

## 基于副本复制和 Bloom Filter 的 P2P 概率路由算法\*

朱桂明<sup>1+</sup>, 郭得科<sup>2</sup>, 金士尧<sup>1</sup>

<sup>1</sup>(国防科学技术大学 计算机学院 并行与分布处理国家重点实验室, 湖南 长沙 410073)

<sup>2</sup>(国防科学技术大学 信息与管理学院 信息系统工程国家重点实验室, 湖南 长沙 410073)

### P2P Probabilistic Routing Algorithm Based on Data Copying and Bloom Filter

ZHU Gui-Ming<sup>1+</sup>, GUO De-Ke<sup>2</sup>, JIN Shi-Yao<sup>1</sup>

<sup>1</sup>(National Laboratory for Parallel and Distributed Processing, School of Computer, National University of Defense Technology, Changsha 410073, China)

<sup>2</sup>(National Laboratory for Information System Engineering, School of Information System and Management, National University of Defense Technology, Changsha 410073, China)

+ Corresponding author: E-mail: guiming.zhu@yahoo.com.cn

**Zhu GM, Guo DK, Jin SY. P2P probabilistic routing algorithm based on data copying and Bloom Filter. Journal of Software, 2011, 22(4): 773-781. <http://www.jos.org.cn/1000-9825/3757.htm>**

**Abstract:** It is hard to optimize query latency, query hit, and query cost at the same time for the resource location of unstructured peer-to-peer network. For this problem, this paper presents a probabilistic routing algorithm called DCBF (data copying and Bloom Filter), which is based on data copying and a Bloom Filter technique. DCBF makes a few copies of each shared resource and places each copy on a random selected node, based on a directed random network. Each node forwards membership information to neighboring nodes with distributed declining Bloom Filters. Analysis and experimental results show that DCBF can make the most of the nodes, use the membership information of resource objects by making only a few copies, and forward membership information with distributed declining Bloom Filter to achieve high query hits with low cost and low latency.

**Key words:** peer-to-peer computing; directed random network; data copying; decaying Bloom Filter; probabilistic routing

**摘要:** 非结构化 P2P 网络资源定位过程中的查询延迟、查准率和查询成本难以同时被优化,为此,提出一种基于副本复制和 Bloom Filter 技术的 P2P 概率路由算法 DCBF(data copying and Bloom Filter).DCBF 基于有向随机网络,对资源对象进行少量的复制,并将各个副本随机路由给网络中的节点;接收副本的节点,以分布式衰减 Bloom Filter 向邻近节点传递副本的成员资格信息.理论分析和实验结果均表明,DCBF 仅需复制少量的副本,通过以分布式衰减 Bloom Filter 传递副本的成员资格信息,使得网络中的绝大多数节点能够感知到副本的成员资格信息,从而使得各个节点能够以极低的查询代价,在较低的路由延迟范围内,高概率地将查询路由到目标节点.

**关键词:** 对等计算;有向随机网络;副本复制;衰减 Bloom Filter;概率路由

\* 基金项目: 国家自然科学基金(60903206, 61070216); 国家重点基础研究发展计划(973)(2007CB310900); 国家高技术研究发展计划(863)(2011AA0123824001); 国防科学技术大学预研基金

收稿时间: 2009-04-22; 修改时间: 2009-07-06; 定稿时间: 2009-10-22

中图法分类号: TP393

文献标识码: A

对等计算(peer-to-peer,简称 P2P)是近年来兴起的一种重要的网络计算技术,它使得处于网络边缘的用户更易共享资源.如何在节点之间高效、快速地定位资源,一直是 P2P 研究的热点与核心问题.基于分布式哈希技术(DHT)的结构化对等网络虽然能够在固定延迟内将查询路由到目标节点,但是哈希方法掩盖了资源和节点的语义信息,使得查询仅能按照精确匹配的方式进行;结构化对等网络存在着由于“热点”访问而导致的负载不均衡问题<sup>[1]</sup>;节点的频繁加入与退出,使得结构化对等网络的维护开销比较大.通常情况下,非结构化对等网络中节点之间的关系比较随意,动态维护开销较小;对节点、资源和查询语义的表示,能够计算出查询与节点及资源的相关性大小,从而使得查询能够以模糊匹配的方式进行.但是,非结构化对等网络存在着资源定位效率低、准确性难以保证的问题<sup>[2]</sup>.如何对非结构化对等网络中的节点进行有效组织,对节点和资源的语义有效传递,使得网络中的任一节点均可以以极低的查询代价、在确定的延迟范围内、高概率地将查询路由到目标节点,是一个颇具挑战性的问题.

针对这一问题,本文提出了一种基于副本复制和 Bloom Filter 的 P2P 概率路由算法 DCBF(data copying and Bloom Filter).DCBF 提出了基于有向随机网络,使用分布式衰减 Bloom Filter 表达、传递节点和资源的语义信息.与双向随机网络相比,这种方法有效抑制了信息回流和多径叠加问题.DCBF 提出了副本复制和分布式衰减 Bloom Filter 相结合的技术表达和传递资源语义,使得网络中绝大多数节点均可感知到资源对象的成员资格信息;由于采用了分布式衰减 Bloom Filter,对任何一个资源而言,DCBF 仅需复制少量的副本,即可使得资源的成员资格信息覆盖网络中绝大多数节点,从而使得网络中任何一个节点发起针对任何一个资源的查询,均可以以极低的查询代价、在确定的路由延迟范围内、高概率地路由到目标节点;并且,由于采用了副本复制技术,能够在一定程度上缓解“热点”资源的访问问题.

