



(12) 发明专利申请

(10) 申请公布号 CN 105656997 A

(43) 申请公布日 2016. 06. 08

(21) 申请号 201510994480. 2

(22) 申请日 2015. 12. 25

(71) 申请人 中国科学院信息工程研究所

地址 100093 北京市海淀区闵庄路甲 89 号

(72) 发明人 安冬

(74) 专利代理机构 北京路浩知识产权代理有限公司 11002

代理人 李相雨

(51) Int. Cl.

H04L 29/08(2006. 01)

H04W 28/14(2009. 01)

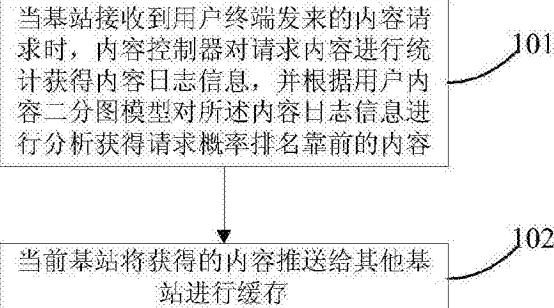
权利要求书2页 说明书8页 附图3页

(54) 发明名称

基于移动用户关系的热度缓存内容主动推送方法

(57) 摘要

本发明提供一种基于移动用户关系的热度缓存内容主动推送方法，包括：当基站接收到用户终端发来的内容请求时，内容控制器对请求内容进行统计获得内容日志信息，并根据用户内容二分图模型对所述内容日志信息进行分析获得请求概率排名靠前的内容；当前基站将获得的内容推送给其他基站进行缓存。本发明通过基站将其缓存的请求热度较高的内容推送给其他基站进行缓存，实现了内容热度“迁移”，充分利用系统资源，不但可以预知用户访问内容的可能性，同时，有利于减轻网络的压力，提高缓存命中率。采用这种技术，是对移动网络缓存系统设计的一种新尝试，能够适应新型移动网络的发展趋势。



1. 一种基于移动用户关系的热度缓存内容主动推送方法,其特征在于,包括:

当基站接收到用户终端发来的内容请求时,内容控制器对请求内容进行统计获得内容日志信息,并根据用户内容二分图模型对所述内容日志信息进行分析获得请求概率排名靠前的内容;

当前基站将获得的内容推送给其他基站进行缓存。

2. 根据权利要求1所述的推送方法,其特征在于,当多个基站处于同一所设基站区域内,从当前基站区域内所有基站的内容日志信息中获得请求概率排名靠前的内容,并推送到其他基站区域内的基站下进行缓存。

3. 根据权利要求1所述的推送方法,其特征在于,根据用户内容二分图模型获得请求热度排名靠前的内容,包括:

根据内容控制器获取的内容日志信息建立用户内容二分图模型;

根据用户内容二分图模型获得基站下各个访问用户的活跃度和关联度,并根据所述活跃度和所述关联度获得用户排名信息;

获取用户排名信息排名靠前的用户在预设时间段内所请求的内容,并设定为暂定推送新内容;

对获得的暂定推送新内容根据内容日志信息和用户排名信息获得概率矩阵

$$\begin{pmatrix} p_{11} & \cdots & p_{1C} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NC} \end{pmatrix},$$

其中,N表示基站下用户个数,C表示请求内容个数,矩阵每行表示任一用

户请求各个内容的概率,若用户对请求内容有访问,则对应概率值为用户排名信息上的数值,反之为0;

根据概率矩阵通过 $p_j = \sum_{i=1}^N p_{ij}$ 获得其他基站用户对暂定推送新内容的请求概率矩阵 (p_1, p_2, \dots, p_C) , 其中, p_j 为其他基站用户对某一暂定推送新内容的请求概率, p_{ij} 为当前基站用户对某一暂定推送新内容的请求概率;

根据请求概率矩阵 (p_1, p_2, \dots, p_C) 获取请求概率排名靠前的内容,设定为推送新内容。

4. 根据权利要求2所述的推送方法,其特征在于,当多个基站处于同一所设基站区域内,根据用户内容二分图模型获得请求热度排名靠前的内容,包括:

根据内容控制器获取的内容日志信息建立用户内容二分图模型;

根据用户内容二分图模型获得基站下各个访问用户的活跃度和关联度,并根据所述活跃度和所述关联度获得用户排名信息;

获取当前基站区域内用户排名信息排名靠前的用户在预设时间段内所请求的内容,并设定为暂定推送新内容;

对获得的暂定推送新内容根据内容日志信息和用户排名信息获得概率矩阵

$$\begin{pmatrix} p_{11} & \cdots & p_{1C} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NC} \end{pmatrix},$$

其中,N表示当前基站区域下用户个数,C表示请求内容个数,矩阵每行表

示任一用户请求各个内容的概率,若用户对请求内容有访问,则对应概率值为用户排名信

息上的数值,反之为0;

