

# 一种新型神经网络结构模型的应用研究

周金荣 黄道 蒋慰孙

(华东理工大学自动化所 上海 200237)

## 摘 要

基于过程先验知识与神经网络结构选择相结合的思想,提出了一类新型神经网络结构,并具体研究了其学习算法.同时给出了一种新的样本抽取方法及其相应的网络训练步骤.这些思想和方法均在复杂尿素生产过程的建模和优化中得到了应用.基于所建神经网络模型知识的专家系统已在某化肥厂得到成功的应用.

**关键词:** 神经网络,先验知识,学习算法.

## 1 引言

近年来, Rumell hart 等人所提出的多层前向神经网络已在控制中得到了较广泛的应用.然而对于各种不同的复杂实际对象、不同的非线性特性,仅用同一种结构的网络加以逼近并不一定适宜.国外已有学者开始了这方面的研究工作<sup>[1-3]</sup>,但尚无工业应用成果.显然,根据不同对象所能获得的机理特征和变量间因果关系来指导神经网络结构的选择及针对工业过程的特点给出新的样本抽取方法,都是神经网络应用于工业过程中所需解决的难题.

基于对尿素、炼油等过程的机理分析结果,在多层前向神经网络的基础上提出了多层复杂前馈结构等具有工程意义的网络模型,并给出了相应的学习算法;同时针对工业过程的数据受各种干扰影响的特点,给出了一种新的样本抽取方法.在此基础上进行了尿素生产过程的建模与优化的应用研究,结果令人满意.

## 2 新型神经网络结构模型及其学习算法

### 2.1 工程背景

对于一个复杂对象,一般或多或少地对其机理和变量间因果关系等先验知识有所掌握.尿素生产是一个涉及气、液、固三相,多种平衡以及高温、高压状态下多个可逆反应的复杂工业过程,以该过程中合成塔各变量间的相互关系为例,其转化率  $\eta$  受系统氨碳比  $a$ 、水碳比  $b$ 、反应温度  $T$  及系统压力  $P$  的影响.而系统压力又受  $a, b, T$  及塔顶气相出

口阀门开度  $V$  的影响。依据机理分析, 系统压力  $P$  在特定条件下可作为优化目标系统转化率的间接反映变量, 各变量间关系如图 1 所示。

当然, 可按照传统前向神经网络建模, 但亦可构造图 2 所示的神经网络结构图。图中虚线所圈部分为代表  $P$  的神经元群, 而直接将输入元氨碳比  $a$ , 水碳比  $b$ ,  $T$ ,  $V$  与输出元  $\eta$  连接起来, 以表征前述分析所得的过程机理。后面的仿真结果表明, 对于同样数量的训练样本, 这种新的网络结构比传统结构能更加快速和精确地实现尿素合成塔

