



## (12) 发明专利

(10) 授权公告号 CN 103019550 B

(45) 授权公告日 2015.08.26

(21) 申请号 201210524614.0

第 2-3 段、22-23 行，附图 4.

(22) 申请日 2012.12.07

CN 101504671 A, 2009.08.12, 说明书第 4 页  
11-18 行, 第 8 页第 4 段.

(73) 专利权人 东软集团股份有限公司

CN 101930475 A, 2010.12.29, 全文.

地址 110179 辽宁省沈阳市浑南新区新秀街  
2 号

US 6654735 B1, 2003.11.25, 全文.

(72) 发明人 邹存璐 蒋理成 韩宇 赵博  
王菊

US 2007276790 A1, 2007.11.29, 全文.

(74) 专利代理机构 北京鸿元知识产权代理有限  
公司 11327

CN 101079824 A, 2007.11.28, 全文.

代理人 林锦辉

CN 102087582 A, 2011.06.08, 全文.

(51) Int. Cl.

审查员 李锦川

G06F 3/0485(2013.01)

G06Q 30/02(2012.01)

(56) 对比文件

CN 102063456 A, 2011.05.18, 说明书第  
0035, 0038, 0043, 0044, 0046, 0047, 0103 段.

CN 101650634 A, 2010.02.17, 说明书第 5 页

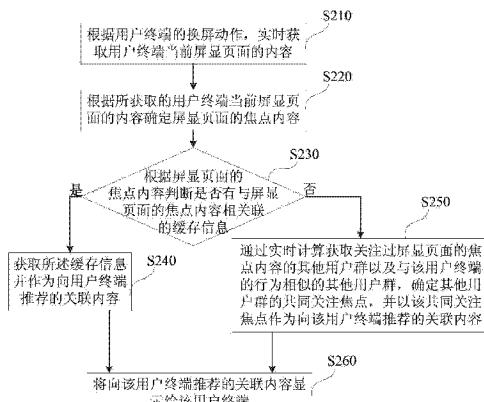
权利要求书 2 页 说明书 6 页 附图 4 页

(54) 发明名称

关联内容实时展现方法及系统

(57) 摘要

本发明提供了一种关联内容实时展现方法及系统，其方法包括：根据用户终端的换屏动作，实时获取用户终端当前屏显页面的内容；然后确定屏显页面的焦点内容；进而判断是否有与屏显页面的焦点内容相关联的缓存信息；如果有，获取缓存信息并作为向用户终端推荐的关联内容显示给用户终端；否则，通过实时计算获取关注过屏显页面的焦点内容的其他用户群以及与用户终端的行为相似的其他用户群，确定其他用户群的共同关注焦点，并以所述共同关注焦点作为向所述用户终端推荐的关联内容显示给用户终端。本发明能够根据用户当前停留以及关注的焦点不同，实时变化显示与该焦点相关以及用户可能喜欢的内容，从而提高用户的体验以及网站的转化率。



1. 一种关联内容实时展现方法,包括 :

根据用户终端的换屏动作,实时获取用户终端当前屏显页面的内容;其中,用户终端的换屏动作包括用户改变屏显页面内容的滚动或者滑动动作;

根据所获取的用户终端当前屏显页面的内容确定所述屏显页面的焦点内容;

根据所述屏显页面的焦点内容判断是否有与所述屏显页面的焦点内容相关联的缓存信息;如果有,获取所述缓存信息并作为向所述用户终端推荐的关联内容显示给所述用户终端;否则,

通过实时计算获取关注过所述屏显页面的焦点内容的其他用户群以及与所述用户终端的行为相似的其他用户群,确定其他用户群的共同关注焦点,并以所述共同关注焦点作为向所述用户终端推荐的关联内容显示给所述用户终端。

2. 如权利要求 1 所述的关联内容实时展现方法,其中,根据用户终端的滚动或者滑动动作获取用户当前显示页面的内容的过程包括 :

根据所述屏显页面中滚动条位置和所述用户终端的屏幕分辨率确定所述用户终端当前屏显页面的内容。

3. 如权利要求 1 所述的关联内容实时展现方法,其中,所述根据所获取的用户终端当前屏显页面的内容确定所述屏显页面的焦点内容的过程包括 :

利用用户终端的点击流数据以及点击位置信息,根据拟合的 log-Normal 函数拟合出用户在当前屏显页面的视线分布;

利用概率理论将视线分布概率最大的位置作为所确定的所述屏显页面的焦点内容。

4. 如权利要求 1 所述的关联内容实时展现方法,其中,在根据所获取的用户终端当前屏显页面的内容确定所述屏显页面的焦点内容之后,还包括 :

将所述屏显页面的焦点内容和用户在所述屏显页面的点击内容关联地存入缓存。

5. 如权利要求 4 所述的关联内容实时展现方法,其中,获取所述缓存信息并作为向所述用户终端推荐的关联内容显示给所述用户终端的过程包括 :

