

IP 网络多组播流相互影响均衡问题研究

赵 曦¹ 陈建阳²

(上海金融学院金融信息技术研究所 上海 201209)¹ (同济大学交通工程系 上海 200019)²

摘 要 自从 20 世纪 90 年代初 IP 组播概念及试验网络出现以后,IP 组播路由一直是互联网路由领域的研究热点,由于互联网规模的急速扩张带来的网络阻塞问题必须在新一代网络中得到较好解决。本文针对动态组播路由条件下多组播流相互影响问题展开研究,建立了多组播流优化均衡模型,并证明模型在总费用的最优解不但存在,而且唯一。简单的算例表明:相对静态组播路由,动态组播路由对网络性能有明显改善,为新一代网络和路由提供算法参考,并扩展到更加复杂网络的路由均衡研究。下一步将在典型的路由设备和具备一定规模的网络上进行验证和应用。

关键词 IP 组播,动态组播路由,网络均衡模型

Study on Equilibrium Problem of Multicast Flow Interacting in IP Network

ZHAO Xi¹ CHEN Jian-yan²

(Shanghai Finance University, Shanghai 201209, China)¹ (Department of Traffic Engineering, Tongji University, Shanghai 200019, China)²

Abstract IP multicast concepts and multicast test network have been proposed in since 90's, the multicast routing is a hot topic in internet routing area. This paper studies interactions of dynamic multicast routing in IP network, builds the interaction equilibrium model of multicast flows and also proves the solution existence and uniqueness of the proposed model. And finally uses a example to show the cost benefits of dynamic multicast routing. The equilibrium model can be extended in solving more complicated network routing problem and is to be experimented on a real network environment supported by common used routers.

Keywords IP multicast, Dynamic multicast routing, Network equilibrium model

1 引言

所谓组播,就是组播源发送数据包到多个接收者的通信,它可以是点到多点也可以是多点到多点的方式。视频点播、网络会议、计算机协同计算、文件发布等都属于这类应用。自从 20 世纪 90 年代初 IP 组播概念及试验网络出现以后,IP 组播路由一直是互联网路由领域的研究热点^[1-6]。组播路由问题有静态和动态之分。静态组播路由是指在组播通信开始之前确定组播成员并根据组播成员分布调用事先存储的路由方案。动态组播路由在组播通信的过程中,考虑组成员的随时加入和离开以及网络状态变化。目前,有关组播网络研究中,主要是静态组播路由算法,既不考虑链路费用随链路流量变化,也不考虑多个组播流的相互影响,或还没有相关网络流量均衡模型。本文重点考虑当多个组播流都采用动态组播路由时,多个组播路由如何相互影响?网络是否能达到均衡?

2 多组播流相互影响均衡模型

在不影响多个组播流都采用动态组播路由均衡最优解的存在和唯一研究的前提下,为了讨论方便,做以下几个基本假设:

- 1) 网络范围只涉及单个 ISP,不考虑边界网络路由。对于多 ISP 网络,通常各 ISP 采用不同的路由策略,需要使用不同的路由模型。
- 2) 路由优化时间间隔大于组播流分组的 RTT,可以认为不同时间间隔的网络行为不相互影响。主要考虑高负载时段的网络行为。
- 3) 当某个组播流流量较大时,路由系统可能采用多个组

播树传输组播流。

4) 所有组播流的组播路由优化目标是相同的,即组播路由广义费用最小。

5) 不考虑拥塞控制对组播路由的影响。

在大型多源组播网络中,如果每个组播源以某种组播路由机制如:最小费用树路由,传送分组(或数据包),网络中各链路的流量将发生变化,相应地链路费用随流量发生变化(如:平均时延),那么当各组播源流量已知,如何找组播路由树及相应的组播流量,以及最终的网络负载(每个链路的流量)。对于网络经营者或网络服务提供者来说,最好是路由的网络费用最小。由于同一组播源到目的结点可能有多个组播树,同一组播源的流量应尽可能分散到各个组播树上。假定每个组播源信息是完备的,即了解网络中所有链路的费用,且每个组播源只有一个组播流,当网络达到平衡时,同一组播源选择的组播树的费用都相等,且比未选择的组播树的费用小。用数学模型描述如下:

$$\mu_{D(r)} - c_k^{D(r)} = \begin{cases} 0, & \text{if } f_k^{D(r)} > 0 \\ \leq 0, & \text{if } f_k^{D(r)} = 0 \end{cases} \quad (1)$$

