软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2022,33(1):211−232 [doi: 10.13328/j.cnki.jos.006218]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

域名系统测量研究综述*

刘文峰, 张 宇, 张宏莉, 方滨兴

(哈尔滨工业大学 网络空间安全学院, 黑龙江 哈尔滨 150001)

通信作者: 张宇, E-mail: yuzhang@hit.edu.cn

摘 要: 域名系统 (domain name system, DNS) 测量研究是深入理解 DNS 的重要研究方式. 从组件、结构、流量、安全 4 个方面对近 30 年 (1992–2019) 的 DNS 测量研究工作梳理出 18 个主题. 首先, 介绍组件测量, 组件有解析器和权威服务器两种, 解析器测量包括公共解析器、开放解析器、解析器缓存、解析器选择策略 4 个主题, 权威服务器包括性能、任播部署、托管、误配置 4 个主题. 其次, 阐述结构测量, 包括桩解析器与解析器的依赖结构、解析器间依赖结构、域名解析依赖结构 3 个主题. 然后, 描述流量测量, 包括查询流量特征、异常根查询流量、流量拦截共 3 个主题. 最后综述了安全测量, 包括 DNSSEC 代价与隐患、DNSSEC 部署进展、加密 DNS 部署、恶意域名检测 4 个主题.

关键词: 域名系统测量; 网络空间测绘; 互联网测量

中图法分类号: TP393

中文引用格式: 刘文峰, 张宇, 张宏莉, 方滨兴. 域名系统测量研究综述. 软件学报, 2022, 33(1): 211–232. http://www.jos.org.cn/1000-9825/6218.htm

英文引用格式: Liu WF, Zhang Y, Zhang HL, Fang BX. Survey on Domain Name System Measurement Research. Ruan Jian Xue Bao/Journal of Software, 2022, 33(1): 211–232 (in Chinese). http://www.jos.org.cn/1000-9825/6218.htm

Survey on Domain Name System Measurement Research

LIU Wen-Feng, ZHANG Yu, ZHANG Hong-Li, FANG Bin-Xing

(School of Cyberspace Science, Harbin Institute of Technology, Harbin 150001, China)

Abstract: Domain name system (DNS) measurement research is an important way to understand DNS. This paper reviews the DNS measurement work during 1992 and 2019 on 18 topics from four aspects of components, structure, traffic, and security. Firstly, in the aspect of components, the four resolver-related topics are on public resolver, open resolver, resolver caching, and resolver selection policy; the four authoritative-server-related topics are on performance, anycast deployment, hosting, and misconfigurations. Secondly, in the aspect of structure, there are three topics: the dependency structure between stub resolvers and resolvers, the dependency structure of resolvers, and the dependency structure of domain name resolution. Then, in the aspect of traffic, there are three topics: query traffic characteristics, abnormal root query traffic, and traffic interception. Moreover, in the aspect of security, there are four topics: DNSSEC cost and risk, DNSSEC deployment, DNS encryption deployment, and malicious domain name detection. Finally, future research topics are discussed.

Key words: DNS measurement; cyberspace surveying and mapping; Internet measurement

1 引 言

域名系统 (domain name system, DNS) 是提供域名解析服务的关键互联网基础设施, 常用于将域名翻译成 IP 地址. DNS 随着互联网的演化不断发展, 规模不断扩大, 根服务器从原有 13 个部署点扩展为 1 000 多个[1]; 截至 2020 年一季度已注册的域名多达 3.66 亿[2]; DNS 功能日渐复杂, DNS 相关的 RFC 有近 200 个[3]. 大量服务已经围绕 DNS 形成一个复杂的生态系统, 测量成为认识和改进 DNS 的必要手段.

* 基金项目: 国家重点研发计划 (SQ2018YFB1800702, 2016YFB0801303)
收稿时间: 2020-07-05; 修改时间: 2020-11-10; 采用时间: 2020-12-04; jos 在线出版时间: 2021-01-15

DNS 测量的研究意义在于: (1) 观察 DNS 中各组件运行情况, 有助于日常维护工作, 例如权威服务器的性能随时间变化状况[4]、随地理位置变化情况[5−7]; 解析器解析选择偏好[8]、持续可用性和有效性[9]. (2) 发现利用 DNS 或针对 DNS 的攻击行为, 有助于了解风险发展趋势以制定相应对策, 例如缓存下毒攻击[10]和 DNS 隐匿信道[11,12]. (3) 考察新技术的应用效果, 有助于推动 DNS 持续改进, 例如任播部署方案[13], DNSSEC 安全扩展方案[14]和 DoT[15]、DoH[16]加密方案.

在 DNS 运行的 30 多年间, 学术界和工业界累计发表了数百篇 DNS 测量文献, 文献 [17–20] 已经对 DNS 安全进行了综述, 但目前缺乏对 DNS 测量研究的综述. 本文将相关工作归纳为以下 4 类: (1) 组件测量, 侧重研究 DNS 组件特征, 见第 2 节. (2) 结构测量, 侧重研究组件间依赖结构, 见第 3 节. (3) 流量测量, 侧重研究 DNS 运行时的状态, 见第 4 节. (4) 安全测量, 侧重研究 DNS 安全协议评估和安全方案部署问题, 见第 5 节. 本文组织结构和覆盖的研究主题如图 1 所示.

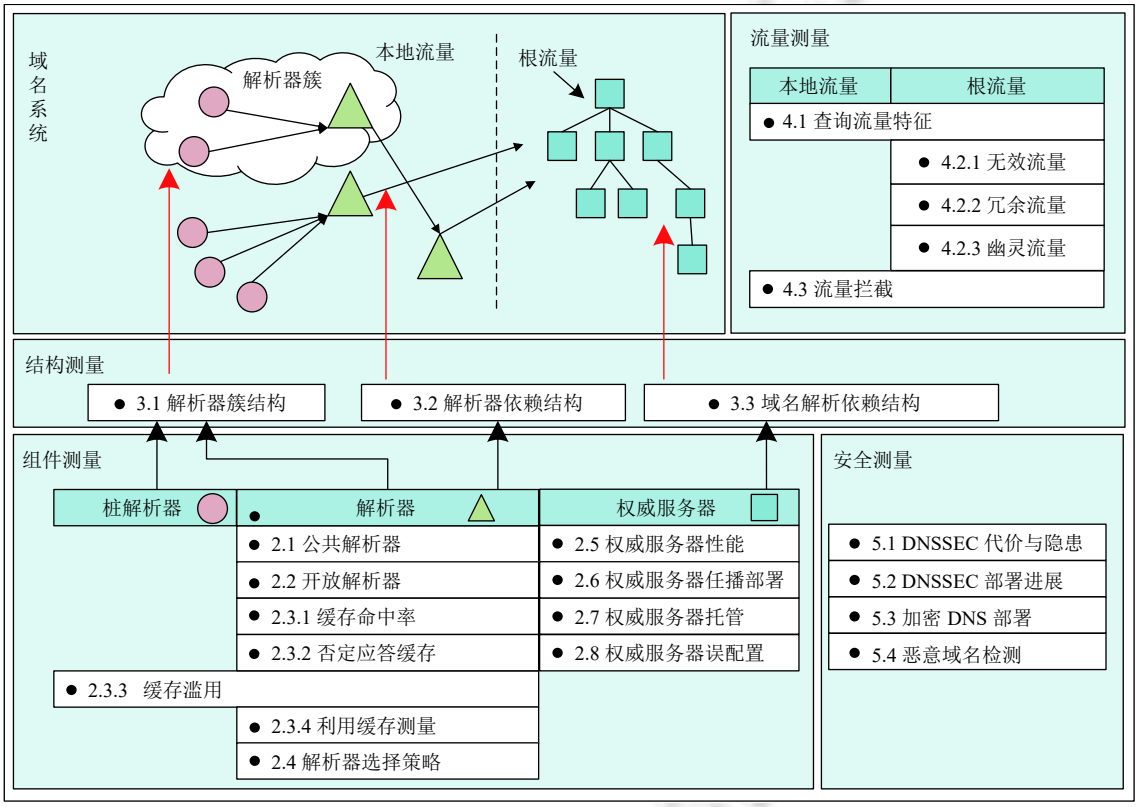

图 1 DNS 测量研究主题示意图

2 组件测量

DNS 中的组件按功能可分为 2 类: 权威服务器和解析器 (又称缓存解析器或递归服务器), 桩解析器 (又称用户或终端主机) 是 DNS 的用户. 桩解析器向解析器发起递归查询后, 等待解析器的应答; 解析器收到查询后, 向各级权威服务器发送迭代查询, 得到的权威应答后缓存并回复桩解析器; 权威服务器负责响应查询, 回复权威应答或子域权威服务器索引. 针对解析器的测量研究关注规模、时延、应答行为、缓存机制和选择机制等方面; 针对权威服务器的测量研究侧重关注性能情况、部署情况、数据有效性和一致性.

2.1 公共解析器

规模、时延: 公共解析器 (public DNS resolver), 又称第三方解析器 (third-party DNS resolver), 是由云服务商、

内容服务商、安全服务商等商业公司部署的对外提供开放域名递归解析服务的服务器. 2010 年前后, 公共解析器的用户在短时间内快速增长, 但与 ISP 解析器相比部署覆盖度仍显不足, 且解析时延偏高. 文献 [21] 于 2011 年 9 月至 2012 年 1 月通过在 VuzeBitTorrent 客户端安置插件获取用户配置的解析器信息和域名解析数据, 并结合 EdgeScope[22]项目从 VuzeBitTorrent 客户端积累的 2010 年 4 月–2011 年 11 月的历史测试数据发现, 截至 2011 年 12 月, 公共解析器服务已经覆盖了 8.6% 的互联网用户, 21 个月内增长 27%. 文献 [23] 于 2011 年通过在用户访问的网页内嵌入 DNS 查询代码的方式测量用户配置的解析器信息, 测量发现英国电信和美国 AT&T 网络中部署解析器的/24 网段数量分别为 377 和 679, 同时期谷歌公共解析器仅在 12 个/24 网段部署. 文献 [24] 于 2012 年使用 PlanetLab 平台测量发现, 谷歌公共解析器只有 46 个不同地理位置的部署点, 数量甚至低于同时期 CDN 服务商 Akamai 的 260 个, 解析器数量过少不利于 CDN 的数据精准分发. 文献 [25] 于 2010 年测量解析器发现, ISP 解析器平均解析时延为 11 ms, 显著低于公共解析器的 24 ms. 文献 [23] 于 2011 年使用 PlanetLab 平台测量发现, 使用公共解析器会给 10% 的用户带来 50–80 ms 的额外解析时延.

2.2 开放解析器

规模: 开放解析器 (open DNS resolver) 是指对外提供开放域名解析的服务器. 公共解析器是它的子集, 仅特指因商业行为主动提供域名解析服务的行为. 当前互联网中绝大部分的开放解析器是由于错误配置、管理不当或被恶意操纵而对外提供域名解析服务, 其数量庞大, 生命周期短, 解析结果存在较大安全隐患. 文献 [26] 于 2013 年 11 月–2014 年 2 月扫描全网共发现 2 300 万–2 550 万开放解析器. 文献 [9] 于 2014 年 1 月至 2015 年 6 月期间扫描全网发现 1 700 万–2 600 万开放解析器, 发现数量与 Open Resolver Project[27]工作相似, 两者数量仅相差 2%.

应答行为、生命周期: 文献 [28] 于 2007 年用易被“钓鱼”的敏感域名测试 60 万开放解析器发现, 2.4% 开放解析器会回复不正确应答. 文献 [9] 于 2014–2015 年测量发现, 52.2% 的开放解析器在被发现后 1 周内关闭, 发现的开放解析器在 1 年后仅 4.0% 存活. 该工作通过对扫描发现的开放解析器使用 155 个代表性域名做解析测试发现, 其中有超过 300 万开放解析器会篡改解析结果. 文献 [29] 设计了针对开放解析器的 18 种缓存注入攻击, 并于 2017 年测试其缓存行为和抗缓存注入能力, 发现 97% 的开放解析器会被至少 1 种攻击方式得手.

2.3 解析器缓存

2.3.1 缓存命中率

与缓存时间的关系: 缓存命中率受域名资源记录的缓存时间 (TTL 值) 影响. 文献 [30] 于 2001 年分析本地社区 DNS 流量发现, 缓存命中率随 TTL 值的变化呈对数分布, 当 TTL 值超过 600 s 后, 继续增加 TTL 值只能提高很少的缓存命中率. 文献 [31] 于 2009 年测量发现解析器的缓存命中率平均在 70%–80% 之间, 当解析器缓存功能关闭时, 权威服务器收到的查询量仅增加了 1 倍, 说明由于操作系统缓存和用户浏览器缓存的存在, 解析器缓存对降低权威服务器负载的作用只占 50%, 剩下的 50% 对查询负载由安装在用户端的缓存机制分担. 测量发现资源记录的平均 TTL 值呈降低趋势, TTL 值小于 5 min 的 A/AAAA 类型应答占比在 2002 年不足 10%[30], 2012–2013 年约为 30%–40%[32–34], 2017 年为 52%[35].

与查询类型的关系: 文献 [33] 于 2012 年通过重放真实 DNS 查询, 分析解析器缓存命中率与查询类型的关系. 测量发现 CNAME 类型查询的缓存命中率最低, 与剔除重放流量中的 CNAME 查询后相比, 使用 CNAME 查询会降低 10% 的缓存命中率, 并会使对根服务器的查询量增加 10 倍, 对顶级域权威服务器的查询量增加 2.8 倍.

2.3.2 否定应答缓存

否定应答缓存: 文献 [36] 于 2019 年测量解析器对否定应答缓存功能的支持情况, 使用 7 174 个 RIPE Atlas 观测点测量 900 个解析器发现, 仍有 12.07% 的解析器未开启或不支持否定应答缓存, 这类解析器在各大洲的占比为 9.87%–14.89%.

