

基于网络层析成像技术的拓扑推断^{*}

赵洪华⁺, 陈 鸣

(解放军理工大学 指挥自动化学院, 江苏 南京 210007)

Topology Inference Based on Network Tomography

ZHAO Hong-Hua⁺, CHEN Ming

(Institute of Command Automation, PLA University of Science and Technology, Nanjing 210007, China)

+ Corresponding author: E-mail: honghua.nj@gmail.com

Zhao HH, Chen M. Topology inference based on network tomography. *Journal of Software*, 2010,21(1): 133-146. <http://www.jos.org.cn/1000-9825/3692.htm>

Abstract: Network tomography is a novel technique of network measurement in recent years. It combines network measurement with statistical inference, which can solve some problems that can not be solved by normal network measurement techniques. Network topology inference is one of the main applications of network tomography, which can infer the topology of network through end to end measurement and does not need the cooperation of internal nodes. In this paper, the technique of topology inference based on network tomography is summarized systematically, and the topology inference techniques and algorithms are compared and analyzed. Finally, the problems of current research and the future directions are discussed.

Key words: network tomography; topology inference; network measurement; correlation

摘 要: 网络层析成像是近年来新兴的一种网络测量技术,该技术结合了网络测量和统计推断,可以解决网络测量的部分难题.网络拓扑推断是网络层析成像技术的重要应用之一,基于网络层析成像技术的拓扑推断通过端到端的测量即可推断网络内部的拓扑结构,而不需要内部节点的合作.系统地总结了基于网络层析成像的拓扑推断技术的研究现状,对国内外研究进展进行了讨论,最后讨论了当前拓扑推断中存在的问题,并讨论了未来的研究方向.

关键词: 网络层析成像;拓扑推断;网络测量;相关性

中图法分类号: TP393 文献标识码: A

由于Internet具有分散性、异构性以及各子网分属不同ISP(Internet service provider)的特点,获得完整的网络拓扑比较困难.例如,IP级拓扑发现中,Internet网络大约有 13%的路由器对traceroute并不响应^[1].同时,出于安全的考虑,许多ISP对自己网络中的路由器作了特殊限制,从外部通过当前的测量方法无法获得网内的拓扑结构.因此,能否通过特殊的方式获得一个网络的拓扑结构,是一个更新的研究课题.

1996年,Vardi提出一个新的术语“网络层析成像(network tomography)”^[2].这是一种采用全新思想的网络测量技术,它是基于医学透视的概念提出来的,利用统计和推断的原理进行网络性能和拓扑结构的推理.虽然网络

* Supported by the National Natural Science Foundation of China under Grant No.90304016 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2007AA01Z48 (国家高技术研究发展计划(863))

Received 2008-08-18; Revised 2009-04-27; Accepted 2009-07-23

层析成像技术的研究还处于萌芽阶段^[3,4],但是它已经被认为是进行大规模网络性能评估、故障诊断和拓扑发现的最有前景的方法之一^[5,6].

网络层析成像技术的主要特点是能够基于端到端的测量技术来获取网络内部的特性,而不需要网络内部节点的合作.网络拓扑推断是网络层析成像技术的最新应用之一,它根据网络中节点性能特性的相关性来推断网络的拓扑结构.因为研究表明,在树状结构的网络中,节点的共享链路越多,节点的性能特性越相近,即相关性越大^[7,8].

基于网络层析成像的拓扑推断主要有 3 个步骤:

- ① 通过端到端的测量技术获得端到端的性能参数;
- ② 根据端到端的性能参数计算出节点间的相关性;
- ③ 根据节点间的相关性重构节点的连接情况从而推断网络的拓扑结构.

基于层析成像技术的拓扑推断分别从以上 3 个方面展开研究,例如,如何较准确、方便地测量端到端的性能参数,如何根据性能参数计算节点间的相关性以及如何通过节点间的相关性高效、准确地推断网络拓扑结构.

基于层析成像技术的拓扑推断最先应用于多播网络,然后才推广到单播网络.目前,多播网络主要存在于内部网络或私有网络,在 Internet 中,只有 IETF 建立的 MBONE 网络是多播网络.由于实际网络中支持多播的设备较少,因此,研究单播网络中的拓扑推断更有实际意义.而且,单播网络中的拓扑推断技术更为复杂.

1 网络层析成像技术的应用

目前,网络层析成像技术的研究主要分为 3 个方面:① 网络拓扑推断,即根据网络端到端的性能特性推断网络内部的拓扑结构^[8,9];② 网络链路级参数估计^[10-14].链路级参数估计通过端到端的网络特性分析网络内部链路级的特性,链路级参数估计包括链路时延、链路丢包率和链路带宽等参数的估计;③ 网络端到端(original-destination)流量估计^[2,15-19]、端到端流量估计是网络链路级参数推理的反过程,目的是利用链路级的流量数据推断网络端到端的流量.

大多数网络层析成像问题可以描述成如下的数学模型:

$$Y_t = AX_t + \varepsilon \quad (1)$$

其中: Y_t 是在时间 t 上的可观察到的测量向量,如数据包的数量或是端到端的时延; A 是路由矩阵; X_t 是在时间 t 上的数据分组的相关参数向量,如分组在链路的平均时延、分组成功传输的概率值或是随机的源端到目的端的流量矩阵; ε 是噪声向量.

网络层析成像技术的链路级参数推断过程是根据 Y_t 、 A 和 ε 来求解 X_t 的数学过程,这种求解方式在数学中被称为逆问题求解.网络层析成像技术中的逆问题求解依赖于矩阵 A 和向量 ε 的特性,可以通过选择合适的迭代方法来获得 X_t .

基于网络层析成像技术的链路级参数估计主要是研究每条链路的时延或丢包率等网络性能特性的估计,例如,通过测量的端到端时延和路由矩阵来估计链路的时延,由于路由矩阵 A 通常不是满秩的,所以在推断过程中,通常采用极大似然估计和蒙特卡洛算法^[20,21],Venkata提出了一种利用贝叶斯函数推断网络内部特性的方法^[22,23].链路级参数估计可以通过端到端的测量技术获得网络内部链路的特性,对于故障定位和性能评价具有重要的意义^[5].

基于网络层析成像技术的端到端流量估计主要是推断有多少网络流量从特定的源节点流向特定的目标节点^[15].流量估计通过记录经过网络节点的数据包的数量来获得测量数据,然后对测量数据进行统计推断得到端到端的流量估计值.这种方法可以直接使用在拥有管理权的网络上,以完成数据的采集工作.

2 基于层析成像技术的拓扑推断

Ratnasamy和McCanne于1999年在网络丢包率测量中观察到了节点相关性的特性,并把基于丢包率测量观

察到的节点相关性首先应用于多播网络拓扑结构的推断^[10].Duffield等人严格证明了基于节点相关性推断网络拓扑结构的正确性,并应用了其他网络性能参数^[8],例如时延、丢包率、链路利用率、时延分布等性能参数.由于采用单一的性能参数推断拓扑时受到网络负载的影响,Duffield在2002年提出了一种在多播网络中结合链路利用率和丢包率计算节点相关性的拓扑推断算法^[24],通过链路利用率和丢包率的结合可以使用的网络负载范围明显大于单一参数的情况,但链路利用率的测量占用较多的网络带宽,并且这种方法只适用于多播网络,所以在实际应用中受到较多限制.

