

复杂非线性系统的多神经网络预测控制

熊智华 王雄 徐用懋

(清华大学自动化系 北京 100084)

摘要 针对工作点变化的复杂非线性系统,提出了一种基于多神经网络的预测控制方法.采用多神经网络建立预测模型,其中子网络的选取采用广义信息熵融合方法,从而提高了预测模型的准确性和可靠性,并实时递推计算子网络输出的连接权值,以适应对象在不同工作区域内变化. *CSTR* 对象的仿真研究表明了该方法的有效性.

关键词 多神经网络, 广义信息熵, 非线性系统, 预测控制

1 引言

复杂工业过程往往呈现出高度非线性,并且由于特性的变化而在多个工作点区域内波动,采用多模型方法来解决是一种较为简单而实用的方法^[1].采用多个神经网络的权连接方式使得每个子模型对输出都有一定的贡献,有利于消除模型切换引起的抖动、避免振荡^[2],因此采用多神经网络(Multiple Neural Networks, *MNN*)构成的预测模型,可以提高预测控制系统的稳定性和鲁棒性.作者在文献[3]中给出了构成多神经网络模型的单个子网络设计指导性条件,并提出了一种基于广义信息熵的融合策略来优选子网络的方法,达到了降低整个模型不确定性、提高模型可靠性的目的.

本文采用基于广义信息熵的融合方法来优选子网络,利用子网络输出与对象输出之间的残差,实时递推计算连接权值,以适应对象在不同工作点区域变化.进而提出了一种基于多神经网络模型的非线性预测控制方法,利用多神经网络预测模型来预报过程未来的输出,采用序贯二次规划算法(SQP)来求解输入输出受约束情况下的非线性优化问题.对一典型 *CSTR* 对象的仿真结果表明了该方法的有效性.

2 建立基于多神经网络的预测模型

模型预测控制(MPC)近年来在工业过程中得到了广泛应用,复杂非线性系统的预测控制也受到了广泛的关注^[4,5].对于复杂非线性过程,预测控制需要解决的问题主要包括预测模型的建立、非线性优化策略等.本文采用基于多神经网络的预测控制方法解决复杂非线性过程因特性的变化而在多个工作点区域内波动的问题,其中采用多神经网络来建立非线性对象的预测模型.

假设所考虑的复杂非线性过程可以采用下列离散时间 NARMAX 模型^[4]描述:

$$y(t)=f(y(t-1),\dots,y(t-n_y),u(t-1),\dots,u(t-n_u))+e(t) \quad (1)$$

其中 u 为输入向量, $u \in R^p$, y 为输出向量, $y \in R^q$, e 为噪声向量, $f(\cdot)$ 为非线性函数. 采用多神经网络模型来逼近该非线性对象, 其结构如图 1 所示^[3].

定义 MNN 的输入向量为 $X(t)=[y(t-1),\dots,y(t-n_y),u(t-1),\dots,u(t-n_u)]^T$, 假定非线性对象在 n 个工作点区域内变化, 则可将样本数据经过聚类分析得到与工作点区域相对应的 n 组样本子集, 采用各样本子集训练、测试得到 n 个子网络模型 $NN_i (i=1,\dots,n)$, 整个 MNN 模型输出 $\tilde{Y}(t)$ 为各子网络模型输出 $Y_i(t)$ 的加权和, 为

$$\tilde{Y}(t) = \sum_{i=1}^n W_i Y_i(t) = \sum_{i=1}^n W_i \tilde{f}_i(X_i(t)) \quad (2)$$

式中 $W_i = \text{diag}(W_{i1}, \dots, W_{iq})$, 为 MNN 模型输出与第 i 个子网络模型输出之间的连接权值阵, $Y_i(t)$ 为子网络的输出, $X_i(t)$ 为 NN_i 的输入向量, 由经聚类分析得到第 i 个样本数据构成, $\tilde{f}_i(\cdot)$ 为 NN_i 实现的非线性关系.

