

6-3-2020

Online Optimal Control Method of Conveyor-serviced Batch Processing System

Wu Song

School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China;

Tang Hao

School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China;

Tan Qi

School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Online Optimal Control Method of Conveyor-serviced Batch Processing System

Abstract

Abstract: In conveyor-serviced batch processing system, jobs arrive dynamically and are stored in a buffer to be further processed by the batch processing machine. Under the cases of non-identical job sizes, online optimal control of the system with a single machine was mainly concerned. *Optimization model of the system was built by using look-ahead range as control variable. The objective is to maximize the job processing rate of the system in infinite horizon. Two decision processes are included in the model, which are batching process and action-selection process. For the batching process, the rule of minimizing process-capacity wasting ratio of the machine during production period was proposed. And for the action-selection process, Q-learning algorithm was employed to derive the optimal look-ahead policy.* The effectiveness of the proposed algorithms was demonstrated through the simulation experiments. Besides, the influence of some batching rules and physical parameters on system performance was showed.

Keywords

batch processing, attribute-different jobs, single machine, online optimization

Recommended Citation

Wu Song, Tang Hao, Tan Qi. Online Optimal Control Method of Conveyor-serviced Batch Processing System[J]. Journal of System Simulation, 2017, 29(4): 730-739.

传送带给料分批加工系统在线优化控制方法

吴松, 唐昊, 谭琦

(合肥工业大学电气与自动化工程学院, 安徽 合肥 230009)

摘要: 传送带给料分批加工系统中, 工件动态到达, 并配置有存放待加工工件的缓冲区, 其加工主体为批处理设备。考虑工件属性差异, 重点研究单机模型的在线优化控制问题。以前视距离为控制变量, 无穷时段内的工件处理率最大为优化目标, 建立了系统的优化模型。针对该模型中的工件分批决策, 提出一种以批处理机加工周期内加工能力浪费比最小为准则的工件分批规则。对于该模型中的行动选择决策, 文中引入 Q 学习优化算法, 以求解最优前视控制策略。通过仿真实验, 对算法的有效性进行了验证, 并分析了不同分批策略及参数对系统性能的影响。

关键词: 批处理; 差异工件; 单机; 在线优化

中图分类号: TP202

文献标识码: A

文章编号: 1004-731X (2017) 04-0730-10

DOI: 10.16182/j.issn1004731x.joss.201704005

Online Optimal Control Method of Conveyor-serviced Batch Processing System

Wu Song, Tang Hao, Tan Qi

(School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

Abstract: In conveyor-serviced batch processing system, jobs arrive dynamically and are stored in a buffer to be further processed by the batch processing machine. Under the cases of non-identical job sizes, online optimal control of the system with a single machine was mainly concerned. Optimization model of the system was built by using look-ahead range as control variable. The objective is to maximize the job processing rate of the system in infinite horizon. Two decision processes are included in the model, which are batching process and action-selection process. For the batching process, the rule of minimizing process-capacity wasting ratio of the machine during production period was proposed. And for the action-selection process, Q -learning algorithm was employed to derive the optimal look-ahead policy. The effectiveness of the proposed algorithms was demonstrated through the simulation experiments. Besides, the influence of some batching rules and physical parameters on system performance was showed.

Keywords: batch processing; attribute-different jobs; single machine; online optimization

引言

半导体芯片制造过程中含有一个老化工序, 即



收稿日期: 2015-06-30 修回日期: 2015-10-30;
基金项目: 国家自然科学基金(61174186, 61374158, 71231004, 61573126, 60404009), 高等学校博士学科点专项科研基金(20130111110007);
作者简介: 吴松(1991-), 男, 安徽六安, 硕士, 研究方向为离散事件动态系统、强化学习。

将芯片放入高温炉中进行预烧, 以检测出有缺陷的芯片。高温炉一次可以预烧多个芯片, 而芯片的预烧时间由芯片类型预先确定。因为老化工序耗时往往是其它工序的几倍至几十倍, 成为整个生产制造过程的瓶颈, 如何提高高温炉利用率对于提高整个生产制造过程的效率至关重要。若将芯片看作工件, 高温炉看作机器, 则该问题就是一个典型的批处理问题。该类问题不仅仅局限于半导体生产领域中,

<http://www.china-simulation.com>

它还广泛存在于现代化生产作业线, 如金属加工、钢铁生产的预加热过程以及生物制药等领域^[1-2]。由于该类问题的重要性和广泛性, 国内外的学者对其进行了深入的研究, 建立了各种环境下的批处理模型, 并设计出了众多求解算法^[3-13]。

文献[3]提出并研究了以极小化最大完工时间为目标, 含有差异工件的单机批处理调度问题, 证明了单机环境下该问题是 NP 难题, 并基于该问题给出了若干个启发式算法, 其中以 FFLPT(First-Fit Longest Processing Time)性能最优。文献[4]对文献[3]的工作进行了扩展, 提出了另外两个启发式算法: SKP(Successive Knapsack Problem)和 BFLPT(Best-Fit Longest Processing Time), 并证明两者的性能要优于 FFLPT。文献[5-6]分别采用模拟退火算法和遗传算法对该问题进行了求解。文献[7]将问题扩展到多客户生产环境, 并提出了相应的启发式算法对问题进行求解。以上研究均是针对事先已知工件的全部参数信息的离线问题。对于工件信息完全未知的情况, 文献[8]利用模型预测控制对该问题的在线优化控制问题进行了研究。文献[9]研究了一类含有不相容工件族的模型, 并给出了该问题相应的启发式算法。文献[10]对包含一台不考虑容量约束的批处理机和一台标准机的并行调度问题进行了研究, 并给出了在线算法。对于工件具有动态到达特性的批处理机调度问题, 文献[11-13]分别采用启发式规则和遗传算法、蚁群算法等元启发式算法对单机和平行机环境下的该类问题进行了求解。

