

移动 P2P 网络预载—缓存策略研究

周欣欣^{1,2}, 卢 喆¹, 邹振婉¹, 宋人杰¹

(1 东北电力大学 信息工程学院, 吉林 吉林 132012;

2 中国矿业大学(北京) 机电与信息工程学院, 北京 100083)

摘 要: 为了提高缓存命中率, 减少用户访问时延, 提高网络服务质量, 针对具有超级节点的半分布式移动 P2P 网络, 提出一种基于兴趣-相关度的移动 P2P 预载策略, 通过对移动节点的历史访问记录统计并进行数据挖掘, 筛选出价值高的数据资源并进行预载; 为了避免频繁预载-缓存而导致的缓存命中率降低, 提出一种基于兴趣-相关度和信息素的缓存替换策略, 该策略充分考虑了用户对数据资源的兴趣趋向性。仿真结果表明, 所提出的预载与缓存替换策略能够有效提高缓存命中率。

关键词: 移动 P2P 网络; 预载; 兴趣-相关度; 缓存替换

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2015)11-0069-05

Strategy of Prefetch and Cache Replacement
for Mobile P2P NetworksZHOU Xin-xin^{1,2}, LU Zhe¹, ZOU Zhen-wan¹, SONG Ren-jie¹

(1 School of Information Engineering, Northeast Dianli University, Jilin 132012, China;

2 School of Mechanical Electronic & Information Engineering, China University of
Mining & Technology, Beijing 100083, China)

Abstract: In order to improve the cache hit rate, reduce the access delay and improve the network quality of service, an interest-based prefetch strategy for semi-distributed mobile P2P network is proposed. The history access records of mobile nodes is counted and the data resources which is with the highest value is prefetched. A cache replacement strategy based on interest-relevancy and pheromone is presented to void the reduction of cache hit rate that is caused by preload-cache frequently. The user interest tendency for data resource is adequately considered in this strategy. The simulation results show that the proposed strategies of prefetch and cache replacement can effectively improve the cache hit ratio.

Key words: mobile P2P network; prefetch; interest-relevancy; cache replacement

1 引言

移动 P2P 网络 (Mobile Peer-to-Peer Networks, MP2P) 作为一种新兴的移动数据共享方式, 为人们提供了一种在无线方式进行数据传输的解决方案^[1]。然而, 由于移动 P2P 网络具有节点移动性强、无线通信环境不稳定以及节点能量有限等特点给移动计算环境中数据的组织与数据传输带来极大的挑战。因此, 如何提高用户数据访问效率、

降低数据访问代价是移动 P2P 网络研究的重要内容。

研究表明^[2-7]预载和缓存技术能够显著降低移动用户数据访问时延, 提高整个移动 P2P 系统工作效率。数据预载通过分析用户当前及历史访问记录, 预测未来可能会访问到的内容并将其提前缓存到本地, 当用户请求访问该数据时, 可以直接从本地缓存中读取该数据。缓存替换用于决定在缓存空间不足时如何替换缓存中的数据项。文献^[2]中针对移动多

媒体应用,提出了具有 QoS 服务质量保证的数据预载模型,采用马尔可夫链对未来可能访问的数据对象进行预测.文献[3]中综合考虑了移动设备内存容量、电池电量、以及数据传输成本等因素对预载的影响.文献[4]中提出了一种面向网络化移动应用的全局适应性数据预载机制,综合考虑了用户应用偏好和节点能量消耗等因素,并通过成本指标和利益指标来评价数据对象的预载成本和获益.以上这些预载策略都是通过对数据对象的访问频率来设计预载价值函数,缺乏对数据对象之间相关性的考虑.对于缓存管理,文献[5-6]中考虑各种影响缓存替换策略的因素,利用价值函数将缓存中价值最低的数据替换出去.文献[7]中则通过让节点度高的节点缓存非热点内容,让节点度低的节点缓存热点内容,从而达到网络负载平衡的目的,该方法仅从数据访问频度来考察资源的流行情况.已有的多数缓存替换算法通常仅考虑到缓存数据大小、使用频率等简单因素,忽略了数据之间的相关性,而用户访问的数据之间往往存在着一定的联系,即数据相关性,因此,重视数据之间的相关性对于预载与缓存替换具有重要意义.

