

文章编号: 1009-6744 (2010) 01-0145-07

基于 Elman 神经网络的道路网短时交通流预测方法

董春娇, 邵春福*, 熊志华, 李娟

(北京交通大学 城市交通复杂系统理论与技术教育部重点实验室, 北京 100044)

摘要: 以道路子网为研究对象, 采用 Elman 神经网络实现道路网多断面交通流短时预测。首先通过提取交通流空间特性对道路网进行划分, 降低道路网整体分析复杂度及解空间维数, 提高交通流预测的计算精度和效率; 其次以实时采集的交通流数据为基础, 并以重构的交通流时间序列作为输入, 采用 Elman 神经网络实现道路网多断面交通流同时预测; 最后, 基于城市快速路多断面交通流量数据对短时交通流预测方法进行验证, 并与 BP 神经网络预测结果进行对比分析。验证结果表明, 本文提出的道路网划分方法能够划分出满足预测需求的子路网, 在划分的子路网上, 应用 Elman 神经网络能够实现道路网多断面同时预测, 且预测效果优于 BP 神经网络。

关键词: 城市交通; 交通流短时预测; 道路网划分; 广义空间距离; Elman 神经网络; BP 神经网络

中图分类号: U491

文献标识码: A

Short-Term Traffic Flow Forecasting of Road Network Based on Elman Neural Network

DONG Chun-jiao, SHAO Chun-fu, XIONG Zhi-hua, LI Juan

(MOE Key Laboratory for Transportation Complex Systems Theory and Technology,
Beijing Jiaotong University, Beijing 100044, China)

Abstract: The methodology of short-term traffic flow forecasting is presented in this paper based on Elman neural network, which sets sub-network as objectives. To simplify road network analysis, and to reduce dimension of solution space, this paper extracts space characteristics of traffic flow and split the road network more scientific and rational based on general space distance. Then, it introduces Elman neural network to forecast traffic flow of multi-sections in the road network whose input vector is constructed by time series of traffic flow. In the last section, the methodology is tested using traffic flow data from the road network, which is compared to the method of BP neural network. The methodology can split the road network into several sub-networks satisfying short-term forecasting demand, and short-term traffic flow forecasting of the road network is realized by Elman neural network. The results are proved to be super to the method of BP neural network.

Key words: urban traffic; short-term traffic flow forecasting; road network splitting; general space distance; Elman neural network; BP neural network

CLC number: U491

Document code: A

收稿日期: 2009-04-15 修回日期: 2009-07-11 录用日期: 2009-08-31

基金项目: 国家自然科学基金项目 (50578015); “973” 国家重点基础研究发展规划项目 (2006CB705505)。

作者简介: 董春娇 (1982 -), 女, 辽宁大石桥人, 博士生。

* 通讯作者: cfshao@center.njtu.edu.cn

1 引言

随着经济的快速发展,城市交通量急剧增加,使得交通拥挤与交通堵塞日趋严重. 20世纪80年代以来,发达国家就开始道路交通系统的管理与控制技术的研发,智能交通系统(Intelligent Transportation Systems, ITS)应运而生. 交通控制与交通诱导是智能交通系统研究的重要领域,实现交通控制和诱导的关键问题之一是实时准确的短时交通流预测.

从20世纪60年代开始,人们就开始把其他领域应用成熟的预测模型用于短时交通流预测领域,开发了多种预测模型和方法. 较早期的预测方法有:自回归模型(AR)、滑动平均模型(MA)、自回归滑动平均模型(ARMA)、历史平均模型(HA)和Box-Cox法等等^[1,2]. 随着该领域研究的逐渐深入,又出现了一批更复杂、精度更高的预测方法. 从表现形式上,可分为基于确定的数学模型方法和无模型算法两大类. 前者包括多元回归模型、ARIMA模型、自适应权重联合模型、Kalman滤波模型、基准函数—指数平滑模型、UTCS-2(3)模型,以及由这些模型构成的各种组合预测模型等;后者则包括非参数回归、KARIMA算法、谱分析法、状态空间重构模型、小波网络、基于多维分形的方法、基于小波分解与重构的方法,以及多种与神经网络相关的复合预测模型等^[3-8]. 基于数学解析模型的方法由于模型本身的局限性,难以处理随机干扰因素对交通流的影响,因而无法反映交通流系统本身的高度不确定性与非线性,预测精度不高;基于知识的智能模型预测方法,通过方法本身的结构机制来获取预测的“经验”、“知识”,以预测下一周期的流量,具有一定的自适应能力,但没有结合交通流本身的特性予以考虑和处理,影响了其预测效果.