根据概率矩阵通过 $p_j = \sum_{i=1}^N p_{ij}$ 获得其他基站区域用户对暂定推送新内容的请求概率矩阵 (p_1, p_2, \dots, p_c) , 其中, p_j 为其他基站区域用户对某一暂定推送新内容的请求概率, p_{ij} 为当前基站区域用户对某一暂定推送新内容的请求概率;

根据请求概率矩阵 (p_1, p_2, \dots, p_c) 获取请求概率排名靠前的内容, 设定为推送新内容。

5. 根据权利要求3或4所述的推送方法, 其特征在于, 对请求概率矩阵 (p_1, p_2, \dots, p_c) 根据 $q_j = p_j / \sum_{i=1}^c p_i$ 进行归一化处理获得归化概率矩阵 (q_1, q_2, \dots, q_c) 。

6. 根据权利要求3或4所述的推送方法, 其特征在于, 所述用户排名为用户活跃度与用户关联度的乘积, 其中, 所述用户活跃度为当前用户请求内容的个数; 所述用户关联度为当前用户与其他用户有共同请求行为的次数之和。

7. 根据权利要求3或4所述的推送方法, 其特征在于, 在建立用户内容二分图模型之前, 对基站下的访问用户确定是否为稳定用户, 其中, 所述稳定用户为基站下的访问用户在预设时间内均有访问记录的用户。

8. 根据权利要求1所述的推送方法, 其特征在于, 若当前推送的内容在当前基站内已缓存, 则推送下一个要推送的内容, 直到所有推送内容均被处理; 反之, 在当前基站缓存。

9. 根据权利要求8所述的推送方法, 其特征在于, 向基站推送的内容按请求概率排名的高低顺序进行推送。

10. 根据权利要求7所述的推送方法, 其特征在于, 所述稳定用户为基站下的访问用户在一周时间内至少四天均有访问记录的用户。

基于移动用户关系的热度缓存内容主动推送方法

技术领域

[0001] 本发明涉及网络通信技术领域,尤其涉及一种基于移动用户关系的热度缓存内容主动推送方法。

背景技术

[0002] 随着LTE无线网络的快速部署和智能终端及移动应用的极大丰富,移动网络数据流量迅速增加,这给现有网络架构带来了很大的技术冲击。为此,在基站侧部署缓存成为应对上述挑战的重要手段之一。当前无线网络基站中使用的缓存机制是传统的“被动式”缓存(reactive caching)模型,即当移动用户发起对内容的请求后基站才有条件的进行缓存。由于该模型被动服务的特点,“被动式”缓存在平抑网络流量方面存在一定的局限性,尤其是在平抑网络高峰时段的流量方面。比如,如果基站根据传输协议发现该内容是可缓存的并且该内容缓存到本地有较大利用价值,基站才会缓存该内容。如果本地缓存空间有限,则需要基于传统算法,例如LFU,LRU算法,对缓存空间进行管理。传统算法需要经过一定时间的计算,某一个内容才有可能被缓存下来。如果网络处于空闲状态,此时如果没有缓存该内容,缓存该内容对网络性能造成的影响会相对较小,但是,如果网络处于忙时,大量的内容请求就会造成网络质量的进一步下降。

[0003] 图1是基站侧缓存的示意图。如图1所示,每个基站下面存在大量访问用户,这些用户每天都会访问互联网内容,互联网内容的获取需要经过基站。由于移动基站具备缓存能力,可实现热门内容的本地化获取,因此可以减轻移动网络的流量压力。传统reactive caching机制,根据最近访问频率、最近访问时间、内容热度等因素,设计相关缓存替换算法决定是否需缓存相应的内容。传统reactive caching模型,是基于用户请求驱动的,即当用户发出实际请求时才会进行相应的缓存查询、回源等操作。因此,难以进一步平抑网络忙时的流量,难以充分利用“忙”、“闲”时段,实现网络流量的均衡化,达到“削峰填谷”的目的。同时,随着缓存能力下沉到距离用户一跳的基站,有必要通过对用户行为的挖掘,进行精确的个性化推送,进一步提升用户体验。这些都是传统的reactive caching机制难以实现的。

发明内容

[0004] 本发明提供一种基于移动用户关系的热度缓存内容主动推送方法,达到了提高网络缓存质量的目的。

[0005] 为了达到以上目的,本发明提供一种基于移动用户关系的热度缓存内容主动推送方法,包括:

[0006] 当基站接收到用户终端发来的内容请求时,内容控制器对请求内容进行统计获得内容日志信息,并根据用户内容二分图模型对所述内容日志信息进行分析获得请求概率排名靠前的内容;

[0007] 当前基站将获得的内容推送给其他基站进行缓存。

[0008] 优选地,当多个基站处于同一所设基站区域内,从当前基站区域内所有基站的内

容日志信息中获得请求概率排名靠前的内容，并推送到其他基站区域内的基站下进行缓存。

- [0009] 优选地，根据用户内容二分图模型获得请求热度排名靠前的内容，包括：
- [0010] 根据内容控制器获取的内容日志信息建立用户内容二分图模型；
- [0011] 根据用户内容二分图模型获得基站下各个访问用户的活跃度和关联度，并根据所述活跃度和所述关联度获得用户排名信息；
- [0012] 获取用户排名信息排名靠前的用户在预设时间段内所请求的内容，并设定为暂定推送新内容；
- [0013] 对获得的暂定推送新内容根据内容日志信息和用户排名信息获得概率矩阵

$$\begin{pmatrix} p_{11} & \cdots & p_{1C} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NC} \end{pmatrix}$$

