

研究论文

利用结构逼近式混合神经网络 实现间歇反应器的建模

曹柳林, 李晓光, 王 晶

(北京化工大学信息科学与技术学院自动化研究所, 北京 100029)

摘要: 提出了一种新的混合神经网络建模方法——结构逼近式混合神经网络。基于此结构建立的混合神经网络可以充分利用已知非线性系统的结构信息, 使神经网络“灰盒”化, 更好地解释和描述系统各变量间的因果关系, 从而提高网络的建模精度和模型的可靠性。本文介绍了这类神经网络的基本特性、拓扑结构和训练方法。报告了一个典型放热液相二级平行间歇反应的建模过程; 并针对间歇反应过程测量滞后的情况, 与两种不同的混合神经网络模型作了比较, 仿真和比较结果证明了方法的有效性。

关键词: 结构逼近式混合神经网络; 间歇反应器; 建模; 串联混合神经网络

中图分类号: TQ 316.2

文献标识码: A

文章编号: 0438-1157 (2008) 04-0958-06

Modeling of batch reactor based on structure approaching hybrid neural networks approach

CAO Liulin, LI Xiaoguang, WANG Jing

(Institute of Automation, School of Information Science and Technology, Beijing University
of Chemical Technology, Beijing 100029, China)

Abstract: A new hybrid neural networks modeling approach, named structure approaching hybrid neural network (SAHNN) is presented in this paper. The characteristics and structure of this approach are introduced. The approach which combines first-principles model with neural networks fully utilizes structural information of known nonlinear system, makes neural networks a “grey-box”, describes and explains the consequence relation of system variables in a better way. A detailed analysis of SAHNN modeling was performed. Considering delayed measurement of outputs, simulation of a chemical batch reactor was performed and comparison of two types of hybrid neural network (HNN) was made. The results of simulation and comparison showed that the approach was a promising tool to model complicated nonlinear system effectively and could be utilized as a vehicle to control and optimize chemical reactors.

Key words: SAHNN; batch reactor; modeling; HNN

引 言

间歇反应过程在化学过程工业中占有十分重要

的地位。间歇过程具有灵活多变的特性, 可以用同一套多用途、多功能的设备生产多种类型的产品, 特别适合聚合物、生物制品和特殊化学品的制备。

2007-05-09 收到初稿, 2007-09-16 收到修改稿。

联系人: 李晓光。第一作者: 曹柳林 (1953—), 女, 教授。

基金项目: 国家自然科学基金项目 (60704011)。

Received date: 2007-05-09.

Corresponding author: LI Xiaoguang. E-mail: 2004080102@grad.buct.edu.cn

Foundation item: supported by the National Natural Science Foundation of China (60704011).

然而由于间歇过程的强非线性、缺少稳态操作条件、反应过程的不确定性、未知扰动、过程变量的限制，不易建立准确的数学模型，进而难于实施自动控制。因此，间歇反应过程建模是现在研究的热点和难点。建立间歇反应过程的数学模型一般通过3种方式：一是利用机理和实验的方法得到系统的数学模型^[1]；二是利用单一和混合神经网络方法来获得过程的黑箱模型^[2]；三是将机理模型和神经网络方法相结合，利用神经网络对机理模型当中的某个时变参数进行估计，得到系统的灰箱模型^[3-5]。在第二、三种建模方法的基础之上，本文提出了一种新的混合神经网络建模方法：结构逼近式混合神经网络（SAHNN）。基于这种方法，对一类典型的间歇反应建立了数学模型，并和两种不同类型的混合神经网络模型作了比较说明，证明了方法的有效性。

1 间歇反应过程

1.1 间歇反应过程描述

间歇反应过程^[6-7]包括间歇反应器和夹套换热系统。典型的间歇过程如图1所示。该过程包括两个并行的高度放热的反应，反应为搅拌充分、液相定容过程。

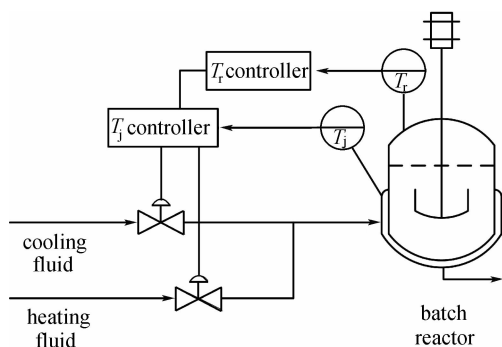
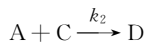
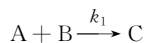