为了提高多播网络拓扑推断的准确性并降低计算的复杂性,Tian等人^[25-27]提出了基于链路空闲率计算Hamming距离的网络拓扑推断算法,基于Hamming距离的拓扑推断适用于网络负载较小的环境.李勇军等人^[28,29]对该种方法进行了改进,通过改变链路之间的关联函数解决了算法在特殊情况下误判的概率.基于Hamming距离的拓扑推断算法只能应用在网络负载较小的情况下,并且需要节点间的时钟同步,所以,在实际应用中也受到较多限制.由于实际网络中支持多播的设备较少,因此,单播网络的拓扑推断更有实际意义和研究价值,而且单播网络的拓扑推断也是拓扑推断中的难点.

2.1 拓扑推断中的测量方法

无论是在单播网络还是在多播网络,在测量端到端性能时都有3个假定条件:① 空间独立性,即数据分组在不同链路上是相互独立的;② 时间独立性,即同一链路上不同数据分组是相互独立的;③ 网络拓扑在测量过程中是固定不变的.

拓扑推断中,为了准确计算节点间的相关性,测量端到端的性能时要求到达不同节点的测量分组在节点共享链路部分经历的网络状况相同,在多播网络中较容易实现到达不同节点的测量分组在节点共享链路部分经历相同的网络状况,而单播网络中的端到端性能测量要复杂得多,需要采用特殊的测量方法.当前,单播网络中使用比较多的几种测量方法是紧接分组对测量方法^[30,31]、“三明治”分组列车测量方法^[32]和四元分组列车测量方法^[33].

2.1.1 紧接分组对测量方法

紧接分组对由两个大小相同的数据分组组成,分组之间具有较小的时间间隔,并且分别到达不同的目标地址,把紧接分组对中两个数据分组的目标地址称为目标节点对.

由于紧接分组对中两个分组的时间间隔较小,因此采用紧接分组对可以实现到达不同目标节点的数据分组在节点的共享链路部分经历的网络状况相同.为了减少测量产生的流量,通常紧接分组对中的分组长度都很小,例如可设为50B.

在采用紧接分组对测量节点的端到端性能时,由测量源节点向目标节点对发送紧接分组对,根据目标节点收到测量分组的情况计算端到端单向时延或丢包率(单向时延需要节点间时钟同步),根据单向时延或丢包率即可计算出两个节点的相关性数值.

图1为在树状结构的网络拓扑中紧接分组对测量节点性能的实例,图中根节点0向目标节点对(3,5)发送紧接分组对,紧接分组对在节点对(3,5)的共享链路部分紧密相邻.

2.1.2 “三明治”分组列车测量方法

“三明治”分组列车由两个短的数据分组和一个长的数据分组组成,两个短分组具有相同的目标地址,长分组具有另一个目标地址.两个短分组之间的时间固定,长分组位于两个短分组中间,长分组的长度要远大于短分组的长度并且要小于网络的最大分组长度.

在“三明治”分组列车中,设长分组的长度为 $s(q)$,短分组的长度为 $s(p)$,设“三明治”分组列车经过的链路为 L_1, L_2, \dots, L_m ,链路的带宽为 b_1, b_2, \dots, b_m ,则在目标节点对共享链路部分长分组影响第2个短分组排队时延的条件如公式(2)所示^[32]:

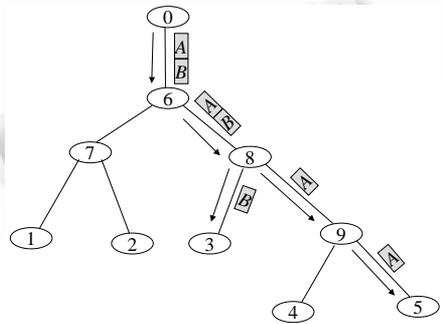


Fig.1 Instance of back-to-back probe

图1 紧接分组对测量实例

$$\frac{s(q)}{s(p)} \geq \frac{b_{i+1}}{b_i}, i = 1, 2, \dots, m \quad (2)$$

设在“三明治”分组列车中两个短分组之间的时间间隔为 d , 到达目标节点后短分组之间的时间间隔为 $d + \Delta d$, 在目标节点对非共享链路部分的背景流量对短分组的时间间隔不产生影响的条件如公式(3)所示^[32]:

$$\frac{s(p)}{d + \Delta d} \leq \min(b_i), m + 1 \leq i \leq n \quad (3)$$

在“三明治”分组列车方法测量节点的性能时, 由测量源节点向测量目标节点对发送“三明治”分组列车, “三明治”分组列车通过长分组在网络中产生排队时延, 使短分组之间的时间间隔增大, 根据目标节点收到短分组的时间间隔计算排队时延. 该方法不需要节点间的时钟同步, 并且排队时延即可作为节点间的相关性.

图2为“三明治”分组列车测量端到端性能的具体实例, 图中在树状结构的拓扑中, 根节点0向目标节点对(3,5)发送“三明治”分组列车, 测量目标节点对(3,5)在共享链路的排队时延.

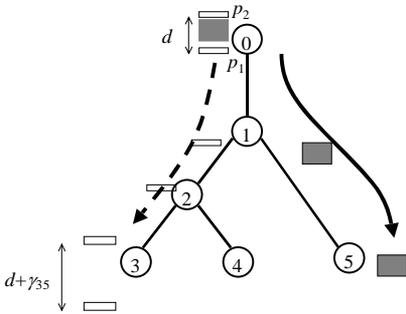


Fig.2 Instance of “sandwich” probe
图2 “三明治”分组列车测量实例

2.1.3 四元分组列车测量方法

四元分组列车由两个紧接分组对组成, 把一个四元分组列车表示为 $(x_1 - y_1, x_2 - y_2)$, 其中: $x_1 - y_1$ 表示四元分组列车中的第1个紧接分组对, $x_2 - y_2$ 表示四元分组列车中的第2个紧接分组对; 分组 x_1, x_2 具有相同的地址, 分组 y_1, y_2 具有相同的地址. 为了减少测量产生的流量, 通常四元分组列车中的分组长度都很小, 例如可设为 $50B$.

在四元分组列车中, 紧接分组对的时间间隔要满足下面的条件: 设四元分组列车经过的链路为 L_1, L_2, \dots, L_m , 各条链路的带宽为 b_1, b_2, \dots, b_m , 设四元分组列车中分组的长度为 $s(p)$, 则四元分组列车内紧接分组对的时间间隔 T_1 要满足公式(4):

$$s(p) / T_1 < \min(b_i), 1 \leq i \leq m \quad (4)$$

如图3所示, 为采用四元分组列车测量时延抖动的具体实例. 在图3中的四元分组列车为 $(B_2 - A_2, B_1 - A_1)$, 分组 B_1 和分组 B_2 的目的节点都为节点3, 分组 A_1 和分组 A_2 的目的节点都为节点2, 在测量源节点分组 B_1 和分组 B_2 及分组 A_1 和分组 A_2 之间的时间间隔固定, 在节点2和节点3记录收到测量分组的时间, 然后计算分组之间的时间间隔变化. 四元分组列车测量方法通过在目标节点计算测量分组的时间间隔变化来计算时延抖动^[34], 因此不需要节点间的时钟同步.

