网络空间反测绘技术研究

刘庆云[1], 李仁杰[1,2], 周 舟[1], 钟友兵[1,2], 石峰源[1,2], 郭 莉[1,2]

[1]中国科学院信息工程研究所 信息内容安全国家工程研发中心 北京 中国 100093
[2]中国科学院大学 网络空间安全学院 北京 中国 100049

摘要 网络空间测绘技术的迅猛发展与广泛推广在极大促进人们对各类网络空间资源及其属性的全方位认知的同时, 也为攻击方绘制“攻击面地图”提供了极大便利, 严重威胁到关键信息基础设施的安全性。在网络攻防动态博弈演进的过程中, 网络空间反测绘是构筑安全防护屏障的重要一环, 是保护各类网络空间资源及其属性不被测绘方探测、分析和绘制的全方位防御过程, 有助于在网络空间博弈中占据更多主动权。网络空间反测绘的核心思想是通过阻断测绘方的探测和扰乱测绘方对探测数据的关联分析, 使得测绘方无法绘制出动态、实时、可靠的网络空间地图, 实现“测不到、测不准、绘不对”的防御效果。本文首先阐述了网络空间测绘的相关概念与现有测绘技术体系, 进而给出了网络空间反测绘的相关概念及定义。然后, 提出了网络空间反测绘技术体系, 从探测行为识别、探测行为防护和测绘分析欺骗三个层次梳理了相关关键技术与研究成果, 并从蜜罐识别与对抗、加密网络流量识别与对抗、去匿名化技术三个方面对网络空间反测绘对抗技术进行探讨。最后, 对当前的研究现状进行总结, 从提升伪装抗识别效能、促进各类防御手段高效协同、增强反测绘智能化三个维度展望了未来研究方向。

关键词 反测绘; 探测; 防护; 欺骗

中图法分类号 TP393 **DOI 号** 10.19363/J.cnki.cn10-1380/tn.2023.08.24

Research on Cyberspace Anti-Surveying and Mapping

LIU Qingyun[1], LI Renjie[1,2], ZHOU Zhou[1], ZHONG Youbing[1,2], SHI Fengyuan[1,2], GUO Li[1,2]

[1] National Engineering Research Centre of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China
[2] School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract The technology of cyberspace mapping has experienced swift growth and widespread adoption. On the one hand, it greatly promotes people's holistic comprehension of various cyberspace resources and their attributes. On the other hand, it introduces substantial ease for adversaries aiming to construct an attack surface map, posing a grave threat to the security of vital information infrastructure. As the ever-changing contest between cyber attacks and safeguards unfolds, Cyberspace Anti-Surveying and Mapping (CASM) plays a pivotal role in erecting a robust security barrier. CASM functions as a comprehensive defensive procedure that protects various cyberspace resources and their attributes from being detected, analyzed, and visualized by adversarie. This approach empowers us to take a more proactive stance in the risk game of cyberspace. The core idea behind CASM is to obstruct the adversaries' detection efforts and prevent them from correlating and analyzing the detected data, thus rendering it incapable for adversaries to construct a dynamic, real-time, and reliable attack surface map, thereby achieving the defensive outcomes known as the “unreachable probe”, “unreal measurement” and “unreliable depiction”. Our initial step is to introduce the pertinent notions of cyberspace mapping technology and then provide related concepts and definitions of CASM from the perspective of prevention, control and early warning. Following this, we propose the technical architecture of CASM and elaborate on key technologies and related research accomplishments of CASM categorized under three levels: “identification of probe behavior”, “protection of probe behavior” and “deception of surveying and mapping”. Additionally, we delve into countermeasures technology of CASM from three aspects: honeypot countermeasures, encrypted network traffic countermeasures, and deanonymization technology. Concluding our discussion, we summarize the current state of research and outline the prospects of future research directions from three dimensions: improving the effectiveness of camouflage, promoting the efficient coordination of diverse defense methods, and enhancing the intelligence of CASM.

Key words anti-surveying and mapping; detection; protection; deception

通讯作者: 周舟, 博士, 高级工程师, Email: zhouzhou@iie.ac.cn。

本课题得到中国科学院战略性先导科技专项(No. XDC02030400)资助。
收稿日期: 2021-04-07; 修改日期: 2021-07-05; 定稿日期: 2023-08-30

1 引言

网络空间是继陆、海、空、天之后的第五大主权领域空间, 已成为国家发展的重要基础。受到地理空间地图的启发, 研究人员正在试图构建一个可以全方位展示网络空间信息的“作战地图”, 并将构建“作战地图”的行为称为“网络空间测绘”。通过网络空间测绘, 研究人员可以得到一个细致全面的“网络空间攻击面地图”。近年来, 在网络空间测绘领域出现了众多的代表性工作, 许多国家在网络空间测绘能力建设方面部署了多个重大项目, 逐步形成对全球资源的探测能力。美国最先进行了国家级的网络空间测绘项目, 例如美国国防部的 X 计划[1]、美国国土局的 SHINE 计划[2]、美国国安局的藏宝图计划[3]。这些计划针对网络空间资源进行了全方位的探测和分析, 构建了全面而深入的网络空间信息数据库。此外, 网络安全研究人员相继开发了 Zmap[4]、Masscan[5]等全网扫描引擎, 为网络空间测绘提供了极大的便利。在工业界也相继出现了 Shodan[6]、Zoomeye[7]、Censys[8]、FOFA[9]等较为成熟的网络空间搜索引擎。

网络空间安全研究是一个攻防双方不断对抗、博弈的过程, 在此过程中, 攻守双方往往是不对等的, 防御方在博弈中往往处于下风。网络空间测绘技术的不断发展进步, 不仅使防御方对网络空间中的各类资源有了更加清晰的了解, 有助于提升自身联网资产的安全性, 还为攻击方提供了极大的便利, 降低了攻击门槛。攻击方不仅可以借助先进的测绘技术对攻击目标进行详细的侦察和分析, 获取攻击目标的开放端口、子域名、系统信息等关键数据, 还可以利用现有的网络空间搜索引擎快速构建全网攻击面, 有效发现和收集攻击资源, 进而扩大攻击范围。为了应对攻击方的测绘行为, 需要站在网络空间测绘的对立面考虑反制措施, 为此, 本文提出了“网络空间反测绘”的概念。利用反测绘技术, 能够有效防范攻击性的测绘行为, 增加攻击成本, 拖延攻击方的行动, 为防御方争取更多的防御时间, 增强防御方的主动性。

面对网络空间测绘技术的迅猛发展, 一些重视相关技术的国家和重要单位开始对网络空间测绘采取谨慎和对抗的态度。为了保护自身信息不被敌对的测绘方获悉, 他们已经开始采取多种反测绘技术手段, 旨在使得测绘方无法通过测绘技术获取到真实、有效的信息, 进而不能得到全面准确的网络空间地图, 实现保护、隐藏己方网络和网络资产的目标。反测绘技术经过长时间的发展, 逐渐从协议伪装、客户端识别、IP 重定向等这些较为被动的防御措施走向更为主动的扰乱或是欺骗。例如发展了几十年的蜜罐技术以及基于蜜罐思想演化而来的各种变体, 再比如近些年来吸引了众多研究人员兴趣的移动目标防御[10](Moving Target Defense, MTD)、拟态防御和网络欺骗[11](Cyber Deception)等主动性更强的防御机制。

研究人员针对上述技术已经开展了相关研究工作, 近些年来也出现了一些有针对性的综述工作。例如, 石乐义等人[12]从发展历史、关键技术、成果分类以及发展趋势等方面对近年来国内外蜜罐领域的研究成果进行了详细梳理; Sengupta 等人[13]从定义、分类、实现方式、评估指标、案例分析等方面对移动目标防御技术进行了系统性的综述; 贾召鹏等人[14]针对网络欺骗技术给出了形式化定义和层次化模型, 总结了发展历程和代表性研究成果, 并介绍了针对欺骗技术的对抗手段以及欺骗技术的发展趋势。此前的综述工作只针对某一特定的技术领域, 但是单一的技术并不能解决所有的网络空间安全问题, 需要这些技术相互关联、相互协作。由于目前在学术界和工业界尚未将这些技术融入到统一的框架中, 未形成对网络空间反测绘相关概念的统一认知, 缺乏对网络空间反测绘技术体系的顶层设计。因此, 本文综合多种防御技术, 围绕网络空间反测绘的概念和技术体系进行研究和探讨, 旨在为网络空间反测绘理论和技术的研究和发展奠定基础。

本文首先阐述了网络空间测绘的相关概念以及现有的测绘技术体系, 进而提出了网络空间反测绘的概念; 然后, 从探测行为识别、探测行为防护、测绘分析欺骗三个维度构建了网络空间反测绘的技术体系, 对这三个技术方向的研究工作进行了详细阐述; 最后, 简要介绍了反测绘对抗技术以及未来的发展方向。

2 网络空间反测绘相关概念

2.1 网络空间测绘定义

网络空间测绘是对网络空间中的各种虚实资源及其属性进行探测、分析和绘制的全过程[15], 其中, 实体资源是网络交换设备、接入设备的集合, 虚拟资源则包括承载在实体资源之上的信息内容、虚拟用户和应用服务。网络空间测绘涉及计算机科学、网络科学、测绘科学、信息科学等领域, 网络探测、网络分析、实体定位、地理测绘和地理信息系统等技术[16], 以期全面掌握网络空间资源的属性和状态,

绘制网络空间资源全息地图。

网络空间测绘分为三个部分: 探测、分析和绘制。通过探测可以获得实体资源和虚拟资源在网络空间的位置、属性、拓扑结构, 并将探测结果作为网络空间测绘的数据基础; 通过分析对探测得到的基础数据进行处理, 设计有效的定位算法和关联分析方法, 将网络空间中的虚实资源映射到地理空间和社会空间; 最后通过绘制将探测和映射的结果绘制成一份动态、实时、可靠的网络空间地图, 以表示网络空间资源的坐标、拓扑、周边环境等信息。

网络空间测绘的研究范围从狭义上主要指互联网环境, 在广义上则包括互联网、电信网、工业控制网等各种类型的网络, 探测对象除互联网资源外, 还包括其他各种网络上的资源[17]。

2.2 网络空间测绘相关技术

网络空间测绘分为探测、分析、绘制三个过程, 本节从探测技术、分析技术、绘制技术三个层次进行阐述。

2.2.1 探测技术

探测是进行网络空间测绘的基础, 根据探测对象的不同, 可以将探测技术分为实体资源探测技术和虚拟资源探测技术。

实体资源探测技术在拓扑层面主要用于网络层拓扑发现, 可以分为 IP 接口级、路由器级、PoP(Point of Presence)级、AS 级四个层次, 具体包括 Sherry 等提出的基于 IP 时间戳选项的路由别名判别方法[18], Spring 等提出的基于 DNS 解析的 PoP 提取算法[19], 以及 Nmap、Zmap、Masscan 等应用广泛的典型扫描工具[20]等。在实际设备层面, 实体资源探测技术则主要用于主动或被动地进行网络设备组件探测识别。其中, 主动探测更有针对性, 能够得到更多、更可靠的信息, 被动探测则更加泛用, 对网络状态的影响较小, 对网络环境的要求也较低, 是目前主流的探测方式。也有研究者提出“协同探测”模式[12], 进行主被动协同、多点协同、协作协同, 以提高探测速度与精度。

在虚拟资源探测技术方面, 主要有内容分析、关联分析和社会信息网络挖掘等文本资源探测技术, 多语言识别、固定音频检索、视频特征表示和语义属性分析等音视频内容探测技术, 以及网络扫描、指纹识别、流量检测等应用和服务探测技术。

2.2.2 分析技术

为了将探测得到的基础数据应用于绘制, 需要对探测结果进行融合分析[16], 即对实体资源和虚拟资源进行属性提取、关联和画像, 并向物理空间和社会空间进行关联映射[13]。

在分析建模方面, 实体资源探测结果分析技术主要包括网络路径重构技术、拓扑分析技术和拓扑语义标注技术, 虚拟资源探测结果分析主要包括特定信息内容快速发现和关联分析技术、特定音频内容的检索与识别技术。在映射方面, 主要将实体资源向地理空间映射, 虚拟资源向社会空间映射。在将实体资源向地理空间映射时, 涉及到网络空间测量与定位技术, 包括实体地标的获取与评估、网络实体定位等。在将虚拟资源向社会空间映射时, 涉及的技术主要有虚拟人物活动地点推断、虚拟群体关系挖掘、虚拟群体轨迹发现等技术。

2.2.3 绘制技术

在将探测结果进行分析映射后, 为了更加形象地对网络空间进行可视化表达, 首先需要建立网络空间时空基准, 并进行地图要素分类与符号设计[16-21], 然后再进行逻辑图绘制、地理信息图绘制乃至全息绘制[13]。

逻辑图绘制技术通过构建拓扑可视化模型将探测得到的网络拓扑可视化; 地理信息图绘制技术则将数据在地理层面上进行可视化表达, 包括地理空间和网络空间的数据同化技术、网络空间信息和地理空间信息的集成可视化技术、网络节点辅助分析技术; 全息绘制则主张进行“叠加绘制”、“时空建模”, 将网络空间多类资源叠加绘制, 并构建虚拟网络空间和实体地理空间统一描述的时空坐标体系。

2.3 网络空间反测绘概念

本文提出了网络空间反测绘的概念, 其中, 反测绘针对的是以发动攻击为目的而进行的测绘行为, 而以防御为目的测绘行为通常是无害且有利于提升自身联网资产的安全性, 因此本文不将这类测绘行为包含在反测绘的范畴之内。本文针对网络空间反测绘的定义给出如下描述。

定义 1. 网络空间反测绘是一种综合利用多种技术手段来保护己方网络空间的各类虚实资源及其属性不被测绘方探测、分析和绘制的全方位防御机制。

具体来说, 网络空间反测绘首先是要通过网络行为分析技术对探测行为进行识别, 进而综合利用多种防护手段来保护网络设备等实体资源以及应用服务、信息内容、用户数据等虚拟资源不被测绘方发现; 其次, 设计并实施有效的欺骗策略, 在设备、网络、应用和数据四个方面对测绘方的探测结果进行伪装与混淆, 误导其对各类虚实资源及其属性的关联分析结果; 最终, 通过对探测行为和分析行为的有效反制, 使得测绘方无法绘制出一个动态、实

时、可靠的网络空间地图。

3 网络空间反测绘关键技术研究进展

网络空间反测绘作为与网络空间测绘的博弈方, 目的是针对网络测绘行为进行识别发现, 并采取特定的措施对测绘行为进行阻断或欺骗, 使得测绘难以进行, 让测绘分析结果偏离真实环境。对应网络空间测绘技术体系的层级划分, 反测绘技术应当首先针对测绘技术的基础, 即探测行为, 进行反制; 其次针对测绘技术体系中的分析过程采取有针对性的欺骗等防御手段; 最终通过对探测和分析行为的遏制, 使得测绘方无法绘制出真实有效的网络空间地图。因此, 本文将基于以下三个方面, 构建网络空间反测绘技术体系: 一是探测行为识别技术, 二是探测行为防护技术, 三是测绘分析欺骗技术。

