智能系统学报
CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

融合图滤波与注意力机制的图异常检测方法

王天昊, 窦浩, 王翔, 毛国君

引用本文:
王天昊, 窦浩, 王翔, 等. 融合图滤波与注意力机制的图异常检测方法[J]. 智能系统学报, 2026, 21(4): 919-931.
WANG TianHao, DOU Hao, WANG Xiang, et al. Graph anomaly detection method integrating graph filtering and an attention mechanism[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(4): 919-931.

在线阅读 View online: https://dx.doi.org/10.11992/tis.202510033

DOI: 10.11992/tis.202510033
网络出版地址: https://link.cnki.net/urlid/23.1538.tp.20260422.1848.006

融合图滤波与注意力机制的图异常检测方法

王天昊[1], 窦浩[1], 王翔[1,2], 毛国君[1,2]
(1. 福建理工大学 计算机科学与数学学院, 福建 福州 350118; 2. 福建理工大学 福建省大数据挖掘与应用技术重点实验室, 福建 福州 350118)

摘 要: 为缓解图神经网络异常检测中普遍存在的异配连接与特征不一致问题, 本文提出了一种结合图滤波与注意力机制的图异常检测方法 (flexible high-low frequency and attention neural network, FHANN)。针对异配连接问题, 设计了图滤波模块, 可以自适应提取高频与低频信号并融合生成节点嵌入; 对于特征不一致问题, 从节点特征的相似度出发, 引入图注意力模块用于学习节点嵌入; 最后将两模块得到的节点嵌入进行融合并用于图异常检测任务。在 3 个真实数据集的实验结果表明, FHANN 在 AUC-ROC(area under the receiver operating characteristic curve) 和 AUC-PR(area under the precision-recall curve) 两个评估指标上平均提高了 1.17% 和 4.25%。本文研究成果可为社交网络分析、金融欺诈检测等复杂图结构数据的异常识别任务提供方法借鉴与技术参考。

关键词: 图数据; 神经网络; 异常检测; 异配连接; 特征不一致; 图滤波; 注意力机制; 频域信息融合
中图分类号: TP391 **文献标志码**: A **文章编号**: 1673−4785(2026)04−0919−13

中文引用格式: 王天昊, 窦浩, 王翔, 等. 融合图滤波与注意力机制的图异常检测方法 [J]. 智能系统学报, 2026, 21(4): 919–931.
英文引用格式: WANG TianHao, DOU Hao, WANG Xiang, et al. Graph anomaly detection method integrating graph filtering and an attention mechanism[J]. CAAI transactions on intelligent systems, 2026, 21(4): 919–931.

Graph anomaly detection method integrating graph filtering and an attention mechanism

WANG TianHao[1], DOU Hao[1], WANG Xiang[1,2], MAO Guojun[1,2]
(1. School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China; 2. Fujian Provincial Key Laboratory of Big Data Mining and Application, Fujian University of Technology, Fuzhou 350118, China)

Abstract: To address the pervasive issues of heterophilic connections and feature inconsistency in graph neural network–based anomaly detection, this paper proposes the flexible high–low-frequency and attention neural network (FHANN), a graph anomaly detection method that integrates graph filtering and attention mechanisms. Specifically, to mitigate the effects of heterophilic connections, a graph filtering module is designed to adaptively capture and fuse high- and low-frequency signals for node representation learning. To address feature inconsistency, a graph attention module is introduced to learn the importance of node features and generate node representations based on feature similarity. The representations produced by the two modules are then fused for anomaly detection. Experiments on three real-world graph anomaly detection datasets demonstrate that FHANN achieves average improvements of 1.17% and 4.25% in area under the receiver operating characteristic curve(AUC-ROC) and area under the precision-recall curve(AUC-PR), respectively. These results verify the effectiveness of the proposed method and provide valuable insights for anomaly detection in complex graph-structured data, including social network analysis and financial fraud detection.

Keywords: graph data; graph neural networks; anomaly detection; heterophilic connection; feature inconsistency; graph filtering; attention mechanism; frequency domain information fusion

随着互联网的快速发展, 在网络上存在各种各样的异常事件[1-3], 例如在社交网络伪装者扮成正常用户发布虚假评论[4] 或实施欺骗活动[5]; 在计算机网络中, 攻击者可以伪装成普通用户通过

收稿日期: 2025−10−27. **网络出版日期**: 2026−04−23.
基金项目: 国家重点研发计划项目(2019YFD0900900/05); 福建省教育厅中青年教师教育科研项目(JAT241070); 福建理工大学科技项目(GY-Z21183).
通信作者: 王翔. E-mail: wxsyhwl@fjut.edu.cn.

远程攻击工具或计算机漏洞对网络中的用户进行病毒植入或数据窃取，造成网络安全事故及信息泄露[6]；在金融网络中，同样存在各种恶意攻击活动和欺骗行为，给用户和机构造成金融损失。为了对网络的异常事件或对象进行检测，传统的深度学习方法将对象用特征矢量进行表示，然后在特征矢量空间上进行异常检测[7-8]。虽然这些方法在异常检测任务中取得较好的效果，但是却忽略了不同实体间存在的交互信息，限制了该类方法的异常检测效果。

在现实世界中实体间普遍存在各种各样的交互，可以为异常检测任务提供额外有价值的信息。这些网络用户可以看成节点，用户间的交互可以看成边，从而可以将实体及其交互关系用图的形式来表示[9]。例如在社交网络中，用户可以看成节点，用户发表的评论内容构成节点的特征，而用户间相互评论及交流等交互行为构成不同用户的边。通过所构建的图结构数据，可以发展基于图的异常检测方法并用于检测网络中存在的异常用户或行为。

近年来，随着图神经网络 (graph neural networks, GNNs) 的迅速发展，出现了许多基于图神经网络的异常检测方法，并取得了很好的效果[10-12]。例如有些工作使用无监督的方法进行图异常检测以克服标签数据少的问题。Ding 等[13] 提出了一种基于图自编码架构的方法用于图异常检测。AnomMAN (anomaly on multi-view attributed networks)[14] 通过将图数据分解为多个属性子图进行异常检测。李贺等[15] 则进一步提出了一个基于图卷积自编码器的多视图属性网络异常检测模型 (anomaly detection on multi-view attribute networks, AMEAN)，通过依据节点类别拆分多视图并引入简化的图卷积，有效提升了多类别属性网络中的异常检测精度。多视图数据光谱聚类融合[16] 是另一重要思路，早期研究利用构建孪生网络 (siamese networks) 与谱聚类相结合的框架处理多视图数据[17-18]，而陈容珊等[19] 的工作虽聚焦于谱聚类，但其利用注意力机制与图卷积网络引导节点表示学习的思想，也为无监督图异常检测中的多视图关系建模提供了重要参考。研究者还发展了许多半监督的图异常检测方法。例如，PC-GNN (pick and choose graph neural network)[20] 通过结合采样方法和图神经网络方法进行异常检测；AOGNN(short for AUC-oriented GNN)[21] 引入了一种基于最大化 ROC 曲线下面积 (area under ROC curve, AUC) 指标的异常检测方法。值得注意的是，跨域学习为半监督异常检测提供了新的思路。苏世玉等[22] 提出的 DCRN(dual correlation reduction network) 框架通过双重分类与重建机制，实现了源域知识向目标域的有效迁移，显著提升了在目标图中识别共享与非共享异常的能力。这些方法大多是直接在空间域对图数据的结构和特征进行处理，无法学习图数据在频域上的模式和特征。针对该问题，近年来人们开始从频域的角度设计图异常检测方法，例如 AMNet(adaptive multi-frequency graph neural network)[23]、BWGNN (beta wavelet graph neural network)[24] 等，这些方法在图异常检测任务上都取得了较好的效果。

虽然现有的基于图神经网络的方法在异常检测任务中取得了较好的效果，但是这些方法还未能充分考虑图异常检测中的异配连接和特征不一致问题。这些问题来源于图神经网络是通过聚合邻居节点的特征进行节点表示学习。这种邻居节点特征聚合的前提是图数据的邻居节点有相似的特征及类别[25]。而在图异常检测中由于异常节点与正常节点普遍存在连接，导致这一前提往往不能成立，这就带来了两个方面的不一致问题：1）异配连接，异常节点可以和正常节点存在连接。这种情况的出现是由于异常和正常节点存在交互，例如在社交网络中伪装用户对普通用户发表虚假评论等。同时，在一个网络中，通常异常节点只是占其中的一小部分。这种情况下，使用 GNNs 直接聚合邻居节点的特征会导致异常节点聚合大量正常节点特征，从而使模型更加难以区分正常节点和异常节点。从图信号的频域角度看，异常节点的信号是高频信号为主而正常节点的信号以低频信号为主[24]。由于 GNNs 的低通滤波特性，将会过滤大部分的高频信息即异常节点的信息，从而阻碍异常节点的检测。因此，从频域角度，在图数据的异常检测中需要同时提取图信号的高频和低频信息。2）特征不一致，在一个网络中可能会有一些节点属于同一类别（正常或异常）但特征有很大差异。例如在社交网络中一个用户对其他两个正常用户发表评论，则这些评论的内容（节点特征）可能有很大差别，因为两个用户可能属于不同角色。这种情况下聚合节点的特征信息，会导致学习到的节点语义特征难以区分，影响模型的性能。

