(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113553446 A
(43)申请公布日 2021.10.26

(21)申请号 202110858461.2

(22)申请日 2021.07.28

(71)申请人 厦门国际银行股份有限公司
地址 361000 福建省厦门市鹭江道8-10号国际银行大厦1-6层

(72)发明人 洪镇宇 张志远 胡涛

(74)专利代理机构 厦门仕诚联合知识产权代理事务所(普通合伙) 35227
代理人 乐珠秀

(51)Int.Cl.
G06F 16/36(2019.01)
G06K 9/62(2006.01)
G06Q 40/00(2012.01)

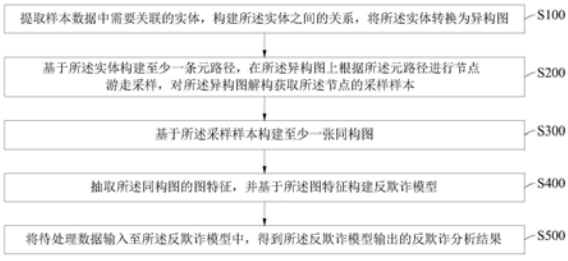
权利要求书2页 说明书9页 附图2页

(54)发明名称

一种基于异构图解构的金融反欺诈方法及装置

(57)摘要

本发明提供了一种基于异构图解构的金融反欺诈方法及装置,涉及金融反诈骗技术领域,该方法包括以下步骤:提取样本数据中需要关联的实体,构建实体之间的关系,将实体转换为异构图;基于实体构建至少一条元路径,在异构图上根据元路径进行节点游走采样,对异构图解构获取节点的采样样本;其中,元路径由对称的实体构成,元路径中的实体对应为元路径的节点;基于采样样本构建至少一张同构图;抽取同构图的图特征,并基于图特征构建反欺诈模型;将待处理数据输入至反欺诈模型中,得到反欺诈模型输出的反欺诈分析结果,本发明充分利用异构图所包含的实体信息,有效的替换掉专家经验制定规约规则,构建出通用的自适应的金融反欺诈解决方案。



CN 113553446 A

1.一种基于异构图解构的金融反欺诈方法,其特征在于,包括以下步骤:

提取样本数据中需要关联的实体,构建所述实体之间的关系,将所述实体转换为异构图;

基于所述实体构建至少一条元路径,在所述异构图上根据所述元路径进行节点游走采样,对所述异构图解构获取所述节点的采样样本;其中,所述元路径由对称的所述实体构成,所述元路径中的实体对应为所述元路径的节点;

基于所述采样样本构建至少一张同构图;

抽取所述同构图的图特征,并基于所述图特征构建反欺诈模型;

将待处理数据输入至所述反欺诈模型中,得到所述反欺诈模型输出的反欺诈分析结果。

2.根据权利要求1所述的基于异构图解构的金融反欺诈方法,其特征在于,所述基于所述实体构建至少一条元路径,在所述异构图上根据所述元路径进行节点游走采样,对所述异构图解构获取所述节点的采样样本步骤中,根据采样的顺序、单节点采样次数以及元路径循环轮数,并根据所述元路径中设置的实体遍历顺序在节点上游走采样。

3.根据权利要求1所述的基于异构图解构的金融反欺诈方法,其特征在于,所述基于所述采样样本构建至少一张同构图步骤中,保留所述采样样本中同类型的所述实体,去除不同类型的所述实体,并基于同类型的所述实体构建至少一张同构图。

4.根据权利要求1所述的基于异构图解构的金融反欺诈方法,其特征在于,所述抽取所述同构图的图特征,并基于所述图特征构建反欺诈模型步骤中,图特征抽取方法包括图嵌入算法、图挖掘算法、图神经网络和聚合方式。

5.根据权利要求1所述的基于异构图解构的金融反欺诈方法,其特征在于,所述实体包括客户相关的身份标识、手机号、地址、公司、联系人标识和银行卡。

6.一种基于异构图解构的金融反欺诈装置,其特征在于,包括:

异构图构建模块(100),用于提取样本数据中需要关联的实体,构建所述实体之间的关系,将所述实体转换为异构图;

解构模块(200),用于基于所述实体构建至少一条元路径,在所述异构图上根据所述元路径进行节点游走采样,对所述异构图解构获取所述节点的采样样本;其中,所述元路径由对称的所述实体构成,所述元路径中的实体对应为所述元路径的节点;

同构图构建模块(300),用于基于所述采样样本构建至少一张同构图;

模型构建模块(400),抽取所述同构图的图特征,并基于所述图特征构建反欺诈模型;

