

一种新的物理网络拓扑发现算法

陈松^① 王珊^② 周明天^①

^①(电子科技大学计算机学院 成都 610054)

^②(电子科技大学物理电子学院 成都 610054)

摘要: 该文提出了一种新的物理网络拓扑发现算法。该算法的基本思想是采用树的形式表示交换机拓扑连接关系。在此理论基础上利用生成树集合演绎树的变化过程来完成拓扑发现,并结合一种典型的交换机连接关系进行了算法推导。理论分析和实际应用结果表明,这种新的拓扑发现算法在发现效率、准确性、有效性等方面都有了很大程度的提高。

关键词: 网络测量; 拓扑发现; 地址转发表; 生成树

中图分类号: TP393.06

文献标识码: A

文章编号: 1009-5896(2010)01-0172-06

DOI: 10.3724/SP.J.1146.2008.01783

A New Physical Network Topology Discovery Algorithm

Chen Song^① Wang Shan^② Zhou Ming-tian^①

^①(College of Computer, University of Electronic Science and Technology of China, Chengdu 610054, China)

^②(College of Physical Electronics, University of Electronic Science and Technology of China, Chengdu 610054, China)

Abstract: This paper presents a new physical network topology discovery algorithm. The tree form is applied to describe the connections between switches. Based on this theory, the topology discovery is completed by using spanning tree set to deduce tree change, and the algorithm referring to a typical connection relationship between switches is derived. Theoretical analysis and practical application indicate that the proposed topology discovery algorithm is advanced in efficiency, accuracy and effectiveness.

Key words: Network measurement; Topology discovery; Address forwarding table; Spanning tree

1 引言

网络拓扑结构发现^[1-3]是网络测量技术的一个重要基础,是进行网络各项性能指标计算或推断,评价网络性能与行为,建立和理解网络行为模型的重要前提条件。物理拓扑发现算法是自动发现交换机之间的连接关系,主要分为两类:一类是基于传统 SNMP 和交换技术的拓扑发现算法^[4,5],这类算法发现准确性高,实际应用最为广泛,但由于受到安全权限的限制,发现范围存在一定的局限性。另外一类是随着 BGP 路由表数据的开放,许多研究者开始对自治系统(Autonomous System, AS)级 Internet 网络拓扑^[6,7]的宏观度量进行统计分析,目前研究工作的局限性主要是:受限于统计数据数量较少或时间跨度较短,不能较为全面地对网络拓扑做多角度分析,尤其是随时间变化的演化分析。

贝尔实验室的 Breitbart 等人提出了基于交换机

地址转发表 AFT(Address Forwarding Table)的物理网络拓扑发现方法^[4,5]。其算法的核心是判断交换机端口直连定理:当且仅当同一子网内两个端口的地址转发条目集合的交集为空,且其并集中包含了该子网中所有交换机的地址条目,则分属两个交换机的一对端口是相连的。为保证地址转发表的完整性, Breitbart^[5]专门设计出一种 mping 程序,通过对原有的 ping 程序进行修改,在给定的目标交换机地址转发表中增加了目的设备的 MAC 地址,从而保证地址转发表的完整性。郑海等^[8]给出了一组判定定理,由上行端口与下行端口判断交换机之间的连接关系;孙延涛等^[9,10]提出了由网络节点的直系关系和旁系关系判断交换机之间的连接关系,这些算法能够利用不完整的地址转发表构造出网络拓扑图。

本文提出了一种新的物理网络拓扑发现算法,判断交换机之间的连接关系,与现有算法相比:(1)提出了一种新的树的演绎方法,用树的演绎方法可以遍历所有的顶点和边,完成拓扑发现,这种推导过程与现有算法相比,时间复杂度大大降低,推导方法也更为容易理解;(2)提出了一套新的判断理论,能够判断边缘节点、边缘连线和进行树的演绎过程

