第23卷 第3期 自 动 化 学 报 Vol. 23, No. 3
1997年5月 ACTA AUTOMATICA SINICA May, 1997

预防阈值控制策略及其实现[1)]

涂菶生
(南开大学计算机与系统科学系 天津 300071)
宋东平 Sheldon XC Lou
(浙江大学工业控制国家重点实验室 杭州 310027) (美国加州大学)

摘 要 对机器故障为一般概率分布的不可靠制造系统,提出一种预防阈值控制策略和实现这一控制策略的参数优化方法. 应用扰动分析法得到性能指标对控制参数的灵敏度估计,并证明了估计的无偏性,又利用随机逼近技术优化控制参数. 仿真结果验证了控制策略的有效性.

关键词 制造系统,扰动分析,无偏估计,随机逼近.

1 引言

近年来,已有不少学者研究了不可靠制造系统的生产控制问题[1-3],它们假设机器的不可靠性由一个有限状态的 Markov 过程刻划,即机器的故障率和修复率均为恒定常数. 文献[2]证明了对一台机器两个机器状态只加工一类产品的系统,阈值(hedging point)控制策略是最优的,且导出了控制参数的精确解. 通常,当机器的故障率和修复率均为恒定常数时,等价于机器每次处于因故障而进行维修的时间和机器每两次相邻故障之间的正常时间都服从参数恒定的负指数分布. 本文将考虑机器每次因失效而维修的时间和机器每两次失效之间的时间分别服从同一随机变量的概率分布,但不要求这两个随机变量为指数分布. 对这类比较一般的系统,本文将说明 HP 控制策略并不是最优的,并推广 HP 策略得到预防阈值(PHP)控制策略. 其控制效果明显优于 HP 策略,同时给出了实现 PHP 策略的一种简单算法.

2 模型与控制策略

考虑一台机器只加工一类产品的制造系统. 机器是不可靠的,它每次处于正常、故障的时间分别服从随机变量为 ξ,η 的分布,且 $0<E\xi=a<\infty$, $0<E\eta=b<\infty$. 机器的生产率记为 $u(t)$,当机器正常时,它可以在$[0,U]$内任意调整,其中 U 为最大生产率. 产品的需求率是恒定的,记为 d;生产累积量与需求累积量之差记为 $x(t)$. 则系统可用 Flow Rate

1) 得到国家自然科学基金和国家教委博士点专项基金资助. 本文曾在 12th IFAC World Congress (Sydney, 1993)上宣读.

收稿日期 1995-09-07

Control Model[1]描述如下：

$$\dot{x}(t) = u(t) - d, \tag{1}$$

$$0 \leqslant u(t) \leqslant \alpha(t) \cdot U \tag{2}$$

其中 $\alpha(t)=1$，当机器正常时；$\alpha(t)=0$，当机器故障时. 目标函数为长期平均费用

$$J = \lim_{T\to\infty} E \frac{1}{T}\int_0^T [c^+ x^+(t) + c^- x^-(t)]\mathrm{d}t, \tag{3}$$

其中 $x^+(t)=\max(x(t),0)$，$x^-(t)=\max(-x(t),0)$，常数 c^+，c^-分别是产品库存和欠缺的费用系数. 这样，问题就化为寻找控制策略 $u(t)$，使得目标函数 J 达到最小. 为使问题有意义，假定系统在长期运行下有足够的生产能力，即满足

$$a \cdot U > (a+b) \cdot d. \tag{4}$$

文[1]提出的 HP 策略如下：如果 $\alpha(t)=1$，则

$$u(t) = \begin{cases} U, & x(t) < h, \\ d, & x(t) = h, \\ 0, & x(t) > h. \end{cases} \tag{5}$$

其中 h 为控制参数，称为阈值. 推广后可得到如下预防阈值(PHP)控制策略：

如果 $\alpha(t_0^-)=0$，且 $\alpha(\tau)=1, \forall\, \tau\in[t_0,t]$，则

$$u(t) = \begin{cases} U, & \text{if } t \leqslant t_0 + s, \quad x(t) < h, \\ d, & \text{if } t \leqslant t_0 + s, \quad x(t) = h, \\ 0, & \text{if } t \leqslant t_0 + s, \quad x(t) > h, \\ U, & \text{if } t > t_0 + s. \end{cases} \tag{6}$$

其中 h,s 为控制参数. 在每一个机器正常的时间段内，区间$[t_0,t_0+s]$称为阈值控制阶段，$t>t_0+s$ 称为预防控制阶段.

