

文章编号: 1001-0920(2005)03-0261-05

模糊神经网络模型混沌混合优化学习算法及应用

秦斌^{1,2}, 吴敏¹, 王欣²

(1. 中南大学 信息科学与工程学院, 湖南 长沙 410083; 2 株洲工学院 电气系, 湖南 株洲 412008)

摘要: 基于混沌优化的思想, 提出一种新的模糊模型的优化学习算法。将模糊推理规则转化为模糊RBF网络模型, 用模糊C均值(FCM)聚类算法和分区效验熵得到模型结构, 用混沌变换序列寻优得到优化的中心初值群, 用FCM获得最优聚类中心, 最后获得模糊神经网络模型。将该方法应用于转炉终点磷含量预报模型, 取得了较好的结果。

关键词: 模糊模型; 径向基网络; 模糊聚类; 混合混沌优化; 终点磷含量

中图分类号: TP183;O224

文献标识码: A

Hybrid chaos optimization algorithm for fuzzy neural network model and its applications

QIN Bin^{1,2}, WU Min¹, WANG Xin²

(1. College of Information Science and Engineering, Central South University, Changsha 410083, China;

2 Department of Electric, Zhuzhou Institute of Technology, Zhuzhou 412008, China Correspondent: Q N Bin,

E-mail: qinbin99@163.com)

Abstract: Based on the idea of chaos optimization, an optimization algorithm for the fuzzy model is presented. The fuzzy model can be represented as a fuzzy RBF neural network model. The structure of the model is determined using the FCM algorithm and the clustering validity criteria. The initial parameter of clustering centers is obtained using synthetical chaos series and is further optimized using the FCM algorithm. The proposed approach is used successfully for the prediction of end phosphorus content in converter.

Key words: fuzzy model; RBF neural network; fuzzy clustering; hybrid chaos optimization; end phosphorus content

1 引言

近年来, 模糊模型在系统识别、数据挖掘、控制和决策中得到了广泛的应用, 建立系统的模糊模型主要分为两个步骤: 1) 模型结构识别; 2) 模型参数优化。模型结构识别将前提空间映射到模糊区间; 参数优化是调整前提和结论参数, 使指定的性能指标最小。利用模糊推理系统与RBF神经网络在一定条件下的等价性, 可将所求问题转化为神经网络的逼近问题。RBF网络的性能依赖于基函数的个数、位置(即中心向量)及其形状。中心向量的学习算法主要分为遗传类算法和非遗传类算法。非遗传类算法

主要有C均值、模糊C均值(FCM)、Kohonen聚类、学习向量量化(LVQ)、误差回传(BP)算法^[1-3]等; 网络结构的学习方法主要有结构修剪法和结构增长法。C均值和Kohonen聚类算法都是启发式过程, 它们对初始值的选取相当敏感^[1]。这不仅影响了网络的模式分类能力, 而且制约了网络的回归能力。

为提高网络的回归能力, 可采用类似于MLP的BP算法来训练网络。然而BP算法的计算复杂度较高, 容易陷入局部最小, 并且忽视了RBF网络所固有的特点。混沌是自然界中一种较为普遍的现象, 它具有随机性、遍历性和规律性等特点, 在一定范围内

收稿日期: 2004-04-19; 修回日期: 2004-07-30

基金项目: 国家863计划基金项目(2001AA11040); 湖南省自然科学基金项目(01JJY3029); 湖南省教育厅基金项目(04G718)。

作者简介: 秦斌(1963—), 男, 湖南永洲人, 教授, 博士生, 从事复杂工业过程建模、智能控制等研究。吴敏(1963—), 男, 广东化州人, 教授, 博士生导师, 从事智能控制、鲁棒控制等研究。

能按自身的规律不重复地遍历所有状态 混沌优化方法在搜索空间不大时效果显著,但当搜索空间大时其效果却不能令人满意^[4].

