互联网 BGP 路由可视及安全检测技术架构与实践

叶朝阳[1]，沈辰[2]，黄明庆[3]，张士聪[1]，刘伊莎[1]
（1. 浙江省新型互联网交换中心有限公司，浙江 杭州 311200;
2. 中国信息通信研究院，北京 100191；3. 华为技术有限公司，北京 100095）

摘　要：边界网关协议（border gateway protocol，BGP）是支撑互联网 50 年来快速发展的核心协议，因早期设计考虑不足一直存在路由劫持、路由泄露等路由安全威胁漏洞。随着互联网应用日益深入，BGP 路由安全问题逐渐引起业界重视，边界网络安全防护意义重大。提出了一种 BGP 路由安全检测架构，通过推理构建全球 BGP 路由知识库实现互联网全局路由可视性，并基于此实现路由劫持、路由泄露等路由安全事件的准实时检测。通过在杭州交换中心部署实践，证明本系统可构造较完整的互联网全局路由知识库、实现较准确和实时的 BGP 路由安全事件检测。

关键词：BGP；路由安全；路由劫持；路由泄露
中图分类号：TP393
文献标识码：A
doi: 10.11959/j.issn.1000−0801.2021263

Architecture and practice of BGP internet routing visibility and security detection

YE Chaoyang[1], SHEN Chen[2], HUANG Mingqing[3], ZHANG Shicong[1], LIU Yisha[1]
1. National (Hangzhou) New-Type Internet Exchange Point, Zhejiang 311200, China
2. China Academy of Information and Communications Technology, Beijing 100191, China
3. Huawei Technologies Co., Ltd., Beijing 100095, China

Abstract: Border Gateway Protocol (BGP) is the de facto inter-domain routing protocol of today's global internet for exchanging routing information. However, it was supposed that all participants were reliable without generating routing security issues by mistakes or on purpose when BGP was designed 50 years ago. As Internet is getting involved in all aspects of our society, internet routing security is becoming the problems that couldn't be ignored anymore. A general architecture was proposed which coved inference of BGP routing knowledge database and provided visibility of global internet routing. Detection of route security events such as routing hijacks and routing leaks were realized. The deployment shows that the system can provide good visibility of internet routing and precise detection of routing security events.

Key words: BGP, internet routing security, routing hijack, routing leak

收稿日期：2021−08−24；修回日期：2021−12−08

1　引言

1.1　BGP 路由安全问题

以 BGP（border gateway protocol）为基础协议的全球互联网经过 50 多年的蓬勃发展，逐步从计算机互联网、消费互联网向产业互联网演进，成为全社会数字化基础设施，对安全可信的路由服务诉求越来越强烈，而早期 BGP 设计所带来的两个矛盾也越来越突出。一方面是缺乏安全可信机制。协议设计之初假设参与各方可信、可靠，没有考虑参与方的无意错误或恶意行为可能造成的路由安全威胁；另一方面是缺乏全局视图。各参与方只掌握自己相关的局部信息，缺乏全局视图以支撑互联网级别的可视运营、分析检测、诊断等，处理相关路由安全事故偏被动，且支撑信息和工具不足。

据路由安全相互协议规范（MANRS）统计，仅仅被监测到的路由安全事故每年高达数千起，其中一些重大安全事故更是造成了全球范围的影响。例如，2019 年 6 月，BGP 优化器进行基于精细路由的流量疏导，下游运营商错误地将该路由泄露给了 Verizon（美国无线运营商，威讯无线），Verizon 进一步扩散后放大该错误进而导致网络严重拥塞，直接影响了约 15%的互联网流量。类似的问题在 2010 年、2019 年也先后发生在中国电信，同样是因二次路由泄露导致网络严重拥塞，但不同的是，当时中国电信曾因被海外媒体污名报道而承受较大压力；2021 年 4 月，VDF 因错误通告原本属于 Google、Microsoft、Akamai、Cloudflare、Fastly 等大型科技公司的 31 000 多个路由，形成大面积路由劫持；2018 年 4 月，犯罪者通过劫持某以太坊钱包网站的 DNS 明细路由，窃取用户访问入口进而实施网络盗取。

所有这些路由安全事故都是因有意或无意的错误作用到 BGP 路由传播链条而形成，按照错误形成方式进行分类包括如下 3 方面。

- 前缀劫持：始发自治系统（autonomous system，AS）通告了本不属于自己的前缀。
- 路径劫持：传播扩散路由时伪造不存在的路径。
- 路由泄露：传播扩散时将路由泄露给了不合理的第三方。

这些路由安全事故可能造成的危害包括：路由黑洞导致的网络访问中断，流量绕行导致的网络拥塞、结算费用增加，更严重的包括利用路由劫持和泄露进行流量侦听、中间人攻击和仿冒攻击等。

1.2　相关研究

业界 BGP 路由安全相关研究可以最早追溯到 BGP 协议诞生之初，包括以下 3 个研究方向。

（1）协议层面的安全防护加固

早期，业界讨论的重点是从协议层面形成安全机制，考虑到性能开销以及现实部署等多方面的因素，目前已经得以应用的安全机制较为有限。比较常见的如 TCP MD5、GTSM（generalized TTL security mechanism）、路由抖动抑制（route flap damping，RFD）等。但此类方案无法解决路由劫持、泄露及路径篡改等当前域间路由安全核心问题。

