

12-13-2019

Energy Efficiency Data Mining and Scheduling Optimization of Discrete Workshop

Yugu Lin

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China;

Wang Yan

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, 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.

Energy Efficiency Data Mining and Scheduling Optimization of Discrete Workshop

Abstract

Abstract: This paper addresses the optimization of energy consumption in discrete workshops and establishes the energy efficiency optimization model of discrete workshops. The relationship between data mining and knowledge discovery is established. Through *scheduling data preprocessing* and *C4.5 decision tree learning algorithm*, the discovery of scheduling knowledge is realized. Energy efficiency optimization calculation is achieved in discrete workshops by *the combination of scheduling knowledge and improved differential evolution algorithm (IDE)*. By comparing with TLBO, GA and PSO, the feasibility of IDE algorithm is verified.

Keywords

energy efficiency, data mining, scheduling knowledge, differential evolution algorithm

Recommended Citation

Lin Yugu, Wang Yan. Energy Efficiency Data Mining and Scheduling Optimization of Discrete Workshop[J]. Journal of System Simulation, 2019, 31(12): 2702-2711.

离散车间能效数据挖掘及调度优化

林雨谷, 王艳

(江南大学, 物联网技术应用教育部工程中心, 江苏 无锡 214122)

摘要: 离散车间的加工过程是由不同零件的加工子过程串联或者并联构成的复杂柔性加工过程。面向离散车间的能效优化, 建立了离散车间能效优化模型; 关联数据挖掘和知识发现关系; 通过对调度数据预处理和 C4.5 决策树学习算法, 实现对调度知识的发现; 将调度知识和改进后的差分进化算法(IDE)相结合用于离散车间的能效优化计算。通过和 TLBO、GA 和 PSO 算法仿真比较验证了 IDE 算法的可行性。

关键词: 能效; 数据挖掘; 调度知识; 差分进化算法

中图分类号: TP391.9 文献标识码: A 文章编号: 1004-731X (2019) 12-2702-10

DOI: 10.16182/j.issn1004731x.joss.19-FZ0257

Energy Efficiency Data Mining and Scheduling Optimization of Discrete Workshop

Lin Yugu, Wang Yan

(Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China)

Abstract: This paper addresses the optimization of energy consumption in discrete workshops and establishes the energy efficiency optimization model of discrete workshops. The relationship between data mining and knowledge discovery is established. Through *scheduling data preprocessing* and *C4.5 decision tree learning algorithm*, the discovery of scheduling knowledge is realized. Energy efficiency optimization calculation is achieved in discrete workshops by *the combination of scheduling knowledge and improved differential evolution algorithm (IDE)*. By comparing with TLBO, GA and PSO, the feasibility of IDE algorithm is verified.

Keywords: energy efficiency; data mining; scheduling knowledge; differential evolution algorithm

引言

在经济高速发展的当今社会, 每年能源消耗的总量在不断的增加, 由此而引起的能源浪费和环境污染的问题也日趋严重, 因此制造业的绿色节能问题亟待解决。离散车间加工过程的调度研究考虑如何提高车间生产效率, 缩短生产周期并

且降低能耗, 调度研究对于车间的实际生产具有现实的意义^[1]。

对于离散车间的优化调度问题的研究成果很多^[2-3]。在知识指导情况下, 针对离散车间的调度问题, 我们有更广阔的范围选取优化目标和规则。因此, 研究如何从原始的调度数据中挖掘调度知识和实现知识发现, 并将调度知识应用于调度方案以优化性能是具有现实意义的。

数据挖掘是知识发现的一个必不可少的步骤, 但这两者可以交互使用。大量的离散车间的调度数据中隐含了许多有价值的调度知识, 通过挖掘和利用, 对管理调度知识和优化决策, 进一步提高调度



收稿日期: 2019-03-08 修回日期: 2019-06-26;
基金项目: 国家自然科学基金(61973138);
作者简介: 林雨谷(1994-), 男, 江苏盐城, 硕士生, 研究方向为智能调度优化; 王艳(1978-), 女, 江苏盐城, 博士, 教授, 研究方向为基于大数据知识自动化的离散制造能耗网络协同优化。

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

• 2702 •

方案的可行性有着很重要的意义。因此, 对于数据的分析和处理成为关键的问题。R. Balasundaram 等提出了一种利用数据挖掘生成调度规则的方法, 通过将学习算法应用到作业车间调度中来发现调度序列^[4]。针对车间作业问题, Yanhong Wang 等提出了一种新的基于数据的机制。该机制基于决策树算法和规则挖掘技术, 可以解决大量信息和突发时间的调度问题^[5]。

本文提出了一种基于数据驱动的离散车间调度的优化方法。

1 离散车间能效优化模型

离散制造系统包括了 2 个阶段, 图 1 分别是加工任务分配阶段和加工阶段, 其加工任务分配阶段是在系统运行前完成的。离散制造系统不同的调度任务, 不同的调度目标, 请求的资源不同导致调度方案也会多种多样。离散制造系统的能耗主要在运输和加工上, 通过合理的调度方案, 降低系统的能耗, 提升耗能的效率, 从而使系统能效能够达到最优。因此, 对系统能耗进行优化, 我们一般选择最小化设备总能耗、最小化完工时间和设备的最大利用率等作为优化目标对调度问题进行求解, 产生优化的调度方案。

