

一种非平稳非线性频谱 占用度时间序列分析方法

魏鸿浩, 贾云峰

(北京航空航天大学电子信息工程学院, 北京 100191)

摘要: 针对传统频谱占用度分析模型由于未考虑序列的非线性非平稳特性, 导致无法准确描述频谱占用度特性的问题, 该文提出将集合经验模式分解 (EMD) 方法与人工神经网络 (ANN) 的方法结合应用于频谱占用度时间序列建模方法中, 采用 EMD + ANN 的频谱占用度序列建模和预测方法. 首先应用 EMD 分解算法把原始频谱占用度时间序列分解成不同尺度的基本模态分量, 再根据不同尺度的基本模态分量分别构建 ANN 模型, 提高了模型针对复杂频谱占用度时间序列的学习能力. 结合实测数据分析, 表明该模型相对传统频谱占用度模型具有更高的拟合和预测精度, 验证了该方法的正确性与有效性.

关键词: 电磁环境; 电磁频谱; 频谱占用度; 集合经验模式分解

中图分类号: TN914 **文献标识码:** A **文章编号:** 0372-2112 (2017)08-2026-05

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2017.08.031

A Method for Analysis of Non-linear and Non-stationary Spectrum Occupancy Time Series

WEI Hong-hao, JIA Yun-feng

(School of Electronics and Information Engineering, Beihang University, Beijing 100191, China)

Abstract: In order to analyze the non-linear and non-stationary spectrum occupancy time series which cannot be directly analyzed base on traditional time series method, a novel prediction modelling method of spectrum occupancy time series based on ensemble empirical mode decomposition (EMD) and Artificial Neural Network (ANN) is proposed. Firstly, the spectrum occupancy time series is decomposed into several intrinsic model function (IMF) so as to make every component stationary. Then in view of the stationary time series, a prediction ANN model is established correspondingly for each IMF. Simulative experiment for practical measured data shows that the proposed method has higher precision in comparison with other methods, i. e., effective to non-linear and non-stationary complicated spectrum occupancy time series prediction.

Key words: electromagnetic environment; electromagnetic spectrum; spectrum occupancy; ensemble empirical mode decomposition (EMD)

1 引言

频谱资源是不可再生资源, 频谱占用度分析是电磁环境监测、频谱管理和认知无线电领域中的重要内容, 无线电频谱占用和空闲两种状态是衡量电磁频谱利用程度与电磁环境性能的重要指标. 受复杂电磁环境、电磁辐射源频谱参数变化、空间中电磁波传播衰落特性以及检测设备的差异性影响, 频谱占用状态表现为与时间、空间、频率、用频检测设备性能差异相关联的

非线性、非平稳随机过程.

当前国内外对于频谱占用度的研究从结合随机过程理论定义频谱占用度开始, 文献[1]中对整个频谱感知过程中的样本相关性进行了研究, 文献[2~4]采用马尔科夫模型利用信道状态转移概率反映信道状态的相关性. 但这种方法在信道状态随时间长度呈指数增长的问题时遇到很大困难. 近些年出现基于时间序列分析构建频谱状态的时间序列模型, 文献[5~7]中采用神经网络模型对非线性频谱占用度进行了预测, 但

由于数据量大,训练速度极为缓慢,并且由于初始参数无法系统选定,不同门限间隔影响具有差异性等问题,预测精度变化较大.文献[8]根据 GSM900/1800 频段和电视频段频谱测试数据集和频谱占用时间序列,建立自回归时间序列(Auto Regressive AR)模型,文献[9~11]中对频谱占用度采用自回归移动平均(Auto Regressive Moving Average, ARMA)模型研究.文献[12]中对频谱占用状态的时域相关特性进行了研究.构建频谱占用状态变化等效模型.但这些基于时间序列分析的方法都要求原始序列是平稳时间序列,对于非平稳的频谱占用度数据通常需要采用差分平稳化的方法,但其平稳化效果并不是很好,这将直接影响模型的拟合和预测精度.

