一种负载均衡网络中内部链路时延推测算法*

焦 利+, 林 宇, 王文东, 金跃辉, 程时端

(北京邮电大学 网络与交换国家重点实验室,北京 100876)

A Novel Algorithm for Link Delay Inference in the Networks with Load-Balancing Routing

JIAO Li+, LIN Yu, WANG Wen-Dong, JIN Yue-Hui, CHENG Shi-Duan

(National Key Laboratory of Networking and Switching, Beijing University of Posts and Telecommunications, Beijing 100876, China)

+ Corresponding author: Phn: +86-10-62282007, 13910743515, E-mail: jiaoli@bupt.edu.cn, http://www.bupt.edu.cn

Received 2004-04-10; Accepted 2004-11-15

Jiao L, Lin Y, Wang WD, Jin YH, Cheng SD. A novel algorithm for link delay inference in the networks with load-balancing routing. *Journal of Software*, 2005,16(5):886−893. DOI: 10.1360/jos160886

Abstract: Engineering a large IP backbone network without a view of internal link state is challenging. Previous algorithms assume that probes experience fixed routes in networks. As there are load-balancing equipments in networks, probes are delivered across random routes. This results in invalidation of the prevous algorithms. New algorithm proposed in this paper uses CGF (cumulate generating function) to infer delay characteristics of the internal link under stochastic routes. Simulation results prove that the algorithm could resolve the delay inference in the networks with load-balancing route. Based on the delay characteristics of the internal link, the bottleneck link can be located.

Key words: network measurement; delay inference; load-balancing; cumulate generating function (CGF)

摘 要: 了解网络内部链路特征对运维大型 IP 网络至关重要.前人在假定固定路由条件下采用端到端主动测量的方式从网络边缘推测网络内部链路行为特征.由于网络中存在导致随机路由的负载均衡设备,使以前的主动测量方法无法实施.采用累计生成函数和随机过程方法解决随机路由条件下的网络内部链路时延推测问题.仿真结果表明,算法可以很好地解决随机路由下的内部链路时延推测问题.根据链路时延分布,可以用来判决瓶

* Supported by the National Natural Science Foundation of China under Grant No.90204003 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant Nos.2002AA103063, 2003AA121220 (国家高技术研究发展计划(863)); the National Grand Fundamental Research 973 Program of China under Grant No.2003CB314806 (国家重点基础研究发展规划(973)); the National Research Foundation for the Doctoral Program of Ministry of Education of China under Grant No.20010013003 (国家教育部博士点基金)

作者简介: **焦利**(1974−),男,河北邯郸人,博士生,主要研究领域为互联网测量,IP 网络管理,包交换网络服务质量;**林宇**(1976−),男,博士,副教授,主要研究领域为 Internet 测量,无线网络,移动组播,TCP 建模;**王文东**(1963−),男,教授,主要研究领域为网络和业务的服务质量(QoS)管理,高速网络(新一代互联网 NGI 和下一代网络 NGN)上实时业务及其协议的研究;**金跃辉**(1965−),女,副研究员,主要研究领域为网络性能测量技术,传输层协议及优化机制,IP 网的服务质量和性能评估体系,移动 IP 技术;**程时端**(1940−),女,教授,博士生导师,主要研究领域为 IP 网络流量控制与拥塞控制,移动 IP 技术,组播技术,移动互联网性能分析,服务质量控制.

颈链路,为网络运维提供极具价值的参考.

