

文章编号: 1001-0920(2011)01-0115-05

基于蚁群算法的小波降噪双阈值优选方法

王传鑫^a, 邵 诚^a, 韩 瑜^b

(大连理工大学 a. 先进控制技术研究所, b. 软件学院, 辽宁 大连 116024)

摘要: 将相关性分析方法和阈值降噪方法相结合, 提出了一种基于蚁群算法的小波降噪双阈值优选方法. 首先根据噪声的分布特点确定合理的寻优区间; 然后根据信号和噪声的小波系数自相关系数的不同特点确定优化目标函数; 最后利用蚁群算法在寻优区间内搜索最优上下阈值. 仿真和压缩机振动故障检测的实际应用表明, 该方法能够选择最优降噪阈值, 有效去除噪声.

关键词: 小波变换; 降噪阈值; 相关性分析; 蚁群算法; 压缩机振动故障检测

中图分类号: TP277

文献标识码: A

Wavelet de-noising double threshold optimization method based on ant colony algorithm

WANG Chuan-xin^a, SHAO Cheng^a, HAN Yu^b

(a. Institute of Advanced Control Technology, b. School of Software, Dalian University of Technology, Dalian 116024, China. Correspondent: WANG Chuan-xin, E-mail: wangchuanxin@yahoo.cn)

Abstract: By combining correlation analysis method and threshold de-noising method, a method based on ant colony algorithm is given for optimizing wavelet de-noising double-threshold. The reasonable optimization interval is determined according to the characteristic of distributed noise. The objective function is chosen according to the difference of autocorrelation coefficient of signal's wavelet coefficient and noise's wavelet coefficient. Finally the optimal upper threshold and lower threshold are calculated by using ant colony algorithm. Simulation and compressor vibration fault detection application results show that the proposed method can optimize the de-noising threshold and denoise effectively.

Key words: wavelet transform; de-noising threshold; correlation analysis; ant colony algorithm; compressor vibration fault detection

1 引言

小波变换理论继承和发展了短时傅里叶变换的局部化思想, 具有良好的时频特性和多分辨率特性, 已广泛应用于信号降噪处理^[1]. 但在许多实际应用中, 现有的降噪方法往往无法取得很好的降噪效果, 因此如何有效降噪、准确提取信号特征, 是目前亟待解决的问题^[2-3].

在小波降噪方法中, 阈值决策降噪法最为常用, 但如何选取合适的阈值一直是焦点问题^[4-6]. Donoho 提出的经典阈值^[4]在渐近意义上具有最优性, 但是随着小波系数的增长, 经典阈值逐渐增大, 有用信号有时也会被当作噪声去除掉. 文献 [5] 提出了双阈值的降噪方法, 其降噪效果优于单阈值降噪方法, 但因其全局阈值的选取不同, 确定出的上下阈值也就不同, 这将直接影响降噪效果. 在小波阈值降噪过程

中, 若阈值选取不当, 则信号的高频部分很容易被误认为是噪声而被去除, 因此需要根据信号和噪声的不同特点进行有效区分.

本文提出了基于蚁群算法的双阈值优选方法, 根据噪声的分布特点确定了合理的寻优区间; 通过对比信号和噪声的小波系数自相关系数的不同表现形式来确定优化目标函数; 利用蚁群算法在寻优区间内根据目标函数计算最优上下阈值. 最后, 通过仿真和实际应用于压缩机振动故障检测的结果验证了该方法能够选取最优降噪阈值, 有效去除噪声.

2 小波降噪原理

考虑到噪声, 信号的观测值可表示为

$$x_i = f_i + \sigma z_i, \quad i = 1, 2, \dots, n. \quad (1)$$

其中: x_i 是信号的观测值, f_i 是信号的真实值, z_i 是满

收稿日期: 2009-10-16; 修回日期: 2010-02-08.

作者简介: 王传鑫(1978-), 男, 博士生, 从事过程建模、压缩机控制系统的研究; 邵诚(1958-), 男, 教授, 博士生导师, 从事先进控制理论等研究.