其中: $\mu_{D(r)}$ 为网络流量平衡时源点 r 到终点集合 $D(r)$ 的组播树的费用; $c_k^{D(r)}$ 为源点 r 到终点集合 $D(r)$ 的第 k 棵组播树的费用; $f_k^{D(r)}$ 为源点 r 到终点集合 $D(r)$ 的第 k 棵组播树上的流量; $D(r)$ 为源点 r 对应的终点集合。

组播树的费用是路由树上所有链路的费用之和,可以表示为

$$NC(R) = \sum_{e \in R} c(e)$$

赵 曦 副研究员,博士,研究方向为互联网应用、数据模型、信息系统;陈建阳 教授,博士,研究方向为网络优化系统工程。

R 表示路由树, e 是路由树上的链路, $c(e)$ 是 e 的费用。

模型(1)用来求解组播网络链路费用随链路流量变化时的网络流量比较困难, 为便于求解, 本文另行构造了最小费用树均衡模型, 并证明它的解与模型(1)是相同的。

在分析多源组播网络路由问题时, 通常将通信网络看成是赋权连通无向图 $G=(V, E)$, 其中 V 为节点集合, 每个节点代表主机或路由器, E 中的边 a 表示连接网络节点的通信链路, 边 a 的费用定义为随链路流量变化的函数 $t_a(x)$ 。可构造最优化极值模型如下:

$$\min Z(X) = \sum_a \int_0^{x_a} [t_a(\omega)] d\omega \quad (2a)$$

$$q_{rD(r)} = \sum_k f_k^{D(r)} \quad \forall r \quad (2b)$$

$$x_a = \sum_r \sum_k \delta_{ak}^{D(r)} f_k^{D(r)} \quad \forall a \quad (2c)$$

$$f_k^{D(r)} \geq 0 \quad \forall k, r \quad (2d)$$

其中: t_a 为链路 a 费用, 这里假设只与链路 a 上流量有关; x_a 为链路 a 上的流量; $q_{rD(r)}$ 为源点 r 到终点集合 $D(r)$ 的流量;

$$\delta_{ak}^{D(r)} = \begin{cases} 1, & \text{如果 } a \text{ 在 } r-D(r) \text{ 的第 } k \text{ 棵树上} \\ 0, & \text{其他} \end{cases}$$

3 最优化极值模型与原模型的等价性

由数学规划的知识可知, 模型(2)的拉格朗日函数为:

$$L(X, \mu) = Z(X) + \sum_r \mu_r (q_{rD(r)} - \sum_k f_k^{D(r)}) \quad (3)$$

式中: $\mu = (\dots, \mu_r, \dots)$ 。

模型(2)的一阶条件等价于使拉格朗日函数式(3)极小的一阶条件, 同时要满足约束条件(2d), 即:

$$\min L(X, \mu)$$

$$\text{s. t. } f_k^{D(r)} \geq 0$$

模型(2)对应的一阶条件为:

$$\frac{\partial L(\cdot)}{\partial f_k^{D(r)}} \cdot f_k^{D(r)} = 0 \quad \frac{\partial L(\cdot)}{\partial f_k^{D(r)}} \geq 0 \quad \frac{\partial L(\cdot)}{\partial \mu_r} = 0 \quad \forall r, k$$

$$f_k^{D(r)} \geq 0$$

$$\text{由于 } \frac{\partial Z(X)}{\partial x_a} = t_a(x_a), \quad \frac{\partial x_a}{\partial f_k^{D(r)}} = \delta_{ak}^{D(r)}$$

