高技术通讯 2010 年 第 20 卷 第 12 期:1247 ~ 1252

doi:10.3772/j.issn.1002-0470.2010.12.007

无结构化 P2P 网络的路由算法①

徐海湄②* ** 卢显良* 齐守青** 彭永祥*

(*电子科技大学计算机学院 成都 610054)

(**解放军重庆通信学院 重庆 400035)

摘 要 针对无结构化 P2P 网络的洪泛搜索与随机漫步机制的盲目性,提出了一种利用 Hash 函数与 M-tree 技术将文件聚类后,再利用路由表完全分布式存储索引指针的新的路由算法。该算法使每个节点的路由表主要记录拥有各类资源的高能力节点指针,并利用概率统计的方法不断地更新路由表项。当节点收到搜索以后,通过查询路由表,只需一跳就进入能以最大概率回应的节点处查找,并能以较低的网络时延命中多个优质资源副本,达到了高速并行下载的目的。仿真试验和数学分析表明该算法有效地减少了盲目搜索造成的网络流量,提高了查找成功率,并且具有越稀缺的资源越容易找到的特性。

关键词 无结构化 P2P 网络,一跳式路由算法(OHRA),洪泛,随机漫步,盲目搜索,搜索成功率

0 引 言

无结构化 P2P 文件共享网络的应用日益广泛,但是网络的各种搜索算法如洪泛(Flood)[1]与随机漫步(Random Walk)[2]等属于盲目搜索,每次搜索都会产生大量的网络流量,对网络造成了很大的负担,影响了系统的可扩展性。而且盲目搜索只适合搜索流行资源,与结构化网络采用分布式 Hash 表(DHT)技术查找资源的方式相比,无结构化网络中稀缺资源的获得比较困难。文献[3]的研究表明,在 Gnutella 中,尽管符合搜索要求的稀缺资源在网络中存在,但是有 18%的搜索得不到任何响应。

当前该领域的研究工作主要集中在利用各种算法提高各类资源的搜索命中率。SQR(scalable query routing)[4]路由机制采用指数衰减的 Bloom Filter 来扩散文件的索引信息,形成以目标资源为圆心的一个感应区域,只要搜索请求进入这个区域,便会被路由到目标处,但是此感应区域的有效半径较短,超出则因为噪音过大,假命中率急剧上升。PoP (a parallel collaborative and probabilistic search mechanism)[5]在 SQR 的基础上进行了改进,采用分布式 Bloom Filter 发布文件信息,每个节点将接收到的 Bloom Filter 向量分成几个子向量,并按照邻居节点的带宽能力扩散。ASAP(advertisement-based search algorithm for unstructured peer-to-peer systems)[6]事先发布文件索引信息,每个节点只缓存与自己兴趣相关的信息,当需要搜索文件时,只需查询本地索引信息,便可以高概率获得搜索结果。但是无结构网络中节点的频繁加入与离开使索引信息的维护费用比较大,因此 Makalu[7]利用网络结构的优化获取较高的搜索命中率。文献[8]将洪泛、随机漫步、索引机制、一跳复制等方法结合在一起使用,并比较了各种搜索算法在规则网与幂律网中的性能参数。

本文提出了一种新的一跳式路由算法(one-hop routing algorithm,OHRA),将任意搜索直接路由到能以最大概率回应的节点处,从而保证在低耗费、低时延的约束条件下提高搜索命中率。

1 完全分布式的文件聚类

文件聚类指的是将无结构化网络中的所有文件资源按照属性归为 N 类,N 越大,文件聚类的划分粒度越细。

当前文件聚类的方法主要是基于文件属性的分类,每个文件属性用一个 n 维向量[9]或者一个二进

① 国家自然科学基金(10577007)与重庆市重点自然科学基金(CSTC, 2007ba2017)资助项目。

② 女,1974 年生,博士生;研究方向:分布式计算,对等网络技术;联系人,E-mail:xuhaimei@uestc.edu.cn

(收稿日期:2009-08-26)