获取所述缓存信息,所述缓存信息包括关联存储的屏显页面的焦点内容和用户在所述屏显页面的点击内容;

根据所述焦点内容对查询到的历史点击过的关联内容的点击次数进行排序;

以所述排序中点击次数最高的预设数量的焦点内容作为向所述用户终端推荐的关联内容显示给所述用户终端。

6. 如权利要求 1 所述的关联内容实时展现方法,其中,实时计算获取关注过所述屏显页面的焦点内容的其他用户群以及与所述用户终端的行为相似的其他用户群,确定其他用户群的共同关注焦点的过程包括 :

获取最近邻居以及从最近邻居中的共同行为中发现用户感兴趣的内容,所述最近邻居包括关注过所述焦点内容的用户群以及行为相似的用户群;

按照共同浏览的用户终端数量对所述焦点内容进行排序;以所述排序中共同浏览的用户终端数量最高的预设数量的焦点内容作为所述共同关注焦点;其中,

利用关注焦点内容的用户群来获取与所述关注焦点内容最相关的内容,利用行为相似的用户群来获取跟该用户最相关的内容。

7. 如权利要求 1 所述的关联内容实时展现方法,其中,所述显示给所述用户终端的推

荐的关联内容始终显示在用户终端的当前屏显页面上。

8. 一种关联内容实时展现系统，包括：

屏显内容获取单元，用于根据用户终端的换屏动作，实时获取用户终端当前屏显页面的内容；其中，用户终端的换屏动作包括用户改变屏显页面内容的滚动或者滑动动作；

焦点内容确定单元，用于根据所述屏显内容获取单元获取的用户终端当前屏显页面的内容确定所述屏显页面的焦点内容；

关联信息缓存单元，用于关联地存储所述屏显页面的焦点内容和用户在所述屏显页面的点击内容；

缓存信息管理单元，用于根据所述屏显页面的焦点内容判断在所述关联信息缓存单元中是否有与所述屏显页面的焦点内容相关联的缓存信息；如果有，获取所述缓存信息并作为向所述用户终端推荐的关联内容；否则，通过实时计算获取关注过所述屏显页面的焦点内容的其他用户群以及与所述用户终端的行为相似的其他用户群，确定其他用户群的共同关注焦点，并以所述共同关注焦点作为向所述用户终端推荐的关联内容；

关联内容展现单元，用于把向所述用户终端推荐的关联内容显示给所述用户终端。

## 关联内容实时展现方法及系统

### 技术领域

[0001] 本发明涉及网页浏览技术领域,更为具体地,涉及一种推荐网络信息的关联内容实时展现方法及系统。

### 背景技术

[0002] 随着社会的进步和技术的发展,人们越来越多地利用网络来获取信息。在这种情况下,通过用户的浏览、搜索、点击和评论等行为,为用户推荐与他们最相关联或可能会喜欢的内容就成了门户网站、网上商城等公司或企业最为关注的焦点内容之一。

[0003] 推荐系统最早在亚马逊的网站上应用,根据以往用户的购买行为,推荐出购买某种产品同时可能购买的其他产品。从亚马逊成功的利用推荐系统提高其电子书店的转化率(网站期望的动作的访问量 / 总访问量,该指标用来衡量网站内容对访问者的吸引程度以及网站的宣传效果)之后,目前个性化推荐系统在互联网产品、电子商务以及社交网站中都得到了广泛的应用。一般来说,推荐算法需要根据大量用户在网站对某种产品使用过程中的行为(浏览,点击,评论等),结合相关的推荐算法(例如协同过滤算法),发现并向用户展现与其最相关联(可能喜欢的)的内容。

[0004] 近年来,推荐算法得到了不断的更新改进,例如协同过滤算法,该算法发现共同浏览过某个页面或与该用户行为相似的用户群体,并从相似用户群体中发现其共同行为并推荐给该用户,从而能够更加有效地向用户推荐可能感兴趣的内容或好友。传统用户行为数据的抓取分为隐性数据与显性数据,显性数据主要包括评价、购买这样直接表现用户喜好的操作;隐性数据主要包括用户的浏览行为、点击链接这样不能直接表达用户喜好的动作。隐性数据有其一定的优越性:首先隐性数据不需要用户额外的一些操作过程,用户体验较好;另外隐性数据由于其规模较大可以有效避免数据造假刷分等欺骗行为。

[0005] 图1示出了利用传统隐性用户行为数据的抓取为用户推荐关联内容的流程。如图1所示,在用户浏览一个URL网页时,即添加该浏览行为到用户浏览行为数据库,用户浏览行为数据库根据该浏览行为获取两方面的信息:①获取浏览该网页的其他用户行为信息;②获取跟该用户行为相似的其他用户行为信息,然后按照共同浏览的用户数对所涉及的网页的URL进行倒排序;显示排名前N个URL内容信息推荐给用户,用户通过点击即可查看推荐的URL内容信息。

