

引文格式: 郭剑, 史耀耀, 胡昊, 等. 基于工业物联网的混流车间机器人自适应调度[J]. 航空制造技术, 2021, 64(5):42-51.

GUO Jian, SHI Yaoyao, HU Hao, et al. Adaptive robot scheduling using SP-MCTS for industrial internet of things-enabled hybrid flow shop[J]. Aeronautical Manufacturing Technology, 2021, 64(5):42-51.

基于工业物联网的混流车间 机器人自适应调度

郭 剑^{1,2}, 史耀耀^{1,2}, 胡 昊¹, 陈 振^{1,2}, 张军锋³, 赵 盼^{1,2}

(1. 西北工业大学机电学院, 西安 710072;

2. 西北工业大学航空发动机高性能制造工业和信息化部重点实验室, 西安 710072;

3. 西安建筑科技大学机电工程学院, 西安 710055)

[摘要] 随着工业物联网技术与人工智能技术深度融合, 物料机器人已广泛应用于物联网车间中。针对车间存在实时动态变化和状况不确定等诸多复杂因素, 本文提出以组分层建树和以实时状态为根节点的 SP-MCTS (Single-player monte-carlo tree search algorithm) 搜索方法实现车间自适应调度决策。该方法将机器人调度问题转化为马尔科夫决策过程 (Markov decision process, MDP), 并详细描述车间状态、动作、奖励和策略的表示方法。在实时调度过程中, 该搜索方法依据工件组分层建树, 搜索中只考虑相邻两组间的状态关系, 从而简化计算难度。在子树搜索中, 应用 SP-MCTS 以实时状态为根节点进行搜索, 同时应用扩展方法和剪支方法进行策略探索和信息累积, 使得在子树内实时状态节点越深, 就越能够快速精准获取最优策略。最后, 通过实际案例模拟分析, 验证了该方法的有效性和优越性。

关键词: 工业物联网; 混流车间; 机器人; 马尔科夫决策过程; SP-MCTS

DOI: 10.16080/j.issn1671-833x.2021.05.042



郭 剑
博士研究生, 研究方向为智能生产与调度、智能机器人。

随着计算机技术、传感器技术、无线通信技术和机器人技术快速发展, 制造业正向高度自动化和智能化方向发展^[1]。自新冠肺炎全球蔓延以来, 5G 技术、大数据中心、工业物联网、人工智能和机器人技术在“抗疫”中发挥了极大作用, 力促我国加速工业“新基建”的投入。在这一背景下, 企业将大量使用移动机器人和自动导引运输车 (Automated guided vehicle, AGV) 进行车间调度和仓库物料配送, 实现智能化生产和供应链管理^[2]。

当前, 制造企业面临着激烈的市场竞争和不断上升的制造成本等方面的挑战, 同时产品需求频繁波动也

给企业带来较大压力^[3]。而物料调度不仅能够优化完工时间, 同时在遵从客户需求下能够找到设备、路径和运送时间等的最佳组合。近几年国内外对于物料机器人调度方面的研究主要集中于传统静态调度问题, 即在所有信息稳定且已知的情况下通过启发算法求解。朱琳等^[4]提出了一种改进的遗传算法进行自动导航小车的任务分配和任务排序优化。Mousavi 等^[5]结合遗传算法和粒子群算法优化柔性车间中的多目标 AGV 调度问题。Gen 等^[6]以最小完工时间为优化目标, 应用基于混合优先级的遗传算法求解 AGV 调度问

题。Jerald 等^[7]提出自适应的遗传算法和蚁群算法,以实现 AGV 调度中的综合惩罚成本和机器空闲时间的最小化。Batur 等^[8]在两种邻域结构下应用模拟退火启发算法求解混流车间 (Hybrid flow shop, HFS) 中的物料机器人调度问题,并通过对比验证了该方法的有效性。Zabihzadeh 等^[9]应用双信息素的蚁群算法和遗传算法分别优化混流车间中多物料机器人调度问题,结果表明遗传算法性能更优。然而,在实际生产中,难以提前获知所有加工任务信息,并且在动态的车间环境中存在诸多不确定信息(如紧急任务,设备故障,任务返工等),因此难以将静态调度用于实际生产中。

近年来,由于无线射频识别和条形码等工业物联网技术已逐步扩展到制造环境中,车间实时数据经多源数据融合处理后形成制造业的大数据环境^[10]。而这些来自各种传感器的实时数据流并不能直接适用于传统的物料处理系统,因此,越来越多的学者开始关注物料机器人实时调度问题,以解决运行环境中动态调度问题。Lee 等^[11]指出在实时数据环境下,制造系统中的物料处理系统将更加高效。Zhang 等^[12]基于 AGV 实时状态提出物料运送的优先策略用于物料调度。Li 等^[13]提出了同时面临多个客户需求时的智能仓库系统 AGV 调度机制。研究表明,在车间调度中使用多种调度规则 (Multiple scheduling rules, MSRs) 比使用单一调度规则能更大程度地提高生产效率^[14],刘想德^[15]应用层次分析法 (Analytic hierarchy process, AHP) 在组合规则中选择调度规则,避免了传统调度人员在制定调度方案时凭借经验或自己偏好选择调度规则。上述文献中,多在实时环境下应用先进先出 (First come first served, FCFS), 最短距离 (Shortest distance first, SDF) 和最早截止时间 (Earliest due date first, EDD) 等规则调度物

料机器人,或基于传统知识库选择规则,因此实时适应性较差。

为解决上述问题,越来越多的学者开始应用强化学习算法求解车间物料调度问题。Wang^[16]提出基于加权的强化学习车间自适应调度方法,通过聚类和动态搜索确定适合策略。另外 Zhang 等^[17]通过融合遥感信息和 AGV 信息实时共享,并优化 AGV 速度避免碰撞,同时提出紧急任务超车策略实现 AGV 调度。Shahrabi 等^[18]引入带有 Q 因子的强化学习算法以提高动态车间调度性能,仿真结果表明该方法相比传统单规则调度和邻域搜索,优化性能较高。

尽管上述文献中物料机器人调度研究取得了一些研究成果,但考虑到在工业物联网下车间调度属于复杂的实时优化问题,且鲜有涉及大批量和多种类任务的动态混流车间调度问题的相关研究。因此本文提出应用 SP-MCTS 的机器人自适应调度方法,是在基于工业物联网的混流车间中以最小化完工时间和延迟率为优化目标,将调度系统实时信息作为输入状态,通过搜索学习快速选出实时状态下适应的调度规则和机器人,然后发送调度指令进行工件运送。最后通过模拟试验对所提方法进行分析评估,验证其可靠性和有效性。