说明.

1)如果 $s=\infty$，则 PHP 策略退化成为 HP 策略，因此，PHP 是 HP 的推广；不管 ξ,η 为什么分布，对 PHP 策略的控制参数 h,s 进行最优化后，其控制效果必定不次于 HP 策略，代价是控制参数搜索空间的增大.

2)如果 $\xi\equiv a,\eta\equiv b$，这时系统变成确定性的. 在性能指标(3)和条件(4)下，不难导出最优控制规律就是 PHP 策略，且 $h=0,s=a-bdc^-/(U-d)(c^-+c^+)$.

3)当 ξ,η 为一般的概率分布时，其最优控制策略应该是反馈形式的，并能预测随机事件的发生，即为带预测的反馈控制. 而 PHP 策略中的预防控制部分事实上起到了预估故障发生的作用. 本文没能给出 PHP 策略最优性的证明，但因为 PHP 策略在更大范围内搜索，包含有预估的思想，因此比 HP 策略更接近于最优控制形式.

3 参数优化实现 PHP 策略

设系统在 $t=0$ 的初值条件为 $x(0)=0,\alpha(0)=1$. 在 PHP 策略控制下，系统的采样路径 $x(t)$ 由随机过程 $\alpha(t)$ 和控制参数 $\theta=(s,h)$ 唯一确定，如图 1 所示.

一次采样路径的目标函数记为 $L_T(\theta,\alpha(t)) = \frac{1}{T}\int_0^T [c^+ x^+(t) + c^- x^-(t)]\mathrm{d}t$，因此有

