软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2015,26(9):2356−2372 [doi: 10.13328/j.cnki.jos.004867]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

网络链路性能参数估计的层析成像方法综述*

潘胜利, 张志勇, 费高雷, 钱 峰, 胡光岷

(宽带光纤传输与通信网技术教育部重点实验室(电子科技大学),四川 成都 611731)

通讯作者: 胡光岷, E-mail: hgm@uestc.edu.cn

摘 要: 网络层析成像能够在网络内部节点不提供测量协作的情况下,根据端到端的测量结果,间接地估计网络内部链路性能参数,是一种重要的网络测量手段,能直接指导网络管理和网络优化,目前受到国内外学术界和工业界广泛的关注.在广泛收集国内外资料的基础上,首先总结了目前网络层析成像使用的主要端到端测量方法和技术;再根据不同参数对链路性能刻画程度的不同,将链路性能参数的网络层析成像方法分为两类:定量参数推断方法和定性参数推断方法;然后,针对不同类型参数的估计问题,概括分析了现有算法的特点;最后指出该类方法未来的研究方向与潜在的应用前景.

关键词: 逆问题;网络层析成像;链路性能评价;端到端测量;相关性;可辨识性

中图法分类号: TP393

中文引用格式: 潘胜利,张志勇,费高雷,钱峰,胡光岷.网络链路性能参数估计的层析成像方法综述.软件学报,2015,26(9): 2356−2372. http://www. jos.org.cn/1000-9825/4867.htm

英文引用格式: Pan SL, Zhang ZY, Fei GL, Qian F, Hu GM. Survey on network tomography for link performance parameter evaluation. Ruan Jian Xue Bao/Journal of Software, 2015,26(9):2356−2372 (in Chinese). http://www.jos.org.cn/1000-9825/4867.htm

Survey on Network Tomography for Link Performance Parameter Evaluation

PAN Sheng-Li, ZHANG Zhi-Yong, FEI Gao-Lei, QIAN Feng, HU Guang-Min

(Key Laboratory of Broadband Optical Fiber Transmission & Communication Networks (UEST of China), Chengdu 611731, China)

Abstract: Network tomography provides the ability to employ the end-to-end measurements to infer the network-internal link performance parameters indirectly without requiring cooperation from the intermediate elements of the network. As a significant alternative to network measurements to be able to guide the network management as well as the network optimization, network tomography receives a plenty of attention both in academia and industry. This survey is based on an extensive collection and reference of research works at home and abroad. First, the measurement schemes exploited by the network tomography are summarized. Next, the corresponding tomographic approaches are divided into two classes regarding at what granularity they describe the link's performance: the quantitative parameter estimation and the qualitative parameter estimation. Then according to inference problems of the different parameters, a general analysis of the existing algorithms is conducted. Lastly, future research areas and potential applications are suggested.

Key words: inverse problem; network tomography; link performance evaluation; end-to-end measurement; correlation; identifiability

Internet 是一个分布式管理的网络,出于竞争以及安全等因素的考虑,各自治部分往往将其链路性能参数等视为敏感信息而不愿意公布与分享.随着近年来 Internet 规模的迅速扩增[1]、各种新式网络应用的大量涌现以及网络经济活动的兴起,Internet 的影响已经深入到人们日常衣食住行的各个方面,这使得对网络高效稳定安全

* 基金项目: 国家自然科学基金(61171091, 61201127, 61301274); 中央高校基本业务费专项基金(ZYGX2012J005, E022050205); 保密通信重点实验室基金(9140C110503140C11054)

收稿时间: 2014-03-26; 修改时间: 2015-01-28; 定稿时间: 2015-06-13

运行的要求变得日益迫切.例如网络视频电话应用 Skype 和在青少年中流行的在线竞技游戏 DotA(defence of the ancient)以及天猫“光棍节”网络抢购盛宴等,用户要求稳定且高质量的网络链路.这使得网络提供商(Internet service provider,简称 ISP)不仅要准确地掌握所辖网络中的链路性能参数,而且还需要他们能够获取对等网络中的链路性能数据,对于确保 SLA(service level agreement)和及时排查和消除网络故障等具有重要的意义.然而在各自治部分高度独立且缺乏网络测量协作的 Internet 中,直接监控对等网络几乎是不可能的.尽管基于类 traceroute 工具可以部分达到某些测量目的,如 CAIDA 等所进行的 Internet 拓扑发现[2],但是 ICMP(Internet control message protocol,控制报文协议)[3]报文的低转发优先级以及网络中大量存在的匿名路由器[4],使得测量得到的链路性能参数(如链路时延和丢包率等)不仅不真实而且不完整,远远达不到网络监控要求.通过端到端路径级测量的网络层析成像方法[5],能够在目标网络内部节点不协作的情况下,依然能够准确有效地获取链路性能参数,成为近 10 年来一个持续受到广泛关注的研究课题.

端到端设计原则[6]使得网络对于用户或者各对等网络而言,具有某种意义的不可见性或者透明性.网络层析成像将数据测量放在网络边缘,不要求网络内部节点进行测量协作.如同医学成像技术中通过对内脏器官多个部位的观察来了解器官内部情况一样,网络层析成像以获得的端到端路径级参数数据来反推网络内部链路级参数,如时延与丢包率等.网络链路性能参数的层析成像问题主要包括有链路时延分布推断、链路丢包率推断、公共拥塞链路检测、网络拥塞链路定位以及链路中立性检测等[5,7−10].由于网络中端到端路径数往往小于网络的链路数,网络层析成像是一个典型欠定性反问题.为了克服欠定性,网络层析成像方法通过采用特殊的测量手段,如多播测量和模拟多播的单播背靠背包测量,来获取报文级的相关性[11];要么就是假定链路性能服从特定分布或具有时空独立性和平稳性等[5].但这些用于克服欠定性的方式与方法反过来又会影响到网络层析成像方法的实用性与健壮性.如何有效地克服欠定性以及去掉不合理假设等,是当今网络层析成像方法研究的难点.

根据目标参数对于链路性能刻画粒度的不同,将网络层析成像方法分为两大类:第 1 类包括有估计出链路性能具体特征,如时延分布与丢包率等的定量参数估计方法;而第 2 类主要包括用以描述链路状态特征,如是否具有公共拥塞链路、链路拥塞与否以及链路是否中立等的定性参数估计方法.第 1 类方法主要是利用报文级相关性,因而对于端到端测量要求比较高;而第 2 类方法主要是利用性能级相关性[11],使用被动测量就能满足数据要求(如图 1 所示).需要指出的是:由于第 1 类方法利用了报文级相关性,使定量参数具有更多的信息量,并使得定量参数对链路性能的描述粒度也相应较高.相比于定量参数估计这种类似于点估计的特点,描述链路性能状态好坏等二元特征的定性参数估计方式则可属于范围估计.不难看出,可以直接通过定量参数推断出定性参数,如:先估计出链路丢包率大小,再与 1%丢包率门限比较评价链路是否拥塞;也可以通过分析链路具体的时延分布情况等来评价判断链路时延性能状况.遗憾的是,定量参数的获取往往依赖于报文级相关性测量,其测量推断方案表现出较高的复杂度,因而测量代价也相应较大.而相比较之下,定性参数的估计可直接基于性能级相关性测量,其对测量噪声的容忍度高,使得测量方案易于设计与实施,更具实用性.当前,除了定性参数推断中出现了部分实用化案例外[12],几乎看不到定量参数推断方法实用化的案例.

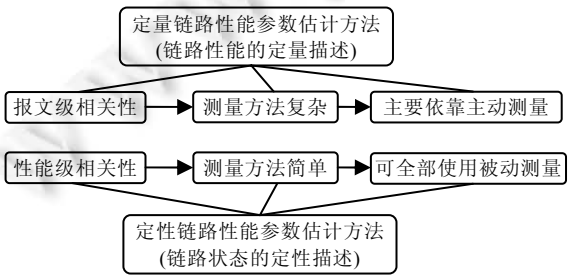


Fig.1 Inference for link performance parameters

图 1 链路性能参数估计

定量与定性链路性能参数关系如图 2 所示.

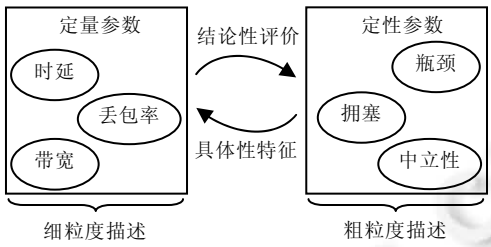


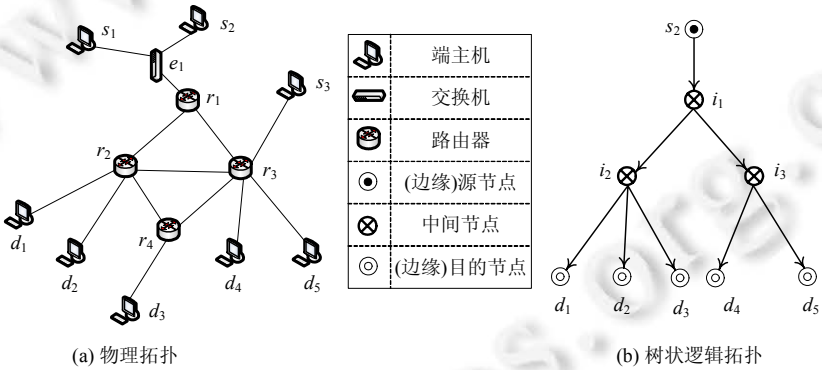
Fig.2 Relationships between quantitative and qualitative link performance parameters

图 2 定量与定性链路性能参数关系图

1 基本模型

1.1 网络模型

基于网络层析成像技术的链路性能参数推断通常要求目标网络的逻辑拓扑已知[5,7,8],如图 3(b)所示.区别于一个网络的物理拓扑,代表一个物理网络中各实体设备之间的物理互联关系(如图 3(a)所示),而一个网络的逻辑拓扑用于描述数据流量如何在网络中传输与流动[13].逻辑拓扑中的节点可以代表一个主机、一个传感器、一个路由器,甚至是一个网络;逻辑拓扑中的逻辑链路可以是一条实际的物理链路,也可以由很多条物理链路所组成,它直接指示数据的传输关系.



