面向工业传感网络的时间序列异常检测综述

吴越，曹国彦
（西北工业大学网络空间安全学院，陕西 西安 710072）

摘　要：随着工业控制系统与信息网络的深度融合，工业关键基础设备的网络化、智能化成为未来工业发展的趋势。工业传感网络作为工业系统网络化的重要组成部分，其数据安全已成为被工业安全领域广泛关注。工业传感网络数据异常影响工业控制系统的物理安全、信息安全和网络安全。工业传感网络异常检测是面向网络攻击及物理故障，通过对复杂、多层次、多尺度的传感时间序列分析，发现隐蔽的异常逻辑及故障原因的方法。总结了工业传感网络异常的成因，系统地综述了工业传感网络异常检测的研究进展，从时序特征、时空多尺度及非结构图表征3个视角，对工业传感网络异常检测的关键技术及典型方法进行分类阐述，分析现有各类方法的发展脉络及主要突破。介绍了用于工业传感网络的数据集和评价指标，及方法的检测效果，并通过对比这些方法的实验结果，说明了各方法的特点及技术侧重，给出了现有工作的应用前景，梳理出当前异常检测方法在实际应用中所面临的挑战。最后提出了工业传感网络异常检测发展趋势及未来的研究方向。

关键词：工业控制系统；传感网络；时间序列；异常检测
中图分类号：TP309
文献标志码：A
DOI：10.11959/j.issn.2096−109x.2024050

Survey of time series anomaly detection for industrial sensor networks

WU Yue, CAO Guoyan
School of Cybersecurity, Northwestern Polytechnical University, Xi’an 710072, China

Abstract: The deep integration of industrial control systems and information networks drives the trend towards networking and intelligence in future industrial development. Industrial sensor networks, crucial for industrial system networking, raise concerns in industrial security, particularly regarding data security. Anomalies in industrial sensor network data impact the physical, information, and network security of industrial control systems. Industrial sensor network anomaly detection, addressing network attacks and physical faults, involves analyzing complex, multi-

收稿日期：2024−01−19；**修回日期：**2024−07−28
通信作者：曹国彦，guoyan.cao@nwpu.edu.cn
基金项目：国家自然科学基金（61803303）；核高基国家重大专项（2017ZX01030-2021）；航空科学基金（20182D53045）
Foundation Items: The National Natural Science Foundution of China (61803303), The National Science and Technology Major Project of the Nuclear HighTech Bases of China (2017ZX01030-2021), The Aeronautical Science Foundation of China (20182D53045)

引用格式：吴越, 曹国彦. 面向工业传感网络的时间序列异常检测综述[J]. 网络与信息安全学报, 2024, 10(4): 17-36.
Citation Format: WU Y, CAO G Y. a Survey of time series anomaly detection for industrial sensor networks[J]. Chinese Journal of Network and Information Security, 2024, 10(4): 17-36.

layered, and multi-scale sensor time series data to discover hidden anomalous logic and fault causes. The causes of anomalies in industrial sensor networks were summarized, research progress in industrial sensor network anomaly detection was reviewed systematically, and key technologies and typical methods were explained categorically from three perspectives: time series features, spatiotemporal multiscale, and non-structured graph representation. The developmental trajectories and major breakthroughs of various existing methods were analyzed and consolidated. Datasets and evaluation metrics currently used for industrial sensor networks were introduced, the detection performance of existing methods was summarized, and through comparative analysis of experimental results, the characteristics and technical focuses of each method were highlighted. The application prospects of existing work were pointed out and the challenges faced by current anomaly detection methods in practical applications were outlined. Future development trends and research directions for industrial sensor network anomaly detection were suggested.

Keywords: industrial control system, sensor network, time series, anomaly detection

0　引言

工业控制系统（industrial control system，ICS）是工业领域控制系统的统称，用于监控、控制和优化工业生产过程，是现代工业生产过程自动化运行不可或缺的工具。为了实时监测工业流程的安全状态，ICS在关键基础设施中部署了大量的传感设备。这些传感设备形成的传感网络具有数据采集、状态感知、信息传输以及安全监测等功能，在工业安全生产中发挥着至关重要的作用。如图1所示，无论是在航天航空、能源化工等国家关键工业基础设施，还是在智能制造、工程生产等企业工程领域，工业传感网络均得到了广泛应用。

工业传感网络是一种专门设计用于工业环境的传感器网络，相较一般的传感网络，工业传感网络具有以下特点。

（1）规模大且复杂

一个工业流程中通常涉及大量且不同种类的传感设备，例如在水处理系统中，涉及液位监测、流量监测、水质监测等各种功能的传感设备，导致了工业传感网络的复杂性。

（2）实时性要求高

工业传感网络需要及时采集和传输数据，以确保对生产过程的快速响应。

（3）安全容错低

工业系统涉及重要的生产和商业信息，因此安全性是工业传感网络的一个重要考虑因素，包

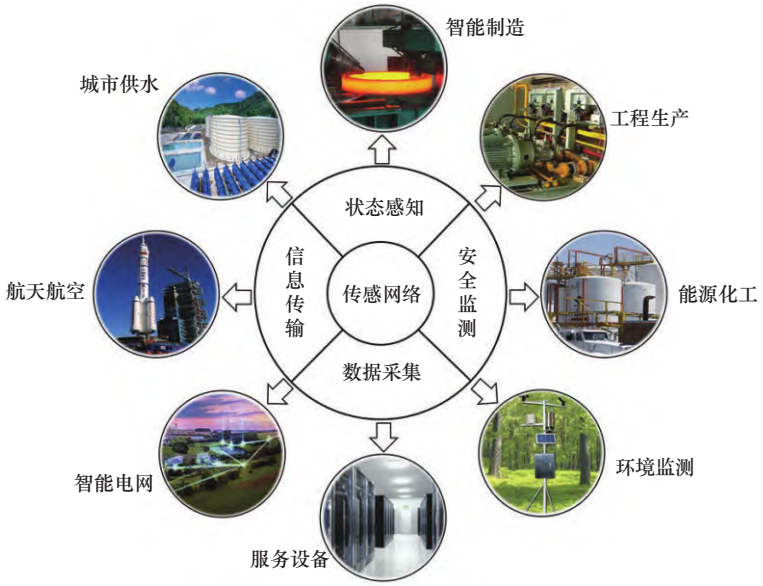

图1　传感网络的特点及应用领域

Figure 1　Characteristics and application areas of sensor networks

括数据的保密性、完整性和可用性等方面的安全性要求。

随着ICS与网络信息技术的深度融合以及工业产业智能网络化的迅速发展，ICS传感网络具有互联程度更高、开放性更强、自主性更智能的特点，这使ICS面临更大的信息安全挑战，尤其是关键工业基础设施系统面临的安全问题层出不穷。例如：2010年，伊朗核电站被Stuxnet病毒入侵，离心机由于铀浓缩过程中的异常动作而损坏[1]；2020年，美国天然气管道遭受攻击，整个天然气管道运营被迫停摆持续两天；2021年，美国佛罗里达州水处理工厂遭受黑客攻击，试图将氢氧化钠的浓度提高到极其危险的水平[2]；2021年伊朗遭受恶意软件Meteor的网络攻击，导致全国加油站的加油系统中断[3]；2022年德国风电整机制造商巨头Enercon遭受网络攻击，欧洲卫星通信大规模中断。ICS传感网络的安全不仅关系到工业生产能否正常进行，更涉及人身财产安全，甚至是国家安全。

工业传感网络采集的传感数据通常记录在ICS的监视控制与数据采集（supervisory control and data acquisition，SCADA）系统中[4]，这些数据可以被建模为多变量时间序列，根据这些数据实现工业传感网络的安全检测已成为ICS安全领域的热门研究内容。时间序列分析作为时序数据异常检测最常用的方法，被广泛用于工业传感网络的异常检测，通过从传感数据中提取有意义的特征，从而诊断过去的行为或者预测未来的趋势[5]。传统的异常检测方法通过采用多种数学和统计模型对时间序列进行分析，包括基于线性模型的方法[6]、基于距离的方法[7]、基于密度的方法[8]等。尽管这些方法具有一定的异常检测能力，但随着ICS设备数据复杂性的增加，这些传统方法在处理多变量时间序列时表现出一定的局限性。为了更有效地解决多变量时间序列的异常检测问题，越来越多的研究开始将机器学习方法引入时间序列分析领域，例如支持向量机[9]、ARIMA[10]等方法。与传统的方法相比，这些方法在异常检测方面均取得了显著的检测效果[11-12]。

随着人工智能技术与算力水平的提升，深度学习无论是在检测精度还是在数据处理能力上，都展示了较好的应用前景。不断涌现的深度学习方法，例如循环神经网络（recurrent neural network，RNN）[13]、卷积神经网络（convolutional neural network，CNN）[14]、图神经网络（graph neural network，GNN）[15]以及Transformer[16]等，在面向工业传感网络的多变量时间序列时，取得了较好的异常检测效果。这些方法在提升检测性能的同时，也使表征互联传感设备的内在耦合关系成为可能。

现有可供工业传感网络安全性研究参考的综述主要集中在ICS安全以及时间序列异常检测两个方面。在ICS安全方面，Umer等[17]对入侵检测进行了多维度的分类，并分别梳理了面向工业控制系统的入侵检测和异常检测的机器学习与强化学习方法。Koay等[18]从实验流程的角度梳理了ICS中的漏洞，并介绍与机器学习方法相关的知识。唐士杰等[19]详细讨论了ICS主流体系结构及相关组件的安全问题，从可用性、完整性和机密性的角度对攻击进行分类。杨婷等[4]从攻击和防御的角度对多个安全研究领域进行分析和综述。在时间序列异常检测方面，Choi等[5]分析了时间序列数据的性质和异常类型，从上下文建模等多个角度整理深度学习方法的应用。Darban等[20]重点从模型特性的角度梳理了目前基于深度学习的异常检测方法。Luo等[21]从应用的角度整合了基于深度学习的异常检测方法，并给出模型选择和训练技术的指导原则。孙海丽等[22]根据技术原理全面阐述了工业物联网异常检测方法，分析这些方法的优缺点。这些文献虽然为工业传感网络异常检测的研究提供了重要的参考价值，但没有考虑将时间序列异常检测方法和工业流程特性进行有效结合。

本文对工业传感网络的异常检测技术进行全方面的综述，分析造成工业传感网络异常的原因，主要包括网络攻击和物理故障两个方面。总结目前适用于工业传感网络的异常检测方法，从检测异常的时间和空间两个方面对这些深度学习方法进行阐述，并分析异常检测方法的实际应用能力。最后，本文对目前异常检测方法在未来工业环境中的应用进行了总结和展望。

1　工业传感网络异常成因

工业传感网络主要用于从工业环境中获取关键基础设施的安全状态信息，并将传感数据反馈到SCADA系统中。ICS通常依赖这些获取的数据来评估和执行控制逻辑，因此，一旦传感器网络中的数据发生异常，就有可能导致系统误判，继而影响系统的正常运行。造成时序数据异常的原因可以分为两类：一类是由网络攻击引起的异常；另一类是由系统硬件故障引起的异常。本节将对这些成因及模型进行综述。

