

文章编号: 1001-0920(2013)12-1791-08

## 服务覆盖网络中一种动态流量工程模型与算法

郑明春<sup>1,2</sup>, 杨寿保<sup>2</sup>, 刘希玉<sup>1</sup>

(1. 山东师范大学 管理科学与工程学院, 济南 250014; 2. 中国科学技术大学 计算机科学与技术学院, 合肥 230037)

**摘要:** 针对服务覆盖网络中的自私路由造成的网络流量失衡将严重影响网络效率和稳定性的问题, 研究如何在覆盖层应用动态流量工程的方法进行流量优化分配. 基于随机动态优化配流理论, 提出一种服务覆盖网络的动态流量工程模型, 并设计了分布式的流量工程算法. 该算法可以折衷控制路由的自私与负载均衡的程度. 模拟实验显示, 所提出的方法较其他方法具有更好的性能, 尤其对于实时动态流量有着较好的适应性.

**关键词:** 服务覆盖网络; 动态流量分配; 流量工程

中图分类号: TP393

文献标志码: A

## A dynamic traffic engineering model and algorithm in service overlay network

ZHENG Ming-chun<sup>1,2</sup>, YANG Shou-bao<sup>2</sup>, LIU Xi-yu<sup>1</sup>

(1. School of Management Science and Engineering, Shandong Normal University, Ji'nan 250014, China; 2. School of Computer Science and Technology, University of Science and Technology of China, Hefei 230037, China. Correspondent: ZHENG Ming-chun, E-mail: zhmc163@163.com)

**Abstract:** The imbalance of traffic distribution in service overlay network caused by selfish routing influences the efficiency and stability of network. Therefore, the paper proposes a novel approach of dynamic traffic engineering to optimize traffic distribution in the overlay layer. A model of dynamic traffic engineering for service overlay network is proposed based on the stochastic dynamic optimization traffic distribution theory. A distributed algorithm is designed according to the model, which can control the trade-off between selfish routing and loading equilibrium. The experiment result shows that the approach has better performance and better adaptability for real time traffic than prior methods.

**Key words:** service overlay network; dynamic traffic assignment; traffic engineering

### 0 引言

服务覆盖网络(SON)是构建在基础网络之上的虚拟网络,由第3方经营或管理,通过向用户提供高质量的数据增值服务而获得利润.近年来,SON的研究与开发已引起众多研究者和网络服务商的关注<sup>[1]</sup>.然而,SON覆盖层实施的自私路由策略往往引起覆盖层流量的失衡,使网络资源不能合理地利用,降低了网络效益<sup>[2]</sup>.另一方面,自私的覆盖层路由与承载层的流量工程(TE)机制,由于目标冲突而产生矛盾,造成双层控制频繁翻动,使得网络不稳定<sup>[3]</sup>.为解决这些问题,一些文献研究了双层协调的方法<sup>[3]</sup>.然而,承载层的流量工程是面向AS域内的,而SON关注的是域间流量的平衡,由于AS域的自治与隐蔽性,而且与SON的覆盖范围不能一致对应,双层协调的方法难以

奏效.针对这些问题,本文提出一种新的覆盖层流量工程方法,通过建立覆盖层全局动态优化模型和算法实现SON资源的优化利用.传统的基础层流量工程优化模型与算法的设计主要依据静态的数学优化模型<sup>[4-5]</sup>,算法的动态自适应性较差,而且这些优化流量工程算法的分布式实现存在许多困难.因此,本文提出的动态TE优化模型和算法将重点解决以下问题:

1) 稳定性和收敛性问题. TE算法分为离线的长周期静态方法和在线的动态自适应方法,前者在剧烈抖动的实时流量情况下效果不佳,后者开销大且其收敛性和稳定性都难以得到保证<sup>[6]</sup>.本文应用Wardrop平衡原则建立动态优化模型,以改善算法的稳定性和收敛性.

2) 全局优化目标的分布式实现问题.全局优化模

收稿日期: 2012-08-15; 修回日期: 2012-12-26.

基金项目: 国家自然科学基金项目(61170038); 山东省自然科学基金项目(ZR2012FM013).

作者简介: 郑明春(1963-),女,教授,从事Overlay网络、流量工程等研究; 杨寿保(1947-),男,教授,博士生导师,从事新一代互联网协议和移动计算等研究.

型往往需通过中心化的代理服务器进行优化决策、全局监控和路径管理,这可能会产生中心瓶颈问题.为此,本文研究了全局优化模型的分布式实现方法.

3) 接入流的 QoS 保证与网络负载均衡之间的矛盾.自私的路由可以提高接入流的 QoS 水平,但会损害网络整体效益,而负载均衡又可能会损伤某些接入流的 QoS.为此,本文提出柔性的负载均衡策略,以使 SON 的服务更具有灵活性.

近年来,人们已开始将广泛应用于路网交通流研究的 Wardrop<sup>[7]</sup>平衡理论引入通信网络的路由优化,目前已成为一种新的研究思路,取得了一些有意义的成果<sup>[4,8-11]</sup>.其中 Fischer 等系统地研究了将 Wardrop 平衡理论应用于 TE,特别是应用于自适应流量分配领域的可行性和有效性.然而,通信网络的时变特点明显,流量变化剧烈,除了表现在高峰期流量的剧烈抖动以外,起、终点之间的最短路径以及数据包在网络中的滞留时间也都具有动态变化的特征.已有的基于静态 Wardrop 平衡理论提出的方法,在处理网络动态流量控制方面同样表现出一些不足.近年来,人们在路网规划研究中,应用动态流量配流理论(DTA)对静态 Wardrop 平衡理论进行了扩展,着眼于网络流量的瞬时变化,解决了路网流量分配的许多动态控制问题<sup>[9]</sup>.