尽管以上这些方法能够在一定程度上减少建模和预测的困难,但是,大部分模型还不能实现路网范围内的交通流预测,一些方法仅仅能够实现单一断面交通流的预测. 本文提出一种基于Elman神经网络的短时交通流预测方法,该方法基于路网划分结果,并以重构的交通流量时间序列作为输入,采用Elman网络实现道路网多断面交通流量同时预测.

2 基于交通流时间特性的道路网划分

智能交通控制和诱导系统需要的是路网范围

内的交通流短时预测,然而,以整个路网作为研究对象,不仅增加建模难度,同时会降低计算效率. 因此,本文在考虑交通流空间特性的基础上,将研究路网划分成满足分析需要的子路网.

断面空间相关性受到各种因素的影响,本文利用广义空间距离来解释断面空间相关性,各影响因素与广义空间距离之间的关系如图1所示.

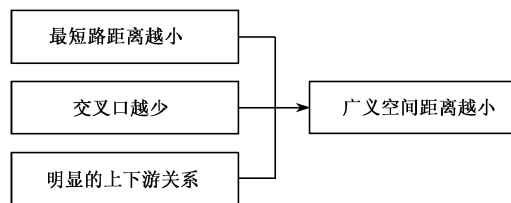


图1 广义空间距离与其影响因素关系图

Fig.1 Relation graph between general space distance and influence factor

由图1可以看出,广义空间距离与各断面之间最短距离和交叉口数目成正比,与断面之间的上下游关系成反比. 根据以上分析,提出用于度量路网中各断面空间相关性的广义空间距离模型如式(1)所示:

$$D_{ij} = \frac{1}{\alpha_{ij} + 1} \left(d_{ij} + \sum_k n_{ij}^k L_{ij}^k \right) \quad (1)$$

式中 D_{ij} —— i, j 断面之间的广义空间距离;

α_{ij} —— i, j 断面之间上下游距离系数,无上下游关系, $\alpha_{ij} = 0$;

d_{ij} —— i, j 断面之间最短距离;

n_{ij}^k —— i, j 断面之间 k 型交叉口个数, k 为交叉口类型变量,按照统计学定类变量的定义方法,令 T 型和 Y 型交叉口的 k 值为 3,十字型交叉口的 k 值为 4,其他类型交叉口的 k 值为 5;

L_{ij}^k —— i, j 断面之间由于 k 型交叉口的存在带来的附加距离, $L_{ij}^k = (k - 1) \min(l_i, l_j)$, 其中 l_i, l_j 分别代表交叉口距 i, j 断面的距离.

广义空间距离的计算就是为了更好的反映各断面之间的空间相关性,相关性强的断面,广义空间距离就小,相关性弱的断面,广义空间距离就大. 根据广义空间距离模型,可以获得度量各断面之间的相对位置参数,引进 R-Q 型因子分析法,对其进行进一步分析.

R-Q 型因子分析法,又称为对应分析(Correspondence Analysis),由法国统计学家 Benzecri 于 20

世纪 70 年代初提出,该方法是在因子分析的基础上发展起来.对应分析从编制两变量的交叉列联表入手,并通过交叉列联表来进一步分析和探索变量之间的关系.主要方法是将定类或定序型变量转换为可度量的分值,并通过降低变量的维度作分值分布图.对应分析结合了因子分析和多维尺度方法的优点.在降低维度方面,对应分析与因子分析类似;在作分布图方面,对应分析与多维尺度方法类似,这些特点是传统的统计方法所不具备的.

基于广义空间距离的 R-Q 型因子分析的主要步骤包括:

Step 1 编制广义空间距离二维列联表并计算二维频率矩阵.

设有断面 $A_i (i = 1, 2, \dots, m), B_j (j = 1, 2, \dots, n)$, 则可以得到 $m \times n$ 广义空间距离的二维列联表 D , 将列联表转化为频率矩阵 F .

$$f_{ij} = \frac{D_{ij}}{D} = \frac{D_{ij}}{\sum_{i=1}^m \sum_{j=1}^n D_{ij}} \quad (2)$$

式中 f_{ij} ——断面 i 和 j 的联合频率.