关键词: 网络测量;时延推测;负载均衡;累积生成函数

中图法分类号: TP393 **文献标识码**: A

网络性能测量在网络运维管理中占有重要的地位,它是进行性能统计、预测和故障诊断的基础.根据是否发送测量探针,网络测量可以分为被动测量[1]和主动测量[2−9]两种.探针是由同一信源发送的携带探测信息的分组序列.主动测量通过向网络发送大量探针,并根据探针所携带的信息来推测网络的情况.另外,根据发送探针分组的类型,还可将主动测量划分为组播探针[5−7]和单播探针[2,4]两种类型.被动测量不向网络发送探针,而是通过监听网络中的分组流来推测网络情况,它不影响网络负荷,但灵活性受到网络运维和技术的限制.根据测量实施的位置,被动测量可分为基于路由器的测量(router-based measurement,简称 RBM)和端到端测量(end-to-end measurement,简称 E2EM,).RBM 主要由路由器中的管理软件来完成内部网络的测量.E2EM 是通过边缘主机的协作来获取网络性能统计,并尽可能地减小对网络运行的影响.ISP(internet service provider)通常采用 RBM 来监测其内部网络.但是,由于各 ISP 之间非协作性,网络内部性能数据一般对外保密;另外,将大量路由器统计的数据传递给网管系统,本身就增加了较大的网络负荷.因此使得对于链路性能监测,在许多场合下更倾向于采用 E2EM 方案.

网络内部链路时延/抖动特征是网络运维的重要参数指标,也是网络测量技术研究的主要指标.网络时延特征和网络吞吐率有着必然的联系,Power 公式很好地描述了二者间的关系: $power=\frac{throuthput^a}{Delay}, 0<a<1$.

现有的时延/抖动推测的工作主要包括:(1) 将连续时延分布进行离散化(或量化)[9,10],使得时延推测转化为便于计算的矩阵操作.这种方法简单,但由于对时延分布进行了量化操作,使得时延推测的精度有所损失;(2) 文献[11]提出非参数化的连续时延分布(nonparametric density estimation)推测方法;(3) 文献[8]提出利用累计生成函数(cumulate generating function)推测时延分布的方法,文献[12]进一步研究了在时延非平稳情况下的时变时延分布推测方法;(4) 文献[13]通过端到端时延抖动来推测链路的时延抖动.

遗憾的是,这些推测方法[8−13]都假定测量探针在网络中总是经历固定路由,而实际网络中常常存在随机路由的情况.比如由于采用负载均衡路由技术或采用负载均衡设备会使同一个测试探针流内的分组在网络中经历不同的路由.在这种情况下,文献[8−13]的算法无法获得正确的测量结果.本文首先建立存在随机路由的网络模型,在此模型基础上建立网络时延推测的数学模型和推测算法,并通过网络仿真进行验证.仿真结果表明,本算法能够很好地适应随机路由条件下的网络内部链路时延推测.

本文第 1 节介绍基本概念.第 2 节描述随机路由网络模型和网络时延推测模型.第 3 节通过仿真来验证第 2 节提出的推测模型,并针对分析结果给出一种应用实例.第 4 节总结全文.

1 基本概念

分组通过一条网络路径的端到端时延包括 3 部分:(1) 排队时延,依赖于路由器缓冲排队所需的时间;(2) 发送时延,依赖于分组包长和链路速率;(3) 电磁波在传输媒介的传播时间.利用 GPS 可实现高精度的端到端时延测量;如果没有类似 GPS 的时钟同步机制,由于测量主机间存在时钟频率偏差[14],一般只能获得排队时延.实际上,定长分组经过一条广域网的网络路径所需的发送时延和传播时延是较小的常数,端到端的时延主要决定于排队时延,并且链路忙闲程度也主要体现为排队时延.文献[8−13]中研究的时延推测也主要针对排队时延.

假定要监测的网络由 m 条内部链路构成,探针沿 n 条路径通过网络.假定 M_i 是第 i 条探针经过路由所包含的链路集合, $i=1,...,n$.假定已知每个探针的路由,则可得探针路由矩阵 $R=(r_{ij})_{n\times m}$,其中, $r_{ij}\in\{0,1\}$ 表示路径和链路间的拓扑关系, $j\in M_i$.当路径 i 包含链路 j 时, $r_{ij}=1$,否则为 0.显然,第 i 次探针经过一条路径的端到端时延 Y_i 等于路径上各链路时延之和:

$$Y_i=r_{i_1}X_{i_1}+...+r_{i_m}X_{i_m}, i=1,...,n \tag{1}$$

其中, X_{ij} 为第i次探针通过链路j时的时延.一般假定链路时延空间独立且分布平稳:链路时延 X_{ij} 相互独立,

$i=1,...,n, j\in M_i$; $\{X_{ij}\}_{i=1}^{n}$ 为随机变量 X_j (链路j的时延)的n个独立同分布的样本.文献[10]采用随机变量 Y_i 的累计生成函数(CGF),

$$K_{Y_i}(t)=\log E[e^{tY_i}],\quad i=1,...,n,\ t\in(-\infty,\infty) \tag{2}$$

对式(1)进行变换,形式如下:

$$K_{Y_i}(t)=\log E[e^{tY_i}]=\log E[e^{t(r_{i1}X_{i1}+...+r_{im}X_{im})}]=\sum_{j=1}^{m}r_{ij}K_{X_j}(t) \tag{3}$$

定义 $K_Y(t)=[K_{Y_1}(t),...,K_{Y_m}(t)]^T$, $K_X(t)=[K_{X_1}(t),...,K_{X_m}(t)]^T$,则

$$K_Y(t)=RK_X(t) \tag{4}$$

其中, $K_{X_i}(t)=\log\int_{-\infty}^{\infty}e^{tx}p_{X_i}(x)\mathrm{d}x$ 与随机变量 X_i 的概率密度分布 $p_{X_i}(x)$ 一一对应,因此,当 R 为满秩矩阵时,可通过 $K_Y(t)$ 推测 $K_X(t)$,还可进一步获得链路时延分布特征.

遗憾的是,文献[8−13]都假定测量探针在网络中总是经历固定路由(此时,R 是常量矩阵),但是在实际网络中,因为随机路由的存在,使各种算法无法获得准确的推测结果.

2 随机路由下的网络时延推测算法

当网络中存在负载均衡设备时,探针将在该设备处以一定的概率经过此设备之后的各路由,假定 M_i 是第 i 条探针可能经过的路由所包含的链路集合, $i=1,...,n$.假定探针 i 通过链路 j 的概率 p_{ij} 已知,由此构成负载均衡概率矩阵 $P=(p_{ij})_{n\times m}$,则第 i 次探针所经历的端到端时延是一个随机变量:

$$Y_i=\sum\nolimits_{j\in M_i}X_{ij}\cdot p_{ij} \tag{5}$$

利用 CGF 可推导如下:

$$\left.\begin{aligned}K_{Y_i}(t)&=\log E[e^{tY_i}]=\log E[e^{t(\sum_{j\in M_i}X_{ij}\cdot p_{ij})}]=\log\left\{\prod_{j\in M_i}E[e^{tX_{ij}}]^{p_{ij}}\right\}\\&=\sum_{j\in M_i}p_{ij}\cdot\log E[e^{tX_{ij}}]=\sum_{j=1}^{m}r_{ij}\cdot p_{ij}\cdot K_{X_j}(t)=Z(i)\cdot K_X(t)\end{aligned}\right\} \tag{6}$$

其中, $Z(i)$ 表示矩阵Z的第i行,矩阵 $Z=(z_{ij})_{n\times m}$ 是路由矩阵R和负载均衡的概率矩阵P的复合矩阵,

$$z_{ij}=r_{ij}\cdot p_{ji} \tag{7}$$

定义链路和端到端时延分布的 CGF 为 $K_X(t)=[K_{X_1}(t),...,K_{X_m}(t)]^T$, $K_Y(t)=[K_{Y_1}(t),...,K_{Y_m}(t)]^T$,则可通过线性方程表示:

$$K_Y(t)=Z\cdot K_X(t) \tag{8}$$

当 $n\geq m$ 时,Z 是满秩矩阵,则式(8)是可逆的,所以 $K_X(t)$ 可以通过 $K_Y(t)$ 按下式计算得到:

$$K_X(t)=(Z^TZ)^{-1}Z^TK_Y(t) \tag{9}$$

令 $B=(Z^TZ)^{-1}Z^T$,则:

$$K_{X_j}(t)=\sum_{i=1}^{n}b_{ji}K_{Y_i}(t) \tag{10}$$

当 Z 不满秩时,从式(8)只能得到链路时延线性组合的 CGF,它们对应于某个网络区域的时延性能.因此要推测网络内部每条链路的 CGF,需选择不同的探针路径覆盖网络,使 $n\geq m$,保证矩阵 Z 满秩.