[0006] 根据图1所示的流程图,可以发现,目前传统隐性用户行为数据的抓取方法只是获得用户点击浏览一个页面的内容,但是由于现在一个页面往往包含了大量不同的内容,用户在浏览的过程中往往只是浏览或者关注了其中很小的一部分内容。

[0007] 并且由于网站UI设计的不同,不同网页用户关注的焦点位置都存在差异。而由于手机、平板电脑等移动设备屏幕尺寸的限制,屏幕显示的内容非常有限,需要用户通过滚动滚轴或者滑动的方式查看一个网页中不同的内容(例如目前热门的社交网站、微博网站等在手机应用上非常广泛,一个页面中包含了很多用户产生的不同内容)。所以浏览同一个URL页面的不同用户有极大可能所浏览以及关注的内容(滚动/滑动页面停留的页面部

分)是不一样的,而上述方法并无法区分这种情况。因此,这种仅仅通过用户点击浏览的一个页面内容的抓取为依据确定推荐内容的方式,会不可避免地导致推荐结果准确度的降低。

[0008] 另外,目前的推荐系统计算出的结果往往都固定显示在页面的一个位置,用户在滚动浏览页面不同位置时,推荐结果的位置以及内容都不会根据用户关注焦点的变化而实时产生相关联的变化,导致用户要么无法看到推荐的内容(推荐内容滚动出屏幕显示范围),要么推荐的内容跟当前关注的焦点无关。

## 发明内容

[0009] 鉴于上述问题,本发明的目的是提供一种能够实时获取用户关注焦点并根据关注焦点的不同实时变化推荐关联的内容的关联内容实时展现方法及系统,以提高用户体验以及网站的转化率。

[0010] 根据本发明的一个方面,提供了一种关联内容实时展现方法,包括:

[0011] 根据用户终端的换屏动作,实时获取用户终端当前屏显页面的内容;根据所获取的用户终端当前屏显页面的内容确定所述屏显页面的焦点内容;根据所述屏显页面的焦点内容判断是否有与所述屏显页面的焦点内容相关联的缓存信息;如果有,获取所述缓存信息并作为向所述用户终端推荐的关联内容显示给所述用户终端;否则,通过实时计算获取关注过所述屏显页面的焦点内容的其他用户群以及与所述用户终端的行为相似的其他用户群,确定其他用户群的共同关注焦点,并以所述共同关注焦点作为向所述用户终端推荐的关联内容显示给所述用户终端。

[0012] 其中,用户终端的换屏动作包括用户改变屏显页面内容的滚动或者滑动动作,所述根据用户滚动以及滑动动作,获取用户当前显示页面的内容,具体指用户当前显示屏所能显示的内容。由于屏幕分辨率尺寸的不同,每个用户一个屏幕页面所能显示出的内容不尽相同,这里需要根据滚动条位置,屏幕分辨率计算出当前用户屏幕所显示出的内容。

[0013] 另一方面,本发明还提供一种关联内容实时展现系统,包括:

[0014] 屏显内容获取单元,用于根据用户终端的换屏动作,实时获取用户终端当前屏显页面的内容;

[0015] 焦点内容确定单元,用于根据所述屏显内容获取单元获取的用户终端当前屏显页面的内容确定所述屏显页面的焦点内容;

[0016] 关联信息缓存单元,用于关联地存储所述屏显页面的焦点内容和用户在所述屏显页面的点击内容;

[0017] 缓存信息管理单元,用于根据所述屏显页面的焦点内容判断在所述关联信息缓存单元中是否有与所述屏显页面的焦点内容相关联的缓存信息;如果有,获取所述缓存信息并作为向所述用户终端推荐的关联内容;否则,通过实时计算获取关注过所述屏显页面的焦点内容的其他用户群以及与所述用户终端的行为相似的其他用户群,确定其他用户群的共同关注焦点,并以所述共同关注焦点作为向所述用户终端推荐的关联内容;

[0018] 关联内容展现单元,用于把向所述用户终端推荐的关联内容显示给所述用户终端。

[0019] 为了实现上述以及相关目的,本发明的一个或多个方面包括后面将详细说明并在

权利要求中特别指出的特征。下面的说明以及附图详细说明了本发明的某些示例性方面。然而,这些方面指示的仅仅是可使用本发明的原理的各种方式中的一些方式。此外,本发明旨在包括所有这些方面以及它们的等同物。

