计算机研究与发展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202440387
61(10): 2404−2416, 2024

动态决策驱动的工控网络数据要素威胁检测方法

王泽鹏　马　超　张壮壮　吴黎兵　石小川
（武汉大学国家网络安全学院　武汉　430072）
（空天信息安全与可信计算教育部重点实验室（武汉大学）　武汉　430072）
（wangzepeng@whu.edu.cn）

Dynamic Decision-Driven Threat Detection Method for Data Elements in Industrial Control Networks

Wang Zepeng, Ma Chao, Zhang Zhuangzhuang, Wu Libing, and Shi Xiaochuan
（*School of Cyber Science Engineering-WHU*, *Wuhan University* 430072）
（*Key Laboratory of Aerospace Information Security and Trusted Computing* (*Wuhan University*), *Ministry of Education*, *Wuhan* 430072）

Abstract　In recent years, the industrial control network has been developing rapidly. The advantages of digitization, intelligence, and automation have brought significant benefits to the industry while also introducing increasingly complex and variable attack threats. In the context of data element security, timely detection and response to industrial control network threats have become an urgent task to be solved. By continuously monitoring and analyzing the data flow in industrial control networks, the problem of industrial control network threat detection can be transformed into a time series anomaly detection problem. However, the existing time-series anomaly detection methods are limited by the quality of industrial control network datasets and are often sensitive to only a single type of anomaly while ignoring other anomalies. Therefore, in this paper, we propose a deep reinforcement learning and data augmentation based threat detection method in industrial control networks (DELTA). DELTA introduces a novel data augmentation selection technique for time series datasets, which allows for the selection of appropriate data augmentation operations sets tailored to different baseline models to enhance the quality of the industrial control network time series datasets. Simultaneously, deep reinforcement learning algorithms (A2C/PPO) dynamically select candidate models from the baseline models at different time points, leveraging multiple types of anomaly detection models to address the issue of sensitivity to single-type anomalies. The experimental results compared with the existing time series anomaly detection models show that DELTA has a significant improvement in accuracy and $F1$ value over all baseline models at an acceptable cost of additional time consumption, which verifies the effectiveness and practicality of the method.

Key words　industrial control networks; data element security; time series; anomaly detection; deep reinforcement learning; data augmentation; model selection

摘　要　近年来，工控网络发展势头迅猛. 其数字化、智能化、自动化的优势为工业带来巨大效益的同时，也面临着愈发复杂多变的攻击威胁. 在数据要素安全的背景下，及时发现和应对工控网络威胁成为一项

收稿日期：2024-05-31；**修回日期**：2024-07-17
基金项目：国家重点研发计划项目（2021YFB3101100）；国家自然科学基金项目（62272352）；湖北省重点研发计划项目（2021BAA039）；湖北省自然科学基金项目（2022CFB012）
This work was supported by the National Key Research and Development Program of China (2021YFB3101100), the National Natural Science Foundation of China (62272352), the Key Research and Development Program of Hubei Province (2021BAA039), and the Natural Science Foundation of Hubei Province (2022CFB012).
通信作者：石小川（shixiaochuan@whu.edu.cn）

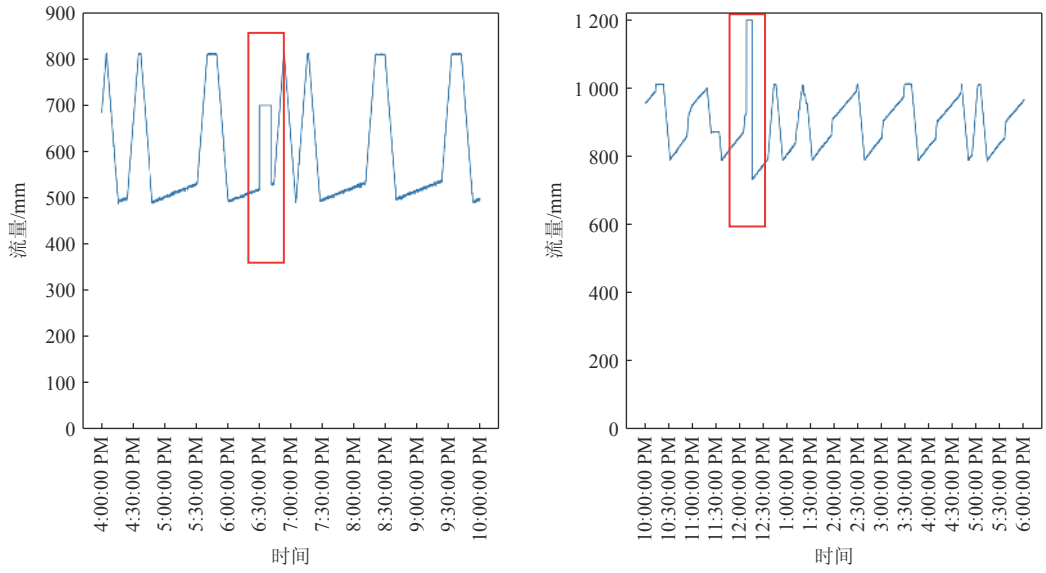
迫切需要得到解决的任务. 通过对工控网络中的数据流进行连续监测和分析，工控网络威胁检测问题可以转化为时间序列异常检测问题. 然而现有时间序列异常检测方法受限于工控网络数据集的质量，且往往仅对单一类型异常敏感而忽略其他异常. 针对上述问题，提出了一种基于深度强化学习和数据增强的工控网络威胁检测方法（deep reinforcement learning and data augmentation based threat detection method in industrial control networks, DELTA）. 该方法提出了一种新的时序数据集数据增强选择方法，可以针对不同的基准模型选择合适的数据增强操作集以提升工控网络时间序列数据集的质量；同时使用深度强化学习算法（A2C/PPO）在不同时间点从基线模型中动态选取候选模型，以利用多种类型的异常检测模型解决单一类型异常敏感问题. 与现有时间序列异常检测模型对比的实验结果表明，在付出可接受的额外时间消耗成本下，DELTA 在准确率和 $F1$ 值上比所有基线模型有明显的提升，验证了方法的有效性与实用性.

关键词 工控网络；数据要素安全；时间序列；异常检测；深度强化学习；数据增强；模型选择

中图法分类号 TP391

工控系统作为现代工业自动化的核心组成部分，广泛应用于能源、制造、交通等关键基础设施领域. 工控网络威胁检测是针对工业控制系统网络中的安全威胁进行监测和识别的重要领域. 随着工业物联网(IoT)和工业 4.0 的快速发展，工控系统的数字化、智能化和互联化程度的不断提高，工控系统的安全性成为了一个日益突出的问题. 特别是在数据要素安全的背景下，工控网络面临着多样化的威胁和越来越多的潜在威胁和攻击风险. 这些威胁不仅涉及传统的网络安全问题，而且包括针对特定数据要素的攻击，如数据篡改、非授权访问等. 网络攻击者可以利用漏洞、恶意软件或未经授权的访问，对工控网络进行入侵，从而导致设备故障、生产中断甚至人身安全风险. 确保工控网络的安全性和稳定性成为保护关键基础设施的重要任务. 因此，开发能够有效检测和应对这些威胁的技术至关重要.

时间序列数据是按时间顺序收集的观测值的序列，如股票价格、气象数据、传感器测量等，已成为各个领域中不可或缺的重要资源. 然而，这些数据中常常存在着异常，即与正常模式不符的异常值或行为. 异常可能是由故障、欺诈、异常事件或未知因素引起的，对于电网、通信网络等工控网络关键基础设施来说及早识别和处理这些异常非常关键. 如图 1 所示，图 1(a)和 1(b)分别为工控安全水处理系统的液位传感器与流量计实时收集的流量信息，当系统受到攻击时，流量数据时间序列会表现出与正常数据时间序列不同的异常模式，同时对系统的正常运行造成不同程度的破坏与影响. 因此，工控网络威胁检



(a) LIT101液位传感器异常　　(b) LIT301流量计异常

Fig. 1 Examples of sensor exceptions in the SWaT dataset

图 1 SWaT 数据集中的传感器异常示例

测可以自然地建模为时间序列异常检测问题，时间序列异常检测也成为了对于工控网络系统数据分析和预测中一个重要而具有挑战性的问题[1-3].