现有的方法大多基于空间域，即直接处理图数据的特征信息或聚合图数据的邻居节点特征，因此无法有效考虑异配连接和特征不一致问题。虽然基于频域的方法可以有效处理异配连接问

题, 然而这些方法没有考虑特征不一致问题同时也无法有效学习局部依赖信息, 影响方法的异常检测效果。

为此, 本文提出一种结合图滤波与注意力机制的异常检测方法(flexible high-low frequency and attention neural network, FHANN), 此方法同时考虑图异常检测任务的异配连接和特征不一致问题, 以提升图异常检测效果。本研究主要贡献如下:

1) 设计了一种图滤波模块以及合适的损失函数, 可以自适应地学习图数据中高频和低频信息以缓解异配连接问题。

2) 引入了基于特征相似度的图注意力模块用于缓解特征不一致问题, 模块可以学习不同节点的重要性信息并进行邻居节点语义特征学习。

3) 构建了基于自适应图滤波与注意力机制的异常检测方法, 并在 3 个真实世界的数据集进行实验。结果表明, 与基准方法相比, 所提方法在 AUC-ROC(area under the receiver operating characteristic curve)和 AUC-PR(area under the precision-recall curve)指标上分别平均提高 1.17% 和 4.25%。

1　图神经网络及图异常检测相关工作

1.1　图神经网络

图神经网络可以有效学习图结构数据的结构和特征, 生成有效的节点表示。现有的方法主要可以分成基于空间域的方法和基于谱域的方法两类。

基于空间域的方法通过聚合邻居节点的特征来学习节点表示。例如 Kipf 等[26] 首次提出图卷积神经网络(graph convolutional network, GCN), 该方法通过聚合一阶邻居节点的信息来更新节点表示。GCN 在处理节点分类、链路预测和异常检测等问题中表现出色, 激发了越来越多的研究者在图神经网络领域进行深入探索。随后, 为了缓解 GCN 带来的过平滑问题, Velickovic 等[27] 提出了图注意力网络(graph attention networks, GAT), 通过引入注意力机制使模型能够自适应地学习邻居节点的权重信息。为了更好地处理大规模图数据, Hamilton 等[28] 提出对节点的邻居进行采样并聚合的方法(graph sample and aggregate, GraphSAGE)。

基于谱域的方法先将图信号变换到频域, 然后在频域中对图信号进行处理。例如 Bruna 等[29] 通过图拉普拉斯矩阵的本征分解定义图信号从空域到频域的变换以及图卷积运算。然而, 这种方法需要对拉普拉斯矩阵进行本征分解, 具有很高的计算复杂度。因此 Defferrard 等[30] 提出通过引入切比雪夫多项式来近似图滤波器, 从而简化图滤波运算。此外, Xu 等[31] 提出图小波卷积神经网络, 通过引入图小波变换和图小波算子实现在频域上的图卷积操作。

虽然一般的图神经网络方法可以直接用于图异常检测任务, 但是图神经网络的消息传递机制要求邻居节点的类别和特征趋于一致。因此, 使用一般的图神经网络进行图异常检测往往效果不佳。

1.2　图异常检测方法

近年来, 随着 GNNs 在图建模中的广泛应用, 研究者们提出许多基于 GNNs 的方法来解决图异常检测问题。例如, Li 等[32] 提出 GAS(GCN-based anti-spam) 模型, 该方法基于图卷积神经网络, 结合异构图和同构图来捕捉异常节点的局部语义和全局语义信息; Gao 等[33] 提出 GDN(graph decomposition network) 模型, 该模型旨在缓解数据集的结构分布偏移, 进而提升异常检测效果; Mesgaran 等[34] 提出的 GFCN(graph fairing convolutional networks for anomaly detection) 方法可以通过跳跃连接有效利用图结构和远距离节点特征进行图异常检测; Roy 等[35] 设计的 GAD-NR(graph anomaly detection via neighborhood reconstruction) 模型基于图结构重建的方式进行图异常检测。这些方法都是在空间域对图数据进行处理, 无法从频域角度学习异常节点信号的模式。近年来一些新颖的研究提出基于频域的异常检测方法, 例如 Chai 等[23] 利用伯恩斯坦多项式模拟多个频带滤波器从而能够自适应地学习图信号的多方面频域信息, 实验结果表明可以学习更有效的节点表示; Gao 等[36] 设计的 GHRN(graph heterophily resistant network) 模型旨在从频谱域的角度缓解异常图数据中的异质性连接问题, 从而更好地捕获图数据中的异常信号; Xu 等[37] 设计了 SEC-GFD (semi-supervised enhanced contrastive graph-based fraud detection) 方法, 通过使用多频带滤波和局部特征约束方法, 能够缓解图异常检测任务中的异质性和标签利用问题; Zheng 等[38] 提出 DSGAD (dynamic spectral graph anomaly detection) 方法, 通过引入可训练的动态小波机制与频带感知的特征融合策略, 提升了图异常检测中对不同频带异常模式的感知能力。基于频域的方法在异常检测中表现出优越的效果是因为它们能够捕捉图数据中不同频率成分的特性。

虽然现有的基于 GNNs 的方法在图异常检测上取得了较好效果, 但是现有的方法还未能充分

考虑图数据的异配连接和特征不一致的问题，限制了方法的异常检测效果。因此，本研究引入图滤波和注意力机制分别用于解决图数据的异配连接和特征不一致问题，以提升图异常检测效果。

2　符号与定义

定义 1　属性图 属性图表示为$\boldsymbol{G}=(V,\boldsymbol{X},\boldsymbol{A})$，$V=\{v_1,v_2,\cdots,v_N\}$表示节点的集合，$N$ 表示节点个数。$\boldsymbol{A}\in\mathbf{R}^{N\times N}$表示图 $\boldsymbol{G}$ 的邻接矩阵，当v_i和v_j两个节点之间存在边时，则$\boldsymbol{A}_{i,j}=1$，否则$\boldsymbol{A}_{i,j}=0$。$\boldsymbol{X}\in\mathbf{R}^{N\times d}$是节点的特征矩阵，$d$ 为节点特征的维度。

定义 2　图异常检测 给定一个属性图 $\boldsymbol{G}$，图上的每个节点都有一个标签$Y\in\{0,1\}$用于表示该节点是否是异常节点，其中标签 0 表示正常节点，标签 1 表示异常节点。其中异常节点的特征通常和正常节点的特征有很大区别，并可以通过特征区分出节点是否异常。在这里本研究关注于属性图上的半监督图异常检测，即给定一部分节点的标签，通过学习的检测器判断给定节点是否是异常节点。

定义 3　图滤波 图信号即图特征矩阵 $\boldsymbol{X}$ 的滤波操作是基于图傅里叶变换空间。定义图拉普拉斯矩阵$\boldsymbol{L}=\boldsymbol{I}-\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$，$\boldsymbol{L}$ 是实对称矩阵，可以进行特征分解为$\boldsymbol{L}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{U}$，这里$\boldsymbol{U}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_N]$和$\boldsymbol{\Lambda}=\mathrm{diag}([\lambda_1,\lambda_2,\cdots,\lambda_N])$分别是 $\boldsymbol{L}$ 的特征向量矩阵和特征值构成的对角矩阵，对应$\boldsymbol{u}_i(i=1,2,\cdots,N)$和$\lambda_i(i=1,2,\cdots,N)$分别是 $\boldsymbol{L}$ 的特征向量和特征值，特征值表示频率大小且从小到大排列，则图信号 $\boldsymbol{X}$ 的傅里叶变换和逆变换分别定义为$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}$和$\boldsymbol{U}\boldsymbol{X}$。给定一个图滤波器$h(\boldsymbol{\Lambda})$，$h(\boldsymbol{\Lambda})$是$\boldsymbol{\Lambda}$的函数，是一个对角矩阵，即$h(\boldsymbol{\Lambda})=\mathrm{diag}([h(\lambda_1),h(\lambda_2),\cdots,h(\lambda_N)])$，则图信号的滤波操作定义为

$$h(\boldsymbol{\Lambda})*\boldsymbol{X}=\boldsymbol{U}h(\boldsymbol{\Lambda})\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}$$

3　图异常检测方法

为了有效进行属性图的异常检测，针对异常检测数据集中存在的异配连接和特征不一致问题，本研究提出一种新颖的基于图滤波和注意力机制的图异常检测模型 FHANN，模型结构如图 1 所示。对于异配连接问题，考虑到异常和正常的节点信号在频域中分别主要表现为高频和低频信号，本研究从图信号滤波的角度出发通过所设计的图滤波器和损失函数可以让模型同时学习到图数据中每一个节点的高频信号和低频信号，然后通过点乘注意力机制将高频信号和低频信号融合起来作为节点表示。对于特征不一致问题，引入基于特征的注意力机制能够自适应学习不同节点特征的重要性，然后根据不同节点特征的重要性进行邻居节点聚合并学习节点表示，由于节点特征重要性可以自适应地捕获不同节点语义特征的差异，因此能够缓解特征不一致问题。

