计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200300078

基于 LRBG 方法的 IP 定位研究

赵 茜 陈曙晖
国防科技大学计算机学院 长沙 410037

摘 要 IP 地址是互联网设备的网络标识,IP 定位根据网络设备的 IP 地址来确定网络设备所在的地理位置。地标是 IP 定位中的一个关键因素,以往的研究采用家庭 PC、服务器或共用路由器作为地标,受 IP 地址动态分配、地标地理分布不均匀和时延-距离换算关系复杂等因素的影响,IP 定位结果不够精确。traceroute 工具可以定位出探测点至目标主机之间的所有路由器,Last-hop Router Based Geolocation(LRBG)方法以 traceroute 路径中的最后一跳路由器(LR)为地标,将 IP 定位问题分解为两步。第一步是以 LR 服务范围内的固定网络宽带用户为参照物,估算 LR 的地理位置。第二步是将 LR 作为地标,根据 LR 与目标主机的网络拓扑关系确定目标主机的地理位置。实验结果表明,LRBG 方法实现了 IP 地址的街道级定位,平均精度为 3.17 km。

关键词:IP 定位;网络测量;位置测量;路由器定位;地标
中图法分类号 TP391

LRBG-based Approach for IP Geolocation

ZHAO Qian and CHEN Shu-hui
College of Computer, National University of Defense Technology, Changsha 410037, China

Abstract IP geolocation determines the geographic location of network devices based on their IP addresses, which are the identifications of Internet devices. Landmark is a key factor in IP geolocation. Prior methods use home PCs, web servers as well as common routers as landmarks, they produce erroneous results due to changeable IP addresses, inconsistent density as well as complicated geometric relations between time delay and distance. Traceroute command is able to find all the routers between a probe and the target host. This paper proposes a new method named Last-hop Router Based Geolocation method(LRBG). The last-hop router in a traceroute path is used as the landmark. The problem is solved by two steps. The first step is to employ the fixed Internet users within the range of a last router's delivery to infer its location. The second step is to identify the geographic location of target host based on the relation between the target host and the last hop router. The experiment results show that the LRBG method achieves street-level geolocation of IP address with an average accuracy of 3.17 km.

Keywords IP geolocation, Network measurement, Position measurement, Router geolocation, Landmarks

1 引言

IP 定位使互联网主机的 IP 地址与地理位置形成映射关系,在网络安全领域发挥着重要作用。例如感知 DoS 攻击中的僵尸主机的地理分布,辅助判断 DoS 攻击规模,辅助定位攻击源的地理位置;溯源 APT 组织,为判定攻击者身份、所属团体、攻击意图提供线索;解析恶意样本中的 IP 地址,为判断样本同源性提供依据等。

根据定位是否依赖 GPS、蜂窝基站、蓝牙等客户端,IP 定位算法分为基于客户端的定位算法和独立于客户端的定位算法。独立于客户端的 IP 定位不具备硬件支持,常用于定位家庭 PC、路由器、服务器或其他非合作环境下的网络主机,定位精度介于几千米至几百千米之间。相关研究表明[1-4],地标在独立于客户端的定位算法中发挥着重要作用。作为 IP 地址和地理位置已知的网络主机,地标可用作目标主机的参照物。部分定位方法直接将目标主机绑定到最近的地标上,例如 GeoPing[5] 将相对时延最接近的地标的位置确定为目标主机的位置。其他方法[6-7] 将地标与目标主机之间的往返时延(RTT)乘以传输速率换算成相对距离,再通过线性回归、矩阵运算、概率计算等方法对地标和目标主机之间的相对位置进行进一步量化。

常用的地标分别是家庭 PC 和网络服务器。家庭 PC 的 IP 地址是动态 IP 地址,频繁地被分配给其他网络主机。采用家庭 PC 作为地标[5] 对研究方法的时效性要求很高。路由器和部分长时间在线的服务器采用静态 IP 地址,变更周期较长。SLG[2] 和 structon[8] 采用万余台架设在本地的网站服务器作为地标,二者的研究发现对于 IP 定位研究产生了深远的影响。服务器作为地标的缺点是:服务器通常集中分布于部分地区,如高新产业园区、高教园区和其他科技较发达地区,在其他区域的数量较少,因此定位一个位于居民小区、乡村或其他区域的目标主机时,服务器与目标主机的实际地理位置较远,往返时延受到时延膨胀、迂回路径等因素的影响,造成距离的计算误差较大。此外,云服务和 CDN 的发展和应用打破了地理位置和 IP 地址的对应关系,减少了可用地标的数量。

本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家重点研发计划(2018YFB0204301)
This work was supported by the National Key Research and Development Program of China (2018YFB0204301).