（2）域间路由可信验证机制

针对 BGP 存在无法对路由信息进行真实性和完整性验证的问题，业界探索形成了通过带外建立可信任权威数据源的方式，推动域间路由从“无条件信任”向“可验证”方式演变。通过在带外建立互联网资源数据签发基础设施——即资源公钥基础设施（resource public key infrastructure，RPKI），为路由器提供可信的 BGP 路由安全验证防护所需的全局资源数据，从 2012 年起陆续发布了 BGP-ROV（route origin verification）和 BGP-PV（path verification）系列 RFC。近期在互联网名称与数字地址分配机构（the Internet corporation for assigned names and numbers，ICANN）/区域互联网注册机构（regional internet registries，RIRs）的

大力推动下，路由源签证（route origin authentication，ROA）数据签发及相应的 BGP-ROV 部署应用近年来获得了较快进展。尽管如此，离真正全面落实 BGP 路由安全协议加固防护仍然有较大距离，存在如下挑战。

- ROA 数据签发覆盖率仍然不足 30%，且在实际网络中 BGP-ROV 方案部署率更低，离全面落实路由起源验证仍然有较长的路要走。
- 围绕路径劫持和路由泄露的安全防护，业界提出了一些方案，但都未形成共识，其主要制约因素有：BGP 的路径迭代验证带来大量计算开销，业界还在探索路径验证能力与计算开销之间的平衡方案；路径验证数据的签发，特别是商业关系数据，相较于 ROA 更涉及运营商的商业隐私。
- RPKI 基础设施全面部署和应用，本身也存在挑战：除相关基础设施的建设投资和运营投入外，RPKI 的中心化签名认证和分发机制还涉及中心化治理风险、与底层基础网络间的数据同步问题等。

（3）BGP 路由安全分析检测

如果说前者是主动的 BGP 路由安全防护，BGP 路由安全分析检测则属于后端被动监测，二者共同构成完整的防护体系。BGP 路由安全分析监测的核心价值包括如下 3 个方面。

- 主动防护只能基于本地路由器接收到的协议报文进行，并不能有效防范发生在外部的针对本网始发前缀的劫持和泄露危害。
- 主动防护方案受部署场景覆盖及所依赖数据的完备性限制，预期在相当长时期内都无法依赖其实现完整的路由安全保障。
- 互联网级别的网络运营监管，需要打造全网可视化视图，进行实时安全威胁分析及态势感知，为进一步的故障定位、诊断、消减等提供支撑。

近 20 多年来，BGP 路由信息收集及安全分析检测一直是互联网生态圈学术研究的热点之一，特别是在美欧地区，逐步建立起了初步的用于全球互联网路由信息采集及安全分析能力。

（1）互联网路由信息采集

传统上，互联网路由运维主要依赖运营商共同发起提供的 Looking Glass Servers。20 世纪末，美国和欧洲分别发起了以 Route Views 和 RIPE RIS 项目为主的 BGP 路由信息采集公共基础设施建设——通过专用采集器（collector）与现网 BGP 路由器采集点（vantage point，VP）建立对等（peer）关系单向获取 BGP 路由信息，并开放提供给业界进行进一步的数据分析利用。截至目前，Route Views 和 RIPE RIS 在全球共计建设了超过 50 多个路由采集器，从全球数百个 AS 提取 BGP 路由信息。

（2）BGP 路由数据分析研究与应用

基于上述基础设施所采集到的路由信息，业界展开了 BGP 路由数据分析相关研究与应用，其中最有代表性的组织是 CAIDA（Center for Applied Internet Data Analysis），其组织和支撑的研究覆盖数据采集、分析、可视化、分层共享等各相关环节以及商业、教育、研究和政府组织等产业生态。概况起来，这一领域的分析研究有两大主要方向：全球互联网基础知识库的推理构建以及路由劫持、路由泄露等路由安全事故的检测分析。其中，知识库包括前缀起源、AS 拓扑、AS 邻居商业关系等，基于此可构造全球互联网级别的全局数据视图，提供互联网运营维护支撑的同时，也为进一步的路由安全检测提供了支撑。总体上业界形成了两大技术路径：以知识库推理为基础的逻辑推理方法和以 AI 算法为核心的路由异常监测大数据分析方法。

1.3 本文研究方向

相对来说，BGP 路由安全分析监测领域在我国总体上还处于空白状态，一个客观原因是当前我国 BGP AS 无论规模还是开放互联互通程度与

美欧相比有着相当大的差别（例如，美国有近 2 万 AS，大量通过互联网交换中心（internet exchange point，IXP）互联互通；我国 AS 数量不到 2 000，主要通过三大运营商交换中心进行互联互通）。面向未来，本文认为有必要加强如下能力的研究和储备。