混流车间中物料调度问题

混流车间中物料机器人的调度

问题可以描述为:多组工件 $j_e = \{1, 2, \dots, n_e\}$ 需经 $O_{j1} \rightarrow O_{j2} \rightarrow \dots \rightarrow O_{js}$ 道工序加工。每道工序 $S = \{1, 2, \dots, s\}$ 内有 $M_k \geq 1$ 个同类并行设备,且工件 j 只能占用工序内一台平行设备加工。 R 个机器人对线内物料进行卸载、运送、装载。由于各工序间不设缓冲工位,物料完成加工后将在设备上等待运送,因此当阻塞发生时机器人选择运送工件的实时调度策略尤为重要。如图 1 所示,物料运送过程满足以下约束:(1) 开始调度时,各组内的工件释放时间相同;(2) 工件 j 在每道工序只能加工一次,且每台设备同一时刻只能加工一个工件;(3) 工件在占用每台设备的时间由两部分组成,即加工时间和等待运送时间;(4) 物料的运输时间由设备的距离决定。

为模型建立需要,引入参数变量用于描述数学模型,见表 1。

混流车间实时调度是依据实时状态选择规则,然后按照规则在车间约束关系下选择运送物料的方法,因此混流车间中物料运送的约束关系如下^[19]:

$$\text{Minimize } C_{\max} = \text{Max } C_{js} \quad (1)$$

$$\forall j \in \{1, 2, \dots, N\}$$

$$\text{Minimize } D_r = N_d / N_T \quad (2)$$

即

$$\text{Minimize } Z = \beta_1 \cdot \text{Minimize } C_{\max} + \beta_2 \cdot \text{Minimize } D_r \quad (3)$$

s. t.

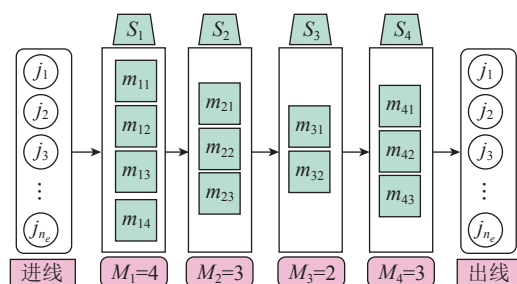


图1 混合流水车间物料流动示意图

Fig.1 Diagram of material flow in hybrid flow shop

表1 符号说明
Table 1 Problem notation

变量	说明	
基本变量	j_e	为工件编号, $j \in \{1, 2, 3, \dots, n_e\}$, e 为组别
	s	为工序编号, 包括进线站和出线站, $s \in \{1, 2, 3, \dots, S\}$
	M_s	为每道工序 s 中同类并行设备数, 进出线站各 n_e 个工位, $m \in \{1, 2, 3, \dots, M_s\}$
	r	为单抓手机器人人数, $r \in \{1, 2, 3, \dots, R\}$
	β_1, β_2	权重因子
参数变量	P_{js}	为工序 s 内工件 j 加工所需时间为 P_{js} , $P_{j1} = P_{js} = 0$
	C_{js}	为工件 j 从工序 s 准备运送至工序 $s+1$ 的时间
	BM	极大数
	$t_{j,s-1,m',s,m}$	为工件 j 从工序 $s-1$ 上加工设备 m' 运送至工序 s 上加工设备 m' 所用时间
	w_{js}	为工件 j 在工序 s 上完成加工后的等待时间
	D_t	延迟率
	N_d	延迟工件总数
	N_T	总工件数
决策变量	Y_{jsm}	如果工序 s 内设备 m 加工工件 j , 值为 1, 否则为 0
	$X_{jj's}$	如果工件 j 在工序 s 段完成加工早于 j' , 值为 1, 否则为 0
	$Z_{j,s,f}$	第 f 次运送操作是工序 s 内夹取工件 j , 运送至下个工序加工设备, 值为 1, 否则为 0

$$C_{js} = C_{j,s-1} + t_{j,s-1,m',s,m} + P_{js} + w_{js} \quad \forall j \in \{1, 2, \dots, n_e\}; s \in \{1, 2, \dots, S\} \quad (4)$$

$$C_j = C_{js} \quad \forall j \in \{1, 2, \dots, n_e\}$$

$$C_{j,s} + BM(2 - (Y_{j,s,m} + Y_{j,s-1,m'})) \geq C_{j,s-1} + t_{j,s-1,m',s,m} + P_{j,s} \quad (5)$$

$$\forall j \in \{1, 2, \dots, n_e\}; s \in \{1, 2, \dots, S\};$$

$$m \in \{1, 2, \dots, E_{j,s}\};$$

$$m' \in \{1, 2, \dots, E_{j,s-1}\}; \quad (6)$$

$$\sum_{m=1}^{M_s} Y_{jsm} = 1 \quad \forall j \in \{1, 2, \dots, n_e\};$$

$$s \in \{1, 2, \dots, S\} \quad (7)$$

$$C_{j',s-1} + BM(4 - Y_{jsm} - Y_{j'sm} - Y_{j,s-1,m'} - X_{jj's}) \geq C_{js} + t_{j,s-1,m',s,m}$$

$$\forall j, j' \in \{1, 2, \dots, n_e\}; j \neq j';$$

$$s \in \{1, 2, \dots, S\};$$

$$m \in \{1, 2, \dots, E_{j,s}\};$$

$$m' \in \{1, 2, \dots, E_{j',s-1}\} \quad (8)$$

$$C_{j,s-1} + BM(3 - Y_{jsm} - Y_{j'sm} - Y_{j,s-1,m'} + X_{jj's}) \geq C_{j',s} - t_{j',s-1,m',s,m}$$

$$\forall j, j' \in \{1, 2, \dots, n_e\}; j \neq j';$$

$$s \in \{1, 2, \dots, S\};$$

$$m \in \{1, 2, \dots, E_{j,s}\};$$

$$m' \in \{1, 2, \dots, E_{j',s-1}\} \quad (9)$$

$$\sum_{f=1}^W Z_{j,s,f} = 1 \quad \forall j \in \{1, 2, \dots, n_e\};$$

$$s \in \{1, 2, \dots, S-1\} \quad (10)$$

$$\left(\sum_{j=1}^n \sum_{s=1}^{S-1} Z_{j,s,f} \right) = 1$$

$$\forall f \in \{1, 2, \dots, W\} \quad (11)$$

$$\left(\sum_{f=1}^{f'} Z_{j,s,f} \right) - BM \times (1 - Z_{j,s-1,f'}) \leq 0$$

$$\forall f \in \{1, 2, \dots, W\}; s \in \{2, \dots, S-1\};$$

$$f' \in \{1, 2, \dots, W\} \quad (12)$$

$$\left(\sum_{f=1}^{f'} Z_{j,s-1,f} \right) - BM(3 - (Y_{j,s,m} + Y_{j',s,m} + Z_{j',s,f'} - X_{j,j',s})) \leq 0$$

