

doi: 10.3969/j.issn.1002-0268.2021.11.015

基于 ARMA-SVR 的短时交通流量 预测模型研究

王博文¹, 王景升¹, 朱茵¹, 王统一², 张泽有¹

(1. 中国人民公安大学 交通管理学院, 北京 100038; 2. 山东科技大学 电气信息学院, 山东 济南 250000)

摘要: 短时交通流量预测是辅助智能交通系统进行决策, 解决拥堵问题, 提高道路通行能力的关键任务。为提高模型对于小样本数据的拟合效果及速度, 并充分挖掘交通流序列中存在的线性与非线性关系, 将对于线性数据具备良好拟合效果, 并且时间复杂度较低的自回归滑动平均 (ARMA) 模型与对于非线性、小样本数据具有计算准确率高、时间复杂度低等优势的支持向量回归 (SVR) 模型进行组合, 提出一种残差优化组合预测模型。采用赤池信息准则对 ARMA 模型进行定阶, 实现交通流量的线性拟合, 并得到相应的残差序列。然后将重构后的残差序列作为 SVR 模型的输入, 对残差序列进行预测, 以补偿交通流量数据中的非线性变化。将 ARMA、SVR、长短期记忆网络、人工神经网络及 ARMA-SVR 加权组合模型作为对照组进行模型评价。结果表明: 样本的时间间隔分别为 5, 10, 15 min 时, ARMA-SVR 残差优化组合模型的均方根误差 (RMSE) 及平均绝对误差 (MAE) 均小于对照组模型, RMSE 降低约 0.378 ~ 7.063, MAE 降低约 0.054 ~ 0.802; ARMA-SVR 残差优化组合模型在不同的样本时间间隔下均具备较高的预测能力、较低的时间复杂度及数据计算成本, 可以满足基于不同样本时间间隔的交通流量预测的需要。

关键词: 智能交通; 交通流量预测模型; 自回归滑动平均模型; SVR 模型; 智能交通

中图分类号: U491.1+4

文献标识码: A

文章编号: 1002-0268 (2021) 11-0126-08

Study on Short-term Traffic Volume Prediction Model Based on ARMA-SVR

WANG Bo-wen¹, WANG Jing-sheng¹, ZHU Yin¹, WANG Tong-yi², ZHANG Ze-you¹

(1. School of Traffic Management, People's Public Security University of China, Beijing 100038, China;

2. School of Electrical Information, Shandong University of Science and Technology, Jinan Shandong 250000, China)

Abstract: Short-term traffic volume forecasting is the key task to assist ITS to make decisions, solve congestion problems and improve road capacity. In order to improve the fitting effect and speed of the model for small sample data, and to fully explore the linear and nonlinear relationships existing in the traffic volume sequence, the ARMA model with good fitting effect for linear data and low time complexity is combined with the support vector regression (SVR) model with advantages of high calculation accuracy and low time complexity for nonlinear and small sample data, and an residual optimization composite forecasting model is proposed. The order of the ARMA model is set by using Akaike information criterion to realize the linear fitting of traffic volume, and the corresponding residual sequence is obtained. Then, using the reconstructed residual sequence as the input of SVR model, and the residual sequence is predicted to compensate for the nonlinear changes in traffic volume data. The ARMA, SVR, long-term and short-term memory networks, artificial neural network, and ARMA-SVR weighted composite forecasting model are used as control group for

收稿日期: 2021-03-15

基金项目: 公安部公安理论及软科学研究计划项目 (2020LLYJGADX020); 中国人民公安大学拔尖创新人才培养经费支持研究生科研创新项目成果 (2021jysky014); 中国人民公安大学公共安全行为科学实验室开放课题基金资助项目 (2020SYS15)

作者简介: 王博文 (1999-), 女, 河南周口人, 硕士研究生. (201621310017@stu.ppsuc.edu.cn)

model evaluation. The result shows that (1) when the sampling time interval is 5, 10, 15 min, the root mean square error (RMSE) and mean absolute error (MAE) of the ARMA-SVR residual optimization composite forecasting model are smaller than those of the control models, and the RMSE and MAE are reduced by about 0.378 – 7.063 and 0.054 – 0.802 respectively; (2) the ARMA-SVR residual optimization composite forecasting model has higher prediction ability, lower time complexity and data calculation cost in different sample time intervals, which can meet the needs of traffic volume prediction based on different sample time intervals.

