

doi: 10.3969/j.issn.1000-5781.2009.03.010

基尼指数遗传算法求解具有共性的调度问题

杨晓梅^{1,2}, 梁吉业¹, 曾建潮², 梁嘉骅¹

(1. 山西大学管理学院, 山西 太原 030006;

2. 太原科技大学复杂系统与计算智能实验室, 山西 太原 030024)

摘要: 通过对不同类型车间作业调度问题的描述, 分析各类问题之间的共性特征, 提出了一种基于基尼指数的遗传算法. 该算法结合进化生态学和决策树算法的思想, 提出了新的遗传调度算子, 在算法操作过程中利用基尼指数值来衡量个体的优劣及种群间的差异, 确定交叉 (或变异) 父代并选择相应的交叉 (或变异) 点, 既维持了群体的多样性, 又能体现遗传算法较强的搜索能力. 仿真结果验证了该算法对于求解具有共性特征的调度问题是非常合适的.

关键词: 遗传算法; 基尼指数; 车间作业调度问题

中图分类号: TP18 **文献标识码:** A **文章编号:** 1000-5781(2009)03-0322-07

Gini-index genetic algorithm for the scheduling problems with similar characteristics

YANG Xiaomei^{1,2}, LIANG Jiye¹, ZENG Jianchao², LIANG Jiahua¹

(1. School of Management, Shanxi University, Taiyuan 030006, China; 2. Complex System and Computer Intelligence Laboratory, Taiyuan University of Science & Technology, Taiyuan 030024, China)

Abstract: Through describing the different job shop scheduling problem and analyzing the similar characteristic between them, a Gini-index genetic algorithm is proposed. This algorithm introduces a new genetic scheduling operator based on the idea of evolving ecology and decision tree algorithm. By calculating the Gini-index among the individuals, the operators can not only measure current population, but also find the crossover (or mutation) parents and the crossover (or mutation) points. It is helpful to raise the quality of genetic algorithm and guarantee population diversity. For the scheduling problems with similar characteristics, the computation results of simulation validate the effectiveness of the proposed algorithm.

Key words: genetic algorithm; Gini-index; job shop scheduling problem

0 引言

作为计算机集成制造系统领域的一个重要内

容, 对于车间作业 (job shop) 调度问题的研究有利于企业降低产品生产成本, 提高机器利用率, 增强企业车间管理水平. 在实际生产中, 车间作业调

收稿日期: 2007-05-08; 修订日期: 2008-06-24.

基金项目: 国家自然科学基金资助项目 (60773133); 山西省青年科研基金资助项目 (2006021019); 山西省自然科学基金资助项目 (2006011065).

度问题所具有的复杂性、约束性、动态性等特点,无疑增加了该问题的求解难度.近年来学者们在利用遗传算法研究车间作业调度问题时,大都以特定的车间作业调度问题为背景,虽然取得了一定的成果,但却因遗传算法的自身缺陷限制了问题的求解性能.本文首先通过对不同类型车间作业调度问题的描述,结合目前遗传调度算法的性质,并借鉴进化生态学原理和决策树算法的属性分裂准则,提出一种基于基尼指数的遗传调度算法.该算法针对不同类型的车间作业调度问题,挖掘出各问题间的共性特征,对其相同的特性采用基于基尼指数的遗传算子操作,同时将基尼指数作为个体多样性和优劣性的衡量指标,从而拓宽了遗传算法求解不同类型车间作业调度问题的思路,主要在以下三个步骤中结合了基尼指数的计算:首先判定群体中的优良个体加以复制,其次从每一群体中选择出需要进行交叉(变异)的个体,然后在确定交叉(变异)父代的交叉点时也引入基尼指数的计算,使得算法从最大限度上维持了群体的多样性,并保留了个体的优良性质.

1 车间作业调度问题的描述

一个传统的车间作业调度问题可描述如下:设 n 个工件在 m 台机器上进行加工,已知工件的加工顺序及每道工序加工时间,根据每个工件的加工工艺要求,应如何安排每台机器上所有工件的加工顺序,以使某种性能指标在满足给定约束条件下达到最优(本文中采取最小完工时间).该问题的约束条件除满足上述已知条件还包括:1)每一时刻每台机器只能加工一个工件;2)每个工件在固定时刻只能在同一台机器上加工;3)生产加工过程不间断.

