

# 数据基础设施抗测绘理论与技术发展研究

薛向阳<sup>1</sup>, 邹宏<sup>1</sup>, 赵进<sup>2</sup>, 周喆<sup>2</sup>, 尚玉婷<sup>1</sup>, 鄂江兴<sup>1,3\*</sup>

(1. 复旦大学大数据研究院, 上海 200433; 2. 复旦大学计算机科学技术学院, 上海 200433;  
3. 国家数字交换系统工程技术研究中心, 郑州 450002)

**摘要:** 当前, 我国正全面深化数字化转型, 推动“数据要素×”行动计划, 这对数据基础设施建设提出了新要求。作为一类新型基础设施, 数据基础设施和传统信息基础设施一样, 应具备抗测绘能力, 防止被单向透明探测。数据基础设施的抗测绘能力成为数据空间安全的基本要求, 直接关系到国家、社会和个人的信息安全。本文概述了数据基础设施测绘及抗测绘的基本概念, 分析了主流抗测绘技术的发展现状、主要特点以及存在的关键问题, 提出了内生安全抗测绘的理论、方法和关键技术, 以解决现有网络架构下数据资产固有可测性这一关键问题。研究建议: 尽快开展数据空间抗测绘基础理论和关键技术研究, 将抗测绘能力纳入国家数据基础设施建设基本要求和基础能力之中, 加快数据基础设施关键技术研发和攻关力度, 高度关注新兴技术带来的“被测绘”风险, 加强数据基础设施抗测绘产业的扶持力度, 推动多学科交叉融合和产教融合, 尽快形成我国在数据空间的抗测绘能力。

**关键词:** 数据空间; 数据基础设施; 抗测绘; 内生安全

**中图分类号:** TP393 **文献标识码:** A

## Advances in Anti-Surveying-and-Mapping Theory and Technologies for Data Infrastructure

Xue Xiangyang<sup>1</sup>, Zou Hong<sup>1</sup>, Zhao Jin<sup>2</sup>, Zhou Zhe<sup>2</sup>, Shang Yuting<sup>1</sup>, Wu Jiangxing<sup>1,3\*</sup>

(1. Institute of Big Data, Fudan University, Shanghai 200433, China; 2. School of Computer Science, Fudan University, Shanghai 200433, China; 3. National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002, China)

**Abstract:** China is deepening its digital transformation and promoting a data-driven action plan, proposing new requirements for the construction of data infrastructure. Like conventional information infrastructure types, data infrastructure needs to possess an anti-surveying-and-mapping capability to avoid being detected in one-way and transparent manners. The anti-surveying-and-mapping capability of data infrastructure has become a fundamental requirement for data space security, directly related to the information security of a country, society, and individuals. This study outlines the basic concepts of data infrastructure surveying-and-mapping and anti-surveying-and-mapping, analyzes the development status, major characteristics, and key problems of mainstream anti-surveying-and-mapping technologies, and proposes the theory, methods, and key technologies of anti-surveying-and-mapping with endogenous security, to solve the key problem of inherent measurability of data assets under the existing network architecture. To this end, the following suggestions are proposed: strengthening research on the basic theories and key technologies regarding data space anti-

**收稿日期:** 2024-09-23; **修回日期:** 2024-11-22

**通讯作者:** \*鄂江兴, 复旦大学计算机科学技术学院教授, 中国工程院院士, 研究方向为计算机与网络技术、网络空间安全;

E-mail: jxwu@fudan.edu.cn

**资助项目:** 国家重点研发计划项目(2022YFB3102901); 中国工程院咨询项目“国家数据空间发展路径与技术体系研究”(2024-XBZD-05), “国家数据空间发展战略研究”(2023-XBZD-16)

**本刊网址:** sscae.engineering.org.cn

surveying-and-mapping; taking anti-surveying-and-mapping capability as the basic requirement of national data infrastructure construction; and increasing support for key technology research, industry incubation, multidisciplinary integration, and industry-university-research collaboration.

**Keywords:** data space; data infrastructure; anti-surveying-and-mapping; endogenous safety

## 一、前言

当前,数据要素已成为与土地、劳动力、资本、技术等并列的生产要素之一,逐步融入生产生活各个环节,深刻影响并重构着经济社会结构<sup>[1-4]</sup>。随着数据要素应用广度和深度大幅拓展,数据空间作为人类活动的新空间正在加快形成。数据空间关注数据的产生、流通、聚合、使用的全过程,从网络连接一切发展为一切皆为数据,推动海量碎片化数据的广泛关联和价值释放。发展数据空间首先要推动数据基础设施建设,数据基础设施是数据空间建设和运行的基本支撑,是数字经济高质量发展的基石。

数据基础设施作为一种新型的信息基础设施,是“人、机、物、地、信息”的重要载体。数据要素存在海量泛在、属性稳定、跨域关联等固有特征,决定了通过对数据要素的分析,能够再现数据基础设施的结构及功能,使得数据基础设施存在被对手单向透明的测绘风险。“基础不牢、地动山摇”,数据基础设施暴露是经济社会发展乃至国家安全的重大风险隐患。一些先进国家在大力发展“数字攻击地图”计划,构建针对数据基础设施的“攻击面地图”<sup>[5,6]</sup>。据统计,截止到2023年3月底,我国互联网协议第4版(IPv4)和互联网协议第6版(IPv6)的暴露面资产总数达到33亿。为此,在我国数据基础设施建设的筹划、起步阶段,需要像建设重要工业基础设施、国防基础设施一样,特别关注抗测绘问题,为提高数据空间安全防御能力提供强有力的技术支撑。

中国工程院于2023年4月设立“国家数据空间发展战略研究”重大咨询项目,用内生安全理论赋能数据基础设施建设<sup>[7-9]</sup>,是该项目的重要研究方向。通过研究发现,内生安全构造具有天然的抗测绘能力,能够建立数据空间的高维表达,更好达成数据基础设施抗测绘中“测不到”“测不准”“测不全”“绘不出”等目标。目前,我国的网络空间抗测绘研究取得了一定进展,但是在数据基础设施抗测绘这一领域,还缺乏相关研究。研究探讨数据基

础设施抗测绘理论与技术发展,既可以弥补我国在该领域的研究空白,也是国家数据空间发展的重大战略需求。

本文在文献调研、专家研讨、定性与定量研究的基础上,提出数据基础设施测绘与抗测绘的基本概念、技术体制、发展路径;分析主要数据基础设施抗测绘技术体制的特点,以及存在的缺陷和问题;通过对不同技术路线的分析比较,提出了内生安全理论赋能数据基础设施抗测绘的理论构想、实现机理、技术路线;结合对国家数据空间建设的思考,提出推进数据基础设施抗测绘技术创新和应用的研究启示与发展建议。

## 二、数据基础设施测绘技术概述

“测绘”一词从字面上可理解为测量和绘图,其目的是把复杂的结构、现象、关系以简洁、清晰、易于理解的方式表达出来。随着人类活动范围从物理空间、网络空间到数据空间的不断拓展,测绘的概念也在迭代演进<sup>[10]</sup>,针对数据基础设施进行探测、感知、映射、表达、呈现的理论和应用技术应运而生。

### (一) 数据基础设施

数据基础设施是从数据要素价值释放的角度出发,在网络、算力等设施的支持下,面向社会提供一体化数据汇聚、处理、流通、应用、运营、安全保障服务的一类新型基础设施,是覆盖硬件、软件、开源协议、标准规范、机制设计等在内的有机整体。数据基础设施主要由算力网络、数据要素平台和数据服务构成。其中,算力网络实现异构计算资源的虚拟化和抽象化,并通过智能调度算法,根据任务需求动态优化资源配置;数据要素平台通过技术组件之间的协同作用,支撑数据的流通与汇集,完成数据的全生命周期管理;数据服务通过计算和网络系统提供的功能与编排,支持数据面向行业的高级分析、应用服务和价值释放,形成生态。如图1所示,算力网络、数据要素平台、数据服务

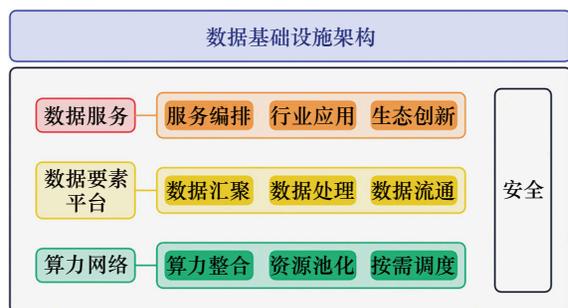


图1 数据基础设施架构

共同构成数据基础设施架构，而泛在的安全问题在数据基础设施中有着形态各异的表现和特征，对数据基础设施的测绘是其中面临的重要风险。

从测绘的视角看，数据基础设施具有以下特征。一是海量泛在。随着物联网的逐步普及、算力网广泛互联，联网的数据资产数量大幅增加，资产的开放暴露面激增。二是属性稳定。实体资源和虚拟资源的属性，包括软硬件信息、服务类别、端口、归属等固化耦合。三是跨域关联。网络空间、地理空间和社会空间进行相互映射，能够以数据资产为中心，通过数据业务流程和使用路径，表达数据交互、流动等行为链条。四是价值涌现。通过海量数据要素汇聚与加工，实现知识的提炼、数据的生成与价值的涌现，扩大了数据和信息可达的边界。数据基础设施的特点如同“双刃剑”，在促进数据价值高效释放的同时，也为测绘提供了诸多“先天便利条件”。