基于以上分析,本文应用 Wardrop 动态流量优化配流理论提出一种动态优化模型.作为自适应 TE 的基础,首先证明模型符合 Wardrop 系统优化的原则;然后对优化函数中的路径代价函数进行边际化,将问题转化为动态用户平衡问题.这样,全局优化问题便可以采用分布式算法实现.此外,根据代价的误差理论,用 Logit 随机加载计算路径权值,利用公式中的平衡系数作为折衷控制负载均衡度和路径代价对流量分配的影响因子,使服务覆盖网络的服务需求更柔性化.

## 1 相关知识

为了便于理解,下面首先对 SON 的概念和动态随机配流理论进行简单的陈述.

### 1.1 服务覆盖网络

本文讨论的 SON 是构建在 IP 协议之上的虚拟骨干网络,由特别选定的位于不同 AS 域(或域间)的路由器构成,通过向 ISP 购买具有确定的 QoS 保证的基础网络的资源,向用户提供高质量的数据转发服务.

### 1.2 Wardrop 平衡、动态优化配流与随机用户平衡

Wardrop 平衡有两个基本原则:用户平衡(UE)原则和系统优化(SO)原则.前者描述的是当所有的参与者(以下称 Agent)都从自身的利益出发作出自私的路径选择,以追求个体利益最优时网络流达到平衡时的

形态,平衡时在起、终点之间所有可供选择的途经中通信流所占用的各条途经上的代价全部相等,而且小于未被选途经上的代价;后者假定所有的 Agent 能够在统一协调指挥下作出路径选择,其平衡时将达到系统的总代价最小的网络流状态.理论分析已经证明:当网络流量较小,系统畅通时,UE 与 SO 的解是相同的;随着网络拥塞程度增大,UE 与 SO 解的差别将增大.然而,如果 UE 中 Agent 考虑自己增加的流量对网络整体性能造成的影响,即 Agent 用边际代价最小作为路径选择的策略,则博弈平衡时网络将达到系统最优<sup>[12]</sup>.本文为 SON 设计一种全局优化且分布式决策算法,使每个 Agent 作为流量控制的独立执行者,因此应用 UE 原则作为 Agent 的决策策略,令边际代价作为其决策的依据,以达到系统最优的目的.详细的 Wardrop 平衡与网络配流理论请参阅文献 [12].

如前所述,DTA 注重流量的瞬时拥挤特征和需求的时变性,考虑了流量的时滞性对配流的影响,与静态配流理论比较,DTA 更真实地反映了网络的时变特征.目前,已有文献报道了一些实用的模型和算法,如文献 [13] 在解决路网应急疏散问题时,提出了一个实用的动态系统优化模型.该模型主要从车辆疏散时间方面考虑,目标是系统中的车辆总疏散时间与车辆在危险区域内滞留时间之和最小.这项成果对本文的工作具有很好的启发意义.

此外,在网络流量分配过程中,由于路径代价的估计存在误差,随机用户平衡更接近实际情况<sup>[12]</sup>,它放松了要求用户准确了解路径代价的假设,虽然每个 Agent 总是选择它认为最优的方案,但从整体看,效用大的路径被选择的概率大,效用小的路径被选择的概率小.基于 Logit 的随机配流模型是建立在所有可选路径的效用均独立且服从 Gumbel 分布的假设基础上,由于该模型易于理解和处理,已广泛应用于静态交通分配研究中,当用户对网络的代价估计得非常准确时,Logit 模型趋向于 Wardrop 用户平衡<sup>[12]</sup>.文献 [14] 将静态的 Logit 模型扩展到动态流量分配中,提出了基于 Logit 的动态随机用户优化模型.该模型建立在多路径随机分配非均衡模型的基础上,且适合大型的多 O-D 对的动态随机网络,具有较好的实用性.

综合上述分析,本文在系统优化模型的基础上采用 Logit 的随机策略进行流量分配.

## 2 动态流量工程模型

### 2.1 动态优化模型

假设 SON 中每个对的入点负责接入服务,节点的决策行为由 SON 路由器中的 Agent 模块执行.对于 SON 服务商而言,为了得到最大利润,会努力使系统

总代价最小, 同时 SON 还必须保证网络性能良好, 尽可能地减少拥塞. 为达到上述目标, 本文设计一种以系统最优为目标的动态优化模型.

定义 SON 网络为有向图  $G(N, A)$ ,  $N$  是节点的集合,  $A$  是有向弧即虚拟链集合,  $a \in A$ .

下面给出所使用的变量符号, 其中  $t$  为某时隙中的时刻:

- $x_a^n(t)$ : 链路  $a$  上的流向终点  $n$  的流量;
- $x_a(t)$ :  $a$  上的流量负荷,  $x_a(t) = \sum_n x_a^n(t), \forall a \in A$ ,

$n$  是终点数;

- $u_a^n(t)$ : 链路  $a$  上发往终点的输入流量速率;
- $v_a^n(t)$ : 链路  $a$  上发往终点的输出流量速率;
- $q^{rs}(t)$ : O-D 对从  $r$  到  $s$  的流量需求;
- $c_a(t)$ : 链路  $a$  上的代价;
- $A(k)$ : 有向弧起点是  $k$  的链路的集合;
- $B(k)$ : 有向弧终点是  $k$  的链路的集合;
- $S_{k_n}(t)$ : 节点  $k$  产生的流向终点  $n$  的速率.

SON 系统最优的动态优化模型如下:

$$J = \min \sum_{a \in A} \int_0^T [c_a(t)u_a(t)]dt. \quad (1)$$

其中: 经过链路  $a$  发往每一个终点  $n$  的输入流量速率  $u_a^n(t)$  是控制变量, 链路  $a$  上发往终点  $n$  的输出流量速率  $v_a^n(t)$  是状态变量. 模型变量满足的约束条件如下.