(a) 物理拓扑 (b) 树状逻辑拓扑

Fig.3 Physical topology and tree logical topology

图 3 物理拓扑与树状逻辑拓扑

用一个有向无环图$\mathcal{G}(V,E)$对网络的逻辑拓扑进行建模,其中:V 代表网络中的节点集合,由边缘源节点 S、边缘目的节点 D 和内部节点 I 组成,即 $V=S\cup D\cup I$;E 代表网络中的逻辑链路集合.特别地,当$\mathcal{G}$是如图 3(b)所示的树型逻辑拓扑时,除根节点外,$\mathcal{G}$中任一节点 $i\in V\backslash S$,都具有唯一一个父节点 $f(i)\in V\backslash D$.对任意正整数 n,递归地定义 $f^n(i)=f^{n-1}(f(i))$.若存在 n 使得 $j=f^n(i)$,则称节点 j 是 i 的一个祖先节点,i 是 j 的一个子孙节点,并用 $j\succ i$ 表示.$j\succeq i$ 则表示$j\succ i$或者$j=i$.令以内部节点$k\in I$为共同祖先节点的边缘目的节点集合为 $D_k=\{d|k\succ d,d\in D\}$.如果$\mathcal{G}$是树状拓扑,那么此时将$\mathcal{G}(V,E)$表示为$\mathcal{T}(V,E)$,且以内部节点 k 为起始根节点的子树表示为$\mathcal{T}_k$(如后文图 6(a)所示).令 $e_{k\to j}\in E$ 表示从节点 k 到节点 j 的一条直连逻辑链路.对于一条从源节点 s 到目的节点 d 的端到端路径$\mathcal{P}_{s\to d}$,有$\{e_{f(k)\to k}|s\succ k\pm d;s\in S,d\in D\}$表示其经过的所有链路构成的集合.

1.2 数学模型

对于以节点 s 为根节点的树状网络,它到某个目的节点 d 的路径对应的端到端时延值,应该该路径上所有链

路的时延值之和.将一个路径的时延测量值记为 $t_{s\to d}$,探测包在这个测量时隙所经历的链路时延值记为 $t_{f(k)\to k}(s\succ k\succeq d)$.那么,有如下方程:

$$t_{s\to d}=\sum_{s\succ k\succeq d} t_{f(k)\to k} \tag{1}$$

令链路 $e_{f(k)\to k}$ 的丢包概率为 $l_{f(k)\to k}$,那么当假设测量周期内各条链路的丢包概率相互独立时,从源节点 s 到目的节点 d 这条路径的丢包概率与其上各条链路的丢包概率具有如下方程关系:

$$1-l_{s\to d}=\prod_{s\succ k\succeq d}(1-l_{f(k)\to k}).$$

其中,$l_{s\to d}$ 表示路径丢包概率.对上式两边取对数,则能将连乘变为与公式(1)一样的线性连加:

$$\log(1-l_{s\to d})=\sum_{s\succ k\succeq d}\log(1-l_{f(k)\to k}) \tag{2}$$

把公式(1)与公式(2)称为测量方程.用一个列向量表示当前网络所有路径在 i 时刻的测量结果,记为 y_i;用一个矩阵表示所有 m 条路径在 T 个时隙的测量结果,记为 $Y_{m\times T}=\{y_1,\ldots,y_i,\ldots,y_T\}$.记 $X_{n\times T}=\{x_1,\ldots,x_i,\ldots,x_T\}$ 为 T 个测量时隙中所有 n 链路相应的参数.$A_{m\times n}=\{a_{i,j}\}$ 表示路由矩阵,由 0 与 1 组成.当第 i 条路径经过了 j 条链路时,$a_{i,j}=1$;反之,则有 $a_{i,j}=0$.记逻辑拓扑近似误差以及由时钟漂移等因素引起的测量误差为 ε,那么网络层析成像可以使用下面的线性模型为

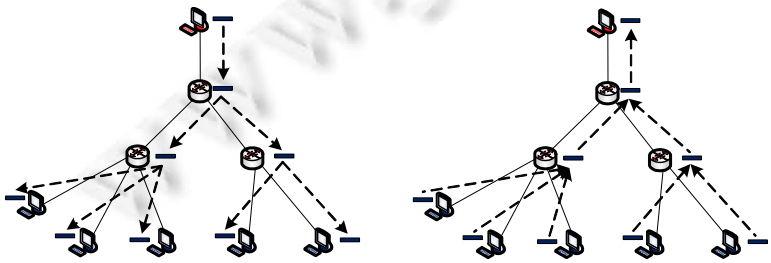
$$Y=AX+\varepsilon \tag{3}$$

2 端到端测量

网络层析成像将目标网络视为一个黑盒子,通常,所有的测量均是围绕网络的端节点进行,这种测量策略被称为端到端测量.端到端测量除了被动监听端节点对之间的数据报文传输外,其更多地是采用主动地在端到端节点之间发送探测报文的方式.根据所采取的路由协议不同,端到端测量方式目前主要分为两种:多播测量与单播测量.出于安全等因素的考虑,当前 Internet 中路由器对于单播的支持高于多播[5].

2.1 多播测量

在配置了多播路由协议的网络中,源节点每次向一组目的节点发送单一的多播探测报文;中间路由节点在接收到多播探测报文后首先对其进行复制,然后分别向相应的下一跳端口转发,如图 4(a)所示.最早的网络层析成像研究者们注意到多播的这种复制转发机制能使得端到端测量数据具有良好的相关性,从而有利于网络内部参数推断.如美国马萨诸塞州大学 MINC(multicast-based inference of network-internal characteristics)工程[14],他们基于多播端到端测量进行了链路丢包率以及链路时延分布等性能参数的估计算法研究.此外,在无线传感器网络中还存在一种反向多播[15]的特殊数据传输方式.反向多播相当于一般多播的逆向过程,即,数据包由多个源节点向同一个目的节点发送.在反向多播网络中,中间结点会周期性地收集与整合其接收到的信息,并向其上一级节点传递,如图 4(b)所示.从端到端测量角度看,反向多播和一般多播方式相同.



(a) 多播 (b) 反向多播

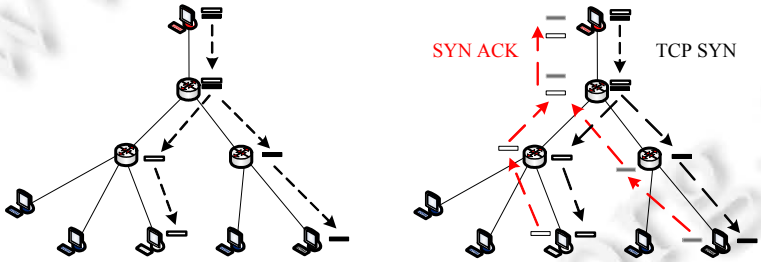
Fig.4 Multicast and reverse multicast

图 4 多播与反向多播

在路由器的路由配置中,多播路由与单播路由往往具有不同的缓存队列[16].当目标网络中多播业务,如多播 Virtual Private Network(VPN)等的流量占主导地位时[17],其估计效果还比较真实;然而一旦目标网络中单播流量处于主导位置时,路由器对多播报文的不同处理,会使得利用多播测量估计出的链路性能参数,与实际相比存在不可忽视的偏差.当前,Internet 中绝大多数流量都是由单播通信产生,因而基于单播测量的网络层析成像所受到的关注也相对较多.虽然如此,但基于多播测量开发的算法也可以用于单播测量中[18],因而在后续介绍相关算法时,将不再特别强调测量数据的获取方式.

2.2 单播测量

路径上探测报文之间的相关性是网络层析成像的基础[19].对于瓶颈链路诊断以及网络拥塞链路定位等,其对端到端测量报文要求较低,一般只需要路径之间的测量数据具有性能级相关性即可;但是对于链路时延与丢包率估计而言,往往要求路径间测量报文在其共享路径上,尽可能经历相同的时延与丢包过程.模仿多播测量,Coates 等人[20]使用背靠背包来进行端到端单播测量.如图 5(a)所示,源节点在一个测量时刻向两个目的节点发送两个紧邻在一起的探测包,那么从源节点离开后,这两个紧邻的探测包在共享路径上将经历相似甚至是完全相同的时延与丢包过程.Duffield 等人[21]提出了类似的模拟多播的探测方法,称为包群探测,每次探测时可以向多于两个的目的节点发送报文;他们还对发包过程做了改进[22]:在每次探测时,可以同时向目的节点发送多个连续的探测包,从而有效地增强了背靠背探测包间的相关性.不管是单播还是多播,路径丢包率是路径上丢失的报文数目与路径的发送报文之间的比值;链路时延(包括排队时延、发送时延、传播时延以及传输时延[23])的主要是关注排队时延,因而路径上总的(排队)时延值通过减去测量周期内的最小测量值来得到.除了这些端到端测量方式外,还有一些能够适用于其他特殊目的,如拓扑推断[24]与流量矩阵估计[25,26]等的测量与监控方案.



