



(12) 发明专利

(10) 授权公告号 CN 101071424 B

(45) 授权公告日 2010.08.25

(21) 申请号 200610061354.2

审查员 赵晓春

(22) 申请日 2006.06.23

(73) 专利权人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市高新科技园南区  
高新南一道飞亚达高科技大厦 5-10 层

(72) 发明人 丁江伟 余祥鑫

(74) 专利代理机构 北京派特恩知识产权代理事

务所(普通合伙) 11270

代理人 蒋雅洁 王黎延

(51) Int. Cl.

G06F 17/30(2006.01)

(56) 对比文件

CN 1645366 A, 2005.07.27, 全文.

US 6701311 B2, 2004.03.02, 全文.

US 2003/0140148 A1, 2003.07.24, 全文.

CN 1471020 A, 2004.01.28, 说明书第 1 页第

26 行到第 4 页最后 1 行.

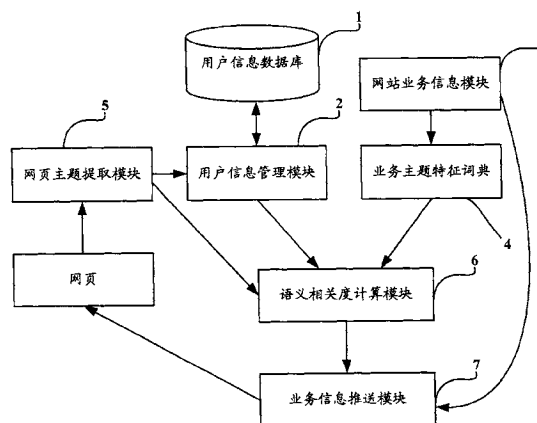
权利要求书 2 页 说明书 5 页 附图 4 页

(54) 发明名称

一种个性化信息推送系统和方法

(57) 摘要

本发明公开了一种个性化信息推送系统,包括:用于得出标识当前网页主题的若干主题词汇的网页主题提取模块、用于对待推送给用户的所有业务信息按照语义相关度进行计算,将计算结果排序的语义相关度计算模块、以及按照语义相关度排序的结果,将符合推送条件的个性化信息呈现给用户的业务信息推送模块。本发明还公开了一种个性化信息推送方法。本发明能够根据用户的静态属性和动态偏好,实现个性化的信息推送服务。



1. 一种个性化信息推送系统,其特征在于,该系统包括:网页主题提取模块、语义相关度计算模块、业务信息推送模块,其中:

网页主题提取模块,对当前浏览网页内容进行语义分析,得出标识当前网页主题的若干带有权重的主题词汇,所述的主题词汇的权重是该词汇按照揭示浏览网页主题的作用不同而赋予的;

语义相关度计算模块,根据当前用户的静态属性、历史行为偏好信息以及用户当前浏览网页的主题词汇信息对待推送给用户的所有业务信息按照语义相关度进行计算,将计算结果排序;

业务信息推送模块,按照语义相关度排序的结果,将符合推送条件的个性化信息呈现给用户。

2. 如权利要求1所述的系统,其特征在于,所述用户的静态属性包括用户的ip地址以及用户在进行网络服务系统注册时所登记的注册信息。

3. 如权利要求1所述的系统,其特征在于,所述的系统还包括:

用户信息数据库,保存用户的静态属性和使用网络的历史行为偏好信息;

用户信息管理模块,用于用户静态属性和历史行为偏好信息的上传和下载,还用于用户历史行为偏好的衰减;

网站业务信息模块,存储可以推送给用户的所有业务信息。

4. 如权利要求1所述的系统,其特征在于,所述的系统还包括:业务主题特征词典(4),按照用户区分度将业务进行分类,每个类别为一个与主题词汇相关的主题词表。

5. 一种个性化信息推送方法,包括下述步骤:

(a) 用户浏览网页时,对网页内容进行主题语义分析,得到当前网页内容中若干带权重的主题词汇,所述的主题词汇的权重是该词汇按照揭示浏览网页主题的作用不同而赋予的;

(b) 根据用户静态属性、历史行为偏好信息和当前网页内容的主题词汇,在所推送的业务信息里计算出最相关的若干条业务信息进行推送。

6. 如权利要求5所述的方法,其特征在于,所述的步骤(a)之前还包括下述步骤:

用户使用网络服务系统时,获取该用户的静态属性、历史行为偏好信息。

7. 如权利要求5所述的方法,其特征在于,所述的用户的静态属性包括用户的ip地址以及用户在进行网络服务系统注册时所登记的注册信息。

8. 如权利要求5所述的方法,其特征在于,所述的用户历史行为偏好信息是通过下述方法得到的:

(a1) 用户在进行当前网络操作时记录行为偏好;

(a2) 用户完成业务时或者在预定时间,根据用户对网络进行操作时记录的行为偏好对该用户的历史行为偏好信息进行更新。

9. 如权利要求8所述的方法,其特征在于,所述的步骤(a2)还包括:将用户的行为偏好设定有权重,用户进行历史行为偏好信息更新时,衰减较早的历史行为偏好权重,保存衰减后的计算结果。

10. 如权利要求5所述的方法,其特征在于,所述的步骤(b)进一步地包括下述步骤:

(b11) 利用当前网页内容的主题词汇获知其所对应的主题词表;

(b12) 根据用户静态属性、历史行为偏好信息和当前网页内容所对应的主题词表在所要推送的业务信息里计算出最相关的若干条业务信息进行推送。

## 一种个性化信息推送系统和方法

### 技术领域

[0001] 本发明涉及一种个性化信息推送方法,特别地,涉及一种基于用户行为分析的个性化信息推送系统和方法。

### 背景技术

[0002] 网络信息服务是指信息服务机构为满足用户的信息需求,通过计算机网络提供经过加工、整理的信息产品和服务的总称,它包括信息内容的加工处理、信息内容的提供、信息内容的存储与获取、网络增值服务、信息咨询服务等方面。目前网络信息的推送方法包括:在线广告、在线购物、上传和下载以及搜索引擎等,其中:

[0003] 在线广告,通常设置在网页界面上的一个固定位置,当需要播放广告的时候,按照预先设定的清单到网络服务器取广告数据,并且每天按照一定的设定播放规则(如:轮流播放等)来进行播放。这种信息推送的方式并没有用户区别,对于每个用户所推送的信息都是一致的。

[0004] 在线购物,通常用户按照网站预设定的类别进行逐级的选择,自主地找到自己喜欢的物品然后支付购买;或者按照所有产品的热度、经典度、最新到货等信息做成概率上顾客容易接受的推荐产品,这种信息推送的方式也是一种大众化的产品推介。

[0005] 上传和下载,通常仅对用户地理位置进行区分选择物理服务器,对所有用户提供毫无二致的信息服务。

[0006] 搜索引擎,通常仅仅是利用概率论进行排序算法,但是推送信息的排序并没有关注使用者的个体属性和行为,也是一种共性的信息。

[0007] 综上,上述信息推送的方式都是将所有的网络使用者为目标的,而没有考虑到信息受众的个体静态属性和动态兴趣偏好的差异,因此经常会出现信息在用户不需要的时候也进行推送,甚至产生不可预料的骚扰,导致用户对网络服务质量的不满意度升高。

### 发明内容

[0008] 本发明的目的即是针对上述现有技术中存在的缺点而提出的一种个性化信息推送系统和方法,以实现针对不同个体具有针对性的信息推送。

[0009] 为达到上述目的,本发明提供一种个性化信息推送系统,包括:网页主题提取模块、语义相关度计算模块、业务信息推送模块,其中:

[0010] 网页主题提取模块,对当前浏览网页内容进行语义分析,得出标识当前网页主题的若干带有权重的主题词汇,所述的主题词汇的权重是该词汇按照揭示浏览网页主题的作用不同而赋予的;

[0011] 语义相关度计算模块,根据当前用户的静态属性、历史行为偏好信息以及用户当前浏览网页的主题词汇信息对待推送给用户的所有业务信息按照语义相关度进行计算,将计算结果排序;

[0012] 业务信息推送模块,按照语义相关度排序的结果,将符合推送条件的个性化信息

呈现给用户。

[0013] 进一步地,所述用户的静态属性包括用户的 ip 地址以及用户在进行网络服务系统注册时所登记的注册信息。

[0014] 进一步地,所述的系统还包括:用户信息数据库,保存用户的静态属性和使用网络的历史行为偏好信息;用户信息管理模块,用于用户静态属性和历史行为偏好信息的上传和下载,还用于用户历史行为偏好的衰减;网站业务信息模块,存储可以推送给用户的所有业务信息。

[0015] 进一步地,所述的系统还包括:业务主题特征词典(4),按照用户区分度将业务进行分类,每个类别为一个与主题词汇相关的主题词表。

[0016] 为达到上述发明目的,本发明还提供一种个性化信息推送方法,包括下述步骤:

[0017] (a) 用户浏览网页时,对网页内容进行主题语义分析,得到当前网页内容中若干带权重的主题词汇,所述的主题词汇的权重是该词汇按照揭示浏览网页主题的作用不同而赋予的;

[0018] (b) 根据用户静态属性、历史行为偏好信息和当前网页内容的主题词汇,在所推送的业务信息里计算出最相关的若干条业务信息进行推送。

[0019] 进一步地,所述的步骤(a)之前还包括下述步骤:

[0020] 用户使用网络服务系统时,获取该用户的静态属性、历史行为偏好信息。

[0021] 进一步地,所述的用户的静态属性包括用户的 ip 地址以及用户在进行网络服务系统注册时所登记的注册信息。

[0022] 进一步地,所述的用户历史行为偏好信息是通过下述方法得到的:

[0023] (a1) 用户在进行当前网络操作时记录行为偏好;

[0024] (a2) 用户完成业务时或者在预定时间,根据用户对网络进行操作时记录的行为偏好对该用户的历史行为偏好信息进行更新。

[0025] 更进一步地,所述的步骤(a2)还包括:将用户的行为偏好设定有权重,用户进行历史行为偏好信息更新时,衰减较早的历史行为偏好权重,保存衰减后的计算结果。

[0026] 进一步地,所述的步骤(b)进一步地包括下述步骤:

[0027] (b11) 利用当前网页内容的主题词汇获知其所对应的主题词表;

[0028] (b12) 根据用户静态属性、历史行为偏好信息和当前网页内容所对应的主题词表在所推送的业务信息里计算出最相关的若干条业务信息进行推送。

[0029] 与现有技术相比,本发明具有如下显著优点:

[0030] (1) 根据用户的静态属性和动态偏好,实现个性化的信息服务,更能符合用户需求,使用户避免无关信息的打扰;

[0031] (2) 将信息推送到最需要、最关注的用户,提升单位信息的推送价值,以及信息提供者的服务质量;

[0032] (3) 本发明适用范围广泛,可以应用于即时通讯系统、电子交易网站系统、bbs 社区、email 系统、手机通讯系统、博客系统等网络服务系统。

[0033] **附图说明**

[0034] 图 1 为本发明具体实施例一的个性化信息推送系统的结构图;

[0035] 图 2 为本发明具体实施例一的个性化信息推送方法的流程图;

