

# TARSB:一种基于相对搜索介数的P2P拓扑演化策略

王义麟 陶 焯 陈建忠 陆桑璐 陈道蓄

(南京大学计算机系计算机软件新技术国家重点实验室 南京 210093)

**摘要** 近几年来,对等网络(P2P networks)得到了很好的发展,然而其巨大的网络开销也逐渐成为阻碍其进一步推广的主要问题。在其最重要的应用领域——P2P文件共享系统中,早期的系统往往基于一种单一的拓扑结构,忽略了成员之间的差异对网络本身的影响,使得其理想的结构与实际网络之间存在很大差距,难以达到预期的搜索效率。针对P2P文件共享系统进行研究,提出了相对搜索介数(Relative Search Betweenness)来刻画各种异构性对搜索过程的综合影响,并设计了一种基于相对搜索介数的拓扑演化策略:TARSB。该策略通过动态自适应地调整覆盖网络结构来平滑各结点在搜索性能上的差异,从而达到提高网络整体搜索性能的效果。实验表明,TARSB策略在保证搜索命中率的情况下,大大降低了搜索报文的数量,并使网络自主演化成一种更加稳定和高效的拓扑结构。

**关键词** 可交互,流媒体,IP组播,应用层组播

## TARSB:Topology Adaptation Based on Relative Search Betweenness in P2P Networks

WANG Yi-lin TAO Ye CHEN Jian-zhong LU Sang-lu CHEN Dao-xu

(State Key Laboratory of Novel Software Technology, Department of Computer Science and Technology,  
Nanjing University, Nanjing 210093, China)

**Abstract** Peer-to-Peer (P2P) applications become very popular in current networks, but the large traffic consuming limits their further spread. In the field of P2P file-sharing system, one of the most important applications of P2P networks, early systems usually obtain a single and flat topology, which ignores the heterogeneity of peers. To give a comprehensive measure of heterogeneity in P2P file-sharing systems, especially in the searching phase, we defined a parameter, called Relative Search Betweenness (RSB). We also proposed a topology adaptation algorithm, to accommodate the natural heterogeneity presented in most peer-to-peer systems with RSB. This algorithm smoothes search capabilities of nodes by adjusting the network topology dynamically, to improve the global search capability of the networks. The simulation results show that, with the same (even better) hit ratio, the traffic of flooding search method declined 10% in the adapted topology. And the network becomes more reliable and effective.

**Keywords** Interactive, Streaming media, IP multicast, Application layer multicast

## 1 引言

Peer-to-Peer(简称P2P)作为一种重要的网络结构模式,与C/S模式相比,其主要优势在于克服了服务器在文件传输时的瓶颈问题,把原本集中的开销分散到整个网络中。著名的P2P文件共享系统有Gnutella, FastTrack, eDonkey等,根据文献[1]的数据,目前这3种系统的用户数目总计已超过800万,而且仍不断有新用户加入。

在P2P网络中,结点之间通过邻居结点表建立起逻辑连接,形成了一种独立于物理链路的拓扑结构(通常用有向图表示),即所谓的覆盖网络(Overlay Networks)。在本文中,P2P网络结构即指的是覆盖网络结构。在非结构化的P2P文件共享系统中,用户对文件的搜索过程大致如下:向邻居结点发送搜索请求,邻居结点检查是否拥有目标文件,若没有则向自

己的邻居结点转发请求。当结点命中搜索请求后,会产生成功应答报文,应答报文沿搜索报文的传递路径返回搜索发起结点。搜索发起结点接受到成功应答后,将从命中的结点处下载文件。

对于非结构化的P2P系统而言,由于缺乏精确的定位信息,搜索报文的转发在本质上是随机的。因此,对于一个搜索任务往往产生多个搜索报文,并发地在网络中传播,最典型的方法就是洪泛法(flooding)。这类方法的最大缺点在于,大量无用的搜索报文充斥着网络,消耗大量的带宽资源,甚至影响其它报文的正常传输。为了克服洪泛法的缺点,最直接的方法就是减少报文的转发量,仅向邻居中的某些可能“最适合”结点提交搜索请求,如以网络带宽<sup>[7]</sup>、结点满意度<sup>[6]</sup>或信任度<sup>[9]</sup>为依据来选择。然而全面评估结点的“适合程度”是相当困难的,不但要考察现实网络中可能出现的各种影响因素,而

到稿日期:2008-04-02 本文受国家高技术研究发展计划863项目(No. 2006AA01Z199),国家自然科学基金(No. 60573106)资助。

王义麟 硕士研究生,研究方向为并行处理和分布式计算,E-mail:wangyl@dislab.nju.edu.cn;陶焯 讲师,研究方向为并行处理和分布式计算;陈建忠 硕士研究生,研究方向为并行处理和分布式计算;陆桑璐 教授,博士生导师,研究方向为并行处理和分布式计算;陈道蓄 教授,博士生导师,研究方向为并行处理和分布式计算。