$$\forall j, j' \in \{1, 2, \dots, n_e\}, j \neq j';$$

$$s \in \{2, \dots, S-1\}; m \in \{1, 2, \dots, E_{j,s}\};$$

$$f' \in \{1, 2, \dots, W\} \quad (13)$$

$$\left(\sum_{f=1}^{f'} Z_{j',s-1,f} \right) - BM(4 - (Y_{j',s,m} + Y_{j,s,m} + Z_{j,s,f'} + X_{j,j',s})) \leq 0$$

$$\forall j, j' \in \{1, 2, \dots, n_e\}, j \neq j';$$

$$s \in \{2, \dots, S-1\}; m \in \{1, 2, \dots, E_{j',s}\};$$

$$f' \in \{1, 2, \dots, W\} \quad (14)$$

上述模型中: 式(1)最小化最大完工时间; 式(2)最小化延迟率; 式(3)最小化综合优化目标; 式(4)每个工件各阶段完工时间; 式(5)各工件的完工时间; 式(6)用来约束工件按照工序加工; 式(7)工件在各工序段中只能在一个加工设备上完成加工; 式(8)和式(9)用来约束在每个工序中同一设备工件的加工顺序; 式(10)和式(11)确保每个机器人在任何时候只能运送一个工件; 式(12)确保每个工件运送过程按照工序的先后顺序进行; 式(13)和式(14)机器人的调度顺序符合工件在同一设备的处理顺序。

马尔科夫决策模型建立

文中将多机器人调度过程转化为马尔科夫决策过程^[20-21], 通过四元组描述为 (S, A, P, R) 。其中 S 为调度车间状态空间, 车间现场状态 $s \in S$; A 表示动作空间, 由调度规则、工序 ID、机器人 ID 组成; $P(s, A, s') = 1$, 一旦动作选择, 状态从 s 转换至 s' 的概率为 1; R 表示在状态 s 下执行动作 a 状态转换至 s' 所得报酬。如图 2 所示, 根据混流车间的调度问题数学约束关系建立物料车间环境, 马尔科夫决策过程即为智能系统根据当前车间环境选择一个机器人运送动作 a 作用于车间, 完成运送后车

间状态发生改变,同时产生一个奖励信号反馈给智能系统。智能体再根据奖励函数和车间状态 s 选择下一个动作 a' 。针对多机器人混流车间调度的实际情况,将车间运行中的状态、动作和报酬函数定义如下。

1 状态

将车间设备特征信息与任务信息组成一个状态变量来引导调度决策,即在时间 t 时状态向量 $S_t = (T_t, F_t, R_t)$, T_t 代表待运送工件特征信息, F_t 代表车间状态信息, R_t 代表机器人特征信息。

(1) T_t 为 $3 \times s$ 矩阵, $T_t(:, 1)$ 至 $T_t(:, s)$ 依次为进线工位至最后一道加工工序的统计信息。 $T_t(1, :)$ 为每道工序中待运送工件数目, $T_t(2, :)$ 代表每道工序待运送工件平均剩余时间, $T_t(3, :)$ 代表每道工序待运送工件中最小剩余时间。

$$T_t(2, k) = \frac{\sum_{l=1}^{T_t(1, k)} t_{k, l}}{T_t(1, k)} \quad (15)$$

$$k \in \{1, \dots, s\}; l = T_t(1, k)$$

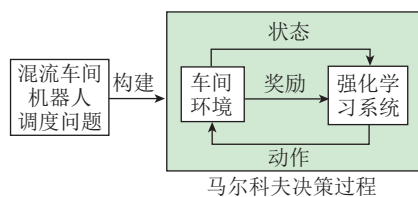


图2 车间调度的马尔科夫决策过程

Fig.2 Diagram of Markov decision process for flow shop scheduling

$$T(3, k) = \min_{k, l} t_{k, l} \quad (16)$$

$$k \in \{1, \dots, s\}; \forall l \in \{1, \dots, T(1, k)\}$$

式中 $t_{k, l}$ 代表在第 k 阶段待运送工件 l 的剩余时间,该参数反映工件的紧迫程度。

(2) F_t 为车间现场状态描述矩阵, F_t 为 $3 \times s$ 矩阵, $F_t(1, :)$ 表示每道工序的空闲设备数,例如 $F_t(1, 1) = 1$ 即为工序 1 中有 1 台空闲设备。 $F_t(2, :)$ 为每道工序中正在加工状态的设备数, $F_t(3, :)$ 代表每道工序中被待运送工件占用的设备数。

(3) R_t 为 $2 \times i$ 矩阵,表示各机器人的工作状态。 i 与机器人 ID 对应, $R_t(1, :)$ 表示机器人的状态, $R_t(1, 2) = 1$ 为 2 号机器人“运送”, $R_t(1, 2) = 0$ 为“空闲”。 $R_t(2, :)$ 为机器人当前位置 ID。

2 动作

动作即为机器人的运送行为,其通过工序 ID,调度规则和机器人 ID 组成向量 $a_i = (Sa_i, Ru_i, Rb_i)$ 。工序 ID 选择调度工件所在工序,调度规则选择运送工件,同时,机器人 ID 确定哪一台机器人被选中用于调度。文中选用 5 种调度规则用于缓冲工位调度,SSO 用于工序内待运送工件调度。调度规则的详细描述如表 2 所示, $Sa_i = \{进线缓冲工位 = 0, 工序 = 1, \dots, s\}$, $Ru_i = \{SPT = 1, LPT = 2, SLACK = 3, SPT/TWK = 4,$

$LPT/TWK = 5, SSO = 6\}$, $Rb_i = \{1 = \text{robot } 1, 2 = \text{robot } 2, \dots, r = \text{robot } r\}$,通过上述动作设计,时间 t 时将在指定规则下分配机器人运送指定工序下的唯一任务。

3 奖励函数

奖励函数被用来对当前行为和调度策略进行评估,由于 SP-MCTS 是通过最终收益回溯逐个更新节点信息,而本文将最小化最大完工时间和延迟率作为综合评价指标,因此奖励函数设计为:

$$Z = (C_1 - C_n)g\beta_1 + (1 - ND)g\beta_2 \quad (17)$$

式(17)中 C_1 为子树内第 1 次搜索至当前所有工件完成加工时所得时间, C_n 为子树内实时状态节点的第 n 次搜索所得时间,其中项 $C_1 - C_n$ 若大于 0 则取实际差值,若小于 0 则取 0 值。 ND 表示当前子树延迟率(延迟工件包括当前搜索中出线延迟的工件和线内剩余时间小于 0 的工件,总数即为子树中的所有工件), β_1, β_2 为权重因子。

