

分布式网页排序算法及其传输模式分析

余 锦 史树明

(清华大学计算机科学与技术系,北京 100084)

E-mail yujin@mails.tsinghua.edu.cn

摘 要 网页规模的飞速发展要求分布式网页排序技术的出现。在分析了分布式环境下网页划分的策略后,基于集中式 PageRank,给出了适于开放系统的 GroupPageRank 算法,接着提出了两个分布式网页排序算法并给出了一些相关理论结果。同时还对传输模式进行了探讨,提出了具有良好扩展性的间接传输模式。最后在真实数据集上进行了实验,验证了实验的结果。

关键词 分布式 网页排序 PageRank 间接传输

文章编号 1002-8331-(2004)29-0182-06 文献标识码 A 中图分类号 TP393

Algorithms for Distributed Page Ranking and Analysis on Transmission Mode

Yu Jin Shi Shuming

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract: Distributed page ranking is needed for the remarkable growth of the web. After the analysis of the strategy for web page partitioning under distributed environment, an algorithm called GroupPageRank is proposed based on the centralized PageRank algorithm. It is also suitable for open system. Then two distributed PageRank algorithms and also some related theoretical results are given. At the same time, we perform discussion on transmission modes. A scalable transmission mode, Indirect Transmission Mode is presented. Finally, we verify some of the discussions by experiments based on real datasets.

Keywords: distributed page ranking, PageRank, indirect transmission

1 引言

网页排序,顾名思义,指对网页进行排序,排定各个网页重要性的方法。它是提高搜索引擎搜索质量的关键技术之一。目前的网页排序算法大都利用网页的链接结构进行计算。如 HITS^[1]算法,它利用网页链接关系为每个网页计算分数;又如 Google 所使用的 PageRank^[2]算法,从网页链接关系推导出了一个矩阵方程,通过迭代计算得到矩阵的特征向量,进而得到每个网页的 rank 值。

现有集中式的搜索引擎由于在扩展性方面的缺陷,难以对网络规模迅猛增长做出即时的反应,越来越需要分布式搜索引擎的出现。事实上,即使是拥有世界最大网页索引库的 Google,在 2003 年也就索引了整个 web 网页总数的 5.4%^[18]。而分布式搜索引擎不论对用户的数目还是网页的规模都具有良好的可扩展性。相应地,为了提高搜索质量,提高查询结果的准确度,网页排序不可缺少。与此同时为了获得系统整体的良好可扩展性和高可用性,分布式网页排序是必然的选择。

进行分布式网页排序最直观想法可能就是将 HITS 或者 PageRank 算法简单地扩展到分布式环境。事实上这并不可行。HITS 和 PageRank 算法,他们都是迭代型的算法。每一步的迭代都需要前一步的计算结果,因此计算的各结点间必须进行同步。而在一个高度分散的分布式环境里,这样的同步通讯很难

做到。况且除了同步开销外,还有网页划分,额外的通讯开销等必须考虑。

最近,结构化对等网络赢得了学术界的青睐。基于它们,出现了很多的大规模分布式系统^[6,13-15],它们在可扩展性,自组织性,容错性等比较出色。该文也试图在结构化对等网络上进行有效的分布式网页排序。在讨论了网页划分的策略后,基于 PageRank^[2]提出了一些分布式的网页排序算法,并且给出了一些相关的有趣的性质和结果。由于通讯开销在分布式网页排序中有着比 CPU 和内存容量更为重要的作用,文章还分析了传输模式并提出了间接传输模式。

余下部分是这样组织的:第二部分回顾和介绍了集中式 PageRank 算法,第三部分讨论了分布网页排序中的种种问题并提出了解决方案。第四部分,对传输模式进行了分析,提出了间接传输模式。第五部分在一个实际的数据集上进行了实验,并对实验结果进行了分析。

2 集中式 PageRank 算法

PageRank^[2]算法根据网页间的超级链接关系来量化网页的重要性。每个网页都有它自己的 rank 值(重要值),如果网页 u 有一个超级链接指向了网页 v ,那么网页 v 将部分得到网页 u 的 rank 值,而网页 v 的 rank 值即是所有像 u 这样的网页所贡

基金项目:国家自然科学基金项目(编号:60173007);国家 863 高技术研究发展计划项目(编号:2001AA111080,2002AA104580)资助

作者简介:余锦(1980-),男,在读硕士研究生,主要研究方向为对等网络,分布式搜索和计算网络。史树明(1976-),男,在读博士研究生,主要研究方向为对等网络,分布式搜索和计算网络。

