

采用正交小波网络的非线性系统辨识方法*

王海清 宋执环 李 平

(浙江大学工业控制技术国家重点实验室, 工业控制技术研究所以杭州, 310027)

摘要: 基于多分辨分析的递阶逼近思想, 采用正交小波网络研究了输入样本空间分布非均匀时非线性系统的辨识问题. 重点讨论了样本非均匀时网络系的设计问题, 并给出了基于该网络系的在线递阶辨识算法. 最后利用正交小波网络分别对非线性静态和动态系统进行了仿真辨识.

关键词: 正交小波网络; 非均匀辨识样本; 多分辨递阶算法

文献标识码: A

Nonlinear System Identification Using Orthogonal Wavelet Networks

WANG Haiqing, SONG Zhihuan and LI Ping

(National Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Zhejiang University · Hangzhou, 310027, P. R. China)

Abstract: Based on the idea of multiresolution hierarchical approximation of wavelet, the identification of nonlinear system with sampled data dispersed irregularly in space is discussed by using orthogonal wavelet networks (OWN). The design criteria of the lattice structure are intensively discussed and justified, and an on-line hierarchical identification algorithm is given. The OWN is used to identify nonlinear static and dynamic systems. Simulation results indicate that the proposed algorithm has superior performance than conventional methods.

Key words: orthogonal wavelet networks; space-irregular sampled data; multiresolution hierarchical algorithm

1 引言(Introduction)

采用网络结构(network-based)的辨识方法是研究非线性系统建模的有力工具之一^[1], 神经网络、模糊自适应网络和近年来出现的小波网络(wavelet networks)得到了广泛的研究和重视^[2-4]. 其中源于小波分析理论的小波网络由于其独特的数学背景, 使得它的分析和设计均有许多不同于其他网络的方面^[4-6]. 而且小波网络的形式和设计方法也是多种多样的: 如文[7]是利用小波函数(或尺度函数)替换普通神经网络中的激励函数; 文[4]则是从多分辨分析的角度利用正交小波基构造网络; 文[5, 6]则重点讨论了高维小波网络的设计问题. 其中以紧支正交小波和尺度函数构造的正交小波网络具有系统化的设计方法, 能够根据辨识样本的分布和逼近误差要求确定网络结构和参数; 此外如正交小波网络还能够明确给出逼近误差估计, 网络参数获取不存在局部最小问题等优点.

本文针对输入样本空间非均匀分布(注意不是指时间上的分布)时的非线性系统建模问题, 利用正

交小波网络进行了研究, 着重讨论了其中网络系(lattice structure)设计和参数辨识的有关算法.

2 多分辨分析与递阶逼近算法(The multi-scale analysis and hierarchical approximation algorithm)

小波可看作是由母小波经过伸缩和平移所产生的一个函数序列. 适当选择有限长的尺度函数 $\phi(x)$ 和小波函数 $\Psi(x)$, 并定义: $\phi_{j,k}(x) = 2^{j/2}\phi(2^jx - k)$, $\Psi_{j,k}(x) = 2^{j/2}\Psi(2^jx - k)$, $j = j_0, j_0 + 1, \dots, j_m$, $k \in \mathbb{Z}$, 一般取最粗尺度 $j_0 = 0$, 最细尺度 j_m 则与采样频率有关. 适当选择 $\phi_{j,k}(x)$ 和 $\Psi_{j,k}(x)$ 可构成 $L^2(\mathbb{R})$ 的一组正交小波基, 而多分辨分析(MRA)则是提供了信号分析和综合的一种分层理论框架, 记 $\phi(x)$ 生成的 $L^2(\mathbb{R})$ 中一个多分辨分析为 $\{V_j\}$, W_j 是关于 V_j 的正交补空间. 于是信号 $f(x) \in V_m$ 的MRA分解表达式为:

$$f(x) = f_a^{(m-1)} + f_d^{(m-1)} = f_a^{(m-2)} + \sum_{j=m-2}^{m-1} f_d^{(j)} = \dots = f_a^{(0)} + \sum_{j=0}^{m-1} f_d^{(j)}, \quad (1)$$

* 基金项目: 国家自然科学基金(青年)基金(69704006)资助项目.

收稿日期: 1999-07-12; 收修改稿日期: 1999-12-27.

