高子宸, 牛学军 . 基于 STL-GRU-Transformer 的短时交通流量预测[J]. 智能计算机与应用,2026,16(6):83-89. DOI:10.20169/j. issn. 2095-2163. 24101803

基于 STL-GRU-Transformer 的短时交通流量预测

高子宸, 牛学军
(中国人民公安大学 交通管理学院, 北京 100038)

摘 要: 随着智能交通系统的发展,如何更准确地预测交通流量这一课题的研究不断深化。针对道路上的短时交通流量的周期性与非线性特征,本文基于 STL-GRU-Transformer 的组合模型提出一种预测短时流量的方法,具体是基于季节性和趋势分解使用局部加权回归 STL 分解方法,将交通流量分解成季节因子序列、趋势因子序列及残差因子序列,使交通流量序列能够更好地被预测模型捕捉其特性和未来趋势。通过结合因子序列特征与模型特性,使用结构简单、计算效率高的门控循环单元 GRU 对 STL 分解出来的季节因子和趋势因子进行预测,同时借助适合处理复杂残差信号的 Transformer 模型对残差因子进行预测。结果表明:STL 分解的方法有助提升预测精度;Transformer 模型在预测残差的过程中展现出优秀性能,大幅提升了模型精度;基于 STL-GRU-Transformer 的组合预测模型表现出优异的预测效果。该组合模型对流量数据集的预测结果的平均绝对误差 MAE 达到 25. 67,相较于单模型预测精度至少提高 5. 98,效果提升 18. 89%。

关键词: 交通流量预测; 神经网络; 集成模型; 深度学习; STL 分解; 门控循环单元; Transformer 模型
中图分类号: U495 **文献标志码**: A **文章编号**: 2095-2163(2026)06-0083-07

Short-term traffic flow prediction based on STL-GRU-Transformer

GAO Zichen, NIU Xuejun
(School of Traffic Management, People's Public Security University of China, Beijing 100038, China)

Abstract: With the development of intelligent transportation systems, improving the accuracy of traffic flow prediction has become a deepening research topic. Considering the periodic and nonlinear characteristics of short-term traffic flow on roads, this paper proposes a method for short-term traffic flow prediction based on a combined STL-GRU-Transformer model. By using seasonal-trend decomposition through locally weighted regression (STL), traffic flow data is decomposed into seasonal, trend, and residual components, allowing the prediction model to better capture the characteristics and future trends of the traffic flow. The GRU model, known for its simple structure and computational efficiency, is used to predict the seasonal and trend components, while the Transformer model, adept at handling complex residual signals, is employed to predict the residual component. The results show that the STL decomposition method significantly improves prediction accuracy, and the Transformer model demonstrates excellent performance in predicting residuals, substantially enhancing the overall model accuracy. The combined STL-GRU-Transformer model achieves a Mean Absolute Error (MAE) of 25. 67 on the traffic dataset, improving prediction accuracy by at least 5. 98 in comparison with single models, which represents an 18. 89% improvement.

Key words: traffic flow forecast; neural network; integration model; deep learning; STL decomposition; gated cycle unit; Transformer model

0 引 言

当代经济高速发展,科学带来的技术革新影响着所有人的生活,交通更是与每一个人息息相关。随着智能交通系统[1]的发展,能否准确预测交通流量并达到出行快捷且安全的目标变得尤为重要。良好的预测效果对道路资源的优化配置和交通应急管理都具有重要意义。在现有的交通流量预测方法中,大致可分为传统统计模型方法和基于机器学习的预测方法。

传统统计学模型中,Stephanedes 等学者[2]于 1980 年首先将历史平均模型(History Average,HA)

作者简介: 高子宸(1998—),男,硕士研究生。主要研究方向:交通管理,大数据技术。
通信作者: 牛学军(1974—),男,博士,副教授。主要研究方向:交通安全。Email:gadxnxj@ sina. com。