Step 2 对频率矩阵进行标准化,得到过渡矩阵.

对二维频率矩阵 F 进行如下变换:

$$z_{ij} = \frac{f_{ij} - f_i f_j}{\sqrt{f_i f_j}} \quad (3)$$

得到过渡矩阵 $Z = (z_{ij})$.

Step 3 利用协方差矩阵进行因子分析,从而达到降维的目的.

可以证明, m 个断面之间的协方差矩阵 $\Sigma_m = ZZ'$ 和 n 个断面之间的协方差矩阵 $\Sigma_n = Z'Z$ 有完全相同的非零特征根,记为 $\lambda_1 > \lambda_2 > \dots > \lambda_k, k \leq \min\{m, n\}$. 设 u_1, u_2, \dots, u_k 为相对特征根 $\lambda_1 > \lambda_2 > \dots > \lambda_k$ 的关于断面 B_j 各水平之间斜方差矩阵的特征向量,则 Zu_1, Zu_2, \dots, Zu_k 就是相对特征根 $\lambda_1 > \lambda_2 > \dots > \lambda_k$ 的关于断面 A_i 各水平之间协方差矩阵 Σ_m 的特征向量. 这样,得到关于 B_j 的 n 个断面因子载荷矩阵 F , 以及关于 A_i 的 m 个断面因子载荷矩阵 G .

$$F = \begin{bmatrix} u_{11} \sqrt{\lambda_1} & u_{12} \sqrt{\lambda_2} & \dots & u_{1m} \sqrt{\lambda_m} \\ u_{21} \sqrt{\lambda_1} & u_{22} \sqrt{\lambda_2} & \dots & u_{2m} \sqrt{\lambda_m} \\ \vdots & \vdots & & \vdots \\ u_{n1} \sqrt{\lambda_1} & u_{n2} \sqrt{\lambda_2} & \dots & u_{nm} \sqrt{\lambda_m} \end{bmatrix}$$

$$G = \begin{bmatrix} Zu_{11} \sqrt{\lambda_1} & Zu_{12} \sqrt{\lambda_2} & \dots & Zu_{1n} \sqrt{\lambda_n} \\ Zu_{21} \sqrt{\lambda_1} & Zu_{22} \sqrt{\lambda_2} & \dots & Zu_{2n} \sqrt{\lambda_n} \\ \vdots & \vdots & & \vdots \\ Zu_{m1} \sqrt{\lambda_1} & Zu_{m2} \sqrt{\lambda_2} & \dots & Zu_{mn} \sqrt{\lambda_n} \end{bmatrix}$$

Step 4 根据因子载荷图绘制 R-Q 型因子分析图. 在一张二维平面图上以公共因子为坐标轴,分别以断面的因子载荷为坐标绘制 R-Q 型因子分析图,从而可以直观地观察各断面之间的空间相关关系.

Step 5 观察图中各断面之间的空间相关关系,如最大子网包括的断面满足预测需求,则停止,否则转到 Step1. 通常情况下,最大子网包括的断面数应不大于 7.

3 基于 Elman 神经网络的交通流短时预测

Elman 神经网络是 Elman 于 1990 年提出的,该模型在前馈网络的隐含层中增加一个承接层,作为延时算子,以达到记忆的目的,从而使系统具有适应时变特性的能力,能直接反映动态过程系统的特性^[9].

Elman 神经网络一般分为 4 层:输入层、中间层(隐含层)、承接层和输出层. 输入层的单元仅起信号传输作用,输出层单元起线性加权作用. 隐含层单元的传递函数可采用线性或非线性函数,承接层又称为状态层,用来记忆隐含层单元前一时刻的输出值并返回给输入,可以认为是一个一步延时算子. Elman 神经网络的特点是隐含层的输出通过承接层的延迟与存储,自联到隐含层的输入. 这种自联方式使其对历史状态的数据具有敏感性,内部反馈网络的加入增加了网络本身处理动态信息的能力,从而达到了动态建模的目的.