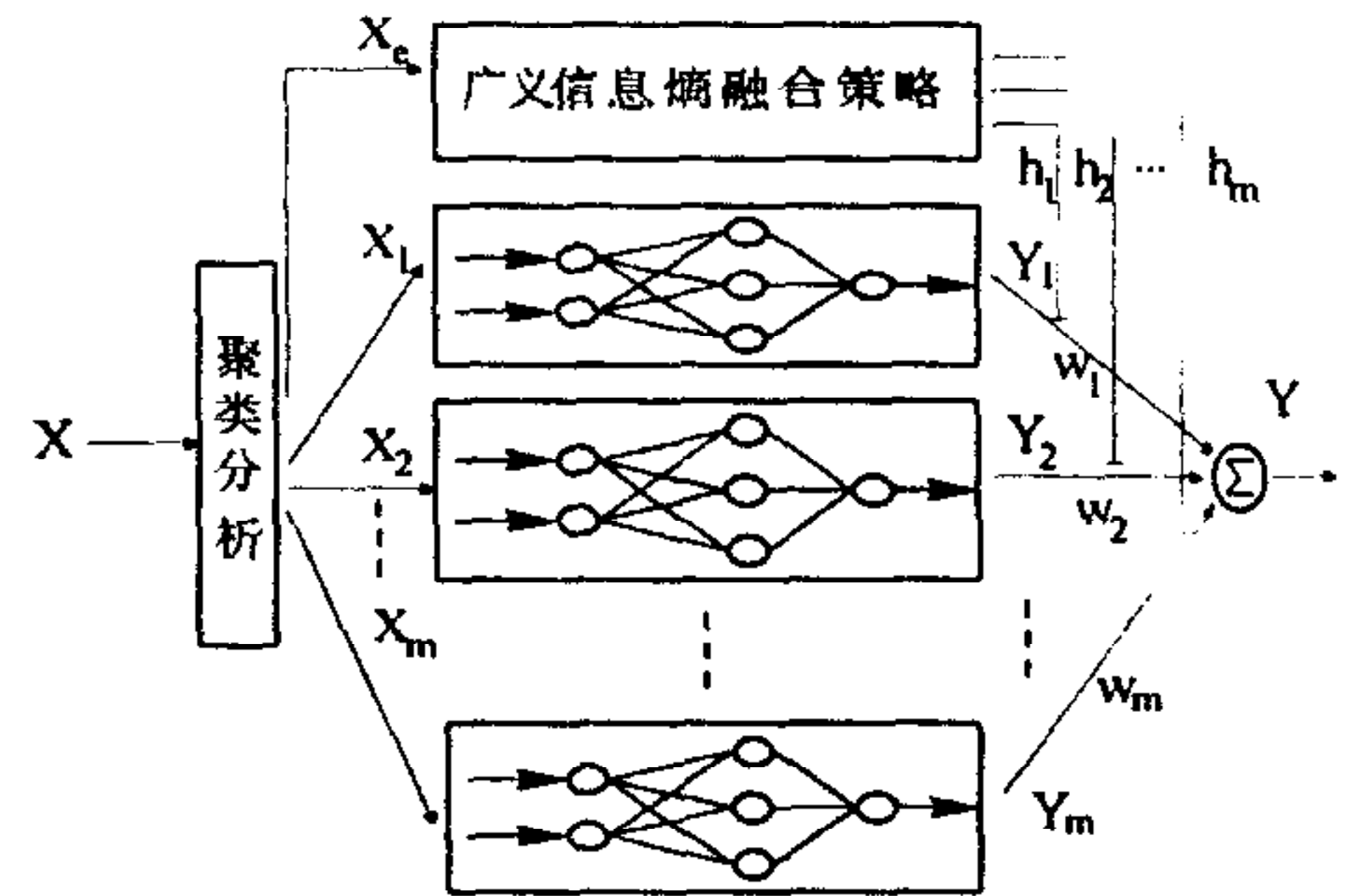


图 1 基于广义信息熵融合策略的 MNN 模型

对于上述 n 个子模型, 往往需要建立 $N (N \geq n)$ 个网络模型并从中选出 n 个子网络, 现有的选取方法是按照一定的准则如误差平方和(SSE)等判断其优劣来选取子网络, 或是根据经验来选择, 这样难以达到满意的效果. 子网络的选取实质上是 n 个子网络的信息融合过程, 其目的是减小整个模型输出的不确定性, 提高模型可靠性, 即设计子网络 $NN_i (i=1, \dots, n)$ 使得由 $Y_i(t)$ 融合得到的 $\tilde{Y}(t)$ 具有最小的不确定性.