图1 间歇反应过程

Fig. 1 Diagram of batch reactor process

反应过程为



其中，A和B分别代表原料，C和D分别代表主产物和副产物； k_1 、 k_2 为反应速率常数，符合Arrhenius方程。

1.2 数学模型

物料平衡方程

$$\frac{dM_A}{dt} = -k_1 M_A M_B - k_2 M_A M_C$$

$$\frac{dM_B}{dt} = -k_1 M_A M_B$$

$$\frac{dM_C}{dt} = +k_1 M_A M_B - k_2 M_A M_C$$

$$\frac{dM_D}{dt} = +k_2 M_A M_C \quad (1)$$

能量平衡方程

$$\frac{dT_r}{dt} = \frac{[-\Delta H_1 (k_1 M_A M_B) - \Delta H_2 (k_2 M_A M_C)] + UA (T_j - T_r)}{(M_A + M_B + M_C + M_D) C_{p_r}} \quad (2)$$

其中，

$$C_{p_r} = \frac{(C_{p_A} M_A + C_{p_B} M_B + C_{p_C} M_C + C_{p_D} M_D)}{M_r}$$

反应速率函数

$$R_1 = \exp\left(k_1^1 - \frac{k_1^2}{T_r + 273.15}\right) M_A M_B \quad (3)$$

$$R_2 = \exp\left(k_2^1 - \frac{k_2^2}{T_r + 273.15}\right) M_A M_C \quad (4)$$

其中， T_r 代表间歇反应的反应温度， T_j 代表夹套温度。间歇过程的动态过程由式(1)、式(2)表示。 M_A 、 M_B 、 M_C 和 M_D 的初始值分别为12、12、0、0 kmol。 T_j 和 T_r 的初始温度均为20℃。反应温度和夹套温度的变化范围为20~120℃，间歇反应器反应温度的参考值设定在92.0℃。其他过程参数见文献[6]。

2 结构逼近式混合神经网络模型

本文提出的结构逼近式神经网络模型是利用B样条神经网络（BSNN）和对角递归神经网络（DRNN）混合而成。网络结构见文献[8-9]。

2.1 结构逼近式混合神经网络

从状态空间描述的角度出发，本文针对非线性动态系统的特点，提出了如图2所示的混合神经网络拓扑图，并将其命名为结构逼近式混合神经网络。

实际化工过程中存在这样一类非线性对象，其非线性因素 $R=f(\mathbf{X})$ 具有稳态的可分离特性，它是系统状态变量的非线性函数，可使用非线性代数方程描述。若把 R 作为系统的内部输入变量来考虑，则这类对象便可转化为线性系统。因此，可以采用线性状态空间描述方法来描述这类非线性系统

$$\begin{cases} \dot{\mathbf{X}} = L_1(\mathbf{R}) + L_2(\mathbf{U}), & \mathbf{R} = f(\mathbf{X}) \\ \mathbf{Y} = L_3(\mathbf{X}) \end{cases}$$

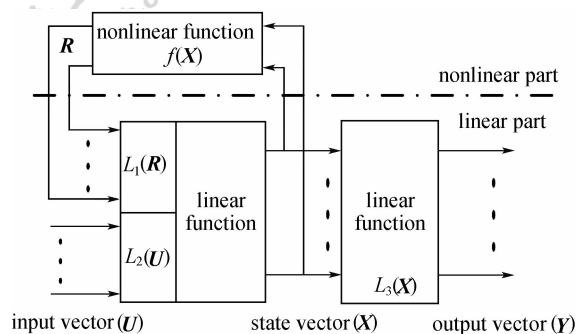


图 2 非线性过程线性描述

Fig. 2 Linear description for nonlinear process

式中 $\mathbf{X}$ 、 $\mathbf{U}$ 、 $\mathbf{Y}$ 分别代表系统的状态向量、输入向量和输出向量； $\mathbf{R}$ 代表非线性部分输出向量； L_1 、 L_2 、 L_3 分别代表 3 个不同的线性函数； f 表示非线性函数。在图 2 中，网络各层输出分别代表实际系统的输入向量、状态向量和输出向量；而每一层中的节点个数又分别对应于各个向量维数，与实际系统中各变量的意义一一对应。若把其中非线性因素作为内部输入变量处理之后，神经网络结构与描述系统的状态空间方程一致，因此称之为结构逼近式混合网络。依照此结构建立起来的混合神经网络可以充分利用已知的非线性系统的结构信息，使神经网络“灰盒”化，更清楚地解释和描述系统输入、输出及中间变量间的因果关系，从而把握各变量间的定性逻辑，提高网络的建模精度。

