域间链路发现与商业关系推断综述

吴天昊
网络科学与网络空间研究院
2020312670
wth20@mails.tsinghua.edu.cn

摘要—随着互联网的日益重要，催生了很多对大规模互联网路由、流量进行分析的需求。而一个完整、准确的互联网拓扑与路由是进行大规模路由、流量分析的基础。发现域间链路并推测商业关系可以帮助我们获得完整准确的互联网拓扑与路由。本文调研了域间链路发现和商业关系推断的相关研究，主要分为以下三个方面。第一部分总结了与域间链路发现与商业关系推断相关的公开数据源，详细说明了各个数据源的内容及其优缺点。第二部分介绍域间链路发现的相关研究，主要调研基于 IP 路径的域间链路发现的相关研究。第三部分介绍域间链路商业关系推断的相关研究。

***Index Terms*—域间链路公开数据源，域间链路发现，域间链路商业关系推断**

I. 引言

如今互联网已经成为服务全球数十亿人民的重要基础设施，个人、公司与企业需要使用互联网来进行通讯、娱乐、线上交易、远程教育、办公，国家与政府需要使用互联网来帮助治理各项社会问题、开展民众参与的活动。随着互联网在政治、经济、生活各个方面的作用日益重要，催生了很多对互联网大规模路由感知、分析的需求，以保障互联网的路由安全。例如，近年来很多国家政府对互联网流量的安全日益重视，他们希望了解本国流量被潜在敌对国家监听的风险 [1]；识别对互联网核心网络或者关键流量有较大影响力的国家 [2] [3]。[4] 在互联网路由分析的基础上采取绕避策略使本国的流量绕开部署的流量劫持设施。[5] 利用互联网大规模路由分析，识别互联网中的瓶颈链路，评估网络遭受路由劫持攻击的风险。[6] 为了抵御分布式拒绝服务攻击（DDOS），在互联网大规模路由感知的基础上设计路由策略，将 DDOS 攻击流量分散在非拥塞的链路上，保护了受攻击的网络。

然而对互联网进行大规模路由感知和分析并非易事，互联网规模庞大，其去中心化和分布式的架构，不存在中央控制器让我们能够获得整个互联网的拓扑和路由信息，我们必须根据目前测量得到的有限信息推断未知的拓扑和路由。目前的互联网由六万多个自治系统（Autonomous system，简称 AS，也被称为自治域）组成[1]，它们之间彼此互联，传输互联网的流量。每一个 AS 都由一个运营商所控制，如 AT&T、Cogent、中国电信、中国联通等。AS 之间会签订商业合同、建立商业关系。域间的路由遵循的路由协议是边界网关协议（Border Gateway Protocol，简称 BGP），BGP 协议是基于策略的路由协议，路由策略受到域间商业关系的影响。发现域间链路可以帮助我们获得一个更完整、准确的互联网拓扑，推断域间链路的商业关系有助于我们对域间路由策略进行更好的建模，推测互联网未能直接测量的路由，获得互联网规模的完整、准确的拓扑与路由信息。

无论是域间链路的发现，还是域间商业关系的推断都需要依赖于各类公开的数据集。这些公开的数据集可以分为以下四类：BGP 路由数据、互联网地区性注册机构（Regional Internet Registry，简称 RIR）的数据、互联网交换中心（Internet Exchange Point，简称 IXP）数据以及 traceroute 数据。但是这些数据有的十分分散，有的可能存在错误或者过时的问题，针对这一问题，本文调研了常见的可以用于域间链路发现和商业关系推断的数据源，并分析了各个数据源可能存在的问题与挑战。特别的，traceroute 数据实际上是 IP 接口级别的拓扑、路由数据，如果简单的映射到 AS 级别的拓扑、路由会带来很多错误，因此本文调研了基于 IP 路径的域间链路发现的相关研究，准确的进行 IP 到 AS 的映射可以帮助我们利用 traceroute 数据发现域间链路。

域间链路商业关系推断的研究自从 2000 年马萨诸

[1]截至 2020 年 12 月

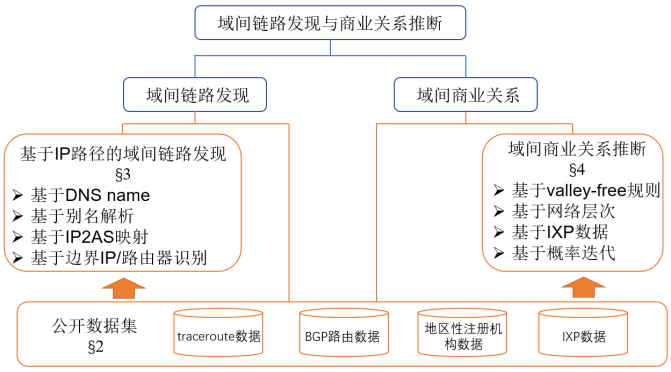

图 1: 综述分类框架

塞大学的 Lixin Gao 教授首次提出已经有了二十多年的历史。从传统的商业关系到复杂的商业关系，商业关系的模型越来越精细，推测的商业关系在实际应用中的效果越来越好。本文将商业关系推测算法分为四类：基于无谷（valley-free）规则、基于 IXP 的数据、基于网络层次、基于机器学习。本文主要围绕上述每一类算法的基本思路，技术演变，优缺点进行详细的分析。

本文的组织架构如下：第二章介绍域间链路发现与商业关系推断所依赖的数据源，第三章介绍边界 IP 地址/路由器归属识别的技术，第四章对域间商业关系推断算法进行详细分析。第五章对本文进行总结。

II. 公开数据源

无论是域间链路发现还是推测域间商业关系，都需要利用公开的数据集作为数据源。对于域间商业关系的推测结果，需要作为 ground truth 的数据集进行验证。本文调研的数据集主要包括四类：第一节介绍 BGP 路由数据、第二节介绍区域性注册机构的数据、第三节介绍 IXP 数据、第四节介绍 traceroute 数据。每一小节首先介绍数据源的数据，其次介绍数据集的应用和存在的问题。第五节对公开数据集进行总结。

A. BGP 路由数据

BGP 路由数据可以分为三类，第一类是 BGP 观测点的数据、第二类是 Route server 的数据、第三类是 Looking glass。(1) 一些组织和机构会在互联网中部署一些路由收集器（Route collector），也成为观测点、测量点，这些收集器会和相邻的 BGP 路由器建立连接，收集它们宣告的 BGP 路由表，最著名的组织包括 Routeviews [7] 和 RIPE RIS [8]，还有一些 ISP 会提供 BGP 测量点，如 Internet2 NOC BGP Rib Dumps [9]。尽管 Routeviews 和 RIPE 会将其测量点观测到的路由数据存储在他们的网站中，但是从数十个测量点数据下载并解压仍然十分麻烦，CAIDA 设计了一个 BGPStreaming 的库 [10]，可以直接调用下载并解压 Routeviews 和 RIPE 指定时间和过滤条件的 BGP 路由数据。(2) 一些 BGP 路由器提供接口可以通过 telnet 或者 ssh 远程连接，利用“show BGP summary”的命令展示 BGP 路由表，这种数据源称之为 Route servers。(3) 一些 BGP 路由器不仅提供了 BGP 数据的查询服务，还支持 ping 或者 traceroute 命令，这种数据源被称为 looking glass。[11]

优点：BGP 路由数据是从 BGP 路由器获得的路由表，其中的路由数据可以说是 AS 拓扑路由的 ground truth。因此 BGP 路由数据中的域间链路是完全可信的真实链路，可信度很高。BGP 路由数据也是域间商业关系推断的基础。每一个 BGP 观测点得到的路由数据几乎包含到达所有 IPv4 前缀的路由，在数据量方面有着一定的规模。

缺点：BGP 路由数据主要存在两个缺点。第一，BGP 数据受到测量点的数量限制和偏差的影响，远离测量点的链路较难被发现。域间链路商业关系推断的研究表明，测量点的数量限制和偏差会影响推测的商业关系的准确性，远离测量点的域间链路的商业关系推测的准确率很低。为了缓解测量点的限制，有研究者提出了 isolario 计划 [12]，希望建立更多的 BGP 收集器，和更多的 BGP 路由器建立联系来扩充 BGP 路由数据。而且 Route server 和 Looking glass 接口分散，缺乏统一的组织管理。第二，BGP 路由是可达性宣告，而不是连通性宣告，AS 可能出于流量工程的考虑添加不属于真实路径上的 AS。而且 BGP 路由表中的数据是经过邻居 AS 的出策略过滤、测量点所在的 AS 的入策略过滤以及路由选择的过程得到的最优路由，因此很难发现备用的路由和链路，这对域间链路的发现和商业关系的推断都造成了阻碍。

每条 BGP 路由的目的 IP 前缀和目的 AS 为我们提供了 IP 地址到 AS 的映射，CAIDA 将 routeviews 的路由数据中的 IP 前缀到 AS 的映射关系存储在它的 Prefix-to-AS mappings 数据库 [13] 中。但是这种简单的 IP2AS 映射只是最简单基础的映射，准确度存在着很多问题，我们将在第三章详细阐述这一问题。

