



(12)发明专利申请

(10)申请公布号 CN 108763556 A

(43)申请公布日 2018. 11. 06

(21)申请号 201810555333.9

(22)申请日 2018.06.01

(71)申请人 北京奇虎科技有限公司

地址 100088 北京市西城区新街口外大街
28号D座112室(德胜园区)

(72)发明人 张心铭 徐雅静 高升 李晓明
耿硕钦 王斌锋

(74)专利代理机构 北京市浩天知识产权代理事
务所(普通合伙) 11276

代理人 宋菲 刘云贵

(51)Int.Cl.

G06F 17/30(2006.01)

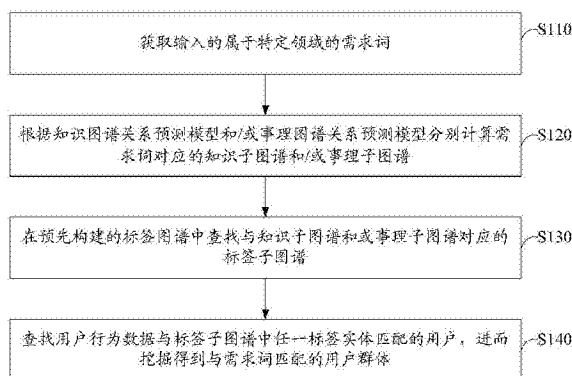
权利要求书2页 说明书14页 附图4页

(54)发明名称

基于需求词的用户挖掘方法及装置

(57)摘要

本发明公开了一种基于需求词的用户挖掘方法及装置。其中,方法包括:获取输入的属于特定领域的需求词;根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词对应的知识子图谱和/或事理子图谱;在预先构建的标签图谱中查找与知识子图谱和/或事理子图谱对应的标签子图谱;查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。本方案通过知识图谱和/或事理图谱,以及标签图谱可快速有效地挖掘得到与需求词匹配的用户群体,避免了现有技术中需人工挖掘与需求词对应的用户群体而造成的效率低下,人工成本高的弊端;并且,本方案挖掘得到的用户群体与需求词具有较高的匹配度。



1. 一种基于需求词的用户挖掘方法,所述方法基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行,所述方法包括:

获取输入的属于特定领域的需求词;

根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱;

在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱;

查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

2. 根据权利要求1所述的方法,其中,在所述获取输入的属于特定领域的需求词之后,所述方法还包括:在底层语料库查找与所述需求词对应的至少一个泛化词;

所述根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱具体为:根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词和所述至少一个泛化词对应的知识子图谱和/或事理子图谱。

3. 根据权利要求2所述的方法,其中,所述在底层语料库查找与所述需求词对应的至少一个泛化词进一步包括:

对所述需求词进行解析,生成所述需求词向量;

根据所述需求词向量与所述底层语料库中各个语料的词向量之间的距离,查找与所述需求词对应的至少一个泛化词。

4. 根据权利要求2所述的方法,其中,所述根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词和所述至少一个泛化词对应的知识子图谱和/或事理子图谱进一步包括:

利用距离计算方法,查找知识图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个知识实体;根据知识图谱关系预测模型对所述至少一个知识实体进行实体推荐,得到所述知识子图谱;

和/或,利用距离计算方法,查找事理图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个事理实体;根据事理图谱关系预测模型对所述至少一个事理实体进行实体推荐,得到所述事理子图谱。

5. 根据权利要求1-4中任一项所述的方法,其中,所述方法还包括预先按照以下步骤构建知识图谱:

构建特定领域的初始知识图谱,将所述初始知识图谱作为训练集,构建知识图谱的实体抽取模型与实体关系分类模型;

爬取特定领域知识类的结构化数据和非结构化文本数据;

根据知识图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构文本数据中抽取知识图谱三元组,对初始知识图谱进行扩充,获得构建后的知识图谱。

6. 根据权利要求1-4中任一项所述的方法,其中,所述方法还包括预先按照以下步骤构建事理图谱:

构建特定领域的初始事理图谱,将所述初始事理图谱作为训练集,构建事理图谱的实

体抽取模型与实体关系分类模型；

爬取特定领域事理类的结构化数据和非结构化文本数据；

根据事理图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构文本数据中抽取事理图谱三元组,对初始事理图谱进行扩充,获得构建后的事理图谱。

7. 根据权利要求1-6中任一项所述的方法,其中,所述方法还包括预先按照以下步骤构建标签图谱:

获取用户行为数据;

提取所述用户行为数据中的用户标签数据;

根据用户与用户标签数据之间的关联关系,建立标签图谱,其中所述标签图谱中标签实体之间的关系是根据用户与用户标签数据之间的关联关系得到的。

8. 一种基于需求词的用户挖掘装置,所述装置基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行,所述装置包括:

获取模块,适于获取输入的属于特定领域的需求词;

计算模块,适于根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱;

查找模块,适于在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱;

挖掘模块,适于查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

9. 一种计算设备,包括:处理器、存储器、通信接口和通信总线,所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信;

所述存储器用于存放至少一可执行指令,所述可执行指令使所述处理器执行如权利要求1-7中任一项所述的基于需求词的用户挖掘方法对应的操作。

10. 一种计算机存储介质,所述存储介质中存储有至少一可执行指令,所述可执行指令使处理器执行如权利要求1-7中任一项所述的基于需求词的用户挖掘方法对应的操作。

基于需求词的用户挖掘方法及装置

技术领域

[0001] 本发明涉及数据处理技术领域,具体涉及一种基于需求词的用户挖掘方法及装置。

背景技术

[0002] 随着科学及社会的不断发展,各类产品及服务的数量呈快速增长。为了提升产品及服务的推广效果,提高供应商收益,通常需根据产品或服务的特性挖掘潜在用户。

[0003] 目前,在挖掘潜在用户过程中,通常是采用人工挖掘的方式,通过与产品或服务有关的专业人员的知识经验,确定出某类或某几类用户为潜在用户,进而确定该类用户或该几类用户为产品或服务对应的用户群体。

[0004] 然而,采用该方法挖掘用户的效率低下,人工成本高,不利于大规模应用;并且,鉴于挖掘人员知识域的限制,导致最终挖掘的用户群体往往局限于某一类用户,无法真实全面地反映出该特定需求所对应的潜在用户。

发明内容

[0005] 鉴于上述问题,提出了本发明以便提供一种克服上述问题或者至少部分地解决上述问题的基于需求词的用户挖掘方法及装置。

[0006] 根据本发明的一个方面,提供了一种基于需求词的用户挖掘方法,该方法基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行,该方法包括:

[0007] 获取输入的属于特定领域的需求词;

[0008] 根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱;

[0009] 在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱;

[0010] 查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

[0011] 根据本发明的另一方面,提供了一种基于需求词的用户挖掘装置,该装置基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行,该装置包括:

[0012] 获取模块,适于获取输入的属于特定领域的需求词;

[0013] 计算模块,适于根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱;

[0014] 查找模块,适于在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱;

[0015] 挖掘模块,适于查找用户行为数据与所述标签子图谱中任一标签实体匹配的用

户,进而挖掘得到与需求词匹配的用户群体。

[0016] 根据本发明的又一方面,提供了一种计算设备,包括:处理器、存储器、通信接口和通信总线,所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信;

