 DOI: https://doi.org/10.12349/iser.v6i10.7907

Research on internal threat detection and response mechanism based on behavior analysis

Hongyu Zhang Qi Sun

Beijing Institute of Space Mechanics & Electricity , Beijing, 100086, China

Abstract

Internal threats, characterized by legitimate identities, covert operations, and high destructive potential, have emerged as a critical risk in organizational information security. These threats may originate from malicious attacks, accidental operations, or unauthorized access, making them difficult to detect through traditional rule-based detection and perimeter defense systems. With advancements in big data and user behavior analytics, behavior-based dynamic detection has become a key approach for internal threat prevention. This study explores a multi-source behavior data-driven monitoring framework, examining the root causes of internal threats, behavioral pattern recognition, user baseline modeling, and anomaly detection mechanisms. It proposes a three-phase response system: pre-incident alerts, real-time coordination, and post-incident investigation. The research demonstrates that integrating model analysis with strategic governance can establish a closed-loop mechanism encompassing prevention, identification, response, and tracking, thereby enhancing the intelligence, precision, and proactivity of organizational security management.

Keywords

internal threat; behavior analysis; anomaly detection; access control; security response mechanism

基于行为分析的内部威胁检测与响应机制研究

张宏宇 孙麒

北京空间机电研究所，中国 · 北京 100086

摘 要

内部威胁因具备合法身份、行为隐蔽且危害性强，已成为组织信息安全防护中的突出风险来源。其既可能源于恶意破坏，也可能来自误操作和权限滥用，传统依赖规则匹配与边界防护的检测模式难以有效识别。随着大数据与用户行为分析技术的发展，基于行为特征的动态检测逐渐成为内部威胁防控的重要方向。本文从内部威胁成因、行为模式识别、用户基线构建与异常检测模型入手，探讨多源行为数据驱动的监测路径，并设计事前预警、事中联动、事后追溯的响应体系。研究认为，模型分析与策略治理相结合，可形成覆盖预防、识别、处置与追踪的闭环机制，有助于提高组织安全管理的智能化、精细化和主动化水平。

关键词

内部威胁；行为分析；异常检测；访问控制；安全响应机制

1 引言

随着组织信息化程度不断加深，内部人员对于系统资源和数据具有天然可达性，这使得内部威胁行为更具有隐蔽性和难以识别性。相比外部攻击，内部威胁更容易突破安全边界防护体系，其行为往往伪装为正常操作，难以通过传统规则匹配或特征检测方法进行判定。近年来，企业数据泄露、业务破坏与敏感信息转移等案例频繁发生，内部人员参与或因权限配置不当而引发的风险呈现增长趋势。单纯依赖身份认证和访问控制已无法满足安全需求，系统需要能够识别“看似合法、实则异常”的操作者行为。行为分析技术的引入使得安全防护能够从静态规则向动态行为模型转变，基于用户行为基线和异常偏移的检测思路，有助于构建更加自适应的内部威胁检测体系。本文围绕内部威胁识别与响应展开系统研究，旨在为组织安全建设提供参考路径。

【作者简介】张宏宇（1988–），男，满族，中国北京人，本科，工程师，从事计算机科学研究。

2 内部威胁的内涵与形成机制

2.1 内部威胁的定义与特征

内部威胁是指具有合法身份和正常访问权限的组织内部主体，在主观故意或非主观意图条件下，对系统安全、数据资产或业务稳定造成潜在或实际损害的行为。与外部攻击相比，内部威胁的危险性更高，其隐蔽性强，往往以“合法操作”作为掩护，行为特征难以通过传统安全防护边界识别。内部人员对系统结构、业务逻辑和关键资源分布具有天然认知优势，因此其异常行为更容易在初期被视为正常工作流程

而难以触发告警。此外，内部威胁表现形式多样，既可能表现为数据泄露、恶意破坏，也可能表现为意外误操作与安全管理疏漏，均可能导致系统性损害。

2.2 内部威胁的类型划分

内部威胁可分为恶意型和非恶意型两类。恶意型内部威胁通常来源于个人私利、组织矛盾、职业报复或外部策反，如主动窃取敏感数据或破坏系统功能；其行为隐蔽且意图明确。非恶意型内部威胁则多由安全意识薄弱、缺乏风险认知、操作失误或权限配置不当引起，例如误删数据、错误共享机密文件等。尽管非恶意行为不具备破坏意图，但由于常发生在高权限操作环节，其造成的损失往往不亚于恶意行为。两类内部威胁均对组织的数据安全和业务连续性构成重大挑战。

