

一种多受限最小代价的动态组播路由算法^{*}

A Multi-Constrained Dynamic Least Cost Multicast Routing Algorithm

石 坚 董天临

(武汉华中科技大学电子与信息工程系 武汉430074)

Abstract The paper presents a new dynamic Multicast Routing 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 multiple multicast communications exist and multicast nodes change dynamically. The scheme not only guarantees the bandwidth and end-to-end delay optimizing the cost of the routing tree, but also effectively controls the algorithm complexity for large size of network.

Keywords Multicast, Bandwidth, Delay, Cost, Message

1. 引言

为确保通信网能提供(QoS)服务质量保证,必须研发有效的基于QoS的路由机制以提供高质量信息传输。一般地,基于QoS的路由要达到两个目标:一是要满足用户的QoS要求,如必须提供足够的带宽、足够小的延时和延时抖动等;二是要优化网络的利用率及代价。近年来,各国学者都开始关注基于QoS的路由问题。由于此类问题属于NP-Complete问题,所以各国学者大都采用启发式方法求解。文[1~3]提出了一些适用于信宿固定情况下的源路由算法。文[4~7]提出了几种动态路由算法,由于求解动态组播树的问题很复杂,大部分学者都将此问题分为两个部分求解:中心点(负责组播树的维护并将会话的状态传给所有网络节点)求解和基于中心点的路由选择,如PIM-SM^[4,5]和CBT^[6,7]算法。本文提出了一种多受限最小代价的动态组播路由算法MDLCMR(Multi-constrained Dynamic Least Cost Multicast Routing)。该算法不必求解中心节点,适合于组播节点的动态变化,是一种快速动态组播路由算法。

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。

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 $M \subseteq V - \{s\}$,寻找满足条件: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:

$$Cost(T) \rightarrow Min(\sum_{(x,y) \in E_T} C(x,y)) \quad (1)$$

$$\sum_{(x,y) \in P_T(s,v)} D(x,y) \leq D_{Max} \quad \forall v \in M \quad (2)$$

$$Width(P_T(s,v)) \geq W_{Min} \quad \forall v \in M \quad (3)$$

其中 $Cost(T)$ 表示组播树 T 的代价,式(1)描述了路由树的代价应达到最小;式(2)、(3)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 T 到信宿 v 的路径。

3. MDLCMR 算法

若网络中有 $|M|$ 个组播通信组,每个通信组有各自的信源 s 、各自的组播成员集 M 及各自的组播连接请求,本文为各组播连接请求设立一唯一的标识号 id ,这样,当网络中有多个不同的请求时不会发生消息冲突。

MDLCMR算法首先选择当前组播树 $T'(T'$ 由信源 s 、已加入组播的信宿以及信宿加入组播树的路径组成)构成初始组播树 $T'=(V_{T'}, E_{T'})$,然后根据信宿的连接请求或退出请求,依照加入和退出操作规则,动态地建立或切断连接。组播树的形成过程其实就是信宿的动态加入和退出过程。MDLCMR算法由连接请求、