时间序列异常检测是一项旨在识别和定位时间序列数据中异常值的任务. 与传统的离群点检测方法不同，时间序列异常检测需要考虑数据的时序性质，即异常值与周围数据点的时间依赖性. 传统的基于阈值或统计方法的离群点检测技术在处理时间序列数据时往往效果不佳，因为它们无法捕捉到时间序列中的复杂模式和关联性. 而基于监督学习的算法可以通过从数据中提取特征并建立分类模型，使其在异常检测任务上表现良好，但其依赖于大量的标签来训练模型，这对于现实世界是很难满足的. 因此，在无监督环境下开发针对时间序列异常检测的专门方法和算法是至关重要的.

此外，由于不同类型的异常检测算法侧重点不同，单一类型的异常检测模型可能会对特定类型的异常敏感，而忽视其他异常. 而传统的集成学习方法难以确定基础模型的权重来发挥组合模型的最优性能. 如何设计一个通用的模型可以针对存在的不同类型异常的数据在检测性能上完全超越其他异常检测模型，成为一个困难且有研究价值的问题.

为解决上述问题，本文提出了一种基于深度强化学习和数据增强的工控网络威胁检测方法(deep reinforcement learning and data augmentation based threat detection method in industrial control networks, DELTA). 首先提出一种可靠的数据增强方法，以提升时序数据集异常数据的多样性. 然后利用增强数据集训练多个不同类型的异常检测基线模型. 最后利用深度强化学习算法在不同时间步数据点从基线模型中动态选取合适的模型用于最终的异常检测. 本文的主要贡献有 3 个方面：

1)提出了一种新的时序数据集数据增强选择方法，提升了时序数据集样本的质量，为后续异常检测模型和基于深度强化学习的模型筛选的训练提供了基础.

2)提出了使用深度强化学习在不同时间步数据点上筛选合适的基准模型进行组合预测的方案，结合不同类型异常检测模型的优势，提高了异常检测的总体性能.

3)在公共和真实的数据集上对 DELTA 方法进行了评估，与其他基线模型相比，DELTA 方法在准确率和 $F1$ 值 2 个关键指标上有显著的提升，在时序数据异常检测任务上展现了明显的优势.

1 相关工作

1.1 无监督环境下的多变量时间序列异常检测

近年来越来越多的研究开始关注异常检测领域并取得了一定的成果[4-5]. 尽管大部分的异常检测算法并非针对时间序列异常检测任务所提出的，但这些算法在时间序列任务上仍有良好的应用.

由于不需要大量的标签数据，且具有推断时间序列之间相关性的能力，基于无监督学习技术的算法在异常检测领域越来越受到研究者的关注[6]. 本节主要对近年来的基于无监督学习的多变量时间序列异常检测方法进行介绍.

Audibert 等人[6]提出了一种基于反向训练自编码器的快速稳定的多变量时间序列异常检测方法 USAD，对抗性训练的架构使得其能够进行快速训练的同时隔离异常，在多方面上提升了异常检测性能. Zhang 等人[7]提出的 MSCRED 联合考虑了时序依赖性、噪声鲁棒性和异常严重程度的可解释性，有效地提升了多变量时间序列异常检测和诊断的能力. LSTM-VAE[8]使用 LSTM 替换 VAE 中的前馈网络，对时序数据的时间依赖性进行建模，得到了比传统方法更好的泛化能力. Su 等人[9]提出了一种使用随机递归神经网络的多变量时间序列异常检测模型 OmniAnomaly，通过随机变量连接和平面归一化流等关键技术来捕获鲁棒的多变量时间序列的正常表示以重构输入数据，并使用重建概率来确定异常.

1.2 基于集成学习的异常检测方法

集成学习通过组合多个基础模型，以减小单个模型带来的误差，提高最终模型对于任务决策的准确率，近年来也愈发受到研究者的关注[10-11].

在异常检测领域，最常用的方法是通过多个基础模型的最大值或平均值来计算异常得分. 一些研究关注基于机器学习和深度学习的集成学习方法[12]. Tama 等人[13]使用 3 种分类器(随机森林、梯度提升器和 XGBoost)作为基础模型，并使用一个广义线性模型作为组合器. Adeyemo 等人[14]结合深度学习和集成学习，将 LSTM 网络与随机森林等算法结合，采取投票集合的平均概率方法解决网络入侵检测问题. Ünlü等人[15]提供了一种基于内部聚类质量度量的加权策略，赋予单个聚类不同的重要性. 这些集合学习研究主要集中在权重计算或不同基础模型的线性组合上.

1.3 时间序列数据增强

如图 2 所示，针对时间序列数据的数据增强方法可以分为 2 大类：基础方法和高级方法[16].

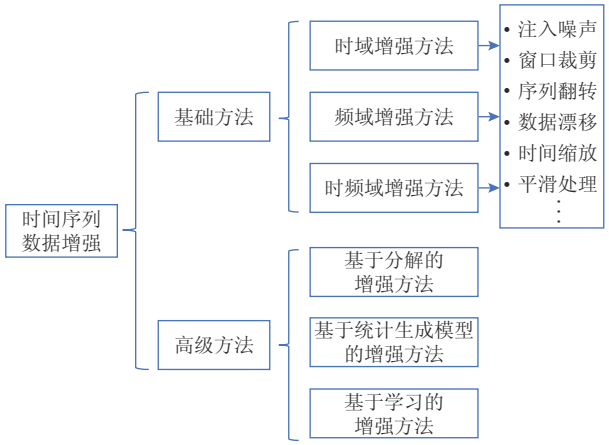


Fig. 2 Classification of time series data augmentation methods

图 2 时间序列数据增强方法分类

基础方法主要是针对时间序列在不同域(时域、频域、时频域)上进行多种变换. 其中，时域变换是应用最广泛的时间序列数据增强方法，该类方法直接对原始输入时间序列进行变换操作，如注入噪声、窗口裁剪、窗口变形、翻转等. 频域和时频域上的变换方法研究相对较少. Gao 等人[17] 同时利用时域上和频域上的数据增强策略，在基于卷积神经网络的时间序列异常检测任务中生成更多的人工标记数据，有效地提升了时间序列异常检测的性能. Steven 等人[18] 采用短傅里叶变换(STFT)对传感器时间序列生成时频特征并进行数据增强. 类似地，Park 等人[19] 在语音时间序列上通过屏蔽翻转 IAAFT 增强时间步长块，大大提高了语音识别神经网络的性能.

高级方法可以分为基于分解的方法、基于统计生成模型的方法和基于学习的方法. 基于分解的时间序列增强方法在时间序列异常检测和预测等相关任务上表现优秀. Park 等人[20] 将时间序列中的每一个数据点分解为趋势信号、周期信号和余数信号 3 种信号的叠加. Gao 等人[17] 提出使用鲁棒性分解方法，利用生成的分解残差对时间序列进行时域和频域上的增强，显著地提高了异常检测的性能. 基于统计生成模型的时间序列增强方法通常是使用统计模型对时间序列进行动态建模. Li 等人[21] 提出了一个耦合扩散概率模型，以在不增加不确定性的情况下增强时间序列数据，并利用 BVAE 实现更简便的推理过程. Kang 等人[22] 使用混合自回归模型来模拟时间序列集，以探究生成的时间序列在特征空间上的多样性和覆盖度. 最近，基于学习的方法展现了对真实数据的特征进行模拟的潜力. Devries 等人[23] 提出在嵌入空间中应用编码输入以产生更可信的生成时序数据，Cheung 等人[24] 则利用自动数据增强来微调隐空间中的 4 种通用数据转换操作，以适应不同模态的数据转换，在时间序列分类问题上展现出了优异的性能. Yoon 等人[25] 提出了一种生成式时间序列模型 TimeGAN，通过具有监督损失和无监督损失的嵌入空间进行对抗性地训练和学习. Fons 等人[26] 使用自动数据增强方法为时间序列数据设计了 2 种样本自适应加权策略，在多个时间序列数据集上的分类问题都展现出了良好的性能.