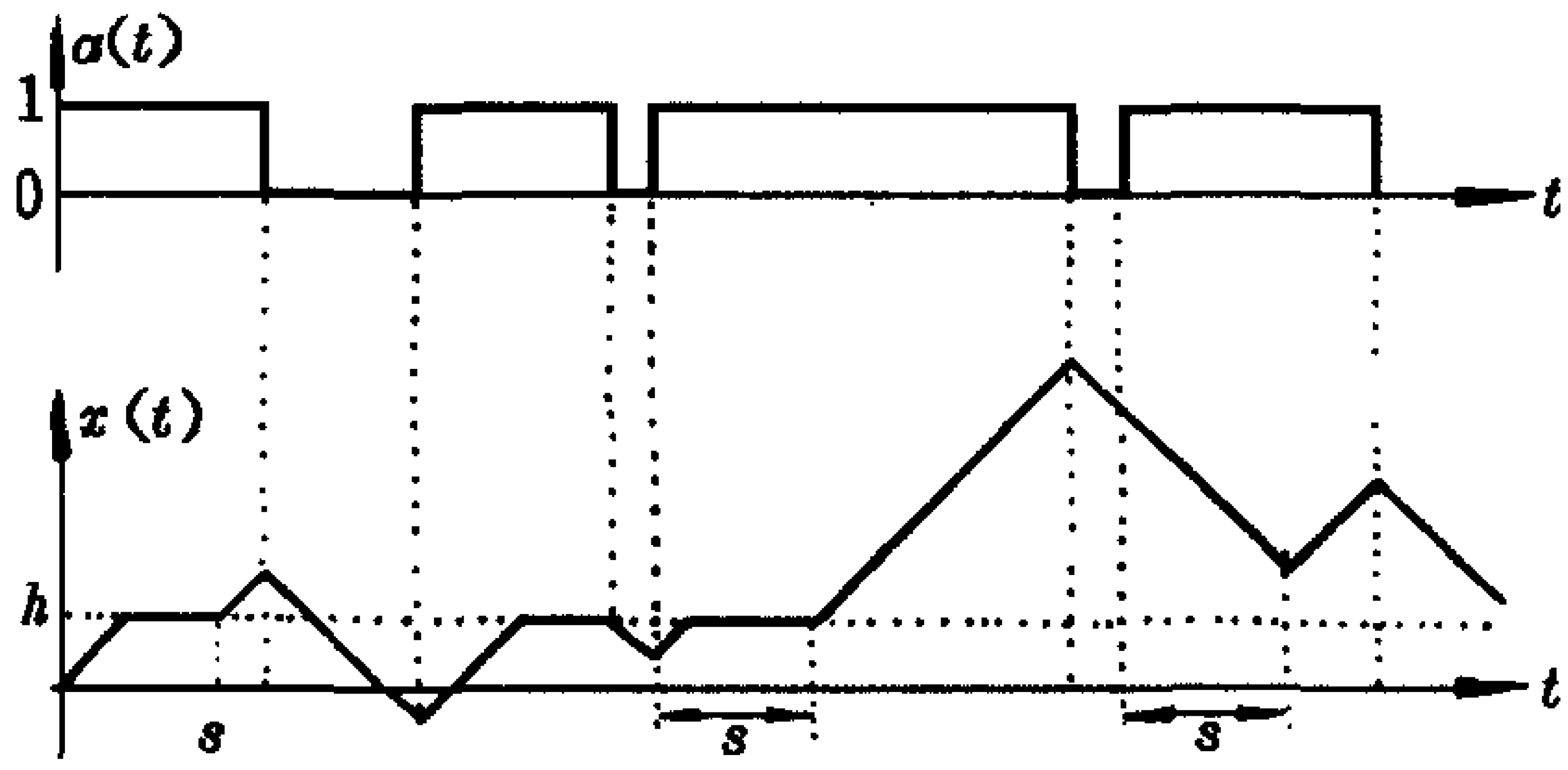

图 1　系统在 PHP 策略下的一次采样路径

$J_T(\theta)=EL_T(\theta,\alpha(t))$. 如果能够求出梯度 $\nabla J_T(\theta)$,则可以优化参数 $\theta=(s,h)$. 扰动分析法(PA)[4]的优点是基于一次采样路径就能给出 $\nabla J_T(\theta)$的估计. 但其可行的前提是要求估计的无偏性.

定理 1. 系统(1),(2)在 PHP 策略控制下,以 $J_T(\theta)=EL_T(\theta,\alpha(t))$为目标函数,则有

$$\text{(a)}\quad \frac{\partial EL_T(\theta,\alpha(t))}{\partial\theta}=E\frac{\partial L_T(\theta,\alpha(t))}{\partial\theta};\text{(b)}\quad E\left|\frac{\partial L_T(\theta,\alpha(t))}{\partial\theta}\right|<\infty.$$

为了证明定理 1,先给出下列事件的定义

1)机器故障——$\alpha(t^-)=1,\alpha(t)=0$;

2)机器修复——$\alpha(t^-)=0,\alpha(t)=1$;

3)机器阈值阻塞——机器正常时,$x(t)$从不等于 h 变为等于 h;

4)机器预防生产——从机器每次修复开始计时,当超过时间 s 时,称为发生机器预防生产事件;

5)$S=\{\omega$:在$[0,T]$时间内,存在某一时刻,同时发生上述不同的两个或两个以上事件$\}$,其中 ω 为系统在 PHP 策略控制下的样本试验.

记$[0,T]$中的事件数为 N,设事件发生的时刻为$\{t_0=0,t_1,t_2,\cdots,t_N=T\}$,称为事件点. 则有以下引理.

引理 1. 如上定义的 N 是一个 ω 的非负随机变量,且 $P\{\omega|N=\infty\}=0$.