化学反应过程的模型描述普遍存在着上述特点，可以被分解成非线性静态和线性动态两部分，因而可以使用结构逼近式混合神经网络来建模。反应器中的非线性因素主要表现在反应速率常数 k 的 Arrhenius 方程中，如式 (3)。该非线性代数方程描述了反应速率函数与反应温度和反应物浓度之间的静态关系，当把此非线性因素 $\mathbf{R} = f(\mathbf{X})$ 作为系统的内部输入变量来考虑，这类系统便转化为一个线性动态系统。结合式 (1) ~ 式 (4) 构建混合神经网络，如图 3 所示，该混合网络由两部分组成，一部分为非线性部分，采用 BSNN 描述式 (3)、式 (4)；其输入为系统的状态向量，输出为反应速率函数；另一部分为动态网络，注意到对于间歇反应，也包括 CSTR 等反应，线性部分的状态空间形式中的状态转移矩阵都是以对角的形式出现，同时考虑到系统的动态行为，采用线性 DRNN 描述式 (1)、式 (2)，该网络由输入层、普通隐含层、带有对角反馈权值的状态隐含层和输出层组成。

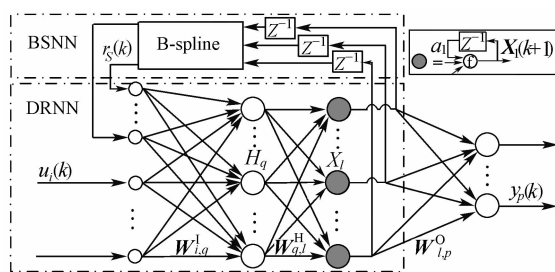


图 3 SAHNN 拓扑图

Fig. 3 Topology for SAHNN

系统的状态空间描述如下

$$\begin{cases} \mathbf{X}(k+1) = \mathbf{A}\mathbf{X}(k) + \mathbf{W}^H \mathbf{W}^I \mathbf{U}(k) \\ \mathbf{Y}(k) = \mathbf{W}^O \mathbf{X}(k) \end{cases}$$

其中， r_s 表示非线性部分第 s 个输出变量； H_q 表示隐含层的第 q 个节点输出； $\mathbf{W}^I \in \mathbb{R}^{h \times m}$ ， $\mathbf{W}^H \in \mathbb{R}^{n \times h}$ ； $\mathbf{W}^O \in \mathbb{R}^{p \times n}$ ； $\mathbf{X} \in \mathbb{R}^{n \times 1}$ ， $\mathbf{Y} \in \mathbb{R}^{p \times 1}$ ， $\mathbf{U} \in \mathbb{R}^{m \times 1}$ ，分别表示 SAHNN 状态向量、输出向量和输入向量，输入向量中包含 BSNN 的输出向量。

系统的状态转移矩阵 $\mathbf{A}$ 为

$$\mathbf{A} = \begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & a_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & a_n \end{bmatrix}^{n \times n}$$

系统的每一个状态变量的数学描述如下

$$\begin{aligned} x_i(k+1) &= a_i x_i(k) + \sum_q H_q \mathbf{w}_{q,i}^H = \\ &= a_i x_i(k) + \sum_q \sum_i^m \mathbf{w}_{i,q}^I \mathbf{w}_{q,i}^H u_i(k) \\ u_i(k) &= \sum_{b=1}^v \mathbf{w}_{i,b}^{\text{BS}} \tilde{B}_b[\mathbf{X}(k)] \quad i = 1, \dots, s, b = 1, \dots, v \end{aligned} \quad (5)$$

式中 $\tilde{B}_b[\mathbf{X}(k)]$ 为 B 样条函数， $\mathbf{w}_{i,b}^{\text{BS}}$ 为其对应的权值。

