|  |  |
| --- | --- |
| **申报编号：** |  |

国家重点研发计划

项目申报书

|  |  |  |  |
| --- | --- | --- | --- |
| 项目名称： |  | | |
|  |  | | |
| 所属专项： |  | | |
| 指南方向： |  | | |
| 推荐单位： |  | | |
|  |  | | |
| 申报单位： |  | （公章） |  |
| 项目负责人： |  | | |

**中华人民共和国科学技术部**

**二Ο一 年 月 日**

**填报说明**

一、填写说明

1、项目申报书分为“国内外现状及趋势分析”、“研究内容及目标”、“申报单位及参与单位研究基础”、“进度安排”、“项目组织实施、保障措施及及风险分析”、“研究团队”、“经费概算”和“指南所要求的附件”八个部分。

申报书的内容将作为项目评审、以及签订任务书的重要依据，申报书的各项填报内容请实事求是、准确完整、层次清晰。

2、请申报单位认真阅读指南，所申报的项目研究内容须对应指南编号、符合指南的要求。

3、申报单位可根据申请指南、申报项目的研究内容自行确定所申报项目的项目名称。项目名称应清晰、准确反映研究内容，项目名称不宜宽泛。

4、通过国家科技管理信息系统按照系统提示在线填写申报书，请按照申报书的基本格式撰写申报书。申报书标题，统一用黑体四号字。申报书正文部分，统一用宋体小四号字填写。正文（包括标题）行距为1.5倍。凡不填写的内容，请用“无”表示。

5、外来语要同时用原文和中文表达，第一次出现的缩略词，须注明全称。

6、申报书中的单位名称，请填写全称，并与单位公章一致。申报书纸质版应与在线填写的电子版版本一致，纸质档需要项目负责人签字，日期如实填写。

填写完成后，请申报单位对所申报信息的真实、完整、有效进行审核。

二、申报说明

申报单位将对申报材料的真实性、完整性负责。

请申报单位审核、确认申报材料后，网上在线提交，并同时将加盖公章一式2份纸质申报材料（含光盘）寄送或提交到指定地点。

所有申报材料要求统一采用胶背方式封装成一册（双面打印），请勿另行添加封页或其他材料。

**项目基本信息表**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 项目名称 | |  | | | | | | | | | |
| 所属专项 | |  | | | | | | | | | |
| 指南方向 | |  | | | | | | | | | |
| 项目类型 | |  | | | | | | | | | |
| 经费概算 | | 总概算 万元，其中申请国家专项经费 万元 | | | | | | | | | |
| 执行周期 | | 年 | | | | | | | | | |
| 申报  单位 | 单位名称 | |  | | | 单位性质 | | | |  | |
| 单位所在地 | |  | | | 组织机构代码 | | | |  | |
| 通信地址 | |  | | | 邮政编码 | | | |  | |
| 推荐  单位 |  | | | | 推荐单位  性质 | | |  | | | |
| 项目  负责  人 | 姓 名 | |  | 性 别 |  | | | 出生日期 | | |  |
| 证件类型 | |  | 证件号码 |  | | | | | | |
| 所在单位 | |  | | | | | | | | |
| 最高学位 | |  | | | | | | | | |
| 职 称 | |  | | | | 职务 | |  | | |
| 电子邮箱 | |  | | 移动电话 | |  | | | | |
| 项目  联系  人 | 姓 名 | |  | 电子邮箱 | |  | | | | | |
| 固定电话 | |  | 移动电话 | |  | | | | | |
| 证件类型 | |  | 证件号码 | |  | | | | | |
| 主要  参加  单位 | |  |  |  |  | | --- | --- | --- | --- | | 序号 | 单位名称 | 单位性质 | 组织机构代码 | |  |  |  |  | | | | | | | | | | | |
| 项目  任务(课题)  分解 | |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | 序号 | 任务(课题)名称 | 任务(课题)  承担单位 | 任务(课题)  负责人 | 建议总经  费(万元) | 建议国拨  经费(万元) | |  |  |  |  |  |  | | | | | | | | | | | |

**申报项目简介**

**第一部分 国内外现状及趋势分析**

1. 国外研究现状及趋势

**1.IP级别的拓扑测量：**

traceroute是接口级拓扑测量中最常用的测量工具，也是Internet接口级别拓扑测量的仅有的原语(primitive)，其他的工具和技术都是traceroute-like的。

最早在1988年，van jacobson实现了ICMP traceroute。在之后的大部分UNIX-like系统发行版本中，traceroute都被实现成了UDP版本。

2001年，Michael Toren用TCP SYN探测包实现了traceroute，一定程度上克服了ICMP，UDP被防火墙过滤的缺陷，然而在网络边缘的主机仍有可能过滤未建立连接的TCP。

针对负载均衡造成的traceroute假路径问题，2006年，法国UPMC的Brice Augustin等发布了Paris traceroute工具，对traceroute进行了改进，通过操纵ICMP的首部，控制探测包在负载均衡存在的情况下只走一条路径。

针对大规模traceroute中观察点的选择问题，2009年，法国Brice Augustin使用了Targeted Probing(定向探测)技术，根据领域知识，精心选择TR的源和目的列表，并辅以LSSR选项，使得TR对IXP的peering的发现率大大提高。[3] 2011年，美国Boston University的B.Eriksson提出了将拓扑测量转化为统计学上的unseen species 问题，然后使用矩阵填充技术来推断未被大规模traceroute发现的设备和链接。[4]

针对大规模traceroute的负载问题，2006年，法国UPMC的Benoit Donnet在文章中提出了far probe技术，在监控点内部使用，并且提出top set技术，在监控点之间使用，降低了他raceroute的冗余和负载。[5]

traceroute大规模拓扑探测最早是从几个观测点出发进行的探测，例如1998年起，法国UoS Pansiot等人进行的研究。[16]

专用大规模拓扑探测平台有CAIDA的Skitter利用web服务器，以及Skitter的后继Arc。[6][7]

此外还有一般意义的分布式测量平台，iPlane[8]，和UW的RocketFuel[17]，都是基于2003年发起的PlanetLab项目的。

除此之外一些公共的Traceroute server，作为拓扑探测的窥镜(looking glass)，由于其测量点的分布覆盖范围广，能够保证较好的完全性，但是由于公共服务器对查询次数的限制，这种方法只适用于小尺度的拓扑测量。

此外还有crowd sourcing测量，不再适用专用的测量平台，而是利用Internet大量的用户基础，通过在浏览器安装插件，让终端用户执行测量工作，如Scriptroute[10]，Dimes[12]，Bitprobe[11]等。这种方法，需要额外的设计权衡，和安全性考虑。

针对traceroute正反向路径的不对称性，2010年，USC的E. Katz-Bassett实现了reverse-traceroute，利用了prespecified timestamp(TS)和 record route options(RR)来识别沿着反向路径的每一跳。[13]

子网发现作为traceroute的扩展，对ip路径上的子网拓扑进行了更近一步的探测。traceroute返回路径中转发路由接口的IP地址，2010年utaxes的M. Tozal 的tracenet可以访问到路由间的子网，并且标注IP地址之间的关系和通过测试获得的子网掩码。类似的工具还有同样是2011年M. Tozal发表的XNET工具。

1. **路由器级别的拓扑测量：**

路由器级别的拓扑测量研究方向可以分为别名解析和渐进路由器发现两种。别名解析是利用数据平面上的Traceroute数据，使用接口合并技术，将同属于一台路由器的ip接口映射到该路由器上。渐进路由器发现利用路由器的MRINFO协议，通过类似OSPF泛洪的方法，通过邻居列表渐进获得网络中路由器的拓扑。

别名解析技术中，又可以分为基于指纹和基于推断两个方向。

2.1基于指纹的接口合并：

2.1.1相同源地址：

法国Jean-Jacques PANSIOT，在1995年开始通过主动探测得到Internet的多播树拓扑，使用的探测手段是类traceroute，首先以自己主机作为源地址，向自己数据库(能够通信，支持源路由选项，能反向DNS查询)的5000台主机进行Traceroute，再从5000台主机采样11台，分别作为源地址再次进行Traceroute。得到的测量数据，在使用相同源地址方法进行别名解析，向可能是别名的ip地址对发送两个探测包，检测到返回包中的源地址发生变化，则认为两个地址是别名。[16]

