

一种基于延时及带宽受限的启发式组播路由算法

石 坚, 董天临, 邹 玲, 贺聿志

(武汉华中科技大学电子与信息工程系, 湖北武汉 430074)

摘 要: 在分析了网络中基于延时和带宽受限的组播路由优化问题的基础上, 本文提出了一种新的启发式算法, 并进行了实验和分析. 结果表明文中构造的路由方案成功地解决了当网络中存在多组组播业务时的 QoS 路由选择问题. 此方案不仅保证了带宽、端到端延时, 优化了路由树的代价, 而且有效地控制了算法的复杂性并可适用于大规模的网络中.

关键词: 组播路由; 带宽; 延时; 代价; 路由树

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2001) 08-1113-04

A Heuristic Algorithm for Delay-Bandwidth-Constrained Multicast Routing

SHI Jian, DONG Tianlin, ZOU Ling, HE Yuzhi

(Dept. of Electronics & Information Eng., Huazhong Univ. of Sci. & Tech., Wuhan, Hubei 430074, China)

Abstract: On the basis of describing the optimization of multicast routing with delay and bandwidth constrained, the paper presents a new heuristic algorithm. Then some experiments have been given to analyze the performances. The result shows that the routing mechanism in this paper successfully solves the QoS routing problems when many-to-many cast session exists. The scheme not only optimizes the routing tree with guarantee of the bandwidth and end to end delay, but also effectively controls the algorithm complexity which is suitable for large size of network.

Key words: multicast; bandwidth; delay; cost; routing tree

1 引言

当今, QoS 路由技术已成为网络支持 QoS 保证的关键技术之一. 国内外许多学者都在从事组播路由算法的研究并取得了成果, 但大多数学者只综合考虑了实时应用对延时、延时抖动方面的要求^[1~3], 而考虑带宽因素的较少. 文[4, 5] 讨论了网络中单播 (unicast) 时如何通过预约带宽使信息在限定时间内传递到信宿的问题, 但他们的方案尚不适用于组播情况. 文[6~8] 研究了基于延时和带宽限制组播路由问题, 它们的基本思路类似, 即先把网络中不满足带宽要求的链路删除, 然后求解基于延时限制最小代价的组播路由树 (以下简称组播树). 它们都存在算法时间复杂度较高的问题. 文[6] 的算法虽每次都能获得可行解, 但文[6] 假设所有组播业务采用相同的延时和带宽限制条件. 文[7, 8] 虽然认为不同组播业务所采用的带宽或延时限制可能不同, 但文[7] 的算法并非每次都可获得可行解; 文[8] 采用的遗传算法是一种随机化搜索技术^[1, 9], 而文中并没有给出有关算法成功率 (获得可行解的概率) 的实验分析与结论.

本文给出了基于延时和带宽限制的代价优化组播路由模型及问题的表述, 并提出了一种新的启发式组播路由算法. 新算法不仅具有极高的成功率, 而且考虑了网络中多个组播业

务对带宽和延时的不同要求. 由于算法求解的是基于延时及带宽限制的最小代价组播树 (Delay-Bandwidth-Constrained Minimum Cost Multicast Routing Tree), 本文称之为 DBCT 算法. 通过分析与实验, 比较已有算法与 DBCT 算法的性能, 结果表明本算法不仅成功地解决了当网络中存在多个组播业务时的 QoS 路由选择问题, 而且有效地控制了算法的时间复杂度.

2 网络模型及问题的表述

通信网络可表示为无向赋权图 $G = (V, E)$. 其中, V 为网络中所有交换节点的集合; E 为任意两相邻节点 x, y 间通信链路 (x, y) 的集合; 对于 $\forall (x, y) \in E$, 均有三个正实数加权重值 $(D(x, y), C(x, y), B(x, y))$, 分别表示 (x, y) 的延时、代价和可用带宽. 对于 $\forall a, b \in V$, 本文设 a, b 间路径 $P(a, b)$ 的延时 $\text{Delay}(a, b) = \sum_{(x, y) \in P(a, b)} D(x, y)$; $P(a, b)$ 的代价 $\text{Cost}(a, b) = \sum_{(x, y) \in P(a, b)} C(x, y)$; $P(a, b)$ 的可用带宽 $\text{Width}(a, b)$ 为 $P(a, b)$ 上的瓶颈带宽, 即:

$$\text{Width}(a, b) = \min_{(x, y) \in P(a, b)} (B(x, y)) \quad (1)$$

在多媒体实时业务的 QoS 传输中, 基于延时及带宽限制的组播路由优化问题可表述为: 给定图 $G = (V, E)$ 、信源 s 、信宿的集合 $M \subseteq V - \{s\}$, 寻找从信源 s 到所有信宿 $v (v \in M)$

