



DOI:10.12404/j.issn.1671-1815.2403486

引用格式: 吴安昊, 卜凡亮, 李志远. 基于图神经网络的有向网络节点影响力计算方法[J]. 科学与技术工程, 2025, 25(12): 5066-5072.

Wu Anhao, Bu Fanliang, Li Zhiyuan. A graph neural network-based method for calculating node influence in directed networks[J]. Science Technology and Engineering, 2025, 25(12): 5066-5072.

基于图神经网络的有向网络节点影响力计算方法

吴安昊, 卜凡亮*, 李志远

(中国人民公安大学信息网络安全学院, 北京 100038)

摘要 使用有向图建模社交网络、利用深度学习计算方法计算结算图中节点的影响力是社交研究领域的重要分支。为解决图神经网络不能很好地依据构建的节点特征计算节点影响力的问题,受 GraphGPS 启发,设计一个热度模块,提出一种融合真实网络中语义、行为、热度信息的节点影响力计算方法。首先依据节点多种中心性和正态分布采样得到的自信息作为节点的初始语义特征;其次,通过图神经网络融合节点特征;再次,通过热度模块学习节点热度信息;最后,融合提取的语义、行为、热度特征,实现节点影响力计算。在 4 个真实网络数据集上进行实验,结果表明添加热度模块的模型能有效的计算节点影响力。

关键词 节点影响力; 图神经网络; 有向网络; 社交网络

中图分类号 TP391.3 TP183; 文献标志码 A

A Graph Neural Network-based Method for Calculating Node Influence in Directed Networks

WU An-hao, BU Fan-liang*, LI Zhi-yuan

(School of Information Network Security, People's Public Security University of China, Beijing 100038, China)

[Abstract] Modeling social networks using directed graphs and calculating the influence of nodes in settlement graphs using deep learning methods are important branches in the field of social research. In order to solve the problem that graph neural networks cannot well calculate the influence of nodes based on constructed node features, inspired by GraphGPS, a heat module was designed to propose a method of calculating the influence of nodes that integrated semantic, behavioral, and heat information in real networks. Firstly, the self-information obtained based on nodes' multiple centrality and orthogonal distribution sampling was used as the initial semantic features of the nodes. Secondly, the node features were fused by graph neural network. Once again, the node heat information was learned by the heat module. Finally, the fusion of the extracted semantic, behavioral, and heat features was implemented to calculate the node influence. Experiments were conducted on four real network datasets. The results show that the model with the addition of the heat module can effectively calculate node influence.

[Keywords] nodal influence; graph neural networks; directed networks; social networks

近些年来,随着对图神经网络分析能力的认识不断深入,其作为一种有效的代表各种领域复杂系统的方法,可以用来解决包括社会科学^[1]、自然科学^[2]和许多其他研究领域^[3-4]在内的相关系统性问题。图具有包含丰富潜在信息的复杂结构^[5]。使用有向图建模社交网络并利用深度学习计算方法计算节点影响力是社交研究领域的重要分支。计算有向网络中节点影响力对识别预测社交群组中的关键节点有广阔的应用前景^[6]。

不同网络结构中节点影响力的评估方式不同^[7],目前适用有向网络的节点影响力评估方式可以分为:基于拓扑的中心性、基于主题敏感的中心

性、控制理论和图采样四大类^[8]。这种分类方式适用的数据集不同,而且对关键节点的定义也有所区别。例如,基于拓扑的中心性主要关注网络的拓扑结构,将能高度影响网络功能作为计算节点影响力的指标^[9];基于主题敏感的中心性更为关注用户发布内容的主题,根据用户发布内容的主题来计算用户在网络中的影响力;控制中心性计算节点影响力则使用控制理论,利用动态影响模型(dynamic influence model),通过用户引导整体意见方向的能力计算节点影响力。图采样方法计算节点影响力更注重节点重建整个网络的能力。

研究表明,有向网络存在多中心交织重叠的复

收稿日期: 2024-05-11; 修订日期: 2025-01-22

基金项目: 中国人民公安大学安全防范工程双一流专项(2023SYL08)