通信作者:赵茜(zhaoqian@nudt.edu.cn)

针对地标选择问题，本文对目标主机的 IP 地址进行 traceroute，采用 traceroute 路径中的最后一跳路由器（Last-hop Router，LR）作为地标。LR 作为地标主要有以下 3 个优势：1）最后一跳路由器采用静态 IP 地址，IP 地址变化周期较长；2）与网络服务器相比，最后一跳路由器的地理分布更均匀；3）最后一跳路由器与目标主机的平均距离为几千米，定位最后一跳路由器后，将 LR 的地理位置作为目标主机的近似位置，可以满足街道级的定位需求。

使用路由器作为地标首先要确定路由器的 IP 地址和地理位置。对于路由器地理位置的研究，文献[9-10]利用路由器的主机名、Whois 等信息对路由器的位置进行推断，例如根据路由器的别名 corerouter1. SanFrancisco. cw. net，将其定位到美国旧金山，该方法只能实现城市级的定位，且对于没有别名的路由器会失效。文献[11]受到 SLG 和 Structon 方法的启发，将网络服务器作为参照物，根据参照物的地理分布推测路由器的位置，可以较准确地定位出服务器分布密集的区域的路由器，对于其他区域的路由器定位精度较低。本文继承和发展文献[11]的思路，首先采集固定宽带用户的地理位置，然后从探测点对宽带用户进行 traceroute，路由器响应探测包后就会暴露自身的 IP 地址以及拓扑关系。最后，根据固定宽带用户的地理分布，基于极大似然估计计算最后一跳路由器的地理位置。

本文的主要贡献有：1）提出使用最后一跳路由器作为地标的方法，解决了 IP 地址变更频率高和地理分布不均匀的问题；2）针对路由器定位问题，提出使用固定宽带用户作为参照物，提出了路由器定位的数据挖掘方法，解决了参照物覆盖不均匀、数量不足等问题，采用极大似然估计解决路由器定位问题，实现路由器的街道级定位。

2 相关工作

独立于客户端的 IP 定位研究，通过数据收集、网络测量和关联分析对网络主机的地理位置进行推测，研究中包含 3 类节点：1）地标，IP 地址和地理位置已知的网络设备；2）目标主机，待定位的网络设备；3）探测点，对地标和目标主机进行网络测量的网络主机。

地标的选取与实验结果有着密切的联系。GeoPing 等[5]首次提出“地标”概念，该方法将部署在美国的 265 台大学主机以及部分互联网用户作为地标，以 14 台服务器为探测点，分别测量每个探测点至地标或目标主机的往返时延向量，将与目标主机时延向量最接近的地标的位置作为目标主机的地理位置，平均误差为 400 km，后续的网络测量方法大多是在该方法的基础上发展而来的。CBG 算法[6]使用 PlanetLab[12]平台的 137 个节点作为地标，将地标间的往返时延换算为相对距离，再分别以地标为圆心、以距离约束为半径画圆，取多点定位的交集作为目标主机的地理范围，在美国实验的平均误差为 95 km。TBG 算法[13]使用 68 个 PlanetLab 服务器和 128 个大学主机作为地标，在 CBG 方法的基础上进行进一步发展，将时延和网络拓扑结构换算为距离约束，精度分别为 180 km 和 67 km。以上方法直接将往返时延转换为距离约束，受时延抖动、排队时延、处理时延等因素的影响，精度较低。

Octant[14]，NBIGA[7]和 Posit[15]等方法使用了概率和统计模型，以减弱时延不确定性对精度的影响。Octant 方法[14]使用了 50 个地标，从探测点发送 traceroute 命令至目标主机，以探测点为起点，推测第二跳节点的最佳区域，并将推测出的节点作为辅助地标，推测下一跳节点，依此类推直至目标节点，平均精度为 35 km。Posit 方法[15]使用极大似然估计得出目标主机的位置，平均精度为 45 km。NBIGA[7]是基于朴素贝叶斯理论的定位算法，该方法将 IP 定位问题转化为机器学习分类问题，使用了 225 个地标，以探测点至地标的时延和路由跃点数为输入，计算地标在不同城市的概率，平均精度约为 50 km。