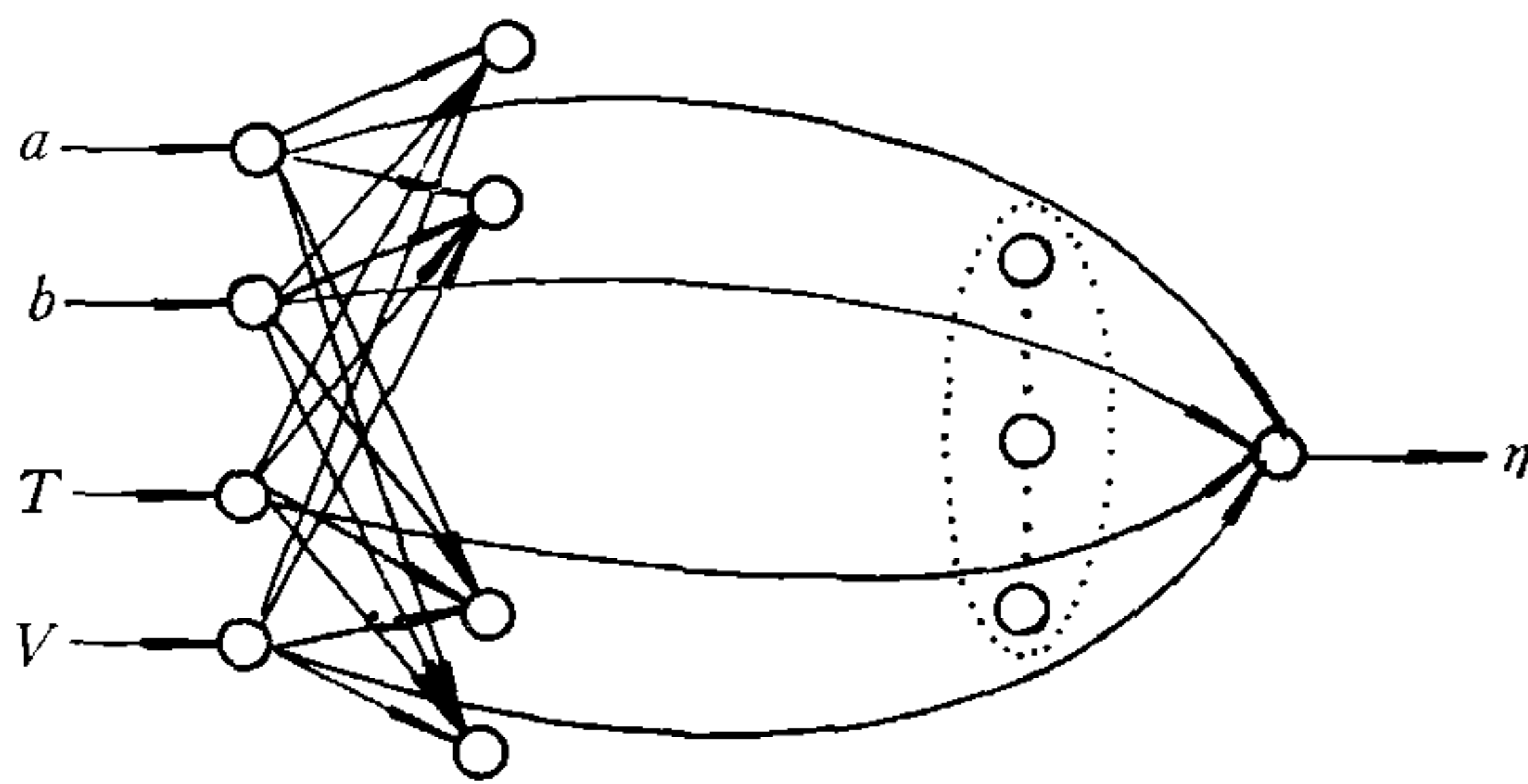


图 2 尿素合成塔新的神经网络模型结构图

## 2.2 一类新型神经网络结构模型

根据 2.1 中典型实例分析结果, 并推广至一般性原理分析, 这里提出一类具有不同连线特征的神经网络。

**定义 1.** 由输入层至输出层相邻两层之间神经元通过前向权值连线相互连接的方式, 称之为简单前馈连接。

**定义 2.** 由输入层至输出层各层的每一神经元通过前向权值连线与后续所有神经元相互连接的方式, 称为复杂前馈连接。

**定义 3.** 由输出层至输入层相邻两层之间后层神经元的输出通过权值连线作为前层各神经元的输入的连接方式, 称之为简单反馈连接。

**定义 4.** 由输出层至输入层后层神经元的输出通过权值连线作为前面所有神经元的输入的连接方式, 为复杂反馈连接。