但在实际生产生活中, 涉及到的此类问题, 一般只能预知未来一段时间内的工件信息。如半导体芯片制造过程包括氧化、蚀刻、老化等多道工序, 由于每道工序的加工时间并非完全可知, 导致后道工序只能预知已完成前道工序的工件的信息^[14]。根据这种情况, Matsui 提出了一种传送带给料生产加工站(Conveyor-Serviced Production Station, CSPS)模型^[15], 并将单站点 CSPS 系统抽象为求解前视距离最优控制问题^[16]。文献[17]在 Matsui 研究的基础上提出模型无关的在线策略迭代学习算法, 对单站点 CSPS 系统的前视距离最优控制问题进行了求

解。但上述文献的研究均是基于机器不具有批量加工能力的基本假设, 对于批处理问题则没有涉及。

在现代化大规模生产环境下, 批量加工和预测控制机制对于提高生产效率, 满足现代化生产需求具有重要的意义。因此含有预测控制机制的批处理系统的最优控制问题, 也极具研究价值。但是, 至今为止针对该类问题的相关研究文献却极其匮乏, 只有部分文献^[11-13]对工件具有动态到达特性的批处理问题进行了研究, 该类问题中, 工件的到达时间仍是已知的, 并非完全的在线优化问题。因此, 本文以工件信息不完全可知的在线批处理问题为研究重点, 首先借鉴文献[17]中相关设定, 建立传送带给料分批加工系统的物理模型和优化模型。然后, 采用模型无关的 Q 学习优化算法, 结合工件分批规则, 求解系统的最优前视控制策略。最后, 通过仿真实验, 对算法的有效性进行了验证, 并分析了不同分批策略及参数对系统性能的影响。

1 问题模型

本文将预测控制机制引入批处理问题中, 提出一种传送带给料分批加工系统物理模型。针对该系统的单机模型, 在考虑工件属性差异的情况下, 以前视距离为控制变量, 无穷时段内的工件处理率最大为优化目标, 建立系统的数学模型。

1.1 物理模型

传送带给料分批加工系统的物理模型如图 1 所示。该物理模型主要由一条输送工件的传送带, 存放工件的缓冲区, 负责捡取工件的机械臂, 以及具有批处理能力的加工主体组成。其中, 机械臂可以完成工件的捡取、移动和放置工作。系统能够沿着工件到达的方向向前观察一段距离, 即前视距离, 并通过传感器得到工件的属性信息。

在每个决策时刻, 系统根据缓冲区中存放的工件情况选择前视距离, 然后对前视距离及缓冲区中所有工件统一进行分批决策, 选出待加工批。当选出的待加工批中的工件全部到达并放入加工主体后, 开始加工, 加工过程不允许中断, 加工完成后,

系统开始新一轮的决策过程。在整个系统运行过程中，若有工件进入前视距离，有以下三种情形：

- 1) 是待加工批中的工件，则将该工件捡取到加工主体中；
- 2) 不是待加工批中的工件，且缓冲区空余量充足，则将该工件捡取到缓冲区中；
- 3) 不是待加工批中的工件，且缓冲区空余量不足，则该工件流失。

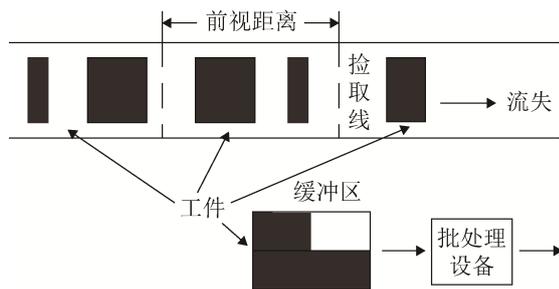


图1 传送带给料分批加工系统物理模型

Fig. 1 Physical model of conveyor-served batch processing system

本文主要研究单机环境下含有差异工件的传送带给料分批加工系统的在线优化控制问题，以提高系统的工件处理率。为了不失一般性且方便研究，本文研究基于以下假设：

(1) 传送带匀速运行，工件按照参数为 λ 的泊松流到达，工件有多种类型，不同种类工件到达的概率相等。

(2) 加工主体为一台批处理机，其容量为 C ，批中的工件尺寸之和不能大于 C ，假定没有尺寸大于 C 的工件存在。在满足容量约束时，任意工件均可放入同一批中进行加工。

(3) 批的加工时间等于该批中所有工件加工时间的最大值，第 n 个决策时刻选取的待加工批 b_n 的加工时间用 P^{b_n} 表示， $P^{b_n} = \max\{p_j | j \in b_n\}$ ，其中 p_j 表示工件 j 的加工时间。批 b_n 的到达时间 R^{b_n} 等于当前决策时刻至该批工件全部到达缓冲区的时间，即 $R^{b_n} = \max\{r_j | j \in b_n\}$ ，其中 r_j 表示工件 j 在当前决策时刻距离到达缓冲区的时间。如果在当前决策时刻，批 b_n 中所有工件均在缓冲区，则 $R^{b_n} = 0$ 。

(4) 工件的捡取时间消耗忽略不计。

1.2 数学模型

假设工件类型总数为 M ，缓冲区容量为 U 。记缓冲区中 i 类型工件的数量 x_i 作为 i 类型工件自身的状态，有 $0 \leq x_i \leq U$ ，其中 $1 \leq i \leq M$ 。整个系统的状态 x 由各个类型工件自身状态联合而成，即 $x = (x_1, x_2, \dots, x_M)$ 。系统所有可能的状态构成一个有限集合 X ，即状态空间。前视距作为系统的控制变量，即行动，用 a 表示，前视距离的范围构成行动集 A 。由于传送带匀速运行，故可将前视距离 a 和其他距离都用等效时间来表示，则 $A = [0, L]$ ，这里 L 为最大前视距离(时间)。 V 是系统的一个确定性平稳策略，它是状态到行动的映射，即在任意决策时刻，系统将从 A 中选取当前状态 x 所对应的行动 a ，这就是状态到行动的映射过程。