第一作者: 吴安昊(2001—),男,汉族,河南信阳人,硕士研究生。研究方向:安全防范工程。E-mail:2766084722@qq.com。

*通信作者: 卜凡亮(1965—),男,汉族,江苏徐州人,博士,教授。研究方向:安全防范工程。E-mail:buanliang@sina.com。

杂情况,计算节点影响力需要考虑多类型中心性构建节点特征^[8]。同时图神经网络(graph neural network, GNNs)最近变得非常流行,并在社交网络^[10],点云分析^[11]和推荐系统^[12]等各种图应用中展示了很好的结果,GNNs 遵循消息传递机制,该机制迭代地在节点之间进行信息交互^[13],能够挖掘节点的潜在表达。对于节点影响力计算任务,现存的图神经网络方法均是通过将节点的中心性作为节点特征,通过 GCNs 计算节点影响力,并且各种方法只是在构建节点语义特征进行改进,在保持模型输入相同的基础上,依据 GraphGPS^[14]的思想改进图神经网络的网络结构,提出基于热度、语义、行为的 Has-GNN(heat-action-semantics based GNN)图神经网络模型,并提出一种网络热度提取模块。所提方法从节点的语义、行为和热度 3 个维度提取信息来计算节点影响力。其中,语义信息包括节点网络特征;行为特征是网络的拓扑结构;热度信息则提取的节点热度特征。设计一种融合热度、语义、行为特征的节点影响力计算模型,在消息传递网络外添加热度提取模块,提取网络节点热度特征,提高节点影响力计算能力。提出一种有向网络热度提取模块,提高模型的学习能力。在 4 个真实数据集验证了 Has-GNN 模型和热度模块的有效性。为节点影响力预测、关键节点识别等研究提供一种新的方法。

1 相关工作

基于机器学习计算有向网络中节点的影响力大致可以分为两步,首先是使用特征工程得到节点的语义特征,然后使用机器学习算法训练和计算网络中节点的标签^[15]。在基于图神经网络的节点影响力方法中,主要使用传统的中心性方法构建节点初始特征,然后使用图神经网络来学习预测节点标签。适用有向网络的传统中心性主要是基于网络拓扑结构。相关工作将简单介绍构建节点特征的中心性和基于图神经网络的节点影响力计算方法。

1.1 基于拓扑结构的中心性

图论和网络分析中的中心性是评价网络中节点重要性的一个基本概念。在社交网络中,中心性度量有助于识别与网络其他节点有广泛联系的节点。度中心性(degree centrality, DC)^[16]是指一个节点在网络中的连接数。在现实世界中,度最高的节点是与网络中的许多其他用户直接交互的用户。这种方法对影响力的定义十分直观,但是没有考虑图的全局结构。紧密中心性(closeness centrality, CC),利用路径长度考虑间接连接,将局部中心性扩展为全局中心性。紧密中心性的基本思想是,紧密

中心性高的节点可以快速地将信息传播给其他节点,根据紧密中心性的基本思想,一个点在网络中的位置比它拥有的链接数量更重要,通过测量网络的传播效应^[17],将具有高紧密中心性的用户看作是网络中的关键节点,然而,紧密中心性对较大距离的节点对或节点对缺失链接非常敏感。另一种拓扑中心性是介数中心性(betweenness centrality, BC)^[16]是一个节点位于网络中其他两个节点之间的最短路径上的次数。紧密中心性和介数中心性都很难应用于大规模网络,在一些横断面和时间网络中是不稳定的。PageRank 中心性^[18]为了减轻部分节点对中心性分数的影响,减少了从这些节点出发的链接的权重。在 PageRank 中,传入链接的权重与其源节点的 PageRank 分数成正比。控制中心性(control centrality)^[19-20]对社会影响力的定义如下:给定两个个体 u, v , 其中 u 对 v 具有影响力,表示为 u 以直接或间接的方式改变 v 的观点^[21]。个体在社交网络中的影响受到个体行为的自动力、个体之间的耦合动力和社会网络的网络结构的影响。控制中心性同时考虑社交网络结构和每个节点的动态,来量化单个节点控制有向加权网络的能力。

1.2 基于图神经网络的计算方法

Wang 等^[22]设计了一种基于(graph convolutional network, GCN^[23])的影响深度学习方法 IDL(influence deep learning),在评估社交网络用户的传播影响时考虑了节点拓扑属性和行为日志。IDL 根据每个节点的社会行为日志作为深度学习模型的输入对每个节点的固定大小的子网络进行采样,并使用预训练的网络嵌入方法来获得每个节点的低维表示。由于 IDL 需要用户的社交行为信息,它只适用于社交网络,难以扩展到其他网络,如交通网络和生物网络。Zhao 等^[24]将传播影响节点识别问题作为分类问题,提出了 InfGCN 算法。InfGCN 使用 BFS(breadth first search)算法对每个节点的邻居网络进行采样,使用节点的 DC、BC、CC 和聚类系数用于构建模型的输入。然后,使用 GCN 的输出用作全连接神经网络的输入来预测每个节点的标签。Yu 等^[25]提出了一种基于 GCN 的高效算法 RCNN,该算法在为每个节点生成输入时只使用度值。虽然 RCNN 算法是有效的,但节点的传播影响可以同时受到其微观层面、社区层面和宏观层面的结构属性的影响,Ou 等^[26]在 RCNN 基础上进一步提出了 M-RCNN,从网络的 3 个层面构建节点特征,但是这种节点特征构建方法使得 M-RCNN 不适合有向网络。在有向图上计算节点影响力可以分为不考虑网络拓扑结构和考虑网络拓扑结构两类方法。其

