

基于关键输入和多输入层高维小波网络¹⁾

李换琴¹ 万百五²

¹(西安交通大学理学院 西安 710049)

²(西安交通大学系统工程研究所 西安 710049)

(E-mail: hqlee@mail.xjtu.edu.cn)

摘要 提出一种基于关键输入和加工工序的多输入层高维小波神经网络结构, 该网络结构是在传统前馈神经网络的基础上, 将一部分输入节点根据实际情况移到神经网络的相关隐层, 关键输入节点不仅与随后一层隐节点相连, 而且与输出节点相连, 更真实地反映了大工业生产过程中变量之间复杂的函数关系. 将该种小波网络模型应用于连铸连轧生产线产品质量建模, 其效果较其他 4 种神经网络为优越.

关键词 关键输入, 小波神经网络, 多输入层, 加工工序, 高维, 质量模型

中图分类号 P321

High Dimension Wavelet Neural Network Based on Key Input and Multi-Input-Layer Structure

LI Huan-Qin¹ WAN Bai-Wu²

¹(Faculty of Science, Xi'an Jiaotong University, Xi'an 710049)

²(Systems Engineering Institute, Xi'an Jiaotong University, Xi'an 710049)

(E-mail: hqlee@mail.xjtu.edu.cn)

Abstract This paper proposes a new architecture of high-dimension input wavelet neural network based on key inputs and work procedures for modeling complex large-scale industrial systems. In the neural network, some input variables are connected directly to the second hidden layer or other later hidden layers according to their actions being early or late in the work procedure. The key input nodes are connected to not only all nodes in the subsequent layer but also the output node. The developed method is applied to build a product quality model for continuous casting furnace and hot rolling mill. Simulation results demonstrate that the proposed network is more efficient and provides a higher accuracy compared with other four neural networks methods.

Key words Key inputs, wavelet neural network, multi-input-layer, work procedure, high dimension, product quality model

1) 国家“863”计划(863-51-945-011)、国家自然科学基金(60274055)和西安交通大学在职博士基金资助
Supported by the National “863” Program(863-51-945-011), the National Natural Science Foundation of P.R.China(60274055) and the Doctoral Foundation of Xi'an Jiaotong University
收稿日期 2003-03-14 收修改稿日期 2003-07-16
Received March 14, 2003; in revised form July 16, 2003

1 引言

包含多道加工工序的生产过程是现代制造业的一个显著特征, 其中一个产品的形成要经过几道加工工序, 每道工序都有若干个参数对最终产品的质量产生影响. 在这些参数中, 有些起主要作用 (称为关键输入变量), 而且不同工序的参数在不同时期发生作用. 由于其输入变量很多, 变量分阶段输入, 常规产品质量的数学模型难以建立. 这一直是世界各企业面临的难题, 也是控制领域研究的热点问题^[1].

文 [2~5] 曾试验过多种网络模型及学习算法. 本文在小神经网络的基础上, 考虑到关键输入和多道加工工序, 提出一种多输入层高维小波神经网络结构, 将这一网络结构用于气瓶钢质量建模, 获得成功.

2 基于关键输入和加工工序的多输入层高维小波神经网络

2.1 网络模型

为描述方便起见, 本文讨论有 k 个关键输入变量、3 输入层高维小波神经网络的情形. k 个关键输入变量分别为 x_1, x_2, \dots, x_k 都在第一输入层. 结论不难推广到 4 输入层或其它多输入层、关键输入变量位于不同输入层的情形. 有 k 个关键输入的 3 输入层高维小波神经网络结构如图 1 所示. 它类似于多层前馈神经网络, 区别在于: 1) 所有输入变量不是在同一层输入, 而是根据变量起作用的时间顺序分 3 层输入; 2) k 个关键输入节点除与

