



(12) 发明专利申请

(10) 申请公布号 CN 103914550 A

(43) 申请公布日 2014. 07. 09

(21) 申请号 201410146060. 4

(22) 申请日 2014. 04. 11

(71) 申请人 百度在线网络技术(北京)有限公司  
地址 100085 北京市海淀区上地十街 10 号  
百度大厦三层

(72) 发明人 陈炜于 刘四维 欧阳显雅  
柳金杜

(74) 专利代理机构 北京清亦华知识产权代理事  
务所(普通合伙) 11201  
代理人 宋合成

(51) Int. Cl.  
G06F 17/30(2006. 01)

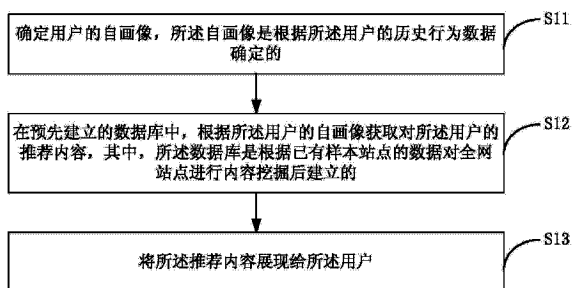
权利要求书2页 说明书11页 附图3页

(54) 发明名称

展现推荐内容的方法和装置

(57) 摘要

本发明提出一种展现推荐内容的方法和装置,该方法包括确定用户的自画像,所述自画像是根据所述用户的历史行为数据确定的;在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的;将所述推荐内容展现给所述用户。该方法能够扩大能够提供的内容的范围。



1. 一种展现推荐内容的方法,其特征在于,包括:  
确定用户的自画像,所述自画像是根据所述用户的历史行为数据确定的;  
在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的;  
将所述推荐内容展现给所述用户。
2. 根据权利要求1所述的方法,其特征在于,所述确定用户的自画像,包括:  
获取用户的历史行为数据,并根据所述用户的历史行为数据确定所述用户的自画像,所述用户的历史行为数据包括:用户在站外的历史行为数据,和/或,用户在站内的历史行为数据。
3. 根据权利要求2所述的方法,其特征在于,所述获取用户的历史行为数据,包括:  
根据预先设置的存储在用户本地终端上的数据 cookie,获取用户的历史行为数据;和/或,  
根据用户的账户以及日志中记录的与所述账户对应的历史行为数据,获取用户的历史行为数据。
4. 根据权利要求2所述的方法,其特征在于,所述历史行为数据包括如下项中的至少一项:  
已浏览内容,已选择内容,已注册信息。
5. 根据权利要求1所述的方法,其特征在于,还包括:  
实时获取用户的当前行为数据;  
根据所述当前行为数据在所述数据库中重新获取推荐内容;  
将所述重新获取的推荐内容实时展现给所述用户。
6. 根据权利要求1至5任一项所述的方法,其特征在于,还包括:建立所述数据库,所述建立所述数据库包括:  
确定已知样本站点;  
获取所述已知样本站点中的内容标签;  
根据所述内容标签对全网站点进行内容挖掘,获取包含所述内容标签的其他站点;  
将所述已知样本站点的内容以及所述其他站点的内容抓取保存到所述数据库中。
7. 根据权利要求6所述的方法,其特征在于,所述确定已知样本站点,包括:  
对垂直领域进行行业统计,确定已知样本站点;或者,  
根据配置信息,确定已知样本站点;或者,  
根据有配置行业的软件或服务,确定已知样本站点。
8. 根据权利要求6所述的方法,其特征在于,所述全网站点包括预先收录的多个站点。
9. 一种展现推荐内容的装置,其特征在于,包括:  
确定模块,用于确定用户的自画像,所述自画像是根据所述用户的历史行为数据确定的;  
获取模块,用于在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的;  
展现模块,用于将所述推荐内容展现给所述用户。

10. 根据权利要求 9 所述的装置,其特征在于,所述确定模块具体用于:

获取用户的历史行为数据,并根据所述用户的历史行为数据确定所述用户的自画像,所述用户的历史行为数据包括:用户在站外的历史行为数据,和/或,用户在站内的历史行为数据。

11. 根据权利要求 10 所述的装置,其特征在于,所述确定模块进一步具体用于:

根据预先设置的存储在用户本地终端上的数据 cookie,获取用户的历史行为数据;和/或,

根据用户的账户以及日志中记录的与所述账户对应的历史行为数据,获取用户的历史行为数据。

12. 根据权利要求 10 所述的装置,其特征在于,所述确定模块获取的所述历史行为数据包括如下项中的至少一项:

已浏览内容,已选择内容,已注册信息。

13. 根据权利要求 9 所述的装置,其特征在于,还包括:

更新模块,用于实时获取用户的当前行为数据;根据所述当前行为数据在所述数据库中重新获取推荐内容;将所述重新获取的推荐内容实时展现给所述用户。

14. 根据权利要求 9 至 13 任一项所述的装置,其特征在于,还包括:用于建立所述数据库的建立模块,所述建立模块具体用于:

确定已知样本站点;

获取所述已知样本站点中的内容标签;

根据所述内容标签对全网站点进行内容挖掘,获取包含所述内容标签的其他站点;

将所述已知样本站点的内容以及所述其他站点的内容抓取保存到所述数据库中。

15. 根据权利要求 14 所述的装置,其特征在于,所述建立模块进一步具体用于:

对垂直领域进行行业统计,确定已知样本站点;或者,

根据配置信息,确定已知样本站点;或者,

根据有配置行业的软件或服务,确定已知样本站点。

16. 根据权利要求 14 所述的装置,其特征在于,所述建立模块进一步具体用于:预先收录的多个站点确定为所述全网站点。

## 展现推荐内容的方法和装置

### 技术领域

[0001] 本发明涉及通信技术领域,尤其涉及一种展现推荐内容的方法和装置。

### 背景技术

[0002] 随着互联网的发展和普及,互联网中的信息飞速增长。在互联网海量信息中,用户通常需要很长时间才能找到需要的信息。为了方便用户获取信息,垂直领域的内容推荐应运而生。垂直领域的内容推荐是指对信息进行分类,向用户提供某类专门的信息,例如女性资讯、母婴资讯、健康资讯、IT 资讯等不同的垂直领域。

[0003] 目前,垂直领域中提供的内容通常是由编辑人员编辑生成的,能够提供的内容不够全面和丰富。

### 发明内容

[0004] 本发明旨在至少在一定程度上解决相关技术中的技术问题之一。

[0005] 为此,本发明的一个目的在于提出一种展现推荐内容的方法,该方法可以扩大能够提供的内容的范围。

[0006] 本发明的另一个目的在于提出一种展现推荐内容的装置。

[0007] 为达到上述目的,本发明第一方面实施例提出的展现推荐内容的方法,包括:确定用户的自画像,所述自画像是根据所述用户的历史行为数据确定的;在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的;将所述推荐内容展现给所述用户。

[0008] 本发明第一方面实施例提出的展现推荐内容的方法,通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验。

[0009] 为达到上述目的,本发明第二方面实施例提出的展现推荐内容的装置,包括:确定模块,用于确定用户的自画像,所述自画像是根据所述用户的历史行为数据确定的;获取模块,用于在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的;展现模块,用于将所述推荐内容展现给所述用户。

[0010] 本发明第二方面实施例提出的展现推荐内容的装置,通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验。

[0011] 为达到上述目的,本发明第三方面实施例提出的客户端设备,包括:壳体、处理器、存储器、电路板和电源电路,其中,电路板安置在壳体围成的空间内部,处理器和存储器设置在电路板上;电源电路,用于为客户端设备的各个电路或器件供电;存储器用于存储可

执行程序代码；处理器通过读取存储器中存储的可执行程序代码来运行与可执行程序代码对应的程序，以用于执行以下步骤：确定用户的自画像，所述自画像是根据所述用户的历史行为数据确定的；在预先建立的数据库中，根据所述用户的自画像获取对所述用户的推荐内容，其中，所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的；将所述推荐内容展现给所述用户。

[0012] 本发明第三方面实施例提出的客户端设备，通过根据已知样本站点对全网站点进行内容挖掘，可以将不同站点的内容聚合到当前站点，扩大数据库中内容的范围，为用户提供更全面丰富的数据；通过根据自画像获取相应内容，可以为用户提供个性化内容，提升用户体验。

[0013] 本发明附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本发明的实践了解到。

### 附图说明

[0014] 本发明上述的和 / 或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

[0015] 图 1 是本发明一实施例提出的展现推荐内容的方法的流程示意图；

[0016] 图 2 是本发明实施例中一种展现结果的示意图；

[0017] 图 3 是本发明实施例中另一种展现结果的示意图；

[0018] 图 4 是本发明另一实施例提出的展现推荐内容的方法的流程示意图；

[0019] 图 5 是本发明另一实施例提出的展现推荐内容的装置的结构示意图；

[0020] 图 6 为本发明另一实施例提出的展现推荐内容的装置的结构示意图。

### 具体实施方式

[0021] 下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，仅用于解释本发明，而不能理解为对本发明的限制。相反，本发明的实施例包括落入所附加权利要求书的精神和内涵范围内的所有变化、修改和等同物。

[0022] 图 1 是本发明一实施例提出的展现推荐内容的方法的流程示意图，该方法包括：

[0023] S11：确定用户的自画像，所述自画像是根据所述用户的历史行为数据确定的。

[0024] 其中，用户的自画像可以表明用户的行为特征，例如，用户搜索过什么内容，用户浏览过什么内容等。

[0025] 通过用户的自画像可以为用户提供更个性化的内容。

[0026] 用户的自画像可以根据用户在站外的历史行为数据确定，和 / 或，用户的自画像也可以根据用户在站内的历史行为数据确定。

[0027] 例如，用户登录第一网站后，如果用户在第一网站中并没有历史行为，本实施例可以根据用户在其他网站，例如第二网站中的历史行为确定自画像。和 / 或，用户在第一网站中有历史行为时，也可以根据用户在第一网站中的历史行为确定自画像。

[0028] 进一步的，用户的历史行为数据可以根据预先设置的 cookie，也就是存储在用户

本地终端上的数据获取,和 / 或,根据用户的账户从后台服务器记录的日志中获取。

[0029] 例如,当用户没有采用账户登录时,可以根据预先设置的 cookie 获取用户在站外和 / 或站内的历史行为数据,当用户采用账户登录时,可以根据账户获取用户在站外和 / 或站内的历史行为数据。具体的,预先设置的 cookie 是提供第三方服务的 cookie,通过该预先设置的 cookie 可以获取用户在该第三方服务覆盖的站点以及主站的用户行为数据,例如,对于百度,可以设置百度 cookie,之后可以获取用户在百度站点以及提供百度服务的其他站点上的历史行为数据。

