

递归-权威侧部署加密DNS协议的隐私 收益评估方法及测量分析

段丽莹¹, 李瑞烜², 刘西蒙¹, 邵俊³, 刘保君²

(1. 福州大学计算机与大数据学院, 福建 福州 350108;

2. 清华大学网络科学与网络空间研究院, 北京 100084; 3. 浙江工商大学计算机科学与技术学院, 浙江 杭州 310018)

摘要: 加密DNS协议最初被设计用于保护用户和递归解析器之间(用户-递归侧)的DNS通信隐私。目前, 加密DNS协议获得了广泛应用。然而, 递归解析器和权威服务器之间(递归-权威侧)的DNS通信仍遭受着大量隐私威胁。历经4年的标准化进程, 互联网工程任务组在2024年2月正式发布RFC 9539, 提出可利用加密DNS协议来保障递归-权威侧的DNS通信隐私。聚焦于在递归-权威侧部署加密DNS协议所带来的隐私收益, 提出评估互联网域名隐私收益的方法。针对243万个流行域名和4万个敏感域名, 结合1 058个顶级域名的区域文件, 分析权威服务器所托管的域名数量, 从而判定目标域名的隐私收益等级。测量结果表明, 超过90%的域名可获得递归-权威侧部署加密DNS协议的隐私保护, 但是6.28%的敏感域名无法从递归-权威侧部署加密DNS协议中获得隐私收益。此外, 一些高流行度的域名也没有获得足够的隐私收益。相较于大型域名托管商, 小型域名托管商可为域名提供更高的隐私收益。将域名部署于仅托管单个域名的权威服务器上会极大地损害递归-权威侧加密DNS协议的隐私保护效果, 管理人员应重新审视域名托管服务。

关键词: 域名系统; 加密DNS; 隐私保护; 互联网测量

中图分类号: TP393

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2024067

Measurement and evaluation for privacy benefits of deploying encrypted DNS protocol between recursive and authoritative servers

DUAN Liying¹, LI Ruixuan², LIU Ximeng¹, SHAO Jun³, LIU Baojun²

1. College of Computer and Data Science, Fuzhou University, Fuzhou 350108, China

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

3. School of Computer Science and Technology, Zhejiang Gongshang University, Hangzhou 310018, China

收稿日期: 2024-05-30; 修回日期: 2024-08-22

通信作者: 刘保君, lbj@mail. tsinghua. edu.cn

基金项目: 国家重点研发计划青年科学家项目(2023YFB3105600)

Foundation Item: The National Key R&D Program for Young Scientists of China (2023YFB3105600)

引用格式: 段丽莹, 李瑞烜, 刘西蒙, 等. 递归-权威侧部署加密DNS协议的隐私收益评估方法及测量分析[J]. 网络与信息
安全学报, 2024, 10(5): 71-80.

Citation Format: DUAN L Y, LI R X, LIU X M, et al. Measurement and evaluation for privacy benefits of deploying en-
crypted DNS protocol between recursive and authoritative servers[J]. Chinese Journal of Network and Information Security,
2024, 10(5): 71-80.

Abstract: The encrypted DNS protocol was originally designed to protect DNS communication privacy between users and recursive resolvers (user-recursive side). Currently, encrypted DNS communication has been widely deployed. However, DNS communications between recursive resolvers and authoritative servers (recursive-authoritative side) still faced significant privacy threats. To address this issue, the Internet Engineering Task Force (IETF) officially released RFC 9539 in February 2024, which utilized the encrypted DNS protocol to protect DNS communication privacy on the recursive-authoritative side. Focusing on the privacy benefits of deploying the encrypted DNS protocol on the recursive-authoritative side, a method to evaluate the privacy benefits of domain names was proposed. The method defined three levels of privacy benefits by analyzing the number of domain names hosted by authoritative servers of the target domain name. Combined with the zone files of 1058 top-level domains, the privacy benefit level was determined for 2.43 million popular domain names and 40 thousand sensitive domain names. The results showed that over 90% of domain names could achieve privacy protection through the deployment of encrypted DNS on the recursive-authoritative side. However, 6.28% of sensitive domain names could not benefit from such deployment. In addition, some popular domain names also did not gain privacy benefits. Compared to large domain hosting providers, smaller providers could offer higher privacy benefits for domain names. Administrators were advised not to deploy domains on authoritative servers that hosted only a single domain name, which significantly compromised the privacy protection effectiveness of encrypted DNS protocol deployment on the recursive-authoritative side.

