Computer Science and Application 计算机科学与应用, 2024, 14(11), 60-69
Published Online November 2024 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2024.1411216

Hans汉斯

面向“实战化”的动车组网络安全主动防御体系研究与实践

熊凯舟[1,2]，周泽岩[3]，郑雪洋[2]，赵　薇[3]

[1]中国铁道科学研究院研究生部，北京
[2]中国铁道科学研究院集团有限公司机车车辆研究所，北京
[3]中国铁道科学研究院集团有限公司电子计算技术研究所，北京

收稿日期：2024年10月12日；录用日期：2024年11月11日；发布日期：2024年11月20日

摘　要

随着人工智能技术、互联互通等网络化新技术在动车组中的广泛应用，动车组网络控制系统内外部接口大量增加，随之带来内部与外界对系统网络环境的恶意攻击威胁日益增加。作为动车组列车的“大脑”，网络控制系统的信息安全尤为重要。本文针对网络控制系统信息网络的安全隐患及防护现状，提出了集安全保障体系、安全技术体系、安全标准体系的多个维度构成的动车组网络控制系统安全主动防御体系，对防御体系的关键技术展开了深入研究，提出了基于智能对抗的漏洞利用流量数据扩展技术、基于TRDP协议的智能分析与恶意数据识别技术等多个关键技术实现方法，为有效保障动车组列车运行安全提供了体系和技术支撑。

关键词

动车组，网络控制系统，主动防御，智能分析

Research and Practice of EMU Network Security Active Defense System for “Actual Combat”

Kaizhou Xiong[1,2], Zeyan Zhou[3], Xueyang Zheng[2], Wei Zhao[3]

[1]Department of Postgraduate, China Academy of Railway Sciences, Beijing
[2]Locomotive & Car Research Institute, China Academy of Railway Sciences Co. Ltd., Beijing
[3]Institute of Computing Technologies, China Academy of Railway Sciences Co. Ltd., Beijing

Received: Oct. 12th, 2024; accepted: Nov. 11th, 2024; published: Nov. 20th, 2024

文章引用: 熊凯舟, 周泽岩, 郑雪洋, 赵薇. 面向“实战化”的动车组网络安全主动防御体系研究与实践[J]. 计算机科学与应用, 2024, 14(11): 60-69. DOI: 10.12677/csa.2024.1411216

Abstract

With the extensive application of new networking technologies such as artificial intelligence technology and interworking in EMU, the internal and external interfaces of EMU network control system have increased greatly, and the threat of internal and external malicious attacks on the system network environment has increased day by day. As the "brain" of EMU trains, the information security of network control system is particularly important. Aiming at the security risks and protection status of the network control system information network, this paper proposes an EMU network control system security active defense system composed of multiple dimensions including security guarantee system, security technology system and security standard system, and conducts in-depth research on the key technologies of the defense system. Several key technology implementation methods are proposed, such as vulnerability exploitation traffic data extension technology based on intelligent countermeasure, intelligent analysis based on TRDP protocol and malicious data identification technology, which provide system and technical support for effectively guaranteeing the safety of EMU trains.

Keywords

EMU, Network Control System, Active Defense, Intelligent Analysis

1. 引言

随着现代铁路运输技术的飞速发展，动车组成为高效、便捷的出行方式。网络控制系统作为动车组的核心控制单元，其安全性直接关系到列车运行的安全与稳定。尤其在人工智能和物联网技术广泛应用于动车组的背景下，网络控制系统面临的安全挑战更加严峻[1]。本文研究旨在构建一个有效的动车组网络控制系统安全防御体系，通过主被动防御技术相结合，确保网络控制系统的安全性和稳定性。