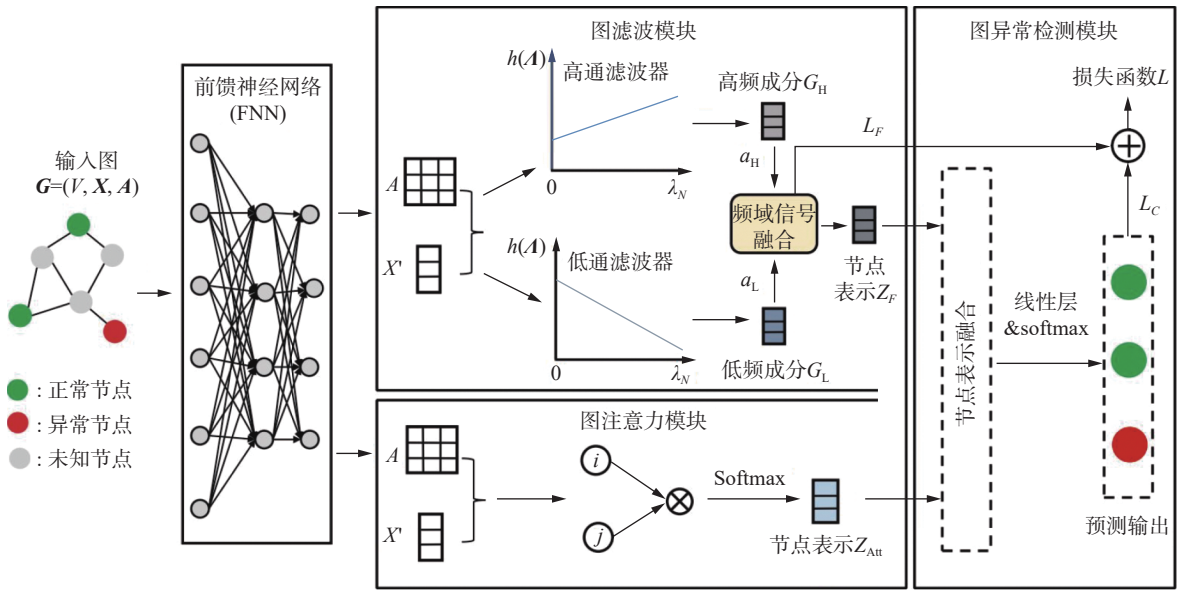


图 1　FHANN 模型框架

Fig. 1　The overall architecture of FHANN

FHANN 方法的伪代码算法如算法 1 所示。

算法 1　FHANN 方法的算法伪代码

输入　属性图$\boldsymbol{G}=(V,\boldsymbol{X},\boldsymbol{A})$；

输出　预测结果$\boldsymbol{P}\in\mathbf{R}^{N\times 2}$。

1）特征变换$\boldsymbol{X}'\leftarrow\boldsymbol{X}$；

2）对输入特征进行图滤波；

3）频域信号融合；

4）For i=1,2,⋯,N do

5）搜索邻居节点集合N_i。

6）For $j \in N_i$ do

7）计算注意力系数P_{ij}。

8）End

9）计算节点表示；

10）End

11）融合图滤波模块和图注意力模块的节点表示；

12）进行反向传播并训练异常检测模型；

13）预测结果$\boldsymbol{P} \in \mathbf{R}^{N\times 2}$

3.1　图滤波模块

在这里本研究首先设计两个图滤波器$h_i(\boldsymbol{\Lambda})$, $i \in \{H,L\}$分别用于捕获图信号的高频和低频信息, 在下文中将使用下标H和L分别表示高频和低频, 图数据上的每个节点可以通过图滤波器的可学习参数捕获到不同频率的信息。正如前面所讨论, 不同类型的节点具有不同的频率成分, 即异常节点的高频成分占大多数而正常节点低频成分占大多数。为了建模不同类型节点的频率差异, 本研究进一步使用节点层面的点乘注意力机制并通过该注意力机制构建出体现不同类型节点频率差异的损失函数, 通过构建的损失函数模型能够自适应地学习到不同类型节点的频域成分。最后通过学习到的注意力系数, 可以将不同频率成分的节点信号融合起来作为图滤波模块的节点表示。学习到的节点表示由于同时包含高频和低频信息, 同时通过损失函数自适应地学习到不同频率成分的信息, 模型可以通过不同频率成分的多少有效区分不同节点的类别。

为了更有效地进行输入图信号滤波操作, 使用前馈神经网络 (feedforward neural networks, FNN) 将节点特征矩阵$\boldsymbol{X} \in \mathbf{R}^{N\times d}$进行转化以提取输入信号更加有效和抽象的特征表示, FNN 中的非线性激活函数可以学习输入信号的非线性关系增强模型的非线性拟合能力, 特征转换过程为

$$\boldsymbol{X}' = \mathrm{FNN}(\boldsymbol{X})$$

式中: $\boldsymbol{X}' \in \mathbf{R}^{N\times d'}$为转化后的特征矩阵, d'为特征转化的输出维度, 这里 FNN 是包含两个全连接层和一个 RELU( ) 激活函数的神经网络。为了同时捕获节点信号的高频和低频信息, 使用设计的两个图滤波器$h_i(\Lambda), i \in \{H,L\}$对转化后的特征矩阵$\boldsymbol{X}'$进行滤波, 得到相应的图滤波信号。

$$G_i = h_i(\boldsymbol{\Lambda}) * \boldsymbol{X}' = \boldsymbol{U}h_i(\boldsymbol{\Lambda})\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}', i \in \{H,L\} \quad (1)$$

这里G_{H}和G_{L}表示经过两个滤波器滤波后的输出频域信号如前面的定义$h_i(\boldsymbol{\Lambda}) = \mathrm{diag}([h_i(\lambda_1), h_i(\lambda_2),\cdots,h_i(\lambda_N)])$是两个图滤波器同时也是拉普拉斯矩阵$\boldsymbol{L}$本征值序列的函数, 在模型中它们是可学习的参数, 从而使滤波器可以自适应地学习图信号的高低频成分。

由于式 (1) 的图滤波计算涉及到矩阵的特征分解会带来很大的计算量, 为了简化计算使用线性高通和低通滤波器分别提取图信号的高频和低频信息, 即

$$h_{\mathrm{H}}(\boldsymbol{\Lambda}) = (\boldsymbol{I} + \tilde{\boldsymbol{\Lambda}}) \text{和} h_{\mathrm{L}}(\boldsymbol{\Lambda}) = (\boldsymbol{I} - \tilde{\boldsymbol{\Lambda}})$$

在这里对本征值序列进行了归一化, 即

$$\tilde{\boldsymbol{\Lambda}} = \frac{2}{\lambda_N}\boldsymbol{\Lambda} - \boldsymbol{I}$$

式中: $\boldsymbol{I}$为单位矩阵, λ_N为拉普拉斯矩阵$\boldsymbol{L}$的最大本征值。由于图滤波器$\boldsymbol{I}+\tilde{\boldsymbol{\Lambda}}$和$\boldsymbol{I}-\tilde{\boldsymbol{\Lambda}}$的频率响应函数分别是$1+\lambda_i$和$1-\lambda_i$, 即高通和低通滤波器, 因此可以有效提取图信号的高频和低频信息, 这里λ_i表示频率。由于不同的图信号具有不同的频域响应分布, 因此需要设计一个能自适应学习频域分布的图滤波器。然而, 图滤波器$\boldsymbol{I}+\tilde{\boldsymbol{\Lambda}}$和$\boldsymbol{I}-\tilde{\boldsymbol{\Lambda}}$的不同频率幅度即频率响应函数是固定的, 无法自适应学习图信号不同频率成分。因此在图滤波中要对图滤波器进行参数化, 在这里引入可学习参数矩阵$\boldsymbol{W}_{\mathrm{H}}$和$\boldsymbol{W}_{\mathrm{L}}$对图滤波器进行参数化以用于学习不同的频率成分。将图滤波器代入式 (1) 并注意到$\boldsymbol{U}^{\mathrm{T}}\tilde{\boldsymbol{\Lambda}}\boldsymbol{U} = \frac{2}{\lambda_N}\boldsymbol{L} - \boldsymbol{I} = \tilde{\boldsymbol{L}}$, 高低频图滤波操作为

$$G_{\mathrm{H}} = h_{\mathrm{H}}(\boldsymbol{\Lambda}) * \boldsymbol{X}' = (\boldsymbol{I} + \tilde{\boldsymbol{L}})\boldsymbol{X}'\boldsymbol{W}_{\mathrm{H}}$$

$$G_{\mathrm{L}} = h_{\mathrm{L}}(\boldsymbol{\Lambda}) * \boldsymbol{X}' = (\boldsymbol{I} - \tilde{\boldsymbol{L}})\boldsymbol{X}'\boldsymbol{W}_{\mathrm{L}}$$

式中: $G_i \in \mathbf{R}^{N\times d}, i \in \{H,L\}$为图滤波的输出表示, $\boldsymbol{W}_i \in \mathbf{R}^{d\times d'}, i \in \{H,L\}$为可学习的参数矩阵用于学习高频和低频信息。为了增强模型的非线性拟合能力, 经过图滤波后应用 RELU() 激活函数。