1.1　网络攻击

初期，工业领域将关键基础设备运行于隔离内网中，未充分考虑网络攻击可能带来的安全问题。然而，随着工业基础设施不断地暴露在互联网络中，由工控系统设备互联网络引发的安全事件呈现日益增多的趋势，引起了广泛关注。工业传感网络中的网络攻击主要分为3种类型，分别是拒绝服务（denial of service，DoS）攻击，虚假数据注入（false data injection，FDI）攻击以及被动攻击。表1给出了造成异常的网络攻击模型。

1.1.1　拒绝服务攻击

DoS攻击主要通过消耗资源的方式来阻止或干扰传感设备之间的通信，属于可用性攻击的一种形式。通常情况下，工业传感网络所面临的DoS攻击针对网络多层结构目标展开，其主要目的在于干扰SCADA系统中的传感设备数据采集，进而影响对工业流程的判断与决策[23]。在无线工业传感网络场景中，由于设备受限于有限的能量预算，DoS攻击成为有效阻碍网络通信的手段[24]。为了最大程度地影响DoS攻击对系统数据估计的效果，Zhang等[25]提出了一种基于线性二次高斯控制成本函数的攻击策略模型。该模型基于相对理想的假设，即攻击者仅能发起有限次数的攻击，且所有数据丢失均可追溯至攻击行为。为了突破这一假设的局限性，Qin等[26, 27]考虑在自然丢包场景下的DoS攻击模型，然而，解决该问题的同时引入了传感设备具有自主计算能力的新假设，而传感设备在真实工业环境中通常不具备自主计算能力。因此，Zhang等[28]设计并提出DoS攻击的最优调度模型，专门考虑设备传感器不具备自主计算能力的情况。

1.1.2　虚假数据注入攻击

FDI攻击代表了最典型的完整性攻击，其主要手段涉及改变或删除传感设备的测量数据，从而导致系统做出错误的控制决策。FDI攻击也被称为欺骗性攻击，攻击者利用网络漏洞截取原本正确的数据并传输经过修改的数据。由于篡改的数据在传感网络中以蔓延的方式传播，所以FDI攻击比DoS攻击更加难以被检测[29]。

FDI攻击主要分为基于时间驱动的模型和基

表1　造成异常的网络攻击模型

Table 1　Network attack models that cause anomalies

相关工作	异常类型	应用领域	简要描述	涉及安全属性		
				可用性	完整性	机密性
Muraleedharan et al[23]	DoS	智能读卡系统	传感器通过蠕虫洞进行伪装	●	○	○
Peng et al[24]	DoS	网络物理系统	干扰传感器和远程间的通信	●	○	○
Zhang et al[25]	DoS	无线网络系统	最大化二次高斯代价函数	●	○	○
Qin et al[26, 27]	DoS	网络物理系统	研究未攻击下的数据丢包	●	○	○
Zhang et al[28]	DoS	网络物理系统	多传感器共享公共通信信道	●	○	○
Wu et al[31]	FDI	网络物理系统	使用时变权重矩阵增加难度	○	●	○
Mousavinejad et al[32]	FDI	网络化控制系统	用传感器测量数据更新预测	○	●	○
Wu et al[33]	FDI	网络物理系统	在智能动态传感中降低质量	○	●	○
Zhang et al[34]	FDI	无人地面飞行器	用Kullback-Leibler约束	○	●	○
Li et al[35]	FDI	智能电网	保守和侵略两种拓扑攻击	○	●	○
Wang et al[37]	被动	水声传感网络	建立能量和位置的消耗模型	○	○	●
Yuan et al[38]	被动	网络物理系统	待反馈随机算法最大化传输	○	○	●

于事件驱动的模型[30]。Wu 等[31]采用连续时间模型对 ICS 建模，设计了两种最佳位置切换策略以实施 FDI 攻击。后来为了摆脱对通信通道理想性的假设，越来越多的攻击模型开始关注攻击的隐蔽性，提高 FDI 的攻击效率[32, 33]。Zhang 等[34]采用 Kullback-Leibler 散度作为最大化效用函数，设计了一个高度隐蔽的最佳欺骗攻击模型。基于事件驱动的策略将 ICS 建模成离散事件模型，着重关注传感设备的攻击事件序列[35, 36]。通过此模型，攻击者能够更有针对性地影响系统的决策过程，使攻击更难以被察觉。因此，对于 FDI 攻击的研究逐渐演变为对系统动态特性和攻击行为隐蔽性的深入探讨，以应对这一类具有高度欺骗性的网络攻击。

1.1.3　被动攻击

被动攻击是指攻击者未经授权，试图侵犯传感数据的保密性、窃取信息以及查看数据的行为，通常这种行为涉及系统的机密性。常见的被动攻击形式主要包括窃听和欺骗，与 DoS 攻击和 FDI 攻击等主动攻击相比，被动攻击更为隐蔽，难以及时被检测。

攻击者通常通过窃听技术来获取传感设备与控制器之间的通信信息，从而迅速而不被察觉地获取敏感数据。为了有效应对这一威胁，研究者已经提出了多种方法，以保护传感网络免受窃听攻击的威胁[37, 38]。近年来，不透明度的概念引起了广泛关注，该概念可有效验证系统的保密性[39-41]。通过引入不透明度，系统能够更好地隐藏敏感信息，增加攻击者获取数据的难度，从而提高系统抵抗被动攻击的能力。欺骗攻击通常也被用来伪造工业关键设施的定位，以非法访问并获取机密资源[42]。这种类型的攻击往往通过伪装成合法用户或设备来误导系统，使攻击者能够绕过安全措施。为了对抗欺骗攻击，需要进一步深入了解攻击者的行为模式，并制定相应的欺骗检测和防御机制，以确保系统的完整性和机密性得到有效的保护。

1.2　物理故障

物理故障通常是由系统的复杂性和设备的异构性造成的，有多种因素会造成传感设备的损坏。值得关注的是，物理设备遭受攻击或自然因素引发的损害通常会导致更严重的后果，且无法通过软件防护等手段避免。物理故障主要分为物理破坏和固件攻击两种形式。

1.2.1　物理破坏

在 ICS 中，传感网络通常呈多级结构，而不同的传感集群往往处于各自独特的工作环境中。在恶劣环境中运行的传感设备通常会受到较大的损耗，甚至可能发生物理性的严重破坏。传感设备的损耗和破坏会导致其采集的数据产生一定的误差，进而影响系统的逻辑决策。例如，航天器上的传感器在遇到环境中的意外情况时可能导致数据的不准确性[43]。相对而言，人为引起的破坏更加显著，传感设备因操作不当或人为破坏而造成的故障与损坏更具有危害性，不易被及时察觉，从而造成传感数据的持续异常。

1.2.2　固件攻击

可编程逻辑控制器（programmable logic controller，PLC）是专门为 ICS 设计的可编程存储器，其主要用途之一是控制工业现场的物理设备[44]。

固件攻击通常作用于 PLC 底层 CPU 的操作系统，攻击者利用固件中的漏洞对传感设备进行攻击，从而控制这些设备。当传感设备接收到错误的逻辑信号时，就会偏离正常的工作状态。错误的操作同样会缩短设备的使用寿命，甚至直接导致设备损坏，从而引发不可预测的后果。在物理层面上，这种攻击形式加剧了系统的脆弱性，对工业控制系统的安全性和稳定性构成潜在威胁。

2　工业传感网络异常检测方法

针对 ICS 传感网络的异常检测方法大多基于数据驱动的时间序列分析方式实现，这些方法通过融合新的技术实现高精度检测的目的。面对工业传感网络异常的时空特性，本节采用时间和空间两个尺度的检测对这些方法分类，分为单时间序列模型、多尺度序列模型以及图结构模型，工业传感网络异常检测方法汇总如图 2 所示。图 3 给出了基于深度学习的异常检测方法发展历程。

2.1　单时间序列模型

单时间序列模型聚焦时间维度上的相关性捕

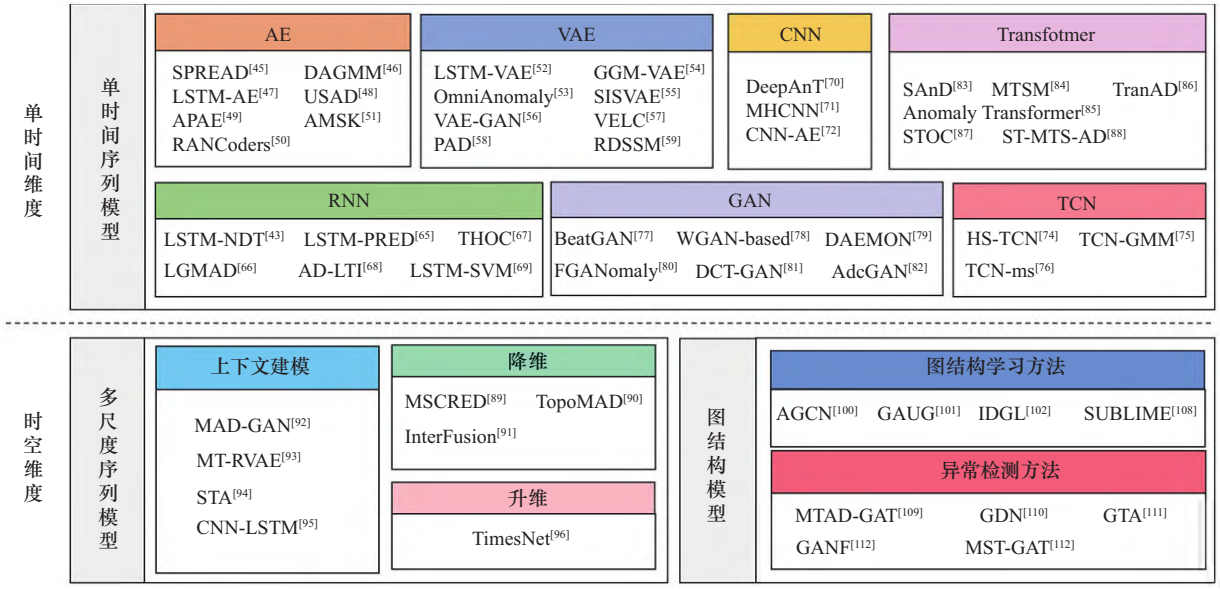

图2　工业传感网络异常检测方法汇总

Figure 2　Summary of anomaly detection methods in industrial sensor networks

捉，通过深度学习的方法捕捉时间上的特征，构建处于正常行为下的时间序列模型，进而基于重构和预测的策略来探索时间序列的异常情况。这种模型可以检测到异常发生的具体时间，但未充分考虑异常在空间上的特征。本节利用多种深度学习方法分别对单时间序列模型进行介绍。