方法论述

1 多层子树设计

SP-MCTS 是一类利用模拟运行过程来评价非终端状态的搜索树算法。模拟过程中选择执行动作,直到达到一个终端状态,并基于多次模

表2 调度规则

Table 2 Scheduling rules

名称	规则描述	优点	缺点
SPT	选择加工时间最短的任务	提高车间完工效率	不能保证紧急件即时加工
LPT	选择加工时间最长的任务	提高设备利用率	不能及时处理紧急件
SLACK	选择松弛量最小的工件优先调度	保证紧急件完工,降低延迟率	难以保证生产效率
SPT/TWK	选择工序加工时间与总加工时间比值最小的工件	保证当前工序的完工率,提高后续工序设备利用率	不能及时处理紧急件
LPT/TWK	选择工序加工时间与总加工时间比值最大的工件	保证当前工序的设备利用率,保证后续完工率	不能及时处理紧急件
SSO	选择后继工序加工时间最短的工件	确保下个工序的设备利用率	不能及时处理紧急件,仅考虑下个工序设备利用率

拟中平均奖励,以此评估每个动作的潜力^[22]。针对工业物联网下混流车间多机器人调度问题提出基于SP-MCTS算法的多层子树自适应优化方法,如图3所示。图3中根据各子任务和当前车间环境依次建立多个子树(TR_1, TR_2, \dots, TR_n)进行策略优化,如子树 TR_1 进行实时优化时,是以车间当前状态作为根节点,组内所有工件完成加工作为目标状态,应用SP-MCTS算法进行优化,并取使节点收益最大的分支规则作为运送动作。根据实时状态依次进行优化,最终取最优路径节点 S_{43} (组间临界实时状态,即进线工位清空,工件全部进线加工时)作为下一个子树 TR_2 的根节点原始环境,此时下一组进入缓冲区的工件与当前车间环境建立

子树 TR_2 根节点,从而依次优化直至各组工件完成加工。

如图4子树模拟优化过程,工序内和进线出线缓冲区的数字“1,2,3”分别代表工件等待、加工、和出线3种状态,各工序的平行设备数为“2 2 1 2”。在子树 TR_1 内,若实时车间为状态 s_1 ,优化完成后执行策略 $\pi(s)$ 至 s_2 ,然后保留节点 s_2 下的子树信息,丢弃其他节点信息。此时等待下一个实时调度状态信息与 s_2 进行对比,若相同则以根节点 s_2 继续优化(子树节点信息可累积使用)。依次搜索和节点更新,最终直至该组调度完成。

2 基于SP-MCTS强化学习

结合SP-MCTS算法寻优特性,设计适应于文中混流车间调度问题

的选择、扩展、模拟和回溯方案^[23],如图5所示。

2.1 选择

选择步骤始于当前根节点 s_0 ,终止于树中叶节点 s_L ,其选择策略根据UCT^[24](Upper confidence bounds applied to trees)算法构建,其改进如下:

$$Q^\oplus(s, a) = Q(s, a) + c \cdot \sqrt{2 \ln N(s) / N(s, a)} + \sqrt{(\sum q(s, a)^2 - N(s, a) \times Q(s, a)^2 + D) / N(s, a)} \quad (18)$$

$$\pi(s) = \operatorname{argmax}_a Q^\oplus(s, a) \quad (19)$$

式(18)中 $N(s)$ 代表状态 s 被访问的次数, $N(s, a)$ 表示状态 s 时执行运送动作 a 的次数, $Q(s, a)$ 表示节点 s 处选择运送动作 a 的平均收益, $\sum q(s, a)^2$ 表示到目前为止在状态 s 处执行动作 a 获得收益 $Q(s, a)$ 的平方和。式(19)中 $\pi(s)$ 代表状态 s 下的最优策略。

2.2 扩展与模拟

如果搜索至叶节点 s_L 且该节点遍历次数大于 P (P 为节点扩展临界值)时,应用选择策略后执行动作 a 得到状态 s_{CL} 。此时将状态 s_{CL} 扩展为树中叶节点且节点信息初始为 $\{N(s_L, a)=0, Q(s_L, a)=0, N(s_{CL})=0\}$ 。

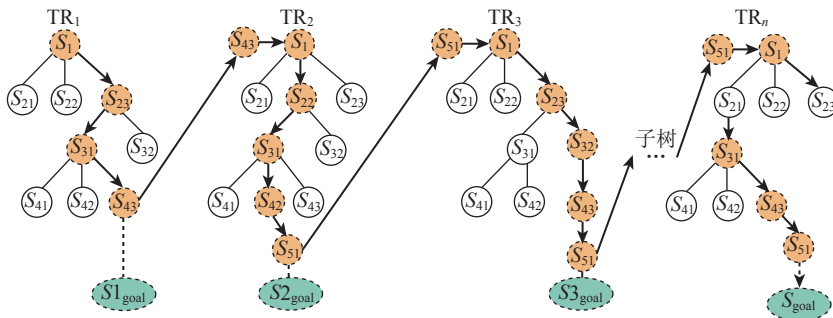


图3 基于SP-MCTS算法的多层子树示意图

Fig.3 Diagram of multi-layer subtree based on SP-MCTS algorithm

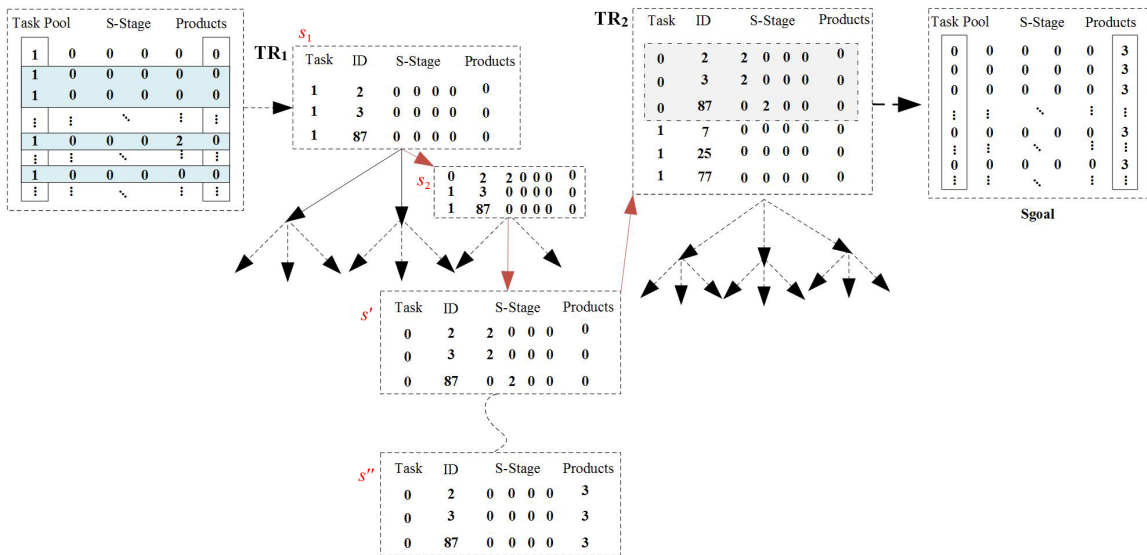


图4 子树优化过程

Fig.4 Subtree optimization process

如果搜索至叶节点 s_L 且该节点遍历次数小于 P 时, s_L 将作为初始的模拟节点。在模拟过程中, 每一次执行模拟策略前需对线内工件评估紧迫度, 若剩余时间 $t_{kr} < 2T_r$ (t_{kr} 任务 k 的剩余时间, T_r 为车间中设备最大距离间机器人运送时间), 则执行 LLF 作为模拟策略, 否则执行既定的 3 种模拟策略 SPT, LPT 和随机策略 (Random)。