Key words: ITS; traffic volume prediction model; autoregressive moving average (ARMA) model; SVR model; intelligent transport

0 引言

交通流量预测是智能交通系统 (Intelligent Traffic System, ITS) 的重要组成部分, 准确的交通流量预测模型可以辅助交通信号配时、拥堵疏导等任务, 是智能交通领域研究的热点问题。

早期的交通流预测任务通常采用传统统计模型, 通过使用线性模型挖掘交通流序列中存在的线性关系^[1-3]。如 EMAMI 等^[4]研发了一种卡尔曼滤波器对交通流进行了有效的预测, 为短期交通流量预测提供了一种低成本手段。KUMAR 等^[5]提出季节性自回归移动平均 (Autoregressive Integrated Moving Average, ARIMA) 模型, 利用历史数据和实时数据对早晚高峰时段的交通流量进行短期预测, 得到了良好的预测效果。但是, 随着居民出行需求的提高, 交通流所具备的非线性特征愈发明显, 许多机器学习的算法, 如支持向量机, 及深度学习算法, 如等及长短时记忆 (Long Short-Term Memory, LSTM)、卷积神经网络 (Convolutional Neural Network, CNN)、图神经网络 (Graph Neural Networks, GNN) 及其改进算法, 被应用于短时交通流预测^[6-18], 这类模型多为非线性模型, 以期望充分挖掘交通流序列的非线性关系。邹宗民等^[19]使用粒子群算法对支持向量回归 (Support Vector Regression, SVR) 进行参数寻优, 并将优化后的模型用于交通流预测任务中。相较于传统统计算法, 提高了模型的预测准确性, 相较于深度学习算法, SVR 在小样本交通流序列预测中表现较好。温惠英等^[20]引入双长短期记忆网络, 将基准模型拆分为两个方向, 并通过试验证明了模型对于非线性交通流数据具备良好的预测和泛化能力。李磊等^[21]提出一种 CNN 和 LSTM 的结合方法, 有效提取了交通流的时空特征, 并通过大量试验证明了非线性模型对于交通流序列的预测情况较为良好。GUO 等^[22]构建一种基于注意力机制的时空图卷积网

络, 对交通流序列进行了准确的预测。单一的传统统计模型或浅层机器学习模型不能够同时挖掘交通流序列中存在的线性及非线性关系。基于深度学习的模型, 尤其是基于 GNN 及其改进算法的模型, 对于试验样本为大样本量的依赖性较高, 数据计算成本高, 并且时间复杂度相较于传统统计模型及浅层机器学习算法高。短时交通流预测任务作为辅助 ITS 进行决策的关键工作, 需要预测模型对道路情况进行快速、准确的判断, 对于预测实时性的要求极高。

因此, 为解决小样本、快速短时交通流预测问题, 并充分挖掘交通流序列中存在线性与非线性关系, 本研究将对于线性数据具备良好拟合效果, 并且时间复杂度较低的自回归滑动平均 (Auto Regressive Moving Average, ARMA) 模型与对于非线性、小样本数据的处理具有计算准确率高、时间复杂度低等优势优势的 SVR 模型进行组合, 达到提高交通流预测模型预测准确率及降低模型时间复杂度的目的。

1 交通流预测方法

1.1 理论模型

1.1.1 ARMA 模型

ARMA 模型是目前应用最广的线性平稳时间序列预测模型, 模型的形式为:

$$y_t = c + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \quad (1)$$

式中, y_t 为被预测变量; y_{t-1} 为与被预测变量相关的滞后阶; c 为常数; ε_t 为白噪声序列, 服从均值为 0 的正态分布; α_i 为自相关系数 ($i=1, 2, \dots, p$); θ_i 为移动平均系数 ($i=1, 2, \dots, q$); p, q 分别为自回归、移动平均过程中的滞后阶数。

参数 p, q 的选择决定了 ARMA 模型的拟合效果。当样本量 n 固定时, 使赤池信息准则 (Akaike Information Criterion, AIC) 达到最小值的 (p, q)

即为最佳参数组合, AIC 的计算公式为:

$$A = 2n - 2\ln L, \quad (2)$$

式中, n 为样本个数; L 为似然函数。

1.1.2 SVR 模型

SVR 是将支持向量机 (Support Vector Machine, SVM) 应用于回归问题, 目的在于建立一个最优的超平面, 使训练样本距离最优超平面的误差最小, 此时模型达到最佳拟合效果。

对于非线性数据集 ($i = 1, 2, \dots, m$), x_i, y_i 分别为输入和输出。把样本从低维空间映射到高维空间, 超平面公式表示为:

$$f(x_i) = \omega^T \Phi(x_i) + b, \quad (3)$$

式中, ω 为超平面的法向量; $\Phi(x_i)$ 为 x_i 映射到高维空间后的特征向量; b 为位移项。

模型的训练过程即参数 ω, b 的寻优过程, 最终得到一组参数 ω, b 使 $f(x_i)$ 最接近 y_i , 将问题转化为凸二次规划问题, 表示为:

$$\begin{aligned} \min_{\omega, b} & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) \\ \text{s. t.} & \begin{cases} f(x_i) - y_i \leq \varepsilon + \xi_i \\ y_i - f(x_i) \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (4)$$

式中, C 为惩罚项, C 用于平衡结构风险和经验风险, 即 $\frac{1}{2} \|\omega\|^2$ 和 $\sum_{i=1}^m (\xi_i + \xi_i^*)$; ξ_i, ξ_i^* 为松弛变量; ε 为 y_i 与 $f(x_i)$ 的偏差。