2008-12-25 收到, 2009-03-09 改回

国家自然科学基金(60572092)和国家“十一五”科技支撑计划项目(2006BAH02A0407)资助课题

通信作者: 陈松 chensong_mail@163.com

推导；(3)文中的算法支持将交换机物理拓扑关系表述成多棵树的形式，支持存在缺失连接关系的拓扑结构发现；(4)提出了交换机拓扑连接关系发现不仅需要地址转发表，也需要生成树。因为地址转发表获取的仅仅是连接的对方端口的 MAC 地址信息和本端虚拟端口编号，虚拟端口映射到本地实际端口必须通过生成树 (BRIDGE-MIB: dot1dBase PortTable: 1.3.6.1.2.1.17.1.4) 的映射关系才能够获取。

2 物理拓扑发现算法

2.1 算法基础

文中在给出了连接关系的定义方法和 MAC 地址集合定义方法，在此基础上进行推导。

定义 1 对于同一子网：若第 i 台是交换机表示为 S_i ，交换机 S_i 第 j 个端口表示为 S_{ij} ，若第 i 台是主机则表示为 H_i 。

定义 2 对于交换机任意端口 S_{ij} ， P_{ij} 表示交换机 S_i 的地址转发表中通过端口 S_{ij} 学习到的其余交换机端口和主机的集合，称 P_{ij} 为交换机可达路径集合。

定义 3 对于交换机任意端口 S_{ij} ， $L < S_{ij}, S_{kl} >$ 表示交换机 S_i 通过端口 S_{ij} 与交换机 S_k 的 S_{kl} 端口连接， $L < S_{ij}, H_k >$ 表示交换机 S_i 通过端口 S_{ij} 与主机 H_k 相连。

定义 4 对于交换机任意端口 S_{ij} ，若学习到的端口 MAC 地址中有且只有另外一台交换机的端口 S_{kl} 的 MAC 地址，则交换机 S_k 称为边缘交换机，对应的交换机连线 $L < S_{ij}, S_{kl} >$ 称为边缘连线。

定义 5 集合 V 表示所有交换机端口可达路径 P_{ij} 的集合，即 $\{P_{11}, P_{12}, \dots, P_{1m}, P_{21}, P_{22}, \dots, P_{2n}, P_{i1}, P_{i2}, \dots, P_{ik}, \dots\}$ 。

定义 6 集合 E 表示交换机所有连接关系的集合。

2.2 算法框架

连通不含回路的无向图称为树，用 $G = \langle V, E \rangle$ 表示。拓扑图可以用一棵或者几棵树表示。用 n 表示 G 中的顶点数， m 为 G 中的边数，则 $n = m + 1$ 。由树的特性可以得知：度数为 1 的顶点称为树叶，度数大于 1 的顶点称为分支点。如果一个分支点有且只有一条边连接其它分支点，其余边连接的顶点均为树叶，则称该分支点为可演化分支点。

本文定义的树的演化过程如图 1 所示：对无向图 G_1 删除所有的树叶以及与树叶相连的边，得到无向图 G_2 ，结论如下：

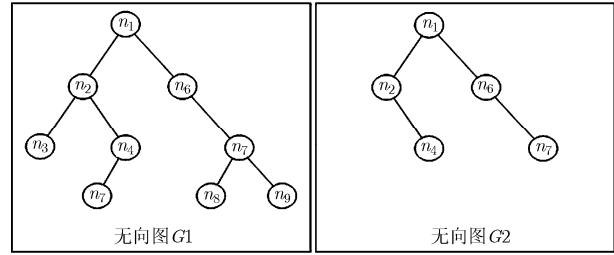


图 1 树的演化过程

(1)所有可演化分支点均变成树叶

证明 由于可演化分支点有且只有一条边连向其它分支点，其余边均与树叶相连。因此，在去掉与树叶相连的边之后，可演化分支点的顶点度数也变为 1，演化成树叶。

(2)得到的子图仍然是树

证明 (a)连通性：假设删除所有树叶 n_i 以及与树叶相连的边 m_j ，子图变成了非连通的。则 n_i 必定与其它节点相连，也 n_i 本身是树叶相矛盾。因此，可以得出得到的子图是连通的；(b)不存在回路：由于在删除过程中并没有增加新的边，则不会构成新的回路。