收稿日期: 2024-10-18

应用于城市交通控制系统,王晓全等学者[3]结合广义自回归条件异方差、李洪娟等学者[4]通过建立灰色系统模型进行预测,Kim 等学者提出了 ARIMA 模型来预测高速公路流量预测[5],Vythoulkas[6]提出了基于卡尔曼滤波理论的交通流量预测模型,周伟[7]使用组合时间序列分解技术对短时交通流量进行预测。在基于机器学习的预测方法中,循环神经网络(Recurrent Neural Network,RNN)[8]因其短期记忆,可以处理顺序数据并识别历史数据中的模式,也能对不同长度的输入数据进行处理,但同时也存在梯度下降消失的问题。Hochreiter 等学者[9]于 1997 年提出长短期记忆神经网络(Long Short - Term Memory,LSTM)。LSTM 是一种特殊类型的 RNN,解决了 RNN 梯度消失的问题,但由于结构复杂,其计算成本更高,从而导致训练时间更长。Chung 等学者[10]提出了 RNN 的新变体门控循环单元模型(Gated Recurrent Unit,GRU),解决了传统 RNN 中的梯度消失和梯度爆炸问题。同 LSTM 相比,其拥有较低的计算复杂度和良好的性能。Transformer 模型由 Vaswani 等学者[11]于 2017 年提出,是一种基于注意力机制的神经网络架构。在时间序列预测领域,通过完全依赖自注意力机制来捕捉输入序列中的长距离依赖关系,克服了 RNN 和 LSTM 在处理长序列时的局限性。在近几年的交通流预测模型研究中,组合模型被广泛用于预测。陈宇等学者[12]通过结合卡尔曼滤波与 LSTM 自编码器的递归框架预测高速公路实时交通流量,王莉莉等学者[13]提出 GA、RBF 和改进 Cao 方法结合、周海赟等学者[14]提出 SSARX-NARX 模型、闫少华等学者[15]提出小波优化 GRU-ARMA 模型、王宁等学者[16]提出 PSO-LSTM 模型、王代君等学者[17]提出 Bayes-ARIMA 模型、徐厚生等学者[18]提出的 Bi-GRU-Atten 模型,都是作为组合模型去预测。也有用同种模型堆叠预测或多次分解,如廉清云等学者[19]堆叠 LSTM,陈昆等学者[20]对时间序列二次分解再组合。文献中的预测模型多是运用模型间的组合,其核心是传统统计学方法与循环神经网络的优质结合,对交通流预测准确率的提升效果为 4%~12%。采用这种方法在一定程度上提高了预测准确性,但仍存在长期依赖性弱、模型泛化能力不足等问题,这类模型鲜少采用注意力模型对数据进行挖掘,对数据内部的深层逻辑规律挖掘度不够。本文据此提出一种基于 STL-GRU-Transformer 的短时交通流量预测模型,通过 STL 分解,将交通流量数据分解为趋势、季节和残差因子,即分离成规律性数据与残差数据,采用 GRU 和 Transformer 模型分别预测,通过 Transformer 模型对残差数据做进一步挖掘,并结合小波阈值重构整合预测结果,最终达到提高预测的泛化能力和鲁棒性的目的。本研究亮点在于提出了一种新的模型组合方法,并通过实验验证其优越性。

1 基本理念

1.1 STL 分解

STL 是一种时间序列分解方法,用于将时间序列分解为趋势因子、季节因子和残差因子三个部分。具体采用了局部回归(LOESS)方法,通过加权移动平均来估计每个数据点的值,从而实现对时间序列数据的平滑处理和拟合。STL 的优点在于可以灵活处理数据中的非线性趋势,可解释性和适应性强,适用于各种类型的时间序列数据。其计算公式为:

$$Y_t = T_t + S_t + R_t \tag{1}$$

其中,Y_t 表示 t 时刻的流量值;T_t、S_t、R_t 分别表示该时刻的趋势因子值、季节因子值和残差因子值。

1.2 GRU 模型

GRU 是 RNN 模型的一种变体。同经典的 LSTM 相比,LSTM 通过引入记忆单元和门机制有效解决了 RNN 梯度消失和梯度爆炸的问题,但 LSTM 的结构复杂,计算量较大,而 GRU 通过引入更新门和重置门,在解决了这些问题的同时保持了较低的计算复杂度。作为 LSTM 的简化版本,GRU 通过更新门和重置门简化结构,仍能有效捕捉长程依赖关系。GRU 的结构如图 1 所示。