设 T_n 为第 n 个决策时刻($T_0=0$)，此时系统状态记为 $x(T_n)$ 。在策略 V 下，设系统选择行动 $a(x(T_n))$ 。系统将 $a(x(T_n))$ 范围内未到达的工件以及缓冲区中已到达的工件统一进行分批决策，选出待加工批 b_n 。当批 b_n 中的工件全部到达并放入批处理机后，开始加工，待加工完成后，进入下一决策时刻 T_{n+1} ，此时系统状态记为 $x(T_{n+1})$ 。如果在当前决策时刻，缓冲区中没有工件，系统将一直等待传送带上的下一个工件到达，捡取至缓冲区后进入下一决策时刻。

在选定一个合适的工件分批规则 δ 后，传送带给料分批加工系统的优化目标就是在策略空间 Ω 上寻找一个最优的前视控制策略 V^* ，使得系统在该策略的控制下无穷时段内的工件处理率 P_δ 最优，

$$P_\delta = \lim_{T \rightarrow \infty} \frac{P_\delta(T)}{Q(T)} \quad (1)$$

其中： $P_\delta(T)$ 表示系统在采用工件分批规则 δ 的情况下 T 时段内加工的工作量； $Q(T)$ 表示 T 时段内到达系统的工作量。下面给出工件量的定义。

定义 1 一个工件的工件量等于其尺寸与加工时间的乘积。

当系统稳态运行时，在一段较长的时间内，到达的工件量期望数一定，若流失的工件量越少，则

系统的工件处理率也就越高。记 $f_{\delta}(x(T_n), a(x(T_n)))$ 为系统采用分批规则 δ , 在状态 $x(T_n)$ 时采取行动 $a(x(T_n))$, 转移到下一状态 $x(T_{n+1})$ 之前的单位时间等效期望代价, 其值与系统流失的工件量有关。系统在无穷水平情况下平均准则的性能代价表示为:

$$\eta_{\delta}^V = E \left[\lim_{N \rightarrow \infty} \frac{1}{T_N} \sum_{n=0}^{N-1} \int_{T_n}^{T_{n+1}} f_{\delta}(x(T_n), a(x(T_n))) dt \right] \quad (2)$$

因此, 在工件分批规则选定的情况下, 系统的优化目标就是找到一个最优策略 V^* 满足 $V^* \in \arg \min_{V \in \Omega} \eta_{\delta}^V$ 。

2 优化控制方法

在传送带给料分批加工系统的在线优化控制问题中, 涉及到两个决策优化过程, 即行动选择决策和工件分批决策。行动选择决策是用来确定哪些工件将参与工件分批决策, 而工件分批决策则是从这些参与决策的工件中挑选出待加工批。如何优化这两个决策过程, 寻找最优前视距离控制策略和工件分批规则是本文主要研究的问题。下面首先研究工件分批规则的选择问题, 然后给出求解最优前视距离控制策略的优化算法。

2.1 分批规则

传送带给料分批加工系统的加工主体是具有批量加工能力的批处理机, 由于考虑工件尺寸差异, 每次形成的待加工批不一定是满批, 导致批处理机可能产生空闲空间。此外, 由于考虑工件加工时间差异, 且批的加工过程不允许中断, 同一批中工件的时间差异也会影响批处理机的利用率。同时, 当选出的待加工批中存在部分工件未到达缓冲区时, 从当前决策时刻到待加工批内工件全部到达且放入批处理机这段时间内, 批处理机一直处于空闲状态, 导致了时间上的浪费。

因此, 根据系统特点, 本文将上述三类浪费用工件的加工时间与工件尺寸的乘积这一量纲来统一度量, 提出批处理机加工能力浪费比的概念。

图 2 是从当前决策时刻 T_n 到下一决策时刻 T_{n+1} 时间内批处理机工作状态的示意图。图中 C 为批

处理机的容量(后文简称机器容量); 工件 $1, 2, \dots, n'$ 为当前批 b_n 中的工件, $s_1, s_2, \dots, s_{n'}$ 和 $p_1, p_2, \dots, p_{n'}$ 分别为工件 $1, 2, \dots, n'$ 的尺寸和加工时间; R^{b_n} 为批 b_n 的到达时间。区域 I 表示由于等待批 b_n 中的工件全部到达缓冲区所导致的浪费; 区域 II 表示由于批 b_n 中的工件尺寸之和小于 C 所导致的浪费; 区域 III 表示由于批 b_n 中的工件加工时间差异所导致的浪费。由此, 给出如下定义:

定义 2 一个决策周期内批处理机加工能力浪费比等于该决策周期内由于 I 类、II 类和 III 类浪费而少加工的工件量与该决策周期内批处理机理论上最大可加工工件量的比值。

假设第 n 个决策周期内批处理机加工能力浪费比为 φ_n , 根据定义 2, 从图 2 中可以得出:

$$\varphi_n = 1 - \frac{s_1 p_1 + \dots + s_{n'} p_{n'}}{C(T_{n+1} - T_n)} = 1 - \frac{\sum_{j \in b_n} s_j p_j}{C(R^{b_n} + P^{b_n})} \quad (3)$$

式中: P^{b_n} 为批 b_n 的加工时间, 其值等于 $p_{n'}$, $(\sum_{j \in b_n} s_j p_j) / (C(R^{b_n} + P^{b_n}))$ 为第 n 个决策周期内批处理机的利用率(后文简称机器利用率)。这里给出无穷时段内的机器利用率 C_{δ} 的表达式:

$$C_{\delta} = \lim_{T \rightarrow \infty} \frac{P_{\delta}(T)}{C(T)} \quad (4)$$

