•聚焦国家重点研发计划•

DOI:10.12454/j.jsuese.202600179

本刊网刊

面向城市轨道交通系统的网络安全智能防御关键技术研究构想与成果展望

李贝贝[1],王耀萱[1],杜卿芸[1],冯家成[1],黄晓波[2]

(1.四川大学 网络空间安全学院,四川 成都 610065;2.宁波和利时信息安全研究院有限公司,浙江 宁波 315301)

摘　要:近年来,工业数字化与智能化加速推进,网络空间边界不断向物理世界延伸,城市轨道交通系统作为国家关键信息基础设施正面临日益严峻的网络安全威胁。城轨工控系统网络环境呈现广域、动态、异构等复杂特点,网络覆盖范围广、运行状态动态变化、设备类型与通信协议多样,对网络安全防御的及时性、准确性、自适应能力及协同管控能力提出了更高要求。传统网络安全防御技术在此环境下具有较大的盲目性和被动性,及时、准确发现和识别广域网络威胁难度大,防御方式静态被动,难以随网络环境与威胁变化动态调整,且对多类异构安全设备的统一管控与协调能力较弱。因此,仍面临网络威胁监测准确性和时效性低、防御方式静态被动、统一管控与协调能力弱等技术挑战。为此,本文聚焦于复杂城轨工控系统网络环境,以“人工免疫”和“大模型”等智能化技术为创新手段,围绕广域网络威胁发现时效性与准确性、动态网络威胁主动防御及异构网络安全设备协同管控等关键科学问题,分别开展三项关键技术及一套原型系统研究。通过基于免疫识别的城轨工控系统网络安全风险监测技术研究,融合多源异构数据并构建统一免疫表征,结合层次聚类、检测器耐受训练与簇内快速应答,利用自体与非自体检测器的双重验证机制,实现对工控网络威胁的快速发现与精准识别;基于免疫进化的城轨工控系统网络威胁主动防御技术研究,利用克隆增殖与层次差分进化开展威胁基因演化,结合随机攻防博弈和深度学习,研究防御策略生成与自适应动态优化,实现对工控网络的主动安全防护;基于大模型的城轨工控系统网络安全设备协同管控技术研究,汇聚云边端异构威胁情报,利用大模型的知识推理与生成能力形成跨域联动管控方案,实现兼容多类异构网络安全设备的协同防御。在此基础之上,构建面向城市轨道交通系统的网络安全智能防御原型系统,对研究成果展开技术验证。本文提出的面向城市轨道交通系统的网络安全防御新技术,突破了传统网络安全以“打补丁”为主的被动防御技术瓶颈,为国家关键信息基础设施的安全可靠运行提供技术支撑。

关键词:城市轨道交通系统;工业控制系统;网络威胁;主动防御;人工免疫;大模型

中图分类号:TP393.08　　**文献标志码:**A　　**文章编号:**2096-3246(2026)06-0001-14

近年来,国家关键信息基础设施往往承载电力、交通、通信等关键公共服务,一旦遭受攻击,可能导致运营中断、客流滞留甚至安全事故,直接威胁城市运行秩序与公共安全,并可能引发跨行业、跨区域的连锁反应。2024年,美国联邦调查局发布的网络犯罪报告中汇总了263 455条网络威胁事件信息,其中,涉及关键基础设施部门的事件达4 878条[1],反映出针对关键信息基础设施的攻击活动具有持续性与现实危害。在“俄乌”冲突等地缘政治事件影响下,交通、电力等关键信息基础设施在网络对抗中的“前线”地位进一步凸显,逐步演化为地缘政治博弈中的关键脆弱点与战略支点。

在此背景下,城市轨道交通系统作为国家关键信息基础设施的重要组成部分,承载着高密度客流运输与关键民生保障功能[2]。随着“智慧地铁”[3]、“云信号系统”[4]等新形态快速落地,城轨工控系统架构正由传统相对封闭的专用系统向“云边协同、车地一体”的互联化架构演进,业务系统与云平台、控制中心、车站及区间网络通过IP互联程度显著提升。已有综述指出,互联化城轨系统在提升运行效率和管理能力的同时,也扩大了网络攻击暴露面,使网络安全与运营可靠性之间的耦合关系更加紧密[5]。

收稿日期:2026－03－13　　**修回日期:**2026－10－14　　**网络出版日期:**2026－10－20

基金项目:国家重点研发计划项目(2024YFB3108300)

作者简介:李贝贝(1992—),男,研究员,博士. 研究方向:工业控制系统安全. E-mail: libeibei@scu.edu.cn

这种架构演进使城轨网络安全问题不同于一般信息系统，也不同于常规工业控制系统中的单点设备安全问题。一方面，跨域互联为攻击者提供了横向移动路径，使其可能由终端设备渗透至控制层、业务层乃至云平台，形成跨站点、跨线路、跨子系统的复杂攻击链；另一方面，新旧系统并存、多厂商设备与专有协议混杂，使资产识别、脆弱性管理和安全配置一致性维护更加困难，风险难以统一识别与持续治理。已有案例表明，加拿大、英国等国家相继出现售票系统瘫痪、列车调度系统遭恶意控制等安全事件，说明城轨系统遭受网络攻击后，其影响不仅表现为数据泄露或服务中断，还可能波及列车运行组织、信号控制、综合监控、车站机电和乘客服务等关键业务，进而对运营秩序和公共安全产生连锁影响[6]。因此，城轨系统具有广域互联、实时运行、业务耦合紧密、低容错和高公共安全影响等特点，其网络安全防护不能仅依赖静态规则、边界隔离和事后处置，而需要具备面向多源异构数据的持续感知能力、面向未知威胁的快速识别能力、面向动态攻击过程的主动防御能力，以及面向多站点、多线路和多设备场景的协同管控能力。

针对上述风险，全球城市轨道交通网络安全投入持续上升。2025 年发布的全球城轨网络安全市场报告显示，全球包括地铁在内的铁路网络安全市场规模已达 89.6 亿美元，投入重点聚焦远程监控、安全分析与网络管理等解决方案[7]。中国国内在标准规范与安全防护体系建设层面亦同步推进，已发布 GB/T 26718—2024[8]、GB/T 46739—2025[9]等相关城轨安全标准，对安全分区与通道划分、生命周期安全管理及安全防护给出了框架性指导。然而，面向具体城轨业务场景的风险监控、动态防御、设备协同管控方面仍存在短板：现有监测手段多依赖静态特征规则与单点告警，难以有效识别多阶段、强隐蔽攻击链，易产生误报、漏报并导致响应滞后；防御处置仍以固定策略与补丁式加固为主，缺乏随业务状态与威胁动态调整的机制；同时，跨车站、跨线路、跨子系统的安全能力分散，设备统一管控响应不足，难以形成高效联动与闭环对抗。

综上，城市轨道交通系统网络环境具有广域分布、实时运行、业务耦合紧密、设备高度异构和安全影响外溢等特征。现有以静态规则、边界防护和被动响应为主的安全机制难以适应城轨系统中持续演化、跨域传播和隐蔽渗透的网络威胁。因此，亟须面向城市轨道交通系统构建新一代网络安全智能防御技术方案。人工免疫系统具有自体/非自体区分、未知异常识别、免疫记忆、动态进化和分布式协同等特点，与城轨网络安全中持续监测、未知威胁发现、主动防御和协同管控的需求具有较高契合度。基于此，本文引入人工免疫与大模型（LLM）等智能化技术，围绕广域威胁监测、动态主动防御和异构设备协同管控等方向展开研究构想，以期推动城轨网络安全防御由被动补丁式防护向智能主动防御转变。

1 国内外研究现状

近年来，国内外众多学者围绕城市轨道交通系统网络安全开展了大量研究。本文主要从网络安全风险监测、安全防御、设备协同管控 3 个方面对相关研究进行梳理。总体而言，机器学习、深度学习与联邦学习等智能方法已被广泛用于威胁识别与检测，并逐步延伸至防御策略生成与跨域协同管控。相关研究也开始面向城轨的列控系统、通信系统与综合监控系统等典型业务场景推进，推动防护方案由“单点算法”向“跨层、多域协同”的一体化智能防御框架演进。

1.1 网络安全风险监测技术

在网络安全风险监测方面，现有研究主要围绕城市轨道交通系统的网络数据表征、异常检测、入侵识别和告警预警等环节开展。相关研究正由传统基于规则、流量统计特征和协议字段的检测方法，逐步向机器学习、深度学习、联邦学习、轻量化神经网络和人工免疫等智能化方法发展，以提升复杂工控环境下的威胁识别能力和未知攻击发现能力。

早期研究主要依赖网络流量特征、协议字段和过程量测数据构建检测模型。随着工业控制系统和城轨系统网络化程度不断提升，相关研究开始引入深度学习和联邦学习等方法。Li 等[10]提出基于深度联邦学习的工业控制系统（ICS）入侵检测方法，在保护数据隐私的同时提高了多类型攻击检测效果。Cao 等[11]面向铁路入侵检测提出 RailDet 方法，通过动态入侵区域和轻量级神经网络兼顾检测精度与实时性。Walita 等[12]将深度学习与统计分析结合，用于 ICS 与关键基础设施异常检测。丁朝晖等[13]构建多维工控系统网络安全风险监测预警系统，引入基于广义径向基函数的 AI 行为建模，实现对已知与未知威胁的高精度识别与预警。郭伟杰等[14]面向无人值守 ICS 建立多目标优化框架，对现场端与监控端的信息安全策略和功能安全策略进行量化权衡与联合决策，从而在满足安全约束的同时提升整体风险控制能力。Zhu 等[15]系统梳理了机器学习在城市轨道交通状态感知和安全监测中的应用，指出深度学习、联邦学习等新范式有望构建端–云协同与多模态融合的风险感知框架。Huang 等[16]受人工免疫与差分进化启发，构建一种面向未知网络攻击检测的系统，提出面向未知网络攻击检测的双向演化检测方法，提升了识别未知威胁的能力。

