

# 大规模 IPv6 网络拓扑发现探测目标点的构建和选取<sup>\*</sup>

刘振山 罗军勇 王清贤

(国家数字交换系统工程技术研究中心 郑州 450002)

**摘要** 探测目标点的数量及分布位置对网络拓扑发现的覆盖率和探测效率有直接的影响,因此成为网络拓扑发现研究领域一直关注的焦点。本文针对目前大规模网络拓扑发现目标点相关研究现状,提出了两种目标点集合构建方法,在对目标点冗余分析的基础上提出了相应的探测目标点选取方法。最后通过对 CERNET-2 骨干网络的实际测试,说明了本文提出的这些方法可以推广至国家级规模的 IPv6 网络拓扑发现应用中。

**关键词** 目标点,广度优先搜索,顶级前缀,扩展前缀

## Construct and Select Methods of Probing Targets Set in Large Scale Topology Discovery for IPv6 Networks

LIU Zhen-shan LUO Jun-yong WANG Qing-xian

(National Digital Switching System Engineering & Technology R&D Center, Zhengzhou 450002, China)

**Abstract** The numbers and locations of probing targets problem affect the coverage and efficiency of topology discovery directly, so they are being focus and difficulties in the domain of topology discovery. This paper raised 2 kinds of methods on the construction of probing target sets based on the background of research of probing target for large scale topology discovery. This paper also raised 2 kinds of probing targets select methods based on the research of probing redundancy. In the end, this paper proposed the results of topology discovery for the part of the backbone of CERNET-2 and proved these proposed methods could be extended to large scale topology discovery for IPv6 networks.

**Keywords** Probing target, BFS, Top prefix, Span prefix

一直以来, Internet 路由器级拓扑发现技术大都采用具有普遍适用性的 Traceroute 路径探测机制。在这种机制下,探测目标点(Traceroute 的目的地址)的数量及分布位置直接影响拓扑发现的覆盖率和执行效率,因而成为首要考虑的因素。用于 IPv4 网络拓扑发现的目标点集合构建方法,主要有以下四种:

(1) 根据地址类别,选用不同的网络前缀长度,进行启发式扫描,将应答地址整理成探测目标地址集合<sup>[1]</sup>。

(2) 在给定区间范围内,按照一定的抽样法则抽取探测地址,形成探测目标地址集合。其中比较有代表性的是哈尔滨工业大学姜誉博士提出的“桩网络覆盖,两阶段随机采样”方法<sup>[2]</sup>。

(3) 根据搜集的 Web 服务器列表,形成目标地址集合<sup>[3]</sup>。

(4) 根据 RouteView BGP 表中的目的网络和 AS\_PATH 选择 IP 地址进行探测,称为有导向探测<sup>[4,5]</sup>,这种目标点选取方法的完备性取决于 BGP 路由表中路由信息的完全性及覆盖范围的完全性。

Barford 等人<sup>[6]</sup>所进行的边际效用分析认为增加探测目标点比增加探测源的收益更大。若抽样率过低,则会引起拓扑发现覆盖率缺失;若抽样率过高,则会引起探测冗余显著增加。

现有 IPv6 网络拓扑发现技术主要用于网络测量和网络管理两个方面。用于网络测量的单源 IPv6 拓扑发现,如 Atlas<sup>[7]</sup>,要求事先获知目标网络大致拓扑结构及目标网络内充足的 IPv6 有效地址;多源 IPv6 拓扑发现,如 scamper<sup>[8]</sup>等,要

求在全球不同的位置部署 20 多台探测引擎,来发现这些探测引擎之间的 IPv6 路由器级拓扑结构。大规模网络拓扑发现往往无法事先获知目标网络的大致拓扑结构和充足的探测目标点数量;采用分布式部署方式来实现拓扑发现的代价非常高昂,并且探测引擎之间的通讯维护极为繁琐。最重要的是 IPv6 海量地址空间的特性,使得抽样率与探测效率的矛盾始终无法找到一个折衷的解决方案,因此基于启发扫描方式和抽样方式的目标点集合构建方法同样不适用于 IPv6 网络拓扑发现。针对这种情况,本文从面向国家或地区级的大规模 IPv6 网络拓扑发现的目标点集合构建及目标点选取两个方面展开论述,提出相应的目标点集合构建和选取方法,并通过实际测试验证了相关结论和方法的合理性。

