第 1 章 引 言

1.1 研究背景

域名（domain name）是互联网上识别和定位计算机的层次结构式的字符标识。实现域名与 IP 地址等互联网资源之间的映射，是进行几乎所有互联网活动的首要条件。截至 2021 年第三季度，全球域名保有量已超过 3.6 亿个[1]，构成了国际互联网的关键基础资源。维护全球唯一的域名空间（domain name space）、健全互联网域名的管理和功能体系，是保障国际互联网稳定运行和防止国际互联网分裂的重要前提。《“十四五”信息通信行业发展规划》[2] 将规范域名的注册使用、加强域名服务安全保障能力建设纳入发展重点，聚焦全面增强互联网基础管理能力，着力防范遏制重特大网络安全事件。

互联网域名体系由一系列机构、主机、协议和规范构成，根据对应功能划分为三个层面：一是域名注册（register）层面，主要包含互联网名称与数字地址分配机构（Internet Corporation for Assigned Names and Numbers，ICANN）、域名注册机构和注册数据服务器等，用于实现域名注册信息的收集和存储，对应域名的分配功能；二是域名解析（resolve）层面，主要包含域名服务器和域名协议等，用于实现域名与互联网资源之间映射关系的维护和查询，对应域名的检索功能；三是域名监管（administer）层面，主要包含政府部门和网络安全研究机构等，用于实现对涉及互联网攻击行为的域名进行检测和阻断，对应域名的撤销功能。各层面实体共同维护全球唯一的域名空间，保障相关功能的高效平稳运行。

互联网域名体系面临安全威胁。长期以来，互联网域名安全事件频发，催生多种现实网络安全威胁，暴露出互联网域名体系中存在的一系

列脆弱性。在域名注册层面，注册机构对域名持有人信息（例如姓名、电话、邮寄地址等）的广泛收集面临隐私泄露风险，曾诱发多起大规模注册数据窃取事件[3-4]。2021 年，“匿名者”黑客组织从域名注册商 Epik 处窃取了超过 1500 万条域名注册人的电子邮件地址，相关个人信息面临被滥用的安全风险[5]。在域名解析层面，普通域名报文采用基于用户数据报协议（user datagram protocol，UDP）的明文传输模式，缺乏消息保密性和完整性保障，易遭受报文劫持[6-7] 和流量嗅探[8] 攻击。由斯诺登曝光的多项美国国家安全局秘密计划[9-10] 亦证实，针对域名报文的大规模劫持和嗅探已被应用于国家层面的网络安全行动。在域名监管层面，域名滥用问题日益突出，大量域名被用于僵尸网络控制调度[11-12]、钓鱼欺诈[13-14]、展示违法信息[15-16] 等互联网攻击行为。然而，长期以来的学术研究和域名监管实践大多集中于对多种滥用行为的识别和检测，缺乏统一技术在全球互联网范围内实现对恶意域名的访问阻断。

近年来，为治理域名安全风险、保障域名空间的稳健，在互联网域名体系中设计引入了一系列安全技术。具体地，根据对应体系层面，将已形成标准化最佳安全实践（best security practice）的三方面安全技术总结如下：

1. 域名注册隐私保护技术

根据域名注册的有关规定[17-18]，域名持有人需提供准确的姓名、电话、邮寄地址等个人信息，构成域名注册数据。随着数据窃取事件的频发和国家层面数据安全法律的相继出台，有必要对域名注册数据进行隐私保护。相应地，ICANN 于 2018 年制定并公布《通用顶级域（generic top-level domain，gTLD）注册数据临时规范》[19]，明确全球域名注册机构需对注册数据中包含的个人信息进行访问控制和匿名化处理。此外，域名注册数据长期以来被大量网络安全基础研究依赖，广泛应用于网络欺诈检测、网络犯罪溯源等业务。公开资料[20-22] 虽曾论证隐私保护技术对网络安全业务产生的负面影响，但缺乏系统性的定量分析。全球域名注册机构是否按规范妥善保护个人信息、隐私保护技术如何影响网络安全基础研究，目前仍不清楚。

