

一种动态优化的知识超图网络地铁客流预测

张阳,王冲聪,陈麟

(福建理工大学人工智能与交通工程学院,福建福州 350118)

摘要: 针对地铁站点间高阶关系问题,提出一种动态优化知识超图网络地铁客流预测模型。该模型结合多尺度时空卷积块和时空变换器网络构成高阶信息融合模块,通过动态知识图谱嵌入技术构建地铁网络的动态超图卷积结构,利用多尺度时空卷积块提取地铁站点和客流量的时空特征,并借助高阶信息融合模块实现时空特征的融合。采用基于知识超图的时空变换器网络训练地铁站点的动态嵌入模式,并构建不同时间窗口的动态相似矩阵,以更精确地捕捉站点的动态特征变化。在南昌地铁真实客流数据集上的实验表明,模型在工作日和周末两个数据集上的均方根误差分别为18、27,平均绝对误差分别为14、23;和传统GCN预测模型进行对比,工作日平均绝对百分误差降低了13%,周末平均绝对百分误差降低了12%,显示出更优的预测性能。

关键词: 知识超图;地铁;客流量预测;卷积网络

中图分类号: U293.13

文献标志码: A

文章编号: 2097-3853(2026)03-0266-08

A dynamic optimized knowledge hypergraph network subway passenger flow prediction

ZHANG Yang, WANG Chongcong, CHEN Lin

(School of Artificial Intelligence and Transportation Engineering, Fujian University of Technology, Fuzhou 350118, China)

Abstract: A dynamic optimization knowledge hypergraph network model was proposed for subway passenger flow prediction, which addresses the high-order relationship between subway stations by integrating a high-order information fusion module of multi-scale spatio-temporal convolutional blocks and spatio-temporal transformer networks. The model constructs a dynamic hypergraph convolutional structure of the subway network through dynamic knowledge graph embedding technology, utilizes the multi-scale spatio-temporal convolutional block to extract spatio-temporal features of subway stations and passenger flow, and fuses these spatio-temporal features through a high-order information fusion module. Finally, a method based on a hypergraph spatiotemporal transformer network is employed to train the dynamic embedding patterns of subway stations and construct dynamic similarity matrices for different time windows, thereby more precisely capturing the dynamic changes in station features. Experiments conducted on real passenger flow datasets of Nanchang Metro show that the model achieves a root mean square error (RMSE) of 18 and 27, and a mean absolute error (MAE) of 14 and 23 on weekday and weekend datasets, respectively. Compared with traditional GCN prediction models, the mean absolute percentage error (MAPE) is reduced by 13% on weekdays and 12% on weekends, showing better prediction performance.

Keywords: knowledge hypergraph; subway; passenger flow prediction; convolutional network

地铁站点间的客流关联往往蕴含着复杂的高阶关系,这些关系不仅影响周边商业发展,也关系

着客流分布与交通效率。因此,从高阶关联角度对地铁客流量进行准确预测有助于交通规划者优

收稿日期:2025-03-25

基金项目:福建省自然科学基金资助项目(2023J01946)

第一作者简介:张阳(1983—),男,湖北武汉人,教授,博士,研究方向:铁路客流预测。

化站点布局、改进调度策略,对提高交通效率有较强的研究意义和应用价值。

当前,地铁客流量预测大多基于深度学习模型,通过深层次的神经网络层自适应地挖掘交通客流的逻辑关系,具有较好的空间特征学习能力。其中,图结构因其能够有效挖掘交通流深层次的特征成为客流预测的常用表示方法。卷积神经网络(convolutional neural network, CNN)^[1]可以将时间序列数据转换成图像数据的形式进行特征提取和预测,但难以直接应用于复杂地铁站台图。图卷积神经网络(generalized connection network, GCN)^[2]能够利用站点之间的连接关系传递信息、提取空间特征,且能够适应地铁客流量随时间变化的特点处理动态图数据,从而更好地捕捉客流量的时序变化规律,是目前应用最广泛的预测算法之一。

然而,地铁客流受到多种复杂因素的影响,包括站点间的联系、换乘关系、特定时间段的人流模式等,普通图结构难以表达地铁站点间的多重复杂关系。为了深入挖掘高阶相互作用,知识图谱和超图等高阶异构图结构逐渐受到关注。例如,毛慧慧等人^[3]为了结合动态知识图谱嵌入、残差网络和长短期记忆网络进行短期地铁客流量预测方法;罗秋雨^[4]等人基于地铁网络拓扑设计知识图谱嵌入的超参数选择;Wang^[5]等人引入外部因素构建地铁交通知识图谱以提升预测性能;Wang^[6]等人提出动态超图神经网络框架以挖掘

隐藏关系;Wan 等人^[7]则使用时间点超图学习时间点关联,并通过多静态超图与信息融合模块整合时空特征。上述研究均表明,知识图谱和超图能够更充分地刻画站点间的高阶关联,从而显著提高预测精度和效率。

