

基于过程神经元网络模型的间歇反应优化计算

蔺淑倩 王晶* 曹柳林

(北京化工大学 信息科学与技术学院, 北京 100029)

摘要: 针对间歇过程中最优操作轨线经常从生产经验中人工获取的问题,提出了一种基于过程神经网络(process neural network, PNN)模型的自动计算方法。利用 PNN 独特的时间聚合算子特点,建立反应操作变量与产品质量之间的关系。基于此时间函数模型,通过产品质量性能指标优化计算,获得最佳的控制轨线,作为实际生产的设定曲线。最后,MPCE 仿真平台的间歇反应实验说明了该方法的有效性。

关键词: PNN 模型; 控制轨线优化计算; MPCE 仿真平台

中图分类号: TP273

引言

间歇反应过程在助剂、制药和染料行业生产过程中普遍存在,具有典型的代表意义^[1]。间歇反应运行时间有限,没有稳定的工作点,且运行具有一定的重复性等,其优化问题通常是以提高产品的质量、产量或缩短运行时间等为目标,获得最佳操作轨线,进而跟踪最佳操作轨线实施最优控制。在实际工业生产中,采用最优温度控制曲线作为最佳操作轨线尤为普遍。因此研究如何获取间歇反应过程的最优温度控制曲线显得尤为有意义^[2-5]。

多数工厂实际生产中,通常是工人师傅根据自己多年积累的经验找到一条针对某个具体间歇反应的温度控制曲线。这种依靠经验寻找温度控制曲线的方面明显费时费力,因此希望找到一种更简捷、快速的数学方法来获取最佳的温度控制曲线。近年我国何新贵院士等人提出过程神经元网络这个概念,为寻找最优温度控制曲线提供了理论基础^[6-7]。

本文首次将过程神经元网络(PNN)建模方法用于间歇过程的最优控制曲线的计算中,其原因有两点。一是过程神经元网络,它克服了传统神经网络对输入的限制,输入和权值均可以为时变函数(或称过程),是人工神经网络在时域上的一种扩

展,从物理真实性上来看,更符合间歇反应过程对操作曲线的要求,即模型结构与实际过程要求一致。此外,间歇过程非线性严重,有限时间通常存在缓慢升温(反应诱发)、快速升温(反应剧烈)和恒温(反应结束)三种不同的状态;而且间歇过程关注反应结束时的产品质量,要求模型必须具有很好的长效预报能力,从这两个特点来看,PNN 比传统神经网络的建模效果会更好。

1 温度控制曲线的优化计算

1.1 间歇反应的优化问题

间歇生产过程的最优化问题所要达到的目标并不着重于最佳条件(即设定值)的选择,或者间歇反应过程的中间状态值的优良与否,而在于间歇生产过程结束时间歇单元是否获得最大的产品产率或是产品产量。经典的间歇反应优化问题可表述为

$$\min_{u(t)} J = \phi(x(t_f))$$

其中, $u(t)$ 为控制变量, t_f 为终点时刻, $\phi(x)$ 为优化目标函数,若只考虑过程结束时的产品质量要求,通常可以用下式表示

$$\phi(t) = \int_{t_0}^{t_f} [f(u(t)) - g(t_f)]^2 dt$$

该指标是关于产品质量函数 $f(u(t))$ 与给定的过程结束时($t=t_f$)的质量指标 $g(t_f)$ 之间的距离度量,优化控制的目的在于选取最优的温度操作曲线 $u(t)$,使得聚合反应结束时的产品质量尽可能接近给定的质量指标。

对间歇过程温度进行优化控制的第 1 步就是建立沿时间分布的反应温度与产品质量之间的关系模

收稿日期: 2009-06-25

基金项目: 国家自然科学基金(60704011/60974031)

第一作者: 女,1982年生,硕士生

* 通讯联系人

E-mail: jwang@mail.buct.edu.cn

型 $f(u(t))$ 。下面具体说明反应温度与产品质量之间的关系模型的特殊性。如图 1 所示,在每一批次的生产中,系统的输入是整个反应过程中沿时间的温度函数 $u(t)$,输出是过程结束时的产品质量(如转化率),可以被认为是依赖于输入时变过程的泛函数。