2. 动车组网络安全主动防御体系架构与设计

2.1. 安全防御体系架构

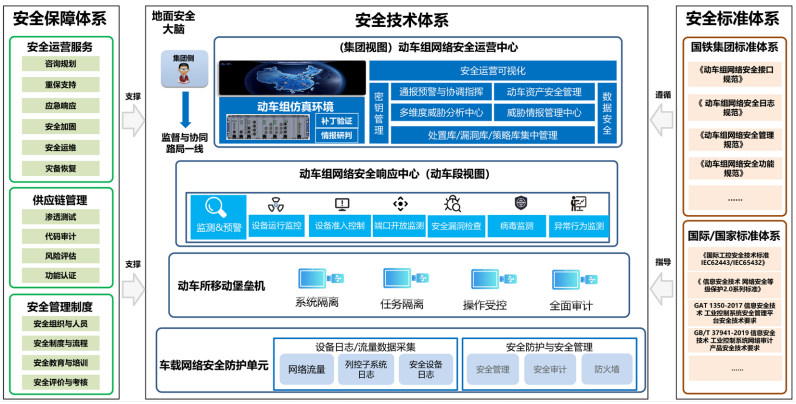

Figure 1. Defense architecture diagram
图 1. 防御体系架构图

动车组网络控制系统安全防御体系分为 3 大架构：安全保障体系、安全技术体系、安全标准体系，各体系之间相对独立又相互协同，提供良好的稳定性和扩展性(如图 1 所示)。

2.1.1. 安全保障体系

主要包括安全运营服务、供应链管理和安全管理制度三个部分，安全运营服务支持咨询规划、重保支持、应急响应、安全加固、安全运维、容灾恢复等方面的服务保障；供应链管理支持渗透测试、代码审计、风险评估、功能认证等方面的管理保障[2]；安全管理制度支持安全组织与人员、安全制度与流程、安全教育与培训、安全评价与考核等方面的管理制度服务，为安全技术体系提供保障支撑。

2.1.2. 安全技术体系

主要包括动车组网络安全运营中心、动车组网络安全响应中心、车载网络信息安全平台三个部分，动车组网络安全运营中心支持动车组车载系统网络安全仿真、动车组安全运营态势展示、通报预警与协调指挥、动车资产安全管理、多维度威胁分析、联动威胁情报动态响应等安全能力；动车组网络安全响应中心支持设备运行监控、设备准入控制、安全漏洞检查、病毒监测、异常行为监测等安全能力；车载网络信息安全平台支持网络安全集中管理、边界防护、入侵检测与安全审计等安全能力，实现对车载网络的分区分域、物理隔离、可控传输和全面审计，提升铁路核心业务数字化和智能化水平[3]。

2.1.3. 安全技术体系

通过《动车组网络安全接口规范》《动车组网络安全日志规范》《动车组网络安全管理规范》《动车组网络安全功能规范》等标准体系支撑安全体系建设；通过《国际工控安全技术标准 IEC62443/IEC65432》《信息安全技术网络安全等级保护 2.0 系列标准》《信息安全技术网络安全等级保护 2.0 系列标准》《GB/T37933-2019 信息安全技术工业控制系统专用防火墙技术要求》《GAT 1350-2017 信息安全技术工业控制系统安全管理平台安全技术要求》《GB/T 37941-2019 信息安全技术工业控制系统网络审计产品安全技术要求》等标准指导安全体系的建设。