[0017] 所述存储器用于存放至少一可执行指令,所述可执行指令使所述处理器执行上述基于需求词的用户挖掘方法对应的操作。

[0018] 根据本发明的再一方面,提供了一种计算机存储介质,所述存储介质中存储有至少一可执行指令,所述可执行指令使处理器执行上述基于需求词的用户挖掘方法对应的操作。

[0019] 根据本发明提供的基于需求词的用户挖掘方法及装置,首先获取输入的属于特定领域的需求词,并根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词对应的知识子图谱和/或事理子图谱;进一步在预先构建的标签图谱中查找与知识子图谱和/或事理子图谱对应的标签子图谱;最终查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。本方案通过知识图谱和/或事理图谱,以及标签图谱可快速有效地挖掘得到与需求词匹配的用户群体,避免了现有技术中需人工挖掘与需求词对应的用户群体而造成的效率低下,人工成本高的弊端;并且,本方案挖掘得到的用户群体与需求词具有较高的匹配度,可真实全面地反映该需求对应的潜在用户。

[0020] 上述说明仅是本发明技术方案的概述,为了能够更清楚了解本发明的技术手段,而可依照说明书的内容予以实施,并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂,以下特举本发明的具体实施方式。

附图说明

[0021] 通过阅读下文优选实施方式的详细描述,各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的,而并不认为是对本发明的限制。而且在整个附图中,用相同的参考符号表示相同的部件。在附图中:

[0022] 图1示出了根据本发明一个实施例提供的基于需求词的用户挖掘方法的流程示意图;

[0023] 图2示出了根据本发明另一个实施例提供的基于需求词的用户挖掘方法的流程示意图;

[0024] 图3a示出了根据本发明另一个实施例提供的知识图谱示意图;

[0025] 图3b示出了根据本发明另一个实施例提供的事理图谱示意图;

[0026] 图3c示出了根据本发明另一个实施例提供的标签图谱示意图;

[0027] 图4示出了根据本发明一个实施例提供的基于需求词的用户挖掘装置的结构框图;

[0028] 图5示出了根据本发明一个实施例提供的一种计算设备的结构示意图。

具体实施方式

[0029] 下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例,然而应当理解,可以以各种形式实现本公开而不应被这里阐述的实施例

所限制。相反,提供这些实施例是为了能够更透彻地理解本公开,并且能够将本公开的范围完整的传达给本领域的技术人员。

[0030] 图1示出了根据本发明一个实施例提供的基于需求词的用户挖掘方法的流程示意图,该方法基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行。如图1所示,该方法包括:

[0031] 步骤S110,获取输入的属于特定领域的需求词。

[0032] 其中,本步骤中的需求词并不局限于直接输入的特定领域的需求词,其还包括通过从输入的需求信息中提取出的属于特定领域的需求词。并且,本发明对输入的具体方式不做限定,例如,其可以为手动输入、语音录入、肢体识别录入等输入方式中的一种或多种的结合。

[0033] 步骤S120,根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词对应的知识子图谱和/或事理子图谱。

[0034] 其中,本实施例提供的基于需求词的用户挖掘方法是基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行的。知识图谱为结构化的语义知识库,用以描述物理世界中的概念及其相互关系,其基本组成单位是“知识实体-关系-知识实体”三元组,知识实体之间通过关系相互联结,构成网状的知识结构。而事理图谱是以“事件实体-关系-事件实体”三元组为核心,描述事件的发展、因果等关系的逻辑知识库。本实施例对具体的知识图谱关系预测模型和/或事理图谱关系预测模型的训练方法不做限定,本领域技术人员可自行选择相应的方法构建知识图谱关系预测模型和/或事理图谱关系预测模型。

[0035] 具体地,本步骤中根据训练得到的知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词对应的知识子图谱和/或事理子图谱。例如,根据训练得到的知识图谱关系预测模型在知识图谱中确定与需求词对应的知识实体,并将包含有与该知识实体相关联的知识实体的知识子图谱作为与该需求词对应的知识子图谱;和/或,根据训练得到的事理图谱关系预测模型在事理图谱中确定与需求词对应的事件实体,并将包含有与该事件实体相关联的事件实体的事理子图谱作为与该需求词对应的知识子图谱。

[0036] 通过本步骤,可获得与该需求词关联的知识子图谱和/或事理子图谱,避免直接根据需求词在标签图谱中挖掘得到的用户群体的全面性及准确度较低的弊端。

[0037] 步骤S130,在预先构建的标签图谱中查找与知识子图谱和/或事理子图谱对应的标签子图谱。

[0038] 其中,标签图谱是根据用户行为数据而构建,通过对用户行为数据抽取标签化信息后,获取标签之间的关联关系,构建“标签实体-关系-标签实体”三元组,从而建立标签图谱。

[0039] 并且,根据步骤S120中获得的知识子图谱和/或事理子图谱,通过相应的查找方法,查找出与知识子图谱和/或事理子图谱对应的标签子图谱。例如,可根据知识子图谱和/或事理子图谱中的各个知识实体或事件实体与标签图谱中的标签实体的相似度或关联性,确定与知识子图谱和/或事理子图谱对应的标签子图谱,本实施例对查找与知识子图谱和/或事理子图谱对应的标签子图谱的具体方式不做限定,本领域技术人员可根据实际的业务需求自行设定。

[0040] 步骤S140,查找用户行为数据与标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

[0041] 根据步骤S130获得的标签子图谱,查找用户行为数据与标签子图谱中任一标签实体匹配的用户。例如,标签子图谱包含“360金融”标签及“游戏”标签,则在用户行为数据中查找安装或使用360金融应用和/或游戏应用的用户群体作为与需求词匹配的用户群体。

[0042] 本实施例首先获取输入的属于特定领域的需求词,并根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词对应的知识子图谱和/或事理子图谱;进一步在预先构建的标签图谱中查找与知识子图谱和/或事理子图谱对应的标签子图谱;最终查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。本方案通过知识图谱和/或事理图谱,以及标签图谱可快速有效地挖掘得到与需求词匹配的用户群体,避免了现有技术中需人工挖掘与需求词对应的用户群体而造成的效率低下,人工成本高的弊端;并且,本方案挖掘得到的用户群体与需求词具有较高的匹配度,可真实全面地反映该需求对应的潜在用户。

[0043] 图2示出了根据本发明另一个实施例提供的基于需求词的用户挖掘方法的流程示意图。如图2所示,该方法包括:

[0044] 步骤S210,构建特定领域的知识图谱和/或事理图谱。

[0045] 具体地,在构建特定领域的知识图谱过程中,可先构建特定领域的初始知识图谱:其中,初始知识图谱在构建过程中,首先需设定该特定领域中的各个知识实体及对应的关系类型,如设定知识实体“银行”、知识实体“中国银行”及两者的对应关系为包含与被包含的关系,其次根据爬取的公开数据(包含结构化的知识图谱三元组数据及非结构化的文本数据),人工提取非结构化文本数据中的知识实体及知识实体间的关系,从而形成知识图谱三元组数据,并根据该结构化的知识图谱三元组数据以及根据非结构化的文本数据生成的知识图谱三元组数据整合为初始知识图谱;进一步地,在构建初始知识图谱之后,将该初始知识图谱作为训练集,构建知识图谱的实体抽取模型与实体关系分类模型,其中,构建知识图谱的实体抽取模型与实体关系分类模型的具体方法本发明不做限定,例如,可采用BiLSTM+CRF模型构建知识图谱的实体抽取模型,以及根据BiGRU+2attention模型来构建知识图谱的实体关系分类模型;进一步地,在构建知识图谱的实体抽取模型与实体关系分类模型之后,爬取特定领域知识类的结构化数据和非结构化文本数据,并利用构建的实体抽取模型与实体关系分类模型在该结构化数据和非结构文本数据中抽取知识图谱三元组,对初始知识图谱进行扩充,获得构建后的知识图谱。例如,图3a中示出了构建的金融领域知识图谱中的部分图谱,从图3a中可看出,知识图谱中的节点为各个知识实体,知识实体之间为对应的知识实体关系,如知识实体“储蓄”、“汇兑”、“借贷”均为知识实体“银行”的业务属性;而知识实体“分期”属于知识实体“借贷”中的一类。