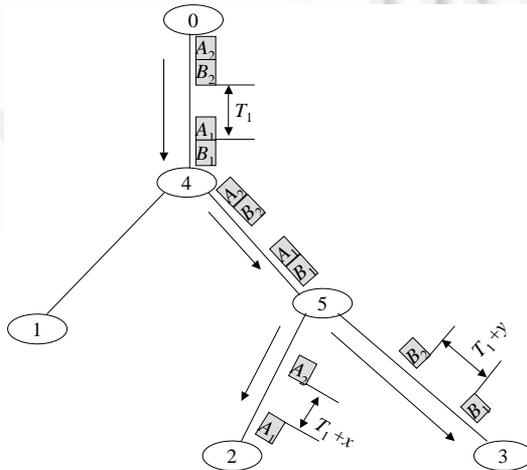


Fig.3 Instance of 4-packets train probe
图3 四元分组列车测量方法实例

2.1.4 测量方法分析

紧接分组对测量方法、“三明治”分组列车测量方法和四元分组列车测量方法都是单播网络中拓扑推断的测量方法,3种测量方法测量的网络性能参数、实现难度及适用的网络环境等都存在差别,表1对比了3种测量方法。

Table 1 Comparison of three measurement methods

表1 3种测量方法的对比

Measurement method	Time synchronization	Parameters	Environment	Complexity
Back-Back packets	Y	One way delay	Moderately loaded	Normal
Back-Back packets	N	Loss	Highly loaded	Normal
Sandwich packets	N	Queue delay	Lightly loaded	Complex
4-Tuple packets train	N	Delay variation	Moderately loaded	More complex

单播网络中的性能测量每次只能测量一个目标节点对,而拓扑推断需要任意两个节点间的相关性值,因此需要测量所有的目标节点对.假设目标节点的数量为 N ,则目标节点对的数量为 $N \times (N-1)$,为 $O(N^2)$ 数量级.当目标节点数较多时,需要测量的节点对数量非常大.拓扑推断中,为了准确计算节点间的相关性,在实际计算中采用多次测量计算均值的方式来提高计算的准确度,通常采用几百或上千次的测量值.因此,单播网络中的测量产生较多的流量。

我们以紧接分组对测量方法来分析测量产生的流量,其他方法与紧接分组对测量方法类似.假设采用紧接分组对测量方法对于每个目标节点对测量50组数据,每组20个样本值,紧接分组对中每个分组的长度为50B,则分组对的大小为100B,测量源向每个节点对发送的测量流量约为100KB.设节点数为 N ,测量产生的流量为 $flux$,那么,

$$flux = 100 \times N \times (N-1) \text{ (KB)} \quad (5)$$

通过公式(5)可以发现,在单播网络拓扑推断中,测量所产生的流量非常大,假设 $N=20$,则测量产生的流量约为40MB,此流量已对网络产生影响.当目标节点较多时,测量流量会影响到网络的正常运行.因此,减少测量流量是一个非常重要的问题。

2.2 节点相关性计算

基于层析成像技术的拓扑推断根据测量的节点端到端的性能计算节点间的相关性,对于节点间的相关性采用下面的定义。

定义1. 设 $T=(V,L)$ 表示树状拓扑,其中, V 为节点集, L 为链路集, 0 表示根节点同时也是测量源, R 为叶节点集同时也是测量的目标节点集合, $R=\{1, \dots, i, \dots, j\}$.令 $a(i,j)$ 为叶节点 i 和叶节点 j 的父节点, $n_i \prec n_j$ 表示 n_i 是 n_j 的子孙,则 $i \prec a(i,j)$, $j \prec a(i,j)$.如果 $a(k,j) \prec a(i,j)$,则节点 k 和 j 的共享链路比节点 i 和 j 的共享链路要多。

定义2. 节点间的相关性值是一种在树状结构的网络拓扑中,反映叶节点在网络中共享链路特性的数值,叶节点的共享链路越多,节点间相关性值越大,反之亦然。

在单播网络的拓扑推断中采用的节点间相关性参数主要有排队时延^[32]、时延协方差^[35]、成功传输率^[31]、时延抖动协方差^[33],每种参数都只适应于相应的网络环境^[33,35]。

2.2.1 排队时延

排队时延由“三明治”分组列车方法测量得到,采用排队时延推断网络的拓扑结构在网络负载较小时效果最好,采用“三明治”分组列车方法测量排队时延的方法如下:

设在测量源节点“三明治”分组列车中两个短数据包的间隔为 T ,由测量源节点向测量目标节点对 (i,j) 发送“三明治”分组列车,短分组的目标节点为 i ,长分组的目标节点为 j .设共发送 N 个“三明治”分组列车,在目标节点 i 共收到 k 个“三明治”分组列车中的短分组对,收到短分组的时间为 $\{T_{1,1}, T_{1,2}, T_{2,1}, T_{2,2}, \dots, T_{k,1}, T_{k,2}\}$,则节点对 (i,j) 的相关性计算如公式(6)所示:

$$\bar{T}_{i,j} = \sum_{m=1}^k (T_{m,2} - T_{m,1}) / k \quad (6)$$

由排队时延的测量可知,在树状结构的网络拓扑中,叶节点对的共享链路越多,则排队时延越大.通过“三明治”方法测量排队时延的计算较为简单,而且由于只计算时间间隔,所以不需要节点间的时间同步.

2.2.2 时延协方差

根据紧接分组对的方法测量的端到端单向时延可以计算节点间的时延协方差,时延协方差可作为节点间的相关性.采用时延协方差推断网络拓扑结构在网络负载适中时效果最好,采用紧接分组对测量端到端单向时延计算时延协方差的方法如下:

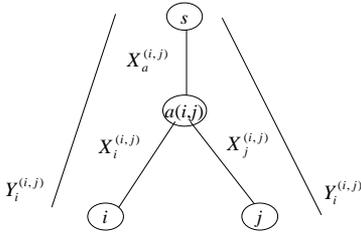


Fig.4 Correlation based on end to end delay
图4 根据端到端时延计算相关性

如图4所示,设紧接分组对测量的目标节点对为*(i,j)*,*a(i,j)*为节点*i*和*j*的父节点, $Y_i^{(i,j)}$, $Y_j^{(i,j)}$ 表示通过紧接分组对测量的到达节点*i*和*j*的端到端单向时延向量. $X_a^{(i,j)}$, $X_i^{(i,j)}$, $X_j^{(i,j)}$ 分别表示链路($S \rightarrow a(i,j)$)、链路($a(i,j) \rightarrow i$)和链路($a(i,j) \rightarrow j$)的单向时延向量.

根据紧接分组对测量的目标节点对*(i,j)*的端到端时延,并根据测量的假定条件,分组在不同链路是相互独立的,可以计算目标节点对*(i,j)*的时延协方差,如公式(7)所示:

$$Cov(Y_i^{(i,j)}, Y_j^{(i,j)}) = Cov(X_a^{(i,j)} + X_i^{(i,j)}, X_a^{(i,j)} + X_j^{(i,j)}) = Var(X_a^{(i,j)}) \quad (7)$$

把图4的情况一般化,设($S \rightarrow a(i,j)$)由链路 L_1, L_2, \dots, L_n 组成, $X_1^{(i,j)}, X_2^{(i,j)}, \dots, X_n^{(i,j)}$ 为链路 L_1, L_2, \dots, L_n 的单向时延向量,根据假定的条件各链路是相互独立的,则

$$Var(X_a^{(i,j)}) = Var(X_1^{(i,j)} + X_2^{(i,j)} + \dots + X_n^{(i,j)}) = Var(X_1^{(i,j)}) + Var(X_2^{(i,j)}) + \dots + Var(X_n^{(i,j)}) \quad (8)$$

由公式(7)、公式(8)可知,在树状结构的网络拓扑中,叶节点对的共享链路越多,则叶节点对的时延协方差越大.

2.2.3 成功传输率

成功传输率可以通过紧接分组对的方法测量得到,采用成功传输率推断网络拓扑在网络负载较大丢包较多时效果最好,成功传输率的计算方法如下所示.

