

求解多目标组合优化的改进 Pareto 适应度遗传算法

杨开兵^{1,2}, 刘晓冰²

YANG Kai-bing^{1,2}, LIU Xiao-bing²

1.大连工业大学 信息科学与工程学院, 辽宁 大连 116034

2.大连理工大学 CIMS 中心, 辽宁 大连 116024

1.College of Information Science and Engineering, Dalian Polytechnic University, Dalian, Liaoning 116034, China

2.CIMS Center, Dalian University of Technology, Dalian, Liaoning 116024, China

E-mail: kaibing@126.com

YANG Kai-bing, LIU Xiao-bing. Improved Pareto Fitness Genetic Algorithm for multi-objective combinatorial optimization. Computer Engineering and Applications, 2009, 45(8): 44-46.

Abstract: Combining Pareto Fitness Genetic Algorithm (PFGA) with local search, an Improved Pareto Fitness Genetic Algorithm (IPFGA) for multi-objective combinatorial optimization is proposed. In the proposed algorithm, a local search procedure is applied to each solution generated by genetic operations. The concept of Pareto dominance is used in the local search procedure, and the idea of crowding distance in the external population is used to perform elitism. The experimental results show that the IPFGA, compared with the known PFGA, has faster convergent speed.

Key words: multi-objective combinatorial optimization; Pareto Fitness Genetic Algorithm (PFGA); local search

摘要: 将 Pareto 适应度遗传算法 (PFGA) 与局部搜索相结合, 提出了一种用于求解多目标组合优化问题的改进算法 IPFGA, 该算法基于 Pareto 支配关系对遗传操作产生的每一个个体进行局部搜索, 并采取在外部群体中引入拥挤距离的精英选择策略。实验结果表明, 与 PFGA 相比, IPFGA 有更快的收敛速度。

关键词: 多目标组合优化; Pareto 适应度遗传算法; 局部搜索

DOI: 10.3778/j.issn.1002-8331.2009.08.014 文章编号: 1002-8331(2009)08-0044-03 文献标识码: A 中图分类号: TP301

1 引言

多目标组合优化是实际中广泛存在的 NP 难问题。由于目标间的冲突性以及计算的复杂性, 使得多目标组合优化问题的求解尤为困难。近年来, 利用进化算法解决多目标问题已经成为一个研究热点, 出现了许多优秀的多目标进化算法, 例如 MOGA^[1]、SPEA^[2]、PAES^[3]、NAGA-II^[4] 等。最近 Elaoud 等人^[5] 提出了一个 Pareto 适应度遗传算法 PFGA (Pareto Fitness Genetic Algorithm), 其新颖之处在于个体的适应度函数采用基于 Pareto 支配关系的双倍排序策略以及自适应密度评估技术, 不仅能够体现个体间的 Pareto 最优性, 还能够体现它们在整个搜索空间的分布, 能够更好地防止遗传漂移, 保持群体多样性, 在求解 6 个具有不同 Pareto 前沿特征的测试函数上取得了较好的优化效果。然而, PFGA 也存在着不足, 比如局部搜索能力较差等。

为增强局部搜索能力, 本文提出了一种改进的算法 IPFGA (Improved Pareto Fitness Genetic Algorithm)。该算法将局部搜索融入到 PFGA 中, 对进化群体中的每一个体采取基于 Pareto 支配关系的局部搜索, 并在外部群体中引入拥挤距离进行精英

选择, 以提高算法的搜索效率。通过对一组多目标流水车间问题的求解, 并与 PFGA^[6] 比较, 结果表明, IPFGA 有更快的收敛速度和更好的优化效果。

2 问题描述

一般地, 多目标组合优化问题可描述如下^[6]:

$$\text{Minimize } z=f(\mathbf{x})=(f_1(\mathbf{x})=z_1, \dots, f_r(\mathbf{x})=z_r)$$

$$\text{Subject to } \mathbf{x} \in \Omega$$

其中, $f(\mathbf{x})$ 为具有 r 个分量的目标向量, \mathbf{x} 是离散的决策向量, Ω 是可行解的集合, 点 $z=f(\mathbf{x})$ 是解 $\mathbf{x} \in \Omega$ 在目标空间中的像。