(a) 背靠背探测 (b) 网络雷达探测

Fig.5 End-to-End measurement schemes

图 5 端到端测量方案

事实上,无论是多播探测还是单播探测,虽然均不要求网络内部(路由)节点进行测量协作,但是都隐含假设了网络端节点是处于受控范围的.此外,在当前网络层析成像研究工作中,大部分都没有考虑端到端测量数据中噪声的影响,并且默认单源端到端方式所测量的拓扑为树型拓扑[27].实际网络中,由于端节点时钟漂移的不同以及网络路由的变化等原因,会使得测量数据以及树型拓扑近似程存在不同程度上的偏差.一般说来,上述端到端测量方案具有如下几处不足:

① 需要(目的)端节点协作测量;

② 往往要求端节点时钟同步;

③ 不考虑测量噪声;

④ 忽略了路由对测量的影响.

Tsang 等人[28]根据 TCP 协议建立连接时需要进行三次握手[3]这一特点,创造性地改良了背靠背探测机制,提出一种能够能够不需要端节点协作与时钟同步测量方法,称为网络雷达(network radar).如图 5(b)所示:源节点向两个目的节点发送两个紧邻的 TCP SYN 报文,目的节点在接受到此请求连接的报文后会给源节点发送一个

SYN-ACK 报文;然后,源节点统计出各条路径的 RTT(round trip time,往返时延)测量值.由于路径间报文的相关性还是主要由它们经历的共享路径引起的,反向路径其实在某种意义上可以看最后一跳链路的一个延伸.换言之,将网络雷达得到的测量数据用于估计链路时延分布与丢包率时,相当于在测量数据中引入了噪声,这会使得最终的估计结果存在偏差,因而可能不适用于这类链路性能参数的估计.但由于网络雷达使用了高层协议(应用层),其在一定程度上能避免一些基于五元组流(由具有相同源目的 IP 地址、源目的端口号以及协议号的报文所组成)的流量均衡策略所造成的多路由影响[29].除了以上具体的端到端测量机制研究外,Gu 等人[30]针对单播时延网络层析成像,基于费舍尔信息矩阵理论提设计出一种最优探测方案;Xu 等人[31]基于压缩感知理论,研究了如何选择端到端测量路径才能够有效地达监控网络内部链路的目的;Jaggard 等人[32]则系统地研究了如何才能有效地设计一个探测方案.

3 定量参数推断方法

3.1 链路丢包率推断

3.1.1 基于链路丢包时域独立性假设

假设单源树状拓扑$\mathcal{G}(V,E)$,令 $l_e=\{l_{f(k)\to k}|k\in I\cup D\}$表示内部链路丢包概率集合,$n(d)(d\in D)$表示从根节点 s 到目的节点 d 这条路径上接收到的探测包数目.不考虑时钟漂移以及丢包的时域相关性等因素,利用公式(3)提供的等式约束关系,得到似然函数 $\mathcal{L}(l_e)=\sum_{d\in D}n(d)\log(1-l_{s\to d})$.将 l_e 看成是缺失数据,进而将其变成一个数据缺失的 MLE(maximum likelihood estimation)问题[5].由于$\mathcal{L}(l_e)$与 l_e 的关系复杂,直接求导往往不可行,此时可以借助于 EM(expectation maximization)算法进行求解[5,7,20,33−36].EM 算法是利用 E 步与 M 步进行迭代求解的,其中 E 步计算出每条内部链路成功传输报文的期望数目.虽然 EM 算法具有较好的性能,但有可能会陷入局部最优解,而且其收敛速度会随着网络规模的增大而急剧增大.

Lawrence 等人[5]提出了一系列基于最小二乘思想的链路丢包率估计方法:即使是使用数量较少的探测包,IRWLS(iteratively-reweighted least square)估计器的性能也能非常接近 MLE.而且基于最小二乘可以很容易地就计算出估计器的渐进方差.相比于 MLE,基于最小二乘的方法计算效率更高.注意到:多播测量中,一旦当前节点的兄弟节点(具有共同父节点的节点)接收到了报文,即使当前节点未收到报文,也可以断定他们的父节点接收到了报文.Cáceres 等人[33]提出了一种自顶向下的链路丢包率估计方法.同样基于多播测量,清华大学的 Su 等人[37]提出一种自底向上的链路丢包率显式估计算法 FPA(fast path-based approach).证明了 FPA 估计量的一致性,并且得到与 MLE 一样的渐进方差.基于显式估计算法的计算时间受网络规模影响较小,而 MLE 的计算时间随网络规模的增大而显著增加.

3.1.2 基于链路丢包时间相关性假设

假设链路丢包独立,其实是将丢包过程看成是一个 Bernoulli 模型[7].而 Guo 等人[38]认为,丢包过程往往具有时域相关性,提出使用 Gilbert 丢包模型来代替传统的 Bernoulli 丢包模型.在 Gilbert 模型中,当前报文的传输状态(成功传输与被丢弃)与前一报文的传输状态有关联.通过 Gibbs 采样获取各条链路的 Gilbert 模型参数,然后进行链路丢包率的估计.Arya 等人[39]结合子树划分思想,进一步将链路的丢包关联长度推广到任意 k 个.他们证明了,链路的平均丢包长度可以通过链路传输概率以及连续两个包被丢弃的概率计算得到.

文献[40]指出,马尔可夫链模型能够很好地模拟链路丢包过程.文献[41]在一个深度为 4、链路条数为 16 的树型网络中进行模拟实验,展示了在 k-MC(k 阶马尔可夫链)丢包模型、Gilbert 丢包模型以及 Bernoulli 丢包模型下估计 $p_{4/3}$(表示前面连续 3 个报文都丢失额情况下,第 4 个报文被丢弃的概率,见表 1)的平均相对误差大小.

Table 1 Compute $p_{4/3}$ in different loss models

表 1 在不同的丢包模型中计算 $p_{4/3}$

k-MC	Gilbert	Bernoulli
$p_{4/3}$	$p_{2/1}$	p_1

文献[41]得出,用 k-MC 丢包模型在所有链路上的估计结果均具有最小的相对误差,说明使用 k-MC 丢包模型比 Bernoulli 丢包模型以及 Gilbert 丢包模型能更好地刻画链路丢包过程.

3.2 链路时延分布推断

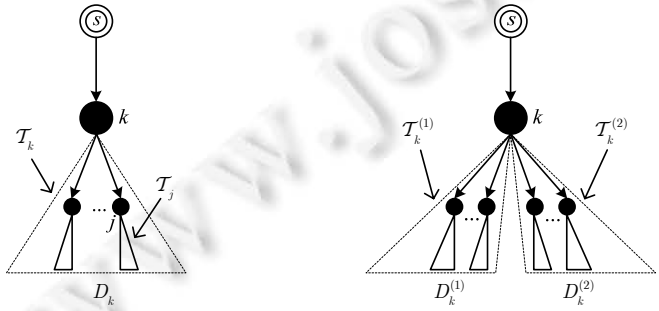
3.2.1 参数化方法

参数化方法通过假定链路时延分布服从有限的特定分布,如高斯分布、泊松分布以及混合分布等,因而其模型参数个数常常是固定的[42].文献[43]将链路时延分布建模成离散时延模型,即,链路的时延仅具有几个有限的离散状态,丢包被视为一个离散时延状态[18,44,45].采用一个反卷积过程,从最小时延状态开始向上推断出所有链路的离散时延分布.事实上,他们估计方法不是 MLE 而是一类基于矩估计的方法[46],此外,直接计算 MLE 通常不可行[17].Liang 等人[36]通过原树型网络其划分成多个子树网络(如图 6(b)所示),通过忽略各个子树之间的相关性,将原问题转化为多个子问题.然后,在每个子树上进行 MLE,并得到原问题的最大伪似然估计(maximum pseudo likelihood estimate,简称 MPLE).文献[47]认为:采用相同的时延区间去离散时延状态时,如果时延区间太小,对于高带宽低时延的链路的刻画粒度太细;而如果时延区间太大的话,对于低带宽高时延的链路的粒度又显得又太粗糙.因此,需要采用可变离散区间来对链路时延进行建模.Arya 等人[17]认为,链路丢包具有时域相关性,把丢包率看成一个离散时延状态并采用子树划分技术后,将先前估计链路丢包率的工作[39]推广到了时域相关的链路离散时延分布估计上.

Vardi 首次提出了网络层析成像的概念[26],虽然是针对流量矩阵估计问题,但是她将路径级流量建模成服从泊松分布的思路是可以应用于链路时延分布估计的.Shih 和 Hero 结合离散分布模型与连续分布模型,提出一种链路连续时延分布估计的方法[46].对于某条链路 e,$\delta(x)$表示报文通过其时具有零时延的概率,$\phi(x;\theta_{e,m})$表示第 m 个具有参数集合$\theta_{e,m}$连续分布成分在时延为 x 的概率.假设有 k_e 个连续分布混合成分,其混合比例 $\sum_{m=0}^{k_e+1} a_{e,m}=1$,那么时延分布的概率密度函数表示如下:

$$f_e(x)=a_{e,0}\delta(x)+\sum_{m=1}^{k_e} a_{e,m}\phi(x;\theta_{e,m}) \tag{4}$$

公式(4)中的连续分布$\phi(x;\theta_{e,m})$可以是高斯分布、指数分布以及均匀分布等.文献[46]用高斯混合混合分布进行了链路时延建模时 matlab[48]与 ns2[49]仿真结果,可以发现,混合模型可以比较好地逼近原有的时延分布.文献[50]针对链路时延建模中的非参数模型以及参数模型中的混合分布模型等合理性进行了理论分析,并指出它们都能够很好地逼近真实的时延分布.除了高斯混合模型,使用大量时延说明了指数模型以及指数混合模型在链路时延建模上的合理性以及 EM 算法在估计参数方面的有效性.