在2000年，CAIDA的Ken Keys 实现了iffinder别名解析工具，采用了相同源地址方法。这种方法的局限性在于，Traceroute的回复受制于路由器的应答策略。

2.1.2 IPID计数器：

2002年华盛顿大学的Neil Spring提出了RocketFuel项目，作为ISP拓扑测量的引擎。使用了别名解析工具Ally(现在已经不再维护，而是作为Scriptroutede 一部分)。Ally所用到的原理就是IPID计数器。[17]原理是向两个候选别名ip分别发探测包，根据回复包的IPID差值，若小于一定阈值，在向先回复的ip发送第三个包，如果三个包的回复IPID满足升序，且IPID差很小，则认为两个ip是别名。

为了降低时间复杂度O(n^2)，Ally依据别名应当有相近TTL的原理，对TTL进行聚类，并使用逆向DNS查询进行分类，实现对候选别名集合的剪枝。

局限性，有些IPID计数器增加幅度不同，导致false negative。时间复杂度，不剪枝会很大O(n^2)，剪枝会不准确。[20]

2008年美国马里兰大学的Adam Bander，提出了RadarGun工具，改进了Ally工具，不再像Ally一样逐对探测别名，而是逐个多次探测ip列表，找到ID增速都一致的ip认为是别名。RadarGun，把复杂度降到了O(n)，并且解决了false negative问题。然而，如果对ip的探测间隔控制不好，ID会发生wrap around(折回)，导致新的错误判断。[19]

到2016年CAIDA使用的基于IPID的别名解析工具是MIDAR (Monotonic ID based Alias Resolution)，与Ally，RadarGun相比，MIDAR准确性更高，而且能够在大规模网络测量中实现产品化。在CAIDA的Macroscopic Internet Topology Data Kit (ITDK)项目中，MIDAR和iffinder共同使用作为别名解析工具。

主动探测的方法缺点，首先需要引入探测流量，其次主动探测不能控制中间路由器的行为，路由器可能故意不回复，ip实现可能不使用全局IPID计数器(为了隐藏自己的网络结构或者为了安全性考虑)。

2.1.3 IP选项：

IP选项别名解析的思路是，使用ip包首部中的ip选项，结合traceruote，获得路径的更多信息用来推断别名。

2008年UW的Neil Spring等提出了一种利用了RR选项(Record Route option)的别名解析工具DisCarte，使用析取逻辑规划的模型，将RR数据和TR数据对齐。2012年又将DisCarte整合到了Sidecar中。RR是ip首部的一个数组，路径上的路由器会在RR中插入自己的ip地址，协议设计最多能存储9个ip地址。RR选项可以克服TR的一些问题，发现不回复的路由器别名，隐藏路由器(不递减TTL)，匿名路由器(递减TTL但不回复icmp echo)。缺点是，最多支持9跳，RR包可能被过滤或丢弃，还有最重要的是，TR数据和RR数据对齐的复杂性，导致这种方法只适用于小范围的拓扑测量。[23]

此外，TS选项(pre-specified timestamp)也可以被用作别名解析。发包者在TS选项中指定4个ip，要求对应ip的设备给选项按照给出的ip顺序(即ip在前的写完后面的设备才能写)写入时间戳。将候选ip别名交叉写入TS如ABAB，BABA。回复中如果，将4个ip都写入了时间戳，并且时间戳数值相同(一个路由器只有单个时钟)，则认为AB互为别名。[24]

2.2基于推断的接口合并：

2.2.1基于DNS-name：

相近的DNS-name认为是别名。单独使用缺陷很明显：需要ISP对路由器进行系统的命名，需要认为干预，在AS边界非常不准确。[22]

DNS-name在RocketFuel中，主要是用来路由器标识和地理位置映射，DNS标识相对于AS地址空间标识的优势是，AS地址空间对于不进行BGP广播的邻居而言，只用自己内部的地址编号，这样边界不容易被发现，此外同属于两个AS的路由器可能被编号成任意一个AS中的地址，还有DNS能区分出非路由器设备，避免AS造成的混淆。标识路由器时，利用DNS逆向查询得到的DNS名字，使用undns工具从中提取ISP标识路由器的字段。地理位置映射方面，利用ISP植入到DNS名字中的信息来推断路由器的地理位置，从而更好的识别POP和骨干，深入了解ISP的结构特征。[17]

这种标识路由器的局限性：准确度会受到DNS数据更新不及时的影响，完整度会由于ISP发布的DNS数据不完整而受到影响。[20]

2.2.2基于图的启发法：

基于两条规则，对TR的有向图进行推断。

共同后继：适用于路由器点对点连接的情况(相对于多点接入网络和交换结构网络)，由于后继接口标识了唯一一个路由器，所以拥有共同后继的接口可以认为是别名。

相同TR：相同TR中的所有接口都不是别名，因此一个具有n个节点的TR路径可以消除n^2个别名对可能。

基于图的方法只是用来做预处理用，减少下一步主动探测的数量。[20]

2.2.3基于子网的分析法：

2006年M.H.Gunes提出一种基于子网的别名解析算法。思路是先用双向TR数据建立图，根据ip中子网划分的机制，找出分别在两个路径上的，且在同子网中的接口对，二者之间的连接认为是点对点连接，对于点对点连接之间的接口认为是别名。[21]

2.3渐进路由器发现：

渐进路由器发现方法利用了支持MRINFO协议的路由器保存有邻居路由器列表，通过类似于OSPF的渐进路由器发现方法，逐步发现网络中所有路由器。

2004年，法国Jean-Jacques PANSIOT使用自主开发的基于mrinfo协议的mrinfo-rec工具，从单一路由器开始探测，直到08年结束。平均每日探测10万接口（不包括无法路由到的IP地址，特殊IP地址，隧道，无效接口等），发现1万左右路由器，超过850个AS。mrinfo-rec的局限性在于，适用范围只限制于支持多播的路由器，ICMP报文可能被过滤。[25]

2011年，PANSIOT针对对mrinfo-rec的问题做了改进，并发布了新的MERLIN工具。

1. **POP级别的拓扑测量：**

截至目前，对于POP级拓扑的研究有三种主要的方式，最常用也是最基础的方法是基于traceroute工具设计并实现的，通过traceroute可以获取域际路由跳转信息，进一步能够掌握接口的IP信息，拓扑图也根据地理上的聚类工作进行绘制，代表性的研究如Rocketfuel[17]、iPlane[8]、Tian等人的工作[26]、Feldman等人的工作.[27]。第二种方法完全绕开traceroute，使用ping工具基于ICMP回应延迟来判断当前访问的节点与本地的距离，Yoshida et al.[28]拥有遍布日本的服务器，通过一定数量的机器收集的信息可以绘制出网络拓扑图。第三种方法依赖于不同的网络供应商官方提供的数据信息，也就是静态或动态的基站、路由分布图，制作精准的POP级网络拓扑图，Topology Zoo[8]、Internet Atlas[9]项目做了相关工作。

1. **网络扫描技术：**

网络扫描是指探测局域或者互联网范围服务，搜索漏洞或渗透IT设备方法的行为。国外在该领域研究较早研究水平较高，新的研究成果层出不穷。在网络扫描技术方面从简单的全连接扫描，到半连接扫描再到秘密扫描，从单源扫描到分布式扫描[31]。文献[34]介绍了利用IP报头的时间戳选项进行主动探测的技术可行性评估研究，在扫描工具方面有业内常用的功能强大的nmap[32],也有最近出现的以快著称45分钟扫描整个IPv4空间的zmap[33]，该工具通过对传统扫描技术的发包速度和状态保存等多个优化使其速度达到nmap的1300多倍。