且需要定义一个参量来体现各种因素综合作用的效果。本文提出了一种基于相对搜索介数的评估方法及相应的网络结构调整算法。与以往的利用某一方面异构特性的方法不同,本文定义的相对搜索介数是多种异构性对搜索过程影响的综合反应,并且可以实时反映结点当前的搜索能力。此外,本文所介绍的基于相对搜索介数的拓扑演化策略具有很强的灵活性,能更准确和高效地发现搜索能力高的结点。调整之后的网络结构在搜索性能提高的同时,也大大降低了带宽的开销。

本文首先回顾了利用异构性调整网络结构的相关工作,第2部分给出相对搜索介数的定义,第3部分详细介绍拓扑演化策略 TARSB,第4部分是实验模拟和性能分析,最后是总结部分。

## 2 相关工作

最初的 P2P 系统为了体现对等性,假设网络中的所有成员都是平等的,这种平等性隐含地包括结点的带宽、CPU 能力和存储能力,又包括网络文件的种类、数量和分布情况。基于这种假设的系统有 Gnutella(0.4)<sup>[2]</sup>, Chord<sup>[3]</sup>, CAN<sup>[4]</sup> 等。随着 P2P 系统的普及,大量真实数据<sup>[14-17]</sup>表明, P2P 网络中存在着大量的不平衡情况,如结点的连接度呈类似 Power-Law 分布<sup>[16]</sup>,用户的在线时间呈 Pareto 分布<sup>[17]</sup>,而文件分布<sup>[15]</sup>以及结点物理属性的区别<sup>[14]</sup>也是非常明显的。这一系列结果表明原先的平等性假设过于粗糙。为了构建合理的 P2P 结构,进一步提高 P2P 网络的性能,必须充分考虑结点间的异构特性。

目前利用异构性来对网络结构进行优化的工作,大致可分为3类:

1) 物理异构性,包括网络带宽、CPU 处理能力、存储延时以及结点间的物理距离等。这些参数的特点是能直接反映出结点在物理性能上的优劣,是最直观、最易于利用的异构特性<sup>[5-8]</sup>。文献[5]关注的是结点的网络带宽,定义结点的容量为一段时间内该结点所能处理的报文数量,其核心思想是当发生过载时,结点主动把搜索频度最高的邻居重定向到其它的空闲结点上,从而减轻自身负载。此外,现行的大多数超结点(supernode)结构,如 Gnutella0.6<sup>[2]</sup>,也都利用物理信息作为选择超结点的标准。

2) 逻辑异构性,包括结点的度、结点间的距离(跳数)等覆盖网络拓扑信息和结点的历史信息(邻居结点被成功命中的次数、在线时间等等)。这些量能间接代表结点的性能,但往往脱离了物理网络或是过分依赖历史信息,所以并不完全可靠<sup>[9-11]</sup>。文献[9]中根据结点历史信息,对每一个交互过的结点计算其本地信任度(local trust value),在最大链接数受限的情况下,以此值作为建立和中断链接的依据,从而使使得总的网络可信度达到最大。

3) 语义异构性,即文件固有的异构性,包括文件的内容、大小、热门度以及副本数量等等。这些异构特性是面向具体应用的,为文件共享系统所独有。文献[12]提出了利用结点上的文件内容对结点进行分类,从而按照内容来构建网络结构,即形成语义覆盖网(Semantic Overlay Networks)。然而,由于在现实环境中很难明确划分结点的语义信息,此类方法目前还没有取得突破性的进展。

以上的工作大多专注于异构性的某一个方面,不足之处

在于忽略了各种异构性的叠加效果。举例而言,网络带宽(物理异构性)是评价结点优劣的一个重要指标,许多优化算法<sup>[5-7]</sup>都把带宽大的结点安排在网络中比较重要的位置上。但如果忽略了该结点的动态性(逻辑异构性),即如果这些处于核心位置的结点很不稳定,则其对网络整体性能的破坏将远大于普通结点。又如,一些机器性能很好(物理异构性)但文件资源相对稀少(语义异构性)的结点,在忽略语义异构性的优化算法的指导下,逐渐演化成为网络的主干。这无形中可能把那些拥有丰富文件资源但物理性能并不出色的结点推向了网络的边缘,从而增加了搜索的长度,也增大了网络开销。各种异构性之间的相互影响错综复杂,寻求一个对每一种异构性对网络的独立影响进行刻画,再将其效果进行叠加的方法到目前为止仍是个难题。因此,本文直接考察结点的搜索效率,即各种异构性对于搜索的叠加效果,从而绕开了上述难题。文献[9]与本文的设计思路比较相近,但它是由每个结点根据自身的搜索历史对交互过的结点进行评估,无法对结点的搜索性能形成一个公正客观的评价,而且该文默认结点对网络中的所有结点都有一个初始的评价(一旦初始评价为0,则永远不会与其进行交互)。而本文是通过搜索介数形成对结点搜索性能的全局有效的评估,对于从来未发生过交互的备选结点,本地结点可以根据其搜索介数来进行选择。