献的 rank 值的和。直观上,如果一个网页有很高的 rank 值,要么是有很多超级链接指向了它,要么是指向它的网页中有一些具有很高的 rank 值。

假定 n 为网页总数, $R(u)$ 表示网页 u 的 rank 值, $d(u)$ 表示网页 u 的出度 (即网页 u 指向别的网页的链接数), B_v 为所有存在超级链接指向网页 v 的网页构成的集合。那么网页 v 的 rank 值计算如下:

$$R(v) = c \sum_{u \in B_v} \frac{R(u)}{d(u)} + (1-c) E(v) \quad (1)$$

上式右边第二项是为了防止 rank sink^[2]现象。

方阵 A 的行和列对应于各个网页 u , 若存在超级链接指向 v , 令 $A_{u,v} = 1/d(u)$, 否则 $A_{u,v} = 0$ 。那么公式 (1) 可以改写为如下矩阵形式:

$$R = cAR + (1-c)E \quad (2)$$

那么 PageRank 可以按如下方式计算。

$R_0 = S$

loop

$$R_{i+1} = AR_i$$

$$d = \|R_i\|_1 - \|R_{i+1}\|_1$$

$$R_{i+1} = R_{i+1} + dE$$

$$\delta = \|R_{i+1} - R_i\|_1$$

算法 1: PageRank 算法

3 分布式网页排序

不能简单套用算法 1 来进行分布式的网页排序,这主要有两方面的原因 (1) 由于每个结点仅仅拥有一部分的网页,因此如 $\|R_i\|_1$ 这样的全局性的操作将十分耗时 (2) 每一次迭代都需要前一次迭代的结果,因此各个结点间必须进行同步;

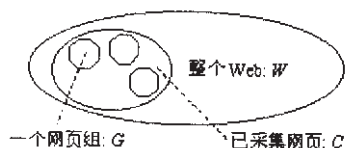


图1 几个网页集合关系图

分布式环境下,几个网页集合的关系如图1所示。其中大椭圆代表整个 Web 中的网页集合;小椭圆为所有已采集的网页组成的集合;小八边形则表示各个结点所负责的网页的集合,称为一个网页组。

很直观地,在进行分布式网页排序前,先得决定如何将 C (已采集的网页) 划分成一个个的 G (网页组)。

3.1 网页划分

主要有三种网页划分策略 (1) 随机划分 (2) 根据网页 URL 的哈希码 (3) 根据网页所在网站的哈希码。由于网页内容会出现变化,采集器需要重新采集一些网页,对于随机划分来说,这将导致不同时期对网页划分产生不同的结果,将大大影响系统性能。而不同的结点上的网页往往存在链接,因此进行网页排序时,需要在不同的结点间传输网页的 rank 值,不同的划分方案将导致不同的网络通信开销。文[16]结果表明网页中的 90% 的链接都指向本身的站点,因此 (3) 与 (2) 相比将产生更低的网络通信开销。

3.2 GroupPageRank 算法

对于单个结点,也不能简单套用算法 1 来对它上面的网页组进行排序。因为算法 1 将所处理的网页集作为一个封闭系统,而这里,每个结点上的网页集却是开放的系统,它们与其他结点上的网页往往存在链接。因此提出了 GroupPageRank 算法来解决网页组这样开放系统下的网页排序。

对网页组来说,主要存在三种链接:内部链接、传入链接和传出链接。内部链接指结点组内部网页间的链接;传入链接则指其它结点组中网页指向本结点组中网页的链接;传出链接为本结点组中网页指向其它结点组中网页的链接。另外,为避免 Rank Sink^[2]现象,与文[8]一样,加上了虚拟链接使所有网页构成一个全连接图。图2示意了各种链接,由于虚拟链接数量太多,为清晰起见,图中只画出了结点组内部网页间的虚拟链接,而结点组间网页的虚拟链接被略去。接下来分别考虑这4种链接。

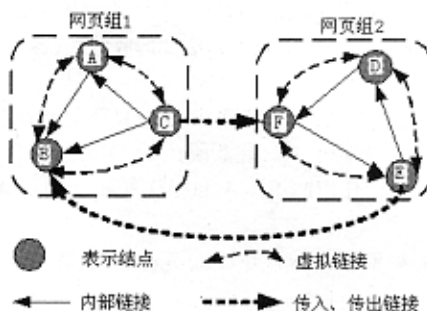


图2 网页组及各种链接图

对于网页组 G 中任意网页 u 和 v , 用 $R(u)$ 和 $d(u)$ 分别表示 u 的 rank 值和出度, B_v 表示由 G 中所有存在链接指向 v 的网页构成的集合。假设 G 中的每个网页 u , 它 rank 值中 $\alpha R(u)$ 的部分通过实际的链接传递 (即通过内部链接和传出链接), $\beta R(u)$ 部分通过虚拟链接传递 ($\alpha + \beta = 1$)。