2 基于深度强化学习和数据增强的工控网络威胁检测方法（DELTA）

本文所提出的 DELTA 整体框架如图 3 所示.

整体流程主要分为 3 步：首先，在数据处理阶段，DELTA 利用模型性能评分迭代评估生成数据增强操

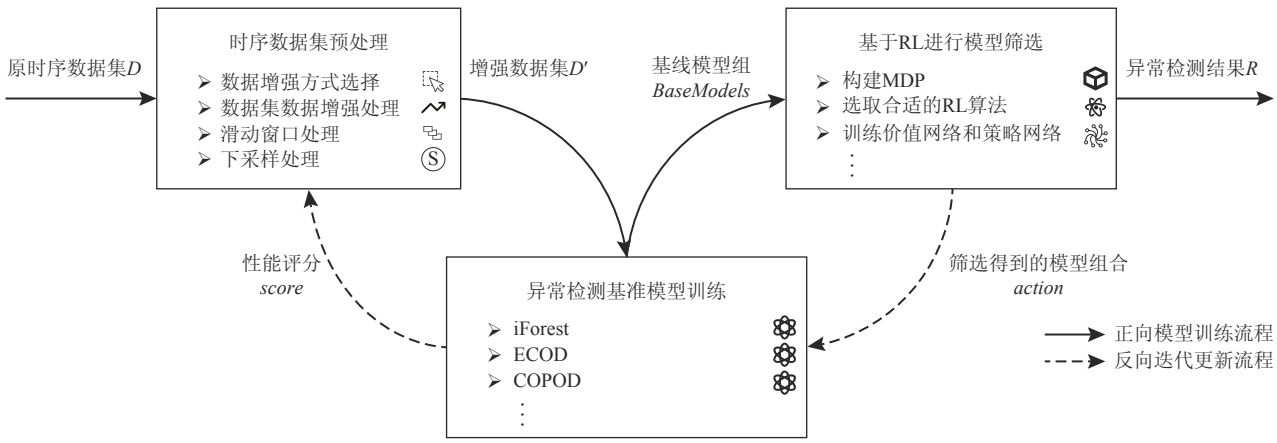


Fig. 3 Overall framework diagram of DELTA

图 3 DELTA 整体框架图

作组合,针对不同的模型使用不同的操作组合对时间序列数据集进行数据增强,以提升数据集质量,再使用滑动窗口和下采样技术处理得到最终用于训练的增强数据;其次,使用增强数据集对多种不同类型的异常检测基线模型进行训练,同时生成每个基线模型对于数据集每个测试点的预测分数、异常检测器自身的异常阈值以及基于异常阈值得到的预测标签;最后将多种信息集成到用于强化学习建模的马尔可夫决策过程中,利用强化学习算法(A2C/PPO)在不同时间步对数据集上的不同数据点从训练好的异常检测基线模型中动态选择候选模型进行最终的预测.

2.1 基于数据增强的时序数据集预处理

为保证时间序列数据集数据的多样性,提升数据集数据质量,在进行模型训练和筛选之前,本文提出先使用数据增强技术对时序数据集进行增强的思路.

参考已有典型工作[17,24],本文采取了多种经验证有效的时序数据增强方式,包括注入噪声、数据漂移、时间缩放、平滑处理、时间剪切等多种类型的增强手段,并对多种增强方式进行组合,对于每一个异常检测基准模型,根据增强前后的异常检测器的检测性能来迭代更新每轮选取的增强方式,最终为每一个异常检测基准模型生成一个最佳的增强方式组合,进而用于后续的数据处理与模型训练.

数据增强方式选择的流程如算法 1 所示.

算法 1. 时序数据集数据增强选择方法.

输入:数据增强可选操作集 DA,训练集 D_{train},测试集 D_{test},异常检测基准模型 M,数据增强操作数量 n,迭代轮数 T;

输出:数据增强操作方式组合 DAC.

① $M = TrainModel(M, D_{\text{train}})$; /*使用原始训练集 D_{train} 训练模型 M*/

② $S_0 = TestModel(M, D_{\text{test}})$; /*使用测试集 D_{test} 测试,得到初始性能得分 S_0*/

③ $DA = DA \cup \{\text{None}\}$; /*为数据增强可选操作集增加一个空操作*/

④ $DAC = \{da_1, da_2, \cdots, da_n\}$; /*利用 DA 中的可选操作随机初始化数据增强方式组合*/

⑤ $DAC_{\text{tmp}} \leftarrow DAC$; /*初始化一个临时数据增强方式组合*/

⑥ for $i = 1$ to T

⑦ $D_{\text{train}} = Aug(D_{\text{train}}, DAC_{\text{tmp}})$; /*使用 DAC_{tmp} 对数据集进行数据增强处理*/

⑧ $M = TrainModel(M, D_{\text{train}})$; /*利用增强后的数据集 D_{train} 训练异常检测基准模型 M*/

⑨ $S_i = TestModel(M, D_{\text{test}})$; /*在测试集 D_{test} 上评估模型 M 的性能,得到性能得分 S_i*/

⑩ if $S_0 < S_i$ then

⑪ $DAC \leftarrow DAC_{\text{tmp}}$; /*更新最优数据增强方式组合*/

⑫ $S_0 \leftarrow S_i$;

⑬ else

⑭ $DAC_{\text{tmp}} \leftarrow DAC$; /*回退为上一轮的数据增强方式组合*/

⑮ end if

⑯ $da' = RandomChoice1(DA)$; /*从 DA 中随机选择一个增强操作 da'*/

⑰ $j = RandomChoice2\left(DAC_{\text{tmp}}\right)$; /*从 DAC_{tmp} 中随机选择一个位置 j*/

⑱ $DAC_{\text{tmp}}\left[j\right] \leftarrow da'$;

⑲ end for

具体来说,首先使用原始训练集训练得到一个异常检测模型,并评估其在测试集上的初步性能.然后引入一个包含空操作的数据增强操作集,并从中随机选择初始的数据增强组合.在每次迭代中,算法使用当前的数据增强组合处理训练数据,并重新训练模型,然后在测试集上评估模型的性能.如果新模型的性能超过了之前的最佳性能,算法将更新最优的数据增强组合;否则,保持上一轮的数据增强组合不变.此外,算法在每轮迭代中随机更换一个数据增强操作,以探索更多可能的操作组合.通过设定的迭代轮数,算法最终确定一组能够最大化模型性能的数据增强操作组合.

此外,在上述数据增强流程的基础上,还可以采取 3 种可选的措施来降低算法复杂度:

1)控制数据增强操作集中选取的数据增强操作数量 n.由于初始数据增强操作集的选取是经验性的,而针对不同的数据集及异常检测模型,不同数据增强操作的作用程度是不同的,往往只有少数的操作发挥效果明显,因此通过限制 n 的大小来得到多个最优操作的集合是更高效的操作.

2)控制数据集的采样率.在对数据增强操作集评估时,由于需要训练并测试一个完整的异常检测模型,时间成本较大.因此可以通过控制数据集的采样率,选取一定比例的数据及增强后的数据组成新的训练数据集用于训练.尽管数据集样本数量的减少会导致最终性能的下降,但由于增强前后使用相

同规模的数据集进行训练与检测，性能的优劣变化可以反映数据增强操作的有效与否，只需保证每次数据集的采样率保持一致，即可大大提升训练效率，降低算法复杂度.

3）子策略自适应调整. 根据性能评估的结果，可以动态地调整子操作的更新频率. 如果某一个子操作连续几轮表现较好，可以减少其更新频率，以节省计算资源. 相反，如果某一个子操作连续几轮表现不佳，可以增加其更新频率，以尝试更多的改进. 这样可以根据子操作的表现情况自适应地调整更新策略，提高效率.

同时，在数据预处理阶段，除了进行时序数据增强操作外，还进行了滑动窗口和下采样处理.

