流水车间调度问题的启发式算法研究

唐　聃，黄　健

(成都信息工程学院软件工程学院　成都　610225)

【摘要】针对以完工时间最小为目标的流水车间调度问题，对问题的定义进行了全新的推导和分析，从数学的角度进一步挖掘出问题的本质特征。在控制第一台机器和最后一个工件加工时间的基础上，尽量压缩每个工件在加工前的等待时间，以提高算法的实际效果。模拟实验的结果表明，新的启发式算法具有很好的性能，使用新算法计算得到的调度序列平均质量以及算法本身的稳定性方面均明显优于与之具有相当算法复杂度的其他启发式算法。

关　键　词　流水车间;　启发式算法;　完工时间;　生产调度

中图分类号　TP391.7　　文献标志码　A　　doi:10.3969/j.issn.1001-0548.2013.06.021

Research of Heuristic Algorithm for Flow Shop

TANG Dan and HUANG Jian

(Software Engineering College, Chengdu University of Information Technology　Chengdu　610225)

Abstract　For the flow shop scheduling problem which aims to minimize makespan, this paper gives a new derivation about its mathematical definition. A new heuristic method is proposed to shorten the waiting time of each job as much as possible on the basis of reducing the processing time of the first machine and last job. The result of simulation experiments shows that, the new heuristic algorithm has good performance, and the average quality and the stability of scheduling sequences generated by new method are significantly better than other heuristic algorithms with same complexity.

Key words　flow shop;　heuristic algorithm;　makespan;　production scheduling

生产调度是当前制造业企业信息化的一个研究热点，也是理论研究中最困难的问题之一，良好的调度策略将极大提高生产体系的运行效率并增加生产效益。流水车间(flow shop)调度问题是当前很多以流水线方式生产的制造业车间调度的抽象模型，也被证明是一个典型的NP完全问题[1]，因此具有很高的理论研究价值和实践价值。总完工时间(makespan)是流水车间调度问题中的一个非常重要的性能指标，总完工时间最小能使资源更加有效利用、任务更迅速传递及在制品库存最小。对流水车间调度问题的分析一般分为两种思路，求得最优解和取得次优解。求最优解一般使用的是动态规划法、分支定界法等方法，该类方法从某种意义上来说都是属于穷举法，只是在穷举的过程中根据一些计算结果排除某些明显不必要的计算。但是由于这些算法的搜索空间会随着工件数的增加呈指数式急剧增长，所以它们在求解大规模问题上有较大的局限性。正是在这种情况下，产生了以求可行解或次优解为目的的启发式算法。启发式算法是相对于最优算法提出的，可作如下定义：一个基于直观或者经验构造的算法，在可接受的花费(时间、空间等)下，给出待解决组合优化问题的每一个实例的一个可行解，该可行解与最优解的偏离程度不一定事先预计[2]。启发式算法以其计算量小、算法简单并且能得到较好的解而吸引了众多的研究者。启发式算法一般可以划分为元启发式算法[3-5]和构造式启发式算法[6-10]。通常情况下，元启发式算法获得的调度优于构造式启发式算法，但却需要更多的机器时间和空间，在满足制造业生产车间实时性的需求上有一定的困难，因此，本文的研究重点是构造式启发式算法。目前经典和常用的构造式启发式算法有Palmer[6]、Gupta[7]、CDS[8]、RA[9]和NEH[10]算法，其中又以NEH算法的性能最佳。此外，近年来也不断有新的启发式算法提出[11-13]，但多数是基于NEH的算法思想或是对其的改进。然而，NEH算法在实现过程中需要通过多次计算拟定工件序列的makespan并进行比

收稿日期： 2012 – 01 – 04；修回日期： 2012 – 05 – 28
基金项目：四川省科技支撑计划(2010GZ2314)
作者简介：唐聃(1982 –)，男，博士，主要从事算法分析方面的研究.

