

基于组合模型的短时交通流量预测算法

芮兰兰 李钦铭*

(北京邮电大学网络与交换技术国家重点实验室 北京 100876)

摘要: 交通流量预测是实现智能交通技术的核心问题,及时准确地预测道路交通流量是实现动态交通管理的前提,短时交通流量的预测是交通流量预测的重要组成部分。该文针对十字路口的短时交通流量预测问题设计了基于交通流量序列分割和极限学习机(Extreme Learning Machine, ELM)组合模型的交通流量预测算法(Traffic Flow Prediction Based on Combined Model, TFPBCM)。该算法首先采用K-means对交通流量数据在时间上进行序列分割,然后采用ELM对各个序列进行建模和预测。仿真实验证明,与单一的BP(Back Propagation)神经网络和ELM相比,该组合模型算法建模时间为BP的1/10, ELM建模时间的4倍,均方误差为BP的1/50, ELM的1/20, 该组合模型算法决定系数 R^2 更接近于1, 模型可信度更高。

关键词: 短时交通流量; K均值算法; 极限学习机; 组合模型算法

中图分类号: TP393.07

文献标识码: A

文章编号: 1009-5896(2016)05-1227-07

DOI: 10.11999/JEIT150846

Short-term Traffic Flow Prediction Algorithm Based on Combined Model

RUI Lanlan LI Qinming

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Traffic flow prediction is a key problem of realizing intelligent transportation technology. Forecasting traffic flow in time and accurately is the precondition to realize the dynamic traffic management. Short-term traffic flow prediction is an important part of traffic flow prediction. In this paper, the Traffic Flow Prediction Based on Combined Model (TFPBCM) based on traffic flow sequence partition and Extreme Learning Machine (ELM) is designed for the short time traffic flow forecasting. The algorithm divides the traffic flow into different patterns along a time dimension by K-means, and then models and forecasts for each pattern by ELM. The proposed algorithm is compared with Back Propagation (BP) and ELM. The combined model algorithm on modeling time is 1/10 of BP, but is 4 times ELM. Its MSE is 1/50 of BP and 1/20 of ELM. The combined model algorithm's coefficient of determination (R^2) is close to 1, so the credibility of the model is higher than others.

Key words: Short-term traffic flow; K-means algorithm; Extreme Learning Machine (ELM); Traffic Flow Prediction Based on Combined Model (TFPBCM)

1 引言

随着国内城市的高速发展,交通拥堵,交通效率低下造成了交通事故的频发,这已成为各个城市发展面临的共同问题^[1]。及时做好交通流量预测对各个主干道路进行调整成为了解决城市交通拥堵问题

的重要途径之一。交通流量预测能有效缓解交通拥堵,降低事故发生率,给出出行者舒适安全的交通环境。道路交通变化是一个实时、非线性、非平稳的随机过程,统计时段越短,交通流量短时变化的随机性和不确定性越强。由于十字路口是城市交通的重要组成部分,交通情况复杂,干扰因素多,其交通能力不足的问题已成为交通网络的瓶颈^[2,3]。因此,十字路口短时交通流量预测的研究是实现城市交通智能化的一个关键问题。

文献[3]提出基于交通流量序列分割和BP神经网络相结合的组合预测模型,但该模型预测时间长,预测精度较低。文献[4-6]利用BP神经网络预测短时交通流量,然而该类算法易陷入局部极小值,泛

收稿日期: 2015-07-14; 改回日期: 2016-01-08; 网络出版: 2016-03-11

*通信作者: 李钦铭 liqinming@bupt.edu.cn

基金项目: 国家自然科学基金创新研究群体科学基金(61121061), 国家自然科学基金(61302078, 61372108), 北京高等学校青年英才计划项目(YETP0476)

Foundation Items: Funds for Creative Research Groups of China (61121061), The National Natural Science Foundation of China (61302078, 61372108), Beijing Higher Education Young Elite Teacher Project (YETP0476)

