

输入随机的不可靠串行生产线的反馈控制*

张 伟 孙优贤

(浙江大学工业控制技术研究所 杭州 310027)

摘 要 对于具有随机输入和随机需求的一类串行生产线系统,利用转移率一致化技术和随机动态规划方法,给出了输入率和服务率分配的最优反馈控制策略,指出系统的最优控制具有邦-邦形式的开关结构。仿真例子验证了方法的有效性。

关键词 串行生产线,一致化技术,随机动态规划

分类号 O 221

Feedback Control of Serial Production Line with Unreliable Workstations and Random Input

Zhang Wei, Sun Youxian

(Zhejiang University)

Abstract The considered serial production line consists of possibly unreliable workstations, whose inputs may be random. The optimal input and service feedback control policy are derived by using the uniformization technique and stochastic dynamic programming. It is shown that the optimal policy is of bang-bang type. Simulation examples are given to illustrate the results.

Key words serial production line, uniformization technique, stochastic dynamic programming

1 引 言

不可靠制造系统的最优控制问题是生产规划和管理中重要的研究问题之一。Gershwin 做了大量的开创性研究工作,提出了流率控制模型(Flow rate control model)、阈值(Hedging point)控制策略、递阶结构框架^[1]。Kumar 最先建立了严格的数学分析方法,具有重要的理论意义^[2]。Sethi 在实时递阶调度和渐近最优反馈控制等方面取得了重要成果^[3]。

串行生产线是一类应用广泛的基本而典型的制造系统。它是由一组工作站和一组内部缓冲区间隔连接组成的线形网络。文献[4]针对机器不可靠、随机加工时间、随机产品需求等多种不确定因素并存的串行生产线系统,给出了一种基本的数学描述和分析方法,并给出了最优反馈控制策略。该方法的创新之处在于能够保持系统的离散事件本质,进行实时动态最优控制分析。但该文假定系统具有无限的

工件源,从而限制了该方法的实际应用。在实际生产中,工作源的输入往往是随机的或受到某些条件限制的,无限的工作源会造成生产成本的浪费。针对这一问题,本文考虑了具有随机工件输入的一类串行生产线系统,给出了输入率和服务率的最优反馈控制策略。

2 问题描述

考虑输入随机的一类串行生产线,它是由一组工作站($WS_1, WS_2, \dots, WS_n$)和一组内部缓冲区间($B_1, B_2, \dots, B_n$)间隔连接而组成,如图 1 所示。其中方框代表工作站,圆形代表缓冲区。工件从外部到达工作站 WS_1 ,完成加工后进入缓冲区 B_1 ,然后在工作站 WS_2 上加工,如此前进,直到在 WS_n 上完成最后一道工序变成产品存放在缓冲区 B_n 。当需求未满足时,产品就离开缓冲区 B_n 。

2.1 假设条件

假设系统满足如下条件:

1) 串行生产线的工件源输入是输入率为 k 的

* 国家自然科学基金项目(69635010)

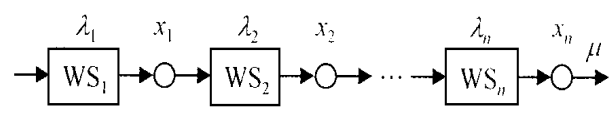


图 1 串行生产线

poisson 过程, 其中输入率 k 是可以调节的, $k \in [0, \bar{K}]$ 。

2) 工作站 $WS_i (i = 1, 2, \cdots, n)$ 对工件的加工时间是随机的, 服从参数为 λ 的负指数分布, 其中服务率 λ 是可控的, 当工作站正常时, 它可在 $[0, \bar{\lambda}]$ 中任意调整。

