

# 带运输的混合流水车间调度问题的改进遗传算法

许可<sup>1a,2</sup>, 叶彩霞<sup>1b,2</sup>, 孙文娟<sup>1a,2</sup>

(1. 沈阳理工大学 a. 理学院, b. 自动化与电气工程学院, 沈阳 110159; 2. 辽宁省兵器工业智能优化与控制重点实验室, 沈阳 110159)

**摘要:** 为实现分布式制造环境中上下游工序和机器间的协同生产, 研究了带有运输的混合流水车间调度问题。以包含加工时间、运输时间和加工等待时间的完工时间最小为目标, 建立了带有运输约束的混合流水车间调度模型, 基于 Q-learning 设计了改进的遗传算法 (QGA) 求解该模型。在该算法中, 首先基于工件序号设计编码和遗传算子等遗传操作; 然后根据种群适应度函数构建种群的状态集合, 以交叉概率和变异概率的取值作为动作, 以最佳个体适应度和种群平均适应度作为奖励; 最后采用 Q-learning 对交叉和变异参数进行智能调整, 提高算法的收敛速度与全局搜索能力。仿真实验结果表明, 与改进的遗传算法 (GA-TS) 相比, 本文 QGA 的最大完工时间平均减少了 2.0%, 收敛速度提升了 18.1%。

**关键词:** 混合流水车间调度; 运输时间; 强化学习; 遗传算法

**中图分类号:** TP301.6 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2024.02.002

## Improved Genetic Algorithm for Scheduling Problem of Hybrid Flow Shop with Transportation

XU Ke<sup>1,2</sup>, YE Caixia<sup>1,2</sup>, SUN Wenjuan<sup>1,2</sup>

(1. Shenyang Ligong University, Shenyang 110159, China; 2. Liaoning Key Laboratory of Intelligent Optimization and Control for Ordnance Industry, Shenyang 110159, China)

**Abstract:** In order to realize the cooperative production between upstream and downstream processes and machines in distributed manufacturing environment, the scheduling problem of hybrid flow shop with transportation is studied. A scheduling model of hybrid flow shop with transportation constraints is established to minimize the completion time including processing time, transportation time and processing waiting time. An improved genetic algorithm (QGA) based on Q-learning is designed to solve the model. Firstly, in this algorithm, genetic operations such as coding and genetic operators are designed based on the job sequence number. Secondly, the state set of the population is constructed according to the fitness function of the population, the values of the crossover probability and mutation probability are taken as the action, and the best individual fitness and the average population fitness are used as rewards. Finally, Q-learning is used to intelligently adjust the crossover and mutation parameters, the convergence speed and global search ability of genetic algorithm can be improved. The simulation results show that compared with genetic algorithm, the makespan of QGA in this paper is reduced by 2.0% on average, and convergence speed is increased by 18.1%.

收稿日期: 2023-07-07

基金项目: 辽宁省“百千万人才工程”资助项目(2021921089); 辽宁省教育厅高等学校基本科研项目(LJKQZ2021057, LJKZ20260); 辽宁省“兴辽英才计划”项目(XLYC2006017)

作者简介: 许可(1982—), 女, 副教授, 研究方向为优化理论与算法、生产调度与物流优化、目标识别等。

**Key words:** hybrid flow shop scheduling; transportation time; reinforcement learning; genetic algorithm

混合流水车间调度问题 (hybrid flow shop scheduling problem, HFSP) 将传统的流水车间和柔性生产线相结合, 研究在具有多个设备、多种不同类型任务和各种约束条件的制造环境中如何合理安排工作任务和分配时间, 以实现某种性能指标的优化。HFSP 广泛存在于玻璃、纺织、造纸和钢铁等制造业领域<sup>[1]</sup>, 该问题已被证明属于 NP-hard 问题<sup>[2]</sup>, 其求解难度较大。在分布式制造环境下, 各工序的加工制造设备可能处于不同的地理位置, 在考虑设备制造能力的同时, 也要兼顾运输时间对调度结果的影响。对企业来说, 研究考虑运输时间的 HFSP 具有重要的现实意义。

目前, 许多学者开展了 HFSP 研究。轩华等<sup>[3]</sup>研究了带运输时间的多阶段动态可重入 HFSP, 目标函数为最小化总加权完成时间, 设计了两种改进的拉格朗日松弛算法, 提高了解的质量和运行时间。袁帅鹏等<sup>[4]</sup>研究了带有序列相关准备时间和阶段间运输时间的混合流水车间成组调度问题, 以最小化最大完工时间为目标, 设计了一种协同进化文化基因算法, 并通过对比实验证明了算法的有效性。耿凯峰等<sup>[5]</sup>研究了带工序跳跃的绿色混合流水车间机器人和自动引导车联合调度问题, 提出了改进的文化基因算法, 实现了最大完工时间和总能耗同时最小化。Hidri 等<sup>[6]</sup>研究了带有转移时间和运输时间的两阶段 HFSP, 以最小化最大完工时间为优化目标, 提出了一种由两阶段组成的启发式方法, 该方法能够在可接受的计算时间内提供最优解或近似最优解。Li 等<sup>[7]</sup>针对混合流水车间生产运输一体化调度问题, 设计了一种有效的遗传禁忌搜索算法, 达到了最大完工时间的最小化。Yuan 等<sup>[8]</sup>研究了考虑阻塞和运输时间的二机流水车间成组调度问题, 设计了协同进化遗传算法, 实现了最小化最大完工时间。上述研究均属于将运输和生产协同考虑的并行同速机的 HFSP, 而在不相关并行机的 HFSP 方面研究则较少, 且目前的求解方法也多为智能优化算法和精确算法, 求解质量和求解效率都不高, 易陷入局部最优。