下面说明如何通过测量数据计算 $K_Y(t)$.令 N_i 是路径 i 上收集到的探针数, $i=1,...,n$,设 Y'_{ik} 是路径 i 上的第 k 个探针测得的时延.令 Y_{ik} 表示路径上各链路的累计排队时延,.一般假定在数据集合 $\{Y'_{ik},k=1,...,N_i\}$ 中最小的时延是分组在路径各个链路上都没有排队的时延,因此端到端的排队时延可以近似为

$$Y_{ik}=Y'_{ik}-\min\{Y'_{ik},k=1,...,N_i\} \tag{11}$$

定义 $M_{Y_i}(t)=e^{K_{Y_i}(t)}$ 的估值为

$$\hat{M}_{Y_i}(t)=\frac{1}{N_i}\sum_{k=1}^{N_i}e^{tY_{ik}} \tag{12}$$

这里, $\hat{M}_{Y_i}(t)$ 是矩量母函数(moment generating function) $M_{Y_i}(t)=e^{K_{Y_i}(t)}$ 的无偏差估值.这样,可采用最小二乘法(LS)从向量 $\hat{M}_{Y_i}(t)=[\hat{M}_{Y_1}(t)...\hat{M}_{Y_n}(t)]^T$ 中得到向量 $K_Y(t)$ 的估值.这给我们一个假象:好像可以如下式用矩法(method-of-moments,简称 MoM)来估计 $K_{X_j}(t)$:

$$\hat{K}'_{X_j}(t)=\sum_{i=1}^{n}b_{ji}\hat{M}_{Y_i}(t) \tag{13}$$

但是因为 log 函数的非线性,这种估值方法将引入较大偏差.我们采用线性化方法[9]来校正式(13)中 $K_{X_j}(t)$ 的估值偏差:

$$\left.\begin{aligned}\hat{K}'_{X_j}(t)&=\log\left\{\prod_{i=1}^{n}(\hat{M}_{Y_i}(t))^{b_{ji}}\right\}\\&=\log\left\{\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)]-\left(\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)]-\prod_{i=1}^{n}(\hat{M}_{Y_i}(t))^{b_{ji}}\right)\right\}\\&=\log\left\{\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)]\left(1-\left(1-\frac{\prod_{i=1}^{n}(\hat{M}_{Y_i}(t))^{b_{ji}}}{\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)]}\right)\right)\right\}\\&=\log\left\{\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)]\right\}+\log\left\{\left(1-\left(1-\frac{\prod_{i=1}^{n}(\hat{M}_{Y_i}(t))^{b_{ji}}}{\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)]}\right)\right)\right\}\\&=K_{X_j}(t)+\log(1-\omega_j)=K_{X_j}(t)-\omega_j-\frac{1}{2}\omega_j\end{aligned}\right\} \tag{14}$$

其中, $\omega_j=\left(1-\frac{\prod_{i=1}^{n}(\hat{M}_{Y_i}(t))^{b_{ji}}}{\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)]}\right)$.