（1）我国网间互联架构持续优化，工信部于 2019 年正式批复国家（杭州）新型 IXP，交换中心的运营主体——浙江省新型互联网交换中心有限责任公司，于 2020 年成功组建，鼓励企业使用自主 AS/IP 接入网络，打造开放互联的网络新生态。运营商之间的骨干直连点仅面向基础电信运营商开放，交换中心则面向更多企业和机构提供流量交换服务，如本地互联网接入服务商、互联网内容提供商、云服务商、工业互联网企业、科技网、教育网等，汇聚了大量具备自主 AS/IP 的网络，网络之间通过 BGP 对等互联。因此，在多边复杂流量交换的背景下，一旦发生 BGP 安全事件，将会影响大量企业和流量，交换中心的网络安全防护意义重大。

（2）互联网级别的 BGP 路由安全威胁感知能力。当前国内主要聚焦于单网/单设备的自身配置触发的 BGP 路由安全防范，而随着互联网在社会生活中的深入渗透，互联网级安全威胁感知能力愈发重要。BGP 入口消息潜在安全威胁的检测能力，以帮助基础电信运营商尽快发现和界定故障、避免因进一步消息扩散而导致的放大效应；针对自身重要业务前缀在全球范围的安全威胁感知能力。在跨国业务和交易日益发展的背景下，存在对这类重要前缀可能发生在外部的劫持和泄露感知的需求。

（3）随着我国信息化战略推进，加快 IXP 部署脚步以支持更开放高效的云网融合架构已提上议事日程，未来我国 BGP AS 规模及互联互通复杂度必然将大幅度提升。随着越来越多关乎国计民生的业务走向线上，相应的全局可视化视图及事故检测分析能力必然成为该新型数字化基础设施的核心能力要求。

本文主要方向如下。

（1）本研究采用基于知识库推理的逻辑推理技术路线。通过推理算法构建起全球互联网三大核心知识库：前缀起源、AS 拓扑及邻居商业关系。其中，前缀起源知识库可覆盖 100%全球骨干路由，AS 拓扑及邻居商业关系知识库可覆盖全球互联网较高层级 AS（覆盖度依赖于路由信息采集器部署密度和位置。一般来说，AS 层级（tier）或传送度（transit degree）越高，其对应相关知识库覆盖程度越好）。相对于路由安全主动防护所依赖的 RPKI 数据库中约 30%覆盖率的 ROA 数据（对应前缀起源知识库）、受标准进度影响还是空白的 AS 路径相关数据（对应 AS link 拓扑及邻居商业关系知识库），本研究所提供的互联网知识库是对资源公钥基础设施（resource public key infrastructure， RPKI）资源签名记录的极大补充，尽管暂不能直接应用到协议主动防护（需确保记录 100%可信），但对构建全球互联网可视化视图、BGP 路由安全事件的准实时检测仍然有巨大意义，甚至未来通过知识库记录可信度管理，可直接作为 RPKI 数据的补充而应用于主动防护。

（2）本研究基本算法以时空稳定度、贝叶斯概率推理为核心推理构建上述三大知识库，但要进一步提升准确度，必须对相关干扰因素进行深入研究。主要有两类干扰因素，一方面，BGP 所允许的合法“例外”存在，例如，因任播（anycast）、前缀聚合、分布式拒绝服务攻击（distributed denial of service，DDoS）攻击防护等需要而存在的合法多起源冲突；除一般的 lateral peering、transit provider 商业关系之外存在的 sibling、partial transit、hybrid 等特殊商业关系；一些个别私下协商长期合法存在的 valley-path（违反 BGP 路由扩散路径的一般原则——valley-free）。另一方面，因 BGP 路由采集点覆盖不足导致的 BGP 基础路由信息采样偏差，相应地可能使得一些数据特征提取可能偏差。本研究通过针对性的合法例外场景

识别和分析、数据偏差下的数据调测处理，较大幅度提升了推理结果的准确度，均达到 99%以上。

（3）推理分析结果的验证校准是评估算法、持续推动算法优化的关键一环，除了抽样式人工比对验证，对于类似于本案例的较大规模数据分析输出，基于一定规模的基准库进行整体量化评估非常重要。本研究的三大知识库中，前缀起源基准库可基于业界已经具备的 RPKI ROA 数据集构建，但 AS link 拓扑、邻居商业关系基准库业界还缺乏公认标准，因此本文对如何构建 AS link 拓扑、邻居商业关系基准库进行了初步尝试。其中，AS link 拓扑基准库的构建本文基于两个核心判断：BGP 路由信息采集点（vantage point，VP）所输出的直接邻居是可靠可信，且在 full-feeding 模式下（采集器与目标采集点的 BGP 邻居关系为 C2P，通告所有路由）能采集到该 VP 采集点 AS 所有邻居；而 AS 邻居商业关系基准库的构建，则是通过部分 ISP 在其网站公开的 BGP 团队属性路由策略邻居关系信息语义挖掘完成。

2 BGP 路由安全分析检测系统

2.1 系统总体架构

BGP 路由安全分析检测系统总体架构如图 1 所示，共有三大关键模块组成：数据采集与预处理、互联网知识库推理、路由安全事故检测。

本系统整体采用逻辑推理方法，核心是互联网知识库推理构建，在此基础上经相对简单的推理即可实现路由安全事故检测。相较于一些采用 AI 算法直接基于大数据完成路由安全事故检测的研究，本系统所基于的知识库逻辑推理，一方面更符合运维人员工作逻辑，有利于事故检测的同时提供更多具体的故障定位支撑；另外，其所构建的互联网全局知识库，类似于为互联网打开了交通地图，为数字世界基础设施的高效运维提供了重要底座。