以上方法大多采用几十至几百个地标，而 SLG 方法则极大提高了地标密度。SLG[2]算法使用了美国境内七万余台网络服务器作为地标，将与目标主机相对时延最小的地标的位置作为目标的位置，突破性地将 IP 定位精度提升至街道级，平均精度为 2 km[4]，是目前已知的精度最高的方法。

NCR-geo 算法[11]采用共用路由器定位目标主机，根据网络服务器到共用路由器之间的时延推测出共用路由器的大致地理区域，将共用路由器的地理区域作为目标 IP 的位置。“共用路由器”是网络测量发现的地标与目标主机之间的相同的路由器。与最后一跳路由器相比，共用路由器可能是 traceroute 路径中的任意一跳，路由器距离目标主机越远，时延与距离之间的换算关系越复杂。若“共用路由器”与目标主机相距多跳，则实验结果的误差较大。

3 定位最后一跳路由器与目标 IP

采用最后一跳路由器作为地标可以解决动态 IP 地址不稳定、服务器的地理分布不均匀以及时延-距离转化关系复杂等问题。LRBG 方法将定位目标主机分为两步，第一步是定位 LR，第二步根据 LR 与目标主机的拓扑关系定位目标主机。

在第一步中，首先采集 LR 服务范围内的 WiFi 位置，根据 WiFi 的地理分布基于极大似然估计推测 LR 最有可能的位置。详细步骤如图 1 所示。

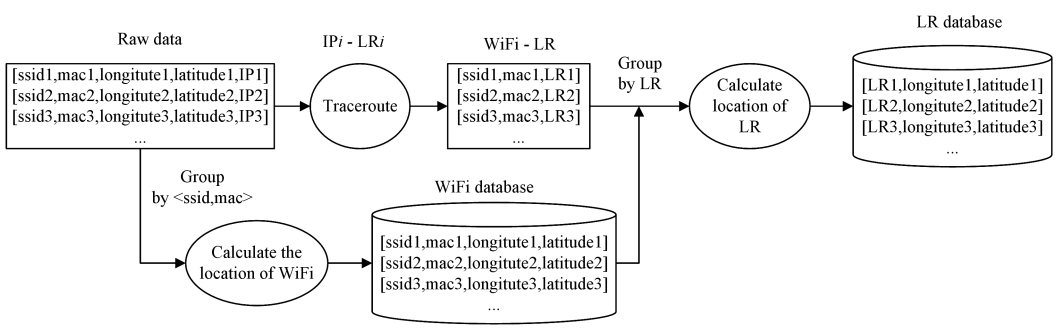


图 1 定位 LR

Fig. 1 LR geolocation

(1)收集安装了 WiFi 的固定网络宽带用户的数据信息，包括 IP 地址、SSID、MAC 地址和所在地理位置的经纬度坐标。

(2)对固定宽带用户的 IP 地址进行 traceroute 并提取 traceroute 路径中的 LR，将 LR 相同的固定宽带用户分为一组。

(3)通过每组宽带用户的分布计算 LR 的位置，并将 LR 的 IP 地址、经度、纬度和半径保存至数据库中。

3.1　数据收集

LR 服务的设备包括一定范围的所有手机终端、网络服务器、固定宽带用户等。固定宽带用户最适用于作为定位 LR 的参照物，原因如下：

(1)手机终端通过蜂窝网络接入互联网时，数据传输依次经过基站、S-GW、PGW 和路由器，不适用于定位 LR。

(2)网络服务器在城市高新产业园区、高教园区、城市中心等地区分布密集，在其他地区的数量较少，地理分布不均匀，只适用于定位少量 LR。

(3)《2019 年中国互联网发展报告》指出“截至 2018 年底，全国固定宽带用户量达 4.07 亿”，与前两者相比，固定宽带用户数量较多、与 LR 位置较近、地理分布更均匀，因此适用于作为定位 LR 的参照物。