由此可以得知：得到的子图是连通图且不存在回路，仍然是一棵树。

(3)依次递归调用，可以不重复地遍历所有的顶点和所有的边，直至发现顶点退出。

2.3 算法步骤

(1)遍历子网内所有交换机端口的地址转发表，将交换机学习到的 MAC 地址集合根据桥接 MIB 转换为交换机端口的集合 P_{ij} ；

(2)将所有得到的 P_{ij} ，构造生成树集合 V ；

(3)发现交换机与主机的连接关系：取出交换机只包含主机 MAC 地址的集合 $P_{ij}\{H_k\}$ 或者 $P_{ij}\{H_k, \dots, H_l\}$ ，如果是 $P_{ij}\{H_k\}$ ，则交换机端口 S_{ij} 与主机 H_k 相连，删除集合 P_{ij} ，并将 H_k 在其余的生成树集合中去掉，添加连接关系 $\langle S_{ij}, H_k \rangle$ ；如果是 $P_{ij}\{H_k, \dots, H_l\}$ ，则主机 H_k, \dots, H_l 通过 HUB 与交换机相连，删除集合 P_{ij} ，并将主机 H_k, \dots, H_l 在其余的生成树集合中去掉，添加连接关系 $\langle S_{ij}, H_k \rangle \dots \langle S_{ij}, H_l \rangle$ ；

(4)发现边缘交换机、边缘连线：

(a)取交换机一元关系集合 $P_{ij}\{S_{kl}\}$ ，由定理得出交换机 S_i 通过端口 S_{ij} 与交换机 S_k 的 S_{kl} 端口是直接相连的， S_k 为边缘交换机， $\langle S_{ij}, S_{kl} \rangle$ 是边缘连线；

(b)删除一元集合 P_{ij} ；删除边缘交换机 S_k 所有端口在生成树集合中 V 中的元素；删除边缘交换机 S_k 所有端口的在生成树中所有的集合；

(c)添加交换机连接关系 $\langle S_{ij}, S_{kl} \rangle$ ；

(d)循环处理下一个交换机一元集合,直至该轮一元集合处理完毕。

(5)按照步骤(4)继续处理生成树集合,直到生成树集合为空,则发现算法程序退出。

2.4 算法推导分析

本文给出了一种典型的拓扑连接关系进行拓扑发现算法推导,给定的待发现网络中有 7 台交换机, 3 台主机, 其中 2 台主机通过 HUB 与交换机相连, 1 台主机直接与交换机相连, 拓扑连接关系图如图 2 所示。

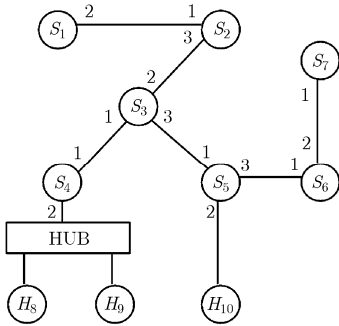


图 2 拓扑连接关系图

步骤 1 产生生成树集合

由地址转发表得出的生成树集合 V 如下:

- $P_{12}\{S_{21}, S_{32}, S_{41}, S_{51}, S_{61}, S_{71}, H_8, H_9, H_{10}\}$
- $P_{21}\{S_{12}\}$
- $P_{23}\{S_{32}, S_{41}, S_{51}, S_{61}, S_{71}, H_8, H_9, H_{10}\}$
- $P_{31}\{S_{41}\}$
- $P_{32}\{S_{23}, S_{12}\}$
- $P_{33}\{S_{51}, S_{61}, S_{71}, H_{10}\}$
- $P_{41}\{S_{31}, S_{23}, S_{12}, S_{51}, S_{61}, S_{71}, H_{10}\}$
- $P_{42}\{H_8, H_9\}$
- $P_{51}\{S_{33}, S_{23}, S_{12}, S_{41}, H_8, H_9\}$
- $P_{52}\{H_{10}\}$
- $P_{53}\{S_{61}, S_{71}\}$
- $P_{61}\{S_{53}, S_{33}, S_{23}, S_{12}, S_{41}, H_8, H_9, H_{10}\}$
- $P_{62}\{S_{71}\}$
- $P_{71}\{S_{62}, S_{53}, S_{33}, S_{23}, S_{12}, S_{41}, H_8, H_9, H_{10}\}$