反欺诈处理模块(500),用于将待处理数据输入至所述反欺诈模型中,得到所述反欺诈模型输出的反欺诈分析结果。

7.根据权利要求6所述的基于异构图解构的金融反欺诈装置,其特征在于,所述解构模块(200)中根据采样的顺序、单节点采样次数以及元路径循环轮数,并根据所述元路径中设置的实体遍历顺序在节点上游走采样。

8.根据权利要求6所述的基于异构图解构的金融反欺诈装置,其特征在于,所述同构图构建模块(300)中保留所述采样样本中同类型的所述实体,去除不同类型的所述实体,并基于同类型的所述实体构建至少一张同构图。

9.一种电子设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运

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至5任一项所述基于异构图解构的金融反欺诈方法的步骤。

10.一种非暂态计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求1至5任一项所述基于异构图解构的金融反欺诈方法的步骤。

一种基于异构图解构的金融反欺诈方法及装置

技术领域

[0001] 本发明涉及金融反诈骗技术领域，尤其涉及一种基于异构图解构的金融反欺诈方法及装置。

背景技术

[0002] 在互联网金融反欺诈领域，与传统的反欺诈不同，互联网上的各类交易行为是虚拟的，因此对于互联网金融反欺诈，对数据的处理尤为关键。尤其是金融贷款、交易欺诈等反金融欺诈领域通常需要对用户的信息进行核查，以进行欺诈风险管控。比如说，是金融贷款、交易欺诈等反欺诈金融领域通常需要对用户的信息进行核查、分析与挖掘，以进行欺诈风险管控。现有的贷款业务核查模式通常是采用人工结合专家系统或人工智能模型，上述两种模式在现实中存在一些弊端，例如借助专家系统存在一定的主观性、审查要素间缺乏总体关联性、复杂关联计算只能抽取历史跑批数据而无法实时计算、对审批对象缺乏整体性认识等等。

[0003] 知识图谱目前大量运用于分析领域，传统知识图谱运用方法主要是通过计算最短路径、扩散汇聚、节点重要性等抽取节点特征。近年来对于异构图的研究比较多，异构图的利用主要分为两种方式，一种是采用图表示算法，直接计算异构图中节点的相关特征，其所包含的问题是，图表示算法一般难以做到实时更新、实时计算整张图上的特征，且图表示算法一般消耗的资源较多，在应用角度存在诸多难点；第二种是采用专家经验制定一些规则，对异构图中的关系或者节点进行归约，从而将其转化为同构图进行下一步计算，但是专家经验在不同产品、不同场景下并不是通用的，且针对不同产品可能已知的规则都不适用，涉及到开发新规则需要耗费大量的人力物力，且专家经验一般只关注于已有实体，对于未知实体的加入需要重新开发规则，流程繁琐且也可能存在误差，并且大部分归约后仅使用一个实体相关的同构图，忽略了其它实体同构图可能隐含的信息，归约规则大部分也作为商业机密，难以进行实际的运用。

[0004] 综上，异构图虽然可以用于互联网金融反欺诈，但是如何充分利用异构图所包含的实体、如何有效的替换掉专家经验制定规约规则，构建出通用的自适应的金融反欺诈解决方案时目前反诈骗行业亟待解决的重要课题。

发明内容

[0005] 有鉴于此，本发明提供一种基于异构图解构的金融反欺诈方法及装置，用以解决现有技术中异构图利用过程中的缺陷，实现充分利用异构图所包含的实体信息，有效的替换掉专家经验制定规约规则，构建出通用的自适应的金融反欺诈解决方案。

[0006] 基于上述目的，本发明提供了一种基于异构图解构的金融反欺诈方法，包括以下步骤：

[0007] 提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图；

[0008] 基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本；其中，所述元路径由对称的所述实体构成，所述元路径中的实体对应为所述元路径的节点；

[0009] 基于所述采样样本构建至少一张同构图；

[0010] 抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型；

[0011] 将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果。

[0012] 可选的，所述基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本步骤中，根据采样的顺序、单节点采样次数以及元路径循环轮数，并根据所述元路径中设置的实体遍历顺序在节点上游走采样。

[0013] 可选的，所述基于所述采样样本构建至少一张同构图步骤中，保留所述采样样本中同类型的所述实体，去除不同类型的所述实体，并基于同类型的所述实体构建至少一张同构图。

[0014] 可选的，所述抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型步骤中，图特征抽取方法包括图嵌入算法、图挖掘算法、图神经网络和聚合方式。

