

DOI:10.19651/j.cnki.emt.2210372

基于奇异谱分析与LSSVM算法的列车 无线网络控制时延预测方法

窦顺坤¹ 李常贤² 张丽艳³

(1.大连交通大学自动化与电气工程学院 大连 116028; 2.大连交通大学轨道交通装备设计与制造技术国家地方联合工程研究中心 大连 116028; 3.大连交通大学计算机与通信工程学院 大连 116028)

摘要:无线网络控制是推进高速列车智能化的有利因素,无线网络时延作为一种典型的时间序列,存在随机性强、波动性大等问题导致预测难度大,针对这些问题提出了一种奇异谱分析-改进粒子群优化LSSVM的无线网络时延预测模型。首先将获取到的时延序列通过Cao方法确定窗口长度,再将时延序列通过奇异谱分析得到一系列子序列,将各子序列采用混沌粒子群优化后的LSSVM模型进行预测,最后将所有子序列预测值进行叠加得到最终预测结果,仿真结果表明,该模型MAPE、MSE及MAE相比小波分解模型分别降低了2.8%、1.055、0.44;相比EMD分解模型分别降低了7.4%、3.377、1.118;相比CEEMD分解模型分别降低了6.2%、2.568、0.974,精度明显高于其他模型。

关键词:无线网络时延;奇异谱分析;混沌粒子群;LSSVM

中图分类号: TP18 **文献标识码:** A **国家标准学科分类代码:** 510.40

Prediction method of train wireless network control delay based on singular spectrum analysis and LSSVM algorithm

Dou Shunkun¹ Li Changxian² Zhang Liyan³

(1. School of Automation and Electrical Engineering, Dalian Jiaotong University, Dalian 116028, China;
2. National and Local Joint Engineering Research Center for Rail Transit Equipment, Dalian Jiaotong University, Dalian 116028, China;
3. School of Computer and Communication Engineering, Dalian Jiaotong University, Dalian 116028, China)

Abstract: Wireless network control is a favorable factor to promote the intelligence of high-speed trains. As a typical time series, wireless network delay has strong randomness, large volatility and other problems leading to difficult prediction. In view of these problems, a wireless network delay prediction model with singular spectrum analysis-improved particle swarm optimization and LSSVM is proposed. The length of the window was first determined by the Cao method, the delay sequences were analyzed by singular spectral analysis to obtain a series of subsequences. Each subsequence was predicted using the LSSVM model optimized for the chaotic particle swarm. Finally, all the subsequence predicted values were superimposed to obtain the final prediction results, the simulation results show that the average absolute percentage error (MAPE), mean squared error (MSE) and average absolute error (MAE) are 2.8%, 1.055 and 0.44 lower respectively compared with the wavelet decomposition model. Compared with the EMD decomposition model, 7.4%, 3.377 and 1.118 decreased, respectively. Compared with the CEEMD decomposition model, it was reduced by 6.2%, 2.568, and 0.974, respectively. The accuracy was significantly higher than that in the other models.

Keywords: wireless network delay; singular spectrum analysis; chaos particle groups; LSSVM

0 引言

列车网络控制(train communication network, TCN)是现行高速列车网络控制的主要方式,为推进高速列车智能

化发展,2015年欧盟启动了Roll2 Rail项目,并将无线TCN列为其中一项研究计划,无线TCN系统采用纯无线或无线与有线相结合的传输方式,提供了更加便利的设备接口。但由于迄今为止并没有专为列车无线通信划分的通

频段,同时无线信号的传输在开放的环境下进行,因此易受到外界干扰,从而产生难以确定的时延波动。当时延较大时往往会使得控制响应不及时,对列车的安全运行造成隐患,从而制约着无线 TCN 的发展。

为解决上述问题,除采用专用的通信频段外,对时延的预测与补偿可以作为另一方向。时延作为典型的时间序列,一些学者对此展开了研究,赵景波等^[1]采用麻雀搜索算法优化神经网络对控制系统中的时延进行了预测。汪知宇等^[2]采用改进粒子群算法(particle swarm optimization,PSO)优化最小二乘支持向量机(least squares support vector machine, LSSVM)模型用于列车网络时延的预测。时维国等^[3]分别采用自回归模型(autoregressive, AR)模型、相空间重构^[4]等方式对网络控制中的时延进行了预测。由于时延序列存在非线性强、随机度高等问题,对此,田中大等^[5]采用经验模态分解(empirical mode decomposition, EMD)将时延序列中的信息分解到不同的时间尺度,但在文中明确表示,大量信息包含在第 1 个分量中,因此导致第一分量预测难度依然较高。王金海等^[6]运用近似熵原理,将熵值相近的序列叠加处理以降低运算量。小波分解作为另外一种常见分解算法被广泛用于时间序列的处理^[7],但在分解的过程中小波基的选择过于困难,且不同的小波基对于最终的分解结果影响较大。奇异谱分析(singular spectrum analysis, SSA)方法是一种基于特征值的分析处理方法^[8],可以有效的将序列中隐含的信息分解到不同特征值对应的特征向量中,避免了上述分解方法存在的问题,因此可用于时延序列的预测,但窗口长度的选择往往根据个人经验进行选取,对此可借鉴相空间重构理论,采用 Cao 方法对窗口长度进行选取。对于预测模型的选择可采用 LSSVM 模型^[9],LSSVM 在处理小样本、非线性、高纬度序列时有较好表现,不存在如神经网络模型易陷入局部最优^[10]、回归建模复杂度高等问题^[11],因此可用于时延预测,对于预测中的参数选取问题,可采用智能算法进行获取。