对网页 v 来说, 它的 rank 值可以来自内部链接, 虚拟链接或者传入链接, 分别用 $I(v)$, $V(v)$ 和 $X(v)$ 来表示。用 2 中同样的方法容易知道:

$$I(v) = \alpha \sum_{u \in B_v} \frac{R(u)}{d(u)} \quad (3)$$

对于虚拟链接, 假设所有的虚拟链接传递的 rank 值都相同, 即网页通过虚拟链接给每个网页 (包括它本身) 都传递了相同的 rank。因此:

$$V(v) = \sum_{u \in W} \frac{\beta R(u)}{w} = \frac{\beta}{w} \sum_{u \in W} R(u) = \beta E(v) \quad (4)$$

这里 W 表示整个 web, $w = |W|$ 。 $E(v)$ 是整个 web 中网页的平均 rank 值, 这与 (1) 式中的意义相同。为简单起见, 假设对网页组中的所有网页 $E(v) = 1$ 。个性化网页排序^[5,9]即是通过 $E(v)$ 进行定制实现的。

因此, 网页组中的所有网页的 rank 可以计算如下:

$$R(v) = I(v) + V(v) + X(v) = \alpha \sum_{u \in B_v} \frac{R(u)}{d(u)} + \beta E(v) + X(v) \quad (5)$$

或者:

$$R = AR + (\beta E + X) \quad (6)$$

这里方阵 A 的行和列对应于各个网页, 如果 u 存在内部链接指向 v , 令 $A_{u,v} = \alpha/d(u)$, 否则 $A_{u,v} = 0$ 。令 $Y(v)$ 表示通过传出链接传递的 rank 值, 则有:

$Y=BR$

这里方阵 B 满足：

$$B_{u,v} = \begin{cases} \frac{\beta}{d(u)} & d(u) > 0 \\ 0 & d(u) = 0 \end{cases}$$

集中式 PageRank 算法与所提出的这个改进算法(Group-PageRank)主要区别在于：前者基于封闭系统，而后者用于开放系统，允许 rank 值流入、流出系统。

function $R^* = \text{GroupPageRank}(R_0, X)$

```
repeat
     $R_{i+1} = AR_i + \beta E + X$ 
     $\delta = \|R_{i+1} - R_i\|_1$ 
until  $\delta > \epsilon$ 
return  $R_i$ 
```

算法 2 GroupPageRank 算法

以下的几个定理则保证了算法 2 的收敛性(参见文[7]中的证明)。

定理 1 以 x_0 为初始值的迭代 $x = Ax + f$ 收敛当且仅当 $\rho(A) < 1$ 。这里 $\rho(A)$ 指矩阵 A 的谱半径。

定理 2 对于任何的矩阵 A 和矩阵范数 $\|\cdot\|$ ，都有 $\rho(A) \leq \|A\|$ 。

定理 3 如果 $\|A\| < 1$ ，并且 $x_m = Ax_{m-1} + f$ 收敛于 x^* ，那么

$$\|x^* - x_m\| \leq \frac{\|A\|}{1 - \|A\|} \|x_m - x_{m-1}\|$$

对于算法 2，根据定理 2 有 $\rho(A) \leq \|A\|_\infty \leq \alpha$ ，再根据定理 1，那么迭代收敛。而定理 3 表明可以用 $\|x_m - x_{m-1}\|$ 作为迭代的终止条件。

3.3 分布式 PageRank 算法

在 P2P 环境下进行分布式网页排序，由于各个结点上的网页组间存在着相互的链接关系，因此各结点必须周期性地交换网页的 rank 值，直至整个系统中的所有网页 rank 值都收敛。

DPR1 和 DPR2 都是用来进行分布式网页排序的算法，分别如算法 3 和算法 4 中所示。两者都含有一个主循环，每次循环中，由于传入链接已经传入了别的网页组的新的 rank 值，因此先刷新 X 的值，接着通过迭代计算出向量 R ，最后计算新的 Y 并将它送给别的结点。

