

更新知识图谱的方法及装置

本申请要求于 2022 年 03 月 23 日提交中国国家知识产权局专利局、申请号为 202210290077.1、发明名称为“更新知识图谱的方法及装置”的中国专利申请的优先权，
5 其全部内容通过引用结合在本申请中。

技术领域

本说明书一个或多个实施例涉及计算机技术领域，尤其涉及更新知识图谱的方法及装置。
10

背景技术

知识图谱 (Knowledge Graph) 是以图模式描述真实世界中的各种实体及其关系的一种语义网络。通过知识图谱结合专家经验及先验数据，可以解释图谱中关系、规则的正确性，以及推理图中未出现的关系、规则。通过知识图谱可以进行与实体的关联关系相关的业务
15 处理。近年来也出现了一下知识图谱平台，作为以知识图谱为核心能力的中台，面向各种业务提供知识管理、知识推理、知识服务的能力，以及与这些能力相配套的图谱解决方案。

发明内容

本说明书一个或多个实施例描述了一种更新知识图谱的方法及装置，用以解决背景技术提到的一个或多个问题。
20

根据第一方面，提供一种更新知识图谱的方法，所述方法包括对知识图谱进行多轮增量更新，其中，一轮增量更新包括：获取该轮增量更新的初始知识图谱；进行更新步骤，包括重复执行的实时更新操作和满足预设的增量更新条件的情况下的增量更新操作，其中，该实时更新操作包括：响应于接收到新的业务数据，利用接收的业务数据对前一实时更新
25 操作中更新后的知识图谱进行更新，该增量更新操作包括：利用该轮增量更新期间产生的业务数据对所述初始知识图谱进行更新，以作为下一轮增量更新的初始知识图谱。

在一个实施例中，所述实时更新操作、所述增量更新操作均包含以下实体链指过程：确定是否存在至少 2 个节点对应的业务主体具有相同特性；在存在的情况下，针对实体链指结果还执行以下实体归一过程：将具有相同特性的节点合并为一个节点，并且具有相同
30 特性的各个节点相应的实体描述信息叠加后作为合并后的节点的实体描述信息。

在一个实施例中，在该轮增量更新是首轮增量更新的情况下，该轮增量更新的初始知识图谱基于对利用全量业务数据构建的知识图谱的实体链指结果进行实体归一得到；在该轮增量更新不是首轮增量更新周期的情况下，该轮增量更新的初始知识图谱基于对前一轮增量更新中的初始知识图谱的增量的实体链指结果进行实体归一得到。

5 在一个实施例中，所述对利用全量业务数据构建的知识图谱全量的实体链指结果通过以下方式获取：针对利用全量业务数据构建的知识图谱中的各个节点分别获取其对应的实体描述信息；根据各个节点各自对应的实体描述信息提取各个节点分别对应的各个特征向量；检测各个特征向量两两之间的相似性；根据两两特征向量的相似性是否满足预定同质条件，识别相应的两两节点是否具有相同特性。

10 在一个实施例中，所述初始知识图谱包括第一节点，针对所述第一节点的第一业务数据为当前接收的新的业务数据，所述响应于当前业务中产生新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新包括：利用所述第一业务信息更新所述第一节点的第一实体描述信息；从更新后的第一实体描述信息中提取第一特征向量；比较所述第一特征向量与其他各个节点的各个其他特征向量一一对应的各个相似性；基于
15 各个相似性是否满足预定同质条件，得到是否存在与所述第一节点具有相同特性的其他节点实时的实体链指结果；基于该实时的实体链指结果对前一实时更新操作中更新后的知识图谱进行更新。

20 在一个实施例中，所述方法还包括：将当前接收的新的业务数据作为增量数据添加至当前增量数据集；所述利用该轮增量更新期间产生的业务数据对所述初始知识图谱进行更新包括：利用当前增量数据集中的各条增量数据进行针对该轮增量更新的初始知识图谱增量的实体链指；利用增量的实体链指结果更新所述初始知识图谱。

在一个实施例中，所述增量更新条件包括：预定周期到达，或者该轮增量更新期间产生的业务数据条数达到预定条数。

25 在一个实施例中，在该轮增量更新不是首轮增量更新的情况下，所述更新步骤还包括：获取基于前一轮增量更新中满足预设的增量更新条件之后的实时更新操作中得到的各个实时的更新结果；根据各个实时的更新结果更新该轮增量更新的初始知识图谱。

在一个实施例中，所述实体描述信息包括属性信息、连接信息中的至少一项。

在一个实施例中，所述特征向量包括以下中的一项，或以下中的多项经嵌入得到的向量：文本语义向量、轨迹向量、图结构向量、图表征向量。

30 在一个实施例中，实时的实体链指过程通过在线搜索引擎完成，基于实时的实体链指

更新当前知识图谱通过在线图存储引擎完成；所述利用增量的实体链指结果更新所述初始知识图谱包括：通过数据转存机制，将所述增量的实体链指结果同步至在线搜索引擎及在线图存储引擎，从而完成所述增量的实体链指结果对该轮增量更新期间内产生的各个实时的实体链指结果的替换，从而利用增量的实体链指结果更新所述初始知识图谱。

- 5 在一个实施例中，在增量数据中涉及的第二业务主体在该轮增量更新的初始知识图谱中不存在相对应的节点的情况下，所述增量更新操作还包括：在该轮增量更新的初始知识图谱中增加与所述第二业务主体相对应的第二节点；基于增加所述第二节点后的知识图谱进行增量的实体链指。

- 10 在一个实施例中，在该轮增量更新为首轮增量更新的情况下，该轮增量更新的首次实时更新操作为：利用接收的业务数据对该轮增量更新的初始知识图谱进行更新。

根据第二方面，提供一种更新知识图谱的装置，所述装置包括：

获取单元，配置为在各轮增量更新中获取初始知识图谱；

- 15 更新单元，配置为在各轮增量更新中进行包括重复执行的实时更新操作和满足预设的增量更新条件的情况下的增量更新操作的更新步骤，其中，该实时更新操作包括：响应于接收到新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新，该增量更新操作包括：利用该轮增量更新期间产生的业务数据对所述初始知识图谱进行更新，以作为下一轮增量更新的初始知识图谱。

根据第三方面，提供了一种计算机可读存储介质，其上存储有计算机程序，当所述计算机程序在计算机中执行时，令计算机执行第一方面的方法。

- 20 根据第四方面，提供了一种计算设备，包括存储器和处理器，其特征在于，所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现第一方面的方法。

