GAD：基于拓扑感知的时间序列异常检测

戚琦，申润业，王敬宇
（北京邮电大学网络与交换国家重点实验室，北京 100876）

摘 要：为了解决网络中节点设备异常检测、智能运维、根因分析等问题，针对链路时延、网络吞吐率、设备内存使用率等时序数据，提出了一种基于图的门控卷积编解码异常检测模型。考虑网络场景的实时性需求以及网络拓扑连接关系对时序数据的影响，基于门控卷积对时序数据并行提取时间维度特征并通过图卷积挖掘空间依赖关系。基于时空特征提取模块组成的编码器对原始输入时序数据编码后，卷积模块组成的解码器用于重构时序数据。原始数据和重构数据间的残差进一步用于计算异常分数并检测异常。在公开数据和模拟仿真平台上的实验表明，所提模型相对于目前的时间序列异常检测基准模型具有更高的识别准确率。

关键词：智能运维；异常检测；时间序列；时空卷积
中图分类号：TP393
文献标识码：A
doi: 10.11959/j.issn.1000−436x.2020113

GAD: topology-aware time series anomaly detection

QI Qi, SHEN Runye, WANG Jingyu
State Key Laboratory of Networking and Switching Technology,
Beijing University of Posts and Telecommunications, Beijing 100876,China

Abstract: To solve the problems of anomaly detection, intelligent operation, root cause analysis of node equipment in the network, a graph-based gated convolutional codec anomaly detection model was proposed for time series data such as link delay, network throughput, and device memory usage. Considering the real-time requirements of network scenarios and the impact of network topology connections on time series data, the time dimension features of time series were extracted in parallel based on gated convolution and the spatial dependencies were mined through graph convolution. After the encoder composed of the spatio-temporal feature extraction module encoded the original input time series data, the decoder composed of the convolution module was used to reconstruct the time series data. The residuals between the original data and the reconstructed data were further used to calculate the anomaly score and detect anomalies. Experiments on public data and simulation platforms show that the proposed model has higher recognition accuracy than the current time series anomaly detection benchmark algorithm.

Key words: AIOps, anomaly detection, time series, spatio-temporal convolution

1 引言

随着智能化的 B5G 及 6G 网络的部署，分析网络中的数据并学习其经验和规则可用于优化网络管理，这促使智能运维（AIOps, artificial intelligence for IT operations）技术在网络领域的发展。通过利用 AIOps 技术，可以部署异常检测、异常定位、根因分析、故障修复等辅助功能，从而提高运维效率[1]。基于对网络全局状态的精确掌控，软件定义网络（SDN, software defined networking）提供了网络自

收稿日期：2020–03–03；**修回日期**：2020–05–24
基金项目：国家重点研发计划基金资助项目（No.2018YFB1800502）；国家自然科学基金资助项目（No.61671079，No.61771068）；北京市自然科学基金资助项目（No.4182041）
Foundation Items: The National Key Research and Development Program of China (No.2018YFB1800502), The National Natural Science Foundation of China(No.61671079，No.61771068), The Natural Science Foundation of Beijing (No.4182041)

动化管理、增量部署、网络可编程化等功能[2]。然而，随着网络规模的扩展及信息交换的复杂化，维持这种网络的动态特性需要可靠的运维机制。因此，本文提出基于 AIOps 技术来实现 SDN 中的闭环控制。如图 1 所示，数据收集模块周期性地收集拓扑状态及时序数据，并上传至智能控制模块进行异常检测。在对每条链路及设备的状态信息检测后，SDN 控制器根据异常识别的检测结果通过下发流表的方式实现数据分组的调度，如图 1 中的数据转发模块所示。基于流表的调度规则可以实现路由优化及故障修复，从而维持网络正常的传输功能。

作为 AIOps 的关键组成部分，基于时间序列的异常检测模型通过对网络中的链路时延、网络吞吐率、设备内存使用率等关键绩效指标（KPI, key performance indicator）进行建模，从而检测出网络中可能存在的异常。传统的异常检测模型通常采用自回归移动平均（ARIMA, autoregressive integrated moving average）模型、指数平滑（ES, exponential smoothing）、自回归（AR, auto regressive）模型等时间序列分析方法对输入序列进行建模，以及通常采用基于规则的统计方法（例如 3-sigma 规则）用于计算阈值并检测异常[3]。由于传统方法需要人工进行参数和阈值调整，使用机器学习进行异常检测逐渐成为有效的解决方案。有监督的方法利用传统异常检测方案中的序列建模作为特征提取器，并将异常检测问题看作二分类任务。Liu 等[4]研发了一套系统来辅助运维人员标注异常，并提出基于随机森林自动选取特征提取器参数及阈值的算法。Wang 等[5]提出了一种基于标签筛选及再学习的方法来挖掘 KPI 中连续的异常间隔，并通过训练深度神经网络（DNN, deep neural network）作为异常检测分类器。然而，基于有监督的方法需要标记不同类型的 KPI 标签，这需要耗费大量的人力资源。

在无监督方法中，Malhotra 等[6]采用预测的方法，通过长短期记忆（LSTM, long short term memory）网络对序列波动模式建模并预测下一时刻的 KPI 值，将预测值和实际值之间的残差作为检测异常的依据。Hundman 等[7]提出一种基于 LSTM 并结合上下文控制信息进行时序建模的方法，该方法通过时序数据预测来检测异常。然而此类方法对噪声敏感，通常识别出大量的假阳性结果。基于重构

