

# 基于 Manhattan 距离的网络拓扑推测方法研究<sup>\*</sup>

李勇军 蔡皖东 王 伟 田广利

(西北工业大学计算机学院 西安 710072)

**摘要** 网络拓扑推测是网络断层扫描研究内容之一,是推测网络内部链路性能的前提条件。目前的网络拓扑推测方法主要是基于网络性能参数的极大似然估计方法,随着网络规模的增加,计算量相对较大,还有些需要先验知识,从而影响在实际网络中的应用。为了克服这些问题,作者提出了一种新的网络拓扑推测方法,通过直接计算节点间测量数据的 Manhattan 距离,对节点进行分组,从而推测网络的逻辑拓扑结构。该方法计算简单且准确推测的概率收敛速度快,在实际网络环境中应用价值。

**关键词** 网络断层扫描,网络拓扑推测,Manhattan 距离

## Network Topology Identification Based on Manhattan Distance

LI Yong-Jun CAI Wan-Dong WANG Wei TIAN Guang-Li

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

**Abstract** Network Identification, one of the studies of the network tomography, is the proposition of the network link-level performance inference. The present methods rely either on the network performance or on the posterior distribution, and the time spent on the identification increase as the size of the network, which may restrict the technique to be used in practice. To overcome the above problems, we propose a fast approach to identify the logical network topology in this paper. Compared with the previous methods, the proposed one only needs to calculate the Manhattan distance based on the measurements to identify the network topology, which saves more time than the present ones, and the time spent on the identification do not increase sharply as the size of the network. From the simulation study, we find that the fraction of correctly classified trees fast converge, and accurately identify the trees even under the conditions that just hundreds of probe packets are injected. So the proposed method is very promising in the real network.

**Keywords** Network tomography, Network topology identification, Manhattan distance

近年来,国际上提出了一种称为网络断层扫描(Network Tomography,简称 NT)的测量技术,根据网络外部(即网络边界)的测量来分析和推断网络的内部性能<sup>[1]</sup>,解决了因网络安全和商业利益等原因,有些自治系统并不对外开放,难以获取其内部性能参数的问题。网络断层扫描的主要研究集中在网络内部性能参数的推断上面,这类研究有一个重要的前提条件就是要推测的网络拓扑结构是已知的,这个前提在很多情况下是不能够满足的。因此根据网络边缘的测量推测网络内部的逻辑拓扑结构显得极为重要。

现有网络拓扑推断方法的基本思想是根据在网络边缘的节点采集到的测量数据,推测端到端的性能参数,该性能参数应该是节点间共享链路数量的单调递增函数,以此性能参数作为拓扑推测的依据。例如延时相关性<sup>[1]</sup>,节点间共享的链路数越多,链路延时相关性就越大(假设非共享链路之间不相关)。常用的推测方法有极大似然方法、贝叶斯方法和分组方法。极大似然函数方法<sup>[2~4]</sup>,在已知端到端的报文丢失情况下,为所有可能的逻辑拓扑构建一个似然函数,找到一个拓扑结构使似然函数取得最大值。这种方法计算量会随着网络节点数量的增加而急剧增加。贝叶斯方法<sup>[2]</sup>把拓扑和链路概率看作是已知先验概率的随机变量,定义一个能够反映错误分类的函数,然后根据验后分布选择一个使得错误分类函数较小的网络拓扑,这与似然估计有点类似,从所有可能的拓扑中

选择一个最优。分组方法<sup>[2]</sup>首先定义一个基于端到端的链路报文丢失率的距离函数,选择距离最近的节点作为兄弟节点,该方法相对于前两种方法计算相对简单,也不需要先验知识。

本文提出了一种简单快速的网络拓扑推测方法,与前三种方法不同之处在于采用 Manhattan 距离作为函数评估节点间的关系,距离函数直接建立在观测到的测量数据上面,不需要计算额外的参数(如分组方法<sup>[2]</sup>中的报文丢失率)。推测方法基于在多播的情况下,探测报文在共享链路上的丢失情况是相同的这一事实,共享链路越多 Manhattan 距离越小。选择节点间 Manhattan 距离最小的作为兄弟节点。本文首先给出了报文丢失模型,在此基础上详细描述了网络拓扑推测算法,然后用仿真的方法证明了该算法的可行性。

