DOI:10.20033/j.1003-7241.(2026)01-0106-05

智能云运维系统混合蜜罐入侵自动检测方法

张一平[1,2], 徐 雯[1,2], 杨小艺[2], 饶 煜[1,2]

(1.中石油煤层气有限责任公司,北京 100028; 2.中联煤层气国家工程研究中心有限责任公司,北京 100095)

摘 要:为了降低系统入侵虚警率,本文提出利用混合蜜罐入侵的自动检测方法。利用操作系统、网络服务和应用程序等作为蜜罐设备,并通过计算系统各节点的破坏因子进行优化蜜罐部署,结合入侵行为特征的约束条件函数,求取入侵特征惯性权值,进而优化入侵行为特征,基于此,采用神经网络构造入侵检测模型,并将当前入侵行为与特征库中的非正常行为进行比对,根据匹配阈值,判断待测数据是否为入侵数据。实验结果表明,该方法的虚警率始终保持在0.3%以下,能够有效降低系统的入侵检测虚警率。

关键词:智能云运维系统;混合蜜罐;神经网络;入侵行为;特征聚类

中图分类号: TP277;TN915.08 **文献标志码**: A **文章编号**: 1003-7241(2026)01-0106-05

Intelligent cloud operation and maintenance system hybrid honeypot intrusion automation detection method

ZHANG Yiping[1,2], XU Wen[1,2], YANG Xiaoyi[2], RAO Yu[1,2]

(1. PetroChina Coalbed Methane Company Ltd., Beijing 100028, China;
2. China Union Coalbed Methane National Engineer Research Center Company Ltd., Beijing 100095, China)

Abstract: To reduce the false alarm rate of system intrusion, this paper proposes an automatic detection method utilizing a hybrid honeypot approach. By employing operating systems, network services, and applications as honeypot devices, and optimizing the deployment of honeypots through calculating the damage factors of each system node, combined with the constraint condition functions of intrusion behavior characteristics, the inertia weight of intrusion features is determined. This, in turn, optimizes the intrusion behavior characteristics. Based on this, a neural network is used to construct an intrusion detection model, which compares current intrusion activities with abnormal behaviors in the feature database. According to the matching threshold, it judges whether the data to be tested is intrusion data. Experimental results show that the false alarm rate of the method proposed in this paper consistently remains below 0.3%, indicating that the method can effectively reduce the false alarm rate of system intrusion detection.

Keywords: intelligent cloud operation and maintenance system; mixed honey pot; neural networks; invasion behavior; feature clustering

智能云运维系统作为云计算技术的重要组成部分,承担着保障企业数据安全和系统稳定运行的重要职责[1]。传统的入侵检测方法,如基于签名的检测和基于行为的检测,虽然能够在一定程度上识别已知的攻击模式,但往往不适合新的和变体的攻击方法。特别是在云环境中,由于资源的动态分配和服务的快速部署,传统的检测方法在实时性、准确性和可扩展性方面面临着重大挑战。因此,如何有效地检测和防御针对云运维系统的入侵行为成为业界关注的焦点。

文献[2]基于双域配对策略实时监控物联网所有活动,并触发入侵检测机制保护系统安全。该方法具有适应性较广的特点,不仅适用于传统的物联网环境,还可以适应各种复杂多变的场景,该策略的实施取决于环境中设备和系统是否正常运行和通信可能导致在某些情况下,系统可能会将正常活动误判为入侵行为,从而导致高误报率。文献[3]提出一种控制区域网络图注意网络(control area network graph attention network,CAN-GAT)模型,利用图卷积、图注意、CAN-GAT网络和图神经网络(graph neural network,GNN)构建异常检测框架,具有良好的检测速度性能。但是该模型框架需要大量的标记数据进行训练,这在实际应用中可能是不可行的。文献[4]设计应用深度自编码网络的局域网空间入侵监测系统。先采用库函数以及FHW抓包方法抓取数据包并提取数据特征;结合支持向量回归预测算法构建入侵监测模型,完成入侵数据监测,但是训练数据不够多样化或代表性不足,模型的泛化能力会受到限制,无法有效识别新型或变种攻击。文献