[0046] 同理,在建立特定领域的事理图谱过程中,可构建特定领域的初始事理图谱,将初始事理图谱作为训练集,构建事理图谱的实体抽取模型与实体关系分类模型;爬取特定领域事理类的结构化数据和非结构化文本数据,根据事理图谱的实体抽取模型与实体关系分类模型在结构化数据和非结构文本数据中抽取事理图谱三元组,对初始事理图谱进行扩充,获得构建后的事理图谱。图3b为构建的金融领域的事理图谱中的部分图谱,从图3b中可看出事件“结婚”会导致事件“买房”和/或“旅行”的发生,而事件“买房”将导致事件“贷款”

的发生。

[0047] 步骤S220, 构建标签图谱。

[0048] 不同于步骤S210中的知识图谱及事理图谱, 本步骤建立的标签图谱并非基于特定领域数据而建立, 而是基于用户行为数据而构建。具体地, 通过获取用户行为数据, 提取用户行为数据中的用户标签数据, 根据用户与用户标签数据之间的关联关系, 建立标签图谱, 其中, 标签图谱中标签实体之间的关系是根据用户与用户标签数据之间的关联关系得到的。可选的, 在提取用户标签数据之后, 可进一步地根据用户标签数据获得相应的描述标签, 根据描述标签与用户标签的对应关系以及用户与用户标签数据之间的关联关系建立标签图谱。如图3c所示, 通过对用户行为数据的统计分析得出, 90%的用户同时安装有安居客应用及汽车之家应用, 而安装有汽车之家应用的用户中有80%安装有蚂蚁花呗, 则建立标签“安居客”与“汽车之家”的关联关系, 以及建立标签“汽车之家”与“蚂蚁花呗”的关联关系, 并且, 针对于标签“安居客”及标签“房天下”可获得相应的描述标签“买房”, 则进一步建立标签“买房”与标签“安居客”及“房天下”的关联关系; 同理, 建立“汽车之家”及“瓜子网”与对应的描述标签“买车”的关联关系, 以及, 建立“瓜子网”、“蚂蚁花呗”、及“京东白条”与对应的描述标签“分期”的对应关系。从图中可看出, 本步骤构建的标签图谱中并不记录用户信息(如用户ID, 用户名称等), 而是将根据用户行为数据获得的标签数据呈现至标签图谱中, 从而避免因将用户信息加入标签图谱而引发的数据处理量增加, 处理速度降低等弊端。

[0049] 步骤S230, 获取输入的属于特定领域的需求词。

[0050] 其中, 本步骤中的需求词并不局限于直接输入的特定领域的需求词, 其还包括通过从输入的需求信息中通过语义分析等方法提取出的属于特定领域的需求词。并且, 本发明对输入的具体方式不做限定, 例如, 其可以为手动输入、语音录入、肢体识别录入等输入方式中的一种或多种的结合。

[0051] 步骤S240, 在底层语料库查找与需求词对应的至少一个泛化词。

[0052] 具体地, 预先构建底层语料库。其中, 底层语料库的具体构建方法本领域技术人员可自行设置, 本实施例对此不做限定。例如, 可根据知识图谱、事理图谱, 以及标签图谱的原始语料数据构建底层语料库。可选的, 为便于后续泛化词的查找效率, 可将底层语料库中的语料数据转换为相应的词向量。例如, 可对原始语料数据进行数据清洗, 并对清理后的语料数据进行分词操作, 通过词向量训练模型(如CBOW word2vec词向量训练模型)将分词结果生成词向量, 获得包含词向量的底层语料库。

[0053] 进一步地, 在底层语料库中查找与该需求词对应的至少一个泛化词。在具体的实施过程中, 可预先对该需求词进行解析, 生成对应的需求词向量, 并根据需求词向量与底层语料库中各个语料的词向量之间的距离, 查找与需求词对应的至少一个泛化词。具体地, 可将需求词向量与底层语料库中各个语料的词向量之间的距离由小至大进行排序, 并获取位于排序位列中前n个词向量, 将该n个词向量对应的词确定为与该需求词对应的泛化词。例如, 需求词为“贷款”, 则通过底层语料库可获取与“贷款”词向量距离最近的4个词向量, 将该4个词向量对应的词“借贷”、“借钱”、“借款”、以及“贷钱”作为与“贷款”对应的泛化词。

[0054] 步骤S250, 根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词和至少一个泛化词对应的知识子图谱和/或事理子图谱。

[0055] 具体地, 利用距离计算方法, 查找知识图谱中与需求词向量和至少一个泛化词向

量的距离最小或者小于预设阈值的至少一个知识实体,并根据知识图谱关系预测模型对至少一个知识实体进行实体推荐,得到知识子图谱。其中,知识图谱关系预测模型可以为基于机器学习方法,以知识图谱中的三元组数据为训练集,通过RPEM (representation prediction embedding model) 模型而构建。例如,在图3a所示的知识图谱中查找与需求词“贷款”词向量的距离最小的知识实体“借贷”,并根据知识图谱关系预测模型对该知识实体“借贷”进行实体推荐,将与知识实体“借贷”关联的知识实体“分期”、“银行”、“民间借贷”的组成的子图谱确定为与需求词对应的知识图谱子图谱。

[0056] 和/或,利用距离计算方法,查找事理图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个事理实体;根据事理图谱关系预测模型对至少一个事理实体进行实体推荐,得到事理子图谱。其中,事理图谱关系预测模型可基于机器学习方法,以事理图谱中的三元组数据为训练集,通过RPEM模型而构建。例如,在图3b所示的事理图谱中查找与需求词“贷款”距离最近的事件实体“贷款”,进而根据事理图谱关系预测模型对该事件实体“贷款”进行实体推荐,将包含有与事件实体“贷款”相关联的事件实体“买房”的子图谱确定与需求词对应的事理图谱子图谱。

[0057] 步骤S260,在标签图谱中检索与知识子图谱和/或事理子图谱对应的标签子图谱。

[0058] 具体地,在标签图谱中检索与知识子图谱对应的标签子图谱时,可利用距离计算方法,查找标签图谱中与知识子图谱和/或事理子图谱中的各个知识实体和/或各个事件实体的距离最小或者小于预设阈值的至少一个标签实体;根据标签图谱关系预测模型对至少一个标签实体进行实体推荐,得到标签子图谱。其中,标签图谱关系预测模型可基于机器学习方法,以标签图谱中的“标签实体-关系-标签实体”三元组数据为训练集,通过RPEM模型而构建。