最近几年来开始出现用经验模式分解(Empirical Mode Decomposition, EMD)处理非线性、非平稳信号的数据.该方法以瞬态频率为表征信号交变的基本量,以本征模态函数(Intrinsic Mode Function, IMF)为基函数的时频分析方法.克服了传统方法中的全局性和非自适应性.该方法本质上讲是对一个信号进行平稳化处理,将信号中不同尺度波动和趋势进行分解,文献[13]中对 EMD 方法相对于其他方法的优越性进行了详细研究.但是 EMD 方法有一个显著地缺点存在模态混叠现象,降低了预测模型对各个分量的自适应性,进而影响到模型的预测精度.集合经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)方法把高斯白噪声加入到信号中,通过噪声辅助消除信号在时间尺度上的间隙性,避免由于 IMF 的不连续而造成的模态混叠现象.本文针对频谱占用状态的非线性、非平稳性的特征,将 EEMD 方法和人工神经网络(Artificial Neural Network)相结合,首先对实测频谱占用度序列运用 EEMD 技术进行预处理,将原始数据分解成若干个 IMF,在对每个 IMF 分量利用 ANN 分别建模,再对每个 IMF 分量利用 ANN 进行预测,然后利用各个分量的预测值重构出原始频谱占用度序列的预测序列,从而提高了预测精度,经过对比证明该模型的拟合和预测精度均优于其他模型

2 理论基础

2.1 EMD

经验模式分解是和相应的希尔伯特黄变换是由 N. E. Huang 等人提出一种新的数据分析处理方法,该方法从本质上讲是将信号进行平稳化处理,将信号中的不同尺度的波动和趋势逐渐进行分解,从而产生不同特征尺度的数据序列,这些分量称作本征模式函数 IMF, N. E. Huang 认为,对数据信号 EMD 分解时,得到的 IMF 必须满足以下 2 个条件^[14]:

(1)对于整个时间序列信号来说,其极点个数和过零点个数目必须相等或者至多相差一点.

(2)任意一点,由最大值和最小值所构成的包络线的平均值必须为 0.

对于序列 $A(t)$ 进行经验模式分解的过程如下:

(1)首先利用三次样条插值函数分别连接序列 $A(t)$ 中所有的局部最大值和局部最小值,构成上包络线 $b(t)$ 和下包络线 $c(t)$,求得局部均值序列 $m(t) = (a(t) + b(t))/2$.从而得到原始序列与包络线均值序列的差值: $A(t) - m(t) = h(t)$.

(2)检查序列 $h(t)$ 是否满足 IMF 的两个条件,如果满足执行步骤(3),如若不满足则将 $h(t)$ 替换为新的信号 $A(t)$,重复执行步骤(1).

(3)令 $C_i(t) = h(t)$,即 $h(t)$ 即为分解得到的第 i 个 IMF 分量 $C_i(t)$ 和信号的剩余部分 $r_i(t) = A(t) - C_i(t)$.

(4)检查是否满足分解停止条件,如果满足,则执行步骤(5);否则则令 $A(t) = r_i(t)$ 返回步骤(1), $i = i + 1$.

(5)分解完成,得到基本模式分量 $C_1(t), C_2(t), \dots, C_n(t)$ 和残余分量 $r_n(t)$,最终得到基于经验的模式解:

$$A(t) = \sum_{i=1}^n C_i(t) + r(t) \quad (1)$$

式中:原始信号被分解为 n 个基本模式分量 $C_1(t) \sim C_n(t)$ 和一个残差序列 $r(t)$,基本模式分量包括原始数据序列经分离后得到的不同尺度信息,残差序列反映了原始信号序列的平均趋势.

N. E. Huang 给出了上述过程的结束标准,定义了标准偏差:

$$SD = \frac{1}{T} \int_0^T \frac{|h_k(t) - h_{k-1}(t)|^2}{|h_{k-1}(t)|^2} \quad (2)$$

这个标准使得 $h(t)$ 充分接近 IMF 的要求,又能控制筛分的次数,从而所得到的 IMF 分量充分保留原始信号中的幅值调制信息.

2.2 EEMD

由于 EMD 在信号分析中边缘效应和尺度混合的问题存在,造成各种尺度振动模态的混合,甚至可以使个别 IMF 失去物理意义,进而造成模型对信号分量的适应性下降明显^[15],严重影响预测精度.