Keywords: domain name system, encrypted DNS, privacy protection, internet measurement

0 引言

域名系统 (DNS, domain name system) 作为互联网基石, 主要功能是将用户读写友好的域名转换为机器可识别的 IP 地址。客户端向递归解析器发送域名解析请求, 递归解析器随后执行迭代查询, 并最终将权威服务器的响应返回给客户端。然而, 域名系统最初被设计为基于 UDP 的明文传输, 导致 DNS 流量极易被嗅探和篡改^[1]。为降低 DNS 威胁, 4 种加密 DNS 协议应运而生, 并被确立为互联网国际标准, 包括 DNS over TLS (DoT)^[2]、DNS over HTTPS (DoH)^[3]、DNS over QUIC (DoQ)^[4] 及 DNS over HTTP/3 (DoH3)^[5]。加密 DNS 协议允许客户端和 DNS 服务器通过协商的加密信道来传输 DNS 数据, 从而有效保护用户的隐私和安全。

加密 DNS 协议最初被设计为部署在用户和递归解析器之间 (用户-递归侧), 已经得到了众多 DNS 供应商、客户端以及操作系统的采用^[6-7]。然而, 递归解析器与权威服务器之间的通信 (递归-权威侧), 仍面临诸多安全和隐私威胁^[8]。从 2020 年开始, 保护递归-权威侧的 DNS 通信安全和隐私逐步成为互联网社区的工作重点。2024 年 2 月, 互联网工程任务组 (IETF) 正式发布 RFC 9539^[9], 规定递归解析器和权威服务器可建立

DoT 或 DoQ 会话来加密传输 DNS 报文。值得注意的是, 递归-权威侧的加密 DNS 协议不要求递归解析器验证权威服务器的身份, 因此只能用来防范递归解析器和权威服务器通信路径上的流量监听, 而无法提供身份验证和防篡改功能。

本文聚焦于递归-权威侧部署加密 DNS 协议的隐私保护效果, 提出一种评估互联网域名隐私收益等级的方法。该方法通过分析目标域名的权威服务器所托管的域名数量, 定义 3 个隐私收益等级, 从高到低分别为隐私受益、隐私受损和隐私泄露。具体而言, 处于隐私泄露等级的域名, 其所有权威服务器都仅托管一个域名, 这使得路径观察者仅根据权威服务器的 IP 地址就可直接推断出递归解析器所查询的域名。此外, 如果域名的部分权威服务器仅托管一个域名, 则该域名属于隐私受损等级。相反, 隐私受益等级的域名的所有权威服务器都托管了多个域名, 可完全获得递归-权威侧加密 DNS 协议提供的隐私保护。

本文扫描了 243 万个流行域名和 4 万个敏感域名的权威服务器, 并结合 1 058 个顶级域名的区域文件来判定域名的隐私收益等级。测量结果显示, 超过 90% 的域名可享受到递归-权威侧加密 DNS 协议所提供的隐私保护。然而, 即使在递归-权威侧部署了加密 DNS 协议, 仍有 6.28%

的敏感域名存在隐私泄露的风险。进一步分析发现,流行度更高的域名并非可获得更多的隐私收益。本文发现3个流行度排名前100的域名属于隐私泄露等级。此外,许多大型域名托管商的权威服务器上部署的域名的隐私收益有限,反而小型域名托管商可为域名提供更高的隐私保护效果。总之,为有效发挥加密DNS协议保护递归-权威侧通信隐私的能力,管理人员应将域名部署于托管多个域名的权威服务器。

1 相关工作

1.1 加密DNS技术的发展和部署

域名系统最初被设计为基于UDP的明文通信,这导致其面临众多的安全和隐私威胁^[10]。为降低这一风险,互联网社区提出了对DNS报文进行加密的解决方案。目前,已有4项加密DNS协议成为互联网国际标准,分别为DoT、DoH、DoQ以及DoH3。

加密DNS技术的部署和运行状况是互联网社区的关注重点,许多研究人员开展了大范围测量工作。2019年,Lu等^[11]首次报道了递归解析器部署DoT和DoH协议的情况,并通过被动数据揭示了真实用户的加密DNS流量变化。之后,Li等^[7]提出了全面发现加密DNS服务器域名的方法,并长期监控了加密DNS服务的部署趋势和服务集中化现象。Kosek等^[12]开展了29周的扫描工作,指出部署DoQ协议的服务器数量正在稳步上升。

此外,一些研究工作还对加密DNS服务的安全配置进行了深入探究。文献[13-14]测量了部署DoT和DoH协议的服务器对TLS版本和证书的配置情况,发现一些服务器采用废弃的TLS版本和自签名证书。Li等^[7]探究了服务器和客户端的加密DNS的实现,揭示了加密DNS服务在证书相关安全机制部署方面的系列缺陷。

上述研究主要关注用户-递归侧的加密DNS协议,也有学者对递归-权威侧的加密DNS协议部署进行了初步分析。具体来说,Li等^[7]监控了流行域名的权威服务器对DoT和DoH协议的采用,并描绘了其数量的长期演变趋势。

1.2 隐私收益评估方法

加密DNS技术可保障互联网用户的隐私安全,众多研究工作探究了其隐私保护效果。具体

来说,Houser等^[15]提出了一种DoT流量指纹分析方法,通过对数据包的时序模式建模,从而推断用户访问的网站。研究结果表明,即使DNS通信被加密,信息泄露的风险仍然存在。Bana-daki等^[16]利用机器学习方法识别统计特征,有效检测出了DoH流量中的恶意DNS活动。Hu等^[17]发现攻击者可通过分析DoQ流量中的查询间隔等关键特征,推断用户访问网站的类别。这表明即使是最新的DNS加密协议,也无法完全避免隐私泄露问题。

