

BGP 最优路径选择中的瓶颈区域的研究*

刘亚萍[†], 龚正虎, 朱培栋

(国防科学技术大学 计算机学院, 湖南 长沙 410073)

Research on the Bottleneck Area of Optimal BGP Route Selection

LIU Ya-Ping[†], GONG Zheng-Hu, ZHU Pei-Dong

(School of Computer, National University of Defense Technology, Changsha 410073, China)

+ Corresponding author: Phn: +86-731-4573692, E-mail: ypliu@nudt.edu.cn, http://www.nudt.edu.cn

Received 2004-05-24; Accepted 2004-11-03

Liu YP, Gong ZH, Zhu PD. Research on the bottleneck area of optimal BGP route selection. *Journal of Software*, 2005,16(5):946-959. DOI: 10.1360/jos160946

Abstract: Optimal BGP route selection on traffic demand is one of the problems in interdomain traffic engineering. Determining bottleneck area will give important heuristic information to the problem. As the problem of determining bottleneck area is NP-hard, a bottleneck area predicting algorithm on traffic demand in polynomial time is proposed, which deals with interdomain peering links and intradomian links simultaneously. Moreover, this paper also analysis the relationships between the traffic and the bottleneck area, as well as the relationships between the topology and the bottleneck area. Simulation results show the accuracy of the algorithm is more than 90%. In addition, a conclusion is drawn from the simulation that the topology is a very important factor in determining the bottleneck area.

Key words: interdomain traffic engineering; BGP route selection; linear programming; load balance; bottleneck

摘要: 基于流量需求的 BGP 最优路径选择是域间流量工程研究的一个问题.其中瓶颈区域的判定可为域间流量工程的决策过程提供重要的启发信息.然而,瓶颈区域的判定是 NP 难问题.在同时考虑域内链路和域间链路的前提下,提出多项式时间的基于流量需求的瓶颈区域的预测算法.在此基础上,系统地研究了流量、拓扑结构与瓶颈区域间的关系.模拟实验表明,预测算法的准确性超过 90%,研究结果表明,拓扑结构是决定瓶颈区域的重要因素.

关键词: 域间流量工程;BGP 路径选择;线性规划;负载均衡;瓶颈

中图法分类号: TP393 文献标识码: A

Internet 中的域间路由协议一般采用 BGP(border gateway protocol)协议,当目的网络存在多出口路径时,BGP 允许根据策略设置选择最优路径^[1-4],如根据最高本地优先值优选路由、优选由本路由器始发的路由、

* Supported by the National Grand Fundamental Research 973 Program of China under Grant No.2003CB314802 (国家重点基础研究发展规划(973)); the National Natural Science Foundation of China under Grant No.90204005 (国家自然科学基金)

作者简介: 刘亚萍(1973 -),女,广西桂林人,博士生,讲师,主要研究领域为域间路由,网络管理;龚正虎(1945 -),男,教授,博士生导师,主要研究领域为高速路由与交换技术,网络管理;朱培栋(1971 -),男,博士,副教授,主要研究领域为域间路由,主动网络.

用最短的 AS(autonomous systems)路径优选路由、用最低的 MED(multi-exit discriminator)优选路由等。CISCO 也定义了类似的路由选择过程^[5]。但是,上述选择过程未考虑域间的流量,从而有可能导致因为出口选择不当,而引起 ISP(Internet service provider)的域间或域内链路的拥塞。

文献[6-8]讨论了基于流量需求的 BGP 路由选择,但是它们均假设瓶颈链路是域间的链路。这种假设简化了问题的求解,但是与实际的网络运行环境并不完全相符。例如,图 1 中 AS_0 表示一个 ISP 的拓扑图,1,2,3,4,5,6 表示路由器标识,其中,1,2,3 代表运行 BGP 协议的边界路由器,501~504 分别表示域间链路的标识,401 表示路由器 4 和 5 之间的链路标识,402 表示路由器 4 和 6 之间的链路标识。假设域间链路的容量均为 100M,域内链路 401 和 403 是 25M,其他均为 100M。假设当前只存在从 501 链路进入到目的网络为 P_1 的流 f_1 ,其流量为 50M 和从 502 链路进入到目的网络为 P_2 的流 f_2 ,其流量为 50M。并且去 P_1, P_2 均可通过路由器 2 的 503 链路或路由器 3 的 504 链路。如果根据文献[6-8]中的假设,只考虑域间链路的流量平衡,则 f_1 对应的出口选择 503; f_2 对应的出口选择 504。这时,无论域内路由如何选取(假定采用常用路由协议),域间链路的最大利用率是 50%,域内链路的最大利用率是 200%,将导致域内链路 401 或者 403 发生拥塞。

如果同时考虑域内链路和域间链路,根据第 1.3.1 节的瓶颈区域预测算法,调用 `compute_out_max` 可得到域间链路最大利用率的估计最大值是 100%,调用 `compute_in_max` 得到域内链路最大利用率的估计最大值是 400%,调用 `compute_out_min` 得到域间链路最大利用率的估计最小值是 50%,调用 `compute_in_min` 得到域内链路最大利用率的估计最小值是 80%。根据算法描述的步骤(4),可以得到瓶颈区域在域内链路。上述情况下,最优的选择应该是 f_1, f_2 对应的出口均选择 504。此时,域间链路的最大利用率是 100%,域内链路的最大利用率是 100%。

所以,对于一个传输 ISP,在实际的网络运行环境中,基于流量需求的 BGP 最优路径选择问题需要同时考虑瓶颈链路有可能出现在域间链路,也有可能出现在域内链路。测试研究也表明,瓶颈链路有可能出现在域间链路和域内链路^[9]。因而,根据流量需求、ISP 拓扑结构、BGP 路由信息,预测瓶颈链路的区域对于解决基于流量需求的 BGP 路径选择问题有很大的指导意义。如果首先能确定瓶颈链路的区域一定是域间链路,或者一定是域内链路,或者瓶颈链路的区域按照一定概率位于域间链路或域内链路,则 BGP 路径选择问题可分解为若干子问题来解决,从而使得该问题的求解与实际相符,并对 ISP 更具实际意义。

当前,国内的研究工作大多集中在域内流量工程,如利用优化链路权值优化域内流量^[10],以及基于 MPLS(multi protocol label switching)的流量工程等。但是对域间流量工程以及本文所提出的问题研究得较少。

所以,本文主要针对传输 ISP,来研究基于主要跨域流量需求下(约 10%的流占有所有跨域流量需求量的 89%^[11]),瓶颈区域如何确定的问题,并研究与分析流量变化、拓扑结构、路由变化与瓶颈区域的关系。

本文第 1 节提出基于流量需求的瓶颈区域的预测算法,第 2 节是该算法的分析与模拟结果,第 3 节分析流量变化、拓扑结构、路由与瓶颈区域的关系以及模拟结果,第 4 节总结研究的结论,并给出今后的主要研究重点。附录部分给出实验环境和详细的实验数据。

本文的主要贡献是提出了基于流量需求的瓶颈区域的预测算法,模拟实验表明,预测算法的准确性较好。本文还分析了流量、拓扑、路由与瓶颈区域间的关系,模拟结果与研究结论对于基于流量需求的 BGP 路由选择问题、域间流量工程以及网络规划等领域具有实际指导意义。

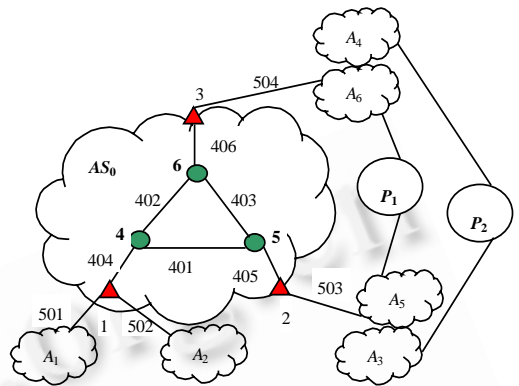


Fig.1 The topology of an ISP
图 1 某个 ISP 的拓扑图

1 基于流量需求的瓶颈区域的预测算法

1.1 问题描述

定义 1. 瓶颈链路. 链路利用率最大的链路称为瓶颈链路. 瓶颈链路不一定唯一.