还有一类特殊的数据类型，被称为 BGP commu-

nity。BGP community 是 BGP 路由的一个属性，通常由 32 位的数字组成。高 16 位通常是 AS 号，低 16 位是一个有着特殊语义的值。BGP community 主要分为三类，第一类是 inbound community，主要用于自治域标记其接收到的路由；第二类是 outbound community，主要用于流量工程；第三类是 blackhole community，用于帮助路由安全。inbound community 类的 BGP community 数据对域间链路发现和商业关系推测很重要。inbound community 又可以有四种：一种是用来标记路由是从哪个 IXP 接收到的；第二种是标记与宣告路由的邻居的商业关系，如 AS209 会用 13570 的 community 值来标记它从它的客户那里接收到的路由；第三种是用来标记接收到路由的地理位置；第四种是用来标记从哪个邻居收到路由 [14]。因此 BGP community 数据既可以帮助我们发现域间链路，也可以提供域间链路的商业关系的信息。

优点：BGP community 数据和 BGP 路由数据一样，具有较高的置信度。而且其提供的域间链路商业关系的信息，几乎是商业关系推断的研究中唯一可信的 ground truth。

缺点：BGP community 数据没有统一的规范，community 值的含义几乎完全由各个 AS 自行定义。一些 ISP 和 IXP 会在其网站中公开他们的 BGP community 数据的语义，如 [15] [16]。

B. 地区性注册机构数据

地区性互联网注册机构（RIR）是负责将 IP 地址块、AS 号等资源分配给互联网服务提供商（Internet Service Provider, ISP）的国际组织。目前的互联网共有五大地区性互联网注册机构（RIR），分别是欧洲 IP 地址注册中心 RIPE（Reseaux IP Europeans），拉丁美洲和加勒比海 Internet 地址注册中心 LACNIC（Lation American and Caribbean Internet Address Registry）、美国 Internet 编号注册中心 ARIN（American Registry for Internet Numvers）、非洲网络信息中心 AFRINIC（Africa Network Information Centre）、亚太地址网络信息中心 APNIC（Asia Pacific Network Information Centre）。每一个 RIR 负责相应的大洲的 IP 地址、AS 号的分配，并在其 whois 数据库记载相应的注册的机构信息。其中的 aut-num 数据库记载了 RIR 分配的 AS 的信息，由自治域的管理者维护。其中可能包含了 AS 的路由策略，包括对邻居 AS 的入策略规则和出策略规则，入策略规则指的是 AS 会接受哪些它的邻居宣告给它的路由，出策略规则指的是 AS 会将路由宣告给它的哪些邻居。路由策略信息不仅能够帮助我们发现 AS 与邻居间的链路，而且还能够提供域间商业关系的信息。其中的一个规则 ANY 代表 AS 会接受来自某个邻居的全部路由或是将全部路由宣告给某个邻居，这通常代表着 P2C 关系或者是 C2P。除了路由策略以外，aut-num 数据库还记载了自治域所属的组织机构的信息，这有助于我们识别自治域之间的兄弟关系（我们会在第四章详细介绍域间商业关系），CAIDA 将 RIR 数据库中 AS 与组织机构的对应关系记载在它的 AS2Org 数据库 [17] 中。RIR 数据还有可能包含第一节中提到的 BGP community 数据的语义信息。

优点：第一，数据权威。RIR 记载的是 AS 的运营人员注册的信息，相比于其他数据源的数据更权威、更可信。第二，包含隐藏的链路。绝大多数的测量数据都受到测量点限制与偏差的影响，但是 RIR 作为一个中心式的机构，提供了很多测量点无法发现的隐藏链路与备用链路。

缺点：RIR 数据的缺点主要有两个：一个是格式复杂，RIR 的数据库中的数据是通过路由策略规范语言（Routing Policy Specification Language，简称 RPSL），这种语言的格式规范复杂，难以进行统一的分析；第二是数据滞后。RIR 中记载自治域信息的数据依赖于每个自治域的管理者主动更新，很多 AS 的信息已经有十多年没有更新，相应的域间链路及其路由策略信息很有可能已经过时或改变。使用 RIR 数据通常需要考虑完整性和及时性的 trade-off [18]。正因如此，有研究设计工具来提取和挖掘 RIR 中的有效信息 [19]。

C. IXP 数据

互联网交换中心，简称 IXP，是互联网的重要基础设施，两个 AS 可以在 IXP 建立对等（peer to peer）的商业关系，传输流量，降低传输流量的代价。

PeeringDB [20] 是一个公开的、由用户维护的数据库，记载了 IXP 的地理位置以及 AS 的互连信息。AS 的互连信息由 AS 的管理人员自行维护，定期更新。互连信息包括 AS 在哪些公开和私有的 IXP 建立互连关系。除此之外，PeeringDB 记录了在各个公开的和私有的 IXP 建立互连关系的 AS。PeeringDB 宣称有记录的

AS 数量大于三分之一。有研究的分析表明 PeeringDB 的数据准确率很高而且更新及时 [21]。CAIDA 每个月都会收集 PeeringDB 的数据并存储在它的 peeringdb 数据库中 [22]。

Packet Clearing House (PCH) projects [23] 是负责为互联网关键基础设施（包括 IXP 和域名系统）提供运营支持和安全性的国际组织，他们在全球约三分之一的公共 IXP 上部署 Looking Glass 和路由数据收集设施，AS 可以通过 PCH 与 202 个 IXP 建立互连关系。PCH 维护了一个记载全球 1070 个 IXP 的目录，记录了 IXP 的地理位置、流量以及 IXP 网站的 url，IXP 的网站一般会记录在该 IXP 互连的 AS，有的还会公开 AS 的互连策略。

Euro-IX [24] 与 PCH 类似，它是为欧洲的 IXP 整合资源，以协调技术标准，制定通用程序，共享和发布统计数据以及其他有用信息的组织。Euro-IX 维护了一个 IXP 信息的数据库 IXPDB，这个数据库记载了 632 个 IXP 的信息、15665 个 AS 在这些 IXP 的互连信息以及 49 个 IXP 的 route server 的信息。除此之外，Euro-IX 还提出了一套统一的 BGP community 语义规范，并推荐 IXP 采用这一规范。

优点：IXP 的数据有两个优点。第一，学界公认互联网中的 P2P 链路很难被发现，且 P2P 商业关系很容易被推测错误。而 IXP 数据可以帮助我们发现域间的 P2P 链路，因此具有非常大的价值。第二，IXP 数据虽然是由 AS 或者 IXP 的管理者更新和维护，相比于 RIR 的数据，更新频率很快，应用价值很高。

缺点：不同的 AS 在 IXP 有着不同的互连策略，两个 AS 与同一个 IXP 建立互连关系，并不代表这两个 AS 彼此互连，需要根据 AS 在 IXP 的互连策略或者采取进一步的探测方法来确认 AS 之间的互连关系。

D. traceroute 数据

traceroute 是获得 IP 接口级别路由的工具，执行 traceroute 命令的主机被称为测量点（Vantage Point，简称 VP），VP 向目标 IP 地址发送 TTL 递增的 ICMP 报文（有的版本的 traceroute 采用 UDP 或 TCP 报文），获得 VP 到目的 IP 经过的路由器响应的接口 IP 序列。CAIDA 的 ark 项目从全球数十个测量点 24 小时不间断的对 IPv4 的全部/24 前缀地址空间的 IP 地址进行探测并记录探测结果 [25]。利用 BGP 路由数据的 IP2AS 映射，我们可以通过 traceroute 数据发现域间的链路。RIPE Altas 提供全球大量的 VP 可供研究人员使用，但是使用 VP 需要消耗 credit，因此存在着限制。

数据集分类	数据集	可信度	及时度	受测量点偏差影响	集中式管理	应用范围
BGP路由数据	BGP观测点的数据 Route server Looking Glass BGP community	高	高	是	部分是	域间链路发现 域间商业关系推断的原始数据和验证集
地区性注册机构数据	RIR的WHOIS数据库	中	低	否	是	推测域间的兄弟关系 提供路由策略辅助进行商业关系推测 包含部分BGP community的语义信息
IXP数据	PeeringDB PCH projects Euro-IX	高	高	否	是	发现P2P链路
traceroute数据集	CAIDA的Ark数据集	中	高	是	否	发现细粒度的域间链路 推断复杂的域间商业关系

图 2: 公开数据集总结

优点：利用 traceroute 数据发现域间链路主要有三个优点。第一，traceroute 路由数据相比于 BGP 路由数据，更能反映在真实互联网中的路由情况。第二，traceroute 可以提供更细粒度的域间链路数据，相比于 BGP 路由数据只能发现域间的单条链路，traceroute 数据可以帮助我们发现自治域之间的多条链路，以及链路两端 IP 地址。IP 地址相比于 AS 更容易进行地理定位，可以为发现的域间链路提供更多的信息。第三，VP 的部署要比 BGP 路由收集器的部署更容易。