制比特串[10]表示，然后用 Euclidian Distance 或者 Hamming Distance 来计算两个文件的相似性，但是这两种方法不适用于全局网络文档资源的聚类。本文提出两种技术嵌套使用实现文件资源的聚类：(1)相同文件的鉴定与聚类；(2)相似文件的聚类。

1.1 相同文件的鉴定与聚类

相同文件即同一个文件的多个副本，基于 Hash 算法[11]实现聚类。

Hash 函数就是把任意长度的输入变成固定长度的输出。数学表述为：$h = H(f)$，其中 $H()$ 表示单向 Hash 函数，f 为任意长度明文，h 为固定长度的输出值，称为数字摘要。

用 Hash 函数计算文档的数字摘要，数字摘要相同的文档则判定它们的内容是相同的，被视为相同文件。由于 Hash 函数具有不可逆运算的特性，即使网络上的病毒文件伪装成流行文件的名字，通过比较摘要信息很容易将其识别出来，因此具有一定的抗攻击特性。

根据下面给出的算法 1 将相同文件以索引指针的形式关联在一起，形成一个多副本关联网络。关联方法采用主动探测[12]方法和基于搜索反馈结果[12,13]的方法。主动探测指节点主动发起搜索，将具有相同数字摘要的文件相互关联。基于搜索反馈的方法指的是节点充分利用每次的搜索结果将摘要相同的副本相互关联。

$$\begin{cases} h_1 = H(f_1), h_2 = H(f_2) \cdots h_i = H(f_i) \\ if(h_1 = h_2 = \cdots = h_n) \\ then(f_1 = f_2 = \cdots = f_n) \\ i = 1,2,\cdots,n \end{cases}$$

算法 1 相同文件的聚类算法

下面分析一下利用 Flood 方法搜索到一个副本与获取到多个副本的搜索效率。

定理 1：设文件 A 通过算法 1 计算后完全相同的副本数量为 r，r 随机分布在 n 个节点的无结构化 P2P 网络中。设用 Flood 方法遍历了 x 个节点后才能保证收集到所有副本，$p(x)$ 表示成功收集齐 r 个副本的概率，则

$$p(x) = \binom{x-1}{r-1}(\frac{r}{n})^r(1-\frac{r}{n})^{x-r} \quad x = r, r+1, \cdots \tag{1}$$

证明：为了决定 $p(x)$，假设 s_1 代表搜索了 $x-1$ 个节点后，已经收集到了 $r-1$ 个副本，并且 s_2 表示在第 x 个节点处收集到了第 r 个副本的事件，由于 s_1 与 s_2 事件相互独立，所以

$$p(x) = P(s_1 \cap s_2) = P(s_1)P(s_2) \tag{2}$$

因为 $p(s_1)$ 符合参数为 $(x-1, r-1)$ 的二项式分布，即

$$p(s_1) = \binom{x-1}{r-1}(\frac{r}{n})^{r-1}(1-\frac{r}{n})^{x-r} \quad x = r, r+1, \cdots \tag{3}$$

而 $p(s_2)$ 为

$$p(s_2) = \frac{r}{n} \tag{4}$$

将式(3)与式(4)带入(2)得到式(1)。

定理 2：对于文件 A 的 r 个副本随机分布在 n 个节点的网络中，设用 Flood 遍历了 x 个节点后才能保证收集到所有副本，需要遍历 y 个节点才能保证找到一个副本，则 x 的数学期望为 n，y 的数学期望为 n/r，即 $E(x) = n, E(y) = n/r$。