足 $N(0,1)$ 分布的噪声, σ 是噪声的标准差. 对信号的观测值作离散小波变换, 可得

$$W_{j,k} = W_{j,k}(f) + W_{j,k}(z). \quad (2)$$

其中: $W_{j,k}(f)$ 是信号真实值的小波系数, $W_{j,k}(z)$ 是噪声的小波系数. 对于信号, 其小波变换系数随着尺度的增大而增大; 对于噪声, 其小波变换系数随着尺度的增大而减小. 小波降噪的主要原理是根据信号和噪声在不同尺度上所表现的特征不同, 通过设定阈值将它们区分开.

3 基于蚁群算法的双阈值优选方法

3.1 寻优区间

白噪声服从标准差为 σ 的零均值正态分布, 其主要分布在 $[-3\sigma, 3\sigma]$ 之内^[1], 因此绝对值小于 3σ 的小波系数可以认为主要由噪声的小波系数组成. 根据这一原则, 可以确定寻优区间上限为

$$T_{\max} = 3\sigma. \quad (3)$$

其中 σ 可根据下式估计^[4]:

$$\sigma \approx \text{Mid}/0.6745, \quad (4)$$

Mid 是最小尺度上的小波系数按顺序排列后中间位的取值.

取某层上绝对值最小的小波系数作为寻优区间下限, 即

$$T_{\min} = |C_{j,\min}| > 0. \quad (5)$$

其中: $C_{j,\min}$ 表示小波分解第 j 层中绝对值最小的且不为零的小波系数. 因此, 寻优区间确定为 $[T_{\min}, T_{\max}]$, 以此来合理减小寻优区间, 提高搜索速度.

3.2 优化目标函数

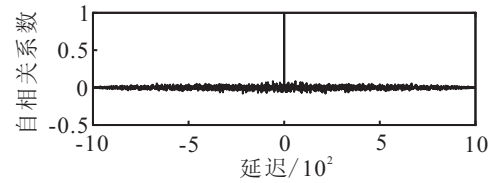
3.2.1 下阈值优化目标

白噪声是一个纯随机过程, 离散白噪声的自相关序列 $A_z(m)$ 具有如下特点^[2]:

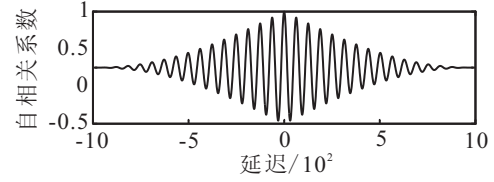
$$A_z(m) = \begin{cases} 1, & m = 0; \\ 0, & m \neq 0. \end{cases} \quad (6)$$

其中 m 是延迟, 随着 m 的增加 $A_z(m)$ 迅速衰减到零, 如图 1(a) 所示. 白噪声的小波系数亦有此特点, 因此最优的下阈值应使小于此阈值的小波系数尽可能满足式 (6). 由于其自相关系数的分布是左右对称的, 可以只考虑其右半部分的系数. 将其右半部分 $(N-1)$ 个系数看作 $(N-1)$ 维空间内的一点, 由于其值均非常小, 这一点到原点的距离也非常小, 可将下阈值优化问题表示为

$$\min F_{\min} = \sqrt{\sum_{i=N+1}^{2N-1} A_z^2(i)}. \quad (7)$$



(a) 白噪声的自相关系数



(b) 信号的自相关系数

图1 白噪声和信号的自相关系数

3.2.2 上阈值优化目标

信号的小波系数自相关序列可表示为

$$A_s(m) = \begin{cases} 1, & m = 0; \\ f(m), & m \neq 0, |f(m)| < 1. \end{cases} \quad (8)$$

