工业互联网确定性内生安全防御技术综述

张志为[1,2]，汤贵源[1,3]，沈玉龙[1,2]，陈莹玉[1,2]，习宁[1,3]，任保全[4]，焦成伟[5]，何吉[1,2]，张涛[1,2]，马建峰[1,3]

（1.西安电子科技大学陕西省网络与系统安全重点实验室，陕西 西安 710126；2.西安电子科技大学计算机科学与技术学院，陕西 西安 710126；3.西安电子科技大学网络与信息安全学院，陕西 西安 710126；4.军事科学院系统工程研究院，北京，100091；5.华为技术有限公司，广东 深圳 518129）

摘要：针对工业互联网严苛运行约束与传统外挂式静态边界防护范式间存在的结构性失配问题，开展确定性内生安全防御研究。围绕系统内在的实体、行为、结构与环境四类确定性，提出了确定性内生安全防御范式；系统综述内生身份管理、任务行为确保、未知攻击研判、主动安全防御与安全效能评测五类关键使能技术，构建了涵盖理论基础、体系模型与四维立体化技术架构的内生安全防御体系。对比分析揭示了现有内生安全实践在跨路线协同与场景适配上的共性瓶颈；在智能制造与工业5G场景中分析表明，所提架构能够支撑从理论机制到工程防护的转化，以系统内在确定性对抗威胁不确定性，为解决工业互联网安全跨域联动与动态演化瓶颈提供可行路径。

关键词：工业互联网；确定性内生安全；任务行为确保；未知攻击研判

中图分类号：TP309

文献标志码：A

Survey on deterministic endogenous security defense techniques for industrial Internet

Zhang Zhiwei[1,2], Tang Guiyuan[1,3], Shen Yulong[1,2], Chen Yingyu[1,2], Xi Ning[1,3], Ren Baoquan[4], Jiao Chengwei[5], He Ji[1,2], Zhang Tao[1,2], Ma Jianfeng[1,3]

1. Shaanxi Key Laboratory of Network and System Security, Xidian University, Xi'an 710126, China
2. School of Computer Science and Technology, Xidian University, Xi'an 710126, China
3. School of Cyber Engineering, Xidian University, Xi'an 710126, China
4. Systems Engineering Institute, Academy of Military Science, Beijing 100091, China
5. Huawei Technologies Co., Ltd., Shenzhen 518129, China

Abstract: To address the structural mismatch between strict operational constraints of the Industrial Internet and the conventional bolt-on, static-perimeter security paradigm, research on deterministic endogenous security defense was conducted. A deterministic endogenous security defense paradigm was proposed based on four categories of intrinsic determinism: entity, behavior, structure, and environment. Five key enabling technologies were systematically surveyed: endogenous identity management, task-behavior assurance, unknown-attack inference, active security defense, and security effectiveness evaluation. These were used to construct an endogenous security defense architecture covering prerequisites, system models, and four-dimensional technical implementation. The common bottlenecks of existing endogenous

收稿日期：XXXX–XX–XX；修回日期：XXXX–XX–XX

通信作者：沈玉龙，ylshen@mail.xidian.edu.cn

基金项目：国家重点研发计划（No. 2023YFB3107500）；国家自然科学基金（No. 62572371, No.92467201）；山东省重点研发计划（No. 2024CXPT039, No. 2025CXPT089）

Foundation Items：The National Key R&D Program of China (Grant No. 2023YFB3107500), National Natural Science Foundation of China (Grant No. 62572371, Grant No.92467201), Key R&D Program of Shandong Province (Grant No. 2024CXPT039, Grant No. 2025CXPT089)

security practices in cross-path coordination and scenario adaptation were revealed through comparative analysis. Scenario analyses in intelligent manufacturing and industrial 5G illustrate that the proposed architecture can support the translation from theoretical mechanisms into engineering protection by leveraging intrinsic system determinism to counter threat uncertainty, thereby providing a viable path to addressing the cross-domain coordination and dynamic-evolution bottlenecks of industrial Internet security.

Key words: industrial Internet, deterministic endogenous security, task-behavior assurance, unknown-attack inference

0 引言

工业互联网作为新一代信息通信技术与现代工业体系深度融合的关键基础设施[1]，正持续承载制造强国、网络强国、数字中国战略的落地。党的二十届四中全会审议通过的《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》将“建设现代化产业体系、巩固壮大实体经济根基”置于“十五五”战略任务首位；2026年发布的《政府工作报告》进一步明确深化“人工智能+”战略、打造“5G+工业互联网”升级版的部署路径[2-4]。美国、德国、欧盟等主要经济体也围绕先进制造与工业数字化持续推进战略部署[5-7]，工业互联网已成为全球产业竞争与关键技术博弈的焦点。

伴随规模化扩张，面向工业系统的网络攻击正加速演化为跨“数字—物理”两域的现实威胁。2026年伊朗关联APT攻击者针对美国关键基础设施可编程逻辑控制器进行攻击活动[8]，该事件表明，面向工业互联网的网络攻击已可由网络入口深入控制层，直接扰动工业自动化过程；2025年披露的美国家安全局针对国家授时中心的渗透活动表明[9]，关键支撑系统已成高级持续性威胁的主要目标，一旦被破坏，影响将由对时间、控制等基准的共性依赖级联传导至全体系。Dragos统计显示，2025年全球针对工业组织的勒索团伙与受害企业数量同比攀升，制造业受害比例超三分之二[10]。上述事实反映出，工业互联网面临的安全威胁由个案性、局部性的数字资产损失，跃升为关乎生产连续性、关键基础设施稳定与物理过程安全的复合风险。

针对当前威胁演化态势，面向传统信息系统的安全防护范式日显乏力[11]：零散的策略更新与外挂式补救，难以跟上威胁跨域演化与工业场景深度耦合所形成的系统性冲击。新型威胁与传统安全防护范式之间呈现出结构性失配，然而，其失效并非局部策略缺陷，而是范式逻辑与场景特性之间的结构性错配。通过梳理现有防护体系在工业互联网环境下的主要暴露点[12-13]，可将其共性缺陷归纳为安全假设失配、安全目标失配与安全机制失配三类，如表1所示。究其根源，传统范式以“可信内网—静态边界—外挂组件—事后响应”为设计内核，其被动依赖已知特征与事后修补的防御逻辑，与工业互联网强实时、严连续、跨层耦合与持续演化的运行特性根本背离[11]，无法有效支撑工业互联网在复杂动态环境下的安全需求，亟需在防御范式层面寻求结构性突破。

突破上述结构性失配的关键，在于将安全能力从外挂式部署回归至系统自身的运行规律中，即确立“安全能力内生”的理念[14]，安全能力不应作为外挂组件被附加于系统之上，而应从系统本体的可观测、可建模、可约束特征中生长出来。

围绕该理念，以拟态防御等为代表的内生安全范式逐步成形[14-16]，主张利用系统内在规律生成安全能力，实现从“对抗威胁不确定性”向“利用系统内在确定性”的范式迁移。就其切入角度而言，现有工作大致沿四条主线演进：其一，围绕实体确定性构建持续可信基础，以零信任[17]、标识内生安全[18]与可信计算[19]等为代表，构建持续验证与可度量的信任基座，以缓解身份与边界模糊带来的

表1 传统安全防护范式在工业互联网场景中的失配分析

失配类型	核心含义	突出矛盾	导致后果
安全假设失配	前置假设偏离场景现实	可信内网、静态边界、已知特征等假设在IT/OT贯通与威胁演化下失效	边界防御失效，规则更新滞后
安全目标失配	防护目标偏离生产需要	机密性与资产保护目标难以覆盖生产连续性与功能安全	防护优先级错位，物理后果难抑制
安全机制失配	响应机制偏离运行形态	事后响应与补丁修复难以适配长寿命、不可停机的装备形态	响应滞后，补丁难闭环，风险累积

风险；其二，基于运行行为确定性实施过程约束与异常识别，以拟态安全[15,20]、仿生内生安全[21]等为代表，通过冗余异构执行与功能语义建模，抑制对业务逻辑的篡改与越权；其三，从结构确定性出发增强系统韧性与主动防御能力，以移动目标防御[22]、仿生内生安全[21]与 AI 安全箍[23]等为代表，将自适应、反馈与进化机制嵌入系统结构，提升对未知攻击的原生韧性；其四，依托运行环境确定性推动安全能力原生嵌入，以云原生安全[24]为代表，在系统构建阶段即植入安全属性，以替代外挂式安全组件。然而，上述路线对确定性的利用仍具分散性与局部性，未能在内生依据与体系形态上形成统一表达，仍缺乏将工业互联网多维度确定性协同转化为内生安全能力的系统化理论与技术体系。

工业互联网装备类型相对固定、工艺流程规范、通信关系稳定，在实体、行为、结构与环境四个维度上呈现出可识别、可建模的稳定规律，为利用系统内在确定性承载安全能力提供了客观基础。本文

据此将研究主线收敛至工业互联网确定性内生安全，开展系统性综述与体系化构建。主要贡献如下：

（1）构建了分类模型与关键使能技术体系。从时间、威胁与属性三维视角解构现有技术路线，建立涵盖身份管理、行为确保、攻击研判、主动防御与效能评测的闭环使能链条；

（2）揭示了现有方案的演进规律与共性缺陷。对国内外研究与实践方案进行系统归类，提炼传统防护在假设、目标与机制层面的结构性失配根因，指出跨路线协同不足与场景适配受限等关键瓶颈；

（3）提出了确定性内生安全防御范式并构建了分层体系架构。阐明以系统内在确定性对抗威胁不确定性的范式迁移逻辑，依托实体、行为、结构与环境的四维确定性，构建从理论基础到技术实现的分层体系架构，并在智能制造与工业 5G 场景中分析理论机制向工程防护转化的实践路径。

第 1 章梳理工业互联网安全攻防演进脉络，归纳支撑内生安全落地的五类关键使能技术；第 2 至 6 章分别阐述内生身份管理、任务行为确保、未知攻击研判、主动安全防御与安全效能评测的研究进展，其中前四类技术分别依托实体、行为、结构与环境四维确定性生成安全能力，安全效能评测对所生成能力进行验证与度量；第 7 章在五类使能技术基础上构建确定性内生安全体系架构；第 8 章面向典型场景检验架构由理论机制向工程防护的转化路径，并展望开放问题；第 9 章总结全文。

1　工业互联网安全攻防对抗演进

在工业互联网安全攻防演进过程中，传统安全范式长期以边界防护、外挂组件和事后响应为主，在一定时期内发挥了重要作用。但随着工业互联网

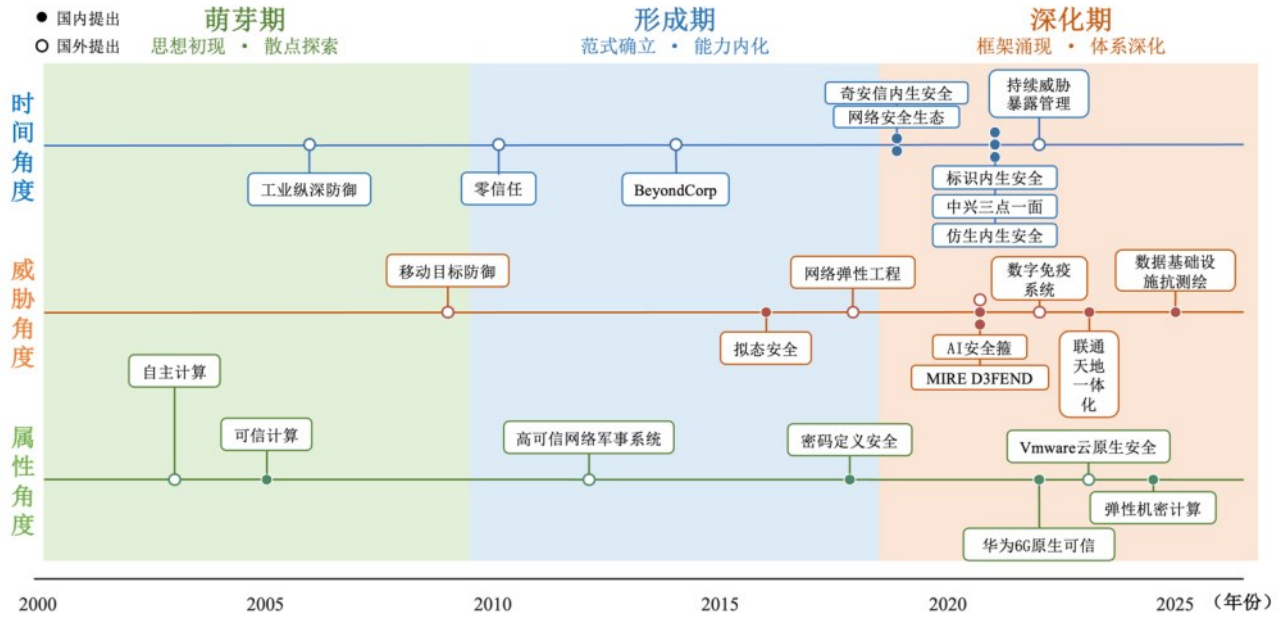

图1　内生安全演进时间线

向高动态、强耦合、复杂交互方向发展，其安全能力外置、响应滞后和防护被动等问题日益突出。在此背景下，“安全能力内生”作为一种新的安全理念被提出，并逐步演化为内生安全技术范式。基于此，从理念演进与实现基础两个层面对内生安全发展脉络进行系统梳理。

1.1 “安全能力内生”理念的探索与分类

“安全能力内生”是在工业互联网复杂场景与传统防护困境的共同驱动下逐步成形的安全理念。其核心在于促使安全机制从外部附加转化为系统的内在运行机制，与业务流程和运行环境协同演化。

从发展脉络看，安全建设经历了由外挂式防护向系统内嵌式防护，再向能力内生式防护的转变。早期防护依赖边界清晰的假设，以防火墙、入侵检测等外部机制为主。随着工业互联网系统复杂度和耦合度不断提高，安全能力逐步由外围下沉至身份认证、行为监测等内部环节。近年来相关研究更加强调在系统设计、构建、运行和演化全过程中组织安全能力，推动安全从“运行后补充”转向“构建时融入”，从外部附加属性转向系统内在特征。

围绕“安全能力内生”理念，现有内生安全研究形成了多条技术路线，如图1所示。考虑到不同路线在切入角度上的差异，本文从时间、威胁和属性三个视角对其进行归纳。表2从研究视角、确定性要素与关键使能技术三个维度对上述路线进行了对照梳理。