(a) 分别起源于节点k与j的两颗子树 (b) 一棵子树被划分成两颗子树

Fig.6 Subtree and subtree partition[17]

图 6 子树与子树划分[17]

3.2.2 非参数化方法

非参数化方法是指参数的数目或者自由度是随着测量数据个数变化的[51,52].Tsang 等人[42]在同文献[43]一

样的离散模型的基础上,将离散的状态个数 K 扩展到了大于等于测量数据个数 N,有 $K=2^{\lceil\sqrt{N}\rceil}$.由于状态数目增多使得对时延状态的刻画有可能存在过适的问题,他们提出了一种多尺度最大惩罚似然估计方法(multiscale maximum penalized likelihood estimation,简称 MMPLE)来解决对链路时延状态过度拟合的问题.定义链路 i 在状态 j 的一个离散概率质量函数(probability mass function,简称 PMF) $p_{i,j}=\int p(t)\mathrm{d}t$,其积分区间为$[(j)/K,(j+1)/K]$.令 N_i 表示第 i 条链路上的测量数据个数,那么有 MMPLE 的最优化目标为

$$\arg\min_{p}\log l(y\mid p)-\sum_{i}\frac{1}{2}\log(N_i)\times\#i \tag{5}$$

其中,y 表示的端到端路径时延测量结果,$\#i$ 是链路 i 的概率质量函数 p_i 的非负 Haar 小波系数[53].一般而言,利用 MMPLE 估计出来的链路时延概率能够刻画出非常细腻的时延分布波形特征.

基于 MLE 的链路时延估计方法实际上是求得一个模型分布,使得其与测量数据 Y 的经验分布之间的 K-L 距离[55]最小.文献[54]不直接利用 Y 的经验分布,而是借助公式(3)首先推导出路径与链路时延概率分布之间的特征函数的乘积关系.由于一个概率分布之间与其特征函数是一对傅立叶互逆变换的关系,因此一旦确定了特征函数,就等同于确定了概率分布.

文献[54]然后通过将链路时延建模成有限均匀分布混合模型 $f_j(x)=\sum_{e=1}^{k_j}p_{je}\kappa_{je}(x)$,通过最小化 Y 的经验分布与模型分布之间的差别来求得相应的链路混合时延模型参数.而且当测量数据趋于无穷大时,文献[54]所提出的估计器能够和 MLE 一样是渐进有效的.

链路性能定量参数推断的主要算法归纳见表 2.除了上述链路丢包率与时延分布估计等参数外,还有部分对(端到端路径)瓶颈链路可用带宽进行估计的工作[56,57].上述网络链路定量参数估计一般而言,是面向常规有线网络.文献[58]针对支持网络编码的网络,提出一种有向算法(orientation algorithm,简称 OA)来设置网络中各节点的编码方式,使得源节点向目的节点发送的数据报文能够覆盖整个网络;然后,集合置信传播(belief propagation,简称 BP)算法估计链路丢包率.不过,当前支持网络编码的网络较少,限制了其使用范围.同样针对支持网络编码网络,文献[59,60]将文献[58]中技术应用到了无线传感器网络的链路丢包率估计.此外,还有如文献[61]将网络层析成像技术应用与覆盖网络的测量与监控[62].

Table 2 Inference algorithms for quantitative parameters of link performance

表 2 链路性能定量参数推断算法

Name or related paper	Inference parameters	Algorithm strategy
IRWLS	Loss	Least square
MVWA[18]	Loss & delay	MLE
Ref.[37]	Loss	Explicit estimator
Ref.[31]	Loss	Bayesian inference
Ref.[57]	Delay	Discrete model & least square
MML[46]	Delay	MLE
MMPLE	Delay	MLE & time-frequency method
WCF[54]	Delay	Moment method
Ref.[14]	Loss & delay	Discrete model & explicit estimator

4 定性参数推断方法

4.1 公共拥塞链路检测

4.1.1 基于时频域分析

当两条路径经历了公共拥塞链路时,它们的路径性能将会具有很强的相关性.Kim 等人[63]基于时延相关性,提出利用小波降噪的方法来进行公共拥塞链路的识别.记第 i 条路径与第 j 条路径的(单向)路径时延为 $Y_i=\{y_i^t\}$ 与 $Y_j=\{y_j^t\}$,那么其时延相关系数为

$$XCOR_{ij;n}=\frac{\sum_{t=1}^{n}(y_i^t-\bar{Y}_i)(y_j^t-\bar{Y}_j)}{\sqrt{\sum_{t=1}^{n}(y_i^t-\bar{Y}_i)^2\sum_{t=1}^{n}(y_j^t-\bar{Y}_j)^2}} \tag{6}$$

其中,n 为测量时隙数.当这两条路径不共享拥塞链路时,认为它们不具有相关性,有 $XCOR_{ij;n}=0$.如果两条路径经历了公共拥塞链路,那么它们的相关性主要由拥塞链路所决定.Kim 等人发现:当网络轻载时,路径时延波形同常见的噪声信号波形相似,有着比较小的幅度;而当重载时(有拥塞时),路径时延波形具有很强的脉冲特性.他们将由拥塞引起时延看成是信号 $s(t)$,然后,背景流量引起的时延看成是加性噪声 $n(t)$,将路径时延 $y(t)$建模成:

$$y(t)=s(t)+n(t) \tag{7}$$

选取小波基为 $\psi_{i,j}(t)=2^{-i/2}\psi(2^{-i}t-j)$,对应的小波系数为 X_j^i,那么有:

$$Y(t)=\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}X_j^i\psi_{i,j}(t) \tag{8}$$

通过将小波系数小于设定门限的信号成分去掉,得到提出的信号为

$$\hat{s}(t)=\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}d_T(X_j^i)\psi_{i,j}(t) \tag{9}$$

然后,使用经过公式(9)去噪后的信号来进行公式(6)中的路径时延相关性计算.如文献[63]所指出:当时钟漂移达到 1s 时,基于小波去噪的方法依然能够具有比较良好的检测率.此外,还有其他一些基于时频分析的检测方法[64]以及基于 PCA(principal component analysis)分析[65]的方法等,它们往往是要求测量数据拥有良好的相关性,或者说需要测量数据具有高信噪比,因而算法健壮性稍稍不及小波去噪方法.

4.1.2 基于统计推断

利用统计推断手段进行公共拥塞链路检测主要有基于时延与基于丢包两大类.当两条路径之间存在公共拥塞链路时,此时两条路径将表现出空间相关性,主要表现为路径间的时延与路径丢包率具有关联性.如果两个去往不同路径且相隔很近的探测包经过共同的拥塞链路,它们将经历相似的时延过程;而当前面第 1 个探测包被丢弃时,后面第 2 个探测包被丢弃的概率往往也较高.Rubenstein 等人[66]在两条路径上分别发送探测流,其中,各路径探测流包的发送间隔服从泊松分布.记路径 1 上面第 j 个探测包 $p_{1,j}$ 丢失了为 $L_{1,j}=0$;同理,$L_{2,i}=0$ 表示路径 2 上第 i 个探测报文 $p_{2,i}$ 丢失了.如果报文 $p_{1,j}$ 与报文 $p_{2,i}$ 到达目的节点的时间是紧邻的,那么有 $a(p_{1,j},p_{2,i})=1$.对于具有相同源节点的两条路径(倒 Y 结构)而言,文献[66]分别在单条路径与两条路径的测量报文中计算相邻两个报文在第 1 个报文丢失后第 2 个报文也被丢失的条件概率(公式(10)):

$$\begin{cases}Ma=\Pr(L_{2,i}=0\mid L_{2,i-1}=0)\\ Mx=\Pr(L_{2,i}=0\mid L_{1,j}=0,a(p_{1,j},p_{2,i})=1)\end{cases} \tag{10}$$

而当两条路径具有共同的目的节点时,则有上述概率表述(公式(11)):

$$\begin{cases}Ma=\Pr(L_{2,i}=0)\\ Mx=\Pr(L_{2,i-1}=0\mid L_{1,j-1}=0,a(p_{1,j},p_{2,i})=1)\end{cases} \tag{11}$$

记探测包 $p_{1,j}$ 所测得的路径时延值为 $D_{1,j}$,那么基于时延协方差可以有如上相应的度量为

$$\begin{cases}Ma=C(\{(D_{2,i},D_{2,i+1})\})\\ Mx=C(\{(D_{2,i},D_{1,j}):a(p_{2,i},p_{1,j})=1\})\end{cases} \tag{12}$$

其中,$C(\{(D_{2,i},D_{2,i+1})\})$表示如公式(6)所计算的路径时延协方差.当计算完上述度量值时,通过 $Mx>Ma$ 的关系来判断两条路径拥有公共拥塞链路;否则没有.一般而言,基于时延的方法比基于丢包率的方法更具健壮性[66].

