

大规模服务覆盖网拓扑设计

张 栋^① 吴春明^① 姜 明^② 熊 伟^①

^①(浙江大学人工智能研究所 杭州 310027)

^②(杭州电子科技大学计算机学院 杭州 310018)

摘 要: 针对大规模服务覆盖网拓扑设计方法构建效率低的问题, 该文提出了以带宽容量为约束条件, 基于多商品流的线性规划模型和实现算法, 以降低构建算法的时间与空间复杂度。仿真结果表明, 该算法在大规模网络情况下, 构建效率和资源利用率有显著提高。

关键词: 服务覆盖网; 拓扑设计; 多商品流

中图分类号: TP393.01

文献标识码: A

文章编号: 1009-5896(2010)04-0841-05

DOI: 10.3724/SP.J.1146.2009.00481

Topology Design for Large Scale Service Overlay Networks

Zhang Dong^① Wu Chun-ming^① Jiang Ming^② Xiong Wei^①

^①(AI Institute of Zhejiang University, Hangzhou 310027, China)

^②(Computer Science College of Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: Considering the problem of low efficiency on topology design for large scale Service Overlay Networks(SON), a linear programming model is proposed based on multi-commodity flow and algorithm with bandwidth capability constraints, which reduces the time complexity and the space complexity. The simulation results demonstrate that the proposed algorithm can improve the efficiency of constructing and resource usage.

Key words: Service Overlay Networks(SON); Topology design; Multi-commodity flow

1 引言

这些年来, Internet 网络上的服务和应用不断丰富和扩展。多媒体、网络游戏、视频会议等实时应用的发展对互联网提供具有 QoS 保证的要求越来越迫切。传统互联网基于“尽力而为”的服务模式, 难以适应当前业务多样化的发展趋势和对服务质量的要求。IETF 先后提出 IntServ/RSVP^[1] 和 DiffServ^[2] 的网络体系结构, 但要对互联网进行大规模的升级改造, 难以广泛应用和有效解决互联网网络传输的服务质量问题。同时, 互联网上存在着许多不同的自治域, 每个自治域管理自己的传输、性能和内部的路由机制, 这也给互联网提供端到端的 QoS 保证带来困难。

服务覆盖网(SON)是近年来提出的在“尽力而为”的互联网上改进服务质量的有效方式, 是一种能够穿透自治域的覆盖网, 同时提供端到端的 QoS 的保证, 为实时多媒体传输、VOIP 等对 QoS 要求较高的网络应用提供传输服务质量保证^[3,4]。

SON 是通过预留带宽的方式来提供具有 QoS

保证的网络服务。SON 服务提供商根据服务需求、终端用户的规模等设置节点服务器, 同时向 ISP(Internet 服务提供商)租用带宽保证。终端用户和 SON 节点通过提供带宽预留的 ISP 相连, 并通过向 SON 服务提供商付费来享受 Internet 增值业务。

SON 部署是一个资本密集型投资。节点部署, ISP 带宽预留租用等都需要庞大资金投入, 拓扑设计就成为 SON 构建的关键问题。选择良好的拓扑构建算法, 能够满足 SON 服务提供商的运营需求, 提供给终端用户具有端到端 QoS 保证的 Internet 增值服务, 同时考虑最优的拓扑结构、最少的构建费用和带宽预留的租用费用。

由于求解 SON 拓扑设计可归约为最小 Steiner Forest 问题, 而最小 Steiner Forest 是一个著名的 NP-hard 问题^[5], 因此 SON 拓扑设计是一个 NP-hard 问题^[6], 通常采用线性规划^[6,7]、启发式算法^[8,9]——模拟退火算法和贪心算法等近似和约束条件来求解。本文的目标是设计在最少费用情况下满足所有终端用户的流量需求, 特别是在大规模 SON 网络情况下的拓扑构建算法。

此前已有相关的研究关注于服务覆盖网的拓扑设计问题, 如文献[6-8,10]。文献[7]提出 Full-Coverage SON Design model(FCSD), 解决覆盖所

2009-04-07 收到, 2009-10-29 改回

国家自然科学基金(60773182)和国家 863 重大项目(2008AA01A323)资助课题

通信作者: 吴春明 wuchunming@zju.edu.cn

有终端用户情况下的网络费用最小化问题,同时提出 Revenue-Maximization SON Deign (RMSD) model 解决在考虑收益最大化的前提下,构建服务覆盖网拓扑的问题。文献[6]根据节点权重和边权重的费用模型(TDNEW),通过整数线性规划方法求解 SON 拓扑设计。文献[8]是解决在假设带宽保证下的 SON 节点和 ISP 之间连通的拓扑设计问题。文献[10]是在物理承载网架构中基于流量均衡实现虚拟网的拓扑设计。