通过本说明书实施例提供的方法和装置，在为当前业务提供基于知识图谱的数据支持过程中，采用在线、离线相结合的方式更新知识图谱。首先可以基于利用全量业务数据离线构建的初始知识图谱进行全量的实体链指，以初始化知识图谱作为冷启动的知识图谱。

- 25 之后，针对冷启动的知识图谱进行多轮增量更新。在单轮增量更新期间，一方面，基于实时产生的业务数据提供在线实时的知识图谱更新，另一方面，按照预设的增量更新条件，在增量更新条件满足时，按照当前轮增量更新期间新增的业务数据提供离线的知识图谱增量的实体链指，并用离线增量的实体链指结果代替实时的实体链指结果更新当前轮增量更新初始的知识图谱。如此，各个轮次的增量更新循环往复，既通过在线实时的实体链指保
30 证了知识图谱数据更新的实时性，又通过离线增量的实体链指确保数据无遗漏的准确性，

从而使得基于相应知识图谱的相关业务处理结果更准确有效。

附图说明

5 为了更清楚地说明本发明实施例的技术方案，下面将对实施例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它的附图。

- 图 1 示出根据本说明书的一个具体实施场景示意图；
- 图 2 示出根据本说明书的一个更新知识图谱的具体实施架构示意图；
- 10 图 3 示出根据本说明书一个实施例的针对初始知识图谱全量的实体链指的方法流程图；
- 图 4 示出根据本说明书一个实施例的更新知识图谱的方法流程图；
- 图 5 示出根据一个实施例的用于更新知识图谱的装置的示意性框图。

15 具体实施方式

下面结合附图，对本说明书提供的技术方案进行描述。

为了更清楚地理解本说明书的技术方案，首先结合一个具体实施场景描述本说明书中的技术方案提出的技术背景。

图 1 示出了本说明书的一个具体实施架构。该实施架构涉及一个基于知识图谱进行业务处理的场景。在图 1 示出的实施架构中，业务服务器可以为各个用户在相应终端上进行的相关业务（例如搜索业务、查询业务、收付款业务、导航业务等），提供相应业务支持。计算平台可以与业务服务器交互数据。其中，计算平台可以是与业务服务器连接的其他计算机、设备、服务器等，也可以是业务服务器的一部分，或者说设于业务服务器，在此不做限定。在一个具体例子中，计算平台可以是一个知识图谱服务平台，用于作为以知识图谱服务为核心能力的中台，面向各种业务提供知识管理、知识推理、知识服务的功能支持，以及

20 与这些功能相配套的图谱解决方案。

25 与这些功能相配套的图谱解决方案。

单个业务主体可以通过预先在业务服务器注册的账号进行相关业务。单个业务主体可以是进行预定业务的独立实体，如一个自然人、一个商户、一个企业等。账号例如通过唯一的用户标识（如手机号、银行卡号等）描述。实践中，可能会出现一个业务主体（账号

30 的实际使用者或控制者）注册一个或多个用户标识的情形。如图 1 中，作为业务主体的用

户 1 注册有账号 1、账号 2，用户 2 注册有账号 3、用户 3 注册有账号 4 等等。

假设相关业务基于知识图谱进行，知识图谱可以通过采集各个用户标识对应的业务数据而构建，初始构建的知识图谱中，单个用户标识可以作为一个业务主体对应单个节点。基于前述的一个业务主体注册有多个账号的情形，还可以基于各个节点的特征数据执行全量的实体链指操作，并将同一业务实体控制的不同用户标识的节点进行实体归一，从而更新相应知识图谱并保存在计算平台，以供业务服务器使用。

更进一步地，业务服务器可以从计算平台获取知识图谱中的相关数据进行业务处理。而业务处理过程中产生的业务数据，可以传递给计算平台。为了更好地为实时业务提供数据服务，知识图谱需要持续更新。于是，计算平台根据这些业务数据，可以对知识图谱执行实体链指操作，从而根据新的业务数据，修正知识图谱中的实体归一结果，进而更新知识图谱。

其中，实体链指从业务应用的角度，可以推理出知识图谱中任意两个节点对应的业务主体是否具有相同的特性。具有相同的特性通常标志着对应同一个业务主体。如两个用户是否属于同一个家庭、两个收钱码是否属于同一个店铺、两个账号是否属于同一个自然人，等等。其中，这里的同一个家庭、同一个店铺、同一个自然人各自代表着一个业务主体，两个用户、两个收钱码、两个账号在具有相同特性的情况下，可以对应着统一业务主体。实体链指的目标通常是实体归一，即基于实体链指的结果，进一步通过实体描述信息（如属性信息、连接关系信息等）的合并处理方式，处理“被识别为具有相同特性”的多个业务主体（节点），以得到唯一的业务主体（节点）。归一前“被识别为具有相同特性”的业务主体对应的多个节点上的描述信息（如连接关系、属性信息等），都会挂载到归一后的业务主体（即节点）上。

基于实体链指和实体归一操作，可以针对知识图谱进行知识的融合。常规技术中，针对知识图谱进行知识融合的更新通常为离线批量处理或者在线实时处理。离线批量更新例如按照预定周期（如一天）更新，存在着时效性差的问题，而在线实时处理可能因为网络问题、数据不全面问题等，存在着融合失败的可能性，如消息拥塞时，融合目标（需融合的某个节点）尚未记录进入知识图谱，就无法链指上融合目标，长期积累导致知识图谱的可用性降低，业务处理准确度下降。

有鉴于此，本说明书针对知识图谱的更新过程提出改进，得到可用性更高的知识图谱数据，使相应业务处理的准确度、有效性提高。如图 1 示出的实施场景中，对知识图谱执行实体链指、实体归一操作以通过更新的业务数据对知识图谱的部分进行改进。为此，本

说明书提供一种离线、在线结合的知识图谱更新方案。

图 2 示出了本说明书的技术架构。如图 2 所示，在本说明书的实施架构下，知识图谱融合过程可以包括三种实体链指过程，全量的实体链指、实时的实体链指和增量的实体链指。实体链指的目的是融合知识图谱中的知识。因此，在实体链指结果中存在至少 2 个节点对应的业务实体具有相同特性的情况下，可以确定具有相同特性的节点对应的业务实体是同一个业务实体，从而进行实体归一操作。否则，如果实体链指结果中不存在任何 2 个节点对应的业务实体具有相同特性，则不进行实体归一操作。也就是说，实体归一操作是基于实体链指的结果进行或不进行的，因此，图 2 仅标示出实体链指的示意，而未标注实体归一操作。为了描述方面，图 2 中将全量的实体链指、实时的实体链指和增量的实体链指分别称为全量链指、实时链指和增量链指。