[0030] 在后台服务器中可以保存日志,该日志中可以记录用户账户与历史行为数据的对应关系,因此,当用户采用账户登录一网站后,可以从日志中获取与该账户对应的站外和 / 或站内的历史行为数据。

[0031] 历史行为数据可以包括如下项中的至少一项:已浏览内容,已选择内容,已注册信息。其中,已选择的内容可以是向用户展现多个内容,由用户在其中选择感兴趣的内容,或者,向用户展现多个问题,用户回答这些问题等;已注册信息可以是在注册时要求用户填写的年龄、性别、职业等信息。

[0032] S12:在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的。

[0033] 例如,用户的自画像表明用户对“美容”感兴趣,那么可以在数据库中获取与“美容”相关的内容作为推荐内容。

[0034] 与现有技术通过编辑得到垂直领域内容不同的是,本实施例获取的数据是根据已有样本站点对全网站点进行内容挖掘后得到的。例如,已有样本站点是第一站点,本实施例中,不仅可以将第一站点的内容抓取到数据库中,还可以根据第一站点的内容抓取全网其它站点中的内容,例如,从第一站点中获取相关标签(tag),例如,“美容”、“瘦身”等,之后采用这些标签对全网其它站点进行内容挖掘,例如其它站点中的第二站点中也有这些标签,那么可以将第二站点中的内容也抓取到数据库中。

[0035] 上述的全网站点可以是预先收录的多个站点,例如,已知样本站点是女性相关站点,而收录的全网站点不限于女性相关站点,也可以包括军事、IT 等其他垂直领域,或者,也不限于垂直领域,也可以是包括多种内容的站点。

[0036] S13:将所述推荐内容展现给所述用户。

[0037] 由于不同用户感兴趣的内容可能不同,对应的历史行为数据也就不同,相应的自画像也就不同,因此,根据用户的自画像可以向不同用户展现不同的推荐内容,满足用户个性化需求。

[0038] 例如,假设第一用户对“化妆品”感兴趣,展现的推荐内容可以如图 2 所示,假设第二用户对“婚纱”感兴趣,展现的推荐内容可以如图 3 所示。

[0039] 本实施例通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验。

[0040] 图 4 是本发明另一实施例提出的展现推荐内容的方法的流程示意图,该方法包括:

[0041] S41:建立垂直领域的站点的数据库。

[0042] 其中,该数据库采用内容挖掘的方式对全网站点的内容进行内容聚合得到。

[0043] 例如,该垂直领域的站点是女性资讯的站点,可以对已知样本站点,例如网易的女性频道,从这些网站中挖掘出一些标签,例如,美容,瘦身等,之后再用这些标签去全网的其他站点进行内容挖掘,当某一站点也包含美容、瘦身这些标签后,将该站点的内容也加到数据库中。

[0044] 另外,上述的已知样本站点可以是预先配置的,例如,配置网易的女性频道;或者,上述的已知样本站点也可以是进行行业统计后确定的,例如,通过统计女性用户对站点的搜索和访问等情况,可以统计得到女性资讯领域的大站,之后将统计得到的站点作为已知样本站点,又或者,可以通过某个有配置行业的软件或者服务得到已知样本站点,例如百度统计,HA0123等。

[0045] 再者,上述的全网站点可以是预先收录的,例如百度收录库中的站点,由于百度收录的数据量很大,可以提高覆盖面,进而可以提高获取内容的范围。

[0046] S42:用户登录到该垂直领域的站点。

[0047] 其中,该登录可以是指用户输入账户密码登录,或者,登录也可以是指用户不输入账户密码,仅是访问该站点。

[0048] S43:确定该用户的自画像。

[0049] 其中,假设用户登录的该垂直领域的站点称为第一站点,当用户在该第一站点中没有历史行为数据时,例如,首次登录该第一站点时,可以根据预先设置的 cookie 获取用户在该第一站点之外的历史行为数据,例如获取用户在第二站点的历史行为数据,其中,预先设置的 cookie 是提供第三方服务的 cookie,通过该预先设置的 cookie 可以获取用户在该第三方服务覆盖的站点以及主站的用户行为数据,例如,对于百度,可以设置百度 cookie,之后可以获取用户在百度站点以及提供百度服务的其他站点上的用户行为数据。或者,用户采用账户登录时,例如,用户的账户为第一账户,那么可以从日志中获取该第一账户对应的站外历史行为数据,之后可以根据获取的站外历史行为数据确定自画像。

[0050] S44:根据用户的自画像,在数据库中获取推荐内容。

[0051] 例如,根据用户的自画像可以获知用户感兴趣的内容,例如,用户的历史行为数据表明搜索或浏览的大部分是美容相关的内容,那么可以在数据库中获取美容方面的内容;又或者,用户的历史行为数据表明对瘦身感兴趣,那么可以在数据库中获取瘦身方面的内容。

[0052] S45:将推荐内容展现给用户。

[0053] 例如,一个用户对美容感兴趣,那么可以将美容相关的内容展现给该用户,而另一用户对瘦身感兴趣,那么可以将瘦身相关的内容展现给该另一用户。

[0054] 进一步的,展现的内容可以实时更新。即,本实施例的方法还可以包括:

[0055] S46:实时获取用户的当前行为数据。

[0056] 例如,通过历史行为数据向用户展现美容相关内容后,用户可能不再浏览美容相关内容,还可能浏览其他内容,例如浏览瘦身相关内容。可以理解的是,上述向用户展现的内容不限于用户感兴趣的内容,还可以展现其他内容,因此,用户还可以浏览其他内容。例如,用户对美容感兴趣,网页上可以大部分展现美容相关内容,另外的小部分,如“您还可能感兴趣的内容”部分展现其他内容,这样就可以实现用户对其他内容的浏览。

[0057] S47:根据所述当前行为数据在所述数据库中重新获取推荐内容。

[0058] 例如,用户实时浏览的是瘦身内容,那么可以重新从数据库中获取瘦身相关内容,而不再是美容相关内容。

[0059] S48:将所述重新获取的推荐内容实时展现给所述用户。

[0060] 例如,可以将展现的美容相关内容实时更新为瘦身相关内容,以满足用户当前个性化需要。

[0061] 本实施例通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验;本实施例通过行业统计得到已知样本站点,可以扩大样本数量,进而扩大数据库中的内容,为用户提供更丰富的信息;本实施例通过实时更新展现的内容,可以满足用户实时的个性化需求,提升用户体验。

[0062] 图5是本发明另一实施例提出的展现推荐内容的装置的结构示意图,该装置50包括确定模块51、获取模块52和展现模块53。

[0063] 确定模块51用于确定用户的自画像,所述自画像是根据所述用户的历史行为数据确定的;

[0064] 其中,用户的自画像可以表明用户的行为特征,例如,用户搜索过什么内容,用户浏览过什么内容等。

[0065] 通过用户的自画像可以为提供更个性化的内容。

[0066] 用户的自画像可以根据用户在站外的历史行为数据确定,和/或,用户的自画像也可以根据用户在站内的历史行为数据确定。

[0067] 例如,用户登录第一网站后,如果用户在第一网站中并没有历史行为,本实施例可以根据用户在其他网站,例如第二网站中的历史行为确定自画像。和/或,用户在第一网站中有历史行为时,也可以根据用户在第一网站中的历史行为确定自画像。

[0068] 进一步的,用户的历史行为数据可以根据预先设置的 cookie,也就是存储在用户本地终端上的数据获取,和/或,根据用户的账户从后台服务器记录的日志中获取。

[0069] 例如,当用户没有采用账户登录时,可以根据预先设置的 cookie 获取用户在站外和/或站内的历史行为数据,当用户采用账户登录时,可以根据账户获取用户在站外和/或站内的历史行为数据。

[0070] 具体的,预先设置的 cookie 是提供第三方服务的 cookie,通过该预先设置的 cookie 可以获取用户在该第三方服务覆盖的站点以及主站的用户行为数据,例如,对于百度,可以设置百度 cookie,之后可以获取用户在百度站点以及提供百度服务的其他站点上的历史行为数据。

[0071] 在后台服务器中可以保存日志,该日志中可以记录用户账户与历史行为数据的对应关系,因此,当用户采用账户登录一网站后,可以从日志中获取与该账户对应的站外和/或站内历史行为数据。

[0072] 历史行为数据可以包括如下项中的至少一项:已浏览内容,已选择内容,已注册信息。其中,已选择的内容可以是向用户展现多个内容,由用户在其中选择感兴趣的内容,或者,向用户展现多个问题,用户回答这些问题等;已注册信息可以是在注册时要求用户填写的年龄、性别、职业等信息。



[0073] 一个实施例中,所述确定模块具体用于:

[0074] 获取用户的历史行为数据,并根据所述用户的历史行为数据确定所述用户的自画像,所述用户的历史行为数据包括:用户在站外的历史行为数据,和/或,用户在站内的历史行为数据。

[0075] 一个实施例中,所述确定模块进一步具体用于:根据预先设置的存储在用户本地终端上的数据 cookie,获取用户的历史行为数据;和/或,根据用户的账户以及日志中记录的与所述账户对应的历史行为数据,获取用户的历史行为数据。

[0076] 一个实施例中,所述确定模块获取的所述历史行为数据包括如下项中的至少一项:已浏览内容,已选择内容,已注册信息。

[0077] 获取模块 52 用于在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的。

[0078] 例如,用户的自画像表明用户对“美容”感兴趣,那么可以在数据库中获取与“美容”相关的内容作为推荐内容。

[0079] 与现有技术通过编辑得到垂直领域内容不同的是,本实施例获取的数据是根据已有样本站点对全网站点进行内容挖掘后得到的。例如,已有样本站点是第一站点,本实施例中,不仅可以将第一站点的内容抓取到数据库中,还可以根据第一站点的内容抓取全网其它站点中的内容,例如,从第一站点中获取相关标签(tag),例如,“美容”、“瘦身”等,之后采用这些标签对全网其它站点进行内容挖掘,例如其它站点中的第二站点中也有这些标签,那么可以将第二站点中的内容也抓取到数据库中。

[0080] 上述的全网站点可以是预先收录的多个站点,例如,已知样本站点是女性相关站点,而收录的全网站点不限于女性相关站点,也可以包括军事、IT 等其他垂直领域,或者,也不限于垂直领域,也可以是包括多种内容的站点。

