

doi:10.19665/j.issn1001-2400.20230804

## 面向多维时间序列异常检测的时空图卷积网络

王静<sup>1,2</sup>, 何苗苗<sup>3</sup>, 丁建立<sup>3</sup>, 李永华<sup>3</sup>

(1. 中国民航大学 安全科学与工程学院, 天津 300300;

2. 中国民航大学 信息安全测评中心, 天津 300300;

3. 中国民航大学 计算机科学与技术学院, 天津 300300)

**摘要:** 针对现有多维时间序列异常检测模型对局部和全局时空依赖性捕获能力不足的问题, 提出一种基于时空图卷积网络的多维时间序列异常检测模型。首先, 在时间维度上利用扩张因果卷积和多头自注意力机制, 分别捕获短期和长期时间依赖性, 并且引入通道注意力来学习不同通道的重要性权重; 其次, 在空间维度上利用静态图学习层根据节点嵌入构建静态图邻接矩阵, 旨在捕获多维时间序列数据的全局空间依赖性, 同时利用动态图学习层构建一系列演化的图邻接矩阵, 旨在建模局部动态的空间依赖性; 最后, 联合优化重构模型和预测模型, 通过重构误差和预测误差计算异常分数, 然后比较阈值和异常分数的关系, 进而检测异常。在MSL、SMAP和SWaT三个公开数据集上的实验结果表明, 该模型在异常检测性能指标F1分数方面优于OmniAnomaly、MTAD-GAT和GDN等相关的基线模型。

**关键词:** 图卷积网络; 时空依赖; 多维时间序列; 异常检测

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 1001-2400(2024)03-0170-12

## Spatial-temporal graph convolutional networks for anomaly detection in multivariate time series

WANG Jing<sup>1,2</sup>, HE Miaomiao<sup>3</sup>, DING Jianli<sup>3</sup>, LI Yonghua<sup>3</sup>

(1. College of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China;

2. Information Security Evaluation Center, Civil Aviation University of China, Tianjin 300300, China;

3. College of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China)

**Abstract:** To address the problem that the existing multivariate time series anomaly detection models have an insufficient ability to capture local and global spatial-temporal dependencies, a multivariate time series anomaly detection model based on spatial-temporal graph convolutional networks is proposed. First, in the temporal dimension, the short-term and long-term temporal dependencies in time series data are captured by using dilated causal convolution and multi-headed self-attention mechanisms, respectively. And the channel attention is introduced to learn the importance weights of different channels. Second, in the spatial dimension, a graph adjacency matrix is constructed by the static graph learning layer according to the node embedding, which is used to model the global spatial dependencies. Meanwhile, a series of evolutionary graph adjacency matrices is constructed by using the dynamic graph learning layer, so as to capture the local dynamic spatial dependencies. Finally, the reconstruction model and the prediction model are jointly

收稿日期: 2023-06-14

网络出版时间: 2023-11-21

基金项目: 国家自然科学基金民航联合基金重点课题(U2033205); 中国民航大学信息安全测评中心开放基金资助(ISECCA-202006)

作者简介: 王静(1980—), 女, 副教授, E-mail: j\_wang@cauc.edu.cn

何苗苗(1998—), 女, 中国民航大学硕士研究生, E-mail: 1782690867@qq.com

丁建立(1963—), 男, 教授, E-mail: jlding@cauc.edu.cn

通信作者: 李永华(1974—), 女, 讲师, E-mail: yh-li@cauc.edu.cn

网络出版地址: <https://doi.org/10.19665/j.issn1001-2400.20230804>

optimized, and the anomaly score is calculated by the reconstructed error and the prediction error. Then, the relationship between the threshold and the anomaly score is compared to detect the anomaly. Experimental results on three public datasets, MSL, SMAP, and SwaT, show that the model outperforms the relevant baseline models such as OmniAnomaly, MTAD-GAT, and GDN in terms of the anomaly detection performance metric F1 score.

**Key Words:** graph convolutional networks; spatial-temporal dependencies; multivariate time series; anomaly detection

## 1 引言

网络物理系统中通常使用传感器监控系统的运行状态,随着传感器数量的快速增加,产生了海量的多维时间序列(Multivariate Time Series, MTS)数据<sup>[1]</sup>,例如,在水处理系统中会使用不同的传感器采集水位、流速、水压以及阀门状态等监测数据。通过对时间序列数据进行高效、准确的异常检测,可以帮助工作人员快速定位异常,并且及时执行干预措施,降低安全风险和经济损失<sup>[2]</sup>。

早期的时间序列异常检测方法主要基于统计学习和机器学习<sup>[3]</sup>,例如自回归综合移动平均和支持向量机等,但这些方法无法处理复杂且非线性的多维时间序列数据。随着深度学习技术的快速发展,研究人员将生成对抗网络(Generative Adversarial Network, GAN)<sup>[4]</sup>、变分自编码器(Variational Autoencoder, VAE)、循环神经网络(Recurrent Neural Network, RNN)<sup>[5]</sup>及其变体等深度神经网络广泛应用于多维时间序列异常检测领域。例如,文献[6]提出了结合VAE和长短期记忆网络(Long Short-Term Memory, LSTM)的MTS异常检测模型,利用LSTM作为VAE架构的推断网络和生成网络来捕获时间相关性。文献[7]提出了一种结合门控循环单元(Gated Recurrent Unit, GRU)和VAE的随机循环神经网络OmniAnomaly,该模型利用GRU捕获MTS的时间依赖性,利用随机变量连接技术对潜在空间中随机变量的时间相关性进行建模,根据重构概率检测异常。文献[8]提出基于生成对抗网络(Generative Adversarial Network, GAN)的无监督MTS异常检测模型,使用LSTM作为GAN的生成器和鉴别器,从而捕获时间依赖性,同时在检测过程中利用生成器的重构误差和鉴别器损失计算异常分数。文献[9]提出的无监督异常检测(UnSupervised Anomaly Detection, USAD)模型包括一个编码器和两个解码器,采用两阶段对抗训练框架来放大含有异常输入的重构误差,避免模型无法区分接近正常样本的异常样本。

