中国科学: 信息科学 2023 年 第 53 卷 第 9 期: 1784–1801
SCIENTIA SINICA Informationis
论文

基于启发式时空图神经网络的多变量时序异常检测

姜羽[1], 陈华[2*], 张小刚[1*], 王炼红[1], 王鼎湘[1]

1. 湖南大学电气与信息工程学院, 长沙 410081
2. 湖南大学信息科学与工程学院, 长沙 410081
* 通信作者. E-mail: chua@hnu.edu.cn, zhangxg@hnu.edu.cn

收稿日期: 2022–11–03; 修回日期: 2023–01–18; 接受日期: 2023–02–24; 网络出版日期: 2023–09–14

国家自然科学基金 (批准号: 62273139, 62171184, 62106072) 资助项目

摘要 针对信息物理系统的多变量时序数据的异常检测是预防系统故障、保证安全生产的必要手段. 由于系统变量间的强耦合性和传播效应, 设计异常检测算法时应考虑系统变量间的耦合特性、传播有向性和因果时滞性, 从系统结构变化的角度检测早期异常. 本文提出一种端到端的启发式时空图神经网络 (heuristic spatio-temporal graph neural network, HST-GNN) 用于多变量时序数据的异常检测. 首先, 考虑变量间关系的有向性和集群性, 设计一种有向相似性函数和基于启发式聚类算法的图结构学习算法, 对多变量时序数据进行图建模以学习变量间的空间耦合关系; 其次, 使用门控卷积注意单元和多头图注意层作为时空图注意模块, 从时空层面同时捕获系统的非线性因果时序和空间耦合深度特征; 最后, 量化系统的图结构特征, 将其作为时空图网络提取的传感器深度特征的补充, 输入自编码器中, 从系统级别和传感器级别来检测异常. 本文在 4 个公共数据集上验证了 HST-GNN 的性能. 实验结果表明, 稀疏有向的图结构有利于系统耦合特性的提取, 从系统和传感器级别检测异常增加了模型对不显著的早期异常的敏感度.

关键词 多变量时序数据, 无监督异常检测, 启发式图结构, 时空图注意网络, 系统级图结构特征

1 引言

工业系统的信息物理系统 (cyber-physical systems, CPSs) 通常采用大量传感器来采集数据, 这些数据以多变量时序的形式呈现, 反映 CPS 的运行状态. 及时准确的多变量时序异常检测可对潜在事件发出警报, 这对 CPS 系统的健康运行非常重要.

根据建模时多变量时序数据是否有标签, 异常检测算法可以分为有监督和无监督两类. 有监督的异常检测方法训练时需要使用标签, 模型只能检测出已知的异常类型, 应用范围有限, 因此异常检测研究主要集中在无监督方法上. 常见的无监督多变量时序异常检测方法可分为三类: 基于统计学

引用格式: 姜羽, 陈华, 张小刚, 等. 基于启发式时空图神经网络的多变量时序异常检测. 中国科学: 信息科学, 2023, 53: 1784–1801, doi: 10.1360/SSI-2022-0425
Jiang Y, Chen H, Zhang X G, et al. Multivariable time series anomaly detection using heuristic spatio-temporal graph neural network (in Chinese). Sci Sin Inform, 2023, 53: 1784–1801, doi: 10.1360/SSI-2022-0425

的方法 [1,2]、基于机器学习的方法 [3,4] 和基于深度学习的方法 [5~16]. 基于统计学的方法主要利用高斯 (Gauss) 分布或泊松 (Poisson) 分布等建立模型, 将模型低概率区域中的数据视为异常数据, 由于算法需要充分的数据和先验知识, 因此对于机理不可知的现实系统并不适用. 基于机器学习的方法主要采用主成分分析法或聚类的方法学习变量间的关联关系, 其提取的耦合关系较为简单, 难以处理具有复杂耦合关系的高维系统.

近年来, 得益于深度学习出色的非线性拟合能力, 基于自编码器 (autoencoder, AE) [7,8]、循环神经网络 (recurrent neural network, RNN) [11~13]、生成对抗网络 (generative adversarial network, GAN) [11~13] 的异常检测方法取得了长足的进步. 但上述方法在处理具有潜在成对耦合特性的多变量时序数据时面临困难, 模型的可解释性差. 最近, 图神经网络 (graph neural network, GNN) 展现了对成对变量建模的卓越能力. 一些研究使用 GNN 提取传感器对的耦合特征, 对多变量时序进行图建模来实现异常检测, 取得了良好的检测结果 [6,14,15]. 然而基于 GNN 的多变量时序异常检测仍然存在下列问题:

(1) GNN 需要图结构作为应用基础, 而多变量时序数据的图结构通常是未知的. 因此, 一些研究采用全连接的方式对多变量时序数据进行图建模, 得到无向图; 一些研究采用 ϵ- 半径 [17] 或 k- 近邻 [18] 来创建相对简洁的图结构. 而由于传感器的关联关系是有向的 [18], 甚至是单向的, 使用无向图表示变量间的有向关系, 会将不存在的关系引入图结构; 由于变量间的耦合关系存在强弱区别, 具有一定的集群特性, 当使用不恰当的 ϵ 和 k 进行静态图建模时会忽略变量的集群特性, 使得构建出的图结构存在孤立节点或冗余, 导致必要的耦合关系丢失或将不存在的关系引入图结构, 为后续耦合特征的提取带来误差 [19].

(2) 在 CPS 系统中, 不同的传感器具有不同的特性, 对于系统状态变化的响应速度不同. 当系统异常运行时, 上述情况尤为明显. 而目前用于提取非线性时间特性的 RNN/LSTM (long short term memory) 等难以捕捉长期依赖关系, 而 Transformer 虽然能提取长期依赖, 但其只能对时序数据进行逐点级的点积注意力, 没有考虑到各节点间的因果时滞效应.

(3) 目前研究仅在传感器级别进行异常检测, 当传感器级别的早期异常十分微小, 异常状态与正常状态区别不大时, 现有的异常检测方法可能会产生漏检.

针对上述问题, 考虑系统变量间的强耦合性和传播效应, 本文提出一种从系统和传感器级别检测异常的多变量时序异常检测模型 —— 启发式时空图神经网络 (heuristic spatio-temporal graph neural network, HST-GNN). HST-GNN 通过有向相似性函数和启发式聚类算法的图结构学习算法, 对多变量时序进行图建模, 使用时空图注意模块捕获系统的非线性因果时序和空间耦合深度特征, 从系统级别和传感器级别检测时序数据异常. 首先, 考虑系统变量间的关系是有向的, 设计了一种有向的相似性函数计算变量对的相关性, 同时考虑图结构的集群特性, 提出一种基于启发式聚类算法的图结构学习算法. 其次, 提出了一种包含门控卷积注意力单元 (gated convolutional attention unit, GCAU) 和多头图注意层 (multihead graph attention network, MGAT) 的时空图注意模块网络 (spatio-temporal graph attention module, ST-GAM), 从时间和空间维度同时捕捉多变量时序数据的特征: GCAU 采用具有局部上下文敏感的卷积注意力 (convolution attention, CA) [20], 通过因果卷积产生查询向量 **quary** 和键 **key**, 以便将局部上下文更好地纳入注意机制, 解决变量间的因果时延问题, MGAT 通过图注意力机制精确捕捉空间层面变量对之间的耦合特性 [21]. 最后, 针对传感器级偏差会忽略早期异常检测的问题, 选择平均路径长度等复杂网络测度量化图结构, 将其作为表征当前时刻的系统运行和信息传输状态的系统级特征, 与 ST-GAM 捕捉到的时空深度特征一起送入 AE 中, 进行多变量时序无监督异常检测.

本文的主要贡献如下:

(1) 提出了一种连通有向图结构学习算法对多变量时序进行图建模, 解决了 GNN 在多变量时序分析中图建模时存在传感器孤岛 (单节点或子图) 或结构密集的问题.

(2) 设计了一个 ST-GAM 模块, 通过带有局部上下文敏感的 CA 的门控注意力单元 (gate attention unit, GAU) [22] 和 MGAT 结构, 同时挖掘存在时延的变量时间维度上的非线性行为和变量间在空间维度上的耦合关系, 对时序数据的异常突变更为敏感.

(3) 提出了一种基于复杂网络拓扑量化特征的系统级特征, 作为 ST-GAM 捕捉到的时空特征传感器级特征的补充, 使网络对传感器的早期异常更为敏感.

本文的结构如下: 第 2 节给出了关于多变量时序异常检测和图神经网络的综述. 第 3 节详细介绍了提出的异常检测模型 HST-GNN. 第 4 节给出了实验验证, 对实验结果进行分析和讨论. 第 5 节给出结论.

2 相关工作

2.1 多变量时序异常检测

在常见的多变量时序无监督异常检测方法中, 基于统计学的方法需要已知全部或部分被监控系统的先验知识 (例如数据满足某种分布或概率模型), 通过判断数据是否满足先验知识来实现异常检测: 例如, Benkabou 等 [2] 提出使用泊松分布拟合时间序列在各个时间戳的依赖性概率. 基于模型的方法需要已知全部或部分被监控系统的先验知识, 通过判断先验知识是否满足实现异常检测, 不适用于模型未知、不可预测和高度可变的现实系统.

