



(12) 发明专利

(10) 授权公告号 CN 113961811 B

(45) 授权公告日 2024. 04. 05

(21) 申请号 202111262925.X

G06F 16/36 (2019.01)

(22) 申请日 2021.10.28

G06F 40/35 (2020.01)

G06F 16/955 (2019.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 113961811 A

(56) 对比文件

(43) 申请公布日 2022.01.21

CN 111475625 A, 2020.07.31

CN 113407729 A, 2021.09.17

(73) 专利权人 平安科技(深圳)有限公司

CN 113434703 A, 2021.09.24

KR 20210040891 A, 2021.04.14

地址 518000 广东省深圳市福田区福田街
道福安社区益田路5033号平安金融中
心23楼

审查员 李伟腾

(72) 发明人 童鑫 任笑笑

(74) 专利代理机构 深圳众鼎专利商标代理事务
所(普通合伙) 44325

专利代理师 谭果林

(51) Int. Cl.

G06F 16/9535 (2019.01)

权利要求书2页 说明书10页 附图4页

(54) 发明名称

基于事件图谱的话术推荐方法、装置、设备
及介质

(57) 摘要

本发明涉及智能推荐技术领域,本发明公开了一种基于事件图谱的话术推荐方法、装置、设备及介质,所述方法包括:获取用户的用户信息和行为轨迹,并对用户信息进行用户画像,获得用户的用户类别;爬取预设时间段内的热门事件,并进行要素识别,识别出事件图谱,以及对各事件图谱进行类别识别,得到各热门事件的事件类别;获取服务类型推荐模型,对行为轨迹进行预测,得到用户的推荐结果;从所有热门事件中获取相匹配的热门事件,并记录为推荐事件;运用TextRank算法,对推荐事件进行文本抽取,得到摘要文本;基于摘要文本和相应推荐模板,生成与用户对应的推荐话术。因此,本发明实现了自动匹配出推荐事件,更新话术内容,提高了话术推荐的及时性。



1. 一种基于事件图谱的话术推荐方法,其特征在于,包括:

获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别;

爬取预设时间段内的热门事件,并对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别;

获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果;

从所有所述热门事件中获得与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所述热门事件记录为推荐事件;

运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本;

基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术;

所述对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别,包括:

运用文本关键要素识别算法,对各所述热门事件进行关键要素提取,得到与各所述热门事件对应的多个三元组;

基于与各所述热门事件对应的所有所述三元组,构建各所述热门事件的事件图谱;

运用图卷积网络,对各所述事件图谱进行类别识别,得到各所述热门事件的事件类别。

2. 如权利要求1所述的基于事件图谱的话术推荐方法,其特征在于,所述对所述用户信息进行用户画像,获得所述用户的用户类别,包括:

对所述用户信息进行图谱构建,得到用户图谱;

基于所述用户图谱,对所述用户进行用户画像,得到所述用户类别。

3. 如权利要求1所述的基于事件图谱的话术推荐方法,其特征在于,所述爬取预设时间段内的热门事件,包括:

获取预设时间段内的各URL信息;

基于URL的类别规则,从获取的所有所述URL信息中筛选出与新闻类别对应的URL信息,并将筛选出的所述URL信息记录为待提取URL信息;

基于收集的热搜文本,运用网络爬取技术,爬取各所述待提取URL信息所对应的页面数据中与所述热搜文本匹配的文本内容,并将匹配的文本内容标记为所述热门事件。

4. 如权利要求2所述的基于事件图谱的话术推荐方法,其特征在于,所述通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果,包括:

通过所述服务类型推荐模型对所述行为轨迹进行关注服务的时长分析,得到所述用户的服务认知结果;

获取所述用户图谱中的交际用户,爬取与所述交际用户关联的服务事件;

根据所述服务认知结果和所有所述服务事件,对所述行为轨迹进行意图分析,分析出与所述用户对应的所述推荐结果。

5. 如权利要求1所述的基于事件图谱的话术推荐方法,其特征在于,所述运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本,包括:

对所述推荐事件进行分词及虚词去除处理,得到可选分词;

运用TextRank算法,对各所述可选分词进行词性标注,并构建可选关键词图,对所述可选关键词图进行关系抽取,得到多个抽取关键词;

将各所述抽取关键词按照所述推荐事件中的文本顺序进行拼接,得到所述摘要文本。

6.一种基于事件图谱的话术推荐装置,其特征在于,包括:

获取模块,用于获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别;

爬取模块,用于爬取预设时间段内的热门事件,并对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别;

预测模块,用于获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果;

匹配模块,用于从所有所述热门事件中获得与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所述热门事件记录为推荐事件;

抽取模块,用于运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本;

生成模块,用于基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术;

所述爬取模块包括:

第一识别单元,用于运用文本关键要素识别算法,对各所述热门事件进行关键要素提取,得到与各所述热门事件对应的多个三元组;

构建单元,用于基于与各所述热门事件对应的所有所述三元组,构建各所述热门事件的事件图谱;