## 1 报文丢失模型

### 1.1 树模型

二叉树在描述逻辑网络拓扑中具有普遍适用性,本文采用二叉树  $T=(V, L)$  表示逻辑网络拓扑,其中  $V$  代表拓扑中的逻辑节点,  $L$  表示逻辑连接,一条逻辑连接可能包含一条或者几条物理连接。一个典型的树模型如图 1 所示。假设探测报文从根节点被多播到叶节点,根节点用  $0 \in V$  表示,叶节点用  $R \subset V$  表示。 $\forall k \in V \setminus R \cup \{0\}$ ,称  $k$  为内部节点。任意一个内部节点  $k$  的孩子用  $c(k)$  表示,任何一个非根节点  $k$  的父

<sup>\*</sup> 基金项目:航天科技创新基金。李勇军 博士生,主要从事网络安全、网络性能测量的研究。

节点用  $p(k)$ , 链路  $(p(k), k) \in L$  用  $k$  表示。

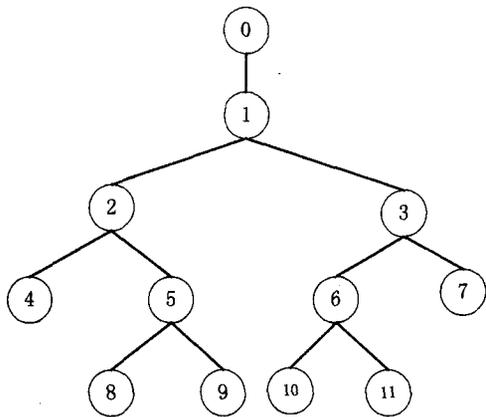


图1 网络拓扑

### 1.2 报文丢失模型

假设树  $T$  中链路  $k$  的报文丢失符合独立的 Bernoulli 模型。用随机过程  $X = (X_k)_{k \in V}$  描述节点接收探测报文的情况, 如果报文到达了节点  $k$ , 则  $X_k = 1$ ; 否则  $X_k = 0$ 。根据树的特性有以下关系:

$$- \text{If } X_k = 0, X_j = 0, \forall j \in c(k) \quad (1)$$

$$- \text{If } X_j = 1, \forall j \in c(k), X_k = 1 \quad (2)$$

对  $n$  个探测报文来说, 每个节点  $k$  就维护了一个大小为  $n$  的 0-1 序列, 用  $\{X_k^{(n)}\}, k \in V$  表示, 其中 0 表示对应的探测报文没有到达该节点, 1 表示到达了该节点。  $n(X_i = 1)$  表示节点  $i$  接收到的探测报文数量。

## 2 推测方法

网络断层扫描测量方法是在网络边缘进行的, 通过端端的测量来推测网络拓扑结构, 不需要网络内部设备的协作。探测报文从节点  $0 \in V$ , 自顶向下被多播到  $T$  中的每个节点, 只有在叶节点才可以观测到报文丢失情况。为了推测网络逻辑拓扑结构, 需要先定义节点间的距离函数和推测网络内部节点的报文丢失情况。

### 2.1 Manhattan 距离函数

定义 1 设  $X_i, X_j$  分别为节点  $i$  和节点  $j$  的报文丢失情况,  $X_i = \{X_i^{(n)}\}, i \in V$  和  $X_j = \{X_j^{(n)}\}, j \in V, |X_i| = |X_j| = N$ 。由 Manhattan 距离公式<sup>[5]</sup>, 节点  $i$  和节点  $j$  的 Manhattan 距离均值 (Average Manhattan Distance, AMD) 为

$$D_{ij} = \frac{1}{N} \sum_{k=1}^N |X_i^{(k)} - X_j^{(k)}| \quad (3)$$

由定义 1 可知当且仅当  $\forall k \in \{1, 2, \dots, N\}, X_i^{(k)} = X_j^{(k)}$  也就是  $X_i$  和  $X_j$  的不匹配数为 0 时,  $D_{ij} = 0$ 。在估算节点间的 AMD 时, 不失一般性采用如图 2 所示的报文传输模型, 节点  $p(l)$  有两个子节点  $l$  和  $P(i)$ , 节点  $p(i)$  有两个子节点  $i$  和  $j$ 。