基于 Elman 神经网络的交通流短时预测非线性状态空间表达式如式(4)、式(5)和式(6)所示.

$$y(k) = g(w^3 x(k)) \quad (4)$$

$$x(k) = f(w^1 x_c(k) + w^2(q(k-1))) \quad (5)$$

$$x_c(k) = x(k-1) \quad (6)$$

其中 y, x, q, x_c 分别表示 m 维输出节点向量、 n 维中间层节点单元向量、 r 维经过相空间重构技术处理过的子路网上多断面交通流量时间序列和 n 维反馈状态向量. w^3, w^2, w^1 分别表示中间层到输出层、输入层到中间层、承接层到中间层的连接权

值. $g(\cdot)$ 为输出神经元的传递函数, 是中间层输出的线性组合. $f(\cdot)$ 为中间层神经元的传递函数, 常采用 S 函数. Elman 网络学习指标函数采用误差平方和函数如式(7)所示.

$$E(w) = \sum_{k=1}^n [y_k(w) - \tilde{y}_k(w)]^2 \quad (7)$$

其中 $\tilde{y}_k(w)$ 为目标输出向量.

在 Elman 神经网络中, 稳定性与其联想记忆的能力密切相关, 因此, 稳定性是一个重要的研究方向. 为了提高 Elman 神经网络的稳定性, 本文采用 traindx 作为训练函数, 该函数为动量及自适应的梯度递减训练函数 (Gradient Descent Backpropagation with Adaptive Learning Rate); 以梯度下降动量学习函数 Learngdm 作为学习规则, 利用神经元的输入和误差、权值或阈值的学习速率和动量常数, 来计算权值或阈值的变化率; 分别以 tansig 函数、purelin 函数作为隐层节点传递函数和输出层传递函数.

用训练样本集合对 Elman 网络模型的参数进行标定, 包括隐层节点数量和权值. 对于具有一个隐层的 Elman 网络, 其输出、输入节点数是由所研究的问题决定的, 只有隐层节点的数量是需要特殊确定的. 隐层节点过多, 会造成 Elman 网络结构庞大, 学习时间长; 隐层节点过少, 则难以保证预测对象的精度要求. 因此, 选择合适的隐层节点个数是 Elman 网络设计的重要因素. 如果能够为 Elman 网络选择出最佳的隐层神经元个数, 则会在保证预测精度的同时, 减少 Elman 网络学习和再学习的时间. 目前, 尚无普遍适用的隐层节点确定方法, 绝大部分的 Elman 网络都采用经验方法确定隐层节点数量.

根据短时交通流量所具有的混沌特性^[10], 确定 Elman 网络输入节点为 5, 输出节点为 1, 隐层节点数量由试算法确定. 根据已有经验, 初步设定隐层节点数范围为 16 ~ 36, 然后用测试样本集合对 Elman 网络模型进行测试, 评价其预测的效果, 选择出最优模型.

4 实测数据分析验证

以北京市三环路上 20 个断面的数据作为研究对象, 根据交通流数据预处理技术, 对城市快速路上的交通流数据进行预处理, 整理 20 个断面每隔 2 分钟的断面交通流量数据. 研究路网及断面布置如图 2 所示.

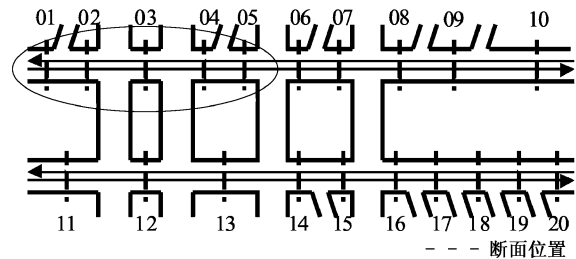
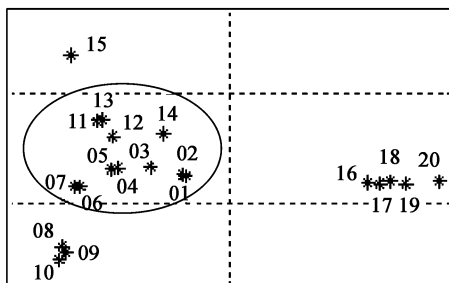