其中, $f_d^{(j)} = \sum_k c_{j,k} \phi_{j,k}(x)$, $f_d^{(j)} = \sum_k d_{j,k} \Psi_{j,k}(x)$ 为 $f(x)$ 在子空间 V_j 和 W_j 上的投影, 分别称为轮廓和细节, 系数 $c_{j,k} = \langle f, \phi_{j,k} \rangle$ 和 $d_{j,k} = \langle f, \Psi_{j,k} \rangle$. 结合 $\phi_{j,k}(x)$ 和 $\Psi_{j,k}(x)$ 的定义, 显然式(1)在结构上类似于具有一个隐层的神经网络, 隐层的激励函数由各尺度下的 $\phi_{j,k}(x)$ 和 $\Psi_{j,k}(x)$ 组成, 如图1所示. 从图中可见网络共有两种节点: 尺度节点 $\phi_{j,k}(x)$ 和小波节点 $\Psi_{j,k}(x)$.

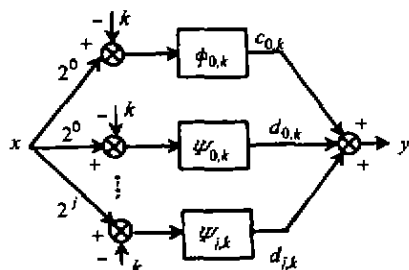


图1 正交小波网络结构图

Fig. 1 The diagram of orthogonal wavelet networks

若 $\phi_{j,k}(x)$ 和 $\Psi_{j,k}(x)$ 均为紧支正交的, 则有

$$\begin{cases} \langle \phi_{0,k}, \phi_{0,k'} \rangle = \delta_{kk'}, \\ \langle \Psi_{j,k}, \Psi_{j',k'} \rangle = \delta_{jj'} \cdot \delta_{kk'}, \end{cases} \quad \text{且 } \langle \phi, \Psi \rangle = 0. \quad (2)$$

当输入样本 x 是空间分布均匀时, 可由 x 的采样间距和数据长度定下网格系及最大分解层次 $j_m^{[1]}$. 然后用最小二乘法 (LS) 先求出最粗尺度时 $\phi_{0,k}$ 的系数 $\{c_{0,k}\}$, 之后再根据逼近误差逐层用 LS 求出 $\Psi_{j,k}$ 的系数 $\{d_{j,k}\}$, 直到满足逼近精度要求为止. 且由正交条件式(2)知系数 $\{c_{0,k}\}$ 和 $\{d_{j,k}\}$ 的确定是互不影响的, 整个求解过程呈递阶的形式进行下去. 上述即为正交小波网络的递阶逼近算法基本思想^[4,8]. 采用递阶逼近算法使得正交小波网络在逼近函数的过程中不存在局部最小的问题, 而且可以建立网络结构与逼近精度之间的明确关系, 这与神经网络和模糊网络相比, 显然是一个很大的优点.

3 非均匀输入样本下正交小波网络的分析与设计 (The analysis and design of orthogonal wavelet networks with irregular input samples)

由于正交小波函数难以获得封闭的表达式, 因此要求输入数据在空间分布均匀, 所有计算均在各层的网格点上进行, 故上述正交小波网络的递阶逼近算法在样本空间分布均匀时是很有效的, 可是一般情况下样本 x 都是空间分布非均匀的, 这样就存

在两个困难: 首先不均匀点处的小波函数 (或尺度函数) 的值需借助细分迭代方程 (Subdivision) 近似求出, 这就需要对近似化带来的误差进行分析, 以保证将逼近精度控制在一定的范围内. 而且此时在各层用 LS 求解系数时要对矩阵求逆, 计算量增大; 其次是小波网络逼近算法一般采用的都是离线辨识^[4,5,7], 而离线的单输入小波网络实际上只能实现对静态系统的辨识. 对动态系统模型同一输入在不同的时刻会对应不同的输出, 文[4]对此采取的方法是用几个不同的输出值取平均作为逼近值, 这样做显然是很不准确的, 必会有较大的误差. 下面就分别对这两方面的问题进行分析.

3.1 MRA 网格系的确定 (The design of lattice structure for MRA)

要解决样本空间分布非均匀对网络设计带来的影响, 关键在于给出一个合适的网格系, 以作为整个计算的框架. 这就需要结合待辨识对象的信息来考虑. 不失一般性, 作如下假设:

假设 1 设输入样本空间 $\{x_i\}_{i=1}^N$ 和输出 $\{y_i\}_{i=1}^N$ 分别属于 R 上的有界闭集 $U = [u_{\min}, u_{\max}]$ 和 $Y = [y_{\min}, y_{\max}]$, 其中映射 $f: \{x_i\} \rightarrow \{y_i\}$ 为一非线性映射.