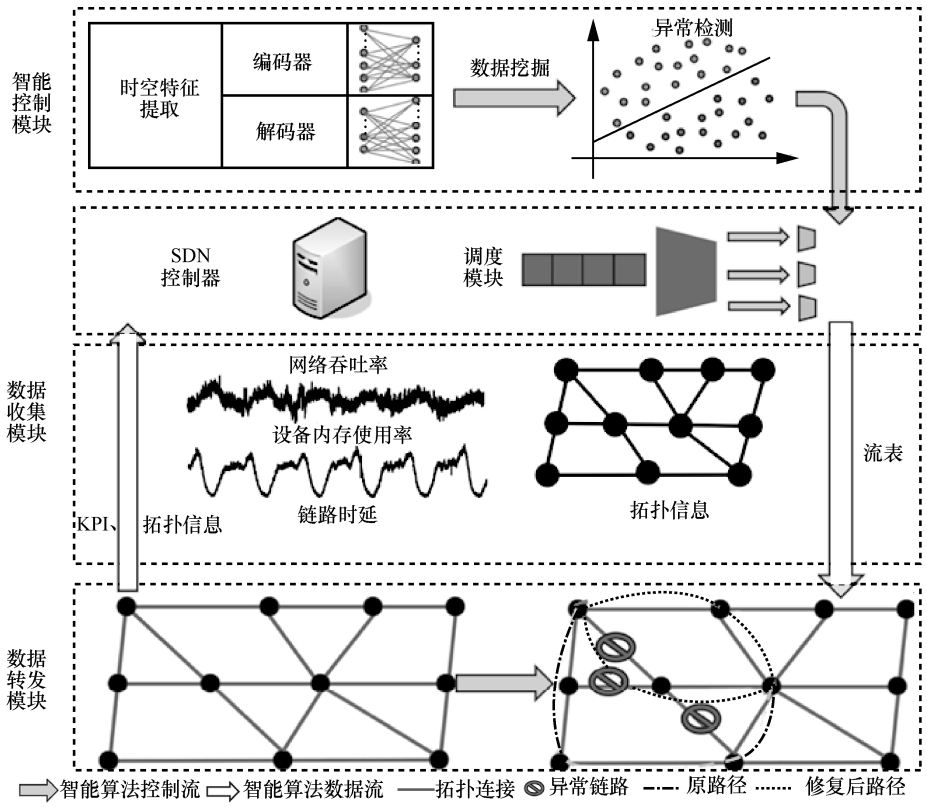


图 1　基于 AIOps 的网络运维架构

的方法可以提取出序列的低纬度特征，从而过滤噪声的干扰，具有较强的稳健性。Hawkins 等[8]率先将自动编码器用于重构时序数据并进行异常检测。Malhotra 等[9]提出了基于 LSTM 的编码–解码结构，并使用编码器的最终隐状态作为解码器的初始状态，以相反的顺序重构时间序列。Xu 等[10]提出了一种改进的变分自编码器模型，并给出了编码后低维度特征上核密度估计的合理解释。尽管这些模型具有较好的性能，但难以满足网络场景中对异常检测的实时性要求，同时难以针对网络中的拓扑连接关系进行关联分析。

针对网络拓扑中大规模 KPI 异常检测的场景，本文提出了一种基于图的门控卷积编解码异常检测（GAD，graph-based gated convolution codec for anomaly detection）模型。考虑网络场景的实时性需求及网络中设备拓扑连接关系对 KPI 数据的影响，GAD 模型采用门控卷积网络（Gated-CNN，gated convolutional network）并行提取时间特征，采用图卷积（GCN, graph convolutional network）提取空间依赖关系。在由 Gated-CNN 和 GCN 组成的编码器提取时空特征后，基于多维卷积的网络构成解码器，以重构原始拓扑中的 KPI，重构后的时序 KPI 与原始 KPI 的残差用于检测异常。本文的贡献总结如下。

1) 鉴于 SDN 缺少可靠的维持网络动态特性的运维机制，提出基于 AIOps 辅助 SDN 运维的架构，以优化网络管理，实现 SDN 的闭环控制。

2) 考虑到网络场景中节点间的交互，提出一种基于网络拓扑结构的编解码异常检测模型，通过提取节点间的空间特征，以挖掘详细的节点连接状态信息。

3) 将网络拓扑中节点的 KPI 序列组成多维矩阵，不同于已有异常检测模型仅能对单独 KPI 检测异常，GAD 模型通过门控卷积实现时间维度上的并行特征提取，提高了复杂网络场景的实时性。

4) 在公开数据集及 SDN 模拟仿真平台进行实验，结果表明 GAD 模型性能优于已有的异常检测基准模型。

2 网络中时序数据异常检测

2.1 问题定义

对网络中时序数据进行异常检测，主要是根据每个设备节点的 KPI 序列（例如链路时延、网络吞吐率、设备内存使用率等）的波动规律及网络拓扑中节点间的连接关系，对时序数据在不同时刻存在的异常波动进行检测。每个网络设备都有相应的反映运行状态的 KPI 曲线，如图 2 所示，其中异常数据已在图中标出。本文将拓扑中时序数据定义为无权图 $G=(V,E)$，其中，V 为网络中节点集合，$|V|=n$ 表示节点的个数；E 为边的集合，表示网络设备节点间的连接关系。令 $x_i(t_j)$ 表示节点 i 在 t_j 时刻的 KPI 值，则 $X(t_j)=[x_1(t_j),x_2(t_j),\cdots,x_n(t_j)]$ 表示网络中所有节点在 t_j 时刻的 KPI 值。本文的目标是对滑动窗口为 L 的时序集合 $X=[X(t_1),X(t_2),\cdots,X(t_L)]$ 进行时空特征编码及解码重构，其中，重构数据和原始数据的残差用于检测时序集合 X 是否存在异常波动。