基于以上分析,本文选取奇异谱分析对无线时延序列进行处理,设计了一种基于 SSA-PSO-LSSVM 的无线时延预测方法,旨在对列车无线控制环节产生的时延进行精确预测,从而保障列车的安全运行。该方法首先将实验平台所获取的时延通过 Cao 方法选取适合的窗口长度^[12],然后依据选取的窗口长度对时延序列进行奇异谱分析;然后将分解后的子序列输入到 PSO-LSSVM 模型进行预测,同时为避免粒子早熟等问题,在普通粒子群算法中加入了混沌搜索与自适应权值调整;最终将各分量预测结果进行叠加得到最终预测结果。最终结果表明,该模型具有较高的预测精度,可为之后的时延补偿提供基础,为列车智能化提供参考。

1 算法原理

1.1 奇异谱分析

SSA 是一种对非线性时间序列进行分解重构的方法,

其本质是采用奇异值分解对序列进行处理分析,将原始序列中的大量信息分解到各子序列中,从而降低原始序列的复杂度,常与时间序列预测模型结合使用,可以有效降低预测难度。其主要步骤包括:构建轨迹矩阵、奇异值分解、分组和对角平均化等,具体流程如下。

将长度为 n 的时间序列 $T = \{t_1, \dots, t_n\}$ 构造为以下形式的 $L \times N$ 阶轨迹矩阵 \mathbf{X} :

$$\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_K] = \begin{bmatrix} t_1 & t_2 & \dots & t_K \\ t_2 & t_3 & \dots & t_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ t_L & t_{L+1} & \dots & t_N \end{bmatrix} \quad (1)$$

式中: $\mathbf{X}_i = [t_i, t_{i+1}, \dots, t_{i+L-1}]^T$, $K = N - L + 1$, L 为窗口长度,通常 $L \leq N/2$ 。

对轨迹矩阵 \mathbf{X} 进行奇异值分解,记 $\mathbf{S} = \mathbf{X}\mathbf{X}^T$, 其特征值为 $\lambda_1, \lambda_2, \dots, \lambda_L$, 对应的特征向量为 u_1, u_2, \dots, u_L , 那么对 \mathbf{X} 的奇异值分解可表示为以下形式:

$$\mathbf{X} = \mathbf{E}_1 + \mathbf{E}_2 + \dots + \mathbf{E}_d \quad (2)$$

式中: $\mathbf{E}_i = \sqrt{\lambda_i} \mathbf{U}_i \mathbf{V}_i^T$, $\mathbf{V}_i = \mathbf{X}^T \mathbf{U}_i \sqrt{\lambda_i}$, 其中 $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_d$ 为主成分, d 为非零特征值的个数, $\sqrt{\lambda_i}, \mathbf{U}_i, \mathbf{V}_i$ 依次为轨迹矩阵的第 i 个奇异值、左奇异向量、右奇异向量。

通过特征值或贡献率可对获得的子序列进行分析评价,贡献率越大,重构序列时该序列所占比例越大,贡献率 G 可通过特征值按以下方式求得:

$$G_i = \frac{\lambda_i}{\sum_{i=1}^d \lambda_i} \quad (3)$$

设 \mathbf{Z} 为一个 $L \times K$ 阶的矩阵,其中元素为 z_{ij} , 令 $L^* = \min(L, K)$, $K^* = \max(L, K)$, $N = L + K - 1$, 若 $L < K$, 则让 $z_{ij}^* = z_{ij}$, 反之, $z_{ij}^* = z_{ji}$ 。此时可采用对角平均的方式重构序列:

$$y_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} z_{m, k-m+2}^*, & 1 \leq k \leq L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} z_{m, k}^*, & L^* \leq k \leq K^* \\ \frac{1}{N-n} \sum_{m=k}^{N-K^*+1} z_{m, k-m+2}^*, & K^* \leq k \leq N \end{cases} \quad (4)$$