化能力差,预测精度低。文献[7-9]提出一种单隐层前向神经网络训练算法——极限学习机(Extreme Learning Machine, ELM),该算法降低了网络参数调节时间,泛化能力强,然而,当要处理的数据量过大时,隐含层神经元个数难以确定,过大或过小都会带来较大误差。文献[10]提出利用相邻数周中相对应时间的交通流比值代替原始数据,建立基于历史数据的 Kalman 滤波交通流预测模型,但没有解决输出延迟的问题。文献[11]提出一种基于经验模态分解(EMD)和神经网络的短时交通流量预测方法,通过 EMD 把交通流量分解成不同模态,利用神经网络对分解后的各分量进行预测,再将预测值累加得到最终的预测结果。文献[12]提出基于混沌粒子群优化 CPSO 小波神经网络 WNN 的短时交通流量,利用混沌粒子群算法优化小波神经网络的模型参数,克服传统小波神经网络采用梯度下降法易陷入局部极值和引起震荡效应现象的缺陷。文献[13]提出一种结合卡尔曼滤波与支持向量机的预测模型,该模型采用预测误差平方和与相关系数极大化准则智能选取预测方式,利用支持向量机的稳定性与卡尔曼滤波的实时性。文献[14,15]分别提出基于频谱分析能够实时更新的短时交通流预测算法和基于回归的 Hermite 神经网络的经验模式分解预测模型,都利用了模态函数来预测短时交通流量。文献[16]提出了一种基于指数平滑法和 Levenberg Marquardt 的混合算法来训练神经网络预测交通流量。指数平滑法用于交通流量数据的预处理。文献[17]提出了一种新的基于非参数模型的短期交通流量预测方法,该模型基于分类和回归树。文献[18]提出基于 ARIMA 模型的高速公路交通流量预测方法,但是 ARIMA 模型对数据的要求高,其中异常和缺失数据对模型具有较大影响,不适合进行模型的自适应调整。

针对十字路口短时交通流量预测问题,本文提出基于交通流量时间序列分割和极限学习机组合模型的交通流量预测算法(Traffic Flow Prediction Based on Combined Model, TFPBCM)。该算法首先采用 K-means 算法对交通流量进行聚类,通过引入时间基线将整个数据域按时间进行序列分割,然后采用极限学习机(ELM)对各个分割序列进行建模和预测。该算法的创新点,第一,在序列分割阶段引入时间基线的概念,通过聚类和时间基线对数据进行划分,保证划分后的数据块之间在流量值和时间上具有更多相似性,减小由于数据差异性较大带来的高误差;第二,将 K-means 与 ELM 组合运用,与单一 BP 神经网络和 ELM 相比,该组合模型建模时间(t)为 BP 的 1/10, ELM 建模时间的 4 倍,均方

误差(MSE)为 BP 的 1/50, ELM 的 1/20, 该组合模型算法决定系数(R^2)更接近于 1, 模型可信度更高。该预测模型基于前馈型神经网络,相比 ARIMA 模型自适应调整能力更强,与 Kalman 滤波交通流预测模型相比不存在输出延迟的问题。

本文后续部分分为:第 2 节是问题描述;第 3 节为算法设计;第 4 节详述了组合预测模型算法;第 5 节是仿真实验;第 6 节是本文的总结。

2 问题描述

十字路口的简化模型如图 1 所示,十字路口的交通流量相对于单点路段上的交通流量有更大的波动性,尤其是短时流量序列。能否及时、准确地预测十字路口短时交通流量成为亟待解决的问题。

当前,基于某十字路口,每 5 min 统计一次该时间段内十字路口经过的汽车总数,记录下来,用 $S_{[n]}$ 表示,利用数据分割函数 θ 对车流量数据 $S_{[n]}$ 进行分割, $M_1, M_2, \dots, M_k$ 为分割后的数据段, θ 用 K-means 算法逼近,如式(1)所示。

$$\theta(S_{[n]}) = \{M_1, M_2, \dots, M_k\} \quad (1)$$

例如, $S_{[n]}$ 为 7:30~10:00 之间的车流量数据,通过 K-means 算法将该组数据划分成了时间为 7:00~7:50, 7:55~9:05, 9:10~10:00 对应的数据段。规定 5 min 为 1 个时间单位,对每一数据段进行预测, $f(t)$, $f(t-1)$, $f(t-2)$, $f(t-3)$, $f(t-4)$ 为当前数据段预测方向上的前 5 个时间单位的流量值,利用流量预测模型 φ 预测下一个时间单位的交通流量 $f(t+1)$, Φ 用 ELM 神经网络逼近,如式(2)所示。

$$f(t+1) = \varphi(f(t), f(t-1), f(t-2), f(t-3), f(t-4)) \quad (2)$$

例如,当前为 7:00~7:50 时间范围内的数据段, $f(t)$, $f(t-1)$, $f(t-2)$, $f(t-3)$, $f(t-4)$ 分别为 7:00~7:05, 7:05~7:10, 7:10~7:15, 7:15~7:20 和 7:20~7:25 时间范围内的车流量数据,将其作为 ELM 神经网络输入,预测 7:25~7:30 时间范围内的车流量数据即 $f(t+1)$ 。