本文针对移动 P2P 网络自身特点,充分考虑数据对象之间的相关性,将预载和缓存有机结合起来,提出移动 P2P 网络预载和缓存替换策略,达到提高缓存命中率、降低用户访问时延、提升网络服务质量的目的.

2 基于兴趣-相关度的 MP2P 预载模型

本文提出的预载策略主要针对半分布式移动 P2P 网络,超级节点负责域内节点和资源的管理以及转发普通节点资源查找请求等,普通节点功能比较单一,只负责自身动态数据的管理.整个网络组成以超级节点为中心的域.预载策略是基于兴趣-相关度的,其核心思想是首先超级节点根据域内普通节点的历史访问记录进行统计和分析,通过计算数据对象之间的兴趣-相关度,建立数据相关度矩阵,为后续计算提供数据支持;其次通过对同一节点在一段时间内连续发起的访问进行统计,构建访问跳转矩阵,然后利用该矩阵和数据挖掘技术计算数据对象的支持度和置信度;最后构造预载效用函数,选取价值最高的数据,将其提前缓存到该超级节点所在域组中几个能力较强的节点,域内其他用户下一次发起请求时,可以直接从本域内找到该数据副本,从而达到优化网络服务的目的.

用无向图 $G = (V, E)$ 来表示移动 P2P 网络模型,其中 $V = \{VS, VN\}$ 表示网络中所有节点集和,它包含两个子集 VS 和 VN ,其中 $VS = \{vs_1, vs_2, \dots, vs_l\}$ 表示网络中超级节点集合, $VN = \{n_1, n_2, \dots, n_h\}$ 表示网络中普通节点集合. E 表示网络中边的集合.假设网络中所有移动节点都具有唯一的标识符,以节点 s_i 为超级节点的域内移动节点集合表示为 $N = \{n_1, n_2, \dots, n_p\}$.网络中数据对象集合表示为 $D = \{d_1, d_2, \dots, d_q\}$.

2.1 兴趣-相关度定义

定义 1 兴趣度是指节点对资源感兴趣的程度,可以通过用户对数据对象的访问频率来衡量,即 $I_{n_i}(d_a) = \frac{\text{节点 } n_i \text{ 对数据对象 } d_a \text{ 的访问次数}}{\text{节点 } n_i \text{ 发起的总的访问次数}}$,其中, $n_i \in N, d_a \in D$. $I_{n_i}(d_j)$ 越大,则说明节点 n_i 对 d_a 越感兴趣.

定义 2 假设 $I_{s_j}(d_a) = \{I_{n_1}(d_a), I_{n_2}(d_a), \dots, I_{n_p}(d_a)\}$ 表示超级节点 s_j 对其域内普通节点 $\{n_1, n_2, \dots, n_p\}$ 对数据对象 d_a 的兴趣度集合.

定义 3 对于数据对象 d_a 与 d_b 的兴趣-相关度定义按式(1)进行,如下:

$$R(d_a, d_b) = \frac{\sum_{i=1}^p |(I_{n_i}(d_a) - I_{s_j}(d_a))(I_{n_i}(d_b) - I_{s_j}(d_b))|}{\sqrt{\sum_{i=1}^p ((I_{n_i}(d_a) - I_{s_j}(d_a))^2 + (I_{n_i}(d_b) - I_{s_j}(d_b))^2)}} \quad (1)$$

其中, $R(d_a, d_b) \in [0, 1], \forall a, b \in [1, q], s_j$ 为超级节点.式(1)值越大,表示数据对象 d_a 与 d_b 的兴趣-相关度越高,反之表示相关度越低.对于网络中全部数据对象 $D = \{d_1, d_2, \dots, d_q\}$,利用式(1)可以计算出任意两个数据对象之间的兴趣-相关度,从而建立其一个 $q \times q$ 的兴趣-相关度矩阵 $R_{q \times q}$,其中,行和列均为数据对象,矩阵元素 R_{ab} 为数据对象 d_a 和 d_b 的兴趣-相关度,即 $R(d_a, d_b)$.