1. **IP地理定位技术：**

IP定位技术：代表性探究机构与项目包括：（1）、CAIDA在美国国土安全的资助下，对各种IP定位算法进行了分析[35]；（2）、IETF成立了专门的工作组来讨论internet定位技术的标准、隐私保护等相关问题并提出了相应的草案[36]；（3）、W3C在下一代网络标准中将定位服务作为一个标准接口写入了HTML5规范[37]等。

主要技术进展：（1）基于推测的定位算法：一般通过查询whois数据库等来过去该IP地址的主机名，所在街道或者通过IP地址段地理位置来推测该IP设备的位置。典型算法有IP2LL, NetGeo[38], GeoTrack[39], VisualRoute, Neotrace, GTrace[40], Quova, MaxMind, EdgeScape, TraceWare, SoftWare77, IPligence, HostIP, IPInfoDB等。

1. 基于时延的定位算法：通过测量目标主机到测量点的时延来估测主机位置,为了提高定位精度,往往结合网络拓扑信息来进行定位,典型的定位算法和系统有Geoping, Shortest Ping[42]，Constraint-based Geolocation（CBG)[41], Topology-based Geolocation (TBG)[42], NBIGA[43] ,Posit, Geo-RX, Spotter[44] 等；
2. 综合定位，当基于推测、时延等算法不能单独地对主机进行准确定位,为此,一些研究人员提出了综合的定位方法。典型的算法有Octant[45], TTG[46], GeoWeight 这些定位算法同时运用上述2~3 种算法对待定位主机进行定位,然后进行交互验证以确定主机位置

[1] M. C. Toren, tcptraceroute: an implementation of traceroute using TCP SYN packets

[2] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and

R. Teixeira, Avoiding Traceroute Anomalies with Paris Traceroute, in Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. ACM, 2006, pp. 153-158.

[3] B. Augustin, B. Krishnamurthy, and W. Willinger, “IXPs: Mapped?” in Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference. ACM, 2009, pp. 336–349.

[4] B. Eriksson, P. Barford, J. Sommers, and R. Nowak, DomainImpute: Inferring unseen components in the Internet, in INFOCOM, 2011 Proceedings IEEE, pp, 171-175.

[5] B. Donnet, P. Raoult, T. Friedman, and M. Crovella, Deployment of an algorithm for large-scale topology discovery, Selected Areas in Communications, IEEE Journal on, vol. 24, no. 12, pp. 2210-2220, 2006.

[6] CAIDA, Macroscopic Topology Measurements Project and the Skitter infrastructure, http://www.caida.org/tools/measurement/skitter/.

[7] CAIDA, Macroscopic topology measurements project and the archipelago measurement

infrastructure, http://www.caida.org/projects/ark/, 2011.

[8] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and

A. Venkataramani, iPlane: An information plane for distributed services, in Proceedings of the 7th symposium on Operating systems design and implementation. USENIX Association, 2006, pp. 367-380.

[9] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy, Lord of the links: a framework

for discovering missing links in the Internet topology, IEEE/ACM Transactions on Networking

(ToN), vol. 17, no. 2, pp. 391-404, 2009.

[10] N. T. Spring, D. Wetherall, and T. E. Anderson, Scriptroute: A Public Internet Measurement Facility, in USENIX Symposium on Internet Technologies and Systems, 2003.

[11] T. Isdal, M. Piatek, A. Krishnamurthy, and T. Anderson, Leveraging BitTorrent for end host measurements, in Passive and Active Network Measurement. Springer, 2007, pp. 32-41.

[12] Y. Shavitt and E. Shir, DIMES: Let the Internet measure itself, ACM SIGCOMM Computer

Communication Review, vol. 35, no. 5, pp. 71-74, 2005.

[13] E. Katz-Bassett, H. V. Madhyastha, V. K. Adhikari, C. Scott, J. Sherry, P. Van Wesep, T. E. Anderson, and A. Krishnamurthy, Reverse traceroute, in NSDI, vol. 10, 2010, pp. 219-234.

[17] M. Tozal and K. Sarac, Tracenet: an Internet topology data collector, in Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM, 2010, pp. 356-368.

[15] P. Marchetta, V. Persico, E. Katz-Bassett, and A. Pescap´e, “Dont trust traceroute (completely),” in ACM CoNEXT Student workshop, 2013

[16] J.-J. Pansiot and D. Grad, On routes and multicast trees in the Internet, ACM SIGCOMM Computer Communication Review, vol. 28, no. 1, pp. 41-50, 1998.

[17]N. Spring, R. Mahajan, and D. Wetherall Measuring ISP topologies with Rocketfuel, ACM SIGCOMM Computer Communication Review, vol. 32, no. 4, pp. 133-145, 2002.

[18] W. Chen, Y. Huang, B. F. Ribeiro, K. Suh, H. Zhang, E. d. S. e Silva, J. Kurose, and D. Towsley, Exploiting the IPID field to infer network path and end-system characteristics," in Passive and Active Network Measurement. Springer, 2005, pp. 108-120.

[19] A. Bender, R. Sherwood, and N. Spring, Fixing ally's growing pains with velocity modeling, in Proceedings of the 8th ACM SIGCOMM conference on Internet measurement. ACM, 2008, pp. 337-342.

[20] N. Spring, M. Dontcheva, M. Rodrig, and D. Wetherall, How to resolve IP aliases, Univ. Michigan, UW CSE Tech. Rep, pp. 04-05, 2004.

[21] M. H. Gunes and K. Sarac, Analytical IP alias resolution, in Communications, 2006. ICC'06. IEEE International Conference on, vol. 1. IEEE, 2006, pp. 459-464.

[22] R. Govindan and H. Tangmunarunkit, Heuristics for Internet Map Discovery, in INFOCOM 2000, vol. 3. IEEE, 2000, pp. 1371-1380.

[23] R. Sherwood, A. Bender, and N. Spring, Discarte: a disjunctive Internet cartographer," in ACM SIGCOMM Computer Communication Review, vol. 38, no. 4. ACM, 2008, pp. 303-314.

[24] W. de Donato, P. Marchetta, and A. Pescape, A hands-on look at active probing using the IP prespeci ed timestamp option, in Passive and Active Measurement. Springer, 2012, pp. 189-199.

[25] J.-J. Pansiot, P. M´ erindol, B. Donnet, and O. Bonaventure, “Extracting intra-domain topology from mrinfo probing,” in Passive and Active Measurement. Springer, 2010, pp. 81–90.

[26] Y. Tian, R. Dey, Y. Liu, and K. W. Ross, “China's Internet: Topology mapping and geolocating," in INFOCOM, 2012 Proceedings IEEE. IEEE, 2012, pp. 2531-2535.

[27] D. Feldman ,Y. Shavitt,N. Zilberman, “A structural approach for PoP geo-location," in INFOCOM IEEE Conference on Computer Communications Workshops, 2010. IEEE, 2010, pp. 1-6.

[28] K. Yoshida, Y. Kikuchi, M. Yamamoto, Y. Fujii, K. Nagami, I. Nakagawa, and H. Esaki, “Inferring PoP-level ISP topology through end-to-end delay measurement," in Passive and Active Network Measurement. Springer, 2009, pp. 35-44.

[29] S. Knight, H. X. Nguyen, N. Falkner, R. Bowden,and M. Roughan,”The Internet topology zoo," Selected Areas in Communications, IEEE Journal on, vol. 29, no. 9, pp. 1765-1775, 2011.

[30] R. Durairajan, S. Ghosh, X. Tang, P. Barford, and B. Eriksson,”Internet Atlas: A Geographic Database of the Internet," 2013.

[31] Bou-Harb, E. et al. 2014. Cyber scanning: A comprehensive survey. IEEE Communications Surveys and Tutorials. 16, 3 (2014), 1496–1519.

[32] Lyon G F. Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning[M]. Insecure, 2009.

[33]Durumeric Z, Wustrow E, Halderman J A. ZMap: fast internet-wide scanning and its security applications[C]// Usenix Conference on Security. USENIX Association, 2013:605-620.

[34] W. de Donato, P. Marchetta, and A. Pescap´e, “A hands-on look at active probing using the IP prespecified timestamp option,” in Passive and Active Measurement. Springer, 2012, pp. 189–199.