针对以上问题,本文提出基于交通流量序列分割和 ELM 组合模型的交通流量预测算法 TFPBCM, 具体见第 3 节。

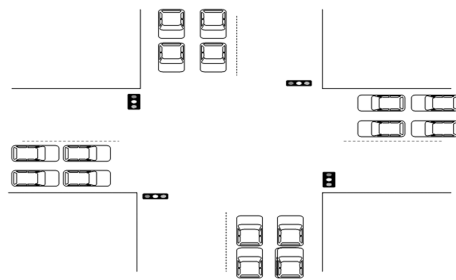


图 1 十字路口简化模型

3 算法设计

TFPBCM 算法首先对交通流量数据进行聚类, 利用聚簇中心计算得到一组时间基线, 时间基线将数据域按时间进行序列分割, 然后利用极限学习机对各个分割序列进行建模和预测。由于数据之间具有差异性, 通过聚类和时间基线对数据进行划分, 保证了划分后的数据块之间具有更多的相似性, 减小了由于数据差异性较大带来的高误差。极限学习机降低了网络参数调节时间, 泛化能力强, 预测精度高。

通过均方误差、决定系数和建模时间 3 个指标来评价预测算法, 当有一组交通流量数据到来时, 平台就会按表 1 中的方法进行预测。

表 1 TFPBCM 算法流程

输入: 聚类数目 k 和交通流量数据 $S_{[n]}$, 神经网络隐含层神经元个数 n 和激活函数 sigmoid。	
输出: 未来连续时间单位的交通流量值 T 。	
Begin	
(1)	初始化神经网络输入层到隐含层连接权值 ω
(2)	初始化隐含层神经元阈值 b
(3)	选择 k 个数据对象作为聚簇中心
(4)	While(1)
(5)	计算所有对象 $S_{[n]}$ 与各个簇心的距离并将对象划分到距离其最近的簇
(6)	重新计算每个聚簇的中心对象
(7)	If 新的聚簇中心与前一次聚簇中心相比无变化
(8)	Then
(9)	聚簇形成, 得到最新聚簇的质心 o 和质量 r
(10)	Break
(11)	End-If
(12)	End While
(13)	计算时间基线 t_w^*
(14)	利用 t_w^* 划分交通数据流量 X 得到 k 段数据, $D_1, D_2, \dots, D_k$
(15)	For i from 1 to k
(16)	利用 D_i 中的训练集计算隐含层输出矩阵 H
(17)	利用 D_i 中的训练集计算隐含层和输出层连接权值 K
(18)	利用 D_i 中的测试集计算未来连续时间单位的交通流量 T
(19)	End For
(20)	End

$$\mathbf{X} = [S_1 \ S_2 \ S_3 \ S_4 \ S_5]$$

$$\mathbf{H}(\omega_1, \omega_2, \dots, \omega_n, b_1, b_2, \dots, b_n, s_1, s_2, \dots, s_5) = \left[\begin{array}{cccc} g(\omega_1 \cdot s_1 + b_1) & g(\omega_2 \cdot s_1 + b_2) & \cdots & g(\omega_n \cdot s_1 + b_n) \\ g(\omega_1 \cdot s_2 + b_1) & g(\omega_2 \cdot s_2 + b_2) & \cdots & g(\omega_n \cdot s_2 + b_n) \\ \vdots & \vdots & \ddots & \vdots \\ g(\omega_1 \cdot s_5 + b_1) & g(\omega_2 \cdot s_5 + b_2) & \cdots & g(\omega_n \cdot s_5 + b_n) \end{array} \right]_{5 \times n} \quad (4)$$