本文中的固定网络宽带用户特指安装了 WiFi 的固定宽带用户。收集方法为：使用手机连接 WiFi 后，利用自研 APP 获取 WiFi 的 IP 地址、SSID 和 MAC 地址。将手机 GPS 客户端获取的经纬度作为 WiFi 的经纬度。表 1 列出了收集数据的格式，〈SSID，MAC〉二元组唯一确定一个 WiFi 热点。对于被多次连接的 WiFi，对测得的经纬度求平均值，$Lat_{\text{mid}}=\frac{\sum long_i}{N}$，$Lat_{\text{mid}}=\frac{\sum lat_i}{N}$（$N$ 表示数据条数）。

表 1　固定网络宽带用户的信息

Table 1　Information of fixed Internet users

SSID	MAC	IPaddress	longitude	latitude	date
facha	0c:d8:6c:*:*:*	222.*.195.134	118.714262	31.971691	2019-6-20 8:23
facha	0c:d8:6c:*:*:*	114.*.240.39	118.714359	31.971674	2019-7-28 9:35
facha	0c:d8:6c:*:*:*	121.*.154.14	118.714274	31.971688	2019-10-9 10:08

3.2　计算 LR 的服务范围

IP 地址为 58.*.201.206 的 LR 及其服务范围内的 WiFi 分布如图 2 所示。根据实际分布，LRBG 方法将 WiFi 样本的经度和纬度简化为近似服从二维正态分布，并假设最后一跳路由器位于正态分布中心，LR 的服务半径为 LR 至最远 WiFi 的距离，使用极大似然方法估计正态分布中心的经纬度值。以 x，y 表示 WiFi 的经纬度坐标，μ_1 和 μ_2 表示正态分布中心。由于经度值 x_i 和纬度值 y_i 相互独立（$\rho=0$），概率密度为：

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

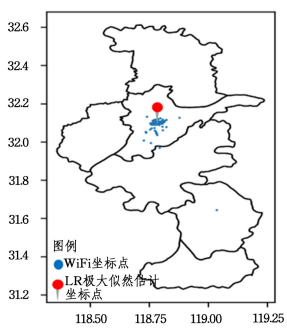


图 2　同一个 LR 的 WiFi 位置分布图

Fig. 2　Distribution of WiFi users within same last router

设 WiFi 样本的集合为 $D=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，且样本是独立同分布的。设 $\mu=(\mu_1,\mu_2)$，联合概率密度 $p(D|\mu)$ 相对于 $\{x_1,x_2,\cdots,x_n\}$ 的似然函数为：

$$l(\mu)=p(D|\mu)=p(x_1,x_2,\cdots,x_n|\mu)=\prod_{i=1}^{n}p(x_i|\mu)$$

设 $\hat{\mu}$ 是参数空间中使似然函数 $l(\mu)$ 最大的值，也就是 μ 的极大似然估计。

$$\hat{\mu}=\arg\max_{\mu} l(\mu)=\arg\max_{\mu}\prod p(x_i|\mu)$$

对极大似然函数取对数再求导，记梯度因子为 $\nabla_\mu=\left(\frac{\partial}{\partial\mu_1},\frac{\partial}{\partial\mu_2}\right)^{\mathrm{T}}$，$\hat{\mu}$ 的值为如下方程的解。

$$\nabla_\mu \ln l(\mu)=\sum\nabla_\mu \ln p(x_i|\mu)=0$$

本文在南京、香港和台湾各部署一台测量服务器，开发了基于 C# 的远程调度程序，实时调度测量服务器对目标主机的 IP 地址进行 traceroute，而后基于固定网络宽带用户的分布计算 LR 的位置，具体方法如算法 1 所示。

算法 1　计算 LR 的位置和半径

输入：$LR[n]=\{LR_1,LR_2,\cdots,LR_n\}$，$LocGroup_i[m]=\{Loc_1,Loc_2,\cdots,Loc_m\}(i=1,2,\cdots,n)$)//LR[n]为 LR 的 IP 地址列表，$LocGroup_i$为第 i 个 LR 服务范围内的固定网络宽带用户的坐标