Harfoush 等人[67]认为,这种将路径时延与丢包分开来处理的手段并没有很好地利用好测量数据中的有用信息,提出了将时延与丢包联合起来处理的方法认为,丢失的探测报文仍可以被视为经历了隐藏的时延状态,并使用隐式马尔可夫模型来对端到端路径时延进行建模,进而实现在单条路径对拥塞链路进行识别.值得注意的是:文献[66,67]中进行端到端测量时,均不再要求时钟同步.虽然文献[66]涉及到不同的路径时延,但是在计算协

方差时认为时钟漂移是常数,故不会影响协方差计算结果.

4.2 网络拥塞链路定位

4.2.1 基于单时隙测量

拥塞链路一般是指丢包率超过一定数值(≥1%)的链路,而当路径的丢包率超过 5%时,将路径判定为故障链路[11].Padmanabhan 等人[68]用丢包率作为衡量链路是否拥塞的性能指标,提出了 3 种估计链路丢包率的方法:线性最优化、随机采样和基于 Gibbs 采样的贝叶斯估计.线性最优化方法以路径传输率和链路传输率的对数之间的线性关系为约束,链路传输率的对数和为最小化目标建立一个约束最优化问题并求解;随机采样方法则根据路径丢包率与链路丢包率之间的关系确定链路丢包率的大致取值范围,再在该范围内随机采样,获得各条链路的丢包率样本,以样本均值作为链路丢包率的估计值;基于 Gibbs 采样的贝叶斯估计方法则利用比较成熟的 Gibbs 采样方法[69]生成一个马尔可夫链,使得该马尔可夫链的稳态分布正好是链路丢包率的后验分布.因此,Gibbs 采样过程稳定后所生成的样本服从链路丢包率的后验分布,这些样本的均值可视为链路丢包率的后验估计.以上 3 种方法中,随机采样方法由于采样过程比较粗糙,估计精度在三者当中最差,线性最优化对测量误差比较敏感[70,71].

Duffield 将他前期的工作[72]加以完善,建立了一个用于诊断拥塞链路的网络层析成像新框架[11].由于该框架建立在布尔代数域[73],因此通常称为布尔网络层析成像或者二进制网络层析成像.在该框架下,链路和路径的性能参数首先被映射为正常或者拥塞两个状态(分别用 0 和 1 表示),然后,根据链路性能参数模型的可分离性导出链路状态和路径状态之间的关系:一条路径是正常的当且仅当它经过的所有链路都是正常的.基于该框架,Duffield 针对树状网络提出了一种快速而简单的算法 SCFS(smallest consistent failure set algorithm).图 7 是 SCFS 的操作示意图,其中,正常路径经过的链路用虚线表示.Dhamdhere 等人[12]将 SCFS 算法推广到了一般网络情形.Nguyen 等人[74]注意到:在某些网络场景下,链路性能参数并不严格满足可分离性,他们基于最大后验准则提出一种启发式方法对路径状态进行修正后再使用 SCFS 算法.此外,Matsuda 等人[75]利用压缩感知方法进行网络高丢包率链路的定位.Zarifzadeh 等人结合传统层析成像方法和布尔层析成像方法的优点,提出一种称为 Range Tomography 的框架[76].在该框架下,他们不仅能够识别出链路状态,而且在得知链路拥塞时能估计出链路丢包率范围,进而获取链路的具体拥塞程度.

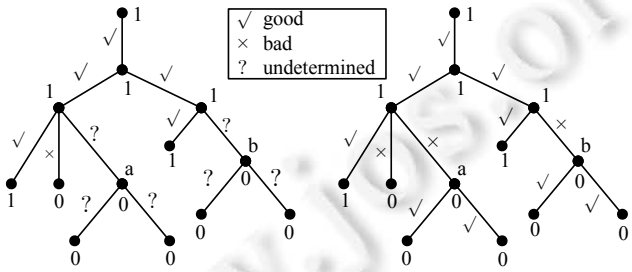


Fig.7 An illustration for SCFS[11]

图 7 SCFS 示意图[11]

4.2.2 基于多时隙测量

Nguyen 等人在布尔网络层析成像的框架下证明了链路的先验拥塞概率可以通过多个测量时隙的数据唯一确定,并提出一种基于矩阵求逆的方法 CLINK(congested LINK identification)对其求解;然后,将链路的先验拥塞概率和后续测量时隙的端到端数据结合,可以获得链路状态的最大后验估计[77].相较于 SCFS 算法,这种方法的准确度更高,特别是当网络中拥塞链路较多时,它具有更高的检测率.洛桑理工学院的 Ghita 等人将以上方法推广到更一般的场景中,发现并证明了网络中某些链路的状态不相互独立时,链路先验拥塞概率可辨识的充要条件[78],提出一种只需对少量端到端路径进行测量即可求得链路先验拥塞概率的方案[79].这一工作极大地推广

了故障链路诊断方法的适用范围.

CLINK,SCFS 以及 MCMC 算法的精度见表 3.

Table 3 Accuracy of the CLINK, SCFS and MCMC algorithms[77]

表 3 CLINK,SCFS 以及 MCMC 算法精度[77]

Topology	CLINK		SCFS		MCMC[82]	
	DR (%)	FPR (%)	DR (%)	FPR (%)	DR (%)	FPR (%)
Babarasi-Albert[80]	92.0	0.8	60.0	0.9	70.0	5.0
Waxman[80]	91.2	0.7	62.0	0.6	73.4	1.3
Hierarchical[80]	87.3	0.9	57.0	1.1	76.7	1.9
PlanetLab[81]	90.3	1.1	61.0	0.7	69.4	2.4

Nguyen 和 Ghita 等人在文献[82,83]中利用链路传输率和路径传输率的线性模型直接求解链路传输率的近似值.这种方法的求解目标和前文提及的最优化方法[68,70,71]类似,不同之处在于:他们认为链路丢包率的方差和相应的丢包率成正比,丢包率方差小的链路相应的丢包率也就小.由多个测量时隙的路径丢包率方差估计出链路丢包率方差后,在后续时隙中,将方差小于某门限的链路丢包率近似为 0,再求解其他链路的丢包率近似值.

这种方法的本质是:在保持端到端路径数不变的情况下,通过减少链路数目来解决链路丢包率估计的欠定性问题.

网络拥塞链路定位算法见表 4.

Table 4 Algorithms for congested link identification

表 4 网络拥塞链路定位算法

诊断策略	计算出链路性能参数(丢包率)的近似值,然后再判断链路状态	直接根据路径状态推断链路状态
单个时隙测量	文献[70,71,76]	文献[12,72−74]
多个时隙测量	文献[82,83]	文献[77−79]

4.3 链路中立性检测

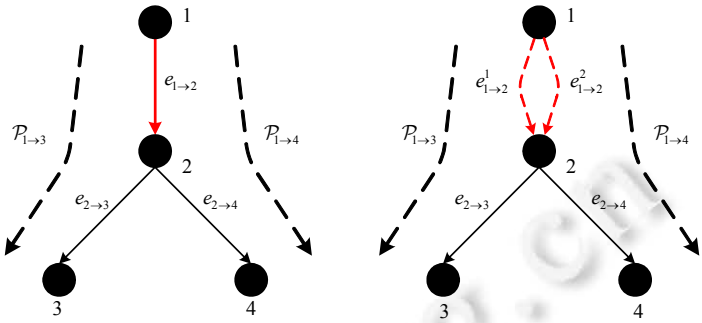
单播网络层析成像技术的一个基本假设是,不同路径在相同链路上经历的性能参数服从相同分布[5,7].当网络中所有链路都遵守中立性原则[84],平等对待所有经过它的流量时,该假设通常成立.然而随着互联网技术及应用的发展,网络中立性原则不再受到保障.例如,相比购买 2M 带宽的用户,购买 10M 带宽的用户能够获得更佳的在线高清视频直播体验.此外,随着网络 OTT(over the top)业务如微信(WeChat)[85]等的逐渐流行,运营商的传统核心业务(如语音、短信)正遭受挤压,从而使其寻求管控 OTT 流量.当前,关于网络是否应该保持中立性的争论还在继续.然而,一旦网络不具有中立性,现有大部分网络层析成像技术将面临着全面失效的命运.

文献[86−89]提出在同一条路径上传输不同类型的探测流,并对它们的性能参数(如时延、丢包率等)进行测量,从而检测该路径经过的链路是否按照流的类别区分服务.尽管这些方法能够检测网络路径上是否存在非中立现象,但并没有试图去定位具体是由哪条或哪些链路引起的.如图 8(a)所示:如果链路 $e_{1\to2}$ 是非中立的,那么路径$\mathcal{P}_{1\to3}$与$\mathcal{P}_{1\to4}$均能被检测到经历了非中立性链路.将非中立链路 $e_{1\to2}$ 抽象成两条虚拟链路 $e^1_{1\to2}$ 与 $e^2_{1\to2}$ (如图 8(b)所示),则有表 5 中的列满秩路由矩阵.这样的非中立性网络是较难进行非中立性链路定位的.在不考虑误差因素的情况下,Zhang 等人[10]首次系统地对链路中立性检测的层析成像方法进行了理论分析,发现:如果目标网络非中立且满足某些特定条件时,能够通过判断公式(3)的有解性等手段来定位非中立链路.