如果对任意的 $j \in \{1, 2, \dots, r\}$, 有 $z_j=f_j(\mathbf{x}) \leq z'_j=f_j(\mathbf{x}')$, 并且至少存在一个 j , 有 $z_j < z'_j$, 则称点 $z=(z_1, \dots, z_r)$ 支配点 $z'=(z'_1, \dots, z'_r)$ 。

如果解 \mathbf{x} 的像支配解 \mathbf{x}' 的像, 则称解 \mathbf{x} 支配解 \mathbf{x}' 。

如果不存在 $\mathbf{x} \in \Omega$ 使得 $z=f(\mathbf{x})$ 支配 $z'=f(\mathbf{x}')$, 则称解 $\mathbf{x}^* \in \Omega$ 是 Pareto 最优解 (或非劣解)。

Pareto 最优解组成的集合称为 Pareto 最优集。Pareto 最优集在目标空间的像称为 Pareto 前沿 (或非劣前沿)。

基金项目: 国家自然科学基金 (the National Natural Science Foundation of China under Grant No.70572098)。

作者简介: 杨开兵 (1967-), 女, 讲师, 研究方向: 多目标优化、计算机集成制造; 刘晓冰 (1956-), 男, 教授, 博士生导师, 研究方向: 计算机集成制造、企业信息化。

收稿日期: 2008-01-21

修回日期: 2008-04-16

3 改进的 Pareto 适应度遗传算法

PFGA 采用简单遗传算法的框架, 在每次迭代中, 对当前群体的每一个体赋予一个适当的适应度值, 然后进行选择、交叉和变异, 产生子代群体, 进行下一次迭代。在进化过程中还随机地从外部群体中选择一定数目的优秀个体进入子代群体中, 参加下一次进化。本文在 PFGA 的基础上, 改进了精英选择方式和适应度赋值方式, 并融入了局部搜索操作, 以增强局部搜索能力, 提高算法的搜索效率。

3.1 适应度赋值

PFGA^[6]基于 Pareto 支配关系, 采取双倍排序策略和自适应密度评估方法, 对群体中的个体进行适应度赋值, 先根据 Pareto 支配关系计算每一个体的排序数, 再根据个体周围的拥挤状况计算个体的密度值, 最后综合确定个体的适应度值。具体作法如下:

(1) 计算当前群体 P 中每个个体 i 的伪排序数 $R'(i)$:

$$R'(i) = |\{j \in P, j \succ i\}| \quad \forall i \in P$$

其中, 符号“ \succ ”表示 Pareto 支配关系, 即 $R'(i)$ 表示当前群体 P 中支配个体 i 的个体数。

(2) 计算个体 i 的排序数 $R(i)$:

$$R(i) = R'(i) + \sum_{j \in P, i \prec j} R'(j) \quad \forall i \in P$$

即个体 i 的排序数 $R(i)$ 等于个体 i 的伪排序数与支配个体 i 的所有个体的伪排序数之和。

(3) 将目标空间根据群体规模 N_{pop} 划分成 $n_c \times n_c \times \dots \times n_c$ 个网格区域, n_c 表示每维目标空间的网格数, n_c 的确定方法为: 设 $\sqrt[n]{N_{pop}}$ 的整数部分为 a , 小数部分为 r , 则当 $r=0$ 时, $n_c=a$, $r \neq 0$ 时, $n_c=a+1$ 。每一个体都位于某个网格区域内, 将每个个体所在的网格区域内的个体数作为该个体的密度值。

(4) 个体的适应度为:

$$fitness(i) = \frac{1}{\exp(R(i)) \times D(i)}$$

其中 $R(i)$ 表示个体 i 的排序数, $D(i)$ 表示个体 i 的密度值。

这种适应度赋值方式, 既保证了群体朝着 Pareto 前沿逼近, 又能在一定程度上保持群体多样性。为进一步提高群体多样性, 本文在上述适应度赋值的基础上, 采取将群体中的极端个体赋予最大的适应度, 以增加被选择的机会, 避免 Pareto 前沿过于狭窄, 使产生的非劣前沿分布范围更广。