^{*}教育部高等学校骨干教师资助计划项目。

加入操作及退出操作三部份组成。

3.1 连接请求

连接请求操作由申请加入组播(multicast)的网络节点完成。当G中某节点v申请加入组播时,先使用本文随后提出的多受限最小代价路径算法MCLCP(Multi-Constrained Least Cost Path)找出由v到当前组播树T'的满足(1)、(2)和(3)式的路径P,然后沿P及T'向信源发送连接请求消息Add(id,s,v,P),MCLCP算法描述如下:

```
MCLCP(G,T',v,DMax,WMin)
For 每个V中的节点u do //初始化
  Cost(u,v)=∞; Delay(u,v)=∞; Poi[u]=NIL; //指针
  Poi[u]用于记录节点u的父节点
  Q=V; R=φ; L=t; Cost(v,v)=0; Delay(v,v)=0;
(1) While L=t do //寻找满足带宽及延时要求的最小代价路径
(2) 从Q中弹出具有最小Cost(u,v)的节点u; R=R+(u);
(3) if u∈VT then 计算路径PT(s,u)的延时Delay(PT(s,u));
(4) if Delay(u,v)≤DMax-Delay(PT(s,u)) then
  While Poi[u]≠NIL do
    B(u,Poi[u])=B(u,Poi[u])-WMin; //修改P上各链路的可用带宽
    P=P+{(u,Poi[u])}; L=f; //记录路径P
(5) for 每个与u相邻的节点w do
(6) if (Cost(w,v)>Cost(u,v)+Mb(w,u)·C(w,u)) and
  (Delay(u,v)+D(w,u)≤DMax) then
  Cost(w,v)=Cost(u,v)+Mb(w,u)·C(w,u);
  Delay(w,v)=Delay(u,v)+D(w,u);
  Poi[w]=u; //计算w到v的代价和延时并记录此路径
(7) if (u=s) or Cost(u,v)=∞ then //寻找满足带宽约束的最小延时路径
(8) Q=Q+R;
(9) While L=t do
(10) 从Q中弹出具有最小Delay(x,v)的节点x;
(11) if Delay(x,v)>DMax then 路径不存在,v与s协商放宽延时限制条件; L=f;
(12) if x∈VT then 计算路径PT(s,x)的延时Delay(PT(s,x));
(13) if Delay(x,v)≤DMax-Delay(PT(s,x)) then
  While Poi[x]≠NIL do B(u,Poi[x])=B(u,Poi[x])-WMin; //修改P上各链路的可用带宽
  P=P+{(u,Poi[x])}; L=f;
  else if x=s then 路径不存在,v与s协商放宽延时限制条件; L=f;
(14) for 每个与x相邻的节点y do
  if (Delay(y,v)>Delay(x,v)+Mb(x,y)·D(x,y)) then
    Delay(y,v)=Delay(x,v)+Mb(x,y)·D(x,y); Poi[y]=x;
```

3.2 加入操作

信源接收到加入请求信息后开始加入操作。

不失一般性,讨论以G中任一未加入组播的信宿v为例。设此时的初始组播树为T',v到节点u的路由路径为P(u为P与T'相交的连接点),T'中s到u的路由路径为P_T(s,u)。

加入操作的过程即信宿连接到组播树的过程。为确保树的特性,在信宿加入组播树时,必须对组播树进行维护。组播树的维护由信源s完成,在s接收到连接请求消息Add(id,s,v,P)后,s先将T'与路径P合并构成G的子图G',然后用Kruskal算法求G'的最小延时树T,最后将信宿v加入到注册成员表并沿T给出

确认消息ACKA(id,s,v)。

对于T上的任意一节点x,当x接收到ACKA(id,s,v)后,若x≠v且x是T的叶子节点,则x把自己的状态置为CLOSED,否则x把自己的状态置为ESTAB从而完成加入操作。经过加入操作,节点v的加入路由路径为P,最终路由路径为P_T(s,u)+P;组播树为T。

3.3 退出操作

退出操作由组播树上的任意一个非信宿的叶节点(状态为CLOSED的叶节点)执行。当某信宿v欲退出组播时,若v不是T的叶节点,则v仅向信源s发送退出请求消息Remove(id,s,v)即可,否则v先释放先前分配的资源,然后将自己的状态改为CLOSED。退出操作包含删除请求及链路删除两个模块:

(1)删除请求模块由T中任一叶节点x执行:

if x处于CLOSED状态 and x是T的叶节点 then
沿路径P_T(s,x)向其父节点w发送链路删除消息Delete(id,x,w)。

(2)链路删除模块由收到链路删除消息的节点执行:

if 节点w收到链路删除消息Delete(id,x,w) then
B(w,x)=B(w,x)-W_{Min}; //修改被删除链路的可用带宽
删除T中的链路(w,x)后向s发送链路已删除消息Deleted(id,s,w,x)。