近年来, 随着机器学习的快速发展, 将强化学习 (reinforcement learning, RL) 与智能优化算法相结合用于解决相关领域的调度问题受到广泛关注。尹爱军等<sup>[9]</sup>研究了柔性作业车间多目标优化

调度问题, 针对算法多样性不足、易于早熟及局部收敛的缺点, 提出了一种基于 RL 的改进非支配排序遗传算法。张韵等<sup>[10]</sup>提出了一种基于机器学习的多策略并行遗传算法, 通过引入能自主感知环境的 RL, 实现遗传算法 (GA) 中重要参数交叉概率的自学习。李润佛等<sup>[11]</sup>为解决港口拥堵问题, 提出了一种基于 RL 的自学习遗传算法 (SLGA), 减少了船舶在港等待时间, 提高了港口通航效率。Shahrabi 等<sup>[12]</sup>针对考虑随机作业到达和机器故障的动态作业车间调度问题, 使用 Q-learning 算法来提高可变邻域搜索 (VNS) 算法的性能。Cao 等<sup>[13]</sup>针对柔性作业车间调度问题, 通过 RL 算法构建了布谷鸟搜索算法的自适应参数控制方案, 提出了一种基于知识的布谷鸟搜索算法。Alipour 等<sup>[14]</sup>将 GA 和 RL 相结合, 用于求解旅行商问题 (TSP), 取得了较优的结果。Chen 等<sup>[15]</sup>针对柔性作业车间调度问题, 将 GA 作为基本优化方法, 采用 RL 对其关键参数进行智能调整。上述研究中将机器学习与智能优化算法相结合, 提高了算法求解性能及求解质量。

带运输的 HFSP 是一个复杂的组合优化问题, 需要综合考虑加工时间、运输时间和优先级约束, 以实现最优的调度方案。GA 稳定性好、通用性强, 是求解 HFSP 最常用的算法。GA 中交叉概率和变异概率的设置直接影响算法的搜索能力和收敛速度, 针对不同的调度问题, 交叉概率和变异概率等参数通常采用实验分析<sup>[16]</sup>或经验来确定。目前, 在 HFSP 的求解中, 关于 GA 参数智能调整的研究很少。本文以最小化最大完工时间为目标, 建立带运输的 HFSP 优化模型, 提出一种基于 Q-learning 的改进遗传算法 (QGA) 求解该模型, 采用 Q-learning 算法智能调整 GA 的交叉概率和变异概率, 提高算法的求解效率和求解质量。

## 1 带运输的 HFSP 模型

现有  $n$  个工件需要在  $L$  道工序上进行加工, 工件序号为  $j(j=1, 2, \dots, n)$ , 工序序号为  $l(l=1, 2, \dots, L)$ , 第  $l$  道工序中有  $m_l$  台不相关并行机器, 机器序号为  $k(k=1, 2, \dots, m_l)$ , 全部工序共有  $M$  台机器, 各机器位于不同地点。混合流水车间调度流程如图 1 所示。

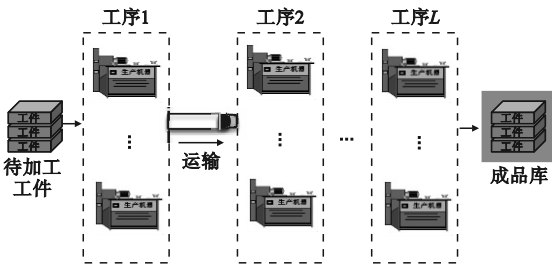


图 1 车间调度流程图

Fig. 1 Flow chart of shop scheduling

为简化模型,进行如下假设:

- 1) 工序间缓冲区的容量无限;
- 2) 相邻工序间存在运输阶段,且不受运输工具的制约;
- 3) 工件一旦开始加工或运输,不可中断;
- 4) 所有工件在零时刻已经到达并可以开始加工。

为实现最大完工时间最小化,需要决策每个工件在每道工序的生产机器以及各工件在该机器上的生产顺序。

为准确描述 HFSP,定义如下参数和决策变量: $b_{j,l}$ 表示工件  $j$  在工序  $l$  的开始加工时间; $e_{j,l}$ 表示工件  $j$  在工序  $l$  的加工完成时间; $p_{j,l}^k$ 表示工件  $j$  在工序  $l$  的机器  $k$  上的加工时间; $t_k^{k'}$ 表示工件从当前工序的机器  $k$  运输到下一道工序的机器  $k'$  的时间; $X_{j,l}^k$ 表示决策变量,若工件  $j$  在工序  $l$  选择机器  $k$  加工, $X_{j,l}^k$ 为 1,否则为 0; $\sigma$ 表示一个可行的整体调度方案; $C_{\max}$ 表示所有工件的最大完工时间。

