

## 基于双重分类和重建的跨域图异常检测

苏世玉, 于炯, 李姝, 酒世承

### 引用本文

苏世玉, 于炯, 李姝, 酒世承. 基于双重分类和重建的跨域图异常检测[J]. 计算机科学, 2025, 52(8): 374-384.

SU Shiyu, YU Jiong, LI Shu, JIU Shicheng. [Cross-domain Graph Anomaly Detection Via Dual Classification and Reconstruction](#) [J]. Computer Science, 2025, 52(8): 374-384.

---

### 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

**Similar articles recommended (Please use Firefox or IE to view the article)**

#### [随时间持续演化的流图神经网络](#)

Continuously Evolution Streaming Graph Neural Network

计算机科学, 2025, 52(8): 118-126. <https://doi.org/10.11896/jsjcx.241000186>

#### [基于动态阈值伪标签筛选的深度图对比聚类算法](#)

Deep Graph Contrastive Clustering Algorithm Based on Dynamic Threshold Pseudo-label Selection

计算机科学, 2025, 52(8): 100-108. <https://doi.org/10.11896/jsjcx.240700112>

#### [时空图神经网络在PM<sub>2.5</sub>浓度预测中的应用综述](#)

Review on Application of Spatial-Temporal Graph Neural Network in PM<sub>2.5</sub> Concentration Forecasting

计算机科学, 2025, 52(8): 71-85. <https://doi.org/10.11896/jsjcx.240700153>

#### [基于干扰样本分布优化的工控异常检测改进SVM模型](#)

Improved SVM Model for Industrial Control Anomaly Detection Based on Interference Sample

Distribution Optimization

计算机科学, 2025, 52(7): 388-398. <https://doi.org/10.11896/jsjcx.240500100>

#### [基于跨模态超图优化学习的多模态情感分析](#)

Cross-modal Hypergraph Optimisation Learning for Multimodal Sentiment Analysis

计算机科学, 2025, 52(7): 210-217. <https://doi.org/10.11896/jsjcx.240600127>

# 基于双重分类和重建的跨域图异常检测

苏世玉<sup>1</sup> 于炯<sup>2</sup> 李姝<sup>3</sup> 酒世承<sup>1</sup>

1 新疆大学软件学院 乌鲁木齐 830000

2 新疆大学信息科学与工程学院 乌鲁木齐 830000

3 新疆大学计算机科学与技术学院 乌鲁木齐 830000

(107552201726@stu.xju.edu.cn)

**摘要** 跨域图异常检测通过带标签的源图辅助无标签目标图,提升了异常节点检测的准确性,进而有效降低了无监督图异常检测中的高误报率。尽管已有多种领域自适应方法被相继提出,但图数据复杂的拓扑结构与节点属性之间的关系使得源图与目标图之间的特征难以对齐;此外,图异常节点的多样性进一步增加了域对齐后的检测难度。为了解决上述问题,提出了一种新的跨域图异常检测框架,即双重分类和重建网络(Dual Classification and Reconstruction Network, DCRN)。该网络采用重建策略进行领域自适应,通过联合优化结构和属性的共享编码器、异常分类器和解码器,使共享编码器能够有效捕捉源图与目标图之间复杂的拓扑结构和节点属性关系,实现特征对齐与知识迁移。在对目标图进行异常检测的过程中,DCRN 结合异常分类器和解码器的检测结果,识别与源图相似的异常节点以及仅存在于目标图中的特有异常,从而提升了模型的检测效果。在 4 个真实数据集上的实验表明,与 10 种基线方法相比,DCRN 的 AUC-ROC 和 AUC-PR 平均提升了 4.5% 和 20.5%,且 FAR 指标降低了 16.13%。这些结果表明 DCRN 能够有效地检测目标图中的异常节点。

**关键词:** 异常检测;属性图;域自适应;图神经网络;知识迁移

中图分类号 TP391

## Cross-domain Graph Anomaly Detection Via Dual Classification and Reconstruction

SU Shiyu<sup>1</sup>, YU Jiong<sup>2</sup>, LI Shu<sup>3</sup> and JIU Shicheng<sup>1</sup>

1 School of Software Engineering, Xinjiang University, Urumqi 830000, China

2 College of Information Science and Engineering, Xinjiang University, Urumqi 830000, China

3 School of Computer Science and Technology, Xinjiang University, Urumqi 830000, China

**Abstract** Cross-domain graph anomaly detection improves the accuracy of detecting anomalous nodes by leveraging labeled source graphs to assist in detecting anomalies in unlabeled target graphs, effectively reducing the high false positive rate in unsupervised detection. Aligning features between source and target graphs remain challenging due to the complex relationships between graph topology and node attributes, and the diversity of anomalous nodes further complicates detection. To address this, a Dual Classification and Reconstruction Network (DCRN) is proposed. DCRN employs a reconstruction-based strategy for domain adaptation, optimizing shared structure and attribute encoders, anomaly classifiers, and decoders. This enables the model to capture complex topological and attribute relationships between source and target graphs, achieving effective feature alignment and knowledge transfer. DCRN combines classifier and decoder results to identify both shared and unique anomalies in the target graph, enhancing detection accuracy and robustness. Experiments on four real-world datasets show that DCRN outperforms 10 baseline algorithms, with an average improvement of 4.5% in AUC-ROC, 20.5% in AUC-PR, and a 16.13% reduction in FAR, demonstrating its effectiveness in detecting anomalous nodes in target graphs.

**Keywords** Anomaly detection, Attribute graph, Domain adaptation, Graph neural network, Knowledge transfer

## 1 引言

属性图是一种强大的数据表示形式<sup>[1]</sup>,因复杂的拓扑结构和属性特征,被广泛应用于社交媒体网络<sup>[2]</sup>、引文网络<sup>[3]</sup>和

金融交易网络<sup>[4]</sup>等领域。在这些属性图中,经常存在与其他节点显著不同的异常节点,它们可能引发一系列现实世界中的问题。例如,社交网络中的恶意用户可能散布虚假信息、推广广告或发动网络攻击;金融网络中的欺诈行为可能导致客

到稿日期:2024-10-28 返修日期:2025-03-12

基金项目:国家自然科学基金(62262064)

This work was supported by the National Natural Science Foundation of China(62262064).

通信作者:于炯(yujiong@xju.edu.cn)

户和商家蒙受重大经济损失。因此,检测属性图中的异常节点变得尤为重要。

由于标记异常点的过程高度劳动密集且依赖专业领域知识,目前大多数图异常检测(Graph Anomaly Detection,GAD)方法主要采用无监督方式<sup>[5-8]</sup>。但是,这些无监督方法无法充分利用标签信息来指导学习过程,通常具有较高的误报率,因此在实际应用中受到限制。

跨域图异常检测(Cross-domain Anomaly Detection,CD-GAD)为这一问题提供了有效的解决方案<sup>[9]</sup>。CD-GAD旨在利用带有标记的源域图的异常节点信息,来辅助未标记的目标域图的异常检测。通过调整源图和目标图之间的域差异,源图中的异常知识可以为目标图上的异常检测提供重要依据,显著提升了模型对异常节点的鉴别能力,进而提高了检测性能,降低了误报率。

在实际应用中,CD-GAD技术展现出巨大的潜力。例如,Yelp的酒店和餐饮服务因用户群和行为重叠,会共享类似的欺诈性评论模式。跨域图异常检测,可以将酒店服务中的标记异常评论模式迁移到餐饮服务领域,减少对目标域的标记需求<sup>[9]</sup>。在金融行业中,不同的银行或支付平台可能共享部分客户和交易数据,但每个平台的异常行为有所不同。跨域图异常检测方法可以通过在一个平台上训练异常检测模型来帮助识别另一个平台中的潜在欺诈行为<sup>[10]</sup>。

虽然跨域图异常检测显著提高了对属性图中异常节点的识别能力,但要在不同领域之间进行有效的特征对齐和异常检测,仍面临以下2个主要挑战。

1)图数据的复杂性显著增加了跨域特征对齐的难度。图数据由拓扑结构和属性特征共同构成,这两种模式之间的交互决定了异常模式的多样性<sup>[11]</sup>。具体而言,图中的异常节点可划分为3类:结构异常、属性异常和混合异常<sup>[12]</sup>。如图1所示,结构异常指节点在其邻域内结构不一致但属性一致;属性异常指节点在其邻域内结构一致但属性不一致;混合异常则指节点在结构和属性上均表现为不一致。在跨域场景中,不同领域间的拓扑结构和属性特征分布存在显著的异质性,导致图异常的多样性,从而对跨域特征对齐提出了更高的要求。

但是,现有的跨域图异常检测方法<sup>[9,13-14]</sup>在域自适应过程中往往仅关注结构特征的对齐,忽略了属性特征在捕捉异常模式中的重要性。这种方法虽然能够在一定程度上检测结构异常,但对于依赖属性特征的异常(如属性异常和混合异常)的检测能力有限,导致对异常模式的识别不够全面。

2)图异常的多样性增加了目标图异常检测的复杂性。属性图中通常包含结构异常、属性异常和混合异常,这使得在跨域图异常检测中,源图和目标图中的异常类型分布可能存在显著差异。具体而言,目标图可能同时包含与源图异常模式一致的共享异常,也可能存在仅在目标图中出现的非共享异常<sup>[9]</sup>。这种多样性增加了跨域图异常检测的难度,特别是在如何全面且有效地检测目标图中所有类型的异常时,依然面临挑战。