2.2. 部署模式

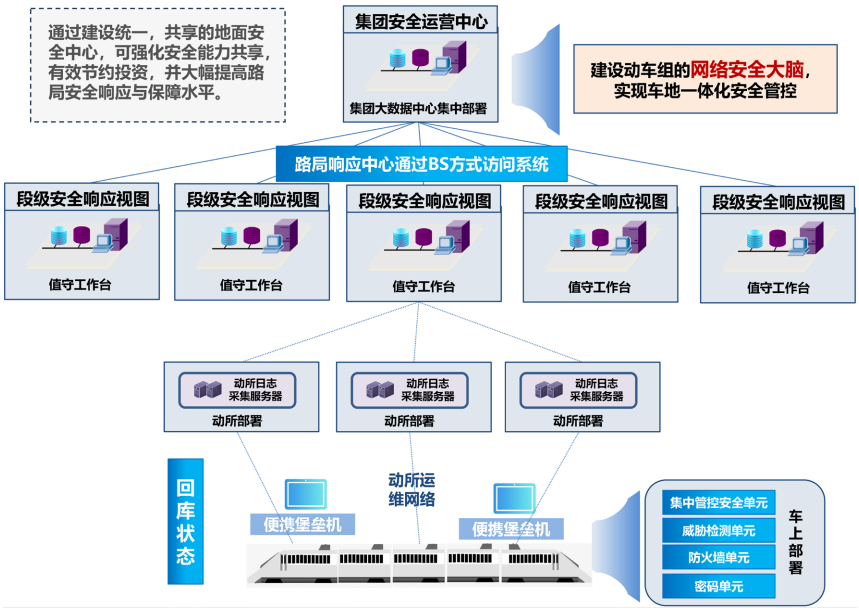

Figure 2. Deployment pattern diagram
图 2. 部署模式图

动车组安全保障和安全技术系统部署模式如图 2 所示。

1) 在集团安全生产网区域内部署地面安全运营中心平台，实现对路局动车组的统一安全数据的监测与分析，为各路局提供统一的安全管理保障。

2) 在路局安全生产网区域内部署安全响应中心，实现对动车组网络安全实时运行状态的监测，并结合驻场安全运维保障团队，实现车上高危风险的实时处置与日常响应。

3) 在动车所部署安全日志采集服务器及便携堡垒机，实现运维人员操作升级。

4) 在动车组机房部署车载网络安全平台设备，如防火墙，安全审计、安全管理等车载安全装置，实现对动车组网络边界防护、入侵防御、行为审计、威胁预警等能力，实时回传动车组网络安全状态到地面平台进行监控分析，保障动车组稳定运行。

3. 基于“实战化”网络安全主动防御体系的关键技术

3.1. 基于智能对抗的漏洞利用流量数据扩展技术

3.1.1. 融合多种算法的类别均衡技术

在处理不平衡数据集时，地面态势系统考虑采用类别均衡算法进行处理，分阶段利用 SMOTE (Synthetic Minority Over sampling Technique)算法和 Tomek Link 算法进行综合采样。第一阶段，利用 SMOTE 算法增加数据集中少数类样本的数量，解决数据集不平衡的问题；第二阶段，利用 Tomek Link 算法来处理边缘模糊化的数据，解决数据重叠所导致样本分类不准确的问题。

1) 过采样

过采样中的 SMOTE 算法是基于随机过采样方法的一种改进方案，由于随机过采样采取简单复制样本的策略来增加少数类样本，这样容易产生模型过拟合的问题，即使得模型学习到的信息过于特别而不够泛化，SMOTE 算法的基本思想是对少数类样本进行分析，并根据少数类样本人工合成新样本添加到数据集中[4]，算法流程如下：

对于少数类中的每一个样本 x，以欧式距离为标准计算它到少数类样本集中所有样本的问题，得到 k 近邻。

根据样本不均衡比例，设置一个采样倍率 $\sigma \in [0,1]$。对每一个少数类样本 x，从其 k 近邻中随机选择个样本，假设选择的近邻为 x_i。

对于每一个随机选出的近邻 x_i，分别与原少数样本构建新样本 $x_{new}^i = x + \sigma^*(x_i - x)$，得到 N 个新样本其中必须有 $k > N$，但 k 到底取多大，还需要根据数据集反复测试确定。

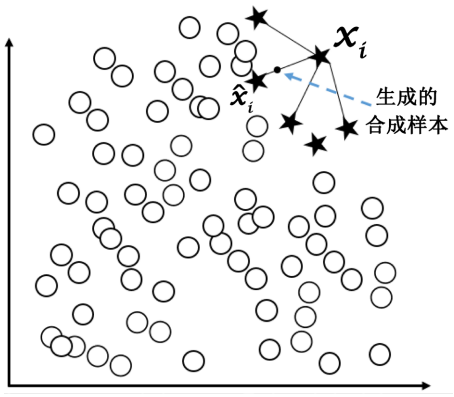

Figure 3. Presentation of the synthetic samples generated by the SMOTE algorithm
图 3. SMOTE 合成样本的展示

因此，SMOTE 算法的思想是合成新的少数类样本，合成的策略是对每个少数类样本 x，从它的最近邻中随机选一个样本 $\hat{x}$，然后在 x、$\hat{x}$ 之间的连线上随机选一点作为新合成的少数类样本，合成过程如图 3 所示。