对于不同的具体应用对象, 可适当地选择每个神经元的连接方式, 以使所构成的网络能最佳地描述该具体对象的机理特征。

**定义 5.** 由若干层神经元以复杂前馈连接方式构成的神经网络, 称为多层复杂前馈型网络, 其结构图如图 3 所示。

类似地, 还可给出具有其它复杂连接特征的复杂网络结构的定

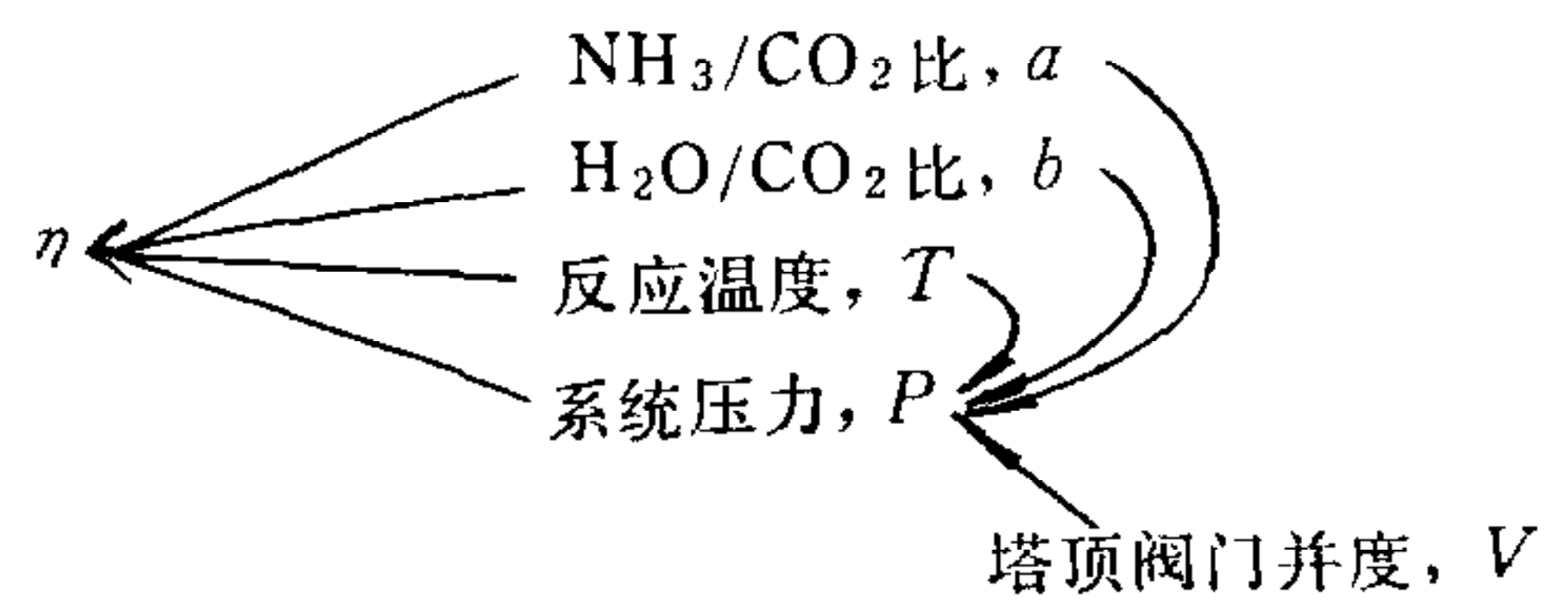


图 1 尿素合成塔各变量因果关系图

这种新的网络结构比传统结构能更加快速和精确地实现尿素合成塔的模式化。

同样, 对作者曾研究过的炼油厂润滑油系统的生产调度问题进行了依据过程机理加以选择神经元间的连接的研究, 得到了相同的结论。显然, 依据机理分析结果构造各种不同结构的神经网络将使神经网络理论更加实用化。