[0059] 距离来说,若步骤S250中确定的知识图谱子图谱中包含知识实体“分期”,确定的事理图谱子图谱中包含事件实体“买房”,则确定图3c所示标签图谱中的标签实体“分期”及“买房”,并进一步通过标签图谱关系预测模型对标签实体“分期”及“买房”进行实体推荐,确定包含与标签“分期”相关联的标签“瓜子网”、“京东白条”及“蚂蚁花呗”,以及与标签“买房”相关联的标签“安居客”、“房天下”和/或“汽车之家”的子图谱确定为对应的标签子图谱。

[0060] 步骤S270,查找用户行为数据与标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

[0061] 根据步骤S260获得的标签子图谱,查找用户行为数据与标签子图谱中任一标签实体匹配的用户,从而挖掘得到与需求词匹配的用户群体。

[0062] 可选的,在挖掘得到与需求词匹配的用户群体之后,可进一步地为该用户群体分配与该需求词对应的标签,从而便于根据该标签快速地检索相应的用户。

[0063] 本实施例首先获取输入的属于特定领域的需求词,并在底层语料库中查找与需求词对应的至少一个泛化词,从而提高最终挖掘的用户的全面性及准确性;并根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词对应的知识子图谱和/或事理子图谱;进一步在预先构建的标签图谱中查找与知识子图谱和/或事理子图谱对应的标签子图谱;最终查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。本方案通过知识图谱和/或事理图谱,以及标签图谱可快速

有效地挖掘得到与需求词匹配的用户群体,避免了现有技术中需人工挖掘与需求词对应的用户群体而造成的效率低下,人工成本高的弊端;并且,本方案挖掘得到的用户群体与需求词具有较高的匹配度,可真实全面地反映该需求词对应的潜在用户。

[0064] 图4示出了根据本发明一个实施例提供的基于需求词的用户挖掘装置的结构框图。该装置基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行。如图4所示,该装置包括:获取模块41、计算模块42、查找模块43、以及挖掘模块44。

[0065] 其中,获取模块41,适于获取输入的属于特定领域的需求词。

[0066] 计算模块42,适于根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱。

[0067] 查找模块43,适于在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱。

[0068] 挖掘模块44,适于查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

[0069] 可选的,该装置还包括:泛化模块(图中未示出),适于在所述获取模块获取输入的属于特定领域的需求词之后,在底层语料库查找与所述需求词对应的至少一个泛化词;

[0070] 计算模块42进一步适于:根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词和所述至少一个泛化词对应的知识子图谱和/或事理子图谱。

[0071] 可选的,泛化模块进一步适于:对所述需求词进行解析,生成所述需求词向量;

[0072] 根据所述需求词向量与所述底层语料库中各个语料的词向量之间的距离,查找与所述需求词对应的至少一个泛化词。

[0073] 可选的,计算模块42进一步适于:

[0074] 利用距离计算方法,查找知识图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个知识实体;根据知识图谱关系预测模型对所述至少一个知识实体进行实体推荐,得到所述知识子图谱;

[0075] 和/或,利用距离计算方法,查找事理图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个事理实体;根据事理图谱关系预测模型对所述至少一个事理实体进行实体推荐,得到所述事理子图谱。

[0076] 可选的,该装置还包括:第一建立模块(图中未示出),适于预先按照以下步骤构建知识图谱:

[0077] 构建特定领域的初始知识图谱,将所述初始知识图谱作为训练集,构建知识图谱的实体抽取模型与实体关系分类模型;

[0078] 爬取特定领域知识类的结构化数据和非结构化文本数据;

[0079] 根据知识图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构化文本数据中抽取知识图谱三元组,对初始知识图谱进行扩充,获得构建后的知识图谱。

[0080] 可选的,该装置还包括:第二建立模块(图中未示出),适于预先按照以下步骤构建事理图谱:

[0081] 构建特定领域的初始事理图谱,将所述初始事理图谱作为训练集,构建事理图谱的实体抽取模型与实体关系分类模型;

- [0082] 爬取特定领域事理类的结构化数据和非结构化文本数据；
- [0083] 根据事理图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构化文本数据中抽取事理图谱三元组,对初始事理图谱进行扩充,获得构建后的事理图谱。
- [0084] 可选的,该装置还包括第三建立模块(图中未示出),适于预先按照以下步骤构建标签图谱:
- [0085] 获取用户行为数据;
- [0086] 提取所述用户行为数据中的用户标签数据;
- [0087] 根据用户与用户标签数据之间的关联关系,建立标签图谱,其中所述标签图谱中标签实体之间的关系是根据用户与用户标签数据之间的关联关系得到的。
- [0088] 可选的,该装置还包括:语料库构建模块(图中未示出),适于根据知识图谱、事理图谱,以及标签图谱的原始语料数据构建底层语料库。
- [0089] 可选的,该语料库构建模块进一步适于:对所述原始语料数据进行数据清洗,并对清理后的语料数据进行分词操作;
- [0090] 通过词向量训练模型将分词结果生成词向量,获得包含词向量的底层语料库。
- [0091] 可选的,该装置还包括:标签配置模块(图中未示出),适于在所述挖掘模块在挖掘得到与需求词匹配的用户群体之后,为所述用户群体中的所有用户配置与所述需求词对应的标签。
- [0092] 本装置首先获取输入的属于特定领域的需求词,并根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算需求词对应的知识子图谱和/或事理子图谱;进一步在预先构建的标签图谱中查找与知识子图谱和/或事理子图谱对应的标签子图谱;最终查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。本装置通过知识图谱和/或事理图谱,以及标签图谱可快速有效地挖掘得到与需求词匹配的用户群体,避免了现有技术中需人工挖掘与需求词对应的用户群体而造成的效率低下,人工成本高的弊端;并且,本方案挖掘得到的用户群体与需求词具有较高的匹配度,可真实全面地反映该需求对应的潜在用户。
- [0093] 根据本发明一个实施例提供了一种非易失性计算机存储介质,所述计算机存储介质存储有至少一可执行指令,该计算机可执行指令可执行上述任意方法实施例中的基于需求词的用户挖掘方法。
- [0094] 图5示出了根据本发明一个实施例提供的一种计算设备的结构示意图,本发明具体实施例并不对计算设备的具体实现做限定。
- [0095] 如图5所示,该终端可以包括:处理器(processor)502、通信接口(Communications Interface)504、存储器(memory)506、以及通信总线508。
- [0096] 其中:
- [0097] 处理器502、通信接口504、以及存储器506通过通信总线508完成相互间的通信。
- [0098] 通信接口504,用于与其它设备比如客户端或其它服务器等的网元通信。
- [0099] 处理器502,用于执行程序510,具体可以执行上述基于需求词的用户挖掘方法实施例中的相关步骤。
- [0100] 具体地,程序510可以包括程序代码,该程序代码包括计算机操作指令。
- [0101] 处理器502可能是中央处理器CPU,或者是特定集成电路ASIC(Application

Specific Integrated Circuit), 或者是被配置成实施本发明实施例的一个或多个集成电路。计算设备包括的一个或多个处理器, 可以是同一类型的处理器, 如一个或多个CPU; 也可以是不同类型的处理器, 如一个或多个CPU以及一个或多个ASIC。

[0102] 存储器506, 用于存放程序510。存储器506可能包含高速RAM存储器, 也可能还包括非易失性存储器(non-volatile memory), 例如至少一个磁盘存储器。