步骤 2 发现交换机与主机的连接关系

取交换机只包含主机 MAC 地址的集合 $P_{42}\{H_8, H_9\}$, $P_{52}\{H_{10}\}$, 删除集合 P_{42} , P_{52} , 并将 H_8, H_9, H_{10} 从其余集合中去掉, 添加连接关系 $\langle S_{42}, H_8 \rangle$, $\langle S_{42}, H_9 \rangle$, $\langle S_{52}, H_{10} \rangle$ 。

该轮搜索结束后, 搜索得到的连接关系结合 E

为: $\langle S_{42}, H_8 \rangle$, $\langle S_{42}, H_9 \rangle$, $\langle S_{52}, H_{10} \rangle$ 。
该轮搜索结束后, 得到的拓扑结构如图 3。

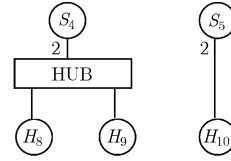


图 3 步骤 2 得到的拓扑结构图

该轮搜索结束后, 待处理的生成树集合如下:

- $P_{12}\{S_{21}, S_{32}, S_{41}, S_{51}, S_{61}, S_{71}\}$
- $P_{21}\{S_{12}\}$
- $P_{23}\{S_{32}, S_{41}, S_{51}, S_{61}, S_{71}\}$
- $P_{31}\{S_{41}\}$
- $P_{32}\{S_{23}, S_{12}\}$
- $P_{33}\{S_{51}, S_{61}, S_{71}\}$
- $P_{41}\{S_{31}, S_{23}, S_{12}, S_{51}, S_{61}, S_{71}\}$
- $P_{51}\{S_{33}, S_{23}, S_{12}, S_{41}\}$
- $P_{53}\{S_{61}, S_{71}\}$
- $P_{61}\{S_{53}, S_{33}, S_{23}, S_{12}, S_{41}\}$
- $P_{62}\{S_{71}\}$
- $P_{71}\{S_{62}, S_{53}, S_{33}, S_{23}, S_{12}, S_{41}\}$

步骤 3 第 1 轮发现边缘交换机, 边缘连线

取交换机一元关系集合 $P_{21}\{S_{12}\}$, $P_{31}\{S_{41}\}$, $P_{62}\{S_{71}\}$, 删除集合 P_{21} , P_{31} , P_{62} , 并将 S_1, S_4, S_6 的所有端口及所有端口的集合从生成树集合中去掉, 添加连接关系 $\langle S_{21}, S_{12} \rangle$, $\langle S_{31}, S_{41} \rangle$, $\langle S_{62}, S_{71} \rangle$ 。

该轮搜索结束后, 搜索得到的连接关系结合 E 为: $\langle S_{42}, H_8 \rangle$, $\langle S_{42}, H_9 \rangle$, $\langle S_{52}, H_{10} \rangle$, $\langle S_{21}, S_{12} \rangle$, $\langle S_{31}, S_{41} \rangle$, $\langle S_{62}, S_{71} \rangle$ 。

该轮搜索结束后, 得到的拓扑结构如图 4。

该轮搜索结束后, 待处理的生成树集合如下:

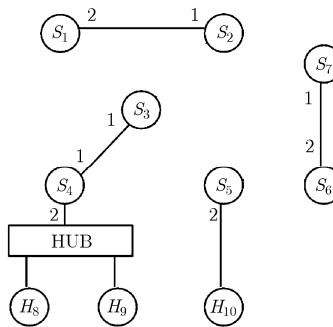


图 4 步骤 3 得到的拓扑结构图

- $P_{23}\{S_{32}, S_{51}, S_{61}\}$
- $P_{32}\{S_{23}\}$
- $P_{33}\{S_{51}, S_{61}\}$
- $P_{51}\{S_{33}, S_{23}\}$
- $P_{53}\{S_{61}\}$
- $P_{61}\{S_{53}, S_{33}, S_{23}\}$