现有的异常检测方法<sup>[15-16]</sup>通常通过训练一个域适应

分类器来检测目标图上的异常节点。若源图与目标图中异常类型相似(即共享异常),则此类方法能够取得较好的检测效果。然而,当目标图中出现源图中未见过的异常类型(即非共享异常)时,现有方法往往无法有效识别这些异常,导致检测性能下降,难以准确捕捉目标图中的所有异常节点。

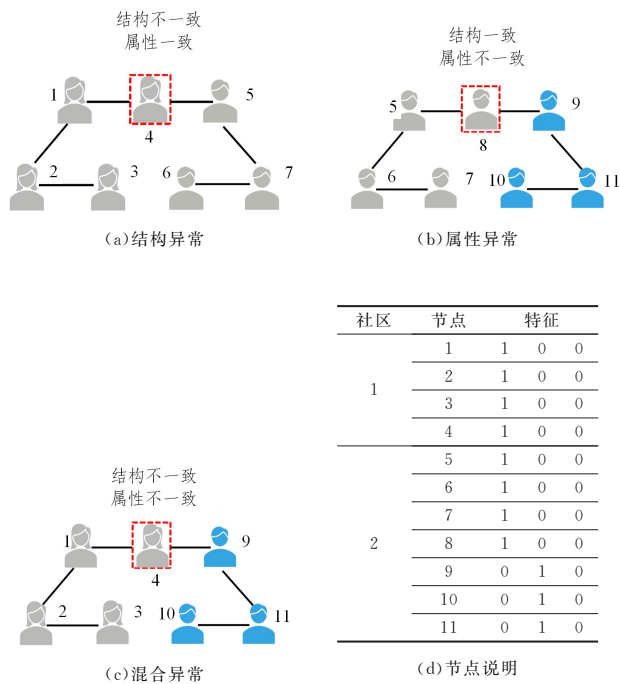


图1 属性图中的3种异常类型

Fig. 1 Three types of anomalies in the attributed networks

为了解决上述问题,本文提出了一种双重分类和重建网络(DCRN)。在域自适应阶段,DCRN不仅在结构特征上进行域自适应,还引入了属性特征的域自适应,能够同时对源图和目标图的结构与属性特征进行有效对齐,进而更好地捕捉跨域图中的复杂异常模式。在异常检测阶段,DCRN在通过异常分类器检测目标图中的共享异常的基础上,结合重建误差评估机制来增强对非共享异常的检测能力,从而显著提高模型的检测精度和鲁棒性。本文的主要贡献如下:

- 1)提出了DCRN框架,利用结构与属性共享编码器实现了源图与目标图的特征对齐与知识迁移;
- 2)在异常检测中,联合分类器和解码器,能够同时检测目标图中的共享异常和非共享异常;
- 3)在多个真实数据集上的广泛实验验证了模型的有效性,结果表明DCRN在跨域图异常检测任务中具有显著的性能优势。

## 2 相关工作

### 2.1 图异常检测

近年来,图异常检测方法<sup>[17-19]</sup>受到了广泛关注。由于异常标签稀缺,因此这类方法通常以无监督方式进行。传统方法如ALAD<sup>[20]</sup>,Radar<sup>[21]</sup>和SGASD<sup>[22]</sup>等,通过使用统计模型、矩阵分解和残差分析等技术提取节点的结构和属性模式,进行后续的异常检测。尽管这些非深度学习技术能够有效捕捉图的拓扑结构和节点属性信息,但其计算成本较高,在大规模

模现实网络中应用受限。

随着图神经网络(Graph Neural Networks, GNNs)的快速发展,基于深度学习的图异常检测方法取得了显著进展。例如, Dominant<sup>[23]</sup>, AnomalyDAE<sup>[24]</sup>和 ComGA<sup>[25]</sup>等方法利用深度自编码器技术,通过节点的重建误差来评估异常性。此外, CoLA<sup>[26]</sup>, SL-GAD<sup>[27]</sup>和 Sub-CR<sup>[28]</sup>等方法采用图对比学习,通过构造正负样本对,基于节点嵌入的差异来评估节点的异常性。然而,这些方法大多侧重于单域图异常检测,在 CD-GAD 方面的研究相对有限。

## 2.2 无监督域适应

UDA 通过将不同但相关的源域和目标域数据映射到一个共享的特征空间,并在该特征空间中最小化两者的分布差异,来实现跨域适应<sup>[29-30]</sup>。目前主流的方法分为 3 类:基于差异的方法、基于重建的方法和基于对抗的方法。

基于差异的方法通过最小化源域和目标域在特征空间中的分布差异来实现,常用度量包括最大均值差异<sup>[31]</sup>、高阶中心矩差异<sup>[32]</sup>、对比域差异<sup>[33]</sup>和 Wasserstein 度量<sup>[34]</sup>。这些方法通过直接减少特征分布之间的差异,确保模型能在两个域中保持一致性,进而实现域适应。

基于重建的方法使用自编码器模型提取具有迁移能力的特征。例如, DRCN<sup>[35]</sup>将源域标签嵌入编码器的学习过程中,通过同步训练编码器和分类器,使得编码的特征对类别具有更好的区分能力。DSN<sup>[36]</sup>则通过领域共享编码器和领域专属编码器提取特征,并通过正交约束保证两者彼此独立。同时,它通过重建原始输入,确保提取的特征信息完整且适用于跨域任务。

基于对抗的方法受到生成对抗网络(GAN)<sup>[37]</sup>的启发,通过训练一个域鉴别器来度量源域和目标域间的分布差异。经典的 DANN<sup>[38]</sup>对抗性训练域鉴别器以区分源域和目标域,同时训练特征提取器以混淆鉴别器,学习到域不变特征。CDAN<sup>[39]</sup>进一步在 DANN 基础上整合分类信息,增强特征提取和域鉴别的协同效果。

针对跨域图异常检测, COMMANDER<sup>[9]</sup>通过对抗学习实现域对齐,并使用在源域数据上训练的异常分类器来检测目标图中的异常节点。ACT<sup>[13]</sup>则结合对比学习与 Wasserstein<sup>[34]</sup>度量进行域对齐,随后通过偏差学习检测目标图的异常节点。然而,这些方法在领域对齐过程中仅关注图的网络结构特征,未考虑到属性特征在不同领域中的重要性和差异,导致模型在属性特征的域适应方面表现不佳,从而影响异常检测的准确性。

## 2.3 符号说明

表 1 对本文使用的符号进行了定义。

本文使用斜体字母代表符号、函数名以及变量,用斜体加粗字母代表矩阵和向量。用  $G=(V, E, \mathbf{X})$  表示一个图结构,其中  $V=\{v_1, v_2, \dots, v_n\}$  和  $E=\{e_1, e_2, \dots, e_m\}$  分别表示节点集合和边集合,  $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{n \times d}$  表示节点的  $n$  个  $d$  维属性矩阵。该属性图可以简化表示为  $G=(\mathbf{X}, \mathbf{A})$ , 其中  $\mathbf{A}=\{0, 1\}^{n \times n}$  为属性图的邻接矩阵, 当  $A_{i,j}=1$  时, 表示节点  $v_i$  和  $v_j$  之间存在一条边, 否则  $A_{i,j}=0$ 。

表 1 符号及其定义

Table 1 Symbols and their definitions

符号	意义
$G=(\mathbf{X}, \mathbf{A})$	属性图
$\mathbf{X}$	$G$ 的属性矩阵
$\mathbf{A}$	$G$ 的邻接矩阵
$V$	$G$ 的节点集合
$E$	$G$ 的边集合
$Z_s^s$	通过结构共享编码器得到的源图嵌入
$Z_t^s$	通过结构共享编码器得到的目标图嵌入
$Z_s^t$	通过属性共享编码器得到的源图嵌入
$Z_t^t$	通过属性共享编码器得到的目标图嵌入
$N_{\text{ano}}$	源图中正常节点个数
$N_{\text{nor}}$	源图中目标节点个数
$\alpha$	控制结构特征和属性特征在总损失中的权重比例
$\lambda$	控制分类器和解码器在总损失中的权重比例
$\theta$	结构解码器中的加权矩阵,用于加大缺失边缘的重建误差惩罚
$\eta$	属性解码器中的加权矩阵,用于加大缺失属性的重建误差惩罚

## 3 DCRN

### 3.1 问题定义

问题:给定一个源域属性图  $G_s=(\mathbf{X}_s, \mathbf{A}_s)$  和一个目标域属性图  $G_t=(\mathbf{X}_t, \mathbf{A}_t)$ , 其中源域为有标签的属性图,目标域为无标签的属性图。假设源图  $G_s$  和目标图  $G_t$  共享相同的特征空间,但在节点和边上没有重叠。目标是学习一个异常检测模型,该模型能够从源图  $G_s$  中概括标记信息,并将其迁移至目标图  $G_t$  以检测目标图上的异常节点。也就是说,模型的目标是学习一个评分函数:  $S: V_i \rightarrow y_i \in \mathbb{R}$ , 该函数能够结合源域的异常判别知识,并根据阈值  $\theta$  对样本  $x_i$  进行分类<sup>[24]</sup>。

$$y_i = \begin{cases} 1, & \text{if } S(V_i) \geq \theta \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

