



(12) 发明专利申请

(10) 申请公布号 CN 103108049 A

(43) 申请公布日 2013. 05. 15

(21) 申请号 201310054790. 7

(22) 申请日 2013. 02. 20

(71) 申请人 杭州东信北邮信息技术有限公司  
地址 100191 北京市海淀区知春路 9 号坤讯大厦 7 层

(72) 发明人 廖建新 刘红明 周立娟 李萍  
李曲 赵贝尔

(51) Int. Cl.

H04L 29/08 (2006. 01)

H04M 1/725 (2006. 01)

G06F 17/30 (2006. 01)

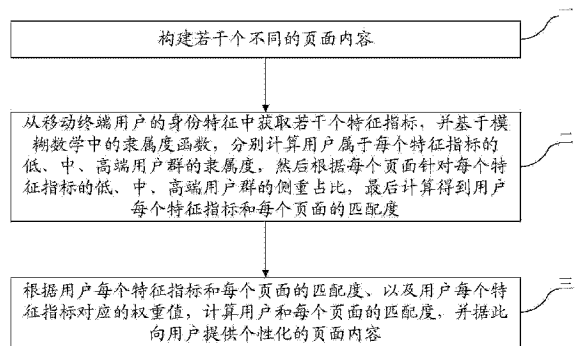
权利要求书3页 说明书7页 附图2页

(54) 发明名称

一种为移动终端用户提供个性化页面的方法

(57) 摘要

一种为移动终端用户提供个性化页面的方法,包括有:步骤一、构建若干个不同的页面内容;步骤二、从移动终端用户的身份特征中获取若干个特征指标,并基于模糊数学中的隶属度函数,分别计算用户属于每个特征指标的低、中、高端用户群的隶属度,然后根据每个页面针对每个特征指标的低、中、高端用户群的侧重占比,最后计算得到用户每个特征指标和每个页面的匹配度;步骤三、根据用户每个特征指标和每个页面的匹配度、以及用户每个特征指标对应的权重值,计算用户和每个页面的匹配度,并据此向用户提供个性化的页面内容。本发明属于网络应用技术领域,能根据移动终端用户的身份特征来为用户提供个性化页面,还能进一步应用于图书阅读的门户网站。



1. 一种为移动终端用户提供个性化页面的方法,其特征在於,包括有:

步骤一、构建若干个不同的页面内容;

步骤二、从移动终端用户的身份特征中获取若干个特征指标,并基于模糊数学中的隶属度函数,分别计算用户属于每个特征指标的低、中、高端用户群的隶属度,然后根据每个页面针对每个特征指标的低、中、高端用户群的侧重占比,最后计算得到用户每个特征指标和每个页面的匹配度;

步骤三、根据用户每个特征指标和每个页面的匹配度、以及用户每个特征指标对应的权重值,计算用户和每个页面的匹配度,并据此向用户提供个性化的页面内容。

2. 根据权利要求 1 所述的方法,其特征在於,所述步骤二中,从移动终端用户的身份特征中获取的特征指标包括但不限于月通讯费用、短信量、GPRS 流量或年龄。

3. 根据权利要求 1 所述的方法,其特征在於,所述步骤二中,用户属于第  $i$  个特征指标  $tz_i$  的低端用户群的隶属度  $ptz_{i\_low}(x_i)$  的计算公式如下:

$$ptz_{i\_low}(x_i) = \begin{cases} \frac{x_i}{a}, & 0 < x_i \leq a \\ 1, & a < x_i \leq b \\ \frac{c-x_i}{c-b}, & b < x_i \leq c \\ 0, & x_i > c \end{cases};$$

用户属于第  $i$  个特征指标  $tz_i$  的中端用户群的隶属度  $ptz_{i\_mid}(x_i)$  的计算公式如下:

$$ptz_{i\_mid}(x_i) = \begin{cases} 0, & 0 < x_i \leq b \\ \frac{x_i-b}{c-b}, & b < x_i \leq c \\ 1, & c < x_i \leq d \\ \frac{e-x_i}{e-d}, & d < x_i \leq e \\ 0, & x_i > e \end{cases};$$

用户属于第  $i$  个特征指标  $tz_i$  的高端用户群的隶属度  $ptz_{i\_high}(x_i)$  的计算公式如下:

$$ptz_{i\_high}(x_i) = \begin{cases} 0, & 0 < x_i \leq d \\ \frac{x_i-d}{e-d}, & d < x_i \leq e \\ 1, & e < x_i \leq f \\ \frac{g-x_i}{g-f}, & f < x_i \leq g \\ 0, & x_i > g \end{cases};$$

其中,  $x_i$  是用户对应于第  $i$  个特征指标  $tz_i$  的值,区间参数  $a, b, c, d, e, f, g$  是大于 0 的实数。

4. 根据权利要求 3 所述的方法,其特征在於,当特征指标是月通讯费用时,所述区间参数的值设置为:  $a=25, b=40, c=45, d=60, e=80, f=125, g=800$ ;当特征指标是短信量时,所述