3) 产品的需求过程是一个参数为 μ 的 poisson 过程, 即产品的需求间隔时间服从参数为 μ 的负指数分布, 是不可控的。

4) 工作站 $WS_i (i = 1, 2, \cdots, n)$ 具有 Markov 的故障和修复过程, 故障率和修复率分别为 ξ_i 和 η_i 。

工作站 $WS_i (i = 1, 2, \cdots, n)$ 的状态用 $\alpha(t)$ 表示, $\alpha(t) = 1$ 表示工作站处于正常状态, $\alpha(t) = 0$ 表示工作站处于故障状态。工作站发生故障时不能加工工件, 所以仅当工作站正常时才采取控制决策。根据假设条件 1), 假想在 WS_1 前有一虚拟工作站 WS_0 和虚拟缓冲区 B_0 , 该工作站不会发生故障并且具有无限的工作件源, 工作站对工作件源的服务率即为输入到工作件源库的工作件的输入率 k 。

记 $x_i(t)$ 为 t 时刻缓冲区 $B_i (i = 0, 1, \cdots, n - 1)$ 内的工件数, 显然 $x_i(t)$ 为非负整数, $x_n(t)$ 为工作站 WS_n 在 t 时刻的生产累积量与需求累积量之差, $x_n(t) \in Z$; 正值表示该类产品有库存, 负值表示该类产品欠缺。定义向量

$$X(t) = (x_0(t), x_1(t), \cdots, x_n(t))$$

$$U(t) = (k, \lambda_1, \cdots, \lambda_n)$$

$$\alpha(t) = (\alpha_0(t), \alpha_1(t), \cdots, \alpha_n(t))$$

2.2 控制目标

当工作站正常时, 采取控制决策动态地调整工件的输入速率, 同时动态地决定生产工件的速率, 或者决定是否让工作站空闲。工件的输入速率和工作站的服务率决定了工件的库存大小以及产品的供应情况。工件和产品的库存会造成成本的浪费。我们的目标是使生产尽可能接近需求, 且保持尽量少的中间产品和成品的库存, 使生产费用和库存或缺所造成的成本代价尽可能低。即寻找最优策略, 使得下列费用函数达到最小

$$V(X, \alpha) = \min E \int_0^\infty e^{-\beta t} g(X(t)) dt \quad (1)$$

其中, β 是折扣因子, $g(X(t))$ 是对中间库存、产品库存与欠缺的惩罚函数, 一般取

$$g(X(t)) = c_0 x_0 + c_1 x_1 + \cdots + c_n^+ x_n^+ + c_n^- x_n^-$$

其中 $x_n^+ = \max\{x_n, 0\}$, $x_n^- = \max\{-x_n, 0\}$, c_n^+ , c_n^- 分别是产品库存和欠缺的惩罚系数, c_i 是中间产品库存的惩罚系数, $i = 0, 1, \cdots, n - 1$ 。

定义映射 T_i, D_i 和 R_i 如下

$$T_i \alpha = (\alpha_1, \cdots, \alpha_{i-1}, \sigma, \alpha_{i+1}, \cdots, \alpha_n)$$

$$\sigma_i \equiv \alpha_i + 1 \bmod 2, \quad i = 1, 2, \cdots, n$$

$$R_0 X = (x_0 + 1, x_1, x_2, \cdots, x_n)$$

$$R_i X = (x_0, x_1, \cdots, x_{i-1} - 1, x_i + 1, x_{i+1}, \cdots, x_n) \\ x_{i-1} > 0, \quad i = 1, 2, \cdots, n$$

$$R_i X = X, \quad x_{i-1} = 0, \quad i = 1, 2, \cdots, n$$

$$R_{n+1} X = (x_0, x_1, \cdots, x_{n-1}, x_n - 1)$$

$$D_i X = (x_0, x_1, \cdots, x_{i-1}, x_i - 1, x_{i+1}, \cdots, x_n) \\ i = 0, 1, \cdots, n$$

称 $R_i X$ 和 $T_i \alpha$ 为系统的事件。其中 $R_i X$ 表示一个工件在工作站 $WS_i (i = 0, 1, \cdots, n)$ 上加工的完成, $R_{n+1} X$ 表示一个产品因需求而离开系统, $T_i \alpha$ 表示第 i 个工作站 WS_i 发生故障和修复。显然 $D_n X = R_{n+1} X$ 。