使用拉格朗日乘子法计算得到:

$$\omega = \sum_{i=1}^m (a_i^* - a_i) \Phi(x_i), \quad (5)$$

式中 a_i, a_i^* 为拉格朗日对偶问题的解。

使用核函数代替线性方程中的线性项可以使原来的线性算法非线性化, 即能做非线性回归, 此时引进核函数达到了提升维度的目的, 也可以有效地控制过拟合。不同的核函数会对 SVR 的性能产生影响。径向基核函数 (RBF) 的抗干扰能力及适应能力较强, 表示为:

$$K(X, X') = \exp(-\gamma \|X - X'\|^2), \quad (6)$$

式中, $K(X, X')$ 为满足 Mercer 条件的核函数; γ 为核参数, 决定映射样本在特征空间的分布; X 为输入样本构成的特征向量; X' 为 RBF 核函数的中心。

计算得到 SVR 回归模型为:

$$f(x) = \sum_{i=1}^m (a_i^* - a_i) K(X, X') + b. \quad (7)$$

RBF 的参数 C (惩罚项) 和 g (核函数方差) 的选取对于模型的结果存在较大影响。

1.2 两种组合预测方法

1.2.1 ARMA-SVR 加权组合模型

CRITIC 法是一种比熵权法和标准离差法更好的客观权重赋权法, 它基于评价指标的对比强度和指标之间的冲突性来综合衡量指标的客观权重。

假设有 n 个待评价样本, p 项评价指标, 形成原始指标数据矩阵:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}, \quad (8)$$

式中 x_{ij} 为第 i 个样本的第 j 项评价指标的值。

指标的对比强度以标准差的形式来表现:

$$\begin{cases} \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \\ S_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}} \end{cases} \quad (9)$$

式中, $\bar{x}_j$ 为第 j 项指标的均值; S_j 为第 j 项指标的标准差。

在 CRITIC 法中使用标准差来表示各指标的内取值的差异波动情况, 标准差越大表示该指标的数值差异越大, 越能反映出更多的信息, 该指标本身的评价强度也就越强, 应该给该指标分配更多的权重。

指标的冲突性用相关系数进行表示:

$$R_j = \sum_{i=1}^p (1 - r_{ij}), \quad (10)$$

式中 r_{ij} 为评价指标 i 和 j 之间的相关系数。

使用相关系数来表示指标间的相关性, 与其他指标的相关性越强, 则该指标就与其他指标的冲突性越小, 反映出相同的信息越多, 所能体现的评价内容就越有重复之处, 一定程度上也就削弱了该指标的评价强度, 应该减少对该指标分配的权重。

第 j 个评价指标所包含的信息量 C_j 表示为:

$$C_j = S_j \sum_{i=1}^p (1 - r_{ij}) = S_j R_j, \quad (11)$$

所以第 j 个指标的客观权重 W_j 为:

$$W_j = \frac{C_j}{\sum_{j=1}^p C_j}. \quad (12)$$

ARMA-SVR 加权组合模型原理如图 1 所示。

(1) 使用 ARMA 模型对样本进行建模, 得到预测结果 P_{ARMA} 。

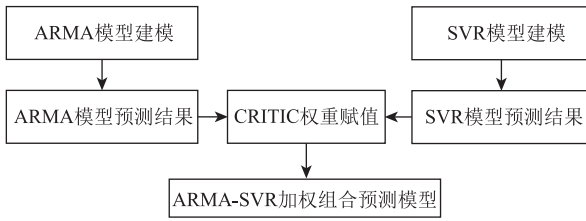


图 1 ARMA-SVR 加权组合模型的建模流程

Fig. 1 Modeling process of ARMA-SVR weighted composite model

(2) 使用 SVR 模型对样本进行建模, 得到预测结果 P_{SVR} 。

(3) 使用 CRITIC 赋权法对 ARMA 与 SVR 模型的预测结果进行客观赋权, 分别得到权重 W_{ARMA} 和 W_{SVR} 。

(5) 将 ARMA 模型预测结果 P_{ARMA} 与 SVR 模型预测结果 P_{SVR} 进行加权相加, 得到 ARMA-SVR 加权组合模型的预测结果表示为:

$$C_{ARMA_SVR} = W_{ARMA}^* P_{ARMA} + W_{SVR}^* P_{SVR} \quad (13)$$

1.2.2 ARMA-SVR 残差优化组合模型

交通流存在线性和非线性的特性, 而 ARMA 和 SVR 模型分别对线性模型和非线性模型的处理上具备优势, 理论上将两个模型的优势相结合, 能够达到提高模型效果的目的。假设时间序可视为线性自相关部分与非线性残差相结合的结果, 表示为:

$$P_t = L_t + R_t, \quad (14)$$

式中, L_t 为线性自相关部分; R_t 为非线性残差。

ARMA-SVR 残差优化组合模型原理如图 2 所示。

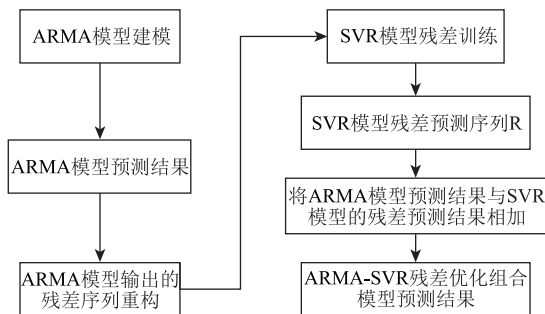


图 2 ARMA-SVR 残差优化组合模型的建模流程

Fig. 2 Modeling process of ARMA-SVR residual optimization composite model

由图 2 得, ARMA-SVR 残差优化组合模型的基本原理为:

(1) 使用 ARMA 模型对样本的线性部分进行建模, 得到预测结果 ARMA, 进而得出残差序列 R 。

(2) 以特定的输入步长对残差序列进行重构。

(3) 将重构后的残差序列 R^* 作为 SVR 模型的输入。

(4) 使用 SVR 模型对残差序列的非线性部分进行建模, 得到残差序列预测结果。

(5) 将残差序列预测结果与 ARMA 模型预测结果相加, 得到 ARMA-SVR 残差优化组合模型的预测结果。

1.3 评价指标

回归模型中, 常用均方误差 (Mean Absolute Error, MAE)、平均绝对误差 (Mean Squared Error, MSE)、均方根误差 (Root Mean Squared Error, RMSE) 作为模型的评价指标, 本研究使用 MAE 和 RMSE 作为评价指标, 模型的 MAE 和 RMSE 越小, 预测效果越好。

2 实例验证

2.1 数据来源

本试验采用内蒙古包头市某路段 2020 年 9 月的地磁交通流量数据, 每组数据的单次采样时间间隔为 5 min, 包括单向所有车道的交通流量总和, 样本总量为 8 640。

2.2 数据分析与预处理

将数据分别处理为以 5, 10, 15 min 为时间间隔的数据。3 个样本集的样本数量分别为 8 640 个, 4 320 个, 2 880 个。将 3 个样本集分别按照 8:2 的比例划分训练集与测试集。处理后的交通流量变化趋势如图 3 所示。

2.3 模型建立

2.3.1 ARMA 模型

平稳性及白噪声检验。对样本进行平稳性检验及白噪声检验详见表 1。

表 1 交通流量序列的平稳性检验

Tab. 1 Stationarity test of traffic volume sequence

样本的时间间隔/min	t 统计量	平稳性检验				白噪声检验	
		置信度			P 值	LB	P 值
		1%	5%	10%		统计量	
5	-9.45	-3.43	-2.86	-2.57	4.65E-16	6 135.459	0
10	-11.44	-3.43	-2.86	-2.57	6.11E-21	3 245.015	0
15	-11.79	-3.43	-2.86	-2.57	1.00E-21	2 323.877	0

注: LB 统计量为白噪声检验的检验统计量。

由表 2 得, 平稳性检验的 t 统计量值小于 1%, 5%, 10% 这 3 个水平的值, 且 P 值小于 0。白噪声

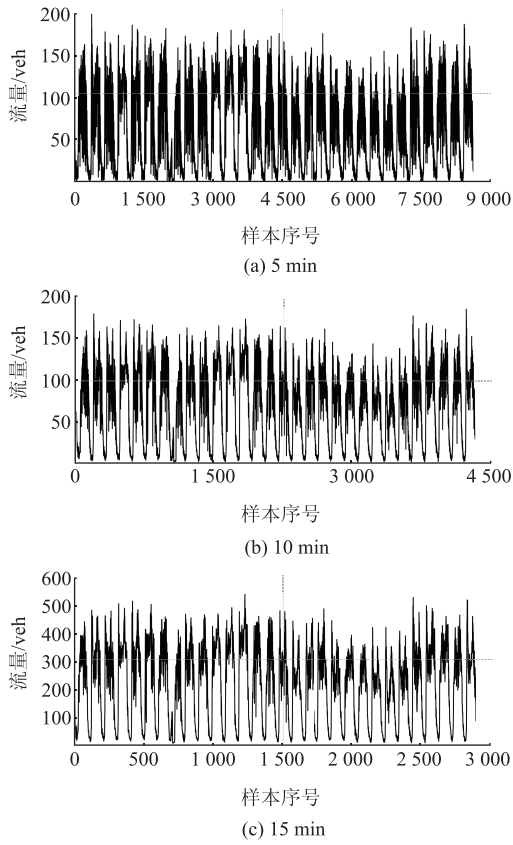


图 3 不同样本时间间隔下的交通流量的变化趋势
Fig. 3 Variation trends of traffic volume at different sample time intervals

检验的 P 值小于 0, 所以 3 个样本集均为平稳非白噪声序列。

ARMA 模型参数确定。多次试验得到, 当样本的时间间隔分别为 5, 10, 15 min 时 $p = 4, q = 3$, AIC 值达到最小, 最佳参数组合为 (4, 3)。

对 3 个 ARMA 模型的参数进行显著性检验, 详见表 2。由表 2 可得, 3 个模型均满足显著性小于 0.05, 说明参数的取值具备合理性。上述建模流程得到 ARMA 模型预测结果。