在上述过程中 L 的大小需要主动设置,且 L 的大小直接影响着分解的优劣,其中 L 越大越有利于数据中成分的提取,但同时会产生大量的子序列,不利于数据的处理; L 越小产生的子序列越少,有利于观测序列的周期性波动,但往往导致数据分解不充分。针对此现象一些学者使用 Cao 方法对嵌入维度进行选取,并取得了良好的分解效果。

1.2 LSSVM 算法

支持向量机(support vector machine, SVM)相比于神经网络具有较强的泛化能力,适用于强非线性序列,但存在计算时间长、计算量大等缺点,因此 Suykens 和 Vandewalle

提出了LSSVM算法,LSSVM算法将SVM中的不等式约束转化为等式约束,采用误差平方和损失函数代替SVM中的二次规划问题,从而转化为线性矩阵求解问题,极大程度上降低了计算量。

对于给定的数据集 $\{x_i, y_i\}$, 其中 $1 < i < n, x_i$ 为输入集, y_i 为输出集, 通过非线性函数 $\varphi(x)$ 将 n 维输入集对应的映射到高维空间, 此时, 在LSSVM上的回归估计函数可以表示为:

$$y(x) = \omega \cdot \varphi(x) + b \quad (5)$$

式中: ω 为权值向量, b 为偏置向量。根据风险最小化原理, 回归问题可以表示为如下的约束优化问题:

$$\begin{cases} \min J(\omega, \zeta) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} \gamma \sum_{i=1}^n \zeta_i^2 \\ \text{s. t. } y_i = \omega^T \cdot \varphi(x) + b + \zeta_i \end{cases} \quad (6)$$

式中: γ 为正则化参数, ζ_i 为误差向量。

为求解上述优化问题, 可通过构建Lagrangian函数进行求解, 此时转换为:

$$L = J(\omega, \zeta) - \sum_{i=1}^n \alpha_i [\omega^T \cdot \varphi(x_i) + b + \zeta_i - y_i] \quad (7)$$

式中: α_i 为拉格朗日乘数。

根据Karush-Kuhn-Tucker条件, 通过对 ω, ζ, α 求偏导, 令其偏导数等于0, 可将 ω, ζ 约去, 进一步得到下式:

$$\begin{bmatrix} 0 & l^T \\ l & K(x_i, x_j) + \gamma^{-1} \mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (8)$$

式中: \mathbf{I} 为单位阵, $l = [1, 1, \dots, 1]^T$, $K(x_i, x_j)$ 为核函数, 通常选择径向基核函数(radial basis function, RBF):

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (9)$$

最终回归估计函数形式如下:

$$y(x) = \sum_{j=1}^n \lambda_j K(x, x_j) + b \quad (10)$$

式中: x 为输入样本集, x_j 为核函数中心, $\lambda_j \neq 0$ 。

对于LSSVM预测而言, γ 与 σ 两个参数选取的优劣直接影响着模型的预测性能, 因此本文将LSSVM的预测值与真实值的均方误差作为改进PSO算法的适应度函数, 从而对两个参数进行迭代寻优。

1.3 改进粒子群优化算法

PSO算法由Kennedy和Eberhart在1995年提出, 是一种群体智能的优化算法, 算法实现简单、需要调节参数较少从而被广泛应用, 但传统粒子群优化算法存在局部搜索能力较弱、易陷入局部最优、后期收敛速度慢等缺点, 对此本文在传统PSO中加入了混沌思想用来改善算法性能, 传统粒子群算法可参考相关文献[13]。

混沌运动对初始条件敏感, 且具有随机性、遍历性、不重复性, 因此在粒子群优化中加入混沌思想有利于提高种群收敛速度与全局搜索能力, 具体步骤如下:

首先将全局最优粒子的位置 g_{best} 按以下方式映射到

Logistic方程的定义域 $[0, 1]$ 上:

$$y^{(t)} = \frac{g_{best}(t) - x_{min}}{x_{max}} \quad (11)$$

式中: $y^{(t)}$ 为经 t 次迭代映射后的相对位置, x_{max} 和 x_{min} 分别为粒子范围的上界与下界。

将映射后的粒子采用Logistic方程混沌化:

$$y^{(t+1)} = u y^{(t)} (1 - y^{(t)}) \quad (12)$$

式中: u 为混沌系数, 组织范围通常为 $[0, 4]$, 本文取 $u = 4$ 。

最后, 将经多次迭代的混沌变量解映射到原区间, 构成新的全局最优值。在粒子更新时, 惯性权值 ω 影响着粒子移动的速度与粒子的位置, ω 较大有利于粒子全局搜索, ω 较小则有利于粒子局部搜索, 因此, 对 ω 单一取值不利于种群的收敛。对此, 本文采用线性权值递减的方式处理惯性权值, 从而保障迭代过程能够快速收敛。