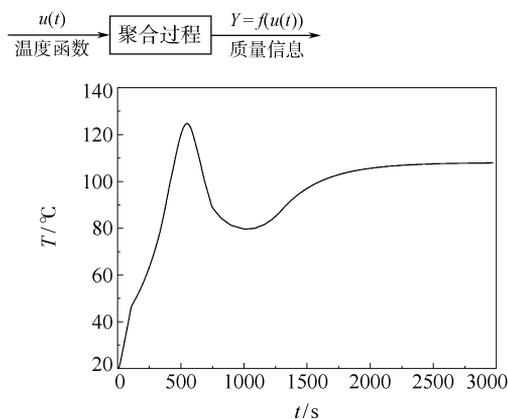


图 1 间歇反应温度分布与产品质量模型示意图
Fig. 1 Temperature distribution of batch reaction and product quality model

间歇反应过程具有较大的时变性、不确定性和高度非线性,很难建立精确的机理模型,而过程神经网络本质上是一类泛函模型,与此过程完全相符,因此采用 PNN 对上述过程进行建模。

1.2 过程神经元建模

如图 2 所示的过程神经网络为 3 层结构(输入层、隐含层、输出层),中间隐含层为过程神经元,包含加权、聚合和激励 3 部分^[3]。

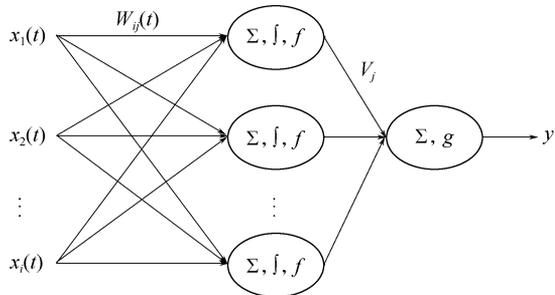


图 2 含一个隐含层的过程神经网络
Fig. 2 Configuration of the process neural network configuration

过程神经网络的输入函数一般有两种类型:
①基于系统理论模型或基于统计规律的解析函数,这一类可以直接根据固定公式进行傅立叶基函数变换;
②通过实验采样得到的与时间相关的离散数据序列,这时需要先将离散数据序列拟合成连续函数形式,再作傅立叶基函数变换。

利用过程神经网络建模具体如下。

第 1 步:连续输入函数的基函数展开

首先假设过程神经网络的输入空间为 $[0, 2\pi]$,在此空间内将连续的输入函数 $f(t)$ 按标准正交基

$$M = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos t, \frac{1}{\sqrt{\pi}}\sin t, \dots, \frac{1}{\sqrt{\pi}}\cos nt, \frac{1}{\sqrt{\pi}}\sin nt, \dots \right\}$$

展开成如下有限项形式

$$f(t) = \sum_{i=0}^L a_i e_i(t) = \frac{1}{\sqrt{2\pi}}a_0 + \frac{a_1}{\sqrt{\pi}}\sin t + \frac{a_2}{\sqrt{\pi}}\cos t + \dots, t \in [0, 2\pi] \quad (1)$$

其中, $e_i(t) \in M, i=0, 1, 2, 3, \dots$

第 2 步:初始化权值和阈值

权值的初始化与传统神经网络权值有一定区别,其中网络连接权函数 $W_{ij}(t)$,也应该表示为同组基函数的有限项展开形式

$$W_{ij}(t) = \sum_{l=0}^L w_{ij}^l e_l(t) \quad (2)$$

由于输出层神经元是非时变的普通神经元,因此权值 $V_j(t)$ 可以取 $[0, 1]$ 之间的随机数。隐含层神经元和输出层神经元阈值初始化方法与 $V_j(t)$ 类似。

第 3 步:输入函数系数的归一化

由于输入函数和权值函数在傅立叶正交基展开下的系数差别很大,而网络中的各神经元均采用 S 型函数作为激励函数,因此为了避免神经元输出陷入饱和,需要对输入函数系数作归一化预处理。在 MATLAB 软件中有归一化处理的函数 premax(), 此函数将系数归一化到 $[-1, 1]$ 之间。

第 4 步:计算网络输出和误差

针对隐含层为 n 个神经元,输出层为 m 个神经元的 3 层神经网络的输入输出映射关系可表示为

$$y = g \left\{ \sum_{j=1}^m v_j f \left(\sum_{i=1}^n \left[\int_0^T (W_{ij}(t)x_i(t)) dt - \theta_j \right] \right) - \theta \right\} \quad (3)$$

