

文章编号:1001-9081(2008)08-1951-03

# 用 column generation 算法规划网络编码业务

宋运吉, 王 晟, 王 雄

(电子科技大学 宽带光纤传输与通信网络技术教育部重点实验室, 成都 610054)

(songyunji@uestc.edu.cn)

**摘要:**网络编码能够有效降低网络中关键边的资源消耗,改善网络的负载均衡。但是普通的启发式路由算法通常只能为单个业务寻找最优路由,无法优化网络的整体性能。运用 column generation 算法对网络编码业务进行规划,为松弛系数赋予具体的物理含义,并据此进行路径更新,有针对性地为每个业务寻找路由。与启发式算法相比, column generation 从整体上提高了网络的吞吐量,改善了网络的负载均衡。同时,与普通 ILP 算法相比, column generation 算法无需计算大量备选路径,且函数始终处于收敛状态,不会产生振荡,求解总时间缩短了 23.5%,总代价优化 2.5%。

**关键词:**网络编码;整数线性规划;网络吞吐量

**中图分类号:** TP393.02 **文献标志码:** A

## Network coding traffic design using column generation techniques

SONG Yun-ji, WANG Sheng, WANG Xiong

(Key Laboratory of Broadband Optical Fiber Transmission and Communication Networks,  
University of Electronic Science Technology of China, Chengdu Sichuan 610054, China)

**Abstract:** Network coding can reduce the cost of key links and improve the load balance of the network. But heuristic routing algorithms cannot find the global optimal solution for the network. We used column generation to solve the Network Coding based traffic programming problem. First, we relaxed the constraints of the problem by using Lagrange Relaxation, then defined the physical meaning of each relaxation variable. At last, used relaxation variables updating the multicast graph in iteration. Compared with the heuristic routing algorithm, column generation can improve the network throughput, compared with Integer Linear Programming (ILP). It can compute less inventory routings and accelerate convergence.

**Key words:** network coding; Integer Linear Programming (ILP); network throughput

### 0 引言

对于单播业务,依靠路由算法就能够获得其最大流。但是对于多播业务,中间节点仅仅对数据进行路由、转发的方式很难获得网络多播的最大流。2000年,R. Ahlswede等提出了网络编码的基本概念<sup>[1]</sup>,其核心的思想是节点对收到的数据进行编码,克服传统的数据传输模式,为实现网络多播最大流奠定了基础。

引用经典的蝴蝶模型来说明线性网络编码<sup>[2]</sup>。图1(a)中标出了图中每条链路的容量,源节点 $S$ 和 $t_1, t_2$ 之间的最大流分别是2。从源节点 $S$ 向 $t_1, t_2$ 多播两个比特 $b_1$ 和 $b_2$ ,传统的多播算法只能建立一棵多播树,无法同时传输 $b_1, b_2$ 。当引入网络编码后,这个问题就能得到解决。如图1(b)所示,源节点 $S$ 分别向链路 $(S, 1)$ 和 $(S, 2)$ 发送数据 $b_1$ 和 $b_2$ ;节点1,2分别对数据 $b_1$ 和 $b_2$ 进行复制转发;当节点3收到 $b_1, b_2$ 之后,使用异或运算 $b_1 \oplus b_2$ 对数据进行编码,再将数据分别传输至 $t_1, t_2$ ;节点 $t_1$ 在收到 $b_1$ 和 $b_1 \oplus b_2$ 之后,通过 $b_2 = b_1 + (b_1 \oplus b_2)$ 解码计算出 $b_2$ ,同理,节点 $t_2$ 也能解码出 $b_1$ 和 $b_2$ 。

与传统多播路由不同,网络编码的路由算法需要为业务建立多播图。文献[3,4]中指出,网络编码同时传输 $k$ 单位的数据的充要条件是多播图中每个目的节点与源节点之间存在 $k$ 条链路分离的路径。也就是说,在有向图 $G(V, E)$ 中,若网

络编码路由的度为 $k$ ,则所有节点的度必须大于等于 $k$ ,且节点两两之间的最小割也必大于等于 $k$ 。但目前大多数仿真用拓扑中,节点的最小度都小于或等于3,因此,在目前的研究中,启发式路由算法通常取 $k = 2$ 。启发式算法的一般做法是在有向图 $G(V, E)$ 中,分别为多播业务的源节点与每个目的节点分别找一条链路分离的路径对,而当不同的路径对在同一条链路上重合时,在此链路的入节点处对来自不同路径对的信息进行编码后传输。