如上的车间作业调度问题中,每个工件各道工序的加工机器是唯一确定的,而若在满足上述条件的基础上,每个工件在固定时刻可在多个可选机器中选其一进行加工,则该问题为柔性车间作业调度问题.

2 基于基尼指数的遗传调度算法

2.1 目前遗传调度算法的分析

鉴于车间作业调度问题具有复杂性、约束性、

动态性的特点^[1],在问题求解过程中不仅应考虑如何降低问题的复杂性或提高算法的求解性能,还需满足问题的约束条件及其动态变化状态.免费午餐定理表明,没有任何一个算法适用于求解所有的问题^[2],所以在搜索问题过程中还应考虑具体问题的特点.针对不同类型的车间作业调度问题,有必要从各问题间提取相同的特征,对其采取相同的操作方法,从而简化问题的复杂性.

近年来模拟退火算法、禁忌搜索算法等都曾被应用于调度领域^[3,4],但随着问题规模的扩大,算法计算量呈指数增长.自从1985年遗传算法被 Davis^[5]首次用于解决调度问题以来,该算法在车间作业调度问题的求解中得到成功应用,但遗传算法自身的缺陷无疑会制约算法的求解质量,如搜索前期若有超大个体出现易使算法陷入早期收敛,搜索过程中当群体的多样性较差,即个体间的差异较小会使算法停滞不前.目前群体的多样性虽然有一些衡量标准^[6-8],但进化个体间差异仅通过适应值的大小和个体的编码结构来区别,如何借助个体间的适应值,在后天进化过程中通过学习得到更有价值的信息,有利于提高遗传算法求解调度问题的能力.

2.2 基于基尼指数的遗传调度算法

2.2.1 基于基尼指数的遗传调度算法的分析

作为一个经济学指标,1912年 Gini 提出以基尼指数体现全社会贫富阶层之间整体差异程度^[9].鉴于基尼指数提出的初衷是用来刻画随机变量取值的散布程度,该指标在进化生态学理论中被作为群落多样性的一个度量指标,应用于生态学的理论研究^[10].1984年,Reiman 将基尼指数引入决策树算法中作为属性分裂的评价标准,从当前可选属性中选取基尼指数值最小的属性作为扩展属性^[11].具体操作为:首先假设集合 T 包含 N 个类别的样本,其基尼指数值为 $Gini(t) = 1 - \sum_{j=1}^N P_j^2$ (其中 P_j 为类别 j 在 t 节点处的相对频率).当 $Gini(t) = 0$ 时,此节点处所有样本均属于同一类别,若 $Gini(t)$ 越大,表示这一属性若被作为分裂属性,则对应的样本呈均匀分布,即说明节点分裂质量越好.

常见的遗传算法中,个体间的差异依赖于对个体适应值的计算或对个体基因模式间的差异加以比较,可是对于车间作业调度问题而言,同一个最优解所对应的最优调度方案并不唯一,因此多样性的度量如果仅凭个体适应值的大小或基因模式间的不同决定,无法真正反映个体间的差异.考虑到每一代群体中的个体均可被划为优个体和劣个体两类,本文在复制过程中引入基尼指数,判定个体的优劣及群体多样性,随后在交叉(或变异)过程中利用基尼指数确定交叉(或变异)父代和交叉点,提出一种基于基尼指数的遗传调度算法.

2.2.2 基于基尼指数的遗传调度算法的描述

遗传算法中的关键问题是编码方法和遗传算子的操作.其中编码方法的设定必须结合问题的特点并直接影响交叉算子、变异算子的操作.虽然传统车间作业调度问题和柔性车间作业调度问题所代表的生产环境不同,后者在确定正待加工的工序应加工的机器号后,求解过程与前者相同,因此这两个问题的性质有着相同之处.本文对于前一问题采用基于主动调度的编码方法,该编码方法依据自然数的编码形式确定待加工工序的优先加工顺序,后一问题采用分部编码方法^[12],分部编码方法以两行矩阵编码的形式表示,依据矩阵第一行的向量选择出该工序所加工的机器,第二行仍利用基于主动调度的编码方法,因此即使这两个问题的特点不完全相同,但在遗传过程中仍可对其中相似的染色体结构采取同样的算子操作.具体算法如下所示.

