Parameter Optimization of a Control Policy for Unreliable Manufacturing Systems *

SONG Dongping
(Institute of Industrial Control Technology, Zhejiang University · Hangzhou, 310027, PRC)

TU Fengsheng
(Depantent of Computer and System Sciences, Nankai University · Tianjin, 300071, PRC)

Sheldon X. C. Lou
(California State University, San Marcos, USA)

Abstract: We study a tandem two machine system producing one part type. The machines are subject to failures. A policy called surplus control which features two threshold levels, each for one machine, is used to regulate the production. We propose a simple algorithm based on perturbation analysis techniques to estimate the gradient of a cost functional. It is proved that this estimate is unbiased. Examples are given to demonstrate the algorithm.

Key words: manufacturing system; perturbation analysis; unbiased estimate; threshold control

1 Introduction

Threshold control policies have been proven to be effective in regulating production of manufacturing systems with unreliable machines[1~3]. An important feature of such policies is simplicity. To optimize the system performance, one only has to calculate the optimal threshold values and this is done "off-line". The "on-line" control actions are simple: when certain variables that represent the system states, e. g. Work-In-Process (WIP) levels at different stages, exceed their corresponding threshold values, the production of those stages should stop, otherwise, they should work as hard as they can. Kanban system is a noted representative.

In this paper, we consider one of such policies dubbed surplus control proposed in [4]. Under this policy, the states of a manufacturing system are surplus levels, as defined in the next section. To select the optimal threshold values so that a cost functional can be minimized. one has to calculate the gradient of the cost functional with respect to the threshold values. The gradient may then be used to minimize the cost functional. We propose a very simple algorithm based on perturbation analysis techniques[5,6] to estimate the gradient. We show that our estimate is unbiased. That is, The expectation of the estimate is equal to the true gradient.

* This research is partly supported by NSFC, "863" Project in China and URIF, NSERC and CIDA.

Part of this paper appeared the Proceedings on the 31th IEEE CDC.

Manuscript received Jul. 24, 1995, revised May 15, 1996.

2 The Model and the Policy

We consider a tandem two machine system producing one product (see Fig. 1). The machines are unreliable, each having two states, up and down. Two buffers, one placed after the first machine M_1 and the other after M_2, are used to store parts. The production rates u_1 and u_2 are control variables that have to be properly regulated to minimize a cost functional defined later in this section.

Fig. 1 A tandem two machine system

The system can be described by the following equations

$$\dot{x}_i(t)=u_i(t)-d, \quad i=1,2, \tag{1}$$

$$x_1(t)\geqslant x_2(t), \tag{2}$$

$$0\leqslant u_i(t)\leqslant U_i\alpha_i(t), \quad i=1,2. \tag{3}$$

The machines are unreliable. The state of $M_i, i=1,2$, is described by $\alpha_i(t)$, a finite state Markov process called machine process. $\alpha_i(t)=1$ if M_i is up and $\alpha_i(t)=0$ if M_i is down. The up times and down times are exponentially distributed with average up time being $1/p_i$ and average down time $1/r_i$ for $M_i, i=1,2$. Denote $\alpha(t):=(\alpha_1(t),\alpha_2(t))$. The production rate of machine M_i is $u_i(t)$ and can be adjusted to any value between 0 and U_i if M_i is up. For the simplicity, we assume $U_1=U_2=U$. Further, d is a constant demand rate. $x_1(t)$ and $x_2(t)$ are called surpluses for M_1 and M_2 respectively. They are defined, as seen from Eqs. (1), as the difference between the cumulative production at M_i and the cumulative demand for the system, $i=1,2$. The initial conditions are $x_1(t)=x_2(t)=0$ and $\alpha_1(t)=\alpha_2(t)=1$. The objective is to minimize the average cost during a finite period T, as defined in the following equations

$$J_T=E\,\frac{1}{T}\int_0^T[c_1b(t)+c_2^+x_2^+(t)+c_2^-x_2^-(t)]\mathrm{d}t$$

where $b(t)=x_1(t)-x_2(t)$, $x_2^+(t)=\max(x_2(t),0)$ and $x_2^-(t)=\max(-x_2(t),0)$. Constants c_1 and c_2^+ are inventory cost coefficients for M_1 and M_2 respectively, and c_2^- is the backlog cost coefficient for M_2.