本文基于混沌优化的思想,提出了模糊模型的一种新的学习算法 将模糊模型转化为RBF网络,根据FCM和分区熵确定最佳规则数,中心向量的学习规则采用FCM与混沌优化的混合算法 根据所选混沌映射的概率分布,提出一种基于混沌的合成序列 混合混沌优化过程分为两个阶段进行:1)在变量的取值范围内依次遍历经过的各点,以进化的思想保留较好点作为候选最优;2)以候选最优群为中心,采用FCM优化算法进行细搜索寻找最优,采用最小二乘法优化权值 与其他聚类算法相比,该算法可消除对初始值选取的敏感性,避免陷入局部最小,充分利用网络隐层的聚类功能,改善网络的回归能力

2 模型结构

改进的T-S模糊推理规则定义为

$$R_i: \text{If } X_i \text{ is } F^i \text{ Then } y = G_i(X). \quad (1)$$

其中: X 为模型的输入向量,由输入变量构成; y 为模型的输出变量; $G_i(X)$ 表示规则 R_i 的局域函数,可取单值、模糊集和线性函数

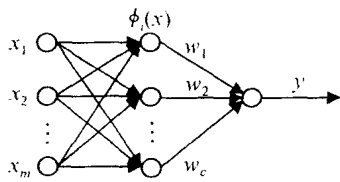


图1 RBF神经网络模型

当采用单值模糊化、中心平均反模糊化和 T 范数运算时,由式(1)定义的模糊推理系统在满足一定条件时^[5],可用图1所示的分层归一化RBF网络系统来实现 RBF网络是一种品质优良的网络,已证明它可在任意精度下逼近任意的非线性函数,且不存在局部最小问题 正则化RBF神经网络由单隐层前向网络组成,其中隐层单元输出 $R_i(x)$ 定义为

$$R_i(x) = \exp\left[-\frac{(x - c_i)^2}{\sigma_i^2}\right]. \quad (2)$$

模糊推理的输出为

$$y = \frac{\sum_{i=1}^c G_i \mu_F(X)}{\sum_{i=1}^c \mu_F(X)} = \frac{\sum_{i=1}^c y_i \mu_F(X)}{\sum_{i=1}^c \mu_F(X)} = \frac{\sum_{i=1}^c y_i R_i(x)}{\sum_{i=1}^c R_i(x)} = \sum_{i=1}^c w_i \phi_i \quad (3)$$

因此,可将改进的T-S模糊推理系统规则数和模糊子域划分转换为:1)确定分层模糊RBF隐层节点的数量,也就是模糊规则数;2)隐层中心向量和宽度;

3)输入隶属函数和输出权重 其中中心向量学习是关键

3 模糊RBF神经网络学习算法

3.1 由模糊聚类确定模糊规则数

模糊 C 均值聚类是用隶属度确定每个数据点属于某个聚类的方法 构造样本集 $P_j = \{x_{1,j}, x_{2,j}, \dots, x_{m,j}, y_j\}$, FCM 将 n 个向量 $P_j (j = 1, 2, \dots, n)$ 分为 C 个模糊组 设样本 P_j 属于模糊集合的隶属度为 u_{ij} ,按归一化规定,有

$$\sum_{i=1}^C u_{ij} = 1, \forall j = 1, 2, \dots, n. \quad (4)$$

为衡量聚类的质量,常采用非相似性能指标为聚类准则函数,即

$$J(P, V) = \sum_{i=1}^C J_i = \sum_{j=1}^n \sum_{k=1}^m u_{ij}^m d_{ik}^2, \quad (5)$$

其中 d_{ik} 表示第 k 个样本到第 i 个中心点的欧氏距离 由于式(5)是有关 C 的单调函数,引入改进的Bezdek分区熵^[6]来衡量FCM聚类数指标

$$H(U, C) = -\frac{1}{n} \sum_{j=1}^n \max_{i \in C} (u_{ij}) \log(\max_{i \in C} (u_{ij})). \quad (6)$$

聚类中心向量一旦确定,宽度 σ_i 便可通过最接近的 p 个中心距离的平均值来计算,使与 σ_i 有关的目标函数极小化,即

$$E\sigma = \frac{1}{2} \sum_{i=1}^p \left[v_j - v_{j, \text{nearest}} \left(\frac{v_i - v_j}{\sigma_i} \right) - r \right]^2, \quad (7)$$

其中 r 为超调常数

由FCM算法确定模糊规则数的步骤如下^[11]:

Step 1: 确定聚类最大类别数 C , 预置允许误差 ϵ , 初始化最优中心点 v_i^{*F} 和最优指标 J^{*F} ; 初始化中心点 $v_{i,0} = v_i^{*F}$, 选择迭代变量 $t = 1$ 和迭代上限 t_{\max} .