输出：(LR_point，LR_radius)

```
1. 初始化：InitList(&LR_radius)；InitList(&LR_point)；
   //创建线性表分别存储 LR 的坐标和半径
2. for (i=0;i<ListLength(LR);i++)
   {
     if ListLength(LocGroup_i)>1
     {//列表长度大于 1
       averagePoint(LocGroup_i,i,&LR_point)
       //求 LRi 的坐标，并将结果保存在 LR_point 的第 i 个元素。
       maxRadius(LocGroup_i,i,LR_point,&LR_radius)
         //以 LR 至最远固定网络宽带用户的距离为半径，并将结果
           保存在 LR_radius 的第 i 个元素中。
     }
   }
```

3.3　定位目标主机

LRBG 方法采用与目标主机直连的 LR 的地理位置作为目标主机的位置。然而网络中的少部分路由器由于安全策略配置不响应 traceroute 请求，因此导致 LR 数据缺失。为了解决这一问题，本文将定位目标主机分为 4 种情况进行讨论。

本文对固定宽带用户的 traceroute 路径中的相同路由器和相同链路进行聚合，筛选出频次较高的节点和链路，绘制南

京地区互联网骨干拓扑结构，如图 3 所示。

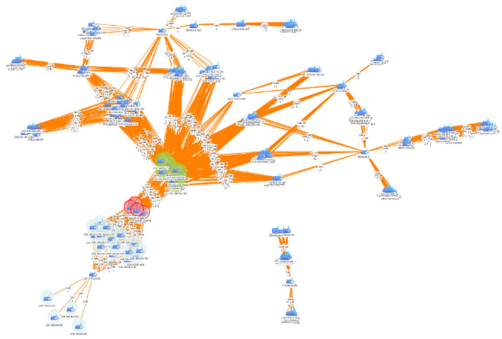


图 3 南京地区部分网络骨干拓扑图

Fig. 3 Topology of some network backbone in Nanjing

测量服务器 traceroute 目标主机的 IP 地址后，提取 traceroute 路径中的 LR，以图 4 中的局部放大图为例，图 5 为图 4 的邻接表存储结构。定位目标主机可分为以下 4 种情形。

(1)traceroute 路径中 LR 存在，且在 LR 数据库中能查询到该 LR 的记录，将该 LR 的地理位置作为待测 IP 的地理位置。

(2)traceroute 路径中 LR 存在，但 LR 数据库中没有记录，如部分路径为“183. *. 19. 146－＞183. *. 25. 205－＞183. *. 222. 1－＞221. *47. 5”，则以数据库中 183. *. 25. 205 的后继节点的位置平均值为待测 IP 的地理位置。

(3)若 traceroute 路径中的 LR 缺失，如部分路径为“183. *. 19. 146－＞183. *. 25. 205－＞ * －＞221. *47. 5”，则搜索 183. *. 25. 205 的后继节点，以后继节点的位置平均值为待测 IP 的地理位置。

(4)traceroute 路径中连续两跳及以上均缺失，如部分路径为“183. *. 19. 146－＞ * －＞ * －＞221. *47. 5”，以 183. *. 19. 146 的两跳后继节点平均值为待测 IP 的地理位置，依次类推。

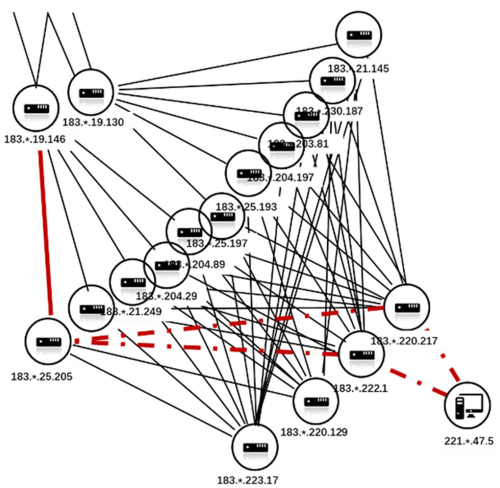


图 4 局部放大图

Fig. 4 Partial enlargement

183.*.19.146	183.*.25.205	183.*.20.141	...	183.*.25.197
183.*.19.130	183.*.25.193	183.*.204.197	...	183.*.21.245
183.*.25.205	183.*.223.17	183.*.220.129	183.*.222.1	183.*.220.217
183.*.21.145	183.*.223.17	183.*.220.129	183.*.221.1	183.*.220.227
...				
183.*.220.217	221.*.47.5			