基于机器学习的方法认为异常数据在某些方面孤立于主体数据, 通常包括基于距离、聚类的方法等. 距离法把与相邻数据点的距离过大的数据点看作异常数据. Qiu 等 [3] 提出了使用主成分分析方法构建检测模型, 通过正常和异常实例之间的距离判断异常是否发生. 基于聚类的方法把数据分类, 正常数据实例属于数据中的一个集群, 而异常则不属于任何集群, 通过估计每个数据点与聚类的关系判断异常. Li 等 [4] 使用扩展模糊 C 均值聚类算法拟合多变量时序中的可用结构, 检测多变量时间序列的振幅和形状的异常. 基于机器学习的方法以相对简单的方式对传感器之间的相互关系进行建模, 通过样本内在的属性判断是否异常, 难以综合考虑样本之间的关系, 对数据的趋势变化不敏感, 难以处理时序数据. 因此, 在处理具有复杂非线性关系的现实系统时, 基于机器学习的方法存在很大的局限 [10, 23].

近些年随着神经网络的发展, 许多学者提出了基于 AE, RNN, GAN 和 LSTM 的异常检测模型. 基于 AE 的方法通过编码器学习时序数据的表达, 再通过解码器重构尽可能接近原始输入的数据, 学习时序数据值编码. Zong 等 [7] 利用深度 AE 为每个输入数据点生成低维表示和重建误差, 将误差送到高斯混合模型中, 同时联合优化深度自动编码器和混合模型的参数, 构建出深度自动编码高斯混合模型 DAGMM. Abdulaal 等 [9] 提出对预训练的 AE 的潜在表示进行频谱分析, 并提取出主要频率输入到后续网络中, 学习得到原始多变量的同步表示, 得到异常检测模型 RCoders-RSCoders. 基于 RNN 或 LSTM 的方法利用 RNN 提取前后时间戳的信息, 捕捉时序数据的特征. Su 等 [8] 提出一种用于多变量时间序列异常检测的随机循环神经网络 OmniAnomaly, 该模型利用随机变量连接和平面归一化流算法学习多变量时序的鲁棒正态表示, 通过表示重构输入数据并利用重构概率确定异常. 基于 GAN 的方法通过学习数据的分布对多变量时序进行建模. 李丹等 [11] 提出了一种用 LSTM-RNN 作为 GAN

的生成器和判别器的无监督多变量异常检测方法 MAD-GAN. Deng 等 [12] 提出了一种时空图卷积对抗网络进行交通异常情况检测, 利用图卷积门控循环单元训练生成器和判别器, 经过对抗训练后, 生成器和判别器可以独立用作检测器, 其中生成器对正常的交通动态模式进行建模, 判别器提供随时空特征变化的检测标准; Yoo 等 [13] 提出了由卷积时空存储器和时空注意机制组成的卷积循环重建网络, 对锡膏打印机异常检测.

最近, 一些研究将 GNN 用于多变量时间序列的异常检测: 如蔡美玲等 [10] 引入了图注意力层以自动学习变量间的复杂依赖关系, 将考虑了耦合特征的时序输入到嵌入 Transformer 的 GAN 中进行异常检测; Deng 等 [14] 提出的 GDN 将学习到的图结构与多变量时序同时输入到多层图注意网络, 得到下一时刻的各个变量的预测值, 考虑预测值与真实值间的 L_2 距离判断异常是否发生; 与 GDN 类似, Zhao 等 [15] 提出的 MTAD-GAT 使用两个并行的图注意层以学习多变量时序在时空两维度的复杂依赖关系, 比较预测值与真实值以检测异常.

2.2 图神经网络

图数据可以有效、直观地表达节点间的关系. 近年来, GNN 作为一类处理图数据的神经网络, 尤其是图卷积网络在动作识别 [24]、时序预测 [25] 和故障诊断 [26] 等领域被广泛应用. 虽然图卷积网络 (graph convolution network, GCN) 可以更精确地对变量对之间的相关性进行建模, 如 Feng 等 [25] 使用扩散图卷积网络, 利用炼钢过程中元素浓度之间的相关性来准确预测终点成分, 但没有充分考虑时序数据的非线性时序特性. 时空图卷积网络 (spatio-temporal graph convolution network, ST-GCN) 作为一种引入了时序特征分析的 GNN, 能够同时提取时空两域的特征信息, 更加适合处理多变量时序数据 [17,27~29]. 目前 ST-GCN 一般采用 GCN 或 GAT 提取空间依赖, 使用 CNN (convolutional neural network)/RNN 获取时域特征.

由于使用 GCN 或 GAT 提取空间依赖特征时, 需要明确的图结构. 因此, 当使用 GCN 或 GAT 对多变量时序进行处理时, 产生了一些多变量时序图建模方法. 一些图建模方法直接生成一个全连接的图 [10,15], 即任意两变量间均存在连边. 而全连接的图结构扭曲了真实的网络拓扑结构, 会为后续的 GNN 或 GAT 学习过程提供误导信息 [19]. 鉴于此, 最近一些方法使用 ϵ- 半径 [19] 或 k- 近邻 [17] 的方法生成简洁的图结构: 其中, k- 近邻图结构生成算法选择与节点 v_i 最邻近的 k 个邻居节点进行连接, 如 Deng 等 [14] 和 Wu 等 [18] 使用 k- 近邻的方法对多变量时序进行图建模; ϵ- 半径图结构生成算法选择与节点 v_i 距离小于 ϵ 的所有邻居建立连边, 如 Khodayar 等 [17] 使用互信息计算任意两变量间的相似性, 使用 ϵ- 半径算法选择距离每个变量小于阈值半径 ϵ 的节点进行连接, 从而构建出无向图结构.

3 所提模型

本文提出的 HST-GNN 模型的任务是提取 t 时刻前长度为 h 的历史数据 $\boldsymbol{X}_t = [\boldsymbol{x}_{t-h+1}, \ldots, \boldsymbol{x}_t]$ 的深度时空特征及其对应的系统级结构特征. 深度时空特征由时空深度特征提取模块进行提取, 该模块以 $\boldsymbol{X}_t$ 为输入, t 时刻后长度为 h 的数据 $\boldsymbol{X}_p = [\boldsymbol{x}_t, \ldots, \boldsymbol{x}_{t+h-1}]$ 为输出, 训练学习多变量时序数据的深度时空耦合特征, 作为传感器结构特征. 图结构特征由图结构特征提取模块进行提取, 该模块以当前输入数据 $\boldsymbol{X}_t$ 的图结构为输入, 提取其平均路径长度 $\overline{D}(\boldsymbol{G})$、聚类系数 $\mathrm{CC}(\boldsymbol{G})$、密度 $\rho\boldsymbol{G}$、节点度熵 $\mathrm{NDE}(\boldsymbol{G})$ 作为其系统级特征. 然后将传感器级特征和系统级特征串联起来得到串联特征 $[\boldsymbol{x}_t, \boldsymbol{x}_{t+1}, \boldsymbol{x}_{t+h-1}, \overline{D}(\boldsymbol{G}), \mathrm{CC}(\boldsymbol{G}), \rho(\boldsymbol{G}), \mathrm{NDE}(\boldsymbol{G})]$, 送到 AE 中进行拟合学习. 如果重建数据和采集到的真实系统运行数据一样

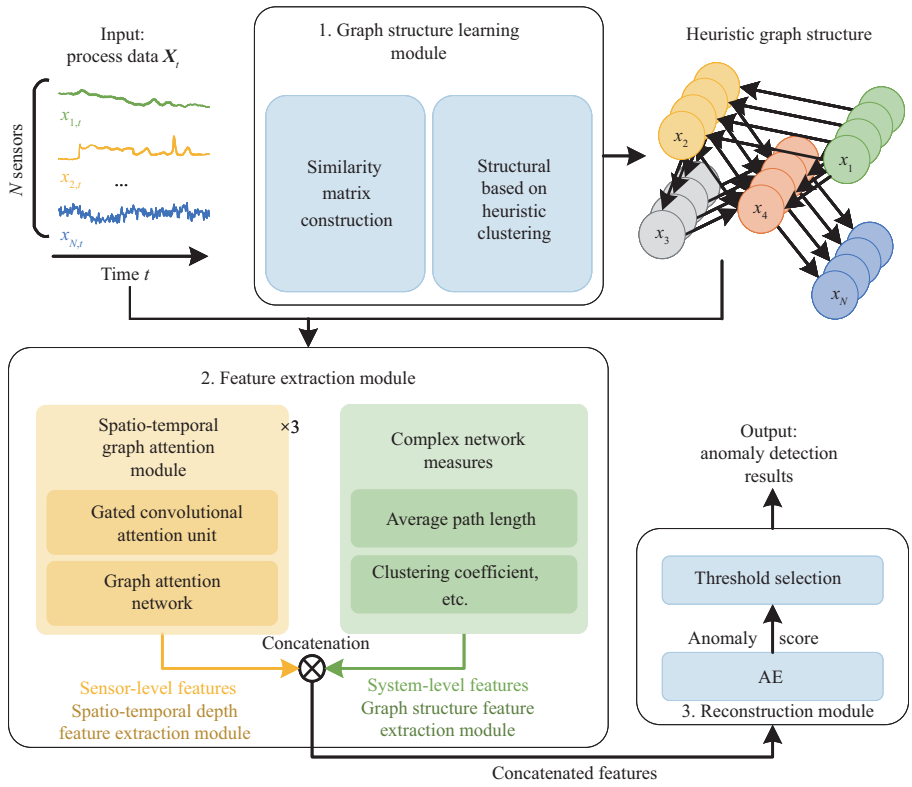

图 1　(网络版彩图) HST-GNN 框架示意图

Figure 1　(Color online) HST-GNN framework

或者接近, 表明 t 时刻系统处于正常运行状态, 否则表明 t 时刻系统处于异常工作状态. 模型的输出 output 是一组二进制标签, 表示每个测试时间是否异常. 当 $\text{output}(t)=1$ 时表示时间 t 发生异常, 否则, 表示系统运行状态正常.