引理 2. 满足条件 S 的概率为零,即 $P\{\omega|\omega\in S\}=0$.

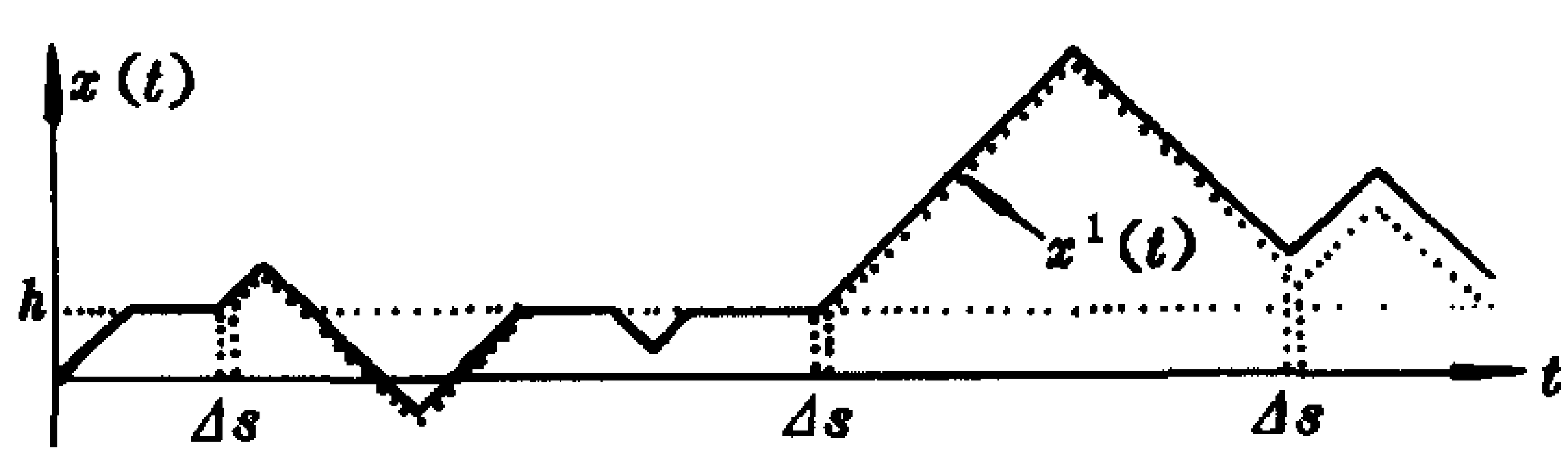

图 2　s 扰动前后的路径

固定参数 h,让 s 作微小扰动变为 $s+\Delta s$,其中 Δs 为充分小的正数. 图 2 给出了满足 $\omega\notin S$且 $N<\infty$的某次采样过程的标称路径 $x(t)$的扰动路径 $x'(t)$. 因为试验 $\omega\notin S$ 且 $N<\infty$,所以对充分小的 Δs,总可以保证标称轨迹 $x(t)$与扰动轨迹 $x'(t)$的确定性相似. 显然,两相邻事件之间的扰动量是不变的. 记 Δx_i 为第 i 个事件与第 $i+1$ 个事件之间 $x(t)$的扰动量,即 $\Delta x_i=x_i'(\cdot)-x_i(\cdot)$,并置初值 $\Delta x_0=0$,则可得 s 作微小扰动时系统的扰动规则如下:

性质 1. 1)设 t_i 为预防生产事件发生的时刻,如果 $x(t_i^-)=h$,则 $x(t)$ 将叠加产生扰动 $-\Delta s(U-d)$,即 $\Delta x_i=\Delta x_{i-1}-\Delta s(U-d)$;如果 $x(t_i^-)>h$,则 $\Delta x_i=\Delta x_{i-1}-\Delta s\cdot U$;如果 $x(t_i^-)<h$,则 $\Delta x_i=\Delta x_{i-1}$. 2)设 t_i 为阈值阻塞事件发生的时刻,则 $\Delta x_i=0$. 3)若 t_i 为其它事件发生的时刻,则 $\Delta x_i=\Delta x_{i-1}$.