2.3 回溯

当模拟至终止状态时, 信息更新从模拟的初始叶节点 s_L 回溯至根节点 s_0 。更新各遍历节点信息:

$$N(s) \leftarrow N(s) + 1 \quad (20)$$

$$N(s, a) \leftarrow N(s, a) + 1 \quad (21)$$

$$Q(s, a) \leftarrow Q(s, a) + \frac{z - Q(s, a)}{N(s, a)} \quad (22)$$

$$\sum q(s, a)^2 \leftarrow \sum q(s, a)^2 + z^2 \quad (23)$$

式(20)更新节点 s 被访问的总次数。式(21)记录 s 状态下动作 a 被执行次数, 式(22)更新节点 s 执行动作 a 的平均收益。式(23)计算收益求平方和。

3 系统架构

如上所述, 文中提出了基于 SP-MCTS 求解工业物联网的混流车间多机器人调度问题, 其体系架构如图 6 所示。

该架构包含 4 个子模块, 基于工业物联网的混流车间, 系统状态模块, SP-MCTS 模块和规则与执行动作

作模块。在机器人完成运送或空闲时且有新工件组进入缓冲区或状态改变时, 调度过程立即触发。此时, 工业物联网采集混流车间内各制造资源实时状态, 并且发送实时信息至系统状态模块以提取关键状态信息, 这些信息包括任务信息和设备信息。上述这些关键信息最终被传输至 SP-MCTS 模块用于优化调度策略。数据经 SP-MCTS 模块处理后, 输出数据至规则与执行动作模块。最终, 规则与执行动作模块根据规则选择调度工件并规划机器人传输动作, 从而实现基于工业物联网混流车间自适应调度的闭环反馈机制。

4 基于 SP-MCTS 系统优化过程

SP-MCTS 系统自适应优化过程可分为扩展方法和剪支方法。如图 4 所示, 以实时状态 s_1 为根节点, 多次迭代后建立了 TR_1 , 满足停止条件时, 则依据策略 $\pi(s)$ 执行动作 a 得到实时调度状态 s_2 , 此时以 s_2 为根节点继续搜索, 实时状态依次建树的过程称为扩展过程。选择动作 a 时, 树内则保留节点 s_2 以下分支, 剪掉其他分支即剪支过程, 这样依次搜索直至目标状态 s_{goal} 。其求解混流车间多机器人调度问题的流程图如图 7 所示, SP-MCTS 算法伪代码如算法 1 所示 (表 3)。

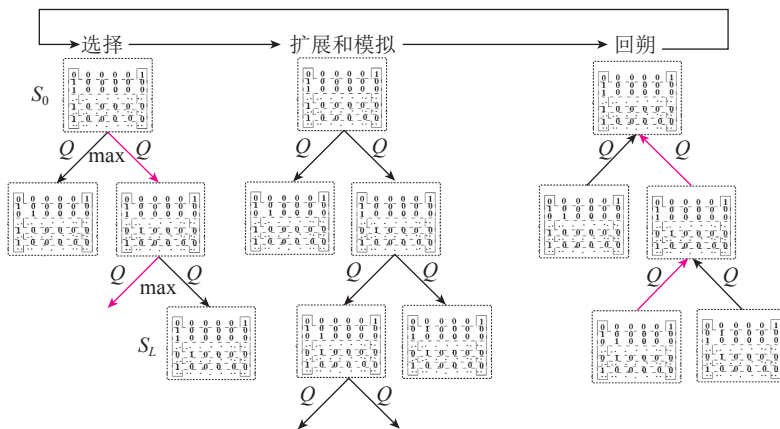


图5 SP-MCTS算法原理图

Fig.5 Schematic diagram of SP-MCTS algorithm

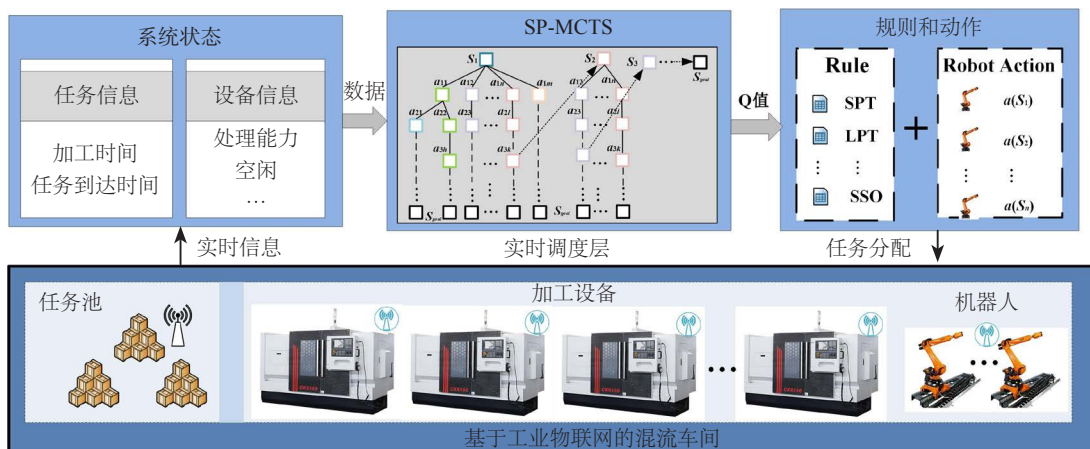


图6 应用SP-MCTS的多机器人调度方法架构

Fig.6 Architecture of multi-robot scheduling approach using SP-MCTS

方法验证与讨论

为评估所提方法的有效性,在此提出基于工业物联网混流车间多机器人调度的案例用于模拟验证。

1 案例描述

应用西门子 Tecnomatix 数字化平台作为案例模拟软件,如图 8 所示,混流车间主要由 9 台加工设备以及 4 道工序组成,其中 9 台加工设备、进线缓冲工位、出线缓冲工位均匀分布于机器人轨道两侧,且有两台机器人用于物料运送。