[0081] 展现模块 53 用于将所述推荐内容展现给所述用户。

[0082] 由于不同用户感兴趣的内容可能不同,对应的历史行为数据也就不同,相应的自画像也就不同,因此,根据用户的自画像可以向不同用户展现不同的推荐内容,满足用户个性化需求。

[0083] 例如,假设第一用户对“化妆品”感兴趣,展现的推荐内容可以如图 2 所示,假设第二用户对“婚纱”感兴趣,展现的推荐内容可以如图 3 所示。

[0084] 本实施例通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验。

[0085] 图 6 为本发明另一实施例提出的展现推荐内容的装置的结构示意图,该装置 50 还包括更新模块 54。

[0086] 更新模块 54 用于实时获取用户的当前行为数据;根据所述当前行为数据在所述数据库中重新获取推荐内容;将所述重新获取的推荐内容实时展现给所述用户。

[0087] 例如,通过历史行为数据向用户展现美容相关内容后,用户可能不再浏览美容相关内容,还可能浏览其他内容,例如浏览瘦身相关内容。可以理解的是,上述向用户展现的内容不限于用户感兴趣的内容,还可以展现其他内容,因此,用户还可以浏览其他内容。例

如,用户对美容感兴趣,网页上可以大部分展现美容相关内容,另外的小部分,如“您可能感兴趣的内容”部分展现其他内容,这样就可以实现用户对其他内容的浏览。

[0088] 例如,用户实时浏览的是瘦身内容,那么可以重新从数据库中获取瘦身相关内容,而不再是美容相关内容。之后,可以将展现的美容相关内容实时更新为瘦身相关内容,以满足用户当前个性化需要。

[0089] 一个实施例中,该装置还包括:用于建立所述数据库的建立模块 55,所述建立模块 55 具体用于:确定已知样本站点;获取所述已知样本站点中的内容标签;根据所述内容标签对全网站点进行内容挖掘,获取包含所述内容标签的其他站点;将所述已知样本站点的内容以及所述其他站点的内容抓取保存到所述数据库中。

[0090] 其中,该数据库采用内容挖掘的方式对全网站点的内容进行内容聚合得到。

[0091] 例如,该垂直领域的站点是女性资讯的站点,可以对已知样本站点,例如网易的女性频道,从这些网站中挖掘出一些标签,例如,美容,瘦身等,之后再这些标签去全网的其他站点进行内容挖掘,当某一站点也包含美容、瘦身这些标签后,将该站点的内容也加到数据库中。

[0092] 一个实施例中,所述建立模块 55 进一步具体用于:对垂直领域进行行业统计,确定已知样本站点;或者,根据配置信息,确定已知样本站点;或者,根据有配置行业的软件或服务,确定已知样本站点。

[0093] 例如,配置网易的女性频道;或者,上述的已知样本站点也可以是进行行业统计后确定的,例如,通过统计用户对女性站点的搜索和访问等情况,可以统计得到女性资讯领域的大站,之后将统计得到的站点作为已知样本站点,具体可以用百度统计获取各个垂直领域的已知样本站点。

[0094] 一个实施例中,所述建立模块 55 进一步具体用于:预先收录的多个站点确定为所述全网站点。

[0095] 例如,上述的全网站点可以是预先收录的,例如百度收录库中的站点,由于百度收录的数据量很大,可以提高覆盖面,进而可以提高获取内容的范围。

[0096] 本实施例通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验;本实施例通过行业统计得到已知样本站点,可以扩大样本数量,进而扩大数据库中的内容,为用户提供更丰富的信息;本实施例通过实时更新展现的内容,可以满足用户实时的个性化需求,提升用户体验。

[0097] 本发明实施例还提供了一种客户端设备,该客户端设备包括壳体、处理器、存储器、电路板和电源电路,其中,电路板安置在壳体围成的空间内部,处理器和存储器设置在电路板上;电源电路,用于为客户端设备的各个电路或器件供电;存储器用于存储可执行程序代码;处理器通过读取存储器中存储的可执行程序代码来运行与可执行程序代码对应的程序,以用于执行以下步骤:

[0098] S11':确定用户的自画像,所述自画像是根据所述用户的历史行为数据确定的。

[0099] 其中,用户的自画像可以表明用户的行为特征,例如,用户搜索过什么内容,用户浏览过什么内容等。

[0100] 通过用户的自画像可以为用户提供更个性化的内容。

[0101] 用户的自画像可以根据用户在站外的历史行为数据确定,和 / 或,用户的自画像也可以根据用户在站内的历史行为数据确定。

[0102] 例如,用户登录第一网站后,如果用户在第一网站中并没有历史行为,本实施例可以根据用户在其他网站,例如第二网站中的历史行为确定自画像。和 / 或,用户在第一网站中有历史行为时,也可以根据用户在第一网站中的历史行为确定自画像。

[0103] 进一步的,用户的历史行为数据可以根据预先设置的 cookie,也就是存储在用户本地终端上的数据获取,和 / 或,根据用户的账户从后台服务器记录的日志中获取。

[0104] 例如,当用户没有采用账户登录时,可以根据预先设置的 cookie 获取用户在站外和 / 或站内的历史行为数据,当用户采用账户登录时,可以根据账户获取用户在站外和 / 或站内的历史行为数据。具体的,预先设置的 cookie 是提供第三方服务的 cookie,通过该预先设置的 cookie 可以获取用户在该第三方服务覆盖的站点以及主站的用户行为数据,例如,对于百度,可以设置百度 cookie,之后可以获取用户在百度站点以及提供百度服务的其他站点上的历史行为数据。