2.3.2 SVR 模型

数据归一化。使用 SVR 模型之前, 将数据做归一化操作, 以此提高模型的收敛速度和预测能力。本研究采用 Rescaling, 即 Min-Max 归一化, 将交通流量数据归一化至 $[0, 1]$ 区间, 表示为:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{9}$$

式中, x' 为归一化后的交通流量数据; x 为原始交通流量数据; $x_{\min}$ 为样本的最小值; $x_{\max}$ 为样本的最大值。

表 2 ARMA 模型参数的显著性检验

Tab. 2 Significance test of ARMA model parameters

样本的时间 间隔/min	参数	回归系数	标准误差	t 统计量	显著性
5	常量	71.36	5.07	14.07	0.00
	α_1	0.84	0.06	14.68	0.00
	α_2	0.61	0.07	8.76	0.00
	α_3	-0.74	0.06	-12.30	0.00
	α_4	0.27	0.02	16.13	0.00
	θ_1	-0.42	0.06	-7.26	0.00
	θ_2	-0.73	0.06	-12.93	0.00
	θ_3	0.50	0.04	12.03	0.00
10	常量	142.74	10.17	14.04	0.00
	α_1	0.56	0.06	9.11	0.00
	α_2	-0.52	0.08	-6.72	0.00
	α_3	0.88	0.05	18.80	0.00
	α_4	0.17	0.02	14.13	0.00
	θ_1	-0.15	0.06	-2.36	0.02
	θ_2	0.73	0.06	11.64	0.00
	θ_3	-0.38	0.04	-10.78	0.00
15	常量	215.08	4.78	45.04	0.00
	α_1	0.66	0.04	15.06	0.00
	α_2	1.27	0.04	31.62	0.00
	α_3	-0.67	0.04	-15.30	0.00
	α_4	-0.28	0.04	-6.91	0.00
	θ_1	-0.12	0.05	-2.66	0.01
	θ_2	-0.97	0.01	-178.56	0.00
	θ_3	0.15	0.04	3.32	0.00

注: α_i 为 ARMA 模型的自相关系数 ($i = 1, 2, \dots, p$); θ_i 为 ARMA 模型的移动平均系数 ($i = 1, 2, \dots, q$)。

$$Y^* = \hat{y}[\hat{y}_{\max} - \hat{y}_{\min}] + \hat{y}_{\min}, \tag{10}$$

式中, Y^* 为反归一化后的交通流量预测值; $\hat{y}$ 为模型输出的交通流量预测值。 $\hat{y}_{\max}$ 为模型输出的交通流量预测值的最大值; $\hat{y}_{\min}$ 为模型输出的交通流量预测值的最小值。

本研究中 SVR 模型的建立均基于 sklearn 机器学习库, 模型的参数为默认值。使用训练好的 SVR 模型对测试集进行预测, 对 SVR 预测结果进行反归一化后, 上述建模流程得到 SVR 模型预测结果。

2.3.3 ARMA-SVR 加权组合预测模型

在模型相互补偿的过程中, 需要确定各个模型输出结果的权重, 使用 CRITIC 权重法对 ARMA 和 SVR 预测的交通流量进行融合。模型的权重详见表 3。

表 3 CRITIC 权重法赋予模型的权重

Tab. 3 Weights given to model by the CRITIC weighting method

采样时间间隔/min	W_{ARMA}	W_{SVR}
5	0.510	0.490
10	0.497	0.503
15	0.477	0.523

将 ARMA 模型预测结果与 SVR 模型预测结果进行加权组合, 得到 ARMA-SVR 加权组合预测模型的预测结果。

2.3.4 ARMA-SVR 残差优化组合模型

ARMA 模型得到的残差如图 4 所示。

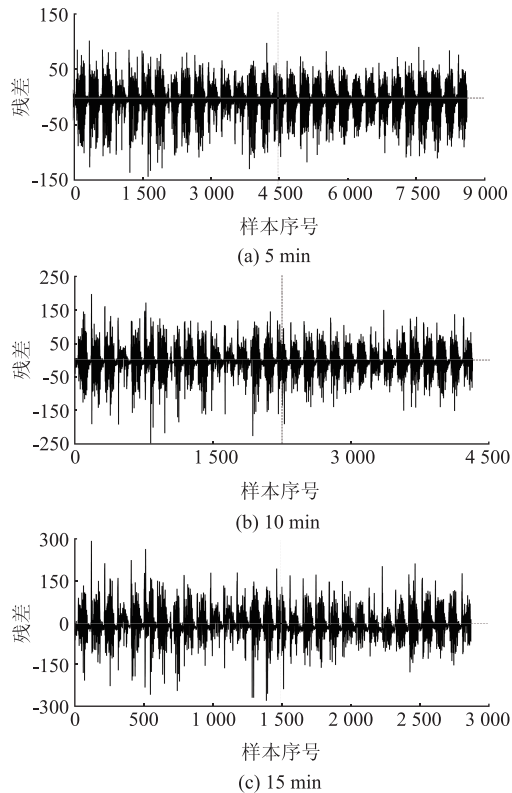


图 4 不同样本时间间隔下 ARMA 模型得到的残差

Fig. 4 Residues obtained by ARMA model at different sample time intervals

首先对通过 ARMA 模型得到残差进行归一化处理, 将残差值规约在 $[0, 1]$ 。然后按照 8:2 的比例划分训练集与测试集。最后使用 SVR 对重构后的残差序列进行预测。