如图5所示,设紧接分组对测量的目标节点对为*(i,j)*,*a(i,j)*为节点*i*和*j*的父节点.设 $P_a^{(i,j)}$, $P_i^{(i,j)}$, $P_j^{(i,j)}$ 分别表示链路($S \rightarrow a(i,j)$)、链路($a(i,j) \rightarrow i$)和链路($a(i,j) \rightarrow j$)的成功传输率,设 $S_i^{(i,j)}$ 为测量源节点*S*到叶节点*i*的成功传输率, $S_j^{(i,j)}$ 为测量源节点*S*到叶节点*j*的成功传输率, $S^{(i,j)}$ 表示紧接分组对的两个分组都成功到达目标节点的概率.设在节点*s*发送的紧接分组对数量为*N*,在目标节点*i*收到测量分组的数量为 N_i ,在目标节点*j*收到测量分组的数量为 N_j ,在目标节点*i*和目标节点*j*收到相同的紧接分组对中测量分组的数量为 N_{ij} .

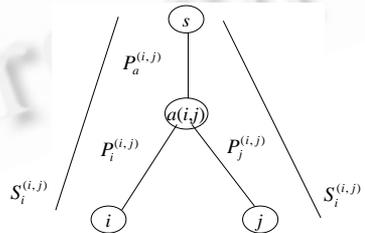


Fig.5 Correlation based on loss
图5 根据丢包率计算相关性

根据上面的分析,我们可以得到如公式(9)所示的结果:

$$\left. \begin{aligned} S_i^{(i,j)} &= \frac{N_i}{N}, S_j^{(i,j)} = \frac{N_j}{N}, S^{(i,j)} = \frac{N_{ij}}{N} \quad (N \rightarrow \infty) \\ P_a^{(i,j)} &= S_i^{(i,j)} \times S_j^{(i,j)} / S^{(i,j)} = \frac{(P_a^{(i,j)} \times P_i^{(i,j)}) \times (P_a^{(i,j)} \times P_j^{(i,j)})}{P_a^{(i,j)} \times P_i^{(i,j)} \times P_j^{(i,j)}} \end{aligned} \right\} \quad (9)$$

设链路($S \rightarrow a(i,j)$)由链路 L_1, L_2, \dots, L_n 组成, $P_1^{(i,j)}, P_2^{(i,j)}, \dots, P_n^{(i,j)}$ 分别为链路 L_1, L_2, \dots, L_n 的成功传输率,则

$$P_a^{(i,j)} = P_1^{(i,j)} \times P_2^{(i,j)} \times \dots \times P_n^{(i,j)} \quad (10)$$

由于 $P_x^{(i,j)} \leq 1$,根据公式(10),在树状网络拓扑中叶节点对的共享链路越多则成功传输率 $P_a^{(i,j)}$ 越小,在计算

中可以通过对数计算对 $P_a^{(i,j)}$ 进行转换.

2.2.4 时延抖动协方差

时延抖动可以通过四元分组列车测量方法得到,根据测量的时延抖动即可计算时延抖动协方差.采用该参数推断网络拓扑在网络负载适中时效果最好,时延抖动协方差的计算方法如下:

如图 6 所示,设四元分组列车的目的节点对为 (i,j) , $a(i,j)$ 为节点 i 和 j 的父节点, $Y_i^{(i,j)}, Y_j^{(i,j)}$ 表示通过四元分组列车测量的到达节点 i 和 j 的端到端时延抖动向量, 设 $J_a^{(i,j)}, J_i^{(i,j)}, J_j^{(i,j)}$ 分别表示链路 $(S \rightarrow a(i,j))$ 、链路 $(a(i,j) \rightarrow i)$ 和链路 $(a(i,j) \rightarrow j)$ 的时延抖动向量, 计算节点 i 和节点 j 的时延抖动协方差.

由于分组在不同链路是相互独立的, 因此, $Cov(J_i^{(i,j)}, J_a^{(i,j)}) = 0$, $Cov(J_a^{(i,j)}, J_j^{(i,j)}) = 0, Cov(J_i^{(i,j)}, J_j^{(i,j)}) = 0$,

$$Cov(Y_i^{(i,j)}, Y_j^{(i,j)}) = Cov(J_a^{(i,j)} + J_i^{(i,j)}, J_a^{(i,j)} + J_j^{(i,j)}) = Var(J_a^{(i,j)}) \quad (11)$$

设链路 $(S \rightarrow a(i,j))$ 由链路 L_1, L_2, \dots, L_n 组成, $J_1^{(i,j)}, J_2^{(i,j)}, \dots, J_n^{(i,j)}$ 分别为链路 L_1, L_2, \dots, L_n 的时延抖动向量, 由于各链路相互独立, 则

$$Var(J_a^{(i,j)}) = Var(J_1^{(i,j)} + J_2^{(i,j)} + \dots + J_n^{(i,j)}) = Var(J_1^{(i,j)}) + Var(J_2^{(i,j)}) + \dots + Var(J_n^{(i,j)}) \quad (12)$$

由公式(11)和公式(12)可知,在树状网络拓扑中,叶节点对的端到端时延抖动的协方差是目的节点对共享链路的时延抖动方差之和,节点对的共享链路越多,则端到端时延抖动的协方差越大.

2.2.5 相关性参数分析

排队时延通过“三明治”分组列车测量方法得到.当网络负载较小时,端到端排队时延完全由“三明治”分组列车中的长分组产生,无背景流量的干扰,因此测量的结果较为准确,推断网络拓扑的正确性最高;当网络负载适中时,网络中的背景流量较多,背景流量影响排队时延的概率增大,“三明治”分组列车测量的效果变差;当网络负载较大时,背景流量影响排队时延的概率最大,并且由于丢包导致测量的样本数减少,“三明治”分组列车的测量结果最差.

时延协方差通过紧接分组对测量的单向时延计算得到.当网络负载适中时,测量分组在共享链路的时延变化较大,节点间共享链路的时延方差较大;在网络负载较小时,测量分组在共享链路的时延变化较小,节点间共享链路的时延协方差较小;而在网络负载较大时丢包较多,影响到紧接分组对的测量.所以,时延协方差在网络负载适中时能够准确地反映节点共享链路的情况,推断网络拓扑的正确性最高.

成功传输率通过紧接分组对测量的丢包率计算得到.当网络负载适中或网络负载较小时,网络中的丢包较少,成功传输率不能准确地反映节点共享链路的情况;只有在网络负载较大时丢包较多,丢包率较大.采用成功传输率可以较好地反映节点间共享链路的情况.

时延抖动协方差通过四元分组列车方法测量的时延抖动计算得到.当网络负载较小时,由于网络的流量较小导致测量分组的时延抖动较小,因此,计算的时延抖动协方差的数值也较小;而在网络负载适中时,测量分组的时延抖动较大,计算的时延抖动协方差也较大,效果最好;在网络负载较大时,由于丢包较多,导致样本数减少,影响到测量的效果.

Meng等人在NS2 仿真中分别使用了时延协方差、成功传输率和排队时延推断网络拓扑结构^[31],分别在不同网络负载的情况下观察了3种相关性参数推断拓扑结构的效果,仿真中采用的拓扑结构如图7所示.Zhao等人在NS2 仿真中分别使用了时延协方差和时延抖动协方差^[33],在网络负载适中和网络负载较大的情况下对比了采用时延抖动协方差和时延协方差推断拓扑结构的效果,仿真中采用的拓扑结构如图8所示.