2) 欠采样

欠采样通过删掉一部分多类样本的数量来实现样本均衡，最直接的方法是随机去掉一些多数类样本来减小多数类的规模。这样做的缺点也很明显，有可能删得太过份，把多数类别样本中一些重要的信息也会删除掉。为了避免这种情况，可以采用一些成熟的欠采样算法，如 Tomek Link 算法或 ENN (Edited Nearest Neighbours)算法，欠采样的大致效果如图 4 所示。

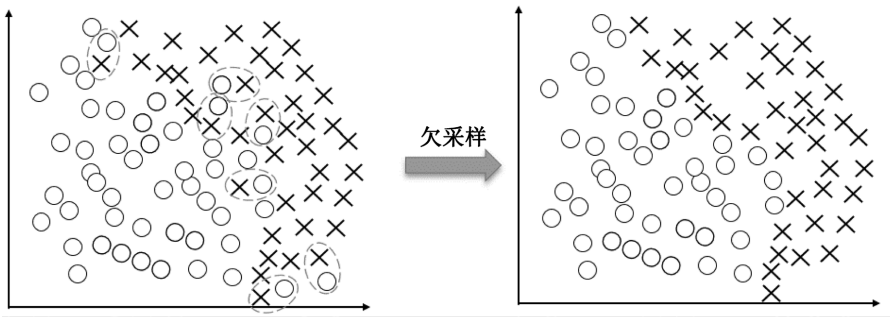

Figure 4. Undersampling effect
图 4. 欠采样效果

Tomek Link 表示不同类别之间距离最近的一对样本，即这两个样本互为最近邻且分属不同类别。假设 x_i、x_j 属于两个不同的类别，如过不存在第三个样本 x_l 使得 $d(x_i,x_l)<d(x_i,x_j)$ 或者 $d(x_j,x_l)<d(x_i,x_j)$，那么这两个样本就称为了一个 Tomek Link，当两个样本为 Tomek Link 的时候，则要么其中一个是噪音，要么两个样本都在边界附近。这样通过移除 Tomek Link 就能“清洗掉”类间重叠样本，使得互为最近邻的样本皆属于同一类别，从而能更好地进行分类[5]。

针对多数类的一个样本，可采用 ENN 方式处理，如果其 K 个近邻点有超过一半不属于多数类，则这个样本会被剔除。

3.1.2. 单侧标签平滑的特征拓展技术

对于损失函数，希望尽可能的使用预测概率分布去拟合真实概率分布，而拟合 one-hot 的真实概率函数无法保证模型的泛化能力，容易造成过拟合，同时也会造成模型过于相信预测的类别。

标签平滑是一种损失函数的修正，已被证明是非常有效的训练深度学习网络的方法。在几乎所有的情况下，使用标签平滑训练可以产生更好的校准网络，从而更好地泛化，最终对不可见的生产数据产生更准确的预测。因此，标签平滑能够在训练时立即假设标签可能存在错误，避免“过分”相信训练样本的标签。

one-hot 编码将各个分类类别进行扩展，用二进制向量表示分类变量。各个分类值进行 one-hot 编码后的特征，其实每一维度的特征都可以看作是连续的特征。因此，one-hot 编码一定程度上提供了训练数据中类别之间的关系[6]。

一般设定标签平滑参数 ϵ，对于二分类，将 one-hot 中的 1 和 0 替换为 $1-\epsilon$ 和 ϵ，对于多分类，只需要把 one-hot 中所有的 1 替换为 $1-\epsilon$，把所有的 0 替换为 $\epsilon/(K-1)$ 即可，其中 K 是类别的数量，因此，采用标签平滑后的真实概率分布变化为：

$$p_i = \begin{cases} 1-\epsilon, if(i=y) \\ \frac{\epsilon}{K-1}, if(i\neq y) \end{cases} \tag{1}$$

标签平滑化具有很多好处，特别对于 GAN (Generative Adversarial Nets)而言，能够让判别函数不会给出太大的梯度信号，也能防止算法走向极端样本的陷阱。在 GAN 中，可以采用单侧标签平滑，即设定标签平滑参数 ϵ，只将 one-hot 中的 1 替换为 $1-\epsilon$，而 0 不变。这种做法可以避免判别器的极端预测行为，如果判别器通过学习来预测一个极端大的逻辑值，也就是对某些输入的输出概率接近于 1 时，它将被惩罚并被鼓励回到一个较小的逻辑值上去。