1) 多样性度量

遗传算法的进化过程中,种群间群体和群体内个体间相互竞争、互为作用而得到问题的解决方案.本文采取基尼指数度量这两种多样性,并在一个群体中所有个体被划分为优个体和劣个体两类,设一个体的适应值为 $value_i$, 群体规模为 $pop\ size$, 令 P_{i1} 、 P_{i2} 分别表示个体作为优、劣个体的相对概率,有 $P_{i1} = value_i / \sum (value)$, $P_{i1} = 1 - P_{i2}$, 使得 $Gini(i) = 1 - P_{i1}^2 - P_{i2}^2$, 则 $Gini(i)$ 值可用来度量个体间的差异.在此基础上计算群体的多样性为 $diver_i = 1 - \sum_{i=1}^{pop\ size} P_{i1}^2 - \sum_{i=1}^{pop\ size} P_{i2}^2$. 复制时

首先计算个体的 $Gini(i)$ 值,以此作为个体的复制概率得到复制后代.这一方法有助于衡量算法过程中的种间、种内多样性,从一定程度上还能够引导算法淘汰劣个体,有效保留群体中的优个体.

2) 交叉(或变异)父代及交叉(及变异)点的选择

常见交叉(或变异)父代和交叉点的选择过程都有相应的随机操作,如选择交叉(或变异)父代时,设交叉概率为 P_c ,对于每一个个体选取一个随机数 r_i ,若 $r_i < P_c$,则该个体作为交叉(或变异)父代,令交叉(或变异)个体两两组合,针对随机选择的交叉点,利用遗传算子操作产生后代个体,该方法具有一定的任意性,无法判定后代是否朝着有利进化的趋势演化,为此本文提出了基于基尼指数的交叉(或变异)父代及交叉点的选择方法.

(a) 基于基尼指数的交叉操作

进行交叉操作前,首先应选择出交叉父代,具体方法是选取 $pop\ size$ 个随机数 ra .比较每一个随机数和 P_c 的大小,若有 k_c 个随机数满足 $ra_i < P_c$,则本代应进行交叉的个体总数为 k_c .从当代群体中随机选择三个个体两两组合,将基尼指数值最小的两个组合个体作为所选择的交叉父代,重复如上过程直到选择出所有的交叉父代.鉴于本文的染色体编码表示待加工工序的优先加工顺序,染色体在解码时仅利用了个体中部分基因的排列顺序,因此有必要加大各基因位的变化,才能从一定程度上提高最优解的搜索效率.依据编码特点,执行优先调度次序互换的交叉操作:首先,将交叉点分为适于交叉和不适于交叉两类,在随机选定两个待交叉点位置的基础上,设这两个待交叉点依次为 $dotc_1$, $dotc_2$, 染色体长度为 $chrom\ size$, 其中一个交叉点及其对应基因值在另一父代的基因位置把两个交叉父代分割为 s_1 , s_2 , s_3 , s_4 个基因,且

$$s_1 + s_2 = chrom\ size - 1, s_3 + s_4 = chrom\ size - 1$$

另一个交叉点及其对应基因值在另一父代的基因位置把两个交叉父代分割为 s_5 , s_6 , s_7 , s_8 个基因,且

$$s_5 + s_6 = chrom\ size - 1, s_7 + s_8 = chrom\ size - 1$$

令 $t_i = \frac{s_i}{\text{chrom size}}$ ($i = 1, 2, \dots, 8$), 则

$$\text{Gini}_1 = 1 - \sum_{i=1}^4 \left(\frac{t_i}{2}\right)^2, \text{Gini}_2 = 1 - \sum_{i=5}^8 \left(\frac{t_i}{2}\right)^2$$

选择 $\min(\text{Gini}_i)$ 所对应的交叉点作为最终确定的交叉点, 通过群体多样性的衡量, 对于连续多代 diver 差异较小时, 互换该交叉位前后的基因段, 得到相应的交叉父代, 否则对于连续多代 diver 差异较大的情况, 说明群体仍保持着很好的多样性, 则仅互换二交叉位的基因后产生相应的交叉父代.

