

# 柔性作业车间调度系统的优化与应用

闫镜宣 王志兵

湖南工业大学计算机学院 湖南株洲 412007

**摘要:** 柔性作业车间调度系统由多个工件、多台机器以及一系列加工任务构成。与传统作业车间不同,柔性作业车间中的每个工件能够在多台不同的机器上进行加工,且加工时间和加工顺序会因机器的不同而存在差异。因此,柔性作业车间管理成为一项极为复杂的任务。

针对这一问题,本文提出一种在负荷需求、多种设备以及工件生产等综合因素影响下的柔性作业车间实时优化运行策略。该策略首先基于强化学习框架,将柔性作业车间中生产与运输活动之间的运行问题建模为马尔可夫决策过程。为有效利用柔性作业车间的互联结构,在近端策略优化算法(Proximal Policy Optimization, PPO)的基础上,本文设计了一种图注意力近端策略优化算法(Graph Attention Proximal Policy Optimization, GT-PPO)。该算法利用注意力机制和图神经网络学习柔性作业车间节点的相关性,以便学习各类环境下不同时段柔性作业车间的最优调度策略。

本次实验对比多种算法,结果显示近端策略优化算法PPO等算法在最大完工时间和相对百分比偏差方面有不同程度提升,其中图注意力机制GT-PPO(即Graph Attention Proximal Policy Optimization, GT-PPO)的最大完工时间提升高达35%,相对百分比偏差降低33%,性能极为优秀,显著优于其他算法,具有突出性能优势。

**关键词:** 柔性作业车间调度;近端策略优化算法;优化策略;马尔可夫决策

## 一、柔性作业车间调度系统及优化模型

### (一) 问题描述

给定一组工件和一组机器,每个工件由一系列的操作组成,每个操作可以在多台机器上进行加工。目标是确定每个操作在哪个机器上进行加工以及加工的顺序,使得某个或某些性能指标达到最优<sup>[1]</sup>。设柔性作业车间中有几个工件,记为 $J = \{J_1, J_2, \dots, J_n\}$ ,每个工件 $J_i$ 由一系列工序组成,记为 $O = \{O_{i1}, O_{i2}, \dots, O_{im}\}$ ,其中 $m$ 为工件 $J_i$ 的工序数量。同时,有 $k$ 台机器,记为 $M = \{M_1, M_2, \dots, M_k\}$ <sup>[2]</sup>。

决策变量的设置对于准确描述问题至关重要。其中,

$x_{ijh} = \begin{cases} 1, & \text{如果工件 } J_i \text{ 的工序 } O_{ij} \text{ 在机器 } M_h \text{ 上加工, 否则,} \\ 0, & \end{cases}$  它表示了工序与机器之间的分配关系。

$y_{ijj'hh'} = \begin{cases} 1, & \text{如果工件 } J_i \text{ 的工序 } O_{ij} \text{ 在机器 } M_h \text{ 上先于} \\ & \text{工序 } O_{ij'} \text{ 加工, 否则, 用于确定不同工件的工序在同一机器} \\ & \text{上的加工先后顺序。} \\ 0, & \end{cases}$   $C_{ij}$ 表示工件 $J_i$ 的工序 $O_{ij}$ 的完工时间,是衡量调度方案优劣得重要指标之一。

目标函数通常以最小化最大完工时间为目标,即<sup>[3]</sup> $\min C_{max} = \min(\max_j C_{ij})$ 这意味着要找到一种调度方,使得所有工件的最后一道工序的完工时间尽可能早,从而提高生产效率,降低生产成本。

约束条件则进一步规范了调度方案的可行性。机器约束要求对于任意机器 $M_h$ ,在同一时刻只能加工一个工

件的一个工序,即 $\sum_{i=1}^n \sum_{j=1}^m x_{ijh} \leq 1$ 。这确保了机器资源的合理利用,避免了资源冲突。工序顺序约束规定对于同一工件的相邻工序,后一道工序的开工时间必须晚于前一道工序的完工时间,即 $C_{ij} + t_{ijh} c_{ijh} \leq C_{i(j+1)}$ ,其中 $t_{ijh}$ 表示工件 $J_i$ 的工序 $O_{ij}$ 在机器 $M_h$ 上的加工时间。决策变量约束确保 $x_{ijh}, y_{ijj'hh'} \in \{0, 1\}$ ,使得决策变量具有明确的取值范围,便于问题的求解。