通过图滤波器捕获到每个节点不同频率的特征后, 需要进行频域信号融合即融合高频和低频信号。考虑到不同类别的节点高频和低频成分不同, 引入点乘注意力计算不同频率信息的重要性得分, 然后通过重要性得分将高频和低频的信号融合起来作为节点表示, 重要性得分$\boldsymbol{e}_i \in \mathbf{R}^{N\times 1}$计算过程为

$$\boldsymbol{e}_i = \tanh[G_i\boldsymbol{W}^p + \boldsymbol{X}'^{\mathrm{T}}\boldsymbol{W}^q], i \in \{H,L\}$$

式中: $\boldsymbol{X}'$为 FNN 转换后的特征矩阵, T为转置操作, $\boldsymbol{W}^p \in \mathbf{R}^{d\times d'}$和$\boldsymbol{W}^q \in \mathbf{R}^{N\times 1}$都是可学习的参数矩阵, tanh() 为激活函数。为了使模型更易于训练且使重要性得分在 0 ~ 1, 将重要性得分使用 Softmax() 归一化得到高频和低频成分的注意力值$\boldsymbol{a}_i \in \mathbf{R}^{N\times 1}$。

$$\boldsymbol{a}_i = \mathrm{Softmax}(\boldsymbol{e}_i) = \frac{\exp(\boldsymbol{e}_i)}{\exp(\boldsymbol{e}_{\mathrm{H}}) + \exp(\boldsymbol{e}_{\mathrm{L}})}, i \in \{H, L\}$$

可以看出计算得到每一个节点 $j(j = 1,2,\cdots,N)$ 都有两个注意力值 $\boldsymbol{a}_{\mathrm{H}}^j$ 和 $\boldsymbol{a}_{\mathrm{L}}^j$，它们都是标量，分别用于表示该节点图信号的高频和低频成分的大小。例如，如果节点 j 是异常节点，其信号主要是高频成分则 $\boldsymbol{a}_{\mathrm{H}}^j$ 值更大，而如果节点 j 是正常节点，其信号主要是低频则 $\boldsymbol{a}_{\mathrm{L}}^j$ 值更大。计算得到注意力值后，将图信号的高频和低频信息融合起来，则第 j 个节点的输出为 $\boldsymbol{Z}_F^j \in \mathbf{R}^{d'}$。

$$\boldsymbol{Z}_F^j = \sum_i \boldsymbol{a}_i^j \boldsymbol{G}_i^j$$

式中：$\boldsymbol{G}_i^j \in \mathbf{R}^{d'}, i \in \{H, L\}$ 为第 j 个节点的信号经过图滤波器后的输出，计算所有节点的图信号后得到所有节点的输出 $\boldsymbol{Z}_F \in \mathbf{R}^{N \times d'}$。

使用图滤波器 $\boldsymbol{G}_{\mathrm{H}}$ 和 $\boldsymbol{G}_{\mathrm{L}}$ 可以通过监督的方式学习图信号的高低频信息。在这里为了使模型能够更好学习不同类型节点的频域信息，可以对注意力值进行约束并构建一个基于注意力值的损失函数作为优化目标函数。所设计的损失函数通过对注意力值的优化可以更好捕获到异常节点和正常节点的主要高频信息和低频信息。这项损失函数 L_F 的公式为

$$L_F = \sum_j \max(0, \boldsymbol{r}_j(\boldsymbol{a}_{\mathrm{L}}^j - \boldsymbol{a}_{\mathrm{H}}^j))$$

式中：j 为节点的序号；$\boldsymbol{r}_j$ 为固定的系数，即当节点是异常节点时，$\boldsymbol{r}_j = 1$；当节点是正常节点时，$\boldsymbol{r}_j = -1$。训练时训练集节点的标签已知，因此 $\boldsymbol{r}_j$ 是已知的。$\boldsymbol{a}_{\mathrm{L}}^j$ 和 $\boldsymbol{a}_{\mathrm{H}}^j$ 分别是低频信息和高频信息的注意力值。可以容易看到当某一节点 j 是异常节点时，通过损失函数的优化使 L_F 变小，会使 $\boldsymbol{a}_{\mathrm{L}}^j$ 减小而 $\boldsymbol{a}_{\mathrm{H}}^j$ 增大，从而学习到更多的高频信息；而当某一节点 j 是正常节点时，通过损失函数的优化会使 $\boldsymbol{a}_{\mathrm{H}}^j$ 减小而 $\boldsymbol{a}_{\mathrm{L}}^j$ 增大，从而学习到更多的低频信息。

3.2　图注意力模块

由于图滤波模块可以有效捕获图信号的不同频率成分，可以有效缓解异配连接的问题，但是图数据集还存在特征不一致问题。由于特征不一致问题是因为同一类别节点可能存在较大的特征差异，直接聚合这些节点的特征会使学习到的节点语义特征难以区分，影响模型的性能。从另一个角度看，在图神经网络聚合节点信息时这些节点特征的重要性不同，需要通过监督学习的方式自适应地学习不同节点特征的重要性信息，从而缓解特征不一致问题。

为了自适应地学习不同节点特征的重要性信息，设计了基于特征相似度的图注意力模块，用于学习节点的特征表示。由于特征相似度可以表示不同节点特征的差异程度，使用特征相似度可以有效获取节点特征的重要性信息。对于输入的图数据邻接矩阵 $\boldsymbol{A}$ 和经过 FNN 转化的特征矩阵 $\boldsymbol{X}' = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N\}$，如果节点 j 和节点 i 根据邻接矩阵 $\boldsymbol{A}$ 存在连接，则节点 j 到节点 i 的注意力系数 $\boldsymbol{S}_{ij}$ 的计算公式为

$$\boldsymbol{S}_{ij} = \beta \cdot \mathrm{Attn}(\boldsymbol{x}_i, \boldsymbol{x}_j), j \in N_i$$

式中：β 为一个标量，是每个传播层的可学习注意力参数；$\boldsymbol{x}_i \in \mathbf{R}^r$ 和 $\boldsymbol{x}_j \in \mathbf{R}^t$ 分别为节点 i 和节点 j 的特征向量。 Attn() 表示特征注意力函数定义为

$$\mathrm{Attn}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j / (\|\boldsymbol{x}_i\| \|\boldsymbol{x}_i\|)$$

式中：$\|\cdot\|$ 表示取二范数，N_i 为节点 i 和节点 i 的一阶邻居节点集合。为了使模型更易于训练同时将注意力系数限定在 0 ~ 1，使用 Softmax() 函数对注意力系数做归一化。

$$\boldsymbol{P}_{ij} = \mathrm{Softmax}(\boldsymbol{S}_{ij}) = \frac{\exp(\boldsymbol{S}_{ij})}{\sum_{u \in N_i} \exp(\boldsymbol{S}_{iu})}$$

式中：$\boldsymbol{P}_{ij}$ 为归一化的注意力系数，在计算中对于节点 i 只计算它的一阶邻居节点及自身的注意力系数，该注意力系数表示节点 i 和它的邻居节点的语义特征重要性程度，也表示语义特征的差异，可以用于进一步学习节点表示以缓解特征不一致问题。则对于第 i 个节点，根据归一化的注意力系数节点表示计算为

$$\boldsymbol{Z}_{\mathrm{Att}} = \sigma\left(\sum_{j \in N_i} \boldsymbol{P}_{ij} \boldsymbol{x}_j \boldsymbol{W}_{\mathrm{Att}}\right)$$

式中：$\sigma()$ 为激活函数；$\boldsymbol{W}_{\mathrm{Att}} \in \mathbf{R}^{d' \times d'}$ 是一个所有节点共享的参数矩阵，用于学习节点更抽象的语义特征。在计算所有节点的表示后可以得到图数据的节点特征矩阵 $\boldsymbol{Z}_{\mathrm{Att}} \in \mathbf{R}^{N \times d'}$。

3.3　图异常检测模块

通过图滤波模块和图注意力模块学习到节点表征 $\boldsymbol{Z}_F$ 和 $\boldsymbol{Z}_{\mathrm{Att}}$ 之后，为了能够同时缓解异配连接的问题和特征不一致问题，需要将学习到的两种特征表示融合起来，为此先将两种表示进行拼接然后使用可学习参数矩阵 $\boldsymbol{W}_Z \in \mathbf{R}^{2d' \times d'}$ 进一步学习不同输入表示节点层面的重要性信息，得到最终的节点表示 $\boldsymbol{Z} \in \mathbf{R}^{N \times d'}$。

$$\boldsymbol{Z} = \mathrm{concat}(\boldsymbol{Z}_F, \boldsymbol{Z}_{\mathrm{Att}}) \boldsymbol{W}_{\boldsymbol{Z}}$$

式中 concat() 为拼接操作，通过拼接并进一步学习后的节点表示 $\boldsymbol{Z}$ 融合了图滤波模块和图注意力模块的语义特征，能够缓解图异常检测中存在的

异配连接问题和特征不一致问题。最终的节点表示 $\boldsymbol{Z}$ 先进行一次线性变换后再通过 Softmax() 归一化函数处理得到最终预测结果 $\boldsymbol{P}\in\mathbf{R}^{N\times 2}$, 其中 $\boldsymbol{W}_p\in\mathbf{R}^{d'\times 2}$ 为可学习参数矩阵, $\boldsymbol{b}_P\in\mathbf{R}^{N\times 2}$ 为偏置。

$$\boldsymbol{P}=\text{Softmax}(\boldsymbol{Z}\boldsymbol{W}_p+\boldsymbol{b}_p)$$