基于上述说明,建立如下数学模型。

$$\min C_{\max} = \max \{ p_j(\sigma) + t_j(\sigma) + w_j(\sigma) \} \quad (1)$$

$$\text{s. t. } p_j(\sigma) = \sum_{l=1}^L \sum_{k=1}^{m_l} p_{j,l}^k X_{j,l}^k \quad (2)$$

$$t_j(\sigma) = \sum_{l=1}^{L-1} \sum_{k=1}^{m_l} \sum_{k'=1}^{m_{l+1}} t_k^{k'} X_{j,l}^k X_{j,l+1}^{k'} \quad (3)$$

$$w_j(\sigma) = b_{j,1} + \sum_{l=1}^{L-1} b_{j,l+1} - e_{j,l} - t_j(\sigma) \quad (4)$$

$$0 \leq b_{j,l} \leq e_{j,l} \quad (5)$$

$$e_{j,l} + t_j(\sigma) \leq b_{j,l+1} \quad (6)$$

$$\sum_{k=1}^{m_l} X_{j,l}^k = 1 \quad (7)$$

$$e_{j,l} = b_{j,l} + \sum_{k=1}^{m_l} X_{j,l}^k p_{j,l}^k \quad (8)$$

模型各方程中:式(1)表示目标函数;式(2)表示工件的加工时间;式(3)表示工件的运输时间;式(4)表示工件在机器前等待加工时间;式(5)表示工件在每道工序的完工时间必须大于其相应工序

的开始加工时间,且所有工件开始时间不为负;式(6)表示工件在前一道工序完工之后需经过运输方能在下一道工序开始加工;式(7)表示每道工序中一个工件只能选择一台机器加工;式(8)表示工件完工时间等于工件开始时间与其加工时间之和。

## 2 基于 Q-learning 的改进遗传算法

传统 GA 中交叉概率( $P_c$ )和变异概率( $P_m$ )等参数需要依靠实验分析或经验确定,参数数值过大时,会使求解速度过慢,不易寻得最优解,数值过小时,很难产生新的个体。针对该问题,本文采用 Q-learning 算法对参数进行智能化调整,根据种群当前状态和过去状态智能选择  $P_c$  和  $P_m$ ,并对未来状态进行预测,从而提高求解效率和精度。

### 2.1 编码与解码

在第一道工序上采用基于工件序号的编码方式表示工件的加工优先级,在其余工序上采取先来先服务的原则。每条染色体包含  $n$  个基因位,分别对应  $n$  个工件的序号。解码基于最早完工时间解码算法,计算工件在某道工序所有加工机器上的完工时间,选择完工时间最少的机器进行加工。

编码与解码过程示例如下。

设有三个工件依次经过 2 道工序进行加工,共有 4 台加工机器,每道工序的机器数量均为 2 台。各工件在机器上的加工时间分别为: $p_{1,1}^1 = 6, p_{2,1}^1 = 5, p_{3,1}^1 = 4, p_{1,1}^2 = 5, p_{2,1}^2 = 5, p_{3,1}^2 = 5, p_{1,2}^1 = 4, p_{1,2}^2 = 5, p_{3,2}^1 = 2, p_{2,2}^1 = 3, p_{2,2}^2 = 4, p_{3,2}^2 = 2$ 。机器间的运输时间分别为: $t_1^1 = 3, t_2^1 = 1, t_1^2 = 2, t_2^2 = 5$ 。如果染色体编码为[3 1 2],解码结果如图 2 所示,图中二元组表示工件序号与工序序号。

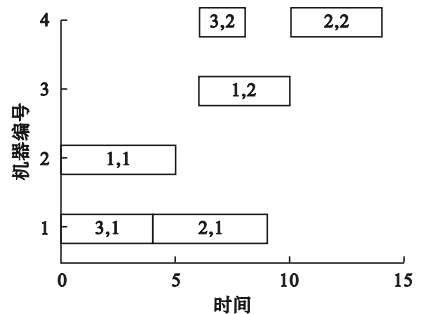


图 2 解码结果示意图

Fig. 2 Diagram of the decoding result

## 2.2 适应度函数和遗传操作

适应度函数在 GA 中起着关键作用,其值越高,个体就越接近最优解,反之,则个体越差。适应度函数计算式为

$$f(x'_i) = \frac{1}{1 + C(x'_i)} \quad (9)$$

式中: $x'_i$  表示第  $t$  次迭代时的第  $i$  个个体; $C(x'_i)$  表示第  $t$  次迭代时个体  $i$  的最大完工时间,分母上加 1 有助于平衡适应度函数的取值范围。

采用转盘赌选择、部分匹配交叉和两点交换变异的方法对染色体进行选择、交叉和变异操作。

## 2.3 Q-learning 算法

Q-learning 是一种经典的 RL 算法,属于基于价值函数的方法。在使用 Q-learning 算法求解问题时需要考虑状态、行动、报酬和策略四个要素, $Q$  值更新计算中需综合考虑智能体过去所经历的状态、动作和奖励,公式为

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)] \quad (10)$$

式中: $Q(s_t, a_t)$  表示在状态  $s_t$  采取行动  $a_t$  时的  $Q$  值; $r_t$  为即时奖励; $\alpha$  为学习率; $\max_{a'} Q(s_{t+1}, a')$  表示在状态  $s_{t+1}$  时采取行动  $a_{t+1}$  得到  $Q$  值表中期望最高的  $Q$  值; $\gamma$  为折扣率。

### 2.3.1 状态特征

GA 中判断种群好坏的指标通常包括平均适应度、多样性及最佳个体适应度。平均适应度反映了种群整体的适应能力;多样性反映了种群内个体之间的差异;最佳个体适应度反映了种群中最佳个体的适应性水平。平均适应度越高、多样性越高、最佳个体适应度越高,说明种群的适应能力越强。因此,为更好地评价种群的表现,选择种群的平均适应度、多样性以及最佳个体适应度作为种群的状态集合。