[0015] 可选的，所述实体包括客户相关的身份标识、手机号、地址、公司、联系人标识和银行卡。

[0016] 本发明提供了一种基于异构图解构的金融反欺诈装置，包括：

[0017] 异构图构建模块，用于提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图；

[0018] 解构模块，用于基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本；其中，所述元路径由对称的所述实体构成，所述元路径中的实体对应为所述元路径的节点；

[0019] 同构图构建模块，用于基于所述采样样本构建至少一张同构图；

[0020] 模型构建模块，抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型；

[0021] 反欺诈处理模块，用于将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果。

[0022] 可选的，所述解构模块中根据采样的顺序、单节点采样次数以及元路径循环轮数，并根据所述元路径中设置的实体遍历顺序在节点上游走采样。

[0023] 可选的，所述同构图构建模块中保留所述采样样本中同类型的所述实体，去除不同类型的所述实体，并基于同类型的所述实体构建至少一张同构图。

[0024] 本发明还提供一种电子设备，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上述任一种所述基于异构图解构的金融反欺诈方法的步骤。

[0025] 本发明还提供一种非暂态计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现如上述任一种所述基于异构图解构的金融反欺诈方法的步骤。

[0026] 从上面所述可以看出，本发明提供的基于异构图解构的金融反欺诈方法及装置，通过提取不同类型的实体，并充分利用了各实体的连接关系构建异构图，根据元路径的不

同，在游走采样的过程中具备获取周围或较远同类实体关联性的能力，通过对异构图结构能够自动地构建出每个实体对应的同构图，并能够科学的分配节点间的连接权重，极大提升了知识图谱使用时的自动化程度，在实际项目中提升了字段的利用率，针对能够加入知识图谱的实体类别都能够快速的构建对应的同构图。同时，在保留一定的异构图表达能力时，又不丧失实时性。对于新节点加入，只需要更新进异构图，从新节点按照预设参数进行遍历，再将不同实体节点更新入对应的同构图。若是在同构图中加入节点属性，既采用属性图的方式进行计算，则完全能够做到实时计算，相较于图嵌入或现有技术中的同构图使用方法具备更强的实时性与表达能力，该基于异构图解构的金融反欺诈方法及装置中构建的反欺诈模型充分利用异构图所包含的实体信息，有效的替换掉专家经验制定规约规则，构建出通用的自适应的金融反欺诈解决方案。

附图说明

[0027] 为了更清楚地说明本发明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发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

[0028] 图1是本发明提供的基于异构图解构的金融反欺诈方法的流程示意图；

[0029] 图2是本发明提供的基于异构图解构的金融反欺诈装置的结构示意图；

[0030] 图3是本发明提供的电子设备的结构示意图。

具体实施方式

[0031] 为使本发明的目的、技术方案和优点更加清楚明白，以下结合具体实施例，并参照附图，对本发明进一步详细说明。

[0032] 需要说明的是，除非另外定义，本发明实施例使用的技术术语或者科学术语应当为本公开所属领域内具有一般技能的人士所理解的通常意义。本公开中使用的“第一”、“第二”以及类似的词语并不表示任何顺序、数量或者重要性，而只是用来区分不同的组成部分。“包括”或者“包含”等类似的词语意指出现该词前面的元件或者物件涵盖出现在该词后面列举的元件或者物件及其等同，而不排除其他元件或者物件。“连接”或者“相连”等类似的词语并非限定于物理的或者机械的连接，而是可以包括电性的连接，不管是直接的还是间接的。“上”、“下”、“左”、“右”等仅用于表示相对位置关系，当被描述对象的绝对位置改变后，则该相对位置关系也可能相应地改变。

[0033] 作为本发明的一个优选实施例，本发明提供本发明提供了一种基于异构图解构的金融反欺诈方法，包括以下步骤：

[0034] 提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图；

[0035] 基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本；其中，所述元路径由对称的所述实体构成，所述元路径中的实体对应为所述元路径的节点；

[0036] 基于所述采样样本构建至少一张同构图；

[0037] 抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型；

[0038] 将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果。

[0039] 本发明提供了一种基于异构图解构的金融反欺诈装置，包括：

[0040] 异构图构建模块，用于提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图；

[0041] 解构模块，用于基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本；其中，所述元路径由对称的所述实体构成，所述元路径中的实体对应为所述元路径的节点；