其中,  $y_i$  表示样本  $x_i$  的标签, 0 为正常样本, 1 为异常样本。

### 3.2 DCRN 框架概述

如图 2 所示, DCRN 是一个端到端的联合表示学习框架,其核心思想是将源图中标记的异常检测知识迁移至目标图,从而在目标图上训练异常节点检测模型。DCRN 分为域自适应和异常检测两个主要阶段,涵盖了 6 个关键组件。

首先,在域自适应阶段, DCRN 通过基于重构的域自适应策略,分别利用结构共享编码器和属性共享编码器对源图与目标图的网络结构和节点属性进行编码,生成对应的节点嵌入和属性嵌入。随后,结构异常分类器和属性异常分类器通过有监督学习提取源图的异常检测知识,结构解码器和属性解码器则通过无监督的方式对目标图的网络结构和节点属性进行重建,评估重构误差以支持后续的异常检测。通过编码器、分类器和解码器的联合优化, DCRN 能够有效对齐源图与目标图的结构和属性特征,将其映射到域不变的特征空间,从而实现特征对齐与知识迁移。

在完成域对齐后, DCRN 利用结构共享编码器和属性共享编码器将目标域的结构特征和属性特征映射到与源域一致的特征空间。随后,源域数据训练的结构异常分类器和属性异常分类器负责检测目标图中的共享异常,结构解码器和属

性解码器则通过评估重构误差辅助检测目标图中的非共享异常。通过融合分类器与解码器的检测结果,DCRN 能够更加全面地捕捉目标图中的异常模式。

相比现有方法,DCRN 在结构特征域自适应的基础上,引

入属性特征的域自适应,实现了更全面的跨域特征对齐。此外,通过分类器与解码器的联合检测机制,DCRN 能够同时检测共享异常和非共享异常,显著提升了异常检测的精度和鲁棒性。

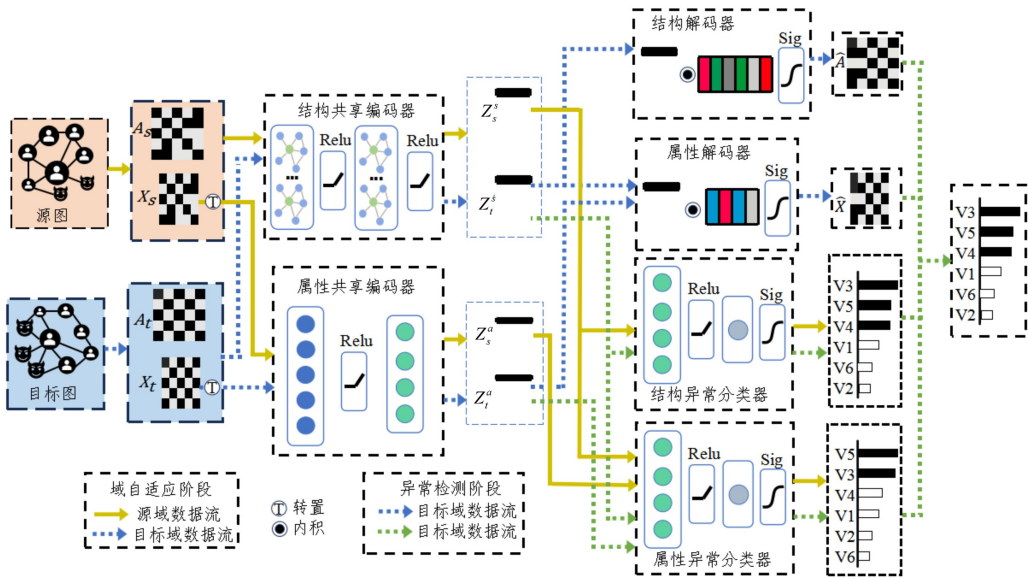


图2 DCRN 的跨域重构对齐与异常检测机制

Fig. 2 Reconstruction-aligned cross-domain anomaly detection in DCRN

### 3.3 属性图上的域适应

在训练过程中,DCRN 首先通过结构共享编码器和属性共享编码器对源图与目标图的网络结构与节点属性分别进行编码,生成节点嵌入和属性嵌入。随后,利用结构异常分类器和属性异常分类器对源图的嵌入进行有监督的异常分类,同时通过结构共享编码器和属性共享编码器对目标图的嵌入进行无监督重建。通过对编码器、分类器与解码器的联合优化,结构共享编码器和属性共享编码器能够逐步将源图和目标图的网络结构与节点属性映射到域不变的特征空间,从而实现特征对齐与知识迁移。

#### 1) 结构共享编码器

结构共享编码器将属性图的邻接矩阵和属性矩阵作为输入,利用图卷积神经网络(Graph Convolutional Network, GCN)学习源图和目标图的节点表示。编码器采用两层 GCN 提取网络结构特征,其卷积操作定义如下:

$$\tilde{H}^1 = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \tilde{X} \mathbf{W}^{s(1)}) \quad (2)$$

$$\mathbf{Z}_s^s = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \tilde{H}^1 \mathbf{W}^{s(2)}) \quad (3)$$

其中,  $\tilde{A} = \mathbf{A} + \mathbf{I}$ ,  $\tilde{D} = \Sigma \tilde{A}$ ,  $\mathbf{I}$  是单位矩阵,  $\sigma(\cdot)$  为激活函数(如 ReLU 和 Tanh),  $\tilde{X}$  为属性矩阵,  $\mathbf{W}^{s(1)}$  和  $\mathbf{W}^{s(2)}$  分别为第一层和第二层的网络权重矩阵,  $\mathbf{Z}_s^s$  为结构共享编码器的最终输出。通过对源图和目标图的结构特征进行编码,结构共享编码器将这些特征映射到统一的特征空间。源图和目标图的结构嵌入分别表示为  $\mathbf{Z}_s^s$  和  $\mathbf{Z}_t^s$ 。

#### 2) 属性共享编码器

属性共享编码器仅以属性图的属性矩阵为输入,通过多层感知机(Multilayer Perceptron, MLP)学习源图和目标图的属性表示。具体过程如下:

$$\tilde{\mathbf{Z}}^a = \sigma(\mathbf{X}^T \mathbf{W}^{a(1)} + \mathbf{b}^{a(1)}) \quad (4)$$

$$\mathbf{Z}_s^a = \tilde{\mathbf{Z}}^a \mathbf{W}^{a(2)} + \mathbf{b}^{a(2)} \quad (5)$$

其中,  $\mathbf{X}$  为输入的属性矩阵,  $\mathbf{W}^{a(1)}$  和  $\mathbf{W}^{a(2)}$  分别为两层全连接层网络的权重矩阵,  $\mathbf{b}^{a(1)}$  和  $\mathbf{b}^{a(2)}$  为偏置项,  $(\cdot)^T$  表示矩阵的转置运算,  $\mathbf{Z}_s^a$  为属性共享编码器的最终输出。源图和目标图的属性嵌入分别表示为  $\mathbf{Z}_s^a$  和  $\mathbf{Z}_t^a$ 。

#### 3) 结构异常分类器

结构异常分类器用于处理通过结构共享编码器生成的源域节点嵌入  $\mathbf{Z}_s^s$ , 辅助结构共享编码器捕捉源域和目标域的结构不变特征。该分类器由两层非线性特征变换组成,定义如下:

$$\mathbf{H}_s^s = \sigma(\mathbf{Z}_s^s \mathbf{W}_c^{s(1)} + \mathbf{b}_c^{s(1)}) \quad (6)$$

$$\bar{\mathbf{y}}_s^s = \text{sigmoid}(\mathbf{H}_s^s \mathbf{W}_c^{s(2)} + \mathbf{b}_c^{s(2)}) \quad (7)$$

其中,  $\mathbf{W}_c^{s(1)}$  和  $\mathbf{W}_c^{s(2)}$  分别为第一层和第二层的权重矩阵,  $\mathbf{b}_c^{s(1)}$  和  $\mathbf{b}_c^{s(2)}$  为偏置矩阵,  $\text{sigmoid}(\cdot)$  为 sigmoid 激活函数,  $\bar{\mathbf{y}}_s^s$  为结构异常分类器预测的节点标签矩阵。

结构异常分类器的损失函数为加权二元交叉熵损失:

$$\mathcal{L}_c^s = - \left( \frac{N_{\text{nor}}}{N_{\text{ano}}} y \ln \bar{y}_s^s + (1-y) \ln (1-\bar{y}_s^s) \right) \quad (8)$$

其中,  $N_{\text{nor}}$  和  $N_{\text{ano}}$  分别表示源图中正常节点和异常节点的数量,  $y$  是源图节点的真实标签,  $\bar{y}_s^s$  是结构异常分类器对源图节点的预测标签,  $\mathcal{L}_c^s$  是结构异常分类器的损失。该损失函数通过加权调整不同类别节点的影响,缓解了类别不平衡问题对模型训练的影响。

#### 4) 属性异常分类器

属性异常分类器处理由结构共享编码器和属性共享编码器生成的源域节点嵌入  $\mathbf{Z}_s^s$  和属性嵌入  $\mathbf{Z}_s^a$ , 辅助结构共享编码器和属性共享编码器捕捉源域与目标域的结构