不对 one-hot 中的 0 标签进行平滑处理也是很重要的。假设以 α 代替正分类目标，以 β 代替负分类目标，那么最优的判别模型函数将变为：

$$D^*(x) = \frac{\alpha P_{data}(x) + \beta P_{model}(x)}{P_{data}(x) + P_{model}(x)} \tag{2}$$

在该式子中，当 β 为零时，那么通过 α 的平滑仅仅是按比例缩小判别模型的最优值。而当 β 非零时，最优判别模型的函数形状会发生变化。特别是在 P_{data} 近似为零且 P_{model} 很大的区域中，来自 P_{model} 的错误样本几乎没有动力靠近数据。因此，地面态势系统只将正标签平滑为 α，而将负标签依旧设为 0。

3.1.3. 试验对比和结果分析

为了证明本文提出类别均衡算法支持漏洞利用流量数据检测的有效性与可行性，选择 SMOTE 算法和 Tomek Link 算法进行综合采样对比。实验结果如表 1 所示。第一阶段，利用 SMOTE 算法增加数据集中少数类样本的数量，解决数据集不平衡的问题；利用 Tomek Link 算法来处理边缘模糊化的数据，解决数据重叠所导致样本分类不准确的问题。第二阶段利用标签平滑算法实现损失函数的修正，有效的训练深度学习网络，更好的校准网络，从而更好地泛化，最终对不可见的生产数据产生更准确的预测，提升检测率。综上所述，基于支持类别均衡算法和标签平滑算法的漏洞利用检测方法在车载网络系统的通信流量数据异常检测中具有一定的优势。

Table 1. Comprehensive sampling comparison between SMOTE algorithm and Tomek Link algorithm
表 1. SMOTE 算法和 Tomek Link 算法综合采样对比

序号	测试内容	优化前	优化后
1	BPS CVE 代码注入/执行攻击	208 条检出 123 条检出率 59.13%	208 条检出 188 条检出率 90.38%
2	某护网入侵攻击库	1097 条检出 779 条检出率 71.01%	1097 条检出 1015 条检出率 92.52%

3.2. 基于 TRDP 协议的智能分析与恶意数据识别技术

为了有效应对不符合简单逻辑关系与规律的网络安全威胁行为，有必要构建一种高度专业化的统计分析模型，该模型专注于对网络中 TRDP (Train Real-time Data Protocol)协议数据的深入与智能化挖掘分析，分析过程依托机器学习技术，利用其内置的高级机器学习模型来实现对数据的深度分析。系统采用了隐马尔可夫模型(Hidden Markov Model, HMM)作为主流智能分析工具，用以增强对复杂网络行为的识别与预测能力。

隐马尔科夫模型(HMM)是一种统计分析模型，它的状态不能直接观察到，但能通过观测向量序列分析到所处状态，每个观测都是通过某种概率密度分布表现为各种状态，每一个观测向量是由一个具有相应概率密度分布的状态序列产生，如图 5 所示。隐马尔科夫模型是一个双重随机过程，具有一定状态数的隐马尔科夫链和显示随机函数集。通过对数据的所有访问记录分析，不难发现：普通用户的正常请求

虽然不一定完全相同，但总是彼此相似；攻击者的异常请求总是彼此各有不同，同时又明显不同于正常请求。基于这个前提，通过动车组列车收集到的大量正常数据，构建起一个能表达所有正常值的正常模型，那么一切不满足于该正常模型的参数值，即为异常。

