

控制与决策

Control and Decision

基于FWADE-ELM的短时交通流预测方法

陈如清, 李嘉春, 俞金寿

引用本文:

陈如清, 李嘉春, 俞金寿. 基于FWADE-ELM的短时交通流预测方法[J]. *控制与决策*, 2021, 36(4): 925-932.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.1103>

您可能感兴趣的其他文章

Articles you may be interested in

基于局部搜索的反向学习竞争粒子群优化算法

Opposition-based learning competitive particle swarm optimizer with local search

控制与决策. 2021, 36(4): 779-789 <https://doi.org/10.13195/j.kzyjc.2019.1150>

基于多维泰勒网的超前d步预测模型

d-step-ahead predictive model based on multi-dimensional Taylor network

控制与决策. 2021, 36(2): 345-354 <https://doi.org/10.13195/j.kzyjc.2019.0722>

基于改进萤火虫算法的区域交通信号配时优化

Timing optimization of regional traffic signals based on improved firefly algorithm

控制与决策. 2020, 35(12): 2829-2834 <https://doi.org/10.13195/j.kzyjc.2019.1835>

基于改进堆叠自动编码器的循环冷却水系统工艺介质温度预测控制方法

Predictive control method of process medium temperature in circulating cooling water system based on improved stacked auto encoders

控制与决策. 2020, 35(12): 2835-2844 <https://doi.org/10.13195/j.kzyjc.2019.0694>

基于强化学习的小型无人直升机有限时间收敛控制设计

Finite time control based on reinforcement learning for a small-size unmanned helicopter

控制与决策. 2020, 35(11): 2646-2652 <https://doi.org/10.13195/j.kzyjc.2019.0328>

基于FWADE-ELM的短时交通流预测方法

陈如清^{1†}, 李嘉春², 俞金寿³

(1. 嘉兴学院 机电工程学院, 浙江 嘉兴 314001; 2. 嘉兴学院 数理与信息工程学院, 浙江 嘉兴 314001;
3. 华东理工大学 自动化研究所, 上海 200237)

摘要: 受道路环境和人为因素影响, 实际交通系统可视为一个复杂的非线性动力系统, 交通流数据具有较强的非线性、时变性和易受随机噪声影响等特征. 针对复杂环境下的短时交通流预测问题, 提出一种基于烟花差分进化混合算法-极限学习机的短时交通流预测方法. 采用奇异谱分析方法滤除原始交通流数据中包含的噪声成分, 降噪后的交通流数据用于训练极限学习机(ELM)网络预测模型. 进行相空间重构, 利用C-C算法确定ELM网络的结构和关键参数. 通过融合烟花算法和差分进化算法提出一种烟花差分进化混合算法, 可有效提高基本算法的整体优化性能. 将改进的混合优化算法用于优化ELM网络的权阈值(结构为9-11-1, 维数为110), 建立短时交通流预测模型. 测试与应用结果表明, 所构建的短时交通流预测模型具有较高的预测精度和较强的泛化能力(均方误差为7.75, 平均绝对百分比误差为0.0867), 预测值与实际值的拟合程度较好.

关键词: 智能交通系统; 短时交通流预测; 极限学习机; 奇异谱分析; 混合优化算法

中图分类号: TP391; U491

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.1103

开放科学(资源服务)标识码(OSID):



引用格式: 陈如清, 李嘉春, 俞金寿. 基于FWADE-ELM的短时交通流预测方法[J]. 控制与决策, 2021, 36(4): 925-932.

Short-term traffic flow forecasting based on hybrid FWADE-ELM

CHEN Ru-qing^{1†}, LI Jia-chun², YU Jin-shou³

(1. College of Mechanical and Electrical Engineering, Jiaying University, Jiaying 314001, China; 2. College of Mathematics, Physics and Information Engineering, Jiaying University, Jiaying 314001, China; 3. Research Institute of Automation, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Affected by road condition and human factors, the practical traffic system can be considered as a complex nonlinear dynamical system and the traffic flow data have the properties of strong non-linearity, time-variety and susceptibility to random noises. To solve the problems of short-term traffic flow forecasting in complex environment, a prediction method based on the hybrid FWADE-ELM is proposed. The singular spectrum analysis (SSA) technique is used to filter the noise existing in the original traffic flow data and then the ELM neural network prediction model is trained with the preprocessed data. The structure and key parameters of ELM network are determined using the C-C algorithm after reconstructing phase space. A hybrid optimization method integrating the fireworks algorithm with the differential evolution algorithm is developed to improve the optimal performance of the basic algorithms. The ELM network forecasting model is built and its weights and biases (9-11-1 structure and 110 dimensions) are optimized using the proposed FWADE hybrid algorithm. The results of property testing and practical application show that this short-term traffic flow forecasting model has higher forecasting accuracy and better generalization ability, and the predicted values comply well with the actual values.

Keywords: intelligent traffic system; short-term traffic flow forecasting; extreme learning machines; singular spectrum analysis; hybrid optimization algorithm

0 引言

城市道路交通流预测是智能交通系统关键技术之一, 是实现交通实时控制与诱导的基础和前提. 短

时交通流预测是利用动态获取的若干道路交通流状态数据的时间序列, 实现对未来时段交通流状态数据的预测. 短时交通流的实时准确预测对路网交通状

收稿日期: 2019-07-31; 修回日期: 2019-10-21.

基金项目: 浙江省基础公益研究计划项目(LGG18F030011); 国家自然科学基金项目(61603154).