区间参数的值设置为 :a=5, b=10, c=15, d=30, e=60, f=140, g=800 ;当特征指标是 GPRS 流量时,所述区间参数的值设置为 :a=5, b=10, c=18, d=30, e=35, f=55, g=400 ;当特征指标是年龄时,所述区间参数的值设置为 :a=5, b=10, c=15, d=30, e=60, f=140, g=80。

5. 根据权利要求 1 所述的方法,其特征在于,所述步骤二中,预先设置每个页面针对每个特征指标的低、中、高端用户群的侧重占比,且每个页面针对每个特征指标的低、中、高端用户群的侧重占比之和为 1。

6. 根据权利要求 1 所述的方法,其特征在于,所述步骤二中,用户第  $i$  个特征指标  $tz_i$  和页面  $j$  的匹配度  $p_j tz_i(x_i)$  的计算公式如下:

$p_j tz_i(x_i) = ptz_{i\_low}(x_i) * p_{j\_tz_i\_l} + ptz_{i\_mid}(x_i) * p_{j\_tz_i\_m} + ptz_{i\_high}(x_i) * p_{j\_tz_i\_h}$ , 其中,  $x_i$  是用户对应于第  $i$  个特征指标  $tz_i$  的值,  $ptz_{i\_low}(x_i)$ 、 $ptz_{i\_mid}(x_i)$ 、 $ptz_{i\_high}(x_i)$  分别是用户属于第  $i$  个特征指标  $tz_i$  的低、中、高端用户群的隶属度,  $p_{j\_tz_i\_l}$ 、 $p_{j\_tz_i\_m}$ 、 $p_{j\_tz_i\_h}$  分别是页面  $j$  针对特征指标  $tz_i$  的低、中、高端用户群的侧重占比。

7. 根据权利要求 1 所述的方法,其特征在于,所述步骤三中,用户和页面  $j$  的匹配度的计算公式如下:  $p_j = \sum_{i=1}^n \omega_i * p_j tz_i(x_i)$ , 其中,  $\omega_i$  是用户第  $i$  个特征指标  $tz_i$  对应的权重值,  $n$  是从用户身份特征中获取的特征指标总数,  $p_j tz_i(x_i)$  是用户第  $i$  个特征指标  $tz_i$  和页面  $j$  的匹配度。

8. 根据权利要求 1 所述的方法,其特征在于,所述步骤三还包括有:

从用户和所有页面的匹配度中挑选最大值,并将匹配度最大值所对应的页面内容提供给移动终端用户。

9. 根据权利要求 1 所述的方法,其特征在于,当应用于图书阅读门户网站时,根据页面的图书内容对用户和页面  $j$  的匹配度进行调整,所述步骤三还包括有:

步骤 A1、将页面  $j$  上所有图书内容包含的不重复的内容标签构成页面  $j$  内容标签集,并查询用户在一段时间内访问过的图书信息,然后根据页面  $j$  内容标签集中每个内容标签在用户所有访问过的图书内容包含的内容标签中出现的频率,逐一计算页面  $j$  内容标签集中每个内容标签的用户-内容标签权重;

步骤 A2、调整用户和页面  $j$  的匹配度:  $p_j = \alpha * p_j + (1-\alpha) * \lambda * \sum_{\mu=1, \dots, n_j} R_{j\mu} * tr_{\mu}$ , 其中,  $\alpha$  是用户身份特征的权重,  $1-\alpha$  是用户阅读行为偏好的权重,  $\lambda$  是图书内容对于页面区分识别的重要性度量权重,  $n_j$  是页面  $j$  内容标签集中的内容标签总数,  $tr_{\mu}$  是页面  $j$  内容标签集中第  $\mu$  个内容标签的用户-内容标签权重,  $R_{j\mu}$  是页面  $j$  内容标签集中第  $\mu$  个内容标签的权重。

10. 根据权利要求 1 所述的方法,其特征在于,当应用于图书阅读门户网站时,根据页面的推荐位链接对用户和页面  $j$  的匹配度进行调整,所述步骤三还包括有:

步骤 B1、查询用户在一段时间内对推荐位链接的访问信息,并根据页面  $j$  的每个推荐位链接在用户所有访问过的推荐位链接中出现的频率,逐一计算页面  $j$  的每个推荐位链接的权重;

步骤 B2、调整用户和页面  $j$  的匹配度:  $p_j = \alpha * p_j + (1-\alpha) * \beta * \sum_{v=1, \dots, m_j} lr_v$ , 其中,  $\alpha$  是用户身份特征的权重,  $1-\alpha$  是用户阅读行为偏好的权重,  $\beta$  是推荐位链接对于页面区分识别的

重要性度量权重,  $m_j$  是页面  $j$  的推荐位链接总数,  $l_{r_v}$  是页面  $j$  的第  $v$  个推荐位链接的权重。

11. 根据权利要求 1 所述的方法, 其特征在于, 当应用于图书阅读门户网站时, 根据页面的活动对用户和页面  $j$  的匹配度进行调整, 所述步骤三还包括有:

步骤 C1、查询用户在一段时间内主动参与的活动信息, 并根据页面  $j$  的每个活动在用户所有参与过的活动中出现的频率, 逐一计算页面  $j$  的每个活动的权重;