缺点：第一，与 BGP 路由数据类似，traceroute 数据也受到测量点的数量限制和偏差的影响。第二，traceroute 最初只是一个网络诊断工具，并不是为了进行网络拓扑路由发现。因此 traceroute 得到的 IP 路由存在着很多问题。第三，IP 地址到 AS 的映射并不简单，仅仅通过 BGP 路由数据进行映射会带来很多错误。正因为利用 traceroute 技术进行域间链路发现面临如此多的挑战，很多研究提出启发式算法以消除利用 traceroute 数据发现域间链路的错误，我们将在第三章详细介绍挑战和目前研究的解决方案。

E. 总结

图 2总结了域间链路发现和商业关系推断所涉及到的公开数据集，总的来说，不同的数据集在可信度、及时度、是否受测量点偏差影响、是否集中式管理等方面有着很大的差异，不同的数据集的应用范围不同。公开数据集既是域间链路发现和商业关系推断的原始数据，

也是验证集。在对未知的链路和链路的商业关系进行推断时，灵活使用各类公开数据集，融合使用多源数据，能够大幅度提升推测的准确性和覆盖率。

III. 基于 IP 路径的域间链路发现

在第二章中我们在介绍公开数据集时提到 traceroute 对于域间链路发现有着重要的意义。一方面，traceroute 测量点的部署要比 BGP 收集器的部署容易的多，这意味着我们能在更多的测量点测量数据；另一方面，traceroute 提供的 IP 级别的数据为我们提供了更细粒度的域间链路信息。但是从 traceroute 得到的 IP 路径发现域间链路并非易事。挑战主要来自两个方面：一是 traceroute 本身的问题，如无响应、响应接口没有统一的规范等问题；二是 IP 地址到 AS 的映射问题，尤其是位于自治域边缘路由器的接口地址的映射非常困难。在本节中，总结了利用 IP 路径进行域间链路发现的相关研究，围绕基本思路和技术演变进行分析。在第一节中介绍基于 DNS name 的方法，在第二节中介绍基于别名解析的方法，在第三节中介绍基于 IP2AS 的方法，在第四节中介绍基于边界 IP/路由器识别的方法。

A. 基于 *DNS name*

IP 地址对应的 DNS 域名中蕴含着很多信息，其中可能包含了 IP 地址所属的 ISP 或者 AS 的信息，比如我们可以从域名 sl-bb11-nyc-3-0.sprintlink.net 得出该 IP 地址是属于 sprintlink 这个 ISP 的。Spring 等人 [26] 提出了一个著名的且使用广泛的 DNS 域名信息挖掘的工具 UNDNS，利用正则表达式提取域名中的有效信息。

使用 DNS name 识别 IP 地址的归属 AS 最大的优点在于：由于 ISP 之间对 IP 地址的域名规范不同，DNS name 可以帮助我们精准的定位网络边缘，例如两个网络之间的域间链路的两端 IP 地址通常属于一个 AS 的地址空间，通常被识别为同一个归属 AS，但是两个 IP 地址的 DNS name 却有所不同。

但是利用 DNS name 识别 IP 地址的归属 AS 存在三个问题。第一，不是所有的 IP 地址的域名信息中都包含 IP 地址的归属 ISP 和 AS，也就是说仅凭 DNS name 进行 IP 到 AS 的映射是不全的，这是基于 DNS name 的方法所遇到的最主要问题。第二，不同的 ISP 对域名规范不同。[27] [28] 指出域名没有统一的规范，很难提取包含复杂命名规则的域名中的有效信息。为了解决这一问题，[28] 提出了一个在 DNS 主机名中查找特定信息的自动化过程，他通过编写字符串字典，生成通用的域名规则，获得了 1398 个自治域的域名规则，但是这个数量相比于全球互联网的高达六万多个的自治域数量仍然不尽人意。第三，域名不会经常更新，这会导致 IP 地址信息的滞后，以及映射的错误 [29]。

总的来说，使用 DNS name 进行 IP 到 AS 的映射的准确度很高，具有较高的可信度，但是 DNS name 并不能解决所有 IP 地址的映射问题，我们需要新的方法扩展可以映射的 IP 地址的范围，在之后的研究中，DNS name 通常会作为验证集来验证其算法的准确性。

B. 基于别名解析

在 IP 到 AS 的映射过程中还存在着一个中间的粒度——路由器级别的粒度。我们可以将 IP 地址聚合成路由器，再研究路由器的归属 AS。将 IP 地址聚合成路由器的技术被称为别名解析技术，迄今为止，很多研究都在改善别名解析的准确性 [26] [30] [31] [32] [33]。

Chang 等人 [34] 首次利用别名解析的路由器拓扑发现域间链路。他们利用 Mecator 工具 [32] 对 traceroute 得到的 IP 地址进行别名解析。在得到的路由器级别的拓扑上识别路由器的归属 AS，他们提出了简单的启发式条件推断路由器的归属。启发式条件的假设如果一个路由器的一个端口是属于 AS_1 的，那么这个路由器就是属于 AS_1 或者 AS_1 的邻居 AS $peer(AS_1)$，同理一个端口是属于 AS_2 的，那么这个路由器就是属于 AS_2 或者 AS_2 的邻居 AS $peer(AS_2)$。假设一个路由器的端口是属于 $\{AS_1,AS_2,...,AS_n\}$ 的，那么这个路由器就应该属于 $\{AS_1, peer(AS_1)\} \cap \{AS_2, peer(AS_2)\} \cap \cdots \cap \{AS_n, peer(AS_n)\}$，如果交集的结果只有一个 AS，那么就推测这个路由器属于这个 AS，如果有多个就采用多数表决的原则推测路由器的归属。

Huffaker 等人 [35] 利用 MIDAR [30] 和 kapar [33] 两个工具进行别名解析。对于每个路由器，利用 BGP 路由表将每个接口 IP 地址映射到 AS。利用五步的启发式算法对路由器的归属 AS 进行推测。第一，如果一个路由器的所有接口 IP 地址都属于一个 AS，那么这个路由器就属于这个 AS；第二，如果一个路由器绝大多数的接口 IP 地址都属于一个 AS，那么这个路由器就属于这个 AS；第三，考虑邻居路由器的归属 AS 情

况，如果一个路由器绝大多数的邻居路由器都属于一个 AS，那么这个路由器就属于这个 AS；第四，如果邻居路由器所属的 AS 之间为 provider to customer 的商业关系（在第四章中我们将详细讲述自治域间的商业关系），那么就推测路由器属于 customer 的 AS，因为通常 provider 会对外宣告他的部分 customer 的 IP 地址空间。第五，考虑邻居路由器所属的 AS 的度数，度数较小的 AS 通常为 customer，因此路由器属于度数最小的邻居路由器所属的 AS。

[36] 希望发现域间的对等链路。他们利用测量点探测可能的域间对等链路设施，将 IP 路径利用 MIDAR 聚合成路由器，推测可能的域间链路。这些链路会作为约束设施搜索（Constrained Facility Search, CFS）算法的输入。约束设施搜索算法会不断迭代，缩小可能的候选域间对等链路设施，直到推测出唯一的域间对等链路设施。

总结来看，基于别名解析的方法依赖于别名解析的准确性。目前为止，还没有一个足够准确的别名解析算法可供精准的推测域间链路。而且，基于别名解析的这些研究没有考虑 traceroute 的固有问题对推测结果的影响；在 IP 到 AS 的映射过程中，采用的都是 BGP 路由表这种最简单的映射方式，而这种映射存在很多挑战。不过，将 IP 地址聚合成路由器的方法的确避免了 IP 链路推测错误造成的影响，在之后的研究中，一些工作同样沿用别名解析技术，在路由器级别的拓扑上推测域间链路。

C. 基于 *IP2AS* 映射

为了解决 traceroute 本身的问题以及 IP 到 AS 的映射错误对域间链路发现带来的挑战，Mao 等人 [37] 将 BGP 路径和 traceroute 路径产生分歧的原因进行分类，并分别提出启发式算法，识别映射错误的情况并予以改进。BGP 路径和 traceroute 路径产生分歧的原因共有六类。第一类是由于 traceroute 中的某一跳没有响应。针对这一问题，作者的解决方案是，如果没有响应的跳的前后两跳的 IP 地址被映射到同一个 AS，那么这一跳的 IP 也应该映射到这个 AS。第二类是由于 IP 地址无法映射到对应的 AS。类似的，如果这个 IP 地址前后两个 IP 被映射到同一个 AS，那么这一跳的 IP 也应该映射到这个 AS；如果不一样，则使用 DNS name 将这个无法映射的 IP 地址映射到 AS。第三类是多源 AS 的 IP 地址，即一个 IP 地址可以映射到多个 AS。作者认为只要能映射到其中一个 AS 就算是映射正确。第四类是 IXP 的 AS。IXP 所属的 AS 通常不会在 BGP 路径中出现，但是在经过 IXP 的 traceroute 的路径中会出现 IXP 的 AS，导致 BGP 路径和 traceroute 路径的分歧。解决方案是识别 traceroute 中的扇型结构，通常这种结构意味着 traceroute 路径经过了一个 IXP，当发现这种结构后会将 IXP 的 AS 移除，并将两侧的 AS 修改成全连接的结构。第五类是兄弟关系的 AS，互为兄弟关系的两个 AS 会共享地址空间。假设 AS B 和 D 是互为兄弟关系的两个 AS，在实际 BGP 路径中可能只经过一个 B，而在 traceroute 得到的路径映射后的 AS 路径会经过 B 和 D。因此作者识别这种错误，并移除多余的 AS。第六类是 provider 的 AS 可能会宣告部分 customer 的 IP 地址空间，当 traceroute 路径经过 customer AS 时会出现映射错误。这种错误的一个表现是 traceroute 映射后的 AS 路径中出现环路，因此作者识别出现环路的路径，修正错误的 IP 到 AS 映射。