责任编辑: 魏秀琨.

[†]通讯作者. E-mail: ruqing2002@163.com.

态分析、交通网络规划和交通优化控制都具有十分重要的作用^[1]。交通流预测建模方法研究具有重要的理论价值和实际意义^[2]。

根据原理不同,现有交通流预测建模方法可分为参数方法和非参数方法两类^[3-4]。参数方法主要包括历史平均、线性和非线性参数回归、卡尔曼滤波、自动回归和指数平滑等预测方法。基于经典参数预测方法,建立了多种改进的交通流预测模型。Xie等^[5]融合小波变换方法与卡尔曼滤波方法,提出了一种短时交通流预测模型,在交通流变化剧烈的情况下具有良好的预测效果;Pan等^[6]对SCTM框架进行扩展,通过分析交通流时间与空间的相关性实现对交通状态的实时预测;Tchrakian等^[7]基于谱分析提出了一种短时交通流预测方法并构建相应的实时预测模型;基于现有自适应预测算法,Wang等^[8]提出了一种改进的自适应车流量预测算法预测常规交通状况下进入路网的车流量;Li等^[9]采用Granger因果检验方法发现交通流数据中存在的依赖关系,提出了基于分层算法的交通流预测方法。非参数方法主要包括非参数回归、人工神经网络和支持向量机等预测方法或多模型组合预测方法^[1,10-12]。Boto-Giralda等^[13]通过离散小波变换去除交通流数据的噪声,构建了自组织神经网络交通流预测模型;Chan等^[14]融合指数平滑与LM(levenberg-marguardt)算法提出了一种改进的人工神经网络交通流预测方法;Han等^[15]通过分析交通流数据的时空特性对交通流预测的影响,提出了基于深度学习的短时交通流预测模型;Jeong等^[16]基于加权支持向量回归在线学习方法,建立了短时交通流预测模型;Lippi等^[17]基于时间序列分析和监督学习方法,建立了两种支持向量机回归预测模型,结合交通流量的周期性进行了短期交通流预测;Hu等^[18]提出了一种基于SVR(support vector regression)的交通流预测模型,并用PSO算法优化SVR模型的关键参数,具有较高预测精度;Bing等^[19]针对实际交通系统具有非线性动态特性问题,基于混沌理论和局域预测方法,构建了基于PSO优化的组合核相关向量机预测模型。参数和非参数交通流预测模型在实时性、精确性和复杂性等方面各有优缺点。总体而言,参数方法的特点是模型简单且计算量较小,但存在模型学习能力不足、难以处理模型内在不确定性及交通数据突变影响预测精度等不足。非参数方法则采用数据挖掘和人工智能等方法根据历史数据揭示数据变化规律,建立输入输出映射关系。其缺点是预测精度严重依赖历史数据的数量和质量,且模型参数选取较为困难。

受城市道路交通环境、天气状况和道路行人等

不确定因素影响,实际交通流数据具有较强的非线性、时变性和易受随机噪声影响等特征。现有交通流预测建模方法存在模型参数难以确定、预测精度较低和泛化能力较差等不足,难以满足复杂情况下的短时交通流预测要求。针对以上问题,本文提出一种基于烟花差分进化混合算法-极限学习机的短时交通流预测建模方法。采用奇异谱分析方法滤除原始交通流数据中的噪声成分,用降噪及相空间重构后的交通流数据训练ELM模型,将提出的烟花差分进化混合优化算法用于优化ELM权阈值并建立短时交通流预测模型。仿真和实验结果表明,该预测模型是有效可行的。

1 ELM神经网络

极限学习机(extreme learning machines, ELM)是一种用于求解单隐层前馈神经网络的算法^[20]。ELM神经网络训练过程随机设定输入层和隐含层的连接权值及阈值,通过对隐含层的求逆及与训练输出的相乘计算隐含层与输出层间的权值,无需调整网络参数即可完成网络训练。与传统神经网络学习算法相比,在保证学习精度前提下,ELM具有训练参数更少、学习速度更快和泛化性能更好等优点^[21]。

单隐层前馈神经网络如图1所示。

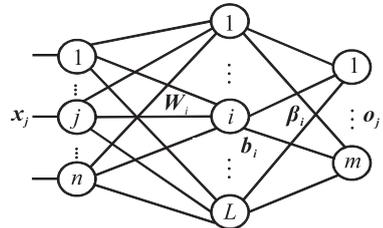


图1 单隐层前馈神经网络

在图1中:输入层、隐含层及输出层的点数分别为 n 、 L 和 m 。给定 N 组训练样本数据集 $\{x_j, t_j\}$, $x_j = [x_{j1}, x_{j2}, \dots, x_{jn}]^T \in R^n$, $t_j = [t_{j1}, t_{j2}, \dots, t_{jm}]^T \in R^m$,网络的输入输出关系可表示为

$$\sum_{i=1}^L \beta_i g(W_i \cdot x_j + b_i) = o_j, \quad j = 1, 2, \dots, N. \quad (1)$$

其中: $g(\cdot)$ 为激活函数, $W_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ 和 β_i 分别为输入层与隐含层和隐含层与输出层的权值, b_i 为隐含层的阈值。学习过程中需调整的权阈值总数为 $L(n+1)$ 。

