文章编号：2095-2163(2020)01-0249-04 中图分类号：TP393.08 文献标志码：A

一种高效的末跳路由发现技术

刘 洋，方滨兴
（哈尔滨工业大学 计算机科学与技术学院，哈尔滨 150001）

摘 要：本文提出了一种高效的目标主机末跳路由器发现方法，此方法相比于 traceroute 发包量显著降低。通过向目标发送精心构造的 UDP 大端口包，能够一次探测获取网络距离，两次探测即可获取末跳路由器。考虑互联网存活主机中仅有 20%对 UDP 大端口进行响应，本文又提出二分法进行网络距离计算，进而获取末跳路由器信息。与 traceroute 相比，能够稳定有效地降低发包量，不受网络距离的影响。

关键词：末跳路由发现；网络距离；traceroute

An efficient end-hop router discovery method

LIU Yang, FANG Binxing
(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

【Abstract】 This paper proposes an efficient target host end-hop router discovery method. Compared with traceroute, this method significantly reduces the amount of packets sent. By sending a carefully constructed UDP big port packet to the target, it is possible to detect the network distance with one packet and obtain the end-hop router with two packets. Considering that only 20% of the Internet surviving hosts respond to UDP big ports packet, this paper proposes a binary method for network distance calculation, and obtains the last hop router information. Compared with traceroute, it can stably and effectively reduce the amount of packets, regardless of the network distance.

【Key words】 end-hop router discovery; network distance; traceroute

0 引 言

末跳路由器是测量点到目标 IP 地址的 IP 路径上，与目标 IP 直接相连的最后一跳路由器。高效末跳路由器发现技术，实现用尽可能少的探测包发现到目标的末跳路由器。

高效末跳路由器发现包含了对网络距离的快速获取，可以为实现更高效的 traceroute 探测[1-3]提供依据。例如在已知到目标的网络距离 D 时，traceroute 可以让 TTL 从 D 开始反向测量，遇重复探测 IP 提前停止，这样能显著地降低测量冗余；或者一次性同时发送 TTL 从 1 到 D 的所有探测包，在不引入额外测量负载的同时，显著减少测量时间（O(D^2) 到 O(D)）。此外，因为终端 IP 的末跳路由器和目标 IP 的拓扑邻近性，则有助于对二者进行 IP 地理定位[4-5]。

获取末跳路由器最朴素的办法是进行完整的 traceroute 测量，但是仅就发现末跳路由器这一目的来说，对中间路由的测量并无必要。如果已知到目标的距离，发送一个探测包能够获取末跳路由器。文献[6]中采用向目标发送大端口侦测包的方法，根据回复的字段来推断到目标的距离。但是，考虑到往返路径的不对称性等原因，基于大端口侦测包的方法用于网络距离预测可能存在偏差。此外，并不是所有终端都会对测量包给予回复，事实上，本文实验所用到的存活主机仅有 20%左右的主机会做出应答，这也为获取所有末跳路由器带来了困难。

针对上述问题，本文结合网络距离预测和二分策略的 traceroute 技术，提出了更通用的高效末跳路由器发现方法。本文的安排如下：首先，研究了基于 ICMP 端口不可达报文的网络距离估计方法的设计实现，讨论了一种基于二分法的网络距离及末跳路由获取方法。然后给出实验结果与分析。最后，对本文的工作进行总结与展望。

1 基于 ICMP 端口不可达报文的网络距离计算

1.1 从 ICMP 端口不可达报文提取网络距离

ICMP 协议是互联网中报文消息控制协议，通过调研发现，向目标发送 UDP 大端口报文，目标会返回端口不可达报文，ICMP 端口不可达报文结构

基金项目：国家重点研发计划(2016YFB0801303-2)。

作者简介：刘 洋(1995-)，男，硕士研究生，主要研究方向：网络拓扑测量；方滨兴(1960-)，男，博士，教授，博士生导师，主要研究方向：并行计算、网络安全、拓扑发现。