Step 2: 取 $c = 2, 3, \dots, C$.

Step 3: 对于 $t = 1, 2, \dots, t_{\max}$:

1) 计算隶属度

$$m_t = m_0 - t\Delta m, \quad (8)$$

$$u_{ik,t} = \frac{1}{\sum_{l=1}^c \left(\frac{d_{lk}}{d_{ik}} \right)^{\frac{2}{m_t-1}}}. \quad (9)$$

其中 m_0 是大于1的常数,且 $m = 1$; d_{ik} 表示第 k 个样本到第 i 个中心点的欧氏距离

2) 按下式更新中心向量:

$$v_{i,t} = \frac{\sum_{j=1}^n u_{ij}^{m_t} P_j}{\sum_{j=1}^n u_{ij}^{m_t}} \quad (10)$$

Step 4: 计算

$$E = |v_{i,t} - v_{i,t-1}|^2, \quad (11)$$

若 $E < \epsilon$ 或达到最大迭代上限,则计算 $J(P, v_{i,t})$; 若 $J(P, v_{i,t}) < J^{*F}$, 则 $v^{*F} = v_{i,t}$; 若 $J^{*F} = J(P, v_{i,t})$,

则计算各聚类的 σ ; 否则, 重复 Step 3

Step 5: 计算 $H(U, c)$, 如果 $H(c-1) = H(c)$, 并且 $H(c-1) = H(c-2)$, 则停止, 确定模糊规则数为 $c-1$; 否则, 转 Step 2

经过上述迭代运算, 可确定隐含节点数和初步聚类中心向量 将聚类中心向量 V 分成 X 和 Y 两部分, 则为输入聚类权向量和输出权向量 式(9)和(10)是必要条件而非充分条件, 不能确保 FCM 收敛于最优解, 算法的性能依赖于初始聚类中心, 因此有必要进一步优化

3.2 混沌混合优化

对于 $J(X, V)$ 连续对象的全局优化问题, 选择 Logistic 映射产生的混沌变量

$$z_{i,t+1} = 4z_{i,t}(1.0 - z_{i,t}),$$

$$z_{i,t} \in [0, 1], i = 1, 2, \dots, m + 1 \quad (12)$$

混沌点集的概率密度为

$$p(z) = \frac{1}{\pi\sqrt{z(1-z)}} \quad (13)$$

当 Δz 较小时, Logistic 映射落入区间 $(z - \Delta z/2, z + \Delta z/2)$ 的概率为

$$p\{z - \Delta z/2 \leq Z \leq z + \Delta z/2\} = p(z)\Delta z \quad (14)$$

式(14)说明 Logistic 映射落入中心区间的概率较小, 而在边缘较大, 单纯提高迭代步数并不能显著提高映射的遍历程度 为此, 本文提出一种基于混沌混合的优化方法, 其主要思想是: 首先通过变换或多个变量组合, 将混沌变量映射到另一区间, 产生变换混沌序列, 概率密度分布满足在优化点处局部寻优的要求, 如呈中心大两边小; 然后在寻优区间将序列放大进行混沌初步优化, 得到初始优化值群; 最后与 FCM 方法相结合, 利用梯度法得到最优解

如取变换

$$z_i = z_{i1} + z_{i2}, \quad (15)$$

那么新的序列的概率分布(迭代 10^4 步时)如图 2 所示 中间密度大, 能较好地局部寻优, 加快收敛速度; 两边密度相对较小, 用于跳出局部最小

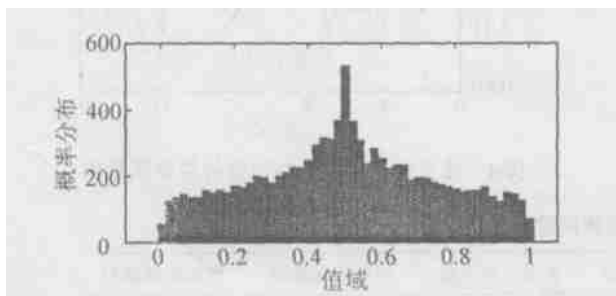


图 2 基于混沌混合序列的概率分布
混沌混合优化的步骤如下:

Step 1: 根据前一步确定聚类类别数 C , 预置允许误差 $\epsilon, r = 0$; 初始化混沌初值 $z_{i,0}$ 为 $[0, 1]$ 区间 $2C$ 个相异的初值, 且 $z_{i,0}$ 不等于映射不动点; 按式(15)生成 $z_{i,0}$, 预置最优混沌参数初值 $z_i^{*C} = z_{i,0}$ 和进化种群个数 $s(s = 5)$, 取前一步得到的模型参数为优化参数 v_i^{*C} 初值

Step 2: 将 $z_{i,t}$ 以前一步得到的初值为中心, 线性映射到优化变量取值区间, 即

$$v_{i,t} = v_i^{*C} + \lambda(r)\beta\sigma^c(z_{i,t} - 0.5),$$

$$i = 1, 2, \dots, C. \quad (16)$$

其中: $\lambda(r)$ 为退火系数, $\lambda(r) \in (0, 0.5)$; β 为宽度系数, 一般取 $\beta = 0.5 \sim 1.0$

Step 3: 若 $J(p, v_{i,t}) < J_s^{*C}$, 则将 $J(p, v_{i,t})$ 按大小插入 J_s^{*C} 队列; 若 $J_s^{*C} = J(p, v_{i,t})$, 则 $z_i^{*C} = z_{i,t}$ 比较该点与队列中前一点和后一点的距离, 若与前点距离小于某预定的正数, 则删除该点; 若与后点距离小于某预定的正数, 则删除后点; 否则继续

Step 4: $t = t + 1$, 混沌变量迭代按式(12)和(15)计算 $z_{i,t+1}$

Step 5: 重复 Step 2 ~ Step 4, 直到一定步数内 J_s^{*C} 保持不变或达到最大迭代上限为止 得到的最优 $v_{i,s}^{*C}$ 即为混沌算法得到的最优聚类中心解群, J_i^{*C} 为算法得到的最优解

Step 6: 以候选优化群为初值, 执行 FCM 算法求得新的最优聚类中心 v^* , 并计算 σ^*

聚类后, 权值训练算法运用一组输入向量 $\{X_t | X_t \in R^m, t = 1, 2, \dots, n\}$ 及其输出 Y_t 进行迭代的 LSM, 使误差函数

$$E = \sum_{i=1}^n \Phi(X_i)W - Y_i^2 \quad (17)$$

最小, 从而达到调节权值 w_i 的目的 w_i 的调整公式为

$$w_{t+1} = w_t + \eta\Phi(X_t)(Y_t - F(X_t)), \quad (18)$$

其中 t 表示迭代次数往复训练, 直到误差满足要求为止

4 在转炉终点磷含量预报中的应用

4.1 终点磷含量预报模型

转炉的动态、静态控制过程, 主要是以终点碳含量和终点钢水温度作为控制对象, 而没有直接对终点磷含量进行预报和控制 随着冶金企业日趋激烈的竞争和对钢材质量要求的提高, 需要对终点磷含量进行预报和控制^[7] 转炉炼钢是十分复杂的非线性过程, 单纯依靠传统的建模方法非常困难, 也不十分有效 基于数据的模糊神经网络技术是将模糊逻辑和神经网络结合起来, 二者取长补短, 具有强大的函数映射能力 它能够通过输入输出数据对过程进行

有效的学习,因而在解决基于知识的非线性系统建模领域具有良好的应用前景 终点磷含量预报模型结构如图 3 所示



图 3 终点磷含量预报模型结构

根据机理分析和实验,选择吹氧量、石灰重量、轻烧重量、萤石重量、钢水重量、废钢重量和矿石重量、钢水的硅、锰、磷含量、钢水温度、上一炉终点磷含量、参考炉次的终点磷含量作为测量的二次变量,炉次的终点磷含量 Y_r 作为模型输出量 神经网络模型只考虑了影响吹炼磷含量的主要因素,并且只考虑了静态因素 然而,目标值还与炉况、操作人员的操作习惯、其他未考虑扰动等因素有关,因此必须考虑网络模型学习后残差的补偿^[8]