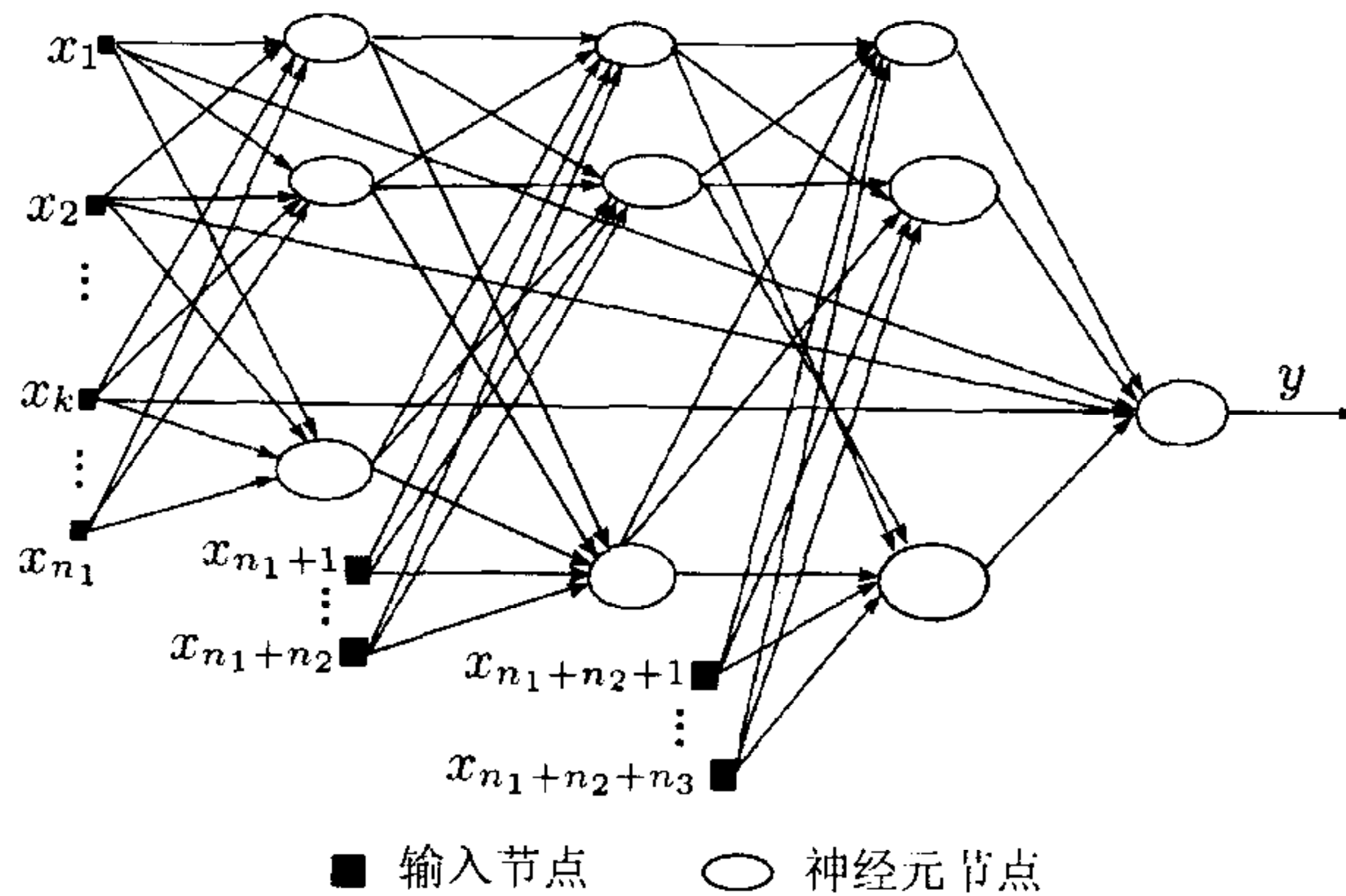


图 1 基于关键输入和三输入层高维小波神经网络结构图

Fig. 1 The architecture of three-input-layer wavelet neural network with key inputs

后一层隐节点连接外还与输出节点相连. 输入层 (第 0 层) 有 n_1 个输入节点, 第 1 层有 m_1 个隐节点和 n_2 个输入节点, 第 2 层有 m_2 个隐节点和 n_3 个输入节点, 第 3 层有 m_3 个隐节点, 第 4 层为线性输出层. 其中 $d_{ij}^{(q)}$ 表示第 $q-1$ 层输入节点 i 与后一层神经元 j 的连接权值 ($q = 1, 2, 3$), $w_{ij}^{(t)}$ 表示第 t 层神经元与第 $t+1$ 层神经元 j 的连接权值, $w_i^{(t)}$ 表示第 t 层神经元 i 的阈值, $s_i^{(t)}$ 表示第 t 层神经元 i 的输入, $y_i^{(t)}$ 表示第 t 层神经元 i 的输出 ($t = 1, 2, 3, 4$). 第 i 个关键输入节点与输出层的连接权为 u_i , 各神经元的激励函数 $f(x)$ 根据文 [4] 取为 4 阶 B 样条小波函数, 即

$$f(x) = \begin{cases} 0, & x < 0 \\ x^3/6, & 0 \leq x < 1 \\ (-3x^3 + 12x^2 - 12x + 4)/6, & 1 \leq x \leq 2 \\ f(4-x), & \text{其它} \end{cases}$$