第二识别单元,用于运用图卷积网络,对各所述事件图谱进行类别识别,得到各所述热门事件的事件类别。

7.一种计算机设备,包括存储器、处理器以及存储在所述存储器中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计算机程序时实现如权利要求1至5任一项所述基于事件图谱的话术推荐方法。

8.一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,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任一项所述基于事件图谱的话术推荐方法。

基于事件图谱的话术推荐方法、装置、设备及介质

技术领域

[0001] 本发明涉及智能推荐技术领域,尤其涉及一种基于事件图谱的话术推荐方法、装置、设备及介质。

背景技术

[0002] 近年来,随着互联网的快速发展,信息正呈爆炸式增长,如何从大量的信息中筛选出用户感兴趣的内容成为互联网领域的研究重点,因此,信息推荐技术在近几年也取得了比较大的进步,其可向用户推荐感兴趣的信息或产品,以满足用户需求。

[0003] 现有推荐信息或者产品时,往往是通过分析用户的历史浏览数据,根据历史浏览数据筛选出用户感兴趣的信息或者产品,因历史浏览数据只能表征用户整体的兴趣偏向,难以将推荐信息或者产品跟当前热门事件相结合,无法准确地就热门事件向用户进行推荐合适或者感兴趣的信息或者产品,因此,推送内容信息或者产品时,容易与用户当前的关注度相差加大,无法与热门事件相结合容易忽略消息或者产品对自身的重要程度,从而错失关注或者重要的消息或者产品。

发明内容

[0004] 本发明提供一种基于事件图谱的话术推荐方法、装置、设备及介质,实现了自动匹配出推荐事件,并相应地输出推荐话术,做到与时俱进地更新话术内容,提高了话术推荐的准确性和及时性。

[0005] 一种基于事件图谱的话术推荐方法,包括:

[0006] 获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别;

[0007] 爬取预设时间段内的热门事件,并对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别;

[0008] 获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果;

[0009] 从所有所述热门事件中获得与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所述热门事件记录为推荐事件;

[0010] 运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本;

[0011] 基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术。

[0012] 一种基于事件图谱的话术推荐装置,包括:

[0013] 获取模块,用于获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别;

[0014] 爬取模块,用于爬取预设时间段内的热门事件,并对所有所述热门事件进行要素

识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别;

[0015] 预测模块,用于获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果;

[0016] 匹配模块,用于从所有所述热门事件中获取与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所述热门事件记录为推荐事件;

[0017] 抽取模块,用于运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本;

[0018] 生成模块,用于基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术。

[0019] 一种计算机设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述基于事件图谱的话术推荐方法的步骤。

[0020] 一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,所述计算机程序被处理器执行时实现上述基于事件图谱的话术推荐方法的步骤。

[0021] 本发明提供的基于事件图谱的话术推荐方法、装置、设备及介质,该方法通过获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别;爬取预设时间段内的热门事件,并对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别;获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果;从所有所述热门事件中获取与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所述热门事件记录为推荐事件;运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本;基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术,如此,能够结合当前热门事件、识别出用户类别和预测出用户的推荐服务,自动匹配出推荐事件,并相应地输出推荐话术,做到与时俱进地更新话术内容,推荐与用户贴合的或者潜在的关注信息或者产品,提高了话术推荐的准确性和及时性。

附图说明

[0022] 为了更清楚地说明本发明实施例的技术方案,下面将对本发明实施例的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性的前提下,还可以根据这些附图获得其他的附图。

[0023] 图1是本发明一实施例中基于事件图谱的话术推荐方法的应用环境示意图;

[0024] 图2是本发明一实施例中基于事件图谱的话术推荐方法的流程图;

[0025] 图3是本发明一实施例中基于事件图谱的话术推荐方法的步骤S20的流程图;

[0026] 图4是本发明一实施例中基于事件图谱的话术推荐方法的步骤S30的流程图;

[0027] 图5是本发明一实施例中基于事件图谱的话术推荐方法的步骤S50的流程图;

[0028] 图6是本发明一实施例中基于事件图谱的话术推荐装置的原理框图;

[0029] 图7是本发明一实施例中基于事件图谱的话术推荐装置的爬取模块的原理框图；
[0030] 图8是本发明一实施例中计算机设备的示意图。

具体实施方式

[0031] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0032] 本发明提供的基于事件图谱的话术推荐方法,可应用在如图1的应用环境中,其中,客户端(计算机设备或终端)通过网络与服务器进行通信。其中,客户端(计算机设备或终端)包括但不限于为各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备。服务器可以是独立的服务器,也可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、内容分发网络(Content Delivery Network, CDN)、以及大数据和人工智能平台等基础云计算服务的云服务器。

[0033] 在一实施例中,如图2所示,提供一种基于事件图谱的话术推荐方法,其技术方案主要包括以下步骤S10-S60:

[0034] S10,获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别。

[0035] 可理解地,所述用户在应用程序软件中能够接收到话术推荐的消息或者产品,在用户打开所述应用程序软件时,自动触发话术推荐请求,从而获取所述用户的所述用户信息和所述行为轨迹,所述用户信息为与所述用户相关的基础信息,所述用户信息为与姓名、年龄、性别、联系电话等等各基础属性相关的信息,所述用户信息包括交际用户,所述交际用户包括与用户关联的血缘关系成员,以及与用户在应用程序中存在交互消息的成员,所述行为轨迹为所述用户在所述应用程序中的访问轨迹,所述用户画像为通过用户画像模型对所述用户信息中的各个元素进行图谱构建,构建出所述用户的用户图谱;基于该用户图谱,对所述用户图谱中的各个元素进行评分,根据评分结果进行新老用户特征识别,识别出所述用户的用户类别,所述用户类别包括新用户、断点老用户、潜力老用户、忠实老用户。

[0036] 在一实施例中,所述步骤S10中,即所述对所述用户信息进行用户画像,获得所述用户的用户类别,包括:

[0037] 对所述用户信息进行图谱构建,得到用户图谱。

[0038] 可理解地,按照预设过滤规则对所述用户信息进行实体过滤,得到可选实体,并以所述用户所对应的标识码为中心,各所述可选实体均与该中心连接,并在连接线上标明中心和与连接的所述可选实体之间的关系,各所述可选实体中为赋予一个实体值,所述实体值为根据该可选实体所对应的关系而计算出的值,例如:所述可选实体为使用应用程序软件的时长,其关系为时间关系,则根据该时长折算成相对应的分值,比如1000分钟折算成100分值,所述可选实体为购买的历史产品,其关系为买卖关系,则根据购买的历史产品的属性赋予相应的属性值,所述可选实体为该用户的交际用户,其关系为交际关系(夫妻、同事、联系频繁等等交际关系),则根据该交际用户作为中心,构建该交际用户的用户图谱,并根据该交际用户的用户图谱中的可选实体所对应的实体值进行求和,再对求和的结果进行

预设衰减函数的衰减获得该交际用户对应于该用户的标识码的实体值,通过可选实体、关系和实体值构成可选三元组,从而构建出以该用户所对应的标识码为中心的所述用户图谱。

[0039] 其中,所述预设过滤规则可以根据需求设定,比如预设过滤规则可以为从所述用户信息中筛选出该用户使用该应用程序软件的时长、购买的历史产品和与交际用户相关的实体。

[0040] 基于所述用户图谱,对所述用户进行用户画像,得到所述用户类别。

[0041] 可理解地,所述用户画像的过程为对所述用户图谱中的各可选三元组进行特征提取,根据提取的各可选三元组所对应的特征进行评分,即针对不同的关系对各实体值进行加权,并对加权后的实体值进行深度学习的评分处理,得到评分结果,再根据评分结果进行新老用户特征识别,识别出所述用户的用户类别,所述用户类别包括新用户、断点老用户、潜力老用户、忠实老用户。

[0042] 其中,深度学习的评分处理为运用深度学习的方式学习各个关系的不同实体值所对应的权重的,并根据该权重进行评分的处理过程。

[0043] 本发明实现了通过对所述用户信息进行图谱构建,得到用户图谱;基于所述用户图谱,对所述用户进行用户画像,得到所述用户类别,如此,能够通过用户图谱以及基于该用户图谱的用户画像,综合用户信息中各个有用实体进行科学地评分,自动识别出用户的用户类别,提高了用户类别识别的准确性和可靠性。

[0044] S20,爬取预设时间段内的热门事件,并对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别。

[0045] 可理解地,所述爬取预设时间段的热门事件的过程为运用网络爬取技术,从互联网的预设网站清单中获取预设时间段内的各URL信息;根据预设的URL的类别规则,从所有所述URL信息中解读出与新闻类别所对应位置相匹配的URL信息,并记录为待提取URL信息;基于所述预设时间段内收集的历史用户所搜索的热搜文本,从各所述待提取URL信息所对应的页面数据中爬取与所述热搜文本匹配的文本内容,并将匹配的文本内容标记为所述热门事件的过程,所述热门事件为当前预设时间段内更新的事件。

[0046] 其中,所述网络爬取技术是一种按照一定的规则,自动地抓取万维网信息的程序或者脚本从而获取所需的信息的技术,所述要素识别为运用文本关键要素识别算法,对各所述热门事件进行关键的要素进行识别的过程,所述文本关键要素识别算法为提取所述热门事件中的实体及实体所对应的实体类别和概念类别,根据提取出的实体及实体类别和概念类别,构成一个三元组的算法,从而可以构建出多个所述三元组;基于与各所述热门事件对应的所有所述三元组,构建各所述热门事件的事件图谱的过程;所述类别识别为通过基于图卷积网络的识别模型,对各所述事件图谱进行类别特征的聚合,得到各所述热门事件的事件类别的过程,所述事件类别包括产品、风险工具、家庭财富风险、社会风险等等。

[0047] 在一实施例中,所述步骤S20中,即所述爬取预设时间段内的热门事件,包括:

[0048] 获取预设时间段内的各URL信息。

[0049] 可理解地,所述预设时间段可以根据需求设定,比如一天、两天、一周等等,所述URL (Uniform Resource Locator,统一资源定位器) 信息为互联网上各网页的唯一的网址