1) 状态约束

$$\frac{dx_a^n(t)}{dt} = u_a^n(t) - v_a^n(t). \quad (2)$$

式 (2) 的含义是流量负荷率等于进入链路的速率减去离开链路的速率. 因为  $a$  是虚拟链, 它可能包含大量的实链, 所以排队可能发生在任何一段实链中, 产生时延.

2) 非负约束. 对于  $\forall a, n, t$ , 有  $x_a^n(0) = 0, x_a^n(t) \geq 0, u_a^n(t) \geq 0$  和  $v_a^n(t) \geq 0$ .

3) 节点流量守恒约束

$$\sum_{a \in A(k)} u_a^n(t) = S_{k_n}(t) + \sum_{a \in B(k)} v_a^n(t), \quad (3)$$

$$\sum_{a \in A(n)} u_a^n(t) = 0, \quad \forall n, t. \quad (4)$$

**定理 1** 数学模型的解与 Wardrop 平衡的系统最优原则是等价的.

**证明** 建立 Hamilton 函数

$$H[x(t), u(t), \lambda(t)] = \sum_{a \in A} [c_a(t)u_a(t)] + \sum_{a \in A} \sum_{n \in N} \lambda_a^n(t)(u_a^n(t) - v_a^n(t)). \quad (5)$$

其中:  $\lambda_a^n(t)$  是协态变量;  $x(t), \mu(t)$  和  $\lambda(t)$  表示向量. 其广义拉格朗日函数

$$L[x(t), u(t), \lambda(t), \mu(t)] =$$

$$H(\cdot) + \sum_k \sum_n \mu_k^n \left[ S_{k_n}(t) + \sum_{a \in B(k)} v_a^n(t) - \sum_{a \in A(k)} u_a^n(t) \right] + \sum_n \mu_n^n(t) \sum_{a \in A(n)} [-u_a^n(t)], \quad (6)$$

其中  $\mu_k^n(t) (k = 1, 2, \dots, n)$  是拉格朗日乘子.

由 Pontryagin 最小值定理可得

$$\dot{\lambda}_a^n(t) = -\frac{\partial L(\cdot)}{\partial x_a^n(t)} = -\left[ u \frac{\partial c_a(t)}{\partial x_a^n(t)} - \frac{\partial v_a^n(t)}{\partial x_a^n(t)} \lambda_a^n(t) + \underline{\mu}_k^n(t) \frac{\partial v_a^n(t)}{\partial x_a^n(t)} \right], \quad \forall n, t, a, \quad (7)$$

其中 “ $\underline{\quad}$ ” 表示  $k$  与  $a$  的关系是  $a \in B(k)$ .

令  $\theta = \frac{\partial v_a^n(t)}{\partial x_a^n(t)}$ , 则

$$\dot{\lambda}_a^n(t) = -u_a^n(t) \frac{\partial c_a(t)}{\partial x_a^n(t)} + \theta \lambda_a^n(t) - \theta \underline{\mu}_k^n(t) = -u_a^n(t) \frac{\partial c_a(t)}{\partial x_a(t)} + \theta(\lambda_a^n(t) - \underline{\mu}_k^n(t)), \quad (8)$$

根据 KKT 条件, 有

$$\frac{\partial L(\cdot)}{\partial u_a^n(t)} = \lambda_a^n(t) - \underline{\mu}_k^n(t) + c_a(t) \geq 0, \quad \forall n, t, a,$$

其中 “ $\underline{\quad}$ ” 表示  $k$  与  $a$  的关系是  $a \in A(k)$ , 且

$$u_a^n(t) \left[ \frac{\partial L(\cdot)}{\partial u_a^n(t)} \right] = 0, \quad \forall n, t, a; \quad (9)$$

$$\frac{\partial L(\cdot)}{\partial \mu_k^n(t)} =$$

$$S_{k_n}(t) + \sum_{a \in B(k)} v_a^n(t) - \sum_{a \in A(k)} u_a^n(t) = 0; \quad (10)$$

$$\partial L(\cdot) / \partial \mu_n^n(t) = 0. \quad (11)$$

如果

$$\lambda_a^n(t) - \underline{\mu}_k^n(t) + c_a(t) > 0, \quad u_a^n(t) = 0; \quad (12)$$

$$\lambda_a^n(t) - \underline{\mu}_k^n(t) + c_a(t) = 0, \quad u_a^n(t) > 0. \quad (13)$$

则由式 (8) 可以推出

$$\begin{aligned} \lambda_a^n(t) - \underline{\mu}_k^n(t) + c_a(t) &= \\ \frac{1}{\theta} \left[ \lambda_a^n(t) + u_a^n(t) \frac{\partial c_a(t)}{\partial c_a(t)} + \theta c_a(t) \right] &= \\ c_a(t) + \frac{\partial c_a(t)}{\partial c_a(t)} u_a^n(t) + \frac{\lambda_a^n(t)}{\theta}. \end{aligned} \quad (14)$$

在 SON 中,  $x_a^n(t)$  是  $t$  时刻滞留在链路中的流量, 即

$$x_a^n(t) = v_a^n(t) d_a^n(t), \quad (15)$$

其中  $d_a^n(t)$  是  $t$  时刻链路  $a$  上流向终点  $n$  的流量的时延. 显然, 它与  $c_a(t)$  成正比.