，其中，N表示基站下用户个数，C表示请求内容个数，矩阵每行表示任一用

户请求各个内容的概率，若用户对请求内容有访问，则对应概率值为用户排名信息上的数值，反之为0；

- [0014] 根据概率矩阵通过 $p_j = \sum_{i=1}^N p_{ij}$ 获得其他基站用户对暂定推送新内容的请求概率矩阵 (p_1, p_2, \dots, p_C) ，其中， p_j 为其他基站用户对某一暂定推送新内容的请求概率， p_{ij} 为当前基站用户对某一暂定推送新内容的请求概率；

- [0015] 根据请求概率矩阵 (p_1, p_2, \dots, p_C) 获取请求概率排名靠前的内容，设定为推送新内容。
- [0016] 优选地，当多个基站处于同一所设基站区域内，根据用户内容二分图模型获得请求热度排名靠前的内容，包括：

- [0017] 根据内容控制器获取的内容日志信息建立用户内容二分图模型；
- [0018] 根据用户内容二分图模型获得基站下各个访问用户的活跃度和关联度，并根据所述活跃度和所述关联度获得用户排名信息；
- [0019] 获取当前基站区域内用户排名信息排名靠前的用户在预设时间段内所请求的内容，并设定为暂定推送新内容；
- [0020] 对获得的暂定推送新内容根据内容日志信息和用户排名信息获得概率矩阵

$$\begin{pmatrix} p_{11} & \cdots & p_{1C} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NC} \end{pmatrix}$$

，其中，N表示当前基站区域下用户个数，C表示请求内容个数，矩阵每行表

示任一用户请求各个内容的概率，若用户对请求内容有访问，则对应概率值为用户排名信息上的数值，反之为0；

- [0021] 根据概率矩阵通过 $p_j = \sum_{i=1}^N p_{ij}$ 获得其他基站区域用户对暂定推送新内容的请求概率矩阵 (p_1, p_2, \dots, p_C) ，其中， p_j 为其他基站区域用户对某一暂定推送新内容的请求概率， p_{ij} 为当前基站区域用户对某一暂定推送新内容的请求概率；
- [0022] 根据请求概率矩阵 (p_1, p_2, \dots, p_C) 获取请求概率排名靠前的内容，设定为推送新

内容。

[0023] 优选地，对请求概率矩阵(p_1, p_2, \dots, p_c)根据 $q_j = p_j / \sum_{i=1}^c p_i$ 进行归一化处理获得归化概率矩阵(q_1, q_2, \dots, q_c)。

[0024] 优选地，所述用户排名为用户活跃度与用户关联度的乘积，其中，所述用户活跃度为当前用户请求内容的个数；所述用户关联度为当前用户与其他用户有共同请求行为的次数之和。

[0025] 优选地，在建立用户内容二分图模型之前，对基站下的访问用户确定是否为稳定用户，其中，所述稳定用户为基站下的访问用户在预设时间内均有访问记录的用户。

[0026] 优选地，若当前推送的内容在当前基站内已缓存，则推送下一个要推送的内容，直到所有推送内容均被处理；反之，在当前基站缓存。

[0027] 优选地，向基站推送的内容按请求概率排名的高低顺序进行推送。

[0028] 优选地，所述稳定用户为基站下的访问用户在一周时间内至少四天均有访问记录的用户。

[0029] 由上述技术方案可知，本发明通过基站将其缓存的请求热度较高的内容推送给其他基站进行缓存，实现了内容热度“迁移”，充分利用系统资源，不但可以预知用户访问内容的可能性，同时，有利于减轻网络的压力，提高缓存命中率。采用这种技术，是对移动网络缓存系统设计的一种新尝试，能够适应新型移动网络的发展趋势。

附图说明

- [0030] 图1为现有技术中基站侧缓存示意图；
- [0031] 图2为本发明实施例1和实施例2提供的所示推送方法的流程示意图；
- [0032] 图3为本发明实施例1提供的获取推送新内容的流程示意图；
- [0033] 图4为本发明实施例2提供的获取推送新内容的流程示意图；
- [0034] 图5为本发明实施例1和实施例2所述用户内容二分图模型示意图；
- [0035] 图6为本发明实施例2基站区域分类示意图。

具体实施方式

[0036] 下面结合附图和实施例，对本发明的具体实施方式作进一步详细描述。以下实施例用于说明本发明，但不用来限制本发明的范围。

[0037] 实施例1：

[0038] 图2示出了本发明一实施例提供的一种基于移动用户关系的热度缓存内容主动推送方法，包括：

[0039] 101、当基站接收到用户终端发来的内容请求时，内容控制器对请求内容进行统计获得内容日志信息，并根据用户内容二分图模型对所述内容日志信息进行分析获得请求概率排名靠前的内容；