2 应用 SP-MCTS 进行多机器人调度

应用 TensorFlow 软件和西门子 Tecnomatix 数字化平台联合模拟工业物联网下混流车间多机器人实时调度系统,如图 9 所示。

该系统软件可分为两部分,分别是在 TensorFlow 平台使用 python 开发的 SP-MCTS 程序和使用西门子 Tecnomatix 平台模拟软件开发的模拟程序。其整个模拟程序分为以下几个子模块:设备管理、任务管理、状态管理、通信模块和调度指令模块。在模拟车间生产过程中,设备管理模块负责车间内加工设备、机器人和各类传感器的信息管理;任务管理模块负责所有工件的管理;状态处理模块作为关键的信息处理模块,其作用是处理从设备管理和任务管理模块发送的实时设备和工件信息,并将提取的关键信息发送至通信模块。通信模块是在 SP-MCTS 程序和模拟程序间建立一个通信网络,实时传输状态信息和调度指令信息。SP-MCTS 优化策略根据当前状态优化调度策略,将调度策略发送至调度指令模块。最后,调度指令模块根据策略协调规则和机器人执行调度任务。

3 案例描述

混流车间内的设备、任务、机器人以及计算机模拟环境等参数设置如下:

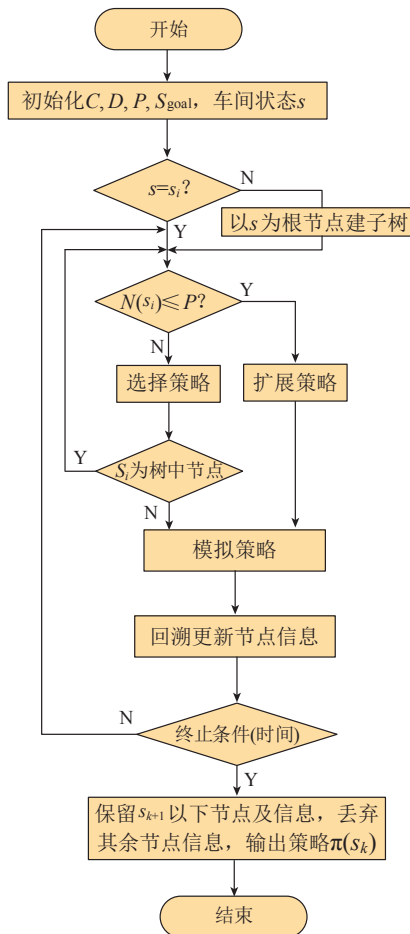


图7 混流车间多机器人自适应调度流程
Fig.7 Multi-robot scheduling process with SP-MCTS

(1) 总的加工任务数为 500, 进线缓冲工位有 20 个工位, 根据紧迫度确定工件的进线次序, 如果存在紧迫度相同的产品, 则按交货期由早到晚排序^[23]。从而将 500 个任务分成 25 组, 依次送入进线缓冲工位。

(2) 工件在每个工序的加工时间服从 (10, 50) 的均匀分布。

(3) 为使问题简单化, 文中在机器人运输过程中不考虑最优路径和机器人碰撞冲突问题, 同时工件运输时间服从 (0.5, 1) 均匀分布。

(4) 参数 (C, D) 设置为 (0.5, 10000)^[25], $P=3, \beta_1=0.5, \beta_2=2$ 。

(5) 计算机模拟环境为 i7-6700 CPU@3.40GHz, 16G RAM。

依据上述参数取一组车间加工参数, 应用 SP-MCTS 算法对模拟车

表3 算法1: 基于SP-MCTS算法的优化主程序

Table 3 Algorithm 1: Optimized main procedure based on SP-MCTS

1: 初始化参数和目标状态 :C, D, P, Sgoal	
状态输入 s	
	If 当前任务组为最后一组任务
	终止状态和临界状态均为 Sgoal
	else 当前任务组工件全部进入车间加工时为临界状态 s', 线内工件全部加工完成为目标状态 s"
	end
	If s==si (si 为子树中节点)
	k=i
	执行步骤 1
	else 以 s 作为根节点建立子树
	i=1, k=1, si=s
	end
	While 停止条件不满足时 do
3:	步骤 1: 选择策略
4:	if N(si) <= P then
	执行步骤 3
	else 应用式 (19) 选择调度策略 a.
	执行策略 a, i=i+1
	更新状态 si
	end
	步骤 2: 扩展策略
8:	if si 为子树叶节点 then
9:	return 步骤 1
10:	else 扩展 si 为树中叶节点, 执行步骤 3
11:	end
	步骤 3: 模拟策略
	叶节点 si 做为模拟起始节点
	选择模拟策略模拟至终止节点得到报酬 Z
	步骤 4: 回溯
31:	从叶结点 si 到 sk 应用公式 (20) → (23) 对路径上的节点信息更新, i=k。
32:	if 停止条件不满足时
33:	return 步骤 1
34:	else 依据式 (19) 选择策略 pi(sk) 得到节点 sk+1, 节点 sk+1 及以下树节点及信息保留, 其余扩展节点及信息丢弃, 输出策略 pi(sk)。
35:	end if
51	end while

间中物料进行实时调度。如图 10 所示,以第一组进线任务的路径节点与最大收益值为例,可以看出在前期搜索中随着节点经验信息依次累加,收

益数值快速提高,后期搜索中随着状态空间不断缩小,树节点越深其收益数值变化越趋平缓,且数据波动越趋收窄。因而说明应用 SP-MCTS 在实时优化中能够快速准确地选择最优策略进行物料调度。

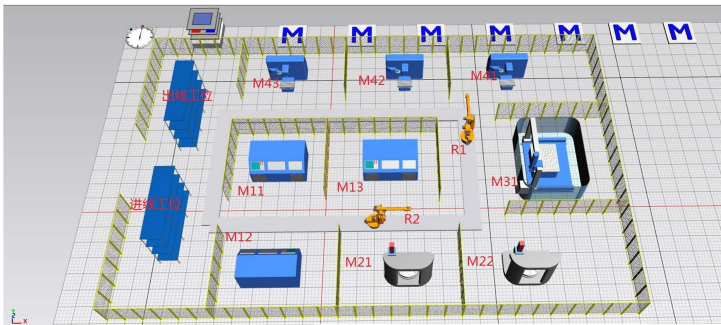


图8 混流车间多机器人调度模拟环境

Fig.8 Simulation of multi-robot scheduling on shop floor environment

4 结果分析

在此引入 RLVNS^[18]、AHP^[15]、SPT、LPT 和 SLACK 与本文所提方法进行对比,在相同的环境下处理随机生成的 500 个任务,通过完工时间和延迟率进行综合比较。

随机产生 10 组案例车间参数应用各方法分别优化,然后取平均值进行比较,如图 9 所示。多机器人调度方法比较示意图见图 11,可以直观看出 5 种方法在完工时间和延迟率两个指标下的比较结果。与 SPT、LPT 和 SLACK 规则相比,SP-