### 作者简介:

闫镜宣,性别:男,民族:汉,籍贯:湖南省株洲市,单位全称:湖南工业大学计算机学院,单位所在省市和邮编:湖南省株洲市,412007,学历:硕士,研究方向:工业大数据。

王志兵,性别:男,民族:汉,籍贯:湖南省株洲市,单位全称:湖南工业大学计算机学院,单位所在省市和邮编:湖南省株洲市,412007,学历:硕士,研究方向:工业互联网。

## (二) 约束条件

### 1. 机器约束

在柔性作业车间中, 每台机器在同一时刻只能加工一个工件的一个工序。这是因为机器的资源是有限的, 不能同时处理多个任务。用数学公式表示为:

$\sum_{i=1}^n \sum_{j=1}^m x_{ijh} \leq 1$ , 其中  $n$  为工件数量,  $m$  为工件的工序数量,  $h$  表示机器编号。这个约束确保了机器资源的合理分配, 避免了过度使用和冲突。

### 2. 工序顺序约束

对于同一工件的相邻工序, 后一道工序的开工时间必须晚于前一道工序的完工时间。这是由于生产工艺的要求, 工件的加工必须按照特定的顺序进行。用数学公式表示为:  $C_{ij} + t_{ijh} x_{ijh} \leq C_{i(j+1)}$ , 其中  $C_{ij}$  表示工件  $J_i$  的工序  $O_{ij}$  的完工时间,  $t_{ijh}$  表示工件  $J_i$  的工序  $O_{ij}$  在机器  $M_h$  上的加工时间<sup>[4]</sup>。这个约束保证了工件的加工顺序符合生产流程, 避免了工序的混乱和错误。例如, 如果一个工件的第一道工序在某台机器上加工完成后, 下一道工序只能在第一道工序完成后才能开始<sup>[5]</sup>。

### 3. 决策变量约束

决策变量  $x_{ijh}$  和  $y_{ijj'hh}$  必须在特定的取值范围内, 即  $x_{ijh}, y_{ijj'hh} \in \{0, 1\}$ 。其中  $x_{ijh}$  表示工件  $J_i$  的工序  $O_{ij}$  是否在机器  $M_h$  上加工,  $y_{ijj'hh}$  表示工件  $J_i$  的工序  $O_{ij}$  在机器  $M_h$  上是否先于工件  $J_{i'}$  的工序  $O_{i'j'}$  加工<sup>[9]</sup>。这个约束确保了决策变量的明确性和可操作性, 便于问题的求解<sup>1</sup>。例如, 当  $x_{ijh}=1$  时, 表示工件  $J_i$  的工序  $O_{ij}$  在机器  $M_h$  上加工; 当  $x_{ijh}=0$  时, 表示不在该机器上加工。

### 4. 时间约束

在实际生产中, 所有工序的完工时间必须在规定的时间范围内。这是为了满足客户的需求和生产计划的要求。用数学公式表示为  $C_{ij} \leq T_{max}$ , 其中  $C_{ij}$  表示工件  $J_i$  的工序  $O_{ij}$  的完工时间,  $T_{max}$  为最大允许完工时间。这个约束确保了生产进度的可控性, 避免了延误和超时。例如, 如果某个工件的工序完工时间超过了最大允许完工时间, 那么可能会影响整个生产计划的执行, 导致客户不满和经济损失。

## (三) 马尔可夫决策过程

### 1. 状态空间

状态可以用一个向量表示, 包括当前正在加工的工件、机器的状态、剩余加工时间等信息。设状态为  $S$  可以表示为  $S = (\omega_1, \omega_2, \dots, \omega_n, m_1, m_2, \dots, m_k, t_1, t_2, \dots, t_l)$ , 其, 其中  $\omega_i$  表示工件  $i$  的状态 (如正在加工的工序、已完成的工序等),  $m_j$  表示机器的状态 (如是否