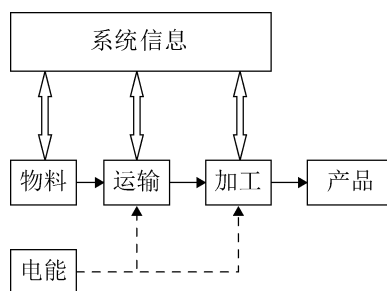


图 1 离散制造系统能耗流量图

Fig. 1 Energy consumption flow diagram of discrete manufacturing system

建立精确的能效优化模型是对离散车间进行能耗优化的前提。我们需要将离散车间的实际问题映射到数学模型去解决, 通过对加工过程建立更加清晰的数学表达式, 更加合理的且具有实际意义的

对能效优化的问题进行求解。

1.1 离散制造系统的调度问题描述

离散制造系统的调度问题可以做如下描述: n 个工件需要通过 m 台设备进行加工, 不同的设备对于不同的工件所加工的时间时不同的, 每个工件有多道工序。 $J = \{J_1, J_2, J_3 \dots J_i, \dots J_n\} (i \in [1, n])$, $M = \{M_1, M_2, M_3 \dots M_i, \dots M_n\} (i \in [1, m])$, J 为设备集, M 为机器集。且能效优化的约束条件有: 不同的待加工工件具有不同的工艺路线; 每个工件的每道工序一旦被加工, 需要一次性完成, 中途不能被中断(机器坏机除外); 第 j 台机器 M_j 在任意时刻能且只能加工某个工件其中的一道工序; 在初始时刻所有工件的加工优先级相同; 不同工件的不同工序具有相同的加工顺序优先级; 同一工件的各道工序之间允许等待。能效优化参数的定义如下

k : 表示设备号;

p : 表示每个工件的最大工序数;

$M = \{M_k | 1 \leq k \leq m\}$: 表示设备集;

$J = \{J_i | 1 \leq i \leq n\}$: 表示工件集;

O_{ij} : 表示工件 J_i 的第 j 道工序;

$M_{ij} = \{M_k | X_{ijk} = 1\}$: 表示工序 O_{ij} 的可用设备集;

P_{ij} : 表示工件 J_i 的第 j 道工序能在多台设备上加工;

R_c : 表示 O_{ij} 和 O_{ab} 可以在设备 M_k 上加工;

t_i^E : 表示完成工件 i 最后工序加工的时刻;

t_{ijk} : 表示工序 O_{ij} 在设备 M_k 上加工的时间;

$T_{S_{ijk}}$: 表示工序 O_{ij} 在设备 M_k 上加工的开始时间;

$T_{E_{ijk}}$: 表示工序 O_{ij} 在设备 M_k 上加工的结束时间;

$T_{s_{i,j}}$: 表示工件 J_i 的第 j 道工序的加工时间;

$T_{E_{i,j-1}}$: 表示工件 J_i 的第 $j-1$ 道工序的结束时间;

$T_{S_{k,l+1}}$: 表示设备 M_k 在加工第 $l+1$ 道工序的开始时间;

$T_{E_{k,l}}$: 表示设备 M_k 在加工第 l 道工序的结束

时间;

tw_{ijk} : 表示在加工工序 O_{ij} 前设备 M_k 的等待

时间;

P_k : 表示设备 M_k 的等待功率;

e_{ijk} : 表示 O_{ij} 在设备 M_k 上的耗能;

EM_k : 表示设备 M_k 上加工工件的耗能;

E : 表示加工所有工件的总耗能。

1.2 能效优化约束条件数学表达

(1) 机器约束

$$T_{S_k,l+1} > T_{E_k,l} \quad (1)$$

式(1)保证了设备 M_k 在加工时,能且只能加工某一工件的某一工序。

(2) 过程约束

$$T_{E_{ijk}} - T_{S_{ijk}} = t_{ijk}, P_{ijk} = 1 \quad (2)$$

式(2)保证了某个工件的某道工序一旦开始加工,不可以中断。

$$T_{S_i,j} > T_{E_i,j-1} \quad (3)$$

式(3)保证了同一工件的加工工序的先后顺序。

(3) 工艺约束

$$T_{S_i,j} > T_{E_i,j-1}, P_{ijk} = P_{i(j-1)k} = 1 \quad (4)$$

式(4)保证了同一个工件的不同工序按照工艺路线加工。

1.3 能效优化目标函数

根据离散车间加工过程的特点,本文建立了以设备加工的总能耗和完工时间为优化目标的目标函数对离散车间的能效进行优化。离散车间的机器设备的工作过程包括设备的等待过程和设备的加工过程,因此在生产过程中,设备在加工前的等待能耗和加工过程中的加工能耗为离散车间的加工总能耗。本文用 f_1 表示最小化生产能耗,用设备的最小化最大完工时间 f_2 表示完工时间:

$$f_1 = \min E$$

$$E = EM + \min \sum_{i=1}^n \sum_{j=1}^p tw_{ijk} \times P_k$$

$$EM = \sum_{i=1}^n \sum_{j=1}^p e_{ijk}$$

$$f_2 = \min \{ \max(t_1^E, t_2^E \dots t_n^E) \}$$

2 调度数据挖掘

数据挖掘是知识发现(Knowledge Discovery in Databases, KDD)的一个步骤。数据挖掘的任务包括分类、回归聚类和依赖模型等。在分类任务中,通常以 If-then 规则的形式表示发现的知识,能够让我们对发现知识的理解变得更加容易。

2.1 离散车间调度数据挖掘框架

调度数据挖掘重点是如何建立调度数据和知识之间的映射关系,将传统的离散车间调度后的大量的数据隐含了许多的有价值的知识,经过挖掘可以对离散车间调度有指导和优化作用。本文应用数据挖掘的目的根据调度数据不同的属性,基于决策树对知识的挖掘,预测离散车间产品在机器上的加工顺序。如图 2 所示是调度数据挖掘的框架,通过数据预处理将连续属性离散化,对数据不同的属性进行取值,采用决策树方法可以从决策树中获得的有价值的规则知识更新到知识库,运用到作业车间调度问题的方案求解。

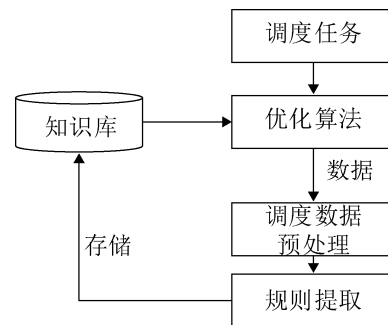


图 2 调度数据挖掘框架

Fig. 2 Scheduling data mining framework

2.2 调度数据预处理

连续属性离散化是数据挖掘领域研究的一个重要方面,它的要求是需要做到最大程度化不改变信息的本质意义,最小化减少信息的损失。本文运用 Chi2 算法将适当的结构化连续属性离散化,从而获得了一个通用模型。该模型可以适用于大量的离散车间调度问题。Chi2 算法用到统计量 χ^2 , 其具体步骤包括两部分:

step 1: 初始化, 根据要离散的属性对实例进行排序, 每个实例属于一个区间;

step 2: 预先设定一个阈值, 计算每一对相邻区域的统计量 x^2 , 然后将阈值之下的区间合并, 阈值之上的区间保持分区。

x^2 的计算方法为:

$$x^2 = \sum_{i=1}^m \sum_{j=1}^k \frac{(A_{ij} - E_{ij})^2}{E_{ij}} \quad (5)$$

$$E_{ij} = R_i \times C_j / N \quad (6)$$

式(5)和(6)中, k 为类别数, $m=2$, A_{ij} 为 i 区间中 j 类样本的个数, R_i 为 i 区间中样本数, C_j 为 j 类样本的个数, N 为总样本个数。统计量 x^2 反映的是节点相邻的两个区间 j 类样本的分布均匀程度。

本文针对离散车间, 在满足交货期的前提下, 以能效为优化目标, 应用数据挖掘技术, 依据以下属性, 对离散车间调度工件在机器上加工顺序进行预测。其属性包括剩余加工时间(RPT)、等待时间(WT)、交货期(DT)和工件的加工时间(TLJ)。Chi2 算法对等待时间(WT)的离散化过程如表 1~2 所示, 显著性水平为 0.9。

表 1 WT 离散化: Chi2 初始化

Tab. 1 WT discretization: Chi2 initialization

Intervals	Class frequency		x^2
1	23	0	0.6
2	31	1	2.6
3	43	7	15.7
4	18	24	1.4
5	23	25	6.8
6	3	29	3.1
7	7	19	3.51
8	2	25	0.8
9	0	7	1.7
10	3	10	-

表 2 WT 离散化结果

Tab. 2 WT discretization results

Intervals	Class frequency		x^2
1~3	93	8	32.9
4~5	41	49	23.2
6~10	15	90	-

2.3 基于 C4.5 算法的决策树数据挖掘

决策树的结构是树结构, 是一种十分常见的分类方法, 其中每个内部节点表示一个属性上的测试, 每个分支表示一个测试输出, 每个叶节点表示一种类别。如何选择最优划分属性是决策树的关键点, 随着算法过程不断进行, 决策树能够形成纯度较高的分支节点和叶节点。

决策树分类的性能很大程序受样本集好坏的影响。在传统的原始的调度数据前, 选取支持向量机对原始的数据进行分类学习, 根据分类学习所得到的判别函数对随机选取一组属性特征向量进行判断, 从而可以构成一个新的样本集, 并且和部分的原始数据集一同构成新的样本集。