时间视角关注安全能力的全过程嵌入与持续演化，阐明安全能力如何实现全生命周期伴生。主张安全建设应贯穿研发、部署、运行与维护各阶段，将一次性核验改造为贯穿全生命周期的连续过程。

威胁视角关注未知威胁、内源风险与复杂攻击下的主动应对，阐明安全能力如何适应未知威胁与复杂攻击。该类路线不依赖完整的攻击特征，借助系统内部的冗余结构、动态变化与协同机制构造对未知扰动的天然抑制。

属性视角关注安全能力的前置植入与稳定支撑，阐明安全能力如何形成系统内生属性。该类路线在系统构建阶段即将可信根、原生特征与基线度量纳入设计目标，使安全成为系统的内生属性。

尽管内生安全范式为突破被动防御困境提供了多元路径，但各类技术路线往往仅侧重于实体、行为、结构或环境中的单一维度特征，未能将内在确定性的多维协同转化为体系化防御合力。

1.2 内生安全的关键使能技术

不同视角技术路线虽各有侧重，但实践表明，依托内在确定性优化安全机制是提升防御效能的有效途径。综合现有技术共性与工业场景约束，可将支撑内生安全范式落地的关键使能技术概括为内生身份管理、任务行为确保、未知攻击研判、主动安全防御与安全效能评测五类。其中，前四类技术分别依托实体、行为、结构与环境四维确定性生成安全能力，安全效能评测则面向所生成的安全能力实施验证与度量。

内生身份管理：面向系统参与主体，其任务是建立可持续验证的身份与信任基础，为访问控制与审计追溯提供支撑。工业互联网设备类型明确、角色关系稳定，为构建持续认证与细粒度授权提供了天然条件。利用实体内在确定性，构建不可伪造且持续可验的确定身份。

任务行为确保：面向系统运行过程，其任务是依据工艺流程与业务语义对行为施加实时约束，使执行处于预期范围。不同于依赖异常特征的被动监测，其强调从工序与任务的内在规则出发识别违规操作。依托行为内在确定性，约束运行过程始终符合预期的确定行为。

未知攻击研判：面向异常识别与归因，其任务是将攻击发现从依赖外部特征匹配转向基于系统内部规律的综合研判。未知攻击虽可规避已知规则，但其引发的状态偏离会在系统层面留下可观测线索。基于结构内在确定性，从偏离规律中因果推演确定威胁。

主动安全防御：面向异常处置与响应，其任务是通过阻断、反制与策略演进抑制攻击影响，保障系统受扰时维持基本运行。工业装备长期在线、停机代价高的情况，要求防御需与系统运行同步生效。融合环境内在确定性，驱动防御机制持续输出确定安全效能。

安全效能评测：面向安全能力的验证与度量，其任务是检验防御能力是否真正源于系统内在确定性，以及检测、研判、阻断与恢复各环节能否在工业运行约束下闭环生效。工业互联网工艺流程规范、运行基线稳定，为构建可复现的评测环境与过程化指标体系提供了客观依据。

这五类技术相互依赖，内生身份管理与任务行

表 2　　内生安全相关技术的研究视角、确定性要素与关键使能技术对照

内生安全技术	研究视角	确定性要素				关键使能技术			
		实体	行为	结构	环境	身份管理	行为确保	攻击研判	主动防御
工业纵深防御□[11]	时间	○	○	★	○	○	○	○	○
零信任□[17]	时间	★	○	○	○	★	○	—	—
BeyondCorp□[25]	时间	★	○	○	○	★	○	—	—
奇安信网管内生安全[26]	时间	○	★	○	○	○	★	★	○
标识内生安全[18]	时间	★	—	★	—	★	—	○	—
持续威胁暴露管理□[27]	时间	○	○	★	○	○	—	★	○
中兴三点一面[28]	时间	★	★	○	○	★	★	★	○
仿生内生安全[21]	时间	—	★	★	○	—	★	★	○
网络安全生态[29]	时间	—	○	★	○	—	—	○	★
移动目标防御□[22]	威胁	—	★	★	○	—	—	○	★
拟态安全[20]	威胁	—	○	★	—	—	★	★	★
网络弹性工程□[30]	威胁	○	★	★	○	○	★	★	★
MITRE D3FEND□[31]	威胁	○	★	★	○	★	★	★	★
AI 安全箍[23]	威胁	○	★	★	—	—	★	★	★
数字免疫系统□[32]	威胁	—	○	★	○	—	○	○	★
联通天地一体化安全[33]	威胁	★	—	★	○	○	—	—	—
数据基础设施抗测绘[34]	威胁	○	○	★	○	—	—	○	★
自主计算□[35]	属性	○	○	★	○	—	○	○	★
可信计算[19]	属性	★	—	★	★	★	—	—	○
高可信网络军事系统□[36]	属性	○	—	★	★	○	★	○	★
VMware 云原生安全□[37]	属性	○	○	★	★	○	★	○	○
密码定义安全[38]	属性	★	—	○	○	★	○	—	○
华为 6G 原生可信[39]	属性	★	○	★	○	★	○	○	★
弹性机密计算[40]	属性	○	—	★	★	★	—	○	○

研究视角分为“时间、威胁、属性”三类，单选标注；确定性来源（实体 / 行为 / 结构 / 环境）与关键使能技术（身份管理 / 行为确保 / 攻击研判 / 主动防御）采用多选标注，★ = 主要归属，○ = 次要涉及，— = 未明确涉及；“□”表示该技术为国外提出或源自国外研究。

为确保依托确定性要素保障主体可信与过程合规，未知攻击研判与主动安全防御完成威胁的识别与处置，安全效能评测对上述能力的内生性与闭环性进行验证与反馈。

2　内生身份管理

身份是工业互联网系统中访问与交互的逻辑起点，是安全能力由外部防护转向系统内部生成的重要支撑。身份管理作为访问控制的前置环节，是工业互联网内生安全体系中连接主体可信和责任追溯的基础机制。

围绕身份管理，现有内生安全研究已从不同确定性来源出发形成了具有代表性的实践脉络。以零信任[17]与中兴三点一面[28]为代表，将身份置于防护体系核心，通过多因子动态评估与跨层联动认证，将信任延展至访问全周期；以标识内生安全[18]与可信计算[19]为代表，从网络通信层与硬件底层切入，借助标识绑定或可信根度量，将身份锚定于系统结构特征；以华为 6G 原生可信[39]与 VMware 云原生安全[37]为代表，将身份能力嵌入基础设施控制点，实现身份与架构一体化交付。代表性实践的对比情况见表 3。

2.1　内生身份构建

身份构建是身份管理生命周期的起点，身份质

表3　代表性内生安全实践的身份管理特征对比

代表性实践	物性绑定	动态授权	免硬件改造	场景适配
零信任	—	●	●	时延敏感
中兴三点一面	○	●	●	电信级系统
标识内生安全	○	○	●	多标识网络
可信计算	●	—	—	需硬件改造
华为6G原生可信	○	●	●	6G基础设施
VMware云原生安全	—	●	●	云原生负载

●表示充分具备/明确支持，○表示部分具备/间接支持，—表示未明确涉及。

量直接影响认证、授权与审计的可靠度。工业互联网装备规模大、寿命周期长且异构程度高，传统身份构造方式易复制且难以迁移。以实体物性特征、功能属性和行为模式等内在确定性为依据，构建难伪造且可持续核验的内生身份，已成为该方向的研究主线。

硬件物性特征是当前常用的身份依据。物理不可克隆函数利用芯片制造中的不可控工艺偏差生成不可复制的激励响应映射[41]，可构造工业设备硬件级身份指纹。相关研究近年在阵列结构、抗侧信道攻击和低功耗实现方面推进[41]，并与消息认证码、密钥协商协议结合，形成面向内生身份的轻量化方案。此类方案对硬件工艺与运行工况敏感，在缺少专用安全芯片的存量装备上迁移代价较高。而射频指纹[42]与信道指纹[43]将身份依据扩展至电磁空间，通过射频前端器件差异带来的辐射特征或多径变化中的链路特征，借助统计建模或深度学习实现设备级辨识[44]。

另一类研究将身份依据由静态物性扩展至动态语义，以设备在工序中的调用序列、通信对象集合和指令时序分布等任务级稳定特征补充身份表征[48]。该类方法不依赖专用硬件，部署门槛较低，但对工艺切换和任务更替敏感，需与物性特征融合以保持稳定。

内生身份构建主要沿三条方向演进：表征维度由单一物性扩展为物性、功能与行为的多维融合；建模方法由特征工程转向深度学习与表示学习；部署形态由集中注册转向轻量化和端侧化。但在面向工业场景的多源特征加权策略、长周期运行中的身份自适应更新机制、跨厂商设备身份互认协议仍缺乏统一可行的技术框架。

2.2　隐式认证

隐式认证将认证过程以业务无感方式嵌入系统运行，避免频繁显式交互引入额外时延或操作中断。其关键是将认证依据由用户主动出示的凭证，转向系统运行中可持续验证的内在信号。

面向资源受限装备的轻量级认证是该方向的基础路径。基于对称密码的挑战应答机制、基于椭圆曲线密码的相互认证协议[45]、基于无证书公钥体制的密钥协商方案，均以降低计算和通信代价为目标。近年来，属性基加密[46]与部分同态加密[47]被引入工业认证场景，用于合并认证与细粒度授权，或在不暴露原始特征的条件下完成隐式比对。该类研究可在同等安全强度下降低开销，但本质上仍是显式凭证体系的轻量化改造，未摆脱对预置密钥管理的依赖。

行为特征持续认证与物理层信号认证代表认证依据的迁移。行为类研究[48]通过建模设备调用时序和交互模式，在会话运行过程中完成隐式持续验证。物理层认证[49]利用无线信道快变特性与射频器件辐射差异建立链路级身份验证机制，认证过程与通信过程耦合，攻击者难以在不改变物理收发链路的条件下伪造身份。基于信道状态信息的链路指纹[43]、基于到达角的空间指纹[49]、基于信号轮廓的辐射指纹[42]也已形相对完整的方案族。面向风险的自适应认证[50]进一步以实时风险评估驱动认证强度和因子组合，在低风险阶段依赖背景信号持续核验，在风险上升或情境切换时引入额外因子，使认证强度与设备状态变化协同。

在工业互联网约束下，相关工作仍存在瓶颈：行为模型跨工段迁移时易出现特征漂移；物理层方法在强电磁干扰环境下鲁棒性不足；风险驱动的动态调节缺乏统一度量基准；认证强度与控制回路时延裕度之间尚未形成可工程化的取舍框架。

2.3　权限无感控制

权限控制是身份管理结果在系统层面的落实环节。传统的基于角色的静态授权和基于固定属性策略的方案，难以刻画身份、任务与环境之间的动态联动关系，因此催生形成以无感性和连续性为特征的权限无感控制机制。

以身份状态驱动动态权限控制是当前的基本路径。属性基访问控制[63]将主体、客体和环境属性

组合为可在运行时求值的授权策略，并把设备指纹、认证强度和信任评估等动态属性注入决策流程，使权限与身份实时可信度匹配[51]。基于智能合约的分布式访问控制[52]将授权逻辑沉淀为链上可验证规则，支持跨域主体在缺少中心化授权服务时实现细粒度管控，但链上计算代价和策略更新时延仍限制规模化部署。以信任评估驱动的持续授权机制[53]进一步将权限控制由“一次授予、长期有效”改造为与任务生命周期同步的连续过程，在权限生效期间持续评估主体行为和环境条件，一旦偏离即触发权限降级、紧急隔离或熔断撤销，或将零信任[17]持续验证结果转化为运行时执行动作。

面向跨层协同的权限联动机制将权限控制扩展到多域多系统场景[54]。工业无线网络需实现从物理层认证到应用层授权的跨层策略协同，可借助统一身份实现多层策略联动更新[54]。在云边端协同场景中，基于策略引擎的跨域联动和基于区块链的策略共识机制，可支撑组织边界间的权限一致性[52]。

现有研究主要存在三类不足：工业互联网大规模实体下，精细粒度权限易引发策略爆炸；异构系统间的信任度量缺乏统一语义；跨域协同机制对策略一致性和更新时延的保障不足，在工艺节拍紧约束场景下更为突出。如何在保障生产连续性的前提下实现策略语义压缩、跨域信任统一映射和权限联动节拍同步，仍是该方向需回应的问题。

3　任务行为确保

任务行为刻画工业互联网实体在生产任务中的执行轨迹，以判断异构实体是否按既定任务语义运行。该环节承接身份可信结果，支撑过程合规控制，并为后续安全研判提供行为证据。

现有内生安全实践已对任务行为确保从多方向展开探索。零信任[17]强调对访问行为和上下文状态进行持续裁决；VMware 云原生安全[37]侧重在工作负载和网络控制点中嵌入运行时防护；中兴三点一面[28]通过流量、日志、资产和业务行为融合支撑协同监测；奇安信网管内生安全[26]面向边缘接入场景强化状态监测和违规阻断；AI 安全箍[23]则通过行为边界、风险触发和安全接管约束智能体运行。表 4 对相关实践进行归纳。

3.1　行为建模方法

行为建模是将设备操作与控制过程抽象为可计算且可验证的行为表示，为运行监测与攻击研判等环节提供统一依据。工业任务行为受工艺顺序和控制逻辑的强约束，建模对象不仅要刻画单点动作，还需把握任务流程及其运行上下文[57]。相关研究可沿两个维度展开：按目标分为正常行为建模与异常行为建模，按数据组织方式分为单源建模与多源关联建模。

正常任务行为建模从稳定工况数据中抽取设备和任务的知识数据，用于形成任务执行的预期边界[57]。有限状态机以状态集合及转换关系表达设备启动、运行、暂停、故障和恢复等过程，可用于识别偏离预期状态迁移的异常操作[55]。Petri 网及其扩展模型以库所、变迁和令牌流动描述任务步骤、触发条件和资源状态，可利用工业控制系统输入/输出信号刻画工业过程运行逻辑[56]。状态感知异常检测进一步将过程状态和控制上下文纳入行为边界表达，通过状态相关阈值约束不同工况下的行为变化[57]。与上述正常边界建模不同，攻击或异常行为建模以已知违规操作、故障样本或攻击样本为对象，通过监督式模型学习不合规行为的特征边界[58]。