基于流量需求的瓶颈区域的预测是指已知跨域的流量需求矩阵、BGP 路由表(到同一目的地存在多条路由)、ISP 拓扑结构、域内路由算法的前提下,预测瓶颈链路存在的区域.显然,当到同一目的地均只存在 1 条路由时,该预测是没有意义的.预测瓶颈区域的目的是为了解决多出口路径下,BGP 路由选择问题提供启发信息.如果 BGP 的路由表中存在到同一目的地只有 1 条路由的情况,那么根据域内路由算法,易求出到这些目的地的流量涉及的链路资源的占用情况.因此,预测算法可以首先求出所有到同一目的地只有 1 条路由的流量涉及的链路资源,然后再根据余下的流量需求矩阵进行瓶颈链路的预测.为讨论方便,不妨假定 BGP 的路由表包含的路由均是到目的地存在多条出口的情况;流量需求矩阵包含的流量也均是到目的地存在多条出口的情况.并且 BGP 路由选择的要求是:由相邻域的某入口链路而来,去往某一目的地的流只能选择一条出口,且从某入口链路来的所有流量需求不能超过入口链路的容量.

1.2 问题分析

基于流量需求的 BGP 的出口链路选择问题是一种指派问题,可归结为 0-1 整数规划问题,是 NP 难问题^[6,7],而判定瓶颈链路是处于域内链路或域间链路甚至判定是否存在瓶颈链路处于域间链路(或域内链路)的问题也是一种 0-1 整数规划问题,下面以判定是否存在瓶颈链路处于域间链路的问题为例给出了其抽象描述.

设 X 表示 n 维向量,每个分量的取值只能是 0 或 1,每个分量的物理意义表示进入该域的某个流是否需要选择某个出口链路,如果选择,取值为 1,否则为 0. $U_l(X)$ 表示根据某种选择方式计算出链路 l 的利用率,则判定是否存在瓶颈链路处于域间链路的问题可以表示为:

求 X , 它满足如下约束条件:

$$r > 0, X = (x_i), x_i \in \{0, 1\},$$

$$\text{对域内的任意链路 } j, U_j(X) < r,$$

$$\text{对域间的任意链路 } k, U_k(X) \leq r.$$

由于 0-1 整数规划问题是 NP 难问题^[6,7,12],所以判定瓶颈链路的区域问题也是 NP 难问题,因此,只能通过近似算法预测瓶颈链路的区域.如果预测算法能够判定瓶颈链路一定是在域内或域间,则称瓶颈区域是确定区域,否则称瓶颈区域处于不确定区域.

1.3 瓶颈区域预测算法

为讨论方便,首先引入一些符号与定义,见表 1.用 out_{\max} 表示域间链路最大利用率的最大值, out'_{\max} 表示域间链路最大利用率的估计最大值, out_{\min} 表示域间链路最大利用率的最小值, out'_{\min} 表示域间链路最大利用率的估计最小值, in_{\max} 表示域内链路最大利用率的最大值, in'_{\max} 表示域内链路最大利用率的估计最大值, in_{\min} 表示域内链路最大利用率的最小值, in'_{\min} 表示域内链路最大利用率的估计最小值, A 表示相邻 AS 的集合, B 表示域间链路的集合, C 表示链路容量, L 表示域内链路的集合, P 表示“跨域流量”目的网络地址前缀的集合, $Out(P_k)$ 表示目的为 P_k 的可行出口链路的集合, $In(A_h)$ 表示从相邻 A_h 域可进入本域中的入口链路的集合, $t(h, i, k)$ 表示从相邻 h 域、入口链路 i 来,去往目的地 k 的流量, N 表示边界路由器的集合, $R(b_i)$ 表示域间某条链路 b_i 对应的边界路由器.函数 f 表示将流量映射为某个出口链路的函数,设对某个流量 $t(h, i, k)$,若 $f(h, i, k) = j$,则可用下述 0-1 变量表示为

$$x_{h,i,k}^j = 1, j \in Out(k),$$

$$x_{h,i,k}^{j'} = 0, \forall j' \in Out(k), j' \neq j,$$

$$\sum_j x_{h,i,k}^j = 1, \forall j \in Out(k).$$

函数 f 表示流量与域内链路映射的函数,例如 $f'_{n_1, n_2}(l)$ 表示从节点 n_1 到 n_2 之间流量为 1 的流在域内链路 l

上流过的流量,且 n_1, n_2 是边界路由器.

$t'(n_1, n_2)$ 表示从节点 n_1 到 n_2 之间的流量,它满足如下关系:

$$t'(n_1, n_2) = \sum_{h,i,k} \left. \begin{aligned} &Y_i^{n_1} \times Y_{f(h,i,k)}^{n_2} \times t(h, i, k) \\ &\{h, i, k, n_1, n_2, | h \in A, i \in \text{in}(h), k \in P, n_1, n_2 \in N, \text{若} R(i) = n_1, Y_i^{n_1} = 1, \text{否则} Y_i^{n_1} = 0\} \end{aligned} \right\} \quad (1)$$

Table 1 Notations and definitions

表 1 符号表示与定义

Notation	Description
out'_{\max}	The real max value of the max peering link utility
out'_{\min}	The estimated max value of the max peering link utility
out_{\max}	The real min value of the max peering link utility
out'_{\min}	The estimated min value of the max peering link utility
in_{\max}	The real max value of the max intra-domain link utility
in'_{\max}	The estimated max value of the max intra-domain link utility
in_{\min}	The real min value of the max intra-domain link utility
in'_{\min}	The estimated min value of the max intra-domain link utility
A	$\{A_1, A_2, \dots, A_q\}$, set of AS neighbors
B	$\{b_1, b_2, \dots, b_n\}$, set of peering links
$C(j)$	The capacity of link j
L	Set of intra-domain links
P	$\{P_1, P_2, \dots, P_m\}$, set of network prefixes for transit routing
$\text{Out}(P_k)$	Set of egress peering links for P_k
$\text{in}(A_h)$	Set of ingress peering links from neighbor A_h
$t(h, i, k)$	Traffic from neighbor h through ingress peering link i destined for prefix k
N	Set of border routers
$R(b_i)$	The border router whose peering link is b_i
f	Function that maps traffic to an egress point
f'	Function that maps traffic to the intra-domain links
$s(j)$	The link utility of the link j
$t'(n_1, n_2)$	The traffic from n_1 to n_2

链路利用率 $s(j)$ 分为域间链路利用率和域内链路利用率.域间链路利用率可根据式(2)计算,域内链路利用率可根据式(3)计算.