由信息理论^[6]可知, 当 Y_1, \dots, Y_n 的相关性最小, 即 Y_1, \dots, Y_n 相互独立时, 融合得到的模型输出其不确定性达到最小. 为了保证 MNN 模型能够从对象尽量多的信息中辨识非线性对象, 提高整个模型的可靠性、鲁棒性, 子网络模型 NN_i 的结构选取必须满足如下选取准则:

准则 1 子网络 NN_i 之间相关性应尽量小, 包含: (1)子网络的拓扑结构和训练算法等构成机理应尽量不同; (2)子网络应采用不同的样本数据集 $X_i (i=1, \dots, n)$ 进行训练, 使得各子网络模型能够获得关于非线性对象尽可能多的信息.

准则 2 采用不同于训练样本数据集 X_i 的融合样本 X_e 来优选子网络, 构成多神经网络模型.

根据上述子网络的设计准则, 作者在文献[3]中提出了子网络的广义信息熵、整个多神经网络模型的整体信息熵的概念, 并详细推导了基于广义信息熵优选子网络模型的方法. 可知若子网络 NN_i 广义信息熵 h_i 越小, 则决策水平越高, 即决策的不确定性越小, 其可靠性越高; MNN 模型的整体信息熵 H 较小的模型其决策水平较高, 可靠性高. 因此用不确定性较小的子网络模型代替不确定性较大的子网络模型, 可以降低模型的整体信息熵, 提高整个 MNN 模型的决策水平和可靠性.

由此得到基于广义信息熵融合策略的 MNN 模型建模方法. 首先根据非线性对象的

多个工作点, 将对象的样本数据通过聚类分析的方法得到 n 组样本子集, 在此子集内采用不同的网络拓扑结构和训练算法来建立 n 个子网络模型, 然后在信息融合样本集 X_c 内计算各子网络模型的预报输出, 并计算各子网络模型的广义信息熵值 h_i , 通过优选其中广义信息熵值 h_i 小的模型, 并校验其整体信息熵值 H , 从而确定构成整个 MNN 模型的子网络模型. 子网络可以采用 BP 网络等, 子网络的权值参数采用相应的训练算法来得到. 最后计算连接权值矩阵 W_i , 从而建立完整的 MNN 模型.

为了 MNN 模型能够实时地反映被控对象在不同工作区域内变化, 利用子网络输出与对象输出之间的误差均方根值 $R_i(t)$, 实时递推计算连接权值矩阵 $W_i(t)$ ^[7].

由式(2)可知 $W_i(t)=\text{diag}(W_{i1}(t), \dots, W_{iq}(t))$, 采用如下递推式:

$$W_i(t) = \frac{W_i'(t)W_i'^T(t)}{\sum_{j=1}^n W_j'(t)W_j'^T(t)}, \quad W_i'(t) = \frac{\exp(-S_i(t)/2V^2)W_i(t-1)}{\text{tr}[\sum_{j=1}^n \exp(-S_j(t)/2V^2)W_j(t-1)]} \quad (3)$$

其中 $S_i(t)=R_i^T(t)R_i(t)$, $R_i(t)=(r_{i1}(t), \dots, r_{iq}(t))^T$, ($i=1, \dots, n$), $\text{tr}(\cdot)$ 为矩阵求迹函数, V 是控制连接权值收敛速率的参数. 由于系统在多个工作区域内变化, 各子网络模型都有可能趋近系统的理想模型, 因此初始连接权值设置为均值, 即 $W_i(0)=1/n(i=1, \dots, n)$.

