文章编号:1005-0523(2025)03-0077-10

参数优化的图卷积门控循环网络地铁客流预测

张 阳,李露玢,陈燕玲

(福建理工大学交通运输学院,福建 福州 350118)

摘要:充分挖掘地铁网络相关站点间客流的空间关联性对地铁客流预测精度的提升有积极作用。由于地铁各站点之间的空间相关性难以学习并传递,捕捉并量化客流数据的空间规律十分困难。提出一种改进的图卷积门控循环神经网络地铁客流预测模型,通过整合多元时空数据提升模型处理不同数据类型的能力,采用基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法动态调整模型结构参数,以优化门控循环神经网络的隐层结构。实验结果表明,在工作日模型的预测精度明显高于周末,相较于周末,工作日的均方根误差、平均绝对误差、平均绝对百分误差分别降低了13个百分点、12个百分点、0.08个百分点。参数优化门控循环神经网络的隐层结构可以获得更好的收敛效果,预测精确度更高。

关键词:门控循环神经网络;图卷积运算;注意力机制;莱维飞行扰动策略;地铁客流预测

中图分类号:U293.5 **文献标志码:**A

本文引用格式:张阳,李露玢,陈燕玲. 参数优化的图卷积门控循环网络地铁客流预测[J]. 华东交通大学学报,2025,42(3):77-86.

Parameter Optimization of Graph Convolution Gated Recurrent Neural Network for Subway Passenger Flow Prediction

Zhang Yang, Li Lubin, Chen Yanling

(School of Transportation, Fujian University of Technology, Fuzhou 350118, China)

Abstract: Fully exploiting the spatial correlation of passenger flow between related stations in the subway network has a positive effect on the improvement of subway passenger flow prediction accuracy. Capturing and quantifying spatial patterns in passenger flow data is difficult due to the difficulty of learning and transferring spatial correlations between metro stations. An improved graph-convolution gated recurrent neural network (GC-GRU) metro passenger flow prediction model was proposed to enhance the model's ability to handle different data types by integrating multivariate spatio-temporal data. The spider wasp optimisation (SWO) algorithm based on Tent chaotic mapping and Levy flight disturbance strategy was used to dynamically adjust the model structural parameters in order to optimize the hidden layer structure of the gated recurrent neural network. The experimental results show that the prediction accuracy of the model is significantly higher on weekdays than on weekends, and the root mean square error, mean absolute error, and mean absolute percentage error are reduced by 13 percentage points, 12 percentage points, and 0.08 percentage points, respectively, during weekdays compared to weekends. Dynamic optimization of the hidden structure of gated recurrent networks can lead to better convergence of the prediction model and higher prediction accuracy.

收稿日期:2024-03-08

基金项目:福建省自然科学基金项目(2023J01946)

Key words: gated recurrent neural network; graph convolution operation; attention mechanism; Levy flight disturbance; subway passenger flow prediction

Citation format: ZHANG Y, LI L B, CHEN Y L. Parameter optimization of graph convolution gated recurrent neural network for subway passenger flow prediction[J]. Journal of East China Jiaotong University, 2025, 42(3): 77−86.

随着我国城市化进程的加快,城市人口逐年增长,地铁面临着越来越大的客流压力,合理的列车调度及客流管控策略受到地铁运营方的重视。较为精准的客流量预测能为列车调度及客流分配提供必不可少的基础数据,成为了地铁优化运营过程中的重要一环,也是目前的一个热点研究方向。

传统客流预测方法是基于统计理论的预测模型,主要有历史平均法[1]、时间序列模型[2]、卡尔曼滤波模型[3]和灰色模型。常见的时间序列预测模型有差分整合移动平均自回归模型(ARIMA)[4]、季节性差分自回归滑动平均模型(seasonal auto-regressive integrated moving average, SARIMA)[5]等。但是统计理论模型不能够较好的适应客流数据因外部环境改变带来的突变。随着机器学习的发展,出现了以支持向量回归(SVR)为代表的预测模型[6]。这类模型能够获取比统计模型更高的预测精度,但由于其浅层结构给学习能力带来的局限性,还不能很好的捕捉波动性较大的客流变化规律,实用性有待提升。不同于机器学习模型,深度学习预测模型含有多个隐层神经网络,能够较好的挖掘并传递客流量在深层次的关联性,比较有代表性的有门控循环神经网络(GRU)[7]、长短时记忆神经网络(LSTM)[8]等多种模型。当选择合适的输入步长时,会有更好的预测精度[9]。但是以上模型更多考虑了时间特征,缺乏对空间特征的学习。指数平滑法、极限学习机、GCN-GRU[10-11]、GCN-SBULSTM[12]等多种组合模型优化了对空间特征的学习能力。如Zhu等[13]使用知识融合单元(KF Cell)和图卷积网络,学习交通信息和外部因素之间的相关性的问题,具有良好的预测效果。

尽管这些研究从网络拓扑中提取了地铁站点之间的空间关系,但没有充分考虑到以站点为起讫点的空间特征影响以及站点周边的用地特征,还不能全面的提取和量化客流数据的空间关联性。同时,由于在使用组合模型进行预测时,其参数较难以确定,影响预测精度。一些优化算法被用于优化预测模型的参数,降低模型预测误差,如结合超图、粒子群优化算法[14-15]、蝙蝠算法优化[16]等。由于预测模型训练过程较为复杂,导致经典优化算法在寻找最优解时易陷入局部最优解且收敛速度较慢,使得模型在预测过程中运行效率较低,间接影响到预测结果的准确性和实时性。

为了解决上述问题,提出一种结构优化的图卷积门控循环神经网络地铁客流预测方法(SWO-AGGRU)。构建图卷积门控循环神经网络,根据地铁网络的不同客流特征,使用多个输入端进行地铁客流量时空关联性的学习。为了更好的提取空间特性和处理客流突变现象,将门控循环神经网络中的线性运算用图卷积运算替代,并将图卷积网络中的邻接静态矩阵根据特征学习更改为动态特征矩阵,同时构建注意力机制模块对不同客流特征的输出数据进行权重分配。考虑到预测模型的隐层结构参数设置会对预测结果产生较大影响,一种基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法被提出,用于进一步优化预测模型的隐层结构,以期能够进一步提升预测精度。