2.2 GAD 模型框架

本文基于卷积编解码网络实现网络中时间序列的异常检测，通过 Gated-CNN 及 GCN 组成的编码器来提取网络中 KPI 的时空特征，随后基于卷积

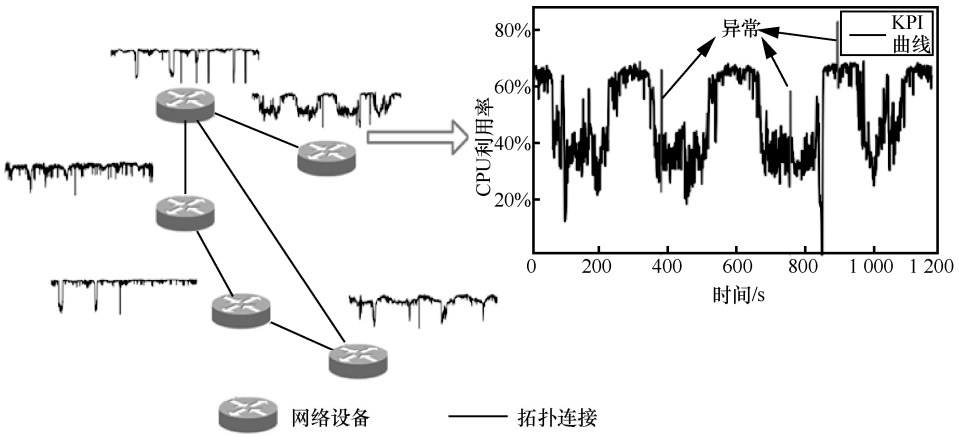


图 2 网络拓扑中时间序列异常检测

网络的解码器用以重构原始 KPI，从而评估异常分数以检测异常。

GAD 模型如图 3 所示，分为 4 个模块：预处理模块、编码器模块、解码器模块和异常检测模块。预处理模块负责对 KPI 数据进行归一化处理，并基于滑动窗口提取时序片段。编码器模块基于时序信息及网络的拓扑连接关系，提取 KPI 的时空特征，以建立低维度表征。解码器模块基于多尺度的卷积网络对时空特征解码，从而重构原始输入 KPI。异常检测模块基于重构值与真实值的相异程度计算异常分数，从而检测异常。由于异常数据难以重构，因此通过正常 KPI 训练模型，当测试集 KPI 的异常分数超过特定阈值时，模型将其判定为异常。

2.3　时间特征建模

基于捕获时间维度特征，递归神经网络（RNN，recurrent neural network）被广泛用于时间序列异常检测领域。但在网络场景中，针对每个不同的 KPI 序列曲线，RNN 都需要训练模型并检测，这无法满足网络中实时性的需求。因此，本文采用基于门控的卷积网络[11]并行地对每个时间序列进行编码，同时门控单元用于捕获时序数据的时间特征。门控卷积网络的主要结构与卷积网络相似，也是通过在卷积操作中加入门控机制从而捕获长期记忆。如图 4(a) 所示，门控卷积网络由一个因果卷积网络和另一个经过 sigmod 激活函数处理的卷积网络组成。对于图 G 中滑动窗口为 L 的 KPI 时序数据构成的输入 $\boldsymbol{X} \in \boldsymbol{R}^{n\times m\times L}$，门控卷积网络可表示为

$$\boldsymbol{h}(\boldsymbol{x}) = (\boldsymbol{XW} + \boldsymbol{b}) \Theta \sigma(\boldsymbol{XV} + \boldsymbol{c}) \tag{1}$$

其中，n 为 G 中 KPI 的节点数量，m 为特征维度，L 为滑动窗口的长度，σ 表示 sigmod 激活函数，Θ

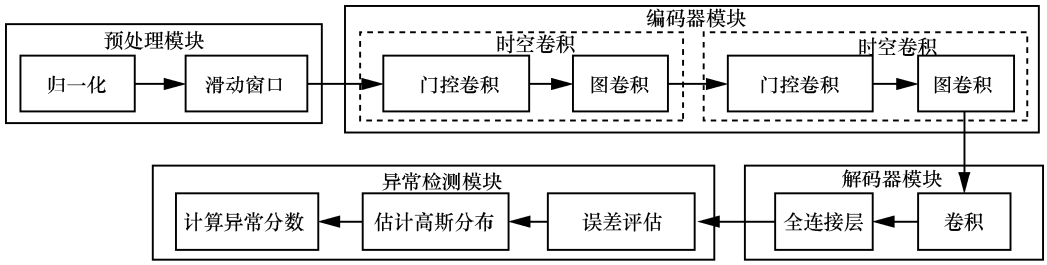


图 3　GAD 模型

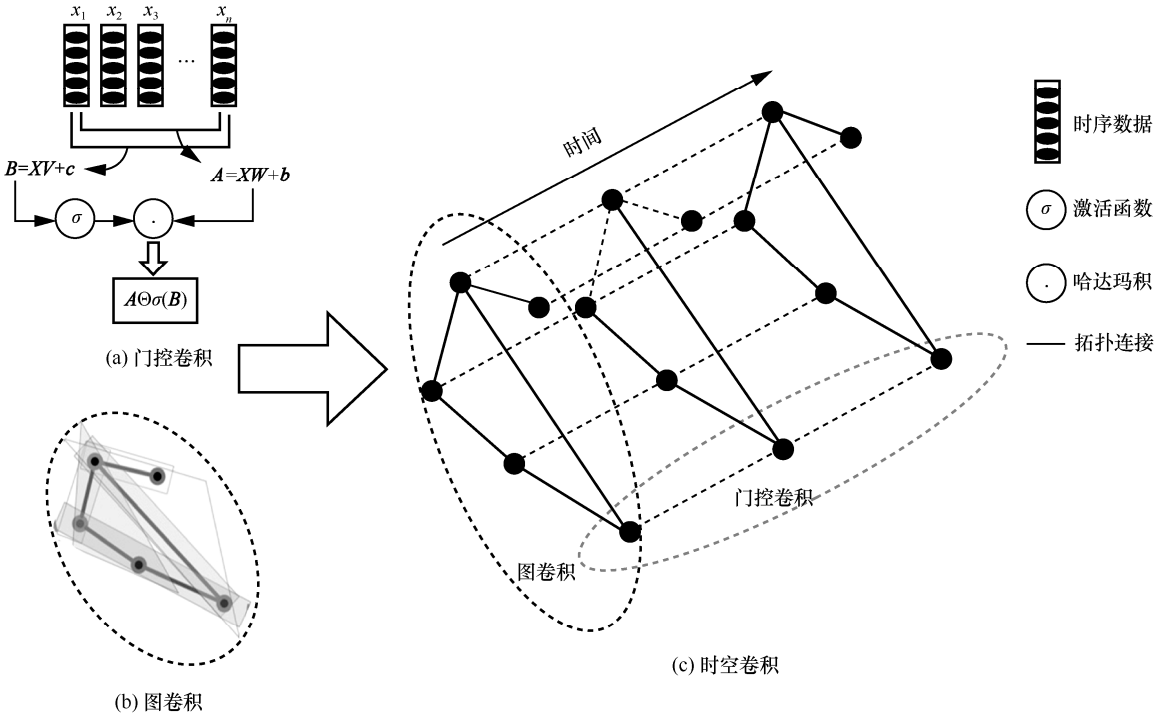


图 4　时空特征提取模块

表示哈达玛乘积，$\boldsymbol{W}$、$\boldsymbol{V}$ 表示权重矩阵，$\boldsymbol{b}$、$\boldsymbol{c}$ 表示偏置。通过堆积门控组件，GAD 模型可以保留并传递时间维度依赖信息。