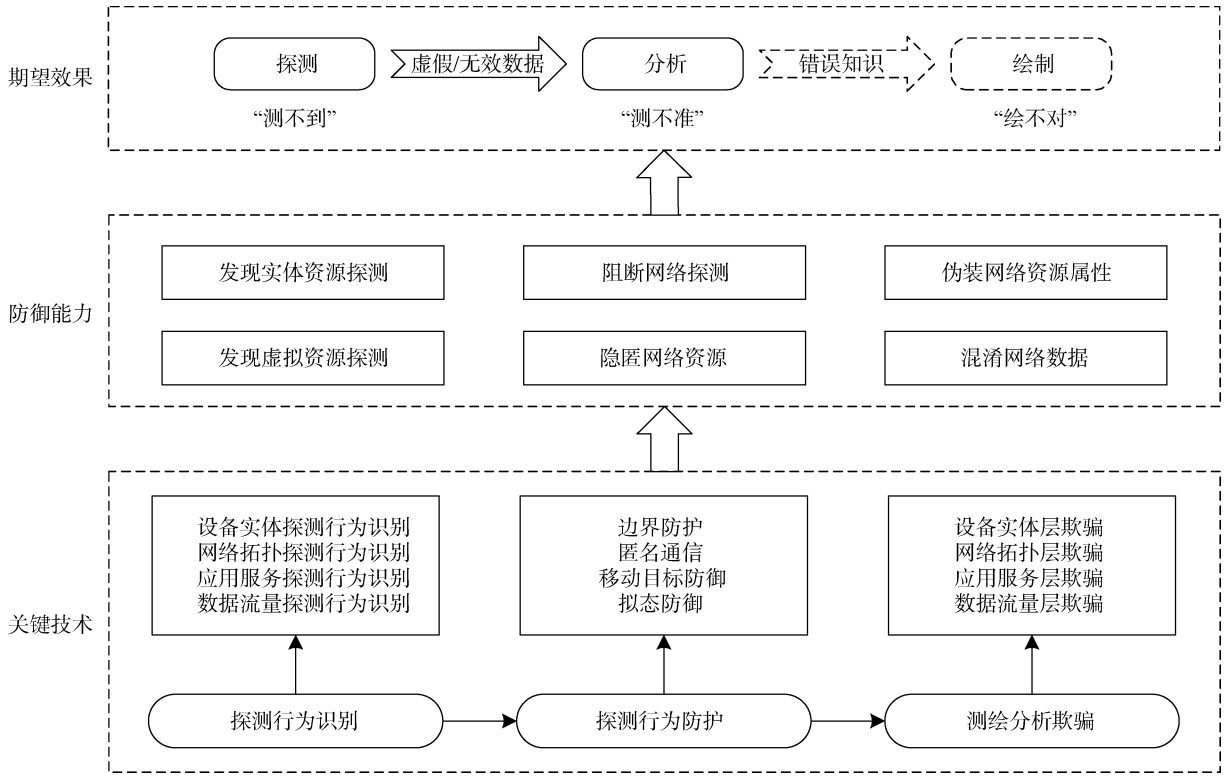


图 1 网络空间反测绘技术体系

Figure 1 Technical Architecture of Technology of Cyberspace Anti-Surveying and Mapping

3.1 网络空间反测绘技术体系

网络空间反测绘技术体系如图 1 所示, 表 1 对网络空间反测绘技术体系各阶段的关键技术进行了总结。为了实现反测绘, 首先要识别探测行为, 进而通过多种防护手段阻断探测行为, 让测绘方“测不到”。但是, 测绘方的探测行为并不是完全可防可监测的, 因此就需要考虑让测绘方“测不准”, 即对测绘技术体系中的探测结果分析这一关键步骤进行有针对性的混淆或欺骗。以“测不到”和“测不准”为基础, 最终使测绘方绘制的网络空间地图与真实的网络环境不符, 达到“绘不对”的效果。

作为反测绘的第一阶段, “测不到”旨在使测绘方无法发现我方的各类网络空间资源, 减少对外暴露的攻击面, 提升自身安全性。为了实现“测不到”的目标, 首要任务就是对测绘方的探测行为进行识别, 发现测绘方的探测意图和目标。按照被探测资源所在的层次, 本文将探测行为识别划分为三大类: 设备探测行为识别、网络拓扑探测行为识别以及虚拟资源探测行为识别; 其中, 虚拟资源探测又被细分为应用服务探测和数据流量探测两类。利用多种探测识别技术, 可以为抵御针对各类虚实资源的探测行为提供防御基础。如何准确而有效地识别探测行为是此阶段需要解决的关键问题, 不仅要充分了解各类探测行为的探测原理和行为模式, 还需要对其产生的网络流量进行细致的特征分析。

在识别探测行为的基础上, 防御方通过边界防护技术、匿名通信技术、移动目标防御技术、拟态防御技术等防护手段, 获得阻断探测行为和隐匿网络空间资源等防御能力, 达到让测绘方“测不到”的效果。在阻断探测行为阶段, 基于边界的防护往往很

难拦截所有的探测行为, 且存在误拦截的情况; 匿名通信技术则需要不断加强自身的机密性和安全性, 应对各种主/被动攻击, 尤其是近些年来机器学习或深度学习技术在用户分析和流量识别方面取得了不错的效果, 为匿名通信带来了不小的挑战; 移动目标防御强调目标的可变性, 且需要变化是可管理、可持续、快速、多样且难以预测的, 存在着跳变多样性不足、跳变同步依赖第三方实体等问题; 拟态防御技术则要设计多变体执行架构, 面临着执行效率、协同管理、安全性和透明性等关键问题。

除了对探测行为进行识别与防护外, 更进一步地, 需要具备让测绘方“测不准”的能力。作为反测绘的第二阶段, “测不准”旨在主动对测绘方实施干扰、混淆和欺骗, 使得测绘方即使能够探测到我方目标, 也无法准确判断目标属性和弱点。在该阶段, 需要针对测绘分析阶段采取更主动的反制措施, 通过蜜罐/蜜网、流量加密混淆和匿名化等技术对测绘分析实施欺骗, 实现伪装网络资源属性和混淆网络数据等防御能力, 使测绘方得到虚假或无效的探测结果, 进而误导攻击方的判断, 使其不能正确地对探测结果进行建模和关联分析, 无法分析出真实的网络空间资源属性。有效的伪装和混淆技术是欺骗能够成功的关键, 而目前的欺骗技术研究还处于相对初级阶段, 欺骗的动态化、层次化、智能化等方面仍有较大的提升空间, 在应对高级持续威胁时的防御效果也有待提高。

最后, 基于“测不到”和“测不准”这两阶段的防御, 便可实现使攻击方“绘不对”的最终目标。攻击方不但无法成功实施探测, 而且探测所得为虚假信息, 因此, 其最终无法获取目标网络的真实属性, 无法掌握真实有效的攻击面地图。

表 1 反测绘各阶段及其关键技术

Table 1 States and Key Technologies of Cyberspace Anti-Surveying and Mapping

反测绘阶段	关键技术
探测行为识别	探测工具指纹识别, 流量特征分析, 日志分析, 访问序列分析, 行为建模
探测行为防护	访问控制, 入侵检测与防御, 匿名通信, 移动目标防御, 拟态防御
测绘分析欺骗	操作系统混淆, 蜜罐, 蜜网, 诱饵路由, 诱饵链接, 诱饵漏洞, 流量加密混淆, 匿名化技术

3.2 探测行为识别技术

网络空间测绘中, 网络探测是指对网络空间的各类资源及其属性进行探测。网络探测方式多种多样, 根据探测基础数据来源不同, 网络探测行为可以分为主动探测行为、被动探测行为和基于网络空间搜索引擎的探测行为。主动探测行为是指通过主动向目标发送探测数据包, 从响应的数据包中提取可以标识目标状态的数据信息, 如主机发现、端口扫描等; 被动探测是指通过网络设备采集目标网络流量, 分析流量中的特殊字段或指纹特征, 提取关键信息, 实现目标的探测; 基于网络空间搜索引擎的探测行为是指利用网络空间搜索引擎进行目标探测以及信息挖掘。根据探测目标不同, 探测行为可以分为设备探测、网络拓扑探测和虚拟资源探测, 虚拟资源探测可以进一步分为应用服务探测和数据流量探测。

网络探测行为的识别是从反测绘方的角度出发, 发现目标正在遭受的探测扫描, 识别探测行为。根据探测目标划分, 本文将网络探测行为识别分为设备实体探测行为识别、网络拓扑探测行为识别、应用服务探测行为识别和数据流量探测行为识别。

3.2.1 设备实体探测行为识别

随着互联网及物联网的快速发展, 联网设备越来越多, 设备探测的方式也变化多样。了解设备探测方式及原理对于识别该类探测行为至关重要。目前, 最为主流的设备探测技术依然是基于指纹实现的, 如 Shodan、Zoomeye、Censys 以及各类漏洞扫描产品[22], 基本原理是对目标设备进行网络扫描, 根据响应包中的多个字段信息构造设备指纹, 如开放端口号、会话方式及协议等, 同指纹库进行匹配, 确定探测的设备类别。因此, 通过识别指纹探测行为, 可以有效识别相应设备的探测行为。

为了细粒度地识别设备探测行为, 可根据设备指纹库将设备分类, 基于不同类别设备的探测方法的差异性, 识别探测数据包。根据网络空间大多数设备已有相应的分类命名规则, 确定每类设备的网络扫描方法, 识别网络扫描数据包的格式, 提取特征, 归纳聚合, 建立设备探测行为指纹库。扫描程序发出的请求数据包和常规请求数据包之间往往存在内容上的差异, 这些差异可以被用作识别探测行为的指纹。例如, Nmap 在进行操作系统扫描时, 发送的 UDP 数据包中默认使用字符“C”填充数据部分, 而在 ICMP 数据包中使用“0”填充, 同时, 扫描时默认的 TCP 窗口大小字段固定为“1024”; Zmap 在扫描时使用固定的 IPID“54321”以及固定的窗口大小“65535”。针对部分未知类别的设备, 通过深入分析设备探测行为的模式, 将其与已知类别设备的探测行为模式相比较, 采用相似性度量方法, 将其归类

为最相似的类别, 从而识别针对该设备的探测行为。

设备实体探测行为的识别通常依赖于探测工具指纹的有效性, 然而由于一些探测工具的开源特性, 其指纹信息可以随意修改, 使得已有的指纹信息失效, 影响识别系统的稳定性和有效性。而相似性度量等流量特征分析技术不确定性较大, 误报和漏报情况较为突出。

3.2.2 网络拓扑探测行为识别

随着网络拓扑规模的日益扩大和结构的日益复杂, 网络拓扑探测对网络空间安全的影响越来越大。一方面, 对于网络运维人员, 掌握网络的拓扑结构, 了解节点设备信息, 可以更好地诊断网络故障。另一方面, 对于攻击方而言, 获得目标网络的拓扑信息可以更有指导性地实施入侵行为, 因此识别网络拓扑探测行为具有重要的意义。

网络拓扑探测是为了发现网络拓扑结构。常用的拓扑探测方法有类 Traceroute 的网络拓扑发现方法[23], 该类方法依赖于探测程序返回的路径信息, 如常用的 Traceroute 工具。此外, 还有基于 ARP、SNMP、OSPF 等协议的拓扑发现方法。然而这些方法存在一些不足, 例如, 基于SNMP的探测受限于路由设备访问权限, 无法适用于大规模的网络拓扑探测; 采用 ARP 的拓扑探测方法只能用于局域网中。针对单个路由器具有多个 IP 地址的情况, 可以采用别名解析和递归路由器探测技术进行拓扑发现。

网络拓扑探测是根据回传给探测源的报文发现目标拓扑结构。相应地, 通过提取发送的探测报文和回传报文的特征, 如协议类型和数据包格式, 识别拓扑探测行为。例如, 在使用 Traceroute 程序进行拓扑路径探测时, 程序发出的初始 ICMP 数据包中的TTL 值为 1, 此后每收到一个来自路径上路由器的响应就将后续探测包中的TTL值加1, 直到被探测的目标节点收到 TTL 值为 1 的 ICMP 数据包并向探测源节点回复 ICMP“端口不可达”差错报文后结束探测。在探测过程中, 路径上的同一路由器所看到的 ICMP 数据包中的 TTL 值是由 1 开始递增的。因此, 通过分析 Traceroute 程序发出的ICMP包序列特征及其收到的ICMP超时或不可达报文, 即可识别该拓扑探测行为。对于多节点协同探测的探测方式, 需要采用多点协同关联分析的手段来识别探测行为。

网络拓扑探测行为识别面临着网络拓扑、网络配置等属性动态变化, 以及网络监测节点难以全面覆盖的情况, 在探测报文和回传报文捕获时存在不准确、不全面的问题, 导致识别准确性和稳定性不足。

3.2.3 应用服务探测行为识别

应用服务是网络空间虚拟资源的重要组成部分, 针对应用服务的探测方式多种多样, 有端口扫描、漏洞扫描等。端口扫描是向目标系统发送探测数据包, 旨在探测网络主机是否存在开放端口和其他可用服务; 漏洞扫描是在端口扫描的基础上, 将扫描结果同漏洞库数据进行匹配, 进一步识别应用服务类型、版本等信息。

为了识别应用服务探测行为, 首先应分析系统端口的开放状态, 并且记录探测方访问应用服务时留下的信息, 进而通过分析探测数据包、提取报文特征、建立数学模型、数据可视化等方式, 并结合统计学方法、人工智能算法或其他启发式方法识别探测行为。Bhuyuan 等人[24]将端口扫描分为单源扫描和分布式扫描两大类, 并详细介绍了如何识别这两类扫描行为; Bou-Harb 等人[25]则专注于分布式扫描的识别, 从统计学方法、算法方法、数学方法和启发式方法四个方面进行了详细的阐述。为了识别应用层的探测行为, 我们可以通过 HTTP 的状态码、URL 长度、User-Agent 字段等特征来区分探测程序发出的请求和正常请求。例如, Web 漏洞扫描器 Awvs(Acunetix Web Vulnerability Scanner)在请求的 URL、Headers、Body 等字段包含了“by_wvs”、“acunetix_wvs_security_test”等能代表自己的特征字符串; Sqlmap 在 User-Agent 中包含了“sqlmap”这个特征字符串。

与设备实体探测行为的识别类似, 应用服务探测行为识别依赖于对扫描工具的指纹构建, 但通常有经验的攻击方会基于开源的扫描工具进行二次开发, 刻意改变默认字段、请求负载等指纹特征, 从而使得原有的指纹信息失效。扫描工具的行为模式、流量统计信息等特征也在时刻变化, 加大了统计学方法、人工智能算法或其他启发式方法的识别难度。

3.2.4 数据流量探测行为识别

在网络空间中, 除了应用服务之外, 虚拟资源还包括用户信息、文本/音视频等数据和网络流量, 本文将其统称为数据流量。针对这些虚拟资源的探测手段主要有数据爬取和被动流量分析等。数据爬取主要是采用爬虫的方式主动爬取公开的网页信息, 如用户账号信息及社交媒体信息等可获取的数据内容; 被动流量分析则是以镜像等方式采集目标系统的流量, 综合使用模式匹配、数据建模、关联分析等技术识别虚拟资源。

针对数据爬取类探测行为的识别技术可分为四类: 基于访问日志的分析识别[26]、基于访问模式的分析识别[27-28]、基于访问行为的分析识别[29-30]和基于

访问序列的分析识别[31]。现有的网络爬虫技术变换多样, 模仿合法请求的能力较强, 基于访问日志的分析识别很容易被绕过, 进而失去识别能力; 基于访问模式的分析识别依赖于访问频率、规律等信息, 检测手段过于粗糙, 存在较多的漏报; 基于访问行为的分析识别依赖于可靠的特征属性构建和有效的分类算法, 需要较多的专家知识; 基于访问序列的分析识别属于静态检测机制, 面对不断更新迭代的数据爬取技术难以达到预期效果。而对于被动流量分析, 由于其隐蔽性强且不存在与系统的交互行为, 往往难以察觉, 缺乏有效的识别方式。