对于退出操作,非叶子信宿退出时,组播树拓扑不作任何改变;叶子信宿v退出时,v沿路径P_T(s,v)向其父节点w发送链路删除消息Delete(id,v,w),父节点w接收到删除消息后,如果w不只是v的中间节点,删除链路(w,v)即可,否则删除链路(w,v)后,继续沿路径P_T(s,w)向父节点发送删除消息。这样一来,新的组播树T=T-P_T(u,v),其中u为v从T上断开的节点。由于MDLCMR算法包含删除请求及链路删除模块,因此,在加入操作中不必剪去T中非信宿的叶节点。

4. 实验研究

实验中所使用的通信网络是全双工的有向网络,实验网络的拓扑生成部分借鉴了Salama博士的拓扑生成算法^[6],该算法保证节点平均度为4,这与现实的网络环境很接近。链路带宽的初值均为44.736Mbps(DS3),实验中信源、信宿随机选择,信宿v的延时约束D_{Max}在[30ms,160ms]均匀分布,实时业务所需的带宽为[10Mbps,40Mbps]间的随机数,链路的延时正比于节点间的欧拉距离,在[5,50ms]上分布,链路的代价在[1,200]上均匀分布。

4.1 动态组播组的构造

为了有效地测试动态组播路由算法的性能,我们提出了一个动态组播组构造方案,该方案不仅保证了组播组的信宿成员是不断更新的,而且保证了组播成员更新的随机性。在这里,“随机性”包括两个方面的内容:①节点加入和退出组播组这两个事件的出现是随机的;②加入或退出组播组的节点是随机选择的。动态组播组构造方案可描述如下:

(1)给定任意一个网络 $G=(V, E)$, 首先在网络 G 中随机选择一个不参与更新的永久节点 s 作为信源或汇聚点, 然后随机选择任意几个节点作为信宿。

(2)对于 $\forall u \in V$, 令节点 u 可能成为组播节点的概率 $PR_u(|M|)$ 为:

$$PR_u(|M|) = \frac{\eta \cdot (|V| - |M|)}{\eta \cdot (|V| - |M|) + (1 - \eta) \cdot |M|} \quad (4)$$

式(4)中,“ η ”为预先输入的实验所需组播规模期望值, $0 < \eta < 1$ 。

(3)在网络 G 中随机选择一个节点 u , 给定任意一个随机数 ζ ($0.5 \leq \zeta \leq 1.5$), 若 $PR_u(|M|) \cdot \zeta > 0.5$, 则令节点 u 为组播节点; 对于 $\forall v \in M$, 若 $(1 - PR_v(|M|)) \cdot \zeta > 0.5$, 则令节点 v 为非组播节点。

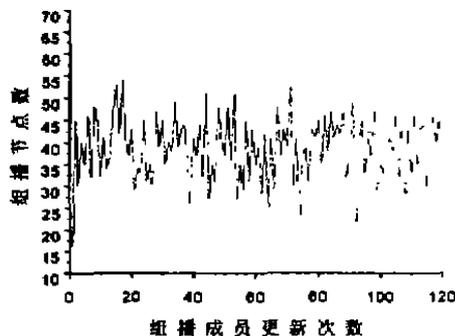


图1 组播节点数与更新次数之间的关系

根据上述动态组播组构造方案可以构造随机的动态组播组。图1所示的是在一个具有200个节点、节点的平均度为3.876的网络中, 当 $\eta = 20\%$ 时, 某组播组中的组播节点数与更新次数之间的关系。可以看出, 虽然组播成员数目的变化是随机的, 但其平均数还是保持在40个节点左右, 即组播规模仍能维持在20%。显然, 本节给出的动态组播组构造方案可以满足实验要求。

为验证 MDLCMR 算法的性能, 我们对算法的消息复杂度、路由请求平均成功率、平均代价不精确度及算法的运行时间等几个方面进行了检验, 并将 MDLCMR 算法与 LDT、KPP、BSMA 算法作了比较。所有实验结果为多次实验所取的平均值。

4.2 路由请求平均成功率

路由请求的平均成功率 θ_{req} 定义如下:

• 98 •

$$\theta_{req} = N_{ack} / N_{req}$$

其中, N_{ack} —路由请求成功的平均数, N_{req} —总共产路由请求的平均数。图2给出了当成功的路由请求被定义为算法构成的组播树满足延时及带宽约束, 且 D_{max} 固定为180ms 时, BSMA、KPP、LDT 算法和 MDLCMR 算法路由请求平均成功率的实验结果。可以看出, MDLCMR 算法的路由请求平均成功率要远高于其它算法, 且随网络中组播通信组的个数的增加, MDLCMR 算法的优势愈加明显。这是因为, 其它算法没有考虑带宽约束条件, 使得多个通信组共享少数链路造成这些链路过于拥挤, 从而导致了路由请求失败。这证明了 MDLCMR 算法不仅具有极高的路由发现能力, 而且可适用于解决在网络中存在多个组播通信组时的路由选择问题, 有效地提高了网络的利用率。

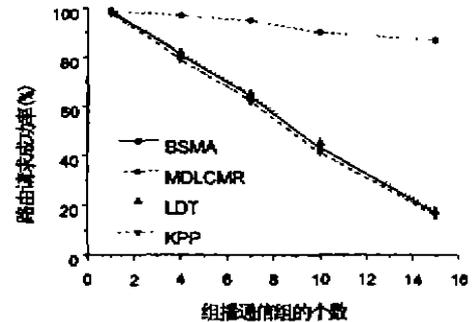


图2 各算法的成功率随组播业务比较

4.3 平均消息复杂度

图3显示了网络节点数为100时, MDLCMR 算法的平均消息数随组播规模变化的曲线。可以看出, 组播规模的增大并不会导致 MDLCMR 算法的平均消息数恶性膨胀。这表明 MDLCMR 算法不会产生消息风暴, 可以在大型网络中应用。

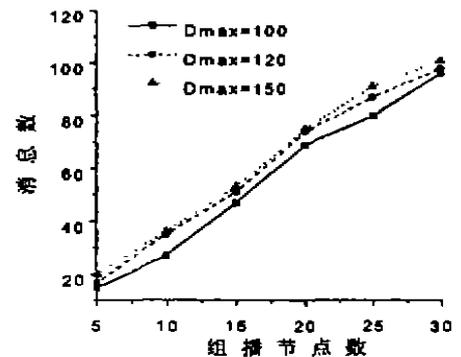


图3 MDLCMR 平均消息变化曲线

4.4 平均代价不精确度

由于 BSMA 被认为是优化性能最好的算法^[9], 因

此该项仿真实验以 BSMA 算法为参考。为便于讨论，我们定义了一个比较因子——平均代价不精确度。对于满足延时约束的组播树 T，代价不精确定义为：

$$\varphi(p) = (cost(T) - cost(BSMA)) / cost(BSMA)$$

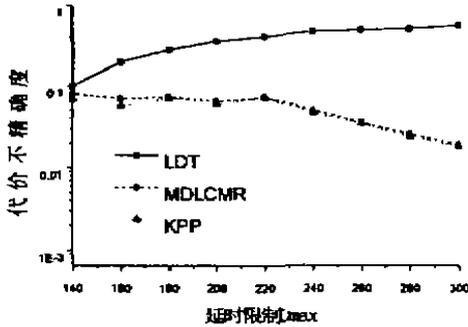


图4 各种算法的代价不精确度比较

实验比较了 MDLCMR 算法与 KPP 及 LDT 算法的平均代价不精确度。在节点数一定情况下，对于特定延时约束，每种算法重复 15 次，取代价计算平均值，按上述公式计算平均代价不精确度。图 4 给出了当网络节点数等于 150 时的情况。可以看出，随着延时条件的增大，LDT 算法平均代价不精确度一直增大；而 MDLCMR 算法的平均代价不精确度与 BSMA 和 KPP 算法相差不大，总能保持在 10% 之内，证明 MDLCMR 算法精确性很高。