综上,本研究提出了一种动态优化的知识超图网络地铁客流预测模型,通过动态知识图谱嵌入技术构建地铁网络的动态超图卷积结构,利用多尺度时空卷积块提取时空特征,并通过基于质心的反向学习策略的混沌粒子群算法动态优化重要参数、提升算法性能,以期在充分挖掘地铁各站点间高阶关系的同时有效捕捉时空关系,实现更有效的地铁客流量预测。

1 面向地铁网络的动态知识超图卷积网络

为了深入挖掘地铁站点间的高阶关系,本研究采用知识图谱与超图相结合的方法对地铁网络进行建模,通过动态知识图谱嵌入技术构建动态超图卷积结构,并结合多尺度时空卷积块与高阶信息融合模块,实现时空特征的同步提取与融合。在此基础上,引入基于超图的时空变换器网络,学习站点的动态嵌入表示并生成动态相似矩阵,从而提升客流预测的准确性与模型适应性。图 1 为用于客流预测的动态超图卷积网络客流预测模型框架示意图。

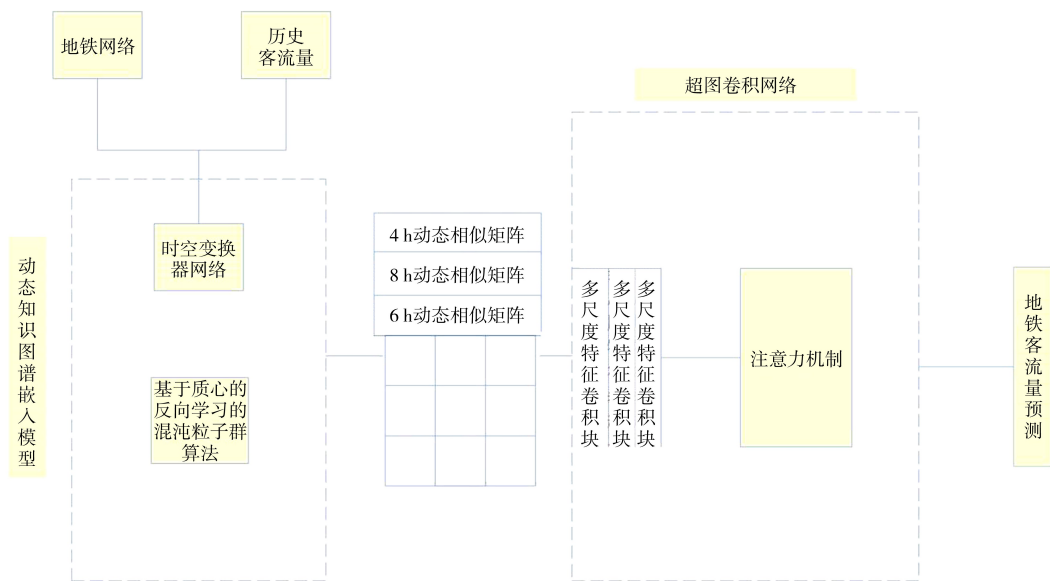


图 1 超图卷积网络客流预测模型框架示意图

Fig.1 Hypergraph convolutional network passenger flow prediction model framework diagram

模型使用动态知识图谱嵌入技术将地铁站点及其间的高阶关系映射到低维向量空间,从而能够有效捕捉地铁网络中各站点之间的动态交互关系并减少运算量。动态知识图谱四元组为 (h, r, t, τ) , 嵌入模型表示如式(1)所示。

$$e_{(h,r,t,\tau)} = e_h + e_r + e_t + e_\tau \quad (1)$$

式中, h, r, t 和 τ 分别表示头实体、关系、尾实体和时间戳, e_h, e_r, e_t 和 e_τ 分别是头实体、关系、尾实体和时间戳的嵌入向量。

将动态知识图谱四元组映射至地铁客流预测场景, 由于地铁客流预测的核心实体为站点, 重点关系为各站点间的客流流量模式, 动态性体现为站点间的客流流量随时间而变化, 所以本研究的具体映射方式是将头实体设为站点, 关系为站点客流, 尾实体为具体客流量数, 时间戳为训练和预测的时间段, 通过式(1)构建出站点客流知识图谱动态四元组。

为了更准确地捕捉交通特征, 本研究使用时空转换器网络训练地铁站点的动态嵌入模式, 并构建动态相似矩阵。在每个时间窗口内, 应用时空变换器, 如式(2)~式(5)所示。

$$PE_{(pos, 2i)} = \sin\left(\frac{pos}{10\,000^{\frac{2i}{d}}}\right) \quad (2)$$

$$APE_{(pos, 2i+1)} = \cos\left(\frac{pos}{10\,000^{\frac{2i}{d}}}\right) \quad (3)$$