由式 (15) 可得到

$$\frac{\partial v_a^n(t)}{\partial x_a^n(t)} = \frac{1}{d_a^n(t)} = \theta. \quad (16)$$

由约束方程 (2) 可知

$$\frac{\partial c_a(t)}{\partial v_a^n(t)} = \frac{\partial c_a(t)}{\partial u_a^n(t)}. \quad (17)$$

代入式(14)可得到

$$\begin{aligned} \lambda_a^n(t) - \underline{\mu}_k^n(t) + c_a(t) = \\ d_a^n(t)\lambda_a^n(t) + c_a(t) + \frac{\partial c_a(t)}{\partial u_a^n(t)}u_a^n(t). \end{aligned} \quad (18)$$

令

$$\varepsilon(t) = d_a^t(t)\dot{\lambda}_a^n(t), \quad (19)$$

则式(18)可以写成

$$\begin{aligned} \lambda_a^n(t) - \underline{\mu}_k^n(t) + c_a(t) = \\ \varepsilon(t) + c_a(t) + u_a^n(t)\frac{\partial c_a(t)}{\partial u_a^n(t)}. \end{aligned} \quad (20)$$

显然, 式(20)等号右边是  $t$  时刻边际阻抗, 记作  $\tilde{c}_a(t)$ . 令

$$\varphi_p(t) = \sum_{a \in A(p)} \tilde{c}_a(t), \quad \forall p. \quad (21)$$

其中:  $p$  是从起点到终点的一条路径,  $A(p)$  是路径  $p$  上的链路的集合.

$$\varphi_p(t) = \sum_{i=1}^m [\lambda_{a_i}^n(t) + c_{a_i}(t) - \underline{\mu}_{k_i}^n(t)], \quad (22)$$

其中  $m$  是路径  $p$  中包含的链路数, 由式(13)得到.

若  $u_a^n(t) > 0$ , 则有

$$\lambda_{a_i}^n(t) + c_{a_i}(t) = \underline{\mu}_{k_{i-1}}^n(t). \quad (23)$$

由式(22)推出

$$\varphi_p(t) = \mu_{k_0}^n(t) - \mu_n^n(t), \quad (24)$$

即, 若  $u_a^n(t) > 0$ , 则

$$\varphi_p(t) = \inf\{\varphi_b(t), \forall b \in P_{kn}\}, \quad (25)$$

其中  $P_{kn}$  是  $k$  到  $n$  的路径的集合. 得到

$$\begin{aligned} \varphi_1(t) = \varphi_2(t) = \dots = \varphi_j(t) \leq \varphi_{j+1}(t) \leq \dots \leq \varphi_n(t), \\ u_a^n(t) > 0, a \in \{A(p) | p = 1, 2, \dots, j\}, \\ u_a^n(t) = 0, a \in \{A(p) | p = j+1, j+2, \dots, n\}, \end{aligned} \quad (26)$$

其中  $n$  是  $P_{kn}$  的路径数.

以上证明过程说明, 模型如果有最优解, 则其最优解代表的流量形态表示在流量使用的路径上, 其边际阻抗相等且小于未被使用的路径, 显然这与 Wardrop 第 2 原则(SO)是吻合的.

## 2.2 模型的意义与进一步讨论

模型的意义有以下几点: 1) 模型(1)是系统优化, 如果按照模型进行流量控制, 则将实现系统代价最小的优化目标. 2) 由于 Wardrop 平衡是收敛的和稳定的, 应用该模型进行流量优化从理论上能够保证系统的收敛性和稳定性. 3) 模型中考虑了流量在链路中的滞留, 对于动态流量需求有较好的适应性, 实验结果也验证了这一点. 4) 模型的证明过程说明, 若每个虚拟节点用最小边际代价值作为选择路径的策略, 则系统会趋向系统最优, 也是边际代价意义上的动态用户最

优. 这样又可以将系统最优转化为用户平衡问题, 使 Agent 可以独立地进行流量分配, 为分布式算法的设计打下了基础.

此外, 本文考虑了如下两个更进一步的问题:

1) 每个虚拟节点在实际选择下一跳时, 只能根据它们当时所理解(估计)的最小代价选择路径, 即

$$C_{k,p}^n(t) = \psi_{k,p}^n(t) + \sigma_{k,p}^n(t), \quad \forall k, p, n. \quad (27)$$

其中:  $C_{k,p}^n(t)$  是路径上估计的最小代价;  $\psi_{k,p}^n(t)$  是从  $k$  点开始的路径  $p$  上的实际代价;  $\sigma_{k,p}^n(t)$  是误差项, 表示 Agent 的估计偏差, 如果每一个处于出发点或任意一个中间节点(决策点)上的 Agent 都选用当时它所估计的从该决策点直到终止点的最短代价路径, 则网络会趋向随机用户平衡的状态. 达到稳定状态时, 路径  $p$  在时刻  $t$  被选中的概率就是所有该 O-D 对路径中决策者估计的瞬时代价最小的概率, 用数学形式表示为

$$\rho_p^{rs}(t) = \Pr[C_p^{rs}(t) \leq C_q^{rs}(t), \forall q], \quad \forall r, s. \quad (28)$$

如果给出路径估计的代价概率分布, 则每一条路径被选中的概率便可以确定, 且每条路径的流量  $x_p^{rs}(t)$  为

$$x_p^{rs}(t) = q^{rs}(t)\rho_p^{rs}(t). \quad (29)$$

其中  $q^{rs}(t)$  是 O-D 对  $rs$  之间的流量需求, 通常采用下式表示瞬时代价:

$$C_p^{rs}(t) = \psi_p^{rs}(t) - \frac{1}{\theta}\varepsilon_p^{rs}(t), \quad \forall r, s, p, \quad (30)$$

其中  $\theta$  是一个非负的参数, 对  $\varepsilon_p^{rs}(t)$  起调节作用. 假设  $\varepsilon_p^{rs}(t)$  是相互独立的 Gumbel 变量, 则分布函数

$$F(w) = \Pr(\xi_k \leq w) = \exp[-e^{-\mu(\varepsilon-n)}]. \quad (31)$$

根据效用理论, 路径选择概率的 Logit 公式可表示为

$$\Pr_p^{rs}(t) = \frac{\exp[-\theta\psi_p^{rs}(t)]}{\sum_q \exp[-\theta\psi_q^{rs}(t)]}, \quad \forall r, s, p. \quad (32)$$

当  $\theta \rightarrow \infty$  时,  $\text{Var}[C_p^{rs}(t)] \rightarrow 0$ , 说明估计的瞬时代价趋于准确值, Agent 的选择策略等价于动态用户最优; 当  $\theta \rightarrow 0$  时,  $\Pr_p^{rs}(t) = 1/n$ ,  $n$  是 O-D 对之间路径个数, 此时估计误差大, 流量完全按等概率分配. 由以上分析可见, 用 Logit 公式作为用户路径选择策略既可以较好地反映真实网络的流量形态, 又不会偏离系统的优化目标.