## （二）数据基础设施测绘技术的内涵特征

网络空间测绘是对网络空间中各种虚实资源及其属性进行探测、分析和绘制的全过程<sup>[5,6,10,11]</sup>。借鉴网络空间测绘的概念，数据基础设施测绘可以理解为，通过数据要素化、数据分析、数据可视化等技术手段，将数据基础设施以及数据资源的基本属性与物理空间、网络空间、数据空间乃至社会空间、认知空间的关联关系进行建模和表达，实现全要素、全链条的全息映射和可视化地图展现，反映区域或国家数据基础设施以及资源的空间结构、时空分布、部署状态、演进变化和人类意图等。可见，数据基础设施测绘与网络空间测绘虽然具有相关性，但是其关注点仍具有区分度，一是关注基础设施的部署和配置，尤其是关键算力节点、要素连

接关系等基本状态；二是关注演进和变化，尤其是增量速度、分布以及地理空间特性；三是关注与新兴产业之间的关系，例如对自动驾驶、智能制造、智慧医疗甚至是国家关键能力之间的耦合关系；四是关注与社会活动以及认知范畴的连接态势，以此对国家社会行为变化进行分析预测。

## （三）数据基础设施测绘技术发展情况

数据基础设施测绘技术的发展历程可以简要概括为三个阶段。首先是萌芽阶段，大约在20世纪90年代，随着互联网的初步发展和最初数据基础设施的建设，对其进行测绘的需求开始显现。这一时期的测绘主要聚焦于数据基础设施的物理设备识别<sup>[12-14]</sup>。21世纪初，数据基础设施初具规模，并进入飞速发展阶段。在这一阶段，随着互联网的快速扩张，数据基础设施测绘技术的内涵和外延得到了极大的丰富，不仅限于物理设备的测绘，还开始涉及更广泛的数据资源和属性<sup>[5,6,10,11]</sup>。这一时期，国际上，美国等发达国家率先推进网络空间测绘技术的应用。例如，美国先后推出了三个代表性的网络空间测绘计划<sup>[5,6]</sup>：2006年美国国家安全局的“藏宝图”（TreasureMap）计划，旨在建立一个近实时交互式的全球互联网地图；2012年美国国防部启动网络战发展项目“X计划”，生成网络空间作战态势图；2015年美国国土安全部启动SHodan Intelligence Extraction（SHINE）计划，建立美国本土网络空间关键基础设施信息数据库。除了政府部门的研究计划，也有开源或者商业的网络安全搜索引擎，例如Shodan、Censys、ZoomEye等，这些搜索引擎利用非侵入式的探测，实现对基础设施设备、逻辑拓扑、网络服务等进行探测和分析，生成包括资源基本信息、网络属性、位置属性、连通属性等信息。

当前，数据基础设施测绘的含义得到进一步的补充和完善。研究者们开始关注数据基础设施、数据资源与物理空间、网络空间、数据空间乃至社会空间、认知空间的相互映射和融合分析，数据基础设施测绘理论体系逐渐构建完善，技术也趋于成熟。

数据基础设施测绘的关键技术分为探测、分析、绘制三个过程。云计算、人工智能、大数据分析等技术的不断创新和发展<sup>[15-19]</sup>，也为数据基础设施测绘技术的融合和创新带来了更多的机遇。

探测是测绘的基础，是指通过主动、被动方式获取网际互联网协议（IP）存活性、端口、拓扑、操作系统、服务等方面的资源测量数据。当前探测的主要趋势集中在如何在规模资源环境以及反探测约束下，提升探测的精度和效率。例如，利用协同探测模式提升效率<sup>[20]</sup>，基于互联网控制报文协议（ICMP）限速侧信道来加速 IPv6 的网络测量<sup>[21]</sup>。

分析是从探测结果中提取资源及其属性，并进行分析建模和关联映射，实现对资源属性识别、拓扑结构还原分析和重要目标画像等。当前的研究进展主要集中在以下几方面：利用 MapReduce、Spark 等大数据框架进行数据分析；利用深度学习、多源数据融合等技术实现对数据基础设施资源及其属性的精准画像<sup>[22]</sup>；利用知识图谱等形式呈现画像<sup>[23]</sup>。

通过绘制并建立时空模型，对探测和分析结果进行可视化表达，包括要素可视化表达、关系可视化描述和事件可视化分析等。当前的研究进展主要集中在引入地理学理论构建网络空间可视化体系<sup>[24]</sup>，以及形成多尺度、多要素融合的网络空间统一描述<sup>[25]</sup>。

这些测绘技术被应用于算力网络、数据要素平台以及数据服务的深度分析和大规模监控，有效提高了数据基础设施测绘的效率和准确性。通过对关键资源全面、系统、连续的观测和度量，可以揭示数据基础设施的关键资产、运行规律和相互作用关系，从而对数据基础设施进行精准画像，但同时也给防范非法测绘带来了更多的挑战。

### 三、数据基础设施抗测绘技术体系分析

数据基础设施抗测绘是数据基础设施测绘的反向技术，是数据基础设施建设的“保底”技术，用于保护各类数据基础设施抵御非法测绘，即对测绘行为进行干扰、欺骗、遏制、阻断，让测绘分析结果偏离真实情形，防止数据基础设施的结构、状态被外部恶意探测，使得测绘方无法绘制出真实的数据空间地图。数据基础设施抗测绘技术可以梳理为抗跨域关联、抗拓扑发现和抗设备探测三类，如表 1 所示。图 2 为数据基础设施抗测

表 1 数据基础设施抗测绘技术体系

技术类型	作用层次	抗测绘技术	
抗设备探测	算力节点	迷惑与陷阱	设置迷惑和陷阱节点，诱导攻击者进行错误判断和攻击，从而保护真实节点
		计算混淆	对节点信息进行混淆处理，使攻击者难以获取准确的设备信息
		频谱跳变	通过伪装设备的通信频率，隐藏真实设备的存在
	网络基础设施	分布式防御	采用分布式防御机制，增强网络的整体抗攻击能力
边缘计算		将计算任务分散到网络边缘，减少集中式计算的安全风险	
弹性计算		通过弹性计算技术，动态调整计算资源，提高系统的灵活性和抗攻击能力	
抗拓扑发现	数据流通与汇聚	动态路由技术	通过动态调整路由路径，防止攻击者通过固定路径追踪和获取数据
		分段与隔离	将数据流分段处理，以物理介质或传输信道隔离不同的数据，增加数据追踪和关联的难度
	网络拓扑	虚拟网络技术	通过虚拟化技术，创建虚拟网络隔离实际网络架构，防止拓扑发现和攻击
抗跨域关联	数据服务	流量混淆	通过混淆网络流量，掩盖真实的通信模式，增加攻击者的分析难度
		IP 黑 / 白名单	使用 IP 地址的黑 / 白名单机制来控制访问权限，黑名单中的 IP 地址将被拒绝访问，而白名单中的 IP 地址则被允许访问
		查询混淆	通过混淆查询请求来保护数据隐私，使外部无法通过分析查询模式来推测出实际的数据
	数据运行	数据模糊化	对数据进行模糊化处理，即使数据被获取，攻击者也难以还原出原始数据
		用户匿名化	通过匿名化处理，移除或隐匿用户的身份信息，从而保护用户隐私
		数据遮罩	通过对数据进行处理和掩盖，使其在分析和共享时不会泄露个人信息
	数据处理	访问控制	通过严格的访问控制机制，限制未授权用户对数据的访问，确保数据安全
同态加密		一种加密方法，允许在加密数据上进行计算，而不需要解密数据，从而保护数据隐私	
		差分隐私	在数据交换时避免传输原始数据，每次传输内容为与之前数据相比的差异部分，从而保护个体数据的隐私，同时允许对数据集进行统计分析