该交叉方法利用基于基尼指数值的交叉父代及交叉点选择策略, 不仅提高了优良个体的利用价值, 而且在交叉过程中倾向于父代染色体内的基因相互竞争得到有效的交叉后代, 即避免了不可行调度方案的产生, 又有利于扩大求解问题的空间.

(b) 基于基尼指数的变异操作

变异过程仍采用基于基尼指数的变异算子操作, 首先选取 pop size 个随机数 x_2 , 计算 $x_2 < P_m$ 的个体数目 k_m , 若该数值不等于 0, 则说明当代有 k_m 个个体被作为变异父代, 随机从多样性最大的那一代个体和当代个体选择 $k_m \times 2$ 个个体, 以基尼指数值排在前面 k_m 个的个体作为变异父代. 针对每一个变异父代, 随机选择两个变异点, 若两个变异点恰恰为第一个和最后一个基因, 则采取倒

序变异操作, 否则该染色体被分为三个基因段, 每个基因段内包括若干个基因, 根据每两段组合基因的基尼指数值判定应进行互换的两段基因, 互换这两个基因段, 从而得到所需的变异后代.

该变异方法在执行过程中考虑到变异点分割后的基因段中的基因数目, 通过基尼指数值确定得出较优的互换基因段, 不仅操作方便, 而且所得到的变异后代确实能够改变工序的优先调度次序, 有助于维持群体多样性.

3 仿真结果及分析

为了验证该算法的适用性和有效性, 并具体对这些问题染色体编码和解码后的调度方案加以描述, 本文利用 MATLAB 语言, 在 Intel 1.0 的计算机上进行编程, 将该算法应用于不同类型的车间作业调度问题. 算法所采用的参数设置为: 种群规模 $\text{pop size} = 20$, 交叉概率 $P_c = 0.85$, 变异概率 $P_m = 0.15$, 截止代数 $\text{maxgen} = 300$, 针对每个问题算法执行 30 次.

1) 传统车间作业调度问题 (以 F06、La06 和 La14 问题为例, 问题的数据集可参看文献 [1])

(a) F06 问题. 通过对该问题进行仿真可得到问题的最优解 55 s 其中一个最优解所对应调度方案的工件和机器甘特图分别如图 1、图 2 所示.

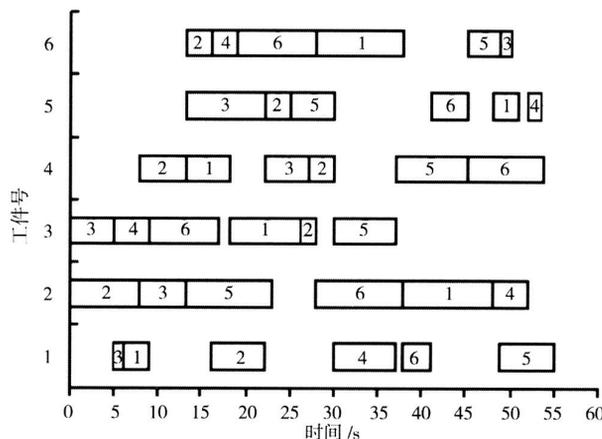


图 1 F06 的工件甘特图

Fig 1 The job Gantt chart for F06

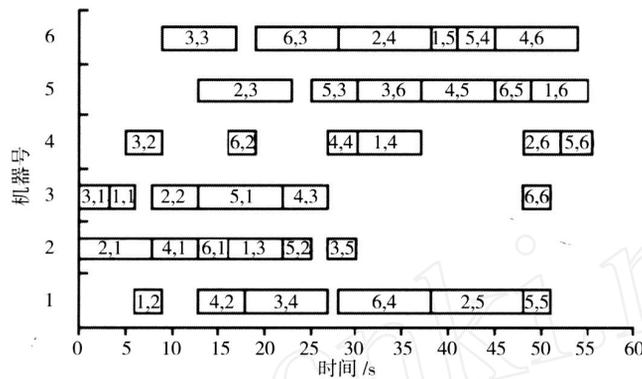


图 2 F06的机器甘特图

Fig 2 The machine Gantt chart for F06

(b) 不同规模车间作业调度问题. 对不同规模的传统车间作业调度问题进行仿真后可得到表 1 所示的结果.