因此在 EMD 分解方法的基础上, N. E. Huang 等人于 2005 年提出了一种新的利用噪声辅助的新的分析方法,即 EEMD 法.该方法引入了白噪声,并在此基础上进行集合平均,从而避免了尺度混叠问题,使得分解得到的 IMF 保持了物理意义上的唯一性. EEMD 的具体步骤如下:

(1) 将白噪声序列添加到原始数据序列之中。

(2) 将添加了白噪声的信号序列按照 EMD 方法分解成多个 IMF。

(3) 重复迭代步骤(1)和(2), 每次添加不同的白噪声序列, 最后将每次分别得到的多组 IMF 的平均值作为最终分解结果。

(4) 通过在原始数据中加入一定比例的白噪声后, 进行经验模态分解, 最后进行集合平均, 使得加入的白噪声相互抵消, 保留了原始信号序列的信息, 可以在同一个 IMF 分量中比较辨识信号的不同尺度, 有效减少模态混叠问题, 进而有效改善 EMD 方法。

2.3 基于 EEMD 和神经网络的信号预测

利用神经网络对时间序列进行预测, 可以充分利用神经网络具有的并行处理、自适应、自组织、联想记忆、容错性和鲁棒性等特点. 可以从样本中自动学习, 逼近样本函数. 利用神经网络进行时间序列预测, 是指利用时间序列的前 K 个值 $X_n, X_{n-1}, \dots, X_{n-k+1}$ 去预测之后的 m 个值 $X_{n+1}, X_{n+2}, \dots, X_{n+m}$, 用公式表示如下:

$$X_{n+1}, X_{n+2}, \dots, X_{n+m} = F(X_n, X_{n-1}, \dots, X_{n-k+1}) \quad (3)$$

本文采用的预测模型方法步骤如图 1 所示

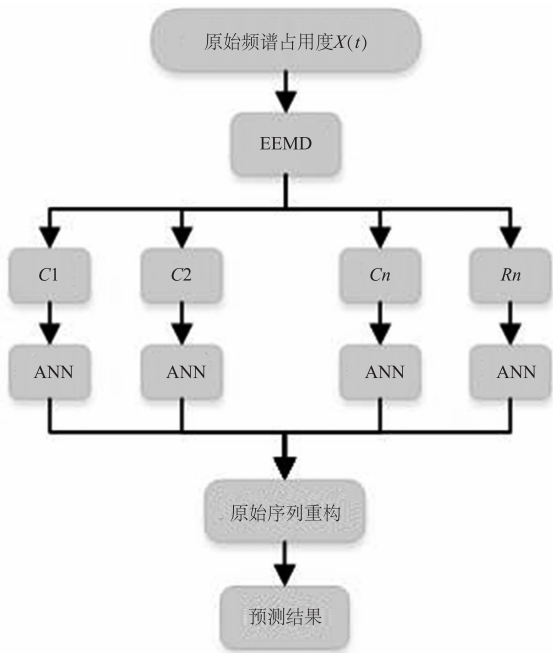


图1 频谱占用度时间序列建模流程

(1) 采用 EEMD 对原始频谱占用度序列进行分解, 得到不同尺度的 IMF, 分别包含了原始序列从高到低不同频率段的成分。

(2) 将各个 IMF 分量, 分别采用不同的神经网络进行预测。

(3) 将所有网络预测得到的预测值重构出原始频谱占用度序列的预测序列。

3 电磁频谱监测数据处理

本论文监测数据来自对北京航空航天大学新主楼周边电磁环境实测数据, 日期从 2016 年 3 月 1 日 00:00:00 持续到 2016 年 3 月 30 日 00:00:00, 检测设备由 Agilent E4408B 频谱分析仪、数据采集软件和 CS-AOS30-3000V 有源全向天线组成, 频率分辨率为 100Khz, 采样时间间隔为 10s, 每天的采样点数为 8640 个。

设 $M(f, t)$ 表示监测设备在频点 f , 时刻 t 所测无线电信号的场强值 (dBmV/m), 用“1”表示该信道被占用, “0”表示该信道空闲, η 表示频谱占用度噪声门限值, 那么

$$B(f, t) = \begin{cases} 1, & M(f, t) \geq \eta \\ 0, & M(f, t) < \eta \end{cases} \quad (4)$$

定义频谱占用度 SOR (Spectrum Occupancy Rate) 为某时刻监测频段内无线电业务占用度的信道个数与该频段包括的信道总数的比值, 即

$$\text{SOR}(t) = \frac{\sum_{f \in S} B(f, t)}{N} \quad (5)$$

其中 S 表示某监测频段频率分布, N 表示该业务所包换的信道个数, 与采样分辨率有关, 所以 SOR 随着时间先后的不同取值, 构成频谱占用度序列, 并且 $\text{SOR} \in [0, 1]$, 根据实测数据, 选取某天 88Mhz - 108Mhz, 860Mhz - 900Mhz, 450Mhz - 470Mhz 频谱占用度序列如图 2 所示。