注意到结点间是异步运行分布式 PageRank 算法。也就是说各个结点可以在不同的时间启动网页排序程序，以不同的速度运行，睡眠或挂起一段时间，甚至可以关机。实际上，可以在任何操作之后加入延迟。

```
function DPR1(X)
     $R_0 = S$ 
     $X = 0$ 
    loop
         $X_{i+1} = \text{Refresh } X$ 
         $R_{i+1} = \text{GroupPageRank}(R_i, X_{i+1})$ 
        Compute  $Y_{i+1}$  and send it to other nodes
        Wait for some time
    while true
}
```

算法 3 分布式 PageRank 算法 DPR1

DPR1 和 DPR2 两算法的区别在于输入变量 X 的刷新频率和输出变量 Y 的更新频率上。DPR1 中，是在利用算法 2 进行若干步迭代后得到收敛的 R 向量后，才对 Y 进行更新并传

(7)

给其他网页组。而 DPR2 中，每个结点总是使用最新的 X 向量，并且对 R 迭代一次后就马上更新 Y 并将其传给其他网页组。

```
function DPR2(X)
     $R_0 = S$ 
     $X = 0$ 
    loop
         $X_{i+1} = \text{Refresh } X$ 
         $R_{i+1} = AR_i + \beta E + X_{i+1}$ 
        Compute  $Y_{i+1}$  and send it to other nodes
        Wait for some time
    while true
```

算法 4 分布式 PageRank 算法 DPR2

3.4 收敛性分析

对于 DPR1 有(由于篇幅关系这里证明略)

定理 3 对于静态的链接图，每个结点运行 DPR1 算法得到的迭代序列 $\{R_1, R_2, \dots\}$ 单调递增。

定理 4 对于静态的链接图，每个结点运行 DPR1 算法得到的迭代序列 $\{R_1, R_2, \dots\}$ 有上界。

单调递增的有界序列收敛，由定理 3 和 4，可以马上知道算法 DPR1 收敛。

当 $R_0 = 0$ 时，定理 3 和 4 同样适用于算法 DPR2。此时将每个网页看成一个网页组，此时 DPR2 就相当于 DPR1 了。

现在可能会有疑问：“DPR1、DPR2 他们收敛得到的值和集中式 PageRank 算法得到的值相同吗？”，实验说明答案是正确的。

4 传输模式

进行分布式网页排序时，由于结点所负责的网页组间往往存在相互链接，因此，结点间必然需要进行点到点的通讯。这一部分就是要在传输模式角度，研究如何减少网页排序中的通讯开销。而如压缩等这样可以为每个模式采用的通用方法，这里不做讨论，将在后续工作中完成。

根据约定：共有 N 个结点，负责计算 W 个网页的 rank 值。另 l 表示每个链接的平均大小， r 表示查询目标结点消息的大小，每个查询平均需要 h 个 hops。使用站点哈希对网页进行划分，而文[16]结果表明每个网页大概只有一个链接指向别的站点，因此 W 个网页中约有 W 个站间链接，而在同一个结点上的站间链接，则不需要在网络上传输。因此若哈希均匀，约有 $(N-1) * W / N$ 个站间链接需要在网络上传输。

4.1 直接传输模式

图 3 显示了一些假想的结点(小圈表示)和他们间的相互通讯(带箭头直线表示)情况。称这种传输模式为直接传输模式。每个结点直接把交互的数据传给目标结点，当站间链接比较多时，直接传输模式通讯图即是一个全连接图，如图 3 所示。

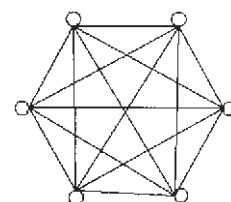


图 3 直接传输模式

当结点数很多(比如大于 1000)，一个结点不大可能知道

其它所有结点的情况。实际上,在 P2P 网络^[6,13-15]中,每个结点通常只有其他几十个结点的信息(称这些结点为该结点的邻居结点)。若源结点 S 想给目标结点 D 发送消息,它必须先知道 D 的地址和端口。在结构化的 P2P 网络中,这是通过查询操作完成的。

因此,若 W 足够大,每个迭代步中有如下关系式。

消息数：
$$S_{di} = (h+1)N^2 \tag{8}$$

其中 h 即为约定中的每次查询所需的平均 hop 数。

数据传输量为：

$$D_{di} = \frac{N-1}{N}lW + hrN^2 \tag{9}$$

基于现有的结构化 P2P 网络 h 为 $O(\log N)$ (CAN^[13] 较为特殊为 $O(d)$) ,因此,直接传输模式下消息数为 $O(N^2 \log N)$,对 N 的可扩展性较差。