信息,各所述URL信息都有发布时间,根据各所述URL信息的发布时间,筛选出所述预设时间段内的URL信息。

[0050] 基于URL的类别规则,从获取的所有所述URL信息中筛选出与新闻类别对应的URL信息,并将筛选出的所述URL信息记录为待提取URL信息。

[0051] 可理解地,所述URL的类别规则中包括了新闻类别在URL信息中所在位置的规则,基于所述URL的类别规则,从所有所述URL信息中筛选出新闻类别的URL信息,即筛选出news相关的URL信息,news是一种与新闻相关的URL类型,并将筛选后的所述URL信息确定为所述待提取URL信息。

[0052] 基于收集的热搜文本,运用网络爬取技术,爬取各所述待提取URL信息所对应的页面数据中与所述热搜文本匹配的文本内容,并将匹配的文本内容标记为所述热门事件。

[0053] 可理解地,所述热搜文本为历史用户搜索的关键词内容,或者用户点击量高的URL信息所对应的文本内容,可以通过所述热搜文本运用网络爬取技术查找出相匹配的URL信息,并截取出匹配的URL信息所对应的页面数据中的文本内容,将其标记为所述热门事件。

[0054] 其中,网络爬取技术又称为网页蜘蛛或者网络机器人,是一种按照一定的规则,自动地抓取万维网信息的程序或者脚本从而获取所需的信息的技术。

[0055] 本发明实现了通过获取预设时间段内的各URL信息;基于URL的类别规则,从获取的所有所述URL信息中筛选出与新闻类别对应的URL信息,并将筛选出的所述URL信息记录为待提取URL信息;基于收集的热搜文本,运用网络爬取技术,爬取各所述待提取URL信息所对应的页面数据中与所述热搜文本匹配的文本内容,并将匹配的文本内容标记为所述热门事件,如此,可以在海量的URL信息中爬取出有用的热门事件,去除无用的URL信息,提高了后续的话术推荐的准确性和精度。

[0056] 在一实施例中,如图3所示,所述步骤S20中,即所述对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别,包括:

[0057] S201,运用文本关键要素识别算法,对各所述热门事件进行关键要素提取,得到与各所述热门事件对应的多个三元组。

[0058] 可理解地,从所述热门事件中识别出事件的实体,即识别出热门事件中的时间、地点、人物、事件类型等元素,按照三元组格式对每个元素进行转换,得到多个元素的三元组。

[0059] 其中,所述关键要素提取为运用文本关键要素识别算法,对各所述热门事件进行关键的要素进行抽取的过程,所述文本关键要素识别算法为提取所述热门事件中的实体及实体所对应的实体类别和概念类别,根据提取出的实体及实体类别和概念类别,构成一个三元组的算法,所述三元组格式为实体、关系和属性,例如:一个热门事件中的内容包含有“今天,墨西哥地震伤亡5人”,则从该内容可以提取出的三元组有(时间、地震、当天日期)、(地点、地震、墨西哥)、(人物、地震、伤亡5人)、(事件类型、地震、自然灾害)和(事件概念、地震、意外风险),可通过地震将各三元组的关系联系起来,如此,可以对所述热门事件中的各个实体进行构建多个三元组,如此,能够将各三元组之间进行联系起来。

[0060] S202,基于与各所述热门事件对应的所有所述三元组,构建各所述热门事件的事件图谱。

[0061] 可理解地,以一个所述热门事件作为中心,从该热门事件关键抽取出的三元组连

接至该中心,以及将各三元组之间相同类型或者相关元素连接起来,并根据连接后的三元组构成该热门事件的所述事件图谱。

[0062] S203,运用图卷积网络,对各所述事件图谱进行类别识别,得到各所述热门事件的事件类别。

[0063] 可理解地,所述图卷积网络为通过深度学习技术将事件图谱中的三元组的节点(实体或属性)或者边(关系)映射为向量空间中的点,进而可以对向量空间中的点进行聚类、分类等处理的网络,将每个所述热门事件的所述事件图谱进行类别识别,所述类别识别为对输入的事件图谱中的各个三元组进行向量的聚类,从而可以根据汇聚的点的密集程度,确定出密集点的事件类别,例如:通过三元组量化后的向量进行卷积,并根据卷积后的结果进行内积,计算出各事件类别的落入的数量,并汇总出概率分布情况,从而识别出该热门事件的时间类别,进而获得各个所述热门事件的事件类别。

[0064] 本发明实现了通过运用文本关键要素识别算法,对各所述热门事件进行关键要素提取,得到与各所述热门事件对应的多个三元组;基于与各所述热门事件对应的所有所述三元组,构建各所述热门事件的事件图谱;运用图卷积网络,对各所述事件图谱进行类别识别,得到各所述热门事件的事件类别,如此,能够运用文本关键要素识别算法,识别出各个元素的三元组,并构建各个热门事件的事件图谱,通过图卷积网络能够更加准确地识别出热门事件之间的关联性,从而确定出热门事件的事件类别。

