

文章编号:1671-6833(2011)04-0060-04

融合 BP 神经网络与 ARIMA 的短时交通流预测

曾庆山, 全书鹏, 靳志强

(郑州大学 电气工程学院, 河南 郑州 450001)

摘 要:为了能在交通管理中提前采取措施规避可能存在的交通拥挤或堵塞,提出了一种高效可靠的短时交通流预测算法.首先采用 BP 神经网络与自回归求和滑动平均(ARIMA)两种方法分别建立单项预测子模型,再以 BP 神经网络作为最优非线性组合模型的逼近器,建立组合预测模型,对单项预测子模型的预测值进行融合,由此得到最终的预测结果.通过 MATLAB 与 SPSS 平台对实测交通流量数据进行仿真分析,结果表明,该种组合预测方法是切实可行的.

关键词:智能交通;短时交通流;预测;时间序列;神经网络

中图分类号: TP183

文献标志码: A

0 引言

交通流预测是支撑交通控制系统、交通诱导系统等 ITS 子系统的关键技术之一.交通管理与控制效果的好坏,在于是否能够根据当前路网交通流变化规律,合理推测未来一定时间范围内交通流的新动向,从而提前采取有效措施规避可能存在的交通拥挤或堵塞.因此,对短时交通流预测技术的研究具有重要意义和应用价值.短时交通流预测是对交通流的下一步预测,即在时刻 t 对下一决策时刻($t + \Delta t$)的交通流做出短期实时预测,预测时间跨度一般不超过 15 min^[1].几十年来,世界各国的专家和学者研究出了各种预测模型用于短时交通流预测,主要可以分为基于统计方法的模型、交通仿真模型、神经网络模型、综合模型等七类模型.任何一个模型都有其优缺点及适用条件^[2-3],由于单个模型有其局限性,所以用综合模型进行预测是交通流预测领域的发展趋势.

随着国内 ITS 发展对交通信息预测精度及实时性要求的不断提高,组合预测理论成了交通科技工作者们关注的热点.文献[4]将神经网络集成技术引入交通量预测,通过训练多个神经网络并将各网络输出进行合成,从而提高学习系统的

泛化能力.文献[5]根据交通流量具有周相似的特性,构造了周相似序列,提出用指数平滑法和神经网络分别进行周相似序列与残差预测的方法.由于没有考虑到交通流量时间序列既具有明显的短期相关性,又具有很大的随机性、不确定性的特征,这些方法在交通流预测中具有其局限性.为此,笔者提出了一种融合 BP 与 ARIMA 的非线性组合预测算法来对城市道路的交通流进行短时预测,从而进一步改善短时交通流量的预测效果.

1 单项预测子模型

1.1 BP 神经网络模型

由于 BP 网络具有拟合任意非线性函数的功能,因此可以利用 BP 网络来代替复杂的非线性数学关系式,实现对现实交通系统的较准确刻画.基于 BP 网络的交通流量预测算法可归纳如下^[6-7]:

- (1)原始数据采集及预处理;
- (2)设计网络结构,确定网络层数以及各层神经元的个数;
- (3)利用原始数据构造训练样本对;
- (4)设计学习算法训练神经网络;
- (5)对于训练满足要求的 BP 网络,输入一组已知数据得到相应的交通信息预测值.

收稿日期:2011-01-11;修订日期:2011-04-17

基金项目:河南省自然科学基金资助项目(0611054100).

作者简介:曾庆山(1963-),男,湖北武汉人,郑州大学副教授,博士,研究方向为复杂系统的分析、建模与控制, E-mail: qszeng@zzu.edu.cn.

基于 BP 网络的交通流量预测模型如图 1 所示,选用 3 层结构的 BP 网络. 利用前 5 个时刻的交通流量检测值来预测下一时刻的交通流量. BP 网络的隐层神经元个数按经验 $(2n + 1)$ 选取,在这里即为 11 个.