滑动窗口处理是将时序数据转换为以大小固定的窗口为单位的样本序列，帮助模型捕捉到时序数据中的局部模式和关联性. 假设训练集或测试集中的一个时间序列$X=\{x_1,x_2,\cdots,x_n\}$，本文将其处理为一组新的样本集合：

$$X_w=\{(x_1,x_2,\cdots,x_w),(x_2,x_3,\cdots,x_{w+1}),\cdots,(x_{n-w+1},x_{n-w+2},\cdots,x_n)\},\tag{1}$$

其中，w为滑动窗口的大小，每个样本的大小为w，样本X_i的标签由该样本最后一个时间点的数据确定，当该时间点的原始数据为异常样本时，则样本X_i的标签设置为1（异常），否则设置为0（正常）. 滑动窗口操作将更多的时间信息作为输入特征，以便于模型更好地学习和训练.

下采样操作是通过减少采样频率或样本数量来减小数据集的规模，由于时序数据集规模较大（包含大量时间步的多维信息）. 本文采取和文献[6, 27]一致的下采样率，每5个时间步进行1次平均值计算，这样下采样处理后的数据集规模为原数据集的20%，可以显著减少计算和存储需求，减少模型训练和推理的计算负担，从而提高效率. 同时某些时序数据集可能存在长期依赖关系，即当前样本的预测可能与过去较远的样本相关. 在这种情况下，下采样可以通过降低采样频率，减少时间步长之间的间隔，以更好地捕捉长期依赖关系，从而提高模型的性能.

2.2 多类别异常检测基准模型训练与关键信息提取

经过数据增强和滑动窗口、下采样处理后的时序数据集将用于多类别异常检测基准模型的训练. 一个多类别异常检测基准模型的训练流程为：

在无监督环境下，经过数据增强、滑动窗口处理及下采样处理后的训练样本集合表示为$X_{\text{train}}=\{(x_1,x_2,\cdots,x_w),\cdots,(x_{n-w+1},\cdots,x_n)\}$，其中$n$为训练样本数量，在无监督环境下，训练数据集全部由正常数据组成；经过滑动窗口处理的测试数据集可以表示为$X_{\text{test}}=\{(x_1,x_2,\cdots,x_w),\cdots,(x_{m-w+1},\cdots,x_m)\}$，其中$m$为测试样本数量，测试数据集由正常数据和少量异常数据组成. 在训练阶段，使用增强后的训练集数据对异常检测模型进行训练，学习捕获正常时序数据的特征；在测试阶段，训练好的异常检测模型将对测试集数据的每个测试样本输出一个异常值分数，用于表征该样本为异常样本的可能性，异常值分数越高则该样本为异常的可能性越大，将该异常值分数与每一个基准模型预先设置的一个经验阈值来决定最终的预测标签为1（异常）或为0（正常）.

为了后续利用深度强化学习进行模型筛选，DELTA尝试从异常检测基准模型中提取到更多深层次的信息以利于后续强化学习的训练. 受文献[27]的启发，DELTA使用2个分数来表征模型对预测结果的自信程度：距离阈值置信分数和预测共识置信分数.

距离阈值置信分数反映了模型预测输出异常值分数超出预设异常阈值的程度，置信分数越高则表示该模型对于该样本被认定为异常的自信程度越高，同时本文对该分数进行了最大最小值尺度缩放，将其归一化为[0,1]的范围以便于训练，表示为：

$$s_{\text{distance}}=\frac{s-pt}{max_s-min_s},\tag{2}$$

其中，s为模型预测输出异常值分数，pt为预设异常阈值，max_s为模型预测输出异常值分数的最大值，min_s为模型预测输出异常值分数的最小值.

预测共识置信分数反映了所有基准模型对于样本预测结果所达成一致的程度，置信分数越高则表示该样本被所有基准模型认定为同一类型的共识程度越高，表示为：

$$s_{\text{consensus}}=\begin{cases}\dfrac{\sum_{i=0}^{m-1}(pl_i==1)}{m}, & \text{if } l_{\text{output}}\geqslant d,\\[2ex] \dfrac{\sum_{i=0}^{m-1}(pl_i==0)}{m}, & \text{if } l_{\text{output}}<d,\end{cases}\tag{3}$$

其中，m为基准模型的数量，pl_i为第$i+1$个基准模型的预测标签，l_{output}为经过模型筛选后输出的加权标签，$d\in[0,1]$为一个超参数，作为控制模型最终输出为异常或正常的阈值，d越大则模型输出为异常的条件更严格，当$d=1$时，意味着当且仅当全部基准模型输出的预测标签均为异常时，模型最终才输出为异常.

此外, 利用后续经过使用强化学习模型筛选得到的结果, 还可以提取到 3 种类型的关键信息: 加权归一化预测分数、加权归一化异常阈值、加权预测标签.

加权归一化预测分数是经过模型筛选后对原基准模型预测分数进行加权处理得到的新的预测分数, 反映了经过模型筛选后样本异常值的预测结果, 表示为:

$$s_{scaled}=\frac{\sum_{i=0}^{m-1} scaled_s_i\times a[i]}{m}, \tag{4}$$

其中, m为异常检测基准模型的数量, $scaled_s_i$为第$i+1$个基准模型经过归一化处理得到的原始预测分数, $a[i]$为模型筛选时是否选择该模型用于后续预测.

加权归一化异常阈值是经过模型筛选后对原基准模型的多个异常阈值 pt 进行加权处理得到的新的异常预测阈值, 反映了经过模型筛选后对样本标签预测为异常与否的新的阈值, 表示为:

$$pt_{scaled}=\frac{\sum_{i=0}^{m-1} scaled_pt_i\times a[i]}{m}, \tag{5}$$

其中, $scaled_pt_i$为第$i+1$个基准模型经过归一化处理得到的原始异常预测阈值.

加权预测标签是经过模型筛选后对测试集数据点标签预测进行加权处理得到的新的标签, 表示为:

$$l_{\text{output}}=\frac{\sum_{i=0}^{m-1} l_i\times a[i]}{m}, \tag{6}$$

其中, l_i为第$i+1$个基准模型的原始预测标签.

2.3 基于深度强化学习的异常检测模型筛选

异常检测基准模型训练完毕后, DELTA 进一步将模型选择问题建模为马尔可夫决策过程[28], 对应的元素具体描述为:

1)状态. 状态空间由多个异常检测基准模型(具体使用的基准模型将在实验部分进行介绍)训练后提取得到的关键信息所组成, 包括加权归一化预测分数、加权归一化异常阈值、加权预测标签、预测距离置信分数和预测共识置信分数共 5 项关键信息, 具体信息的定义已在 2.2 节介绍.

2)动作. 动作空间设置为多维离散空间$A=\{a_1, a_2,\cdots,a_n\}$, 维数$n$为异常检测基准模型的数量, 每一维的动作是离散的二值变量, 表示当前状态下是否选择该模型用于后续的异常检测(“1”为选取, “0”为不选取).

3)状态转移矩阵. 在该问题环境下, 状态按照数据集中的时间戳的顺序进行确定性的转移, 即对于数据集中连续的 2 个时间步对应产生的状态s_t和s_{t+1}, 对应的状态转移概率$P(s_{t+1}|s_t)=1$, 而对于其他情况下$P(s_{t+1}|s_t)=0$.

4)奖励. 奖励函数是根据当前时间步模型的预测标签与真实标签的对比确定的, 真阳性(true positive, TP)表示预测标签和真实标签均为异常; 真阴性(true negative, TN)表示预测标签和真实标签均为正常; 假阳性(false positive, FP)表示预测标签为异常、真实标签为正常; 假阴性(false negative, FN)表示预测标签为正常、真实标签为异常, 根据这 4 种情况构建奖励函数, 对真阳性和真阴性给予正奖励反馈, 对假阳性和假阴性给予负奖励反馈, 具体表示为:

$$R=\begin{cases} r_1, & TP,\\ r_2, & TN,\\ -r_3, & FP,\\ -r_4, & FN, \end{cases} \tag{7}$$

其中, r_1, r_2, r_3, r_4均为正值.

5)折扣因子. 在该问题环境下, 无需对未来奖励进行衰减处理, 因此折扣因子设置为 1, 对应的回报函数为无折扣累积回报.