## 附图说明

- [0020] 通过参考以下结合附图的说明及权利要求书的内容,并且随着对本发明的更全面理解,本发明的其它目的及结果将更加明白及易于理解。在附图中:
- [0021] 图 1 为利用传统隐性用户行为数据的抓取为用户推荐关联内容的流程示意图;
- [0022] 图 2 为根据本发明实施例的关联内容实时展现方法的流程示意图;
- [0023] 图 3 为根据本发明实施例的在滚动或者滑动操作下进行关联内容实时展现的流程示意图;
- [0024] 图 4 为根据本发明实施例的关联信息缓存模块的具体实现过程示意图;
- [0025] 图 5 为根据本发明实施例的关联内容实时展现系统的逻辑结构示意图。
- [0026] 在所有附图中相同的标号指示相似或相应的特征或功能。

## 具体实施方式

[0027] 其中,用户终端的换屏动作包括用户改变屏显页面内容的滚动或者滑动动作,可以通过鼠标或者触摸屏等输入装置实现。比如,用户在使用 PC 浏览器时利用鼠标滚轴或者鼠标拖拽滚轴查看一个网页中不同位置内容的操作,用户在使用手机、平板电脑等移动设备中使用触摸屏等设备滑动页面查看不同位置的内容等。

[0028] 其中,实时计算获取关注过该焦点内容的其他用户群以及与该用户行为相似的其他用户群,从而获取其他用户群的共同关注焦点;其具体实现主要包括获取最近邻居(包括关注过该焦点内容的用户群以及行为相似的用户群)以及从最近邻居中的共同行为中发现用户可能感兴趣的内容。其中,利用关注焦点内容的用户群来获取跟焦点内容最相关的内容,而利用行为相似的用户群来获取跟该用户最相关的内容。

[0029] 具体地,作为示例,图 3 示出了在滚动或者滑动操作下进行关联内容实时展现的流程示意图。如图 3 所示,在用户根据浏览需要打开一个 URL 页面(步骤 S310)后,首先获取用户终端当前显示页面的内容,具体指用户当前显示屏所能显示的内容(S320)。由于各个用户终端的屏幕分辨率以及屏幕尺寸的不同,每个用户一个屏幕页面下所能显示出的内容不尽相同,可以根据滚动条位置、屏幕分辨率计算出当前用户屏幕所显示出的内容。

[0030] 然后,在步骤 S330 中,根据所获取的当前页面显示内容获取用户关注的

[0031] 焦点内容,具体是指在一个页面下用户最有可能聚焦的位置。由于页面的内容的多样性和用户关注的焦点的不确定性,用户在一个页面中关注的焦点并不是平均分布的,而是成 log-Normal 分布。用户浏览一个网页的视线分布可以通过用户的点击动作来拟合,用户点击位置与用户关注的焦点是相同的,所以利用用户的点击流数据以及点击位置信息,可以有效地拟合出用户在不同页面视线分布,从而利用概率理论发现用户最有可能关注的焦点。

[0032] 随着网页设计的改版变化,用户关注的焦点分布有可能会产生变化,利用实时点击流数据的拟合方法,可以实时获得最为准确的焦点分布。其具体计算方法为,假设用户关

注的焦点概率可以利用如下 log-Normal 分布, 这里需要利用 Multivariate Log-Normal 分布 :

$$F(X, D, \mu) = (2\pi)^{-\frac{P}{2}} * |D|^{-\frac{1}{2}} * [X_1, X_2, \dots, X_P]^{-1} * \exp\left(-\frac{(\ln X - \mu)'^D (\ln X - \mu)}{2}\right)$$

[0033]

[0034] 其中, P 是 X 的维数 (在显示屏幕中, 维数为 2, X 是屏幕点的坐标其中左下角为原点 (0,0)), D 和  $\mu$  分别是 X 概率分布方程的参数,  $\ln X = [\ln x_1, \ln x_2, \dots, \ln x_p]$ 。所以在屏幕 (2 维空间) 上的 log-normal 分布为 :

[0035]

$$F(x, y) = (2\pi)^{-1} * |D|^{-\frac{1}{2}} * [x, y]^{-1} * \exp\left(-\frac{([\ln x, \ln y] - \mu)'^D ([\ln x, \ln y] - \mu)}{2}\right)$$

[0036] 从上述公式可以发现, log-normal 的概率分布图取决于两个参数 D 和  $\mu$ , 所以 log-normal 的函数拟合, 要根据已有的用户点击位置, 来估算出上述公式中的参数, 此处用最大似然估算方法得到 :

[0037] 设已有用户 n 个点击坐标集为  $[(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)]$ , 分别计算出该集合的平均值以及方差,

$$[0038] \quad \bar{x} = \frac{\sum_i x_i}{n}, \bar{y} = \frac{\sum_i y_i}{n}$$

$$[0039] \quad s_{xx} = \text{var}(x) = \sum_i (x_i - \bar{x})^2 / n, s_{yy} = \text{var}(y) = \sum_i (y_i - \bar{y})^2 / n$$

$$[0040] \quad s_{xy} = s_{yx} = \text{var}(x, y) = \sum_i (x_i - \bar{x})(y_i - \bar{y}) / n$$

