基于单播丢包层析技术的探测包发送机制①

彭　任，黎文伟

(湖南大学 软件学院，长沙 410082)

摘　要：网络层析技术通过端到端测量就能推测网络的内部性能，无须网络中间节点配合，因此被广泛的应用于网络测量及诊断。对基于单播的网络丢包率层析成像模型进行了详细的探讨，并针对现有的单播层析发包方式的包组相关性缺陷提出了一种改进发包方法。通过 NS2 的仿真，证明了改进的发包机制能更加准确的估算网络丢包率。

关键词：网络层析成像；端到端测量；单播探测包；丢包率

Probe-Sending Methods Based on Unicast Network Loss Tomography

PENG Ren, LI Wen-Wei

(Software School, Hunan University, Changsha 410082, China)

Abstract: Network tomography is a new technique to infer the network internal performance or logic topology only by end to end measurements. In this paper, in order to strengthen the correlation between probes and indeed improve inference accuracy, we propose a new probe-sending methods to get the internal loss rates. NS2 simulations demonstrate the feasibility of this tomography method.

Key words: network tomography; end-to-end measurement; Unicast probes; loss rate

TCP/IP 网络体系结构与协议的巨大成功来源于其开放性，但正是这种开放性使得互联网成为了目前高度异构的复杂系统,给其控制和管理带来了一系列的困难。为了更好的管理、控制和设计网络，就必须运用适当的网络测量技术，准确了解网络的特性和状态参数。网络层析成像技术[1][2]是目前国际学术界备受关注的网络测量技术之一，它是一种基于端到端测量的技术，通过发送多种探测包给指定的接收端，观测并分析接收端所获得的数据，并运用统计学方法来推断各种网络内部参数，包括链路参数和拓扑结构等，从而实现与网络内部结构或协议无关的网络测量。

网络层析技术分为单播测量和多播测量两种方式，多播测量对网络造成的负载相对较小，但并非所有的网络都支持多播，为了弥补多播网络层析成像在现有网络支持上的不足，研究人员开始进行了一些单播网络层析成像的研究[3]，建立了基于单播测量的网络层析成像模型。

本文对包括背靠背包对方式和 3 包组方式单播丢包率层析方法进行了详尽的分析，现有的发包方式在处理测量数据的时候，近似的将包组中的数据包到达分支节点的同一性，即条件概率 β 的值近似为 1 的理想期望值，因而存在由此因素引入的测量误差。为了尽可能的提高丢包率层析的精确度，减小这种误差的影响，本文提出了一种改进的类似于 4 包组的丢包率估算方法来评估网络的丢包性能,并通过仿真实验，评估该方法在估算网络丢包率参数时的有效性。

1　网络丢包层析模型

1.1 网络丢包层析基本模型

首先讨论基本的简单二叉树网络拓扑结构。如图 1 所示，该树的节点 1 对应源节点。节点 3 和 4 对应一组接收节点，中间节点 2 为内部路由器，也是链路 2 和链路 3 的分支节点，链路 1 为数据包从源节点 1 到接收节点 3 和 4 经过的共享链路。在网络层析过程

① 基金项目:国家自然科学基金(60703097)
收稿时间:2011-02-19;收到修改稿时间:2011-03-18

中，中间节点 2 的数据是完全不可见的。

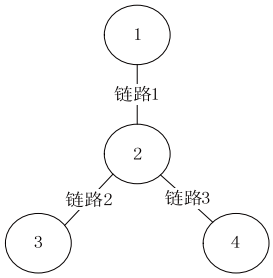


图 1 简单二叉树拓扑

本文定义 T=(V,L)来描述由节点 V 和链路集合 L 构成的逻辑树，节点集 V 由源节点 O，接收节点 D 和内部节点组成。P(O,D_i）表示源节点 O 到接收节点 Di 的路径。假设 O 向 D_i 发送的探测包数为 ni,其中 Di 接收到的探测包数为 m_i,那么 P(O,D_i）上探测包的成功传输概率为 $p_i=m_i/n_i$，定义 a_j 为链路 P(O,D_i）上子链路的成功传输概率，R 为链路数，则有:

$$p_i = \prod_{j=1}^{R} a_j \quad \text{或者} \quad \lg p_i = \sum_{j=1}^{R} a_j \tag{1}$$