4 基于 EEMD + ANN 建模和性能分析

4.1 EEMD 分解

这里以 88Mhz - 108Mhz 频谱占用度序列为例, 首先对原始时间占用度序列首先归一化, 然后进行 EEMD 分解, 每组白噪声序列的标准差均设为 0.1, 集合平均次数均为 500 次, 总共产生 13 个 IMF ($C_1(t) \sim C_{12}(t)$) 和一个剩余分量 $r(t)$, 如图 3 所示, 第一组为原始序列, 其余为序列分解后的分量, 展示了不同的频率和振幅. 为了对比, 选取 $C_6(t), C_7(t), C_8(t)$ 本征模态函数分别展示 EMD 和 EEMD 两种分解结果, 如图 4 所示, 可以看到这些分量相对于原时间序列变化比较平稳, 而且也比 EMD 分解的分量更加平稳, 有效的减少了模态混叠问题。

4.2 EEMD + ANN 预测模型

对原始时间序列经 EEMD 分解产生的模式分量, 通过一个分离的神经网络对被一个 IMF 进行预测, 最后将每个神经网络的预测值重构, 得到预测序列. 这里采用 15 - 30 - 1 三层网络结构, 进行单步预测, 每预测一个点后, 再把其输入模型, 来预测下一个点, 预测序列