Intuitively, our objective is to adjust dynamically the production rates u_1 and u_2 (of course only when the machines are up) so that both inventory and backlog cost (negative x_2) can be minimized. The control policy is called surplus control and is defined as[4]:

$$u_1(t)=\begin{cases}U_1, & x_1(t)<h_1,\\ d, & x_1(t)=h_1,\\ 0, & x_1(t)>h_1,\end{cases}\qquad u_2(t)=\begin{cases}U_2, & x_2(t)<h_2, b(t)>0,\\ d, & x_2(t)=h_2, b(t)>0,\\ 0, & x_2(t)>h_2,\\ u_1(t), & x_2(t)\leqslant h_2, b(t)=0,\end{cases}$$

where h_i is a predetermined threshold, $i=1,2$.

Apparently, after the SC policy is selected, to optimize the system performance, i.e., to minimize J_T, the only thing one needs to do is to select the optimal threshold levels h_1 and h_2. As we explained earlier, the purpose of this paper is to derive the gradient of J_T with respect to h_1 and h_2 using perturbation analysis (PA) technique. One may then use the gradient to calculate the optimal threshold values.

3 Perturbation Analysis

Since the initial condition is given ($x_1(0)=x_2(0)=0$, $\alpha_1(0)=\alpha_2(0)=1$), the sample path $x(t)=(x_1(t), x_2(t))$ is uniquely determined by the machine process $\alpha(t)$ and the threshold levels $h=(h_1,h_2)$. Such a sample path is shown in Fig. 2, where $q_0, q_1, \cdots$ and $s_0, s_1, \cdots$ denote the transition instances of $\alpha_1(t)$ and $\alpha_2(t)$ respectively. Each such transition is called an event and $t_0, t_1, \cdots$ denote the instances when events take place.

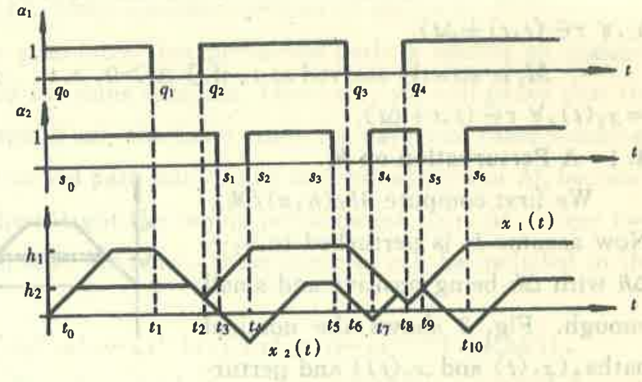

Fig. 2 Sample paths

The PA approach is to observe one sample path (called nominal path), which is obtained either from experiment or simulation under a set of parameter values (in this paper, h_1 and h_2) and to derive the gradient of the cost functional. Let us denote the cost functional obtained from a single sample path as $L_T(h,\alpha)$ i.e. $L_T(h,\alpha)=\frac{1}{T}\int_0^T[c_1b(t)+c_2^+x_2^+(t)+c_2^-x_2^-(t)]dt$. Thus, $J_T(h)=EL_T(h,\alpha)$. As stated earlier, we would like to compute $\partial J_T(h)/\partial h$. We will show that based on a single sample path, a simple algorithm which consists of integrals of indicator functions will provide $\partial L_T(h,\alpha)/\partial h$ which will serve as the estimate of $\partial J_T(h)/\partial h$. Then, one important issue is to show that this estimate is unbiased[7]. It will be shown that this is indeed true for our system. That is

Theorem 1 For the system described by Eqs. (1)~(3) with SC policy, we have

$$E\frac{\partial L_T(h,\alpha)}{\partial h}=\frac{\partial}{\partial h}EL_T(h,\alpha) \text{ and } E\left|\frac{\partial L_T(h,\alpha)}{\partial h}\right|<\infty.$$

The proof is given in the next section.