将输入函数和权值函数的基展开形式带入上式,由于基底 $e_k(t)$ 的完备正交性,上式可简化如下

$$y = g \left\{ \sum_{j=1}^m v_j f \left[\sum_{i=1}^n \sum_{l=1}^L a_{il} w_{ij} - \theta_j \right] - \theta \right\} \quad (4)$$

其中, θ_j 为隐含层节点 j 的阈值, θ 为输出节点阈值。 L 是按函数拟合精度要求需要展开的基底个数。假设给定 K 个样本,则误差函数可表示为

$$E = \sum_{k=1}^K (y_k - d_k)^2 = \sum_{k=1}^K \left(g \left(\sum_{j=1}^m v_j f \left(\sum_{i=1}^n \sum_{l=1}^L a_{il}^k w_{ij}^l - \theta_j \right) - \theta \right) - d_k \right)^2 \quad (5)$$

式(5)中 d_k 表示第 k 次训练的期望输出。

第 5 步:调整网络权值和阈值

采用梯度下降法的学习算法,取激励函数 f, g 均为 S 型函数,则网络权值和阈值的学习规则为

$$v_j = v_j + \alpha \Delta v_j \quad w_{ij}^l = w_{ij}^l + \beta \Delta w_{ij}^l \\ \theta = \theta + \eta \Delta \theta \quad \theta_j = \theta_j + \gamma \Delta \theta_j$$

$$\Delta v_j = -2 \sum_{k=1}^K [(g(z_k) - d_k) g'(z_k) f(u_{kj})]$$

$$\Delta w_{ij}^l = -2 \sum_{k=1}^K [(g(z_k) - d_k) g'(z_k) v_j f'(u_{kj}) a_{il}^k]$$

$$\Delta \theta_j = -2 \sum_{k=1}^K [(g(z_k) - d_k) g'(z_k) v_j f'(u_{kj}) (-1)]$$

$$\Delta \theta = -2 \sum_{k=1}^K [(g(z_k) - d_k) g'(z_k) (-1)] \quad (6)$$

其中 $\alpha, \beta, \gamma, \eta$ 为学习速率, u_{kj} 和 z_k 定义为

$$u_{kj} = \sum_{i=1}^n \sum_{l=1}^L a_{il}^k w_{ij}^l - \theta_j$$

$$z_k = \sum_{j=1}^m v_j f \left(\sum_{i=1}^n \sum_{l=1}^L a_{il}^k w_{ij}^l - \theta \right)$$

第 6 步:循环结束判断

如果训练次数 $N > N_{\max}$, 则表明训练结束,输出训练结果,否则循环次数加 1,返回第 4 步继续训练。训练结束后保存网络权值和阈值^[8]。

2 生产橡胶硫化促进剂的温度控制仿真实验

橡胶硫化促进剂的生产是比较复杂的间歇反应过程,其温度控制包括 3 个阶段:加热升温阶段,过渡阶段和保温恒定阶段。以往的实际生产中,温度控制曲线由反复实际生产累积经验中获得。现在希望通过建模寻优方法,找到一条满足要求的温度控制曲线。

2.1 仿真实验平台简介

在 MPCE-1000 的教学仿真平台上对橡胶硫化促进剂进行仿真实验。如图 3 所示,平台由 3 部分组成。

(1) 盘台。包括过程设备、阀门、按钮开关和插接孔。

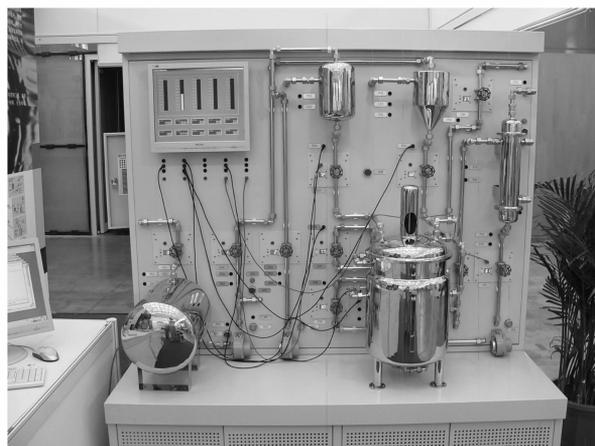


图 3 MPCE-1000 仿真实验平台

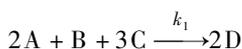
Fig. 3 The MPCE-1000 simulation platform