[35]CAIDA.CAIDA’s geolocation tools comparison.2010.http://www.caida.org/projects/cybersecurity/geolocation/

[36]IETF Geopriv Workgroup.draft-ietf-geopriv-uncertainty.2014.http://tools.ietf.org/wg/geopribv/

[37]Robin B,Travis L,Silvia P,Edward O.Html5 draft.2012.http://dev.w3.org/html5/spec-author-view/

[38]Moore D, Periakaruppan R, Donohoe J, Claffy K. Where is the world is netgeo.caida.org? In: Proc. of the INET 2000. 2000.<http://www.caida.org/publications/papers/2000/inet_netgeo/inet_netgeo.html>

[39]Padmanabhan VN, Subramanian L. An investigation of geographic mapping techniques for Internet hosts. In: Proc. of theACM SIGCOMM. San Diego: ACM Press, 2001. 173−185. [doi: 10.1145/964723.383073]

[40]Periakaruppan R, Nemeth E. Gtrac A Graphical Traceroute Tool. Usenix LISA, 1999.

[41]Gueye B, Ziviani A, Crovella M, Fdida S. Constraint-Based geolocation of Internet hosts. ACM/IEEE Trans. on Networking, 2006,14(6):1219−1232. [doi: 10.1109/TNET.2006.886332]

[42]Katz-Bassett E, John J, Krishnamurthy A, Wetherall D, Anderson T, Chawathe Y. Towards IP geolocation using delay and topology measurements. In: Proc. of the IMC 2006. Rio de Janeiro: ACM Press, 2006. 71−84. http://homes.cs.washington.edu/~arvind/papers/geoloc.pdf [doi: 10.1145/1177080.1177090]

[43]Eriksson B, Barford P, Sommersy J, Nowak R. A learning-based approach for IP geolocation. In: Proc. of the PAM. 2010. [doi:10.1007/978-3-642-12334-4\_18]

[44]Laki S, Mátray P, Hága P, Sebök T, Csabai I, Vattay G. Spotter: A model based active geolocation service. In: Proc. of the

INFOCOM. Shanghai: IEEE Press, 2011. 3173−3181. [doi: 10.1109/INFCOM.2011.5935165]

[45]Wong B, Stoyanov I, Sirer EG. Octant: A comprehensive framework for the geolocalization of Internet hosts. In: Proc. of the NSDI2007 Symp. Cambridge: ACM Press, 2007. 313−326. <http://www.cs.cornell.edu/People/egs/papers/octant-nsdi.pdf>

[46]Wang Y, Burgener D, Flores M, Kuzmanovic A, Huang C. Towards street-level client-independent IP geolocation. In: Proc. of the8th USENIX Conf. on Networked Systems Design and Implementation. Berkeley, 2011. https://www.usenix.org/legacy/event/nsdi11/tech/full\_papers/Wang\_Yong.pdf

1. 国内研究现状及趋势
2. IP级别和路由器级别拓扑测量：

（包括国内从事该方面研究的主要机构及其研究方法、研究水平；国内外研究方法、研究进展的对比分析；该方面研究可能取得的突破及预期成果。限2000字以内。）

国内研究现状：

2007年，哈尔滨工业大学[1]以建立中国IP级网络拓扑图景为目标,首先评价了传统的目标抽样法,提出了3项改进技术,并测量了中国网络,然后运用复杂网络研究方法提取并比较分析了拓扑特征,最后采用IP2AS技术统计了自治域规模。

2009年，哈尔滨工业大学[2]开发了fastrace，其中所采用的新技术可在保证完整性的同时将探测负载降低一半。采用了多项当时领先的测量技术，包括单进程多任务探测引擎，多类型包混合穿透技术，多类型包标记，子网二分探测，反向traceroute等等。从安全中心提供12个测量点对全国互联网IP级拓扑实施测量，发现81万节点，137万链接，规模是同期国际测量项目skitter的5倍。

[1] 张宇，方滨兴，张宏莉. 中国IP级网络拓扑测量与分析. 通信学报. 2007

[2] 张宇，大规模网络拓扑测量与分析研究，计算机系统结构， 2009

1. **POP级别拓扑测量：**

中国最大的两个网络供应商是中国移动和中国联通，他们拥有高性能的国家主干网络，连通着全国各地各区域，在各个省、直辖市为终端用户提供服务，此外中国境内第三大网络供应商是CERNET，主要侧重于国际商业、学术网络服务的提供。由于中国境内没有公共的可利用的BGP路由，所以对于IP地址的搜寻工作只能通过国家周边的路由获得。此外中国的网络现状缺乏进行大规模扫描、测量的条件。例如中国境内对PlanetLab项目的支持不够，节点很少，以至于不能进行大规模的测量工作，另外路由接口大多是未知的，地理位置也不能推断。[26]

**3.AS级别拓扑测量：**

2011年，哈尔滨工业大学[6]，提出了一种系统框架，用来量化traceroute测量AS级别拓扑时的纰漏，并且得出了AS级别纰漏的原因主要不是因为traceroute的问题而是因为路由器使用了peering关系的邻居的ip地址。

[6] Y. Zhang, R. Oliveira, Y. Wang, S. Su, B. Zhang, J. Bi, H. Zhang, and L. Zhang, A framework to quantify the pitfalls of using traceroute in AS-level

topology measurement, Selected Areas in Communications, IEEE Journal on, vol. 29, no. 9, pp. 1822-1836, 2011.

**5.IP地理定位：**

网络犯罪侦察的IP定位技术

国内IP定位技术:1.北京邮电大学计算机学院王婷等人， 针对IP 定位数据库如何及时更新IP 定位数据并同时保证其定位准确率和覆盖率的问题，提出一种基于关联规则挖掘的IP 定位方法IPGEL。IPGEL 方法在提高IP 定位准确率和覆盖率的同时有效提高了已有IP 定位数据库的可信度[1]．

中国矿业大学计算机科学与技术学院闫子骥等人，为实现快速精确的IP 定位，提出了基于SNMP 协议，并结合哈希查找技术，通过对相关网络设备ARP 及MAC 地址表的一系列计算的定位方法 。通过实验表明，该方法相比人工定位及全网搜索的IP定位方法，执行效率更高，定位更精确[2].

北京邮电大学计算机学院宋建等人，通过分析国内主流的IP 地址库，提出了一种基于IP 地址库之间差异对比来评估其可信程度的方法。该成果为找出评估IP 地址库的可信度、提高通过IP 地址库进行IP 定位的准确度提供了可参考的方法[3]。

[1]王婷，宋俊德，宋美娜． 一种基于关联规则挖掘的IP 定位方法［J］． 东南大学学报: 自然科学版，2015， 45( 4) : 657-662． ［doi: 10．3969 /j． issn． 1001 － 0505． 2015． 04． 008］

[2] 闫子骥， 安计勇.一种高效的定位方法。安徽电子信息职业技术学报，[doi:1671－802X(2014)03－0010－05]

[3] 宋建,许可,宋美娜. 一种评估国内IP地址库可信度的方法

**第二部分 研究内容及目标**

一、项目目标及考核指标

（一）申报项目与所属指南方向的关联关系

（包括项目与所属指南方向的匹配性，对指南方向目标的支撑作用。限500字以内。）

（二）**项目目标及考核指标、评测方式/方法**

**IP级别和路由器级别拓扑测量模块指标、评测方法：**

通过收集系统运行的结果得到拓扑数据集，将结果与现有测量工具得到的数据集进行交叉验证，和iffinder等工具得到的数据集、IRR提供的数据集，比较路径和节点的发现率。和CAIDA的ally，MIDAR得到的Arc数据集进行交叉验证以评估系统方法测量结果的完整性，准确性。

此外，使用相同的运行环境，对工具的性能进行测试，评估系统测量的效率。

**POP级别拓扑测量模块指标、评测方法：**

通过统计POP级拓扑测量的结果，可以获得全球POP数量以及分布情况，与先前的多个项目结果作对比，即可判断准确性。可视化处理能够筛选显示区域，显示自治域内POP的互联情况，进一步显示详细的地理信息。