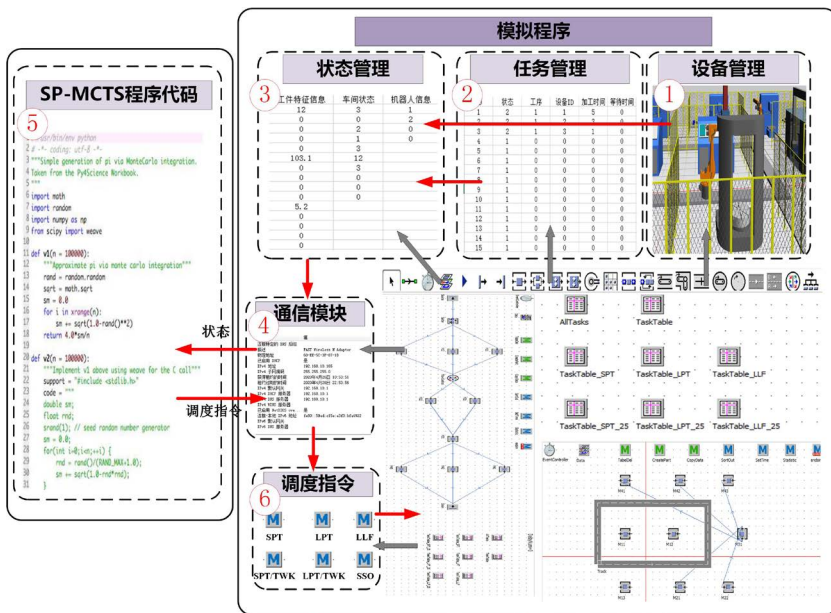


图9 混流车间多机器人调度系统

Fig.9 Implementation of the SP-MCTS based multi-robot scheduling system

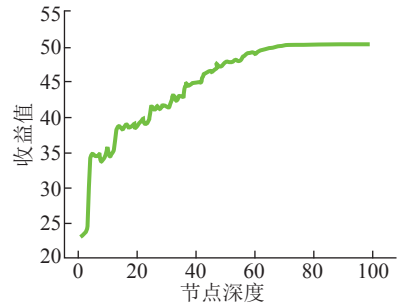


图10 子树内节点深度与收益值关系

Fig.10 Relationship between node depth and revenue value in subtree

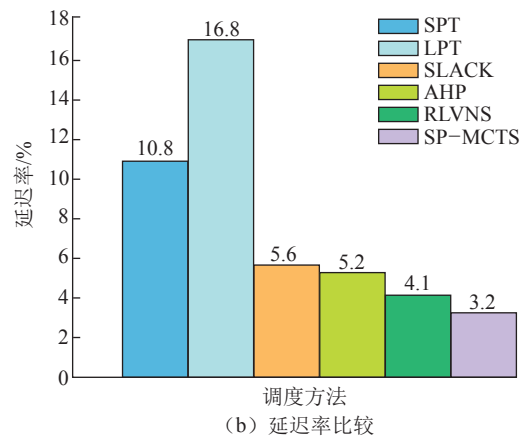
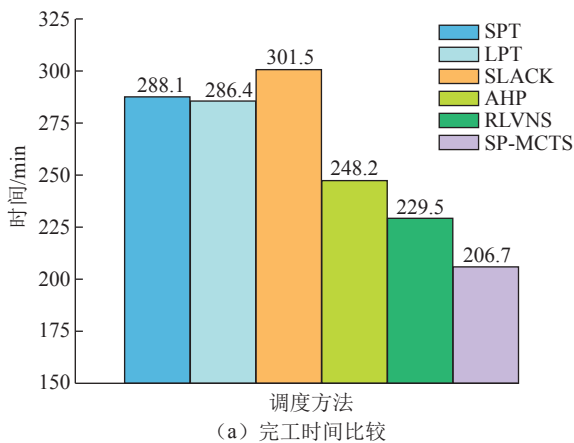


图11 各方法比较结果示意图

Fig.11 Diagram of comparison results of scheduling methods

MCTS 算法调度所得完工时间分别降低了 28.3%、27.8% 和 31.4%，同时延迟率分别降低了 70.4%、81% 和 42.9%；而相比于 AHP 和 RLVNS 方法，完工时间分别降低了 16.7% 和 9.9%，延迟率分别降低了 38.5% 和 22%。由此看出，单一的 SPT、LPT 和 SLACK 规则调度，虽然能快速响应，但其适应性差且调度质量难以保证，而应用 SP-MCTS 算法搜索寻优后建立信息网络，其结合现场状况在多种规则中选择适应当前状态的最优调度策略，从而得到更好的求解质量。

基于 SP-MCTS 解的质量优于多规则组合 AHP 方法和强化学习 RLVNS 方法，可知，与多规则组合 AHP 方法相比，SP-MCTS 算法的适应性更强；而 RLVNS 方法只考虑第一工序的邻域搜索学习，不能够分辨工序之间工件调度时的信息差异，存在明显的局限性。因此，该模拟结果验证了在工业物联网下应用 SP-MCTS 对混流车间进行多机器人调度的有效性和优越性。

结论

为最小化工业物联网混流车间中多机器人调度的完工时间和延迟率，本文提出基于 SP-MCTS 算法的自适应调度方法。由于工件分组进入缓冲工位，因而采用分层子树搜索机制，应用 SP-MCTS 以实时状态为节点进行搜索和信息积累，然后根据现场情况快速选择合适的规则和物料机器人，从而达到了自适应调度目的。主要结论如下：

(1) 分层子树优化，使得各组工件从一个完整的根节点出发同时兼顾相邻子树，简化了模型难度。

(2) 状态依次作为根节点搜索，能够根据累积的节点信息，精确快速定位动作分支，提高了搜索效率。

(3) 在模拟实验中与 AHP 方法和调度规则比较，验证了 SP-MCTS

方法的适应性；同时 SP-MCT 求解的完工时间和延迟率均优于 RLVNS 方法，验证了 SP-MCTS 方法在求解混流车间调度问题时的搜索理念的优越性。

本文调度方法架构简明，易于实现，为进一步解决工业物联网下车间实时调度问题提供了参考。下一步，将结合深度学习方法建立各规则的概率网络和价值网络，从而使调度系统成为一种有意识的智慧大脑。

参考文献

[1] ZHANG Y F, ZHANG G, QU T, et al. Analytical target cascading for optimal configuration of cloud manufacturing services[J]. Journal of Cleaner Production, 2017, 151: 330-343.

[2] DEMESURE G, DEFOORT M, BEKRAR A, et al. Decentralized motion planning and scheduling of AGVs in an FMS[J]. IEEE Transactions on Industrial Informatics, 2018, 14(4): 1744-1752.

[3] ZHANG Y F, GUO Z G, LV J, et al. A framework for smart production-logistics systems based on CPS and industrial IoT[J]. IEEE Transactions on Industrial Informatics, 2018, 14(9): 4019-4032.

[4] 朱琳, 范秀敏, 何其昌. 柔性生产系统配料区多自动导航小车调度优化[J]. 计算机集成制造系统, 2012, 18(6): 1168-1175.