3.2 精英选择策略

PFGA 使用两个群体: 进化群体和外部群体。外部群体用来保存进化过程中产生的所有非劣解。在每次迭代中, 从外部群体中随机地选择 N_{elite} 个个体, 加入到下一代群体中参加进化。这种选择方式难以保证选择到分散程度较好的 N_{elite} 个个体。为此本文基于文[7]提出改进的精英选择策略, 即在外部群体中, 计算每个个体与其他个体在目标空间中的欧几里德距离, 将其中的最小值作为该个体的拥挤距离, 在每次迭代中, 从外部群体中选择 N_{elite} 个拥挤距离最大的个体作为精英, 加入到下一代群体中, 当外部群体中的个体数不足 N_{elite} 时, 不足的部分再次从外部群体中按拥挤距离由大到小依次选择。这种策略不但操作简单, 而且能够选择密度较小的外部群体中的个体, 更好地维持群体多样性, 提高算法的收敛速度。

3.3 多目标局部搜索

研究表明, 进化算法和局部搜索的混合, 在求解多目标优化问题, 特别是求解多目标组合优化问题中, 常常表现出胜过

“纯”进化算法的较快的收敛速度和较高的全局搜索能力^[8]。本文将局部搜索融入到 PFGA 中, 对进化群体中的每一个体采取基于 Pareto 支配关系的局部搜索。局部搜索过程应用到遗传操作产生的每一个新个体上, 也应用到从外部群体中选择的精英个体上。设 x 是实施局部搜索的初始解, 从 x 的邻域 $N(x)$ 中搜索更好的解, 不妨设为 y 。再将 y 作为初始解, 重复上述搜索过程, 这样会得到一个改善的解。为了减少局部搜索的时间, 在搜索过程中限制检查的邻域解的个数, 即设定搜索步长 N_{iter} , 如果从当前解的邻域中检查了 N_{iter} 个解而没找到更好的解, 则搜索停止。具体步骤如下:

(1) S 是实施局部搜索的解的集合, $|S|$ 是 S 中元素的个数, $S = \{x_1, x_2, \dots\}$, 令 $k=1$ 。

(2) 确定局部搜索的当前解 x_k , 令 $t=1$ 。

(3) 产生一个解 $y \in N(x_k)$, 检查 y 与 x_k 的支配关系。

(4) 若 y 支配 x_k , 则用 y 替代 x_k , 转步骤 2; 若 y 不支配 x_k , 且 $t < N_{iter}$, 则 $t=t+1$, 转步骤 3; 若 y 不支配 x_k , 且 $t \geq N_{iter}$, 接步骤 5。

(5) 如果 $k < |S|$, 则 $k=k+1$, 返回步骤 2; 否则, 停止。

3.4 IPFGA 的算法过程

(1) 算法参数设定: 群体规模 N_{pop} , 交叉概率 P_c , 变异概率 P_m , 最大精英数 N_{elite} , 局部搜索的搜索步长 N_{iter} 。

(2) 初始化: 随机产生规模为 N_{pop} 的初始群体, 计算每个个体的目标值。初始化外部群体 $ND_1 = \emptyset$ 。

(3) 求非劣解: 求出当前群体 P_t 中的所有非劣解。

(4) 修正外部群体 ND_t : 将 P_t 中的非劣解加入到 ND_t 中, 再从 ND_t 中删去被支配的个体。

(5) 适应度赋值: 按 3.1 节的方法计算 P_t 中所有个体的适应度。

(6) 选择、交叉和变异: 从 P_t 中选择两个不同的父代, 依概率 P_c 进行交叉得到子代, 子代依概率 P_m 进行变异, 共产生 $N_{pop} - N_{elite}$ 个个体, 形成子代群体 P_{t+1} 。