对函数式(3)两边分别求关于 $f_k^{D(r)}, \mu_{D(r)}$ 的偏导数, 得:

$$\begin{aligned} \frac{\partial L(\cdot)}{\partial f_k^{D(r)}} &= \sum_a \frac{\partial Z(\cdot)}{\partial x_a} \frac{\partial x_a}{\partial f_k^{D(r)}} - \mu_{D(r)} \\ &= \sum_a \delta_{ak}^{D(r)} t_a(x_a) - \mu_{D(r)} \\ &= c_k^{D(r)} - \mu_{D(r)} \end{aligned}$$

$$\frac{\partial L(\cdot)}{\partial \mu_{D(r)}} = q_{D(r)} - \sum_k f_k^{D(r)}$$

因此, 模型(2)对应的一阶条件为:

$$f_k^{D(r)} (c_k^{D(r)} - \mu_{D(r)}) = 0 \quad \forall r, k$$

$$c_k^{D(r)} - \mu_{D(r)} \geq 0 \quad \forall r, k$$

$$q_{D(r)} = \sum_k f_k^{D(r)} \quad \forall r, k$$

$$f_k^{D(r)} \geq 0 \quad \forall r, k$$

可知最优化极值模型(2)的最优解与模型(1)是相同的。

4 最优解的唯一性

模型(2)的最优解不但存在, 而且唯一, 下面进行证明。

从组播源点选择最短路径到第一个目的地, 然后从第一个目的地分别选择到其余目的地的最短路径, 将流量分配到这些边上, 得到模型(1)的一个可行解, 这说明模型的最优解存在。

函数 $Z(X)$ 关于 x_a, x_b 的二阶偏导数为:

$$\frac{\partial^2 Z(\cdot)}{\partial x_a \partial x_b} = \begin{cases} \frac{\partial t_a}{\partial x_a}, & x_a \text{ 为 } a \text{ 上的流量, } x_b = x_a \\ 0 & \text{其他} \end{cases}$$

$Z(X)$ 的 Hessian 矩阵为:

$$\nabla^2 Z(X) = \text{diag}(\frac{dt_1(x_1)}{dx_1}, \frac{dt_2(x_2)}{dx_2}, \dots, \frac{dt_A(x_A)}{dx_A})$$

一般链路费用函数是链路流量的增函数, 即

$$\frac{dt_a}{dx_a} > 0 \quad \forall a$$

因此 $Z(X)$ 的 Hessian 矩阵是正定的, 模型(2)的最优解是唯一存在的。

5 求解

1) 用最小费用树近似算法将流量分配到网络上, 求得初始可行解 $X^{(1)} = (\dots, x_a^{(1)}, \dots)$, 并令 $n=1$ 。

2) 在 $X^{(n)}$ 已知的条件下, 计算边 a 费用 $t_a^{(n)}$, 利用最小费用树近似算法分配方法重新进行流量分配, 得到新的流量分配结果 $Y^{(n)} = (\dots, y_a^{(n)}, \dots)$ 。由此确定 $X^{(n)}$ 处的目标函数下降方向 $Y^{(n)} - X^{(n)}$;

3) 寻找最优搜索步长

$$\text{令: } X_n(a) = X^{(n)} + \alpha(Y^{(n)} - X^{(n)}), (0 \leq \alpha \leq 1)$$

通过求解 $\min \sum_a \int_0^{x_a^{(a)}} [t_a(\omega)] d\omega$ 来确定最优搜索步长, 并记所求得的最佳步长为 α_n ;

4) 令 $X^{(n+1)} = X_n(\alpha_n)$, 如果 $|X^{(n+1)} - X^{(n)}| \leq \epsilon$, 停止计算; 否则, 令 $n=n+1$, 转向步骤 2 重复进行。

6 算例

在如图 1 所示的组播网络中, 有 7 个节点和 22 个链路, 各链路的费用函数为:

$$f_i = a_i + \frac{b_i}{c_i - x} \quad (4)$$

其中第一项表示链路传输费用, 第二项表示链路平均时延费用, 表 1 是各链路费用函数参数值。表 2 列出各组播流相关参数。

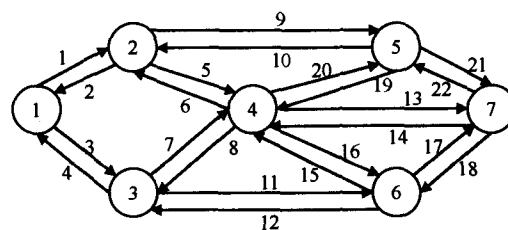


图 1 组播网络拓扑结构

表 1 组播网络各链路费用函数参数表

链路编号	a_i	b_i	c_i	链路编号	a_i	b_i	c_i
1	5	40	20	12	5	45	25
2	5	40	20	13	5	50	20
3	4	40	20	14	5	50	20
4	4	40	20	15	4	50	20
5	4	50	30	16	4	50	20
6	4	50	30	17	7	60	30
7	5	50	30	18	7	60	30
8	5	50	30	19	5	60	30
9	6	45	25	20	5	60	30
10	6	45	25	21	5	60	30
11	5	45	25	22	5	60	30