2.2 建立访问跳转矩阵

由于本文提出的预载策略是基于历史访问记录的,现抽取超级节点 s_j 30 秒内的历史转发记录如表 1 所示,假设域内有 3 个普通节点 n_1, n_2, n_3 ,网络中数据对象有 6 个,分别是 $\{d_1, d_2, \dots, d_6\}$.

首先,超级节点 s_j 对其域内普通节点对网络中数据资源的历史访问记录进行统计,具体方法是,当一个节点一次访问结束后(此处采用基于空白点(gap-based)的方法^[8]来判定一次访问是否结束),对这次被访问的数据对象进行统计.以表 1 为例,空白点时间域为 5 秒,建立的节点历史访问记录如下:

表 1 超级节点 s_j 的历史转发记录

节点	数据对象	时间	节点	数据对象	时间	节点	数据对象	时间
n_1	d_2	0:01	n_3	d_4	0:07	n_2	d_4	0:24
n_2	d_4	0:01	n_1	d_1	0:15	n_1	d_3	0:27
n_3	d_1	0:02	n_1	d_3	0:17	n_1	d_4	0:28
n_1	d_4	0:03	n_1	d_4	0:18	n_1	d_6	0:29
n_1	d_6	0:06	n_2	d_1	0:20	n_2	d_5	0:29
n_2	d_5	0:06	n_2	d_3	0:21	n_2	d_6	0:30
n_2	d_6	0:07	n_1	d_6	0:21			

N_1 发起一次连续访问: $\{d_2, d_4, d_6\}$; N_2 发起一次连续访问: $\{d_4, d_5, d_6\}$; N_3 发起一次连续访问: $\{d_1, d_4\}$; N_1 发起一次连续访问: $\{d_1, d_3, d_4, d_6\}$; N_2 发起一次连续访问: $\{d_1, d_3, d_4, d_5, d_6\}$; N_1 发起一次连续访问: $\{d_3, d_4, d_6\}$. 为了统计域内节点对资源的兴趣度, 不考虑具体由哪个节点发起的访问, 只对访问的数据对象进行统计, 则对这 6 次访问记录统计结果如下: $C_1: \{0, 1, 0, 1, 0, 1\}$; $C_2: \{0, 0, 0, 1, 1, 1\}$; $C_3: \{1, 0, 0, 1, 0, 0\}$; $C_4: \{1, 0, 1, 1, 0, 1\}$; $C_5: \{1, 0, 1, 1, 1, 1\}$; $C_6: \{0, 0, 1, 1, 0, 1\}$, 其中, C_i 表示第 i 次访问, 每一列分别表示数据项 $\{d_1, \dots, d_6\}$, 若访问了 d_j 项数据, 则该项标记为 1, 否则标记为 0. 根据以上的 6 次访问记录, 分别统计 d_j 到 $d_k (j, k \in [1, 6], j \neq k)$ 的跳转情况, 形成跳转矩阵, 如图 1 所示. 例如, 上述 6 次访问中, 在一次访问中所访问的数据资源同时包含了 d_1, d_3 共有 2 次, 因此, 第 1 行第 3 列元素值是 2, 其他依次类推. 跳转矩阵主要为了后续计算支持度、置信度等提供数据支持.

	d_1	d_2	d_3	d_4	d_5	d_6
d_1	0	0	2	3	1	2
d_2		0	0	1	0	1
d_3			0	3	1	3
d_4				0	2	5
d_5					0	2
d_6						0

图 1 跳转矩阵

2.3 基于兴趣—相关度的 MP2P 预载模型

定义 4 支持度表示两个数据对象 d_a 和 d_b 同时出现的概率, 定义如下:

$$\text{Sup}(d_a, d_b) =$$

$$p(d_a, d_b) = \frac{\text{同时访问数据对象 } d_a \text{ 和 } d_b \text{ 的次数}}{\text{总的访问次数}}.$$

式中, $\forall d_a, d_b \in D$. 同时访问数据 d_a 和 d_b 的次数, 可通过跳转矩阵直接获得. $\text{Sup}(d_a, d_b)$ 值越大, 说明数据对象 d_a 和 d_b 同时出现的概率越高, 其相关性就越高.