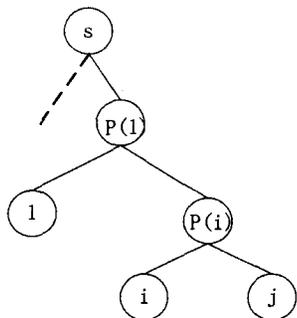


图2 报文传输模型

假设丢失探测报文  $k$  是在节点 0 到节点  $p(i)$  之间的某条链路上发生的, 由性质(1)可知  $X_i^{(k)} = X_j^{(k)}$ , 也就是说丢失在链路  $p(i)$  或其上游链路上的报文对计算  $D_{ij}$  不产生影响,  $D_{ij}$  仅仅反映了探测报文在链路  $i$  和  $j$  的丢失差异。另外随着探测报文数量的增加, 报文在链路上的丢失数量会逐渐收敛于发送报文的数量乘以链路丢失率。因此任意两个节点之间的 AMD 也会随着探测报文数量的增加, 而逐渐收敛。另一方面  $\lim_{n \rightarrow \infty} D_{p(i)} > \lim_{n \rightarrow \infty} D_k$ 。本文正是利用这一性质对网络节点进行分组, 从而实现网络逻辑拓扑的推测。

### 2.2 内部节点报文丢失情况推测

计算任何两个节点的 AMD 时, 需要两节点的报文接收情况是已知的。因此推测网络内部拓扑结构时, 需要推测网络内部节点接收到探测报文的情况。

假设节点  $i$  和  $j$  的链路丢失情况是已知的, 分别为  $\{X_i^{(n)}\}$  和  $\{X_j^{(n)}\}$ , 如图 2 所示。现在要推测的是节点  $p(i)$  的报文丢失情况  $\{X_{p(i)}^{(n)}\}$ 。根据节点  $p(i)$  与其子节点之关系可知, 节点  $p(i)$  收到的报文一部分成功到达了节点  $i$  或  $j$ , 还有一部分同时丢失在链路  $i$  和  $j$ 。对于前者来说可以从  $\{X_i^{(n)}\}$  和  $\{X_j^{(n)}\}$  得到, 对于后者是不可推测或不可观测到的。因为节点  $i$  或  $j$  没有收到报文, 一种情况是丢失在链路  $i$  和  $j$  上, 还有一种情况是丢失在链路  $p(i)$  或者其上游链路上, 这两种情况是难以区分的。用  $\Delta_{ij}$  表示同时丢失在链路  $i$  和  $j$  上的报文数量, 则节点  $p(i)$  接收到的报文数量为:

$$n(X_{p(i)} = 1) = n(X_i = 1) + n(X_j = 1) - n(X_i = 1 \wedge X_j = 1) + \Delta_{ij}$$

$\Delta_{ij}$  相对于其他部分较小, 因此有,

$$\lim_{n \rightarrow \infty} n(X_{p(i)} = 1) = n(X_i = 1) + n(X_j = 1) - n(X_i = 1 \wedge X_j = 1)$$

节点  $p(i)$  的报文丢失情况  $X_{p(i)}$  可以用公式(4)进行估测,

$$X_{p(i)}^{(n)} = X_i^{(n)} \vee X_j^{(n)} \quad (4)$$

### 2.3 推测算法

基于二叉树的网络逻辑拓扑方法首先从节点集合中, 选择两个 AMD 最小的节点作为兄弟节点, 构建父节点, 并推测父节点的报文丢失情况, 在节点集合中用父节点代替兄弟节点, 重复上述过程直至节点集的大小为 1。下面的算法详细描述了网络逻辑拓扑结构的推测方法。

1. 输入: 叶节点集合  $R$ , 叶节点接收报文情况  $(X_k^{(i)})_{i \in k} \dots^n$ 。
2.  $R' = R, L' = \Phi, D' = \Phi$
3. for  $i \in R'$  do  
 for  $j \in R'$  do  
 if  $D_{ij} \notin D'$  then  
 计算  $D_{ij}$ ,  
 $D' \leftarrow D' \cup \{D_{ij}\}$
4. while  $|R'| > 1$  do  
 选择  $D'$  中的最小值, 其对应的两个节点分别为  $u_1, u_2 \in R'$ ,  
 $u = p(u_1)$   
 $X_u^{(n)} = X_{u_1}^{(n)} \vee X_{u_2}^{(n)}, n = 1, \dots, N$   
 $R' \leftarrow R' \setminus \{u_1, u_2\}$   
 $R' \leftarrow R' \cup \{u\}$   
 $L' \leftarrow L' \cup \{(u, u_1), (u, u_2)\}$   
 $D' \leftarrow D' \setminus \{D_{u_1 u_2}\}$   
 for  $i \in R'$  do  
 if  $D_{iu} \notin D'$  then  
 计算  $D_{iu}$ ,  
 $D' \leftarrow D' \cup \{D_{iu}\}$
5.  $L' \leftarrow L' \cup \{(0, u)\}$
6. 输出  $L'$