此外,Hoang等^[18]评估了加密DNS协议和ESNI(encrypted server name indication)机制在用户网络浏览隐私方面的保护效果,关注的是用户-递归侧的加密DNS隐私保护能力。其研究方法的核心在于使用主机 k -匿名度和IP地址轮换度两项指标,评估攻击者能否根据IP地址推断出用户访问的网站。研究结果表明,许多域名通过加密DNS协议和ESNI机制所获得的隐私保护效果甚微。

相较于之前的研究工作,本文关注的重点是递归-权威侧部署加密DNS协议所带来的隐私收益。在递归-权威侧加密DNS协议标准化的初期,本文研究可推动域名系统隐私保护的发展进程。

2 域名隐私收益等级评估方法

自加密DNS协议在2016年成为国际标准以来,用户-递归侧的加密DNS协议得到了快速部署。然而,递归-权威侧的DNS明文传输风险仍是安全社区未解决的问题。IETF在2024年发布的RFC 9539有望成为上述问题的解决方案。本节首先介绍递归-权威侧的加密DNS协议,然后提出评估域名隐私收益等级的方法。

2.1 递归-权威侧加密DNS协议

正如文献[8]所述,递归-权威侧的DNS通信会泄露大量互联网用户隐私。例如,一些网络应用会将用户姓名嵌入查询域名中,导致攻击者可从递归解析器的查询流量推断出用户的网络行为。随着加密DNS协议的快速发展,递归-权威侧的DNS隐私保护逐步成为互联网社区的关注重点之一。

2024年2月,IETF正式发布RFC 9539来保护递归解析器和权威服务器之间的DNS通信隐私。与用户-递归侧的加密DNS协议工作原理相似,递归解析器可与权威服务器进行DoT和

DoQ 会话协商, 并利用加密信道来传输 DNS 数据包。在 TLS 会话协商过程中, 递归解析器可通过验证权威服务器的证书来确保权威服务器身份的真实可信。

此外, RFC 9539 规定权威服务器可使用自签名证书, 递归解析器不得因服务器的证书无效而拒绝建立加密通信。因此, 递归-权威侧的加密 DNS 协议无法有效抵御主动攻击者的威胁, 而主要用于保障递归解析器和权威服务器之间的 DNS 通信隐私。这一规定主要是为了减少部署加密 DNS 协议所需的成本和性能开销, 以及提升递归-权威侧加密 DNS 协议与现有域名系统的兼容性。另外, RFC 9539 还规定递归解析器在与权威服务器进行 TLS 会话协商时, 在 SNI 字段中不得包含权威服务器的域名信息。

2.2 加密 DNS 隐私收益等级

在递归-权威侧部署加密 DNS 协议后, 路径上的攻击者所能监听到的有效明文信息只剩递归解析器和权威服务器的 IP 地址, 而无法看到任何 DNS 报文内容。然而, 如果路径上的攻击者根据 IP 地址就可推断出递归解析器是否在查询目标域名, 则本文认为递归-权威侧的加密 DNS 协议并不能为该域名提供隐私保护。本文基于 k -

匿名度概念, 提出评估递归-权威侧部署加密 DNS 协议的隐私收益方法。 k -匿名度是一个在数据隐私领域被广泛应用的概念, 用来衡量个体数据的匿名化程度。该评估方法的核心思想如下所述。

本文设定攻击者只具有基本的路径流量监听和数据分析能力, 而不具备高深的加密流量分析能力。在上述条件下, 递归-权威侧部署加密 DNS 协议的隐私保护效果, 取决于攻击者根据权威服务器的 IP 地址推断递归解析器正在查询的域名的难易程度, 而这又取决于权威服务器的 IP 地址上托管的域名数量。本文采用 k -匿名度表示权威服务器的 IP 地址上托管的域名数量。 k 值越大, 意味着权威服务器的 IP 地址上托管的域名越多。因此, 攻击者推断递归解析器所查询域名的难度就越大, 该域名从递归-权威侧部署加密 DNS 协议所获得的隐私收益就越多。而在低 k -匿名度的情况下, 一个权威服务器的 IP 地址只与一个或极少数的域名关联, 路径观察者可以轻易推断出递归解析器所查询的域名, 导致加密 DNS 协议无法起到预期的隐私保护效果。