表 4　代表性内生安全实践的任务行为确保特征对比

代表性实践	语义任务建模	行为表征层	约束管控方式	监测与联动对象
零信任	○	访问层级	持续访问约束	策略引擎动态裁决
VMware 云原生安全	○	负载层级	工作负载防护	控制点内嵌
中兴三点一面	○	跨越多层	协同监测与编排	云网安算联动
奇安信网管内生安全	—	管控层级	状态监测与违规阻断	管控过程内嵌
AI 安全箍	●	智能体	触发限制与安全接管	智能体与决策过程

●表示充分具备/明确支持，○表示部分具备/间接支持，—表示未明确涉及。

单源建模通常基于主机日志[78]、网络流量[1]、控制指令[66]或传感器数据[67]中的某一类数据源构建行为表示。例如，基于控制网络数据流或物理传感信号的预测模型，可用于识别运行过程中的异常行为[59]。多源融合建模则将主机行为、网络交互、控制过程和物理状态进行联合表示，常借助图结构[60]或跨系统表示学习方法[61]刻画实体关系与跨层依赖。中心性感知图卷积网络将工业控制系统组件关系建模为图结构，并利用中心性信息增强图卷积网络对组件关联关系的利用能力[60]；多层图注意力方法通过构建物理过程图和控制器图，学习过程变量与控制器行为之间的关联关系[61]。该类研究表明，多维关系建模有助于增强对工业任务状态的表达能力。

现有行为建模能够对任务执行状态进行有效表征，但仍有局限：时间约束刻画不足，难以反映任务节拍；对任务变更和上下文变化适应有限，易依赖稳定工况下的规则或样本分布；多层多域的耦合表达不足，影响跨层解释和约束执行。仍需构建支持动态切换且具可解释性的任务行为模型。

3.2 行为持续监测

行为持续监测是任务行为确保在运行时的观测环节，以及时识别偏离预期任务过程的行为变化。与周期性审计不同，它强调对运行状态的连续采集与实时分析，并借助上下文关联，为行为约束与安全研判以及后续处置提供过程证据。

现有行为观测主要依托主机日志[78]、网络流量[1]和控制过程[59]数据。主机侧通常关注进程、文件、系统调用和服务状态[78]，网络侧关注通信对象、协议字段、连接频率和流量变化，控制侧关注指令序列、传感器反馈和执行器响应。围绕上述数据，行为模式匹配、异常聚类和时间片特征提取等方法常用于在线行为变化识别[59]。状态感知异常检测将物理过程状态与控制上下文纳入判断过程，可通过状态相关阈值识别偏离预期过程行为的异常[57]。

由于主机日志、网络流量和控制过程数据分别反映系统运行的不同侧面，单一数据源通常只能捕获局部行为变化，难以揭示多维度之间的关联关系。因此，多源数据关联监测成为行为持续监测的重要方向[59]。该类方法将离散日志、流量特征、控制变量和状态事件组织为统一关系表达，以刻画任务行为的连续性和跨层依赖。图结构关联分析是其中较为典型的实现方式，可将系统实体、操作事件和交互关系映射为图结构，描述行为之间的相互关系[78]。中心性感知图卷积网络和多层图注意力网络等方法表明，图结构方法能够为工业控制系统中的组件关联建模、过程变量关系学习和行为变化解释提供支撑[60-61]。

现有行为持续监测已由单点日志审计扩展为多源在线分析，但多源数据的语义对齐仍存在难度，行为变化与任务后果之间的关联表达也不充分。监测结果如何联动约束处置和任务恢复，仍需结合运行上下文和图结构推理机制进一步完善。

3.3 行为安全约束

行为安全约束面向任务执行过程中的运行边界控制。保障实体操作处于正确任务阶段和有效交互路径之中。当任务遭受扰动时，安全约束需支撑系统实施隔离降级以抑制风险。与一般访问授权不同，该环节强调将行为建模和持续监测结果转化为过程合规控制能力。

基于角色的访问控制通过角色权限映射限定主体可执行操作，基于属性的访问控制将主客体特性和运行环境纳入授权决策，动态策略引擎进一步根据运行上下文调整访问边界[62-63]。零信任架构通过策略引擎和执行点实现持续访问控制，使访问授权由一次性裁决转向运行过程中的持续验证[17]。在工业任务场景中，这类机制可与任务阶段、工艺状态、通信对象和链路条件结合，用于约束访问授权、控制指令、交互路径和任务操作的合法性。

围绕任务受扰后的恢复，现有方法从风险抑制和功能维持两方面展开。IEC 62443 的区域与通道模型通过安全区域划分和通信通道控制限制系统间交互范围，为工业控制系统中的隔离控制和故障扩散抑制提供工程依据[64]。软件定义网络（software defined networking, SDN）可通过流表下发、流量重定向和动态隔离实现异常通信链路的快速调整，为任务链路恢复和网络层重构提供支撑[65]。在工业无线网络中，链路质量、接入状态、通信时延和信道扰动也可作为恢复决策依据，用于触发备用策略，从而降低异常行为对任务连续性的影响。

行为安全约束已从基础访问授权扩展到面向任务阶段的过程控制。但现有方法在规则模板依赖、跨系统策略一致性、任务语义表达和实时恢复代价

方面仍存在不足。面对动态环境波动与跨域协同需求，如何平衡操作合法性、约束实时性和任务连续性，仍是后续研究重点。

4　未知攻击研判

未知攻击研判旨在解决攻击先验信息匮乏带来的研判问题。为克服上述挑战，研判工作需甄别异常的攻击属性，推演威胁扩散路径及影响范围。有别于依赖已知特征匹配的传统入侵检测，该环节核心在于利用系统内部可观测的确定性知识，驱动从异常发现经风险推演至示踪溯源的连贯闭环。

现有实践从多路径支撑未知攻击研判。仿生内生安全[21]侧重由正常状态、异常偏离和环境反馈识别未知异常；拟态安全[20]依托动态异构冗余与输出裁决识别未知扰动；AI 安全箍[23]面向智能体行为设置风险触发和安全接管边界；中兴三点一面[28]通过流量、日志、资产和行为数据融合支撑跨层态势感知。表 5 对上述实践进行了对比。

表 5　代表性内生安全实践的未知攻击研判特征对比

代表性实践	研判依据	免攻击先验	反馈自调节
仿生内生安全	状态偏离	●	●
拟态安全	输出裁决	●	●
AI 安全箍	行为边界	●	○
中兴三点一面	多源融合	○	○

●表示充分具备/明确支持，○表示部分具备/间接支持，—表示未明确涉及。

4.1　攻击实时检测

攻击实时检测旨在运行过程中即时发现异常状态。工业互联网检测对象横跨物理设备与上层任务，需协同提取网络通信特征与物理传感反馈等多维数据，以研判攻击对生产过程的实质性影响。

现有实时检测方法可归纳为规则驱动和数据驱动两类。规则驱动方法依托专家知识、协议规范、任务流程进行检测[66]。相关研究通过挖掘控制通信行为提取规格模型[66]，或从过程数据中学习物理不变量[67]，以识别偏离约束的异常交互。

数据驱动方法借助机器学习与时序预测等技术从多源运行数据中学习行为模式并识别偏离[59]。例如，状态感知检测通过引入过程上下文构建动态阈值，相关方法利用超过 120 GB 的历史运行数据建立状态模型，相较对照的状态型检测方法，在水位残差阈值为 30 mm 时将误报告警数减少 99.21%，同时将攻击检测时间缩短 14.50%[57]；过程感知分析可捕捉隐蔽攻击引发的时序结构偏移，面向长期潜伏攻击的在线分析，任务引导的进程溯源图分割方法在 DARPA TC 的 Trace 子集上取得了 99.99% 的准确率[68]；融合控制流与传感数据的建模方法则能有效识别跨层异常[59]。

规则驱动方法具备可解释性优势，数据驱动方法擅长挖掘复杂模式。然而，当前检测仍受限于规则维护代价高、跨场景泛化弱、标注样本缺失和任务语义表达不足等问题。后续研究需深化领域规则与驱动模型的融合，以提升未知攻击检测的适应性与可解释性。

4.2　风险推演预测

风险推演预测面向前向分析，推断攻击的后续步骤及影响范围。与实时检测关注已发异常不同，该环节侧重研判威胁跨越信息层与物理层的潜在扩散路径，从而为主动防御的资源调度与优先级决策提供支撑。

针对工业环境下的前向预测需求，现有研究从多维度展开探索：在逻辑层面剖析攻击路径的横向扩散，在状态层面刻画攻防演化的阶段特征，在数据层面挖掘风险传播的关联规律，在物理层面估测过程受扰的偏差程度。由此，形成了攻击路径推演、概率状态预测、数据驱动演化预测及物理状态估计四类核心路线。

攻击路径推演主要利用资产、漏洞与权限状态刻画从入口到目标的潜在路径，适于分析横向移动过程[69]。相关研究将概率攻击图引入企业网络风险分析[69]，并拓展至运营技术系统的风险传播建模[70]；贝叶斯攻击图则进一步引入条件概率量化事件依赖，可以支撑电力信息物理系统的攻击路径推断[71]。

概率状态预测侧重刻画攻击阶段演化，通过隐马尔可夫模型与贝叶斯网络等将攻击表示为状态转移序列，依据观测事件推断当前阶段并预测后续行为[73]。贝叶斯网络因擅长表达条件依赖，被用于动态更新风险判断[72]；结合连续时间贝叶斯网络，可实现攻击风险的时间演化评估[73]。

数据驱动方法利用时序模型与图学习从告警、流量与过程变量中学习演化规律[75]。机器学习模型可基于历史模式预测后续攻击[74]；图神经网络

则通过建模设备关系与运行依赖，支撑异常传播与潜在风险识别[75]。

针对强物理约束场景，状态估计与残差分析通过比对预测与观测值判断传感器或执行器攻击。基于预测模型与控制不变量的方法可有效识别物理过程异常[76]；基于树结构的状态估计与安全控制也成为应对网络攻击的关键技术[77]。

现有风险推演研究在路径表达、演化建模与扰动分析上已取得进展，但在动态接入、跨域融合和实时适配方面仍存在不足。后续需将任务行为确定性作为约束，结合图推理与在线状态估计，构建可解释的推演方法。

4.3　攻击示踪溯源

攻击示踪溯源面向后向分析，旨在逆向还原攻击的初始入口。区别于实时检测与前向推演预测，该环节聚焦挖掘行为间的因果关联与路径完整性，为责任认定与防御策略优化提供依据。

现有攻击示踪溯源通常沿“因果建模—路径重建—可信增强”链条展开。因果关系建模为溯源分析提供基础结构，将系统实体（例如，进程、文件、套接字、服务和设备）表示为节点，各类操作表示为边，构建溯源图[78]。溯源图以有向时序图表达主客体间控制与数据流，连接跨时间因果事件，支撑威胁检测与来源回溯[78-79]。

攻击路径重建旨在从大规模因果图中提取入口与关键节点。全量回溯产生的依赖图规模庞大，需进行路径筛选。相关研究通过反向传播依赖影响并结合数据流与时间属性，筛选高相关路径[80]；或从恶意节点出发聚合邻接边分数以重构攻击摘要，提升归因解释力[81]。

可信增强旨在提升溯源稳定性与语义完整性。上下文感知的图规约可削减粗粒度日志引入的虚假依赖[82]；基于知识图谱的方法通过抽取攻击技术关系或统一内外部知识，支撑攻击链描述与场景分析[83-84]。路径评分与图规约可有效缓解日志缺失与语义不完整问题。

现有溯源已形成较完整的分析框架，但在工业场景仍受日志缺失、时序误差与跨域语义不一致制约，无线环境中的链路波动与动态接入进一步增加还原难度。后续需结合任务语义与多源知识对齐，构建可解释的溯源机制。

5　主动安全防御

主动安全防御是内生安全体系的处置与演进环节，旨在通过阻断、隔离、重构与诱捕等机制抑制风险扩散，维持关键任务连续运行。

从发展脉络看，主动防御实践已由仿生免疫式识别与攻击面动态变化，向可信状态度量和云网协同运营扩展。移动目标防御[22]通过动态改变配置提升攻击不确定性；拟态安全[15,20]依托冗余裁决压缩扰动空间；人工智能安全箍[23]面向工业智能组件提供运行时约束；可信计算[19]通过完整性度量支撑核验；中兴三点一面[28]则强调云网安算协同处置。表 6 对上述实践进行归纳对比。

5.1　攻击快速阻断

攻击快速阻断是主动防御的前端处置环节，旨在攻击迹象显露或风险确认后，快速限制攻击源与异常链路活动，抑制威胁扩散。

现有阻断研究主要围绕流量阻断、分区隔离与动态重构三类方法展开。流量阻断侧重对异常连接或报文进行快速限制，主要通过访问冻结、策略下发与流量过滤等实现[65]。SDN 凭借集中控制与可编程特性，被广泛用于工控系统入侵响应，可通过流量重定向与策略重配置实现快速阻断[65]；基于 SDN 的自动化响应方案可由主机或网络告警触发，对特定组件实施增强监测或隔离[85]。

分区隔离侧重限制横向移动与故障扩散。IEC

表 6　代表性内生安全实践的主动安全防御特征对比

代表性实践	核心处置机制	业务无感处置	策略自演进	未知攻击适应
移动目标防御	攻击面变换	○	●	●
拟态安全	执行体隔离	●	●	●
AI 安全箍	行为接管	○	○	○
可信计算	节点隔离	○	○	—
中兴三点一面	策略阻断	○	●	○

●表示充分具备/明确支持，○表示部分具备/间接支持，—表示未明确涉及。

62443 标准的区域与通道模型将系统划分为安全区域并控制域间通信，为网络分段与扩散抑制提供工程依据[64]。在此基础上，节点隔离、微隔离与安全域切换可将异常设备从生产网络剥离，降低对关键链路的影响[37]。该方法适用于边界清晰的场景，但在多任务协同下需处理策略一致性问题。

动态重构强调阻断攻击时维持任务执行。移动目标防御通过地址、端口与路径的动态变换及服务迁移改变攻击条件，提升系统抗中间人攻击韧性[86]；低时延网络属性随机化可主动缓解侦察攻击[87]。工业无线网络中则可通过信道切换与备用链路启用降低伪装接入或恶意流量的干扰。