## 1 相关工作

Bloom Filter 是一种结构简单而空间复杂度很小的随机数据结构<sup>[3]</sup>,它主要用来对集合进行精确表示,并代替被表示的集合回答类似“元素  $x$  是否在集合  $A$  中”的判定.Bloom Filter<sup>[3,4]</sup>被广泛用于非结构化对等网络<sup>[5-8]</sup>、无线传感器网络<sup>[9]</sup>、Ad Hoc 网络<sup>[10-13]</sup>以及 Mesh 网络<sup>[14]</sup>中实现概率路由.其共同的思想<sup>[15]</sup>是,每个节点将其所共享的全部资源对象用一个 Bloom Filter 表示,并将其传播到给定逻辑跳步内(例如  $h$  跳)的节点;节点把从相同邻居链路传递进来的 Bloom Filter 集合作为该邻居链路的路由条目.如果一个节点需要将查询消息路由到距离其  $h$  跳步内的潜在目的节点,则通过检查每个邻居链路的路由条目就能挑选出正确的路由方向,并转发给相应的邻居节点.这些方法的缺点是,每个节点保存其  $h$  跳步内所有节点的 Bloom Filter 需要消耗大量的存储空间,逐一扫描每个 Bloom Filter 会导致路由决策的延迟比较大.Kumar 等人首次提出了依指数衰减的 Bloom Filter 表达、传递节点和资源的成员资格元信息,并在此基础上提出了非结构化对等网络的弱状态路由机制<sup>[16]</sup>.每个节点产生的本地 Bloom Filter 以传播距离为指数进行衰减传递,每个节点将从相同邻居传递进来的 Bloom Filter 进行“并”运算,并将产生的联合 Bloom Filter 作为该邻居的路由条目.查询路由时,将查询消息沿着拥有目标对象的成员资格信息量最大的方向传递,而且在查询消息的整个传递路径上,关于其目标对象的成员资格信息越来越多<sup>[16]</sup>.文献[15]指出,由于噪声的原因,依据文献[16]的弱状态路由机制进行路由,查询消息往往以一定的概率沿着错误方向传递.为此,文献[15]假定任意节点  $h$  范围内的节点的网络拓扑结构为度为  $C$  的树型网络,且在不考虑信息回流和多径叠加问题的基础上,对于线性衰落和指数衰落模型,根据 Bloom Filter 的比特数个数  $m$ 、哈希函数个数  $k$ 、衰落参数  $d$ 、最大衰落传输距离  $h$  以及各节点共享资源的个数  $n$ ,经细致建模,求得单播弱状态路由和有效弱状态路由的成功概率的下界;反之,根据给定查询成功率的下界  $\sigma$ ,可以约束参数  $m, k, d, h$  和  $n$  的取值.BubbleStorm<sup>[17]</sup>提出了一种副本复制和 Bubblecast 技术,使得在副本和查询消息数均为  $O(\sqrt{N})$  量级的情况下,能够以非常高的概率命中目标对象.其中,  $N$  为网络规模.

相对于 BubbleStorm,DCBF 能够在更大程度上降低副本和消息的数量,且仍然能够在确定延迟范围内,以高

概率将查询消息路由到目标节点.

## 2 DCBF 算法

DCBF 算法的基本思想是,对于任何一个共享资源对象,复制一定数量的副本,将每个副本随机路由到网络中的某个节点放置,并采用依指数衰减的 Bloom Filter 来传递副本对象的成员资格信息,每个节点将从相同邻居节点传递进来的 Bloom Filter 的并运算结果作为该邻居节点的路由条目;查询路由时,将查询消息沿着目标对象成员资格信息量增大的方向进行.DCBF 试图通过这种方法使得任何资源对象的成员资格信息均可覆盖网络中的绝大多数节点,从而使任何节点针对任何资源对象的查询均可在确定延迟范围内高概率地完成.

由于文献[15]没有给出解决信息回流和多径叠加问题的方法,DCBF 明确提出了基于有向随机网络模型传递表达节点语义的衰落 Bloom Filter,以有效抑制信息回流和多径叠加问题.

### 2.1 基于有向随机网络的路由条目更新

弱状态路由机制能否将查询消息沿着正确的方向传递,取决于节点路由条目中关于目标对象成员资格信息量的增大趋势是否如实反映了各节点与目标对象所在节点的距离远近关系问题.双向随机网络中,节点路由条目的更新存在着信息回流、多径叠加和噪声问题,往往使得查询消息沿着错误的方向传递.对于有向随机网络而言,节点路由条目的更新沿着单个方向衰减进行,资源对象成员资格信息量在最短路径上以最小的衰减到达各节点.通过设定资源对象成员资格信息的最大传递距离小于随机网络任意两个节点之间距离的期望值,节点从非最短路径上收到关于目标对象的成员资格信息量的期望值接近为 0,从而能够有效抑制信息回流和多径叠加的问题,使得各节点能够沿着最短路径将查询消息路由到目标节点.