3.1　模型描述

本文提出的端到端的异常检测算法 HST-GNN 包括 3 个主要组成部分, 如图 1 所示.

(1) 图结构学习模块: 通过构建的有向相似性函数和设计的基于启发式聚类的图结构学习算法, 学习变量间的图结构.

(2) 特征提取模块: 该模块包含时空深度特征提取模块和图结构特征提取模块. 时空图注意力网络预测模块由输入层、3 个本文设计的 ST-GAM 模块和一个全连接层构成. 时空深度特征提取模块以学习到的图结构和多变量时序数据 $\boldsymbol{X}_t$ 为输入, 输出为当前数据 $\boldsymbol{X}_t$ 的时空深度耦合特征. 图结构特征提取模块以当前输入数据 $\boldsymbol{X}_t$ 的图结构为输入, 提取两者的平均路径长度、聚类系数等复杂网络测度作为其图结构特征.

(3) 基于 AE 的重构模块: 使用多层感知机 (multilayer perceptron, MLP) 作为编码器和解码器. 以时空深度耦合特征与图结构特征串联后的特征时序作为编码器的输入, 将串联特征压缩为低维向量, 随后将低维向量传递给解码器重构出输入数据, 使输出与输入尽量相同, 通过重构损失判别系统是否发生异常, 给出检测结果.

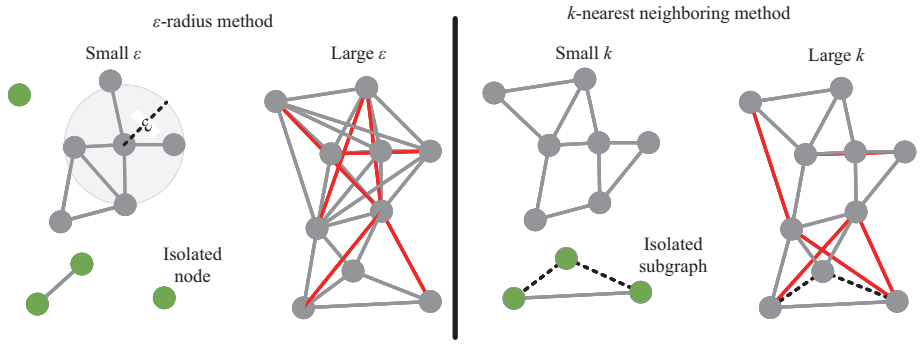

图 2 (网络版彩图) 基于 ϵ- 半径或 k- 近邻的图结构学习算法的不足

Figure 2 (Color online) Disadvantages of graph structure learning algorithms based on ϵ-radius and k-neighbor

3.2 图结构学习模块

本小节使用图结构学习模块将多变量时序表示为图 $\boldsymbol{G}=(\boldsymbol{V},\boldsymbol{E},\boldsymbol{W})$, 其中, $\boldsymbol{V}$ 为节点集合, 每一个节点代表一个变量, $\boldsymbol{E}$ 是边集, 每条边表示边两端的节点间存在耦合关系, $\boldsymbol{W}$ 为权重集, 每个权重代表每一条边上的耦合强度. 图结构学习模块包括相似矩阵构建和结构学习两部分.

3.2.1 相似矩阵构建

相似矩阵 $\boldsymbol{S}$ 或距离矩阵 **Dis** 构建函数是量化两两变量之间的相似性或差异性的函数. 本文提出了一个有向相似度计算函数 S 表示变量间耦合关系的强弱:

$$S(\boldsymbol{X}_t)=\mathrm{SoftPlus}(\tanh(\boldsymbol{X}_t\boldsymbol{W}_1)\tanh(\boldsymbol{W}_2\boldsymbol{X}_t)-\tanh(\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{X}_t^{\mathrm{T}})\tanh(\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{W}_2^{\mathrm{T}})), \tag{1}$$

其中, $\boldsymbol{X}\in\mathbb{R}^{h\times N}$ 表示输入到网络的 N 个节点 t 时刻前长度为 h 的历史数据, $\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$ 是需要学习的模型参数. 针对矩阵 $\tanh(\boldsymbol{X}_t\boldsymbol{W}_1)\tanh(\boldsymbol{W}_2\boldsymbol{X}_t)-\tanh(\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{X}_t^{\mathrm{T}})\tanh(\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{W}_2^{\mathrm{T}})$ 中小于零的元素, SoftPlus 则将其映射成接近于 0 的值, 否则映射成大于 1 的值. 因此, 保证了相似性矩阵 $\boldsymbol{S}$ 的有向性, 即节点 v_i 到节点 v_j 与节点 v_j 到节点 v_i 的耦合强度不同. 距离矩阵 **Dis** 与相似矩阵 $\boldsymbol{S}$ 的转换公式 [19] 为

$$\mathrm{dis}_{ij}=\sqrt{s_{ii}+s_{jj}-s_{ij}}, \tag{2}$$

其中, s_{ij} 表示节点 v_i 和 v_j 间的相关性, dis_{ij} 表示节点 v_i 和 v_j 间的距离.

3.2.2 图结构学习

图结构学习算法根据距离矩阵元素 dis_{ij} 的大小, 判断节点 v_i 和 v_j 之间是否需要添加连接边. 典型的图结构学习算法包括 k- 近邻和 ϵ- 半径算法. 它们使用统一的标准处理图结构的密集和稀疏区域, 这将导致最后学习得到的图结构中存在孤立的节点或子图, 或者得到冗余的图结构, 如图 2 所示.

较低的阈值 ϵ 和邻居数 k 可能会造成孤立节点或子图 (图中标绿), 导致必要的耦合关系丢失; 而为保证连通性, 设置较大的 k 和阈值 ϵ 将连接耦合关系弱的节点 (图中标红), 会导致图结构冗余. 针对上述问题, 本文提出基于启发式聚类的图结构学习算法. 使用该算法可以构建出一个连通、稀疏的图结构, 并保持原始数据集的集群性, 其步骤如算法 1 所示. 其中步骤 4 的目的是确保最终网络是全连通的.

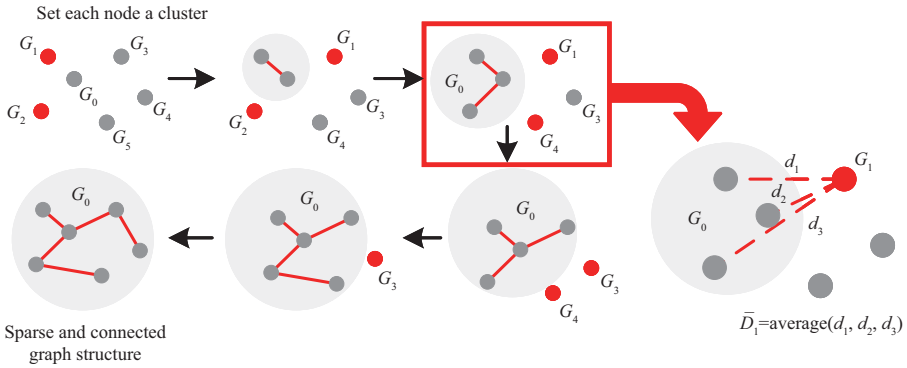

图 3 (网络版彩图) 基于启发式聚类的图结构构建流程示意图

Figure 3 (Color online) Graph structure construction process based on heuristic clustering