从上述规则可知,扰动量 $\Delta x_i(i=1,2,\cdots,N)$ 是 Δs 的线性函数,即 $\Delta x_i/\Delta s(i=1,2,\cdots,N)$ 与 Δs 无关. 因为预防生产事件发生的条件为机器连续保持正常的时间超过 s,所以在 $[0,T]$ 中预防生产事件发生的累加次数不超过 $[T/s]$. 因而,存在常数 const,使得 $|\Delta x_i|<\text{const}\cdot\Delta s$ 对 ω 一致成立. 由性质 1 和标称轨迹与扰动轨迹的相似性,不难得到

$$\begin{aligned}\frac{\partial L_T(h,s,\alpha(t))}{\partial s}&=\lim_{\Delta s\to 0}\frac{L_T(h,s+\Delta s,\alpha(t))-L_T(h,s,\alpha(t))}{\Delta s}\\&=\lim_{\Delta s\to 0}\frac{1}{T\cdot\Delta s}\sum_{i=1}^{N}[(c^+\Delta x_i(t_i-t_{i-1})^+-c^-\Delta x_i(t_i-t_{i-1})^-+0(\Delta s)]\\&=\frac{1}{T}\sum_{i=1}^{N}[(c^+\frac{\Delta x_i}{\Delta s}(t_i-t_{i-1})^+-c^-\frac{\Delta x_i}{\Delta s}(t_i-t_{i-1})^-],\end{aligned}\tag{7}$$

其中 $(t_i-t_{i-1})^+$ 为 (t_{i-1},t_i) 中 $x(t)\geqslant 0$ 的区间长度,$(t_i-t_{i-1})^-$ 为 (t_{i-1},t_i) 中 $x(t)<0$ 的区间长度,且上式与 Δs 无关.

引理 3. 系统(1),(2)在 PHP 策略控制下,有

$$L_T(h,s+\Delta s,\alpha(t))-L_T(h,s,\alpha(t))=\frac{\partial L_T(h,s,\alpha(t))}{\partial s}\cdot\Delta s+0(\Delta s,\alpha(t))\quad a.s.,\tag{8}$$

其中 $\frac{\partial L_T(h,s,\alpha(t))}{\partial s}$ 由(7)式给出,且 $0(\Delta s,\alpha(t))$ 满足

$$\lim_{\Delta s\to 0}E\frac{|0(\Delta s,\alpha(t)|}{\Delta s}=0.\tag{9}$$

证明. 由引理 1 和引理 2,只需考虑 $N<\infty$ 和不满足条件 S 的样本. 此时,如果先给定 ω,则(7)式显然成立,亦即

$$\lim_{\Delta s\to 0}|0(\Delta s,\alpha(t))|/\Delta s=0.\tag{10}$$

如果 Δs 先给定,我们总可以写成(7)和(8)式的形式,但不能保证(10)式对 ω 一致成立. 一致性遭到破坏的原因在于两个相邻事件的时间间隔可能小于 Δs(导致 $x(t)$ 扰动产生的不充分或扰动消失的不完全). 令 $\Omega=\{\omega|\lim_{\Delta s\to 0}|0(\Delta s,\alpha(t))|/\Delta s=0\}$,记 $\bar{\Omega}$ 为 Ω 的补集,则有 $\lim_{\Delta s\to 0}P\{\omega|\omega\in\bar{\Omega}\}=P\{\omega|\omega\in S\}=0$. 另一方面,尽管(10)式的一致性遭到破坏,注意到 $|\Delta x_i|<\text{const}\cdot\Delta s$ 对 ω 一致成立,所以对 $N<\infty$ 和不满足条件 S 的 ω 一致地有 $|0(\Delta s,\alpha(t))|/\Delta s\leqslant C$,其中 C 为一个常数. 于是