3.3 探测行为防护技术

探测行为防护是反测绘的第二阶段, 在该阶段防御方需要同时具备阻断测绘方的探测行为和隐匿己方网络空间资源的能力。一方面, 使测绘方的探测受阻, 难以获取到分析建模阶段所需的数据, 另一方面, 使测绘方仅能够获取到动态、随机、多样的无效数据, 进而, 使得测绘方陷入数据不足或数据无效的困境, 使其无法实现网络路径重构、拓扑分析、资源映射等目标。本节将可以用来实现探测行为防护的主要技术概括为四个方面: 基于边界的防护、基于匿名通信技术的防护、基于移动目标防御技术的防护以及基于拟态防御技术的防护。

3.3.1 基于边界的防护

测绘方与反测绘方可以看作是攻防的两端, 测绘方作为攻击端可以通过多种方式对防御方的基础设施、网络服务等关键资源进行探测和分析, 旨在构建全面的资源数据库, 实现精细化的画像。因此, 作为防守侧的反测绘方可以将测绘方的探测行为看作是一种异常行为, 进而可以借鉴防火墙、入侵检测与防御等基于边界的防护技术, 有效发现并阻断探测行为。

1) 访问控制

为了不让测绘方探测到, 反测绘方就需要尽量减少其网络对外的暴露面。最简单的方式就是对外部网络的访问实施严格的访问控制, 把住进入自身网络的必经通道。访问控制技术可以有效地阻止外来威胁进入内部系统。防火墙作为一种最典型的访问控制形式, 可以将自身网络和外部网络强制隔离, 其安全设计原理是基于包过滤与应用代理技术。防火墙可以通过配置访问控制列表(ACL)控制数据流的流入和流出, 同时结合 NAT/PAT 技术隐藏内网设备的 IP 地址。通过使用多种防火墙技术, 可以增加测绘方的探测难度, 进而延缓其测绘进度。

2) 入侵检测与防御

随着测绘方可利用的技术手段逐渐多样化, 网络探测趋于规模化、复杂化, 探测的精确性和有效性得到了极大提升, 给防御方带来了越来越多的挑战, 单独使用传统的防火墙技术已无法有效应对如此规模化、复杂化的威胁, 还需要结合入侵检测(Intrusion Detection)与入侵防御(Intrusion Protection)技术进行有效应对。入侵检测系统(Intrusion Detection System, IDS)旨在识别传统防火墙无法识别的恶意流量和非授权的访问, 对于充分保证网络的机密性、完整性和可用性至关重要。根据其识别恶意行为的原理的不同, IDS 可以分为两大类: 基于签名的入侵检测系统(Signature-based Intrusion Detection Systems, SIDS)和基于异常的入侵检测系统(Anomaly-based Intrusion Detection Systems, AIDS)。其中 SIDS 在一些文献中也被称作是基于知识的入侵检测系统或是基于误用的入侵检测系统[32]。Snort[33]和 NetSTAT[34]是较为典型的 SIDS。根据识别异常行为的模式, AIDS 又进一步可以分为三大类: 基于统计的、基于知识的以及基于机器学习/深度学习的。

入侵防御系统(Intrusion Protection System, IPS)是一种主动保护系统, 是 IDS 的进一步发展。IPS 能够监视网络或网络设备间的数据传输, 即时中断、调整或隔离一些不正常或是具有伤害性的行为。

基于边界的防护在阻断探测行为方面虽然有着不错的效果, 但是仍有不足之处。在制定访问控制策略、确定网络隔离的力度等方面仍然需要较多的人为干预, 通常依赖于已有的经验, 自动化、智能化不足; IDS 和 IPS 则难以适用于所有的网络环境, 其鲁棒性不断受到变化多端的逃避技术的挑战, 且在面对数据加密的情况时也难以保证可靠性。

3.3.2 基于匿名通信技术的防护

为了充分地应对测绘方的探测行为, 可以考虑使用匿名通信技术作为防御手段, 通过隐匿通信双方的真实地址、路由等信息, 让测绘方无法探测到通信的实体、内容和属性。此外, 匿名网络中的资源无法通过普通的网络空间搜索引擎进行检索, 使得测绘方对于掌握全面、准确的网络空间资源信息显得无能为力。

根据传输方式的不同, 可以将匿名通信技术分为两大类: 基于广播/多播的匿名通信(DCnets[35]、Horders[36]等)和基于重路由(rerouting)的匿名通信(Mix-nets[37]、Onion Routing[38-39]、Tor[40]等)。在基于广播/多播的匿名通信中, 发送方向接收方发送消息时会将数据同时转发到多个或者所有的节点, 而不是显式地做出特定的路由决策。在基于重路由的匿名通信中, 根据消息发送方的行为, 又可以分为

source-routed 和 hop-by-hop 两种模式。source-routed 模式允许发送方预先指定想要经过的所有节点的集合, 而 hop-by-hop 模式只需要发送方指定第一跳的节点, 后续路由则由各个中间节点自己进行决策。这两种匿名通信方式使得探测方无法获取到通信的实体信息以及链路信息。

DCnets 是由 Chaum[35]在 1988 年首次提出的概念, 其基于安全多方计算等密码学和信息论机制保证匿名性。Horders[36]是第一个利用多播路由匿名接收数据的协议。Mix-nets 的概念最早被 Chaum 应用于提供匿名电子邮件服务。Mix-nets 通过引入一种叫做 mixes 的代理路由数据包, 结合公钥加密算法, 在数据传输路径上的每个 mixes 之间对数据包进行加密, 使得测绘方很难将发送方和接收方关联起来。Onion Routing 由一组称为洋葱路由器(Onion Router, OR)的节点组成, 是一种由几个相互连接的 OR 组成的基于 TCP 的覆盖网络(Overlay Network), 通过每个 OR 层层加密的方式中继用户数据, 每个 OR 只能与其前一跳和后一跳进行加密通信, 因此只有第一跳(守卫节点)知道真实的用户信息, 最后一跳(退出节点)知道真实的目的地址。

Tor 是对 Onion Routing 的一种成功的扩展, 提供了更好的安全性、更高的效率以及更简单的部署方式, 同时其完全开源且免费。作为第二代 Onion Routing, Tor 是最著名和最广泛使用的低延迟匿名网络。隐藏服务(Hidden Service, HS)是基于 Tor 网络的主要的应用之一。隐藏服务是 Tor 网络中的一些特殊的主机提供的服务, 这些主机可以在完全匿名(即保持其 IP 地址隐藏)的情况下, 向其他主机提供某种服务, 且保持客户端和服务端的完全匿名。由于这些隐藏服务并不会直接暴露在公共互联网上, 因此给测绘方的探测带来了极大的困难, 测绘方也无法轻易获取到服务端地址等关键信息。

近些年, 还出现了基于 P2P(Peer-to-Peer)网络的匿名通信技术, 如 Freenet[41-42]和 I2P[43-44], 此类技术提供了更高的弹性和可扩展性, 进一步增大了测绘方的探测难度。

匿名网络和匿名通信技术的应用给网络空间测绘带来了巨大的挑战, 虽然提供了较好的匿名性, 隐藏了通信实体、路由等信息, 但是仍然面临着被各类去匿名化技术攻破的风险, 主要体现在网络流量分析和针对通信节点的攻击两方面[45], 本文将在第 4 节对这两类问题展开探讨。为了应对各类威胁, 研究人员也提出了多种安全增强技术, 如流量整形[46-47]、流量填充[48-49]、背景流量[50-51]等技术。

3.3.3 基于移动目标防御技术的防护

移动目标防御[10]是美国针对防御方目前处在劣势地位这样的实际情况而提出的一种动态防御技术。《Cybersecurity game-change research and development recommendations》[52]一文中将移动目标防御的内涵描述为: “期望能够创建、分析、评估和部署多样化的、随时间持续变化的机制和策略, 以增加攻击方实施攻击的复杂度和成本, 降低系统脆弱性曝光和被攻击的几率, 提高系统的弹性”。

从上述定义可以推断出, 在防御过程中也可以利用 MTD 技术, 构建一种动态的、多样化的、不断变化的网络空间目标环境, 以系统的随机性和不可预测性对抗测绘方的探测行为。MTD 可以在多个层面上部署应用, 包括网络层、运行环境层和软件层, 表 2 给出了各层的一些代表性工作。

1) 基于通信网络变换的防护

通信网络的变换主要是对 IP 地址和端口号进行变换, 这种方式使探测方捕捉到的 IP 地址和端口号随时间而失效, 导致无法再被探测到。

动态网络地址及端口变换。Kewley 等人[53]提出的动态网络地址转换(Dynamic Network Address Translation, DYNAT)是一种能够对 IP 数据包和 TCP 数据包头进行变换的机制。通过一种基于时间的机制, 周期性地变换 IP 数据包头中目的地址的主机地址和 TCP 数据包中的目的端口。具体过程为: 在客户端使用 TCP/IP DYNAT shim 软件, 进行主机地址和目的端口的变换, 由于目的网络地址没有变换, 数据包可以到达目的网关, 在目的网关处, 使用相应的翻译方法对变换过的主机地址和目的端口进行还原。Antonatos 等人[54]提出的网络地址随机化(Network Address Space Randomization, NASR)是一种针对蠕虫的局域网地址跳变技术, 在探测方已经收集到目标的主机的地址信息但还未开始采取进一步行动时, 防御方对所有主机在网络地址空间中进行地址随机化, 赋予每个主机一个地址空间中的随机地址, 使得探测方已经确定的探测目标的地址失效, 进而阻断蠕虫的成功释放。Lee 等人[55]提出一种周期性改变应用服务使用的 UDP/TCP 端口号的端口跳变技术(Port Hopping), 使得测绘方无法获得应用服务使用的端口号。Sifalakis 等人[56]借鉴了扩频无线电通信的思想, 提出了网络地址跳变技术(Network Address Hopping), 通过建立多个不同的信道将通信数据隐藏在多条不同的路由中。胡毅勋等人[58]提出了一种基于 OpenFlow 的网络层移动目标防御方案, 该方案通过在网络域内的 OpenFlow 交换机实现了

表 2 移动目标防御技术在网络空间反测绘技术体系中的应用

Table 2 Application of MTD in Technical Architecture of Cyberspace Anti-Surveying and Mapping

层级	参考文献及年份	技术要点	防御单元
网络层	Kewley 等[53], 2001	动态网络地址转换	网关
	Antonatos 等[54], 2007	网络地址随机化	主机
	Lee 等[55], 2004	端口跳变技术	主机
	Sifalakis 等[56], 2005	网络地址跳变技术	主机
	Atighetchi 等[57], 2003	IP 和端口号的变换	主机
	胡毅勋等[58], 2017	基于 OpenFlow 的 IP 和端口跳变	OpenFlow 交换机
	Sharma 等[59], 2018	IP 多路复用	主机
	Narantuya 等[60], 2019	多 SDN 控制器	主机
	Dunlop 等[61], 2011	IPV6 下的地址旋转	主机
	Hao 等[62], 2020	IPV6 下的无地址公共服务器	服务器
运行环境层	Al-Shaer[63], 2011	生成虚拟网络层	网络
	Bangalore 等[64], 2009	自清洗入侵容忍	虚拟机
	Nguyen 等[65], 2011	云环境下的自清洗入侵容忍	虚拟机
	Huang 等[66], 2011	移动攻击面	虚拟机
	Ahmed 等[67], 2016	分布式系统下的变换	虚拟机
	Alavizadeh 等[68], 2018	云环境下的转换和多样性	虚拟机属性
	Carter 等[69], 2014	操作系统变换	操作系统
	Thompson 等[70], 2014	操作系统变换	操作系统
软件应用层	Roeder 等[71], 2010	代码变换	代码
	Jangda 等[72], 2015	代码变换	代码
	Mahmood 等[73], 2016	代码动态加载	应用服务
	Thompson 等[74], 2016	流量动态重定向	应用服务
	Christodorescu 等[75], 2011	多种随机化方法	应用服务
	Vikram 等人[76], 2013	HTML 元素随机化	HTML 表单控件

IP 跳变技术,通过域边缘的 OpenFlow 交换机充当网关在域间实现了端口跳变技术。APOD(Application that Participate in their Own Defense)[57]项目主要是在用户主机内和目的端口所在的网关处实现对 IP 和端口号的周期性变换, 使得探测方截获到数据包后获得的 IP 地址和端口号仅在当前周期内有效。Sharma 等人[59]提出了一种基于 SDN 环境的随机虚拟 IP 多路复用技术(Flexible Random Virtual IP Multiplexing, FRVM), 使主机可以具有多个随机的、随时间变化的虚拟 IP 地址, 通过多路复用和解复用实现主机的虚拟地址与真实地址的动态映射。Narantuya 等人[60]则提出了基于多个 SDN 控制器的 MTD 架构, 可以与 FRVM 等技术结合实现更加灵活、高效、安全的 IP 地址变换。

IPv6。移动目标 IPv6 防御(MT6D)[61]是在 IPv6 环境下实现的移动目标防御。此方法通过使用 IPv6 庞大的地址空间, 在两台主机进行会话的期间进行地址旋转。MT6D 方法中的动态地址变换为 IPv6 环境下的反测绘提供了一种较好的解决方案。Hao 等人[62]所提出的无地址公共服务器(Addressless Public Server)方案中, 引入了 IPv6 地址加密变换, 能够有效防御针对服务器的网络端口扫描。

可变网络。可变网络(Mutable Network, MUTE)[63]是一种能够生成虚拟网络层, 并将其覆盖到现有网络层之上的方法。在这样的虚拟网络层中, 网络主机会获得动态生成的虚拟 IP 地址、端口号和目的路由, 对于这些数据, 还会通过加密的方式对其进行保密, 能够对探测方的探测、识别行为进行有效的防护。

基于通信网络变换的防护虽然可以达到隐藏真实 IP 和端口、实现地址跳变等防御目标, 但是在抵御主动攻击、跳变频率和用户体验等方面仍存在着一些不足。例如, DYNAT 虽然可以达到隐藏真实 IP 地址的目的, 但是其无法抵御主动扫描攻击。NASR 在地址的跳变速率和变化随机性上存在一定的局限性, 且对终端来说是不透明的。FRVM 虽然可以抵御主动扫描和嗅探, 但是并没有充分保证合法用户的服务质量, 高速、动态的地址变换也存在导致通信中断的情况。此外, 现有的基于 SDN 技术的地址跳变技术往往是在单一的 SDN 控制器下实现的, 在安全性和性能方面存在单点故障的情况, 在面对大型网

