

# 一种基于时延限制的分布式组播路由选择算法

李 博

(广东水利电力职业技术学院 电教计算机信息中心, 广东 广州 510635)

摘要: 基于最短路径算法提出了主树的概念, 并且利用组播树的代价和时延限制的定义给出了一种分布式组播路由选择算法. 可以在时延限制的前提下使组播树的代价接近最优. 最后, 说明了本文方法的有效性.

关键词: 时延限制组播树; 分布式路由选择算法; 主树; 实时通信

中图分类号: TP301.6

文献标识码: A

文章编号: 1001-411X(2003)04-0096-04

组播是指从某一源节点向一组目的节点同时发送消息, 实时组播就是所有的目的节点必须在一个特定的时间限制内接收到消息的一种组播方案. 当前有许多通信都依赖于实时组播服务, 例如: 交互式视频会议系统、分布式网络游戏系统、共享编辑系统等. 在实时通信系统中, 需要在传输数据之前建立从源节点到目的节点的一种逻辑连接. 连接建立时, 要在连接上的每个网络节点预留充足的资源(例如网络带宽、缓冲器等), 才能在传输数据时保证用户要求的服务质量. 建立组播连接过程一般分为 2 步: 路由选择和连接配置. 路由选择是指找到源节点连接到所有目的节点的路由选择树; 连接配置是指在树中每个节点上配置新的连接的操作过程, 包括预留网络资源、在路由表中注册连接(为了把进来的分组交换到相应的输出链路上等).

在面向连接的通信中, 一旦建立连接, 所有的业务量都沿着预先选择的路由传送, 直到连接拆除为止. 因此, 在组播连接期间, 选择一个网络代价最小的路由选择树是很重要的. 一般来说在网络中找到这样的组播树问题称作 Steiner 树问题<sup>[1]</sup>, 它是 NP-完全问题<sup>[2]</sup>. 为了找到 Steiner 树的近似解, 有人已经提出了一些启发式算法<sup>[3,4]</sup>. 然而, 在实时组播连接中应用启发式算法主要有如下两个困难:

1) 不能以完全分布的方式进行路由选择, 这是因为大多数启发式算法都是集中式的. 集中式算法必须由一个中心节点负责整个组播树的计算任务, 而且中心节点必须知道整个网络的信息. 在大规模网络中, 就会显露出许多缺点, 例如: 对错误的容忍度(尤其当中心节点出错时)、中心节点会负担过重的计算负载、在保持网络信息更新时将产生较高的通信代价和不准确的路由选择信息等.

2) 时延限制下计算最优的路由选择树可能遇到

NP 问题. 因为, 对于实时的连接不仅要求最小的网络代价, 还要求从源节点到所有目的节点的时延不能超过某个指定的界限. 然而, 在组播路由选择中, 要求最小的网络代价通常和时延限制要求互相矛盾. 目前, 建立连接的方法通常把路由选择操作和连接配置操作分开, 即首先建立路由选择树, 然后再沿着树分别为每个节点完成配置操作. 这不仅增加了网络通信的代价, 而且增加了建立连接所用的时间.

本文针对实时组播连接提出一种分布式路由选择算法, 目的是克服以上缺点. 在时延限制下, 找到近似最优的路由选择树. 因此, 算法使用了较少的信息和较短的收敛时间来建立组播连接.

## 1 系统模型和问题阐述

### 1.1 问题的定义和系统模型

组播就是把消息传送给组播要求中的所有目的节点. 以减少网络中数据分组的复制次数为出发点, 来设计组播路由选择树的结构(分组只在树的分枝点出现复制). 组播树定义如下:

对于给定的图  $G(V, E)$ , 给定一个源节点  $s \in V$ , 一个目的节点集合  $D$  包含于  $V$ , 且  $s$  不属于  $D$ , 组播路由选择树就是图  $G(V, E)$  的子图. 其中根节点为  $s$ , 包含  $D$  中的所有节点, 而且可以包含  $V-D$  中任意的子集合, 同时该子图的叶节点只能是  $D$  中的节点.