尽管现有风险监测研究已在特征建模、异常检测和未知威胁识别方面取得一定进展，但面向城市轨道交通复杂场景仍存在不足。首先，真实城轨攻击样本和公开基准数据集仍较为有限，许多模型主要依赖仿真环境或实验室数据进行训练和验证，导致其在真实运营环境中的泛化能力仍需进一步检验[17−18]。其次，深度学习方法虽然能够提升检测精度，但其黑盒特性削弱了告警结果的可解释性，而轨道安全告警往往与列车运行控制、信号系统处置和运维决策相关，对告警原因和处置依据的可解释性要求更高[19]。此外，城轨系统具有高度异构性，关键通信过程涉及厂商私有协议和特定工业网络协议，通用入侵检测系统难以充分解析其业务逻辑[20]。随着综合监控系统、企业 IT 网络和移动互联网平台逐步融合，攻击路径也由单点攻击扩展为跨 IT/OT 边界的复杂攻击链，现有研究对多阶段攻击过程的协同检测能力仍显不足[21]。最后，车载控制单元、轨旁设备和边缘节点等现场设备计算资源有限，而基于通信的列车控制系统（CBTC）等系统对实时性和可靠性要求较高，复杂深度模型、图计算模型或 Transformer 模型在现场部署时可能面临计算开销和响应时延约束[22]。因此，如何在可控成本下构建高效、可解释、可部署的多源异构风险监测模型，并兼顾威胁发现的时效性与识别准确性，仍是亟须突破的关键问题。

1.2　网络安全防御技术

在网络安全防御方面，相关研究主要围绕纵深防护、跨层联动防御、移动目标防御、强化学习策略优化和主动缓解机制展开，由传统的边界隔离和静态规则防护，逐步向智能化、动态化和跨层协同防御方向发展。

早期的研究主要以边界隔离、分区防护和纵深防御为基础。刘磊等[23]从站控设备安全、防护链路与网络边界等视角提出 ICS 信息安全多层防御体系，为关键基础设施场景下的分层防护提供了参考。Li 等[24]面向轨道交通工业控制系统提出网络安全防御技术，通过网络分区、边界隔离和纵深防护降低攻击面。Li 等[25]设计了由安全监测、攻击防御和应急处置组成的智能工业控制网络安全防御系统，强调面向业务流程的全生命周期安全管控。上述研究为轨道工控系统的基础防护体系建设提供了重要支撑，但整体仍较依赖静态边界、预设规则和人工配置，面对内部威胁、远程运维接入、云边协同架构及高级持续性攻击时，防护能力仍存在一定局限。

随着攻击行为呈现多阶段、隐蔽化和动态演化特征，相关研究开始关注跨层联动和主动防御机制。Zhong 等[26]提出高效并行强化学习方法，用于工业控制系统跨层防御机制设计，通过联合优化物理过程层、控制层和网络层的防御动作，提高系统对复杂攻击的鲁棒性。Zhou[27]和 Zhang[28]等分别在工业物联网和车联网场景中引入移动目标防御，通过动态迁移和服务扰动增强系统抗攻击能力。陈大等[29]进一步将数据冲突与分布式拒绝服务攻击防御统一建模，利用资源自适应分配机制提升防御联动能力。这类方法推动安全防御由被动阻断向动态调整和主动缓解转变，但现有研究多面向通用工控系统、工业物联网或车联网场景，对城市轨道交通系统中强实时性、强确定性和业务连续性约束考虑仍不充分。

在城市轨道交通场景下，CBTC、列车自动监控、防护系统及 SCADA 等核心系统的网络化与互联化程度不断提升，城轨系统逐渐由相对封闭的专用系统演变为典型的信息物理融合系统，网络攻击可能进一步影响物理控制过程和系统运行安全[30]。在技术体系层面，边界隔离、分区防护、工业防火墙、DMZ 隔离及 VLAN 划分等方法仍是工程部署中的基础手段。然而，这类方法主要依赖静态信任边界，面对内部威胁、远程运维接入、云边协同架构及高级持续性攻击时，防护能力存在一定局限[31]。同时，工业协议安全研究逐渐从传统深度包检测转向协议语义理解与行为建模，通过对 Modbus TCP、IEC 60870−5−104 等典型工业协议进行深度解析，实现异常指令识别和白名单控制[32]。但现有方法多聚焦于单协议、单子系统或单类攻击场景，对跨协议语义关联、跨系统攻击链识别和业务逻辑约束下的联动防御支持仍然不足。因此，如何在不影响列车运行安全和业务连续性的前提下，对潜在攻击行为、攻击路径和威胁模式进行推理与预判，并生成可动态调整的自适应防御策略，是当前亟须解决的关键问题。

1.3　网络安全设备管控技术

在网络安全设备管控方面，随着城市轨道交通系统规模的扩张和设备类型的激增，面向海量异构安全设备与业务系统进行统一管控与协同成为研究热点。相关研究经历了由工控数据集成与设备运维管理，向异构平台协同控制，再向智能化策略编排与跨域联动处置逐步演进的过程。

早期的研究主要关注工控数据集成、设备状态管理和基础数据统一维护。Trunzer 等[33]提出一种面向自动化生产系统异构数据源的灵活架构，简化了数据集成、处理和跨组织边界共享流程，降低了工控数据分析复杂度。美国罗克韦尔自动化公司提出 ThinManager 系统[34]，面向工业网络环境提供终端设备管理和内容交付能力，提升了工业终端设备的集中管理水平。在城市轨道交通场景中，施聪等[35]构建轨旁信号设备全生命周期数据管理体系，从数据模型、安全管理和双向反馈

等方面统一组织设备设计、制造与运维阶段的信息。黄苏苏等[36]提出区域中心–局管平台–本地工作站的3级电子地图数据管理架构，实现跨线路基础数据的统一存储、分发与权限管控。上述研究为城轨设备状态管理、基础数据维护和运维支撑提供了重要基础，但整体仍偏重数据管理和设备运维，对网络威胁情报融合、跨设备安全联动和动态处置策略生成的支持相对有限。

随着城轨工控网络向“云–边–端”融合架构发展，传统依赖人工配置和静态规则的管控模式逐渐暴露出自动化不足、响应滞后和策略一致性难以保障等问题。相关报告也指出，铁路网络安全建设正在更加关注服务连续性、恢复能力和跨系统安全运营能力[37]。因此，相关研究开始面向异构平台协同、跨域访问控制和安全数据共享开展探索。Zhang等[38]将联邦学习与深度强化学习结合，选择数据质量较高的设备节点参与模型聚合，在隐私保护的同时降低通信开销，为大规模异构安全设备的协同调度提供了思路。Lu等[39]提出面向ICS传感器的区块链安全数据存储协议，通过群签名、智能合约和代理重加密实现云端外包数据的匿名认证与细粒度访问控制，在降低传感器计算、通信开销的同时强化了多角色数据访问管理能力。Koo等[40]进一步提出面向工业物联网异构平台互操作的访问控制框架，通过区块链和元数据注册机制实现跨平台用户属性共享、资源访问控制与安全互操作。与此同时，零信任架构通过持续验证、永不信任的机制，将安全防护重点由静态边界隔离转向动态身份认证、行为审计和细粒度访问控制，为异构平台间的微隔离和动态授权提供了重要支撑[41]。

近年来，大语言模型（LLM）的发展进一步推动城轨安全管控向自主化与智能化演进。LLM可用于网络安全中的威胁情报分析、漏洞分析、恶意行为识别、安全运营辅助和自动化响应等任务，为安全运营中心提升告警研判和事件处置效率提供了新的技术路径[42]。在工业物联网和关键基础设施场景下，LLM也可辅助理解多源异构安全数据，生成安全分析结果和处置建议，从而增强复杂系统的智能化防护能力[43]。因此，相比传统依赖人工编排的安全策略，LLM有望对跨线路、跨子系统的日志、告警、威胁情报和处置规则进行语义关联分析，并辅助生成联动处置策略。总体来看，未来城轨网络安全管控将逐步形成以数据驱动和人工智能为核心的智能协同体系，为关键基础设施稳定运行提供更加自动化的安全保障。因此，如何提高异构安全设备之间的协同和全局优化能力，形成具有高效决策能力的跨域联动管控方案，是当前亟须解决的重要问题。

为进一步明确本研究构想与现有城市轨道交通网络安全体系、工程方案和相关研究之间的关系，选取城轨网络安全标准规范、工业互联网信息管控指南、工程化安全防护案例及典型论文研究进行横向对比，见表1。