证明：由定理 1 得

$$P(x) = \binom{x-1}{r-1}(\frac{r}{n})^r(1-\frac{r}{n})^{x-r}, x = r, \cdots, n \tag{5}$$

$$\begin{aligned} E(x) &= \sum_{x=r}^{n} xp(x) \\ &= \sum_{x=r}^{n} x\binom{x-1}{r-1}(\frac{r}{n})^r(1-\frac{r}{n})^{x-r} \\ &= \sum_{x=r}^{n} r\binom{x}{r}(\frac{r}{n})^r(1-\frac{r}{n})^{x-r} \\ &= r(\frac{r}{n})^r\sum_{x=r}^{n}\binom{x}{r}(1-\frac{r}{n})^{x-r} \\ &= r(\frac{r}{n})^r\Big[1 + \binom{r+1}{1}(1-\frac{r}{n})^1 \\ &\quad + \binom{r+2}{2}(1-\frac{r}{n})^2 + \ldots\Big] \\ &= r(\frac{r}{n})^r\frac{1}{(\frac{r}{n})^{n+1}} = n \end{aligned} \tag{6}$$

由式(5)得，当 $r=1$ 时就变为找到一个副本需要遍历的节点数为 y 的概率

$$P(y) = (\frac{r}{n})(1-\frac{r}{n})^{y-1}, y = 1, \cdots, n$$

同理推导出

$$E(y) = \sum_{y=1}^{n} yp(y) = \sum_{y=1}^{n} y(\frac{r}{n})(1-\frac{r}{n})^y = \frac{n}{r} \tag{7}$$

由式(6)和式(7)得出 $E(y) = \frac{1}{r}E(x)$，这说明相同副本相互关联网络的建立意味着用搜索到一个副本

的网络费用可以找到多个副本,这样便于并行下载文件,从而大大缩短时延,提高了下载成功率。

1.2 相似文件基于 M-tree 结构的聚类

相似文件按照 M-tree 结构(如图 1 所示)聚类,例如 $M=16$,即每一层分为 16 类,以 16 进制表示为 0 – F,按照资源总类划分 k 层,则资源最多可以划分为 16^k 类。例如,图 1 中 ID 号为 01F 的文件资源表示小说类。相似文件的聚类同 1.1 采用基于搜索反馈的方法与主动探测的方式相互关联。相似文件中完全相同的文件又按照算法 1 关联在一起。

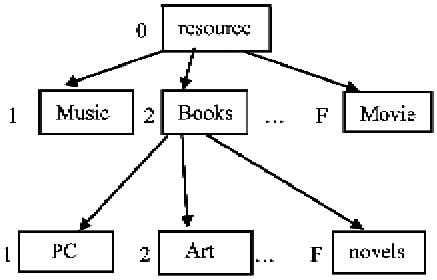

图 1 相似文件聚类的树形结构

2 一跳式路由算法 OHRA

2.1 路由表的建立

路由关联索引表(routing associative index table, RAIT),简称路由表,即本地节点搜索文档的关联索引表。每个节点保存的路由表如表 1 所示,路由表中的第 i 行 $X_{ij}(i=1,\cdots,N;\ j=1,\cdots,k)$ 表示 k 个具有 i 类资源的节点指针,k 个节点按照权重,即搜索命中率排列。搜索命中率根据以往的搜索结果统计分析得出,见 2.3 节。

无结构化网络是动态变化的,当新节点加入系统时,路由表是空的,在节点的生存期内,节点对自身的搜索结果进行统计分析,填充相应的路由表项。为了尽快获得更多路由信息,新节点可以主动发起搜索请求,定位系统中各类资源的 k 个节点,或者向邻居节点请求路由信息,每隔一段时间再根据网络的变化情况,不断更新路由表项(见 2.3 节)。

表 1 路由关联索引表

资源 ID	关联节点列表
1	$X_{11}, X_{12},\cdots,X_{1k}$
2	$X_{21}, X_{22},\cdots,X_{2k}$
…	…
N	$X_{N1}, X_{N2},\cdots,X_{Nk}$

路由表存储费用评估:每个节点使用 $k\cdot N\cdot S$ 字节来存储路由表大小。S 表示每个节点指针所占据的字节数,N 表示文件资源种类。在实际应用中,节点可以根据自身能力决定 k 的大小,从而使存储费用线性递减或者递减。