中前者主要将节点特征作为输入,不考虑节点在网络中的拓扑结构,如MLP(multi-layer perceptron),这种方法由于缺少网络的拓扑信息,无法充分拟合网络数据,但是网络结构简单不需要多次训练就能达到较好的效果。后者主要就是图神经网络,如基于光谱的图神经网络ChebNet^[27](chebyshev network)、GCN^[23],基于空间的图神经网络GraphSAGE^[28](graph sample and aggregate)、GAT^[29](graph attention network)、HopGNN^[30],以及适用有向网络的深度学习方法(digraph inception convolutional networks, DiGCN)^[31]等,这些方法考虑了网络的拓扑结构,但是这些方法使得节点的信息只能在网络中传播扩散,使得节点之间差距越来越小,容易忽略节点原本的特征。

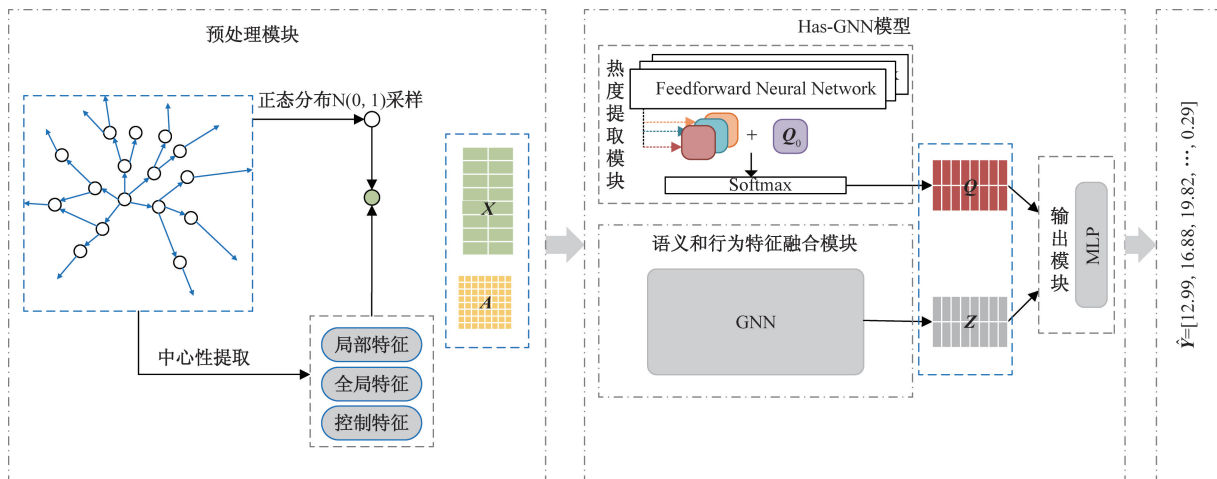
2 方法

Has-GNN模型的主要结构如图1所示,主要包

含热度特征提取模块和图神经网络。整个模型框架包括:①预处理模块,提取节点中心性特征,并采样获得节点自身的信息,初步融合语义和行为特征;②语义和行为特征融合模块,通过图神经网络融合预处理得到的节点信息和图拓扑结果;③热度特征提取模块,独立于图神经网络提取网络中节点热度特征;④输出模块,融合热度、语义、行为特征,输出计算结果。

2.1 预处理模块

特征提取模块如图2所示,主要提取节点的行为特征。从图2中抽取局部特征、全局特征、和控制特征形成行为特征向量。根据单一中心性不能针对不同特点的有向图自适应的计算节点影响力^[8],将节点的拓扑中心和控制论中的控制中心性(control centrality)作为节点特征提取出来,并分别提取网络拓扑的局部中心和全局中心,扩充节点特征。



Feedforward Neural Network 为前馈神经网络;GNN 为图神经网络;MLP 为多层感知器; X 为节点特征矩阵; A 为网络邻接矩阵; Softmax 为激活函数; Q_0 为可学习的节点初始热量; Z 为主干网络学习到的节点的潜在表示; \hat{Y} 为模型预测结果