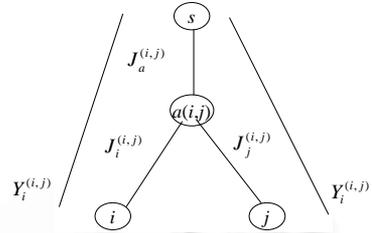


Fig.6 Correlation based on jitter

图6 根据时延抖动计算相关性

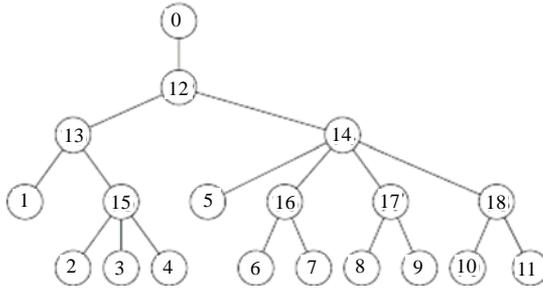


Fig.7 Simulation topology 1
图 7 仿真拓扑结构 1

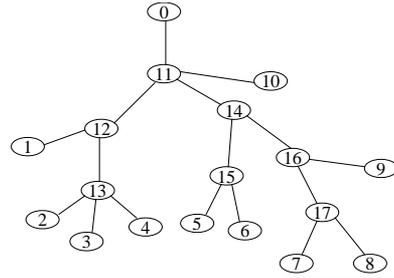


Fig.8 Simulation topology 2
图 8 仿真拓扑结构 2

图 9 为采用图 7 所示的网络拓扑,在网络负载较小、网络负载适中和网络负载较大的情况下,分别采用 3 种相关性参数推断网络拓扑结构的情况(*Sandwich*为排队时延,*Cov*为时延协方差,*Loss*为成功传输率).其中,横坐标 N_1 为测量的组数, N_2 为每组测量的样本数,纵坐标为平均树编辑距离^[36].图 9(a)表示平均时延,图 9(b)表示平均时延方差,图 9(c)表示平均丢包率.

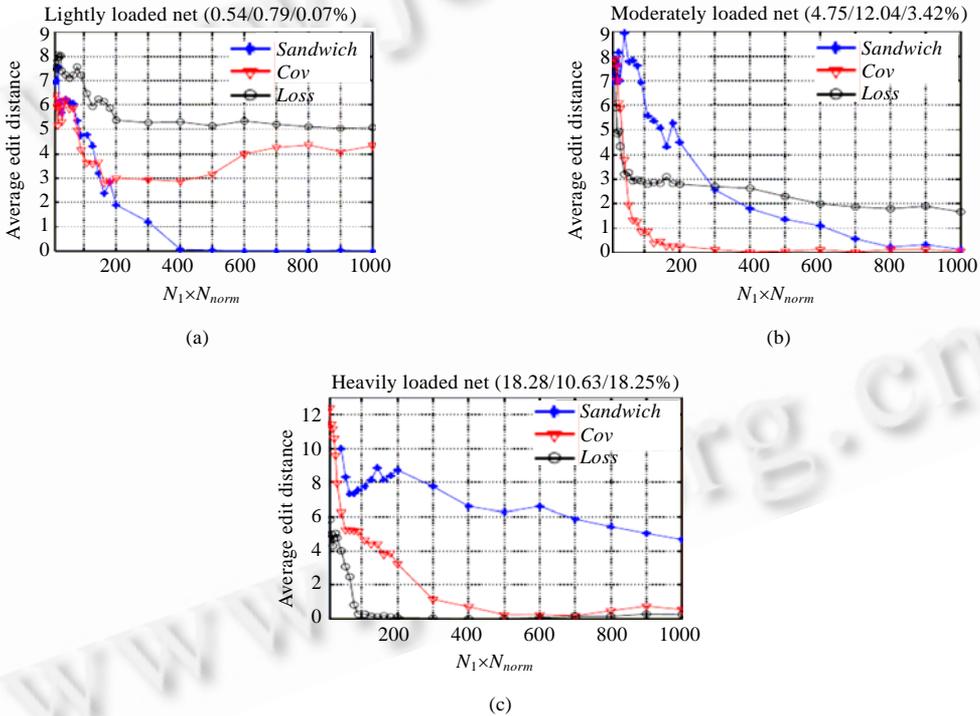


Fig.9 Topology inference based on different parameters-1
图 9 不同参数推断拓扑的情况-1

由图 9 可知:在网络负载较小时,采用排队时延推断拓扑结构的效果最好,采用时延协方差和成功传输率效果较差;在网络负载适中时,采用时延协方差推断拓扑结构的效果最好,排队时延次之,成功传输率最差;在网络负载较大时,采用成功传输率推断拓扑结构的效果最好,时延协方差次之,排队时延最差.

图 10 为采用图 8 所示的网络拓扑,在网络负载适中和网络负载较大时,采用时延抖动协方差(Internet packet delay variation,简称 IPDV)和时延协方差(one way delay,简称 OWD)推断拓扑结构的情况.

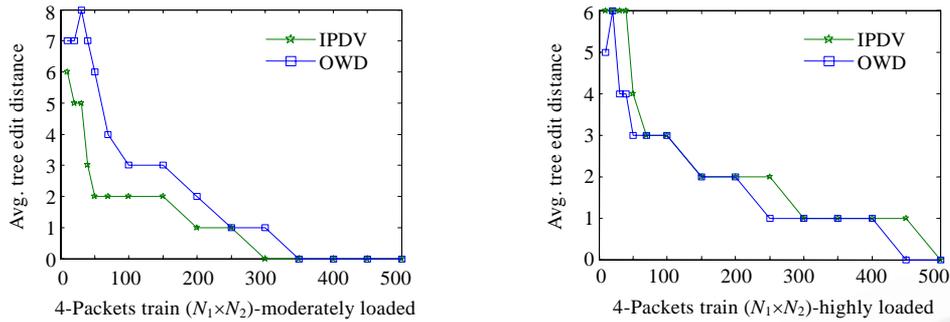


Fig.10 Topology inference based on different parameters-2

图 10 不同参数推断拓扑的情况-2

由图 10 可知:在网络负载适中时,采用时延抖动协方差推断拓扑结构的效果好于采用时延协方差推断拓扑结构的效果;而在网络负载较大时,采用时延抖动协方差推断拓扑结构的效果不如采用时延协方差推断拓扑结构的效果,但时延抖动的测量不需要节点间的时钟同步。

文献[30]提出了一种单播网络拓扑推断中根据不同的权值综合利用成功传输率和时延协方差计算节点相关性的方法,该方法可以适用于网络负载适中和网络负载较大的情况,比单一的采用成功传输率或时延协方差推断拓扑结构的效果要好,该方法的缺点是仍不能适用于网络负载较小的情况。

2.3 拓扑推断算法

网络拓扑推断算法根据节点间相关性值来推断网络的拓扑结构,当前的拓扑推断算法主要有 DBT 算法(deterministic binary tree classification algorithm)、LBT(likelihood based hierarchical binary tree)算法、MLT(maximum likelihood tree)算法和分层算法 HTE(hierarchical tree estimation)。

2.3.1 DBT 算法

Duffield 首先在多播网络中设计并应用了拓扑推断算法,称为 DBT 算法^[8,25]。DBT 算法是一种自底向上二叉树拓扑推断算法,该算法根据节点间的相关性值逐步合并相关性最大的两个节点以生成最终的拓扑树。DBT 算法的思想如下:

- (1) 从叶节点 R 中选择一对最相似的节点,设为节点 i, j ;
- (2) 合并两个节点并添加新的节点,设为 k ;
- (3) 更新节点 k 和其他叶节点的相关性;
- (4) 重复步骤 1~步骤 3,直到合并所有的节点。