定理 1 对于满足假设 1 条件的非线性系统 $y_i = f(x_i)$, 其 MRA 分解系数具有以下性质:

$$1) |c_{j,k}| \leq 2^{-j/2} \cdot y_{\max} \cdot |\phi|_k = |c_j|_{\max} \text{ 和 } |d_{j,k}| \leq 2^{-j/2} \cdot y_{\max} \cdot |\Psi|_k = |d_j|_{\max};$$

$$2) \lim_{j \rightarrow +\infty} |d_{j,k}| = 0, \text{ 其中, } |\phi|_k = \int_U |\phi(x-k)| dx \text{ 和 } |\Psi|_k = \int_U |\Psi(x-k)| dx.$$

证 由系数 $c_{j,k}$ 定义

$$c_{j,k} = \langle f(x), \phi_{j,k} \rangle = 2^{j/2} \int_U f(x) \phi(2^j x - k) dx,$$

于是有

$$|c_{j,k}| \leq 2^{j/2} \cdot |f|_{\max} \cdot \int_U |\phi(2^j x - k)| dx = 2^{-j/2} \cdot y_{\max} \cdot \int_U |\phi(x-k)| dx, \quad (3)$$

由于 ϕ 是紧支正交的尺度函数, 故 $\phi(x) \in L^1(R)$. 不妨记: $|\phi|_k = \int_U |\phi(x-k)| dx$ 为一有界常数 (可近似认为与 k 无关), 从而可得尺度 j 时系数 $|c_{j,k}|$ 的上界: $|c_j|_{\max} = 2^{-j/2} \cdot y_{\max} \cdot |\phi|_k$. 同理可证系数 $|d_{j,k}|$ 的上界: $|d_j|_{\max} = 2^{-j/2} \cdot y_{\max} \cdot |\Psi|_k$, 其中 $|\Psi|_k = \int_U |\Psi(x-k)| dx$ 为

一有界常数,性质1)得证.

由性质1)易见尺度 $j \rightarrow +\infty$ 时,系数 $|d_{j,k}| \rightarrow 0$, 即性质2)成立.

定理2 对于满足假设1条件的非线性系统 $y_i = f(x_i)$, 若欲使非均匀样本近似至网格点引起的网络计算误差为小于 e_Δ , 则最粗尺度 ($j=0$) 下的网格间距 Δ 满足下式:

$$\Delta \leq \frac{2e_\Delta}{l_i \cdot |c_0|_{\max} \cdot |\phi'|_{\max}}, \quad (4)$$

其中 l_i 为支集包含样本点 x_i 的尺度函数的项数; $|c_0|_{\max}$ 定义见定理1的性质1); $|\phi'|_{\max}$ 为尺度函数在其支集内导数绝对值的上界.

证 记第 i 时刻的输入样本为 $x_i \in U$, 在 $j=0$ 层距点 x_i 最近的网格点为 x_i , 即 $|x_i - x_i| \leq \Delta/2$. 且有 l_i 个尺度函数项满足: $\{\phi_{0,k}(x_i)\}_{k=1}^{l_i} \neq 0$. 于是将 x_i 近似为网格点 x_i 带来的计算误差为

$$\begin{aligned} & |f_a^{(0)}(x_i) - f_a^{(0)}(x_i)| = \\ & |c_{0,1} \cdot [\phi_{0,1}(x_i) - \phi_{0,1}(x_i)] + \cdots + \\ & c_{0,l_i} \cdot [\phi_{0,l_i}(x_i) - \phi_{0,l_i}(x_i)]| \leq \\ & [|c_{0,1}| + \cdots + |c_{0,l_i}|] \cdot \frac{\Delta}{2} \cdot |\phi'|_{\max} \leq e_\Delta. \end{aligned} \quad (5)$$

上式中: $|\phi'|_{\max} = \max\{|\phi'(\xi_1)|, \cdots, |\phi'(\xi_{l_i})|\}$, $\{\xi_1, \cdots, \xi_{l_i}\}$ 为各相应区间中的某点. 又由 $|c_0|_{\max}$ 定义有: $|c_{0,1}| + \cdots + |c_{0,l_i}| \leq l_i \cdot |c_0|_{\max}$, 将其代入式(5)即可知式(4)成立.

上述证明需保证 $\phi'(x)$ 存在且有界, 以 Daubechies 小波为例, 由于 ${}_4\phi(x), {}_4\Psi(x) \in C^{1.275}$, 故实际计算时取 4 阶以上 db 小波即可.

