维普资讯 http://www.cqvip.com

基于多源网络断层扫描的拓扑推断技术研究

田广利　蔡皖东

（西北工业大学计算机学院　西安 710072）

摘　要　网络断层扫描(NT)技术可以在内部节点不配合的情况下推断网络逻辑拓扑。多源 NT 比单源 NT 可以获取更多的拓扑和链路性能信息。四种 2-by-2 结构的逻辑拓扑的推断是多源 NT 的核心问题。基于四种 2-by-2 结构特点的分析，将问题转换为判定两条路径是否有共享链路的问题，并提出一种新的三明治探测技术，以端到端延时或者丢包率是否变化作为依据，解决了四种 2-by-2 结构的拓扑推断问题。仿真结果表明，此技术具有很高的正确性和有效性。

关键词　网络断层扫描，拓扑推断，端到端测量，三明治探测

Topology Inference Based on Multiple Source Network Tomography

TIAN Guang-li　CAI Wan-dong

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract　Network tomography techniques can infer the logical topology of network without the cooperation of nodes. Multiple source network tomography can obtain more information about the topology detail and the link performance than single source network tomography. How to differ from each other between the four 2-by-2 structures is the core technique of the multiple source network tomography. Based on the analysis of structures, the problem can be converted into how to judge whether two paths have a shared link. A new sandwich probing method is proposed to resolve the problem based on checking the variety of the end-to-end packets delay or loss. The simulation results indicate correctness and effectualness of the method.

Keywords　Network tomography, Topology inference, End-to-end measurements, Sandwich probing

1　引言

拓扑和链路状态信息是网络监测、优化和管理的基础。然而，由于网络的分布式特性，要从网络的所有元素中直接取得测量数据是不现实的。另外，由于对安全性要求的提高，人们不能依靠 Internet 网络元素对特殊用途的消息（如 ICMP 超时时戳）做出响应。网络断层扫描（Network Tomography, NT）技术是近年来国际上提出的一种新的网络测量技术，它将医学上的计算机断层扫描（Computed Tomography, CT）思想引入到网络测量中，根据网络外部（即网络端点或边界）的端到端测量来分析和推断网络的内部性能和拓扑结构。拓扑推断技术可以分为三类：(1)基于 traceroute 的网络层（ISO 第三层）拓扑推断；(2)基于 SNMP 的链路层（ISO 第二层）拓扑推断；(3)基于 NT 技术的逻辑拓扑推断。前两种拓扑推断依赖网络内部节点配合，基于 NT 技术的拓扑推断只能得到逻辑拓扑。一条逻辑链路可能包含多条物理链路，一个逻辑节点可能是一个子网，比物理拓扑丢失了一部分信息，且稳定性差一些，但它是前两种技术的补充。文献[1, 2]给出 NT 技术的一些综述。NT 技术的探测方式一般是基于单个源节点，多个接收节点大部分问题已得到解决。而探测方式采用多源节点的 NT 技术未得到广泛研究。多源节点 NT 技术的一个显著优点是可以利用多源节之间的合作来构造探测包，理论上可以得到更多的拓扑结构和更准确的链路信息，如图 1 所示，可以部分弥补逻辑拓扑相对物理拓扑缺少的信息。但同时由于 NT 技术的推断属于逆问题，要获取更多的内部链路信息，增加推断的难度和准确性，所以需要构造巧妙的探测包方式和推断算法。文献[3]提出了将多个单源节点推断的逻辑拓扑合并的算法。文献[4,5]改进了文献[3]的算法，提出了多源、多目的节点的 NT 技术，采用的探测包结构是半随机双包对结构，利用探测包到达目的节点的顺序区分出共享式结构与非共享结构。但此方法无法区分三种非共享结构，而非共享结构更能体现多源 NT 技术的优点。文献[5]同时给出 M-by-N 结构分解成 2-by-2 结构的算法和 2-by-2 结构合并的算法。本文的研究集中在多源 NT 技术中 2-by-2 结构的拓扑推断问题。