$$\begin{aligned}\lim_{\Delta s\to 0}E\frac{|0(\Delta s,\alpha(t)|}{\Delta s}&=\lim_{\Delta s\to 0}E\frac{|0(\Delta s,\alpha(t)|}{\Delta s}\Big|_{\Omega}+\lim_{\Delta s\to 0}E\frac{|0(\Delta s,\alpha(t)|}{\Delta s}\Big|_{\bar{\Omega}}\\&\leqslant 0\cdot P\{\omega|\omega\in\Omega\}+C\cdot\lim_{\Delta s\to 0}P\{\omega|\omega\in\bar{\Omega}\}=0.\end{aligned}\tag{11}$$

证毕.

固定参数 s,让 h 扰动为 $h+\Delta h$,其中 Δh 为充分小的正数. 定义 I 为从机器首次发生

阈值阻塞直到 T 的区间. 记 $l(I)=\int_0^T i\{t\in I\}\mathrm{d}t$, 其中 $i\{\cdot\}$ 为示性函数. 类似可得到如下性质 2 和引理 4.

性质 2. 从机器首次发生阈值阻塞直到 T 的时间区间内 $x(t)$ 扰动量都为 Δh, 其它时间为 0.

引理 4. 系统(1),(2)在 PHP 策略控制下,有

$$L_T(h+\Delta h,s,\alpha(t))-L_T(h,s,\alpha(t))=\frac{\partial L_T(h,s,\alpha(t))}{\partial h}\cdot\Delta s+0(\Delta h,\alpha(t))\quad a.s.,\tag{12}$$

其中

$$\frac{\partial L_T(h,s,\alpha(t))}{\partial h}=\frac{1}{T}[c^+l(I^+)-c^-l(I^-)]\quad a.s.,\qquad \lim_{\Delta h\to 0}E\frac{|0(\Delta h,\alpha(t)|}{\Delta h}=0.\tag{13),(14}$$

定理 1 的证明. 将(8)式两边同除以 Δs, 取数学期望, 再让 $\Delta s\to 0$, 立即得到结论(a); 而(b)可由(7)式直接得到. 对(12),(13)式类似处理可证关于控制参数 h 的结论.　证毕.

设 $\alpha_1(t),\alpha_2(t),\cdots,\alpha_k(t)$ 为 $\alpha(t)$ 的一系列采样, 用蒙德卡罗法作估计 $\frac{\partial\bar{J}_T}{\partial\theta}=\frac{1}{k}\sum_{i=1}^{k}\frac{\partial L_T(\theta,\alpha_i(t))}{\partial\theta}$, 由强大数定律和定理 1 得 $\lim_{k\to\infty}\partial\bar{J}_T/\partial\theta=E\partial L_T(\theta,\alpha(t))/\partial\theta=\partial J_T/\partial\theta$ $a.s.$. 因此 $\partial\bar{J}_T/\partial\theta$ 是 $\partial J_T/\partial\theta$ 的无偏估计. 既然得到了目标 函数对于控制参数的梯度估计, 就可以用随机逼近技术优化参数. 一般的随机逼近最优化算法如下[5]:

设 $J_T(\theta)=EL_T(\theta,\alpha(t))$, $\theta_{k+1}=\theta_k-\beta_k d_k$. 其中 d_k 为 $\nabla J_T(\theta_k)$ 的一个估计; $\{\beta_k\}$ 为一列非负的步长系数, 当 $k\to\infty$ 时, 有 $\beta_k\to 0$. 对于本文给出的系统, 用扰动分析法得到了 $\nabla J_T(\theta_k)$ 的一致无偏估计, 即 $d_k=\nabla\bar{J}_T=(\partial\bar{J}_T/\partial h,\partial\bar{J}_T/\partial s)$ 就可以逐步优化参数 h,s, 实现 PHP 策略.

4　仿真例子和结论

采样数 $k=100$, $T=360$, $U=2$, $d=1$, $c^+=1$, $c^-=5$. 记 $\overline{J_T}=\frac{1}{k}\sum_{i=1}^{k}L_T(\theta,\alpha_i(t))$. 则有表 1 仿真结果.