2.2 数据采集与预处理

2.2.1 数据采集

数据采集主要涉及以下 3 类。

（1）VP 点 BGP 路由信息，包括 RIB 信息以及 BGP update 消息。这些信息通常由 Route Views 和 RIPE RIS 等公共采集器通过与路由器建立 BGP 会话的方式获取，并每隔一段时间压缩为 MRT 文件存档，供外部使用者通过 HTTP get 获取。

（2）交换中心 BGP 路由信息，使用服务器与 RS（route server）建立 BGP 邻居，基于开源路由软件采集路由前缀、起源 AS 以及 AS path 等关键数据。

（3）其他辅助数据源，如 Peering DB、IRR DB、RPKI ROA、Tier 1 ASN 清单、保留 ASN 号清单等。这类数据大多无须实时提取，但仍需要根据各自特点保证数据时效性。

2.2.2 数据预清洗与预整理

数据预清洗最重要的目的是过滤掉无效数据，以避免这些数据流入后续处理环节、影响分析结果

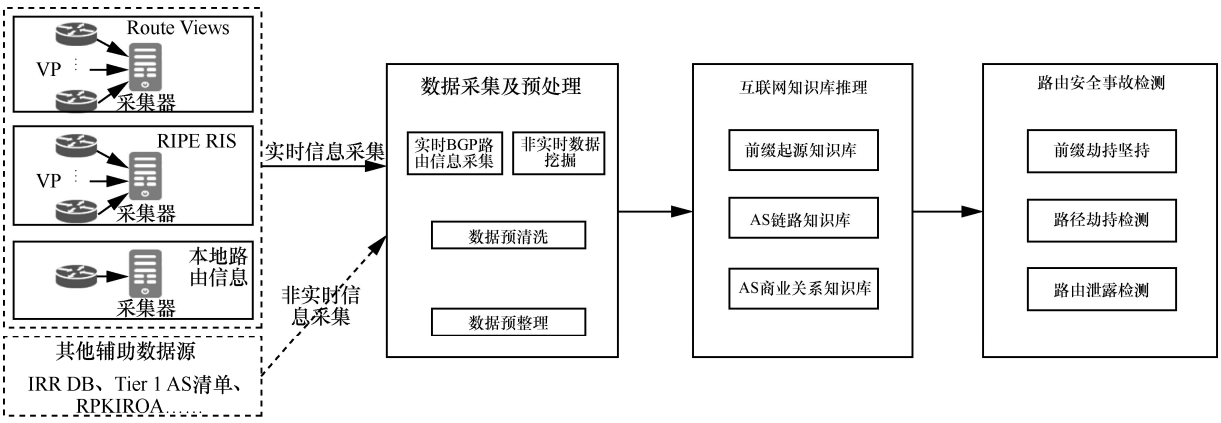


图 1 BGP 路由安全分析检测系统总体架构

的准确性。例如，根据是否存在 AS 环路、是否涉及保留 ASN、是否出现了非连续 Tier1 ASN 等，可以将一些 AS path 作为无效数据提前加以识别清洗。

因涉及持续大量的数据更新和迭代计算，数据预整理成为保障高效、准实时分析输出必不可少的重要环节。本系统从原始 BGP 路由中提取前缀起源、AS path 信息及其相关特征（如时间特征、空间特征，其中时间特征为观测到该记录的时戳信息，空间特征为观察到该记录的观测点信息）进行数据库整理，并建立时间跨度滑窗机制（窗口大小可配置，1 周到 3 个月不等）持续进行数据更迭。如此可大幅度提升系统 I/O 效率、大幅度提升迭代计算的实时性。

2.3　互联网路由全局知识库推理

互联网路由全局知识库就像是数字世界基础设施的交通地图，是互联网运维的重要支撑，其全局知识库主要包括：前缀起源（站点）、AS Link 拓扑（路网）、AS 商业关系（方向指示）三大知识库组成。因互联网 BGP 各参与方只掌握局部信息（本地及由邻居独立选择后部分通告的信息），如何能基于这些局部信息推理构建出较完整的互联网全局视图一直是业界的重点研究方向之一。

2.3.1　前缀起源知识库推理构建

前缀起源信息记录特定前缀的始发 AS 信息，其知识库推理构建算法包括如下两个关键部分。

（1）基于时空稳定度，构造前缀起源基础数据库

这里时间稳定度指的是在特定时间窗内，能持续观测到该前缀起源的时间模型；空间稳定度指的是能观测到该前缀起源的观测点分布情况。基于时空稳定度模型，系统可每对前缀起源记录进行稳定度打分，并得到相对稳定的前缀起源数据库记录。