2. 域名解析安全增强技术

域名协议标准形成于 1987 年，采用基于 UDP 的明文传输模式，具有出色的运行性能。为弥补普通域名协议安全特性的缺失，设计提出了一系列域名解析安全增强技术：加密域名协议（包含基于传输层安全的域名协议（DNS-over-TLS，DoT）[23] 和基于超文本传输安全的域名协议（DNS-over-HTTPS，DoH）[24]）在客户端和域名服务器间建立加密信道实现域名报文传输，提供消息保密性保护；域名签名协议（即域名系统安全扩展（domain name system security extensions，DNSSEC）协议[25]）通过数字签名算法实现对响应报文的验证，提供消息完整性保护；域名报文随机性增强方案（包含随机源端口和消息序号[26]，以及域名 0x20 编码[27]）通过提升响应报文伪造难度有效缓解多种域名劫持攻击。各相关安全协议和方案，特别是起步较晚的协议和方案的实际部署应用态势如何及其在互联网环境中存在怎样的现实缺陷，目前仍不清楚。

3. 域名监管中的查封技术

域名作为互联网的关键基础资源，在被应用于合法业务的同时，也被大量互联网攻击行为所利用。针对日益频发的域名滥用行为，引入了查封（take-down，又称 seizure）技术[28] 对恶意域名进行阻断和撤销。具体地，监管机构维护域名黑洞（sinkhole），将恶意域名强制解析至安全的受控主机，使得其原始指向的恶意互联网资源（例如钓鱼欺诈网站）在全球互联网范围内均无法被访问。然而，有关域名查封行为的公开资料极少，可能导致域名被重复查封等不良后果[29]；监管机构对恶意域名的认定规则不明，例如美国司法部曾于 2021 年以散播虚假消息为由查封伊朗媒体域名[30-31]，引发广泛争议。互联网中有多少域名已被监管机构查封及对于恶意域名的认定规则和管理维护是否存在安全漏洞，目前仍不清楚。

安全技术测量研究具有现实意义。互联网域名体系在各组成层面引入的安全技术，形成了一系列互联网标准、行业规范和最佳安全实践，构成了遏制域名安全风险的核心前提。然而，相关协议和方案的部署应用态势和现实缺陷、能否有效治理域名安全风险，目前仍不清楚。对互联网域名体系安全技术进行大规模、系统性的测量研究，具有以下方面的实际意义：一是通过测量相关协议和方案的部署应用现状，可以分析其近年来成功得到普及或长期推广不力的具体原因；二是通过识别域名服务

提供者的错误配置和管理漏洞，能够发现各相关协议和方案的现实缺陷，揭示不规范的部署应用面临的安全风险；三是基于测量研究结果，能够从协议和规范制定者、域名管理和服务提供者、互联网用户、网络安全研究者等相关实体的角度为进一步治理域名安全风险、保障互联网域名体系的稳健提供规范建议。

1.2 研 究 内 容

本书以互联网域名体系安全技术作为研究对象，对相关协议和方案的部署应用现状和现实缺陷进行测量研究，旨在为进一步治理域名安全风险提供规范建议。当前，国际互联网域名空间内有数亿域名投入使用，由全球数千域名注册机构、数百万域名服务器和其他实体参与维护并提供功能，构成全球最大的开放分布式数据库。从这一研究角度出发，本书总结出如下关键科学问题：如何通用表征分布式网络数据库查询检索过程的安全与隐私属性，并分析其脆弱性？基于上述问题，根据互联网域名体系框架和安全技术对应的体系层面，总结得出本书的以下主要研究内容：

1. 互联网域名体系技术框架

本书对互联网域名空间进行建模，基于集合论和代数理论给出有关概念和定义。基于分布式网络数据库的思想，建立互联网域名体系的技术框架，对各层面实体的主要功能、面临的安全脆弱性和引入的安全技术进行形式化描述。进一步地，提出本书的主要研究问题，为后续研究内容提供理论支撑。

2. 域名注册隐私保护技术测量研究

在域名注册层面，关于域名注册隐私保护技术的规范于 2018 年提出，供全球域名注册机构参照执行。然而，尚未有工作系统性分析全球域名注册机构对于数据访问控制的执行情况和域名注册数据缺失对网络安全基础研究的负面影响。本书采用数据驱动的设计思路，提出并实现了基于文本相似性特征的数据隐私合规性分析系统，对全球域名注册机构的数据隐私保护技术进行测量研究。在此基础上，通过对近年发表的网络