现有阻断研究仍存在场景局限，且多以流量过滤与区域隔离为主，难以应对隐蔽性强的新型攻击，阻断手段的主动性与多样性仍需提升。

5.2　反制协同打击

反制协同打击是快速阻断后的主动压制环节，在合法授权与安全边界内，通过欺骗诱捕与资源牵制削弱攻击者侦察与横向移动能力。

欺骗诱捕是典型路径，主要通过部署蜜罐、伪装服务与仿真环境引导攻击者进入可控交互空间。其中，网络欺骗机制通过诱导攻击者暴露战术意图提升主动感知能力[88]。面向信息物理系统与工业控制场景，蜜罐与蜜网可吸引攻击者访问虚假系统辅助防御[89]，高交互诱捕环境则能提供更丰富的行为观测能力[90]。

资源牵制侧重延长攻击者侦察与验证过程。Tarpit 技术利用虚假主机表象与 TCP 流控维持连接并限制传输，以拖慢扫描攻击[91]；在物联网场景中，该机制也可占用受感染设备交互过程以降低恶意活动效率[92]。此外，欺骗资源迫使攻击者投入额外时间区分真实与诱饵目标，在攻击链早期引入摩擦并提高成本[88]。工业互联网场景中，该机制表现为会话拖延与虚假工艺数据牵制，但需与真实生产链路隔离以避免反制影响控制业务。

工程实现上，反制协同打击需与阻断、溯源及策略演进联动。攻击进入诱捕环境后，系统结合日志回溯与路径识别触发访问控制与流量调度；诱捕所得样本与行为序列可反馈至检测模型优化防御。因此，反制协同打击是以欺骗诱捕与资源牵制为核心、多控制点联动为支撑的主动防御过程。

尽管反制协同打击已向诱捕、牵制与联动结合的机制演进，但欺骗环境可信度、反制授权边界与任务连续性约束仍是工业场景面临的突出挑战。后续需结合溯源结果与任务拓扑构建可控低干扰的反制机制，使反制动作服务于资源消耗与风险隔离。

5.3　智能演进免疫

智能演进免疫是主动防御的长期提升环节，以将攻击事件与运行反馈转化为可持续更新的防御知识与恢复能力。

围绕防御能力的持续更新，现有研究主要形成人工免疫学习、自愈韧性恢复、可信基线更新与策略自适应优化等路线。

人工免疫系统借鉴生物免疫的自体/非己识别与克隆选择机制，构建可识别异常并持续更新规则的防御模型。通过抗原/抗体编码与演化模式设计，实现异常行为检测与自适应模式识别[93-94]，进而支撑工业场景未知威胁下的规则演化与模型更新。

自愈与韧性恢复是另一重要方向。攻击或故障后，关键任务恢复需通过状态诊断与功能重构维持系统可用性，而非仅依赖人工修复。冗余配置与恢复控制是提升信息物理系统抗扰能力的关键[95]，自感知与自诊断自愈也是工业物联网适应复杂环境的发展方向[96]。

可信基线更新为演进提供校验基础。可信计算通过完整性度量与远程证明验证设备状态，防止被篡改节点重新参与任务。可信计算在工控系统中可用于增强设备与控制系统的可信运行基础[97-98]。该机制为不可信节点排除与恢复后再接入提供依据。

策略自适应优化推动防御走向闭环更新。移动目标防御正从随机变化转向自适应策略，以平衡安全收益与切换代价[99]。防御系统可依据攻击行为实时调整诱捕配置以提升欺骗效果[100]。工业互联网场景中，策略优化需结合研判结果与处置反馈，避免将噪声或误报固化为防御知识。

人工免疫学习转化异常记忆为更新知识，自愈韧性恢复维持任务连续性，可信基线与策略优化则支撑状态维护与闭环演进。然而现有演进机制仍面临在线学习稳定性不足与策略更新安全性难保障等问题。后续需融合研判结果与处置反馈构建可审计回滚的演进机制。

6　安全效能评测

安全效能评测是内生安全体系中连接技术能力

与工程验证的关键环节。评测结果既用于验证前述技术的功能与性能，也为后文体系架构中的能力边界刻画和闭环优化提供依据。区别于一般信息系统安全测试，内生安全效能评测不仅关注系统能否抵御攻击，还关注防御能力是否真正源于系统内部，以及从威胁发现到系统恢复的全过程能否形成闭环。

6.1 效能指标构建

现有评测标准与指标主要聚焦三类问题：评测活动的规范化、安全能力的量化与攻击下任务承载的度量。

在规范化层面，ISO/IEC 15408[101]、ISO/IEC/IEEE 29119[102]、IEEE 1012[103]等通用标准从评估流程、保护轮廓与验证活动等角度建立了产品级安全评估骨架，NIST SP 800-53[104]与 SP 800-115[105]进一步细化了控制基线与技术性测试要求；面向工业场景，IEC 62443[64]以区域划分与安全等级为核心给出行业基准，国内 GB/T 28448[106]与 GB/T 36466[107]等在等级保护与工控风险评估方向上构成对应规范。此类标准为评测活动提供了通用骨架，但其面向静态安全功能声明设计，对能力是否内生等机制性属性缺乏直接刻画。

在量化层面，相关研究普遍以检测率、误报率、漏报率、F1 与 AUC 为基础指标度量识别能力，并结合检测时延、阻断时延、吞吐量与资源占用等维度考察安全机制对运行代价的影响；面向工业场景，相关工作进一步引入提前预警时间、控制变量恢复时间与过程偏离量等过程化指标，使评测从识别精度向过程承载延伸。然而，上述指标仍主要围绕识别和响应二阶段，对闭环防御与内生性本身缺乏专门刻画。

在任务承载层面，MITRE ATT&CK[108]与 D3FEND[31]通过攻击战术与对抗性防御的结构化语义为攻击覆盖度分析提供了通用语言，NIST SP 800-160[30]与 MITRE 网络弹性度量框架[109]进一步引入任务可用性曲线与关键功能保持度等弹性指标。围绕内生安全的特殊关注点，相关研究对传统指标进行了扩充：移动目标防御[22]研究关注攻击面缩减率与变换熵，拟态防御[14-16]研究使用共模成功率与一致性输出概率，零信任[17]研究引入信任评分稳定性与权限漂移量。这一方向呼应了内生安全所强调的能力闭环，但跨技术路线的指标尚未形成统一标准。

6.2 评测方法设计

测试环境与评测方法主要聚焦三类问题：评测结果的可复现性、复杂攻击的可执行性与内生机制的可验证性。

在可复现层面，工业控制公开数据集与测试床构成最主要的实验依托。SWaT[110]等覆盖水处理、水分配、化工过程与电力调度等典型场景，在不同程度上引入多变量、多阶段攻击；MiniCPS 等仿真工具[111]与相关综述[112]进一步对各类测试床的覆盖范围与差异进行了系统比较。这类资源使攻击检测与异常识别研究具备了一定的横向可比性，但其物理过程类型与攻击注入方式相对集中，跨场景迁移仍然受限。

在复杂攻击层面，攻击图、攻击库与靶场构成主要技术手段。攻击图方法将系统状态与攻击动作组织为有向图，用于路径枚举与关键节点识别[69]；基于 MITRE ATT&CK[108]的自动化攻击仿真工具支持原子攻击的组合执行；网络靶场与红蓝对抗以攻防博弈方式检验闭环防御能力[113]。三者在覆盖度、可组合性与对抗真实性上形成互补，是评测从已知攻击复现走向未知与组合攻击验证的主要支撑，但其在工控协议、实时控制与物理后果维度上的扩展仍不充分。

针对未知攻击，评测需要明确“未知”的参照对象。若攻击样本、规则或特征已进入被测系统的知识库，则判定属于已知攻击；若攻击目标或类型可以描述，但具体载荷、参数和步骤从未向被测系统公开，可视为“已知的未知”；若攻击方式及其组合路径均超出被测系统先验（例如，零日攻击），则可视为“未知的未知”。这一划分描述的是测试信息对被测系统的可见性，“未知”是相对于被测系统而言，测试组织仍应掌握攻击过程并保留复现条件，以保证评测结果可核验。

在方法上，可在现有数据集、测试床与网络靶场体系内增加样本盲测约束。被测系统只接收正常运行数据、业务规则和必要接口说明，不向其提供攻击样本、特征标签或处置脚本。对“已知的未知”，可通过样本留出、参数变异和原子攻击组合形成未见用例；对“未知的未知”，采用红蓝对抗，由攻击方动态构造多阶段攻击链。评测重点在于其能否在缺少攻击先验时发现确定性偏离、限制影响

扩散并恢复关键任务。然而，该类盲测流程、未知度分级和对抗强度控制仍有待进一步规范，但可作为内生安全效能评测的优先发展方向。

在内生机制层面，主流路线已发展出与机制耦合的专用评测方法：拟态防御[14-16]围绕异构冗余执行体使用共模测试、白盒插桩与故障注入验证裁决机制，移动目标防御[22]借助攻击者建模与博弈分析评估变换策略对攻击成功率的抑制作用，可信计算[19]通过逐层度量与远程证明验证度量链可信性并已在电力监控等场景获得实地验证，零信任[17]与持续认证以行为数据集与会话级注入实验评估身份漂移与重认证策略。

6.3　评测实践分析

现有评测实践在内生安全场景下的不足可归结为三方面，并由此牵引出相应的发展方向。

首先是指标体系侧重检测性能，对内生性与闭环性的刻画不足。大量工作仍以识别和响应二阶段指标为核心，对安全能力是否真正源自系统内部确定性以及检测、研判、阻断、恢复是否构成完整闭环，缺乏专门度量。其次是测试场景与攻击样本覆盖偏窄。公开数据集集中于若干典型工业过程，对智能制造、能源互联网、工业无线网络与无人集群等新兴场景覆盖有限；多数样本以单点异常或单阶段攻击为主，对复杂组合攻击的链式过程仍难以充分覆盖。最后是真实性与工程性之间存在权衡。硬件在环测试还原度高但成本与风险较高，数字孪生与仿真覆盖灵活但与真实系统存在偏差，真实产线测试又受停机代价制约；跨技术路线的横向对比与端到端工程指标评估也较稀缺。

此外，从范式定位看，NIST 网络安全框架以风险治理为核心，通过结果导向的功能体系帮助组织识别和管理网络安全风险，但不规定具体技术路径。NIST SP 800-82 强调在性能、可靠性与安全约束下保护运营技术系统[11]，IEC 62443 则通过区域与通道模型和全生命周期要求形成工业控制安全的工程基线[64]。零信任不因网络位置或资产归属授予隐式信任，而是依靠持续验证与最小权限保护资源[17]。MITRE ATT&CK 与 D3FEND 分别组织攻击行为知识和防御技术知识，为威胁建模及防御覆盖分析提供通用语义[31,108]。

与上述范式相比，确定性内生安全的设计起点是工业系统内部可验证的实体、行为、结构与环境规律，核心原则是将这些规律转化为伴随系统运行的约束和恢复机制，以系统确定性压缩威胁不确定性的作用空间。两者并非替代关系。国际主流框架可为确定性内生安全提供风险治理边界、工程控制基线和攻防知识语义，确定性内生安全则为相关要求在工业任务中的持续执行提供内生机制，形成“外部规范牵引、内部机理支撑”的互补关系。

针对上述不足，相关研究已开始向场景化测试床建设、可组合攻击复现框架、数字孪生与硬件在环协同的多保真度评测，以及面向网络弹性与任务承载的工程化指标体系等方向延伸。沿此趋势，内生安全效能评测有望由围绕单一机制的功能验证，逐步走向覆盖未知攻击适应性、闭环防御完整度与业务连续性影响的多维综合评估。

7　工业互联网确定性内生安全架构

前文围绕安全内生理念的演进脉络，系统梳理了内生身份管理、任务行为确保、未知攻击研判、主动安全防御与安全效能评测五类关键使能技术的研究进展。已有研究在不同维度上呈现出共性主线，但仍以分散的技术形态存在，缺乏将其组织为统一框架的体系化表达。本章在此基础上从安全能力内生与工业互联网安全需求的适配关系出发，阐述确定性内生安全的成立前提、体系模型和技术架构。

7.1　安全能力内生与工业互联网安全需求的适配

工业互联网在生产连续性、控制实时性、功能正确性以及与物理过程的强耦合性方面具有区别于一般信息系统的固有约束，安全机制一旦割裂业务过程或在控制回路上引入不可接受的时延，即便具备较强的攻击识别能力也难以为生产过程所容纳。传统安全防护范式与工业互联网运行特性存在结构性失配，需要与系统自身运行规律深度融合、伴随业务过程协同生成的安全能力。

然而，现有内生安全技术虽已在身份、行为、结构、环境等方面提供了有益探索，其体系化程度仍显不足。零信任[17]、可信计算[19]、拟态防御[14-16]、移动目标防御[22]、原生可信[39]等代表性路线分别从不同确定性来源切入，对系统内在规律的具体内涵尚缺乏一致表达，使得不同方法之间的能力边界与可组合性难以判断；此外，多数路线在身份、行为、攻击研判与主动防御之间未形成显式

的闭环关系，能力之间彼此孤立，难以共同支撑识别、约束、研判、阻断与反制的连贯链路。

上述差距指向一个底层问题：要使安全能力内生获得体系化支撑，需要为其确定统一的成立前提与组织主线。该前提需要具体落到系统自身可被刻画与利用的稳定规律之上。本文据此将研究对象进一步收敛为确定性内生安全，即以工业互联网系统内在的实体确定性、行为确定性、结构确定性与环境确定性为统一基础，将安全能力从外挂式、经验式与规则式防护，转向基于系统自身规律的内生生成与闭环演化。确定性内生安全并非安全能力内生之外的另概念，而是其在工业互联网场景下的体系化收束，为后续前提条件、体系模型与技术架构的讨论提供了逻辑入口。

图2给出了传统安全机理与确定性内生安全机理在攻防博弈过程中的对照示意。传统机理跟随攻击事件被动推进，其防御行为发生在攻击实施之后；内生安全机理则将安全能力的组织过程前置至攻击发生之前，依托系统自身可表征、可建模、可量化与可机读的内生信息形成防护基础，使防御起点由攻击侧迁移至系统侧。