文献[6]是在假设带宽没有容量限制情况下,采用整数线性规划的最优算法和选择最短路径方式的近似算法来进行拓扑构建,并得出最优算法的构建完成时间级在小时级别上。现实情况下,网络之间的带宽是有限的,随着网络构建的进行,可用带宽将随之减少。目前 SON 拓扑设计最优化模型构建耗时较长^[6,7],同时缺少带宽容量限制的约束条件,难以满足现实需要。本文针对上述问题,提出在带宽容量限制的约束下,基于多商品流模型的大规模 SON 拓扑设计的线性规划最优算法,同时降低算法的时间复杂度。

2 SON 拓扑设计模型

SON 的拓扑设计实质是在 ISP 提供的底层物理网络中构建一个子网,该子网能够提供指定节点间的数据传输服务。本文主要关注在 SON 节点已经确定的情况下如何在 SON 入口节点和出口节点之间寻找合适的路径,通过租用链路上 ISP 提供的带宽,构建覆盖网络,使构建费用最小化。

2.1 问题的基本数学模型

记底层 ISP 提供的物理网络为带权无向图 $G(V, E)$, V 为节点集合, E 为边集合。 $\forall v \in V$, v 表示网络中的节点; $\forall e \in E$, e 表示网络中的节点间的连接。边权函数 b 定义在 E 上, $\forall e \in E$, $b(e)$ 表示边 e 所对应连接能够承载的最大流量; 流量函数 x 定义在 E 上, $\forall e \in E$, $x(e)$ 表示边 e 上的流量大小。费用函数 f 定义在 E 上, $\forall e \in E$, $f(e)$ 表示向 ISP 在边 e 所表示的连接上购买单位流量带宽所需的费用; 费用函数 f' 定义在 V 上, $\forall v \in V$, $f'(v)$ 表示在 v 所对应的网络节点部署 SON 节点所需费用。SON 的拓扑设计则是在 $G(V, E)$ 中构建一个子图 $G'(V', E')$, 这个子图应当满足一组 SON 约束条件, 这组条件通常可以描述为

$$R = \{(s_1, t_1, d_1), (s_2, t_2, d_2), \dots, (s_l, t_l, d_l)\}$$

其中对 $1 \leq i \leq l$, (s_i, t_i, d_i) 表示节点 s_i 和节点 t_i 间所需带宽大小为 d_i 的网络传输能力。这里使用 $x(e)$ 的值来定义需要构建的子图 G' , 即

$$x(e) = \begin{cases} 0, & e \text{ 不在 } G' \text{ 中} \\ \alpha, & e \text{ 上的流量} \end{cases}, \quad \alpha \in (0, b(e)] \quad (1)$$

设计目标是通过费用函数度量的子图 G' 的费用尽量小。则构建 G' 的费用 C 为

$$C = \sum_{e \in E} f(e) \cdot x(e) + \sum_{v \in V'} f'(v) \quad (2)$$

假设 SON 构建需求中已经确定服务节点, 则构建需求的节点费用 c 是确定的, 即

$$c = \sum_{v \in V'} f'(v) = \rho_{G'(V', E')} \quad (3)$$

其中 $\rho_{G'(V', E')}$ 是关于 $G(V, E)$ 的子图 $G'(V', E')$ 的常量。于是构建 G' 的费用可简单的表示为

$$C = \sum_{e \in E} f(e) \cdot x(e) + \rho_{G'(V', E')} \quad (4)$$

SON 的拓扑设计目标是使得 SON 构建费用最小化, 即

$$\text{Minimize } \sum_{e \in E} f(e) \cdot x(e) + \sum_{v \in V'} f'(v) \quad (5)$$

式(4)中, $\rho_{G'(V', E')}$ 为常量, 故式(5)可简化表示为

$$\text{Minimize } \sum_{e \in E} f(e) \cdot x(e) \quad (6)$$

由定义, 式(6)中的 $f(e)$ 是与 ISP 所提供的物理网络相关而与构建无关的函数, 因此, SON 拓扑设计实质是在针对一个 SON 的构建需求 R , 求一个定义在边集 E 上的函数 x , 满足式(6)的最小化需求, 同时满足 SON 构建需求 R , 即对于函数 x , 应当满足如下条件:

$\forall (s_i, t_i, d_i) \in R$, 存在 $G'(V', E')$ 的子图 $G'_i(V'_i, E'_i)$, 在这个子图中, 从 s_i 到 t_i 的最大流满足:

$$f_{\max}(G'_i(V'_i, E'_i), s_i, t_i) \geq d_i \quad (7)$$

对于式(6), 式(7)所限定的 x 进行求解, 朴素的方法是将 $x(e)$ 的值离散化, 即将式(1)定义的 $x(e)$ 的取值由 $[0, b(e)]$ 转变为

$$x(e) \in N \cap [0, b(e)] \quad (8)$$

式(8)为自然数集。于是 x 所属解空间即

$$X = \{x \mid \forall e \in E, x(e) \in N \cap [0, b(e)]\} \quad (9)$$

其中整数区间 $N \cap [0, b(e)]$ 可根据实际情况增大增值跨度, 如采用 100 作为增量值, 即 $x(e)$ 的取值范围变为

$$U = \{0, 100, \dots, 100 \cdot k, \dots, b(e)\} \quad (10)$$

于是, 对 x 的求解就转变为搜索解空间 X , 寻找满足由式(7)约束条件的一个函数 x 。但采用目前已知的各种求解方法, 对 x 的搜索都将是规模巨大的, 其解空间的规模有: $\prod_{e \in E} \frac{[b(e)]}{\kappa}$ (其中 $[x]$ 是高斯函

数, κ 是区间增量), 例如, 考虑一个 100 条连接, 边容量约为 100 Mbps, 区间增量为 1 Mbps, 则实际的解空间规模就有:

$$\prod_{i=1}^{100} \frac{100MB}{1MB} = 100^{100} = 10^{10000}$$

2.2 基于线性规划的 SON 拓扑设计模型

考虑 SON 构建需求 R 中一对节点的约束 (s_i, t_i, d_i) , 对于该约束需要满足式(7)所指定的条件, 将整个 R 中的约束统一起来, 实际上是一个 Multi-Commodity Flow(多商品流)问题的约束条件。因此, 问题的求解可以采用解决多商品流的线性规划模型。

为便于模型数学描述, 引入如下定义:

对网络 $G(V, E)$, 定义一个与之匹配的有向带权图 $G^*(V, E^*)$, 其中, E^* 是将 E 的每条无方向的边转化为带有方向两条边, 如图 1 所示:

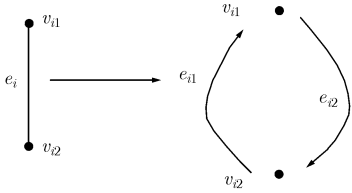


图 1 有向带权图转化示意

G^* 的边集上的容量函数定义为

$$b^*(e_i) = b(e_j), \quad e_i \in E^*, e_j \in E \quad (11)$$

其中 e_i 是由 e_j 生成的两条边之一。类似地, 定义 G^* 的费用函数为

$$f^*(e_i) = b(e_j), \quad e_i \in E^*, e_j \in E \quad (12)$$

在 $G^*(V, E^*)$ 中, 定义 $|V| \times |E^*|$ 的矩阵 \mathbf{A} , 其元素 a_{ij} 定义为

$$a_{ij} = \begin{cases} 1, & \text{节点 } v_i \text{ 是边 } e_i \text{ 的起点} \\ -1, & \text{节点 } v_i \text{ 是边 } e_i \text{ 的终点} \\ 0, & \text{其它} \end{cases} \quad (13)$$

关联矩阵 \mathbf{A} 描述了图 $G^*(V, E^*)$ 的点与边之间的连接关系。

因扩展的网络定义 $G^*(V, E^*)$ 中的边是有向的, 所以将原 R 中的每一对需求相应扩展为两对, 即

$$(s_i, t_i, d_i) \rightarrow (s_i, t_i, d_i), (t_i, s_i, d_i)$$

新定义的节点对需求 (s_i, t_i, d_i) 表示节点 s_i 和节点 t_i 之间所需带宽大小为 d_i 的网络传输能力, 其集合记为 R^* 。如无特别说明, 下面的节点对需求均指扩展后的节点需求。对每一个新定义的节点对需求 (s_i, t_i, d_i) , 定义一个在 V 上的函数 r_i , 其值为

$$r_i(v) = \begin{cases} d_i, & v = s_i \\ -d_i, & v = t_i \\ 0, & \text{其它} \end{cases} \quad (14)$$