2.4 空间特征建模

CNN 可以捕获空间特征，但不适用于非欧几里得空间，例如网络设备的连接关系所构成的图结构。如图 4(b)所示，本文采用 GCN[12]挖掘空间维度上节点间的依赖关系。在谱图论中，图 G 的空间特征可由拉普拉斯矩阵表示为 $\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{A}$。其中，$\boldsymbol{A}$ 表示邻接矩阵，$\boldsymbol{D}$ 表示由节点度构成的对角矩阵。拉普拉斯矩阵的特征值分解为 $\boldsymbol{L}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$，其中，$\boldsymbol{\Lambda}=\mathrm{diag}(|\lambda_1,\lambda_2,\cdots,\lambda_n|)$ 表示由 n 个特征值组成的对角矩阵，$\boldsymbol{U}=(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n)$ 是由单位特征向量组成的矩阵。以 t_j 时刻图上的 N 维向量 $\boldsymbol{f}=\boldsymbol{X}(t_j)$ 为例，图的傅里叶变换表示为 $\hat{\boldsymbol{f}}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{f}$，其傅里叶逆变换为 $\boldsymbol{f}=\boldsymbol{U}\hat{\boldsymbol{f}}$。图卷积运算可通过傅里叶域中的线性算子替代经典卷积算子来实现，根据卷积核 $\boldsymbol{h}$，图卷积表示为

$$\begin{aligned}&\boldsymbol{f}*_G\boldsymbol{h}=\boldsymbol{U}((\boldsymbol{U}^{\mathrm{T}}\boldsymbol{h})\Theta(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{f}))=\\&\boldsymbol{U}\mathrm{diag}(h(\lambda_1),\cdots,h(\lambda_n))\boldsymbol{U}^{\mathrm{T}}\boldsymbol{f}\end{aligned}\tag{2}$$

其中，$h(\lambda_l)=\sum_{i=1}^{N}h(i)u_l(i)$。式(2)可以理解为对 $\boldsymbol{h}$、$\boldsymbol{f}$ 进行傅里叶变换到谱域，再对 $\boldsymbol{h}$、$\boldsymbol{f}$ 变化结果的哈达玛积进行傅里叶逆变换。用 $\mathrm{diag}(\theta_l)$ 替换 $\mathrm{diag}(h(\lambda_l))$，则 $\boldsymbol{f}*_G\boldsymbol{h}=\boldsymbol{U}\boldsymbol{g}_\theta(\boldsymbol{\Lambda})\boldsymbol{U}^{\mathrm{T}}\boldsymbol{f}$，其中 $\boldsymbol{g}_\theta(\boldsymbol{\Lambda})=\mathrm{diag}(\theta_1,\theta_2,\cdots,\theta_n)$ 表示卷积核[13]。鉴于 $\boldsymbol{U}$、$\boldsymbol{g}_\theta(\boldsymbol{\Lambda})$ 及 $\boldsymbol{U}^{\mathrm{T}}$ 的矩阵乘积在大规模图运算具有较高的计算复杂度，本文使用切比雪夫多项式递归计算卷积核[14]，即 $\boldsymbol{g}_\theta(\boldsymbol{\Lambda})=\sum_{k=0}^{k-1}\beta_k\boldsymbol{T}_k(\boldsymbol{\Lambda})$，其中，$\beta_k$ 表示切比雪夫多项式的系数，则 $\boldsymbol{f}*_{\boldsymbol{G}}\boldsymbol{h}=\sum_{k=0}^{k-1}\beta_k\boldsymbol{T}_k(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}})\boldsymbol{f}=\sum_{k=0}^{k-1}\beta_k\boldsymbol{T}_k(\hat{\boldsymbol{L}})\boldsymbol{f}$，其中 $\hat{\boldsymbol{L}}=\frac{2\boldsymbol{L}}{\boldsymbol{\lambda}_{\max}-\boldsymbol{I}}$。切比雪夫多项式的递归定义为 $\boldsymbol{T}_k(\hat{\boldsymbol{L}})=2\hat{\boldsymbol{L}}\boldsymbol{T}_{k-1}(\hat{\boldsymbol{L}})-\boldsymbol{T}_{k-2}(\hat{\boldsymbol{L}})$，其中，$\boldsymbol{T}_0(\hat{\boldsymbol{L}})=\boldsymbol{I}$，$\boldsymbol{T}_1(\hat{\boldsymbol{L}})=\hat{\boldsymbol{L}}$。根据切比雪夫多项式的性质，时间复杂度从 $O(n^2)$ 降低为 $O(K|E|)$。切比雪夫多项式的展开与以每个节点为中心的 0 到第 $(K-1)$ 阶邻居的卷积运算相对应。图卷积运算后，每个节点将通过其 $0\sim(K-1)$ 个邻居的值进行更新。