2.2 路由效率分析

定理 3 n 个节点的网络中,X_i 个节点拥有 i 类文件资源,y 个节点拥有文件 B 的副本,文件 B 属于 i 类资源,即 $y\subset X_i$,利用路由表 1,用 Flood 的方法搜索文件 B,则搜索到文件 B 需要的平均步数即 TTL 为 $\dfrac{\ln(\frac{(k-2)x_i}{ky}+1)}{\ln(k-1)}$,而不用路由表 1,搜索到文件 B 需要的平均步数 TTL 为 $\dfrac{\ln(\frac{(k-2)n}{ky}+1)}{\ln(k-1)}$.

证明: i 类文件资源的聚类中每个节点含有文件 B 的概率为 $p=\frac{y}{x_i}$;设利用路由表 1 需要遍历 ζ 个节点后才能搜索到文件 B 的副本,而不利用路由表 1 则需要遍历 β 个节点才能发现文件 B 的副本。

设遍历 h 个节点发现文件 B 的概率为 $p(\xi=h)$,则 $P(\xi=h)=(1-p)^{h-1}p$,令 $q=1-p$,则数学期望 $E(\zeta)$ 为

$$E(\xi)=p+2pq+3q^2p+\cdots+nq^{n-1}p$$
$$=(1+2q+3q^2+\cdots+nq^{n-1}+\cdots)p$$
$$n=1,2,\cdots,\infty$$

令

$$S_n=1+2q+3q^2+\cdots+nq^{n-1}$$
$$qS_n=q+2q^2+3q^3+\cdots+nq^n$$

推导出

$$(1-q)S_n=1+q+q^2+\cdots+q^{n-1}-nq^n$$
$$S_n=\frac{1-q^n}{(1-q)^2}-\frac{nq^n}{1-q}$$

因为 $0<p<1$,知 $0<q<1$,则 $\lim\limits_{n\to\infty}q^n=0$,因此得出 $1+2q+3q^2+\cdots+nq^{n-1}=\lim\limits_{n\to\infty}S_k=\dfrac{1}{(1-q)^2}$

从而 $E(\xi)=\frac{x_i}{y}$;同理可以推导出 $E(\beta)=\frac{n}{y}$。

因此可以得出利用路由表 1 平均需要遍历 x_i/y 个节点能搜索到文件 B,而不使用路由表 1 只用传统的 k 路 Flood 方法平均需要遍历 n/y 个节点才能搜索到文件 B。

设利用路由表 1 和 Flood 的方法,需要 $L=$TTL 步才能遍历 $E(\xi)$ 个节点,则

$$k+k(k-1)+\cdots+k(k-1)^{L-1}=x_i/y$$

由 $\frac{k[(k-1)^L-1]}{k-2}=\frac{x_i}{y}$ 推出

$$L=\frac{\ln(\frac{(k-2)x_i}{ky}+1)}{\ln(k-1)} \tag{8}$$

同理可以得出不利用路由表1,则Flood需要的步数为

$$L=\frac{\ln(\frac{(k-2)n}{ky}+1)}{\ln(k-1)} \tag{9}$$

证毕。

比较式(8)与式(9)可以看出,文件聚类后,利用路由表1,则搜索的TTL主要取决于 x_i/y,与网络规模 n 无关,而且拥有同类资源的节点数量 x_i 越少,副本数量 y 越多,L 越少。说明利用路由表1后,只需遍历更少的节点,以更低的时延就可以搜索到目标资源,而且稀缺资源也象流行资源一样容易地被找到。而由式(9)看出,不使用路由表1,搜索命中率随着网络规模的扩大而大幅度降低。

2.3 路由表的更新与维护

每个节点对历史搜索的反馈结果进行统计分析,以定期更新路由表。

为了描述路由更新算法,首先引入几个概念:

T: 路由表更新周期;

S_u :在第 i 个时间周期 T 内对节点 u 发出的搜索请求的反馈节点集合;