不变特征和属性不变特征。该分类器由两层全连接层构成,定义如下:

$$\mathbf{H}_c^a = \sigma(\mathbf{Z}_s^a (\mathbf{Z}_s^a)^\top \mathbf{W}_c^{a(1)} + \mathbf{b}_c^{a(1)}) \quad (9)$$

$$\bar{\mathbf{y}}_s^a = \text{sigmoid}(\mathbf{H}_c^a \mathbf{W}_c^{a(2)} + \mathbf{b}_c^{a(2)}) \quad (10)$$

其中,  $\mathbf{W}_c^{a(1)}$  和  $\mathbf{W}_c^{a(2)}$  为两层全连接层的权重矩阵,  $\mathbf{b}_c^{a(1)}$  和  $\mathbf{b}_c^{a(2)}$  为偏置矩阵,  $\bar{\mathbf{y}}_s^a$  为预测的节点标签矩阵。

属性异常分类器的损失函数同样采用加权二元交叉熵损失:

$$\mathcal{L}_c^a = - \left( \frac{N_{\text{nor}}}{N_{\text{ano}}} y \ln \bar{y}_s^a + (1-y) \ln(1-\bar{y}_s^a) \right) \quad (11)$$

同样地,  $N_{\text{nor}}$  和  $N_{\text{ano}}$  表示源图中正常节点和异常节点的数量,  $y$  是源图节点的真实标签,  $\bar{y}_s^a$  是属性异常分类器对源图节点的预测标签,  $\mathcal{L}_c^a$  是属性异常分类器的损失。

### 5) 结构解码器

结构解码器用于处理目标域经过结构共享编码器得到的结构嵌入  $\mathbf{Z}_t$ , 通过计算节点嵌入间的相似性来重建原始网络的邻接矩阵  $\mathbf{A}_t$ , 从而辅助结构共享编码器捕捉源域和目标域的结构不变特征。具体而言, 结构解码器以节点嵌入  $\mathbf{Z}_t$  作为输入, 重建邻接矩阵:

$$\tilde{\mathbf{A}}_t = \text{sigmoid}(\mathbf{Z}_t (\mathbf{Z}_t)^\top) \quad (12)$$

其中,  $(\cdot)^\top$  表示矩阵的转置运算,  $\tilde{\mathbf{A}}_t$  为重建之后的邻接矩阵。

结构解码器的重建损失定义如下:

$$\mathcal{L}_t^s = \|\tilde{\mathbf{A}}_t - \mathbf{A}_t\|_{\text{F}}^2 \quad (13)$$

$$\theta_{i,j} = \begin{cases} 1, & \text{if } A_{i,j} = 0 \\ \theta, & \text{otherwise} \end{cases} \quad (14)$$

其中,  $\odot$  是 Hadamard 乘积,  $\|\cdot\|_{\text{F}}^2$  表示 Frobenius 范数,  $\mathcal{L}_t^s$  为结构解码器的重建误差。  $\theta > 1$  用于对现实世界中边缘缺失导致的非零元素重建误差施加更大的惩罚。通过这种加权方式, 模型能够更好地应对由于数据稀疏性或边缘缺失引发的误差问题, 从而提高结构解码器的重建精度。

### 6) 属性解码器

属性解码器用于处理目标域经过结构共享编码器和属性共享编码器得到的结构嵌入  $\mathbf{Z}_t$  和属性嵌入  $\mathbf{Z}_t^a$ , 通过重建目标图的属性矩阵  $\mathbf{X}_t$  来辅助结构共享编码器和属性共享编码器捕捉源域和目标域的结构不变特征和属性不变特征。属性解码器以  $\mathbf{Z}_t$  和  $\mathbf{Z}_t^a$  为输入, 重建属性矩阵:

$$\tilde{\mathbf{X}}_t = \text{sigmoid}(\mathbf{Z}_t (\mathbf{Z}_t^a)^\top) \quad (15)$$

属性解码器的重建损失定义如下:

$$\mathcal{L}_t^a = \|\tilde{\mathbf{X}}_t - \mathbf{X}_t\|_{\text{F}}^2 \quad (16)$$

$$\eta_{i,j} = \begin{cases} 1, & \text{if } A_{i,j} = 0 \\ \eta, & \text{otherwise} \end{cases} \quad (17)$$

与结构解码器类似, 上式中  $\odot$  是 Hadamard 乘积,  $\|\cdot\|_{\text{F}}^2$  表示 Frobenius 范数,  $\mathcal{L}_t^a$  为属性解码器的重建误差。  $\eta > 1$ , 用于对现实世界中某些缺失属性导致的非零元素重建误差施加更大的惩罚。这种加权方式能够帮助模型更有效地减小小属性缺失带来的影响, 从而提高属性解码器的重建精度。

## 3.4 模型训练

为了实现源图和目标图的结构域自适应和属性域自适

应, 模型 DCRN 的每个组件都需要专门的训练目标。完整函数定义如下:

$$\mathcal{L} = \lambda(\alpha \mathcal{L}_c^a + (1-\alpha) \mathcal{L}_c^s) + (1-\lambda)(\alpha \mathcal{L}_t^a + (1-\alpha) \mathcal{L}_t^s) \quad (18)$$

其中,  $\mathcal{L}_c^a$ ,  $\mathcal{L}_c^s$ ,  $\mathcal{L}_t^a$ ,  $\mathcal{L}_t^s$  分别为属性异常分类器、结构异常分类器、属性解码器、结构解码器的损失, 参数  $\lambda$  用于平衡分类器和解码器的损失,  $\alpha$  控制结构和属性损失的比重。

算法 1 给出了 DCRN 的训练过程。通过最小化专用目标函数, DCRN 实现了从源图到目标图的域逐步转移, 使结构共享编码器和属性共享编码器能够逐步将源图和目标图的网络结构与节点属性映射到域不变的特征空间, 从而实现特征对齐与知识迁移。DCRN 的所有参数均通过标准反向传播算法进行优化。

### 算法 1 DCRN 训练过程

输入: 目标图  $G_t$ , 源图  $G_s$ , 源图中正常节点数量  $N_{\text{nor}}$ , 源图中异常节点数量  $N_{\text{ano}}$ , 训练轮数 epoch, 超参数  $\alpha, \lambda, \eta, \theta$

输出: 结构共享编码器、属性共享编码器、结构异常分类器、属性异常分类器的相关网络参数

1. 初始化网络参数

2. for  $i \leftarrow 1$  to epoch do

2.1. 使用结构共享编码器, 通过式(3)生成源图和目标图的节点嵌入  $\mathbf{Z}_s^a$  和  $\mathbf{Z}_t^a$ ;

2.2. 使用属性共享编码器, 通过式(5)生成源图和目标图的属性嵌入  $\mathbf{Z}_s^a$  和  $\mathbf{Z}_t^a$ ;

2.3. 使用结构异常分类器, 通过式(7)计算源图节点嵌入  $\mathbf{Z}_s^a$  对应的异常分数  $\bar{y}_s^a$ ;

2.4. 使用属性异常分类器, 通过式(10)计算源图属性嵌入  $\mathbf{Z}_s^a$  对应的异常分数  $\bar{y}_s^a$ ;

2.5. 使用结构解码器, 通过式(12)重建目标图的邻接矩阵  $\tilde{\mathbf{A}}_t$ ;

2.6. 使用属性解码器, 通过式(15)重建目标图的属性矩阵  $\tilde{\mathbf{X}}_t$ ;

2.7. 通过式(18)计算模型的整体损失;

2.8. 梯度下降更新模型参数;

3. 结束循环。

### 3.5 跨域异常检测

经过域对齐后, 源图和目标图的网络结构和节点属性被映射到域不变的特征空间。DCRN 能够直接利用在源图上训练的结构异常分类器和属性异常分类器, 检测目标图中的共享异常。同时, 结构解码器和属性解码器通过评估重建误差, 辅助检测目标图中的非共享异常, 并融合分类器与解码器的检测结果, 识别目标图上的异常节点。

#### 1) 分类器异常检测

分类器的检测结果由结构异常分类器和属性异常分类器组成, 两者协同工作, 检测目标图中的共享异常节点。

$$C_{v_i} = \alpha \bar{y}_i^a + (1-\alpha) \bar{y}_i^s \quad (19)$$

其中,  $\bar{y}_i^a$  和  $\bar{y}_i^s$  是目标域数据通过结构异常分类器和属性异常分类器得到的目标图的属性异常和结构异常,  $\alpha$  用于平衡属性异常分类器和结构异常分类器的权重,  $C_{v_i}$  为分类器预测节点  $v_i$  的异常分数。

#### 2) 解码器异常检测

解码器的检测结果由结构解码器和属性解码器共同

组成,它们分别对目标图的邻接矩阵和属性矩阵进行重构,并将结构与属性的重构误差作为异常分数。

$$O_v = \alpha \mathcal{L}_r^a + (1-\alpha) \mathcal{L}_r^s \quad (20)$$

$$R_{v_i} = \frac{O_{v_i} - \min_i O_{v_i}}{\max_i O_{v_i} - \min_i O_{v_i}} \quad (21)$$

其中, $O_v$ 是目标图网络结构和节点属性的重构误差,对其进行归一化得到 $R_{v_i}$ ,并将 $R_{v_i}$ 作为解码器预测节点 $v_i$ 的异常分数。