网络学习目标为最小化神经网络的输出误差,即

$$\sum_{j=1}^N \|o_j - t_j\| \rightarrow 0. \quad (2)$$

即存在 β_i , W_i 和 b_i ,使

$$\sum_{i=1}^L \beta_i g(W_i \cdot x_j + b_i) = t_j, \quad j = 1, 2, \dots, N. \quad (3)$$

用矩阵表示为

$$H\beta = T. \tag{4}$$

其中: H 为隐层输出, β 为输出权重, T 为期望输出, 即

$$H = \begin{bmatrix} g(W_1 \cdot x_1 + b_1) & g(W_2 \cdot x_1 + b_2) & \dots & g(W_L \cdot x_1 + b_L) \\ g(W_1 \cdot x_2 + b_1) & g(W_2 \cdot x_2 + b_2) & \dots & g(W_L \cdot x_2 + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g(W_1 \cdot x_N + b_1) & g(W_2 \cdot x_N + b_2) & \dots & g(W_L \cdot x_N + b_L) \end{bmatrix}_{N \times L}, \tag{5}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \beta_2^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m}, \quad T = \begin{bmatrix} T_1^T \\ T_2^T \\ \vdots \\ T_N^T \end{bmatrix}_{N \times m}$$

通过单隐层神经网络的训练获得一组最优 W_i^* , b_i^* 和 β_i^* , 使

$$\|H(W_i^*, b_i^*)\beta_i^* - T\| = \min_{W, b, \beta} \|H(W_i, b_i)\beta_i - T\|, \tag{6}$$

其中 $i = 1, 2, \dots, L$. 式(6)相当于将损失函数 E 最小化.

$$E = \sum_{j=1}^N \left(\sum_{i=1}^L \beta_i g(W_i \cdot x_j + b_i) - t_j \right)^2. \tag{7}$$

基于梯度下降法的传统算法在求解此类问题的迭代过程中需要调节所有参数, 而ELM则只需确定权值 W_i 和阈值 b_i , 便可唯一确定隐层输出 H . 于是, 将单隐层神经网络的训练转化为线性系统 $H\beta = T$ 的求解, 输出权重 β 可由下式确定.

$$\beta^* = H^\dagger T. \tag{8}$$

其中: H^\dagger 为矩阵的 Moore-Penrose 广义逆, 可采用奇异值分解法、正交分解法或正交投影法等方法求解.

与传统神经网络算法相比, ELM 具有较明显的优势, 但也存在固有缺陷: 1) 输入层与隐含层的连接权值及隐含层的阈值均采用随机方式设定, 具有一定的盲目性, 易导致 ELM 网络的稳定性变差; 2) 训练数据中若存在噪声或离群点, 易导致隐含层的输出矩阵出现病态问题, 使 ELM 网络的鲁棒性变差、泛化性能降低.

2 基于烟花差分进化混合算法的 ELM 网络

2.1 基本烟花算法

烟花算法 (fireworks algorithm, FWA) 是受烟花燃放时爆炸产生火花这一现象的启发而提出的, 其凭借

着较强的鲁棒性能和寻优能力, 受到业界和广大学者的密切关注. 通过对现有烟花算法的深入研究分析, 提出了多种改进或混合算法, 也给出了一些成功的应用案例^[22-23].

FWA 的寻优过程, 依次采用了爆炸算子、变异算子、映射规则 and 选择策略. 基于下式, 计算烟花爆炸半径及爆炸火花数量:

$$R_i = R \times \frac{f(x_i) - y_{\min} + \varepsilon}{\sum_{i=1}^N (f(x_i) - y_{\min}) + \varepsilon}, \tag{9}$$

$$S_i = M \times \frac{y_{\max} - f(x_i) + \varepsilon}{\sum_{i=1}^N (y_{\max} - f(x_i)) + \varepsilon}. \tag{10}$$

其中: R_i 为烟花 x_i 的爆炸半径, R 为最大爆炸幅度的常数, y_{\min} 和 y_{\max} 分别为目前烟花个体中最好和最差的个体适应度, S_i 为烟花 x_i 的爆炸火花数目, M 用于限制火花总数的常数, ε 为一个防止除零的微小量. 为控制烟花爆炸过程产生火花的数量, 对各烟花的爆炸火花数目作相关限制, 即

$$S_i = \begin{cases} \text{round}(a \times M), & S_i < aM; \\ \text{round}(b \times M), & S_i > bM, a < b < 1; \\ \text{round}(S_i), & \text{otherwise.} \end{cases} \tag{11}$$

其中: a, b 均为常数, round 函数用于四舍五入取整.

位置迁移是 FWA 中烟花成员在解空间位置的改变, 即

$$ex_{ik} = x_{ik} + R_i \times \text{rand}(-1, 1). \tag{12}$$

其中: $R_i \times \text{rand}(-1, 1)$ 为位置偏移量; x_{ik} 为第 i 个烟花成员的 k 维分量, ex_{ik} 为 x_{ik} 爆炸后产生的爆炸火花.

高斯变异操作的目的是增加 FWA 的突变性和多样性, 即

$$mx_{ik} = x_{ik} \times e. \tag{13}$$

其中: $e \sim N(1, 1)$, mx_{ik} 为生成的变异火花.