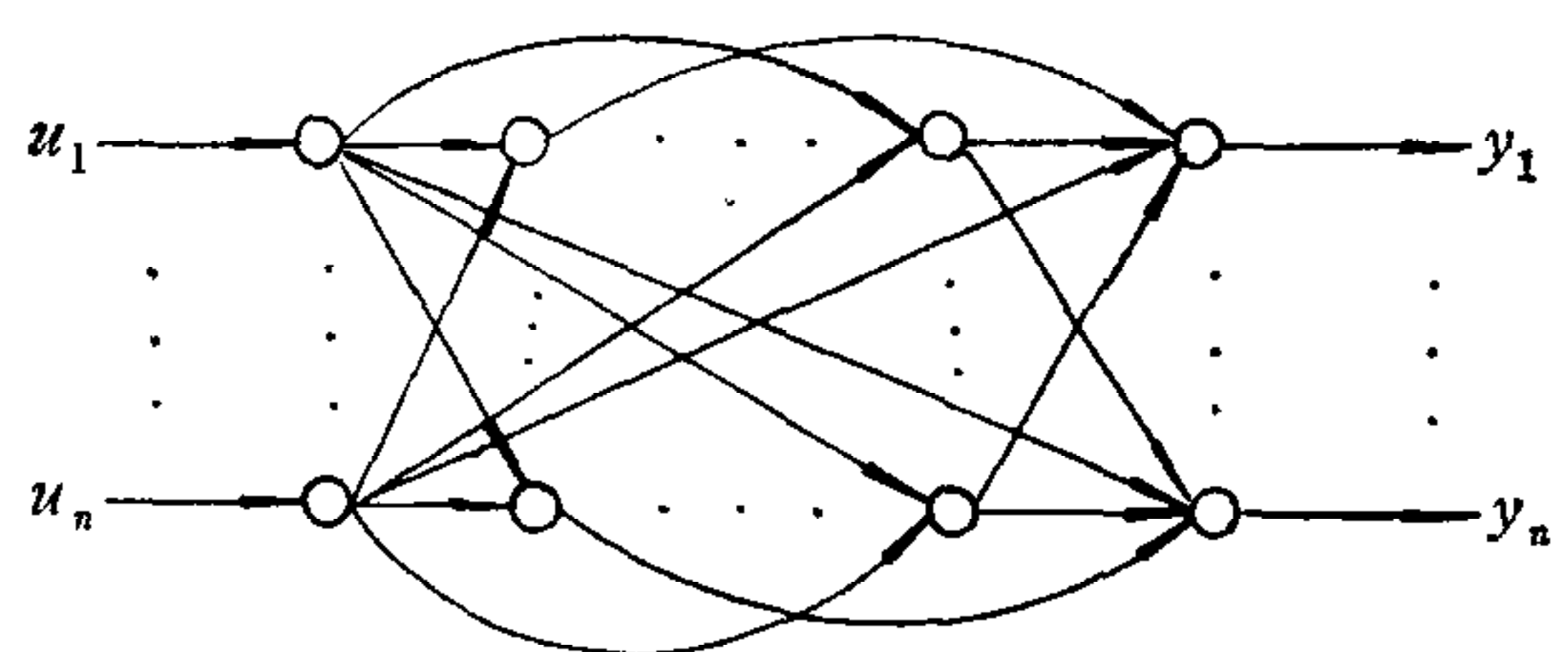


图 3 多层复杂前馈型网络结构示意图



义。对某一具体的研究对象而言,这类网络结构的简化形式通常较为简单。

### 2.3 多层复杂前馈型网络的学习算法

设  $W_{jil}$  表示第 1 层的第  $j$  神经元到第  $k$  层的第  $i$  神经元的权值;  $N_{ik}$ ,  $O_{ik}$ ,  $t_{ik}$  分别为第  $k$  层第  $i$  个神经元的输入、输出和阈值。

每一神经元为一非线性环节,取为 Sigmoid 函数。学习目标选为

$$\min |E = \frac{1}{2} \sum_R \sum_{i=1}^n (y_i^d - y_i)^2. \quad (2-1)$$

再设  $U_1, U_2, \dots, U_m$  为神经网络的输入,且各自到第一层神经元的连接权值为

$$W_{ji0} = \begin{cases} 1, & \text{当 } j = i \text{ 时,} \\ 0, & \text{当 } j \neq i \text{ 时,} \end{cases}$$

则

$$\begin{aligned} N_{ik} &= \sum_{i_{k-1}} W_{i_{k-1}, i, k-1, k} O_{i_{k-1}, k-1} + \sum_{i_{k-2}} W_{i_{k-2}, i, k-2, k} O_{i_{k-2}, k-2} + \dots \\ &\quad + \sum_{j_0} W_{j_0, i, 0, k} O_{j_0, 0} + t_{ik} \\ &\quad (k = 1, 2, \dots, L; L \text{ 为神经网络的层数}). \end{aligned} \quad (2-2)$$

**定义 6.** 神经元的阈值、非线性系数及与之相连的各连接权值统称为神经元的相关参数。

**定义 7.** 第  $k$  层中的所有神经元的相关参数的全部统称为第  $k$  层的网络参数值。

**定理 1.** 以式 (2-1) 为目标的具有  $L$  层的多层复杂前馈型网络中第  $k$  层的第  $i$  个神经元的相关的修正量,只与自身的相关参数值、网络的输出偏差量以及从第  $k+1$  层至第  $L$  层的网络参数值、参数修正量有关,而与从第 1 层至第  $k-1$  层的网络参数值、参数修正量无关。

**证明.** 由“误差反传”思想,可得各参数的具体修正公式:

$$\begin{cases} \frac{\partial E}{\partial W_{jil}} = \frac{\partial E}{\partial N_{ik}} \cdot \frac{\partial N_{ik}}{\partial W_{jil}} = \frac{\partial E}{\partial N_{ik}} \cdot O_{jl}, \\ \frac{\partial E}{\partial t_{ik}} = \frac{\partial E}{\partial N_{ik}} \cdot \frac{\partial N_{ik}}{\partial t_{ik}} = \frac{\partial E}{\partial N_{ik}}, \\ \frac{\partial E}{\partial \theta_{ik}} = \frac{\partial E}{\partial O_{ik}} \cdot \frac{\partial O_{ik}}{\partial \theta_{ik}} = \frac{\partial E}{\partial O_{ik}} \cdot \frac{O_{ik}}{\theta_{ik}}, \\ \frac{\partial E}{\partial \beta_{ik}} = \frac{\partial E}{\partial O_{ik}} \cdot \frac{\partial O_{ik}}{\partial \beta_{ik}} = \frac{\partial E}{\partial O_{ik}} \cdot \left[ O_{ik} \left( 1 - \frac{O_{ik}}{\theta_{ik}} \right) N_{ik} \right]. \end{cases} \quad (2-3)$$

其中  $l = 0, 1, \dots, k-1; k = 1, 2, \dots, L$ .