络环境时可扩展性不足。

2) 基于执行环境变换的防护

执行环境的变换主要指对程序运行时所在系统的软硬件、操作系统、配置文件等进行变换, 使得测绘方难以探测到系统执行环境的特征。

虚拟机变换。Bangalore 等人[64]提出的自清洗入侵容忍(Self-Cleansing Intrusion Tolerance, SCIT)使用虚拟机技术对一个服务器的正常初始状态生成多个虚拟服务器, 每一个虚拟服务器在互联网上提供很短时间的服务(通常不到一分钟)后就会脱机, 然后被另一个拥有初始状态的虚拟服务器替代。Nguyen 等人[65]将 SCIT 技术扩展到云环境中, 用来增强在云环境中部署的应用程序和服务的安全性。移动攻击面(Moving Attack Surfaces, MAS)[66]同样使用了虚拟机技术生成多个虚拟 Web 服务器, 为不同的虚拟 Web 服务器赋予了不同的软件组合, 这就导致了虚拟服务器进行替换时, 攻击面也发生改变。由于测绘方面对的是一个不断变化的虚拟环境, 因此极大增加了探测难度。Ahmed 等人[67]提出的 Mayflies 是一种应用于基于虚拟云平台的分布式系统的 MTD 技术, 它以特定时间间隔为周期不断改变分布式系统中的虚拟机配置, 实现了分布式系统的基础结构的不断变换, 单个虚拟机仅在较短的时间间隔内运行, 之后会被具有不同特征的新虚拟机替换。Alavizadeh 等人[68]提出了一种在云环境上使用转换、多样性以及两种方法相结合的 MTD 方案。转换是指不断变换承载虚拟机的服务器, 多样性是指不断变换虚拟机的软硬件属性, 与 Mayfles 类似, 该方案也可以较好地防御针对云环境的探测和攻击。

操作系统变换。Carter 等人[69]设计了一个可以在不同操作系统之间切换的 MTD 系统, 并基于博弈论分析了操作系统的迁移模式, 得到了最优迁移策略。Thompson 等人[70]提出了多操作系统旋转的 MTD(Multiple Operating System Rotational Environment MTD, MORE MTD), 在固定应用服务 IP 地址并保证服务可靠性的前提下, 为其随机分配不同配置的操作系统, 这种方案可以使测绘方得到不断变化操作系统信息, 且针对单一操作系统只有很小的探测时间窗口。

基于执行环境的变换技术虽然具备多样性、随机性、冗余性等优势, 但是这些特性在增强安全性的同时, 也会为防御带来附加成本。过度频繁的变换对性能提出了不小的挑战, 关系到系统的稳定性, 可能会导致服务中断的情况出现, 降低系统的可用性, 影响用户体验。为了提升系统可用性, 往往还会创建虚拟机、操作系统的副本或是提供相同功能的组件来提升系统冗余性, 这也会带来额外的开销, 甚至进一步扩大系统的暴露面。

3) 基于软件应用变换的防护

软件应用变换主要是对软件的代码、接口等要素进行变换, 使测绘方难以探测到运行的软件应用的属性。

代码变换与动态加载。Roeder 等人[71]提出的主动混淆(Proactive obfuscation)技术对一个可执行文件生成多个副本, 每一个副本提供相同的服务, 副本代码却各不相同, 通过编译、加载或运行时执行等方式运行不同副本, 大大增加了服务的多样性, 这种多样性大大增加了测绘方所需的工作量, 导致测绘方无法准确识别正在使用的软件副本。Jangda 等人[72]提出了一种使用 Java 字节码实时编译器(Just-In-Time compiler)对代码进行变换的方法, 通过在代码中随机插入 NOP 指令实现安全性和多样化。Mahmood 等人[73]提出了一种在物联网中实现 MTD 的方案, 该方案中各个物联网设备从云端下载并运行随机化后的代码, 运行结束后就直接删除下载的代码。由于应用程序多样且动态变化, 测绘方很难判断物联网设备上正在运行的是哪一个应用程序。

流量动态重定向。Thompson 等人[74]将 MORE MTD 扩展到 Web 应用程序, 提出了应用程序动态旋转(Dynamic Application Rotation Environment, DARE)的 MTD 方案。使用两个最常见且免费 Web 服务器(Apache 和 Nginx)运行同一应用程序, 以随机的时间间隔将访问流量重定向到不同的服务器上。与 MORE MTD 类似, 这种变换策略增加了测绘方的探测成本, 并降低了被漏洞利用的可能性。

随机化。端到端软件多样化(End-to-end software diversification)[75]组合多种随机化方法, 通过对指令集、应用程序编程接口(API)、代码组件等进行随机化, 增加系统复杂度, 让探测方更难以确定系统的准确信息。Vikram 等人[76]提出了一种应对网络爬虫的软件变换防御方式(NOMAD), 对 HTML 表单中控件的 name/id 参数值进行随机化, 令网络爬虫无法获得正确的 name/id 值, 从而避免其对服务器大规模的探测请求。

与前文所述的两类变换技术类似, 基于软件应用变换的防护技术具有多样性、随机化和冗余性等优势, 在一定程度上有助于隐藏软件应用的各类属性, 但是在实现这些特性时也存在着增加防御成本和提升系统复杂度等问题, 需要在安全性和可用性之间寻求平衡。

3.3.4 基于拟态防御技术的防护

针对网络安全的不平衡态势及本源问题, 邬江兴院士提出了网络空间拟态防御(Cyber Mimic Defense, CMD)思想[77-80], 旨在为解决网络空间不同领域相关应用层次上的未知漏洞、后门或病毒木马等不确定性威胁, 提供具有普适性的创新防御理论和方法。拟态防御是以动态异构冗余(Dynamic Heterogeneous Redundancy, DHR)形态的广义鲁棒控制架构为基础, 引入了多模裁决、策略调度、负反馈控制、多维动态重构以及输入与输出代理等多种机制, 进而提供高可靠、高可用、高可信的融合式防御功能。与移动目标防御思想类似, 拟态防御也会强调多样性、动态性、随机性, 但拟态防御的思想相比移动目标防御更加广泛, 功能也更为全面。拟态防御作为一种全新的防御机制, 获得了国内研究人员的广泛关注[81-91]。

在反测绘体系当中, 引入拟态防御技术对探测行为防护有着很大的帮助。拟态防御技术提供了很强的动态性, 不仅大大增加了测绘方的探测难度, 还进一步增强了被探测系统的不确定性; 此外, 利用拟态防御技术还可以隐匿系统的内部结构和特性, 减少网络空间资源的暴露面; 同时, 拟态防御技术通过引入多执行体机制实现异构冗余, 可以使探测得到的反馈信息被扰乱, 这不仅有助于减小探测的成功率, 还有利于更进一步地防御测绘分析。

马海龙等人[81]基于拟态防御技术, 在路由器体系架构上引入异构冗余功能执行体, 通过动态调度机制实现了主动防御。实际测试结果表明, 在防范针对系统信息的扫描探测方面, 仅凭借一两次扫描探测难以准确识别目标系统的相关信息, 即使增加了扫描探测的次数和频度, 也不一定能够获得准确的目标系统信息。因此, 该系统可以明显降低探测者获取目标系统信息的准确度, 大大增加探测花费的时间, 显著提升探测者获取目标系统信息的难度; 在防范针对系统漏洞的扫描探测方面, 动态异构冗余机制可以在不消除系统漏洞或脆弱点前提下改变其特性, 使探测者很难在短时间内探测出目标系统的漏洞信息。

仝青等人[82]针对 Web 服务器系统的安全问题, 提出了基于拟态防御模型的拟态防御 Web 服务器, 在文件层、SQL 脚本层、服务器软件层、虚拟机操作系统层以及物理机操作系统层这五个层次实现了拟态防御模型。使用 Nikto、Nmap 等工具对该 Web 服务器进行了多次扫描和探测, 结果表明该拟态 Web 服务器应对扫描探测行为时能够变换系统的指纹信息, 呈现不确定性。

魏帅等人[83]以拟态技术为基础, 提出了面向工控领域的拟态安全处理机架构, 利用三余度异构冗余处理机作为设计基础, 仿真验证结果验证了该系统可以有效抵御针对该系统的扫描探测, 探测者只能得到拟态处理机对外呈现的信息, 而无法探测到系统全貌。

张铮等人[84]提出了一种适用于拟态防御架构的 Web 服务器测试方法, 并在 Web 服务器拟态防御原理验证系统中开展了实验, 结果表明, 在 Web 拟态防御系统中, 执行体的异构冗余与动态特性使得探测者远程扫描的结果不确定, 从而增加了探测者的扫描难度, 扰乱了扫描的可锁定性。

拟态防御技术经过多年的发展在抵御探测等多种类型的攻击方面取得了不错的效果, 有拟态 Web 服务器、拟态路由器、拟态域名服务器等系列产品设备落地, 不仅涉及到操作系统、应用服务等软件层面, 还涉及到交换机、路由器等硬件层面, 具有高可用、高可靠、安全可信等优势。然而, 在针对软件自身的安全防御方面还有所不足, 缺少拟态软件构造经验[89]。同时, 多执行体调度算法作为拟态防御的关键技术, 现有研究在调度时机、调度数量以及执行体异构度等方面仍有优化的空间。

3.4 测绘分析欺骗技术

随着网络空间测绘技术的不断发展, 测绘所采取的探测技术逐渐变得成熟且有效, 作为反测绘方也需要应对不断进化的探测技术, 这大大增加了防御的难度。采用前述的防御技术可以有效阻断绝大多数的探测行为, 更进一步地, 可以考虑将反测绘手段应用于测绘技术的第二阶段, 即对抗测绘分析行为, 让测绘方“测不准”。网络欺骗技术作为一种主动防御技术, 很适合被用来对抗测绘分析行为。利用测绘方需要依赖第一阶段探测得到的信息进行下一步分析这个特点, 防御方可以通过采取欺骗措施构造出一系列的虚假信息, 达到干扰、混淆测绘方的探测结果的目的, 使其在探测阶段得到的信息从源头上就出现偏差, 进而误导其对探测数据的关联分析, 让测绘方绘制出一幅错误的网络空间地图。

网络空间安全领域中, “欺骗”(Deception)一词来源于军事上对欺骗的定义。Yuill[132]在 2007 年提出了“计算机安全欺骗”(computer-security deception)的概念, 其定义为“为误导攻击方而采取的有计划的行动, 从而使他们采取(或不采取)有助于计算机安全防御的具体行动”, 该定义也是最广为接受的定义之一。此后, Almeshekah 等人[133]在《Cyber Deception》

表 3　网络空间反测绘技术体系中的欺骗技术

Table 3　Deception Techniques in Technical Architecture of Cyberspace Anti-Surveying and Mapping

层级	参考文献及年份	技术要点	欺骗单元
设备实体层	Murphy 等[92], 2010	操作系统混淆	系统注册表参数
	Zhao 等[93], 2017	操作系统混淆	系统流量指纹
	Rrushi 等[94], 2016	建立虚拟网卡	网络流量
	游建舟等[95], 2020	物联网蜜罐	物联网设备
网络拓扑层	Rowe 等[96], 2006	假蜜罐	配置参数
	Spitzner[97], 2003	蜜网	虚拟主机和拓扑
	Provos[98-99], 2003/2004	Honeyd 蜜罐	虚拟主机和拓扑
	Abbasi 等[100], 2009	Gen-III 蜜网	蜜网网关
	Han 等[101], 2016	HoneyMix 蜜网	蜜网网关
	Karlin 等[102], 2011	诱饵路由	路由
	Brewer 等[103], 2010	诱饵链接	资源链接
应用服务层	Jhon 等[104], 2011	Web 应用蜜罐	Web 服务
	Mphago 等[105], 2015	Web 应用蜜罐	Web 服务
	贾召鹏等[106], 2018	基于协同机制的蜜罐簇	Web 服务
	Kippo[107], 2014	SSH 蜜罐	SSH 服务
	Virvilis 等[108], 2014	虚假配置文件	配置文件
	Julian[109], 2002	随机增加响应延迟	服务性能
	Michael 等[110], 2002	软件诱饵	应用程序响应
	Araujo 等[111], 2014	未打补丁的诱饵	应用程序漏洞
	Araujo 等[112], 2015	未打补丁的诱饵	应用程序漏洞
	Vollmer 等[113], 2014	创建虚假网络实体	蜜罐流量
	Albanese 等[114], 2015	人为操作输出流量	流量指纹
	Obfsproxy[115], 2012	流量伪装	流量类型
	Dust[116], 2011	流量随机化	流量统计特征
	Obfs[117-119], 2011/2013/2014	流量随机化	流量统计特征
	FTE[120], 2013	流量拟态	流量指纹
	CensorSpoofer[121], 2012	流量拟态	流量指纹
数据流量层	CloudTransport[122], 2014	隧道加密	流量加密负载
	Meek[123], 2015	隧道加密	流量加密负载
	Barradas 等[124], 2020	隐蔽信道	流量加密负载
	Li 等[125], 2016	黑盒分析推断流量识别规则	流量内容数据
	胡永进等[126], 2020	基于对抗样本生成欺骗流量	流量特征
	Conti 等[127], 2011	虚拟专用社交网络	用户信息
	Danezis 等[128], 2010	匿名消息和 IP 语音通信	通信关系
	Beato 等[129], 2014	利用 P2P 网络隐藏用户交互	社交关系
	Angel 等[130], 2016	隐藏内容和元数据	用户数据
	Abraham 等[131], 2020	增强匿名提交广播	用户数据

一书中, 在 Yuill 的定义基础上增加了“混淆(Confusion)”的概念, 将欺骗定义为“为误导和/或混淆攻击方而采取的有计划的行动, 从而使他们采取(或不采取)有助于计算机安全防御的具体行动”。可以看出欺骗技术的精髓就是防御方采取主动措施去误导或混淆对手, 本文中将这类技术统称为“网络欺骗”。

有关网络欺骗技术的分类, 前人也有很多相关的研究。Almeshekah 等人[134]从网络欺骗技术的粒度出发, 将欺骗技术划分为: 决策(Decision)、响应(Response)、服务(Service)、活动(Activity)、弱点(Weakness)、性能(Performance)、配置(Configuration)和数据(Data)。Gartner[135]从应用欺骗技术的层次出发将网络欺骗划分为四个层次: 网络层(Network)、端点层(Endpoint)、应用程序层(Application)和数据层(Data)。贾召鹏等人[14]则从网络欺骗的作用点出发将网络欺骗分为四层: 设备层、网络层、数据层和应用层。

本文从反测绘的角度出发, 在表 3 中对反测绘有关的网络欺骗技术进行了总结, 按照网络欺骗技术被应用的层次将这些已有的研究工作划分为四个层次: 设备实体层、网络拓扑层、应用服务层和数据流量层。

3.4.1 设备实体层测绘对抗

设备实体层测绘对抗主要是使用操作系统混淆、物联网设备蜜罐和虚假蜜罐进行欺骗, 这些技术不仅可以让测绘方得到虚假的设备信息, 误导测绘方的数据分析, 还可以捕获测绘方的探测流量。

OSfuscate[92]是一种能够在 Windows 操作系统上运行的操作系统混淆工具, 该工具对操作系统的注册表值进行修改, 使得指纹识别工具不能正确地收集到操作系统指纹信息。Zhao 等人[93]提出使用软件定义网络建立一个指纹跳变系统来混淆主机操作系统指纹, 通过对数据包中的字段信息进行修改, 实现改变数据包指纹的效果, 从而防止网络指纹攻击。Rrushi 等人[94]提出了一种在操作系统中建立虚拟网卡的方法, 在该方法中虚拟网卡上的网络流量都是虚假的, 从而延缓攻击进度。