[0105] 在后台服务器中可以保存日志,该日志中可以记录用户账户与历史行为数据的对应关系,因此,当用户采用账户登录一网站后,可以从日志中获取与该账户对应的站外和 / 或站内的历史行为数据。

[0106] 历史行为数据可以包括如下项中的至少一项:已浏览内容,已选择内容,已注册信息。其中,已选择的内容可以是向用户展现多个内容,由用户在其中选择感兴趣的内容,或者,向用户展现多个问题,用户回答这些问题等;已注册信息可以是在注册时要求用户填写的年龄、性别、职业等信息。

[0107] S12':在预先建立的数据库中,根据所述用户的自画像获取对所述用户的推荐内容,其中,所述数据库是根据已有样本站点的数据对全网站点进行内容挖掘后建立的。

[0108] 例如,用户的自画像表明用户对“美容”感兴趣,那么可以在数据库中获取与“美容”相关的内容作为推荐内容。

[0109] 与现有技术通过编辑得到垂直领域的内容不同的是,本实施例获取的数据是根据已有样本站点对全网站点进行内容挖掘后得到的。例如,已有样本站点是第一站点,本实施例中,不仅可以将第一站点的内容抓取到数据库中,还可以根据第一站点的内容抓取全网其它站点中的内容,例如,从第一站点中获取相关标签(tag),例如,“美容”、“瘦身”等,之后采用这些标签对全网其它站点进行内容挖掘,例如其它站点中的第二站点中也有这些标签,那么可以将第二站点中的内容也抓取到数据库中。

[0110] 上述的全网站点可以是预先收录的多个站点,例如,已知样本站点是女性相关站点,而收录的全网站点不限于女性相关站点,也可以包括军事、IT 等其他垂直领域,或者,也不限于垂直领域,也可以是包括多种内容的站点。

[0111] S13':将所述推荐内容展现给所述用户。

[0112] 由于不同用户感兴趣的内容可能不同,对应的历史行为数据也就不同,相应的自画像也就不同,因此,根据用户的自画像可以向不同用户展现不同的推荐内容,满足用户个性化需求。

[0113] 例如,假设第一用户对“化妆品”感兴趣,展现的推荐内容可以如图 2 所示,假设第二用户对“婚纱”感兴趣,展现的推荐内容可以如图 3 所示。

[0114] 本实施例通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验。

[0115] 另一实施例中,处理器通过读取存储器中存储的可执行程序代码来运行与可执行程序代码对应的程序,以用于执行以下步骤:

[0116] S41':建立垂直领域的站点的数据库。

[0117] 其中,该数据库采用内容挖掘的方式对全网站点的内容进行内容聚合得到。

[0118] 例如,该垂直领域的站点是女性资讯的站点,可以对已知样本站点,例如网易的女性频道,从这些网站中挖掘出一些标签,例如,美容,瘦身等,之后再这些标签去全网的其他站点进行内容挖掘,当某一站点也包含美容、瘦身这些标签后,将该站点的内容也加到数据库中。

[0119] 另外,上述的已知样本站点可以是预先配置的,例如,配置网易的女性频道;或者,上述的已知样本站点也可以是进行行业统计后确定的,例如,通过统计女性用户对站点的搜索和访问等情况,可以统计得到女性资讯领域的大站,之后将统计得到的站点作为已知样本站点,又或者,可以通过某个有配置行业的软件或者服务得到已知样本站点,例如百度统计,HA0123等。

[0120] 再者,上述的全网站点可以是预先收录的,例如百度收录库中的站点,由于百度收录的数据量很大,可以提高覆盖面,进而可以提高获取内容的范围。

[0121] S42':用户登录到该垂直领域的站点。

[0122] 其中,该登录可以是指用户输入账户密码登录,或者,登录也可以是指用户不输入账户密码,仅是访问该站点。

[0123] S43':确定该用户的自画像。

[0124] 其中,假设用户登录的该垂直领域的站点称为第一站点,当用户在该第一站点中没有历史行为数据时,例如,首次登录该第一站点时,可以根据预先设置的 cookie 获取用户在该第一站点之外的历史行为数据,例如获取用户在第二站点的历史行为数据,其中,预先设置的 cookie 是提供第三方服务的 cookie,通过该预先设置的 cookie 可以获取用户在该第三方服务覆盖的站点以及主站的用户行为数据,例如,对于百度,可以设置百度 cookie,之后可以获取用户在百度站点以及提供百度服务的其他站点上的用户行为数据。或者,用户采用账户登录时,例如,用户的账户为第一账户,那么可以从日志中获取该第一账户对应的站外历史行为数据,之后可以根据获取的站外历史行为数据确定自画像。

[0125] S44':根据用户的自画像,在数据库中获取推荐内容。

[0126] 例如,根据用户的自画像可以获知用户感兴趣的内容,例如,用户的历史行为数据表明搜索或浏览的大部分是美容相关的内容,那么可以在数据库中获取美容方面的内容;又或者,用户的历史行为数据表明对瘦身感兴趣,那么可以在数据库中获取瘦身方面的内容。