[0041] 则参数 D 跟  $\mu$  可以利用一下公式估算而得 :

$$[0042] \quad \mu = [\mu_x, \mu_y] = \left[ \exp(\bar{x} + \frac{s_{xx}^2}{2}), \exp(\bar{y} + \frac{s_{yy}^2}{2}) \right]$$

[0043]

$$\begin{aligned} D &= \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix} \\ &= \begin{bmatrix} \exp(2\bar{x} + \frac{s_{xx}^2}{2})(\exp(s_{xx}^2) - 1) & \frac{\exp\left[(\bar{x} + \bar{y}) + (s_{xx}^2 + s_{yy}^2)\right]}{2}(\exp(s_{xy}^2) - 1) \\ \frac{\exp\left[(\bar{x} + \bar{y}) + (s_{xx}^2 + s_{yy}^2)\right]}{2}(\exp(s_{xy}^2) - 1) & \exp(2\bar{y} + \frac{s_{yy}^2}{2})(\exp(s_{yy}^2) - 1) \end{bmatrix} \end{aligned}$$

[0044] 利用上述拟合的 log-Normal 函数, 就可以获得用户屏幕可能关注的焦点概率分

布,概率最大的位置就是用户最可能关注的位置。

[0045] 在获取了用户关注的焦点内容之后,进一步判断在焦点信息关联缓存中是否有与该焦点内容相关联的缓存信息(步骤S340);如果有,则从焦点信息关联缓存中获取缓存信息(步骤S351);如果没有,实时计算获取关注过该焦点内容的其他用户群(S352)以及与该用户行为相似的其他用户群(S353),从而获取其他用户群的共同关注焦点。

[0046] 其中,在焦点信息关联缓存中判断是否有与该焦点内容相关联的缓存信息,其主要作用是为了提高推荐内容的获取速度。由于滚动或者滑动关联推荐对于实时的要求非常高,如果推荐的关联内容不能及时推送展现给用户,用户很有可能将屏显内容滚动或者滑动到页面的其他位置,导致推荐的内容与当前的显示的内容存在延时或者用户无法查看到推荐的内容,为了解决该问题,采用关联信息缓存模块来提高推荐信息的获取速度。

[0047] S360:根据上面描述的存储结构进行查询,根据焦点信息查询到历史点击过的关联内容,并根据点击次数进行倒排序;

[0048] S370:根据比对方法发现的关联内容的推荐指数进行排序显示,显示推荐的关联内容,其具体内容为根据比对方法发现的关联内容的推荐指数进行排序显示(推荐指数越高排序越靠前),同时显示的关联内容不会随着用户的滚动滑动操作而产生位置变化,其内容会始终显示在用户的当前页面;

[0049] S380:根据系统推荐其估算出的焦点相关内容,用户发现自己感兴趣的推荐内容并点击链接查看其具体细节;

[0050] S390:判断用户滑动/滚动操作,具体包括用户在使用PC浏览器时利用鼠标滚轴或者鼠标拖拽滚轴查看一个网页中不同位置内容的操作,也包括用户在使用手机、平板电脑等移动设备中使用触摸屏等设备滑动页面查看不同位置的内容。如果检测到滑动/滚动操作则重复以上步骤,否则持续进行实时用户滑动/滚动操作检测。

[0051] 使用关联信息缓存模块存储用户关注的焦点内容并判断焦点内容是否有缓存的关联信息;如果有,获取缓存信息;如果没有,实时计算获取关注过该焦点内容的其他用户群以及与该用户行为相似的其他用户群,从而获取其他用户群的共同关注焦点。本发明中采用关联信息缓存模块来提高推荐信息的获取速度,该模块的具体实现步骤主要分为缓存信息写入以及缓存信息查询。图4示出了根据本发明实施例的关联信息缓存模块的具体实现步骤。

[0052] 如图4所示,根据本发明实施例的关联信息缓存模块的关联信息缓存以及关联信息查询的过程如下:

[0053] 步骤S411-S412:对于缓存信息写入,主要根据用户对于推荐内容的历史点击行为,其具体步骤可以描述为,当用户滚动浏览一个页面时,系统根据估算出的用户可能关注的焦点推荐出相关内容,用户发现自己感兴趣的推荐内容会点击链接查看其具体细节,这时把点击的内容跟用户当前关注的焦点内容进行关联存储(一条记录包含焦点内容作为Key,点击内容作为Value),同时统计该点击内容的历史点击次数(包括其他用户的点击次数),当点击次数超过一定阈值后(例如超过10次),就把该关联存储内容(焦点内容作为Key,点击内容以及其点击次数作为Value)放入到缓存中以备推荐内容的快速查询获取,同时为了保证缓存大小不会无限制扩大,当缓存容量超过一定阈值后,点击次数最低的记录会从缓存中移除;

