

并行机调度遗传算法研究

何敬东¹, 黄向², 严太山²

(1.湖南汨罗新市中学, 湖南 汨罗 414400; 2.湖南理工学院计算机系, 湖南 岳阳 414000)

【摘要】为解决接近实际生产状况的并行机调度问题, 将工件延迟与提前完成成本、机器闲置时间成本、机器启动成本列入考虑因素。在遗传算法的设计上, 提出了一套染色体的组成方式, 设计了适应度函数和进化机制, 使其能较好地表示问题的解空间并提升求解效率, 具有一定的实用价值。

【关键词】并行机调度; 遗传算法; 适应度

【中图分类号】TP 273

【文献标识码】A

【文章编号】1008-1151(2008)06-0054-02

(一) 前言

并行多机调度问题是 NP-Hard 问题, 它既有丰富的研究内容, 同时在生产调度, 机械制造等方面有广泛的实际应用价值。但是在实际操作中, 人们常常按经验法则(如处理时间短优先调度或到期日早优先调度)进行调度, 这样得到的调度结果一般不能满足实际环境的需求, 而且当并行机数和工件数增加时, 问题的复杂度成指数增加, 传统方法不适合解决此类问题。目前的并行机调度问题研究大部分是集中于简化问题, 然后寻找最优解或次优解。

如何使并行机调度问题更接近实际生产状况并找到次优解成为本问题研究热点之一。遗传算法由于其固有的全局搜索与收敛特性, 能够较好地解决此类优化组合问题。用遗传算法解决并行机调度问题已经有相关研究, 其中 Cheng 提出一套方案, 利用特殊符号(*)和工件号组成染色体, 特殊符号为分隔符, 将工件分成不同的组, 每个工件组指派到不同的并行机中。本文对 Cheng 的方案进行了改进, 并且把工件延迟与提前完成成本、机器闲置时间成本和机器设备启动成本等列入考虑因素, 使多机调度问题的解决更接近实际的生产状况, 目的在于求得考虑实际因素下的次优调度结果, 以降低存货存储、交货延迟的违约、设备及人力等成本。

(二) 遗传算法与并行机调度问题

1. 遗传算法

遗传算法是一种模仿自然界中优胜劣汰自然选择规则的随机搜索方法, 该算法是从一系列的初始点(被称为初始种群)开始, 通过循环执行选择、交叉和变异等遗传操作, 逐步得到问题接近全局的最优解。

2. 并行机调度问题

考虑并行机调度中的实际需求定义如下并行机调度问题: 设有 $m(m>1)$ 台并行机 M_1, M_2, \dots, M_m , 需要加工 n 个工件, F 时调度工作结束, 每个工件订单零时刻到达, 可在任一机器上完成加工。工件一旦分配给某机器加工就必须加工完毕,

不许中断。工件 i 在机器 j 上加工前准备时间为 Z_{ij} , 加工时间为 T_{ij} , 工件 i 的到期时间为 Q_i , 机器 j 一天的工作时间为 W_j , 启动成本为 S_j , 闲置成本权重比例为 R_j , 则工件 i 的完成日期 C_i 为本身加工前准备时间 Z_{ij} (j 为工件 i 的加工机器)、加工时间 T_{ij} 与机器 j 中 i 工件之前所有工件消耗时间之和, 提前完成时间 B_i 为 $\max(0, Q_i - C_i)$, 延迟时间 D_i 为 $\max(0, C_i - Q_i)$, 机器 j 的闲置时间为 I_j (F 减 j 上最后一个工件完工时间)。问如何安排每台机器上所加工的工件及加工顺序使成本最小。