马尔可夫决策过程构建完毕后, 便可以选取合适的深度强化学习算法来进行学习训练. 由于动作空间为多维离散空间, 本文选择使用了 A2C(advantage Actor-Critic)和 PPO(proximal policy optimization)[29] 两种深度强化学习算法进行实验. 但值得强调的是, 本文所提出的方法框架具有良好的算法兼容性, 用于训练使用的强化学习算法类型并不受限, 理论上任何适用于多维离散动作空间下的强化学习算法都可适用于本文所提出的方法框架.

为展现框架的通用性, 选取的 2 种算法均是经过时间检验证实有效的经典强化学习算法, 二者均属于基于策略梯度的强化学习算法. 其中, A2C 使用基于价值函数的优势估计来更新策略, 可以减少更新的方差, 有良好的训练稳定性和收敛性; PPO 则通过对策略更新的幅度进行限制, 保证新旧策略之前的更新幅度不会过大, 同样有助于提高训练的稳定性和收敛性.

依赖于神经网络强大的拟合能力, 2 种算法均使用神经网络对价值函数和策略函数进行拟合, 分别用V_ω和π_θ表示. 对于价值函数, 二者均使用均方误差作为价值函数的损失函数, 使用梯度下降法进行更新, 其损失函数表示为:

$$Loss = MSE(r+\gamma V_{\omega}(s'), V_{\omega}(s)). \tag{8}$$

对于策略函数，A2C 采取最大化策略梯度的方式来进行策略网络参数的更新：

$$\theta = \theta + \alpha \nabla_{\theta} \log \pi_{\theta}(s_t, a_t) A_{\omega}(s_t, a_t), \tag{9}$$

其中A_{ω}为优势函数，用于反映当前状态s_t采取动作a_t相对于平均程度的优势大小，α为学习率.

而 PPO 则采取重要性采样、目标函数裁剪等多种优化手段对策略网络参数进行更新，其损失函数表示为：

$$Loss = \sum_{(s_t,a_t)} (\min(ratio \cdot A_{\omega}(s_t,a_t), clip(ratio, 1-\varepsilon, 1+\varepsilon) \cdot A_{\omega}(s_t,a_t))), \tag{10}$$

其中$ratio = \dfrac{\pi_{\theta^{\text{new}}}(a_t|s_t)}{\pi_{\theta^{\text{old}}}(a_t|s_t)}$，$clip()$为裁剪函数，参数$\varepsilon$为超参数，三者共同控制策略更新幅度不会过大或过小.

3 实验结果与分析

3.1 实验设置

3.1.1 数据集

本文选取用于进行工控网络威胁检测实验评估的数据集为安全水处理数据集(SWaT)[30]，该数据集由新加坡理工大学网络安全中心 iTrust 所采集并开源，其模拟了一个工业水处理厂的日常运行情况，捕获了试验台内 51 个传感器和执行器的运行数据，包含了 7 天的正常运行数据和 4 天的异常(被攻击)运行数据. 具体来说，SWaT 是一个具有 51 个特征的多源时间序列数据集，包含 495 000 个时间步的正常数据作为训练集，包含 449 919 个时间步的异常数据作为测试集，其中测试集中异常数据占比为 12.13%. 该数据集可以真实反映一个现实工控环境下的标准物理系统的运行过程.

3.1.2 对比模型

本文选取了 8 种不同类型的时间序列异常检测模型作为对比模型用于实验性能评估. 其中 7 种模型均基于无监督学习算法，具体介绍为：

1)iForest(isolation forest)算法[31]. iForest 算法是异常检测最常用的算法之一，其利用异常样本与正常样本的数量及特征值差异很大的特点，能够在更少的随机特征划分次数中有效地分离出异常样本. 主要优点是它不需要依赖复杂的统计模型和假设，且具有较低的计算复杂度，对于高维数据和大规模数据集也能够有效地进行异常检测.

2)DIF (deep isolation forest for anomaly detection)[32]. DIF 在 iForest 算法的基础上，利用随意初始化的神经网络将原始数据映射到随机表示集合中，随后应用随机轴平行切割来执行数据分割. 这种表示方案有利于在原始数据空间中进行高自由度的分割，从而促进随机表示和基于随机分割的隔离之间的独特协同作用. 同时，作为一种集成学习方法，该模型还提出了一种高效计算表示集合方法，以减小计算开销.

3)LODA(lightweight on-line detector of anomalies)[33]. LODA 是一种轻量级的基于集成学习的异常检测模型，主要通过一组非常弱的检测器组成的集合产生一个检测能力更高的检测器.

4)VAE(variational autoencoder)[34]. VAE 是一种基于变分推理的生成模型，同时结合了深度学习和概率模型二者的优点，其通过最小化重构误差和最大化隐分布的似然函数学习得到数据的真实分布.

5)beta-VAE (beta variational autoencoder)[35]. beta-VAE 作为 VAE 算法的一种变体，在 VAE 的基础上通过引入一个 beta 超参数，以控制损失项中 KL 散度的重要程度，从而控制模型表示能力与解纠缠能力之间的平衡，当 beta 为 1 时 beta-VAE 退化为原始的 VAE.

6)ECOD(empirical cumulative distribution for outlier detection)[36]. ECOD 是一种基于经验累积分布的异常值检测算法. ECOD 首先计算数据每个维度上的经验累积分布，再利用经验累积分布来估计每个维度上的尾部概率，最后通过汇总在各维度上估计的尾部概率来计算每个数据点的异常评分.

7)COPOD(copula-based outlier detection)[37]. COPOD 是一种基于 Copula(一种用于建模多变量分布的函数)的多变量数据异常检测算法. COPOD 使用 Copula 来描述多变量数据中的依赖结构，并计算每个样本在 Copula 模型下的概率密度函数，最后得到每个数据点的异常评分.

此外，本文同样选取了一种基于强化学习的异常检测模型 RLMSAD[27] 作为对比模型，RLMSAD 同样是一种使用强化学习来进行异常检测模型选择的方法，但是不同于本文提出的方法，RLMSAD 仅使用 DQN 算法来对一维动作空间的模型进行选择，且每次只选择一个异常检测模型进行异常评分的预测，没有将所有异常检测模型提供的关键信息完全利用.

3.1.3 评价指标

本文使用准确率(P)、召回率(R)和$F1$值作为异常检测性能的评估指标：

$$\begin{cases} P=\dfrac{TP}{TP+FP}, \\ R=\dfrac{TP}{TP+FN}, \\ F1=\dfrac{2\times P\times R}{P+R}, \end{cases} \tag{11}$$

其中, 异常样本被视为正样本, 正常样本被视为负样本. TP表示异常样本被模型正确预测为异常的样本个数, FP表示正常样本被模型错误预测为异常的样本个数, FN表示异常样本被模型错误预测为正常的样本个数. 准确率P表示被模型检测为异常的样本中实际异常样本的百分比, 召回率R表示实际异常样本中被模型预测为异常的样本的百分比, $F1$值则结合了准确率和召回率两者的优势.

3.1.4 实验环境与参数设置

本文实验环境为 Python 3.8、PyTorch 2.0.1、cuda 12.1, 服务器配置为 Intel Xeon Gold 6138 CPU、RTX-4090、128 GB 内存.

对比模型中, iForest 算法使用 Scikit-Learn[38] 库中的 IsolationForest 进行实现, DIF, LODA, VAE, beta-VAE, ECOD 和 COPOD 采取 PyOD[39] 工具箱的官方实现. 其中, 训练好的 iForest, DIF, LODA, ECOD 和 COPOD 模型构成 DELTA 模型的基线模型组, 参与后续的训练. RLMSAD 基于官方代码库[40] 实现, 所有算法的超参数均与论文或官方代码库中的默认配置保持一致.

为公平比较 2 种基于强化学习的模型, 除部分参数外(将在后续实验部分进行讨论), 本文提出的 DELTA 模型参数与 RLMSAD 模型保持一致, 使用的基准强化学习算法均来自于 Stable-baselines3[41](一个通用的标准强化学习算法库), DELTA 模型的相关参数设置如表 1 所示, 其中滑动窗口值过大或过小均会影响模型最终性能, 为保障性能和对比公平性, 参数值参考文献 [27] 的设置. 同时为了减少误差, 以下的所有实验中强化学习模型都随机初始化运行 5 次, 统计评估指标的平均值和标准差结果.