在训练过程中, 使用交叉熵损失函数 L_c 和基于注意力值差异的损失函数 L_F 作为模型训练的损失函数, 同时计算过程中引入超参数 γ 用于控制损失函数 L_F 的比例, 计算公式为

$$L=L_c+\gamma L_F$$

4　实验及结果分析

4.1　实验数据集和基线方法

实验使用 Elliptic、Yelp、Weibo 和 inj_cora 共 4 个不同领域的图异常检测公共数据集用于验证此模型的有效性。其中, inj_cora 数据集是用于 4.11 节特征不一致对模型效果的影响分析。表 1 给出 4 个数据集的统计信息, 包括节点的个数、边的个数、特征的个数以及异常值在整个数据集中的占比。

表 1　数据集的统计信息

Table 1　Statistics of datasets

数据集	节点数	边数	特征数	异常值占比/%
Elliptic	46 564	73 248	93	9.76
Yelp	45 954	3 846 979	32	14.53
Weibo	8 405	407 963	400	10.33
inj_cora	2 708	11 060	1 433	5.10

1) Elliptic[23]: 该数据集是一个从比特币交易网络中获取的部分真实交易数据。其中, 节点表示比特币的交易, 分为合法和非法两类。合法类别包括交易所、矿工、金融服务提供商等; 非法类别包括欺诈、恶意软件、洗钱、庞氏骗局等。边代表比特币的交易流。

2) Yelp[23]: 该数据集包含对美国几个州的酒店和餐馆的虚假评论和真实评论, 其中节点代表用户的评论。如果评论是由同一用户发布的, 则他们之间存在一条边; 同一个月内同一产品下的两条评论存在一条边; 同一产品下的评论与用户评论时给产品打的星级产生关系则存在一条边。

3) Weibo[35]: 该数据集来自微博平台, 是多用户发布帖子的社交网络。其中用户作为节点, 若用户在另一用户发布的贴子下发布评论则产生一条边。若用户在特定的时间内发布评论则视为可疑事件, 制造至少 5 次可疑事件, 则视为该用户为异常用户, 其余用户则被归类为正常用户。

4) inj_cora: 该数据集是引文网络, 来源于 PyGOD 仓库 (https://github.com/pygod-team/data)。该数据集一共有 2 708 个节点分成 7 个类别, 同时数据集包含一定比例的合成异常节点, 适合用于特征不一致的影响分析。

为了验证 FHANN 模型的有效性, 本研究将该模型与 3 类不同类型的模型进行了对比分析, 分别是一般的图神经网络模型(GCN[26]、GAT[27]、GraphSAGE[28])基于空域的图异常检测方法(GAS[32]、GDN[33]、GFCN[34])和基于频域的图异常检测方法(AMNet[23]、BWGNN[24]、GHRN[36]、SEC-GFD[37]、DSGAD[38])。

4.2　评价指标

AUC-ROC: 该指标是一种常用于二分类问题的性能评估指标, 通过计算不同分类阈值下的真正例率(true positive rate, TPR)和假正例率(false positive rate, FPR), 绘制出 ROC 曲线。其中, 横轴表示假正例率, 纵轴表示真正例率。最终利用曲线下的面积来衡量模型的分类性能, 取值范围在 0 ~ 1, AUC-ROC 为 1 表示完全识别, 为 0.5 则表明该模型性能等同于随机猜测。

AUC-PR: 该指标是基于精确率 (precision) 和召回率 (recall)曲线下的面积来评估模型的性能。精确率是指被分类器判定为正例的样本中实际为正例的比例, 召回率是指实际为正例的样本中被分类器判定为正例的比例。AUC-PR 取值范围在 0 ~ 1, 值越接近 1 代表性能越好。

4.3　实验设置

实验的对比模型和 FHANN 模型均在一个配置有 Tesla A40-48G GPU 和 Intel Xeon Gold 6 326 CPU 的服务器上运行。为了保证公平性, 所有实验均在相同的实验条件下进行。对于 GCN[26]、GAT[27] 和 GraphSAGE[28] 模型的实现, 使用 PyTorch Geometric 库提供的函数。为了优化模型性能, 针对每个数据集调整学习率和 dropout 超参数。对于其他的图异常检测模型, 使用原始作者提供的源代码和超参数进行实验。每个数据集独立运行 10 次实验并记录平均结果和标准差。

为公平比较, 本文保持由 Chai 等[23] 对 Yelp 和 Elliptic 数据集的划分比例, 即 Yelp 和 Elliptic 数据集的训练集、验证集和测试集的比例分别为 7∶1∶2 和 4.5∶3.5∶2。inj_cora 使用 PyGOD 提供的划分比例即训练集、验证集和测试集分别为 140、500 和 1 000 个节点。对于 Weibo 数据集, 训练集、验证集和测试集的比例随机划分为 6∶3∶1。

在实验中，本文对学习率和超参数进行微调以获得最优结果。模型的隐藏层维度对 inj_cora 数据集设为 128，其他数据集为 64。Yelp 的学习率设为5×10^{-3}，权重衰减参数设为5×10^{-4}，γ设为 0.5；Elliptic 数据集的学习率设为5×10^{-5}，权重衰减参数设为1×10^{-5}，γ设为 0.5；Weibo 和 inj_cora 数据集的学习率设为5×10^{-4}，权重衰减参数设为1×10^{-4}，γ设为 0.1。所有数据集的训练 Epoch 为 400，以确保收敛。

4.4　性能比较

实验结果如表 2 ~ 3 所示。FHANN 在 3 个真实世界数据集上的异常检测性能均取得了最优表现。在 Yelp、Elliptic 和 Weibo 数据集上，FHANN 的 AUC-ROC 相较最佳方法分别提升 2.15%、0.40% 和 0.97%，AUC-PR 提升 6.06%、2.32% 和 4.36%。从表 2 ~ 3 中可看出，3 种一般图神经网络的效果较差，而其他的异常检测模型效果更好。这表明异常检测任务与简单的节点分类问题有很大的不同，这是因为在图异常检测中需要重点考虑异配连接和特征不一致问题，而一般节点分类不需要考虑这些问题。从表 2 ~ 3 中可以看出，基于频域的方法 DSGAD、GHRN、BWGNN、AMNet 和 SEC-GFD 效果比 FHANN 差。这些方法通过捕获图信号的不同频带特征能够在一定程度上缓解异配连接问题，但这些方法没有考虑特征不一致问题，因此效果更差。

表 2　FHANN 与不同基线模型的 AUC-ROC 对比
Table 2　Comparison of different baseline methods on AUC-ROC scores　%

方法	Yelp	Elliptic	Weibo
GCN	60.05±0.76	83.63±2.23	92.34±1.60
GAT	80.88±1.48	87.29±1.00	92.19±2.52
GraphSAGE	52.03±1.97	85.41±1.34	94.92±2.16
GAS	77.90±0.84	85.48±1.83	94.81±0.87
GDN	79.60±0.92	85.83±0.36	91.61±0.54
GFCN	77.74±0.11	85.84±1.42	95.58±1.05
BWGNN	81.24±0.53	88.37±0.57	93.63±0.62
AMNet	84.72±0.58	87.65±0.78	96.38±1.61
GHRN	84.34±0.45	88.32±0.68	95.63±2.16
SEC-GFD	84.55±0.58	86.66±0.73	94.80±1.16
DSGAD	85.35±0.67	87.52±0.88	95.86±1.00
FHANN	**87.50±0.32**	**88.77±1.54**	**97.35±1.28**

注：加粗表示最佳，下划线表示次佳。

表 3　FHANN 与不同基线模型的 AUC-PR 对比
Table 3　Comparison of different baseline methods on AUC-PR scores　%

方法	Yelp	Elliptic	Weibo
GCN	23.57±0.70	41.42±5.16	87.22±2.39
GAT	45.69±2.92	50.06±11.13	87.46±2.78
GraphSAGE	15.56±0.77	50.02±5.02	85.46±4.19
GAS	36.17±1.44	47.46±8.87	89.97±1.93
GDN	45.11±2.67	66.19±5.54	84.97±1.26
GFCN	41.26±0.23	45.53±8.45	91.54±1.26
BWGNN	48.37±1.01	45.21±5.50	90.04±0.85
AMNet	51.88±1.33	69.50±4.79	90.41±3.00
GHRN	47.95±1.16	46.49±10.35	91.23±2.85
SEC-GFD	56.53±1.39	66.59±3.71	88.95±2.10
DSGAD	58.51±1.29	53.39±10.10	92.75±1.12
FHANN	**64.57±0.73**	**71.82±4.64**	**95.90±1.35**