空闲、正在加工哪个工件的工序等),  $t_k$  表示剩余加工时间等相关参数。

### 2. 动作空间

动作是在每个决策时刻可以选择的操作, 如选择哪个工件进行加工、选择哪台机器进行加工等。设动作  $a$ , 可以表示为  $a = (i, j)$ , 其中  $i$  表示选择的工件编号,  $j$  表示选择的机器编号。动作空间的定义决定了调度策略的可行范围。

### 3. 奖励函数

奖励函数用于衡量每个动作的优劣, 通常根据目标函数来设计。例如, 当最大完工时间减少时给予正奖励, 反之给予负奖励。设奖励函数为  $r(s, a, s')$ , 其中  $s$  为当前状态,  $a$  为采取的动作,  $s'$  为下一个状态。可以设计为

$$r(s, a, s') = \begin{cases} +k & \text{如果采取动作 } a \text{ 后最大完工时间减少} \\ -k & \text{如果采取动作 } a \text{ 后最大完工时间增加} \\ 0 & \text{其他情况} \end{cases}$$

其中  $k$  为一个正的常数。奖励函数的设计直接影响到调度策略的优化方向。通过合理地设置奖励函数, 可以引导算法朝着最优的调度方案进行搜索。

### 4. 目标函数

目标是最大化累计奖励, 即通过选择最优的动作序列, 使得长期的奖励最大化。设累计奖励为  $R = \sum_{t=0}^{\infty} \gamma^t r(s_t, a_t, s_{t+1})$ , 其中  $\gamma$  为折扣因子,  $0 < \gamma < 1$ ,  $s_t$  表示时刻  $t$  的状态,  $a_t$  表示时刻  $t$  采取的动作。目标函数可以表示为  $\frac{\max}{\pi} E[R | \pi]$ , 其中  $\pi$  表示策略, 即从状态到动作的映射,  $E[R | \pi]$  表示在策。

## 二、基于PPO的柔性作业车间调度策略优化

### (一) PPO

近端策略优化 (PPO) 是一种先进的强化学习算法, 具有高效、稳定的特点。它通过优化策略网络, 使得智能体在与环境交互的过程中能够学习到最优的策略。设策略网络为  $\pi_{\theta}(a | s)$ , 表示在状态  $s$  下采取动作  $a$  的概率, 其中  $\theta$  为网络参数。PPO 通过最大化目标函数来更新参数  $\theta$ , 目标函数通常采用期望回报的形式:

$$J(\theta) = \sum_{t=0}^{\infty} \gamma^t R(s_t, a_t | \theta) \quad (\text{公式1})$$

其中  $\gamma$  为折扣因子,  $R(s_t, a_t | \theta)$  为在状态  $s_t$  采取动作  $a_t$  后的回报, 取决于策略网络参数  $\theta$ 。

### (二) 图注意力

设图结构  $G = (V, E)$ , 其中  $V$  为节点集合 (代表工件和机器),  $E$  为边集合。对于节点  $i$ , 其特征向量为点

$h_i$ 。通过图注意力机制计算注意力系数：

$$e_{ij} = a(Wh_i, Wh_j) \quad (\text{公式2})$$

其中  $a$  为注意力函数，通常是一个前馈神经网络， $W$  为可学习的权重矩阵。然后对注意力系数进行归一化：

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in N(i)} \exp(e_{ik})} \quad (\text{公式3})$$

其中  $N(i)$  表示节点  $i$  的邻居节点集合。最后得到节点  $i$  的新特征表示：

$$h_i^i = \sigma\left(\sum_{j \in N(i)} a_{ij} Wh_j\right) \quad (\text{公式4})$$

其中  $\sigma$  为激活函数。

### (三) GT-PPO 算法

GT-PPO 算法结合了图注意力机制和 PPO 算法。首先利用图注意力机制计算工件与机器之间的关系权重，然后将这些权重融入到 PPO 的策略网络中。具体来说，在计算策略网络的输出时，考虑图注意力机制得到的权重：

$$\pi_{\theta}(a|\delta) = \frac{\exp(f_{\theta}(s, a) \cdot \omega(s))}{\sum_{a' \in A} \exp(f_{\theta}(s, a') \cdot \omega(s))} \quad (\text{公式5})$$

其中  $f_{\theta}(s, a)$  为策略网络的原始输出， $\omega(s)$  为根据图注意力机制得到的状态权重向量， $A$  为动作空间。通过这样的方式，GT-PPO 算法能够更好地利用工件与机器之间的关系，提高调度的效率和质量。