2.3.3 缓存滥用

检测方法: 缓存滥用是指违反 DNS 规范延长资源记录缓存时间的行为, 也称为缓存延期. 总结已有测量 DNS 缓存滥用行为的测量文献 [34,37–39] 发现, 检测 DNS 缓存滥用行为的方法可分为 3 类, 详见表 1: (1) 基于用户流量的检测方法[34,38], 判断的依据是比较用户收到应答报文与后续 TCP/UDP 请求的时间差是否超过 TTL 值, 这种方法可以发现用户 (包含桩解析器和安装在用户终端的应用程序, 下同) 的缓存滥用行为. (2) 基于服务端流量的检

测方法[39], 判断的依据是查看用户是否会被 DNS 应答引导至特定的服务器, 未能向预设目标发起请求的用户将被判定存在缓存滥, 用这种方式可以发现解析器和用户的缓存滥用行为; (3) 使用直接验证解析缓存的方法[37], 判断的依据是比较解析器收到的应答报文中资源记录的 TTL 与缓存值的 TTL 值是否一致.

表 1 缓存滥用行为的检测方法对比

测量方法	类别	文献列表	可检测的缓存滥用对象	
			解析器	用户(桩解析器和应用程序)
基于用户流量检测	被动测量	[34,38]	—	Y
基于服务端流量检测	被动测量	[39]	Y	Y
验证解析器缓存检测	主动测量	[37]	Y	—

滥用行为测量: 文献 [39] 于 2004 年采用基于服务端流量的检测方法测量发现, 47% 的用户使用延长缓存时长的 A 记录; 14% 解析器存在缓存延期现象, 其中 25% 的解析器缓存延期达 5 小时. 文献 [38] 于 2012 年采用基于用户流量的检测方法测量发现, 在缓存滥用行为最严重的 ICSI 社区中, 30.1% 的 TCP/UDP 连接使用了延长缓存时长的 A/AAAA 记录中的 IP 地址. 文献 [34] 于 2013 年在文献 [38] 的基础上用相同的方法发现, CCZ 社区中 40% 资源记录的缓存延期达长 1 天以上, 资源记录的平均延长时间为 168 s. 文献 [37] 于 2013 年采用直接验证解析器缓存的检测方法测量 25 000 个解析器发现, 35% 解析器平均缓存延期超过 1000 s, 其中缓存行为可分为 4 类: (1) 18% 解析器写入缓存的 TTL 值大于应答报文中的 TTL 值; (2) 9% 解析器延缓 TTL 值在缓存中的衰减速度; (3) 7% 解析器在缓存中不衰减 TTL 值; (4) 1% 解析器增加本该随时间衰减 TTL 值.

2.3.4 利用缓存测量

热度估算: 利用解析器缓存估算域名热度, 即通过测量域名在解析器缓存中的存在时间, 估算使用该域名的用户规模, 是一种服务无关、协议无关且不泄露用户隐私的通用测量估算方法, 由文献 [40] 于 2003 年首先提出. 原理如图 2 所示, 通过对目标域名缓存命中情况进行周期性探测, 计算出缓存失效间隔后结合查询分布规律推算目标域名的热度, 域名热度越高则缓存在解析器中的失效间隔越短. 文献 [41] 于 2008 年测量发现域名查询与时间的关系符合指数分布, 并将该分布结果应用于文献 [40] 提出的估算模型, 在本地局域网内估算域名 www.google.com 的热度值, 估算误差率 (|估计值−实际值|/实际值) 为 10.58%. 文献 [42] 使用 2010–2014 年的 9 份来自解析器的数据发现, 不同地理位置的用户群体的查询分布规律存在差异, 提出基于不同地理区域权值的用户查询超指数分布模型, 该模型的拟合准确率为 85.9%, 优于指数分布的 32.3%、伽马分布的 37.5%、韦布尔分布的 63.1%. 利用该模型估算远程网络中访问了恶意域名的受感染主机数量证实, 误差率约为指数分布模型的一半.

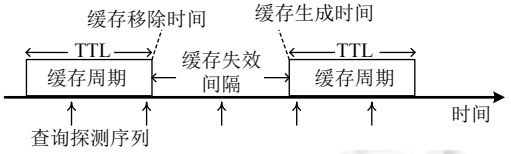

图 2 利用解析器缓存估算域名热度的测量方法示意

时延估算: 文献 [43] 于 2002 年提出估算任意主机间时延的 King 方法, 其原理如图 3 所示. 以估算主机 A、B 两点间时延为例, 首先选取邻近 A、B 最近的 1 个公共解析器 A_Resolver 和 1 个权威服务器 B_Auth; 然后构造 2 个查询, 使第 1 个查询确保 A_Resolve 缓存不命中且 B_Auth 应答 (步骤 1、2、3、4), 第 2 个查询确保 A_Resolve 缓存命中 (步骤 1、4), 通过计算两次查询时延的时间差计算 A_Resolver 和 B_Auth 间的时延, 以此作为 A、B 两点间时延的估算值. 但方法存在如下局限性: (1) B_Auth 可能有多个部署在不同地理位置的从服务器, 干扰准确性; (2) A_Resolve 是转发解析器时, 时延估算结果将会偏大; (3) 会给解析器 A_Resolver 插入大量不必要的缓存记录造成污染. 文献 [44] 于 2008 年提出 T-King 方法对 King 存在上述 3 问题进行了改进, T-King 测量方法只需要消耗 King 测量方法一半带宽且减少了 King 方法造成的缓存污染几个数量级. 同时, T-King 维护了一个域名服务器列表, 排除了干扰准确性的权威服务器和解析器.

丢包率估算: 文献 [45] 在 King 方法的基础上提出 Queen 方法估算任意主机间的丢包率. Queen 首先使用

King 方法估计任意两点间的时延, 然后利用解析器内置的重传机制, 进而从观察到的过长时延推断解析器与权威服务器之间的丢包重传次数, 计算丢包率. 该文献于 2009 年使用 Queen 方法测量由 PlanetLab 组建的 260 条网络路径的丢包率, 并与真实丢包率比较发现, 对 85% 的路径的丢包率估算值与真实值误差小于 1%.

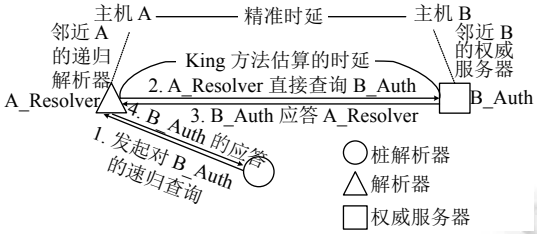

图 3 利用解析器缓存估算任意主机间时延的测量方法示意

2.4 解析器选择策略

时延偏好: DNS 规范要求“解析器迭代查询时从权威服务器列表中选择认为最优的目标发起查询”(RFC2182[46]), 多数解析器将时延作为判断最优权威服务器的依据. CAIDA 于 2003 年通过重放查询日志, 测量主流解析器软件在 4 种不同网络环境下的选择策略发现[47]: (1) BIND9 依据时延选择权威服务器, 在所有根服务器和顶级域权威服务器处于无时延、固定时延和网络不可达环境下都能将查询均匀分配到所有权威服务器, 且当不同的根服务器处于网络时延线性增加的不同环境下时, 其被选择的频率随时延线性增加呈指数下降趋势. 文献 [8] 于 2017 年利用 8 500 个 RIPE Atlas 观测点, 测量真实环境下的解析器选择策略发现, 59%–69% 的解析器通过时延选择权威服务器, 且时延差越大, 解析器选择偏好越明显. 测量还发现 75%–96% 的被测解析器会周期检查所有可用权威服务器的时延.

次优选择: 次优选择行为指解析器在未能选择时延最低的权威服务器发起查询. 文献 [47] 于 2003 年仿真测量发现: (1) DJBDNS 始终均匀选择权威服务器; (2) Win2000 和 Win2003 会倾向于随机选择某一个和某几个权威服务器, 平均选择和随机选择都会导致解析器做出次优选择. 文献 [48] 于 2012 年在仿真环境中, 通过重放真实 DNS 查询测量解析器的次优选择行为发现: (1) DNSCache 解析器未将时延作为选择依据; (2) Unbound 解析器在时延低于 400 ms 的权威服务器中随机选择; (3) BIND 9.8 会将 23% 的查询发向未响应权威服务器, BIND 9.7 为 11%; (4) PowerDNS 和 Unbound 解析器都需要较长时间发现已经降低了时延的权威服务器, 其中 Unbound 最坏情况下需要 15 min.

2.5 权威服务器性能

权威服务器性能测量研究分为 2 类: (1) 测量权威服务器性能随时间波动状况; (2) 测量不同地理位置用户感知的权威服务器性能差异.

时间波动: 文献 [4] 于 2000–2001 年使用本地观测点, 被动测量根服务器和顶级域权威服务器的性能随时间的变化情况, 通过比较查询量、响应时延、丢包率等指标随时间的波动情况, 推测造成根性能表现异常的原因. 文献 [49] 于 2002–2005 年使用 100 多个观测点主动测量根服务器的响应时延和丢包率. 文献 [50] 于 2002–2017 年使用 5 个本地观测点被动测量本地发向根的查询与应答, 分析 13 个根的查询量、应答时延、丢包率随时间变化情况. 文献 [51] 于 2008 年测量顶级域 cn 的 6 个权威服务器, 共 15 个部署点的性能随时间变化情况.

地理差异: 文献 [5] 于 2002 年通过 75 个不同网络的观测点解析 Larbin crawler 数据集[52]中的 10 万域名, 比较不同地理位置用户感知的权威服务器性能差异. 测量发现不同地理位置的域名解析平均时延在 0.95–2.31 s 之间, 极值相差 2.4 倍, 其中, 根查询时延在 0.063–1.41 s 之间; 顶级域查询时延在 0.037–0.89 s 之间. 文献 [7] 于 2003 年被动测量各大洲用户对根服务器的感知性能差异发现, 基于用户感知时延可将所测的 10 个根服务器聚类到 4 个组中, 同组内的根服务器性能差异小到可以忽略不计. 测量发现欧洲和亚洲的根服务器负载过重导致时延过大, 解析器会因此舍近求远访问美洲的根服务器, 导致用户对根的感知时延偏高. 文献 [6] 于 2013 年使用 19 593 个公共解析器测量根在不同洲的性能差异发现, 根在五大洲时延均小于 100 ms, 但时延差异依旧明显, 北美洲和欧洲最优而非洲最差, 两者最有 6 倍时延差.

2.6 权威服务器任播部署

性能、可用性、稳定性: F 根于 2002 年率先使用跨地域任播部署方案, 随后两年内 C、K、M 根也实施任播

部署. 文献 [53] 于 2005 年从服务端视角被动测量 C、F、K 根的任播部署点并对其评估发现, 查询 C 根和 F 根的解析器中只有不到 2% 会在任播节点间跳转 (flipping), 表明任播部署对根的查询稳定性影响很小. 文献 [54] 于 2004 年通过 400 个 PlanetLab[55]探测点对比已任播部署的 F、K 根和未任播部署的 B 根在性能、可用性、稳定性上的差异. 测量发现: (1) 任播根节点的平均查询时延显著降低, 非任播根的平均时延为 115 ms, 而任播根的平均时延为 75 ms. (2) 任播根节点出现未应答查询, 导致解析不可用的比例小于 0.9%, 与未任播部署的根持平. (3) 任播根节点与解析器间稳定性良好, K 根和 F 根只有不超过 0.006% 的解析器在任播根节点间跳转的现象.

抗 DDoS 攻击: 文献 [56] 于 2015 年测量根任播节点在 DDoS 攻击下 (100 倍常规查询量) 的抗攻击能力, 使用从 RIPE Atlas 测量平台[57]、根服务器运营商[1]和 BGPmon 项目[58]获取的 30 天公开数据集, 分析 13 个根的上百个任播节点受到 DDoS 攻击时的情况. 测量结果显示, 任播根面临 DDoS 攻击时, 少数负责吸收攻击的任播根节点的会不堪重负 (丢包率>95%), 以保护大部分任播根节点免受 DDoS 攻击的影响 (丢包率<1%), 总体上降低了受影响用户数量.

任播节点发现: 文献 [59] 于 2011–2012 年使用测量平台 PlanetLab[55]和 Netalyzr[60]对比多种发现任播节点方法, 提出两种优化发现方案: (1) 基于 CHAOS+Traceroute 探测非合作任播节点; (2) 利用 30 万公共解析器作观测点扩大探测区域. 优化后的测量方案可正确枚举 93% 的 F 根任播节点, 评估后认为该探测方法召回率达到 80% 时需要至少 10 000 的观测点. 测量还发现超过 72% 顶级域已经采用任播部署, AS112 项目[61]中超过 26% 的任播节点已经失效.

2.7 权威服务器托管

DNS 托管服务商在全球范围内部署 DNS 权威服务器为域名权威提供安全经济的域名解析服务.

同步效率: 文献 [62] 于 2018 年使用 1000 个 RIPE Atlas 节点测量 Google 等 8 个托管服务商的资源记录更新效率发现, 权威服务器间的资源记录的更新传播时延通常为十几秒, 时延最高的 Google Cloud 需 50 s.

服务器共享: 托管服务的流行使权威服务器逐渐向少数服务商集中. 文献 [63] 于 2013–2018 年测量热度前 100 万域名的二级域权威服务器的共享程度发现, 91%–93% 的二级域存在共享权威服务器的情况, 一组权威服务器最多被 9 000 个二级域共享. 这种集中的趋势对 DNS 鲁棒性造成威胁.