[0042] 同构图构建模块，用于基于所述采样样本构建至少一张同构图；

[0043] 模型构建模块，抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型；

[0044] 反欺诈处理模块，用于将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果。

[0045] 通过该基于异构图解构的金融反欺诈方法及装置，通过提取不同类型的实体，并充分利用了各实体的连接关系构建异构图，根据元路径的不同，在游走采样的过程中具备获取周围或较远同类实体关联性的能力，通过对异构图结构能够自动地构建出每个实体对应的同构图，并能够科学的分配节点间的连接权重，极大提升了知识图谱使用时的自动化程度，在实际项目中提升了字段的利用率，针对能够加入知识图谱的实体类别都能够快速的构建对应的同构图。同时，在保留一定的异构图表达能力时，又不丧失实时性。对于新节点加入，只需要更新进异构图，从新节点按照预设参数进行遍历，再将不同实体节点更新入对应的同构图。若是在同构图中加入节点属性，既采用属性图的方式进行计算，则完全能够做到实时计算，相较于图嵌入或现有技术中的同构图使用方法具备更强的实时性与表达能力，该基于异构图解构的金融反欺诈方法及装置中构建的反欺诈模型充分利用异构图所包含的实体信息，有效的替换掉专家经验制定规约规则，构建出通用的自适应的金融反欺诈解决方案。

[0046] 下面结合附图对本发明基于异构图解构的金融反欺诈方法及装置的较佳实施例进行说明。

[0047] 请参阅图1，该方法包括以下步骤：

[0048] S100、提取样本数据中需要关联的实体，构建实体之间的关系，将实体转换为异构图。

[0049] 在步骤S100中，针对反欺诈项目，实体包括客户相关的身份标识(ID，后续简称ID)、手机号(MOBILE，后续简称M)、地址(ADDRESS，后续简称ADDR)、公司(CORPORATION，后续简称CORP)、联系人标识(LINKMAN，后续简称LINK)、银行卡(BANKCARD，后续简称BC)等，并构建实体之间的关系，关系包括“身份标识-持有-手机”、“身份标识-居住于-地址”、“公司-位于-地址”，“身份标识-就职-公司”、“身份标识-拥有-银行卡”、“身份标识-联系-联系人标识”、“联系人标识-持有-手机号”、等。

[0050] S200、基于实体构建至少一条元路径，在异构图上根据元路径进行节点游走采样，对异构图解构获取关联节点的采样样本；其中，元路径由对称的实体构成，元路径中的实体对应为元路径的节点。

[0051] 元路径(meta-path)是一种实体与实体间的路径形式,描述了实体间的一种综合关系,在步骤S200中采用基于元路径策略的随机游(random-walk)思想,能够通过节点上的游走采样到节点间的关系。具体的,根据实体采样的顺序、单节点采样次数、元路径循环轮数等参数,之后根据元路径中设置的实体遍历顺序在节点上游走采样。以反欺诈项目为例,可构建元路径如ID-M-ID、ID-ADDR-CORP-ADDR-ID、CORP-ADDR-CORP等。

[0052] 在本实施例中,多条元路径中的实体类型能够覆盖想要生成的同构图的实体类型就可以。

[0053] S300、基于采样样本构建至少一张同构图。

[0054] 在步骤S300中,保留采样样本中同类型的实体,去除不同类型的实体,并基于同类型的实体构建至少一张同构图,保留的同类别实体根据采样顺序是一个序列,之后结合ngram思想或其它方式分配关系的权重,也可以选择不分配权重。其中,若是采用分配权重的做法,例如对于id1id2id2样本来说,若是采用2gram,可拆分为(id1,id2),(id2,id2),对于同类别实体可以选择添加自环或是去除;对于3gram来说,可组成(id1id2,id2),此时对于第二个id2来说与之前出现的id1与第一个id2相关,分解为(id1,id2)与(id2,id2),并根据次数赋予相应权值。对于id1id2id3样本来说,采用3gram,则生成(id1id2,id3),id3对于id1与id2都形成一条关系,并赋予相应权重。对于同一类实体,采用不同元路径采样生成的图结构及关系权值是不同的,在2gram中,可以简单的把同样连接的次数累加作为权值,以判别节点间关系的远近。在3gram中,在每个3gram关系中,以第一、二个节点为核心,第三个节点为后续节点,拆分为对应的二元组之后,构成对应的同构图。选择越大的n能够在一定程度上学习远距离节点的关联性。最后针对同样方法构建的同构图再做节点的归一化,若是针对一个实体有多个方法构建的同构图,权重与边可在归一化后进行叠加。因此在同一类实体所构建的不同的同构图上,可以分别进行利用或是对每张图权值归一化后将关系进行合并,因此,最终步骤S300得出多个实体的一张或多张同构图。