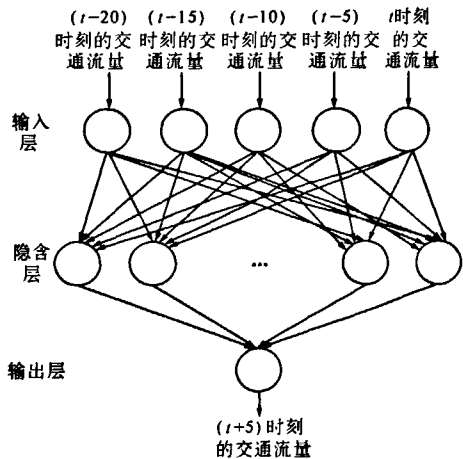


图 1 BP 网络预测模型结构图
Fig.1 Structure of BP model

1.2 ARIMA 预测子模型

具有如下结构的模型称为求和自回归移动平均模型,简记为 ARIMA(p, d, q) 模型:

$$\Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t$$
$$E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t$$
$$Ex_s \varepsilon_t = 0, \forall s < t$$

式中: B 为延迟算子; $\nabla^d = (1 - B)^d$, $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$.

ARIMA 模型的实质是差分运算与 ARMA 模型的组合. ARMA 模型的相关性特征如表 1 所示, 据此便可进行 ARMA 模型的识别.

表 1 ARMA 模型的相关性特征
Tab.1 Correlation of ARIMA model

模型	自相关系数	偏自相关系数
AR(P)	拖尾	P 阶截尾
MA(q)	q 阶截尾	拖尾
ARMA(p, q)	拖尾	拖尾

用 ARIMA 模型对观察序列进行建模的流程如图 2 所示.

ARIMA 模型的建立通过统计学软件平台 SPSS 16.0 来进行. 在该软件的 Time Series 模块中, 依据上文所述建模流程和 ARMA 模型的相关性特征, 便可拟合出合适的 ARIMA 模型, 进而实现交通流量的预测.

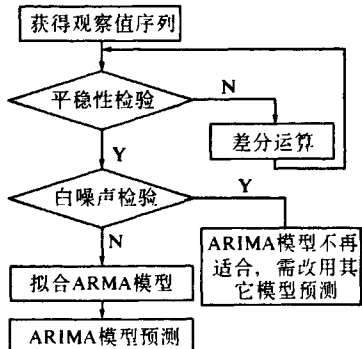


图 2 ARIMA 模型建模流程图
Fig.2 Flow chart to model ARIMA

2 组合预测模型

组合预测就是把不同的预测模型组合起来, 综合利用各种预测方法所提供的信息, 以适当的方式得出组合预测模型和最佳预测结果, 从而达到提高预测精度和增加预测可靠性的效果. 其本质就是综合利用多种预测方法的特点, 将分项预测信息加以适当融合得出更为准确的预测结果.

对于短时交通流量这一微观交通信息序列 $X = (x_1, x_2, L, x_N)^T$, 选取与其动力学特征相对应的 K 种预测方法 (K 值视预测对象的具体特点而定), 其中, 通过第 i 种方法对第 j 期的预测值为: $f_{ij} (i = 1, 2L, K; j = 1, 2L, N)$, 记向量 $F_j = (f_{1j}, f_{2j}, L, f_{Kj})$. 则可以构造出预测对象的最优非线性组合预测模型 $x_j = g(F_j)$, 对单项预测子模型的预测结果进行融合, 并使其满足性能指标最小化: $\min J = \sum_{j=1}^N (x_j - g(F_j))^2$. 基于上述思想, 笔者采用一种融合了 BP 神经网络与 ARIMA 的非线性组合预测算法, 即在使用 BP 神经网络和 ARIMA 两种方法对其进行预测的基础上, 再次使用 BP 神经网络来逼近上述非线性函数关系 $x_j = g(F_j)$, 从而进行短时交通流量的组合预测, 实现预测值和实际值的最佳逼近. 组合预测模型的结构如图 3 所示:

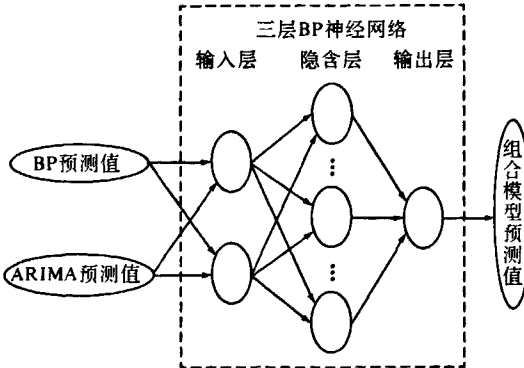


图 3 组合预测模型结构图
Fig.3 Structure of combination model

在利用组合模型进行预测时,在一段时间内,可以认为 BP 神经网络与 ARIMA 两种子模型的加权系数是不变的.每隔一段时间,重新训练一次 BP 神经网络,便可动态地调整各子模型的权重.即组合模型中对两个单项预测值的加权融合是在滚动变化的,从而就克服了全天交通流量变化规律差别较大的问题.