4.5 算法运行时间

图 5 示出了 BDCT、BSMA、KPP 算法和 LDT 算法在各种网络规模情况下的运行时间。可以看出，BDCT 算法的计算时间与 LDT 算法差别不大，明显低于 BSMA 及 KPP 算法的求解时间，且随着网络节点的增多，BDCT 算法的优势更加明显。这说明在大规模网络的组播路由树求解中，从运算时间上考虑本算法比 BSMA 及 KPP 算法优越很多。

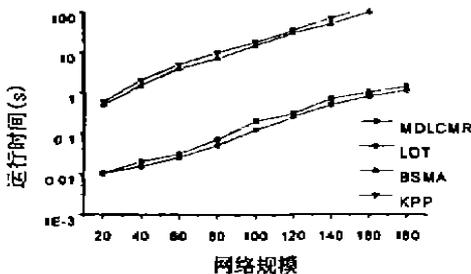


图5 各种算法的运行时间比较

4.6 算法动态性能

图 6 显示了在图 1 所示情况下 MDLCMR 路由树的代价随信宿更新而变化的情况。由图 6 可见，虽然 MDLCMR 路由树的代价随信宿更新次数的增加有所增大，但正如虚线所示，其增幅较小且增长缓慢，完全适合于解决动态组播路由问题。

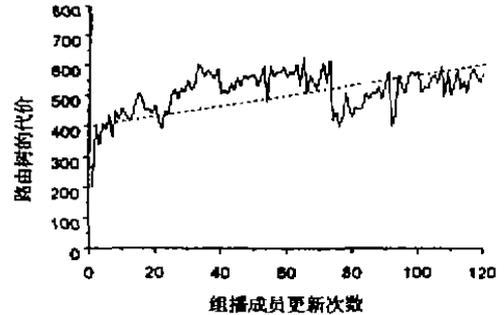


图6 MDLCMR 树的代价随信宿更新变化情况

结论 基于延时及带宽受限的最小代价组播路由问题是典型的优化问题，数学上也称多约束 Steiner 树问题，此类问题属 NP 完全问题^[2,9]。本文提出的 MDLCMR 算法，通过节点间的消息传递，巧妙地计算出了此 NP 完全问题的近似解。仿真实验表明 MDLCMR 算法是一种简单、快速、精确、健壮动态启发式算法，该算法的思想也容易扩展到其它具有 NP 难度的 QoS 路由问题。

参考文献

- 1 Kempella V, Pasquale J, Polyzos G. Multicast Routing for Multimedia Communication. *IEEE/ACM Transactions on Networking*, 1993, 1(3): 286~293
- 2 Zhu Q, Parsa M, Garcia-Luna-Aceves J. A Source-Based Algorithm for Delay-Constrained Minimum-Cost Multicasting. In: *Proc. of IEEE INFOCOM'95*, 1995, 377~385
- 3 Rom C R, Shavitt Y. Multi-Path Routing Combined with Resource Reservation. *IEEE INFOCOM'97*, Japan, April 1997 92~100
- 4 Deering S, et al. Protocol Independent Multicast-Sparse Mode (PIM-SM); Protocol Specification. Internet Draft, June 1996
- 5 Deering S, Estrin D, Farinacci D, et al. The PIM Architecture for Wide-Area Multicast Routing. *IEEE/ACM Transactions on Networking*, 1996, 4(2): 153~162
- 6 Kadane J, Knight G. Comparison of Dynamic Multicast Routing Algorithms for Wide-Area Packet Switched (Asynchronous Transfer Mode) Networks. *IEEE INFOCOM'95*, 1995, 212~219
- 7 Ballardie A, Francis P, Crowcroft J. Core Based Trees (CBT); an Architecture for Scalable Inter-Domain Multicast Routing. In: *Proc. of ACM SIGCOMM'93*, 1993, 85~95
- 8 Salama H F, Reeves D S, Viniotis Y. A Distributed Algorithm for Delay-Constrained Unicast Routing. *IEEE INFOCOM'97*, Japan, April 1997
- 9 Chen S, Nahrstedt K. On Finding Multi-Constrained Paths. *IEEE ICC98*, June 1998