当  $k = L$  时,

$$\frac{\partial E}{\partial N_{iL}} = \frac{\partial E}{\partial y_i} \cdot \frac{\partial y_i}{\partial O_{iL}} \cdot \frac{\partial O_{iL}}{\partial N_{iL}}. \quad (2-4)$$

由于

$$E = \frac{1}{2} \sum_R \sum_{i=1}^n (y_i^d - O_{iL})^2, \quad O_{iL} = \frac{\theta_{iL}}{1 + e^{-\beta_{iL} N_{iL}}}, \quad (2-5)$$

因此

$$\begin{cases} \frac{\partial E}{\partial N_{iL}} = -(y_i^d - O_{iL}) \cdot O_{iL} \cdot \left(1 - \frac{O_{iL}}{\theta_{iL}}\right) \beta_{iL}, \\ \frac{\partial E}{\partial O_{iL}} = -(y_i^d - O_{iL}). \end{cases} \quad (2-6)$$

显然,将式(2-6)代入式(2-3)中可得  $L$  层各神经元参数修正量,这些修正量只与网络输出偏差  $(y_i^d - O_{iL})$  和第  $L$  层的网络参数值有关,而与前面各层的网络参数值无关。因此,对  $L$  层定理成立。

当  $k = L - 1$  时,

$$\frac{\partial E}{\partial O_{i,L-1}} = \sum_{j_L=1}^n \frac{\partial E}{\partial N_{j_L L}} \cdot \frac{\partial N_{j_L L}}{\partial O_{i,L-1}} = \sum_{j_L=1}^n \frac{\partial E}{\partial N_{j_L L}} \cdot W_{i,j_L,L-1,L}. \quad (2-7)$$

当  $k = 1, 2, \dots, L - 2$  时,

$$\frac{\partial E}{\partial O_{ik}} = \sum_{m=k+1}^L \sum_{j_m} \frac{\partial E}{\partial N_{j_m m}} \cdot \frac{\partial N_{j_m m}}{\partial O_{ik}} \quad (m = k + 1, \dots, L), \quad (2-8)$$

可得

$$\frac{\partial E}{\partial O_{ik}} = \sum_{m=k+1}^L \sum_{j_m} \frac{\partial E}{\partial N_{j_m m}} \cdot W_{ij_m k m}. \quad (2-9)$$

显然,当  $k = 1, 2, \dots, L - 1$  时,有

$$\frac{\partial E}{\partial O_{ik}} = \sum_{m=k+1}^L \sum_{j_m} \frac{\partial E}{\partial N_{j_m m}} \cdot W_{ij_m k m}. \quad (2-10)$$

其中  $\frac{\partial E}{\partial N_{j_m m}}$  可由  $\frac{\partial E}{\partial N_{j_L L}}$  的值逐层反传递推求得。

此外

$$\frac{\partial E}{\partial N_{ik}} = \frac{\partial E}{\partial O_{ik}} \cdot \frac{\partial O_{ik}}{\partial N_{ik}} = \frac{\partial E}{\partial O_{ik}} \cdot O_{ik} \left(1 - \frac{O_{ik}}{\theta_{ik}}\right) \beta_{ik}. \quad (2-11)$$

将式(2-10)和式(2-9)代入式(2-3),可得第  $k$  层各神经元的参数修正量,它们只与本身的相关参数值以及第  $k + 1$  层至第  $L$  层的网络参数值、参数修正量有关。

因此,对于  $k = 1, 2, \dots, L$  层的神经元,定理成立。式(2-1)~(2-11)就构成了多层复杂前馈网络的学习算法。类似地,可推导得到其它带反馈的复杂网络的学习算法,其结论与定理 1 相同。

上述学习算法与传统 BP 算法的区别,在于网络参数修正采用了多层误差信息同时向前反向传播。它们的递推性能以及针对具体问题的网络结构的简化处理,使得学习过程变得简单。

### 3 新的样本抽取方法及其网络学习方法

针对实际工业过程输入输出变量一般总是受一些不可测、不可控扰动影响的特点<sup>[4]</sup>。这里,考虑样本数据库由两部分样本组成,一种是包含过程动态信息的以周期  $T$  为间隔的



采样样本,即第 I 类样本:

$$(\tilde{u}_i(T), \tilde{y}_j(T)), (\tilde{u}_i(2T), \tilde{y}_j(2T)), \dots, (\tilde{u}_i(KT), \tilde{y}_j(KT))); \quad (3-1)$$

另一种则是表征随机系统的稳态信息的一段时间区域上的均值样本,即第 II 类样本:

$$(E_{k+1}\{\tilde{u}_i(t)\}, E_{k+1}\{\tilde{y}_j(t)\}), (E_{k+2}\{\tilde{u}_i(t)\}, E_{k+2}\{\tilde{y}_j(t)\}), \dots, (E_R\{\tilde{u}_i(t)\}, E_R\{\tilde{y}_j(t)\}). \quad (3-2)$$

这里  $i = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ ;  $T$ —采样周期;  $K$ —第 I 类样本的总数目;  $R$ —整个样本的数目.