[0040] 102、当前基站将获得的内容推送给其他基站进行缓存。

[0041] 需要说明的是，根据地域原因，会在不同的位置设置基站，在每个基站环境下会存在很多的访问用户进行网络访问请求。因此，对于上述步骤101所述的获取请求率排名靠前

的请求内容,如图3所示,包括:

[0042] A1011、根据内容控制器获取的内容日志信息建立用户内容二分图模型。在本步骤中,内容控制器用于记录基站下用户HTTP访问的内容日志信息。在日志文件中,每一条记录需要包含用户ID、内容URL、服务基站IP、内容大小、内容类型、请求开始时间和请求结束时间等字段。这些信息存储在内容控制器中,以便后续计算使用。在本步骤中,用户二分图模型的建立就是建立在内容日志信息基础上。在这里,内容控制器可以根据内容日志信息建立模型,也可以在基站具有处理能力的情况下,使基站自身建立所需模型。有关二分图模型已是公知常识,如图5所示,为基于内容日志信息的二分图模型。在图5中,大写英文字母A, B,C,D表示用户,小写英文字母a,b,c,d,e,f表示内容,二者之间的连线表示某一个用户请求过某一个内容,从图5中容易看出用户A请求内容列表为{a,b,d},个数为3个,用户B请求内容列表为{a,c,f},个数为3个,用户C请求内容列表为{b,e},个数为2个,用户D请求内容列表为{d,f},个数为2个。

[0043] 需要说明的是,在建立用户内容二分图模型之前,还可以先对基站下的访问用户确定是否为稳定用户。对于每个基站都存在稳定用户(经常性访问的用户)和临时访客用户(临时访问用户)。为了减少临时访问用户对系统建模的影响,需要求解每个基站下的活跃稳定用户,这里的稳定用户是指在当前基站下用户在多天之内有访问记录的情况。在这里,可以选定一周的数据来计算稳定用户数量:比如基站下的稳定用户定义为在过去一周时间之内或在一周时间内不少于4天都有访问记录的用户。

[0044] A1012、根据用户内容二分图模型获得基站下各个访问用户的活跃度和关联度,并根据所述活跃度和所述关联度获得用户排名信息。在本步骤中,需要计算每个访问用户的活跃度和关联度。由图5可以得知,不同用户在一段时间之内所请求的个数是不同的,同时对于同一个内容,会发生不同用户请求的情况。

[0045] 在这里,定义用户活跃度(User Activity)为用户二分图中的出度为: $UA(u)=d(u)$ 。其中,u表示某个用户,d表示出度,在图5中,对于用户而言,其只有出度,所有的入度为0,同理,对于内容而言,其只有入度,所有的出度为0。对于用户A来讲,其活跃度为3,其他用户依次是,用户B活跃度为3,用户C活跃度为2,用户D活跃度为2。注意这里即使用户多次重复访问同一内容,在二分图中其依然只是关系表示,并不会影响出度值,也就是说如某用户多次访问同一内容,在计算出度的时候,只记为发生了一次这样的请求,在这里标记的是关系,而不是次数。

[0046] 在这里,定义用户关联度(User Correlation)为该用户和所有用户有共同请求行为的次数之和,即:

$$[0047] UC(u) = \sum_{c=1}^C \sum_{v=1}^N content(u, v, c)$$

[0048] 其中v不等于u,C表示内容总数,N表示稳定用户总数,其中content为记录内容被用户访问的列表,记录的是某个内容被哪些用户访问过,计算公式如下:

$$[0049] content(u, v, c) = \begin{cases} 1, & \text{如果列表含有用户 } u, v \\ 0, & \text{如果列表不含用户 } u, v \end{cases}$$

[0050] 在这里,共同请求行为是指两个用户都请求某个内容。如图5所示,用户A和用户B

有共同请求行为,次数为1,仅仅都请求过a内容,用户A和用户C有共同请求行为,次数为1,仅仅都请求过b内容,用户A和用户D有共同请求行为,次数为1,仅仅都请求过d内容,由此可以得出用户A的用户关联度为用户A和所有用户有共同行为的次数之和:1+1+1=3。

[0051] 综上,用户排名(UR)为用户活跃度(UA)与用户关联度(UC)的积,即:UR(u)=UA(u)×UC(u)。

[0052] A1013、获取用户排名信息排名靠前的用户在预设时间段内所请求的内容,并设定为暂定推送新内容。在本步骤中,暂定推送新内容在这里仅是为了区分而命名的。其可以直接作为推送内容进行推送,也可以在其里面选择所需的内容进行推送。具体内容可在下面步骤中详细介绍。本步骤中所说的在排名信息排名靠前的用户所请求的内容均被设定为暂定推送新内容。