## 3 相对搜索介数的定义

在复杂网络(Complex Networks)中,结点的介数(Betweenness)<sup>[13]</sup>被定义为网络中所有最短路径中通过该结点的数量。设结点  $i, j$  间的最短路径数为  $C(i, j)$ ,其中通过某一个结点  $k$  的数量为  $C_k(i, j)$ ,则结点  $k$  的介数为  $\sum_{i, j \in N} (C_k(i, j)/C(i, j))$ 。从定义可以看出,介数的大小体现了该结点在网络中的重要程度。

借用介数的思想,针对 P2P 网络的搜索问题,我们可以定义结点的搜索介数(Search Betweenness)。

定义1(最短搜索路径, Shortest Search Path) 在请求发起结点与请求命中结点之间已知的最短路径。

不失一般性,假设某个结点发出的搜索请求沿多条路径最终都到达了目标结点,此时最先到达的搜索请求所经过的路径即是此两点间的最短搜索路径。若假设相邻结点间的传输时延相同,则最短搜索路径上的中间结点数应是所有路径中最少的之一。

定义2(结点的搜索介数, Search Betweenness) 通过该结点的最短搜索路径的数量。

其含义是:网络中所有搜索成功的路径中通过该结点的数量。这里的搜索成功路径即是搜索发起结点与搜索命中结点之间的最短路径(假设结点会丢弃重复的报文且不发生拥塞现象)。结点的搜索介数越大,表明该结点搜索成功的次数越多,在一定程度上反映出该结点本身或附近的结点(TTL 跳数之内)成为源结点的可能性比较大。对于搜索介数的计算也十分方便,只要统计通过该结点的成功应答报文的数量即可。

介数是定义在固定网络结构下的,而搜索介数所面对的环境是动态变化的 P2P 网络,不同的应用背景使得二者在计算方法上有所区别。其一,复杂网络领域对于介数的计算存在多种方法<sup>[20]</sup>,其中当  $i, j$  两点间新增一条最短路径时,位于

这条路径上的各个中间结点的介数要加上相应的权值,而端点  $i, j$  本身的介数可以加上这个权值,也可保持不变。但对于搜索介数,除了中间结点外,端点  $i$ (搜索发起结点)和  $j$ (命中结点)的搜索介数也要增加,这是因为在动态网络环境中,某个结点可能在这次搜索任务中是端点,在下次搜索任务中就成为中间结点,只有把作为端点的情况也包括在内,才能体现出该结点每一次搜索的效果,为其他结点在邻居选择时提供可信的参考依据。其二,对于介数,若  $i, j$  两点间有  $k$  条最短路径,则在最终统计时,每条路径上的结点只能加上  $1/k$  的权值。而对于搜索介数,由于重复的报文会被自动丢弃,因此  $i, j$  两点间只会有一条最短路径,不会出现平分权值的情况。其三,对于一次搜索而言,目标文件可能存在于多个结点上,此时每一条成功路径上的结点都应该加上权值。

在实际应用中,搜索成功的绝对次数显然还与搜索的发起次数有关,而搜索的发起次数又与搜索报文的数量有很大关系(对一个结点而言,一个搜索报文就相当于一次搜索),因此对于高度动态性的 P2P 网络而言,我们需要定义结点的相对搜索介数。

**定义 3**(相对搜索介数, Relative Search Betweenness) 一段时间内通过该结点的搜索介数与搜索请求报文数之比。在网络比较稳定的情况下,可用通过该结点的平均成功应答报文数与平均搜索请求报文数之比来近似该值,即

$$\text{相对搜索介数}(RSB) = \frac{\text{平均成功应答报文数}(r)}{\text{平均搜索请求报文数}(s)} \quad (1)$$

在一个相对稳定的网络结构中,结点的 RSB 是由其最大跳数之内所有相关结点的文件种类和数量所决定的。以这些相关结点到起始结点的最大跳数  $k$  为划分标准,可把其划分为  $k+1$  层。设通过结点  $N_0$  的所有搜索报文中还能进行  $i$  跳搜索的数量为  $s_i (i=0, 1, \dots, k)$ , 第  $i$  层的第  $j$  个结点的命中率为  $h_{i,j}$ , 由第  $i$  层返回的成功应答数量为  $r_i$ 。根据式(1)可得:

$$RSB_{N_0} = \frac{r_{N_0}}{s_{N_0}} = \frac{r_0 + r_1 + \dots + r_k}{s_0 + s_1 + \dots + s_k} \quad (2)$$

假设通过  $N_0$  的搜索报文的 TTL 均匀分布,即  $s_0 = s_1 = \dots = s_k = \frac{s_{N_0}}{k+1}$ , 则有

$$r_0 = h_0 \cdot (s_0 + s_1 + \dots + s_k) = h_0 \cdot s_{N_0}$$

$$r_1 = \sum_j h_{1,j} \cdot (s_1 + s_2 + \dots + s_k) = \sum_j h_{1,j} \cdot \frac{k \cdot s_{N_0}}{k+1}$$

...