Table 5 Routing matrix of equivalent network in Fig.8

表 5 图 8 中所示等价网络的路由矩阵

	$e^1_{1\to2}$	$e^2_{1\to2}$	$e_{2\to3}$	$e_{2\to4}$
$\mathcal{P}_{1\to3}$	1	0	1	0
$\mathcal{P}_{1\to4}$	0	1	0	1



(a) 原始网络 (b) 等价网络

Fig.8 Non-neutrality network

图 8 非中立性网络

5 存在的问题与研究展望

网络层析成像自 1996 年被提出后[26],其研究也涉及到网络测量的方方面面.纵观网络层析成像技术研究的发展历程,一个重要的内在动力是技术的实用化.在当前网络层析成像研究中,网络拓扑的准确性是影响最终网络性能参数推断结果精度的一个重要因素[90],因其直接关系到网络层析成像基本系统模型的准确与否.此外,相比于大量网络层析成像推断算法的研究,当前对数据测量的研究相当少.文献[32]指出:目前已有的测量方案大多很少考虑测量频率、网络测量覆盖、通信与计算负担等对最终参数推断结果的影响,要么就只考虑极少数因素.此外,测量数据本身所含的噪声也通常被忽略,而这将会影响到最终推断结果的准确度与可信度[27].

综上所述,目前基于网络层析成像技术的链路性能参数估计研究中存在的问题主要有:

(1) 网络端到端测量方案的可操作性以及其测量数据的有效性与准确性不能很好地满足实际网络测量需要.如采用多播测量方式,除要求严格的时钟同步外,还要求端节点提供分布式测量协作的支持等.此外,缺乏对网络层析成像中测量噪声的讨论以及建模研究;

(2) 拓扑模型近似存在误差.如实际端到端测量路径可能并不形成一颗树,还有如流量均衡带来的端到端测量路径模糊性问题.拓扑的准确度将直接关联系统模型误差的大小;

(3) 链路时空独立性假设不合理.TCP 通信流的链路丢包与时延都具有很强的时间相关性.另外,由于网络管理策略的相似以及二层物理设备共享等因素,链路之间还会具有空间相关性;

(4) 网络链路统计性能平稳性等假设存在缺陷.如果测量周期足够小,往往可以认为链路性能在测量周期内是平稳的.而如背靠背探测等,当网络规模较大时其测量周期往往较长,此时的链路性能是否平稳需要依实际情况来考量[23,45].

网络层析成像技术的实用化是网络层析成像未来发展的一个时代要求.主要包括 3 方面:(1) 研究更具实用性与有效性的测量机制;(2) 去除不合理假设,按照实际应用需要来改进层析成像数学模型,研究兼具计算速度与精度的推断算法;(3) 拓宽网络层析成像技术的应用范围.根据以上 3 个方面,具体的研究方向分为:

① 消除网络雷达[28]测量时反向路径对测量数据的噪声影响.

网络雷达不需要端节点的协作以及时钟同步,而且其测量复杂度低,易于实施与操作管理.如果能够对测量数据中由反向路径(如图 5(b)所示)引入的噪声进行合理建模,并将其有效地消除,网络雷达将可以胜任当前几乎所有的网络层析成像测量任务.这势必将极大地推动网络层析成像技术的实用化.

② 提升网络测量数据的准确性与有效性.

由于网络流量的高突发性,使得链路性能参数的往往具有非平稳特性,而当前模拟多播测量的单播测量机制常常假设在其测量周期内链路性能参数平稳.但当网络规模很大时,背靠背探测等需要跨越非常大的测量周期,而这将很难保证平稳性等假设的合理性.因而,如何根据链路性能平稳性确立最佳测量周期以及综合考虑网

络测量覆盖与测量频率等问题,设计更加准确有效的网络测量方式,将会是未来一大研究热点.

③ 对端到端测量数据噪声和网络层析成像系统模型噪声进行合理地数学建模.

端到端测量因为需要端节点的协作以及时钟同步等,测量数据中往往存在着误差;对基于树型拓扑建模的网络层析成像而言,如果测量路径并不形成树型结构,那么系统模型将存在建模误差.研究如何合理对它们进行噪声建模,将会有效地提高网络层析成像算法的准确度与可靠性.

④ 链路性能参数时空相关性建模.

如在 TCP 通信流中,链路的丢包与时延等将会具有时间相关性;而当网络的管理策略相似或者逻辑链路之间共享了某些二层物理设备时,链路性能将会具有空间相关性.这就需要根据相应的网络测量方式对链路性能参数进行合理的时空相关性建模,以期提升网络层析成像模型对目标网络描述的准确度.

⑤ 拓宽网络层析成像技术的应用范围.

除了主要面对计算机网络,更为广泛的一些网络科学问题也可以尝试利用网络层析成像技术进行解决,如社交网络[91]中社团发现问题等.此外,在公共交通运输领域[92]可以借鉴当前网络拥塞链路定位的思路,将其应用于定位与预测拥塞道路.根据定位与预测结果控制红绿灯时间长短,进而达到均衡各道路车流量的目的.

6 总 结

网络链路性能参数是重要的网络性能状态指标,准确地掌握这些参数对于网络的高效管理等具有极其重要的作用.在当前类 traceroute 工具所测(环回路径性能)数据包含有低转发优先级的 ICMP 报文测量部分以及 Internet 中大多数网络不提供测量协助的背景下,基于端到端测量的网络层析成像测量技术被提出来.为了克服网络层析成像方法所遇到的欠定性问题,端到端测量往往被用来获取报文级相关性,但 Internet 对多播路由协议的支持度较低,从而使得模拟多播测量的单播测量方式被大量应用到实际端到端测量中.通过端到端路径级测量数据,网络层析成像技术能够在不需要网络内部节点提供测量协助的情况下依然有效地推断出目标网络的链路级性能参数,如链路丢包率、链路拥塞状态等.根据目标参数对链路性能的刻画程度,链路性能参数网络层析成像方法主要分为定量参数估计方法与定性参数估计方法.由于定性参数主要利用测量数据性能级相关性,使得其估计结果具有相对较好的稳定性.针对不同类型的链路性能参数的不同网络层析成像估计方法,总结分析了统计推断方法、时频分析方法以及矩估计等方法在链路性能参数推断层析成像问题上的应用.指出当前网络层析成像技术存在的问题,并结合网络层析成像技术的发展历程展望其未来研究几个重要方向以及潜在的应用场景.

References:

[1] Chiara C, Enrico G, Luciano L, Dmitri K. Evolution of the Internet *k*-dense structure. IEEE/ACM Trans. on Networking, 2014, 22(6):1769−1780. [doi: 10.1109/TNET.2013.2282756]

[2] CAIDA: Cooperative association for Internet data analysis. http://www.caida.org

[3] Forouzan BA, Fegan SC. TCP/IP Suite. 4th ed., Columbus: McGraw-Hill Inc., 2010.

[4] Donnet B, Friedman T. Internet topology discovery: A survey. IEEE Communications Surveys & Tutorials, 2007,9(4):56−69. [doi: 10.1109/COMST.2007.4444750]

[5] Lawrence E, Michailidis G, Nair V, Xi B. Network tomography: A review and recent developments. Ann Arbor, 2006,1001: 48109−1107.

[6] RFC3439. http://www.ietf.org/rfc/rfc3439.txt

[7] Coates M, Hero III AO, Nowak R, Yu B. Internet tomography. IEEE Signal Processing Magazine, 2002,19(3):47−65. [doi: 10. 1109/79.998081]

[8] Qian F, Hu GM. A survey of network tomography technologies. Computer Science, 2006,33(9):12−16 (in Chinese with English abstract).

[9] Li YJ, Cai WD, Wang W. A survey on network tomography. Computer Engineering, 2006,32(13):91−93 (in Chinese with English abstract).

[10] Zhang ZY, Mara OS, Argyraki K. Network neutrality inference. In: Krishnamurthy A, Ratnasamy S, eds. Proc. of the ACM Conf. on SIGCOMM. Chicago: ACM Press, 2014. 63−74. [doi: 10.1145/2619239.2626308]

[11] Duffield NG. Network tomography of binary network performance characteristics. IEEE Trans. on Information Theory, 2006, 52(12):5373−5388. [doi: 10.1109/TIT.2006.885460]

[12] Dhamdhere A, Teixeira R, Dovrolis C, Diot C. NetDiagnoser: Troubleshooting network unreachabilities using end-to-end probes and routing data. In: Kurose J, Schulzrinne H, eds. Proc. of the 2007 Conf. on Emerging Networking EXperiments and Technologies (CoNEXT). New York: ACM Press, 2007. 18. [doi: 10.1145/1364654.1364677]

[13] Logical topology. https://en.wikipedia.org/wiki/Logical_topology

[14] MINC: Multicast-Based inference of network-internal characteristics. http://gaia.cs.umass.edu/minc

[15] Mao YY, Kschischang FR, Li BH, Pasupathy S. A factor graph approach to link loss monitoring in wireless sensor networks. IEEE Journal on Selected Areas in Communications, 2005,23(4):820−829. [doi: 10.1109/JSAC.2005.843557]

[16] Cisco. Configure IP multicast routing. http://www.cisco.com/en/US/docs/ios/12_2/ip/configuration/guide/1cfmulti.html

[17] Arya V, Duffield NG, Veitch, D. Temporal delay tomography. In: Merrill D, Saha D, eds. Proc. of the IEEE INFOCOM. Phoenix: IEEE Communications Society, 2008. 276−280. [doi: 10.1109/INFOCOM.2008.64]

[18] Bu T, Duffield NG, Presti F, Towsley D. Network tomography on general topologies. ACM SIGMETRICS Performance Evaluation Review, 2002,30(1):21−30. [doi: 10.1145/511399.511338]

[19] Castro R, Coates M, Liang G, Nowak R, Yu B. Network tomography: Recent developments. Statistical Science, 2004,19(3): 499−517. [doi: 10.1214/088342304000000422]

[20] Coates M, Nowak R. Network loss inference using unicast end-to-end measurement. In: Mauri JL, ed. Proc. of the ITC Conf. on IP Traffic, Modeling and Management. Monterey: IEEE Computer Society, 2000. 28.