[0053] A1014、对获得的暂定推送新内容根据内容日志信息和用户排名信息获得概率矩阵。

阵 $\begin{pmatrix} p_{11} & \cdots & p_{1C} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NC} \end{pmatrix}$,其中,N表示基站下用户个数,C表示请求内容个数,矩阵每行表示任一

用户请求各个内容的概率,若用户对请求内容有访问,则对应概率值为用户排名信息上的数值,反之为0。以图5为例,用户A的排名值为9,用户B的排名值为6,用户C的排名值为2,用户D的排名值为4。则概率矩阵为:

$$[0054] \quad \begin{bmatrix} 9 & 9 & 0 & 9 & 0 & 0 \\ 6 & 0 & 6 & 0 & 0 & 6 \\ 0 & 2 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 & 0 & 4 \end{bmatrix}$$

[0055] A1015、根据概率矩阵通过 $p_j = \sum_{i=1}^N p_{ij}$ 获得其他基站用户对暂定推送新内容的请求

概率矩阵(p_1, p_2, \dots, p_c),其中, p_j 为其他基站用户对某一暂定推送新内容的请求概率, p_{ij} 为当前基站用户对某一暂定推送新内容的请求概率。在本步骤中,以步骤A1014获得概率矩阵为基础,其他基站用户对暂定推送新内容的请求概率矩阵为(17,11,6,13,2,10)。

[0056] A1016、根据请求概率矩阵(p_1, p_2, \dots, p_c)获取请求概率排名靠前的内容,设定为推送新内容。在本步骤获取排名靠前的内容之前,还可以对请求概率矩阵(p_1, p_2, \dots, p_c)根

据 $q_j = p_j / \sum_{i=1}^c p_i$ 进行归一化处理获得归化概率矩阵(q_1, q_2, \dots, q_c)。进行归一化处理之后,

可以从概率矩阵根据高低排列获得所需推送的内容。需要说明的是,概率排名越靠前表示该内容请求热度越高。

[0057] 若当前推送的内容在当前基站内已缓存,则推送下一个要推送的内容,直到所有推送内容均被处理;反之,在当前基站缓存。

[0058] 实施例2:

[0059] 图2示出了本发明一实施例提供的一种基于移动用户关系的热度缓存内容主动推送方法,包括:

[0060] 101、当基站接收到用户终端发来的内容请求时,内容控制器对请求内容进行统计

获得内容日志信息，并根据用户内容二分图模型对所述内容日志信息进行分析获得请求概率排名靠前的内容；

[0061] 102、当前基站将获得的内容推送给其他基站进行缓存。

[0062] 需要说明的是，根据地域原因，会对基站进行类型划分，是为了利用不同类型基站内容热度形成时间上的差异，从而将在某类基站上先行涌现的新热门内容“迁移”到其他基站上。如图6所示，可以看出分出了两个基站区域，即住宅区（包括居民区、校园等）和商业园区（包括工业园、CBD等），说明热度预测机制。对于上述两种类型的基站来讲，新热门内容的涌现时间会略有差异。举例来说，如果在早高峰时段，住宅区的新热门内容的涌现时间会略早于商业园区，那么，我们就可以在此期间将住宅区新涌现的热门内容提前推送到商业园区基站上。因此，对于上述步骤101所述的获取请求率排名靠前的请求内容，如图4所示，包括：

[0063] B1011、根据内容控制器获取的内容日志信息建立用户内容二分图模型。在本步骤中，内容控制器用于记录基站下用户HTTP访问的内容日志信息。在日志文件中，每一条记录需要包含用户ID、内容URL、服务基站IP、内容大小、内容类型、请求开始时间和请求结束时间等字段。这些信息存储在内容控制器中，以便后续计算使用。在本步骤中，用户二分图模型的建立就是建立在内容日志信息基础上。在这里，内容控制器可以根据内容日志信息建立模型，也可以在基站具有处理能力的情况下，使基站自身建立所需模型。有关二分图模型已是公知常识，如图5所示，为基于内容日志信息的二分图模型。在图5中，大写英文字母A，B，C，D表示用户，小写英文字母a，b，c，d，e，f表示内容，二者之间的连线表示某一个用户请求过某一个内容，从图5中容易看出用户A请求内容列表为{a,b,d}，个数为3个，用户B请求内容列表为{a,c,f}，个数为3个，用户C请求内容列表为{b,e}，个数为2个，用户D请求内容列表为{d,f}，个数为2个。

[0064] 需要说明的是，在建立用户内容二分图模型之前，还可以先对基站下的访问用户确定是否为稳定用户。对于每个基站都存在稳定用户（经常性访问的用户）和临时访客用户（临时访问用户）。为了减少临时访问用户对系统建模的影响，需要求解每个基站下的活跃稳定用户，这里的稳定用户是指在当前基站下用户在多天之内有访问记录的情况。在这里，可以选定一周的数据来计算稳定用户数量：比如基站下的稳定用户定义为在过去一周时间之内或在一周时间内不少于4天都有访问记录的用户。