注 1) 对于给定的紧支正交小波 (如 db 小波), 可先由定理1 求出最粗尺度下系数的上界 $|c_0|_{\max}$, 然后再由定理2 求出满足误差要求的 $j=0$ 层的网格分布;

2) 定理1 说明随着尺度的变细, MRA 分解系数趋于 0. 故可采用类似于小波收缩算法 (wavelet shrinkage algorithm)^[1,9] 的思想, 设置系数的阈值 (threshold), 由定理1 的性质1) 就可定出最细尺度 j_m , 至此网格系就已确定. 如图2所示为 $j_m=2$ 的一个网格系示意图. 注意到上述证明过程并不需要样本 $\{x_i, y_i\}$ 的具体信息, 所以尺度函数族 $\{\phi_{0,k}\}$ 和小波函数族 $\{\Psi_{j,k}\}_{j=0}^{j_m}$ 可事先求出存于计算机中;

3) 严格地说近似误差 e_Δ 还应包括在各层细节部分中的误差, 但由于网格间距以 $\Delta/2^{j+1}$ 减小, 且

由性质2) 知其影响是很小的. 另外在下述的在线递阶辨识算法中可以看出这种影响将进一步得到减小, 因此由式(4)获得的网格估计在实际计算中已能满足要求.

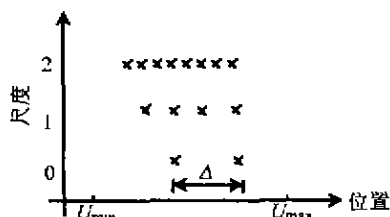


图2 正交小波网格系示意图

Fig. 2 The lattice structure of orthogonal wavelet networks

3.2 在线递阶辨识算法 (On-line hierarchical identification algorithm)

如前所述, 对于动态模型采用离线算法效果较差, 必须研究在线算法. 由于输入样本非均匀, 每一时刻支集内含有数据点的 $\{\phi_{0,k}\}$ 和 $\{\Psi_{j,k}\}_{j=0}^{j_m}$ 均在变化, 使得系数的辨识算法不能获得递推的形式, 但是注意到此时数据点近似后均位于网格系上, 而不是象文[4]中近似到由细分方程求出的中间格点上. 因此用递阶算法求解系数时已无需对矩阵求逆 (因为采用的是正交小波), 计算量是很小的. 为便于以后的阐述, 作如下定义:

定义1 对于紧支正交小波 Ψ 和尺度函数 ϕ , 定义其支集分别为:

$$\begin{aligned} S_{j,k} &= \{x_i \in U: \Psi_{j,k}(x_i) \neq 0, 0 \leq j \leq j_m\}, \\ S'_{0,k} &= \{x_i \in U: \phi_{0,k}(x_i) \neq 0\}. \end{aligned} \quad (6)$$

对 $\forall x_i \in U$, 定义 L_i 为支集包含 x_i 的尺度函数和小波函数的指标集

$$L_i = \{(j,k) \cup (0,k): x_i \in S_{j,k} \text{ 或 } x_i \in S'_{0,k}\}, \quad (7)$$

其中下标 L_i 表示指标集中的项数为 L_i 项.

采用如下的在线递阶算法实现对非线性系统的辨识:

步骤1 按 3.1 节建立网格系, 采用固定长度为 N 的辨识数据进行在线辨识. 但 N 不能太小, 以减弱边界效应的影响, 但太大会降低跟踪的速度, 其具体值需结合小波滤波器的长度选取;

步骤2 设置系数 $|c_{0,k}|$ 和 $|d_{j,k}|$ 的初值. 根据定理1 确定的各层次系数的上界, 采用随机算法设置初值;

步骤3 采用递阶的方式逐层确定当前时刻各基函数的系数. 由于数据非均匀性, 一般均有: $N >$

L_i , 故可充分利用此冗余性, 在逐层求解系数时选取与该层网络最相匹配的数据点, 以进一步减小近似带来的误差. 将求出的系数保存, 用于小波网络计算下一时刻的输出.

可见不同时刻小波网络的系数和节点数均是在变化的, 因此图 1 所示的是一种“动态”的网络, 这是与神经网络^[2]和模糊网络^[3]相比的又一显著不同之处. 而且正交小波网络还可以给出每一时刻的逼近误差的估计公式, 这与离线算法是类似的^[8], 故不再作详细讨论.