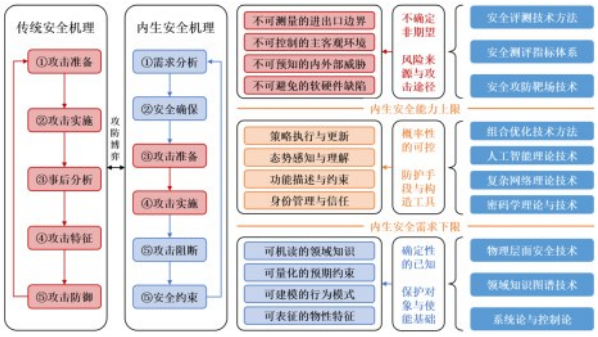

图2　工业互联网确定性内生安全攻防博弈与作用关系

7.2　确定性内生安全的前提

工业互联网装备类型相对固定、控制流程层次清晰、任务执行高度规范、通信关系长期稳定等结构化特征，使系统在多个维度上呈现出可被识别、表征与建模的稳定规律，为将系统自身确定性作为安全能力依托提供了客观条件。本文从实体、行为、结构与环境四类来源刻画其具体内涵，使之与本文所提出的四类关键使能技术形成可对应的语义基座，整体关系如图3所示。

实体确定性源于系统实体的有限性。工业互联网中的参与主体，在数量规模、种类规格与指标规范上具有显著有限性，其身份、角色、权限、物性特征与逻辑属性长期保持稳定，主体特征可被表征。

图3　工业互联网系统确定性来源

行为确定性源于系统运行的规律性。工业过程以工序节拍、控制周期与任务编排为支撑，状态转移与行为模式呈现出可建模的规律性，使任务序列、数据流向与业务交互在时间与语义上具备可描述与可约束特征。

结构确定性源于系统功能的固定性。工业装备的软硬形态、输入输出与安装配置长期保持稳定，承载功能的工业网络拓扑、功能模块关系、控制回路层次与业务依赖关系在结构层面呈现出较强的可观测性与可约束性，使系统约束可被量化。

环境确定性源于系统生态的可控性。可信根、时序基准、通信通道、资源约束与工况参数等运行条件，其上下依赖、内外交互与运维管理普遍处于可控范围，使运行条件具备可测量与可验证的基线特征。

确定性指在给定工艺规则、任务约束与运行边界条件下，存在可归纳、可表征、可建模、可验证的稳定规律，攻击与扰动恰因偏离上述规律而被察觉。可表征的物性特征、可建模的行为模式、可量化的预期约束与可规约的领域知识等四维内生信息，可分别看作支撑实体、行为、结构与环境四类确定性的具体信息形态，为后续体系模型与技术架构的构建提供可操作的前提。

此外，为统一实体、行为、结构与环境四类确定性的安全语义，将其视为安全不变量在不同系统对象上的投影。实体不变量表征设备物性特征与逻辑身份之间稳定且可验证的映射，支撑通信认证和入网鉴别。行为不变量表征系统运行的模式规律及其约束边界，支撑协同任务的安全确保。结构不变量表征设备软硬件组件的有序构成及其度量状态，支撑可信启动和终端管控。环境不变量表征系统运行条件与预期基线之间持续可验证的一致性，为其他三类不变量的成立与保持提供边界条件。形式

上，将时刻t的系统状态表示为：

$$\mathcal{X}_t = \left(E_t, B_{0:t}, G_t, \mathcal{E}_t\right) \tag{1}$$

其中，E_t表示实体特征及身份状态，$B_{0:t}$表示截至时刻t的行为轨迹，G_t表示软硬件组成及其功能依赖关系，$\mathcal{E}_t$表示系统所处的运行环境。设$\mathcal{D}_{\mathcal{E}}$为满足身份与特征唯一对应关系的实体允许域，$\mathcal{D}_{\mathcal{S}}$为满足组成关系和功能依赖约束的结构允许域，$\mathcal{D}_{\mathcal{E}}$为满足预期运行基线的环境允许域，$\mathcal{D}_{\mathcal{B}}\left(E_t, G_t, \mathcal{E}_t\right)$为给定实体、结构与环境条件下的行为允许域，则四类安全不变量可定义为：

$$I_E(t) = 1\left\{E_t \in \mathcal{D}_{\mathcal{E}}\right\} \tag{2}$$

$$I_B(t) = 1\left\{B_{0:t} \in \mathcal{D}_{\mathcal{B}}\left(E_t, G_t, \mathcal{E}_t\right)\right\} \tag{3}$$

$$I_S(t) = 1\left\{G_t \in \mathcal{D}_{\mathcal{S}}\right\} \tag{4}$$

$$I_{\mathcal{E}}(t) = 1\left\{\mathcal{E}_t \in \mathcal{D}_{\mathcal{E}}\right\} \tag{5}$$

其中，$1\{\cdot\}$为示性函数，条件成立时取值为1，否则取值为0。在本文框架下，系统处于确定性安全状态的联合判据为：

$$I_{\text{sec}}(t) = I_E(t) \wedge I_B(t) \wedge I_S(t) \wedge I_{\mathcal{E}}(t) = 1 \tag{6}$$

相应的允许安全域可表示为：

$$\mathcal{D}_{sec} = \left\{\mathcal{X}_t \mid I_{\text{sec}}(t) = 1\right\} \tag{7}$$

由于行为允许域$\mathcal{D}_{\mathcal{B}}\left(E_t, G_t, \mathcal{E}_t\right)$由实体状态、结构关系与环境条件共同确定，四类确定性虽然在分析对象上可以区分，但并非严格正交。环境给出系统运行的边界，结构限定实体的组织与交互关系，实体在既定条件下承担任务执行，行为轨迹则反映系统的实际演化状态。因此，四类确定性具有相互支撑的层次依赖关系。

不变量发生偏离表明系统状态离开允许安全域$\mathcal{D}_{sec}$而进入异常状态。异常可进一步区分为良性异常与恶意异常。设备故障、传感器误差、正常工况切换、维护操作以及环境变化等非对抗性因素引发的偏离属于良性异常。这类偏离通常局限于单一或有限维度，其残差可被已知物理模型或过程约束解释，也可通过将合法操作与典型故障模式纳入相应允许域后，使状态重新满足联合判据。而网络攻击引发的偏离属于恶意异常，其典型特征是多维不变量同时或协同失效，跨层状态演化无法用已有物理过程或合法任务转移解释，表现出对抗性残差模式。框架以四类不变量的联合保持为判定基准，综合考察偏离的维度一致性、残差可解释性以及是否可落入扩展允许域，从而实现良性异常与恶意异常的归因分类。防御与告警机制重点针对攻击造成的恶意异常实施响应，以在保持对真实攻击敏感性的同时，有效抑制良性异常导致的误报，保障工业系统的生产连续性。

7.3　确定性内生安全体系模型

与传统安全机制的被动防御不同，确定性内生安全体系模型利用系统的内在属性和特征开展安全防御，通过确定性的系统安全防御抵御不确定性的威胁攻击，体系模型主要由“四维内生信息、四类对象、三个区划、两条界限、一套测评”组成，如图4所示。将工业互联网系统抽象为四元组$S = (E,B,T,V)$，其中E为参与实体集合，B为运行行为集合，T为结构关系集合，V为环境条件集合。四类确定性可相应表述为：实体确定性指对任意$e \in E$存在稳定的特征映射$\varphi: e \to F_e$，将实体映射为可表征的物性与逻辑特征向量F_e；行为确定性指系统行为轨迹$b(t)$落于由工艺规则与任务约束界定的可达状态空间ΩB之内；结构确定性指功能依赖关系图$G = (N,L)$（N为功能节点集合，L为节点间依赖关系集合）在观测窗口内保持拓扑与层次稳定；环境确定性指运行条件参数v满足可验证的基线区间$v \in \left[v_{min}, v_{max}\right]$。可表征的物性特征、可建模的行为模式、可量化的预期约束与可规约的领域知识四维内生信息，分别构成上述四类确定性的信息载体。

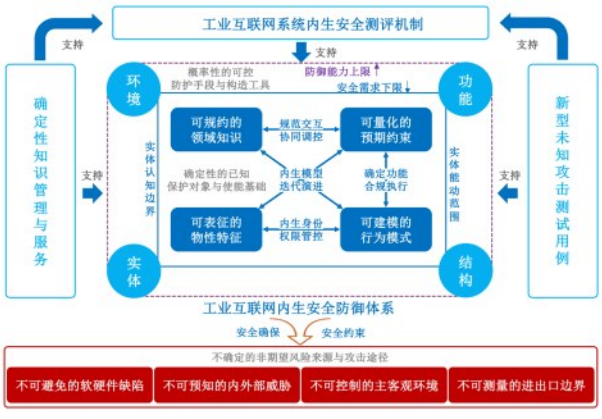

图4　工业互联网内生安全防御机理

（1）四维内生信息。可规约的领域知识是机器理解知识、表示知识的关键，是构建确定性知识库

的基础。可量化的预期约束是根据可规约的领域知识形成的判断系统运行是否正常的依据，是异常检测、攻击研判的前提。可表征的物性特征是指可以被量化和描述的内在物理或逻辑属性，是构建不可复制、不可伪造的实体内生安全身份的基础。可建模的行为模式是指实体在一定的时间和空间范围所表现出来的可形式化表示的规律性，是攻击研判、主动防御的依据。

（2）四类对象。实体指特征明确的设备、组件或服务等。结构指具有明确目标的确定性实体功能。环境指工业互联网系统的应用环境，会对实体表现和功能实现造成影响。行为指实体为实现既定功能而表现出的交互模式。针对实体和功能明确的系统，运行条件、运行状态、运行结果等信息随即确定，得到确定的系统运行规律。

（3）三个区划。确定性区域是内生安全防御的目标和对象，是内生安全防御的起点和基础。概率性区域是内生安全防御的作用范围，防御效果受到内外部因素的影响。不确定性区域是威胁攻击的来源，攻击主客体、时间等知识和情报信息未知。

（4）两条界限。安全需求确定工业互联网内生安全能力下限，是维持系统功能性能正常运行的最低要求。而内生安全防御技术也不是无条件安全的，内生安全能力上限是内生安全防御所能对抗的攻击强度的最大能力。

（5）一套测评。针对工业互联网的组合攻击方法与内生安全防御功能反馈评估测评机制，基于对抗技术的专用原子攻击设计组合方案，制定安全防护功效多维测评指标体系，实现高威胁攻击测试用例自动化组合攻击，为内生安全防御评估提供功能性能的定量和定性测评方案。

整个体系模型围绕安全约束和安全确保的需求与目标，设计基于确定性和不确定性博弈的内生安全构造框架，实现对不确定性非预期因素引发的已知和未知攻击的主动防御，确保任务符合系统约束地可靠执行。

7.4　确定性内生安全技术架构

确定性系统通过其本源特性抵御不可预测的非确定性威胁。非确定性威胁指难以完全预测或预先识别的风险，通常来源于系统未知的攻击或不可预见的环境变化，这类威胁具有高度的随机性和复杂性，可能绕过传统防护措施。基于工业互联网的内生安全体系模型，分析确定性系统与非确定性威胁之间的内在关系，提出了以内生身份确保、任务行为确保、未知攻击研判和主动安全防御为核心的四维立体化内生安全构架，具体框架如图5所示。

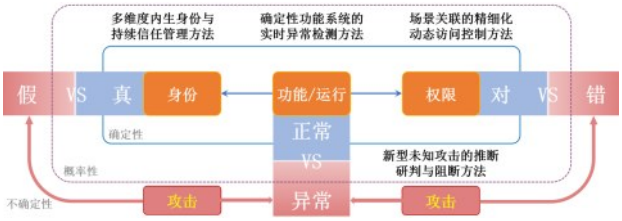

图5　工业互联网系统确定性内生安全技术架构

整个架构以确定性系统抵御非确定性威胁为基本立场，以真实身份、合规行为与正确权限所构成的确定性区域作为防御的起点与基础，以虚假身份、异常行为与错误权限所构成的不确定性区域视为威胁的来源，最终内生安全能力在两者之间的概率性区域生效，并由安全需求下限与防御能力上限界定其作用范围。

图6将上述总体逻辑沿设备、功能、网络与系统四个层面纵深展开，以知识维度、行动维度与能力维度三条主线贯穿。其中，设备层面对应点域视野，由设备特征集合、异构复合鉴权与实体物性测量构成，承担物理身份与基础属性的内生采集；功能层面对应线域视野，由服务能力列表、任务自洽保护与行为逻辑理解构成，承担工艺与任务层面的合规约束；网络层面对应面域视野，由网络性能指标、主动演进防御与交互态势感知构成，承担跨节点协同与态势研判；系统层面对应体域视野，由知识体系、推理决策与外显信息构成，承担全局学习、研判与防御反制的闭环。

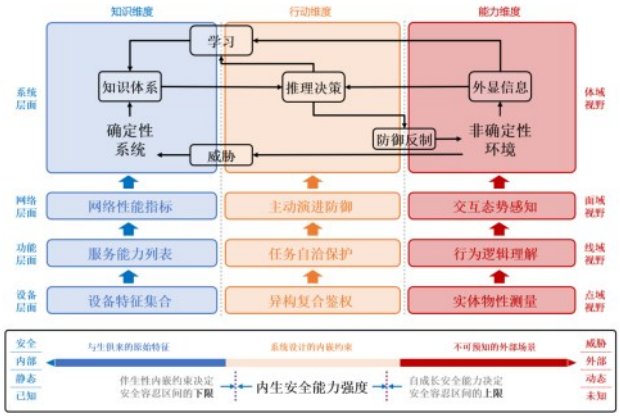

图6　工业互联网系统确定性内生安全防御架构

三条维度之间，知识维度沉淀系统的伴生性内嵌约束，对应安全容忍区间的下限；能力维度刻画系统对外部动态未知威胁的应对水平，对应安全容忍区间的上限；行动维度则在两者之间承担系统设计的内嵌约束与主动演进，决定内生安全能力强度的实际取值。该层次化展开使确定性内生安全的技术能力具备由点至体、由静至动、由已知至未知的递进形态。

8　典型应用场景与展望

工业互联网场景中，不同系统的业务约束和安全需求存在差异，确定性内生安全架构的落地需要结合具体场景进行适配。本章围绕工业互联网领域中的内生安全应用展开，讨论其由理论机制向场景能力的转化路径。

8.1　智能制造确定性安全防护实例

智能制造正由集中式刚性产线向柔性化形态演化，可移动智能装备与无人化作业单元开始大规模部署，边缘协同设备和柔性生产系统也逐步融入现代生产流程。智能制造场景具备强自治与高动态特性，传统静态防护难以兼顾身份可信、行为合规与链路受控。本节给出确定性内生安全适配方案，由状态认知、异常推理与攻击识别三个环节依次衔接。