Table 1 List of the Major Parameters

表 1 主要参数列表

参数	取值
PPO 学习率	3E–4
A2C 学习率	7E–4
折扣因子γ	0.99
裁剪因子ε	0.2
批次大小	64
最大仿真步数	2 048
参数更新轮数	10
梯度裁剪最大值	0.5
滑动窗口值	12

此外, 根据使用的深度强化学习基础算法不同, 本文基于 PPO 算法与 A2C 算法分别提出了 2 种基于深度强化学习和数据增强的异常检测方法 DELTA-P 和 DELTA-A.

3.2 实验结果分析

3.2.1 总体表现

首先展示本文方法与其他对比模型在多变量时间序列数据集 SWaT 上的总体异常检测性能, 结果如表 2 和图 4 所示.

Table 2 Overall Performance Results

表 2 总体性能结果 %

模型类型	模型	准确率	召回率	$F1$ 值
无监督学习	iForest	67.30	62.79	64.97
	DIF	69.23	59.55	67.66
	LODA	66.62	63.47	65.00
	VAE	66.55	63.40	64.94
	beta-VAE	67.53	64.33	65.89
	ECOD	71.42	68.04	69.69
	COPOD	71.62	68.23	69.88
强化学习	RLMSAD	71.75±0.39	**68.31±0.06**	69.99±0.19
	DELTA-A (本文)	**88.84±4.93**	63.07±3.51	**73.57±1.72**
	DELTA-P (本文)	88.40±2.64	46.68±6.34	60.88±5.72

注: 黑体数值表示最优结果, "±"后的值表示多次实验的标准差.

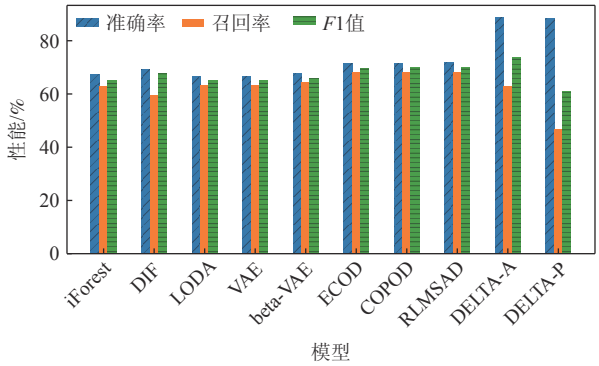


Fig. 4 Overall performance results

图 4 总体性能结果

由表 2 和图 4 结果, 可以观察到 3 个结论:

1)不同类型模型之间的对比. 基于强化学习的异常检测模型总体性能明显优于基于无监督学习的异常检测模型, 这得益于前者使用了强化学习算法对不同时间步数据点进行了不同的异常检测模型筛选, 进而提升了检测准确率.

2)不同强化学习模型筛选框架之间的对比. 本文提出的模型 DELTA-A 相较于同样基于强化学习的 RLMSAD 模型在准确率和 $F1$ 值上有明显的提升，但在召回率上有一定程度的下降，这得益于 DELTA 在不同时间步数据点进行异常检测模型选择时结合了更多数量的异常检测模型，提高了检测准确率，但同时会导致对部分异常过于谨慎，从而降低了召回率.

3)不同强化学习算法之间的对比. 基于 A2C 算法的 DELTA-A 模型在准确率、召回率和 $F1$ 值上都优于基于 PPO 算法的 DELTA-P 模型，从召回率上比较大的差距可以得出，这可能是由于 PPO 的多次迭代更新会导致训练过程稳定性较差，相反在模型筛选问题上 A2C 则更容易收敛到更好的结果.

3.2.2 数据增强对方法性能的影响

本节探究 DELTA 中数据增强模块对异常检测性能的作用，该实验分为 2 部分：数据增强模块对基准模型的影响实验，以及数据增强模块对最终的 DELTA 模型性能的影响实验.

首先，利用数据增强模块对不同类型的异常检测基准模型进行处理，以探究所提出数据增强模块的通用性与有效性. 实验结果如表 3 所示.

Table 3 Impact of Data Augmentation Module on the Performance of Baseline Model

表 3 数据增强模块对基准模型性能的影响 %

模型	准确率	召回率	$F1$ 值
iForest w/o DA	67.30	62.79	64.97
iForest	**69.69**	**66.39**	**68.00**
DIF w/o DA	69.23	59.55	67.66
DIF	**69.70**	**66.40**	**68.01**
ECOD w/o DA	71.42	68.04	69.69
ECOD	**72.22**	**68.81**	**70.47**
COPOD w/o DA	71.62	68.23	69.88
COPOD	**72.03**	**68.62**	**70.28**

注：黑体数值表示最优结果.

可以发现，在 4 种无监督异常检测模型上利用所提出的数据增强模块进行数据增强处理后，所有模型的性能都有不同程度的提高. 其中，基于树形结构的 iForest 和 DIF 经过数据增强后得到的数据可以更好地覆盖数据空间，有助于构建更加全面和稳健的隔离树模型；而 ECOD 和 COPOD 算法的性能提升相对较小，其原因是此类算法通过对数据分布进行建模，更关注于捕捉原始数据的整体分布特征，而对增强数据的依赖较小.

再利用数据增强处理后的基准模型作为用于 DELTA 进行模型选择的基准模型，得到最终数据增强模块对于 DELTA 模型性能的影响对比结果，结果如表 4 所示.

Table 4 Impact of Data Augmentation Module on the Performance of DELTA Model

表 4 数据增强模块对 DELTA 模型性能的影响 %

模型	准确率	召回率	$F1$ 值
DELTA-A	**88.84±4.93**	**63.07±3.51**	**73.57±1.72**
DELTA-A w/o DA	87.44±7.34	55.04±18.75	64.51±20.27
DELTA-P	**88.40±2.64**	**46.68±6.34**	**60.88±5.72**
DELTA-P w/o DA	85.30±3.50	44.15±14.28	56.93±13.62

注：黑体数值表示最优结果，“±”后的值表示多次实验的标准差.

由表 4 可以发现，基于 A2C 算法的 DELTA-A 模型与基于 PPO 算法下的 DELTA-P 模型，在添加数据增强模块后都会使模型的异常检测性能有明显的提升，这得益于数据增强模块对于时序数据集增加样本多样性、提高模型鲁棒性上的贡献，证明了数据增强模块对于 DELTA 的重要性.

3.2.3 奖励函数设置对方法性能的影响

本节探究了强化学习算法中不同奖励函数的设置对于算法的影响，以 DELTA-A 为例，主要针对 FP 和 FN 两个部分的奖励函数进行对比实验，不同值对实验性能影响的结果如图 5 和图 6 所示.

由图 5 可以发现，当 FP 的惩罚值增大时，检测结果召回率下降，准确率普遍上升. 这是由于加大对 FP 的惩罚，会使模型在预测时更加保守，即模型更倾向于将待测样本预测为正常，而只会在对预测结果足够有信心时才会报告异常，从而使得预测准确率上升，但同时这会导致模型对于一些较小的异常情况过于谨慎，使得对异常情况的覆盖会变小，从而

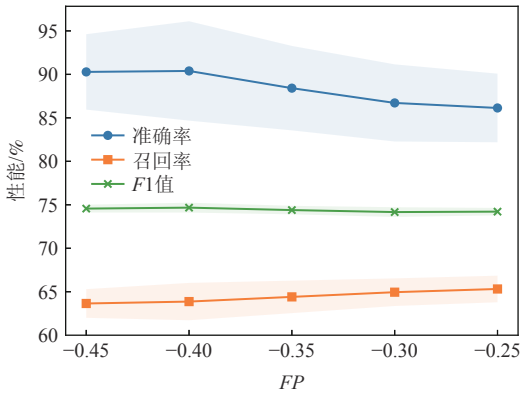


Fig. 5 Impact of different FP on DELTA model's performance

图 5 不同 FP 对 DELTA 模型性能的影响

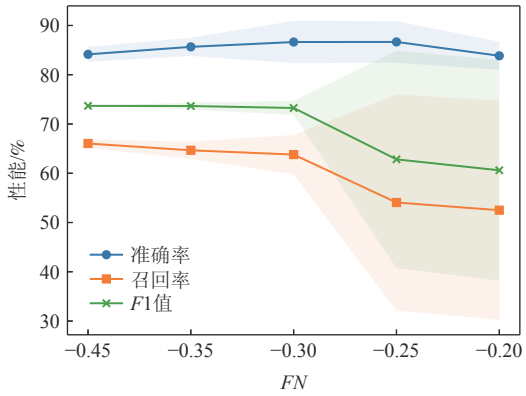


Fig. 6 Impact of different *FN* on DELTA model's performance

图 6 不同 *FN* 对 DELTA 模型性能的影响

使得召回率降低.