实验采用多元线性回归建立残差与各影响因素之间关系的模型,将 Q 个样本作为样本集,以此对训练好的网络给出的预测值进行补偿 设 $\Delta y(q)$ ($q = 1, 2, \dots, Q$) 为吹炼实际终点磷含量与神经网络给出的预测值的残差, $x_i(q)$ ($q = 1, 2, \dots, Q$) 为神经网络的输入变量,从而建立线性回归补偿模型

$$\Delta Y = XA. \tag{19}$$

其中

$$\begin{aligned} \Delta Y &= [\Delta y(1), \Delta y(2), \dots, \Delta y(Q)]^T; \\ X &= [X_0, X_1, \dots, X_p], X_0 = [1, 1, \dots, 1]^T, \\ X_i &= [x_i(1), x_i(2), \dots, x_i(Q)]^T, i = 1, 2, \dots, P; \end{aligned}$$

$A = [a_0, a_1, \dots, a_p]^T$ 为回归系数,可应用最小二乘法求得

$$A = (X^T X)^{-1} X^T \Delta Y. \tag{20}$$

残差补偿可采用新息模型,即利用前 Q 次的输入变量和预测残差来预测第 $Q + 1$ 次的残差 $\Delta y(Q + 1)$,再加上 RBF 模型输出 Y_r ,得到终点磷含量测量值 在得到实际的 $Q + 1$ 次残差信息后,将其和第 $Q + 1$ 次的神经网络输入变量加入样本集,去掉第 1 个样本的信息,始终保持 Q 个样本

对于终点磷含量模型,由生产实际获取样本数

据,先用 FCM 聚类算法进行聚类,确定中心点并计算相关参数,聚类后对输入变量的重要等级进行效验,应用上述混沌混合算法确定中心点,训练 RBF 神经网络权值,最后得到终点磷含量模糊 RBF 网络模型 对于一个新的过程输入,先由式(4)计算其隶属度,再由式(3)网络输出和式(19)计算残差补偿,从而完成在此输入作用下过程输出的估计

4.2 数据处理与模型结果

考虑到外部干扰等因素,输入数据不能直接作为软测量仪表输入,必须进行预处理 预处理主要有数据变换和误差处理 先对测量数据进行归一化处理,设输入数据 X 的最大值和最小值分别为 X_{max} 和 X_{min} ,则归一化值为 $X = (X - X_{min}) / (X_{max} - X_{min})$. 误差分为随机误差和过失误差,对于随机误差,采用移动平均值数字滤波;对于过失误差,采用数据协调技术^[9]

1) 输入变量等级效验的主要思想是^[10]:除规则前提部分需效验的变量外,去除所有其他变量,并计算有关这个变量的模糊输出,通过计算 m 个输入-输出曲线进行效验,将影响较小的输入变量删除 经过效验,删除了钢水重量和废钢重量两个影响较小的输入变量

2) 采用模糊神经网络模型和改进的 RBF 神经网络算法,将实验数据分成训练数据和测试数据两部分,按 FCM 和 FCM 加混沌混合优化两种方法,在自行编制的软测量模型软件上运行,得到系统的辨识模型

3) 终点磷含量模型训练误差及检验误差数据如表 1 所示,部分模型测试数据输出相应于实际曲线的拟合如图 4 所示 由图可见,本文提出的混沌混合优化方法优于 FCM 方法,模型较好地反映了过程的变化和趋势,可应用于控制系统的分析和设计

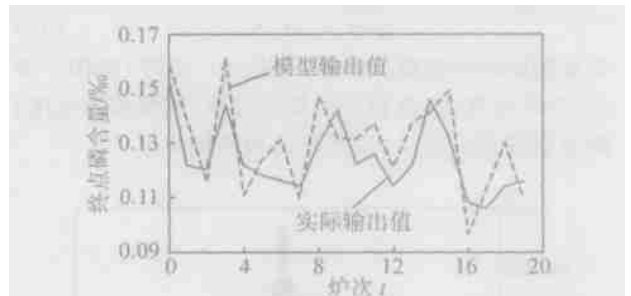


图 4 终点磷含量模型输出值与实际测量值

表 1 终点磷含量模型训练误差及检验误差

模 型	预报命中率 (± 0.004)	预报命中率 (± 0.003)	主中心点个数	训练数据规模	检验数据规模
本文方法	90.1%	87.3%	8	80	20
FCM 方法	86.7%	84.2%	8	80	20

5 结 论

本文将T-S模糊模型转化为RBF神经网络,提出一种新的模糊神经网络混沌混合优化方法。首先采用混沌变换获得所需概率分布的序列,应用混沌变换序列优化获得初步优化群;然后应用FCM获得最优中心,解决了神经网络模型复杂性与逼近精度和泛化性能的矛盾,增强了模型的预测能力。实际应用结果表明,采用该方法能达到较好的建模效果,但算法较为复杂,计算量较大。如何简化计算,使算法既能较快收敛又能保持全局最优,将是今后研究的重点。