2.3 内部威胁产生的组织与技术原因

内部威胁的形成具有多因素驱动特征，既包含制度性问题，也涉及技术层面的漏洞。在制度层面，一些组织存在岗位权限交叉、审批流程缺乏制衡、监督机制薄弱等问题，使内部人员能够在系统中获取超出实际工作需要的广泛权限，形成潜在风险。在文化层面，如果缺乏安全意识教育和责任认同机制，员工对信息安全的重要性和敏感性认识不足，容易出现误操作和违规行为。在技术层面，日志记录不完善、审计链不连续、权限控制策略粗放等问题，使内部异常行为难以被及时捕捉和追踪。因此，内部威胁治理应从组织制度、人员行为和技术体系三个层面协同推进，构建全方位防控体系。

3 基于行为分析的内部威胁检测模型构建

3.1 用户行为基线模型的建立

用户行为基线模型旨在通过长期、连续的行为数据采集来刻画个体在正常工作状态下的行为模式，从而为异常识别提供参考基准。在模型构建过程中，需要对用户在不同业务系统中的访问频率、登录时间段、常用操作路径、数据读取与修改习惯、文件交互方式以及常用终端与网络环境等多维特征进行统计与建模。通过对行为的时间序列特征、操作事件间依赖关系及资源调用逻辑进行分析，可形成具有稳定性和差异性的个体行为画像。基线模型强调“因人建模”而非统一模板，能够有效适应岗位职责、工作方式和技能水平差异。同时，可以将群体行为均值、岗位行为模板与个体行为基线进行分层融合，提高模型在多场景下的适用性。当用户实时行为与基线模型出现偏离时，即可作为风险判断的重要依据，为进一步异常检测提供前置判断条件。

3.2 多源行为数据融合机制

内部威胁识别需要突破单一日志数据维度，将不同系统产生的操作痕迹进行统一采集与关联分析，构建完整、可还原的用户行为链路。常见行为数据包括服务器系统日志、数据库审计日志、网络流量与会话记录、终端指令集、文件操作信息、外设使用轨迹以及身份认证与权限变更日志。通过引入数据清洗、时间戳对齐、行为事件统一建模等技术，可以将异构数据源标准化处理，实现跨系统行为关联。数据融合的关键在于建立“人—资源—行为—场景”的多维关联关系，使系统能够识别出看似独立但在逻辑上具有联动性的行为事件。例如，若某用户在非工作时段远程登录并下载大量敏感文件，再结合其近期岗位调整信息，系统可提高预警等级。通过多源融合，可显著减少审计盲区，提高异常检测精度。

3.3 异常行为判定方法

异常行为判定的核心在于识别用户实时行为与基线模型之间的偏离程度，并综合行为情境要素进行风险判定。在具体方法上，可基于阈值偏离模型，对访问次数、操作速率或数据调用规模的异常增量进行快速识别；也可采用聚类与密度分析方法，将行为映射到高维空间识别“离群点”；对于操作序列较为复杂的情形，可引入深度学习序列模型（如 LSTM、GRU）来捕捉行为模式变化趋势，从而识别潜在威胁。同时，需要引入情境特征进行综合判断，包括访问地点可信度、终端环境是否异常、操作目标是否超出岗位权限范围、行为发生时间是否偏离常态等。情境增强能够降低误报，提高判断的准确性与解释性。通过算法与场景结合，可实现内部威胁的实时、精确识别。

4 内部威胁监测系统设计与实现路径

4.1 体系架构设计思路

内部威胁监测系统的体系架构应在整体上实现分层设计、模块解耦与功能协同，以保证系统的灵活扩展与稳定运行。系统可划分为数据采集层、行为分析层、模型判定层与响应联动层。数据采集层负责从操作系统日志、应用系统日志、网络流量监控、终端指令记录、文件操作记录、认证授权系统等多源通道中进行统一数据收集，并对数据进行格式化、时间同步与匿名化预处理，保证数据质量与可用性。行为分析层对采集的数据进行特征提取和行为序列建模，构建用户行为基线画像与资源访问模式图谱，为后续模型识别提供输入。模型判定层使用阈值检测、聚类分析、序列学习模型或情境增强模型对实时行为进行异常判定，识别潜在内部威胁风险。

4.2 行为分析模型的动态更新机制

行为分析模型并非一成不变，而是需要随组织业务与人员行为变化持续演进。如果模型长期不更新，将出现行为刻画失真、阈值偏移或风险识别精度下降等问题，导致误报或漏报。为保证模型有效性，应建立动态更新机制：一方面，通过周期性重新训练，将最新采集到的行为数据纳入模型学习样本中，更新行为基线特征分布，使模型能够反映当前组织的真实行为状态；另一方面，通过增量学习或在线学习方式，使模型能够在检测运行中持续吸收新行为模式，尤其适