攻击影响: 文献 [64] 利用 OpenIntel[65,66]平台测量攻击对用户选择托管服务商的影响发现, 2016 年 5 月针对 NS1 的攻击和 10 月针对 Dyn 的攻击会令用户同时使用多个托管解析服务商以分担风险.

2.8 权威服务器误配置

区数据误配置: 文献 [67] 于 2003 年测量 3 类误配置: 跛脚授权 (lame delegation)、冗余不足 (diminished server redundancy) 和循环依赖 (cyclic zone dependency), 发现 1.8 万区 (zone) 中存在上述 3 类误配置的占比分别为 15%、15%、2%. 文献 [68] 于 2007 年测量 494 万二级域 (占 com 和 net 的 6.6%) 的区文件并从中检测到 4 类误配置: (1) 使用无效、过时、未分配的资源记录类型, 占 0.008%; (2) 域名结尾缺少点号导致后缀重复, 占 0.12%; (3) 包含已授权给子域的域名, 占 0.48%; (4) 包含重复的域名, 占 0.14%. 文献 [69] 于 2009 年分析 300 万域名的误配置情况发现: (1) 父域胶水记录中配置的子域权威服务器有 2.5% 不可用, 有 1.2% 只能做出非权威应答; (2) 0.095% 的区存在循环依赖. 文献 [70] 于 2009 年测量 6 个顶级域下 1.06 亿二级域的发现, 权威服务器中 1.7% 为孤儿服务器, 即在子域出现但在父域并不存在的权威服务器, 孤儿服务器平均只有 8–9 天的生存期, 不到 2% 的孤儿服务器存活超过 30 天.

软件误配置: 文献 [71] 测量发现部分权威服务器未禁止非安全动态更新, 即权威服务器不限制动态更新请求的 IP 地址、也不验证身份密钥. 该工作于 2016 年利用 DNSDB[72]和 Project Sonar Data Repository[73]两个数据集测量 2 860 万个域的权威服务器的动态更新行为发现: (1) 1 877 个域 (占比 0.065%) 的 188 个权威服务器可以成功被任意请求者更新, 这些域集中在政府、大学等机构; (2) 误配置域中有 66.2% 被托管日本网络运营商的基础设施上, 托管误配置域最多的 10 个运营商覆盖了 88.6% 的误配置域.

综上, 组件测量研究主题多样, 表 2 对组件测量的代表性工作进行了总结, 发现了 DNS 中多方面的问题. 在解析器方面, 公共解析器规模和性能还有待提升, 开放解析器普遍存在安全隐患. 解析器中存在不支持否定应答缓存以及普遍的缓存滥用行为. 在权威服务器方面, 任播部署尚未解决根服务器和顶级域服务器服务的地理不均衡问题. 域名托管服务趋势削弱了 DNS 整体鲁棒性, 误配置则会直接影响 DNS 的可用性.

表 2 DNS 组件测量相关工作比较

节序号、主题 (测量对象)	文献	观测点类型	观测点数量	主被动	◆平台 ★工具 ● 数据	实验时间 (周期)	数据公开	关键结论
2.1 公共解析器	[21]	C	10^4	主	—	2011–2012 (127天)	否	公共解析器的平均解析时延是ISP解析器的2倍
	[23]	C	10^6	主	◆PlanetLab	※2011 (8周)	否	观测点与公共解析器的平均距离比ISP解析器远1000 km
2.2 开放解析器	[9]	C	—	主	—	2014–2015	否	数百万开放解析器存在拦截篡改、重定向等恶意行为
	[29]	C	10^4	主	—	2015–2017	否	97%开放解析器存在至少1种缓存注入漏洞
2.3.1 解析器缓存命中率与否定应答缓存	[30]	R^*-A	—	被	—	2001–2002 (20天)	否	TTL值超过10min时, 增加TTL值只能少量提升的缓存命中率
	[31]	A	—	被	★tcpdump	2009(7天)	否	TTL大幅降低不会显著增加本地查询量.
	[36]	C	10^4	主	◆RIPE Atlas	—	否	12.07%的解析器不支持否定应答缓存
2.3.2 解析器缓存滥用	[34]	*C-R	—	被	—	2011–2012 (14月)	否	13.7%的TCP连接使用过期解析结果
	[39]	A	—	被	—	2003	否	14%解析器存在缓存滥用行为
	[37]	C	10^2	主	◆PlanetLab	2012–2013 (131天)	是	81%解析器存在缓存滥用行为
2.3.3 利用解析器缓存的测量	[40]	C	—	主	—	2003 (1周)	否	本地局域网内估算域名www.google.com的热度值, 估算误差率为10.58%
2.4 解析器选择策略	[47]	*R-A	—	被	★NeTraMet	2003 (3月)	否	不同解析器的选择策略迥异, BIND基于RTT选择, DJBDNS均匀选择 WindowsDNS随机选择
	[8]	C, A	10^4	混	◆RIPE Atlas ●DITL-2017	2017	是	绝大多数解析器倾向于将低时延作为选择策略的依据
2.5 权威服务器性能	[4]	*R-A	—	被	★NeTraMet	2000–2001	否	测量网络内的用户偏好使用A根和F根
	[5]	C	10^2	主	★dnsperf ●Larbin	2002 (3月)	否	递归解析过程中查询根服务器和顶级域服务器耗时共占整个递归解析耗时的60%
	[7]	A, R-*A	—	混	★CAIDA skitter	2002	是	全球根服务器地理分布不均, 欧洲和亚洲根性能不足, 美洲根性能过剩
	[6]	R	10^4	主	—	※2013	否	非洲和南美查询根服务器时延远高于欧洲和北美地区
2.7 托管权威服务器	[62]	C, R	10^3	主	◆PlanetLab ◆RIPE Atlas	2018	是	DNS托管服务商的数据更新时延小于1 min
	[63]	A	—	被	—	2013–2018	否	91%–93%二级域至少与其它1个二级域共享权威服务器
	[64]	C	—	主	◆OpenIntel	2016	否	DDoS攻击令用户改变DNS托管服务商
2.6 根的任播部署	[53]	A	—	被	★tcpdump	2006 (2天)	是	52%C根用户, 35%F根用户, 29%C根用户查询被引导至最近任播根
	[54]	C	10^2	主	◆PlanetLab	2004 (20天))	否	37%–80%用户查询被引导至最近任播根
	[56]	C, A	10^4	混	◆RIPEAtlas ◆BGPmon	2013–2017	是	DDoS攻击中, 吸收攻击流量的任播根丢包率最高达95%, 保护其它任播根仅有1%丢包率
	[59]	C, R	10^5	注	◆Netalyzr ◆PlanetLab	2011–2012	否	提出一种枚举任播服务中所有部署点的方法, 该方法可枚举93%的F任播根
2.8 权威服务器误配置	[67]	A, *R-A	—	混	●ISC reverse zone data	2003 (16天)	是	存在跛脚授权、冗余不足、循环依赖3种误配置的区占15%, 15%, 2%
	[70]	R	—	被	—	2009 (15天)	否	1.7%的二级域权威服务器是孤儿服务器

注: C为客户端、R为解析器、A为权威服务器, *X-Y为X与Y信道中靠近*的一侧, ※表示论文发表时间

3 结构测量

DNS 组件间的依赖关系形成 3 类依赖结构, 如图 4所示: (1) 桩解析器依赖递归解析器形成“解析器簇”结构; (2) 转发解析器与递归解析器之间的依赖结构; (3) 由授权依赖和别名依赖形成的域名解析依赖结构.

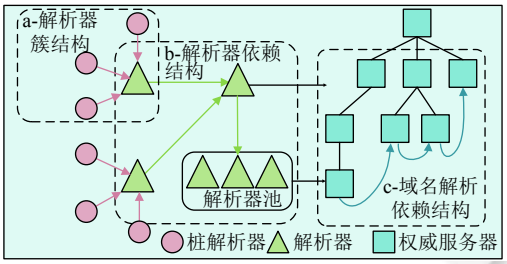

图 4 DNS 组件间形成的 3 种依赖结构

3.1 解析器簇结构

解析器簇结构 (图 4) 由桩解析器和为其提供解析服务的递归解析器组成. 相关研究关注解析器簇结构带来的负载隐藏问题 (hidden load problem)[74]和发起人识别问题 (originator problem)[75].

负载隐藏问题: 解析器簇规模对权威服务器不可见, 导致服务点无法准确预估负载. 如图 5 所示, 不同解析器簇中桩解析器数量差异巨大, 服务点负载不均衡. 文献 [76] 于 2013 年通过大规模投放广告测量 1 137 万桩解析器组成的 27 万个解析器簇发现: (1) 解析器簇规模分布差异巨大, 超过 90% 解析器簇仅覆盖 1% 桩解析器, 少数大象簇拥有大量桩解析器; (2) 解析器簇活跃度分布同样差异巨大, 最大的 10 个大象簇发出超过 80% 的桩解析器查询.

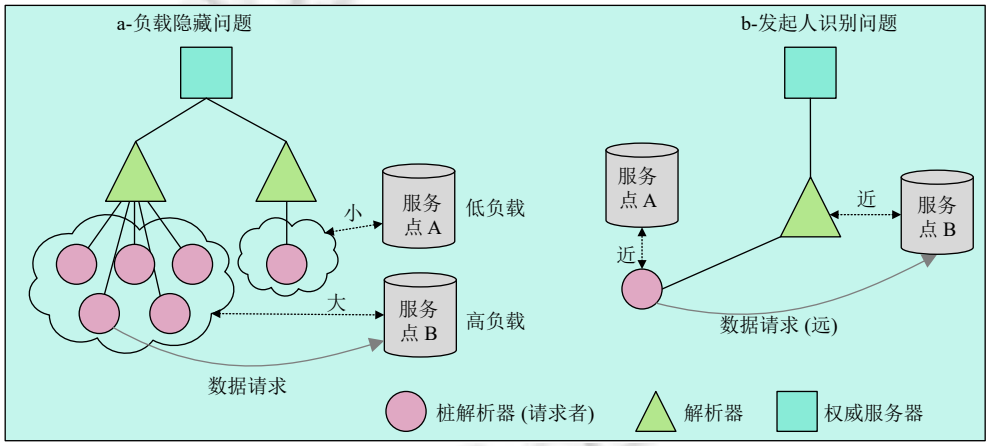

图 5 解析器簇结构引发的“负载隐藏”问题和“发起人识别”问题

发起人识别问题: 桩解析器位置对权威服务器不可见, 导致桩解析器未被引导至最优服务点. 如图 5 所示, 服务点用解析器位置估算与桩解析器的距离时偏差较大. 文献 [77] 于 2002 年从 AS 级和 BGP 前缀级测量解析器簇内桩解析器与解析器的网络位置关系发现, 仅有 64% 的桩解析器与解析器在同一自治域. 文献 [23] 于 2011 年测量 ISP 发现桩解析器与 ISP 解析器的 (80 百分位数) 距离为 253 km, 与公共解析器的距离为 1 358 km. 文献 [21] 于 2011 年 9 月至 2012 年 1 月测量发现, 桩解析器与解析器网络距离对 HTTP 请求时延有明显的影响, 使用 ISP 解析器的时延只有使用公共解析器的一半, 与文献 [25,78,79] 结论相似. 文献 [76] 于 2013 年测量发现, 15% 的解析器簇内有超过 50% 的桩解析器与解析器在不同自治域, 10% 的解析器簇内桩解析器与解析器都在不同自治域.

通过使解析器在 DNS 查询中附加桩解析器 IP 地址的/16 或/24 前缀后 (edns-client-subnet, ECS)[80], 上述两问

题得到较好的解决. 文献 [81] 测量发现 Akamai 等内容分发服务商使用 ECS 方案后能显著提升数据分发的准确性, 用于评估数据分发性能的映射距离、RTT/RTTB、数据下载时间等指标均有改善.

3.2 解析器依赖结构

部分网络服务商仅在本地部署转发器, 将域名解析委托给上游的解析器. 随着域名解析任务被多级委托, 解析器之间形成复杂的依赖结构, 如图 6 所示. 间接解析器专门用于为转发器提供服务且不为桩解析器提供服务.

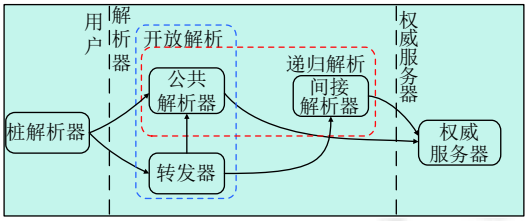

图 6 解析器依赖结构示意图

解析器间依赖: 文献 [28] 于 2007 年测量解析器之间的依赖关系发现, 扫描 IPv4 地址空间测得的 1700 万提供开放解析服务的解析器中有 96.4% 是转发器, 其解析依赖于 58 万递归解析器, 其中 71% 转发器仅依赖于 1 个递归解析器. 文献 [37] 于 2012–2013 年通过高频率扫描 IPv4 地址空间发现, 开放解析器数量是先前工作的[82,83]两倍 (33 M vs. 15 M), 50% 的间接解析器被超过 10 个转发器使用, 80% 的间接解析器被超过 1 个转发器使用.

解析器池: 解析器池由一组共享缓存但各自独立提供服务的解析器组成, 见图 4. 解析器池利于提高缓存利用率和命中率. 文献 [25] 于 2010 年测量发现, 部分解析器池的缓存共享配置不当, 导致对同一域名连续查询时解析时延并未降低. 文献 [76] 于 2013 年测量发现 40 万个解析器池, 覆盖解析器总量的 13%, 为超过 90% 的桩解析器提供服务.