的组播树 $T = (V_T, E_T)$, ($T \subseteq G$), 并满足条件:

$$\text{Cost}(T) = \min \left(\sum_{(x,y) \in E_T} C(x,y) \right) \quad (2)$$

$$\sum_{(x,y) \in P_T(s,v)} D(x,y) \leq D_{\max}, \forall v \in M \quad (3)$$

$$\text{Width}(P_T(s,v)) \geq W_{\min}, \forall v \in M \quad (4)$$

其中公式(2)描述了路由树的代价应达到最小; 式(3)、(4)描述了实时业务的 QoS 要求, D_{\max} 为实时业务所要求延时的上限值, W_{\min} 为应用所需求的最小带宽. $P_T(s,v)$ 为图 G 中从信源 s 经组播树到信宿 v 的路径.

3 DBCT 算法

定义 1 对于 $\forall (x,y) \in E$, 定义 (x,y) 的带宽指示函数 $MB(x,y)$ 为:

$$MB(x,y) = \begin{cases} 1, & B(x,y) \geq W_{\min} \\ \infty, & B(x,y) < W_{\min} \end{cases} \quad (5)$$

定义 2 对于 $\forall (x,y) \in E$, 给定远大于 1 的正实数 δ , 定义 (x,y) 的长度 $\text{Len}(x,y)$ 为:

$$\text{Len}(x,y) = \begin{cases} C(x,y)/(D_{\max} \cdot \delta - D(x,y)), & D_{\max} + \varepsilon - D(x,y) > 0 \\ \infty, & D_{\max} + \varepsilon - D(x,y) \leq 0 \end{cases} \quad (6)$$

定义 3 对于 $\forall a \in V$, 从信源 s 出发经中间节点 b 到节点 a 路径上的距离 $\text{Dis}(a)$ 定义为:

$$\text{Dis}(a) = MB(b,a) \cdot \text{Len}(b,a) + \text{Dis}(b) \quad (7)$$

本文提出的 DBCT 算法不仅通过定义带宽指示函数和判断是否 $\text{Dis}(a) < \infty$ 来保证带宽及延时约束, 而且使用 $\text{Dis}(a)$ 的值来衡量从 s 到节点 a 的最优代价路径. 它采用树枝的逐渐延伸方式构造组播树: 在 G 中依次选取从 s 到 v 距离最小的路径加入 T . 算法的伪代码描述如下:

```
DBCT( $G, s, M, D_{\max}, W_{\min}$ ) //输入图  $G = (V, E)$ 、信源  $s$ 、信宿集  $M$  及  $D_{\max}$  和  $W_{\min}$ .
1 for 每个节点  $a \in V$  do //指针  $\text{Poi}[a]$  用于指向  $P(s, a)$  中节点  $a$  的父节点.
2    $\text{Dis}(a) = \infty$ ;  $\text{Poi}[a] = \text{NIL}$ ;  $\text{Delay}(s, a) = \infty$ ;
3 for 每条链路  $(x, y) \in E$  do //按定义 1、2 计算链路  $(x, y)$  的长度和  $MB(x, y)$  的值.
4   if  $D_{\max} \cdot \delta - D(x, y) \geq 0$  then  $\text{Len}(x, y) = C(x, y)/(D_{\max} + \varepsilon - D(x, y))$  else  $\text{Len}(x, y) = \infty$ ;
5   if  $B(x, y) \geq W_{\min}$  then  $MB(x, y) = 1$  else  $MB(x, y) = \infty$ ;
6  $Q = V$ ;  $V_T = \{s\}$ ;  $E_T = \emptyset$ ;  $\text{Dis}(s) = 0$ ;  $\text{Delay}(s, s) = 0$ ;  $\delta = 1.5$ ;
7 while  $Q \neq \emptyset$  and  $M \neq \emptyset$  do
8   从  $Q$  中弹出具有最小  $\text{Dis}(a)$  的节点  $a$ ; //选择距信源  $s$  有最小距离的节点  $a$ .
9   if  $\text{Dis}(a) = \infty$  then 算法停止;
10  if  $a \in M$  then //若  $a$  是信宿, 则将位于  $a$  到组播树  $T$  间最小;
11     $b = a$ ; //距离路径上的所有节点及链路加入  $T$ ;
12    while ( $\text{Poi}[b] \neq \text{NIL}$ ) and  $b \notin V_T$  do //并修改此
```