**定理 1.** 假设有向随机网络  $G$ ,其网络规模为  $N$ ,节点出度为  $C$ ,令  $G$  中任意两个节点之间距离的期望值为  $AVG$ ,则当资源对象成员资格信息传递的最大距离小于  $AVG$  时,节点从非最短路径上接收到的资源对象的成员资格信息量为 0.

证明:由于  $G$  中任意两个节点之间距离的期望值为  $AVG$ ,因此,任意节点  $u$  从非最短路径上接收到节点  $v$  共享资源对象成员资格信息量,经距离为  $AVG$  的衰减传递;由于资源对象成员资格信息传递的最大距离小于  $AVG$ ,从而有节点  $u$  从非最短路径上接收到节点  $v$  共享资源对象成员资格信息量为 0.  $\square$

基于有向随机网络的节点接收和转发 Bloom Filter 的算法如算法 1 所示.节点将从上行(UP)邻居节点  $src$  接收到的  $bloomfilter$  与上行邻居节点表中  $src$  原路由条目的  $srcBloomFilter$  进行“并”运算,并将运算结果作为  $src$  节点新的路由条目.将衰减后的  $bloomfilter$  传递给所有下行(DOWN)邻居节点,由其作进一步处理.

**算法 1.**  $rcv\_snd(Node\ src, BloomFilter\ bloomfilter)$ .

1. Get  $src$  node Bloom Filter  $srcBloomFilter$  from UP neighbors list;
2.  $srcBloomFilter \leftarrow srcBloomFilter \cup bloomfilter$ ; //与  $src$  节点的原路由条目并运算
3.  $put(src, srcBloomFilter)$ ; //更新  $src$  节点的路由条目
4.  $bloomfilter \leftarrow bloomfilter/d$ ; //衰减  $bloomfilter$
5. for all neighbor  $i$  from DOWN neighbors list //更新所有下行邻居节点的路由条目
6.  $rcv\_snd(i, bloomfilter)$ ;

需要说明的是,节点的下行(DOWN)邻居是从网络中随机选择的.节点无法选择上行(UP)邻居节点,节点的上行邻居是由邻居节点向该节点(作为其下行邻居)发送路由条目更新时所获取的.

### 2.2 基于 Bloom Filter 的副本复制

为获取较高的查准率, BubbleStorm 的副本和查询消息数均取为  $O(\sqrt{N})$  量级.在保证查询准确率的前提下, DCBF 试图进一步降低副本和查询消息的数量,使得任何节点发起的针对任何对象的查询均可以高概率地在确定的延迟内完成. DCBF 将任意共享资源对象复制一定数量的副本,将各个副本随机放置在网络中的节点.即,首先在网络中随机行走  $3(1+\log N)$  步长<sup>[17]</sup>以随机选择网络中的节点,到达某节点后,由该节点出发以随机路由的方

式查找尚未被该资源对象副本覆盖的节点  $u$  放置副本;之后,节点  $u$  产生该副本的 Bloom Filter,并将其依指数衰减的方式向其他节点传递;在传递 Bloom Filter 的过程中,若接收节点是已覆盖节点,则由该节点以 Bloom Filter 不衰减和消息生命期不递减的方式向该节点所有尚未覆盖的邻居节点传递.节点接收和转发带有副本元信息的信息算法如算法 2 所示,依指数衰减传递副本的 Bloom Filter 的算法如算法 3 所示.

**算法 2.**  $resRcvSnd(Node\ node, Message\ message)$ .

1. if (message is alive||node is covered by Bloom Filter of the copy
2.     corresponding to the copy meta-information taken by the message)
3.     decrease message time to live;
4.     select random neighbor  $v$  from DOWN neighbor list;
5.      $resRcvSnd(v, message)$ ;
6. else
7.     keep the copy corresponding to the copy meta-information taken by the message;
8.     for each neighbor  $v$  from DOWN neighbor list
9.          $issueResBF(v, itemBloomFilter)$ ;     //issue the Bloom Filter of the copy

**算法 3.**  $issueResBF(Node\ node, BloomFilter\ bf)$ .

1. if (node covered by this bf)
2.     for (each node  $v$  not covered by this bf from DOWN neighbor list)
3.          $issueResBF(v, bf)$ ;
4. else
5.     update incoming node's Bloom Filter with  $bf$ ;
6.     decrease the life time of  $bf$ ;
7.      $bf \leftarrow bf/d$ ;
8.     for (each node  $v$  from DOWN neighbor list)
9.          $issueResBF(v, bf)$ ;

### 3 DCBF 算法分析

对于网络规模为  $N$ 、节点出度为  $C$  的有向随机网络,给定副本的最大传递距离为  $H$ ,可以通过下述定理求得资源对象成员资格信息覆盖的节点个数与副本个数之间的关系.

**定理 2.** 给定随机有向图  $G=(V, E)$ ,  $V$  为节点集合,  $E$  为边集合,  $|V|=N$ . 假设任意节点  $u, v \in V$ , 存在从  $u$  指向  $v$  的有向边的概率为  $p$ , 令  $V_i^u$  是与节点  $u$  的距离小于或等于  $i$  的节点集, 定义  $|V_0^u|=0$ , 则有  $|V_i^u|=N(1-e^{-p|V_{i-1}^u|})$ . 其中,  $i \geq 2$  且  $|V_1^u|=N(1-e^{-p})$ .

