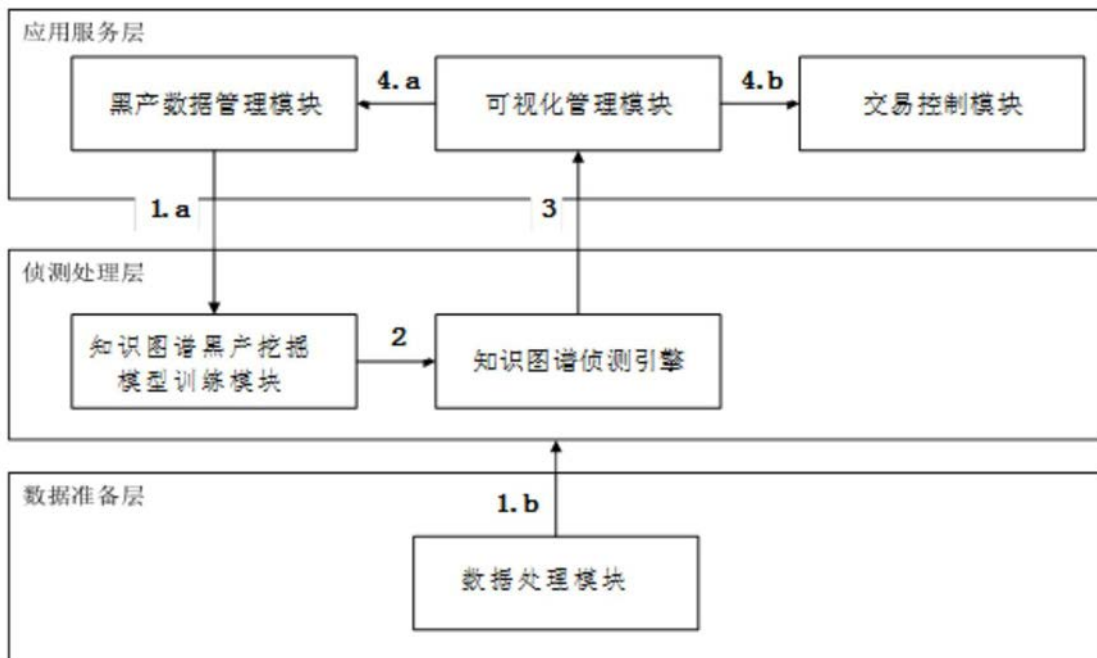


图6

700

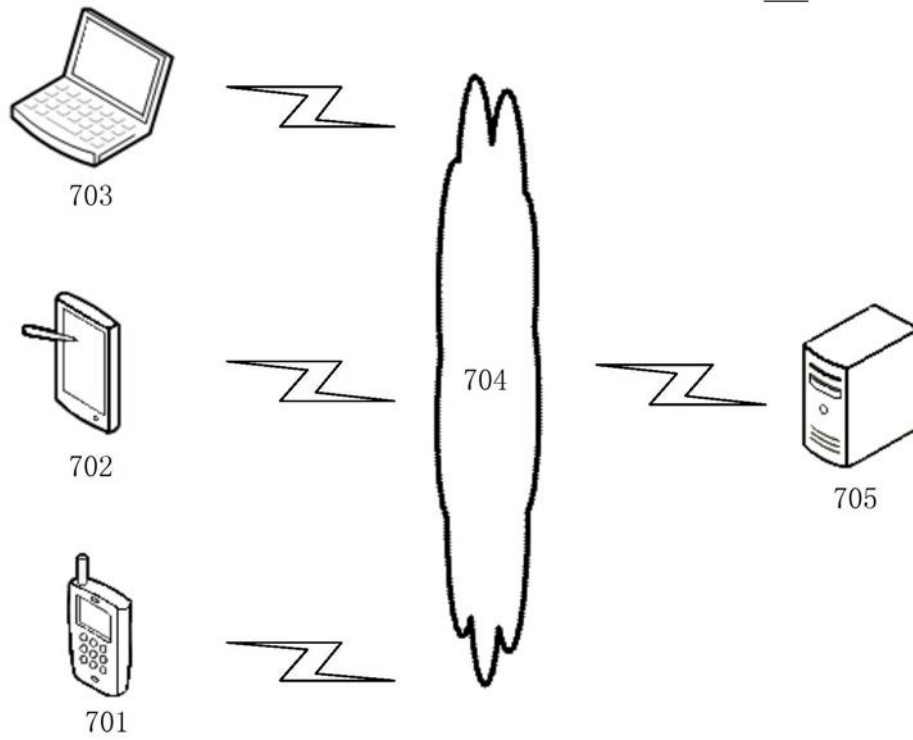


图7

800

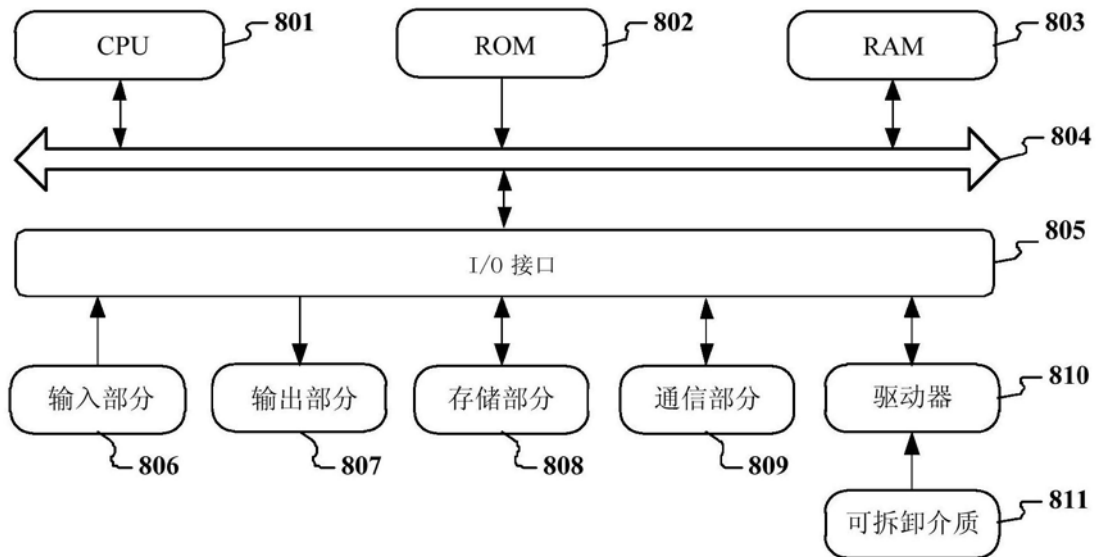


图8