通过对链路上探测包成功传输概率的计算，可得到网络路径 P(O,D_i）的丢包率为(1-p_i)。

网络丢包率测量过程中，在网络路径成功传输概率与路径上子链路成功传输概率之间建立方程，可以得到网络丢包层析成像的线性模型:

$$Y = Ax + \varepsilon \tag{2}$$

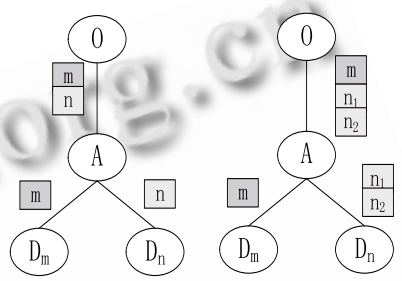
式(2)中 $Y=(Y_1, \cdots, Y_Z)^T$ 是端到端测量得到的 Z 条路径的成功传输概率的对数：$X=(X_l, \cdots, X_F)T$ 是需要估计的 F 条链路的传输概率的对数；A 为路由矩阵，一般是数值为 0 或 1 矩阵，1 表示测量路径经过某链路，否则为 0。ε 为测量数据 Y 产生的加性噪声，也可能是 X 关于其均值产生的随机偏差。如果忽略噪声的影响，则线性模型可写为：

$$Y = Ax \tag{3}$$

1.2 单播层析模型

多播层析技术在许多研究中已有了很好的阐述,本节重点分析单播层析成像。单播层析中使用比较多的是背靠背包对的方法[4][5]和改进的 3 包组方法[6]。

背靠背包对法由两个大小相同的数据包组成一包组(包对)，包组内的数据包一般都比较小，以保证包组内的数据包传输的连续性。源节点连续发送由 2 个背靠背数据包组成的包组，假设这对数据包分别发给不同的接收端，那么这个包组必然具有一段共享的内部网络链路。由于现今的 Internet 广泛应用的是尾部丢弃队列管理方式，如果两个紧邻的数据包先后到达节点的路由缓冲区，其中一个数据包没有被丢弃，即缓冲区未满，则另一个数据包也极大可能不会因为缓冲区溢出而被丢弃，也就是说间隔时间足够小的连续紧邻的两个数据包在该节点的可达性基本一致。



(a) (b)

图 2 包对发包和 3 包组发包

如图 2(a)所示，从源节点 O，发送一对背靠背包对<m,n>,接收节点为 D_m 和 D_n。两条网络路径 P(O,D_m) 和 P(O,D_n)分支节点为 A，则其中一条网络路径 P(O,D_m)上的子链路 P(A,D_m)（非共享链路）的成功传输概率 p(m|n)可表示为：

$$p(m \mid n) = \beta(m \mid n)\frac{r_{m,n}}{r_n} \tag{4}$$

式(4)中，β(m|n)表示包对的相关性，即数据包 n 被分支节点 A 接收的情况下数据包 m 也被 A 接收的概率,$r_{m,n}$ 表示包对<m,n>都被 D_m 和 D_n 接收到的包对数，rn 表示 D_n 接收到包对中数据包 n 的个数。由此，在对中间节点 A 一无所知的情况下观测源端和接收端数据即可大致推断出内部子链路 P(A,D_m)的链路丢包率。

在文献[6]中，作者采用一种改进的 3 包组的发包方式，通过扩展包对的方法来增强探测包间的相关性，提高了最终估计的准确度。如图 2(b)所示，3 包组的发包方式选择接收节点对 D_m 和 D_n，源节点每次发送探测包组<m,n_1,n_2>，数据包 m 的接收节点是 Dm,数据包 n_1 和 n_2 的接收节点是 D_n。假设分支节点为 A，类似于式(4)，可以得到网络子链路 P(A,D_m)的成功传输概率：

$$p(m \mid n_1, n_2) = \beta\left(\mathrm{m} \mid \mathrm{n}_1, \mathrm{n}_2\right)\frac{r_{m,n1,n2}}{r_{n1,n2}} \tag{5}$$

式(5)中 β(m|n_1,n_2)描述了 3 包组的相关性，表示 3 包组

中<n_1,n_2>被分支节点 A 成功接收的情况下，数据包 m 被 A 接收的概率，$r_{m,n1,n2}$ 表示包组中<m,n_1,n_2>都被 D_m 和 D_n 接收到的 3 包组个数，rn_1,n_2 表示 D_n 接收到 3 包组中数据包<n_1,n_2>的个数。式(5)中相关性表达式 β(m|n_1,n_2)大于式(4)中 β(m|n)[6] 。

2　改进的4包组单播层析模型

2.1 包对和 3 包组发包方法的缺陷

如 1.2 节所述，在 Internet 广泛应用尾部丢弃队列管理方式的情况下，可以大致认为包对或者 3 包组在共享的网络内部链路上具有相同的传输特性，他们同时被成功传输或者传输失败，所以，一般采用背靠背包对方式或 3 包组方式进行丢包率计算的时候，把 β(m|n)或者 β(m|n_1,n_2)都近似的取为 1。但在实际网络中，这个值是小于 1 的，尽管 β(m|n_1,n_2)相比较更接近于 1，这个忽略处理依然存在，也就是说存在一个数据包被分支节点 A 接收而另外一个传输失败的情况，以 3 包组发包方式为例，存在这样的可能性：在共享链路中，数据包<n_1,n_2>成功到达分支节点 A 而数据包 m 在共享链路上传输失败。针对上述单播层析发送包方式的包组相关性缺陷，本文对 3 包组的发包方式进行了改良，使相关性 β 的值尽可能的趋向于我们的期望值 1，从而降低测量误差。