步骤 C2、调整用户和页面  $j$  的匹配度:  $p_j = \alpha * p_j + (1-\alpha) * \gamma * \sum_{\eta=1, \dots, d_j} ar_{\eta}$ , 其中,  $\alpha$  是用户身份特征的权重,  $1-\alpha$  是用户阅读行为偏好的权重,  $\gamma$  是活动对于页面区分识别的重要性度量权重,  $d_j$  是页面  $j$  的活动总数,  $ar_{\eta}$  是页面  $j$  的第  $\eta$  个活动的权重。

## 一种为移动终端用户提供个性化页面的方法

### 技术领域

[0001] 本发明涉及一种为移动终端用户提供个性化页面的方法,属于网络应用技术领域。

### 背景技术

[0002] 在网络信息日益剧增的当下,如何准确把握移动终端用户诉求、快速将符合用户喜好的信息以直线距离推送到用户面前显得尤为重要。如何为移动终端用户提供个性化页面,目前也提出了一些相关的解决方案:

[0003] 专利申请 CN201010179183. X (申请名称:一种移动终端个性化页面推送方法和装置,申请日:2010-05-19,申请人:深圳市五巨科技有限公司)公开了一种移动终端个性化页面推送方法和装置,所述方法包括用户行为统计分析过程和页面推送过程,所述页面推送过程包括:步骤 21:用户点击页面,移动终端连接至远程服务器;步骤 22:获取用户访问的头信息,分析出其中带有的标签;步骤 23:获取与上述标签关联度最高的一个或者多个关联标签;步骤 24:根据上述标签和关联标签中的关键字,在资源服务器中寻找对应的资源,搜索出用户需求的目标内容和个性化推送内容组成最终的页面反馈给用户。该技术方案主要用于根据用户的使用偏好(即访问记录)来推送个性化的网页页面,由于用户存在有年龄、个性、知识层次等不同的身份特征,其喜好也随身份的不同而千差万别,该技术方案并不能根据用户的身份特征(例如月通讯费用、短信量、GPRS 流量或年龄等能表征用户身份特征的特征指标)来提供个性化页面。

[0004] 因此,如何根据移动终端用户的身份特征,来为用户提供个性化页面,仍是一个未解决的技术问题。

### 发明内容

[0005] 有鉴于此,本发明的目的是提供一种为移动终端用户提供个性化页面的方法,能根据移动终端用户的身份特征中来为用户提供个性化页面。

[0006] 为了达到上述目的,本发明提供了一种为移动终端用户提供个性化页面的方法,包括有:

[0007] 步骤一、构建若干个不同的页面内容;

[0008] 步骤二、从移动终端用户的身份特征中获取若干个特征指标,并基于模糊数学中的隶属度函数,分别计算用户属于每个特征指标的低、中、高端用户群的隶属度,然后根据每个页面针对每个特征指标的低、中、高端用户群的侧重占比,最后计算得到用户每个特征指标和每个页面的匹配度;

[0009] 步骤三、根据用户每个特征指标和每个页面的匹配度、以及用户每个特征指标对应的权重值,计算用户和每个页面的匹配度,并据此向用户提供个性化的页面内容。

[0010] 与现有技术相比,本发明的有益效果是:由于用户群的划分具有模糊性和不确定性,本发明采用模糊数学中隶属度函数作为区分用户群的方法,从用户身份特征中获取月

通讯费用、短信量、GPRS 流量和年龄等多个特征指标,并计算用户属于每个特征指标的低、中、高端用户群的隶属度,也就是说,通过隶属度来衡量用户隶属于某个页面的程度指标,隶属度的大小对应概率的大小,即隶属度越大概率越大,从而能根据用户的身份特征来为用户提供与其喜好最相符的个性化页面,实现了对用户喜好内容的直观引导;对于图书阅读的门户网站来说,本发明还进一步通过用户过去对页面上图书内容、推荐位链接或活动的访问纪录来分析用户的阅读行为偏好,并根据用户的阅读行为偏好来调整用户和页面的匹配度,从而能在用户身份特征基础上,进一步整合用户的阅读数据,综合界定出最符合用户喜好的图书阅读页面。

### 附图说明

[0011] 图 1 是本发明一种为移动终端用户提供个性化页面的方法流程图。

[0012] 图 2 是图 1 步骤三中,应用于图书阅读门户网站时,根据页面的图书内容对用户和页面的匹配度进行调整的具体流程图。

[0013] 图 3 是图 1 步骤三中,应用于图书阅读门户网站时,根据页面的推荐位链接对用户和页面的匹配度进行调整的具体流程图。

[0014] 图 4 是图 1 步骤三中,应用于图书阅读门户网站时,根据页面的活动对用户和页面的匹配度进行调整的具体流程图。

### 具体实施方式

[0015] 为使本发明的目的、技术方案和优点更加清楚,下面结合附图对本发明作进一步的详细描述。

[0016] 如图 1 所示,本发明一种为移动终端用户提供个性化页面的方法,包括有:

[0017] 步骤一、构建若干个不同的页面内容;