1 面向地铁网络的门控循环神经网络客流预测模型

城市地铁客流数据不仅在时间上存在关联,还受多起讫站点,空间距离、土地利用性质等因素影响,因而在一定范围内相关联的空间上,客流数据间也有着彼此的联系。该空间联系影响因素多且较微观,准确找寻其变化规律非常困难,因此,一种新的面向地铁网络的图卷积门控循环神经网络客流预测模型被提出。为了能够同时对多类数据进行处理,级联3个图卷积门控循环神经网络,同时学习地铁客流量的时序变化、多起讫地铁站点客流量变化以及地铁网络中相邻站点的历史客流量数据,

并利用注意力机制融合学习结果，找寻地铁客流量数据的时空特征，得出预测结果。面向地铁网络的图卷积门控循环神经网络客流预测模型框架示意图如图1所示。图1中GCN为线性运算，σ为Sigmoid激活函数，$\boldsymbol{H}^{t-1}$为$t-1$时刻的隐藏层状态，$\boldsymbol{x}_{t\gamma}$，$\boldsymbol{x}_{t\zeta}$，$\boldsymbol{x}_{t\lambda}$为$t$时刻隐藏节点的输出，$H^t$为传递给下一个节点的隐藏层状态。

模型中将地铁网络拓扑结构视为一个未加权有向图，各地铁站点视为一个节点，表示如下

$$G=(V,E,\boldsymbol{A}) \tag{1}$$

式中：V为顶点集，即为地铁网络中的各地铁站点，$V=\{V_1,V_2,\cdots,V_n\}$为顶点个数，即为研究范围内地铁站点的个数；E为边集；$\boldsymbol{A}\in\{0,1\}^{|V|\times|V|}$为节点连通性的邻接矩阵。

为了能够更好的学习地铁客流量的空间关联性，使用图卷积网络[17]学习地铁客流量空间相关性，其特征学习公式如下

$$\boldsymbol{B}=\boldsymbol{D}^{-1}\boldsymbol{A} \tag{2}$$

$$D_{ii}=\sum_j A_{ij} \tag{3}$$

$$\boldsymbol{X}_{*\mathrm{G}}g(\boldsymbol{\theta})=\sum_{i=0}^{k}(\boldsymbol{B})^i\boldsymbol{X}\boldsymbol{\theta}^i \tag{4}$$

式中：$\boldsymbol{B}$为规范化关系矩阵；$\boldsymbol{D}$为地铁站点的对角矩阵；$g(\theta)$为滤波器参数；$\boldsymbol{X}$为输入值；*G为图形卷积运算。

在时间为t时，图卷积运算中特征矩阵如下

$$\boldsymbol{V}_t\in\mathbb{R}^{V\times d} \tag{5}$$

式中：d为输入特征的尺寸（包括两个特征：地铁站点客流流出及流入，即d=2）。

$$\boldsymbol{H}^{l+1}=\sigma(D_{ii}^{-\frac{1}{2}}\boldsymbol{B}D_{ii}^{-\frac{1}{2}}\boldsymbol{H}^l\boldsymbol{W}^l) \tag{6}$$

式中：W^l为包含该层的参数。

$$f(X,\boldsymbol{A})=\sigma(\boldsymbol{B}\mathrm{ReLU}(\boldsymbol{B}X\boldsymbol{\omega}_0)\boldsymbol{\omega}_1) \tag{7}$$

式中：$\boldsymbol{\omega}_0\in\mathbb{R}^{P\times H}$为输入层到隐藏层的权重矩阵；$\boldsymbol{\omega}_1\in\mathbb{R}^{H\times T}$为从隐藏层到输出层的权重矩阵；$P$为特征矩阵的长度；$\boldsymbol{H}$为隐藏层的层数；$f(X,\boldsymbol{A})\in\mathbb{R}^{N\times T}$为输出，$T$为模型预测长度；ReLU为修正线性单位；$l$为层数。

由于GCN中的邻接矩阵是静态的，因此不能够较好的学习地铁客流网络的空间特征。为了能够动态学习地铁网络中客流量的空间特征，需要对其进行改进，通过网络训练更新迭代关系矩阵，动态学习空间特征。其中关系矩阵计算如下

$$\boldsymbol{A}^m=\varphi(\boldsymbol{A}^{m-l},\cdots,\boldsymbol{A}^{m-1}) \tag{8}$$

$$\boldsymbol{A}^{m+1}=\phi^m(\boldsymbol{A}^m) \tag{9}$$

$$[\boldsymbol{X}_{t+1},\cdots,\boldsymbol{X}_{t+T}]=f(G;(\boldsymbol{X}_{t-n},\cdots,\boldsymbol{X}_{t-1},\boldsymbol{X}_t)) \tag{10}$$

$$\boldsymbol{Z}^{m+1}=\boldsymbol{Z}^m{}_{*\mathrm{G}}g_\theta^m=\sum_{i=0}^{k}(\boldsymbol{A}^m)^i\boldsymbol{Z}^m\boldsymbol{\theta}_i^m \tag{11}$$

式中：$\boldsymbol{A}^m$为站点之间的第m层关系；l为$\boldsymbol{A}^{m+l}$的所在层数；$\boldsymbol{A}^0$随着整个网络的训练进行迭代更新，其余层数根据映射函数φ进行更新；ϕ^m为$m+1$层上的映射函数；$\boldsymbol{\theta}_i^m$为卷积参数；n为预测时间粒度个数；$\boldsymbol{Z}^m$为图卷积上第m层站点的特征；k为扩散网络的层数。