[0065] S30,获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果。

[0066] 可理解地,所述服务类型推荐模型为针对不同的用户类别而训练完成的用于基于用户的行为轨迹而聚类预测出用户所需的推荐服务,所述推荐结果体现了用户接收的各推荐服务的概率分布结果,所述推荐服务包括产品、风险工具、家庭财富风险、社会风险等,所述推荐结果中推荐服务的全集与所述事件类别的全集相同。

[0067] 其中,所述预测的过程可以为通过所述服务类型推荐模型对所述行为轨迹进行关注服务的时长分析,通过用户关注的服务的时长进行确定所述用户对关注的服务的认知程度;然后获取所述用户图谱中的交际用户,爬取与所述交际用户关联的服务事件,通过关联的服务事件可以挖掘出所述用户的潜在关注服务,通过结合用户的服务认知程度以及潜在关注服务,对所述用户的行为轨迹进行意图分析,从而分析出与所述用户对应的所述推荐结果的过程。

[0068] 在一实施例中,如图4所示,所述步骤S30中,即所述通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果,包括:

[0069] S301,通过所述服务类型推荐模型对所述行为轨迹进行关注服务的时长分析,得到所述用户的服务认知结果。

[0070] 可理解地,所述时长分析可以通过运用Mapreduce方法,计算用户在应用程序中的各个行为轨迹的访问点所停留的时长,对各个访问点的精确时长进行分析,从而衡量出用户对访问的服务相关的服务认知程度,即根据各个用用户的关注服务的准确的时长,将该准确的时长进行折算,确定出该用户对关注服务的所述服务认知结果。

[0071] 其中,所述Mapreduce方法为将复杂的、运行于大规模集群上的并行计算过程高度地抽象到了两个函数:Map和Reduce,采用“分而治之”策略,一个存储在分布式文件系统中

的大规模数据集,会被切分成许多独立的分片(split),这些分片可以被多个Map任务并行处理,再对Map任务并行处理的结果进行排序、合并等操作最后得出汇总数据的过程,即对每个用户在访问关注的服务的响应时间以每秒的日志记录进行去重,再进行每分钟的关注服务的条数进行累加,然后计算每小时的关注服务的时长汇总,最后将相同用户下的关注服务的每小时的时长进行合并的过程,其中每一步都运用Map和Reduce函数进行处理。

[0072] S302,获取所述用户图谱中的交际用户,爬取与所述交际用户关联的服务事件。

[0073] 可理解地,根据各所述交际用户,运用网络爬取技术,从互联网中爬取出各所述交际用户所关联的服务事件,所述服务事件为所述交际用户在应用程序中访问过的服务,或者已经关注了的服务。

[0074] S303,根据所述服务认知结果和所有所述服务事件,对所述行为轨迹进行意图分析,分析出与所述用户对应的所述推荐结果。

[0075] 可理解地,所述意图分析为对所述行为轨迹中与所述服务认知结果和各所述服务事件的交集识别,识别出交集的部分,并对交集的部分进行意图预测,预测出所述用户潜在关注的事件类别,所述意图分析的方式可以通过训练完成的意图识别模型对输入的服务认知结果、所述服务事件和所述行为轨迹进行向量转换并进行向量交集提取,对提取的向量交集进行卷积,从而提取出交集特征,并根据交集特征进行预测的方式处理。

[0076] 本发明实现了通过所述服务类型推荐模型对所述行为轨迹进行关注服务的时长分析,得到所述用户的服务认知结果;获取所述用户图谱中的交际用户,爬取与所述交际用户关联的服务事件;根据所述服务认知结果和所有所述服务事件,对所述行为轨迹进行意图分析,分析出与所述用户对应的所述推荐结果,如此,能够通过时长的分析以及交际用户的服务事件,侧面反映出用户后续的意图或者潜在的关注服务,可以通过意图分析出推荐结果。

[0077] S40,从所有所述热门事件中获得与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所述热门事件记录为推荐事件。

[0078] 可理解地,根据用户类别和事件类别之间关联规则,其中该关联规则为预设的规则,比如:新用户类别关联家庭财富风险类事件和风险工具的事件类别,通过与所述用户类别对应的关联规则,可以对所述用户的所述推荐结果中进行过滤,只过滤出与所述用户类别相应关联规则的事件类别,作为所述用户最终关注的推荐结果,将其作为待匹配的推荐结果,从所有所述热门事件中各所述热门事件的所述事件类别的概率分布与所述待匹配的推荐结果中的事件类别进行匹配,可以匹配出各所述热门事件与所述待匹配的推荐结果之间的匹配值。

[0079] 其中,将最大的匹配值所对应的所述热门事件作为所述推荐事件。

[0080] S50,运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本。

[0081] 可理解地,所述TextRank算法为把文本拆分成字或词作为网络节点,组成词汇网络图模型,将词语间的相似关系看成是一种推荐或投票关系,使其可以计算每一个词语的重要性的算法,所述文本抽取的过程为对所述推荐事件进行分词及虚词去除处理,得到可选分词,运用所述TextRank算法,对各所述可选分词进行词性标注,并构建可选关键词图,对所述可选关键词图进行关系抽取,将各关系抽取后的可选关键词图按照所述推荐事件中的文本顺序进行拼接,得到所述摘要文本的过程,所述摘要文本为对所述推荐事件进行内