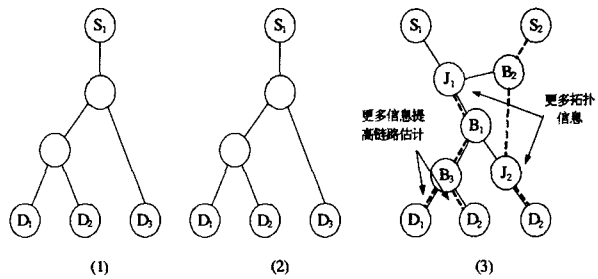

图 1　多源节点探测方式的 NT 技术的优点

2　2-by-2 结构分析和探测包设计

单源节点探测方式推断出的逻辑拓扑是一个逻辑树，可以分解为 1-by-2 型式，如图 2 所示。图中 B 节点表示分叉节点(branching nodes)，J 节点表示汇合节点(joining nodes)。多源的逻辑拓扑可以分解成 1-by-2 和 2-by-1 或者是 2-by-2 结构。而 2-by-1 结构的 NT 推断问题很难解决，所以 2-by-2 结构逻辑拓扑模型是多源 NT 技术研究的基础。根据分叉节点和汇合节点的位置不同，2-by-2 结构可以分为四种结构，如图 3 所示。其中(a)结构中两源节点具有共同的汇聚节点和分叉节点，称为共享式结构，具有两个中间节点和 5 个链路。

田广利　博士研究生，主要研究方向为无线网络通信和网络安全；蔡皖东　教授，博士生导师，主要从事计算机网络、分布式计算、多媒体通信、网络信息安全与对抗等学科方向的研究。

维普资讯 http://www.cqvip.com

后三种称为非共享式结构，具有 4 个中间结点和 8 个链路。所以非共享式结构比两个 1-by-2 探测获取更多的拓扑信息，比共享结构更能体现多源 NT 技术的优点。

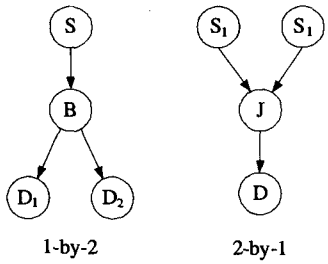

图 2　两种基本的探测方式

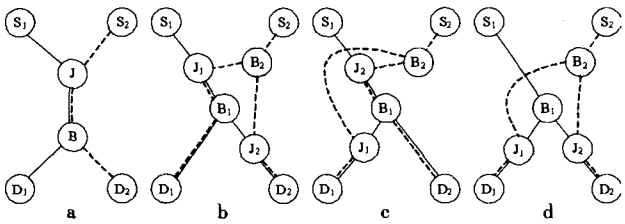

图 3　四种 2-by-2 拓扑结构

本文采用的符号含义如表 1 所示。

表 1　符号表

符号	含义
$P_{X_iY_j}$	节点 X_i 到节点 Y_j 的路径
$\xi_{X_iY_j}$	节点 X_i 到节点 Y_j 的逻辑链路
$\rho^k_{X_iD_j}$	节点 X_i 发送到节点 Y_j 的第 k 组数据包
$d_{X_iY_j}(n)$	节点 X_i 到节点 Y_j 的第 n 个数据包的延时
$d^q_{X_iY_j}(n)$	节点 X_i 到节点 Y_j 的第 n 个数据包的动态延时(主要是排队延时和传输延时)
$d^c_{X_iY_j}(n)$	节点 X_i 到节点 Y_j 的固定延时(主要是传播延时和处理延时)

通过分析四种 2-by-2 结构可以发现：在(a)结构中，路径 P_{S1D1} 和路径 P_{S2D2} 有共享链路 ξ_{JB}，路径 P_{S1D2} 和路径 P_{S2D1} 也有共享链路 ξ_{JB}；在(b)结构中，路径 P_{S1D1} 和路径 P_{S2D2} 无共享链路，路径 P_{S1D2} 和路径 P_{S2D1} 有共享链路 ξ_{J1B1}；在(c)结构中，路径 P_{S1D1} 和路径 P_{S2D2} 有共享链路 ξ_{J1B1}，路径 P_{S1D2} 和路径 P_{S2D1} 无共享链路；在(d)结构中，路径 P_{S1D1} 和路径 P_{S2D2} 无共享链路，路径 P_{S1D2} 和路径 P_{S2D1} 也无共享链路。所以区分四种 2-by-2 结构的关键转化为如何判断两条路径是否有共享链路的问题。若两条路径有共享链路，一条路径的数据流会对另一条路径的端到端性能造成影响，在网络负载较轻的情况下，延时变化较大；在负载较重的情况下，丢包率变化较大。若无共享链路，两条路径的数据流对彼此的端到端性能无任何影响。衡量一条路径的性能的变化，必须使用两组数据包：一组数据包保证在传输过程中不受另外一条路径数据包的影响，另外一组数据包则要尽量受到影响。为此本文设计了一种类似三明治的探测包，如图 4 所示。