然而,上述异常检测方法只考虑了多维时间序列数据在时间维度上的依赖性,但没有考虑其空间维度上的依赖关系。近年来,图神经网络<sup>[10]</sup>(Graph Neural Networks, GNN)在复杂时空数据建模方面取得了重大进展,由于MTS本质上也是时空数据,因此研究人员开始将GNN应用于MTS异常检测任务。文献[11]提出一种基于图注意力网络的MTS异常检测模型图偏差网络(Graph Deviation Network, GDN),该模型根据节点嵌入向量之间的余弦相似度构建有向图结构,用于描述不同变量之间的相关性,然后利用基于图注意力网络的方法预测未来的值,根据预测误差确定异常。文献[12]提出的基于图注意力网络的多维时间序列异常检测(Multivariate Time-series Anomaly Detection via Graph Attention Network, MTAD-GAT)模型采用两个并行的图注意力层,其中面向特征的图注意力层旨在建模空间依赖性,面向时间的图注意力层用于捕获不同时间戳之间的依赖关系。文献[13]提出的基于Transformer的图学习异常检测(Graph Learning with Transformer for Anomaly Detection, GTA)模型将Transformer<sup>[14]</sup>和图卷积网络(Graph Convolutional Networks, GCN)<sup>[15]</sup>相结合,用于建模MTS中的时间和空间依赖性。

但是现有基于图神经网络的多维时间序列异常检测模型仍然存在以下两个问题:

- (1) 现有方法通常忽略了同时捕获多维时间序列数据的短期时间依赖性和长期时间依赖性;
- (2) 现有方法主要通过构建固定的图结构来捕获全局空间依赖关系,忽略了多维时间序列中不同变量之间的潜在依赖关系可能随时间动态变化,导致模型不能充分挖掘全局和局部动态的空间特征。

针对以上问题,文中提出一种基于时空图卷积网络的多维时间序列异常检测模型(Multivariate Time

Series Anomaly Detection via Spatial-Temporal Graph Convolutional Networks, MTSAD-STGCN)。该模型主要由两个时空模块组成,从时间维度和空间维度同时提取局部和全局特征。首先,在时间模块中利用堆叠的扩张因果卷积提取局部时间特征,并且引入通道注意力对提取的特征从通道维度调整重要程度并分配权重,同时利用多头自注意力机制捕获长期时间依赖性;其次,在空间模块中设计了静态图学习层和动态图学习层,在没有任何先验知识的情况下从数据中学习静态和动态图邻接矩阵,分别建模全局空间依赖关系和随时间变化的局部空间依赖关系,接着将两种类型的图邻接矩阵集成到门控图卷积模块中,提高模型挖掘局部和全局空间特征的能力;最后,在 MSL、SMAP、SWaT 公开数据集上开展实验,实验结果表明文中提出的模型和基线模型相比具有更高的 F1 分数。

## 2 MTSAD-STGCN 模型设计

### 2.1 问题定义

将多维时间序列数据集表示为  $X = \{x^{(1)}, x^{(2)}, \dots, x^{(T)}\} \in \mathbf{R}^{N \times P}$ , 其中  $N$  为变量的数目,  $P$  为历史时间步的长度。在第  $t$  个时间步,  $x^{(t)} \in \mathbf{R}^N$  表示所有变量的取值。对于长时间序列, 使用滑动步幅为 1, 窗口大小为  $w$  的滑动窗口划分原始时间序列, 每个滑动窗口输入的样本记为  $\chi = \{(t-w), (t-w+1), \dots, (t-1)\} \in \mathbf{R}^{N \times w}$ , 将其作为模型输入。异常检测模型的任务是产生一组二进制标签  $\text{label} \in \mathbf{R}^L$ , 其中  $L$  是测试集的长度,  $\text{label}^{(t)} \in \{0, 1\}$ ,  $1 \leq t \leq L$ , 表明测试集的第  $t$  个时间步是否存在异常, 其中 1 表示异常, 0 表示正常。文中使用的数学符号如表 1 所示。

表 1 符号描述

符号	描述	符号	描述
$N$	变量的数目	$C$	通道数
$P$	历史时间步的长度	$\gamma$	超参数, 计算异常分数时的组合比例
$w$	滑动窗口的大小	$Z$	时间模块的输出特征
$L$	测试集的长度	$H$	空间模块的输出特征
$\chi$	滑动窗口输入	$\text{score}^{(t)}$	$t$ 时刻的异常分数

### 2.2 模型整体框架

MTSAD-STGCN 模型的整体框架如图 1 所示, 该模型主要由一个  $1 \times 1$  卷积层、两个堆叠的时空模块、联合优化模块和异常判定模块组成。首先, 模型的输入样本  $\chi \in \mathbf{R}^{N \times w}$  通过  $1 \times 1$  的二维卷积实现通道维度的扩充, 得到初始特征表示  $S \in \mathbf{R}^{N \times w \times C}$ 。然后, 将其作为时空模块的输入, 每个时空模块包括时间模块和空间模块, 时间模块旨在捕获短期和长期时间依赖性, 空间模块旨在捕获局部和全局空间依赖性, 同时, 为了避免梯度消失问题, 每个时空模块中加入了残差连接。接着, 将提取的时空特征送入基于预测的模型和基于重构的模型中, 得到预测值和重构值。最后, 在异常判定模块中根据阈值和异常分数判断是否发生异常。

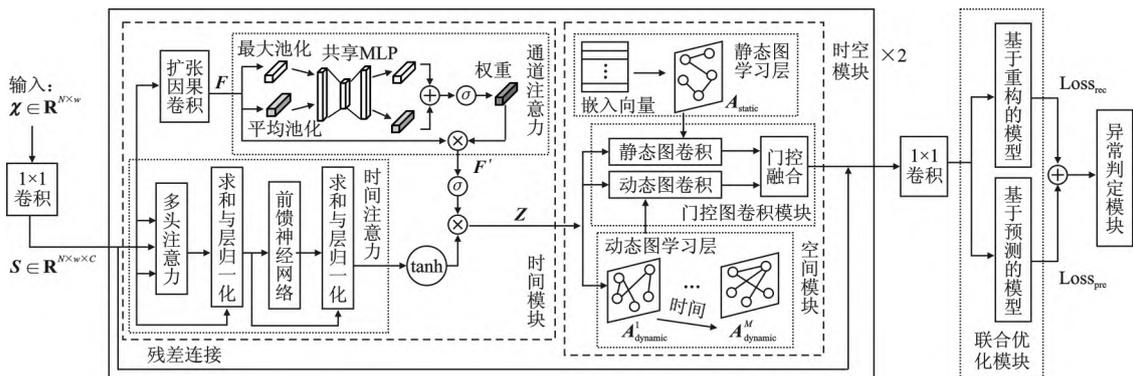


图 1 模型整体框架

### 2.3 时间模块

为了使模型可以同时捕获多维时间序列数据中不同时间戳之间的短期和长期依赖,文中设计了一个时间模块,如图1所示,该模块包括时间注意力分支,以及扩张因果卷积和通道注意力组成的分支,前者用于提取时序数据中的全局时间特征,后者则用于提取局部时间特征。