其中，全量链指通常针对知识图谱中的全部数据进行，可以看作是当前知识图谱的初始化过程。全量数据通常数据量级较大，如 10 万亿条数据，因此，全量链指通常在使用知识图谱提供数据服务前一次性执行。但不排除在可选的实现方式中，全量链指按照预定的全量链指条件进行，比如每隔半年或一年进行一次全量链指操作。全量链指操作通常为离线执行的操作。

实时链指和增量链指均可以看作对增量数据的链指操作。通常，实时链指的数据量级较小，通常针对增加的单条业务数据进行，增量链指的数据量级远大于实时链指的数据量级，但小于全量链指的数据量，如针对 10 万条业务数据进行。其中，如图 2 所示，针对初始的知识图谱经过离线的全量链指操作后，可以将经由实体归一的知识图谱作为初始化的当前知识图谱作为线上数据库进行相关业务处理。在业务处理过程中，可能不断产生新的业务数据，例如，一个具体业务为张三向李四的转账业务，则张三和李四对应知识图谱中的节点属性或连接属性发生变化，如从无连接变为有连接。针对这样一条实时业务数据，可以实时监测张三、李四的特征变化，并将变化后的特征与其他节点进行比较，以挖掘变化后张三、李四分别对应的两个节点是否和其他节点的特征变得相似。该过程即为实时链指过程，根据以上示例可知，实时链指为在线过程，且根据实时链指结果可以进行实体归一操作或不进行实体归一操作。如图 2 所示，知识图谱可以在业务数据更新过程中不断基于实时链指结果更新。这种更新可以包括节点对应的实体描述信息的更新或者节点特征向量的更新等。

增量链指可以按照预定的增量更新条件进行，例如，每天定时（如 0 点）进行，或者按照业务数据的产生数量（例如每 10 万条数据）进行。增量更新条件每满足一次，可以

进行一轮次的增量更新。增量数据往往是多条实时业务数据的积累数据。增量链指操作完成后，可以替换当前轮次增量更新期间针对知识图谱基于实时链指的更新结果。例如，当前知识图谱记为 T ，针对各条业务数据的实时链指分别记为 $\delta_1, \delta_2, \dots, \delta_t$ 等，在第 t 次实时更新后的知识图谱记为 $T + \delta_1 + \delta_2 + \dots + \delta_t$ 。此时，进行增量链指，假设增量数据记为 t ，则增量链指结果可以记为 Δ_t ，利用增量链指结果更新的知识图谱例如记为 $T + \Delta_t$ 。此时，相当于用 Δ_t 替换 $\delta_1 + \delta_2 + \dots + \delta_t$ 。增量更新的知识图谱可以作为下一轮增量更新的初始知识图谱。增量链指可以是离线的实体链指过程。

如此，经过离线的全量链指结果对当前知识图谱的初始化，以及后续增量更新轮次内在线的实时链指更新和离线的增量链指更新，使得当前知识图谱兼顾实时性和数据准确性，从而保持其高可用性。

下面详细描述本说明书的技术构思。

首先需要说明的是，本说明书所涉及的知识图谱可以是任何业务场景下的知识图谱，例如：描述商户/企业之间的相互关系的商户图谱，知识图谱中的各个节点分别对应各个商户/企业，具有关联关系的两个商户/企业对应的两个节点之间通过连接边连接；描述消费偏好的知识图谱，各个节点可以对应商户、消费者、商品等，消费者消费过的商户，相应的两个节点间通过连接边连接，同样，消费者购买过的商品、商户经营的商品，相应节点之间均可以连接边表达其连接关系。

图 3 示出了根据本说明书一个实施例的针对知识图谱全量的实时链指流程。该流程的执行主体可以是具有一定计算能力的计算机、设备、服务器。更具体地，如可以是图 1 中的计算平台。图 3 示出的知识图谱全量的实体链指流程可以用于初始时针对全量业务数据的知识融合。该流程可以在知识图谱更新过程中终身仅执行一次。在一些可能的实施例中，也可以每经过一个较长的时间间隔，如半年、一年、五年等，执行一次。

如图 3 所示，该针对知识图谱全量的实体链指流程可以包括：步骤 301，针对利用全量业务数据构建的知识图谱中的各个节点分别获取其对应的实体描述信息，其中，该知识图谱包括全量业务数据中各个业务主体一一对应的各个节点，以及连接两两节点的连接边，用于描述业务主体之间的连接关系；步骤 302，根据各个节点各自对应的实体描述信息，提取各个节点分别对应的各个特征向量；步骤 303，基于各个特征向量检测两两节点之间的相似性；步骤 304，根据两两特征向量的相似性是否满足预定同质条件，识别相应的两两节点是否具有相同特性。

首先，在步骤 301，针对利用全量业务数据构建的知识图谱中的各个节点分别获取其

对应的实体描述信息。

5 这里的知识图谱可以是根据初始全量业务数据构建的知识图谱，例如，根据线下商户的收款账户等商户数据构建的知识图谱。初始的知识图谱可以包括各个业务主体一一对应的各个节点，以及连接两两节点的连接边，用于描述业务主体之间的连接关系。假设商户图谱中，单个收款账户作为一个业务主体在知识图谱中对应一个节点。两个收款账户之间具有关联关系，则相应的两个节点之间通过连接线连接。这里的关联关系例如可以包括但不限于转账、注册人身份信息（如姓名、电话号码）一致、相互关注、互为通讯录好友，等等。

10 其中，构建初始的知识图谱的业务数据可以根据线上抓取、线下统计等各种方式获取。初始的知识图谱可以根据全量的业务数据预先构建，也可以根据全量的业务数据在当前流程中构建，在此不做限定。

可以理解，节点对应的实体描述信息用于对节点对应的业务主体进行描述。实体描述信息可以包括业务主体自身的属性信息、业务主体与其他业务主体相关联的连接信息中的至少一项。属性信息可以是描述相应的单个业务主体（如单个收款账号）的各种属性的信息，如对应于商户的业务主体的属性信息可以包括以下中的至少一项：注册时间、注册地点、绑定的银行卡、交易设备、登录手机号，等等。与其他节点之间的连接关系描述出节点对应的实体之间的关联关系。