### 3) 联合检测

模型结合异常分类器和解码器的检测结果,选取两者中较大的作为节点的异常分数。这是因为分类器和解码器分别在不同维度捕捉异常信息:分类器侧重于检测共享异常,解码器通过重构误差更敏感地捕捉非共享异常。因此,选择较大的检测结果能够确保当一方出现误差或遗漏时,另一方可以弥补不足,最大限度地发挥分类器和解码器在检测共享与非共享异常上的优势,从而提升异常检测的准确性和鲁棒性。

$$S_{v_i} = \max(C_{v_i}, R_{v_i}) \quad (22)$$

其中, $\max(\cdot)$ 表示取两者中的较大值, $S_{v_i}$ 为节点 $v_i$ 的最终异常分数。

算法2给出了DCRN的异常检测过程。在异常检测过程中,模型只需要输入目标域数据,异常分类器和解码器进行联合检测,最终得到目标图的异常分数。

#### 算法2 DCRN 推理过程

输入:目标图 $G_t$ ,超参数 $\alpha, \lambda, \eta, \theta$

输出:目标图异常分数 $S$

1. 初始化目标图节点的异常分数 $S$
2. 使用结构共享编码器,通过式(3)得到目标图节点嵌入 $Z_t^s$
3. 使用属性共享编码器,通过式(5)得到目标图的属性嵌入 $Z_t^a$
4. 使用结构异常分类器,通过式(7)计算目标图节点嵌入 $Z_t^s$ 对应的异常分数 $\bar{y}^a$
5. 使用属性异常分类器,通过式(10)计算目标图属性嵌入 $Z_t^a$ 对应的异常分数 $\bar{y}^s$
6. 使用结构解码器,通过式(13)得到目标图结构结构重构误差 $\mathcal{L}_r^s$
7. 使用属性解码器,通过式(16)得到目标图的属性重构误差 $\mathcal{L}_r^a$
8. 通过式(19)得到目标图经过异常分类器得到的异常分数 $C$
9. 通过式(21)得到目标图经过重构误差得到的异常分数 $R$
10. 通过式(22)得到目标图的最终异常分数 $S$

## 4 实验设置

为了验证DCRN的有效性,本章详细介绍了实验设置,包括所使用的数据集、基线方法、参数设置及评估指标。随后,展示了DCRN在异常检测任务中的实验结果,并将其与当前最先进的方法进行了对比分析。此外,还进行了消融实验、参数敏感性分析和可视化分析,以全面评估所提出方法的有效性。

### 4.1 数据集

选择4个公开的图异常检测数据集<sup>[9-13]</sup>,即YelpHotel, YelpRes, YelpNYC, Amazon,并组成了4组跨域数据集,其中 $A \Rightarrow B$ 表示数据集A和B互为源域和目标域。表2列出了各数据集的统计信息。

表2 数据集信息

Table 2 Dataset information

Dataset	Dim	Nodes	Avg Deg	Ration
YelpRes	8000	5012	41.79	0.0499
YelpHotel	8000	4322	23.55	0.0578
YelpNYC	10000	21040	78.81	0.0475
Amazon	10000	18601	28.30	0.0390

YelpHotel $\Leftrightarrow$ YelpRes:这两个数据集均来自Yelp平台,分别代表芝加哥地区的酒店和餐厅业务领域。该数据集中,用户作为节点,当两个用户对同一家酒店或餐厅进行评论时,创建一条边。通过Yelp的反欺诈过滤机制,用户被分为异常用户和正常用户,这些分类标签可作为真实标签,用于异常检测模型的评估。

YelpNYC $\Leftrightarrow$ Amazon:YelpNYC数据集同样来源于Yelp平台,Amazon数据集则来自电子商务平台。在Amazon数据集中,若用户对两个或多个众包任务产品发表评论,则该用户被标记为异常用户;否则,用户被视为正常用户。这两个数据集来自不同平台,存在较大的领域差异,适合用于研究不同域差异对模型性能的影响。

### 4.2 评估指标

为了全面评估异常检测模型的性能,本文选取了AUC-ROC, AUC-PR<sup>[40]</sup>和FAR<sup>[30]</sup>这3个指标。每个指标针对模型性能的不同方面进行评价,共同构成对异常检测模型的全方位评估框架。其中,AUC-ROC侧重于对整体分类能力的衡量,AUC-PR专注于评估模型对少数正例的识别性能,而FAR则用于评估模型对负例的误报情况。这种多角度的评估方式,能够从分类能力、正例识别能力和鲁棒性3个维度全面反映模型性能。

各个指标的计算涉及到混淆矩阵的相关概念,其组成部分如表3所列。

表3 混淆矩阵

Table 3 Confusion matrix

真实类别	预测类别	
	正样本	负样本
正样本	TP	FN
负样本	FP	TN

TP代表真实类别为正样本而预测类别为正样本的数量;FP代表真实类别为负样本而预测类别为正样本的数量;FN代表真实类别为正样本而预测类别为负样本的数量;TN代表真实类别为负样本而预测类别为负样本的数量<sup>[40]</sup>。

AUC-ROC(Area Under the Receiver Operating Characteristic Curve)用于评估分类模型的性能。ROC曲线的横轴是假阳性率(False Positive Rate, FPR),纵轴是真阳性率(True Positive Rate, TPR),曲线下的面积即为AUC-ROC值,其取值范围为0~1。

AUC-ROC值反映模型将一个正样本排在一个负样本之前的概率。AUC-ROC值越接近1,模型的分类性能越好。

$$TPR = \frac{TP}{TP + FN} \quad (23)$$

$$FPR = \frac{FP}{FP + TN} \quad (24)$$

由于 AUC-ROC 对正负样本不均衡的影响较小,因此在异常检测任务中,AUC-ROC 是衡量模型分类能力的重要指标。

AUC-PR(Area Under the Precision-Recall Curve)则更适用于处理类别高度不平衡的数据集。PR 曲线以召回率(Recall)为横轴,精确率(Precision)为纵轴,通过计算 PR 曲线下的面积衡量模型对正例样本的识别性能,其取值范围为0~1。

AUC-PR 专注于模型在正例上的表现,值越高表明模型在少数正例样本的识别上表现越好,尤其是在异常检测场景中,当正例(异常)样本比例极低时,AUC-PR 能更直观地反映模型性能。

$$Precision = \frac{TP}{TP + FP} \quad (25)$$

$$Recall = \frac{TP}{TP + FN} \quad (26)$$

与 AUC-ROC 相比,AUC-PR 更专注于少数类样本(即异常样本)的检测精度和覆盖度。在异常检测任务中,AUC-PR 能更有效地衡量模型减少误报(提高精确率)与捕捉更多异常样本(提高召回率)的能力。因此,AUC-PR 是衡量模型在异常检测中的精度与效果的重要指标。

FAR(False Alarm Rate)用于评估模型对负例样本(正常样本)的误分类情况。它反映的是在所有真实负例中被错误分类为正例(异常样本)的比例。FAR 的取值范围为0~1,值越低说明模型对负例的分类能力越强,从而减少误报带来的干扰。

$$FAR = \frac{FP}{FP + TN} \quad (27)$$

FAR 在异常检测中尤为重要,特别是在实际应用中,高误报率会显著增加系统成本和用户干扰。因此,FAR 是衡量异常检测模型鲁棒性的重要补充指标。

### 4.3 实验设计

将 DCRN 与 10 种相关方法进行了比较,包括无监督图异常检测方法和跨域图异常检测方法。每组实验使用 5 个不同的随机种子并计算其均值和标准差。具体而言,比较的 8 种无监督图异常检测方法分别为 SCAN<sup>[6]</sup>, ONE<sup>[12]</sup>, AnomalyDAE<sup>[24]</sup>, CoLA<sup>[26]</sup>, ANOMALOUS<sup>[41]</sup>, DONE<sup>[42]</sup>, AdONE<sup>[42]</sup> 和 OCGNN<sup>[43]</sup>, 2 种跨域图异常检测方法为 COMMANDER<sup>[9]</sup> 和 ACT<sup>[13]</sup>。

所提模型在 PyTorch 框架下实现,使用 Adam<sup>[44]</sup> 优化器进行训练。结构共享编码器、属性共享编码器、结构异常分类器和属性异常分类器均由两层全连接层组成,具体的网络结构参数如表 4 所列。学习率设定为 0.0005,训练过程持续 150 轮。

表 4 网络参数

Table 4 Network parameters

网络架构	结构共享 编码器	属性共享 编码器	结构异常 分类器	属性异常 分类器
第一层	256	256	128	128
激活函数	Relu	Relu	Relu	Relu
第二层	128	128	1	1
激活函数	Relu	—	Sigmoid	Sigmoid

对于 4 个数据集,超参数  $\eta$  和  $\theta$  均统一设置为 1 和 200,参数  $(\alpha, \lambda)$  的设置如下:YelpHotel $\rightarrow$ YelpRes 为 (0.7, 0.4), YelpRes $\rightarrow$ YelpHotel、YelpNYC $\rightarrow$ Amazon 和 Amazon $\rightarrow$ YelpNYC 均为 (0.8, 0.4),其中 A $\rightarrow$ B 表示在目标数据集 B 上使用源域数据集 A 进行图异常检测。主要实验在配备 NVIDIA A100 GPU 和 Intel<sup>(R)</sup> Xeon<sup>(R)</sup> Gold 6248R CPU 的平台上进行,CUDA 驱动程序版本为 515.65.01。