#### 2.3.1 时间注意力层

Transformer中的多头自注意力机制克服了循环神经网络由于梯度消失或爆炸而无法捕获长期时间依赖性的问题,因此,笔者利用多头自注意力机制提取MTS数据的全局时间特征。首先通过可学习的参数矩阵 $W_i^Q$ 、 $W_i^K$ 、 $W_i^V$ 将输入 $S$ 分别线性映射为查询矩阵 $Q_i$ 、键矩阵 $K_i$ 和值矩阵 $V_i$ 。然后利用缩放点积注意力计算第 $i$ 个子头的自注意力结果 $O_i$ :

$$O_i = \text{Attention}(Q_i, K_i, V_i) = \text{softmax}\left(\frac{Q_i K_i^T}{(d_k)^{1/2}}\right) V_i \quad (1)$$

接着将 $h$ 个单头自注意力机制的结果拼接起来,并通过线性映射矩阵 $W^O$ 将其投影回原始空间。随后计算结果被送入前馈神经网络,得到输出 $\xi \in \mathbf{R}^{N \times w \times C}$ 。

#### 2.3.2 扩张因果卷积

扩张因果卷积<sup>[16]</sup>在标准卷积操作的基础上引入控制跳跃距离的扩张因子,其感受野随着层深度的增加呈指数级扩大,灵活的感受野大小使得模型可以挖掘更丰富的时间特征表示。同时扩张因果卷积引入了因果属性,确保当前时刻的输出值仅与历史数据有关,避免了未来数据的泄露。因此,文中利用在时间维度上的扩张因果卷积提取MTS数据的局部时间特征。扩张因果卷积模块的输入为 $S$ ,在卷积过程中通过padding策略保持时间维度长度不变,扩张卷积的输出定义为

$$F = \text{ReLU}(\Phi *_{d} S) \quad (2)$$

其中, $\Phi \in \mathbf{R}^{K \times C \times C}$ 表示卷积核, $*_{d}$ 表示膨胀因子为 $d$ 的扩张因果卷积操作, $\text{ReLU}(\cdot)$ 是激活函数,输出特征 $F \in \mathbf{R}^{N \times w \times C}$ 。

#### 2.3.3 通道注意力

文中引入通道注意力<sup>[17]</sup>建模不同通道之间的相关性,其结构如图1所示。通道注意力模块首先在时间维度和空间维度上对扩张因果卷积层输出的特征 $F \in \mathbf{R}^{C \times N \times w}$ 进行全局平均池化和全局最大池化,得到平均池化特征 $F_{\text{avg}}^c \in \mathbf{R}^{C \times 1 \times 1}$ 和最大池化特征 $F_{\text{max}}^c \in \mathbf{R}^{C \times 1 \times 1}$ 。然后将两个特征经过具有共享权重的两层全连接层,接着使用sigmoid激活函数将不同通道的注意力权重限定在 $(0, 1)$ 范围内。最后将得到的通道注意力权重 $M_c \in \mathbf{R}^{C \times 1 \times 1}$ 和原始输入特征 $F$ 相乘,得到通道注意力模块的输出 $F'$ 。整个过程的计算公式如下:

$$F' = M_c(F) \otimes F \quad (3)$$

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) = \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \quad (4)$$

式中, $W_0 \in \mathbf{R}^{C \times C}$ 和 $W_1 \in \mathbf{R}^{C \times \frac{C}{r}}$ 分别表示第一个和第二个全连接层的权重参数, $r$ 是缩减比率, $\sigma$ 表示sigmoid激活函数, $\otimes$ 代表对应元素相乘。

#### 2.3.4 时间特征融合

为了更好地控制信息流,文中采用了与文献[18]类似的门控机制,其由两个并行的激活函数构成,tanh激活函数被作为一个滤波器,sigmoid激活函数用于控制传递到下一个模块的信息量。具体而言,将通道注意力模块的输出 $F'$ 和时间注意力层的输出 $\xi$ 经过两个不同的激活函数,然后进行逐元素相乘,融合后的特征作为时间模块的输出。公式定义如下:

$$Z = \sigma(F') \odot \tanh(\xi) \quad (5)$$

式中, $\sigma(\cdot)$ 是sigmoid激活函数, $\odot$ 表示哈达玛积, $Z \in \mathbf{R}^{N \times w \times C}$ 是融合后的时间特征。

### 2.4 空间模块

为了挖掘多维时间序列数据中不同变量之间隐藏的局部和全局空间依赖关系,文中设计了空间模块,如

图 1 所示,该模块由静态图学习层、动态图学习层以及门控图卷积模块组成。具体而言,首先,静态和动态图学习层分别从数据中学习静态和动态图邻接矩阵,旨在从不同的角度反映不同变量之间的相关性。其次,在门控图卷积模块中,采用两个并行的图卷积网络根据静态和动态图邻接矩阵分别捕获局部和全局空间特征。最后,利用门控机制自适应地融合两个图卷积模块的输出。

#### 2.4.1 静态图学习层

静态图学习层旨在在没有任何先验知识的情况下,以数据驱动的方式学习自适应的图邻接矩阵,用于建模变量之间的全局空间依赖关系。文中利用两个随机初始化的节点嵌入矩阵构建静态图结构<sup>[18]</sup>,计算过程如下:

$$\begin{aligned} \mathbf{M}_1 &= \tanh(\mathbf{E}_1 \boldsymbol{\theta}_1) \quad , \\ \mathbf{M}_2 &= \tanh(\mathbf{E}_2 \boldsymbol{\theta}_2) \quad , \\ \mathbf{A}_{\text{static}} &= \text{softmax}(\text{ReLU}(\mathbf{M}_1 \mathbf{M}_2^T - \mathbf{M}_2 \mathbf{M}_1^T)) \quad , \end{aligned} \quad (6)$$

式中, $\mathbf{E}_1 \in \mathbf{R}^{N \times e}$  表示源节点嵌入字典, $\mathbf{E}_2 \in \mathbf{R}^{N \times e}$  表示目标节点嵌入字典,在模型训练期间通过随机梯度下降算法学习,其中  $N$  是变量的数量, $e$  是嵌入向量维度。 $\boldsymbol{\theta}_1$ 、 $\boldsymbol{\theta}_2$  是模型参数,使用 ReLU 激活函数消除节点之间的弱连接,softmax 函数用于归一化邻接矩阵,静态邻接矩阵  $\mathbf{A}_{\text{static}} \in \mathbf{R}^{N \times N}$ 。