定义 5 置信度表示数据对象 d_a 出现时, d_b 出现的概率有多大, 定义如下:

$$\text{Con}(d_a, d_b) =$$

$$p(d_a | d_b) = \frac{\text{同时访问数据对象 } d_a \text{ 和 } d_b \text{ 的次数}}{\text{访问包含资源 } d_a \text{ 的次数}}$$

式中, $\forall d_a, d_b \in D$. 同理同时访问数据 d_a, d_b 的次数, 也通过跳转矩阵获得. 若置信度为 100%, 则说明数据对象 d_a 和 d_b 可以捆绑缓存, 反之, 若置信度较低, 则说明这两个数据对象之间的关联性不大.

本文综合考虑了数据对象之间的兴趣—相关度、支持度、置信度以及数据对象的大小、路径等因素, 对于节点已经访问了数据对象 d_a , 建立了数据对象 d_b 的预载效用函数, 即式(2):

$$\text{Max}F(d_a, d_b) = \omega(R(d_a, d_b) \times \text{Sup}(d_a, d_b) \times \text{Con}(d_a, d_b)) \times \text{Path}\varphi_{d_b} / \text{Size}\phi_{d_b} \quad (2)$$

式中, $R(d_a, d_b)$ 为资源 d_a 和 d_b 的兴趣相关相度, $\text{Sup}(d_a, d_b)$ 、 $\text{Con}(d_a, d_b)$ 分别为资源 d_a 和 d_b 的支持度和置信度, Size 为当前考虑是否要预载的数据对象 d_b 的大小, Path 是源节点到数据对象 d_b 的路径距离, 可以用跳数来度量. 其中 ω, φ, ϕ 分别为全局影响因子, 取值范围均为 $[0, 1]$, 可以根据实际情况进行选取. 从预载资源效用函数 F 可以看出, 效用函数值与资源相关度、支持度、置信度和路径成正比, 与将要预载的数据对象大小成反比. 根据预载效用函数式(2) 选出价值最高的数据对象进行预载, 若缓存未满, 直接将该数据对象放入缓存中, 若缓存已满, 根据下文提出的缓存替换策略确定是否要缓存该数据对象.

3 基于兴趣—相关度的缓存替换策略

预载得到的数据对象必须存放到缓存中才能供移动用户所使用, 由于移动设备缓存容量有限, 若频繁进行预载—缓存操作可能会使许多经常访问的数据被替换出去, 从而导致缓存—预载抖动问题, 降低缓存命中率. 因此, 当缓存已满时, 对于预载的数据对象, 如何替换已有的缓存内容, 将是非常重要的.

传统的缓存替换算法通常利用对数据对象的访问频率或者最近的访问时间等做为替换标准,很少有关注意到数据之间的相关性等方面的信息.本文提出的缓存替换算法利用了蚁群算法中蚂蚁通过信息素留存寻找最优路径的机制,并综合考虑数据对象之间的相关性,数据块的大小,数据块的使用次数和频度等因素来设计缓存价值函数,当需要缓存替换时,利用缓存价值函数将综合价值最低的数据对象替换出缓存.该算法不仅能体现出数据对象对于本地和网络的重要性,也能很好地解决缓存-预载抖动问题.