表1 城市轨道交通网络安全相关研究对比

Tab.1 Comparison of related studies on urban rail transit cybersecurity

代表性研究	类型	面向对象	主要内容	主要局限
DB31/T 1499—2024[44]	标准	城市轨道交通网络安全建设	规定城轨网络安全整体架构、集中防护、安全区域边界、访问控制、基础设施安全等要求	侧重规范性要求和安全基线，未展开具体智能检测、主动防御和协同管控方法
T11/AII 022—2021[45]	应用指南	城轨工业互联网信息管控系统	定义城轨工业互联网信息管控系统体系架构，支撑数据汇聚、综合利用和运维管理	侧重信息管控和数据综合利用，对网络威胁识别与主动防御机制支撑不足
威努特—飞腾轨道交通信号系统网络安全综合解决方案[46]	工程方案	城轨信号系统	面向控制中心、车辆段、停车场和设备集中站等关键节点，部署工控安全监测审计、主机防护、日志审计和统一安全管理能力	主要面向信号系统安全防护与审计，对跨线路、跨子系统的智能联动防御支持有限
启明星辰地铁综合监控系统安全防护项目[47]	工程案例	地铁综合监控系统、信号系统等	按照等级保护要求开展入侵检测、访问控制、漏洞检查、安全配置检查和工控安全统一管理	工程建设和安全运营落地，公开资料对动态防御算法和智能策略生成机制披露有限
RailDet入侵检测方法[11]	论文研究	铁路/城轨入侵检测场景	通过动态入侵区域和神经网络提升入侵检测精度与实时性	主要面向检测环节，对主动防御、策略生成和设备联动支持不足
轨道交通工控网络安全防御[24]	论文研究	轨道交通工业控制系统	通过网络分区、边界隔离和纵深防御降低攻击面	依赖静态边界和预设策略，对攻击演化和自适应防御支持不足
CBTC跨层防御研究[32]	论文研究	边缘智能赋能的CBTC系统	面向中间人攻击提出跨层防御方案，结合通信层和控制层信息降低攻击影响	主要面向特定攻击类型和CBTC场景，对多源威胁识别和跨设备协同管控覆盖不足
本文	研究构想	轨道交通工业控制系统	融合人工免疫和大模型，构建覆盖风险监测、未知威胁识别、主动防御和设备协同管控的智能防御体系	当前属于阶段性研究构想与成果展望，后续开展第三方实验验证

由表1可以看出，现有工作已从网络安全规范、安全基线建设、信息管控、工程防护、入侵检测和跨层防御等方面为城轨网络安全建设提供了重要的支撑[11,24,32,44–47]。然而，相关工作多数仍侧重规范性要求、数据管控、单系统工程防护、特定场景检测或局部防御能力，对城轨场景下多源异构数据融合、未知威胁识别、动态主动防御及跨站点、跨线路、跨设备协同管控的支撑仍显不足。因此，本文进一步面向城市轨道交通复杂业务场景，构建覆盖风险监测、主动防御和设备协同管控的一体化网络安全智能防御技术路线。

2　研究内容

2.1　关键科学问题

城轨工控系统网络环境具有广域、动态、异构等复杂特点。与一般信息系统相比，城轨系统的网络安全问题不仅关乎数据和系统可用性，还直接关联列车运行组织、站场设备控制和公共安全[48]。与传统工业控制系统相比，城轨系统又具有多线路、多站点、多子系统联动运行的特点，攻击行为可能在控制中心、车站、车辆段、区间设备和云平台之间跨域传播，形成多阶段、强隐蔽、难溯源的攻击链。因此，面向城轨场景的网络安全智能防御需要突破“看得准、判得早、防得住、协同快”的关键难题[49]。

人工免疫技术为解决上述问题提供了新的思路，其自体/非自体区分机制可用于刻画城轨系统正常行为与异常行为之间的边界[50]，免疫识别机制可用于未知威胁发现，免疫进化机制可用于动态攻击行为推演和防御策略更新，免疫记忆与协同响应机制可用于支撑跨站点、跨线路和跨设备的联动防御。然而，将人工免疫技术应用于城轨网络安全仍需突破多源异构数据免疫表征、动态自体边界建模、未知威胁识别、威胁基因演化和跨域协同管控等关键科学问题。基于此，本研究凝练三大关键科学问题：

1）广域城轨工控系统网络威胁发现时效性与准确性问题。城轨系统覆盖控制中心、车站、车辆段、区间设备、列车网络和云平台等多类对象，网络数据来源广、类型多、语义差异大。攻击者可能利用多系统互联关系实施跨域渗透和横向移动，使单点告警难以及时反映整体威胁态势。因此，如何面向流量、日志、告警、协议指令和设备状态等多源异构数据，构建统一的免疫表征方法和高效风险监测模型，兼顾威胁发现的时效性与识别准确性，是实现城轨网络安全智能防御需要解决的首要科学问题。

2）动态城轨工控系统网络威胁主动防御问题。城轨系统运行状态随运行图、客流、设备工况和运维作业持续变化，正常行为边界具有动态性。同时，网络威胁可能呈现长期潜伏、分阶段渗透和跨系统传播等特征，传统依赖静态规则和事后处置的防御方式难以及时应对。如何借鉴免疫进化机制，对潜在攻击行为、攻击路径和威胁模式进行推理与预判，并结合系统运行约束生成动态防御策略，由被动响应向主动防御转变，是实现城轨网络安全智能防御需要解决的根本科学问题。

3）异构城轨工控系统网络安全设备协同管控问题。城轨系统中，安全设备和业务系统分布在不同线路、车站、车辆段和控制中心，设备类型、厂商接口、通信协议和管理策略差异明显，容易形成安全能力分散、告警割裂和处置不一致等问题。如何在“云–边–端”多层架构下融合跨域威胁情报，协调风险监测、主动防御、防火墙、密码机等多类安全设备，实现统一分析、策略生成、联动下发和执行验证，是实现城轨网络安全协同防御需要解决的关键科学问题。

通过以上剖析可见，围绕复杂城轨工控系统网络环境开展网络威胁机理与智能防御技术研究，具有重要的理论价值和工程应用意义。这不仅是破解广域威胁监测、主动防御与协同管控等关键难题的必由之路，也是系统性提升城市轨道交通网络安全防护能力的重要基础。

2.2　主要研究内容

如图1所示，围绕上述面临的三大关键科学问题，以复杂城轨工控系统网络环境为核心，结合人工免疫与大模型等智能化技术，开展面向城市轨道交通系统的网络安全智能防御关键技术研究，包括基于免疫识别的城轨工控系统网络安全风险监测技术、基于免疫进化的城轨工控系统网络威胁主动防御技术、基于大模型的城轨工控系统网络安全设备协同管控技术三大关键技术研究，借此发展和丰富城轨工控网络安全防御的新技术、新方法；在此基础之上，通过研发一套城轨工控系统网络安全智能防御原型系统并进行应用，来对成果进行技术验证，并根据结果再对技术及方法进行改进。

2.2.1　基于免疫识别的城轨工控系统网络安全风险监测技术

针对城轨工控系统中网络覆盖范围广、设备通信距离长、数据海量多源等特点，现有方法存在对多源异构数据的学习能力差、难以挖掘数据之间的内在联系和规律的问题，导致网络攻击的响应机制滞后、检测误报率高，难以满足网络环境中实时监测的需求。

为此，本研究通过模拟人体免疫系统快速响应病毒入侵的过程，构建适应城轨工控免疫监测模型，实现对网络威胁的快速发现和精准识别，具体包括：

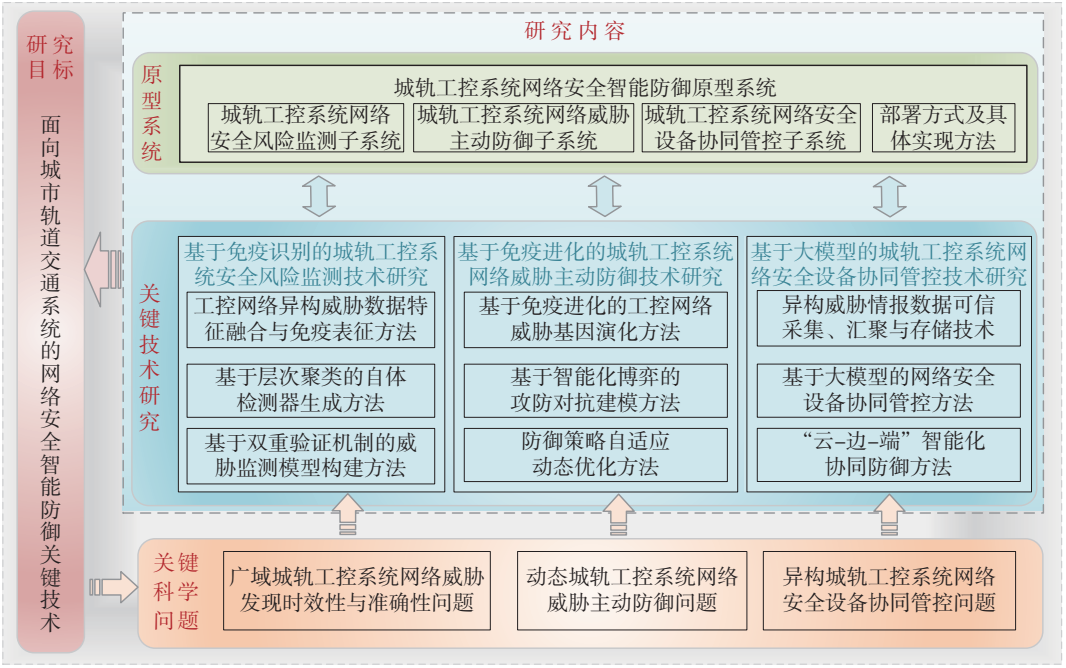


图1 主要研究内容

Fig. 1 Main research contents

1)研究工控网络异构威胁数据特征融合与免疫表征方法,构建统一的免疫特征空间,为后续模型构建提供高质量数据基础;