较，因而基于或改进的NEH算法的启发式算法均存在一个问题，即算法复杂度将远大于其他构造式启发式算法。

本文着重研究以最小makespan为目标的流水车间调度启发式算法，研究内容主要包含两部分：1) 针对问题本身的定义推导和分析，进一步挖掘问题的自身特征；2) 根据问题的自身特征和对多种已有启发式算法研究基础上提出了一种新的启发式算法，该算法具有较低的算法复杂度。

1 问题的定义和描述

制造业企业的生产调度，主要用于调配资源、合理安排作业顺序，其简化数学模型可以表述为安排n个工件在m台机器上加工顺序的问题。本文研究的以最小makespan为目标的流水车间调度问题可做如下描述。

n个工件在m台机器上加工，每个工件需要经过m道工序，每道工序要求不同的机器，n个工件通过m台机器的顺序相同，它们在每台机器上的加工顺序也相同。定义$O_{i,j}$为第j个工件在第i台机器上操作，$p_{i,j}$为$O_{i,j}$的执行时间，$c_{i,j}$表示$O_{i,j}$的完成时间，其中，$i=1,2,\cdots,m$；$j=1,2,\cdots,n$。问题的求解目标是确定n个加工任务在每台机器上的最优加工顺序，使所有加工任务全部完工的时间最短。该问题通常作如下假设：每个加工任务在机器上的加工顺序为$1,2,\cdots,m$；每台机器同时只能进行1个加工任务；1个加工任务不能同时在不同的机器上进行；各任务在加工完后立即送下一道工序；任务在机器上开始加工，必须一直进行到该工序完工，中途不允许停下来插入其他任务；所有任务在0时刻已准备就绪，机器调整时间包括在加工时间内；允许任务在工序之间等待；允许机器在任务未到达时闲置。最大完成时间makespan($C_{\max}$)也即是最后一个操作的完成时间$c_{m,n}$，而本文问题的目标是找到一个排列π，使得makespan最小。

下面将对满足上述条件的流水车间makespan作出定义推导，首先定义符号如下：

C_i是工件数量为i时的makespan，$1\leqslant i\leqslant n$，当$i=n$时，定义$C_i=C_n=C$。$t_{i,j}$为第j个工件在第i台机器上加工前的等待时间，其中$t_{i,1}=0$，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$。D_i为第n个工件在第i台机器上加工前的等待时间，其中$D_1=0$，$1\leqslant i\leqslant m$。$\boldsymbol{P}$为n个工件在m台机器上的加工时间矩阵：

$$\boldsymbol{P}=\begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,n} \end{pmatrix}$$

式中，$p_{i,j}$表示第j个工件在第i台机器上的加工时间，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$。当工件数量为1，即$n=1$时，makespan为该工件在所有机器上的加工时间之和：

$$C_1=p_{1,1}+p_{2,1}+\cdots+p_{m,1}=\sum_{i=1}^{m}p_{i,1}$$

工件数量为1时，工件在每个机器上加工前的等待时间为0。当工件数量为2，即$n=2$时，makespan为第一个工件在最后一台机器上的加工完成时间、第二个工件在最后一台机器上的加工时间与第二个工件在最后一台机器上的加工前等待时间之和：

$$C_2=C_1+p_{m,2}+D_2\text{，}\quad D_2=t_{2,n}$$

可按如下方式进行计算：

$$t_{2,2}=\max(p_{2,1}-p_{1,2},0)$$
$$t_{2,3}=\max(t_{2,2}+p_{2,2}-p'_{1,3},0)$$
$$\vdots$$
$$t_{2,n}=\max(t_{2,n-1}+p_{2,n-1}-p'_{1,n},0)$$

其中，$p'_{i,j}=p_{i,j}+\max(p_{i-1,j}-p_{j,i-1},0)$，$2\leqslant i\leqslant m$，$1\leqslant j\leqslant n$。

以此类推，对于n个工件和m台机器的流水车间，makespan的计算为：

$$C_n=C_{n-1}+p_{m,n}+D_n$$

其中，$D_n=t_{m,n}$的推导为：

$$t_{m,2}=\max(p_{m,1}-p'_{m-1,2},0)$$
$$t_{m,3}=\max(t_{m,2}+p_{m,2}-p'_{m-1,3},0)$$
$$\vdots$$
$$t_{m,n}=\max(t_{m,n-1}+p_{m,n-1}-p'_{m-1,n},0)$$

