

短文

解决并行多机提前/拖后调度问题的 混合遗传算法方法¹⁾

刘 民 吴 澄

(清华大学自动化系 北京 100084)

摘 要 研究了带有公共交货期的并行多机提前/拖后调度问题. 提出了一种混合遗传算法方法, 以便于确定公共交货期和每台机器上加工的任务代号及其加工顺序, 即找到一个最优公共交货期和最优调度, 使加工完所有任务后交货期安排的成本、提前交货成本和拖后交货成本的总和最小. 数值计算结果表明了该混合遗传算法优于启发式算法, 并能适用于较大规模并行多机提前/拖后调度问题. 算法计算量小, 鲁棒性强.

关键词 并行多机, 遗传算法, 提前/拖后, 交货期安排, 调度问题.

HYBRID GENETIC ALGORITHM METHOD FOR IDENTICAL PARALLEL MACHINE EARLINESS/TARDINESS SCHEDULING PROBLEM

LIU Min WU Cheng

(Department of Automation, Tsinghua University, Beijing 100084)

Abstract Identical parallel machine earliness/tardiness scheduling problem with common due-date is studied. A kind of hybrid genetic algorithm is presented to determine the common due-date, job code, and processing orders of every processed job on each machine, that is, find an optimal common due-date and optimal scheduling policy for minimizing the total cost of assignment of due-date, early delivery and tardy delivery. Numerical computations show that the hybrid genetic algorithm has advantage over heuristic procedure, and is fit for larger scale identical parallel machine earliness/tardiness scheduling problem. The algorithm proposed has low computational complexity and strong robustness.

Key words Identical parallel machine, genetic algorithm, earliness/tardiness, due-date assignment, scheduling problem.

1) 国家八六三高技术计划基金(863-511-9600-008)和教育部博士点基金(9500322)资助项目.

收稿日期 1998-04-28 收到修改稿日期 1999-03-16

1 引言

随着 JIT(Just In Time)在日本获得成功,以准时生产为目标的提前/拖期调度问题已成为一个活跃的研究领域^[1].在带有公共交货期约束的提前/拖后调度问题中,若公共交货期未知,则需要同时确定最优公共交货期和最优调度方案,关于这类问题的文献报道不多,且大多针对单机情况,主要是一些启发式算法^[2,3],但启发式算法随着问题规模的扩大,求得解的质量往往不高.本文提出一种解决具有公共交货期的并行多机提前/拖后调度问题的混合遗传算法,并用实例进行了数值计算,计算结果表明该方法优于启发式算法,且运算时间短、鲁棒性强.

2 问题的描述

对具有待定的公共交货期的并行多机提前/拖后调度问题可描述如下:有 n 个相互独立的任务, m 台完全相同的机器,每个任务都有确定的加工时间和公共交货期,且均可由 m 台机器中的任一机完成加工任务.要找一个最优公共交货期和最优调度,即确定公共交货期和每台机器上加工的任务代号及其加工顺序,使加工完所有任务后交货期安排的成本、提前交货成本和拖期交货成本的总和最小.

设 m_j 为任一机器 j 上加工的任务总数, $k(j, i)$ 为机器 j 上第 i 次加工的任务代号,对任意的 $j \in [1, m]$, 有 $i \in [1, m_j]$, $k(j, i) \in [1, n]$. 设 $T(k(j, i))$ 为任务 $k(j, i)$ 的加工时间, d 为任务的公共交货期,则任务 $k(j, i)$ 的完工时间可表示为 $\sum_{i_0=1}^i T(k(j, i_0))$. 对任何一个任务 $k(j, i)$, 拖后时间为 $T_1(k(j, i)) = \sum_{i_0=1}^i T(k(j, i_0)) - d > 0$, 提前时间为 $E_1(k(j, i)) = d - \sum_{i_0=1}^i T(k(j, i_0)) > 0$.

调度目标是分配 n 个任务给 m 台机器,使加工完所有任务后由下式给出的惩罚函数 $f(\delta, d)$ 为最小,式中 δ 对应一个调度方案, α, β, γ 为非负常数,它分别表示交货期安排、提前交货和拖后交货的单位成本. 惩罚函数为

$$f(\delta, d) = \sum_{j=1}^m \sum_{i=1}^{m_j} \{ \alpha d + \beta E_1(k(j, i)) + \gamma T_1(k(j, i)) \} = \\ n \alpha d + \sum_{j=1}^m \sum_{i=1}^{m_j} \{ \beta E_1(k(j, i)) + \gamma T_1(k(j, i)) \}.$$