路径上每条链路的可用带宽.

```
13    $V_T = V_T + \{b\}$ ;  $E_T = E_T + \{(\text{Poi}[b], b)\}$ ;
14    $B(\text{Poi}[b], b) = B(\text{Poi}[b], b) - W_{\min}$ ;  $b = \text{Poi}[b]$ ;
15    $R = Q$ ; //计算  $s$  到图  $G$  中所有没被访问节点  $x$  的距离,
16   while  $R \neq \emptyset$  do //并通过指针  $\text{Poi}[x]$  记录下路径  $P(s, x)$ .
17     从  $R$  中弹出首节点  $x$ ;
18      $\text{Delay}(s, x) = \text{Delay}(s, a) + D(a, x)$ ; *
19     if  $\text{Dis}(x) > MB(a, x) \cdot \text{Len}(a, x) + \text{Dis}(a)$  and  $\text{Delay}(s, x) \leq D_{\max}$  then
20        $\text{Dis}(x) = MB(a, x) \cdot \text{Len}(a, x) + \text{Dis}(a)$ ;  $\text{Poi}[x] = a$ ;
21 //按定义 3 计算  $\text{Dis}(x)$  的最小值.
```

4 性能分析

事实上, 可将 DBCT 算法分为三个部分, 算法的第 1~ 6 步做初始化; 第 7~ 14 步用于求组播树 $T = (V_T, E_T)$, 使得在优化了组播树总代价的基础上, $\forall v \in V_T$, 有 $P(s, v)$ 符合延时及带宽约束; 第 15~ 19 步用于计算 s 到图 G 中所有没被访问节点 x 的距离.

性能 1: DBCT 算法的时间复杂度为 $O(|V|^2)$

在 DBCT 算法的第 7~ 8 步, 算法需依次在 $|V| - i$ ($i = 0, 1, 2, \dots, (|V| - 1)$) 个未访问的节点中选择有最小距离的节点, 共需进行 $|V|^2/2$ 次操作; 在第 10~ 14 步, 最多需进行 $|V|^2$ 次操作即可将 M 个信宿中的一个加入组播树 T , 所以算法最多需进行 $|M| \cdot |V|$ 次操作以将所有信宿加入组播树 T ; 在第 15~ 19 步, 算法需依次调整从信源 s 到 $|V| - j$ ($j = 1, 2, 3, \dots, (|V| - 1)$) 个未被访问节点的距离, 最多需进行 $|V|^2/2$ 次操作. 所以, DBCT 算法的时间复杂度为 $O(|V|^2)$.

推论 1: 与文[3, 6, 7]所给出的算法相比, DBCT 算法具有更快的运算速度.

文[6]所给出的算法的时间复杂度为 $O(|V|^3)$; BSMA 算法^[3]的时间复杂度为 $O(k|V|^3 \log |V|)$, 其中, k 为一常数, k 的值越大, 则 BSMA 算法的优化性能越好^[3]; 文[7]所给出的算法时间复杂度为 $O(|V| \cdot |V|^2 \log_2(B \cdot D \cdot |V|))$, 其中, $|V| \leq |V|$, B 和 D 分别为 E 中链路带宽的最大值和链路延时的最大值^[7]. 由于 DBCT 算法的时间复杂度为 $O(|V|^2)$, 所以, 与文[3, 6, 7]所给出的算法相比, DBCT 算法具有更快的运算速度.

性能 2: 当网络中存在多个组播通信组(以下称通信组), 且各通信组有各自的时延及带宽要求时, DBCT 算法可用于为每个通信组分别构造一棵符合各自时延及带宽约束的组播树.

分析 DBCT 算法第 10~ 14 步可看出, DBCT 算法考虑了带宽的预留. 当算法为某个通信组构造组播树时, 算法不仅为此通信组预留了带宽, 而且计算出了在预留带宽后各链路所剩

* 对于图 G 中任意给定的两节点 a, x , 若链路 (a, x) 不存在, 则本文令 $D(a, x) = \infty$.

余的可用带宽. 因此, 当网络中存在 m 个通信组时, 仅需调用 m 次 DBCT 算法即可为每个通信组分别构造一符合各自延时及带宽约束的组播树. 此时, 算法的时间复杂度为 $O(m \cdot |V|^2)$.

设集合 $A = \{A_1, A_2, \dots, A_{m-1}, A_m\}$, 其中 $A_i (i = 1, 2, \dots, m-1, m)$ 中存放的是第 i 个通信组的参数, 参数包括此通信组的信源 s 、信宿集 M 、通信所要求的最小带宽 W_{\min} 及延时上限 D_{\max} . 则用 DBCT 算法为多个通信组分别构造组播树的过程可描述为:

while $A \neq \phi$ do

从 A 中弹出首个通信组的参数 $\{s, M, D_{\max}, W_{\min}\}$;

// 依次取出各通信组的参数.