[0055] S400、抽取同构图的图特征,并基于图特征构建反欺诈模型。

[0056] 在步骤S400中,同构图中图特征的抽取可以使用图嵌入算法(例如line、node2vec、metapath2vec等)计算每个节点的特征,以抽取节点的基本属性,如度、二阶邻居数、网络结构形态等,或采用传统的图计算方法,例如pagerank、中心性计算、最短路径、节点的各类中心性、节点相似度等,也可采用属性图的方式,对节点赋予属性特征,通过节点聚合操作计算节点的局部特征,常见的聚合操作包括和、众数、最小值、最大值等;还可以通过图神经网络对解构后的图进行建模计算,抽取图中的节点特征或全图特征,再进行后续利用。

[0057] S500、将待处理数据输入至反欺诈模型中,得到反欺诈模型输出的反欺诈分析结果。

[0058] 在步骤S500中,反欺诈模型构建时,包括有监督方法与无监督方法,有监督方法其中一个最优实现是基于梯度提升树的lightgbm算法,无监督方法包括传统聚类算法及基于深度学习的自编码器单分类器方法。基于lightgbm的算法主要使用的是由进件信息或其它表格数据中抽取的特征,再结合每个进件信息(在金融反欺诈领域,待处理数据主要包括进件信息、征信信息、交易信息、采购的三方信息等)中已知的贷款人、电话、地址等信息,在已构建的同构图中进行相关特征的抽取。传统聚类算法其中一个最优实现是K-means算法,主

要区别是使用更加严格的特征筛选方式，挑选出几个进件信息关键特征、结合少量图特征，根据最终特征维度再采取降维或剔除特征的方法，最后使用K-means进行聚类。基于深度学习的自编码器方法则是采用自编码器(Autoencoder)架构，采用重构误差的方式，只学习正常样本的表格数据特征与图特征，采用单分类器思想，将欺诈样本作为异常进行看待。单分类器除了使用自编码器之外，还可以使用单分类支持向量机(One Class SVM)、孤立森林(Isolation Forest)。

[0059] 该方法通过步骤S100提取不同类型的实体，并充分利用了各实体的连接关系构建异构图，通过步骤S200根据元路径的不同，在游走采样的过程中具备获取周围或较远同类实体关联性的能力，通过步骤300对异构图结构能够自动地构建出每个实体对应的同构图，并能够科学的分配节点间的连接权重，极大提升了知识图谱使用时的自动化程度，在实际项目中提升了字段的利用率，针对能够加入知识图谱的实体类别都能够快速的构建对应的同构图。同时，在保留一定的异构图表达能力时，又不丧失实时性。对于新节点加入，只需要更新进异构图，从新节点按照预设参数进行遍历，再将不同实体节点更新入对应的同构图。若是在同构图中加入节点属性，既采用属性图的方式进行计算，则完全能够做到实时计算，相较于图嵌入或现有技术中的同构图使用方法具备更强的实时性与表达能力，该方法中构建的反欺诈模型充分利用异构图所包含的实体信息，有效的替换掉专家经验制定规约规则，构建出通用的自适应的金融反欺诈解决方案。

[0060] 下面对本发明提供的基于异构图解构的金融反欺诈装置进行描述，下文描述的基于异构图解构的金融反欺诈装置与上文描述的基于异构图解构的金融反欺诈方法可相互对应参照。

[0061] 请参阅图2，该装置包括：

[0062] 异构图构建模块100，用于提取样本数据中需要关联的实体，构建实体之间的关系，将实体转换为异构图。

[0063] 在异构图构建模块100中，针对反欺诈项目，实体包括客户相关的身份标识(ID，后续简称ID)、手机号(MOBILE，后续简称M)、地址(ADDRESS，后续简称ADDR)、公司(CORPORATION，后续简称CORP)、联系人标识(LINKMAN，后续简称LINK)、银行卡(BANKCARD，后续简称BC)等，并构建实体之间的关系，关系包括“身份标识-持有-手机”、“身份标识-居住于-地址”、“公司-位于-地址”，“身份标识-就职-公司”、“身份标识-拥有-银行卡”、“身份标识-联系-联系人标识”、“联系人标识-持有-手机号”等。

[0064] 解构模块200，用于基于实体构建至少一条元路径，在异构图上根据元路径进行节点游走采样，对异构图解构获取关联节点的采样样本；其中，元路径由对称的实体构成，元路径中的实体对应为元路径的节点。

[0065] 元路径是一种知识图谱上的节点遍历方式，在解构模块200中基于随机游走思想，能够通过节点上的游走采样到节点间的关系。具体的，根据实体采样的顺序、单节点采样次数、元路径循环轮数等参数，之后根据元路径中设置的实体遍历顺序在节点上游走采样。以反欺诈项目为例，可构建元路径如ID-M-ID、ID-ADDR-CORP-ADDR-ID、CORP-ADDR-CORP等。