证明:由于对于任意节点  $u, v \in V$ , 存在从  $u$  指向  $v$  的有向边的概率为  $p$ , 则任意节点出度的期望值为  $Np$ . 由定义,  $|V_0^u|=0$ , 由于节点  $u$  出度的期望值为  $Np$ , 则  $V_1^u$  相当于从  $V$  中独立进行  $Np$  次的节点选择, 于是有

$$|V_1^u|=N \left( 1 - \left( 1 - \frac{1}{N} \right)^{Np} \right) \approx N \left( 1 - e^{-\frac{Np}{N}} \right) = N(1 - e^{-p}).$$

而  $V_2^u$  相当于从  $V$  中独立进行  $|V_1^u|Np$  次的节点选择, 则  $|V_2^u|=N \left( 1 - e^{-\frac{|V_1^u|Np}{N}} \right) = N(1 - e^{-p|V_1^u|})$ . 依此类推,

$|V_i^u|=N(1 - e^{-p|V_{i-1}^u|})$ ,  $i \geq 2$ . 定理 2 得证. □

**定理 3.** 给定随机有向图  $G=(V, E)$  及其子图  $G'=(V', E')$ , 其中,  $V' \subseteq V, E' \subseteq E$ .  $G$  中任意节点  $u, v \in V$ , 存在从  $u$  指向  $v$  的有向边的概率为  $p$ . 定义  $V'$  的边界点集合  $Border(V')=\{v' | v' \in V', \langle v', v \rangle \in E - E'\}$ . 若将从  $v \in V - V'$  到  $v' \in Border(V')$

有向边 $(v,v')$ 理解为从 $v$ 到 $\{u \in V-V' | (v,u) \in E-E'\}$ 集中所有点的有向边集,则有向图 $G-G'$ 中任意节点 $u',v' \in V-V'$ ,存在从 $u'$ 指向 $v'$ 的有向边的概率 $p'$ 为 $p'=p+|Border(V')|(p^2-p^3)$ .

证明:将从 $v \in V-V'$ 到 $v' \in Border(V')$ 的有向边 $(v,v')$ 理解为从 $v$ 到 $\{u \in V-V' | (v,u) \in E-E'\}$ 集中所有点的有向边集,则对于 $G-G'$ 新增加边的期望值为 $((|V|-|V'|) \times |Border(V')| \times p) \times ((|V|-|V'|) \times p) - (|V|-|V'|)^2 \times p \times |Border(V')| \times p \times p$ .由边的总数相等得到, $(|V|-|V'|)^2 p' = (|V|-|V'|)^2 p + (|V|-|V'|)^2 |Border(V')| p^2 - (|V|-|V'|)^2 |Border(V')| p^3$ ,从而有 $p'=p+|Border(V')|(p^2-p^3)$ .定理3得证.  $\square$

**定理4.** 给定随机有向图 $G=(V,E)$ , $G$ 中任意节点 $u,v \in V$ ,存在从 $u$ 指向 $v$ 的有向边的概率为 $p,|V|=N$ .按照如下副本分发过程计算副本覆盖的节点数:

**Step 1.** 任选一个节点 $u$ ,向所有与该节点距离小于或等于 $H$ 的节点传递副本,若节点在先前的迭代过程中被副本覆盖,则该节点不计入副本的传递距离,且由该节点将副本传递给其所有尚未被副本覆盖的邻居节点;

**Step 2.** 重复 Step 1,直到分发完所有副本.

令 $S_i$ 表示发布 $i$ 个副本覆盖节点的个数, $p_i$ 表示按照上述过程发布 $i$ 个副本后尚未被副本覆盖的任意节点 $u,v$ 概念上存在从 $u$ 指向 $v$ 的有向边的概率, $\theta(n,p,h)$ 表示有向随机网络节点规模为 $n$ 、从任一节点 $u$ 存在指向 $v$ 的有向边的概率为 $p$ 、与任意节点距离小于或等于 $h$ 的节点个数,则有 $S_i=S_{i-1}+\theta(N-S_{i-1},p_i,H),p_i=p+|Border(S_{i-1})|(p^2-p^3)$ .其中, $p_0=p,S_0=0$ .

证明:根据 $p_i,S_i$ 和 $\theta(n,p,h)$ 表达的含义,显然有 $p_0=p,S_0=0,S_i=S_{i-1}+\theta(N-S_{i-1},p_{i-1},H)$ .根据定理3,则有 $p_i=p+|Border(S_{i-1})|(p^2-p^3)$ .定理4得证.  $\square$

对于节点数为 $N$ 、节点出度为 $C$ 的有向随机网络,给定副本的最大传递距离为 $H$ ,可以通过定理4求得发布任意个数副本所覆盖的节点数及其对应的覆盖率;给定覆盖率,亦可通过定理4,采取迭代试探的方式计算出所需发布副本的个数.

**定理5.** 假设网络规模为 $N$ ,随机选择 $r$ 个节点放置副本,则进行 $g$ 次独立查询命中目标的概率为 $p=1-e^{-rxg/N}$ .

证明:由于网络规模为 $N$ ,有 $r$ 个节点放置副本,则一次查询命中目标的概率为 $r/N$ ,则进行 $g$ 次独立查询没有一次命中目标的概率为 $(1-r/N)^g \approx e^{-rxg/N}$ ,从而进行 $g$ 次独立查询命中目标的概率为 $p=1-e^{-rxg/N}$ .定理5得证.  $\square$

由定理5,在给定副本覆盖的情况下,以及已知查询消息的最大路由跳数,可以计算出DCBF的查准率.由于DCBF仅需少量的副本,通过分布式衰减 Bloom Filter 即可覆盖网络中绝大多数节点.因此,仅需少数的查询消息即可获得较高的命中率.