$$r_k = \sum_j h_{k,j} s_k = \sum_j h_{k,j} \cdot \frac{s_{N_0}}{k+1}$$

代入式(3),得:

$$RSB_{N_0} = \frac{r_0 + r_1 + \dots + r_k}{s_{N_0}} = \frac{h_0 \cdot s_{N_0} + \sum_j h_{1,j} \cdot \frac{k \cdot s_{N_0}}{k+1} + \sum_j h_{2,j} \cdot \frac{(k-1) \cdot s_{N_0}}{k+1} + \dots + \sum_j h_{k,j} \cdot \frac{s_{N_0}}{k+1}}{s_{N_0}} = h_0 + \frac{k}{k+1} \cdot \sum_j h_{1,j} + \frac{k-1}{k+1} \cdot \sum_j h_{2,j} + \dots + \frac{1}{k+1} \cdot \sum_j h_{k,j} \quad (3)$$

式(3)是在网络结构稳定的情况下结点 RSB 的理论值。当网络不稳定时,实际的 RSB 会略微滞后于其理论值,但仍能动态反映出结点当前的相对搜索能力。

定义搜索介数和相对搜索介数的最终目的是通过刻画结

点的搜索能力来分析结点的异构性,从而减小异构性对网络搜索性能的影响。如前所说,结点的异构性大体上可以分为 3 种:物理异构性、逻辑异构性和语义异构性,而搜索介数和相对搜索介数正是这 3 种异构性综合作用的反映。例如在其它因素相同的情况下,结点存储空间越大、所共享的文件越多(物理异构性),则被命中的机会就越大,RSB 就越大。同样,结点越稳定(逻辑异构性),经过它的成功应答报文就越多,RSB 也越大。再如,结点上拥有的文件的受欢迎程度越高(语义异构性),被搜索的概率就越大,RSB 也会得到提高。此外还有许多异构因素能通过 RSB 反映,限于篇幅原因就不一一列举了。下面我们将通过平滑相对搜索介数来提高网络的整体搜索性能。

## 4 基于相对搜索介数的 P2P 拓扑演化策略——TARSB

### 4.1 设计思想

相对搜索介数能动态反映出结点当前的搜索能力,对于任何一个网络优化算法,其最终结构都应该使得网络总 RSB 达到最大。当结点数量相对稳定时,网络的平均 RSB 达到最大。同时,相对搜索介数的差异是结点自身 3 种异构性综合作用的结果,这种差异性在现实 P2P 网络中是无法避免的,而作为一种用于搜索的网络结构,需要尽量平滑这种差异。基于上述考虑,我们提出一种通过平滑相对搜索介数来调整网络结构,降低网络开销的拓扑演化策略 TARSB(Topology Adaption based on RSB)。该策略有两个目标:增大网络的平均 RSB 和减小各结点 RSB 的差异。

TARSB 的设计思路如下:结点的 RSB 是成功报文与搜索报文的比值,其中成功报文的数量取决于出度大小和出边所连接的邻居,而搜索报文的数量取决于入度和自身的搜索频率。因此,可以通过调整结点的出边和入边以及替换邻居来调整结点的 RSB,从而平滑结点与其周围结点 RSB 的差异,同时还要保证网络的平均 RSB 在每一次平滑操作后都能得到提高。

### 4.2 TARSB 算法

#### 4.2.1 参数定义

为了能够进行结构调整,结点需要比较现有邻居结点和备选结点的搜索性能,因此在邻居表中除了结点的物理地址,还需要诸如 RSB 等一些性能参数。一个基本的结点描述信息如表 1 所列,其中 ID, Node Address 和 RSB 分别是目标结点的全局标识符、物理地址和当前 RSB 值,Indegree 是该结点当前的入度。MaxIndegree 是该结点的最大入度,仅当入度未达到最大值时,才可将该结点作为新邻居。通常在进行结构调整时,需要通过发送报文来更新后 3 个参数。获取结点信息的途径有很多,可以通过类似 Gnutella 中的 Ping&Pong 机制,也可以由结点本身向外广播。而本文为了减少网络中报文数量,则是通过成功应答报文来收集其路径上的所有结点信息。

表 1 结点的基本信息

ID	Node Address	RSB	Indegree	MaxIndegree
X	xxx	xx	xx	xx

在本文中,我们利用熵(entropy)来监控结点的相对搜索

介数。设结点  $N_0$  有  $n$  个邻居结点 ( $N_1 \sim N_n$ ), 则结点  $N_0$  的局部熵定义为:

$$entropy_{N_0} = \sum_{i=0}^n \left( \frac{RSB_i}{\sum_{j=0}^n RSB_j} \times \log \frac{RSB_i}{\sum_{j=0}^n RSB_j} \right) \quad (4)$$