3 算法

构造适用于并行多机提前/拖后调度问题的混合遗传算法

1) 编码设计

本文中的染色体由三个基因串组成,其中,每台机器上加工的任务代号依次排列作为

第一个基因串, 每台机器上加工的任务总数依次排列作为第二个基因串, 第三个基因串是用二进制表示的公共交货期. 即染色体编码为 $k_{11}, k_{12}, \dots, k_{1m_1}, k_{21}, k_{22}, \dots, k_{2m_2}, \dots, k_{m_1}, k_{m_2}, \dots, k_{mm_m} | m_1, m_2, \dots, m_j, \dots, m_m | l_1, l_2, \dots, l_p, \dots, l_r$, k_{ji} 为机器 j 上第 i 次加工的任务代号, m_j 为在机器 j 上加工的任务总数, 其中 $j \in [1, m], i \in [1, m_j], k_{ji} \in [1, n], k_{ji}$ 取互不相同的自然数, $l_1, l_2, \dots, l_p, \dots, l_r$ 为公共交货期 d 的二进制代码, 即 $l_p \in \{0, 1\}, r$ 为正整数, 可根据需要选定. 由于所有 n 个任务必须在 m 台机器上加工完, 故有 $\sum_{j=1}^m m_j = n$. 用本文的编码方法, 染色体的总位数为 $n+m+r$.

2) 产生初始种群

随机产生 N 个 $n+m+r$ 位的个体, 即产生 N 个编码为 $k_{11}, k_{12}, \dots, k_{1m_1}, k_{21}, k_{22}, \dots, k_{2m_2}, \dots, k_{m_1}, k_{m_2}, \dots, k_{mm_m} | m_1, m_2, \dots, m_j, \dots, m_m | l_1, l_2, \dots, l_p, \dots, l_r$ 的初始种群. 显然用随机的方法产生初始种群时, 很难满足个体的第二个基因串的约束条件 $\sum_{j=1}^m m_j = n$, 为此本文提出一种产生个体的初始可行解第二个基因串的方法.

首先对每一个个体的第二个基因串, 按保持负载均衡的思想, 取 $m_j = \text{int}\left(\frac{n}{m}\right), j \in [1, m-1], m_m = n - (m-1)\text{int}\left(\frac{n}{m}\right)$, int 表示取整运算. 其次, 在每个个体的第二个基因串中随机挑选两个不同的基因座, 并对这两个基因座的基因值一减一增相同的随机正整数, 以保持两个基因值的和不变, 当变换后的两个基因值之一将小于1时, 不进行此次变换, 这样就能保证随机变换后, 仍能满足 m_j 是正整数, 且 $\sum_{j=1}^m m_j = n$, 用此方法变换适当次数, 就能得到不同个体第二个基因串.

由于产生每个个体第二个基因串采用的变换次数也是随机的, 所以个体的多样性就可得到保证. 用本文方法产生的个体的第二个基因串不仅能自动满足约束条件, 从而大大减小了产生初始可行解的计算量, 同时有利于保持机器负载均衡, 产生较优的初始可行解.

3) 适应度函数和选择策略

本文问题的目标函数为

$$f = nad + \sum_{j=1}^m \sum_{i=1}^{m_j} \{\beta E_1(k(j,i)) + \gamma T_1(k(j,i))\}.$$

由于本文是求取问题的最小值, 故对目标函数采用指数定标方式来得到适应度函数, 即取适应度函数为 $f = ae^{-bf}$, 其中 a, b 为正实数.

选择策略采用“轮盘赌”式的正比选择法^[4].

4) 交叉

考虑到每个个体的三个基因串分别代表不同的含义, 为了使交叉时满足调度时的约束要求, 本来可采用分三段独立交叉. 但对第二个基因串进行交叉将很难满足约束条件 $\sum_{j=1}^m m_j = n$, 即很难得到可行解. 为了提高交叉的效率, 本文仅对第一个、第三个基因串进行交叉操作, 对第二个基因串则不进行交叉操作. 在配对的两个个体的第一个基因串 A_i 和 B_i 之间进行交叉时, 由于要求交叉后形成的每个新基因串中, 每个基因的值互不相

同,所以不能采用普通的一点或多点交叉方法.本文采用文献[5]中的交叉方法来完成第一个基因串 A_i 和 B_i 之间的交叉.

对第三个基因串,由于编码为0—1编码,故可采用普通的一点交叉方式即可.

5) 变异

为了使变异时,满足调度时的约束要求,本节采用分三段独立变异.

对于染色体 $A_i|B_i|C_i$,变异过程可用下式表示:

$$A_i \uparrow \Rightarrow A'_i, \quad B_i \uparrow \Rightarrow B'_i, \quad C_i \uparrow \Rightarrow C'_i.$$

变异后的染色体为: $A'_i|B'_i|C'_i$,其中符号 \uparrow 表示进行变异操作.

a. 由于在对基因串 A_i 进行变异时,要求变异后得到的新基因串中每个基因的值互不相同,所以不能采用普通的变值变异方法,本文采用如下方法来完成基因串 A_i 的变异.

在对基因串 $A_i=123\ 456\ 789$ 变异时,选取随机数 $r_2=1$ 或 $r_2=2$.