物联网设备被广泛部署在工业控制、智慧城市等多个关键领域, 由于其数量庞大且很多直接暴露在互联网中, 设备安全性相对不足, 因此成为了黑客重点攻击的目标。相应的, 各类工控蜜罐、消费级物联网蜜罐、信息物理系统蜜罐相继涌现。游建舟等人[95]对上述三类物联网蜜罐进行了详细的综述, 指出由于架构封闭性、诱饵多样性、物理融合性这三大特点, 难以实现通用的高交互物联网蜜罐。

Rowe 等人[96]提出了虚假蜜罐的欺骗方式, 通过修改计算机系统中一些比较常见的指标, 令攻击方认为计算机系统是一个蜜罐, 从而使其放弃对计算机系统的攻击。Rowe 等人在实验中采取了将计算机系统的 VMWare 暴露出来以及在文件系统中加入 Honeynet 项目中的蜜罐工具目录等方式实现了虚假蜜罐欺骗。

设备层欺骗不仅有助于实现对攻击的诱导和欺骗, 还有助于威胁的发现、捕获和分析。其不足体现在设备模仿的有效性和交互性都有待增强, 欺骗的真实性和可控性之间难以平衡; 使用实体设备的成本较高且不易扩展, 若基于虚拟化技术实现则对系统设计、执行环境等要素有较高的技术要求。同时, 对信息物理系统蜜罐的研究仍处于概念研究阶段, 并未出现完整通用的信息物理系统蜜罐方案[95]。

3.4.2 网络拓扑层测绘对抗

网络拓扑层测绘对抗能够利用蜜网、诱饵路由等技术伪造出网络拓扑、路由以及网络上的主机, 使得测绘方获得虚假的网络结构和主机信息。

Honeynet[97]是一个使用蜜罐构建出的标准网络, 一般位于防火墙等访问控制设备之后, 测绘方在绕过访问控制设备后可以探测到构建出的蜜网, 将蜜网误认为是真实的网络。Provos[98-99]提出了可以欺骗 Nmap 等探测工具的 Honeyd 蜜罐, 使用此蜜罐能够对网络主机不同操作系统的网络协议栈进行仿真, 构建出虚拟主机的操作系统指纹、TCP 开放端口、UDP 开放端口, 同时, Honeyd 还能够虚拟出路由器, 构建出一个虚拟网络拓扑结构。

蜜网架构 Gen-III[100]在蜜网中增加了一个称为“蜜墙(Honeywall)”的定制防火墙, 其作为蜜网网关将通过的流量引到已有的蜜网当中。Han 等人[101]提出了基于软件定义网络(Software Defined Network, SDN)的 HoneyMix 蜜网, 该蜜网使用组播的方式将恶意流量发送给多个提供相同服务的蜜罐, 这样可以最大限度使用蜜罐资源。同时, HoneyMix 动态选择最好的响应连接并对响应中携带的指纹信息进行清除。此外, HoneyMix 可以及时对泄漏蜜罐进行隔离并迅速重启一个新的蜜罐。

值得注意的是, 网络层的欺骗也可以用作逃避检测的技术, 例如诱饵路由(Decoy Routing)[102], 也可以称其为折射网络(Refraction Networking)[136]。诱饵路由被设计为利用诱饵目的地来规避基于 IP 地址的网络过滤, 在用户和诱饵目的地之间的路径上实现了支持隐蔽信道的诱饵路由器。因此, 用户能够通过隐蔽通道访问被过滤的 IP 地址, 而测绘方无法发现真实的网络路由情况。

网络层的欺骗技术作用范围较为广泛, 可以隐藏网络的真实拓扑以及网络内的主机, 但是仍有一些值得注意问题。首先, Honeyd 这类蜜罐虽然部署成本较低, 但是交互性不强, 容易被识别; 密网技术则需要以较多的资源为基础, 维护成本较高; 折射网络在实际应用中往往需要 ISP 级别的网络部署能力, 且难以抵御被动流量分析。

3.4.3 应用服务层测绘对抗

在网络空间测绘中, 针对应用服务层的测绘分析主要是针对 Web、SSH 等应用层服务展开的, 应用层的欺骗包括了与特定应用服务(如 Web 应用程序或数据库等)相关联的欺骗技术, 通过伪装应用服务, 响应正常或非正常的请求, 迷惑测绘方对该应用服务的认知。

Brewer 等人[103]提出了一种嵌入诱饵链接的 Web 应用程序, 这些链接对普通用户是不可见的, 但是

会被爬虫程序和网络机器人触发, 对非正常访问进行欺骗。John 等人[104]针对 Web 服务器易受攻击的问题开发了名为“heat-seeking”的 Web 应用服务蜜罐, 可以达到主动诱骗对手的效果。Glastopf[105]是一个使用软件模拟方式实现的 Web 应用蜜罐, 通过伪装成真实服务器的响应达到欺骗效果。为了提高蜜罐系统欺骗的成功率, 贾召鹏等人[106]提出了“蜜罐簇”的概念, 通过在 Web 蜜罐系统中部署多个不同应用的蜜罐组成蜜罐簇, 并通过设计蜜罐簇协同算法使得整个蜜罐簇作为一个 Web 蜜罐发挥作用。Kippo[107]是一款可以伪装成 SSH 网络服务的中交互蜜罐软件, 可以提供伪造的操作系统 Shell 供测绘方访问。还有的研究工作通过故意设置错误的 Web 服务器配置等虚假信息迷惑对手, 例如 Virvilis 等人[108]通过使用包含虚假信息的 robots.txt 实施主动欺骗。

为了应对针对应用服务可能存在的弱点或漏洞的测绘分析行为, 我们可以通过伪装漏洞或随机响应常见的漏洞扫描的方式实施欺骗。Julian[109]通过随机增加延迟的方式模拟系统饱和, 达到欺骗探测方的目的。Michael 等人[110]引入了智能软件诱饵的概念, 该诱饵可以检测并响应可疑行为的模式。针对测绘方可以通过分析应用服务或服务的响应来判断应用服务上是否存在特定漏洞这一情况, Araujo 等人[111-112]提出了 Honey-patches 方案, 其将传统的安全补丁改造成 Honey-patches, 当系统检测到漏洞探测行为时, 会将探测流量重定向到一个未打补丁的诱饵之上, 在诱饵处让对方误认为系统存在某个漏洞。通过这种使用诱饵的方法, 可以利用伪造的数据或系统设置向对方提供虚假的信息, 扰乱其对应用服务的分析, 同时还可以使防御方在诱饵处收集有关探测行为的信息。

应用服务层的欺骗技术通常需要考虑欺骗环境的交互性, 低交互蜜罐往往采取虚拟化部署, 仅模拟单一服务, 较为容易识别; 高交互蜜罐往往以真实系统为基础构建, 功能更加多样, 暴露的信息也会更多, 且存在被攻击方完全控制并当作跳板的风险, 需要配合有效的监控措施。同时, 漏洞伪装、随机响应、流量重定向等手段一定程度上也会影响到服务的可用性, 增大了欺骗系统被识别的机率。

3.4.4 数据流量层测绘对抗

数据流量层测绘对抗是指针对测绘方可以获取到的网络流量、用户信息等数据采取有针对性的混淆、加密、伪装等技术手段的过程, 可以使测绘方的分析结果偏离真实的数据特征。本节将从网络流量分析对抗和社交网络匿名化两方面对数据流量层的测绘对抗进行阐述。

1) 网络流量分析对抗

网络流量分析技术作为一种常用的针对数据流量层的测绘技术, 经常被用来对网络流量进行分类以确定其所属的应用类型(音视频、邮件等)或者进行操作系统、网络服务等信息的指纹识别, 进而帮助测绘方根据分类结果采取进一步的行动。

Vollmer 等人[113]提出了一种自配置蜜罐的新方法, 可以被动地检查工业控制网络中的网络流量, 并且可以通过感知环境进行主动调整, 创建用来实施诱骗的网络实体。针对嗅探、网络流量分析等主被动方式进行远程主机操作系统指纹识别或特定服务指纹识别的测绘分析行为, Albanese 等人[114]通过对输出流量进行人为修改, 使其类似于使用不同操作系统的主机生成的流量, 抵御了操作系统指纹识别; 通过在某些数据包离开主机或网络之前拦截并修改它们, 进而改变服务的 banner(从 banner 信息可以获取到软件开发商、软件名称、服务类型、版本号等敏感信息), 抵御了服务指纹识别。

流量混淆是一种可以有效应对流量分析的办法。文献[48,51,137-138]均通过修改流量的通信特征实现对流量分析行为的欺骗, 但是这种方式的伪装和欺骗能力有限, 对抗性不足。研究人员也提出了许多其他的流量伪装与混淆方案, 其中的代表工作为 Obfsproxy[115](Obfuscated proxy)。Obfsproxy 可以将流量伪装成 HTTP 流量或者即时通讯软件流量。姚忠将等人[139]对流量混淆技术进行了详细的综述, 将流量混淆技术分为三类, 即随机化(Dust[116]、Obfs[117-119]等)、拟态(FTE[120]、CensorSpoofer[121]等)和隧道(CloudTransport[122]、Meek[123]等)。此外, Li 等人[129]提出了黑盒分析流量分类规则的方法, 通过多次测试可以推断出流量分析识别规则, 进而有针对性地修改通信数据包内容, 达到逃避检测的目的。Barradas 等人[124]基于 WebRTC 开发了名为的 Protozoa 的通信工具, 可以将数据隐藏在视频通话的多媒体流数据中, 相比此前的同类隐蔽通信技术[121,140-144], 该工具同时兼顾了高性能和强抗分析能力, 在保证高吞吐量的同时还可以抵御最先进的被动流量分析攻击。

随着机器学习和深度学习技术的快速发展, 这两类技术也被广泛应用于网络流量分类领域, 并将分类准确率大大提高, 为测绘方提供了十分便利的条件。相应的, 针对应用了机器学习和深度学习的流量分类技术的对抗或欺骗技术也随之出现。例如, Szegedy 等人[145]最早提出的用于图像分类对抗领域

的生成对抗样本技术也被应用于网络流量分类对抗领域。胡永进等人[126]针对流量分类问题, 从防御的角度出发, 提出了一种基于对抗样本的网络欺骗流量生成方法, 通过在正常的网络流量中增加扰动, 形成欺骗流量的对抗样本, 导致以深度学习模型为基础的流量分类方法出现错误。

2) 社交网络匿名化技术

用户信息收集、用户画像以及社交图分析等技术是网络测绘领域的研究热点。社交网络中出现的匿名化技术提供了一种在使用社交网络的同时又充分隐藏真实个人信息及用户间交互行为的手段, 让测绘方无法发现用户的真实身份并且难以分析出用户间的社交关系。

虚拟专用社交网络(Virtual Private Social Network, VPSN)这类技术为用户提供了可行的隐私保护, 防止针对用户的分析行为。Conti 等人[127]提出了一种针对 Facebook 实现的 VPSN, 名为 FaceVPSN。FaceVPSN 允许用户在 Facebook 中上传伪造的个人信息, 但是会将真实信息发送给用户的朋友, 并通过浏览器插件的形式更新网页中的虚假信息为用户的真实信息。

用户除了通过伪造个人信息实现匿名化, 还会期望隐藏其与其他用户的交互信息。Danezis 等人[128]提出了一种基于中继的匿名机制 Drac, 提供了实时即时消息传递和 IP 语音通信的匿名性和不可观察性。Beato 等人[129]提出了一个名为虚拟友谊(Virtual Friendship)的系统, 该系统基于去中心化的 P2P 网络构建, 可以隐藏社交网络中的用户交互。Angel 等人[130]提出了名为 Pung 的通信系统, 该系统被证明可以隐藏所有内容和元数据。Abraham 等人[131]首次为匿名提交广播(Anonymous Committed Broadcast)提供了可伸缩和完全健壮解决方案的系统, 同时保证了匿名性、健壮性和可用性。

值得注意的是, 匿名化技术不仅是学术界研究的热点, 在产业界也出现了许多实际应用。例如遨游网推出的“百变邮箱”服务[146], 其本质上是一种邮件转发服务, 它可以为用户的真实邮箱设置任意数量的影子邮箱, 影子邮箱中收到的邮件都将会被转发到真实邮箱, 达到完全隐藏用户的真实邮件账号的效果。此外, “万变 IP”等各种网络代理工具层出不穷, 为各类用户提供任意地区的 IP 接入, 进一步隐匿了用户的真实 IP, 给用户画像带来了极大的挑战。

数据流量层的欺骗实现了指纹、行为、流量等多个层面的混淆和伪装, 可以有效抵御多种主被动测绘分析手段, 但仍有一些需要注意的问题。网络流量分析对抗技术虽然取得了不错的效果, 但是针对其的分析手段也层出不穷, 尤其是各类基于深度学习的流量分析技术。类似的, 社交网络的去匿名化技术也在不断发展, 现有技术面临着极大的挑战。此外, 数据流量欺骗技术与其他欺骗技术一样, 存在着执行效率、性能和安全性、可用性之间的矛盾。

4 反测绘对抗技术

反测绘技术的发展在一定程度上也促进了测绘技术的进步, 为了增强测绘的有效性, 安全研究人员开始从攻击方的角度出发开展一系列的研究工作对抗反测绘技术, 本节将从蜜罐、数据伪装欺骗、匿名化这三大类技术出发, 介绍相应的对抗工作。针对蜜罐等相关技术, 研究人员开展了大量反蜜罐技术的研究工作, 包括对蜜罐的识别和欺骗。针对网络流量加密与混淆等数据伪装欺骗技术, 网络流量识别领域涌现了大量研究成果。针对匿名化技术, 如何实现去匿名化也是当今的研究热点。

1) 蜜罐识别与对抗

基于蜜罐的运行环境和其承载的业务系统并不能与真实世界完全保持一致的特性, 研究人员开发了许多有效的蜜罐识别技术。

Fu 等人[147]利用软件和硬件的时钟频率不同的特点, 通过远程测量网络连接延迟并基于延迟构建分类器, 实现了对 Honeyd 虚拟网络的识别。Mukkamala 等人[148]提出了基于网络响应时间的网络层蜜罐指纹构建和识别的方法, 在理想条件下仅使用 ICMP ECHO 请求即可获得高达 95%的检测率。此外, Mukkamala 等人还提出了一种应用层级别的识别方式, 基于蜜罐往往只提供单一服务而真实的网络主机会提供较为完整的网络服务的区别, 构建了应用层服务指纹, 可以识别 SSH 等较为简单的蜜罐系统。高交互蜜罐往往基于真实的操作系统实现, 操作系统可以运行在虚拟化平台上或是真实的硬件之上, 此类蜜罐往往需要特定监控软件的配合, 而引入监控软件后就给蜜罐识别带来了新的途径, 即能够通过识别蜜罐监控系统发现蜜罐环境[149]。

2) 加密网络流量识别与对抗

网络通信加密化已经成为不可阻挡的趋势, 加密网络流量的规模呈现爆炸式的增长, 这种变化给测绘方提出了新的挑战。针对这种情况, 近年来, 研究人员结合机器学习和深度学习等技术, 通过对网络流量进行属性提取、特征提取、特征选择和学习、训练分类器等一系列处理, 形成了许多较为有效的加密网络流量识别方法[150-160], 下面给出了该领域近