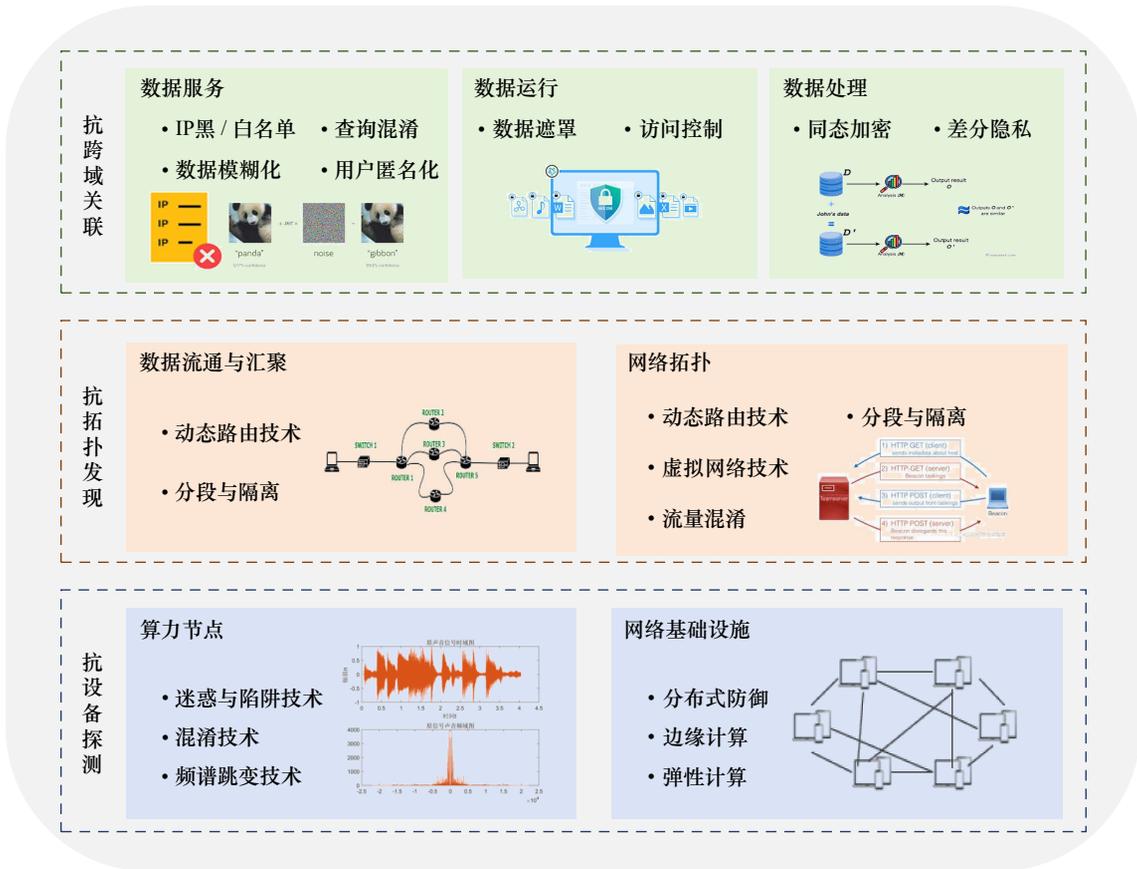


图2 数据基础设施抗测绘技术体系示意图

绘技术体系示意图。

### (一) 关键数据设备抗探测技术

抗设备探测是数据基础设施抗测绘技术体系的重中之重，主要应用在算力网络层的算力节点与网络设施之上，以保护算力网络等关键基础设施。抗设备探测主要集迷惑与陷阱技术<sup>[26]</sup>、混淆技术<sup>[27]</sup>、频谱跳变等技术于一体，为算力节点提供抗测绘保护。抗设备探测也需要满足网络设施的抗测绘需求，网络设施抗探测技术包括分布式防御、弹性计算、边缘计算等技术，重点在于通过增强网络的动态性，提升测绘攻击的成本。抗设备探测技术由于其问题定义的具体性、防护目标的统一性，在保护算力网络层设备方面，具备隐蔽性、动态性、复杂性的优势，同时也存在两方面的不足，一是兼容性问题，针对数据基础设施的不同设备需要因地制宜、因势利导，并在安全性和设备性能之间进行权衡；二是有效性问题，随着新的探测手段不断出现，现有抗探测技术可

能面临失效或被绕过的风险。

### (二) 关键数据节点抗拓扑发现技术

抗拓扑发现旨在隐藏真实拓扑结构，防止攻击者通过观测不同数据要素之间的连接关系复现数据基础设施的整体架构。抗拓扑发现技术可用于隐藏数据基础设施结构，包括算力节点设施和数据要素平台层中数据的流通、汇聚过程。算力网络的抗拓扑发现技术手段包括虚拟网络<sup>[28]</sup>、动态路由、流量混淆<sup>[29]</sup>、分段与隔离等技术，通过隐藏网络物理拓扑、动态改变网络路由、引入混淆等手段增加拓扑探测的难度。数据要素流通与汇聚的抗拓扑发现包括动态路由、分段与隔离等技术，可根据算力网络层或数据要素层应用场景的不同而调整。当前抗拓扑发现技术已具备了隐蔽性强、实时响应能力强、多层次防护能力等，但也存在配置复杂的问题，尤其是动态路由和虚拟拓扑等技术，无形中增加了网络配置和管理的复杂性，也增加了网络监控和故障诊断的难度。

### (三) 跨域隔离抗关联技术

抗跨域关联技术如图3所示，主要是防止攻击者在数据服务提供过程中，通过关联不同域（系统、平台、数据源）之间的服务数据来获取数据基础设施的地理空间、网络空间以及社会空间属性。抗跨域关联是数据服务抗测绘的重点，需要防止攻击者关联服务接口、请求内容和服务处理逻辑等信息，进行测绘与攻击。当前抗跨域关联方法主要包括数据屏蔽、查询混淆、分段检索、IP黑/白名单<sup>[30]</sup>、数据匿名化与模糊化<sup>[31]</sup>等一系列技术。值得注意的是，随着大模型技术的发展，人工智能逐渐在数据基础设施中占据重要地位，可以提供一系列智能数据服务，帮助用户降本增效。然而，大模型的应用往往涉及海量的用户训练数据与交互信息、模型参数等用户隐私信息与敏感数据，智能数据服务的安全对数据服务抗跨域关联技术提出了更高要求。比如，在对以海量数据为基础的数据服务实施“抗跨域关联”时，需要考虑不同域、不同源数据的多样性，为各种数据提供恰当的防护手段。

## 四、数据基础设施抗测绘技术面临的共性问题和挑战

在数据基础设施的抗测绘技术体制中，抗设备

探测、抗拓扑发现与抗跨域关联三大攻防角度并驾齐驱，在提升数据基础设施安全性与可靠性方面发挥重要作用。但是，数据基础设施具有诸多新的特点，导致抗测绘技术面临一些共性难题。这些共性难题在不同形态数据基础设施的抗测绘研究上表现各异，而传统的网络安全技术在测绘安全问题上缺乏强有力的针对性应对能力，对数据基础设施抗测绘研究构成了严峻挑战。

### (一) 数据基础设施抗测绘的共性问题

数据空间的可测性由数据生产要素参与社会生产经营活动的需求决定，是网络空间开放性、数据资源固有可测性与测绘行为隐性性综合作用的结果。数据空间内分布的数据基础设施、应用与服务参数、对象属性与行为日志、模型数据、数据分析挖掘应用等数据实体和属性以及它们之间的关系、作用与行为模式构成数据空间资源集，按照其存在形态与时空性质又可以划分为实体资源、虚拟资源和虚实混合资源。数据空间依赖这些资源对外提供持续、稳定、可用的服务，而服务属性要求数据空间持续对外暴露地址、协议、通信时延等信息与对外产生交互的接口，使得数据空间具有难以避免的可测性。换言之，数据空间完全不可测要求当且仅当数据空间终止一切空间内系统或模型与空间外主

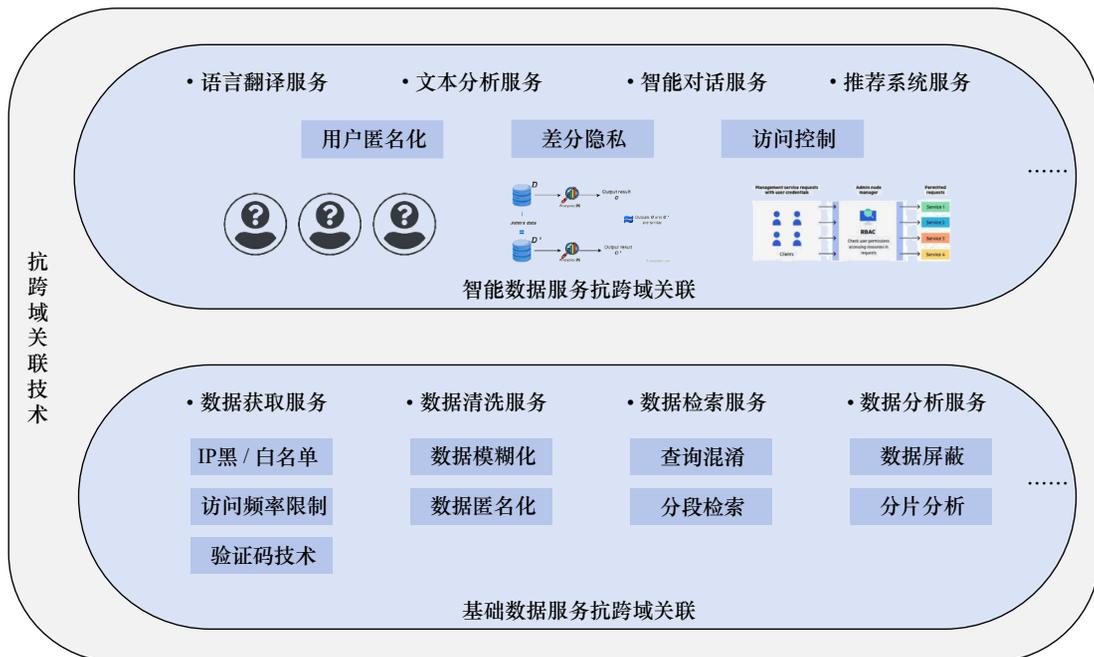


图3 抗跨域关联技术示意图

体的作用和信息流，成为与自然界物理空间和人类社会空间隔绝的“孤岛”。只要数据空间依托物理实体存在，在人类社会中寻找实践应用与社会历史价值实现，数据空间就具有固有的可测性。造成数据空间资源固有可测性的具体因素可以归纳为如下几个方面。