[0036] 图 3 为本发明具体实施例二的个性化信息推送系统的结构图；

[0037] 图 4 为本发明具体实施例二的个性化信息推送方法的流程图。

### 具体实施方式

[0038] 下面结合附图和具体实施例对本发明作进一步的介绍,但不作为对本发明的限定。

[0039] 实施例一：

[0040] 图 1 所示为本发明实施例一的个性化信息推送系统的结构图。从图 1 中看出,该系统包括:用户信息数据库 1、用户信息模块 2、网站业务信息模块 3、业务主题特征词典 4、网页主题提取模块 5、语义相关度计算模块 6、业务信息推送模块 7,其中:

[0041] 用户信息数据库 1,用于保存用户的静态属性,该静态属性包括用户的 ip 地址和用户在进行网络服务系统注册时所登记的注册信息,例如:性别、年龄、住址、兴趣爱好等;还保存用户使用网络的历史行为偏好信息,包括:用户浏览论坛和网页的地理偏好、语言偏好、类别偏好、产品品牌偏好等;

[0042] 用户信息模块 2,是用户信息数据库 1 的接口单元,用于用户静态属性和当前行为偏好信息的采集上传,历史行为偏好权重的衰减以及用户已经保存的静态属性和历史行为偏好信息的下载;

[0043] 网站业务信息模块 3,存储可以推送给用户的所有业务信息,包括:广告业务、购物业务、下载业务等信息,所有业务信息能够根据用于搜索的词汇获取;

[0044] 业务主题特征词典 4,按照用户区分度将业务进行分类,每个类别为一个与主题词汇相关的主题词表,该主题词表中的词汇可以作为该类别信息在网站业务信息模块 3 中进行信息搜索的搜索词汇;例如:将广告按照汽车产品进行分类,同时在以“汽车”为主题词汇时,能够将汽车产品这一类别的主题词表中的词汇都作为用于搜索的词汇;

[0045] 网页主题提取模块 5,对当前浏览网页内容进行语义分析,得出标识当前网页主题的若干主题词汇,这些主题词汇按照揭示主题的作用不同分别被赋予不同的权重;

[0046] 语义相关度计算模块 6,获取当前用户的静态属性、历史行为偏好信息以及用户当前浏览网页的主题词汇等信息,利用上述信息对待推送给用户的所有业务按照语义相关度进行计算,将计算结果排序;

[0047] 业务信息推送模块 7,按照语义相关度排序的结果,将符合推送条件的个性化信息呈现给用户,例如:广告投放,下载服务器链接实际 ip 地址等。

[0048] 用户使用网络服务系统时,获取并保存该用户的 ip 地址以及其在注册过程中所登记的性别、年龄、住址、兴趣爱好等静态属性,在浏览网页过程中获取并保存用户使用网络的行为偏好信息;当用户再次使用该网络服务系统时,该系统可以进行个性化信息的推送,如图 2 所示,该方法包括下述步骤:

[0049] 步骤 101,用户登陆已经访问过的网络服务系统时,获取当前用户的静态属性、历史行为偏好信息,为个性化信息的推送提供选择依据,其中个性化信息可以包括:业务类别的选择、物理服务器的选择等;

[0050] 在本步骤中,如果没有已经存储的历史行为偏好信息时,只获取用户的静态属性;

[0051] 步骤 102, 用户浏览网页信息时, 对网页信息内容进行主题语义分析, 得到当前网页内容中按照权重排列的若干主题词汇, 并记录当前用户操作的当前行为偏好;

[0052] 步骤 103: 根据当前网页内容的主题词汇获取其对应的主题词表;

[0053] 步骤 104, 根据用户静态属性、历史行为偏好信息和当前网页内容的主题词表中的词汇, 在所推送的业务信息里计算出最相关的若干条业务信息推送给用户;

[0054] 步骤 105, 如果所推送的信息满足用户需要, 则用户对所推送的信息进行相关服务请求, 服务器根据用户的请求进行相应处理, 为用户提供相应的业务服务;

[0055] 步骤 106, 当用户所请求的业务结束时, 根据本次网络为用户提供业务过程的信息对该用户的历史行为偏好信息进行更新, 并衰减较早的历史行为偏好权重, 保存衰减后的计算结果。