若 $r_2=1$,则采用交换变异方式,即在第一个基因串中随机选取两个基因座,交换其值.

$$A_i = 123456789, \quad A'_i = 127456389;$$

若 $r_2=2$,则采用逆转变异方式,即在第一个基因串中随机选取两个基因座,把两个基因座之间的基因值逆转

$$A_i = 12\ 3456\ 789, \quad A'_i = 12\ 6543789.$$

b. 由于对第二个基因串 B_i 不采用交叉操作,所以对基因串 B_i 设计合适的变异算子显得尤为重要.又由于所用编码方式使基因串 B_i 在变异时须满足 m_j 是正整数,且 $\sum_{j=1}^m m_j = n$,这是用一般的变值变异方法很难满足的,与本文在产生初始种群中所述的理由相同,为了提高变异的效率,对 B_i 采用一种适用本文问题的特殊的变异方法,即在 B_i 中随机挑选两个不同的基因座,并对这两个基因座的基因值一减一增相同的随机正整数,以保持两个基因值的和不变,当变换后的两个基因值之一将小于1时,不进行此次变换,这样就能保证随机变换后,仍能满足 m_j 是正整数,且 $\sum_{j=1}^m m_j = n$,用此方法变换适当次数,就能得到变异后的新基因串 B'_i .

c. 对个体的第三个基因串 C_i 进行变异时,由于采用0—1编码,故可直接用普通的变值变异方法,即对 C_i 中的每一位基因,以概率 P_m 进行0—1变换.

6) “进化交叉”操作

由于遗传算法的最大优势在于全局搜索能力,而局部搜索能力较差.考虑到本文问题的复杂性和搜索空间之大,为了提高遗传算法的搜索效率,本文通过引入“进化交叉”操作算子,将启发式方法混合到遗传算法中,以提高算法的局部搜索能力,构成全局搜索和局部搜索相结合的优化搜索算法,即混合遗传算法.“进化交叉”操作的具体步骤如下:

首先用启发式算法^[2]求得问题的一个可行解,并把它用本文的编码方式表示成染色体串的形式,在每一次交叉、变异操作之后,将交叉、变异得到的个体与用启发式算法得到的可行解对应的个体以概率 P_{ec} 进行交叉,其中 P_{ec} 为“进化交叉”概率.

4 数值计算及分析

我们在586微机上用实例对以上算法(简称 HGA)进行了数值计算.并与文献[2]提出

的启发式算法求得的解进行了比较. 对每一组 m, n 的组合, 加工时间用随机方法产生. 对每种规模的问题, 表1中数据为程序运行20次的结果.

表1 数值计算结果

	问题规模 $n \times m$	启发式方法		HGA			
		最优值 f_{best}	公共交货期 d	最优值 f_{best}	平均值 f_{arg}	运行时间	公共交货期 d
1	8×3	113	6	106	107.4	4s	8
2	20×6	1 218	25	1 170	1 182	7s	28
3	30×8	1 762	22	1 697	1 703	9s	22
4	40×10	2 865	30	2 775	2 775.8	10s	30

从数值计算结果可看出, 即使对 8×3 的小规模问题, 启发式算法也不能得到最优解 (该问题最优解为106), 但 HGA 能得到该问题的最优解. 而对较大规模调度问题, HGA 算法得到的解的性能明显优于启发式算法, 所需的计算时间也很短. 另外, HGA 能得到满足要求的多组解. 算法对任务加工时间的变化也不敏感, 即加工时间有微小的变动, 也能得到满足要求的解.

本文为遗传算法在具有公共交货期的并行多机提前/拖后调度问题上的应用作了深入研究, 提出的混合遗传算法有很好的性能, 且适应于较大规模的并行多机提前/拖后调度问题.

参 考 文 献

- 1 Baker K R, Scudder G D. Sequencing with earliness and tardiness penalties: a review. *Operations Research*, 1990, **38**(1): 22~35.
- 2 Cheng T C E. A heuristic for common due-date assignment and job scheduling on parallel machines. *J. Opt. Res. Soc.*, 1989, **40**(12): 1129~1135.
- 3 Glover F. Future paths for integer programming and links to artificial intelligence. *Comput & Ops Res.*, 1986, **13**(5): 533~549.
- 4 刘民, 吴澄, 蒋新松. 用遗传算法解决并行多机调度问题. *系统工程理论与实践*, 1998, **18**(1): 14~17.
- 5 LIU Min, WU Cheng, JIANG Xinsong. Genetic algorithm method for minimizing the number of tardy jobs in identical parallel machine scheduling problem. *Chinese Journal of Electronics*, 1998, **7**(2): 188~192.

刘 民 男, 1965年生. 清华大学博士毕业, 现为清华大学自动化系国家 CIMS 工程技术研究中心讲师. 目前感兴趣的领域为复杂制造系统的建模、分析和智能优化.

吴 澄 男, 中国工程院院士、国家863计划自动化领域首席科学家、清华大学自动化系教授、博士生导师、国家 CIMS 工程技术研究中心主任. 目前感兴趣的领域为复杂制造系统的建模、分析和智能优化.