其中 $f(m)$ 是 m 的非线性函数. 信号的自相关系数之间存在较强的连续性和关联性, 如图 1(b) 所示. 最优的上阈值应使大于此阈值的小波系数尽可能满足式 (8). 假设某阈值能够将信号中的噪声全部去除, 余下的小波系数的相关系数右半部分构成的向量在 $(N-1)$ 维空间内至原点的距离为常数 C . 但若某阈值未能完全去除信号中的噪声, 余下的小波系数的自相关性将会由于噪声的加入而减弱, 表现在自相关系数上便是使某些系数减小, 从而使其相关系数右半部分构成的向量在 $(N-1)$ 维空间内至原点的距离小于常数 C . 因此, 可将上阈值优化问题表示为

$$\max F_{\max} = \sqrt{\sum_{i=N+1}^{2N-1} A_s^2(i)}. \quad (9)$$

综合考虑式 (7) 和 (9), 得双阈值优化目标函数为

$$\min F = \frac{F_{\min}}{F_{\max}}. \quad (10)$$

式 (10) 的意义是在确定上下阈值时, 应使被移除的小波系数的自相关性越小越好, 同时应使被保留的小波系数的自相关性越大越好.

3.3 蚁群算法寻优

双阈值优化问题是双变量全局寻优, 其约束条件为 $T'_{\min} < T'_{\max}$, T'_{\max} 和 T'_{\min} 分别为最优上下阈值. 在小波分解的每一层上, 需要分别计算最优的上下阈值, 因此计算量较大. 蚁群算法是一种新型启发式进化算法, 具有较强的全局优化能力和较快的收敛速度, 并且具有并行计算的能力^[7-8], 非常适合解决这类双变量全局优化问题, 所以本文采用蚁群算法来对上下阈值寻优. 寻优的目标是让蚁群在寻优区间 $[T_{\min}, T_{\max}]$ 内寻找下阈值 T'_{\min} 和上阈值 T'_{\max} , 使目标函数 (10) 的函数值尽可能小, 具体操作步骤如下:

Step 1: 取自变量 T'_{\min} 和 T'_{\max} , 每个自变量的取值为4位有效数字, 小数点后3位. 将每个自变量的每个数位标记于 X 轴上, 每个数位可能的取值 $[0, 9]$ 标记于 Y 轴上, 如图2所示.

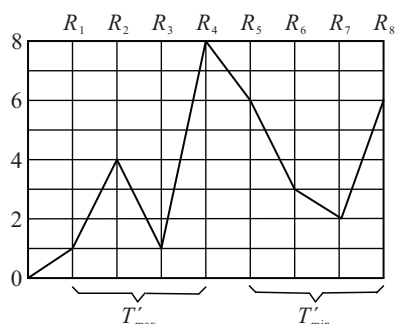


图2 蚁群算法的节点和路径

Step 2: 初始化蚁群数量 m , 并设置禁忌表为

$$T_{\min} \leq T'_{\min} < T'_{\max} \leq T_{\max}.$$

Step 3: 进入循环, 每次循环蚁群从原点出发, 每只蚂蚁根据路径上信息素的浓度和禁忌表来决定行走方向, 目的地为图2中最右侧的节点. 第 k 时刻第 p 只蚂蚁由 R_{i-1} 上的节点 $(x_{i-1}, y_{(i-1)j})$ 向 R_i 上的节点 (x_i, y_{ij}) 移动的概率为

$$P_p(x_i, y_{ij}, k) = \frac{\tau(x_i, y_{ij}, k)}{\sum_{g=1}^N \tau_g(x_i, y_{ij}, k)}, \quad (11)$$

其中 $\tau(x_i, y_{ij}, k)$ 为 k 时刻节点 (x_i, y_{ij}) 的信息素.