### 1. 数据可达性

数据空间内存在的数据需要能够为数据空间内的对象以及数据空间外访问数据空间服务与功能的实体所用。因此作为社会生产要素、在人类社会中具有使用价值的虚拟数据，对于数据空间内和跨空间的存在必然在逻辑层面可达。由于数据同步与传输依托数据基础设施间的存储媒介运输、有线光电缆、无线电波、量子纠缠等物理通信机制，因而逻辑层面的可达性决定了基础设施在物理空间中的连通性。故而，能够作为生产要素参与社会再生产的虚拟数据与数据基础设施，在物理层与逻辑层都具有固有的可达性。

### 2. 属性稳定性

为保证数据空间的服务属性，数据空间内的资源在可达的基础上，还需要保证服务的持续可用。数据服务的持续性要求数据存储设施与传输链路在空间上维持相对稳定的介质种类、位置、结构等物理参数与逻辑拓扑关系，同时其属性在连续的时间跨度上不变或呈现出可控的、可预测的、有限范围内的变化。在此基础上，持续可用的服务还要求其在认知域上的相对稳定性，即提供服务的对象和服务内容在一段时期的不同时刻要能以相同或相似的处理逻辑认知和利用。服务属性不要求数据空间各维度的特性严格不改变，但能够在人类社会生产、生活中解释和应用的数据，其特征在空间、时间和认知上整体是相对稳定的。

### 3. 跨域耦合性

数据域与物理域、认知域紧密耦合、不可分割，而这种耦合关系成为测绘方“发现、测量、收集、分析”认知闭环的桥梁，为测绘方源源不断地提供“顺藤摸瓜”式测绘启发。一方面，根据唯物论的观点，数据域中的资源无法摆脱对客观存在的依赖，数据信息本身要以字符、电位、自旋态等形式记录在存储媒介中，数据的交流传输也需要依靠物理设备和链路；另一方面，数据作为人类社会活动的产物，需要在特定历史背景中寻求解释，而数

据中记录的人类认知也需要归宿于人类历史实践。简而言之，数据域与物理域、认知域的耦合分别为数据空间提供了存在基础与价值基础，作为人类生产活动空间之一的数据空间必然具有跨域的耦合性。

数据基础设施测绘不同于网络攻击，它不以渗透、提权、控制和破坏为直接目的，而是利用数据基础设施的可达性、稳定性和关联性等特点，通过正常的网络与数据服务手段，获取所需的信息。这种测绘行为隐蔽性强，难以察觉且防御难度高，传统的网络安全技术往往难以有效应对。

## （二）数据基础设施抗测绘面临的挑战

面对数据基础设施的可测性及其带来的安全隐患，抗测绘技术的研究变得至关重要。现行的网络安全技术在测绘领域往往难以奏效，而传统的网络空间抗测绘中的抗探测、抗发现和抗关联等技术在数据基础设施环境中发挥出的效果也不尽如人意。数据基础设施的复杂性和多样性使传统技术无法针对特殊的安全需求提供有力的防御措施。

1. 数据基础设施的动态性和多变性增加了抗测绘技术的复杂性

数据基础设施不仅涉及到固定的物理设备和传输链路，还包括虚拟化和云计算等新兴技术，这些技术的动态特性使测绘行为更难以预测和防范。同时，虚拟化和云计算环境中的资源调度和管理也对抗测绘技术提出了更高的要求。

2. 数据基础设施的跨域耦合性增加了抗测绘的难度

数据基础设施不仅与物理空间紧密结合，还涉及到认知空间和社会空间，这种多域耦合性使攻击者可以通过跨域关联的方法，利用一个域中的信息获取其他域中的敏感数据。这种跨域关联性的存在，极大地增加了测绘行为的复杂性和隐蔽性。

3. 数据基础设施服务的持续性和高可用性要求抗测绘技术不能影响正常的服务质量

抗测绘技术需要在保障服务稳定和用户体验的前提下，实现对测绘行为的有效防御，这对技术的性能和可靠性提出了很高要求。传统的防护措施往往会对系统性能产生负面影响，难以满足这一要求。

4. 测绘技术的不断演进和发展使抗测绘技术面临持续挑战

随着大数据、人工智能和物联网技术的发展, 测绘技术也在不断进步, 攻击者可以利用更加先进的手段进行测绘, 传统的防护措施需要不断更新和优化, 才能有效应对新的威胁和挑战。

综合来看, 数据基础设施抗测绘技术的研究具有重要意义和紧迫性。系统开展数据基础设施抗测绘理论和技术研究, 对维护数据空间安全和提升数据基础设施的安全性、可靠性具有重要作用。这不仅是技术上的挑战, 更是关系到数据安全和隐私保护的重要课题。通过不断创新和优化抗测绘技术, 才能构建一个更加安全和可靠的数据基础设施环境。

## 五、基于内生安全的数据基础设施抗测绘理论分析

我国从2013年起就开始研究实现数字产品设计的内生安全理论与拟态防御技术<sup>[32,33]</sup>, 2016年率先完成了不依赖先验知识的拟态防御原理验证系统。内生安全理论基于“构造决定安全”的思想, 运用“结构编码”形成的环境加密“双盲效应”, 可以不依赖(但可融合)关于攻击者的先验知识和附加安全措施, 实现高可信、高可靠、高可用三位一体的内生安全网络弹性设计<sup>[34,35]</sup>。从内在机理看, 数据基础设施测绘本质上是对各种结构的再现和表达。由此, 数据基础设施抗测绘的核心问题是, 对“结构”的隐匿或抗发现。基于内生安全的数据基础设施抗测绘理论和技术, 旨在运用内生安全理论方法, 突破现有网络体系架构下数据基础设施和资源固有可测性制约这一难题, 探索抗测绘的新路径和新方法。

### (一) 数据空间内生安全

信息世界是客观存在于人类社会历史中的抽象存在, 而数据空间是信息世界的重要组成部分。在当今人类社会, 数据已然成为信息记录与传输的最主要媒介, 数据空间作为支撑信息世界的基石, 正发挥着越来越重要的作用。数据空间的基础性作用决定了数据空间安全研究的重要性与战略意义。然而在安全防御实践中, 不确定性或“未知的未知”威胁成为新挑战, 在缺乏先验知识的条件下, 各种

“外挂”或附加式安全产品难以从根本上避免或杜绝跨越数据安全防御边界条件的攻击。数据空间安全研究困境的成因在于现有安全范式的局限性, 在现有范式下的安全方案与产品, 无论其设计多么精细完善, 总囿于先验依赖的思维视角与“附加防御”方法论的限制, 未能深入到数据空间的构成要素中, 故而在面临日趋繁杂的“未知的未知”威胁类型时难以有效应对。为了解决数据空间安全研究困境, 需要将内生安全理念注入到数据空间安全体系中, 强调从系统设计之初就将安全性融入到系统的架构、机制、场景和运行规律等各个层面, 而非仅仅依赖于附加的外挂式安全组件, 以此构建数据空间安全新范式。数据空间内生安全改变了传统的被动式防御模式, 从外挂到内生, 从被动到主动, 赋予数据空间内的模型与系统自主应对各种已知和未知威胁的能力, 实现“开箱即用”的默认安全效果。

### (二) 研究构想

基于内生安全的数据基础设施抗测绘研究, 重点是针对数据空间边界、数据设施和资产的天然可测、数据服务多样化等特性, 厘清抗测绘机理, 提出基于动态异构冗余构造的抗测绘理论, 通过研究“结构编码”“环境加密”“工程化融合”等技术方法, 赋予数据基础设施内生抗测绘能力, 降低受保护数据基础设施的被探测概率, 阻止测绘方获取到全面、准确、真实、实时、可用的测绘结果, 实现数据基础设施全方位、长效、普适的抗测绘能力。与一般意义上的针对数据基础设施的攻击类型不同, 测绘攻击不直接破坏攻击目标、致使其异常或失效, 而是窃取攻击目标的内容、属性、统计特征等信息, 故而难以被防御方直接地、实时地、全面地感知。尽管测绘攻击行为不具备直接破坏性, 测绘方对攻击目标信息的非法获悉和利用, 将对目标安全构成作用层次更深、潜在危害性更大的威胁。与传统抗测绘技术相比, 内生安全赋能的抗测绘技术具有三个方面的效能增益。一是具备应对潜在深层安全威胁的能力, 通过动态异构冗余构造, 保护数据基础设施的属性信息不被测绘方发现、收集和分析, 使测绘方无法测得可认知、可利用的信息。二是能为不同应用场景赋能, 针对工业互联网、物联网、第六代移动通信、卫星网络等不同场景, 以及未来可能出现的各种新兴应用场景, 给出场景通

用的内生安全架构与场景适应的内生安全抗测绘构造方法。三是具有自主进化特性，数据基础设施抗测绘能力能够自主动态提升，在测绘 / 抗测绘的演进和博弈中形成自我发现、自动修复和自主平衡能力。内生安全的数据基础设施抗测绘构造能够针对数据基础设施固有的可测机理来提供测绘防护，使测绘安全成为数据基础设施的一个自然属性，通过系统内在安全机制实现更为主动、灵活和持久的“测不了、测不准、绘不对”防御效果。