#### 2.4.2 动态图学习层

在多维时间序列中,不同变量之间的相关性很可能会随着时间的推移而动态变化,仅使用静态图结构无法捕获这种局部动态空间依赖关系。因此,文中设计了一个动态图学习层,如图 2 所示,其核心思想是将多维时间序列中的变量视为图中的节点,利用自注意力计算节点之间的空间相关性强度。

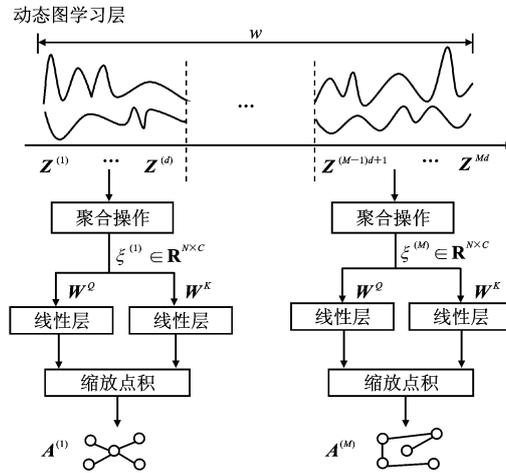


图 2 动态图学习层

首先将动态图学习层的输入特征  $\mathbf{Z} \in \mathbf{R}^{\omega \times N \times C}$  沿着时间维度划分为  $M$  个片段  $\{\mathbf{Z}^m\}_{m=1}^M$ , 然后对每个片段的特征进行聚合,得到变换后的特征  $[\xi^1, \xi^2, \dots, \xi^m, \dots, \xi^M]$ , 第  $m$  个片段的聚合过程如下:

$$\xi^m = \text{AGGREGATE}(\mathbf{Z}^{((m-1)d+1, md)}) \quad , \quad (7)$$

式中, $d$  为每个片段的时间间隔, $\omega$  为滑动窗口大小, $\omega = d \times M$ ,  $\mathbf{Z}^{((m-1)d+1, md)} \in \mathbf{R}^{d \times N \times C}$ , 聚合后的特征  $\xi^m \in \mathbf{R}^{N \times C}$ ,  $\text{AGGREGATE}(\cdot)$  表示通过均值运算实现的聚合操作,用于将时间维度降低为 1。

在跨时间维度的聚合之后,采用缩放点积计算每个片段不同节点之间的空间相关性,如下所示:

$$\alpha_{i,j}^m = \text{softmax}\left(\frac{(\xi_i^m \mathbf{W}^Q) \cdot (\xi_j^m \mathbf{W}^K)^T}{(d_{\text{model}})^{1/2}}\right), \quad \text{for } 1 \leq i, j \leq N \quad , \quad (8)$$

式中, $N$  表示变量数目, $\xi_i^m, \xi_j^m \in \mathbf{R}^C$ ,  $C$  表示通道维度, $d_{\text{model}} = C$ ,  $\mathbf{W}^Q \in \mathbf{R}^{C \times C}$ ,  $\mathbf{W}^K \in \mathbf{R}^{C \times C}$  分别是查询权重和键权重。 $\alpha_{i,j}^m$  是注意力分数,用于度量第  $m$  个片段节点  $i$  和节点  $j$  之间的空间相关性强度,值大表示强相关性,值小则表示弱相关性。所有节点之间的注意力分数构成第  $m$  个时间片段的动态空间依赖权重矩阵  $\mathbf{A}_{\text{dynamic}}^m \in$

$\mathbf{R}^{N \times N}$ , 表示如下:

$$\mathbf{A}_{\text{dynamic}}^m = \begin{bmatrix} \alpha_{1,1}^m & \cdots & \alpha_{1,N}^m \\ \vdots & & \vdots \\ \alpha_{N,1}^m & \cdots & \alpha_{N,N}^m \end{bmatrix} \quad (9)$$

最终, 动态图学习层的输出为一系列邻接矩阵  $\mathbf{A}_{\text{dynamic}} = [\mathbf{A}_{\text{dynamic}}^1, \mathbf{A}_{\text{dynamic}}^2, \dots, \mathbf{A}_{\text{dynamic}}^M]$ 。由于动态图结构  $\{\mathbf{A}_{\text{dynamic}}^m\}_{m=1}^M$  的学习过程依赖于输入  $Z$ , 所以当输入随着时间变化, 邻接矩阵  $\mathbf{A}_{\text{dynamic}}$  也随之变化。

### 2.4.3 门控图卷积模块

为了提取时序数据的局部和全局空间特征, 文中设计了基于图卷积网络的门控图卷积模块, 如图 3 所示, 门控图卷积模块由静态图卷积、动态图卷积和门控融合机制组成。

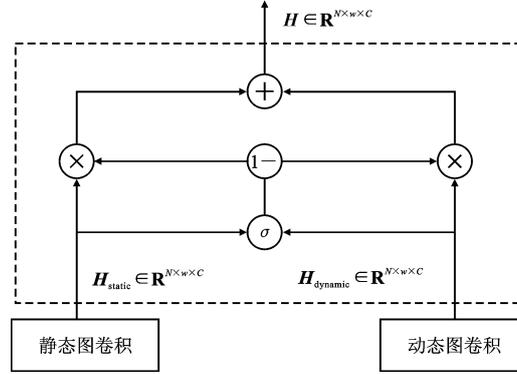


图 3 门控融合

GCN 被广泛用于处理非结构化数据, 旨在通过聚合邻居节点的特征信息得到中心节点新的表征, 计算过程如下:

$$\begin{aligned} \tilde{\mathbf{A}} &= \tilde{\mathbf{D}}^{-1}(\mathbf{A} + \mathbf{I}_N) \quad , \\ f(\mathbf{Y}, \mathbf{A}) &= \sigma(\tilde{\mathbf{A}}\mathbf{Y}\mathbf{W}_1) \quad , \end{aligned} \quad (10)$$

式中,  $\mathbf{A}$  表示图邻接矩阵,  $\mathbf{I}_N$  表示单位矩阵,  $\tilde{\mathbf{A}}$  表示归一化的邻接矩阵,  $\tilde{\mathbf{D}}$  为度矩阵。

静态图卷积模块将时间模块的输出  $Z \in \mathbf{R}^{N \times w \times C}$  和静态图学习层学习到的静态图邻接矩阵  $\mathbf{A}_{\text{static}} \in \mathbf{R}^{N \times N}$  作为输入, 卷积过程如下:

$$\mathbf{H}_{\text{static}} = f(\mathbf{Z}, \mathbf{A}_{\text{static}}) \quad , \quad (11)$$

式中,  $\mathbf{H}_{\text{static}} \in \mathbf{R}^{w \times N \times C}$  表示静态图卷积层提取的全局空间特征。