## 三、算例分析

### (一) 柔性作业车间调度场景设置

设置不同规模的柔性作业车间调度场景，包括工件数量、机器数量、操作复杂度等。通过改变这些参数，可以评估算法在不同情况下的性能。

### (二) 算法及算例设置

将 GT-PPO 算法与其他传统算法进行比较，设置相同的初始条件和终止条件。通过多次运行算法，取平均值作为最终结果，以提高结果的可靠性。模型如下：

假设对 GT-PPO 算法、传统算法 1、传统算法 2、等分别运行  $N$  次；对于 GT-PPO 算法，每次运行得到的最大完工时间记为  $T_{gtppo}^i$ ，其中  $i=1, 2, \dots, N$ 。

(1) 对于传统算法 1，每次运行得到的最大完工时间记为  $T_{trad1}^i$ ，其中  $i=1, 2, \dots, N$ 。

(2) 对于传统算法 2，每次运行得到的最大完工时间记为  $T_{trad2}^i$ ，其中  $i=1, 2, \dots, N$ 。

因此：

(1) GT-PPO 算法的平均最大完工时间为：

$$\bar{T}_{gtppo} = \frac{1}{N} \sum_{i=1}^N T_{gtppo}^i$$

(2) 传统算法 1 的平均最大完工时间为：

$$\bar{T}_{trad1} = \frac{1}{N} \sum_{i=1}^N T_{trad1}^i$$

(3) 传统算法 2 的平均最大完工时间为：

$$\bar{T}_{trad2} = \frac{1}{N} \sum_{i=1}^N T_{trad2}^i$$

通过比较  $\bar{T}_{gtppo}$ 、 $\bar{T}_{trad1}$  和  $\bar{T}_{trad2}$  的大小，可以评估 GT-PPO 算法相对于传统算法的性能优劣。

如果  $\bar{T}_{gtppo} < \bar{T}_{trad1}$  且  $\bar{T}_{gtppo} < \bar{T}_{trad2}$ ，则说明 GT-PPO 算法在相同初始条件和终止条件下，得到的平均最大完工时间更短，性能更优。

### (三) 算例结果及分析

#### 1. 初始条件

设初始状态为一组工件、多台机器以及一系列加工任务的初始分配情况。可以用一个向量表示为： $S_a = (\omega_1^0, \omega_2^0, \dots, \omega_n^0, m_1^0, m_2^0, \dots, m_k^0, t_1^0, t_2^0, \dots, t_l^0)$ ，其中  $\omega_i^0$  表示工件  $i$  的初始状态， $m_j^0$  表示机器  $j$  的初始状态， $t_k^0$  表示初始剩余加工时间等相关参数。

#### 2. 终止条件

可以设置为当所有工件的所有工序都完成加工，或者达到最大运行时间等。例如，当满足  $\forall i, j; C_{ij} \leq T_{max}$  且所有机器都处于空闲状态时，算法终止。其中  $T_{max}$  为最大允许完工时间。通过以上数学模型和评估方法，可以有效地比较 GT-PPO 算法与其他传统算法在柔性作业车间调度中的性能表。