2 无线时延预测方法

2.1 数据来源

本文数据来源为自行搭建的基于无线通信的列车网络控制半实物实验平台, 平台采用无线与有线相结合的方式, 其中列车骨干网为有线方式传输, TCN通过无线接入点与客户端结合的方式实现无线方式传输, 平台包含完整的控制系统, 由中央控制单元、牵引控制单元模拟设备、输入输出模块、辅助控制单元模拟设备、司机显示屏、监控模块等组成, 数据发送端口共计56个, 通信频段为802.11n协议下的5GHz, 带宽为20MHz, 发送端采用定向天线, 接收端采用全向天线, 所传输数据为列车实时以太网数据(train real-time data protocol, TRDP), 不考虑外界信息涌入情况下, 负载率约为20%。在发送端对数据包进行标记并记录发送时间, 当接收端传来回复信息时停止计时, 并记为一次往返时延, 时间标记通过Wireshark软件查看并计算获取。具体实验平台如图1所示。



图1 无线通信列车网络控制半实物实验平台

2.2 预测流程及模型建立

在预测过程中, LSSVM的核函数与正则化参数直接影响着预测结果的精度, 因此采用改进的PSO算法进行参数寻优, 具体步骤如下:

- 1) 初始化 LSSVM 及改进 PSO 的粒子速度和位置。
- 2) 将 LSSVM 的预测值和实际值的均方误差作为改进 PSO 的适应度函数, 将适应度作为评价标准, 依据速度和位置更新公式更新最优粒子的速度及位置。
- 3) 对最优粒子的位置进行混沌优化。
- 4) 将步骤 2)、3) 进行迭代, 当最优粒子不变或达到最大迭代次数时停止, 输出 LSSVM 所需的最优参数。

具体的无线时延预测模型如图 2 所示。

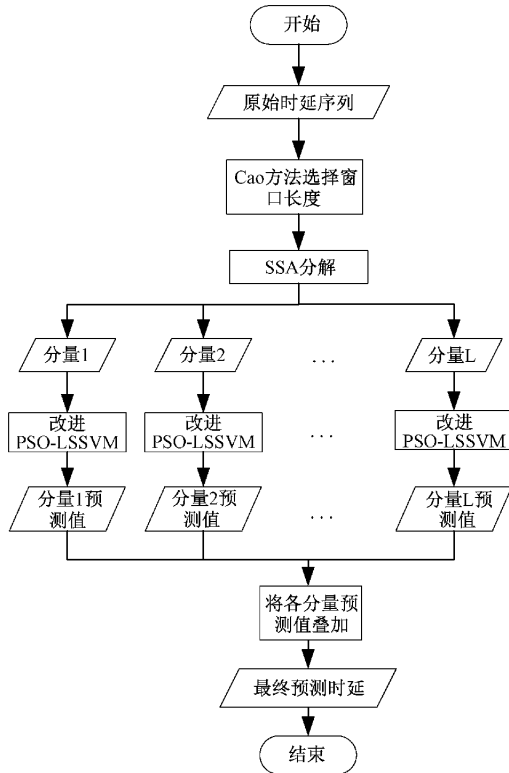


图 2 无线时延预测模型

2.3 评价指标

为了对预测结果更加直观的分析, 本文选取均方误差 (mean square error, MSE)、平均绝对误差 (mean absolute error, MAE)、平均绝对百分比误差 (mean absolute percentage error, MAPE) 作为衡量预测性能的标准, 3 种标准表示形式如下:

$$MSE = \frac{1}{N} \sum_{j=1}^N (y_j - y_j^*)^2 \quad (13)$$

$$MAE = \sqrt{\frac{1}{L} \sum_{k=1}^L |y_j - y_j^*|} \quad (14)$$

$$MAPE = \frac{1}{L} \sum_{k=1}^L \left| \frac{y_j^* - y_j}{y_j^*} \right| \times 100\% \quad (15)$$

式中: N 为测试的样本数, y_j 为时延的实际值, y_j^* 为时延的预测值。

样本熵 (sample entropy, SampEn) 通过衡量序列中产生新信息概率的大小对序列的复杂度做出评判^[14], 样本熵

的值越大, 则序列中包含的信息越多, 序列越复杂, 可作为时延序列复杂度衡量标准。

3 实验及仿真

3.1 时延数据的奇异谱分析

通过对原始序列进行 SSA 分析, 可大幅降低序列复杂度, 从而保障预测的顺利进行。数据来源以搭建的实验平台中牵引制动部分为背景进行获取, 牵引控制系统由牵引控制单元与中央控制单元组成, 包含 20 个数据发送端口, 所发送的数据包大小均为 1 432 byte。当发送周期为 32 ms、无外界干扰时, 时延情况如图 3(a) 所示; 当发送周期为 64 ms、外界存在不确定干扰时, 时延情况如图 3(b) 所示。