$$s(j) = \sum_{h,i,k} \frac{X_{h,i,k}^j \times t(h, i, k)}{C(j)}, \{h, i, k, j | h \in A, i \in \text{in}(h), k \in P, j \in \text{Out}(k)\} \quad (2)$$

$$s(l) = \sum_{n_1, n_2} \frac{t'(n_1, n_2) \times f'_{n_1, n_2}(l)}{C(l)}, \{n_1, n_2, l | n_1, n_2 \in N, l \in L\} \quad (3)$$

以引言中的示例为例,如果 f_1, f_2 对应的出口均选择 504,则表 1 中定义的一些符号的取值如下:

$\text{out}'_{\max}=1, \text{out}'_{\min}=0.5, \text{in}_{\max}=4, \text{in}_{\min}=1, P=\{P_1, P_2\}, \text{Out}(P_1)=\{503, 504\}, \text{in}(A_1)=\{501\}, R(501)=1,$

$A=\{A_1, A_2, A_3, A_4, A_5, A_6\}, B=\{501, 502, 503, 504\}, L=\{401, 402, 403, 404, 405, 406\}, N=\{1, 2, 3\},$

$\text{out}'_{\max}=1, \text{out}'_{\min}=0.5, \text{in}'_{\max}=4, \text{in}'_{\min}=0.8$ (根据第 1.3.1 节中的算法).

$f(A_1, 501, P_1)=504, f'_{1,3}(402)=1, f'_{1,3}(401)=0, t'(1, 3)=100M, s(504)=100\%, s(402)=100\%.$

定义 2. 网络状态.根据坐标轴上域间链路最大利用率的区间与域内链路最大利用率的区间的关系,定义网络的 6 种状态,见表 2,示意图如图 2 所示.

定义 3. 网络估计状态.用估计值代替实际值定义的网络状态,称为网络估计状态.

Table 2 Network state

表 2 网络状态

State	Satisfied inequality
0	$\text{in}_{\max} \leq \text{out}_{\min}$
1	$\text{in}_{\max} > \text{out}_{\min}$ and $\text{in}_{\max} \leq \text{out}_{\max}$ and $\text{in}_{\min} < \text{out}_{\min}$
2	$\text{in}_{\max} > \text{out}_{\min}$ and $\text{in}_{\max} \leq \text{out}_{\max}$ and $\text{in}_{\min} > \text{out}_{\min}$
3	$\text{in}_{\max} > \text{out}_{\max}$ and $\text{in}_{\min} > \text{out}_{\min}$ and $\text{in}_{\min} < \text{out}_{\max}$
4	$\text{in}_{\max} > \text{out}_{\max}$ and $\text{in}_{\min} < \text{out}_{\min}$
5	$\text{in}_{\min} > \text{out}_{\max}$

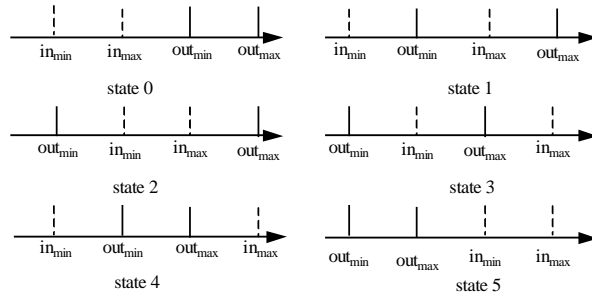


Fig.2 Network state

图 2 网络状态

1.3.1 瓶颈区域预测算法

瓶颈区域预测算法描述.

(1) 调用 compute_out_max()计算 out'_{max} ,调用 compute_out_min()计算 out'_{min} ,调用 compute_in_max()计算 in'_{max} ,调用 compute_in_min()计算 in'_{min} 的值.

(2) 根据定义,求网络估计状态.

(3) 如果估计状态是 0,则瓶颈区域在域间链路.如果估计状态是 5,则瓶颈区域在域内链路.

(4) 如果估计状态是 1,且交叉部分的线段所占比例小于 γ (当前取 $\gamma=12\%$),则瓶颈区域在域间链路的概率接近 1;如果估计状态是 3,且交叉部分的线段所占比例小于 γ ,则瓶颈区域在域内链路的概率接近 1;其他情况下,瓶颈区域处于不确定区域.

compute_out_max 算法描述.

(1) 对于所有的 $t(h,i,k)$,令对任意 $j \in out(k)$, $X^j_{hik} = 1$.

(2) 根据定义计算每条域间链路的链路利用率 $s(j)$.

(3) 计算域间链路的最大链路利用率 out'_{max} .

compute_out_min 算法描述.

用线性规划的方法求出 z , $out'_{min} = z$.

下面的(4)表示规划目标是求最小值,式(5)~式(7)表示需要满足的约束方程.

$$\text{求 } z=r \text{ 的最小值} \tag{4}$$

对任意属于域间链路集合 B 的链路 j ,满足如下条件:

$$\forall j \in B, \sum_{h,i,k} \frac{X^j_{h,i,k} \times t(h,i,k)}{C(j)} \leq r; i \in in(h), h \in A, k \in P, j \in Out(k) \tag{5}$$

$$\sum_j X^j_{h,i,k} = 1 \tag{6}$$

$$X^j_{h,i,k} \geq 0, r \geq 0 \tag{7}$$

compute_in_max 算法描述.

(1) 根据域内路由算法 f' ,计算出域内每条链路 l 上所有能承载的源地址和目的地址是 BGP 节点的流量对 $FL(l) = \{ \langle n_1, n_2 \rangle, \langle n'_1, n'_2 \rangle, \dots \} (n_1, n_2, n'_1, n'_2 \in N)$.

(2) 对每条域间链路 l 按下述步骤求链路利用率:

(2-1) $t=0$;

(2-2) 对每个 $t(h,i,k)$,若存在 $j \in out(k)$ 且满足 $X^j_{hik} = 1, \langle R(i), R(f(h,i,k)) \rangle \in FL(l)$,则 $t=t+t(h,i,k)$;

(2-3) $s(l)=t/C(l)$.

(3) 计算域内链路最大链路利用率 in'_{max} .

compute_in_min 算法描述.

(1) 对任意 $t(h,i,k)$,若存在 $j \in out(k)$ 且 $R(i)=R(j)$,则令 $X^j_{hik} = 1$,其他 $j' \in out(k)$,令 $X^{j'}_{hik} = 0$.

(2) 对不存在(1)中情况的所有 $t(h,i,k)$ 作如下处理:

(2-1) 求 $j \in \text{out}(k)$ 对应的所有 $R(j)$, 令 $RR(R(i),k) = \{R(j)\}$.

(2-2) 若 RR 元素数目大于 1, 则引入新节点 n' (用 $D(k)$ 表示), 增加 RR 集合中各点到 n' 的边, 边的容量为无穷大, 令 $N' = N \cup \{n'\}$, 如果图中已存在同样的点 n' , 则不引入新节点, 不增加边, 执行(2-4).

(2-3) 若 RR 元素数目等于 1, 不妨记 RR 集合中的点为 n (仍可用 $D(k)$ 表示), 执行(2-4).

(2-4) $t'(R(i), D(k)) += t(h,i,k)$.