接着，在步骤 302，基于各个节点各自对应的实体描述信息，提取各个节点分别对应的各个的特征向量。

20 从节点的实体描述信息中提取特征向量的过程，是将实体描述信息数字化的过程。也就是说，用抽象的数据表示实体信息，从而便于计算机处理这些信息。基于单个节点对应的实体描述信息，可以提取相应的特征向量。在本说明书实施例中，节点的特征向量可以包括文本语义向量、基于位置（LBS，Location-Based Service）的轨迹向量、图结构向量、图表征向量等等中的至少一项，用于描述相应业务实体。

25 其中，文本语义向量可以通过文本描述相应业务主体的信息中提取的语义信息。例如，商户的经营范围等，语义向量可以是分词后得到的各个词汇分别对应的各个词向量的融合向量，例如各个词向量进行拼接或嵌入（embedding）等方式融合得到的向量。

LBS 向量可以表示基于位置的轨迹信息。具体而言，可以按照时间顺序采集相应业务主体的位置信息，从而构建其轨迹向量。例如，向前采样预定个数（如 5 个）的位置点，或者采样预定时间段（如采样时间前 24 小时）内的位置点，依次排列构成轨迹向量。作

为示例，一个商户经过依次的 5 个最新位置点为 L1、L7、L6、L5、L3，则可以对应位置向量 (L1, L7, L6, L5, L3)。位置点的采集方式和业务主体有关，在业务主体对应着具有通信功能的终端设备的情况下，可以通过相应终端设备采集相应位置点，在业务主体可以对应与电子设备无关的其他载体（如纸质二维码）的情况下，可以通过使用该载体的其他终端设备采集相应位置点，在此不再赘述。

图结构向量可以用于描述单个节点与其他节点之间的连接关系。例如，对于知识图谱中的单个节点，基于其在知识图谱中涉及的各个连通路径构建单个图结构向量、利用其在知识图谱的邻接矩阵中对应的一行或一列元素构成的向量作为图结构向量，等等。

图表征向量可以是经过图模型处理知识图谱得到的表征向量。这种情况下，单个节点的图表征向量可以融入自身特征和其邻居节点的特征，因此，既包含相应业务主体的属性信息，又包含相应业务主体与其他业务主体的连接信息。

在其他实施例中，基于节点对应的实体描述信息，还可以提取其他描述向量，在此不再一一例举。利用这些描述性向量中的一项或多项，可以从一个或多个维度描述相应业务主体。在单个业务主体的描述向量为 1 个的情况下，可以将相应的 1 个描述向量作为相应单个节点的特征向量。在单个业务主体的描述向量有多个的情况下，可以将多个描述向量的拼接向量或嵌入 (embedding) 向量，作为相应单个节点的特征向量。其中，嵌入向量可以通过神经网络处理得到，或者对各个描述向量加权、求平均等得到，在此不做限定。

这样，可以得到各个节点的特征向量。特征向量描述了节点所对应业务主体的各种信息，为了检测两两业务主体是否具有相同特性，可以经由步骤 303，基于两两特征向量检测两两节点之间的相似性。

在一个实施例中，可以通过向量的匹配度衡量两个向量的相似性。匹配度例如可以按照匹配一致的元素数量和元素总数量确定。例如，在两个特征向量的维度一致的情况下，可以基于匹配一致的元素数量与向量维数的比值确定两个特征向量的匹配度。如一个具体例子中，两个特征向量的维度均为 10 维，其中有 8 个元素匹配一致，则可以确定其匹配度为 80%。在两个特征向量不一致的情况下，可以基于匹配一致的元素数量与预先约定的较大或较小向量维数的比值确定两个特征向量的匹配度。例如，两个特征向量的维度分别为 10 维、8 维，其中有 8 个元素匹配一致，以较小向量维度相比，则可以确定其匹配度为 100%。

在另一个实施例中，可以通过向量的相似度衡量两个向量的相似性。向量的相似度例如通常可以通过诸如杰卡德 (Jaccard) 系数、余弦相似度、皮尔逊相似度、欧几里得距离、

KL 散度 (Kullback–Leibler divergence, 相对熵) 之类的参数进行衡量。两个向量的相似度可以与杰卡德 (Jaccard) 系数、余弦相似度、皮尔逊相似度等中的一项正相关, 或与欧几里得距离、KL 散度等中的一项负相关。

其中, 以 Jaccard 系数为例, 两个向量 A、B 的相似度例如可以描述为:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad \text{其中, } |A \cap B| \text{ 表示两个向量 A 和 B 中相同元素的个数, } |A \cup B| \text{ 表示两个向量 A 和 B 中合并相同元素后总的元素的个数。}$$

值得说明的是, Jaccard 系数的计算方式不要求两个向量 A、B 的维数必然相等, 因此具有更强的普适性。而余弦相似度、皮尔逊相似度、欧几里得距离、KL 散度之类的方法通常更适用于相同元素的集合 (如相同维数的向量) 之间的相似性衡量。

步骤 304, 根据两两特征向量的相似性是否满足预定同质条件, 识别相应的两两节点是否具有相同特性。

可以理解, 检测两两节点之间相似性的目的是为了进行实体链指, 即判断两个节点之间是否具有相同特性 (对应同一个业务主体)。判断条件可以预先设定, 这里记为预定同质条件。根据向量相似性的衡量方式不同, 预定同质条件可以为, 向量匹配度超过预定匹配度阈值, 或者, 向量相似度超过预定相似度阈值, 等等。

值得说明的是, 针对单个特征向量与两个以上的特征向量满足预定同质条件的情况下, 另外的两个以上的特征向量两两之间不一定均满足预定同质条件。此时, 可以在两个特征向量的相似性满足预定同质条件的情况下, 则认为相应的两个节点对应的业务主体为同一个。如此, 单个特征向量与两个以上的特征向量满足预定同质条件的情况下, 可以确定这些节点均具有相同特性, 对应同一业务主体。作为示例, 假设节点 a 对应的特征向量 Ia 与节点 b 对应的特征向量 Ib 满足预定条件, 节点 b 对应的特征向量 Ib 与节点 c 对应的特征向量 Ic 满足预定同质条件, 由于可以得到节点 a 与节点 b 对应同一个业务主体、节点 b 与节点 c 对应同一业务主体的识别结果, 因此, 不论节点 a 对应的特征向量 Ia 与节点 c 对应的特征向量 Ic 是否满足预定同质条件, 都可以确定节点 a、b、c 均对应同一业务主体, 如对应同一商户、同一消费者等。