3.3 域名解析依赖结构

传递依赖风险: 域名之间通过授权 (NS) 和别名 (CNAME) 产生依赖关系. 当被依赖的域名不安全时, 威胁将会传递给依赖它的域名, 称为传递依赖风险. 文献 [84] 于 2005 年首先指出该类风险并通过域名授权图 (delegation graph) 予以量化. 该文献分析 196 个顶级域下共 59 万域名的解析依赖关系后发现: (1) 每个域名解析平均依赖 46 个权威服务器; (2) 6.5% 的域名依赖超过 200 个权威服务器; (3) 30% 的域名解析依赖存在瓶颈, 劫持该域名最少只需控制 2 个权威服务器. 文献 [69] 构建服务器依赖结构图 (server dependency graph) 并于 2010 年分析 ODP/SC08 数据集[85]内 310 万域名发现, 解析域名所需最小权威服务器数量的平均值为 3.48, 使域名解析不可用的最小权威服务器数量平均值为 2.34. 文献 [86] 将域之间的依赖关系分为 4 种类型: 一般依赖、显示依赖、关键依赖和基本依赖, 并分别对这 4 中依赖关系构建域名依赖图, 并于 2013 年测量热度前 100 万域名发现, 在关键依赖图中存在依赖循环的域为 168 个 (0.02%), 显示依赖图中为 7 504 个 (0.72%), 一般依赖图中为 60 598 个 (5.74). 文献 [87] 于 2019 年测量热度前 100 万域名的解析依赖关系发现, 超过半数的域名, 若解析该域名的权重最高的权威服务器被劫持时, 该域名被劫持的概率超过 50%.

域外依赖: 文献 [88] 认为, 域外 (out-of-bailiwick) 权威服务器存在潜在的不可控因素, 使用域外权威服务器提升解析鲁棒性的同时会降低解析安全性. 该文献于 2012 年分析 ODP/SC08 数据集[85]发现, 60% 域名的域外权威服务器所在区的数量占全部权威服务器所在区的一半. 文献 [89] 于 2018 年测量热度前 100 万域名发现, 99% 域名依赖域外权威服务器, 41% 的区严重依赖 2 个以上其他区.

综上, DNS 结构测量工作发现了依赖结构中蕴含的负面效应. 表 3 对结构测量相关工作进行了总结和比较, 例如解析器与桩解析器之间的依赖结构会带来负载隐藏问题和发起人识别问题, 给基于 CDN 的数据精准分发带来阻碍. 解析器之间的依赖结构将域名系统复杂化, 当解析器之间存在多级依赖时, 域名解析服务的运维难度随之增加. 域名之间的依赖关系则会使得域名出现解析的传递依赖风险, 造成解析可用性问题.

表 3 DNS 结构测量相关工作比较

节序号、主题 (测量对象)	文献	观测点类型	观测点数量	主被动	◆平台 ★工具 ● 数据	实验时间 (周期)	数据公开	关键结论
3.1 解析器簇结构	[76]	C, R	10^7	主	—	2011 (28天)	否	90%的解析器簇仅覆盖1%桩解析器, 最大簇包含12.9万桩解析器
3.2 解析器依赖结构	[28]	R, *C-R	—	混	—	2007	否	1 500万开放解析器中, 96.4%只转发解析器依赖58万递归解析器
3.3 域名解析依赖结构	[84]	C	—	主	—	2004	否	1/3域名的权威服务器中存在已知漏洞
	[69]	C	—	主	●ODP ●SC08	2009	是	使域名解析不可用所需的最小权威服务器平均值为2.34

注: C为客户端、R为解析器、A为权威服务器, *X-Y为X与Y信道中靠近*的一侧, ※表示论文发表时间

4 流量测量

流量测量通常以被动测量的方式获取原生状态下的 DNS 流量, 以展示 DNS 的当前运行状态. 本节将 DNS 流量测量工作分为 3 类: (1) 流量统计特征分析; (2) 异常根流量分析; (3) 流量拦截现象.

4.1 查询流量特征

重头长尾特征: 查询类型、域名热度、解析器查询频率等指标符合 Pareto/Zipf 分布规律 $P(x) = C/x^{\alpha}$, 具有重头、长尾的分布特征, 具体表现为:

(1) 文献 [30,33,34] 分别于 2001、2011、2013 年测量用户与解析器间流量发现, 占比最高的 3 种查询类型占比 90%–99%, 远超其他尾部类型, 它们分别在 2001 年是 A、PTR、MX 类型[30], 2011 年是 A、AAAA、MX 类型[33], 201 年和 2013 年是 A、AAAA、PTR 类型[34].

(2) 文献 [30] 于 2001 年测量 30.2 万域名发现, 单个域名的被查询的次数符合 Zipf 分布 (α=0.91), 前 10% 的域名占查询总数的 68%, 后 46% 的域名每个仅有 1 次查询.

(3) 文献 [90] 于 2004 年测量解析器与 8500 个 Akamai 权威服务器间流量发现, 95% 解析器发出的查询之和仅占服务端查询总量的 10%; 所有负责解析网站的权威服务器中, 80% 权威服务器收到的查询总量仅占权威服务器总查询量的 5%.

(4) 文献 [32] 于 2012 年测量 600 个解析器与权威服务器间流量发现, 每秒查询量前 10 的头部解析器的查询总量是后 400 的尾部解析器查询总量的 10 万倍; 文献 [91,92] 分别于 2006、2007 年通过测量根流量发现, 查询数量前 1% 的解析器发出查询总量的 80%.

4.2 根异常查询流量

异常根流量主要包括 3 类: (1) 无效查询, 根给出否定应答的查询; (2) 冗余查询, 过度重复的查询, 对根造成负担; (3) 幽灵查询, 在根改变 IP 地址后, 旧 IP 地址仍意外收到的查询. 一个查询可同时符合上述 3 类特征. 结合文献 [93,91] 于 2002 和 2006–2008 年的测量结论发现, 根流量中绝大部分是异常查询, 而正常有效的查询占比分别只有 2.15% (2002)、2.1% (2006)、2.3% (2007)、1.9% (2008).

4.2.1 无效流量

无效查询分类: 无效查询造成根否定应答的原因包括以下 3 类, 见表 4: (1) 查询类型未定义, 查询未定义的网络类型或资源记录类型; (2) 查询内容无效, 包括查询 IP 地址的 A 类型资源记录, 查询域名中包含无效字符, 查询的顶级域尚未由 IANA 授权[94], 查询私有 IP 地址的域名; (3) 查询的 TCP/IP 头部字段错误, 例如使用私有 IP 地址做源 IP 地址, 使用端口号 0 做源端口号.

无效查询占比: 文献 [95,93] 分别于 2001、2002 年测量 F 根发现, 无效查询占总查询量的 25% 和 23%, 占比最大的 3 种无效查询类型是: (1) 查询未授权的顶级域; (2) 查询 IP 地址的 IP 地址 (A-for-A 查询); (3) 私有 IP 地址发起的查询或查询私有 IP 地址的主机名, 分别占总查询量的 12.5%、7.0%、1.6%. 文献 [91] 分析 2008 年的 8 个

根的日测数据 DITL[96]发现, 无效查询占比在 20%–25% 之间, 其中 F 根上无效查询占比与文献 [95,93] 于 2001 和 2002 年的测量结果相似, 最多的 3 种无效查询占比分别为 25.5%、3.6%、0.2%. 文献 [97] 于 2004 年测量根服务器发现 A-for-A 类型的无效查询占根查询流量的 7%, 文献 [32] 于 2012 年测量 600 个解析器发向根的查询发现, A-for-A 类型的无效查询已减少至无效流量的 0.4%, 未注册顶级域查询占比增加至无效流量的 53%.

表 4 根上无效查询类型统计 (%)

类别	无效查询类型(描述)	[93](2002)	[95](2001)	[97](2004)	[91](2008)
查询类型未定义	Unused query class (未定义查询网络类型)	—	0.024	—	0.1
	Unused query type (未定义查询记录类型)	—	2.15	—	—
查询内容无效	A-for-A (查询IP地址的IP地址)	12.0–18.7	7.03	~7	2.7
	Invalid char (查询中存在无效字符)	—	1.94	—	0.1
	Invalid TLD (查询未授权顶级域)	19.6–29.7	12.5	15–20	22
	RFC 1918 related A?(查询私有IP地址的域名)	0.03–3.6	1.61	1–3	0.4
TCP/IP头部字段错误	RFC 1918 related(查询报文使用私有IP地址做源地址)	2.0–11.5	1.61	1–3	—
	Source Port Zero (查询报文端口号为0)	<0.1	—	—	—

无效查询成因: 综合相关文献推测无效查询成因如下: (1) 文献 [95] 于 2001 年测量发现, F 根上 DNS 动态更新请求激增, 推测与当时新发布的 Windows 2000 操作系统有关. 文献 [98] 于 2002 年对发向 F 根 (AS112 节点) 和 K 根 (伦敦节点) 上的来自私有 IP 地址的动态更新请求分析更新频率、更新时段和源端口的规律发现, 一个主要原因是 Windows 2000 和 Windows XP 操作系统默认开启对 in-addr.arpa 域的更新请求, 当本地网络配置不当时这些查询会泄露到公网的根上. 为缓解私有 IP 地址查询给根的压力, ICANN 于 2002 年启动 AS112 项目[99]设置黑洞服务器群吸收发向根的对私有 IP 地址的查询. (2) 文献 [93] 于 2002 年测量发现某数据源持续查询非 IPv4 地址的主机名, 占当日 in-addr.arpa 区查询总量的 99.96%, 推测由解析器软件缺陷导致. (3) 文献 [100] 于 2013 至 2014 年测量 A 根和 J 根发现, 每天有 5.05 万个对 i2p、onion 等专用于隐匿网络的伪顶级域的查询, 约占每日查询总量的 0.1%, 推测由浏览器代理配置错误或恶意软件导致.

4.2.2 冗余流量

冗余查询是指 name、type、class 字段相同的查询. 根据 TXID 字段是否也相同可进一步分为两类. 文献 [95,93,91] 于 2001、2002、2008 测量 F 根发现, TXID 相同的冗余查询占比为 25.4% (2002)、15.6% (2008). TXID 不同的冗余查询占比为 62%~85% (2001)、44.9% (2002)、44.9% (2008). 结合相关测量文献, 分析造成冗余查询的原因如下.

软件缺陷: 文献 [101] 于 1991 年测量 A 根 98.5 万查询并对其去重后发现最小必要查询数为 1.57 万. 分析发现这是由解析器软件 (BIND 4.8.3) 缺陷造成的: 解析器无法检测循环依赖, 解析器无法检测失效的权威服务器, 解析器未缓存否定应答. 文献 [95] 于 2001 年测量 F 根发现, 1 月 24 日二级域 microsoft.com 的权威服务器出现故障, windows 解析器默认不开启否定应答缓存, 导致对该域名的查询占根当日查询量的 9.1% (该域名常规查询量占 0.003%). 文献 [32] 于 2012 年测量发现解析器过度查询现象普遍存在, 例如 Unbound 解析器的超时等待时间过短, 仅等待一个 RTT 周期即重复查询; GoogleDNS、OpenDNS 等公共解析器对短时间内相同递归查询请求所触发的迭代查询未做限制.

网络配置不当: 文献 [93] 于 2002 年测量发现某网络在 24 小时内对 F 根发送了 2300 万次查询, 占当天 F 根查询总量的 15.3%. 通过联系该网络的管理员排查原因发现, 该网络中的包过滤器阻止来自网络外的 DNS 应答, 网络内的解析器由于查询失败而重复查询.

恶意扫描攻击: 文献 [95] 于 2001 年测量 F 根发现了 2 个针对根的攻击源. 一个攻击源利用根实施反射攻击, 攻击者通过伪造源 IP 地址向根发送大量冗余查询, 迫使根频繁对 209.67.50/24 网段内的主机做出应答. 另一个攻击源向根连续发送大量不重复的反向域名解析查询, 推测攻击者正在扫描 IP 地址空间, 尝试发现部署了对外服务

的 IP 地址.

4.2.3 幽灵流量

幽灵查询: 根服务器切换 IP 地址后, 旧 IP 地址仍意外收到的查询. 出于部署任播根节点的需求, D 根于 2013 年 1 月 3 日切换 IP 地址, 并为保证服务平稳过渡, 在 4 个月过渡期内保证新旧 IP 地址都提供服务. 文献 [102] 测量发现, 旧根会收到大量幽灵查询, 新根与旧根的总查询量增至原来的 1.5 倍并持续了 3 个月; 在 IP 地址切换 4 个月后仍有 63% 的用户持续查询旧根. 文献 [102] 分析幽灵查询特征发现: (1) 幽灵查询的域名和权威服务器局限在小范围内; (2) 旧根的有效查询率明显高于新根, 很少出现因拼写错误导致的无效查询. 这些现象表明幽灵查询可能来自直接通过 IP 地址查询的扫描脚本或攻击程序. 幽灵查询同样出现在 2002 年 J 根切换 IP 地址[103], 以及在 2004 年 B 根切换 IP 地址时[104].

4.3 流量拦截

DNS 流量拦截行为普遍存在, 相关测量研究发现许多地区的网络中存在该现象[105–108]. 文献 [109] 使用 8 000 个 Atlas 节点[57]测量 2 275 个自治域内的域名解析结果, 并结合时延信息、服务器信息和路由信息, 发现 10 个伪装成根服务器的 DNS 代理和 1 个的假根服务器. 文献 [9,110] 发现运营商拦截并篡改权威服务器应答, 运营商会将无效解析篡改为广告网站的 IP 地址实现获利. 文献 [111] 利用全球 148 478 个观测点测量发现, 8.5% (259/3 047) 的自治域存在拦截并篡改权威服务器应答流量的行为.