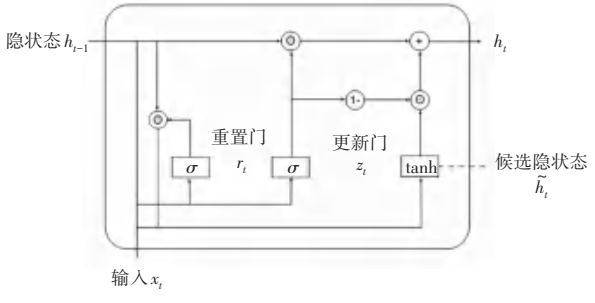


图 1 门控循环单元结构图

Fig. 1 Diagram of Gated Recurrent Unit structure

GRU 完全依赖隐藏状态作为记忆,因此有着更简单的架构。其更新门决定了前一时间步的状态在当前时间步中保留的信息,而重置门则控制前一时间步的状态在生成候选激活时产生的影响。以下是 GRU 的主要公式。

更新门 z_t 的作用是控制当前输入 h_t 和前一时刻的隐藏状态 h_{t-1} 之间的混合程度。其公式为：

$$z_t = \sigma(\boldsymbol{W}_z \cdot [h_{t-1}, x_t]) \tag{2}$$

其中，z_t 表示更新门；$\boldsymbol{W}_z$ 表示权重矩阵；h_{t-1} 表示前一时间步的隐藏状态；x_t 表示当前时间步的输入；σ 表示 Sigmoid 函数。

重置门 r_t 则决定丢弃上一时间步的状态信息数据。其公式为：

$$r_t = \sigma(\boldsymbol{W}_r \cdot [h_{t-1}, x_t]) \tag{3}$$

其中，r_t 表示重置门，$\boldsymbol{W}_r$ 表示权重矩阵。

当前时间步的候选隐藏状态计算公式为：

$$\tilde{h}_t = \tanh(\boldsymbol{W} \cdot [r_t \odot h_{t-1}, x_t]) \tag{4}$$

其中，$\tilde{h}_t$ 表示候选的隐藏状态；$\boldsymbol{W}$ 表示权重矩阵；“$\odot$”表示逐元素乘积；tanh 表示双曲正切函数。

最终的隐藏状态 h_t 是前一步状态 h_{t-1} 和候选隐藏状态 $\tilde{h}_t$ 的线性插值。其公式为：

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{5}$$

1.3　Transformer 模型

Transformer 模型是一种基于注意力机制的深度学习架构，通过引入自注意力机制，实现并行处理和长距离依赖的建模，在自然语言处理任务中可以显著提升性能。Transformer 模型由编码器（Encoder）和解码器（Decoder）组成，每个部分均由多个相同的层堆叠而成。Transformer 的典型结构如图 2 所示。

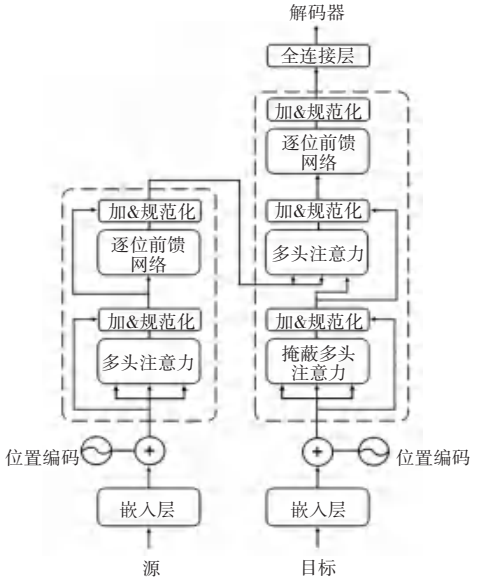


图 2　Transformer 模型结构图

Fig. 2　Diagram of Transformer model structure

Transformer 模型每个编码器层均包含 2 个主要子层：多头自注意力机制（Multi-Head Self-Attention）和前馈神经网络（Feed Forward Neural Network，FFNN）。多头注意力机制的公式为：

$$\text{Attention}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{6}$$

其中，$\boldsymbol{Q}$ 表示查询矩阵；$\boldsymbol{K}$ 表示键矩阵；$\boldsymbol{V}$ 表示值矩阵；d_k 表示键向量的维度。

前馈神经网络首先计算隐藏层的输入，使用线性变换和偏置，对隐藏层的输入应用激活函数 ReLU，计算得出输出层的输入，再经线性变换和偏置，得到最终输出，其公式为：

$$\text{FFNN}(x) = \text{ReLU}(x\boldsymbol{W}_1 + \boldsymbol{b}_1)\boldsymbol{W}_2 + \boldsymbol{b}_2 \tag{7}$$