2.1.1　基于自动编码器的方法

自动编码器（autoencoder，AE）是用于数据压缩和降维的神经网络模型，通过编码器和解码器实现数据的压缩和还原。基于AE的异常检测方法主要集中在对时间序列的建模上，使用编码器和解码器构建系统的正常运行模型。编码器负责将时间序列数据压缩到低维空间，而解码器将这一低维空间还原至原始的高维空间。其数学原理如下所示：

$$z = e_\theta(x),\quad \hat{x} = d_\delta(z) \tag{1}$$

其中，$e_\theta(\cdot)$表示编码操作，θ为其可设置的超参数。$d_\delta(\cdot)$表示解码操作，其拥有和$e_\theta(\cdot)$相反的神经网络层关系，δ是解码器中的超参数。AE的编码器和解码器通常在同一个损失函数上进行训练，通过设置损失函数来识别异常数据。损失函数一般设置为原始数据和重构数据的均方误差：

$$\text{Loss} = \left\| x - \hat{x} \right\|_2 \tag{2}$$

基于AE的异常检测方法主要集中在对时间序列的建模上，从而构建系统的正常运行模型。SPREAD[45]提出一种基于稀疏循环神经网络的异常检测方法，该方法通过输入层的稀疏前馈连接层实现逐点降维，将编码器与循环神经网络相结合，在端到端的学习环境中对系统的正常行为进行建模。DAGMM[46]使用深度自编码高斯混合模型实现异常检测，该模型使用深度AE为每个输入数据生成低维表示和重建误差，并馈送到高斯混合模型中，利用单独的估计网络促进混合模型的参数学习。LSTM-AE[47]基于LSTM开发了无监督的自动编码器模型，通过滑动窗口实现实时的异常检测，可以确定发生异常的时间间隔。USAD[48]是基于AE的多元时间序列无监督异常检测方法，其利用编码器–解码器架构的对抗性训练突破了这两种技术的局限性。APAE[49]提出了一种近似投影自动编码器，该方法最先研究了异常检测自动编码器对不同类型对抗性攻击的脆弱性，提高了模型在对抗性攻击影响下的鲁棒性。RANCoders[50]利用傅里叶变换将AE低维空间的时域信息转为频域信息，再将这些频域信息作为网络层中的先验知识，实现学习多变量输入的同步表示的目的。AMSL[51]引入自监督学习模块提高了无监督异常检测的泛化能力。通过自适应记忆融合网络，分别通过全局和局部记忆模块学习共同和特定的特征，基于卷积自动编码器框架以端到端方式进行训练。

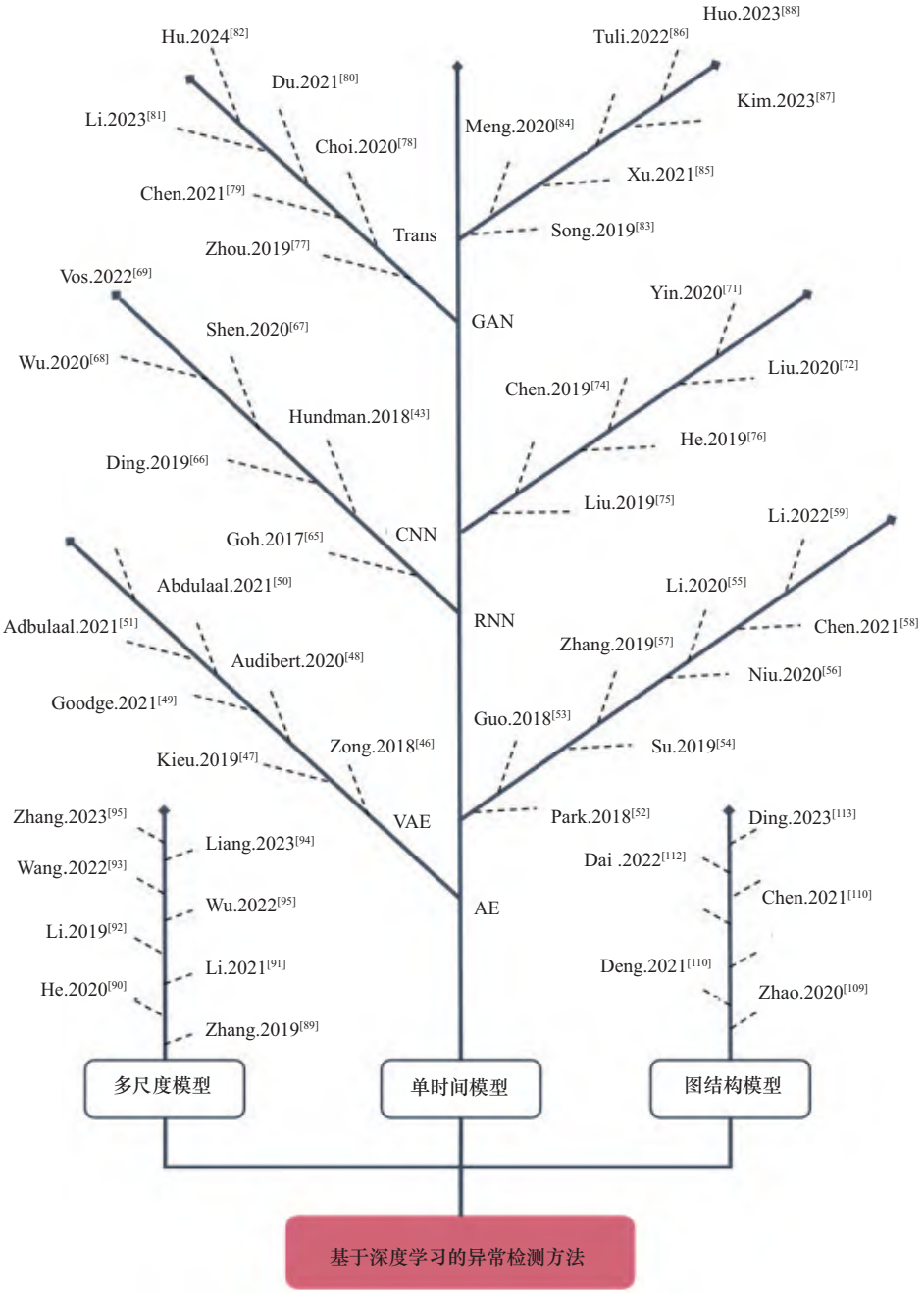

图3　基于深度学习的异常检测方法发展历程

Figure 3　The development history of anomaly detection methods based on deep learning

2.1.2　基于变分自动编码器的方法

变分自编码器（variational autoencoder，VAE）和AE一样都是时间序列的单值映射问题，但是VAE反映的是分布的映射关系，解决了AE中非正则化低维空间的问题，给编码器添加了合适的噪声，具有更好的泛化能力。VAE不再简单地将编码器输出作为解码器输入，而是从标准正态分布中生成类似但不同于原始数据的新样本。VAE共计算两组编码：一组为均值编码$m=\{m_1,m_2,\cdots,m_i\}$；另一组为控制噪声干扰的方差编码$\sigma=\{\sigma_1,\sigma_2,\cdots,\sigma_i\}$。损失函数在除去AE的误差计算之外，还引入新的计算误差方法，其计算公式如下：

$$\text{Loss}=\sum_{i=1}^{n}(e^{\sigma_i}-(1+\sigma_i)+(m_i)^2) \tag{3}$$

LSTM-VAE[52]使用LSTM和VAE融合信号

并重建其预期分布，设计基于重建的异常分数和基于状态的阈值实现异常检测。OmniAnomaly[53]提出一种用于多元时间序列的新型随机循环神经网络，该方法根据组成单变量的时间序列的重构概率为方法提供解释，对于各种设备都具有较好的鲁棒性。GGM-VAE[54]采用门控循环单元（gated recurrent unit，GRU）捕捉时间序列数据之间的相关性，在编码器的低维空间使用高斯混合先验来表征多模态数据，设置重建概率和阈值来实现异常检测。SISVAE[55]使用一种变分正则化器从异常的时间序列中捕获正常模式，通过对生成模型的非频化输出进行惩罚来提高鲁棒性。VAE-GAN[56]基于生成式对抗网络（generative adversarial network，GAN）实现异常检测，该方法联合编码器、生成器和鉴别器寻找从高维空间到低维空间的最佳映射，充分发挥了 VAE 和 GAN 的特点。VELC[57]通过在 VAE 的低维空间中添加约束网络，更好地生成了与训练样本相似的潜在变量。其使用长短期记忆（long short-term memory，LSTM）网络作为 VAE 框架的编码器和解码器，更好地处理了时间序列。PAD[58]使用 LSTM 在 VAE 重建的时间序列中获取纯净输入，使对预测任务的异常和噪声具有鲁棒性，同时联合模型预测器和异常检测器，解决了系统状态预测和异常检测两个问题。RDSSM[59]设计两个转换模块，同时考虑时间依赖性和不确定性，并通过基于密度估计的检测技术应对随时间变化的波动性噪声。

2.1.3　基于循环神经网络的方法

RNN 因其特殊的神经网络结构被广泛地应用于处理序列数据。RNN 相较传统的神经网络，其隐藏层的输入同时包括输入层的输出以及上一时刻隐藏层上的输出，具有一定的记忆能力。为解决 RNN 的长期依赖问题，即梯度消失和梯度爆炸现象。近年来，RNN 的改进算法不断出现，目前 LSTM[60]和 GRU[61]更常用。为了进一步改进这些方法的性能，设计了很多策略对其进行优化[62-64]。

LSTM-NDT[43]针对航天器异常检测描述了使用 LSTM 的高效性能，并保持了整个系统的可解释性。对于生成的预测模型设计了一种非参数、动态的阈值方法评估残差，解决了数据阈值相关的多样性、非平稳性和噪声问题。LSTM-PRED[65]使用 LSTM 对时间序列进行建模，学习时间依赖性，并采用累积和的方式进行异常检测，而不是计算每个传感器的阈值。LGMAD[66]通过 LSTM 模型评估每个单变量传感时间序列的实时异常，并采用高斯混合模型对可能的异常进行多维联合检测。为了提高算法效率，该方法还提出健康因子的概念描述系统的健康水平。但该方法仅考虑低维时间序列的异常检测，没有进一步考虑高维时间序列的异常检测问题。THOC[67]对时间进行了分类，通过使用具有跨层连接的扩张循环神经网络同步多个尺度的时间动态，通过分层聚类过程获得多个超球面，并添加时域中的自我监督任务完成端到端的训练，促进表示学习。AD-LTI[68]利用底层预测模型对局部序列进行预测，对不同预测的集合进行加权，并按时间顺序排列。通过构建帧到序列的 GRU 网络，同时分解每个样本通道的时间序列提供的季节性特征来扩展输入。季节性特征的整合有效地减少了异常样本的负面影响。LSTM-SVM[69]采用 LSTM 架构与支持向量机相结合，将多个传感器测量数据的异常样本与正常样本进行分离。

2.1.4　基于卷积神经网络的方法

CNN 是一类常见的深度神经网络，最常用于计算机视觉图像的工作中。基于其共享权重架构和平移不变性的特征，该方法也被称为移动不变的神经网络。其隐藏层结构可以分为卷积层、池化层、非线性激活层和全连接层。与其他神经网络算法相比，CNN 使用相对较少的预处理。DeepAnT[70]使用深度卷积神经网络构建时间序列预测器，采用时间序列窗口预测下一个时间戳的数据，然后将预测的数据传递给异常检测模块检测异常。借助 CNN 有效参数共享的特性，该方法在相对较小的数据集上也可以完成训练，同时具有良好的泛化能力。CNN-AE[71]使用两阶段滑动窗口学习时间序列更好地表示，然后结合 CNN 和 LSTM 对数据的特征进行提取。其实验结果证明，两阶滑动窗口的设计会给 CNN 提取的特征附加时间依赖性，这有助于 LSTM 提取更多的时间特征。TCQSA[72]将 CNN 与 LSTM 模型