**AS级别拓扑测量模块指标、评测方法：**

AS级拓扑测量的成果将以图形用户界面的形式呈现出来。用户可以在图形用户界面上设置监控点和采集点的控制参数，包括采集点的采集时间、采集频率等信息。同时，用户可以观看AS级拓扑结构图，并与之交互获得AS的详细信息(AS号、IP前缀和地理位置等信息)和AS间连接的详细信息(连接两侧的AS、连接的商业关系等信息)。而且，用户能够在当前AS拓扑结构中发生异常时得到警报，并能够在拓扑图中观察到发生异常的AS信息。

考核指标：最终获得的BGP路由信息与部分已知AS间的拓扑结构达到基本吻合。

评测方式：使用采集点从部分已知的AS边界路由获得路由信息，并与采集系统处理后的BGP路由信息进行一致性比较，评估系统测量结果的完整性和准确性。

**网络扫描模块指标、评测方法：**

该项目的研究目标是在分析现有的网络扫描技术的基础上，针对大规模网络扫描，提出一种新的大规模网络快速扫描系统，并加以实现。做到功能与速度的结合，在扫描覆盖率与准确率相差不大的情况下，大幅提高扫描的速度。

评估方法为：使用相同指定配置的主机或服务器，在相同的网络环境下，对相同数目的目标集使用nmap等现有扫描工具和该项目成果进行多种扫描方式的扫描，记录实验结果中的覆盖率，正确率和完成所耗时间，并进行比较。

考核指标为：在上述实验评估中，在覆盖率和正确率相差不超过5%的情况下，扫描速度与其它工具相比提高3倍以上。

**IP地理定位模块指标、评测方法：**

IP定位模块最终达到的指标:

已知IP地址，可以得到其相对应的地理地址，获得一张可以极大程度的映射IP地址与地理位置的地图。

测评方式分如下几方面：

算法复杂度:复杂度是衡量算法计算开销的重要指标,开销越小,则算法速度越快,越易实现和部署

定位精度：定位精度是衡量定位算法优劣的最重要的指标,不同粒度的定位精度,决定了算法的应用范围;

主动测量：主动测量会给网络产生扰动,从而影响定位精度;而被动测量或查询的方式则对网络影响较小;

扩展性：主要从定位系统是否存在中心化设施、部署成本以及是否容易增量部署几个方面进行衡量.分为差、中、较好3 个层次;

部署方式：在定位过程中,算法的设计确定了定位系统的部署方式.采用集中式部署借助于服务器能够提高系统的收敛速度,但是不易于扩展;而分布式系统则无需基础设施的支持,具有较好的扩展性.

（三）项目成果的呈现形式及描述

**IP级别和路由器级别拓扑测量模块：**

作为一个测量工具开源发表，并部署到PlanetLab，国家安全中心的监测点，进行长时间大规模拓扑测量，并进行深度路径分析。测量和分析得到的数据将周期性发布到ftp服务器以供下载。

**POP级别拓扑测量模块：**

本模块意在全国部署多个节点，并且结合PlanetLab所提供节点进行全球的POP级网络拓扑测量，并且将所获得的数据进行可视化处理，在地图中显示自治域内POP分布情况，以及出入口IP地址，进一步结合地理信息系统，显示POP区域内的详细地理信息。

**AS级别拓扑测量模块：**

AS级拓扑测量的成果将以图形用户界面的形式呈现出来。用户可以在图形用户界面上设置监控点和采集点的控制参数，包括采集点的采集时间、采集频率等信息。同时，用户可以观看AS级拓扑结构图，并与之交互获得AS的详细信息(AS号、IP前缀和地理位置等信息)和AS间连接的详细信息(连接两侧的AS、连接的商业关系等信息)。同时，用户能够在当前AS拓扑结构中发生异常时得到警报，并能够在拓扑图中观察到发生异常的AS信息。

**网络扫描模块：**

项目成果以Linux C程序编写的软件形式呈现，在Linux环境下运行，使用精简高效的命令行形式进行操作，软件模块划分清晰，功能完备，提供丰富的功能命令集。针对不同的扫描方式设计多种扫描模块。提供多种输出格式，满足不同的结果形式要求。

**IP地理定位模块：**

在定位过程中，提出一个IP地址，可以在一定反应时间内得到相应的地理位置，定位精度在一定的范围内。我们可以拥有一张覆盖面相对较大，位置较准确的IP地址与地理地址映射地图。

二、项目研究内容、研究方法及技术路线

（一）项目的主要研究内容

**1.IP级别和路由器级别拓扑测量模块：**

虽然现有的拓扑发现技术和工具都已经比较成熟，但是在这些工具和数据集的使用上会存在许多问题。对于这种Internet大数据范畴的问题，测量的好坏不单纯由大数据分析数量决定，比如在使用BGP路由数据测量AS级拓扑时，对于测量的不完全性是不能单一用增加测量数量来解决的。此外，还需要领域知识和工程经验的应用，比如traceroute测量接口级拓扑结构时，人工选择的观察点会提高发现率。除此之外，还要有对测量工具和数据集合的限制的明确认识，盲目的使用，不与研究目的契合，不可能得出正确的研究结果。

此外，缺乏一种深度路径监测平台，对网络尤其是特定国家和地区网络的拓扑进行深度监测的平台，通过多点、实时、主动、定向的监测来观察特定服务所依赖的网络路由状态，发现并分析导致服务性能不稳定甚至瘫痪的根源，及时准确地向应用服务商提出告警通知，并为应急处理与故障恢复提供完整可信的数据支撑。缺乏一个便于网络管理员，或者普通用户使用的可视化监控界面。

本项目在现有的ip接口级的测量工具和技术的基础上，对测量数据进行深度路径分析

最终要实现的功能是，首先，将集成了多种接口级别测量技术的类traceroute工具部署到PlanetLab，国家安全中心的监测点，进行长时间大规模拓扑测量，并将测量结果与CAIDA最新的数据集进行交叉验证，比较测量得到的拓扑数据的完整性和准确性。

其次，将ip级别的拓扑作为输入，进行深度路径分析，执行接口合并得到路由器级别的拓扑，并进行子网发现，获得路径经过的子网的更详细的拓扑信息。

**2.POP级别拓扑测量模块：**

“Point of Presence”简称POP，直译为网络接入点。在POP级开展网络拓扑研究具有一定的优势。每一个POP都是若干个路由的聚集代表[1]，这些路由分布于同一区域。通过POP的互联，构建了AS的主干网络[2]，正因如此，网络供应商和其客户才能够建立通信。在这一层级中，一个节点便是一个归属于AS的POP，每一个节点都对应着其拥有者的地理信息，两个节点之间的链路代表了两个区域中的路由有一条物理连接。网络供应商所提及的POP一般指代一座真实的大楼，楼中部署了很多路由，或者是一个大型的都市，这也充分体现了自治域中的分层设计原则，这样的设计有助于增加网络的可维护性。目前获取基本数据的方法有三中：traceroute、ping、网络供应商官方数据。先前的研究者们通过三种不同的思路，分别进行了POP级拓扑研究工作，各有优缺点，绘制出的POP拓扑图准确性参差不齐。本项目旨在通过对已有工具的算法进行改进，以至于获得更为精确的互联网中POP拓扑图，

1. **AS级别拓扑测量模块：**

AS级拓扑模块：该模块用于从不同的数据源收集BGP路由信息，推断AS间的商业关系，并推断出AS级拓扑的层次结构。同时，分析BGP路由信息发现网络中的性能与安全问题。

模块主要实现一下功能：

（1）定期从BGP踪迹收集器、路由服务器、窥镜和因特网选路注册机构中采集BGP路由信息，使用分析技术对得到的BGP路由信息进行处理，排除和纠正其中一些不一致的路由信息，将最终的BGP路由信息保存。

