

【文章编号】 1004-1540(2010)01-0052-07

精英策略的改进非支配遗传算法

张晓菲, 张火明

(中国计量学院 计量测试工程学院, 浙江 杭州 310018)

【摘要】 在研究 NSGA-II 算法的基础上, 提出了一种新的精英策略. 它既可以使种群中优良个体保持到下一代, 又可以维护种群的多样性, 避免算法陷入局部最优解. 此外, 对 NSGA-II 算法的交叉算子和变异算子进行了重新选择, 交叉算子采用模拟二进制交叉算子, 变异算子采用非均匀变异算子. 将它们引入 INSGA-II 算法中, 在优化前期可以进行大范围搜索, 而后期可以集中搜索某个重点区域, 从而可以提高解的精确性. 通过经典函数测试, 并将 INSGA-II 算法与 NSGA-II 算法的结果进行比较. 结果表明, 基于精英策略的改进算法 INSGA-II 不仅取得了较好的结果, 而且分布性也得到了有效的提高.

【关键词】 多目标优化; 精英策略; 非支配排序遗传算法

【中图分类号】 TP18

【文献标识码】 A

Improved non-dominated sorting genetic algorithm II based on the elitist strategy

ZHANG Xiao-fei, ZHANG Huo-ming

(College of Metrology and Measurement Engineering, China Jiliang University, Hangzhou 310018, China)

Abstract: A new elitist strategy based on NSGA-II was proposed, which could not only make excellent individuals in population maintained until the next generation, but also maintain the population diversity and avoid the algorithm into a local optimal solution. In addition, the crossover operator and mutation operator on NSGA-II algorithm were re-selected. SBX and the non-uniform mutation operator were introduced in order to realize INSGA-II search in a wide range in the early stage of optimization, but could realize a focus search on key areas in the latter stage. So they could improve the solution accuracy. The results of the classical function by INSGA-II were compared to the ones by NSGA-II, which indicates that the improved algorithm based on the elitist strategy INSGA-II has achieved good results, and also that the distribution has been effectively improved.

Key words: multi-objective optimization; elitist strategy; NSGA

在大部分情况下, 对于一个多目标优化问题, 各个子目标有可能是相互冲突的, 一个子目标性能的改善可能会引起另一个子目标性能的降低,

也就是说, 若要同时使这多个子目标都达到最优值是不可能的, 只能在它们中间进行协调和折中处理, 使各个子目标函数都尽可能地达到最优^[1],

【收稿日期】 2010-01-27

【基金项目】 国家自然科学基金青年基金资助项目(No. 10602055)

【作者简介】 张晓菲(1985-), 女, 山西运城人, 硕士研究生. 主要研究方向为智能优化算法, 测试计量技术及仪器.

这就是多目标优化问题的本质所在. 多目标优化同单目标优化^[2,3]不同, 多目标问题最优解集中的元素就全体目标而言是不可比较的, 一般称为 Pareto 最优解集^[4]. 所谓 Pareto 最优解集, 就是对于一些解不可能进一步优化某一个或几个目标而其他目标不至于劣化, 因此也称为非劣最优解集. 直到 1985 年才出现第一个多目标演化算法 (MOEA)—基于向量评估的遗传算法 (VEGA), 但这不是基于 Pareto 最优概念的演化算法. 1993 年到 1995 年, 多目标遗传算法 (MOGA)、非支配排序遗传算法 (NSGA)、小组决胜遗传算法 (NPGA) 的出现成为基于 Pareto 最优概念的多目标演化算法研究工作的奠基石^[5].

非支配排序遗传算法 (non-dominated sorting genetic algorithm, NSGA) 是由 Srinivas 和 Kalyanmoy^[6] 于 1995 年第一次提出的, 它首次将非支配排序的概念引入了多目标优化领域, 并取得了较好的效果. 但是 NSGA 本身存在许多不足之处, 主要在于: 1) 计算复杂性较高; 2) 缺乏精英策略; 3) 需要特别指定共享半径. 这使得它在处理高维、多模态等问题时, 难以得到满意的结果. 因此, Kalyanmoy 等人^[7] 提出了 NSGA-II 解决上述问题.