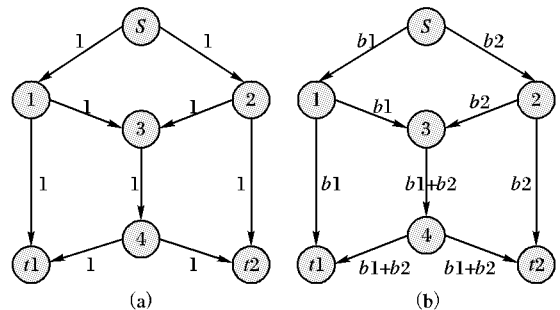


图1 网络编码的蝴蝶模型

网络编码的最主要目的就是实现多播的最大流传输,改善网络的吞吐量<sup>[5-7]</sup>。但是,在文献[5]中指出,网络编码与普通多播方式相比,传输单位数据所消耗的资源相差不大。这是因为网络编码虽然能通过编码降低关键边的资源消耗,

收稿日期:2008-02-28;修回日期:2008-04-10。

基金项目:国家973计划项目(2007CB307100);国家自然科学基金资助项目(60472008);四川省青年科技基金项目(05ZQ026-002)。

作者简介:宋运吉(1983-),男,四川成都人,硕士研究生,主要研究方向:WDM网络多播算法及机制;王晟(1971-),男,四川成都人,教授,博士生导师,主要研究方向:光纤宽带网传输技术;王雄(1980-),男,四川攀枝花人,博士研究生,主要研究方向:光网络多播算法。

但是同多播树相比,多播图能支持传输更多数据,同时其消耗的网络资源也更多。通过仿真计算发现,其传输单位数据消耗的资源与普通多播相若。文献[5]同时指出,网络编码的主要优势在于减少图中关键边的资源消耗,使得网络负载均衡,从而提高网络的吞吐量。

传统的启发式路由算法只能为单个业务寻找最小代价路由,而从整体上却无助于改善网络性能。普通的整数线形规划(ILP)方法虽然能够解决这个问题,但是其问题规模和求解时间都偏大。column generation<sup>[8-9]</sup>是一种基于问题分解的方法,其原理是用拉格朗日松弛约束条件后,通过分析最优解,为松弛系数之间建立关系,使得各个系数之间不再相互独立。再为松弛系数赋予具体的物理含义,并根据松弛系数的物理含义,在每次迭代过程中有针对性地更新路由,减小了问题规模。相比于 ILP,其求解速度更快,且能够解决部分 ILP 无法处理的大规模规划问题。

## 1 用 column generation 求解

### 1.1 模型描述

假设有向图  $G(V, E)$  中,存在  $D$  组多播业务,每组业务包含了不同数量的目的节点,将网络编码业务表示为优化模型,我们首先定义如下符号和变量。

$e = 1, 2, \dots, E$ : 常量,表示图中的有向边。

$d = 1, 2, \dots, D$ : 常量,表示业务。

$c_e$ : 常量,有向边  $e$  的代价。

$u_e$ : 常量,有向边  $e$  的容量。

$\delta_{dnp}$ : 布尔变量,当其值为 1 时,表示有向边  $e$  被业务  $d$  的第  $n$  个目的节点的第  $p$  条路径对所占占用;否则其值为 0。

$W_{de}$ : 布尔变量,当业务  $d$  占用了有向边  $e$  时,其值为 1;否则其值为 0。

$\theta_{dnp}$ : 布尔变量,当其值为 1 时,表示业务  $d$  的第  $n$  个目的节点使用第  $p$  条路径对进行数据传输;否则其值为 0。

在以上定义中,我们将一个多播业务按目的节点数量分解为多个单播业务,每个单播业务存在  $p$  对备选路径,但是其只能选择其中某一条路径对传输数据,而多播图则可看作是由此业务的  $n$  组目的节点与源节点的路径对重叠所得。这样定义,可以使得多播图的描述更加灵活,不必因为多播图中某一段的改动而重新计算整个多播图,减小了计算量。由此,我们的优化模型描述为:

$$\min \sum_d \sum_e W_{de} \cdot c_e \quad (1)$$

Sub to:

$$\sum_p \theta_{dnp} = 1 \quad (2)$$

$$\sum_d W_{de} \leq u_e \quad (3)$$

$$\sum_n \sum_p \theta_{dnp} \delta_{dnp} \leq L W_{de} \quad (4)$$

$$W_{de} \leq \sum_n \sum_p \theta_{dnp} \delta_{dnp} \quad (5)$$

在上述优化模型中,优化目标(1)描述了网络中所有业务消耗的总代价最小;约束(2)表示每对源宿节点对只能从  $p$  条备选路径中择一条为其路由;约束(3)为每条链路上的容量约束;约束(4)中  $L$  为任意大的整数,约束(4)、(5)合起来约束了变量  $W_{de}$  与  $\theta_{dnp}$  之间的关系,表示了多播图是由此业务的  $n$  条目的节点与源节点的路径对重叠所得。