设种群在第  $t$  代的平均适应度为  $F^t$ 、第  $t$  代种群的多样性为  $D^t$ 、种群在第  $t$  代最佳个体适应度为  $M^t$ 、种群状态值为  $S^*$ ,计算方法如式(11)~(14)所示。

$$F^t = \sum_{i=1}^N f(x'_i) / N \quad (11)$$

$$D^t = \sum_{i=1}^N \left| f(x'_i) - \sum_{i=1}^N f(x'_i) / N \right| \quad (12)$$

$$M^t = \max f(x'_i) \quad (13)$$

$$S^* = w_1 \frac{F^t}{F^1} + w_2 \frac{D^t}{D^1} + w_3 \frac{M^t}{M^1} \quad (14)$$

式中: $N$  为种群中个体数量; $w_1$ 、 $w_2$ 、 $w_3$  分别表示三个指标的权重, $w_1 + w_2 + w_3 = 1$ 。本文采用种群的平均适应度和多样性反映种群的整体状态,有利于获得更优秀的个体,因此  $w_1$  和  $w_2$  值比  $w_3$  值相对重要,取  $w_1 = 0.35$ 、 $w_2 = 0.35$ 、 $w_3 = 0.3$ 。

状态的划分对收敛性影响很大:划分过多,寻优耗时较长;划分过少,确定的方案较差。为此,基于 GA 求解 HFSP 的适应度和  $S^*$  值时,划分为 10 个状态,状态向量表示为  $S = [s(1), s(2), \dots, s(10)]$ , $S^*$  的间隔为 0.1。如状态为  $s(1)$ , $S^* \in [0, 0.1)$ ;如状态为  $s(2)$ , $S^* \in [0.1, 0.2)$ 。依此类推。

### 2.3.2 动作

每一代智能体都会选择不同的动作得到合适的  $P_c$  和  $P_m$ ,其中  $P_c$  的取值范围为 0.4~0.9,动作间隔为 0.1,动作集中包含 10 个动作。如动作设为  $a_1$  时, $P_c \in [0.4, 0.5)$ ,则  $P_c$  从该区间内随机选择一个值;动作设为  $a_2$  时, $P_c \in [0.5, 0.6)$ 。 $P_m$  的选择与上述类似,取值范围为 0.01~0.21,间隔为 0.04。如动作设为  $a_1$  时, $P_m \in [0.01, 0.05)$ ,则  $P_m$  从该区间内随机选择一个值。

### 2.3.3 奖励

奖励通过最佳个体适应度和种群平均适应度来设计。调整  $P_c$  后的奖励  $r_c$  与调整  $P_m$  后的奖励  $r_m$  计算见式(15)与式(16)。

$$r_c = \frac{\max f(x'_i) - \max f(x'_i)^{-1}}{\max f(x'_i)} \quad (15)$$

$$r_m = \frac{\sum_{i=1}^N f(x'_i) - \sum_{i=1}^N f(x'_i)^{-1}}{\sum_{i=1}^N f(x'_i)} \quad (16)$$

### 2.3.4 探索与利用

$\varepsilon$ -贪心策略是一种同时考虑探索和开发的行动选择策略,即以概率  $\varepsilon$  随机选择一个动作,以概率  $1 - \varepsilon$  选择  $Q$  值最大的动作,表达式为

$$A(s) = \begin{cases} \operatorname{argmax}_a Q(s, a), & P = 1 - \varepsilon \\ a(\text{Randomly}), & P = \varepsilon \end{cases} \quad (17)$$

式中: $A(s)$  表示状态为  $s$  时选择的动作; $P$  为选取动作的概率。

## 2.4 算法流程

将 Q-learning 与 GA 结合,设计如图 3 所示的 QGA 框架,包括环境、学习模块和强化过程三部分。

采用基于 Q-learning 算法设计的 QGA 求解考虑运输时间的 HFSP,具体步骤如下。

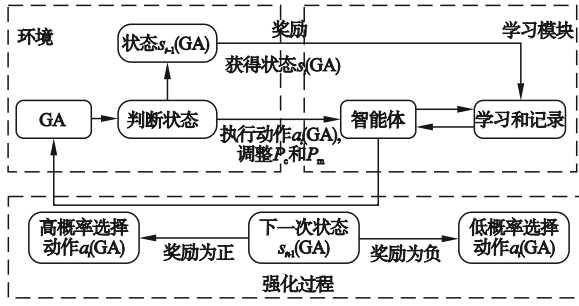


图 3 QGA 框架

Fig. 3 Framework of QGA

**步骤 1** 初始化参数,包括调度问题参数、GA 参数、Q-learning 算法参数。

**步骤 2** 初始化  $Q$  值表,包括状态集、动作集、策略。

**步骤 3** 随机生成初始种群,设置当前迭代次数  $t=0$ 。

**步骤 4** 计算当前个体的适应度值,并采用轮盘赌方法对父代染色体进行选择。

**步骤 5** 获取所有个体在第  $t$  次迭代时的状态  $s_t$ ,使用  $\epsilon$ -贪心策略根据  $Q$  值表选取动作,获得调整后的  $P_c$  和  $P_m$ 。

**步骤 6** 计算从状态  $s_t$  采取行动  $a_t$  到状态  $s_{t+1}$  所获得的奖励  $r(s_t, a_t, s_{t+1})$ ,并根据式 (14) 更新  $Q$  值表。

**步骤 7** 使用更新后的参数进行交叉和变异操作,生成新的个体。

**步骤 8** 判断是否达到最大迭代次数,如果是,输出结果,结束计算,否则令  $t=t+1$ ,转到步骤 4。

### 3 仿真实验与分析

采用仿真实验验证本文算法的有效性。实验硬件环境为: Intel (R) Core (TM) i5-7200U CPU 2.50 GHz、20 G 内存。采用 PyCharm2017.3.2 编译器、python3.7 解释器软件编程实现。