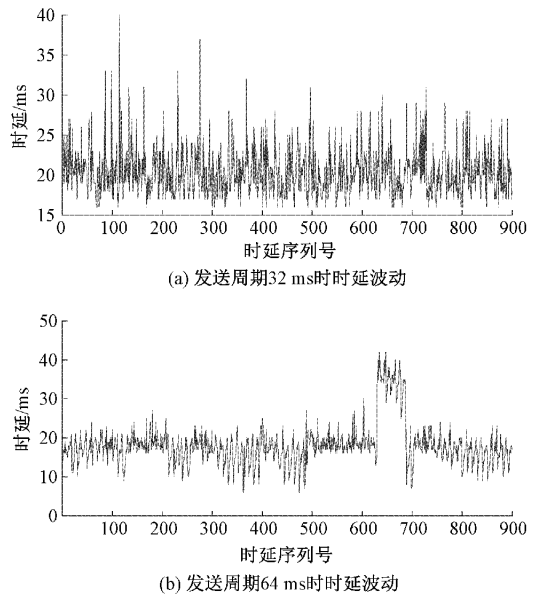


图 3 不同条件下的无线时延序列

由图 3 可以看出, 时延序列呈现高频波动性与随机性, 且无明显规律性; 当出现干扰时, 时延值迅速上升, 干扰消失时, 时延又在极短的时间回落, 由此可以看出无线时延值极易受到外界干扰。通过对两个序列进行样本熵测试, 得到序列 a 与序列 b 的 SampEn 分别为 2.204 与 2.017, 这表明无线时延序列中包含大量信息、序列复杂, 采用常见的预测模型将难以对其进行拟合。

通过 SSA 分析的方式可以将原始序列中的大量信息分解到不同的特征向量中, 从而降低预测难度。分解的关键在于窗口长度的进行选取, 文献[12]中表明奇异谱分析中窗口长度选取与相空间重构存在相似之处, 因此, 窗口长度可以通过 Cao 方法进行选取。

通过 Cao 方法测试得到两序列最佳窗口长度均为 8, 分别对两序列进行分解, 序列 a 分解后前 4 个分量的波动情况如图 4(a) 所示, 序列 b 分解后前 4 个分量的波动情况如图 4(b) 所示。对比图 3、4 可以看出原序列时延值大致

在各自分量1的趋势上往复波动,其他分量则呈高频波动性。为对各分量进行定量分析,文中选取了贡献率、与原始序列的 Pearson 相关系数^[15]、样本熵作为分析标准,并进行了统计,最终得到结果如表1所示。

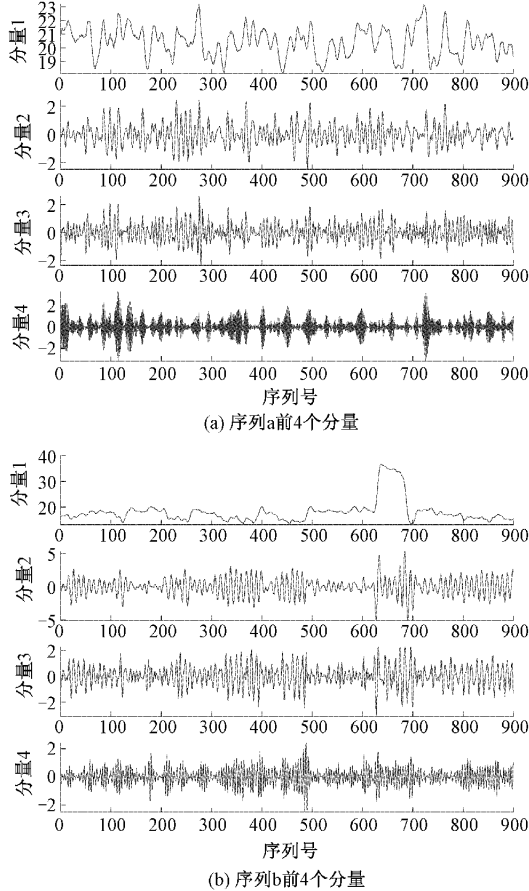


图4 无线时延序列SSA处理

表1显示,两序列所产生的分量1均占据极高的贡献率,因此可以认为所获得的首个分量能在较大程度上表现出原序列的趋势,明显不同之处在于与原序列的相关性,通过上述分析可知,分量1表现为原序列的趋势分量,当序列较为稳定时,如序列a,此时时延在一个固定范围内波动,波动特性是影响曲线上限的主要因素,因此趋势分量与原始序列的相关性较低。当序列存在突发干扰时,如序列b,此时突变取代波动成为影响曲线上限的主要因素,更能表现原序列的外在特征,因此与原序列呈现出高度相关性。而通过对比发现,各分量样本熵与原序列相比有明显降低,均处于1以下,这是一个较低的数值,因此可以表明,经SSA分析后,信息被分解到各分量中,预测难度大幅下降。

为表明SSA分析对于所获取无线时延的优越性,本文选取了文献[5]中EMD分解模型、互补集合经验模态分解(complementary ensemble empirical mode decomposition, CEEMD)模型^[16]以及文献[7]中小波分解模型进行对比,