综上, 流量测量工作通常以被动测量的方式发现 DNS 实际运行中的规律及问题. 表 5 对流量测量相关工作进行了总结和比较. DNS 查询呈现 Pareto/Zipf 分布规律, 意味着具有重头长尾的分布特征, 这一规律有助于对 DNS 相关的服务进行优化. 根服务器上比例高达 98% 的异常流量暴露了当前 DNS 系统中若干亟待解决的问题, 例如无效查询问题[91,93,95,97]和冗余查询问题[32,91,101]. 广泛存在的 DNS 流量拦截行为在实现特定网络服务目标的同时, 也破坏了互联网的端到端原则.

表 5 DNS 流量测量相关工作比较

节序号、主题 (测量对象)	文献	观测点类型	观测点数量	主被动	◆平台 ★工具 ● 数据	实验时间 (周期)	数据公开	关键结论
4.2.1 根无效流量	[32]	*R-A	10^2	被	●SIE	2012 (14天)	是	根服务器的99%无效查询时无效顶级域查询
	[91]	A	10^2	被	●DITL	2006–2008	是	根服务器收到的98%查询都属于无效查询
	[98]	A	—	被	—	2002 (34天)	否	绝大多数私有IP地址动态更新源来自MS OS 2000和XP
	[100]	A	—	被	—	2013–2014 (127天)	否	A根和J根上对i2p顶级域查询占4.6%.
4.2.2 根冗余流量	[95]	A	—	被	★tcpdump	2001 (10天)	否	F根上冗余流量最高达85%
	[93]	A	—	被	★tcpdump	2002 (24 h)	否	网络中间盒丢弃来自根的应答, 导致对根的重复查询
	[101]	R-*A	—	被	—	1991 (2月)	否	根的查询量是必要查询量的61倍
4.2.3 根幽灵流量	[102]	A	—	被	★tcpdump	2013	否	F根切换IP地址后, 原根服务器查询量增加
4.3 流量拦截	[109]	C	10^3	主	◆RIPE Atlas	2014	否	结合Hostname.bind, Traceroute, BGP信息发现10个拦截根查询的伪装根

注: C为客户端、R为解析器、A为权威服务器, *X-Y为X与Y信道中靠近*的一侧, ※表示论文发表时间

5 安全测量

本节从测量的角度综述 DNS 安全协议的部署可行性、展示 DNS 安全协议的部署应用状况、阐述识别恶意域名的检测技术. 本节将安全测量工作分 4 类: (1) 分析 DNSSEC 的代价与隐患; (2) 展示 DNSSEC 的部署现状; (3) 分析加密 DNS 的应用现状; (4) 阐述测量识别恶意域名的方法.

5.1 DNSSEC 代价与隐患

DNSSEC[14]采用依托于 DNS 的层级结构认证和数字签名等密码学技术保护权威服务器应答的真实性和完整性, 但同时增加了 DNS 的复杂性和运行成本. DNSSEC 中的数字签名会增加 DNS 服务器计算开销和应答报文长度. 本节介绍 DNSSEC 额外时延、计算和带宽的测量工作, 介绍利用 DNSSEC 实施的放大攻击的可行性测量评估工作.

性能代价: 文献 [112] 从解析时延和 CPU 负载两方面对比不同 DNS 软件启用 DNSSEC 时产生的额外开销, 测量发现启用 DNSSEC 使用户解析时延将增加 35%; BIND 权威服务器对每个查询的平均处理时间增加 25%; NSD 权威服务器平均增加 8%; BIND 解析器对每个查询的平均处理时间增加 116%; Unbound 解析器平均增加 253%. 文献 [113] 在仿真测量环境下, 通过重放根的真实查询流量并强制所有查询开启 DNSSEC 选项对权威服务器性能的影响, 发现只给根服务器增加 5% 的内存负载、4%–5% 的 CPU 负载、10% 的带宽负载. 文献 [114] 通过 DNS 仿真测量框架 LDplayer 测量解析器强制启用 DNSSEC 功能后对当前根的应答流量的影响, 将占比 72% 的 DNSSEC 查询提升至 100% 时, 根应答流量将会从 225 Mb/s 提升至 296 Mb/s, 增幅 31%; 若根的 ZSK (区签名密钥) 长度从 1024 位提升到 2048 位, 根的应答流量将增加 32%.

放大攻击隐患: 攻击者利用 DNSSEC 长应答报文实施分布式反射拒绝服务攻击. 文献 [115] 测量 DNSSEC 对应答报文长度的影响发现, 应用最广泛 RSA 数字签名算法在最坏情况下能令应答报文增长至 12.7 倍, ECC 数字签名算法增长至 6.7 倍. 文献 [102] 测量 D 根服务器时发现周期性流量尖峰现象, 推测 D 根服务器正在被用于实施反射拒绝服务攻击. 文献 [116] 测量 1 404 个部署 DNSSEC 的权威服务器发现, 平均放大倍数为 54.6, 最高的前 10% 的放大倍数达到 98.3. 文献 [117] 测量 6 个顶级域下 (com、net、org、nl、se、uk) 约 250 万个部署 DNSSEC 的二级域发现, 直接利用权威服务器实施放大攻击的效果比使用开放解析器更好, 可实现 50 至 179 的放大倍数. 从实施攻击的持续性和难易度来看, 权威服务器通常比解析器具有更高的带宽, 且攻击者具有更广泛的二级域选择范围.

5.2 DNSSEC 部署进展

DNSSEC 的部署依赖于各级权威服务器、解析器以及网络中间盒 (middle-box, 如防火墙或 NAT 设备) 制造商和运营商之间协作. 本节首先介绍了 DNSSEC 在多个方面的部署进展, 而后介绍了 DNSSEC 渐进部署带来的信任孤岛问题和 DNSSEC 误配置带来的问题.

权威服务器部署进展: DNSSEC 于 2005 年 10 月开始在国家顶级域 se 部署, 2010 年 6 月开始在通用顶级域 org 部署, 同年 7 月在根上部署[118]. 文献 [68] 于 2007 年测量 1.28 亿二级域 (占当时二级域注册总数的 58%) 发现仅有 161 个二级域的区文件中包含 DNSKEY 资源记录, 部署比例为 0.003%. 文献 [119] 于 2010 年测量发现顶级域 com、net、org 下二级域部署的比例为 0.022%. 文献 [120] 于 2015 年 2 月至 2016 年 11 月测量发现顶级域 org 下二级域部署比例最高, 为 1%; 顶级域 com 其次, 为 0.7%. 文献 [121] 测量发现, 截至 2016 年底已有 89% 的通用顶级域和 47% 的国家顶级域部署. 文献 [122] 于 2016 年 3 月至 9 月测量发现, 热度前 100 万域名所在的 90% 的顶级域和 1.66% 的二级域已经部署. 文献 [123] 于 2017 年测量发现, 应用 DNSSEC 的二级域已达到 640 万个, 部署二级域数量最多的 3 个顶级域分别是 com、nl、se. 测量还发现 ECDSA 这种在保证加密强度但能使密钥更短的数字签名算法更加流行, 近两年使用率增加 8%. ICANN 于 2020 年 4 月的统计结果[124]显示 1 513 个顶级域中的 90.9% 已经部署.

解析器部署进展: 文献 [125] 于 2010–2011 年测量 org 权威服务器发现, 约 8%–10% 的解析器部署 DNSSEC. 文

献 [126] 于 2011 年测量 35 000 个解析器发现, 其中 97% 发起 DNSSEC 查询但不验证数字签名. 文献 [120] 于 2015–2016 年通过大规模投放广告并广告网页中嵌入 JS 程序实施测量, 发现 4 472 个解析器中有 82% 查询 DNSSEC 相关资源记录但不验证数字签名. 文献 [127] 于 2012 年提出 Check-Repeat 方法判断解析器是否 DNSSEC 验证, 在顶级域 com、net 权威服务器上使用该方法测得, 12% 的查询由支持 DNSSEC 验证的解析器产生.

中间盒部署进展: DNSSEC 应答报文可能超过以太网最大传输单元, 此时若中间盒不能正确传输 IP 分片将导致解析失败. 文献 [110] 于 2011 年通过大规模投放广告并在广告网页中嵌入 JS 测量程序测量用户与解析器间的链路发现, 由于中间盒无法发送 IP 分片, 9% 的 DNSSEC 应答因缺失分片无法被解析器验证.

信任孤岛: DNSSEC 依赖自根向下建立的认证链提供安全防护, 因父区未部署 DNSSEC 导致已部署 DNSSEC 但无法认证的子区称为 DNSSEC 信任孤岛. 文献 [128] 于 2008 年测量发现有 97.4% 的区为孤岛. 文献 [129] 于 2011 年测量发现有 76.6% 的区为孤岛, 其中 97.5% 的孤岛的规模为 1, 即该区的父区和子区均未部署. 文献 [122] 于 2016 年 3 月至 9 月测量热度前 100 万域名发现, 90% 的顶级域和 1.66% 的二级域已经部署 DNSSEC, 其中 19.46% 的二级域为孤岛. 文献 [130] 于 2015 至 2016 年测量发现, 注册服务商 (registrar) 是阻碍 DNSSEC 部署规模增长的重要因素, 排名前 26 的注册服务商中只有 2 个完全支持 DS 等 DNSSEC 资源记录的上传, 这两个注册服务商覆盖了顶级域 com、net、org、nl、se 下超过 50% 的部署 DNSSEC 二级域.

DNSSEC 误配置: 文献 [131] 于 2013 年测量顶级域 bg、br、co、se 下的二级域, 并结合公开数据集 DNS-Census-2013[132]发现, 存在 DNSSEC 误配置的二级域占比 4%~26%, 误配置将导致这些二级域无法被解析器验证.

5.3 加密 DNS 部署

文献 [133] 于 2017–2019 年通过部署在 116 个国家和地区的 12 万测量点测量 DoT (DNS-over-TLS)[15]加密解析器行为, 以及通过被动测得的流量分析 DoH (DNS-over-HTTPS)[16]加密解析器行为. 测量发现: (1) 2018 年 7 月至 2018 年 12 月, 加密 DNS 流量增长了 56%; (2) 在 1 500 个提供 DoT 解析的 IP 地址中, 大型公共解析器服务商, 如 Cloudflare、CleanBrowsing 等覆盖了 75% 的 IP 地址; (3) Cloudflare 的 DoT 解析器解析失败比例约为 1%, 远低于其普通解析器的平均值 16.46%.

5.4 恶意域名检测

恶意域名是指攻击发起者和攻击执行者间用于建立隐秘信道的域名. 文献 [17] 从 DNS 安全防护的角度着重综述了检测技术原理, 本节侧重于对检测实验结果的分析.

基于域名注册行为检测: 通过发现域名注册行为的异常, 例如使用可疑注册服务商、高频率批量注册等行为识别恶意域名. 文献 [134] 于 2012 年分析恶意域名在 com 域的注册规律发现, 恶意域名倾向于使用少量固定的注册服务商, 仅占新注册域名数量 20% 的 10 个注册服务商产生了多达 70% 的新垃圾邮件域名. 域名注册行为测量相关研究成果对此具有潜在应用价值. 例如, 文献 [135] 于 2015–2016 年测量 5 个顶级域下 740 万域名注册行为发现, 规模大的顶级域 (com、net、org) 下二级域名的再注册比例要明显高于规模小的顶级域 (biz、name); 高龄二级域的再注册比率明显高于低龄二级域.

基于域名查询量变化检测: 恶意域名的注册时间与首次使用时间间隔比正常域名的间隔更短. 文献 [136] 于 2011 年测量恶意域名查询量变化发现, 超过 55% 的恶意域名注册后的 1 天内会被用于攻击, 3–4 天内查询量出现爆发增长. 文献 [137] 分析 SIE 数据集[138]证实, 超过 73% 的恶意域名会在注册当天投入使用, 95% 在注册 2 天内使用, 而正常域名仅有 33% 会在当天投入使用.

基于权威服务器特征检测: 权威服务器的 IP 地址特征可以作为检测恶意域名的依据. 文献 [139] 于 2007 年测量 3 360 个实施诈骗活动的恶意域名发现, 恶意域名的 IP 地址及其权威服务器的 IP 地址的轮换速度远高于正常域名, 恶意域名使用的 IP 地址分布范围更广, 40% 的恶意域名中每个域名使用的 IP 地址覆盖 300 多个/24 网络. 文献 [140] 测量发现恶意域名的权威服务器普遍存在 2 个特征: (1) 权威服务器所在域通常注册时间小于 1 年; (2) 恶意域名通常会选择自己域下的权威服务器进行解析 (又称“自解析”), 而非选择域外的权威服务器. 文献 [136] 于 2011 年

测量发现, 与正常域名相比恶意域名的权威服务器集中在少量 IP 地址段与自治域内, 其中一个自治域包含了 30% 的恶意域名的至少一个 IP 地址.

计算域名信誉值检测: 文献 [11] 综合使用上述多个维度的特征计算域名的信誉值, 并基于信誉值发现恶意域名, 用计算信誉值的特征包括: 域名字面特征、IP 地址特征、TTL 值特征和查询量变化特征. 该工作从 4 800 万域名中筛选 30 万疑似恶意域名, 最终从中确定了 1.76 万恶意域名. 文献 [12] 相较文献 [11] 额外引入了域名黑名单作为修正信誉计算模型的依据, 该工作于 2009 年测量 2 700 万域名, 检测准确率为 96.8%. 文献 [141] 基于域名查询者的空间分布和用户特征, 再结合域名黑名单作为计算信誉值的依据, 该工作于 2010 年测量 1200 万域名, 并从中检测出 2.7 万恶意域名, 检测准确率为 98.4%. 文献 [142] 基于域名依赖关系和域名查询者的空间分布作为计算域名信誉值的依据, 该工作测量主干网中解析器与权威服务器的通信流量, 检测出 5 533 个恶意二级域.