$R_i(u,v)$: 在第 i 个时间周期 T 内,节点 u 对搜索反馈节点 v 的评价,$0 \leqslant R_i(u,v) \leqslant 1$;

$q_i(u,v)$:第 i 个时间周期 T 内节点 u 从节点 v 收到的有效命中搜索数量;

β :衰减指数,$0<\beta<1$;令

$$\phi(u,x)=q(u,x)/\sum_{y\in S_u}q(u,y)$$

$$\phi_{\max}=\max\{\phi(u,x)\mid x\in S_u\}$$

则节点 u 对节点 v 的评价更新过程为

$$R_i(u,v)=\beta R_{i-1}(u,v)+(1-\beta)\frac{\phi(u,v)}{\phi_{\max}} \tag{10}$$

根据式(10)每隔一个周期 T, 节点 u 将取前 k 个权重较大的节点更新路由表1中的行表项,从而完善相似文件的聚类。根据算法1,返回相同副本的节点以 $R=1$ 更新相应的路由表项,从而完善相同文件的聚类。

路由表的更新与维护费用 oh 估计:由上述分析可以得出,每个节点的路由表的更新维护费用 oh 与 $\frac{kN}{T}$ 成正比,即 $oh \propto \frac{kN}{T}$,其中 k 表示路由表的每一行包含的节点指针的个数,N 表示文件资源的种类,T 表示路由表更新的周期。更新周期越短,更新的越频繁,oh 越大。更新周期的长短取决于网络的搅动情况与资源的访问频率。结构化网络利用DHT方式快速定位资源所需要的路由费用为 $O(\log n)$,其中 n 为网络中节点的个数,因此随着网络规模的扩展,路由费用快速增长。而OHRA算法的路由费用完全由每个节点独立承担,与 n 无关,不会随着网络规模的增长快速增长,因此具有可扩展性。

3 仿真实验

OHRA的设计目的是在无结构化网络中低时延、低开销地获得多个优质的目标文件。因此基于开放源代码的仿真器 PeerSim[14],采用 Java 编码实现了基于 Gnutella(Ver0.4)协议的无结构化网络环境。仿真程序在单机(CPU:双核,3.16G;RAM :2GB)上运行。网络中各个节点既是服务器又是客户机,每个节点发出的查询信息以 Flood 方式向邻居节点扩散,TTL 值被用来控制消息扩散的跳数,TTL值越大说明查询遍历的节点数量越大,消息开销越大,查询效率越低,所以本文的仿真实验用查询消息终止前遍历的节点数目衡量消息开销。本文将OHRA与经典的盲目搜索算法 Flood 和当前较新的研究成果 SQR,PoP 进行了比较。

每个实验做10次,取平均值。实验分四组进行。

第一组:10^4 个节点的仿真系统中,设置10类资源,每类资源随机分布在 10^3 个节点中,每类资源各取一个样本,分别设置10个副本,20个副本随机分布在网络中。对于每一个样本资源,随机选择节点发起搜索,直到成功搜索到至少一个副本为止。其中路由表1中 k 均取值为10,消息开销用搜索结束时遍历的节点数目衡量。由图2看出,在相同副本情况下,用OHRA的消息开销远小于不用OHRA的Flood搜索算法。而副本越多,消息开销越小。这是因为OHRA将搜索集中在以最高概率回应的节点处,而盲目搜索 Flood 是随机遍历节点的。因此OHRA只需要遍历少量节点就能找到目标,节省了网络带宽,这也验证了定理3的分析。

第二组:10^4 个节点的系统中,取一个 i 类资源

的文件样本 A, i 类资源随机分布在 10^3 个节点上,A 设置 10 个副本,随机分布。随机发起搜索,搜索到所有副本为止。搜索到的副本数目与消息开销之间的关系用图 3 表示。因为样本 A 的多个副本相互关联,OHRA 算法搜到一个副本很快就可以找到其它副本,因此消息开销变化不大。而 Flood 产生的消息开销几乎是线性增长。这验证了定理 2 的分析。