[0065] B1012、根据用户内容二分图模型获得基站下各个访问用户的活跃度和关联度，并根据所述活跃度和所述关联度获得用户排名信息。在本步骤中，需要计算每个访问用户的活跃度和关联度。由图5可以得知，不同用户在一段时间之内所请求的个数是不同的，同时对于同一个内容，会发生不同用户请求的情况。

[0066] 在这里，定义用户活跃度（User Activity）为用户二分图中的出度为： $UA(u) = d(u)$ 。其中， u 表示某个用户， d 表示出度，在图5中，对于用户而言，其只有出度，所有的入度为0，同理，对于内容而言，其只有入度，所有的出度为0。对于用户A来讲，其活跃度为3，其他用户依次是，用户B活跃度为3，用户C活跃度为2，用户D活跃度为2。注意这里即使用户多次重复访问同一内容，在二分图中其依然只是关系表示，并不会影响出度值，也就是说如某用户多次访问同一内容，在计算出度的时候，只记为发生了一次这样的请求，在这里标记的是关系，而不是次数。

[0067] 在这里,定义用户关联度(User Correlation)为该用户和所有用户有共同请求行为的次数之和,即:

$$[0068] UC(u) = \sum_{c=1}^C \sum_{v=1}^N content(u, v, c)$$

[0069] 其中v不等于u,C表示内容总数,N表示稳定用户总数,其中content为记录内容被用户访问的列表,记录的是某个内容被哪些用户访问过,计算公式如下:

$$[0070] content(u, v, c) = \begin{cases} 1, & \text{如果列表含有用户 } u, v \\ 0, & \text{如果列表不含用户 } u, v \end{cases}$$

[0071] 在这里,共同请求行为是指两个用户都请求某个内容。如图5所示,用户A和用户B有共同请求行为,次数为1,仅仅都请求过a内容,用户A和用户C有共同请求行为,次数为1,仅仅都请求过b内容,用户A和用户D有共同请求行为,次数为1,仅仅都请求过d内容,由此可以得出用户A的用户关联度为用户A和所有用户有共同行为的次数之和:1+1+1=3。

[0072] 综上,用户排名(UR)为用户活跃度(UA)与用户关联度(UC)的积,即:UR(u)=UA(u)×UC(u)。

[0073] B1013、获取当前基站区域内用户排名信息排名靠前的用户在预设时间段内所请求的内容,并设定为暂定推送新内容。在本步骤中,是要对每个基站区域内的用户进行排名,从而获得所有基站的排名信息。比如:基站区域A内设置4个基站,每个基站下均有5个访问客户,在本步骤中,需要在4*5个客户进行排名,获得排名信息。然后确定排名靠前的用户。所述暂定推送新内容在这里仅是为了区分而命名的。其可以直接作为推送内容进行推送,也可以在其里面选择所需的内容进行推送。具体内容可在下面步骤中详细介绍。本步骤中所说的在排名信息排名靠前的用户所请求的内容均被设定为暂定推送新内容。

[0074] B1014、对获得的暂定推送新内容根据内容日志信息和用户排名信息获得概率矩阵

$$\begin{pmatrix} p_{11} & \cdots & p_{1C} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NC} \end{pmatrix}, \text{其中,} N \text{表示区域基站下用户个数,} C \text{表示请求内容个数,矩阵每行表示}$$

任一用户请求各个内容的概率,若用户对请求内容有访问,则对应概率值为用户排名信息上的数值,反之为0。上述概率矩阵的计算与实施例1的计算相同,在此不再赘述。

[0075] B1015、根据概率矩阵通过 $p_j = \sum_{i=1}^N p_{ij}$ 获得其他区域基站用户对暂定推送新内容的

请求概率矩阵(p_1, p_2, \dots, p_C),其中, p_j 为其他区域基站用户对某一暂定推送新内容的请求概率, p_{ij} 为当前区域基站用户对某一暂定推送新内容的请求概率。

[0076] B1016、根据请求概率矩阵(p_1, p_2, \dots, p_C)获取请求概率排名靠前的内容,设定为推送新内容。在本步骤获取排名靠前的内容之前,还可以对请求概率矩阵(p_1, p_2, \dots, p_C)根

$$\text{据 } q_j = p_j / \sum_{i=1}^C p_i \text{ 进行归一化处理获得归化概率矩阵} (q_1, q_2, \dots, q_C) \text{。进行归一化处理之后,}$$

可以从概率矩阵根据高低排列获得所需推送的内容。需要说明的是,概率排名越靠前表示该内容请求热度越高。

[0077] 若当前推送的内容在当前基站内已缓存,则推送下一个要推送的内容,直到所有

推送内容均被处理；反之，在当前基站缓存。

[0078] 本发明面向移动网络中基站侧部署缓存的场景，实现在不同类型的基站之间进行热度迁移。由于各个基站所辖的用户的差异性和内容的动态变化性，热门内容涌现的时间会有一定的差异性，同时热门内容热度的形成需要一定的时间长度。本发明所提热度预测迁移是指，利用上述差异性，将某个或某些基站已经缓存的新的热门内容，推送至其它尚未缓存这些热门内容的基站上。