空间安全相关文献进行收集和分类，定量分析域名注册隐私保护技术对网络安全基础研究产生的制约。

3. 域名解析安全增强技术测量研究

在域名解析层面，相关安全增强协议和方案于 2005 年至 2018 年陆续形成，近年来均得到了工业界的大力推广和广泛的软件实现。本书提出并实现了针对域名解析安全增强技术的主被动方法结合的大规模测量系统，对各相关协议和方案的部署应用情况、服务质量和性能开销进行分析和对比。在此基础上，本书通过发起全球大规模域名解析交互，识别相关协议和方案在应用中暴露的现实缺陷和安全风险。

4. 域名监管中的查封技术测量研究

在域名监管层面，针对域名查封的一般规范于 2012 年形成。然而长期以来，监管机构的域名查封行为存在高度不透明性，对恶意域名的认定标准不清晰；外界极难准确判断恶意域名的当前状态及相关安全风险是否已被缓解。本书采用数据驱动的设计思路，提出并实现了基于域名状态转移图的域名查封行为挖掘与关联系统。在此基础上，本书通过观察被查封域名的类别和历史状态，分析各机构的域名监管策略，揭示了域名查封技术中存在的安全漏洞。

1.3 主 要 贡 献

本书的主要贡献包含以下三方面：

1. 针对域名注册隐私保护技术，提出基于文本相似性特征的数据隐私合规性分析方法，证实注册数据访问控制的广泛应用对网络安全基础研究产生普遍制约

本书首次提出并实现了基于文本相似性特征的数据隐私合规性分析系统，共分析了全球 256 个域名注册机构，证实超过 85%的机构已按照现行规范要求对其管辖的个人信息进行访问控制。同时发现部分机构尚未部署隐私保护技术，或者设置过于严格的规则导致数字证书签发、网站漏洞披露等网络安全业务无法进行。此外，由于技术规范预留的执行

时间过短，全球机构普遍超前保护了所有域名的注册数据，导致大规模的公开基础数据损失。本书对 2005 年以来发表的 4304 项学术论文进行的定量分析表明，网络安全基础研究对公开域名注册数据的依赖程度呈现逐年上升的趋势；然而，以域名注册数据作为输入的研究方案中，高达 69%的工作将受到隐私保护技术的制约，需要进行局部调整、重新设计或寻找替代数据源。本书根据主要结论，向政策和规范制定者、域名注册机构以及网络安全研究者等方面提出了具体建议。

相关研究成果发表于 ISOC NDSS 2021 会议（网络空间安全领域国际顶级学术会议，TH-CPL 列表推荐 A 类会议）。研究成果得到多家权威网络安全机构报道，例如欧洲理事会网络犯罪项目办公室（Cybercrime Programme Office of the Council of European，C-PROC）[32]、瑞典国家计算机安全响应中心（Computer Emergency Response Center of Sweden，CERT-SE）[33] 等。研究成果同时推动形成了针对现行隐私保护规范的改进提案，部分结论被引用于网络安全研究机构致 ICANN 关于域名注册数据管理的调查报告和建议函[34]。

2. 针对域名解析安全增强技术，设计实现主被动方法结合的大规模测量平台，揭示域名安全协议的全球部署态势以及普遍存在的服务配置缺陷

本书提出并实现了主被动方法结合的大规模测量系统，分析了相关安全增强协议和方案的部署应用现状。首次对加密域名协议进行测量研究，发现其自形成标准以来的部署应用规模扩展迅速，域名服务器数量和协议流量均存在明显增长。在服务质量和性能开销方面，本书证实加密域名协议的整体查询成功率高于普通域名协议，且在复用连接时仅带来毫秒级别的额外查询时间。针对域名签名协议，本书发现其部署规模虽有小幅增长，但截至 2022 年（即协议标准形成 15 年后），域名签名率仅为 3.4%，仍然低于预期。针对域名报文随机性增强方案，发现占比超过 99%的递归域名服务器都使用随机源端口和消息序号，具备一定的消息完整性保障。此外，各协议在实际部署时普遍存在使用无效数字证书、连接超时管理不当等配置缺陷，可能导致域名解析失败或安全防护功能失效，需要及时进行修正。根据主要结论，向协议设计者、服务提供者以及互联网用户等方面提出了具体建议。