4 The Proof of Theorem 1

For any fixed T, there exists an event number N, s.t. $t_N\leqslant T\leqslant t_{N+1}$. Apparently, N is a non-negative random variable depending on machine process $\alpha(t)$. Denote the number of state changes of M_i in $[0,t_N]$ by N_i. Let $\xi_1^{i+1}=q_{2i+1}-q_{2i}$, $\eta_1^{i+1}=q_{2i+2}-q_{2i+1}, i\geqslant 0$, $2i+2\leqslant N_1$; $\xi_2^{i+1}=s_{2i+1}-s_{2i}$, $\eta_2^{i+1}=s_{2i+2}-s_{2i+1}, i\geqslant 0, 2i+2\leqslant N_2$. So $\xi_i^{i+1}\sim\xi_i$ and $\xi_i^{i+1}\sim\eta_i$, where ξ_i and η_i are the random variables describing the up times and the down times for $M_i, i=1,2$. Assume $\{x^0,x^1,\cdots,x^N\}$ be the sequence of system state with respect to $\{t_0,t_1,\cdots,t_N\}$. Denote $v=U-d$, then

Lemma 1 For the random variable N s.t. $EN<\infty$.

Proof Because $\alpha_i(t)$ is an alternative renewed process, T is finite, $0<E\xi_i, 0<E\eta_i, i=1,2$, by renewed theory[8], the result is clear. Q.E.D.

To further facilitate the proof of Theorem 1, we need the following definitions:

- $x_i(t)$ reaches h_i at t: if $x_i(t)=h_i$ and $\exists\ \Delta t>0$, s. t. $x_i(\tau)<h_i, \forall\ \tau\in(t-\Delta t,t)$.
- M_2 is starved at t: if $x_1(t)=x_2(t)$ and $\exists\ \Delta t>0$, s. t. $x_1(\tau)>x_2(t), \forall\ \tau\in(t-\Delta t,t)$.
- $x_i(t)$ strictly reaches h_i at t: if $\exists\ \Delta t>0$, s. t. $x_i(\tau)<h_i, \forall\ \tau\in(t-\Delta t,t)$ and $x_i(\tau)=h_i, \forall\ \tau\in(t,t)+\Delta t)$.
- M_2 is strictly starved at t: if $\exists\ \Delta t>0$, s. t. $x_1(\tau)>x_2(t), \forall\ \tau\in(t-\Delta t,t)$ and $x_1(\tau)=x_2(t), \forall\ \tau\in(t,t+\Delta t)$.

4.1 A Perturbation on h_2

We first compute $\partial J_T(h,\alpha)/\partial h_2$. Now assume h_2 is perturbed to $h_2+\Delta h$ with Δh being positive and small enough. Fig. 3 shows the nominal paths, ($x_1(t)$ and $x_2(t)$) and perturbation path $x_2'(t)$ with such a perturbation and some machine process $\alpha^0(t)$ s. t. $N<\infty$. In denoted by dotted curve if it deviates from $x_2'(t)$. We are interested in those time intervals in which the nominal and the perturbed paths do not overlap. Apparently, a perturbation on h_2 does not affect $x_1(t)$. Denote those time intervals in which $x_2(t)\neq x_2'(t)$ by $\varphi_i, i=1,2,\cdots$. It is easy to observe that φ_i starts when $x_2(t)$ strictly reaches h_2 and terminates when M_2 is starved. Denote $I_2=\{t\,|\,t\in\varphi_i, i=1,2,\cdots\}$, $I_2^+=\{t\,|\,t\in\varphi_i$ and $x_2(t)\geqslant 0\}$, $I_2^-=I-I_2^+$; $n_2=$ the number of intervals in I_2, $n_2^-=$ the number of intervals in I_2^-, $n_2^{IS}=$ the times that M_2 becomes starved from idle state (M_2 is up but not operation) in I_2, $n_2^{OS}=$ the times that M_2 becomes starved from operation state in I_2; $l(I_2)=\int_0^T l\{t\in I_2\}\mathrm{d}t$, where $l\{\cdot\}$ is an indicator function. We then have