3 模型的仿真实验

应用数据 1 来源于检测器采集的全天交通流量,选择包含高峰时段在内的 7:30-16:00 之间的数据,数据间隔为 5 min.取其中 7 月 18 号的路段交通流量进行分析预测.

3.1 仿真实验过程

方法 1:单独使用 ANN 模型进行预测.把 7 月 18 号采集到的早 7:30 到下午 16:00 之间的交通流量数据看成一时间序列,由于时间间隔为 5 min,则该序列长度为 102.把前 78 个数据作为训练样本,后 24 个数据作为测试样本,即用前 78 个数据来训练建立 BP 神经网络模型,预测随后的 24 个数据.其中神经网络的输入数据为 $t-20$ 、 $t-15$ 、 $t-10$ 、 $t-5$ 、 t 时刻的交通流量;输出数据为 $t+5$ 时刻的交通流量.

方法 2:单独使用 ARIMA 模型进行预测.首先对原始数据进行平稳化处理,然后通过时间序列的自相关函数和偏自相关函数图以及 BIC 准则,确定模型参数 (p, d, q) ,进而进行预测.这一过程使用 SPSS 软件中的 ARIMA 子程序来实现.采用 7:35 到 14:00 的实测数据来预测下一时刻 14:05 的交通流量,然后利用 7:40 到 14:05 的实测数据预测 14:10 的流量.依次类推,经过滚动预测,计算 14:00 到 16:00 的预测数据,共计 24 个预测值.

方法 3:使用组合模型进行预测.组合预测模型中 BP 网络的训练样本来自于单项预测子模型.取步长为固定值 8,即认为在一定时间内各子模型的权重系数不变,基于此,每隔 8 个时段,重新训练一次组合模型中的 BP 神经网络,便可进行组合模型的滚动预测.由于数据有限,这里只拿 14:00 到 16:00 的车流量数据来进行组合模型的创建与验证分析.取前 16 组预测值和对应时刻的实测值来训练建立 BP 神经网络,然后输入单项预测子模型的“预测值对”,便可得到相应时刻组合预测模型的预测结果.

3.2 性能评估与可行性分析

使用绝对百分比误差(APE)和平均绝对百分比误差(MAPE)来评估算法的预测性能

$$APE = \frac{|y_i - \hat{y}_i|}{y_i} \times 100\%$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{y_i} \times 100\%$$

式中: \hat{y}_i 是交通流量的预测值; y_i 则是交通流量的实测值; n 为预测点的个数.

用以上 3 种方法(ANN、ARIMA、组合模型)得到的预测结果与实测数据的对比见图 4,3 种方法的百分比绝对误差见图 5,平均百分比绝对误差见表 2.

由图 4、图 5 和表 2 可得:组合模型的预测误差与单一子模型相比有了明显减小,同时组合模型的 MAPE 最小,这表明组合模型的预测效果最好.由于在 15:35 这一时刻,交通流量时间序列出现了明显的波动,使得该点附近的预测误差偏大,这也刚好体现出现实交通流系统的复杂、不确定性特征.由表 2 可见,组合预测模型的平均绝对百分比误差在 5% 以内,即总体上来说平均误差在可接受的范围之内.