定义 $R_m(t)$ 、 $W_m(t)$ 为与非线性系统具有理想匹配程度的模型误差均方根值和对应的连接权值矩阵. 由于 $W_i(t)$ 为对角矩阵, 因此有 $W_i^T(t)W_i(t)=W_i^2(t)$. 根据式(3)可得

$$\frac{W_i(t)}{W_m(t)} = \exp\left[-\sum_{k=0}^{t-1} \frac{2^k}{V^2} (S_i(t-k) - S_m(t-k))\right] \left[\frac{W_i^2(0)}{W_m^2(0)} \right] \quad (4)$$

可得
$$\frac{W_i(t)}{W_m(t)} = \exp\left[-\sum_{k=0}^{t-1} \frac{2^k}{V^2} (S_i(t-k) - S_m(t-k))\right] \cdot \mathbf{I} \quad (5)$$

由于 $\lim_{t \rightarrow \infty} S_i(t) \rightarrow S_m(t)$, 可知
$$\lim_{t \rightarrow \infty} W_i(t)/W_m(t) = \mathbf{I} \quad (6)$$

因为
$$\sum_{i=1}^n W_i(t) = W_m(t) \left[\mathbf{I} + \sum_{i=1, i \neq m}^n W_i(t)/W_m(t) \right] = \mathbf{I} \quad (7)$$

则可得
$$\lim_{t \rightarrow \infty} W_m(t) = \mathbf{I} \quad (8)$$

从式(6)、(8)可知与非线性系统具有理想的匹配程度的子网络模型连接权值矩阵随时间将趋近于 \mathbf{I} , 其余子网络模型连接权值矩阵将趋近于零矩阵, 即近似于采用单一子网络模型来描述非线性系统. 参数 V 的值越小, 连接权值矩阵收敛越快, 但是 V 值太小则会造成过分依赖于单一子网络模型的作用, 从而使系统的鲁棒性下降.

3 预测控制算法

基于多神经网络的预测控制问题可描述为极小化下列滚动优化性能指标^[5]:

$$\min_{\Delta u(t+i)} J(\mathbf{u}) = \sum_{i=1}^P \|\mathbf{y}_r(t+i) - \mathbf{y}_m(t+i|t)\|_{\mathbf{R}}^2 + \sum_{i=0}^M \|\Delta \mathbf{u}(t+i)\|_{\mathbf{Q}}^2 \quad (9)$$

且满足如下约束条件:

$$\mathbf{u}_{\min} \leq |\mathbf{u}| \leq \mathbf{u}_{\max}, \quad \Delta \mathbf{u}_{\min} \leq |\Delta \mathbf{u}| \leq \Delta \mathbf{u}_{\max}, \quad \mathbf{y}_{\min} \leq |\mathbf{y}| \leq \mathbf{y}_{\max} \quad (10)$$

其中 y_r 为输出参考轨迹, $y_m(t+i|t)$ 为 t 时刻预测未来 $t+i$ 时刻模型输出向量, 即为 $\tilde{Y}(t+i)$, R 、 Q 为加权矩阵, 且都为正定矩阵, P 、 M 分别为预报时域和控制时域.

当考虑复杂过程中由于操作条件和物理限制, 对控制量 u 和输出 y 引入约束后, 由于多神经网络预测模型是一个非线性模型, 使得滚动优化成为非线性优化问题. 该优化问题可行的方法是非线性规划法, 而序贯二次规划法 (SQP) 是比较常用且有效的方法^[5,8]. SQP 将非线性规划问题的求解转化为一系列二次规划问题进行求解. 对于非线性约束优化问题, 由于约束条件的限制, 当考虑搜索方向和搜索步长后, 优化问题往往通过较少的迭代过程就可以得到解决, 因此具有较好的实用性.

由此可得到基于 MNN 模型预测控制的基本算法, 为在 t 时刻根据 MNN 模型预报系统输出, 采用 SQP 算法求解非线性优化问题, 得到控制序列 $u(t)$, 施加于系统进行控制, 在 $t+1$ 时刻根据当前的系统实际输出值修正预测模型的输出, 再重新求解非线性优化问题.