显然,采样样本的学习实际上是对较大范围内工况的映射,而均值样本的学习则侧重于稳态工作点附近的信息记忆,这使神经网络模型的适应性得到了提高.

## 4 尿素生产过程神经网络建模和优化研究

利用尿素生产过程机理分析结合本文的研究结果,分别建立了尿素合成塔、气提塔等的神经网络模型.以合成塔为例,其优化目标是选择一组合适的操作条件,使二氧化碳转化率最高.为此构造了图 4 所示多层复杂前馈网络.

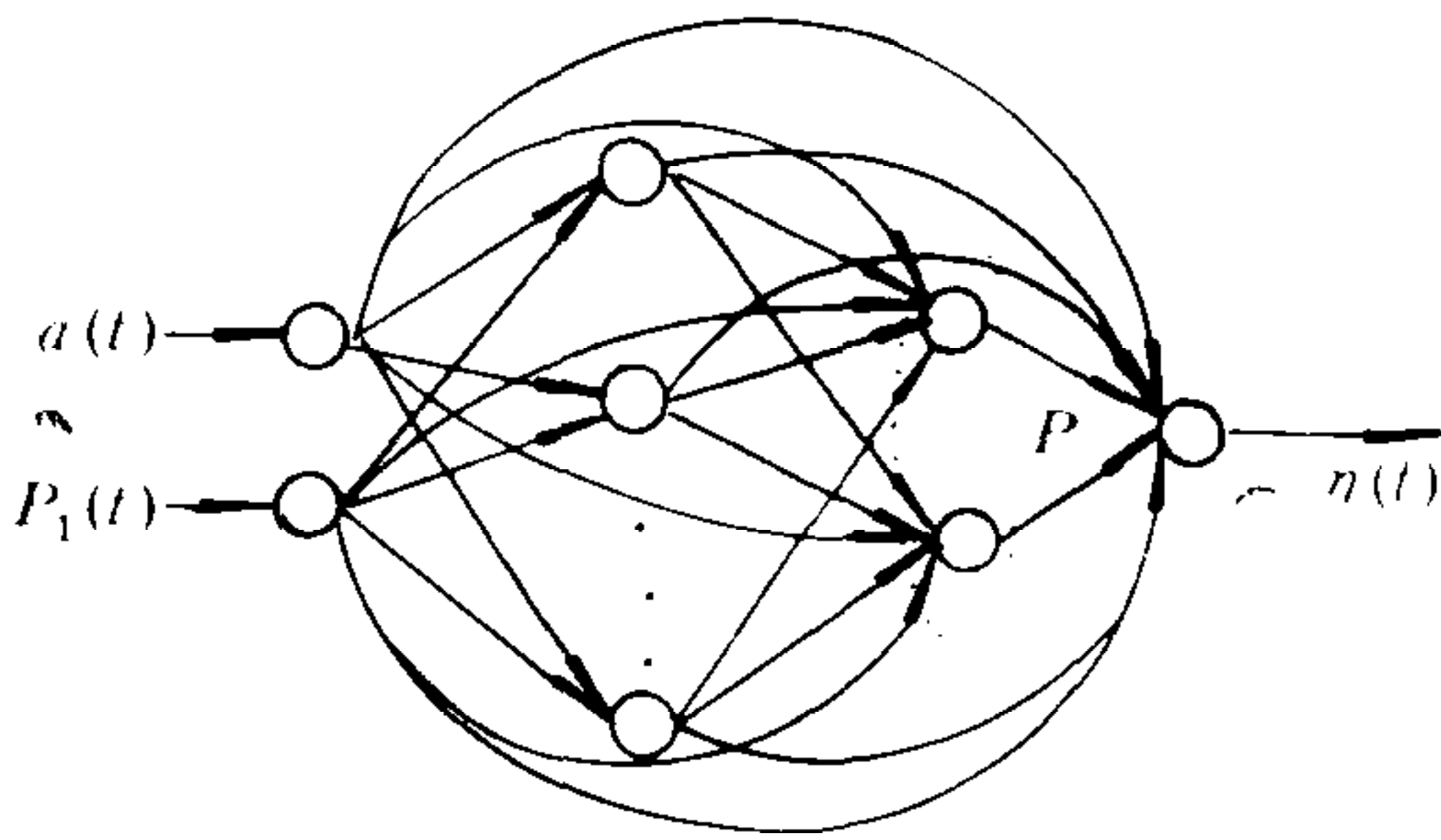


图 4 合成塔多层复杂前馈网络结构图

网络结构选为 2-8-3-1, 其中虚线所圈部分为代表压力  $P(t)$  的神经元群. 利用前面所推导的学习算法,以  $\{(a(t), P_1(t)), P(t), \eta(t)\}$  的相应的 40 组采样样本和 10 组均值样本对上述网络进行训练, 这里考虑以第二隐层的三个神经