(3) 根据当前的域内路由算法, 计算节点间的流量需求矩阵 $t'(n_1, n_2)$ 中所有满足 $n_2 \in N$ 的流量涉及的域内链路资源 $\text{used}(l), l \in L$.

(4) 根据新的拓扑图 G' , 节点间的流量需求矩阵 $t'(n_1, n_2) (n_2 \in N' - N)$, 以及当前的域内路由算法和选路的约束条件, 利用线性规划的方法, 计算最大链路率的最小值 z' , $\text{in}'_{\min} = z'$. 引入变量 $G'_{R(i), D(k)}^r, r \in RR(R(i), k)$. 下面所列式子中, 式(8)表示规划目标是求最小值, 式(9)~式(11)表示需要满足的约束方程.

$$\text{求 } z' = w \text{ 的最小值} \tag{8}$$

$$\forall l \in L, \frac{\text{used}(l) + \sum_{n_1, n_2} t'(n_1, n_2) G'_{R(i), D(k)}^r f'_{n_1, r}(l)}{C(l)} \leq w; n_1, n_2 \in N', R(i) = n_1, D(k) = n_2 \tag{9}$$

$$\sum_r G'_{R(i), D(k)}^r = 1 \tag{10}$$

$$G'_{R(i), D(k)}^r \geq 0, w > 0 \tag{11}$$

以引言中的示例为例, 调用 `compute_out_max`, 易得 $\text{out}'_{\max} = 100\%$. 调用 `compute_out_min` (例如利用 MATLAB 提供的解 LP 问题的函数), $\text{out}'_{\min} = 50\%$. 调用 `compute_in_max`, 易得 $\text{in}'_{\max} = 400\%$. 根据 `compute_in_min`, 新的拓扑图 G' 如图 3 所示 (增加了新节点 n'). 此时 $t'(1, n') = 100M$, 易求出 $\text{in}'_{\min} = 80\%$. 根据预测算法, 网络估计状态是 3, 且交叉部分的线段所占比例是 5.7%, 小于 γ , 瓶颈区域在域内链路.

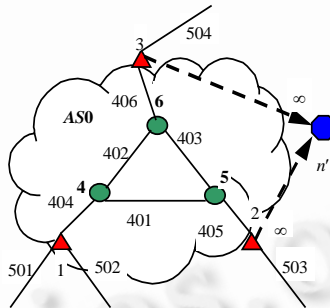


Fig.3 G'
图 3 图 G'

1.3.2 瓶颈区域预测算法的正确性证明

引理 1. $\text{out}'_{\max} = \text{out}_{\max}, \text{in}'_{\max} = \text{in}_{\max}, \text{out}'_{\min} \leq \text{out}_{\min}, \text{in}'_{\min} \leq \text{in}_{\min}$.

证明: (1) 证明 $\text{out}'_{\max} = \text{out}_{\max}$.

(反证法):

根据 out'_{\max} 的求解算法, 显然 $\text{out}'_{\max} \geq \text{out}_{\max}$. 假设 $\text{out}'_{\max} \neq \text{out}_{\max}$, 则 $\text{out}'_{\max} > \text{out}_{\max}$. 不妨设 out_{\max} 对应的出口选择方式为 x , out'_{\max} 对应的瓶颈链路是 b , 且有 n 条流选择的出口是 b .

由此, 令这 n 条流对应的变量 $X_{hik}^b = 1$, 其他流的出口任意选择, 显然, 这是一种出口选择方式, 不妨记该选择方式为 y . 在 y 选择方式下域间链路利用率的最大值等于 out'_{\max} . 而 $\text{out}'_{\max} > \text{out}_{\max}$, 与 out_{\max} 的定义矛盾. 因此假设不成立, out'_{\max} 一定等于 out_{\max} .

(2) 证明 $\text{in}'_{\max} = \text{in}_{\max}$.

同理, 可用反证法证明 $\text{in}'_{\max} = \text{in}_{\max}$.

(3) 证明 $out'_{min} \leq out_{min}, in'_{min} \leq in_{min}$.

out_{min} 的计算与 out'_{min} 的不同之处是 out_{min} 是一种特殊的线性规划,0-1 规划,即将 out'_{min} 的约束条件(7)改为 $X_{h,i,k}^j \in \{0,1\}$.由于 out'_{min} 是求出的理想最优值,所以 $out'_{min} \leq out_{min}$,同理 $in'_{min} \leq in_{min}$.

瓶颈区域预测算法的正确性证明.

证明:证明该算法的正确性就是要证明如果网络的估计状态是 0,则实际状态一定是 0,如果网络的估计状态是 5,则实际状态一定是 5.根据引理 1,网络的估计状态与实际状态存在如表 3 所示的关系(将 $in_{max} = out_{max}, in_{min} = out_{min}$ 也作为状态 4 的一个特例),因此结论成立.

Table 3 Relationship between estimated state and state

表 3 估计状态与实际状态关系

Estimated state	state
0,1,2	0
1,2	1
1,2	2
3,4	3
3,4	4
3,4,5	5

2 瓶颈区域预测算法的分析与模拟实验

2.1 瓶颈区域预测算法的复杂性分析

out'_{max}, in'_{max} 的求解显然是多项式时间, out'_{min}, in'_{min} 是用线性规划的方法求解.线性规划问题是多项式时间可求解的问题,因此瓶颈区域预测算法的复杂性是多项式时间的.

2.2 瓶颈区域预测算法的准确性分析

2.2.1 网络状态预测的准确性分析

假设 out'_{min} 与 out_{min} 间的误差用 α 表示, in'_{min} 与 in_{min} 间的误差用 β 表示.设 α 服从正态分布 $N(\mu_1, \sigma_1^2)$, β 服从正态分布 $N(\mu_2, \sigma_2^2)$.由于 out'_{min}, in'_{min} 均采用类似的线性规划的方法求解,因此不妨假定有相同的误差精度 $\mu_1 = \mu_2 = \mu$.用 e_state 表示估计状态, $state$ 表示实际状态, P 表示概率.分 3 种情况进行分析.

(1) $state=0$

令 $a_1 = out_{min} - in_{max}, a_2 = in_{max} - in_{min}$,则

$$P(e_state = 0 | state = 0) = P(\beta < a_1) = \begin{cases} \phi((a_1 - \mu) / \sigma) & (a_1 > \mu) \\ 1 - \phi((\mu - a_1) / \sigma) & (a_1 < \mu) \end{cases}$$

$$P(e_state = 1 | state = 0) = \begin{cases} \phi((a_1 + a_2) / (\sigma_1^2 + \sigma_2^2)^{1/2}) - \phi((a_1 - \mu) / \sigma) & (a_1 > \mu) \\ \phi((a_1 + a_2) / (\sigma_1^2 + \sigma_2^2)^{1/2}) - \phi((\mu - a_1) / \sigma) - 1 & (a_1 < \mu) \end{cases}$$

$$P(e_state = 2 | state = 0) = P(\beta - \alpha > a_1 + a_2) = 1 - \phi((a_1 + a_2) / (\sigma_1^2 + \sigma_2^2)^{1/2}).$$

所以发生状态预测不准确的概率为 $1 - \phi((a_1 - \mu) / \sigma) (a_1 > \mu)$ 或 $\phi((\mu - a_1) / \sigma) (a_1 < \mu)$.