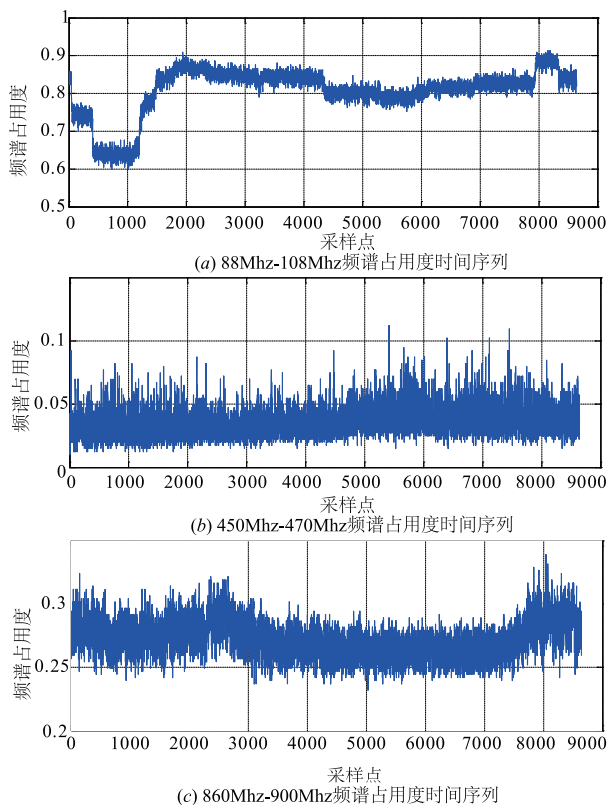


图2 无线电业务频谱占用度时间序列

与原始序列对比如图 5 所示,Targets 表示原始频谱占用度序列值,Outputs 为预测序列值,Errors 为预测误差。

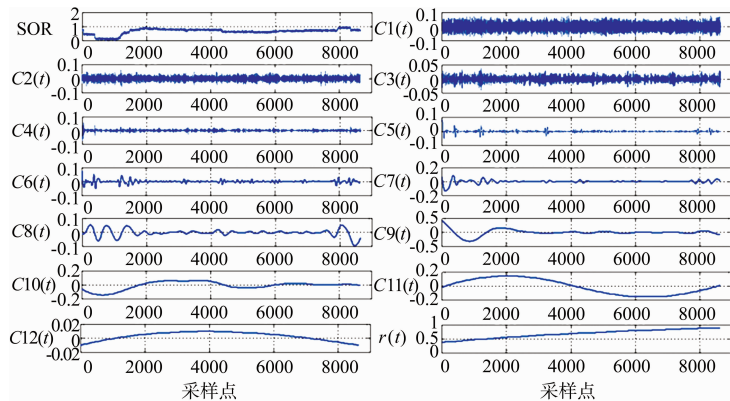


图3 频谱占用度序列EEMD分解

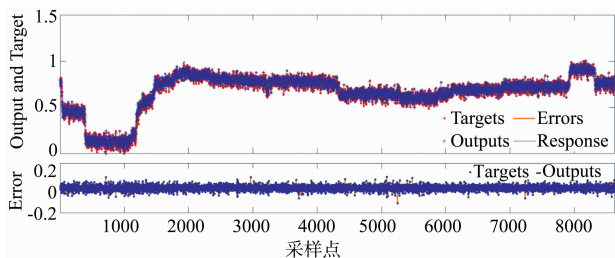


图5 EEMD-ANN预测结果与实际值误差对比

为了比较本文方法的预测效果,对所选 88MHz - 108MHz 频谱占用度序列数据,分别利用 EEMD + ANN、EMD + ANN、EMD + ARMA、EEMD + ARMA、ARMA 方法对所选频谱占用度数据进行预测,图 6 显示单步预测时针对不同数目的模型输入序列值,每种模型的均方误差曲线.图中 RMSE 表示均方值误差,Step 表示输入序列点数目,从图中可以看出虽然一个输入序列时误差比较接近,但是实际应用中几乎不会使用单点输入序列单步预测,当随着输入序列点数目增加,本文提出的 EEMD + ANN 方法预测精度明显优于其他方法,并且均方误差值趋于稳定。

为了验证本文方法的适用性,表 1 给出采取五种方法对不同频段频谱占用度序列进行预测时的最小均方误差值从表中可以看出,本文提出的 EEMD + ANN 方法预测精度明显优于其他方法。

表 1 最小均方误差值

| 建模方法 | 88 - 108MHz | 860 - 900MHz | 450 - 470MHz | 平均 RMSE |
|-------------|-------------|--------------|--------------|---------|
| ARMA | 0.0401 | 0.1063 | 0.1077 | 0.847 |
| EMD + ARMA | 0.0274 | 0.0695 | 0.0744 | 0.0571 |
| EEMD + ARMA | 0.0235 | 0.0663 | 0.0716 | 0.0538 |
| EMD + ANN | 0.0218 | 0.0661 | 0.0648 | 0.0492 |
| EEMD + ANN | 0.0167 | 0.0577 | 0.0616 | 0.0453 |

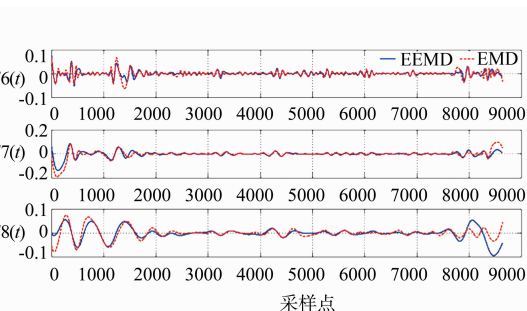


图4 频谱占用度序列EMD与EEMD分解对比

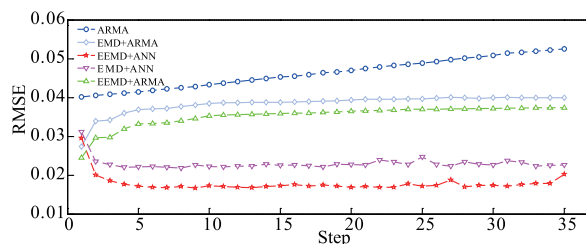


图6 EEMD+ANN与其他建模方法预测对比

5 结论

由于电磁频谱占用度时间序列具有非平稳性非线性的特点,采用传统时间序列分析方法,预测精度很难保证.为此,本文提出 EEMD + ANN 算法对频谱占用度序列进行拟合预测,将非平稳频谱占用度序列采用 EEMD 平稳化处理分别采用 ANN 方法建模,最后通过数据重构,得到预测结果,提高了单步预测模型精度.此方法针对其他分析方法体现了更高的预测和拟合精度,并通过实测电磁频谱数据验证了该方法的正确性.