#### 3.1 初始条件与参数选择

在仿真实验中,取工件数量  $n=40$ ,工序数量  $L=3$ ,每道工序的机器数量分别为  $m_1=3$ 、 $m_2=2$ 、 $m_3=4$ 。工件在每台机器上的加工时间  $p_{j,l}^k$  在 1~8 之间随机生成,各机器间的运输时间  $t_k^k$  在 1~5 之间随机生成。

采用不同参数进行仿真计算,通过比较目标函数值确定 GA 的最佳参数取值。计算结果如表 1 所示。

表 1 不同 GA 初始参数下的仿真计算结果

Table 1 Simulation results with different GA initial parameters

序号	$N$	$P_m$	$P_c$	$C_{max}$
1	30	0.01	0.9	75
2	30	0.05	0.9	77
3	30	0.10	0.9	75
4	80	0.01	0.8	76
5	80	0.05	0.8	75
6	80	0.10	0.8	74
7	130	0.01	0.7	73
8	130	0.05	0.7	74
9	130	0.10	0.7	75

由表 1 可见,目标函数值  $C_{max}$  最小为 73,据此选取种群规模为 130、交叉概率为 0.7、变异概率为 0.01。变异概率过高会使算法近似于随机搜索,难以得到最优解,种群数量过小会使算法陷入局部最优。

为确定 Q-learning 算法的参数取值,采用不同参数进行仿真计算,结果如表 2 所示。

表 2 不同 Q-learning 算法初始参数下的仿真计算结果

Table 2 Simulation results with different initial parameters of Q-learning algorithm

序号	学习率	折扣率	贪婪率	$C_{max}$
1	0.10	0.99	0.1	74
2	0.05	0.99	0.1	73
3	0.01	0.99	0.1	73
4	0.10	0.95	0.3	74
5	0.05	0.95	0.3	74
6	0.01	0.95	0.3	72
7	0.10	0.90	0.5	75
8	0.05	0.90	0.5	74
9	0.01	0.90	0.5	73

由表 2 可见,目标函数值  $C_{max}$  最小为 72,据此选择学习率为 0.01、折扣率为 0.95、贪婪率为 0.3。学习率越低,保留之前的学习经验越多,越容易得到最优解;折扣率越高,越重视以前的经验,越容易得到最优解;贪婪率居中选取为优。

图 4 为上述最佳实验参数下算法迭代到 400 代时目标函数值的收敛情况。由图 4 可见,在迭代次数为 150 时,  $C_{max}$  值已经趋于稳定,故本次实验最大迭代次数选择为 150。

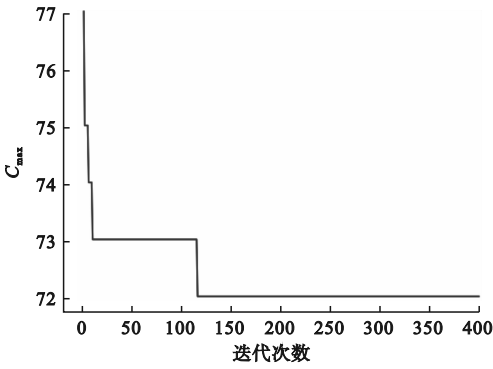


图4 C<sub>max</sub>的迭代收敛图

Fig.4 The iterative convergence plot of C<sub>max</sub>

### 3.2 实验结果分析

为验证本文 QGA 的可行性,在两组不同生产规模(工件、工序及机器数量各不相同)下进行实验。实验设置如表 3 所示,表 4 为不同生产规模下重复进行 10 次实验得到的最优 C<sub>max</sub> 值。

表 3 实验设置

Table 3 Experimental setup

实验编号	<i>n</i>	<i>L</i>	<i>M</i>	机器数量分布
S1	12	4	10	3,3,2,2
S2	9	8	13	2,3,1,1,1,3,1,1

表 4 不同生产规模下的实验结果

Table 4 Experimental results under different sizes

实验编号	C <sub>max</sub>									
	1	2	3	4	5	6	7	8	9	10
S1	31	32	31	31	31	31	31	31	31	31
S2	94	94	94	94	94	95	94	94	94	94

由表 4 可知,在不同规模下达到最优值的概率均为 90%,证明了本文算法的稳定性。图 5 和图 6 为对应两种规模下的最优调度甘特图,可以看到,机器的空闲时间较少,说明机器的利用率高,能够获得较小的最大完工时间,验证了本文 QGA 算法的可靠性。