对于每个消息,它在发送端和接收端都必须经过网络协议栈。因此它必须两次进出内核空间,引起两次的核心态和用户态的上下文切换,开销较大。

4.2 间接传输模式

为了减少消息的数目,提出了一种可扩展的传输模式:间接传输模式。在间接传输模式中,更新后的网页 rank 值并不直接传送给它的目标结点,而是经过中间的几个结点最后到达目标结点。换句话说,间接传输模式沿着结构化 P2P 网络中路由的通路进行传输消息-初看起来似乎有点违背 P2P 的精神。

图 4 显示了间接传输的核心思想。图中,结点 B 需要将它更新后的网页的 rank 值传给别的结点,它并不是直接将数据传给它的目标结点(在查询到目标结点的地址后),它将数据打包分别传给他的邻居。当结点 A 接收到一些包(从它的邻居 B、C、D 和 E 后),便将这些包解开,根据其中消息的目标结点重新组合形成新的包。这些新的包接着被分别传给 A 的各个邻居结点。包含了网页的 rank 值信息的信息在经历了一系列的打包和解包的过程后将到达它的目标结点。

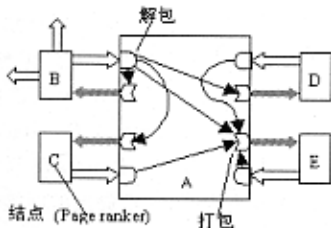


图 4 间接传输模式原理图

图 5 则分别展示了当每个结点拥有两个和 3 个邻居情形下采用间接传输模式的结点通讯图。

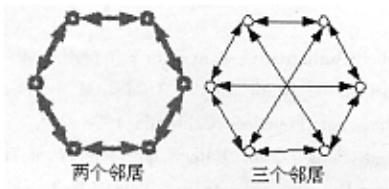


图 5 间接传输模式

同样,若 W 足够大,每个迭代步中有如下关系式。

消息数：
$$S_u = gN \tag{10}$$

这里 g 是一个常数,因为在间接传输模式下,每个结点并非收到消息就立即开始发送消息,而是周期性地发送消息,如果迭代间隔为 T ,而发送周期为 t ,那么显然 $g=T/t$ 。

数据传输量：

$$D_u = \frac{N-1}{N}hW \tag{11}$$

从中看到,使用间接传输模式,数据仅仅在邻居结点间传输。因此,每个迭代步仅需要 $O(N)$ 个消息。

5 实验与分析

数据集：实验中采用的链接图是从 Google 的程序设计比赛数据^[3]中产生的,它由“edu”域下 100 个不同的站点上的一些网页组成的集合。其中大概有 1M 个网页,15M 个链接。虽然有点小,但是它已经是目前能够获得的最大的数据集了。

5.1 算法实验

实验设定：通过 $T_u(u, m)$ 来模拟不同结点上计算的异步性,它表示第 $m-1$ 轮迭代结束到开始第 m 轮迭代开始的时间延迟。实验中,对于给定的 u , $T_u(u, m)$ 服从指数分布,每个网页组的平均延迟时间是 $[T_1, T_2]$ 中的随机数(这里的 T_1 and T_2 是可以调整的参数)。为了模拟潜在的网络故障,假定向量 Y 在向其他的网页组传递失败的概率为 p 。文章进行了多组模拟,对应不同的 T_1, T_2, p and K (K 是网页组的个数或者结点的个数)。

令 R, R^* 分别表示由分布式 PageRank 和集中式 PageRank 计算得到的 rank 值,它们的相对误差则为 $\|R - R^*\| / \|R^*\|$,用它来表征这两个算法间的差别。图 6 表明这种相对误差随着时间的推移而不断变小。

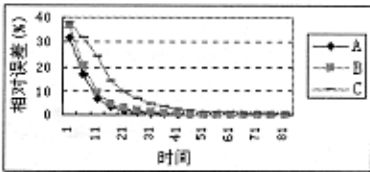


图 6 分布式 PageRank 收敛到集中式 PageRank 得到的 rank 值 ($K=1000, A: p=1, T_1=0, T_2=6; B: p=0.7, T_1=0, T_2=6; C: p=0.7, T_1=0, T_2=15$)