对于任意资源对象,假设DCBF复制的副本数为 $a$ ,在每个副本放置之前,带有副本元信息的信息在网络中随机行走的距离约为 $3(1+\log N)$ ,从而为放置副本而产生的消息数约为 $3a(1+\log N)$ .通常情况下,针对资源对象的查询是经常性的行为,DCBF算法为单个资源对象的副本放置而产生的约为 $3a(1+\log N)$ 个消息代价,使得每次查询仅需少量的消息个数即可获得较高的查询命中率,具有较高的效率.

## 4 模拟实验

我们在 Peersim<sup>[18]</sup>上实现了DCBF算法.实验主要考察基于有向随机网络的信息传递抑制信息回流和多径叠加问题的有效性、覆盖率与副本个数之间的关系、查准率与副本个数和路由跳步数之间的关系.衡量的指标有无噪率(noise-free rate)、最优路由率(shortest routing rate)、覆盖率(covered rate)和查准率(query hit rate).其中,无噪率定义为仅从最短路径上接收资源对象成员资格信息的节点个数占所有接收到资源对象成员资格信息节点个数的比例,覆盖率定义为被资源对象成员资格信息覆盖的节点个数占节点总数的比例,最优路由率定义为能以最短路径路由到目标节点的节点个数占目标资源对象成员资格信息覆盖节点个数的比例,查准率为命中目标节点的查询次数占查询总数的比例.以下各组实验仅发布1个资源对象及该资源对象的副本,即实验不考虑由于多个资源对象成员资格信息量在节点路由条目 Bloom Filter 中的叠加而引起的噪声问题;查询路由时,将查询消息沿着目标对象成员资格信息量增大的方向转发.

针对基于有向随机网络的信息传递对抑制信息回流和多径叠加问题的有效性进行了一组实验(I),衡量指

标有无噪率、最优路由率和覆盖率.图 1 和图 2 显示了不同网络规模、节点出度和传递距离下的覆盖率、无噪率和最优路由率.其中:

图 1 给出了网络规模  $N$ 、节点出度  $C$  分别为  $(N=10000, C=4)$ ,  $(N=10000, C=5)$ ,  $(N=10000, C=6)$ ,  $(N=2000, C=5)$  和  $(N=5000, C=5)$ , 发布 1 个副本时覆盖率、无噪率和最优路由率与传递距离  $H$  之间的关系.图 1 表明:不同网络规模、不同节点出度的有向随机网络,其覆盖率、无噪率和最优路由率关于传递距离均有类似的趋势;当传递距离  $H$  为  $\lfloor \text{AVG} \rfloor$  时,具有较大的覆盖率、较高的无噪率和较优的最优路由率;当传递距离为  $\lfloor \text{AVG} \rfloor - 1$  时,覆盖率在 0.1 附近,无噪率在 0.9 附近,最优路由率则接近 1.

图 2 给出了网络规模  $N$  为 10 000、节点出度  $C$  取 3~9,网络规模  $C$  取 1 000~10 000 以及节点出度为 5,传递距离分别为  $\lfloor \text{AVG} \rfloor$ ,  $\lfloor \text{AVG} \rfloor - 1$  和  $\lfloor \text{AVG} \rfloor - 2$  时的无噪率.如图 2 所示,当传递距离为  $\lfloor \text{AVG} \rfloor$  时,无噪率在 0.6 附近;当传递距离为  $\lfloor \text{AVG} \rfloor - 1$  时,无噪率在 0.9 附近;而当传递距离取  $\lfloor \text{AVG} \rfloor - 2$  时,无噪率接近 1.因此,限制传递距离小于  $\lfloor \text{AVG} \rfloor$ ,即限制最大传递距离在有向随机网络中任意两个节点之间距离的期望值以内,基于有向随机网络的信息传递能够有效抑制信息回流和多径叠加的问题.

针对 DCBF 算法覆盖率与副本个数之间的关系,进行了一组实验(II),实验结果如图 3 所示.图 3 显示了网络规模  $N$  为 10 000、节点出度  $C$  为 4 和 5 时,覆盖率与副本个数之间关系的实验和理论结果.实验结果表明,DCBF 算法的资源对象的覆盖率与副本个数之间关系的实验结果与理论值基本吻合,且当传递距离取  $\lfloor \text{AVG} \rfloor - 1$  时,发布少量的副本即可覆盖绝大多数节点.

针对查准率与副本个数和路由跳步数之间的关系,进行了一组实验(III).实验中,每个节点发起一次针对目标对象的查询.图 4 显示了网络规模  $N$  为 10 000、节点出度  $C$  为 5、传递距离  $H$  为 4 和 5 时,不同副本个数对应的查准率与最大路由跳数之间的关系.实验结果表明,以  $\lfloor \text{AVG} \rfloor$ ,  $\lfloor \text{AVG} \rfloor - 1$  传递副本的成员资格信息量,仅需发布少量副本即可以较小的查询延迟获得较高的查准率.