Step 4: 每次循环后, 根据

$$\begin{aligned} T'_{\max} &= y_{1j} \times 10^0 + y_{2j} \times 10^{-1} + y_{3j} \times 10^{-2} + \\ &\quad y_{4j} \times 10^{-3}, \\ T'_{\min} &= y_{5j} \times 10^0 + y_{6j} \times 10^{-1} + y_{7j} \times 10^{-2} + \\ &\quad y_{8j} \times 10^{-3} \end{aligned} \quad (12)$$

求出第 p 只蚂蚁经过的路径所代表的自变量取值, 将自变量带入目标函数求出目标函数值. 根据目标函数值, 按下式对节点的信息素进行更新:

$$\Delta\tau_p(x_i, y_{ij}) = \begin{cases} \frac{Q}{F_p}, & \text{蚂蚁 } p \text{ 选择 } (x_i, y_{ij}); \\ 0, & \text{蚂蚁 } p \text{ 不选择 } (x_i, y_{ij}). \end{cases} \quad (13)$$

其中: Q 是信息素常数, F_p 是第 p 只蚂蚁所选择的输出带入目标函数所对应的函数值. 按下式再次更新信息素:

$$\begin{aligned} \tau(x_i, y_{ij}, k+m) &= \\ &(1-\rho)\tau(x_i, y_{ij}, k) + \rho\Delta\tau(x_i, y_{ij}, k), \\ \Delta\tau(x_i, y_{ij}, k) &= \sum_{p=1}^m \Delta\tau_p(x_i, y_{ij}). \end{aligned} \quad (14)$$

其中: $\Delta\tau_p(x_i, y_{ij})$ 是第 p 只蚂蚁在本次循环中, 在节点 (x_i, y_{ij}) 上留下的信息素; $\Delta\tau(x_i, y_{ij}, k)$ 是本次循

环中, 节点 (x_i, y_{ij}) 上的信息素总和; $\rho(0 \leq \rho < 1)$ 是信息素的持久性, $(1-\rho)$ 表示信息素的消逝程度.

Step 5: 再次循环, 直到全部蚂蚁收敛到同一条路径或达到最大循环次数.

Step 6: 根据式 (12) 得出最优上下阈值.

3.4 阈值处理规则

根据蚁群算法优化的结果, 由下式确定小波系数取值:

$$W'_{j,k} = \begin{cases} 0, & |W_{j,k}| < T'_{\min}; \\ \frac{T'_{\max}}{T'_{\max} - T'_{\min}} (|W_{j,k}| - T'_{\min}) \text{sgn}(W_{j,k}), & T'_{\min} \leq |W_{j,k}| \leq T'_{\max}; \\ W_{j,k}, & |W_{j,k}| > T'_{\max}. \end{cases} \quad (15)$$

其中: $W_{j,k}$ 和 $W'_{j,k}$ 分别为通过阈值处理前后的小波系数, T'_{\min} 和 T'_{\max} 是蚁群算法计算出的最优下阈值和上阈值. 大于上阈值的小波系数主要由信号组成, 直接保留; 小于下阈值的小波系数主要由噪声组成, 将其系数置零; 对处于上下阈值之间的小波系数作收缩处理.

4 仿真分析

选取 Matlab 中的 heavy sine 和 mishmash 两种不同特点的染噪信号进行降噪仿真. 基小波采用 sym8 小波, 分解层数为 5 层. 对每一种信号分别采用固定阈值、Rigorous SURE 阈值、传统双阈值以及本文提出的优化双阈值对其降噪. 这 4 种阈值对 heavy sine 信号的降噪效果见图 3, 对 mishmash 信号的降噪效果见图 4.

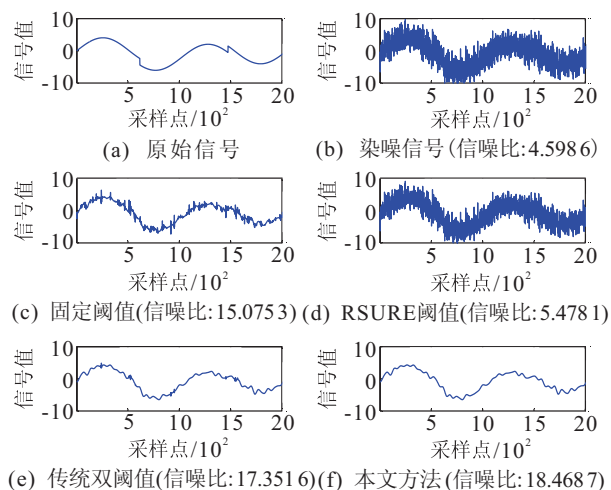


图3 4种阈值对 heavy sine 信号的降噪效果

由图3可见, 对于 heavy sine 这种连续性较强的信号, 固定阈值的降噪效果一般, 如图3(c)所示, 因为信号和噪声的小波系数的特征比较明显, 所以通过选取合适的阈值便可有效降噪, 但其降噪后信号的