2.5 异常评分与判定机制

在对网络 KPI 的空间和时间特征进行编码后，用卷积网络构成的解码器模型来重构 KPI 序列。节点 i 在 t_j 时刻的重建误差 $e_{i,t_j}=|x_i(t_j)-x_i'(t_j)|$，其中，$x_i(t_j)$ 表示节点 i 在 t_j 时刻的 KPI 数据，$x_i'(t_j)$ 表示模型重构的 KPI 数据。$\boldsymbol{e}_i=[e_{i,t_1},e_{i,t_2},\cdots,e_{i,t_n}]$ 表示节点 i 的残差矩阵。本文算法的异常评分与判定机制如算法 1 所示。本文将正常 KPI 分为 4 组：训练集 S_N、验证集 V_{N1}、验证集 V_{N2} 及测试集 T_1。异常 KPI 分为 2 组：验证集 V_{N3}、测试集 T_2。通过 S_N 训练模型，并由 V_{N1} 概括 GAD 模型获得误差向量后，本文基于最大似然来评估每个节点 KPI 的正态分布 $\mathbf{N}(\boldsymbol{u}_i,\boldsymbol{\Sigma}_i)$。每个 KPI 节点 i 的异常分数矩阵为

$$\mathbf{score}_i=(\boldsymbol{e}_i-\boldsymbol{u}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i(\boldsymbol{e}_i-\boldsymbol{u}_i)\tag{3}$$

给定阈值 $\mathrm{threshold}_i$，如果节点 i 在 t_j 时刻的异常分数 $\mathrm{score}_i[t_j]>\mathrm{threshold}_i$，则 KPI 节点 i 在时刻 t_j 的数值被判定为异常点，否则判定为正常点。随后，V_{N2} 和 V_{N3} 来评估 KPI 节点的最佳阈值，以使 F1-score=$\frac{2PR}{P+R}$ 最大化，其中，P 表示精度，R 表示召回率。

算法 1 异常评分与判定机制

输入 $S_N,V_{N1},V_{N2},T_1,V_{N3},T_2$

输出 智能检测结果

1) $\mathrm{reprocess}(S_N,V_{N1},V_{N2},T_1,V_{N3},T_2)$

2) $\mathrm{train}(S_N)$ // 通过 S_N 预训练模型

3) $\mathrm{error}\leftarrow\mathrm{loss}(V_{N1},\mathrm{reconstruction}(V_{N1}))$ // 通过 V_{N1} 计算误差向量

4) for $i\leftarrow1$ to n do // 评估正态分布的参数 $\boldsymbol{u}_i$ 以及 $\boldsymbol{\Sigma}_i$

5) $\boldsymbol{u}_i\leftarrow\mathrm{mean}(\mathbf{error}_i)$

6) $\boldsymbol{\Sigma}_i\leftarrow\mathrm{cov}(\mathbf{error}_i)$

7) end for

8) $\boldsymbol{e}\leftarrow\mathrm{loss}(V_{N2}\cup V_{N3},\mathrm{reconstruction}(V_{N2}\cup V_{N3}))$

9) for $i\leftarrow1$ to n do

10) $\mathbf{score}_i\leftarrow(\boldsymbol{e}_i-\boldsymbol{u}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i(\boldsymbol{e}_i-\boldsymbol{u}_i)$ //对每个节点 i 计算异常分数向量 $\mathbf{score}_i$

11) $\mathrm{threshold}_i\leftarrow\mathrm{findBestThreshold}(\mathbf{score}_i)$ // 选取使 F1-score 取最优的 $\mathrm{threshold}_i$

12) end for

13) $\boldsymbol{e}\leftarrow\mathrm{loss}(T_1\cup T_2,\mathrm{reconstruction}(T_1\cup T_2))$

14) for $i \leftarrow 1$ to n do

15) $\mathbf{score}_i \leftarrow (\boldsymbol{e}_i - \boldsymbol{u}_i)^{\mathrm{T}} \boldsymbol{\Sigma}_i (\boldsymbol{e}_i - \boldsymbol{u}_i)$

16) $\mathbf{result}_i \leftarrow \mathrm{getResult}(\mathbf{score}_i, \mathrm{threshold}_i)$

17) end for

2.6 复杂度分析

异常检测模块主要由 3 个部分组成。1) 只包含正常数据的训练集 S_N 预先用于训练 GAD 模型，验证集 V_{N1} 通过 GAD 模型计算误差向量，每个节点的误差向量用于评估该节点的正态分布参数 $\boldsymbol{u}_i$ 及 $\boldsymbol{\Sigma}_i$。该部分时间复杂度为 $O(g+nl')$，其中 n 表示网络中节点的数目，l' 表示数据集的时间点总数，g 表示 GAD 模型时间复杂度，$g=O\left(\sum_{r=1}^{d} m_r^2 k_r^2 c_{r-1} c_r\right)$，$d$ 表示网络深度，m_r 表示第 r 个卷积核输出边长，k_r 表示第 r 个卷积核边长，c_r 表示网络中第 r 个卷积操作输出通道数，c_{r-1} 表示网络中第 r 个卷积操作输入通道数。2) 验证集 V_{N2}、V_{N3} 通过 GAD 模型重构误差向量后，根据复杂度分析中 1)部分在每个节点构建的正态分布 $\boldsymbol{u}_i$ 及 $\boldsymbol{\Sigma}_i$，计算该节点的异常分数向量。通过在不同的异常阈值中取值并计算 F1-score，把该节点取得最大 F1-score 的对应阈值作为最佳阈值。该部分时间复杂度为 $O\big(g+n(l'+t)\big)$，其中 t 表示阈值搜索范围。3) 测试集 T_1 和 T_2 通过 GAD 模型计算误差向量并获得异常分数后，基于复杂度分析中 2)部分的最佳阈值来判断该节点在不同时刻是否为异常点。该部分时间复杂度为 $O(g+nl')$。综上所述，异常检测模块总的时间复杂度为 $O\big(g+n(l'+t)\big)$。

3 实验验证

本节将对 GAD 模型进行性能分析，分别在公开数据集及模拟仿真平台进行对比实验，并与以下几种异常检测模型进行对比。

1) Autoencoder[8]。自动编码器，由多层线性感知组成的编解码网络模型，是数据重构的一种经典方法。

2) EncDec-AD（encoder-decoder scheme for anomaly detection）[9]。将 LSTM 作为编码器和解码器，把编码器的最终状态值用作解码器的初始状态，以相反的顺序重构时间数据的异常检测方法。

3) DAGMM（deep autoencoding gaussian mixture model）[15]。通过联合训练自动编码器及高斯混合模型并计算重构残差，实现对概率分布建模的异常检测方法。