其中，$\boldsymbol{W}_1$ 表示第一层的权重矩阵；$\boldsymbol{W}_2$ 表示第二层的权重矩阵；$\boldsymbol{b}_1$ 表示第一层的偏置向量；$\boldsymbol{b}_2$ 表示第二层的偏置向量。

ReLU 激活函数定义为：

$$\text{ReLU}(z) = \max(0, z) \tag{8}$$

解码器部分与编码器结构类似，相比之下增设第三子层，用于对编码器的输出进行注意力计算。

不同于 RNN 的循环结构，Transformer 完全基于注意力的机制，无法天然捕捉序列中的位置关系，因此引入位置编码以保留输入序列中各个元素的位置信息，并将位置维度分为奇偶数维度。

偶数维度位置编码采用正弦函数，其式为：

$$\text{PE}_{(\text{pos},2i)} = \sin\left(\frac{\text{pos}}{10\,000^{2i/d_{\text{model}}}}\right) \tag{9}$$

奇数维度位置编码采用余弦函数，其式为：

$$\text{PE}_{(\text{pos},2i+1)} = \cos\left(\frac{\text{pos}}{10\,000^{2i/d_{\text{model}}}}\right) \tag{10}$$

其中，PE 表示位置编码，用于给定位置和维度的编码值；pos 表示序列中元素的位置索引；i 为维度索引，表示位置编码向量中的第几个维度；d_{model} 表示 Transformer 模型中的隐藏层向量的维度。这些位置编码函数确保每个位置在每个维度都有唯一的编码值，从而帮助 Transformer 模型捕捉输入序列的顺序信息。

2　STL-GRU-Transformer 模型构建

在交通预测领域中，通常采用组合模型或优化算法寻优，其核心旨在探索传统统计学方法与循环神经网络间的联系，或是通过增加注意力机制模块处理噪声影响，这类方法都具有局限性，虽能提高预测准确性，但提升程度有限。本文根据日常交通流量所呈现出的非线性与周期性，考虑使用不同特性

模型,分别预测经 STL 分解出的因子数据,通过将表层规律性数据与残差数据分离,深度挖掘数据内部的逻辑规律以提高预测精确度。最后使用小波阈值重构预测结果。本文提出的 STL－GRU－Transformer 预测模型结构如图 3 所示。

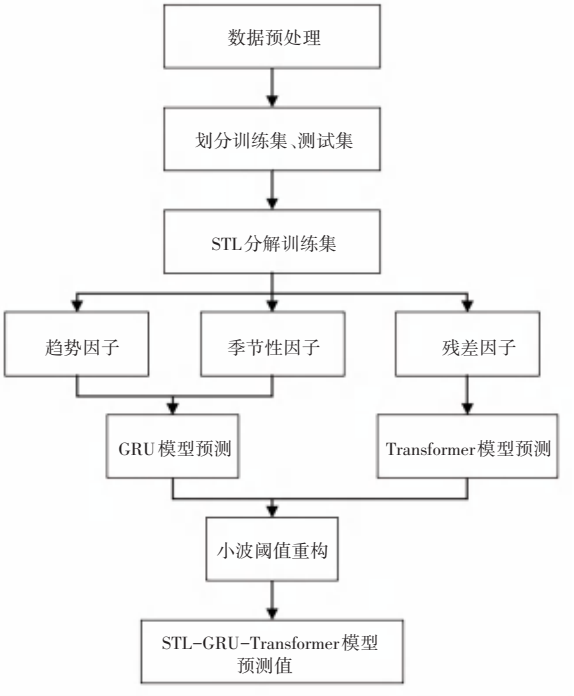


图 3 STL-GRU-Transformer 预测模型结构图

Fig. 3 Structure diagram of STL-GRU-Transformer prediction model

该组合模型在输入交通流量后,通过 STL 分解将时间序列分解为季节因子、趋势因子和残差因子三部分。因 GRU 在循环神经网络中相较于 RNN 与 LSTM 等模型更具结构简单、计算效率高和易捕捉关键点的特点,考虑选用 GRU 模型对具备较强规律性的季节因子和趋势因子进行预测;残差因子是去除趋势和季节因子后的残余数据,其趋势往往呈现出以 0 为均值的随机波动特征,但其内部仍然有着与交通流特征相适应的数据特征待挖掘,同时 Transformer 模型拥有强大的自注意力机制,能够捕捉长距离依赖关系和复杂的时序模式,适合找到复杂残差信号之间的关联点,因此考虑使用 Transformer 模型进行预测。在模型针对不同因子完成预测后,实验选用小波阈值重构方法对预测结果进行重构,通过小波分解将信号分解成不同频率的子带系数,根据选定的阈值对小波系数加以处理,将低于阈值的系数置零或缩减,高于阈值的系数保留或缩减。最后,通过逆小波变换将处理后的小波系数进行信号重构,得到去噪后的预测结果。