(三) 算法设计

在 Cheng 的方案基础上, 提出了一个染色体的改进的编码组成方式, 重新设计了适应度函数和进化机制。

1. 染色体编码与适应度函数设计

并行机调度问题的染色体必须表示两个方面的信息: 一方面确定工件加工所在的机器, 另一方面确定每台机器上工件的加工顺序。本算法中染色体编码时用自然数序列表示工件排列, 用含数字下标的分隔符*_i ($0 < i \leq m$) 表示不同机器, 并把整个排列看成一个循环。例如把 8 个工件分配到 4 台并行机的某染色体为 (1 *₂ 4 5 *₁ *₃ 2 3 6 7 *₄ 8), 其意义是: 最后一个分隔符之后与第一个分隔符之前的工件 8, 1 按顺序分配给机器 2, 工件 4, 5 分配给机器 1, 机器 3 没有分配工件, 不用启动, 工件 2, 3, 6, 7 分配给机器 4。这种染色体编码方式不仅明确表明了每个工件所在的机器号, 而且包含了每个工件在每台机器上的加工顺序。染色体意义简单明了。

遗传算法演化过程的目的在于使适应度函数最小化, 如果用 B, D, S, I 分别表示提前、延迟、重启和闲置成本, 则演化的目的是要找到最好的排列使得:

$F(B^*, D^*, S^*, I^*) = \min \{F(B, D, S, I)\}$, 适应度函数为 F ,

$$F(B, D, S, I) = \sum_i [\alpha(B_i) + \beta(D_i)] + \sum_j (\gamma_j * S_j + R_j * I_j)$$

B_i, D_i, S_j, R_j, I_j 意义同上, $\alpha(B_i), \beta(D_i)$ 分别为提前和延迟完

【收稿日期】2008-04-13

【作者简介】何敬东(1978-), 女, 湖南汨罗人, 汨罗新市中学教师, 研究方向为遗传算法、现代教育技术; 黄向(1977-), 男, 湖南汨罗人, 湖南理工学院计算机系讲师, 硕士, 研究方向为分布式网络、软件理论; 严太山(1968-), 男, 湖南祁东人, 湖南理工学院计算机系讲师, 博士研究生, 研究方向为进化算法、神经网络。

工成本函数， y_j 的值在机器 j 上有工件分配时等于 1，否则为 0。

2. 种群初始化

初始种群的个体分为两类，第一类个体随机生成；另一类个体可以按一定经验法则生成，比如可以把处理时间作为考虑因素，处理时间越短越先安排，确保在最短的时间里做尽可能多的工件，或者按到期日期为标准，越早到期的工件越早安排，可以使大部分工件如期完成。

3. 选择

选择过程的目的是为了从当前群体中选出优良个体，使他们有机会作为父代将其优良的个体信息传递给下一代，使子代向最优解逼近^[1]。判断个体优良与否的标准是各自的适应值。每个个体的选择概率采用基于排序的适应度分配(rank-based fitness assignment)方法，种群按适应度进行排序，个体的生存概率使用 Baker 提出的线性排序计算公式

求得：
$$p_i = \frac{1}{N} [\eta^+ - (\eta^+ - \eta^-) \frac{i-1}{N-1}]$$
，其中 i 为个体排序序号， N 为种群大小， $1 < \eta^+ < 2$ ， $\eta^- = 2 - \eta^+$ 。然后用轮盘赌选择法(roulette wheel selection)来决定每个个体的选择份数，选择后的 N 个个体送到配对库，以备配对繁殖。

4. 交叉

交叉是结合来自父代种群的信息产生新的个体，依据个体基因编码表示方法的不同有多种交叉算法。这里采用两点顺序交叉(OX)，顺序交叉能够保留排列并融合不同排列的有序结构单元。如有下面两个父个体，交叉操作步骤如下：

p1: (1 *₂ 4 5 *₁ *₃ 2 3 6 7 *₄ 8)

p2: (4 1 *₂ 5 *₁ 3 6 *₃ 2 7 *₄ 8)

(1) 随机选择两个交叉点“|”，保留交叉点之间的中间段不变。

q1: (x x | 4 5 *₁ *₃ 2 | x x x x x)

q2: (x x | *₂ 5 *₁ 3 6 | x x x x x)