为了能够更好地学习地铁客流量的空间关联

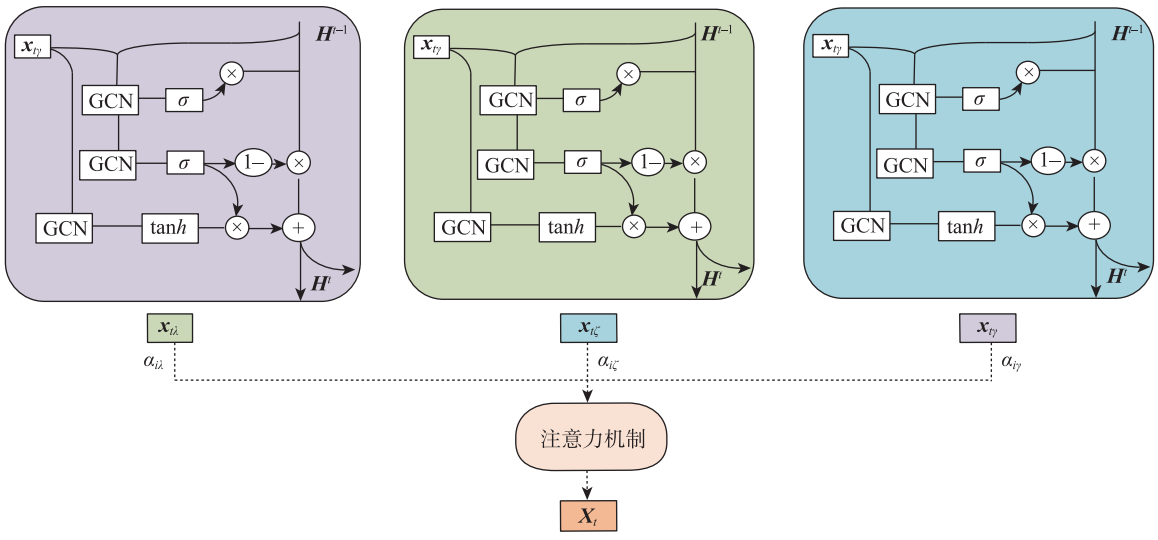


图1　图卷积门控循环神经网络模型框架示意图

Fig. 1　Schematic framework of convolutional gated recurrent neural network model

性，本模型在GRU深度学习框架的基础上，使用图卷积运算代替GRU中线性运算来增强算法的空间信息处理能力，以学习目标地铁站客流量的时序变化、多起讫地铁站点客流量以及地铁网络中相邻站点的历史客流量数据。模型中使用3个输入端以学习以上3类数据，每个输入端通过相应的地铁客流量数据学习得出一个输出结果，最终使用注意力机制对以上3类数据进行权重量化，更为客观的描述以上3类数据对客流量的影响程度，从而实现多特征融合。图卷积门控循环网络模型（GGRU）中时空特征的学习与传递公式如下

$$\boldsymbol{r}^t=\sigma(\beta_{\mathrm{r}}*\boldsymbol{W}_z\cdot[\boldsymbol{h}^t,\boldsymbol{H}^{t-1}]+\boldsymbol{b}_{\mathrm{r}}) \tag{12}$$

$$\boldsymbol{u}^t=\sigma(\beta_{\mathrm{u}}*\boldsymbol{W}_z\cdot[\boldsymbol{h}^t,\boldsymbol{H}^{t-1}]+\boldsymbol{b}_{\mathrm{u}}) \tag{13}$$

$$\boldsymbol{c}^t=\tan h(\beta_{\mathrm{c}}*\boldsymbol{W}_z\cdot[\boldsymbol{h}^t,(\boldsymbol{r}^t\circ\boldsymbol{H}^{t-1}+\boldsymbol{b}_{\mathrm{c}})]) \tag{14}$$

$$\boldsymbol{H}^t=\boldsymbol{u}^t\circ\boldsymbol{H}^{t-1}+(1-\boldsymbol{u}^t)\circ\boldsymbol{c}^t \tag{15}$$

式中：$\boldsymbol{h}^t$和$\boldsymbol{H}^t$分别为在时间为t时GCN模块和GRU模块的输出；$\circ$为哈达玛积；σ为Sigmoid激活函数；重置门$\boldsymbol{r}^t$用于信息更替，对不必要的信息进行遗忘；更新门$\boldsymbol{u}^t$为可以控制时间步长t处卷积GRU的输出；β_{r}，β_{u}，β_{c}均为相应的滤波器参数；$\boldsymbol{b}_{\mathrm{r}}$，$\boldsymbol{b}_{\mathrm{u}}$和$\boldsymbol{b}_{\mathrm{c}}$均为偏差。

该图卷积门递归神经网络最终输出为$\boldsymbol{x}_{t\gamma}$，$\boldsymbol{x}_{t\zeta}$，$\boldsymbol{x}_{t\lambda}$，使用3个模型输出的结果组成向量$\{\overline{\boldsymbol{x}_{t\gamma}}\}$，$\{\overline{\boldsymbol{x}_{t\zeta}}\}$，$\{\overline{\boldsymbol{x}_{t\lambda}}\}$，将其作为注意力机制模块中输入值，如下

$$\boldsymbol{S}_i=\overline{\boldsymbol{x}_{t\gamma}}\times(\overline{\boldsymbol{x}_{i\gamma}})^{\mathrm{T}}+\overline{\boldsymbol{x}_{t\zeta}}\times(\overline{\boldsymbol{x}_{i\zeta}})^{\mathrm{T}}+\overline{\boldsymbol{x}_{t\lambda}}\times(\overline{\boldsymbol{x}_{i\lambda}})^{\mathrm{T}} \tag{16}$$

$$\alpha_i=\mathrm{soft\,max}(\boldsymbol{S}_i)=\frac{\exp(\boldsymbol{S}_i)}{\sum_{i=1}^{N}\exp(\boldsymbol{S}_i)} \tag{17}$$

$$Attention_k=\sum_{i=P}^{k=t-1}\alpha_i\boldsymbol{x}_{i\gamma}+\sum_{i=P}^{k=t-1}\alpha_i\boldsymbol{x}_{i\zeta}+\sum_{i=P}^{k=t-1}\alpha_i\boldsymbol{x}_{i\lambda} \tag{18}$$

$$\hat{\boldsymbol{x}}_i=\sum_{k=1}^{t}Attention_k\boldsymbol{x}_{i\gamma} \tag{19}$$

式中：$\boldsymbol{S}_i$为权重矩阵；α_i为3类数据各占权重；$k\in[P,t-1]$，P为当前时间步长。

2 结构优化的图卷积门控循环神经网络地铁客流预测模型

在图卷积门控循环神经网络地铁客流预测模型中，不同隐层结构维度的选择会对模型的预测精度产生较大的影响，导致同一框架而不同参数选择下模型的预测性能相差较大。为了能够进一步提升该模型的预测精度，提出一种基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法，对客流预测模型的隐层结构维度进行优化，使其能够动态合理的选择最优的隐层结构。