通过结合 GRU 的高效率与 Transformer 的捕捉长距离关联优势,再加上小波阈值重构方法来提升组合模型结果的鲁棒性,使该组合模型能更准确预测交通流量,提升预测性能。

3 实验分析

3.1 数据集来源及参数设置

数据集选用浙江省温州市机场大道高架路上某检测点数据,时间跨度为 2024 年 1 月 1 日 0 时起至 2024 年 2 月 29 日 24 时,每 5 min 计算一次流量数据,实际数据 17 044 条,经预处理与补全后共 17 280 条流量数据。仿真实验中选取其中 80% 数据作为训练集,剩下 20% 数据作为测试集,即选取 2024 年 1 月 1 日 0 时至 2024 年 2 月 17 日 24 时、共 13 824 条流量数据作为训练集,余下的 3 456 条数据作为测试集。

本研究运行环境为 Python 3.11, 采用 PyTorch 1.13.1 作为深度学习框架,实验中硬件配置为 Core i9-13900K。GRU 模型的输入维度 input_size 和输出维度 out_size 均为 1,隐藏层维度 hidden_size 为 50,使用单层 GRU,训练过程中使用 MSE 损失函数和 Adam 优化器,学习率 learning_rate 为 0.001,迭代 epoch 为 200 次。Transformer 模型的输入维度 n_input 为 24,线性层的维度 d_model 为 128,多头注意力机制头数 nhead 为 8,编码器和解码器层数均为 3,训练时同样使用 MSE 损失函数和 Adam 优化器,学习率 learning_rate 为 0.001,迭代 epoch 为 200 次。小波阈值重构方法中使用 Daubechies 小波,阈值基于信号的标准差计算,进行软阈值处理。

3.2 评价指标

对预测模型的效果的评价采用误差评价,误差越小则说明模型预测精度越高,效果越好。误差评价指标使用平均绝对误差(Mean Absolute Error, MAE)、均方误差(Mean Square Error,MSE)、均方根误差(Root Mean Square Error,RMSE),计算公式为:

$$E_{\mathrm{MAE}} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{11}$$

$$E_{\mathrm{MSE}} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{12}$$

$$E_{\mathrm{RMSE}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{13}$$

其中,y_i 和 $\hat{y}_i$ 分别表示实际值和预测值。

3.3 基于 STL-GRU-Transformer 模型预测

在交通流量数据经预处理与补全后,将所分割出的训练集数据进行 STL 分解,分解结果如图 4 所

示。观察所分解出的 3 条因子曲线，发现于 2024 年 1 月 1 日至 2024 年 2 月 15 日期间，趋势因子和季节性因子的曲线表现出较稳定周期变化，但自 2 月 8 日之后，呈现出趋势因子数值显著下降、季节因子振幅逐渐减小的态势。考虑该变化是因春节假期，民众放假、务工人员返乡或出游等因素导致此区域交通流量缩减。而残差因子在该情况下仍呈现出以 0 为均值的波动特征，反向证明 STL 分解仍具有效性。经观察发现，残差因子的周期变化同季节因子的周期变化相吻合，展现出工作日较稳定，多呈现正残差、周末残差波动大，多呈现负残差、正负残差峰值能达到 200 veh 等结构特点。