收稿日期:2024-06-20;**录用日期**:2024-07-01

基金项目:国家自然科学基金(61202494)

作者简介:张一平(1986-),男,硕士,中级工程师,研究方向:人工智能技术在网络安全中的研究与应用。

引用本文:张一平, 徐雯, 杨小艺, 等. 智能云运维系统混合蜜罐入侵自动检测方法[J]. 自动化技术与应用, 2026,45(1):106-109,113. (ZHANG Yiping, XU Wen, YANG Xiaoyi, et al. Intelligent cloud operation and maintenance system hybrid honeypot intrusion automation detection method[J]. Techniques of Automation and Applications, 2026,45(1):106-109,113.)

[5]基于预处理的数据,选择与入侵分类结果高度相关的特征,使用随机森林训练多个基分类器,从而检测到入侵行为。该方法可以处理复杂的网络环境和多种攻击方法,鲁棒性较强。但是该方法参数调整困难,参数调整不当可能导致模型性能下降。文献[6]通过对收集和存储光纤通信网络的流量数据,进行小波阈值去噪处理。应用粗糙集理论和决策树算法构建入侵信号提取与检测模型,根据信号相位差得到异常入侵定位检测结果。该方法结合了数据挖掘和机器学习等领域的现有知识,提高了检测效率。然而,此方法无法及时识别新的和未知的攻击行为,导致检测精度偏低。综上所述,尽管已经存在许多系统入侵检测方法,但仍存在一些挑战,如误报率高、模型泛化能力不足、参数调整困难等。因此,继续研究和改进智能云运维系统入侵检测技术仍然具有重要的意义。

针对以上分析,为提高智能云运维系统的入侵检测准确率,本文旨在探索混合蜜罐入侵自动检测方法在智能云运维系统中的实现和应用。通过对相关技术和算法的深入研究,为云运维系统的安全保护提供了新的思路和方法。

1 智能云入侵自动检测方法

1.1 混合蜜罐环境部署

混合蜜罐环境通过模拟真实的服务和应用程序,诱使攻击者进行攻击,从而收集和分析攻击者的行为、工具和方法,可以为智能云运维系统的入侵检测提供丰富的样本和情报[7],帮助检测系统更加准确地识别出潜在的威胁和攻击行为,进而制定相应的防御策略和措施。其核心价值在于监测、分析系统的攻击活动,为后续入侵检测提供信息[8]。

在智能云运维系统和混合蜜罐服务中,假设系统运行期间所有的访问请求均为合法用户,攻击的均为黑客,且蜜罐的目标是检测黑客的攻击,则当第 i 个黑客没有被攻击到,其攻击收益可以表示为

$$b_i = \sum_{i=1}^{n} \frac{\alpha_0 \times g_s}{\sqrt{\left\| \frac{t \times j_0}{x_u} \right\|^2}} \tag{1}$$

式中,n 表示攻击类型数量;α_0 表示离散因子,用于描述入侵数据的离散程度,0 表示初始状态标识;g_s 表示攻击代价,s 表示攻击类型序号;t 表示时间常数;j_0 表示所有访问中攻击的概率,x_u 表示第 u 次访问代表的调整因子。

利用操作系统、网络服务和应用程序等具有迷惑性的蜜罐环境作为目标设备,并通过星形连接方式将其与真实网络连接,以模拟真实的业务环境[9]。根据攻击者的特点,计算分流后系统 λ 的入侵概率为

$$p_\lambda = \frac{1}{m+1} \times \exp\left(- \frac{b_i \times n_0}{S_u} \right) \tag{2}$$

式中,m 表示系统配置的蜜罐数量;b_i 表示第 i 个分流节点;n_0 表示递增凹函数,S_u 表示分流系数。

基于上式计算系统每个节点的入侵破坏因子,表达式为

$$d_k = \sum_{k=1}^{m_0} \frac{p_\lambda \times g_t}{\frac{a_t \times h_c}{s_y \times r}} \tag{3}$$