2.1 一种基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法

模型初始参数的选取影响优化算法寻优时的收敛速度，从而在一定程度上影响模型的运行效率。在传统的蜘蛛黄蜂算法[18]中，初始化种群采用随机生成的方法，耗时长，随机性大。使用Tent混沌映射代替随机参数生成蜘蛛黄蜂的初始化种群，使得算法能够在搜索空间中生成具有良好多样性的初始解，以提高算法的全局搜索能力、收敛速度和收敛精度。采用式（20）所示的Tent混沌映射来产生初代蜘蛛黄蜂群体，如下

$$\begin{cases}SW_{i,n+1}=\alpha SW_{i,n}, & 0\leqslant SW_{i,n}\leqslant\dfrac{1}{2}\\ SW_{i,n+1}=\alpha(1-SW_{i,n}), & \dfrac{1}{2}<SW_{i,n}\leqslant 1\end{cases} \tag{20}$$

式中：α为控制混沌形态的参数；i为混沌变量序号；n为种群序号。通过给式（20）赋D个0~1范围内的初值，可得到D个混沌变量$SW_{1,1}$，$SW_{1,2}$，…，$SW_{1,D}$。

按照公式依次迭代，即可得到含有N个蜘蛛黄蜂个体的初始群体，如下

$$\boldsymbol{SW}_{\mathrm{pop}}=\begin{bmatrix}SW_{1,1} & SW_{1,2} & \cdots & SW_{1,D}\\ SW_{2,1} & SW_{2,2} & \cdots & SW_{2,D}\\ \vdots & \vdots & \ddots & \vdots\\ SW_{N,1} & SW_{N,2} & \cdots & SW_{N,D}\end{bmatrix} \tag{21}$$

在搜索阶段，初代雌性黄蜂需寻找蜘蛛以喂养幼虫。由于其搜索空间有限且易陷入局部最优，因此将莱维飞行扰动策略用于改进蜘蛛黄蜂优化算法。通过莱维飞行扰动策略产生随机步长，提高其全局寻优能力及收敛速度，在一定程度上解决传统蜘蛛黄蜂优化算法收敛速度较慢且易陷入局部最优的问题。采用莱维飞行扰动策略生成随机步长，莱维飞行服从Levy分布，如式（22）所示。扰动步长S_{p}的计算公式如下

$$Levy=t^{-\lambda},1<\lambda\leqslant 3 \tag{22}$$

$$S_i=\frac{u}{|v|^{1/\beta}} \tag{23}$$

$$u\sim N(0,\sigma_u^2),v\sim N(0,\sigma_v^2) \tag{24}$$

$$\sigma_u=\left\{\frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}}\right\}^{\frac{1}{\beta}},\ \sigma_v=1 \tag{25}$$

$$S_{\mathrm{P}}=\tau S_i x^j \tag{26}$$

式中：S_i 为利用 Mantegna 模拟的随机搜寻步长，$i\in\{1,2,\cdots,n\}$，$j\in\{1,2,\cdots,D\}$；u 和 v 均为正态分布；$\beta=1.5$；Γ 为标准 Gamma 函数；x^j 为第 j 个雌黄蜂的位置。雌黄蜂搜索位置步长扰动 $v_{i,j}$ 定义为

$$v_{i,j}=v_{i,j}S_{\mathrm{P}} \tag{27}$$

2.2　结构优化的门控循环神经网络地铁客流预测实现步骤

利用城市交通IC卡采集所需要的刷卡时间、刷卡站点等相关客流量数据信息，对数据进行分析和预处理。根据模型输入数据要求，将地铁网络中的客流量数据分为以下3类：历史交通客流数据、临近周边重要站点的历史客流数据以及历史数据中以预测站点为起讫点的客流量数据，而后进行地铁客流量预测。利用基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法优化模型中的参数，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）与平均绝对百分误差（MAPE）之和，如下

$$Loss=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i-x_i)^2}+\frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i-\hat{x}_i}{x_i}\right| \tag{28}$$