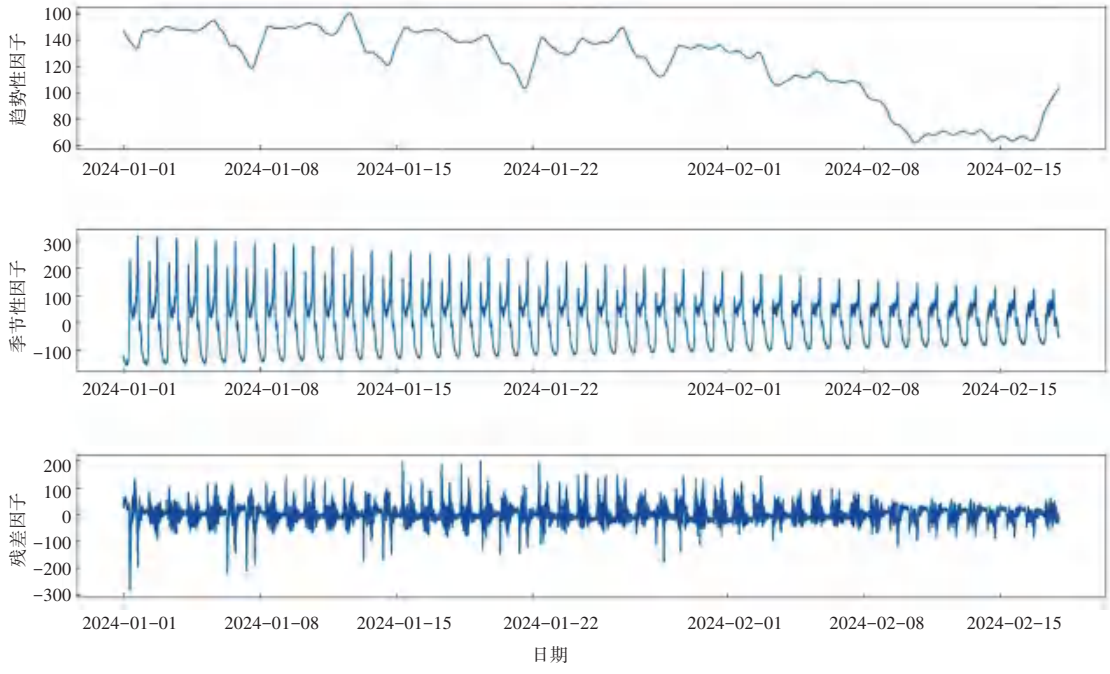


图 4　STL 分解流量序列结果

Fig. 4　Flow sequence results of STL decomposition

在对数据集进行分割后，于 Python 内对 GRU 与 Transformer 模型先进行定义，后将季节因子和趋势因子输入 GRU 模型，将残差因子输入 Transformer 模型，根据历史数据迭代训练预测未来的 3 456 条数据。随后匹配预测长度与测试集长度，对比预测值与真实值。为分批展示模型预测效果，先输出 GRU 模型重构效果，如图 5 所示。图 5 中，预测值与真实值展现出高度拟合，误差极小，说明 GRU 模型成功抓取季节因子与趋势因子的关键特征，迭代学习后实现高精度预测。再结合 Transformer 模型预测结果，通过小波阈值重构输出，并生成组合模型重构效果图，如图 6 所示。由图 6 可知，组合模型在趋势上已经能成功拟合交通流趋势，即使训练集和测试集中或多或少受假期多变流量影响，结果仍具有较高拟合度。虽然预测时高峰期峰值与真实值略有差距，后续还可以通过优化算法寻找最优参数设置，进一步拟合来提高预测效率。

3.4　实验结果及分析

预测模型生成并输出数据后，将输出结果以误差评价指标计算，为验证组合模型优越性，本文选用 ARIMA、BP、RNN、LSTM、GRU 和 Transformer 六个模型作为对照。本文所选取的对比模型基本参数与迭代次数基本一致，预测结果见表 1。分析表 1 可知，深度学习模型预测精度普遍优于传统统计学模型；在循环神经网络模型中，GRU 模型优于 LSTM，LSTM 优于 RNN；Transformer 模型虽然是注意力机制模型，但同样适用于时间序列的预测；STL-GRU-Transformer 的预测精度最高。

再以 STL-GRU 模型和 STL-Transformer 模型结果作为对照，生成组合模型预测精度，见表 2。通过对比表 2 和表 1 结果可知，经 STL 分解的交通流量序列，能让预测模型更好发现数据内在规律，进行特征捕捉。通过误差数值对比可见组合模型优于单模型，STL-GRU 和 STL-Transformer 模型较其原始模型，MAE 分别提升了 3.16 和 5.41。STL-GRU-Transformer 模型预测精度的进一步提升也证明该组合模型的优越性，使最终的 MAE 较单模型结果至少提高 5.98，精度提升 18.89%。