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (4)$$

$$\text{FFN} = \max(0, x\mathbf{W}_1 + b_1)\mathbf{W}_2 + b_2 \quad (5)$$

式中, $APE_{(pos, 2i+1)}, PE_{(pos, 2i)}$ 是位置编码, pos 是位置, i 是维度索引, d 是嵌入维度; $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 分别是查询、键和值向量; d_k 是键向量的维度, Attention 为注意力机制, softmax 为归一化函数, $\mathbf{W}_1, \mathbf{W}_2, b_1, b_2$ 是可训练的权重和偏置参数, FFN 为前馈神经网络层。

为了捕捉地铁站点间的动态特征变化, 本研究在不同时间窗口内构建动态相似矩阵, 相似度公式为

$$S_{i,j,k} = \frac{h_{i,k} \cdot h_{j,k}}{\|h_{i,k}\| \|h_{j,k}\|} \quad (6)$$

式中, $S_{i,j,k}$ 为相似度矩阵, $h_{i,k}$ 和 $h_{j,k}$ 分别表示时间窗口 k 内的站点 i 和站点 j 的嵌入向量。

动态知识图谱嵌入模型可以轻松扩展以包含

新的站点或关系类型, 在此基础上构建超边关系, 可以更精细地表达地铁站点间高阶关系的动态变化。超边上的每个实体都附有时间戳, 以捕捉实体间关系的动态变化, 如式(7)所示。超边 e 与其连接的顶点及其与时间戳的关系表示如式(8)所示。

$$\text{KHG} = (\mathbf{V}, \mathbf{E}, \mathbf{T}) \quad (7)$$

$$e = \{(v_1, t_1), (v_2, t_2), \dots, (v_n, t_n)\} \quad (8)$$

式中, KHG 为超图集合, \mathbf{V} 是顶点集合, \mathbf{E} 是超边集合, \mathbf{T} 是时间戳集合。每条超边 $e \in \mathbf{E}$ 可以连接两个或两个以上顶点, 表示为 $e = \{v_1, v_2 \dots v_n\}$, 其中 $v_i \in \mathbf{V}$, 每个顶点 v_i 都有一个与之相关的时间戳 $t_i \in \mathbf{T}$, 表示该顶点在时间 v_i 的状态, (v_i, t_i) 表示顶点 v_i 在时间戳 t_i 的状态。

将超图中的超边 e_i 定义为地铁的第 i 条线路, 而 v_i 定义为第 i 个站点, 则超图的关联矩阵 $\mathbf{I}(v_i, e_j)$ 可以用式(9)定义。

$$\mathbf{I}(v_i, e_j) = \begin{cases} 1, & v_i \in e_j \\ 0, & v_i \notin e_j \end{cases} \quad (9)$$

对于超图中的每一个节点 v , 节点的度 $d(v_i)$ 可以定义为与节点 v_i 相连的所有超边权重之和, 如式(10)所示。

$$d(v_i) = \sum w(e_j) \cdot \mathbf{I}(v_i, e_j) \quad (10)$$

式中, $d(v_i)$ 为节点 v_i 的度, $w(e_j)$ 为边 e_j 的权重。

超边的度则可以定义为与其相连的所有顶点的数量, 定义式如式(11)所示。

$$\delta(e_i) = \sum \mathbf{I}(v, e_i) \quad (11)$$

式中, $\delta(e_i)$ 为超边 e_i 的度。

利用超图的关联矩阵和度矩阵, 可根据式(12)得到超图的邻接矩阵。

$$\mathbf{A}^l = \mathbf{I}\mathbf{W}^l\mathbf{T}^l - \mathbf{D}_v \quad (12)$$

式中, \mathbf{A}^l 为邻接矩阵, \mathbf{W} 为权重矩阵, \mathbf{D} 是度矩阵。

为了捕捉地铁网络中的高阶关系, 本研究采用超图卷积层。假设超图中有 k 个超边, 每个超边连接多个节点。通过堆叠多个超图卷积层, 将初始低维度向量逐层转换为高维特征表示。每一层的输入是上一层的输出。超图卷积如式(13)所示。

$$H^{(l+1)} = \sigma(D^{-\frac{1}{2}}\mathbf{A}H^{(l)}\mathbf{W}^{(l)}D^{-\frac{1}{2}}) \quad (13)$$

式中, $H^{(l)}$ 是第 l 层的节点嵌入, \mathbf{A} 是超图的邻接矩阵, $\mathbf{W}^{(l)}$ 是第 l 层的可训练权重矩阵, \mathbf{D}_k 是超图