(7) 精英选择: 按 3.2 节的方法从集合 ND_t 中选出 N_{elite} 个个体, 加入到子代群体 P_{t+1} 中。

(8) 局部搜索: 应用 3.3 节的方法对子代群体 P_{t+1} 中的每一个个体进行局部搜索, 得到改善后的子代群体 P_{t+1}^* 。

(9) 停止准则: 若停止条件满足, 则输出外部群体, 否则 $P_t = P_{t+1}^*$, 返回步骤 3。

4 实验结果

4.1 测试问题

Flow shop 问题是一种典型的组合优化问题, 置换 Flow shop 调度可以描述为 n 个工件按照同一加工顺序在 m 台机器上依次进行处理, 要求确定使预定目标函数最小的工件加工顺序。这意味着问题的解为 n 个工件的一个全排列, 其搜索空间的规模是 $n!$ 。本文的优化目标是使最大完成时间 $f_1(\text{make span})$ 和最大拖后时间 $f_2(\text{maximum tardiness})$ 达到最小。测试算例^[9]包括机器数 m 为 10 和 20, 工件数 n 为 20、40、60 和 80 的 6 个规模不同的问题, 每一问题表示成 $n \times m$ 。算法均由 Visual B6.0 编程实现, 在 CPU 为 1.0 GHz、内存为 256 M 的个人电脑上进行。

4.2 性能指标

本文采用 Fieldsend 等^[9]提出的 C 指标来比较两个集合的收敛速度。设 A, B 是两个待比较的集合, C 的定义如下:

$$C(A, B) = \frac{| \{ b \in B : \exists a \in A, a > b \} |}{|B|}$$

$C \in [0, 1]$, 表示 B 的非劣解被 A 的非劣解支配的个数占 B 的非劣解总数的比率。当 B 中任何点都被 A 中某些点支配时, $C(A, B)=1$ 。当 B 中所有点都不被 A 中任何点支配时, $C(A, B)=0$ 。一般 $C(A, B) \neq C(B, A)$ 。

4.3 实验结果

采用基于工件加工顺序的自然数编码, 统一设定参数: 群体规模 $N_{pop}=30$, 交叉概率 $P_c=0.9$, 变异概率 $P_m=0.3$, 精英解的数目 $N_{elit}=3$ 。两种算法使用相同的遗传操作, 即两点交叉和插入变异, 局部搜索采用随机选择两个工件交换其加工排序的邻域结构且搜索步长 N_{iter} 为 2。计算终止条件均设定为评价 100 000 个解。每种算法在相同的初始条件下独立运行 20 次。表 1 给出了 C 指标的统计结果以及每种算法的平均运行时间。由表 1 可知, IPFGA 具有更快的收敛速度和较短的运行时间。

表 1 实验结果

问题	C 指标的统计结果		运行时间/s	
	$C(A, B)$	$C(B, A)$	IPFGA	PFGA
20×10	0.947 4	0	25	33
40×10	0.954 5	0	50	70
40×20	1.000 0	0	85	106
60×10	1.000 0	0	74	107
60×20	1.000 0	0	129	185
80×20	1.000 0	0	178	251

说明: C 指标评价中, A 表示 IPFGA 所得的非劣解集, B 表示 PFGA 所得的非劣解集。

图 1(a)和(b)以问题 20×10 和 40×20 为例, 比较了不同算

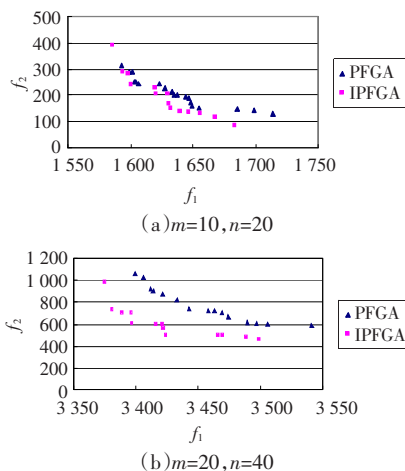


图 1 不同算法得到的近似集比较

(上接 25 页)