## 3 仿真分析

为了验证算法的有效性, 用 ns2<sup>[6]</sup> 进行了仿真实验。用 12 个节点的二叉树表示网络拓扑, 如图 1 所示。图中每条链路的带宽为 1M、传输延时 10ms、采用 DropTail 策略丢弃报

文、队列大小为 10。另外为了在仿真过程观测到报文丢失,仿真过程引入了 error model,设定链路 1、2、3、5 和 6 的报文丢失概率在 [0.05, 0.1] 区间内随机产生,其余链路的报文丢失的概率在区间 [0.01, 0.06] 区间内产生。探测报文由 Exponential 产生,报文大小为 200, idle time 为 200ms, burst time 为 300ms, 产生速率为 32k。报文从节点 0, 自顶向下被多播到网络中每个节点,仿真时间长度从 2s 到 100s 不等,仿真时间越长意味着发送的探测报文越多。其对应关系如表 1 所示。

表 1 仿真时间与报文数量关系

时间/(s)	5	20	50	80	100
报文数量	86	261	560	1000	1243

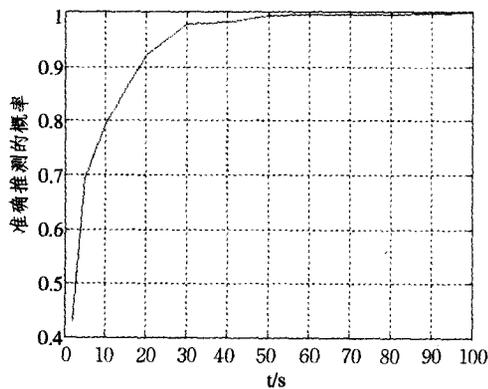


图 3 正确推测次数与仿真时间的关系

每个仿真时间重复 1000 次,统计其中正确推测网络逻辑

拓扑的次数。图 3 中显示了准确推测出如图 1 所示逻辑拓扑比率随时间的变化关系。从中不难发现随着仿真时间的增加,或者说随着发送探测报文数量的增加,能够正确推测网络拓扑的次数也在增加,正确推测的比率逐渐收敛于 1。在仿真时间为 10s 时,发送的报文数量大约 130 左右,正确推测网络的次数达到 792 次,正确推测的比率接近 0.8,因此可以通过发送较少的报文推测出比较准确的网络逻辑拓扑结构。

**结束语** 基于 Manhattan 距离的网络拓扑推测方法是通过端到端的测量,推测网络的逻辑拓扑结构,不需要网络内部设备的协作。相对于现有推测方法,计算量较小,并且不会随着网络规模的增加,计算量而急剧增加。通过仿真验证了该方法准确推断率的收敛速度较快,即使在发送较少报文(150 个探测报文)的情况下,也可以得到较高的准确推断率(接近 0.8)。为了使推测的拓扑更加接近物理拓扑结构,需要扩展该算法,进一步研究基于一般树型结构的网络拓扑推测方法。

参考文献

- 1 Mark C, Hero III A O, Robert N, Yu Bin. Internet Tomography. IEEE Signal Processing Magazine, 2002, 19(3): 47~65
- 2 Duffield N G, Horowitz J, Lo Presti F, Towsley D. Multicast Topology Inference from Measured End-to-end Loss. IEEE Transactions on Information Theory, 2000, 48(1): 26~45
- 3 Castro R, Coates M, Gadhok M. Maximum Likelihood Network Topology Identification from Edge-based Unicast Measurements; [Rice University, Tech Rep; TREE-0107]. 2001
- 4 Castro R, Coates M. Maximum Likelihood Identification of Network Topology from End-to-end Measurement; [Rice University. Tech Rep; TREE-0109]. 2002
- 5 Han Jiawei, Kamber M 著. 范明, 孟小峰, 等译. 数据挖掘-概念与技术. 北京: 机械工业出版社, 2001
- 6 Information Sciences Institute of University of Southern California. The Network Simulator 2. www.isi.edu/nsnam/ns2, 2005