此时可以推导出：

$$\begin{gathered}\text{makespan}=C=C_n=C_{n-1}+p_{m,n}+D_n=\\ C_{n-2}+p_{m-1,n}+p_{m,n}+D_{n-1}+D_n=\\ C_{n-3}+p_{m-2,n}+p_{m-1,n}+p_{m,n}+D_{n-2}+D_{n-1}+D_n=\cdots=\\ p_{1,1}+p_{2,1}+\cdots+p_{m,1}+p_{m,2}+p_{m,3}+\cdots+\\ p_{m,n}+D_1+D_2+\cdots+D_n=\\ \sum_{i=1}^{m}p_{i,1}+\sum_{j=2}^{n}p_{m,j}+\sum_{k=1}^{n}D_k\end{gathered}\tag{1}$$

而以最小makespan为目标的n个工件和m台机器的流水车间调度问题可以归纳为需找一个工件序列，使得第n个工件在第m台机器上的完成时间尽量接近或等于

$$\min\left(\sum_{i=1}^{m}p_{i,1}+\sum_{j=2}^{n}p_{m,j}+\sum_{k=1}^{n}D_k\right)$$

2 启发式算法的提出

由式(1)可知，流水车间makespan由3部分组成：第一个工件在所有机器上的加工时间，最后一台机器上除第一个工件外所有工件的加工时间和各工件加工前的等待时间。根据这一的特点，本文基于控制第一台机器和最后一个工件加工时间的基础上，尽量压缩每个工件在加工前的等待时间的思想，提出一种新的启发式算法，描述如下。

算法中使用到的符号P、m、n和$p_{i,j}$的含义同第一节所述，其他符号定义如下：

$\boldsymbol{P}_j$为第j个工件，即矩阵$\boldsymbol{P}$的第j行($1\leqslant j\leqslant n$)，定义如下：

$$\boldsymbol{P}_j=\begin{pmatrix}p_{1,j}\\p_{2,j}\\\vdots\\p_{m,j}\end{pmatrix}$$

$\overline{S}_j$为第j个工件顺序在m个机器上加工的时间斜度($1\leqslant j\leqslant n$)，定义为：

$$\overline{S}_j=\sum_{i=1}^{m}(2i-m-1)p_{i,j} \tag{2}$$

$\mathrm{sum_}c_j$为$\boldsymbol{P}_j$中不包含最后一个元素的其他元素之和，定义为：

$$\mathrm{sum\_}c_j=\sum_{i=1}^{m-1}p_{i,j} \tag{3}$$

$\mathrm{sum_}f_j$为$\boldsymbol{P}_j$中不包含第一个元素的其他元素之和，定义为：

$$\mathrm{sum\_}f_j=\sum_{i=2}^{m}p_{i,j} \tag{4}$$

对矩阵$\boldsymbol{P}$任意两列$\boldsymbol{P}_i$和$\boldsymbol{P}_j$($1\leqslant i,j\leqslant n,i\neq j$)的前后顺序按如下方法确定：

1) 将$\boldsymbol{P}_i$和$\boldsymbol{P}_i$中的元素分别代入式(2)，计算得到$\overline{S}_i$和$\overline{S}_j$。

2) 当$\overline{S}_i*\overline{S}_j\leqslant 0$时，若$\overline{S}_i>0$，则$\boldsymbol{P}_i$置于$\boldsymbol{P}_j$之前，反之则$\boldsymbol{P}_j$应在$\boldsymbol{P}_i$之前。

3) 当$\overline{S}_i,\overline{S}_j>0$时，将$\boldsymbol{P}_i$和$\boldsymbol{P}_j$代入式(3)，计算得到$\mathrm{sum_}c_i$和$\mathrm{sum_}c_j$。若$\mathrm{sum_}c_i<\mathrm{sum_}c_j$，则$\boldsymbol{P}_i$置于$\boldsymbol{P}_j$之前；若$\mathrm{sum_}c_i>\mathrm{sum_}c_j$，则$\boldsymbol{P}_j$在$\boldsymbol{P}_i$之前；若$\mathrm{sum_}c_i=\mathrm{sum_}c_j$，则去掉$\boldsymbol{P}_i$和$\boldsymbol{P}_i$的最末一个元素后返回步骤1)。