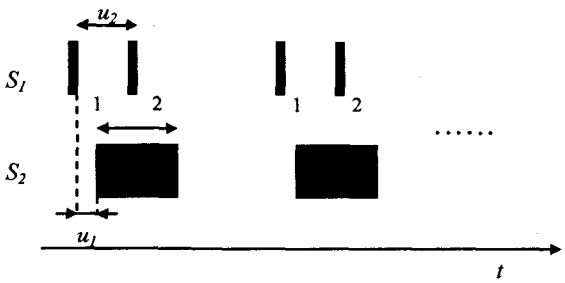

图 4　探测包结构示意图

不失一般性，我们下面只研究如何判断路径 P_{S1D1} 和路径 P_{S2D2} 是否有共享链路。源节点 S_1 向目的节点 D_1 发送两组小数据包，用于比较性能的变化，两组数据包的间隔为 u_2s。源节点 S_2 向目的节点 D_1 发送一组大数据包(为避免突发数据造成丢包，使用连续发送的数据包代替)，用于对第二组数据的性能造成影响，其持续时间为 Δs，其和第一组数据间隔为 u_1s。当网络负载较轻的时候，采用延时的变化作为依据较为合理；当网络负载较重时，采用丢包率的变化作为依据较合理。

假设路径 P_{S1D1} 和路径 P_{S2D2} 有共享链路 ξ_{JB}。在无丢包的情况下，两组数据的平均延时的差 Ω 为

$$\begin{aligned}\Omega &= \frac{1}{N}\sum_{n=1}^{N}(d^2_{S1D1}(n)-d^1_{S1D1}(n))\\ &= \frac{1}{N}\sum_{n=1}^{N}(d^c_{S1D1}(n)+d^{q2}_{S1J}(n)+d^{q2}_{JB}(n)+d^{q2}_{BD1}(n)-\\ &\quad d^c_{S1D1}(n)-d^{q1}_{S1J}(n)-d^{q1}_{JB}(n)-d^{q1}_{BD1}(n))\\ &= \sum_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}(d^{q2}_{JB}(n)-d^{q1}_{JB}(n))+E(d^q_{S1J})-E(d^q_{S1J})\\ &\quad +E(d^q_{BD1})-E(d^q_{BD1})\\ &= \sum_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}(d^{q2}_{JB}(n)-d^{q1}_{JB}(n))\end{aligned} \tag{1}$$

所以，两组数据包的平均延时为数据包 ρ_{S2D2} 中在逻辑链路 ξ_{JB} 上的所有节点的输入队列中。排队在数据包 ρ^2_{S1D1} 前面的部分的排队时间与处理时间的和。同时，若路径 P_{S1D1} 和路径 P_{S2D2} 无共享链路，根据公式(1)知 $\Omega=0$。网络负载较重时，Ω 的值与端到端的相对值较小，随机误差增加。我们可以使用两组数据的丢包率的差 Ψ 作为依据。当路径 P_{S1D1} 和路径 P_{S2D2} 无共享链路时，$\Psi \underset{N\to\infty}{=} 0$，否则 $\Psi \underset{N\to\infty}{>} 0$。

为提高判断路径有无共享链路的准确率和减少探测时间，需要精心设计 u_2、u_1 和 Δ 的值。假设源节点 S_1 和 S_2 具有近似时间同步，误差范围为 $(-c,c)$。为满足数据包 $\rho^1_{S1D1}(n)$ 不受数据包 $\rho_{S2D2}(n)$ 的影响，则需要满足

$$0+d_{S1J}+c+q_{\max}<u_2+d_{S2J} \tag{2}$$

其中 $q_{\max}$ 为数据包在节点的最大排队时间，一般为 10ms 级，可以忽略。同时，由于无法测量 d_{S1J} 和 d_{S2J}，但可以测量 d_{S2D2}，且 $d_{S1J}-d_{S2J}<d_{S1J}<d_{S1D1}$，所以放宽不等式条件得 $\min u_2=\max d_{S2D2}+c$。我们可以通过测量得到 $t_1=Ed_{S1D1}$。为减少探测时间，一般取 $u_2=3t_1+c$，可以保证绝大部分数据包 $\rho^1_{S1D1}(n)$ 不受数据包 $\rho_{S2D2}(n)$ 的影响。同样，为满足数据包 $\rho^2_{S1D1}(n)$ 受数据包 $\rho_{S2D2}(n)$ 的影响，且数据包 ρ_{S2D2} 中排队在数据包 ρ^2_{S1D1} 前面的部分大于一个常数 l，设 l 大小的数据包对应着数据包 $\rho_{S2D2}(n)$ 持续时间 Δ 中的时间为 h，则有

$$\begin{cases}u_1+d_{S1J}-(u_2+d_{S2J})>h+c\\ u_2+\Delta+d_{S2J}-(u_1+d_{S1J})>c\end{cases} \tag{3}$$