（1）复合状态空间构建

无人化协同作业单元具有较强的自治能力，编队结构随任务动态调整，物理状态与网络行为之间存在显著耦合，状态表达具有高维与跨域特性，单一维度难以刻画其运行态势与潜在威胁。为支撑任务过程的可验证性，需将任务状态、网络行为与设备物理状态加以融合，构建具备跨层交互能力的统一状态空间。

依托任务执行记录与领域知识形成正常状态转移空间，结合异常数据与推演规则定义异常状态分支，融合得到支持行为链追踪与异常演化路径分析的复合状态空间，如图7所示。

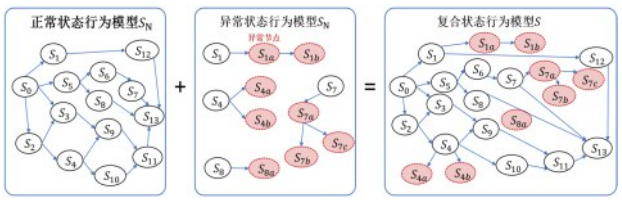

图7　复合状态转移模型

将系统状态进一步划分为任务层、网络层与设备层三个映射维度，构成三层异构耦合的复合状态空间结构，如图8所示。任务层、网络层与设备层分别刻画任务执行、通信交互与设备本体属性，三层之间通过关联规则建立超图映射，实现跨层高阶关联建模。

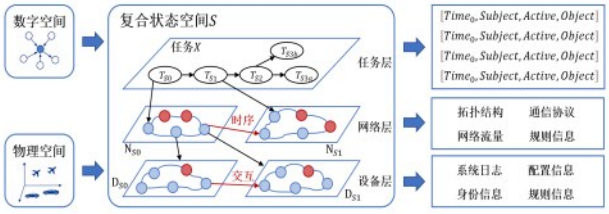

图8　复合状态空间模型

（2）多维度未知攻击研判

跨层复合状态下，异常表现为多源异构的信息形式，仅凭单点行为偏离难以判明攻击意图。为支撑确定性内生安全架构中的攻击研判与阻断闭环，需对多源异常进行结构化规约与语义关联，并基于状态空间开展图驱动的攻击识别。依据多维异常间的因果与时序特性构建异常因果链图，识别同源实体并关联交叉操作行为，生成结构与时序耦合的异常子图，刻画局部实体在特定行为路径下的异常传播特征，如图9所示。将异常子图嵌入复合状态空间进行图匹配，已知攻击直接判定，未知攻击则触发增量匹配机制进行动态风险赋权与趋势预测，如图10所示。

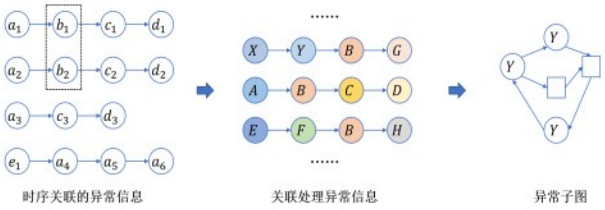

图9　异常子图生成

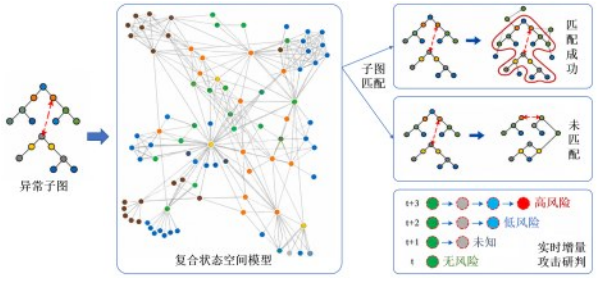

图10　攻击研判流程

（3）可组合网络攻击库

确定性内生安全架构的实战防护能力可通过对抗式训练与系统化验证形成闭环。可组合网络攻击库面向智能制造装备的多维攻击面，将多类攻击工具及其典型用例按模块化与参数化方式组织，覆盖不同攻击阶段，按装备系统的层次结构划分为上层应用与底层系统，如图 11 所示。各类工具按攻击链需求进行组合，形成可重复、可量化的对抗验证能力，为内生安全系统提供闭环评估支撑。

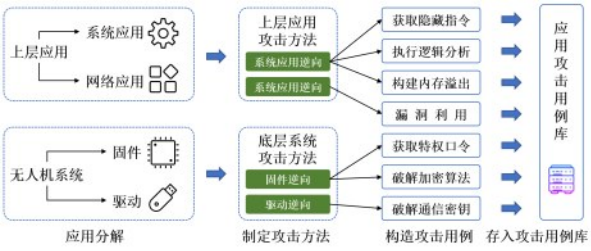

图 11　应用攻击用例库的构建方法

综上，复合状态空间提供语义底座，多维异常融合与图驱动研判承载关键能力，可组合攻击库支撑对抗验证闭环，使四维确保机制在智能制造场景可以贯通，讨论说明了确定性内生安全架构由理论向场景化防护的转化可行性。

8.2　工业 5G 安全通信实例

工业 5G 是支撑智能制造、控制系统协同和无人化作业单元等业务的重要通信基础设施。工业 5G 网元部署、切片划分与接入策略可控，软硬件功能与运行行为稳定，确定性前提显著；同时面临终端身份多样、跨切片访问频繁与攻击潜伏路径长等挑战，与确定性内生安全架构的适用条件相吻合。工业 5G 安全通信场景下，四维内生安全体系转化为相互衔接的能力组合，构成闭环内生安全防护链条，典型应用形态如图 12 所示。

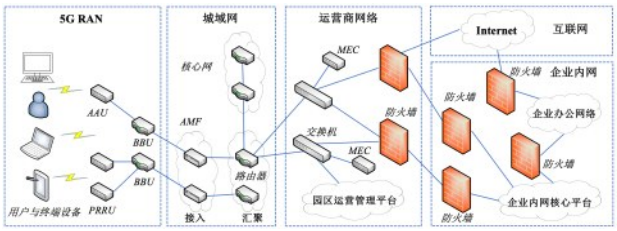

图 12　5G 领域典型应用场景

（1）内生身份与信任管理

工业 5G 网元的身份由实体的物性特征与逻辑特征协同构建，物性特征源自器件物理构造的差异性，主要包括器件指纹和信道特征；逻辑特征则基于实体行为的可观测性，包括行为特征与 I/O 时序特征，如图 13 所示。在多维特征基础上，动态构建多维内生身份，并引入持续信任管理机制，实现贯穿业务全过程的信任评估与持续认证。

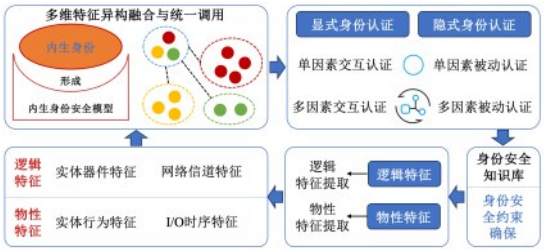

图 13　内生身份与信任管理技术

（2）精细化动态访问控制

针对工业 5G 中时空连续变换的业务场景，构建场景关联的多因素精细化动态访问控制模型，如图 14 所示。结合认证因素与任务全生命周期实施细粒度与动态持续的访问控制，实现切片内及跨切片场景下资源利用率与权限边界的协同优化。

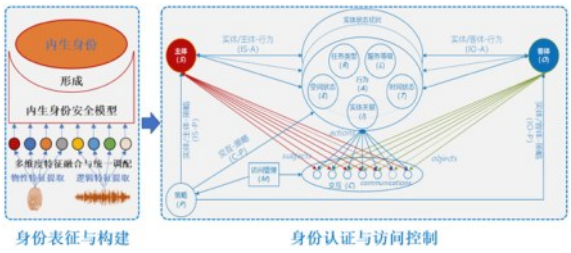

图 14　精细化动态访问控制技术

（3）实时异常行为检测

在工业 5G 网元上部署粗细两级行为信息收集器，并以此构建网络系统的确定性知识库，从中提取系统功能与行为的安全约束，进而对实时行为进行合规判别，如图 15 所示。基于确定性知识库对正常网络行为进行建模，将实时行为与约束模型进行对比匹配，在线捕获通信链路上的行为偏离。

（4）攻击推理研判与阻断

构建攻击语义知识库为异常解析提供先验指导，支撑可回溯的未知攻击研判，如图 16 所示。异常事件与知识库比对以判定已知攻击；未命中则进入知识推理流程，结合确定性任务知识研判任务破坏程度。一旦研判成立，则对受攻击网元进行精

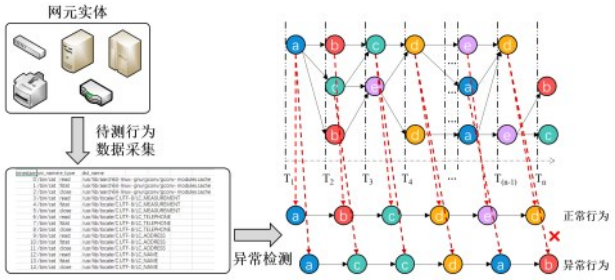

图15　实时异常行为检测技术

确调控与隔离，形成运行时阻断能力，使安全通信能力在攻击防御维度上完成闭环。

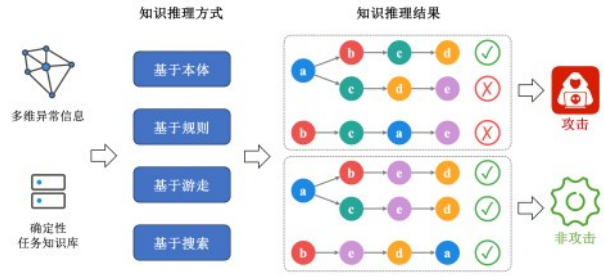

图16　攻击推理研判技术

总体而言，工业5G场景以确定性前提为支点，将四维内生安全能力嵌入关键环节，实现身份可信、权限可控、行为可验证与攻击可阻断，为切片隔离与可信接入等场景提供工程化支撑。

8.3　开放问题与未来研究方向

智能制造与工业5G两类典型场景的实例表明，确定性内生安全架构具备从理论模型向场景能力迁移的可行性，后续演进可在以下方向继续深化。

一是面向更广义工业互联网形态的能力适配。现有方案通常围绕特定工艺过程、设备类型和网络形态建立确定性模型，迁移至其他行业时容易出现知识语义不一致和模型参数失效等问题。后续应将四维确定性中具有共性的表达与行业特有约束适当分离，形成通用模型与场景配置相结合的适配方式。通用模型用于描述实体、行为、结构和环境之间的基本关系，场景配置则用于承载具体工艺规则、任务边界和连续性要求。在此基础上，研究跨场景知识映射、合法工况切换下的模型校准以及新增设备接入后的增量更新，使确定性模型能够适应系统的正常演化，避免将合法变化误判为安全偏离。表7对若干典型与新兴工业互联网场景进行初步映射。此外，四维确定性架构的场景通用性需要在新场景中重新识别可验证实体、可约束行为、稳定结构和运行基线，并据此构造异常判定与恢复路径。

二是确定性指标体系与效能评测的体系化建设。现有评测仍较多依赖检测准确率或单项性能指标，尚不足以判断安全能力是否真正来源于系统内在确定性，也难以反映防护机制对生产任务的实际保障作用。后续应建立由不变量保持状态向任务承载能力递进的指标体系，将实体可信、行为合规、结构完整和环境稳定的保持程度，与任务保持度、系统恢复时间及安全机制开销建立关联。评测环境可采用由数字孪生到硬件在环再到真实系统的逐级校准方式，使仿真环境中的规模化测试能够得到真实生产过程的验证，并提升不同技术路线之间的可比较性。

三是认知能力与对抗能力的协同演进。可组合攻击库与知识推理机制能够为未知威胁下的策略生成提供依据，但模型更新过程也可能引入错误知识，并产生不符合工业任务约束的防御动作。后续需要研究研判结论向处置策略的安全转换机制，使系统能够依据任务影响程度选择适当的约束、隔离或恢复方式，并利用执行结果持续修正确定性知识模型。策略演进过程还应具备可解释和可回滚能

表7　新兴工业互联网场景的四维确定性适配

场景	四维确定性适配	防御关注
能源互联网与工控电力	实体：设备身份、行为：调控序列 结构：电网拓扑、环境：运行基线	APT潜伏与调控偏离
轨道交通	实体：车地设备、行为：行车联锁 结构：线路依赖、环境：时空窗口	供应链与远程运维风险
空天地一体化网络	实体：星地节点、行为：协同任务 结构：动态拓扑、环境：时空信道	伪装接入与跨域横向移动
工业大模型驱动的智能决策体系	实体：模型智能体、行为：决策链 结构：工具依赖、环境：任务边界	模型污染与决策越界

力，避免因过度响应破坏控制节拍或引发新的任务偏离，从而形成认知、处置与验证相互闭合的安全演进过程。

四是标准化建设、生态共建与产业落地路径的协同推进。我国工业互联网安全标准化工作已由顶层体系规划逐步进入分对象防护、分类定级和能力评价阶段。《工业互联网综合标准化体系建设指南（2021 版）》为相关标准研制提供了总体框架[114]。2024 年发布的 GB/T 44462.1—3 分别面向应用工业互联网的工业企业、平台企业和标识解析企业提出网络安全防护要求[115-117]，2026 年发布的第 4 部分进一步补充了数据防护要求[118]。同期发布的 YD/T 6825 系列和 YD/T 6826 则开始形成与上述防护要求相配套的定级方法和防护能力评价方法[119-122]。国际上，ISA/IEC 62443 也在既有生命周期安全要求基础上，进一步完善面向资产所有者的安全项目要求和工业自动化控制系统安全保护方案[64,123-124]。

在推进内生安全体系落地方面，零信任与拟态防御代表两种不同但可以协同的落地方式。零信任以持续身份核验、最小权限和策略执行点为核心，适合在存量系统中按照业务优先级渐进部署。邬江兴院士团队提出的拟态防御则通过动态异构冗余、输出裁决和反馈清洗将防护机制嵌入系统结构。相关工程成果已由拟态构造网络设备向芯片级载体延伸。未来，应以典型工业场景为牵引，在标准化评测与产业生态支撑下，逐步形成可验证、可复制且可规模化推广的确定性内生安全落地路径。

9 结束语

本文围绕工业互联网内在的实体、行为、结构与环境四类确定性，系统综述了内生身份管理、任务行为确保、未知攻击研判、主动安全防御与安全效能评测五类关键使能技术，并构建了涵盖前提条件、体系模型与四维立体化技术架构的确定性内生安全防御体系。智能制造与工业 5G 两类典型场景的探索表明，所提架构具备由理论机制向工程防护转化的可行性。未来工作将面向能源互联网、空天地一体化网络及工业大模型驱动的智能决策系统，进一步推进通用语义底座、体系化效能评测与“认知—对抗”协同演进。

参考文献：

[1] 刘奇旭, 肖聚鑫, 谭耀康, 等. 工业互联网流量分析技术综述[J]. 通信学报, 2024, 45(8): 221-237.
LIU Q X, XIAO J X, TAN Y K, et al. Survey of industrial Internet traffic analysis technology[J]. Journal on Communications, 2024, 45(8): 221-237.