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

提出的结构优化的图卷积门控循环神经网络地铁客流预测主要实现步骤如下（图2）。

步骤1：采集地铁网络中的客流样本数据，对数据进行预处理，去除数据中的噪声以及无关数据。将所得地铁客流量数据按5 min聚合，使其符合预测模型输入数据格式要求。

步骤2：构建图卷积门控循环神经网络，设置3个图卷积门控神经网络输入层、图卷积门控神经网络中间层以及注意力策略输出层，初步设置显隐层结构，以及相关参数指标。

步骤3：使用历史 n 个时间序列数据作为输入，并使用图卷积网络捕获城市道路网络的拓扑结构，进行地铁客流量空间信息学习，通过网络迭代动态学习客流量的空间特征。

步骤4：将时间序列输入门控循环单元模型，通

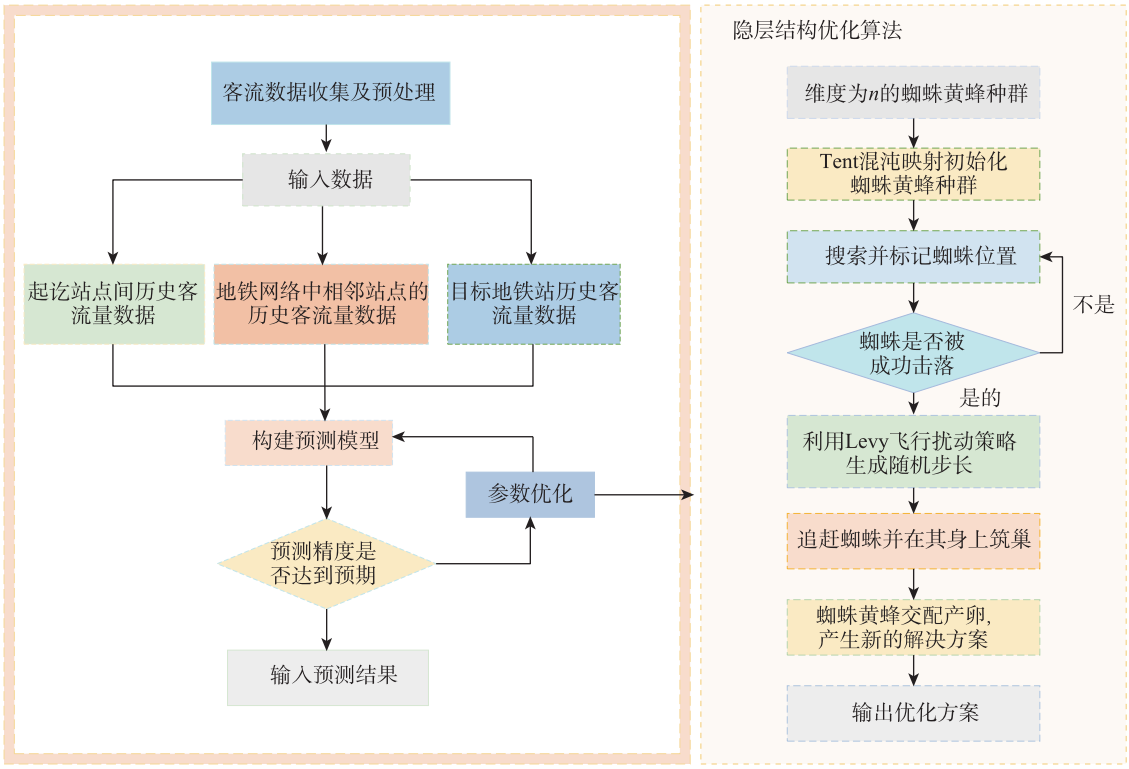


图2　结构优化的图卷积门控循环神经网络地铁客流预测流程图

Fig. 2　Flowchart of metro passenger flow prediction by structurally optimised graph convolutional gated recurrent neural networks

过单元间的信息传输获得动态变化，以捕获时间特征，通过完全连接层得到结果。学习3类不同数据并使用注意力机制量化3类数据对客流量影响的权重大小，根据该权重分配进行客流量预测。

步骤5：对图卷积门控神经网络进行结构与优化，基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法，根据式(20)~式(27)对算法进行优化，以期寻找到最优的超参数。

步骤6：根据优化过的超参数，建立最终面向地铁网络的结构优化的门控单元卷积神经网络地铁客流预测模型，利用训练好的地铁客流预测模型对实际客流量数据进行预测，以验证所提模型的预测效果。

3 实验结果分析

3.1 研究范围

采用上海地铁网络2022年4月份客流量数据，上海市共有地铁线路20条，地铁站点508座。以人民广场站为例，人民广场站为1号线、2号线、8号线换乘站点。实验样本采集间隔为5 min，主要为人民广场站点的历史交通客流数据、邻近周边重要站点的历史客流数据以及历史数据中以人民广场站点为起讫点的客流量数据。选取一个月的地铁客流数据作为3个输入端的训练样本，选取2022年4月26日(周末)、4月27日(工作日)和4月28日(工作日)为验证集，用于验证提出的预测模型的预测精确度。人民广场站周边节点分布如图3所示。

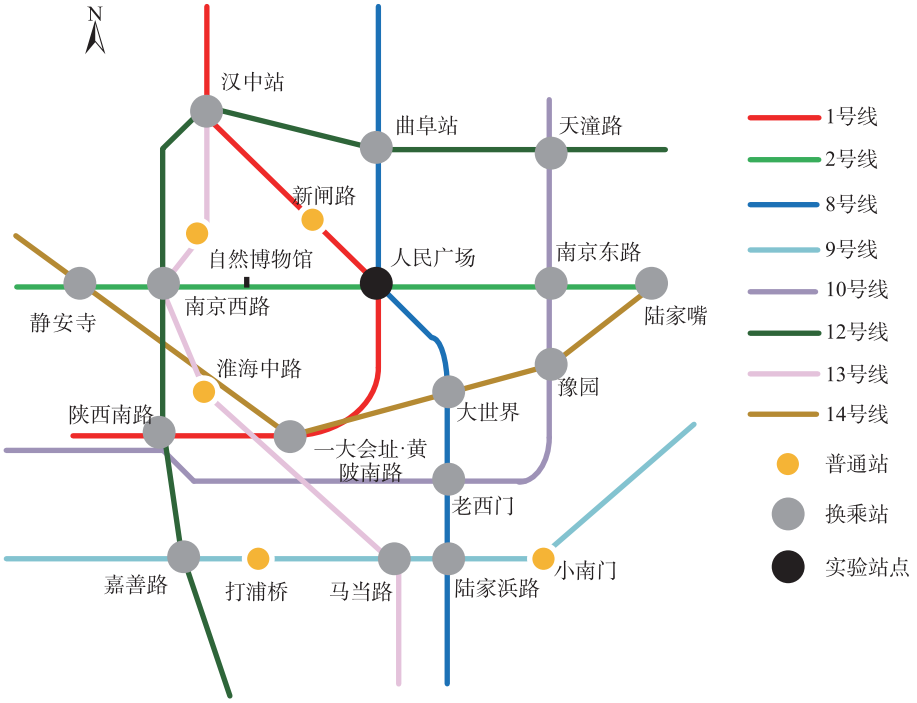


图3 人民广场站点周边节点分布图

Fig. 3 Node distribution around the People's Square Station

3.2 指标选取

为了能够更为客观的评价地铁交通客流预测的准确性，指标选取平均绝对误差(MAE)[19]、均方根误差(RMSE)[20]、平均绝对百分误差(MAPE)以及相关系数R^2，公式如下

$$MAE=\frac{1}{N}\sum_{i=1}^{N}|\hat{x}_i-x_i|\times 100\% \tag{29}$$

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i-x_i)^2}\times 100\% \tag{30}$$

$$MAPE=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{x}_i-\bar{x}}{x_i}\right|\times 100\% \tag{31}$$

$$R^2=\frac{\sum_{i=1}^{N}(\hat{x}_i-\bar{x})^2}{\sum_{i=1}^{N}(x_i-\bar{x})^2} \tag{32}$$

式中：i为样本数量；x_i为时刻i的真实值；$\hat{x}_i$为时刻i的预测值；$\bar{x}$为客流量真实值的平均值；N为预测数据的总量。

3.3　基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法与经典优化算法对比

为了更准确地评估基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法对模型预测精确度的影响，将该方法与使用萤火虫算法优化预测模型、使用蝙蝠算法优化预测模型进行对比。分别对工作日和周末进行预测，其指标选取见3.2节，工作日不同时间段各指标对比如图4所示，周末不同时间段各指标对比如图5所示。

由图4和图5可以得知，SWO可以通过20次的迭代达到最佳效果，而FA和BA则是需要45次和70次的迭代达到效果。SWO的收敛速度和优化性能均优于经典的优化算法。由此可见，SWO对模型的预测性能起到了积极作用。