进一步地, 对于初始构建的知识图谱中对应同一业务主体的各个节点可以进行实体归一。即合并为一个节点, 并将相应实体描述信息 (如属性信息、连接信息等信息) 进行融合。例如上面的例子中, 节点 a、b、c 合并为节点 a', 同时, 节点 a、b、c 的属性信息和连接信息均归属于节点 a'。比如, 节点 a 与节点 e、d 连接, 节点 b 与节点 d、h 连接, 节点 c 与节点 g 连接, 则合并后得到的节点 a' 与节点 e、d、h、g 均具有连接关系。

5 在一个可选的实施例中，对应同一业务主体的各个节点的属性信息、连接信息等实体描述信息归一化过程中，还可以通过特征向量的融合实现。例如，通过对应同一业务主体的各个节点的特征向量的平均、加和、取中位数、嵌入（embedding）等之一的方式对相应多个节点（如节点 a、b、c）的各个特征向量进行融合，融合后的特征向量作为描述归一后的节点对应的业务实体信息的特征向量。

如此，可以将初始构建的知识图谱中各组对应到同一业务主体的节点分别合并归一，形成初始的全量知识图谱。

10 初始的全量融合知识图谱可以作为初始增量更新轮次的初始知识图谱提供线上业务的图谱服务，并循环更新。如前文所述，循环更新由如图 2 所示的离线增量更新循环和在线实时更新循环配合进行。图 4 示出了使用知识图谱提供线上业务的图谱服务过程中，更新知识图谱的流程。该流程的执行主体是可与业务服务器实时交换数据的任意具有计算能力的计算机、设备、服务器，如图 1 中的计算平台。更进一步地，其可以与图 3 所示流程的执行主体一致，也可以不一致。可以理解，知识图谱上线后，其实体链指过程可以按增量更新轮次进行。为了描述方便，图 4 示出的实施流程以其中一个增量更新轮次为例进行描述。

20 如图 4 所示，本说明书一个实施例提供的更新知识图谱的流程中，针对一轮增量更新可以包括：步骤 401，获取该轮增量更新的初始知识图谱；步骤 402，进行更新步骤，包括重复执行的实时更新操作和满足预设的增量更新条件的情况下的增量更新操作，其中，该实时更新操作包括：响应于接收到新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新，该增量更新操作包括：利用该轮增量更新期间产生的业务数据对该轮增量更新的初始知识图谱进行更新，以作为下一轮增量更新的初始知识图谱。

首先，经过步骤 401，获取该轮增量更新的初始知识图谱。

25 当前轮次增量更新的初始知识图谱是当前增量更新轮次初始的知识图谱。该初始知识图谱可以是基于对初始时通过全量业务数据构建的知识图谱的全量链指结果确定的。具体地，在首轮增量更新期间，该初始知识图谱可以是利用图 3 示出的实体链指流程进行全量数据的实体链指更新的知识图谱，在非首轮增量更新期间，该初始知识图谱可以是在利用图 3 示出的实体链指流程进行全量链指更新的知识图谱基础上，经过若干轮次增量更新后得到的知识图谱。或者说，是前一轮次增量更新后得到的知识图谱。

30 该初始知识图谱可以用于为当前业务提供知识图谱的数据支持。例如，在当前业务处

理过程中，可以从当前知识图谱中获取业务主体的属性数据、关联关系数据中的至少一项。当前业务可以是与当前知识图谱相关的各种业务。例如，在当前知识图谱为商户图谱的情况下，各个节点分别对应各个收款账号，当前业务可以为权益激励业务，单个商户在 24 小时内完成 50 笔收款则即时给予预定积分、红包或现金等的奖励。如此，当前业务可以在商户发生收款业务的情况下，从知识图谱获取收款次数相关的属性数据等。

接着，在步骤 402，进行更新步骤。

根据本说明书的技术构思，该更新步骤是基于前述的初始知识图谱进行更新的步骤。该更新步骤可以包括重复执行的实时更新操作和满足预设的增量更新条件的情况下的增量更新操作。

- 10 可以理解，当前业务进行过程中，还可以产生新的业务数据。例如，在利用商户图谱进行权益激励业务的情况下，在一次收款业务中，针对收款方，可以产生收款金额、支付方、支付时间、收款地点等业务数据。新的业务数据可能对知识图谱中节点的属性信息等产生影响。例如，收款次数增加、收款轨迹改变、关联关系改变等。甚至还可能增加节点数量（例如出现新的注册账户）。为了满足业务的实时性需求，可以针对新产生的业务数据
- 15 据进行实时的实体链指操作。

- 可以理解，实时的实体链指操作是在业务处理过程中针对实时业务数据进行的，其是对知识图谱局部进行的实体链指。更具体地，针对当前业务数据所涉及的节点进行。例如，当前业务包括第一业务，针对第一业务产生的第一业务数据所涉及的第一节点，按照第一业务数据修改第一节点相应的实体描述信息。然后，针对第一节点基于修改后的实体描述
- 20 信息提取其对应的特征向量，如记为第一特征向量。接着将该第一特征向量与其他各个节点分别对应的各个其他特征向量进行相似性比较，从而确定是否有与经过信息更新后的第一节点具有相同特性的其他节点，以完成实时的实体链指。

- 进一步地，基于实时产生的新的业务数据，在所涉及的节点被识别为与其他若干节点具有相同特性的情况下，这些节点可能对应同一业务主体。则还可以将对应同一业务主体的各个节点合并归一（执行实体归一）。例如，检测到第一节点与第二节点、第三节点均具有相同特性，则可以认为他们均对应同一业务主体，可以将第一节点、第二节点、第三节点归并为一个节点（如第一节点），三者的实体描述信息进行合并作为归并后的节点（如第一节点）对应的实体描述信息。另一方面，在所涉及的节点被识别为与其他若干节点均不具有相同特性的情况下，记录该实时的实体链指结果，以及针对第一节点融合第一业务
- 25 数据后的实体描述信息，而无需实体归一操作。
- 30

如此，可以对当前知识图谱进行实时更新，并使用更新后的知识图谱进行后续业务处理。并且，在不断产生新的业务数据的情况下，实时的实体链指结果可以叠加。其中，知识图谱的实时的实体链指操作可以通过诸如 ha3、Probase、知心、知立方之类的基于知识图谱的在线搜索引擎进行。在一次搜索过程中，在线搜索引擎可以将知识图谱中的知识联
5 系起来，反馈给用户更精准的检索结果，并且可以收集业务处理结果，例如用户是否选择所反馈的信息等。另外，实体归一例如可以通过 geabase、gstore 之类的在线图存储引擎完成，例如将具有相同特性的各个节点的节点标识修改一致，并将各个节点对应的实体描述信息都与修改后的节点标识对应存储。

