

中国重点域名权威服务器服务状态测量与分析

聂乐遥^{1,2}, 李城龙^{1,2,3}, 董 聪², 宋光磊², 张 辉^{1,3}, 樊琳娜⁴, 杨家海^{1,2,3}

¹清华大学网络科学与网络空间研究院 北京 中国 100084

²中关村实验室 北京 中国 100094

³泉城实验室 济南 中国 250215

⁴国防科技大学信息通信学院 武汉 中国 430019

摘要 域名系统权威服务器的可用性对于网络服务的正常运行至关重要, 权威服务器不可用会直接导致域名不可用, 影响用户访问网站、发送电子邮件等正常网络活动。俄乌冲突爆发后, 域名系统的服务可用性受到了地缘政治的影响, 导致以.ru为代表的俄罗斯相关域名的服务运营被迫迁往本国基础设施。然而, 目前还没有公开研究全面分析中国域名权威服务器的配置和部署情况, 缺少对中国重点域名权威服务器的可用性风险评估研究。面对断网停服等全球网络空间安全威胁, 迫切需要对国家网络空间重点领域域名开展态势感知。本文从政府、教育、企业三个领域着手收集整理了中国重点领域的域名作为数据集, 然后对上述重点领域域名进行了全面的态势测量, 包括基础设施静态属性测量、基础设施共享测量、权威服务器部署方式测量、无效权威服务器测量、基础设施境外依赖度测量和顶级域名境外依赖度测量, 从配置有效性、部署方式、地区依赖三方面进行数据分析, 揭示了基础设施托管域名过多形成单点瓶颈、权威服务器无法提供权威响应、IPv6网络中无法解析中国重点域名的风险更高的情况, 研究还发现部分中国域名的权威服务器完全依赖于境外基础设施, 存在潜在的安全隐患。本文从综合评价分析的角度, 提出了量化的中国重点域名权威服务器可用性风险评分指标。最后基于总体风险的分析, 提出了中国重点域名的管理建议, 以强化中国重点领域域名的健康运营。

关键词 域名系统; 权威域名服务器; 网络测量; 依赖度; 风险评估

中图分类号 TP393 DOI号 10.19363/J.cnki.cn10-1380/tn.2025.11.01

Measurement and Analysis of Service Status for Key Domain Authoritative Nameservers in China

NIE Leyao^{1,2}, LI Chenglong^{1,2,3}, DONG Cong², SONG Guanglei², ZHANG Hui^{1,3},

FAN Linna⁴, YANG Jiahai^{1,2,3}

¹ Institute for Network Sciences and Cyberspace, Tsinghua University, Beijing 100084, China

² Zhongguancun Laboratory, Beijing 100094, China

³ Quan Cheng Laboratory, Jinan 250215, China

⁴ College of Information and Communication, National University of Defense Technology, Wuhan 430019, China

Abstract In the context of rapid Internet development, the availability of Domain Name System (DNS) authoritative servers is essential for the normal functioning of network services. The unavailability of these servers directly results in domain inaccessibility, impairing users' ability to access websites, send emails, and engage in other routine online activities. Following the outbreak of the Russo-Ukrainian War, the availability of DNS services has been affected by geopolitical dynamics, necessitating the migration of services associated with Russian domains, particularly those ending in .ru, to domestic infrastructures. However, to date, there is no comprehensive public research that analytically examines the configuration and deployment of authoritative DNS servers for Chinese domains, nor studies that assess the availability risks associated with these servers. In response to global cybersecurity threats, such as service disruptions and complete network outages, there is an urgent need to conduct situational awareness for domain names within critical national cyberspace sectors. This paper initiates its study by compiling a dataset of domain names from the government, educational, and business sectors in China. We proceed to conduct a comprehensive situational assessment of these domains, measuring static infrastructure attributes, infrastructure sharing, authoritative server deployment methods, inactive authoritative servers, dependencies on foreign infrastructure, and top-level domain dependencies on foreign elements. The study undertakes a three-fold data analysis encompassing configuration efficacy, deployment modalities, and regional dependencies to highlight the risks associated with infrastructure hosting bottlenecks, authoritative servers' failure to provide responses, and

通信作者: 董聪, 博士, 助理研究员, Email: dongcong@zgclab.edu.cn

本课题得到中关村实验室“十四五”重点研究计划项目(No. 2022YFB3105000), 泉城实验室重点项目(No. QCLZD 202304-2), 山东省实验室项目(No. SYS202201)资助。

收稿日期: 2024-03-05; 修改日期: 2024-05-06; 定稿日期: 2025-10-14

higher risks of unresolved key Chinese domain names in IPv6 networks. It also uncovers that some Chinese domain authoritative servers are entirely reliant on foreign infrastructures, presenting potential security vulnerabilities. From an integrated evaluation perspective, this paper proposes quantified risk scoring metrics for the availability of authoritative servers for China's key domain names. Finally, based on a comprehensive risk analysis, management recommendations are suggested to enhance the robust operation of domain names in China's key sectors.

Key words domain name system; authoritative server; network measurement; overseas dependency; risk assessment

1 引言

域名系统(Domain Name System, DNS)是现代互联网最重要的信息基础设施之一,是一个分布在根域名服务器、多个层级权威域名服务器和解析器上的分布式数据库,提供方便阅读和记忆的域名和 IP 地址之间的映射。域名系统中发挥重要作用的权威域名服务器(后文简称权威服务器)负责存储域名的权威 DNS 记录,由域名所有者控制和部署。权威服务器不可用会直接导致域名不可用,影响用户访问网站、发送电子邮件等正常网络活动^[1-4],所以权威服务器的服务可用性对于确保网络服务的正常运行至关重要^[5-6]。

权威服务器的可用性容易受到地缘政治的影响,俄乌冲突爆发之后,美国外汇资产控制办公室开始在其制裁实体名单上列出俄罗斯企业网站^[7]。乌克兰副首相请求 ICANN 吊销 .ru(俄罗斯英文顶级域名)、.рф(俄罗斯俄文顶级域名)和 .su(苏联顶级域名,仍在使用中)域名,并关闭位于俄罗斯联邦的 DNS 根服务器实例^[8]。2022 年 3 月俄罗斯当局要求所有国有网站和服务仅限于使用该国的 ISP、DNS 运营商和托管服务提供商^[9]。可见,俄罗斯域名的相关基础设施在事件发生后被迫迁回俄罗斯境内。受此启发,深入了解我国域名基础设施的服务状态,有助于保持中国域名系统的健康稳定运行,规避中国域名在特定国际形势下的不可用风险。然而,当前缺少对中国域名权威服务器的配置和部署的全面态势分析工作,同时缺乏对针对中国重点域名的监控。面对断网停服等全球网络空间安全威胁,迫切需要对国家网络空间重点领域域名开展态势感知。

现有权威服务器态势测量工作主要围绕配置验证^[10-16]、部署方式^[17-19]、地区分布^[4]三个角度展开,但上述研究选择的测量目标局限在常用的通用顶级域名区域文件和 Alexa 前 100 域名^[20],测量范围也局限在 IPv4 网络中,缺乏 IPv6 网络中域名服务状态的研究,因而无法有效反映我国域名系统的整体态势。并且过去的研究缺乏对权威服务器的系统性整体测量。由此,本文展开对中国重点域名权威服务器服务状态的测量。

本文构建了中国重点域名数据集,并对其展开一个季度的测量实验,从配置测量(无效权威服务器)、部署方式测量(服务提供商采用情况、自建权威服务器)、地区依赖测量三方面进行数据分析。本文的测量揭示了 NS 记录无法解析、权威域名服务器无法提供权威响应的情况。研究发现部分中国域名的权威服务器完全依赖于境外基础设施,存在潜在的安全隐患。最后本文针对发现的现象给出管理建议。本文的贡献点如下。

(1) 整理了中国重点领域域名数据集,包含教育、政府、企业的 6766 个中国域名。

(2) 从无效服务器配置、基础设施部署方式、基础设施境外依赖度、IPv6 网络等角度,分析重点域名的基础设施服务状态。

(3) 给出中国重点域名权威服务器可用性风险评分指标,并提出针对中国重点域名的管理建议,如例行检查服务器配置、使用国内域名基础设施等。

本文第 2 节介绍权威域名服务器测量的相关方法和研究现状。第 3 节介绍中国重点领域域名数据集的构建流程。第 4 节介绍重点域名权威服务器服务状态测量实验。第 5 节围绕权威服务器的基础设施共享、部署方式、无效服务器、境外依赖度等对探测结果进行分析,并形成综合风险评分指标。第 6 节给出管理建议。