可以得到式(13)误差校正后的估值:

$$\hat{K}_{X_j}(t)=\sum_{i=1}^{n}b_{ji}\log(\hat{M}_{Y_i}(t))+\hat{E}[\omega_j]+\frac{1}{2}\hat{E}[\omega_j^2] \tag{15}$$

其中, $\hat{E}[\bullet]$ 表示我们用矩法估计的经验平均:

$$\hat{E}[\omega_j]=1-\frac{\prod_{i=1}^{n}\hat{E}[(\hat{M}_{Y_i}(t))^{b_{ji}}]}{\hat{M}_{X_j}(t)} \tag{16}$$

$$\hat{E}[\omega_j^2]=1-\frac{2\prod_{i=1}^{n}\hat{E}[(\hat{M}_{Y_i}(t))^{b_{ji}}]}{\hat{M}_{X_j}(t)}+\frac{\prod_{i=1}^{n}\hat{E}[(\hat{M}_{Y_i}(t))^{2b_{ji}}]}{\hat{M}^2_{X_j}(t)} \tag{17}$$

$\hat{M}_{X_j}(t)$ 是链路 j 时延的矩量母函数,

$$\hat{M}_{X_j}(t)=\prod_{i=1}^{n}E^{b_{ji}}[\hat{M}_{Y_i}(t)] \tag{18}$$

我们采用滑动窗口法[10]求得 $E[(\hat{M}_{Y_i}(t))^{b_{ji}}]$,窗口尺度为 W,滑动步长为 S,定义 $N_w=\frac{N_i-W}{S}$ 为窗口增量,则:

$$E[(\hat{M}_{Y_i}(t))^{b_{ji}}]=\sum_{l=1}^{N_w}\frac{1}{N_w}\left(\frac{1}{W}\sum_{k=(l-1)S+1}^{(l-1)S+W}e^{ty_{ik}}\right)^{b_{ji}} \tag{19}$$

用同样的方法可以求得 $E[(\hat{M}_{Y_i}(t))^{2b_{ji}}]$:

$$E[(\hat{M}_{Y_i}(t))^{2b_{ji}}]=\sum_{l=1}^{N_w}\frac{1}{N_w}\left(\frac{1}{W}\sum_{k=(l-1)S+1}^{(l-1)S+W}e^{ty_{ik}}\right)^{2b_{ji}} \tag{20}$$

这样,我们就可以通过测量的端到端时延结果数据集合获得各个链路时延分布的累计生成函数,进一步可获得各个链路时延的分布函数.

3 仿真结果

为验证第 2 节中所述算法的有效性,我们采用 OPNET 建立如图 1 所示的网络模型,网络链路 L_2,L_3,L_4,L_5 的带宽为 4Mb/s,延迟为 50ms;边缘链路 L_1 设计为 2Mb/s,时延为 50ms;边缘链路 L_6 设计为 3Mb/s,时延为 10ms.链路采用 Drop-Tail 模型(有限长缓冲的 FIFO 队列),队列缓冲为 150 个包,探针为 40 字节的 UDP 包,探针在相应的源端以泊松过程产生,平均到达时间为 16ms,速率为 20Kb/s.背景流量采用指数到达的 ON-OFF 模型的 UDP 流和 FTP 流.

在路由器 R_1 处采用路由负载均衡,两条路由 $R_1\rightarrow R_2\rightarrow R_4$ 和 $R_1\rightarrow R_3\rightarrow R_4$ 分担的流量比例为 1:1,所有的背景流和探针流在 R_1 处以概率 0.5 分布在这两条路径上.我们通过 7 条路径发送探针,按序号排列的路由和相应的探针随机路由矩阵分别为

(1) $R_0\rightarrow R_1\rightarrow R_2/R_3\rightarrow R_4\rightarrow R_6$;
(2) $R_0\rightarrow R_1\rightarrow R_2$;
(3) $R_0\rightarrow R_1\rightarrow R_3$;
(4) $R_0\rightarrow R_1\rightarrow R_2/R_3\rightarrow R_4$;
(5) $R_1\rightarrow R_2/R_3\rightarrow R_4$;
(6) $R_3\rightarrow R_4\rightarrow R_6$;
(7) $R_2\rightarrow R_4\rightarrow R_6$.