(2) 把父个体 2 从第 2 个交叉点作排列，到达表尾时返回表头继续，得到排列*₃ 2 7 *₄ 8 4 1 *₂ 5 *₁ 3 6，减去父个体 1 中已有基因部分 4 5 *₁ *₃ 2 得到排列顺序 7 *₄ 8 1 *₂ 3 6，再将这个顺序从第 2 个交叉点开始复制给子个体 1，以此决定子个体 1 对应位置的未知码 x ，生成子个体 1 即 q1: (3 6 4 5 *₁ *₃ 2 7 *₄ 8 1 *₂)，同样产生子个体 q2: (*₃ 2 *₂ 5 *₁ 3 6 7 *₄ 8 1 4)。

5. 变异

变异是子代基因按小概率产生的变化。同样可以采用多种方法实现基因变异，如可以采用交换两个基因位置实现变异，但是这种方法个体变异较小，不利于新个体的生成，特别是当交换的两个基因都是数字时，新个体只是改变了两个工件的加工位置。这里使用另外一种方法实现变异操作即首先随机选取父代中的某优秀染色体，对照染色体基因个数($m+n$)随机生成一个二进制序列，将染色体中对应 0 的基因选为一组，对应 1 的基因选为另一组，两组基因重新组合得到子个体。变异操作如下：

染色体 1 *₂ 4 5 *₁ *₃ 2 3 6 7 *₄ 8

二进制序列 0 1 0 1 0 0 1 1 1 0 1 1

0 对应基因 1 4 *₁ *₃ 7

1 对应基因 *₂ 5 2 3 6 *₄ 8

子个体 1 4 *₁ *₃ 7 *₂ 5 2 3 6 *₄ 8

6. 种群进化与停止准则

标准遗传算法若在群体选择前(或后)保留当前最优值，则最终能收敛到全局最优值。本算法通过改进标准遗传算法，在群体选择作用前保留当前最优解，则保证本算法在某些情况下收敛到全局最优解。

停止准则一般预先设定一个最大的繁殖代数 N_{max} ，在繁殖过程中进行到最大的繁殖代数则停止运行。

综上所述，本并行机行排序问题的遗传算法伪码如下：

```
begin
k:=0;
初始化 p(k); //p(k) 为第 k 代群体;
求 p(k) 中个体的适应值并排序;
while (k<Nmax) do
begin
k:=k+1;
从 p(k-1) 中选择 p(k); //选择
对 p(k) 中的个体依概率进行交叉运算; //交叉
对 p(k) 中的个体依概率进行变异运算; //变异
求 p(k) 中个体的适应值并排序;
保留当前群体最优解;
end
end
```

(四) 结果

设有 45 个工件要在 3 台机器上加工。在工作提早与延迟完成成本方面，按照工件的重要性将提前与延迟完成成本公式分为三类：

对重要性不高的工件成本公式为：

$\alpha(B_i)=2$, $\beta(D_i)=3$

对重要性一般的工件成本公式为：

$\alpha(B_i)=2B_i+2$, $\beta(D_i)=3D_i+3$

对优先权高的工作成本公式为：

$\alpha(B_i)=2B_i^2+2B_i+2$, $\beta(D_i)=3D_i^2+3D_i+3$

工件在不同机器上的加工时间，加工前准备时间，机器启动和闲置成本表略。种群大小为 50，最大遗传代数为 80，交叉概率为 0.6，变异概率为 0.01。按本文与 Cheng 所提算法分别演算，所得结果如表 1 所示：

表 1 实验结果对比

	本文算法	Cheng 算法
标准差/平均适应值	2.16%	5.32%
最佳适应值	215	216
最差适应值	224	232

由表 1 可知，本文算法有助于使最后解答的染色体之间的差异性小，每次运算后的解接近最优解。(下转第 52 页)