容压缩的文本内容。

[0082] 在一实施例中,如图5所示,所述步骤S50中,即所述运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本,包括:

[0083] S501,对所述推荐事件进行分词及虚词去除处理,得到可选分词。

[0084] 可理解地,所述分词处理为对所述推荐事件中的每个字或者词语进行最小单元的划分,分出多个字或者词语的处理过程,所述虚词去除处理为对分词后的各个字或者词语进行虚词识别和/或停用词识别,将识别出的虚词和/或停用词进行去除的处理过程,将剩余的字或者词语记录为所述可选分词。

[0085] S502,运用TextRank算法,对各所述可选分词进行词性标注,并构建可选关键词图,对所述可选关键词图进行关系抽取,得到多个抽取关键词。

[0086] 可理解地,所述TextRank算法为把文本拆分成字或词作为网络节点,组成词汇网络图模型,将词语间的相似关系看成是一种推荐或投票关系,使其可以计算每一个词语的重要性的算法,运用所述TextRank算法对各所述可选分词进行词性标注,其中,词性标注为标注可选分词的词性类别,以汉语为例,汉语的词类系统有18个子类,包括7类体词,4类谓词、5类虚词、代词和感叹词等等,从而根据标注的词性标注组件词汇网络图模型,即所述可选关键词图。

[0087] 其中,所述关系抽取为对构建后的所述可选关键词图中的各可选分词进行两两之间关系的出现频次进行抽取的过程,将出现两两之间关系出现频次排序靠前的可选分词记录为所述抽取关键词。

[0088] S503,将各所述抽取关键词按照所述推荐事件中的文本顺序进行拼接,得到所述摘要文本。

[0089] 可理解地,所述拼接的过程为对各所述抽取关键词所对应的在所述推荐时间中的文本位置的顺序进行首尾拼接的过程,从而得到所述摘要文本。

[0090] 本发明实现了通过对所述推荐事件进行分词及虚词去除处理,得到可选分词;运用TextRank算法,对各所述可选分词进行词性标注,并构建可选关键词图,对所述可选关键词图进行关系抽取,得到多个抽取关键词;将各所述抽取关键词按照所述推荐事件中的文本顺序进行拼接,得到所述摘要文本,如此,通过TextRank算法,能够准确地抽取出推荐事件中的关键词,能够简要拼接出言简意赅的文本内容,减少用户阅读的厌烦度。

[0091] S60,基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术。

[0092] 可理解地,获取与所述用户类别对应的推荐模板,将所述摘要文本自动填充至所述推荐模板中的相应位置上,从而生成所述推荐话术,并在所述应用程序软件的界面上展现该推荐话术,以向所述用户推荐。

[0093] 本发明实现了通过获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别;爬取预设时间段内的热门事件,并对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别;获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果;从所有所述热门事件中获得与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所

述热门事件记录为推荐事件;运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本;基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术,如此,能够结合当前热门事件、识别出用户类别和预测出用户的推荐服务,自动匹配出推荐事件,并相应地输出推荐话术,做到与时俱进地更新话术内容,推荐与用户贴合的或者潜在的关注信息或者产品,提高了话术推荐的准确性和及时性。

[0094] 应理解,上述实施例中各步骤的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不对本发明实施例的实施过程构成任何限定。

[0095] 在一实施例中,提供一种基于事件图谱的话术推荐装置,该基于事件图谱的话术推荐装置与上述实施例中基于事件图谱的话术推荐方法一一对应。如图6所示,该基于事件图谱的话术推荐装置包括获取模块11、爬取模块12、预测模块13、匹配模块14、抽取模块15和生成模块16。各功能模块详细说明如下:

[0096] 获取模块11,用于获取用户的用户信息和行为轨迹,并对所述用户信息进行用户画像,获得所述用户的用户类别;

[0097] 爬取模块12,用于爬取预设时间段内的热门事件,并对所有所述热门事件进行要素识别,识别出各所述热门事件的事件图谱,以及对各所述事件图谱进行类别识别,得到各热门事件的事件类别;

[0098] 预测模块13,用于获取与所述用户类别对应的服务类型推荐模型,通过获取的所述服务类型推荐模型对所述行为轨迹进行预测,得到所述用户的推荐结果;

[0099] 匹配模块14,用于从所有所述热门事件中获取与所述推荐结果匹配的所述事件类别所对应的所述热门事件,并将获得的所述热门事件记录为推荐事件;

[0100] 抽取模块15,用于运用TextRank算法,对所述推荐事件进行文本抽取,得到摘要文本;

[0101] 生成模块16,用于基于所述摘要文本和与所述用户类别对应的推荐模板,生成与所述用户对应的推荐话术。

[0102] 在一实施例中,如图7所示,所述爬取模块12包括:

[0103] 第一识别单元21,用于运用文本关键要素识别算法,对各所述热门事件进行关键要素提取,得到与各所述热门事件对应的多个三元组;

[0104] 构建单元22,用于基于与各所述热门事件对应的所有所述三元组,构建各所述热门事件的事件图谱;

[0105] 第二识别单元23,用于运用图卷积网络,对各所述事件图谱进行类别识别,得到各所述热门事件的事件类别。

[0106] 关于基于事件图谱的话术推荐装置的具体限定可以参见上文中对于基于事件图谱的话术推荐方法的限定,在此不再赘述。上述基于事件图谱的话术推荐装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中,也可以以软件形式存储于计算机设备中的存储器中,以便于处理器调用执行以上各个模块对应的操作。

[0107] 在一个实施例中,提供了一种计算机设备,该计算机设备可以是客户端或者服务端,其内部结构图可以如图8所示。该计算机设备包括通过系统总线连接的处理器、存储器、

网络接口和数据库。其中,该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括可读存储介质、内存储器。该可读存储介质存储有操作系统、计算机程序和数据库。该内存储器为可读存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种基于事件图谱的话术推荐方法。

[0108] 在一个实施例中,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,处理器执行计算机程序时实现上述实施例中基于事件图谱的话术推荐方法。

[0109] 在一个实施例中,提供了一种计算机可读存储介质,其上存储有计算机程序,计算机程序被处理器执行时实现上述实施例中基于事件图谱的话术推荐方法。

[0110] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读取存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本发明所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限,RAM以多种形式可得,诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双数据率SDRAM(DDRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

[0111] 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,仅以上述各功能单元、模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能单元、模块完成,即将所述装置的内部结构划分成不同的功能单元或模块,以完成以上描述的全部或者部分功能。

[0112] 以上所述实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围,均应包含在本发明的保护范围之内。