则网络的输出为

$$\begin{aligned} y &= \sum_{i=0}^{m_3} w_{i1}^{(3)} y_i^{(3)} + \sum_{i=1}^k u_i x_i \quad (y_0^{(3)} = w_1^{(4)}, w_{01}^{(3)} = -1) \\ y_j^{(3)} &= f(s_j^{(3)}) \quad (j = 1, 2, \dots, m_3) \\ s_j^{(3)} &= \sum_{i=0}^{m_2} w_{ij}^{(2)} y_i^{(2)} + \sum_{i=n_1+n_2+1}^{n_1+n_2+n_3} d_{ij}^{(3)} x_i \quad (y_0^{(2)} = w_j^{(3)}, w_{0j}^{(2)} = -1) \\ y_j^{(2)} &= f(s_j^{(2)}) \quad (j = 1, 2, \dots, m_2) \\ s_j^{(2)} &= \sum_{i=0}^{m_1} w_{ij}^{(1)} y_i^{(1)} + \sum_{i=n_1+1}^{n_1+n_2} d_{ij}^{(2)} x_i \quad (y_0^{(1)} = w_j^{(2)}) \\ y_j^{(1)} &= f(s_j^{(1)}) \quad (j = 1, 2, \dots, m_1) \\ s_j^{(1)} &= \sum_{i=0}^{n_1} d_{ij}^{(1)} x_i \quad (x_0 = w_j^{(1)}, w_{0j}^{(1)} = -1) \end{aligned}$$

2.2 学习算法

记所有权阈值构成的向量为 θ , 则网络的输出可以表示为

$$y = y(\theta)$$

若给定 N 组训练样本

$$(x_1^{(k)}, \dots, x_{n_1}^{(k)}, x_{n_1+1}^{(k)}, \dots, x_{n_1+n_2+n_3}^{(k)}, y^{(k)}) \quad k = 1, 2, \dots, N$$

第 k 组样本对应的网络输出值为 $y_k(\theta)$, 则第 k 组训练样本的拟合误差为

$$e_k(\theta) = y^k - y_k(\theta)$$

定义总误差为

$$E(\theta) = \frac{1}{2} \sum_{k=1}^N (y^{(k)} - y_k(\theta))^2 = \frac{1}{2} \sum_{k=1}^N (e_k(\theta))^2$$

利用链导法则容易推出 $E(\theta)$ 对各个权阈值的偏导数. 本文采用误差反传学习算法. 反传学习算法中, 最常用的是 BP 算法, 标准的 BP 算法是最速下降法, 它对权值的修正以负梯度方向为搜索方向, 实现起来最简单, 但速度太慢, 常常导致网络训练时间过长, 训练精度下降甚至网络难以收敛. 牛顿法与拟牛顿法收敛较快, 但每次迭代都需要存贮 $n \times n$ 的矩阵, 因此很难应用于大规模问题. 共轭梯度法在算法的简便性、所需存贮量等方面均与最速下降法差别不大, 而收敛速度比最速下降法要快. 所以本文利用非线性共轭梯度法训练神经网络, 经过多次试验, 基于 GoldStein 非精确线搜索的 Dai-Yuan(DY) 非线性共轭梯度法^[6] 具有较快的收敛速度和拟合精度, 本文利用该方法训练神经网络.

3 实际应用

连铸连轧过程是典型的包含多道加工工序的大工业生产过程. 成品气瓶钢的质量主要与炼钢过程中的 23 个化学元素、加热炉中 2 个参数及轧制过程中 7 个参数的变化有关, 其中碳、锰、钛 3 个元素起关键作用. 在现场收集到的数据中, 经过数据预处理, 选出气瓶钢生产过程中记录的 960 组数据. 本文以这些数据作为样本点, 对其归一化处理后, 随机选取其中 600 组为建模样本, 其余 360 组为测试样本, 以断裂延伸率为质量指标建立气瓶钢的神经网络质量模型.