2.2 改进的发包方式

已有的研究表明[7]，网络的丢包具有突发性的特点，可能导致一串连续的数据包被丢弃。针对丢包的连续性特征，本文采用可以近似的理解为“4 包组法”的改进发包方式，如图 3 所示，在 3 包组发包的基础上，加入 1 个同样发往接收节点 Dn 的数据包记为 n0，也就是近似于发送连续的<n0,m,n1,n2>4 包组。

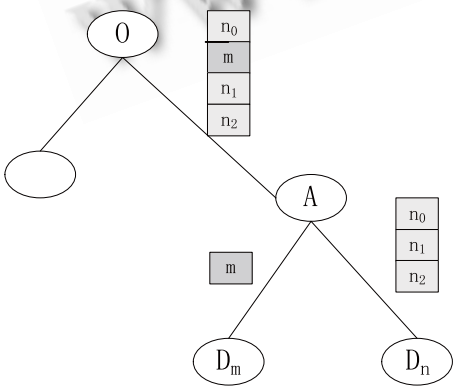


图 3　改进的 4 包组发包

假设从源节点 O 向接收节点 D_m 和 D_n 发送 4 包组<n_0,m,n_1,n_2>，P(O,D_m)和 P(O,D_n)分支节点为 A。网络子链路 P(A,D_m)的成功传输概率可记为：

$$p(m \mid n_0, n_1, n_2) = \beta(m \mid n_0, n_1, n_2)\frac{r_{m, n0, n1, n2}}{r_{n0, n1, n2}} \tag{6}$$

式(6)中，β(m|n_0,n_1,n_2)描述了 4 包组的相关性，表示已知 4 包组中<n_0,n_1,n_2>被分支节点 A 接收的情况下，数据包 m 也被分支节点 A 接收的概率，$r_{m,n}0,_{n}1,_{n}2$ 表示包组<m,n_0,n_1,n_2>都被 D_m 和 D_n 接收到的 4 包组个数，$r_{m,n}0,_{n}1,_{n}2$ 表示 D_n 成功接收到 4 包组中数据包<n_0,n_1,n_2>的个数。

结合 2.1 节的内容对 4 包组的丢包情况进行分析，在发送 4 包组时，发包顺序是交错的 n_0-m-n_1-n_2,首先向 D_n 发送数据包 n_0,然后向 D_m 发送数据包 m,最后再向 D_n 发送数据包 n_1 和 n_2。4 包组<n_0,m,n_1,n_2>在 P(O,A)共享链路上是连续传输的，如文献[8]所述，若共享链路发生连续丢包，此时如果 4 包组中 m 包被丢弃，则和 m 紧邻的 n_0 和 n_1 中必然至少有一个被丢弃，也就是说，对于 4 包组的共享链路 P(O,A),不存在 m 被丢弃而<n_0,n_1,n_2>成功传输的情况，因此,相关性表达式 β(m|n_0,n_1,n_2)的值比 3 包组中的 β(m|n_1,n_2)更加接近于 1,即改进的发包机制应该可以进一步减少单播丢包层析中由于端到端测量不准确引入的误差，提高精度。

3　仿真及数据分析

本文采用了 NS2 网络仿真工具[8]来验证改进的发包方式的性能。首先搭建由图 4 所示的网络拓扑,仿真实验大概的模拟现实网络中的情况，背景流量由多个

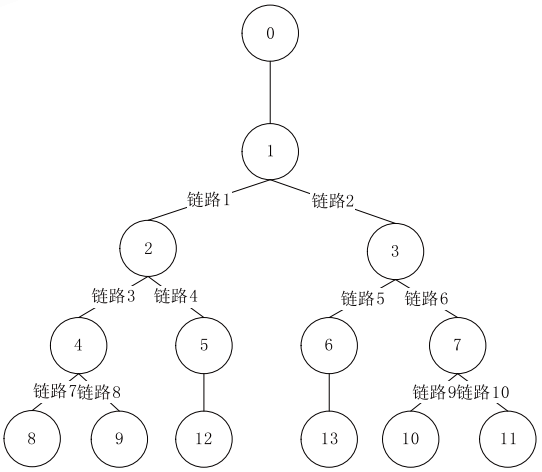


图 4　网络拓扑

TCP 数据流（包含长时和短时 TCP 流）和 UDP 数据流叠加而成，网络中所有的链路均为尾部丢弃队列管理方式，队列长度设置为 50。因为 Internet 中具有内部链路带宽较大，边缘链路带宽较小的特点，故仿真中链路参数设置为:内部链路带宽为 50Mbps，延时为 100ms；边缘链路带宽为 20Mbps,延时为 50ms。