$$Z=\begin{array}{c} \\ P_1\\P_2\\P_3\\P_4\\P_5\\P_6\\P_7\end{array}\begin{array}{c}\begin{array}{cccccc}L_1&L_2&L_3&L_4&L_5&L_6\end{array}\\\begin{pmatrix}1&0.5&0.5&0.5&0.5&1\\1&1&0&0&0&0\\1&0&1&0&0&0\\1&1&0&1&0&0\\0&0.5&0.5&0.5&0.5&0\\0&0&0&0&1&1\\0&0&0&1&0&1\end{pmatrix}\end{array} \tag{21}$$

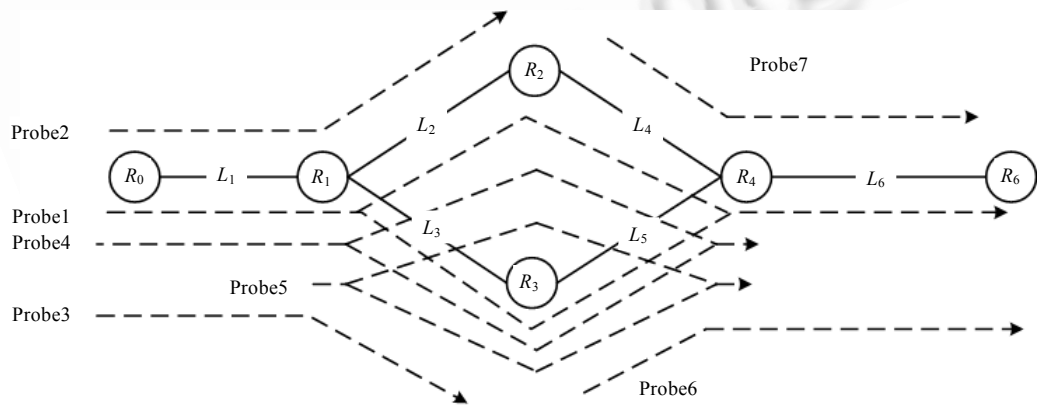

Fig.1 Route of probes (load-balancing in R_1)
图 1 探针路由(在 R_1 处负载均衡)

每个路径采集 N 个探针,共采集 $7\times N$ 个探针,我们通过减去 N 个探针时延中的最小值来估算每个探针的排

队时延.这是对排队时延的带偏差的估值,因为探针时延的最小值是探针发送时延加传输时延带偏差的估值,偏差以 1/N 递减.为了求得式(19)、式(20)的期望值,我们设定窗口尺寸 W=2N/3,步长 S 为 10 个探针时延采样.取 N=1500,–200<t<200.

图 2 对比了采用仿真统计的链路 2 的时延数据直接计算的链路时延 CGF 累计生成函数与通过端到端时延统计数据推测出来的链路时延 CGF 函数的结果(包括式(13)所示的有偏估计和式(20)所示的无偏估计).从图 2 中可以看出,推测的 CGF 与实际的 CGF 非常接近,而无偏估计比有偏估计更为精确.类似地,图 3 比较了链路 4 时延的仿真统计结果和推测结果的比较.同样,两种估计都较好地与仿真直接统计结果相吻合,并且无偏估计优于有偏估计.

我们还进一步比较了利用仿真直接统计计算出来的链路时延的均值和方差与通过推测链路时延 CGF 获得的链路时延均值和方差,结果如图 4 所示.从图 4 中可以看出:根据仿真直接统计得出链路时延均值和方差与推测出的链路时延的均值和方差符合得很好.表 1 列出了两种方法结果的均方误差(mean square error,简称 MSE),可以看到偏差校正后误差较小.

实际上,这种时延推测技术可以被应用于瓶颈链路的检测.因为链路的 CGF 是链路时延概率密度函数的傅里叶变换的对数,它保留了链路时延的统计信息,我们可以利用 CGF 估算时延分布的多项特性.可以利用 Chernoff bound 公式来判决瓶颈链路.由 Chernoff bound 公式:

$$P(X_j \geq \delta) \leq e^{-t\delta} E[e^{tX_j}] = P_j \tag{22}$$

我们定义瓶颈为链路时延超过某个时延阈值δ的概率超过预定义的概率阈值 P,通过适当选取时延阈值δ和接近 1 的概率阈值 P,按照 $\mathrm{Max}_{j=1,\ldots,m}P_j>P$ 即可挑选瓶颈链路.