[0054] 步骤 S420 :缓存信息查询,主要根据上面描述的存储结构进行查询,根据焦点信息查询到历史点击过的关联内容,并根据点击次数进行倒排序。

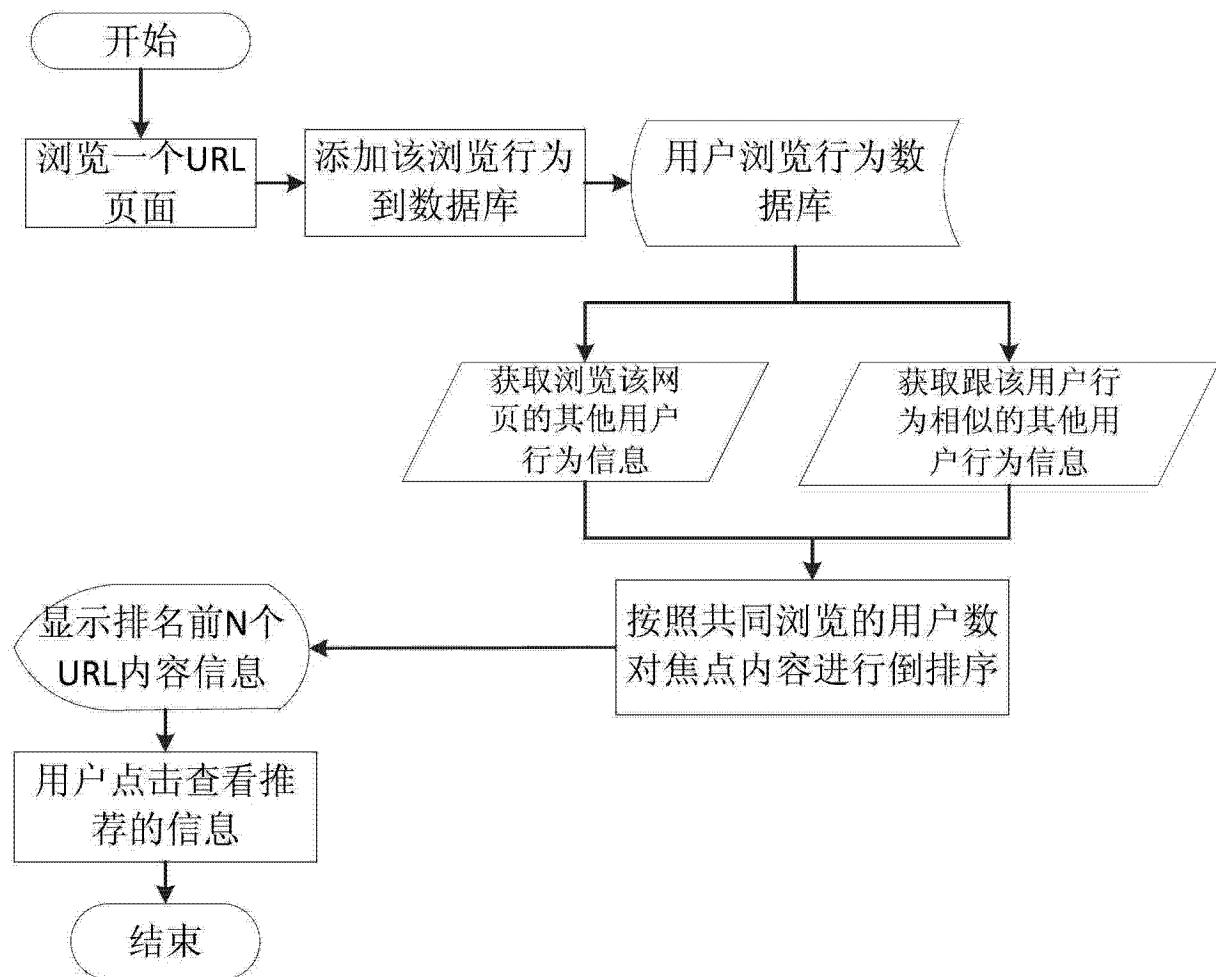