### 1.2 松弛变换

用 column generation 算法为业务计算路由,必须先使用

拉格朗日松弛改变优化目标和约束条件,使得各个松弛系数之间存在固定的关系,再为松弛系数赋予合理的物理含义。首先,需要对原式进行变换。在 1.1 节所描述的模型中,因为优化目标的存在,所以式(5)对结果没有影响,可以省略。而在式(4)中,  $L$  为一个大的整数,若令  $L_{de} = \sum_n \sum_p \theta_{dnp} \delta_{dnp}$ , 那么式(4)将变为:

$$\sum_n \sum_p \theta_{dnp} \delta_{dnp} = L_{de} W_{de} \quad (6)$$

用式(6)代替式(4),再对约束(2)、(3)、(6)进行拉格朗日松弛之后,优化目标将发生改变,问题变为:

$$H(W, \theta, \lambda, \pi, \varphi) = \min_{W, \theta} \max_{\lambda, \pi, \varphi} \sum_d \sum_e W_{de} \cdot c_e + \sum_d \sum_e \pi_{de} \left( \sum_n \sum_p \theta_{dnp} \delta_{dnp} - L_{de} W_{de} \right) + \sum_d \sum_n \lambda_{dn} \left( 1 - \sum_p \theta_{dnp} \right) + \sum_e \varphi_e \left( \sum_d W_{de} - u_e \right) \quad (7)$$

Sub to:

$$\pi_{de} \geq 0 \quad (8)$$

$$\lambda_{dn} \geq 0 \quad (9)$$

$$\varphi_e \geq 0 \quad (10)$$

将式(7)中所有包含变量的项进行合并,将可以得到式(11)。

$$\min_{W, \theta} \max_{\lambda, \pi, \varphi} \sum_d \sum_e W_{de} (c_e + \varphi_e - \pi_{de} L_{de}) + \sum_d \sum_n \sum_p \theta_{dnp} \left( \sum_e \pi_{de} \delta_{dnp} - \lambda_{dn} \right) + \sum_d \sum_n \lambda_{dn} - \sum_e \varphi_e u_e \quad (11)$$

若假设  $H(W^*, \theta^*, \lambda^*, \pi^*, \varphi^*)$  为松弛变换后优化问题的最优解,那么根据松弛问题最优解的性质,  $H(W^*, \theta^*, \lambda^*, \pi^*, \varphi^*)$  为鞍点,当  $\lambda^*, \pi^*, \varphi^*$  为固定值时,任意的  $W, \theta$  都必然使函数上升;反之,  $W^*, \theta^*$  固定时,任意的  $\lambda, \pi, \varphi$  都必然使函数上升。即:

$$H(W, \theta, \lambda^*, \pi^*, \varphi^*) \geq H(W^*, \theta^*, \lambda^*, \pi^*, \varphi^*) \geq H(W^*, \theta^*, \lambda, \pi, \varphi) \quad (12)$$

同时,由于最优解必定是在可行域之中,我们可推断出最优解具有以下性质:

当  $\sum_d W_{de}^* < u_e^*$  时,  $\varphi_e^* = 0$ ; 当时  $\sum_d W_{de}^* > u_e^*$ ,  $\varphi_e^* = \infty$ ;  
当  $\sum_d W_{de}^* = u_e^*$  时,  $\varphi_e^* \geq 0$ ;

当  $\sum_n \sum_p \theta_{dnp}^* \delta_{dnp}^* < L_{de}^* W_{de}^*$  时,  $\pi_{de}^* = 0$ ; 当  $\sum_n \sum_p \theta_{dnp}^* \delta_{dnp}^* > L_{de}^* W_{de}^*$  时,  $\pi_{de}^* = \infty$ ; 当  $\sum_n \sum_p \theta_{dnp}^* \delta_{dnp}^* = L_{de}^* W_{de}^*$  时,  $\pi_{de}^* \geq 0$ ;