表 1　PA 法优化 PHP 和 HP 仿真结果

例	PHP					HP		
	s	h	$\partial\bar{J}_T/\partial s$	$\partial\bar{J}_T/\partial h$	$\bar{J}_T$	h	$\partial\bar{J}_T/\partial h$	$\bar{J}_T$
1	8.33	0	−0.0080	0.6755	0.2770	1	0.0139	1.1625
2	16.5	1	−0.0069	−0.0356	1.8029	1.3	0.0179	1.8437
3	7.85	0	0.0081	0.5368	0.7684	1.1	0.0128	1.5198
4	$+\infty$	1.7	—	−0.0051	3.3431	1.7	−0.0051	3.3431

注. 表中例 1 为 $\xi=10,\eta=2$ 的确定性分布. 例 2 为 $\xi-U(0,20)$, $\eta-U(0,4)$ 的均匀分布. 例 3 为截尾正态分布, 即 $\xi'-N(10,1)$, $\eta'-N(2,1)$ 为正态分布. 其中 ξ,η 定义为 $\xi=0$, 当 $\xi'<0$; $\xi=\xi'$, 当 $\xi'\in[0,$

20];$\xi=20$,当 $\xi'>20$. $\eta=0$,当 $\eta'<0$;$\eta=\eta'$,当 $\eta'\in[0,4]$;$\eta=4$,当 $\eta'>4$. 例 4 为 $\xi-0.1e^{-0.1t}$,$\eta-0.5e^{0.5t}$ 的指数分布,此时 PHP 退化成 HP 策略,即 $s=+\infty$.

仿真结果表明,PHP 策略明显优于 HP 策略,特别是当系统的随机性不太严重时,如例 1 和例 3. 用 PA 法估计参数灵敏度,然后优化控制参数的方法,广泛适用于一般的随机系统,而且比单纯的仿真法节约了大量的工作量,但 PA 法必须要求 $T<\infty$.

参 考 文 献

[1] Kimemia J G, Gershwin S B. An algorithm for the computer control of production in flexible manufacturing systems. *IIE Trans.*, 1983, **15**(4):353－362

[2] Bielicki T, Kumar P R. Optimality of zero-inventory policy for manufacturing systems. *Operations Research*, 1988, **36**(4):532－541.

[3] Sharifnia A. Production control of a manufacturing system with multiple machine states. *IEEE Trans.*, *Automat. Contrl.* 1988, **33**(7):620－625

[4] Ho Y C, Cao X R. Pterturbation analysis of discrete-event dynamic systems, Boston:Kluwer, 1991.

[5] Rubinstein R Y. Monte carlo optimization, simulation and sensitivity of queuing networks. Wilay & Sons, 1986.

PREVENTIVE HEDGING POINT CONTROL AND ITS REALIZATION

TU FENGSHENG

(*Dept. of Computer and System Science, Nankai Univ.*, *Tianjin* 300071)

SONG DONGPING

(*National Lab. of Ind. Control Tech.*, *Zhejiang Univ.*, *Hangzhou* 310027)

SHELDON XC LOU

(*California State University*)

Abstract We consider a failure-prone machine whose up and down times follows general probability distributions. A policy called preventive hedging point control, as well as a parameter optimization algorithm, is proposed. Using perturbation analysis technique, the sensitivity estimate of the performance index with respect to control parameters is derived, and it is proved that the estimate is unbiased. Then, the stochastic approximation method is applied to determine the optimal control parameters. Simulations examples show the effectiveness of the policy.

Key words Manufacturing system, perturbation analysis, unbiased estimate, stochastic approximation.

涂奉生 简介见本刊第 21 卷第 6 期.

宋东平 1967 年生. 1992 年在南开大学计算机与系统科学系获硕士学位,现为浙江大学工业控制技术国家重点实验室讲师,在职博士生. 感兴趣的研究方向有柔性生产系统、离散事件动态系统.