#### 3. 算例结果

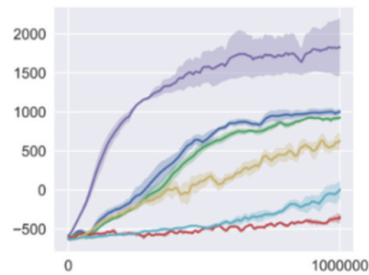


图 (1) PPO 算法结果

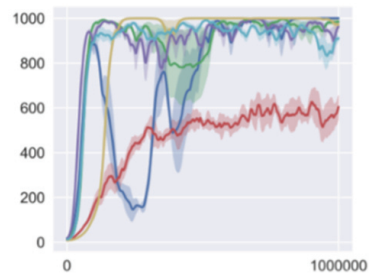


图 (2) GT-PPO 算法结果

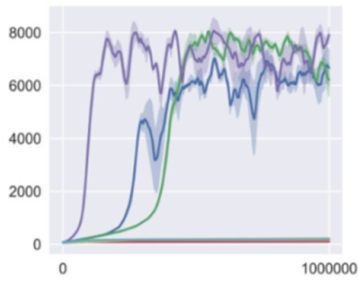


图 ( 3 ) A2C 算法结果

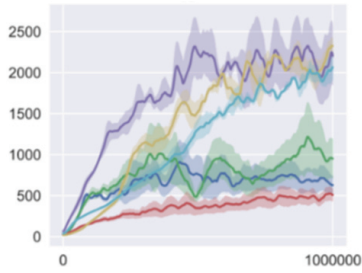


图 ( 4 ) A3C 算法结果

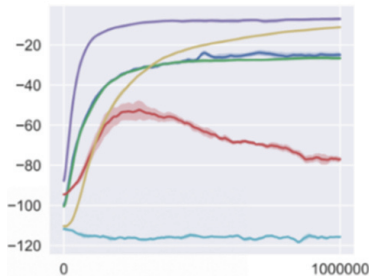


图 ( 5 ) DDPG 算法结果

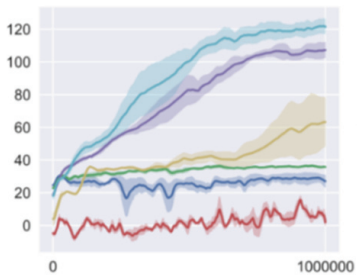


图 ( 6 ) SAC 算法结果

表 3-1 PPO、GT-PPO、A2C、A3C、DDPG、SAC  
6 种算法对比结果

算法	最大完工时间提升情况	相对百分比偏差降低情况	性能评价
近端策略优化算法 PPO	11%	16%	/
图注意力机制 GT-PPO	35%	33%	性能优秀
优势演员-评论家算法 A2C	7%	23%	/
异步优势演员-评论家算法 A3C	13%	16%	/
深度确定性策略梯度 DDPG	17%	14%	/
软性演员-评论家算法 SAC	7%	13%	/

本次实验对6种算法进行了对比分析。其中，近端策略优化算法PPO的最大完工时间提升了11%，相对百分比偏差降低了16%。优势演员-评论家算法A2C的最大完工时间提升了7%，异步优势演员-评论家算法A3C提升了13%，深度确定性策略梯度DDPG提升了17%，软性演员-评论家算法SAC提升了7%。而图注意力机制GT-PPO的最大完工时间提升高达35%，相对百分比偏差降低了33%，表现出极为优秀的性能。

### 结束语

柔性作业车间调度系统由于其独特的结构和复杂性，对管理提出了极高的挑战。本文针对这一问题提出的柔性作业车间实时优化运行策略，尤其是设计的图注意力近端策略优化算法，在实验中展现出了卓越的性能。

通过与多种算法的对比，可以观察到，GT-PPO算法在最大完工时间提升和相对百分比偏差降低方面表现极为突出，显著优于近端策略优化算法、优势演员-评论家算法、异步优势演员-评论家算法、深度确定性策略梯度和软性演员-评论家算法等其他算法。这一成果为柔性作业车间的管理提供了更为有效的解决方案，有助于提高生产效率、降低成本，并增强企业在市场中的竞争力。

### 参考文献

- [1]张惺卓,王鼎鑫,李富娜.智能边缘与AI融合推动产业创新升级[N].滨城时报,2024-08-08(001). DOI: 10.28019/n.cnki.nbfs.2024.001528.
- [2]高玉芳,姬智,赵康健,等.LEO星座边缘计算网络中的动态计算卸载策略[J].通信学报,2024,45(07):61-69.
- [3]Hou W,Zhang S.Assembly line balancing and optimal scheduling for flexible manufacturing workshop[J].Journal of Mechanical Science and Technology,2024,38(6):2757-2772. DOI:10.1007/s12206-024-2206-2.
- [4]Li X,Yang Q,Qin W.Robust optimization scheduling of microgrid considering a tiered carbon trading mechanism and flexible loads[J].IOP Publishing Ltd,2024. DOI:10.1088/1742-6596/2788/1/012035.
- [5]Zhao Y,Luo X,Zhang Y.The application of heterogeneous graph neural network and deep reinforcement learning in hybrid flow shop scheduling problem[J].Computers&Industrial Engineering,2024,187(Jan.):109802.1-109802.17.