式中: $P_{\delta}(T)$ 表示系统在采用工件分批规则 δ 的情况下 T 时段内加工的工件量; $C(T)$ 表示 T 时段内批处理机理论上最大可加工工件量。

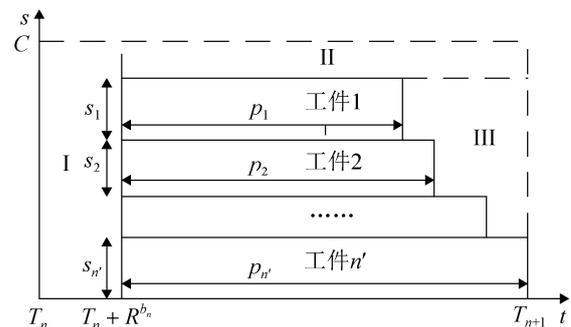


图 2 第 n 个决策周期内批处理机工作状态示意图
Fig. 2 Operation of batch processing machine in n th decision period

从式(3)中可以看出, 极小化 φ_n 等价于极大化该决策周期内的机器利用率。但是该决策周期内的

机器利用率达到最高,也仅表示系统在这一个决策周期内的局部最优,无法保证在一段较长的时间范围内系统的全局最优。如果将一个决策周期内批处理机加工能力浪费比最小作为系统的分批规则,在针对批处理机 I 类浪费的优化问题上将会与求解前视距离控制策略的优化算法产生冲突,对系统的优化控制产生影响。

综上所述,本文将批处理机加工周期内加工能力浪费比最小作为系统的分批规则,即仅对批处理机的 II 类和 III 类浪费进行优化,这样便可与求解前视距离控制策略的优化算法有机结合,从而协同优化批处理机的 I 类、II 类和 III 类浪费。下面给出批处理机加工周期内加工能力浪费比的定义。

定义 3.3 批处理机加工周期内加工能力浪费比等于批处理机加工过程中由于 II 类和 III 类浪费而少加工的工件量与批处理机加工过程中理论上最大可加工工件量的比值。

根据定义 3 和图 2 可以得出,批处理机第 n 个加工周期内加工能力浪费比 φ'_n 的表达式如下:

$$\varphi'_n = 1 - \frac{s_1 p_1 + \dots + s_n p_n}{C(T_{n+1} - T_n - R^{b_n})} = 1 - \frac{\sum_{j \in b_n} s_j p_j}{CP^{b_n}} \quad (5)$$

2.2 优化算法

在本文研究的传送带给料分批加工模型中,工件动态到达,且有多种尺寸类型,每种类型工件的加工时间也不同。与仅含单一类型工件的模型相比,本文所研究的模型在复杂性上大大提高。此外,本文研究的是在线情况下的优化控制问题,其复杂性远大于离线情况。因此,运用理论算法将会遇到“建模难”和“维数灾”的问题。本文引入模型无关的 Q 学习优化方法,求解最优前视控制策略。

Q 学习算法是通过仿真或观测系统运行,不断学习并逼近状态-行动对的函数值而对问题进行求解。学习过程中,系统每次状态转移,将记录一个观测样本 $\langle x(T_n), a(x(T_n)), w(T_n), l(T_n) \rangle$, 其中 $w(T_n)$ 表示系统在状态 $x(T_n)$ 的逗留时间, $l(T_n)$ 表示在 $[T_n, T_{n+1})$ 时间内流失的工件量。根据此样本,系统

采用分批规则 δ , 可计算 T_n 到 T_{n+1} 这段时间内产生的代价 $f'_\delta(x(T_n), a(x(T_n)))$, 其计算公式如下:

$$f'_\delta(x(T_n), a(x(T_n))) = Kl(T_n) \quad (6)$$

式中: K 为流失单位工件量的惩罚代价。根据文献[18], 可得系统的即时差分公式如下:

$$d_\delta(T_n) = f'_\delta(x(T_n), a(x(T_n))) + \min_{a \in A} Q_\delta(x(T_{n+1}), a) - w(T_n)\bar{\eta}_\delta - Q_\delta(x(T_n), a(x(T_n))) \quad (7)$$

式中: $\bar{\eta}_\delta$ 表示系统的平均代价学习值。于是,系统的 Q 值更新公式如下:

$$Q_\delta(x(T_n), a(x(T_n))) := Q_\delta(x(T_n), a(x(T_n))) + \gamma_\delta(x(T_n), a(x(T_n)))d_\delta(T_n) \quad (8)$$

式中: $\gamma_\delta(x(T_n), a(x(T_n)))$ 表示系统的 Q 值学习步长,一般与系统访问状态-行动对的次数有关。具体的学习算法流程如图 3 所示。