在动态图卷积模块中, 首先将时间模块的输出  $Z \in \mathbf{R}^{N \times w \times C}$  沿着时间维度进行切片, 划分成  $M$  个片段  $\{\mathbf{Z}^m\}_{m=1}^M$ , 将每一个片段的特征和其对应的动态邻接矩阵  $\mathbf{A}_{\text{dynamic}}^m$  一起送入图卷积层, 定义为

$$\mathbf{H}^m = f(\mathbf{Z}^{(m-1)d+1:md}, \mathbf{A}_{\text{dynamic}}^m) \quad , \quad (12)$$

式中,  $d$  为时间间隔,  $w$  为滑动窗口大小,  $w = d \times M$ ,  $\mathbf{H}^m \in \mathbf{R}^{d \times N \times C}$  表示在第  $m$  个片段图卷积的输出。最后将所有片段的输出  $\{\mathbf{H}^m\}_{m=1}^M$  沿着时间维度进行拼接, 得到动态图卷积层的最终输出  $\mathbf{H}_{\text{dynamic}} \in \mathbf{R}^{w \times N \times C}$ , 具体形式如下:

$$\mathbf{H}_{\text{dynamic}} = \text{concat}(\mathbf{H}^1, \mathbf{H}^2, \dots, \mathbf{H}^M) \quad (13)$$

采用门控机制融合提取到的动态局部空间特征  $\mathbf{H}_{\text{dynamic}} \in \mathbf{R}^{w \times N \times C}$  和全局空间特征  $\mathbf{H}_{\text{static}} \in \mathbf{R}^{w \times N \times C}$ , 如图 3 所示, 表示如下:

$$\begin{aligned} \alpha &= \sigma(\mathbf{H}_{\text{static}}\mathbf{W}_1 + \mathbf{H}_{\text{dynamic}}\mathbf{W}_2 + b) \quad , \\ \mathbf{H} &= \alpha \odot \mathbf{H}_{\text{static}} + (1 - \alpha)\mathbf{H}_{\text{dynamic}} \quad , \end{aligned} \quad (14)$$

式中,  $\mathbf{W}_1, \mathbf{W}_2 \in \mathbf{R}^{C \times C}$  为可学习的参数,  $b$  为偏置项,  $\sigma(\cdot)$  为 sigmoid 激活函数,  $\alpha$  为门控值, 融合后的空间特征  $\mathbf{H} \in \mathbf{R}^{w \times N \times C}$ ,  $\odot$  表示哈达玛积。

## 2.5 联合优化模块

联合优化模块包括基于预测的模型和基于重构的模型,前者旨在预测下一个时刻每个变量的值,后者旨在从高维时间序列中挖掘潜在在表征。在训练过程中,同时更新两个模型的参数,模型的损失函数包含两个优化目标,定义如下:

$$\text{Loss} = \text{Loss}_{\text{pre}} + \text{Loss}_{\text{rec}} \quad , \quad (15)$$

式中,  $\text{Loss}_{\text{pre}}$  表示预测模型的损失函数,  $\text{Loss}_{\text{rec}}$  表示重构模型的损失函数。

### 2.5.1 基于预测的模型

利用 MLP 预测下一个时间步的值,损失函数定义如下:

$$\text{Loss}_{\text{pre}} = \left( \sum_{i=1}^N (x_i^{(w+1)} - \hat{y}_i^{(w+1)})^2 \right)^{1/2} \quad (16)$$

式中,  $N$  是变量的数目,  $w$  是窗口大小,  $x_i^{(w+1)}$  表示在  $w+1$  时刻第  $i$  个时间序列的真实值,  $\hat{y}_i^{(w+1)}$  表示第  $i$  个时间序列在  $w+1$  时间步的预测值。

### 2.5.2 基于重构的模型

从重构的角度出发,将模型视为自编码器(AutoEncoder, AE)架构,从而获得原始输入的重构输出,损失函数定义为

$$\text{Loss}_{\text{pre}} = \sum_{t=1}^w \|x_t - \hat{x}_t\|^2 \quad , \quad (17)$$

式中,  $x_t$  表示在  $t$  时刻所有变量的真实值,  $\hat{x}_t$  表示重构值。

## 2.6 异常判定模块

### 2.6.1 异常分数

首先根据预测误差和重构误差,计算在当前时刻的异常分数,计算公式如下:

$$\text{score}^{(t)} = \frac{\sum_{i=1}^N (x_i^{(t)} - \hat{y}_i^{(t)})^2 + \gamma(x_i^{(t)} - \hat{x}_i^{(t)})^2}{1 + \gamma} \quad , \quad (18)$$

式中,  $x_i^{(t)}$  为第  $i$  个变量在  $t$  时刻的真实值,  $\hat{y}_i^{(t)}$  为  $t$  时刻的预测值,  $\hat{x}_i^{(t)}$  为重构值,  $N$  为变量的数目,  $\gamma$  为超参数,用于优化预测误差和重构误差的组合比例,在 3.7 节中,将给出不同  $\gamma$  值对模型效果的影响。

### 2.6.2 阈值选择

异常检测阈值的选择方法如下:首先将测试集中所有时刻的异常分数作为网格搜索的阈值范围,然后计算相应的 F1 分数,找到使 F1 分数最大的异常分数,将其作为测试集的异常检测阈值  $\epsilon^{[19-20]}$ 。如果当前时刻  $t$  的异常分数  $\text{score}^{(t)}$  大于阈值  $\epsilon$ ,则将该时刻标记为异常,否则标记为正常。

## 3 实 验

### 3.1 数据集

为了验证模型的有效性,文中在 3 个公开的异常检测数据集进行实验,分别是 MSL(Mars Science Laboratory rover)、SMAP(Soil Moisture Active Passive satellite)和 SWaT(Secure water treatment)。MSL 和 SMAP 是美国宇航局提供的航天器遥感数据集<sup>[21]</sup>,SWaT 是从水处理厂系统收集的监测数据<sup>[22]</sup>,各数据集的详细信息如表 2 所示。

表 2 各数据集的详细信息

数据集	特征维度	训练集样本数	测试集样本数	测试集异常样本比例/%
MSL	55	58 317	73 729	10.72
SMAP	25	135 183	427 617	13.13
SWaT	51	496 800	449 919	11.98

### 3.2 评估指标

文中选择精确率 Precision、召回率 Recall 和 F1 分数作为异常检测模型性能的评估指标<sup>[23]</sup>,计算方法如下:

$$\begin{cases} \text{Precision} = \frac{TP}{TP+FP} \\ \text{Recall} = \frac{TP}{TP+FN} \\ \text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \end{cases}, \quad (19)$$