图 2 道路网及道路断面布置图

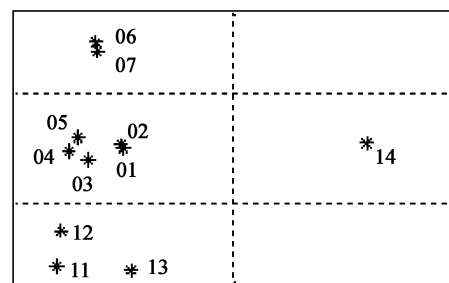
Fig.2 Cross-sections in road network

4.1 基于交通流空间特性的道路网划分

根据计算得到的广义空间距离矩阵, 应用 R-Q 型因子分析法, 得到各断面空间划分示意图. 如图 3(a) 所示, 第一次划分结果中最大子路网包括断面 01、02、03、04、05、06、07、11、12、13、14, 断面数量太多, 不能满足交通流短时预测的需求, 因此以其为研究对象, 进行第二次划分, 划分结果如图 (b) 所示.



(a) 第一划分结果



(b) 第二次划分结果

图 3 基于交通流空间特性的道路网划分结果

Fig.3 Result of road network splitting based on traffic flow space characteristics

经过两次划分,将包括 20 个断面的道路网划分成 7 个子网,最大的子网仅包括 5 个断面,最小的子网包括一个断面,由于最大子网包括的断面数已经满足 Elman 网络的输入需求,因此停止划分,选择最大子网包括的断面 01、02、03、04、05 作为研究对象,进行交通流短时预测研究.

4.2 基于 Elman 神经网络的道路网交通流短时预测

根据道路网划分结果,选取断面 01 - 05 为研究对象,根据道路网划分结果,选取断面 01 - 05 为研究对象,应用相空间重构技术确定交通流量时间序列输入,相空间嵌入维数 $m = 5$,时间延迟 τ 选

取数据上传时间间隔,即 2 min. 为了达到道路网范围内多断面同时预测的目的,构建具有 25 个输入向量,5 个输出向量的 Elman 网络进行预测,训练输入样本是 $25 \times 2\,999$ 的矩阵,输出样本是 $5 \times 2\,999$ 的矩阵;测试输入样本是 25×596 的矩阵,输出样本是 5×596 的矩阵,通过试算法,确定隐层节点数是 21. 选取 5 个断面中具有代表性的 4 个断面的预测结果如图 4 所示. 由图 4 可知,预测结果与实际交通量拟合较好,能够实时地反映、跟踪交通流状态变化.