式中,m_0 表示系统节点数目,g_t 表示蜜罐的配置成本;a_t 表示蜜罐总的配置预算;h_c 表示黑客利用 c 种攻击蜜罐服务的概率;s_y 表示第 y 种攻击距离函数;r 表示混合系数;d_k 表示第 k 个节点的破坏因子。

d_k 可以反映节点的破坏程度。数值越大,该节点的破坏程度越大。因此,可将蜜罐依次部署在破坏因子较大的节点处[10],形成混合蜜罐环境,从而完成混合蜜罐环境的部署。

1.2 系统入侵行为特征优化

在混合蜜罐部署环境中,对智能云运维系统的入侵行为特征进行优化主要是通过分析网络流量数据,识别并细化能够指示潜在入侵行为特征[11],便于快速且准确地检测出恶意行为,并及时响应安全威胁。

首先采用标准化处理方式对系统内的入侵样本数据进行缩放,计算公式为

$$f_t = \frac{d_k \times \tau_0}{g_d \times p_w / b_\alpha} \tag{4}$$

式中,d_k 表示系统第 k 个节点的破坏因子;τ_0 表示数据的灰阶特征量,p_w 表示第 w 个系统节点的破坏因子对应的关联参数;b_α 表示第 α 个灰阶度对应的基准值参数尺度因子;g_d 表示第 d 个灰阶特征维度的原始数据值的平均贴近度,其计算方法为

$$g_d = \frac{2A_0 - f_u}{\sqrt{|\boldsymbol{C}_t|}} / h_t \tag{5}$$

式中,A_0 表示数据样本条数,f_u 表示第 u 个样本的数据属性特征数量;$\boldsymbol{C}_t$ 表示系数矩阵;h_t 表示朗德因子。

在数据缩放基础上,构造入侵行为特征的约束条件函数[12],表达式为

$$v_t = f_t \times \sum_{t=1}^{N} \left\| \frac{y_1}{r_\beta} \right\|^2 \tag{6}$$

式中,f_t 表示入侵行为特征约束项;N 表示数据采样次数;y_1 表示径向基函数,初始状态 1;r_β 表示第 β 次采样对应的惩罚因子。

利用权函数构造 2 维解空间,将原始数据映射到此空间内[13],获取数据的初始惯性权值为

$$w_0 = \frac{\frac{v_t \times \varepsilon_0}{e_1 / k_\varphi}}{B_s} + \frac{j_f}{\lambda_w} \tag{7}$$

式中,ε_0 表示核宽参数;e_1 表示运算误差,初始状态 1;k_φ 表示第 φ 个解空间维度解空间的值域;B_s 表示采样 s 的似真度函数;j_f 表示第 j 个原始数据样本模拟参数;λ_w 表示第 w 个解空间维度对应的正则化动量因子。

由此可采用下式提取入侵行为特征的隶属度[14]为

$$Y_t = w_0 \times \mu_0 \times q_g \tag{8}$$

式中,μ_0 表示特征变量;q_g 表示第 g 个入侵样本类内均值。

根据样本的类内散度 b_1 和类间散度 b_2，优化入^侵行为特征，表达式为

$$G_t = \frac{z_0(b_1 + b_2)}{| V_p \times \boldsymbol{c}_\varphi |} \tag{9}$$

式中，z_0 表示样本类内期望方差；V_p 表示第 p 个样本维度；c_φ 表示第 φ 个正对角加权矩阵；G_t 表示优化后的入侵行为特征。

在混合蜜罐部署环境基础上，构造入侵行为特征的约束条件函数，并计算其惯性权值，结合入侵特征的隶属度优化入侵行为特征，便于后续实现入侵检测。

1.3 系统入侵检测输出

通过上述方法优化入侵行为特征，采用神经网络方法构造入侵检测模型，并将当前入侵行为与特征库中的非正常行为进行比对，根据匹配阈值，判断待测数据是否为入侵数据[15]。系统入侵检测过程示意图如图 1 所示。