[0103] 程序510具体可以用于使得处理器502执行以下操作:

[0104] 获取输入的属于特定领域的需求词;

[0105] 根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱;

[0106] 在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱;

[0107] 查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户, 进而挖掘得到与需求词匹配的用户群体。

[0108] 在一种可选的实施方式中, 程序510具体可以用于使得处理器502执行以下操作:

[0109] 在底层语料库查找与所述需求词对应的至少一个泛化词;

[0110] 根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词和所述至少一个泛化词对应的知识子图谱和/或事理子图谱。

[0111] 在一种可选的实施方式中, 程序510具体可以用于使得处理器502执行以下操作:

[0112] 对所述需求词进行解析, 生成所述需求词向量;

[0113] 根据所述需求词向量与所述底层语料库中各个语料的词向量之间的距离, 查找与所述需求词对应的至少一个泛化词。

[0114] 在一种可选的实施方式中, 程序510具体可以用于使得处理器502执行以下操作:

[0115] 利用距离计算方法, 查找知识图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个知识实体; 根据知识图谱关系预测模型对所述至少一个知识实体进行实体推荐, 得到所述知识子图谱;

[0116] 和/或, 利用距离计算方法, 查找事理图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个事理实体; 根据事理图谱关系预测模型对所述至少一个事理实体进行实体推荐, 得到所述事理子图谱。

[0117] 在一种可选的实施方式中, 程序510具体可以用于使得处理器502执行以下操作:

[0118] 构建特定领域的初始知识图谱, 将所述初始知识图谱作为训练集, 构建知识图谱的实体抽取模型与实体关系分类模型;

[0119] 爬取特定领域知识类的结构化数据和非结构化文本数据;

[0120] 根据知识图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构化文本数据中抽取知识图谱三元组, 对初始知识图谱进行扩充, 获得构建后的知识图谱。

[0121] 在一种可选的实施方式中, 程序510具体可以用于使得处理器502执行以下操作:

[0122] 构建特定领域的初始事理图谱, 将所述初始事理图谱作为训练集, 构建事理图谱的实体抽取模型与实体关系分类模型;

[0123] 爬取特定领域事理类的结构化数据和非结构化文本数据;

[0124] 根据事理图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构

文本数据中抽取事理图谱三元组,对初始事理图谱进行扩充,获得构建后的事理图谱。

[0125] 在一种可选的实施方式中,程序510具体可以用于使得处理器502执行以下操作:

[0126] 获取用户行为数据;

[0127] 提取所述用户行为数据中的用户标签数据;

[0128] 根据用户与用户标签数据之间的关联关系,建立标签图谱,其中所述标签图谱中标签实体之间的关系是根据用户与用户标签数据之间的关联关系得到的。

[0129] 在一种可选的实施方式中,程序510具体可以用于使得处理器502执行以下操作:

[0130] 根据知识图谱、事理图谱,以及标签图谱的原始语料数据构建底层语料库。

[0131] 在一种可选的实施方式中,程序510具体可以用于使得处理器502执行以下操作:

[0132] 对所述原始语料数据进行数据清洗,并对清理后的语料数据进行分词操作;

[0133] 通过词向量训练模型将分词结果生成词向量,获得包含词向量的底层语料库。

[0134] 在一种可选的实施方式中,程序510具体可以用于使得处理器502执行以下操作:

[0135] 为所述用户群体中的所有用户配置与所述需求词对应的标签。

[0136] 在此提供的算法和显示不与任何特定计算机、虚拟系统或者其它设备固有相关。各种通用系统也可以与基于在此的示教一起使用。根据上面的描述,构造这类系统所要求的结构是显而易见的。此外,本发明也不针对任何特定编程语言。应当明白,可以利用各种编程语言实现在此描述的本发明的内容,并且上面对特定语言所做的描述是为了披露本发明的最佳实施方式。

[0137] 在此处所提供的说明书中,说明了大量具体细节。然而,能够理解,本发明的实施例可以在没有这些具体细节的情况下实践。在一些实例中,并未详细示出公知的方法、结构和技术,以便不模糊对本说明书的理解。

[0138] 类似地,应当理解,为了精简本公开并帮助理解各个发明方面中的一个或多个,在上面对本发明的示例性实施例的描述中,本发明的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而,并不应将该公开的方法解释成反映如下意图:即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说,如下面的权利要求书所反映的那样,发明方面在于少于前面公开的单个实施例的所有特征。因此,遵循具体实施方式的权利要求书由此明确地并入该具体实施方式,其中每个权利要求本身都作为本发明的单独实施例。

[0139] 本领域那些技术人员可以理解,可以对实施例中的设备中的模块进行自适应性地改变并且把它们设置在与该实施例不同的一个或多个设备中。可以把实施例中的模块或单元或组件组合成一个模块或单元或组件,以及此外可以把它分成多个子模块或子单元或子组件。除了这样的特征和/或过程或者单元中的至少一些是相互排斥之外,可以采用任何组合对本说明书(包括伴随的权利要求、摘要和附图)中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述,本说明书(包括伴随的权利要求、摘要和附图)中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

[0140] 此外,本领域的技术人员能够理解,尽管在此所述的一些实施例包括其它实施例中所未包括的某些特征而不是其它特征,但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如,在权利要求书中,所要求保护的实施例的任意之一

都可以以任意的组合方式来使用。

[0141] 本发明的各个部件实施例可以以硬件实现,或者以在一个或者多个处理器上运行的软件模块实现,或者以它们的组合实现。本领域的技术人员应当理解,可以在实践中使用微处理器或者数字信号处理器(DSP)来实现根据本发明实施例的装置中的一些或者全部部件的一些或者全部功能。本发明还可以实现为用于执行这里所描述的方法的一部分或者全部的设备或者装置程序(例如,计算机程序和计算机程序产品)。这样的实现本发明的程序可以存储在计算机可读介质上,或者可以具有一个或者多个信号的形式。这样的信号可以从因特网网站上下下载得到,或者在载体信号上提供,或者以任何其他形式提供。

[0142] 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制,并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中,不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了若干装置的单元权利要求中,这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

[0143] 本发明公开了:A1.一种基于需求词的用户挖掘方法,所述方法基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行,所述方法包括:

[0144] 获取输入的属于特定领域的需求词;

[0145] 根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱;

[0146] 在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱;

[0147] 查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

[0148] A2.根据A1所述的方法,其中,在所述获取输入的属于特定领域的需求词之后,所述方法还包括:在底层语料库查找与所述需求词对应的至少一个泛化词;

[0149] 所述根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词对应的知识子图谱和/或事理子图谱具体为:根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词和所述至少一个泛化词对应的知识子图谱和/或事理子图谱。

[0150] A3.根据A2所述的方法,其中,所述在底层语料库查找与所述需求词对应的至少一个泛化词进一步包括:

[0151] 对所述需求词进行解析,生成所述需求词向量;

[0152] 根据所述需求词向量与所述底层语料库中各个语料的词向量之间的距离,查找与所述需求词对应的至少一个泛化词。

[0153] A4.根据A2所述的方法,其中,所述根据知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词和所述至少一个泛化词对应的知识子图谱和/或事理子图谱进一步包括:

[0154] 利用距离计算方法,查找知识图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个知识实体;根据知识图谱关系预测模型对所述至少一个知识实体进行实体推荐,得到所述知识子图谱;