[0018] 步骤二、从移动终端用户的身份特征中获取若干个特征指标,并基于模糊数学中的隶属度函数,分别计算用户属于每个特征指标的低、中、高端用户群的隶属度,然后根据每个页面针对每个特征指标的低、中、高端用户群的侧重占比,最后计算得到用户每个特征指标和每个页面的匹配度;

[0019] 步骤三、根据用户每个特征指标和每个页面的匹配度、以及用户每个特征指标对应的权重值,计算用户和每个页面的匹配度,并据此向用户提供个性化的页面内容。

[0020] 所述步骤一中,可以根据实际业务需要构建多个页面内容,例如 4 个页面。

[0021] 所述步骤二中,由于用户身份可以是中学生(包括初高中生)、大学生、外来务工人员、中高端商务人士(例如白领和一些社会精英)等,这些不同的用户身份在年龄构成、消费水平、上网需求等方面会有较大的差别,因此从移动终端用户的身份特征中获取的特征指标可以是月通讯费用、短信量、GPRS 流量或年龄等。

[0022] 由于用户群的划分具有模糊性和不确定性,如白领和外来务工人员、白领和高端商务人士没有明确的区分界限;其次用户身份特征与页面也没有明确的所属关系,如大学生并非一定要看奋斗励志的内容,高端商务人士也并非只看经管历史的内容。因此,本发明采用模糊数学中隶属度函数作为区分用户群的方法,也就是说,通过隶属度来衡量用户隶属于某个页面的程度指标,隶属度的大小对应概率的大小,即隶属度越大概率越大。

[0023] 步骤二中,用户属于第  $i$  个特征指标  $tz_i$  的低端用户群的隶属度  $ptz_{i\_low}(x_i)$  的计算公式如下:

$$[0024] \quad ptz_{i\_low}(x_i) = \begin{cases} \frac{x_i}{a}, & 0 < x_i \leq a \\ 1, & a < x_i \leq b \\ \frac{c-x_i}{c-b}, & b < x_i \leq c \\ 0, & x_i > c \end{cases};$$

[0025] 用户属于第  $i$  个特征指标  $tz_i$  的中端用户群的隶属度  $ptz_{i\_mid}(x_i)$  的计算公式如下:

$$[0026] \quad ptz_{i\_mid}(x_i) = \begin{cases} 0, & 0 < x_i \leq b \\ \frac{x_i-b}{c-b}, & b < x_i \leq c \\ 1, & c < x_i \leq d \\ \frac{e-x_i}{e-d}, & d < x_i \leq e \\ 0, & x_i > e \end{cases};$$

[0027] 用户属于第  $i$  个特征指标  $tz_i$  的高端用户群的隶属度  $ptz_{i\_high}(x_i)$  的计算公式如下:

$$[0028] \quad ptz_{i\_high}(x_i) = \begin{cases} 0, & 0 < x_i \leq d \\ \frac{x_i-d}{e-d}, & d < x_i \leq e \\ 1, & e < x_i \leq f \\ \frac{g-x_i}{g-f}, & f < x_i \leq g \\ 0, & x_i > g \end{cases};$$

[0029] 其中,  $x_i$  是用户对应于第  $i$  个特征指标  $tz_i$  的值,区间参数  $a, b, c, d, e, f, g$  是大于 0 的实数,其值根据业务实际情况来设定。

[0030] 例如当特征指标是月通讯费用时,所述区间参数的值可以设置为:  $a=25, b=40, c=45, d=60, e=80, f=125, g=800$ ,  $x_i$  是用户的月通讯费用值,即用户属于月通讯费用的

低端用户群的隶属度的计算公式如下:  $ptz_{i\_low}(x_i) = \begin{cases} \frac{x_i}{25}, & 0 < x_i \leq 25 \\ 1, & 25 < x_i \leq 40 \\ 9 - \frac{x_i}{5}, & 40 < x_i \leq 45 \\ 0, & x_i > 45 \end{cases}$  用户属于月

$$\text{通讯费用的中端用户群的隶属度的计算公式如下: } ptz_{1\_mid}(x_1) = \begin{cases} 0, & 0 < x_1 \leq 40 \\ \frac{x_1}{5} - 8, & 40 < x_1 \leq 45 \\ 1, & 45 < x_1 \leq 60 \\ 4 - \frac{x_1}{20}, & 60 < x_1 \leq 80 \\ 0, & x_1 > 80 \end{cases};$$

用户属于月通讯费用的高端用户群的隶属度的计算公式如下:

$$ptz_{1\_high}(x_1) = \begin{cases} 0, & 0 < x_1 \leq 60 \\ \frac{x_1}{20} - 3, & 60 < x_1 \leq 80 \\ 1, & 80 < x_1 \leq 125 \\ \frac{800 - x_1}{675}, & 125 < x_1 \leq 800 \\ 0, & x_1 > 800 \end{cases}.$$

[0031] 同样,当特征指标是短信量时,所述区间参数的值可以设置为:a=5, b=10, c=15, d=30, e=60, f=140, g=800;当特征指标是GPRS流量时,所述区间参数的值可以设置为:a=5, b=10, c=18, d=30, e=35, f=55, g=400;当特征指标是年龄时,所述区间参数的值可以设置为:a=5, b=10, c=15, d=30, e=60, f=140, g=80。