（2）多起源 AS（multiple origin AS，MOAS）冲突清理

MOAS 指的是特定前缀有多个起源 AS，多起源冲突清理就是需要区分哪些是合法的多起源冲突、哪些是潜在的前缀劫持。系统首先排查出相对稳定的 MOAS 记录作为长期合法存在的多起源纳入知识库，并进一步对瞬态 MOAS 冲突进行清理，识别合法 MOAS 冲突。具体清理方法依赖于不同的合法 MOAS 冲突场景，例如，anycast 服务的多起源通告；较常见的 provider AS 为 customer AS 代为始发通告，或者进一步的路由聚合；sibling AS（指从属于同一运营商，二者之间互为 transit provider 的 AS）之间互相始发对方的前缀；DDoS 防护服务商利用类似前缀劫持方式引导流量等场景。对于合法 MOAS 冲突，可进行标注后计入知识库；而未被识别合法 MOAS 冲突的瞬态记录，则作为疑似前缀起源劫持而被排除在知识库记录之外。

2.3.2　AS link 拓扑知识库推理构建

本知识库记录格式为：<Prefix, Origin AS, 时间维度信息, 空间维度信息>。其中，时间维度信息为推理分析窗口时间范围内（一般为“数周”时长的滑窗）观测到对应前缀起源记录的次数；空间维度信息推理分析窗口时间范围内能够观测到该起源记录的采集点数量。

AS link 拓扑知识库的推理构建总体上也是基于时空稳定度，也即所观测到的 AS link 的时间稳定度和空间稳定度。进一步的记录清洗处理主要需要应对好两个挑战。一方面，对于相对不稳定的 AS link 记录处理，如何将 backup link、正常的拓扑调整（如增删链路）等存在一定瞬态表现的记录与存在 link 伪造的路径劫持记录进行区分。另一方面，如何避免 IXP RS 等不同的处理方式对 AS link 记录的干扰（RS 在进行 BGP 控制消息处理时最新的推荐做法是不在 AS Path 中记录 IXP 自身的 ASN，但较早存在不少类似情况，进而对 AS link 的判断形成一定干扰）。

本知识库记录格式为：<AS1, AS2, 时空稳定度>。其中，AS1 与 AS2 为互为邻居的两个 AS，

不区分先后。时空稳定度为推理分析窗口时间内综合评估空间和时间维度信息得到的 AS 邻居关系稳定度评估值。

2.3.3 AS 邻居商业关系知识库推理构建

AS 邻居商业关系知识库推理构建相对来说是所有知识库构建中挑战最大的，其主要体现在两个方面。其一是 AS 邻居商业关系的多样性。除了最为常见的 P2C（provider to customer）和 P2P（peer to peer）商业关系，还有多种复杂商业关系：sibling to sibling（属于同一运营商，表现为互为 transit）、hybrid（两个 AS 之间存在两条或以上 link，且不同 link 采用了不同商业关系）、partial transit（相对于 full transit，其并不将上游 transit provider 路由进一步下发给 customer）等。其二是商业关系及其路由扩散策略尽管有一般性要求和规范，但本质上由运营商协商和定义，基于一般性规范推导的结果必然存在偏差。

（1）一般商业关系知识库推理构建

AS 邻居商业关系推理构建相关研究大多建立在 Lixia Gao 的 Valley-free 理论之上，即认为有效的 AS path 应该是“任意数量 C2P link + 0/1 个 P2P link + 任意数量 P2C link”的组合。ProbLink 是业界最新有关一般商业关系推理的研究。首先，基于 AS rank 算法推理得到 Top Clique（Tier 1 ASes）；然后，依据 Valley-free 原则推导 P2C 链路，继而将剩余链路标记为 P2P 链路，形成初始邻居关系推理结果。受观测限制（观测点数量及观测点位置）以及部分存在的 Valley-path 情况，上述推理必然存在冲突或错误。针对这种情况，ProbLink 提取 AS path 主要特征（如 link triplet、non-path、distance to clique、vantage point、co-located IXP 等），通过朴素贝叶斯概率推理来进一步求解 link 邻居商业关系的最大可能性。

本系统主要以 ProbLink 为基础进行一般商业关系推理，除了过程中的贝叶斯参数调测，主要进行了如下方面的调整。

- 互联网 Tier 1 AS 清单可公开获取且相对来说变更不是太频繁，本系统尝试用静态输入的 Tier 1 AS 清单取代 Top Clique 推理算法，发现推理结果有更好的准确率表现。
- 朴素贝叶斯算法要求特征相互独立，但 ProbLink 所提取的特征实际上并不完全独立，且在特征提取时没有考虑到 P2P 观测量会远小于 P2C 的情况。本系统特征组合、特征增加、特征参数等进行了调试优化，发现可以进一步提升算法准确率。
- 为了提升算法效率，同时也避免错误的逆推理（把原本正确的结果推导为错误的结果），本文对进入第二步概率推理的 link 范围进行了进一步的筛选：基于第一步所得到的推理结果，如果某 link 出现在不同观测点 AS path 数量足够，对其出现违反 Valley-free 的情况进行信用（credit）计分，最终筛查出相对可信的推理结果，不用重复进行概率推理。