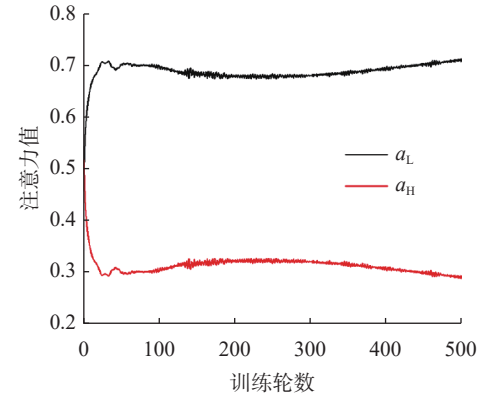
注：加粗表示最佳，下划线表示次佳。

另外，从表 2 中可以看出基于空域的方法 GAS、GDN 和 GFCN 的效果普遍比基于频域的方法要差。这是因为基于空域的方法直接对图数据的结构和特征进行处理，无法学习图信号不同频率的信息，对异配连接问题和特征不一致问题的处理效果不佳。因此，为有效解决异配连接与特征不一致问题，FHANN 模型通过图滤波模块捕获图数据的高频信息和低频信息有效缓解图异常检测的异配连接问题；同时，引入图注意力模块，针对性缓解特征不一致问题。最终，使模型能够有效提升图异常检测效果。

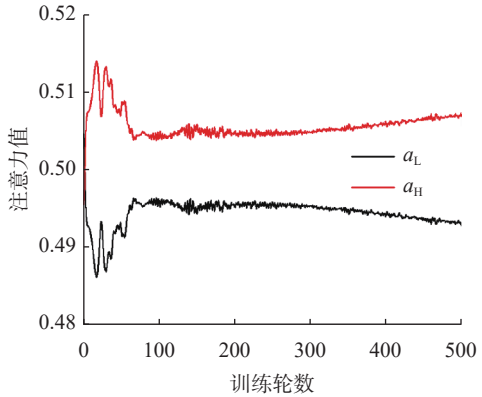
4.5　图滤波器注意力值变化趋势

为了验证所提出的图滤波模块在捕获图信号的高频和低频信息的有效性，在这里画出正常节点和异常节点的两个图滤波器的注意力值随训练 Epoch 的变化曲线。其中注意力值的计算是将所有正常节点和异常节点的注意力值（$\boldsymbol{a}_{\mathrm{H}}$和$\boldsymbol{a}_{\mathrm{L}}$）相加取平均得到。图 2 给出了 Elliptic 数据集上正常节点和异常节点的注意力值随训练 Epoch 的变化曲线。从图 2(a) 中可以看出正常节点的高频滤波器注意力值$\boldsymbol{a}_{\mathrm{H}}$随着训练轮数的增加不断下降直至收敛，而低频率波器的注意力值$\boldsymbol{a}_{\mathrm{L}}$随着训练轮数的增加不断上升直至收敛，从注意力值的变化可以看出模型可以更多地学习到正常节点的低频信息，这与正常节点信号的频域特征符合。从图 2(b) 可以看到异常节点的高频率波器注意力

值$\boldsymbol{a}_{\mathrm{H}}$随着训练轮数的增加先急剧下降, 然后慢慢上升直至收敛, 而低频率波器注意力值$\boldsymbol{a}_{\mathrm{L}}$显示出相反的趋势, 从注意力值的变化可以看出模型可以学习到更多的高频信息, 这与异常节点信号的频域特征相符。综上, 从图滤波器注意力值变化趋势可以看出, 提出的图滤波模块可以有效捕获异常节点和正常节点的高频信息和低频信息, 提升模型的异常检测效果。



(a) 正常节点



(b) 异常节点

图 2　图滤波器注意力值随训练轮数变化趋势

Fig. 2　The trends of attention value of graph filters

4.6　消融实验分析

为了验证所提出的图滤波模块和图注意力模块的有效性, 本节对这两个模块进行消融实验。对 FHANN 模型移除其中一个模块并保持其他部分不变得到两个变体。

1) FHANN-A: 它消除了 FHANN 模型中的图滤波模块, 保留其他部分包括 FNN 和图注意力模块进行异常检测。

2) FHANN-G: 它消除了 FHANN 模型中的图注意力模块, 保留其他部分包括 FNN 和图滤波模块进行异常检测。

表 4 和表 5 分别给出了消融实验的结果。从表 4 和表 5 中可以看出 FHANN 的效果明显超过了其他两个变体 FHANN-A 和 FHANN-G。与 ANFAN-A 相比, FHANN 模型在 AUC-ROC 和 AUC-PR 上平均值分别提高了 5.15% 和 13.01%, 验证了提出的图滤波模块的有效性。与 FHANN-G 相比, FHANN 模型在 AUC-ROC 和 AUC-PR 上平均值分别提高了 2.30% 和 5.22%, 验证了提出的注意力机制在图异常检测中的有效性。本文提出的方法通过融合图滤波模块和图注意力模块的信息有效缓解了图异常检测中存在异配连接问题和特征不一致问题, 有效提升图异常检测的效果。

表 4　FHANN 模型及其变体在 AUC-ROC 上的对比

Table 4　Comparison of FHANN model and its variants on AUC-ROC scores　%

模型	Yelp	Elliptic	Weibo
FHANN-A	81.38±0.61	82.27±1.11	94.52±1.40
FHANN-G	85.55±0.29	84.96±0.45	96.21±1.27
FHANN	**87.50±0.32**	**88.77±1.54**	**97.35±1.28**

注: 加粗表示最佳。

表 5　FHANN 模型及其变体在 AUC-PR 上的对比

Table 5　Comparison of FHANN model and its variants on AUC-PR scores　%

模型	Yelp	Elliptic	Weibo
FHANN-A	50.00±1.14	55.72±7.14	87.53±5.39
FHANN-G	58.15±0.54	64.55±1.53	93.94±1.56
FHANN	**64.57±0.73**	**71.82±4.64**	**95.90±1.35**

注: 加粗表示最佳。

4.7　训练损失函数分析

为了评估所提出模型的训练稳定性和收敛性, 将 FHANN 模型的训练损失趋势与 AMNet、GFCN、SGC-GFD 和 DSGAD 共 4 种基线方法进行对比, 并在 Yelp 和 Elliptic 数据集上进行了实验。图 3 给出了不同模型在 Yelp 和 Elliptic 数据集上的训练曲线。其中图 3(a) 给出了 Yelp 数据集上损失函数的变化趋势, 而图 3(b) 则描绘了 Elliptic 数据集上的损失变化情况。从图 3 中可以观察到, FHANN 模型的训练曲线保持相对稳定, 表明该模型具有较好的训练稳定性。此外, 该模型的收敛速度较快, 损失函数能够迅速达到最小值, 这说明 FHANN 在与其他方法相比时, 具备更高的收敛效率和训练效率。这主要归因于 FHANN 包含图滤波模块和基于注意力机制的节点表示模块, 能够有效缓解异配连接和特征不一致问题, 从而有效提升异常检测的性能。

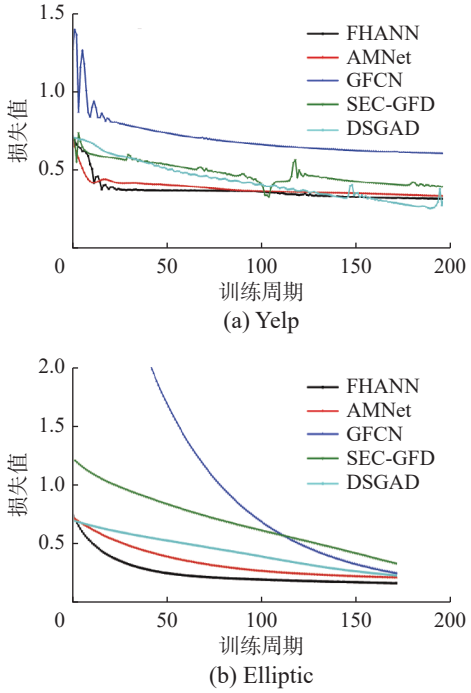


图 3　在 Yelp 和 Elliptic 数据集上的训练损失曲线

Fig. 3　The trends of loss value on the Yelp、Elliptic datasets

4.8　超参数γ分析

损失函数中涉及的超参数γ用于平衡L_C和L_F损失的贡献。为了分析损失函数L_F在学习图信号不同频率分量的影响，本文进行了参数敏感性实验。使用 Yelp、Elliptic 和 Weibo 数据集进行研究，并记录了不同γ值下 AUC-PR 得分的变化趋势。图 4 是实验结果，可以看出参数γ对性能有较大影响。对于 3 个数据集，随着参数γ的增加，AUC-PR 值最初呈上升趋势，随后开始下降。

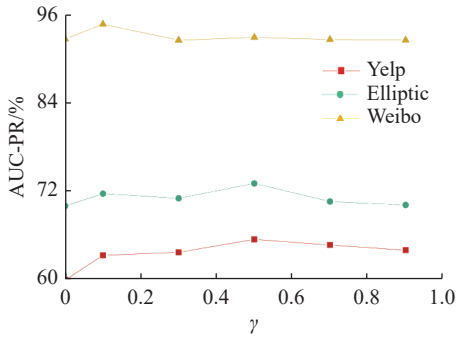


图 4　在 Yelp、Elliptic 和 Weibo 数据集上 AUC-PR 得分随超参数γ变化的曲线

Fig. 4　The trends of AUC-PR score with the hyperparameter γ on the Yelp、Elliptic and Weibo datasets