并简记:

$$r_{ij} = r(i, v_j), \quad 1 \leq i \leq l, 1 \leq j \leq |V| \quad (15)$$

其中 v_j 的排序与定义矩阵关联 \mathbf{A} 时的排序相一致。

原先定义在 E 上函数 x , 扩展定义域为

$$U^* = (N \cap [1, l]) \times E^* \quad (16)$$

式(16)中, l 值的意义是多商品流问题中的商品数目, 即扩展后的节点对数目。 $\forall (e_i, k) \in U^*$, $x^*(e_i, k)$ 表示为满足 SON 需求中节点对 (s_i, t_i, d_i) 约束而需要在边 e_i 所对应的连接上购买的带宽, 并简记:

$$x_{ij} = x^*(i, e_j), \quad 1 \leq i \leq l, 1 \leq j \leq |E^*| \quad (17)$$

其中 e_j 的排序与定义矩阵关联 \mathbf{A} 时排序相一致。

定理 1 在网络 $G^*(V, E^*)$ 上定义的需求集合 R^* , 等价于

$$\mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i, \quad 1 \leq i \leq l \quad (18)$$

其中 $\mathbf{x}_i, \mathbf{r}_i$ 为

$$\left. \begin{aligned} \mathbf{x}_i &= (x_{i1}, x_{i2}, \dots, x_{i|E^*|}) \\ \mathbf{r}_i &= (r_{i1}, r_{i2}, \dots, r_{i|V|}) \end{aligned} \right\} \quad (19)$$

证明 $\forall i \in N(1 \leq i \leq l)$, 不失一般性, 令

$$\left. \begin{aligned} r_{i\alpha} &= d_i, & 1 \leq \alpha \leq |V| \\ r_{i\beta} &= -d_i, & 1 \leq \beta \leq |V| \\ r_{i\theta} &= 0, & 1 \leq \theta \leq |V|, \theta \neq \alpha, \beta \end{aligned} \right\} \quad (20)$$

(1) 若 \mathbf{x}_i 满足式(14), 则 $\mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i$

对 $\mathbf{A} \cdot \mathbf{x}_i^T$ 所得的向量的第 k 个元素 $\sum_{j=1}^{|E^*|} a_{kj} \cdot x_{ij}$,

考虑 \mathbf{A} 中元素定义的以及 \mathbf{x}_i 中元素排序于 \mathbf{A} 的对应关系, 当且仅当编号为 j 的边与编号为 k 的端点有

邻接关系时, $a_{kj} \neq 0$, 则 $\sum_{j=1}^{|E^*|} a_{kj} \cdot x_{ij}$ 表示编号为 k 的

端点上流量的代数和(流出为负, 流入为正)。于是, 若 \mathbf{x}_i 满足式(14)所描述的需求, 则有

$$\text{当 } k = \alpha \text{ 时, } \sum_{j=1}^{|E^*|} a_{kj} \cdot x_{ij} = d_i,$$

$$\text{当 } k = \beta \text{ 时, } \sum_{j=1}^{|E^*|} a_{kj} \cdot x_{ij} = -d_i,$$

$$\text{当 } k \neq \alpha, \beta \text{ 时, } \sum_{j=1}^{|E^*|} a_{kj} \cdot x_{ij} = 0,$$

也即 $\mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i$ 。

(2) 若 $\mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i$ 则: \mathbf{x}_i 满足式(20)

对于(1)的证明显然是可逆的，所以在满足 $\mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i$ 时，同样可以得到若 \mathbf{x}_i 满足式(14)所描述的约束条件。 证毕

可以给出构建大规模 SON 网络的线性规划模型：

$$\left. \begin{aligned} & \text{Minimize} \quad \sum_{i=1}^l \sum_{e_j \in E^*} f^*(e_j) \cdot x_{ij} \\ & \text{s.t.} \quad \sum_{i=1}^l x_{ij} \leq b^*(e_j), \quad \forall e_j \in E^* \\ & \quad \mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i, \quad i = 1, \dots, l \\ & \quad \mathbf{0} \leq \mathbf{x}_i, \quad i = 1, \dots, l \end{aligned} \right\} \quad (21)$$