### 4.4 实验结果

在 4 种不同的域转移场景下,将所提出的框架与其他基线方法进行性能对比实验。实验结果如表 5、图 3 和图 4 所示(无监督图异常检测方法仅使用目标域数据进行检测)。表 5 列出了使用 5 个不同的随机种子在 10 种基准模型和 DCRN 在 4 个数据集上训练所得的 AUC-ROC 和 AUC-PR 的均值和方差,其中加粗的值代表当前数据集中表现最优的指标,下划线为次优指标。图 3 则展示了 11 种方法在 4 个数据集上的平均 AUC-ROC 和 AUC-PR 得分的可视化结果,其中 8 种无监督方法用青色点表示,2 种跨域方法用蓝色点表示,DCRN 则用红色点标注。图 4 为 11 种方法在 4 个数据集上的 FAR 指标。

由图 3 可以看出,DCRN 在 4 个数据集上的平均 AUC-ROC 和 AUC-PR 得分均为最高,表明其在跨域异常检测任务中有着优越的性能。同时,两种跨域异常检测方法 COMMANDER 和 CoLA 的表现也明显优于 8 种无监督图异常检测算法,进一步证明了跨域方法在应对域间差异时的有效性。

由表 5 可见,DCRN 在所有跨域数据集上的 AUC-ROC 和 AUC-PR 指标相较于最佳基线性能均有显著提升。具体而言,AUC-ROC 和 AUC-PR 在 YelpHotel $\rightarrow$ YelpRes 跨域数据集中分别提升了 0.041 和 0.088,在 YelpRes $\rightarrow$ YelpHotel 跨域数据集中分别提升了 0.121 和 0.218,在 YelpNYC $\rightarrow$ Amazon 跨域数据集中分别提升了 0.037 和 0.204,而在 Amazon $\rightarrow$ YelpNYC 跨域数据集中则分别提升了 0.014 和 0.039。其中,AUC-PR 指标在数据集 YelpRes $\rightarrow$ YelpHotel 和 YelpNYC $\rightarrow$ Amazon 上相比次优模型 ACT 的提升最为显著,这得益于 DCRN 在域自适应阶段同时采用了结构特征和属性特征的域自适应策略,并在异常检测阶段结合分类器与重构误差评估机制,有效识别共享异常和非共享异常。这种方法使 DCRN 能够更全面地捕捉目标图中的异常节点,显著提升了模型的检测性能。

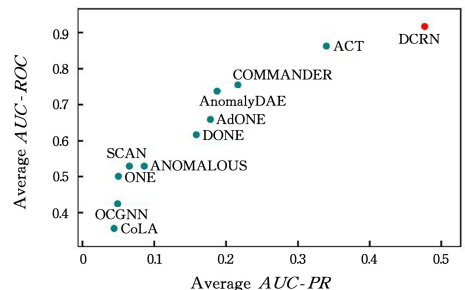


图 3 不同算法在 AUC-ROC 和 AUC-PR 上的比较(电子版为彩图)

Fig. 3 Comparison of different algorithms on AUC-ROC and AUC-PR

表 5 不同数据集上 AUC-ROC 和 AUC-PR 的比较  
Table 5 Comparison of AUC-ROC and AUC-PR on different datasets

Method	YelpHotel→YelpRes		YelpRes→YelpHotel		YelpNYC→Amazon		Amazon→YelpNYC	
	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR
SCAN <sup>[6]</sup>	0.500±0.000	0.049±0.000	0.726±0.000	0.118±0.000	0.366±0.000	0.040±0.000	0.529±0.000	0.053±0.000
ONE <sup>[12]</sup>	0.507±0.017	0.052±0.003	0.504±0.021	0.059±0.003	0.496±0.005	0.040±0.006	0.498±0.010	0.048±0.002
AnomalyDAE <sup>[24]</sup>	0.818±0.035	0.283±0.019	0.707±0.012	0.226±0.005	0.701±0.049	0.079±0.013	0.726±0.045	0.160±0.025
CoLA <sup>[26]</sup>	0.274±0.055	0.046±0.006	0.411±0.026	0.049±0.004	0.233±0.031	0.025±0.001	0.507±0.035	0.055±0.014
ANOMALOUS <sup>[41]</sup>	0.498±0.014	0.046±0.001	0.535±0.013	0.128±0.007	0.466±0.005	0.034±0.003	0.619±0.001	0.136±0.001
DONE <sup>[42]</sup>	0.649±0.051	0.211±0.015	0.436±0.003	0.175±0.004	0.831±0.003	0.128±0.002	0.550±0.013	0.119±0.008
AdONE <sup>[42]</sup>	0.746±0.020	0.253±0.007	0.453±0.009	0.173±0.006	0.606±0.008	0.154±0.008	0.834±0.003	0.133±0.003
OCGNN <sup>[43]</sup>	0.500±0.000	0.049±0.000	0.500±0.000	0.058±0.000	0.200±0.000	0.040±0.000	0.500±0.000	0.048±0.000
COMMANDER <sup>[9]</sup>	0.774±0.007	0.268±0.006	0.690±0.009	0.210±0.007	0.859±0.007	0.242±0.019	0.699±0.006	0.145±0.001
ACT <sup>[13]</sup>	0.892±0.015	0.330±0.018	0.804±0.006	0.287±0.006	0.925±0.004	0.497±0.020	0.830±0.002	0.243±0.003
DCRN	<b>0.933±0.007</b>	<b>0.418±0.016</b>	<b>0.925±0.001</b>	<b>0.505±0.001</b>	<b>0.962±0.012</b>	<b>0.701±0.055</b>	<b>0.848±0.012</b>	<b>0.282±0.027</b>

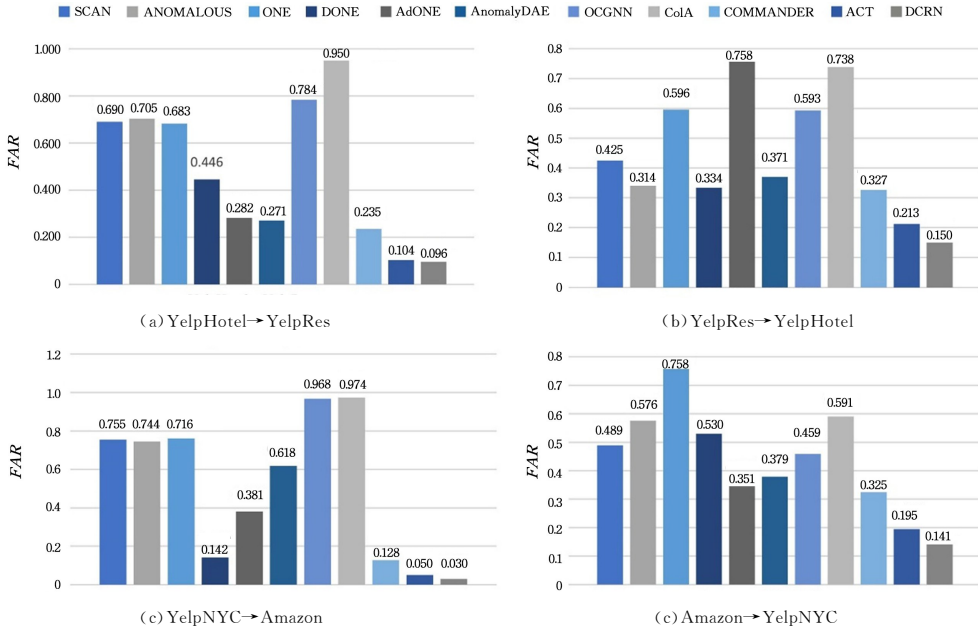


图 4 不同数据集上 FAR 的比较

Fig. 4 Comparison of FAR on different datasets

由图 4 可见,DCRN 在所有跨域数据集上的 FAR 指标均显著低于其他基线方法,表明其在图异常检测任务中具有更低的误报率。具体而言,在 YelpHotel→YelpRes 跨域数据集中,DCRN 的 FAR 仅为 0.096,相较于最佳基线方法 COMMANDER 降低了 0.139;在 YelpRes→YelpHotel 跨域数据集中,FAR 为 0.15,较最佳基线方法降低了 0.063;在 YelpNYC→Amazon 跨域数据集中,DCRN 的 FAR 为 0.03,显著低于最佳基线方法的 0.128;而在 Amazon→YelpNYC 跨域数据集中,DCRN 的 FAR 为 0.141,较最佳基线方法降低了 0.054。这些结果表明,DCRN 在跨域场景下能够有效减少误报率,提升异常检测的可靠性。此外,在处理跨域图数据的异常检测任务时,DCRN 能够更加有效地减少误报率,为实际应用提供了更具鲁棒性的解决方案。

#### 4.5 超参分析

本节探讨不同权重参数  $\alpha$  和  $\lambda$  对模型性能的影响,通过比较不同设置下的 AUC-ROC 和 AUC-PR 值,评估模型对这两个参数的敏感度。在跨域数据集 YelpHotel→YelpRes 上,针对  $\alpha$  和  $\lambda$  分别属于  $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$  的组合进行了实验。实验结果如图 5 和图 6 所示。