[0155] 和/或,利用距离计算方法,查找事理图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个事理实体;根据事理图谱关系预测模型对所述至少一个事理实体进行实体推荐,得到所述事理子图谱。

[0156] A5.根据A1-A4中任一项所述的方法,其中,所述方法还包括预先按照以下步骤构建知识图谱:

[0157] 构建特定领域的初始知识图谱,将所述初始知识图谱作为训练集,构建知识图谱的实体抽取模型与实体关系分类模型;

[0158] 爬取特定领域知识类的结构化数据和非结构化文本数据;

[0159] 根据知识图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构化文本数据中抽取知识图谱三元组,对初始知识图谱进行扩充,获得构建后的知识图谱。

[0160] A6.根据A1-A4中任一项所述的方法,其中,所述方法还包括预先按照以下步骤构建事理图谱:

[0161] 构建特定领域的初始事理图谱,将所述初始事理图谱作为训练集,构建事理图谱的实体抽取模型与实体关系分类模型;

[0162] 爬取特定领域事理类的结构化数据和非结构化文本数据;

[0163] 根据事理图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构化文本数据中抽取事理图谱三元组,对初始事理图谱进行扩充,获得构建后的事理图谱。

[0164] A7.根据A1-A6中任一项所述的方法,其中,所述方法还包括预先按照以下步骤构建标签图谱:

[0165] 获取用户行为数据;

[0166] 提取所述用户行为数据中的用户标签数据;

[0167] 根据用户与用户标签数据之间的关联关系,建立标签图谱,其中所述标签图谱中标签实体之间的关系是根据用户与用户标签数据之间的关联关系得到的。

[0168] A8.根据A2所述的方法,其中,所述方法还包括:

[0169] 根据知识图谱、事理图谱,以及标签图谱的原始语料数据构建底层语料库。

[0170] A9.根据A8所述的方法,其中,所述根据知识图谱、事理图谱,以及标签图谱的原始语料数据构建底层语料库进一步包括:

[0171] 对所述原始语料数据进行数据清洗,并对清理后的语料数据进行分词操作;

[0172] 通过词向量训练模型将分词结果生成词向量,获得包含词向量的底层语料库。

[0173] A10.根据A1所述的方法,其中,在所述挖掘得到与需求词匹配的用户群体之后,所述方法还包括:为所述用户群体中的所有用户配置与所述需求词对应的标签。

[0174] 本发明还公开了:B11.一种基于需求词的用户挖掘装置,所述装置基于经过训练得到的特定领域的知识图谱关系预测模型和/或事理图谱关系预测模型而执行,所述装置包括:

[0175] 获取模块,适于获取输入的属于特定领域的需求词;

[0176] 计算模块,适于根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分

别计算所述需求词对应的知识子图谱和/或事理子图谱；

[0177] 查找模块,适于在预先构建的标签图谱中查找与所述知识子图谱和/或事理子图谱对应的标签子图谱；

[0178] 挖掘模块,适于查找用户行为数据与所述标签子图谱中任一标签实体匹配的用户,进而挖掘得到与需求词匹配的用户群体。

[0179] B12.根据B11所述的装置,其中,所述装置还包括:

[0180] 泛化模块,适于在所述获取模块获取输入的属于特定领域的需求词之后,在底层语料库查找与所述需求词对应的至少一个泛化词；

[0181] 所述计算模块进一步适于:根据所述知识图谱关系预测模型和/或事理图谱关系预测模型分别计算所述需求词和所述至少一个泛化词对应的知识子图谱和/或事理子图谱。

[0182] B13.根据B12所述的装置,其中,所述泛化模块进一步适于:

[0183] 对所述需求词进行解析,生成所述需求词向量；

[0184] 根据所述需求词向量与所述底层语料库中各个语料的词向量之间的距离,查找与所述需求词对应的至少一个泛化词。

[0185] B14.根据B12所述的装置,其中,所述计算模块进一步适于:

[0186] 利用距离计算方法,查找知识图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个知识实体;根据知识图谱关系预测模型对所述至少一个知识实体进行实体推荐,得到所述知识子图谱；

[0187] 和/或,利用距离计算方法,查找事理图谱中与所述需求词向量和至少一个泛化词向量的距离最小或者小于预设阈值的至少一个事理实体;根据事理图谱关系预测模型对所述至少一个事理实体进行实体推荐,得到所述事理子图谱。

[0188] B15.根据B11-B14中任一项所述的装置,其中,所述装置还包括:

[0189] 第一建立模块,适于预先按照以下步骤构建知识图谱:

[0190] 构建特定领域的初始知识图谱,将所述初始知识图谱作为训练集,构建知识图谱的实体抽取模型与实体关系分类模型；

[0191] 爬取特定领域知识类的结构化数据和非结构化文本数据；

[0192] 根据知识图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构化文本数据中抽取知识图谱三元组,对初始知识图谱进行扩充,获得构建后的知识图谱。

[0193] B16.根据B11-B14中任一项所述的装置,其中,所述装置还包括:

[0194] 第二建立模块,适于预先按照以下步骤构建事理图谱:

[0195] 构建特定领域的初始事理图谱,将所述初始事理图谱作为训练集,构建事理图谱的实体抽取模型与实体关系分类模型；

[0196] 爬取特定领域事理类的结构化数据和非结构化文本数据；

[0197] 根据事理图谱的实体抽取模型与实体关系分类模型在所述结构化数据和非结构化文本数据中抽取事理图谱三元组,对初始事理图谱进行扩充,获得构建后的事理图谱。

[0198] B17.根据B11-B16中任一项所述的装置,其中,所述装置还包括

[0199] 第三建立模块,适于预先按照以下步骤构建标签图谱:

[0200] 获取用户行为数据；

- [0201] 提取所述用户行为数据中的用户标签数据；
- [0202] 根据用户与用户标签数据之间的关联关系，建立标签图谱，其中所述标签图谱中标签实体之间的关系是根据用户与用户标签数据之间的关联关系得到的。
- [0203] B18. 根据B12所述的装置，其中，所述装置还包括：
- [0204] 语料库构建模块，适于根据知识图谱、事理图谱，以及标签图谱的原始语料数据构建底层语料库。
- [0205] B19. 根据B18所述的装置，其中，所述语料库构建模块进一步适于：
- [0206] 对所述原始语料数据进行数据清洗，并对清理后的语料数据进行分词操作；
- [0207] 通过词向量训练模型将分词结果生成词向量，获得包含词向量的底层语料库。
- [0208] B20. 根据B11所述的装置，其中，所述装置还包括：
- [0209] 标签配置模块，适于在所述挖掘模块在挖掘得到与需求词匹配的用户群体之后，为所述用户群体中的所有用户配置与所述需求词对应的标签。
- [0210] 本发明还公开了：C21一种计算设备，包括：处理器、存储器、通信接口和通信总线，所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信；
- [0211] 所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执行如A1-A10中任一项所述的基于需求词的用户挖掘方法对应的操作。
- [0212] 本发明还公开了：D22一种计算机存储介质，所述存储介质中存储有至少一可执行指令，所述可执行指令使处理器执行如A1-A10中任一项所述的基于需求词的用户挖掘方法对应的操作。