[0032] 为保证页面设置尽可能满足多数用户差异化的需求,即所构建的多个页面在每个特征指标的高、中、低端用户群上将有所侧重和倾向,比如页面一侧重于高月通讯费用、低短信量、高GPRS流量和中高年龄段的用户群,这样一方面可以提升用户的价值认同,另一方面也可以考虑到同类用户的趋众心理。因此,可以根据实际业务需要,采用专家分析法来预先设置每个页面针对每个特征指标的低、中、高端用户群的侧重占比,且每个页面针对每个特征指标的低、中、高端用户群的侧重占比之和为1,也就是说,  $p_{j\_tz_i\_l} + p_{j\_tz_i\_m} + p_{j\_tz_i\_h} = 1$ , 其中  $p_{j\_tz_i\_l}$ 、 $p_{j\_tz_i\_m}$ 、 $p_{j\_tz_i\_h}$  分别是页面j针对特征指标  $tz_i$  的低、中、高端用户群的侧重占比。例如下表1所示,构建4个页面:页面一、页面二、页面三和页面四,并从用户身份特征中获取特征指标为:月通讯费用、短信量、GPRS流量及年龄,则根据实际业务需要,预先设置表1中4个页面分别针对月通讯费用、短信量、GPRS流量及年龄的低、中、高端用户群的侧重占比,由于页面一侧重于高月通讯费用的用户群,则  $p_{1\_tz_1\_l} = 0.2$ ,  $p_{1\_tz_1\_m} = 0.2$ ,  $p_{1\_tz_1\_h} = 0.6$ 。

[0033]



指标维度		页面一	页面二	页面三	页面四
月通讯费用	低端用户群	$p_{1\_tz_1\_l}$	$p_{2\_tz_1\_l}$	$p_{3\_tz_1\_l}$	$p_{4\_tz_1\_l}$
	中端用户群	$p_{1\_tz_1\_m}$	$p_{2\_tz_1\_m}$	$p_{3\_tz_1\_m}$	$p_{4\_tz_1\_m}$
	高端用户群	$p_{1\_tz_1\_h}$	$p_{2\_tz_1\_h}$	$p_{3\_tz_1\_h}$	$p_{4\_tz_1\_h}$
GPRS 流量	低端用户群	$p_{1\_tz_2\_l}$	$p_{2\_tz_2\_l}$	$p_{3\_tz_2\_l}$	$p_{4\_tz_2\_l}$
	中端用户群	$p_{1\_tz_2\_m}$	$p_{2\_tz_2\_m}$	$p_{3\_tz_2\_m}$	$p_{4\_tz_2\_m}$
	高端用户群	$p_{1\_tz_2\_h}$	$p_{2\_tz_2\_h}$	$p_{3\_tz_2\_h}$	$p_{4\_tz_2\_h}$
短信量	低端用户群	$p_{1\_tz_3\_l}$	$p_{2\_tz_3\_l}$	$p_{3\_tz_3\_l}$	$p_{4\_tz_3\_l}$
	中端用户群	$p_{1\_tz_3\_m}$	$p_{2\_tz_3\_m}$	$p_{3\_tz_3\_m}$	$p_{4\_tz_3\_m}$
	高端用户群	$p_{1\_tz_3\_h}$	$p_{2\_tz_3\_h}$	$p_{3\_tz_3\_h}$	$p_{4\_tz_3\_h}$
年龄	低端用户群	$p_{1\_tz_4\_l}$	$p_{2\_tz_4\_l}$	$p_{3\_tz_4\_l}$	$p_{4\_tz_4\_l}$
	中端用户群	$p_{1\_tz_4\_m}$	$p_{2\_tz_4\_m}$	$p_{3\_tz_4\_m}$	$p_{4\_tz_4\_m}$
	高端用户群	$p_{1\_tz_4\_h}$	$p_{2\_tz_4\_h}$	$p_{3\_tz_4\_h}$	$p_{4\_tz_4\_h}$

[0034] 所述步骤二中,用户第  $i$  个特征指标  $tz_i$  和页面  $j$  的匹配度  $p_j tz_i(x_i)$  的计算公式如下:

[0035]  $p_j tz_i(x_i) = ptz_{i\_low}(x_i) * p_{j\_tz_i\_l} + ptz_{i\_mid}(x_i) * p_{j\_tz_i\_m} + ptz_{i\_high}(x_i) * p_{j\_tz_i\_h}$ , 其中,  $x_i$  是用户对应于第  $i$  个特征指标  $tz_i$  的值,  $ptz_{i\_low}(x_i)$ 、 $ptz_{i\_mid}(x_i)$ 、 $ptz_{i\_high}(x_i)$  分别是用户属于第  $i$  个特征指标  $tz_i$  的低、中、高端用户群的隶属度,  $p_{j\_tz_i\_l}$ 、 $p_{j\_tz_i\_m}$ 、 $p_{j\_tz_i\_h}$  分别是页面  $j$  针对特征指标  $tz_i$  的低、中、高端用户群的侧重占比。

[0036] 所述步骤三中,用户和页面  $j$  的匹配度的计算公式如下:

[0037]  $p_j = \sum_{i=1}^n \omega_i * p_j tz_i(x_i)$ , 其中,  $\omega_i$  是用户第  $i$  个特征指标  $tz_i$  对应的权重值,  $n$  是从用户身份特征中获取的特征指标总数。

从用户和所有页面的匹配度中挑选最大值,并将匹配度最大值所对应的页面内容提供给移动终端用户,从而实现个性化的网页页面。

[0038] 值得一提的是,本发明还可以进一步应用于图书阅读的门户网站中。对于图书阅读的门户网站来说,不同页面内容的主要差异包括图书内容、推荐位链接和活动,其中:

[0039] (1)、图书内容:可以根据图书主题、情节建立内容标签库,并通过内容标签库中的若干个内容标签对页面上用户可访问的图书内容进行标识;

[0040] (2)、推荐位链接:用户可以通过点击页面的推荐位链接,跳转至一级页面;

[0041] (3)、活动:用户可以主动参加页面上不同形式内容的活动。

[0042] 当本发明应用于图书阅读的门户网站中时,本发明还可以在用户的身份特征基础上,进一步通过分析用户过去对页面上的图书内容、推荐位链接或活动的访问纪录来分析用户的阅读行为偏好,并根据用户的阅读行为偏好来调整用户和页面的匹配度。

[0043] 如图 2 所示,所述步骤三中,应用于图书阅读门户网站时,本发明还可以根据页面的图书内容对用户和页面 j 的匹配度进行调整,进一步包括有:

[0044] 步骤 A1、将页面 j 上所有图书内容包含的不重复的内容标签构成页面 j 内容标签集,并查询用户在一段时间(如近 3 个月)内访问过的图书信息,然后根据页面 j 内容标签集中每个内容标签在用户所有访问过的图书内容包含的内容标签中出现的频率,逐一计算页面 j 内容标签集中每个内容标签的用户-内容标签权重,即  $tr_{\mu} = \frac{t_{\mu}}{T}$ ,其中  $tr_{\mu}$  是页面 j 内容

标签集中第  $\mu$  个内容标签的用户-内容标签权重,  $t_{\mu}$  是在用户所有访问过的图书内容包含的内容标签中第  $\mu$  个内容标签的出现次数, T 是用户所有访问过的图书内容包含的所有内容标签的出现总次数,其中图书内容包含的内容标签是从内容标签库中选取,并用于对图书内容进行标识的关键词;

[0045] 步骤 A2、调整用户和页面 j 的匹配度:  $p_j = \alpha * p_j + (1-\alpha) * \lambda * \sum_{\mu=1, \dots, n_j} R_{j\mu} * tr_{\mu}$ , 其中,

由于用户身份特征和阅读行为偏好对页面的影响程度不同,因此  $\alpha$  是用户身份特征的权重,  $1-\alpha$  是用户阅读行为偏好的权重,  $\lambda$  是图书内容对于页面区分识别的重要性度量权重,  $n_j$  是页面 j 内容标签集中的内容标签总数,  $tr_{\mu}$  是页面 j 内容标签集中第  $\mu$  个内容标签的用户-内容标签权重,  $R_{j\mu}$  是页面 j 内容标签集中第  $\mu$  个内容标签的权重,由于不同页面的图书内容可能存在交叉,为保证页面区分,不同页面的同一图书内容的权重可能不同,即同一内容标签的权重可能不同,  $R_{j\mu}$  的值根据实际业务情况来设定。

[0046] 如图 3 所示,所述步骤三中,应用于图书阅读门户网站时,本发明还可以根据页面的推荐位链接对用户和页面 j 的匹配度进行调整,进一步包括有:

[0047] 步骤 B1、查询用户在一段时间(如近 3 个月)内对推荐位链接的访问(即点击)信息,并根据页面 j 的每个推荐位链接在用户所有访问过的推荐位链接中出现的频率,逐一计算

页面 j 的每个推荐位链接的权重,即  $lr_v = \frac{l_v}{L}$ ,其中  $lr_v$  是页面 j 的第 v 个推荐位链接的权重,

$l_v$  是在用户所有访问过的推荐位链接中第 v 个推荐位链接的访问次数, L 是用户对所有推荐位链接的访问总次数;

[0048] 步骤 B2、调整用户和页面 j 的匹配度:  $p_j = \alpha * p_j + (1-\alpha) * \beta * \sum_{v=1, \dots, m_j} lr_v$ , 其中,  $\alpha$  是

用户身份特征的权重,  $1-\alpha$  是用户阅读行为偏好的权重,  $\beta$  是推荐位链接对于页面区分识别的重要性度量权重,  $m_j$  是页面 j 的推荐位链接总数,  $lr_v$  是页面 j 的第 v 个推荐位链接的权重。

[0049] 同理,如图 4 所示,所述步骤三中,应用于图书阅读门户网站时,本发明还可以根据页面的活动对用户和页面 j 的匹配度进行调整,进一步包括有:

[0050] 步骤 C1、查询用户在一段时间(如近 3 个月)内主动参与的活动信息,并根据页面 j 的每个活动在用户所有参与过的活动中出现的频率,逐一计算页面 j 的每个活动的权重,