由图 6 可以发现, 当 *FN* 的惩罚值增大时, 检测结果召回率上升, 准确率普遍下降. 这是由于加大对 *FN* 的惩罚, 错过实际异常的成本会变得很高, 从而会使模型在预测时更加激进, 即模型更倾向于将待测样本预测为异常, 以避免错过大量的真实异常样本, 进而由于覆盖到更多的异常样本, 预测召回率随之上升, 但同时这也会导致预测准确率的下降.

3.2.4 超参数研究

本节探究了模型中的一个关键超参数——阈值 *d* 对模型最终性能的影响, 以 DELTA-A 为例, 不同阈值 *d* 对实验性能影响的结果如图 7 所示.

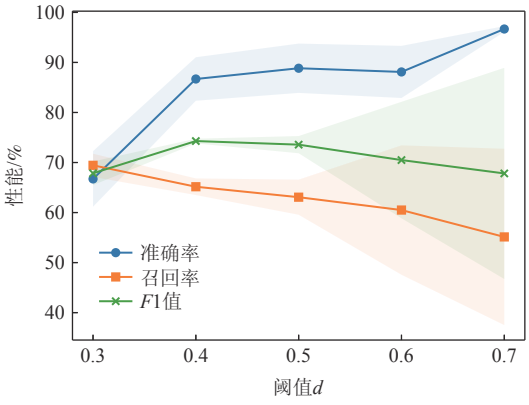


Fig. 7 Impact of threshold *d* on the performance of DELTA model

图 7 阈值 *d* 对 DELTA 模型性能的影响

由图 7 可以发现, 随着阈值 *d* 的上升, 异常检测结果的准确率会上升, 召回率会下降. 这是由于当阈值 *d* 越高, 模型对于作为基线的异常检测模型的需求就越高, 即只有当更多数量的基线异常检测模型均将当前时间步样本点预测为异常时, 模型最终的预测输出结果才为异常, 因此模型最终预测的准确率会提升, 同时由于模型对异常样本的判定更为严格, 使得轻微的异常很难被所有类型的基准模型判定为异常, 进而使得模型最终预测的召回率降低.

3.2.5 模型效率研究

本节通过对比不同异常检测模型的训练时间与推理时间来探究模型之间的效率差异, 由于相较于传统方法, 深度学习类方法(VAE, beta-VAE)训练代价十分大, 故在模型效率研究实验中只以效率更高的传统方法作为对比方法进行比较. 以 DELTA-A 为例, 与其他异常检测模型的时间消耗对比实验结果如图 8 所示.

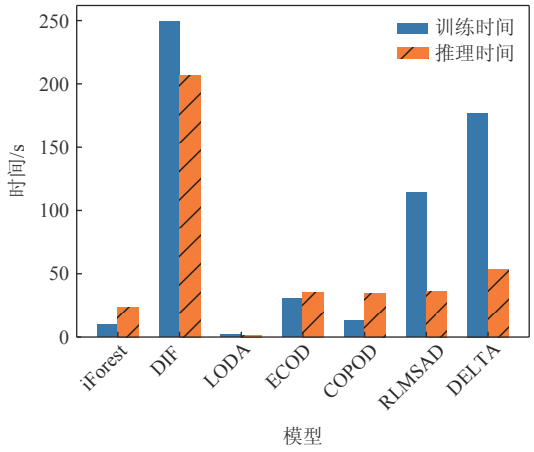


Fig. 8 Comparison of training time and inference time of different models

图 8 不同模型的训练时间和推理时间对比

由图 8 可以发现, 强化学习类的算法模型(RLMSAD 和 DELTA)由于需要进行额外的模型筛选步骤, 比传统的算法模型要花费更多的时间进行模型的训练和推理, 但换来了十分明显的性能提升. 而相比于 RLMSAD 仅支持单维动作空间环境, DELTA 由于需要在多维动作空间中进行筛选, 会付出更大的时间代价, 但在可接受的范围内(少于基于集成学习的 DIF 的时间花费)同样可以获得更高的性能(准确率提高 17.09%, *F*1值提高 3.58%).

4 结 论

本文提出了一种基于深度强化学习和数据增强的工控网络威胁检测模型选择方法 DELTA, 该方法首先使用一种时序数据集数据增强选择方案来选取一组合适的数据增强策略, 进而对时序数据集进行数据增强, 再利用增强数据集对多种类型的异常检测基准模型进行训练, 最后利用深度强化学习算法

在不同时间步数据点下选择利用合适的基准模型进行组合预测. 与其他异常检测模型相比, DELTA 在可接受程度的额外时间成本消耗下, 在准确率和 $F1$ 值上有明显的提高, 在时间序列异常检测问题上展现出更大的潜力.

此外, DELTA 仍有可以改进的地方. 为降低算法复杂度, 本文所使用的数据增强集是基于传统的数据增强策略所组成的, 然而近年来基于对抗生成网络进行数据增强的思路得到了广泛应用, 未来可以尝试使用对抗生成网络等先进技术进行时序数据增强的尝试. 此外, 如何设计高效的数据增强选择方案和合适的强化学习算法以提升模型效率成为未来值得探索的方向.

作者贡献声明: 王泽鹏负责设计方案、完成实验及撰写论文; 马超对论文模型方法提出指导意见; 张壮壮和吴黎兵对论文进行修改和完善; 石小川提出选题和研究内容.

参 考 文 献

[1] Wu H S. A survey of research on anomaly detection for time series[C]//Proc of 2016 13th Int Computer Conf on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). Piscataway, NJ: IEEE, 2016: 426−431

[2] Chalapathy R, Chawla S. Deep learning for anomaly detection: A survey[J]. arXiv preprint, arXiv: 1901.03407, 2019

[3] Xu Lijuan, Wang Bailing, Yang Meihong, et al. Multi-mode attack detection and evaluation of abnormal states for industrial control network[J]. Journal of Computer Research and Development, 2021, 58(11): 2333−2349 (in Chinese)
(徐丽娟, 王佰玲, 杨美红, 等. 工业控制网络多模式攻击检测及异常状态评估方法[J]. 计算机研究与发展, 2021, 58(11): 2333−2349)

[4] Xi Liang, Wang Yong, Zhang Fengbin. Anomaly detection algorithm based on FCM with adaptive artificial Fish-Swarm[J]. Journal of Computer Research and Development, 2019, 56(5): 1048−1059 (in Chinese)
(席亮, 王勇, 张凤斌. 基于自适应人工鱼群 FCM 的异常检测算法[J]. 计算机研究与发展, 2019, 56(5): 1048−1059)

[5] Chen Bofeng, Li Jingdong, Lu Xingjian, et al. Survey of deep learning based graph anomaly detection methods[J]. Journal of Computer Research and Development, 2021, 58(7): 1436−1455 (in Chinese)
(陈波冯, 李靖东, 卢兴见, 等. 基于深度学习的图异常检测技术综述[J]. 计算机研究与发展, 2021, 58(7): 1436−1455)

[6] Audibert J, Michiardi P, Guyard F, et al. USAD: Unsupervised anomaly detection on multivariate time series[C]//Proc of the 26th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2020: 3395−3404

[7] Zhang Chuxu, Song Dongjin, Chen Yuncong, et al. A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data[C]//Proc of the AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2019, 33(1): 1409−1416

[8] Park D, Hoshi Y, Kemp C C. A multimodal anomaly detector for robot-assisted feeding using an LSTM-based variational autoencoder[J]. IEEE Robotics and Automation Letters, 2018, 3(3): 1544−1551