表1 经SSA分析后各分量参数

| 序列 | 分量 | 贡献率/ % | Pearson 相关系数 | 样本熵 |
|------|------|-----------|-----------------|-------|
| 序列 a | 分量 1 | 98.14 | 0.473 | 0.614 |
| | 分量 2 | 0.32 | 0.571 | 0.735 |
| | 分量 3 | 0.29 | 0.604 | 0.918 |
| | 分量 4 | 0.28 | 0.431 | 0.609 |
| | 分量 5 | 0.25 | 0.561 | 0.885 |
| | 分量 6 | 0.24 | 0.595 | 0.956 |
| | 分量 7 | 0.24 | 0.555 | 0.723 |
| | 分量 8 | 0.24 | 0.539 | 0.791 |
| 序列 b | 分量 1 | 97.36 | 0.850 | 0.542 |
| | 分量 2 | 1.01 | 0.484 | 0.717 |
| | 分量 3 | 0.58 | 0.456 | 0.955 |
| | 分量 4 | 0.27 | 0.284 | 0.708 |
| | 分量 5 | 0.23 | 0.295 | 0.886 |
| | 分量 6 | 0.20 | 0.256 | 0.677 |
| | 分量 7 | 0.20 | 0.259 | 0.704 |
| | 分量 8 | 0.18 | 0.183 | 0.500 |

并将样本熵以及缺失对应分量情况下的MSE作为对比项,MSE越大表明此分量对原序列的组成贡献越大。各方法对序列分解后都存在一个包含信息较多的主要分量,如上述SSA中的趋势分量、小波分解后的近似分量、EMD及CEEMD分解后的残差,主要分量缺失时MSE极大,为便于展示,缺失主要分量时的MSE本文不予给出。

以序列a为例进行对比分析,各方法分解后结果对比如图5所示。其中小波分解选则db3小波基,分解层数为7;EMD分解由系统自主选择;CEEMD分解时,设置噪声幅值为0.2,添加次数为100次,分解层数为7。将小波分解、EMD分解、CEEMD分解的第8个分量设置为主要分量。由图中可以看出各方法主要分量的样本熵均处于较低水平,因此尽管其对原序列影响较大,但预测难度低,不会对最终预测结果造成较大影响。观察发现,经SSA获得的各序列样本熵均处于[0.6, 1]之间,对比原始序列样本熵2.204有明显降低,且MSE较小,该方法有利于降低预测难度。其他3种方法分解后获得的首个分量样本熵均在2左右,所对应的MSE均超过4,这表明该分量预测难度较高,且预测不准确时对最终预测结果影响较大,经对比发现其他分量样本熵与MSE远低于首个分量,因此可以认为首个分量是造成预测误差的主要因素。

3.2 仿真与结果分析

本文仿真通过MATLAB2020b进行编写。为探究不同方法对最终预测结果的影响,选取小波-LSSVM预测模型、CEEMD-LSSVM预测模型、EMD-LSSVM预测模型与本文进行对比,其中小波分解选则db3小波基,分解层数为7;EMD分解由系统自主选择;CEEMD分解时,设置噪声幅