用于人员岗位变动频繁、业务流程更新较快的组织场景。此外，模型更新需与权限调整、岗位职责变更、系统应用升级等管理信息联动，避免模型与业务脱节。

4.3 模型可解释性与人工复核机制

内部威胁检测系统不仅应具备识别能力，还应具备结果可解释性，以支持安全管理员对异常事件进行快速判断与决策。可解释性可通过规则链回溯、特征贡献度分析与行为路径可视化等方式实现，使异常判定能够明确说明“何处偏离、偏离程度、风险原因”。同时，在涉及人员行为判断的场景中，过度依赖自动化模型可能导致误报或误判，因此需要引入人工复核机制。人工复核可由安全管理员、业务专家与权限管理人员共同完成，对高风险事件进行综合评估，避免因模型偏差导致不必要的业务中断或人员误伤。

5 内部威胁响应机制的构建策略

5.1 事前预警与访问最小权限策略

事前预警的核心在于通过制度性约束与动态风险识别降低内部威胁发生的可能性。在权限管理方面，应坚持“最小权限”原则，将访问权限与岗位职责精准绑定，避免“一人多权”“权限超配”等现象。同时，应建立权限生命周期管理机制，实现权限申请、审批、分配、使用与回收的全流程可控，尤其对离职、转岗、长期未使用及外包账户实施定期清理与动态收缩。在行为准入层面，可以通过设定敏感资源访问白名单、操作指令合法性校验与异常登录特征识别等手段，在行为发生前对其合法性进行判定。针对关键岗位与高风险系统，可引入情境风险评估机制，将时间、地点、设备可信度、网络环境等参数纳入访问许可判断逻辑，提升安全决策的精细化水平。

5.2 事中联动响应机制

在内部威胁检测中，事中响应的及时性直接决定风险扩散程度。一旦系统识别到行为偏离用户基线模型或与已知攻击模式存在高度相似性，应启动自动化安全响应机制。联动响应可根据威胁等级进行分级触发，例如：对轻微异常行为执行提示与二次确认；对中度风险执行多因素身份验证、访问限速或切换只读模式；对重大风险行为如大批量数据外传、越权访问核心系统等，则迅速执行会话隔离、账号冻结或自动切断网络连接，阻断攻击链条的进一步发展。同时，各系统之间应建立事件联动接口，实现终端检测系统（EDR）、身份管理平台（IAM）、网络流量分析系统（NTA）与日志审计系统（SIEM）的信息共享，使响应不再是单点动作，而是跨系统协同防护。此外，事中响应还需建立人工介入机制，由安全管理员对高风险事件进行溯源分析和策略再调整，确保响应措施既迅速有效，又不过度干预正常业务运行。

5.3 事后取证与审计追溯体系

内部威胁事件的事后处理不仅是对损失的补救，更是对系统安全能力的再构建。因此，事后取证与审计追溯必须在系统运行中具备可执行性与可证明性。首先，应建立统一日志审计平台，将操作日志、访问日志、系统日志、终端行为日志和网络流量信息统一归档与时间同步，确保全链路行为可还原。其次，利用行为链回溯模型可实现从异常行为触发点向前追踪，重建完整攻击路径、资源访问路径与主客体关系图谱，为事件定性和责任判定提供依据。在取证过程中，可采用哈希校验、数字签名等方式保证日志与证据材料的完整性和可靠性。事件处置完成后，应将本次事件的行为特征、系统反应效果和潜在制度漏洞纳入威胁场景知识库，用于模型再训练和策略优化，从而推动系统形成自我演化能力。通过持续复盘与反向优化，组织可以在不断经验积累中形成长期稳定的安全韧性，使内部威胁防控从被动防御迈向主动治理。

6 结语

内部威胁检测与响应是组织安全体系的重要组成部分，其核心在于由静态防护向动态安全转型。基于行为分析构建内部威胁检测体系，可以有效提升威胁识别的准确性和及时性。未来，应在算法模型自适应性、隐私保护机制与组织协同治理方面进一步加强，以构建“可预警、可防御、可响应、可追溯”的全链路内部安全治理体系。

参考文献

[1] 孙小双,王宇.基于多源数据的内部威胁检测技术综述[J].计算机应用与软件,2024,41(09):1-8+40.

[2] 侯瑞.基于多源信息特征分析的内部威胁检测研究[D].北京邮电大学,2024.

[3] 刘明瑾.基于行为日志的内部威胁检测方法研究[D].哈尔滨工程大学,2024.

[4] 葛鼎威.面向内部威胁检测的多域信息融合深度模型研究[D].福州大学,2023.

[5] 顾兆军,郭靖轩.基于角色异常行为挖掘的内部威胁检测方法[J].计算机工程与设计,2020,41(10):2740-2746.