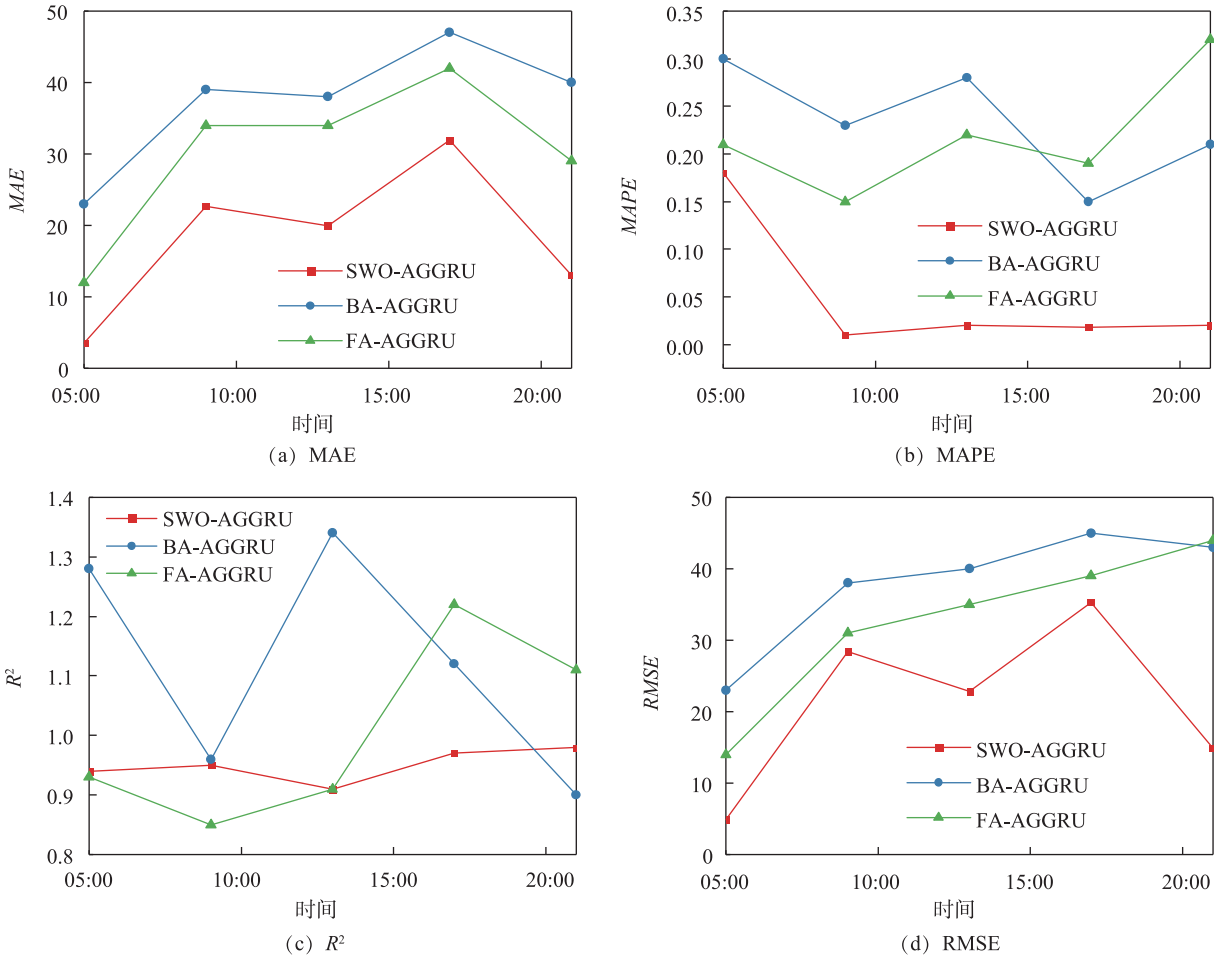


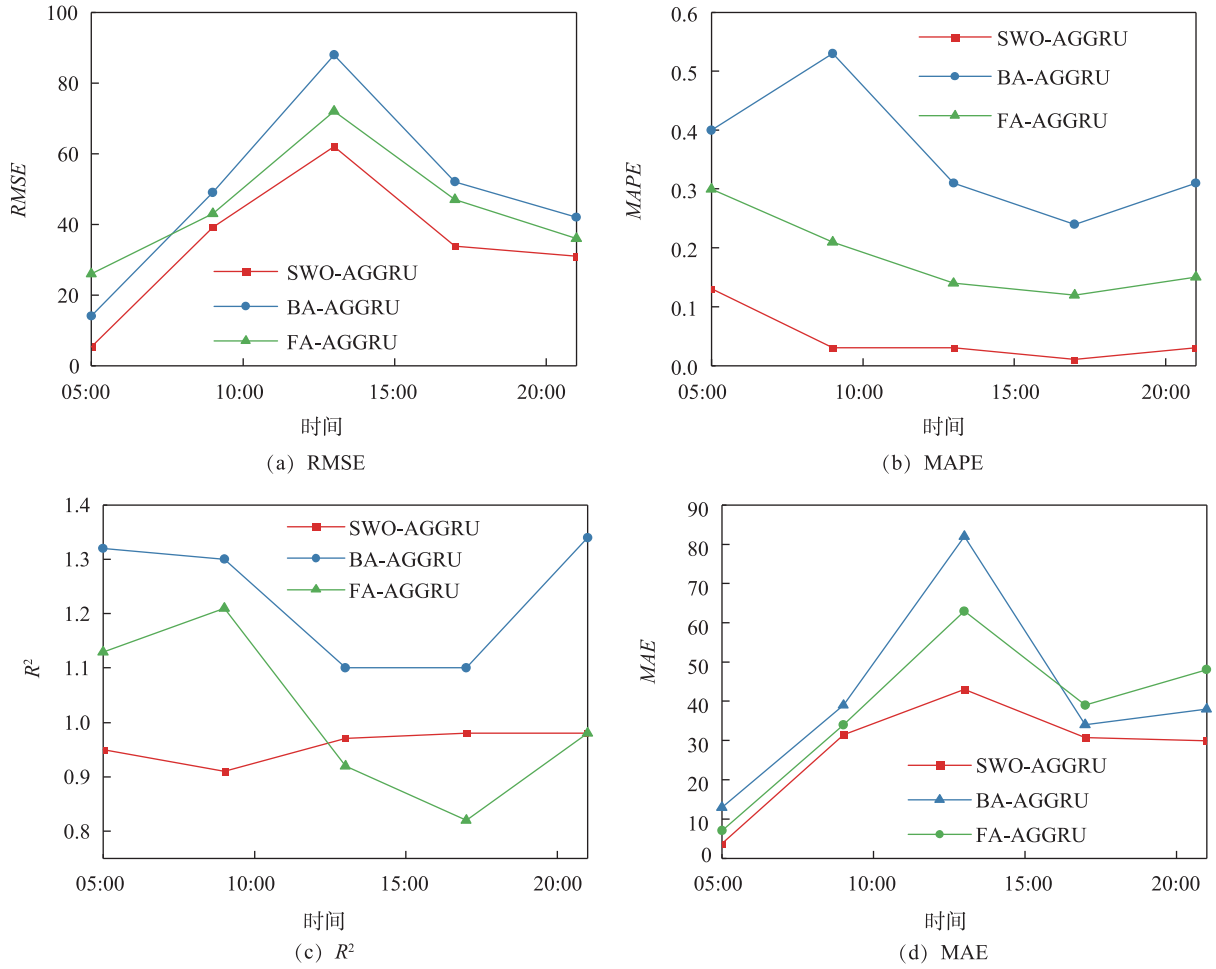
图4　工作日不同时段各指标对比

Fig. 4　Comparison of indicators in different periods of working days

3.4　结构优化的门控循环神经网络对预测性能的影响

为了能够更好地学习地铁客流量的时空特征，同时动态优化预测模型的参数设置，提出一种基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化的AGGRU预测模型结构动态优化方法。为评估基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法、图卷积网络以及注意力机制对模型预测精确度的影响，本次实验将SWO-AGGRU模型、AGGRU模型以及传统GRU模型进行对比，以验证其对预测精度的影响。图6为工作日的预测结果与周末预测结果对比，将提出的SWO-AGGRU与经典深度学习预测模型[21-22]以及相关组合预测模型[23]进行对比，具体指标选取如3.2所示。各指标对比如表1和表2所示。

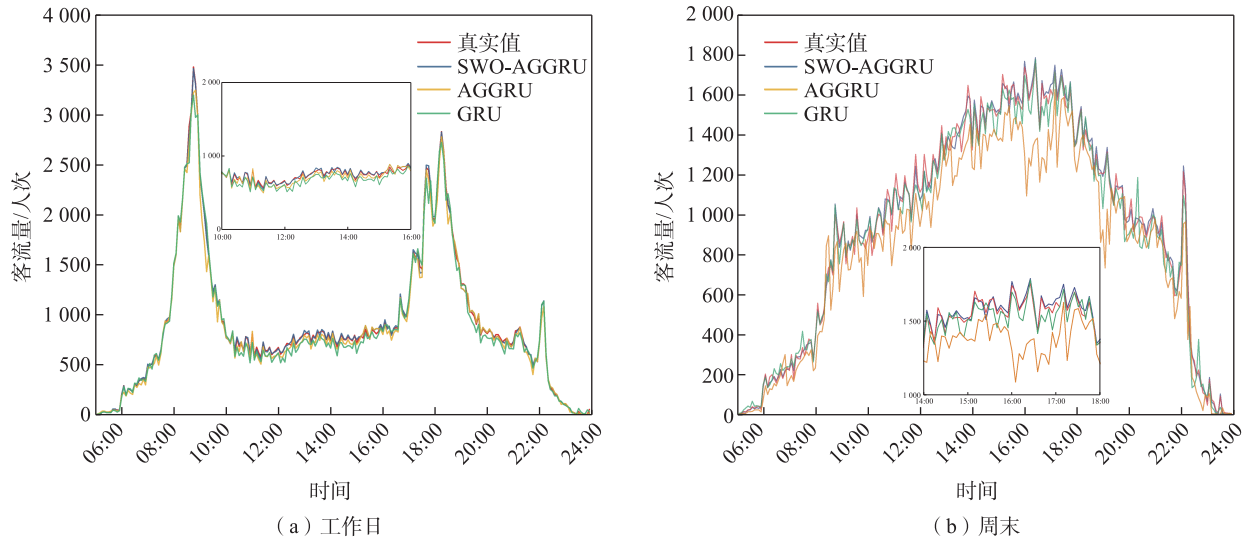
由表1、表2和图6可以得知，在工作日时间段与



（a）RMSE （b）MAPE （c）R^2 （d）MAE

图5 周末不同时段各指标对比

Fig. 5 Comparison of indicators in different periods on weekends



（a）工作日 （b）周末

图6 预测结果对比

Fig. 6 Comparison of prediction results

表1　工作日预测结果对比