定义 6 $\tau_{d_a}(t)$ 表示缓存中数据对象 d_a 在 t 时刻残留的信息素浓度.当数据对象刚进入缓存时,信息素浓度相同,当随着时间的推移,数据对象上的信息素会逐渐挥发,采用残留因子 $\rho(\rho \in (0,1))$ 进行描述, x 为距离上一次被访问的时间间隔,按式(3)

进行信息素更新:

$$\tau_{d_a}(t) = \begin{cases} \tau_{d_a}(t) \times \rho^x + 0.1 & \text{当缓存的数据对象 } d_a \text{ 被使用时} \\ \tau_{d_a}(t) \times \rho^x & \text{当缓存的数据对象 } d_a \text{ 未被使用时} \end{cases} \quad (3)$$

当缓存数据对象时,对新进入的数据对象 d_a 的信息素初始化为 $\tau_{d_a}(t) = 1$.当某个数据对象被使用时,则利用式(3)对其信息素浓度进行更新,则随着命中次数的增加,该数据对象的信息素浓度越浓.

定义 7 $\eta_{d_a}(t)$ 表示数据对象 d_a 的重要程度,

$\eta_{d_a}(t) = \frac{1}{\text{Size}(d_a)} \times N_{d_a}(t)$, 其中 $\text{Size}(d_a)$ 表示数据对象 d_a 的大小, $N_{d_a}(t)$ 表示数据对象 d_a 在 t 时刻被访问的总次数.

定义 8 $P_a(t)$ 表示 t 时刻缓存的数据对象 d_a 的缓存价值函数,按式(4)进行计算:

$$P_a(t) = \begin{cases} \frac{\sum_{d_b \in D} R(d_a, d_b) \alpha(t) \eta_{d_a}(t) \tau_{d_a}(t)}{\sum_{\substack{d_c \in D \\ d_c \neq d_d}} \sum_{d_d \in D} R(d_a, d_b) \alpha(t) \eta_{d_c}(t) \tau_{d_c}(t)} \\ 1 \end{cases} \quad (4)$$

当缓存已满时,首先按着式(3)更新全部数据的信息素浓度,然后按照式(4)对缓存中缓存的全部数据对象计算其价值,将价值最低的数据被替换出缓存.

4 仿真实验与结果分析

采用 JAVA 语言,编程实现了本文提出的策略,对其性能进行了分析.实验主要考查了预载命中率和缓存命中率两项指标.预载命中率是成功预测用户即将访问资源的概率;缓存命中率是节点自身缓存提供的数据服务与用户请求的服务的比值.

假设整个网络数据资源为 1 000 个,其访问概率服从 zipf 分布.在测试预载命中率时,起始访问记录数据设为 100 条,每访问完一次将该记录加入到历史记录中,最高访问记录为 1 000 条.实验中暂不考虑数据大小与路径的影响,将数据大小与路径均为设为 1.将本文提出的预载策略 IPS (Interest-relevancybased PrefetchStrategy) 与文献

缓存已满,需要执行替换

缓存未满,不需要执行替换

[9]中提出的 CMIP(Cache-Miss-Initiated Prefetch)预载策略进行性能比较,如图 2 所示,随着历史记录不断增大,两个策略均呈现出比较稳定的预载命中率,但是本文策略 IPS 的预载命中率均显著高于文献[9]中的预载命中率.当用户访问数据出现突变时的预载命中率如图 3 所示,其对 IPS 影响不大,而 CMIP 的预载命中率下降幅度较大,这是由于 IPS 是基于历史访问记录进行统计,当出现个别数据访问突变时,仍然能通过历史记录获得较稳定的预载命中率.可见 IPS 策略性能优于 CMIP 策略,尤其是在模拟现实中访问突变的情况下仍然能够获得较高的命中率.