2)研究基于层次聚类的检测器生成方法,实现非自体检测器快速耐受训练、动态层次化识别和簇内快速应答,提升对多类型威胁的响应;

3)研究基于双重验证机制的网络威胁监测模型构建方法,利用自体检测器实现威胁的快速发现,并且,进一步通过非自体检测器实现威胁的精准识别。

2.2.2 基于免疫进化的城轨工控系统网络威胁主动防御技术

现有城轨工控系统网络安全主要依赖于被动防御和静态防御,在面对网络架构庞杂、网络环境与威胁动态多变的网络环境时,攻击者可以通过分析防御系统的漏洞或弱点,找到绕过或突破的途径,且大多数系统在攻击发生后才采取防御措施,无法实现对攻击的主动响应。为此,本研究围绕基于免疫进化的城轨工控系统网络威胁主动防御技术开展系统研究,具体包括:

1)研究基于免疫进化的工控网络威胁基因演化方法,实现对潜在攻击行为的推理与预判,为后续博弈建模过程提供威胁模式和攻击知识;

2)研究基于智能化博弈的攻防对抗建模方法,通过神经网络的推理和博弈分析,智能化生成最优防御策略,实现对网络威胁的主动防御;

3)研究防御策略自适应动态优化方法,通过结合威胁基因的动态演化方法与防御策略的动态博弈方法,不断调整和部署防御策略,从而提高城轨工控系统对网络威胁的持续主动应对能力。

2.2.3 基于大模型的城轨工控系统网络安全设备协同管控技术

传统的工控系统设备往往是单点运作,缺乏协同和全局优化的能力,在面对网络威胁数据和网络安全设备异构、分散的环境时,大多数管控策略需要人工设计与维护,难以有效协调海量异构零散的数据和设备;在面对日益复杂威胁时,往往显得力不从心,难以形成协同联动,导致防御碎片化,难以提供全局性综合防御策略。为此,本研究面向“云-边-端”多层架构开展城轨工控系统网络安全设备协同管控技术研究,具体包括:

1)研究异构威胁情报数据可信采集、汇聚与存储技术,通过在城轨工控系统“云-边-端”各层次设置不同数据聚合节点,对网络威胁情报数据进行反复融合压缩采集,构建高完备性的城轨工控网络威胁情报库;

2)研究基于大模型的网络安全设备协同管控方法,利用大模型的知识推理和生成能力实现对异构威胁情报数据的分析理解,形成跨域联动管控方案;

3)研究“云-边-端”智能化协同防御方法,在发生安全事件时基于联动管控方案形成处置策略,实现对异构网络威胁情报数据和网络安全设备的智能化协同管控。

2.2.4 城轨工控系统网络安全智能防御原型系统

基于上述理论和方法,研发城轨工控系统网络安全智能防御原型系统,包括基于免疫识别的城轨工控

系统网络安全风险监测子系统、基于免疫进化的城轨工控系统网络威胁主动防御子系统和基于大模型的城轨工控系统网络安全设备协同管控子系统；结合应用场景的工控系统生产运行要求和安全防护需求，研究上述原型系统的部署方式及具体实现方法，包括软硬件环境搭建、通信规约制定、性能提升优化、人机交互界面设计等系统集成技术；在轨道交通业务环境下开展应用，并根据应用结果对技术和方法进行改进和提高。后续将在原型系统完善后，构建面向城市轨道交通系统的网络安全仿真验证环境，并委托第三方测试机构开展系统仿真、攻防验证和性能测试，重点评估威胁检测准确率、响应时延、动态防御策略生成效果和多设备协同处置能力，以验证免疫识别、免疫进化与攻防联动机制在城轨场景下的适用性和有效性。

3　主要技术路线

3.1　基于免疫识别的城轨工控系统网络安全风险监测技术路线

传统的网络防护技术在面对网络覆盖范围广、设备通信距离长、数据海量多源的网络环境时，存在对多源异构数据学习能力差、监测模型部署成本高等问题，难以满足城轨网络威胁实时监测的需求。鉴于此，提出一种基于免疫识别的城轨工控网络安全风险监测技术，其研究技术路线如图 2 所示。

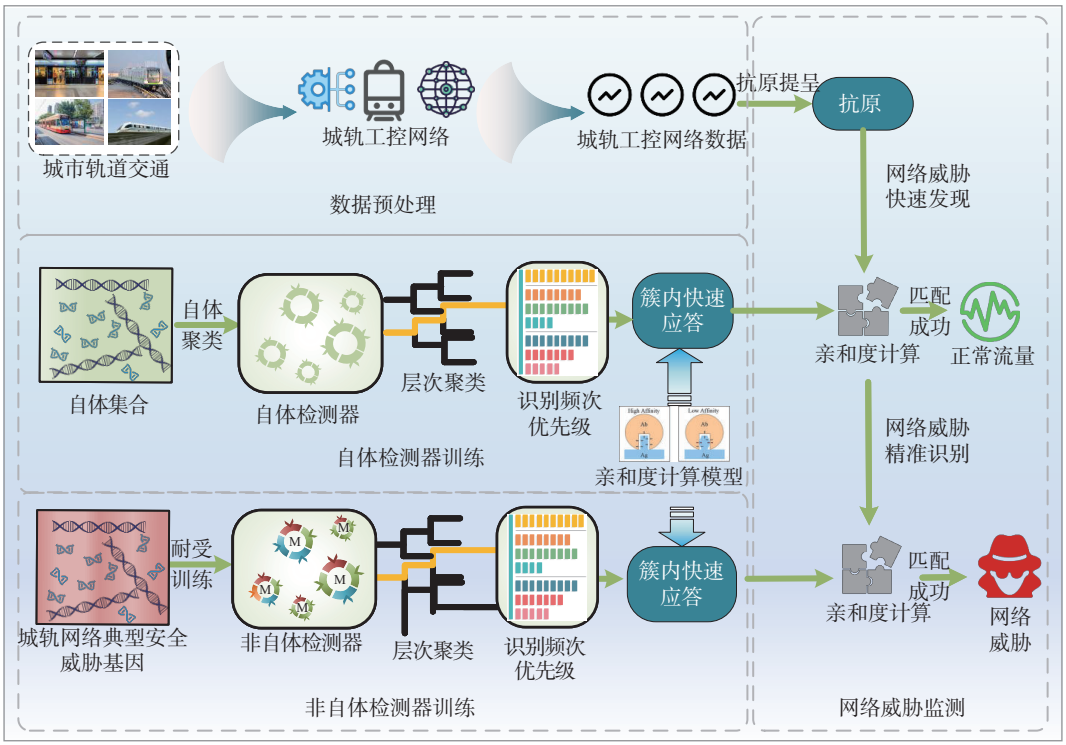


图 2　基于免疫识别的网络安全风险监测技术路线

Fig. 2　Technical route of immune-recognition-based cybersecurity risk monitoring

首先，在工控网络异构威胁数据特征融合与免疫表征阶段，对城轨工控网络不同数据源的数据进行采集和抗原提呈。其中，抗原提呈是人工免疫系统中的工程化类比，指将网络安全数据加工为可被免疫检测器识别和匹配的特征对象。然后，自体集合经过自体聚类后训练成自体检测器，典型安全威胁基因经过耐受训练后生成非自体检测器，由此为后续双重验证机制提供正常、威胁两类参照基准与判别依据。最后，在基于双重验证机制的威胁监测模型构建阶段，分两步完成威胁发现与识别。

第 1 步，将待识别抗原与自体检测器集合进行亲和度计算，区分自体（正常流量）与存活抗原（可能存在威胁），实现威胁的快速发现。

第 2 步，将第一步筛选出的存活抗原与非自体检测器进一步进行亲和度计算，精准识别城轨网络数据中存在的威胁。为避免在大规模检测器集合上逐一比对带来的计算开销，对检测器按表征相似性进行分簇。同时，聚类结构可随新抗原与检测器的更新动态调整，保障在线监测的实时性。在该过程中，采用层次聚类技术和簇内快速应答方法实现对监测过程效率的优化。层次聚类技术可以实现检测器的动态化聚类，最终结果产生的合并路径为一棵二叉树，自末端逐层判断簇间距离是否大于设定阈值，若在某一层的根结点处，簇间距离大于设定阈值，则该层已达到需要的聚类精度，则可将后续的叶子节点归到当前根节点，形成一个检测器簇集合。簇内快速应答方法使用

亲和度计算，将亲和度数值保证在0–1之间并保留计算结果的直观性，依次计算抗原和两个检测器的亲和度，并与预先设定的阈值相匹配。

3.2 基于免疫进化的城轨工控系统网络威胁主动防御技术路线

传统城轨工控系统的防御手段以静态、被动防御为主，在动态演化的网络威胁环境下，攻击者可通过挖掘防御体系存在的漏洞与薄弱点，绕过甚至突破安全防护机制，难以应对复杂多变的网络攻击，无法满足城轨工控系统高可靠、高实时性的安全保障需求。鉴于此，提出一种基于免疫进化的城轨工控系统主动防御技术，其研究技术路线如图3所示。