图1 Has-GNN 结构

Fig. 1 Structure of Has-GNN

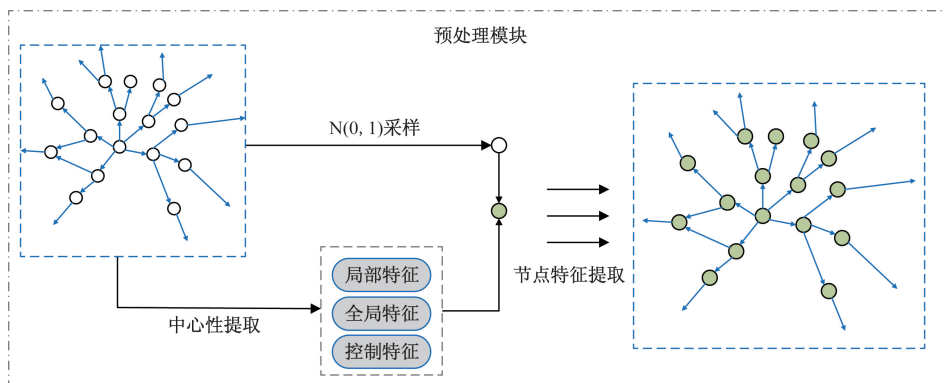


图2 节点特征提取

Fig. 2 Node feature extraction

(1)局部属性。有向图中节点的度中心性(degree centrality)是指一个节点在网络中的连接数。有向图中节点的度中心性的计算公式为

$$k_i = \sum_{j=1}^N a_{ij} + \sum_{j=1}^N a_{ji} - 1 \quad (1)$$

式(1)中: k_i 为节点*i*的度值; a_{ij} 和 a_{ji} 分别为节点*i*同节点*j*的出度和入度; N 为有向图的节点数量。

节点的度中心性计算简单、可解释性强。

(2)全局属性。计算有向图中节点的紧密中心性,计算公式为

$$C(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)} \quad (2)$$

式(2)中: $d(v,u)$ 为*v*和*u*之间的最短路径距离; n 为可以到达节点*u*的节点数。

计算有向图中节点位于网络中其他两个节点之间的介数中心性,计算公式为

$$C(v) = \sum_{s,t \in V} \left[\frac{\sigma(s,t|v)}{\sigma(s,t)} \right] \quad (3)$$

式(3)中: V 为节点集合; $\sigma(s,t)$ 为(*s,t*)的最短路径的数量; $\sigma(s,t|v)$ 为节点*s*到节点*t*的最短路径中经过点*v*的数量。

Pagerank通过添加缩放因子,使其惩罚与从高质量节点链接的低质量节点和奖励节点相关联的节点,减轻了具有许多传出链接的节点的影响,更加关注传入链接的质量,而不是数量。

控制属性来源于控制理论,将网络表示为线性动态系统控制中心性(control centrality)的计算公式为

$$\mathbf{C}^{(i)} \equiv [\mathbf{b}^i, \mathbf{A}\mathbf{b}^i, \mathbf{A}^2\mathbf{b}^i, \dots, \mathbf{A}^{N-1}\mathbf{b}^i] \quad (4)$$

$$C(i) = \text{rank}[\mathbf{C}^{(i)}] \quad (5)$$

式中: $\mathbf{C}^{(i)}$ 为点*i*的控制矩阵; $\mathbf{b}^i \equiv [0, 0, \dots, b_i, \dots]^T$,且 $b_i \neq 0$; $\mathbf{A} \in \mathbf{R}^{N \times N}$ 为有向图的邻接矩阵; N 为节点数量; $C(i)$ 为控制矩阵 $\mathbf{C}^{(i)}$ 的阶数。

针对每个网络的节点,从数据生成和分布的角度,使得整个网络节点的数据服从同一分布,并依据中心极限定理,通过在高斯分布中采样获得节点的自信息。

节点*i*的初始语义特征为

$$\mathbf{x}_i = [d_c(i), c_c(i), b_{cc}(i), c_{lc}(i), p_r(i), n(i)] \quad (6)$$

式(6)中: $d_c(i)$ 、 $c_c(i)$ 、 $b_{cc}(i)$ 、 $c_{lc}(i)$ 分别为依据式(1)~式(5)计算得到的节点*i*的度中心性、紧密中心性、介数中心性、控制中心性; $p_r(i)$ 为节点*i*的pagerank值; $n(i)$ 为节点*i*在 $N(0,1)$ 采样得到的自信息。

2.2 语义和行为特征融合模块

语义和行为特征融合模块则是1.3节中介绍的计算方法使用的图神经网络,将构建的节点初始特征和有向网络的行为特征 \mathbf{A} 输入图神经网络,可表示为

$$\mathbf{Z} = \text{GNN}(\mathbf{X}, \mathbf{A}) \quad (7)$$

式(7)中: \mathbf{X} 为节点的特征矩阵; \mathbf{A} 为网络的邻接矩阵;GNN为图神经网络。

选取有向图神经网络DiGCN^[31]作为语义和行为特征融合模块。

2.3 热度特征提取模块

依据GraphGPS^[14]的思想,图神经网络可以利用位置编码和结构编码作为软偏差来提高模型的表达性,同时构建的节点语义特征包含网络的结构编码,节点提出一个热度模块学习节点的热度特征。网络热度特征的定义为

$$\begin{cases} \mathbf{Q} = \text{softmax}(L_1 \odot L_2 \odot L_3) + \mathbf{Q}_0 \\ L_1 = \tanh(\mathbf{X}\mathbf{W}_c + \mathbf{b}_c) \\ L_2 = \tanh(\mathbf{X}\mathbf{W}_m + \mathbf{b}_m) \\ L_3 = \mathbf{A}\mathbf{W}_l + \mathbf{b}_l \end{cases} \quad (8)$$