算法 1 Graph structure learning algorithm based on heuristic clustering

输入: Input batch data $\boldsymbol{X}_t \in \mathbb{R}^{N\times n\times TS}$;
输出: Graph structure $\boldsymbol{A}_t \in \mathbb{R}^{N\times N}$;
count $\Leftarrow 1$; $\boldsymbol{A}_t \Leftarrow \boldsymbol{O} \in \mathbb{R}^{N\times N}$;
Calculate $\boldsymbol{S}(\boldsymbol{X}_t)$;
Calculate **Dis**$(\boldsymbol{X}_t)$;
while count of cluster >1 **do**
  Identify the two clusters G_1 and G_2 closest to G_0;
  Calculate the average distances $\overline{D}_1$ and $\overline{D}_2$;
  for i in 1: 1: count **do**
    if $\boldsymbol{d}_i \leqslant d_{\text{thre}}$, $d_{\text{thre}} = \max(\overline{D}_1,\overline{D}_2)$ **then**
      $\boldsymbol{a}_{i,0} \Leftarrow 1$;
      Add G_i to G_0;
      count $\Leftarrow$ count $+ 1$;
      count of cluster $\Leftarrow$ count of cluster $- 1$;
    end if
  end for
  Updata **Dis**;
end while

以 $k = 2, 6$ 个节点为例构建其图结构, 其构建流程示意图如图 3 所示. 初始化一个完全断开的图结构, $\boldsymbol{A}_t = \boldsymbol{O} \in \mathbb{R}^{N\times N}$, 其中 $\boldsymbol{O}$ 为全零矩阵, 此时一个节点就是一个簇. 根据式 (1) 计算得到相似性矩阵, 并使用式 (2) 将相似性矩阵转化为距离矩阵 **Dis**. 根据距离矩阵识别出距离簇 G_0 最近的两个簇 G_1 和 G_2 并将其标红, 计算簇 G_1 和 G_2 相对于簇 G_0 的平均相异度 $\overline{D}_1$ 和 $\overline{D}_2$, 比较 G_1 与 G_0 中各个节点的距离与距离阈值 $\max(\overline{D}_1,\overline{D}_2)$ 间距离的大小, 若小于等于该阈值, 则将对应的两节点相连接, 重复上述步骤, 直至所有变量节点均属于同一个簇 G_0. 最终得到稀疏且连通的图结构.

3.3 特征提取模块

3.3.1 时空深度特征提取

(1) 时间卷积层. 本文提出使用 GCAU 作为 ST-GAM 模块中的时间卷积层, 其组成如图 4 所示. GCAU 将 CA 与 GAU 相结合以自适应地捕获各时间戳之间的相关性. CA 使用长度大于 1 的一维卷

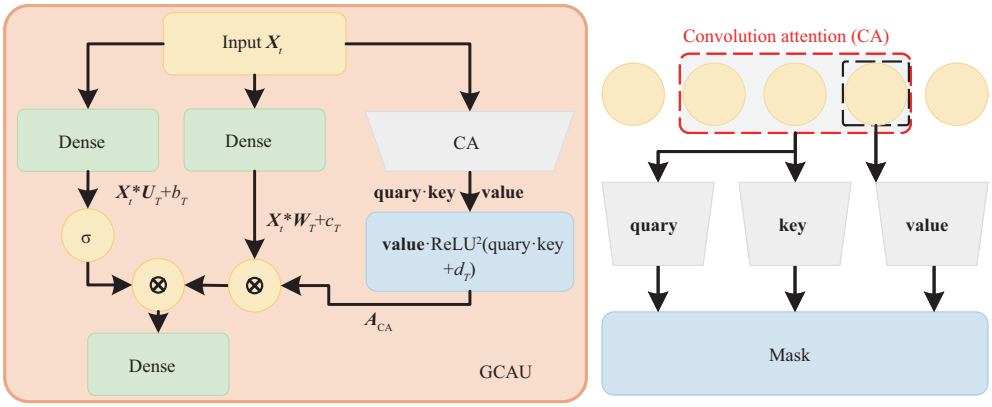

图 4 (网络版彩图) GCAU 的结构示意图

Figure 4 (Color online) Structure of GCAU

积核计算变量查询向量 **quary**、键 **key** 和值 **value**, 使得注意力集中在局部上下文的变化上, 从而可以匹配更多相关的特征. GCAU 层的计算公式为

$$h_{\mathrm{GCAU}}(\boldsymbol{X}_t) = (\boldsymbol{A}_{\mathrm{CA}} \otimes (\boldsymbol{X}_t\boldsymbol{W}_T) + \boldsymbol{c}_T)) \otimes \sigma(\boldsymbol{X}_t\boldsymbol{U}_T + \boldsymbol{b}_T), \tag{3}$$

其中, $\boldsymbol{X}_t$ 和 $h_{\mathrm{GCAU}}(\boldsymbol{X}_t)$ 是 GCAU 的输入和输出, 选择对于突变更为敏感的加权线性单元作为激活函数 σ, $\boldsymbol{W}_T$, $\boldsymbol{U}_T$, $\boldsymbol{b}_T$ 和 $\boldsymbol{c}_T$ 是可学习参数, $\otimes$ 是哈达玛积. $\boldsymbol{A}_{\mathrm{CA}} \in \mathbb{R}^{n\times n}$ 是根据输入 $\boldsymbol{X}_t$ 计算出的卷积注意力:

$$\boldsymbol{A}_{\mathrm{CA}} = \frac{1}{H_{\mathrm{CA}}}\mathbf{value} \cdot \mathrm{ReLU}^2(\mathbf{quary} \cdot \mathbf{key} + \boldsymbol{d}_T), \tag{4}$$

其中, H_{CA} 是注意力的头数, $\mathbf{quary} = \mathbf{key} = \sum_{i=1}^{k-1} f(i) \cdot \boldsymbol{x}_{s-i}$, $V = \boldsymbol{X}_t\boldsymbol{W}_{\mathrm{aff}}$, $\boldsymbol{x}_{s-i}$ 为时序数据 $\boldsymbol{X}_t$ 中位置为 $s-i$ 的元素, $f: 0, 1, \ldots, k-1 \Leftarrow R$ 为卷积核, k 为卷积核的大小, $\boldsymbol{d}_T$ 是可训练的参数, 如图 4 中红色虚线框所示, $\boldsymbol{W}_{\mathrm{aff}}$ 为仿射变换矩阵, 其计算过程沿用 Transformer 中 **value** 的计算方法.

(2) 空间卷积层. 为了充分利用节点及其邻域的信息, 本文使用多头自注意力来构建 GAT 层. GAT 可以同时建模长期依赖和短期依赖, 并行计算不同子空间的注意力, 完成时空深度特征的提取, 如图 5 所示. 具体计算公式如下:

$$h_1^{k+1} = \|_{\mathrm{head}=1}^{H} \sigma\left(\sum_{j\in N(v_i)} \alpha(\boldsymbol{h}_i^k, \boldsymbol{h}_j^k)\boldsymbol{h}_j^k\right), \tag{5}$$

其中, $\|$ 为矩阵拼接操作, $\boldsymbol{h}_i^k$ 和 $\boldsymbol{h}_j^k$ 是第 k 层节点 v_i 和 v_j 的特征变量, α 为注意力函数. 多头图注意力机制对于同一个节点, 分别计算 H 次注意力, 并以矩阵拼接的方式合并 H 次注意力矩阵. 注意力函数的计算公式如下:

$$\mathrm{att}_{ij} = \mathrm{LeaklyReLU}(\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{W}_G x_{i,t} \parallel \boldsymbol{W}_G x_{j,t})), \tag{6}$$

$$\alpha_{ij} = \frac{\exp(\mathrm{att}_{ij})}{\sum_{k\in N(i)} \exp(\mathrm{att}_{ij})}, \tag{7}$$

其中, $\boldsymbol{e}^{\mathrm{T}}$ 和 $\boldsymbol{W}_G$ 是可学习的参数, $x_{i,t}$ 是 t 时刻节点 v_i 的特性向量, $x_{j,t}$ 是 t 时刻节点 v_j 的特性向量, $N(i)$ 表示节点 v_i 的邻居数量.