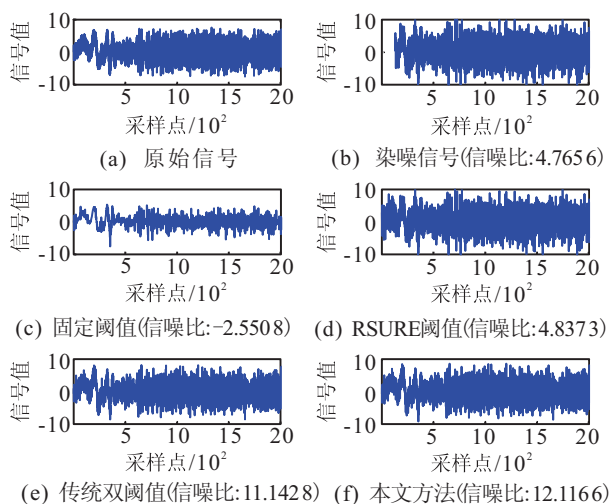


图 4 4 种阈值对 mishmash 信号的降噪效果

吉布斯效应比较明显; RSURE 阈值无法有效降噪, 如图 3(d) 所示, 因为 RSURE 阈值是基于无风险估计方法选取阈值, 阈值选取比较保守, 所以无法有效降噪; 传统双阈值方法的降噪效果在很大程度上取决于其全局阈值, 但最优的全局阈值往往无法取得, 因此其降噪效果较好, 如图 3(e) 所示; 本文方法根据信号和噪声小波系数的不同特征来去除噪声, 合理减小了寻优区间, 并对上下阈值进行优化选取, 因此获得了最好的降噪效果, 如图 3(f) 所示。

由图 4 可见, 对于 mishmash 这种类似于振动信号的高频脉冲信号, 固定阈值无法有效降噪, 如图 4(c) 所示, 因为这类信号连续性不强, 信号高频部分的系数很容易被当作噪声去除; RSURE 阈值由于采用无风险估计方法选取阈值, 虽然无法有效降噪, 但可以保留信号的特征, 如图 4(d) 所示, 这也说明了降噪过程中必须对某层中全部的小波系数作全局考虑; 对于 mishmash 信号, 传统双阈值方法和本文方法均取得了较好的降噪效果, 如图 4(e) 和图 4(f) 所示, 这也说明双阈值降噪方法要优于单阈值降噪方法。由于本文方法对阈值优化选取, 因此获得了更高的信噪比。

表 1 为本文采用蚁群算法、遗传算法和粒子群算法分别用于优选双阈值各项指标的比较。采用的信号为图 3 中的 heavy sine 信号, 染噪信号采用 sym8 小波分解为 5 层, 因此需要进行 5 次优化来计算每层中的最优上下阈值。蚁群算法的蚁群数、粒子群算法的粒子群数和遗传算法的种群数均取为 20, 最大循环次数选为 50 次。由表 1 可见, 本文采用的蚁群算法具有最短的优化时间和最高的信噪比, 而且 5 次优化都在最

表 1 3 种优化算法用于优选双阈值的比较

	蚁群算法	粒子群算法	遗传算法
优化时间	4 m 29 s	5m02s	5m58s
信噪比	18.537	18.264	18.474
到达最大循环次数	0	1	2