将第 6 个时间单位的流量值 $\mathbf{Y}$ 作为输出, 训练模型, 对于隐含层与输出层之间的连接权值 K , 可

4 算法详述

针对十字路口短时交通流量预测问题, 本文提出了基于交通流量时间序列分割和 ELM 组合模型的交通流量预测算法。

交通流量数据是按时间顺序排列的一组数据, 用 $S_{[n]}$ 表示, 本节选择 K-means 聚类算法作为序列分割的基础算法, 对交通流量序列进行分割。对交通流量数据进行聚类, 假设得到 k 个簇集, 每个簇集的质心为 o , 质量为 r 。质心 o 的横坐标为时间 t , 纵坐标为该时段的平均车流量, 质量 r 为簇集中对象的个数 m 。按照质心横坐标 t 的先后顺序进行排序得到 $t_1, t_2, \dots, t_k$, 对应的簇集质量为 $m_1, m_2, \dots, m_k$ 。假设时间基线为 t_w^* , 如式(3)所示。

$$t_w^* = t_w + \left| \frac{m_w}{m_{w+1} - m_w} \right| \times (t_{w+1} - t_w), w \in \{1, 2, \dots, k-1\} \quad (3)$$

由于交通流数据是按时间顺序排列的一组数据, 利用时间基线将交通流量数据分割为 k 段, 记做 $D_1, D_2, \dots, D_k$ 。

对于分割后的交通流量序列, 需要选择一种预测模型对序列分别预测。本节采用极限学习机神经网络算法, 网络结构如图 2 所示, 该模型具有学习速度快、泛化性能好的特点。隐含层输出矩阵 H 和隐含层与输出层之间的连接权值 β 是 ELM 模型训练和预测过程中两个重要的变量, 本节首先利用最近到来的历史交通流量数据进行模型训练, 得到 H 和 K , 然后利用 H 和 K 对即将到来的交通流进行流量预测。

假设隐含层神经元为 n , 选择激活函数 $g(x)$ 为 sigmoid。依次处理 k 段数据, 即 $D_1, D_2, \dots, D_k$ 。在 D_i 中, 最近到来历史数据的前 5 个时间单位交通流量数据 X 作为模型的输入, 计算得到隐含层输出矩阵 H , 如式(4)所示, 其中, ω 为输入层到隐含层的连接权值, b 为隐含层神经元的阈值, 可随机获得。

以通过最小二乘法求解, 如式(5)所示, 其中, H^+ 为隐含层输出矩阵 H 的 Moore-Penrose 广义逆, $\mathbf{Y}^T$ 是

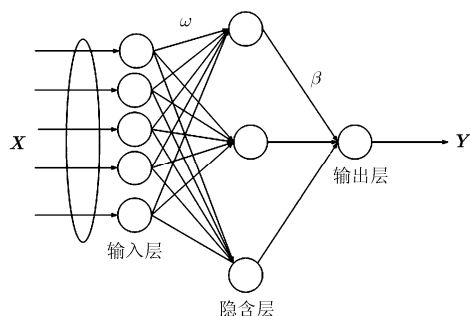


图2 ELM网络结构

Y 的转置。至此，为每段数据完成了模型的训练过程。

$$Y = [s_6], \quad K = H^+ \cdot Y^T \quad (5)$$

在预测阶段，将预测方向上的前 5 个时间单位的交通流量数据作为模型的输入，对未来到来的时间单位的流量值 T 进行预测，如式(6)所示。

$$T = [t_1 \ t_2 \ \cdots \ t_s]_{m \times s},$$

$$t_j = \begin{bmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{mj} \end{bmatrix}_{m \times 1} = \begin{bmatrix} \sum_{i=1}^n \beta_{i1} \cdot g(\omega_i x_j + b_i) \\ \sum_{i=1}^n \beta_{i2} \cdot g(\omega_i x_j + b_i) \\ \vdots \\ \sum_{i=1}^n \beta_{im} \cdot g(\omega_i x_j + b_i) \end{bmatrix}_{m \times 1}, \quad j = 1, 2, \dots, s \quad (6)$$

TFPBCM 算法首先采用 K-means 算法对道路交通流量进行聚类，通过引入时间基线将整个数据域按时间进行序列分割，然后再采用 ELM 对各个分割序列进行建模和预测。该算法在建模速度、模型可信度和预测精度上，都有较大的提高。

5 仿真实验

仿真实验采用 Matlab 分别对 BP 神经网络、ELM 和 TFPBCM 算法进行训练和仿真，并且从建模时间(t)，均方误差(mse)和决定系数(R^2) 3 方面出发，对 3 种模型进行评估。