（2）根据最终的BGP路由信息推断AS间的商业关系，并通过商业关系推断AS级拓扑的层次结构，并进行可视化显示。

（3）对BGP路由信息进行安全分析，在发生性能和安全问题时发出警报，并在层次结构图中的相应位置标注警报信息。

**4.网络扫描模块：**

该项目主要研究内容是在分析现有的网络扫描技术的基础上，针对大规模网络扫描，提出一种新的大规模网络快速扫描系统，并加以实现。现有的扫描技术有：全连接扫描（Open Scan），半连接扫描（Half-Open Scan），隐秘扫描（Stealth Scans），存活性扫描（Sweep Scans）等。现有的网络扫描工具基本都使用了上述的扫描方法，而且也存在一些不足。Namp的扫描功能比较强大，可以实现端口探测，版本检测和系统检测等，但是在大规模网络应用上存在速度较慢的问题；而近年出现的zmap扫描速度有了很大提升，但是其主要针对的是单端口扫描，功能相对较少。将二者优势结合将具有巨大意义。

**5.IP地理定位模块：**

IP定位模块：该模块应用IP定位技术，通过设备的IP地址来确定其地理位置。

主要功能如下：

（1）当已知网络设备或者网络用户的IP地址时，可以得到一定精度范围内的相应的地理位置。

（2）网络安全应用：通过用户的位置信息对其身份进行验证，进行入侵检测，通过将用户的IP地址与地理地址相关联，减少IP哄骗攻击，在用户访问控制中，根据分析用户身份的真实性并依据访问控制策略，响应或拒绝用户的通信请求。

（3)社交网络：通过将社交网络成员的位置或运动轨迹与他们的行动、兴趣进行关联，分析用户的信息等。

达成目标:将IP地址与地理位置一一映射，对有相应威胁的IP地址进行监测，对IP地址进行画像。

（二）项目拟采取的研究方法

1、项目研究拟解决的问题，及拟采用的方法、原理、机理、算法、模型等

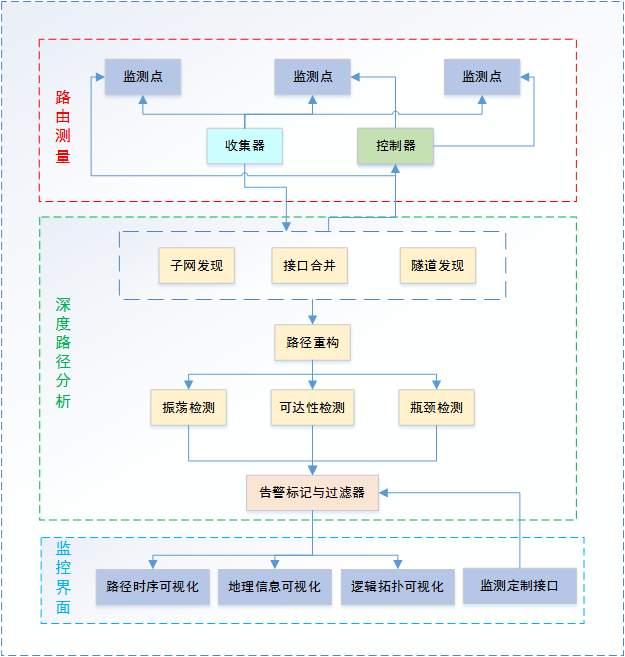
（研究方法、技术路线可用图表描述。限800字以内）

**1.1 ip拓扑测量模块：**

ip拓扑测量模块解决的问题是，对互联网ip和路由器级别的拓扑进行测绘(mapping)。

ip路径测量使用类traceroute技术，利用icmp的ttl跳数限制机制。在此基础上，操纵ip选项字段和icmp首部，增加对探测包经过路径的控制，记录路径的更多信息，并克服负载均衡带来的困难。

ip接口合并，通过综合探测法和推断法，使用推断法对ip路径进行预处理，得到候选别名对，对于存疑的别名对，发送精心构造好的主动探测包(设置了RR选项和TS选项)，根据包的回复，使用相同源地址，IPID计数器等路由器的指纹，对是否是别名进行有根据的判断(相同源地址方法的原理是向可能是别名的ip地址对发送两个探测包，检测到返回包中的源地址发生变化，则认为两个地址是别名)（IPID计数器方法的原理是向两个候选别名ip分别发探测包，根据回复包的IPID差值，若小于一定阈值，在向先回复的ip发送第三个包，如果三个包的回复IPID满足升序，且IPID差很小，则认为两个ip是别名）。

****

IP拓扑测量模块可以划分成如下三个子模块：

路由测量模块，深度路径分析模块，监控界面模块。

路由测量模块由监测点，收集器，控制器三个子模块组成。监测点部署了主动IP路径探测程序的机器，将测量结果提交到收集器。收集器收集多监测点测量结果并进行初步过滤与汇总。控制器对监测点实施控制，包括测量任务下达、状态监测等。

深度路径分析模块又可以拆分成路径分析，路径重构，问题检测，过滤告警四个子模块。路径分析包括：子网发现、接口合并与隧道发现，根据测量数据进行子网发现、接口合并与隧道发现等测量分析工作，必要时按需通过控制器执行新测量任务。路径重构基于上述测量数据，对通往目标的路径进行重构，力图推测IP网络实际拓扑与路径结构。问题检测模块包括：振荡检测、可达性检测与瓶颈检测，基于重构后路径数据，对路径频繁振荡、不可达黑洞以及瓶颈链路进行检测。

过滤告警模块包含告警标记与过滤器，路径分析中一旦发现异常，则对异常相关数据进行标记，提供异常事件背景和辅助信息，并通过过滤器来丢弃用户认定的误报消息。

最后监控界面模块设计成4个可视化子模块，分别是路径时序可视化、地理信息可视化、逻辑拓扑可视化、监测定制接口。路径时序可视化，按时序显示关键路由指标变化，例如丢包率，更新频率。地理信息可视化，在地图中显示关键网络资源位置及该资源相关数据，例如路由器节点位置，路径的地理轨迹。逻辑拓扑可视化，在二维平面显示关键网络资源位置及该资源相关数据，例如多源到单目标的树形结构，瓶颈节点相对位置。

监测定制接口，为用户提供定制所关心目标IP地址，配置告警过滤器等功能，或按需进行实时的IP路径探测任务。

**1.2 POP级别拓扑测量：**

项目研究参考先前已有的六个项目，分别为Rocketfuel、iPlane、Y. Tian等人对于中国网络拓扑的研究、K.Yoshida等人对日本网络拓扑的研究、Topology Zoo项目、Atlas项目。

Rocketfuel项目规划了六个模块，首先Egress Discovery是为了发现出口路由，从而获取独立的前缀，进一步得到待问询IP。Tasklist Generation使用BGP Table（源自RouteView）的数据生成一张直接访问列表，通过与前一步的信息比较，去掉重复项。Path Reduction通过记录已访问的出入口，减少重复请求次数，降低传输通道的压力。Execution engine处理复杂的事项，平衡并限制各个服务器的工作。Traceroute parser从提取表示路由接口、链路的IP地址信息。Alias Resolution通过分析域名名称、各节点回应延迟来寻找alias（属于同一个路由的IP addresses）。

iPlane的工作流程概括的来说有三个部分组成，首先手机一定区域内的路由信息，然后通过探测子网来发现候选的路由间IP对，再通过位置已知的大量的节点对所有接口发送ICMP请求，从反馈结果中确定接口的分布情况，将PoP侦测问题转为了通过聚类、测量等方法对接口进行划分从而获得PoP布局的问题，通过计算节点边界进一步确定POP的划分。