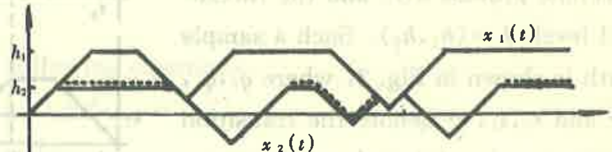

Fig. 3 Nominal and perturbed paths

$$\begin{aligned}&L_T(h_1,h_2+\Delta h,\alpha^0(t))\\&=L_T(h_1,h_2,\alpha^0(t))+\frac{1}{T}[-l(I_2)c_1+l(I_2^+)c_2^+-l(I_2^-)c_2^-]\Delta h+0(\Delta h,\alpha^0(t)).\end{aligned}\tag{4}$$

where

$$0(\Delta h,\alpha^0(t))=\frac{\Delta h^2}{T}\left[(c_1-c_2^+)\left(\frac{n_2}{2v}+\frac{n_2^{IS}}{2d}+\frac{n_2^{OS}}{2U}\right)+(c_1+c_2^+)\left(\frac{\bar{n}_2^-}{2v}+\frac{n_2^-}{2d}\right)\right].\tag{5}$$

with $\bar{n}_2^-=n_2^-$ or n_2^--1 depending upon the value of $\dot{x}_2(T)$. Denote

$$\frac{\partial L_T(h_1,h_2,\alpha^0(t))}{\partial h_2}=\frac{1}{T}[-l(I_2)c_1+l(I_2^+)c_2^+-l(I_2^-)c_2^-].\tag{6}$$

Apparently, (6) is independent of Δh. Now (4) can be written as

$$L_T(h_1,h_2+\Delta h,\alpha^0(t))-L_T(h_1,h_2,\alpha^0(t))=\frac{\partial L_T(h_1,h_2,\alpha^0(t))}{\partial h_2}\Delta h+0(\Delta h,\alpha^0(t))\tag{7}$$

Lemma 2 The system (1)~(3) with the perturbation of h_2, we have

$$L_T(h_1,h_2+\Delta h,\alpha(t))-L_T(h_1,h_2,\alpha(t))=\frac{\partial L_T(h_1,h_2,\alpha(t))}{\partial h_2}\Delta h+0(\Delta h,\alpha(t)),\ \text{a. s.}\tag{8}$$

where

$$\frac{\partial L_T(h_1,h_2,\alpha(t))}{\partial h_2}=\frac{1}{T}[-l(I_2)c_1+l(I_2^+)c_2^+-l(I_2^-)c_2^-],\quad \text{a. s.}\tag{9}$$

$$\lim_{\Delta h\to 0} E\frac{|0(\Delta h,\alpha(t))|}{\Delta h}=0. \tag{10}$$

Proof For any machine process $\alpha(t)$ s. t. $N<\infty$, we can write the formulas such as (8) and (9). Now we should prove (10). Only consider samples of $\alpha(t)$ s. t. $N<\infty$.

Now that Δh is given, we cannot guarantee that perturbed path is similar to nominal path. That means Eq. (7) may not hold for some samples. However, we will prove that the probability of these samples trends to zero if $\Delta h\to 0$. Let's study the nature of these samples, if M_2 becomes starved some time in perturbed path but it isn't in nominal path or M_2 becomes down while it is being perturbed but hasn't got the whole perturbation gain Δh, these two cases both will result in failure of Eqs. (5). Anyway, these samples can be included in the following set

$$\Omega_1=\{\omega|\alpha(t) \quad \text{s. t.} \quad t_i-t_{i-1}\in f_i(x^0,x^1,\cdots,x^{i-1}), f_i(x^0,x^1,\cdots,x^{(i-1)})+R\Delta h)\},$$

where R is a constant and f_i is a linear function which depends on the history of $(x(t),\alpha(t)$. If $\omega\notin\Omega_1$, then $0(\Delta h,\alpha(t))$ in (8) is something such as Eq. (5), using $n_2\leqslant N$ and $n_2^-\leqslant N$, we have: $|0(\Delta h,\alpha(t))|\leqslant C\cdot\Delta h^2\cdot N$, where C is a constant which is independent on $\alpha(t)$. By Lemma 1, that yields

$$\lim_{\Delta h\to 0} E\frac{|0(\Delta h,\alpha(t))|}{\Delta h}\Bigg|_{\bar\Omega_1}=0. \tag{11}$$

If $\omega\in\Omega_1$ then $|0(\Delta h,\alpha(t))|\leqslant C\Delta h^2N+\bar C\Delta h$, where $\bar C$ is also a constant. That following

$$E\frac{|0(\Delta h,\alpha(t))|}{\Delta h}\Bigg|_{\Omega_1}=C\cdot\Delta h\cdot EN+\bar C\cdot P(\Omega_1). \tag{12}$$

Recall the assumption that ξ_i and η_i both are exponentially distributed random variables, and distribution functions are consistent right continuous, thus, $\lim_{\Delta h\to 0}P(\Omega_1)=0$. We then get

$$\lim_{\Delta h\to 0} E\frac{|\Delta h,\alpha(t))|}{\Delta h}\Bigg|_{\Omega_1}=0. \tag{13}$$

From (11) and (13), Lemma 2 holds.