使用 SVR 模型以 2 个时间步长读取训练集的残差序列进行训练, 然后使用训练好的 ARMA-SVR 残差优化组合模型对测试集进行预测。对残差预测结果进行反归一化后, 将 SVR 模型残差预测结果与 ARMA 模型预测结果相加, 得到 ARMA-SVR 残差优

化组合模型的预测结果。

2.4 模型评价

为验证 ARMA-SVR 残差优化组合模型的优越性, 将 ARMA 模型、SVR 模型、长短期记忆网络 (Long Short-Term Memory, LSTM) 模型、人工神经网络 (Artificial Neural Networks, ANN) 模型及 ARMA-SVR 加权组合模型作为对照组进行模型评价。

LSTM 和 ANN 模型的建立均基于 Keras 深度学习网络框架, 激活函数均为线性整流函数 Relu, 损失函数为 MSE, 优化器为 Adam, 默认神经网络隐层数设置为 1, 神经元数量为 8, 通过 compile 方法完成学习过程的配置, 训练的迭代次数为 100, 当监测到 loss 停止改进时, 结束训练。

使用 Min-Max 归一化, 将交通流量数据归一化至 $[0, 1]$ 区间, 分别使用训练好的 LSTM 和 ANN 模型以 $t-1$ 及 t 时刻的交通流量值预测 $t+1$ 时刻的交通流量值。

不同模型在样本的时间间隔分别为 5, 10, 15 min 时对于测试集的拟合评价结果详见表 4。

表 4 交通流量预测模型的评价结果

Tab. 4 Evaluation result of traffic volume forecasting model

模型	RMSE			MAE		
	5 min	10 min	15 min	5 min	10 min	15 min
ARMA	23.726	40.094	54.547	4.058	5.410	6.386
SVR	26.103	43.178	55.343	4.367	5.555	6.468
ANN	28.625	46.532	59.586	4.779	5.877	6.743
LSTM	26.351	43.651	58.499	4.398	5.785	6.741
ARMA-SVR (CRITIC)	23.681	40.948	54.754	4.056	5.539	6.411
ARMA-SVR (Residual)	23.303	39.469	54.005	3.977	5.331	6.332

由表 4 得, 在预测准确率方面, 对于小样本数据, 相较于传统统计模型及机器学习模型, 深度学习模型的预测准确率较低。当采样间隔为 5 min 时, ARMA-SVR 加权组合模型的 RMSE, MAE 值均低于 ARMA, SVR, LSTM, ANN 的单一模型, 模型拟合效果较 4 个单一模型好。在不同的采样时间间隔下, ARMA-SVR 残差优化组合预测模型的预测准确率和稳定性较其他模型均有所提升, RMSE 降低约 0.378 ~ 7.063, MAE 降低约 0.054 ~ 0.802, 可以为不同采样时间间隔下交通流量预测提供方法参考。除此之外, 通过大量试验得到, 在时间复杂度方面, ARMA 和 SVR 的两种组合模型的时间复杂度及设备的配置远远低于深度学习模型 ANN 及 LSTM。相关