相结合，同时提取时间序列中的整体变化趋势和局部特征，通过注意力机制实现异常的检测。

CNN 在许多任务上的表现优于 RNN，因为其避免了递归网络梯度爆炸等。为了更好地将 CNN 应用于时间序列数据，Bai 等[73]提出时间卷积网络（temporal convolutional network，TCN）架构。该方法利用扩张卷积神经网络适应序列数据，有很多的研究者将 TCN 应用到时间序列异常检测中。HS-TCN[74]设计了一种半监督式的分层堆叠 TCN，采用分层模型考虑了智能家居场景中传感数据的特点，采用堆叠的方法剔除异常值，较大地提升了检测精度。TCN-GMM[75]应用 TCN 提取时间序列的特征，将高维时间序列映射到低维空间，并通过低维空间的特征捕捉实现异常检测。在高斯混合模型的基础上引入贝叶斯推理，以较低的计算复杂度有效地检测异常。TCN-ms[76]在正常序列上训练 TCN 模型，用来预测多个时间步长。使用多元高斯分布拟合预测误差并计算点的异常分数，并提出多尺度融合的方法来提高性能。

2.1.5　基于生成式对抗网络的方法

GAN 主要由两个神经网络构成，生成网络负责生成模拟数据，判别网络负责判断数据源自数据集还是机器生成。通过两个网络之间的对抗，两个模型均能达到较好的效果。生成器可以生成与原始数据非常相似的图像，判别器为了判断正确也会不断训练。BeatGAN[77]使用 GAN 方法执行正则化，使模型可以稳定地重建数据。通过设计一种扭曲方法来利用时间序列特征增强用于训练的数据，实验证明该方法具有很高的稳定性。WGAN-based[78]将多元时间序列转换为二维图像，使用 GAN 的生成器学习一系列图像到下一个距离图像的映射，既可以分析时间序列数据的时间关联，也可以通过卷积滤波分析多变量之间的相关性。DAEMON[79]使用两个判别器对抗性地训练自动编码器以学习多元时间序列的正常模式，然后使用重建误差来检测异常。该方法通过使用对抗生成方法对隐藏变量和重构数据进行正则化，来保证鲁棒性。FGANomaly[80]在训练鉴别器之前用伪标签过滤可能的异常样本，而在训练编码器解码器形式的 GAN 生成器时，设计自适应加权损失替换 MSE 作为训练目标。DCT-GAN[81]设计多个生成器和单个辨别器缓解模式崩溃的问题。每个生成器由一个扩张卷积网络和一个 Transformer 块组成，可以获得时间序列的细粒度和粗粒度信息，有效地提高了泛化能力。AdcGAN[82]采用 Dropout 近似表示模型的不确定性，使用概率分布代替点估计作为预测结果，从观测的差异和模型的不确定性两个方面来衡量异常。

2.1.6　基于 Transformer 的方法

近年来，Transformer 模型被广泛应用于各种领域，比较热门的 BERT 和 GPT 都基于该模型。Transformer 是基于自注意力机制的机器学习模型，通常由编码器和解码器构成，与之前提到的方法不同，其完全依靠注意力机制来实现所有功能。可以捕获输入向量之间多种维度上的相关系数，相较 RNN 可以做到更好的并行训练。SAnD[83]首次利用带掩码的自注意力机制进行时间的建模，采用位置编码和密集插值策略合并时间序列。通过设计多任务变体，以联合推断具有多个诊断任务的模型，证明了 Transformer 在时间序列分析中的有效性。MTSM[84]基于 Transformer 架构建立时间序列模型，使用注意力机制并行更新时间步长，并使用掩码策略提前检测异常。其在航天器数据上的实验结果表明，与 LSTM 对时间序列建模相比，Transformer 大大减少了时间消耗，且精度没有明显下降。Anomaly Transformer[85]使用 Transformer 对时间序列的先验关联和系列进行建模，体现了时间之间的关联差异，并提出极小极大策略放大关联差异的正常与异常可区分性，推导出基于关联的异常检测方法。TranAD[86]使用 Transformer 的序列编码器进行快速推理，通过设计基于焦点分数的自我调节来实现强大的多模态特征提取和对抗训练以获得稳定性，实验表明该方法在检测和诊断方面有较出色的性能。STOC[87]使用多个 Transformer 编码器和一个卷积层解码器，实现同时考虑时间序列的全局趋势和局部变化来检测异常。ST-MTS-AD[88]借用 Transformer 编码器输出的长时依赖特征在隐空间中传播随机变量间的长时依赖性，采用门控转换函数生成随机变量的先验分布，实现随机变量间的非线性转换。

2.2　多尺度序列模型

与单时间序列模型不同，很多方法在考虑时间相关性的同时也考虑变量之间的相关性，本文定义这些方法为多尺度序列模型。这些方法从时间和空间两个维度对时间序列数据进行建模，不仅可以发现异常，而且可以定位发生异常的具体传感设备。这些方法同样也基于 2.1 中提到的神经网络结构，形式主要分为降维方式、上下文建模和升维方式。

2.2.1　降维方式

降维的方式主要基于 AE 和 VAE 实现。例如 MSCRED[89]使用一种多尺度卷积循环编码器–解码器实现异常检测，该方法首先构建多尺度特征矩阵来表征不同时间步长的多级别系统状态，开发基于注意力的卷积 LSTM 捕获时间模式，同时采用卷积编码器构建传感器间相关性的特征图。TopoMAD[90]基于 VAE 实现了对异常数据之间空间和时间依赖性的稳健表示，该方法将图神经网络与 LSTM 相结合作为 VAE 的基本结构，使用随机 VAE 模型以完全无监督的方式进行异常检测。InterFusion[91]采用具有显式低维度量和时间嵌入的分层变分自动编码器共同学习鲁棒的时间序列表示，并基于马尔可夫链蒙特卡洛方法提出一种多元时间序列的新型异常检测方法，对异常进行时间和空间上的度量。

2.2.2　上下文建模方式

上下文建模的方式是通过捕捉时间序列中的上下文关系来实现异常检测。MAD-GAN[92]使用 LSTM 构建生成器和鉴别器，在 GAN 框架中捕获时间序列分布的时间相关性。在数据处理时，不是独立处理每个数据流，而是同时捕获变量之间的潜在相互作用；并提出新的异常分数计算方法 DR-Score 用于异常检测。MT-RVAE[93]改进了 Transformer 的位置编码，使用全局时间编码将时间序列信息和周期信息添加到数据中，更好地捕获序列中的长期依赖性。同时，通过多个尺度的特征融合算法，弥补了数据采样过程中丢失的细节信息，以获得鲁棒性的特征表达。STA[94]采用基于 LSTM 的时间注意力机制自适应地学习序列的时间相关性，同时使用多图注意力网络捕获空间相关性，在时间和空间维度上实现异常检测。CNN-LSTM[95]构建由时间注意力和空间注意力组成的时空注意力模块，通过学习正常样本分布实现异常检测。

2.2.3　升维方式

之前的异常检测方法通常基于一维时间序列建模，即只关注数据随时间维度的变化。一维时间序列通常具有较少的特征，直接在一维时间序列上进行分析面临较大的挑战。为了突破一维时间序列的局限性，TimesNet[96]将时间序列数据转换为二维变量，通过快速傅里叶变换分析时间序列的周期性，提出 TimesBlock 自适应地发现多周期性，并将其多周期的时间序列融合。TimesNet 将时间序列转换为二维数据也为应用 2D 卷积方法提供了条件，其在异常检测等 5 个主流领域达到了先进水平。

2.3　图结构模型

随着将神经网络应用于欧式结构数据的成功，研究人员开始关注非欧氏结构数据。图作为最常见的非欧几里得结构，其良好的应用前景得到了广泛关注。近年来，随着 GNN 的快速发展，更多的研究将 ICS 中的传感网络考虑成图的结构进行建模。该结构在考虑时间相关性的同时，也考虑了传感设备之间的相关性。基于图结构的异常检测方法取得了最为显著的检测效果，这些方法主要使用 GNN 的变体：图卷积神经网络（graph convolutional network，GCN）[97]和图注意力神经网络（graph attention network，GAT）[98]。本节内容主要分为图结构学习以及基于图的异常检测方法两个模块。

2.3.1　图结构学习

ICS 的传感网络结构通常是未知的，图结构学习成为 ICS 使用 GNN 的主要挑战之一，不正确或有噪声的图结构可能会导致 GNN 学习到的表示效果较差。下面整理了图结构学习的方法，但不限于工业控制系统的应用。图结构学习模型通常由 3 个阶段组成：图构建、图结构建模和消息传播[99]。基于度量的方法对每个节点嵌入向量，并采用节点对之间的度量函数来计算节点之间边的权重。例如，AGCN[100]通过计算节点对特征之间的广义马哈拉诺比斯距离来学习图结构，并使用高斯函数来细化结构。GAUG[101]使

用内积来计算边的权重，解决了在正常和噪声环境中推理边操作的增强问题。而 IDGL[102]使用余弦相似度来衡量边的权重，并使用自适应图正则化来控制学习图的质量。除了度量函数方法，很多方法采用神经网络来推断节点表示的边权重。这些方法通常使用神经网络来编码节点嵌入的交互关系，利用多层感知器生成图邻接矩阵[103-105]。也有很多方法直接将邻接矩阵视为可学习的参数，与神经网络的参数一起使用优化器优化[106, 107]。近年来，随着图对比学习的快速发展，SUBLIME[108]采用图对比学习实现了无监督的图结构学习。

2.3.2　基于图的异常检测方法

鉴于 ICS 传感设备分布式的特征，近年来很多方法将传感设备时间序列建模成图结构。图结构的引入不再单独考虑时间上的周期性变化，通过对图关系中边的建模捕捉传感设备之间的相互影响关系。基于图的异常检测方法取得了较为显著的成效，例如，MTAD-GAT[109]使用两个图注意力层对时间序列进行建模，面向特征的图注意力层负责捕捉多个特征之间的因果关系，面向时间的图注意力层强调时间维度的依赖关系。通过联合基于预测的模型和基于重构的模型优化时间序列的表示，使用联合目标函数对模型进行优化。GDN[110]将图结构学习方法与基于时间序列的预测模型相结合，对每个传感器嵌入向量，使用图注意力神经网络学习传感器之间的关系，基于相邻传感器之间的注意力系数来预测传感设备的未来行为。GTA[111]基于 Gumble-Softmax 抽样方法学习传感器之间单向图结构，提出一种名为影响传播卷积的新图卷积网络结构描述节点之间的异常信息流，并使用多分支注意机制改进 Transformer 解决了异常检测中的二次复杂性问题。GANF[112]使用归一化流进行异常检测。通过贝叶斯网络对变量之间的因果关系进行建模，这种结构允许将所有图节点的联合密度分解为每个节点的条件密度的乘积，降低了处理的难度，并使用增广流模型检测多个时间序列数据的异常。MST-GAT[113]使用多通道的时空图注意力网络解决了捕捉单变量时间序列之间时空关系的问题，增强了对检测到的异常的可解释性。