当组播消息时, 源节点  $s$  (组播树的根) 把消息沿着树传送到下一代节点. 这些子节点把消息分组发送到它的下一代节点, 直到树中属于集合  $D$  的所有节点都收到源节点发送的数据分组. 如果  $|D|=1$ , 转化为点对点通信问题. 如果  $|D|=|V|-1$ , 转化为广播问题.

使用连接图  $G(V, E)$  建立网络模型. 图中的节

点表示网络路由器, 边界表示节点之间的通信链路. 每个链路  $l \in E$  有 2 个参数:  $c(l)$  和  $d(l)$ .  $c(l)$  表示链路  $l$  的通信代价(一般指跳数).  $d(l)$  表示传送信息时在链路  $l$  上经历的时延(包括排队时间、传输时间). 图 1 是一个简单的网络图例. 为清楚起见, 使用同一整数来表示图中的代价值和链路的时延参数(这不影响图的一般性).

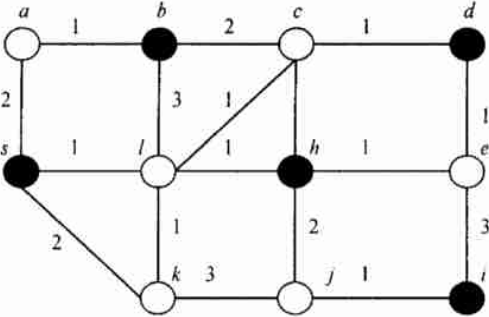


图 1 简单的网络图例

Fig.1 A sample of net

根据树的性质, 组播消息时, 只有通过组播树的每一个边, 才能到达所有的目的节点. 因此, 把消息组播到所有目的节点的网络代价对应于树中所有链路的代价的和. 我们定义组播路由选择树  $T$  的代价[记为  $NC(T)$ ] 为:

$$NC(T)=\sum_{l \in T} c(l), \tag{1}$$

组播路由选择的一个重要目标就是最小化路由选择树的代价(即找到最优的路由选择树). 通过选择最优组播路由选择树, 可以最小化组播树的代价.

另一个就是传送消息的时延限制, 即从源节点到任何目的节点的通信时延不能超过某个上限值. 时延上限的精确值依赖于特定的应用. 设  $P(s, u)$ ,  $u \in D$ , 是沿着路由选择树从源节点  $s$  到节点  $u$  的路径,  $\Delta$  是时延的上限值, 满足

$$u \in D : \sum_{l \in P(s, u)} d(l) \leq \Delta, \tag{2}$$

网络代价在整个组播树是可加的, 而时延限制要求只能沿着组播树从源节点到目的节点的某个路径上是可加的. 我们的目标是在式(2)的条件下, 找到具有最优网络代价的组播路由选择树. 另外, 为了保证满足不等式(2)的组播树的存在性, 必须满足如下条件

$$u \in D : \sum_{l \in SDP(s, u)} d(l) \leq \Delta, \tag{3}$$

其中,  $SDP(s, u)$  是从源节点到任何节点  $u$  的最短时延路径. 否则满足时延限制  $\Delta$  的组播路由选择树可能不存在.

1.2 满足时延限制的最优路由选择树

在组播路由选择中, 代价要求常常和时延要求

相互矛盾. 具有最优网络代价的 Steiner 树也许到 Steiner 树的最远目的节点有较长的时延. 在实时通信中, 实际上不必最小化到所有目的节点的平均时延. 通常要求到所有目的节点的时延在某个限制之内. 我们可以找到一些时延限制路由选择算法如: 代价一时延启发式和代价启发式算法. 代价一时延启发式算法把链路的代价和时延转换成权值, 通过权值来优化网络. 代价启发式算法只是在满足  $D[v] < \Delta$  条件下, 选择至树的代价最小的节点  $v$ , 并把该节点连接到树中. 这 2 种算法后来被扩展到分布式算法<sup>[9]</sup>. 在其扩展的分布式算法中, 选择连接边界的策略是一样的. 因此提出完全分布的、能建立时延限制的最优树组播路由选择算法.