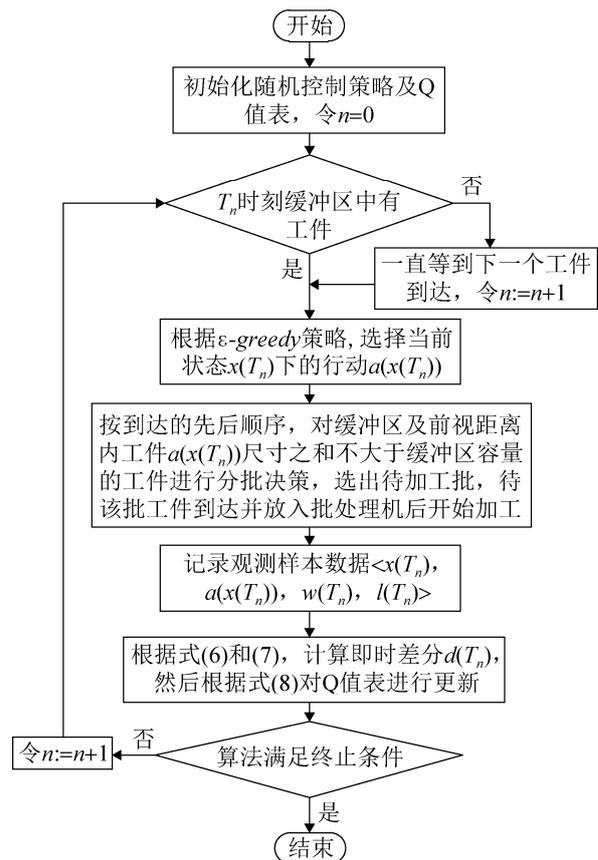


图 3 算法流程图

Fig. 3 Flowchart of proposed algorithm

其中, ϵ -greedy 策略是为了使 Q 学习算法的搜索能够根据现有的 Q 值表进行局部寻优, 同时又

能够有较好的探索能力, 实现“探索”和“开采”的平衡。探索率 ε 取值范围为(0,1)。根据 ε -greedy 策略选择行动时, 系统有 $1-\varepsilon$ 的概率选择当前状态下的最优行动, 有 ε 的概率从当前状态下的可行动中随机选择一个行动。

从式(7)和(8)中可以看出, Q 学习算法中的 Q 值更新时不仅考虑当前行动所产生的一步代价, 还考虑了当前行动对系统后续的影响, 避免算法陷入局部最优。因此, 通过 Q 学习算法求得的前视距离最优控制策略中每个状态所对应的行动不一定是该状态下的最优行动。若 Q 值的更新仅考虑当前行动所产生的代价, 则求得的前视距离控制策略就是一步贪心策略, 该策略中每个状态所对应的行动均是该状态下的最优行动。

3 仿真实验

传送带给料分批加工系统在线优化控制问题的研究, 不仅涉及到算法和工件分批规则的选择问题, 还涉及到机器容量 C 、缓冲区容量 U 、最大前视距离 L 、代价 K 、工件到达率 λ 、工件尺寸 s 以及工件加工时间 p 等参数的设置问题。下文通过仿真实验, 首先验证 Q 学习算法的求解性能, 并将其求得的最优策略与一步贪心策略进行对比。然后对几种工件分批规则和一些主要参数对系统性能的影响情况进行了分析总结。最后将本文中提出的最优策略和传统的零前视策略进行了对比。表 1 列出了部分系统参数以及算法参数的值。

表 1 参数设置表
Tab. 1 The value of parameters

参数	C	U	L	K	λ	ε	s	p
设定值	10	20	10	1	0.5	0.3	(4,3,2)	(1,5,10)

表 1 中, L 等价于时间表示, 且与 p 的单位相同, 均为秒。仿真中, 将行动区间[0,10]等间距离散化为有限行动集{0,0.2,0.4,...,10}。本文所有仿真实验均在 Microsoft Windows 7(CPU 3.40 GHz, 7.90 GB 内存)环境下使用 Matlab 2012a 平台编程实现。在 Q 学习算法运行初始, 令系统在任意状态下选择

任意可行行动的概率都相等, 每学习 2 万步进行一次 Monte-Carlo 策略评估。具体方法为, 根据当前学得策略仿真系统运行 10 万步, 共进行 20 次独立实验, 然后估计相关性能值。此外, 表 1 中设定的参数均是考虑实际生产加工情况, 并通过多次实验, 最终选取的。以下的仿真实验中, 如无特别说明, 则对应仿真参数均按照表 1 设置, 且实验结果均是通过 Monte-Carlo 策略评估后得到的值。

3.1 系统的优化效果

图 4 是系统分别在一步贪心策略和 Q 学习算法求得的最优策略控制下的工件处理率优化曲线。从图中可以看出, 随着学习步数的增加, 系统在两种策略下的工件处理率, 均从初始不到 0.9 逐渐提高至 0.935 左右, 并最终收敛, 可见两种策略都能有效提高系统的工件处理率。且 Q 学习算法由于从系统的整体最优进行考虑, 其优化效果要优于一步贪心策略。图 5 为对应图 4 仿真得到的系统分别在两种策略下的性能代价优化曲线, 其变化规律正好与图 4 的工件处理率优化曲线反相对应, 验证了代价选取的合理性。本文其它仿真实验, 如无特别说明, 其实验结果均为采用 Q 学习算法求得的最优策略后得到的值。

表 2 是在不同的 λ 下, 系统采用三种策略得到的工件处理率及机器利用率。其中, 零前视策略只对缓冲区中的工件进行决策, 可以有效减少 I 类浪费; 最大前视策略在每次决策时均采取最大前视距离, 能够有效减少 II 类和 III 类浪费。当 λ 较小时, 主要是 I 类浪费, 从表中可以看出, 此时零前视策略明显优于最大前视策略。随着 λ 的增大, 主要浪费由 I 类浪费逐渐变为 II 类和 III 类浪费, 此时零前视策略的优势逐渐被最大前视策略取代。而在不同的 λ 下, Q 学习算法都能不断地学习优化, 综合考虑 I 类、II 类和 III 类浪费, 其求得的最优策略的优化效果也明显优于其它两种策略。说明了 Q 学习算法可以有效地求解该类问题。

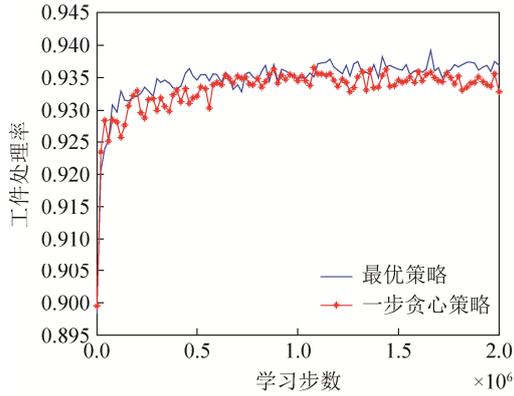


图 4 系统在最优策略和一步贪心策略下的工件处理率
Fig. 4 Job processing rate of system controlled by optimal policy and one-step greedy policy