度矩阵, σ 是非线性激活函数。

图卷积方法虽然能够基于节点组成的邻接矩阵来捕获节点之间的相互关联,但无法直接应用于超图。为了充分挖掘时间和空间特征,本研究引入多尺度时空卷积块(multi-scale spatiotemporal convolutional block, MSCB)。对于给定的节点嵌入,多尺度时空卷积块通过不同尺度的卷积核进行卷积操作,如式(14)所示。

$$\mathbf{H}^{(S)} = \sum_{i=1}^S \text{Conv}_{k_i}(\mathbf{H}) \quad (14)$$

式中, Conv_{k_i} 表示卷积核为 k_i 的卷积操作, S 表示不同尺度的数量。

为了进行时空特征的同步融合,采用注意力机制整合不同尺度的特征表示。多尺度时空卷积块的输出为 $\mathbf{H}^{(S)}$, 高阶信息融合模块表示如式(15)~式(17)所示。

$$\mathbf{Q}_S = \mathbf{W}_Q \mathbf{H}^{(S)}, \mathbf{K}_S = \mathbf{W}_K \mathbf{H}^{(S)}, \mathbf{V}_S = \mathbf{W}_V \mathbf{H}^{(S)} \quad (15)$$

$$\alpha_s = \frac{\exp(\mathbf{Q}_S \cdot \mathbf{K}_S^T / \sqrt{d_k})}{\sum_{s'=1}^S \exp(\mathbf{Q}_{S'} \cdot \mathbf{K}_{S'}^T / \sqrt{d_k})} \quad (16)$$

$$\mathbf{H}^{(j)} = \sum_{S=1}^S \alpha_S \mathbf{V}_S \quad (17)$$

式中, $\mathbf{Q}_s, \mathbf{K}_s, \mathbf{V}_s$ 分别表示查询矩阵、键矩阵和值矩阵, α_s 是注意力权重, d_k 是键矩阵的维度, \mathbf{W}^0 是输出权重矩阵, $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V$ 为可训练权重矩阵, $\mathbf{H}^{(j)}$ 表示融合后的特征表示。

2 动态知识超图卷积网络模型优化方法

为了能够合理动态调整参数,有效提升模型的预测精度,本研究根据超图卷积层结构维度特征,提出一种基于质心的反向学习策略的混沌粒子群算法(T-MFO),对预测模型的结构维度进行优化,使其能够逐步选择最优的参数。

2.1 基于质心的反向学习策略的混沌粒子群算法

初始解对算法的计算量和收敛速度有很大影响,所以需要提高初始种群的质量。标准 PSO 算法^[8]随机生成初始解,可能出现集聚的低质量个体,这些个体远离全局最优解,会增加迭代的次数和计算量,太远的距离还会导致粒子陷入局部最优解。为提高初始种群的质量,提高收敛速度,本研究采用 Tent 混沌映射并结合方向学习策略^[9]

生成初始解。

Tent 混沌映射在变量空间中按一定规则生成混沌序列,再将其映射到目标函数定义域内,生成初始种群。由此生成的初始种群可遍历定义域并具有良好的均匀性,也可使初始解具有更好的多样性。Tent 映射的数学表达式如式(18),为防止周期不稳定导致映射陷入不动点,在基础 Tent 映射公式中添加了随机数。

$$x_{i+1} = \begin{cases} \frac{\lambda x_i + r(0,1)}{n}, & 0 \leq x \leq \lambda \\ \frac{(1-\lambda)(1-x_i) + r(0,1)}{n}, & \lambda \leq x \leq 1 \end{cases} \quad (18)$$

式中, λ 为控制参数,一般为 0.5, i 为序列号, $r(0,1)$ 为 $(0,1)$ 之间的随机数。在目标函数的定义域内进行混沌映射,包括 3 个步骤。

步骤 1: 随机生成一组 n 个元素向量 $\mathbf{a}_1 = (a_{11}, a_{12}, \dots, a_{1n})$, 向量中每个元素介于 $(0,1)$ 之间。

步骤 2: 根据式(18)求得第 j 列元素的混沌序列 $(a_{1j}, a_{2j}, \dots, a_{mj})$, 重复 n 次可得混沌序列矩阵。

步骤 3: 将混沌矩阵映射至适应度函数定义域内,得到初始解矩阵,映射公式如式(19)。

$$x_{ij} = x_{\max} - (x_{\max} - x_{\min}) a_{ij} \quad (19)$$

式中, x_{\max} 与 x_{\min} 分别为适应度函数定义域的上下界, a_{ij} 为 i 行 j 列的混沌数。

得到初始解后,利用质心反向学习策略求得初始解 $\{x_i\}$ 的反向序列,对二者进行比较,选取其中优者作为新的初始解 $\{x_i^*\}$, 以此提高初始种群的质量。假设每个粒子都具有质量,每个粒子的初始值就是粒子的质量,根据式(20)可得第 i 维粒子的质心。

$$M_i = \frac{1}{n} \sum_{j=1}^n x_{i,j} \quad (20)$$

式中, M_i 为第 i 维粒子的质心, n 为该维粒子个数。

再根据第 i 维的质心按公式(21)求得第 i 维粒子的反向序列,与正向序列对比,选取其中适应度函数更优的 n 个粒子作为第 i 维初始解集合。

$$x_{ij}^* = 2M_i - x_{ij} \quad (21)$$

式中, x_{ij}^* 为粒子 x_{ij} 的反向粒子。

惯性权重 w 表示迭代前后粒子的速度保留量,对算法的迭代速度和精度有很大影响。标准 PSO 算法中的 w 保持不变,前后期速度变化幅度一致,但在优化过程中,对速度变化的要求不同。