Y. Tian等人提出的POP级中国网络拓扑探测利用15个主要节点进行大规模扫描的方法（针对中国的互联网），其中也利用了很多网络数据中心的数据。开发的两个工具基于traceroute，名为nested IP block partitioning、collaborative tracerouting。与iPlane相比，这两种工具能够发现更多的接口、链接并且整个过程中使用了更少的traceroute 请求。其中IP块划分操作先建立二叉树，将块细化，这里的块以202.85.208.0/20为例，如果对整个域进行traceroute操作是非常耗费时间的。该方法可以将块细分，寻找独立的子网域，从而得到粒度更细的多个分块，在此基础上再进行拓扑测量能提高效率。Traceroute过程为了保证能够得到有效地回应包，选择一个块中的第二个IP地址作为请求目标，通常情况下，这些地址是防火墙地址，对于traceroute请求会给予回应。其次对于每个节点设置reach\_set集合，保存所能到达的IP地址，对于每个IP设置source\_set集合，保存所有能到当前位置的IP地址，在发送问询请求前，查询两张表，即可判定先前的记录中，能否构造一条访问路径，若可以的话，则不进行traceroute操作，最大限度节约工作量。

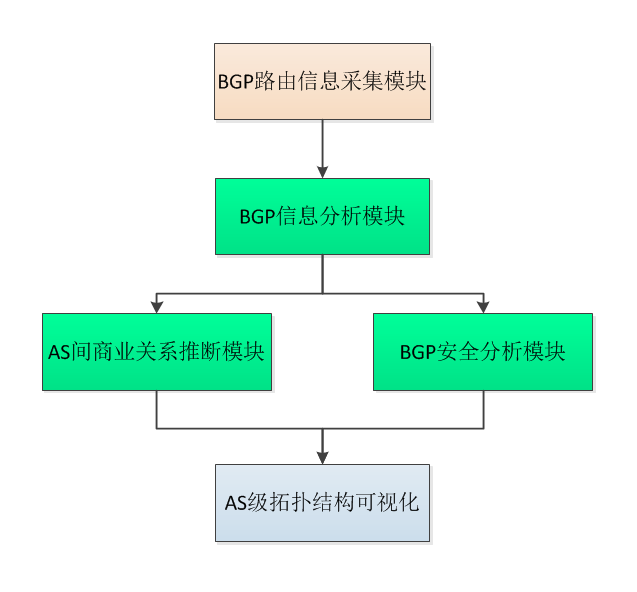
K.Yoshida等人提出了一种不依赖于ISP核心网络的前提下，探寻ISP怎样在POP级别部署其网络的新技术。为了推断出ISP网络的拓扑属性，稳重部署了测量用户间通讯延迟并且在地图中显示PoP级别的交互路径的系统。该方法建立在延迟存在于光导纤维、铜线传输存在延迟的基础上。网络中的传输速率受到最低传输速率的限制，这也是延迟的主要来源。为了消除大多数情况下连接延迟对于总延迟的影响，作者努力的探测核心网络的延迟，将延迟细化为诸多部分，从而更加精准的判断每个阶段的延迟信息。文中简洁的描述了用户间的通讯延迟，并给出了具体公式：delay(src, dst) = adsrc + addst + CD(src, dst) + Esrc,dst，其中，src和dst表示连接至互联网的节点，并且二者相互通讯。Delay（src，dst）表示了源端与目的端通讯延迟，adsrc使src建立连接的延迟，CD（src，dst）使ISP核心网络中发生的延迟，Esrc，dst是测量发生错误时产生的延迟。

Topology Zoo是一个新的数据集合——“互联网拓扑结构集（the Internet Topology Zoo）”。截至目前已统计232个独立的网络结构，而且容量是在持续增长的。该数据集提供了非常准确的数据，可以用来分析未来网络的发展，可以为其他级别的研究提供数据支持，可以模拟真实网络环境来检验新的算法。网络运营商提供的地图是该研究的直接数据来源，交互式的地图如Flash版的地图，可以显示更多的细节，包括连接的类型与传输速率，经纬度信息等。为了建立互联网拓扑结构数据集，研究者们手动下载这些图（涉及到动态地图时会变得复杂一些），然后利用开源工具“yEd”，直接覆盖原图绘制网络元素如路由和POP，并且添加注释信息如节点名称、边界信息。最后处理的结果会被发布到网站上，以GML格式存储。因为数据源的准确性有保证，所以统计结果是非常准确的，但是源数据发生更新时，会导致历史数据失去准确性。

Atlas项目与Topology Zoo相似，目前包含了约9500个PoP区域和全球270个网络结构13500条链路连接。这一项目的创新性在于显示PoP的同时，将所属区域的其他信息如飓风预警、地震预警、公路状况、社会媒体、公共基础设施、天气预报等一并做了收集并显示出来。

**1.3 AS级别拓扑测量模块：**

项目拟解决的问题：从不同的数据源收集BGP路由信息，推断AS间的商业关系，并推断出AS级拓扑的层次结构。同时，分析BGP路由信息发现网络中的性能与安全问题。



AS级别拓扑测量模块可以划分为如下5个子模块：

1. BGP路由信息采集模块：定期从BGP踪迹收集器、路由服务器、窥镜和因特网选路注册机构等数据源采集BGP路由信息。
2. BGP信息分析模块：对采集到的BGP路由信息进行处理，首先将不同来源的不同格式的BGP路由数据进行统一格式化；之后，清除损坏信息、私有地址信息、过度聚合前缀等；然后，对其中不一致的路由信息使用交叉验证的方法获得正确的信息，并排除无效的路由信息；最后，将分析后的BGP信息存入数据库中。
3. AS间商业关系推断模块：使用“无谷底”路径模式(Valley-Free Path Model)分析BGP路由信息，从而推断每条路径上的AS间的商业关系，并生成拓扑层次结构图。
4. BGP安全分析模块：分析IP前缀起源，综合历史数据对前缀劫持事件进行检测，在发现前缀劫持时间时发出安全警报；对AS间的连通性变化进行跟踪检测，在出现网络连接异常是发出安全警报。
5. AS级拓扑结构可视化：显示AS级拓扑层次结构图；在发生安全警报时，在结构图的相应位置显示警报信息。

**1.4 网络扫描模块：**

该项目旨在解决网络扫描速度和扫描功能的矛盾，在保持多样的扫描功能的同时，相对目前的扫描工具，对扫描速度进行优化和提高。

拟采用namp和zmap等成熟工具的主要技术，并进行结合和优化，进行良好的模块划分和整体架构设计，更好的整合二者的优势。

**1.5 IP地理定位模块：**

采用多种IP定位方法相结合互相交叉验证。

借鉴TTG模型，将定位算法进行如下分层设计：

用CBG算法通过时延与地理距离来估测主机位置,同时考虑时延抖动、测量开销以及路径的非对称性等因素。

在第一层的范围内，引入规划理论与概率估计来计算主机的位置，选择合适的数学方法如POIST或建立更为准确的模型,提高定位精度。

对现有IP地址与地理位置映射的数据库进行交叉验证，提炼数据。

2、项目研究方法（技术路线）的可行性、先进性分析

（限500字以内）

**2.1 IP拓扑测量：**

可行性：以fastrace为基础，使用Paris traceroute克服负载均衡造成的假路径，targeted probing增加隐藏路径的发现率。接口合并的实现，参考CAIDA的iffinder和MIDAR别名解析工具，将其使用的同源地址和IPID技术添加到fastrace中。子网发现已有tracenet和Xnet等成熟的工具。

先进性：使用state-of-the-art的测量工具和技术，依赖于大规模测量平台的支持，能够最有效的进行持续拓扑测量。

**2.2 POP级别拓扑测量：**

可行性：正因为网络供应商不希望对外公开所管理的自治域中的路由、接口、子网等信息，所以当前的研究都遇到了很大的阻力。基于traceroute、ICMP请求的探测方法都会依赖于分布广泛的节点服务器，这样才能获取较为准确的数据，才能够将地理信息与POP结合，最终绘制出POP级别的拓扑图。

先进性：基于官方数据的研究相对于其他方法，不能很好的应对数据失效带来的影响，另一方面来看，由于数据的来源有保证，所以得到的拓扑图在一定期限内的准确性最高。

**2.3 AS级别拓扑测量：**

可行性：本方法从多个可靠的BGP信息源处获得BGP路由信息，保证了获得信息的安全性和正确性。对于不一致的信息使用交叉验证的方式，保证了数据的公认性。使用了Valley-Free性质分析AS间商业关系，保证了推断的拓扑层次结构的可靠性。