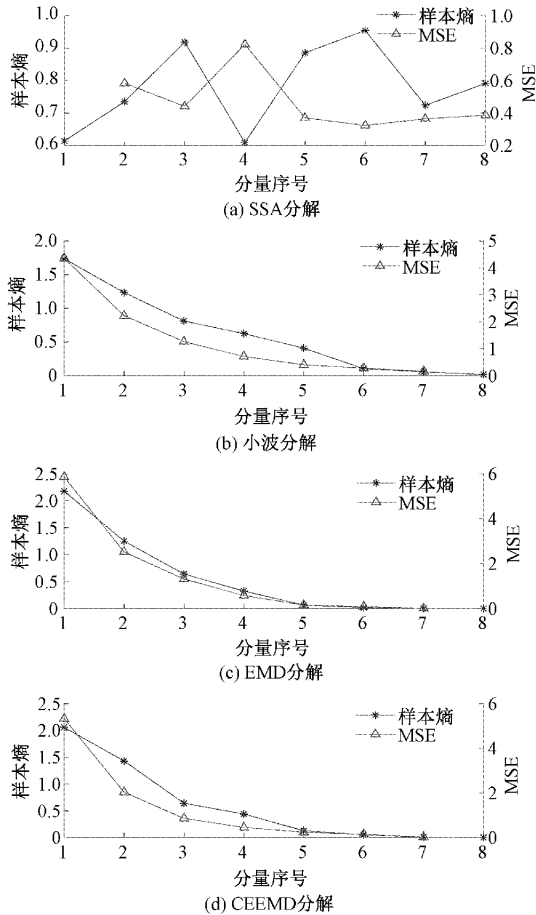


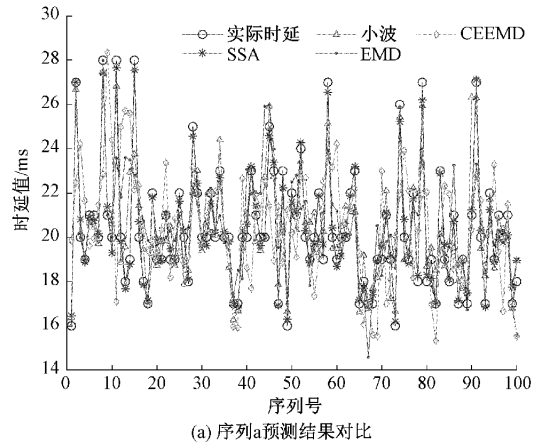
图 5 各方法分解序列 a 结果对比

值为 0.2, 添加次数为 100 次, 分解层数为 7。为保证公平性, 模型参数均由改进后的 PSO 算法进行选取。将前 800 组数据作为训练集, 后 100 组数据最为测试集, 所有改进 PSO 算法中种群规模均设为 20, 学习因子 $c_1 = c_2 = 2$, 最大迭代次数为 200, 混沌参数 $u = 4$, LSSVM 中嵌入维度为 10, γ 的取值范围设为 $[0, 1\ 000]$, σ 的取值范围设为 $[0, 100]$ 。其中本文模型经优化后各分量 LSSVM 最优参数如表 2 所示。最终预测结果如图 6 所示, 预测误差对比如表 3 所示。

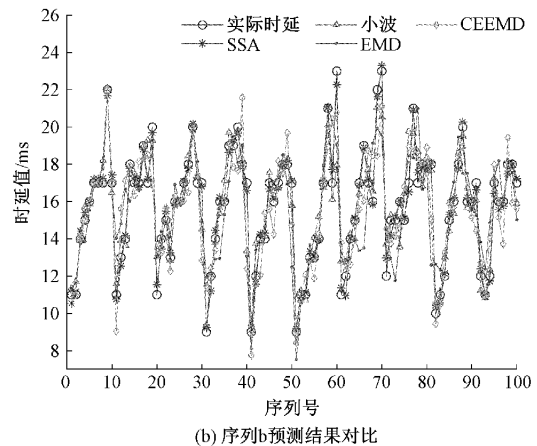
由图 6 可知, 本文模型对于不同情况下产生的时延序列均具有较好的拟合效果。表 3 显示了不同预测模型的预测误差, 对于序列 a 而言, 本文模型 MAPE、MSE 及 MAE 相比小波分解模型分别降低了 3.5%、1.846、0.716, 相比 EMD 分解模型分别降低了 5.3%、3.455、1.068, 相比 CEEMD 分解模型分别降低了 11%、10.12、2.255; 对于序列 b 而言, 本文模型 MAPE、MSE 及 MAE 相比小波分解模型分别降低了 2.8%、1.055、0.44, 相比 EMD 分解模型分别降低了 7.4%、3.377、1.118, 相比 CEEMD 分解模型分别降低了 6.2%、2.568、0.974。由此可知, 该模型对于不同情况下产生的时延预测精度均明显高于其他预测模型。EMD 分解及 CEEMD 分解两种方式获得的分量复杂

表 2 优化后各分量最优参数

| 序列 | 分量 | γ | σ |
|------|------|----------|----------|
| 序列 a | 分量 1 | 984.2 | 1.23 |
| | 分量 2 | 610.62 | 33.09 |
| | 分量 3 | 471.34 | 8.42 |
| | 分量 4 | 987.42 | 6.39 |
| | 分量 5 | 29.48 | 12.4 |
| | 分量 6 | 964.2 | 9.27 |
| | 分量 7 | 177.41 | 0.3 |
| | 分量 8 | 845.34 | 12.08 |
| 序列 b | 分量 1 | 448.63 | 2.12 |
| | 分量 2 | 919.57 | 0.33 |
| | 分量 3 | 467.44 | 0.014 |
| | 分量 4 | 27.28 | 0.16 |
| | 分量 5 | 344.24 | 8.14 |
| | 分量 6 | 996.28 | 10.26 |
| | 分量 7 | 142.11 | 0.14 |
| | 分量 8 | 477.38 | 2.17 |



(a) 序列 a 预测结果对比



(b) 序列 b 预测结果对比

图 6 各方法时延预测结果对比

度不平均, 大量信息包含在第 1 个分量中, 导致第一分量的预测难度较高; 小波分解同样存在信息分解不平均的问题,

表3 各方法预测误差对比

| 序列 | 方法 | 本文 | 小波 | EMD | CEEMD |
|------|------|-------|-------|-------|--------|
| 序列 a | MAPE | 0.017 | 0.052 | 0.070 | 0.127 |
| | MSE | 0.183 | 2.029 | 3.628 | 10.303 |
| | MAE | 0.345 | 1.061 | 1.413 | 2.600 |
| 序列 b | MSE | 0.123 | 1.178 | 3.500 | 2.691 |
| | MAE | 0.285 | 0.725 | 1.403 | 1.259 |
| | MAPE | 0.019 | 0.047 | 0.093 | 0.081 |