前期距离粒子距离全局最优解较远,粒子速度保留量应较大,提升收敛速度和局部探索能力,后期粒子距离全局最优解较近,速度保留量应较小,提升算法收敛的精度,减少迭代次数。故本研究按照粒子与最优解之间的距离对惯性权重进行改进。为方便表示当前粒子位置与全局最优位置之间的距离,定义距离因子如式(22)所示。

$$D(x_{ij}) = \mu \frac{F_{OA} - F_i}{F_{OA} - F_1} \quad (22)$$

其中, $D(x_{ij})$ 为粒子 x_{ij} 的距离因子, F_{OA} 为预测最优的适应度函数, F_1 为初始适应度函数, μ 为控制参数, F_i 为第 i 次迭代时的适应度函数。

利用距离因子和逻辑回归模型对惯性权重进行改进,逻辑回归模型首尾比较于线性模型更为平缓,有利于精细化调整。改进后的惯性权重如式(23)。

$$\omega_{ij} = \omega_{\max} - \frac{\lambda}{1 + \exp[-10D(x_{ij}) + 5]} (\omega_{\max} - \omega_{\min}) \quad (23)$$

式中, ω_{\max} 、 ω_{\min} 分别为惯性权重的最大值和最小值, ω_{ij} 为第 ij 个粒子的惯性权重, λ 为调整参数。按上式调整的惯性权重,在粒子距离最优解较远时较大,随着距离减小惯性权重非线性减小,利于平衡粒子探索和开发的能力。

2.2 动态知识超图卷积网络地铁客流预测步骤

利用城市居民地铁出行的刷卡数据,对数据进行分析 and 预处理。去除噪声及无关数据后,按 5 min 为间隔将所得地铁客流量数据聚合,使其满足预测模型的输入数据格式要求,然后进行地铁客流量预测。利用基于质心的反向学习策略的混沌粒子群算法优化动态知识图谱嵌入模型中的更新策略,以损失函数作为最小寻优目标,损失函数为预测值和实际值之间的均方根误差(RMSE)与均绝对百分误差(MAPE)之和^[10],如式(24)所示。

$$\text{Loss} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i)^2} + \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (24)$$

式中, i 为样本数量, x_i 为在时刻 i 的真实值, \hat{x}_i 为在时刻 i 的预测值, N 为预测数据的总量。

步骤 1:采集地铁网络中的客流样本数据,对数据进行分析和处理,去除噪声,每五分钟聚合一次,数据满足输入格式。

步骤 2:构建动态知识图谱嵌入模型,以捕捉地铁站点间的复杂关系和动态变化特征。

步骤 3:构建多尺度时空卷积块,以提取地铁站点和客流量的时空特征。通过卷积操作捕获不同时间尺度下的时空模式,增强模型对多样化时空特征的表达能力。该模块利用图注意力机制整合了地铁站点间的高阶关系,从而提升了时空特征的表达效果。

步骤 4:采用基于超图的时空变换器网络方法,训练地铁站点的动态嵌入模式。通过构建不同时间窗口的动态相似矩阵,捕捉站点的动态特征变化,提高预测精度。

步骤 5:对动态知识图谱嵌入模型的更新策略进行优化,利用基于质心的反向学习策略的混沌粒子群算法寻找到参数的全局最优解,以期寻找到最优的参数。

步骤 6:对优化过后的地铁客流预测模型进行训练,再利用训练好的预测模型对实际的客流量进行预测,并与实际数据进行对比验证模型精度。

3 实验分析

3.1 研究范围

本研究以南昌地铁八一广场站为研究对象,如图 2 所示。

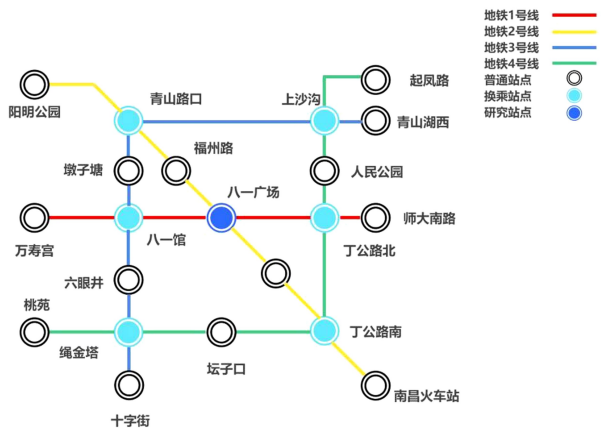


图 2 八一广场站及附近站点示意图

Fig.2 Schematic diagram of Bayi Square Station and nearby stations

南昌地铁共包含 4 条线路、121 个站点,实验数据涵盖八一广场站和其周边 20 个站点的历史客流,以及以该站为起讫点的客流数据。数据采集时间范围为每日 6:00 至 24:00,间隔为 5 min。使用 3 月份数据作为训练集,对 T-MFO-DHGCN