选择策略是烟花算法对爆炸后烟花种群成员的选择方法. 候选成员集合由烟花成员、爆炸火花和高斯变异火花组成. 集合中的最优个体最先被选定为下一代烟花成员, 再采取“轮盘赌”方式依次从候选成员中挑选其余的烟花成员, 集合成员 x_i 被选中概率为

$$p(x_i) = \frac{R(x_i)}{\sum_{x_j \in K} R(x_j)},$$

$$R(x_i) = \sum_{x_j \in K} d(x_i - x_j) = \sum_{x_j \in K} \|x_i - x_j\|, \tag{14}$$

其中 $R(x_i)$ 为 x_i 与 S 中各个体之间的距离和。

2.2 FWADE-ELM算法

差分进化算法(differential evolution algorithm, DE)是一种基于群体的启发式搜索算法,具有原理简单、全局搜索能力强及鲁棒性好等特点。DE算法的进化流程主要包括变异、交叉和选择等3个操作。基于DE搜索机制进行迭代计算,保留优良个体并淘汰劣质个体,引导搜索方向朝最优解逼近。FWA算法具有较强的通用性且易与其他优化算法融合,但在处理复杂优化问题时具有算法复杂度较高、搜索时间较长和易陷入局部最优等不足。DE算法采用种群全局搜索策略,可有效减小遗传操作的计算复杂性。此外,DE算法特有的记忆能力可使其动态跟踪当前的搜索状况并及时调整其搜索策略,具有较好的全局优化性能。

为充分发挥两种优化算法的优势并提高算法的整体优化性能,本文将烟花算法与差分进化算法有机结合,提出一种烟花差分进化混合优化算法(fireworks algorithm-differential evolution algorithm, FWADE)。FWADE算法优化过程的初始阶段采用标准DE算法搜索,对于每个目标向量,通过变异和交叉运算产生试验向量。若试验向量的适应度优于目标向量,则将其作为下一代候选解;否则,算法进入FWA阶段,利用烟花算法搜索机制产生新的候选解。FWADE算法优化过程中融入了FWA搜索阶段,通过在种群中形成扰动以维持种群的多样性并产生最优解。整个迭代过程中,不断淘汰劣质个体并选择优良个体,引导种群朝全局最优解的方向逼近。

FWADE算法的具体实现步骤描述如下:

step 1: 算法的初始化。根据待优化问题确定种群规模 N_1 , 最大迭代次数 N_2 , 个体成员的初始值 $X_i(k)$; 设置FWA算法爆炸半径调节常数 R , 爆炸火花数调节常数 M , 爆炸火花数下限系数 a 和上限系数 b , 高斯变异火花数 G 等参数初始值; 确定DE算法的比例因子 F 和交叉率 CR 等参数初始值; 令迭代次数 $k = 1$ 。

step 2: 根据待优化问题的数学模型, 计算各个体成员的适应度值 $f(X_i(k))$ 。

step 3: DE算法搜索阶段: 1) 变异操作。随机选择种群中两个不同的个体, 参照DE算法对种群中每个目标向量 $X_i(k)$ 进行变异操作产生变异向量 $V_i(k)$ 。2) 交叉操作。对每对目标向量 $X_i(k)$ 和对应的变异向量 $V_i(k)$ 进行交叉操作, 形成新的试验向量 $U_i(k)$ 。3) 选择操作。根据待优化问题的数学模型, 评价每个试验向量与对应目标向量的适应度值 f 。若 $U_i(k)$ 的适应度值更优, 则作为下一次迭代的候选解并转至 step 5; 否则进入 step 4, 即FWA算法搜索阶段。

step 4: FWA算法搜索阶段。根据FWA算法优化机制, 参照FWA算法的式(9)~(14)对种群执行搜索, 计算种群的适应度值, 择优产生新的候选解。

step 5: 对种群中每个个体进行相同操作。个体序号 $i = i + 1$, 若 $i \leq N_1$, 则转至 step 3, 否则转至 step 6。

step 6: 更新全局最优解和迭代次数。根据新产生的候选解, 评估每个个体的适应度并更新全局最优解, 更新迭代次数 $k = k + 1$ 。

step 7: 算法终止条件(即群体的最优适应度值满足要求或迭代次数达到最大迭代次数)判断。若满足, 则终止条件转 step 8, 否则重复 step 3 ~ step 7。

step 8: 输出FWADE算法的全局最优解。

将本文提出的FWADE算法用于优化ELM网络的连接权值和阈值, 构建FWADE-ELM拟合模型。模型参数优化过程中充分发挥了FWA算法和DE算法的各自优势, 多次迭代不断选取更优的输入权重 W_i 和隐层阈值 b_i 矩阵。经过上述优化过程最终获得最优群成员位置(即最优输入权重 W_i^* 和隐层阈值 b_i^* 矩阵), 再根据ELM的广义逆计算输出权重 β 。

2.3 算法测试

为验证本文FWADE算法优化ELM神经网络的效果, 进行仿真测试分析。测试对象为 $\sin C$ 函数和高斯正态函数 ($-10 \leq x \leq 10$), 函数表达式如下:

$$f_1(x) = \begin{cases} \frac{\sin x}{x}, & x \neq 0; \\ 0, & x = 0. \end{cases} \quad (15)$$

$$f_2(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}(x-\mu)^2}, \quad \sigma = 1, \mu = 1. \quad (16)$$

训练样本和测试样本数量分别为4000组和1000组, 数据服从区间 $[-10, 10]$ 上的均匀分布。此外, 为验证拟合模型的鲁棒性能, 在训练样本中添加了 $[-0.2, 0.2]$ 的白噪声。仿真过程引入DE算法, FWA算法和FWADE算法, 分别建立DE-ELM, FWA-ELM和FWADE-ELM神经网络拟合模型, 与基本ELM拟合模型形成对比。