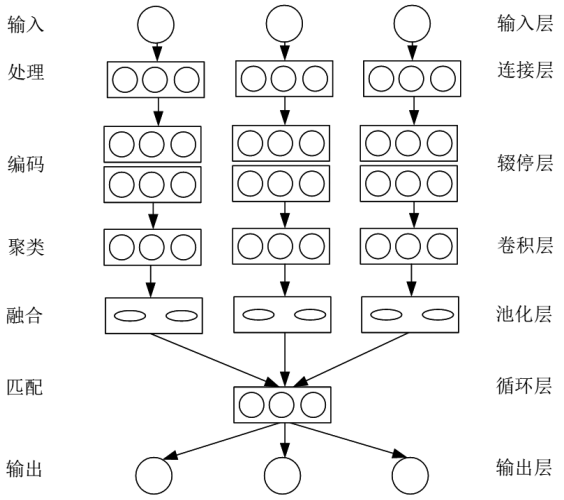

图 1 入侵检测模型

Fig. 1 Intrusion detection model

将优化后的入侵行为特征输入至文中，对其进行预处理，以消除数据冗余，并在辍停层对其编码，将第 y 个字符型数据转化为数值型数据为

$$D_y = G_t \times \boldsymbol{\kappa}_c \times \boldsymbol{e}_h \tag{10}$$

式中，$\boldsymbol{\kappa}_c$ 表示第 c 个编码向量；$\boldsymbol{e}_h$ 表示第 h 个白化矢量。

在卷积层，对入侵特征进行聚类分析，得到融合特征属性[16]为

$$E = \frac{D_y}{\rho_x} \tag{11}$$

式中，ρ_x 表示第 x 个卷积核尺寸。

对捕捉的入侵融合特征进行分解，并将其与特征库中的入侵特征进行比对，得到两者之间的匹配系数为

$$\omega = \frac{E \times r_{\tilde{\omega}}}{Q_t \times l_w} / \delta_e \tag{12}$$

式中，$r_{\tilde{\omega}}$ 表示第 w 次检测密度比；Q_t 表示决策函数；l_w 表示第 w 个检测范围；ω 表示匹配值；δ_e 表示特征库中的第 e 个入侵特征。

根据智能云运维系统的入侵检测需求和实际运行条件，设置匹配阈值 τ，并将求得的匹配值 ω 与 τ 相比较，若 $\omega > \tau$，则判定当前指令为异常行为，否则为正常运行行为。至此，完成智能云运维系统混合蜜罐入侵自动检测方法的设计。

2 实例论证分析

为验证本文提出的智能云运维系统混合蜜罐入侵自动化检测方法在实际应用中的效果，设计仿真对比实验，并根据实验结果分析该方法的检测效果。

2.1 实验准备

本次实验以某大型智能云运维系统为研究对象。该系统是一个基于大数据、云计算和边缘计算的综合运维管理平台。该系统能够对 IT 系统的各种指标进行实时监测、预测分析和自动化处理，实现故障预警、快速定位、自动修复等功能，其基本结构如图 2 所示。

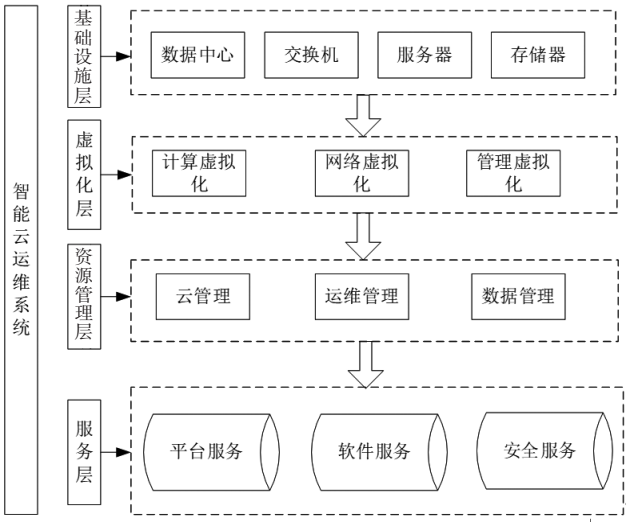

图 2 某大型智能云运维系统的基本结构

Fig. 2 Basic architecture of a large-scale intelligent cloud ops system

该系统中包含大量设备的实时数据、历史数据和日志数据。其中，实时数据通过系统内置的监控工具实时收集，包括设备状态、网络负载和服务器性能；历史数据存储在系统的数据库中，包括故障记录、服务器运行时间以及网络流量；日志数据用于分析系统的运行状况，包括系统日志和应用日志。