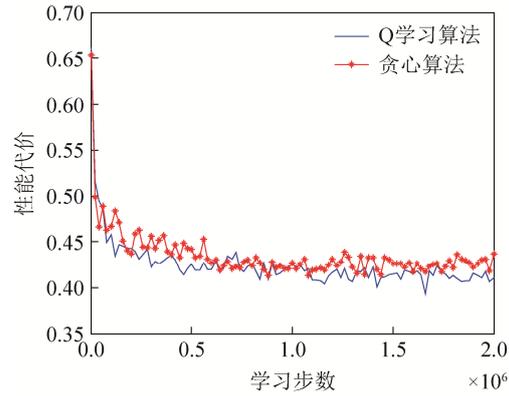


图 5 系统在最优策略和一步贪心策略下的性能代价
Fig. 5 Performance cost of system controlled by optimal policy and one-step greedy policy

表 2 不同的 λ 下三种控制策略的优化效果比较

Tab. 2 Performance comparison of three different control policies under different λ

控制策略	工件到达率	$\lambda=0.2$	$\lambda=0.4$	$\lambda=0.6$	$\lambda=0.8$	$\lambda=1.0$	$\lambda=1.2$
最优策略	工件处理率	1.000	0.975	0.874	0.743	0.627	0.538
	机器利用率	0.261	0.507	0.683	0.771	0.815	0.841
零前视策略	工件处理率	0.998	0.961	0.846	0.714	0.604	0.518
	机器利用率	0.260	0.501	0.661	0.744	0.784	0.810
最大前视策略	工件处理率	999.0	954.0	0.837	0.717	0.616	0.534
	机器利用率	0.260	0.495	0.654	0.748	0.801	0.835

3.2 分批规则对系统性能的影响

表 3 是在不同的 λ 下，系统分别采用四种分批规则，得到的六组仿真算例的工件处理率。其中，规则一即本文所采用的批处理机加工周期内加工能力浪费比最小分批规则；规则二是一个决策周期内批处理机加工能力浪费比最小分批规则；规则三是 FBLPT(First Batch Longest Processing Time)分批规则，其原理是将参与决策的工件按加工时间非增的顺序编号，以机器容量为约束，自前向后形成待加工批，可以有效较少

III 类浪费；规则四是先到达先加工分批规则，可以有效较少 I 类浪费。从表中可以看出，由于规则一可以与 Q 学习算法有机结合，从而协同优化 I 类、II 类和 III 类浪费，其优化效果明显优于仅考虑 III 类浪费的规则三和仅考虑 I 类浪费的规则四。规则二虽然也综合考虑了 I 类、II 类和 III 类浪费，但由于它在针对 I 类浪费的优化问题上与 Q 学习算法相冲突，其优化效果在部分算例情况下与规则一仍有少许差距。所以本文其它仿真实验采用的分批规则均为规则一。

表 3 不同的参数下四种分批规则的优化效果比较

Tab. 3 Performance comparison of four different batching rules under different parameters

工件加工时间		(1,5,10)	(2,10,3)	(3,4,7)	(5,5,5)	(6,5,2)	(8,6,4)
规则一	$\lambda=0.5$	0.937	0.901	0.980	0.980	0.955	0.851
	$\lambda=1.0$	0.625	0.572	0.691	0.662	0.617	0.469
规则二	$\lambda=0.5$	0.929	0.900	0.978	0.980	0.953	0.848
	$\lambda=1.0$	0.622	0.570	0.690	0.661	0.615	0.467
规则三	$\lambda=0.5$	0.735	0.736	0.897	0.971	0.942	0.814
	$\lambda=1.0$	0.381	0.374	0.476	0.633	0.545	0.418
规则四	$\lambda=0.5$	0.765	0.785	0.922	0.966	0.923	0.798
	$\lambda=1.0$	0.404	0.420	0.529	0.633	0.536	0.414

3.3 系统参数对性能的影响

图 6 是 λ 分别为 0.3、0.6、0.9、1.2 时, 系统的工件处理率随缓冲区容量 U 的变化曲线。横向比较可以看出, 当 $U \leq 20$ 时, 系统的工件处理率随 U 的增加而迅速增加, 而当 $U > 20$ 时, 系统工件处理率随 U 增加而增加的幅度迅速减小, 直至趋于平缓。纵向比较可以看出, 随着 λ 的增大, 曲线越来越平缓。由此可见, 通过增加 U 可以在一定程度上提高系统的工件处理率, 其提高的幅度与 λ 呈负相关。

虽然通过增加 U 可以在一定程度上提高系统的工件处理率, 但随着 U 增加, 系统的状态也会急剧增加, 这势必会使 Q 学习算法的收敛速度大大降低。通过图 7 也可以看出, 当 λ 较大时, U 的增加也会使平均每次分批决策所需时间迅速增加, 当增加到一定程度, 将必须考虑分批决策时间对系统的影响, 这也会使系统的复杂性提高。