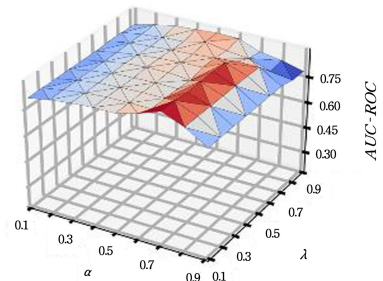


图 5 超参对 AUC-ROC 值的影响

Fig. 5 Effect of hyperparameters on AUC-ROC value

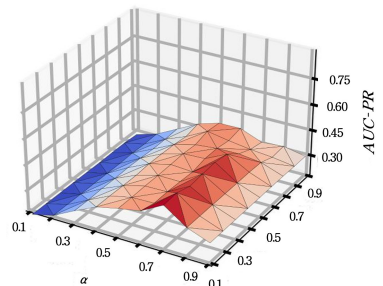


图 6 超参对 AUC-PR 值的影响

Fig. 6 Effect of hyperparameters on AUC-PR value

可以看出,  $\alpha$  和  $\lambda$  的值分别为 0.7 和 0.4 时, AUC-ROC 和 AUC-PR 的值达到相对较高的水平。这表明在 YelpHotel $\rightarrow$ YelpRes 数据集中, 属性特征在异常检测中起到了更为重要的作用, 且存在较多的非共享异常。

#### 4.6 可视化分析

为了验证所提出方法的有效性, 对模型在学习过程中的节点嵌入进行了可视化, 使用主成分分析 (Principal Component Analysis, PCA) 对其进行降维处理。

图 7 展示了 DCRN 模型在跨域数据集 Amazon $\rightarrow$ YelpNYC 上进行域对齐前后, 源图和目标图节点嵌入的变化情况, 具体包括初始状态 (见图 7(a))、训练至 50 轮 (见图

7(b)) 和 100 轮 (见图 7(c)) 时的中间状态, 以及最终对齐状态 (见图 7(d))。在初始阶段, 源图和目标图的节点嵌入在特征空间中呈现明显的分离, 随着训练轮次的增加, 分布逐渐靠拢, 最终实现了较好的对齐效果, 这表明 DCRN 能够有效处理复杂的图结构数据, 实现源图与目标图在特征空间上的对齐。从域对齐的最终状态可以观察到, 目标图中存在与源图特征分布相似的共享异常, 以及与正常节点分布明显偏离存在于目标图的非共享异常。对于共享异常, DCRN 通过基于源图训练的异常分类器进行有效检测; 对于非共享异常, 由于其与正常节点存在显著偏差, DCRN 可以通过解码器的重构误差对这些异常节点进行准确识别。

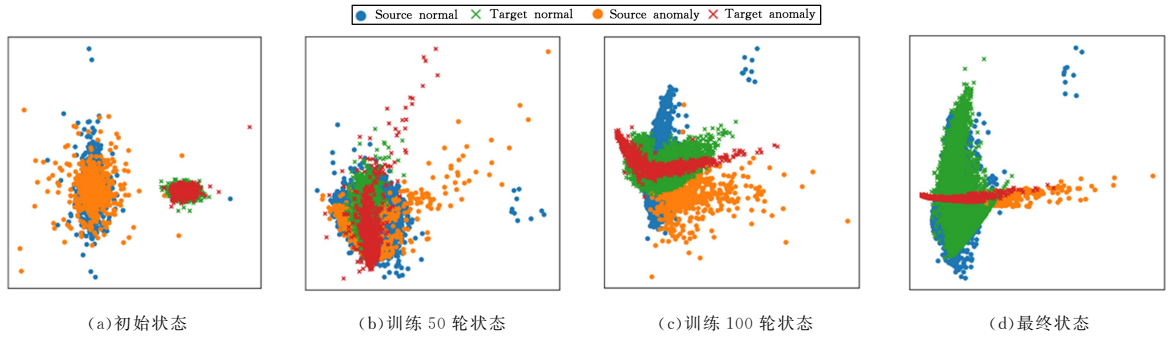


图 7 DCRN 联合学习期间节点表示空间的可视化

Fig. 7 Visualization of the node representation space during DCRN joint learning

#### 4.7 消融实验

为了验证所提方法的有效性, 以及模型各个模块的作用, 设计了 4 种消融实验。消融实验的详细设置如表 6 所列, 表中符号  $\checkmark$  表示该实验包含相应模块,  $\times$  表示不包含相应模块。

表 6 消融实验方法包含的模块

Table 6 Ablation experimental methodology includes modules

方法	结构共享	属性共享	结构	属性	结构异常	属性异常
	编码器	编码器	解码器	解码器	分类器	分类器
DCRN	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Wo_Stru	$\times$	$\checkmark$	$\times$	$\checkmark$	$\times$	$\checkmark$
Wo_Attr	$\checkmark$	$\times$	$\checkmark$	$\times$	$\checkmark$	$\times$
Wo_Dec	$\checkmark$	$\checkmark$	$\times$	$\times$	$\checkmark$	$\checkmark$
Wo_Clf	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\times$	$\times$

表 7 DCRN 模块消融实验结果的比较

Table 7 Comparison of DCRN module ablation experimental results

Method	YelpHotel $\rightarrow$ YelpRes		YelpRes $\rightarrow$ YelpHotel		YelpNYC $\rightarrow$ Amazon		Amazon $\rightarrow$ YelpNYC	
	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR
DCRN	<b>0.933<math>\pm</math>0.007</b>	<b>0.418<math>\pm</math>0.016</b>	<b>0.925<math>\pm</math>0.001</b>	<b>0.505<math>\pm</math>0.001</b>	<u>0.962<math>\pm</math>0.012</u>	<b>0.701<math>\pm</math>0.055</b>	<b>0.848<math>\pm</math>0.012</b>	<b>0.282<math>\pm</math>0.027</b>
Wo_Dec	<u>0.931<math>\pm</math>0.001</u>	0.416 $\pm$ 0.030	0.612 $\pm$ 0.076	0.272 $\pm$ 0.019	<b>0.962<math>\pm</math>0.078</b>	0.648 $\pm$ 0.027	0.441 $\pm$ 0.003	0.122 $\pm$ 0.020
Wo-Clf	0.392 $\pm$ 0.455	0.170 $\pm$ 0.171	0.693 $\pm$ 0.230	<u>0.321<math>\pm</math>0.184</u>	0.602 $\pm$ 0.472	0.427 $\pm$ 0.373	0.590 $\pm$ 0.158	0.222 $\pm$ 0.062
Wo-Stru	0.787 $\pm$ 0.001	0.322 $\pm$ 0.017	0.705 $\pm$ 0.002	0.259 $\pm$ 0.003	0.856 $\pm$ 0.001	0.174 $\pm$ 0.005	0.689 $\pm$ 0.001	0.143 $\pm$ 0.003
Wo-Attr	0.794 $\pm$ 0.011	0.129 $\pm$ 0.003	<u>0.824<math>\pm</math>0.032</u>	0.288 $\pm$ 0.012	0.646 $\pm$ 0.014	0.136 $\pm$ 0.018	0.410 $\pm$ 0.033	0.129 $\pm$ 0.022

方法 Wo\_Dec 表示在异常检测阶段, 仅将异常分类器的检测结果作为模型的最终异常分数; 方法 Wo\_Clf 则表示仅使用解码器的重构误差作为最终检测结果。从表 7 可以看出, 这两种方法的检测性能均低于完整模型 DCRN。这表明, 单独依赖分类器或解码器的检测机制无法全面捕捉目标图中的所有异常模式。相比之下, DCRN 通过融合异常分类器和解码器的检测结果, 能够更有效地检测共享异常和非共享异常, 显著提升了模型的检测精度和鲁棒性。

最终实验结果如表 7 所列, 其中加粗数据为当前数据集的最优指标, 下划线数据为当前数据集的次优指标。

具体来说, 方法 Wo\_Stru 和 Wo\_Attr 分别表示在域自适应过程中仅考虑结构特征的域自适应和仅考虑属性特征的域自适应。从表 7 可以看出, 这两种方法在所有数据集上的整体指标均低于完整模型 DCRN。这表明, 单独进行某一模态的域自适应不足以捕捉跨域图数据中的复杂异常模式。相比之下, DCRN 同时对结构特征和属性特征进行联合域自适应, 使得源图与目标图在多模态特征空间中实现了更全面的特征对齐, 从而能更好地应对复杂的图结构数据。

消融实验结果表明, DCRN 结构与属性特征的联合域自适应和分类器与解码器的联合检测机制, 在模型性能提升中发挥了关键作用。联合域自适应使得源图与目标图的多模态特征得以有效对齐, 从而更好地捕捉跨域图中的复杂异常模式; 而分类器与解码器的联合检测机制能够更加全面地检测共享异常和非共享异常, 显著提高了检测结果的可靠性和准确性。这些实验结果验证了 DCRN 在跨域图异常检测任务中的优越性, 说明了本文提出的模块设计在复杂图数据中的

实际应用价值。

**结束语** 本文提出了面向跨域图异常检测的深度分类和重建网络。在域自适应过程中,DCRN通过结构共享编码器和属性共享编码器,实现源图和目标图的结构域适应以及属性域适应,全面捕获图中的异常模式;在异常检测过程中,DCRN联合分类器和解码器的检测结果,识别源图中存在的共享异常和非共享异常,实现了在目标域上的精确异常检测。实验结果表明,DCRN在多个真实数据集上的性能显著优于现有的基线算法,充分验证了其在跨域图异常检测任务中的有效性。