The proof of the theorem 1 Take the expectation of both sides of Eqs. (8), then let $\Delta h\to 0$, from Eqs. (9) and (10), we obtain

$$E\frac{\partial L_T(h,\alpha)}{\partial h_2}=\frac{\partial}{\partial h_2}EL_T(h,\alpha) \text{ and } E\left|\frac{\partial L_T(h,\alpha)}{\partial h_2}\right|\leqslant c_1+c_2^++c_2^-<\infty. \tag{14}$$

4.2 A Perturbation h_1

Now we turn to $\partial L_T(h,\alpha)/\partial h_1$. Similar to the case of h_2, the perturbed paths $x_1'(t)$ and $x_2'(t)$ can be constructed from the nominal paths $x_1(t)$ and $x_2(t)$, as shown in Fig. 4.

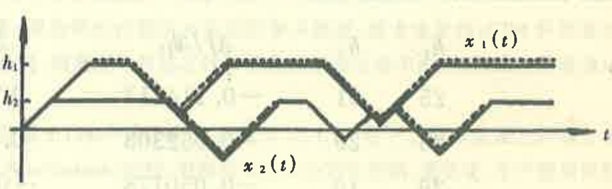

Fig. 4 Sample paths when h_1 is perturbed

Denote $I_1=\{t\,|\,x_1'(t)\neq x_1(t)\}$. It can be easily seen that I_1 is the time interval from the instance when $x_1(t)$ first strictly reaches h_1 to T. Also denote those intervals in which $x_2(t)\neq x_2'(t)$ by $\varphi_i, i=1,2,\cdots$. An interval φ_i for some i starts when $x_1(t)$ is strictly starved and ends when $x_2(t)$ reaches h_2 and $t\in I_1$. Define $I_2=\{t|t\in\varphi_i, i=1,2,\cdots\}$, I_2^+, I_2^-, $l(I_2)$, $l(I_1)$, $l(I_2^+)$ and $l(I_2^-)$ are similarly defined as in the precious section. We have

Lemma 3 The system (1)~(3) with h_1 perturbed, we have

$$L_T(h_1+\Delta h,h_2,\alpha(t))-L_T(h_1,h_2,\alpha(t))=\frac{\partial L_T(h_1,h_2,\alpha(t))}{\partial h_1}\Delta h+0(\Delta h,\alpha(t)),\quad \text{a. s.} \tag{15}$$

where

$$\frac{\partial L_T(h_1,h_2,\alpha(t))}{\partial h_1}=\frac{1}{T}[l(I_1)c_1-l(I_2)c_1+l(I_2^+)c_2^+-l(I_2^-)c_2^-],\quad \text{a. s.} \tag{16}$$

$$\lim_{\Delta h\to 0}E\frac{|\Delta h,\alpha(t))|}{\Delta h}=0. \tag{17}$$

The proof is similar to that of Lemma 2 and is omitted.

By Lemma 3, the assertion of Theorem 1 for h_1 holds. Thus, we have completed the proof of Theorem 1.

5 Numerical Examples

We now use Monte Carlo technique to generate a sequence of machine processes $\alpha^1(t)$, $\alpha^2(t),\cdots,\alpha^k(t)$. The strong law of large number and our theorem yield

$$\lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\frac{\partial L_T(h,\alpha^i)}{\partial h}=E\frac{\partial L_T(h,\alpha)}{\partial h}=\frac{\partial J_T}{\partial h},\quad \text{a. s.}$$

Thus, $\frac{\partial J_T}{\partial h}:=\frac{1}{k}\sum_{i=1}^{k}\frac{\partial L_T(h,\alpha^i)}{\partial h}$ is an unbiased estimate of $\frac{\partial J_T}{\partial h}$. Denote $\frac{\partial J}{\partial h}=\left(\frac{\partial J}{\partial h_1},\frac{\partial J}{\partial h_2}\right)$, where $\frac{\partial J}{\partial h_j}=\frac{1}{k}\sum_{i=1}^{k}\frac{\partial L_T(h,\alpha^i)}{\partial h_j}$, $j=1,2$ and $\bar{J}=\frac{1}{k}\sum_{i=1}^{k}L_T(h,\alpha^i)$. We now present two examples.