### (三) 抗测绘机理分析

数据基础设施抗测绘关注的非破坏性、非感知性测绘攻击行为，与内生安全问题中不易察觉、难以预测的“未知的未知”问题在内质与外显形态上高度契合，因而将内生安全引入抗测绘研究，开展内生安全的数据基础设施抗测绘研究，有助于开辟数据基础设施抗测绘的新路径。图4给出了具有内生安全特性的数据集成设施抗测绘机理。

数据空间开放性要求数据基础设施具有服务属性，具体表达为服务可达性、属性稳定性、跨域耦合性，从而使数据基础设施及其资源必然能够从物理域和认知域出发，经由相对稳定的线路和方法被人类访问、解释和利用。故而，对于非内生安全设计的数据基础设施，无论其防御措施如何精密，测绘者总能利用其服务属性，通过蛛丝马迹获取到数据空间内高价值目标的测绘信息。数据基础设施的固有可测性无法通过有限的设计规则以链式、旁

路、分支等作用逻辑简单组合消除或规避，需要以内生安全的视角对数据空间防御框架进行“釜底抽薪”式重构，将动态、异构、冗余的构造方法作用于数据空间要素，以内生测绘安全打破固有可测性的制约。内生安全思想为数据基础设施赋予动态性、异构性和冗余性，以扰动和阻止针对数据基础设施的测绘攻击。在数据基础设施中，其动态特性可以有效阻止测绘方通过试错攻击和渗透测试获取信息，降低数据空间的可感知性和可预测性，打破测绘方对目标的认知闭环。异构的数据存储与访问方式增加了数据要素之间、数据要素与外部实体之间的关联复杂度，从而提高了访问和测量的成本，降低了其可达性和属性稳定性。异构性是将设备或服务的功能作为样本，采用不同的实现方式构造出功能等价的异构单元。这些异构单元以冗余化的方式运行，并由动态调度逻辑选择合适的异构单元投入使用，从而在数据空间中提供服务。此外，冗余化的功能和部件通过统计学上的扰动作用进一步增加了测绘难度，降低了特定目标被探测的概率，阻碍了测绘方获取准确、全面的信息。同时，拟态控制基元确保了拟态化系统对用户的透明性，使拟态化系统在服务维度上与非拟态化系统无异。因此，内生安全构造能够契合数据空间抗测绘研究非破坏性、非感知性测绘行为的需求，以内生的安全构造扰动和阻止测绘攻击，从而突破数据资源的固有可测性。动态异构冗余构造对数据资源固有可测性的突破是全方位的，其对服务可达、属性稳定和跨域

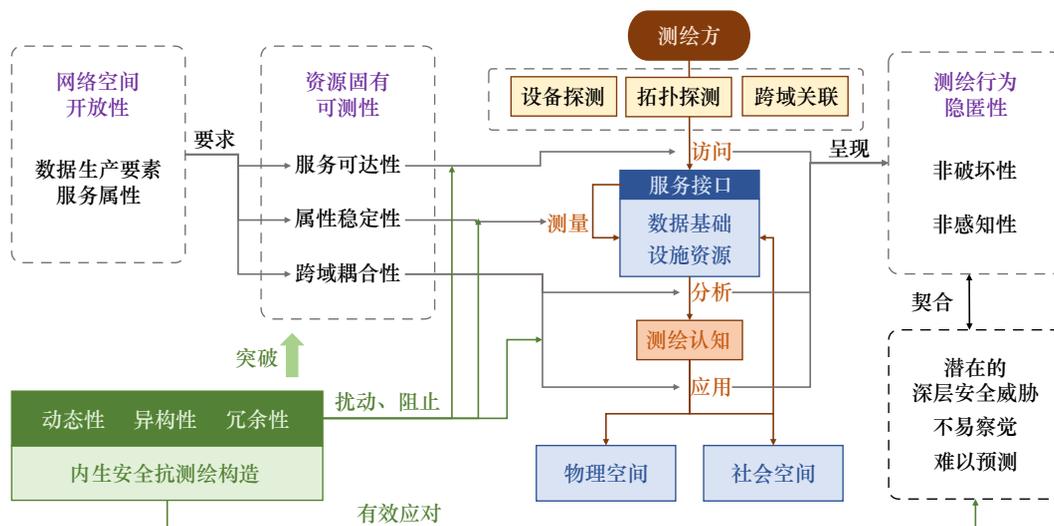


图4 内生安全的数据基础设施抗测绘机理

耦合的抑制效应，在数据资产、数据基础设施运行环境和抗测绘技术协同运作中均能发挥出良好效能，具体表现为以下三个方面。

数据资产结构编码化。动态异构冗余（即 DHR）架构作为一种设计安全框架<sup>[9]</sup>，可以直接应用于各类数据基础设施。对于需要保护的数据基础设施硬件、链路或应用服务等各种数据资产，以其自身功能为样本，以不同实现方式构造出功能等价的异构单元，将这些异构单元冗余化运行，并配以相应的动态编码逻辑选择投入使用的异构单元，并组成瞬态拓扑结构对外提供服务。在结构编码化的数据基础设施中，动态性能降低结构时间轴的耦合性，减少不同时期对结构观测的拼接成功率，达到“测不准”的目标；异构的数据存储与访问方式能够复杂化数据空间内要素之间以及数据空间内要素与空间外实体的关联模式，从而增加数据空间目标的访问测量成本，降低其可达性与属性稳定性；冗余化的功能和部件也可以在统计学意义上对测绘方产生扰动，动态配置的冗余化异构部件能够进一步增加测绘难度，降低特定目标的被探测概率，阻碍测绘方获取准确、全面的测绘认知。由此，将数据基础设施置于动态异构冗余的控制框架中，能够在保证服务属性的同时降低数据基础设施的服务可达性、属性稳定性和跨域耦合性，从而突破数据基础设施的固有可测性，提升数据空间抗测绘安全性。

运行环境密文化。不依赖先验知识的 DHR 架构具有密码学意义上的结构加密性质（凡是不依赖先验知识的安全防御技术都具有加密性质），这使得 DHR 架构赋能的内生安全目标系统，除了本征服务是透明的之外，其他的非期望功能都会被结构加密屏蔽。理论上，当 DHR 构造内出现差模或有感共模状态时，当前运行环境将作类似“一次一密”的“一次差模一次重构”的“完美安全”变换。攻击表面除了以明文形态呈现的本征服务恒定不变外，当前运行环境内所有可利用的软硬件资源，包括防御者未知或攻击者已知的漏洞后门、病毒木马等，也会因为不知原因的“一次差模扰动”而无一例外地再次实施“结构变换”操作。数据运行环境密文化从结果上让攻击者“致盲”，通过构造加密环境减弱原始数据的可达性，在保证数据本征服务的同时，解除数据与可认知信息域的耦合，从

而保护数据基础设施不被测绘方探测、分析和利用。可以看出，内生安全在数据基础设施抗测绘研究中具有高度普适性和有效性，DHR 框架既能够作用于数据基础设施，也能够用以组织数据抗测绘算法，在提升测绘安全的同时也能增强数据基础设施的功能健壮性、算法正确性与抗破坏能力。内生安全的抗测绘思路并非动态性、异构性与冗余性的机械叠加，而是要求系统构造或算法同时具备三者的完全相交表达，此时系统基于构造的内源性效应，即使在缺乏先验知识的前提下，仍能管控构造内基于任何未知漏洞后门或未知测绘方式的差模测绘攻击。在动态性、异构性和冗余性完全相交表达的作用下，数据基础设施构成要素能够规避服务连通性、属性稳定性和跨域耦合性带来的测绘风险，突破数据基础设施固有可测性的制约，具备内生的测绘安全属性，为现有防御范式存在的认知安全防护不全面、资产适应性不足、局限于单一业务场景以及静态依赖性等问题提供解决思路。

抗测绘技术多样化融合。DHR 安全机制的有效性虽然不依赖（但不排除）其他抗测绘措施或技术手段，但在架构内通过合理编排或配置不同安全机理的抗测绘安全产品，可以获得指数量级的安全增益，这是 DHR 构造固有的融合特性使然。如同“钢筋骨架”“混凝土砧料”与建筑物强度和韧度之间的关系，钢筋构型、材料性质以及混凝土砧料配方和搅拌工艺等决定期望的建筑质地。DHR 架构在这里起到的是“钢筋骨架构型”的作用，倘若再按需选择诸如钢筋材料、混凝土砧料与搅拌工艺之类的传统抗测绘技术，就是在运行环境的执行体或场景中策略性、差异化的配置一些对应用服务无感的附加式或嵌入式的抗测绘措施，可等效提升运行环境内的异构度，因而通过 DHR 架构特性获得指数量级的安全增益。

## 六、基于内生安全的数据基础设施抗测绘理论关键技术与实践

### （一）关键技术

内生安全的数据基础设施抗测绘研究既包含于内生安全研究，也包含于抗测绘研究，通过内生安全机制抵御针对数据基础设施的测绘攻击，阻止测绘方对数据空间进行准确、全面、实时的认知，赋

予数据基础设施内生的认知安全性。内生安全的数据空间抗测绘研究域如图5所示。

数据基础设施内生安全研究强调将安全机制融入到数据基础设施的系统或模型设计中，以减少对外部安全措施的依赖，使安全成为数据基础设施系统架构不可分割的一部分。在抗测绘领域，内生安全技术强调通过增强数据基础设施自身的抵抗能力和恢复能力来对抗数据基础设施测绘。接下来，详细介绍几种关键的内生安全数据基础设施抗测绘技术，结合被动防御与主动防御两类防护思路，针对