本文采用基于 C4.5 算法的决策树来发现离散车间调度数据的知识。常用的属性选择标准有基尼不纯度指数(Gini Impurity Index)、基于熵的信息增益(Information Gain)和信息增益比(Information Gain Ratio)。C4.5 算法根据基于熵的信息增益比的标准选择各节点的属性。

给定训练集 D_t 中有 n 个类, 其熵的计算公式如式(7)所示。

$$E(D) = -\sum_{j=1}^n p(j|D) \log p(j|D) \quad (7)$$

式中: $p(j|t)$ 为给定 j 类在训练集 D 中的概率, 在式(7)中, 如果 $p=0$, 则定义 $0 \log 0=0$ 。通常, 式(7)中的对数以 2 或 e 为底, 这时熵的单位分别为比特或纳特。熵与 D 取值无关, 只依赖于 D 的分布。

给定父节点 x 根据属性 a 分为 k 个子节点, 对应集合为 $x_1, x_2 \dots x_k$, 则信息增益为:

$$G = Entropy(x) - [\sum_{j=1}^k \frac{x_j}{x} Entropy(j)] \quad (8)$$

在父节点求熵是以训练集 D 类别为随机变量, 下式中是以属性 a 为随机变量, 按照此属性的属性值对训练集 D 进行划分, 从而求得熵:

$$E_a(x) = -\sum_{j=1}^n \frac{n_j}{n} \log \frac{n_j}{n}$$

在每一个节点计算信息增益比, 在候选属性中

找出信息增益高于平均水平的属性,然后在这些属性中再选择信息增益率最高的属性。信息增益比定义如式(9)所示。

$$GRatio = \frac{Gain}{E_a(x)} \quad (9)$$

算法具体步骤如下:

step 1: 对数据集进行预处理,选取属性并确定属性值及其类别值,从数据集中确定包含类的训练集 D 和测试集 T ;

step 2: 将训练集对应为父节点 x ,父节点选择熵最小的属性;

step 3: 计算特征集 A 对于训练集的信息增益比,按照上文所述选择方式选择属性 a ,将 D 分为非空子集 $D_i(i=1,2,\dots,n)$,若子集只含有一个类别或信息增益比小于阈值 δ ,结束分类,得到树 T ;

step 4: 对于节点 i ,以 D_i 为训练集,以 $A-a$ 为特征集,递归调用 step 3,得到子树 T_i ;

以 6×6 的离散车间调度问题为例,从原始调度数据中选取调度数据集,通过离散化处理样本数据集中的连续数据。选取剩余加工时间(RPT)、等待时间(WT)、机器负载(ML)和工件的加工时间(TLJ)作为决策树输入属性。其数据结构表如表 3 所示。

表 3 数据结构表

Tab. 3 Data structure table

属性	RPT	WT	ML	TLJ
属性值	L 低	H 高		
类别值	0 先执行	1 次执行		

由决策树提取 “If-then” 的规则为如表 4 所示。

表 4 If-then 规则

Tab. 4 If-then rules

序号	if	then
1	WT=L,TLJ=H,	0
2	WT=L,TLJ=L,RPT=L	0
3	WT=H	1
4	WT=H,TLJ=M	1
5	WT=H,TLJ=M,RPT=M	1
6	WT=H,TLJ=L,RPT=L	0

3 知识和改进的差分进化算法融合的离散车间能效优化

3.1 基本差分进化算法

差分进化(Differential Evolution, DE)算法是一种高效的全局优化算法,其基于群体进化,进化流程类似于遗传算法,主要包含变异、交叉和选择操作。在差分进化算法迭代寻优过程中,选择变异操作是从父代个体间选择两个个体进行向量作差生成差分向量,与另一个个体求和生成实验个体;然后将父代个体与实验个体进行交叉操作,从而生成新的子代个体;最后选择操作是将父代个体和子代个体之间符合要求的个体保存到下一代群体中。

(1) 初始种群

在 DE 算法中,定义种群个体为 X ,种群个体总数为 N ,每个个体假定有 D 维,算法随机生成种群公式为:

$$x_{i,j} = l_j + r \cdot (u_j - l_j), \quad j = 1, 2, 3, \dots, D$$

式中: r 为一个随机数,均匀分布于 $[0,1]$ 区间内; u_j 和 l_j 分别为个体第 j 维变量取值的上界和下界。

(2) 变异操作

在 DE 算法中,定义种群个体为 X ,种群个体总数为 N ,每个个体假设有 m 维,第 G 代第 i 个体为 $X_{i,j}^G = (x_{i,1}^G, x_{i,2}^G, \dots, x_{i,D}^G)$ 。第 G 次迭代中,随机的从种群中选择 3 个体,分别为 $X_{r_1}^G$ 、 $X_{r_2}^G$ 和 $X_{r_3}^G$,且 $r_1 \neq r_2 \neq r_3$,则具体变异操作为:

$$V_i^{G+1} = X_{r_1}^G + F \cdot (X_{r_2}^G - X_{r_3}^G)$$

式中: F 为缩放因子,对差分向量进行缩放,取值范围为 $[0,2]$,一般取值为 0.5。

(3) 交叉操作

交叉操作是通过概率的方式随机生成新的个体,具体操作为:

$$U_i^G = \begin{cases} V_i^{G+1}, & \text{if } r \leq Cr \text{ or } j = rand(1, D) \\ X_i^G, & \text{otherwise} \end{cases}$$