2) 模型(1)仅从成本最小的角度对系统进行优化, 没有考虑 SON 的收益. 如果将多个目标加入模型, 将会造成模型过于复杂而难以实施. 根据经验, 如果系统能在单位时间内转发更多的数据包, 则 SON 的收益会提高. 因此, 只要在总边际代价最小化的同时提高网络利用率, 就能达到提高 SON 收益的目的. 为此, 让 SON 节点主动执行基于 Logit 公式的负载均衡的路由策略, 利用偏差系数作为系统负载均衡的控制

参数(称  $\theta$  为平衡系数), 一方面可保持系统最优的优化趋势, 同时又可以合理地收敛于动态随机优化状态.

### 2.3 链路代价的定义和度量

设  $r$  是 SON 的一个虚拟节点, 位于 O-D 对  $s \rightarrow d$  之间, 表示为  $s \rightarrow r \rightarrow d$ . 从  $r$  到  $d$  也有若干路径, 若  $r$  到  $d$  的所有可能的下一跳节点的集合是  $N(r, d)$ , 则对  $\forall v_i \in N(r, d)$ , 用  $r \rightarrow v_i$  表示  $r$  与  $v_i$  的虚拟链路, 简记为  $a_i$ . 若用  $c_{a_i}(t)$  表示链路  $a_i$  在  $t$  时刻的代价, 则有如下定义:

$$c_{a_i}(t) = \omega_1 e_{a_i}(t) + \omega_2 d_{a_i}(t), \quad (33)$$

其中  $e_{a_i}(t)$  是 SON 向 ISP 购买  $a_i$  链的成本. 假设对于每个确定的链,  $e_{a_i}(t)$  是已知的,  $d_{a_i}(t)$  是时延相关的代价,  $\omega_1$  和  $\omega_2$  是权. 对于  $d_{a_i}(t)$  的测量, 考虑到用探测会给网络增加额外的负荷, 本文设计了如下公式作为瞬时时延代价的度量:

$$d_{a_i}(t) = \frac{\alpha t_a^0}{1 - \frac{u_{a_i}(t)}{B_{a_i}(t)}}. \quad (34)$$

其中:  $t_a^0$  是网络空闲的时延代价, 称为自由代价;  $B_{a_i}(t)$  是  $t$  时刻链路有效带宽;  $\alpha$  是系数;  $u_{a_i}(t)$  是  $t$  时刻链路  $a_i$  的输入速率. 根据式(1), 用  $l_{a_i}(t)$  表示链路  $a_i$  上的边际代价, 有

$$l_{a_i}(t) = \frac{\partial c_{a_i}(t)}{\partial u_{a_i}(t)} u_{a_i}(t) + c_{a_i}(t). \quad (35)$$

结合式(34)和(35), 得到链路代价

$$l_{a_i} = \omega_1 e_{a_i}(t) + \omega_2 d_{a_i}(t) \frac{B_{a_i}(t)}{B_{a_i}(t) - u_{a_i}(t)}. \quad (36)$$

定义从任意节点  $r$  到终点  $s$  的下一跳是  $v_i$ , 其路径代价  $L(r, s, v_i)$  是从节点  $r$  到终点  $d$  的加权代价之和, 即

$$L(r, d, v_i) = l(r, v_i) + \sum_j h_j L(v_i, d, v_j). \quad (37)$$

其中:  $l(r, v_i)$  是链路  $r \rightarrow v_i$  的代价, 由式(35)决定; 路径代价  $L(v_i, d, v_j)$  是一个递归函数;  $h_j$  是权;  $v_i \in N(v_i, d)$ ;  $v_j$  是  $v_i$  的再下一跳. 令

$$h_j = \frac{\exp[-\theta L(v_i, d, v_j)]}{\sum_{v_k \in N(v_i, d)} \exp[-\theta L(v_i, d, v_k)]}. \quad (38)$$

其中:  $\theta$  是误差估计因子,  $\sum_i h_j = 1$ .

根据上一节的讨论, 如果所有的 Agent 都追求从当前节点到终点的最小边际代价路由选择, 则系统会趋于系统最优, 即每个 Agent 用式(37)作为路径代价的判据, 系统会趋向于系统最优. 此外, 由式(38)可以看出: 当  $\theta \rightarrow 0$  时,  $h_j \rightarrow 1/n$ , 说明各条路径的广义代价对各条路径的影响相等; 当  $\theta \rightarrow \infty$  时, 具有最小的广义代价的路径  $m$  的权  $h_m \rightarrow 1$ , 而其他路径的  $h_m \rightarrow 0$ , 说明这时路径  $m$  的广义代价对整条路径的权具

有绝对的影响力. 可见,  $\theta$  值可以作为流量分布的调节因子, 它控制着网络负载均衡的程度, 以下称为平衡因子.

### 3 流量工程算法

结合前面几节的讨论, 给出 SON 网络流量调节算法 DyGo, 它运行在 SON 虚拟节点上, 假设底层已经设置了基础路由算法, 如 OSPF 等. 以图 1 所示的情景为例, 对于任意虚拟决策节点  $r$ , 假设目标节点为  $d$ ,  $r$  根据 IP 层的路由信息首先确定可选的下一跳节点的集合  $V = \{v_1, v_2, \dots, v_n\}$ ; 从  $r$  到终点  $d$  有  $n$  条路径,  $r$  节点应用式(37)选择具有最小代价的路径, 决定下一跳  $v_i$ ; 在每一个反馈周期  $T_p$  结束时, 链路信息从终点  $d$  向前传递, 为了减少网络震荡, 令虚拟节点之间交换路径信息的反馈周期  $T_p$  大于路由器交换信息包的时间, 而流量控制器进行流量控制周期  $T_d$  又大于反馈周期  $T_p$ , 以下令  $T_d \geq 5T_p$ .