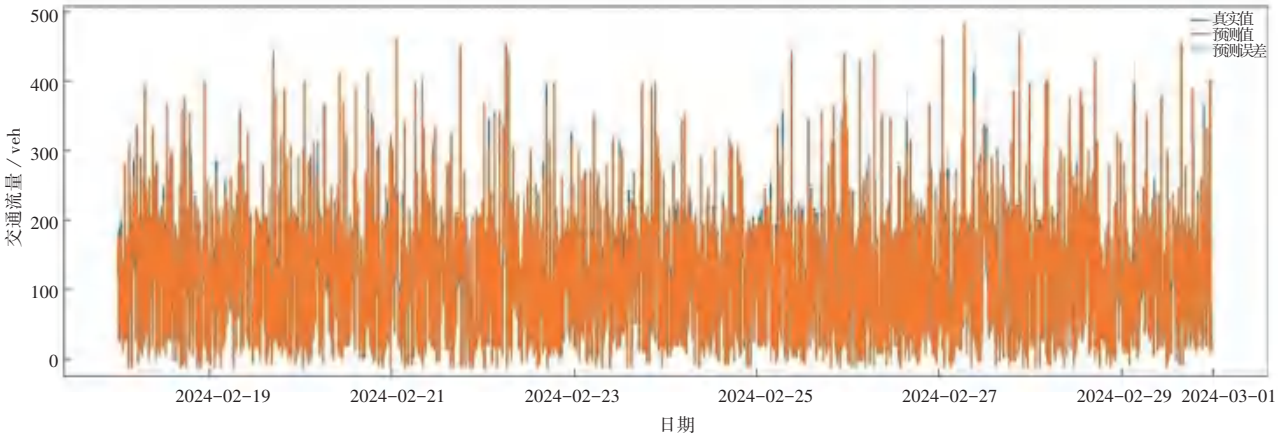


图 5 GRU 模型部分预测效果图

Fig. 5 Partial prediction effect of GRU model

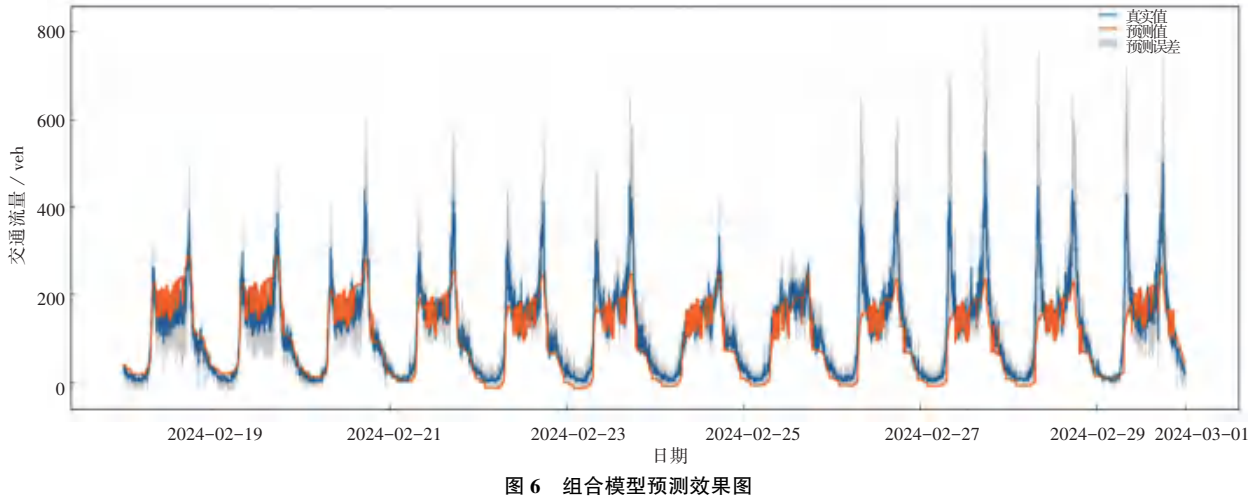


图 6 组合模型预测效果图

Fig. 6 Prediction effect of combined model

表 1 单模型预测精度表

Table 1 Prediction accuracy table of single model

预测模型	MAE	MSE	RMSE
ARIMA	34.69	2 474.59	49.75
BP	34.16	2 393.40	48.92
RNN	34.10	2 400.43	48.99
LSTM	32.81	2 363.09	48.61
GRU	32.06	2 181.55	46.71
Transformer	31.65	2 141.54	46.28
STL-GRU-Transformer	25.67	1 244.84	35.28