模型进行训练,并分别预测 3 月 28 日(周五)与 3 月 29 日(周六)的客流量,通过对比预测值与实际值以验证模型精度。

3.2 数据处理

在客流预测建模过程中,设备故障或乘客操作不当会导致刷卡数据异常,从而影响模型的预测能力,所以需要异常数据进行清洗。异常数据的类型主要有以:完全重复的数据、字节丢失的数据和进出站时间出现明显错误的的数据,如进出站时间一致或进出站时间间隔过长(>3h)。对于第一类重复数据,删除冗余记录;对于第二类缺失数据,基于历史相似信息进行填补;对于第三类时间错误数据,由于其占比低且修正复杂,直接利用 python 识别并删除。

数据清洗完成后,按 5 min 为间隔进行聚合。相较于更长的时间段,如 10 min 或 15 min, 5 min 为间隔能获取更细致的客流变化信息,且更具有随机性,而过长的时间段可能导致重要信息丢失;而更短的时间间隔,如 4 min 或 1 min,会导致噪声过大以及计算量激增。

3.3 指标选取

为了更准确判断模型的预测能力,本研究采用平均绝对误差(mean absolute error, MAE)^[11],均方根误差(root mean square error, RMSE)^[12],平均绝对百分误差(MAPE)^[13]和相关系数 R^2 ^[14] 作为评价指标。

$$MAE = \frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i) \times 100\% \quad (25)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i)^2} \times 100\% \quad (26)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{x}_i - x_i}{x_i} \right| \times 100\% \quad (27)$$

$$R^2 = \frac{\sum_{i=1}^N (\hat{x}_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (28)$$

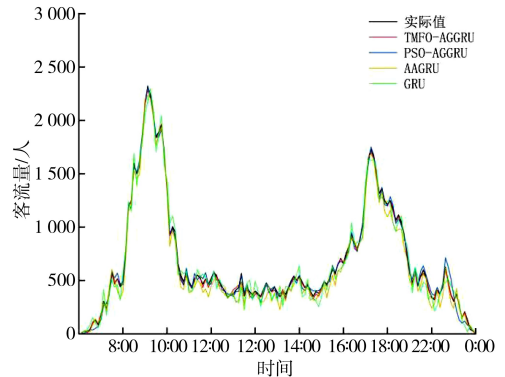
式中, i 为样本数, x_i 为实际值, \hat{x}_i 为在时刻 i 的预测值, \bar{x} 为客流量实际值的平均值, N 为预测数据的总量。

3.4 结构优化的图卷积网络对预测性能的影响

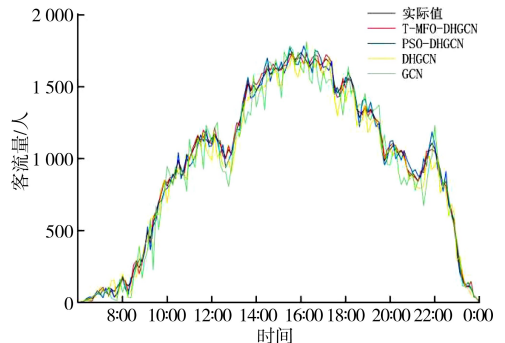
为检验优化算法和知识超图对图卷积网络预测能力的影响,本研究分别用 T-MFO-DHGCN、

PSO-DHGCN、DHGCN 以及传统 GCN^[15] 模型对地铁客流量进行预测,检验 3 种模型的预测能力。

本次实验利用历史数据对模型进行训练,并分别对工作日和周末地铁客流进行预测并评估,预测结果如图 1 所示。评估结果如表 1、表 2 所示。



(a) 工作日预测结果对比



(b) 周末预测结果对比

图 3 预测结果对比

Fig.3 Comparison of prediction results

表 1 工作日预测结果对比

Tab.1 Comparison of weekday prediction results

| 评价指标 | RMSE | MAE | R^2 | MAPE |
|-------------|------|-----|-------|------|
| T-MFO-DHGCN | 18 | 14 | 0.99 | 0.04 |
| PSO-DHGCN | 42 | 32 | 0.98 | 0.10 |
| DHGCN | 66 | 53 | 0.98 | 0.11 |
| GCN | 76 | 63 | 0.97 | 0.17 |