当  $c_e + \varphi_e^* - \pi_{de}^* L_{de}^* > 0$  时,  $W_{de}^* = 0$ ; 当  $c_e + \varphi_e^* - \pi_{de}^* L_{de}^* \leq 0$  时,  $W_{de}^* = 1$ ;

当  $\sum_e \pi_{de}^* \delta_{dnp}^* > \lambda_{dn}^*$  时,  $\theta_{dnp}^* = 0$ ; 当  $\sum_e \pi_{de}^* \delta_{dnp}^* \leq \lambda_{dn}^*$  时,  $\theta_{dnp}^* = 1$ 。

根据以上的性质,我们可以为系数  $\pi_{de}, \lambda_{dn}, \varphi_e$  赋予具体的物理含义:

$\varphi_e$  为有向边  $e$  的超载系数,当边  $e$  上的业务超出其容量时,  $\varphi_e$  为  $+\infty$ , 表示其超载,否则  $\varphi_e$  为 0。

$\pi_{de}$  为有向边  $e$  在业务  $d$  中的权重,其初始值为  $c_e$ , 当业务  $d$  中有  $L_{de}$  条源宿路径对都需要占用有向边  $e$  时,其权重变为  $c_e/L_{de}$  ( $L_{de} > 0$ )。

$\lambda_{dn}$  表示业务  $d$  的目的节点  $n$  与源之间的最小权重路径

对。当  $\sum_e \pi_{de} \delta_{dnp} = \lambda_{dn}$  时,表示  $p$  为最小权重路径对,同时  $\theta_{dnp} = 1$ ,表示业务只占用最小权重的路径对。

将系数赋予实际的物理含义之后,算法在迭代过程中将具有两大优势。

1) 系数拥有物理含义之后,可以在每次迭代中通过公式计算出来,而不是如一般 ILP 算法一样通过步长和收敛方向估算所得。如此,函数不会因为系数步长过大而产生振荡,或者因系数步长过小而收敛缓慢。

2) 普通 ILP 求解,需先利用  $k$ -路由算法计算出  $k$  条备份路由。若计算的路由太少,求解的结果常不理想,而计算的路由过多,则会导致计算量过大,求解困难。在 column generation 算法中,可以根据系数  $\pi_{de}$  计算新路由,进行有针对性的业务路径更新,尽量避免关键边上过载的问题。不仅计算量减小,而且更容易得到最优解。

### 1.3 迭代求解

1) 将多播业务分解为多个单播业务,并为所有源宿节点对计算一条路径对,再根据路径信息初始化变量  $\theta_{dnp}$  和  $W_{de}$ 。

2) 将  $\theta_{dnp}$  和  $W_{de}$  作为常量代入  $\max_{\lambda, \pi, \varphi} H(W, \theta, \lambda, \pi, \varphi)$ , 求解出系数  $\pi_{de}, \lambda_{dn}, \varphi_e$  的值。

3) 根据权重  $\pi_{de}$  对所有源宿节点对进行路径更新,利用最优路径对算法找出一条权重最小的路径对,若此路径对的权重小于原来所有的路径对,则保留此路径对,将其添加进  $\delta_{dnp}$  中,否则就舍弃此路径对。

4) 将之前求解出的系数  $\pi_{de}, \lambda_{dn}, \varphi_e$  的值作为常量代入  $\min_{W, \theta} H(W, \theta, \lambda, \pi, \varphi)$ , 求解  $W$  和  $\theta$  值。

5) 重复 2~4 步,直到无路径更新,且目标函数的上下限小于设定门限,或者迭代次数达到上限后自然结束迭代过程。

## 2 仿真及结果分析

我们在 C++ 下搭建仿真平台,多播业务均通过随机生成,然后分别于 NSFNET(14 节点,21 条链路)和 AMERICAN NET(24 节点,43 条链路)两个拓扑上进行仿真,假设拓扑中每条链路的带宽为 100 Mbps,对比了启发式算法、一般 ILP 规划(用 CPLEX 软件进行计算求解)和 column generation 算法的性能。结果图 2 所示。