3 最优反馈控制策略

令 $Z_2 = \{0, 1\}$, $I = \{(X, \alpha) | X \in Z^n \times Z, \alpha \in Z_2^n\}$ 为系统的状态空间; (X, α) 为系统的状态; $U = \{u(t) = (k, \lambda_1, \lambda_2, \cdots, \lambda_n) | k \in [0, \bar{K}], \lambda_i \in [0, \bar{\lambda}]\}$ 为容许控制集; $(R_i X, \alpha)$ 和 $(X, T_i \alpha)$ 表示事件发生时的状态转移规律, 是时间离散的且依赖于系统的状态和控制。工作站 WS_i 可控的必要条件是其处于正常状态, 即 $\alpha_i(t) = 1$ 。当且仅当事件 $R_i X$ 或 $T_i \alpha$ 发生时, 系统的状态发生转移, 不同的系统状态下有不同的转移率。仅当系统状态发生转移且工作站正常时才采取控制决策。系统的部分状态转移图如图 2 所示。

下面利用一致化技术^[5], 把上述非一致化转移率的连续时间 Markov 链最优决策问题, 转化成一个等价的具有一致化转移率的离散时间 Markov 最优决策问题。

定义一致化转移率为

$$v = \mu + K + \sum_{i=1}^n (\xi_i + \eta_i + \bar{\lambda})$$

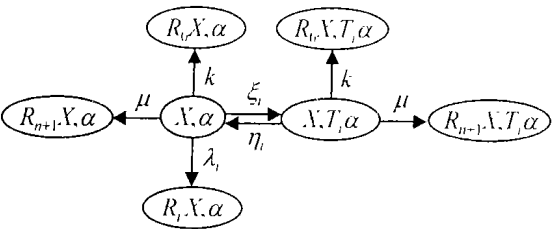


图 2 部分状态转移图

对于任意给定的控制 $u \in U$, 单步转移概率函数 $P(\bullet | \bullet, u)$ 如下

$$\begin{aligned} P(Y|(X, \alpha), u) &= \mu/v \\ Y &= (R_{n+1}X, \alpha) \\ P(Y|(X, \alpha), u) &= k/v \\ Y &= (R_0X, \alpha) \\ P(Y|(X, \alpha), u) &= \lambda_i/v \\ Y &= (R_iX, \alpha), \alpha = 1, x_{i-1} > 0 \\ i &= 1, 2, \cdots, n \\ P(Y|(X, \alpha), u) &= \xi_i/v \\ Y &= (X, T_i\alpha), \alpha = 1 \\ P(Y|(X, \alpha), u) &= \eta_i/v \\ Y &= (X, T_i\alpha), \alpha = 0 \\ P(Y|(X, \alpha), u) &= \\ 1 - (k + \mu + \sum_{i=1}^n (\lambda_i + \xi_i + \eta_i)) / v \\ Y &= (X, \alpha) \\ P(Y|(X, \alpha), u) &, \text{ otherwise} \end{aligned}$$

设 $0 = t_0 < t_1 < \cdots < t_n < \cdots$ 为离散状态转移时刻序列, 记 $X_k = X(t_k)$ 为第 k 次转入的状态, $u_k = u(t_k)$ 为第 k 次采取的控制。则有 $X(t) = X_k; u(t) = u_k$ 当 $t \in [t_k, t_{k+1})$ 。对给定的控制策略, 目标函数可化为

$$\begin{aligned} E^u \int_0^\infty e^{-\beta t} g(X(t)) dt &= \\ E^u \sum_{k=0}^\infty \int_{t_k}^{t_{k+1}} e^{-\beta t} g(X_k) dt &= \\ \frac{1}{\beta + v} \sum_{k=0}^\infty \theta E g(X_k) \end{aligned} \tag{2}$$

其中 $\theta = v/(\beta + v)$ 。于是问题转化为一个等价的离散时间 Markov 链最优决策问题, 并且具有非负无界的单步费用函数和无限可数的状态空间。由此得到如下结论:

定理 1 令 $V_0(X, \alpha) \equiv 0$, 则对于 Bellman 方程 $V_k(X, \alpha)$, 有

$$\lim_{k \rightarrow \infty} V_k(X, \alpha) = V(X, \alpha)$$

$\forall (X, \alpha) \in I = \{(X, \alpha) | X \in Z^+ \times Z, \alpha \in Z^2\}$
其中 $V(X, \alpha)$ 由 (1) 式定义。

证明 由 (2) 式和随机动态规划得 Bellman 方程, 有

$$\begin{aligned} V_{k+1}(X, \alpha) &= \frac{1}{\beta + v} \min_{u \in U} \{g(X) + \mu V_k(D_n X, \alpha) + \\ &k V_k(R_0 X, \alpha) + \sum_{i=1, \alpha_i=1, x_{i-1} > 0}^n \lambda_i V_k(R_i X, \alpha) + \\ &\sum_{i=1, \alpha_i=1}^n \xi_i V_k(X, T_i \alpha) + \sum_{i=1, \alpha_i=0}^n \eta_i V_k(X, T_i \alpha) + \\ &[v - (\mu + k + \sum_{i=1, \alpha_i=1, x_{i-1} > 0}^n \lambda_i + \\ &\sum_{i=1, \alpha_i=1}^n \xi_i + \sum_{i=1, \alpha_i=0}^n \eta_i)] V_k(X, \alpha)\} = \\ &\frac{1}{\beta + v} \{g(X) + \mu V_k(D_n X, \alpha) + \\ &\sum_{i=1, \alpha_i=1}^n [\xi_i V_k(X, T_i \alpha) + \eta_i V_k(X, \alpha)] + \\ &\sum_{i=1, \alpha_i=0}^n [(\xi_i + \bar{\lambda}_i) V_k(X, \alpha) + \eta_i V_k(X, T_i \alpha)] + \\ &K \min[V_k(R_0 X, \alpha), V_k(X, \alpha)] + \\ &\sum_{i=1, \alpha_i=1}^n \bar{\lambda}_i \min[V_k(R_i X, \alpha), V_k(X, \alpha)]\} \end{aligned}$$

其中 $1\{\bullet\}$ 为指示函数。根据上式和文献[5] 的结果可推得定理 1 的结果。

定理 2 系统的最优反馈控制策略 U^* 具有 bang - bang 形式的开关结构, 即对 $(X, \alpha), U^* = (k^*, \lambda_i^*, \cdots, \lambda_n^*)$, 其中

$$\begin{aligned} k^* &= \begin{cases} K, & V(R_0 X, \alpha) \leq V(X, \alpha) \\ 0, & \text{otherwise} \end{cases} \\ \lambda_i^* &= \begin{cases} \bar{\lambda}_i, & V(R_i X, \alpha) < V(X, \alpha) \\ x_{i-1} > 0, & \alpha = 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

由 Bellman 方程表达式 $V_k(X, \alpha)$ 可直接证得。

该定理指出, 系统的最优控制具有 bang - bang 形式的开关结构, 工件的输入率和工作站的服务率分别在最大值和 0 之间进行切换。在实际控制问题中, 可根据定理 1 离线地计算出系统在不同状态下的最优值函数, 然后采用定理 2 所给出的控制策略。另外一种可行的方案是在线地计算 $V_k(X, \alpha), V_k(R_0 X, \alpha)$ 和 $V_k(R_i X, \alpha)$, 并对其进行比较, 然后采用定理 2 中的控制策略。根据上述两个定理可以得出该方案为渐近最优的。

表 1 $k = 40$ 时的部分最优值函数

值函数	(10, 10, - 10)	(2, 2, - 2)	(1, 1, - 1)	(2, 2, 2)	(40, 40, 20)
$V_{40}(X, 1, 1)$	80. 006 1	10. 462 9	3. 067 37	7. 699 86	54. 197 2
$V_{40}(X, 1, 0)$	82. 119 2	12. 040 5	4. 064 67	8. 296 19	63. 534 4
$V_{40}(X, 0, 1)$	80. 980 3	12. 359 3	4. 512 9	8. 978 24	71. 044 3
$V_{40}(X, 0, 0)$	83. 095 8	14. 582 4	6. 332 15	9. 949 73	85. 562 6