2 基于时延限制的分布式主树组播算法

2.1 分布启发式算法的要求与假设

在创建组播树的算法过程模仿 Prim's MST 算法<sup>[6]</sup>, 而且和分布式最短路径算法结合. 假设每个节点都包含最短路径的信息(按照代价标准)和到其他节点的时延信息(这些信息被保存在局部路由表中). 可以基于代价和时延标准运行 Bellman Ford 距离矢量算法来完成算法<sup>[7, 8]</sup>. 首先通过定义主树, 然后通过顺序增加每个目的节点来创建组播路由选择树, 并且随着树的生长, 把路由选择树的连接信息注册在每个树中节点的连接表中. 可以证明算法是无环路的.

最优组播路由选择树就是连接所有目的节点且代价最小的路由选择树.

定义 给定源节点  $s$  和目的节点集合  $D$ , 存在节点  $u \in D$ , 如果  $u$  和源节点  $s$  利用某种规则建立了连接, 当且仅当  $u$  是唯一的且是第一个连接到源节点的目的节点时, 称由源节点  $s$  和节点  $u$  以及他们连接路径上的所有节点和链路组成的树为“主树”. 以下用  $MTN$  表示主树;  $m$  表示主树中节点的集合.

假设 为每个主树节点赋予一个主树路由表 Mutable, Mutable 对应 3 项:  $MT.c$ 、 $MT.tag$ 、 $MT.del$ , 分别表示他们到源节点的代价、作为主树节点的标志、到源节点的时延. 对于网络中的每个节点  $u$  都有一个局部路由选择表 Route Table, Route Table 包含 2 项:  $RT[u].c$ 、 $RT[u].d$ ; 其中  $RT[u].c$  表示节点  $u$  到其局部节点的最短路径的代价;  $RT[u].d$  表示路径上的时延. 每个节点要记住当收到目的节点请求时所积累的时延  $Du$ .

2.2 算法的详细描述

当某个节点收到一个带有参数  $D$  和  $\Delta$  的打开

组播连接的请求时, 该节点称为源节点. 首先, 任意选择某一目的节点  $u$ , 并在满足时延限制  $\Delta$  的条件下, 利用最小时延路径  $SDP(s, u)$  和节点  $u$  建立连接<sup>[9]</sup>. 源节点通过路径  $SDP(s, u)$  向目的节点  $u$  发出指令: 任命路径  $SDP(s, u)$  为主树, 即刷新  $SDP(s, u)$  路径上的每个节点  $v$  的路由表项  $RT[v].tag = MTN; RT[u].c = MT. c; RT[u].d = MT. del$ . 下面考虑其他目的节点  $k \in D - u$  的连接建立过程, 节点  $k$  向距离它最近的(最小代价)节点  $k_{near}$  发送组播连接请求信息, 请求信息中包括源节点地址  $S_{address}$  和节点  $k$  到节点  $k_{near}$  的时延  $D_k$ . 如果  $k_{near}$  为主树节点, 则计算是否满足:  $D_k + MT. del \leq \Delta$ . 如果满足, 则向源节点发送节点  $k$  已经建立连接的消息; 否则, 节点  $k_{near}$  向目的节点  $k$  发送无法建立连接的拒绝消息, 当节点  $k$  接收到拒绝消息后, 立即选择次代价最短路径重复以上建立连接过程, 直到建立连接为止. 根据前面的假设可知: 在有限的时间内, 一定可以找到满足时延要求的路径和源节点建立连接. 不失一般性, 如果某个目的节点和主树中某个节点  $w$  建立了连接, 其中  $w \in s \cup M$ , 如果  $w$  为源节点  $s$ ,  $s$  直接向目的节点发送确认连接消息. 如果  $w \in M$  且  $w \neq s$ , 则节点  $w$  不仅要向目的节点发送建立信息, 还要向源节点发送 fork 信息, 即说明该节点是一个分支节点, 将来转发组播分组时, 该节点要完成一个信息的复制操作, 并指明分组流输出的方向, 因此也就把建立连接过程和连接配置过程结合起来了<sup>[10-11]</sup>. 当源节点收到所有目的节点的连接建立信息时, 立即向所有的目的节点进行组播传送数据信息.