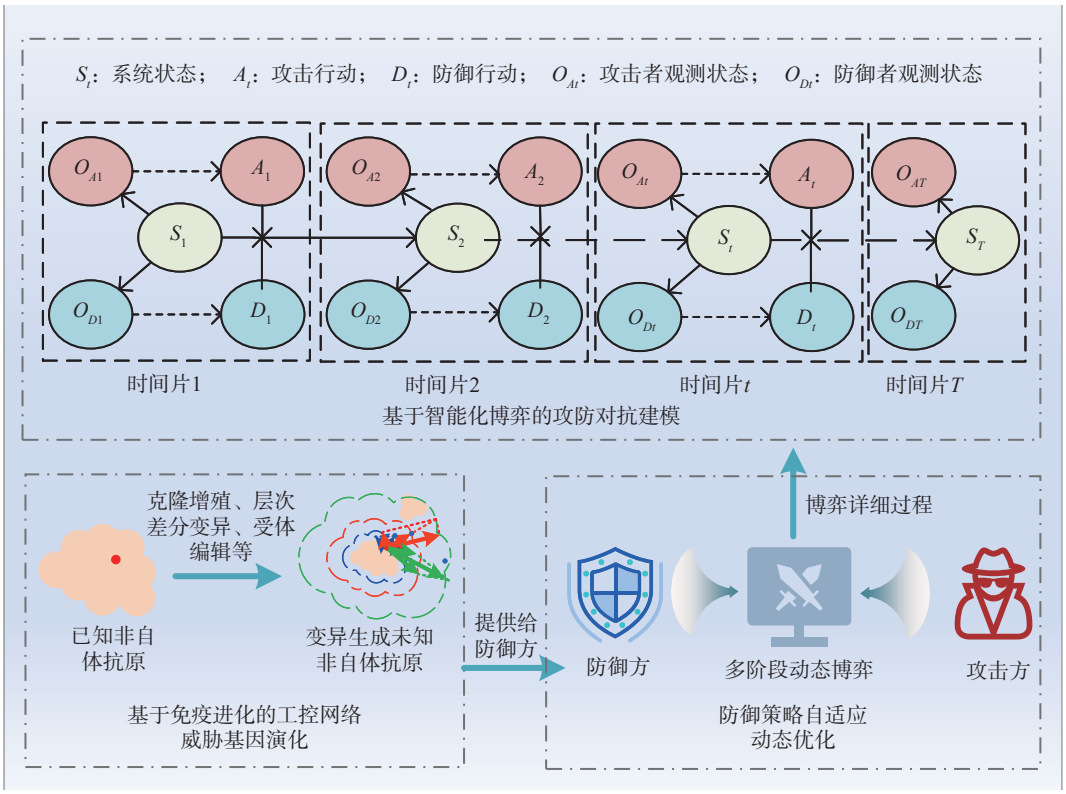


图3 基于免疫进化的网络威胁主动防御技术路线

Fig. 3 Technical route of immune-evolution-based proactive network threat defense

首先，基于免疫进化的网络威胁基因动态演化过程，对边缘自体抗原采用层次差分进化算法，通过克隆增殖、层次差分变异等操作，得到与自体区域存在偏差的未知非自体抗原。在基于智能化博弈的攻防对抗建模阶段，使用随机博弈模拟攻击者、防御者，将攻防过程离散为多个连续的状态，双方在每个状态下选择相应动作。在此基础上，分别为攻击者和防御者定义状态–动作响应函数，用来刻画在给定状态和当前动作下可以获得的综合收益。该收益由两部分组成：一部分是即时奖励，反映当前时刻攻击或防御行为在收益、代价、暴露风险等方面的直接效果；另一部分是期望未来奖励，通过状态转移概率对后续可能到达的各状态收益进行加权求和，反映当前决策对后续博弈进程的长期影响。对这两类奖励进行折扣累积，分别得到攻击方和防御方的状态–策略价值函数，用于评价在所有可能状态序列上某一策略的长期综合收益水平。此外，根据系统运行状态和参与者类型，定义攻击者和防御者的行动选择概率函数。防御者在不同系统状态和不同防御类型下，以一定概率选择不同的防御动作；攻击者则在观测到的系统状态和既定攻击类型下，以对应概率选择不同的攻击动作。该方法能够有效捕捉攻防博弈中的动态变化，考虑到不同条件下的决策风险与回报，确保参与者的行为更符合现实网络安全环境。最终，防御方在上述收益与概率框架下借助深度循环Q网络开展持续学习，根据实时数据不断更新对状态–动作价值的估计，自适应调整防御策略，使网络安全防护措施能够随攻防态势的变化而动态优化。

3.3 基于大模型的城轨工控系统网络安全设备协同管控技术路线

传统工控系统网络安全设备管控技术缺乏协同和全局优化的能力，在面对网络安全设备异构、分散的环境时，智能化程度较低，难以满足城轨工控场景下一体化、动态化的安全管控需求。对此，提出一种基于大模型的城轨工控系统网络安全设备协同管控技术，其研究技术路线如图4所示。

首先，在威胁情报数据采集阶段，模拟人体分布

式免疫细胞威胁感知架构，利用嗅探器等手段从网络流量、安全日志等数据源采集数据，并设置层次化数据融合压缩中心，对海量网络情报数据进行清洗、聚合与降维，构建城轨工控网络威胁情报库。然后，根据收集到的网络威胁情报库对选择的大模型进行数据预训练，并通过提示词交互自动化生成跨域联动管控方案，其管控的网络安全设备包括风险监测系统、主动防御系统，以及系统中的防火墙、密码机等。

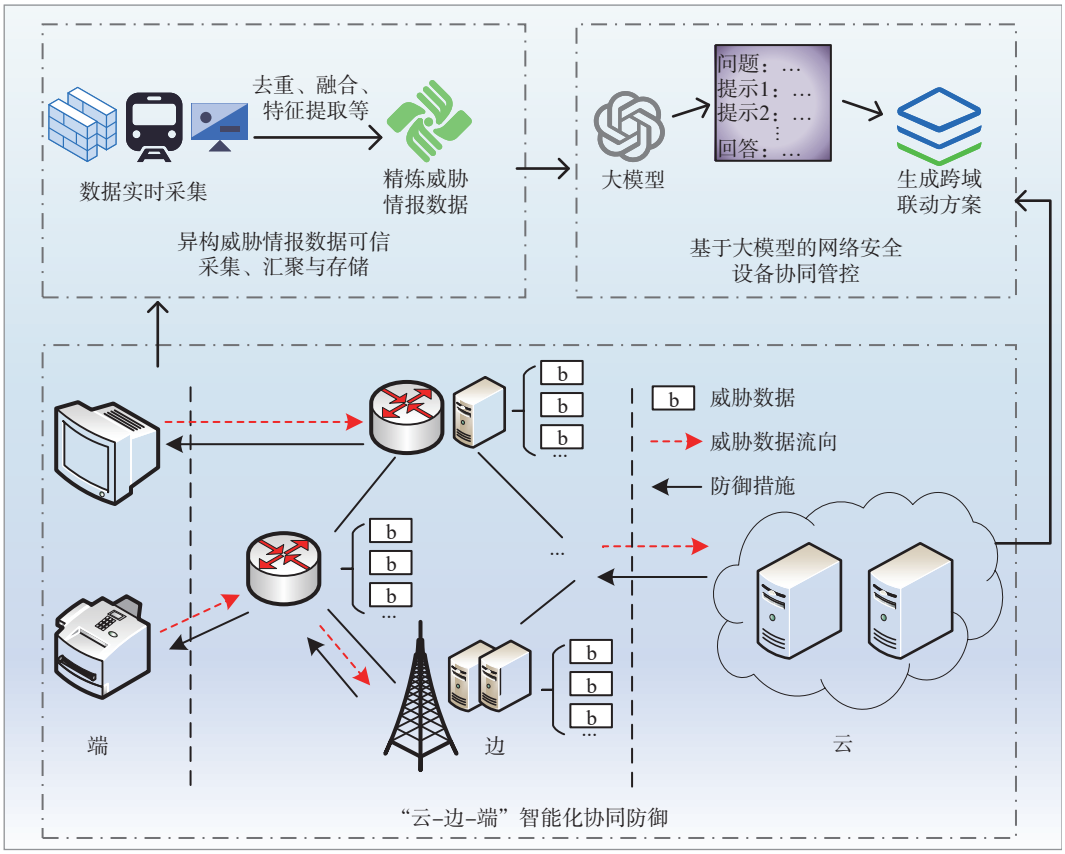


图 4　基于大模型的网络安全设备协同管控技术路线

Fig. 4　Technical route of large-language-model- based collaborative management and control of cybersecurity devices