可以推出

$$\begin{aligned} x_i(k+1) &= a_i x_i(k) + \sum_q \sum_{i=s+1}^m \mathbf{w}_{i,q}^I \mathbf{w}_{q,i}^H u_i(k) + \\ &= \sum_q \sum_{i=1}^s \sum_{b=1}^v \mathbf{w}_{i,q}^I \mathbf{w}_{q,i}^H \mathbf{w}_{i,b}^{\text{BS}} \tilde{B}_b[\mathbf{X}(k)] = \\ &= a_i x_i(k) + \sum_q \sum_{i=s+1}^m d_{q,i,i} u_i(k) + \sum_q \sum_{i=1}^s \sum_{b=1}^v g_{q,i,b,i} \tilde{B}_b[\mathbf{X}(k)] \end{aligned}$$

当系统的状态变量是可测的，而且非线性系统已经转化为权向量的线性形式，构成增广矩阵，所以系统参数 θ 和估计参数 $\mathbf{W}_{i,p}^O$ 可以通过递推最小二乘的方法进行训练和辨识。

θ_l 表示为

$$\begin{aligned} \theta_l = & [a_l, d_{1,1,l}, d_{1,2,l}, \dots, d_{1,m,l}, \dots, d_{2,1,l}, d_{2,2,l}, \dots, \\ & d_{2,m,l}, \dots, d_{h,1,l}, d_{h,2,l}, \dots, d_{h,m,l}, g_{1,1,l}, \dots, g_{1,v,l}, \dots, \\ & g_{1,m,1,l}, \dots, g_{1,m,B,l}, \dots, g_{h,1,1,l}, \dots, g_{h,m,v,l}]^T \\ \varphi_l(k) = & \{x_l(k), u_{s+1}(k), \dots, u_m(k), \dots, u_{s+1}(k), \dots, \\ & u_m(k), \tilde{B}_1[\mathbf{X}(k)], \dots, \\ & \tilde{B}_v[\mathbf{X}(k)], \dots, \tilde{B}_1[\mathbf{X}(k)], \dots, \tilde{B}_v[\mathbf{X}(k)]\}^T \\ x_l(k+1) = & \varphi_l^T(k) \theta_l \end{aligned}$$

系统的输出方程由式 (6) 表示

$$y_p(k) = \sum_i^n W_{i,p}^O x_i(k) \quad (6)$$

对于B样条神经网络,一旦其样条个数与阶次确定下来,输入输出关系就确定了,它的训练主要是依据混合模型的误差、选择样条函数的个数与阶次的过程。

2.2 间歇反应过程建模

将反应速率方程式 (3)、式 (4) 分别代入式 (1)、式 (2) 中,间歇过程的动态方程为式 (7),其状态空间形式如式 (8) 所示。

$$\begin{aligned} \frac{dM_A}{dt} &= -R_1 - R_2 \\ \frac{dM_B}{dt} &= -R_1 \\ \frac{dM_C}{dt} &= +R_1 - R_2 \\ \frac{dM_D}{dt} &= +R_2 \\ \frac{dT_r}{dt} &= \frac{UA(T_j - T_r)}{M_r C p_r} + \frac{1}{M_r C p_r} (-\Delta H_1 R_1 - \Delta H_2 R_2) \end{aligned} \quad (7)$$

$$\begin{aligned} \begin{bmatrix} \dot{M}_A \\ \dot{M}_B \\ \dot{M}_C \\ \dot{M}_D \\ \dot{T}_r \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-UA}{M_r C p_r} \end{bmatrix} \begin{bmatrix} M_A \\ M_B \\ M_C \\ M_D \\ T_r \end{bmatrix} + \\ & \begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ \frac{-\Delta H_1}{M_r C p_r} & \frac{-\Delta H_2}{M_r C p_r} & \frac{UA}{M_r C p_r} \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ T_j \end{bmatrix} \end{aligned} \quad (8)$$