步骤 4 第 2 轮发现边缘交换机，边缘连线

取交换机一元关系集合 $P_{32}\{S_{23}\}$ ， $P_{53}\{S_{61}\}$ ，删除集合 P_{32} ， P_{53} ，并将 S_2 ， S_6 的所有端口和所有端口的集合从生成树集合中去掉，添加连接关系 $\langle S_{32}, S_{23} \rangle$ ， $\langle S_{53}, S_{61} \rangle$ 。

该轮搜索结束后，搜索得到的连接关系结合 E 为： $\langle S_4, H_8 \rangle$ ， $\langle S_4, H_9 \rangle$ ， $\langle S_5, H_{10} \rangle$ ， $\langle S_{21}, S_{12} \rangle$ ， $\langle S_{31}, S_{41} \rangle$ ， $\langle S_{62}, S_{71} \rangle$ ， $\langle S_{32}, S_{23} \rangle$ ， $\langle S_{53}, S_{61} \rangle$ 。

该轮搜索结束后，得到的拓扑结构如图 5。

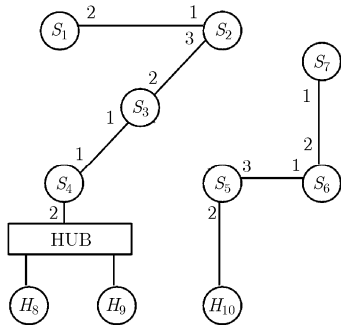


图 5 步骤 4 得到的拓扑结构图

该轮搜索结束后，待处理的生成树集合如下：

- $P_{33}\{S_{51}\}$
- $P_{51}\{S_{33}\}$

步骤 5 第 3 轮发现边缘交换机，边缘连线

取交换机一元关系集合 $P_{33}\{S_{51}\}$ ， $P_{51}\{S_{33}\}$ ，删除集合 P_{33} ， P_{51} ，并将 S_5 ， S_3 的所有端口和所有端口的集合从生成树集合中去掉，添加连接关系 $\langle S_{33}, S_{51} \rangle$ ，由于连线 $\langle S_{51}, S_{33} \rangle$ 与 $\langle S_{33}, S_{51} \rangle$ 是同一连线，就不必再添加。

该轮搜索结束后，搜索得到的连接关系结合 E 为： $\langle S_4, H_8 \rangle$ ， $\langle S_4, H_9 \rangle$ ， $\langle S_5, H_{10} \rangle$ ， $\langle S_{21}, S_{12} \rangle$ ， $\langle S_{31}, S_{41} \rangle$ ， $\langle S_{62}, S_{71} \rangle$ ， $\langle S_{32}, S_{23} \rangle$ ， $\langle S_{53}, S_{61} \rangle$ 。

该轮搜索结束后，得到的拓扑结构如图 6。

该轮搜索结束后，待处理的生成树集合为空，发现算法程序退出。

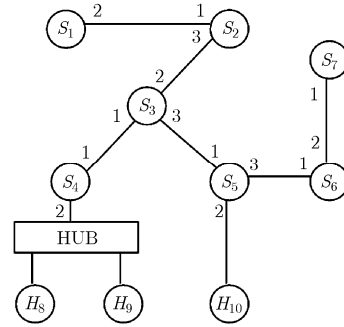


图 6 步骤 5 得到的拓扑结构图

3 算法分析

从前面的推导可以得出，算法发现的连接关系结果同实际结果是一致的。另外，算法在发现效率和有效性方面同 Breitbart 等^[4]提出的算法相比也有了很大程度提高。

3.1 发现效率分析

Breitbart 提出的交换机连线判断算法^[5]，判断每对端口的连线情况时，每个端口的集合将与其它所有端口集合作比较，算法的时间复杂度为 $O(n^2)$ ；本文采用了一种递减的拓扑推导方法，需要的推导步骤为拓扑图中存在的最长路径的 $1/2$ ，算法的时间复杂度为 $O(n)$ 。因此，本文的算法在大型网络发现中，效率上将会有数量级提高。

3.2 算法实际应用分析

本文设计了一套拓扑管理系统，对网络层拓扑发现采用 IETF RFC-1213 MIB 获取路由表的方式，对物理拓扑的发现采用了本文提出的拓扑发现算法。整个拓扑系统框架分为三层(如图 7 所示)：平台支撑层、数据处理层、数据展现层。