元的输出为压力  $P(t)$ , 采样样本训练的目标函数定为 ( $M$  为加权系数)

$$\min E = \frac{1}{2} \sum_{k=1}^{40} (P(kT + \tau) - \hat{P}(k))^2 + \frac{M}{2} \sum_{k=1}^{40} (\eta(kT + \tau) - \hat{\eta}(k))^2. \quad (4-1)$$

对均值样本可类似进行定义. 网络中各参数将以该偏差  $E$  对各参数的偏导数作为修正的依据. 为此本文提出一种新的分层分步迭代学习算法, 具体步骤如下:

Step 1. 初始化, 置各参数初值;

Step 2. 对图 4 中由输入神经元到第二隐层的三个输出元的内层所含的三层传统网络以下列目标函数进行训练, 直至满足要求为止,

$$\min E_1 = \frac{1}{2} \sum_{k=1}^{40} (P(kT + \tau) - \hat{P}(k))^2; \quad (4-2)$$

Step 3. 以上述内层网络参数及其它网络参数计算

$$E_2 = \frac{M}{2} \sum_{k=1}^{40} (\eta(kT + \tau) - \hat{\eta}(k))^2, \quad (4-3)$$

若  $E_2$  满足要求, 则转 Step 5, 否则转 Step 4;

Step 4. 以目标函数

$$\min E_2 = \frac{M}{2} \sum_{k=1}^{40} (\eta(kT + \tau) - \hat{\eta}(k))^2 \quad (4-4)$$

对上述多层复杂前馈神经网络参数进行修正, 返回 Step 2;

Step 5. 训练结束, 对网络进行测试, 并计算总误差  $E$ , 若不满足要求, 则返回 Step 1.

作为比较, 对于相同的样本, 选择了一个结构为 2-10-9-2 的四层传统前向神经网络进行了建模研究, 其输入为  $a(t), P_1(t)$ , 输出为  $P(t), \eta(t)$ , 训练目标函数为

$$\min E = \frac{1}{2} \sum_{k=1}^{40} (P(kT + \tau) - \hat{P}(k))^2 + \frac{1}{2} \sum_{k=1}^{40} (\eta(kT + \tau) - \hat{\eta}(k))^2. \quad (4-5)$$

对随机获取的 10 组测试样本进行的测试结果如表 1 所示.

表 1 网络测试数据及结果

(其中  $P(t), \eta(t)$  为样本值;  $P^*(t), \eta^*(t)$  为传统网络计算值;  $P^*(t), \eta^*(t)$  为前馈网络计算值)

$k$	$a(t)$	$P_1(t)$	$P(t)$	$P^*(t)$	$P^*(t)$	$\eta(t)$	$\eta^*(t)$	$\eta^*(t)$
1	2.818	0.409	15.11	15.1232	15.1198	0.517	0.51841	0.51822
2	2.837	0.371	14.84	14.8571	14.8523	0.538	0.53667	0.53817
3	2.851	0.385	14.76	14.7441	14.7551	0.551	0.55001	0.55176
4	2.867	0.376	14.58	14.5992	14.5978	0.563	0.56146	0.56234
5	2.883	0.389	14.21	14.1936	14.2301	0.572	0.57143	0.57346
6	2.895	0.379	13.81	13.8074	13.8114	0.577	0.57872	0.57609
7	2.906	0.397	13.92	13.9091	13.9311	0.574	0.57398	0.57532
8	2.921	0.374	14.12	14.1351	14.1299	0.563	0.56214	0.56421
9	2.943	0.387	14.70	14.6873	14.6915	0.542	0.54299	0.54213
10	3.041	0.401	14.91	14.9310	14.9022	0.511	0.50954	0.50932

此时图 4 所示多层复杂前馈网络学习的计算机耗时为传统网络耗时的 1/3 强, 而由表 1 的结果可知其精度却更高. 类似地, 有关尿素汽提塔等神经网络建模也得到了同样的结果. 这一结果说明了所提出的多层复杂前馈网络的有效性及其工程应用意义. 这些模型为在某大型化肥厂所开发的尿素生产实时专家系统提供了可靠的过程知识<sup>[5]</sup>. 该项研究已通过国家级鉴定, 经济效益极为显著.

基于上述前馈网络模型, 我们还利用最速下降法解决了尿素合成塔的操作优化问题, 得到了相应于负荷变化的最优  $a(t)$  与  $P_1(t)$ .