3　实验介绍及结果分析

本节将主要对工业传感网络的异常检测进行实验相关内容的介绍以及分析主流的异常检测方法的结果。

3.1　数据集

本文汇总了在异常检测方法中常用的 5 个数据集，这些数据集均表示一定的工业流程。

3.1.1　SWaT 数据集

SWaT（secure water treatment）[114]数据集模拟了大型现代工厂的水处理过程，涵盖分层通信网络、PLC、人机交互界面和 SCADA 系统等组件。该数据集的系统有 6 个阶段，分别对应水处理的不同功能模块，每个阶段包含不同数量的传感设备和执行设备，具体包括液位传感器、流量传感器、水质传感器等多种传感器。该数据集旨在为网络物理系统防御机制的设计和评估提供真实的数据基础。为了模拟 ICS 的攻击场景，该系统人员使用系统的攻击模型来从网络层攻击该系统[115]。

该模型针对系统中的关键基础设施发动攻击，主要分为 4 个类型：单级单点、单级多点、多级单点、多级多点。数据收集过程持续了 11 天，前 7 天系统正常运行，后 4 天对系统发起了 41 次攻击，这 41 次攻击的攻击类型和作用的设备各不相同。

3.1.2　WADI 数据集

WADI（water distribution tested）[116]数据集是 SWaT 数据集的扩展，其中包括更多的配水管，模拟了一个城市大型供水网络。系统主要由 PLC 控制的三级和 RTU 控制的两级组成，每个 PLC 和 TRU 都使用传感器来估计系统状态，并使用执行器进行控制。该系统主要由 3 个阶段组成，对应于供水系统的供水网络、水质网络和回水网络。该系统使用的攻击模型[117]包含攻击意图和目标组件，针对系统的无线网络进行系统攻击。一共收集连续 16 天的数据，前 14 天系统正常运行，最后两天对系统发起 15 次攻击。

3.1.3　SMAP 数据集

SMAP（soil moisture active passive）[43]数据集是美国 NASA 的地球观测卫星之一，主要功能是测量地球表面的土壤湿度。其传感设备主要分

为两类：一类是作为主动传感设备的L波段雷达；另一类是作为被动传感设备的L波段微波辐射计。两者各自产生3 km和36 km的土壤水分产品，并协同生产9 km的土壤水分产品。SMAP数据集共有55个实体传感器，每个传感器分别有25个维度的特征。除了时间变化值，其他数据的值都为0或1，表示相关命令是否发送成功。

3.1.4 MSL数据集

MSL（mar science laboratory）[43]数据集同样是美国NASA的任务之一，旨在探索火星表面，特别是撞击坑内的情况。MSL任务的重要组成部分是Curiosity火星车，通过该设备实现图像、光谱等多项科学测量结果的收集。相较SMAP的55个实体，MSL共有27个实体，且每个实体都有55个维度的特征。除去时间值外的特征值均为0或1。

3.1.5 SMD数据集

SMD（server machine dataset）[54]数据集的数据源自大型互联网的服务器，主要记录该公司28个不同机器在5周时间内的数据。数据每分钟收集一次，28个不同的机器对应于28个不同的传感设备，每个设备有38个维度的特征，这些特征值均用浮点数表示。SMD数据集主要分为两个相等的子集，即训练集和测试集，测试集中的异常也通过人工进行标注。

3.2 评价指标

评价指标是衡量异常检测方法性能的重要组成部分。不正确的指标可能会有偏见的评估，甚至影响方法的可靠性和通用性[118]。在基于物理建模的攻击检测与定位方法中，在对模型进行攻击评估的同时，也对模型的精确率、召回率和$F1$分数进行评估，这3个评价指标基于以下4个计算性能的指标。

- 真阳性（true positive，TP）：表示正常检测到的异常实例数。
- 真阴性（true negative，TN）：表示正确检测到的正常实例数。
- 假阳性（false positive，FP）：表示错误检测到的异常实例数。
- 假阴性（false negative，FN）：表示错误检测到的正常实例数。

（1）精确率

精确率用于度量正确预测异常数据的能力。其公式如下：

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100\% \tag{4}$$

（2）召回率

召回率用于度量预测异常数据的正确率。其公式如下：

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\% \tag{5}$$

（3）$F1$分数

$F1$分数是综合考虑精确率和召回率的调和平均数，其公式如下：

$$F1 = \frac{2PR}{P + R} \tag{6}$$

面对工业传感网络异常的空间属性，大多数方法忽略了方法在空间方面的检测能力，导致目前并没有被广泛认可的评价指标。冯兆文等[119]对异常检测定位方法进行了探索，提出定位成功率（locating success rate，Lsr）的评价方式，即定位成功率表示定位成功的异常与所有正确预测的异常的比值。

$$\mathrm{Lsr} = \frac{\mathrm{SL}}{\mathrm{TP}} \times 100\% \tag{7}$$

其中，SL表示正确预测并成功定位的异常，该指标考虑了成功检测到的异常，并对其进行定位成功率的分析，然而，它忽略了诸如定位误差率等其他重要因素。因此，关于异常定位的评价指标仍有改进的空间。

3.3 方法性能对比

在异常检测领域的方法中，通常采用准确率，召回率以及$F1$分数作为评价异常检测能力的指标。针对这3个评价指标，本文对目前广泛使用的异常检测方法进行了汇总，并对它们的结果进行了分析。

表2汇总了主流的13种异常检测方法在精确率上的结果比较。如表2所示，不同的方法在不同的数据集上表现有所差异。因SWaT和WADI数据集传感器和执行器结构的复杂性，单时间序列模型在这两个数据集上的表现较差，因WADI数据集传感器和执行器结构的复杂性，单时间序列模型在数据集上的表现存在差异。例如，与GDH、GTA等先进方法相比，USAD、LSTM-

NDT 等方法在该数据集上的表现有一定的差距。多尺度以及图结构模型，因其具有捕捉空间特征的能力，在 SWaT 和 WADI 数据集上具有较好的检测性能，同时在其他数据集上也具有较好的表现。同时，通过对比发现大多数方法在 SWaT 数据集上取得了较高的准确率，但在 WADI 数据集上的精确率较低，例如 USAD 方法，在 SWaT 数据集上具有最高的精确率 99.77%，但在 WADI 数据集上仅有 18.73% 的精确率。

针对上述情况，本文研究了数据集特点以及方法的实现过程，发现在 SWaT 数据集中存在一个容易检测且数据样本量占比较高的一次网络攻击，该攻击造成了异常检测方法在 SWaT 数据集上的高精确率，但在其他数据集上，尤其是在 WADI 数据集上，未表现出在 SWaT 数据集上的优异表现。

表 3 统计了这 13 种异常检测方法在召回率上的实验结果。大多数方法均没有较为稳定的召回率表现，USAD 因采用了编码解码器的对抗形式训练，放大异常输入的重建误差。导致其在 SWaT 数据集上取得了较高的精确率，但在召回率的结果中仅有 68.79%，低于在其他数据集上的表现。在可统计的数据中，基于图结构建模的 GTA 具有最稳定的精确率和召回率。在 SMAP 和 MSL 数据集的召回率的数值达到了所有数据集的领先水平，这是因为在这两个数据集的异常特征较明显，在检测异常时不易发生误判。

表 4 对这些异常检测方法在 $F1$ 分数上的实验结果，从表 4 中可以发现，基于图结构的异常检测方法具有较好的检测稳定性，例如 GTA 在 4 种数据集上的 $F1$ 分数均超过了 0.8。GDN 在整体检测效果最差的 WADI 数据集上取得了第二高的 $F1$ 分数。目前较先进的检测方法 TimesNet 在 5 个数据集上也取得了显著的检测效果。

通过这 3 个实验结果的汇总及研究，本文总结了一些关键信息。目前的异常检测方法并未具备较好的普适性，在不同数据集上的表现存在差异。数据的稀疏程度对异常检测方法的有效性有着重要影响。不同的建模方式决定了异常检测方法的实现机理，使在处理不同稀疏程度数据时各自具备独特的优势。多尺度以及基于图结构的方法因其在空间维度上的建模能力，在应对不同数据集时更具有稳定性，且具有较高的检测性能。

4　工业异常检测应用面临的挑战及建议

4.1　面临的挑战

面向未来 ICS 高互联性、强开放性、自主智能化的特点，对工业传感网络的安全性有了更高

表 2　异常检测方法在精确率上的实验结果

Table 2　Experimental results on precision of anomaly detection methods

方法	精确率				
	SWaT	WADI	SMAP	MSL	SMD
DAGMM	27.46%	87.79%	58.45%	54.12%	91.03%
LSTM-VAE	96.24%	87.79%	85.51%	52.57%	—
USAD	**99.77**%	18.73%	74.80%	79.49%	90.60%
OmniAnomaly	97.82%	31.58%	81.30%	78.48%	88.81%
LSTM-NDT	77.78%	1.38	85.23%	62.88%	97.36%
Anomaly Transformer	72.51%	—	91.85%	79.61%	88.91%
TranAD	97.60%	35.29%	80.43%	90.38%	92.62%
MSCRED	99.92%	25.13%	81.75%	89.12%	72.76%
MAD-GAN	98.97%	41.44%	80.49%	85.17%	**99.91**%
TimesNet	88.31%	—	**92.52**%	83.92%	88.66%
MTAD-GAT	97.18%	28.18%	89.06%	87.54%	82.10%
GDN	99.35%	**97.50**%	74.80%	**93.08**%	71.70%
GTA	88.10%	83.61%	91.76%	91.17%	—

表3　异常检测方法在召回率上的实验结果

Table 3　Experimental results on recall of anomaly detection methods

方法	召回率				
	SWaT	WADI	SMAP	MSL	SMD
DAGMM	69.52%	26.99%	90.58%	99.34%	99.14%
LSTM-VAE	59.91%	14.45%	63.66%	95.46%	—
USAD	68.79%	82.96%	96.27%	99.12%	99.74%
OmniAnomaly	69.57%	65.41%	94.19%	99.24%	**99.85**%
LSTM-NDT	51.09%	78.23%	73.26%	100.0%	84.40%
Anomaly Transformer	**97.32**%	—	58.11%	87.37%	82.23%
TranAD	69.97%	82.96%	**99.99**%	**99.99**%	99.74%
MSCRED	67.70%	73.19%	92.16%	98.62%	99.74%
MAD-GAN	63.74%	33.92%	82.14%	89.91%	84.40%
TimesNet	96.24%	—	58.29%	86.42%	83.14%
MTAD-GAT	69.57%	80.12%	91.23%	94.40%	92.15%
GDN	68.12%	40.19%	98.91%	98.92%	99.74%
GTA	88.10%	**83.61**%	91.76%	91.17%	—

表4　异常检测方法在*F*1分数上的实验结果

Table 4　Experimental results on *F*1 score of anomaly detection methods

方法	*F*1分数				
	SWaT	WADI	SMAP	MSL	SMD
DAGMM	0.39	0.36	0.71	0.67	0.95
LSTM-VAE	0.74	0.25	0.73	0.68	0.83
USAD	0.81	0.31	0.84	0.88	0.95
OmniAnomaly	0.82	0.42	0.87	0.88	0.95
LSTM-NDT	0.62	0.03	0.79	0.77	0.90
Anomaly Transformer	0.83	—	0.71	0.83	0.85
TranAD	0.82	0.50	0.89	0.95	**0.96**
MSCRED	0.81	0.37	0.87	0.94	0.84
MAD-GAN	0.77	0.37	0.81	0.87	0.91
TimesNet	0.92	—	0.72	0.85	0.86
MTAD-GAT	0.81	0.42	**0.90**	0.91	0.87
GDN	0.81	0.57	0.85	**0.96**	0.83
GTA	0.91	0.84	**0.90**	0.91	—