(上接第 13 页)

吐量也在增加。这种增加是由于多用户分集所带来的效果,由于各个用户经历独立的衰落过程,通过对多个用户按照行列式准则(21)式进行选择,能最大程度地利用多用户信道状态的“随机性”,用户数越多、信道变化越随机,每个时刻就越有可能选择到一个接近其信道“峰值”状况的用户进行传输,所以随着用户数的增加,系统吞吐量也在增加。另一方面,随着用户数的增加,系统吞吐量越来越接近“峰值”,多用户分集所带来的增益也越来越小。例如在 8dB 处,4 个用户相对单

用户情况,系统吞吐量的增加约为 1260 bits/OFDM symbol, 8 用户相对 4 用户约为 753 bits/OFDM symbol; 而 16 用户相对 8 用户增加只有大约 273 bits/OFDM symbol。

**结论** 本文提出了一种基于多用户 MIMO-OFDM 系统的自适应跨层空间子信道分配算法,采用链路层的 T-ARQ 技术,以获取最大的系统吞吐量为目标,充分利用所有可用的空间子信道,在满足恒定的总发射功率和一定的 QoS(误包率和延迟)的条件下,推导了子载波分配准则,并给出了相应的自适应算法流程。仿真结果表明,上述自适应算法在提高系统吞吐量方面具有明显的优势,可以很好地应用于未来的宽带无线通信。

参考文献

- 1 Cimini JR L J. Analysis and simulation of digital mobile channel using orthogonal frequency division multiplexing [J]. IEEE Transactions on Communications, 1985, 33(7): 665~675
- 2 Foschini G J, Gans M J. On limits of wireless communication in a fading environment when using multiple antennas [J]. Wireless Personal Communications, 1998, 6(3): 311~335
- 3 Raleigh G G, Cioffi J M. Spatio-temporal coding for wireless communication [J]. IEEE Transactions on Communications, 1998, 46(3): 357~366
- 4 Wong K K, Cheng R S K, Letaief K B, et al. Adaptive antennas at the mobile and base stations in an OFDM/TDMA system [J]. IEEE Transactions on Communications, 2001, 49(6): 195~206
- 5 Li J, Ltaief K B, Ma Z, et al. Spatial multiuser access with MIMO smart antennas for OFDM systems [A]. In: IEEE VTC [C], Atlantic, Oct 2001. 1553~1557
- 6 Keller H L. Adaptive multicarrier modulation: A convenient framework for time frequency processing in wireless communications [J]. Proc IEEE, 2000, 88(5): 611~640
- 7 Hughes-Hartogs D. Ensemble Modem Structure for Imperfect Transmission Media [EB/OL]. U. S. Patents 4,679,227 (Jul. 7, 1987); 4,731,816 (Mar. 15, 1988); 4,833,706 (May 23, 1989)

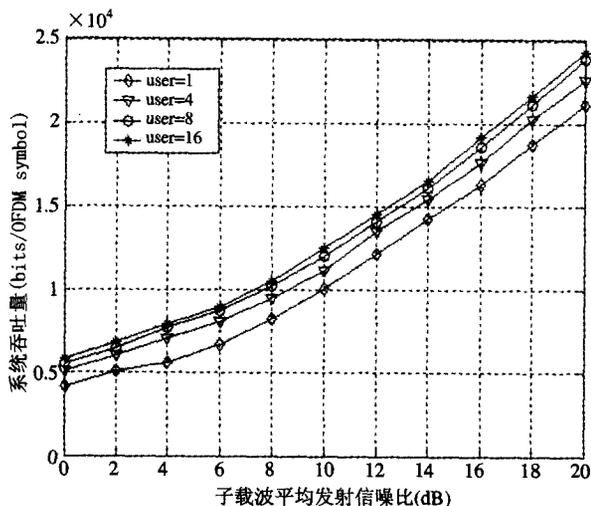


图 5 不同用户数下的系统吞吐量