参考文献

- [1] A D Spaulding and G H Hagn. On the definition and estimation of spectrum occupancy[J]. IEEE Transactions on Electromagnetic Compatibility, 1977, 19(3): 269 – 280.
- [2] López-Benítez M and Casadevall F. Discrete-time spectrum occupancy model based on Markov chain and duty cycle models[A]. IEEE International Symposium on Dynamic Spectrum Access Networks[C]. Aachen, Germany, IEEE, 2011. 90 – 99.
- [3] DALTA D, WYGLINSKI A M, MINDEN G J. A spectrum surveying framework for dynamic spectrum access networks[J]. IEEE Transactions on Vehicular Technology, 2009, 58(8): 4158 – 4168.
- [4] HAMID Eltom, SITHAMPARANATHAN Kandeepan, BILL Moran and ROBIN J Evans. Spectrum occupancy prediction using a hidden Markov model[A]. Signal Processing and Communication Systems[C]. Cairns, QLD, IEEE, 2015. 1 – 8.
- [5] S Bai, X ZHOU, F Xu. Soft decision spectrum prediction based on back-propagation neural networks[A]. IEEE International Conference on Computing, Management and Telecommunications[C]. DaNang, Vietnam; IEEE, 2014. 128 – 133.
- [6] X Xing, T Jing, W Cheng, Y Huo, X Cheng. Spectrum prediction in cognitive radio networks [J]. IEEE Wireless Communications, 2013, 20(4): 90 – 96.
- [7] BAI Suya, ZHOU Xin, XU Fanjiang. Spectrum prediction based on improved-back-propagation neural networks[A]. International Conference on Natural Computation [C]. Zhangjiajie, China; IEEE, 2015. 1066 – 1011.
- [8] GORCIN A, CELEBI H, KHALID A Q, et al. An autoregressive approach for spectrum occupancy modeling and prediction based on synchronous measurements[A]. IEEE 22th International Symposium on Personal, Indoor and Mobile Radio Commission[C]. Toronto, Canada; IEEE, 2011. 705 – 709.
- [9] 王磊, 谢树果, 苏东林, 等. 基于时间序列分析的频谱异常自主检测和稳健估计方法[J]. 电子学报, 2014, 42(6): 1055 – 1060.
- WANG Lei, XIE Shuguo, SU Donglin, et al. An autonomous detection and robust estimation method of spectrum anomaly based on time series analysis[J]. Acta Electronica Sinica, 2014, 42(6): 1055 – 1060. (in Chinese)
- [10] Yarkan S, Arslan H. Binary time series approach to spectrum prediction for cognitive radio[A]. IEEE 66th Vehicular Technology Conference [C]. Baltimore, MD, USA: IEEE, 2007. 1563 – 1567.
- [11] WANG Zhe, Salous S. Time series ARIMA model of spectrum occupancy for cognitive radio[A]. IET Seminar on Cognitive Radio and Software Defined Radios [C]. London; IET, 2008. 1 – 4.
- [12] 王磊, 苏东林, 谢树果, 等. 基于 EGARCH 过程的电磁频谱占用状态波动特性分析[J]. 电子与信息学报, 2012, 34(11): 2767 – 2773.
- WANG lei, SU Donglin, XIE Shuguo, et al. Electromagnetic spectrum occupancy state volatility analysis based on egarch process[J]. Journal of Electronics and Information Technology, 2012, 34(11): 2767 – 2773. (in Chinese)
- [13] HUANG N E, SHEN Z, LONG S, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London Series A, 1998, 454: 903 – 995.
- [14] ZHU Z H, SUN Y L, JI Yu. Short-term load forecasting based on empirical mode decomposition and least square support vector machine[J]. Relay, 2007, 35(8): 38 – 40.
- [15] 刘岱, 庞松岭, 骆伟. 基于 EEMD 与动态神经网络的短期负荷预测[J]. 东北电力大学学报, 2009, 29(6): 20 – 26.
- LIU Dai, PANG Songling, LUO Wei. Power system short-term load forecasting based on EEMD and dynamic neural network[J]. Journal of Northeast Dianli University (Natural Science Edition), 2009, 29(6): 20 – 26. (in Chinese)

作者简介



魏鸿浩(通信作者) 男, 1991年2月生于辽宁省朝阳市, 现为北京航空航天大学电子信息工程学院硕士研究生, 主要研究方向为电磁频谱管理、电磁兼容、电磁频谱异常识别。

E-mail: whhbuaa@foxmail.com

贾云峰 男, 1975年4月生于湖北省荆门市, 2002年在国防科技大学获博士学位, 2003–2005年于中科院电子所微波器件中心进行博士后研究, 现为北京航空航天大学电子信息工程学院硕士生导师, 北航电磁兼容技术研究所设计部部长, 主要从事微波电磁场与电磁兼容预设计、评估和验证等研究工作。