根据对间歇反应过程的分析并结合式 (7),选

取夹套温度 T_j 作为操纵变量,即 $[u_1] = [T_j]$,其维数为 1; 结构逼近式神经网络的输出向量和状态向量相同,即 $\mathbf{Y} = [M_A \ M_B \ M_C \ M_D \ T_r]^T$,其维数为 5,根据式 (6) 可知, $\mathbf{W}^O$ 为维数为 5 的单位矩阵; 另外,系统包括两个化学反应速率,即在混合网络中,非线性部分的输出向量维数为 2,该输出向量作为 SAHNN 的内部输入变量,和系统的操纵变量构成混合神经网络的输入向量 $\mathbf{U} = [u_1, u_2, u_3] = [T_j, R_1, R_2]$,维数为 3。

根据文献 [6] 的间歇反应过程,该过程的反应时间定为 200 min,着重体现放热、非线性动态行为,200 min 后反应进入拟稳态。过程的输出变量的定义域分别为: $M_A \in [0, 12]$, $M_B \in [0, 12]$, $M_C \in [0, 8]$, $M_D \in [0, 2]$, $T_r \in [20, 120]$; 输入变量的定义域为: $T_j \in [20, 120]$ 。根据式 (3)、式 (4) 可知: $R_1 = f(M_A, M_B, T_r)$, $R_2 = f(M_A, M_C, T_r)$ 。结合式 (5),SAHNN 中,BSNN 分别由 k 时刻的系统状态作为输入,反应速率函数作为其输出,选用阶次为 3,内插点为 4,即样条的基函数为 7。DRNN 选用 3 个输入节点,10 个隐含层节点,5 个输出节点。

为了说明建模方法的有效性,将其和两种不同类型的混合神经网络模型 (HNN) 作了比较说明: HNN1^[8] 由 BSNN 和 DRNN 串联构成,采用了较 BP 算法收敛更快、精度更高的递推预报误差算法 (RPE),预测当前时刻网络的输出值,并采用交叉检验的方法优化网络结构参数; HNN2^[5] 将机理方程和神经网络相结合,利用神经网络对过程中不可测的时间参数 (反应速率) 做预估,再将预估值带入机理方程中,获得混合神经网络模型。

对于 3 种混合神经网络模型,采用如下策略产生训练样本。训练过程的采样时间为 0.1 min,输入变量 T_j 为持续激励的覆盖整个操作域的随机信号,数据长度为 2000,对训练样本数据作归一化处理,网络训练采用相同的精度作为收敛标准。SAHNN 在 400 个样本左右达到收敛精度, HNN1 和 HNN2 则分别需要 1300 和 1800 个样本完成训练,SAHNN 的收敛速度较 HNN 快 3~4 倍。

测试过程的采样时间取 1 min,输入变量 T_j 为持续激励的实际操作过程信号^[6-7],数据长度为 200,对于相应的输出信号,即各个反应物和反应温度分别加入 $[0, 0.01]$ 和 $[0, 0.5]$ 的随机信

号模拟测量噪声。

SAHNN 4 个浓度输出与 HNN2 的比较结果见图 4，温度与 HNN1 的比较见图 5，3 种模型输出的均方差如表 1 所示。仿真和比较的结果表明：SAHNN 具有收敛速度快、预测精度高和泛化能力强的特点。特别需要指出的是，SAHNN 模型在系统输出不能实时测量的情况下，使用网络模型输出代替系统输出参与测试，仍能获得很好的建模效果，其精度为其他两种的 4~9 倍，这对于工业过程中产品质量指标不能在线测量的建模过程，尤其具有积极的意义。

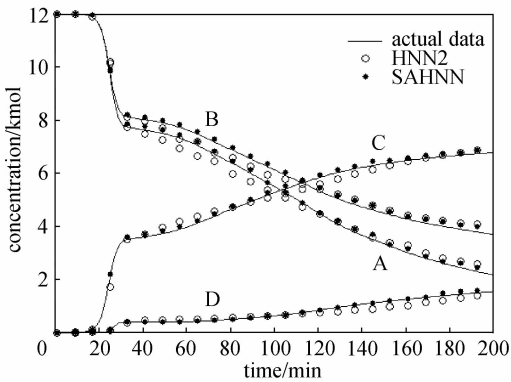


图 4 SAHNN 和 HNN2 物质浓度预测结果比较
Fig. 4 Prediction for concentration
by SAHNN and HNN2