大循环次数内收敛。

5 压缩机振动故障信号检测

压缩机振动故障是影响压缩机机组稳定运行的关键因素, 对压缩机振动信号的检测和分析能够及早发现潜在隐患、及时确定故障点, 对机组安全和人员安全以及降低维护成本都具有十分重要的意义^[9-10]。压缩机故障主要有转子不平衡、转子不对中、基座松动、叶片故障等。压缩机的振动信号中包含丰富的压缩机运行状态信息, 对其进行频谱分析就能够判断故障类型和故障点。例如转子不平衡故障, 其振动的频率成分单一且明朗, 主要表现为转子的基频等于转子的旋转频率。

在某化工厂裂解气压缩机控制系统升级改造过程中, 测其基座振动信号, 采样频率为 10 kHz。由于测量时没有采取任何抗干扰措施, 测量信号中含有较大的背景噪声, 如图 5(a) 所示。图 5(b) 是其频谱图, 由于噪声太大, 从频谱图中只能大致判断这台压缩机的基频大约为 50 Hz 外, 看不出其他任何关于压缩机运行状态的特征。采用 sym8 小波将染噪信号分解 5 层后, 用本文方法对其进行降噪处理, 降噪后信号如图 5(c) 所示, 再对其进行频谱分析, 如图 5(d) 所示。从图 5(d) 中可以明显看出 1/2 倍频、1 倍频 (50 Hz)、3 倍频、5 倍频均有明显的表现, 呈现出典型转子不对中的特征。在对其大修时发现其转子不均匀结垢情况严重, 从而导致压缩机转子不对中, 通过清除结垢可以解决该问题。

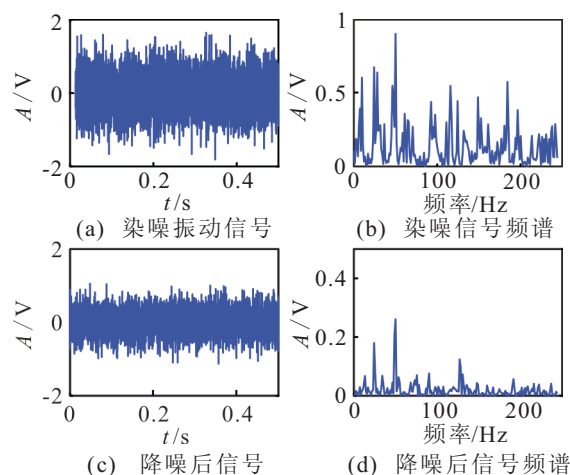


图 5 压缩机振动信号和其频谱

6 结 论

本文提出了基于蚁群算法的双阈值优选方法, 优化区间根据噪声的分布特点而确定, 优化目标则根据信号和噪声小波系数的自相关系数的不同表现形式来确定。利用蚁群算法进行双阈值的寻优, 仿真和实际应用结果说明本文方法能够有效降噪, 在振动故障检测中具有良好的应用前景。

参考文献(References)