[0079] 为了提高热度迁移准确性，本发明提出了将基于图模型的相关算法可以应用到热度预测模型中，在这里使用图模型中的二分图概念来帮助建立热度预测模型。用户会大量请求内容，用户和内容之间的关系实际上是一个二分图，在用户请求基站下内容的过程中，一定会产生一些用户大量请求内容，而一部分用户只是少量请求内容的情况，同时用户习惯存在共性和差异性，基于以上逻辑可以通过相关算法挖掘用户的历史信息，建立内容热度预测模型，充分挖掘历史数据的潜在价值。该模型实现了内容热度“迁移”，充分利用系统资源，不但可以预知用户访问内容的可能性，同时，有利于减轻网络的压力，提高缓存命中率。采用这种技术，是对移动网络缓存系统设计的一种新尝试，能够适应新型移动网络的发展趋势。

[0080] 此外，本领域的技术人员能够理解，尽管在此所述的一些实施例包括其它实施例中所包括的某些特征而不是其它特征，但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如，在下面的权利要求书中，所要求保护的实施例的任意之一都可以以任意的组合方式来使用。

[0081] 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制，并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中，不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了若干装置的单元权利要求中，这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

[0082] 本领域普通技术人员可以理解：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各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，或者对其中部分或者全部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权利要求所限定的范围。