(2) 上位机。装有 MPCE Studio 控制组态软件,通过 VB 编程与该组态软件相联,实现对仿真对象的数据处理和先进控制。

(3) 下位机。是一台工控机,内置仿真模型计算包,模拟真实系统数据并输出。上下位机之间通过网线进行数据通讯。

该过程属于间歇缩合反应,反应装置包括两台高位计量罐和一个搅拌釜式反应器。主控制阀为蛇管冷却水阀,操纵变量为反应器蛇管冷却水入口流量。这里夹套冷却水阀和夹套冷却水入口流量不作为主要控制手段,在仿真实验中均为零。对生成物浓度影响最大的反应釜温度作为被控变量。

反应釜内的主反应表达式为^[10]



其中, A、B、C 分别代表原料邻硝基氯苯($C_6H_4ClNO_2$)、硫化钠(Na_2S)及二硫化碳(CS_2), D 代表主产物橡胶硫化促进剂, k_1 代表反应速率常数。根据表 1 每釜反应原料用量以及主反应表达式可以计算得到每釜主产物的理想产量是 283.5 kg, 每个批次结束时的测量得到的主产物量除以理想产量即是该批次反应的转化率值,作为产品质量指标。

表 1 每釜反应原料用量

Table 1 The dosage of raw material consumed per reactor

序号	名称	纯物质量/kg	工业用量/L
1	多硫化钠	942	1800
2	邻硝基氯苯	237	185
3	二硫化碳	140	125

2.2 建模实验结果

建立橡胶硫化促进剂间歇反应的过程神经网络

模型,必须首先获得反应的实验数据。在仿真平台的上位机上进行仿真实验。参数设定如下:1 个批次的反应时间为 2970 s,采样间隔为 10 s。通过 16 次的仿真实验,得到 16 组实验数据(其中 1 组如图 4 所示)。然后利用 VB 程序导出实验数据,对每组实验数据逐一进行拟合,傅立叶展开,以及归一化处理。

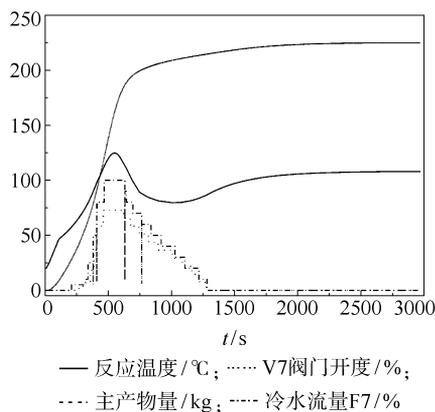


图 4 样本数据示例
Fig. 4 Data samples

因为输入数据(温度)是通过实验采样得到的与时间相关的离散数据序列,因此需要先将其拟合成连续函数形式。多次试验表明,在采样周期内将其拟合成 12 阶多项式,完全满足拟合精度要求。然后以 201 个基底展开,得到相应的傅立叶级数。抽取各项傅立叶级数的系数值,归一化到 $[-1, 1]$ 之间,作为训练网络的输入数据。同时将每个批次终点时刻得到的产品质量转换成产品的转化率,再将转化率数值归一化到 $[0, 1]$ 之间,作为网络的输出数据。但如果输入是基于实际或者是基于统计规律的时间解析函数,则可以直接进行傅立叶展开而省略拟合的步骤,这将使 PNN 建模更为简化。

过程神经网络结构选择 1-25-1,激励函数均为 S 型函数。为了简化计算,阈值均设为零,选择前八组实验数据作为训练样本,后八组数据作为泛化样本,训练网络过程如 1.2 节所述,训练 8532 次后退出循环。从图 5 可以看出,过程神经网络的训练输出和泛化输出与各自样本的期望输出都基本吻合。网络训练和泛化的平均方差如图 6 所示。

当调整网络结构参数,例如隐含层神经元个数,学习速率等将得到不同的训练结果,其中不同的隐含层神经元个数所构成的网络训练结果比较如表 2。经过反复试验比较,选择 25 个神经元作为隐含