## 5 结论

尿素生产过程, 是一个复杂的、多变量的、干扰众多的连续过程. 对如此困难的实际对象, 本文证明了经过改进的神经网络方法可以有效地解决其中关键设备的建模和优化问题. 这一方法同样可应用于其它复杂工业过程.

## 参 考 文 献

- [1] Haesloop D, Holt B R. A neural networks structure for system identification. *Proc. of 1990 American Control Conference*, 3: 2460—2461.
- [2] Lambert J M, Hecht-Nielson R. Application of Feedforward and recurrent neural networks to



- chemical plant predictive modeling. Proc. of IJCNN, 1991, 1:373—378.
- [3] Parlos A G et al. Application of the recurrent Multilayer perceptron in modeling complex process dynamics. *IEEE Trans. NN*, 1994, 5(2):255—266.
- [4] 周金荣,黄道,蒋慰孙. 尿素生产专家系统的开发研究. 华东理工大学学报,1994,20(2): 205—209.
- [5] 周金荣,黄道,蒋慰孙. 基于神经网络的随机动态系统优化新方法. 华东化工学院学报,1992,18(增刊): 27—32.

## THE APPLICATION STUDY OF A KIND OF NOVEL NEURAL NETWORKS MODELS

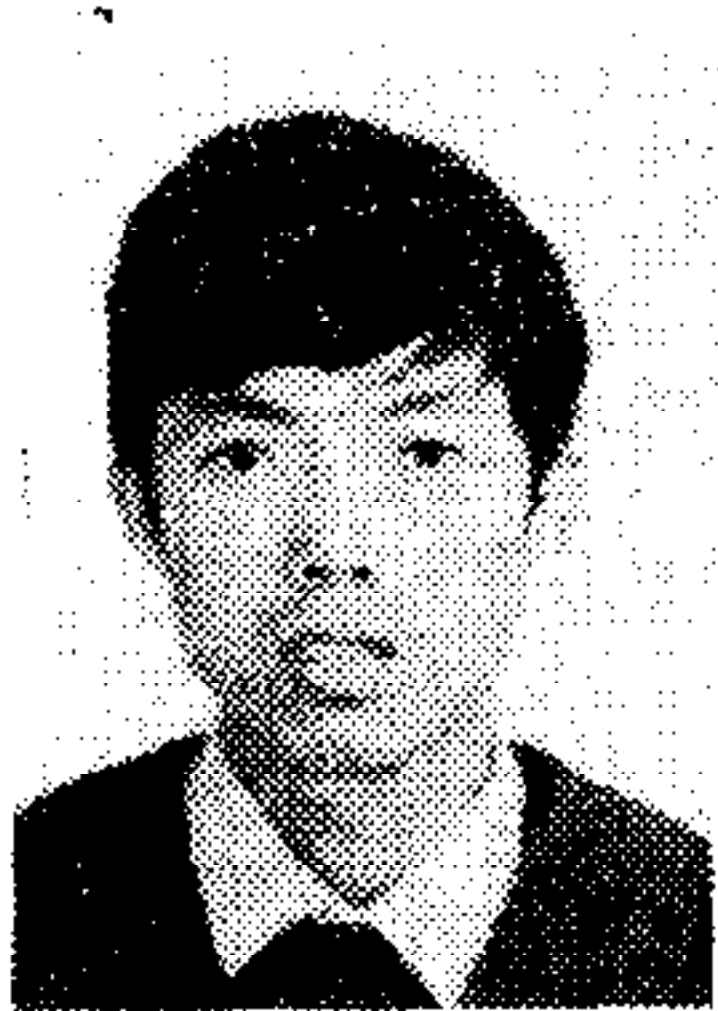
ZHOU JINRONG HUANG DAO JIANG WEISUN

(Res. Inst. of Automatic Control, East China Univ. of Sci. & Tech., Shanghai 200237)

### ABSTRACT

In this paper, a kind of novel neural networks models are proposed via the combination between mechanism analysis and determination of neural networks structure. The corresponding learning algorithms sampling pattern and training pattern are given in details are the patterns. All ideas and methods are effective and verified through the application in a complex urea production process.

**Key words:** Neural networks, mechanism knowledge, learning algorithm.



**周金荣** 1965年生于江苏省姜堰市。1993年毕业于华东理工大学自控系,获博士学位。现为华东理工大学自动化研究所副教授。主要研究兴趣是复杂化工过程计算机控制、神经网络理论与工业应用以及工业过程专家系统开发等。



**黄道** 1947年生于江苏常州。1985年毕业于华东理工大学自控系,获博士学位。现为华东理工大学自动化研究所所长、教授。长期从事过程数学模型和计算机控制方面的研究,主要学术方向是计算机集成生产系统和工业过程智能控制等。

**蒋慰孙** 照片、简介见本刊第18卷第1期。