(1)平台支撑层：提供一个大型分布式系统所必需的各种底层支撑功能。这些功能不仅可以应用于网管系统，理论上也可以应用于各类的大型分布式管理系统。平台支撑层的功能主要包括 4 个部分：J2EE 应用服务器；J2EE 应用服务器的扩展功能；标准网管协议栈(SNMP、ICMP 等)；图形展现工具

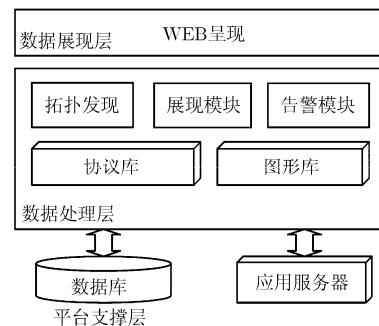


图 7 拓扑管理系统功能架构

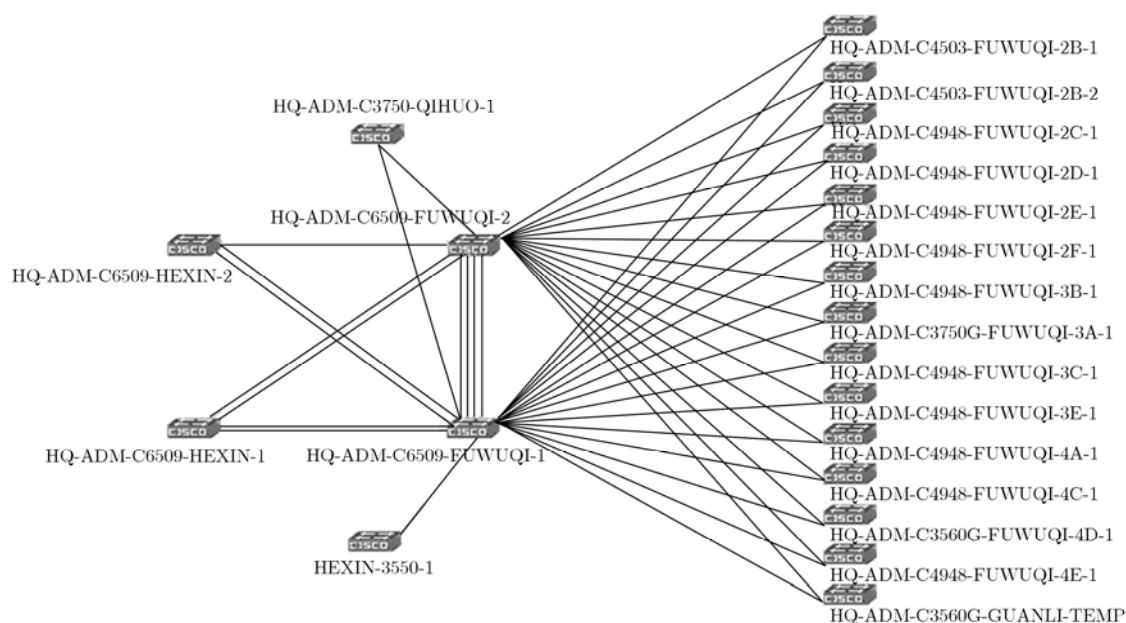


图 8 物理拓扑发现结果

包。

(2)数据处理层：拓扑发现模块调用协议库中的 SNMP 协议库进行拓扑发现，并将发现结果保存到数据库中；展现模块调用图形库的元素，将拓扑数据以最直观的方式呈现给客户，而且能够进行拓扑元素的布局；告警模块处理 Trap 和 Syslog 事件，并通知展现模块将有告警的节点进行变色。

(3)数据展现层：将拓扑数据库中保存的数据转换成拓扑图片，并生成 web 页面，可以方便用户通过浏览器访问。

贝尔实验室 Breitbart 提出的拓扑发现算法以及在此基础上延续的研究，其基本技术是建立在较老的网络拓扑结构基础上。近年来网络技术和网络规模发展都很快，如防火墙、无线局域网等。本文将提出的算法在大规模实际网络中进行了验证，针对某知名金融机构全国运营网络，在北京全国网管中心对本文拓扑系统进行了部署，发现中心和各省级城市的设备连接关系，并对故障节点进行告警。结果证明，本文提出的拓扑发现算法能够有效发现子网内交换机的拓扑连接关系，如图 8 所示。