[0056] 其中: 历史行为偏好权重的初始设定值为系统设定的固定值, 其衰减的计算过程可以采用如下方法, 如果本次历史行为偏好更新的时间距离上次更新时间已经超过了历史行为偏好的衰减半周期, 将较早的历史行为偏好权重衰减为原来的  $1/2$ , 然后增加新的带有权重的偏好信息; 如果本次历史行为偏好更新的时间距离上次更新时间在历史行为偏好的衰减半周期内, 不需要对历史行为偏好进行衰减, 直接增加带有权重的新偏好计算结果, 并保存。

[0057] 历史行为偏好的衰减半周期与用户使用网络服务的频度相关, 可以进行设定, 例如: 可以将用户每访问 10 次网络设定为一个衰减半周期。

[0058] 从本实施例可以看出, 利用本发明所提出的系统和方法具有显著的优点如下:

[0059] (1) 根据用户的静态属性和动态偏好, 实现个性化的信息服务, 更能符合用户需求, 使用户避免无关信息的打扰;

[0060] (2) 将信息推送到最需要、最关注的用户, 提升单位信息的推送价值, 以及信息提供者的服务质量。

[0061] 实施例二:

[0062] 个性化信息推送系统也可以减少上述实施例一所介绍的系统中的业务主题特征词典 4, 此时其系统结构图如图 3 所示。

[0063] 利用该系统来实现个性化信息的推送方法包括图 4 所述的下述步骤:

[0064] 步骤 201, 用户登陆已经访问过的网络服务系统时, 获取当前用户的静态属性、历史行为偏好信息, 为个性化信息的推送提供选择依据, 其中个性化信息可以包括: 业务类别的选择、物理服务器的选择等;

[0065] 步骤 202, 用户浏览网页信息时, 对网页信息内容进行主题语义分析, 得到当前网页内容中按照权重排列的若干主题词汇, 并记录当前用户操作的当前行为偏好;

[0066] 步骤 203, 根据用户静态属性、历史行为偏好信息和当前网页内容的主题词汇, 在所推送的业务信息里计算出最相关的若干条业务信息推送给用户;

[0067] 步骤 204, 如果所推送的信息满足用户需要, 则用户对所推送的信息进行相关服务请求, 服务器根据用户的请求进行相应处理, 为用户提供相应的业务服务;

[0068] 步骤 205, 当用户所请求的业务结束时, 根据本次网络为用户提供业务过程的信息对该用户的历史行为偏好信息进行更新, 并衰减较早的历史行为偏好权重, 保存衰减后的计算结果。

[0069] 可见,实施例二同样具有实施例一所具有的优点,也能够根据用户的静态属性和动态偏好,实现个性化的信息服务,但是与实施例一相比,只是在计算最相关的若干条信息时的根据是用户静态属性、使用网络的历史信息和当前主题词汇;此时当前主题词汇的数量明显少于主题词表的词汇数量,由于词汇较稀疏,所以被选择的推送信息会相对于上一实施例少,且目标性也不如实施例一强。

[0070] 在另一实施例中,对用户的历史行为偏好信息进行更新可以设定为定时更新,而不是在每次业务结束时,例如:每三天按照三天来的业务日志更新用户信息数据库 1 中的信息,虽然减少了更新的次数,但是对用户的历史行为偏好数据的更新准确性会更强。