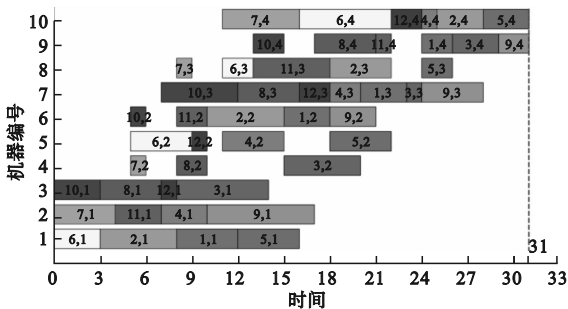


图 5 *n* = 12 和 *L* = 4 的调度甘特图

Fig.5 Schedule Gantt chart with *n* = 12 and *L* = 4

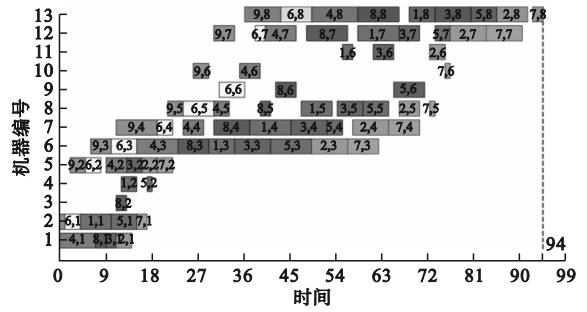


图 6 *n* = 9 和 *L* = 8 的调度甘特图

Fig.6 Schedule Gantt chart with *n* = 9 and *L* = 8

图 7 和图 8 分别为不同平均适应度下 *P<sub>c</sub>* 和 *P<sub>m</sub>* 的变化情况。由图 7 可以看到,随着平均适应度的增加,*P<sub>c</sub>* 增大,说明平均适应度高时,种群中包含一些较优秀的个体。为保留优秀基因,选择较高的 *P<sub>c</sub>*,以在交叉操作中更频繁地将优秀个体的基因信息进行传递。由图 8 可见,随着平均适应度的增加,*P<sub>m</sub>* 降低。这是因为此时种群已经优化,通过适当降低 *P<sub>m</sub>*,防止种群过早收敛到局部最优解。

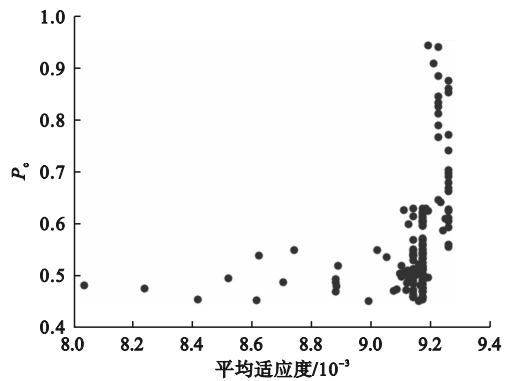


图 7 *P<sub>c</sub>* 的取值变化图

Fig.7 The graph of the value change of *P<sub>c</sub>*

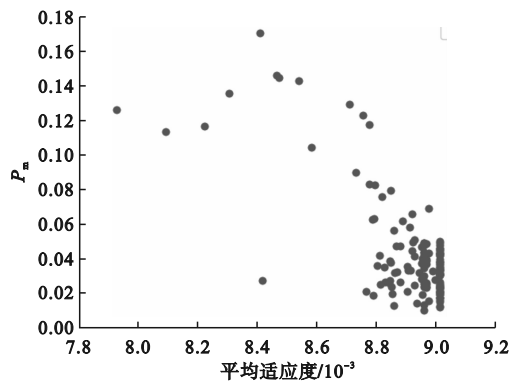


图 8 *P<sub>m</sub>* 的取值变化图

Fig.8 The graph of the value change of *P<sub>m</sub>*

### 3.3 实验对比

为验证本文算法的优势,在不同生产规模下,

采用基于禁忌搜索改进的遗传算法 (GA-TS) 和 QGA 进行实验对比分析。规模参数为: $n$  取 10、25、50、100; $L$  取 3、5、8; $M$  取 9、10、13, 机器随机分布在每道工序。当  $n < 50$  时视为小规模, 否则视为大规模, 表 5 和表 6 分别为小规模 and 大规模下的对比结果。表中: $T_1$  为迭代 1 次所需要的时间; $T_{optimal}$  为首次达到最优结果所需的时间。为避免实验误差, 每组实验重复 5 次取平均值。

表 5 小规模下的实验结果对比

Table 5 Comparison of experimental results at a small scale

$(n, L, M)$	本文 QGA			GA-TS		
	$C_{max}$	$T_1/s$	$T_{optimal}/s$	$C_{max}$	$T_1/s$	$T_{optimal}/s$
(10, 3, 9)	25	0.24	<b>1.71</b>	25	0.14	2.57
(10, 5, 10)	<b>40</b>	0.35	<b>6.22</b>	41	0.17	9.56
(10, 8, 13)	<b>85</b>	0.54	<b>4.89</b>	86	0.30	6.90
(10, 3, 13)	17	0.26	<b>2.85</b>	17	0.18	4.00
(10, 5, 9)	61	0.35	<b>1.74</b>	61	0.20	2.16
(10, 8, 10)	<b>82</b>	0.52	<b>6.27</b>	83	0.28	11.28
(25, 3, 9)	<b>45</b>	0.54	28.62	47	0.25	4.26
(25, 5, 10)	<b>57</b>	0.79	<b>40.94</b>	60	0.38	50.89
(25, 8, 13)	<b>137</b>	1.30	87.41	144	0.60	58.54
(25, 3, 13)	<b>25</b>	0.56	<b>15.17</b>	26	0.27	17.11
(25, 5, 9)	<b>142</b>	0.80	80.33	144	0.38	59.62
(25, 8, 10)	<b>158</b>	1.24	118.12	164	0.60	65.62