仿真实验时以网络层析中广泛使用的 3 包组以及本文改进的 4 包组两种不同的方式从源节点 0 发送 12000 个数据包，单个探测包大小为 50Byte。图 5 显示的是不同链路的链路丢包率估计值的性能对比，图 6 显示了不同链路的链路丢包率估计相对误差的性能对比。

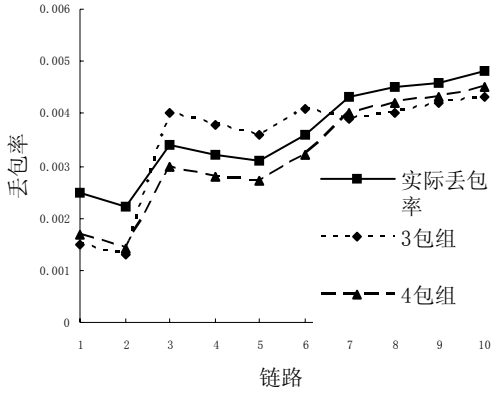


图 5 不同链路丢包率估计对比图

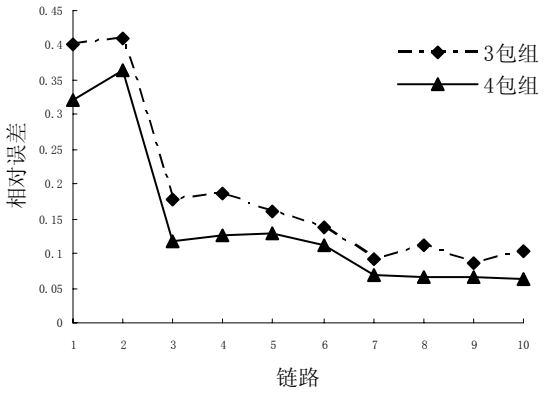


图 6 不同链路丢包率估计相对误差对比图

观察图 5 和图 6，可以清晰的看到随着共享链路的增加，两种发包方式的精确度都有较大程度的提升，这是因为共享链路越长，包组数据包之间的相关性越高,相应的测量误差就会降低，这个结论在以前很多文献中都得到了证明[3,4,6]。分析对比图 5 和图 6 中 3 包组和 4 包组发包方式的性能，由于改进的发包方式能很好的消除本文 2.1 节所述缺陷造成的误差影响，其测量精度方面的优势十分明显，并且随着链路级数的增多其测量误差值也在缩小，如本文仿真拓扑中的链路 7 至链路 10，其测量的相对误差值已经降低至 5%左右，因此，改进的发包方式十分适合多链路级的网络系统丢包率推断。

4 结语

网络层析技术具有广阔的应用前景，在实际的网络应用中,单播测量方式因具备多播不可比拟的优势受到越来越多的关注。由于单播探测包不具备多播探测包的良好相关性，因此增强单播探测包组间数据包的相关性一直是单播测量的研究热点。

本文对基于单播测量的网络丢包率层析成像方法进行了详细的阐述,并在广泛应用的 3 包组发包方式的基础上，提出了一种类似 4 包组发包的改进思路,改善了包组数据包间的关联性能。NS2 的仿真结果论证了改进的丢包率层析方法的优势。

参考文献

1 Claffy K, Monk TE, McRobb D. Internet tomography. Nature, 1999, January.

2 钱峰,胡光岷.网络层析成像研究综述.计算机科学,2006,33(9):12–17.

3 Coates M, Nowak R. Network loss inference using unicast end-to-end measurement. ITC Seminar on IPTraffic, Measurement and Modelling, Monterey, CA, September, 2000. 2384–2394.

4 Zhao H, Chen M, Qiu XF, Zhang GM. Multiple parameters network topology inference based on tomography. Journal of Beijing University of Posts and Telecommunications, 2008, 31(4):24–28.

5 赵洪华,陈鸣.基于网络层析成像技术的拓扑推断.软件学报,2010,21(1):133–146.

6 Dulfield NG, Presti FL, Paxson V, et al. Network loss tomography using striped unicast probes. IEEE/ACM Trans. on Networking, 2006,14(4):697–710.

7 Michael S, Borella D, Uludag SS, et al. Internet Packet Loss: Measurement and Implications for End-to-End QoS. 2002. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=721868&tag=1

8 于斌,孙斌,温暖,等.NS2 与网络模拟.北京:人民邮电出版社,2007.