在仿真过程中, 4种方法均采用结构为1-8-1的ELM神经网络和Sigmoid激励函数; ELM权阈值优化算法的适应度函数为由训练样本计算的均方根误差(RMSE); DE, FWA和FWADE算法的群成员规模均为25, 群成员维数(即待优化的ELM网络权阈值个数)为 $8 \times (1 + 1) = 16$, 最大迭代次数设为500次。其余参数设置情况为: FWA和FWADE算法中爆炸半径调节常数 $R = 40$, 爆炸火花数调节常数 $M = 64$, 爆炸火花数下限系数 $a = 0.04$, 爆炸火花数上限系数 $b = 0.8$, 高斯变异火花数 $G = 40$; DE算法中杂交参数 $CR = 0.8$, 缩放比例因子 $F = 0.6$ 。采用ELM、DE-ELM、FWA-ELM和FWADE-ELM四种拟合模型分

表1 4种拟合模型泛化性能对比(100次实验)

测试对象	RMSE(训练集)				RMSE(测试集)			
	ELM	DE-ELM	FWA-ELM	WADE-ELM	ELM	DE-ELM	FWA-ELM	WADE-ELM
$f_1(x)$	0.2785	0.2116	0.1957	0.1385	0.0586	0.0391	0.0284	0.0071
$f_2(x)$	0.2316	0.1892	0.1765	0.1219	0.0522	0.0291	0.0206	0.0057

别对两个测试函数随机进行100次独立实验,测试结果(训练样本和测试样本的RMSE)如表1所示。

在训练集中人为添加一定程度的白噪音,测试结果中训练集的RMSE整体要大于测试集。分析表明:由于基本ELM模型采用随机方式选取网络的权阈值且训练样本中存在一定程度的白噪声,导致模型的泛化性能比较差;DE-ELM模型采用DE算法优化ELM的权阈值,模型泛化性能得到一定程度提高,但处理噪声环境下的函数优化问题时整体性能不佳,较大程度依赖于控制参数(如种群规模、交叉概率和缩放因子等)的选取;FWA-ELM模型采用FWA算法优化ELM网络的权阈值,通过烟花爆炸方式能以较快速度选取最优权阈值矩阵,具有较小的训练误差和测试误差;FWADE-ELM模型使用FWADE算法确定ELM网络的权阈值,进一步提高了ELM拟合模型的泛化性能,具有最优的拟合精度和鲁棒性能。

3 FWADE-ELM短时交通流预测模型

3.1 短时交通流预测建模

实际交通系统是一个复杂的非线性动力系统,采集到的原始交通流数据具有较强的非线性、时变性和易受随机噪声影响等特征。针对复杂情况下的短时交通流预测问题,本文提出一种基于烟花差分进化混合算法-极限学习机的短时交通流预测建模方法:1)考虑到采集到的原始短时交通流数据易受随机噪声的影响,直接用于建模将降低模型的预测精度。原始数据预处理时,采用奇异谱分析(singular

spectrum analysis, SSA)方法滤除交通流时间序列的噪声成分,得到降噪后的重构时间序列,用于短时交通流预测建模。2)根据ELM的基本原理,ELM能有效改进传统BP(back propagate)算法及SVM(support vector machine)方法,提升学习效率并简化学习参数的设定,具有较快的学习速度和良好的泛化性能,有利于处理短时交通流非线性非平稳的时间序列数据,构建基于时间序列的短时交通流预测模型。3)ELM随机产生的输入权值和隐层阈值矩阵导致其输出权值矩阵的非唯一性,如何获取最优的输出权值矩阵是得到ELM最优输出值的关键。作为一种新型进化算法,FWA具有良好的收敛精度和寻优能力。融合经典DE算法提出一种烟花差分进化混合优化算法,可充分发挥两种优化算法的优势并提高单一算法的整体优化性能,用于优化ELM网络的权阈值。

综上所述,FWADE-ELM交通流预测模型的建立过程为:1)利用SSA方法对现场采集的原始交通流时间序列数据进行降噪处理,作为交通流预测模型的建模数据;2)进行相空间重构,基于C-C算法确定ELM网络的结构、时间延迟和嵌入维数等关键参数;3)以降噪及相空间重构后的交通流时间序列作为ELM网络的训练样本,采用FWADE混合优化算法确定ELM网络输入层与隐含层之间的连接权值和隐层神经元的阈值;4)以未经训练的交通流时间序列数据作为ELM网络模型的测试样本,评估优化后ELM网络模型的预测性能,具体建模流程如图2所示。

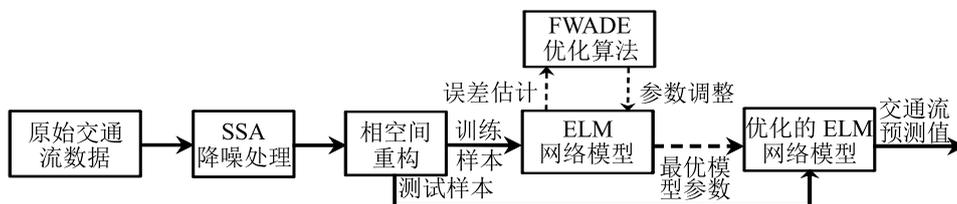


图2 FWADE-ELM交通流预测建模流程

3.2 基于SSA的噪声数据处理

SSA是近年来兴起的一种适用于复杂非线性时间序列的数据处理方法^[24]。应用SSA方法可从包含噪声的原始信号中提取有用信息、滤除尖锐震荡的趋势成分,将包含预报信息的成分重构时间序列,实现减少或消除噪声的目的^[25]。本文利用SSA方法滤