同样，放宽不等式条件得 $\min\Delta=2c+h+2\max\{d_{S1D1},d_{S2D2}\}$，$\min u_1=u_2+h+c+\max\{d_{S1D1},d_{S2D2}\}$。通过测量得到 $t_2=Ed_{S2D2}$，一般取 $\Delta=2c+h+6\max\{t_1,t_2\}$，$u_1=u_2+h+c+3\max\{t_1,t_2\}$，$h=0.1$s 即可。

3　仿真与结果

采用仿真工具 ns-2 对以上的方法进行验证。我们进行四组仿真实验，网络拓扑分别为四种 2-by-2 结构，四种结构中探测数据包和背景流量设置完全相同。两组小探测数据包

(下转第 142 页)

维普资讯 http://www.cqvip.com

说明该属性选择标准太粗略；(2)属性选择时，有可能选择的属性会产生很多叶子结点，这样，有些叶子结点上包含的实例数太少，以至这些叶子结点的泛化能力太小而没有意义，同时在预测时是比较费时。针对这些不足，我们给出了一种新的属性选择标准，即基于伴随正区域的属性选择标准。在新的属性选择标准中，有如下优点：(1) 继承了基于正区域的属性选择标准的优点，即优先考虑一个属性的正区域大小，这样能保证生成的决策树中，叶子结点深度较小。(2)在正区域大小相同的情况下，再考虑属性分出来的叶子数目(注意所选属性分出的叶子的决策属性值若相同则合并叶子)和叶子所覆盖的实例数，这样能保证生成的决策树的叶子数目较小，且生成的叶子具有较大的泛化能力。因而用新的属性选择标准生成的决策树一般具有叶子数目较少，叶子的平均深度也较小，且叶子具有较强的和泛化能力。最后，用一实例说明了新的属性选择标准的优越性。

参考文献

[1] Quinlan J R. Induction of decision trees. Machine Learning, 1986,1:81-106

[2] Quinlan J R. Simplifying decision trees[J]. International Journal of Man-Machine Studies, 1987, 27(3):221-234

[3] 洪家荣，丁明峰，李星原，等. 一种新的决策树归纳学习算法. 计算机学报，1995,18(6): 470-474

[4] 刘小虎，李生. 决策树的优化算法. 软件学报，1998,9(10):797-800

[5] Pawlak Z. Rough sets. International Journal of Computer and Information Sciences, 1982,11:341-356

[6] 王国胤. Rough 集理论与知识获取. 西安：西安交通大学出版社，2001

[7] 张文修，吴伟志，等. 粗糙集理论与方法[M]. 北京：科学出版社，2003

[8] Wang Guoyin. Rough set theory and knowledge acquisition[M]. Xi An: Xi An JiaoTong University Press, 2001

[9] Zhang Wenxiu, WuWeizhi. Rough set theory and method[M]. Beijing: Science press, 2003

[10] 王名扬，卫金茂，伊卫国. 基于粗集理论的新决策树剪枝方法[J]. 东北师范大学(自然科学版)，2005, 37(3): 28-32

[11] 朱红. 基于 Rough Set 的最简决策树确定算法的研究[J]. 计算机工程与应用，2003,41(13):129-131

[12] 赵翔，向一丹，刘同明，祁云嵩. 一种基于粗糙集的决策树生成算法[J]. 华东船舶工业学院学报(自然科学版)，2005,19(4): 73-76

[13] 管红波，田大纲. 基于属性重要性的决策树规则提取算法[J]. 系统工程与电子技术，2004,26(3):334-337

[14] 常志玲，周庆敏，杨清莲. 基于粗糙集理论的决策树构造算法[J]. 南京工业大学学报，2005,27(4):80-83

[15] Jin MaoWei. Rough set based approach to selection of node [J]. International Journal of Computational Cognition, 2003,1(2):25-40

[16] Wang Mingyang, Jin Maowei, Yi Weiguo. A new decision tree pruning method based on RST [J]. Journal of Northeast Normal University (Natural Science Edition), 2005, 37(3): 28-32

[17] Zhu Hong. The Research of the simplest decision tree production algorithm based on rough set [J]. Journal of Computer Engineering and Application, 2003,41(13):129-131

[18] Zhao Xiang, Xiang Yidan, Liu Tongming, et al. An algorithm for decision tree construction based in rough sets [J]. Journal of East China Shipbuilding Institute (Natural Science Edition), 2005,19(4): 73-76