即  $ar_n = \frac{a_n}{A}$ ,其中  $ar_n$  是页面 j 的第 n 个活动的权重,  $a_n$  是在用户所有参与过的活动中第

n 个活动的参与次数, A 是用户对所有活动的参与总次数;

[0051] 步骤 C2、调整用户和页面 j 的匹配度： $p_j = \alpha * p_j + (1-\alpha) * \gamma * \sum_{\eta=1, \dots, d_j} ar_{\eta}$ ，其中， $\alpha$  是用户身份特征的权重， $1-\alpha$  是用户阅读行为偏好的权重， $\gamma$  是活动对于页面区分识别的重要性度量权重， $d_j$  是页面 j 的活动总数， $ar_{\eta}$  是页面 j 的第  $\eta$  个活动的权重。

[0052] 为了更清楚的解释本发明方法的技术效果，申请人提取了部分用户的 BOSS 身份数据和近 3 个月的阅读访问互动数据。应用本发明方法后，分别计算得到用户与 4 个构建页面的匹配度，如下表所示，从每个用户与 4 个页面的匹配度中挑选最大值，并向用户提供所述最大值所对应的页面内容。

[0053]

用户	页面一	页面二	页面三	页面四	提供页面
User1	0.1600	0.1867	0.6000	0.0533	页面三
User2	0.3014	0.1791	0.5129	0.0067	页面三
User3	0.6177	0.2727	0.1096	0.0000	页面一
User4	0.3902	0.4061	0.1852	0.0185	页面二
User5	0.2087	0.3740	0.2957	0.1216	页面二
User6	0.3420	0.1961	0.4508	0.0110	页面三
User7	0.4920	0.2320	0.2760	0.0000	页面一
User8	0.2653	0.3800	0.0295	0.3253	页面二
User9	0.1600	0.1867	0.5834	0.0700	页面三
User10	0.1200	0.1600	0.7200	0.0000	页面三
User11	0.1226	0.2200	0.1574	0.4999	页面四
User12	0.0147	0.2619	0.0073	0.7161	页面四
User13	0.2737	0.3002	0.2303	0.1958	页面二
User14	0.4898	0.0327	0.4775	0.0001	页面一