4 结束语

网络拓扑发现是人们认识、了解、研究网络行为的重要基础，交换机连接关系的判断是一个技术难点。本文研究了物理网络交换机连接关系发现算法，提出了一种新的物理网络拓扑发现算法，判断交换机之间的连接关系。与现有算法相比，首次提出了一种新的树的演绎方法，用树的演绎方法可以

遍历所有的顶点和边完成拓扑发现，这种推导过程与现有算法相比，时间复杂度大大降低，推导方法也更为容易理解。最后在实际环境中对自动发现算法的有效性进行了验证。

参考文献

- [1] Donnet B and Friedman T. Internet topology discovery: A survey[J]. *IEEE Communications Surveys & Tutorials*, 2007, 9(4): 56-69.
- [2] Laurent B and Traian M. Integrated genetic algorithm and goal programming for network topology design problem with multiple objectives and multiple criteria[J]. *IEEE/ACM Transactions on Networking*, 2008, 16(3): 680-690.
- [3] Jin Xing, Tu Wan-qing, and Chan S. Scalable and efficient end-to-end network topology inference[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2008, 19(6): 837-850.
- [4] Breitbart Y, Garofalakis M, and Martin C, et al. Topology discovery in heterogeneous IP networks[C]. *IEEE International Conference on Computer Communication 2000 Proceedings*, Tel, Aviv, March, 1-3. 2000: 265-274.
- [5] Breitbart Y, Garofalakis M, and Jai B, et al. Topology discovery in heterogeneous IP networks: The net inventory system [J]. *IEEE/ACM Transactions on Networking*, 2004, 12(3): 401-414.
- [6] 张宇, 方滨兴, 张宏莉. 中国 AS 级拓扑测量与分析[J]. *计算机学报*, 2008, 31(4): 611-618.
Zhang Yu, Fang Bin-xing, and Zhang Hong-li. Chinese AS-level topology measurement and analysis[J]. *Chinese Journal of Computers*, 2008, 31(4): 611-618.

- [7] 张昕, 赵海, 王莉菲等. AS 级 Internet 拓扑分析[J]. 通信学报, 2008, 29(7): 50-61.
Zhang Xin, Zhao Hai, and Wang Li-fei, *et al.*. Analysis on the Internet AS-level topology[J]. *Journal on Communications*, 2008, 29(7): 50-61.
- [8] 郑海, 张国清. 物理网络拓扑发现算法的研究[J]. 计算机研究与发展, 2002, 39(3): 264-268.
Zheng Hai and Zhang Guo-qing. An algorithm for physical network topology network topology discovery[J]. *Journal of Computer Research and Development*, 2002, 39(3): 264-268.
- [9] 孙延涛, 吴志美, 石志强. 基于地址转发表的交换式以太网拓扑发现方法[J]. 软件学报, 2006, 17(12): 2565-2576.
Sun Yan-tao, Wu Zhi-mei, and Shi Zhi-qiang. A method of topology discovery for switched ethernet based on address forwarding tables[J]. *Journal of Software*, 2006, 17(12): 2565-2576.
- [10] 孙延涛, 石志强, 吴志美. 交换式以太网物理拓扑结构的自动发现[J]. 计算机研究与发展, 2007, 44(2): 208-215.
Sun Yan-tao, Shi Zhi-qiang, and Wu Zhi-mei. Automatic discovery of physical topology in switched ethernets[J]. *Journal of Computer Research and Development*, 2007, 44(2): 208-215.
- 陈 松: 男, 1976 年生, 博士, 研究方向为分布对象计算、计算机网络技术等.
- 王 珊: 女, 1975 年生, 博士, 讲师, 研究方向为分布对象计算、计算机技术等.
- 周明天: 男, 1939 年生, 教授, 博士生导师, 主要研究领域为计算机网络技术、计算机安全、分布式对象计算.