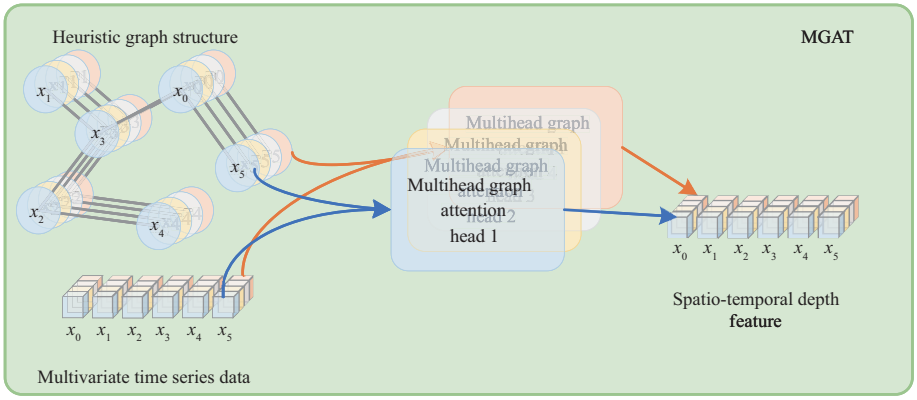

图 5　(网络版彩图) MGAT 的结构示意图

Figure 5　(Color online) Structure of MGAT

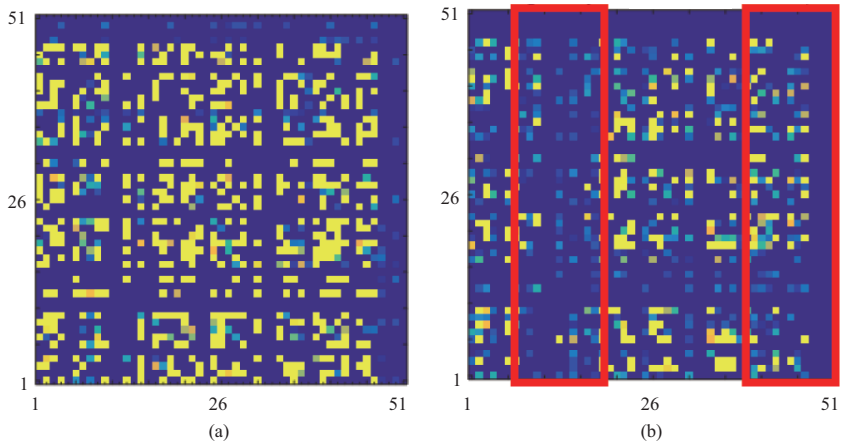

图 6　(网络版彩图) 安全水处理数据图结构对应的相似性矩阵的热图

Figure 6　(Color online) Heat map of the similarity matrix corresponding to the graph structure of SWaT. (a) Normal operating condition; (b) abnormal operating condition

3.3.2　图结构特征提取

以工业水处理厂数据集 SWaT [30] 为例, 本文使用相似性函数 S 计算在正常工况和异常工况下变量间耦合关系, 并进行热图可视化, 如图 6(a) 和 (b) 所示. 安全水处理数据集共有 51 个变量, 因此, 图 6(a) 和 (b) 显示的热图大小为 51×51. 热图中越亮的位置, 代表对应的两变量之间的耦合关系越强. 在系统发生异常时, 传感器 6~16, 41~51 与其他传感器间的关联关系极大减弱 (红框框出部分). 对比正常运行与异常运行时的热图可知, 在两种工况下系统传感器之间的耦合关系是明显不同的.

由此可知, 相比于传感器级的偏差, 在节点变量 v_i 发生异常的早期, v_i 与其他变量间的相关性会发生快速且明显变化. 因此, 系统级的图结构特征得分能更快体现出异常的发生.

基于此, 本文选择计算时空图注意力网络提取的深度特征对应的图结构的复杂网络测度作为图结构特征, 设计了图结构特征提取模块. 该模块计算 t 时刻输入到网络的观测值 $\boldsymbol{Y}$ 的图结构 $\boldsymbol{G}$ 的平均路径长度 $\overline{D}(\boldsymbol{G})$ [31]、聚类系数 $\mathrm{CC}(\boldsymbol{G})$ [32]、密度 $\rho(\boldsymbol{G})$ [33]、节点度熵 $\mathrm{NDE}(\boldsymbol{G})$ [34] 等复杂网络测度作

为 t 时刻的图结构特征:

$$\mathrm{CNM}(\boldsymbol{G}) = \|_{\mathrm{head}=1}^{4} \left(\overline{D}(\boldsymbol{G}), \mathrm{CC}(\boldsymbol{G}), \rho(\boldsymbol{G}), \mathrm{NDE}(\boldsymbol{G})\right), \tag{8}$$

其中, $\overline{D}(\boldsymbol{G})$, $\mathrm{CC}(\boldsymbol{G})$, $\rho(\boldsymbol{G})$, $\mathrm{NDE}(\boldsymbol{G})$ 代表了图结构的网络传输效率、聚集强度、节点连接强度、节点度数的不成比例程度, 用于量化系统级结构特征, 具体定义式如下.

(1) 平均路径长度. 平均路径长度 APL 定义为图 $\boldsymbol{G}$ 上两个节点之间距离的平均值, 其表达式为

$$\overline{D}(\boldsymbol{G}) = \frac{1}{\frac{1}{2}N(N-1)} \sum_{i \geqslant j} d_{ij}, \tag{9}$$

其中, N 是节点数. 网络的 $\overline{D}(\boldsymbol{G})$ 主要用于衡量网络的紧凑性以及相互通信的效率和性能. $\overline{D}(\boldsymbol{G})$ 越集中和紧凑, 通信效率就越高.

(2) 平均聚集系数. 平均聚类系数 $\mathrm{CC}(\boldsymbol{G})$ 可以用图 $\boldsymbol{G}$ 中所有节点的局部聚类系数的平均值来表示:

$$\mathrm{CC}(\boldsymbol{G}) = \frac{1}{N} \sum_{i}^{N} \mathrm{CC}(i) = \frac{1}{N} \sum_{i}^{N} \frac{e_i}{k_i(k_i-1)/2}, \tag{10}$$

其中, $\frac{k_i(k_i-1)}{2}$ 是 k_i 个相邻节点之间可能的边数, e_i 是节点 v_i 的 k_i 个相邻节点之间的实际边数. $\mathrm{CC}(\boldsymbol{G})$ 是整个网络中相互连接的邻居节点的平均比例, 它揭示了网络图的局部连通性属性. 如果 $\mathrm{CC}(\boldsymbol{G}) = 1$, 则表示网络是全连接的. 如果 $\mathrm{CC}(\boldsymbol{G})$ 接近于 0, 则表示网络的结构比较松散.

(3) 密度. 网络密度 $\rho(\boldsymbol{G})$ 是网络中实际边数与可容纳边数上限的比值, $\rho(\boldsymbol{G})$ 越大, 网络结构越密集, 其表达式为

$$\rho(\boldsymbol{G}) = \frac{M}{N(N-1)}, \tag{11}$$

其中, N 是图上节点的数量, M 是图上边的数量.

(4) 节点度熵. 基于熵的复杂网络度量可以描述网络上每条边的影响和网络中节点的特征. 网络熵通过对节点和链路等网络参数创建概率度量来定义节点的重要程度, 节点度熵 $\mathrm{NDE}(\boldsymbol{G})$ 的定义如下:

$$\mathrm{NDE}(\boldsymbol{G}) = -\sum p_i \log p_i = -\sum \frac{\sum_{j=1}^{N} \mathbf{De}_j}{\mathbf{De}_i} \log \frac{\sum_{j=1}^{N} \mathbf{De}_j}{\mathbf{De}_i}, \tag{12}$$

其中, $\mathrm{NDE}(\boldsymbol{G})$ 是网络的平均信息量的度量, p_i 是节点 v_i 的度数的概率描述, $\mathbf{De}$ 为节点 v_i 的度.

3.4 基于 AE 的重构模块

遵循典型的 AE 框架 [35], 本文选择多层感知器以构建编码器和解码器, 如图 7 所示. AE 旨在学习以尽可能低的误差重建输入观测值, 训练好的 AE 能够生成真实样本, 捕捉了系统在正常状态下的数据分布. 具体公式如下:

$$\mathrm{argmin}_{f,g} \langle [\delta(\boldsymbol{X}_f, D(E(\boldsymbol{X}_f; \boldsymbol{W}_e); \boldsymbol{W}_d))] \rangle, \tag{13}$$

其中, $E(\boldsymbol{X}_f; \boldsymbol{W}_e)$ 为编码器, $\boldsymbol{X}_f$ 是编码器的输入, 为前一步特征提取模块提取出的特征, $D(\boldsymbol{E}; \boldsymbol{W}_d)$ 是解码器, $\boldsymbol{E}$ 为编码向量, $\boldsymbol{W}_e$ 和 $\boldsymbol{W}_d$ 为多层感知器的可学习参数.

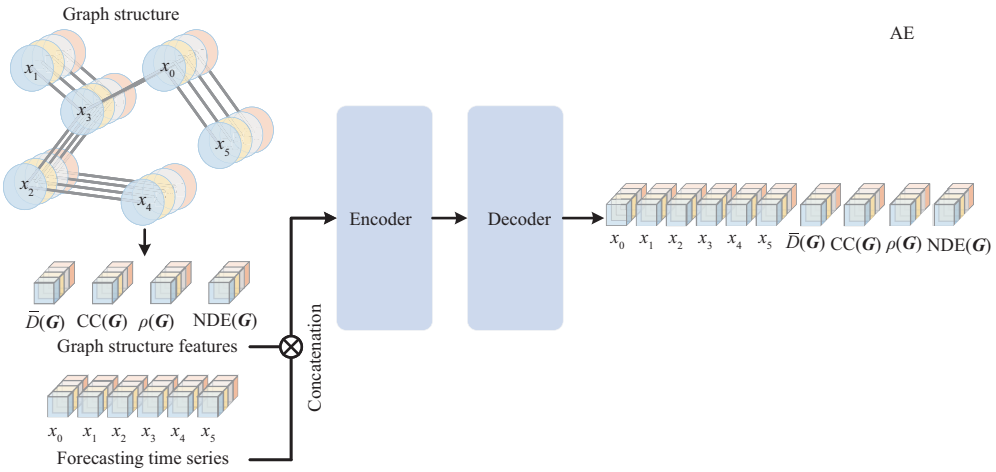

图 7 (网络版彩图) AE 的结构示意图

Figure 7 (Color online) Structure of AE

3.5 模型训练

在模型训练与测试中使用 CUDA 11.3 的 Pytorch 1.12.0 版本实现该方法. 滑窗大小为 24, 步幅为 4, 模型输入时序长度为 64, GAT 的维数为 2, 头的数目为 6, AE 有 3 个 MLP 层构成的编码器和解码器. 此外还应用了 Dropout 策略, 将 Dropout 率设为 0.3. 学习率设为 1E−3, epoch 设为 500, batch size 设为 64. 利用 Adam 优化算法对每个权重参数获取网络误差的梯度 [36], 通过参数更新过程得到新的权重. 用于模型训练的损失函数 [11] 定义如下:

$$\text{Loss} = \sum |\boldsymbol{X}_t - \boldsymbol{X}_p| + \sum |\boldsymbol{X}_f - D(\boldsymbol{X}_f)|, \tag{14}$$