表 1 基尼指数的遗传算法对传统车间作业调度问题的仿真结果

Table 1 The results of Gini-index GA for general job shop

问题	问题规模 (n, m)	最优值	最差值	平均值	平均收敛代数
La06	15, 5	926	945	936.9	129.0
La14	20, 5	1 292	1 310	1 302.1	25.8

2) 柔性车间作业调度问题

(a) 柔性调度问题 (3 ×6). 以文献 [13] 中 3 个工件在 6 台机器上加工的测试数据集为例, 该柔性车间作业问题中, 各工件每一工序均可同时选择在所有机器中进行加工. 通过对问题进行仿真得到问题的最优解 18 s, 收敛率 100%, 所求得的其中一个最优染色体的工件甘特图如图 3 所示.

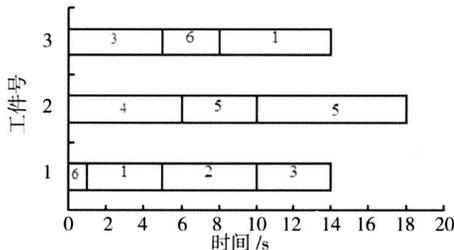


图 3 最优染色体的工件甘特图

Fig 3 The job Gantt chart of optimal chromosome

虽然文献 [13] 的算法也产生相同的最优解, 但由其所列出的解决方案可以看出, 通过禁忌搜索算法得到的第二个工件的第三道工序的加工时间为 7 s 由原数据集可知, 该工序所对应的加工时间应为 8 s 因此该方案实际并非问题最优解.

而本文所提出的算法求出的最优解决方案符合各项约束条件, 确实是该问题的最优解.

(b) 柔性车间作业调度问题 (4 ×6) 的求解. 该柔性车间作业问题中所采用的数据集如文献 [14] 所示, 各工件每一工序可选择在部分机器中进行加工. 其中, 总的加工工序数是 12 道工序, 通过对该问题加以仿真可得到表 2 所示的结果.

表 2 基尼指数的遗传算法对柔性车间作业调度问题的仿真结果

Table 2 The results of Gini-index GA for 4 ×6 FJSP

最优值	最差值	平均值	收敛率 (%)	平均收敛代数	CPU 时间 (s)
17	18	17.5	100	51.7	20.4

与文献 [12] 中所示算法相比较, 本文算法从收敛率和收敛速度上有了极大的改善, 算法的执行时间较快, 所得到的其中一个最优调度方案如图 4 图 5 所示.

图 4 所示的最优调度解优于文献 [14], 即与文献 [15] 同样都能得到最优调度解为 17 s 平均完成时间为 11.75 s 的仿真结果; 由图 5 可知, 本文算法所得到的平均机器完成时间为 7.83 s 而文献 [15] 中所示的两个最优调度方案的平均机器完成时间分别为 8 s 和 11 s 因此相对于文献 [15] 中的算法而言, 本文算法从一定程度上提高了机器的利用率.