另一方面，实时产生的业务数据未必能完全及时地通过实时的实体链指操作进行更新。
10 例如，在一次业务过程中，涉及的两个业务主体，如为账号 A 和账号 B，业务内容为账号 A 向账号 B 进行了转账业务，这两个业务主体仅有一个业务主体（如账号 B）在当前知识图谱中对应相应节点（如节点 b），而另一个节点在当前知识图谱中未对应相应节点。此时，对于未对应相应节点业务主体，其数据不能实时添加在到当前知识图谱，因此仅通过实时的实体链指可能错过相关数据。

15 为此，还可以将当前业务产生的业务数据作为增量数据记载到当前增量数据集。这里的当前增量数据集可以是用于记录当前轮次增量更新中的增量数据的数据集。该增量数据集可以是具有预定标识的数据集，例如具有和当前增量更新周期对应的标识（如 t），也可以是按照预先确定的增量存储位置存储，在此不做限定。

增量更新条件可以是对知识图谱进行增量更新的触发条件，其可以根据具体业务预先
20 设定。在一个实施例，增量更新条件可以为经过预定时间间隔或预定周期到达，例如，预定时间间隔为 24 小时，则每满 24 小时，增量更新条件满足。在另一个实施例，增量更新条件为累积业务数据条数达到预定条数，如 10 万条，则增量数据集中每增加 10 万条增量数据，增量更新条件满足。

在增量更新条件满足的情况下，可以利用增量数据进行增量的实体链指。增量的实体
25 链指的方式与实时的实体链指类似，区别在于，增量的实体链指针对多条业务数据进行，涉及更多的节点，且可以以离线的方式进行。如，增量的实体链指过程中可以获取增量数据集中的离线数据进行操作，该过程与当前的线上业务分离。

具体地，在增量的实体链指过程中，可以针对各条增量数据相关的若干节点进行。例如，可以将增量数据中包含的业务主体的描述信息改变数据等补充到相应节点（如 100 个
30 节点），并重新提取这些节点的特征向量。然后针对这些节点中的单个节点，将重新提取

的特征向量与其他节点的特征向量比较相似性，从而将相似性满足相似条件的节点确定为具有相同特性，可能对应同一业务主体。

为了确保知识图谱更新的一致性，利用增量的实体链指结果，可以在当前轮次的初始知识图谱上进行数据更新，更新后的知识图谱作为下一轮增量更新的初始知识图谱。

- 5 具体地，可以用增量的实体链指结果替换该轮次增量更新期间的实时的实体链指结果。从而，在利用增量的实体链指结果中存在具有相同特性的两两业务实体的情况下，以增量的实体链指结果进行实体归一形成新的知识图谱。增量的实体链指结果替换该轮次增量更新期间的实时链指结果可以通过数据转存（如 dump）机制进行。具体地，将增量的实体链指结果同步至在线检索引擎（如 ha3）及在线图存储引擎（如 geabase），从而完成增量
- 10 的实体链指结果对当前轮次增量期内产生的各个实时的实体链指结果的替换。

值得说明的是，增量的实体链指结果中，可能存在至少两个节点具有相同特性，则可以根据增量的实体链指结果进行实体归一操作。在可选的实施例，一轮增量更新期间产生的业务数据的增量链指结果也可能是不存在任何两个节点具有相同特性，此时，则不需要进行合并节点的实体归一操作。

- 15 可以理解，增量的实体链指往往需要处理远超过单次实时的实体链指的业务数据，因此，由于增量的实体链指的数据量较大，增量的实体链指的耗时也常常远大于实时的实体链指耗时，例如为 30 分钟、1 小时。在知识图谱的线上服务期间，该耗时不可忽略。换句话说，在增量的实体链指过程中，业务处理仍在进行，新的业务数据仍可能产生，实时的实体链指可能持续进行。

- 20 因此，为了确保知识图谱数据的实时性，根据一个可能的的设计，在更新初始知识图谱之后，还可以在当前的初始知识图谱上累加增量更新条件满足之后产生的若干条实时的实体链指结果。例如，当前轮次增量更新针对的增量数据为 γ_1 至 γ_T ，则本次增量的实体链指针对增量数据 γ_1 至 γ_T 进行。增量的实体链指结果如记为 Δ_T ，当前知识图谱 T 基于增量的实体链指结果 Δ_T 更新后为 $T + \Delta_T$ 。在本次增量的实体链指过程中，又产生了实时业务
- 25 数据 γ_{T+1} 至 γ_{T+s} ，当前知识图谱可能还继续经由实时链指进行实时更新，例如经过 s 次实时链指 δ_{t+1} 、 δ_{t+2} ... δ_{t+s} 等。则为了适应后续业务，当前知识图谱从逻辑上，还应该具有 s 次实时链指的结果。实时链指 δ_{t+1} 、 δ_{t+2} ... δ_{t+s} 等相当于在当前次的增量链指后进行的实时链指。则在更新后的知识图谱上，还可以在当前的知识图谱 $T + \Delta_t$ 上，增加 s 次实时链指结果，得到知识图谱 $T + \Delta_t + \delta_{t+1} + \delta_{t+2}$... δ_{t+s} 以进行后续业务处理。也就是说，
- 30 基于增量的实体链指结果更新后的知识图谱 $T + \Delta_t$ ，可以作为下一轮次增量更新的初始知

识图谱，而为了确保业务处理的正常进行，在该初始知识图谱上增加上述的 s 次实时链指结果。而实时业务数据 γ_{T+1} 至 γ_{T+s} ，可以作为下一增量更新周期的增量数据。在下一轮增量更新期间，假设增量链指结果为 Δ_{2t} ，可以用于替换知识图谱 $T+\Delta_t$ 之后的所有实时链指数据，得到知识图谱 $T+\Delta_t+\Delta_{2t}$ ，作为再下一周期的初始知识图谱。

- 5 仅就当前轮次增量更新而言，假设存在前一轮增量更新期间 $T-1$ ，则在步骤 401 中，获取该轮增量更新的初始知识图谱之后，步骤 402 的更新步骤中还可以包含叠加前一增量更新周期 $T-1$ 的增量更新条件满足之后产生实时业务数据（如 γ_1 至 γ_m ， m 小于 t ）的实时的实体链指结果（如 δ_1 至 δ_m ）的操作。