其中, $\boldsymbol{X}_t - \boldsymbol{X}_p$ 是特征提取模块的预测误差, 用于训练特征提取模块, $\boldsymbol{X}_f - D(\boldsymbol{X}_f)$ 为解码器的重建误差, 用于训练基于 AE 的重构模块, $\boldsymbol{X}_f$ 为特征提取模块提取出的串联特征.

在测试过程中, 本文使用 Loss 作为各个时间戳的异常得分, 如果得分超过固定阈值 A_T, 则将该时间标记为异常. 基于极值理论的选择 POT 方法 (peak over threshold) 在原始数据分布非常复杂的情况下, 不需要进行数据分布的假设, 仍然有可能估计出异常事件的发生. 因此, 本文选择 POT 方法进行异常阈值的选择, 并使用带参数的广义帕累托分布 (generalized Pareto distribution, GPD) 拟合极值概率分布的尾部, 进行阈值 A_T 的选择 [8]. 具体公式如下所示:

$$A_T \simeq \text{th} - \frac{\hat{\beta}}{\hat{\gamma}} \times \left(\left(\frac{qh}{n_{\text{th}}} \right)^{-\hat{\gamma}} - 1 \right), \tag{15}$$

$$F(\text{sc}) = P(\text{th} - \text{Sc} | \text{Sc} > \text{th}) \sim \left(1 + \frac{\gamma s}{\beta} \right)^{-\frac{1}{\gamma}}, \tag{16}$$

其中, Sc 是某时刻的异常得分, th 是初始设定的异常得分阈值, β 和 γ 是 GPD 分布的形状和尺度参数, $\hat{\beta}$ 和 $\hat{\gamma}$ 是由极大似然方法估计出的 β 和 γ 的近似, q 是异常得分小于阈值的期望概率, sc 是 GPD 分布的低分位数, h 是异常得分的数据长度.

表 1 数据集的基本统计数据

Table 1 Basic statistics of four datasets

Dataset	Variable number	Training set size	Testing set size	Abnormal ratio (%)
PSM	25	105984	26497	27.80
SMD	12×38	28479	28479	4.16
SWaT	51	496800	449919	11.98
WADI	123	1048571	172801	5.99

4 实验验证

本节将验证所提出的网络 HST-GNN 的检测性能. 所有实验均在 Windows 系统上进行 (CPU: Intel(R) Core (TM) i9-10900K CPU @ 3.70 GHz, GPU: NVIDIA GeForce RTX 3090).

4.1 数据集

本文使用的 4 个数据集包括来自网络服务器的数据集 PSM [9] 和 SMD [37], 来自工业水处理厂的数据集 SWaT [30] 和 WADI [38]. 数据集的基本统计数据列于表 1 中. SWaT 和 WADI 数据集中的训练集只包含正常数据, 测试数据集则包含正常和异常两种状态的数据集. PSM 和 SMD 数据集中训练集可能包含少量异常数据, 测试数据集则包含正常和异常两种状态的数据集. 本文将训练集的前 80% 作为模型的训练集, 将训练集的后 20% 作为模型的验证集. 值得注意的是, SMD 中部分机器在数据采集期间发生了服务变化, 导致训练和测试数据遵循不同的正常模式. 因此, 本文只使用 SMD 数据集中 12 个不受漂移影响的机器的数据. 后续在比较模型性能时, 本文将计算模型在 12 个 SMD 机器数据上的平均性能指标作为模型的最终表现.

4.2 指标

本文使用精度 (precision, Prec), 召回率 (recall, Rec) 和 F1 指标 (F1-Score, F1) 评价模型的检测性能. Prec 表示所有被预测为正的样本中实际为正的样本的概率, Rec 表示实际为正的样本中被预测为正样本的概率, F1 是精确率和召回率的调和平均, F1 越高所评价模型的检测性能越高.

$$\mathrm{Prec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{17}$$

$$\mathrm{Rec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{18}$$

$$\mathrm{F1} = \frac{2 \times \mathrm{Prec} \times \mathrm{Rec}}{\mathrm{Prec} + \mathrm{Rec}}, \tag{19}$$

其中, TP, FP 和 FN 是真阳性、假阳性和假阴性的数量.

4.3 实验结果

为了证明 HST-GNN 的有效性, 本文将其与 7 个多变量时序无监督异常检测方法进行比较, 所有比较的方法都使用滑动窗口输入, 对比模型的简要描述在表 2 中列出.

表 3 列出了上述 8 个模型在 4 个数据集上的检测结果. 最优结果加粗并加下划线, 次优结果加粗. HST-GNN 在 4 个数据集上的 F1 都在 0.85 以上, 优于所有对比模型. 由于 WADI 的变量数目较多且异常率较低, 相比而言数据更为不平衡, 因此大多数多变量时序无监督异常检测模型在 WADI 上的检

表 2　对比模型描述

Table 2　Description of the comparison model

Model	Descriptions
AE [1]	An autoencoder-based model detecting anomalies via reconstruction error
DAGMM [7]	A density-based Gaussian model detecting anomalies via the estimated density
MAD-GAN	A GAN-based model detecting anomalies via reconstruction and identification errors
OmniAnomaly [8]	An RNN-based model detecting anomalies via distribution fitting
RCoders-RSCoders [9]	An autoencoder-based detecting anomalies via reconstruction error
GDN [12]	A GNN-based model detecting anomalies via prediction error
MTAD-GAT [15]	A GNN-based model detecting anomalies via prediction error

表 3　8 个异常检测模型在 4 个公共数据集上的结果

Table 3　Results of eight anomaly detection models on four public datasets

Model	PSM			SMD			SWaT			WADI		
	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
AE	0.6679	0.8285	0.7347	0.8825	0.8037	0.8280	0.7263	0.5263	0.6100	0.3435	0.3435	0.3400
DAGMM	**0.9349**	0.7003	0.8206	0.6730	0.4989	0.5730	0.2746	0.6952	0.3900	0.5444	0.2699	0.3600
MAD-GAN	0.7032	0.8382	0.7700	0.8966	0.7914	0.8200	**0.9897**	0.6374	0.7700	0.4144	0.3392	0.37006
OmniAnomaly	0.8839	0.7446	0.8083	0.8368	**0.8682**	**0.8522**	0.8142	0.8430	**0.8283**	**0.9947**	0.1298	0.2296
RCoders-RSCoders	0.9016	**0.8873**	**0.8924**	0.5122	0.6692	0.4769	0.2248	**0.9264**	0.2752	0.3749	0.3548	0.3481
GDN	0.9232	0.8582	0.8766	0.6658	0.6151	0.6301	**0.9935**	0.6812	0.8100	**0.9750**	0.4019	0.5700
MTAD-GAT	0.5869	0.7525	0.6947	**0.9019**	0.8316	0.8518	0.9342	0.6774	0.7900	0.7896	**0.8533**	**0.8200**
HST-GNN	**0.9759**	**0.9634**	**0.9696**	**0.9996**	**0.9999**	**0.9998**	0.8998	**0.9896**	**0.9357**	0.8597	**0.8787**	**0.8640**

测结果相较于其他数据集稍显逊色. 但是, HST-GNN 在该数据集上取得了最优结果. 这些实验结果证明了 HST-GNN 具有一定的优越性. 接下来, 我们具体分析 HST-GNN 能够取得更好结果的原因.