## 1 大规模 IPv6 网络拓扑发现探测目标点集合的构建

### 1.1 网络拓扑发现范围的地理限定

国家或地区级网络拓扑发现的已知条件是分配给该国家或地区的顶级 IPv6 前缀列表。这些顶级 IPv6 前缀中有一些已经分配给不同机构,而有些还处于未分配状态。根据 IPv6 前缀列表的分配信息,可以界定 IPv6 网络拓扑发现的规模。比如目前华东地区 CERNET-2 正在使用的顶级前缀为 2001:250/32, 2001:da8/32, 那么根据 2001:250/32, 2001:da8/32 可以确定拓扑发现的地址范围,这就相当于将拓扑发现的规模大致限定在华东地区内。

### 1.2 基于 IPv6 前缀停止集的探测目标点集合构建方法

美国俄勒冈州立大学的 Route Views 项目部署了 12 台路由器,在其上运行 GNU-Zebra 路由软件的路由信息搜集

<sup>\*</sup> 本研究得到国家 863 高技术研究发展计划资助(基金编号:2006AA01Z409)。刘振山 博士研究生,主要研究方向为 IPv6 网络拓扑发现与网络特征分析、网络复杂度理论;罗军勇 教授,硕士生导师,主要研究方向为网络安全协议分析、分布式计算;王清贤 教授,博士生导师,主要研究方向为网络安全、计算复杂性理论。

器。这些路由器与 ISP 或者交换节点的骨干路由器建立 BGP, 或者 MPBGP 对等关系, 搜集来自对等体路由器的 MP-BGP 更新消息, 每隔 15min 就定期将收到的 Update 数据包存成 MRT 格式的文件, 同时每隔 2h 将自身的 MPBGP 路由表存成 MRT 格式的文件。由于该项目提供的数据实时性好, 对等体路由器覆盖范围广, 因而成为世界各国研究互联网发展的基础数据。我们采集了 RouteView 的 IPv6 路由器的数据源(大小为 1.9G), 选取 2006 年 12 月 1 日 00:00 至 2006 年 12 月 30 日 24:00 的所有 MPBTDs (360 个文件) 和 MP-BUDs (2880 个文件) 作为数据源, 并将全球不同 IPv6 域内的前缀信息整理成表 1 格式。

表 1 IPv6 地址前缀与 AS 号的对应表

IPv6 Address	前缀长度	AS
2001:1200:0:0:0:0:0:0	32	16531
2001:1210:0:0:0:0:0:0	32	2549
2001:12e8:0:0:0:0:0:0	32	16397
.....	.....	.....

输入一个 IPv6 顶级前缀可以通过表 1 匹配出该前缀所属的 AS 号。比如, 通过查询得到 2001:250/32 属于 AS 4538, 2001:da8/32 属于 AS 23910。然后根据 AS 号, 按字符最大长度匹配的原则再次查询表 1, 可以获得顶级 IPv6 前缀之下的扩展前缀。比如, 通过查询表 1 可以获知属于 AS 23910 的一个顶级前缀 2001:da8/32 的扩展 IPv6 前缀有: 2001:da8:3000/48, 2001:da8:3002/48, 2001:da8:4002/48 等。通过这种方法能够获取目标网络内的一部分 IPv6 前缀, 但有些被聚合的 IPv6 前缀无法通过表 1 查询得到。

通过上面这种方式获取的目标 IPv6 网络拥有的前缀集合可表示为:  $P = (P_1 P_2 P_3 \dots P_m / m | 16 < m < 64)$ ,  $m$  表示前缀长度,  $P_i$  表示比特位, 将集合  $P$  中所有元素后面补充 128-m 比特的 0, 得到集合  $F = (P_1 P_2 P_3 \dots P_m \underbrace{0 \dots 0}_{128-m} / m | 16 < m < 64)$ , 将  $F$  中各元素作为目标地址, 进行路径探测, 将返回主机不可达信息 (Destination host unreachable) 的节点添加进探测目标点集合  $R$  中。