- [1] 臧玉萍, 张德江, 王维正. 小波分层阈值降噪法及其在发动机振动信号分析中的应用[J]. 振动与冲击, 2009, 28(8): 57-60.
(Zang Y P, Zhang D J, Wang W Z. Per-level threshold de-noising method using wavelet and its application in engine vibration analysis[J]. J of Vibration and Shock, 2009, 28(8): 57-60.)
- [2] 王维, 张英堂, 任国全. 小波阈值降噪算法中最优分解层数的自适应确定及仿真[J]. 仪器仪表学报, 2009, 30(3): 526-530.
(Wang W, Zhang Y T, Ren G Q. Adaptive selection and simulation of optimal decomposition level in threshold de-noising algorithm based on wavelet transform[J]. Chinese J of Scientific Instrument, 2009, 30(3): 526-530.)
- [3] 张荣祥, 李正强, 郑世杰. 基于遗传算法的双阈值小波去噪方法研究[J]. 传感器与微系统, 2007, 26(6): 20-22.
(Zhang R X, Li Z Q, Zheng S J. Research of wavelet de-noising method by double-threshold based on genetic algorithm[J]. Transducer and Microsystem Technologies, 2007, 26(6): 20-22.)
- [4] Donoho D L. De-noising by soft-thresholding[J]. IEEE Trans on Information Theory, 1995, 41(3): 613-627.
- [5] Gao H Y, Bruce A G. Waveshrink and semisoft shrinkage[R]. Seattle: STAT Sci Division of Mathsoft Inc., 1995.
- [6] Cary S, Cajetan M A, Phil H, et al. An approach to vibration analysis using wavelets in an application of aircraft health monitoring[J]. Mechanical Systems and Signal Processing, 2007, 21(3): 1255-1272.
- [7] 刘利强, 戴运桃, 王丽华. 蚁群算法参数优化[J]. 计算机工程, 2008, 34(11): 208-210.
(Liu L Q, Dai Y T, Wang L H. Ant colony algorithm parameters optimization[J]. Computer Engineering, 2008, 34(11): 208-210.)
- [8] 张培林, 钱林方, 曹建军, 等. 基于蚁群算法的支持向量机参数优化[J]. 南京理工大学学报, 2009, 33(4): 464-468.
(Zhang P L, Qian L F, Cao J J, et al. Parameter optimization of support vector machine based on ant colony optimization algorithm[J]. J of Nanjing University of Science and Technology, 2009, 33(4): 464-468.)
- [9] 施俊侠, 杨兆建, 李春生. 离心式压缩机的状态监测与故障诊断技术[J]. 中国机械工程, 2008, 19(9): 1099-1102.
(Shi J X, Yang Z J, Li C S. Condition detecting and trouble diagnosis technology on centrifugal compressor[J]. China Mechanical Engineering, 2008, 19(9): 1099-1102.)
- [10] 曾庆生, 王湘江. 基于小波能量谱和粗糙集的离心式压缩机振动故障诊断[J]. 中南大学学报, 2009, 40(3): 706-710.
(Zeng Q S, Wang X J. Fault diagnosis of centrifugal compressor vibration based on wavelet power spectrum and rough set theory[J]. J of Central South University, 2009, 40(3): 706-710.)
- ~~~~~
- (上接第114页)
- [9] 添玉, 黄道, 刘东波. 一种风险规避供应链的协调研究[J]. 华东理工大学学报, 2009, 35(1): 132-138.
(Tian Y, Huang D, Liu D B. Supply chain coordination with a risk-averse retailer[J]. J of East China University of Science and Technology, 2009, 35(1): 132-138.)
- [10] 素寒生, 金以慧. 非对称信息下供需链中供应商的回购决策分析[J]. 控制与决策, 2004, 19(3): 335-338.
(Suo H S, Jin Y H. Supplier's optimal buy back decision under asymmetric information in a two-stage supply chain[J]. Control and Decision, 2004, 19(3): 335-338.)
- [11] 李善良, 朱道立. 不对称信息下供应链线性激励契约委托代理分析[J]. 计算机集成制造系统, 2005, 11(12): 1758-1762.
(Li S L, Zhu D L. Principal-agent analysis of supply chain incentive contract with asymmetric information[J]. Computer Integrated Manufacturing Systems, 2005, 11(12): 1758-1762.)
- [12] 汪峻萍, 王圣东. 非对称信息下以制造商为核心的供应链协调模型[J]. 系统工程学报, 2008, 23(1): 60-66.
(Wang J P, Wang S D. A coordination model on the manufacturer core supply chain under asymmetric information[J]. J of Systems Engineering, 2008, 23(1): 60-66.)
- [13] Baron D, Myerson R. Regulating a monopolist with unknown costs[J]. Econometrica, 1982, 50(4): 911-930.
- [14] Albert Y H. Supplier-buyer contracting: asymmetric cost information and cutoff level policy for buyer participation[J]. Naval Research Logistics, 2001, 48(1): 41-64.
- [15] Corbett C. Stochastic inventory systems in a supply chain with asymmetric information: cycle stocks, safety stocks and consignment stock[J]. Operations Research, 2001, 49(4): 487-500.