根据系统的基本需求和运行条件，选取低交互蜜罐、高交互蜜罐、虚拟蜜罐等混合型蜜罐作为服务环境，并利用虚拟化软件 VirtualBox 和 VMware 创建虚拟机作为蜜罐环境，结合系统节点的攻击破坏因子，完成混合蜜罐环境的部署。混合蜜罐配置信息如表 1 所示。

根据上述过程完成混合蜜罐环境的部署，并在此环境中进行智能云运维系统的入侵检测测试。

2.2 实验说明

使用基于 VMware vSphere 的云计算平台，模拟真实的云环境，根据系统结构和混合蜜罐环境，建立包括内部网络、DMZ 区域和外部网络的网络拓扑结构。使用 ELK 堆

栈收集和分析蜜罐生成的日志数据,并基于 Python 编写自动检测脚本,为每个蜜罐创建一个独立的索引,用于后续入侵行为的分析与检测。

表 1　混合蜜罐配置信息

Tab. 1　Configuration of hybrid honeypot

蜜罐类型	部署位置	伪装服务	监控工具
虚拟机蜜罐	私有云	SSH	Suricata IDS
物理设备蜜罐	DMZ 区域	FTP	
高交互蜜罐	DMZ 区域	HTTP	Snort
容器化蜜罐	容器云平台	MySQL	Zeek
低交互蜜罐	私有网络	SMTP	Bro
云服务蜜罐	公有云	POP3	Elastic Stack(ELK)
邮件蜜罐	私有云	RDP	Mobile Threat Defense
SSH 蜜罐	容器云平台	Telnet	OSSEC
数据库蜜罐	私有网络	DNS	Graylog
移动设备蜜罐	DMZ 区域	NTP	Kiwi Syslog
高交互蜜罐	私有网络	Telnet	ELK Stack

在建立系统入侵检测模型过程中,模型的基本参数设置如下。输入层节点 $N'=100$;输出层节点 $M'=2$,代表两种类别的概率;学习率 $a_g=0.01$;权重衰减系数 $s_y=0.02$;早停次数 $S=20$;匹配阈值 $_e\tau=60\%$;正则化权重 $\eta=0.3$。

基于以上参数设定与实验过程,对本文设计的方法进行性能验证与测试。

2.3　系统入侵检测结果与分析

实验中引入文献[2](方法 1)、文献[3](方法 2)作为本文方法的对比方法。分别采用 3 种方法对该智能云运维系统进行入侵检测。每组实验进行 25 次,以提高实验结果的可靠性,并在 5 组实验中,随机设置异常样本数据,分别统计基于 3 种方法下的异常样本检测数量,结果如图 3 所示。

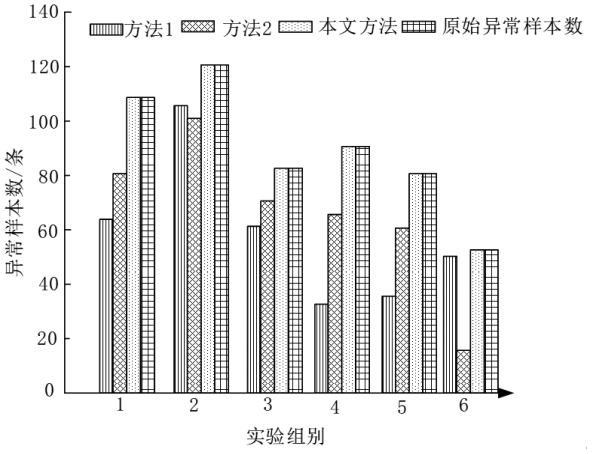

图 3　系统入侵检测结果

Fig. 3　Intrusion detection results of the system

根据图 3 可以看出,在 5 组实验中,对于不同异常样本数据,本文方法展现了较高的检测性能,能够完全检测出系统中的异常样本,显著优于对照组方法。由此说明了本文方法的优越性。