相关研究成果发表于 ACM IMC 2019 会议（互联网测量领域国际顶级学术会议，TH-CPL 列表推荐 A 类会议），同时获得会议最佳论文奖提名和社区贡献奖提名。研究成果有力推动了域名安全协议的部署与应用，获得由国际互联网研究任务组（Internet Research Task Force，IRTF）颁发的应用网络研究奖（Applied Networking Research Prize，ANRP）。研究成果同时促进了域名协议标准的完善，转化为国家通信行业标准一项[①]；部分结论被国际互联网标准文档 *RFC* 9076: *DNS Privacy Considerations*[8] 和 ICANN 域名安全和稳定性规范文档 *SAC* 109: *The Implications of DNS over HTTPS and DNS over TLS*[35] 引用。

3. 针对域名监管中的查封技术，提出基于域名状态转移图的域名查封行为挖掘与关联方案，发现相关监管机构缺乏指导规范且存在严重的安全漏洞

本书首次提出并实现了基于域名状态转移图的域名查封行为挖掘与关联系统，共检出 179 个活跃的域名黑洞和 206 199 个被查封的二级域名，证实了基于域名黑洞的查封行为已成为一种常见的域名监管实践。对监管机构进行分类，发现网络安全公司（例如 Microsoft、AnubisNetworks 等）是目前最主要的参与域名查封业务的监管机构，其维护的域名黑洞数量占比超过 45%。本书还对被查封域名进行分类分析，发现用于僵尸网络调度的算法生成（domain generation algorithm，DGA）域名为当前监管机构的重点打击对象，占所有被查封域名的 80.6%。然而，监管机构对恶意域名的认定标准和释放条件及依赖的网络基础设施均存在较大差异。域名查封行为呈现自发性特点，反映出指导规范的缺乏。此外，本书发现部分监管机构（例如 Netscout 公司、美国联邦调查局 FBI 等）对域名黑洞的管理存在严重的安全漏洞，其维护的多个域名黑洞已过期并可被任何个人或机构接管。通过实际的过期域名黑洞接管实践，证实网络攻击者能够以较低成本控制大量活跃的僵尸网络傀儡机。根据主要结论，向政策和规范制定者以及域名监管机构等方面提出了具体建议。

① 《域名系统解析数据加密传输技术要求》，中华人民共和国通信行业标准（标准号：YD/T 4712—2024，2024 年 3 月发布）。

1.4 组 织 结 构

全书共包含七章，按照图 1.1 所示的结构进行组织。第 1 章为引言，论述互联网域名体系安全技术的研究背景、本书的研究内容和主要贡献，并给出全书的组织结构。第 2 章基于互联网域名体系的各组成层面，对近年来发表的相关安全研究进行梳理和总结。第 3 章建立互联网域名体系技术框架，给出有关概念和问题的形式化表示，为本书的主要研究内容提供理论基础。第 4 章针对域名注册隐私保护技术提出合规性分析系统，对数据访问控制规则的应用情况进行测量研究。第 5 章针对域名解析安全增强技术实现大规模测量平台，对相关协议和方案的部署现状进行测量研究。第 6 章针对域名监管中的查封技术提出挖掘与关联方案，对域名查封行为规模进行测量研究。第 7 章总结本书的主要内容，并对未来的研究工作进行展望。

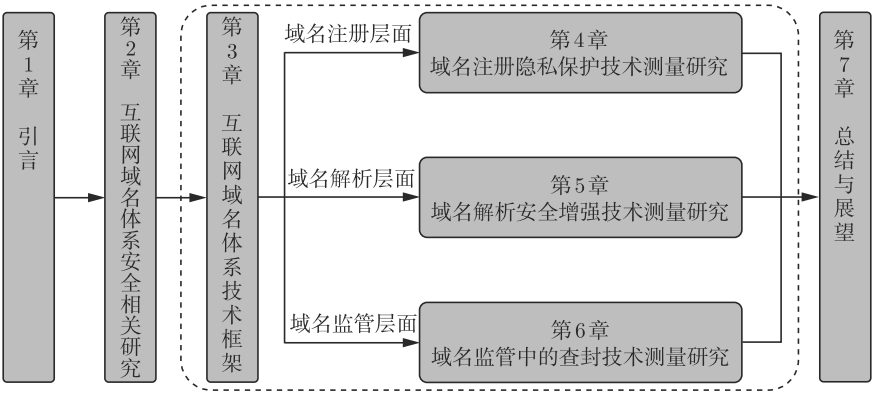

图 1.1 全书组织结构