DBT 算法中重要的步骤是计算节点间的相关性值,节点间的相关性值由计算函数 $B(G)$ 得到, $B(G)$ 的计算公式如下:

$$B(\{i, j\}) = \frac{\gamma(i)\gamma(j)}{\gamma(i) + \gamma(j) - \gamma(\{i, j\})} \quad (13)$$

其中, i, j 为接收节点并且 $i \neq j$, γ 为节点的性能参数(包括丢包率、时延等)。当插入新节点 k 时(k 由节点 i 和 j 生成),节点 k 与其他节点间的相关性值由计算函数 $U(G)$ 得到, $U(G)$ 的计算公式如下:

$$U(\{k, m\}) = \frac{\gamma(\{i, j\})\gamma(m)}{\gamma(\{i, j\}) + \gamma(m) - \gamma(\{i, m\})} \quad (14)$$

DBT 算法是拓扑推断中最早出现的拓扑推断算法,该算法的缺点是推断的拓扑结构都是二叉树结构,但该算法实现较为简单,是拓扑推断研究的先驱,为网络拓扑发现提供了新的技术和方法。

2.3.2 LBT 算法

Coates 在网络拓扑推断中对 DBT 算法进行了改进,提出了 LBT 算法^[37]。LBT 算法也是一种自底向上的拓扑推断算法,通过逐步合并节点以生成最终的拓扑树。LBT 算法与 DBT 算法的不同点在于节点间相关性值的计算

函数不同,LBT算法在节点间的相关性值计算中加入了概率密度函数,并采用了求解最大值的方法(通常为极大似然估计求解).通过概率密度函数和最大值求解,LBT算法在推断过程中能够选择局部范围内的最优解,并且推断的拓扑结构扩展为普通的拓扑树,LBT算法的相关性计算公式如下:

$$\gamma_{ij} = \arg \max(f_{ij}(x_{ij} | \gamma) + f_{ji}(x_{ji} | \gamma)) \quad (15)$$

其中, i,j 为接收节点并且 $i \neq j$, γ 为性能参数, f 为计算的概率密度函数.LBT算法与DBT算法相比提高了推断的准确性,把二叉树结构扩展为普通的拓扑结构.但LBT算法的实现和节点间相关性值的计算函数都较为复杂.

2.3.3 MLT 算法

MLT算法是Coates提出的一种全局搜索的拓扑推断算法^[32],MLT算法的基本思想是,根据节点间相关性值,在所有可能树结构集合 F 中找出与节点间相关性值最可能匹配的树结构 T^* ,在MLT算法中采用评估函数来确定选出的树是否为最优的,评估函数的计算公式如下:

$$L_i(x|T) = \log p(x|T, \hat{\mu}(T)) - \lambda n(T) \quad (16)$$

其中, x 为测量并计算获得的相关性值参数, $n(T)$ 为树 T 中的链路数, $\lambda \geq 0$, $\hat{\mu}(T)$ 为测量的平均时延, $p(x|T, \hat{\mu}(T))$ 为在树 T 中获得测量结果 x 的概率.

MLT中最复杂的部分是在所有的拓扑树中搜索最优的树,当节点数较多时,搜索的空间较大,为 $O(n!)$.因此,计算所有可能树的概率值是不现实的.MLT算法认为,拓扑树的分布呈尖峰形,所以搜索过程只在一部分树结构集合中.搜索算法采用反向马尔可夫链蒙特卡洛算法,简称MCMC(Markov chain Monte Carlo computation)^[38],搜索算法的基本思想如下:

- (1) 选择一个随机的开始状态 $s_0=(T_0, \mu_0)$;
- (2) 进入下一个状态 s_1 (添加,剪枝或 μ 跳),得到最小值

$$\min \left\{ 1, \frac{p(T_1, \mu_1 | x)q(s_0 | s_1)}{p(T_0, \mu_0 | x)q(s_1 | s_0)} \times J_{f(s_1, s_0)} \right\} \quad (17)$$

- (3) 重复步骤 1、步骤 2 的过程,直到建立一个树结构的拓扑.

在公式 17 中, $q(s_i|s_j)$ 表示从状态 s_j 到状态 s_i 的概率, $J_{f(s_i, s_j)}$ 表示从状态 s_j 到状态 s_i 的映射行列式.

2.3.4 HTE 算法

Meng等人提出了一种基于图论的分层推断算法(HTE)^[35],该算法的思想是,根据目的节点对的相关性把节点划分成不同的集合,然后逐级合并节点集合并加入新的节点,直到最后成为一个大的集合.算法描述如下:

- (1) 设接收节点集为 G , $\chi(G)=\{\gamma_{ij}; i < j, i, j \in G\}$ 为相关性集合;
- (2) 根据 γ_{ij} 划分接收节点集,划分的原则为:一个划分中的节点对相关性的最高,一个划分中的节点对相关性的要大于划分之间节点对的相关性;
- (3) 在每个划分中添加一个新节点 u ,计算出新节点之间的相关性,在计算新节点的相关性值时可以用每个划分中的一个节点作为代表,根据新计算的节点间相关性值重新进行划分;
- (4) 重复步骤 1~步骤 3,直到所有节点属于一个集合.

分层算法推断的拓扑结构不再是二叉树结构,比DBT算法推断的要准确;与MLT算法相比,分层算法的计算复杂度要低得多,分层算法的难点是节点类的划分.

2.3.5 拓扑推断算法分析

当前的拓扑推断算法基本上都基于上述 4 种算法,在 4 种拓扑推断算法中,算法的实现难度和推断拓扑结构的正确性都存在较大差别,表 2 对比了 4 种拓扑推断算法.

在 4 种拓扑推断算法中,推断准确性最高的是 MLT 算法,而 MLT 算法最大的缺点是实现难度较大;推断准确性最低的为 DBT 算法,但 DBT 算法最大的优点是实现简单;LBT 算法和 HTE 算法介于 DBT 算法和 MLT 算法之间,LBT 算法实现的难点是概率函数的选择和计算,HTE 算法的实现难点是节点类的划分.

Meng等人采用如图 7 所示的拓扑结构,在仿真环境下对比HTE,DBT和LBT这几种算法推断拓扑结构的情况,如图 11 所示^[35].在图 11 中,左图为 3 种算法推断拓扑结构的正确性对比,右图为 3 种算法推断拓扑的准确率

对比,其中,横坐标为复杂性指数.由图 11 可知,在 3 种拓扑推断算法中:在正确性和准确率方面,HTE算法最好;LBT算法略优于DBT算法;而MLT算法的复杂度较高,在如图 7 所示的拓扑中实现难度较大.