的要求。尽管目前的异常检测方法取得了不错的精度，但在工业实际应用中仍面临较多挑战。

（1）缺少用于验证方法有效性的数据集

异常检测方法的实际应用能力需要在大量的数据集上进行实验，同时方法的有效性也需要不同的工业流程种类作为支撑。在现实世界中，工业部门及企业通常不愿意共享其内部数据供研究人员使用，导致异常检测缺少所需的数据资源。尽管模拟和仿真实验在一定程度上可以模拟真实工业控制系统，但其无法完全再现真实规模和实际攻击环境。工业控制系统环境并不是一成不变的，其随着网络的快速发展不断进行迭代[120]。

这种环境下产生的多变量时间序列也具有变量变化和维度增加等区别。因此，验证方法有效性需要大量且多样的实际数据集。

另外，获取有标签的异常数据也是一个挑战。工业控制系统大多情况下处于正常运行状态[121]，这意味着获取有标记的异常数据具有一定的难度。然而，所有基于深度学习的方法均需要具有标签的数据来验证方法的有效性，即使选择半监督式或无监督式学习方法，获取有标签的数据仍然不容忽视。标签往往需要昂贵的人力资源成本，因此标签数据较少。正因如此，数据集的缺失阻碍了ICS 异常检测的发展。

面对这一问题，最简单有效的方法是构建如SWaT 和 WADI 数据集的工业试验平台，通过网络攻击模型和物理破坏造成传感数据的异常，并将其完整地记录在系统文件中。这些数据在一定程度上还原了真实ICS 的攻击环境，但相比真实工业环境，攻击环境仍较简单。因此，需要工业部门及企业提供更多的真实工业数据。

（2）方法不具有较强的可解释性

方法的可解释性可以描述为用户可以理解的造成结果的原因[122]，通常基于统计与分析的传统方法具有更好的可解释性。尽管基于深度学习的方法在异常检测领域取得显著进展，但其缺乏合理的可解释性。基于深度学习的方法往往侧重于实现高性能，很少涉及其解释性的探讨和研究[123]。这使实验结果数据仅具有直观的意义，未能解释为何模型会产生特定的检测效果，以及模型的决策过程。物理过程显著的工业控制系统对异常检测模型的可解释性有更高的要求，其需要对异常的发生进行解释，使方法符合一定的实际规律。

本文通过对主流的异常检测方法可解释性的探究，认为改进方法可解释性的关键问题在于建立系统正常运行状态下的数学物理模型。目前的异常检测方法侧重于从数据的相关性中发现异常，而忽略了工业流程所描述的物理规律。为解决这一问题，物理信息融合网络（physics-informed neural network，PINN）[124, 125]可以作为一个有价值的选择。

PINN 是一种将物理方程作为约束的神经网络范式，不仅可以像传统神经网络一样学习训练数据样本的分布规律，还能够学习到描述系统的数学方程和物理定律。近年来，PINN 在动态模拟系统物理信息方面显示出了优势[126-134]。

（3）异常检测的实时性要求更高

相较普通的数据驱动领域，工业场景中的异常事件可能对生产流程、安全性和效率产生重大影响，因此，对异常检测的实时性有更高的要求。实际的工业场景通常涵盖更复杂和多变的攻击环境，当前的异常检测方法通常会训练参数体量较大的模型来达到更高的精确度，但模型的复杂性同时也导致方法的实时性不佳。

为了解决这一问题，迫切需要研究更轻量级的异常检测模型，以适应资源受限的工业环境。轻量级的模型在工业应用中具有明显的优势。首先，它们不需要大规模的计算资源，这有助于在实时性要求严格的工业场景中高效运行。其次，轻量级模型的部署和维护相对简单，减轻了技术人员的负担，同时提高了系统的可操作性。在提高异常检测方法性能的同时，也应当考虑其实际应用能力。

（4）缺乏异常定位的能力及评价标准

异常定位能力可以有效地确定异常发生的位置，有助于安全人员对异常事件进行及时的处理，具有较高的实际应用价值。虽然目前的异常检测方法已经开始考虑变量之间的关系，特别是通过系统构建图结构，从时间和空间的尺度上捕获系统的特征。然而，在异常定位方面，这些方法尚未得到充分的实验评估，在异常定位理论上存在一定的空白，迫切需要提出一个被广泛认可的异常定位评价指标。主要涉及异常定位的准确率、召回率、综合性的 $F1$ 分数。通过这样的评价指标，研究人员和从业者可以更准确地评估不同方法在异常定位方面的性能，从而更好地支持系统管理和决策。在工业控制系统的安全性研究中，异常定位评价指标的引入将有助于推动异常检测方法的发展，提高系统的鲁棒性和安全性。

4.2　展望

目前已有的ICS 互联设备安全性研究主要集中于数据样本的异常检测，通过生成或预测的数据标签来匹配原始数据的数据标签。但在真实的

工业环境中，一次简单的网络攻击可能会造成大量的数据异常，这意味着面向数据样本的异常检测高精度并不一定代表着检测到更多的攻击威胁。例如，在SWaT数据集中存在一些容易被检测且持续时间较长的异常数据，导致大多数异常检测方法在SWaT数据集上具有较好的异常检测效果，但在WADI等其他数据集上表现并不理想。相比匹配的精确度，识别异常成因更为重要。

数据中的异常往往是由攻击者的攻击或物理形式上的破坏造成的，为了更准确地描述方法的检测能力，后续研究可将检测的重心转移到攻击的发现上，从异常数据中分析攻击发生的时间和类型，使方法具有识别攻击行为的能力。

5　结束语

本文针对工业传感网络的异常检测技术进行了综述。概述了造成工业传感网络异常的原因，从网络攻击模型和物理故障方式两个方面进行了全面的阐述。本文创新性地从时间和空间两个维度对主流的异常检测方法进行了汇总，重点从对系统的建模方式、技术要点的角度对这些方法的实用性进行了分析和综述。介绍了工业传感网络的实验相关内容，具体包括广泛使用的数据集以及评价指标。汇总了主流的13种异常检测方法在这些评价指标上的性能表现，并对结果进行了分析。最后指出传感网络异常检测工业实际应用的局限性，并从技术角度对突破这些局限性提出了建议。

参考文献：

[1] FALLIERE N, MURCHU L O, CHIEN E. W32. stuxnet dossier[J]. White Paper, 2011, 5(6): 29.

[2] BACKMAN K. When intrusions don’t align: a new water watering hole and oldsmar[EB].

[3] LAKSHMANAN R. Cyber attack in iran reportedly cripples gas stations across the country[EB].

[4] 杨婷, 张嘉元, 黄在起, 等. 工业控制系统安全综述[J]. 计算机研究与发展, 2022, 59(5): 1035-1053.
YANG T, ZHANG J Y, HUANG Z Q, et al. Survey of industrial control system security[J]. Journal of Computer Research and Development, 2022, 59(5): 1035-1053.

[5] CHOI K, YI J, PARK C, et al. Deep learning for anomaly detection in time-series data: review, analysis, and guidelines[J]. IEEE Access, 2021: 120043-120065.

[6] SHYU M L, CHEN S C, SARINNAPAKOR K, et al. A novel anomaly detection scheme based on principal component classifier [C]//Proceedings of IEEE Foundations and New Directions of Data Mining Workshop. IEEE Press, 2003: 172-179.

[7] ANGIULLI F, PIZZUTI C. Fast outlier detection in high dimensional spaces[C]//Proceedings of European Conference on Principles of Data Mining and Knowledge Discovery. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002: 15-27.

[8] BREUNIG M M, KRIEGEL H P, NG R T, et al. Lof: identifying density-based local outliers[C]//Proceedings of 2000 ACM SIGMOD International Conference on Management of Data. 2000: 93-104.

[9] SCHLKOPF B, PLATT J C, SHAWE-TAYLOR J, et al. Estimating the support of a high-dimensional distribution[J]. Neural Computation, 2001, 13(7): 1443-1471.

[10] CONTRERAS J, ESPINOLA E, NOGALES F J, et al. ARIMA models to predict next-day electricity prices[J]. IEEE Transactions on Power Systems, 2003, 18(3): 1014-1020.

[11] KIM D, YANG H, CHUNG M, ET AL. Squeezed convolutional variational autoencoder for unsupervised anomaly detection in edge device industrial internet of things[C]//Proceedings of 2018 IEEE International Conference on Information and Computer Technologies (ICICT). IEEE, 2018: 67-71.

[12] CHEN T, GUESTRIN C. Xgboost: a scalable tree boosting system [C]//Proceedings of 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016: 785-794.

[13] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization[EB]. arXiv Preprint arXiv: 1409.2329, 2014.

[14] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [EB]. arXiv Preprint arXiv: 1803.01271, 2018.

[15] GORI M, MONFARDINI G, SCARSELLI F, et al. A new model for learning in graph domains[C]//Proceedings of 2005 IEEE International Joint Conference on Neural Networks, 2005, 2: 729-734.

[16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017: 30.

[17] UMER M A, JUNEJO K N, JILANI M T, et al. Machine learning for intrusion detection in industrial control systems: applications, challenges, and recommendations[J]. International Journal of Critical Infrastructure Protection, 2022, 38: 100516.

[18] KOAY A M Y, KO R K L, HETTEMA H, et al. Machine learning in industrial control system (ICS) security: current landscape, opportunities and challenges[J]. Journal of Intelligent Information Systems, 2023, 60(2): 377-405.

[19] 唐士杰, 袁方, 李俊, 等. 工业控制系统关键组件安全风险综述[J]. 网络与信息安全学报, 2022, 8(3): 1-17.
TANG S J, YUAN F, LI J, et al. Review on security risks of key components in industrial control system[J]. Chinese Journal of Network and Information Security, 2022, 8(3): 1-17.

[20] DARBAN Z Z, WEBB G I, PAN S, et al. Deep learning for time series anomaly detection: a survey[EB]. arXiv Preprint arXiv: 2211.05244, 2022.

[21] LUO Y, XIAO Y, CHENG L, et al. Deep learning-based anomaly detection in cyber-physical systems: progress and opportunities[J]. ACM Computing Surveys (CSUR), 2021, 54(5): 1-36.

[22] 孙海丽, 龙翔, 韩兰胜, 等. 工业物联网异常检测技术综述[J]. 通

信学报, 2022, 43(3): 196-210.
SUN H L, LONG X, HAN L S, et al. Overview of anomaly detection techniques for industrial internet of things[J]. Journal on Communications, 2022, 43(3): 196-210.

[23] MURALEEDHARAN R, OSADCIW L A. Cross layer denial of service attacks in wireless sensor network using swarm intelligence [C]//Proceedings of IEEE 40th Annual Conference on Information Sciences and Systems. IEEE, 2006: 1653-1658.

[24] PENG L, CAO X, SUN C, et al. Energy efficient jamming attack schedule against remote state estimation in wireless cyber physical systems[J]. Neurocomputing, 2018, 272: 571-583.