同时，为提升方案可解释性与可追溯性，对每次生成的管控方案同步输出关联信息，并记录全流程日志以支撑审计与效果评估。根据运行反馈对大模型进行持续优化，在参数更新过程中，采用基于温度参数的词采样策略，先根据模型对各候选词给出的得分构造概率分布，再通过调整温度参数来控制该分布的尖锐或平滑程度，从而在更稳定、更确定的生成结果和更多样、更具探索性的生成结果之间进行权衡。最后，采用“云–边–端”智能化协同防御方法处理工控系统安全事件，其中：云端部分主要实现全局的安全策略制定与优化；边缘设备实现实时处理局部数据，进行预处理和初步防御；终端设备主要负责实时生成数据并执行具体的防护措施。通过云端分析威胁源头，自动生成修复措施和威胁消除措施，并通过边缘设备执行。

3.4　城轨工控系统网络安全智能防御原型系统架构

城轨工控系统网络安全智能防御原型系统架构如图 5 所示。原型系统包括 3 个子系统：基于免疫识别的城轨工控网络安全风险监测子系统、基于免疫进化的城轨工控系统主动防御子系统、基于大模型的城轨工控系统网络安全设备协同管控子系统。该原型系统具备快速发现和精准识别城轨工控网络威胁、主动生成动态防御策略、网络安全设备的集中管理和协同防御等能力。

风险监测子系统面向流量、日志、告警与协议指令等多源数据开展特征融合与免疫表征，构建并更新免疫检测器，输出异常评分，实现对隐蔽攻击的快速发现与持续跟踪。主动防御子系统在满足行车安全与业务连续性约束的前提下，基于威胁态势自适应演化检测器与防御动作库，并结合智能博弈机制开展主动推演，生成面向潜在未知威胁的防御策略。设备协同管控子系统依托大模型的语义理解与推理能力，将高层策略转化为可执行的设备配置与编排任务，完成跨站点、跨子系统的策略一致性校验、联动下发与执行验证，并提供可解释的处置依据与审计记录。通过上述模块协同，该原型系统具备网络威胁发现、主动防御、设备协同管控能力，可以维护网络正常运行秩序，保护工业控制系统网络资源安全，提升复杂城轨环境下系统的安全韧性与持续保障能力。

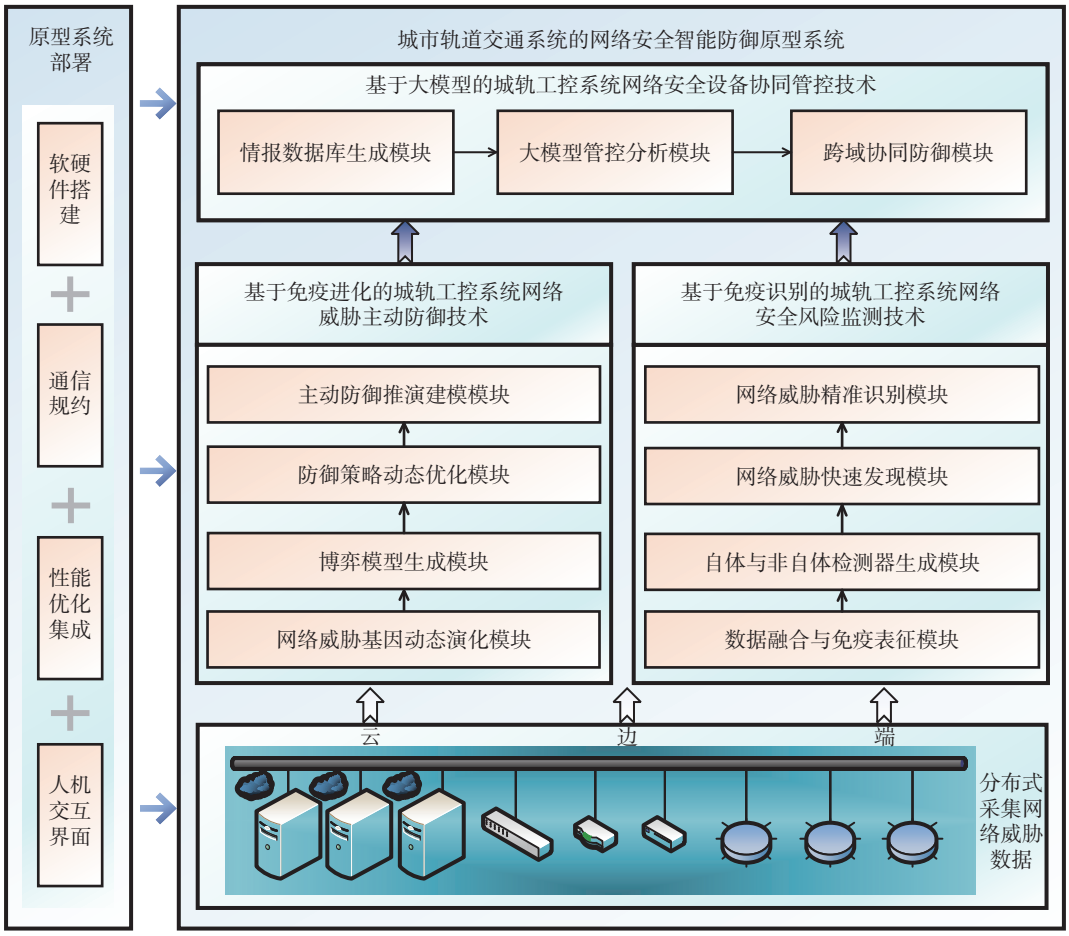


图5 原型系统架构

Fig. 5 Architecture of prototype system

4 结论与展望

本研究面向复杂城轨工控系统网络环境中的广域网络威胁发现与识别、动态主动防御与异构设备协同管控等关键科学问题，通过人工免疫与大模型等创新手段，支撑城市轨道交通网络安全智能防御技术体系构建。主要结论如下：1）提出基于免疫识别的城轨工控系统网络安全风险监测方法，提高对隐蔽威胁和跨域攻击的发现时效性与识别精度；2）构建基于免疫进化的城轨工控系统网络威胁主动防御技术框架，增强系统对持续性攻击与多阶段攻击的抵御能力；3）提出基于大模型的城轨工控系统网络安全设备协同管控方法，并研发网络安全智能防御原型系统，为城市轨道交通系统构建可持续演进的网络安全防线提供可行路径。

展望未来，随着城轨系统向“云边协同、车地一体”的互联架构持续演进，网络安全与行车安全、运营可靠性之间的联系将更加紧密。相关研究可进一步围绕网络攻击跨域传播及其对实际业务的影响机理，探索兼顾安全防护与连续运行的整体防御体系，推动网络安全技术与城轨业务需求深度融合。同时，面向典型城轨业务场景，建立具有代表性的评价体系，结合仿真验证与实际应用，明确不同运行条件下的技术适用范围，为智能防御技术的部署与推广提供依据。通过理论研究、技术创新与工程实践的相互促进，推动城轨网络安全防护由局部能力建设向系统整体安全保障发展，为其他关键信息基础设施的智能化安全防护提供借鉴。

参考文献：

[1] Internet Crime Complaint Center(IC3). Federal Bureau of Investigation Internet crime report 2024[R/OL].https://www.ic-3.gov/AnnualReport/Reports/2024_IC3Report.pdf.

[2] Peng Yihua, Liu Mingyuan, Gao Shuai, et al. Current status and future development of cyberspace security in rail transit industry[J]. Strategic Study of CAE, 2023, 25(6): 137–149.[彭轶华,刘明远,郜帅,等. 轨道交通行业网络空间安全现状与未来发展[J]. 中国工程科学,2023,25(6):137–149.]

[3] Dong Zhenyang, Wang Chenhui. Research on the intelligent dispatching optimization strategy of subway line network based on big data[J]. Intelligent Building & Smart City, 2025(7): 172–174.[董振洋,汪晨晖. 基于大数据的地

铁线网智慧调度优化策略研究[J]. 智能建筑与智慧城市,2025(7):172−174.]

[4] Lv Huihuan. Analysis and discussion of urban rail transit signal system based on cloud architecture[J]. Construction & Design for Engineering,2025(1):116−118.[吕辉环 . 基于云架构的城轨信号系统分析与探讨[J]. 工程建设与设计,2025(1):116−118.]

[5] Soderi S,Masti D,Lun Y Z. Railway cyber-security in the era of interconnected systems: A survey[J]. IEEE Transactions on Intelligent Transportation Systems,2023,24(7):6764−6779.

[6] Thaduri A,Aljumaili M,Kour R,et al.Cybersecurity for eMaintenance in railway infrastructure: Risks and consequences[J]. International Journal of System Assurance Engineering and Management,2019,10(2):149−159.

[7] The Business Research Company. Railway cybersecurity global market report 2025[EB/OL]. [2026−03−03]. https://www.thebusinessresearchcompany.com/report/railway-cybersecurity-global-market-report.

[8] 国家市场监督管理总局,国家标准化管理委员会 . 城市轨道交通安全防范系统技术要求:GB/T 26718—2024[S]. 北京:中国标准出版社,2024.

[9] 国家市场监督管理总局,国家标准化管理委员会 . 城市轨道交通网络信息系统安全基本要求:GB/T 46739—2025[S]. 北京:中国标准出版社,2025.

[10] Li Beibei, Wu Yuhao, Song Jiarui, et al. DeepFed: Federated deep learning for intrusion detection in industrial cyber-physical systems[J]. IEEE Transactions on Industrial Informatics,2021,17(8):5615−5624.

[11] Cao Zhiwei,Qin Yong,Xie Zhengyu,et al.An effective railway intrusion detection method using dynamic intrusion region and lightweight neural network[J].Measurement, 2022,191:110564.