2 相关工作

本节从配置测量、部署方式测量、地区依赖度角度,梳理了过去针对权威服务器的测量工作。

在权威服务器基础设施配置测量方面,Deccio 等人^[11]和 Jiang 等人^[12]都提出用于分析域名依赖关系的图模型,捕捉了 DNS 区域之间不同的依赖关系,并构建了多种依赖图来描述如何关联上述依赖关系,提出域名与域名之间影响的度量标准,对超过 100 万个 DNS 区域进行调查,呈现了 DNS 中区域依赖关系的整体情况。Akiwate 等人^[14]和 Somme-se 等人^[15]的研究阐述了无效权威服务器的危害和委派不一致对负载均衡的影响,并指出攻击者可以抢注权威服务器主机名,然后劫持受害者子域名解析权限。Li 等人^[16]提出分析和评估权威服务器的配置风险的方

法,发现 10%的域名权威服务器配置不当。然而,上述研究测量的角度比较单一,没有系统性地测量权威服务器存在的风险。

在基础设施部署方式测量方面,Bates 等人^[17]和 Allman^[18]通过对顶级域名的健壮性和集中性测量表明,随着时间的推移,顶级域级别的 DNS 基础设施呈现共享化和集中化的趋势。Zembruzki 等人^[19]在二级域名层级分析了 Alexa 前 100 万数据集中所有最受欢迎域名的权威域名服务器基础设施网络拓扑位置,结果显示,有超过 1.2 万个权威域名服务器共享着第三方 DNS 提供商的相同基础设施,如果发生攻击事件,这些权威 DNS 服务器面临遭受牵连损害的风险将大大增加。可见,测量重点域名的服务基础设施共享情况和集中度是必要的,受这些工作的启发,本文也将共享度测量结果纳入我们的风险评估体系中。

在域名权威服务器对地区依赖度测量方面,Akiwate 等人^[4]分析了俄罗斯域名对俄罗斯境外基础设施的依赖程度,绝大多数俄罗斯网站(70%)完全托管在俄罗斯,完全使用俄罗斯境内域名服务器。对于那些使用非俄罗斯基础设施的俄罗斯网站,几乎所有受影响的网站都很快找到了新的俄国境内供应商。但 Akiwate 的工作没有对域名进行领域分类讨论,没有从微观层面分析行业受到冲击的程度,本文整理了教育、政府、企业三个领域的中国重点域名,深入地分析各个测量指标在不同领域的区别。

当前 DNS 测量项目是基于 DNS 系统每日快照生成的,van Rijswijk-Deij 等人^[21]设计了一个 DNS 测量系统 OpenINTEL,具有被动采集数据、大数据存储和分析功能,但是这类项目不会检查区域文件中列出的权威服务器的可用性,其测量对象是通用顶级域的区域文件中的域名,没有对特定区域、特定领域的域名展开配置、部署方式的测量以及分析。因此,为了获得更具体的数据,本文设计了针对权威服务器的主动测量流程。

3 中国重点域名数据集

本工作讨论中国重点域名的权威服务器服务状态,首先需要构建重点关注的域名集。中国重点域名在本文指教育、政府和商业领域中排名靠前、注册在中国的具有代表性的域名。下面就域名集的构建予以介绍。

3.1 初步收集

本节介绍中国重点域名数据集的原始数据来源和初步收集结果。重点域名收集关键步骤如图 1 所

示。本文通过三种来源收集原始域名数据,将在后续章节筛选出属于中国的域名。第一个来源是 2023 年 4 月 18 日 Tranco 数据库^[22]的域名数据集。Tranco 是 LePochat 等人^[22]提供的网络安全或互联网测量领域的研究人员经常使用的网站排名。第二个来源是 2023 年 7 月 19 日由 CERNET(China Education and Research Network, 中国教育和科研计算机网)收集的 edu.cn 的域名列表,一共有 4202 个域名。第三个来源是在 2023 年 7 月 25 日,以中央人民政府门户网站为起始网页,用爬虫的方式获取的网页 URL。

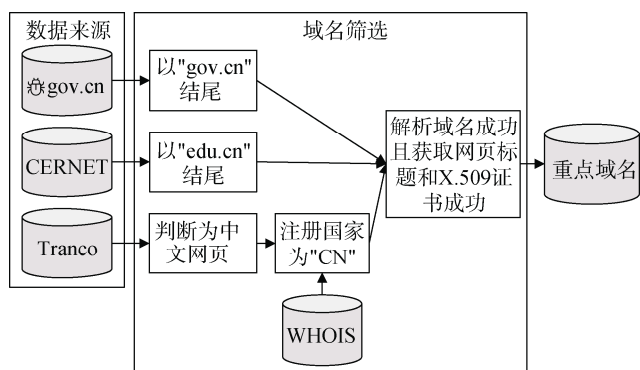


图 1 中国重点域名收集关键步骤

Figure 1 The Steps of Collecting Key Domains in China

(1) 教育域名: 本文以 CERNET 提供的 4202 个教育域名为基础进一步筛选重点教育域名。因为一部分教育网网站只对校园局域网内部师生开放访问权限,这些域名公网无法访问,所以不纳入重点监控范围。本文通过客户端访问这些域名,将可以公开成功解析网页的域名加入数据集,最后得到 1852 个可用教育域名。

(2) 政府域名: 本文以爬虫的方式获取以 "gov.cn" 顶级域下的关联域名作为政府域名。中国政府域名是以 "gov.cn" 结尾的域名,为了获取尽可能多的中国政府域名,本文设计了一个基于 Python 的网页爬虫,用于爬取目标网站 <https://www.gov.cn/> 的下级链接,爬取深度设置为 3 层。为了收集每个网页上的超链接,本文使用了 BeautifulSoup 库来解析网页内容。使用递归函数和先入先出队列来实现多层级爬取。URL 由协议、域名、文件路径组成,本文仅提取 URL 的域名部分,去重后得到 2309 个政府网站域名。

(3) 企业域名: 本文以 Tranco 排名前 100 万域名数据集为基础进一步筛选中国重点企业域名。本文首先访问这些域名对应的网站,判断它们是否是中文网页。判断方法包括: ① 域名后缀。观察域名的顶级域名后缀。中文网页常用的域名后缀包括 ".cn"(中

国)、"com.cn"、"org.cn"等。②内容分析: 访问域名对应的网页, 对网页内容进行分析。可以检查页面上的文字内容、导航菜单、页面标题等。如果大部分内容都是中文字符或中文文本, 那么该网页是中文网页。③语言标识: 查看网页源代码中的语言标识。在网页的 HTML 代码中, 可以通过"lang"属性或"meta"标签中的"content"属性来指定网页所使用的语言。如果语言标识指示为中文("zh-CN"表示简体中文), 那么该网页是中文网页。通过以上三个方法, 最后, 一共得到 39097 个中文网页域名。我们在 3.2 节, 采用 WHOIS 查询注册国家的方法来进一步筛选中国域名。除 Tranco 数据集中的企业域名外, 本文还手动收集了所有国企和央企域名的二级域名共 928 个^[23]。

3.2 交叉验证中国所属域名

在 WHOIS 数据库中, 域名所有者的国家信息通常可以通过查看"Registrant Country"字段来获取。使用 WHOIS 查询工具, 输入域名进行查询, 如果注册人信息显示为中国地区, 那么判断该域名为中国域名。其中由于".cn"国家顶级域名由中国政府相关部门管理, 所以以".cn"为后缀的域名可以直接判定为中国域名。其他国家结尾默认不算中国域名。通过 WHOIS 查询, 最后一共得到 22463 个 WHOIS 注册国家为中国的域名。为了进一步精选重点域名, 确保其中不包含非重点领域域名以及非中国公司域名, 我们与网站进行 SSL/TLS 握手获取其 X.509 证书, 只保留证书的国家代码字段为"CN"的网站域名, 最后经过手工检查一共得到 6766 个有效域名。

3.3 域名分类和重要性说明

前面得到的有效域名的细分领域下包含 1852 个教育域名(edu.cn 结尾), 2309 个政府域名(gov.cn 结尾)和 2605 个企业域名。重点域名数据集的顶级域名数量统计排名前十名见表 1, 其中, 政府顶级域名 gov.cn 最多, 占重点域名总数的 34.13%; 教育顶级域名 edu.cn 的数量排在第二, 占重点域名总数的 27.37%; com 的数量排在第三, 占重点域名总数的 17.94%, 这是由于本文关注的重点企业倾向于使用.com 作为顶级域。