[25] ZHANG H, CHENG P, SHI L, et al. Optimal DoS attack scheduling in wireless networked control system[J]. IEEE Transactions on Control Systems Technology, 2015, 24(3): 843-862.

[26] QIN J, LI M, SHI L, et al. Optimal denial-of-service attack scheduling with energy constraint[J]. IEEE Transactions on Automatic Control, 2015, 60(11): 3023-3028.

[27] QIN J, LI M, SHI L, et al. Optimal denial-of-service attack scheduling with energy constraint over packet-dropping networks[J]. IEEE Transactions on Automatic Control, 2017, 63(6): 1648-1663.

[28] ZHANG J, SUN J, LIN H. Optimal DoS attack schedules on remote state estimation under multi-sensor round-robin protocol[J]. Automatica, 2021, 127: 109517.

[29] ZHANG X M, HAN Q L, GE X, et al. Networked control systems: a survey of trends and techniques[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(1): 1-17.

[30] DUO W, ZHOU M C, ABUSORRAH A. A survey of cyber attacks on cyber physical systems: recent advances and challenges[J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9(5): 784-800.

[31] WU G, SUN J, CHEN J. Optimal data injection attacks in cyber-physical systems[J]. IEEE Transactions on Cybernetics, 2018, 48 (12): 3302-3312.

[32] MOUSAVINEJAD E, YANG F, HAN Q L, et al. A novel cyber attack detection method in networked control systems[J]. IEEE Transactions on Cybernetics, 2018, 48(11): 3254-3264.

[33] WU S, GUO Z, SHI D, et al. Optimal innovation-based deception attack on remote state estimation[C]//Proceedings of 2017 American Control Conference (ACC). IEEE, 2017: 3017-3022.

[34] ZHANG Q, LIU K, XIA Y, et al. Optimal stealthy deception attack against cyber-physical systems[J]. IEEE Transactions on Cybernetics. 2019, 50(9): 3963-3972.

[35] LI B, LU R, CHOO K K R, et al. On reliability analysis of smart grids under topology attacks: a stochastic petri net approach[J]. ACM Transactions on Cyber-Physical Systems, 2018, 3(1): 1-25.

[36] MEIRA-GÓES R, KANG E, KWONG R H, et al. Synthesis of sensor deception attacks at the supervisory layer of cyber-physical systems[J]. Automatica. 2020, 121: 109172.

[37] WANG K, GAO H, XU X, et al. An energy-efficient reliable data transmission scheme for complex environmental monitoring in underwater acoustic sensor networks[J]. IEEE Sensors Journal, 2015, 16(11): 4051-4062.

[38] YUAN L, WANG K, MIYAZAKI T, et al. Optimal transmission strategy for sensors to defend against eavesdropping and jamming attacks[C]//Proceedings of 2017 IEEE International Conference on Communications (ICC). IEEE, 2017: 1-6.

[39] YIN X, LI Z, WANG W, et al. Infinite-step opacity and k-step opacity of stochastic discrete-event systems[J]. Automatica. 2019, 99: 266-274.

[40] AN L, YANG G H. Opacity enforcement for confidential robust control in linear cyber-physical systems[J]. IEEE Transactions on Automatic Control, 2019, 65(3): 1234-1241.

[41] YANG S, HOU J, YIN X, et al. Opacity of networked supervisory control systems over insecure communication channels[J]. IEEE Transactions on Control of Network Systems, 2021, 8(2): 884-896.

[42] DEB D, CHAKRABORTY S R, LAGINENI M, et al. Security analysis of MITM attack on SCADA network[C]//Proceedings of Machine Learning, Image Processing, Network Security and Data Sciences: Second International Conference, MIND 2020, Silchar: SingaporeSpringer, 2020: 501-512.

[43] HUNDMAN K, CONSTANTINOU V, LAPORTE C, et al. Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding[C]//Proceedings of 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 387-395.

[44] WU H, GENG Y, LIU K, et al. Research on programmable logic controller security[C]//Proceedings of IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2019, 569(4): 042031.

[45] GUGULOTHU N, MALHOTRA P, VIG L, et al. Sparse neural networks for anomaly detection in high-dimensional time series[C]// Proceedings of AI4IOT Workshop in Conjunction with ICML, IJCAI and ECAI. 2018: 1551-3203.

[46] ZONG B, SONG Q, MIN M R, et al. Deep autoencoding gaussian mixture model for unsupervised anomaly detection[C]//Proceedings of International Conference on Learning Representations(ICLR). 2018.

[47] KIEU T, YANG B, GUO C, et al. Outlier detection for time series with recurrent autoencoder ensembles[C]//Proceedings of International Joint Conference on Artificial Intelligence(IJCAI). 2019: 2725-2732.

[48] AUDIBERT J, MICHIARDI P, GUYARD F, et al. USAD: unsupervised anomaly detection on multivariate time series[C]//Proceedings of 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020: 3395-3404.

[49] GOODGE A, HOOI B, NG S K, et al. Robustness of autoencoders for anomaly detection under adversarial impact[C]//Proceedings of 29th International Conference on International Joint Conferences on Artificial Intelligence. 2021: 1244-1250.

[50] ABDULAAL A, LIU Z, LANCEWICKI T. Practical approach to asynchronous multivariate time series anomaly detection and localization[C]//Proceedings of 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021: 2485-2494.

[51] ZHANG Y, WANG J, CHEN Y, et al. Adaptive memory networks with self-supervised learning for unsupervised anomaly detection [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(12): 12068-12080.

[52] 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.

[53] GUO Y, LIAO W, WANG Q, et al. Multidimensional time series anomaly detection: a GRU-based gaussian mixture variational auto-encoder approach[C]//Proceedings of Asian Conference on Machine Learning. PMLR, 2018: 97-112.

[54] SU Y, ZHAO Y, NIU C, et al. Robust anomaly detection for multivariate time series through stochastic recurrent neural network[C]// Proceedings of 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019: 2828-2837.

[55] LI L, YAN J, WANG H, et al. Anomaly detection of time series with smoothness-inducing sequential variational autoencoder[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(3): 1177-1191.

[56] NIU Z, YU K, WU X. LSTM-based VAE-GAN for time-series anomaly detection[J]. Sensors. 2020, 20(13): 3738.

[57] ZHANG C, LI S, ZHANG H, et al. VELC: A new variational auto-encoder based model for time series anomaly detection[J]. arXiv Preprint arXiv: 1907.01702, 2019.

[58] CHEN R Q, SHI G H, ZHAO W L, et al. A joint model for it operation series prediction and anomaly detection[J]. Neurocomputing, 2021, 448(8): 130-139.

[59] LI L, YAN J, WEN Q, et al. Learning robust deep state space for unsupervised anomaly detection in contaminated time-series[J]. IEEE Transactions on Knowledge and Data Engineering. 2022, 35 (6): 6058-6072.

[60] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.

[61] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[EB]. arXiv Preprint arXiv: 1406.1078, 2014.

[62] LE Q V, JAITLY N, HINTON G E. A simple way to initialize recurrent networks of rectified linear units[EB]. arXiv Preprint arXiv: 1504.00941, 2015.

[63] JING L, GÜLEHRE A, PEURIFOY J, et al. Gated orthogonal recurrent units: on learning to forget[J]. Neural Computation, 2019, 31 (4): 765-783.

[64] CHEN L, CHEN D, YANG F, et al. A deep multi-task representation learning method for time series classification and retrieval[J]. Information Sciences, 2021, 555: 17-32.

[65] GOH J, ADEPU S, TAN M, et al. Anomaly detection in cyber physical systems using recurrent neural networks[C]//Proceedings of 2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE). IEEE, 2017: 140-145.

[66] DING N, MA H X, GAO H, et al. Real-time anomaly detection based on long short-term memory and gaussian mixture model[J]. Computers & Electrical Engineering. 2019, 79: 106458.

[67] SHEN L, LI Z, KWOK J. Timeseries anomaly detection using temporal hierarchical one-class network[J]. Advances in Neural Information Processing Systems, 2020, 33: 13016-13026.

[68] WU W, HE L, LIN W, et al. Developing an unsupervised real-time anomaly detection scheme for time series with multi-seasonality[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 34 (9): 4147-4160.

[69] VOS K, PENG Z, JENKINS C, et al. Vibration-based anomaly detection using LSTM/SVM approaches[J]. Mechanical Systems and Signal Processing, 2022, 169: 108752.

[70] MUNIR M, SIDDIQUI S A, DENGEL A, et al. DeepAnt: a deep learning approach for unsupervised anomaly detection in time series [J]. IEEE Access, 2018, 7: 1991-2005.

[71] YIN C, ZHANG S, WANG J, et al. Anomaly detection based on convolutional recurrent autoencoder for IoT time series[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 52 (1): 112-122.

[72] LIU F, ZHOU X, CAO J, et al. Anomaly detection in quasi-periodic time series based on automatic data segmentation and attentional LSTM-CNN[J]. IEEE Transactions on Knowledge & Data Engineering, 2022, 34(06): 2626-2640.

[73] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [EB]. arXiv Preprint arXiv: 1803.01271, 2018.

[74] CHENG Y, XU Y, ZHONG H, et al. HS-TCN: a semi-supervised hierarchical stacking temporal convolutional network for anomaly detection in IoT[C]//Proceedings of 2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC). IEEE, 2019: 1-7.

[75] LIU J, ZHU H, LIU Y, et al. Anomaly detection for time series using temporal convolutional networks and Gaussian mixture model [C]//Proceedings of Journal of Physics: Conference Series. IOP Publishing, 2019, 1187(4): 042111.

[76] HE Y, ZHAO J. Temporal convolutional networks for anomaly detection in time series[C]//Proceedings of Journal of Physics: Conference Series. IOP Publishing, 2019, 1213(4): 042050.

[77] ZHOU B, LIU S, HOOI B, et al. BeatGAN: anomalous rhythm detection using adversarially generated time series[C]//Proceedings of International Joint Conference on Artificial Intelligence (IJCAI). 2019: 4433-4439.

[78] CHOI Y, LIM H, CHOI H, et al. GAN-based anomaly detection and localization of multivariate time series data for power plant[C]// Proceedings of 2020 IEEE International Conference on Big Data And Smart Computing (BigComp). IEEE, 2020: 71-74.

[79] CHEN X, DENG L, HUANG F, et al. Daemon: unsupervised anomaly detection and interpretation for multivariate time series [C]//Proceedings of 2021 IEEE 37th International Conference on Data Engineering (ICDE). IEEE, 2021: 2225-2230.

[80] DU B, SUN X, YE J, et al. GAN-based anomaly detection for multivariate time series using polluted training set[J]. IEEE Transactions on Knowledge and Data Engineering. 2021. 35(12): 12208-12219.

[81] LI Y, PENG X, ZHANG J, et al. DCT-GAN: Dilated convolutional transformer-based GAN for time series anomaly detection[J]. IEEE Transactions on Knowledge and Data Engineering. 2023, 35(4): 3632-3644.

[82] 胡智超, 余翔湛, 刘立坤, 等. 基于上下文生成对抗网络的时间序列异常检测方法[J]. 哈尔滨工业大学学报. 2024, 56(5): 1-11.
HU Z C, YU X Z, LIU L K, et al. Time series anomaly detection with contextual generative adversarial network[J]. Journal of Harbin Institute of Technology. 2024, 56(5): 1-11.