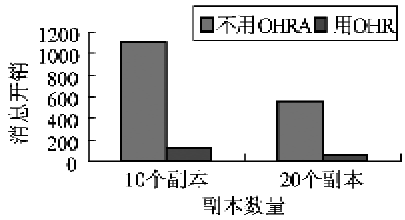

图 2　不同副本情况下搜索成功的消息开销

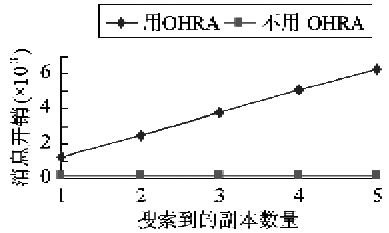

图 3　消息开销随搜索到的副本数量的变化情况

第三组:10^4 个节点的系统中,拥有 i 类资源的节点数量从 200,400,600,800,1000 变化, i 类文件 B 有 5 个副本随机分布在网络中,随机发起搜索,直到成功搜索到至少一个副本为止,比较消息开销的变化情况。由图 4 看出,消息开销与拥有同类资源的节点数量有关,与网络规模无关,资源越稀缺,消息开销越少。这是因为文件聚类以后,拥有同类资源的节点数量越少,则遍历所需要的时延就越少。

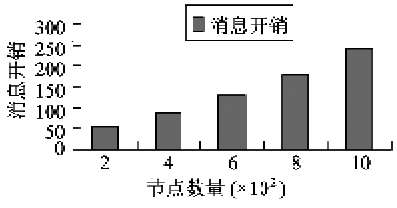

图 4　消息开销随拥有同类资源的节点数量的变化情况

第四组:10^4 个节点的系统中,取一个 i 类资源的文件样本 C, 分别使用 SQR,PoP 机制扩散文件的索引信息,然后令 $L1$ 表示任意节点与样本 C 的距离,用 hops 表示, HR (Hit Rate)表示搜索成功率。相同的时延条件下比较一下 OHRA 与 SQR,PoP 的搜索成功率随 $L1$ 的变化情况。由图 5 可以看出,随着 $L1$ 的增大,SQR,PoP 的搜索成功率迅速降低,这是因为两者都采用 Bloom Filter 扩散文件索引信息, $L1$ 增大后,噪音增大,假命中率增大。OHRA 由于定期更新的路由表,同类文件的关联定时更新,受 $L1$ 的影响并不大,搜索命中率一直保持在 80% 以上。

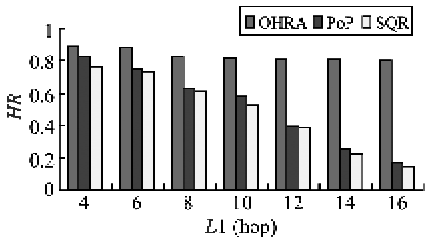

图 5　搜索成功率随 $L1$ 的变化情况

4　结　论

无结构化 P2P 系统采用盲目搜索造成了大量的网络流量,影响了系统的可扩展性。本文提出了一种新的一跳式路由算法 OHRA,以路由表的形式动态地实现了相似文件资源与相同文件资源的双重聚类。每个节点只需维护少量的路由信息,就能以更高的成功概率、更少的消息开销发现更多的相同副本。每一次的搜索结果被充分地利用,定期更新路由表,使网络的搜索性能逐渐加强。数学分析与实验结果表明,OHRA 减少了网络流量,提高了搜索成功率。

参考文献

[1] Klingberg T, Manfredi R. The Gnutella protocol specification v4.0. http://dss.clip.com/Gnutellaprotocol0.4.pdf, 2007-07-09

[2] Chawathe Y, Ratnasamy S, Breslau L, et al. Making Gnutella-like P2P systems scalable. In: Proceedings of 2003 Conference on Applications, Technologies, Architectureds and protocols for Computer Communications, Karlsruhe, Germany, 2003. 407-418