式中,TP(True Positive)表示模型正确检测到的异常样本数量,FP(False Positive)表示正常样本被模型错误判断为异常的样本数量,FN(False Negative)表示异常样本被模型判断为正常的样本数量。F1 分数为精确率和召回率的调和平均值,是衡量异常检测模型性能的主要指标。

### 3.3 实验设置

实验的软件和硬件环境如下:Python 3.6、Pytorch 1.5、CUDA 10.2、CPU 为 Intel(R) Xeon(R) Gold 5215 CPU @2.50GHz,显卡为 NVIDIA Tesla V100。将训练数据按 8:2 的比例划分为训练集和验证集,其中训练集和验证集都是正常数据,测试集存在标记为异常的样本。

模型将原始数据的滑动窗口数据作为输入,将滑动窗口的大小  $w$  设置为 20,滑动步幅为 1。节点嵌入向量的维度  $e$  是 128,多头注意力机制的头数  $h$  被设置为 8。实验采用 Adam 优化器,初始学习率为 0.001,批处理大小  $\text{batch\_size}$  为 64,训练轮数  $\text{epoch}$  为 100,通过在验证集上进行网格搜索,超参数  $\gamma$  设为 0.8。

### 3.4 对比实验与结果分析

为了验证模型在异常检测任务中的性能,在三个公开的数据集上将文中提出的 MTS 异常检测模型 MTSAD-STGCN 与七个典型的异常检测模型进行比较,包括 DAGMM<sup>[24]</sup>、LSTM-VAE<sup>[6]</sup>、OmniAnomaly<sup>[7]</sup>、USAD<sup>[9]</sup>、GDN<sup>[11]</sup>、MTAD-GAT<sup>[12]</sup>和 MST-GAT<sup>[25]</sup>对比结果如表 3 所示,其中最佳的 F1 分数以粗体显示,次佳的 F1 分数以下划线显示。

由表 3 可知,MTSAD-STGCN 模型在所有数据集上均获得最佳的 F1 分数,表明该模型具有良好的泛化能力。具体来说,MTSAD-STGCN 模型在 MSL、SMAP 和 SWaT 数据集上的 F1 分数分别为 0.9661、0.9539、0.8828,比基线模型的最佳结果分别提高了 3.89%、4.85%和 4.73%。DAGMM 模型在三个数据集上的 F1 分数均为最低,原因是它的输入仅包含一个观测值,而不是历史时间序列窗口,因此该模型忽略了对时间依赖关系进行建模,而时间依赖性对于多维时间序列异常检测任务是至关重要的。MTSAD-STGCN 模型在三个数据集上的 F1 分数比 LSTM-VAE 模型分别提高了 28.81%、22.41%、14.43%,LSTM-VAE 的缺点是循环神经网络限制了模型建模长时依赖关系的能力。在文中提出的 MTSAD-STGCN 模型中,扩张因果卷积用于建模短期时间依赖性,同时时间注意力层可以有效地捕获长期时间依赖性。MTSAD-STGCN 模型在 MSL、SMAP 和 SWaT 数据集上的 F1 分数比 OmniAnomaly 分别提高了 6.71%、11.05%、10.06%,比 USAD 模型分别提高了 3.89%、9.05%、9.11%。虽然这两种方法都是基于序列重构,但是 USAD 通过两个自编码器之间的对抗性训练来放大重构误差,从而检测到接近正常样本的异常样本,而 OmniAnomaly 无法做到这一点。OmniAnomaly 模型和 USAD 模型的共同局限性在于它们均无法捕获多元时间序列中变量之间的潜在依赖关系。

GDN、MTAD-GAT 和 MST-GAT 均为基于图神经网络的模型,在三个数据集上实现了比大多数基线方法更好的异常检测性能,这表明图结构可以有效地建模时间序列数据中不同变量之间的关系。文中提出的 MTSAD-STGCN 模型在三个数据集上相较于 GDN 模型均表现出更高的 F1 分数,GDN 模型的局限性在于它只考虑了空间相关性,而忽略了对时间依赖性进行建模,而 MTSAD-STGCN 模型能够在建模空间相关性的同时捕获短期和长期时间依赖性,实现了更好的异常检测效果。在 MSL、SMAP 和 SWaT 数据集上,MTSAD-STGCN 模型的 F1 分数比 MTAD-GAT 模型分别提高了 5.77%、5.26%、7.19%,比 MST-

GAT 模型分别提高了 4.63%、4.85%、4.73%，MTAD-GAT 模型和 MST-GAT 模型虽然都同时考虑了时间和空间相关性，但是固定的图结构难以对随时间演变的空间相关性进行建模。相比之下，文中提出的模型通过静态和动态图学习层捕获局部变化的空间依赖性和全局空间依赖性，显著提高了异常检测的性能。

表 3 MTSAD-STGCN 模型与 7 种模型的性能对比

Model	MSL			SMAP			SWaT		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
DAGMM	0.541 2	0.993 4	0.700 7	0.584 5	0.905 8	0.710 5	0.274 6	0.695 2	0.393 7
LSTM-VAE	0.525 7	0.954 6	0.678 0	0.855 1	0.636 6	0.729 8	0.962 4	0.599 1	0.738 5
OmniAnomaly	0.886 7	0.911 7	0.899 0	0.741 6	0.977 6	0.843 4	0.982 5	0.649 7	0.782 2
USAD	0.881 0	0.978 6	0.927 2	0.769 7	0.983 1	0.863 4	0.985 1	0.661 8	0.791 7
GDN	0.913 5	0.861 2	0.886 6	0.893 2	0.887 2	0.890 2	0.993 5	0.681 2	0.808 2
MTAD-GAT	0.875 4	0.944 0	0.908 4	0.890 6	0.912 3	0.901 3	0.971 8	0.695 7	0.810 9
MST-GAT	0.950 6	0.891 0	0.919 8	0.912 6	0.898 3	0.905 4	0.987 3	0.724 1	0.835 5
Ours	0.981 1	0.951 6	0.966 1	0.935 9	0.972 6	0.953 9	0.986 6	0.798 7	0.882 8

### 3.5 消融实验

在 MSL、SWaT、SMAP 数据集上设计了一系列消融实验，从而验证 MTSAD-STGCN 模型关键组件的有效性。将 MTSAD-STGCN 模型多种变体命名如下：

- (1) V1: 移除时间模块中扩张因果卷积和通道注意力组成的分支，仅保留时间注意力分支。
- (2) V2: 移除时间注意力层。
- (3) V3: 移除动态图学习层，在空间模块中仅使用静态图学习层建模空间相关性，以验证动态图学习层的有效性。
- (4) V4: 移除静态图学习层，保留动态图学习层，其余组件不变。
- (5) V5: 移除重构模型。
- (6) V6: 移除预测模型。