收稿日期：2019-05-30

如图 1 所示。图 1 中左侧为 ICMP 报文的结构,分别由 IP 头、ICMP 头以及 ICMP 负载组成。其中,ICMP 负载部分包含探测源发送给目标的原始报文数据。从原始报文的 IP 头部,就能够提取到生存时间 *TTL*,研究将这个 *TTL* 值定义为 *raw_ttl*。探测源发送初始 *TTL* 值为 *init_ttl* 的 UDP 报文,该报文到达目标时生存时间 *TTL* 值减小到 *raw_ttl*。因此,*init_ttl* 减去 *raw_ttl* 的值是中间路由器的数目,而路由器数目加 1 也就是源到目标的网络距离。至此,基于现有理论可知,一共发送 2 个探测包可获取末跳路由器信息。

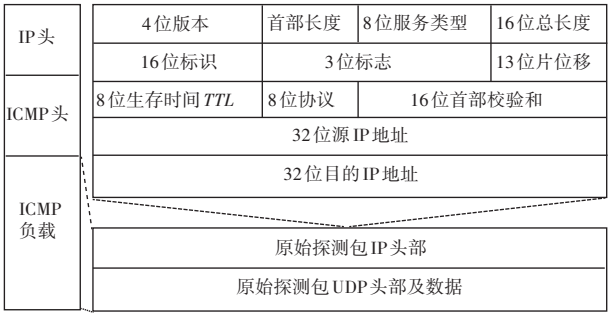

图 1 ICMP 端口不可达报文结构

Fig. 1 ICMP port unreachable packet structure

1.2 末跳路由器发现设计与实现

为了捕获目标返回的 ICMP 消息和末跳路由器信息,需要 2 个监听器(Listener)分别监听 ICMP 端口不可达报文和 ICMP 生存时间超时报文。该方法的时序图如图 2 所示。

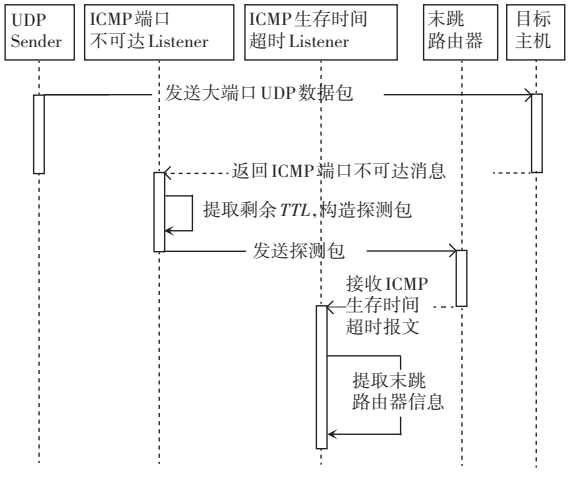

图 2 基于 ICMP 端口不可达获取末跳路由器时序图

Fig. 2 Obtaining the end-hop router timing diagram based on the ICMP port unreachable

具体设计流程可表述为:首先,UDP Sender 构造 UDP 大端口探测包发送给目标主机,若目标主机指定端口未开放,返回 ICMP 端口不可达报文,此报文的数据部分填充为原始 UDP 报文;当本地 ICMP 端口不可达 Listener 捕获到目标主机返回的报文后,从报文的 ICMP 负载部分提取 UDP 大端口探测包中的生存时间 *raw_ttl*,根据 *raw_ttl* 计算网络距离;向目标主机发送 *TTL* 为网络距离减 1 的探测包,此探测包到达目标主机的末跳路由器时 *TTL* 值刚好减为零,触发末跳路由器返回生存时间超时报文,本地 ICMP 生存时间超时 Listener 会捕获末跳路由器返回的报文,从报文中提取末跳路由器信息。

2 基于二分法的网络距离估计

文中前一节提出基于 ICMP 端口不可达报文的末跳路由发现方法高效而稳定,唯一的不足就是并非全部目标主机会响应 UDP 大端口报文。因此,需要寻找一种普适性的方法,这种方法既要发出尽可能少的探测包,又能够获取全部存活主机的末跳路由器。为此,本节提出基于二分法的网络距离估计,相对于传统 traceroute 来说有效减少发包数量。对此部分,本文将给出阐释分述如下。