DBCT($G, s, M, D_{\max}, W_{\min}$); // 调用 DBCT 算法构造符合延时及带宽约束的组播树.

性能 3: DBCT 算法所构造的组播树 (以下简称 DBCT 树) 定能符合延时及带宽约束.

根据算法的第 7~14 步知: 对于 $\forall v \in V_T$, 有 $\text{Dis}(v) < \infty$.

分析算法的第 1~6 及 15~19 步知: 对于 $\forall a \in V$ 及 $\forall P(s, a) \subseteq G$, 当 $\text{Delay}(s, a) > D_{\max}$ 时, $\text{Dis}(a) = \infty$; 当 $\text{Delay}(s, a) \leq D_{\max}$ 时, 算法根据式 (5)、(6) 及定义 3 来计算 $\text{Dis}(a)$.

又因为, 对于 $\forall a \in V$ 及 $\forall P(s, a) \subseteq G$, 若 $\text{Width}(s, a) \geq W_{\min}$, 则由式 (2.1) 及定义 1 可推出: 对于 $\forall (x, y) \in P(s, a)$, 有 $MB(x, y) = 1$; 若 $\text{Width}(s, a) < W_{\min}$, 则由式 (2.1) 及定义 1 可推出: $\exists (x, y) \in P(s, a)$, 使 $MB(x, y) = \infty$.

因此, 根据算法第 19 步及定义 3 可推出: 对于 $\forall a \in V$ 及 $\forall P(s, a) \subseteq G$, 当且仅当 $\text{Width}(s, a) \geq W_{\min}$ 并且 $\text{Delay}(s, a) \leq D_{\max}$ 时, $\text{Dis}(a) < \infty$; 否则 $\text{Dis}(a) = \infty$.

因此, DBCT 树定能符合延时及带宽约束.

5 实验分析

实验网络的拓扑生成部分借鉴了 Salama 博士的拓扑生

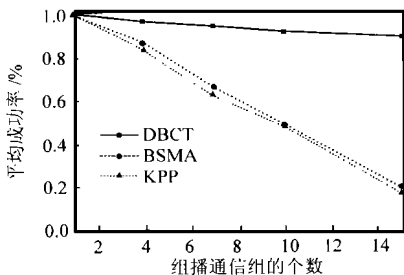


图1 各种算法的平均成功率比较

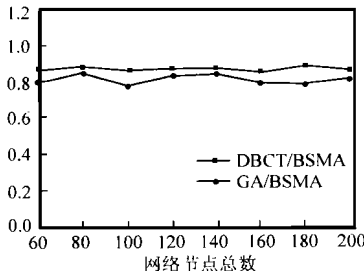


图2 各种算法的优化性能比较

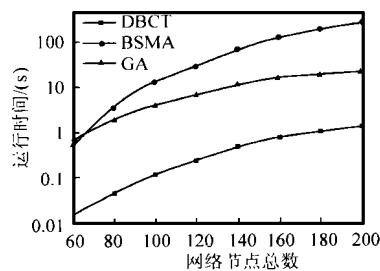


图3 各种算法的运行时间比较

5.3 算法运行时间

图3示出了DBCT、BSMA^[3]和GA^[8]算法在各种网络规模情况下的运行时间. 可以看出, DBCT 算法的运行时间明显低于 BSMA 及 GA 的运行时间, 且随着网络节点的增多 DBCT 算法的优势更加明显. 这说明在解决大规模网络中的组播路由问题时, 从运算时间上考虑本算法比 BSMA 及 GA 优越很多, 更适用于大规模的网络通信.

成算法^[10], 该算法保证结点平均度为 4, 这与现实的网络环境很接近. 实验网络中各链路带宽均为 44.736Mbps (DS3); 链路的延时正比于节点间的欧拉距离, 在 $[5, 50\text{ms}]$ 上分布; 链路的代价在 $[1, 100]$ 上分布; 实验中组播信源及信宿随机选择, 组播规模 $(|M|/|V|)$ 在 $5\% \sim 20\%$ 之间变化; 延时约束 D_{\max} 在 $[100\text{ms}, 500\text{ms}]$ 上分布; 实时业务所需的带宽为 $[5\text{Mbps}, 40\text{Mbps}]$ 间的随机数.

为验证 DBCT 算法的性能, 我们对算法的平均成功率、优化性能及运行时间进行了检验, 并将 DBCT 算法与几种现有算法作了比较. 所有实验结果为多次实验所取的平均值.