Table 2 Comparison of inference algorithm

表 2 拓扑推断算法的比较

Algorithm	Binary tree	Computing complex	Space complex	Apply complex	Correctness
DBT	Y	$O(n)$	$O(n)$	Simple	Normal
LBT	N	$O(n)$	$O(n)$	More difficult	More correct
MLT	N	Between $O(n)$ and $O(n!)$	$O(n)$	Difficult	Correct
HTE	N	$O(n)$	$O(n)$	More difficult	More correct

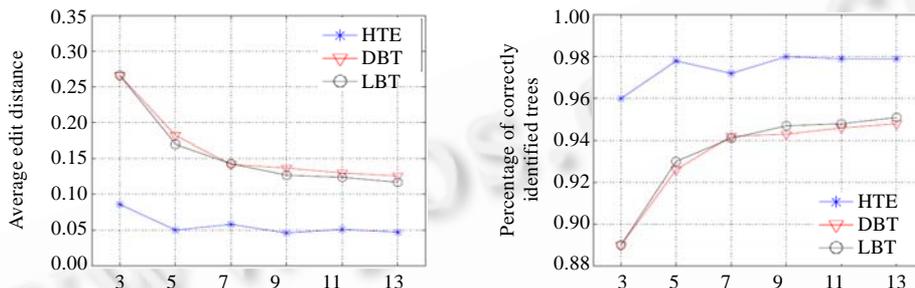


Fig.11 Topology inference based on different algorithms

图 11 不同推断算法推断拓扑的情况

以上所有推断算法都是采用单测量源的方式测量网络端到端性能,然后推断网络拓扑结构,所以推断结果都是树状结构的网络拓扑.Coates和Rabbat等人^[39]提出了一种在多播网络中采用多测量源拓扑推断的思想,该思想通过把多测量源多目标地址的结构分解为 2-to-2 的方式来推断内部的结构,但是这种思想与实际应用还存在较大的差距,并不能实用.

3 存在的问题及未来的研究方向

无论是在单播网络还是在多播网络,当前基于层析成像技术的拓扑推断都只能推断网络的逻辑拓扑结构而不能推断出实际的物理结构.网络的逻辑拓扑结构是网络中节点的逻辑连接情况,逻辑拓扑中的节点是两个叶节点在网络中分离时的节点,如图 12 所示为物理拓扑和逻辑拓扑的区别.

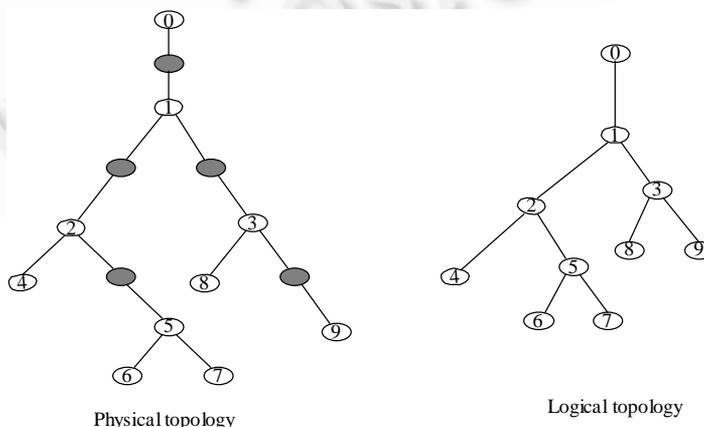


Fig.12 Physical topology and logical topology

图 12 物理拓扑结构与逻辑拓扑结构

在实际网络中以单播网络为主,因此基于多播的拓扑推断应用范围较少.而当前在单播网络中的拓扑推断还存在较多的问题,主要表现为以下几个方面:

- (1) 当前拓扑推断中的测量方法受到较多的限制.例如:“三明治”分组列车测量方法只适用于网络负载较小的情况,当网络负载较大时,推断拓扑结构的误差较大;紧接分组对方法测量端到端时延需要节点间的时钟同步,而时钟同步的代价较高;
- (2) 当前的测量方法在测量过程中产生较多的测量流量.由于单播网络中的测量方法需要测量所有的节点对之间的性能,测量的节点对数量为 $O(n^2)$ 级,因此,当节点数量较多时产生的测量流量非常多;
- (3) 当前的拓扑推断算法采用单一的网络性能参数计算节点间的相关性,采用单一性能参数的拓扑推断受到网络负载的影响较大.

网络拓扑推断是层析成像技术的重要应用之一,但由于该技术正处于起步阶段,还有较多的问题需要深入研究,主要有以下几个方面:

- (1) 减少拓扑推断技术的限制和问题.

对单播网络拓扑推断中的测量方法和节点相关性计算方法进行改进以提高其实用性.例如:对当前的测量方法进行改进,减少不必要的测量;在计算节点相关性时采用自适应综合利用多种参数的方式,以适应网络负载的变化;由于无论是在单播网络还是在多播网络,推断的拓扑结构都是逻辑拓扑,而在实际应用中,物理拓扑更为我们所需要,因此,能否推断物理拓扑结构也是一个需要研究的课题.

- (2) 采用多测量源的网络拓扑推断技术.

现有的基于网络层析成像技术的拓扑推断都是基于单测量源、多接收节点的方式,因此推断的拓扑结构都是树状的拓扑结构,如何把单测量源的拓扑推断发展为多测量源的拓扑推断是一个需要研究的问题.在多测量源的拓扑推断中需要判断拓扑的内部节点是否为同一个节点,由于对网络的内部结构一无所知,所以多测量源节点推断的网络拓扑的合并是下一步研究的关键问题.

- (3) 拓扑推断技术在实际网络中的应用.

研究网络层析成像技术的初衷是对 Internet 这种大规模网络进行性能监测和拓扑发现.然而,现有的网络测量方法和拓扑推断方法基本上都是在仿真网络环境下进行实验研究的.如何在考虑拓扑推断算法的实用性和精确性的同时把现有的基于网络层析成像技术的拓扑推断移植到大规模的、分布式的、网络负载较大的网络环境下,是当前拓扑推断研究面临的关键问题.

- (4) 拓扑推断技术在无线网络中的应用.

当前,无线网络的使用越来越广泛,人们对于无线网络的需求也日益增加.但是,基于网络层析成像技术的拓扑推断主要以有线网络为研究背景.目前,已经有研究人员对无线网络中的层析成像技术进行研究^[40,41].而无线网络不同于有线网络,无线网络的传输协议和节点特性等都与有线网络具有较大的区别.因此,如何将现有的网络拓扑推断技术应用到无线网络上会成为未来研究的重点内容.

4 总 结

网络层析成像技术是一种全新思想的网络测量技术,该技术结合统计推断的数学方法来解决网络测量的相关问题.网络拓扑推断技术是层析成像技术中的一个重要研究内容,本文讨论了基于网络层析成像技术的拓扑推断研究中的相关技术,论述了当前拓扑推断中的测量方法、相关性计算方法和拓扑推断方法,对比分析了拓扑推断中的测量方法、节点相关性计算方法和拓扑推断算法,最后讨论了当前拓扑推断技术中存在的问题和技术难点,并指出了下一步的研究方向.

References:

- [1] Donnet D, Friedman T. Internet topology discovery: A survey. IEEE Communications Surveys and Tutorials, 2007,9(4):2-15.