节点  $c$  路由表的刷新算式:

$MT. c = 2, \quad MT. tag = MTN, \quad MT. del = 2$

以图 1 作为网络例子, 选择节点  $s$  为组播源节点, 目的节点集合  $D = \{b, d, h, i\}$ . 在这里我们假设时延和代价取一个值.

图 2 表示目的节点  $d$  根据满足时延限制条件 Bellman Ford 算法形成的主树<sup>[8]</sup>, 并且主树中每个节点的局部路由表按照路由表的刷新算式来刷新. 当节点  $i$  开始建立连接时, 首先将连接请求发送给节点  $j$  (基于最短路径算法), 节点  $j$  纪录  $i \rightarrow j$  的时延,  $j$  再基于同样的算法把连接请求发送到  $h$ , 并纪录  $i \rightarrow h$  的时延, 即,  $delay_{i \rightarrow h} = delay_{i \rightarrow j} + delay_{j \rightarrow h}$ . 节点  $h$  类似节点  $j$  的操作, 把连接请求发送到节点  $l$ , 并计算  $i \rightarrow l$  的时延,  $delay_{i \rightarrow l} = delay_{i \rightarrow h} + delay_{h \rightarrow l}$ . 因为节点  $l$  是主树节点, 节点  $l$  首先提取出到目的节点的时延  $MT[l].del$ , 然后计算是否满足  $delay_{i \rightarrow l} + MT[l].del \leq \Delta$ . 如果满足, 节点  $l$  向源节点发送节点  $i$

进入组播树的消息分组 fork 信息, 另外向目的节点发送连接确认消息; 否则, 向目的节点  $i$  发送连接拒绝消息<sup>[12-13]</sup>; 目的节点如果接到拒绝消息, 就寻找代价次优的路径重复以上过程.

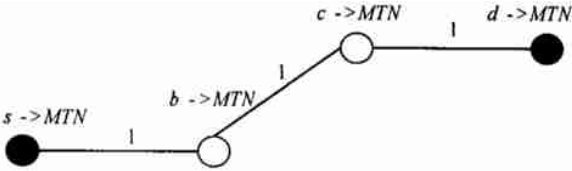


图 2 选择目标的节点  $d$  建立的主树

Fig. 2 The main tree of select object point  $d$

以上是一种完全分布式的算法. 对于一个连接建立有 5 种控制信息: (1) 对应于应用的组播初始信息, 即确定源节点  $s$  和目的节点集合  $D$ ; (2) 确定主树信息; (3) 发送作为分支节点的 fork 信息; (4) 发送连接请求信息; (5) 连接确定信息. 这些信息相互合作来建立组播连接树.

3 算法的讨论

定理 1 时延限制路由选择树中没有“环路”, 假设基本的最短路径路由是无环路的.

证明 主树形成路由显然没有环路. 当某个目的节点  $i$  基于基本的最短路径算法连接到主树中的某个节点  $w$  时, 首先目的节点  $i$  到  $w$  之间是没有环路. 因此, 节点  $i$  到源节点的连接满足基本的最短路径算法, 所以也一定没有环路. 定理 1 同时证明了本文提出的算法没有象并行算法讨论的死锁问题<sup>[14]</sup>.

定理 2 总能找到时延限制组播路由选择树, 只要它存在.