研究人员可将模型的预测结果用于 ITS 的短期决策任务中,如交叉口的信号配时。

不同模型在样本的时间间隔分别为 5, 10, 15 min 时,在测试集上的拟合情况如图 5 所示。

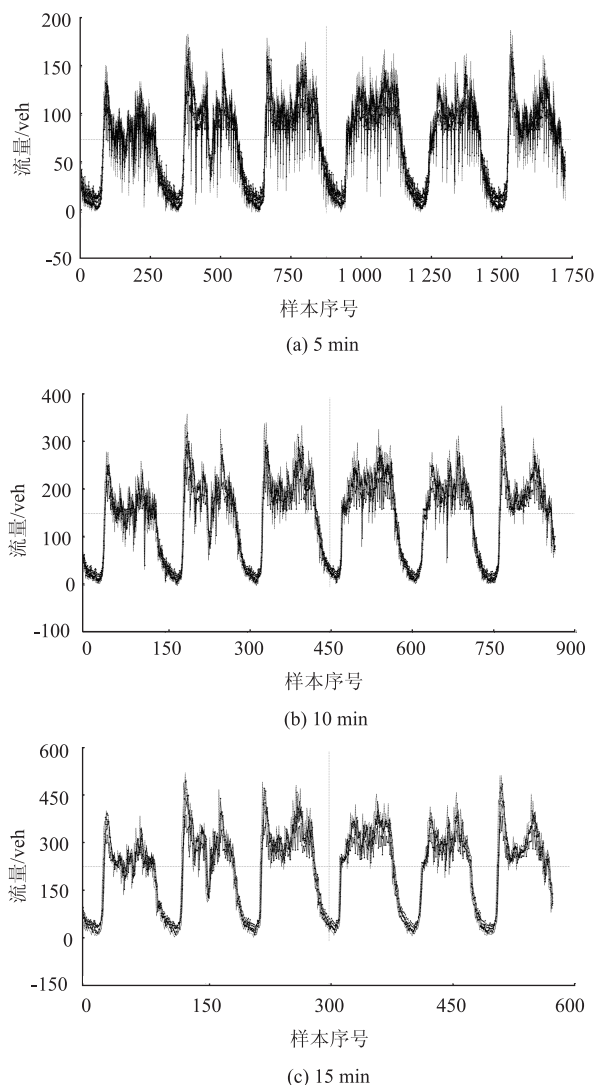


图5 不同时间间隔下交通流量预测模型在测试集上的拟合情况

Fig. 5 Fitting condition of traffic volume prediction model on test set at different time intervals

3 结论

本研究将对于线性数据具备良好拟合效果的 ARMA 模型与对于非线性数据处理具有优势的 SVR 模型进行组合,充分挖掘了交通流序列中的线性及非线性关系,提高了模型预测的准确率。并通过试验证明,对比 ARMA, SVR, LSTM, ANN 及 ARMA-SVR 加权组合预测模型,ARMA-SVR 残差优化组合模型在小样本的短时交通流预测任务中表现出预测精度高、计算速度快及数据计算成本低的优势。并

且在样本的时间间隔分别为 5, 10, 15 min 时 ARMA-SVR 残差优化组合模型均能较好地反映交通流量序列中的线性及非线性关系,可以满足基于不同样本时间间隔的交通流流量预测的需要,将该模型应用于支持 ITS 的决策需求,可以提升 ITS 决策速度及精确度,同时能够降低 ITS 决策系统的建设成本。因为交通流序列往往受多因素的影响,本研究未考虑多因素对预测目标的影响,之后的研究可以通过扩展数据获取的范围,将速度、占有率、天气状况、能见度及事故信息等元素纳入预测模型。

参考文献:

References:

- [1] 刘潭秋,孙湘海,钟翔. 基于三制度 SETAR 模型的短期交通流预测 [J]. 公路交通科技, 2010, 27 (10): 122 - 127, 142.
LIU Tan-qi, SUN Xiang-hai, ZHONG Xiang. Short-term Traffic Flow Forecasting Based on a Three-regime SETAR Model [J]. Journal of Highway and Transportation Research and Development, 2010, 27 (10): 122 - 127, 142.
- [2] 盛慧,王炜,胡晓健. 基于改进型 CTM 模型的道路交通流量预测 [J]. 公路交通科技, 2009, 26 (增 1): 49 - 52.
SHENG Hui, WANG Wei, HU Xiao-jian. Traffic Volume Forecast Based on Improved Cell Transmission Model [J]. Journal of Highway and Transportation Research and Development, 2009, 26 (S1): 49 - 52.
- [3] 孙占全,潘景山,张赞军,等. 基于主成分分析与支持向量机结合的交通流预测 [J]. 公路交通科技, 2009, 26 (5): 127 - 131.
SUN Zhan-quan, PAN Jing-shan, ZHANG Zan-jun, et al. Traffic Flow Forecast Based on Combining Principal Component Analysis with Support Vector Machine [J]. Journal of Highway and Transportation Research and Development, 2009, 26 (5): 127 - 131.
- [4] EMAMI A, SARVI M, BAGLOEE S A. Using Kalman Filter Algorithm for Short-term Traffic Flow Prediction in a Connected Vehicle Environment [J]. Journal of Modern Transportation, 2019, 27 (3): 222 - 232.
- [5] KUMAR S V, VANAJAKSHI L. Short-term Traffic Flow Prediction Using Seasonal ARIMA Model with Limited Input Data [J]. European Transport Research Review, 2015, 7 (3): 1 - 9.
- [6] 韩志聪,樊彦国,吴会胜,等. 基于 GA-SVR 模型的短期交通流量预测方法研究 [J]. 公路交通科技, 2017, 34 (1): 130 - 136.
HAN Zhi-cong, FAN Yan-guo, WU Hui-sheng, et al.