当  $entropy_{N_0}$  小于某个阈值  $\sigma$  时, 表明这  $m+1$  个结点的相对搜索介数差异比较大, 需要对其进行调整。

#### 4.2.2 结构调整操作

结构调整操作具体可分为以下 4 种。

**增加最大入度 (increase maximum-indegree):** 当  $RSB_0 \gg RSB_i (i=1, \dots, n)$  时, 表明结点  $N_0$  具有很高的搜索能力, 从全局优化的角度来看, 结点  $N_0$  应该允许更多的接入结点 (即结点  $N_0$  只接受其搜索请求, 但不向其发出搜索请求)。这样做一方面可以提高新接入结点的 RSB, 同时也降低了结点  $N_0$  的 RSB。该操作仅当结点  $N_0$  的入度已达到最大值时才执行。采用逐步调整最大入度的原因主要是为了避免结点同时接纳过多的新邻居而对其搜索性能造成巨大影响。

**删除出边 (remove an out-edge):** 当  $RSB_k \ll RSB_i (i=0, \dots, n \text{ 且 } i \neq k)$  时, 删除结点  $N_0$  与结点  $N_k$  的连接。这样做的理由是, 当结点  $N_k$  的 RSB 很小时, 可以认为其对于结点  $N_0$  的搜索基本起不到有益的作用, 反而会产生大量无用的搜索报文, 浪费网络资源, 因此必须将其删除。这样, 结点  $N_0$  的局部熵也会有所提高。

**替换出边 (replace an out-edge):** 当  $RSB_0 \ll RSB_i (i=1, \dots, n)$  时, 从备选结点中选出一个结点  $x$ , 替换贡献度最小的那个邻居结点。这种情况所反映的问题是, 尽管邻居结点都具有很好的搜索能力, 但它们所覆盖的文件与结点  $N_0$  的搜索兴趣存在较大的差异, 并不能对结点  $N_0$  的搜索起到很大的帮助, 因此需要替换成与结点自身请求更匹配的新邻居。

**增加出边 (add a new out-edge):** 当  $entropy_{N_0}$  小于阈值  $\sigma$ , 且不属于以上 3 种情况时, 即  $RSB_k \gg RSB_i (i=0, \dots, n \text{ 且 } i \neq k)$  时, 从备选结点中选出一个结点  $x$ , 增加为新邻居。需要注意的是, 此时不能盲目除去结点  $N_k$ , 而是应该通过提高自身的 RSB 来缩小差距。为了起到缩小差异的作用, 新结点的 RSB 不能过大。

在现实系统中, 判断某个 RSB 远大于 (或小于) 其它 RSB 的标准, 需要根据系统的实际情况来决定。本文中采用的方法是设定一个参数  $\varphi$ , 当某个 RSB 大于 (或小于) 这  $n+1$  个结点的 RSB 均值的  $\varphi$  倍时, 即认为其远大于 (或小于) 其它的 RSB。

## 5 模拟实验及结果讨论

### 5.1 实验环境及评估标准

我们选用 Peersim1.0<sup>[19]</sup> 作为实验平台, 在其原有基础上扩展了结点、报文和文件等类。各种分布规律和参数如下: 文件的访问遵循 Zipf 分布, 结点的动态性遵循 Pareto 分布, 结点的文件数量按照文献[14]中的分布规律, 结点的接收缓存分为两种 (100 报文/周期和 200 报文/周期)。实验中的结点数为 5000, 文件数为 8000。搜索采用洪泛法, TTL 设为 4。总实验时间为 3000 个时间周期。网络初始化时, 采用随机的方式给每个结点分配 8 个邻居, 即初始出度和入度都是 8, 结点的最大入度设为 20。通过先前的一些实验结果, 我们把调

整的阈值  $\sigma$  设为 2.3。

评估标准包括文件的平均命中率和平均搜索长度、网络的平均通信开销和平均 RSB。此外, 我们还考察了 TARSB 对结构的调整能力, 以及 4 种调整操作的比例。

### 5.2 实验结果

#### 5.2.1 网络演化

TARSB 算法的中心思想是通过调整网络结构, 具体说是通过调节结点的出度和入度, 来提高网络的搜索性能。因此我们首先考察了 TARSB 算法改变网络结构的能力。对于同一个随机生成的网络拓扑, 分别在第 0, 100, 200 和 500 个周期执行 TARSB 算法, 考察算法起始时间对演化效果的影响。图 1 为到模拟结束时 (第 3000 个周期) 结点的入度分布情况。从图上可以看出算法起始时间对演化效果并无明显影响, 其分布可大致分为 3 个阶段。

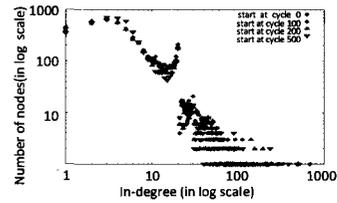


图 1 入度分布情况

第一阶段, 入度为 0~5, 结点的数量呈上升趋势, 且这部分结点的数量约占结点总数的 50%。这一阶段与文献[16, 18] 的结论并不相同, 其原因主要是目前还没有合适的方法来体现不同类型的结点在动态性上的差异。在模拟实验时, 网络结点中的大部分是“纯消费”结点, 能对搜索起到帮助的“优秀”结点只占很少部分, 再加上高动态性的影响, 使得当前在线的“优秀”结点的数量更加稀少, 从而无法为所有的“纯消费”结点提供直接的链接。因此, 为了保持一定的搜索性能, 结点只能通过与某些“纯消费者”建立链接, 间接链接到“优秀”结点上。而在现实网络中, “优秀”结点与“纯消费”结点在动态性方面会有比较大的区别, “纯消费”结点可以与足够多的“优秀”结点直接建立链接。目前针对结点动态规律的差异性还没有公认分布规律, 因此在模拟时只能让“优秀”结点与“纯消费”结点都遵循同样的动态规律。这就是在实验中出现这一反常阶段的一个主要原因。