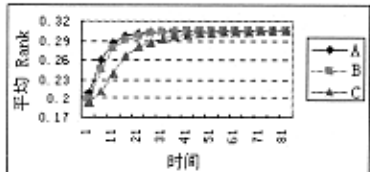


图 7 DPR1 产生的单调增 rank 序列 ($K=100, A: p=1, T_1=0, T_2=6; B: p=0.7, T_1=0, T_2=6; C: p=0.7, T_1=0, T_2=15$)

图 7 展示了由 DPR1 算法产生的 rank 序列的单调性,注意到收敛时,各个网页平均的 rank 才 0.3。这得归功于很大部分链接都是指向了数据集以外的地方。实际上,总共 15M 个链接中只有 7M 个指向了数据集的内部。

图 8 表明了各个算法的收敛性。从中看到 DPR1 收敛得要比 DPR2 快。DPR1 需要的迭代次数甚至比集中式的网页排序算法还要少。从图上还可得出另外一个结论:网页规模对网页排序算法的收敛速度的影响较小。

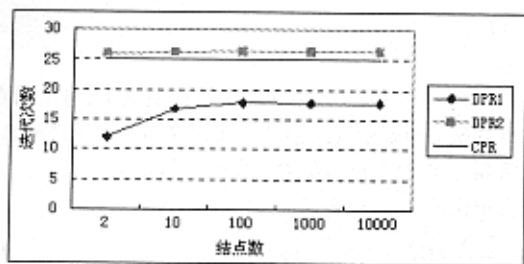


图8 不同网页排序算法间的比较

CPR 指集中式 PageRank。平均误差的阈值为 0.01%
($p=1, T_1=15, T_2=15$)

5.2 传输实验

实验设定：由于数据集中只有 1M 个网页，100 个网站，为了突显间接传输的特性，按照网页哈希来进行网页划分，相当于有 1M 个站点，每个站点一个网页。底层的结构化 P2P 网络则采用 Chord，为了模拟各个结点的异步性，假设一个迭代步中各个结点准备好局部收敛的 rank 值的时间服从正态分布。为了方便，设查询消息和结点间链接大小均为一个单位。

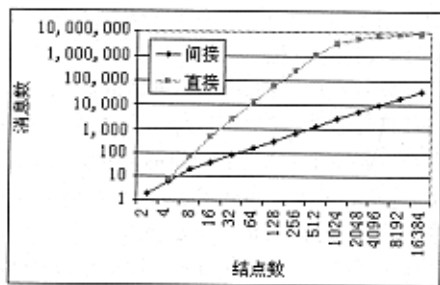


图9 迭代步中消息数对数图

由于直接传输和间接传输两种模式下消息数目绝对数量相差很多，无法用一个普通刻度图明显表示二者的差别和趋势，因此采用对数图。如图 9 所示，间接传输与直接传输在对数图下两条直线概率相差较大，表明二者数量上相差很大。当结点数较大时，每个结点上接收的网页将很少（因为网页总数只有 1M），而每个网页上的链接数有限，因此很多结点并不需要与所有的其他结点通讯，此时，消息数就不是与 N^2 成正比了，这也是直接模式对应曲线后端趋于水平的缘故。而间接传输由于是周期性向它的邻居传输消息，因此它没有这个现象。

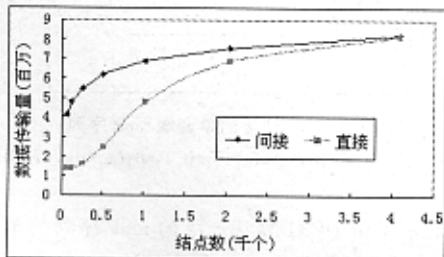


图10 迭代步中数据传输量对比图

初看公式(9)和(11)，似乎未必有 D_{ii} 比 D_{ij} 大，而这是从图 10 看出的明显结论。实际上，从另外一个角度就能清楚地认识这个问题。根据间接传输模式的核心思想，对于每一个结点间的链接，间接传输过程相当于该链接进行了一次路由。对任意

结点 A 和 B，若 A 上有 n 个网页指向了 B，那么间接模式下相当于每个链接都路由了一次，而直接模式下只路由一次，因此间接模式因为 A 指向 B 的链接消耗的带宽约为直接模式下的 n 倍，因此看到了在图 10 中，一开始间接曲线比直接曲线高得多的缘故，当结点数很大时，任意两结点间链接数变得很少，因此间接模式和直接模式趋于重合，结点数越多这个现象越明显。