表 1 重点域名集顶级域名数量统计

Table 1 Statistics of Top-Level Domain Numbers

排序	顶级域名	数量	比例	排序	顶级域名	数量	比例
1	gov.cn	2309	34.13%	6	org.cn	60	0.89%
2	edu.cn	1852	27.37%	7	net	50	0.74%
3	com	1214	17.94%	8	ac.cn	39	0.58%
4	cn	602	8.90%	9	net.cn	27	0.40%
5	com.cn	561	8.29%	10	org	19	0.28%

本文构建的域名数据集能够覆盖我国各领域的重要实体。在教育领域, 本文从中国教育和科研计算机网络机构获取了 edu.cn 顶级域名的域名列表, 该机构负责管理所有 edu.cn 域名, 其官网^[24]明确写到高等院校、中小学校和各级依法设立的教育科研机构均可在 edu.cn 下申请域名, 注册域名单位必须是我国依法设立的各级教育和科研单位, 申请人需提供一份关于本单位中英文全称、域名以及网络管理及技术人员的签章证明。本文的探针位于校园网外部, 为避免大量无效解析, 本文仅收集了公网可解析域名且可访问主站的教育域名。政府领域域名涵盖范围包括中央人民政府主站、地方政府网站(省级、自治区级、直辖市级、地市级、区县级), 包括全国人大、全国政协、国家监察委员会、最高人民法院、最高人民检察院的重要主站和地方性分站。企业领域域名包含了 928 个央国企域名, 另包含电信、汽车、科技等热门行业网站, 这些站点提取自 Tranco 排名前 100 万数据集, 而 Tranco 集成了 4 个知名域名排名数据集, 分别是 Alexa^[20](基于全球流量, 根据访问量和页面浏览量进行排名)、思科 Umbrella^[25](基于抽样和标准化的 OpenDNS 的 DNS 流量, 域名的排名基于发出 DNS 查询的唯一 IP 的数量)、Majestic^[26](基于 120 天内爬行约千亿级别 URL 的网页数据里域名网站的后向链接)、Quantcast^[27](基于一个月内访问站点的人数)。

4 中国重点领域域名权威服务器测量

权威服务器是域名系统的重要组成和基础设施, 我们将对中国重点域名的权威服务器进行服务状态测量, 揭示中国重点域名的运营态势。本节介绍权威服务器服务状态的测量流程及实验结果的数据结构。

4.1 流程设计

本文计划探测收集的数据包括 6766 个中国重点域名集中每个域名的 NS 记录(NS 记录的值是权威域名服务器的主机名)以及下属所有权威服务器主机名的 A 记录和 AAAA 记录。

本文采用 Docker 的形式封装代码及其运行环境(Ubuntu 系统、Python3), 可以方便地将代码部署在任意节点运行。本文将代码运行在一台 1 核 2GB 的虚拟专用服务器上。从 2023 年 9 月 4 日开始, 以一天为周期, 在每天的 20 点进行一次重点域名扫描。本文使用了 ZDNS 工具^[28]进行扫描探测。每隔 24 小时执行一次探测任务, 遍历重点域名列表每一个域名, 并按照如下流程扫描。

- (1) 获取被测域名的二级域名(SLD)。
- (2) ZDNS 工具迭代式查询 SLD 的 NS 记录, 提取响应报文里的字段('answers_ns', 'name', 'status', 'timestamp')。
- (3) 遍历'answers_ns'中每一个子项, 迭代变量设为 HOSTNAME: 1)ZDNS 工具迭代式查询 HOSTNAME 的 A 记录, 提取响应的'data。answers'重新命名为'answers_nsa'。遍历每一个 A 记录里的 IPv4 地址, 请求解析 SLD 的 NS 记录, 并存储响应报文。2)ZDNS 工具迭代式查询 HOSTNAME 的 AAAA

记录, 提取响应的'data。answers'重新命名为'answers_ns4a'。遍历每一个 AAAA 记录里的 IPv6 地址, 请求解析 SLD 的 NS 记录, 并存储响应报文。

本文采用并行化来应对单线程因等待服务器响应耗时过多的问题。采用这种并行查询的方式, 可以使得本文的单次查询时间从 9~10 小时缩短到 30~40 分钟, 单次探测任务的效率提高了 18~20 倍。

4.2 测量结果

探测任务收集的原始数据为 JSON 文件, 本文将其转换为结构化数据以便分析(见表 2)。

表 2 结构化测量数据
Table 2 Structured Measurement Data

字段名	含义	类型	示例数据
name	域名	字符串	"example.com.cn"
status	查询返回的状态码	字符串	"NOERROR"
timestamp	时间戳	字符串	"2023-08-30T12:00:00Z"
answers_ns	域名的 NS 记录	列表(字符串)	["dns10.hichina.com.", "dns9.hichina.com."]
answers_nsa	NS 记录进一步查询的 A 记录	列表(元组)	[("dns10.hichina.com.", "120.76.107.51"), ("dns10.hichina.com.", "139.224.142.11"), ("dns9.hichina.com.", "120.76.107.42"), ("dns9.hichina.com.", "39.96.153.52")]
answers_ns4a	NS 记录进一步查询的 AAAA 记录	列表(元组)	[("dns10.hichina.com.", "2408:4009:501::18"), ("dns9.hichina.com.", "2408:4009:501::17")]

5 实验分析

本节将从权威服务器基础设施静态属性、共享次数、部署方式、无效配置、境外依赖度等 5 个方面进行分析, 最后综合以上分析给出中国重点域名权威服务器可用性风险评分指标。分析结果表明域名集所采用的权威服务器基础设施大部分位于中国境内, 并且偏向于采用第三方 DNS 服务特点。但是本文同样观察到当前存在着单个基础设施上托管过多域名、IPv6 网络中无效权威服务器偏多、企业域名对境外基础设施依赖度高的问题。

我们使用 IP-API 在线服务^[29]的测量结果对每个 IP 地址进行地理定位, 用以定位每个域名的 DNS 基础设施所在地理位置。由于获取域名列表和实际开始周期性探测的日期存在时间差, 一些域名在这个期间过期, 变得无法解析, 以 2023 年 12 月 19 日的日志举例, 响应状态码有 NOERROR(正常响应, 6754 个)、NXDOMAIN(域名不存在, 10 个)、ITERATIVE_TIMEOUT(迭代查询超时错误, 2 个)。

5.1 基础设施静态属性分析

为了更好地呈现 DNS 基础设施的分布情况, 本节统计了这些基础设施 IP 的静态属性。

本文收集了权威域名服务器基础设施的 IP 集合(分为 IPv4 和 IPv6), 一共有 3860 个 IPv4 地址, 1974

个 IPv6 地址。本次测量得到的 IPv4/IPv6 权威服务器分别分布在 17/12 个国家, 62/38 个城市, 164/76 个自治系统, 435/81 个组织。可以看出 IPv4 基础设施的分布比 IPv6 更为分散和广泛, 一个潜在的原因是 IPv4 在互联网中的应用时间更长。表 3 统计了基础设施的地区数量, 教育域名和政府域名 94%以上的权威 IP 地址都在中国, 分布在中国教育网(CERNET)自治系统最多; 而企业域名权威有 10.76%的 IPv4 地址和 11.54%的 IPv6 地址定位在境外, 主要是分布在 Cloudflare 和亚马逊的自治系统。

5.2 权威服务器基础设施共享

基础设施过多共享, 会形成网络瓶颈, 可能导致单点故障。本节统计了重点域名集中 IPv4/IPv6 权威服务器基础设施(以一个 IP 标识一个基础设施)被共享给多少个域名使用, 来揭示风险, 以被共享数量在 2~10, 11~100, 101~500, 501 以上的区间聚合, 见表 4。

表 4 显示 2987 个 IPv4 基础设施和 1642 个 IPv6 基础设施为单独的域名提供服务, 而其余基础设施或多或少地为多个域名提供权威解析服务, 通常是 DNS 解析服务提供商维护和管理这些共享基础设施, 它们需要管理多个区域文件, 从而增加了基础设施管理的复杂度和风险。值得注意的是, 支持超过 500 个中国重点域名的权威解析服务的 IPv4 基础设施有