4) LSTM-NDT（LSTM with the nonparametric dynamic thresholding）[7]。一种基于 LSTM 并结合上下文控制信息进行时序建模，通过时序数据预测来检测异常的方法。

3.1 数据集介绍及评价指标

本文采用文献[16]中的洛杉矶高速路数据集，该数据集由 207 个节点统计的 34 272 个不同时刻的车辆流量数据组成。本文按滑动窗口提取时序集合，并注入异常数据。为评估模型对异常检测的准确性，本文采用精确率（P）、召回率（R）、F1-score 及 AUCPR（the area under the PR curve）衡量模型的准确度作为评价指标。F1-score=$\dfrac{2PR}{P+R}$，AUCPR 为 P-R 曲线围成的面积。鉴于 F1-score 和 AUCPR 能兼顾精确率及召回率，故将其作为衡量模型性能的主要标准。

3.2 实验结果及分析

本实验主要包含 3 个部分。 1)验证性实验：通过与当前的基线模型对比来证明 GAD 模型性能的优越性； 2)对比实验：通过与 GAD 模型的不同变种进行比较，以时间、空间特征提取模块的有效性；3) 参数影响：分析不同的滑动窗口参数对 GAD 模型检测结果的影响。

1) 验证性实验

首先对比 GAD 模型与其他模型的异常检测结果，对比结果如表 1 所示。相对于挖掘时序周期波动的检测模型，AutoEncoder 模型和 DAGMM 模型在重构时序数据时并不能充分挖掘时间维度的信息，故这 2 种模型对时序数据的异常检测效果并不理想。而 EncDec-AD 模型和 LSTM-NDT 模型通过对时序的周期性波动建模，具有较好的识别效果。从识别结果可以看到， LSTM-NDT 模型在识别精确率上具有优势，但相对于 GAD 模型具有较低的召回率，在实际中该模型对异常数据有着较高的警告阈值，综合评估性能并不占优势。当涉及基于拓扑的时序异常检测场景时，这些对比模型没有考虑到异常的传播过程，忽视了空间维度的特征。例如，某个设备的吞吐率异常可能随着数据分组在链路传播，进而影响到相邻的节点。从表 1 可以看出，GAD 模型的 F1-score 为 0.969 8，AUCPR 为 0.972 2，均优于其他模型，这意味着针对网络中 KPI 的异常检测场景，基于空间以及时间特征联合建模，GAD 模型可以挖掘

时序数据的低维特征，从而精确地重构原始时序数据。

表 1　　各模型的异常检测结果

模型	精确率（P）	召回率（R）	F1-score	AUCPR
AutoEncoder	0.288 1	0.636 5	0.352 2	0.267 5
EncDec-AD	0.995 7	0.905 7	0.941 3	0.921 9
DAGMM	0.734 0	0.858 7	0.723 2	0.699 2
LSTM-NDT	0.999 9	0.914 7	0.948 8	0.920 7
GAD	0.983 3	0.962 0	0.969 8	0.972 2

为了详细说明比较结果，将 GAD 模型和其他对比模型在特定 KPI 片段进行详细比较，如图 5 所示。当异常分数高于阈值时，则该点被判定为异常点。针对具有明显异常波动的时序数据（图 5 各子图中第二个和第三个灰色背景标注），EncDec-AD、LSTM-NDT 和 GAD 等模型都给出了较高的异常分数，与正常波动的异常分数相比具有明显的区分。但对于那些持续时间短、振幅低的异常波动（如图 5 各子图中第一个灰色背景标注的异常片段），基于时间维度的特征很难识别异常序列，给出了较低的异常分数，GAD 模型联合时间维度特征及空间维度特征可以检测出大多数的异常点，给出的异常分数结合阈值可明显区分异常点及正常点。

2)对比实验

为进一步说明门控机制和空间特征提取模块在异常检测中的有效性，本文考虑了 GAD 模型的 3 种变体：1) $\text{GAD}^{\text{conv2}}$ 模型：相对于 GAD 模型，编解码模块去除了图卷积模块。2) $\text{GAD}^{\text{conv1}}_{\text{GCN1}}$ 模型：相对于 GAD 模型，编解码模块中仅采用单层的门控卷积以及图卷积用于特征提取。3) $\text{GAD}^{\text{conv1}}$ 模型：相对于 GAD 模型，编解码模块仅由单层的门控卷积用于特征提取。

如图 6 所示，$\text{GAD}^{\text{conv2}}$ 模型的性能相对于 $\text{GAD}^{\text{conv1}}$ 有明显提升，同时 GAD 模型的性能优于 $\text{GAD}^{\text{conv1}}_{\text{GCN1}}$，表明了通过层叠多层门控卷积网络，模型可以提取时序在时间维度上的深层次特征，从而重构序列的周期性波动趋势。同时，观察到 GAD 模型相对于去除 GCN 模块的 $\text{GAD}^{\text{conv2}}$ 模型及 $\text{GAD}^{\text{conv1}}$ 模型，性能有显著提高，表明利用网络拓扑中节点间空间特征，GAD 模型可获取 KPI 在拓扑上的传播过程，从而更准确地获取时序数据的波动趋势，提高异常检测性能。

3) 参数影响

滑动窗口长度 L 在重构序列中起着重要作用。较短的窗口长度不能反映时间维度的特征信息，而较长的滑动窗口可能导致时序数据在较大的时间跨度上缺乏敏感性，并且增加了重构序列的难度。图 7 显示了异常检测性能随滑动窗口长度 L 的变化情况。可以看出，随着滑动窗口长度的增加，GAD