Yu Zhang 等人 [38] 系统地对 IP2AS 的错误匹配的原因进行了研究，他们将发生错误匹配的原因分为了七类：一是由于数据平面的路径与控制平面的路径不一致，如 BGP 路由聚合、默认路由、BGP 多条会话、网络隧道、二层交换机、异常路由以及路由误配置。第二个原因是 traceroute 的路径中的某一跳或某几跳没有相应。三是由于一些 IP 地址没有办法映射到相应的 AS，这可能是私有地址所导致的，也可能是前缀没有被所有者宣告。四是多宿主 AS 的前缀导致的，如果前缀是多宿主 AS，那么这个前缀会被多个 AS 所宣告。五是前缀属于某个 IXP 的 AS，这个前缀有可能不出现在 BGP 路由表中。六是兄弟关系的 AS，彼此之间可能不存在明确的 IP 地址的边界。其中一个 AS 拥有的前缀可能会被兄弟关系的 AS 所宣告。第七个原因，也是最重要的原因是边界 IP 与路由器的识别的困难。如图 3所示，由于一条域间链路的两个接口 IP 属于同一个子网前缀，因此他们只会被其中一个 AS 所宣告，在图中的例子，链路 L2 的两个接口 IP 都呗 AS3 所宣告。由于 traceroute 的响应接口没有强制的规范，因此路由器可能会采用任意一个接口地址响应 traceroute 的探测报文。在图中的例子里，AS2 的边界路由器返回了 L2 一侧的接口地址 IP2，而这个 IP 地址被错误的映射到了 AS3，导致域间链路推测错误。研究结果表明大约

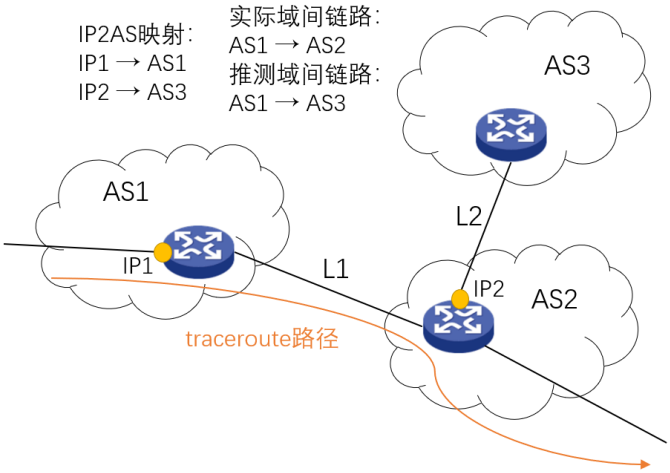

图 3: 第三方 IP 地址导致边界 IP 识别困难

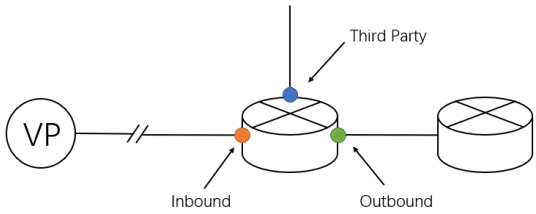

图 4: 三种响应的接口 IP 类型

60% 的 IP 路径到 AS 路径的错误匹配都是由于路由器使用它的邻居 AS 的 IP 地址进行回复所导致的，而之前的研究都没有考虑这一问题，因此在这之后催生了关于边界 IP/路由器归属识别的研究，我们将在第四节详细讲述。

D. 基于边界 IP/路由器映射

图 4展示了路由器响应 traceroute 探测报文的接口 IP 地址的三种类型。接收 traceroute 探测报文的接口地址被称为 inbound 地址，靠近 traceroute 下一跳路由器的接口地址被称为 outbound 地址，路由器的其他接口被称为 third party 地址。

Marchetta 等人 [39] 首次提出边界第三方 IP 地址的识别方法。作者利用 IP Timestamp option 来检测链路中的第三方 IP 地址。在网络探针中设置 IP option 中的预设 IP 地址为代检测地址，通过网络探针返回的 stamp 个数来判断是否为第三方 IP 地址。这样的方法建立在一个较为严格的假设——路由器会回复、引用、标记网络探针 IP option 中所预设置的 IP 地址，然而这个假设是没有保障的。所以作者在进行实验前对数据集进行测试和过滤，从 PREDICT 里 14K 个 AS 的 327K 个地址中，选择了包含约 443K 个地址的约 12M 链路，通过实验，作者发现由 traceroute 推断出的 AS 拓扑中，约有 17% 包含第三方 IP 地址，所以这些 AS 拓扑都可能是不准确的。

由于此前 Marchetta 等人提出的利用 IP Timestamp option 来检测第三方 IP 地址的方法非常依赖于路由器处理 IP Timestamp option 时的一致性，而一致性取决于各个路由器厂商、服务提供商。所以，Luckie 等人希望通过更加严格的实验来验证该方法的鲁棒性。因为路由器之间的连接通常是点对点的连接，属于同一个/30 或是/31 子网，所以如果能确定 traceroute 中连续两跳之间的连接是在一个/30 或是/31 网络下，则可以确定接受数据包的 IP 地址一定是路由器的 in-bound 端口，该地址自然不是第三方 IP 地址。作者首先在 8 个 CAIDA Ark 测量点上各随机挑选 10,000 个地址进行 traceroute 测试，发现仅有不到 20% 的地址会引用 IP Timestamp option。通过作者此前提出的 Scamper 工具 [40]，在所有响应 IP Timestamp option 的 traceroute 结果中筛选出了 197,335 个 IP 连接中的 86,152 个点对点连接。使用 Marchetta 等人提出的方法在这些点对点连接中进行测试，发现约 80% 的 inbound 端口被判定为第三方 IP 地址，说明此前 Marchetta 等人提出了利用 IP Timestamp option 来检测链路中的第三方 IP 地址的方法并不可靠。从实验结果来看，很多路由器会响应 ICMP 探针中的 IP Timestamp option，但却不对 UDP 探针作出响应，这就导致已有方法错误的将很多非第三方 IP 地址判定为第三方 IP 地址。

Marder 等人 [41] 提出了一种边界 IP 地址的识别算法——MAP-IT。由于同一条链路的两端的接口地址经常位于同一个/30 或者/31 子网，因此对每个接口的另一端的接口地址进行推测。记录每个接口地址的前继 IP 地址 N_b 和后继 IP 地址 N_f。并基于前向和后向的方向，将接口地址分成两半（interface halves，简称 IH）。利用四步骤的推断算法，对域间链路进行推断。第一步，利用 IP2AS 的映射直接推测一侧 IH 域间链路。第二步，利用 IP2AS 的映射推测另一侧 IH 的域间链路，如果与第一步的结果不同，则修改第一步得到的结果。第三步，对推测结果中矛盾的地方进行修正。第四步，相邻的域间链路推测。发生在前继接口和后继接口都推测为是域间链路，这时认为靠近 traceroute 观测点的推测是域间链路。

Luckie 等人 [42] 提出了一个著名的边界路由器的识别算法——bdrmap。这一推测算法基于 traceroute 得到的数据。如果想要推测一个网络的边界路由器，则必须在这个网络中包含一个测量点。traceroute 的目的 IP 为 BGP 路由数据的 BGP atom。利用 paris traceroute，将得到的 IP 地址通过 Ally 和 Mercator 两种别名解析技术聚合成路由器，利用 prefixscan 算法验证 IP 地址是 inbound 接口地址。接着利用生成的路由器拓扑，从精确到模糊，逐步骤的推测边界路由器，一共有八个步骤。第一步，识别和 VP 位于同一个 AS 的路由器。第二步，推测防火墙的边界路由器。第三步，识别没有宣告的 IP 地址。第四步，利用 IP2AS 的映射推测边界路由器。第五步，利用 AS 商业关系推断边界路由器。第六步，利用 IP2AS 的映射推测模棱两可的情况下的边界路由器。第七步，推测边界路由器额外的别名。第八步，利用 TTL 超时信息推测边界路由器。

bdrmap 使用从一个测量点得到的数据，通过别名解析聚合成路由器级别的拓扑，利用特殊的启发式算法推测边界路由器，bdrmap 只为第一个 AS 的边界的路由器推测所属 AS，并且在每个感兴趣的网络中都需要一个观测点。MAP-IT 利用一个迭代的过程对由多个测量点得到的 IP 接口级别的拓扑进行推测，MAP-IT 缺乏对边缘网络和低可见链路的推测算法，比如没有后继的跳的路由器可能是防火墙导致的，而且没有使用别名解析聚合路由器。Marder 等人在 bdrmap 和 MAP-IT 算法的基础上，融合二者的优点，提出了 bdrmapIT 算法 [43]，采用 bdrmap 的启发式算法以及 MAP-IT 的递归的架构，可以从任意数量的测量点推测可见路由器的归属。bdrmapIT 算法主要分为三个阶段，一个是构建拓扑图，第二个是推测最后一跳所属 AS，第三是标记推测的路由器和接口地址。首先，作者利用原始的 traceroute 数据构造了一个路由器级别的拓扑图。第二阶段是标记最后一跳，以防止防火墙对结果的影响。第三阶段是通过一个迭代的过程修正结果。bdrmapIT 算法是目前最好的边界路由器的识别算法，其代码开源。