[21] Duffield NG, Presti FL, Paxson V, Towsley D. Inferring link loss using striped unicast probes. In: Sengupta B, ed. Proc. of the IEEE INFOCOM. Anchorage: IEEE Communications Society, 2001. 915−923. [doi: 10.1109/INFCOM.2001.916283]

[22] Duffield NG, Presti FL, Paxson V, Towsley D. Network loss tomography using striped unicast probes. IEEE/ACM Trans. on Networking (TON), 2006,14(4):697−710. [doi: 10.1109/TNET.2006.880182]

[23] Fei GL. Research on unicast end-to-end measurements based network performance parameters estimation methods [Ph.D. Thesis]. Chengdu: UEST of China, 2012 (in Chinese with English abstract).

[24] Zhao HH, Chen M. Topology inference based on network tomography. Ruan Jian Xue Bao/Journal of Software, 2010,21(1): 133−146 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3692.htm [doi: 10.3724/SP.J.1001.2010.03692]

[25] Zhou JJ, Yang JH, Yang Y, Zhang H. Research on traffic matrix estimation. Ruan Jian Xue Bao/Journal of Software, 2007, 18(11):2669−2682 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/18/2669.htm [doi: 10.1360/jos182669]

[26] Vardi Y. Network tomography: Estimating source-destination traffic intensities from link data. Journal of the American Statistical Association, 1996,91(433):365−377. [doi: 10.1080/01621459.1996.10476697]

[27] Krishnamurthy A, Singh A. Robust multi-source network tomography using selective probes. In: Greenberg A, Sohraby, K, eds. Proc. of the IEEE INFOCOM. Orlando: IEEE Communications Society, 2012. 1629−1637. [doi: 10.1109/INFCOM.2012.6195532]

[28] Tsang Y, Yildiz M, Barford P, Nowak R. Network radar: Tomography from round trip time measurements. In: Proc. of the ACM SIGCOMM on Internet Measurement. Oregon: ACM Press, 2004. 175−180. [doi: 10.1145/1028788.1028809]

[29] Zhang X, Philips C. A survey on selective routing topology inference through active probing. IEEE Communications Surveys & Tutorials, 2012,14(4):1129−1141. [doi: 10.1109/SURV.2011.081611.00040]

[30] Gu Y, Jiang GF, Singh V, Zhang YP. Optimal probing for unicast network delay tomography. In: Mandyam G, Westphal C, eds. Proc. of the IEEE INFOCOM. San Diego: IEEE Communications Society, 2010. 1−9. [doi: 10.1109/INFCOM.2010.5461920]

[31] Xu WY, Mallada E, Tang A. Compressive sensing over graphs. In: Ni LM, Zhang WJ, eds. Proc. of the IEEE INFOCOM. Shanghai: IEEE Communications Society, 2011. 2087−2095. [doi: 10.1109/INFCOM.2011.5935018]

[32] Jaggard AD, Kopparty S, Ramachandran V, Wright RN. The design space of probing algorithms for network-performance measurement. In: Harchol-Balter M, ed. Proc. of the ACM SIGMETRICS. Pittsburgh: ACM Press, 2013. 105−116. [doi: 10.1145/2465529.2465765]

[33] Cáceres R, Duffield NG, Horowitz J, Towsley D. Multicast based inference of network internal loss characteristics. IEEE Trans. on Information Theory, 1999,45(7):2462−2480. [doi: 10.1109/18.796384]

[34] Lawrence E. Flexicast network delay tomography [Ph.D. Thesis]. Michigan: University of Michigan, 2005.

[35] Xi BW. Estimating internal link loss rates using active network tomography [Ph.D. Thesis]. Michigan: University of Michigan, 2004.

[36] Liang G, Yu B. Maximum pseudo likelihood estimation in network tomography. IEEE Trans. on Signal Processing, 2003,51(8): 2043−2053. [doi: 10.1109/TSP.2003.814464]

[37] Su H, Li Y, Lin S, Jin D, Zeng L. Inference of link loss rates by explicit estimation. IET Communications, 2010,4(5):540−550. [doi: 10.1049/iet-com.2008.0602]

[38] Guo D, Wang XD. Bayesian inference of network loss and delay characteristics with applications to TCP performance prediction. IEEE Trans. on Signal Processing, 2003,51(8):2205−2218. [doi: 10.1109/SSP.2003.1289509]

[39] Arya V, Duffield NG, Veitch D. Multicast inference of temporal loss characteristics. Performance Evaluation, 2007,64(9): 1169−1180. [doi: 10.1016/j.peva.2007.06.002]

[40] Yajnik M, Sue M, Kurose J, Towsley, D. Measurement and modelling of the temporal dependence in packet loss. In: Doshi B, Omidyar G, Rosenberg C, eds. Proc. of the IEEE INFOCOM. New York: IEEE Communications Society, 1999. 345−352. [doi: 10.1109/INFCOM.1999.749301]

[41] Fei GL, Hu GM. Accurate and effective inference of network link loss from unicast end-to-end measurements. IET Communications, 2012,6(17):2989−2997. [doi: 10.1049/iet-com.2011.0779]

[42] Tsang Y, Coates M, Nowak R. Network delay tomography. IEEE Trans. on Signal Processing, 2003,51(8):2125−2136. [doi: 10.1109/TSP.2003.814520]

[43] Presti FL, Duffield NG, Horowitz J, Towsley D. Multicast-Based inference of network-internal delay distributions. IEEE/ACM Trans. on Networking, 2002,10(6):761−775. [doi: 10.1109/TNET.2002.805026]

[44] Lawrence E, Michailidis G, Nair VN. Network delay tomography using flexicast experiments. Journal of the Royal Statistical Society, 2006,68(5):785−813. [doi: 10.1111/j.1467-9868.2006.00567.x]

[45] Coates M, Nowak R. Sequential monte carlo inference of internal delays in nonstationary data networks. IEEE Trans. on Signal Processing, 2002,50(2):366−376. [doi: 10.1109/78.978391]

[46] Shih MF, Hero III AO. Unicast-Based inference of network link delay distributions with finite mixture models. IEEE Trans. on Signal Processing, 2003,51(8):2219−2228. [doi: 10.1109/TSP.2003.814468]

[47] Duffield NG, Horowitz J, Presti L, Towsley D. Network delay tomography from end-to-end unicast measurements. In: Proc. of the Evolutionary Trends of the Internet. Berlin: Springer-Verlag, 2001. 576−595. [doi: 10.1007/3-540-45400-4_37]

[48] Matlab. http://www.mathworks.com

[49] The network simulator version 2 (ns2). http://www-isi.edu/nsnam/ns/

[50] Xia Y, Tse D. Inference of link delay in communication networks. IEEE Journal of Selected Areas in Communications, 2006, 24(12):2235−2248. [doi: 10.1109/JSAC.2006.884022]

[51] Scott DW. Multivariate Density Estimation: Theory, Practice and Visualization. New York: Wiley & Sons, 1922. [doi: 10.1002/9780470316849]

[52] Airoldi E, Faloutsos C. Recovering latent time-series from their observed sums: Network tomography with particle filters. In: Kim W, Kohavi R, eds. Proc. of the 10th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Seattle: ACM Press, 2004. 30−39. [doi: 10.1145/1014052.1014059]

[53] Amara G. An introduction to wavelets. IEEE Computational Science & Engineering, 1995,2(2):50−61. [doi: 10.1109/99.388960]

[54] Chen A, Cao J, Bu T. Network tomography: Identifiability and Fourier domain estimation. IEEE Trans. on Signal Processing, 2010, 58(12):6029−6039. [doi: 10.1109/TSP.2010.2068294]

[55] Cover TM, Thomas JA. Elements of Information Theory. New York: Wiley & Sons, 2012.

[56] Lakshminarayanan K, Padmanabhan VN, Padhye J. Bandwidth estimation in broadband access networks. In: Lombardo A, ed. Proc. of the 4th ACM SIGCOMM Conf. on Internet Measurement. Taormina: ACM Press, 2004. 314−321. [doi: 10.1145/1028788.1028832]

[57] Croce D, Mellia M, Leonardi E. The quest for bandwidth estimation techniques for large-scale distributed systems. ACM SIGMETRICS Performance Evaluation Review, 2010,37(3):20−25. [doi: 10.1145/1710115.1710120]

[58] Gjoka M, Fragouli C, Sattari P, Markopoulou A. Loss tomography in general topologies with network coding. In: Waring D, Krishnaswamy D, eds. Proc. of the IEEE GLOBECOM. Washington: IEEE Communications Society, 2007. 381−386. [doi: 10.1109/GLOCOM.2007.78]

[59] Lin Y, Liang B, Li B. Passive loss inference in wireless sensor networks based on network coding. In: Zuckerman DN, Anderson R, ed. Proc. of the IEEE INFOCOM. Rio de Janeiro: IEEE Communications Society, 2009. 1809−1817. [doi: 10.1109/INFCOM.2009.5062101]

[60] Mansouri VS, Wong VWS. Link loss inference in wireless sensor networks with randomized network coding. In: Makki K, Znati T, Meyerson M, eds. Proc. of the IEEE GLOBECOM. Miami: IEEE Communications Society, 2010. 1−6. [doi: 10.1109/GLOCOM.2010.5683336]

[61] Fragouli C, Boudec JY, Widmer J. Network coding: An instant primer. ACM SIGCOMM Computer Communication Review, 2006, 36(1):63−68. [doi: 10.1145/1111322.1111337]

[62] Chen Y, Bindel D, Song HH, Katz RH. Algebra-Based scalable overlay network monitoring: Algorithms, evaluation, and applications. IEEE/ACM Trans. on Networking, 2007,15(5):1084−1097. [doi: 10.1109/TNET.2007.896251]

[63] Kim MS, Kin T, Shin Y, Lam SS, Powers EJ. A wavelet-based approach to detect shared congestion. IEEE/ACM Trans. on Networking, 2008,16(4):763−776. [doi: 10.1109/TNET.2007.905599]

[64] Abduljabbar MM. Shared congestion detection: A comparative study. Journal of Engineering, 2012,18(9).