图 1 展示了评估递归-权威侧部署加密 DNS 协议的隐私收益等级分类原理。用户首先向递归

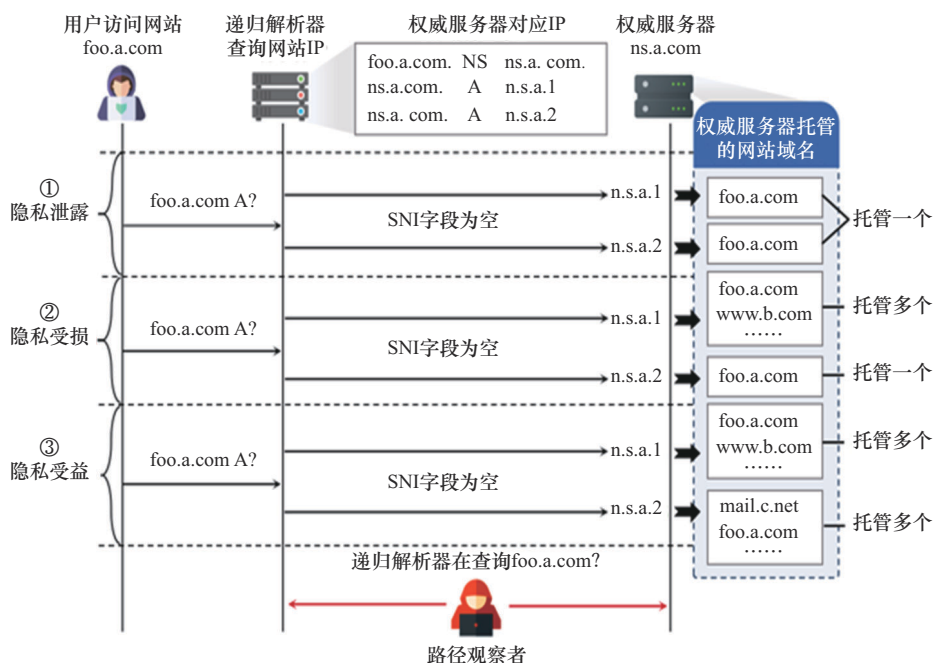


图 1 隐私收益等级分类原理

Figure 1 Principles of privacy benefit level classification

解析器查询目标域名 (foo.a.com) 的 IPv4 地址。之后, 递归解析器会向 foo.a.com 的权威服务器 (n.s.a.1 和 n.s.a.2) 查询目标域名的 IPv4 地址, 并根据协议标准将 SNI 字段设置为空, 以防止路径观察者通过 SNI 字段内容得知递归解析器所连接权威服务器的域名。根据各权威服务器所托管域名的数量, 本文定义 3 种隐私收益等级, 分别为隐私泄露、隐私受损和隐私受益。

(1) 隐私泄露

目标域名的所有权威服务器都仅托管了递归解析器正在查询的目标域名, 即 k -匿名度均为 1。在这种情况下, 路径观察者可以轻松根据权威服务器的 IP 地址确定递归解析器正在查询的域名。这种一一对应的关系使得递归-权威侧的加密 DNS 协议无法起到任何隐私保护效果。因此, 本文将这类域名的隐私收益等级定义为隐私泄露。

(2) 隐私受损

目标域名的部分权威服务器托管了多个域名, 而其余的权威服务器上仅托管了目标域名。在这种情况下, 一部分权威服务器的 k -匿名度值大于 1, 另一部分权威服务器的 k -匿名度为 1。虽然在递归解析器查询高 k -匿名度的权威服务器时, 域名可以获得加密 DNS 协议所带来的隐私保护, 但是递归解析器与 k -匿名度为 1 的权威服务器的通信仍然无法获得任何隐私收益。因此, 本文将这类域名的隐私收益等级定义为隐私受损。

(3) 隐私受益

目标域名的所有权威服务器都托管了多个域名, 即 k -匿名度都大于 1。在这种情况下, 递归解析器与任意权威服务器进行通信, 目标域名都可以获得递归-权威侧加密 DNS 协议的隐私保护。因此, 本文将这类域名的隐私收益等级定义为隐私受益。 k -匿名度值越大, 每台权威服务器上托管的域名数量就越多, 目标域名被路径观察者推断出的概率也就越低。即使路径观察者获取了权威服务器的 IP 地址, 也几乎不可能推断出递归解析器正在查询的域名。

总之, 本文提出的评估方法可揭示递归-权威侧部署加密 DNS 协议的隐私收益上限, 隐私等级分类标准和判定方法可为评估其他隐私保护协议提供参考, 如 ECH (encrypted client hello) [19]

等。此外, 本文的隐私收益量化方法能够直观地展示域名隐私的保护效果, 帮助互联网社区感知递归-权威侧部署加密 DNS 协议的效益, 更好地推动隐私保护技术的发展和创。另外, 隐私风险是多维度和连续的, 本文所采用的隐私收益评估方法不足以覆盖全部的现实场景。例如, 攻击者可利用加密流量分析技术破坏递归-权威侧加密 DNS 协议的隐私保护。