4) 当$\overline{S}_i,\overline{S}_j<0$时，将$\boldsymbol{P}_i$和$\boldsymbol{P}_j$代入式(4)，计算得到$\mathrm{sum_}f_i$和$\mathrm{sum_}f_j$。若$\mathrm{sum_}f_i>\mathrm{sum_}f_j$，则$\boldsymbol{P}_i$在$\boldsymbol{P}_j$之前；若$\mathrm{sum_}f_i<\mathrm{sum_}f_j$，则$\boldsymbol{P}_j$在$\boldsymbol{P}_i$之前；若$\mathrm{sum_}f_i=\mathrm{sum_}f_j$，则去掉$\boldsymbol{P}_i$和$\boldsymbol{P}_j$的首个元素后返回步骤1)。

按照上述方法将矩阵$\boldsymbol{P}$中的各列进行比较，可以唯一确定出一个工件最终的排列顺序。

如有一个$F_4\,\|\,F_{\max}$问题，工件在各个机器上的加工时间矩阵为：

$$\boldsymbol{P}=\begin{pmatrix}8&5&7\\7&4&8\\1&2&5\\5&7&6\end{pmatrix}$$

按照本文的方法，排序过程如下：首先计算各个列的$\boldsymbol{P}_j$的参量$\overline{S}_j$($j=1,2,3$)，得到$\overline{S}_1=-15$，$\overline{S}_2=4$，$\overline{S}_3=-6$。比较上述3个参量，$\overline{S}_2>0$，$\overline{S}_1<\overline{S}_3<0$，可以确定$\boldsymbol{P}_1$在$\boldsymbol{P}_2$和$\boldsymbol{P}_3$之后，且$\boldsymbol{P}_2$在$\boldsymbol{P}_3$之前，至此可以得出$\pi=\boldsymbol{P}_2,\boldsymbol{P}_3,\boldsymbol{P}_1$为算法的解，按此排列计算得到其makespan为36，在该例中这也是最优解。

3 实验与分析

为了说明启发式方法的实际效果，定义最优率Y，表明在特定样本规模中，使用某种启发式算法得出最优解占所有样本数量的比例，Y越大说明该启发式算法性能越好。此外，参照文献[8]而定义出最优偏差率$E=(\mathrm{MK}-\mathrm{MKp})\times 100/\mathrm{MKp}$，其中，MK为使用启发式方法排序后工件的实际makespan，MKp为该组工件的最优makespan。最优偏差率越小则表明该启发式算法对工件顺序的排列越好，而$E=0$则说明此次对工件的排列为最优。本文将给出新算法的测试数据并对比新算法与当前各经典启发式算法。测试所用到的工件在各机器上的加工时间矩阵$\boldsymbol{P}$中的各个元素$p_{i,j}$为0到9的随机整数($1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$)，且测试用例的规模从$n=3$，$m=3$到$n=30$，$m=20$，可代表大多数实际的生产车间。表1给出了多个以样本规模20为单位的测试数据，显示了机器数量和工件数量的变化对新算法得出结果的影响。

从表1容易分析出本文的启发式算法的基本性能，而表2则对新启发式算法和当前经典的流水车间启发式算法的性能进行了对比。而从表2中不难看出，本文的方法无论是从结果的最优性还是平均最优偏差率上均明显高于与之算法复杂度相当的经典流水车间启发式算法Pamler、RA和Gupta。