[0071] 需要说明的是,本发明适用于即时通讯系统、电子交易网站系统、bbs 社区、email 系统、手机通讯系统、博客系统等各种网络服务系统。



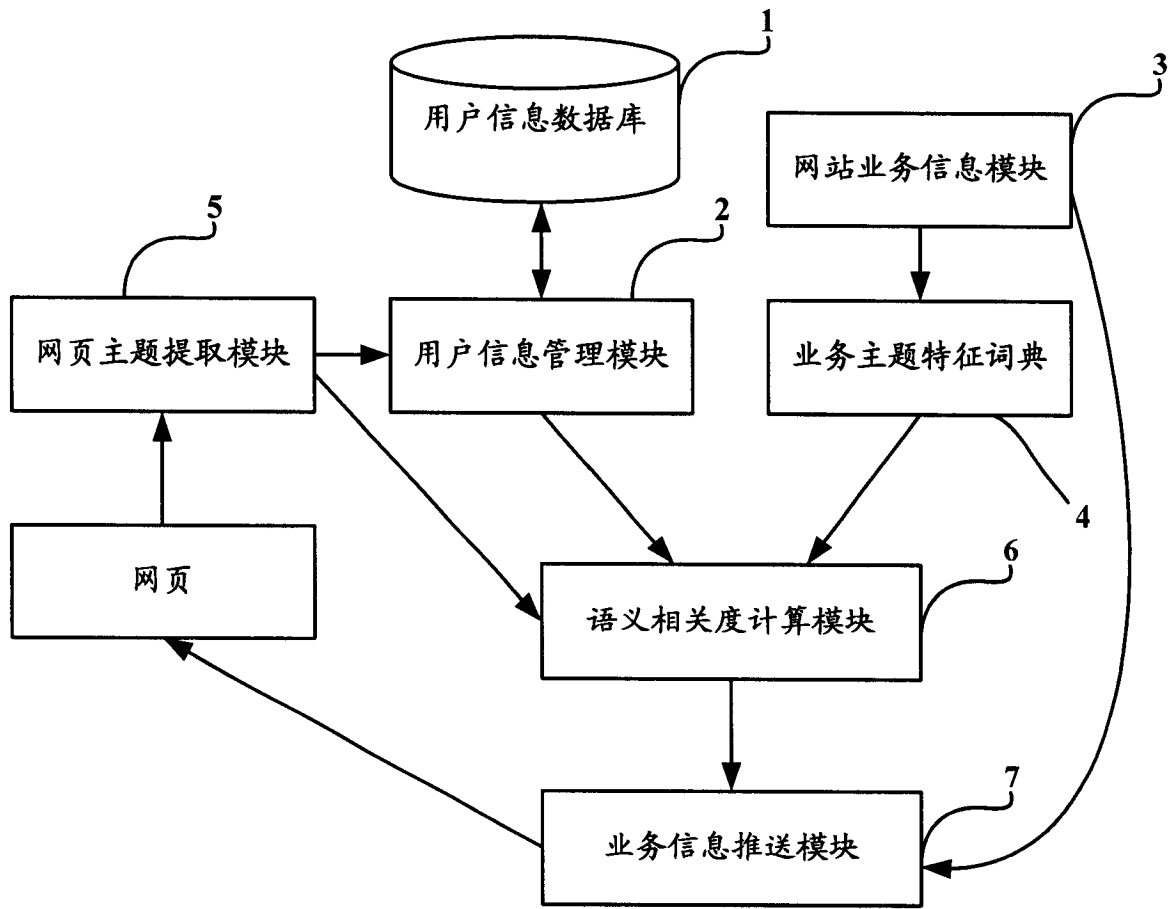


图 1

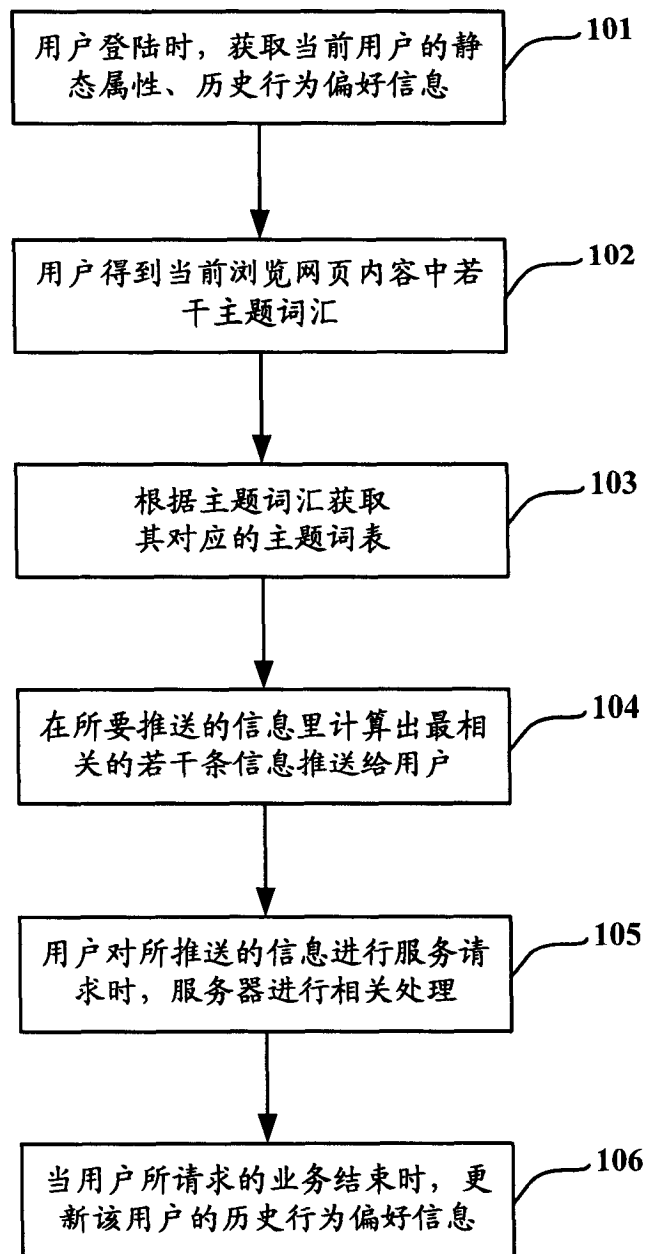


图 2