5.1 平均成功率

算法的平均成功率 θ 定义为: $\theta = N_s/N_t$. 其中, N_s 为算法成功地找到满足延时及带宽约束的组播树的平均次数; N_t 等于组播业务总数除以实验次数.

图1给出了当网络节点数为 180 时, KPP、BSMA^[3]和 DBCT 算法的平均成功率的实验结果. 可见, DBCT 算法的平均成功率要远高于其他算法, 且随网络中通信组个数的增加, 其优势愈加明显. 这是因为, KPP 及 BSMA 没有考虑带宽约束, 使得多个通信组共享少数链路造成这些链路拥塞, 从而导致了路由失败. 此外, 虽然文^[7]的方案考虑了带宽约束, 但其算法的平均成功率在 90% 左右^[7], 因此, DBCT 算法的平均成功率也高于文^[7]中算法的平均成功率. 由此可知: DBCT 算法不仅具有极高的路由发现能力, 而且适用于解决当网络中存在多个组播业务时的路由选择问题.

5.2 优化性能比较

由于在求解基于延时限制最小代价的组播路由算法中, BSMA^[3]被认为是优化性能超群的算法, 所以该实验以 BSMA 算法为参考. 图2给出了 DBCT 和 GA^[8]算法分别与 BSMA 算法对组播树代价的优化性能的比较结果. 可见, 虽然 DBCT 的优化性能比 GA 略差, 但比 BSMA 算法要好 15% 左右. 这说明与 GA 一样, DBCT 算法也具有很好的优化性能.

6 结论

基于延迟及带宽受限的最小代价组播路由问题是典型的优化问题, 此类问题属 NP 完全问题^[4-8]. 本文提出了一种新的启发式算法——DBCT, 算法通过计算每个节点的距离, 巧妙地计算出了具有 NP 难度的在多个约束条件下代价最小组播路由问题的近似解. 在最坏情况下, 算法的时间复杂度为 $O(|V|^2)$. 相对于其他算法, DBCT 有以下优点: 简单性、快速性、

成功率高、支持多点到多点的组播通信、能有效地提高网络的利用率。该算法也容易扩展到其他具有 NP 难度的 QoS 路由问题,能适用于大规模的网络路由计算。

参考文献:

- [1] F Xiang, L Junzhou. QoS routing based on genetic algorithm [J]. Computer Communication, 1999, 22: 1394- 1399.
- [2] M Parsa, Q Zhu, J J Garcia-Luna-Aceves. An Iterative algorithm for delay constrained minimum cost multicasting [J]. IEEE/ACM Transactions on Networking, 1998, 6(4): 461- 474.
- [3] Zhu Q, Parsa M, Garcia-Luna-Aceves J. A source based algorithm for delay constrained minimum cost multicasting [A]. in Proc. of IEEE INFOCOM' 95 [C], 1995: 377- 385.
- [4] Nageswara S V Rao, et al. QoS routing via multiple paths using bandwidth reservation [A]. in proceedings of INFOCOM [C], 1988: 11- 18.
- [5] Z Wang, J Crowcroft. Bandwidth delay based routing algorithms [A]. IEEE Globecom' 95 [C], 1995.
- [6] Youngseok Lee, Yongjiun. A bandwidth and delay constrained minimum cost multicast routing algorithm [A]. ICCOM-11, Taipei [C], Jan. 1997.
- [7] Adil I Erzin. Polynomial algorithm for bandwidth delay constrained multicast routing problem [DB/OL]. URL: <http://citeseer.nj.nec.com/ cachedpage/ 29205/ 2>

- [8] WANG Zhengying, SHI Bingxin. Genetic algorithm for minimal multicast routing with bandwidth and delay constraints [DB/OL]. URL: <http://www.csu.edu.cn/ci/sub07/zheng02/>.
- [9] 陈国良, 王熙法, 庄镇泉, 等. 遗传算法及其应用 [M]. 北京: 人民邮电出版社, 1996, 6.
- [10] H F Salama, D S Reeves, Y Viniotis. A distributed algorithm for delay-constrained unicast routing [A]. IEEE INFOCOM' 97 [C], Japan, April 1997.

作者简介:



石 坚 男. 1966 年出生于湖南省. 1987 年毕业于华中科技大学(原华中理工大学)计算机专业, 1993 年在华中科技大学机械学院获硕士学位, 现于华中科技大学电子与信息工程系攻读博士学位, 从事数据通信与网络路由技术方面的研究。



董天临 男. 1965 年毕业于华中工学院电机工程系. 1982 年和 1985 年分别获美国 Polytechnic Institute of New York 的 Electrophysics 硕士和博士学位. 博士生导师, 教授, 主要研究领域包括光纤通信与网络技术等。