在本节中的上述工作主要为了解决第三方 IP 地址带来的挑战，但是路由器还可能返回与接收 traceroute 探测报文相对的端口 IP 地址，这个 IP 地址被称为 outbound 的 IP 地址。尤其是在 L3VPN 中，路由器很可能以 outbound 的 IP 地址响应 traceroute 的探测。这会带来三个挑战，一是连续两跳的 IP 地址属于同一个子网，但是这两个 IP 地址不是属于同一个链路，如何识别是一个挑战。二是在 traceroute 过程中可能出现路由循环，如果返回 Outbound 地址，路由循环很难识别。三是与 IXP 子网相邻的路由器如果返回 outbound 地址，难以识别归属 AS。四是某些情况下，由于探测的限制，某些路由器的回复可能没收到，导致 outbound 地址处于最后一跳。Marder 等人提出了 vrfinder 算法 [44]，识别 traceroute 过程中出现的 outbound IP 地址，针对上述四个挑战的情况设计启发式条件，识别 IP 地址的归属。目前 vrfinder 算法已经集成到了 bdrmap-IT 的工具中。

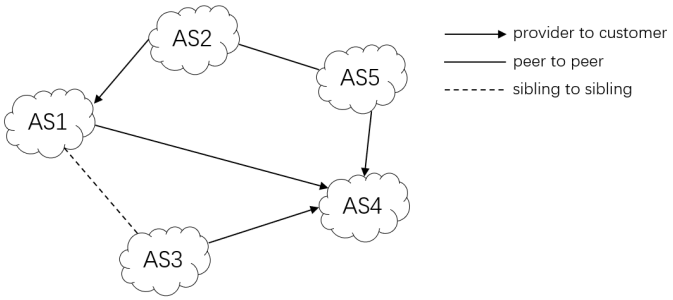

图 5: 传统的域间商业关系

IV. 域间商业关系推断

自治域之间运行的路由协议是 BGP 协议，BGP 协议是基于策略的路由协议，自治域的运营商会按照自治域之间签订的商业合同制定路由策略。而域间的商业关系就是对自治域间的商业合同和路由策略的建模。传统的域间商业关系有三种：(1) 提供商与客户关系（provider to customer，简称 P2C 关系）：提供商为客户传输流量，保证客户流量的全球可达性，客户向提供商付钱。(2) 对等关系（peer to peer 关系，简称 P2P 关系）：建立对等关系的双方互相为对方的客户提供传输服务。(3) 兄弟关系（sibling to sibling，简称 S2S 关系）：建立兄弟关系的双方互相为对方传输所有的流量，这一商业关系通常发生在同一个公司的子公司所运营的自治域之间。三种传统的商业关系如图 5所示：

自治域之间的商业关系是自治域的运营商的商业秘密，运营商不会主动公开它与其他运营商之间的商业关系，我们只能根据原始的 BGP 路由数据，推测自治域的路由策略，进而推测自治域之间的商业关系。域间商业关系推断算法主要分为四类：第一类是基于 valley-free 规则，在第一节中介绍；第二类是基于网络

(a) 第一类 VF 路径　　(b) 第二类 VF 路径

图 6: valley-free 规则

层次，将在第二节中介绍；第三类是基于 IXP 数据，在第三节中介绍；第四类是基于概率迭代的方法，将在第四节介绍。第五节对域间商业关系推断研究进行了总结。

A. 基于 *valley-free* 规则

Gao Lixin [45] 是第一位研究 AS 之间商业关系推断的学者。她提出 AS 之间存在着商业关系，且商业关系分为三类，分别是 provider to customer, peer to peer 和 sibling to sibling。她的算法基于对 BGP 的出策略的一个假设——valley-free 规则，这一规则指的是一个 AS 不会将来自他 provider 和 peer 的路由宣告给它的 provider 或 peer。在这个规则的影响下，一条 AS 路径可以分为两类。第一类 valley-free 路径是由三部分组成的，首先是零条或者数条 customer to provider 的链路，其次是一条 peer to peer 链路，最后是零条或者数条 provider to customer 的链路组成，如图 6(a)所示。第二类 valley-free 路径由两个部分组成，首先是零条或者数条 customer to provider 的链路，其次是零条或者数条 provider to customer 的链路，如图 6(b)所示。基于 valley-free 规则，只要找到位于山峰的最大的 provider，就可以推断出整个路径每条链路的商业关系。作者假设 AS 的规模可以用它的度数来衡量，规模越大的 AS 度数越高，彼此之间建立 p2p 关系的 AS 之间的度数相似。因此，作者提出的算法分析每条 AS 路径中度数最大的 AS 将两侧的路径上的链路识别为传输关系，对于任意两个 AS，如果在绝大多数情况下都发现一个 AS 为另一个 AS 提供传输服务，那么就认为这两个 AS 之间是 p2c 关系，如果在很多情况下发现两个 AS 都为对方提供传输服务，那么就认为两个 AS 是兄弟关系，如果两个 AS 的度数相似，那么就推测是 p2p 关系。推测出的结果通过 AT&T 的与其他邻居 AS 的商业关系作为验证集来验证结果，结果发现，有 99.1% 的推测结果都得到了证实。她还使用了 WHOIS 数据来验证推测出的兄弟关系，结果表明推测出的超过一半的兄弟关系都可以在 WHOIS 数据库中被验证。

Subramanian 等人 [46] 在 [45] 的基础上将 AS 商业关系推断问题正规化为一个称为 ToR(Type of Relationship) 的组合最优化问题：给定一个从 BGP 路由数据中得到的 AS 拓扑，为每一个链路分配一个商业关系使得 BGP 路径中的满足 vall-free 规则的路径数量最大。作者推测 ToR 问题是一个 NP 完全的问题，需要通过启发式算法来解决。Gao 的算法有一个不足之处在于，由于她采用的原始数据是 BGP 路由数据，BGP 路由数据中缺失很多底层的 peer to peer 链路，导致底层的 AS 之间的 P2P 商业关系难以发现和推测。为了解决这一挑战，他们提出了一个推测 AS 商业关系的启发式算法——SARK 算法，这一算法有一个很强的假设，那就是绝大多数的测量点都位于网络的最高层，因此，从这些观测点观测到的 AS 路径都是下坡的路径，都是 p2c 的链路。以测量点为根构建一棵路由树，按照从叶节点到根节点的顺序对 AS 进行 rank 排序，rank 大的 AS 是 rank 小的 AS 的 provider。他们利用多个测量点的数据，每个测量点都根据这个测量点的数据对每个 AS 进行排序，这样每对 AS 在每个测量点有着不同的 rank 顺序。接着作者采用多数决原则 (majority vote)，如果一条链路两个 AS A 和 B 在绝大多数的测量点都发现 A 的 rank 比 B 大，在很少的测量带你发现 B 比 A 大，则推测 A 是 B 的 provider。如果一条链路的两个 AS 在绝大多数的测量点发现 rank 相同，或者观测到 Rank(A)>Rank(B) 和观测到 Rank(A)<Rank(B) 的测量点数量相似，那么就认为 A 和 B 是 P2P 关系。作者衡量算法准确性的方法是统计根据推测的结果有多少路由违反 valley-free 规则，这一评价标准显然是不合理的。

Battista 等人 [47] 分析了 ToR 问题的复杂度。他们首先研究了一个简化的 ToR 决定问题，即判断是否可以为每个链路分配一个商业关系，使得满足 valley-free 规则的路径数量不小于 k。作者将这一问题规约为一个 2SAT 问题，可以在多项式的时间内求解。接着，作者将 ToR 问题规约为一个 MAX2SAT 问题，从数学上证明了 ToR 问题本质上是一个 NP 完全问题；而且证明了当每条路径的路径长度较小（小于一个常数）时，ToR 问题仍然是一个 NP 难问题。因此需要通过启发式的算法进行求解。作者提出了一个两阶段的启发式算