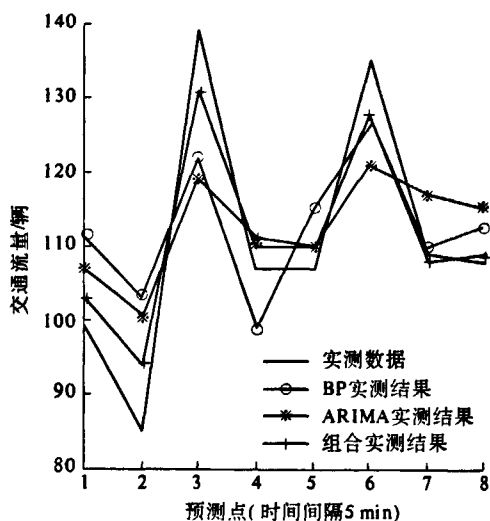


图 4 预测结果与实测数据对比图

Fig. 4 Comparison of forecasting and actual data

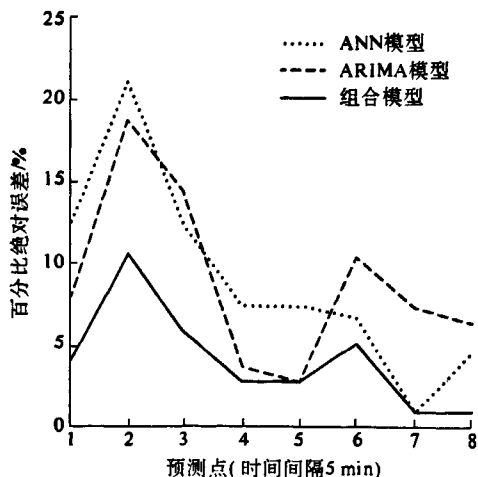


图 5 3 种预测方法百分比绝对误差对比图

Fig. 5 Comparison of three model's APE

表 2 3 种方法的平均百分比绝对误差对比
Tab.2 Comparison of three model's MAPE

预测方法	ANN 模型	ARIMA 模型	组合模型
MAPE/%	9.086 9	9.003 3	4.127 5

4 结论

现实交通流系统是一个非线性的大系统,加之来自环境、行人等随机干扰的影响,交通流的变化过程具有复杂性和不确定性.单纯使用某一种预测方法,难以得到满意的预测结果.笔者采用了一种基于 BP 神经网络的非线性组合预测算法,在可预测性分析的基础上,选取一组对应的预测方案和方法,通过 BP 网络对二者的预测值进行融合,从而得到最终的预测结果,在一定程度上改善了短时交通流量的预测效果.影响交通流这一非线性系统的因素很多,且前期子模型的选取对组合模型的实施起着至关重要的作用,因此,如何选取合适的前期预测子模型以及子模型的个数,使预测结果更加精确是一个值得进一步研究的问题.

参考文献

[1] SMITH B L,DEMETSKY M J. Traffic flow forecasting: comparison of modeling approaches [J]. Journal of Transportation Engineering,1997,123(4):261-266.

[2] 王进,史其信.短时交通流预测模型综述[J].中国公共安全·学术卷·智能交通,2005,6(1):92-98.

[3] MANOEL C N. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions [J]. Expert Systems with Applications, 2009(36): 6164-6173.

[4] 张益,陈淑燕,王炜.短时交通流量时间序列智能复合预测方法概述[J].公路交通科技,2006,23(8):139-142.

[5] 谭满春,李英俊,关占荣,等.周相似特性下交通流组合预测方法研究[J].计算机工程与应用,2007,43(33):193-195.

[6] 陈雪平,曾盛,胡刚.基于 BP 神经网络的短时交通流预测[J].公路交通技术,2008,6(3):115-117.

[7] 李亦芳,程万里,刘建厅.基于人工神经网络与回归分析的水质预测[J].郑州大学学报:工学版,2008,29(1):106-109.

Short-term Traffic Flow's Forecasting by Fusing BP Neural Network and ARIMA

ZENG Qing-shan, QUAN Shu-peng, JIN Zhi-Qiang

(School of Electric Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: In order to take measures in advance to avoid traffic congestion or traffic jamming that possibly exists in traffic management, a short-term traffic flow forecasting algorithm by fusing BP net and ARIMA is put forward in this paper. Firstly, the BP neural network and autoregressive integrated moving average (ARIMA) are used to establish the single forecasting sub-model separately. Then BP net is used as the approximator of the optimal nonlinear combination model to establish the combination forecasting model that can fuse the predicted value from the single forecasting sub-model. Lastly, the final forecasting result is obtained. By using MATLAB and the SPSS platform, the simulation based on the actual traffic flow data is realized, and the results show that the combination forecasting technique is practical and feasible.

Key words: intelligent transportation; short-term traffic flow; forecasting; time series; neural network