说明: 由于探测报文的目标地址为  $P_1 P_2 P_3 \dots P_m \underbrace{0 \dots 0}_{128-m}$ , 探测报文被中间路由器按照最大前缀匹配选路原则或默认路由选路原则, 转发到与  $P_1 P_2 P_3 \dots P_m / m$  完全匹配的一段链路中, 由于  $P_1 P_2 P_3 \dots P_m \underbrace{0 \dots 0}_{128-m}$  为无效的 IPv6 地址, 因此拥有该前缀的 IPv6 路由器返回主机不可达信息。

### 1.3 基于 IPv6 Web 站点搜索的探测目标点集合构建方法

从 2007 年 3 月至 5 月, 我们对 Cernet-2 上运行的 Web 站点进行搜集和整理。整个搜集工作分为两个阶段。

第一个阶段: 搜集了接入 Cernet-2 的不同院校和科研机构的 IPv6 门户网站, 在这个阶段共采集到 147 个 IPv6 Web 站点, 对这些站点进行了域名解析, 获取其 IPv6 地址, 得到 102 个有效 IPv6 Web 站点地址。

第二个阶段: 设计了支持有限空间域内的广度优先搜索算法的 IPv6 Web 网页分析和提取工具。将第一阶段获取的 102 个有效 IPv6 Web 站点作为搜索种子进行广度优先搜索, 并实时解析出不同超链接的 IPv6 地址。我们参考《第 19 次中国互联网络发展报告》提供的国内 IPv6 顶级前缀列表来限制广度优先搜索的范围。在 2007 年 6 月初进行的 IPv6 Web

站点搜索工作中最终解析出有效 IPv6 地址 748 个(包括起始的 102 个搜索种子), 我们将这些地址记录到探测目标点集合  $D$  中。这种目标点构建方法依赖于 IPv6 Web 站点搜索种子的完整性以及这些 IPv6 Web 站点的之间的连接度等属性。

## 2 大规模 IPv6 网络拓扑发现探测目标点的选取

大规模网络拓扑发现的原则是在已知探测目标点数量有限的前提下, 获得目标网络内尽量完整的网络拓扑结构, 同时确保拓扑发现过程遵循低负载、高效和弱感知性的原则。由于 IPv6 网络自身海量地址空间的特性使得拓扑覆盖率和探测效率之间的矛盾更加尖锐, 因此如何在已获取的探测目标点集合中选取合适的探测目标点就成为大规模 IPv6 网络拓扑发现首要考虑的问题。探测目标点选取的基本准则是尽量减少探测冗余的同时不引起拓扑发现覆盖率的缺失。下面就针对不同的冗余情况的分析提出相应的探测目标点选取方法。

前面我们分别通过基于 IPv6 前缀停止集和基于 IPv6 Web 站点搜索的目标点集合构建方法获得了两个探测目标点集合  $R$  和  $D$ 。  $R$  集合由不同路由器的接口对应的 IPv6 地址组成,  $D$  由不同终端主机对应的 IPv6 地址组成。假设探测源点用  $H$  来表示,  $H$  发出的探测报文到达  $R$  中的某个目标点  $r_i$  的路径用  $Path(H, r_i)$  表示, 到达  $D$  中的某个目标点  $d_i$  的路径用  $Path(H, d_i)$  表示, 网络拓扑发现过程中探测目标点引起的冗余主要有以下两种情况:

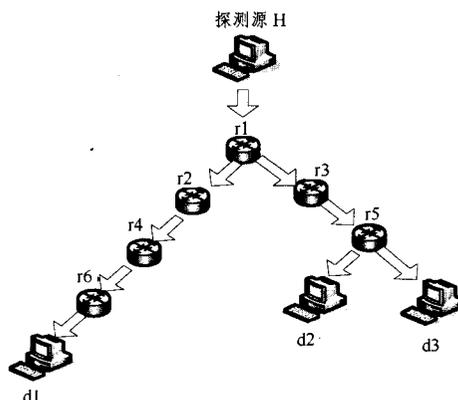


图 1 基本探测路径显示

(1)  $Path(H, d_1)$  和  $Path(H, d_2)$  存在相互包含关系的情况

如图 2 所示,  $Path(H, d_2) = Path(H, d_3)$ , 因此  $d_2$  和  $d_3$  之中只能选一个作为探测目标点, 否则引起探测冗余。因此需要将  $D$  集合中共同拥有 64 位前缀的 IPv6 地址进行去冗处理, 使得  $D$  集合中的 IPv6 地址不重复出现相同的 64 位前缀, 得到集合  $D'$ 。