4 典型 CSTR 对象的仿真分析

考虑如图 2 所示的典型 pH CSTR 非线性对象, Li & Biegler^[8] 推导出其动态方程为:

$$\begin{aligned} dx_1/dt &= u_1 + u_2 - 0.2 x_1^{1/2} \\ dx_2/dt &= (24.9 - x_2) u_1 x_1^{-1} + (0.1 - x_2) u_2 x_1^{-1} - x_2/(1 + x_2)^2 \\ y_1 &= x_1 \quad y_2 = x_2 \end{aligned} \quad (11)$$

其中输出变量 y_1 和 y_2 为反应器液位 h 和浓度 C_B , 输入变量 u_1 和 u_2 为流量 F_1 和 F_2 , 其约束条件为

$$\begin{aligned} 0 < u_1, u_2 < 1, 0 < |\Delta u_1| < 0.3, 0 < |\Delta u_2| < 0.5 \\ 0 < y_1 < 100, 0 < y_2 < 14 \end{aligned}$$

pH CSTR 对象在 pH=7 附近具有很强的非线性特性, 为了充分激励对象的动态特性, 采用伪随机序列 (PRBS) 信号作为输入变量 F_1 和 F_2 的信号, 得到 2000 组训练样本数据, 其中 F_1 和 F_2 在 0~1 之间随机变化, pH 值则在 2 到 12 之间变化. 根据样本数据, 采用基于广义信息熵融合策略的方法建立多神经网络模型, 构成预测模型. 首先以样本输出为基准, 采用 K-means 聚类分析算法对样本数据进行分类, 用样本子集分别训练子网络. 在此采用多层前向网络构成子网络 NN_i , 拓扑结构选用了 (4-10-2) 和 (4-20-2) 两种结构, 分别采用改进的 L-M 算法和 PRPE 算法^[3,4] 来训练, 这样训练得到了 12 个子网络模型. 采用广义信息熵融合方法, 优选了其中 3 个子网络构成 MNN 模型, 为表 1 中有 * 标记的子网络, 而且整个 MNN 模型的可靠性大于 90%. 若任意选用其他子网络得到 MNN 模型的可靠性都小于 90%. 由此可知, 基于广义信息熵融合方法的 MNN 模型具有更好的可