3 域名隐私收益测量

本节讲述域名和权威服务器数据集的收集方法, 以及判定域名隐私收益等级的过程。此外, 本节还将讨论测量实验过程中应遵守的伦理道德要求。

3.1 域名数据收集

本文旨在评估互联网域名的隐私收益等级, 依据之前关于隐私收益评估的论文[20], 主要关注两类域名数据集, 分别为流行域名列表和敏感域名列表。流行域名通常是访问量大、服务用户众多的域名, 如大型社交媒体平台、搜索引擎, 它们的隐私保护策略影响的用户范围较广。敏感域名通常涉及隐私敏感信息, 如金融、医疗、政府等领域的域名, 通常是路径观察者的重点关注目标。评估流行域名可以了解加密 DNS 协议对主流互联网服务和应用场景的隐私保护效果, 而对敏感域名的分析可以感知加密 DNS 协议在高风险和高隐私要求场景下的作用。通过结合两种类型的域名, 本文研究可提供更丰富的数据支持和更全面的视角。

为确保流行域名列表的全面性, 本文分别收集了 Umbrella、Majestic 以及 Tranco 流行域名列表的前 100 万个域名。其中 Majestic 列表的域名数据可能偏向于老旧网站, 对新兴网站的捕捉不足。Umbrella 列表基于被动 DNS 查询数据来构建, 收集了全球范围内最常被访问的域名。Tranco 列表综合了多个域名列表的优点, 可捕捉域名流行度的及时变化。此外, 流行域名列表中可能包括自动生成或无实际网页内容的域名, 本研究并没有将它们进行排除。为充分利用各个流行域名列表的优点, 并减少域名列表偏差所带来的影响, 本文将 3 个列表合并去重得到一个新的流行域名列表, 共包含 2 434 543 个域名。对

于敏感域名列表, 本文从公开数据源^[21]收集了 39 129 个敏感域名, 包含重要政府机构、新闻媒体等各种类型。

3.2 隐私收益等级判定

判定域名的隐私收益等级的关键在于获取其权威服务器上托管的域名数量。本文首先实现批量 DNS 解析工具, 利用该工具向公共 DNS 递归解析器查询所有流行域名和敏感域名的权威服务器。之后, 本文重复扫描 10 次, 并删除了无法获得权威服务器的域名, 后续数据分析中不会涉及这些域名。

为了全面获取权威服务器上托管的域名数量, 本文通过 ICANN 官方提供的区域文件访问服务来下载顶级域名 (TLD) 的区域文件。这些区域文件包含了域名与权威服务器的映射关系。区域文件直接由管理顶级域名的权威机构生成且定期更新, 这保证了本文所搜集数据的全面性和实时性。本文共获得了 1 058 个顶级域名的区域文件, 其中包含了超过 2 亿个互联网域名和其权威服务器的数据。之后, 本文通过计算区域文件中每台权威服务器所托管的域名数量, 来获得流行域名和敏感域名的权威服务器的 k -匿名度。最后, 根据第 2.2 节所描述的隐私收益等级评估方法, 本文可判定所有流行域名和敏感域名的隐私收益等级。此外, 为进一步深入分析测量数据, 本文还通过 IP 信息查询工具^[7]获取了权威服务器所属的供应商信息。

先前的隐私收益评估方法主要是基于机器学习或基于信息映射技术路线的。具体来说, Banadaki 等^[16]利用机器学习来评估 DoH 协议对域名解析的隐私保护效果, 可以较准确地识别潜在的隐私泄露威胁, 反映出 DoH 协议的隐私保护能力的下限。然而, 基于机器学习的方法对攻击者的能力要求较高, 可被真实利用的场景相对比较局限。此外, Hoang 等^[18]通过 IP 地址和域名的映射关系评估了加密 DNS 协议对用户网络浏览隐私的保护效果, 可揭示加密 DNS 协议的隐私保护能力的上限。本文采用基于信息映射的方法, 聚焦于递归-权威侧的加密 DNS 通信, 全面且具体地量化了部署加密 DNS 协议所带来的基本隐私保护效果。

为了避免测量实验对真实互联网和目标主机造成压力, 本文严格遵守互联网测量研究的伦理道德要求, 精心设计测试实验来减少潜在风险。具体来说, 本文严格限制扫描速率, 确保发送的数据包不会给服务器的正常运行带来影响。在获取到目标服务器返回的响应后, 立刻断开与服务器的连接, 而不会发送任何额外数据。除此之外, 本文还为扫描服务器部署在线网页, 说明实验意图和联系方式来允许主机退出后续的测量实验。

4 域名隐私收益分析

本节首先展示递归-权威侧部署加密 DNS 协议给互联网 DNS 带来的整体隐私收益; 然后, 从域名流行度和域名托管商两个角度深入评估域名隐私收益。

4.1 域名隐私收益等级评估

在去除无法获得权威服务器的域名后, 本文共对 1 647 169 个流行域名和 19 116 个敏感域名展开评估。表 1 展示了流行域名和敏感域名中不同隐私收益等级的占比。1 582 509 (96.07%) 个流行域名和 17 766 (92.94%) 个敏感域名属于隐私受益等级, 它们的权威服务器都托管了多个域名, 递归-权威侧加密 DNS 协议能够有效增强这些域名的隐私保护。糟糕的是, 7 792 (0.48%) 个流行域名和 1 201 (6.28%) 个敏感域名属于隐私泄露等级, 这些域名的权威服务器都只托管了一个域名。即使递归-权威侧部署了加密 DNS 协议, 路径上的观察者仍可通过简单分析 IP 地址, 识别出递归服务器是否在查询目标域名。此外, 敏感域名通常是路径嗅探者关注的重点, 这些域名的所有者应在高 k -匿名度的权威服务器上部署域名, 以更好地获得递归-权威侧部署加密 DNS 协议的隐私保护。