2.4　对比结果与分析

为综合体现本文方法的可行性,利用虚警率指标对 3 种方法的性能进行定性评估。较低的虚警率表明相应方法能够准确检测出系统的异常与入侵行为,检测准确度较高。对比结果如表 2 所示。

表 2　基于不同方法的入侵检测虚警率结果

Tab. 2　False positive rate results of intrusion detection based on different methods

攻击样本数量/条	检测虚警率/%		
	方法 1	方法 2	本文方法
100	0.47	0.36	0.20
200	0.59	0.41	0.11
400	0.45	0.35	0.21
800	0.66	0.44	0.14
1 600	0.58	0.42	0.06
3 200	0.60	0.59	0.22
6 400	0.71	0.55	0.19

由表 2 中的数据可知,本文设计的方法对于系统入侵检测虚警率更低,当系统攻击样本数量在 100~6 400 条范围内时,本文方法的虚警率始终保持在 0.3% 以下,而方法 1 和方法 2 的虚警率普遍较高,从而进一步证明了本文方法在系统入侵检测准确度方面的优越性。

3　结论

混合蜜罐入侵自动化检测方法在智能云运维系统中的实现过程融合了先进的网络安全技术和数据分析技术。通过模拟真实的系统环境,引诱并自动分析潜在的网络攻击行为。该方法不仅提高了入侵检测的准确性,也为云运维系统的安全稳定提供了有力保障。同时,该方法的研究对推动网络安全技术创新发展具有重要意义,为构建更安全可靠的云运维系统提供了重要支撑。

参考文献

[1]施永辉, 杨丽敏, 代琪, 等. Spark 框架下改进 TrAdaBoost 分布式入侵检测算法研究[J]. 中国电子科学研究院学报, 2023, 18(12): 1129-1137, 1145.

[2] MA W, LIU R, LI K, et al. An adversarial domain adaptation approach combining dual domain pairing strategy for IoT intrusion detection under few-shot samples[J]. Information Sciences: An International Journal, 2023, 44(36): 526-527.

[3] XIAO J, YANG L, ZHONG F, et al. Robust anomaly-based intrusion detection system for in-vehicle network by graph neural network framework[J]. Applied Intelligence, 2023, 53(3): 3183-3206.

[4]李芳. 应用深度自编码网络的局域网空间入侵监测系统设计[J]. 自动化技术与应用, 2024, 43(3): 91-95.

[5]要丽娟, 郭银芳. 基于集成学习的光纤光栅传感网络入侵行为检测[J]. 激光杂志, 2023, 44(11): 147-151.

[6]杜广周, 唐坤剑. 基于大数据驱动的光纤通信网络异常入侵检测研究[J]. 激光杂志, 2023, 44(11): 116-120.

[7]白万荣, 魏峰, 郑广远, 等. 基于 TCN-BiLSTM 的入侵检测算法研究[J]. 计算机科学, 2023, 50(S2): 941-948.

[8]刘晋钢, 刘晋霞, 曹小凤. 深度学习下增量式网络入侵实时检测算法仿真[J]. 计算机仿真, 2023, 40(11): 375-378.

[9]王心怡, 行鸿彦, 侯天浩, 等. 基于演化博弈的无线传感器网络入侵检测研究[J]. 电子测量与仪器学报, 2023, 37(10): 97-105.

[10]周明月, 常明航, 付殿臣, 等. 一种基于 EA-BinGRU 算法的网络入侵检测算法[J]. 计算机应用与软件, 2023, 40(11): 321-326.

(下转第 113 页)