除交通流时间序列的噪声成分,将SSA降噪交通流数据用于训练ELM神经网络,并采用FWADE算法优化ELM模型的关键参数。

噪声数据的处理原理为,将从现场采集到的含有噪声成分交通流参数时间序列 $Y_N = [y_1, y_2, \dots, y_N]$ 转化为轨迹矩阵 X ;对矩阵 XX^T 进行奇异值分解,得

到 L 个特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$ 及对应的特征向量;将每个特征值代表的信号进行分析组合,重构出新的时间序列 $G = [g_0, g_1, \dots, g_{N-1}]$. 处理过程包括嵌入、奇异值分解、分组与重构4个步骤.

重构过程保留了原始交通流时间序列中前 m 个较大奇异值的成分,舍弃了那些由噪声引起的较小奇异值成分. 原始交通流时间序列经过 SSA 的过滤处理,得到降噪后的重构时间序列 G ,用于建立短时交通流预测模型.

本文的实验数据来自某市交通管理局 UTC/SCOOT 系统采集的市内一段长度约 12 km 快速路的短时交通流数据,通过布设在道路上的 95 个主线和 72 个匝道地感线圈检测器获取交通流量等 3 项交通参数. 采集时间为 2013 年 7 月份的连续 5 个星期一,即 7 月 1 日、7 月 8 日、7 月 15 日、7 月 22 日和 7 月 29 日,数据采集时间间隔为 5 min,每天可得到 288 组短时交通流数据.

SSA 降噪处理时,考虑交通流时间序列具有周期分量,在窗口长度 L 与周期成正比的情况下可以获得较好的可分离性. 处理过程中设 L 为 288,与交通流时间序列的日变化对应. 图 3 为交通流时间序列的奇异谱.

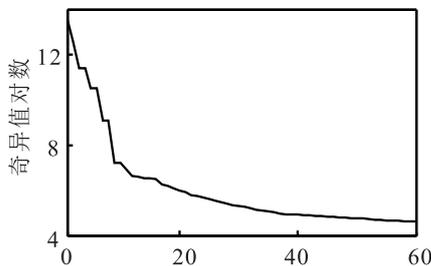


图 3 交通流时间序列的奇异谱

分析表明:前 37 个奇异值对应重构序列的累积贡献率达 98.78%,后 251 个仅为 1.22%,可视为噪声成分. 进行相空间重构,采用 C-C 算法计算时间延迟 $\tau = 15$ 及嵌入窗宽 $\tau_w = 123$. τ_w 计算公式为

$$\tau_w = (m - 1)\tau. \tag{17}$$

由此得出嵌入维数 $m = 9$. 将降噪及相空间重构后的交通流时间序列(即 7 月 1 日、7 月 8 日、7 月 15 日和 7 月 22 日共 4 天的交通流数据)作为 ELM 网络的训练样本,7 月 29 日的原始交通流数据作为 ELM 网络的测试样本. 确定交通流预测建模时 ELM 网络的拓扑结构为 9-11-1,即输入层节点数为 9,对应相空间重构后的交通流时间序列的嵌入维数;模型隐含层节点数的确定目前尚无统一理论依据,仿真过程中经多次试验权衡后最终设置为 11;输出层节点数为 1,对应 ELM 网络的输出值,即交通流量的预测值,学习过程中需优化调整的权阈值总数为 $11 \times (9 + 1) = 110$.

分别采用基本 DE 算法、基本 FWA 算法和本文提出的 FWADE 算法优化 ELM 网络的权阈值,构建 DE-ELM、FWA-ELM 和 FWADE-ELM 三种交通流预测模型,与基本 ELM 交通流预测模型形成对比. 建模过程中,4 种方法均采用结构为 9-11-1 的 ELM 神经网络和 Sigmoid 激励函数;ELM 网络权阈值优化算法的适应度函数为由训练样本计算的均方根误差 (RMSE);DE、FWA 和 FWADE 算法的群成员规模均为 40,群成员维数(即待优化 ELM 网络权阈值的个数)为 110,最大迭代次数设为 5 000 次,其余参数设置情况为:FWA 和 FWADE 算法中爆炸半径调节常数 $R = 210$,爆炸火花数调节常数 $M = 220$,爆炸火花数下限系数 $a = 0.04$,爆炸火花数上限系数 $b = 0.8$,高斯变异火花数 $G = 60$;DE 算法中杂交参数 $CR = 0.6$,缩放比例因子 $F = 0.5$. 训练结束保存最佳权阈值用于 ELM 网络模型预测交通流. 表 2 为 4 种预测模型的训练和测试结果的分析对比, RMSE 代表均方根误差, MAPE 代表平均绝对百分比误差. 实验过程中也分析了隐含层节点数对 ELM 网络模型预测结果的影响,表 3 和表 4 为隐含层节点数分别选取 9 和 13 而其余参数相同的情况下,4 种模型交通流预测结果对比.