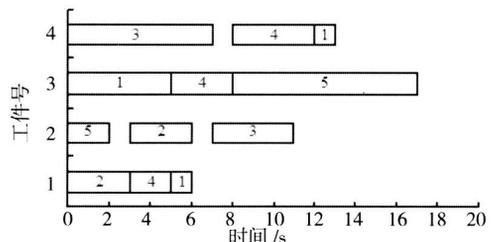


图 4 最优染色体的工件甘特图

Fig 4 The job Gantt chart of optimal chromosome

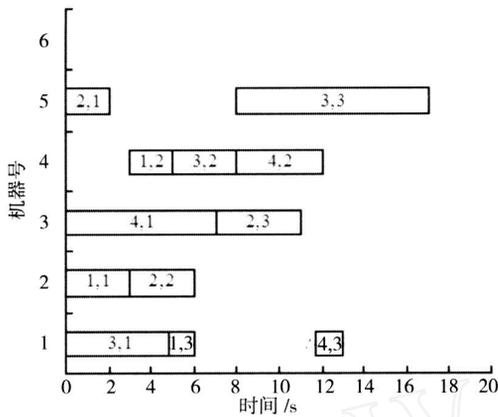


图 5 最优染色体的机器甘特图

Fig 5 The machine Gantt chart of optimal chromosome

(c) 柔性车间作业调度问题 (6 × 6) 的求解. 以文献 [16] 中的测试数据为例对算法进行验证. 为了便于比较, 本例以最小完成时间作为性能评价指标, 种群规模 $pop\ size = 40$, 交叉概率 $P_c = 0.85$, 变异概率 $P_m = 0.15$, 截止代数 $maxgen = 40$. 通过仿真得到图 6 所示的最优调度方案.

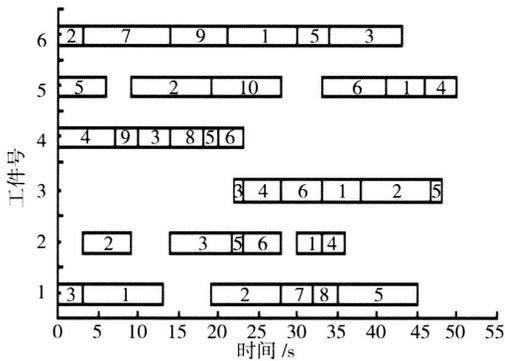


图 6 工件甘特图

Fig 6 The job Gantt chart

图 6 所示调度方案的总完成时间为 245 s, 所对应各工件的完工时间依次为 45 s、36 s、48 s、23 s、50 s 和 43 s 与文献 [16] 的最优调度解 269 s 相比, 本文算法从一定程度上提高了问题的搜索效率, 能够得到更优的调度方案.

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 L. Job Shop Scheduling and Genetic Algorithm [M]. Beijing: Tsinghua University Press, 2003. (in Chinese)
 [2] Wolpert D H, Macready W G. No free lunch theorems for optimization [J]. IEEE Transactions on Evolutionary Computation, 1997, 1(1): 67—82.
 [3] Watson J P, Beck J C, Howe A E, et al. Problem difficulty for tabu search in job-shop scheduling [J]. Artificial Intelli-

上述仿真结果表明, 基于基尼指数的遗传算法借助个体间差异的衡量, 能够适当保留优良基因, 并通过相应的交叉和变异操作保证群体的多样性, 有利于引导算法向有利于搜索到全局最优调度方案的趋势进化. 虽然免费午餐定理的提出证实目前并没有一个通用的问题求解算法, 但对于不同类型的问题, 只要这些问题中具有部分相同的特点, 则可对这些特点所表现出的性质或数据采取类似的算法操作, 本文对于传统车间作业调度问题和柔性车间作业调度问题的特点加以分析后, 对其编码后染色体中具有相同性质的部分 (如传统车间作业调度问题的编码方法和柔性车间作业调度问题所示矩阵编码的第二行向量: 表明各工序优先加工顺序) 采用了同样的基于基尼指数的遗传算子操作. 该方法在设计过程中也充分考虑了问题的特征和遗传算法的性质, 在一定程度上缩短了算法的执行时间, 降低了算法的求解复杂性, 还可被应用于解决其它具有类似特点的实际生产问题.

4 结束语

通过对不同类型车间作业调度问题的描述及目前遗传调度算法求解步骤的分析, 本文从经济学角度出发, 借鉴进化生态学原理和决策树算法的属性选择准则, 提出了基于基尼指数的遗传调度算法. 该算法不仅有效地保证了遗传后代的可行性, 较快地得到最优调度方案, 还有利于企业提高机器利用率, 缩短生产周期, 从而增强企业生产车间管理能力. 此外, 由于该算法所研究的对象是具有共性特征的调度问题, 因此, 也可以将其扩展到对于不确度性调度, 无等待调度等问题间共性特征的分析及其算法设计的研究.