5.1 数据来源

现有两组源数据，第 1 组为北京市平谷区桃园桥交叉十字路口 3 月 16 到 3 月 19 号上午 7:00~10:00 的交通流量数据，如 7:05, 7:10, 7:15 等时刻的数据，单位为(辆/5 min)，共 144 个交通数据，前 3 天的数据用于训练，最后一天的数据用于测试；第 2 组为文献[2]的数据集，2011 年 10 月 25 日至 10 月 27 日 7:00~19:00 的某十字交叉路口的交通数据，单位为(辆/5 min)，只取上午 7:00~10:00 交通数据作为实验数据，共 108 个交通数据，前两天数据用于训

练，最后一天数据用于测试。 k 均值聚类数目为 3，ELM 激活函数 sigmoid。

两组数据的训练集与测试集的分布情况如表 2 和表 3 所示。

表 2 第 1 组训练集与测试集分布情况

时段	训练集样本	测试集样本
7:00~7:50	30	10
7:55~9:05	45	15
9:10~10:00	33	11

表 3 第 2 组训练集与测试集分布情况

时段	训练集样本	测试集样本
7:00~7:50	26	13
7:55~9:05	28	14
9:10~10:00	18	9

5.2 仿真结果

选用 3 层 BP 神经网络结构，对 7:30~10:00 之间的车流量进行预测。预测方向上前 5 个时间单位的车流量数据作为 BP 神经网络的输入，对未来最近一个时间单位的车流量数据进行预测，如将 7:05, 7:10, 7:15, 7:20 和 7:25 时间点的数据作为输入，预测 7:30 时间点上的车流量数据。由经验可知，隐含层神经元个数表示为 $\sqrt{m+n}+a$ ，其中 m 为输入层神经元个数， n 为输出层神经元个数， a 为 1~10 之间的整数。通过决定系数的大小，找到最优隐含层神经元个数，由图 3 可以得出，当隐含层神经元的个数为 9 时，决定系数最接近 1，拟合优度大，自变量对因变量的解释程度高，模型可信度越高，预测结果如图 4，图 5 所示。

第 1 组数据 BP 预测结果的均方误差为 16.9662，决定系数为 0.92642，建模时间为 0.96875。第 2 组数据 BP 预测结果的均方误差为 15.96265，决定系数为 0.92315，建模时间为 0.9217。

ELM 模型输入与 BP 神经网络相同，为 3 层神经网络结构，由经验可知，最优隐含层神经元个数在 25~35 之间，通过决定系数，找到最优隐含层神经元个数，由图 6 可以得出，当隐含层神经元个数为 31 时，决定系数最接近 1，拟合优度大，自变量对因变量的解释程度高，模型可信度越高，预测结果如图 7，图 8 所示。

第 1 组数据 ELM 预测结果的均方误差为 6.4447，决定系数为 0.96733，建模时间为 0.03125。第 2 组数据 ELM 预测结果的均方误差为 8.43261，决定系数为 0.9589，建模时间为 0.03421。