(2) $state=5$

令 $b_1 = in_{min} - out_{max}, b_2 = out_{max} - out_{min}$,与(1)类似,发生状态预测不准确的概率为 $1 - \phi((b_1 - \mu) / \sigma) (b_1 > \mu)$ 或 $\phi((\mu - b_1) / \sigma) (b_1 < \mu)$.

(3) $state \in \{1, 2, 3, 4\}$

定义 4. 相对位移与位移翻转.域间链路最大利用率的最小值减去域内链路最大利用率的最小值,称为相对位移.域间链路最大利用率的估计最小值减去域内链路最大利用率的估计最小值,称为估计相对位移.如果估计相对位移与相对位移的符号相反,称为位移翻转.

若网络实际状态是 1,2,3,4,状态预测不准确一定是发生了位移翻转.

令 $c = |out_{min} - in_{min}|$,

若 $state=1, P(\text{位移翻转}) = P(\alpha - \beta > c) = 1 - \phi(c / (\sigma_1^2 + \sigma_2^2)^{1/2})$,

若 $state=2, P(\text{位移翻转})=P(\beta-\alpha>c)=1-\phi(c/(\sigma_1^2+\sigma_2^2)^{1/2})$,

若 $state=3, P(\text{位移翻转})=P(\beta-\alpha>c)=1-\phi(c/(\sigma_1^2+\sigma_2^2)^{1/2})$,

若 $state=4, P(\text{位移翻转})=P(\alpha-\beta>c)=1-\phi(c/(\sigma_1^2+\sigma_2^2)^{1/2})$.

以上分析说明,如果实际状态是 0 或 5,且 a_1 和 b_1 远大于 μ ,则状态预测不准确的概率很小,但是如果 a_1 和 b_1 小于 μ ,状态预测不准确的概率将会超过 50%.如果实际状态是 1,2,3,4,且 $c>=(\sigma_1^2+\sigma_2^2)^{1/2}$,状态预测不准确的概率将小于 16%,最差情况下 $c=0$,状态预测不准确的概率等于 50%.

2.2.2 瓶颈区域预测的准确性分析

瓶颈区域预测的不准确有两种情况:一是实际瓶颈区域一定在确定区域,但预测瓶颈区域在不确定区域,其概率记为 PP_1 ;二是实际瓶颈区域一定在不确定区域,但预测瓶颈区域在确定区域,其概率记为 PP_2 .第 1 种情况又可分为实际状态是 0,1,3,5 这 4 种,然而 1 与 3 类似,0 与 5 类似,因此第 1 种情况只分析实际状态是 0 或 1.第 2 种情况又可分为实际状态是 1,2,3,4 这 4 种,然而 1 与 3 类似,2 与 4 类似,所以第 2 种情况只分析实际状态是 1 或 2.

(1) 实际瓶颈区域一定在确定区域,但预测瓶颈区域在不确定区域的概率分析.

将网络实际状态为 1 分成 4 种子状态:不妨记瓶颈区域在域间链路,且相交部分所占比例小于 γ 的状态为 X_1 ,瓶颈区域在域间链路,且相交部分所占比例大于 γ 的状态为 X_2 ;瓶颈区域是不确定区域,且相交部分所占比例小于 γ 的状态为 X_3 ,瓶颈区域是不确定区域,且相交部分所占比例大于 γ 的状态为 X_4 .根据瓶颈区域预测算法, PP_1 满足下列等式:

$$PP_1 = P(e_state \neq 0, X_1 | state = 0) \times P(state = 0) + P(e_state \neq X_1 | state = X_1) \times P(state = X_1) + P(e_state \neq X_1 | state = X_2) \times P(state = X_2).$$

根据第 2.2.1 节的分析,当 a_1 远大于 μ 时, $P(e_state \neq 0 | state = 0)$ 很小,当 a_1 与 μ 相当时, $P(e_state \neq 0, X_1 | state = 0)$ 很小,因此 $P(e_state \neq 0, X_1 | state = 0) \times P(state = 0)$ 很小.

若实际状态是 X_1 ,估计状态有 3 种可能性:发生位移翻转、 X_2 和 X_1 .由于实际状态是 X_1 , c 的值一定较大,因此第 1 种可能性很小.第 2 种情况 e_state 是 X_2 可能发生在 μ 远大于相交部分所占比例,这种可能性也较小,因此 $P(e_state \neq X_1 | state = X_1) \times P(state = X_1)$ 很小.

若实际状态是 X_2 ,估计状态也有 3 种可能性:发生位移翻转、 X_2 和 X_1 .由于实际状态是 X_2 , c 值一定较小,发生位移翻转可能性较大.估计状态是 X_2 的可能性也较大.因此 $P(e_state \neq X_1 | state = X_2)$ 较大.所以 PP_1 主要取决于 $P(state = X_2)$.如果 γ 选择合适,满足 $P(state = X_1 | state = 1) \gg P(state = X_2 | state = 1)$,则 PP_1 很小.

(2) 实际瓶颈区域一定在不确定区域,但预测瓶颈区域在确定区域的概率分析.

PP_2 满足下列等式.若实际状态是 2,估计状态是 X_1 ,则发生了位移翻转,而 c 值较小,发生位移翻转的可能性较大,但即使发生位移翻转,交叉部分比例也很大,所以 $P(e_state = X_1 | state = 2)$ 很小.由于 μ_1 等于 μ_2 ,所以 $P(e_state = X_1 | state = X_4)$ 很小, $P(e_state = X_1 | state = X_3)$ 较大.所以 PP_2 主要取决于 $P(state = X_3)$.如果 γ 选择合适, $P(state = X_4 | state = 1) \gg P(state = X_3 | state = 1)$,则 PP_2 很小.

$$PP_2 = P(e_state = X_1 | state = 2) \times P(state = 2) + P(e_state = X_1 | state = X_4) \times P(state = X_4) + P(e_state = X_1 | state = X_3) \times P(state = X_3).$$

综上所述(1)、(2),瓶颈区域预测算法的精确度与 γ 的选择有关.

2.3 模拟实验

设所有内部路由采用 OSPF,到某一目的地的路由是由随机生成的 2~4 个不同的出口形成.流量需求矩阵状态表是指如果存在从某一条域间链路进入到某一目的地的流为 1,否则为 0.流量需求矩阵状态表是随机生成的,且为简化问题,约定一条域间链路的相邻 AS 只有一个.根据流量需求,矩阵状态表随机生成不超过链路容量的流量需求矩阵.分别取链路容量不同的值,进行模拟实验(拓扑图、路由表、流量需求矩阵状态表等详细实验数据参见附录).

2.3.1 实际状态与估计状态比较实验

模拟实验表明(见表 4),根据估计值求出的估计状态接近实际状态.估计状态和实际状态之间的准确度超过 75%.实际状态是 0,估计状态是 2 的概率接近 0;实际状态是 5,估计状态是 4 的概率接近 0.这说明,当实际状态是 0 或 5 时,估计状态发生位移翻转的可能性很小,这是因为此时相对位移的绝对值较大.当表中实际状态为 4 时,状态估计的准确性较低的原因是在实验中出现实际状态为 4 的情况较少,样本空间较小的缘故.模拟实验中状态预测不准确的概率略高于理论分析值,主要原因是统计样本不够大.