2.1 二分法判定网络距离

Traceroute 在收到目标回复时停止探测,在此之前由于探测包设置的生存时间 *TTL* 小于网络距离,导致中间路由器回复 ICMP 生存时间超时消息,包括末跳路由器。因此,研究中可以得出只要满足 2 个条件,可以确定网络距离为 *D*,同时也获取了末跳路由器。对这 2 个条件可做总体概述如下。

(1) 向目标发送初始 *TTL* 等于 *D* 的探测包,目标主机返回响应报文。

(2) 向目标发送初始 *TTL* 等于 *D* 减 1 的探测包,收到生存时间超时报文。

为了尽快满足(1)、(2) 两个条件,探测包生存时间 *TTL* 不必设置成从 1 开始探测。如果发送方收到生存时间超时报文,说明此时探测包设置的 *TTL* 比较小,探测包还未到达目标主机,*TTL* 值就减为 0,需要增大探测包的生存时间。如果发送方收到目标主机返回的报文,说明探测包设置的 *TTL* 值不小于网络距离 *D*,此时减小下次发送探测包的生存时间。研究知道网络中任意 2 个节点的网络距离一般不会超过 30 跳,发送方到目标的网络距离在 1 ~30 之间,因此可以用二分法逐渐快速逼近真实的网络距离。事实上,此方法在获知网络距离的同时,已经获取了末跳路由器信息,因为(2) 中的生存时间超时报文由末跳路由器返回。

2.2　基于二分法的网络距离算法

二分法通过不断缩小初始设置的 *TTL* 范围直到确定真实的网络距离。算法流程如图 3 所示。算法设计流程可表述如下：对于给定的初始 *TTL* 范围[1，30]，每次探测包的初始 *TTL* 取范围的中间值 *mid_ttl*；如果收到生存时间超时报文，记录 *time_exceeded_flag* 为 *mid_ttl*，*TTL* 范围应取右半边[*mid_ttl*，30]；如果收到目标主机回复的报文，说明 *mid_ttl* 大于等于实际网络距离，记录 *echo_reply_flag* = *mid_ttl*，范围应该取左半边[1，*mid_ttl*]，以此类推，直到满足公式(1) 为止。该式可写作如下数学形式：

$$echo\_reply\_flag = time\_exceeded\_flag + 1. \quad (1)$$

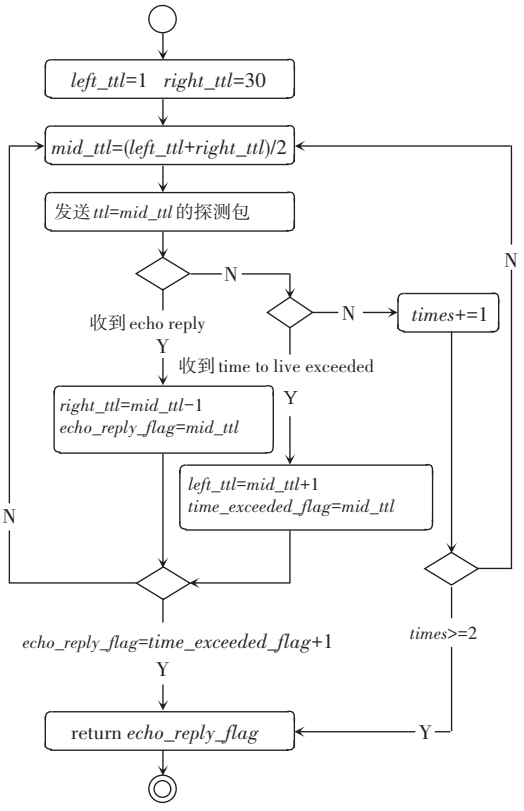

图 3　二分法获取末跳流程图

Fig. 3　Binary method for obtaining end-hop routing

算法设计步骤详见如下：

Step 1　初始化 *left_ttl* = 1，*right_ttl* = 30，*Distance* = − 1；

Step 2　*mid_ttl* = (*left_ttl* + *right_ttl*) > > 1；

Step 3　发送初始 *TTL* = *mid_ttl* 的探测包；

Step 4　若收到 time to live exceeded 回复，*left_ttl* = *mid_ttl* + 1，*time_exceeded_flag* = *mid_ttl*；