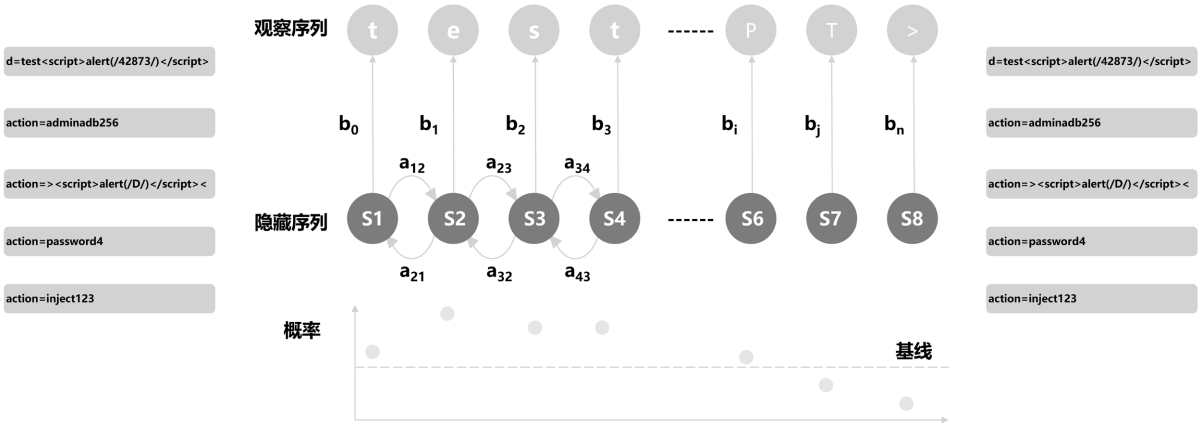

Figure 5. Sequence of observation vectors
图 5. 观测向量序列图

3.3. 基于行为轮廓与人工智能算法的潜藏攻击发现技术

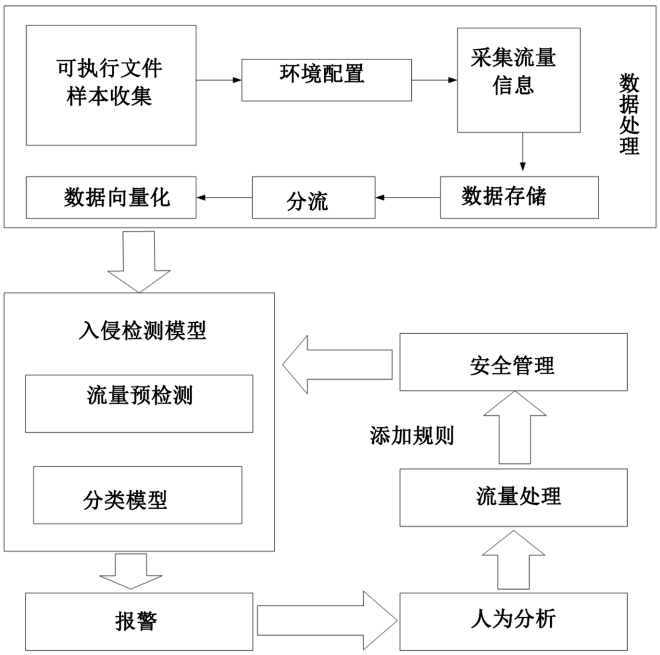

Figure 6. Technical flow of hidden attack discovery
图 6. 潜藏攻击发现技术流程

潜藏攻击发现技术是通过统计方法构建统计结果和异常行为之间的映射关系从而达到检测的目的，在检测异常行为方面，根据在行为特征上的区别来识别是否异常，并且可以通过与人工智能技术相结合，建立卷积神经网络模型应用于动车组列车网络入侵行为检测；模型主要利用卷积核对数据进行卷积来提取局部相关性的特征，从而提高特征提取的准确性，特征提取首先统计整体数据集中的全部行为类型种

类，并对其进行编码处理。其次，利用机器学习算法，设定时间滑动窗口大小并在整个日志信息流中不断的滑动来截取单位时间内的日志信息，对单个时间滑动窗口中包含的日志信息进行提取特征，建立分类模型，从而更好地检测潜藏攻击。基于行为轮廓与人工智能算法的潜藏攻击发现技术实现流程如图 6 所示。

3.4. 基于多维度关联的人机协同威胁研判分析技术

面对多源、海量、异构的安全事件相关数据，单一来源数据的检测方法不能反映安全事件的攻击关系，且往往会有误报和漏报现象，未能充分挖掘安全事件价值链的潜在价值。面对这一挑战，多维度关联分析技术成为强化攻击预警、网络防御及追踪溯源能力的重要手段。该技术将威胁情报、资产信息、漏洞详情、异常行为模式及网络流量数据等多元信息融合，通过大数据技术支持下的多维度关联分析，辨识出攻击活动、攻击意图、攻击趋势，结合人机协同威胁研判，更加深度有效的挖掘安全事件攻击行为的真实性与有效性，解决安全事件价值链无法有效进行人机协同深度研判的问题，最终达到充分挖掘安全事件价值链及减少误报和漏报的目的。多维度关联的人机协同威胁研判分析技术结构如图 7 所示：