[0066] 在本实施例中，多条元路径中的实体类型能够覆盖想要生成的同构图的实体类型就可以。

[0067] 同构图构建模块300，用于基于采样样本构建至少一张同构图。

[0068] 在同构图构建模块300中，保留采样样本中同类型的实体，去除不同类型的实体，并基于同类型的实体构建至少一张同构图，保留的同类别实体根据采样顺序是一个序列，之后结合ngram思想或其它方式分配关系的权重，也可以选择不分配权重。其中，若是采用分配权重的做法，例如对于id1id2id2样本来说，若是采用2gram，可拆分为(id1,id2)，(id2,id2)，对于同类别实体可以选择添加自环或是去除；对于3gram来说，可组成(id1id2，id2)，此时对于第二个id2来说与之前出现的id1与第一个id2相关，分解为(id1,id2)与(id2,id2)，并根据次数赋予相应权值。对于id1id2id3样本来说，采用3gram，则生成(id1id2,id3)，id3对于id1与id2都形成一条关系，并赋予相应权重。对于同一类实体，采用不同元路径采样生成的图结构及关系权值是不同的，在2gram中，可以简单的把同样连接的次数累加作为权值，以判别节点间关系的远近。在3gram中，在每个3gram关系中，以第一、二个节点为核心，第三个节点为后续节点，拆分为对应的二元组之后，构成对应的同构图。选择越大的n能够在一定程度上学习远距离节点的关联性。最后针对同样方法构建的同构图再做节点的归一化，若是针对一个实体有多个方法构建的同构图，权重与边可在归一化后进行叠加。因此在同一类实体所构建的不同的同构图上，可以分别进行利用或是对每张图权值归一化后将关系进行合并，因此，最终步骤S300得出多个实体的一张或多张同构图。

[0069] 模型构建模块400，用于抽取同构图的图特征，并基于图特征构建反欺诈模型。

[0070] 在模型构建模块400中，同构图中图特征的抽取可以使用图嵌入算法(例如line、node2vec、metapath2vec等)计算每个节点的特征，以抽取节点的基本属性，如度、二阶邻居数、网络结构形态等，或采用传统的图计算方法，例如pagerank、中心性计算、最短路径、节点的各类中心性、节点相似度等，也可采用属性图的方式，对节点赋予属性特征，通过节点聚合操作计算节点的局部特征，常见的聚合操作包括和、众数、最小值、最大值等；还可以通过图神经网络对解构后的图进行建模计算，抽取图中的节点特征或全图特征，再进行后续利用。

[0071] 反欺诈处理模块500，用于将待处理数据输入至反欺诈模型中，得到反欺诈模型输出的反欺诈分析结果。

[0072] 在反欺诈处理模块500中，反欺诈模型构建时，包括有监督方法与无监督方法，有监督方法其中一个最优实现是基于梯度提升树的lightgbm算法，无监督方法包括传统聚类算法及基于深度学习的自编码器单分类器方法。基于lightgbm的算法主要使用的是由进件信息或其它表格数据中抽取的特征，再结合每个进件信息(在金融反欺诈领域，待处理数据主要包括进件信息、征信信息、交易信息、采购的三方信息等)中已知的贷款人、电话、地址等信息，在已构建的同构图中进行相关特征的抽取。传统聚类算法其中一个最优实现是K-means算法，主要区别是使用更加严格的特征筛选方式，挑选出几个进件信息关键特征、结合少量图特征，根据最终特征维度再采取降维或剔除特征的方法，最后使用K-means进行聚类。基于深度学习的自编码器方法则是采用自编码器(Autoencoder)架构，采用重构误差的方式，只学习正常样本的表格数据特征与图特征，采用单分类器思想，将欺诈样本作为异常进行看待。单分类器除了使用自编码器之外，还可以使用单分类支持向量机(One Class SVM)、孤立森林(Isolation Forest)。

[0073] 该装置通过异构图构建模块100提取不同类型的实体，并充分利用了各实体的连接关系构建异构图，通过解构模块200根据元路径的不同，在游走采样的过程中具备获取周