些年来一些具有代表性的研究。

Rimmer 等人[161]利用匿名化网络 Tor 的特性, 采用网络流量的包长度序列和方向序列作为深度学习模型的输入, 实现了对 Tor 网页的分类。Taylor 等人[162]利用数据包大小和方向等特征作为侧信道信息, 并结合机器学习技术实现了智能手机 APP 加密流量的指纹识别。针对机器学习需要人工设计特征以及专家知识的问题, Aceto 等人[163]基于深度学习技术实现了自动提取特征的移动应用流量分类器, 对多种深度学习模型进行了复现、剖析和集成。Liu 等人[164]提出了一种基于表示学习的流序列网络(FS-Net), 以双向 GRU 为基本单元, 以编码器-解码器为整体结构, 同时采用重构机制增强了加密流量指纹的表现能力。Zheng 等人[165]设计了一种基于元学习(meta-learning)思想的端到端加密流量分类模型, 称其为基于流的关系网络, 该模型可以从原始流量中学习有代表性的特征, 并且充分考虑了样本不平衡问题, 不仅可以实现出色的分类性能, 还具有很好的泛化性。Ma 等人[166]针对网络中物联网设备流量, 提取了其时空特征, 并结合 CNN 实现了精确的设备指纹识别。

3) 去匿名化技术

如 3.3.2 节所述, 匿名网络面临着网络流量分析和节点攻击两类威胁。此外, 针对匿名社交网络也有相应的去匿名化技术。

Tian 等人[167]将网络流量分析领域的去匿名化工作分为三大类: 推断拓扑结构、推断主机行为和推断节点信息。推断拓扑结构旨在恢复匿名网络的网络拓扑结构, 是一种全局性的去匿名化。Coull 等人[168]利用匿名网络产生的数据包, 基于 K-means 算法实现了子网聚类。推断主机行为旨在发现匿名网络中的主机特征行为等动态信息, 进而对匿名网络中的主机进行识别。Xu 等人[169]提出了称为主导状态分析(Dominant State Analysis)的分析手段, 可以确定匿名网络中每个 IP 的特征行为。推断节点信息旨在获取节点的物理特性、节点间的连通性、路由等静态属性。Paul 等人[170]假设可以向匿名网络中注入足够多的信息, 在网络拓扑不变的情况下可以推断出与 k 条边相关联的节点。

节点攻击指的是针对匿名网络中的目标主机发起的攻击, 旨在使匿名通信服务失效或是暴露主机的隐私信息。例如, 攻击方可以通过攻击 Tor 的中继节点, 实现对隐藏服务和用户的去匿名化。Sniper 攻击[171]是当前对 Tor 网络影响最大的节点攻击, 可以造成中继节点的瘫痪, 实现拒绝服务的攻击效果。Egger 等人[172]针对 I2P 提出了一种去匿名化攻击技术, 可以用来使 I2P 节点泄露自身的节点信息。

在匿名社交网络的去匿名化方面, 研究人员也进行了许多探索。Qian 等人[173]引入了知识图谱技术增强去匿名和推理攻击的能力, 基于已有的用户先验知识构建知识图谱, 制定去匿名化和隐私推断的过程, 并在真实社交网络数据上证实了其有效性。针对在线社交网络(Online Social Network, OSN)的服务提供商会定期发布匿名的 OSN 数据的特点, Gao 等人[174]提出了持久同源性(Persistent Homology)的概念, 可以捕获到动态图中持久化的结构信息, 同时结合结构相似性和属性相似性实现匿名网络中的节点映射。由于不同的社交网络之间的用户存在着重合, 因此可以通过将未知的匿名网络与已知的匿名网络进行匹配的方式对用户进行识别, 这种技术分为有种子和无种子两种情况。Zhang 等人[175]首次将多跳邻域关系整合到无种子的去匿名化过程中, 充分利用了所有节点的多跳邻居信息, 能够以高概率正确地匹配两个网络。

5 总结和展望

随着网络空间规模的与日俱增和网络技术的突飞猛进, 网络空间中承载了越来越多的网络资源, 既包括以硬件基础设施为代表的各类实体资源, 也包括以应用服务、数据为代表的各类虚拟资源。网络空间测绘技术的日新月异使得各类网络空间资源暴露无遗, 这为网络空间带来了极大的威胁。本文首次提出了网络空间反测绘的概念, 为保护网络空间资源提供了思路。首先, 结合网络空间测绘的定义和相关技术, 从防御方的角度出发给出了网络空间反测绘的定义; 其次, 从探测行为识别、探测行为防护和测绘分析欺骗三个方面构建了反测绘技术体系, 并对各个步骤所涉及到的相关工作和防御思路进行了阐述; 最后, 探讨了对抗反测绘技术的手段。从本文对反测绘各阶段关键技术的阐述可以看出, 本文提出的反测绘技术体系涉及到众多技术领域, 这些技术领域各自的思想基础、目标、实现手段、实现效果存在着差异和关联, 反测绘技术体系是多种技术相辅相成、相互融合的结果。

当前网络空间反测绘的相关研究工作已经取得了一定的进展, 但是反测绘技术和测绘技术的发展是一个在博弈中不断更新演进的过程, 现有的反测绘技术仍存在着一些值得深入研究的问题, 本文从以下几个方面展望了进一步的研究工作。

1)进一步提升反测绘技术的伪装能力、抗识别能力, 在执行效率、性能和安全性、可用性之间的寻求

平衡。在各类反测绘对抗技术不断涌现的情况下, 需要研究人员持续投入, 使反测绘技术能够适应形式越来越多样的反蜜罐机制、流量分析技术等威胁。同时, 在提升安全性的同时, 需要兼顾效率问题, 保证系统的可用性, 不影响用户体验。

2)以软件定义网络、网络功能虚拟化、IPv6 等技术为依托, 结合丰富的实物环境构建虚实结合、高效协同、适应大规模部署的防御体系。现有的反测绘技术在单点防御上具有不错的效果, 但是仍然未能充分融合各类防御手段, 缺乏实物环境和虚拟环境间的协同以及多种虚拟环境的协同。因此需要借助软件定义网络、网络功能虚拟化等技术的灵活部署、快速迁移、统一管理等优势, 充分利用软硬件资源, 研究高效协同、灵活多变的防御体系。此外, 还可以进一步利用 IPv6 海量地址空间的优势, 加强其与蜜罐、移动目标防御等技术的融合。

3)综合运用人工智能技术, 构建智能化的反测绘防御体系。机器学习、深度学习等技术通常运用在数据流量的欺骗、伪装等领域, 需要进一步加强与移动目标防御等新兴技术的融合, 构建智能化的技术框架, 实现威胁智能感知、防御策略智能优化、防御单元智能调度、业务环境智能迁移等目标, 使防御效能最大化。

可以预见, 网络空间反测绘技术将会随着测绘手段的不断演进和网络攻防技术的不断进步朝着更加灵活、有效、可靠的方向发展, 必将得到学术界和工业界的持续研究和关注。

致 谢 本课题得到中国科学院战略性先导科技专项(XDC02030400)资助。

参考文献

[1] Nakashima E. With Plan X, Pentagon seeks to spread US military might to cyberspace[J]. *The Washington Post,* 2012, 30.

[2] Grant T. On the military geography of cyberspace[J]. *Leading Issues in Cyber Warfare and Security: Cyber Warfare Secur*, 2015, 2: 119.

[3] Rashid O, Mullins I, Coulton P, et al. Extending Cyberspace[J]. *Computers in Entertainment*, 2006, 4(1): 4.

[4] Durumeric Z, Wustrow E, Halderman J A. ZMap: Fast Internet-Wide Scanning and Its Security Applications[C]. *The 22nd USENIX conference on Security*, 2013: 605-620.

[5] masscan. https://github.com/robertdavidgraham/masscan.

[6] Shodan. https://www.shodan.io/.

[7] ZoomEye. https://www.zoomeye.org/about.

[8] Censys. http://censys.io/.

[9] FOFA. https://fofa.so/.

[10] Jajodia S. *Moving target defense: creating asymmetric uncertainty for cyber threats*[M]. New York: Springer, 2011.

[11] Jajodia S, Subrahmanian V, Swarup V, et al. Cyber deception[M]. Heidelberg: Springer, 2016.

[12] Shi L Y, Li Y, Ma M F. Latest Research Progress of Honeypot Technology[J]. *Journal of Electronics & Information Technology*, 2019, 41(2): 498-508.
(石乐义, 李阳, 马猛飞. 蜜罐技术研究新进展[J]. *电子与信息学报*, 2019, 41(2): 498-508.)

[13] Sengupta S, Chowdhary A, Sabur A, et al. A Survey of Moving Target Defenses for Network Security[J]. *IEEE Communications Surveys & Tutorials*, 2020, 22(3): 1909-1941.

[14] Jia Z P, Fang B X, Liu C G, et al. Survey on Cyber Deception[J]. *Journal on Communications*, 2017, 38(12): 128-143.
(贾召鹏, 方滨兴, 刘潮歌, 等. 网络欺骗技术综述[J]. *通信学报*, 2017, 38(12): 128-143.)

[15] Guo L, Cao Y N, Su M J, et al. Cyberspace Resources Surveying and Mapping: The Concepts and Technologies[J]. *Journal of Cyber Security*, 2018, 3(4): 1-14.
(郭莉, 曹亚男, 苏马婧, 等. 网络空间资源测绘: 概念与技术[J]. *信息安全学报*, 2018, 3(4): 1-14.)

[16] Zhou Y, Xu Q, Luo X Y, et al. Research on Definition and Technological System of Cyberspace Surveying and Mapping[J]. *Computer Science*, 2018, 45(5): 1-7.
(周杨, 徐青, 罗向阳, 等. 网络空间测绘的概念及其技术体系的研究[J]. *计算机科学*, 2018, 45(5): 1-7.)

[17] Zhao F, Luo X Y, Liu F L. Research on Cyberspace Surveying and Mapping Technology[J]. *Chinese Journal of Network and Information Security*, 2016, 2(9): 1-11.
(赵帆, 罗向阳, 刘粉林. 网络空间测绘技术研究[J]. *网络与信息安全学报*, 2016, 2(9): 1-11.)

[18] Sherry J, Katz-Bassett E, Pimenova M, et al. Resolving IP Aliases with Prespecified Timestamps[C]. *The 10th ACM SIGCOMM conference on Internet measurement*, 2010: 172-178.

[19] Spring N, Mahajan R, Wetherall D. Measuring ISP Topologies with Rocketfuel[J]. *ACM SIGCOMM Computer Communication Review*, 2002, 32(4): 133-145.

[20] Hou Y, Chen X, Hao Y, et al. Survey of Cyberspace Resources Scanning and Analyzing[C]. *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 2020: 279-291.

[21] Qi Y F, Bai L F, Tang G, et al. A Summary of Cyberspace Security Mimic Defense Technology[J]. *Cyberspace Security*, 2018, 9(10): 45-49.
(齐云菲, 白利芳, 唐刚, 等. 网络空间测绘概念理解与分析[J].

网络空间安全, 2018, 9(10): 45-49.)

[22] Wang C D, Guo Y B, Zhen S H, et al. Research on Network Asset Detection Technology[J]. *Computer Science*, 2018, 45(12): 24-31.
(王宸东, 郭渊博, 甄帅辉, 等. 网络资产探测技术研究[J]. *计算机科学*, 2018, 45(12): 24-31.)

[23] Zhu X L. *Research on active detection technology for large scale network topology*[D]. Harbin: Harbin Engineering University, 2017.
(朱新立. 大规模网络拓扑主动探测技术研究[D]. 哈尔滨: 哈尔滨工程大学, 2017.)

[24] Bhuyan M H, Bhattacharyya D K, Kalita J K. Surveying Port Scans and Their Detection Methodologies[J]. *The Computer Journal*, 2011, 54(10): 1565-1581.

[25] Bou-Harb E, Debbabi M, Assi C. Cyber Scanning: A Comprehensive Survey[J]. *IEEE Communications Surveys & Tutorials*, 2014, 16(3): 1496-1519.

[26] Burr C, Cristianini N, Ladyman J. An Analysis of the Interaction between Intelligent Software Agents and Human Users[J]. *Minds and Machines*, 2018, 28(4): 735-774.

[27] Geens N, Huysmans J, Vanthienen J. Evaluation of Web Robot Discovery Techniques: A Benchmarking Study[M]. *Advances in Data Mining. Applications in Medicine, Web Mining, Marketing, Image and Signal Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006: 121-130.

[28] Duskin O, Feitelson D G. Distinguishing Humans from Robots in Web Search Logs: Preliminary Results Using Query Rates and Intervals[C]. *The 2009 workshop on Web Search Click Data*, 2009: 15-19.

[29] Tan P N, Kumar V. Discovery of Web Robot Sessions Based on Their Navigational Patterns[M]. Intelligent Technologies for Information Analysis. Berlin, Heidelberg: Springer, 2004: 193-222.

[30] Lagopoulos A, Tsoumakas G, Papadopoulos G. Web Robot Detection: A Semantic Approach[C]. *2018 IEEE 30th International Conference on Tools with Artificial Intelligence*, 2018: 968-974.

[31] Stevanovic D, Vlajic N, An A J. Detection of Malicious and Non-Malicious Website Visitors Using Unsupervised Neural Network Learning[J]. *Applied Soft Computing*, 2013, 13(1): 698-708.

[32] Modi C, Patel D, Borisaniya B, et al. A Survey of Intrusion Detection Techniques in Cloud[J]. *Journal of Network and Computer Applications*, 2013, 36(1): 42-57.

[33] Roesch M. Snort: Lightweight intrusion detection for networks[C]. *Lisa*, 1999, 99(1): 229-238.

[34] Vigna G, Kemmerer R A. NetSTAT: A Network-Based Intrusion Detection System[J]. *Journal of Computer Security*, 1999, 7(1): 37-71.

[35] Chaum D. The Dining Cryptographers Problem: *Unconditional Sender and Recipient Untraceability*[J]. *Journal of Cryptology*, 1988, 1(1): 65-75.

[36] Levine B N, Shields C. Hordes: A Multicast Based Protocol for Anonymity1[J]. *Journal of Computer Security*, 2002, 10(3): 213-240.

[37] Chaum D L. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms[J]. *Communications of the ACM*, 1981, 24(2): 84-90.

[38] Goldschlag D M, Reed M G, Syverson P F. Hiding Routing Information[M]. *Information Hiding*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996: 137-150.

[39] Syverson P F, Goldschlag D M, Reed M G. Anonymous Connections and Onion Routing[C]. *Proceedings of 1997 IEEE Symposium on Security and Privacy*, 2002: 44-54.

[40] Dingledine R, Mathewson N, Syverson P. Tor: The Second-Generation Onion Router[J]. *Proceedings of the 13th USENIX Security Symposium*, 2004.

[41] Clarke I, Sandberg O, Wiley B, et al. Freenet: A Distributed Anonymous Information Storage and Retrieval System[M]. *Designing Privacy Enhancing Technologies*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001: 46-66.

[42] Freenet. https://freenetproject.org/.

[43] Zantout B, Haraty R. I2P data communication system[C]. *Proceedings of ICN*, 2011: 401-409.

[44] I2P Anonymous Network. https://geti2p.net/en/.

[45] Luo J Z, Yang M, Ling Z, et al. Anonymous Communication and Darknet: A Survey[J]. *Journal of Computer Research and Development*, 2019, 56(1): 103-130.
(罗军舟, 杨明, 凌振, 等. 匿名通信与暗网研究综述[J]. *计算机研究与发展*, 2019, 56(1): 103-130.)

[46] Wright C V, Coull S E, Monrose F. Traffic Morphing: An Efficient Defense Against Statistical Traffic Analysis[C]. *NDSS*, 2009: 9.