[12] Walita T, Erba A, Castellanos J H, et al. Blind concealment from reconstruction-based attack detectors for industrial control systems *via* backdoor attacks[C]//Proceedings of the 9th ACM Cyber-Physical System Security Workshop. New York:ACM,2023:36−47.

[13] Ding Zhaohui,Zhang Wei,Yang Guoyu,et al.Research and application of multi-dimensional industrial control system network security risk monitoring and early warning system [J]. Application of Electronic Technique,2023,49(2):76−79.[丁朝晖,张伟,杨国玉,等 . 多维工控系统网络安全风险监测预警系统研究与应用[J]. 电子技术应用,2023,49(2):76−79.]

[14] Guo Weijie, Liu Lu, Du Xin, et al. Multi-objective cooperative decision-making of security and safety strategies for unattended industrial control system[J]. Control and Decision,2024,39(11):3617−3627.[郭伟杰,刘璐,杜鑫,等 . 基于多目标的无人值守工业控制系统安全策略协同决策[J]. 控制与决策,2024,39(11):3617−3627.]

[15] Zhu Li, Chen Cheng, Wang Hongwei, et al. Machine learning in urban rail transit systems: A survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(3):2182−2207.

[16] Huang Hanyuan,Li Tao,Li Beibei,et al.A bidirectional differential evolution-based unknown cyberattack detection system[J]. IEEE Transactions on Evolutionary Computation,2025,29(2):459−473.

[17] Yu Zujun, Wang Hongwei, Chen Feng. Security of railway control systems: A survey, research issues and challenges [J].High-speed Railway,2023,1(1):6−17.

[18] Yadav G,Paul K.Architecture and security of SCADA systems: A review[J]. International Journal of Critical Infrastructure Protection,2021,34:100433.

[19] Wang Yujie, Du Xin, Lu Zhihui, et al. Improved LSTM-based time-series anomaly detection in rail transit operation environments[J]. IEEE Transactions on Industrial Informatics,2022,18(12):9027−9036.

[20] Wang Junyan,Yi Ran,Zhang Bin,et al.Research on application of new intrusion detection technology in urban rail transit signaling system[J].Urban Mass Transit,2022,25(7):43−46.[王俊彦,衣然,张斌,等 . 新型入侵检测技术在城市轨道交通信号系统中的应用研究[J]. 城市轨道交通研究,2022,25(7):43−46.]

[21] Liu Dongdong,Zhou Xingyu,Duo Jianhua,et al.Cybersecurity situational awareness system applicable for urban rail transit CBTC system[J]. Urban Mass Transit,2024,27(10):318−321.[刘懂懂,周星宇,朵建华,等 . 适用于城市轨道交通 CBTC 系统的网络安全态势感知系统[J]. 城市轨道交通研究,2024,27(10):318−321.]

[22] Li Qichang,Zhao Junyi.An intrusion detection method for CBTC systems using blockchain and LSTM[C]//Proceedings of the 2023 3rd Asia−Pacific Conference on Communications Technology and Computer Science(ACCTCS). Shenyang:IEEE,2023:609−612.

[23] 刘磊,陆渝,张志伟,等 . 工业控制系统信息安全多层防御体系模型探析[C]//第五届全国信息安全等级保护技术大会,2016:146−150.

[24] Li Yidong,Zhang Li,Lv Zhuo,et al.Detecting anomalies in intelligent vehicle charging and station power supply systems with multi-head attention models[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(1):555−564.

[25] Li Shengyu, Meng Fanjun, Zhang Dashun, et al. Research on security defense system of industrial control network

[C]//Proceedings of the 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence(ICIBA).Chongqing:IEEE,2021:631−635.

[26] Zhong Kai, Yang Zhibang, Xiao Guoqing, et al. An efficient parallel reinforcement learning approach to cross-layer defense mechanism in industrial control systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2022, 33(11):2979−2990.

[27] Zhou Yuyang, Cheng Guang, Zhao Yuyu, et al. Toward proactive and efficient DDoS mitigation in IIoT systems: A moving target defense approach[J]. IEEE Transactions on Industrial Informatics,2022,18(4):2734−2744.

[28] Zhang Tao, Xu Changqiao, Zou Ping, et al. How to mitigate DDoS intelligently in SD-IoV:A moving target defense approach[J]. IEEE Transactions on Industrial Informatics, 2023,19(1):1097−1106.

[29] Chen Da, Cai Xiao, Sun Yanbin, et al. Optimization of data conf lict and DDoS attack defense mechanisms in industrial control systems based on greedy algorithm[J].Netinfo Security,2025,25(6):943−954.[陈大,蔡肖,孙彦斌,等. 基于贪心算法优化工业控制系统数据冲突与 DDoS 攻击防御机制[J]. 信息网络安全,2025,25(6):943−954.]

[30] Humayed A,Lin Jingqiang, Li Fengjun, et al. Cyber-physical systems security—A survey[J].IEEE Internet of Things Journal,2017,4(6):1802−1831.

[31] Kravchik M, Shabtai A.Efficient cyber attack detection in industrial control systems using lightweight neural networks and PCA[J].IEEE Transactions on Dependable and Secure Computing,2022,19(4):2179−2197.

[32] Li Yang, Zhu Li, Wang Hongwei, et al. A cross-layer defense scheme for edge intelligence-enabled CBTC systems against MitM attacks[J].IEEE Transactions on Intelligent Transportation Systems,2021,22(4):2286−2298.

[33] Trunzer E, Kirchen I, Folmer J, et al. A flexible architecture for data mining from heterogeneous data sources in automated production systems[C]//Proceedings of the 2017 IEEE International Conference on Industrial Technology(ICIT). Toronto:IEEE,2017:1106−1111.

[34] Rockwell Automation.The global leader in thin client management & industrial mobility solutions [EB/OL].[2026−03−03].https://thinmanager.com/.

[35] Shi Cong, Lu Xinyuan, Sun Jie, et al.Full life cycle data management system of trackside signaling equipment[J]. Urban Mass Transit,2023,26(6):239−243.[施聪,陆鑫源,孙杰,等. 轨旁信号设备全生命周期数据管理体系[J]. 城市轨道交通研究,2023,26(6):239−243.]

[36] Huang Susu, Li Gang, Li Ke, et al.Regional rail transit electronic map data management scheme for intercity railway through-operation[J].Urban Mass Transit,2024,27(1):210−214.[黄苏苏,李刚,李克,等. 面向城际铁路直通运行的区域轨道交通电子地图数据管理方案[J]. 城市轨道交通研究,2024,27(1):210−214.]

[37] Cylus. The rail cybersecurity landscape in 2025: Service continuity and recoverability[R/OL]. https://www.cylus.com/resources/the-rail-cybersecurity-landscape-2025.

[38] Zhang Peiying, Wang Chao, Jiang Chunxiao, et al. Deep reinforcement learning assisted federated learning algorithm for data management of IIoT[J]. IEEE Transactions on Industrial Informatics,2021,17(12):8475−8484.

[39] Lu Junqing, Shen Jian, Vijayakumar P, et al.Blockchain-based secure data storage protocol for sensors in the industrial Internet of Things[J].IEEE Transactions on Industrial Informatics,2022,18(8):5422−5431.

[40] Koo J, Kang G, Kim Y G.Access control framework for cross-platform interoperability in the industrial Internet of Things[J]. IEEE Transactions on Industrial Informatics, 2025,21(1):801−810.

[41] Mushtaq S, Mohsin M, Mushtaq M M. A systematic literature review on the implementation and challenges of zero trust architecture across domains[J].Sensors,2025,25(19):6118.

[42] Xu Hanxiang, Wang Shenao, Li Ningke, et al.Large language models for cyber security: A systematic literature review[J/OL].ACM Transactions on Software Engineering and Methodology,2025−09−29.https://doi.org/10.1145/3769676.

[43] Manowska A, Syta J. Application of large language models in the protection of industrial IoT systems for critical infrastructure[J].Applied Sciences,2026,16(2):730.

[44] 上海市市场监督管理局. 城市轨道交通网络安全通用技术规范:DB31/T 1499—2024[S]. 上海:上海市市场监督管理局,2024.

[45] 工业互联网产业联盟. 城市轨道交通工业互联网信息管控系统应用指南:T11/AII 022—2021[S/OL].https://www.aii-alliance. org/uploads/1/20220324/649c458bf0358dcae34e327d81c02ed5.pdf.

[46] 天津飞腾信息技术有限公司. 威努特与飞腾轨道交通信号系统网络安全综合解决方案[EB/OL].[2026−03−03]. https://www.phytium.com.cn/solution/detail/?category=technology&id=387.

[47] 北京启明星辰信息安全技术有限公司. 某市地铁综合监控系统安全防护项目[EB/OL].(2020−11−02)[2026−03−03]. https://www.venustech.com.cn/new_type/gyalfx/20201102/22018.html.

[48] Gao Bing, Bu Bing, Zhang Wei, et al. An intrusion detection method based on machine learning and state observer for train-ground communication systems[J].IEEE Transactions on Intelligent Transportation Systems,2022,23(7):

6608−6620.

[49] Tang Tao,Liu Wentao,Ding Shukui,et al.Urban rail transit FAO system:Technological development and trends[J].Autonomous Intelligent Systems,2022,2(1):25.