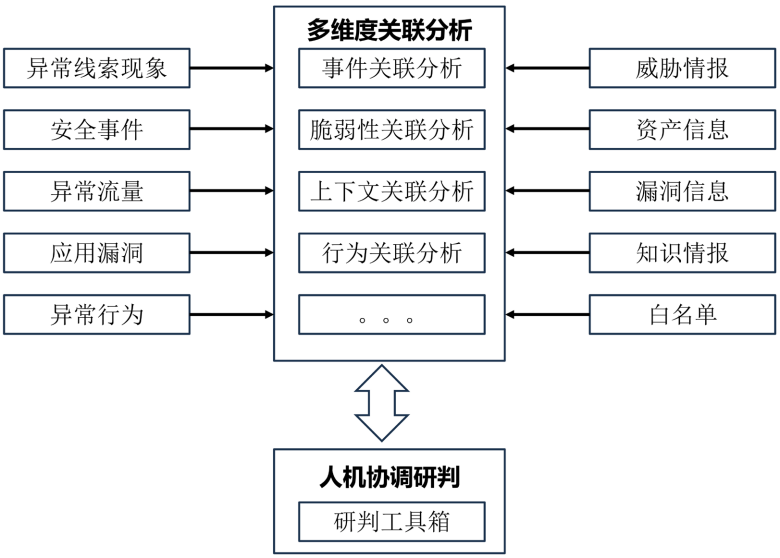

Figure 7. Multi-dimensional relational human-machine cooperative threat analysis structure
图 7. 多维度关联的人机协同威胁研判分析结构

3.5. 基于主动防御体系融合情报驱动与安全分析技术

构建一个融合情报驱动、行为分析的主动防御体系，是确保动车组网络安全的关键，该体系侧重于深度整合多源安全情报，依托集团侧安全运营中心平台，实现对全网态势的实时监控与智能分析，特别是在动车组内部网络区域部署网络信息安全平台，以及时响应威胁情报并执行防御动作。此外，结合动车组网络的分布式特性，推广实施轻量级零信任策略，强化设备监控与身份管理。利用统一身份认证平台，从身份、环境及权限三个维度出发，严格把控访问控制，确保每一次访问均基于验证后的最小权限原则，实现对动车组网络内部“云－管－边－端”的全方位保护，如图 8 所示。这一策略不仅减轻了身份滥用与非授权访问的风险，还通过持续的信任评估机制，动态适应动车组网络环境的变化。

通过云查、多维分析、攻击链分析等分析手段，输出源 IP 大网行为、源 IP 攻击特征、攻击 IP 同源分析等分析结果，输出安全分析的概况及具体安全事件信息，支撑起系统的告警和事件分析[7]。