- [2] Vardi Y. Network tomography: Estimating source-destination traffic intensities from link data. *Journal of the American Statistical Association*, 1996,91:365–377.
- [3] Castro R, Coates M, Liang G, Nowak R, Yu B. Network tomography: Recent developments. *Statistical Science*, 2004,52(3): 499–517.
- [4] Bu T, Duffield N. Network tomography on general topologies. 2002. ftp://gaia.cs.umass.edu/pub/Bu_tomography_02.pdf
- [5] Chen NA, Cao J, Bu T. Network tomography: Identifiability and fourier domain estimation. In: Proc. of the INFOCOMM 2007. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4215800
- [6] Chen Y, Bindel D, Randy H. Tomography-Based overlay network monitoring. In: Proc. of the IMC 2003. Miami Beach, 2003. <http://www.imconf.net/imc-2003/papers/p330-chen1.pdf>
- [7] Ratnasamy S, McCanne S. Inference of multicast routing trees and bottleneck bandwidths using end-to-end measurements. In: Proc. of the IEEE INFOCOM'99. New York, 1999. <http://techreports.lib.berkeley.edu/accessPages/CSD-98-1019.html>
- [8] Duffield N, Horowitz J, Presti FL, Towsley D. Multicast topology inference from end-to-end measurements. In: Proc. of the ITC Seminar on IP Traffic, Measurement and Modelling. Monterey, 2000. <http://gaia.cs.umass.edu/minc/DHLT01-TOIT.pdf>
- [9] Duffield N, Horowitz J, Presti FL, Towsley D. Multicast topology inference from measured end-to-end loss. *IEEE Trans. on Information Theory*, 2002,48(1):26–45.
- [10] Xi B, Michailidis G, Nair V. Estimating network loss rates using active tomography. *Journal of the American Statistical Association*, 2006,101(476):1430–1448.
- [11] Duffield N, Presti FL, Paxson V, Towsley D. Network loss tomography using striped unicast probes. *IEEE/ACM Trans. on Networking*, 2006,14(4):697–710.
- [12] Shih M, Hero A. Unicast inference of network link delay distributions from edge measurements. In: Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. Salt Lake City: IEEE Computer Society, 2001. 3421–3424. <http://ieeexplore.ieee.org/iel5/7486/20359/00940576.pdf?arnumber=940576>
- [13] Lawrence E, Michailidis G, Nair V. Network delay tomography using flexi cast experiments. *Journal of the Royal Statistical Society, Series B*, 2006,68:785–813.
- [14] Hung X, Patrick T. Network loss inference with second order statistics of end-to-end flows. In: Proc. of the IMC 2007. 2007. <http://www.imconf.net/imc-2007/papers/imc70.pdf>
- [15] Singhal H, Michailidis G. Identifiability of flow distributions from link measurements with applications to computer networks. *Inverse Problems*, 2007,23(5):1821–1849.
- [16] Cao J, Davis D, Viel S, *et al*. A scalable method for estimating network traffic matrices from link counts. Technical Report, Bell Labs., 2001.
- [17] Cao J, Vander W, Yu B, Zhu Z. Time-Varying network tomography: Router link data. *Journal of the American Statistical Association*, 2000,95(452):1063–1075.
- [18] Vardi Y. Metrics useful in network tomography studies. *IEEE Signal Processing Letters*, 2004,11(3):353–355.
- [19] Airoidi E, Faloutsos, recovering latent time series from their observer sums: network tomography with particle filters. In: Proc. of the ACM KDD 2004. 2004. 30–39. http://portal.acm.org/ft_gateway.cfm?id=1014059&type=pdf
- [20] Lawrence E, Michailidis G, Nail V. Statistical inverse problems in active network tomography. 2007. <http://arxiv.org/pdf/0708.1079.pdf>
- [21] Liang G, Yu B. Maximum pseudo likelihood estimation in network tomography. *IEEE Trans. on Signal Processing*, 2003,51(8): 2043–2053.
- [22] Pluch P, Wakounig S. Bayesian network tomography and inference. 2007. <http://arxiv.org/abs/cs.NI/0702125v1>
- [23] Padmanabhan VN, Qiu L, Wang HJ. Passive network tomography using Bayesian inference. 2002. <http://www.imconf.net/imw-2002/imw2002-papers/134.pdf>
- [24] Duffield N, Horowitz J, Presti L. Adaptive multicast topology inference. 2001. <http://www.research.att.com/~duffield/papers/dhl01-adapt.pdf>
- [25] Tian H, Shen H. Hamming distance and hop count based classification for multicast network topology inference. In: Proc. of the 19th Int'l Conf. on Advanced Information. 2005. http://d.wanfangdata.com.cn/NSTLHY_NSTL_HY12436052.aspx

- [26] Tian H, Shen H. Analysis on binary loss tree classification with hop count for multicast topology discovery. In: Proc. of the IEEE Consumer Communications and Networking Conf. New York: IEEE, 2004. 164–168. http://d.wanfangdata.com.cn/NSTLHY_NSTL_HY17000577.aspx
- [27] Tian H, Shen H. Discover multicast network internal characteristics based on hamming distance. In: Proc. of the IEEE Int'l Conf. on Communications. Piscataway: IEEE Inc., 2005. 143–147. http://d.wanfangdata.com.cn/NSTLHY_NSTL_HY12425592.aspx
- [28] Li YJ, Cai WD, Wang W, Tian GL. Topology identification based on end-to-end link utilization. Journal of System Simulation, 2006,18(Z2) (in Chinese with English abstract).
- [29] Li YJ, Cai WD, Wang W, Tian GL. Research on network topology identification algorithm based on end-to-end loss performance. Journal on Communications, 2007,28(10):85–91 (in Chinese with English abstract).
- [30] 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 (in Chinese with English abstract).
- [31] Shih MF, Hero III AO. Topology discovery on unicast networks: A hierarchical approach based on end-to-end measurements. 2005. <http://www.eecs.umich.edu/~msim/Publications/cspl-357.ps.pdf>
- [32] Coates M, Castro R, Nowak R. Maximum likelihood network topology identification from edge-based unicast measurements. In: Proc. of the ACM Sigmetric. Marina Del Rey, 2002. <http://www.ece.wisc.edu/~nowak/finver3.pdf>
- [33] Zhao H, Chen M. Network topology inference based on delay variation. In: Proc. of the Int'l Conf. on Advanced Computer Control. Singapore, 2009. <http://portal.acm.org/citation.cfm?id=1510778>
- [34] Demichelis C, Chimento P. IP packet delay variation metric for IPPM. 2002. <http://www.ietf.org/internetdrafts/draft-ietf-ippm-ipdv-09.txt>
- [35] Shih MF, Hero III AO. Hierarchical inference of unicast network topologies based on end to end measurements. IEEE Trans. on Signal Processing, 2007,55(5):1708–1718.
- [36] Philip B. A survey on tree edit distance and related problems. 2005. http://grfia.dlsi.ua.es/ml/algorithms/references/editsurvey_bille.pdf
- [37] Castro R, Coates M, Nowak R. Likelihood based hierarchical clustering. IEEE Trans. on Signal Processing, 2004,52(8):2308–2321.
- [38] Green P. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika, 1995,82: 711–732.
- [39] Rabbat M, Nowak R, Coates M, Nowak R. Multiple source, multiple destination network tomography. In: Proc. of the INFOCOMM 2004. http://www.ieee-infocom.org/2004/Papers/34_4.PDF
- [40] Li YJ, Cai WD, Ji WL, Zhao T. Loss temporal dependency tomography in wireless sensor network. In: Davis GB, Liu NJ, eds. Proc. of the Int'l Conf. on Wireless Communications, Networking and Mobile Computing. Shanghai: IEEE Computer Society, 2007.
- [41] Wang W, Cai WD, Wang BZ, Li YJ, Tian GL. Mobile ad hoc network delay tomography. In: Proc. of the IEEE Int'l Workshop on Anti-Counterfeiting, Security, Identification. Xiamen: IEEE Computer Society, 2007. 365–370. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4244849

附中文参考文献:

- [28] 李勇军,蔡皖东,王伟,田广利.基于端到端链路利用的网络拓扑发现算法研究.系统仿真学报,2006,18(增刊 2).
- [29] 李勇军,蔡皖东,王伟,田广利.基于端到端报文丢失的网络拓扑推断算法研究.通信学报,2007,28(10):85–91.
- [30] 赵洪华,陈鸣,仇小锋,张国敏.Tomography 技术中的多参数网络拓扑推断.北京邮电大学学报,2008,31(4):24–28.



赵洪华(1979—),男,河北吴桥人,博士,讲师,主要研究领域为网络测量,网络拓扑发现.



陈鸣(1956—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机网络,网络测量,分布式计算.