[9] Su Ya, Zhao Youjian, Niu Chenhao, et al. Robust anomaly detection for multivariate time series through stochastic recurrent neural network[C]//Proc of the 25th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2019: 2828−2837

[10] Xin Ruyue, Liu Hongyun, Chen Peng, et al. Robust and accurate performance anomaly detection and prediction for cloud applications: A novel ensemble learning-based framework[J]. Journal of Cloud Computing, 2023, 12(1): 1−16

[11] Zhao Yubo, Guo Ni, Chen Wei, et al. Multi-step ahead forecasting for electric power load using an ensemble model[J]. Expert Systems with Applications, 2023, 211: 118649

[12] Aggarwal C C, Sathe S. Theoretical foundations and algorithms for outlier ensembles[J]. ACM Sigkdd Explorations Newsletter, 2015, 17(1): 24−47

[13] Tama B A, Nkenyereye L, Islam S M R, et al. An enhanced anomaly detection in web traffic using a stack of classifier ensemble[J]. IEEE Access, 2020, 8: 24120−24134

[14] Adeyemo V E, Abdullah A, JhanJhi N Z, et al. Ensemble and deep-learning methods for two-class and multi-attack anomaly intrusion detection: An empirical study[J]. International Journal of Advanced Computer Science and Applications, 2019, 10(9): 520−528

[15] Ünlü R, Xanthopoulos P. A weighted framework for unsupervised ensemble learning based on internal quality measures[J]. Annals of Operations Research, 2019, 276: 229−247

[16] Wen Qingsong, Sun Liang, Yang Fan, et al. Time series data augmentation for deep learning: A survey[C]//Proc of the 30th Int Joint Conf on Artificial Intelligence. Freiburg, German: IJCAI, 2021: 4653−4660

[17] Gao Jingkun, Song Xiaomin, Wen Qingsong, et al. Robusttad: Robust time series anomaly detection via decomposition and convolutional neural networks[C]//Proc of ACM SIGKDD Workshop on Mining and Learning from Time Series (KDD-MiLeTS 2020). New York: ACM, 2020: 1−9

[18] Steven Eyobu O, Han D S. Feature representation and data augmentation for human activity classification based on wearable IMU sensor data using a deep LSTM neural network[J]. Sensors, 2018, 18(9): 2892−2917

[19] Park D S, Chan W, Zhang Yu, et al. SpecAugment: A simple data augmentation method for automatic speech recognition[C]. Proc of. Interspeech New York: ACM, 2019: 2613−2617

[20] Park D S, Chan W, Zhang Yu, et al. RobustSTL: A robust seasonal-trend decomposition algorithm for long time series[C]//Proc of the AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2019, 33(1): 5409−5416

[21] Li Yan, Lu Xinjiang, Wang Yaqing, et al. Generative time series

forecasting with diffusion, denoise, and disentanglement[J]. Advances in Neural Information Processing Systems, 2022, 35: 23009−23022

[22] Kang Yanfei, Hyndman R J, Li Feng. GRATIS: GeneRAting time series with diverse and controllable characteristics[J]. Statistical Analysis and Data Mining: The ASA Data Science Journal, 2020, 13(4): 354−376

[23] Devries T, Taylor G W. Dataset augmentation in feature space[J]. arXiv preprint, arXiv: 1702.05538, 2017

[24] Cheung T H, Yeung D Y. Modals: Modality-agnostic automated data augmentation in the latent space[C]//Proc of the 9th Int Conf on Learning Representations. Washington, DC: ICLR, 2021: 1−18

[25] Yoon J, Jarrett D, Van der Schaar M. Time-series generative adversarial networks[J]. Advances in Neural Information Processing Systems, 2019, 32: 5508−5518

[26] Fons E, Dawson P, Zeng Xiaojun, et al. Adaptive weighting scheme for automatic time-series data augmentation[J]. arXiv preprint, arXiv: 2102.08310, 2021

[27] Zhang J E, Wu Di, Boulet B. Time series anomaly detection via reinforcement learning-based model selection[C]//Proc of 2022 IEEE Canadian Conf on Electrical and Computer Engineering (CCECE). Piscataway, NJ: IEEE, 2022: 193−199

[28] Sutton R S, Barto A G. Reinforcement Learning: An Introduction[M]. Cambridge, MA: MIT press, 2018

[29] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[J]. arXiv preprint, arXiv: 1707.06347, 2017

[30] Goh J, Adepu S, Junejo K N, et al. A dataset to support research in the design of secure water treatment systems[C]//Proc of the 11th Int Conf on Critical Information Infrastructures Security (CRITIS 2016). Berlin: Springer, 2016: 88−99

[31] Liu F T, Ting Kaiming, Zhou Zhihua. Isolation forest[C]//Proc of 2008 8th IEEE Int Conf on Data Mining. Piscataway, NJ: IEEE, 2008: 413−422

[32] Xu Hongzuo, Pang Guansong, Wang Yijie, et al. Deep isolation forest for anomaly detection[J]. IEEE Transactions on Knowledge and Data Engineering, 2023: 12591−12604

[33] Pevný T. Loda: Lightweight on-line detector of anomalies[J]. Machine Learning, 2016, 102: 275−304

[34] Kingma D P, Welling M. Auto-encoding variational Bayes[J]. arXiv preprint, arXiv: 1312.6114, 2013

[35] Higgins I, Matthey L, Pal A, et al. Beta-vae: Learning basic visual concepts with a constrained variational framework[C]//Proc of the 5th Int Conf on Learning Representations. Washington, DC: ICLR, 2017: 1−22

[36] Li Zheng, Zhao Yue, Hu Xiyang, et al. Ecod: Unsupervised outlier detection using empirical cumulative distribution functions[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 35(12): 12181−12193

[37] Li Zheng, Zhao Yue, Botta N, et al. COPOD: Copula-based outlier detection[C]//Proc of 2020 IEEE Int Conf on Data Mining (ICDM). Piscataway, NJ: IEEE, 2020: 1118−1123

[38] Scikit-learn. Scikit-Learn[EB/OL]. 2024 [2024-04-02]. https://github.com/scikit-learn/scikit-learn

[39] Zhao Yue. PyOD[EB/OL]. 2024 [2024-04-02]. https://github.com/yzhao062/pyod

[40] Zhang J E. RLMSAD[EB/OL]. 2022 [2024-04-02]. https://github.com/elisejiuqizhang/RLMSAD

[41] Raffin A, Hill A, Gleave A, et al. Stable-baselines3: Reliable reinforcement learning implementations[J]. Journal of Machine Learning Research, 2021, 22(268): 1−8



Wang Zepeng, born in 1999. PhD candidate. Student member of CCF. His main research interests include reinforcement learning, anomaly detection, and big data analysis.

王泽鹏, 1999 年生. 博士研究生. CCF 学生会员. 主要研究方向为强化学习、异常检测、大数据分析.



Ma Chao, born in 1982. PhD, lecturer. Member of CCF. His main research interests include time series analytics, generative AI, explainable AI, and reinforcement learning.

马　超, 1982 年生. 博士，讲师. CCF 会员. 主要研究方向为时间序列分析、生成式人工智能、可解释性人工智能、强化学习.



Zhang Zhuangzhuang, born in 1994. PhD. Student member of CCF. His main research interests include federated learning security and IoV security.

张壮壮, 1994 年生. 博士. CCF 学生会员. 主要研究方向为联邦学习安全、车联网安全.



Wu Libing, born in 1972. PhD, professor, PhD supervisor. Distinguished member of CCF. His main research interests include network security, Internet of things, machine learning, and data security.

吴黎兵, 1972 年生. 博士，教授，博士生导师. CCF 杰出会员. 主要研究方向为网络安全、物联网、机器学习、数据安全.



Shi Xiaochuan, born in 1984. PhD, associate professor, PhD supervisor. Member of CCF. His main research interests include big data analysis, reinforcement learning, deep learning, and wireless sensor networks.

石小川, 1984 年生. 博士，副教授，博士生导师. CCF 会员. 主要研究方向为大数据分析、强化学习、深度学习、无线传感器网络.