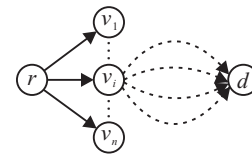


图 1 DyGo 算法路径

DyGo 算法的计算路径如下:

DyGo ( $M_1, C_1$ ) //  $M_1, C_1$  上一个时隙的信息变量

FOR every destination node  $d$  in the routing table DO

FOR all next-hop nodes  $v \in N(r, d)$  DO

Measure available bandwidth on  $r \rightarrow v$  to  $B(r, v)$

Measure rate of  $r \rightarrow v$  to  $u(r, v)$

$$\tilde{d}(r, v) = \frac{\alpha t_0(r, v)}{1 - u(r, v)/B(r, v)}$$

$$d(r, v) = \lambda \tilde{d}(r, v) + (1 - \lambda) d(r, v)$$

$$l(r, v) = d(r, v) \frac{B(r, v)}{B(r, v) - u(r, v)}$$

$$L(r, d, v) = l(r, v) + A(r, d, v)$$

END FOR

$$h_v = \frac{\exp[-\theta L(r, d, v)]}{\sum_{v_j \in N(r, d)} \exp[-\theta L(r, d, v_j)]}$$

$$A(r, d) = \sum_{v \in N(r, d)} h_v L(r, d, v)$$

Append ( $s, \{L(r, d, v) | v \in N(r, d)\}, A(r, d), \{h_v | v \in N(r, d)\}$ ) to  $M_0$

Append ( $\{e(r, d, v), d(r, d, v) | v \in N(r, d)\}$ ) to  $C_0$

END FOR

Send  $M_0, C_0$  to all neighbors  $v$  who store it in  $M_1$  and  $C_1$ .

虚拟节点在每一个控制周期结束时激活 DyGo 算法, 需要注意的是权的计算也是一个递归算法.

## 4 实验分析

本文用 NS2 作为模拟实验工具, 共列出 4 组实验. 为了充分估计本文算法的性能, 分别在简单的拓扑结构和一些典型的商业拓扑结构上进行实验, 并采取一些随机策略以避免同步效应, 虚拟节点之间通过 TCP 连接和交换信息.

### 4.1 路径权与平衡系数的关系

为了测试路径权与平衡系数的关系, 实验采用简单的拓扑结构, 如图 2 所示. 该结构包含 4 个节点、1 个 O-D 对. 节点  $v_1$  作为入点, 有 HTTP 服务器组与之相连; 节点  $v_4$  作为出点, 有客户群与之相连. 网络有两条路径.

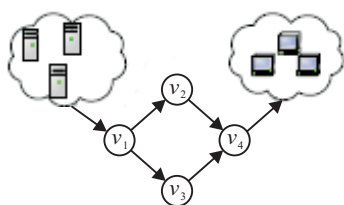


图 2 拓朴结构一

实验首先测试两条路径的权与平衡系数  $\theta$  的关系, 图 3 是当控制窗口为 2s 时两条路径的权与平衡系数的关系. 从图 3 可以看出: 当  $\theta$  从 0.2 到 0.5 变化时, 路径权值会发生很大变化; 当  $\theta$  为 0.2 以下时, 路径权值接近平均值, 说明无论路径代价大小如何, 流量都均衡分配; 当  $\theta$  大于 0.5 时, 平衡效果减弱, 流量分配主要与路径代价的大小相关. 大量实验发现,  $\theta$  大于 6 时, 链路流量分配几乎完全依赖路径代价, 这与式(32)的意义相吻合. 因此, 本文取  $\theta$  值从 0.1 ~ 6 作为流量调节控制的范围.

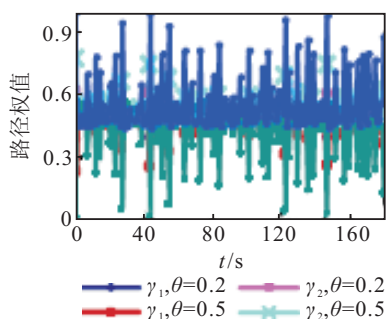


图 3 路径权与平衡系数的关系

### 4.2 平衡系数对网络性能的影响

本组实验测试不同的平衡系数  $\theta$  对网络性能的影响. 采用包含 15 个节点的 SON 拓扑结构, 其中入(出)点 10 个, 内部点 5 个, 链路 27 条, 模拟过程持续 600s, 入点向网络持续随机输入 TCP 流量. 实验观察端节点的平均吞吐量, 以测试网络整体性能, 模拟结果如图 4(a) 所示, 测量时窗  $T = 6s$ . 实验比较覆盖层

加载了 DyGo 算法和没有流量控制的情况, 可以看出: 当  $\theta = 0.2$  时, TCP 平均吞吐量最高; 当  $\theta = 1$  时, 平均吞吐量高于无控制的情况. 平均时延的实验如图 4(b) 所示. 由图 4(b) 可以看出, 当  $\theta = 1$  时, 平均时延比无控制的情况略低; 当  $\theta = 0.2$  时, 平均时延高于无控制的情况. 对比图 4(a) 和图 4(b) 可以看出, 当  $\theta = 1$  时, 网络吞吐量增加, 但平均时延没有明显增加, 本文模拟实验中  $\theta = 1$  是一种比较理想的选择.