[65] Yousaf MM, Welzl M, Yener B. Accurate shared bottleneck detection based on SVD and outliers detection. Technical Report, NSG-DPS-UIBK-01, Innsbruck: University of Innsbruck, 2008.

[66] Rubenstein D, Kurose J, Towsley D. Detecting shared congestion of flows via end-to-end measurement. IEEE/ACM Trans. on Networking, 2002,10(3):381−395. [doi: 10.1109/TNET.2002.1012369]

[67] Harfoush K, Bestavros A, Byers J. Robust identification of shared losses using end-to-end unicast probes. In: Matsushita Y, Ammar M, eds. Proc. of the Int'l Conf. on Network Protocols. Osaka: IEEE Computer Society, 2000. 22−33. [doi: 10.1109/ICNP.2000.896289]

[68] Padmanabhan VN, Qiu L, Wang HJ. Server-Based inference of Internet link lossiness. In: Bauer F, Puigjaner R, eds. Proc. of the IEEE INFOCOM. San Francisco: IEEE Communications Society, 2003. 145−155. [doi: 10.1109/INFCOM.2003.1208667]

[69] Liu J. Monte Carlo Strategies in Scientific Computing. Berlin: Springer-Verlag, 2001. [doi: 10.1007/978-0-387-76371-2]

[70] Lim H, Hou J. Identifying lossy links in wired/wireless networks by exploiting sparse characteristics. Computer Networks, 2007, 51(15):4442−4459. [doi: 10.1016/j.comnet.2007.06.021]

[71] Song H, Qiu L, Zhang Y. NetQuest: A flexible framework for large-scale network measurement. IEEE/ACM Trans. on Networking, 2009,17(1):106−119. [doi: 10.1109/TNET.2008.925635]

[72] Duffield NG. Simple network performance tomography. In: Papagiannaki K, ed. Proc. of the 3rd ACM SIGCOMM Conf. on Internet Measurement. Cluj-Napoca: ACM Press, 2003. 210−215. [doi: 10.1145/948205.948232]

[73] Zhang ZY, Fei GL, Yu FC, Hu GM. Improving the accuracy of boolean tomography by exploiting path congestion degrees. In: Oktug S, Ulema M, Mouftah H, eds. Proc. of the 2012 IEEE Symp. on Computers and Communications. Cappadocia: IEEE Press, 2012. 725−731. [doi: 10.1109/ISCC.2012.6249384]

[74] Nguyen HX, Thiran P. Using end-to-end data to infer lossy links in sensor networks. In: Pascual JD, ed. Proc. of the IEEE INFOCOM. Barcelona: IEEE Communications Society, 2006. [doi: 10.1109/INFOCOM.2006.271]

[75] Matsuda T, Nagahara M, Hayashi K. Link quality classifier with compressed sensing based on l_1-l_2 optimization. IEEE Communications Letters, 2011,15(10):1117−1119. [doi: 10.1109/LCOMM.2011.082911.111611]

[76] Zarifzadeh S, Gowdagere M, Dovrolis C. Range tomography: Combining the practicality of boolean tomography with the resolution of analog tomography. In: Byers J, Kurose J, eds. Proc. of the 2012 ACM SIGCOMM Conf. on Internet Measurement. Boston: ACM Press, 2012. 385−398. [doi: 10.1145/2398776.2398817]

[77] Nguyen HX, Thiran P. The boolean solution to the congested IP link location problem theory and practice. In: Baldwin RL, Qiao C, eds. Proc. of the IEEE INFOCOM. Anchorage: IEEE Communications Society, 2007. 2117−2125. [doi: 10.1109/INFCOM.2007.245]

[78] Ghita D, Argyraki K, Thiran P. Network tomography on correlated links. In: Allman M, ed. Proc. of the 10th ACM SIGCOMM Conf. on Internet Measurement. Melbourne: ACM Press, 2010. 225−238. [doi: 10.1145/1879141.1879170]

[79] Ghita D, KaraKus C, Argyraki K, Thiran P. Shifting network tomography toward a practical goal. In: Cho K, Crovella M, eds. Proc. of the 2007 Conf. on Emerging Networking EXperiments and Technologies (CoNEXT). Tokyo: ACM Press, 2011. 24. [doi: 10.1145/2079296.2079320]

[80] Medina A, Matta I, Byers J. On the origin of power laws in Internet topologies. ACM SIGCOMM Computer Communication Review, 2000,30(2):18−28. [doi: 10.1145/505680.505683]

[81] PlanetLab. http://www.planet-lab.org

[82] Ghita D, Nguyen H, Kurant M, Argyraki K, Thiran P. Netscope: Practical network loss tomography. In: Mandyam G, Westphal C, eds. Proc. of the IEEE INFOCOM. San Diego: IEEE Communications Society, 2010. 1−9. [doi: 10.1109/INFCOM.2010.5461918]

[83] Nguyen H, Thiran P. Network loss inference with second order statistics of end-to-end flows. In: Papagiannaki K, ed. Proc. of the 7th ACM SIGCOMM Conf. on Internet Measurement. San Diego: ACM Press, 2007. 227−240. [doi: 10.1145/1298306.1298339]

[84] Wu T. Network neutrality, broadband discrimination. Journal of Telecommunications and High Technology Law, 2003,2:141−179. [doi: 10.2139/ssrn.388863]

[85] Tencent Inc. WeChat. 2014. http://www.wechat.com.en/

[86] Dischinger M, Marcon M, Guha S, Gummadi K, Mahajan R, Saroiu S. Glasnost: Enabling end users to detect traffic differentiation. In: Castro M, Snoeren AC, eds. Proc. of the USENIX Symp. on Networked Systems Design and Implementation (NSDI). 2010. 405−418.

[87] Kanuparthy P, Dovrolis C. ShaperProbe: End-to-End detection of ISP traffic shaping using active methods. In: Thiran P, Willinger W, eds. Proc of the 2011 ACM SIGCOMM Conf. on Internet Measurement. Berlin: ACM Press, 2011. 473−482. [doi: 10.1145/2068816.2068860]

[88] Lu G, Chen Y, Birrer S, Bustamante FE, Cheung CY, Li X. End-to-End inference of router packet forwarding priority. In: Baldwin RL, Qiao C, eds. Proc. of the IEEE INFOCOM. Anchorage: IEEE Communications Society, 2007. 1784−1792. [doi: 10.1109/INFCOM.2007.208]

[89] Zhang Y, Mao ZM, Zhang M. Detecting traffic differentiation in backbone ISPs with NetPolice. In: Williamson C, ed. Proc. of the 9th ACM SIGCOMM Conf. on Internet Measurement. Chicago: ACM Press, 2009. 103−115. [doi: 10.1145/1644893.1644905]

[90] Ghita D, Argyraki K, Thiran P. Towards accurate and practical network tomography. ACM SIGOPS Operating Systems Review, 2013,47(1):22−26. [doi: 10.1145/2433140.2433146]

[91] Xing EP, Fu W, Song L. A state-space mixed membership block model for dynamic network tomography. The Annals of Applied Statistics, 2010,42(2):535−566. [doi: 10.1214/09-AOAS311]

[92] Dimitrakopoulos G, Demestichas P. Intelligent transportation systems. IEEE Vehicular Technology Magazine, 2010,5(1):77−84. [doi: 10.1109/MVT.2009.935537]

附中文参考文献:

[8] 钱峰,胡光岷.网络层析成像研究综述.计算机科学,2006,33(9):12−16.

[9] 李勇军,蔡皖东,王伟.网络断层扫描技术综述.计算机工程,2006,32(13):91−93.

[23] 费高雷.基于单播端到端测量的网络性能参数估计方法研究[博士学位论文].成都:电子科技大学,2012.

[24] 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. http://www.jos.org.cn/1000-9825/3692.htm [doi: 10.3724/SP.J.1001.2010.03692]

[25] 周静静,杨家海,杨扬,张辉.流量矩阵估算的研究.软件学报,2007,18(11):2669−2682. http://www.jos.org.cn/1000-9825/18/2669.htm [doi: 10.1360/jos182669]

潘胜利(1988－),男,湖北鄂州人,博士生,主要研究领域为网络层析成像.

钱峰(1978－),男,博士,副教授,主要研究领域为网络流分类,网络可视化.

张志勇(1985－),男,博士,讲师,主要研究领域为拥塞链路定位,网络中立性估计.

胡光岷(1966－),男,博士,教授,博士生导师,主要研究领域为复杂多维信号处理技术,网络可视化,网络测量,网络行为学,社交网络.

费高雷(1981－),男,博士,副教授,主要研究领域为网络测量,社交网络.