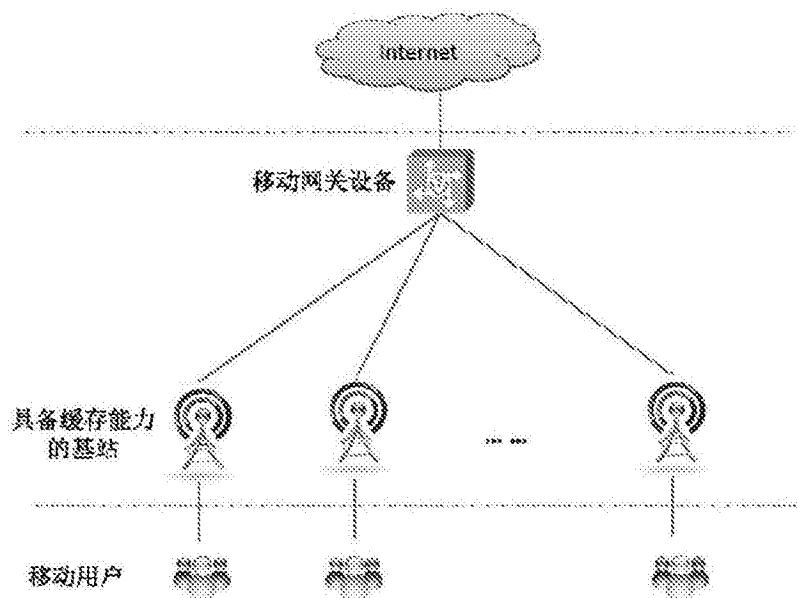


图1

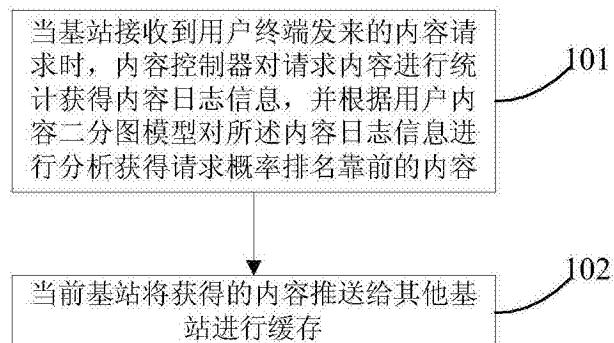


图2

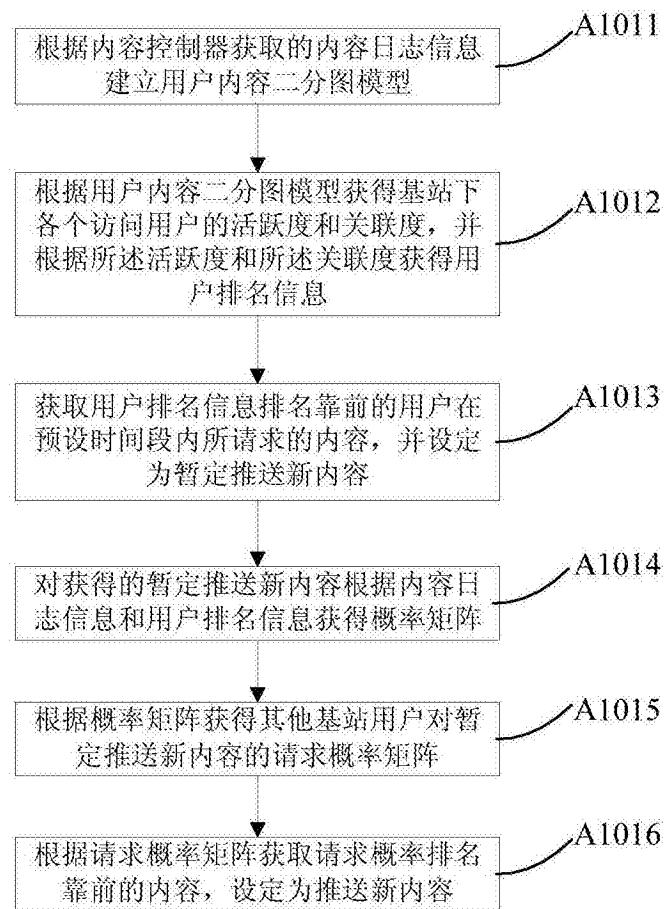


图3

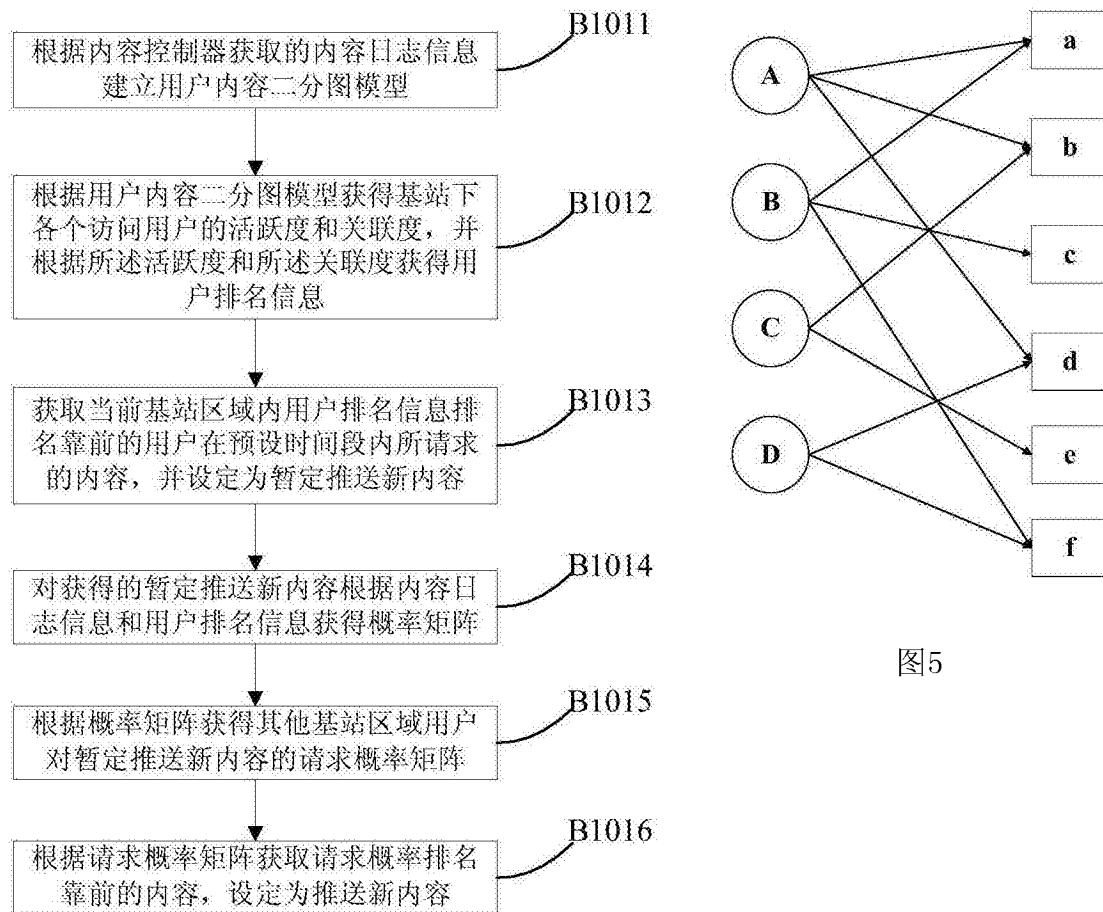


图4

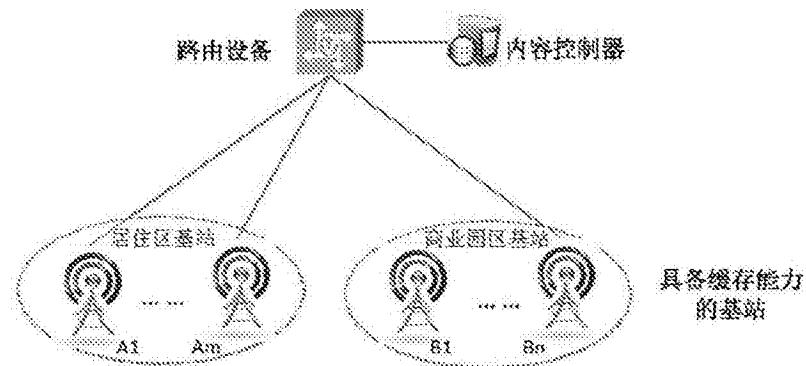


图6