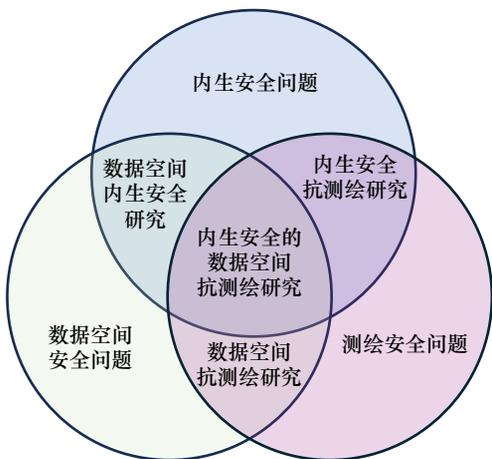


图5 内生安全的数据空间抗测绘研究域

测绘方与数据基础设施中的关键环节，分别给出应对方法。这些技术不仅体现了内生安全理念的实践应用，而且展示了如何将安全措施与数据基础设施的构建和运作深度融合。

抗测绘研究思路可以从测绘方入手，通过识别、拒止测绘方的测绘行为，或者对测绘行为加以诱导和欺骗，达到阻止测绘方探到、测到相应信息。同时，也可以从数据基础设施自身的安全性能入手，隐蔽数据基础设施在数据空间中的坐标，通过设备组织形态变换、等价异构代码生成、可信加密执行环境等技术干扰测绘者的测绘结果，降低数据服务对测绘者的可达性，达到“测不到”的目的；在稳定的数据服务中引入针对测绘者的属性扰动，使测绘者“测不准”“测不全”；隐藏或混淆数据域与物理域、社会域的耦合，为测绘者跨域关联映射增加难度，阻止测绘者绘制出正确的、可用的数据基础设施综合画像，进而避免测绘者的进一步深层次攻击，达成“绘不对”的效果。据此，可以将常用的抗测绘技术分为隐蔽、拒止、干扰、欺骗等类型，如图6所示。此外，动态异构冗余思想中的多模裁决向概率空间推广的方法，以及借助深度学习实现自主防御的抗测绘框架，也是在防御实践中得到检验的行之有效的抗测绘思路。

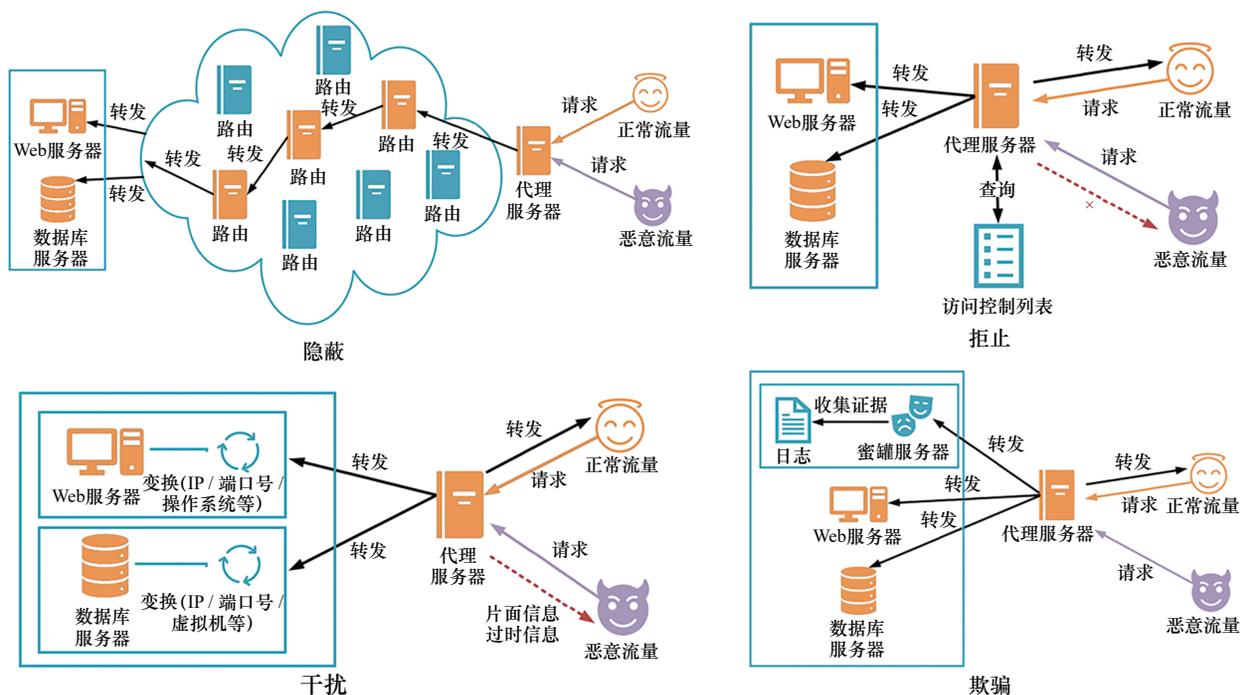


图6 常见的抗测绘思路

测绘行为感知与拒止。探测行为检测旨在识别测绘方的测绘流量与非授权访问行为，基于测绘指纹库、源IP地址与生存时间（TTL）、访问行为特征等维度判断对数据基础设施的服务请求是否为测绘行为。在此基础上，对识别出的测绘行为以防火墙、子网隔离、蜜罐等手段拒绝响应或作出虚假响应的方法就是测绘行为拒止。目前，主流的测绘行为感知与拒止响应以基于规则的匹配和生成方法为主流。在内生安全的数据基础设施抗测绘研究中，需要将基于规则的方法与基于模型推理和生成的方法结合，并根据感知反馈更新规则库与模型参数，实现灵活持久的主动测绘防护。

测绘干扰与数据欺骗。测绘干扰指以动态变化的异构运行环境干扰测绘方判断，使测绘方难以将测绘结果信息转化为可理解、可利用的测绘认知。测绘欺骗指通过在数据基础设施内部署诱饵、虚拟化等手段，引诱测绘方得出错误的测绘结论。测绘干扰与数据欺骗技术同时作用于测绘方与被保护的数据基础设施，结合主动与被动两种防护思路，通过动态异构冗余架构降低测绘结果的真实性、全面性与时效性，达到测不全、测不准的效果。

面向概率空间的拟态裁决方法。在传统的DHR架构中，拟态裁决通过策略性地集成多个独立执行体的输出，进而决策出整个系统正确的输出<sup>[35]</sup>。在数据基础设施抗测绘中，拟态裁决负责做出正确的抗测绘决策。传统的拟态裁决方法主要依赖确定性

裁决算法来确保抗测绘系统的稳定性和可靠性。确定性的裁决方法要求抗测绘系统在面对相同的输入时，可以给出一致的抗测绘决策。然而，随着数据空间的不断演化与扩展，数据空间基础设施在处理相同输入时，往往表现出结果的非一致性，并将这种非一致性蔓延到数据空间基础设施的抗测绘决策中。为了适应这一挑战，需要采用概率拟态裁决方法。相较于传统拟态裁决方法，概率拟态裁决方法主要包括相似度计算和置信度计算两个关键步骤；相似度计算旨在通过量化手段，有效衡量不同抗测绘决策间的相似程度。相似度计算不仅为决策提供了客观的相似性基准，还为后续的置信度计算奠定了基础。置信度计算则是通过置信度计算模型将不同抗测绘决策的相似度转化为每个抗测绘决策的可信度。

自学习主动抗测绘框架。测绘与抗测绘是动态博弈的过程，自学习主动抗测绘框架通过对测绘的学习，实现抗测绘能力的持续提升和自我完善，如图7所示。主动抗测绘策略模块负责根据探测流量的类型，选择并应用合适的主动抗测绘策略。在线抗测绘模块的所有执行结果将被裁决器接收，并运用裁决算法给出最终的判断。对于正常的服务请求流量，框架会将其转发至数据基础设施，以确保服务的正常提供。对于探测流量，框架一方面生成虚假响应，以误导测绘方的知识库，另一方面则收集这些探测流量，构建测绘流量知识库，为学习模块