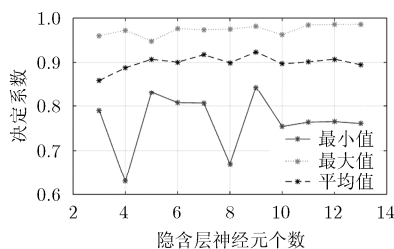


图3 BP 决定系数对比结果

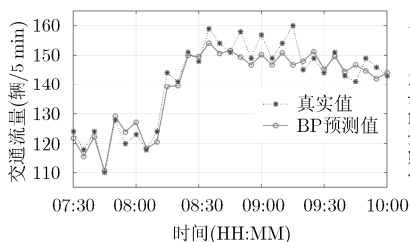


图4 第1组数据 BP 预测结果

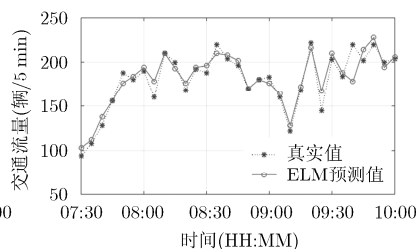


图5 第2组数据 BP 预测结果

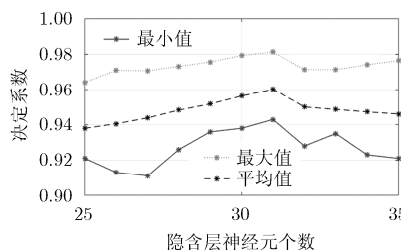


图6 ELM 决定系数对比结果

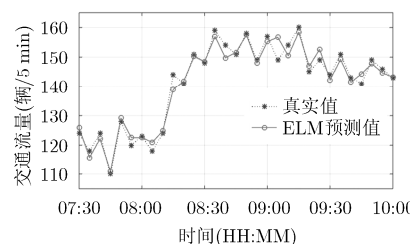


图7 第1组数据 ELM 预测结果

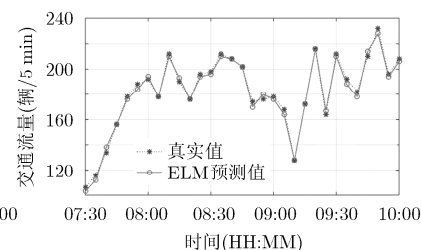


图8 第2组数据 ELM 预测结果

基于组合模型的预测算法利用 K-means 算法对第1组数据前3天的数据分别进行聚类，聚类数目为3，利用每次聚类后的聚类中心计算时间基线，结果如表4所示。

由3次时间基线求均值可得 $t_1=7:53:10$, $t_2=9:07:10$ ，交通流量数据被分割为3个区域：7:00~7:50, 7:55~9:05, 9:10~10:00。对每个数据域使用 ELM 模型进行循环预测仿真，找到的最优隐含层神经元个数分别为7, 19和16, 预测结果如图9, 图10, 图11所示。

第1组数据区域1预测结果的均方误差为0.24276, 决定系数为0.99619, 建模时间为0.015625。第1组数据区域2预测结果的均方误差为0.51139, 决定系数为0.99803, 建模时间为0.046875。第1组数据区域3预测结果的均方误差为0.28868, 决定系数为0.99115, 建模时间为0.03125。

基于组合模型的预测算法利用 K-means 算法对第2组数据前两天的数据分别进行聚类，聚类数目为3，利用每次聚类后的聚类中心计算时间基线，结果如表5所示。

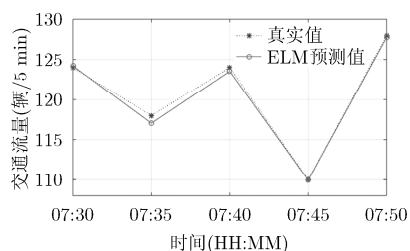


图9 第1组数据区域1预测结果

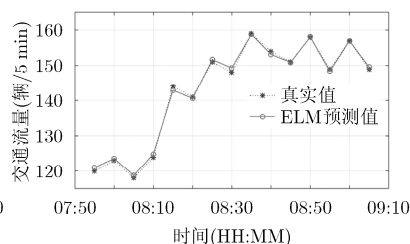


图10 第1组数据区域2预测结果

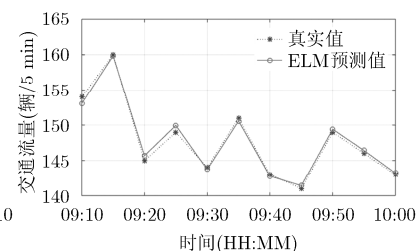


图11 第1组数据区域3预测结果

表4 第1组数据聚类结果

聚类次数	第1次	第2次	第3次
聚类中心	(7:09:00 90.4832)	(7:09:00 92.3636)	(7:09:00 91.7455)
	(7:46:30 120.5321)	(7:46:30 121.5047)	(7:46:30 117.5136)
	(9:07:30 152.3636)	(9:07:30 151.4800)	(9:07:30 153.4218)
时间基线	7:52:30	7:53:00	7:54:00
	9:07:00	9:09:00	9:05:30

表 5 第 2 组数据聚类结果

聚类次数	第 1 次	第 2 次
聚类中心	(7:13:00 156.2826)	(7:13:00 155.4316)
	(8:05:00 189.4513)	(8:05:00 189.5467)
	(9:17:30 198.6437)	(9:17:30 201.3636)
时间基线	8:07:00	8:05:00
	9:18:00	9:17:00

由两次时间基线求均值可得 $t_1=8:05:50$, $t_2=9:17:50$, 交通流量数据被分割为 3 个区域: 7:30~8:05, 8:10~9:15, 9:20~10:00。对每个数据域使用 ELM 模型进行循环预测仿真, 找到的最优隐含层神经元个数分别为 7, 19 和 16, 预测结果如图 12, 图 13, 图 14 所示。

第 2 组数据区域 1 预测结果的均方误差为 0.23768, 决定系数为 0.997039, 建模时间为 0.015825。第 2 组数据区域 2 预测结果的均方误差为 0.52019, 决定系数为 0.99312, 建模时间为 0.041025。第 2 组数据区域 3 预测结果的均方误差为 0.28868, 决定系数为 0.99115, 建模时间为 0.03125。