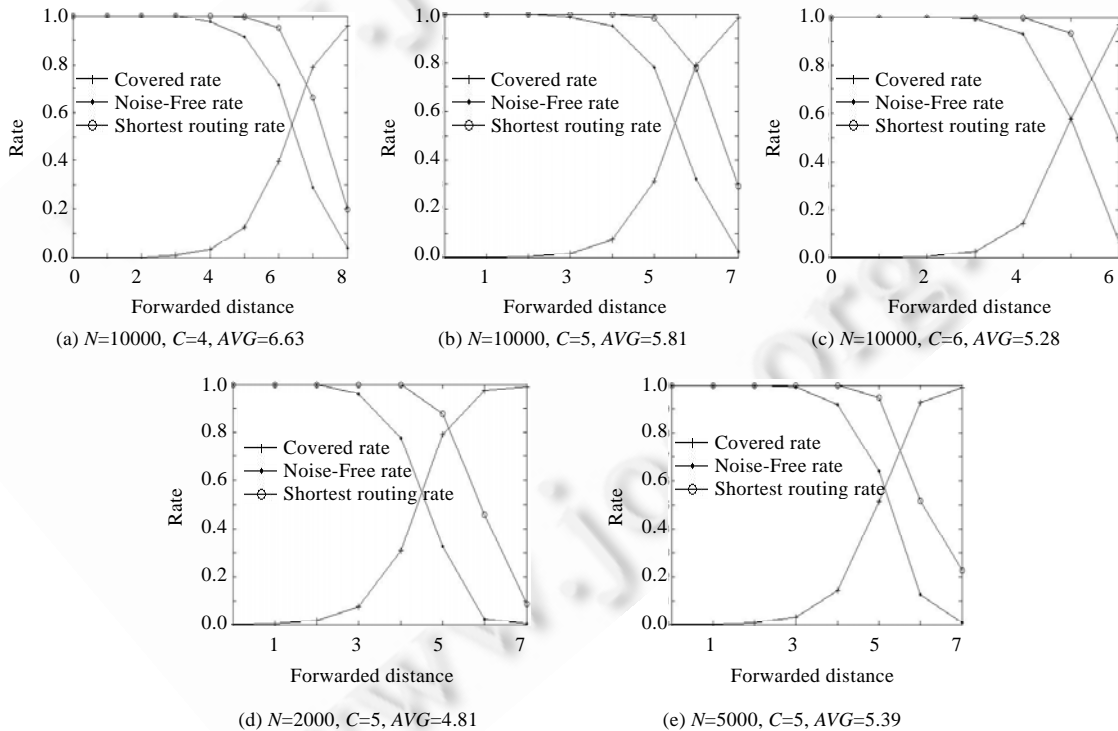


Fig.1 Covered rate, noise-free rate and shortest routing rate are functions of forwarded distance under different network size and out-degree

图 1 不同网络规模和节点出度下的覆盖率、无噪率和最优路由率与传递距离的关系

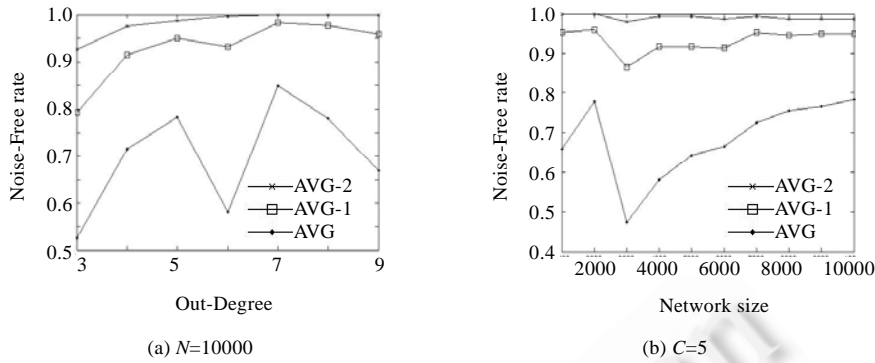


Fig.2 Noise-Free rate is a function of network size and out-degree when forwarded distance is  $\lfloor AVG \rfloor$ ,  $\lfloor AVG \rfloor - 1$  and  $\lfloor AVG \rfloor - 2$

图 2 传递距离为  $\lfloor AVG \rfloor$ ,  $\lfloor AVG \rfloor - 1$  和  $\lfloor AVG \rfloor - 2$  时的无噪率与网络规模和节点出度的关系

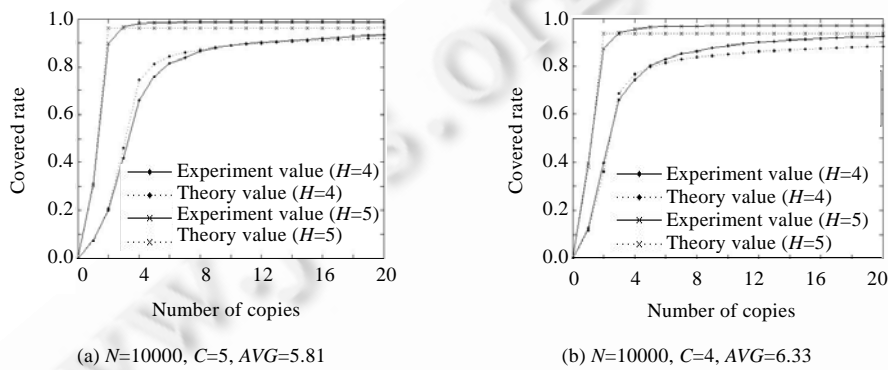


Fig.3 Covered rate is a function of number of copies under different network size, out-degree and forwarded distance