表 6 大规模下的实验结果对比

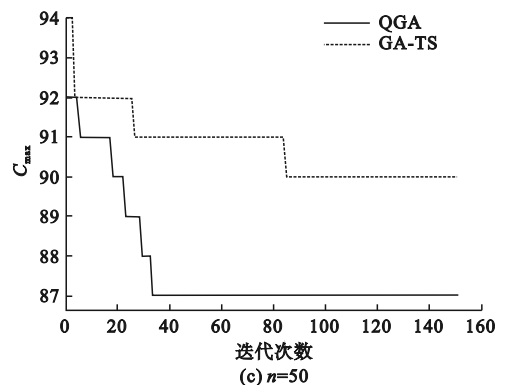
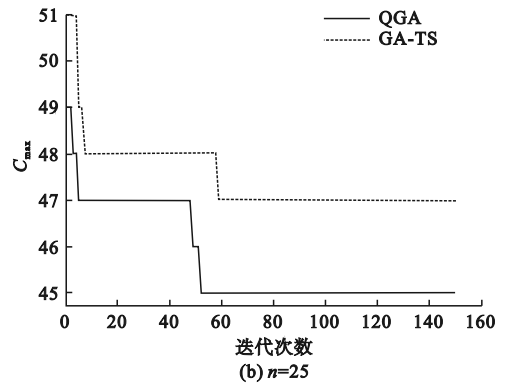
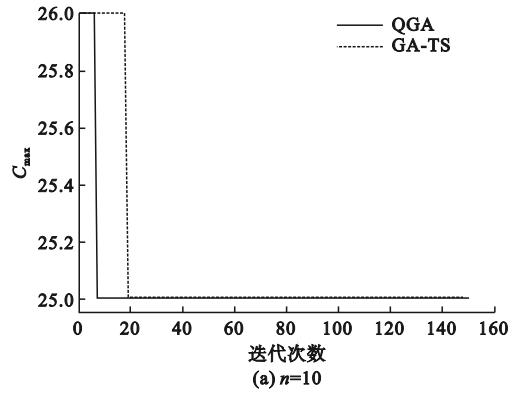
Table 6 Comparison of experimental results at a large scale

$(n, L, M)$	本文 QGA			GA-TS		
	$C_{max}$	$T_1/s$	$T_{optimal}/s$	$C_{max}$	$T_1/s$	$T_{optimal}/s$
(50, 3, 9)	<b>87</b>	1.04	<b>36.56</b>	90	0.50	42.77
(50, 5, 10)	<b>104</b>	1.64	131.36	110	0.77	93.45
(50, 8, 13)	<b>280</b>	2.59	251.42	283	1.19	128.52
(50, 3, 13)	37	1.04	<b>52.17</b>	38	0.51	58.96
(50, 5, 9)	253	1.66	<b>76.30</b>	253	0.79	77.55
(50, 8, 10)	<b>286</b>	2.53	316.83	290	1.18	183.00
(100, 3, 9)	<b>169</b>	2.13	204.67	172	0.99	140.58
(100, 5, 10)	<b>208</b>	3.26	404.82	215	1.54	221.47
(100, 8, 13)	<b>509</b>	5.14	606.76	514	2.47	281.66
(100, 3, 13)	<b>73</b>	2.12	<b>76.22</b>	76	1.01	87.20
(100, 5, 9)	<b>474</b>	3.39	392.93	477	1.59	258.01
(100, 8, 10)	<b>499</b>	4.96	466.68	504	2.50	463.36

由表 5 和表 6 可知, 在 24 个测试实例中, QGA 在 20 个实例中的  $C_{max}$  值都优于 GA-TS, 另外 4 个实例求得的  $C_{max}$  值相同。求解小规模问题时, GA-TS 的  $C_{max}$  平均值为 75, 本文 QGA 的  $C_{max}$  平均值为 73, 减少了 2.7%; 求解大规模问题时, GA-TS 的  $C_{max}$  平均值为 251, 本文 QGA 的  $C_{max}$  平均值为 248, 减少了 1.2%。从整体性能来看, 虽然 GA-TS 求解问题用时较短, 但 QGA 首次达到最优结果所需的时间半数情况下均优于 GA-TS, 收敛速度平均提升了 18.1%。

图 9 为不同工件数量下 QGA 与 GA-TS 两种算法的目标值  $C_{max}$  变化趋势图。

由图 9 可以看到, 在不同工件数量下, QGA 均获得了更好的初始解和最优调度解。在获得相同最优解的情况下, QGA 的收敛速度更快。



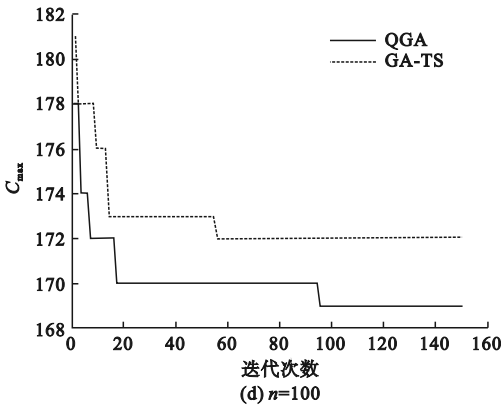


图9 不同工件数量下  $C_{max}$  变化趋势图

Fig.9 Trend plots of  $C_{max}$  changes under different jobs

## 4 结论

以最小化最大完工时间为目标,建立了带有运输的混合流水车间调度模型,基于 Q-learning 设计了改进的遗传算法(QGA)求解调度问题。采用 Q-learning 对 GA 的交叉概率和变异概率进行精确调整,显著提高了求解效率。在不同规模下进行了仿真实验,并与 GA-TS 对比分析,结果表明,QGA 在大多数情况下都能得到较好的解,且收敛性能也优于 GA-TS,验证了其学习效果和优异性能。