表 2 周末预测结果对比

Tab.2 Comparison of weekend prediction results

| 评价指标 | RMSE | MAE | R^2 | MAPE |
|-------------|------|-----|-------|------|
| T-MFO-DHGCN | 27 | 23 | 0.99 | 0.07 |
| PSO-DHGCN | 45 | 38 | 0.98 | 0.11 |
| DHGCN | 70 | 54 | 0.98 | 0.12 |
| GCN | 104 | 78 | 0.92 | 0.19 |

由图 2、表 1 和表 2 可得, T-MFO-DHGCN 模型在工作日与周末的客流预测中, 各项评价指标均优于对比模型。在工作日预测中, T-MFO-DHGCN 的 RMSE 与 MAE 分别为 18 和 14, 优于其他模型。周末预测中其 RMSE 与 MAE 为 27 和 23, 同样是最低值。综上, T-MFO-DHGCN 模型能够有效预测地铁客流, 优化算法和知识超图结构能够提高图卷积网络模型的预测精度, 且相较于标准 PSO 模型, 改进后的粒子群算法对模型预测能力的提升更为明显。

3.5 结构优化的图卷积网络与经典算法对比

为检验 T-MFO-DHGCN 与其他深度学习预测模型^[16-18]之间预测精度的差异, 本研究同时对经典预测模型进行训练和客流预测, 并将结果进行对比。本研究利用地铁客流历史数据对经典预测模型进行训练, 并对 3 月 28 日(周五)与 3 月 29 日(周六)的客流量进行预测, 预测结果如图 3 所示, 评价结果如表 3、表 4 所示。

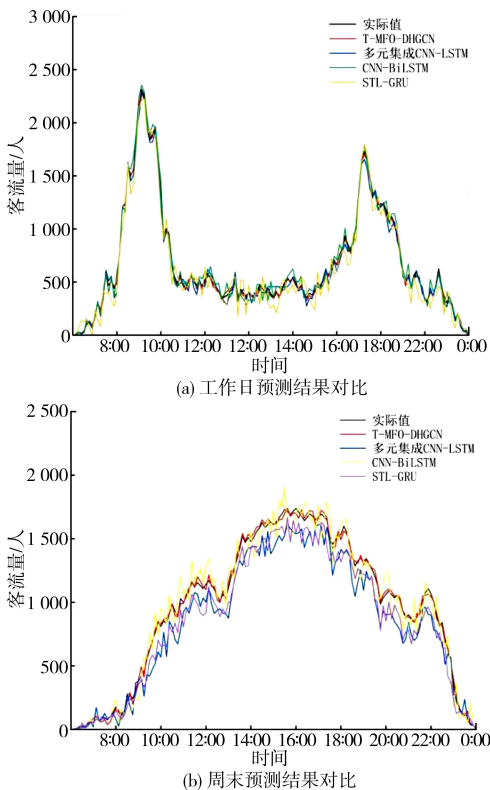


图 4 预测结果对比

Fig.4 Comparison of prediction results

根据图 3、表 3、表 4 可知, 本研究提出的 T-MFO-DHGCN 模型预测结果的评价指标优于其他预测模型, 表现出了更好的预测精度。其中, 相较

于多元集成 CNN-LSTM、CNN-BiLSTM、STL-GRU, T-MFO-DHGCN 模型在工作日预测中的 RMSE、MAE、MAPE 明显更小; 在周末预测中, RMSE 分别减少了 82%、78% 和 73%, MAE 分别减少了 81%、78% 和 73%, MAPE 分别减少了 56%、56% 和 59%。因此, 本研究提出的预测模型能够有效预测地铁客流。而且, 预测模型的工作日预测精度高于周末。分析工作日与周末客流数据可知, 工作日客流受上下班时间和地点限制, 进出站模式较为平稳, 数据波动性小, 且数据集样本量较大, 有利于预测模型获取客流变化规律。而周末客流到离站时间和地点都具有随机性, 数据波动性较大, 预测精度有所降低。但本研究提出的预测模型精度降低幅度最小, 在训练集较少且数据波动性大的情况下仍表现出了较高的预测精度。

表 3 工作日预测结果对比

Tab.3 Comparison of weekday prediction results

| 评价指标 | RMSE | MAE | R^2 | MAPE |
|---------------|------|-----|-------|------|
| T-MFO-DHGCN | 18 | 14 | 0.99 | 0.04 |
| 多元集成 CNN-LSTM | 57 | 39 | 0.98 | 0.09 |
| CNN-BiLSTM | 102 | 85 | 0.86 | 0.25 |
| STL-GRU | 55 | 44 | 0.94 | 0.1 |