Tab.1　Comparison of prediction results on working days

评价指标	SWO-AGGRU	AGGRU	GRU	多元集成CNN-LSTM	CNN-BILSTM	STL-GRU
RMSE	20	61	66	152	214	41
MAE	14	38	43	107	160	26
R^2	0.98	0.94	0.96	0.83	0.77	0.96
MAPE	0.04	0.11	0.37	0.70	0.22	0.11

表2　周末预测结果对比

Tab.2　Comparison of prediction results on weekends

评价指标	SWO-AGGRU	AGGRU	GRU	多元集成CNN-LSTM	CNN-BILSTM	STL-GRU
RMSE	33	40	159	147	139	67
MAE	26	38	123	109	122	46
R^2	0.98	0.95	0.85	0.87	0.94	0.94
MAPE	0.12	0.17	0.27	0.78	0.31	0.15

周末时间段,SWO-AGGRU模型预测精度均优于AGGRU预测模型和传统GRU预测模型。与其他经典深度学习预测模型以及相关改进模型相比,SWO-AGGRU在工作日和周末的预测精度在一定程度上均有所提高。SWO-AGGRU能够兼顾地铁客流量的时空变化、起讫地铁站点客流量以及地铁网络中相邻站点的历史客流量3类数据对客流量预测的影响,同时考虑客流量在时间和空间的相关性,模型精确度优于经典的深度学习预测模型。

4　结论

1）将图卷积网络学习空间关联性融入到GRU的线性运算中,使得模型能够同时学习地铁客流量的时空相关性。采用基于Tent混沌映射和莱维飞行扰动策略的蜘蛛黄蜂优化算法对预测模型的参数进行优化,在一定程度上提高模型的预测精度。