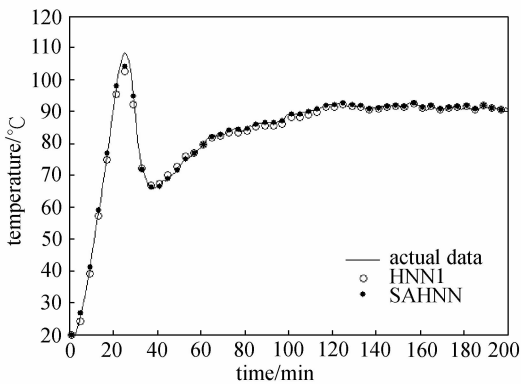


图 5 SAHNN 和 HNN1 反应温度预测结果比较
Fig. 5 Prediction for temperature
by SAHNN and HNN1

表 1 SAHNN 和 HNN 平均误差平方和比较
Table 1 Comparison of MES by SAHNN and HNN

Model	$M_A \times 10^{-3}$	$M_B \times 10^{-3}$	$M_C \times 10^{-4}$	$M_D \times 10^{-4}$	$T_j \times 10^{-1}$
SAHNN	3.245	4.178	8.569	4.712	6.155
HNN1	15.62	22.96	75.34	34.86	31.27
HNN2	13.74	19.12	54.88	21.23	82.51

3 结 论

本文提出了一种新的混合神经网络的建模方法：结构逼近式混合神经网络，利用该方法对一类典型的化工间歇反应过程建立了数学模型，仿真和比较的结果证明该方法的有效性。结构逼近式神经网络充分利用了已知非线性系统的结构信息和特点，使神经网络“灰盒”化，更好地解释和描述系统各变量间的因果关系，同时，根据状态空间描述确定神经网络的结构，更多地获得实际系统的内部信息，从而提高了网络的建模精度和模型的可靠性；建模过程中，将过程的非线性因素（如反应速率）分离出来，将整个过程转化为线性系统，可用状态结构形式描述，结构简单，清晰地表征了系统的结构信息，模型和系统结构的相似性保证了变量间的因果关系的定性趋势，因而保证了逻辑关系的正确性。因此，在建模过程中，所需训练样本少，泛化能力强；考虑到模型最终的表述形式，其增广的参数矩阵为线性形式，可以采用递推最小二乘辨识参数，辨识的速度快、精度高，克服了其他算法，如 BP 等算法收敛慢、容易陷入局部极值等问题；对那些输出变量不能实时测量、只能用分析数据作为训练数据的过程建模，该方法十分有效，有一定的工程意义；混合模型采用了线性与非线性部分的分离，有利于控制策略的制定与实施，相关研究已经证明了方法的有效性。

References

[1] Gentric G, Pla R, Corriou J P. Experimental study of the nonlinear geometric control of a batch emulsion polymerization reactor. *Computers Chem. Engng.*, 1997, **12** (suppl.): 1043-1048

[2] Mohan Krothapally, Srinivas Palanki. A neural network strategy for batch process optimization. *Computers Chem. Engng.*, 1997, **121** (suppl.): 463-468

[3] Klimauskas Casimir C. Hybrid modeling for robust nonlinear multivariable control. *ISA Transactions*, 1998, **37**: 291-297

[4] Molga E J. Neural network approach to support modeling of chemical reactors: problems, resolutions, criteria of application. *Chemical Engineering and Processing*, 2003, **42**: 675-695

[5] Ng C W, Hussain M A. Hybrid neural network — prior knowledge model in temperature control of a semi-batch polymerization processes. *Chemical Engineering and*

- Processing, 2004, **43**: 559-570
- [6] Cott B, Macchietto S. Temperature control of exothermic batch reactor using generic model control. *Industrial and Engineering Chemistry Research*, 1989, **28**: 1177-1184
- [7] Hussain M A, Ng C W, Aziz N, Mujtaba I M. Neural network techniques and application in chemical process control system. *Intelligent System Techniques and Application*, 2002 (5): 326-362
- [8] Cao Liulin (曹柳林), Xu Xinghua (徐兴华), Jiang Pei (江沛). Viscosity prediction for PET process based on hybrid neural networks//Proceeding of the 6th World Congress on Control Automation. Dalian, 2006, **6**: 21-23
- [9] Cao Liulin (曹柳林), Wu Haiyan (吴海燕). MWD modeling and control for polymerization *via* B-spline neural network. *Journal of Chemical Industry and Engineering (China)*(化工学报), 2004, **55** (5): 742-746