ZHU Lin, FAN Xiumin, HE Qichang. Scheduling optimization for multi-AGVs in batching area of flexible production system[J]. Computer Integrated Manufacturing Systems, 2012, 18(6): 1168-1175.

[5] MOUSAVI M, YAP H J, MUSA S N, et al. Multi-objective AGV scheduling in an FMS using a hybrid of genetic algorithm and particle swarm optimization[J]. PLOS One, 2017, 12(3): e0169817.

[6] GEN M, LIN L, JO J B. Hybrid genetic algorithm for designing logistics network, VRP and AGV problems[M]//Studies in computational intelligence. Heidelberg: Springer, 2009: 123-139.

[7] JERALD J, ASOKAN P, SARAVANAN R, et al. Simultaneous scheduling of parts and automated guided vehicles in an FMS environment

using adaptive genetic algorithm[J]. International Journal of Advanced Manufacturing Technology, 2006, 29(5): 584-589.

[8] BATUR G D, EROL S, KARASAN O E. Robot move sequence determining and multiple part-type scheduling in hybrid flexible flow shop robotic cells[J]. Computers & Industrial Engineering, 2016, 100: 72-87.

[9] ZABIHZADEH S S, REZAEIAN J. Two meta-heuristic algorithms for flexible flow shop scheduling problem with robotic transportation and release time[J]. Applied Soft Computing, 2016, 40: 319-330.

[10] ZHANG Y F, ZHANG G, WANG J Q, et al. Real-time information capturing and integration framework of the Internet of manufacturing things[J]. International Journal of Computer Integrated Manufacturing, 2015, 28(8): 811-822.

[11] LEE J, LAPIRA E, BAGHERI B, et al. Recent advances and trends in predictive manufacturing systems in big data environment[J]. Manufacturing Letters, 2013, 1(1): 38-41.

[12] ZHANG Y F, ZHANG G, DU W, et al. An optimization method for shopfloor material handling based on real-time and multi-source manufacturing data[J]. International Journal of Production Economics, 2015, 165: 282-292.

[13] LI Z, BARENJI A V, JIANG J Z, et al. A mechanism for scheduling multi robot intelligent warehouse system face with dynamic demand[J]. Journal of Intelligent Manufacturing, 2020, 31(2): 469-480.

[14] SABUNCUOGLU I. A study of scheduling rules of flexible manufacturing systems: A simulation approach[J]. International Journal of Production Research, 1998, 36(2): 527-546.

[15] 刘想德. 作业车间实时调度若干关键问题研究[D]. 重庆: 重庆大学, 2013.

LIU Xiangde. Research on some critical issues about job shop real time scheduling[D]. Chongqing: Chongqing University, 2013.

[16] WANG Y F. Adaptive job shop scheduling strategy based on weighted Q-learning algorithm[J]. Journal of Intelligent Manufacturing, 2020, 31(2): 417-432.

[17] ZHANG Y F, ZHU Z F, LV J. CPS-based smart control model for shopfloor material

handling[J]. IEEE Transactions on Industrial Informatics, 2018, 14(4): 1764–1775.

[18] 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.

[19] 郭剑, 史耀耀, 张激, 等. 基于 SP-MCTS 算法的混合流水线机器人调度[J]. 计算机集成制造系统, 2019(9):2208–2218.

GUO Jian, SHI Yaoyao, ZHANG Wei, et al. SP-MCTS algorithm for hybrid flow shop scheduling problem with robotic transportation[J]. Computer Integrated Manufacturing Systems, 2019(9):2208–2218.

[20] PUTERMAN M L. Markov decision processes: Discrete stochastic dynamic

programming[M]. John Wiley & Sons, 1994.

[21] BOUTILIER C. Knowledge representation for stochastic decision processes[M]// Artificial intelligence today. Heidelberg: Springer, 1999.

[22] POWLEY E J, COWLING P I, WHITEHOUSE D. Information capture and reuse strategies in Monte Carlo Tree Search, with applications to games of hidden information[J]. Artificial Intelligence, 2014, 217: 92–116.

[23] AUER P, CESA-BIANCHI N, FISCHER P. Finite-time analysis of the multiarmed bandit problem[J]. Machine Learning, 2002, 47(2–3): 235–256.

[24] 谢志强, 常宁宁, 杨静. 基于交货期紧迫度的综合调度算法[J]. 机械工程学报,

2011, 47(22):181–190.

XIE Zhiqiang, CHANG Ningning, YANG Jing. Integrated scheduling algorithm based on delivery dates urgency degree[J]. Journal of Mechanical Engineering, 2011, 47(22): 181–190.

[25] SCHADD M P D, WINANDS M H M, TAK M J W, et al. Single-player monte-carlo tree search for same game[J]. Knowledge-Based Systems, 2012, 34: 3–11.

通讯作者: 史耀耀, 教授、博士, 研究方向为智能制造与物料调度、数控工艺装备(整体叶盘高效强力复合数控铣削)、高速高效数控加工、加工表面光整技术, E-mail: shiyy@nwpu.edu.cn.

Adaptive Robot Scheduling Using SP-MCTS for Industrial Internet of Things-Enabled Hybrid Flow Shop

GUO Jian^{1,2}, SHI Yaoyao^{1,2}, HU Hao¹, CHEN Zhen^{1,2}, ZHANG Junfeng³, ZHAO Pan^{1,2}

(1. School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China;

2. Key Laboratory of High Performance Manufacturing for Aero Engine, Ministry of Industry and Information Technology, Northwestern Polytechnical University, Xi'an 710072, China;

3. School of Mechanical and Electrical Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

[ABSTRACT] With the deep integration of industrial internet of things (IIoT) and artificial intelligence (AI) technology, automated guided vehicles (AGVs) and mobile robots have been widely used in internet of things-enabled floor shop. In view of many complex factors such as real-time dynamic changes and uncertain conditions in the workshop, the SP-MCTS (Single-player Monte-Carlo tree search algorithm) method with each job group as a subtree and real-time state as the root node is proposed to implement adaptive scheduling of workshop. The problem of robot scheduling is formulated as a Markov decision process (MDP) in which state representation, action representation, reward function, and optimal policy, are described in detail. In the real-time scheduling process, the proposed search method establishes a subtree for each job group, and only the state relationship between two adjacent groups is considered in optimization, thereby the calculation difficulty is simplified. In the subtree search process, SP-MCTS is used to search with the real-time state as the root node. At the same time, the expansion method and the pruning method are used to carry out strategy exploration and information accumulation respectively. Therefore, the deeper the real-time status node in the subtree, the faster and more accurate the optimal strategy is obtained. The case study based on a real-world shop is proven and the results validate the effectiveness and superiority of the proposed approach.

Keywords: Industrial internet of things (IIoT); Hybrid flow shop; Robot; Markov decision process; SP-MCTS

(责编 大漠)