表 3 IP 基础设施地区数量统计

Table 3 Regional Statistics of IP Infrastructure

	排序	教育域名			政府域名			企业域名		
		地区	数量	比例	地区	数量	比例	地区	数量	比例
IPv4 基础设施	1	中国	2230	99.78%	中国	632	96.49%	中国	1186	89.24%
	2	美国	2	0.09%	美国	6	0.92%	美国	53	3.99%
	3	英国	1	0.04%	香港	4	0.61%	香港	18	1.35%
	4	日本	1	0.04%	加拿大	4	0.61%	爱尔兰	16	1.20%
	5	香港	1	0.04%	新加坡	3	0.46%	加拿大	16	1.20%
IPv6 基础设施	1	中国	1365	99.78%	中国	223	94.49%	中国	437	88.46%
	2	香港	1	0.15%	美国	3	1.27%	美国	29	5.87%
	3			0.07%	香港	3	1.27%	英国	9	1.82%
	4				加拿大	2	0.85%	香港	5	1.01%
	5				哥斯达黎加	2	0.85%	澳门	3	0.61%

表 4 IP 基础设施被共享次数统计

Table 4 Statistics on the Number of Times IP Infrastructure is Shared

	1	[2,10]	[11,100]	[101,500]	[501,∞)	总数量
IPv4 基础设施	2987	585	146	130	12	3860
IPv6 基础设施	1642	221	82	27	2	1974

12 个, IPv6 基础设施有 2 个, 这代表着它们需要维护超过 500 个域名的区域文件。统计数据中, 单个基础设施最多为 840 个域名提供权威解析服务。按照这样的共享程度, 若该基础设施出现单点故障, 所有托管在它之上的所有域名无法解析的风险将增加, 即使这些域名部署了冗余权威服务器, 共享基础设施的故障仍会带来解析时延增加。定位分析结果显示所有托管超过 500 个域名的基础设施的定位都在中国大陆, 并集中在以下几个自治系统: AS37963 杭州阿里巴巴广告有限公司、AS45090 深圳腾讯计算有限公司、AS132203 腾讯大厦科技中一大道、AS134756 中国电信南京济山 IDC 网络、AS4837 中国联通 169 骨干网、AS56046 中国移动通信集团公司、AS136958 中国联通广东 IP 网络、AS4134 中国电信骨干网、AS4538 中国教育网、AS24348 清华大学 CERNET2 IX。

5.3 权威服务器部署方式

权威域名服务器的部署方式具体有两类: ①自建权威服务器。管理员在自有的服务器主机上安装

DNS 软件(例如, BIND9, Unbound DNS, PowerDNS), 配置各类 DNS 记录, 如 NS 记录、A 记录、CNAME 记录等。在域名注册商的控制台上, 将域名的权威服务器设置为自己搭建的服务器。②使用第三方 DNS 厂商提供的 DNS 解析服务, 如 Cloudflare、Google Cloud DNS 等。在 DNS 服务提供商的控制台上添加 DNS 记录, 并将权威服务器设置为服务提供商的服务器。

本文根据每个域名的 NS 记录中的权威服务器主机名来区分该域名的权威服务器部署方式, 自建权威服务器主机名的域名后缀与待测域名的域名后缀相同, 使用厂商 DNS 服务的权威服务器主机名的后缀则与待测域名的不同。表 5 统计了每个域名的部署权威服务器的方式, 表中的数字代表处于该分类下的域名数量。

从表 5 可以看出, 教育域名采用自建服务器的方式最多, 企业域名次之, 政府域名最少。教育机构更倾向于自建权威服务器, 以满足较高的安全性和自主控制的要求, 保护学术和科研资源。即使使用第

表 5 权威服务器部署方式

Table 5 Deployment Methods of Authoritative Servers

	自建权威服务器	第三方 DNS 服务	两种方式混合采用	域名总量
教育域名	1088 (58.84%)	718 (38.83%)	43 (2.33%)	1849
政府域名	94 (4.08%)	2195 (95.31%)	14 (0.61%)	2303
企业域名	177 (6.80%)	2401 (92.28%)	24 (0.92%)	2602

三方 DNS 服务,也集中于 CERNET 教育网提供的 DNS 解析服务。企业则更倾向于使用专业的 DNS 服务提供商。92.28%的企业域名采用了服务提供商的权威解析服务,通常有如下考虑:服务提供商提供专业的 DNS 解析服务,包括高可用性、全球分布的解析节点、防护服务等,企业可以更专注于业务而不用过多关注基础设施。DNS 服务提供商通常有灵活的付费模式,企业可以根据实际使用情况灵活调整服务规模,不需要投入大量资金购置和维护服务器。本文在表 6 统计了排名前十名的域名托管服务提供商,前十名服务提供商都是中国公司,说明中国域名大多数都采用中国的服务提供商。

表 6 第三方 DNS 服务提供商托管域名数量统计前十

排序	服务提供商	域名数量	比例(%)
1	中国万网	1356	20.04%
2	北京国科云	777	11.48%
3	北京新网数码	654	9.67%
4	腾讯云 DNSPod	513	7.58%
5	阿里 DNS	483	7.14%
6	教育网 DNS	467	6.90%
7	国家工程研究中心	447	6.61%
8	福建三五互联	183	2.70%
9	厦门帝恩思	102	1.51%
10	成都西维数码	97	1.43%

5.4 无效权威服务器

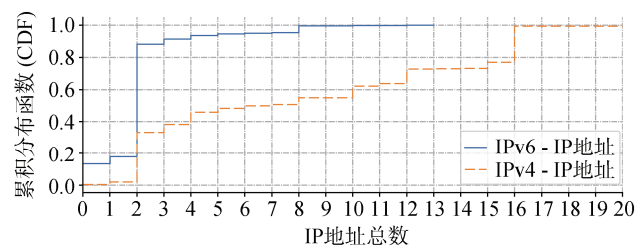
权威服务器的不可用将导致域名承载的网络资源不可用,本节测量每个权威服务器的服务可用性,揭示无效权威服务器在 IPv4 和 IPv6 网络中的态势。当解析器无法从权威服务器获取 DNS 权威响应时^[30],该权威服务器称为无效权威服务器。无效权威服务器有 2 种粒度的统计方式,一种是粗粒度的主机名,主机名以域名形式书写,例如, ns1.qq.com; 另一种是更细粒度的主机 IP 地址,这些 IP 地址是通过主机名的 A 记录或者 AAAA 记录查询得来的。DNS 域名系统中为了提供网络负载均衡,一个粗粒度的主机名可能由多个细粒度的主机 IP 地址托管,这样大量的请求可以指向同一个主机名下的不同主机 IP 地址。

本文推测以下情况会导致权威服务器无效:①当域名所有者修改或删除子区域中的权威服务器,但未在父区域中同步时,导致父区域中的 NS 记录指向不存在或非权威服务器,就会产生无效权威服务器。②配置错误可能导致权威服务器域名指向不存在或非权威服务器。本文通过向权威服务器请求域名的

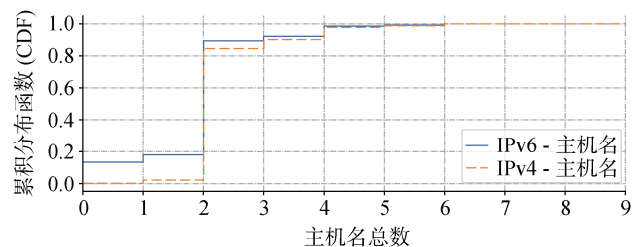
NS 记录,在以下三种情况下判定它是无效服务器的响应结果:①没有响应结果。在这种情况下,请求将重试三次,如果三次都没有响应,则判定为无效权威服务器,从而消除网络数据包丢失的干扰。②响应状态码为"ServFail"或"Refused"。③无法解析该域名的 NS 记录,或者响应里没有权威域名服务器的完整信息。

探测实验获取了子区域 NS 记录数据,首先,本文从宏观上观察重点域名的错误情况,统计受影响的域名数量。宏观来看,这些域名中,有 711 个域名部署了无效 IPv4 权威服务器。有 2046 个域名部署了无效 IPv6 权威服务器。本文从错误日志记录的数据中提取出 928 个<域名,权威服务器主机名,IPv4 地址,响应返回码>四元组,每条四元组代表了一条权威服务器主机名的 A 记录。经过去重,无效 IPv4 地址有 705 个,被解析到这些 IP 地址的权威服务器主机名有 732 个,影响到的域名有 711 个。提取出 3871 个代表了一条权威服务器主机名的 AAAA 记录四元组,经过去重,无效 IPv6 地址共有 911 个,被解析到这些 IP 地址的主机名有 1136 个,影响到的域名有 2046 个(30.24%)。