图1 原始图像

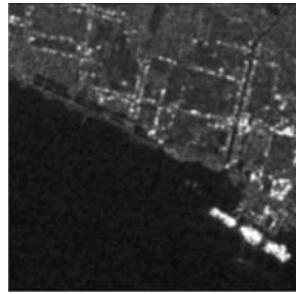


图2 滤波后的图像



图3



图4

图5是用LOG算子对滤波后的图像进行边缘检测,再用得到的边缘和细节信息与图2进行加权融合得到的图像;图6是用LOG算子对原始图像进行边缘检测后,再用得到的边缘和细节信息与图2进行加权融合得到的图像。



图5



图6

表1

	图1	图2	图3	图4	图5	图6
means	51.318	50.315	50.564	52.655	51.308	52.570
varian	40.921	35.794	36.921	44.033	39.310	43.097
MSE		378.69	543.52	777.31	652.11	790.43

从以上两个例子上可以直观的看出,对原始图像进行边缘检测得到的有用信息,比对已经过初步滤波的图像进行边缘检测得到的有用信息要多,而且最终加权融合的图像效果也更好。但前提必须是,原始图像中的斑点噪声不能完全覆盖细节信息。这不只是从定性上来看,从定量上也可以分析出来。如表1所示,与图1(即原始图像)的均值对比,各种方法得到的均值都差不多,但是方差就有较大的差别。图4和图6是在原始图像上直接边缘检测后进行的图像融合,因为没有先平滑图像,所以得到的方差就比图1得到的方差大,效果也不理想。而图2是直接平滑,图3和图5是在平滑后进行边缘检测和图像融合的,得到的方差小,效果也稍好。对均方误差(MSE)来说,图2是最小的,因为只对原始图像进行了平滑,没有进行边缘检测后的融合,而图3和图5是在图2的基础上利用检测到的边缘信息再进行融合,虽然均方误差变大了,但重要的细节信息还是被突出了。图4和图6是在图1的基础上利用检测到的边缘信息再进行融合,均方误差数值较大,反映到图像中就是部分细节信息被“放大”了。由此可见,在原始图像中进行边缘检测比平滑后再进行边缘检测可以得到更为详细的细节信息。

本文利用传统的空域滤波方法,并采取合理的方式进行边缘检测,进行图像的加权数据融合,从而得到质量改善的SAR图像。这种方法可以不必用到图像的统计特性,从而省去了大量的计算,可以满足一般用户的需求。

【参考文献】

- [1] 魏钟铨,等.合成孔径雷达卫星[M].科学出版社,2001.
- [2] 杜培军.RADARSAT 图像滤波的研究[J].中国矿业大学学报,2002.
- [3] 李春升,燕英,陈杰,等.高分辨率星载 SAR 单视图像斑点噪声抑制实现方法[J].电子学报,2000,28(3):13-16.

(上接第55页)

(五) 结束语

本文重新设计了染色体的组成方式,问题的解空间表示得更明确,设计了算法进化机制,求解效率得到提升,考虑了并行机调度问题中的更多实际因素,本算法有一定的实用价值。然而实际影响并行机调度的因素并不限于本文所讨论的范围,另外适应度函数中的权重比例,也需要根据不同工厂的需求做适当的调整。今后将研究影响并行机调度的其它一些因素,使并行机调度问题与实际应用紧密结合,并且考虑把遗传算法与其它算法结合,使遗传算法收敛更快更高效。

【参考文献】

- [1] 何军辉,周泓.求解含调整时间并行机排序问题的遗传算法[J].系统工程理论方法应用,2002,12.
- [2] 欧锦文,施保昌.并行机排序邻域搜索算法设计[J].计算机工程与应用,2003,18.
- [3] Cheng,Runwei . Minmax Earliness/Tardiness Scheduling in Identical Parallel Machine System Using Genetic Algorithms[J].International Conference on computers and industrial Engineering,1995:29:513-517.
- [4] 张聚,李平.基于演化算法的车间作业调度问题的求解方法[J].浙江大学学报,2004,12.
- [5] 王小平,曹立明.遗传算法[M].西安:西安交通大学出版社,2002.