（2）复杂商业关系知识库推理构建

复杂商业关系尽管整体占比较低，但对进一步提升商业关系知识库构建准确度有着较大影响。主要存在如下复杂关系类型。

- sibling：其推理逻辑主要是通过挖掘与 AS 相关的注册登记信息判断，如对应单位、管理员的名称/姓名、地址、邮箱、电话等。
- partial transit：其推断可通过对现有 P2C link 进行进一步 full 或 partial transit 的筛查，对于出现了传递上游 provider 路由到下游 customer 的为 full transit，否则为 partial transit。
- hybrid：hybrid 的推理相对比较复杂，其核心推理支撑是需要判断两个 AS 之间存在两条独立的 link 且位于不同地理位置，直接的方案是借助 traceroute 工具探测得到对

应 link 的不同 IP 地址，再通过 IP 地址信息挖掘获得其 AS 和 POP 地址位置，从而判断是否存在异地双 link 情况。该方案准确性高，但工程部署难度较高（包括 IP 地址 POP 位置挖掘技术挑战及计算量），最终本文采取了基于 PeeringDB 挖掘基于 IXP 的 P2P link，如推理同时存在 P2C，则为 hybrid。

本知识库记录格式为：<AS1, AS2, 邻居关系类型>。其中，AS1 与 AS2 为互为邻居的两个 AS，邻居商业关系类型表达：0 — P2P；-1 — P2C；2 — Sibling；3 — partial transit；4 — hybrid。

2.4　BGP 路由安全（准）实时检测

一旦推理得到上述完整可信的互联网全局知识库，BGP 路由安全检测的逻辑就相对比较简单：从网络中提取 BGP updates 消息，与三大知识库展开分析比对即可快速判断潜在前缀劫持、路径劫持和路由泄露事件。为提高准确度和实时性，本系统进行了两方面的优化。

（1）快速迭代更新知识库，提升知识库时效性、准确性

从 BGP updates 中提取路由信息，除了检测路由安全事情外，还需持续迭代到知识库推理中。互联网是持续更新的系统，如前缀的交易迁移、链路增删、网络建设调整等，系统需要将这些正常的变更进一步与路由安全威胁区分开来，就需要更及时地结合 BGP updates 消息进行知识库迭代管理，而不仅仅基于间隔几个小时采集到的 RIB 信息。

（2）基于本地的实时检测分析

当前 Route Views 和 RIPE RIS 每间隔 5 min 进行一次 update 消息压缩整理供外部提取、检测分析（在考虑后续支持 live stream）。本系统当前可基于获取的本地数据源进行实时的检测分析。

3　实验结果评估

3.1　环境搭建及基准库构建

3.1.1　环境搭建

本系统在国家（杭州）新型互联网交换中心部署，除了通过公网获取 Route Views 和 RIPE RIS 等的全球公共路由信息外，还通过服务器（基于开源路由软件）与交换中心 RS 建立 BGP 邻居关系，单向获取 RS 上的 BGP 路由信息（不反向传递任何路由信息，避免干扰现网），作为本地数据补充。具体实验环境如图 2 所示。

3.1.2　基准库构建

基准库构建是算法调校、结果评估验收非常重要的依据。互联网路由安全分析领域基准库构建一直是业界在探索的难题，特别是跟路径和邻居商业关系相关的基准。

（1）前缀起源基准库

从数据来源来看，当前国际公开权威的前缀

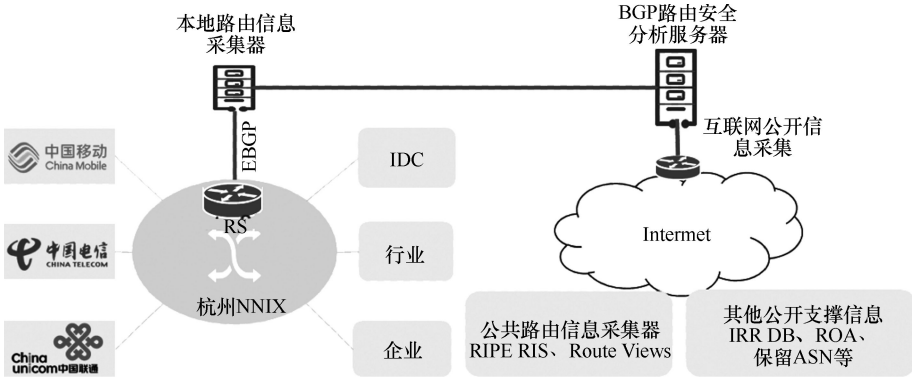


图 2　杭州 NNIX 实验环境

起源数据主要来源于IRR、ROA数据。截至2021年7月，IRR注册覆盖的起源前缀信息90.7万条，ROA签发起源前缀信息33万条。同时，由中国信息通信研究院牵头建设的国内首个路由权威数据源目前已覆盖国内138个自治域网络的路由信息，具备国内超过2.9万条权威前缀信息。

（2）AS link基准库

一般来说，BGP路由VP采集点所在AS如果以full feed方式（将collector作为其customer provider）向collector传递全量路由时，所得到的路由信息应该能全面真实反馈该AS的邻居情况。基于这样的判断，摘取full feed VP采集到的路由信息，以VP AS为原点收集其所有邻居AS信息，作为AS link基准库。本系统通过此方法，建立起了大约25万条link基准库记录。