[19] Guan Hongbo, Tian Dagang. Rule abstracting algoritm by decisive tree based on the importance of attribute [J]. Systems Engineering and Electronics, 2004,26(3): 334-337

[20] Chang Zhiling, Zhou Qingmin, Yang Qinglian. A method based on rough set to construct decision tree [J]. Journal of Najing University of Technology, 2005,27(4):80-83

(上接第 107 页)

类型为 CBR，大小为 10b，大探测包为 10 组间隔很小的 CBR，大小为 2000b，每个 CBR 包被分为 10 个 200b 的 UDP 包。背景流量为 4 组指数达到数据流，分别从节点 S_1 到 D_1，D_2，节点 S_2 到 D_1，D_2。我们进行两组实验：第一组发送速度为 400kb/s，第二组发送速度为 800kb/s。第一组背景流量较小，通过 Ω 来判断拓扑结构；第二组背景流量较大，采用 Ψ 来判断拓扑结构。

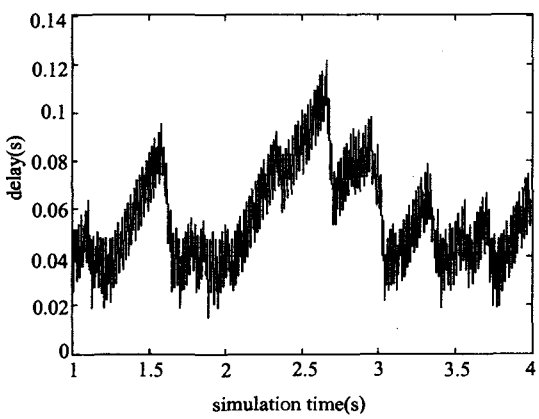

图 5 具有共享链路的端到端延时的相互影响示意图

图 5 是第一组实验中具有共享链路的端到端延时的相互影响示意图。从图中可以看出，两组小数据包的端到端延时在受另一节点发送的大数据包的影响的情况下，具有震荡现象。

实验结果如表 2 所示。在第一组实验中，若两条路径具有共享链路，延时变化大于 20ms，否则小于 2ms，相差大于 10 倍，可选取门限为 10ms。在第二组实验中，若两条路径具有共享链路，丢包率大于 17%，否则小于 3%，相差大于 5 倍，可选取门限为 10%。仿真推断的逻辑拓扑与真实拓扑的设置相符，说明了本文采用探测方法的有效性。

表 2 仿真结果表

拓扑	P_{S1D1} 与 P_{S2D2}				P_{S1D2} 与 P_{S2D1}			
结构	1		2		1		2	
	Ω	Ψ	Ω	Ψ	Ω	Ψ	Ω	Ψ
a	23.2ms	3.0%	35.1ms	22%	24.54ms	3.2%	26.9ms	27%
b	1.4ms	0.1%	5.4ms	2.7%	21.8ms	2.8%	25.6ms	17.6%
c	20.2ms	2.7%	27.2ms	18%	0.8ms	0%	2.9ms	1.7%
d	0.7ms	0.2%	1.4ms	2.3%	1.3ms	0.5%	2.7ms	3.0%

结束语 多源 NT 技术比单源拓扑推断可以获取更多的拓扑信息。本文通过分析四种 2-by-2 结构特征，发现四种 2-by-2 结构的拓扑推断问题可以转化为判断路径是否有共享链路的问题，而具有共享链路的路径的数据流对彼此的性能具有一定的影响。基于此点，提出一种新的三明治的探测方法，并详细研究了探测包的构造方式。通过仿真验证了探测方法的有效性。本文提出的方法对多源 NT 拓扑推断技术研究具有重要的意义。

参考文献

[1] Castro R, Coates M, Liang G, et al. Internet tomography: Recent developments. Statistical Science, 2004,52(3):499-517

[2] Coates M, Nowak R. Sequential monte carlo inference of internal delays in nonstationary communication networks. IEEE Transactions on Signal Processing, 2002,50(2): 366-376

[3] Coates M, Rabbat M, Nowak R. Merging logical topologies using end-to-end measurements// Proc. ACM SIGCOMM Conf. on Internet Measurement. Miami, FL, Oct. 2003

[4] Rabbat M, Nowak R, Coates M. Multiple source, multiple destination network tomography// Proc IEEE Infocom. Hong Kong, Mar. 2004

[5] Rabbat M, Coates M, Nowak R. Multiple Source Internet Tomography. IEEE Journal on Selected Areas in Communications, 2006,24(12):2221-2234