表 1 域名隐私收益等级评估结果

Table 1 Privacy benefits levels assessment result			
域名	隐私受益	隐私受损	隐私泄露
流行域名	96.07%	3.45%	0.48%
敏感域名	92.94%	0.78%	6.28%

域名隐私收益等级的判定依赖于权威服务器的 k -匿名度。图 2 展示了流行域名和敏感域名的

权威服务器的 k -匿名度累计分布函数。结果显示大部分的权威服务器托管了超过1个域名。敏感域名的权威服务器中大约一半托管超过1 000个域名。相较于敏感域名,流行域名的权威服务器上托管的域名数量较少。10%左右的流行域名的权威服务器上仅托管了1个域名。这种一对一的映射关系导致这些权威服务器无法使用加密DNS协议来为用户提供任何隐私保护,极易受到路径观察者的监听和嗅探。此外,大约40%的流行域名的权威服务器上托管的域名数量少于10个。这些权威服务器在部署加密DNS协议后,为域名提供的隐私保护收益相对较少。

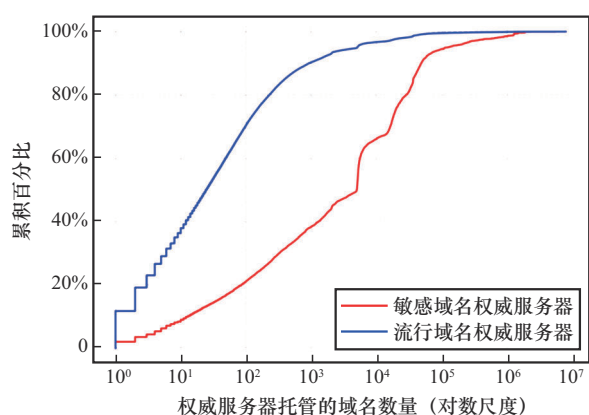


图2 权威服务器 k -匿名度累计分布函数

Figure 2 CDF of k -anonymity for authoritative servers

综上所述,在递归-权威侧部署加密DNS协议后,可增强对大部分域名的隐私保护能力,有效防范递归查询过程中的监听、嗅探,保障用户隐私安全。然而,目前许多域名的权威服务器上托管的域名数量较少,递归-权威侧加密DNS协议的隐私保护未能发挥出最佳效果。

4.2 域名流行度与隐私收益的关系

域名流行度往往与访问量成正比。流行域名的隐私收益等级对互联网用户来说极其重要。本节针对不同流行度域名的隐私收益等级展开分析。

本文所收集的流行域名数据来源于3个独立的列表。为尽可能尊重域名的流行度,本文以域名在Tranco、Majestic及Umbrella这3个流行域名列表的最高排名作为其最终排名。表2展示了不同流行度域名的隐私收益等级。结果表明,高流行度的域名并不一定表现出更高的隐

私收益等级。在排名前100的域名中,隐私受益等级的域名数量为139个(91.45%);而排名前10万的域名中,隐私受益等级的域名数量为159 521个(94.10%)。在云计算服务快速发展的背景下,流行域名的所有者往往有能力自己部署权威服务器或购买专用的域名托管服务,所以与流行域名共享权威服务器的域名数量相对较少。然而,普通的域名所有者通常选择基础的域名托管服务器,这些服务器可能被大量域名所共享。因此,在当前的域名系统中,递归-权威侧加密DNS协议并不会为流行度更高的域名提供更多的隐私保护。

表2 不同流行度域名的隐私收益等级
Table 2 Privacy benefit levels for domains of different popularity

排名	隐私受益	隐私受损	隐私泄露
前100	91.45%	6.58%	1.97%
前1 000	92.43%	5.91%	1.66%
前1万	92.86%	6.24%	0.90%
前10万	94.10%	5.20%	0.70%
前100万	96.07%	3.45%	0.48%

此外,隐私泄露的域名在各个流行度级别中占比均相对较低,大约为1%。在流行度排名前100的域名中,本文发现3个隐私泄露等级的域名,分别为trafficmanager.net、l-msedge.net及s-msedge.net。它们都属于微软公司,如s-msedge.net是Windows 10用于提供连接服务的域名。这可能是微软出于响应性能和安全控制考虑,选择将域名托管于专用的权威服务器,但这也影响域名在递归-权威侧加密DNS通信中的匿名度。