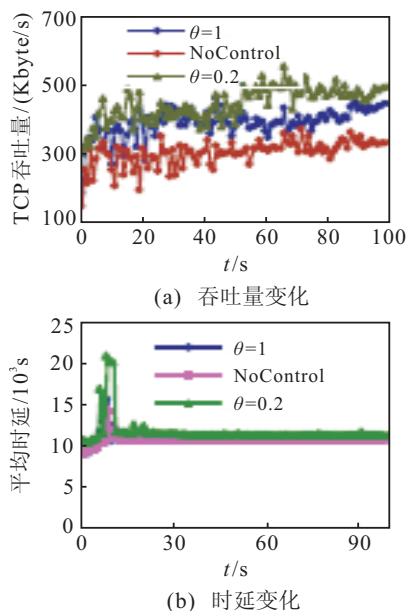


图 4 不同平衡系数下平均 TCP 吞吐量和时延变化

### 4.3 不同拓扑结构下算法的性能

这组实验用于分析复杂的拓扑结构下算法的性能. 实验采用文献 [6] 的实验环境, 用 4 种商用骨干网拓扑结构 (CW, AT, CI, SI) 进行性能测试, 图 5(a) 和图 5(b) 比较了网络在 4 种结构下自私的路由、无覆盖路由和不同平衡系数下 DyGo 算法的平均吞吐量和平均链路利用率水平. 其中 Selfish 曲线表示在 SON 网络中设置自私的覆盖层路由, 虚拟节点在选择下一跳时选择可用带宽最大的链作为下一跳, NoControl 曲线表示覆盖层没有加载路由算法. 这组实验中, DyGo 算法采用链路利用率作为路径代价函数. 由图 5 可以看出: 当  $\theta = 0.5$  时, 覆盖层接近执行完全的负载平衡算法, 网络链路利用率低, 而平均链路吞吐量则接近链路容量; 当  $\theta = 1.5$  时, 链路利用率增加, 但吞吐量降低了; 当  $\theta = 6$  时, DyGo 算法性能已经与自私路由接近, 它导致链路吞吐量降低的原因是由于自私路由可能将流量分配到相同的链路上, 造成部分链路拥塞; 同样, NoControl 是因为覆盖层没有流量控制, 也容易引起链路拥塞, 而且吞吐量比自私的路由还差.

这组实验表明, 在各种不同的拓扑结构下, DyGo 算法的作用是明显的, 当网络服务商需部署负载均衡时, 可以设较小的平衡系数  $\theta$ , 当  $\theta \geq 6$  时, 网络性能

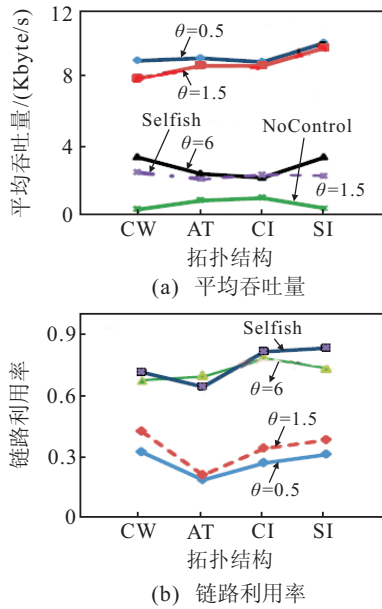


图 5 不同拓扑结构下平均吞吐量和平均链路利用率

已经接近自私路由, 因此认为  $\theta$  在 1.5 ~ 6 之间可以兼顾服务的 QoS 性能和服务商的利益. 将 DyGo 算法与文献 [5] 相比较, 由于文献 [5] 的算法不公开, 这里将图 5(a) 和图 5(b) 与文献 [5] 中的图 4 和图 5 进行对比, 可以看出: 当  $\theta = 0.5$  时, DyGo 算法与文献 [5] 的结果非常接近, 即负载均衡的性能相当. 但是, 由于文献 [5] 的每个覆盖层只面向一类服务, 而且实验中假定每个物理层节点对应一个覆盖层节点, 覆盖层的自私性没有反应出来, 从而 DyGO 算法的适应性和灵活性更好一些. 另一方面, DyGo 算法无需进行路径管理, 可以减少网络负担.

#### 4.4 算法的动态适应性

为了说明算法的动态自适应性, 将所提出的算法与一种典型的在线流量工程算法 MATE<sup>[15]</sup> 进行比较. 这里采用与文献 [15] 完全相同的拓扑结构和实验环境设置, 拓扑结构如图 6 所示. 此拓扑结构中有 3 个 O-D 对, 入点  $I_1$ 、 $I_2$ 、 $I_3$ , 出点  $E_1$ 、 $E_2$ 、 $E_3$ , 每个 O-D 对有两条路径和一个瓶颈路径, 中间节点为 1 ~ 6. 根据文献 [16], MATE 的实验参数如下: 每条链路 45 Mbps, 实验持续 7 200 s, 每个入点有 32 个源, 向网络中发送泊松流, 泊松流的平均间隔是 3 290  $\mu$ s. 每个 IP 包的长度为 257 bytes, 为了测试算法对于抖动流量的反应, 从中间点 1 和节点 3 向网络中注入背景流, 节点 1 连接 15 个泊松源, 在 [0, 7 200] s 中持续向 line1 注入流量. 类似地, 节点 2 在 [0, 3 600] s 和 [3 600, 7 200] s 连接了活跃的泊松源 15 和 30 个, 节点 3 在 [0, 1 800] s 和 [1 800, 7 200] s 分别有 35 和 20 个活跃的泊松源, 平均数据到达间隔为 2 056  $\mu$ s. 图 7 记录了实验中测到的 3 个瓶颈链的丢失率, 与文献 [15] 中图 9 相比, DyGo 算法的丢失率明显降低, 只在 [0, 260] s 和 [3 600, 3 821] s

有 1 个突出的波动, 其他很低, 而且与 MATE 相比, 对于流量抖动的变化, 本文算法反应更快, 突出波动的时段明显少于 MATE, 说明本算法的动态适应性较好. 值得注意的是, 两个算法均以时延作为系统优化的目标. 与 MATE 相比, 本文算法不需在 MPLS 域中实施, 也不需集中控制, 适应性更好.