图 5 邻接表

Fig. 5 Adjacency list

4 实验

4.1 数据集

本文利用自研 APP，在南京地区通过手机实地收集使用了 WiFi 的固定网络宽带用户信息。图 6 给出了南京市建邺区政府附近部分 WiFi 热点在地图上的位置，图 7 给出了收集到的所有固定网络宽带用户在南京市内的分布。由图 7 可知，本文收集到的地标近似覆盖了南京市，可用于全南京市的网络测量和 IP 定位。

图 6 部分 WiFi 热点的位置示意图

Fig. 6 Schematic of some WiFi hotspot locations in Nanjing

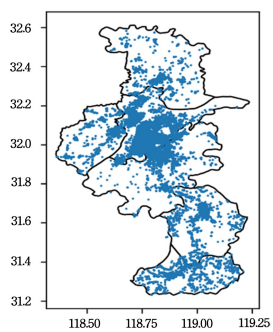
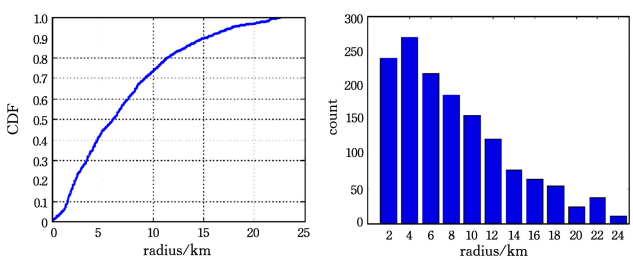


图 7 所有固定网络宽带用户的分布图

Fig. 7 Distribution of all fixed Internet users

4.2 LR 定位

通过 3.2 节中的计算路由器服务范围的方法，利用 LR-BG 方法得出最后一跳路由器 1458 个，其中 90%以上 LR 的计算半径小于等于 15 km。图 8(a)显示了 1458 个路由器的覆盖范围 CDF 函数及分布。图 8(b)显示了路由器计算半径的分布，半径在 2～4 km 和 0～2 km 的路由器的数量最多，分别为 270 个和 240 个。



(a)路由器覆盖范围统计 (b)路由器的半径分布

图 8 LR 的定位结果

Fig. 8 Result of LR geolocation

4.3　IP 定位

SLG 方法是目前已知的精度较高的 IP 定位测量方法，平均精度为 2km。本文共获取 SLG 地标 41 个。SLG 方法将部署在本地的网络服务器作为地标，从网站排行 Alexa 的 547 个南京市域名中爬取“大学”“公司”“政府”类的域名并解析 IP 地址。

采用两种方法获取网站对应的地理位置：1）获取网站页脚提供的通讯地址；2）利用百度地图接口查询公司所在的地址。

可用作地标的网络服务器应满足两个条件：1）服务器有明确的地理位置；2）服务器架设在本地，否则 IP 地址与地理位置不具备对应关系。

在以上 547 个域名中：1）高校类网站中，南京市的 32 所高校和研究机构中有 5 所可用作地标，其余院校包含至少两个校区导致网络服务器地理位置不明确；2）政府类网站中，省级政府网站位于托管机房，南京市各区政府的服务器均不在本地搭建；3）企业网站中，可以通过页面信息或百度地图查找公司地址，网络服务器搭建在本地且有明确地理位置的 36 个。

本文使用 2206 个地理位置已知的固定网络宽带用户的 IP 地址，将 LRBG 方法与 SLG 方法进行实验对比。其中 1748个 IP 地址的 traceroute 路径中包含最后一跳路由器的 IP 地址，其余 458 个通过 3.3 节中的方法，匹配网络拓扑图的前继节点来补充获得。图 9 显示了 LRBG 方法与 SLG 方法的精度对比。

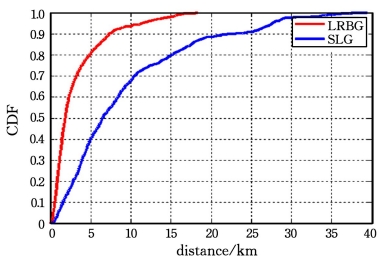


图 9　LRBG 与 SLG 的对比结果

Fig. 9　Comparison results between LRBG and SLG

由图 9 可知，LRBG 方法比 SLG 方法的精度更高且“长尾”较短。LRBG 方法的平均精度为3.17 km，SLG 的平均精度为 9.11 km。该实验的结果说明，最后一跳路由器适用于作为地标进行 IP 定位，LRBG 方法可以较好地适用于街道级 IP 定位的研究。