式中: Cr 为交叉概率,一般取值为 $[0,1]$, r 表示在 $[0,1]$ 随机生成一个数, $rand(1,D)$ 产生一个 $[1,D]$ 之间

的整数, 确保实验个体 U_i^G 中至少有一维分量是取之变异个体 V_i^G , 从而避免与父代个体 X_i^G 相同。

(4) 选择操作

基本的 DE 算法是采用的贪婪的选择方法, 将实验个体与父代个体进行比较, 选择具有较优的目标值的个体进入下一代种群。通常目标值值越小越优, 函数关系由实际模型决定。具体的选择操作为:

$$X_i^{G+1} = \begin{cases} U_i^G, & \text{if } f(U_i^G) \leq f(X_i^G) \\ X_i^G, & \text{otherwise} \end{cases}$$

3.2 改进的差分进化算法

基本的 DE 算法本质是一种连续变量优化算法, 用于求解多维空间中整体最优解, 相关研究学者已经通过实验证明能够有效解决连续问题。离散车间能效优化问题属于离散组合优化问题, 基本的 DE 算法不能应用于该能效优化问题中, 因此本文提出的改进的差分进化算法(IDE)用于离散优化问题, 从而可以有有效的解决该优化问题。具体的改进过程主要有 3 个方面, 第 1 方面: 改进算法的编码方式; 第 2 方面: 引入时间算子 T , 改进算法的变异策略; 第 3 方面: 对变异操作和交叉操作产生的个体进行离散化。

3.2.1 编码与解码设计

在求解最优值问题过程中, 基本的 DE 算法采用的是实数编码方式, 而离散车间优化问题是在离散空间中组合优化的问题, 要想利用差分进化对离散的问题进行组合优化, 必须将其映射到离散编码空间并进行解码。

本文基于工序的方法编码。对于 $n \times m$ 的离散车间调度问题, 用长度为 $n \times m$ 的关于工件号序列表示个体, 每个个体有 n 个工件且重复 m 次, 重复的次数表示当前工件的工序号。以 3×3 调度问题为例, 如表 5 所示。

在个体编码处理之后, 需要计算和评价每个个体的目标适应度值。根据个体排列顺序确定它对应的每道工序, 确定所对应的加工机器以及加工能耗, 当个体解码结束, 所有的加工结束, 求得优化指标的值。

表 5 编码方法

Tab. 5 Encoding methods

个体	2	1	1	3	2	3	1	2	3
工序	2-1	1-1	1-2	3-1	2-2	3-2	1-3	2-3	3-3

3.2.2 改进的变异策略

基本的差分进化算法具有早熟收敛和后期收敛慢的缺点。在变异操作中, 缩放因子 F 影响算法的全局搜索能力和种群的多样性。本文引入一个时间算子 T 来控制算法在不同的迭代时期选择不同的变异策略。具体改进如式(10)所示。

$$T = e^{-\frac{100G}{G_{\max}}} \quad (10)$$

式中: G 为当前迭代代数; G_{\max} 为最大迭代代数。

本文采用的变异策略:

if ($\text{rand} < T$)

$$V_i^{G+1} = X_{r1}^G \oplus F \otimes (X_{r2}^G - X_{r3}^G) \quad (11)$$

else

$$V_i^{G+1} = X_{\text{best}} \oplus F \otimes (X_{r1}^G - X_{r2}^G) \quad (12)$$

式中: X_{r1}^G , X_{r2}^G 和 X_{r3}^G 为种群中随机的 3 个个体, X_{best} 为当前迭代最优个体。 $F \in (0, 1]$ 为缩放因子, 其表达式为:

$$t = 1 - \frac{G}{G_{\max}} \quad (13)$$

$$F = F_0 \cdot 2^t \quad (14)$$

式中: F_0 为缩放因子的初值, 取 0.5。变异操作在迭代初期, 式(11)的变异策略占主导地位并且可以维持个体的多样性, 避免算法早熟, 迭代后期, 缩放因子的值降低, 式(12)的变异策略占主导地位, 能够保留很好的信息并加快收敛速度。

3.3.2 离散差分进化操作

离散车间调度问题采用了离散编码形式来表示问题的解, 那么基本的差分进化算法中连续的差分进化操作已经不适用于求解离散车间的组合优化问题。因此, 我们必须采用适用于离散编码的差分进化操作来合理有效的解决离散车间的调度问题。