本文整理出单个域名的统计信息,分别统计了各个域名下的无效 IP 数量、有效 IP 数量、总 IP 数量,无效主机名数量、有效主机名数量、总主机名数量。图 2(a)分别统计了一个域名在 IPv6 和 IPv4 中的权威主机名进一步解析得到 IP 地址的数量。在 IPv6 权威服务器部署上,902 个域名(13.33%)没有部署 IPv6 权威服务器,大多数域名(4756 个, 70.29%)部署了 2 个 IPv6 权威服务器,在监测的域名中一个域名



(a) IP地址数量



(b) 主机名数量

图 2 权威服务器配置数量累积分布图

Figure 2 CDF of Authoritative Server

最多部署 12 个 IPv6 权威服务器。在 IPv4 权威服务器的部署上, IPv4 权威服务器的数量从 1 个到 50 个均有分布(统计图中只显示到 20 个)。然而, 部署 16 个 IPv4 权威服务器的域名有 1511 个(22.33%)。本文统计了解析到被映射这些 IP 的权威主机名数量, 如图 2(b), 可以看到配置了 A 记录的权威主机名和配置了 AAAA 记录的权威主机名趋势基本一致, 除开没有部署 IPv6 权威服务器的域名, 这些域名大部分在 IPv4(82.47%)和 IPv6(71.39%)网络中部署 2 个权威主机名, 见图 2(b), 这样可以保证一定程度的冗余性, 增加域名解析的鲁棒性。

为进一步分析无效权威服务器的风险, 本文绘制了同样逻辑的累计分布函数图。如图 3(a), IPv6 中, 权威服务器全部有效的域名有 4720 个(69.76%, 对应图中 x 坐标轴为 0 的 y 值), 而 IPv4 中 6055 个域名的权威服务器全部有效(89.49%), 也就是说在解析这些域名时, 不管命中哪一个权威服务器 IP 都可以正常解析, 域名无法解析的风险低。然而, 其余域名或多或少存在命中无效权威服务器 IP 的风险。

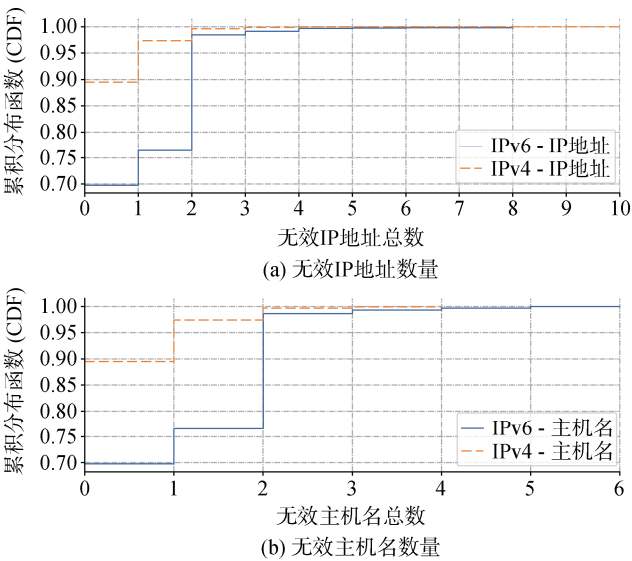


图 3 无效权威服务器配置数量累积分布图
Figure 3 CDF of Invalid Authoritative Server

IPv6 网络中 4.17%(282 个)的域名和 IPv4 中

6.74%(456 个)的域名有 50%的概率命中无效权威服务器, 最坏情况下, 这会造成 2 倍于正常情况下的解析延时。此外, 有 24.68%(1670 个)的域名的所有 IPv6 权威服务器都是无效权威服务器, 该现象说明这 24.68%的域名无法通过 IPv6 网络访问, 即使这些域名的网站管理员的本意可能是建立 DNS 双栈解析通道。值得注意的是, 这部分域名并不包括没有部署 IPv6 权威服务器的域名(13.33%, 902 个), 两者加起来一共有 38.01%(2572 个)的重点域名都无法通过 IPv6 访问。反观 IPv4 权威服务器的情况, 只有 1.37%(93 个)的域名的所有 IPv4 权威服务器都无效, 0.25%(17 个)的域名无法解析出 IPv4 权威服务器的记录(NXDOMAIN 或者未部署)。从本工作实验数据来看, 中国重点域名的 IPv6 域名基础设施服务状态比 IPv4 中的差。

5.5 权威服务器境外依赖度分析

本节将从权威服务器主机名的境外依赖度、IP 基础设施的境外依赖度分析中国重点域名 DNS 基础设施对境外的依赖程度, 重点域名对境外依赖度越低则风险越小。

(1) 权威服务器主机名对境外顶级域名的依赖度。本文提取所有域名的每个权威服务器的顶级域名, 判断它们的权威主机名是否受国家机构控制。数据集依赖的顶级域名总共有 12 个, 48.87%的中国重点域名的权威服务器注册在 .com 顶级域名下, 第二名是我国控制的顶级域名 .cn(41.43%), 然后是 .net(8.02%), .info(0.95%), .biz(0.34%)。如果一个域名的所有权威域名服务器主机名都是在中国顶级域名 (.cn)以外的顶级域名下注册的, 本文判断该域名完全依赖境外顶级域。如果所有权威服务器主机名都是在中国顶级域名下独家注册的, 本文认为该域的顶级域名对境外无依赖。如果所有权威主机名中只有一个子集是中国顶级域名, 本文认为它是部分依赖境外的。

从表 7 境外依赖度的统计结果可以看到, 以 .cn 结尾的权威主机名, 教育行业的域名最多, 几乎所有教育域名都部署了中国自己控制的顶级域下的权

表 7 权威服务器顶级域名境外依赖度
Table 7 Overseas Dependency of Top-Level Domains on Authoritative Servers

	全依赖境外	部分依赖境外	无依赖	总数
教育域名	0 (0.00%)	2 (0.11%)	1847 (99.89%)	1849
政府域名	1615 (70.13%)	290 (12.59%)	398 (17.28%)	2303
企业域名	2188 (84.09%)	108 (4.15%)	306 (11.76%)	2602
重点域名集	3803 (56.31%)	400 (5.92%)	2551 (37.77%)	6754

威服务器, 而企业域名最少, 84.09%的企业域名权威服务器主机名完全依赖外部的顶级域名。教育域名基本上服务于中国用户并且无须部署商业流量分发方案, 中国顶级域名就可以满足服务需求。企业域名通常直接注册在常用的顶级域名下(如.com), 它们既可以自建权威服务器, 也可采用第三方服务提供商提供的基础设施, 在两种情况下, 权威服务器主机名都较少以.cn 结尾。

(2) 权威服务器 IP 地址对境外基础设施的依赖

表 8 权威服务器 IP 地址境外依赖度

Table 8 Overseas Dependency of IP Addresses on Authoritative Servers

	全依赖境外	部分依赖境外	无依赖	总数
教育域名	0 (0.00%)	6 (0.32%)	1843 (99.68%)	1849
政府域名	2 (0.09%)	94 (4.09%)	2202 (95.82%)	2298
企业域名	18 (0.69%)	392 (15.07%)	2192 (84.24%)	2602
重点域名集	20 (0.30%)	496 (7.35%)	6233 (92.35%)	6749