综上, 针对 DNS 安全的测量研究既可用于评估安全协议实施的效果, 也可作为保障 DNS 安全的技术手段. 测量 DNSSEC 发现存在部署进展迟缓的问题, 部署迟缓的原因主要有两个, 其一是部署带来运维上的复杂性和对硬件性能的更高要求; 另一个原因是部署会带来新的安全问题. 测量 DoT、DoH 等加密 DNS 技术发现, 支持加密 DNS 的解析器数量高速增长, 且部署加密 DNS 的解析器通常得到更好的维护, 解析失败率远低于平均值. 测量技术也是保护 DNS 安全的重要手段, 通过测量域名的注册行为、查询量变化、生命周期等特征, 检测实施恶意行为的域名, 为后续的安全应对措施提供支撑. 表 6 对安全测量工作进行了总结和比较.

表 6 DNS 安全测量研究工作比较

节序号、主题 (测量对象)	文献	观测点类型	观测点数量	主被动	◆平台 ★工具 ● 数据	实验时间周期	数据公开	关键结论
5.1 DNSSEC代价与隐患	[113]	A	—	被	★tcpreply	2005	否	K根部署DNSSEC后, CPU/内存/带宽的负载增长5%/5%/10%
	[117]	C	—	主	—	2012–2014	是	DNSSEC用于反射攻击时, 流量放大倍率为50–179
5.2 DNSSEC部署进展	[126]	C	10^5	主	—	※2013 (1周)	否	2.6%解析器完全部署DNSSEC, 97%解析器不完全部署DNSSEC
	[120]	C	—	注	◆OpenIntel	2015–2016 (21周)	是	31%的二级域和12%解析器部署了DNSSEC
	[123]	C	—	主	—	2017 (10天)	否	超过640万二级域部署DNSSEC
	[130]	C	—	主	◆OpenIntel	2015–2016 (17月)	是	前20域名注册服务运营商中支持DNSSEC有关资源记录上传的只有3个
	[122]	C	—	主	—	2016 (7月)	否	90%顶级域、1.66%二级域部署DNSSEC
5.3 加密DNS测量	[133]	CC-*R	10^5	混	★Zmap◆RIPE Atlas●NetFlow ●DNSDB●360 PDNS	2017–2019	是	测量全网发现1500个提供DoT的IP地址, 大型公共解析器服务商, 如Cloudflare、CleanBrowsing等覆盖了75%
5.4 恶意域名检测	[12]	*R-A	—	被	★Notos	2009 (68天)	否	综合BGP、AS、WHOIS等信息判断恶意域名, 准确率96.8%
	[136]	C, A	10^2	混	◆PlanetLab	2011 (30天)	否	55%恶意域名在注册1天后被用于恶意活动
	[134]	A	—	被	—	2012 (8月)	否	50%恶意域名的生命周期小于3个月

注: C为客户端、R为解析器、A为权威服务器, *X-Y为X与Y信道中靠近*的一侧, ※表示论文发表时间

6 未来研究趋势

DNS 测量研究未来仍将会是网络测量领域内的重要研究方向, 其原因在于 DNS 作为关键互联网基础设施是

不可替代的, DNS 技术和生态将持续演进, 并将持续支撑未来的网络应用服务. 下面对未来 DNS 测量领域内值得进一步研究的工作进行探讨.

(1) DNS 测量评估标准化

DNS 测量领域一直存在一个具有挑战性的难题, 即如何根据不同场景提出标准化的测量方法和评估方法. 解决这个问题有利于相关研究成果的比较和分享. 目前 ICANN 等组织尝试建立 DNS 测量方法与评估标准, 工作 [143] 提出评估 DNS 健康状况的 5 个通用评价指标; 文献 [144] 提出“测量名字系统”项目 (MeNSa), 项目提出一套形式化和结构化的测量方法以及一套用于评估 DNS 健康和安全级别的度量标准. 目前, DNS 测量评估标准化仍不成熟, 亟待进一步研究.

(2) IDN (internationalized domain name) 应用测量

IDN 将域名表达形式从 ASCII 字符扩展至更多语种, 例如阿拉伯文字符、中文字符甚至是 emoji 表情字符. IDN 相关标准在 2002 年就已经制定, 但 IDN 域名的注册近几年才开始成为趋势. 对于象形文字, IDN 域名的应用过程中存在同型异义词攻击, 即利用字符的视觉相似性实施欺骗攻击, 文献 [145] 等研究测量实施这类攻击的域名. 未来 IDN 域名的相关风险需要持续关注.

(3) DNS 服务整合 (consolidation) 测量

云服务的兴起使 DNS 服务逐渐向云端迁移. 一方面, 大量域名的托管使得域名解析集中到少数云服务商所控制的服务器上. 文献 [63] 的测量研究表明有 91%–93% 的二级域存在共享权威服务器的情况. 另一方面, 公共解析器高速普及使用户的解析查询流量呈现集中化的发展趋势, 文献 [21] 测量发现截止至 2011 年 12 月, 公共解析器服务已经覆盖了 8.6% 的互联网用户, 21 个月内增长 27%. 上述服务整合的发展趋势违背了 DNS 通过分布式管理提高系统可靠性的初衷, 某个云服务商出现故障会导致大量域名无法解析[64]. 目前, DNS 服务整合的趋势似乎已成定局, 如何测量并评估这种趋势的发展给 DNS 带来的潜在风险是值得研究的方向.

7 结束语

DNS 作为互联网中最复杂生态系统之一, 其经过 30 多年的发展后已经庞大到几乎无法被全面地观察, 每个研究只能以管中窥豹的方式洞察某一细分主题下的“信息碎片”, 全面掌握 DNS 的现状离不开一幅完整“地图”, 本文通过对 DNS 测量工作进行综述尝试勾勒和描绘这幅地图.

References:

[1] Root server lists. https://root-servers.org

[2] Verisign domain name industry brief Q1 2020. https://www.verisign.com/assets/domain-name-report-Q12020.pdf

[3] DNS related RFCs. https://www.statdns.com/rfc

[4] Brownlee N, Claffy K, Nemeth E. DNS Root/gTLD performance measurements. In: Proc. of the Passive and Active Measurement Workshop. USENIX, 2001. 241–256.

[5] Liston R, Srinivasan S, Zegura E. Diversity in DNS performance measures. In: Proc. of the 2nd ACM SIGCOMM Workshop on Internet Measurment. Marseille: ACM, 2002. 19–31. [doi: 10.1145/637201.637204]

[6] Liang JJ, Jiang J, Duan HX, Li K, Wu JP. Measuring query latency of top level DNS servers. In: Proc. of the 14th Int'l Conf. on Passive and Active Network Measurement. Hong Kong: Springer, 2013. 145–154. [doi: 10.1007/978-3-642-36516-4_15]

[7] Lee T, Huffaker B, Fomenkov M. On the problem of optimization of DNS root servers' placement. In: Proc. of the Int'l Workshop on Passive and Active Network Measurement. Springer, 2003.

[8] Müller M, Moura GCM, de O Schmidt R, Heidemann J. Recursives in the wild: Engineering authoritative DNS servers. In: Proc. of the 2017 Internet Measurement Conf. London: ACM, 2017. 489–495. [doi: 10.1145/3131365.3131366]

[9] Kührer M, Hupperich T, Bushart J, Rossow C, Holz T. Going wild: Large-scale classification of open DNS resolvers. In: Proc. of the Internet Measurement Conf. Tokyo: ACM, 2015. 355–368. [doi: 10.1145/2815675.2815683]

[10] US-CERT. Vulnerability note: Multiple DNS implementations vulnerable to cache poisoning. 2008. http://www.kb.cert.org/vuls/id/800113

[11] Bilge L, Kirda E, Kruegel C, Balduzzi M. EXPOSURE: Finding malicious domains using passive DNS analysis. In: Proc. of the 2011

Network and Distributed System Security Symp. 2011. 1–17.

[12] Antonakakis M, Perdisci R, Dagon D, Lee W, Feamster N. Building a dynamic reputation system for DNS. In: Proc. of the 19th USENIX Security Symp. Washington: USENIX, 2010. 273–290. [doi: 10.5555/1929820.1929844]

[13] RSSAC023: History of the root server system. 2016. https://www.icann.org/en/system/files/files/rssac-023-04nov16-en.pdf

[14] Arends R, Austein R, Larson M, Massey D, Rose S. DNS security introduction and requirements. RFC4033, 2005. [doi: 10.17487/RFC4033]

[15] Hu Z, Zhu L, Heidemann J, Mankin A, Wessels D, Hoffman P. Specification for DNS over transport layer security (TLS). RFC7858, 2016. [doi: 10.17487/RFC7858]

[16] Hoffman PE, McManus P. Dns queries over https (DoH). RFC8484, 2018. [doi: 10.17487/RFC8484]

[17] Wang WT, Hu N, Liu B, Liu X, Li SD. Survey on technology of security enhancement for DNS. Ruan Jian Xue Bao/Journal of Software, 2020, 31(7): 2205–2220 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6046.htm [doi: 10.13328/j.cnki.jos.006046]

[18] Ramdas A, Muthukrishnan R. A survey on DNS security issues and mitigation techniques. In: Proc. of the 2019 Int'l Conf. on Intelligent Computing and Control Systems (ICCS). Madurai: IEEE, 2019. 781–784. [doi: 10.1109/ICCS45141.2019.9065354]

[19] Wang Y, Hu MZ, Li B, Yin BR. Survey on domain name system security. Journal on Communications, 2007, 28(9): 91–103 (in Chinese with English abstract). [doi: 10.3321/j.issn:1000-436x.2007.09.015]

[20] Hu N, Deng WP, Su Y. Issues and challenges of internet DNS security. Chinese Journal of Network and Information Security, 2017, 3(3): 13–21 (in Chinese with English abstract). [doi: 10.11959/j.issn.2096-109x.2017.00154]

[21] Otto JS, Sánchez MA, Rula JP, Bustamante FE. Content delivery and the natural evolution of DNS: Remote DNS trends, performance issues and alternative solutions. In: Proc. of the 2012 Internet Measurement Conf. Boston: ACM, 2012. 523–536. [doi: 10.1145/2398776.2398831]

[22] AquaLab. EdgeScope—sharing the view from a distributed Internet telescope. https://aqualab.cs.northwestern.edu/projects/edgescope-sharing-the-view-from-a-distributed-internet-telescope

[23] Huang C, Maltz DA, Li J, Greenberg A. Public DNS system and global traffic management. In: Proc. of 2011 INFOCOM 2011. Shanghai: IEEE, 2011. 2615–2623. [doi: 10.1109/INFCOM.2011.5935088]

[24] Khosla R, Fahmy S, Hu YC. Content retrieval using cloud-based DNS. In: Proc. of the 2012 IEEE INFOCOM Workshops. Orlando: IEEE, 2012. 1–6. [doi: 10.1109/INFCOMW.2012.6193491]

[25] Ager B, Mühlbauer W, Smaragdakis G, Uhlig S. Comparing DNS resolvers in the wild. In: Proc. of the 10th ACM SIGCOMM Conf. on Internet Measurement. Melbourne: ACM, 2010. 15–21. [doi: 10.1145/1879141.1879144]

[26] Kührer M, Hupperich T, Rossow C, Holz T. Exit from hell? Reducing the impact of amplification DDoS attacks. In: Proc. of the 23rd USENIX Security Symp. San Diego: ACM, 2014. 111–125.

[27] Open resolver project. https://archive.nanog.org/sites/default/files/tue.lightning3.open_resolver.mauch_.pdf

[28] Dagon D, Provos N, Lee CP, Lee W. Corrupted DNS resolution paths: The rise of a malicious resolution authority. In: Proc. of 2008 Network and Distributed System Security Symp. 2008.