表 2 4 种模型仿真结果对比(隐含层节点数 11)

预测模型	RMSE(训练)	MAPE(训练)	RMSE(泛化)	MAPE(泛化)
ELM	15.85	0.169 2	17.66	0.209 3
DE-ELM	12.42	0.147 5	15.34	0.161 4
FWA-ELM	9.29	0.101 2	11.21	0.124 9
FWADE-ELM	6.60	0.062 5	7.75	0.086 7

表 3 4 种模型仿真结果对比(隐含层节点数 9)

预测模型	RMSE(训练)	MAPE(训练)	RMSE(泛化)	MAPE(泛化)
ELM	17.92	0.191 4	19.95	0.236 4
DE-ELM	13.83	0.164 7	17.16	0.180 6
FWA-ELM	10.34	0.113 2	12.56	0.139 7
FWADE-ELM	7.26	0.068 7	8.62	0.096 2

表4 4种模型仿真结果对比(隐含层节点数13)

预测模型	RMSE(训练)	MAPE(训练)	RMSE(泛化)	MAPE(泛化)
ELM	15.44	0.165 2	17.27	0.204 7
DE-ELM	12.16	0.144 5	15.08	0.158 7
FWA-ELM	9.10	0.099 1	11.01	0.122 4
FWADE-ELM	6.51	0.061 7	7.66	0.085 7

表2~表4的测试结果表明:1)隐含层节点数越多,预测模型的结构越复杂而误差越小。2)当隐含层节点数小于11时,预测模型的训练误差和泛化误差存在较大变化;当隐含层节点数大于11时,预测模型的训练误差和泛化误差改善不明显。3)相比之下,隐含层节点数对基本ELM模型的影响更大,而对FWADE-ELM模型的影响较小。

综合考虑到ELM网络模型的复杂程度、训练误差和泛化误差等因素,经多次试验权衡后隐含层节点数最终设置为11。4种交通流预测模型(隐含层节点数11)的训练和泛化结果如图4和图5所示。

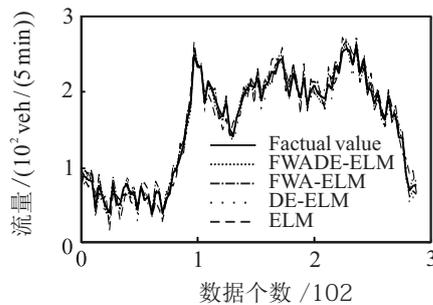


图4 4种交通流预测模型的训练结果(7月22日)

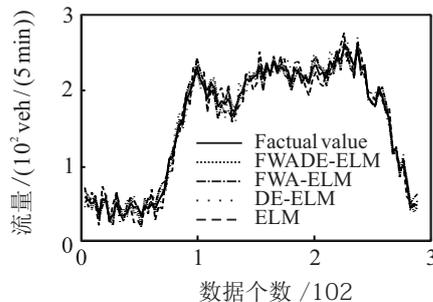


图5 4种交通流预测模型的泛化结果(7月29日)

表2、图4和图5所示的交通流预测结果对比表明:1)DE-ELM模型、FWA-ELM模型和FWADE-ELM模型的短时交通流预测效果均优于基本ELM模型,说明利用优化算法优化ELM的权阈值可提高基本ELM的非线性拟合性能;2)与基本ELM、DE-ELM和FWA-ELM三种交通流预测模型相比,FWADE-ELM短时交通流预测模型的训练误差和泛化误差更低,泛化能力更强,进一步说明采用FWADE算法优化ELM模型的有效性,与2.3节算法测试的结果相似;3)从图4和图5可以看出,FWADE-ELM模型的训练过程符合要求(训练样本的RMSE和MAPE值分别为6.60和0.0625),图5中FWADE-ELM所示模型预测值与实

线所示短时交通流实际值的拟合程度较好(测试样本的RMSE和MAPE值分别为7.75和0.0867),说明FWADE-ELM短时交通流预测模型具有较高的预测精度和较强的泛化能力。

4 结论

1) 本文将时间序列降噪理论与机器学习方法相结合,提出了一种基于FWADE-ELM混合模型的短时交通流预测方法。采用SSA方法对原始交通流数据进行降噪处理,作为短时交通流预测建模的数据;利用C-C算法进行相空间重构,确定ELM网络模型的结构和关键参数;将烟花算法与差分进化算法有机结合,提出一种改进的FWADE混合优化算法,用于优化ELM网络的权阈值。

2) 为验证提出的FWADE-ELM混合模型的有效性,以Sin C函数和高斯正态函数为对象进行了仿真测试,并以国内某城市交通管理部门的实测交通量数据为例进行了实验验证和对比分析。仿真测试和实验结果表明,所提出的FWADE-ELM网络模型的预测值与实际值拟合程度较好,泛化能力较强,整体性能优于基本ELM、DE-ELM和FWA-ELM三种预测模型。

3) 基于改进ELM网络的短时交通流预测建模方法,参数设置较简单,运算效率较高。与神经网络等预测建模方法相比,ELM网络无需在迭代过程中调整网络参数即可完成网络训练,在保证学习精度的前提下能显著提高算法的学习效率并简化学习参数的设定^[20]。采用FWADE算法确定ELM网络的最优结构参数可避免因ELM网络基本参数的盲目选取而导致稳定性变差问题。与SVM方法相比,ELM包含直接特征映射,且神经网络的输入层、隐含层和输出层是连接的,使用岭回归求解输出权重。SVM采用基于核方法的间接特征映射,未考虑特征在神经网络各层的连接,通过求解约束最优化问题得到结果。学习误差相当情况下,ELM的参数设置较少,计算复杂度更低。

4) 受多种内在因素及外在因素的影响制约,实际交通系统是一类复杂的非线性动力系统,表现出较强的非线性和不确定性。交通系统的内在演化规律也具有随机性、动态性、非线性和时空相关性等特征。下一步应结合不同交通参数间存在的时空关联性,研究复杂条件下多变量时间序列的相空间重构及

短时交通流局域预测建模问题.