4 仿真算例 (Simulation examples)

下面是利用正交小波网络作为在线辨识器的两个例子. 仿真中采用 db5 小波, N 取 25.

1) 非线性静态系统辨识.

为便于比较, 采用文[2, 3]中同样的模型. 辨识对象由如下的差分方程描述:

$$y(k+1) = 0.3y(k) + 0.6y(k-1) + g[u(k)], \quad (8)$$

式中未知函数的形式为

$$g(u) = 0.6\sin(\pi u) + 0.3\sin(3\pi u) + 0.1\sin(5\pi u).$$

为辨识这一系统, 采用如下的串-并行模型^[2]:

$$\hat{y}(k+1) = 0.3\hat{y}(k) + 0.6\hat{y}(k-1) + \hat{f}[u(k)], \quad (9)$$

式中 $\hat{f}(\cdot)$ 取为正交小波网络辨识器. 图 3 给出了 $k=700$ 步的辨识结果, 小波网络辨识输出为虚线所示. 正交小波网络很好地逼近了该非线性系统的输出. 同样的辨识对象, 文[2]中的神经网络辨识器具有两个隐层, 共 200 个可调参数. 文[3]中的自适应模糊网络辨识器则采用了 40 条模糊规则, 共 120 个可调参数. 在本例的辨识过程中只在第 $k=184$ 步时 L_i 达最大值 21, 即有 21 个可调参数, 远比上述两种网络的参数少. 在逼近过程中正交小波网络根据设置的精度要求, 逐层采用递阶算法逼近模型, 确保了辨识的效果.

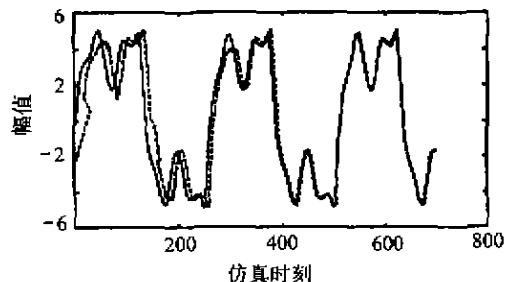


图 3 $k=700$ 的非线性静态系统辨识
Fig. 3 Identification of a nonlinear static system

2) 非线性动态系统辨识.

考虑如下的非线性动态系统^[10]

$$\begin{aligned} y(k) = & 0.5y(k-1) + u(k-2) + 0.1u^2(k-1) - \\ & 0.01y(k-1)y(k-2) + \\ & 0.4y(k-2)u(k-2) + 0.5e(k-1) + \\ & 0.2u(k-1)e(k-2) + e(k). \end{aligned} \quad (10)$$

文[10]中采用小波网络辨识该非线性动态系统, 但没有考虑网络系的设计, 对非均匀的样本点近似到最近的网格点上. 其最大逼近误差达 0.3. 采用本文提出的算法, 由设定的逼近精度 (本例取为 0.05) 定下网格系. 在线逼近时, 网络在各时刻下的迭代次数 (也即网络结构) 与该精度有关. 取控制输入为^[10]:

$$u(k) = 0.6\sin(2\pi k/95) + 0.4\sin(2\pi k/175).$$

辨识结果如图 4 所示. 当 $k > 650$ 后, 动态小波网络很好地逼近了系统的输出. 网络的最大逼近误差 ($k > 650$ 后) 为 0.219. 由于正交小波网络各激励函数均是紧支撑的, 因此具有所谓局部学习^[4] (local learning) 的特点. 而采用在线辨识算法使得在辨识的每一时刻, 只有有限的尺度节点和小波节点实际参与到辨识算法中去, 整个网络随着辨识数据的空间分布而始终处于一种动态的变化过程中. 在某些时刻的逼近误差出现大于设定精度要求可以认为是忽略了小波节点上误差对网络系设计的影响所造成的.