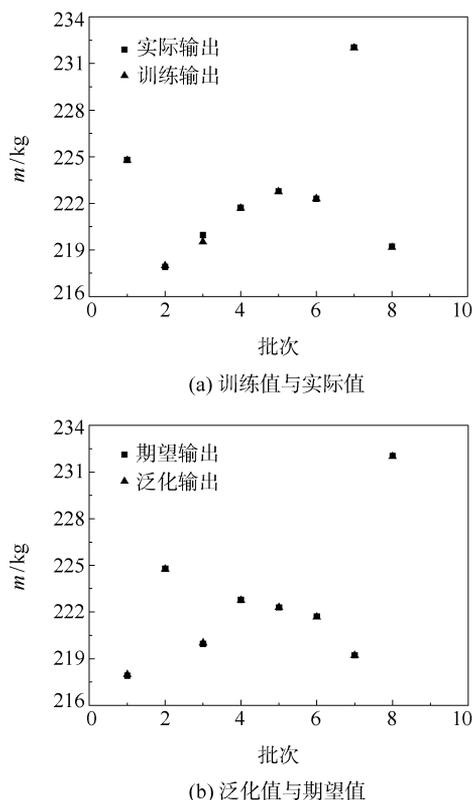


图 5 过程神经网络中训练值与实际值(a),泛化值与期望值(b)比较
Fig. 5 Comparisons of training data with actual data (a), testing data with desire data (b)

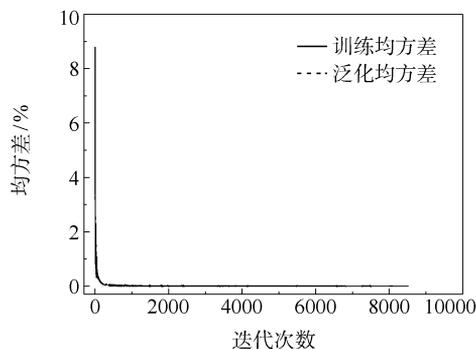


图 6 网络训练与泛化的均方差比较
Fig. 6 Comparison of mean square errors for training and testing

表 2 不同隐含层数目的网络训练结果对比
Table 2 Comparison of training results for networks that have different numbers of neurons in hidden layers

序号	隐含层节点数/个	训练最大误差/%	迭代次数
1	12	0.9919	9672
2	25	0.3352	8532
3	26	0.432	7685

层,学习速率设定为 0.85 完全可以满足训练精度要求,且能得到相对较好的训练结果。

分析仿真建模数据可知(见图 4),此类间歇过程非线性严重,温度从室温(20℃)在短时间内快速上升到近 120℃左右之后,为得到较高转化率必须长时间维持在此温度,使得恒温数据在总体实验数据中过于集中且数量偏多。采用传统的 BP 神经网络很难对此类实验数据进行辨识,多次的仿真实验也验证了这一点。而过程神经网络的输入本质上是温度曲线傅立叶级数展开后的各系数值,避免了数据的单一性,因此能够得到很好的建模结果。

2.3 优化结果

本实验中,由于反应装置、工艺条件等客观因素限制,促进剂最大转化率能达到 80%左右,因此产率优化目标为^[9]

$$\min J = \int_{t_0}^{t_f} [f(u(t)) - 0.8]^2 dt$$

基于上述过程神经网络转化率模型 $f(u(t))$,在无约束条件下求解得到转化率接近 80%的最佳温度控制曲线,如图 7 所示。

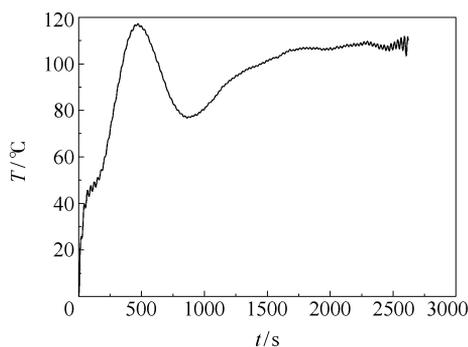


图 7 最优温度控制曲线图

Fig. 7 The optimal temperature control trajectory

事实上,由于间歇反应关注的是终点时刻的产品质量,对中间过程不做要求,因此上述优化问题在不同的模型结构下,可能得到不同的优化解。从图 7 给出的优化曲线可以看出,温度变化经历 3 个阶段:初始平稳升温,从室温上升到 65℃左右,且不需要冷却水控制,让温度上升速度维持在 0.1~0.2℃/s 之间,诱导反应深度进行;中间冷却控制阶段,也称过渡阶段,温度继续升高使反应剧烈进行,此时加入冷却水控制反应釜温度上升过快,防止反应釜超压;最后恒温保持,促使反应充分进行,提高主产物产量。这种基于过程神经网络模型优化计算得到的最佳操作曲线完全符合工艺操作要求,是一种实