（3）邻居商业关系基准库

运营商可通过BGP community传递自定义的相关路由策略，其中包括商业关系相关策略（部分运营商的实践）。但具体策略描述所代表的语义由各运营商自行定义，所以并不能简单通过community相关属性挖掘完成邻居关系基准库的建立。但好消息是，部分运营商会在其website中对其community值所代表的含义给予说明，本系统正是通过挖掘相关网站得到相关语义说明，从而构建出邻居关系基准库。通过这个方法，本文成功挖掘出约7万条商业关系基准库。

基于基准库，本文尝试进行算法查全率和准确率的评估，其中查全率（recall）=TP/(TP+FN)，准确率（precision）=(TP+TN)/样本总数。其中，TP（true positive）代表正确检出的错误数，FN（false negative）代表被漏报的错误，TN（true negative）代表没有被误报的正常结果。

3.2 实验结果

经过3个月的互联网BGP路由信息采集，经过数据预处理后，共得到约100万条前缀起源记录、5 000万条全球AS path记录，并建立了各记录的时间、空间等相关信息，动态管理机制。各知识库推理结果、路由安全检测相关结果如下。

（1）前缀起源知识库及前缀劫持检测

前缀起源支持库推理构建结果如图3所示，共计得到IPv4前缀起源记录1 037 058条，基于ROA基准库的108 354 IPv4前缀，前缀起源匹配度>99%。对前缀劫持检测结果进行抽查对比分析。例如，2020年10月5日检测发现1条前缀劫持，而同期BGPStream发布两条异常。对其中BGPStream有告警而没有告警的记录进行分析发现，该现象的发生是公有云服务商向其他公司转租子前缀，导致出现类似子前缀劫持现象，而本系统通过引入时空稳定度可较好规避此类告警。

图3 前缀起源支持库推理构建结果

（2）AS link知识库及路径劫持检测

AS link知识库推理构建结果如图4所示，推理得到稳定的AS link知识库记录共计有529 337条，但同时还观测到约有1万条稳定度不足的link，这些非稳态link的产生原因需持续观测和进一步分析。在进行准确度评估时发现，基于VP观测点构造的link基准库评估的准确度接近100%，本文认为可能有两方面的原因：根据业界研究，基于link伪造的路径劫持发生比例远低于其他类型路由安全事件，更重要的是，VP观测点所在AS通常位置较重要、运营较规范，出现与VP AS相关的链路伪造可能性更低。受link记录中存在的不稳定情况影响，因暂时还未落实进一步清洗判断（是否为正常的link变更或backup link临时切换），一个直接的影响是基于当前知识库检测到的路径劫持告警可能存在一定的假阳性。

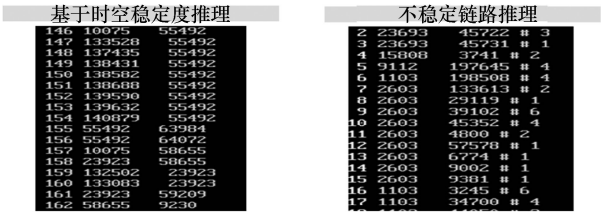


图 4　AS link 知识库推理构建结果

（3）邻居商业关系知识库及路由泄露检测

AS 邻居商业关系知识库推理结果如图 5 所示，P2C 链路 150 228 条，P2P 372 143 条；sibling 319 条，hybrid 和 partial 各约 4 000 条，见表 1，从一般商业关系推理结果基于基准库的评估来看，两种 link 关系类型的查全率和准确率都超过>99%（复杂商业关系因基准库样本不够没有评估），相比于经典 ASRank 和 ProbLink 算法都有不同程度提升。

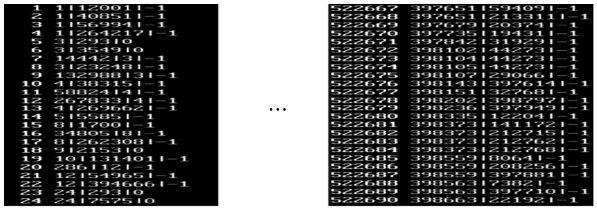


图 5　AS 邻居商业关系知识库推理结果（一般商业关系）

表 1　一般商业关系推理结果评估

链路类型	TP/条	TN/条	FP/条	FN/条	查全率	准确率
P2C	33 870	39 413	94	315	99.08%	99.72%
P2P	39 400	33 872	313	107	99.73%	99.21%

值得注意的是，基于邻居关系知识库推理结论来审视 AS path，共计有 0.5%的 link triplet 三元组、5.7%的 AS path 违反了 valley-free 原则。这其中最主要的因素应该是长期合法 valley-path 的存在，其他原因包括复杂商业关系、一般商业关系推理误差。在进行路由泄露检测时，重点要排除长期合法存在 valley-path 的干扰。

4　结束语

前缀劫持、路径劫持和路由泄露等 BGP 路由安全事故多年来高发不下，每年都会发生若干起全球性重大安全事故。随着互联网日益发展成为全社会数字化核心基础设施，互联网全局路由视图及实时路由安全检测日益重要。相对单系统/单网络的路由安全保障更聚焦于避免自身错误配置导致路由安全，互联网全网级别的路由安全分析检测还可以对入口消息进行检测，发现潜在路由劫持和泄露等安全威胁，可为故障早发现、早定位、早隔离提供实时支撑，避免二次放大造成更严重危害。此外，还可以对重点前缀（如 DNS 等基础设施、关乎国计民生的关键服务设施等）进行全球全网路由安全威胁主动监测，这在业务开展日益国际化、日益分布式的今天意义更为重大。