Table 4 State and estimated state

表 4 实际状态与估计状态

State	e_state	Frequency	e_state	Frequency	e_state	Frequency	Precision
0	0	0.924	1	0.176	2	0.000	0.924
1	1	0.827	2	0.173			0.828
2	1	0.270	2	0.730			0.730
3	3	0.979	4	0.022			0.979
4	3	0.670	4	0.330			0.330
5	3	0.000	4	0.000	5	1.000	1.000

2.3.2 预测算法瓶颈区域的准确度实验

实验表明,预测算法关于瓶颈区域的准确度超过 90%(参见附录 1 的各表),不准确的主要原因是 $P(\text{state}=X_2)$ 较大.即表明当前 γ 选择满足 $P(\text{state}=X_4|\text{state}=1) \gg P(\text{state}=X_3|\text{state}=1)$,但是 $P(\text{state}=X_1|\text{state}=1)$ 并不远大于 $P(\text{state}=X_2|\text{state}=1)$.这是因为根据附录 2 的实验结果, γ 取 0.12 是其均值,但 γ 方差较大,因此 $P(\text{state}=X_2|\text{state}=1)$ 较大. γ 的优化选择有待进一步研究.实验同时表明,瓶颈区域既有可能出现在域间链路,又有可能出现在域内链路,这与文献[9]中对 Internet 实际测量的结果是相吻合的.

2.3.3 瓶颈区域概率与网络实际状态关系实验

分别取不同的拓扑图、路由表,流量需求矩阵表随机生成,并分别对链路容量取不同值(参见附录 2).实验表明,如果状态是 1、3,且交叉部分的线段所占比例小于 γ ($\gamma < 12\%$),则瓶颈链路在域间链路或域内链路的概率接近 1,状态 2、4 下的瓶颈区域概率难以确定.实验结论符合第 2.2.2 节中的理论分析.

3 流量变化、拓扑结构与瓶颈区域的关系

3.1 流量变化与瓶颈链路区域的关系

为简化问题,仅考虑流量变化的两种极端情况:流量单调变化和流量随机变化下对瓶颈区域的影响.

(1) 流量单调变化

实验环境与数据见附录 3,实验表明,当初始流量矩阵位于临界状态(状态 1 或 3)时,易发生瓶颈区域变化;轻载下流量的单调变化,易导致状态变化.中载和重载下,流量的单调变化不易导致状态变化.这是因为轻载下,流量的单调变化易导致较严重的流量分布的不平衡.

(2) 流量随机变化

实验环境与数据见附录 4,实验表明,流量随机变化比流量单调变化易导致状态变化,但是流量随机变化很少导致瓶颈区域在确定的域内链路和域间链路区域间跃变.

上述分析和模拟考虑的是流量变化的极端情况.然而,在实际环境中,根据当前的研究^[2]表明少数的流占用了大多数的跨域流量,并且这些流的流量变化较稳定.因此,重新进行瓶颈区域预测的时机一是流量发生剧烈变化时,二是以较长时间间隔周期进行.当然,在考虑瓶颈区域预测时机时还应考虑路由的抖动性,而在网络负载是重载的情况下,路由稳定性的重要性超过网络负载的均衡性^[13],所以如何检测流量发生剧烈变化以及预测时间间隔的选取是进一步研究与实验的重点.

3.2 拓扑结构与瓶颈区域的关系

为简化问题的讨论,模拟实验仅考虑链路容量的变化.实验结果(实验数据见附录 5)表明,当域间链路容量与域内链路容量之比发生显著变化时,网络状态、瓶颈区域易发生变化.这说明,拓扑结构是决定网络状态的重

要因素。目前,国外一些研究机构正在开展拓扑结构与域间流量工程的研究。研究表明,根据测量的方法,在实际的因特网中,域间链路与域内链路成为瓶颈的概率均等,链路容量小的链路有可能成为瓶颈^[9]。但是,由于网络拓扑结构的复杂性和动态性,它与瓶颈区域的关系还有很多规律有待研究。下一步的工作将结合测量的方法进一步研究在 BGP 最优路径的选择中,拓扑结构与瓶颈区域的关系。

4 总 结

本文主要研究了在基于流量需求的 BGP 最优路由选择中,瓶颈区域如何确定的问题,并分析了流量、拓扑结构与瓶颈区域的关系。本文的研究工作对 BGP 最优路由选择问题的解决有实际的指导意义,它使得对该问题的考虑更贴近实际 ISP。本文的主要研究结论是:(1) 可以用解 LP 的方法来估计域间链路和域内链路最大利用率的最小值,通过域间链路和域内链路最大利用率范围在坐标轴上的关系来定义网络状态,从而确定瓶颈链路可能的区域,因此可为 BGP 路由选择问题提供重要的启发信息。(2) 模拟实验表明,瓶颈区域预测算法的准确性较高,其准确性与 γ 参数有关。(3) 模拟实验表明,拓扑结构是决定网络状态和瓶颈区域的重要因素。

下一步的工作将研究实际域间流量需求的变化、BGP 路由变化、拓扑变化对网络状态与瓶颈区域的影响,以及在实际的网络环境中进行模拟验证。在此基础上,基于瓶颈区域的判定,研究基于流量需求条件下,优化多路径的 BGP 路由选择问题的通用解决方法,并比较文献[7]中提出的迭代随机搜索算法(RRS)、用于 stub ISP 中优化多路径的 BGP 路由选择问题的遗传算法^[8]、优化目标是域内资源占用率最低且满足域间链路容量约束的启发算法^[6]、传统的“热土豆算法”^[4]在通用条件下算法的复杂性和适用程度,从而对该问题形成有效的解决方法。

References:

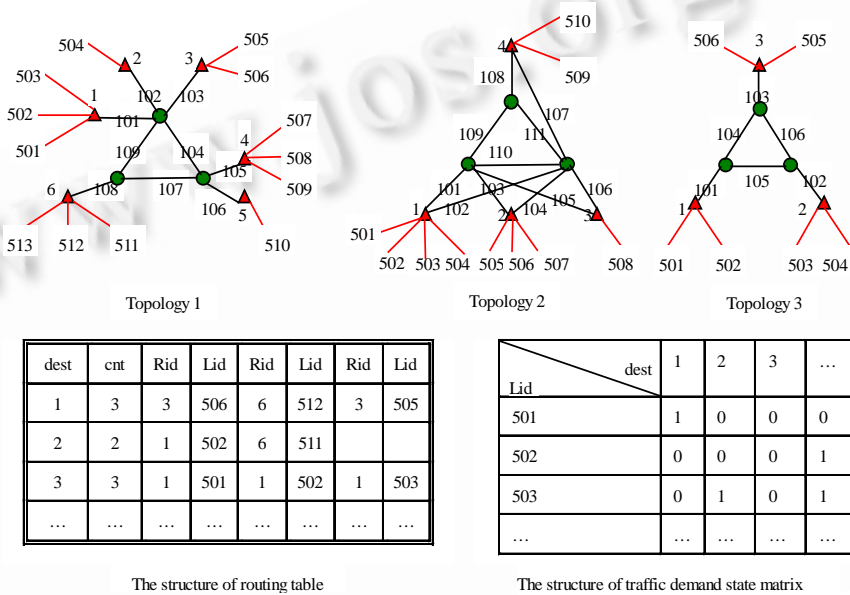
- [1] Quoitin B, Uhlig S, Pelsser C, Swinnen L, Bonaventure O. Interdomain traffic engineering with BGP. *IEEE Communications Magazine*, 2003,41(5):122-128.
- [2] Feamster N, Borkenhagen J, Rexford J. Guidelines for Interdomain traffic engineering. *ACM SIGCOM Computer Communications Review*, 2003,33(5):19-30.
- [3] Awduche D, Chiu A, Elwalid A, Widjaja I, Xiao X. Overview and principles of Internet traffic engineering. Internet Engineering Task Force, RFC3272, May 2002.
- [4] ReKhter Y, Li T. A bodergGateway protocol. Internet Engineering Task Force, RFC1771, March 1995.
- [5] CISCO. BGP best path selection algorithm. 2004. <http://www.cisco.com/warp/public/459/25.pdf>
- [6] Bressoud TC, Rastogi R. Optimal configuration for BGP route selection. In: Proc. of the IEEE INFOCOM2003. San Francisco: IEEE, 2003. 916-926. http://www.ieee_infocom.org/2003/papers/23_01.pdf
- [7] Ye T, Kalyanaraman S. A recursive random search algorithm for large-scale network parameter configuration. In: Proc. of the ACM SIGMETERICS 2003. San Diego: ACM, 2003. <http://www.ecse.rpi.edu/Hompages/shivkuma/research/papers/ols-j.pdf>
- [8] Uhlig S, Bonaventure O, Quoitin B. Interdomain traffic engineering with minimal Bgp configurations. In: Proc. of the 18th International Teletraffic Congress. Berlin: ITC, 2003. http://www.info.ucl.ac.be/people/OBO/papers/itc18_paper030.pdf
- [9] Akella A, Seshan S, Shaikh A. An empirical evaluation of wide-area Internet bottlenecks. In: Proc. of the ACM SIGMETERICS 2003. San Diego: ACM, 2003. 316-327. <http://www.icir.org/vern/imc-2003/papers/p303-akella.pdf>
- [10] Liu H, Bai D, Ding W, Zeng ZM. Internet routing optimization control algorithm with multiple objectives. *ACTA Electronica Sinica*, 2004,32(2):306-310 (in Chinese with English Abstract).
- [11] Fraleigh C, Tobagi F, Diot C. Provisioning IP backbone networks to support latency sensitive Traffic. In: Proc. of the IEEE INFOCOM 2003. San Francisco: IEEE, 2003. http://www.ieee_infocom.org/2003/papers/10_01.pdf
- [12] Cormen TH, Leiserson CE, Rivest R, Stein C. Introduction to Algorithms. 2nd ed, Beijing: Higher Education Press, The MIT Press, 2002.
- [13] Mortier R, Pratt I. Incentive based inter-domain routing. In: Proc. of the Internet Charging and QoS Technology Workshop (ICQT 2003). ACM, 2003. 308-317. http://research.microsoft.com/users/mort/publications/icqt_bgppricing.pdf
- [14] <http://www.cs.bu.edu/brite>. 2002.

附中文参考文献:

[10] 刘红,白栋,丁炜,曾志民.多目标的 Internet 路由优化控制算法.电子学报,2004,32(2):306-310.

附录 模拟实验

模拟实验所需的拓扑图由拓扑产生器 Brite^[14]生成,由于需要验证预测算法的准确性,因此使用的拓扑图较简单.在模拟实验中使用了 3 个拓扑图(如图 4 所示).拓扑图中,三角形表示 BGP 边界路由器,圆形表示域内路由器,用大于 500 的数字表示域间链路标识,用 100~400 间的数字表示域内链路标识.路由表的结构是第 1 列表示目的网络,用 dest 表示;第 2 列表示可选路径的个数,用 cnt 表示;其后的列分别表示出口路由器,用 Rid 表示;出口链路用 Lid 表示;可选路径的个数是 2~4 之间的随机数,出口路由器、出口链路随机生成.流量需求矩阵状态表表示当前有哪些流,其中列表示目的网络、行表示源标识,流量需求矩阵状态表随机生成.流量需求矩阵根据流量需求矩阵状态表随机生成对应流上的流量(示例如图 4 所示).



The structure of routing table

The structure of traffic demand state matrix

Fig.4 The environment of simulation

图 4 模拟环境

附录1 瓶颈区域预测准确度比较实验

采用拓扑图 1 为当前实验所用的拓扑图,分别进行了 3 个实验:实验 1 中所有域间链路的带宽为 100M,所有域内链路的带宽为 125M.实验 2 中所有域间链路的带宽为 100M,所有域内链路的带宽为 250M.实验 3 中所有域间链路的带宽为 200M,所有域内链路的带宽为 100M.表 5~表 7 表示对应 3 个实验的模拟结果.

Table 5 The simulation result about precision of the algorithm (I)

表 5 算法准确度实验的模拟结果(I)

State	Count	e_state	Count	e_state	Count	Precision of e_state	Precision of bottleneck area
0	0	-	-	-	-	-	-
1	119	1	78	2	41	0.655	0.987
2	738	1	199	2	539	0.730	1.000
3	142	3	139	4	3	0.978	0.996
4	3	3	2	4	1	0.333	1.000
5	0	-	-	-	-	-	-
Average precision of e_state						0.756	
Average precision of bottleneck area						0.998	

Table 6 The simulation result about precision of the algorithm (II)

表 6 算法准确度实验的模拟结果(II)

State	Count	e_state	Count	e_state	Count	e_state	Count	Precision of e_state	Precision of bottleneck area
0	791	0	730	1	59	2	0	0.924	0.987
1	209	1	209	2	0	-	-	1.000	0.617
Average precision of e_state								0.940	
Average precision of bottleneck area								0.910	

Table 7 The simulation result about precision of the algorithm (III)

表 7 算法准确度实验的模拟结果(III)

State	Count	e_state	Count	e_state	Count	e_state	Count	Precision of e_state	Precision of bottleneck area
5	1000	5	1000	3	0	4	0	1.000	1.000
Average precision of e_state								1.000	
Average precision of bottleneck area								1.000	

附录2 瓶颈区域概率与网络状态关系的实验

分别采用拓扑图 1 和图 2 进行实验,表 8 是针对拓扑图 1 的模拟结果,表 9 是针对拓扑图 2 的模拟结果.其中 in 表示瓶颈链路在域内的频度,out 表示瓶颈链路在域间的频度,in_out 表示域间和域内链路同时发生瓶颈的频度,crossing ratio 表示在坐标轴上相交线段的比例.

Table 8 The simulation result 1 about the relation between bottleneck area and network state

表 8 瓶颈区域概率与网络状态关系模拟实验结果 1

State	in	out	in_out	Crossing ratio
1	0.213 456	0.785 805	0.000 739	0.654 601
1	0.244 499	0.754 384	0.001 117	0.496 743
1	0.144 578	0.852 503	0.000 000	0.452 701
1	0.000 000	1.000 000	0.000 000	0.146 365
2	0.861 393	0.138 101	0.000 506	0.301 950
2	0.111 931	0.886 625	0.001 444	0.458 937
3	0.944 385	0.000 000	0.055 615	0.395 235
3	0.805 112	0.193 686	0.001 202	0.488 723
4	0.593 750	0.406 250	0.000 000	0.914 634

Table 9 The simulation result 2 about the relation between bottleneck area and network state

表 9 瓶颈区域概率与网络状态关系模拟实验结果 2

State	in	out	in_out	Crossing ratio
1	0.039 397	0.958 000	0.000 200	0.312 977
1	0.007 218	0.992 65	0.000 132	0.179 728
1	0.000 000	1.00 0000	0.000 000	0.101 779
1	0.000 000	1.000 000	0.000 000	0.023 441
1	0.000 000	1.000 000	0.000 000	0.009 226
2	0.532 723	0.466 184	0.001 093	0.458 937
2	0.124 759	0.872 200	0.000 003	0.261 614
2	0.020 936	0.978 656	0.000 408	0.195 682
2	0.000 549	0.999 359	0.000 092	0.159 341
3	1.000 000	0.000 000	0.000 000	0.345 983

附录3 流量单调变化的实验

采用拓扑图 3 进行模拟实验.