证明 假设对于某个组播源  $s$ , 目的节点集合  $D$  以及时延限制  $\Delta$  存在时延限制路由选择树, 即对于给定的  $s, D, \Delta$  对任意  $u \in D$  一定存在满足 (3) 式的路径. 然而, 最短路径又是最小时延路径, 所以最短路径一定满足时延限制. 根据本文提出的算法, 主树是源节点基于最短路径算法从源节点到某一目的节点形成的, 而其他目的节点连接到主树的过程也是基于最短路径算法, 并且以最小时延路径为首选路径, 如果不满足, 就选择次优的路径, 也就是说是从小到大的选择过程. 因此, 即使在最严格的时延要求下 ( $\Delta$  等于按照最短路径到最远的目的节点的时延), 通过我们的算法也能找到满足时延要求的组播树, 证毕.

4 结语

这种完全分布式的时延限制组播路由选择算

法. 除了具有分布式特性外, 还具有较低的网络代价. 另外方法把路由选择和连接配置结合成一个操作过程. 因此削减了连接建立所用的时间. 这方面进一步的理论, 还有待于进一步的研究. 本方法处理大规模网络, 可能要花费很长的时间来建立连接, 所以方法还有待于进一步完善, 这也是今后要做的主要工作.

参考文献:

[ 1 ] GILBERT P. Steiner minimal tree[ J] . SIAMJApplMath, 1968, 16.  
[ 2 ] GAREYMR J. Computer sand intract ability: Aguide to the theory of NP completeness[ M] . SanFrancisco CA: Free Man, 1979. 61— 70.  
[ 3 ] HWANGFK R. Steiner tree problem[ J] . Networks, 1992, 22: 55— 89.  
[ 4 ] WINTERP S J. Path distance heuristic for the Steiner problem in undirected networks[ J] . Algorithmica, 1992, 3: 309— 327.  
[ 5 ] KOPPELLAVP P P. Optimal multicast routing with quality of

service constraint[ J] . J Network SystManagement, 1996, 4 (2): 107— 131.  
[ 6 ] 程诗杰, 施 荣. IP 多播实现及其安全性[ J] . 计算机时代, 2000, (9): 19— 20.  
[ 7 ] 王 晖, 胡晓峰, 库锡树, 等. 一个实用的多点通信部件的设计与实现[ J] . 计算机工程, 1999, 25 (8): 10— 11.  
[ 8 ] 丁 煜, 许林英. IP 多点广播技术及应用[ J] . 计算机工程, 2000, 26(7): 3— 5.  
[ 9 ] 蒋海东, 裘正定. 多播技术在 Internet 上的应用[ J] . 通讯世界, 1999, (10): 38— 40.  
[ 10 ] 刘玉明. IP 多播技术[ J] . 电信科学, 1999, (2): 18— 20.  
[ 11 ] 师 军. Multicast 技术及其应用[ J] . 微型机与应用, 2000, (2): 29— 33.  
[ 12 ] 张 伟. 基于图搜索的网络路由算法复杂度分析[ J] . 辽宁大学学报, 1995, 22(1): 29— 33.  
[ 13 ] 刘 越, 张宝贤, 陈常嘉. IP 多播中的路由技术[ J] . 中国数据通信网络, 2000, (1): 43— 46.  
[ 14 ] 袁有伟, 余江鸿, 湛含辉. 基于神经网络的动态路由算法[ J] . 计算机工程与应用, 2002, (12): 165— 167.

A Distributed Routing Algorithm Based on Delay-Limiting

LI Bo

(Electric Education Centor, Guanglong Technical College of Water Resources and Electric Engineering, Guangzhou 510635, China)

**Abstract:** This paper introduces the concept of main tree based on shortest path algorithm, and by applying the concept of the multicast tree cost and delay-limiting, it maces a presentation of the distributed heuristic routing algorithm which can optimize the cost of multicast tree under the condition of delay-limiting. Finally, it states the validity of the method in this text.

**Key words:** multicast tree based in delay-limiting; distributed routing algorithm; main tree; real-time communication

【责任编辑 李晓卉】