对多变量时序进行异常检测时, 由于当前观测数据与历史数据存在依赖关系, 因此我们在构建模型时应该考虑时间维度的相关性. 由于 DAGMM [7] 没有考虑时间维度的信息, 因此其整体性能最差. GDN [12] 同样也没有利用时间维度的相关性, 但由于图网络能够充分拟合变量间的耦合关系, 因此模型在 WADI 与 PSM 数据集上能够取得次优的结果. 这说明变量间的耦合关系的明确, 对于多变量时序异常检测的成功至关重要. 因此, GDN 和 MTAD-GAT 与本文提出的 HST-GNN 能够在 4 个数据集上取得较好的结果. 但由于 GDN 和 MTAD-GAT [15] 在学习图结构时为了连通度牺牲了稀疏度, 这导致图结构中存在较多扭曲的耦合关系, 在后续的特征提取过程中引入噪声. 在进行高维系统的异常检测时, 扭曲的图结构导致模型结果 Prec 较高, Rec 较低. 具体来说, 在对 WADI 数据集进行实验时, 模型将许多正常状态检测为异常状态, 以保证异常运行的检出. 而基于启发式聚类的图结构学习算法得到的是稀疏且连通的图结构, 因此相比于 GDN 和 MTAD-GAT, 本文提出的 HST-GNN 的检测结果中的 Rec 变高, Prec 相对变低, 此时 F1 有较为明显的提升.

与 DAGMM 相比, AE [1] 和 RCoders-RSCoders [9] 使用顺序的时序数据作为输入, 保留了时间维度的信息, MAD-GAN [11] 使用 LSTM-RNN 作为生成器和鉴别器来捕获时间序列分布的时间相关性, OmniAnomaly [8] 使用随机循环神经网络提取时间依赖性, 使得上述模型的检测性能有所提升. 但是不管输入的滑窗中是否存在异常数据, 上述模型均会执行最好的重构, 这使得当它们无法检测到系统

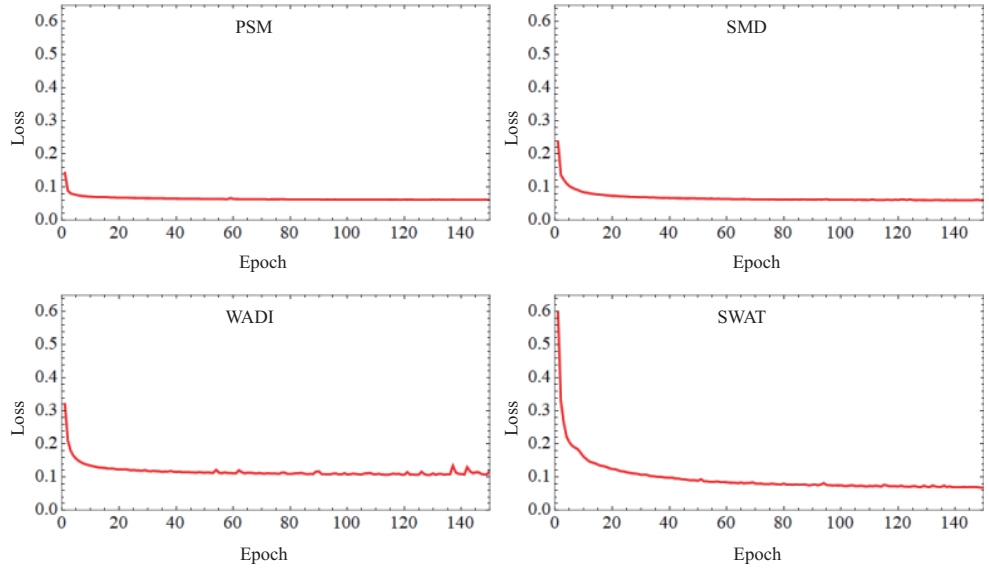

图 8 (网络版彩图) 不同数据集上 HST-GNN 完成 150 个 epoch 的损失

Figure 8 (Color online) HST-GNN losses of 150 epoches on four datasets

早期异常时, 接近正常数据的异常数据. 而本文提出的 HST-GNN 通过引入系统级特征, 允许模型能够检测到早期的异常数据, 并且通过引入基于因果卷积的 CA 使得 GAU 更容易感知到需要注意的早期异常. 因此, HST-GNN 相较于 7 个对比模型的检测精度有所提升.

4.4 模块有效性分析

为了证明使用动态相似性函数, HST-GNN 能够正常收敛. 本文针对 4 个数据集对应的训练集, 计算并画出了不同 epoch 的损失, 如图 8 所示.

为了证明图学习模块和图卷积模块、多头注意力模块、卷积自注意力和结构级偏差在 HST-GNN 模型上的有效性, 进行了消融实验. 我们将不同的模型定义如下:

(1) No-GL&GAT: 去掉图学习模块和图卷积模块.

(2) No-HC: 用完整图代替学习到的图结构.

(3) No-T: 仅使用 GLU 进行时间维度的建模.

(4) No-CA: 去掉卷积注意力, 使用自注意力.

(5) No-SDF: 不使用结构级偏差特征.

(6) No-POT: 不使用 POT 进行阈值选择, 将阈值设定为验证集中的异常得分的最大值.

对于上述模型, 各评价指标测得的测试结果如表 4 所示. 从这些实验结果可得如下结论:

(1) 当去掉图学习模块和图注意模块后, F1 分数下降超过了 0.2, 与不使用变量对耦合特征的模型 AE, DAGMM 等性能相差不大, 尤其是在变量数目较多的 WADI 数据集上, 这说明了模型中的变量间的耦合关系对于多变量时序异常检测的成功非常重要.

(2) 用完整图代替学习到的图结构后, 对于大规模数据集 WADI, HST-GNN 训练一个 epoch 所用时间由 1.22 s 增加到 2.28 s. 全连接的图结构会将不存在的关系引入图结构, 为后续耦合关系的提取带来误差, 降低模型性能.

(3) 传感器具有不同的时变特性, 仅使用 GLU 进行时间维度的建模将导致对待所有传感器使用

表 4　消融实验结果

Table 4　Ablation experiment results

Model	PSM			SMD			SWaT			WADI		
	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Complete model	0.9759	0.9634	0.9696	0.9996	0.9999	0.9998	0.8998	0.9896	0.9357	0.8597	0.8787	0.8640
No-GL&GC	0.7635	0.7940	0.7862	0.7246	0.6517	0.6839	0.6539	0.5861	0.5947	0.5004	0.3864	0.4043
No-HC	0.8629	0.8067	0.8390	0.9096	0.7905	0.8430	0.7246	0.7353	0.8020	0.7736	0.4296	0.5345
No-T	0.8042	0.8607	0.8076	0.8829	0.7994	0.8436	0.9084	0.8140	0.8310	0.7018	0.7682	0.7138
No-CA	0.8714	0.8426	0.8509	0.9072	0.7948	0.8756	0.9315	0.7286	0.8072	0.8104	0.7492	0.7493
No-SDF	0.8916	0.8934	0.8866	0.8952	0.7649	0.8167	0.9065	0.7163	0.7842	0.7636	0.7351	0.7516
No-POT	0.9585	0.8349	0.8769	0.9339	0.9083	0.9245	0.9907	0.8704	0.9125	0.8454	0.8623	0.8562

同样的时间依赖提取方法, 使模型无法充分拟合时序数据在时间维度上的非线性行为, 导致系统的漏检, 指标 Prec 相对于 Rec 的下降更明显.

(4) 去掉 GCAU 的卷积注意力, 仅使用自注意力拟合时序数据在时间维度上的非线性行为, 会导致系统无法及时感知到异常的发生, 导致 HST-GNN 模型实验结果中的 Prec 的下降, 提出卷积自注意力将能够更好地将时序上下文整合到注意力机制中, 对系统变量的突变更为敏感.

(5) 不使用结构级偏差特征作为时空深度特征的补充, 会导致系统无法及时感知到 “轻微” 异常的发生, 导致 HST-GNN 模型实验结果中的 Prec 的下降. 同时也可以说明系统正常运行和异常运行时, 系统各传感器之间的关系不同, 将其作为传感器级别特征的补充, 有利于提高模型的性能.

(6) 通过 POT 进行阈值选择后, 可以提高 Rec, 增强 HST-GNN 对于异常的检出率.

5　结论

本文提出一种基于时空图神经网络模型 HST-GNN, 用于多变量时序数据的异常检测. 我们提出了一种基于启发式聚类的图结构学习算法对多变量时序进行准确图建模, 设计了一个新的 ST-GAM, 它能同时从时间和空间两个维度对时序数据进行建模. 进一步, 提出了一种系统级图结构特征计算方法, 可以更快速地捕捉异常, 提高异常检测的召回率. 基于真实世界数据集的实验结果表明, 所提方法可以生成连通和稀疏性较强的图结构, 有利于系统变量间耦合特性的提取, 并增加了 HST-GNN 对不显著的早期异常的敏感度, 在保证检测模型的准确率的同时, 降低了模型对于早期异常的漏检率, 从而使 Prec, Rec 和 F1 指标相较于其他模型有较为明显的提升.