Step 5　若收到目标主机的回复，*right_ttl* = *mid_ttl* − 1，*echo_reply_flag* = *mid_ttl*；

Step 6　如果 *echo_reply_flag* 等于 *time_exceeded_flag* + 1，*Distance* = *echo_reply_flag*，跳到 Step 7；否则跳到 Step 2；

Step 7　返回网络距离 *Distance*。

3　实验结果及分析

3.1　末跳路由器发现对比实验

实验选取 10 万个目标主机，分别使用 ICMP 端口不可达、二分法和传统 traceroute 方式进行末跳路由器获取，分别对比其发包量以及末跳路由器获取情况，并对实验结果进行分析。

实验结果如图 4 所示。10 万个目标主机中，发现 71 000 个存活主机。其中，traceroute 获取末跳路由器数目为 48 602 个，基于 ICMP 端口不可达方法获取末跳路由器数目为 9 811，二分法获取末跳路由器数目为 43 010 个。相比于传统方法，二分法获取率下降了 11%，基于 ICMP 端口不可达方法获取率仅为 traceroute 的 20%。

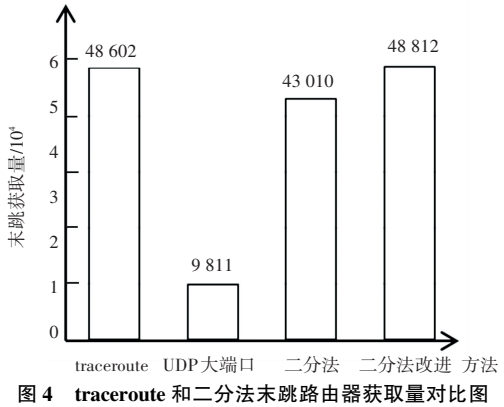

图 4　traceroute 和二分法末跳路由器获取量对比图

Fig. 4　Comparison of traceroute and end-hop router acquisition

在发包量方面，统计实验结果显示 ICMP 端口不可达方法发包量为 19 622 个，二分法发包总数为 250 929，传统 traceroute 发包总数为 675 604。由于每种方法获取的末跳路由器数目不一致，因此不能仅是对比其发包量一项，而应考查平均发包量，即发包量总数除以获取末跳路由器的数目就是平均发包量，测试对比结果如图 5 所示。由图 5 分析可知，ICMP 端口不可达方法平均发包量为 2，符合理论值。二分法平均发包为 5.17 个，这是由二分法区间[1，30]决定的，经过 5 次二分可以将区间范围缩小到 1，因此 5.17 符合二分法理论发包值。而传统 traceroute 是每次探测包 *ttl* 加 1，对于目标主机其发包量即为源到目标的实际网络距离值。13.9 说明 10 万个目标主机中，源到目标主机的平均网络距离为 13.9。二分法相对于传统方法的发包率降低了 63.02%，ICMP 端口不可达方法降低了 85%。

3.2 实验结果分析

针对 traceroute 获取而二分法未获取到目标开展进一步分析后发现,二分法需要每次取区间的中值作为本次探测包的生存时间 *TTL* 值,如果没有收到对此探测包的回应,就无法根据返回包的类型继续缩小区间,二分法就无法进行,导致无法获取网络距离,也就无法获取末跳路由器。这是二分法的末跳路由器获取率低于传统方法的原因。分析上述问题可知,未收到回应报文的原因往往是探测包设置的生存时间小于网络距离,否则发送方会收到来自目标的响应报文。因此,将区间左侧边界增加 1,重新进行二分计算本次探测包设置的 *TTL* 值,以此类推。改进后再经实验测试发现改进后二分法和 traceroute 相比末跳路由器发现率基本一致(见图 5)。