图 1

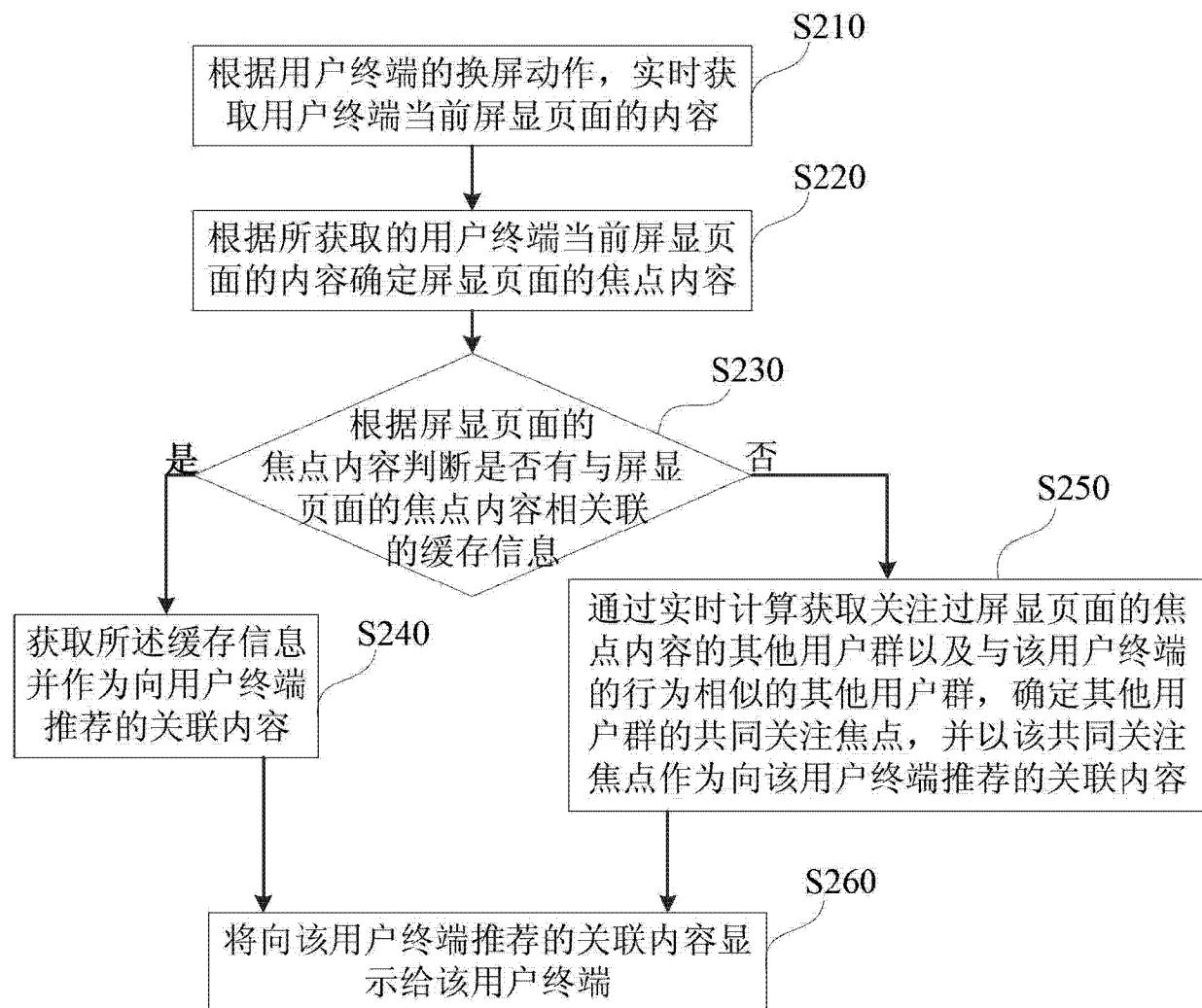


图 2

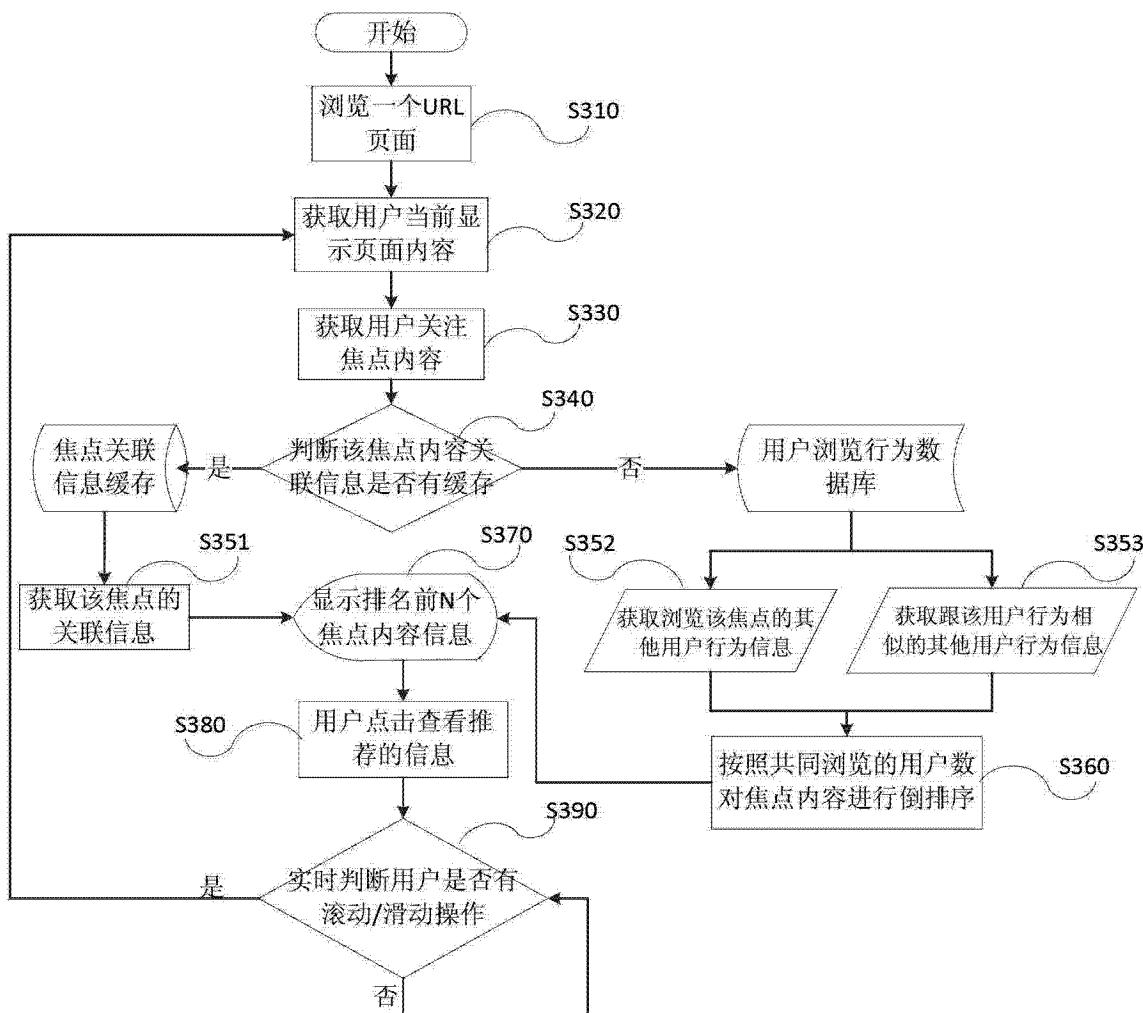


图 3