[29] Klein A, Shulman H, Waidner M. Internet-wide study of DNS cache injections. In: Proc. of the 2017 IEEE Conf. on Computer Communications. Atlanta: IEEE, 2017. 1–9. [doi: 10.1109/INFOCOM.2017.8057202]

[30] Jung J, Sit E, Balakrishnan H, Morris R. DNS performance and the effectiveness of caching. IEEE/ACM Trans. on Networking, 2002, 10(5): 589–603. [doi: 10.1109/TNET.2002.803905]

[31] Bhatti SN, Atkinson R. Reducing DNS caching. In: Proc. of the 2011 IEEE Conf. on Computer Communications Workshops. Shanghai: IEEE, 2011. 792–797. [doi: 10.1109/INFCOMW.2011.5928919]

[32] Gao HY, Yegneswaran V, Chen Y, Porras P, Ghosh S, Jiang J, Duan HX. An empirical reexamination of global DNS behavior. ACM SIGCOMM Computer Communication Review, 2013, 43(4): 267–278. [doi: 10.1145/2534169.2486018]

[33] Fujiwara K, Sato A, Yoshida K. DNS traffic analysis: Issues of IPv6 and CDN. In: Proc. of the 12th IEEE/IPSJ Int'l Symp. on Applications and the Internet. Izmir: IEEE, 2012. 129–137. [doi: 10.1109/SAINT.2012.26]

[34] Callahan T, Allman M, Rabinovich M. On modern DNS behavior and properties. ACM SIGCOMM Computer Communication Review, 2013, 43(3): 7–15. [doi: 10.1145/2500098.2500100]

[35] Almeida M, Finamore A, Perino D, Vallina-Rodriguez N, Varvello M. Dissecting dns stakeholders in mobile networks. In: Proc. of the 13th Int'l Conf. on Emerging Networking Experiments and Technologies. Incheon: ACM, 2017. 28–34. [doi: 10.1145/3143361.3143375]

[36] Shafir L, Afek Y, Bremler-Barr A, Peleg N, Sabag M. DNS negative caching in the wild. In: Proc. of the 2019 ACM SIGCOMM Conf. on Posters and Demos. Beijing: ACM, 2019. 143–145. [doi: 10.1145/3342280.3342338]

[37] Schomp K, Callahan T, Rabinovich M, Allman M. On measuring the client-side DNS infrastructure. In: Proc. of the 2013 Internet Measurement Conf. Barcelona: ACM, 2013. 77–90. [doi: 10.1145/2504730.2504734]

[38] Shue CA, Kalafut AJ, Allman M, Taylor CR. On building inexpensive network capabilities. ACM SIGCOMM Computer Communication Review, 2012, 42(2): 72–79. [doi: 10.1145/2185376.2185386]

[39] Pang J, Akella A, Shaikh A, Krishnamurthy B, Seshan S. On the responsiveness of DNS-based network control. In: Proc. of the 4th ACM SIGCOMM Conf. on Internet Measurement. Taormina: ACM, 2004. 21–26. [doi: 10.1145/1028788.1028792]

[40] Wills CE, Mikhailov M, Shang H. Inferring relative popularity of internet applications by actively querying DNS caches. In: Proc. of the 3rd ACM SIGCOMM Conf. on Internet Measurement. Miami Beach: ACM, 2003. 78–90. [doi: 10.1145/948205.948216]

[41] Rajab MA, Monrose F, Terzis A, Provos N. Peeking through the cloud: DNS-based estimation and its applications. In: Proc. of the 6th Int'l Conf. on Applied Cryptography and Network Security. New York: Springer, 2008. 21–38. [doi: 10.1007/978-3-540-68914-0_2]

[42] Ma XB, Zhang JJ, Li ZH, Li JF, Tao J, Guan XH, Lui JCS, Towsley D. Accurate DNS query characteristics estimation via active probing. Journal of Network and Computer Applications, 2015, 47: 72–84. [doi: 10.1016/j.jnca.2014.09.016]

[43] Gummadi KP, Saroiu S, Gribble SD. King: Estimating latency between arbitrary Internet end hosts. In: Proc. of the 2nd ACM SIGCOMM Workshop on Internet Measurment. Marseille: ACM, 2002. 5–18. [doi: 10.1145/637201.637203]

[44] Leonard D, Loguinov D. Turbo king: Framework for large-scale internet delay measurements. In: Proc. of the 27th Conf. on Computer Communications. Phoenix: IEEE, 2008. 31–35. [doi: 10.1109/INFOCOM.2008.15]

[45] Wang YA, Huang C, Li J, Ross KW. Queen: Estimating packet loss rate between arbitrary internet hosts. In: Proc. of the 10th Int'l Conf. on Passive and Active Network Measurement. Seoul: Springer, 2009. 57–66. [doi: 10.1007/978-3-642-00975-4_6]

[46] Elz R, Bush R, Bradner SO, Patton MA. Selection and operation of secondary DNS servers. RFC 2182, 1997. [doi: 10.17487/RFC2182]

[47] Wessels D, Fomenkov M, Brownlee N, Claffy K. Measurements and laboratory simulations of the upper DNS hierarchy. In: Proc. of the 5th Int'l Workshop on Passive and Active Network Measurement. Antibes Juan-les-Pins: Springer, 2004. 147–157. [doi: 10.1007/978-3-540-24668-8_15]

[48] Yu YD, Wessels D, Larson M, Zhang LX. Authority server selection in DNS caching resolvers. ACM SIGCOMM Computer Communication Review, 2012, 42(2): 80–86.

[49] Cho K, Kato A, Nakamura Y, Somegawa R, Sekiya Y, Jinmei T, Suzuki S, Murai J. A study on the performance of the root name servers. http://mawi.wide.ad.jp/mawi/dnsprobe

[50] Brownlee N. Root/gTLD DNS performance plots. http://www.caida.org/cgi-bin/dns_perf/main.pl

[51] Yuchi X, Wang X, Li XD, Yan BP. DNS measurements at the. CN TLD servers. In: Proc. of the 6th Int'l Conf. on Fuzzy Systems and Knowledge Discovery. IEEE, 2009. 540–545. [doi: 10.1109/FSKD.2009.12]

[52] Sebastien Ailleret. Web crawler généraliste. http://larbin.sourceforge.net

[53] Liu ZQ, Huffaker B, Fomenkov M, Brownlee N, Claffy K. Two days in the life of the DNS anycast root servers. In: Proc. of the 8th Int'l Conf. on Passive and Active Network Measurement. Louvain-la-neuve: Springer, 2007. 125–134. [doi: 10.1007/978-3-540-71617-4_13]

[54] Sarat S, Pappas V, Terzis A. On the use of anycast in DNS. In: Proc. of the 15th Int 'l Conf. on Computer Communications and Networks. Arlington: IEEE, 2006. 71–78. [doi: 10.1109/ICCCN.2006.286248]

[55] PlanetLab. https://developers.planet.com/devtrial/

[56] Moura GCM, de O Schmidt R, Heidemann J, de Vries WB, Muller M, Wei L, Hesselman C. Anycast vs. DDoS: Evaluating the November 2015 root DNS event. In: Proc. of the 2016 Internet Measurement Conf. Santa Monica: ACM, 2016. 255–270. [doi: 10.1145/2987443.2987446]

[57] RIPE Atlas. https://atlas.ripe.net

[58] Yan H, Oliveira R, Burnett K, Matthews D, Zhang LX, Massey D. BGPmon: A real-time, scalable, extensible monitoring system. In: Proc. of the 2009 Cybersecurity Applications & Technology Conf. for Homeland Security (CATCH). Washington: IEEE, 2009. 212–223. [doi: 10.1109/CATCH.2009.28]

[59] Fan X, Heidemann J, Govindan R. Evaluating anycast in the domain name system. In: Proc. of the 2013 IEEE INFOCOM. Turin: IEEE, 2013. 1681–1689. [doi: 10.1109/INFCOM.2013.6566965]

[60] Kreibich C, Weaver N, Nechaev B, Paxson V. Netalyzr: Illuminating the edge network. In: Proc. of the 10th ACM SIGCOMM Conf. on Internet Measurement. Melbourne: ACM, 2010. 246–259. [doi: 10.1145/1879141.1879173]

[61] AS112 Project. 2018. https://www.as112.net

[62] Gao ZY, Venkataramani A. Measuring update performance and consistency anomalies in managed DNS services. In: Proc. of the 2019 IEEE Conf. on Computer Communications. Paris: IEEE, 2019. 2206–2214. [doi: 10.1109/INFOCOM.2019.8737568]

[63] Allman M. Comments on DNS Robustness. In: Proc. of the 2018 Internet Measurement Conf. Boston: ACM, 2018. 84–90. [doi: 10.

1145/3278532.3278541]

[64] Abhishta A, van Rijswijk-Deij R, Nieuwenhuis LJM. Measuring the impact of a successful DDoS attack on the customer behaviour of managed DNS service providers. ACM SIGCOMM Computer Communication Review, 2019, 48(5): 70–76. [doi: 10.1145/3310165.3310175]

[65] Rijswijk-Deij RV, Jonker M, Sperotto A, Pras A. A high-performance, scalable infrastructure for large-scale active DNS measurements. IEEE Journal on Selected Areas in Communications, 2016, 34(6): 1877–1888. [doi: 10.1109/JSAC.2016.2558918]

[66] Open INTEL. Active DNS measurement project. https://openintel.nl

[67] Pappas V, Xu ZG, Lu SW, Massey D, Terzis A, Zhang LX. Impact of configuration errors on DNS robustness. ACM SIGCOMM Computer Communication Review, 2004, 34(4): 319–330. [doi: 10.1145/1030194.1015503]

[68] Kalafut AJ, Shue CA, Gupta M. Touring DNS open houses for trends and configurations. IEEE/ACM Trans. on Networking, 2011, 19(6): 1666–1675. [doi: 10.1109/TNET.2011.2130537]

[69] Deccio C, Sedayao J, Kant K, Mohapatra P. Measuring availability in the domain name system. In: Proc. of the 2010 IEEE INFOCOM. San Diego: IEEE, 2010. 1–5. [doi: 10.1109/INFCOM.2010.5462270]

[70] Kalafut AJ, Gupta M, Cole CA, Chen L, Myers NE. An empirical study of orphan DNS servers in the internet. In: Proc. of the 10th ACM SIGCOMM Conf. on Internet Measurement. Melbourne: ACM, 2010. 308–314. [doi: 10.1145/1879141.1879182]

[71] Korczyński M, Król M, van Eeten M. Zone poisoning: The how and where of non-secure DNS dynamic updates. In: Proc. of the 2016 Internet Measurement Conf. Santa Monica: ACM, 2016. 271–278. [doi: 10.1145/2987443.2987477]

[72] Farsight Security: DNS Database (DNS-DB). https://www.dnsdb.info

[73] Internet wide scan data repository: DNS records (ANY). https://scans.io/study/sonar.fdns

[74] Colajanni M, Yu PS, Cardellini V. Dynamic load balancing in geographically distributed heterogeneous web servers. In: Proc. of the 18th Int'l Conf. on Distributed Computing Systems. Amsterdam: IEEE, 1998. 295–302. [doi: 10.1109/ICDCS.1998.679729]

[75] Rabinovich M, Spatscheck O. Web caching and replication. SIGMOD Record, 2003, 32(4): 107–108. [doi: 10.1145/959060.959079]

[76] Alzoubi HA, Rabinovich M, Spatscheck O. The anatomy of LDNS clusters: Findings and implications for Web content delivery. In: Proc. of the 22nd Int'l Conf. on World Wide Web. Rio de Janeiro: ACM, 2013. 83–94. [doi: 10.1145/2488388.2488397]

[77] Mao ZM, Cranor CD, Douglis F, Rabinovich M, Spatscheck O, Wang J. A precise and efficient evaluation of the proximity between Web clients and their local DNS servers. In: Proc. of the General Track: 2002 USENIX Annual Technical Conf. 2002. 229–242.

[78] Krishnan R, Madhyastha HV, Srinivasan S, Jain S, Krishnamurthy A, Anderson T, Gao J. Moving beyond end-to-end path information to optimize CDN performance. In: Proc. of the 9th ACM SIGCOMM Conf. on Internet Measurement. 2009. 190–201.

[79] Shaikh A, Tewari R, Agrawal M. On the effectiveness of DNS-based server selection. In: Proc. of the 2001 IEEE Conf. on Computer Communications. the 20th Annual Joint Conf. of the IEEE Computer and Communications Society. Anchorage: IEEE, 2001. 1801–1810. [doi: 10.1109/INFCOM.2001.916678]

[80] Contavalli C, van der Gaast W, Lawrence D, Kumari W. Client subnet in DNS queries. RFC 7871, 2016. [doi: 10.17487/RFC7871]

[81] Chen FF, Sitaraman RK, Torres M. End-user mapping: Next generation request routing for content delivery. ACM SIGCOMM Computer Communication Review, 2015, 45(4): 167–181. [doi: 10.1145/2829988.2787500]

[82] Leonard D, Loguinov D. Demystifying service discovery: Implementing an internet-wide scanner. In: Proc. of the 10th ACM SIGCOMM Conf. on Internet Measurement. Melbourne: ACM, 2010. 109–122. [doi: 10.1145/1879141.1879156]

[83] Sisson G. DNS survey: October 2010. http://dns.measurement-factory.com/surveys/201010. [2020-07-04].

[84] Ramasubramanian V, Sirer EG. Perils of transitive trust in the domain name system. In: Proc. of the 5th ACM SIGCOMM Conf. on Internet Measurement. Berkeley: ACM, 2005. 379–384.

[85] Open Directory Project, Web directory of high-quality resources. http://www.odp.org

[86] Jiang J. Research on inconsistent and multiple dependence in the authorization mechanism of internet domain name system [Ph.D. Thesis]. Beijing: Tsinghua University, 2013 (in Chinese with English abstract).

[87] Liu WF, Zhang Y, Li YY, Fang BX. Modeling, measuring, and analyzing the resolution process of popular domains. In: Proc. of the 2019 IEEE Int'l Conf. on Communications (ICC). Shanghai: IEEE, 2019. 1–7. [doi: 10.1109/ICC.2019.8761698]

[88] Deccio C, Sedayao J, Kant K, Mohapatra P. Quantifying dns namespace influence. Computer Networks, 2012, 56(2): 780–794. [doi: 10.1016/j.comnet.2011.11.005]

[89] Jiang J, Zhang J, Duan HX, Li K, Liu W. Analysis and measurement of zone dependency in the domain name system. In: Proc. of the 2018 IEEE Int'l Conf. on Communications (ICC). Kansas City: IEEE, 2018. 1–7. [doi: 10.1109/ICC.2018.8422602]

[90] Pang J, Hendricks J, Akella A, De Prisco R, Maggs B, Seshan S. Availability, usage, and deployment characteristics of the domain name system. In: Proc. of the 4th ACM SIGCOMM Conf. on Internet Measurement. Taormina: ACM, 2004. 1–14. [doi: 10.1145/1028788.

1028790]

[91] Castro S, Wessels D, Fomenkov M, Claffy K. A day at the root of the Internet. ACM SIGCOMM Computer Communication Review, 2008, 38(5): 41–46. [doi: 10.1145/1452335.1452341]

[92] Castro S, Zhang M, John W, Wessels D, Claffy K. Understanding and preparing for DNS evolution. In: Proc. of the 2nd Int'l Workshop on Traffic Monitoring and Analysis. Zurich: Springer, 2010. 1–16. [doi: 10.1007/978-3-642-12365-8_1]

[93] Wessels D, Fomenkov M. Wow, that's a lot of packets. Int'l Workshop on Passive and Active Network Measurement. Springer. 2003.