实验 1:链路容量取值:全部随机取值,初始流量矩阵分别取轻载、中载、重载,然后增加某一条流量的值.模拟结果是未发生状态变化.这是因为拓扑图的近似对称及路由表的对称性对瓶颈区域分布的均匀性起主要因素.

实验 2:将外部链路容量取值为 155M;内部链路容量取值为 100M;初始流量矩阵取值与上述类似.模拟结果是轻载时易发生状态变化.

附录4 流量随机变化的实验

分别采用拓扑图 1 和图 2 进行模拟实验,图 5 表示采用拓扑图 2 进行模拟的结果,图 6 表示采用拓扑图 1 进行模拟的结果,其中实线表示最大链路利用率的变化,虚线表示状态变化,out 与 in 之比表示外部链路容量与内部链路容量的比值,横坐标表示随机变化的次数,纵坐标表示最大链路利用率的值以及状态的取值.

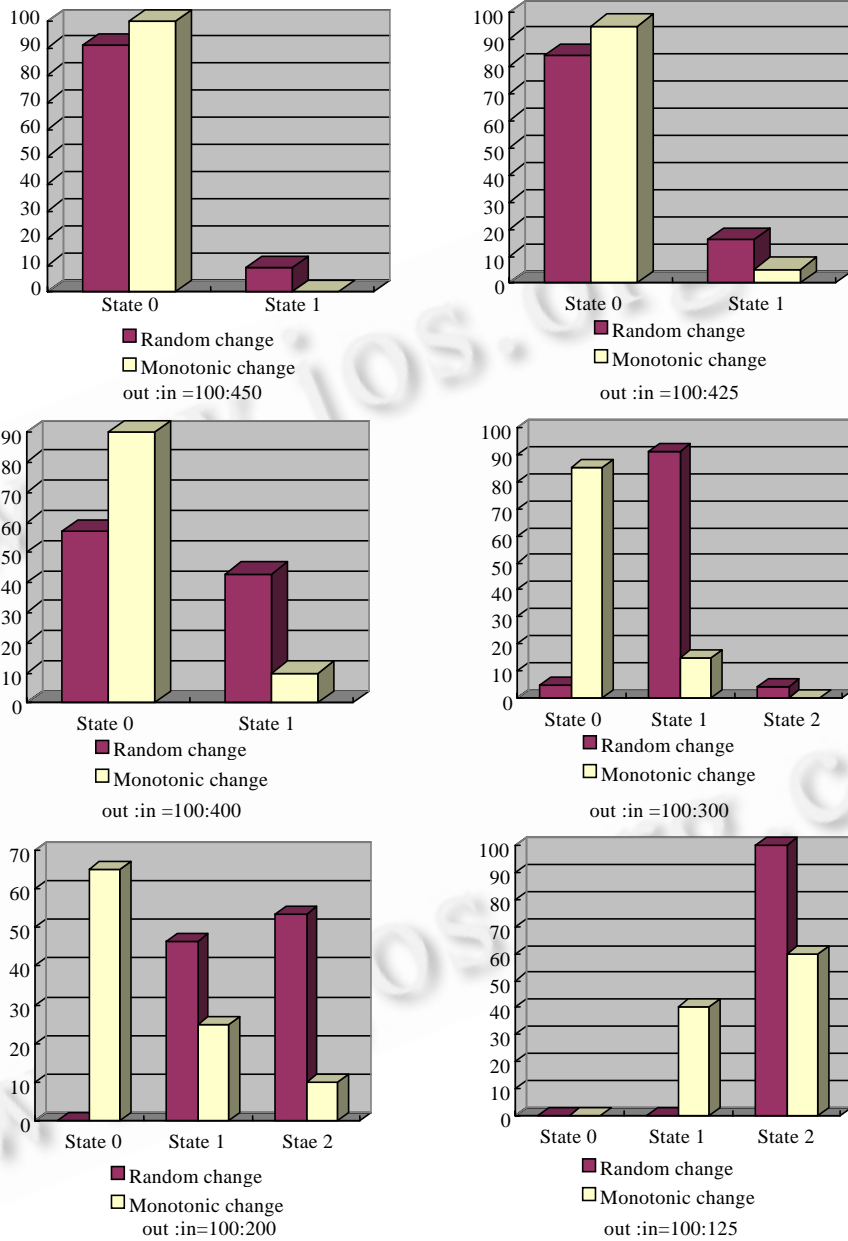


Fig.5 The simulation result about traffic with random change

图 5 流量随机变化模拟结果

附录5 拓扑变化的实验

分别采用 3 个拓扑图进行模拟实验,模拟结果表见 10,并分别对应拓扑图 1~3.其中 out:in 表示域间链路与域内链路容量之比, s 表示状态, f 表示频度, r 表示域内和域间链路容量在 2500,10000,622,1000,100,155 中随机取值.

Table 10 The simulation result about topology change

表 10 拓扑变化模拟实验

out:in	s/f	s/f	s/f	s/f	s/f	s/f
100:300	0/0	1/0	2/15.6%	3/5.4%	4/0	5/77.6%
100:100	0/0	1/0	2/0	3/89.6%	4/0	5/10.4%
200:100	0/0	1/0	2/0	3/56.9%	4/0	5/43.1%
r	0/0	1/0	2/10.9%	3/89.1%	4/0	5/0
100:450	0/91%	1/9%	2/0	3/0	4/0	5/0
100:425	0/84%	1/16%	2/0	3/0	4/0	5/0
100:400	0/57%	1/43%	2/0	3/0	4/0	5/0
100:300	0/5%	1/91%	2/4%	3/0	4/0	5/0
100:250	0/0	1/85%	2/15%	3/0	4/0	5/0
100:200	0/0	1/46.4%	2/53.6%	3/0	4/0	5/0
100:125	0/0	1/0	2/100%	3/0	4/0	5/0
100:100	0/0	1/0	2/100%	3/0	4/0	5/0
125:100	0/0	1/0	2/0	3/100%	4/0	5/0
200:100	0/0	1/0	2/0	3/43%	4/32%	5/5%
400:100	0/0	1/0	2/0	3/0	4/0	5/100%
r	0/0	1/14.9%	2/85.1%	3/0	4/0	5/0
100:300	0/5%	1/95%	2/0	3/0	4/0	5/0
100:200	0/0	1/73%	2/27%	3/0	4/0	5/0
100:150	0/0	1/71%	2/29%	3/0	4/0	5/0
100:125	0/0	1/4%	2/78%	3/16%	4/2%	5/0
100:100	0/0	1/0	2/0	3/97%	4/0	5/3%
250:100	0/0	1/0	2/0	3/0	4/0	5/100%
r	0/0	1/22%	2/78%	3/0	4/0	5/0