式(8)中: \mathbf{Q}_0 为一个可学习的节点初始热量; \odot 为元素乘法运算符; \mathbf{X} 为节点的特征矩阵; \mathbf{A} 为网络的邻接矩阵; \mathbf{Q} 、 L_1 、 L_2 、 L_3 分别为学习到的中间变量; \mathbf{W}_c 、 \mathbf{W}_m 、 \mathbf{W}_l 、 \mathbf{b}_c 、 \mathbf{b}_m 、 \mathbf{b}_l 分别为学习网络的权重和偏置;采用tanh的目的是防止经过学习后得到的热度表示更偏向节点特征,减小节点本身对热度模块的影响,最后将维度扩大至语义模块的输出维度。

热度提取模块如图3所示。

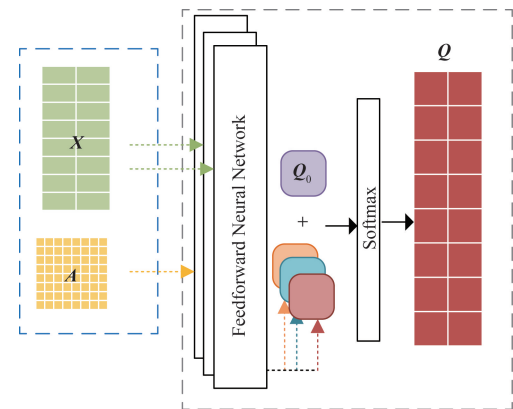


图3 热度特征提取

Fig. 3 Heat feature extraction

2.4 输出模块

在热度信息提取之后,如图4所示,融合热度、语义、行为特征计算函数*g*来生成最终输出。

$$\mathbf{H} = \alpha\mathbf{Z} + \beta\mathbf{Q} \quad (9)$$

$$\hat{\mathbf{Y}} = \text{readout}(\mathbf{H}) \quad (10)$$

式中: α 和 β 为可学习的权重参数,保证能够充分获

得热度、语义、行为三维度特征; H 为网络学习的中间变量; Z 为式(7)的输出; readout 表示操作是一个 MLP。

本文模型的训练目标是最小化 ground truth 值 Y 和预测 \hat{Y} 之间的均方误差 (mean squared error, MSE)。

$$\mathcal{L}_{ce} = \text{MSE}(Y, \hat{Y}) \quad (11)$$

式(11)中: \mathcal{L}_{ce} 为模型优化的损失。

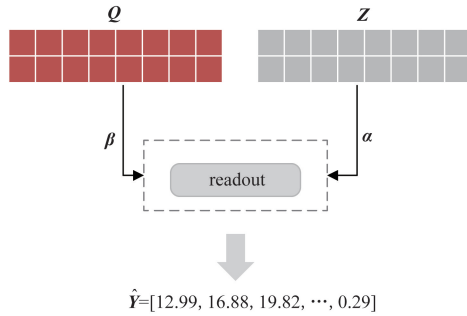


图4 信息融合

Fig. 4 Information fusion

3 实验评估

用于实验的数据集按照训练集占整个数据集的60%、测试集占整个数据集的40%的比例进行划分。训练次数最高为900轮。

3.1 实验数据

选择4个真实的网络来证明所提出的方法的有效性和准确性。4种网络的详细说明如下:①Health有向网络是根据一项调查创建的,每个学生都被要求列出他最好的5位女性朋友和5位男性朋友,一个节点代表一个学生,每个学生与其他学生的交互越多,在网络上的权重越高;②Highschool有向网络包含伊利诺伊州一所小型高中男生之间的友谊,一个节点代表一个男生,在网络中节点得分越高则越容易同其他节点建立联系;③Oz有向网络包含澳大利亚国立大学校园内一个宿舍楼217名住户之间的友谊评级。一个节点代表一个人;④Bit OTC是来自比特币OTC平台的用户之间信任/不信任的网络,用户得分越高说明在平台上越被他人信任。表1中列出了4个网络的统计特征,包括节点数 n 、边数 m 、平均度 $\langle k \rangle$ 、最大度 k_{\max} 。