### 参考文献(References):

- [1] CHENG Q, LIU C F, CHU H Y, et al. A new multi-objective hybrid flow shop scheduling method to fully utilize the residual forging heat[J]. IEEE Access, 2020, 8: 151180 - 151194.
- [2] WANG M Y, SETHI S P, VAN DE VELDE S L. Minimizing makespan in a class of reentrant shops[J]. Operations Research, 1997, 45(5): 702 - 712.
- [3] 轩华, 李冰, 王薛苑, 等. 带运输考虑的多阶段动态可重入混合流水车间调度[J]. 控制理论与应用, 2018, 35(3): 357 - 366.  
XUAN H, LI B, WANG X Y, et al. Multi-stage dynamic reentrant hybrid flowshop scheduling with transportation consideration[J]. Control Theory & Applications, 2018, 35(3): 357 - 366. (in Chinese)
- [4] 袁帅鹏, 李铁克, 王柏林. 带运输时间混合流水车间成组调度的协同进化文化基因算法[J]. 控制理论与应用, 2023, 40(3): 430 - 440.  
YUAN S P, LI T K, WANG B L. Co-evolutionary memetic algorithm for the hybrid flow shop group scheduling with transportation times[J]. Control Theory & Applications, 2023, 40(3): 430 - 440. (in Chinese)
- [5] 耿凯峰, 叶春明. 带工序跳跃的绿色混合流水车间机器与 AGV 联合调度[J]. 控制与决策, 2022, 37(10): 2723 - 2732.

- GENG K F, YE C M. Joint scheduling of machines and AGVs in green hybrid flow shop with missing operations[J]. Control and Decision, 2022, 37(10): 2723 - 2732. (in Chinese)
- [6] HIDRI L, ELSHERBEENY A M. Optimal solution to the two-stage hybrid flow shop scheduling problem with removal and transportation times[J]. Symmetry, 2022, 14(7): 1424.
- [7] LI W M, HAN D, GAO L, et al. Integrated production and transportation scheduling method in hybrid flow shop[J]. Chinese Journal of Mechanical Engineering, 2022, 35(1): 1 - 20.
- [8] YUAN S P, LI T K, WANG B L. A co-evolutionary genetic algorithm for the two-machine flow shop group scheduling problem with job-related blocking and transportation times[J]. Expert Systems with Applications, 2020, 152: 113360.
- [9] 尹爱军, 闫文涛, 张厚望. 面向多目标柔性作业车间调度的强化学习 NSGA-II 算法[J]. 重庆大学学报, 2022, 45(10): 113 - 123.  
YIN A J, YAN W T, ZHANG H W. Reinforcement learning NSGA-II for multi-objective flexible job shop scheduling[J]. Journal of Chongqing University, 2022, 45(10): 113 - 123. (in Chinese)
- [10] 张韵, 钟慧超, 张春江, 等. 基于机器学习的多策略并行遗传算法[J]. 计算机集成制造系统, 2021, 27(10): 2921 - 2928.  
ZHANG Y, ZHONG H C, ZHANG C J, et al. Multi-strategy parallel genetic algorithm based on machine learning[J]. Computer Integrated Manufacturing Systems, 2021, 27(10): 2921 - 2928. (in Chinese)
- [11] 李润佛, 张新宇, 李俊杰, 等. 基于强化学习的自学习遗传算法在船舶调度中的应用[J]. 大连海事大学学报, 2022, 48(3): 20 - 30.  
LI R F, ZHANG X Y, LI J J, et al. Application of self-learning genetic algorithm based on reinforcement learning in ship scheduling[J]. Journal of Dalian Maritime University, 2022, 48(3): 20 - 30. (in Chinese)
- [12] SHAHRABI J, ADIBI M A, MAHOOTCHI M. A reinforcement learning approach to parameter estimation in dynamic job shop scheduling[J]. Computers & Industrial Engineering, 2017, 110: 75 - 82.
- [13] CAO Z C, LIN C R, ZHOU M C. A knowledge-based cuckoo search algorithm to schedule a flexible job shop with sequencing flexibility[J]. IEEE Transactions on Automation Science and Engineering, 2021, 18(1): 56 - 69.
- [14] ALIPOUR M M, RAZAVI S N, FEIZI DERAKHSHI M R, et al. A hybrid algorithm using a genetic algorithm and multiagent reinforcement learning heuristic to solve the traveling salesman problem[J]. Neural Computing and Applications, 2018, 30(9): 2935 - 2951.
- [15] CHEN R H, YANG B, LI S, et al. A self-learning genetic algorithm based on reinforcement learning for flexible job-shop scheduling problem[J]. Computers & Industrial Engineering, 2020, 149: 106778.
- [16] 林飞龙, 陶泽, 王晓晨. 基于遗传算法的混合流水车间调度问题研究[J]. 沈阳理工大学学报, 2020, 39(2): 78 - 82.  
LIN F L, TAO Z, WANG X C. Research on scheduling problem of hybrid flow shop based on genetic algorithm[J]. Journal of Shenyang Ligong University, 2020, 39(2): 78 - 82. (in Chinese)

(责任编辑:宋颖韬)