图 3 不同网络规模、节点出度和传递距离下的覆盖率与副本个数之间的关系

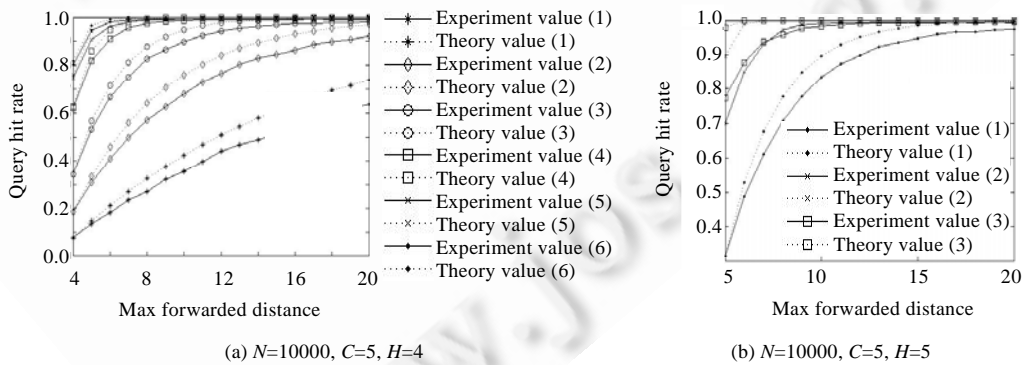


Fig.4 Query hit rate is a function of max forwarded distance of query message under different network size, out-degree, forwarded distance and number of copies

图 4 不同网络规模、节点出度、传递距离和副本个数情况下的查准率与消息最大路由距离之间的关系

针对 DCBF 和 BubbleStorm 两种算法的查准率与副本个数和消息个数(或最大路由跳步数)之间关系的比较,进行了一组实验(IV).图 5 给出了网络规模  $N$  为 10 000、节点出度  $C$  为 5、传递距离  $H$  为 4 时,BubbleStorm

算法的副本个数分别为 1,4,50,100 和 200 时的查准率与消息个数之间的理论结果关系以及 DCBF 算法副本个数分别为 1 和 4 时的查准率与消息个数之间的实验结果关系.从图 5 可以看出,在副本个数、查询消息个数相同且数量较小时,DCBF 算法的查准率远远高于 BubbleStorm 算法;而对于相同的查准率,BubbleStorm 算法的副本个数和查询消息的个数均以较大的幅度大于 DCBF 算法的对应值.相对于 BubbleStorm 算法,为了获取相同的查准率,DCBF 算法能够在更大程度上降低副本和消息的数量.

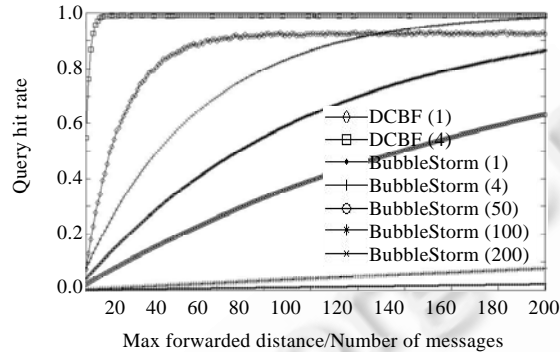


Fig.5 Comparison between DCBF and BubbleStorm

图 5 DCBF 与 BubbleStorm 结果对比

## 5 结 论

针对非结构化 P2P 网络资源定位过程中的查询延迟、查准率和查询成本难以同时被优化的问题,本文提出了一种基于副本复制和 Bloom Filter 的 P2P 概率路由算法 DCBF.算法 DCBF 创新性地基于有向随机网络传递资源对象的分布式衰减 Bloom Filter,并设定最大传递距离略小于任意两节点之间距离的期望值,有效抑制了信息传递过程中的多径、回流和叠加问题;通过使用分布式衰减 Bloom Filter 向邻居节点传递资源对象的成员资格信息,极大地减少了为覆盖相同节点个数而需要的副本数量;通过设置合理的最大传递距离(通常取值略小于任意两节点之间距离的期望值),与 BubbleStorm 算法相比,仅需发布少量的副本个数即可覆盖网络中绝大多数节点.实验和理论均表明,DCBF 算法能够使得任意节点以较小的查询延迟和较低的查询代价获得较高的查询命中率.