参考文献

1　Lv S, Zhu H. A discussion about the test method of outliers in statistical data. J Northeast Normal Univ (Natural Sci Ed), 1993, 3: 5 [吕恕, 朱宏. 统计数据中异常值的检验方法讨论. 东北师大学报: 自然科学版, 1993, 3: 5]

2　Benkabou S E, Benabdeslem K, Kraus V, et al. Local anomaly detection for multivariate time series by temporal dependency based on Poisson model. IEEE Trans Neural Netw Learn Syst, 2022, 33: 6701–6711

3　Qiu W B, Wu Y, Wang G Y, et al. A novel unsupervised anomaly detection based on robust principal component classifier. In: Proceedings of SPIE, Kissimmee, 2003

4　Li J B, Izakian H, Pedrycz W, et al. Clustering-based anomaly detection in multivariate time series data. Appl Soft Computing, 2021, 100: 106919

5 Garg A, Zhang W Y, Samaran J, et al. An evaluation of anomaly detection and diagnosis in multivariate time series. IEEE Trans Neural Netw Learn Syst, 2022, 33: 2508–2517

6 Ni Y M, Chen S C. Continual unsupervised anomaly detection. Sci Sin Inform, 2022, 52: 75–85 [倪一鸣, 陈松灿. 连续无监督异常检测. 中国科学: 信息科学, 2022, 52: 75–85]

7 Zong B, Song Q, Min M R, et al. Deep autoencoding Gaussian mixture model for unsupervised anomaly detection. In: Proceedings of International Conference on Learning Representations (ICLR), 2018

8 Su Y, Zhao Y J, Niu C H, et al. Robust anomaly detection for multivariate time series through stochastic recurrent neural network. In: Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2019

9 Abdulaal A, Liu Z H, Lancewicki T. Practical approach to asynchronous multivariate time series anomaly detection and localization. In: Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2021

10 Cai M L, Wang J X, Liu J P, et al. Transformer-GAN architecture for anomaly detection in multivariate time series. Sci Sin Inform, 2023, 53: 972–992 [蔡美玲, 汪家喜, 刘金平, 等. 基于 Transformer GAN 架构的多变量时间序列异常检测. 中国科学: 信息科学, 2023, 53: 972–992]

11 Li D, Chen D C, Shi L, et al. MAD-GAN: multivariate anomaly detection for time series data with generative adversarial networks. In: Proceedings of International Conference on Artificial Neural Networks (ICANN), 2019

12 Deng L, Lian D, Huang Z, et al. Graph convolutional adversarial networks for spatiotemporal anomaly detection. IEEE Trans Neural Netw Learn Syst, 2022, 33: 2416–2428

13 Yoo Y H, Kim U H, Kim J H. Convolutional recurrent reconstructive network for spatiotemporal anomaly detection in solder paste inspection. IEEE Trans Cybern, 2022, 52: 4688–4700

14 Deng A L, Hooi B. Graph neural network-based anomaly detection in multivariate time series. In: Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI), 2021

15 Zhao H, Wang Y J, Duan J Y. Multivariate time-series anomaly detection via graph attention network. In: Proceedings of IEEE International Conference on Data Mining (ICDM), 2020

16 Wang M, Zhou D, Chen M. Hybrid variable monitoring mixture model for anomaly detection in industrial processes. IEEE Trans Cybern, 2022. doi: 10.1109/TCYB.2022.3228524

17 Khodayar M, Wang J. Spatio-temporal graph deep neural network for short-term wind speed forecasting. IEEE Trans Sustain Energy, 2020, 10: 670–681

18 Wu Z, Pan S, Long G, et al. Connecting the dots: multivariate time series forecasting with graph neural networks. In: Proceedings of International Conference on Knowledge Discovery and Data Mining (KDD), 2020

19 Silva T C. Machine Learning in Complex Networks: Modeling, Analysis, and Applications. Berlin: Springer, 2012

20 Woo S H, Park J C, Lee J Y, et al. CBAM: convolutional block attention module. In: Proceedings of European Conference on Computer Vision (ECCV), 2018

21 Velickovic P, Cucurull G, Casanova A, et al. Graph attention networks. In: Proceedings of International Conference on Learning Representations (ICLR), 2018

22 Hua W Z, Dai Z H, Liu H X, et al. Transformer quality in linear time. In: Proceedings of International Conference on Machine Learning (ICML), 2022

23 Wang F Q, Gao L, Xu T X, et al. Anomaly detection algorithm of UAV flight data based on LSTM-GAN. J Chin Inertial Technol, 2022, 30: 264–271 [王凤芹, 高龙, 徐廷学, 等. 基于 LSTM-GAN 的无人机飞行数据异常检测算法. 中国惯性技术学报. 2022, 30: 264–271]

24 Zhang X, Xu C, Tian X, et al. Graph edge convolutional neural networks for skeleton-based action recognition. IEEE Trans Neural Netw Learn Syst, 2020, 31: 3047–3060

25 Feng L J, Zhao C H, Li Y L, et al. Multichannel diffusion graph convolutional network for the prediction of endpoint composition in the converter steelmaking process. IEEE Trans Instrum Meas, 2021, 70: 1–13

26 Chen Z, Xu J, Peng T, et al. Graph convolutional network-based method for fault diagnosis using a hybrid of measurement and prior knowledge. IEEE Trans Cybern, 2021, 52: 9157–9169

27 Li M S, Chen S H, Zhang Z J, et al. Skeleton-parted graph scattering networks for 3D human motion prediction. In: Proceedings of European Conference on Computer Vision (ECCV), 2022

28 Peng Y S, Ke Q H, Rahmani, H, et al. IGFormer: interaction graph transformer for skeleton-based human interaction

recognition. In: Proceedings of European Conference on Computer Vision (ECCV), 2022
29 Yang H H, Liu Z L, Wu X P, et al. Graph R-CNN: towards accurate 3D object detection with semantic-decorated local graph. In: Proceedings of European Conference on Computer Vision (ECCV), 2022
30 Mathur A P, Tippenhauer N O. SWaT: a water treatment testbed for research and training on ICS security. In: Proceedings of International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater), 2016
31 Leskovec J, Kleinberg J, Faloutsos C. Graph evolution: densification and shrinking diameters. In: Proceedings of International Conference on Knowledge Discovery and Data Mining (KDD), 2007
32 Bloznelis M. Degree and clustering coefficient in sparse random intersection graphs. Ann Appl Probab, 2013, 23: 1254–1289
33 You T, Cheng H M, Ning Y Z, et al. Community detection in complex networks using density-based clustering algorithm and manifold learning. Phys A-Stat Mech Its Appl, 2016, 464: 221–230
34 Pedreschi N, Battaglia D, Barrat A. Important nodes identification based on degree and structural entropy. In: Proceedings of International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), 2020
35 Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. Nature, 1986, 323: 533–536
36 Kingma D P, Ba J. Adam: a method for stochastic optimization. In: Proceedings of International Conference for Learning Representations (ICLR), 2015
37 Ahmed C M, Zhou J, Mathur A P. Noise matters: using sensor and process noise fingerprint to detect stealthy cyberattacks and authenticate sensors in CPS. In: Proceedings of Annual Computer Security Applications Conference (ACSAC), 2018
38 Ahmed C M, Zhou J, Mathur A P. WADI: a water distribution testbed for research in the design of secure cyber physical systems. In: Proceedings of International Workshop on Cyber-Physical Systems for Smart Water Networks (CySWater), 2017

Multivariable time series anomaly detection using heuristic spatio-temporal graph neural network

Yu JIANG[1], Hua CHEN[2*], Xiaogang ZHANG[1*], Lianhong WANG[1] & Dingxiang WANG[1]

1. *College of Electrical and Information Engineering, Hunan University, Changsha* 410081*, China*;
2. *College of Computer Science and Electronic Engineering, Hunan University, Changsha* 410081*, China*
* Corresponding author. E-mail: chua@hnu.edu.cn, zhangxg@hnu.edu.cn

Abstract Anomaly detection for multivariable time series in cyber-physical systems is crucial for preventing system failures and ensuring safe production. The presence of strong coupling between system variables and propagation effects imparts pronounced spatio-temporal characteristics to anomalies. Designing an effective anomaly detection algorithm necessitates consideration of the coupling relationships, propagation directionality, and causal time delays among variables. Furthermore, it is essential to account for the coupling relationships between variables when detecting anomalies from a system perspective. In this study, we propose an end-to-end heuristic spatio-temporal graph neural network (HST-GNN) for the detection of anomalies in multivariate time series (MTS) data. First, we address the directed and clustered inter-variable relationships by designing a directed similarity function and a heuristic clustering algorithm. These components contribute to learning the graph structure required for analyzing MTS data. Subsequently, we employ a combination of gated convolutional attention units and a multihead graph attention layer as a spatio-temporal graph attention module. This module facilitates the simultaneous capture of nonlinear causal temporal and spatially coupled features. Finally, the graph structure features of the system are quantified and used as a complement to the depth features extracted by the spatio-temporal graph attention module. These concatenated features are then fed into an autoencoder for detecting anomalies at both the system and sensor levels. The performance of HST-GNN is verified using four public datasets. Our results highlight the advantages of utilizing a sparse directed graph structure for extracting system coupling characteristics. Additionally, the detection of anomalies at both the system and sensor levels enhances the model's sensitivity to early, albeit potentially insignificant anomalies.

Keywords multivariable time series data, unsupervised anomaly detection, heuristic graph structure, spatio-temporal graph attention network, system-level graph structure features