表1 4个网络的基本信息

Table 1 Basic information of the four networks

网络	n	m	$\langle k \rangle$	k_{\max}
Health	2 539	12 969	10.215	36
Highschool	70	366	10.457	23
Oz	217	2 672	24.627	80
BitOTC	5 881	35 592	12.104	1 298

3.2 评价标准

使用回归模型常用的平均绝对误差 (mean absolute error, MAE)、均方误差 (mean squared error, MSE)、中位绝对误差 (median absolute error, MedAE)、决定系数 R^2 作为评价标准, MAE、MSE、MedAE 的值越低表明方法性能越好, R^2 的值越高表明方法性能越好。

3.3 对比方法

使用如下模型作为基线: MLP; 基于光谱的图神经网络 ChebNet^[27]、GCN^[23]; 基于空间的图神经网络 GraphSAGE^[28]、GAT^[29]; HopGNN^[30]; 以及适用有向网络的深度学习方法 DiGCN^[31]。实验中的所有模型超参数均保持相同设置。本文方法主要提出了热度模块, 提高了图神经网络的表达能力, 并在消融实验验证了本文模块的有效性。

3.4 基线对比

表2列出了在4个真实数据集上的实验结果。可以看出, 添加了热度模块的模型在所有任务中都比原有网络表现良好。

如表2所示, 在Health数据集上, 添加本文模块的神经网络比原本的神经网络在MSE有所降低, 在 R^2 得分上有所增加, 虽然在MAE和MedAE上有所增加, 但是从 R^2 得分上可以看出添加热度模块后, 模型的对数据的解释程度更优秀, 拟合的更好。同时发现, 基于空间的图神经网络相比基于光谱的图神经网络对数据的拟合程度更差, 在4个指标上均比基于光谱的图神经网络效果更差。在表2中, 虽然MLP在各个指标上均优于GCN、SAGE、GAT、HopGNN, 但是从 R^2 指标上可以看出, 这些模型均不同程度的反映了对节点影响力变动的解释程度, 可能原因是这些图神经网络需要更长的训练次数。

如表2所示, 在Highschool数据集上, 可以发现本文方法表现良好, 在绝大多数情况下都是表现最好的网络, 添加所模块后DiGCN的表现 R^2 指标超过了MLP, 能够更好地提高模型对数据的拟合程度。可以看出, 基于空间的图神经网络在4个指标上均比基于光谱的图神经网络效果更差。从表2中可以看出, MLP在相同的循环训练次数上表现优于GCN、SAGE、GAT、HopGNN。

由Oz数据集(表2)可知, 本文方法在 R^2 指标上相比于次优方法DiGCN提高了0.086。在中位绝对误差MedAE上相比于DiGCN有所提高, 但是在平均绝对误差MAE、均方误差MSE上均低于DiGCN, 说明本文方法能够更好地拟合数据。由MLP和GCN的 R^2 指标可知, GCN需要更多的训练次数才能, 达到MLP的水平。

表 2 4 个数据集上基线对比
Table 2 Baseline comparison on four datasets

数据集	方法	MAE	MSE	MedAE	R^2	数据集	方法	MAE	MSE	MedAE	R^2
Health	MLP	4.564	39.649	3.476	0.587	Oz	MLP	11.867	320.263	7.341	0.485
	ChebNet	4.714	40.128	3.833	0.618		ChebNet	12.011	311.341	8.437	0.488
	GCN	6.926	71.448	6.311	0.440		GCN	16.968	485.364	14.606	0.148
	SAGE	8.068	94.014	7.353	0.238		SAGE	18.114	539.945	15.711	0.056
	GAT	8.253	99.528	7.857	0.213		GAT	18.062	537.957	15.706	0.061
	HopGNN	7.853	97.753	7.179	0.040		HopGNN	18.768	578.225	16.210	0.001
	DiGCN	2.886	16.274	2.144	0.830		DiGCN	8.711	252.377	3.465	0.638
	本文方法	2.889	15.964	2.181	0.833		本文方法	8.161	188.511	4.853	0.712
Highschool	MLP	2.297	8.598	2.125	0.782	Bitcoin OTC	MLP	9.191	1 005.267	1.263	0.039
	ChebNet	2.106	6.988	1.994	0.703		ChebNet	11.287	1 055.355	4.805	-0.015
	GCN	1.942	5.517	1.422	0.771		GCN	12.011	1 052.533	5.814	0.007
	SAGE	2.885	12.431	3.307	0.610		SAGE	13.626	1 092.059	7.430	0.009
	GAT	3.327	15.850	2.854	0.492		GAT	14.149	1 134.690	8.374	-0.038
	HopGNN	3.434	20.221	2.231	0.199		HopGNN	9.938	940.998	3.039	0.088
	DiGCN	1.955	5.406	1.784	0.765		DiGCN	9.262	1 053.117	1.766	-0.009
	本文方法	1.839	4.918	1.755	0.785		本文方法	9.061	1 045.657	1.818	0.001