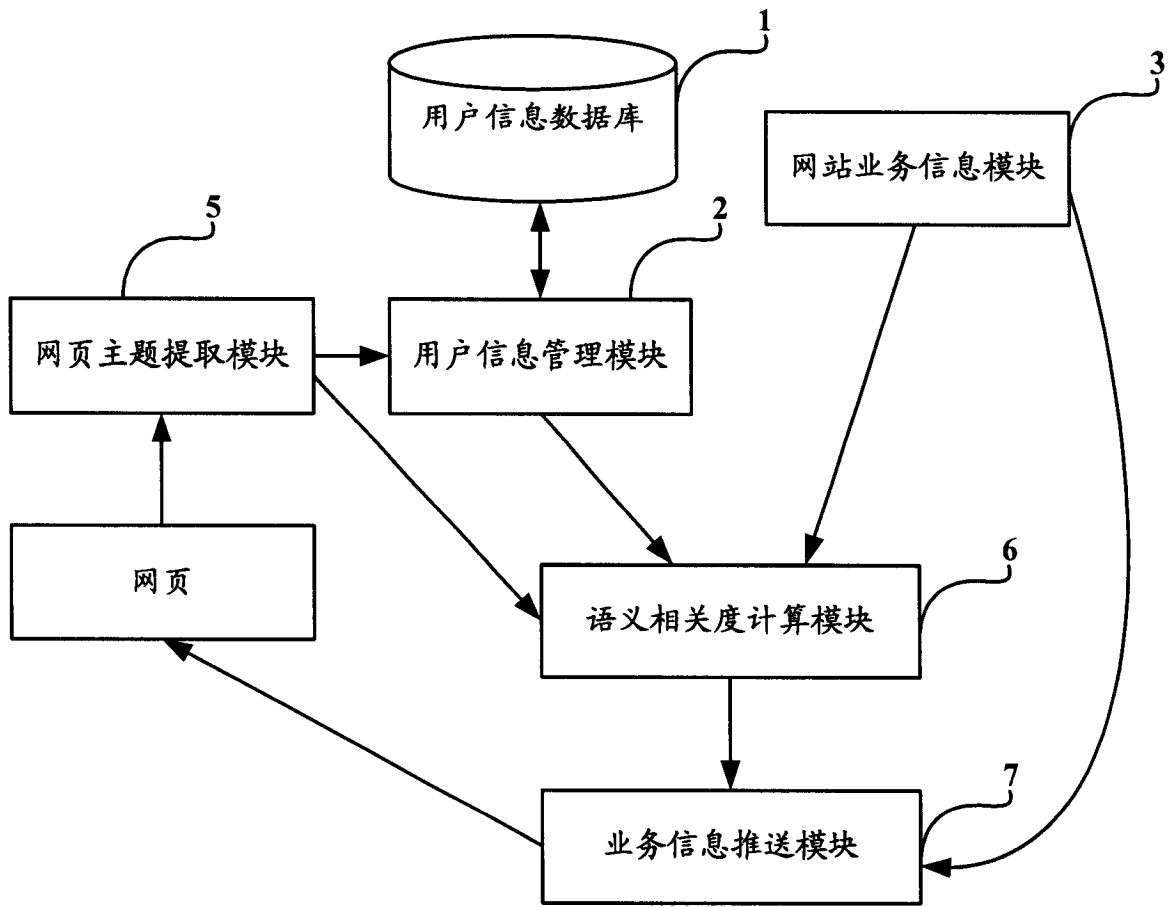


图 3

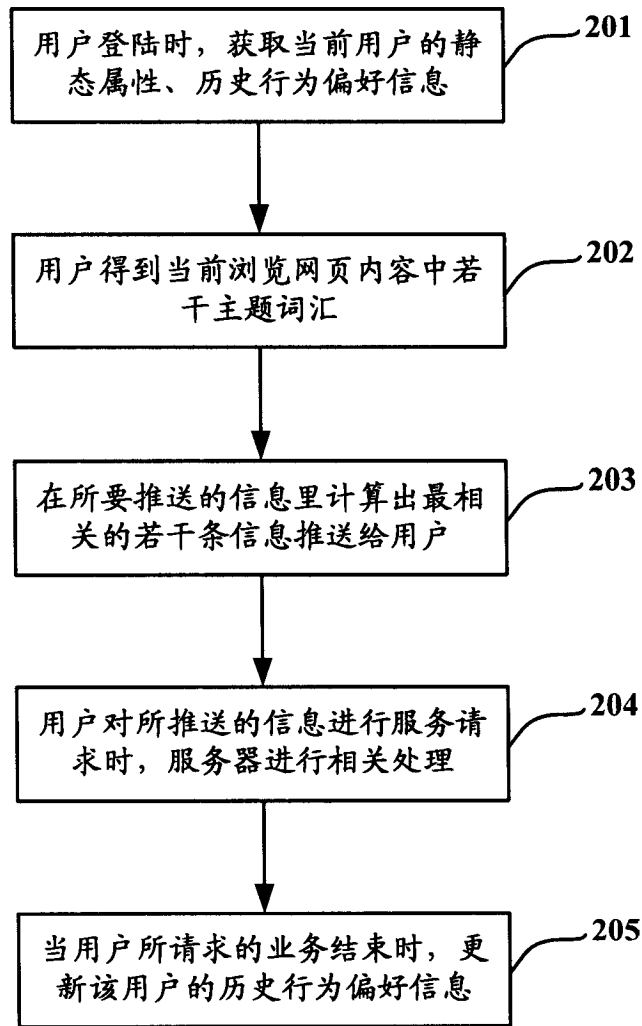


图 4