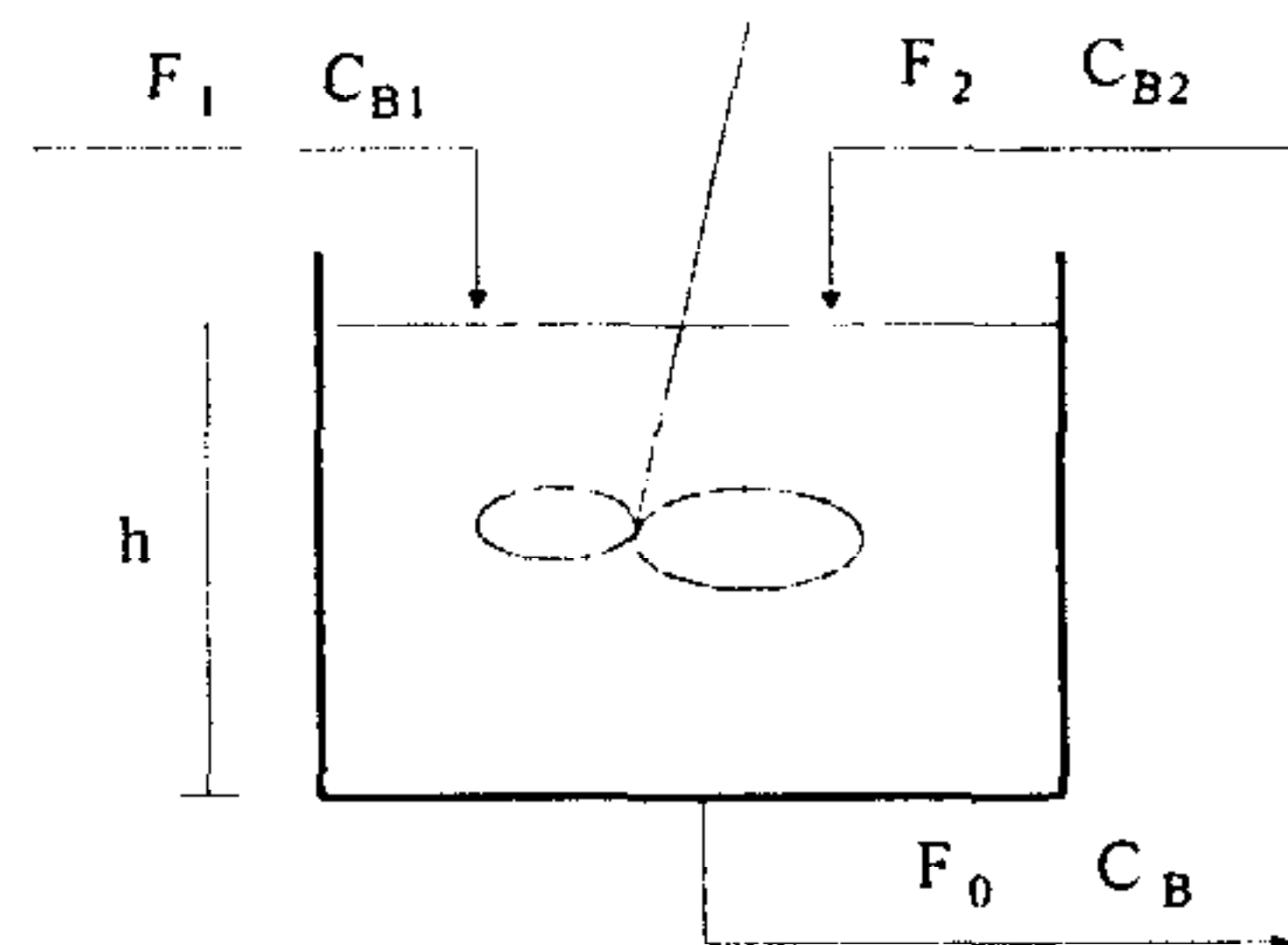


图 2 pH CSTR 对象

表 1 子网络模型 NN_i 的广义信息熵值 $h(\text{nat})$

广义信息熵 h	L-M 算法		PRPE 算法	
	4-10-2	4-20-2	4-10-2	4-20-2
样本子集 1	0.215	0.161	0.132*	0.203
样本子集 2	0.174	0.133*	0.156	0.161
样本子集 3	0.238	0.205	0.246	0.179*
MNN	0.1482		(>90%)	

靠性, 其不确定性也较小.

在选取子网络模型后, 选择 $V=0.48$, $W_i(0)=\text{diag}(1/3,1/3)$, 根据式(3)计算连接权值矩阵, 从而得到多神经网络预测模型. 在此基础上采用预测控制算法对系统实施控制, 其仿真结果如图 3 所示, 其中将基于多神经网络的预测控制结果与基于单神经网络的预测控制结果进行了比较. 从中可知, 两种预测控制方法都能较好地跟踪设定值的变化, 但是由于多神经网络预测模型具有更高的预报精度, 控制性能较好, 响应速度和稳定情况均优于单个神经网络模型.