Example 1 Let $U=2$, $d=1$, $c_1=1$, $c_2^+=2$, $c_2^-=10$, $p_1=p_2=0.1$, $r_1=r_2=0.5$, $T=360$ and $k=100$. $\xi_i\sim 1/p_i e^{-p_i t}$, $\eta_i\sim 1/r_i e^{-r_i t}$, $i=1,2$. We obtain the following simulation results,

h_1	h_2	$\partial J/\partial h_1$	$\partial J/\partial h_2$	$\bar{J}$
4	3.6	−0.091701	0.024428	18.222662
4.3	3.6	0.025260	0.023551	18.213660
4.2	3.5	−0.011006	−0.002164	18.211904

Example 2 Let $r_1=r_2=0.2$, the other parameters remain the same as in Ex. 1. Then the simulation results are

h_1	h_2	$\partial J/\partial h_1$	$\partial J/\partial h_2$	$\bar{J}$
25	21	−0.114213	0.091016	77.607500
26	20	−0.052303	0.080445	77.392441
26	19	−0.050178	−0.003604	77.365894

The above examples show that the gradient estimates indeed lead to the minimization of the cost functional.

Remark It is not difficult to extend the results presented in this paper to N-machine serial production line. The only difference is on the rules of perturbation generation, propagation and change. In addition, the assumption of exponential distributions of machine failure and repair isn't necessary.

References

1 Akella, R. and Kumar, P. R.. Optimal control of production rate in failure prone manufacturing system. IEEE Trans. Automat. Contr., 1986, AC-31(2):116—126

2 Kimemia, J. G. and Gershwin, S. B.. An algorithm for the computer control of production in flexible manufacturing system. IIE Trans., 1983, 15(4):353—362

3 Sheldon, X. C. Lou and Ryzin, G. V.. Optimal control rules for scheduling job shops. Annals of Operation Research, 1989, 17:233—248

4 Yan Hourmin and Sheldon X. C. Lou. Surplus production control and its realization. Proceedings of ASAC, 1991

5 Ho. Y. C. and Cao, X. R.. Perturbation analysis of discrete event dynamic systems. Kluwer Academic Publisher, Norwell, MA, USA, 1991

6 Glasserman, P.. Gradient estimation via perturbation analysis. Kluwer Academic Publisher, Norwell, MA, USA, 1991

7 Cao, X. R.. Convergence of parameter sensitivity estimates stochastic experiment. IEEE Trans Automat. Contr., 1985, AC-30(9):845—953

8 Beaumont, G. P.. Introductory applied probability. Ellis Horward, 1983

不可靠制造系统的一种控制策略的参数最优化

宋东平

（浙江大学工业控制技术研究所·杭州，310027）

涂菶生 Sheldon X. C. Lou

（南开大学计算机与系统科学系·天津，300071） （加拿大多伦多大学）

摘要：本文研究生产单类工件的两台串级机器的系统，机器是不可靠的．系统采用 Surplus Control 策略，该策略以每台机器的一个阈值作为控制参数．基于扰动分析方法，我们提出了一种简单的算法用于估计系统费用函数对于控制参数的梯度，并证明了估计的无偏性，仿真例子验证了算法的有效性．

关键词：制造系统；扰动分析；无偏估计；阈值控制

本文作者简介

宋东平 1967年生．1989年在南开大学数学系获学士学位，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．感兴趣的研究方向有柔性生产系统、离散事件动态系统．

涂菶生 1937年生．1961年毕业于南开大学数学系，现为该校计算机与系统科学系教授，博士生导师，CIM 研究室主任．已发表论文50多篇．主要研究方向为离散事件动态系统、调度理论及其在计算机集成制造系统中的应用，线性系统理论和计算机控制．

Sheldon X. C. Lou 毕业于清华大学电子工程系，分别于1982年和1985年在美国 M. I. T. 电子工程系获硕士和博士学位．现为美国工商管理学院 California State University (San Marcos)教授．主要研究方向为随机控制、最优化、生产规划和控制、系统故障检测和辨识等．