(2)  $Path(H, r_1)$  和  $Path(H, d_1)$  存在相互包含关系的情况

如图 2 所示, 从探测源  $H$  到探测目标点  $d_1$  的路径包含了从  $H$  到  $r_1, r_2, r_4$  和  $r_6$  的路径, 因此在这种情况下选取  $d_1$  作为拓扑发现的探测目标点能够揭示  $H \rightarrow r_1 \rightarrow r_2 \rightarrow r_4 \rightarrow r_6$  的路径。如果除了选取  $d_1$  之外, 还选取了  $r_1, r_2, r_4$  和  $r_6$  中的任意点作为目标点, 都将会引起探测冗余, 如果不选择  $d_1$  作为目标点, 则会引起覆盖率缺失。考虑这种情况, 我们首先用去冗处理后的  $D$  集合作为探测目标点集合, 进行基本路径发现, 得到路由器节点集合  $T$ , 而后从  $R$  集合中删除属于  $T$  的

元素,得到集合  $R'$ 。

### 3 实验结果与分析

#### 3.1 IPv6 网络拓扑发现系统介绍

我们设计的 IPv6 网络拓扑发现系统采用两种探测子系统:

第一种是 IPv6-in-IPv4 多隧道接入分布式探测子系统,见文献[9]。该方式将所有的探测引擎集中部署在一个 C 类子网中(202.196.X.0 网段),利用公共 ISATAP 路由器和 6to4 中继路由器作为执行 IPv6 网络拓扑发现的探测跳板;用一个中央路径存储数据库存储已获取的路径信息,这样可以有效地避免不同探测引擎对同一路径的重复探测。这种集中部署的方式既达到了分布式网络拓扑发现高覆盖率的探测效果,同时避免了高昂的部署费用和繁琐的通讯维护工作,如图 2 所示。

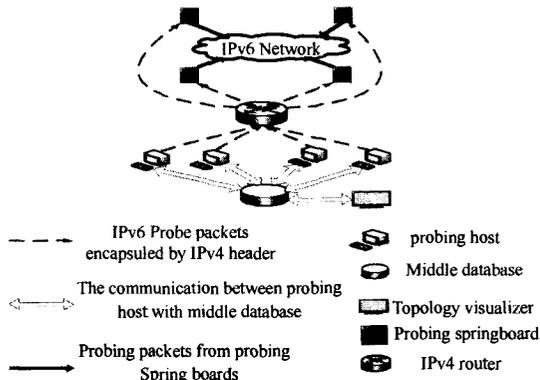


图 2 IPv6-in-IPv4 多隧道接入分布式探测子系统

第二种是纯 IPv6 网络环境下的单源源路由探测子系统。单源源路由探测子系统布置在 CERNET-2 郑州大学接入中心,如图 3 所示。

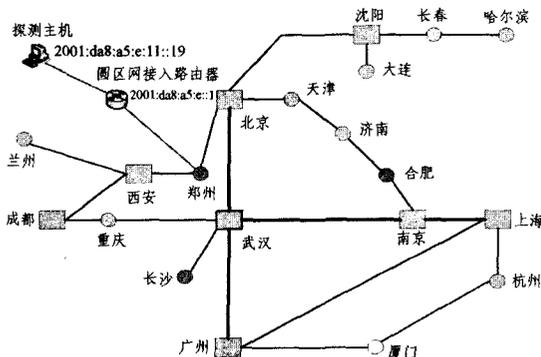


图 3 纯 IPv6 接入方式测试环境

#### 3.2 对比测试与分析

整个测试分为两步:

表 2 利用原始的探测目标点集合得到的测试数据

IPv6-in-IPv4 多隧道接入分布式探测子系统采用的探测跳板数	5
单源源路由探测子系统设置的最大线程数	10
测试总时间(基本路径采集+别名归并)	约 110min
Traceroute 执行总数	13511 次
探测出路由器总数(别名归并后)	177
探测出链路数总数(别名归并后)	319