第二阶段, 入度为 6~20, 结点的数量开始呈下降趋势, 在接近入度 20 的附近减缓下降, 甚至呈上升趋势。这是因为初始的最大入度为 20, 很多结点的入度达到此值后, 由于不满足条件而不能增加最大入度。可以说最大入度是一个区分结点性能的标尺, 只有足够“优秀”的结点才能拥有更大的入度。所以, 在入度 20 的附近聚集了一部分“次优秀”结点, 造成了图中所示的“断档”现象。

第三阶段, 入度为 20 以上, 其结点数量占总数的 5%, 入度之和占总数的 25%。这一阶段的分布规律与文献[16]所描绘的 Gnutella 结点链接分布规律很接近, 也就是通常所认为的 Power-Law 分布。这一部分的结点是通过 TARSB 算法选出来的“优秀”结点, 其优越性可通过比较两个主要的异构特性 (结点拥有的文件数量和结点的接收缓存容量) 来考察, 如表 2 所列。

表2 结点的差异

TARSB 算法起始点	入度	结点数量	平均拥有 的文件数	平均接收缓存 (报文/周期)
周期 0	>20	261	550	183
	≤20	4739	62	149
周期 100	>20	267	554	186
	≤20	4733	64	148
周期 200	>20	309	532	190
	≤20	4691	55	146
周期 500	>20	373	538	192
	≤20	4627	56	140

从结点的入度分布图可以看出, TARSB算法能较好地体现出结点之间的差异性。如果忽略初始最大入度造成的影响, 该图基本符合 Power-Law 分布, 克服了原先无法根据结点特性自主调节连接度的缺点。从下面的性能分析中可以看出, 经过 TARSB算法的调整, 结点确实能够根据实际情况, 通过合理地增加或减少出度来维持自身的搜索性能, 同时降低网络的通信开销。

### 5.2.2 性能分析

TARSB算法的中心思想是利用结点的异构性来提高网络的整体搜索性能, 我们采用平均 RSB 值作为衡量网络整体搜索性能的指标。这一指标受到多种直观性能参数的综合影响, 能够比较全面客观地反映出当前网络的搜索性能。如图 2 所示, 与没进行优化时相比, 执行了 TARSB 算法后, 网络的平均 RSB 值有了显著提高。

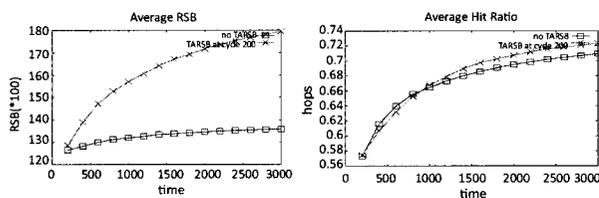


图2 平均RSB比较

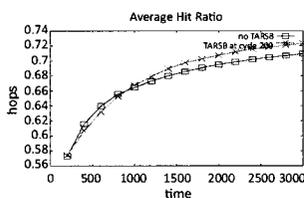


图3 平均命中率

通过分析 4 种调整操作(增加最大入度、删除出边、替换出边以及增加出边)在实验各阶段的执行比例, 可以发现 TARSB 算法是分 3 个步骤来提高网络搜索性能的。

在开始进行调整的 150 个时间周期内, 主要执行的操作是删除出边。在经过一段时间的运行后, 大多数结点会发现原先随机指派的邻居中有许多并不能对搜索起到帮助, 具体表现在它们的 RSB 值与结点自己的 RSB 值相差很大。根据 TARSB 算法, 结点会主动断开这些无用的链接。经过这一阶段的调整, 网络中的无效链路将所剩无几。

接下来的 500 个时间周期内, 占主导地位的操作是增加出边。前一阶段的大量删边会对搜索命中率造成一定程度的影响。为了提高自身的 RSB, 结点需要补充新的邻居。从 TARSB 算法的角度来解释: 通过删除无用结点, 使得剩下的邻居往往具有很高的 RSB, 结点需要通过添加新的邻居来平滑自身 RSB 值与这些高 RSB 值之间的差距。

其后的阶段, 结点基本上可以获得比较稳定的搜索性能, 主要通过替换邻居结点来维持和提高性能。

需要说明的是, 无论处于哪一阶段, 结点都有可能执行 4 种操作中的任一种。以上的 3 个阶段只是对网络演化趋势的大致划分, 并不是每个结点都会经历这 3 个阶段。