在可选的实现方式中，实时业务数据、实时的实体链指结果可以按照预定顺序添加标识的方式按标识存储，以识别增量更新条件满足前后的业务数据、实时的实体链指结果数据等。例如，使用业务产生的时间戳、序列号等作为版本标识。

如此循环更新的知识图谱，结合在线的实时性与离线的准确性，可以得到更高可用性的知识图谱，为相应业务提供支持，以得到更有效的业务结果。例如，更有效地为用户推荐商户、商品，更有效地识别一个自然人、一个商户、一个企业的不同账号，等等。

- 15 回顾以上过程，在为当前业务提供基于知识图谱的数据支持过程中，采用在线、离线相结合的方式更新知识图谱。首先，利用全量业务数据离线构建知识图谱，并进行全量的实体链指、实体归一，以初始化知识图谱。之后，设置增量更新条件，对知识图谱进行各个轮次的循环更新。一方面，基于实时产生的业务数据进行实时链指，提供在线的知识图谱更新，另一方面，按照预设的增量更新条件，在增量更新条件满足时，按照当前轮次增量更新期间内新增的业务数据进行增量的实体链指，从而提供离线的知识图谱更新。然后，将离线增量实体链指结果与在线的实时实体链指结果相融合来更新当前知识图谱。如此，各个增量更新轮次循环往复，即通过在线实时实体链指保证了知识图谱数据更新的实时性，又通过离线增量实体链指确保数据无遗漏的准确性，从而提高知识图谱的数据可用性，使得相关业务处理结果更准确有效。

- 25 根据另一方面的实施例，还提供一种用于更新知识图谱的装置。图 5 示出了根据一个实施例的用于更新知识图谱的装置 500。如图 5 所示，装置 500 可以包括：

获取单元 501，配置为在各轮增量更新中获取初始知识图谱；

- 更新单元 502，配置为在各轮增量更新中进行包括重复执行的实时更新操作和满足预设的增量更新条件的情况下的增量更新操作的更新步骤，其中，该实时更新操作包括：响
30 应于接收到新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱

进行更新，该增量更新操作包括：利用该轮增量更新期间产生的业务数据对初始知识图谱进行更新，以作为下一轮增量更新的初始知识图谱。

其中：在该轮增量更新是首轮增量更新的情况下，该轮增量更新的初始知识图谱基于对利用全量业务数据构建的知识图谱的实体链指结果进行实体归一得到；在该轮增量更新不是首轮增量更新周期的情况下，该轮增量更新的初始知识图谱基于对前一轮增量更新中的初始知识图谱的增量的实体链指结果进行实体归一得到。

在一个实施例中，实时更新操作、增量更新操作均包含以下实体链指过程：确定是否存在至少 2 个节点对应的业务主体具有相同特性；

在存在的情况下，针对实体链指结果还执行以下实体归一过程：将具有相同特性的节点合并为一个节点，并且具有相同特性的各个节点相应的实体描述信息叠加后作为合并后的节点的实体描述信息。

在一个实施例中，装置 500 还可以包括初始化单元（未示出），配置为通过以下方式确定利用全量业务数据构建的知识图谱全量的实体链指结果：

针对利用全量业务数据构建的知识图谱中的各个节点分别获取其对应的实体描述信息；

根据各个节点各自对应的实体描述信息提取各个节点分别对应的各个特征向量；

基于两两特征向量检测两两节点之间的相似性；

根据两两特征向量的相似性是否满足预定同质条件，识别相应的两两节点是否具有相同特性。

在一个可选的实现方式中，初始知识图谱包括第一节点，针对第一节点的第一业务数据为当前接收的新的业务数据，响应于当前业务中产生新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新包括：

利用第一业务信息更新第一节点的第一实体描述信息；

从更新后的第一实体描述信息中提取第一特征向量；

比较第一特征向量与其他各个节点的各个其他特征向量一一对应的各个相似性；

基于各个相似性是否满足预定同质条件，得到是否存在与第一节点具有相同特性的其他节点实时的实体链指结果；

基于该实时的实体链指结果对前一实时更新操作中更新后的知识图谱进行更新。

根据一个可能的设计，更新单元 502 还配置为：

将当前接收的新的业务数据作为增量数据添加至当前增量数据集；

利用该轮增量更新期间产生的业务数据对初始知识图谱进行更新包括:

利用当前增量数据集中的各条增量数据进行针对该轮增量更新的初始知识图谱增量的实体链指;

利用增量的实体链指结果更新初始知识图谱。

- 5 其中, 增量更新条件包括以下中的一项: 预定周期到达、该轮增量更新期间产生的业务数据条数达到预定条数。

在一个实施例中, 在该轮增量更新不是首轮增量更新的情况下, 更新单元 502 进一步配置为:

- 10 获取基于前一轮增量更新中满足预设的增量更新条件之后的实时更新操作中得到的各个实时的更新结果;

根据各个实时的更新结果更新该轮增量更新的初始知识图谱。

其中, 实体描述信息可以包括属性信息、连接信息中的至少一项。

特征向量可以包括以下中的一项, 或以下中的多项经嵌入得到的向量: 文本语义向量、轨迹向量、图结构向量、图表征向量。

- 15 在一个实施例中, 实时的实体链指过程通过在线检索引擎完成, 基于实时的实体链指更新当前知识图谱通过在线图存储引擎完成; 更新单元 502 配置为通过以下方式利用增量的实体链指结果更新初始知识图谱:

- 20 通过数据转存机制, 将增量的实体链指结果同步至在线检索引擎及在线图存储引擎, 从而完成增量的实体链指结果对该轮增量更新期间内产生的各个实时的实体链指结果的替换, 从而利用增量的实体链指结果更新初始知识图谱。

其中, 在增量数据中涉及的第二业务主体在该轮增量更新的初始知识图谱中不存在相对应的节点的情况下, 增量更新操作还包括:

在该轮增量更新的初始知识图谱中增加与第二业务主体相对应的第二节点;

基于增加第二节点后的知识图谱进行增量的实体链指。

- 25 在一个实施例中, 在该轮增量更新为首轮增量更新的情况下, 该轮增量更新的首次实时更新操作为:

利用接收的业务数据对该轮增量更新的初始知识图谱进行更新。

值得说明的是, 图 5 所示的装置 500 与图 4 描述的方法相对应, 图 4 的方法实施例中的相应描述同样适用于装置 500, 在此不再赘述。

- 30 根据另一方面的实施例, 还提供一种计算机可读存储介质, 其上存储有计算机程序,

当所述计算机程序在计算机中执行时，令计算机执行结合图 3 或图 4 等所描述的方法。

根据再一方面的实施例，还提供一种计算设备，包括存储器和处理器，所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现结合图 3 或图 4 等所描述的方法。