参考文献(References)

- [1] Fangce G, John W P, Rajesh K. Predictor fusion for short-term traffic forecasting[J]. *Transportation Research Part C Emerging Technologies*, 2018, 92(7): 90-100.
- [2] Liu L S, Jia N, Lin L, et al. A cohesion-based heuristic feature selection for short-term traffic forecasting[J]. *IEEE Access*, 2019, 7(2): 3383-3389.
- [3] Lin Y L, Dai X Y, Li L, et al. Pattern sensitive prediction of traffic flow based on generative adversarial framework[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 20(6): 2395-2400.
- [4] Ma D F, Sheng B W, Jin S, et al. Short-term traffic flow forecasting by selecting appropriate predictions based on pattern matching[J]. *IEEE Access*, 2018, 6: 75629-75638.
- [5] Xie Y C, Zhang Y L, Ye Z R. Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2007, 22(5): 326-334.
- [6] Pan T, Sumalee A, Zhong R X, et al. Short-term traffic state prediction based on temporal-spatial correlation[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(3): 1242-1254.
- [7] Tchrakian T T, Basu B, O'Mahony M. Real-time traffic flow forecasting using spectral analysis[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2012, 13(2): 519-526.
- [8] Wang Y B, Van Schuppen J H, Vrancken J. Prediction of traffic flow at the boundary of a motorway network[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(1): 214-227.
- [9] Li L, Su X N, Wang Y W, et al. Robust causal dependence mining in big data network and its application to traffic flow predictions[J]. *Transportation Research Part C: Emerging Technologies*, 2015, 58(2): 292-307.
- [10] Liu Z, Guo J H, Cao J D. A hybrid short-term traffic flow forecasting method based on neural networks combined with K-nearest neighbor[J]. *Promet-traffic Transportation*, 2018, 30(4): 445-456.
- [11] Guo J H, Liu Z, Huang W, et al. Short-term traffic flow prediction using fuzzy information granulation approach under different time intervals[J]. *IET Intelligent Transport Systems*, 2018, 12(2): 143-150.
- [12] Bing Q C, Qu D Y, Chen X F, et al. Short-term traffic flow forecasting method based on LSSVM model optimized by GA-PSO hybrid algorithm[J]. *Discrete Dynamics in Nature and Society*, 2018, 2018: 1-10.
- [13] Boto-Giralda D, Diaz-Pernas F J, Gonzalez-Ortega D, et al. Wavelet-based denoising for traffic volume time series forecasting with self-organizing neural networks[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2010, 25(7): 530-545.
- [14] 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.
- [15] Han D X, Chen J, Sun J. A parallel spatiotemporal deep learning network for highway traffic flow forecasting[J]. *International Journal of Distributed Sensor Networks*, 2019, 15(2): 1-12.
- [16] Jeong Y S, Byon Y J, Castro-Neto M M, et al. Supervised weighting-online learning algorithm for short-term traffic flow prediction[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(4): 1700-1707.
- [17] Lippi M, Bertini M, Frasconi P. Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(2): 871-882.
- [18] Hu W B, Yan L P, Liu K Z, et al. A short-term traffic flow forecasting method based on the hybrid PSO-SVR[J]. *Neural Processing Letters*, 2016, 43(1): 155-172.
- [19] Bing Q C, Gong B W, Yang Z S, et al. A short-term traffic flow local prediction method of combined kernel function relevance vector machine[J]. *Journal of Harbin Institute of Technology*, 2017, 49(3): 144-149.
- [20] Huang G B, Wang D H, Lan Y. Extreme learning machines: A survey[J]. *International Journal of Machine Learning and Cybernetics*, 2011, 2(2): 107-122.
- [21] Law A, Ghosh A. Multi-label classification using a cascade of stacked autoencoder and extreme learning machines[J]. *Neurocomputing*, 2019, 358(9): 222-234.
- [22] Li J, Zheng S U, Tan Y. Adaptive fireworks algorithm[C]. *Proceedings of 2014 IEEE Congress on Evolutionary Computation*. Piscataway: IEEE, 2014: 3214-3221.
- [23] Chen R Q, Yu J S. Prediction of dissolved oxygen concentration in wastewater treatment process based on improved FWA-NN[J]. *China Environmental Science*, 2018, 38(10): 3739-3746.
- [24] Portes L L, Aguirre L A. Impact of mixed measurements in detecting phase synchronization in networks using multivariate singular spectrum analysis[J]. *Nonlinear Dynamics*, 2019, 96(3): 2197-2209.
- [25] Silva E S, Hassani H, Ghodsi M, et al. Forecasting with auxiliary information in forecasts using multivariate singular spectrum analysis[J]. *Information Sciences*, 2019, 479(4): 214-230.

作者简介

陈如清(1979—),男,教授,博士,从事复杂工业过程建模与故障诊断等研究, E-mail: ruqing2002@163.com;

李嘉春(1982—),女,讲师,从事智能信息处理的研究, E-mail: 36963992@qq.com;

俞金寿(1939—),男,教授,博士生导师,从事工业过程建模与控制、先进控制等研究, E-mail: jshyu@ecust.edu.cn.

(责任编辑: 闫妍)