5.3 评测指标

从决定系数(R^2), 均方误差(MSE), 建模时间

(t)3 个方面出发对 BP, ELM 和基于交通流量序列分割和极限学习机组合模型进行比较, 如表 6, 表 7 所示。

由以上表 6, 表 7 可得, 基于交通流量序列分割和极限学习机组合模型的预测算法决定系数更接近于 1, 说明该模型拟合优度大, 自变量对因变量的解释程度高, 具有更高的参考价值。该模型具有更小的均方误差, 均方误差为 BP 的 1/50, ELM 的 1/20, 说明预测模型描述实验数据具有更好的精确度。模型相对于 BP 神经网络有较短的建模时间, 约为 BP 神经网络的 1/10, 与 ELM 相比时间略长, 约为 ELM 的 4 倍, 但建模时间在可接受范围内。总体来说, 基于交通流量序列分割和极限学习机组合模型在交通流量预测方面比其他两种模型有更大的优势。

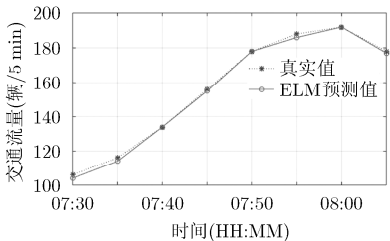


图 12 第 2 组数据区域 1 预测结果

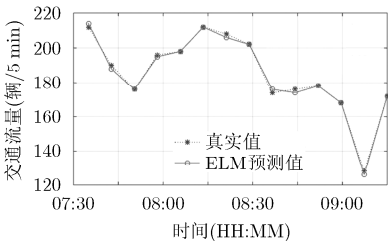


图 13 第 2 组数据区域 2 预测结果

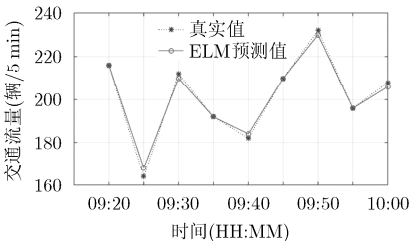


图 14 第 2 组数据区域 3 预测结果

表 6 第1组数据下3种模型比较结果

模型	BP 神经网络	ELM	基于交通流量序列分割和极限学习机组合模型
MSE	16.966200	6.444700	0.347610
R^2	0.926420	0.967330	0.995123
t	0.968750	0.031250	0.133750

表 7 第2组数据下3种模型比较结果

模型	BP 神经网络	ELM	基于交通流量序列分割和极限学习机组合模型
MSE	15.96265	8.43261	0.34835
R^2	0.92315	0.95890	0.99453
t	0.92710	0.03421	0.08410

6 结束语

针对短时十字路口交通流量预测问题, 本文提出了基于交通流量序列分割和极限学习机组合模型预测算法 TFPBCM。仿真实验证明, 与单一的 BP 神经网络和 ELM 相比, 该组合模型算法更优, 结果更精确, 在一定程度上能够满足短时交通流量预测的需要。虽然, TFPBCM 算法模型更优于单一的 BP 神经网络和 ELM, 但是该模型也存在着一些不足, 由于神经网络结构复杂, 导致 TFPBCM 算法模型的网络结构不易选择, 隐含层的节点个数在训练前由用户预先设定, 一般通过多次实验获得最优节点数, 对于一些大型的应用而言, 会有较高的时间复杂度, K-means 算法在初始敏感度和聚类中心选择上都有一定的不足。在后续的工作中, 会进一步研究并讨论这些问题, 找到解决方案, 优化 TFPBCM 算法模型。