[0054] 从运行数据来看，本发明能够解决现有技术的不足，实现了发明目的和较好的技术效果。

[0055] 上述仅为本发明的较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明保护的范围之内。

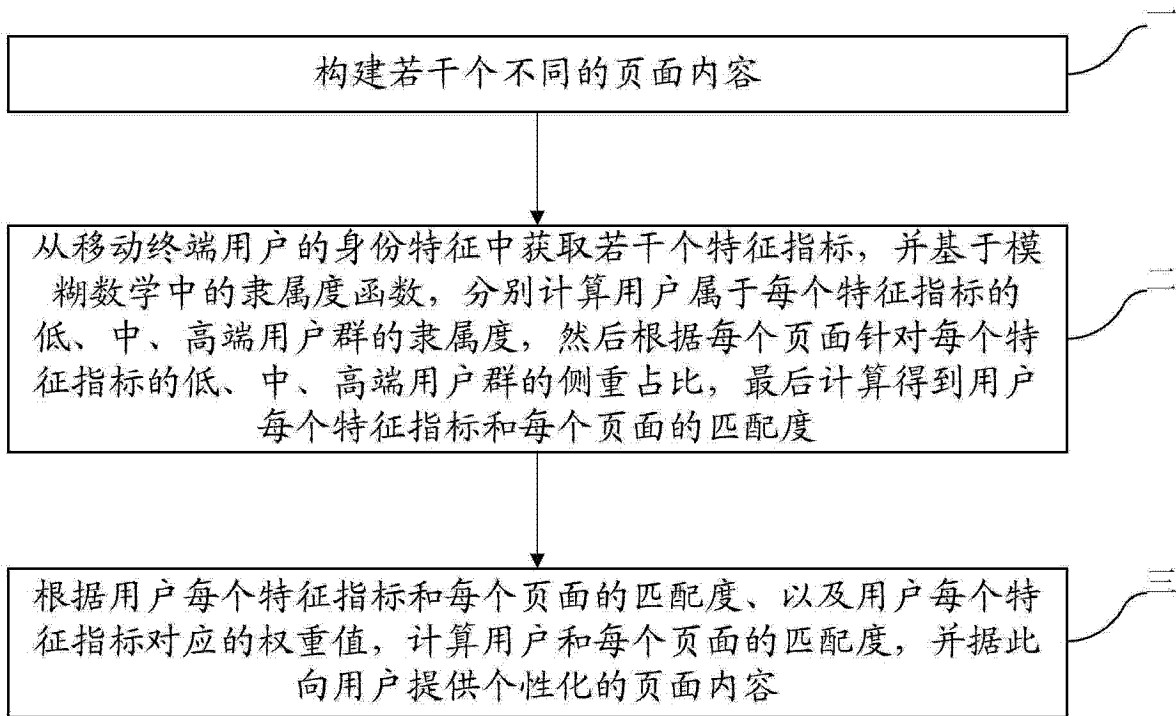


图 1

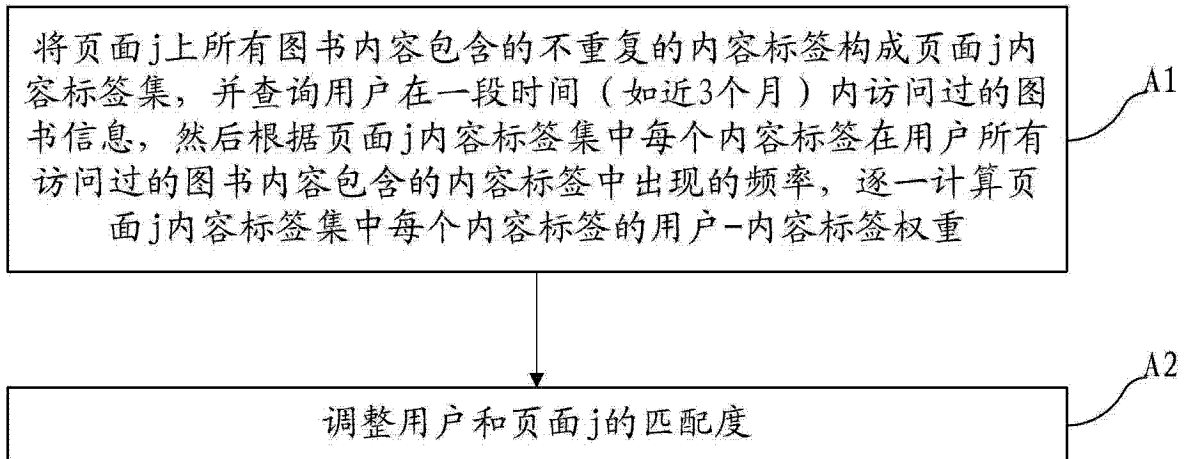


图 2

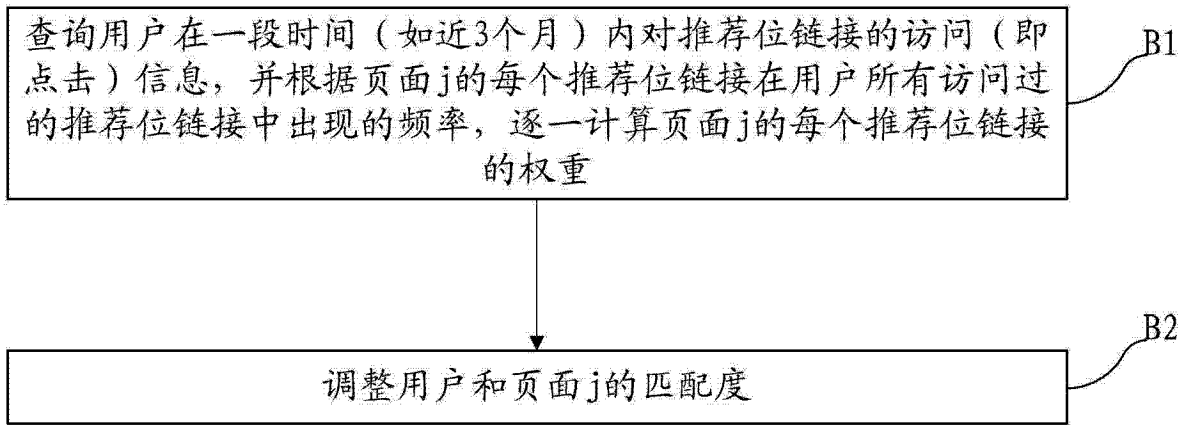


图 3

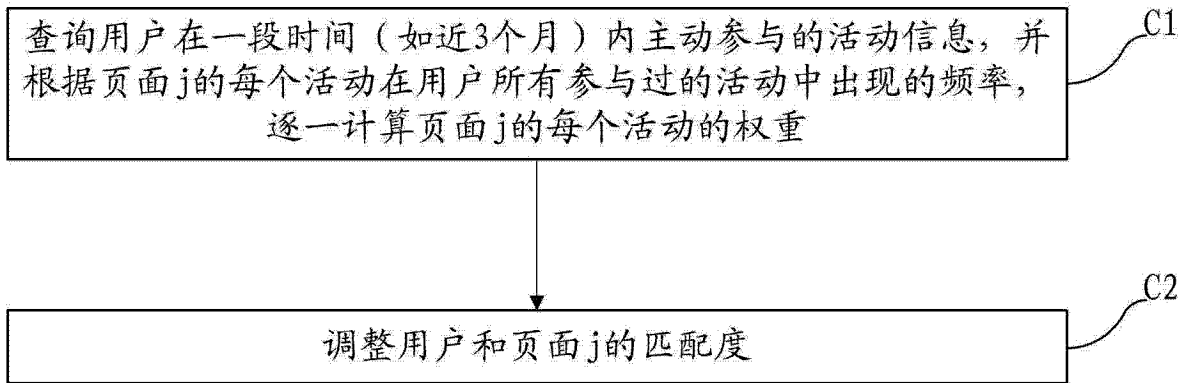


图 4