4 仿真例子

考虑两台工作站组成的串行生产线系统。状态空间限制在 $0 \leq x_0 \leq 40, -40 \leq x_1 \leq 40, -40 \leq x_2 \leq 40$ 。系统参数取为

$\bar{\lambda}_1 = 0.8, \bar{\lambda}_2 = 0.6, \mu = 0.6, \beta = 0.6$
 $\xi_1 = \xi_2 = 0.2, \eta_1 = \eta_2 = 0.6, \bar{K} = 0.8$
 $c_0 = 2, c_1 = 1, c_2^+ = 1, c_2^- = 2$

对于 $(X, \alpha) = (10, 10, -10, 1, 1)$ 和 $(X, \alpha) = (2, 2, 2, 1, 0)$ 两种状态, 采用后一种控制方案进行仿真, 其结果如图 3 和图 4 所示, 图中曲线分别为两种不同状态下的值函数曲线。从图中可以看出, 当 $k = 40$ 时已经取得了较好的收敛结果, 表 1 示出了当 X 取不同值时的部分最优值函数。

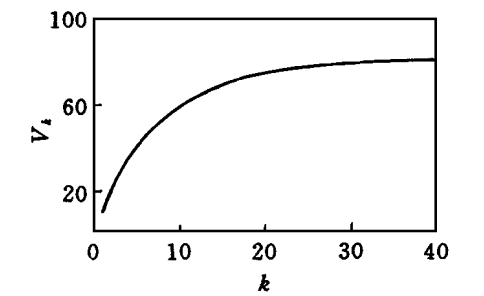


图 3 (10, 10, - 10, 1, 1) 的值函数曲线

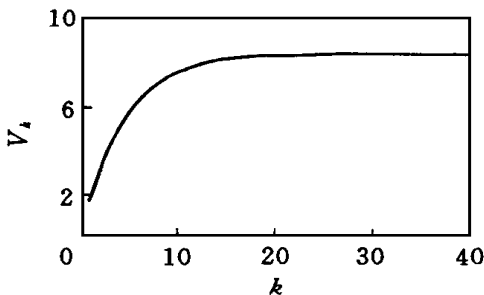


图 4 (2, 2, 2, 1, 0) 的值函数曲线

5 结 语

本文对于机器不可靠、随机加工时间、随机输入和需求等多种不确定因素并存的串行生产线系统, 给出了一种基本的数学描述, 并利用转移率一致化技术和随机动态规划方法, 研究了系统的最优反馈控制问题, 给出了具有开关结构的反馈控制律。该方法能够保持系统的离散事件本质。利用本算法可以在线地决定工件的输入以及加工速率, 合理地决定工件源的储备量, 有效地减少不必要的成本开支。

参 考 文 献

1 Maimon O Z. Gershwin S B. Dynamic scheduling and routing for flexible manufacturing systems that have unreliable machines. Operation Research, 1988, 36(2): 279_ 292

2 Akella R, Kumar P R. Optimal- control of production - rate in a failure prone manufacturing system. IEEE Trans on Automation Control, 1986, 31(2): 116_ 126

3 Sethi S, Zhang Q. Hierarchical production planning in dynamic- stochastic manufacturing systems- asymptotic optimality and error - bounds. J of Mathematical Analysis and Applications, 1994, 181(2): 285_ 319

4 宋东平. 不可靠制造系统的最优反馈控制策略研究. 浙江大学博士学位论文, 1998

5 Bertsekas D P. Dynamic programming: Deterministic and stochastic models. NJ: Prentice- Hall, 1987

作 者 简 介

张 伟 男, 1971 年生。浙江大学工业控制技术研究所博士研究生。研究方向为生产制造系统, 离散事件动态系统和混杂系统。

孙优贤 见本刊 1998 年第 13 卷第 6 期第 680 页。