[47] Chan-Tin E, Kim T, Kim J. Website Fingerprinting Attack Mitigation Using Traffic Morphing[C]. *2018 IEEE 38th International Conference on Distributed Computing Systems*, 2018: 1575-1578.

[48] Dyer K P, Coull S E, Ristenpart T, et al. Peek-a-Boo, I still See You: Why Efficient Traffic Analysis Countermeasures Fail[C]. *2012 IEEE Symposium on Security and Privacy*, 2012: 332-346.

[49] Cai X, Nithyanand R, Wang T, et al. A Systematic Approach to Developing and Evaluating Website Fingerprinting Defenses[C]. *The 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014: 227-238.

[50] Luo X, Zhou P, Chan E W W, et al. HTTPOS: Sealing Information Leaks with Browser-side Obfuscation of Encrypted Flows[C]. *NDSS*, 2011: 11.

[51] Wang T, Goldberg I. Walkie-Talkie: An Efficient Defense Against Passive Website Fingerprinting Attacks[C]. *The 26th USENIX Conference on Security Symposium*, 2017: 1375-1390.

[52] NITRD C. IWG: Cybersecurity game-change research and devel-

opment recommendations[J]. 2013.

[53] Kewley D, Fink R, Lowry J, et al. Dynamic Approaches to Thwart Adversary Intelligence Gathering[C]. *Proceedings DARPA Information Survivability Conference and Exposition II. DISCEX'01*, 2002: 176-185.

[54] Antonatos S, Akritidis P, Markatos E P, et al. Defending Against Hitlist Worms Using Network Address Space Randomization[J]. *Computer Networks*, 2007, 51(12): 3471-3490.

[55] Lee H C J, Thing V L L. Port Hopping for Resilient Networks[C]. *IEEE 60th Vehicular Technology Conference*, *2004*. *VTC2004-Fall*, 2005: 3291-3295.

[56] Sifalakis M, Schmid S, Hutchison D. Network Address Hopping: A Mechanism to Enhance Data Protection for Packet Communications[C]. *IEEE International Conference on Communications*, *2005*. *ICC 2005*, 2005: 1518-1523.

[57] Atighetchi M, Pal P, Webber F, et al. Adaptive Use of Network-Centric Mechanisms in Cyber-Defense[C]. *Sixth IEEE International Symposium on Object-Oriented Real-Time Distributed Computing*, 2003: 183-192.

[58] Hu Y X, Zheng K F, Yang Y X, et al. Moving Target Defense Solution on Network Layer Based on OpenFlow[J]. *Journal on Communications*, 2017, 38(10): 102-112.
(胡毅勋, 郑康锋, 杨义先, 等. 基于 OpenFlow 的网络层移动目标防御方案[J]. *通信学报*, 2017, 38(10): 102-112.)

[59] Sharma D P, Kim D S, Yoon S, et al. FRVM: Flexible Random Virtual IP Multiplexing in Software-Defined Networks[C]. *2018 17th IEEE International Conference on Trust*, *Security and Privacy In Computing and Communications/ 12th IEEE International Conference on Big Data Science and Engineering*, 2018: 579-587.

[60] Narantuya J, Yoon S, Lim H, et al. SDN-Based IP Shuffling Moving Target Defense with Multiple SDN Controllers[C]. *2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks - Supplemental Volume*, 2019: 15-16.

[61] Dunlop M, Groat S, Urbanski W, et al. MT6D: A Moving Target IPv6 Defense[C]. *2011 - MILCOM 2011 Military Communications Conference*, 2012: 1321-1326.

[62] Hao S S, Liu R J, Weng Z, et al. Addressless: A New Internet Server Model to Prevent Network Scanning[J]. *PLoS ONE*, 2021, 16(2): e0246293.

[63] Al-Shaer E. Toward Network Configuration Randomization for Moving Target Defense[M]. Moving Target Defense. New York: Springer, 2011: 153-159.

[64] Bangalore A K, Sood A K. Securing Web Servers Using Self Cleansing Intrusion Tolerance (SCIT)[C]. *2009 Second International Conference on Dependability*, 2009: 60-65.

[65] Nguyen Q L, Sood A. Designing SCIT Architecture Pattern in a Cloud-Based Environment[C]. *2011 IEEE/IFIP 41st International Conference on Dependable Systems and Networks Workshops*, 2011: 123-128.

[66] Huang Y, Ghosh A K. Introducing Diversity and Uncertainty to Create Moving Attack Surfaces for Web Services[M]. Moving Target Defense. New York: Springer, 2011: 131-151.

[67] Ahmed N O, Bhargava B. Mayflies: A Moving Target Defense Framework for Distributed Systems[C]. *The 2016 ACM Workshop on Moving Target Defense*, 2016: 59-64.

[68] Alavizadeh H, Jang-Jaccard J, Kim D S. Evaluation for Combination of Shuffle and Diversity on Moving Target Defense Strategy for Cloud Computing[C]. *2018 17th IEEE International Conference on Trust*, *Security and Privacy in Computing and Communications/ 12th IEEE International Conference on Big Data Science and Engineering*, 2018: 573-578.

[69] Carter K M, Riordan J F, Okhravi H. A Game Theoretic Approach to Strategy Determination for Dynamic Platform Defenses[C]. *The First ACM Workshop on Moving Target Defense*, 2014: 21-30.

[70] Thompson M, Evans N, Kisekka V. Multiple OS Rotational Environment an Implemented Moving Target Defense[C]. *2014 7th International Symposium on Resilient Control Systems*, 2014: 1-6.

[71] Roeder T, Schneider F B. Proactive Obfuscation[J]. *ACM Transactions on Computer Systems*, 28(2)Article No. 4.

[72] Jangda A, Mishra M, De Sutter B. Adaptive Just-in-Time Code Diversification[C]. *The Second ACM Workshop on Moving Target Defense*, 2015: 49-53.

[73] Mahmood K, Shila D M. Moving Target Defense for Internet of Things Using Context Aware Code Partitioning and Code Diversification[C]. *2016 IEEE 3rd World Forum on Internet of Things*, 2017: 329-330.

[74] Thompson M, Mendolla M, Muggler M, et al. Dynamic Application Rotation Environment for Moving Target Defense[C]. *2016 Resilience Week*, 2016: 17-26.

[75] Christodorescu M, Fredrikson M, Jha S, et al. End-to-End Software Diversification of Internet Services[M]. Moving Target Defense. New York: Springer, 2011: 117-130.

[76] Vikram S, Yang C, Gu G F. NOMAD: Towards Non-Intrusive Moving-Target Defense Against Web Bots[C]. *2013 IEEE Conference on Communications and Network Security*, 2013: 55-63.

[77] Wu J X. Cyberspace Mimicry Security Defense[J]. *Secrecy Science and Technology*, 2014(10): 4-9, 1.
(邬江兴. 网络空间拟态安全防御[J]. *保密科学技术*, 2014(10): 4-9, 1.)

[78] Wu J X. Meaning and Vision of Mimic Computing and Mimic Security Defense[J]. *Telecommunications Science*, 2014, 30(7): 2-7.
(邬江兴. 拟态计算与拟态安全防御的原意和愿景[J]. *电信科学*, 2014, 30(7): 2-7.)

[79] Wu J X. Research on Cyber Mimic Defense[J]. *Journal of Cyber*

Security, 2016, 1(4): 1-10.
(邬江兴. 网络空间拟态防御研究[J]. *信息安全学报*, 2016, 1(4): 1-10.)

[80] Wu J X. Mimicry Defense Technology to Build Endogenous Security in National Information Network Space[J]. *Information and Communications Technologies*, 2019, 13(6): 4-6.
(邬江兴. 拟态防御技术构建国家信息网络空间内生安全[J]. *信息通信技术*, 2019, 13(6): 4-6.)

[81] Ma H L, Yi P, Jiang Y M, et al. Dynamic Heterogeneous Redundancy Based Router Architecture with Mimic Defenses[J]. *Journal of Cyber Security*, 2017, 2(1): 29-42.
(马海龙, 伊鹏, 江逸茗, 等. 基于动态异构冗余机制的路由器拟态防御体系结构[J]. *信息安全学报*, 2017, 2(1): 29-42.)

[82] Tong Q, Zhang Z, Zhang W H, et al. Design and Implementation of Mimic Defense Web Server[J]. *Journal of Software*, 2017, 28(4): 883-897.
(仝青, 张铮, 张为华, 等. 拟态防御 Web 服务器设计与实现[J]. *软件学报*, 2017, 28(4): 883-897.)

[83] Wei S, Yu H, Gu Z Y, et al. Architecture of Mimic Security Processor for Industry Control System[J]. *Journal of Cyber Security*, 2017, 2(1): 54-73.
(魏帅, 于洪, 顾泽宇, 等. 面向工控领域的拟态安全处理机架构[J]. *信息安全学报*, 2017, 2(1): 54-73.)

[84] Zhang Z, Ma B L, Wu J X. The Test and Analysis of Prototype of Mimic Defense in Web Servers[J]. *Journal of Cyber Security*, 2017, 2(1): 13-28.
(张铮, 马博林, 邬江兴. web 服务器拟态防御原理验证系统测试与分析[J]. *信息安全学报*, 2017, 2(1): 13-28.)

[85] Li W C, Zhang Z, Wang L Q, et al. The Modeling and Risk Assessment on Redundancy Adjudication of Mimic Defense[J]. *Journal of Cyber Security*, 2018, 3(5): 64-74.
(李卫超, 张铮, 王立群, 等. 基于拟态防御架构的多余度裁决建模与风险分析[J]. *信息安全学报*, 2018, 3(5): 64-74.)

[86] Zhang Y J, Pang J M, Zhang Z, et al. Mimic Security Defence Strategy Based on Software Diversity[J]. *Computer Science*, 2018, 45(2): 215-221.
(张宇嘉, 庞建民, 张铮, 等. 基于软件多样化的拟态安全防御策略[J]. *计算机科学*, 2018, 45(2): 215-221.)

[87] Ding S H, Li J F, Ji X S. Research on SDN Control Layer Security Based on Mimic Defense[J]. *Journal of Cyber Security*, 2019, 4(4): 84-93.
(丁绍虎, 李军飞, 季新生. 基于拟态防御的 SDN 控制层安全机制研究[J]. *信息安全学报*, 2019, 4(4): 84-93.)

[88] Ren Q, Wu J X, He L. Research on Mimic DNS Architectural Strategy Based on Generalized Stochastic Petri Net[J]. *Journal of Cyber Security*, 2019, 4(2): 37-52.
(任权, 邬江兴, 贺磊. 基于GSPN的拟态DNS构造策略研究[J]. *信息安全学报*, 2019, 4(2): 37-52.)

[89] Yao D, Zhang Z, Zhang G F, et al. A Survey on Multi-Variant Execution Security Defense Technology[J]. *Journal of Cyber Security*, 2020, 5(5): 77-94.
(姚东, 张铮, 张高斐, 等. 多变体执行安全防御技术研究综述[J]. *信息安全学报*, 2020, 5(5): 77-94.)

[90] Yao D, Zhang Z, Zhang G F, et al. MVX-CFI: A Practical Active Defense Framework for Software Security[J]. *Journal of Cyber Security*, 2020, 5(4): 44-54.
(姚东, 张铮, 张高斐, 等. MVX-CFI: 一种实用的软件安全主动防御架构[J]. *信息安全学报*, 2020, 5(4): 44-54.)

[91] Yang X, Li H, Wu J X, et al. A Two-Dimension Security Assessing Model for CMDS Combined with Generalized Stochastic Petri Net[J]. *Scientia Sinica* (*Informationis*), 2020, 50(12): 1944-1960.
(杨昕, 李挥, 邬江兴, 等. 融合广义随机 Petri 网的二维拟态安全评估模型[J]. *中国科学: 信息科学*, 2020, 50(12): 1944-1960.)

[92] Murphy S, McDonald T, Mills R. An application of deception in cyberspace: Operating system obfuscation[C]. *Proceedings of the 5th International Conference on Information Warfare and Security*, 2010: 241-249.

[93] Zhao Z, Liu F L, Gong D F. An SDN-Based Fingerprint Hopping Method to Prevent Fingerprinting Attacks[J]. *Security and Communication Networks*, 2017, 2017: 1-12.

[94] Rrushi J L. NIC Displays to Thwart Malware Attacks Mounted from within the OS[J]. *Computers & Security*, 2016, 61: 59-71.

[95] You J Z, Lv S C, Sun Y Y, et al. A Survey on Honeypots of Internet of Things[J]. *Journal of Cyber Security*, 2020, 5(4): 138-156.
(游建舟, 吕世超, 孙玉砚, 等. 物联网蜜罐综述[J]. *信息安全学报*, 2020, 5(4): 138-156.)

[96] Rowe, Duong, Custy. Fake Honeypots: A Defensive Tactic for Cyberspace[C]. *2006 IEEE Information Assurance Workshop*, 2006: 223-230.

[97] Spitzner L. The Honeynet Project: Trapping the Hackers[J]. *IEEE Security & Privacy*, 2003, 1(2): 15-23.

[98] Provos N. Honeyd-a virtual honeypot daemon[C]. *10th DFN-CERT Workshop, Hamburg, Germany*, 2003, 2: 4.

[99] Provos N. A Virtual Honeypot Framework[C]. *USENIX Security Symposium*, 2004, 173(2004): 1-14.

[100] Abbasi F H, Harris R J. Experiences with a Generation III Virtual Honeynet[C]. *2009 Australasian Telecommunication Networks and Applications Conference*, 2010: 1-6.

[101] Han W, Zhao Z M, Doupé A, et al. HoneyMix: Toward SDN-Based Intelligent Honeynet[C]. *The 2016 ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization*, 2016: 1-6.

[102] Karlin J, Ellard D, Jackson A W, et al. Decoy Routing: Toward Unblockable Internet Communication[C]. *FOCI*, 2011.

[103] Brewer D, Li K, Ramaswamy L, et al. A Link Obfuscation Service to Detect Webbots[C]. *2010 IEEE International Conference on Services Computing*, 2010: 433-440.

[104] John J P, Yu F, Xie Y L, et al. Heat-Seeking Honeypots: Design and Experience[C]. *The 20th international conference on World wide web*, 2011: 207-216.

[105] Mphago B, Bagwasi O, Phofuetsile B, et al. Deception in dynamic web application honeypots: Case of glastopf[C]. Proceedings of the International Conference on Security and Management. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing, 2015: 104.

[106] Jia Z P, Fang B X, Cui X, et al. ArkHoney: A Web Honeypot Based on Collaborative Mechanisms[J]. *Chinese Journal of Computers*, 2018, 41(2): 413-425.
(贾召鹏, 方滨兴, 崔翔, 等. ArkHoney: 基于协同机制的 Web 蜜罐[J]. *计算机学报*, 2018, 41(2): 413-425.)

[107] desaster/kippo. https://github.com/desaster/kippo.

[108] Virvilis N, Vanautgaerden B, Serrano O S. Changing the Game: The Art of Deceiving Sophisticated Attackers[C]. *2014 6th International Conference on Cyber Conflict*, 2014: 87-97.

[109] Virvilis N, Vanautgaerden B, Serrano O S. Changing the game: The art of deceiving sophisticated attackers[C]. *2014 6th International Conference On Cyber Conflict. IEEE*, 2014: 87-97.

[110] Julian D P. Delaying-type responses for use by software decoys[R]. NAVAL POSTGRADUATE SCHOOL MONTEREY CA, 2002.