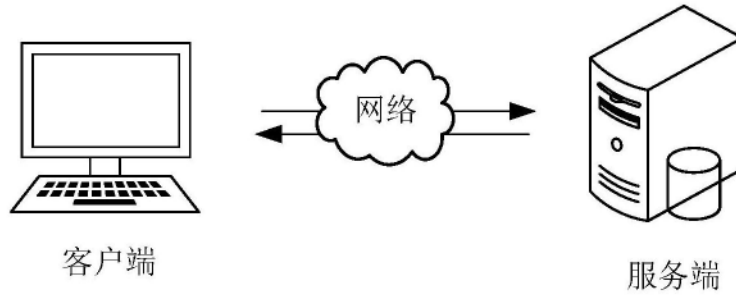


图1



图2

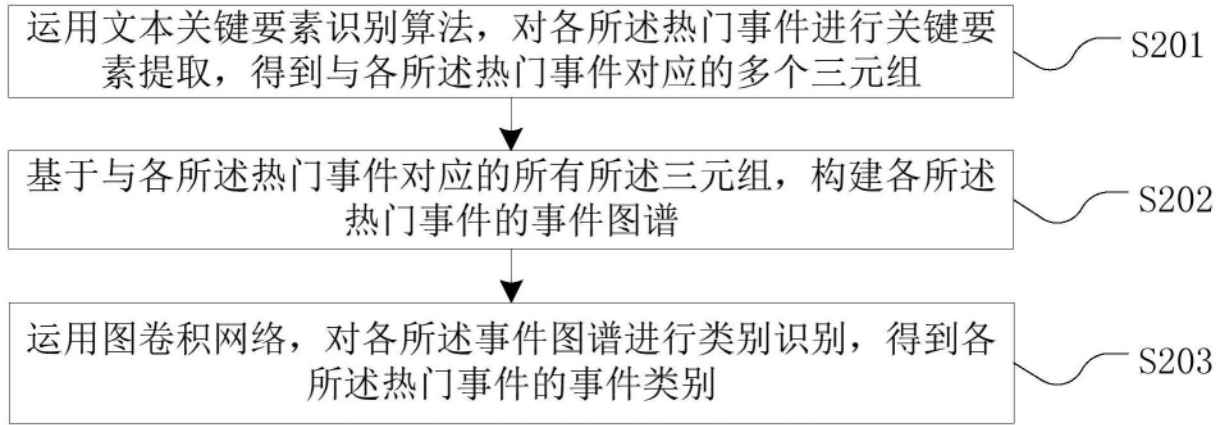


图3

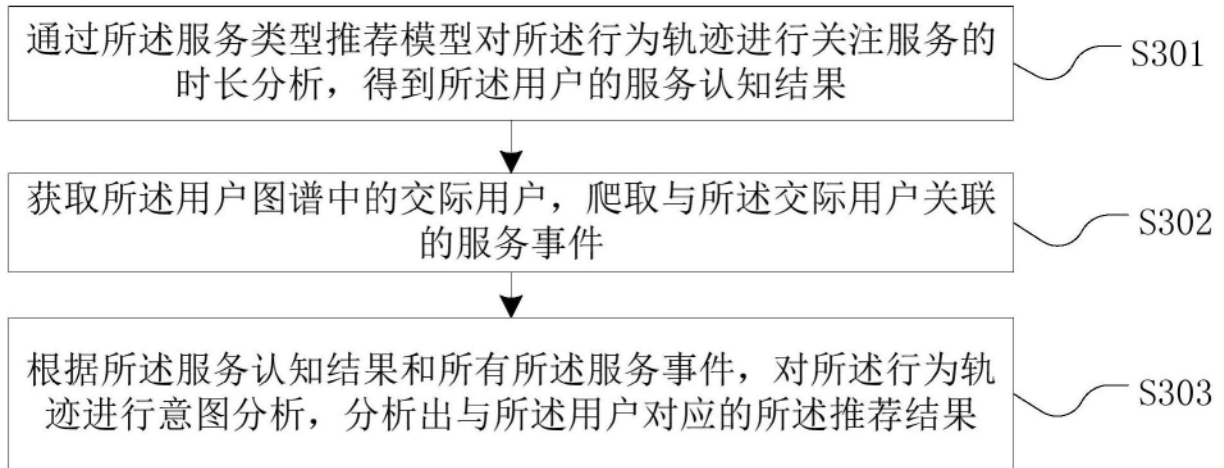


图4

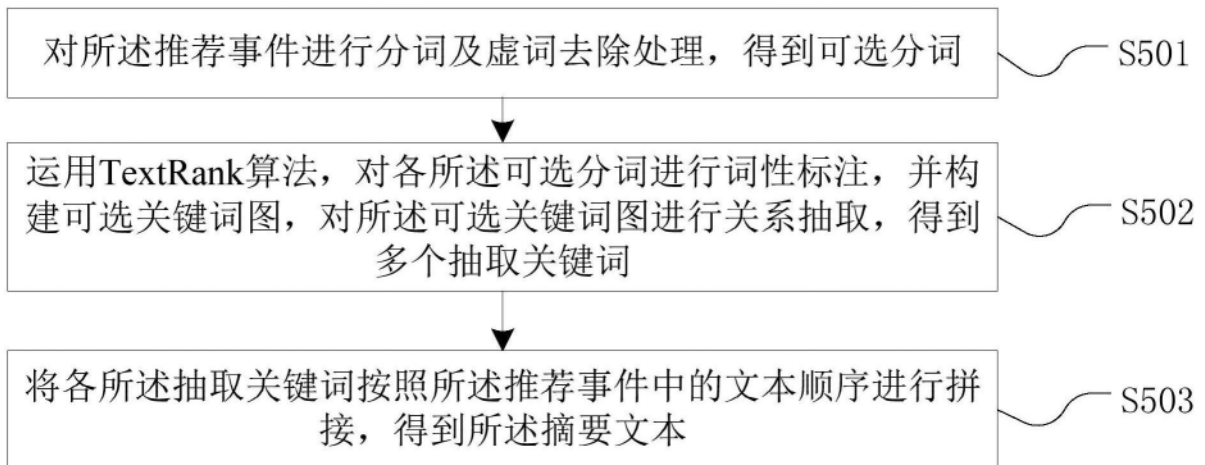


图5

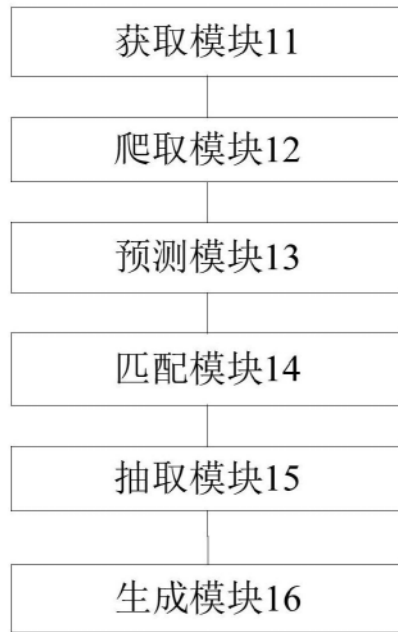


图6



图7

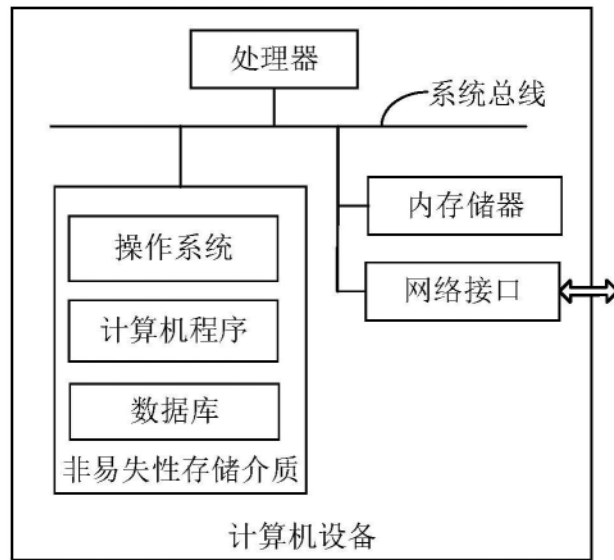


图8