围或较远同类实体关联性的能力，通过同构图构建模块300对异构图结构能够自动地构建出每个实体对应的同构图，并能够科学的分配节点间的连接权重，极大提升了知识图谱使用时的自动化程度，在实际项目中提升了字段的利用率，针对能够加入知识图谱的实体类别都能够快速的构建对应的同构图。同时，在保留一定的异构图表达能力时，又不丧失实时性。对于新节点加入，只需要更新进异构图，从新节点按照预设参数进行遍历，再将不同实体节点更新入对应的同构图。若是在同构图中加入节点属性，既采用属性图的方式进行计算，则完全能够做到实时计算，相较于图嵌入或现有技术中的同构图使用方法具备更强的实时性与表达能力，该方法中构建的反欺诈模型充分利用异构图所包含的实体信息，有效的替换掉专家经验制定规约规则，构建出通用的自适应的金融反欺诈解决方案。

[0074] 图3示例了一种电子设备的实体结构示意图，如图3所示，该电子设备可以包括：处理器(processor)810、通信接口(Communications Interface)820、存储器(memory)830和通信总线840，其中，处理器810，通信接口820，存储器830通过通信总线840完成相互间的通信。处理器810可以调用存储器830中的逻辑指令，以执行基于异构图解构的金融反欺诈方法，该方法包括以下步骤：

[0075] S100、提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图；

[0076] S200、基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本；其中，所述元路径由对称的所述实体构成，所述元路径中的实体对应为所述元路径的节点；

[0077] S300、基于所述采样样本构建至少一张同构图；

[0078] S400、抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型；

[0079] S500、将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果。

[0080] 此外，上述的存储器830中的逻辑指令可以通过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，可以存储在一个计算机可读取存储介质中。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一个存储介质中，包括若干指令用以使得一台计算机设备(可以是个人计算机，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。而前述的存储介质包括：U盘、移动硬盘、只读存储器(ROM，Read-Only Memory)、随机存取存储器(RAM，Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0081] 另一方面，本发明还提供一种计算机程序产品，所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，计算机能够执行上述各方法所提供的基于异构图解构的金融反欺诈方法，该方法包括以下步骤：

[0082] S100、提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图；

[0083] S200、基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本；其中，所述元路径由对称的所

述实体构成，所述元路径中的实体对应为所述元路径的节点；

[0084] S300、基于所述采样样本构建至少一张同构图；

[0085] S400、抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型；

[0086] S500、将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果。

[0087] 又一方面，本发明还提供一种非暂态计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现以执行上述各提供的基于异构图解构的金融反欺诈方法，该方法包括以下步骤：

[0088] S100、提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图；

[0089] S200、基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本；其中，所述元路径由对称的所述实体构成，所述元路径中的实体对应为所述元路径的节点；

[0090] S300、基于所述采样样本构建至少一张同构图；

[0091] S400、抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型；

[0092] S500、将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果。

[0093] 以上所描述的装置实施例仅仅是示意性的，其中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模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性的劳动的情况下，即可以理解并实施。

[0094] 所属领域的普通技术人员应当理解：以上任何实施例的讨论仅为示例性的，并非旨在暗示本公开的范围(包括权利要求)被限于这些例子；在本发明的思路下，以上实施例或者不同实施例中的技术特征之间也可以进行组合，步骤可以以任意顺序实现，并存在如上所述的本发明的不同方面的许多其它变化，为了简明它们没有在细节中提供。

[0095] 本发明的实施例旨在涵盖落入所附权利要求的宽泛范围之内的所有这样的替换、修改和变型。因此，凡在本发明的精神和原则之内，所做的任何省略、修改、等同替换、改进等，均应包含在本发明的保护范围之内。