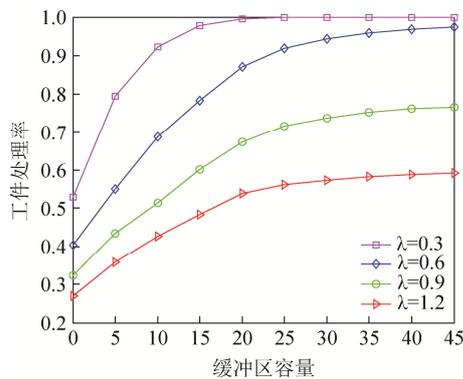


图 6 工件处理率随缓冲区容量变化曲线
Fig. 6 Change of job processing rate vs. buffer capacity

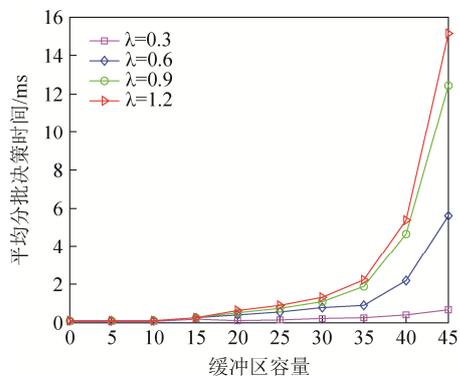


图 7 分批决策时间随缓冲区容量变化曲线
Fig. 7 Change of Batching time vs. buffer capacity

待加工工件具有到达时间、加工时间以及尺寸三个属性。其中, 到达时间由到达率 λ 决定。从图 6 可以看出, λ 与工件处理率成反比。与 λ 不同的是, 由于工件尺寸和加工时间的差异, 工件尺寸及加工时间与处理率之间并不是简单的比例关系。为探讨工件加工时间对于工件处理率的影响, 图 8 是在假定所有工件加工时间都相同, 即避免由于工件加工时间不同而引起 III 类浪费的情况下, 得到的工件处理率随工件加工时间的变化曲线。可以看出, 在 λ 相同的情况下, 工件处理率与工件加工时间呈负相关。

图 9 是在假定到达系统的工件量期望分别为 6 件/秒、9 件/秒以及 12 件/秒的情况下, 得到的工件处理率随工件加工时间方差的变化曲线。工件加工时间的方差越大, 即表示不同类型工件之间的加工时间差异越大, 造成的第 III 类浪费也越大。通过比较可以看出, 工件处理率随工件加工时间的方差增加而递减, 可见工件加工时间的差异越大, 对工件处理率的影响也越大。

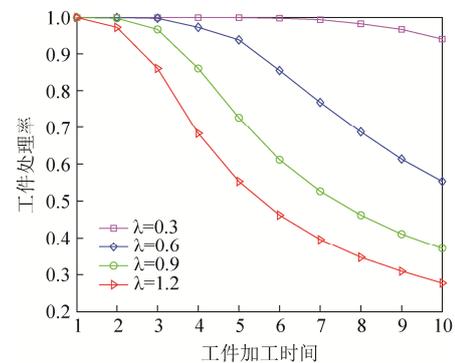


图 8 工件处理率随工件加工时间变化曲线
Fig. 8 Change of job processing rate vs. job processing time

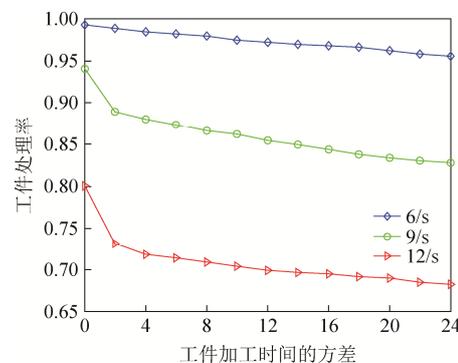


图 9 工件处理率随工件加工时间方差变化曲线
Fig. 9 Change of job processing rate vs. variance of job processing time

表 4 是在假定所有工件的加工时间相同且到达系统的工件量期望均为 9 件/秒的情况下得到的几组工件尺寸类型在两种机器容量 C 下的工件处理率和机器利用率。从表中可以看出, 当 C 相同时, 工件处理率与每组中工件尺寸类型的多寡及工件尺寸的大小没有明确的相关性, 仅与机器利用率呈正相关。再仔细对比每组算例在不同 C 下得到的工件处理率, 可以分析出, 工件处理率和 C 的大小、工件尺寸类型组合与 C 的匹配度呈正相关。

在工件具有动态到达特性的批处理问题中, 其

零等待策略仅对当前已到达的工件进行分批, 而不考虑还未到达的工件, 本文 3.1 节中所考虑的零前视策略本质上与其相同。从表 2 中可以看出, 本文所提出的最优策略相对于零前视策略, 在不同的工件到达率下工件处理率均有一定的提高。为了进一步对比零前视策略和最优策略的优化效果, 表 5 给出了在不同的缓冲区容量、机器容量、工件加工时间、工件尺寸类型下, 两种策略得到的工件处理率。从表中可以明显看出, 在不同的系统参数下, 本文提出的最优策略的优化效果均要优于零前视策略。

表 4 工件尺寸类型及机器容量对系统的影响
Tab. 4 Influence of job size and machine capacity on system

工件尺寸	(5)		(5,4)		(5,4,3)		(5,4,3,2)	
机器容量	C=10	C=12	C=10	C=12	C=10	C=12	C=10	C=12
工件处理率	0.925	0.925	0.863	0.927	0.908	0.945	0.921	0.958
机器利用率	0.834	0.694	0.777	0.696	0.816	0.708	0.830	0.719
工件尺寸	(6)		(6,5)		(6,5,4)		(6,5,4,3)	
机器容量	C=10	C=12	C=10	C=12	C=10	C=12	C=10	12=C
工件处理率	0.655	0.941	0.758	0.913	0.868	0.929	0.868	0.941
机器利用率	0.589	0.706	0.682	0.684	0.783	0.697	0.781	0.706

表 5 不同参数下最优策略和零前视策略的优化效果比较
Tab. 5 Performance comparison of optimal policy and zero look-ahead policy