[50] Li Beibei,Chang Yujie,Huang Hanyuan,et al.Artificial immunity based distributed and fast anomaly detection for Industrial Internet of Things[J]. Future Generation Computer Systems,2023,148:367−379.

Intelligent Cyber Defense Technology for Urban Rail Transit Systems: Conceptual Foundations and Future Prospects

LI Beibei[1], *WANG Yaoxuan*[1], *DU Qingyun*[1], *FENG Jiacheng*[1], *HUANG Xiaobo*[2]

(1. School of Cyber Science and Engineering, Sichuan University, Chengdu 610065, China;
2. Ningbo Hollysys Information Security Research Company Limited, Ningbo 315301, China)

Abstract:

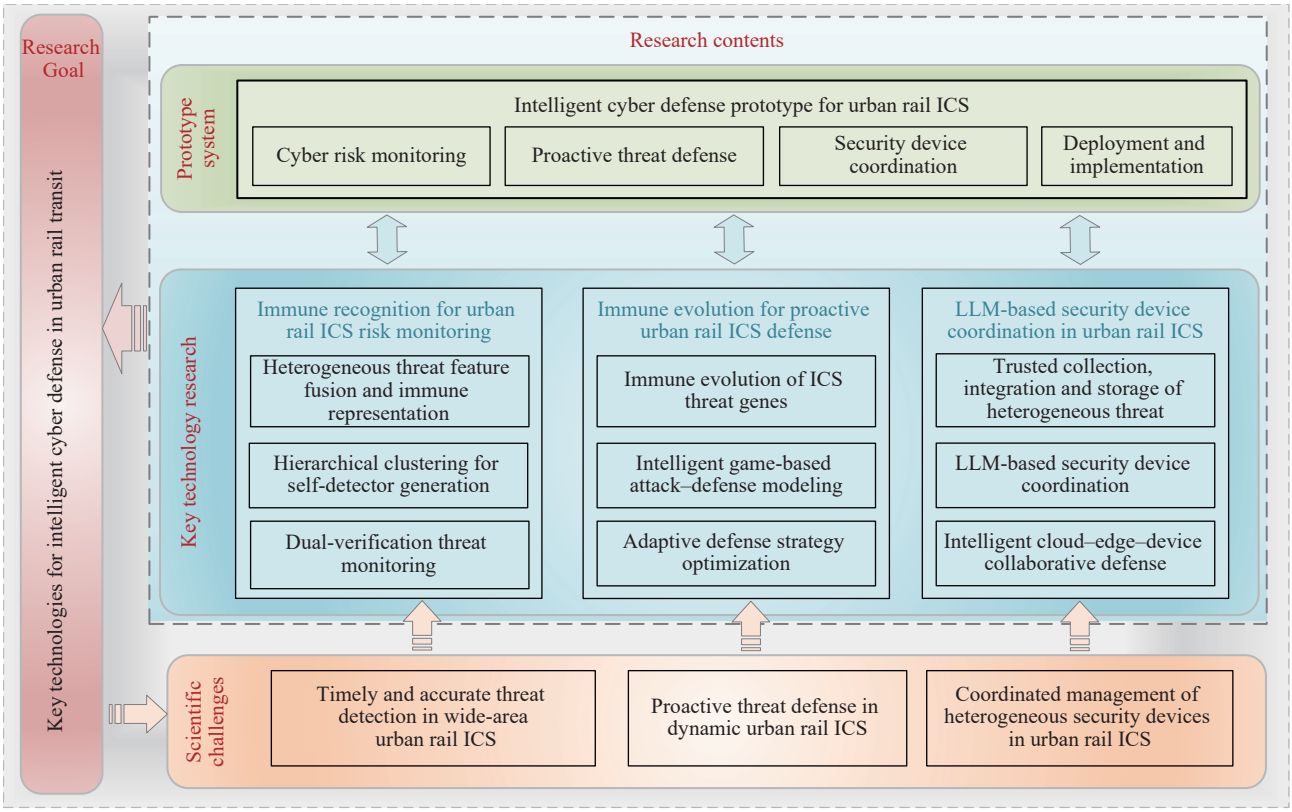
Significance In the recent past, the rapid advancement of industrial digitalization and intelligentization has extended the boundaries of cyberspace into the physical world, exposing urban rail transit systems, as critical national information infrastructure, to increasingly severe cybersecurity threats. The network environment of urban rail industrial control systems is characterized by wide-area deployment, dynamic evolution, and strong heterogeneity. Extensive network coverage, dynamically changing operating conditions, and diverse equipment types and communication protocols impose higher requirements on the timeliness, accuracy, adaptability, and coordination of cybersecurity defenses. Conventional cybersecurity defenses often lack adequate situational awareness and remain reactive in these environments; hence, detection and identification of threats across wide-area networks in a timely and accurate manner is challenging. Static and passive protection mechanisms cannot readily adapt to changing network conditions and threat patterns, whereas the unified management and coordination of heterogeneous cybersecurity devices remain inadequate. Consequently, insufficient accuracy and timeliness in threat monitoring, the limitations of static and passive defense mechanisms, and weak capabilities for unified management and coordination remain major technical challenges. These challenges, together with the interconnection of control centers, stations, train networks, and cloud platforms, increase the risk of threats propagating across different operational domains. Meanwhile, stringent requirements for train operational safety, real-time control, and service continuity must be maintained during security monitoring and defensive operations. Cybersecurity defenses must therefore balance timely threat detection with accurate identification while adapting defensive policies to evolving network conditions and threat patterns. In this context, research on artificial immunity and large language models is necessary to support the recognition of previously unknown threats, the dynamic optimization of defensive policies, and the coordinated management of heterogeneous cybersecurity devices.

Progress Artificial immunity and large language models were integrated into a research framework for complex urban rail industrial control networks to address these challenges. Three principal research problems were identified: timely and accurate wide-area threat detection, proactive defense against dynamically evolving threats, and collaborative orchestration of heterogeneous cybersecurity devices. The proposed framework comprised three core techniques and an integrated prototype system. Specifically, an artificial-immune-recognition-based risk monitoring method was developed to enable rapid threat detection and accurate identification. Multisource heterogeneous data, including traffic, logs, alerts, protocol commands, and equipment states, were used for feature fusion and representation in a unified immune feature space. Hierarchical clustering and detector tolerance training were incorporated into the detector generation scheme, whereas rapid responses within detector clusters were designed to improve monitoring efficiency. A dual-verification mechanism was also introduced, whereby preliminary screening using self detectors was followed by further identification using non-self detectors. An artificial-immune-evolution-based proactive defense framework was subsequently developed to enhance protection against persistent and multistage attacks. Clonal proliferation and hierarchical differential evolution were incorporated into threat gene evolution to facilitate the exploration of potential threat patterns. Stochastic attack-defense games were combined with deep learning in the proposed technical approach to defense policy generation and adaptive dynamic optimization. Successive states, available actions, immediate rewards and costs, and expected future returns were considered in modeling attack-defense interactions. A large-language-model-based collaborative management approach was also developed to coordinate diverse security devices for joint defense. Heterogeneous threat intelligence was organized within a cloud-edge-endpoint architecture to support its collection, aggregation, compression, and storage. The knowledge reasoning and generation capabilities of large language models were leveraged for intelligence analysis and the generation of cross-domain coordination plans. The devices intended for management included risk monitoring systems, proactive defense systems, firewalls, and cryptographic appliances. Policy consistency checks, coordinated distribution, execution verification, and audit records were incorporated into the system design. Building on these developments, a prototype intelligent cybersecurity defense architecture for urban rail transit systems was presented, with system integration, deployment arrangements, and technical validation outlined.

Conclusions and Prospects An integrated research framework for cybersecurity defense in urban rail transit systems was proposed, connecting risk monitoring, proactive defense, and coordinated device management through three complementary technical directions. Artificial immune recognition was introduced to enable timely threat detection and accurate identification, immune evolution was incorporated into adaptive defense policy generation, and large language models were used for heterogeneous threat intelligence analysis and collaborative device management. This framework outlines an approach beyond traditional patch-based passive defense to support the secure and reliable operation of national critical information infrastructure. Future studies should focus on adapting these techniques to actual urban rail operating conditions. Dynamic self-boundary modeling and unknown-threat recognition should be refined to account for changes in operating schedules, equipment conditions, and maintenance activities. Defense policy generation and optimization should be investigated under the constraints of train operational safety, real-time control, and service continuity. The explainability and traceability of generated management plans, cross-device policy consistency, and coordinated execution should also be enhanced. Prototype integration should be advanced, followed by system simulation, attack-defense evaluation, and third-party performance testing. Threat detection accuracy, response latency, dynamic defense policy generation, and coordinated handling across multiple devices should be systematically evaluated. Feedback from practical applications should then be used to refine the proposed technologies and methods.

Key words: urban rail transit system; industrial control system; cyber threat; active defense; artificial immunity; large language model

Graphical abstract:



（编辑　赵　婧）

引用格式: Li Beibei,Wang Yaoxuan,Du Qingyun, et al. Intelligent cyber defense technology for urban rail transit systems: Conceptual foundations and future prospects[J].Advanced Engineering Sciences,2026,58(6):1–14.[李贝贝,王耀萱,杜卿芸,等.面向城市轨道交通系统的网络安全智能防御关键技术研究构想与成果展望[J].工程科学与技术,2026,58(6):1–14.]