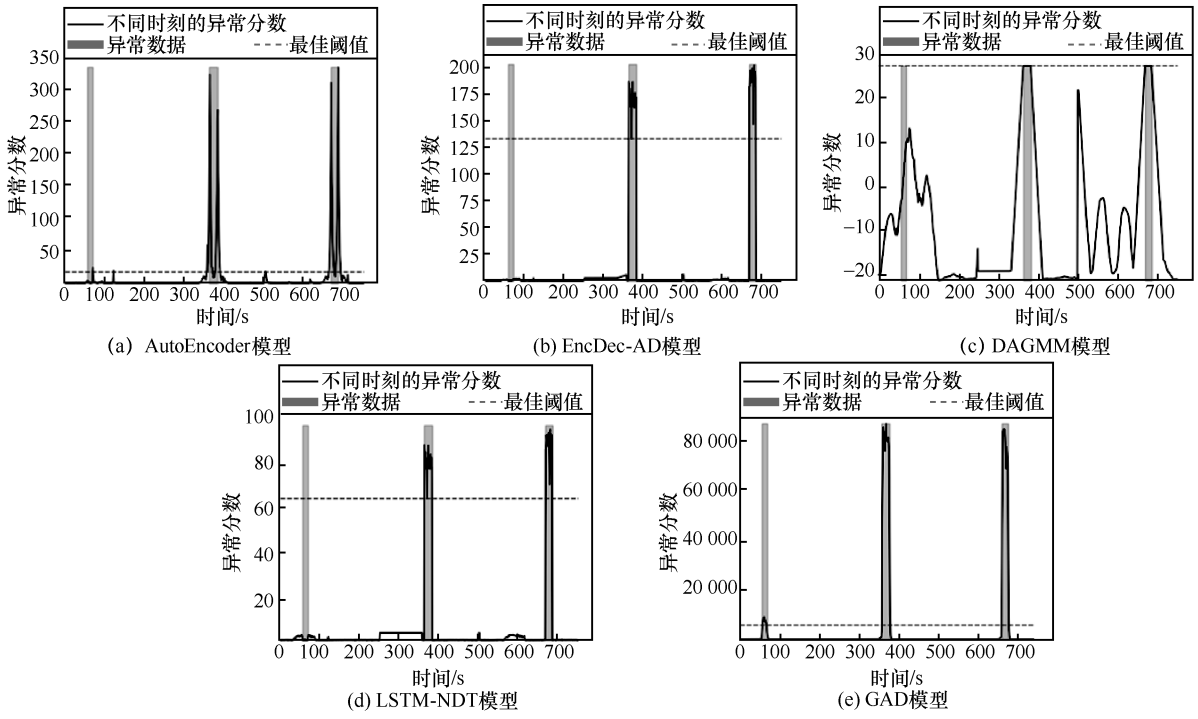
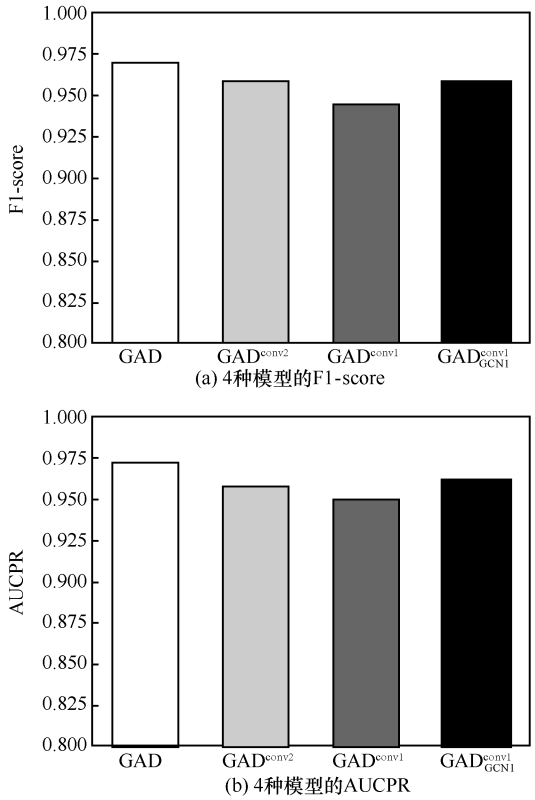


图 5　GAD 与其他模型异常分数对比

模型的性能呈现先逐渐增加又逐渐降低的趋势。综合 F1-score 和 AUCPR 的波动趋势，当滑动窗口大小为 19 时，异常检测能取得较高的性能。



(a) 4种模型的F1-score

(b) 4种模型的AUCPR

图 6　GAD 模型与其模型变种对比

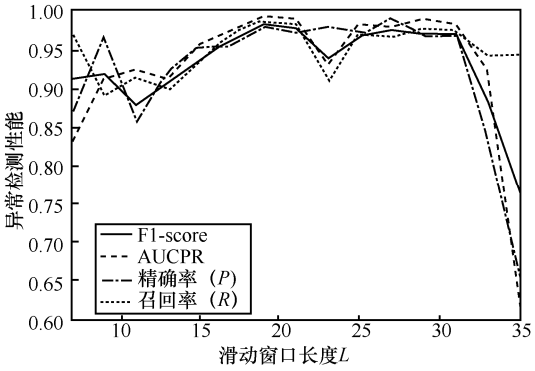


图 7　异常检测性能与滑动窗口长度 L 的关系

4　网络 KPI 异常检测仿真实验

通过搭建网络仿真平台，模拟实际网络中 KPI 的时序波动，从而对 GAD 的性能进行进一步验证。

4.1　仿真实验平台

本仿真实验采用 14 个 Open vSwitch 虚拟交换机并使用 OpenDaylight 作为 SDN 控制器，基于 mininet 搭建网络仿真平台，其网络拓扑如图 8 所示。

针对 KPI 的采集，本文开发了基于 OpenDaylight 的北向插件，周期性地采集网络中的 KPI 时序数据。GAD 模型则通过 REST API（representational state transfer application programming interface）的方式，获取 OpenDaylight 控制器采集到的 KPI 信息。本文采用 tcpreplay 重放 PCAP（packet capture）文件中数据分组，来模拟网络中的正常流量传输模式。同时，为评定 GAD 模型对 KPI 异常的识别能力，本文建立了服务器端程序，通过客户端对服务端的大规模带宽请求，模拟网络中突发性的流量爆发等方式，从而引起 KPI 时序数据的异常波动。