- 5 本领域技术人员应该可以意识到，在上述一个或多个示例中，本说明书实施例所描述的功能可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。

- 10 以上所描述的具体实施方式，对本说明书的技术构思的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所描述仅为本说明书的技术构思的具体实施方式而已，并不用于限定本说明书的技术构思的保护范围，凡在本说明书实施例的技术方案的基础之上，所做的任何修改、等同替换、改进等，均应包括在本说明书的技术构思的保护范围之内。

权 利 要 求 书

1.一种更新知识图谱的方法，所述方法包括对知识图谱进行多轮增量更新，其中，一轮增量更新包括：

获取该轮增量更新的初始知识图谱；

5 进行更新步骤，包括重复执行的实时更新操作和满足预设的增量更新条件的情况下的增量更新操作，其中，该实时更新操作包括：响应于接收到新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新，该增量更新操作包括：利用该轮增量更新期间产生的业务数据对所述初始知识图谱进行更新，以作为下一轮增量更新的初始知识图谱。

10 2. 如权利要求 1 所述的方法，其中，所述实时更新操作、所述增量更新操作均包含以下实体链指过程：确定是否存在至少 2 个节点对应的业务主体具有相同特性；

在存在的情况下，针对实体链指结果还执行以下实体归一过程：将具有相同特性的节点合并为一个节点，并且具有相同特性的各个节点相应的实体描述信息叠加后作为合并后的节点的实体描述信息。

15 3.如权利要求 1 所述的方法，其中：

在该轮增量更新是首轮增量更新的情况下，该轮增量更新的初始知识图谱基于对利用全量业务数据构建的知识图谱的实体链指结果进行实体归一得到；

在该轮增量更新不是首轮增量更新的情况下，该轮增量更新的初始知识图谱基于对前一轮增量更新中的初始知识图谱的增量的实体链指结果进行实体归一得到。

20 4. 如权利要求 3 所述的方法，其中，所述对利用全量业务数据构建的知识图谱全量的实体链指结果通过以下方式获取：

针对利用全量业务数据构建的知识图谱中的各个节点分别获取其对应的实体描述信息；

根据各个节点各自对应的实体描述信息提取各个节点分别对应的各个特征向量；

25 基于两两特征向量检测两两节点之间的相似性；

根据两两特征向量的相似性是否满足预定同质条件，识别相应的两两节点是否具有相同特性。

5. 如权利要求 2 所述的方法，其中，所述初始知识图谱包括第一节点，针对所述第一节点的第一业务数据为当前接收的新的业务数据，所述响应于当前业务中产生新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新包括：

30 据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新包括：

- 利用所述第一业务信息更新所述第一节点的第一实体描述信息；
从更新后的第一实体描述信息中提取第一特征向量；
比较所述第一特征向量与其他各个节点的各个其他特征向量一一对应的各个相似性；
基于各个相似性是否满足预定同质条件，得到是否存在与所述第一节点具有相同特性
- 5 的其他节点实时的实体链指结果；
基于该实时的实体链指结果对前一实时更新操作中更新后的知识图谱进行更新。
6. 如权利要求 2 所述的方法，其中，所述方法还包括：
将当前接收的新的业务数据作为增量数据添加至当前增量数据集；
所述利用该轮增量更新期间产生的业务数据对所述初始知识图谱进行更新包括：
- 10 利用当前增量数据集中的各条增量数据进行针对该轮增量更新的初始知识图谱增量的实体链指；
利用增量的实体链指结果更新所述初始知识图谱。
7. 如权利要求 1 所述的方法，其中，所述增量更新条件包括：预定周期到达，或者该轮增量更新期间产生的业务数据条数达到预定条数。
- 15 8. 如权利要求 1 所述的方法，其中，在该轮增量更新不是首轮增量更新的情况下，所述更新步骤还包括：
获取基于前一轮增量更新中满足预设的增量更新条件之后的实时更新操作中得到的各个实时的更新结果；
根据各个实时的更新结果更新该轮增量更新的初始知识图谱。
- 20 9. 如权利要求 2-5 任一所述的方法，其中，所述实体描述信息包括属性信息、连接信息中的至少一项。
10. 如权利要求 2-5 任一所述的方法，其中，所述特征向量包括以下中的一项，或以下中的多项经嵌入得到的向量：文本语义向量、轨迹向量、图结构向量、图表征向量。
11. 如权利要求 6 所述的方法，其中，实时的实体链指过程通过在线检索引擎完成，
- 25 基于实时的实体链指更新当前知识图谱通过在线图存储引擎完成；所述利用增量的实体链指结果更新所述初始知识图谱包括：
通过数据转存机制，将所述增量的实体链指结果同步至在线检索引擎及在线图存储引擎，从而完成所述增量的实体链指结果对该轮增量更新期间内产生的各个实时的实体链指结果的替换，从而利用增量的实体链指结果更新所述初始知识图谱。
- 30 12. 如权利要求 2 所述的方法，其中，在增量数据中涉及的第二业务主体在该轮增量更

新的初始知识图谱中不存在相对应的节点的情况下，所述增量更新操作还包括：

在该轮增量更新的初始知识图谱中增加与所述第二业务主体相对应的第二节点；

基于增加所述第二节点后的知识图谱进行增量的实体链指。

13. 如权利要求 1 所述的方法，其中，在该轮增量更新为首轮增量更新的情况下，该

5 轮增量更新的首次实时更新操作为：

利用接收的业务数据对该轮增量更新的初始知识图谱进行更新。

14. 一种更新知识图谱的装置，所述装置包括：

获取单元，配置为在各轮增量更新中获取初始知识图谱；

更新单元，配置为在各轮增量更新中进行包括重复执行的实时更新操作和满足预设的

10 增量更新条件的情况下的增量更新操作的更新步骤，其中，该实时更新操作包括：响应于接收到新的业务数据，利用接收的业务数据对前一实时更新操作中更新后的知识图谱进行更新，该增量更新操作包括：利用该轮增量更新期间产生的业务数据对所述初始知识图谱进行更新，以作为下一轮增量更新的初始知识图谱。

15. 一种计算机可读存储介质，其上存储有计算机程序，当所述计算机程序在计算机中

15 执行时，令计算机执行权利要求 1-13 中任一项的所述的方法。

16. 一种计算设备，包括存储器和处理器，其特征在于，所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现权利要求 1-13 中任一项所述的方法。