高流行度域名的隐私收益等级会对递归-权威侧加密DNS通信的隐私保护产生巨大影响。互联网社区应提高流行域名的权威服务器的 k -匿名度,这有助于推动域名系统的隐私保护发展。

4.3 域名托管商对隐私收益的影响

随着云服务的快速发展,共享托管商愈发成为一种流行的域名部署方式。域名托管商维护一批权威服务器,用来存储和分发客户域名的资源记录。托管商的域名托管策略会影响大量域名的隐私收益等级。本节探究在递归-权威侧部署加

密 DNS 协议后, 不同托管商的权威服务器上域名的隐私收益。

本文统计了各域名托管商的权威服务器的 k -匿名度平均值, 其值越大, 说明其上托管的域名所获得的隐私收益越高。表 3 展示了权威服务器的 k -匿名度排名前 5 的流行域名的托管商, 域名排名高的托管商的权威服务器数量相对较少, 这些被托管的域名通常不是特别流行的域名。然而, 这些托管商的权威服务器上托管的域名数量反而相对更多, 因此每个域名都可以获得较高的隐私收益。例如, AS39921 Rackspace Ltd. 虽然只拥有 6 台权威服务器, 但这些服务器平均每台托管了 1 959 个域名。中国的许多域名托管商, 如中国互联网络信息中心 (CNNIC)、中国联合网络通信集团有限公司 (简称“中国联通”)、杭州阿里巴巴广告有限公司, 都拥有 k -匿名度较高的权威服务器。此外, 高 k -匿名度的权威服务器上托管域名的流行度排名并不高, 通常都排在 10 万之后。

对于敏感域名来说, 高 k -匿名度权威服务器的域名托管商和流行域名相似。表 4 展示了权威服务器的 k -匿名度排名前 5 的敏感域名的托管商, 他们大多数为拥有少量权威服务器的小型域名托管商。AS27257 Webair 拥有 8 台权威服务

器, 平均每台服务器上托管了 7 648 个域名, 因此其中的敏感域名可以很好地享受递归-权威侧部署加密 DNS 协议所带来的隐私收益。

受欢迎的大型域名托管商, 如 Cloudflare, 通常拥有数万个权威服务器的 IP 地址, 并且托管着大量流行度较高的域名。然而, 相对于小型托管商, 大型托管商的每台权威服务器上托管的域名数量却相对较少。

总之, 大量流行度较低的域名托管于拥有少量权威服务器的托管商, 可受益于权威服务器的高 k -匿名度。而一些流行度较高的域名托管于拥有大量权威服务器的托管商, 却受制于权威服务器的低 k -匿名度, 无法很好地获得递归-权威侧部署加密 DNS 协议的隐私收益。

加密 DNS 协议作为有效保护递归解析器和权威服务器之间 DNS 数据隐私的重要解决方案, 其保护效果很大程度上依赖于域名的托管商。本文建议域名所有者在 k -匿名度较高的权威服务器上托管域名, 可有效增强域名的隐私保护能力。此外, 域名托管商应加强权威服务器的 IP 地址轮换度。如果权威服务器的 IP 地址经常改变, 那么路径观察者几乎无法实时获取所有目标域名和权威服务器的映射, 因此递归-权威侧加密 DNS 协议提供的隐私保护得以保障。

表 3 权威服务器的 k -匿名度排名前 5 的流行域名的托管商

Table 3 Top 5 popular domain hosting providers with k -anonymity of authoritative servers

域名托管商	k -匿名度平均值	权威服务器数量/台	域名最高排名
AS39921 Rackspace Ltd.	1 959.17	6	169 086
AS37963 CNNIC	1 940.18	102	234 145
AS47836 SC Web Software Development SRL	1 797.5	2	70 623
AS4837 China Unicom Jiangsu Province Network	1 042.43	22	288 224
AS28299 Cyberweb Networks Ltd.	1 030.83	12	281 381

表 4 权威服务器的 k -匿名度排名前 5 的敏感域名的托管商

Table 4 Top 5 sensitive domain hosting providers with k -anonymity of authoritative servers

域名托管商	k -匿名度平均值	权威服务器数量/台	托管域名数量
AS27257 Webair	7 648.0	8	1
AS37963 CNNIC	3 030.4	102	45
AS37963 Hangzhou Alibaba Advertising Co., Ltd.	2 408.98	236	3
AS17816 China Unicom Guangdong Province Network	2 364.5	9	1
AS29066 Host Europe GmbH	2 160.9	594	123

5 结束语

本文系统评估了在递归-权威侧部署加密DNS协议的隐私收益,并提出基于 k -匿名度的域名隐私收益评估方法,通过广泛收集流行域名和敏感域名数据,并结合顶级域名的区域文件数据,全面分析了域名的隐私收益等级,以及不同流行度和不同托管商的域名所获得的隐私保护效果。测量结果表明,超过90%的域名能够从递归-权威侧部署的加密DNS协议中获得隐私收益。然而,仍有大量域名无法通过加密DNS协议来防范递归查询过程中的隐私泄露。相较于大型域名托管商,小型域名托管商可通过递归-权威侧加密DNS协议为域名提供更多的隐私收益。本文建议管理人员选择托管多个域名的权威服务器来部署域名,这可最大化递归-权威侧加密DNS协议的隐私保护效果,提高域名通信数据的隐私性。