同时必须注意到，图 10 是在以网页哈希策略划分网页得出的结果，这意味这每个结点都有较多的链接指向了别的结点，因此才有两模式带宽消耗差距较大的情况。实际上，如果采用站点哈希策略，根据文[16]，一个网页大概只有一个链接指向别的站点，由上面的分析，那么意味着间接模式下与直接模式消耗相当的带宽，但它大幅减少了消息数。

因此，当结点数较少时，每个结点上网页较多，一个结点指向另外一个结点链接也较多，采用直接传输模式较为自然，有着较少的带宽消耗和网络消息数；当结点数较大时，则更适宜采用间接传输模式来减少网络消息数。

6 相关工作

文[1-2]利用网页链接关系开创性地提出了 PageRank 算法计算网页的 rank 值。除此之外，人们还进行了很多相关方面的研究。如 PageRank 的快速计算[4-8]，利用 PageRank 完成个性化和敏感标题搜索[5-9]，使用、扩展它们使之服务于其他任务[10-11]等。但到目前为止，还没有见到别人在分布式网页排序方面的相关讨论。

另一个相关的工作可能就是多机系统下的线性方程求解的并行方法研究。可并行求解线性系统方法主要分为两类：直接法和迭代法。但大部分的方法不适合解决该文的问题，因为它们需要进行矩阵的求逆运算，而 web 链接的巨大规模决定了这种代价过于巨大。如需细节请参看文[12]。

7 结论

由于集中搜索引擎缺乏良好的可扩展性，需要有良好的可扩展性的分布式搜索引擎来满足网页规模飞速发展的需求。为进行分布式网页排序，采用站点哈希策略的网页划分方案有利于减少系统的通讯开销，是个理想的选择策略。由集中式 PageRank 算法改进而来的 GroupPageRank 算法则适用于开放系统的网页排序。文章提出的分布式 PageRank 算法 DPR1，DPR2 都能够收敛，并且与集中式 PageRank 算法收敛于同一值。同时对传输模式分析表明，这里提出的间接传输模式可以获得可扩展的通讯性能。

进一步的工作包括：使用更大的数据集，进行更多的实验以发掘分布式网页排序中更多的有趣的现象。探索更多地减少通讯开销和算法收敛时间的方法。（收稿日期：2004 年 5 月）

参考文献

1. Jon M Kleinberg. Authoritative sources in a hyperlinked environment[C]. In Proceedings of the Ninth Annual ACM/IEEE Symposium on Discrete Algorithms, San Francisco, California, 1998-01
2. Lawrence Page, Sergey Brin, Rajeev Motwani et al. The PageRank citation ranking: Bringing order to the Web[R]. Technical report, Stanford University Database Group, 1998
3. <http://www.google.com>
4. T. H. Haveliwala. Efficient computation of PageRank[R]. Technical Re-

port Stanford University ,1999

5.G Jeh J Widom.Scaling personalized web search[R].Technical Report , Stanford University 2002

6.Rowstron A ,P Druschel.Pastry :Scalable ,distributed object location and routing for largescale peer-to-peer systems[C].In *IFIP/ACM Middleware* ,Heidelberg ,Germany 2001

7.Owe Axelsson.Iterative Solution Methods[M].Cambridge University Press ,1994

8.S D Kamvar ,T H Haveliwala ,C D Manning et al.Extrapolation Methods for Accelerating PageRank Computations[R].Technical Report , Stanford University 2002

9.T H Haveliwala.Topic-sensitive PageRank[C].In :*Proceedings of the Eleventh International World Wide Web Conference* 2002

10.D Rafiei ,A O Mendelzon.What is this page known for?Computing web page reputations[C].In :*Proceedings of the Ninth International World Wide Web Conference* 2000

11.S Chakrabarti ,M van den Berg ,B Dom.Focused crawling :A new approach to topic-specific web resource discovery[C].In :*Proceedings of the Eighth International World Wide Web Conference* ,1999

(上接 59 页)

6 CONCLUSIONS

In this paper ,we have novel algorithms for all-to-all personalized communication 2D/3D Meshes with unequal size in each dimension. And we generalized this idea to multidimensional interconnection networks. The scheme has several interesting features. First ,each node needs to send or receive at most one message at a time. Hence ,each router only needs to be equipped with one injection buffer and one consumption buffer. Second ,the communication is contention-free. The largest advantage is that these algorithms are very simple. Hence the result can be used immediately in many current machines.