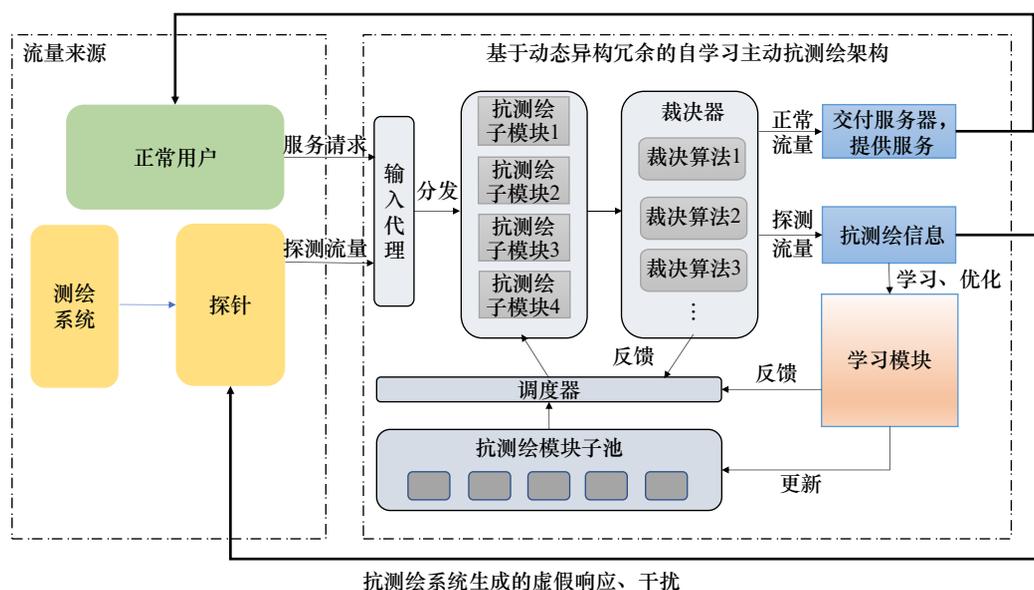


图7 基于动态异构冗余的自学习主动抗测绘架构

提供数据支持。最后，学习模块将不断从测绘流量知识库中学习探测流量的特征，并据此更新和优化抗测绘模块池中的模块，以及调整调度器的调度策略，从而实现对抗测绘能力的持续提升和自我完善。

### （二）实践探索

在数据空间安全实践中，我们对内生安全抗测绘做出了一些前期探索，在Web服务、大模型的训练数据等场景下取得了一些进展。

DHR构造的Web服务器抗测绘。Web服务在数据空间中扮演着至关重要的角色，它不仅提供了数据的访问、共享、交换和集成的核心平台，而且极大地促进了数据的高效利用和价值最大化。一旦Web服务的数据被非法测绘，轻则造成用户隐私泄露，重则可能泄露企业和国家机密。为了应对这一挑战，我们提出了一种基于DHR架构的Web服务器抗测绘方法。该方法将传统的单点式服务器拓展为DHR架构的服务器集群；当面临测绘攻击时，该服务器集群能够利用其动态异构冗余的特性，实现快速响应和自我修复。通过对服务器集群中的节点进行动态上线、下线，既可以使部分节点遭受测绘，也可以通过调度使正确测绘信息快速过时。同时，异构性允许系统针对测绘攻击提供多样化的回应，混淆测绘方捕获的信息。此外，冗余设计确保了关键组件和服务的备份，即使在破坏性攻击下也能迅速恢复，减少系统停机时间。这种基于DHR构造的Web服务器抗测绘方法，不仅提高了系统的抗测绘能力，也为用户提供了一个更加安全、可靠的数据服务平台，有效保护了企业和国家的机密信息，维护了数据空间的认知安全。

大模型训练数据的抗测绘。在大语言模型领域中，大模型训练往往依赖于对海量网络数据的深度挖掘和分析，这一过程本质上也是对网络数据的一种测绘。然而，这种数据测绘往往未经网络用户明示同意，不仅侵犯了用户的隐私权，还容易导致敏感信息的泄露和知识产权的侵犯。为了应对这一挑战，我们提出了一种训练数据抗测绘的方法，旨在从根本上限制和阻止未经授权的训练数据测绘行为。基本思路是首先利用多模态预训练的大模型对网络中的图像数据进行深度的特征提取。这些大模型通常包括卷积神经网络（CNN）和Transformer

架构，能够有效地从图像中提取出丰富的特征信息。接下来，利用聚类算法将这些特征划分为不同的数据簇。常用的聚类算法包括K-均值聚类、层次聚类和谱聚类等，这些算法能够将相似的特征数据聚集在一起，形成明确的数据簇。在此基础上，引入了一种新的不可学习噪声生成技术，这些噪声专门设计用来破坏数据簇的内在结构和模式，使数据簇对于机器学习模型而言变得不可识别、不可学习。这种噪声生成技术可以通过添加随机噪声、噪声扰动和噪声混淆等手段，对数据簇进行干扰，从而使机器学习模型无法准确识别和学习这些数据簇的模式。通过这种训练数据抗测绘的方法，不仅能够有效地抵御有监督学习、无监督学习以及其他各种数据利用手段，还能够绕过商业机器学习平台的数据测绘。

## 七、数据基础设施抗测绘发展建议

目前，数据基础设施抗测绘理论与技术还处于探索阶段，下一步需要继续深化研究和实践，重点包括内生安全原生的数据服务工程开发范式、高效的服务变换技术、数据基础设施的抗测绘评价体系等。鉴于数据基础设施抗测绘对维护数据空间安全、保证我国数字经济健康可持续发展等具有重要意义，就推进发展提出如下建议。

一是尽快开展数据空间抗测绘基础理论和关键技术研究。当前开展的数据基础设施抗测绘研究，只是数据空间抗测绘研究的先导课题。数据空间抗测绘与网络空间抗测绘、地理空间抗测绘在内涵、特点、规律上有明显不同，比数据基础设施抗测绘的范畴更大、复杂性更强。数据空间抗测绘是数据空间安全的基本问题，建议加强这一领域的基础理论研究，为维护数据空间安全提供更有力的支撑。

二是将抗测绘能力纳入国家数据基础设施建设基本要求和基础能力之中。伴随着数据要素化的进程加快，我国数据基础设施建设也将进入快速发展阶段。作为一类重要的基础设施，抗测绘应该成为其基本属性和基础能力。建议在我国数据基础设施建设的筹划部署阶段，就同步考虑抗测绘能力的建设问题，把抗测绘作为基本的建设要求，在“起跑线”上系好“安全带”，避免出现被“单向透明”的局面。

三是加快数据基础设施关键技术研发和攻关力度。目前,我国还没有专门针对数据基础设施抗测绘技术的科研计划安排,相关技术研究还是延续网络空间安全或者网络资源抗测绘的攻关计划。建议国家在数据要素专项工程、数据基础设施建设项目安排上,同步加大数据基础设施抗测绘技术的支持力度,在赋能数据基础设施形成内生抗测绘能力上加强投入,筑牢数据基础设施安全屏障。

四是高度关注新兴技术带来的被测绘风险。近一段时间,大模型和人工智能技术蓬勃发展。研究表明,大模型技术对数据基础设施抗测绘带来了新的挑战。一方面,大模型作为一类数据基础设施的集合,本身存在模型结构及参数泄漏、训练数据集泄漏等风险;另一方面,大模型和人工智能技术为跨境关联提供了新的方法,使抗测绘难度更大。为此,需要加强在大模型技术背景下新型抗测绘技术的研究。

五是加强数据基础设施抗测绘产业的扶持力度。我国在数据资产、网络资产抗测绘技术产业化上具有一定积累,相关产业、行业也得到了发展。当前,建议借助国家数据要素产业发展的重大机遇,借助数据基础设施建设的新潮流,重点培育数据基础设施抗测绘相关产业的发展,通过政府引导、金融支持、场景创设、标准加持等方式,培育新兴产业,通过产业发展促进技术进步与应用。

六是推动多学科交叉融合和产教融合。数据基础设施抗测绘是典型的学科交叉领域,涉及测绘科学、网络科学、数据科学、认知科学等范畴,需要不同学科力量之间的有机融合,仅靠单一学科无法解决其面临的复杂、多样的科学技术发展挑战。应在这一领域设立高校、企业、研究机构共同参加的“协同创新体”或“学术共同体”,集合国家优势力量共同解决我国数据空间面临的重大现实问题,为以数据为关键要素的数字经济高质量发展提供技术支撑。

#### 利益冲突声明

本文作者在此声明彼此之间不存在任何利益冲突或财务冲突。

**Received date:** September 23, 2024; **Revised date:** November 22, 2024

**Corresponding author:** Wu Jiangxing is a professor from School of Computer Science, Fudan University, and a member of Chinese Academy of Engineering. His major research fields include computer network and cyberspace security. E-mail: jxwu@fudan.edu.cn

**Funding project:** National Key R&D Program Project (2022YFB3102901); Chinese Academy of Engineering projects “Research on the Development Path and Technical System of National Data Space” (2024-XBZD-05) and “Research on the Development Strategy of National Data Space” (2023-XBZD-16)