[0127] S45':将推荐内容展现给用户。

[0128] 例如,一个用户对美容感兴趣,那么可以将美容相关的内容展现给该用户,而另一用户对瘦身感兴趣,那么可以将瘦身相关的内容展现给该另一用户。

[0129] 进一步的,展现的内容可以实时更新。即,本实施例的方法还可以包括:

[0130] S46':实时获取用户的当前行为数据:。

[0131] 例如,通过历史行为数据向用户展现美容相关内容后,用户可能不再浏览美容相关内容,还可能浏览其他内容,例如浏览瘦身相关内容。可以理解的是,上述向用户展现的内容不限于用户感兴趣的内容,还可以展现其他内容,因此,用户还可以浏览其他内容。例如,用户对美容感兴趣,网页上可以大部分展现美容相关内容,另外的小部分,如“您还可能感兴趣的内容”部分展现其他内容,这样就可以实现用户对其他内容的浏览。

[0132] S47':根据所述当前行为数据在所述数据库中重新获取推荐内容。

[0133] 例如,用户实时浏览的是瘦身内容,那么可以重新从数据库中获取瘦身相关内容,而不再是美容相关内容。

[0134] S48':将所述重新获取的推荐内容实时展现给所述用户。

[0135] 例如,可以将展现的美容相关内容实时更新为瘦身相关内容,以满足用户当前个性化需要。

[0136] 本实施例通过根据已知样本站点对全网站点进行内容挖掘,可以将不同站点的内容聚合到当前站点,扩大数据库中内容的范围,为用户提供更全面丰富的数据;通过根据自画像获取相应内容,可以为用户提供个性化内容,提升用户体验;本实施例通过行业统计得到已知样本站点,可以扩大样本数量,进而扩大数据库中的内容,为用户提供更丰富的信息;本实施例通过实时更新展现的内容,可以满足用户实时的个性化需求,提升用户体验。

[0137] 需要说明的是,在本发明的描述中,术语“第一”、“第二”等仅用于描述目的,而不能理解为指示或暗示相对重要性。此外,在本发明的描述中,除非另有说明,“多个”的含义是两个或两个以上。

[0138] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本发明的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本发明的实施例所属技术领域的技术人员所理解。

[0139] 应当理解,本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如,如果用硬件来实现,和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列(PGA),现场可编程门阵列(FPGA)等。

[0140] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0141] 此外,在本发明各个实施例中的各功能单元可以集成在一个处理模块中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。

[0142] 上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0143] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

[0144] 尽管上面已经示出和描述了本发明的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本发明的限制,本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