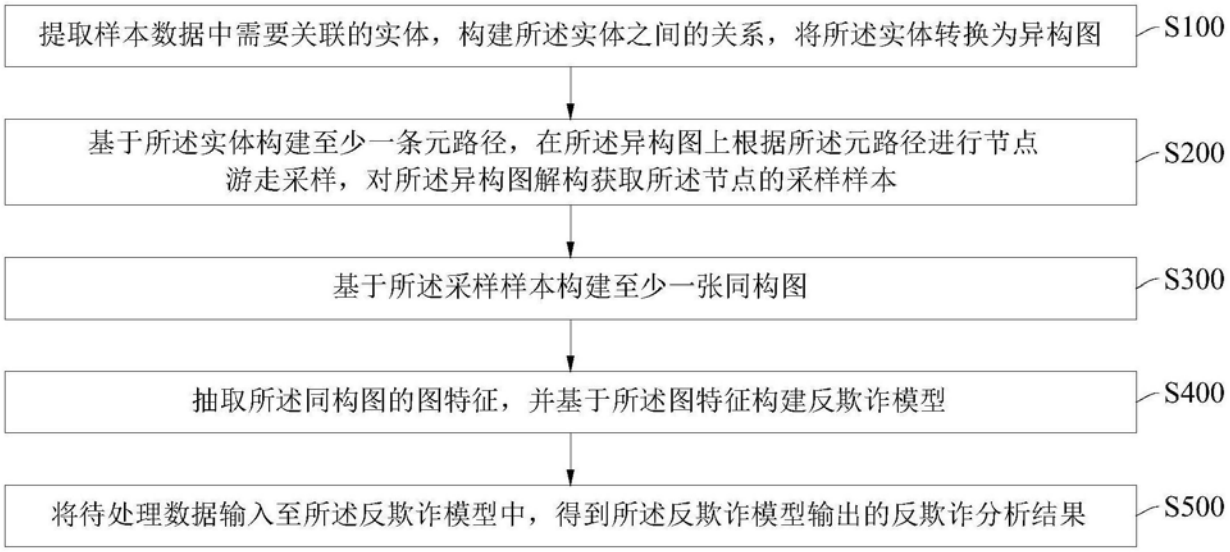

提取样本数据中需要关联的实体，构建所述实体之间的关系，将所述实体转换为异构图
S100
基于所述实体构建至少一条元路径，在所述异构图上根据所述元路径进行节点游走采样，对所述异构图解构获取所述节点的采样样本
S200
基于所述采样样本构建至少一张同构图
S300
抽取所述同构图的图特征，并基于所述图特征构建反欺诈模型
S400
将待处理数据输入至所述反欺诈模型中，得到所述反欺诈模型输出的反欺诈分析结果
S500

图1

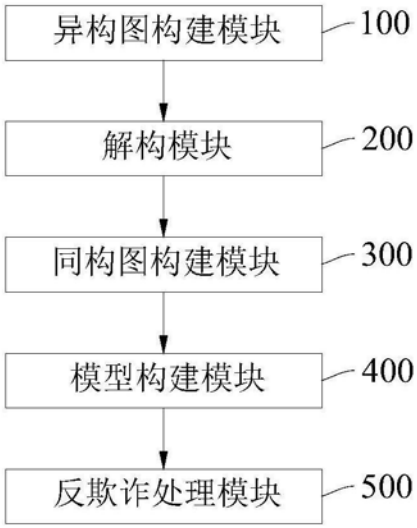

异构图构建模块
100
解构模块
200
同构图构建模块
300
模型构建模块
400
反欺诈处理模块
500

图2

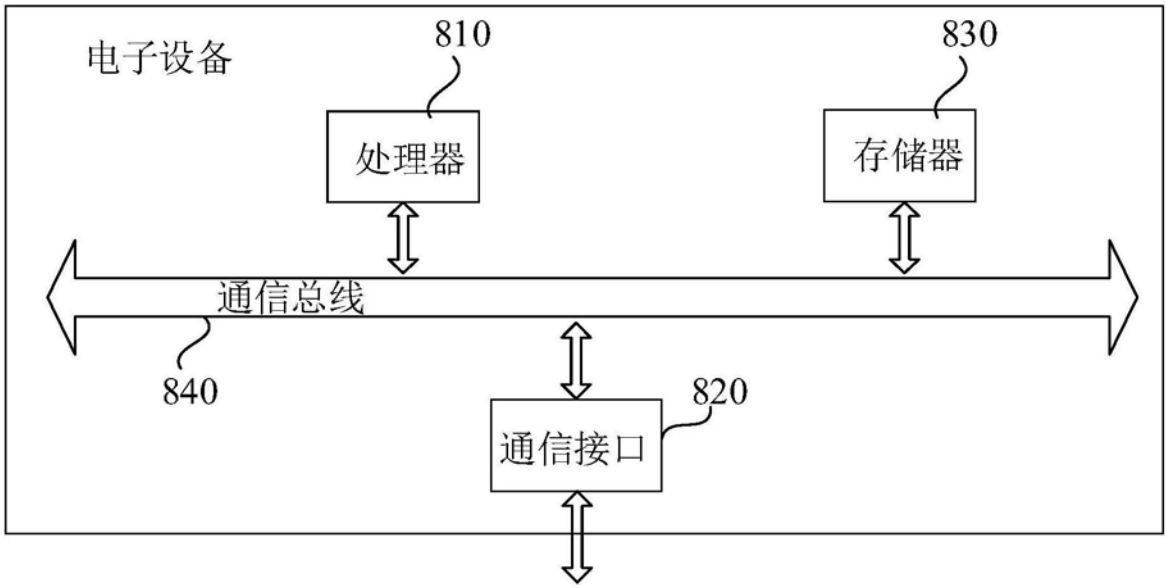


图3