LRBG 方法比 SLG 方法精度高的主要原因是 LRBG 方法采用的地标比较稳定，地理分布均匀且密度更高。该实验同样说明稳定的地标对于 IP 定位研究有着重要的作用。如本文使用的最后一跳路由器在短时间内位置和 IP 地址不会发生变动，而导致 SLG 地标数量下降的原因可能是：随着云服务的兴起，企业和政府部门出于安全和成本的考虑，更倾向于将网络服务器建在云上，从而造成了本地的服务器数量锐减。

结束语　自 2001 年至今，IP 定位研究不断发展，涌现出了许多优秀的定位方法，地标是 IP 定位研究的关键因素，以往研究大多采用家庭 PC 或服务器作为地标。由于地理分布不均匀和 IP 地址频繁变更等原因，地标成为了制约 IP 定位精度进一步提高的瓶颈。本文提出将探测点至目标主机之间的 traceroute 路径中的最后一跳路由器作为地标，其优势在于：1）最后一跳路由器被分配静态 IP 地址，变更周期较长；2）在地理分布上，运营商在接入互联网的地区均部署了路由器。因此，在选择地标时，路由器可以作为优质地标，为解决 IP 定位问题提供了新的思路。

在确定路由器位置时，基于极大似然估计方法，根据固定宽带用户的地理分布反推最后一跳路由器的经纬度值，实现了路由器的高精度定位，其是现有定位方法的有效补充。同时，对于由于安全策略而不响应 traceroute 的情况，LRBG 方法提出了改进措施，能够有效补充缺失的节点。

实验部分对比了 LRBG 方法和目前已知的精度最高的 SLG 方法，在同等条件下，LRBG 的精度是 3.17 km，且长尾较短。

随着 IPv6 网络的普及和推广，面向 IPv6 的定位应用需求将逐步增加，IP 定位技术将面临新的挑战。下一步，我们将继续完善 LRBG 方法，并将其推广应用于更多城市和 IPv6 网络，推动 IP 定位研究的发展。

参 考 文 献

[1] CIAVARRINI G，GRECO M S，VECCHIO A. Geolocation of Internet hosts：Accuracy limits through Cramér-Rao lower bound [J]. Computer Networks，2018(135)：70-80.

[2] YONG W，BURGENER D，FlORES M，et al. Towards street-level client-independent IP geolocation[C]// Usenix Conference on Networked Systems Design & Implementation. 2016.

[3] WANG Z H，ZHANG W D，WEN H，et al. A Comprehensive Survey of IP Geolocation and Evasion[J]. Journal of Cyber Security，2019，4(3)：34-47.

[4] WANG Z F，FENG J，XING C Y，et al. Research on the IP Geolocation Technology[J]. Journal of Software，2014(7)：1527-1540.

[5] PADMANABHAN V N，SUMBRAMANIAN L. An investigation of geographic mapping techniques for internet hosts[J]. Acm Sigcomm Computer Communication Review，2001，31(4)：173-185.

[6] GUEYE B，ZIVIANI A，CROVELLA M，et al. Constraint-Based Geolocation of Internet Hosts[J]. IEEE/ACM Transactions on Networking，2006，14(6)：1219-1232.

[7] ERIKSSON，BRIAN，BANFORD，et al. A Learning-Based Approach for IP Geolocation[C]// Passive & Active Measurement，International Conference. 2010.

[8] GUO C X，LIU Y X，SHEN W C，et al. Mining the Web and the Internet for Accurate IP Address Geolocations[C]// 28th IEEE International Conference on Computer Communications，Joint Conference of the IEEE Computer and Communications Societies(INFOCOM 2009). 2009：19-25.

[9] LAKHINA A，BYERS J W，CROVELLA M，et al. On the geographic location of Internet resources[J]. IEEE Journal on Selected Areas in Communications，2003，21(6)：934-948.

（下转第 326 页）