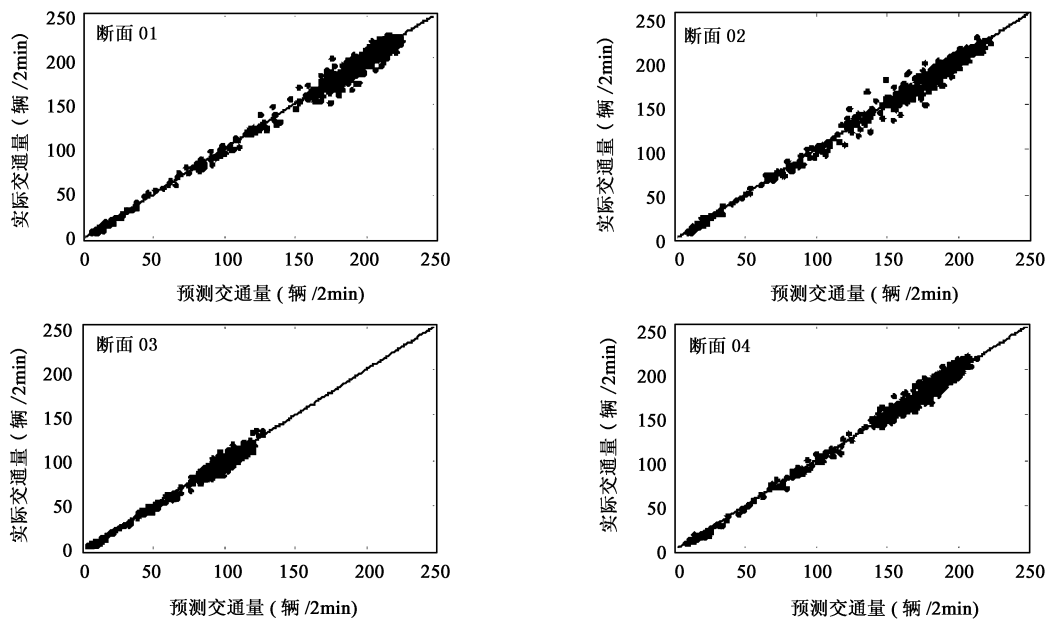


图 4 预测结果与实测值对比

Fig.4 Comparison of forecasting result and real value

4.3 结果分析

交通流短时预测误差结果的对比指标采用平均绝对百分比误差(Mean Absolute Percentage Error, 简称 MAPE,表示预测值与实测值的实际偏差绝对值占实测值百分比的均值)和平均绝对偏差(Mean Absolute Deviation,简称 MAD,表示预测值与实测值的实际偏差绝对值的均值). 前四个道路断面预测百分比误差和绝对偏差如图 5 所示,从误差对比分析可以看出,高峰时刻,预测绝对偏差较大. 平峰时刻,预测百分比误差较大,五个道路断面预测误差对比分析如表 1 所示.

为了证明方法的有效性,应用相同数据,采用 BP 网络进行交通量短时预测,两种方法预测误差

对比如图 6 所示. 由图 6 可以看出,基于 Elman 网络在各断面预测误差波动较小,各断面预测误差平均值小于 BP 网络预测误差.

表 1 道路网多断面交通流短时预测结果评价

Table 1 Results evaluation of short-term traffic volumes forecasting of road network

断面	01	02	03	04	05	平均值
MAPE	3.64%	3.54%	4.70%	3.35%	5.26%	4.10%
MAD	4.58	4.16	3.20	3.94	4.96	4.17

5 研究结论

智能交通控制和诱导系统实现需要的是路网范围内的交通流短时预测,以往的交通流预测仅限

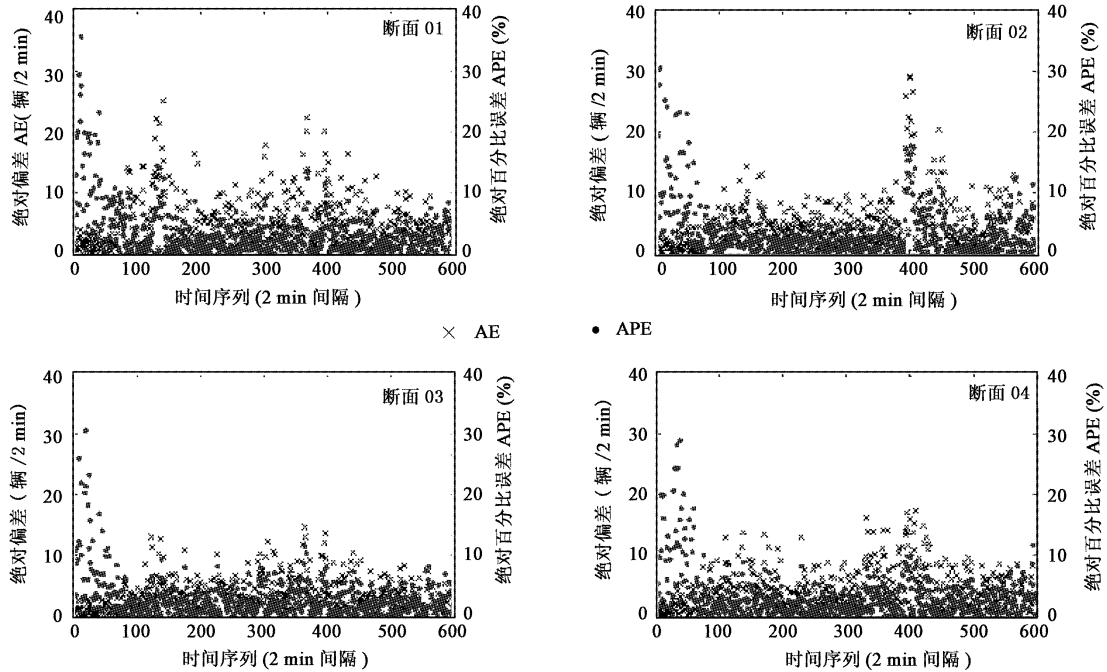


图5 预测误差对比分析

Fig.5 Comparison of forecasting error

于单断面,在实际应用时各断面分别预测增加建模复杂性,计算效率低.本文以路网为研究对象,首先,通过分析交通流空间特性参数,考虑相邻多个断面(包括不同路线)的相互联系,以广义空间距离矩阵为基础,采用R-Q型分析法将研究道路网划分为满足建模需要的子路网;其次,以道路网划分结果为基础,以重构的交通流量时间序列为输入,

采用Elman网络实现道路网多断面交通流量短时预测.应用实例及结果分析表明,在多断面交通流短时预测方面,基于Elman神经网络的交通流短时预测效果优于BP神经网络,预测效果较好,但是交通流时间特性对道路网划分的灵敏度分析及对道路网交通流短时预测影响等方面还需要进一步深入细致的研究.