模型与制造过程之间的约束关系开发的能源信息管理系统, 针对离散制造业能源消耗点分散的特点, 实现了对离散制造业企业能源的精细化管理。通过企业的应用表明了该系统有效、实用, 并且大大降低了企业经营成本。

参考文献:

[1] Spinney P J, Campbell W G. Monte Carlo simulation techniques and electric utility resource decisions[J]. Energy Policy, 1996, 24: 155-163.
 [2] Zaheer-Uddin M, Zheng G R. A VAV system model for simulation of energy management control functions: off normal operation and duty cycling[J]. Energy Convers Manage, 1994, 35: 917-931.

法得到的非劣解集。从图中可以看出, 在相同的算法终止条件下, IPFGA 得到了分布较好的非劣解前沿的近似。

5 结语

本文提出了一种求解多目标组合优化问题的改进的遗传算法 IPFGA。IPFGA 采用改进的 PFGA 适应度赋值技术及精英选择策略, 维护群体多样性, 并基于 Pareto 支配关系对进化的每一个个体进行局部搜索, 以提高算法的局部搜索能力。实验证明, IPFGA 比 PFGA 在求解多目标组合优化问题时有更高的搜索效率和更快的收敛速度。

参考文献:

[1] Fonseca C M, Fleming P J. Genetic algorithm for multiobjective optimization: Formulation discussion and generation[C]//Proceedings of the Conference on genetic algorithm, San Mateo, California, 1993: 416-423.
 [2] Zitzler E, Thiele L. Multiobjective evolutionary algorithm: A comparative case study and the strength Pareto approach[J]. IEEE Transactions on Evolutionary Computation, 1999, 3(4): 257-271.
 [3] Knowles J, Corne D. Pareto archived evolution strategy: A new baseline algorithm for multiobjective optimization[C]//Proceedings of the Congress on Evolutionary Computation, Piscataway, NJ, IEEE, 1999: 98-105.
 [4] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.
 [5] Elaoud S, Loukil T, Teghem J. The Pareto fitness genetic algorithm: Test function study[J]. European Journal of Operational Research, 2007, 177(3): 1703-1719.
 [6] Arroyo J E C, Armentano V A. A genetic local search for multi-objective flowshop scheduling problems[J]. European Journal of Operational Research, 2005, 167: 717-738.
 [7] 林丹, 赵瑞. 基于随机权重的多目标进化算法[J]. 计算机工程与应用, 2006, 42(32): 4-6.
 [8] Ishibuchi H, Yoshida T, Murata T. Balance between genetic search and local search in memetic algorithms for multiobjective permutation flowshop scheduling[J]. IEEE Transactions on Evolutionary Computation, 2003, 7(2): 204-223.
 [9] Fieldsend J E, Everson R M, Singh S. Using unconstrained elite archives for multiobjective optimization[J]. IEEE Transactions on Evolutionary Computation, 2003, 7(3): 305-323.
 [3] Huang J P, Poh K, Ang B W. Decision analysis in energy and environmental modeling[J]. Energy, 1995, 20: 843-855.
 [4] Boonekamp P G M. Monitoring the energy use of households using a simulation model[J]. Energy Policy, 1997, 25: 781-788.
 [5] 刘永生. 用 LonWorks 控制网络构建企业节能平台[J]. 仪器仪表标准化与计量, 2005(6): 12-14.
 [6] 王永川, 陈光明, 李建新. 火电厂能源管理系统方案研究[J]. 电站系统工程, 2003, 19(5): 18-20.
 [7] 姚鹏, 张华忠, 尚艳艳. 无线传感器网络中基于低能量自适应聚类层次协议的多帧通信机制[J]. 计算机应用, 2006, 26(3): 540-542, 557.
 [8] 刘永. 基于 Web 能源管理系统中数据采集与处理方法的研究[D]. 哈尔滨: 哈尔滨工业大学机电工程学院, 2007.