32 维输入变量按照加工工序先后分 3 层输入, 3 个关键输入变量是碳、锰、钛的含量. 在如图 1 所示的网络模型中, 取 $k = 3, n_1 = 23, n_2 = 2, n_3 = 7$. 隐层神经元的个数经多次试验选取为 $m_1 = 30, m_2 = 20, m_3 = 10$. 采用整体训练原则, 学习算法采用误差反传学习与 GoldStein 非精确线搜索的 Dai-Yuan(DY) 非线性共轭梯度法^[6]. 在奔腾 IV 处理器上运行 2 小时 50 分钟, 经过 1258 次迭代, 网络收敛. 网络模型的拟合均方误差为 7.35×10^{-4} , 拟合命中率为 94.0%, 测试均方误差为 0.0011, 测试命中率为 92.3%, 这里命中率指的是网络输出值与实际值相对误差小于 0.05 的样本占总样本的百分比.

针对上述同一问题, 文 [2] 建立的 $32 \times 40 \times 20 \times 1$ 双隐层前馈神经网络模型, 拟合命中率仅为 66%, 测试命中率为 59%; 文 [3] 采用 $32 \times 60 \times 1$ 三层 BP 神经网络结构, 拟合命中率为 75.1%, 测试命中率为 70.5%; 文 [4] 以 $32 \times 60 \times 1$ 高维输入小波网络为模型, 拟合命中率达 80.3%, 测试命中率为 78.5%; 文 [5] 将整个加工过程粗略分为炼钢和轧制两个阶段, 建立的双输入层神经网络模型拟合命中率为 89.3%, 测试命中率为 88.5%.

可见, 小波网络的建模精度略高于传统的 BP 网络的建模精度, 双输入层小波神经网络建模精度远远高于单输入层神经网络. 到目前为止, 基于关键输入和加工工序的 3 输入层小波神经网络建模精度达到最高.

4 结论

1) 对于包含多道加工工序, 影响最终产品质量的参数分阶段输入、且有部分参数比较重要的复杂大工业生产线产品质量建模问题, 选择基于关键输入和加工工序的多输入层高维小波神经网络模型更为合适.

2) 提出的产品质量模型可用于其它加工过程和生产线之类的问题, 如铜板热轧、造纸等生产过程.

3) 提出的网络结构适合下列复合函数类的逼近:

$$y = f(u, x_1, x_2, \dots, x_k), \quad u = g(v, x_{n_2+1}, \dots, x_{n_3}), \quad v = \varphi(t, x_{n_1+1}, \dots, x_{n_2}), \\ t = \psi(x_1, x_2, \dots, x_{n_1}). \quad \text{这里 } x_1, x_2, \dots, x_k \text{ 是关键输入变量, } k < n_3, n_1 < n_2 < n_3.$$

References

- 1 WAN Bai-Wu. Product quality model and quality control model for industries and their application. *Acta Automatic Sinica*, 2002, **28**(6): 1019 ~ 1024 (in Chinese)
- 2 Jia Lei. Neural network quality model with high dimension inputs for production-line products. [Master dissertation]. Xi'an: Xi'an Jiaotong University, 2002 (in Chinese)

- 3 Xing Jinsheng. An investigation on steel plate quality model of large-scale hot rolling mill using KDD technology[Ph D. dissertation]. Xi'an: Xi'an Jiaotong University, 2002 (in Chinese)
- 4 Li Huanqin, Wan Baiwu. A product quality model for large-scale hot rolling mill based on high-dimension input wavelet neural network. *Systems Engineering*, 2002, **20**(5): 55 ~ 58 (in Chinese)
- 5 Li Huanqin, Wan Baiwu. A high-dimension input wavelet neural network based on work procedure and its application. *Control and Decision*, 2003, **18**(S.I.2): 357 ~ 359 (in Chinese)
- 6 Dai Y H, Yuan Y. A nonlinear conjugate gradient with a strong global convergence property. *SIAM Journal of Optimization*, 1999, **10**(1): 177 ~ 182

李换琴 西安交通大学理学院副教授. 分别于 1986 和 1999 年获西安交通大学计算数学专业学士和硕士学位, 2003 年获西安交通大学控制科学与工程博士学位. 现主要从事神经网络、小波神经网络及大工业过程产品质量控制研究.

(**LI Huan-Qin** Received her Master's and Ph.D. degrees all from Xi'an Jiaotong University in 1999 and 2003, respectively. Now she is an associate professor in School of Science, Xi'an Jiaotong University. Her research interests include neural network, wavelet neural network and their application to product quality control.)

万百五 西安交通大学系统工程研究所教授, 博士生导师. 目前主要研究方向是大工业过程稳态递解优化理论与应用, 智能控制、产品质量控制等.

(**WAN Bai-Wu** Professor in Systems Engineering Institute, Xi'an Jiaotong University. Research interests include theory and application of large-scale industrial processes, steady-state hierarchical optimization and intelligent control of large scale systems, product quality control, etc.)