表 3 热度提取模块的消融研究

Table 3 Ablation study of hot extraction module

方法	Health	Highschool	Oz	Bitcoin OTC
DiGCN	0.830	0.765	0.638	-0.009
HopGNN	0.040	0.199	0.001	0.088
DiGCN + Heat	0.833	0.785	0.712	0.001
HopGNN + Heat	0.083	0.354	0.001	0.093

注:Heat 表示模型添加了热度模块。

从表 2 中可以看出,由于数据集的节点数量过于庞大,在训练最多 900 次的情况下各个模型均不能很好地拟合数据,甚至在 ChebNet、GAT、DiGCN 方法上出现了 R^2 指标为负的情况,说明这 3 个指标还没有开始拟合数据,在测试集上预测值和真实值的残差大于真实值的方差,预测值不能很好地解释正式数据的变动情况。其中表现较为优异的 HopGNN 在处理数据时首先对数据进行多跳处理,使得模型能够更早的获得多阶邻居的信息,即节点的行为信息。而添加了热度处理模块后,DiGCN 模型能够获得语义、行为、热度 2 个维度的特征,相比于只能获得语义、行为两个维度特征的原模型能够更早的开始拟合数据。

3.5 消融实验

验证热度提取模块对模型的影响,评估热度提取模块的有效性。消融实验的 R^2 指标结果如表 3 所示。

在 4 个数据集上进行消融实验,评估预处理图神经网络 HopGNN 和专为有向图设计的 DiGCN 两种类型的模型,从实验结果中可以看出,引入热度模块后,模型在 4 个数据集上有更好的表现。DiGCN 在 4 个数据集上 R^2 指标均有所提高;HopGNN 在 Health、Highschool、Bitcoin OTC 3 个数据

集上 R^2 指标分别提高了 0.043、0.155、0.005。这表明 Heat 模块在网络中学习到了节点的热度信息,从而表现得更加鲁棒,有助于验证所提出的热度嵌入模块在节点影响力计算中的实际效果。

4 结论

针对多种类型中心相互作用影响下有向网络中节点影响力计算不准问题,提出一种提高节点计算准确率的计算方法,得出如下结论。

(1)依据 GraphGPS 的思想改进图神经网络的网络结构,提出一种有向网络热度提取模块,融合真实网络中的语义、行为、热度信息计算节点影响力,研究基于图神经网络的有向网络节点影响力计算方法。

(2)在 4 个公开的真实数据集上的实验表明,通过热度提取模块提取网络热度特征,本文方法可获得更高的准确率,能够有效提高模型对网络数据的拟合能力并计算节点影响力。

(3)在未来的工作中,可以考虑使用注意力机制提高热度模块提取特征的能力,进一步提高图神经网络计算节点影响力的准确率。

参 考 文 献

[1] 张永强,李爽,马金龙. 基于复杂网络的不同路由策略下交通流驱动的流行病传播[J]. 科学技术与工程, 2023, 23(4): 1370-1377.
Zhang Yongqiang, Li Shuang, Ma Jinlong. Traffic-driven epidemic spread under different routing strategies based on complex networks [J]. Science Technology and Engineering, 2023, 23(4): 1370-1377.

[2] Alvaro S, Nicolas H, Jost T S, et al. Graph networks as learnable physics engines for inference and control[J]. CoRR, 2018, 80:

- 4467-4476.
- [3] Takuo H, Hidekazu O, Masashi S, et al. Knowledge transfer for out-of-knowledge-base entities; a graph neural network approach [C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. New York: IEEE, 2017: 1802-1808.
- [4] 孔建国, 卢靖宇, 李煜琨, 等. 管制空域复杂网络模型构建及抗毁性分析[J]. 科学技术与工程, 2023, 23(9): 3973-3981.
Kong Jianguo, Lu Jingyu, Li Yukun, et al. A complex network model of controlled airspace and invulnerability analysis[J]. Science Technology and Engineering, 2023, 23(9): 3973-3981.
- [5] 王飞, 魏林琳. 基于复杂网络的空中交通流分形特征分析[J]. 科学技术与工程, 2023, 23(9): 3982-3990.
Wang Fei, Wei Linlin. Fractal characteristics analysis of air traffic flow based on complex network[J]. Science Technology and Engineering, 2023, 23(9): 3982-3990.
- [6] Sigit P, Nyoman I. Social network analysis of Twitter to identify issuer of topic using pagerank[J]. International Journal of Advanced Computer Science and Applications, 2019, 10(1): 107-111.
- [7] Meng L, Xu G, Yang P, et al. A novel potential edge weight method for identifying influential nodes in complex networks based on neighborhood and position[J]. Journal of Computational Science, 2022, 60: 101591.
- [8] Jin B L, Zou M B, Wei Z K, et al. How to find opinion leader on the online social network? [C]//Proceedings of ACM Conference. New York: ACM, 2023: 3956877.
- [9] 韩忠明, 吴杨, 谭旭升, 等. 面向结构洞的复杂网络关键节点排序[J]. 物理学报, 2015, 64(5): 429-437.
Han Zhongming, Wu Yang, Tan Xusheng, et al. Ranking key nodes in complex networks by considering structural holes[J]. Acta Physica Sinica, 2015, 64(5): 429-437.
- [10] Sun J H, Jiang Q H, Lu C W. Recursive social behavior graph for trajectory prediction [C]//Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 657-666.
- [11] Shi W J, Raj R. Point-GNN: graph neural network for 3D object detection in a point Cloud[J]. Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 1708-1716.
- [12] He X, Deng K, Wang X, et al. LightGCN: simplifying and powering graph convolution network for recommendation [C]//Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2020: 639-648.
- [13] Johannes G, Janek G, Stephan G. Directional message passing for molecular graphs[10]. arXiv Preprint, 2020; DOI: 10.48550/arXiv.2003.03123.
- [14] Ladislav R, Mikhail G, Vijay P D, et al. Recipe for a general, powerful, scalable graph transformer [C]//Conference on Neural Information Processing Systems. New York: ACM, 2022: 14501-14515.
- [15] Ahmad A R, Justin M, Mahdi J, et al. A machine learning-based approach for vital node identification in complex networks[J]. Expert Systems with Applications, 2023, 214: 119086.
- [16] Linton C F. Centrality in social networks conceptual clarification [J]. Social Networks, 1979, 1(3): 215-239.
- [17] Yang L, Qiao Y F, Liu Z Z, et al. Identifying opinion leader nodes in online social networks with a new closeness evaluation algorithm[J]. Soft Computing, 2016, 22(2): 453-464.
- [18] Francesco A M, Domingo D. Measuring the academic reputation through citation networks *via* PageRank [J]. Digital Libraries, 2019, 13(1): 185-201.
- [19] Esther G, Miguel R. Controlling centrality: the Inverse ranking problem for spectral centralities of complex networks[J]. Mathematical Methods in the Applied Sciences, 2021, 45(17): 11755-11768.
- [20] Mohammadereza D, Hamid R R, Usman A K. Centrality-based epidemic control in complex social networks[J]. Social Network Analysis and Mining, 2020, 10(1): 1-11.
- [21] Wang J K, Wang X H, Fu L. Evolutionary game model of public opinion information propagation in online social networks [J]. IEEE Access, 2020, 8: 127732-127747.
- [22] Wang F, She J, Ohyama Y, et al. Deep-learning-based identification of influential spreaders in online social networks[C]//IECON 45th Annual Conference of the IEEE Industrial Electronics Society. Lisbon: IEEE, 2019: 8927419.
- [23] Zhang H Q, Lu G, Zhan M, et al. Semi-supervised classification of graph convolutional networks with laplacian rank constraints [J]. Neural Processing Letters, 2021, 54(4): 2645-2656.
- [24] Zhao G, Jia P, Zhou A, et al. InfGCN: identifying influential nodes in complex networks with graph convolutional networks[J]. Neurocomputing, 2020, 414: 18-26.
- [25] Yu E Y, Wang Y P, Fu Y, et al. Identifying critical nodes in complex networks *via* graph convolutional networks [J]. Knowledge-Based Systems, 2020, 198: 105893.
- [26] Ou Y, Guo Q, Xing J L, et al. Identification of spreading influence nodes *via* multi-level structural attributes based on the graph convolutional network [J]. Expert Systems with Applications, 2022, 203: 117515.
- [27] Michael D, Xavier B, Pierre V. Convolutional neural networks on graphs with fast localized spectral filtering[J]. Advances in Neural Information Processing Systems, 2016, 29: 3844-3852.
- [28] William L H, Rex Y, Jure L. Inductive representation learning on large graphs[J]. Advances in Neural Information Processing Systems, 2017, 30: 1024-1034.
- [29] Petar V, Guillem C, Arantxa C, et al. Graph attention networks [C]//International Conference on Learning Representations. New York: IEEE, 2018: 10903.
- [30] Chen J, Li Z L, Zhu Y, et al. From node interaction to hop interaction; new effective and scalable graph learning paradigm[J]. CVPR, 2023, 11: 7876-7885.
- [31] Tong Z, Liang Y, Sun C S, et al. Digraph inception convolutional networks[J]. Neural Information Processing Systems, 2020, 33: 17907-17918.