表 4 周末预测结果对比

Tab.4 Comparison of weekend prediction results

| 评价指标 | RMSE | MAE | R^2 | MAPE |
|---------------|------|-----|-------|------|
| T-MFO-DHGCN | 27 | 23 | 0.99 | 0.07 |
| 多元集成 CNN-LSTM | 144 | 122 | 0.98 | 0.16 |
| CNN-BiLSTM | 127 | 108 | 0.98 | 0.16 |
| STL-GRU | 103 | 85 | 0.94 | 0.17 |

4 结束语

本研究提出一种结合了优化算法的知识超图卷积网络地铁客流预测方法, 将知识超图结构融入图卷积网络, 提高空间信息的处理能力, 并提出基于质心的反向学习策略的混沌粒子群优化算法对预测模型的隐层结构进行优化。利用实际客流数据对提出的模型进行验证可以得出, 基于质心

的反向学习策略的混沌粒子群优化算法相较于传统粒子群优化算法对模型有更好的优化效果,将本文提出的预测方法和部分经典预测模型进行对比,结果表明本研究提出的模型能够有效预测地

铁客流,并且在预测精度上相较于经典模型有一定程度的提高。综上所述,本文提出的模型能够有效捕获站点和客流之间的时空特征,对提升地铁客流预测精度有一定的积极作用。

参考文献:

- [1] 伏军霞. 基于图卷积网络的交通流量预测方法的研究[D]. 南京:南京邮电大学,2022.
- [2] LIU T B, ZHANG J D. An adaptive traffic flow prediction model based on spatiotemporal graph neural network[J]. The Journal of Supercomputing, 2023, 79(14):15245-15269.
- [3] 毛慧慧, 赵小乐, 杜圣东, 等. 基于时序知识图谱嵌入的短期地铁客流量预测[J]. 计算机科学, 2023, 50(7):213-220.
- [4] 罗秋雨, 乐阳, 谷岩岩. 城市地铁出行知识图谱嵌入表达的超参数选择[J]. 地球信息科学学报, 2023, 25(6):1164-1175.
- [5] WANG S, LÜ Y M, PENG Y, et al. Metro traffic flow prediction via knowledge graph and spatiotemporal graph neural network[J]. Journal of Advanced Transportation, 2022:2348375.
- [6] WANG J C, ZHANG Y, WEI Y, et al. Metro passenger flow prediction via dynamic hypergraph convolution networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(12):7891-7903.
- [7] WANG S Q, ZHANG Y J, QI H, et al. Dynamic spatial-temporal hypergraph convolutional network for skeleton-based action recognition[C]//2023 IEEE International Conference on Multimedia and Expo (ICME). Brisbane, Australia: IEEE, 2023:2147-2152.
- [8] KENNEDY J, EBERHART R. Particle swarm optimization[C]//Proceedings of ICNN'95-International Conference on Neural Networks. Perth, WA, Australia: IEEE, 2002:1942-1948.
- [9] ZHU Y M, LAI Y S, MEI C, et al. A whale optimization algorithm based on simulated annealing and adaptive mutation[C]//2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT). Nanjing, China: IEEE, 2024:773-778.
- [10] 张金雷, 陈奕洁, Panchamy Krishnakumari, 等. 基于注意力机制的城市轨道交通网络级多步短时客流时空综合预测模型[J]. 地球信息科学学报, 2023, 25(4):698-713.
- [11] 姚恩建, 周文华, 张永生. 城市轨道交通新站开通初期实时进出站客流量预测[J]. 中国铁道科学, 2018, 39(2):119-127.
- [12] 梁强升, 许心越, 刘利强. 面向数据驱动的城市轨道交通短时客流预测模型[J]. 中国铁道科学, 2020, 41(4):153-162.
- [13] 戚耀, 王晨菡, 吴啸宇, 等. 地铁短时客流预测的 ATT-BiLSTM 模型[J]. 交通科技与经济, 2025, 27(1):89-96.
- [14] 鲁文博, 张永, 李培坤, 等. 自适应多视图融合图神经网络地铁客流预测模型[J]. 交通运输系统工程与信息, 2024, 24(3):194-203.
- [15] 唐郑熠, 黄嘉欢, 王金水, 等. 基于自适应扩散图卷积注意力网络的地铁客流预测[J]. 铁道科学与工程学报, 2024, 21(12):4910-4923.
- [16] ZHANG Y, XIN D R. A diverse ensemble deep learning method for short-term traffic flow prediction based on spatiotemporal correlations[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(9):16715-16727.
- [17] ZHUANG W Q, CAO Y B. Short-term traffic flow prediction based on CNN-BiLSTM with multicomponent information[J]. Applied Sciences, 2022, 12(17):8714.
- [18] 赵建东, 朱丹, 刘佳欣. 基于时间序列分解与门控循环单元的地铁换乘客流预测[J]. 华南理工大学学报(自然科学版), 2022, 50(5):22-31.

(责任编辑:方素华)