[3] Thau L B, Joseph M H, Ryan H, et al. Enhancing P2P file-sharing with an internet-scale query processor. In: Proceedings of the 30th conference on VLDB, Toronto, Canada, 2004. 432-443

[4] Kumar A, Xu J W, Zegura E. Efficient and scalable query routing for unstructured peer-to-peer networks. In: Proceed-

ings of the 2005 IEEE International Conference on Computer Communications, Miami, USA, volume 2, 2005. 1162-1173

[5] Zhan Y M, Li D S, Chen L, Lu X C. Collaborative search in large-scale unstructured peer-to-peer networks. In: Proceedings of the International Conference on Parallel Processing (ICPP 2007), Xi'an, China, 2007. 1-7

[6] Gu P, Wang J, Cai H L. ASAP: an advertisement-based search algorithm for unstructured peer-to-peer systems. In: Proceedings of the International conference on Parallel Processing (ICPP 2007), Xi'an, China, 2007. 1-8

[7] Acosta W, Chandra S. Improving searching using fault-Tolerant oberlay in unstructured P2P Systems. In: Proceedings of the International Conference on Parallel Processing (ICPP 2007), Xi'an, China, 2007. 1-11

[8] Gkansidis C, Mihail M, Saberi A. Hybrid search scheme for unstructured peer-to-peer networks. In: Proceedings of the 2005 IEEE International Conference on Computer Communications, Miami, USA, volume 3, 2005. 1526-1537

[9] Tang C, Xu Z, Dwarkads S. Hybrid global-local indexing for efficient peer-to-peer information retrieval. In: Proceedings of the 1st Symposium on Networked Systems Design and Implementation, San Francisco, USA, 2004. 211-224

[10] Das T, Nandi S, Ganguly N. Community based search on power law networks. In: Proceedings of the 3rd international conference on COMSWARE, Bangalore, India, 2008. 279-282

[11] Nguyen D T, Soh B. Key management for lightweight ad-hoc routing authentication. In: Proceedings of the 4th International Symposium on Wireless Pervasive Computing, Melbourne, Australia, 2009. 1-7

[12] Luo X C, Qin Z G, Geng J, et al. Efficient multi-source location in unstructured P2P systems. In: Proceedings of the International conference on Parallel Processing (ICPP 2007), Xi'an, China, 2007. 61-65

[13] Sun Q X, Garcia-Molina H. SLIC: A selfish link-based incentive mechanism for unstructured peer-to-peer networks. In: Proceedings of the 24th International Conference on Distributed Computing Systems, Tokyo, Japan, 2004. 506-515

[14] The PEERSIM website. http://peersim.sourceforge.net, 2009-4-22.

A novel one-hop routing algorithm for unstructured P2P networks

Xu Haimei[*,**], Lu Xianliang[*], Qi Shouqing[**], Peng Yongxiang[*]

([*] College of Computer Science and Engineering, University of Electronic Science & Technology of China, Chengdu 610054)

([**] College of Chongqing Communications of People's Liberation Army, Chongqing 400035)

Abstract

To improve the current status that the existing blind searching schemes for unstructured P2P networks such as the Flood and the Random Walk incur too much traffic load, the paper presents a novel routing algorithm that can research resources efficiently. The algorithm fully utilize the Hash function, the M-tree technique and the feedbacks of every query to construct clusters of resources and the routing table that maintains pointers of high-capacity nodes. When the peer receives a new query, according to the routing table, it directly forwards the query to nodes with high hit probability. Multiple high-quality replicas of both popular resources and rare resources can be located with the minimum overhead, thus parallel download can be guaranteed. The mathematical analysis and simulation results show that the routing algorithm improves the search efficiency with the high hit rate and the low bandwidth overhead.

Key words: unstructured P2P network, one-hop routing algorithm (OHRA), flooding, random walk, blind searching, hit probability for searching