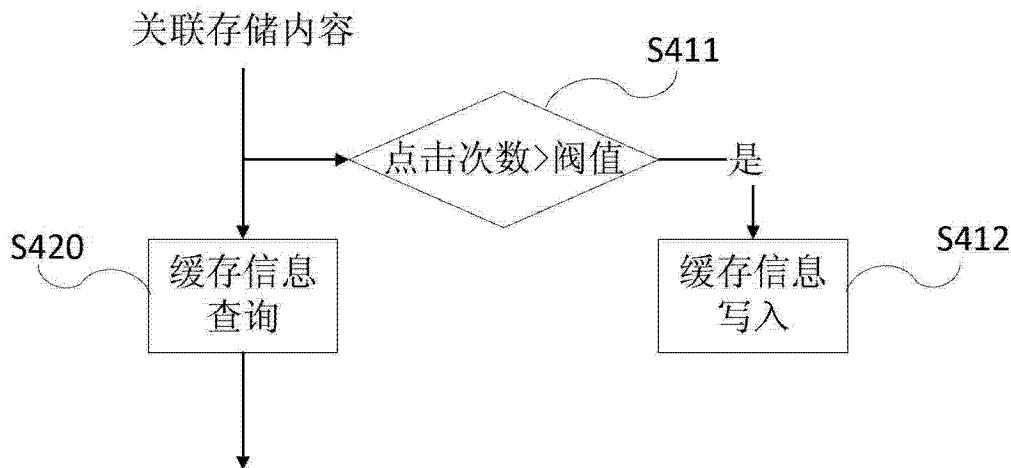


图 4

### 关联内容实时展现系统500

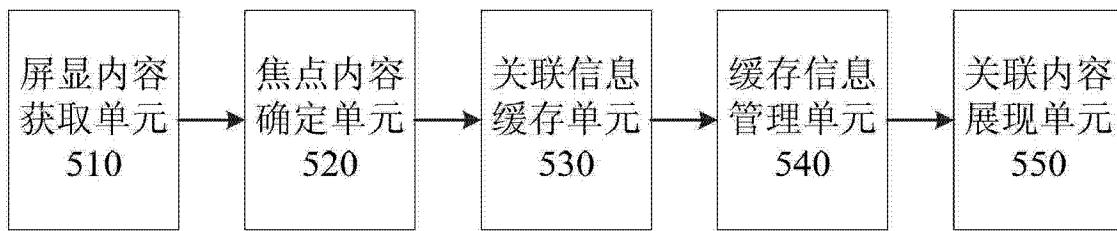


图 5