比较三个重点领域的域名, 对境外域名基础设施的依赖程度, 教育相对最低, 企业相对最高。教育域名基本上是服务位于中国的用户, 对面向境外用户提供服务的需求较少, 而企业存在全球业务需求, 因此维护了境外的权威服务器, 或者企业为了便捷性直接采用了境外的 DNS 解析服务。完全依赖境外 IP 基础设施的政府域名中, 有 1 个域名是依赖香港的基础设施(由中央人民政府驻香港特别行政区联络办公室管理)。有 1 个政府域名(由徐州市纪律检查委员会管理)的权威服务器是委派给了 Cloudflare 的域名服务器, 它们是完全使用第三方 DNS 解析服务而没有在中国境内自建权威服务器。完全依赖境外基础设施的企业域名中, 有 8 个国央企域名, 分别是深圳国际控股有限公司, 华天实业控股集团有限公司, 西安地图出版社有限公司, 山东泰山地勘集团有限公司, 招商局集团有限公司, 中国南光集团有限公司, 湖北省高新产业投资集团有限公司, 青海省投资集团有限公司。中国的外贸出海公司、金融公司为了给境外用户提供网站服务, 也选择了境外基础设施, 如招商局集团有限公司所部署的权威服务器都位于香港。一些小微企业部署了成本较低的境外 DNS 解析服务, 如萤石智能家居。这些完全依赖境外基础设施的域名都采用了境外的第三方 DNS 解析服务, 他们主要采用以下知名的厂商, 如 Cloudflare、亚马逊、Akamai、Cscdns、Markmonitor 等。

部分依赖境外基础设施的重点域名中, 有 30 个域名自建权威服务器。教育域名中, gdsyzx.edu.cn(中学)有 2 个权威服务器部署在美国, 2 个在中国;

度。本文使用 IP-API^[29]查询每个域名的所有权威服务器 IP 地址的定位信息。按依赖境外基础设施的程度分类, 如果一个域名的所有权威域名服务器 IP 都位于境外, 本文认为该域名是完全依赖于境外基础设施的域名。如果所有服务器 IP 都位于中国境内, 本文认为该域名是对境外基础设施无依赖的域名。如果只有子集在中国境内, 本文认为该域名是部分依赖境外基础设施的域名。表 8 记录了三个分类和整体域名集的 IP 地址境外依赖情况。

whcm.edu.cn(音乐学院)的 5 个权威 IP 中有 1 个在英国; hnist.cn(理工学院)的 2 个权威 IP 有 1 个在美国; tsinghua.edu.cn12 个权威 IP 中 2 个定位在香港中文大学(ns5.cernet.net)。国央企域名中, cbhb.com.cn(银行)有 6 个权威服务器主机名, 其中 dns5 的 IPv4 地址定位在韩国; sse.com.cn(交易所)自建的服务器都在上海, 但选用的 DNS 服务提供商有一个权威服务器定位在香港; siic.com(实业公司)2 个权威服务器中 1 个在上海, 1 个在香港; sinosteel.com(钢铁公司)自建的权威都在中国, 但采用的厂商服务中有 2 个 IP 定位在美国亚马逊数据中心; 其他企业如腾讯(tencent.cn), 网易有道(yodao.com), 淘宝(taobao.com)等, 因为业务需要, 在境外也部署有权威服务器。

在此基础上, 本文对支持 IPv4/IPv6 双栈解析的域名进行分析, 分别对权威主机名的 A 记录和 AAAA 记录进行 IP 境外依赖度统计, 如表 9, 对比两个网络下的重点域名集的境外依赖数量, IPv6 下的 IP 境外依赖度相比 IPv4 更小, 得益于企业域名对境外基础设施的依赖更小。

值得注意的是, IPv6 网络中, 有 43 个政府域名的 IPv6 域名基础设施完全依赖境外, 其中 1 个域名是徐州市纪委监委网站, 它采用了 Cloudflare 的服务, 使用了 6 个位于美国的 IPv6 地址。其余 42 个域名被委派给了 2 个 IPv6 基础设施, 2 个 IPv6 地址都属于福建三五互联这个 DNS 服务提供商, 实际上这是较为危险的, 2 个 IPv6 基础设施上至少托管了 42 个政府域名, 一旦基础设施出现问题, 这些政府域名都无法通过 IPv6 访问。完全依赖 IPv6 境外基础设

表 9 IPv4/IPv6 双栈权威服务器 IP 地址境外依赖度

Table 9 Overseas Dependency of Dual-Stack IPv4/IPv6 Authoritative Server IP

		全依赖境外	部分依赖境外	无依赖	总数
IPv4	教育域名	0 (0.00%)	3 (0.20%)	1529 (99.80%)	1532
	政府域名	1 (0.05%)	33 (1.61%)	2016 (98.34%)	2050
	企业域名	9 (0.39%)	286 (12.54%)	1986 (87.07%)	2281
	重点域名集	10 (0.17%)	322 (5.49%)	5531 (94.34%)	5863
IPv6	教育域名	0 (0.00%)	3 (0.20%)	1530 (99.80%)	1533
	政府域名	43 (2.10%)	11 (0.54%)	1996 (97.37%)	2050
	企业域名	40 (1.75%)	48 (2.10%)	2193 (96.14%)	2281
	重点域名集	83 (1.42%)	62 (1.06%)	5719 (97.53%)	5864

施的国央企域名有 20 个, 它们主要采用了三五互联的 DNS 服务。

5.6 综合风险评分

本节综合前文的测量结果(基础设施共享测量、无效权威服务器测量、基础设施境外依赖度测量、顶级域名境外依赖度测量), 给出中国重点域名权威服务器可用性风险评分指标, 一个域名的风险值越高, 该域名的健康运营风险越高。不同的 DNS 解析器软件实现时采用不同的最优权威服务器选择策略, DNSCache、Unbound 和 Windows DNS 均等选择每个权威服务器, BIND9 在初始化状态下随机选择一个权威服务器来启动首选策略^[31]。因此, 本文假设当 DNS 解析器选择权威服务器时, 每个权威服务器主机名被选中的概率是相等的^[16], 在这种条件下, 计算每个域名的风险值。

本文将当前域名记作 d , 其对应的风险评分记作 R_d 。 R_d 由 IPR_d (IP Risk, 权威服务器 IP 风险)和 HR_d (Host Risk, 权威服务器主机名风险)两部分组成。为了计算 IPR_d 和 HR_d , 本文需要计算域名 d 下每个权威服务器的 4 个风险属性值, 用 $a_k(a_1, a_2, a_3, a_4)$ 来标识该属性条件成立与否的布尔值。若该权威服务器 IP 基础设施与其他域名共享, 则 $a_1 = 1$; 若该权威服务器 IP 是无效服务器, 则 $a_2 = 1$; 若权威服务器 IP 依赖境外基础设施, 则 $a_3 = 1$; 若权威服务器主机名依赖境外顶级域名, 则 $a_4 = 1$, 反之亦然。 a_k 见公式(1):

$$a_k = \begin{cases} 1, & \text{条件为真} \\ 0, & \text{条件为假} \end{cases} \quad \text{对于 } k=1,2,3,4 \quad (1)$$

设当前域名 d 配置有 n 个以 IP 标识的权威域名服务器, m 个以主机名标识的权威域名服务器。

域名 d 的 IPR_d 值聚合了负责解析域名的 n 个权威服务器 IP 的中国可用性风险 $a_1 \sim a_3$ 属性[聚合为 A_i , 见公式(2)]和域名的权威服务器 IP 冗余度。对于

IP 冗余度, 以 IP 标识的权威服务器越多, 则风险越小, 所以在公式(3)中表现为 n 的倒数。并且本文主要强调域名的中国可用性, IP 冗余度对 IPR_d 的贡献比起 a_k 风险属性值的贡献更小, 所以在这里本文将 IP 冗余度缩小为 0.1 倍。 IPR_d 的整体公式见公式(3):

$$A_i = \frac{1}{3} \sum_{k=1}^3 a_{k_i} \quad (2)$$

$$IPR_d = \frac{1}{n} \sum_{i=1}^n A_i + 0.1 \times \frac{1}{n} \quad (3)$$

域名 d 的 HR_d 值聚合了负责解析域名的 m 个权威服务器主机名的中国可用性风险 a_4 属性值和域名的权威服务器主机名冗余度。对于主机名冗余度, 风险评估逻辑与 IP 冗余度一致, 在公式(4)中表现为 m 的倒数, 并将系数调整为 0.1。 HR_d 的整体公式如下:

$$HR_d = \frac{1}{m} \sum_{j=1}^m a_{4_j} + 0.1 \times \frac{1}{m} \quad (4)$$

最后, 计算域名 d 的综合中国可用性风险值 R_d , 见公式(5)。本文将定义为 IPR_d 和 HR_d 的加权线性组合, 其中 α 和 β 分别是 IPR_d 和 HR_d 的权重系数。

$$R_d = \alpha IPR_d + \beta HR_d \quad (5)$$

为了客观给出 α 和 β , 本文使用熵权法^[32]为 IPR_d 和 HR_d 分配权重。熵权法是根据各项指标值的信息量程度来确定指标权数的, 这是一种客观赋权法, 避免了人为因素带来的偏差。熵权法基于信息熵的概念, 通过计算每个变量的熵值来衡量其贡献度。熵越大, 变量对于整体系统的贡献越大。权重可以反映每个指标在整个数据集中的相对重要性。本文的观测样本是重点域名集中所有域名, 本文调整了熵权法中数据标准化公式, 使得它更适用于为本文的自变量赋权重。主要有以下 3 个步骤。