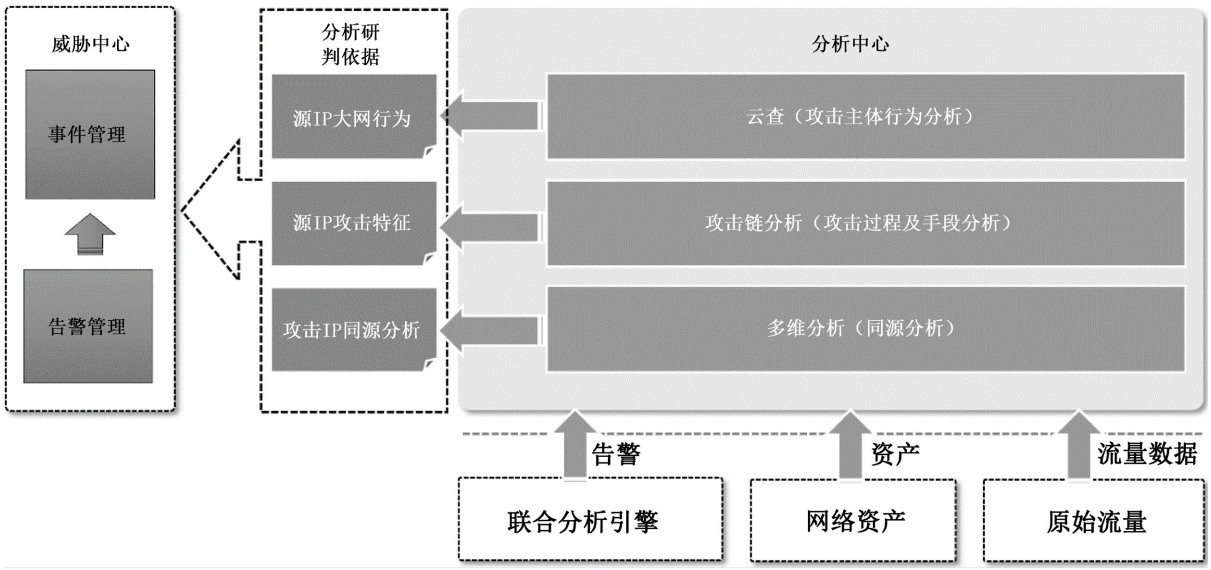

Figure 8. Merging intelligence drive and security analysis
图 8. 融合情报驱动与安全分析

3.6. 基于大数据分析技术的智能问答与响应处置技术

随着网络攻击手法不断演变，这给动车组网络安全防御工作带来了极大的挑战，为了增强对网络对抗演习及重大安全保障任务的安全事件响应能力和质量，可以采取一系列集成化措施。首先，利用自然语言处理大模型技术来构建智能问答系统，该系统能够准确理解用户提出的问题，并提供精准的答案，从而提高问题解决的效率。其次，借助先进的计算能力和推理技术，研发用于网络安全威胁分析与调查的智能分析工具，帮助安全团队更快地识别潜在风险。此外，建立一个全路安全指挥中心，确保人员、系统和流程之间的统一协调与调度，这有助于在紧急情况下迅速调动资源，实施有效的应对措施。最后，通过“大小模型自主协同”技术体系的支持，实现从告警分析、漏洞检测到影响评估以及协同处理等一系列智能化运营功能。

4. 结语

随着数字化、智能化不断推进，中国动车组迎来新一轮技术革命，网络安全工作也成为重要组成部分。为加强动车组列车网络安全建设，需要开展动车组列车实战化、常态化、系统化的主动防御工作。通过构建动车组网络安全主动防御体系，建设动车组网络安全运营中心、动车组网络安全响应中心、车载网络信息安全平台，采用核心关键技术，实现全面、准确、实时捕捉攻击行为、实时预警、有效联动、精准处置，为全面提升动车组网络控制系统的安全防护能力提供了有益的借鉴和参考。

基金项目

本研究获得国家重点研发计划(2022YFB4301101)，中国国家铁路集团有限公司科技研究开发计划K2023J013 (JB)资助。

参考文献

[1] 杨凯. 复兴号 CR200J 型动车组运行安全监测联网应用技术研究[J]. 铁道机车车辆, 2021, 41(3): 15-21.

[2] 李琨, 丁庆行. 信创网络安全测评[J]. 电子技术与软件工程, 2020(18): 246-248.

[3] 刘世文, 马多耀, 雷程, 等. 基于网络安全态势感知的主动防御技术研究[J]. 计算机工程与科学, 2018, 40(6): 1054-1061.

[4] 陈扬, 刘勤明, 梁耀旭. 小样本不平衡设备数据下的机器学习策略研究[J]. 上海理工大学学报, 2022, 44(4): 407-416.

[5] 吴克奇, 崔梦天, 等. 面向软件缺陷数据的协同过滤抽样推荐算法[J]. 西南师范大学学报, 2021, 46(11): 46-55.

[6] 龙腾刚. 基于机器学习的 ACARS 报文解析技术研究[D]: [硕士学位论文]. 成都: 西华大学, 2021.

[7] 刘冬兰, 刘新, 张昊, 等. 基于大数据的网络安全态势感知及主动防御技术研究与应用[J]. 计算机测试与控制, 2019, 27(10): 229-233.