具体而言，当γ值为 0 时，所有数据集的 AUC-PR 均处于较低水平，说明若不引入L_F损失，模型难以充分捕获异常节点的高频特征与正常节点的低频信息，验证了L_F在缓解异配连接问题中的必要性。随着γ逐渐增大，AUC-PR 呈现上升趋势，表明适中的L_F能有效增强模型对异常模式的感知能力。当γ超过某一阈值后，会过度降低分类损失的重要性，导致 AUC-PR 下降。最优γ值因数据集而异，在 Yelp 和 Elliptic 数据集上值为 0.5 时，性能最佳；而在 Weibo 数据集上，当γ值为 0.1 时，AUC-PR 值达到最大。不同数据集在不同的γ上性能最佳这一差异可能与数据集的图结构特性、异常比例以及特征分布有关。因此，在模型优化过程中，需要仔细调整γ值，以获得更好的性能。

4.9　不同注意力机制分析

模型引入注意力方法能够自适应地学习不同节点特征的重要性，从而缓解特征不一致问题。注意力机制有多种变体，为研究不同注意力方法对异常检测的影响，选择图注意力网络[27]和多头自注意力机制[39]两种经典的注意力机制进行研究，并与 FHANN 进行比较。将 FHANN 中的注意力机制模块替换为这两种注意力方法，得到了两个模型变体：1）FHANN-N，该模型在注意力机制模块中使用图注意力网络；2）FHANN-M，该模型在注意力机制模块中使用多头自注意力机制。除了注意力机制模块外，模型的其他部分和参数设置保持不变。这些模型变体在 3 个数据集上独立运行 10 次，并记录 10 次实验的平均 AUC-ROC 和 AUC-PR 结果。实验结果如图 5 所示。

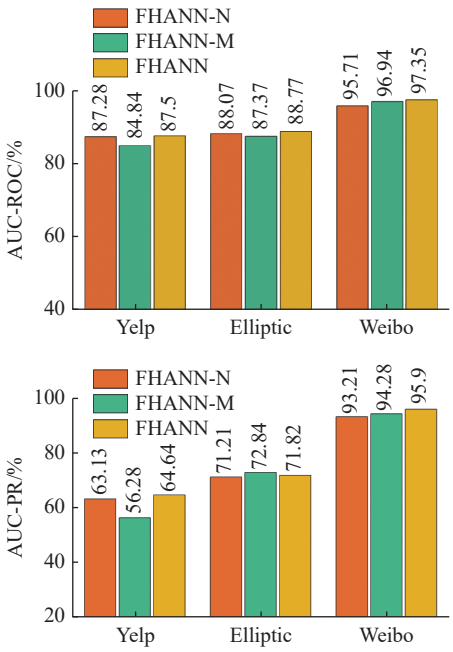


图 5　在 Yelp、Elliptic 和 Weibo 数据集上的不同的注意力机制的对比

Fig. 5　The comparison of different attention mechanisms on the Yelp、Elliptic and Weibo datasets

结果表明，通过使用基于特征相似性的注意力机制，FHANN 模型达到了最佳性能。由于特

征相似性能够更有效地捕捉不同节点之间的特征差异, 因此具有更好性能。而图注意力网络和多头自注意力机制虽然都通过特征进行学习, 但在注意力计算中未能充分捕捉特征相似性信息。因此, 在 FHANN 模型中, 采用了基于特征相似性的注意力机制, 可以更好地缓解特征不一致的问题。

4.10　可视化分析

为了更直观地展示提出的图异常检测模型的效果, 使用 T-SNE 降维技术将模型学习到的节点表示投影到二维并在二维坐标平面上进行可视化。以 Yelp 数据集为例, 可视化如图 6 所示。其中图 6(a) 是未经过模型训练的原始节点特征分布, 图 6(b)、图 6(c)、图 6(d) 分别是经变体 FHANN-A、变体 FHANN-G 和 FHANN 训练后生成的节点表示分布。在二维平面图上, 正常节点和异常节点分别用红色和绿色表示。从图 6 中可以看出, 未训练的节点表示难以区分异常节点和正常节点, 两种类别的节点表示在平面上都呈现出均匀分布。而经过 FHANN 模型训练后的节点表示则能够进行很好地区分, 异常节点的表示明显聚集在一起。从 T-SNE 可视化可以看出所设计的图异常检测模型能够缓解异配连接和特征不一致问题, 得到更有效的节点表示。

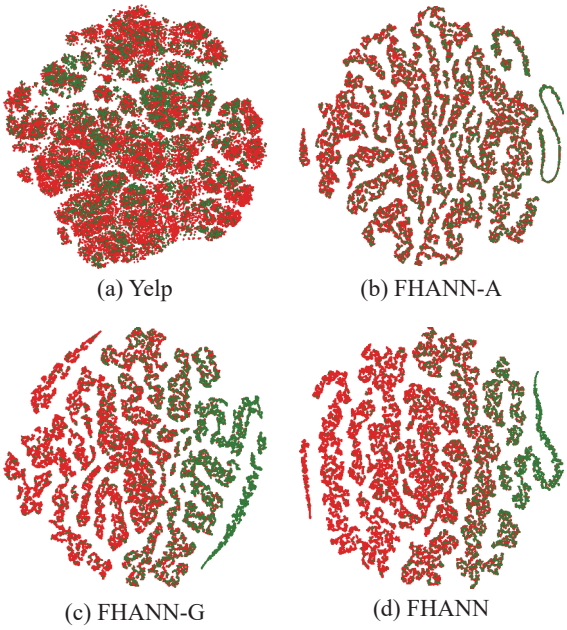


图 6　T-SNE 降维下的正常和异常节点表示的分布

Fig. 6　The representation distribution of normal and anomalous node using T-SNE dimension reduction technique

4.11　节点特征不一致的影响分析

节点的特征不一致对异常检测效果有较大影响, 这是因为图滤波模块捕获图信号的高低频。例如, 在引文网络中, 正常或异常节点本身具有不同类别导致特征差异。此类特征差异会对异常检测的准确性造成干扰。为验证 FHANN 方法在缓解由节点类别差异引起的特征偏差方面的有效性, 本研究开展了相关实验评估。

由于上述实验所用数据集未包含节点类别信息, 所以本研究采用 PyGOD 库提供的公开合成数据集 inj_cora 进行验证。表 1 给出了该数据集的统计信息, 其划分比例遵循标准设置, 训练集、验证集和测试集分别包含 140、500 和 1 000 个样本。该数据集共包含 7 种节点类别, 为了人为放大不同类别节点间的特征差异, 实验首先计算节点特征的均值 u, 随后对类别 1 ~ 4 的节点特征分别叠加不同倍数的 u 值, 从而生成具有不同特征差异程度的数据集。为进行对比分析, 本研究还对比分析了 AMNet、GHRN 和 BWGNN 模型的效果, 并对其参数进行调优。所有实验均独立重复 10 次, 最终结果以均值±标准差的形式记录于表 6 中。

表 6　不同方法在具有不同程度特征差异的数据集上的实验结果（AUC-ROC）

Table 6　Experimental results (AUC-ROC) of different methods on dataset with different degrees of feature disparity

模型	+0	+1u	+2u	+3u
AMNet	68.22±3.48	67.16±3.98	65.18±4.01	64.94±3.58
BWGNN	67.05±3.18	65.40±3.63	64.01±3.58	63.56±3.84
GHRN	71.36±2.11	71.16±2.71	70.67±2.97	70.41±2.50
FHANN	**73.87±3.49**	**72.92±4.05**	**72.25±4.43**	**71.48±3.77**

注: 加粗表示最佳。

从表 6 中可以观察到, 随着不同类别节点间特征差异的增大, 所有模型的性能均出现不同程度的下降。其中, GHRN 与 FHANN 的降幅相对较小, 而 AMNet 和 BWGNN 的下降幅度更为显著, 分别从原来的 68.22% 降至 64.94%、67.05% 降至 63.56%。这可能是因为 GHRN 通过剪枝类间边, 能够聚焦于异常节点本身的模式特征, 从而降低了对其他特征差异的敏感性。对于 FHANN 模型而言, 其将频域信号与注意力机制深度融合能够学习复杂的依赖关系, 从而提升了模型的鲁棒性。

5　结束语

本文研究图异常检测任务中存在的异配连接性问题和特征不一致性问题。为了解决这两个关键问题, 提出了一种新颖的基于图滤波和注意力

机制的异常检测框架。对于异配连接问题，考虑到异常和正常的类别在频域中分别主要表现为高频和低频信号，本研究从图信号滤波的角度出发，通过所设计的图滤波器和合适的损失函数可以让模型同时学习到图数据中每一个节点的高频信号和低频信号，然后通过点乘注意力机制将高频信号和低频信号融合起来作为节点表示。对于特征不一致问题，引入基于特征相似度的注意力机制能够自适应学习不同节点特征的重要性，然后根据不同节点特征的重要性进行邻居节点聚合并学习节点表示。本文提出的模型在多个真实数据集上的实验结果表明相较于现有模型取得了更好的异常效果，验证了提出方法的有效性。在未来工作中，所提出的方法可以用于探索其他更加复杂的图数据异常检测任务，如异构图、动态图等。

参考文献：

[1] 李忠, 靳小龙, 庄传志, 等. 面向图的异常检测研究综述[J]. 软件学报, 2021, 32(1): 167−193.
LI Zhong, JIN Xiaolong, ZHUANG Chuanzhi, et al. A survey on graph-based anomaly detection research[J]. Journal of software, 2021, 32(1): 167−193.