(1) 数据标准化: 首先, 本文需要对数据进行标准化处理, 以消除指标之间的量纲影响和数量级差异。对于正向指标(值越大越好), 标准化公式为

$$r_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \quad (6)$$

其中, r_{ij} 是标准化后的数据, x_{ij} 是第 i 个样本的第 j 个指标的原始值。

(2) 计算信息熵值: 本文计算每个指标的熵值来衡量指标的离散程度, 熵值越大, 说明指标的有效信息越少。熵值的计算公式为

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n r_{ij} \ln r_{ij} \quad (7)$$

$\ln$ 表示自然对数, n 是样本总数, r_{ij} 是第 i 个样本的第 j 个指标的标准化值。如果 r_{ij} 为 0, 则 $r_{ij} \ln r_{ij}$ 定义为 0。通常取 $1/\ln n$, 以确保 $0 \leq e_j \leq 1$ 。

(3) 计算权重: 熵值越大, 指标的信息量越大, 对决策的贡献也越大, 计算各指标的权重 w_j :

$$\begin{cases} w_j = \frac{e_j}{\sum_{j=1}^m e_j} \\ \alpha = w_1, \beta = w_2 \end{cases} \quad (8)$$

本文绘制了中国重点域名权威服务器可用性风险评分的累积分布图, 图 4 标注了覆盖度为 18%, 90% 对应的横轴风险值分别为 0.3681, 0.5087 (Δ : 0.1407)。

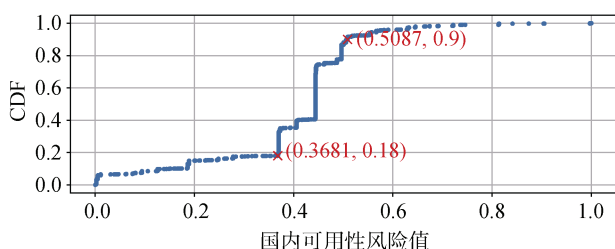


图 4 可用性风险评分的累积分布图
Figure 4 CDF of Availability Risk Scores

首先, 从 0.3681~0.5087, 这一段的覆盖率从 18%~90% 陡升, 涵盖了 70% 的数据集, 我们发现这一段集中了只采用第三方 DNS 解析服务的域名, 它们具有相似的权威服务器部署特性, 所以权威服务器的风险值也相似。例如, 共享基础设施次数最多的两个 DNS 解析服务厂商中, 中国万网的 dns9.hichina.com, dns10.hichina.com 这两台以主机名标识的第三方权威服务器被 1352 个域名共享, 这些域名的风险值都相同, 为 0.4438; 中国新网的 ns15.xincache.com, ns16.xincache.com 两台第三方权威服务器被 563 个域名共享, 与之关联的域名风险值也都为 0.4965。

其次, 观察风险值小于 0.3681 的域名共有 1220 个。其中教育域名有 928 个, 政府域名有 100 个, 央

企域名有 55 个, 其他企业域名有 137 个, 这些域名大部分都是采用自建权威域名服务器的形式且服务器位于中国, 其中少量域名(61 个)选用了第三方 DNS 服务。然后, 我们观察到风险值大于 0.5087 的域名共有 685 个, 其中企业域名有 295 个, 教育域名有 123 个, 政府域名有 182 个, 央企域名有 85 个。一部分域名采用境外第三方 DNS 服务(Cloudflare, Akamai 等), 另一部分采用境外权威服务器 IP, 造成中国可用风险值偏高。

比较前 18% 和后 10% 的风险值对应的域名领域, 可以看到低风险域名通常是带有国家性质的域名, 特别是教育域名, 高风险域名通常是私人性质的企业域名。但值得注意的是, 高风险域名中也包含一部分国家性质的域名, 几乎全部都是依赖福建三五互联的服务器(dns-diy.com), 在本文的数据集里该厂商的服务器有 2 台位于境外, 该厂商提供的服务往往是 DNS 加 CDN 打包网站分发服务, 在它的权威服务器中并未配置二级域名的 NS 记录, 而是直接配置三级(或以上)子域名的 CNAME 记录。这样的配置会导致二级域名在本文的评估体系中风险值偏高, 因为该二级域名权威服务器的子区域文件与顶级域名权威服务器父区域文件配置不一致, 存在 DNS 记录缺失。

6 管理建议

本节根据实验分析结果, 从权威服务器部署、配置、地区依赖等方面给出管理建议。

权威服务器部署方面, ①推荐管理员综合考虑权威服务器的两种部署方式, 在需要较高的安全性和自主控制时选择自建权威服务器, 在业务有较高灵活性、基础设施资金投入有限的情况下选择第三方 DNS 服务。②推荐域名管理员合理规划权威服务器集群的部署策略, 在不同地理位置的不同自治系统中尽可能部署尽可能多的权威服务器, 使用不同的 IP 地址和不同前缀进行不同操作, 以避免陷入单点故障。

权威服务器配置方面, 其错误配置会造成无效权威服务器的问题, 导致域名解析延迟或不可用。本文建议: ①建议管理员仔细研究和熟悉与权威服务器配置相关的 DNS RFC 文件, 并按照这些 RFC 文件的要求配置权威服务器, 确保不存在 NS 记录配置错误的问题, 例如, 私有 IP 错误、权威服务器主机名错误等。②定期检查权威服务器工作状态, 及时排查无效权威服务器, 以防影响域名解析的稳定性。③域名管理者在添加或删除权威服务器时应

当主动检测父区域和子区域 NS 记录是否配置一致, 以避免不正确的负载均衡、增加解析延迟和权威服务器无响应。

境外基础设施依赖方面, 本文发现一部分域名的权威服务器完全部署在境外, 在面对断网停服等潜在极端场景时存在较大的安全风险。本文建议: ①有对外业务的行业如外贸、金融, 企业可按需部署境外权威服务器, 但为保证正常为中国用户提供稳定服务, 应部署一部分中国的权威服务器, 例如, 选择中国的第三方 DNS 厂商或者在中国境内的基础设施上搭建自有权威服务器。②对于教育域名和政府域名, 除服务港澳台地区的域名, 管理员应该将权威服务器尽量迁移到中国境内的基础设施, 以增强我国对重要域名的控制力。除以上两点, 在重点域名数据集筛选过程中, 我们发现并排除了一些不合规网站, 该类网站采用境外基础设施来规避中国的审查, 本文的境外依赖度测量可以为识别不合规网站提供信息, 缩小审查范围。而对于境外基础设施依赖程度高的域名, 建议加强网络监管力度。

从 IPv4/IPv6 发展角度来看, 基于重点域名服务状态的现状, 目前 IPv6 权威服务器对境外基础设施的依赖度更低, 但无法解析域名的风险更高。在我国大力发展 IPv6 网络的背景下, 有 38.01%(2572 个)的重点域名都无法通过 IPv6 访问, 因此 IPv6 在域名系统, 特别是重点领域的应用还需要加强。

7 总结

本研究深入分析了中国重点域名权威域名服务器的服务状态, 强调了健全可靠的域名系统基础设施对国家网络安全和服务稳定的重要性。研究结果揭示了重点域名权威服务器的部署方式、无效配置情况和地区依赖情况。权威服务器部署和配置的现状表明, 我国域名存在对境外基础设施的依赖, 尤其是在政府、教育和企业等关键领域。此外, 无效权威服务器的问题反映了重点域名无法访问或访问延迟增加的潜在脆弱性。

为了降低中国重点域名的国内可用性风险, 本文提出了几项建议, 包括定期验证权威服务器配置, 增强服务器部署的冗余性和多样性, 减少对境外基础设施的依赖。根据重点域名在 IPv6 网络下服务状态不佳的现状, 本文还建议推进 IPv6 下域名系统的完善和普及。