表1 算法最优偏差率E和最有率Y的测试样本数据

m	n	样本规模	最优偏差率E				最优率Y/(%)	平均最优偏差率/(%)
			0%(最优)	$0\% < E \leqslant 5\%$	$5\% < E \leqslant 10\%$	$E > 10\%$		
3	3	20	19	0	1	0	95	0.33
4	3	20	18	2	0	0	90	0.29
5	3	20	17	1	1	1	85	1.25
6	3	20	16	3	1	0	80	1.19
7	3	20	14	5	1	0	70	0.99
8	3	20	14	4	2	0	70	1.15
m	n	样本规模	0%(最优)	$0\% < E \leqslant 5\%$	$5\% < E \leqslant 10\%$	$E > 10\%$	最优率Y/(%)	平均最优偏差率/(%)
3	5	20	17	0	2	1	85	1.10
4	5	20	13	3	3	1	65	2.53
5	5	20	10	7	2	1	50	2.15
6	5	20	9	5	3	3	45	3.70
7	5	20	7	6	5	2	35	4.55
8	5	20	7	8	4	1	35	3.14
m	n	样本规模	0%(最优)	$0\% < E \leqslant 5\%$	$5\% < E \leqslant 10\%$	$E > 10\%$	最优率Y/(%)	平均最优偏差率/(%)
3	7	20	14	4	2	0	70	1.37
4	7	20	9	4	5	2	45	3.99
5	7	20	5	8	6	1	25	3.71
6	7	20	3	6	8	3	15	5.41
7	7	20	3	6	6	5	15	6.66
8	7	20	1	7	9	3	5	7.23

表2 流水车间启发式算法性能对比

算法	复杂度	样本规模	性能指标/(%)	m=3 n=3	m=4 n=3	m=4 n=5	m=3 n=4	m=5 n=4	m=6 n=5	m=6 n=6
Pamler	$n\log(n)+nm$	200	最优率	74.5	66.5	33.0	53.0	42.5	23.0	16.0
			平均最优偏差率	2.13	2.11	4.28	3.66	3.37	4.52	6.28
CDS	$nm^2+nm\log(n)$	200	最优率	93.5	91.5	59.5	77.0	75.0	62.0	39.5
			平均最优偏差率	0.44	0.41	1.66	1.60	0.98	1.42	2.22
RA	$n\log(n)+nm$	200	最优率	81.5	76.0	38.5	66.0	42.0	28.0	16.2
			平均最优偏差率	1.62	1.52	3.91	2.24	3.75	4.04	5.97
Gupta	$n\log(n)+nm$	200	最优率	88.5	70.0	43.5	68.0	44.0	24.5	10.0
			平均最优偏差率	0.85	20.8	4.08	2.98	4.34	6.02	7.30
NEH	n^2m	200	最优率	99.0	97.0	84.5	92.0	88.5	77.5	67.5
			平均最优偏差率	0.03	0.13	0.45	0.32	0.39	0.64	0.93
本文算法	$n\log(n)+nm$	200	最优率	90.0	83.0	47.0	73.0	49.0	31.0	18.5
			平均最优偏差率	0.73	1.00	2.80	2.16	2.94	4.40	5.37

4 结 论

根据以最小makespan为目标的流水车间调度问题的特点，提出了一种新的启发式算法，新方法在控制第一台机器和最后一个工件加工时间的基础上，尽量压缩每个工件在加工前的等待时间，以提高算法的寻优能力。实验表明，新的启发式算法具有很好的性能，其所得调度序列的平均质量及算法稳定性明显优于与之具有相当算法复杂度的其他启发式算法。

参 考 文 献

[1] GAREY M R, JOHNSON D S, SETHI R. The complexity of flowshop and jobshop scheduling[J]. Mathematics of Operations Research, 1976, 1(2): 117-129.

[2] 邢文训, 谢金星. 现代优化计算方法[M]. 北京: 清华大学出版社, 1999.
XING Wen-xun, XIE Jin-xing. Modern optimization methods[M]. Beijing: Tsinghua University Press, 1999.

[3] WANG Ling, ZHANG Liang, ZHENG Da-zhong. An effective hybrid genetic algorithm for flow shop scheduling with limited buffers[J]. Computers and Operations Research, 2006, 33(10): 2960-2971.

[4] KUO I H, HORNG S J, KAO T W, et al. An efficient flow-shop scheduling algorithm based on a hybrid particle swarm optimization model[J]. Expert Systems with Applications, 2009, 36(3): 7027-7032.

[5] FUELLERER G, DOERNER K F, HARTL R F, et al. Ant colony optimization for the two-dimensional loading vehicle routing problem[J]. Computers & Operations Research, 2009, 36(3): 655-673.