表2 组播流参数表

组播流编号	源结点	目标结点	流量
1	1	2,4,5,6,7	7
2	1	3,4,5,6,7	7
3	7	1,2,3,5	6
4	7	1,2,3,6	6
5	2	4,5,6,7	7

表3 静态(互不影响)时的最小费用树

组播编号	最小费用树	费用
1		38.0
2		37.0
3		30.6
4		29.6
5		29.9

表4 静态组播路由分配所有组播流得到的总费用

链路	a_i	b_i	c_i	组播流 1	组播流 2	组播流 3	组播流 4	组播流 5	x_i	$x_i f_i$
1	5	40	20	7			6		13	139.3
2	5	40	20			6			6	47.1
3	4	40	20		7	6			13	126.3
4	4	40	20				6		6	41.1
5	4	50	30	7				7	14	99.8
6	4	50	30						0	0.0
7	5	50	30		7				7	50.2
8	5	50	30						0	0.0
9	6	45	25	7				7	14	141.3
10	6	45	25			6			6	50.2
11	5	45	25		7				7	52.5
12	5	45	25				6		6	44.2
13	5	50	20						0	0.0
14	5	50	20						0	0.0
15	4	50	20						0	0.0
16	4	50	20						0	0.0
17	7	60	30						0	0.0
18	7	60	30						0	0.0
19	5	60	30	7				7	14	122.5
20	5	60	30		7	6			13	110.9
21	5	60	30		7				7	53.3
22	5	60	30	7			6	7	20	220.0

表5 相互影响时的最小费用树

编号	最小费用树	费用
1		39.6
2		37.0
3		30.6
4		29.6
5		32.5

加载所有流的总费用 $\sum_{i=1}^{22} x_i f_i = 1221.0$ 。

结束语 自从20世纪90年代初IP组播概念及试验网络出现以后,IP组播路由一直是互联网路由领域的研究热点,而且随着互联网组播应用的普及和深入,这一研究主题还将升温。本文针对动态组播路由条件下多组播流相互影响问题展开研究,建立了多组播流优化模型,并证明多组播流网络用户最优均衡状态是存在的,且是唯一的,其含义是如果网络满足以下条件时:

- 1)所有路由器的优化目标、优化策略相同;
- 2)路由器能准确了解或测量网络的相关状态参数。

那么IP组播路由系统能使整个网络达到多组播流(用户)最优均衡状态。简单的算例表明,相对静态组播路由,动态组播路由对网络性能有明显改善。本文所建立的路由优化网络模型及研究方法可进一步拓展到复杂路由问题的解析研究。

参考文献

- [1] Kwon M, Fahmy S. Topology-Aware overlay networks for group communication. 2006. <http://www.cs.rit.edu/~jmk/papers/pub.html>
- [2] 胡光岷,李乐民,安红. 动态多播最小生成树算法. 电子与信息学报, 2003, 25(1): 88-93
- [3] Jain S, Mahajan R, Wetherall D, et al. Scalable self-organizing overlay. Technical Report. Washington University, 2000
- [4] Zhang B, Jamin S, Zhang L. Host multicast: A framework for delivering multicast to end users. 2006. <http://www.peer-to-peer.info/bibliography/zhang2002hmt>
- [5] Pendarakis D, Shi S, Verma D, et al. ALMI: An application level multicast infrastructure. 2006. <http://citeseer.ist.psu.edu/pendarakis00almi.html>
- [6] Mathy L, Canonico R, Hutchison D. An overlay tree building control protocol. 2006. <http://portal.acm.org/citation.cfm?id=747484&dl=ACM&coll=&CFID=15151515&CFTOKEN=6184618>
- [7] Banerjee S, Bhattacharjee B, Kommareddy C. Scalable application layer multicast. 2006. <http://portal.acm.org/citation.cfm?coll=GUIDE&dl=GUIDE&id=633045>