(1) 离散变异操作

个体经过变异操作得到的变异个体需要进过离散化处理, 首先采用取整法, 将变异个体的决策量的余数去除, 然后筛选重复的数和超出区间的

数，将多余重复的数和超出区间的数置零，最后将变异个体补全保证个体为合法调度。以 3×2 规模的柔性车间调度模型举例如图 3 所示。

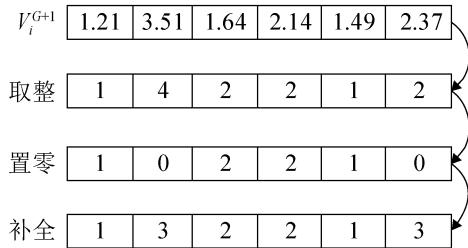


图 3 变异操作离散示意图
Fig. 3 Discrete diagram of variation operation

(2) 离散交叉操作

交叉操作的离散化过程基本类似于变异操作的离散化过程。本文采用的交叉方式具体步骤如下：

step 1: 当 $rand < Cr$ ，则将 $v_{i,j}^{G+1} \in V_i^{G+1}$ 插入到 U_i^{G+1} 的位置 j ，并令 $j = j + 1$ ；

step 2: 否则，将 $x_{i,j}^G \in X_i^G$ 插入到 U_i^{G+1} 的位置 j ，并令 $j = j + 1$ ，如果 $j \leq n$ 则返回 step 1；

step 3: 将 U_i^{G+1} 多余的位置矢量置零并补全，保证得到的试验个体为合法调度。

具体交叉操作离散化示意图如图 4 所示，假设交叉概率设定为 $Cr = 0.3$ 。

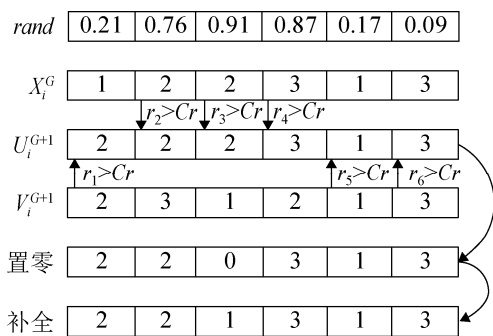


图 4 交叉操作流程示意图
Fig. 4 Cross-operation flow diagram

3.4 知识融合的差分进化算法

基于离散车间调度数据，本文对原始数据进行预处理并利用决策树挖掘调度规则，在机器设备的每个决策点，通过决策树生产的规则，如文中表 4

中的 IF-THEN 规则，判断被比较工序对应的属性值，确定工序的执行顺序，因此可以得到一组有序的工序对。调度规则知识在差分进化算法寻优过程中的运用对离散车间调度的策略选择和优化求解过程起到指导性作用。

本文算法具体步骤为：

step 1: 根据调度任务，从知识库匹配相似的调度数据构成种群 $P1$ ，随机生成种群 $P2$ ，完成种群的初始化；

step 2: 对初始种群所有个体在调度目标上的适应度值进行计算并存储最优解；

step 3: 对所有个体进行离散差分进化操作，计算进化后个体在调度目标上的适应度值，与原来最优解进行比较，更新；

step 4: 判定调度时间是否超过预期时间，若超过，引入决策树对调度顺序重新确定，返回 step 3 继续执行算法，否则执行下一步骤；

step 5: 判断终止条件，如果满足终止条件，则输出调度数据并存储在知识库中，否则，返回 step 3 继续执行算法。

算法大致的流程图如图 5 所示。

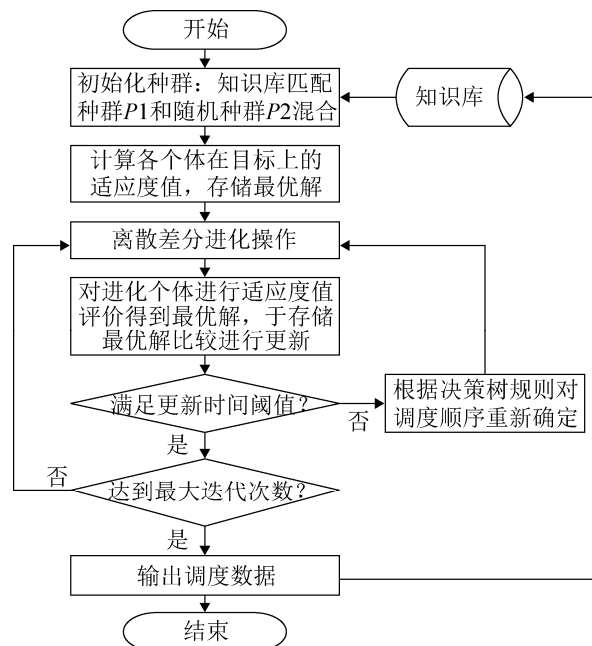


图 5 数据驱动的差分进化算法流程图
Fig. 5 Data driven differential evolution algorithm flow chart

4 实验结果与分析

本文改进的差分进化算法采用 MATLAB 程序编写, 程序运行的环境为 CPU3.20 GHz, 内存 4 GB, 软件为 MATLAB R2016b. 为了验证改进算法的可行性, 选取实际离散车间的真实数据, 选取典型的群智能算法: 教与学算法(TLBO)、遗传算法(GA)和粒子群算法(PSO), 通过实验仿真与改进的差分进化算法进行比较(IDE).