[2] 中共中央. 中共中央关于制定国民经济和社会发展第十五个五年规划的建议[EB/OL]. (2025-10-28)[2026-05-28]. https://www.news.cn/politics/20251028/08920d9f557c432e99459f8f468504db/c.html.

[3] 全国人民代表大会. 中华人民共和国国民经济和社会发展第十五个五年规划纲要[EB/OL]. (2026-03-13)[2026-05-28]. https://www.gov.cn/yaowen/liebiao/202603/content_7062633.htm.

[4] 李强. 政府工作报告: 2026 年 3 月 5 日在第十四届全国人民代表大会第四次会议上[EB/OL]. (2026-03-05)[2026-05-28]. https://www.gov.cn/yaowen/liebiao/202603/content_7062625.htm.

[5] The White House. National strategy for advanced manufacturing[EB/OL]. (2022-10-07)[2026-05-28]. https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/10/National-Strategy-for-Advanced-Manufacturing-10072022.pdf.

[6] Plattform Industrie 4.0. 2030 Vision for Industrie 4.0[EB/OL]. (2019)[2026-05-28]. https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Vision-2030-for-Industrie-4.0.html.

[7] Directorate-General for Research and Innovation. Industry 5.0: towards a sustainable, human-centric and resilient European industry[EB/OL]. (2021-01-05)[2026-05-28]. https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications.

[8] FEDERAL BUREAU OF INVESTIGATION, CYBERSECURITY AND INFRASTRUCTURE SECURITY AGENCY, NATIONAL SECURITY AGENCY, et al. Iranian-affiliated cyber actors exploit programmable logic controllers across US critical infrastructure: AA26-097A[R/OL]. (2026-04-07) [2026-07-20]. https://www.cisa.gov/news-events/cybersecurity-advisories/aa26-097a

[9] 国家互联网应急中心. 关于国家授时中心遭受美国国家安全局网络攻击事件的技术分析报告[R/OL]. (2025-10-19)[2026-05-28]. https://www.cert.org.cn/publish/main/49/2025/20251019142622914870997/20251019142622914870997_.html.

[10] DRAGOS Inc. 2026 OT cybersecurity year in review[R/OL]. Hanover: Dragos, 2026[2026-05-28]. https://www.dragos.com/ot-cybersecurity-year-in-review.

[11] STOUFFER K, PEASE M, TANG C Y, et al. Guide to operational technology (OT) security: NIST SP 800-82 Rev. 3[R]. Gaithersburg: National Institute of Standards and Technology, 2023.

[12] 黄涛, 王郅伟, 刘家池, 等. 工控协议安全研究综述[J]. 通信学报, 2024, 45(6): 60-74.
HUANG T, WANG Z W, LIU J C, et al. Survey on industrial control protocol security research[J]. Journal on Communications, 2024, 45(6): 60-74.

[13] 冯涛, 鲁晔, 方君丽. 工业以太网协议脆弱性与安全防护技术综述[J]. 通信学报, 2017, 38(S2): 185-196.
FENG T, LU Y, FANG J L. Survey on vulnerability and technology of industrial Ethernet protocol[J]. Journal on Communications, 2017, 38(S2): 185-196.

[14] 邬江兴. 论网络空间内生安全问题及对策[J]. 中国科学: 信息科学,

2022, 52(10): 1929-1937.
WU J X. Cyberspace's endogenous safety and security problem and the countermeasures[J]. Scientia Sinica(Informationis), 2022, 52(10): 1929-1937.

[15] 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.

[16] 罗兴国, 仝青, 张铮, 等. 拟态防御技术[J]. 中国工程科学, 2016, 18(6): 69-73.
LUO X G, TONG Q, ZHANG Z, et al. Mimic defense technology[J]. Strategic Study of CAE, 2016, 18(6): 69-73.

[17] ROSE S, BORCHERT O, MITCHELL S, et al. Zero trust architecture: NIST SP 800-207[R]. Gaithersburg: National Institute of Standards and Technology, 2020.

[18] 韦国华, 李挥, 白永杰, 等. 天地一体化内生安全多标识网络体系[J]. 天地一体化信息网络, 2020, 1(2): 66-72.
WEI G H, LI H, BAI Y J, et al. Space-terrestrial integrated multi-identifier network with endogenous security [J]. Space-Integrated-Ground Information Networks, 2020, 1(2): 66-72.

[19] 秦中元, 胡爱群. 可信计算系统及其研究现状[J]. 计算机工程, 2006, 32(14): 111-113.
QIN Z Y, HU A Q. Trusted computing system and its current research [J]. Computer Engineering, 2006, 32(14): 111-113.

[20] 邬江兴. 网络空间拟态防御原理: 广义鲁棒控制与内生安全[M]. 北京: 科学出版社, 2018.
WU J X. Principles of cyberspace mimic defense: generalized robust control and endogenous security[M]. Beijing: Science Press, 2018.

[21] 胡爱群, 方兰婷, 李涛. 基于仿生机理的内生安全防御体系研究[J]. 网络与信息安全学报, 2021, 7(1): 11-19..
HU A Q, FANG L T, LI T. Research on bionic mechanism based endogenous security defense system[J]. Chinese Journal of Network and Information Security, 2021, 7(1): 11-19..

[22] JAJODIA S, GHOSH A K, SWARUP V, et al. Moving target defense: creating asymmetric uncertainty for cyber threats[M]. New York: Springer, 2011.

[23] 方滨兴, 时金桥, 王忠儒, 等. 人工智能赋能网络攻击的安全威胁及应对策略[J]. 中国工程科学, 2021, 23(3): 60-66.
FANG B X, SHI J Q, WANG Z R, et al. AI-enabled cyberspace attacks: security risks and countermeasures[J]. Strategic Study of CAE, 2021, 23(3): 60-66.

[24] THEODOROPOULOS T, ROSA L, BENZAID C, et al. Security in cloud-native services: a survey[J]. Journal of Cybersecurity and Privacy, 2023, 3(4): 758-793.

[25] WARD R, BEYER B. BeyondCorp: a new approach to enterprise security[J]. ;login:, 2014, 39(6): 6-11.

[26] 奇安信战略咨询规划部, 奇安信行业安全研究中心. 内生安全: 新一代网络安全框架体系与实践[M]. 北京: 人民邮电出版社, 2021.

[27] Gartner. Implement a continuous threat exposure management (CTEM) program[R/OL]. Stamford: Gartner Inc., 2025[2026-05-28]. https://www.gartner.com/en/documents/6884566.

[28] 中兴通讯股份有限公司. 2030+ 网络内生安全愿景白皮书[R]. 深圳: 中兴通讯, 2021[2026-05-28]. https://www.zte.com.cn.

[29] 王刚, 胡鑫, 吴昊, 等. 网络生态系统动态演化[M]. 西安: 西安电子科技大学出版社, 2019.
WANG G, HU X, WU H, et al. Dynamical evolution of cyber ecosystem[M]. Xi'an: Xidian University Press, 2019.

[30] ROSS R, PILLITTERI V, GRAUBART R, et al. Developing cyber-resilient systems: a systems security engineering approach: NIST SP 800-160 Vol. 2 Rev. 1[R]. Gaithersburg: National Institute of Standards and Technology, 2021.

[31] KALOROUMAKIS P E, SMITH M J. Toward a knowledge graph of cybersecurity countermeasures[R]. McLean: The MITRE Corporation, 2020.

[32] Gartner. What Is a Digital Immune System and Why Does It Matter? [EB/OL]. Stamford: Gartner Inc., 2022[2026-05-28]. https://www.cdotrends. com/story/17584/what-digital-immune-system-and-why-does-it-matter.

[33] 中国联通研究院, 中国联通网络安全研究院, 下一代互联网宽带业务应用国家工程研究中心. 中国联通基于5G/5G-A网络的天地一体化安全白皮书[R/OL]. 北京: 中国联通, 2023[2026-05-28]. http://221.179.172.81/images/20231129/96001701241780061.pdf.

[34] 薛向阳, 邹宏, 赵进, 等. 数据基础设施抗测绘理论与技术发展研究[J]. 中国工程科学, 2025, 27(1): 72-87.
XUE X Y, ZOU H, ZHAO J, et al. Advance in anti-surveying-and-mapping theory and technologies for data infrastructure[J]. Strategic Study of CAE, 2025, 27(1): 72-87.

[35] KEPHART J O, CHESS D M. The vision of autonomic computing[J]. Computer, 2003, 36(1): 41-50.

[36] FISHER K. HACMS: high assurance cyber military systems[J]. ACM SIGAda Ada Letters, 2012, 32(3): 51-52.

[37] VMware Inc. Protecting applications with VMware NSX[R/OL]. Palo Alto: VMware Inc., 2023[2026-05-28]. https://www.vmware.com/docs/protecting-applications-with-vmware-nsx-advanced-load-balancer.

[38] 郑建华. 新形势下密码研究的思考[J]. 网络安全和信息化, 2018, (10): 38.
ZHENG J H. Thinking about cryptography research under the new situation[J]. Cybersecurity & Informatization, 2018(10): 38.

[39] 华为技术有限公司. 6G: 无线通信新征程白皮书[R/OL]. 深圳: 华为技术有限公司, 2022[2026-05-28]. https://www.huawei.com/cn/huaweitech/future-technologies/6g-white-paper.

[40] 冯登国. 机密计算发展现状与趋势[J]. 信息安全研究, 2024, 10(1): 2-5.
FENG D G. The status and trends of confidential computing[J], Journal of Information Security Research, 2024, 10(1): 2-5.

[41] NING H, FARHA F, ULLAH A, et al. Physical unclonable function: architectures, applications and challenges for dependable security[J]. IET Circuits, Devices & Systems, 2020, 14(4): 407-424.

[42] 闫高丽, 付雪, 王禹, 等. 面向射频指纹信号分析与智能识别的研究综述[J]. 南通大学学报(自然科学版), 2025, 24(2): 1-21.
YAN G L, FU X, WANG Y, et al. A survey on radio frequency fingerprint signal analysis and intelligent identification[J]. Journal of Nantong University(Natural Science Edition), 2025, 24(2): 1-21.

[43] WANG Q H, KANG M Y, YUAN L F, et al. On the application of channel characteristic-based physical layer authentication in industrial wireless networks[J]. KSII Transactions on Internet and Information Systems, 2021, 15(6): 2255-2281.

[44] SHEN G X, ZHANG J Q, MARSHALL A, et al. Deep learning-powered radio frequency fingerprint identification: methodology and case study[J]. IEEE Communications Magazine, 2023, 61(9): 170-176.

[45] ZHAO X W, LI D X, LI H. Practical three-factor authentication proto-

col based on elliptic curve cryptography for industrial Internet of Things[J]. Sensors, 2022, 22(19): 7510.

[46] SHRUTI, RANI S, SAH D K, et al. Attribute-based encryption schemes for next generation wireless IoT networks: a comprehensive survey[J]. Sensors, 2023, 23(13): 5921.

[47] YANG W C, WANG S, CUI H, et al. A review of homomorphic encryption for privacy-preserving biometrics[J]. Sensors, 2023, 23(7): 3566.

[48] LIANG Y J, SAMTANI S, GUO B, et al. Behavioral biometrics for continuous authentication in the Internet-of-Things era: an artificial intelligence perspective[J]. IEEE Internet of Things Journal, 2020, 7(9): 9128-9143.

[49] XIE N, LI Z X, TAN H J. A survey of physical-layer authentication in wireless communications[J]. IEEE Communications Surveys & Tutorials, 2020, 23(1): 282-310.

[50] AL-NAJI F H, ZAGROUBA R. A survey on continuous authentication methods in Internet of Things environment[J]. Computer Communications, 2020, 163: 109-133.

[51] RAGOTHAMAN K, WANG Y, RIMAL B. Access control for IoT: A survey of existing research, dynamic policies and future directions[J]. Sensors, 2023, 23(4): 1805.

[52] WANG W Z, HUANG H K, YIN Z M, et al. Smart contract token-based privacy-preserving access control system for industrial Internet of Things[J]. Digital Communications and Networks, 2023, 9(2): 337-346.

[53] KANG H Z, LIU G, WANG Q, et al. Theory and application of zero trust security: a brief survey[J]. Entropy, 2023, 25(12): 1595.

[54] TRABELSI R, FERSI G, JMAIEL M. Access control in Internet of Things: a survey[J]. Computers & Security, 2023, 135: 103472.

[55] LIU I H, CHOU P W, CHEN N Y, et al. Identifying Industrial Control Systems Anomalies Operation Based on Finite-State Machines[J]. Journal of Advances in Artificial Life Robotics, 2025, 4(3): 146-149.

[56] HUSSAIN M, FIDGE C, FOO E. Discovering a data interpreted Petri net model of industrial control systems for anomaly detection[J]. Expert Systems with Applications, 2023, 230: 120511.

[57] GHAEINI H R, ANTONIOLI D, BRASSER F, et al. State-aware anomaly detection for industrial control systems[C]// Proceedings of the 33rd Annual ACM Symposium on Applied Computing. New York: ACM, 2018: 1620-1628.

[58] ZHAO X, ZHANG L, CAO Y, et al. Anomaly detection approach in industrial control systems based on measurement data[J]. Information, 2022, 13(10): 450.