导致第一分量预测难度较高,尽管该问题通过选择适合的小波基与分解层数将有所改善,但选取过程过于复杂;通过Cao方法选择最佳窗口长度的奇异谱分析方法,能够有效将不同尺度的信息分解到对于特征向量中,各分量预测难度较为平均,有效降低了模型预测难度。同时改进后的粒子群算法避免了粒子早熟或陷入局部最优,为LSSVM模型预测提供了有效保障。

4 结 论

针对无线时延复杂度高、波动性大、预测难度高等问题,以自行搭建的无线网络列车控制系统半实物仿真平台所获取的数据为基础,首先采用SSA算法对原始数据进行了分析,并引入了Cao方法用于解决窗口长度的选取问题,结果表明,与其他分解方法相比,经SSA分析后的子序列样本熵平均,无极端现象;其次通过LSSVM模型对各子序列进行预测,对于LSSVM所需参数,采用了引入混沌因子的PSO算法进行选取,在与其他模型的对比结果中表明,该模型具有更高的预测精度与适用性,进一步表明SSA算法对无线时延的优越性。本文方法为之后的时延补偿提供了基础,也为列车网络控制的无线化提供了参考。但由于分解后增大了计算量,在后续研究中有待改进。

参考文献

[1] 赵景波,朱敬旭辉,邱腾飞,等.基于SSA-BP神经网络的网络控制系统延时预测[J].计算机仿真,2022,39(3):263-267.

[2] 汪知宇,张彤.基于改进LS-SVM算法的列车通信网络时延预测方法[J].城市轨道交通研究,2021,24(1):101-106.

[3] 时维国,邵诚,孙正阳.基于AR模型时延预测的改进GPC网络控制算法[J].控制与决策,2012,27(3):477-480.

[4] 时维国,许超.基于相空间重构与鲁棒极限学习机的时延预测[J].系统工程与电子技术,2019,41(2):417-422.

[5] 田中大,高宪文,李琨.基于EMD与LS-SVM的网络控制系统时延预测方法[J].电子学报,2014,42(5):868-874.

[6] 王金海,史梦颖,张兴华.基于EMD和ApEn特征提取的心律失常分类研究[J].仪器仪表学报,2016,37(S1):168-173.

[7] 谢劭峰,苏永柠,刘春丽,等.基于小波分解与GA-LSSVM的GPS可降水量短临预报[J].大地测量与地球动力学,2019,39(5):487-491.

[8] 马朝永,申宏晨,胥永刚,等.优化奇异谱分解方法在轴承故障诊断中的应用[J].轴承,2022(2):55-60.

[9] 全卫国,郭超宇,赵如意.基于改进麻雀算法优化LSSVM的再循环箱浆液密度预测模型[J].电子测量技术,2022,45(1):70-76.

[10] 王义,刘欣,高德欣.基于BiLSTM神经网络的锂电池SOH估计与RUL预测[J].电子测量技术,2021,44(20):1-5.

[11] 秦全德,黄兆荣,黄凯珊.一种基于局部回归的多尺度碳市场价格预测模型研究[J].运筹与管理,2022,31(1):107-114.

[12] 吴坚,项颂,阎诚,等.基于奇异谱分析的超短期风电功率多步预测[J].可再生能源,2021,39(11):1548-1555.

[13] 陈法法,刘帅,肖文荣,等.混沌粒子群优化RVM的滚动轴承早期故障诊断[J].电子测量与仪器学报,2018,32(8):9-16.

[14] 张韦,张永,骈晓琴,等.基于改进EMD样本熵和SVM的风机滚动轴承故障诊断[J].机电工程技术,2021,50(12):38-41,67.

[15] 姜丰,朱家玲,胡开永,等.Pearson相关系数评价ORC系统蒸发器特性的应用研究[J].太阳能学报,2019,40(10):2732-2738.

[16] 朱永强,王小凡.基于互补型集成经验模态分解和遗传最小二乘支持向量机的交通流量预测模型[J].科学技术与工程,2020,20(17):7088-7092.

作者简介

窦顺坤,硕士研究生,主要研究方向为轨道列车网络控制。

E-mail:doushunkun@163.com

李常贤,博士,教授,主要研究方向为列车网络控制及设备开发等。

E-mail:lichangxianzju@163.com

张丽艳,博士,副教授,主要研究方向为信号与信息处理与嵌入式系统开发等。

E-mail:dalianzhangliyan@163.com