- Study on Methods of Short-time Traffic Volume Forecasting Based on GA-SVR Model [J]. Journal of Highway and Transportation Research and Development, 2017, 34 (1): 130–136.
- [7] 马秋芳. 改进 PSO 优化的 BP 神经网络短时交通流预测 [J]. 计算机仿真, 2019, 36 (4): 94–98, 323.
MA Qiu-fang. BP Neural Network Short-term Traffic Flow Prediction Based on Improved Particle Swarm Optimization [J]. Computer Simulation, 2019, 36 (4): 94–98, 323.
- [8] 吴晋武, 张海峰, 冉旭东. 基于数据约减和支持向量机的非参数回归短时交通流预测算法 [J]. 公路交通科技, 2020, 37 (7): 129–134.
WU Jin-wu, ZHANG Hai-feng, RAN Xu-dong. Nonparametric Regressive Short-term Traffic Flow Forecast Algorithm Based on Data Reduction and SVM [J]. Journal of Highway and Transportation Research and Development, 2020, 37 (7): 129–134.
- [9] JIANG W W, ZHANG L. Geospatial Data to Images: A Deep-learning Framework for Traffic Forecasting [J]. Tsinghua Science and Technology, 2019, 24 (1): 52–64.
- [10] 万玉龙, 李新春, 周红标. 基于 WPD-PSO-ESN 的短期交通流预测 [J]. 公路交通科技, 2019, 36 (8): 144–151.
WAN Yu-long, LI Xin-chun, ZHOU Hong-biao. Prediction of Short-term Traffic Flow Based on WPD-PSO-ESN [J]. Journal of Highway and Transportation Research and Development, 2019, 36 (8): 144–151.
- [11] 张逸飞, 付玉慧. 基于 ARIMA-BP 神经网络的船舶交通事故预测 [J]. 上海海事大学学报, 2020, 41 (3): 47–52.
ZHANG Yi-fei, FU Yu-hui. Prediction of Ship Traffic Accidents Based on ARIMA-BP Neural Network [J]. Journal of Shanghai Maritime University, 2020, 41 (3): 47–52.
- [12] CHAN K Y, DILLON T S, SINGH J, et al. Neural-network-based Models for Short-term Traffic Flow Forecasting Using a Hybrid Exponential Smoothing and Levenberg-Marquardt Algorithm [J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13 (2): 644–654.
- [13] HUANG W H, SONG G J, HONG H K, et al. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15 (5): 2191–2201.
- [14] WANG X X, XU L H, CHEN K X. Data-driven Short-term Forecasting for Urban Road Network Traffic Based on Data Processing and LSTM-RNN [J]. Arabian Journal for Science and Engineering 2019, 44 (4): 3043–3060.
- [15] 熊亨, 戚湧, 张伟斌. 基于 DCGRU-RF 模型的路网短时交通流预测 [J]. 计算机科学, 2020, 47 (5): 84–89.
XIONG Ting, QI Yong, ZHANG Wei-bin. Short-term Traffic Flow Prediction Based on DCGRU-RF Model for Road Network [J]. Computer Science, 2020, 47 (5): 84–89.
- [16] 朱征宇, 刘琳, 崔明. 一种结合 SVM 与卡尔曼滤波的短时交通流预测模型 [J]. 计算机科学, 2013, 40 (10): 248–251, 278.
ZHU Zheng-yu, LIU Lin, CUI Ming. Short-term Traffic Flow Forecasting Model Combining SVM and Kalman Filtering [J]. Computer Science, 2013, 40 (10): 248–251, 278.
- [17] LOPEZ-GARCIA P, ONIEVA E, OSABA E, et al. A Hybrid Method for Short-term Traffic Congestion Forecasting Using Genetic Algorithms and Cross Entropy [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17 (2): 557–569.
- [18] 刘钊, 杜威, 闫冬梅, 等. 基于 K 近邻算法和支持向量回归组合的短时交通流预测 [J]. 公路交通科技, 2017, 34 (5): 122–128, 158.
LIU Zhao, DU Wei, YAN Dong-mei, et al. Short-term Traffic Flow Forecast Based on Combination of K Nearest Neighbor Algorithm and Support Vector Regression [J]. Journal of Highway and Transportation Research and Development, 2017, 34 (5): 122–128, 158.
- [19] 邹宗民, 郝龙, 李全杰, 等. 基于粒子群优化-支持向量回归的高速公路短时交通流预测 [J]. 科学技术与工程, 2021, 21 (12): 5118–5123.
ZOU Zong-min, HAO Long, LI Quan-jie, et al. Short-term Traffic Flow Prediction of Expressway Based on Particle Swarm Optimization-Support Vector Regression [J]. Science Technology and Engineering, 2021, 21 (12): 5118–5123.
- [20] 温惠英, 张东冉. 基于 Bi-LSTM 模型的高速公路交通量预测 [J]. 公路工程, 2019, 44 (6): 51–56.
WEN Hui-ying, ZHANG Dong-ran. Highway Traffic Volume Prediction Based on Bi-LSTM Model [J]. Highway Engineering, 2019, 44 (6): 51–56.
- [21] 李磊, 张青苗, 赵军辉, 等. 基于改进 CNN-LSTM 组合模型的分时段短时交通流预测 [J]. 应用科学学报, 2021, 39 (2): 185–198.
LI Lei, ZHANG Qing-miao, ZHAO Jun-hui, et al. Short-Term Traffic Flow Prediction Method of Different Periods Based on Improved CNN-LSTM [J]. Journal of Applied Sciences, 2021, 39 (2): 185–198.
- [22] GUO S, LIN Y, FENG N, et al. Attention Based Spatial-temporal Graph Convolutional Networks for Traffic Flow Forecasting [C] // Proceedings of the AAAI Conference on Artificial Intelligence. Atlanta: AAAI Press, 2019, 33 (1): 922–929.