在 P2P 文件共享系统中, 命中率和平均搜索长度是衡量

搜索性能的两个直观指标。图 3 和图 4 分别是未执行 TARSB 算法与从周期 200 开始执行 TARSB 算法的两项性能比较。从图 3 可以看出, 当刚开始执行 TARSB 算法时, 网络的平均命中率会有所下降, 经过一段时间的调整后, 逐渐超过未执行优化算法的平均命中率。而平均搜索长度则比未执行优化算法时有了明显的降低。从中可以看出, TARSB 算法对网络的搜索性能确有提高。

网络的通信开销是 P2P 文件共享系统(特别是基于洪泛搜索的系统)所关心的重要指标。从图 5 可以看出, TARSB 算法使得通信开销降低了 12%。

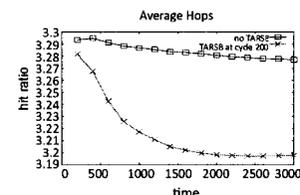


图4 平均搜索长度

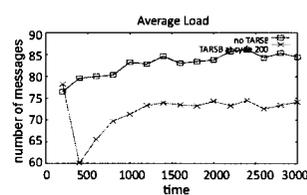


图5 平均通信开销

**结束语** TARSB 算法通过相对搜索介数(RSB)评估结点的搜索性能, 让性能优秀的结点逐渐演化成网络的主干, 而性能差的结点则移动至网络的边缘, 其最终目标是形成与结点异构性相对应的非对称网络结构。本文的贡献在于提供了一种全面衡量结点搜索能力的方法, 克服了以往算法只针对一种异构性的缺陷, 并以此对网络的整体搜索能力实现了进一步的提升, 在保障搜索命中率的前提下, 大大降低了网络的传输开销。同时使得结点具备了进化的智能, 体现出对动态环境良好的适应能力。在今后的研究中, 我们将深入研究各种异构性对 RSB 的影响, 以及主干网络的形成和演化规律。

### 参考文献

- [1] <http://www.Slyck.com>
- [2] <http://rfc-gnutella.sourceforge.net/developer/index.html>
- [3] Stoica I, Morris R, Karger D, et al. Chord: A scalable peer-to-peer lookup protocol for internet applications. IEEE/ACM Transactions on Networking, 2003, 11(1): 17-32
- [4] Ratnasamy S, Francis P, Handley M, et al. A scalable content addressable network // Proceedings of the ACM SIGCOMM, 2001: 161-172
- [5] Lv Q, Ratnasamy S, Shenker S. Can heterogeneity make Gnutella scalable // Proc. of the 1st Int'l Workshop on P2P Systems, Cambridge, 2002
- [6] Chawathe Y, Ratnasamy S, Breslau L, et al. Making Gnutella-like P2P Systems Scalable // Proc. of the ACM SIGCOMM 2003, Karlsruhe, 2003
- [7] Hu Jingfeng, Li Ming, Zheng Weimin, et al. SmartBoa: Constructing p2p Overlay Network in the Heterogeneous Internet Using Irregular Routing Tables, Peer-to-Peer Systems III // Third International Workshop, IPTPS 2004, La Jolla, CA, USA, February 2004: 278
- [8] Waldvogel M, Rinaldi R. Efficient topology-aware overlay network. ACM SIGCOMM Computer Communication Review, 2003, 33(1): 101-106
- [9] Condie T, Kamvar S D, Garcia-Molina H. Adaptive Peer-to-Peer Topologies // Fourth International Conference on Peer-to-Peer Computing (P2P'04), 2004: 53-62

(下转第 89 页)

出速率最接近源数据流的平均速率。

## 4.2 延迟抖动比较

在实验中分别记录了 DLB 整形算法和 JCS 算法的平均延迟和最大延迟,最小延迟以及平均延迟;统计信息如表 2 所示。延迟方差反映了整个过程中延迟的变化波动情况,即反映了抖动的大小。从表中可以得知,DLB 算法和 JCS 算法之间的平均延迟虽然相差不是很大,DLB 整形算法为 4.21ms,而 JCS 算法为 3.18ms,但是它们之间的延迟方差相差却很大,DLB 算法的延迟方差达到了 271.4,而 JCS 算法的延迟方差只有 3.62,这充分说明了 JCS 算法的输出流的延迟比较恒定,也就是说抖动值较小。为了能直观地比较两种算法的延迟抖动,这里给出了试验过程中的抖动值变化,这里的抖动值是利用在算法过程中  $jitter = delay - delay\_ave$  的公式计算的。从抖动变化曲线(如图 6 所示)可以看出,DLB 算法的抖动变化比较大,大约在 -3ms 到 3ms 之间,而 JCS 算法的抖动值大约在 -0.8 到 0.8 之间。

表 2 DLB 算法和 JCS 算法的延迟数据比较

类型	平均延迟/ms	最大延迟/ms	最小延迟/ms	延迟方差
DLB 算法	4.21	7.2	0.85	271.4
JCS 算法	3.18	3.6	2.72	3.62