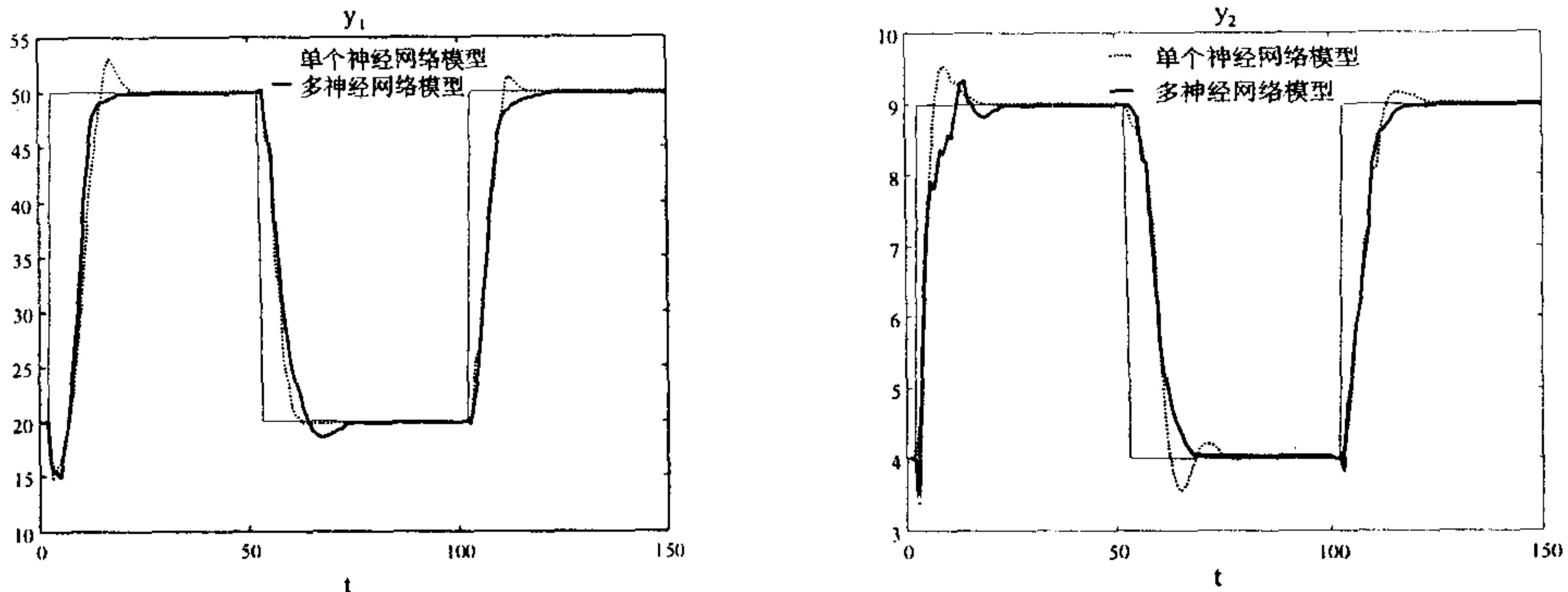


图 3 多神经网络与单个神经网络的预测控制结果中输出变量间的比较

5 结论

本文提出了一种基于多神经网络模型的非线性预测控制方法. 其中多神经网络模型作为预测模型, 通过采用基于广义信息熵的融合方法来优选子网络, 降低了多神经网络预测模型的不确定性, 提高了其准确性和可靠性. 而且实时地递推计算子网络输出的连接权值矩阵, 以适应系统在不同工作点区域内变化. 对于 CSTR 对象的仿真结果表明了该方法的有效性.

参 考 文 献

- 1 席裕庚, 王凡. 非线性系统预测控制的多模型方法. 自动化学报, 1996, 22(4): 456-461
- 2 陈小红, 钱积新. 基于 RBFN 及模糊决策的多模型直接自适应控制. 控制理论与应用, 1999, 16(1):100-104
- 3 熊智华, 王雄, 徐用懋. 基于广义信息熵融合的多神经网络建模方法. 信息与控制, 1999, 28(增刊):491-495
- 4 Chen S., Billings S. A. Neural networks for nonlinear dynamic system modeling and identification. *Int. J. Control*, 1992, 56(2): 319-346
- 5 H. Zhao, J. Guiver, R. Neelakantan, et al. A nonlinear industrial model predictive controller using integrated PLS and neural net state space model. *IFAC'99, Beijing, 1999, Vol N 13-18*
- 6 常迥. 信息理论基础. 北京: 清华大学出版社, 1993
- 7 He W G, Kaufman H., Roy R. Multiple model adaptive control procedure for blood pressure control. *IEEE Trans. on BME*, 1986, 33(1):10-19
- 8 Wei Chong Li, Lorenz T Biegler. Process control strategies for constrained nonlinear systems. *Ind. Eng. Chem. Res.*, 1988, 27: 1421-1433