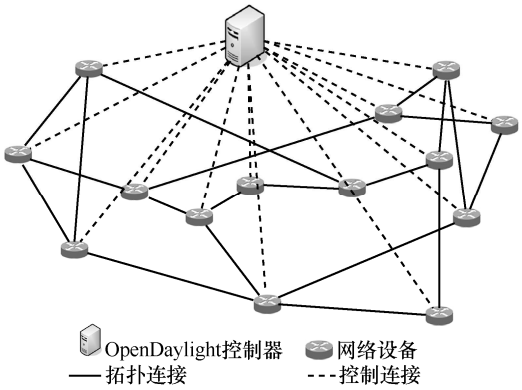


图 8　网络拓扑

4.2　实验结果

通过基于 OpenDaylight 控制器的北向插件，以 5 s 为周期收集一次 KPI 值，共计收集 3 h 的数据并模拟 50 条异常数据，其异常检测结果如表 2 所示。

表 2　　**异常检测结果**

模型	F1-score	AUCPR	响应时间/s
AutoEncoder	0.567 2	0.554 0	0.19
EncDec-AD	0.964 0	0.948 2	1.94
DAGMM	0.689 5	0.680 3	14.45
LSTM-NDT	0.964 8	0.941 9	3.13
GAD	0.974 9	0.981 0	0.25

从表 2 可以看出，本文提出的 GAD 模型在网络 KPI 异常检测方面综合考虑时空维度特征，相对于其他模型具有较高的异常识别精度。在响应时间上，虽然逊色于 AutoEncoder 模型，但相对于其他模型需要对每个 KPI 序列进行单独建模，GAD 模型可以对网络 KPI 进行并行的编解码重构操作，故具有相对较低的响应时间。综上，GAD 模型在网络时序异常检测的识别率上具有较大优势，同时满足

网络中 KPI 检测的实时性。

5 结束语

随着网络规模的扩大及服务间依赖关系的复杂化，运维人员很难对复杂场景中的异常进行精确定位及故障排除。为此，本文研究了基于 SDN 的 AIOps 智能网络运维架构，通过结合时间序列异常检测模型实现网络设备状态的监控。传统异常检测模型通常对单独 KPI 序列进行建模，对网络中的拓扑变化缺乏自适应，也难以满足实时性等需求。针对现有模型的不足，本文提出了基于图的门控卷积编码异常检测模型，通过门控卷积以及图卷积模块提取网络拓扑中 KPI 的时空特征。在由编码器–解码器模型框架重建 KPI 序列后，通过其与原始序列的残差来检测异常。实验表明，所提模型相对于当前的模型具有更好的性能，可以扩展到无人驾驶网络、智能电网、智能机器人集群等基于拓扑图的智能运维场景。

参考文献：

[1] DANG Y, LIN Q, HUANG P. AIOps: real-world challenges and research innovations[C]//2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings. Piscataway: IEEE Press, 2019: 4-5.

[2] KAUR K, GARG S, AUJLA G S, et al. Edge computing in the industrial internet of things environment: software-defined-networks-based edge-cloud interplay[J]. IEEE Communications Magazine, 2018, 56(2): 44-51.

[3] CHOU J S, TELAGA A S. Real-time detection of anomalous power consumption[J]. Renewable and Sustainable Energy Reviews, 2014(33): 400-411.

[4] LIU D, ZHAO Y, XU H, et al. Opprentice: towards practical and automatic anomaly detection through machine learning[C]// Proceedings of the 2015 Internet Measurement Conference. New York: ACM Press, 2015: 211-224.

[5] WANG J, JING Y, QI Q, et al. ALSR: an adaptive label screening and relearning approach for interval-oriented anomaly detection[J]. Expert Systems with Applications, 2019(136): 94-104.

[6] MALHOTRA P, VIG L, SHROFF G, et al. Long short term memory networks for anomaly detection in time series[C]//Proceedings of Presses Universitaires de Louvain. [S.n.: s.l.], 2015:89.

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

[8] HAWKINS S, HE H, WILLIAMS G, et al. Outlier detection using replicator neural networks[C]//International Conference on Data Warehousing and Knowledge Discovery. Berlin: Springer, 2002: 170-180.

[9] MALHOTRA P, RAMAKRISHNAN A, ANAND G, et al. LSTM-based encoder-decoder for multi-sensor anomaly detection[J]. arXiv Preprint, arXiv:1607.00148, 2016.

[10] XU H, CHEN W, ZHAO N, et al. Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in Web applications[C]//Proceedings of the 2018 World Wide Web Conference. New York: ACM Press, 2018: 187-196.

[11] DAUPHIN Y N, FAN A, AULI M, et al. Language modeling with gated convolutional networks[C]//Proceedings of the 34th International Conference on Machine Learning. New York: ACM Press, 2017: 933-941.

[12] GUO S, LIN Y, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2019: 922-929.

[13] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral networks and deep locally connected networks on graphs[J]. arXiv Preprint, arXiv: 1312. 6203, 2013.

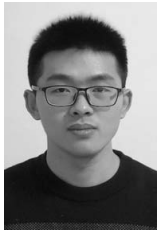
[14] HAMMOND D K, VANDERGHEYNST P, GRIBONVAL R. Wavelets on graphs via spectral graph theory[J]. Applied and Computational Harmonic Analysis, 2011, 30(2): 129-150.

[15] ZONG B, SONG Q, MIN M R, et al. Deep autoencoding gaussian mixture model for unsupervised anomaly detection[C]//IEEE International Conference on Learning Representations. Piscataway: IEEE Press, 2018: 45-64.

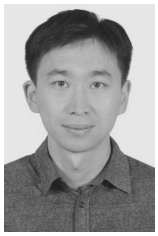
[16] JAGADISH H V, GEHRKE J, LABRINIDIS A, et al. Big data and its technical challenges[J]. Communications of the ACM, 2014, 57(7): 86-94.

[作者简介]

戚琦（1982– ），女，河北廊坊人，博士，北京邮电大学副教授、博士生导师，主要研究方向为智能边缘计算、业务网络智能化、网络资源优化等。



申润业（1996– ），男，安徽六安人，北京邮电大学硕士生，主要研究方向为智能运维、软件定义网络等。



王敬宇（1978– ），男，吉林长春人，博士，北京邮电大学教授、博士生导师，主要研究方向为智能网络、人工智能、云计算、多媒体通信、多路径传输、流量工程等。