DCRN尽管在多个真实数据集上表现出色,但仍存在改进空间。对于结构差异较大的跨域场景,当前模型的共享编码器可能无法全面捕捉源图与目标图之间的异质信息,限制了模型在复杂跨域场景中的表现,未来可以探索更灵活的域适应机制,如基于自适应图编码器或对抗性对齐方法,以提升模型对异质数据的适应能力。此外,DCRN在处理超大规模图数据时的效率问题仍需优化。由于模型采用基于GCN的全图计算方式,内存和计算资源需求显著增加,未来可以尝试引入基于邻居采样的迷你批次训练方法,或利用图切分与分布式训练技术来提升模型的扩展性。同时,属性图中异常类型的多样性可能影响分类器和解码器的鲁棒性。GCN在邻居节点聚合时缺乏灵活性和动态性,限制了其在处理复杂异常模式时的表现。为此,未来的研究可以采用图注意力网络,通过自适应调整邻居节点的权重,提升特征提取的灵活性,并结合对比学习增强异常类型的适应能力。

## 参 考 文 献

- [1] ZHANG Y Y, QIAN Y R, TAO W B, et al. Survey of Attribute Graph Anomaly Detection Based on Deep Learning[J]. *Computer Engineering and Applications*, 2022, 58(19): 1-13.
- [2] HAJLI N, SAEED U, TAJVIDI M, et al. Social bots and the spread of disinformation in social media: the challenges of artificial intelligence [J]. *British Journal of Management*, 2022, 33(3): 1238-1253.
- [3] LU Y J, SHENG L K, ZHANG Y L. Research on structural diversity of citation network oriented to evaluation of literature [J]. *Application Research of Computers*, 2024, 41(2): 408-414.
- [4] LIU H L, LIU Y X, XU J Y, et al. Research Progress in Application of Graph Anomaly Detection in Financial Anti-Fraud[J]. *Computer Engineering and Applications*, 2022, 58(22): 41-53.
- [5] LI Z, JIN X L, WANG Y J, et al. Anomaly Node Detection Method Based on Variational Graph Auto-Encoders in Attribute Networks[J]. *Pattern Recognition and Artificial Intelligence*, 2022, 35(1): 17-25.
- [6] XU X, YURUK N, FENG Z, et al. Scan: a structural clustering algorithm for networks [C] // *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2007: 824-833.
- [7] LI S, YU J, LU Y, et al. Self-supervised enhanced denoising diffusion for anomaly detection [J]. *Information Sciences*, 2024, 669: 120612.
- [8] ZHUANG Y S, LIN S L, LIN Z X, et al. Study on Generative Adversarial Network for Data Anomaly Detection[J]. *Computer Engineering and Applications*, 2022, 58(4): 143-149.
- [9] DING K, SHU K, SHAN X, et al. Cross-domain graph anomaly detection [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 33(6): 2406-2415.
- [10] LI Y, YANG X, GAO Q, et al. Cross-regional Fraud Detection via Continual Learning with Knowledge Transfer [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(12): 7865-7877.
- [11] ZHANG F, FAN H, WANG R, et al. Deep dual support vector data description for anomaly detection on attributed networks [J]. *International Journal of Intelligent Systems*, 2022, 37(2): 1509-1528.
- [12] BANDYOPADHYAY S, LOKESH N, MURTY M N. Outlier aware network embedding for attributed networks [C] // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019: 12-19.
- [13] WANG Q, PANG G, SALEHI M, et al. Cross-domain graph anomaly detection via anomaly-aware contrastive alignment [C] // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2023: 4676-4684.
- [14] LI Z, LIANG S, SHI J, et al. Cross-domain graph level anomaly detection [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(12): 7839-7850.
- [15] QU C, JI F, QIU M, et al. Learning to selectively transfer: Reinforced transfer learning for deep text matching [C] // *Proceedings of the twelfth ACM International Conference on Web Search and Data Mining*. 2019: 699-707.
- [16] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 7167-7176.
- [17] GUO J Y, LI R H, ZHANG Y, et al. Graph neural network based anomaly detection in dynamic networks [J]. *Journal of Software*, 2020, 31(3): 748-762.
- [18] CHEN B F, LI J D, LU X J, et al. Survey of Deep Learning Based Graph Anomaly Detection Methods [J]. *Journal of Computer Research and Development*, 2021, 58(7): 1436-1455.
- [19] MA X, WU J, XUE S, et al. A comprehensive survey on graph anomaly detection with deep learning [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 35(12): 12012-12038.
- [20] LIU N, HUANG X, HU X. Accelerated Local Anomaly Detection via Resolving Attributed Networks [C] // *IJCAI*. 2017: 2337-2343.
- [21] LI J, DANI H, HU X, et al. Radar: Residual analysis for anomaly detection in attributed networks [C] // *IJCAI*. 2017: 2152-2158.
- [22] WU L, HU X, MORSTATTER F, et al. Adaptive spammer detection with sparse group modeling [C] // *Proceedings of the International AAAI Conference on Web and Social Media*. 2017: 319-326.
- [23] DING K, LI J, BHANUSHALI R, et al. Deep anomaly detection on attributed networks [C] // *Proceedings of the 2019 SIAM International Conference on Data Mining*. Society for Industrial

- and Applied Mathematics, 2019:594-602.
- [24] FAN H, ZHANG F, LI Z. Anomalydae: Dual autoencoder for anomaly detection on attributed networks[C]// 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2020). IEEE, 2020:5685-5689.
- [25] LUO X, WU J, BEHESHTI A, et al. Comga: Community-aware attributed graph anomaly detection[C]// Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, 2022:657-665.
- [26] LIU Y, LI Z, PAN S, et al. Anomaly detection on attributed networks via contrastive self-supervised learning[J]. IEEE transactions on neural networks and learning systems, 2021, 33(6): 2378-2392.
- [27] ZHENG Y, JIN M, LIU Y, et al. Generative and contrastive self-supervised learning for graph anomaly detection [J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 35(12):12220-12233.
- [28] ZHANG J, WANG S, CHEN S. Reconstruction enhanced multi-view contrastive learning for anomaly detection on attributed networks[J]. arXiv:2205.04816, 2022.
- [29] HAN X, YUAN S. Unsupervised cross-system log anomaly detection via domain adaptation [C] // Proceedings of the 30th ACM International Conference on Information & Knowledge Management, 2021:3068-3072.
- [30] SU Y, TAN Y, AN S, et al. Semantic-driven dual consistency learning for weakly supervised video anomaly detection[J]. Pattern Recognition, 2025, 157:110898.
- [31] LONG M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks[C]// International Conference on Machine Learning. PMLR, 2015:97-105.
- [32] ZELLINGER W, GRUBINGER T, LUGHOFFER E, et al. Central moment discrepancy (CMD) for domain-invariant representation learning[J]. arXiv:1702.08811, 2017.
- [33] KANG G, JIANG L, YANG Y, et al. Contrastive adaptation network for unsupervised domain adaptation [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019:4893-4902.
- [34] FLAMARY R, COURTY N, TUIA D, et al. Optimal transport for domain adaptation[J]. arXiv:1507.00504, 2016.
- [35] GHIFARY M, KLEIJN W B, ZHANG M, et al. Deep reconstruction-classification networks for unsupervised domain adaptation[C]// Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part IV 14. Springer International Publishing, 2016:597-613.
- [36] BOUSMALIS K, TRIGEORGIS G, SILBERMAN N, et al. Domain separation networks[C]// Proceedings of the 30<sup>th</sup> International Conference on Neural Information Processing Systems, 2016:343-351.
- [37] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// Proceedings of the 28<sup>th</sup> International Conference on Neural Information Processing Systems, 2014:2672-2680.
- [38] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. Journal of Machine Learning Research, 2016, 17(59):1-35.
- [39] LONG M, CAO Z, WANG J, et al. Conditional adversarial domain adaptation[C]// Proceedings of the 32<sup>nd</sup> International Conference on Neural Information processing Systems, 2018:1647-1657.
- [40] KONG L C, LIU G Z. Review of Outlier Detection Algorithms [J]. Computer Science, 2024, 51(8):20-33.
- [41] PENG Z, LUO M, LI J, et al. ANOMALOUS: A Joint Modeling Approach for Anomaly Detection on Attributed Networks[C]// IJCAI. 2018:3513-3519.
- [42] BANDYOPADHYAY S, LOKESH N, VIVEK S V, et al. Outlier resistant unsupervised deep architectures for attributed network embedding [C] // Proceedings of the 13th International Conference on Web Search and Data Mining, 2020:25-33.
- [43] WANG X, JIN B, DU Y, et al. One-class graph neural networks for anomaly detection in attributed networks[J]. Neural Computing and Applications, 2021, 33:12073-12085.
- [44] DIEDERIK P K. Adam: A method for stochastic optimization [J]. arXiv:1412.6980, 2014.



**SU Shiyu**, born in 1998, postgraduate, is a member of CCF (No. W8728G). His main research interests include anomaly detection and attribute graphs.



**YU Jiong**, born in 1964, Ph.D, professor, Ph.D supervisor. His main research interests include network security and distributed computing.

(责任编辑:何杨)