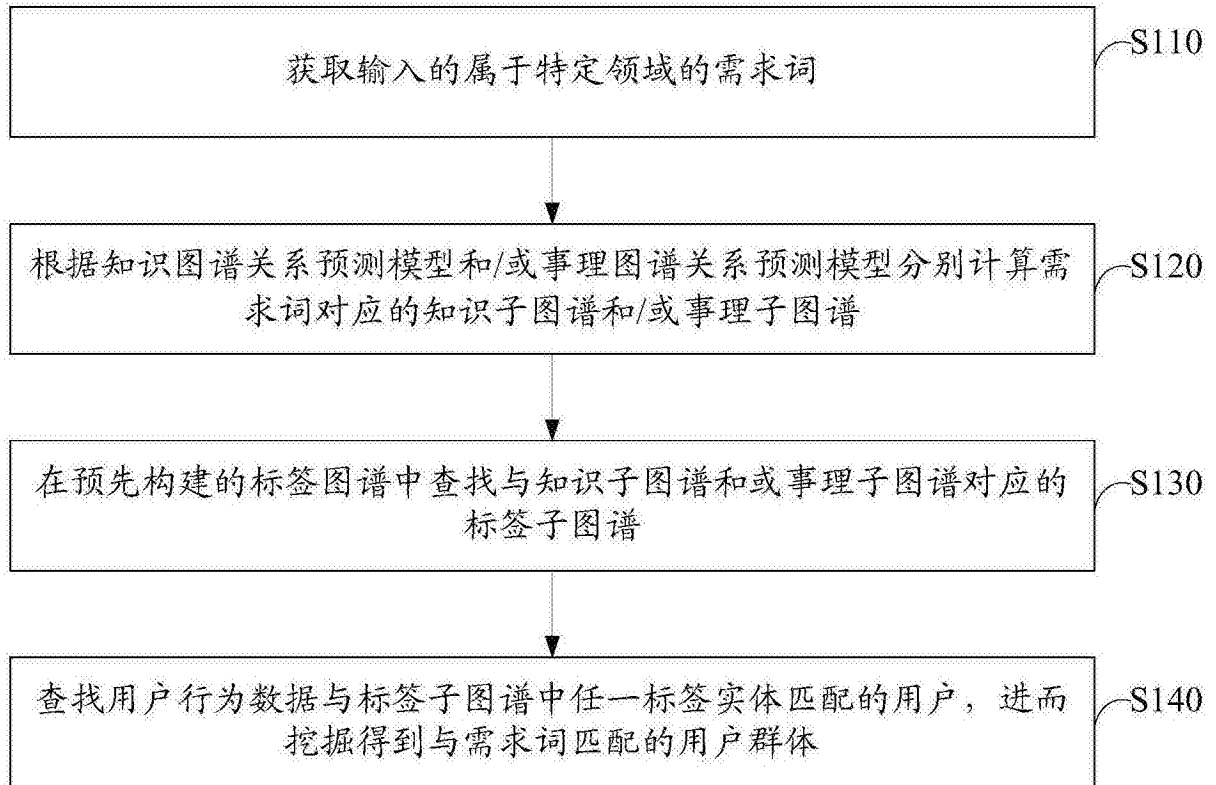


图1

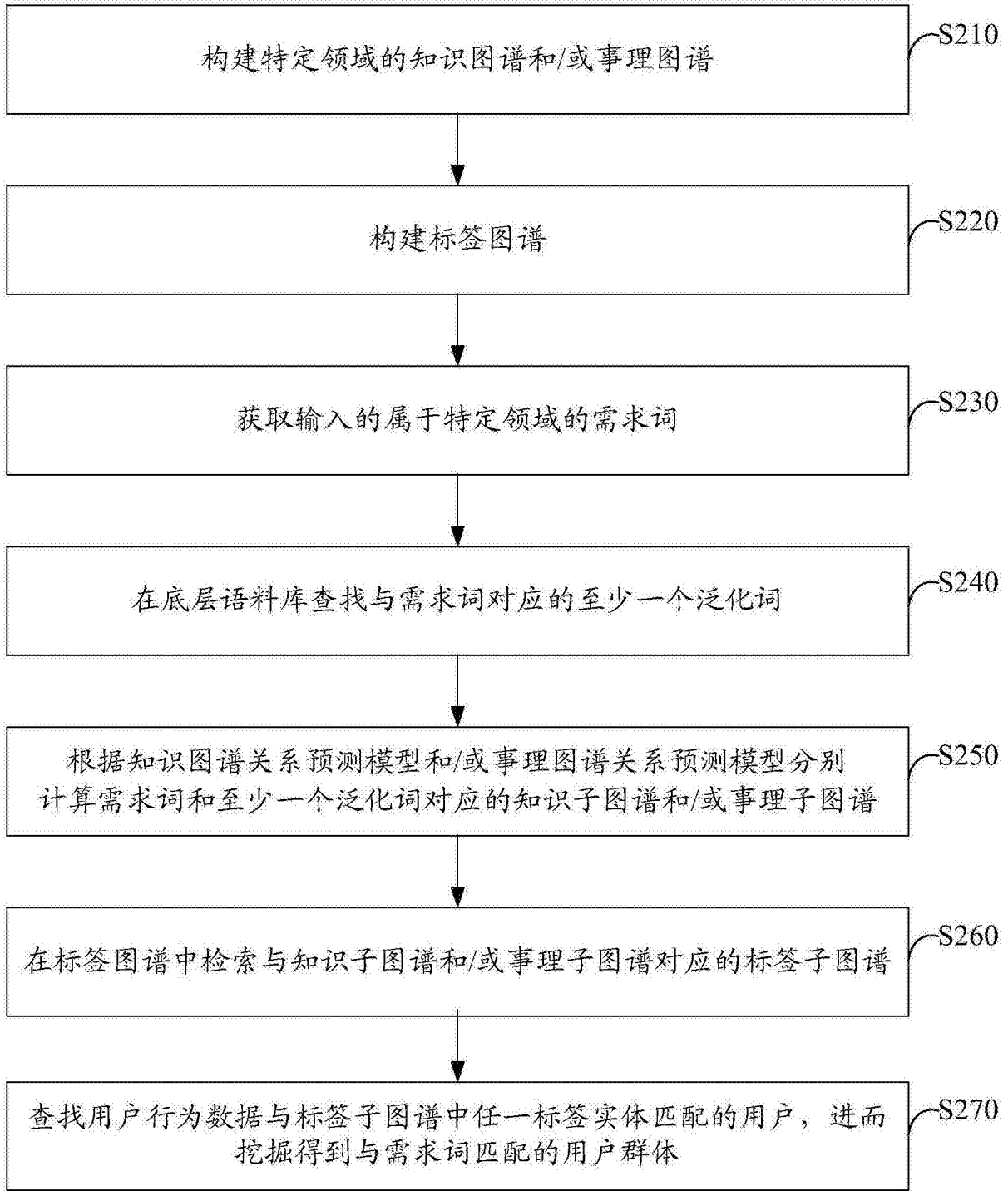


图2

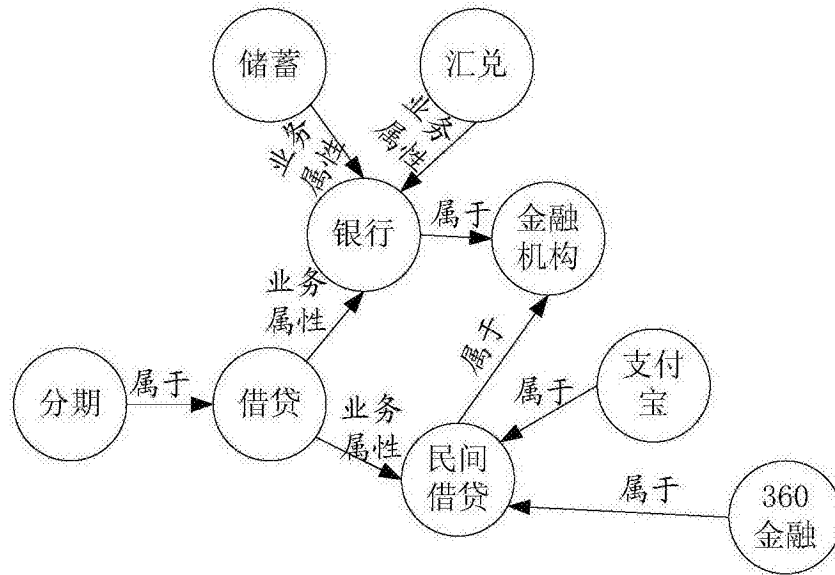


图3a

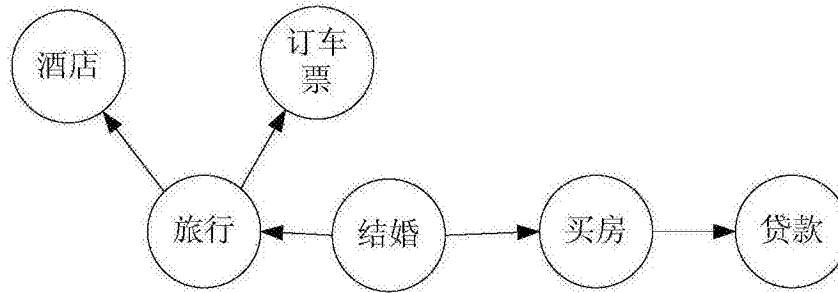