法，第一个阶段通过求解 2-SAT 问题找到一个较大的 k 求解 ToR 决定问题，使得求解出的每条链路的商业关系可以让大于等于 k 的路径都满足 valley-free 规则。对于剩余的路径，检查是否能够加入到集合中且不违反 valley-free 规则。值得注意的是，[47] 只推测 P2C 关系，并且只将域间商业关系推断当成一个数学问题，与实际互联网差别很大。

Shavitt 等人 [48] 利用三种算法识别 AS 拓扑的核心节点，分别是 Jellyfish、K 核分解以及 CAIDA 提供的 AS 排名。在推测核心节点之后，所有经过核心节点的路径中，核心两侧的链路根据 valley-free 规则分别推测为 C2P 关系和 P2C 关系；对于没有经过核心节点的路径，如果路径中存在一条链路的商业关系被推测为 C2P 关系，则这条链路左侧的链路都会被推测为 C2P 关系，同样的，如果路径中存在一条链路被推测为 P2C 关系，那么这条链路右侧的链路都会被推测为 P2C 关系。如果一条链路在推测的过程中出现冲突，则采用多数表决的原则。对于剩余的非推测链路，该算法无法推测其商业关系。

Neudorfer 等人 [49] 首次在 PoP 级别的粒度推测 AS 之间的商业关系。之所以要在 PoP 级别的粒度研究商业关系，是因为 AS 之间在不同的 PoP 点的商业关系可能不相同。他们利用 traceroute 的数据，将 IP 地址映射到 PoP 点和 AS，利用 valley-free 规则对每条链路的商业关系进行投票，投票大于阈值的链路的商业关系被确定的推断；对于无法推测出的链路以及推测出的违反 valley-free 规则的链路，作者进行了异常检测，发现主要有七个原因：一是 IP 地址到 AS 的映射错误；二是链路涉及 IXP 和兄弟关系；三是商业关系的推测错误；四是存在复杂商业关系；五是一些科研机构的 AS 之间的关系与传统的商业关系不同；六是 traceroute 的错误；七是未知的异常。

基于 valley-free 的方法实际上是从路由数据中提取路由策略，再从路由策略推测商业关系，这种方法主要存在四个问题：一是用于推测商业关系的原始路由数据很多存在误配置、毒性路径、路由泄露等问题，这些脏数据会对推测结果造成很大的影响。第二，从 BGP 路由得到的 AS 拓扑丢失很多 P2P 链路，P2P 链路既难以发现，也难以识别。第三，valley-free 规则不广泛成立，valley-free 规则实际上是对域间的出策略进行的抽象和建模，这种建模过于简单，与实际的域间路由策略有所偏离，随着时间的推移，违反 valley-free 规则的路径越来越多。第四，商业关系是对路由策略的建模，但是实际的路由策略要比目前传统的商业关系的粒度更精细。这些挑战使得学界迫切的需要设计新的推测算法，来推测域间的商业关系。

B. 基于网络层次

与以往的方法不同的是，Matthew 等人 [50] 没有使用 valley-free 规则推断商业关系，而是利用网络层次来进行商业关系的推断，这个算法被称为 AS-Rank 算法。基于三个主要假设：(1)AS 会通过建立 provider 关系来实现全球可达性 (2) 最高层次的 AS 彼此之间的对等关系形成一个网状（mesh）结构 (3)p2c 关系没有环路。数据源依旧是 Routeviews 和 RIPE RIS。他们首先对原始的路由数据进行清洗，去除包含 AS loop、保留 AS 号的 AS 以及 IXP 的 AS 的路径。计算每个 AS 的度数以及传输度数（有多少 AS 可能需要它来传输流量）并进行排序，利用 Bron/Kerbosch 最大团算法找到最核心的 AS 并把彼此之间的关系推断为对等关系，将 AS 路径拆分成三元组的形式，按照度数从大到小的顺序、通过十一步的启发式算法推断 AS 之间的 p2c 关系，对于剩余的未知的 AS 链路推断为对等关系。其中最重要的一步（这一步推测出的链路数量占总数的 90%）是将路径拆分为三元组 X Y Z，如果 X 与 Y 的商业关系为 P2P 或者 P2C 则推测 Y 与 Z 的商业关系是 P2C，其实这种推测方法和 valley-free 规则对山峰右侧的路径的商业关系推测方法很类似。之所以不推测山峰左侧的路径，是因为防止路由误配置或者路由泄露（一个 AS 将来自 provider 或者 peer 的路由错误的宣告给它的 peer）导致的错误的 C2P 推测。他们使用从网络管理员获得的信息、RIR 的 WHOIS 数据库和 BGP communities 作为 ground truth 验证推测的准确性，验证推测出的商业关系，其中 c2p 和 p2p 的准确性分别达到了 99.6% 和 98.7%。CAIDA 使用 AS-Rank 算法，每个月推测互联网的域间商业关系，并将结果记录在 CAIDA 的 AS Relationships 数据库 [51] 中，成为很多互联网分析的重要数据来源。

Vasileios 等人 [52] 在 AS-Rank [50] 的基础上推测 AS 之间复杂的商业关系。他们发现除了传统的 p2c，p2p 以及 sibling 关系外，还存在着两种复杂的商业关系；一种复杂的商业关系被称为混合关系，即两个 AS

在不同的 PoP 点具有不同的商业关系（这通常体现在不同目的前缀的路由在不同的 PoP 点宣告，具有不同的路由策略）；另一种复杂的商业关系是部分传输关系，即 provider 不会将来自它的 provider 的路由宣告给它的 customer。首先，他们利用 AS-Rank 的算法推断 AS 之间传统的商业关系。接着，他们将 BGP 路由数据按照目的前缀分组，对于 AS-Rank 推断出的传统的传输关系（P2C 关系），测量对于每一个目的前缀，provider 是否会为 customer 完全传输所有流量，完全传输标记为 FT、部分传输标记为 PT，只为 customer 传输到其他 customer 的流量标记为 P。如果两个 AS 对于某些目的前缀是一种商业关系，对于另一些目的前缀表现为另一种商业关系，这说明两个 AS 可能存在混合的商业关系，在不同的 POP 点商业关系不同。于是他们利用大量 LC 探测点，利用 traceroute 对可能存在混合商业关系的 AS 链路进行定向探测识别 AS 之间的 POP 点，通过 BGP community、PeeringDB、DNS 以及商业数据库进行地理定位，最终得到 AS 之间不同 PoP 点的商业关系。实验结果发现在之前推断出的传统的商业关系的中有 4.5% 是复杂的商业关系。作者采用直接向网络管理员询问、BGP community 和 RIR 的 Whois 数据对推测结果进行验证，推测出的混合关系的准确率为 92.9%，推测出的部分传输关系的准确率为 97.0%。

C. 基于 *IXP* 数据

为了解决传统的商业关系推断中 P2P 链路大量缺失的问题，Vasileios 等人 [53] 提出了一种识别多边对等关系的算法。他利用 IXP 的 Route server 查询 BGP communities 的数据，推测在 IXP 建立的多边对等链路，BGP community 数据记录了在 IXP 互连的 AS 的互连策略，从互连策略中可以推测在 IXP 建立对等关系的 AS 对。除了进行主动的 BGP 查询以外，作者还利用被动的 BGP 数据分析来推断对等链路，利用靠近 IXP 的 BGP 收集器收集到的 BGP community 数据推测在 IXP 互连的对等关系链路。最终发现了 3.6 万条 p2p 链路。在此基础上，他们的跟进工作 [54] 不仅考虑了 IXP 的 router server，还利用了 IXP 及 IXP 的 looking glass 收集路由数据，最终从 13 个欧洲的大型 ISP 推测出了 20.6 万条 p2p 链路，这些 p2p 链路的数量相比于从 BGP 路由数据中推测到的 p2p 链路增加了四倍。[54] 补充了以往的算法所欠缺的大量 P2P 链路，推测得到的结果也被存储在 CAIDA 的 AS Relationships 数据库 [51] 中。