[59] ZHANG Y, LIU J, WANG Z, et al. MICHC-NN: An Offline Industrial Control Anomaly Detection Algorithm Based on Maximum Information Coefficient[J]. IEEE Internet of Things Journal, 2026.

[60] ALY A, MANSOUR E, YOUSSEF A. OCR-APT: Reconstructing APT stories from audit logs using subgraph anomaly detection and LLMs [C]//Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security. 2025: 261-275.

[61] QI J, LUAN Z, HUANG S, et al. LogMoE: Lightweight Expert Mixture for Cross-System Log Anomaly Detection[C]//2025 40th IEEE/ACM International Conference on Automated Software Engineering. IEEE, 2025: 330-341.

[62] FERRAIOLO D F, SANDHU R, GAVRILA S, et al. Proposed NIST standard for role-based access control[J]. ACM Transactions on Information and System Security (TISSEC), 2001, 4(3): 224-274.

[63] HU V C, FERRAIOLO D, KUHN R, et al. Guide to attribute based access control (abac) definition and considerations (draft)[R]. NIST special publication, 2013, 800(162): 1-54.

[64] International Society of Automation. ISA/IEC 62443 Series of Standards[EB/OL]. [2026-06-08]. https://www. isa. org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards.

[65] ETXEZARRETA X, GARITANO I, ZURUTUZA U, et al. Software-defined networking approaches for intrusion response in industrial control systems: a survey[J]. International Journal of Critical Infrastructure Protection, 2023, 42: 100615.

[66] CASELLI M, ZAMBON E, KARGL F. Specification mining for intrusion detection in networked control systems[C]//Proceedings of the 25th USENIX Security Symposium. Berkeley: USENIX Association, 2016: 791-806.

[67] FENG C, PALLETI V R, MATHUR A, et al. A systematic framework to generate invariants for anomaly detection in industrial control systems[C]//Proceedings of the 26th Network and Distributed System Security Symposium. Reston: Internet Society, 2019: 1-15.

[68] ZHANG B, GAO Y, YU C, et al. TAPAS: An Efficient Online APT Detection with Task-guided Process Provenance Graph Segmentation and Analysis[C]//34th USENIX Security Symposium. 2025: 607-624.

[69] SINGHAL A, OU X. Security risk analysis of enterprise networks using probabilistic attack graphs: NISTIR 7788[R]. Gaithersburg: National Institute of Standards and Technology, 2011.

[70] UNGER S, ARZOGLOU E, HEINRICH M, et al. Cybersecurity risk assessment in OT systems using attack graphs[J]. International Journal of Information Security, 2026, 25(1): 34.

[71] SAHU A, DAVIS K. Inferring adversarial behaviour in cyber-physical power systems using a Bayesian attack graph approach[J]. IET Cyber-Physical Systems: Theory & Applications, 2023, 8(2): 91-108.

[72] POOLSAPPASIT N, DEWRI R, RAY I. Dynamic security risk management using Bayesian attack graphs[J]. IEEE Transactions on Dependable and Secure Computing, 2011, 9(1): 61-74.

[73] HOLGADO P, VILLAGRÁ V A, VAZQUEZ L. Real-time multistep attack prediction based on hidden markov models[J]. IEEE Transactions on Dependable and Secure Computing, 2017, 17(1): 134-147.

[74] SHEN Y, MARICONTI E, VERVIER P A, et al. Tiresias: Predicting security events through deep learning[C]//Proceedings of the 2018 ACM SIGSAC conference on computer and communications security. 2018: 592-605.

[75] DENG A L, HOOI B. Graph neural network-based anomaly detection in multivariate time series[C]//Proceedings of the 35th AAAI Conf on Artificial Intelligence. Palo Alto: AAAI Press, 2021, 35(5): 4027-4035.

[76] GIRALDO J, URBINA D, CARDENAS A, et al. A survey of physics-based attack detection in cyber-physical systems[J]. ACM Computing Surveys, 2018, 51(4): 1-36.

[77] Dong F, Wang L, Nie X, et al. DISTDET: A Cost-Effective distributed cyber threat detection system[C]//32nd USENIX Security Symposiu. 2023: 6575-6592.

[78] LI X, JIANG X, WAN H, et al. TeRed: Normal Behavior-Based Efficient Provenance Graph Reduction for Large-Scale Attack Forensics [J]. IEEE Transactions on Information Forensics and Security, 2025.

[79] XU L, ZHAO Z, ZHAO D, et al. TraceCluster: A Lightweight and Adaptive Clustering-based Subgraph Attention Network for APT De-

tection in Provenance Graphs[J]. IEEE Transactions on Information Forensics and Security, 2026.

[80] FANG P, GAO P, LIU C, et al. Back-propagating system dependency impact for attack investigation[C]//Proceedings of the 31st USENIX Security Symposium. Berkeley: USENIX Association, 2022: 2461-2478.

[81] JIANG B, BILOT T, EL MADHOUN N, et al. ORTHRUS: Achieving High Quality of Attribution in Provenance-based Intrusion Detection Systems[C]//34th USENIX Security Symposium. 2025: 7173-7192.

[82] LI J, ZHANG R, LIU J. ProvGRP: A Context-Aware Provenance Graph Reduction and Partition Approach for Facilitating Attack Investigation[J]. Electronics, 2024, 13(1): 100.

[83] LI Z, ZENG J, CHEN Y, et al. AttacKG: Constructing technique knowledge graph from cyber threat intelligence reports[C]//European symposium on research in computer security. Cham: Springer International Publishing, 2022: 589-609.

[84] KURNIAWAN K, EKELHART A, KIESLING E, et al. KRYSTAL: knowledge graph-based framework for tactical attack discovery in audit data[J]. Computers & Security, 2022, 121: 102828.

[85] WANG L, WU D. Seccontrol: Bridging the gap between security tools and sdn controllers[C]//International Conference on Security and Privacy in Communication Systems. Cham: Springer International Publishing, 2017: 11-31.

[86] CHAVEZ A R. Moving target defense to improve industrial control system resiliency[M]//Industrial Control Systems Security and Resiliency: Practice and Theory. Cham: Springer International Publishing, 2019: 143-167.

[87] ETXEZARRETA X, GARITANO I, ITURBE M, et al. Low delay network attributes randomization to proactively mitigate reconnaissance attacks in industrial control systems[J]. Wireless Networks, 2024, 30 (6): 5077-5091.

[88] JAVADPOUR A, JAFARI F, TALEB T, et al. A comprehensive survey on cyber deception techniques to improve honeypot performance[J]. Computers & Security, 2024, 140: 103792.

[89] FRANCO J, ARIS A, CANBERK B, et al. A survey of honeypots and honeynets for Internet of things, industrial Internet of things and cyber-physical systems[J]. IEEE Communications Surveys & Tutorials, 2021, 23(4): 2351-2383.

[90] MAESSCHALCK S, GIOTSAS V, GREEN B, et al. Don't get stung, cover your ICS in honey: how do honeypots fit within industrial control system security[J]. Computers & Security, 2022, 114: 102598.

[91] ALT L, BEVERLY R, DAINOTTI A. Uncovering network tarpits with degreaser[C]//Proceedings of the 30th Annual Computer Security Applications Conf. New York: ACM, 2014: 156-165.

[92] GRIFFIOEN H, DOERR C. Could you clean up the Internet with a Pit of Tar? Investigating tarpit feasibility on Internet worms[C]//2023 IEEE Symposium on Security and Privacy. IEEE, 2023: 2551-2565.

[93] LI Z, WEI X, LI C, et al. Generating detectors from anomaly samples via negative selection for network intrusion detection[J]. Scientific Reports, 2025, 15(1): 36456.

[94] KIM Y J, NAM W, LEE J. Multiclass anomaly detection for unsupervised and semi-supervised data based on a combination of negative selection and clonal selection algorithms[J]. Applied Soft Computing, 2022, 122: 108838.

[95] FENG X, WU J, WU Y, et al. Blockchain and digital twin empowered trustworthy self-healing for edge-AI enabled industrial Internet of things[J]. Information Sciences, 2023, 642: 119169.

[96] SUDHAKAR K, KUMAR A, ARCHANA R A, et al. Anomaly detection based self-healing mechanism using dynamic diffusion spatial-temporal graph convolutional network in industrial IoT[J]. Knowledge-Based Systems, 2026, 331: 114812.

[97] Trusted Computing Group. TCG guidance for securing industrial control systems[R]. Beaverton: Trusted Computing Group, 2022.

[98] BAI J, MA J F, WANG Y D, et al. Survey on application of trusted computing in industrial control systems[J]. Electronics, 2023, 12(19): 4182.

[99] SUN R, ZHU Y, FEI J, et al. A survey on moving target defense: Intelligently affordable, optimized and self-adaptive[J]. Applied Sciences, 2023, 13(9): 5367.

[100] LÓPEZ P B, PÉREZ M G, VASILOMANOLAKIS E, et al. Reactive cyber deception: Stealth-based adaptive redirection to on-demand honeypots with AI-driven data generation[J]. Computer Networks, 2026: 112203.

[101] International Organization for Standardization, International Electrotechnical Commission. ISO/IEC 15408-1:2026 Information security, cybersecurity and privacy protection: evaluation criteria for IT security: Part 1: Introduction and general model[S]. Geneva: ISO, 2026.

[102] International Organization for Standardization. ISO/IEC/IEEE 29119-1:2022 Software and systems engineering — Software testing — Part 1: General concepts[S]. Geneva: ISO, 2022.

[103] IEEE. IEEE Std 1012-2024: IEEE Standard for system, software, and hardware verification and validation[S]. New York: IEEE, 2025.

[104] Joint Task Force. Security and privacy controls for information systems and organizations: NIST SP 800-53 Rev. 5[R]. Gaithersburg: National Institute of Standards and Technology, 2020.

[105] SCARFONE K, SOUPPAYA M, CODY A, et al. Technical guide to information security testing and assessment: NIST SP 800-115[R]. Gaithersburg: National Institute of Standards and Technology, 2008.

[106] 全国信息安全标准化技术委员会 . GB/T 28448-2019 信息安全技术 网络安全等级保护测评要求[S]. 北京: 中国标准出版社, 2019.
National Information Security Standardization Technical Committee. GB/T 28448-2019 Information security technology — Evaluation requirement for classified protection of cybersecurity[S]. Beijing: Standards Press of China, 2019.

[107] 全国信息安全标准化技术委员会 . GB/T 36466-2018 信息安全技术 工业控制系统风险评估实施指南[S]. 北京: 中国标准出版社, 2018.
National Information Security Standardization Technical Committee. GB/T 36466-2018 Information security technology — Implementation guide to risk assessment of industrial control systems[S]. Beijing: Standards Press of China, 2018.

[108] STROM B E, APPLEBAUM A, MILLER D P, et al. MITRE ATT&CK: design and philosophy[R]. McLean: The MITRE Corporation, 2018.

[109] BODEAU D J, GRAUBART R D, MCQUAID R M, et al. Cyber resiliency metrics, measures of effectiveness, and scoring: MTR180314 [R]. Bedford: The MITRE Corporation, 2018.

[110] MATHUR A P, TIPPENHAUER N O. SWaT: a water treatment testbed for research and training on ICS security[C]//Proceedings of the International Workshop on Cyber-physical Systems for Smart Water

Networks. Piscataway: IEEE, 2016: 31-36.
[111] ANTONIOLI D, TIPPENHAUER N O. MiniCPS: a toolkit for security research on CPS networks[C]//Proceedings of the 1st ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy. New York: ACM, 2015: 91-100.
[112] CONTI M, DONADEL D, TURRIN F. A survey on industrial control system testbeds and datasets for security research[J]. IEEE Communications Surveys & Tutorials, 2021, 23(4): 2248-2294.
[113] YAMIN M M, KATT B, GKIOULOS V. Cyber ranges and security testbeds: scenarios, functions, tools and architecture[J]. Computers & Security, 2020, 88: 101636.
[114] 工业和信息化部, 国家标准化管理委员会. 工业互联网综合标准化体系建设指南(2021版)[R/OL]. 2021[2026-07-21].
[115] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 44462.1—2024 工业互联网企业网络安全第1部分:应用工业互联网的工业企业防护要求[S]. 2024.
[116] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 44462.2—2024 工业互联网企业网络安全第2部分:平台企业防护要求[S]. 2024.
[117] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 44462.3—2024 工业互联网企业网络安全第3部分:标识解析企业防护要求[S]. 2024.
[118] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 44462.4—2026 工业互联网企业网络安全第4部分:数据防护要求[S]. 2026.
[119] 工业和信息化部. YD/T 6825.1—2026 工业互联网企业网络安全定级方法第1部分:应用工业互联网的工业企业[S]. 2026.
[120] 工业和信息化部. YD/T 6825.2—2026 工业互联网企业网络安全定级方法第2部分:平台企业[S]. 2026.
[121] 工业和信息化部. YD/T 6825.3—2026 工业互联网企业网络安全定级方法第3部分:标识解析企业[S]. 2026.
[122] 工业和信息化部. YD/T 6826—2026 应用工业互联网的工业企业网络安全防护能力评价方法[S]. 2026.
[123] INTERNATIONAL SOCIETY OF AUTOMATION. ANSI/ISA-62443-2-1-2024 Security for industrial automation and control systems—Part 2-1: Security program requirements for IACS asset owners[S]. Research Triangle Park: ISA, 2024.
[124] INTERNATIONAL SOCIETY OF AUTOMATION. ISA-TR62443-2-2-2025 Security for industrial automation and control systems—Part 2-2: IACS security protection scheme[R]. Research Triangle Park: ISA, 2025.

张志为（1986– ），男，河南开封人，博士，西安电子科技大学副教授、博士生导师，主要研究方向为智能无人系统安全。

汤贵源（1999– ），男，河南平顶山人，西安电子科技大学博士研究生在读，主要研究方向为无人系统安全、无线网络安全。

沈玉龙（1978– ），男，江苏泗洪人，博士，西安电子科技大学教授，主要研究方向为云计算与数据安全、智能网络内生安全等。

何吉（1989– ），男，重庆人，博士，西安电子科技大学副教授、硕士生导师，主要研究方向为物理层安全、智能信息网络内生安全。

张涛（1986– ），男，陕西西安人，博士，西安电子科技大学副教授、硕士生导师，主要研究方向数据安全、大模型安全等。

马建峰（1963– ），男，陕西西安人，博士，西安电子科技大学教授、博士生导师，主要研究方向为无线网络安全、移动智能系统安全等。