图3b

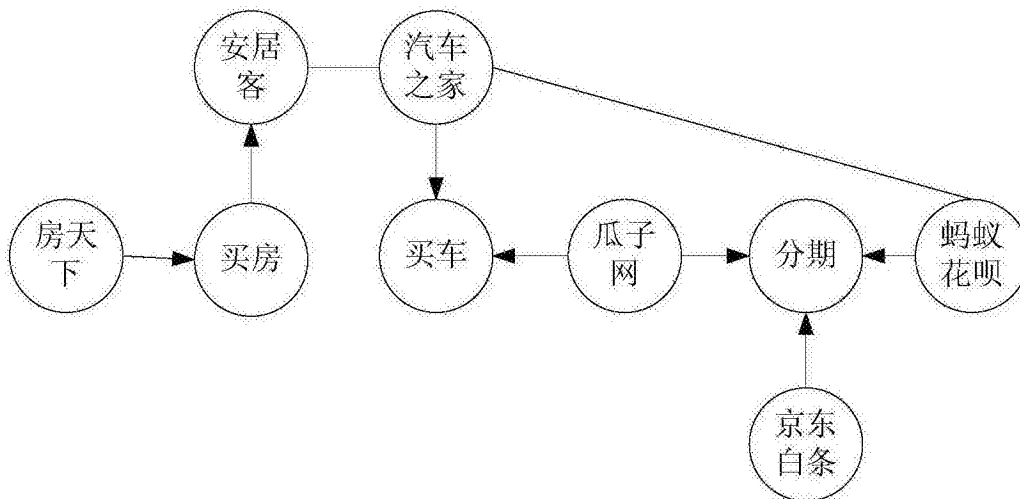


图3c

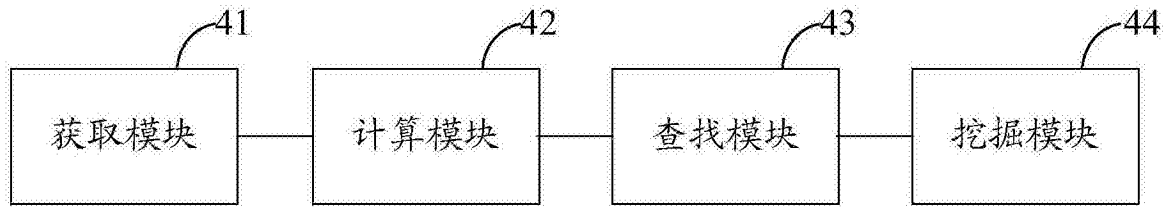


图4

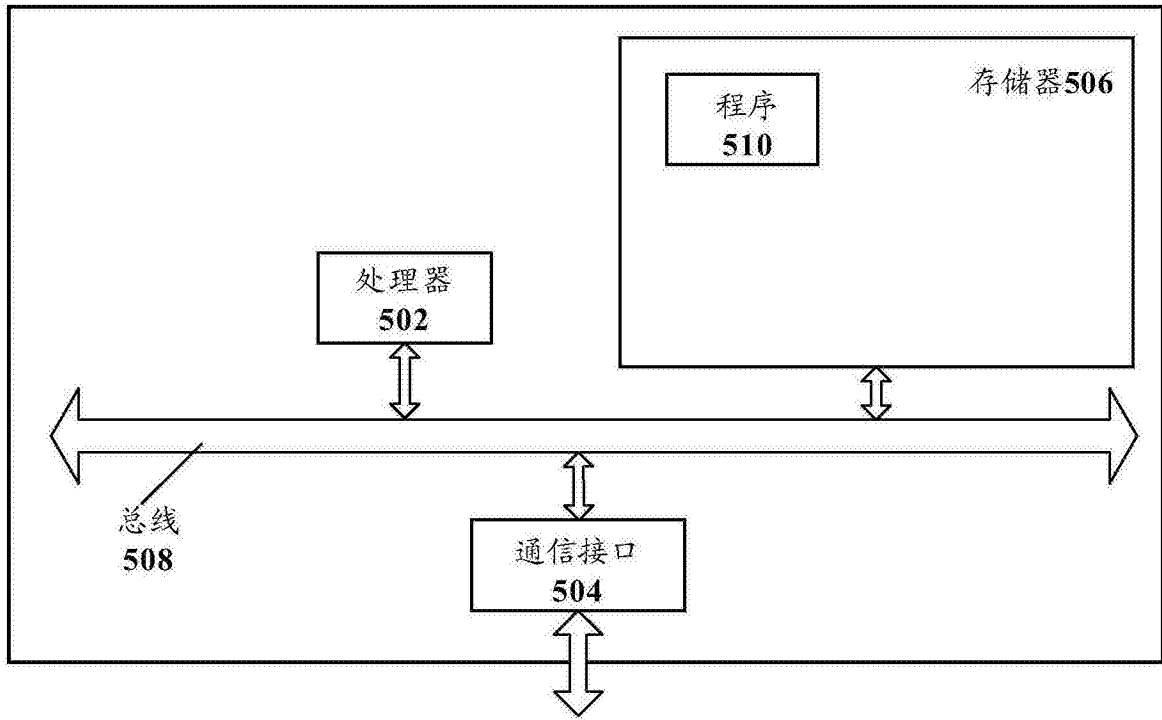


图5