先进性：本方法从多个数据源获取BGP路由信息，并使用了交叉验证的方式提高了获得数据的准确性。同时，使用了先进的安全分析技术对BGP路由信息进行安全分析，保证了分析结果的正确性。

**2.4 网络扫描模块：**

在对nmap和zmap的深入研究后，对二者的核心技术有了深入理解，对zmap的优化技术的核心及实现有了较深的认识，通过系统设计和模块设计可以将二者的优势分别体现，进一步通过良好的整体架构，就可能将二者优势结合，实现大规模网络扫描速度功能及速度的整体优化。

**2.5 IP地理定位模块：**

可行性：对于CBG算法，Gueye等人在2006年提出CBG算法，Ethan等人对算法进行精炼并延伸，相对成熟。对于Poist算法，Laki等人对概率分布进行了更精准的建模，适应环境提高。对于WHOIS数据库，存在大量的IP地址对地理位置的映射数据。

先进性：综合各种定位方法，利用综合定位算法，同时使用基于推测定位及基于时延定位算法，使定位结果更加准确，利用分层设计，逻辑更加严谨。

三、项目任务(课题)分解方案

（一）项目任务（课题）分解情况

（围绕项目目标，根据需要可对项目目标进行任务分解，要求对分解的各任务（课题）能够相对独立表达、独立测度和独立评价）

（二）项目任务（课题）分解的逻辑关系

（建议用图表描述，并简要说明各任务（课题）在项目中的具体作用。限800字以内）

（三）项目各任务（课题）内容

（逐项分段说明各任务（课题）的研究目标、主要研究内容、拟解决的重大科学问题或关键技术、考核指标及评测手段/方法等。每个任务（课题）限1000字以内）

1、任务（课题）1：xxxxx

研究目标：

主要研究内容：

拟解决的重大科学问题或关键技术：

考核指标及评测手段/方法：

......

2、任务（课题）2：xxxxx

......

四、主要预期创新点

（一）项目预期的主要创新点

（围绕基础前沿、共性关键技术或应用试验等层面，简述项目预期的主要创新点。具体内容应包括该项创新的基本形态及其前沿性、时效性等，并说明是否具备方法、理论和知识产权特征。每项创新点的描述限300字以内）

1、创新点1：xxxxx

2、创新点2：xxxxx

五、预期经济社会效益

项目的科学、技术、产业预期指标及社会、经济效益

（限500字以内）

**第三部分 申报单位及参与单位研究基础**

一、申报单位的已有工作基础、研究成果、研究队伍等

（一）申报单位在该研究方向的前期任务承担情况、相关研究成果

（应说明前期任务取得的主要成果、指标及效果，限800字以内）

（二）负责人及项目主要骨干人员的科研水平及主要成果

（包括工作简历、主要学术业绩、近五年主持的与申请项目相关的各类国家科技计划项目情况，奖励、论文、专利等重点成果取得情况，限1200字以内）

（三）申报单位相关科研条件支撑状况

（包括实验平台、大型仪器设备、国家（重点）实验室、国家工程（技术）中心、重大科学工程等研究基地情况，限500字以内）

二、申报团队参与单位的选择原因及其优势

（限800字以内）

三、相关的国际合作与交流

（限300字以内）

**第四部分 进度安排**

（包括项目主要研究任务的研发进度、年度及重点节点安排、中期目标等。可用图表描述，限1000字以内。）

**第五部分 项目组织实施、保障措施及风险分析**

一、**项目组织实施机制**

（包括项目及分任务（课题）的内部组织管理方式、协调机制等，限800字以内）

二、保障措施

（项目实施的外部政策、组织和资源支撑条件，限500字以内）

三、知识产权对策、成果管理及合作权益分配

（限500字以内）

四、风险分析及对策

（包括潜在的技术风险、市场风险、政策风险，实施过程中的制约因素等，以及相应对策措施，限500字以内）

**第六部分 研究团队**

项目参加人员基本情况表

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **填表说明：** 1、专业技术职称类别：A、正高级 B、副高级 C、中级 D、初级 E、其他；  2、投入本项目的全时工作时间（人月）是指在项目实施期间该人总共为项目工作的满月度工作量；累计是指项目组所有人员投入人月之合；  3、项目固定研究人员需填写人员明细；  4、是否有工资性收入：Y、是 N、否；  5、在项目中的角色：A、项目负责人 B、任务（课题）负责人 C、项目骨干 D、其他研究人员；  6、工作单位：填写单位全称，其中高校要具体填写到所在院系。 | | | | | | | | | | | | | |
| 序  号 | 姓名 | 性别 | 出生日期 | 证件类型 | 证件号码 | 专业技术职称 | 职务 | 专业 | 投入本项目的全时工作时间  （人月） | 在项目中的角色 | 在项目中分担的任务 | 是否有工资性收入 | 工作单位 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 固定研究人员合计 | | | | | | |  | ／ | ／ | ／ | ／ |
|  |  | 流动人员或临时聘用人员合计 | | | | | | |  | ／ | ／ | ／ | ／ |
|  |  | 累计 | | | | | | |  | ／ | ／ | ／ | ／ |

**第七部分 经费概算**

项目经费概算（重点体现对于研究任务的匹配性和保障性，如申报项目进入立项程序将编制项目预算，并予以评估）

|  |  |  |  |
| --- | --- | --- | --- |
| **科目名称** | **合计** | **专项经费** | **自筹经费** |
| **一、经费支出** |  |  |  |
| （一）直接费用 |  |  |  |
| 1.设备费 |  |  |  |
| 2.材料费 |  |  |  |
| 3.测试化验加工费 |  |  |  |
| 4.燃料动力费 |  |  |  |
| 5.差旅费 |  |  |  |
| 6.会议费 |  |  |  |
| 7.国际合作交流费 |  |  |  |
| 8.出版/文献/信息传播/知识产权事务费 |  |  |  |
| 9.劳务费 |  |  |  |
| 10.专家咨询费 |  |  |  |
| 11.其他支出 |  |  |  |
| （二）间接费用 |  |  |  |
| 其中：绩效支出 |  |  |  |
| **二、经费来源** |  |  |  |
| （一）申请从专项经费获得的资助 |  |  | / |
| （二）自筹经费来源 |  | / |  |
| 1、其他财政经费拨款 |  | / |  |
| 2、单位自筹经费 |  | / |  |

**第八部分 指南所要求的附件**

在本部分，请**仅附上**申报指南所要求的相关证明材料或文件。

**申请承诺：**

**申报单位承诺书**

本单位依据国家重点研发计划重点专项指南的任务需求，自愿提交申报书，申请承担和完成其中提出的重点专项任务，并承诺如下：

1. 本单位已完全理解申报指南的要求，并按指南要求进行申报；

2. 本单位已就所申报材料的真实性进行审核，所申报材料内容属实，且不存在学术不端行为；

3. 本申报材料符合《中华人民共和国保守国家秘密法》和《科学技术保密规定》等相关法律法规；

4. 如本申报获得国家专项支持，本单位将签署《国家重点研发计划项目任务书》，同意按照国家重点研发计划有关管理要求和专业机构的相关规定，按照《关于进一步加强国家科技计划项目（课题）承担单位法人责任制若干意见》的有关要求，加强对项目的组织管理，完成研究任务。

如有不符，愿意承担相关后果并接受相应的处理。

申报单位法人代表（签字）：

申报单位（盖章）：

年 月 日

**申请承诺：**

**项目负责人承诺书**

作为申报单位的项目负责人，承诺如下：

1.本人已完全理解申报指南的要求，并按指南要求进行申报；

2.本人对申报材料的真实性负责；

3.本人遵守《中华人民共和国保守国家秘密法》和《科学技术保密规定》等相关法律法规；

4.本人严守学术诚信；

5. 如本申报获得国家专项支持，本人愿意作为项目负责人，按照国家重点研发计划有关管理要求和专业机构的相关规定，以及本单位的科研管理规定，负责开展和完成所申报的研究任务。

如有不符，愿意承担相关后果并接受相应的处理。

项目负责人（签字）：

年 月 日