图 4 展示了 MTSAD-STGCN 及其变体在三个数据集上的消融结果，由图 4 可得到以下结论：变体 V1 和变体 V2 的 F1 分数比 MTSAD-STGCN 模型低，这表明捕获多变量时间序列中的长期时间依赖性和短期时间依赖性有助于提高模型异常检测的性能；移除了动态图学习层的变体 V3 会明显降低 F1 分数，因为静态图结构无法感知细粒度的局部空间信息，而动态图结构可以弥补这一缺陷；移除了静态图学习层的变体 V4 其 F1 分数低于 MTSAD-STGCN 模型，这表明在时间序列中建模全局空间依赖的必要性和有效性；变体 V5 和变体 V6 的 F1 分数低于 MTSAD-STGCN 模型，原因是基于重构的模型无法区分接近正常数据的异常数据，而预测模型可以捕获这种异常，这表明联合优化策略可以获得更好的性能。总之，在每个数据集上 MTSAD-STGCN 模型的性能始终优于所有变体，证明模型中的各个组件在提升异常检测性能方面都是有效的。

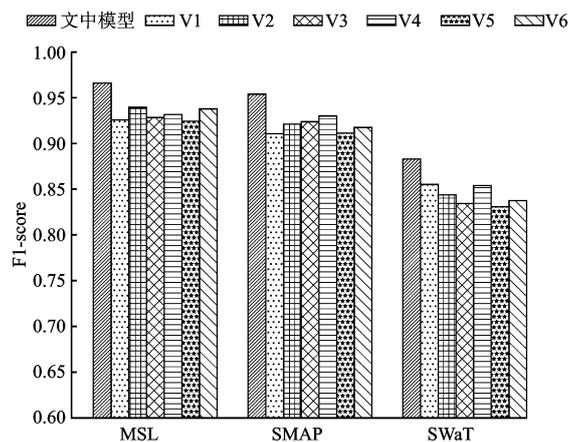


图 4 消融实验结果

### 3.6 可视化分析

为了进一步展示模型异常检测的效果,从 MSL 数据集中选择部分测试集数据进行可视化分析,图 5 展示了可视化结果。其中,图 5(a)中的黑色线条表示每个时刻的异常分数,虚线表示异常检测的阈值;图 5(b)中的黑色线条表示模型检测到的异常;图 5(c)中的虚线表示测试数据集中的实际异常标签。3 个子图的结果直观地证实了文中提出的模型在时间序列异常检测任务中的有效性。

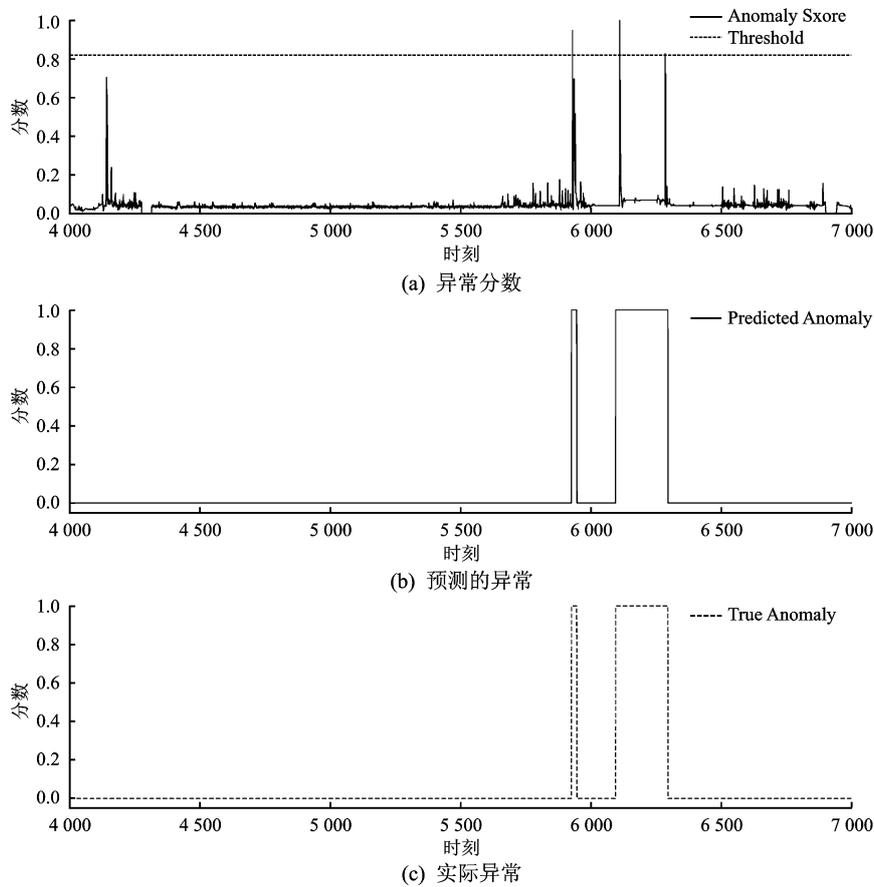


图 5 可视化结果

### 3.7 超参数的敏感性实验

为了研究超参数  $\gamma$  对模型异常检测效果的影响,文中分别在 MSL、SMAP、SWaT 数据集上对  $\gamma$  进行实验,参数  $\gamma$  表示在计算异常分数时预测误差和重构误差的贡献值占比,网格搜索的取值范围为  $\{0.2, 0.4, 0.6, 0.8, 1\}$ 。图 6 展示了在 3 个数据集上不同  $\gamma$  值对应的 F1 分数、精确率和召回率。由图 6 中的结果可看出,当  $\gamma$  为 0.8 时,模型在 3 个数据集上的性能达到最优。总体而言,不同的  $\gamma$  值下 MTSAD-STGCN 模型均取得良好的异常检测性能,表明该模型对  $\gamma$  具有较强的鲁棒性。