2）考虑地铁客流量的时空变化、起讫地铁站点客流量以及地铁网络中相邻站点的历史客流量数据对实验站点的影响,通过构建可同时训练3种与预测节点相关的交通流数据的深度信任网络模型来增强预测的时空关联性。

3）利用实际地铁客流数据进行验证,结果表明与经典深度学习预测模型相比,SWO-AGGRU具有更高的预测精确度和更短的运行时间。

4）通过充分学习客流量的时空特征,提高模型的预测性能以合理地进行列车的调度,从而提高地铁的运输效率与服务质量,降低地铁站点的拥挤程度,为后续地铁的运营管理提供帮助。

参考文献：

[1] STEPHANEDES Y J, MICHALOPOULOS P G, PLUM R A. Improved estimation of traffic flow for real- time control[J]. Transportation Research Record Journal of the Transportation Research Board, 1980, 795: 28-39.

[2] AHMED M S, COOK A R. Analysis of freeway traffic time- series data by using box- Jenkins techniques[J]. Transportation Research Record, 1979, 722: 1-9.

[3] 杨兆升, 朱中. 基于卡尔曼滤波理论的交通流量实时预测模型[J]. 中国公路学报, 1999, 12(3): 63-67.
YANG Z S, ZHU Z. A realtime traffic volume prediction model based on the Kalman filtering theory[J]. China Journal of Highway and Transport, 1999, 12(3): 63-67.

[4] 白丽. 城市轨道交通常态与非常态短期客流预测方法研究[J]. 交通运输系统工程与信息, 2017, 17(1): 127-135.
BAI L. Urban rail transit normal and abnormal short-term passenger flow forecasting method[J]. Journal of Transportation Systems Engineering and Information Technology, 2017, 17(1): 127-135.

[5] CHENG G, HE C L. Analysis of bus travel characteristics and predictions of elderly passenger flow based on smart card data[J]. Electronic Research Archive, 2022, 30(12): 4256-4276.

[6] YANG H F, DILLON T S, CHANG E, et al. Optimized configuration of exponential smoothing and extreme learning machine for traffic flow forecasting[J]. IEEE Transactions on Industrial Informatics, 2019, 15(1): 23-34.

[7] DAI G W, MA C X, XU X C. Short-term traffic flow prediction method for urban road sections based on space-time analysis and GRU[J]. IEEE Access, 2019, 7: 143025-

143035.

[8] SHU W N, CAI K, XIONG N N. A short-term traffic flow prediction model based on an improved gate recurrent unit neural network[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(9): 16654-16665.

[9] 王爽, 黄海超, 石宝存, 等. 基于自相关分析的交通流预测输入步长研究[J]. 华东交通大学学报, 2022, 39(5): 78-85.
WANG S, HUANG H C, SHI B C, et al. Research on time step of traffic flow prediction based on autocorrelation analysis[J]. Journal of East China Jiaotong University, 2022, 39(5): 78-85.

[10] HUANG X, YE Y, YANG X, et al. Multistep coupled graph convolution with temporal-attention for traffic flow prediction[J]. IEEE Access, 2022, 10: 48179-48192.

[11] 徐丽, 符祥远, 李浩然. 基于门控卷积的时空交通流预测模型[J]. 计算机应用, 2023, 43(9): 2760-2765.
XU L, FU X Y, LI H R. Spatial-temporal traffic flow prediction model based on gated convolution[J]. Journal of Computer Applications. 2023, 43(9): 2760-2765.

[12] CHEN P F, FU X D, WANG X. A graph convolutional stacked bidirectional unidirectional-LSTM neural network for metro ridership prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 6950-6962.

[13] ZHU J W, HAN X, DENG H H, et al. KST-GCN: a knowledge-driven spatial-temporal graph convolutional network for traffic forecasting[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(9): 15055-15065.

[14] 李万, 冯芬玲, 蒋琦玮. 改进粒子群算法优化LSTM神经网络的铁路客运量预测[J]. 铁道科学与工程学报, 2018, 15(12): 3274-3280.
LI W, FENG F L, JIANG Q W. Prediction for railway passenger volume based on modified PSO optimized LSTM neural network[J]. Journal of Railway Science and Engineering, 2018, 15(12): 3274-3280.

[15] 付建广, 尤斌, 林毅, 等. 基于粒子群优化与宽度学习系统的地铁客流预测模型[J]. 城市轨道交通研究, 2023, 26(5): 23-26.
FU J G, YOU B, LIN Y, et al. Metro passenger volume prediction model based on particle swarm optimization and broad learning system[J]. Urban Mass Transit, 2023, 26(5): 23-26.

[16] 段中兴, 温倩, 周孟, 等. 基于改进蝙蝠算法优化LSTM网络的短时客流预测[J]. 铁道科学与工程学报, 2021, 18(11): 2833-2840.
DUAN Z X, WEN Q, ZHOU M, et al. Short-term passenger flow prediction based on improved bat algorithm to optimize LSTM network[J]. Journal of Railway Science and Engineering, 2021, 18(11): 2833-2840.

[17] ZHAO L, SONG Y J, ZHANG C, et al. T-GCN: a temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(9): 3848-3858.

[18] ABDEL-BASSET M, MOHAMED R, JAMEEL M, et al. Spider wasp optimizer: a novel meta-heuristic optimization algorithm[J]. Artificial Intelligence Review, 2023, 56 (10): 11675-11738.

[19] 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.
YAO E J, ZHOU W H, ZHANG Y S. Real-time forecast of entrance and exit passenger flow for newly opened station of urban rail transit at initial stage[J]. China Railway Science, 2018, 39(2): 119-127.

[20] 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.
LIANG Q S, XU X Y, LIU L Q. Data-driven short-term passenger flow prediction model for urban rail transit[J]. China Railway Science, 2020, 41(4): 153-162.

[21] 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.

[22] 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.

[23] 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.
ZHAO J D, ZHU D, LIU J X. Metro transfer passenger flow prediction based on STL-GRU[J]. Journal of South China University of Technology(Natural Science Edition), 2022, 50(5): 22-31.



通信作者:张阳(1983—),男,教授,博士,研究方向为铁路客流预测。E-mail:zhang_yang1983@163.com。

(责任编辑:吴海燕)