## References:

- [1] Pong YJ, Fang CT, Yang LW. Keyword search in DHT-based peer-to-peer networks. In: Arora A, ed. Proc. of the 25th IEEE Int'l Conf. on Distributed Computing Systems. Washington: IEEE Computer Society, 2005. 339–348. [doi: 10.1109/ICDCS.2005.44]
- [2] Li DS. Research on peer-to-peer resource location in large-scale distributed systems [Ph.D. Thesis]. Changsha: National University of Defense Technology, 2005 (in Chinese with English abstract).
- [3] Bloom BH. Space/Time tradeoffs in hash coding with allowable errors. Communications of the ACM, 1970,13(7):422–426. [doi: 10.1145/362686.362692]
- [4] Hua N, Zhao HQ, Lin B, Xu J. Rank-Indexed hashing: A compact construction of bloom filters and variants. In: Almeroth K, Ramakrishnan KK, eds. Proc. of the 16th IEEE Int'l Conf. on Network Protocols. Washington: IEEE Computer Society, 2008. 73–82. [doi: 10.1109/ICNP.2008.4697026]
- [5] Hodes TD, Czerwinski SE, Zahorjan BY, Joseph AD, Katz RH. An architecture for secure wide-area service discovery. Wireless Networks, 2002,8(2/3):213–230. [doi: 10.1023/A:1013772027164]
- [6] Reynolds P, Vahdat A. Efficient peer-to-peer keyword searching. In: Schmidt D, Endler M, eds. Proc. of the ACM/IFID/USENIX 2003 Int'l Conf. on Middleware. New York: Springer-Verlag, 2003. 21–40. [doi: 10.1007/3-540-44892-6\_2]



- [7] Rhea SC, Kubiatowicz J. Probabilistic location and routing. In: Lee D, Orda A, eds. Proc. of the IEEE INFOCOM 2002. Washington: IEEE Computer Society, 2002. 1248–1257. [doi: 10.1109/INFOCOM.2002.1019375]
- [8] Bauer D, Hurley P, Pletka R, Waldvogel M. Bringing efficient advanced queries to distributed hash tables. In: Jha S, Hassanein H, eds. Proc. of the 29th Annual IEEE Int'l Conf. on Local Computer Networks. Washington: IEEE Computer Society, 2004. 6–14. [doi: 10.1109/LCN.2004.32]
- [9] Hebden P, Pearce AR. Data-Centric routing using bloom filters in wireless sensor networks. In: Alex H, Begg R, eds. Proc. of the 4th Int'l Conf. on Intelligent Sensing and Information Processing. Washington: IEEE Computer Society, 2006. 72–77. [doi: 10.1109/ICISIP.2006.4286065]
- [10] Hsiao PH. Geographical region summary service for geographical routing. In: Corson MS, Das SR, eds. Proc. of the 2nd ACM Int'l Symp. on Mobile ad hoc Networking & Computing. New York: ACM, 2001. 263–266. [doi: 10.1145/501449.501455]
- [11] Yuen WH, Schulzrinne H. Improving search efficient using bloom filters in partially connected ad hoc networks: A node-centric analysis. Computer Communications, 2007,30(16):3000–3011. [doi: 10.1016/j.comcom.2007.05.055]
- [12] Acer UG, Kalyanaraman S, Abouzeid AA. Weak state routing for large scale dynamic networks. In: Hou J, Ramanathan R, eds. Proc. of the 13th ACM Int'l Conf. on Mobile Computing and Networking. New York: ACM, 2007. 290–301. [doi: 10.1145/1287853.1287888]
- [13] Gilbert R, Johnson K, Wu SM, Zhao BY, Zheng HT. Location independent compact routing for wireless networks. In: Petrioli C, Ramjee R, eds. Proc. of the 1st Int'l Workshop on Decentralized Resource Sharing in Mobile Computing and Networking. New York: ACM, 2006. 57–59. [doi: 10.1145/1161252.1161267]
- [14] Chan CF. Mole: Multi-Hop object location in wireless mesh networks [Ph.D. Thesis]. Hong Kong: Hong Kong University of Science and Technology, 2008.
- [15] Guo DK. Research on peer-to-peer networks based on Kautz digraph and bloom filters [Ph.D. Thesis]. Changsha: National University of Defense Technology, 2008 (in Chinese with English abstract).
- [16] Kumar A, Xu J, Zegura EW. Efficient and scalable query routing for unstructured peer-to-peer networks. In: Knightly E, Makki K, eds. Proc. of the IEEE INFOCOM 2005. Washington: IEEE Computer Society, 2005. 1162–1173. [doi: 10.1109/INFOCOM.2005.1498343]
- [17] Terpstra WW, Kangasharju J, Leng C, Buchmann AP. Bubblestorm: Resilient, probabilistic, and exhaustive peer-to-peer search. In: Taft N, Feldmann A, eds. Proc. of the ACM SIGCOMM 2007. New York: ACM, 2007. 49–60. [doi: 10.1145/1282380.1282387]
- [18] Jelasity M, Montresor A, Jesi GP. The peersim simulator. 2009. <http://peersim.sf.net>

#### 附中文参考文献:

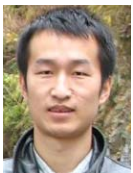
- [2] 李东升. 基于对等模式的资源定位技术研究[博士学位论文]. 长沙: 国防科学技术大学, 2005.
- [15] 郭得科. 基于 Kautz 图和 Bloom 滤波的对等网络研究[博士学位论文]. 长沙: 国防科学技术大学, 2008.



朱桂明(1981—),男,江苏兴化人,博士生,主要研究领域为对等计算,数据中心网络.



金士尧(1937—),男,教授,博士生导师,CCF高级会员,主要研究领域为分布式计算,复杂系统仿真,实时系统,虚拟现实.



郭得科(1980—),男,博士,副研究员,CCF会员,主要研究领域为无线多跳网络,对等计算,Bloom Filters,数据中心互联.