参考文献

- [1] 兰云. 短时交通流量预测研究[D]. [硕士学位论文], 西北工业大学, 2002.
LAN Yun. The study on short time traffic flow prediction[D]. [Master dissertation], Northwestern Polytechnical University, 2002.
- [2] 范腾腾. 城市道路交通流量短时预测研究[D]. [硕士学位论文], 北京交通大学, 2012.
FAN Tengteng. The research of urban road traffic flow short-term prediction[D]. [Master dissertation], Beijing Jiaotong University, 2012.
- [3] 覃明贵. 城市道路交通数据挖掘研究与应用[D]. [博士学位], 复旦大学, 2010.
QIN Minggui. Research and application on the urban traffic data mining[D]. [Ph.D. dissertation], Fudan University, 2010.
- [4] 曹虹. 基于 BP 神经网络的交通流量预测[D]. [硕士学位论文], 长安大学, 2012.
CAO Hong. Traffic flow prediction based on BP neural network[D]. [Master dissertation], Chang'an University, 2012.
- [5] 曹炜. 基于 BP 神经网络的城市交通流量预测与路口分析[D]. [硕士学位论文], 昆明理工大学, 2006.
CAO Wei. The city traffic flow prediction and cross analysis based on BP neural network[D]. [Master dissertation], Kunming University of Science and Technology, 2006.
- [6] 尚宁, 覃明贵, 王亚琴, 等. 基于 BP 神经网络的路口短时交通流量预测方法[J]. 计算机应用与软件, 2006, 23(2): 32-33, 57.
SHANG Ning, QIN Minggui, WANG Yaqin, et al. A BP neural network method for short-term traffic flow forecasting on crossroads[J]. *Computer Applications and Software*, 2006, 23(2): 32-33, 57.
- [7] 翟敏. 极限学习机的自适应网络结构选择方法研究[D]. [硕士学位论文], 西北大学, 2014.
ZHAI Min. Research on adaptive network structure selection method for extreme learning machine[D]. [Master dissertation], Northwest University, 2014.
- [8] 甘露. 极限学习机的研究与应用[D]. [硕士学位论文], 西安电子科技大学, 2014.
GAN Lu. Research and application of extreme learning machine[D]. [Master dissertation], Xidian University, 2014.
- [9] 王智慧. BP 神经网络和 ELM 算法研究[D]. [硕士学位论文], 中国计量学院, 2012.
WANG Zhi-hui. Research on BP neural networks and ELM Algorithms[D]. [Master dissertation], China Jiliang University, 2012.
- [10] 张一. 基于卡尔曼滤波的短时交通流量预测模型研究[D]. [硕士学位论文], 沈阳工业大学, 2014.
ZHANG Yi. Research of short-term traffic volume prediction based on kalman filtering[D]. [Master dissertation], Shenyang University of Technology, 2014.
- [11] 罗向龙. 交通流量经验模态分解与神经网络短时预测方法[J]. 计算机工程与应用, 2010, 46(26): 212-214. doi: 10.3778/j.issn.1002-8331.2010.26.066.
LUO Xianglong. Short-term traffic flow prediction method based on EMD and artificial neural network[J]. *Computer Engineering and Applications*, 2010, 46(26): 212-214. doi: 10.3778/j.issn.1002-8331.2010.26.066.
- [12] 沈永增. 基于混沌粒子群优化小波神经网络的短时交通流预测[J]. 计算机应用与软件, 2014, 31(6): 84-87. doi: 10.3969/j.issn.1000-386x.2014.06.024.
SHEN Yongzeng. Short-time traffic flow forecast based on WNN optimised by CPSO[J]. *Computer Applications and Software*, 2014, 31(6): 84-87. doi: 10.3969/j.issn.1000-386x. 2014.06.024.
- [13] 朱征宇. 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.
ZHU Zhengyu. Short-term traffic flow forecasting model combining SVM and Kalman filtering[J]. *Computer Science*, 2013, 40(10): 248-251.
- [14] TCHRAKIAN T T, BASU B, and O'MAHONY M. Real-time traffic flow forecasting using spectral analysis[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2012, 13(2): 519-526.
- [15] CHEN Syuanyi and CHOU Weiyao. Short-term traffic flow prediction using EMD-based recurrent Hermite neural network approach[C]. 15th International IEEE Conference on Intelligent Transportation Systems, Anchorage, 2012: 1821-1826.
- [16] 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.
- [17] XU Yanyan, KONG Qingjie, and LIU Yuncai. Short-term traffic volume prediction using classification and regression trees[C]. IEEE Intelligent Vehicles Symposium(IV), Gold Coast, 2013: 493-498.
- [18] WILLIAMS B M. Modeling and forecasting vehicular traffic flow as a seasonal stochastic time series process[D]. [Ph.D. dissertation], University of Virginia, 1999: 243-246.

芮兰兰: 女, 1979 年生, 博士, 副教授, 研究方向为网络和业务质量管理, 泛在网络, 大数据等。

李钦铭: 男, 1990 年生, 硕士生, 研究方向为大数据流处理。