表 2 组合模型预测精度表

Table 2 Prediction accuracy table of combined model

预测模型	MAE	MSE	RMSE
STL-GRU	28.90	1 597.84	39.97
STL-Transformer	26.24	1 314.65	36.26
STL-GRU-Transformer	25.67	1 244.84	35.28

4 结束语

本文提出了一种基于 STL-GRU-Transformer 的短时交通流量预测方法,得到如下结论:

(1)STL 分解方法能有效提取交通流量的趋势变换,分离出的季节因子和趋势因子有助于 GRU 模型捕捉关键特征。

(2)Transformer 模型对残差信号的处理有益于提升整体模型预测效果。

(3)通过实验与分析,基于 STL-GRU-Transformer 的组合预测模型表现出最优异的预测效果。结果表明,该组合模型的预测效果的 MAE 为 25.67,MSE 为 1 244.84,RMSE 为 35.28,相较于单模型预测效果至少提升 18.89%。

参考文献

[1] 马寿峰, 贺国光, 刘豹. 智能交通系统中短时交通流预测系统的研究[J]. 预测, 2004(2): 28-34.

[2] 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, 7: 28-39.
[3] 王晓全，邵春福，尹超英，等. 基于 ARIMA-GARCH-M 模型的短时交通流预测方法[J]. 北京交通大学学报，2018，42(4)：79-84.
[4] 李洪娟，吕渭济，张建军. 城市道路交叉口高峰小时流量的灰色预测[J]. 辽宁工程技术大学学报(自然科学版)，2000，19(6)：656-658.
[5] 王进，史其信. 短时交通流预测模型综述[J]. ITS 通讯，2005，7(1)：10-13.
[6] VYTHOULKAS P C. Alternative approaches to short term traffic forecasting for use in driver information systems[C]//Proceedings of the 12th International Symposium on the Theory of Traffic Flow and Transportation. Amsterdam:Elsevier, 1993: 458-505.
[7] 周伟. 组合时间序列分解技术的短时交通流预测研究[D]. 南京：东南大学，2022.
[8] 刘建伟，宋志妍. 循环神经网络研究综述[J]. 控制与决策，2022，37(11)：2753-2768.
[9] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
[10] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv preprint arXiv, 1412.3555, 2014.
[11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st Conference on Neural Information Process System. Long Beach, USA: NIPS Foundation, 2017: 5998-6008.
[12] 陈宇，王炜，华雪东，等. 基于递归框架的高速公路交通流量实时预测方法[J]. 交通信息与安全，2023，41(1)：124-131.
[13] 王莉莉，赵云飞. 基于 GA、RBF 和改进 Cao 方法的空中交通流预测方法[J]. 交通信息与安全，2023，41(1)：115-123.
[14] 周海赟，闫冬梅. 基于 SSARX-NARX 模型的短时交通流预测[J]. 交通信息与安全，2019，37(4)：89-95.
[15] 闫少华，谢晓璇，张兆宁. 基于小波优化 GRU-ARMA 模型的空中交通流量短时预测方法[J]. 交通信息与安全，2022，40(4)：177-184.
[16] 王宁，成利敏，甄景涛，等. 基于 PSO-LSTM 的短时交通流量预测网站设计[J]. 廊坊师范学院学报(自然科学版)，2024，24(1)：29-32.
[17] 王代君，李明，鹿守山. 基于 Bayes-ARIMA 的景区公路短时交通流量预测[J]. 公路，2024，69(4)：225-234.
[18] 徐厚生，郭佳丽. 结合时间注意力机制的 Bi-GRU-Atten 的短时交通流预测[J]. 辽宁工程技术大学学报(自然科学版)，2023，42(6)：763-768.
[19] 廉清云，孙伟，李润生. 基于注意力机制的堆叠 LSTM 短时船舶交通流预测模型[J]. 大连海事大学学报，2024，50(1)：57-65.
[20] 陈昆，曲大义，王少杰，等. 基于二次分解和融合多特征的短时交通流量组合预测模型[J]. 广西师范大学学报(自然科学版)，2023，41(4)：33-46.