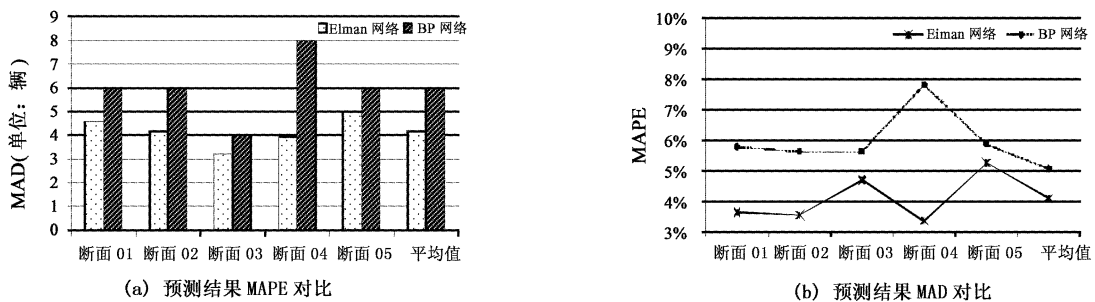


图6 Elman网络与BP网络预测误差对比

Fig.6 Comparison of forecasting error of Elman neural network and BP neural network

参考文献:

- [1] 姚智胜.基于实时数据的道路网短时交通流预测理论与方法研究[D].北京:北京交通大学,2007.[YAO Z S. Research on theories and methods of short-term traffic flow forecasting of road network based on real-time data

[D]. Beijing: Beijing Jiaotong University, 2007.]

- [2] Sherif Ishak, Haitham Al-Deek. Performance evaluation of short-term time-series traffic prediction model[J]. Journal of Transportation Engineering, 2002, 128(6):490-498.
- [3] Chrobok R, Kaumann O, Wahle J, Schreckenberg M. Different methods of traffic forecast based on real date[J].

- European Journal of Operational Research, 2004, 155(3): 558 – 568.
- [4] Sangsoo Lee, Daniel B Fambro. Application of the subset ARIMA model for short-term freeway traffic volume forecasting[J]. Transportation Research Record 1678, 1999: 179 – 188.
- [5] Lingras Pawan, Sharma Satish C, Osborne Phil, et al. Traffic volume time-series analysis according to the type of road use[J]. Computer-Aided Civil and Infrastructure Engineering, 2000, 15(5):365 – 373.
- [6] Anthony Stathopoulos, Matthew G, karlaftis. A multivariate state space approach for urban traffic flow modeling and prediction[J]. Transportation Research PartC, 2003: 121 – 135.
- [7] 姚智胜,邵春福,熊志华,等. 基于主成分分析和支持向量机的道路网短时交通流量预测[J]. 吉林大学学报(工学版), 2008, 38(1):48 – 52. [YAO Z S, SHAO C F, XIONG Z H, et al. Short-term traffic volumes forecasting of road network based on principal component analysis and support vector machine[J]. Journal of Jilin University(Engineering and Technology Edition), 2008, 38(1): 48 – 52(in Chinese).]
- [8] 周小鹏,冯奇,孙立军. 基于最近邻法的短时交通流预测[J]. 同济大学学报(自然科学版), 2006, 34(11): 1494 – 1498. [ZHOU X P, FENG Q, SUN L J. Short-term traffic flow forecasting based on nearest neighbor algorithm[J]. Journal of University(Natural Science), 2006, 34(11): 1494 – 1498(in Chinese).]
- [9] Jeffrey L, Elman. Finding structure in time[J]. Cognitive Science, 1990, 14:179 – 211.
- [10] Chunjiao Dong, Chunfu Shao, Xia Li. Short-term traffic flow forecasting of road network based on spatial-temporal characteristics of traffic flow. WRI World Congress on Computer Science and Information Engineering Proceeding, 2009:645 – 650.