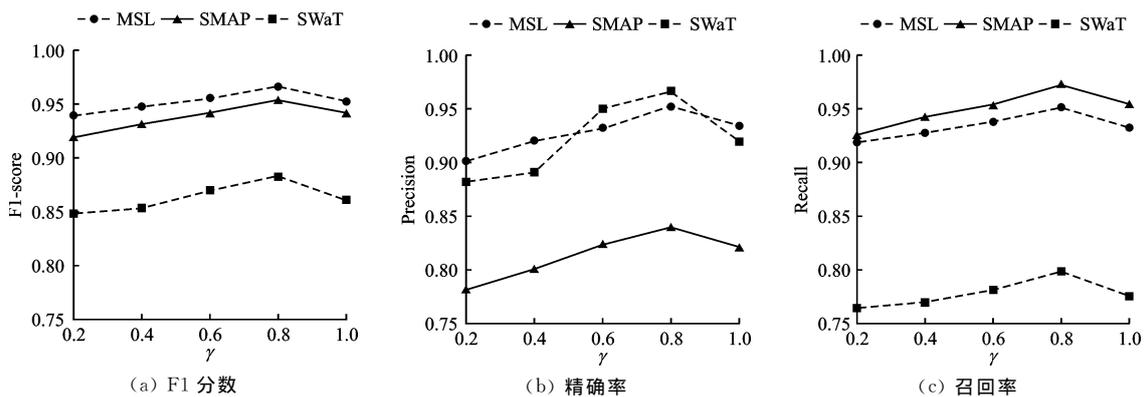


图 6 超参数敏感性实验

## 4 结束语

文中提出了基于时空图卷积网络的多维时间序列异常检测模型 MTSAD-STGCN,该模型设计了时间模块和空间模块,前者旨在捕获多维时间序列的短期和长期时间依赖性,后者旨在学习不同变量之间的动态局部空间依赖关系和全局空间依赖关系。此外,将基于 MLP 的预测模型和基于 AE 的重构模型相结合,可以更全面地检测出异常。在三个公开数据集上的对比实验和消融实验验证了文中方法的有效性以及良好的泛化能力。在未来的工作中,我们将在此模型的基础上研究如何与在线学习策略相结合,进一步提升模型的实用性。

### 参考文献:

- [1] WENIG P, SCHMIDL S, PAPENBROCK T. TimeEval: A Benchmarking Toolkit for Time Series Anomaly Detection Algorithms[J]. Proceedings of the VLDB Endowment, 2022, 15(12): 3678-3681.
- [2] 丁小欧, 于晟健, 王沐贤, 等. 基于相关性分析的工业时序数据异常检测[J]. 软件学报, 2020, 31(3): 726-747.  
DING Xiaou, YU Shengjian, WANG Muxian, et al. Anomaly Detection on Industrial Time Series Based on Correlation Analysis[J]. Journal of Software, 2020, 31(3): 726-747.
- [3] SCHMIDL S, WENIG P, PAPENBROCK T. Anomaly Detection in Time Series: A Comprehensive Evaluation [J]. Proceedings of the VLDB Endowment, 2022, 15(9): 1779-1797.
- [4] 顾兆军, 刘婷婷, 隋翯. 一种 ICS 异常检测的优化 GAN 模型[J]. 西安电子科技大学学报, 2022, 49(2): 173-181.  
GU Zhaojun, LIU Tingting, SUI He. Latent Feature Reconstruction Generative GAN Model for ICS Anomaly Detection[J]. Journal of Xidian University, 2022, 49(2): 173-181.
- [5] YU Y, SI X, HU C, et al. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.
- [6] 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.
- [7] SU Y, ZHAO Y, NIU C, et al. Robust Anomaly Detection for Multivariate Time Series Through Stochastic Recurrent Neural Network [C]// Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2019: 2828-2837.
- [8] LI D, CHEN D, JIN B, et al. MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks [C]// International Conference on Artificial Neural Networks. Heidelberg: Springer, 2019: 703-716.
- [9] AUDIBERT J, MICHIARDI P, GUYARD F, et al. USAD: Unsupervised Anomaly Detection on Multivariate Time Series [C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 3395-3404.
- [10] 吴博, 梁循, 张树森, 等. 图神经网络前沿进展与应用[J]. 计算机学报, 2022, 45(1): 35-68.  
WU Bo, LIANG Xun, ZHANG Shusen, et al. Advances and Applications in Graph Neural Network[J]. Chinese Journal of Computers, 2022, 45(1): 35-68.
- [11] DENG A, HOUI B. Graph Neural Network-Based Anomaly Detection in Multivariate Time Series [C]// Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2021: 4027-4035.
- [12] ZHAO H, WANG Y, DUAN J, et al. Multivariate Time-Series Anomaly Detection via Graph Attention Network [C]// 2020 IEEE International Conference on Data Mining (ICDM). Piscataway: IEEE, 2020: 841-850.
- [13] CHEN Z, CHEN D, ZHANG X, et al. Learning Graph Structures with Transformer for Multivariate Time-Series Anomaly Detection in IoT[J]. IEEE Internet of Things Journal, 2021, 9(12): 9179-9189.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is All You Need [C]// Advances in Neural Information

- Processing Systems (NIPS). San Diego: NIPS, 2017: 6000-6010.
- [15] KIPF T N, WELLING M. Semi-Supervised Classification with Graph Convolutional Networks (2016) [J/OL]. [2016-09-09]. <https://arxiv.org/pdf/1609.02907.pdf>.
- [16] BAI S, KOLTER J Z, KOLTUN V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling (2018) [J/OL]. [2018-03-04]. <https://arxiv.org/pdf/1803.01271.pdf>.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [C]// Proceedings of the European Conference on Computer Vision (ECCV). Heidelberg: Springer, 2018: 3-19.
- [18] WU Z, PAN S, LONG G, et al. Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks [C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 753-763.
- [19] FENG C, TIAN P. Time Series Anomaly Detection for Cyber-Physical Systems via Neural System Identification and Bayesian Filtering [C]// Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM, 2021: 2858-2867.
- [20] XIE J, CUI Y, HUANG F, et al. MARINA: An MLP-Attention Model for Multivariate Time-Series Analysis [C]// Proceedings of the 31st ACM International Conference on Information & Knowledge Management. New York: ACM, 2022: 2230-2239.
- [21] KIM S, CHOI K, CHOI H S, et al. Towards a Rigorous Evaluation of Time-Series Anomaly Detection [C]// Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2022: 7194-7201.
- [22] HAN S, WOO S S. Learning Sparse Latent Graph Representations for Anomaly Detection in Multivariate Time Series [C]// Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2022: 2977-2986.
- [23] ZHANG Z, LI W, DING W, et al. STAD-GAN: Unsupervised Anomaly Detection on Multivariate Time Series with Self-Training Generative Adversarial Networks [J]. ACM Transactions on Knowledge Discovery from Data, 2023, 17(5): 1-18.
- [24] ZONG B, SONG Q, MIN M R, et al. Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection [C]// International Conference on Learning Representations. Piscataway: IEEE, 2018: 1-19.
- [25] DING C, SUN S, ZHAO J. MST-GAT: A Multimodal Spatial-Temporal Graph Attention Network for Time Series Anomaly Detection [J]. Information Fusion, 2023, 89: 527-536.

(编辑: 杨 薇)