本系统采取逻辑推理方法，首先通过推理构建了全球互联网的三大知识库：前缀起源、AS link 拓扑、邻居商业关系，进而基于三大知识库实现（准）实时路由安全事故检测能力。基于此，相当于有了互联网基础性的“交通地图”—路网信息及事故监测。本系统基于时空稳定度、贝叶斯推理等核心算法，通过针对性梳理和识别容易对知识库和检测分析造成干扰的正常场景，包括正常的多起源冲突场景、复杂商业关系场景、backup link 及临时 link 变更场景等，以进一步提升知识库推理和故障检测的准确度；针对数据采集偏差特点、计算开销导致的及时性等，本系统在特征组合、特征设计和调参，提升算法推理准确度的同时降低了工程部署难度。通过基准库构建及评估、与业界最佳实践的分析对比等，证明本系统在查全率、准确率及可用性等方面都达到了业界先进水平。与此同时，本文认为互联网全局路由的可视度、安全检测分析查全率和准确率等都还有较大提升空间。

（1）更加丰富的数据采集生态

从数据推理结果来看，AS link 的可视化程度仍然还有较大差距，特别是 P2P link 的传播特点决定了需要更深更广的采集点覆盖，其提升 P2P 链路可视化程度的同时，也必然可改进 P2C 与

P2P 误判比例，降低路由泄露误判情况。同时，更多的局部和本地信息采集可增加知识库信息的明细程度。

（2）数据面探针基础设施建设

当前的研究聚焦于控制面的大数据分析，未来可考虑进一步结合数据面 ping/traceroute 探针，一方面可以对控制面分析结果给予验证和辅助支撑，例如对前缀可达性和路由路径的验证，对需要结合 POP 位置的相关推理支撑。更核心的是，该数据面探针系统还可提供更广泛的数据面可视化支撑，与本系统控制面可视化共同构成互联网可视化的关键两个要素。

（3）路径劫持检测告警的准确度提升

主要是如何区分非稳态 link 与伪造 link 的研究。实践中，可考虑增加本地或区域邻居信息，以加强跟本地或区域 AS 相关路径伪造的检出准确性。

（4）数据分析积累

这对基于数据分析的研究意义重大。包括互联网也在动态变化，需要持续积累数据、优化算法，不断寻找最优解。BGP 自主对等参与的特点决定了一般规则之外的特例存在，例如 valley-free 是建立在网间流量结算最优的一般规则上，但基于业务体验、网络负载等考虑运营商可部分调整；云服务商在生态中的地位日益提高，基于网络层级的商业结算一般规则也越来越多的出现特例。数据分析需要对这些一般规则之外的特例进行长期积累，进行白名单管理。

参考文献：

[1] ZHAO X L, PEI D, et al. An Analysis of BGP Multiple Origin AS Conflicts[C]// Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement 2001. New York: ACM Press, 2001.

[2] CHIN K W. On the characteristics of BGP multiple origin AS conflicts[C]//Proceedings of 2007 Australasian Telecommunication Networks and Applications Conference. Piscataway: IEEE Press, 2007: 157-162.

[3] LUCKIE M, HUFFAKER B, DHAMDHERE A, et al. AS relationships, customer cones, and validation[C]//Proceedings of the 2013 conference on Internet measurement conference. New York: ACM Press, 2013.

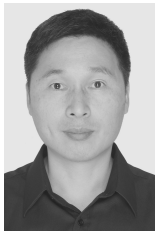
[4] LI Y C, SCOTT C et al. Stable and Practical AS Relationship Inference with ProbLink[C]//Proceedings of 16th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 19). 2019: 581-598.

[5] GIOTSAS V, LUCKIE M, et al. Inferring Complex AS Relationships[C]// Proceedings of the 2014 conference on Internet measurement conference. New York: ACM Press, 2014.

[6] FENG G Y, SESHAN S, STEENKISTE P. PARI: a probabilistic approach to AS relationships inference[EB]. 2019

[7] JIN Z T, SHI X G, YANG Y, et al. TopoScope: recover AS relationships from fragmentary observations[C]// Proceedings of the ACM Internet Measurement Conference. New York: ACM Press, 2020.

[作者简介]



叶朝阳（1976− ），男，浙江省新型互联网交换中心有限责任公司总经理、中国互联网协会互联网互联互通工作委员会副主任委员，主要研究方向为新型互联网交换中心网络架构与协议设计、云网交换等。

沈辰（1989− ），中国信息通信研究院工程师，主要研究方向为互联网网络互联互通、互联网路由安全、互联网测量与性能分析等。

黄明庆（1969− ），男，华为技术有限公司高级 IP 技术研究专家，主要研究方向为网络空间安全、互联网协议架构等。

张士聪（1990− ），男，浙江省新型互联网交换中心有限责任公司技术部经理，主要研究方向为新型互联与网络架构。

刘伊莎（1992− ），女，浙江省新型互联网交换中心有限责任公司 IT 工程师，主要研究方向为新型互联与网络架构信息化。