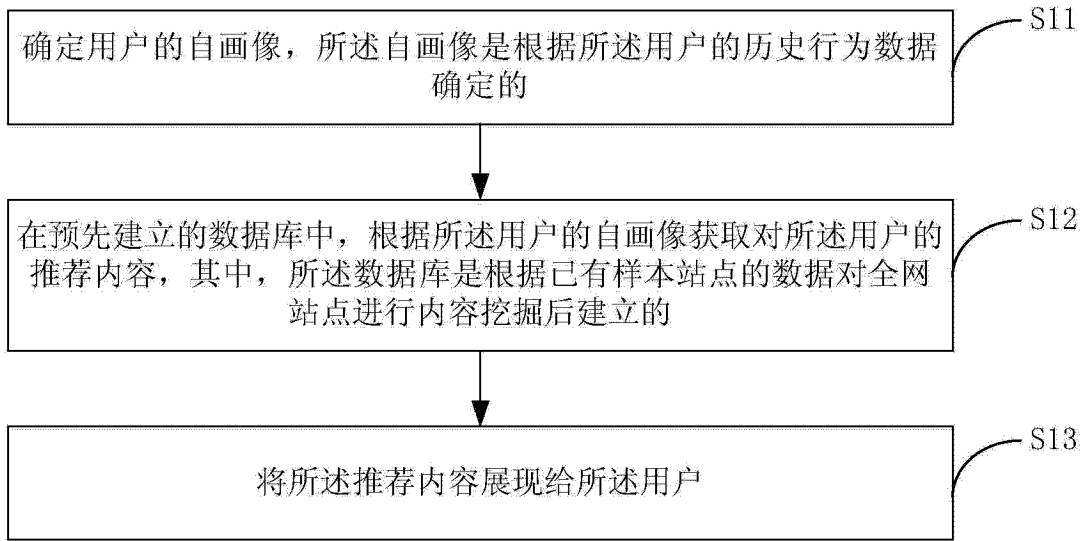


图 1



图 2



图 3

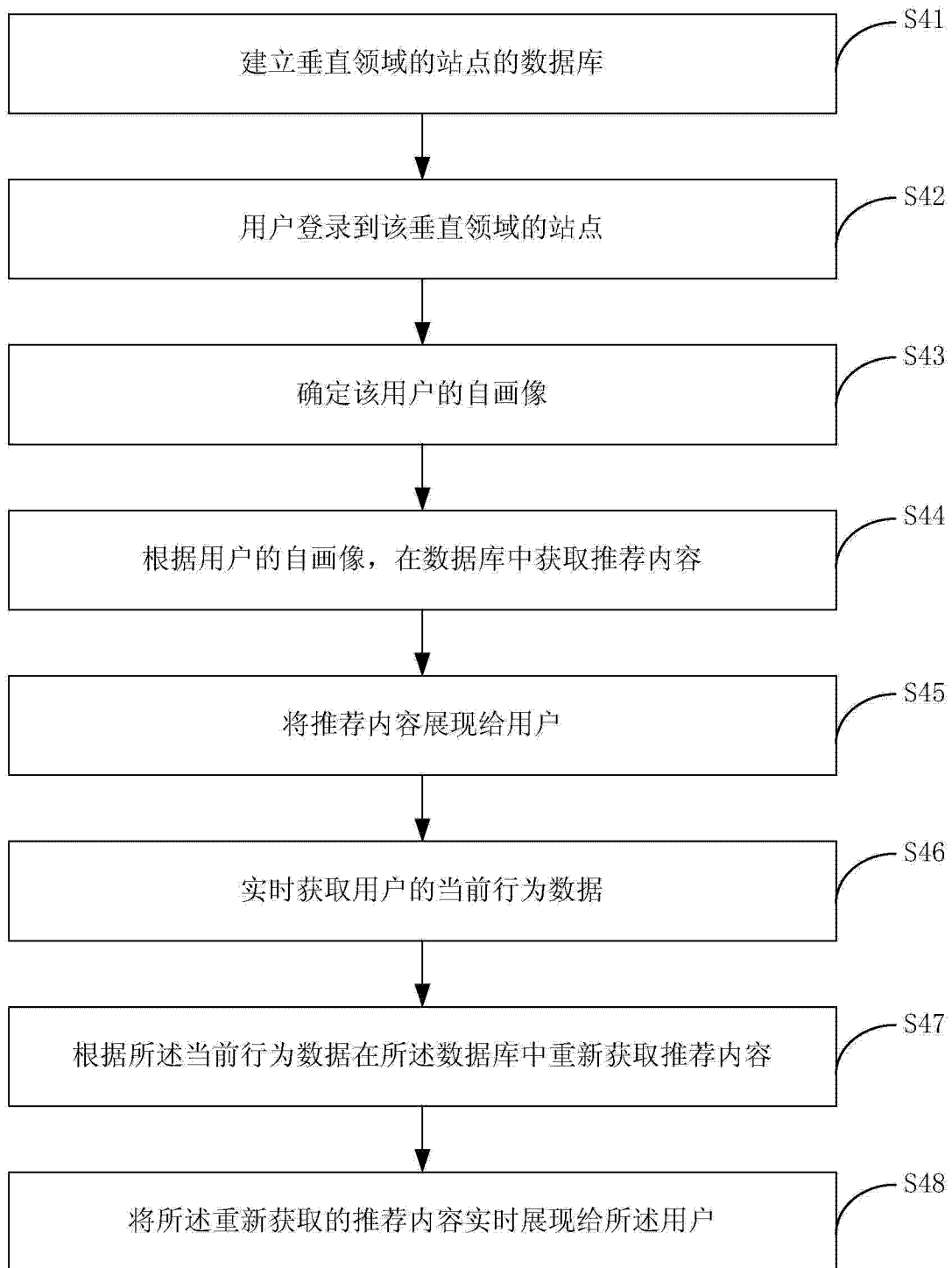


图 4



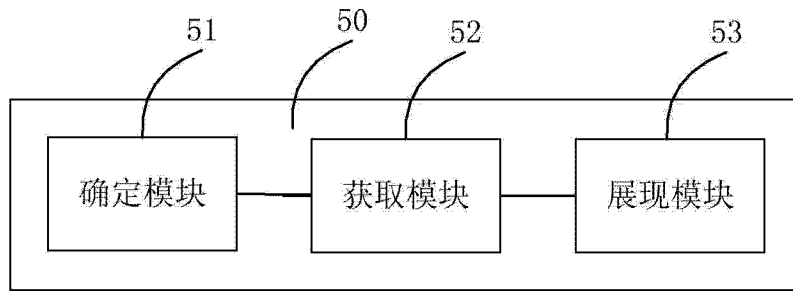


图 5

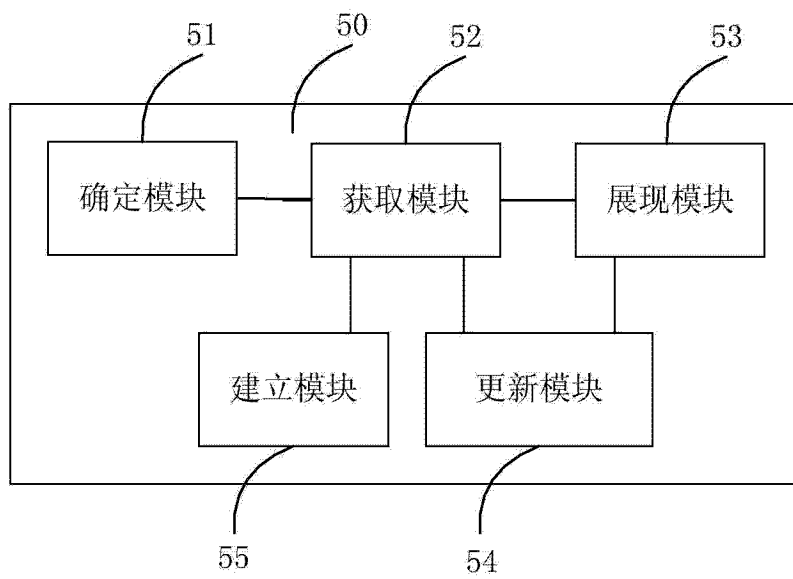


图 6