在式(21)中，边集 E^* 中元素的排序与 \mathbf{A} 的定义相一致。由定理 1，约束 $\mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i$ 与 SON 的构建需求 R^* 等价，从而可知模型的准确性。而式(21)所描述模型的解空间为式(22)所示的一组集合：

$$X_i = \{x_i \mid \mathbf{A} \cdot x_i = \mathbf{r}_i\}, \quad i = 1, \dots, l \quad (22)$$

由于 $\mathbf{A} \cdot \mathbf{x}_i^T = \mathbf{r}_i$ 的限制，将每对节点需求转化为多商品流模型中的一个商品需求，从而将各个需求约束分割开来，使模型的解空间相较于前述的解决方案大大缩小。

3 实验与结果

本文采用随机生成底层物理网络拓扑和需求拓扑的方式，通过 Lingo^[11] 仿真模拟实现拓扑构建方法。图 2 给出一个较小规模的 SON 拓扑构建的图形化实例，描绘的是底层物理拓扑的位置和连接图。

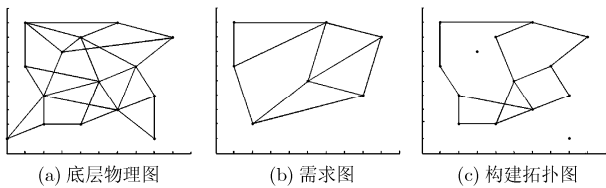


图 2 SON 拓扑构建实例示意图

3.1 构建 SON 占用资源的关系

本实验使用随机生成的底层物理网络拓扑，规模为 60 个节点，304 条连接的大规模网络。为了便于描述问题模型，每组 SON 需求的点对点流量需求均采用 100 ± 20 的随机单位值，底层物理网络的每条链路容量采用 1000 ± 200 的随机单位值。

图 3 所示的曲线中，SON 构建需求网络中包含的节点数目从 10 增加到 45，所占用的底层物理网络资源相对于 SON 节点数目近似线性关系。表 1 提供了图 3 曲线的完整数据。其中， N 为 SON 构建需求网络中的节点数目， L 为需求的点对数目，资源总量是指构建相应覆盖网所需占用的底层物理

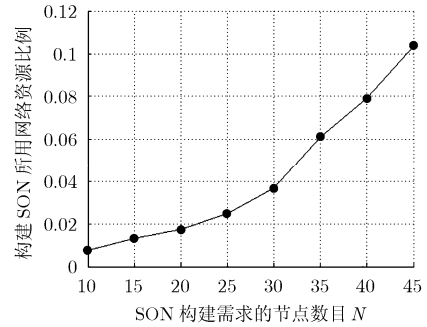


图 3 SON 构建需求节点数和占用资源比例关系图

表 1 SON 构建需求节点数目变化数据

N	L	资源总量	q
10	14	4608	0.00755
15	22	7992	0.01310
20	34	10554	0.01730
25	50	15032	0.02465
30	74	22326	0.03660
35	122	37234	0.06105
40	156	48342	0.07926
45	212	63346	0.10386

网络资源总量， q 为该总量占整个网络承载能力的比例。

从实验结果看，占用资源与 SON 节点数规模间呈近似线性比例关系，数值结果上可以判断构建模型是最优或趋于最优的。

图 4 和表 2 所示的实验以 SON 带宽和底层网络带宽比值 P_d 为变量，研究通过求解模型得到的 SON 所消耗资源，其中，底层物理拓扑仍是前面提到的随机拓扑，SON 构建需求的断点对集合采用一个固定的随机产生的集合。从结果可见，在可接受的范围内，随着 P_d 的增长，构建 SON 所消耗的网络资源大致相对 P_d 呈线性增长。

从接近线性关系的图 4 可以看出本文提出的线

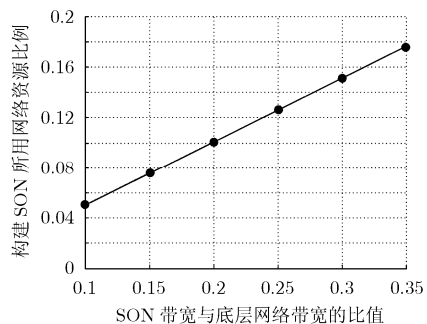


图 4 SON 带宽与底层网络带宽的比值和占用网络资源比例关系图

性规划模型在构建上具有稳定性。具体的数值结果如表 2 所示。

表 2 SON 带宽与底层网络带宽变化数据

P_d	资源总量	占用资源比例
0.10	12826	0.05030
0.15	19239	0.07545
0.20	25652	0.10060
0.25	32065	0.12575
0.30	38478	0.15094
0.35	44892	0.17606