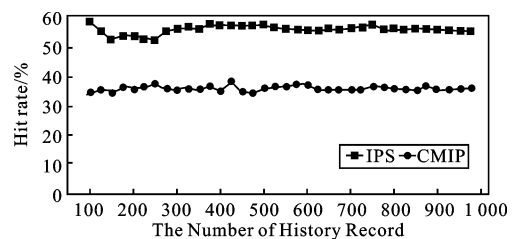


图 2 预载命中率对比

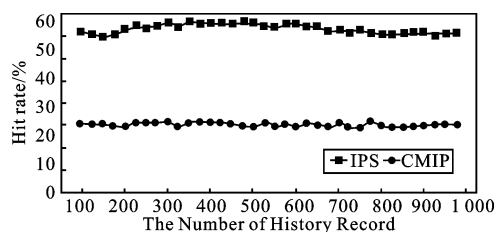


图3 加入数据突变后的预载命中率对比

下面进行缓存命中率的测试,设残留因子 $\rho = 0.9$,缓存大小为 10 MB,网络中数据资源个数为 1 000,且大小相等,均为 1 MB.缓存价值函数 $P_a(t)$ 中参数经过反复实验,选取 $\alpha = 0.9$, $\beta = 0.7$, $\gamma = 0.7$ 这组参数来进行后续实验.将本文提出的 IRP (Interest-Relevancy and Pheromone Cache Replacement Strategy) 策略与 FIFO、LRU 进行对比实验.如图 4 所示,IRP 缓存命中率最高,LRU 次之,FIFO 最低.当访问量在 200 左右,命中率进入平稳阶段.这表明访问概率服从 zipf 分布的环境下,IRP 算法更稳定,可掌握热点资源以及节点对资源的趋向性,算法性能优于另外两种算法.

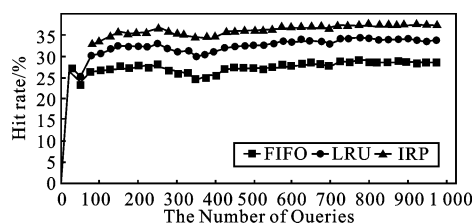


图4 访问量 缓存命中率对比图

5 结束语

本文提出了移动 P2P 网络预载缓存策略,通过对历史访问记录进行数据挖掘,选出预载价值最大的资源进行预载;同时为了避免预载抖动而导致降低缓存命中率,本文还提出了基于兴趣-相关度和信息素的缓存替换算法.本文通过掌握热点资源以及节点对数据资源兴趣的趋向性,有效提升了缓存的命中率,减少了用户的访问延迟,从而提高网络工作性能.在下一步工作中,将重点研究网络通信开销对系统性能的影响以及如何进一步降低通信开销等问题.

参考文献:

- [1] 欧中洪,宋美娜,战晓苏,等.移动对等网络关键技术[J].软件学报,2008,19(2):404-418.
- [2] Tuah N J, Kumar M, Venkatesh S. Resource-aware speculative prefetching in wireless networks[J]. Wireless Networks, 2003,9(1):61-72.
- [3] Vartiainen E, Roto V, Popescu A. Auto-update: a concept for automatic downloading of web content to a mobile device[C]// Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology. Singapore, ACM,2007.
- [4] 张晓薇,曹东刚,田刚,等.网络化移动应用的全局适应性数据预载机制[J].软件学报,2010,21(8):1783-1794.
- [5] 周欣欣,余镇危.基于流行度及最小访问代价的 MP2P 协同缓存优化策略[J].计算机工程与科学,2013,35(8):31-35.
- [6] Chen H, Xiao Y, Shen X. Update-based cache access and replacement in wireless data access[J]. Mobile Computing, IEEE Transactions on, 2006,5(12):1734-1748.
- [7] 高鹏飞.基于非结构化的 P2P 系统的协同缓存策略研究[D].吉林:东北大学,2010.
- [8] Saygin Y, Ulusoy O, Elmagarmid A K. Association rules for supporting hoarding in mobile computing environments[C]// Proceedings of 10th International Workshop on Research Issues in Data Engineering. San Diego, California, IEEE Computer Society,2000:71-78.
- [9] Song Hui, Cao Guohong. Cache-miss-initiated prefetch in mobile environments [J]. Computer Communications, 2005,28(7):370-381.

作者简介:

周欣欣 女,(1977-),博士研究生,讲师.研究方向为计算机网络体系结构、移动 P2P 网络.

卢喆(通讯作者) 男,(1988-),硕士.研究方向为移动 P2P 网络. E-mail:582599827@qq.com.