[83] SONG H, RAJAN D, THIAGARAJAN J J, et al. Attend and diagnose: clinical time series analysis using attention models[C]//Proceedings of AAAI Conference on Artificial Intelligence. 2018. 32(1).

[84] MENG H, ZHANG Y, LI Y, et al. Spacecraft anomaly detection via

transformer reconstruction error[C]//Proceedings of International Conference on Aerospace System Science and Engineering 2019. Springer Singapore, 2020: 351-362.

[85] XU J, WU H, WANG J, et al. Anomaly transformer: time series anomaly detection with association discrepancy[J]. arXiv Preprint arXiv: 2110.02642, 2021.

[86] TULI S, CASALE G, JENNINGS N R. TranAD: deep transformer networks for anomaly detection in multivariate time series data[J]. arXiv Preprint arXiv: 2201.07284, 2022.

[87] KIM J, KANG H, KANG P. Time-series anomaly detection with stacked transformer representations and 1d convolutional network [J]. Engineering Applications of Artificial Intelligence. 2023, 120: 105964.

[88] 霍纬纲, 梁锐, 李永华. 基于随机 Transformer 的多维时间序列异常检测模型[J]. 通信学报, 2023, 44(2): 94-103.
HUO W G, LIANG R, LI Y H. Anomaly detection model for multivariate time series based on stochastic transformer[J]. Journal on Communications, 2023, 44(2): 94-103.

[89] ZHANG C, SONG D, CHEN Y, et al. A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data[C]//Proceedings of AAAI Conference on Artificial Intelligence. 2019. 33(1): 1409-1416.

[90] HE Z, CHEN P, LI X, et al. A spatiotemporal deep learning approach for unsupervised anomaly detection in cloud systems[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 34(4): 1705-1719.

[91] LI Z, ZHAO Y, HAN J, et al. Multivariate time series anomaly detection and interpretation using hierarchical inter-metric and temporal embedding[C]//Proceedings of 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021: 3220-3230.

[92] LI D, CHEN D, JIN B, et al. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks [C]//Proceedings of International Conference on Artificial Neural Networks. Springer, Cham, 2019: 703-716.

[93] WANG X, PI D, ZHANG X, et al. Variational transformer-based anomaly detection approach for multivariate time series[J]. Measurement, 2022, 191: 110791.

[94] 梁李芳, 关东海, 张吉, 等. 基于时空注意力机制的多元时间序列异常检测[J]. 计算机科学, 2023, 50(S2): 450-457.
LIANG L F, GUAN D H, ZHANG J, et al. Spatial-temporal attention mechanism based anomaly detection for multivariate times series[J]. Computer Science, 2023, 50(S2): 450-457.

[95] 张国华, 燕雪峰, 关东海. 基于多模态特征融合的时间序列异常检测[J]. 计算机科学, 2023, 50(S1): 548-554.
ZHANG G H, YAN X F, GUAN D H. Anomaly detection of time-series based on multi-modal feature fusion[J]. Computer Science, 2023, 50(S1): 548-554.

[96] WU H, HU T, LIU Y, et al. TimesNet: temporal 2d-variation modeling for general time series analysis[C]//Proceedings of the Eleventh International Conference on Learning Representations (ICLR). 2023.

[97] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[EB]. arXiv Preprint arXiv: 1609.02907, 2016.

[98] VELIKOVI P, CUCURULL G, CASANOVA A. Graph attention networks[C]//Proceedings of International Conference on Learning Represent- ations(ICLR). 2018.

[99] ZHU Y, XU W, ZHANG J, et al. A survey on graph structure learning: progress and opportunities[J]. arXiv Preprint arXiv: 2103.03036, 2021.

[100] LI R, WANG S, ZHU F, et al. Adaptive graph convolutional neural networks[C]//Proceedings of AAAI Conference on Artificial Intelligence. 2018, 32(1): 3546-3553.

[101] ZHAO T, LIU Y, NEVES L, et al. Data augmentation for graph neural networks[C]//Proceedings of AAAI Conference on Artificial Intelligence. 2021, 35(12): 11015- 11023.

[102] CHEN Y, WU L, ZAKI M. Iterative deep graph learning for graph neural networks: Better and robust node embeddings[J]. Advances in Neural Information Processing Systems, 2020, 33: 19314-19326.

[103] JIANG B, ZHANG Z, LIN D, et al. Semi-supervised learning with graph learning-convolutional networks[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 11313-11320.

[104] LUO D, CHENG W, YU W, et al. Learning to drop: robust graph neural network via topological denoising[C]//Proceedings of 14th ACM International Conference on Web Search and Data Mining. 2021: 779-787.

[105] SUN Q, LI J, PENG H, et al. Graph structure learning with variational information bottleneck[C]//Proceedings of AAAI Conference on Artificial Intelligence. 2022, 36(4): 4165- 4174.

[106] JIN W, MA Y, LIU X, et al. Graph structure learning for robust graph neural networks[C]//Proceedings of 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020: 66-74.

[107] XU H, XIANG L, YU J, et al. Speedup robust graph structure learning with low-rank information[C]//Proceedings of 30th ACM International Conference on Information & Knowledge Management. 2021: 2241-2250.

[108] LIU Y, ZHENG Y, ZHANG D, et al. Towards unsupervised deep graph structure learning[C]//Proceedings of ACM Web Conference 2022. 2022: 1392-1403.

[109] ZHAO H, WANG Y, DUAN J, et al. Multivariate time-series anomaly detection via graph attention network[C]//Proceedings of IEEE International Conference on Data Mining (ICDM). IEEE, 2020: 841-850.

[110] DENG A, HOOI B. Graph neural network-based anomaly detection in multivariate time series[C]//Proceedings of AAAI Conference on Artificial Intelligence. 2021. 35(5): 4027- 4035.

[111] CHEN Z, CHEN D, ZHANG X, et al. Learning graph structures with transformer for multivariate timeseries anomaly detection in iot[J]. IEEE Internet of Things Journal, 2021, 9(12): 9179-9189.

[112] DAI E, CHEN J. Graph-augmented normalizing flows for anomaly detection of multiple time series[EB]. arXiv Preprint arXiv: 2202.07857, 2022.

[113] DING C, SUN S, ZHAO J. MST-GAN: a multimodal spatial-temporal graph attention network for time series anomaly detection[J]. Information Fusion. 2023, 89: 527-536.

[114] GOH J, ADEPU S, JUNEJO K N, et al. A dataset to support research in the design of secure water treatment systems[C]//Proceedings of Critical Information Infrastructures Security: 11th In-

ternational Conference, CRITIS 2016, Paris: Springer International Publishing, 2017: 88-99.

[115] ADEPU S, MATHUR A. An investigation into the response of a water treatment system to cyber attacks[C]//Proceedings of 2016 IEEE 17th International Symposium on High Assurance Systems Engineering (HASE). IEEE, 2016: 141-148.

[116] AHMED C M, PALLETI V R, MATHUR A P. WADI: a water distribution testbed for research in the design of secure cyber physical systems[C]//Proceedings of 3rd International Workshop on Cyber- physical Systems for Smart Water Networks. 2017: 25-28.

[117] ADEPU S, MATHUR A. Generalized attacker and attack models for cyber physical systems[C]//Proceedings of 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC). IEEE, 2016, 1: 283-292.

[118] JUBA B, LE H S. Precision-recall versus accuracy and the role of large data sets[C]//Proceedings of AAAI Conference on Artificial Intelligence. 2019, 33(1): 4039-4048.

[119] 冯兆文, 吴越, 周奇, 等. 面向工业流程的多级图卷积异常检测与定位方法[C]//2022 中国自动化大会,2022.
FENG Z W, WU Y, ZHOU Q, et al. A multi-graph convolution anomaly detection and localization method for industrial process [C]//Proceedings of 2022 The China Automation Congress, 2022.

[120] RAS E, WILD F, STAHL C, et al. Bridging the skills gap of workers in Industry 4.0 by human performance augmentation tools: challenges and roadmap[C]//Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments. 2017: 428-432.

[121] AHMED C M, MR G R, MATHUR A P. Challenges in machine learning based approaches for real-time anomaly detection in industrial control systems[C]//Proceedings of 6th ACM on Cyber-physical System Security Workshop. 2020: 23-29.

[122] ISMAIL A A, GUNADY M, CORRADA B H, et al. Benchmarking deep learning interpretability in time series predictions[J]. Advances in Neural Information Processing Systems, 2020, 33: 6441-6452.

[123] ZHANF X, WANG N, SHEN H, et al. Interpretable deep learning under fire[C]//Proceedings of 29th {USENIX} Security Symposium ({USENIX} Security 20). 2020.

[124] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics informed deep learning (part I): data-driven solutions of nonlinear partial differential equations[EB]. arXiv Preprint arXiv: 1711.10561, 2017.

[125] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics informed deep learning (part II): data-driven discovery of nonlinear partial differential equations[EB]. arXiv Preprint arXiv: 1711.10566. 2017.

[126] ABREU E, FLORINDO J B. A study on a feedforward neural network to solve partial differential equations in hyperbolic-transport problems[C]//Proceedings of Computational Science-ICCS 2021: 21st International Conference, Krakow, Poland, June 16-18, 2021, Proceedings, Part II 21. Springer International Publishing, 2021: 398-411.

[127] KOVACS A, EXL L, KORNDLL A, et al. Conditional physics informed neural networks[J]. Communications in Nonlinear Science and Numerical Simulation, 2022, 104: 106041.

[128] YUCESAN Y A, VIANA F A C. Hybrid physics-informed neural networks for main bearing fatigue prognosis with visual grease inspection[J]. Computers in Industry, 2021, 125: 103386.

[129] VIANA F A C, NASCIMENTO R G, DOURADO A, et al. Estimating model inadequacy in ordinary differential equations with physics informed neural networks[J]. Computers & Structures, 2021, 245: 106458.

[130] JIW, CHANG J, XU H X, et al. Recent advances in metasurface design and quantum optics applications with machine learning, physics-informed neural networks, and topology optimization methods[J]. Light: Science & Applications, 2023, 12(1): 169.

[131] ROY A M, BOSE R, SUNDARARAGHAVAN V, et al. Deep learningaccelerated computational framework based on physics informed neural network for the solution of linear elasticity[J]. Neural Networks, 2023, 162: 472-489.

[132] MISHRA S, MOLINARO R. Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs[J]. IMA Journal of Numerical Analysis, 2022, 42(2): 981-1022.

[133] MOSELEY B, MARKHAM A, NISSEN-MEYER T. Finite basis physics-informed neural networks (FBPINNs): a scalable domain decomposition approach for solving differential equations[J]. Advances in Computational Mathematics, 2023, 49(4): 62.

[134] WU C, ZHU M, TAN Q, et al. A comprehensive study of non-adaptive and residual-based adaptive sampling for physics-informed neural networks[J]. Computer Methods in Applied Mechanics and Engineering, 2023, 403: 115671.

[作者简介]

吴越（1998– ），男，山西临汾人，西北工业大学硕士生，主要研究方向为工业互联网安全。

曹国彦（1986– ），男，陕西榆林人，西北工业大学副教授，主要研究方向为工业控制系统的仿真、控制、计算、通信与安全建模。