4.1 实验数据设置

本次实验采用 6×6 规模的柔性车间调度模型, 共 36 道工序, 每道工序从 8 台机器中选择能加工的机器。算法的参数设置如下: 种群规模为 20, 算法迭代代数设为 3 000, 为了提高算法的准确性, 每个算法运行 50 次, 将每次得到的机器总能耗值相加求平均值。本文的 IDE 中缩放因子 F 如上文所述, 交叉因子 $Cr=0.3$ 。另外由于各个算法的差异性, 将 GA 的参数设置为: $P_f=0.8, P_b=0.8, L_f=1$, 将 PSO 的参数设置为: $C_1=C_2=1.5, W_{max}=0.9, W_{min}=0.4$ 。

4.2 实验结果和分析

本文将 IDE 算法应用到离散制造车间的能效优化, 基于挖掘的规则对工序先确定, 结合实际生产需要从而避免生成不符合离散车间实际生产情况的方案。图 6 为知识驱动的 IDE 算法仿真出的调度甘特图, 图 7 为利用挖掘的调度规则, 通过 DE 仿真的调度甘特图, 图 6~7 通过甘特图分别直观的表现了两种算法的最优调度顺序, 我们可以看出, 知识驱动的 IDE 算法在最大完工时间上有明显的缩短, 减少了加工的等待时间。此外, 知识驱动的 IDE 算法应用在离散车间能效优化使设备利用率更高, 调度方案更优。

将 GA 算法、TLBO 算法和 PSO 算法分别应用到离散车间的能效优化中, 表 5 展示了通过 4 中算法获得的机器总能耗的最优值、平均值以及算法的运行时间。由表 5 可知, IDE 算法能够在较短

时间获得最优目标函数值, 在运行时间上虽然没 PSO 算法和 GA 算短, 但相差不到半秒, 与 TLBO 算法相比有绝对优势。另外, 分析表中数据, IDE 算法应用到离散车间, 车间能耗最低且平均值在 4 中算法中均是最小的, 其寻优能力优于其他 3 种算法。

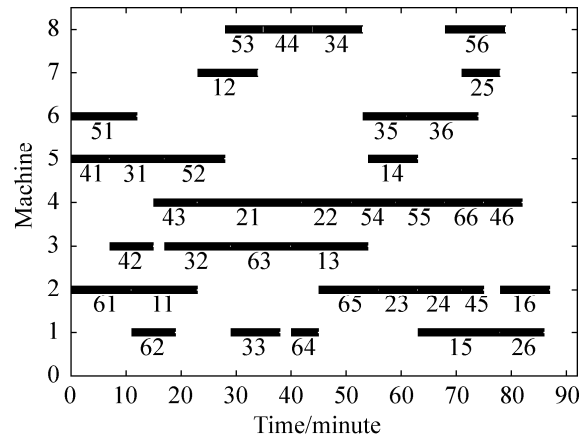


图 6 知识驱动的 IDE 算法最优调度甘特图

Fig. 6 Optimal scheduling Gantt chart of knowledge-driven IDE algorithm

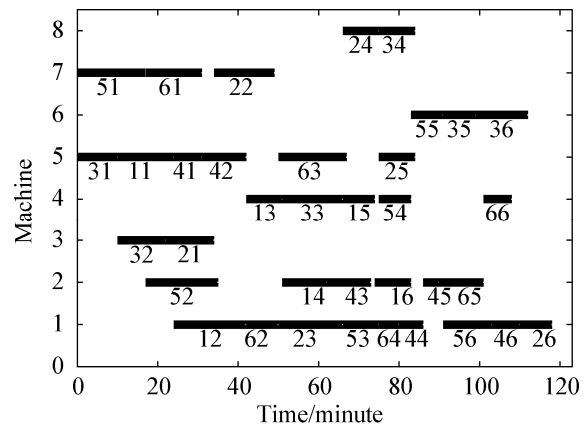


图 7 DE 算法最优调度甘特图

Fig. 7 DE algorithm optimal scheduling Gantt chart

表 5 4 种算法的运行结果比较

Tab. 5 Comparison of running results of 4 algorithms

算法	最优值/kW·h	均值/kW·h	运行时间/s
PSO	167.09	183.89	9.44
GA	166.85	175.51	9.84
TLBO	163.94	176.70	11.90
IDE	157.59	171.96	10.04