#### 参考文献

- [1] 李国杰,程学旗. 大数据研究: 未来科技及经济社会发展的重大战略领域——大数据的研究现状与科学思考 [J]. 中国科学院院刊, 2012, 27(6): 647–657.  
Li G J, Cheng X Q. Research status and scientific thinking of big data [J]. Bulletin of Chinese Academy of Sciences, 2012, 27(6): 647–657.
- [2] 程啸. 论大数据时代的个人数据权利 [J]. 中国社会科学, 2018 (3): 102–122, 207–208.  
Cheng X. Personal data rights in the era of big data [J]. Social Sciences in China, 2018 (3): 102–122, 207–208.
- [3] 许宪春,任雪,常子豪. 大数据与绿色发展 [J]. 中国工业经济, 2019 (4): 5–22.  
Xu X C, Ren X, Chang Z H. Big data and green development [J]. China Industrial Economics, 2019 (4): 5–22.
- [4] 常晓素. 大数据在税收风险管理中的应用探析 [J]. 税务研究, 2019 (6): 78–81.  
Chang X S. Application of big data in tax risk management [J]. Taxation Research, 2019 (6): 78–81.
- [5] 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.  
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.
- [6] 陈涛,程丽君,李明桂,等. 网络空间测绘系统及应用研究 [J]. 通信技术, 2020, 53(11): 2832–2837.  
Chen T, Cheng L J, Li M G, et al. Cyberspace surveying and mapping system and application [J]. Communications Technology, 2020, 53(11): 2832–2837.
- [7] 陈钟,孟宏伟,关志. 未来互联网体系结构中的内生安全研究 [J]. 信息安全学报, 2016, 1(2): 36–45.  
Chen Z, Meng H W, Guan Z. Research on intrinsic security in future Internet architecture [J]. Journal of Cyber Security, 2016, 1(2): 36–45.
- [8] 张伟丽,贺磊. 关于新型内生安全信息基础设施的思考 [J]. 无线电通信技术, 2020, 46(4): 399–404.  
Zhang W L, He L. Consideration of new endogenous security information infrastructure [J]. Radio Communications Technology, 2020, 46(4): 399–404.
- [9] 江伟玉,刘冰洋,王闯. 内生安全网络架构 [J]. 电信科学, 2019, 35(9): 20–28.  
Jiang W Y, Liu B Y, Wang C. Network architecture with intrinsic security [J]. Telecommunications Science, 2019, 35(9): 20–28.
- [10] 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.  
Zhou Y, Xu Q, Luo X Y, et al. Research on definition and technological system of cyberspace surveying and mapping [J]. Com-

- puter Science, 2018, 45(5): 1–7.
- [11] 刘红, 姚旺君, 孙彻, 等. 网络空间测绘系统分类及应用综述 [J]. 信息技术与网络安全, 2021, 40(10): 16–21, 28.  
Liu H, Yao W J, Sun C, et al. Classification and application of cyberspace surveying and mapping system [J]. Information Technology and Network Security, 2021, 40(10): 16–21, 28.
- [12] 宋苑, 卢扬明. 网络安全扫描技术综述 [J]. 广东通信技术, 2004, 24(8): 58–60, 64.  
Song Y, Lu Y M. Overview of network security scanning technology [J]. Guangdong Communication Technology, 2004, 24(8): 58–60, 64.
- [13] 张义荣, 赵志超, 鲜明, 等. 计算机网络扫描技术研究 [J]. 计算机工程与应用, 2004, 40(2): 173–176.  
Zhang Y R, Zhao Z C, Xian M, et al. A study on computer network scanning techniques [J]. Computer Engineering and Applications, 2004, 40(2): 173–176.
- [14] 唐小明, 梁锦华, 蒋建春, 等. 网络端口扫描及其防御技术研究 [J]. 计算机工程与设计, 2002, 23(9): 15–17.  
Tang X M, Liang J H, Jiang J C, et al. Research about technology of port scan and port scan detect [J]. Computer Engineering and Design, 2002, 23(9): 15–17.
- [15] 冯贵兰, 李正楠, 周文刚. 大数据分析技术在网络领域中的研究综述 [J]. 计算机科学, 2019, 46(6): 1–20.  
Feng G L, Li Z N, Zhou W G. Research on application of big data analytics in network [J]. Computer Science, 2019, 46(6): 1–20.
- [16] 姜开达, 李霄, 孙强. 基于网络流量元数据的安全大数据分析 [J]. 信息网络安全, 2014, 14(5): 37–40.  
Jiang K D, Li X, Sun Q. Big data analysis on security based on network traffic metadata [J]. Netinfo Security, 2014, 14(5): 37–40.
- [17] 李晓会, 陈潮阳, 伊华伟, 等. 基于云计算和大数据分析的大规模网络流量预测 [J]. 吉林大学学报(工学版), 2021, 51(3): 1034–1039.  
Li X H, Chen C Y, Yi H W, et al. Large scale network traffic prediction based on cloud computing and big data analysis [J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 51(3): 1034–1039.
- [18] 许子明, 田杨锋. 云计算的发展历史及其应用 [J]. 信息记录材料, 2018, 19(8): 66–67.  
Xu Z M, Tian Y F. The development history and application of cloud computing [J]. Information Recording Materials, 2018, 19(8): 66–67.
- [19] 王勇, 周慧怡, 俸皓, 等. 基于深度卷积神经网络的网络流量分类方法 [J]. 通信学报, 2018, 39(1): 14–23.  
Wang Y, Zhou H Y, Feng H, et al. Network traffic classification method basing on CNN [J]. Journal on Communications, 2018, 39(1): 14–23.
- [20] 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.  
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.
- [21] Pan L, Yang J H, He L, et al. Your router is my prober: Measuring IPv6 networks via ICMP rate limiting side channels [R]. San Diego: Proceedings 2023 Network and Distributed System Security Symposium, 2023.
- [22] 包正晶, 苏马婧, 康彬, 等. 域名画像系统的设计与实现 [J]. 信息技术与网络安全, 2021, 40(6): 1–8.  
Bao Z J, Su M J, Kang B, et al. Design and implementation of domain Name portrait system [J]. Information Technology and Network Security, 2021, 40(6): 1–8.
- [23] 王进, 郝子龙, 温尚国, 等. 一种基于知识图谱的分布式云安全画像和风险预警模型研究 [J]. 通信技术, 2023, 56(10): 1184–1190.  
Wang J, Hao Z L, Wen S G, et al. A distributed cloud security profile and risk precaution model based on knowledge graph [J]. Communications Technology, 2023, 56(10): 1184–1190.
- [24] 高春东, 郭启全, 江东, 等. 网络空间地理学的理论基础与技术路径 [J]. 地理学报, 2019, 74(9): 1709–1722.  
Gao C D, Guo Q Q, Jiang D, et al. The theoretical basis and technical path of cyberspace geography [J]. Acta Geographica Sinica, 2019, 74(9): 1709–1722.
- [25] 郭启全, 高春东, 郝蒙蒙, 等. 发展网络空间可视化技术支持网络安全综合防控体系建设 [J]. 中国科学院院刊, 2020, 35(7): 917–924.  
Guo Q Q, Gao C D, Hao M M, et al. Develop visualization technology of cyberspace to support construction of comprehensive prevention and control system of cyber security [J]. Bulletin of Chinese Academy of Sciences, 2020, 35(7): 917–924.
- [26] 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.  
You J Z, Lyu 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.
- [27] 李凌书, 邬江兴, 刘文彦. SaaS云环境下基于容器指纹匿名的网络欺骗方法 [J]. 信息安全学报, 2022, 7(2): 72–86.  
Li L S, Wu J X, Liu W Y. An anonymous network deception method based on container fingerprint modification for SaaS applications [J]. Journal of Cyber Security, 2022, 7(2): 72–86.
- [28] 吴云坤, 姜博, 潘瑞萱, 等. 一种基于零信任的SDN网络访问控制方法 [J]. 信息网络安全, 2020, 20(8): 37–46.  
Wu Y K, Jiang B, Pan R X, et al. A SDN access control mechanism based on zero trust [J]. Netinfo Security, 2020, 20(8): 37–46.
- [29] 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.  
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.
- [30] 卢先锋, 杨频, 梁刚. 基于动态IP黑名单的入侵防御系统模型 [J]. 计算机工程与设计, 2011, 32(1): 10–13.  
Lu X F, Yang P, Liang G. Model of intrusion prevention system based on dynamic IP blacklist [J]. Computer Engineering and Design, 2011, 32(1): 10–13.
- [31] 王平水, 王建东. 匿名化隐私保护技术研究综述 [J]. 小型微型计算机系统, 2011, 32(2): 248–252.  
Wang P S, Wang J D. Survey of research on anonymization privacy-preserving techniques [J]. Journal of Chinese Computer Systems, 2011, 32(2): 248–252.
- [32] 白紫星, 戴华昇, 宋怡景, 等. 基于多内核的操作系统内生安全技术 [J]. 集成电路与嵌入式系统, 2024, 24(1): 58–63.  
Bai Z X, Dai H S, Song Y J, et al. Endogenous security technology based on multi-kernel operating system [J]. Integrated Circuits and Embedded Systems, 2024, 24(1): 58–63.

- [33] 聂凯君, 曹侯, 彭木根. 6G 内生安全: 区块链技术 [J]. 电信科学, 2020, 36(1): 21–27.  
Nie K J, Cao B, Peng M G. 6G endogenous security: Blockchain technology [J]. Telecommunications Science, 2020, 36(1): 21–27.
- [34] 邬江兴, 邹宏, 薛向阳, 等. 内生安全赋能网络弹性的构想、方法与策略 [J]. 中国工程科学, 2023, 25(6): 106–115.  
Wu J X, Zou H, Xue X Y, et al. Cyber resilience enabled by endogenous security and safety: Vision, techniques, and strategies [J]. Strategic Study of CAE, 2023, 25(6): 106–115.
- [35] 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.  
Wu J X. Research on cyber mimic defense [J]. Journal of Cyber Security, 2016, 1(4): 1–10.
- [36] 王伟, 曾俊杰, 李光松, 等. 动态异构冗余系统的安全性分析 [J]. 计算机工程, 2018, 44(10): 42–45, 50.  
Wang W, Zeng J J, Li G S, et al. Security analysis of dynamic heterogeneous redundant system [J]. Computer Engineering, 2018, 44(10): 42–45, 50.