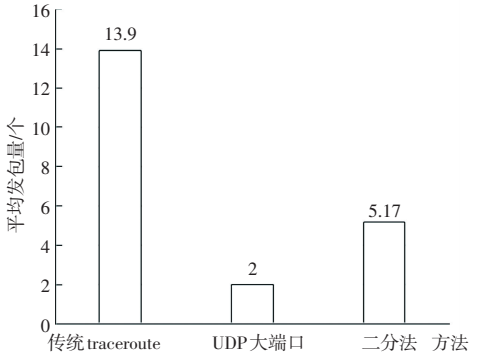

图 5 末跳路由器获取平均发包量对比

Fig. 5 The end-hop router obtains the average amount of sent packets

根据实验结果,ICMP 端口不可达方法末跳获取率仅为 20%左右,分析发现网络中仅有 20%左右的目标对 UDP 大端口报文做出响应,此方法不具有普适性。相反,二分法获取网络距离和末跳的本质和 traceroute 一致,因此末跳路由获取量大致相等。结合 ICMP 端口不可达方法发包少以及二分法末跳获取率高且稳定的特点,研究中将 2 种方法进行整合。先对目标使用 ICMP 端口不可达方法,对于未获取的目标再使用二分法进行获取,理论上整合后的方案平均发包量应该介于 2 种方法之间。同样选取 10 万个目标,对合并后的方案进行末跳路由器获取。统计合并后方案的末跳获取量与 traceroute 相当,但平均发包量为 5.6,反而高于 2 种方法。

分析发现,由于 ICMP 端口不可达方法只对 20%左右的目标有效,但是也需要对每个目标发送 UDP 大端口报文。由于这部分无效的发包,导致平均发包量不降反升。综上论述可知,采用二分法进行末跳路由器发现能够发现较为完整的末跳路由器信息,若不考虑末跳路由器发现的完整性,使用 ICMP 端口不可达方法能够显著降低发包量。

4 结束语

本文提出了一种高效的末跳路由器探测方法。相比于 traceroute 发包量平均降低了 60%,本文方法在最少时仅需要 2 个探测包即可获取末跳路由器。末跳路由器发现只关注目标主机的最后一跳信息,将其作为一项新型的拓扑关系数据,对网络拓扑测量、IP 地理位置定位具有一定的参考价值。在未来网络测量的研究中,这种特殊的网络拓扑信息则有着重要的应用和研究价值。

参考文献

[1] AUGUSTIN B, CUVELLIER X, ORGOGOZO B, et al. Avoiding traceroute anomalies with Paris traceroute[C]//Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement. Rio de Janeiro, Brazil: ACM, 2006: 153-158.

[2] MAO Z M, REXFORD J, WANG J, et al. Towards an accurate AS-level traceroute tool[C]//Proceedings of ACM SIGCOMM. NewYork,USA:ACM, 2003: 365-378.

[3] KATZ-BASSETT E, MADHYASTHA H, ADHIKARI V, et al. Reverse traceroute[C]//The 7th Usenix Symposium on Networked Systems Design and Implementations (NSDI). San Jose, CA, USA:Usenix,2010: 219-234.

[4] BENDALE J, KUMAR J R. Review of different IP geolocation methods and concepts[J]. International Journal of Computer Science and Information Technologies, 2014,5(1):436-440.

[5] WANG Yong, BURGENER D, FLORES M, et al. Towards street-level client-independent IP geolocation[C]// Usenix Conference on Networked Systems Design and Implementation. Berkeley, CA, USA:Usenix Association, 2011:365-379.

[6] MOORS T. Streamlining traceroute by estimating path lengths[C]//2004 IEEE International Workshop on IP Operations and Management. Beijing, China: IEEE, 2004: 123-128.

(上接第 248 页)

[3] 高凯悦. DY 融资租赁公司财务风险评价研究[D]. 西安:西安石油大学,2018.

[4] 路英娥,管红波. 海洋工程装备制造业上市企业财务风险评价[J]. 上海管理科学,2018,40(4):60-64.

[5] OSTROM E. Building trust to solve comments dilemmas: Taking small steps to test an evolving theory of collective action[M]// LEVIN S A. Games, groups and the global good. New York: Springer,2008:211-216.

[6] 张丽哲. 基于大数据的企业财务风险管理体系模型的构建[J]. 中国管理信息化, 2017,20(17):27-28.

[7] 蔡立新,李嘉欢. 大数据时代企业财务风险预警机制与路径探究[J]. 财会月刊,2018(15):38-43.

[8] 宋彪,朱建明,李煦. 基于大数据的企业财务预警研究[J]. 中央财经大学学报,2015(6):55-64.

[9] 赵津怡. 大数据时代的企业财务风险预警研究[J]. 财会学习,2018(21):15-16.