[2] 陈波冯, 李靖东, 卢兴见, 等. 基于深度学习的图异常检测技术综述[J]. 计算机研究与发展, 2021, 58(7): 1436−1455.
CHEN Bofeng, LI Jingdong, LU Xingjian, et al. A survey on graph anomaly detection techniques based on deep learning[J]. Journal of computer research and development, 2021, 58(7): 1436−1455.

[3] CUI Jiafu, YANG Siqi, YI Litai, et al. Recent advances in deep learning for protein-protein interaction: a review[J]. BioData mining, 2025, 18(1): 43.

[4] KUMAR J S, ARCHANA B, KUMAR K M V S. Graph theory: modelling and analyzing complex system[J]. Metallurgical and materials engineering, 2025, 31(3): 70−77.

[5] MOTIE S, RAAHEMI B. Financial fraud detection using graph neural networks: a systematic review[J]. Expert systems with applications, 2024, 240: 122156.

[6] XUAN TUNG N, TUNG GIANG L, DUC SON B, et al. Graph neural networks for next-generation-IoT: recent advances and open challenges[J]. IEEE communications surveys & tutorials, 2026, 28: 2226−2262.

[7] WANG Zhen, LAN Chao. Towards a hierarchical Bayesian model of multi-view anomaly detection[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. Yokohama: International Joint Conferences on Artificial Intelligence Organization, 2020.

[8] PANG Guansong, SHEN Chunhua, CAO Longbing, et al. Deep learning for anomaly detection: a review[J]. ACM computing surveys, 2022, 54(2): 1−38.

[9] WANG Daixin, LIN Jianbin, CUI Peng, et al. A semi-supervised graph attentive network for financial fraud detection[C]//2019 IEEE International Conference on Data Mining. Beijing: IEEE, 2019.

[10] 郭嘉琰, 李荣华, 张岩, 等. 基于图神经网络的动态网络异常检测算法[J]. 软件学报, 2020, 31(3): 748−762.
GUO Jiayan, LI Ronghua, ZHANG Yan, et al. Dynamic network anomaly detection algorithm based on graph neural networks[J]. Journal of software, 2020, 31(3): 748−762.

[11] XU Yiming, PENG Zhen, SHI Bin, et al. Revisiting graph contrastive learning on anomaly detection: a structural imbalance perspective[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(12): 12972−12980.

[12] LIU Ziqi, CHEN Chaochao, YANG Xinxing, et al. Heterogeneous graph neural networks for malicious account detection[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. Torino: ACM, 2018.

[13] DING Kaize, LI Jundong, BHANUSHALI R, et al. Deep anomaly detection on attributed networks[C]//Proceedings of the 2019 SIAM International Conference on Data Mining. Philadelphia: SIAM, 2019.

[14] CHEN Linghao, LI He, ZHANG Wanyuan, et al. AnomMAN: detect anomalies on multi-view attributed networks[J]. Information sciences, 2023, 628: 1−21.

[15] 李贺, 彭以冲, 张万园, 等. 基于图卷积自编码器的多视图属性网络异常检测算法[J]. 中国科学: 信息科学, 2025, 55(2): 269−283.
LI He, PENG Yichong, ZHANG Wanyuan, et al. Detect anomalies on multi-view attributed networks based on graph convolution autoencoder[J]. Scientia sinica informations, 2025, 55(2): 269−283.

[16] 张铭泉, 周辉, 曹锦纲. 基于注意力机制的双 BERT 有向情感文本分类研究[J]. 智能系统学报, 2022, 17(6): 1220−1227.
ZHANG Mingquan, ZHOU Hui, CAO Jingang. Dual BERT directed sentiment text classification based on attention mechanism[J]. CAAI transactions on intelligent systems, 2022, 17(6): 1220−1227.

[17] 于润羽, 李雅文, 李昂. 融合领域特征的科技学术会议语义相似性计算方法[J]. 智能系统学报, 2022, 17(4): 737−743.
YU Runyu, LI Yawen, LI Ang. Semantic similarity computing for scientific and technological conferences[J]. CAAI transactions on intelligent systems, 2022, 17(4): 737−743.

[18] 马甜甜, 杨长春, 严鑫杰, 等. 融合知识图谱和轻量级图卷积网络推荐系统的研究[J]. 智能系统学报, 2022, 17(4): 721−727.
MA Tiantian, YANG Changchun, YAN Xinjie, et al. Research on the fusion of knowledge graph and lightweight graph convolutional network recommendation system[J]. CAAI transactions on intelligent systems, 2022, 17(4): 721−727.

[19] 陈容珊, 高淑萍, 齐小刚. 注意力机制和图卷积神经网

络引导的谱聚类方法[J]. 智能系统学报, 2023, 18(5): 936−944.
CHEN Rongshan, GAO Shuping, QI Xiaogang. A spectral clustering based on GCNs and attention mechanism[J]. CAAI transactions on intelligent systems, 2023, 18(5): 936−944.

[20] LIU Yang, AO Xiang, QIN Zidi, et al. Pick and choose: a GNN-based imbalanced learning approach for fraud detection[C]//Proceedings of the Web Conference 2021. Ljubljana: ACM, 2021.

[21] HUANG Mengda, LIU Yang, AO Xiang, et al. AUC-oriented graph neural network for fraud detection[C]//Proceedings of the ACM Web Conference 2022. Online: ACM, 2022.

[22] 苏世玉, 于卿, 李婧, 等. 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): 375−385.
SU Shiyu, YU Jiong, LI Shu, et al. Cross-domain graph anomaly detection via dual classification and reconstruction[J]. Computer science, 2025, 52(8): 375−385.

[23] CHAI Ziwei, YOU Siqi, YANG Yang, et al. Can abnormality be detected by graph neural networks?[C]//Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. Vienna: International Joint Conferences on Artificial Intelligence Organization, 2022.

[24] TANG Jianheng, LI Jiajin, GAO Ziqi, et al. Rethinking graph neural networks for anomaly detection[C]//Proceedings of the International Conference on Machine Learning. Hyderabad: IMLS, 2022.

[25] AKOGLU L, TONG Hanghang, KOUTRA D. Graph based anomaly detection and description: a survey[J]. Data mining and knowledge discovery, 2015, 29(3): 626−688.

[26] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[C]//Proceedings of the International Conference on Learning Representations. San Juan Capistrano: ICLR, 2016.

[27] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks[C]//Proceedings of the International Conference on Learning Representations. Vancouver: ICLR, 2018.

[28] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs[C]//Proceedings of the Advances in Neural Information Processing Systems. Long Beach: NIPS Foundation, 2017.

[29] BRUNA J, ZREMBA W, SZLAM A, et al. Spectral networks and locally connected networks on graphs[C]// Proceedings of the International Conference on Learning Representations. Aspen: ICLR, 2014.

[30] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering[C]//Proceedings of the Advances in Neural Information Processing Systems. Barcelona: NeurIPS Foundation, 2016.

[31] XU Bingbing, SHEN Huawei, CAO Qi, et al. Graph wavelet neural network[C]//Proceedings of the International Conference on Learning Representations. New Orleans: ICLR, 2019.

[32] LI Ao, QIN Zhou, LIU Runshi, et al. Spam review detection with graph convolutional networks[C]//Proceedings of the 28th ACM International Conference on Information and Knowledge Management. Beijing: ACM, 2019.

[33] GAO Yuan, WANG Xiang, HE Xiangnan, et al. Alleviating structural distribution shift in graph anomaly detection[C]//Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. Singapore: ACM, 2023.

[34] MESGARAN M, BEN HAMZA A. Graph fairing convolutional networks for anomaly detection[J]. Pattern recognition, 2024, 145: 109960.

[35] ROY A, SHU Juan, LI Jia, et al. GAD-NR: graph anomaly detection via neighborhood reconstruction[C]//Proceedings of the 17th ACM International Conference on Web Search and Data Mining. Merida: ACM, 2024.

[36] GAO Yuan, WANG Xiang, HE Xiangnan, et al. Addressing heterophily in graph anomaly detection: a perspective of graph spectrum[C]//Proceedings of the ACM Web Conference 2023. Austin: ACM, 2023.

[37] XU Fan, WANG Nan, WU Hao, et al. Revisiting graph-based fraud detection in sight of heterophily and spectrum[C]//Proceedings of the AAAI conference on artificial intelligence. Vancouver: AAAI, 2024.

[38] ZHENG Jianbo, YANG Chao, ZHANG Tairui, et al. Dynamic spectral graph anomaly detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Philadelphia: AAAI, 2025.

[39] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. Long Beach: NeurIPS, 2017.

作者简介:

王天昊, 硕士研究生, 主要研究方向为人工智能和图神经网络。E-mail: 2241308079@smail.fjut.edu.cn。

王翔, 副教授, 博士, 主要研究方向为人工智能和图神经网络。CCF 专业会员。 E-mail: wxsyhwl@fjut.edu.cn。

毛国君, 教授, 博士, 主要研究方向为人工智能、数据挖掘、大数据和分布式计算。中国人工智能学会专委会常委、国家科学技术奖评审委员、计算机学会生信息物学专委会委员。发表学术论文 100 余篇。E-mail: 19662092@fjut.edu.cn。