系统参数	缓冲区容量			机器容量		
	10	20	30	8	12	16
最优策略	0.768	0.937	0.982	0.895	0.950	0.968
零前视策略	0.694	0.911	0.976	0.876	0.924	0.941
系统参数	工件加工时间			工件加工时间		
	(2,10,3)	(5,5,5)	(10,7,1)	(2,5,3)	(5,1,3)	(2,6,1)
最优策略	0.901	0.980	0.746	0.707	0.925	0.832
零前视策略	0.881	0.972	0.736	0.698	0.897	0.812

4 结论

本文将预测控制机制引入批处理问题中, 提出了一种传送带送料分批加工模型, 并以前视距离为控制变量, 无穷时段内的工件处理率最大为优化目标, 建立了系统的优化模型, 并引入 Q 学习算法求解最优前视控制策略。最后通过仿真实验, 验证了模型的合理性及方法的有效性。在后续的研究中可将单机环境下该系统模型的优化控制问题推广到其它环境, 如并行机环境、车间环境等, 并构造相关学习算法进行求解, 也将是一件有意义的工作。

参考文献:

[1] Germs R. Order Acceptance and Scheduling Policies for a Make-to-order Environment with Family-dependent Lead and Batch Setup Times [J]. International Journal of Production Research (S0020-7543), 2013, 51(3): 940-951.
 [2] 冯大光, 唐立新. 工件具有不同尺寸单台批处理机的调度问题 [J]. 控制工程, 2011, 18(3): 420-423. (Feng Daguang, Tang Lixin. Research on the Scheduling of Single Batching Machine with Non-identical Job Size [J]. Control Engineering of China, 2011, 18(3): 420-423.)
 [3] Uzsoy R. Scheduling a Single Batch Processing Machine with Non-identical Job Sizes [J]. International Journal of

- Production Research (S0020-7543), 1994, 32(7): 1615-1635.
- [4] Dupont L, Dhaenens-Flipo C. Minimizing the Makespan on a Batch Machine with Non-identical Job Sizes: An Exact Procedure [J]. Computers & Operations Research (S0305-0548), 2002, 29(7): 807-819.
- [5] Melouk S, Damodaran P, Chang P Y. Minimizing Makespan for Single Machine Batch Processing with Non-identical Job Sizes Using Simulated Annealing [J]. International Journal of Production Economics (S0925-5273), 2004, 87(2): 141-147.
- [6] Damodaran P, Manjeshwar P K, Srihari K. Minimizing Makespan on a Batch-processing Machine with Non-identical Job Sizes Using Genetic Algorithms [J]. International Journal of Production Economics (S0925-5273), 2006, 103(2): 882-891.
- [7] 谭琦, 陈华平, 杜冰, 等. 面向两客户的差异工件单机批调度问题 [J]. 系统仿真学报, 2013, 25(5): 975-982. (Tan Qi, Chen Huaping, Du Bing, et al. Two-customer Single Batch-processing Machine with Non-identical Job Sizes [J]. Journal of System Simulation (S1004-731X), 2013, 25(5): 975-982.)
- [8] Sizes Azimzadeh F, Palizban H A, Romagnoli J A. Online Optimal Control of a Batch Fermentation Process Using Multiple Model Approach [C]// IEEE Conference on Decision & Control, Tampa, Florida. USA: IEEE, 1998: 455-460.
- [9] Tajan J B C, Sivakumar A L, Gershwin, et al. Control of a Single Batch Processor with Incompatible Job Families and Future Job Arrivals [J]. Semiconductor Manufacturing (S0894-6507), 2011, 24(2): 208-222.
- [10] Fu R Y, Tian J, Yuan J J, et al. Online Scheduling on an Unbounded Parallel-batch Machine and a Standard Machine to Minimize Makespan [J]. Information Processing Letters (S0020-0190), 2014, 114(4): 179-184.
- [11] Chen H P, Du B, Huang G Q. Metaheuristics to minimise makespan on parallel batch processing machines with dynamic job arrivals [J]. International Journal of Computer Integrated Manufacturing (S0951-192X), 2010, 23(10): 942-956.
- [12] 李小林, 杜冰, 谭琦, 等. 分布式环境下批处理机调度启发式算法求解 [J]. 系统工程学报, 2012, 27(1): 99-110. (Li Xiaolin, Du Bing, Tan Qi, et al. Metaheuristics for scheduling batch processing machines in distributed environment [J]. Journal of Systems Engineering, 2012, 27(1): 99-110.)
- [13] Zhou S C, Chen H P, Xu R, et al. Minimising makespan on a single batch processing machine with dynamic job arrivals and non-identical job sizes [J]. International Journal of Production Research (S0020-7543), 2014, 52(8): 2258-2274.
- [14] 胡鸿韬, 江志斌, 陈侃. 半导体芯片制造混合调度方法 [J]. 上海交通大学学报, 2009, 43(3): 460-464. (Hu Hongtao, Jiang Zhibin, Chen Kan. Composite scheduling method in semiconductor wafer fabrication [J]. Journal of Shanghai Jiaotong University, 2009, 43(3): 460-464.)
- [15] Matsui M. A Generalized Model of Convey-serviced Production Station (CSPS) [J]. Journal of Japan Industrial Management Association (S1342-2618), 1993, 44(1): 25-32.
- [16] Matsui M. CSPS Model: Look-ahead Controls and Physics [J]. International Journal of Production Research (S0020-7543), 2005, 43(10): 2001-2025.
- [17] Tang H, Arai T. Look-ahead Control of Conveyor-serviced Production Station by Using Potential-based Online Policy Iteration [J]. International Journal of Control (S0020-7179), 2009, 82(10): 1917-1928.
- [18] Cao X R. Semi-Markov Decision Problems and Performance Sensitivity Analysis [J]. IEEE Transactions on Automatic Control (S0018-9286), 2003, 48(5): 758-769.