[6] PALMER D S. Sequencing jobs through a multi-stage process in the minimum total time— a quick method of obtaining a near optimum[J]. Operational Research Quarterly, 1965(16): 101-107.

[7] GUPTA J. A functional heuristic algorithm for the flowshop scheduling problem[J]. Operational Research Quarterly, 1971(22): 39-47.

[8] CAMPBELL H G, DUDEK R A, SMITH M L. A heuristic algorithm for the n-job, m-machine scheduling problem[J].

[9] DANNENBRING D G. An evaluation of flow shop sequencing heuristics[J]. Management Science, 1977, 23(11): 1174- 1182.

[10] NAWAZ M, ENSCORE E, HAM I. A heuristic algorithm for the m machine, n job flow shop[J]. OMEGA: the International Journal of Management Sciences, 1983, 11(1): 91-95.

[11] 盛立纲, 顾幸生. 混合遗传NEH算法在流水车间调度中的应用[J]. 控制工程, 2010, 17(4): 497-500.
SHENG Li-gang, GU Xing-sheng. Hybrid genetic NEH algorithm for permutation flowshop scheduling problems [J]. Control Engineering of China, 2010, 17(4): 497-500.

[12] 高守玮, 戴杨, 刘媛媛. 对于NEH启发式方法搜索邻域的研究[J]. 控制工程, 2008, 15(2): 217-219.
GAO Shou-wei, DAI Yang, LIU Yuan-yuan. Study on the neighborhood of NEH heuristic[J]. Control Engineering of China, 2008, 15(2): 217-219.

[13] GAO Shou-wei, LEISTEN R, DAI Yang, et al. Break the ties for NEH heuristic in solving the permutation flow shop problems[J]. Journal of Donghua University(English Edition), 2008, 25(3): 258-262.

编 辑 漆 蓉

(上接第899页)

[3] KARAM M A, AMAR F, FUNG A K, et al. A microwave polarimetric scattering model for forest canopies based on vector radiative transfer theory[J]. Remote Sensing of Environment, 1995, 53(1): 16-30.

[4] MCDONALD K C, DOBSON M C, ULABY F T. Using MIMICS to model L-band multiangle and multitemporal backscatter from a walnut orchard[J]. IEEE Trans Geosci Remote, 1990, 48(4): 477-491.

[5] DE ROO R D, DU Yang, ULABY F T, et al. A semi-empirical backscattering model at L-band and C-band for a soybean canopy with soil moisture inversion[J]. IEEE Trans Geosci Remote Sensing, 2001, 39(4): 864-872.

[6] YUEH S H, KONG J A, JAO J K, et al. Branching model for vegetation[J]. IEEE Transactions on Geoscience and Remote Sensing, 1992, 30(2): 390-402.

[7] TSANG L, MANDT C, DING K H. Monte Carlo simulations of the extinction rate of dense media with randomly distributed dielectric spheres based on solution of Maxwell's equations[J]. Opt Lett, 1992, 17(5): 314-316.

[8] ZURK L M, TSANG L, DING K H, et al. Monte Carlo simulations of the extinction rate of densely packed spheres with clustered and non-clustered geometries[J]. Journal of the Optical Society of America A (Optics, Image Science and Vision), 1995(12): 1772-1781.

[9] TOAN T L, RIBBES F, WANG L F, et al. Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results[J]. IEEE Trans Geosci Remote Sensing, 1997, 35(1): 41-56.

[10] OH Y, HONG J Y. Moment method - Monte Carlo simulation of the microwave backscatter of wet-land rice fields[C]//IEEE IGARSS 2007. Busan, Korea: IEEE, 2007: 69-72.

[11] KARAM M A, FUNG A K. Leaf-shape effects in electromagnetic wave scattering from vegetation[J]. IEEE Trans Geosci Remote Sensing, 1989, 27(6): 687497.

[12] JIA Ming-quan, CHEN Yan, TONG Ling, et al. Land-based scatterometer measurements and retrieval of surface parameters using neural networks[C]//IITA Conference on Geoscience and Remote Sensing (IITA-GRS 2008) 2008. Shanghai, China: [s.n.], 2008

编 辑 漆 蓉