3.2 求解运行时间

传统的最优化模型对问题的分解层次较为简化, SON 拓扑构建的求解要花费大量时间。文献[6]中所建立的最优化模型, 基于 ILP(整数线性规划)的模型对于底层网络 30 个节点 77 条边的规模网络进行求解需要花费数小时的时间。本文提出基于多商品流的线性规划模型通过 Lingo 求解, 在底层物理网规模比文献[6]更大的场景下, 构建所需时间也在秒级别。本文的模型是没有采用近似策略的最优化模型, 但时间效率上和文献[6]所提出的近似方法效率相当, 比文献[6]中提到的最优模型的时间效率大大提高。

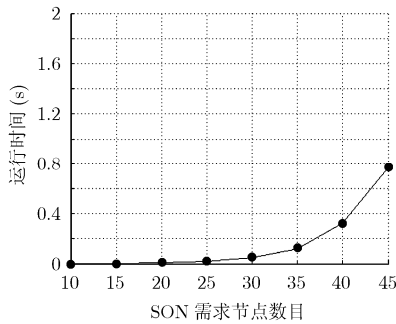


图 5 Lingo 求解运行时间

4 结束语

本文关注大规模服务覆盖网拓扑构建问题, 采用基于多商品流的节点和节点需要提供的网络传输能力为约束条件的线性规划模型, 构建大规模服务覆盖网。本文通过将每对节点的需求看作一个独立的商品流, 再将 SON 的整体拓扑构建统一为一个多商品流的线性规划模型来提高拓扑构建效率。通过

实验数据分析, 考虑带宽容量限制情况下, 采用基于多商品流的最优算法, 大大降低大规模服务覆盖网拓扑构建的时间复杂度。在服务覆盖网的拓扑构建中, 构建需求的实时性、拓扑构建中如费用、负载均衡等要求, 都是我们研究中将要考虑的问题。

参 考 文 献

- [1] Integrated Services Charter. <http://www.ietf.org/html.charters/OLD/intserv-charter.html>.
- [2] Differentiated Services Charter. <http://www.ietf.org/html.charters/OLD/diffserv-charter.html>.
- [3] Liang Jin, Gu Xiao-hui, and Nahrstedt K. Self-configuring information management for large-scale service overlays[C]. IEEE Conference on Computer Communication, Alaska, USA, 2007: 472-480.
- [4] Duan Z, Zhang Z L, and Hou Y T. Service Overlay Networks: SLAs, QoS, and bandwidth provisioning[J]. *IEEE/ACM Transactions on Networking*, 2003, 11(6): 870-883.
- [5] Garey M R and Johnson D S. Computers and Intractability: A Guide to the Theory of NP-Completeness[M]. New York: NY, USA, W. H. Freeman & Co., 1990: 7-9.
- [6] Ling Zhou and Sen A. Topology design of service overlay network with a generalized cost model[C]. Global Telecommunications Conference, New Orleans, LA, 2007: 75-80.
- [7] Capone A, Elias J, and Martignon F. Optimal design of service overlay networks[C]. International Telecommunication Networking Workshop, Italy, 2008: 46-52.
- [8] Vieira S and Liebeherr J. Topology design for service overlay networks with bandwidth guarantees[C]. International Workshop on Quality of Service, Canada, 2004: 211-220.
- [9] Li Z and Mohapatra P. On investigating overlay service topologies[J]. *Computer Networks*, 2007, 51(1): 54-68.
- [10] Zhu Y and Ammar M. Algorithms for assigning substrate network resources to virtual network components[C]. IEEE Conference on Computer Communication, Spain, 2006: 1-12.
- [11] Lingo. LINDO Systems, Inc. <http://www.lindo.com/>.

张 栋: 男, 1981 年生, 博士生, 研究方向为新一代网络、网络虚拟化。
 吴春明: 男, 1967 年生, 教授, 博士生导师, 研究方向为网络服务质量、可重构网络、网络虚拟化。
 姜 明: 男, 1974 年生, 副教授, 研究方向为网络服务质量、可重构网络、网络虚拟化。