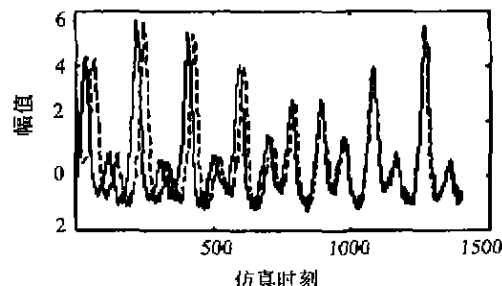


图 4 $k=1400$ 的非线性动态系统辨识
Fig. 4 Identification of a nonlinear dynamic system

5 结束语 (Conclusion)

采用紧支正交小波基的网络形式能够充分地继承和利用紧支正交小波的许多优良特性, 给网络的设计带来许多其他类型网络不具有的优点. 但是由于输入样本的非均匀性和缺乏有效的在线参数获取算法而限制了正交小波网络的应用. 本文对这些问题进行了详细的讨论, 并给出了相应的网络系确定方法以及在线辨识算法. 需要指出的是本文主要集中于单输入小波网络的讨论, 近年来已有文献^[5,6]对高维输入下的小波网络设计问题进行了研究, 一

般采用的均为径向基函数小波,但仍还存在许多问题需要解决.如前所述正交小波网络具有许多优良特性,因此研究可行的高维正交小波网络是一个很有意义的问题.

参考文献(References)

- [1] Sjöberg J, Zhang Q H and Benveniste A, et al. Nonlinear black-box modeling in system identification: a unified overview [J]. *Automatica*, 1995, 31(12): 1691 - 1724
- [2] Narendra K S and Parthasarathy K. Identification and control of dynamical systems using neural networks [J]. *IEEE Trans. Neural Networks*, 1990, 1(1): 4 - 27
- [3] Wang Lixin. *Adaptive Fuzzy Systems and Control: Design and Stability Analysis* [M]. New Jersey: Prentice-Hall, 1994
- [4] Bahavik R B and Stephanopoulos G. Wave-net: a multiresolution, hierarchical neural network with localized learning [J]. *AIChE Journal*, 1993, 39(1): 57 - 81
- [5] Zhang Q. Using wavelet network in nonparametric estimation [J]. *IEEE Trans. Neural Networks*, 1997, 8(2): 227 - 236
- [6] Sanner R M and Slotine J J B. Structurally dynamic wavelet networks for adaptive control of robotic systems [J]. *Int. J. Control*, 1998, 70(3): 405 - 421
- [7] Zhang Q. Wavelet networks [J]. *IEEE Trans. Neural Networks*, 1992, 3(6): 889 - 898
- [8] Wang Haiqing, Song Zhihuan and Li Ping. Reconstruction of high sample-rate data based on wavelet multiresolution approximation [J]. *Information and Control*, 2000, 29(4): 315 - 319 (in Chinese)
- [9] Donoho D. Denoising by soft-thresholding [J]. *IEEE Trans. Inform. Theory*, 1995, 41(3): 613 - 627
- [10] Song Zhihuan. *Wavelet theory and its application to industrial process identification and control* [D]. Hangzhou: Zhejiang University, 1997 (in Chinese)

本文作者简介

王海清 1974 年生, 分别于 1996 年和 1998 年在哈尔滨工业大学获工学学士和硕士学位, 现在浙江大学工业控制技术研究所攻读博士学位, 主要研究方向为小波在工业过程辨识、监测和故障诊断中的应用.

宋执环 1962 年生, 1996 年在浙江大学获工学博士学位, 现为浙江大学工业控制技术研究所副教授, 主要研究兴趣为小波在辨识和控制理论中的应用, 预测控制理论及应用, 流程工业 CIMS.

李平 1954 年生, 1988 年浙江大学获博士学位, 现为浙江大学工业控制技术研究所所长, 博士生导师, 主要研究兴趣为工业过程模型化, 智能控制, 流程工业 CIMS.

(上接第 199 页)

本文作者简介

严洪森 1957 年生, 1982 年在哈尔滨船舶工程学院自动控制系获工学学士学位, 1989 年和 1992 年分别在哈尔滨工业大学电气工程系和控制工程系获工学硕士和博士学位, 1992 年至 1994 年在南京航空航天大学机械工程学科从事博士后研究工作, 现任东南大学自动化研究所教授, 博士生导师, 美国工业工程师学会 (IIE) 高级会员, 主要研究方向为 CIMS 及 FMS 建模, 生产计划, 调度, 控制, 仿真, 并行工程和敏捷制造等. 近年来, 在国内外学术核心期刊和学术会议上发表

表论文 58 篇.

张晓东 1974 年生, 1997 年在山东建筑工程学院机电系获工学学士学位, 2000 年在东南大学自动化研究所获工学硕士学位, 现为东南大学自动化研究所博士生, 主要研究方向为敏捷制造, 车间生产计划等.

朱立峰 1976 年生, 1998 年在四川大学电气工程系获工学学士学位, 现为东南大学自动化研究所硕士生, 主要研究方向为敏捷制造, 车间生产计划等.