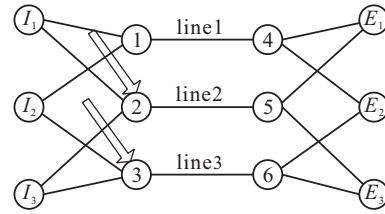


图 6 拓扑结构二

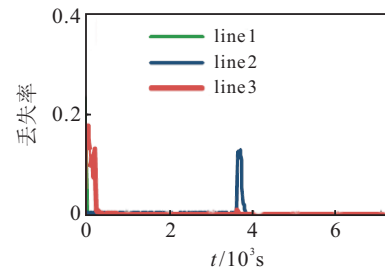


图 7 瓶颈链路丢失率

## 5 结 论

针对服务覆盖网络中由自私路由造成网络流量失衡引起网络整体性能下降的问题, 本文研究了在覆盖层应用动态流量工程的方法进行优化. 基于 Wardrop 平衡的随机动态优化配流理论, 提出了一种服务覆盖网络的动态流量工程优化模型, 并设计了相应的分布式算法. 与传统的流量工程研究方法相比, 本文工作有以下特点: 1) 动态优化模型考虑了流量的时滞, 更好地反映了流量的动态性, 因此可以在较长的控制周期条件下得到短周期的控制效果, 避免了短周期流量控制引起的网络的不稳定性; 2) 用分布式算法实现系统全局最优, 避免了中心化控制的弊端; 3) 利用平衡系数  $\theta$  对路由的自私程度和负载均衡度进行控制, 可以兼顾系统优化与用户服务质量的需求, 避免了不同层目标冲突引起的震荡, 增加了网络服务的灵活性. 理论分析和实验均表明, 本文算法可以有效地改善自私路由引起的网络性能下降问题, 并且对于动态流量具有较好的适应性, 但  $\theta$  最优值的寻找仍需进一步研究.

#### 参考文献(References)

[1] Kurian J, Sarac K. A survey on the design, applications and enhancements of application layer overlay networks[J]. ACM Computing Surveys, 2010, 43(11): 5-44.  
 [2] 王明阳, 毕军, 吴建平. 互联网覆盖路由技术研究[J]. 软件学报, 2009, 20(11): 2988-3000.

- (Wang Y Y, Bi J, Wu J P. Research on Internet overlay routing[J]. *J of Software*, 2009, 20(11): 2988-3000.)
- [3] 王鹏, 罗军舟, 李伟, 等. 基于可信可控网络的流量工程与覆盖网路由的合作博弈模型[J]. *计算机学报*, 2010, 33(9): 1663-1674.
- (Wang P, Luo J Z, Li W, et al. A cooperative game model for traffic engineering and overlay routing based on trustworthy and controllable network[J]. *Chinese J of Computer*, 2010, 33(9): 1663-1674.)
- [4] Kandula S, Katabi D, Davie B, et al. Walking the tightrope: Responsive yet stable traffic engineering[C]. *Proc of ACM SIGCOMM*. Philadelphia, 2005: 253-264.
- [5] Kawahara R, Harada S, Kamiyama N. Traffic engineering using overlay network[C]. *Proc of IEEE Communication*. Kyoto, 2011: 1-6.
- [6] Fischer S, Kammenhuber N, Feldmann A. REPLEX-dynamic traffic engineering based on Wardrop routing policies[C]. *Proc of ACM CoNEXT*. Lisboa, 2006: 100-105.
- [7] Wardrop J. Some theoretical aspects of road traffic research[C]. *Proc of the Institute of Civil Engineers, Part II*. London, 1952: 325-378.
- [8] Vivek R, Kumar P R. Wardrop routing in wireless networks[J]. *IEEE Trans on Mobile Computing*, 2009, 8(5): 636-652.
- [9] Larroca F, Rougier L. Minimum-delay load-balancing through non-parametric regression[J]. *Computer Networks*, 2012, 56(4): 1152-1166.
- [10] Fischer S, Rücke H, Vöcking B. Fast convergence to Wardrop equilibria by adaptive sampling methods[C]. *Proc of the 38th Annual ACM Symposium on Theory of Computing*. New York, 2009: 653-662.
- [11] Fischer S, Vöcking B. Adaptive routing with stale information[J]. *Theory Computer Science*, 2009, 410(36): 3357-3371.
- [12] 陆化普. 交通规划理论与方法[M]. 北京: 清华大学出版社, 2006: 182-183.
- (Lu H P. *Theory and method in transportation planning*[M]. Beijing: Tsinghua University Press, 2006: 182-183.)
- [13] 陈岳明, 萧德云. 基于动态交通分配的路网应急疏散模型[J]. *清华大学学报*, 2009, 49(8): 1102-1105.
- (Chen Y M, Xiao D Y. Dynamic traffic assignment-based method for real-time traffic management during emergency evacuation[J]. *J of Tsinghua University*, 2009, 49(8): 1102-1105.)
- [14] Lim Y, Heydecker B. Dynamic departure time and stochastic user equilibrium assignment[J]. *Transportation Research, Part B*, 2005, 39(2): 97-118.
- [15] Elwalid A, Jin C, Low S H, et al. MATE: MPLS adaptive traffic engineering[C]. *Proc of IEEE INFOCOM*. Anchorage, 2001: 1300-1309.
- [16] Sinha K, Patek S. OpIATE: Optimization integrated adaptive traffic engineering[EB/OL]. (2002-11-12). <http://www.sys.virginia.edu/techreps/2002/sie-020001.pdf>.