[94] IANA. List of valid TLD. http://data.iana.org/TLD/tlds-alpha-by-domain.txt

[95] Brownlee N, Claffy KC, Nemeth E. DNS measurements at a root server. In: Proc. of the 2001 IEEE Global Telecommunications Conf. San Antonio: IEEE, 2001. 1672–1676. [doi: 10.1109/GLOCOM.2001.965864]

[96] DNS OARC. DITL traces and analysis. https://www.dns-oarc.net/oarc/data/ditl/2017

[97] Wessels D. Is your caching resolver polluting the internet? In: Proc. of the ACM SIGCOMM Workshop on Network Troubleshooting: Research, Theory and Operations Practice Meet Malfunctioning Reality. Portland: ACM, 2004. 271–276. [doi: 10.1145/1016687.1016695]

[98] Broido A, Nemeth E, Claffy K. Spectroscopy of private DNS update sources. In: Proc. the 3rd IEEE Workshop on Internet Applications. San Jose: IEEE, 2003. 19–29. [doi: 10.1109/WIAPP.2003.1210282]

[99] Abley J, Sotomayor W. AS112 nameserver operations. RFC7534, 2015.

[100] Jeong SH, Kang AR, Kim J, Kim HK, Mohaisen A. A longitudinal analysis of. i2p leakage in the public DNS infrastructure. In: Proc. of the 2016 ACM SIGCOMM Conf. Florianopolis: ACM, 2016. 557–558. [doi: 10.1145/2934872.2960423]

[101] Danzig PB, Obraczka K, Kumar A. An analysis of wide-area name server traffic: A study of the Internet Domain Name System. ACM SIGCOMM Computer Communication Review, 1992, 22(4): 281–292. [doi: 10.1145/144191.144301]

[102] Lentz M, Levin D, Castonguay J, Spring N, Bhattacharjee B. D-mystifying the D-root address change. In: Proc. of the 2013 Internet Measurement Conf. Barcelona: ACM, 2013. 57–62. [doi: 10.1145/2504730.2504772]

[103] Barber P, Larson M, Kosters M, Toscano P. Life and times of J-root. In: Proc. of the North American Network Operators' Group (NANOG) Archive. 2004.

[104] Manning B. Persistent queries and phantom nameservers. In: Proc. of the CAIDA-WIDE Workshop. 2006.

[105] Khattak S, Javed M, Khayam SA, Uzmi ZA, Paxson V. A look at the consequences of internet censorship through an ISP lens. In: Proc. of the 2014 Internet Measurement Conf. Vancouver: ACM, 2014. 271–284. [doi: 10.1145/2663716.2663750]

[106] Anderson C, Winter P, Roya. Global network interference detection over the RIPE Atlas network. In: Proc. of the 4th USENIX Workshop on Free and Open Communications on the Internet. San Diego: USENIX Association, 2014. 1–8.

[107] Aryan S, Aryan H, Halderman JA. Internet censorship in Iran: A first look. In: Proc. of the 3rd USENIX Workshop on Free and Open Communications on the Internet (FOCI 13). Washington: USENIX Association, 2013. 1–8.

[108] Chaabane A, Chen T, Cunche M, De Cristofaro E, Friedman A, Kaafar MA. Censorship in the wild: Analyzing Internet filtering in Syria. In: Proc. of 2014 Internet Measurement Conf. Vancouver: ACM, 2014. 285–298. [doi: 10.1145/2663716.2663720]

[109] Jones B, Feamster N, Paxson V. Detecting DNS root manipulation. In: Proc. of the 17th Int'l Conf. on Passive and Active Network Measurement. Heraklion: Springer, 2016. 276–288. [doi: 10.1007/978-3-319-30505-9_21]

[110] Weaver N, Kreibich C, Nechaev B, Paxson V. Implications of Netalyzr's DNS measurements. In: Proc. of the 1st Workshop on Securing and Trusting Internet Names (SATIN). 2011.

[111] Liu BJ, Lu CY, Duan HX, Liu Y, Li Z, Hao S, Yang M. Who is answering my queries: Understanding and characterizing interception of the DNS resolution path. In: Proc. of the 27th USENIX Security Symp. Baltimore: USENIX Association, 2018. 1113–1128.

[112] Migault D, Girard C, Laurent M. A performance view on DNSSEC migration. In: Proc. of the 2010 Int'l Conf. on Network and Service Management (CNSM). Niagara Falls: IEEE, 2010. 469–474. [doi: 10.1109/CNSM.2010.5691275]

[113] Kolkman OM. Measuring the resource requirements of DNSSEC. RIPE NCC/NLnet Labs. RIPE NCC, 2005.

[114] Zhu L, Heidemann J. LDplayer: DNS Experimentation at scale. In: Proc. of the 2018 Internet Measurement Conf. Boston: ACM, 2018. 119–132. [doi: 10.1145/3278532.3278544]

[115] Ager B, Dreger H, Feldmann A. Predicting the DNSSEC overhead using DNS traces. In: Proc. of the 40th Annual Conf. on Information Sciences and Systems. Princeton: IEEE, 2006. 1484–1489. [doi: 10.1109/CISS.2006.286699]

[116] Rossow C. Amplification hell: Revisiting network protocols for DDoS abuse. In: Proc. of the 21st Network and Distributed System Security Symp. 2014. [doi: 10.14722/NDSS.2014.23233]

[117] Van Rijswijk-Deij R, Sperotto A, Pras A. DNSSEC and its potential for DDoS attacks: A comprehensive measurement study. In: Proc. of the 2014 Internet Measurement Conf. Vancouver: ACM, 2014. 449–460. [doi: 10.1145/2663716.2663731]

[118] DNSSEC deployment timeline. https://www.dnssec-deployment.org/history/timeline
[119] Sisson G. DNS survey: October 2010. http://dns.measurement-factory.com/surveys/201010/dns_survey_2010.pdf
[120] Chung T, van Rijswijk-Deij R, Chandrasekaran B, Choffnes DR, Levin D, Maggs BM, Mislove A, Wilson C. A Longitudinal, end-to-end view of the DNSSEC ecosystem. In: Proc. of the 26th USENIX Security Symp. Vancouver: USENIX Association, 2017. 1307–1322.
[121] State of DNSSEC deployment 2016. https://www.internetsociety.org/wp-content/uploads/2017/08/ISOC-State-of-DNSSEC-Deployment-2016-v1.pdf
[122] Dai TX, Shulman H, Waidner M. DNSSEC misconfigurations in popular domains. In: Proc. of the 15th Int'l Conf. on Cryptology and Network Security. Milan: Springer, 2016. 651–660. [doi: 10.1007/978-3-319-48965-0_43]
[123] Wander M. Measurement survey of server-side DNSSEC adoption. In: Proc. of the 2017 Network Traffic Measurement and Analysis Conf. Dublin: IEEE, 2017. 1–9. [doi: 10.23919/TMA.2017.8002913]
[124] ICANN. TLD DNSSEC report. http://stats.research.icann.org/dns/tld_report
[125] Guðmundsson Ó, Crocker SD. Observing DNSSEC validation in the wild. In: Securing and Trusting Internet Names (SATIN). 2011.
[126] Lian W, Rescorla E, Shacham H, Savage S. Measuring the practical impact of DNSSEC deployment. In: Proc. of the 22th USENIX Security Symp. Washington: USENIX Association. 2013. 573–588.
[127] Yu YD, Wessels D, Larson M, Zhang LX. Check-Repeat: A new method of measuring DNSSEC validating resolvers. In: Proc. of the 2013 IEEE Conf. on Computer Communications Workshops. Turin: IEEE, 2013. 3147–3152. [doi: 10.1109/INFCOMW.2013.6562861]
[128] Osterweil E, Ryan M, Massey D, Zhang LX. Quantifying the operational status of the dnssec deployment. In: Proc. of the 8th ACM SIGCOMM Conf. on Internet Measurement. Vouliagmeni: ACM, 2008. 231–242. [doi: 10.1145/1452520.1452548]
[129] Yang H, Osterweil E, Massey D, Lu SW, Zhang LX. Deploying cryptography in Internet-scale systems: A case study on DNSSEC. IEEE Trans. on Dependable and Secure Computing, 2011, 8(5): 656–669. [doi: 10.1109/TDSC.2010.10]
[130] Chung T, van Rijswijk-Deij R, Choffnes D, Levin D, Maggs BM, Mislove A, Wilson C. Understanding the role of registrars in DNSSEC deployment. In: Proc. of the 2017 Internet Measurement Conf. London: ACM, 2017. 369–383. [doi: 10.1145/3131365.3131373]
[131] Adrichem NLM, Blenn N, Lúa AR, Wang X, Wasif M, Fatturrahman F, Kuipers FA. A measurement study of DNSSEC misconfigurations. Security Informatics, 2015, 4(1): 1–14. [doi: 10.1186/S13388-015-0023-Y]
[132] DNS Census 2013. https://dnscensus2013.neocities.org
[133] Lu CY, Liu BJ, Li Z, Hao S, Duan HX, Zhang MM, Leng CY, Liu Y, Zhang ZF, Wu JP. An end-to-end, large-scale measurement of DNS-over-encryption: How far have we come? In: Proc. of the 2019 Internet Measurement Conf. Amsterdam: ACM, 2019. 22–35. [doi: 10.1145/3355369.3355580]
[134] Hao S, Thomas M, Paxson V, Feamster N, Kreibich C, Grier C, Hollenbeck S. Understanding the domain registration behavior of spammers. In: Proc. of the 2013 Internet Measurement Conf. Barcelona: ACM, 2013. 63–76. [doi: 10.1145/2504730.2504753]
[135] Lauinger T, Onarlioglu K, Chaabane A, Robertson W, Kirda E. WHOIS lost in translation: (Mis) understanding domain name expiration and re-registration. In: Proc. of the 2016 Internet Measurement Conf. Santa Monica: ACM, 2016. 247–253. [doi: 10.1145/2987443.2987463]
[136] Hao S, Feamster N, Pandrangi R. Monitoring the initial DNS behavior of malicious domains. In: Proc. of the 11th ACM SIGCOMM Internet Measurement Conf. Berlin: ACM, 2011. 269–278. [doi: 10.1145/2068816.2068842]
[137] Spring JM, Metcalf LB, Stoner E. Correlating domain registrations and DNS first activity in general and for malware. In: Securing and Trusting Internet Names. 2011.
[138] Farsight Security. Security Information Exchange (SIE). https://www.farsightsecurity.com/solutions/security-information-exchange
[139] Konte M, Feamster N, Jung J. Dynamics of online scam hosting infrastructure. In: Proc. of the 10th Int'l Conf. on Passive and Active Network Measurement. Seoul: Springer, 2009. 219–228. [doi: 10.1007/978-3-642-00975-4_22]
[140] Felegyhazi M, Kreibich C, Paxson V. On the potential of proactive domain blacklisting. In: Proc. of the 3rd USENIX Conf. on Large-Scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More. USENIX Association, 2010.
[141] Antonakakis M, Perdisci R, Lee W, Vasiloglou N, Dagon D. Detecting malware domains at the upper DNS hierarchy. In: Proc. of the 20th USENIX Conf. on Security Symp. San Francisco: ACM, 2011. 27.
[142] Zhang WW, Gong J, Liu SD, Hu XY. DNS surveillance on backbone. Ruan Jian Xue Bao/Journal of Software, 2017, 28(9): 2370−2387 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5186.htm [doi: 10.13328/j.cnki.jos.005186]
[143] ITHI. ICANN's ITHI project. https://ithi.research.icann.org/about.html
[144] Casalicchio E, Caselli M, Coletta A. Measuring the global domain name system. IEEE Network, 2013, 27(1): 25–31. [doi: 10.1109/MNET.2013.6423188]

[145] Liu BJ, Lu CY, Li Z, Liu Y, Duan HX, Hao S, Zhang ZF. A reexamination of internationalized domain names: The good, the bad and the ugly. In: Proc. of the 48th Annual IEEE/IFIP Int'l Conf. on Dependable Systems and Networks. Luxembourg: IEEE, 2018. 654–665. [doi: 10.1109/DSN.2018.00072]

附中文参考文献:

[17] 王文通, 胡宁, 刘波, 刘欣, 李树栋. DNS安全防护技术研究综述. 软件学报, 2020, 31(7): 2205–2220. http://www.jos.org.cn/1000-9825/6046.htm [doi: 10.13328/j.cnki.jos.006046]

[19] 王垚, 胡铭曾, 李斌, 闫伯儒. 域名系统安全研究综述. 通信学报, 2007, 28(9): 91–103. [doi: 10.3321/j.issn:1000-436x.2007.09.015]

[20] 胡宁, 邓文平, 姚苏. 互联网DNS安全研究现状与挑战. 网络与信息安全学报, 2017, 3(3): 13–21. [doi: 10.11959/j.issn.2096-109x.2017.00154]

[86] 江健. 互联网域名系统授权机制中不一致和多重依赖问题研究[博士学位论文]. 北京: 清华大学, 2013.

[142] 张维维, 龚俭, 刘尚东, 胡晓艳. 面向主干网的DNS流量监测. 软件学报, 2017, 28(9): 2370−2387. http://www.jos.org.cn/1000-9825/5186.htm [doi: 10.13328/j.cnki.jos.005186]

刘文峰 (1992－), 男, 博士生, 主要研究领域为域名系统, 区块链.

张宏莉 (1973－), 女, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为网络信息安全, 信息内容安全.

张宇 (1979－), 男, 博士, 副教授, CCF 高级会员, 主要研究领域为互联网基础设施安全, 网络拓扑测量, 未来网络体系.

方滨兴 (1960－), 男, 博士, 教授, 博士生导师, CCF 会士, 主要研究领域为网络信息安全, 信息内容安全.