D. 基于概率迭代

尽管经典的 AS-Rank 已经达到了很高的准确率，但是在实际应用中 AS-Rank 推测出的商业关系的表现并不好。AS-Rank 的推测结果存在两个问题：一是它依赖的假设有问题。AS-Rank 假设度数最高的 AS 位于互联网的顶层、建立对等关系的 AS 之间的度数相同，研究发现 [55] 度数最高的两个 AS：AS6939 和 AS174 都不是 Tier-1 AS；大约有 14% 的 p2p 链路两端的 AS 的度数差别在 1000 以上，使得 p2p 关系难以与 p2c 关系分开。二是受到测量点的偏差和 snapshot 选择的影响，远离测量点的 AS 链路推测出的商业关系的错误率更高，同一个月的连续三十天的 snapshot 推测出的 AS 商业关系的准确率相差很大。观察发现不同 AS 链路的商业关系推测的难度并不相同，绝大多数的链路的商业关系都可以通过简单的算法推测出来；在过去的研究中每条链路的商业关系按照路由表中路由的顺序进行推测，如果之前的链路推测错误，很有可能会对之后的链路的商业关系推测造成影响。因此，Yuchen Jin 等人 [55] 首次提出了应该用概率迭代的方法推测链路的商业关系，每条链路的商业关系用概率表示，通过多轮迭代达到收敛的稳态。他们提出了一种基于朴素贝叶斯分类的 AS 商业关系推断算法——ProbLink。将每条 AS 链路抽象出三个特征，分别是：(1) 使用这条链路的路径结构 (2) 不使用这条链路的路径结构 (3) 链路两端的 AS 的属性。每条链路对应着一个类型概率的向量，这个向量包括链路是 p2p 关系的概率、链路是 p2c 关系的概率以及链路是 c2p 关系的概率，利用上述特征通过朴素贝叶斯分类算法进行多次迭代，更新每条链路的类型概率向量，最终到达一个收敛的状态停止。该算法相比于 AS-Rank 算法，错误率下降了 1.7 倍；对于几种难以推断的链路，推断的错误率下降了 1.8-6.1 倍；并且与传统的算法相比，受测量点的位置以及 snapshot 的选择的影响较小。最重要的一点，ProbLink 算法的推测结果在路由泄露的检测的应用中准确率上升了 3 到 4 倍。

Feng 等人 [56] 在 [55] 的启发下提出了一个概率模型推测商业关系。他们首先指出 AS 的商业关系是不确定的，这种不确定性取决于四个因素：一是由于推测的结果受到测量点的数量限制的影响。二是 p2p 链路很

难推断。三是主流的基于 valley-free 规则的算法不能回溯或者分支，无法包容 AS 商业关系的不确定性。四是一些假设不可靠，如 Gao 假设度数与 AS 的规模有关等。因此他们提出了一个不确定的 AS 商业关系推断算法，与 [55] 的 coretoleaf 算法类似，他们第一步利用 valley-free 规则进行初步的推断，对于剩余的未知链路采用不确定性的推断方法，每条链路可以用一个向量表示，向量记录了这条链路是各个商业关系的概率。每条链路的商业关系可以用布尔变量表示，链路与链路之间的商业关系存在着依赖。因此作者将不确定的商业关系推断转化为一个最大可满足问题，并利用启发式算法进行求解。

Zitong Jin 等人 [57] 在 [55] 的工作基础上进行了扩展，提出了一个称为 TopoScope 的算法。为了解决作为算法输入的路由数据带有的噪声（如路由泄漏和错配置）以及测量点的数量限制和偏差的影响，他们将路由数据分为几个组，每个组分别使用 [50] 的算法推测链路的商业关系，对于同一个链路，根据每一组的推测结果进行表决，对于 P2C 的链路推测采用多数表决原则，对于 P2P 链路推测采用少数表决原则。与 [55] 类似，这篇论文同样将结果分为可信链路和模糊链路两种，对于模糊链路，这篇论文采用贝叶斯网络和期望值最大化的算法进行推断，因为链路的各个特征不是彼此独立的，特征与特征之间存在依赖。为了避免测量点的限制和偏差导致的链路丢失，作者提出了一种推测丢失链路的算法。他们假设如果 AB 这条链路只被一个观测点观测到，那么这条链路就是一条容易缺失的链路，如果在其他观测点观测到一条 A-C-B 的路径，那么通过这个 A-C-B 的链路特征我们就可以提取缺失链路的特征，进而推测缺失链路。与目前的算法相比，TopoScope 的算法的错误率下降了 2.7-4 倍，发现了大约 30000 个缺失的 AS 链路，但是只有十分之一的链路在验证集中得到了证实。

E. 总结

图 7对域间链路商业关系推断的有关研究进行了总结。可以看到推断域间链路商业关系的研究以及持续 20 年，但仍然是一个热门的研究方向。这是因为推测域间商业关系具有很大的意义。除了能够对域间路由策略进行建模，推测未知的域间路由以外，商业关系还可以用来分析互联网的演化趋势 [58] [59]、识别恶意的 AS [60]、检测路由泄露等等。

分类	方法	推断的商业关系	验证集	准确率
基于valley-free规则	Gao[2000]	P2C+P2P+S2S	AT&T和Whois	99.1%
	SARK[2002]	P2C+P2P	路由数据违反VF的比例	N/A
	Battista[2003]	P2C	路由数据违反VF的比例	N/A
	Dimitropoulos[2009]	P2C+P2P+S2S	询问38个AS的管理人员	96.5%C2P 82.8%P2P 90.3%S2S
	Neudorfer[2013]	P2C+P2P	N/A	N/A
基于网络层次	AS-RANK[2013]	P2C+P2P	询问网络管理员、WHOIS、BGP community	99.6% C2P 98.7% P2P
	Complex[2014]	复杂商业关系	询问网络管理员、WHOIS、BGP community	92.9%和97.0%
基于IXP数据	INFOCOM[2013]	P2P	IXP的Looking Glass查询	94%
	MultiPeering[2013]	P2P	IXP的Looking Glass查询	97.1%-100%
基于概率迭代	ProbLink[2019]	P2C+P2P	BGP community	错误率下降1.7倍
	UNARI[2019]	P2C+P2P	BGP community	98.9%
	TopoScope[2020]	P2C+P2P	BGP community	错误率下降2.7~4倍

图 7: 域间链路商业关系推断总结

在过去的研究中，研究者都试图通过各种验证集证明自己提出的算法的准确性，如 BGP community、RIR 的 whois 数据库，甚至是直接询问管理人员。但是这些验证集都存在着或多或少的问题，比如有研究 [55] 指出使用 BGP communities 作为验证集进行验证是存在偏差的，在验证集中绝大多数的 AS 链路是与 Tier-1 AS 和测量点的 AS 的链路有关，这是因为公开的 BGP communities 的数据大多来自大型 AS，而不是测量点的 AS 经常在 BGP 路由的传播过程中把 BGP communities 丢弃。在第二章中我们介绍 RIR 的 whois 数据库时提到了 whois 数据的滞后性，使用 whois 数据作为验证集可能会存在很多错误。正是由于这些原因，尽管以往的商业关系推断的研究宣称的准确率很高，但是并不代表具有很高的应用价值。推测商业关系的准确性应该与实际应用结合，例如 [55] 的商业关系推测算法就将检测路由泄露的准确率从 20% 增长到 70%-80%。

V. 总结

随着互联网的发展，催生了很多对互联网进行大规模路由、流量分析的需求，迫切需要对互联网拓扑和路由进行完整、准确的感知。但是互联网不存在拥有全局信息的中心组织，我们只能基于各类数据对未知的拓扑路由进行测量和推测。在本文中，我们调研了域间链路发现与商业关系推断的有关研究。首先介绍了与域间链

路发现与商业关系推断相关的公开数据源，详细介绍各类数据源的内容及其优缺点。其次，调研了基于 IP 路径的域间链路发现的相关研究，详细列举了 IP 到 AS 映射的挑战以及学界目前的解决方案。第三部分介绍了域间链路商业关系推断的相关研究，按照算法对研究进行了分类，并按照时间的顺序记述了相关研究的发展脉络、关键挑战。

总结来看，域间链路发现与商业关系推断的研究有三个重要的思想。一是尽可能地运用更多的数据源，从数据源中挖掘可用数据可能会比改进算法对结果的提升大得多。二是多数表决原则。我们在对互联网中的某个实体的属性进行推测时，可能会受到测量数据的偏差影响或者因数据本身的噪声和异常带来错误，这时我们可以将数据进行分组，利用每个分组的数据分别对属性进行推测，最终按照多数表决推测结果，缓解单一数据源的偏差影响以及避免数据本身的噪声和异常的影响，这实际上是一种集成学习的思想。三是迭代原则。在我们对互联网的某个实体的属性进行推测的过程中，很有可能推测的结果会影响这一实体的相邻实体的属性，因此对于单一实体属性的推测尽量避免确定性的推测，而是采用多次迭代的方式对所有实体的属性进行推测，直至收敛，避免确定性推测的错误对其他实体的推测的影响。

获得一个完整而准确的互联网拓扑与路由依然是研究者所追求的目标。我们希望本篇文章可以为本领域之后的研究发展提供帮助。

参考文献

[1] J. A. Obar and A. Clement, “Internet surveillance and boomerang routing: A call for Canadian network sovereignty,” in *TEM 2013: Proceedings of the Technology & Emerging Media Track-Annual Conference of the Canadian Communication Association (Victoria*, 2012.

[2] J. Karlin, S. Forrest, and J. Rexford, “Nation-state routing: Censorship, wiretapping, and BGP,” *arXiv preprint arXiv:0903.3218*, 2009.

[3] A. Edmundson, R. Ensafi, N. Feamster, and J. Rexford, “Nation-state hegemony in internet routing,” in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, 2018, pp. 1–11.

[4] A. Houmansadr, E. L. Wong, and V. Shmatikov, “No Direction Home: The True Cost of Routing Around Decoys.” in *NDSS*, 2014.