References

1.L M Ni ,P K McKinley.A Survey of Wormhole Routing Techniques in

12.Vipin Kumar ,Ananth Grama et al.Introduction to Parallel Computing , Design and Analysis of Algorithms[M].The Benjamin/Cummings Publishing Company

13.Ratnasamy S et al.A Scalable Content-Addressable Network[C].In :*ACM SIGCOMM San Diego ,CA ,USA 2001*

14.Stoica I et al.Chord :A scalable peer-to-peer lookup service for Internet applications[C].In :*ACM SIGCOMM San Diego ,CA ,USA 2001*

15.Zhao B Y ,Kubiatowicz J D ,Josep A D.Tapestry :An infrastructure for fault-tolerant wide-area location and routing[R].Tech Rep ,UCB/CSD-01-1141 ,UC Berkeley EECS 2001

16.Junghoo Cho ,Hector Garcia-Molina.Parallel crawlers[C].In :*Proc of the 11th International World-Wide Web Conference* 2002

17.Jinyang Li ,Boon Thau Loo ,Joseph M Hellerstein et al.On the Feasibility of Peer-to-Peer Web Indexing and Search[C].In :*Proceedings of the 2nd International Workshop on Peer-to-Peer Systems(IPTPS'03)* , 2003

18.Google Press Center .Technical Highlights.http://www.google.com/press/highlights.html

Directed Networks[J].Computer ,1993 ,26(2) :62~76

2.P K McKinley et al.Collective Communication in Wormhole-Routed Massively Parallel Computers[J].Computer ,1995 ,28(12) :39~50

3.Y-C Tseng ,S K S Gupta.All-to-All Personalized Communication in a Wormhole-Routed Torus[J].IEEE Trans on Parallel and Distributed Systems ,1996 ,7(5) :498~505

4.Y-C Tseng et al.Bandwidth-Optimal Complete Exchange on Wormhole-Routed 2D/3D Torus Networks :A Diagonal-Propagation Approach[J].IEEE Trans on Parallel and Distributed Systems ,1997 ,8(4) :380~396

5.Y J Suh ,S Yalamanchili.All-to-All Communication with Minimum Start-Up Costs in 2D/3D Tori and Meshes[J].IEEE Trans on Parallel and Distributed Systems ,1998 ,9(5) :442~458

6.V V Dimakopoulos ,N J Dimopoulos.A Theory of Total Exchange in Multidimensional Interconnection Networks[J].IEEE Trans on Parallel and Distributed Systems ,1998 ,9(7) :639~649

7.Y J Suh ,S Yalamanchili.Configurable Algorithm for Complete Exchange in 2D Meshes[J].IEEE Trans on Parallel and Distributed Systems , 2000 ,11(4) :337~356

(上接 83 页)

价随进化代数变化的情况。尽管算法设计进化代数为 30 代 ,但从图中可以看到算法收敛很快。而且在保证延时约束的同时 ,算法使路径的代价开销达到最小。

5 结论

最近有研究者提出将通用遗传算法应用于解决单播或组播路由问题^[6] ,但绝大多数采用 $N \times N$ (N 为网络节点数)的一维二进制编码或者树型结构编码 ,使得算法的编/解码过程复杂 ,效率比较低 ;或者是随着网络规模增大 ,算法容易陷于未成熟收敛。文章对选播通信服务及其延时受限路由问题进行了研究 ,提出了一种性能良好的选播路由算法 ,利用改进的遗传算法求解问题。算法对个体采用节点序列编码 ,每个个体直接对应网络图中一条可行路径 ,交叉和变异算子直接对这样的路径进行优化。通过仿真 ,表明该算法是有效且切实可行的。

参考文献

1.R Hinden ,S Deering.IPv6 Addressing Architecture[S].RFC 2373 , 1998-07

2.D Xuan ,Weijia Jia ,W Zhao.Routing Protocols for Anycast Messages[J].IEEE Transactions on Parallel and Distributed Systems ,2000 ,11(6) :571~588

3.陈国良 ,王熙法.遗传算法及其应用[M].北京 :人民邮电出版社 ,1996

4.K A DeJong.Analysis of the Behavior of a Class of Genetic Adaptive Systems[D].Ph D thesis.Dep Computer and Communication Sciences ,University of Michigan ,1975

5.B Waxman.Routing of Multipoint Connections[J].IEEE J Select Areas Commun ,1988 ,6(9) :1617~1622

6.F Xiang ,L Junzhou ,W Jieyi et al.QoS routing based on genetic algorithm[J].Computer Communications ,1999 ,22(9) :1394~1399