际可行的优化解。

最后,将求解得到的最优温度曲线作为温度跟踪控制给定曲线,采用 PID 控制方案,进行橡胶硫化促进剂制备的温度控制实验,最终得到的主产物产品转化率为 79.963%,非常接近优化目标值。

3 结论

首次利用过程神经网络的时空二维信息处理能力,有效解决了间歇反应过程中,输入输出与时间有关的建模难点。本文提出的面向过程数据的操作曲线计算方法,为解决实际生产领域的控制难题提供了新的思路,从而为过程神经网络的应用提供了更为广阔的前景。

参考文献:

- [1] 陈治纲, 许超, 邵惠鹤. 间歇过程优化与先进控制综述[J]. 化工自动化及仪表, 2003, 30(3): 1-6.
Chen Z G, Xu C, Shao H H. Summary of batch process optimization and advanced control[J]. Control and Instruments in Chemical Industry, 2003, 30(3): 1-6. (in Chinese)
- [2] Tian Y, Zhang J, Morris J. Optimal control of a batch emulsion copolymerisation reactor based on recurrent neural network models[J]. Chemical Engineering and Processing, 2002, 41: 531-538.
- [3] Zhang J. Batch-to-batch optimal control of a batch polymerisation process based on stacked neural network models[J]. Chemical Engineering Science, 2008, 63: 1273-1281.
- [4] Xiong Z H, Zhang J. A batch-to-batch iterative optimal control strategy based on recurrent neural network models[J]. Journal of Process Control, 2005, 15: 11-21.
- [5] Xiong Z H, Zhang J. Modelling and optimal control of fed-batch processes using a novel control affine feedforward neural network[J]. Neurocomputing, 2004, 61: 317-337.
- [6] 何新贵, 许少华. 输入输出均为时变函数的过程神经网络及应用[J]. 软件学报, 2003, 14(4): 764-769.
He X G, Xu S H. Process neural network with time-varied input and output functions and its applications[J]. Journal of Software, 2003, 14(4): 764-769. (in Chinese)
- [7] 何新贵, 梁久祯. 过程神经网络的若干理论问题[J]. 中国工程科学, 2000, 2(12): 40-44.
He X G, Liang J Z. A number of theoretical issues of process neural network[J]. China Engineering Science,

- 2000, 2(12): 40 - 44. (in Chinese)
- [8] 许少华, 蔡月芹, 谢树民. 基于傅立叶函数基变换的过程神经网络学习算法[J]. 计算机工程与设计, 2006, 27(23): 4413 - 4415.
- Xu S H, Cai Y Q, Xie S M. Learning algorithm of procedure neural networ based on orthogonal base function[J]. Computer Engineering and Design, 2006, 27(23): 4413 - 4415. (in Chinese)
- [9] 苏金明. MATLAB6.1 实用指南[M]. 北京: 电子工业出版社, 2002.
- Su J M. MATLAB6.1 practical guide[M]. Beijing: Electronic Industry Press, 2002. (in Chinese)
- [10] 吴重光. 过程系统仿真技术[M]. 北京: 中国石化出版社, 1997.
- Wu C G. Process system simulation technology[M]. Beijing: China Petrochemical Press, 1997. (in Chinese)

Optimization of a batch reactor based on a process neural network

LIN ShuQian WANG Jing CAO LiuLin

(College of Information Science & Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: The optimal operation trajectory of a batch process is always determined by experienced workers, based on their rich experience. Obviously this method is time-consuming and laborious. In order to overcome this problem, a simple and fast mathematical method is given based on a process neural network (PNN) model. A PNN is used to model batch processes in order to overcome the restrictions on instantaneous synchronization input for a traditional neural network. Based on the PNN, a quadratic form performance index function is presented which allows the optimal operating trajectory to be obtained. The proposed scheme is successfully applied to a simulated batch process.

Key words: PNN model; optimization of operation trajectory; MPCE simulation platform