- gence, 2003, 143(2): 189—217.
- [4] Steinhofel K, Albrecht A, Wong C K. Fast parallel heuristics for the job shop scheduling problem [J]. *Computers and Operations Research*, 2002, 29(2): 151—169.
- [5] Davis L. Job shop scheduling with genetic algorithms [A]. In: *Proceedings of the 1st International Conference on Genetic Algorithm [C]*. New Jersey: Lawrence Erlbaum Associates, 1985. 13—140.
- [6] Ursem R K. Diversity-guided evolutionary algorithms [A]. In: *Proceedings of the 7th International Conference on Parallel Problem Solving From Nature, LNCS 2439 [C]*. Berlin: Springer-Verlag, 2002. 462—474.
- [7] Zhu K Q, Liu Z W. Population diversity in permutation based genetic algorithm [A]. In: *Proceedings of the 15th European Conference on Machine Learning [C]*. Pisa, Italy: Springer-Verlag, 2004. 537—547.
- [8] Wineberg M, Oppacher F. A linear time algorithm for determining population diversity in evolutionary computation [A]. In: *Proceedings of the ASTED International Conference on Intelligent Systems and Control [C]*. Salzburg, Austria: ACTA Press, 2003. 275—286.
- [9] 陈希儒. 基尼系数及其估计 [J]. *统计研究*, 2004, 21(8): 58—60.
Chen X R. Gini coefficient and the estimation method [J]. *Statistical Research*, 2004, 21(8): 58—60. (in Chinese)
- [10] 陶玲, 任珺. 进化生态学的数量研究方法 [M]. 北京: 中国林业出版社, 2004.
Tao L, Ren J. *Quantity Method of Evolutionary Ecology [M]*. Beijing: China Forestry Press, 2004. (in Chinese)
- [11] Raileanu L E, Stoffel K. Theoretical comparison between the Gini index and information gain criteria [J]. *Annals of Mathematics and Artificial Intelligence*, 2004, 41(1): 77—93.
- [12] 杨晓梅, 曾建潮. 遗传算法求解柔性 job shop 调度问题 [J]. *控制与决策*, 2004, 19(10): 1197—1200.
Yang X M, Zeng J C. Solving flexible job shop scheduling problem using genetic algorithm [J]. *Control and Decision*, 2004, 19(10): 1197—1200. (in Chinese)
- [13] Liouane N, Saad I, Hammadi S, *et al*. Ant systems & local search optimization for flexible job shop scheduling production [J]. *International Journal of Computers, Communications & Control*, 2007, 2(2): 174—184.
- [14] Nasr N, Elsayed E A. Job shop scheduling with alternative machine [J]. *International Journal of Production Research*, 1990, 28(9): 1595—1609.
- [15] Moon I, Lee S, Bae H. Genetic algorithm for job shop scheduling problems with alternative routings [J]. *International Journal of Production Research*, 2008, 46(10): 2695—2705.
- [16] 白音, 尚文利, 王成恩. 用遗传算法求解可变机器约束的 job-shop 调度问题 [J]. *信息与控制*, 2004, 33(5): 604—608.
Bai Y, Shang W L, Wang C E. A genetic algorithm for job-shop scheduling problem with alternative machine [J]. *Information and Control*, 2004, 33(5): 604—608. (in Chinese)

作者简介:

- 杨晓梅 (1973—), 女, 陕西临潼人, 博士生, 研究方向: 数据挖掘、智能算法, Email: yxm001_73@sohu.com;
梁吉业 (1962—), 男, 山西晋城人, 教授, 博士生导师, 研究方向: 智能决策、数据挖掘;
曾建潮 (1963—), 男, 陕西大荔人, 教授, 博士生导师, 研究方向: 智能算法、系统工程;
梁嘉骅 (1937—), 男, 广东广州人, 教授, 博士生导师, 研究方向: 系统工程、企业生态.