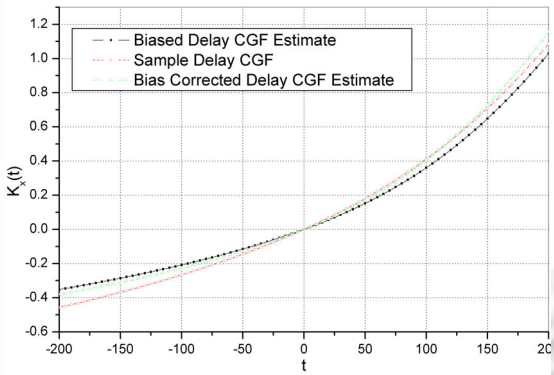

Fig.2 CGF of delay of link 2

图 2 链路 2 时延 CGF

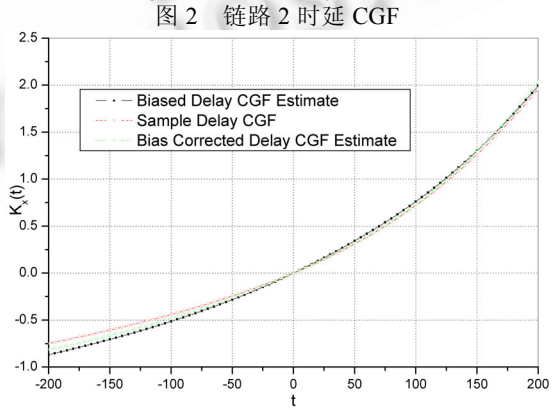

Fig.3 CGF of delay of link 4

图 3 链路 4 时延 CGF

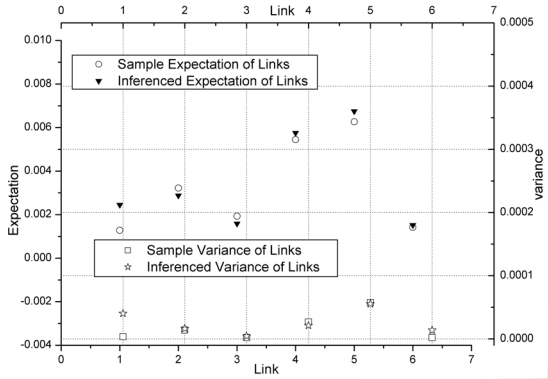

Fig.4 Expectation and variance of each link
图4 各链路时延的均值和方差

在图 4 的仿真环境下,我们在链路 4 上部署较大的背景业务量,通过收集各探针的端到端时延数据获得各个链路的 Chernoff bound 概率(见表 2).若令δ=0.02s,设定 P=0.90,也就是定义“链路时延超过 0.02s 的概率最少为 0.90”为瓶颈链路的判据,即 P_j≥0.90,我们可以确定链路 4 为瓶颈,这和实际仿真时背景业务量的设置完全吻合.

Table 1 Mean square error of $\hat{K}_{X_j}$ (bias-corrected) and $\hat{K}'_{X_j}$ (biased)

表1 $\hat{K}_{X_j}$ (偏差校正)和 $\hat{K}'_{X_j}$ (无校正)的均方误差

Link	1	2	3	4	5	6
MSE of $\hat{K}_{X_j}$	1.827 5	0.002 916	0.005 174	2.121 1	0.417 82	0.443 15
MSE of $\hat{K}'_{X_j}$	2.180 6	0.005 924	0.005 924	2.057 1	0.496 29	0.508 8

Table 2 Probability of Chernoff bound P_j

表2 Chernoff bound概率P_j

Link	1	2	3	4	5	6
P_j	0.005 457	0.061 28	0.727 53	0.998 3	0.589 55	0.825 52

4 结 论

现有的推测方法[8−13]都假定测量探针在网络中总是经历固定路由,而实际网络中存在随机路由的情况使得算法失效.本文着重考察随机路由网络环境下的内部链路时延特征的推测方法.我们利用链路时延分布的累积生成函数,提出了新的推测算法,并通过 OPNET 仿真模型进行了验证.仿真结果与推测结果较好地吻合,这表明新的推测算法能够较好地解决存在随机路由环境下的时延推测问题.