[111] Araujo F, Hamlen K W, Biedermann S, et al. From Patches to Honey-Patches: Lightweight Attacker Misdirection, Deception, and Disinformation[C]. *The 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014: 942-953.

[112] Araujo F, Hamlen K W. Compiler-Instrumented, Dynamic Secret-Redaction of Legacy Processes for Attacker Deception[C]. *The 24th USENIX Conference on Security Symposium*, 2015: 145-159.

[113] Vollmer T, Manic M. Cyber-Physical System Security with Deceptive Virtual Hosts for Industrial Control Networks[J]. *IEEE Transactions on Industrial Informatics*, 2014, 10(2): 1337-1347.

[114] Albanese M, Battista E, Jajodia S. A Deception Based Approach for Defeating OS and Service Fingerprinting[C]. *2015 IEEE Conference on Communications and Network Security*, 2015: 317-325.

[115] Weinberg Z, Wang J, Yegneswaran V, et al. StegoTorus: A Camouflage Proxy for the Tor Anonymity System[C]. *The 2012 ACM conference on Computer and communications security*, 2012: 109-120.

[116] Wiley B. Dust: A blocking-resistant internet transport protocol[J]. *Technical report.* http://blanu. net/Dust. pdf, 2011.

[117] Tor project. Obfsproxy2. https://gitweb.torproject.org/pluggable-transports/obfsproxy.git/tree/doc/obfs2/obfs2-protocol-spec.txt.

[118] Tor project. Obfsproxy3. https://gitweb.torproject.org/pluggable-transports/obfsproxy.git/tree/doc/obfs3/obfs3-protocol-spec.txt.

[119] Yawning. Obfsproxy4. https://github.com/Yawning/obfs4/blob/master/doc/obfs4-spec.txt.

[120] Dyer K P, Coull S E, Ristenpart T, et al. Protocol Misidentification Made Easy with Format-Transforming Encryption[C]. *The 2013 ACM SIGSAC conference on Computer & communications security*, 2013: 61-72.

[121] Wang Q Y, Gong X, Nguyen G T K, et al. CensorSpoofer: Asymmetric Communication Using IP Spoofing for Censorship-Resistant Web Browsing[C]. *The 2012 ACM conference on Computer and communications security*, 2012: 121-132.

[122] Wang Q, Gong X, Nguyen G T K, et al. Censorspoofer: asymmetric communication using ip spoofing for censorship-resistant web browsing[C]. *Proceedings of the 2012 ACM conference on Computer and communications security*, 2012: 121-132.

[123] Fifield D, Lan C, Hynes R, et al. Blocking-Resistant Communication through Domain Fronting[J]. *Proceedings on Privacy Enhancing Technologies*, 2015, 2015(2): 46-64.

[124] Barradas D, Santos N, Rodrigues L, et al. Poking a Hole in the Wall: Efficient Censorship-Resistant Internet Communications by Parasitizing on WebRTC[C]. *The 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2020: 35-48.

[125] Li F F, Kakhki A M, Choffnes D, et al. Classifiers Unclassified: An Efficient Approach to Revealing IP Traffic Classification Rules[C]. *The 2016 Internet Measurement Conference*, 2016: 239-245.

[126] Hu Y J, Guo Y B, Ma J, et al. Method to Generate Cyber Deception Traffic Based on Adversarial Sample[J]. *Journal on Communications*, 2020, 41(9): 59-70.
(胡永进, 郭渊博, 马骏, 等. 基于对抗样本的网络欺骗流量生成方法[J]. *通信学报*, 2020, 41(9): 59-70.)

[127] Conti M, Hasani A, Crispo B. Virtual Private Social Networks[C]. *The first ACM conference on Data and application security and privacy*, 2011: 39-50.

[128] Conti M, Hasani A, Crispo B. Virtual private social networks[C]. *Proceedings of the first ACM conference on Data and application security and privacy*, 2011: 39-50.

[129] Beato F, Conti M, Preneel B, et al. VirtualFriendship: Hiding Interactions on Online Social Networks[C]. *2014 IEEE Conference on Communications and Network Security*, 2014: 328-336.

[130] Angel S, Setty S. Unobservable Communication over Fully Untrusted Infrastructure[C]. *The 12th USENIX conference on Operating Systems Design and Implementation*, 2016: 551-569.

[131] Abraham I, Pinkas B, Yanai A. Blinder — Scalable, Robust Anonymous Committed Broadcast[C]. *The 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2020: 1233-1252.

[132] Abraham I, Pinkas B, Yanai A. Blinder--Scalable, Robust Anonymous Committed Broadcast[C]. *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2020: 1233-1252.

[133] Almeshekah M H, Spafford E H. Cyber Security Deception[M]. Cyber Deception. Cham: Springer, 2016: 23-50.

[134] Almeshekah M H, Spafford E H. Planning and Integrating Deception into Computer Security Defenses[C]. *The 2014 New Security Paradigms Workshop*, 2014: 127-138.

[135] Almeshekah M H, Spafford E H. Planning and integrating deception into computer security defenses[C]. *Proceedings of the 2014 New Security Paradigms Workshop*, 2014: 127-138.

[136] Refraction Networking. https://refraction.network/.

[137] Panchenko A, Niessen L, Zinnen A, et al. Website Fingerprinting in Onion Routing Based Anonymization Networks[C]. *The 10th annual ACM workshop on Privacy in the electronic society*, 2011: 103-114.

[138] Cui W Q, Yu J M, Gong Y M, et al. Realistic Cover Traffic to Mitigate Website Fingerprinting Attacks[C]. *2018 IEEE 38th International Conference on Distributed Computing Systems*, 2018: 1579-1584.

[139] Yao Z J, Ge J G, Zhang X D, et al. Research Review on Traffic Obfuscation and Its Corresponding Identification and Tracking Technologies[J]. *Journal of Software*, 2018, 29(10): 3205-3222.
(姚忠将, 葛敬国, 张潇丹, 等. 流量混淆技术及相应识别、追踪技术研究综述[J]. *软件学报*, 2018, 29(10): 3205-3222.)

[140] Moghaddam H M, Li B Y, Derakhshani M, et al. SkypeMorph: Protocol Obfuscation for Tor Bridges[C]. *The 2012 ACM conference on Computer and communications security*, 2012: 97-108.

[141] Yao Zhongjiang, Ge Jingguo, Zhang Xiaodan, et al. Research Review on Traffic Obfuscation and Its Corresponding Identification and Tracking Technologies[J]. *Journal of Software*, 2018, 10: 3205-3222.
(姚忠将, 葛敬国, 张潇丹, 等. 流量混淆技术及相应识别, 追踪技术研究综述[J]. *软件学报*, 2018, 10: 3205-3222.)

[142] Li S, Schliep M, Hopper N. Facet: Streaming over Videoconferencing for Censorship Circumvention[C]. *The 13th Workshop on Privacy in the Electronic Society*, 2014: 163-172.

[143] McPherson R, Houmansadr A, Shmatikov V. CovertCast: Using Live Streaming to Evade Internet Censorship[J]. *Proceedings on Privacy Enhancing Technologies*, 2016, 2016(3): 212-225.

[144] Barradas D, Santos N, Rodrigues L. DeltaShaper: Enabling Unobservable Censorship-Resistant TCP Tunneling over Videoconferencing Streams[J]. *Proceedings on Privacy Enhancing Technologies*, 2017, 2017(4): 5-22.

[145] Szegedy C, Zaremba W, Sutskever I, et al. Intriguing Properties of Neural Networks[EB/OL]. 2013: arXiv: 1312.6199. https://arxiv.org/abs/1312.6199.

[146] 百变邮箱_虚拟邮箱_临时邮箱_一次性邮箱_傲游浏览器 Maxthon 中国官网. https://www.maxthon.cn/mx5/uumail/.
Fu X W, Yu W, Cheng D, et al. On Recognizing Virtual Honeypots and Countermeasures[C]. *2006 2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing*, 2006: 211-218.

[147] Mukkamala S, Yendrapalli K, Basnet R, et al. Detection of Virtual Environments and Low Interaction Honeypots[C]. *2007 IEEE SMC Information Assurance and Security Workshop*, 2007: 92-98.

[148] Dornseif M, Holz T, Klein C N. NoSEBrEaK - Attacking Honeynets[C]. *Proceedings from the Fifth Annual IEEE SMC Information Assurance Workshop*, 2005: 123-129.

[149] Anderson B, McGrew D. Machine Learning for Encrypted Malware Traffic Classification: Accounting for Noisy Labels and Non-Stationarity[C]. *The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017: 1723-1732.

[150] Barradas D, Santos N, Rodrigues L. Effective Detection of Multimedia Protocol Tunneling Using Machine Learning[C]. *The 27th USENIX Conference on Security Symposium*, 2018: 169-185.

[151] Gu J X, Wang J L, Yu Z W, et al. Walls Have Ears: Traffic-Based Side-Channel Attack in Video Streaming[C]. *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018: 1538-1546.

[152] Zain ul Abideen M, Saleem S, Ejaz M. VPN Traffic Detection in SSL-Protected Channel[J]. *Security and Communication Networks*, 2019, 2019: 1-17.

[153] Rahman M S, Matthews N, Wright M. Poster: Video Fingerprinting in Tor[C]. *The 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019: 2629-2631.

[154] van Ede T, Bortolameotti R, Continella A, et al. FlowPrint: Semi-Supervised Mobile-App Fingerprinting on Encrypted Network Traffic[C]. *Proceedings 2020 Network and Distributed System Security Symposium*, 2020.

[155] Wang X, Chen S H, Su J S. Automatic Mobile App Identification from Encrypted Traffic with Hybrid Neural Networks[J]. *IEEE Access*, 2020, 8: 182065-182077.

[156] Xu S C, Sen S, Mao Z M. CSI: Inferring Mobile ABR Video Adaptation Behavior under HTTPS and QUIC[C]. *The Fifteenth European Conference on Computer Systems*, 2020: 1-16.

[157] Montieri A, Ciuonzo D, Aceto G, et al. Anonymity Services Tor, I2P, JonDonym: Classifying in the Dark[C]. *2017 29th International Teletraffic Congress*, 2017: 81-89.

[158] Acar A, Fereidooni H, Abera T, et al. Peek-a-Boo: I See your Smart Home Activities, even Encrypted![C]. *The 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, 2020: 207-218.

[159] Dong C, Zhang C, Lu Z G, et al. CETAnalytics: Comprehensive Effective Traffic Information Analytics for Encrypted Traffic Classification[J]. *Computer Networks*, 2020, 176: 107258.

[160] Rimmer V, Preuveneers D, Juarez M, et al. Automated Website Fingerprinting through Deep Learning[EB/OL]. 2017: arXiv: 1708.06376. https://arxiv.org/abs/1708.06376

[161] Taylor V F, Spolaor R, Conti M, et al. Robust Smartphone App Identification via Encrypted Network Traffic Analysis[J]. *IEEE Transactions on Information Forensics and Security*, 2018, 13(1): 63-78.

[162] Aceto G, Ciuonzo D, Montieri A, et al. Mobile Encrypted Traffic Classification Using Deep Learning: Experimental Evaluation, Lessons Learned, and Challenges[J]. *IEEE Transactions on Network and Service Management*, 2019, 16(2): 445-458.

[163] Liu C, He L T, Xiong G, et al. FS-Net: A Flow Sequence Network for Encrypted Traffic Classification[C]. *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 2019: 1171-1179.

[164] Zheng W B, Gou C, Yan L, et al. Learning to Classify: A Flow-Based Relation Network for Encrypted Traffic Classification[C]. *Proceedings of The Web Conference 2020*, 2020: 13-22.

[165] Ma X B, Qu J, Li J F, et al. Pinpointing Hidden IoT Devices via Spatial-Temporal Traffic Fingerprinting[C]. *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, 2020: 894-903.

[166] Zheng W, Gou C, Yan L, et al. Learning to Classify: A Flow-Based Relation Network for Encrypted Traffic Classification[C]. *Proceedings of The Web Conference 2020. New York, NY, USA: Association for Computing Machinery*, 2020: 13-22.

[167] Ma X, Qu J, Li J, et al. Pinpointing Hidden IoT Devices via Spatial-temporal Traffic Fingerprinting[C]. *IEEE INFOCOM 2020-IEEE Conference on Computer Communications. IEEE*, 2020: 894-903.

[168] Xu K A, Zhang Z L, Bhattacharyya S. Profiling Internet Backbone Traffic[J]. *ACM SIGCOMM Computer Communication Review*, 2005, 35(4): 169-180.

[169] Paul R R, Valgenti V C, Kim M S. Real-Time Netshuffle: Graph Distortion for On-Line Anonymization[C]. *2011 19th IEEE International Conference on Network Protocols*, 2011: 133-134.

[170] Jansen R, Tschorsch F, Johnson A, et al. The Sniper Attack: Anonymously Deanonymizing and Disabling the Tor Network[C]. *Proceedings 2014 Network and Distributed System Security Symposium*, 2014: 169-180.

[171] Paul R R, Valgenti V C, Kim M S. Real-time Netshuffle: Graph distortion for on-line anonymization[C]. *2011 19th IEEE International Conference on Network Protocols. IEEE*, 2011: 133-134.

[172] Qian J W, Li X Y, Zhang C H, et al. Social Network De-Anonymization and Privacy Inference with Knowledge Graph Model[J]. *IEEE Transactions on Dependable and Secure Computing*, 2019, 16(4): 679-692.

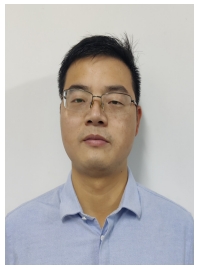
[173] Gao T C, Li F. De-Anonymization of Dynamic Online Social Networks via Persistent Structures[C]. *ICC 2019 - 2019 IEEE International Conference on Communications*, 2019: 1-6.

[174] Zhang J P, Fu L Y, Wang X B, et al. De-Anonymization of Social Networks: The Power of Collectiveness[C]. *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, 2020: 89-98.

刘庆云 于 2005 年在中国科学院大学信息安全专业获得博士学位。现任中国科学院信息工程研究所第三工程部副主任、正高级工程师。研究领域为信息安全、网络空间测绘。Email: f liuqingyun@iie.ac.cn

李仁杰 于 2019 年在南京理工大学软件工程专业获得学士学位。现在中国科学院大学网络空间安全专业攻读博士学位。研究领域为网络安全、网络空间测绘。研究兴趣包括: 入侵检测、网络攻防。Email: lirenjie@iie.ac.cn

周舟 于 2012 年在北京理工大学大学计算机应用技术专业获得博士学位。现任中国科学院信息工程研究所高级工程师。研究领域为网络安全、高性能网络。Email: zhouzhou@iie.ac.cn



钟友兵 于 2021 年在中国科学院大学网络空间安全专业获得博士学位。现任兵器工业计算机应用技术研究所工程师。研究领域为网络流处理。研究兴趣包括: 网络功能虚拟化, 网络流量处理。Email: zhongyoub@sina.com

石峰源 于 2021 年在国际关系学院信息管理与信息系统专业获得学士学位。现在中国科学院大学网络空间安全专业攻读博士学位。研究领域为入侵检测、移动目标防御。研究兴趣包括: 蜜罐、卷积神经网络。Email: shifengyuan@iie.ac.cn

郭莉 于 1994 年在湘潭大学获得硕士学位。现任中国科学院信息工程研究所副所长、正研级高工。研究领域为网络空间安全。Email: guoli@iie.ac.cn