[5] Y. Yang, X. Yin, X. Shi, Z. Wang, J. He, T. Z. Fu, and M. Winslett, “Inter-domain routing bottlenecks and their aggravation,” *Computer Networks*, vol. 162, p. 106839, 2019.

[6] J. M. Smith and M. Schuchard, “Routing around congestion: Defeating DDoS attacks and adverse network conditions via reactive bgp routing,” in *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018, pp. 599–617.

[7] “Routeviews.” [Online]. Available: http://archive.routeviews.org/

[8] “RIPE RIS.” [Online]. Available: https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris/ris-raw-data

[9] “Internet2 NOC BGP Rib Dumps.” [Online]. Available: http://ndb7.net.internet2.edu/bgp/

[10] “BGP Streaming.” [Online]. Available: https://bgpstream.caida.org/

[11] R. Motamedi, R. Rejaie, and W. Willinger, “A survey of techniques for internet topology discovery,” *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1044–1065, 2014.

[12] “isolario.” [Online]. Available: https://www.isolario.it/

[13] “CAIDA Routeviews Prefix-to-AS mappings.” [Online]. Available: http://data.caida.org/datasets/routing/routeviews-prefix2as/

[14] B. Donnet and O. Bonaventure, “On BGP communities,” *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 55–59, 2008.

[15] “AS286 Communities.” [Online]. Available: https://as286.net/AS286-communities.html

[16] “AS9002 bgp communities list.” [Online]. Available: https://lg.retn.net/bgp-communities.html

[17] “CAIDA AS to Organizations.” [Online]. Available: http://data.caida.org/datasets/as-organizations/

[18] B. Zhang, R. Liu, D. Massey, and L. Zhang, “Collecting the internet AS-level topology,” *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 1, pp. 53–61, 2005.

[19] G. Siganos and M. Faloutsos, “Analyzing BGP policies: Methodology and tool,” in *IEEE INFOCOM 2004*, vol. 3. IEEE, 2004, pp. 1640–1651.

[20] “PeeringDB.” [Online]. Available: https://www.peeringdb.com/

[21] A. Lodhi, N. Larson, A. Dhamdhere, C. Dovrolis, and K. Claffy, “Using peeringDB to understand the peering ecosystem,” *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 2, pp. 20–27, 2014.

[22] “CAIDA PeeringDB.” [Online]. Available: http://data.caida.org/datasets/peeringdb-v2/

[23] “Packet Clearing House (PCH) projects.” [Online]. Available: https://www.pch.net/

[24] “Euro-IX.” [Online]. Available: https://www.euro-ix.net/en/

[25] “CAIDA Ark.” [Online]. Available: http://data.caida.org/datasets/topology/ark/ipv4/probe-data/

[26] 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.

[27] J. Chabarek and P. Barford, “What’s in a name? decoding router interface names,” in *Proceedings of the 5th ACM workshop on HotPlanet*, 2013, pp. 3–8.

[28] B. Huffaker, M. Fomenkov, and K. Claffy, “Drop: DNS-based router positioning,” *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 5–13, 2014.

[29] M. Zhang, Y. Ruan, V. S. Pai, and J. Rexford, "How DNS Misnaming Distorts internet topology Mapping." in *USENIX Annual Technical Conference, General Track*, 2006, pp. 369–374.

[30] K. Keys, Y. Hyun, M. Luckie, and K. Claffy, "Internet-scale IPv4 alias resolution with MIDAR," *IEEE/ACM Transactions on Networking*, vol. 21, no. 2, pp. 383–399, 2012.

[31] 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*, 2008, pp. 337–342.

[32] R. Govindan and H. Tangmunarunkit, "Heuristics for Internet map discovery," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064)*, vol. 3. IEEE, 2000, pp. 1371–1380.

[33] K. Keys, "Internet-scale ip alias resolution techniques," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, pp. 50–55, 2010.

[34] H. Chang, S. Jamin, and W. Willinger, "Inferring AS-level Internet topology from router-level path traces," in *Scalability and traffic control in IP networks*, vol. 4526. International Society for Optics and Photonics, 2001, pp. 196–207.

[35] B. Huffaker, A. Dhamdhere, M. Fomenkov *et al.*, "Toward topology dualism: improving the accuracy of AS annotations for routers," in *International Conference on Passive and Active Network Measurement*. Springer, 2010, pp. 101–110.

[36] V. Giotsas, G. Smaragdakis, B. Huffaker, M. Luckie, and K. Claffy, "Mapping peering interconnections to a facility," in *Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies*, 2015, pp. 1–13.

[37] Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz, "Towards an accurate AS-level traceroute tool," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, 2003, pp. 365–378.

[38] 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," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 9, pp. 1822–1836, 2011.

[39] P. Marchetta, W. de Donato, and A. Pescapé, "Detecting third-party addresses in traceroute traces with IP timestamp option," in *International Conference on Passive and Active Network Measurement*. Springer, 2013, pp. 21–30.

[40] M. Luckie, "Scamper: a scalable and extensible packet prober for active measurement of the internet," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 2010, pp. 239–245.

[41] A. Marder and J. M. Smith, "Map-it: Multipass accurate passive inferences from traceroute," in *Proceedings of the 2016 Internet Measurement Conference*, 2016, pp. 397–411.

[42] M. Luckie, A. Dhamdhere, B. Huffaker, D. Clark, and K. Claffy, "Bdrmap: Inference of borders between IP networks," in *Proceedings of the 2016 Internet Measurement Conference*, 2016, pp. 381–396.

[43] A. Marder, M. Luckie, A. Dhamdhere, B. Huffaker, k. claffy, and J. M. Smith, "Pushing the boundaries with bdrmapit: Mapping router ownership at Internet scale," in *Proceedings of the Internet Measurement Conference 2018*, 2018, pp. 56–69.

[44] A. Marder, M. Luckie, B. Huffaker, and k. claffy, "vrfinder: Finding Outbound Addresses in Traceroute," in *Abstracts of the 2020 SIGMETRICS/Performance Joint International Conference on Measurement and Modeling of Computer Systems*, 2020, pp. 55–56.

[45] L. Gao, "On inferring autonomous system relationships in the internet," in *Globecom'00-IEEE. Global Telecommunications Conference. Conference Record (Cat. No. 00CH37137)*, vol. 1. IEEE, 2000, pp. 387–396.

[46] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz, "Characterizing the Internet hierarchy from multiple vantage points," in *Proceedings. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 2. IEEE, 2002, pp. 618–627.

[47] G. Di Battista, M. Patrignani, and M. Pizzonia, "Computing the types of the relationships between autonomous systems," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, vol. 1. IEEE, 2003, pp. 156–165.

[48] Y. Shavitt, E. Shir, and U. Weinsberg, "Near-deterministic inference of AS relationships," in *2009 10th International Conference on Telecommunications*. IEEE, 2009, pp. 191–198.

[49] L. Neudorfer, Y. Shavitt, and N. Zilberman, "Improving AS relationship inference using PoPs," in *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2013, pp. 459–464.

[50] M. Luckie, B. Huffaker, A. Dhamdhere, V. Giotsas, and K. Claffy, "AS relationships, customer cones, and validation," in *Proceedings of the 2013 conference on Internet measurement conference*, 2013, pp. 243–256.

[51] "CAIDA AS Relationships." [Online]. Available: http://data.caida.org/datasets/as-relationships/

[52] V. Giotsas, M. Luckie, B. Huffaker, and K. Claffy, "Inferring complex AS relationships," in *Proceedings of the 2014 Conference on Internet Measurement Conference*, 2014, pp. 23–30.

[53] V. Giotsas and S. Zhou, "Improving the discovery of IXP peering links through passive BGP measurements," in *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2013, pp. 121–126.

[54] V. Giotsas, S. Zhou, M. Luckie, and K. Claffy, "Inferring multilateral peering," in *Proceedings of the ninth ACM conference on Emerging networking experiments and technologies*, 2013, pp. 247–258.

[55] Y. Jin, C. Scott, A. Dhamdhere, V. Giotsas, A. Krishnamurthy, and S. Shenker, "Stable and Practical AS Relationship Inference with Problink," in *16th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 19)*, 2019, pp. 581–598.

[56] G. Feng, S. Seshan, and P. Steenkiste, "UNARI: an uncertainty-aware approach to AS relationships inference," in *Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies*, 2019, pp. 272–284.

[57] Z. Jin, X. Shi, Y. Yang, X. Yin, Z. Wang, and J. Wu, "Toposcope: Recover AS relationships From Fragmentary Observations," in *Proceedings of the ACM Internet Measurement Conference*, 2020, pp. 266–280.

[58] A. Dhamdhere and C. Dovrolis, "Twelve years in the evolution of the internet ecosystem," *IEEE/ACM Transactions on Networking*, vol. 19, no. 5, pp. 1420–1433, 2011.

[59] ——, "Ten years in the evolution of the internet ecosystem," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, 2008, pp. 183–196.

[60] M. Konte, R. Perdisci, and N. Feamster, "Aswatch: An as reputation system to expose bulletproof hosting ases," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 2015, pp. 625–638.