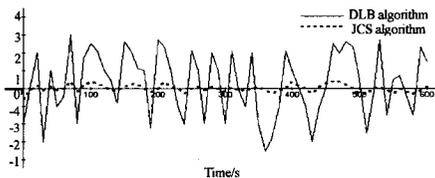


图 6 DLB 算法和 JCS 算法的抖动变化曲线图

从以上试验结果可以看出,JCS 算法不仅能起到平滑突发数据流的作用,而且可以有效降低整形过程中数据缓冲引入的延迟抖动,这样不仅降低了网络的资源利用率,而且对 VBR 流这样的视频应用的服务质量有很大的提高。对于视频流在 Internet 上提高端到端的 QoS 的研究有很大的帮助。

**结束语** 本文提出了一种基于抖动约束的整形算法,与传统的 DLB 算法不同,JCS 算法根据网络节点输入缓冲区充

满程度的变化来动态调整令牌的输入速率,从而平滑输出流的速率和控制输出流的延迟抖动。经实验验证,JCS 算法可以有效地降低 VBR 视频流在整形过程中的延迟抖动,提高 VBR 视频流的服务质量,可以有效地提高网络节点对于 VBR 这样的视频流的支持。

## 参考文献

- [1] Eichler G, Riccio F. AQUILA Network Services [C]// Proceedings of COMCON8. Crete, GREECE, June 2001
- [2] Fernando, Moreiral, Ruela J. A method for traffic scheduling based on token bucket QoS parameters [C]// The 3rd Conf. on Telecommunication. Porto, Portugal, 2001
- [3] Elwalid A, Mitra D. Traffic shaping at a network node: theory, optimum design, admission control [C]// Proceedings of IEEE INFOCOM. 1997; 445-455
- [4] Graf M. VBR video over ATM: Reducing network resource requirements through end system traffic shaping [C]// Proceedings of IEEE INFOCOM. 1997; 48-57
- [5] Bechler M, Riter H, Schafer G. Shaping in end systems attached to QoS-supporting networks [C]// IEEE Symposium on Computers and Communications Proceedings. 2001; 296-301
- [6] Verma D C, Zhang H. Delay Jitter Control for Real-Time Communication in a Packet Switching Network [C]// Proceedings of TriComm'91
- [7] Alam M, Atiquzzaman M, Karim M. Shaping for MPEG video transmission over the next generation Internet [J]. Computer Communications, 2000, 23(14): 1336-1348
- [8] Alam M, Atiquzzaman M, Karim M. Efficient MPEG video traffic shaping for the next generation Internet [C]// IEEE Global Telecommunications Conference v 1 (A) 1999. IEEE, Piscataway, NJ, USA, 1999; 364-368
- [9] Ferrari D. Client Requirements for Real-Time Communication [J]. IEEE Communications Magazine, 1990, 28(11): 65-72
- [10] Naser H, Leon-Gracia A. Performance evaluation of MPEG2 video using guaranteed service over IP-ATM networks [C]// Multimedia Computing and Systems Conference. 1998

(上接第 74 页)

- [10] Wan H, Ishikawa N, Autonomous J H. Topology Optimization for Unstructured Peer-to-Peer Networks // 11th International Conference on Parallel and Distributed Systems (ICPADS'05). July 2005; 488-494
- [11] Darlagiannis V, Mauthe A, Steinmetz R. Overlay Design Mechanisms for Heterogeneous, Large-Scale, Dynamic P2P Systems. Journal of Network and Systems Management, 2004, 12(3): 371-395
- [12] Crespo A, Garcia - Molina H. Semantic Overlay Networks for P2P Systems. Technical Report. Stanford University, January 2003
- [13] Goh K-I, Oh E, Kahng B, et al. Betweenness centrality correlation in social networks. Phys. Rev. E, 2003(67): 017101
- [14] Saroiu S, Gummadi P, Gribble S. A Measurement Study of P2P File Sharing Systems. Technical Report UW-CSE-01-06-02. July 2001
- [15] Chu J, Labonte K, Levine B N. Availability and Locality Meas-

- urements of Peer-to-Peer File Systems // ITCOM: Scalability and Traffic Control in IP Networks II Conferences. July 2002
- [16] Ripeanu M, Foster I. Mapping the Gnutella Network: Macroscopic Properties of Large-Scale Peer-to-Peer Systems. Lecture Notes in Computer Science, Springer-Verlag GmbH, 2002, 2429; 85-93
- [17] Bustemante F E, Qiao Y. Friendships that Last: Peer Lifespan and its Role in P2P Protocols // Intl. Workshop on Web Caching and Distribution. September 2003
- [18] Stutzbach D, Rejaie R, Sen S. Characterizing Unstructured Overlay Topologies in Modern P2P File-Sharing Systems // Proceedings of the ACM SIGCOMM Internet Measurement Conference. New Orleans, October 2005. Technical Report CIS-TR-05-01. University of Oregon, June 2005
- [19] <http://peersim.sourceforge.net/>
- [20] Newman M E J. A measure of betweenness centrality based on random walks. <http://aps.arxiv.org/abs/cond-mat/0309045/>