本文的推测方法假定在探针发送期间网络时延是平稳分布的.在实际应用中,这个条件很难满足,我们可以将时间划分成一系列时间窗口,通过适当选择时间粒度以保证在时间窗口内网络近似平稳分布,然后来近似解决时变网络链路特征的推测.如果内部链路时延在空间上不独立,这就需要建立更为复杂的模型来解决.

另外,在针对大规模网络的实际测量中,构建满秩矩阵 Z 比较困难(这需要很多探针),此时只能获得一些链路时延性能的线性组合 $\sum_{j=1}^{m} r_j K_{X_j}(t)$.注意到,每一组合项对应着网络的某一测量路径区域(或链路组合),我们可以利用瓶颈检测方法确定拥塞区域,然后再在该区域增加测量路径和探针,构造局部满秩矩阵,获取该区域内各个链路的时延特征.如何有效地分解网络区域,以便减小计算量、提高推测精度是我们未来的研究目标.

References:

[1] Cao J, Davis D, Wiel SV, Yu B. Time-Varing network tomography: Router link data. Bell Laboratory Technical Memo, 2000. http:// plan9.belllabs.com/cm/ms/departments/sia/cao/ht-mls/pub.html

[2] Downey B. Using pathchar to estimate Internet link characteristics. In: Proc. of the ACM SIGCOMM'99. 1999.

[3] Bolot J. End-to-End packet delay and loss behavior in the Internet. In: Proc. of the SigComm'93. 1993. 289−298.

[4] Lai K, Baker M. Measuring link bandwidths using a deterministic model of packet delay. In: Proc. of the ACM SIGCOMM 2000. 2000.

[5] Ziotopoulos A, Hero A. Estimation of network link loss rates via chaining in multicast trees. In: ICASSP 2001. 2001.

[6] Caceres R, Duffield NG, Horowitz J. Multicast-Based inference of network-internal loss characteristics. IEEE Trans. on Information Theory, 1999,45(7):2462−2480.

[7] Presti FL, Duffield NG, Horowitz J. Multicast-Based inference of network-internal delay distributions. 1999. http://wwwnet.cs.umass.edu /minc/

[8] Shih M, Hero A. Unicast inference of network link delay distributions from edge measurements. In: Proc. of the IEEE Int'l Conf. Acoust, Speech, and Signal Processing. Salt Lake City, 2001. 3421−3424.

[9] Tsang Y, Coates M. Passive unicast network tomography based on tcp monitoring. Technical Report TR0004, Rice University, 2000.

[10] Coates M, Nowak R. Networks for networks: Internet analysis using graphical statistical models. In: Proc. of the 2000 IEEE Neural Network for Signal Processing Workshop. Sydney, 2000,2:755−764.

[11] Tsang Y, Coates M. Nonparametric Internet tomography. In: Proc. of the IEEE Int'l Conf. Acoust., Speech, and Signal Processing. Orlando, 2002,2:2045−2048.

[12] Coates M. Sequential monte carlo inference of internal delays in nonstationary communication networks. IEEE Trans. on Signal Processing, 2002,50:366−376.

[13] Duffield N, Presti FL. Multicast inference of packet delay variance at interior network links. In: Proc. of the IEEE INFOCOM 2000. Tel Aviv, 2000,3:1351−1360.

[14] Zhang L, Liu Z, Hong C. Clock synchronization algorithms for network measurements. In: Proc. of the IEEE INFOCOM 2002. New York, 2002,1:160−169.