参考文献(References)

- [1] Bezdek J C. Convergence theory for fuzzy C-means: Counterexamples and repairs [J]. *IEEE Trans on Systems, Man and Cybernetics*, 1987, 17(2): 873-877.
- [2] Tsao E C, Bezdek J C. Fuzzy Kohonen clustering networks [J]. *Pattern Recognition*, 1994, 27(5): 757-764.
- [3] Wang L X, Medel J M. Backpropagation fuzzy system as nonlinear dynamic system identifier [A]. *IST IEEE Int Conf on Fuzzy Systems* [C]. San Diego, 1992: 1409-1418.
- [4] 张彤, 王宏伟, 王子才. 变尺度混沌优化方法及其应用 [J]. *控制与决策*, 1999, 14(3): 285-287.
(Zhang T, Wang H W, Wang Z C. Mutative scale chaos optimization algorithm and its application [J]. *Control and Decision*, 1999, 14(3): 285-287.)
- [5] Hunt K J. Extending the functional equivalence of radial basis function networks and fuzzy inference system [J]. *IEEE Trans on Neural Networks*, 1996, (3): 776-781.
- [6] Chen M Y, Linkens D A. A fuzzy modelling approach using hierarchical neural networks [J]. *Neural Comput and Applications*, 2000, (9): 44-49.
- [7] 原隆康. 转炉终点动态磷控制技术开发 [J]. *世界钢铁*, 2000, 2: 53-59.
(Yuan L K. Development of dynamic control technology for the endpoint phosphorus content of converter [J]. *World Iron and Steel*, 2000, 2: 53-59.)
- [8] 谢书明, 高宪文, 柴天佑. 基于灰色模型的转炉炼钢终点预报研究 [J]. *钢铁研究学报*, 1999, 11(4): 9-11.
(Xie S M, Gao X W, Chai T Y. BOF endpoint prediction based on grey model [J]. *J of Iron and Steel Research*, 1999, 11(4): 9-11.)
- [9] 李红军, 秦永胜, 徐用懋. 化工过程中的数据协调及显著误差检测 [J]. *化工自动化及仪表*, 1997, 24(2): 25-32.
(Li H J, Qin Y S, Xu Y M. Data reconciliation and gross error detection in chemical process [J]. *Control and Instruments in Chemical Industry*, 1997, 24(2): 25-32.)
- [10] Lin Y, Cunningham G A. A new approach to fuzzy-neural system modeling [J]. *IEEE Trans on Fuzzy System*, 1995, (3): 190-197.

(上接第260页)

参考文献(References)

- [1] Lawler E L, Lenstra J K, Rinnooy K, et al. *Sequencing and scheduling: A algorithm and complexity* [M]. Amsterdam, 1993.
- [2] Dessouky M I, Deogun J S. Sequencing jobs with unequal ready times to minimize mean flow time [J]. *SIAM J on Computing*, 1981, 10(1): 192-202.
- [3] Chandra R. On $n/1/\sqrt{F}$ dynamic deterministic problems [J]. *Naval Research Logistics Quarterly*, 1979, 26: 537-544.
- [4] Chu C B. Efficient heuristics to minimize total flow time with release dates [J]. *Operations Research Letters*, 1992, 12: 321-330.
- [5] Wang B, Xi Y G, Gu H Y. Rolling horizon heuristic for single machine scheduling problem [A]. *Proc of 2004 World Conf of Intelligent Control and Automation* [C]. Hangzhou, 2004: 2872-2876.
- [6] Suresh Chand, Rodney Traub, Reha Uzsoy. Rolling horizon procedures for the single machine deterministic total completion time scheduling problem with release dates [J]. *Annals of Operations Research*, 1997, 70: 115-125.
- [7] Fang J, Xi Y G. Rolling horizon job shop rescheduling strategy in the dynamic environment [J]. *Int J of Advanced Manufacturing Technology*, 1997, 13: 227-232.
- [8] Ahmadi R H, Bagchi U. Lower bounds for single-machine scheduling problems [J]. *Naval Research Logistics*, 1990, 37: 967-976.