目前本文的研究对象是少量重点域名, 然而这套方法论同样适用于巨量域名集, 形成更全面的域名服务器态势图景, 保障关注的域名安全运营。我们

未来的研究会集中在持续测量境内域名权威服务器服务状态, 观测其长期趋势, 并从更多角度(例如, 权威服务器软件安全、域名解析链路拓扑、域名劫持风险)展开研究, 以防范可能威胁境内域名系统生态系统的地缘政治或网络威胁。

参考文献

- [1] Griffin W, Schlyter J. RFC 4255 Using DNS to Securely Publish Secure Shell (SSH) Key Fingerprints, <http://www.rfc-editor.org/info/rfc4255> 2023,12,2.
- [2] Kucherawy M, Zwicky E. RFC 7489 Domain-based Message Authentication, Reporting, and Conformance (DMARC), <http://www.rfc-editor.org/info/rfc7489> 2023,12,2.
- [3] Hallam-Baker P, Stradling R, Hoffman-Andrews J. RFC 8659 DNS Certification Authority Authorization (CAA) Resource Record. <http://www.rfc-editor.org/info/rfc8659> 2023,12,2.
- [4] Akiwate G, Sommese R, Jonker M, et al. Retroactive Identification of Targeted DNS Infrastructure Hijacking[C]. *The 22nd ACM Internet Measurement Conference*, 2022: 14-32.
- [5] Pappas V, Xu Z G, Lu S W, et al. Impact of Configuration Errors on DNS Robustness[C]. *The 2004 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, 2004: 319-330.
- [6] Ramasubramanian V, Sirer E G. Perils of Transitive Trust in the Domain Name System[C]. *The 5th ACM SIGCOMM Conference on Internet Measurement*, 2005: 35.
- [7] US Department of Treasury. Specially designated nationals and blocked persons list (sdn) human readable lists, <http://home.treasury.gov/policyissues/financial-sanctions/specially-designated-nationals-and-blocked-personslist-sdn-human-readable-lists> 2022,11,20.
- [8] Mykhailo Fedorov. Letter to goran marby, president and ceo of icann, <https://eump.org/media/2022/Goran-Marby.pdf> 2022,11,20.
- [9] Nikita Korolev. Authorities isolate networks, <http://www.kommersant.ru/doc/5249500> 2022,11,20.
- [10] Han D F, Yuan R X, Guan X H. Measurement and Analysis on DNS Configuration Errors in CN Domain[J]. *Computer Engineering*, 2007, 33(2): 105-107.
(韩殿飞, 袁睿翕, 管晓宏. 中国域名服务器配置错误的测量与分析[J]. *计算机工程*, 2007, 33(2): 105-107.)
- [11] Deccio C, Chen C C, Mohapatra P, et al. Quality of Name Resolution in the Domain Name System[C]. *2009 17th IEEE International Conference on Network Protocols*, 2009: 113-122.
- [12] Jiang J, Zhang J, Duan H X, et al. Analysis and Measurement of Zone Dependency in the Domain Name System[C]. *2018 IEEE International Conference on Communications*, 2018: 1-7.
- [13] Zhou D J, Peng J, Lu C Y, et al. Measurement Study on Security of DNS Authoritative Name Servers in CERNET[J]. *Journal of Information Engineering University*, 2019, 20(4): 467-472.
(周东杰, 彭坚, 陆超逸, 等. CERNET 中权威域名服务器安全问题的测量研究[J]. *信息工程大学学报*, 2019, 20(4): 467-472.)
- [14] Akiwate G, Jonker M, Sommese R, et al. Unresolved Issues: Pre-

- valence, Persistence, and Perils of Lame Delegations[C]. *The ACM Internet Measurement Conference*, 2020: 281-294.
- [15] Sommese R, Moura G C M, Jonker M, et al. When Parents and Children Disagree: Diving into DNS Delegation Inconsistency[C]. *Passive and Active Measurement*, 2020: 175-189.
- [16] Li C, Cheng Y N, Zhang Z X, et al. Configuration Anomaly Detection and Resolution Risk Assessment of Authoritative Domain Name Server[J]. *Computers & Security*, 2023, 135: 103426.
- [17] Bates S, Bowers J, Greenstein S, et al. Evidence of Decreasing Internet Entropy: The Lack of Redundancy in DNS Resolution by Major Websites and Services[J]. *SSRN Electronic Journal*, 2018.
- [18] Allman M. Comments on dns robustness[C]. *Proceedings of the internet measurement conference 2018*, 2018: 84-90.
- [19] Zembruksi L, Jacobs A S, Landtreter G S, et al. Measuring Centralization of DNS Infrastructure in the Wild[M]. *Advanced Information Networking and Applications*. Cham: Springer International Publishing, 2020: 871-882.
- [20] Kaggle. Alexa top 1 million sites, <http://www.kaggle.com/cheed-cheed/top1m> 2022,11,20.
- [21] van Rijswijk-Deij R, Jonker M, Sperotto A, et al. A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements[J]. *IEEE Journal on Selected Areas in Communications*, 2016, 34(6): 1877-1888.
- [22] Le Pochat V, Van Goethem T, Tajalizadehkhoob S, et al. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation[EB/OL]. 2018: arXiv: 1806.01156. <https://arxiv.org/abs/1806.01156>.
- [23] State-Owned Assets Certification Centre. State-Owned Assets Inquiry System, <http://www.guoqi.org/certification/> 2023,12,2.
- [24] 中国教育和科研计算机网. CERNET 服务 IP 地址申请和域名注册的方法. https://www.edu.cn/xxh/zt/cernet/200806/t20080627_305382.shtml 2024,4,12.
- [25] Cisco. Cisco umbrella. <http://umbrella.cisco.com/blog> 2023,12,2.
- [26] Majestic. Majestic million csv now free for all, <http://blog.majestic.com/development/majestic-million-csv-daily/> 2023,12,2.
- [27] Quantcast. Quantcast measure, <http://www.quantcast.com/products/measure-audience-insights/> 2023,12,2.
- [28] Izhikevich L, Akiwate G, Berger B, et al. ZDNS: A Fast DNS Toolkit for Internet Measurement[C]. *The 22nd ACM Internet Measurement Conference*, 2022: 33-43.
- [29] IP-API. IP-api.com geolocation api, <http://www.ip-api.com/> 2023,12,2.
- [30] Barr D. RFC 1912 Common DNS Operational and Configuration Errors, <http://www.rfc-editor.org/info/rfc1912> 2023,12,2.
- [31] Yu Y D, Wessels D, Larson M, et al. Authority Server Selection in DNS Caching Resolvers[J]. *ACM SIGCOMM Computer Communication Review*, 2012, 42(2): 80-86.
- [32] Zhu Y X, Tian D Z, Yan F. Effectiveness of Entropy Weight Method in Decision-Making[J]. *Mathematical Problems in Engineering*, 2020, 2020(1): 3564835.



聂乐遥 于 2021 年在北京邮电大学信息安全专业获得学士学位。现在清华大学网络空间安全专业攻读硕士学位。研究领域为网络测绘、域名系统安全。Email: menaknak@qq.com



李城龙 于 2012 年在清华大学计算机科学与技术专业获得博士学位。现任清华大学副研究员。计算机学会(CCF)会员。研究领域为网络空间测绘与态势感知、网络信息安全。Email: lichenglong@tsinghua.edu.cn



董聪 于 2022 年在中国科学院信息工程研究所网络空间安全专业获得博士学位。现任中关村实验室助理研究员。研究领域为网络测量、网络空间安全态势感知。Email: dongcong@zgclab.edu.cn



宋光磊 于 2023 年在清华大学网络空间安全专业获得博士学位。现任中关村实验室助理研究员。计算机学会(CCF)会员。研究领域为 IPv6 活跃地址探测、网络态势感知、网络测量。Email: songgl@mail.zgclab.edu.cn



张辉 于 1998 年在中国科学院计算技术研究所计算机系统结构专业获得硕士学位。现任清华大学高级工程师。研究兴趣包括网络管理、网络测量、网络安全态势感知等。Email: hzhang@cernet.edu.cn



樊琳娜 于 2023 年在清华大学网络空间安全专业获得博士学位。现任国防科技大学信息通信学院讲师。研究领域为网络测量与网络安全。研究兴趣包括物联网、人工智能等。Email: fanlinna17@nudt.edu.cn



杨家海 于 2003 年在清华大学计算机系获得博士学位。现任清华大学教授。计算机学会(CCF)会员。研究领域为互联网体系结构、网络安全与态势感知。Email: yang@cernet.edu.cn