参考文献:

- [1] LIUB, LU C, DUAN H, et al. Who is answering my queries: understanding and characterizing interception of the DNS resolution path[C]//Proceedings of the 2019 Applied Networking Research Workshop (ANRW). New York: Association for Computing Machinery, 2019: 15-16.
- [2] HU Z, ZHU L, HEIDEMANN J, et al. Specification for DNS over transport layer security (TLS)[R]. 2016.
- [3] HOFFMAN P, MCMANUS P. DNS queries over HTTPS (DoH)[R]. 2018.
- [4] HUITEMA C, DICKINSON S, MANKIN A. DNS over dedicated QUIC connections[R]. 2022.
- [5] BISHOP M. HTTP/3[R]. 2022.
- [6] JENSEN T, PASHOV I, MONTENEGRO G. Windows will improve user privacy with DNS over HTTPS[J]. Retrieved May, 2019, 15: 2020.
- [7] LI R, JIA X, ZHANG Z, et al. A longitudinal and comprehensive measurement of dns strict privacy[J]. IEEE/ACM Transactions on Networking, 2023, 31(6): 2793-2808.
- [8] ED T W. DNS privacy considerations[R]. 2021.
- [9] PETIT-HUGUENIN M, SALGUEIRO G, ROSENBERG J. Session traversal utilities for NAT (STUN) extension for third-party authorization[R]. 2021.
- [10] STÉPHANE BORTZMEYER. DNS privacy considerations[R]. 2015.
- [11] LU C Y, LIU B J, LI Z, et al. An end-to-end, large-scale measurement of DNS-over-encryption: how far have we come [C]//Proceedings of the Internet Measurement Conference. New York: ACM, 2019: 22-35.
- [12] KOSEK M, DOAN T V, GRANDERATH M, et al. One to rule them all? a first look at DNS over QUIC[C]//Proceedings of the International Conference on Passive and Active Network Measurement. Cham: Springer, 2022: 537-551.
- [13] BÖTTGER T, CUADRADO F, ANTICHI G, et al. An empirical study of the cost of DNS-over-HTTPS[C]//Proceedings of the Internet Measurement Conference. New York: ACM, 2019: 15-21.
- [14] CHUNG T, VAN RIJSWIJK-DEIJ R, CHOFFNES D, et al. Understanding the security implications of DoT and DoH deployments[C]//Proceedings of the Internet Measurement Conference. New York: ACM, 2020: 347-361.
- [15] HOUSER R, LI Z, COTTON C, et al. An investigation on information leakage of DNS over TLS[C]//Proceedings of the 15th International Conference on Emerging Networking Experiments and Technologies. New York: ACM, 2019: 123-137.
- [16] BANADAKI Y. Detecting malicious DNS over HTTPS traffic in domain name system using machine learning classifiers[J]. Journal of Computer Sciences and Applications, 2020 (8): 46-55.
- [17] HU G N, FUKUDA K. An analysis of privacy leakage in DoQ traffic[C]//Proceedings of the CoNEXT Student Workshop. New York: ACM, 2021: 7-8.
- [18] HOANG N P, AKHAVAN NIAKI A, BORISOV N, et al. Assessing the privacy benefits of domain name encryption[C]//Proceedings of the 15th ACM Asia Conference on Computer and Communications Security. New York: ACM, 2020: 290-304.
- [19] ERIC R, KAZUHO O, NICK S, et al. TLS encrypted client hello [R]. 2024.
- [20] NIAKI A A, CHO S, WEINBERG Z, et al. ICLab: a global, longitudinal internet censorship measurement platform[C]//Proceedings of the 2020 IEEE Symposium on Security and Privacy (SP). Piscataway: IEEE Press, 2020.
- [21] LI R X, LIU B J, LU C Y, et al. A worldwide view on the reachability of encrypted DNS services[C]//Proceedings of the ACM Web Conference 2024 (WWW). New York: ACM, 2024: 1193-1202.

[作者简介]



段丽莹 (2000-), 女, 湖南株洲人, 福州大学硕士生, 主要研究方向为网络安全与互联网测量。



李瑞烜（1999-），男，山西忻州人，硕士，清华大学工程师，主要研究方向为网络安全与互联网测量。



邵俊（1981-），男，浙江金华人，博士，浙江工商大学教授，主要研究方向为网络安全与密码学。



刘西蒙（1988-），男，陕西西安人，博士，福州大学教授、博士生导师，主要研究方向为密码学与人工智能安全。



刘保君（1994-），男，安徽宿州人，博士，清华大学助理教授、博士生导师，主要研究方向为域名安全、网络测量、网络犯罪检测。