图 8 为 IDE, TLBO, GA 和 PSO 四种算法迭

代过程获得的机器总能耗平均曲线,分析图中4条曲线,IDE算法在迭代寻优过程中其收敛速度优于其他3种算法,且在整个寻优过程,其寻优结果也同样明显优于其他3种算法。

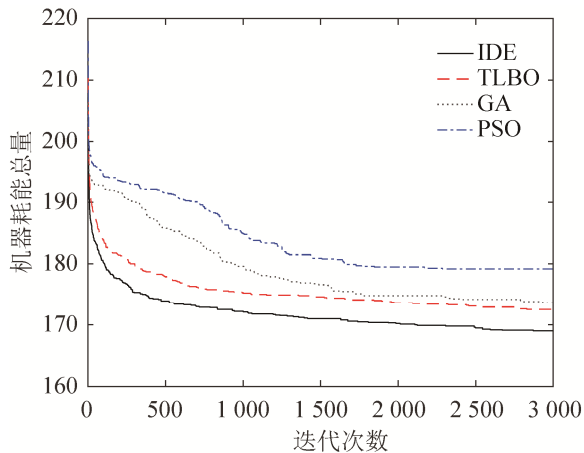


图8 机器总能耗曲线

Fig. 8 Total machine energy consumption curve

5 结论

本文针对离散车间的能效优化提出了一种调度数据环境下数据驱动的差分进化算法。在本文中,差分进化算法变异和交叉操作的离散化处理使其能够应用到离散车间的能耗优化中,并和决策树有效结合,运用调度规则驱动离散车间调度优化,进一步应用于调度问题的求解。经过 Chi2 算法对调度数据进行预处理和 C4.5 决策树对调度数据的挖掘后获得调度规则可以进一步完善调度知识库。选取 TLBO、GA 和 PSO 算法进行仿真对比,数据驱动的差分进化算法具有较好的收敛速度,寻优能力强,排产能耗更低的优点,且提出的调度知识可以指导离散车间调度过程。因此,数据驱动的差分进化算法为离散车间能效优化提供更好的调度方案,具有实际意义。

本文的算法也具有一定的局限性,适用于柔性作业车间的调度问题求解,是针对单目标能耗的问题进行求解,面对多种能耗的求解问题,需要进行进一步的研究。

参考文献:

- [1] 王凌. 车间调度及其遗传算法[M]. 北京: 清华大学出版社, 2003.
Wang Ling. Job shop scheduling and its genetic algorithm[M]. Beijing: Tsinghua University Press, 2003.
- [2] 刘博. 柔性作业车间动态生产调度问题的研究[D]. 沈阳: 东北大学, 2007.
Liu Bo. Research on dynamic production scheduling in flexible job shop[D]. Shenyang: Northeastern University, 2007.
- [3] 王景华, 韩江洪, 刘征宇, 等. 面向对象 Petri 网离散制造系统生产调度建模研究[J]. 系统仿真学报, 2008, 20(15): 4159-4162.
Wang Jinghua, Han Jianghong, Liu Zhengyu, et al. Modeling of production scheduling for a discrete manufacturing system based on object-oriented Petri net[J]. Journal of system simulation, 2008, 20(15): 4159-4162.
- [4] Boulle M. Khiops: A Statistical Discretization Method of Continuous Attributes[J]. Machine Learning (S0885-6125), 2004, 55(1): 53-69.
- [5] Wang Y, Ji D. Data- and Rule-Based Integrated Mechanism for Job Shop Scheduling[J]. International Journal of Computer and Communication Engineering (S2010-3743), 2015, 4(3): 180-186.
- [6] Mohamed A W, Sabry H Z, Khorshid M. An alternative differential evolution algorithm for global optimization[J]. Journal of Advanced Research (S2090-1232), 2012, 3(2): 149-165.
- [7] Wisittipanich W, Kachitvichyanukul V. Differential Evolution Algorithm for Job Shop Scheduling Problem[J]. International Journal of Production Research (S0020-7543), 2011, 50(10): 2757-2773.
- [8] 徐军辉, 王艳. 基于改进教与学算法的离散制造车间能效优化[J]. 系统仿真学报, 2016, 28(12): 3019-3026.
Xu Junhui, Wang Yan. Optimization of energy efficiency of discrete manufacturing workshop based on improved teaching and learning algorithm[J]. Journal of system simulation, 2016, 28(12): 3019-3026.
- [9] 解书琴. 基于多目标混合人工蜂群算法的能效优化调度研究[D]. 武汉: 华中科技大学, 2014.
Xie Shuqin. Study on energy efficiency optimization

- scheduling based on multi-objective hybrid artificial swarm algorithm[D]. Wuhan: Huazhong University of Science and Technology, 2014.
- [10] 王成龙, 李诚, 冯毅萍, 等. 作业车间调度规则的挖掘方法研究[J]. 浙江大学学报(工学版), 2015, 49(3): 421-429.
- Wang Chenglong, Li Cheng, Feng Yiping, et al. Study on the mining method of job-shop scheduling rules[J]. Journal of Zhejiang University (Engineering), 2015, 49(3): 421-429.
- [11] 冯亚. 数据挖掘中决策树分类算法研究与应用[D]. 西安: 西北大学, 2007.
- Feng Ya. Research and application of decision tree classification algorithm in data mining[D]. Xi'an: Northwestern University, 2007.
- [12] Boule M. Khiops: A Statistical Discretization Method of Continuous Attributes[J]. Machine Learning (S0885-6125), 2004, 55(1): 53-69.