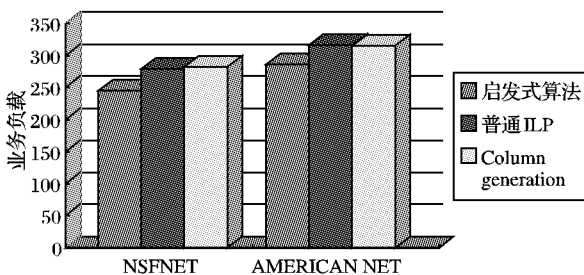


图 2 三种算法下网络的最大业务负载

图 2 是三种不同算法下,网络的最大负载。通过柱形图可以看出,启发式算法对于网络资源的利用率最低,而 ILP 和 column generation 算法则较为接近,均能够提升网络负载 10% 左右。

图 3 是 ILP 规划和 column generation 算法在 NFSNET 中,业务量为 270 时的收敛曲线对比。从图中可以看出,ILP 算法初期一直无法计算出可行解,可以看作函数上限为  $\infty$ ,直到 10s 左右才开始收敛,并逐步接近最优解;column generation 算法的函数始终保持收敛状态,函数曲线一直呈下降趋势。

我们假设当目标函数的上下限之间的 gap < 0.2% 时,函数得到近似最优解,停止迭代,输出结果。那么,普通 ILP 算法的求解时间为 15.3 s,而 column generation 算法求解耗时 11.7 s;并且,普通 ILP 算法求得的最优解为 18 187, column generation 算法的最优解为 17 733,在求解时间和算法性能上, column generation 算法均优于普通 ILP 算法。

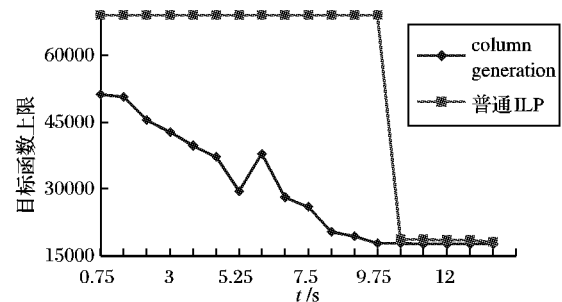


图 3 普通 ILP 与 column generation 算法的收敛曲线对比

## 3 结语

本文将 column generation 算法应用在多播环境中,将多播业务描述为多个单播业务的叠加,使得多播图的描述更加灵活,不必因为多播图中某一段的改动而重新计算整个多播图,减小了业务路由更新的计算量。同时,通过拉格朗日松弛,为松弛系数之间建立了固定的联系,使得各个松弛系数之间不再是相互独立的关系,在迭代过程中,确定其中之一就可以计算出其余的松弛系数。最后,为松弛系数赋予了具体的物理含义,使更新路由有据可依,更加有针对性。仿真数据表明, column generation 算法能有效提高网络的吞吐量,而且迭代时间短,具有较好的性能,在多播规划中还有广阔的发展空间。

### 参考文献:

- [1] AHLWEDE R, CAI N, LI S Y R, *et al.* Network information flow [J]. IEEE Transactions on Inform Theory, 2000, 46(4): 1204 - 1216.
- [2] LI S Y R, YEUNG R W, CAI N. Linear network coding [J]. IEEE Transactions on Information Theory, 2003, 49(2): 371 - 381.
- [3] WU Y. Network coding for multicasting [D]. Princeton: Princeton University, 2005.
- [4] HO T, KARGER D R, MEDARD M, *et al.* Network coding from a network flow perspective [C]// Proceeding of IEEE International Symposium on Information Theory (ISIT). Yokohama, Japan: IEEE Press, 2003: 441.
- [5] WU Y, CHOU P A, JAIN K. A comparison of network coding and tree packing [C]// Proceeding of IEEE International Symposium on Information Theory (ISIT). Chicago, USA: IEEE Press, 2004: 143.
- [6] NOGUCHI T, MATSUDA T, YAMAMOTO M. Performance evaluation of new multicast architecture with network coding [J]. IEICE Transaction Communication, 2003, 86(6): 1788 - 1795.
- [7] AGARWAL A, CHARIKAR M. On the advantage of network coding for improving network [C]// 2004 IEEE Information Theory Workshop. San Antonio: IEEE, 2004: 247 - 249.
- [8] PIORO M, MEDHI D. Routing, flow, and capacity design in communication and computer networks [M]. San Fransisco: Morgan Kaufmann Publishers, 2004: 184 - 192.
- [9] KRITHIKAIVASAN B, SRIVASTAVA S, PIORO M, *et al.* Backup path restoration design using path generation technique [C]// Proceeding of 4th International Workshop on Design of Reliable Communication Networks (DRCN). Banff, Canada: [s. n], 2003: 77 - 84.