第一步,将集合  $D$  作为 IPv6-in-IPv4 多隧道接入分布式探测子系统的探测目标点集合,将集合  $R$  作为单源源路由探测子系统的探测目标点集合,进行网络拓扑发现,并对基本的 Traceroute 执行设置全局加法器,以记录执行总数,得到表 2 数据。

第二步,将集合  $D'$  和  $R'$  分别作为两个拓扑探测子系统的探测目标点集合,进行拓扑发现,同时探测过程两个子系统共享路径存储数据库,得到表 3 数据。

表 3 利用选取后的探测目标点集合得到的测试数据

IPv6-in-IPv4 多隧道接入分布式探测子系统采用的探测跳板数	5
单源源路由探测子系统设置的最大线程数	10
测试总时间(基本路径采集+别名归并)	约 75min
Traceroute 执行总数	9856 次
探测出路由器总数(别名归并后)	163
探测出链路数总数(别名归并后)	297

通过表 2 和表 3 的测试数据进行对比可知,在相同的实验环境和拓扑发现算法下,新的探测目标点集合使得 Traceroute 执行总数和测试时间显著减少。第二步探测出的路由器总数和链路数总数比第一步减少,可能基于以下两个原因:

- (1)网络中某些路由器的工作状态发生了变化。
- (2)设置单一的 ICMPv6 请求超时时间为并不能满足所有的网络状况。

综上,本文提出的目标点集合冗余去除方法在显著减少探测冗余的同时,并不会降低拓扑发现的覆盖率。

**结束语** 本文提出了两种探测目标点集合构建方法,并在冗余分析的基础上提出了探测目标点选取方法。最后通过对比试验说明了本文提出的相关方法在不引起拓扑发现覆盖率缺失的前提下能够显著地减少探测冗余。影响拓扑发现探测冗余和覆盖率的因子还包括探测源的数量及相对于目标网络的分布位置、拓扑探测算法的普适性等亟待解决的问题,而这些问题将作为我们下一步的研究工作的重点。

#### 参考文献

- [1] Govindan R, Tangmunarunkit H. Heuristics for Internet Map Discovery // Proc. IEEE INFOCOM 2000. Tel-Aviv, Israel, Mar, 2000, 3:1371-1380
- [2] YU J, Binxing F, Zengming H. Embellishment on Greedy Algorithm for Set Cover Problem // Proc. ICITA 2004. Harbin, P. R. China, Sydney: Macquarie Scientific Publish, 2004: 167-172
- [3] Huffaker B, Plummer D, Moore D, et al. Topology Discovery by Active Probing // Proc. 2002 IEEE Symposium on Applications and the Internet Workshops (SAINT' 02w). Nara, Japan, 28 Jan. -1 Feb. 2002: 90-96
- [4] Spring N, Mahajan R, Wetherall D. Weasuring ISP Topologies with Rocketfuel. ACM SIGCOMM CCR, 2002, 32(4): 133-145
- [5] Mao Z Q, Rexford J, Wang J. Towards an Accurate AS-Level Traceroute Tool // Proc. ACM SIGCOMM 2003. Karlsruhe, Germany, Aug. 2003, New York: ACM Press, 2003: 365-378
- [6] Barford P, Bestavros A, Byers J, et al. On the Marginal Utility of Network Topology Measurement // Proc. 1st ACM SIGCOMM Workshop on Internet Measurement (IMW 2001). San Francisco, California, Nov. 2001: 5-17
- [7] Waddington D G, Chang Fangzhe, Ramesh V, et al. Topology Discovery for Public IPv6 Networks [J]. ACM SIGCOMM Computer Communications Review, 2003, 33(3): 59-68
- [8] IPv6 Scamper. <http://www.wand.net.nz/~mjl12/ipv6-scamper/>
- [9] Liu Zhenshan, Wang Qingxian, Luo Junyong. A new distributed topology discovery technology for IPv6 networks // Proc. IEEE SoutheastCon, 2007. March 2007: 627- 632