然而 NSGA-II 算法仍存在一些不足之处, 首先, 由于 NSGA-II 算法中的精英策略是将父代与子代合并, 从二倍的种群数量中选取最好解来保持优良种群, 这样容易陷入局部最优解. 为此, 本文提出了改进的精英策略, 来避免早熟. 此外, 对遗传算子进行了重新设计, 在多次试验的基础上采用模拟二进制交叉算子和非均匀变异算子, 使算法的性能更加稳定, 对进化前期的广泛搜索和进化后期的集中搜索进行了平衡, 从而也维护了种群的多样性.

1 多目标优化的基本概念

多目标优化问题一般可描述为下面的数学模型^[8]:

$$\begin{cases} \min y = f(x) = f_1(x), f_2(x), \dots, f_p(x) \\ e(x) = (e_1(x), e_2(x), \dots, e_m(x)) \geq 0 \\ s. t. \quad x \in (x_1, x_2, \dots, x_n) \in X \\ y \in (y_1, y_2, \dots, y_n) \in Y \end{cases} \quad (1)$$

式中: x —决策变量向量; y —目标函数向量; X —决策变量空间; Y —目标函数空间. 约束 $e(x) \geq 0$ 定义了可行域, 可行域 X_f 定义为: $e(x)$ 满足约束条件的决策变量 x 的集合 $X_f = \{x \in X \mid e(x) \geq 0\}$.

定义 1 Pareto 支配关系

设 p 和 q 是进化种群中的任意两个不同的个体, 若满足以下条件, 则称 p 支配 q ^[9].

1) 对每个子目标, p 不比 q 差, 即 $f_k(p) \leq f_k(q)$ ($k = 1, 2, \dots, r$);

2) 至少存在一个子目标, 使 p 比 q 好. 即 $\exists l \in \{1, 2, \dots, r\}$, 使 $f_l(p) < f_l(q)$.

其中, r 为子目标的数量. 此时称 p 为非支配的, q 为被支配的. 表示为 $p \succ q$, “ \succ ”表示支配的关系.

定义 2 Pareto 优化解及 Pareto 优化曲面(线)

决策变量 $x \in X_f$ 相对于集合 $A \subseteq X_f$ 被称为是非支配的, 当且仅当, $a \in A; a \succ x$, 如果相对于可行域 X_f 决策变量 x 是非支配的, 则称 x 是 Pareto 优化解^[9]. 所有的 Pareto 优化解构成了 Pareto 优化解集, 与之相对应的目标函数向量构成了 Pareto 优化曲面(线).

2 非支配排序遗传算法 NSGA-II

NSGA-II 算法在 NSGA 基础上, 主要做了下面几点改进: (1) 运用快速非支配排序法, 使计算复杂度由 $O(MN^3)$ 降为 $O(MN^2)$; (2) 采用拥挤度计算取代指定共享参数来进行分散性估计; (3) 引入精英策略, 防止优良解的丢失.

2.1 快速非支配排序

每个个体需要保存两个量^[10]:

1) 支配个数 n_i . 该量是在可行解空间中可以支配个体 P 的所有个体的数量.

2) 被支配个体集合 S_i . 该量是可行解空间中所有被个体 P 支配的个体组成的集合.

首先, 找到群体中所有 $n_i = 0$ 的个体, 将它们存入当前集 F_1 , 然后对于当前集 F_1 中的每个个体 j , 考察它所支配的个体集 S_i , 将集 S_i 中的每个个体 j 的支配它的个体数量 $n_j - 1$ (因为支配个体 j 的个体已经存入了当前集 F_1), 如果个体 j 的支配它的个体数量 $n_j = 0$, 则将个体 j 存入另一个集 H . 最后, 将得到的当前集 F_1 作为第一级非支配个体集. 继续对集 H 进行上述的分级操作, 直

到所有个体均被分级为止^[11].

2.2 拥挤度计算比较

定义拥挤度为在种群中给定点周围个体的密度,用 i_d 表示,它指出了在个体周围包含个体 i 本身但是不包含其他个体的最小的长方形^[12],如图1.

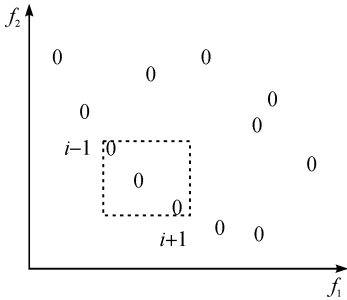


图1 个体 i 的拥挤度

Figure 1 Crowding distance of i

用 L 表示个体密度集合, $l = |L|$, l 为 L 中的个数拥挤度计算的伪代码如下:

对每一个个体 i ,令 $L[i]_d = 0$.

对每一个目标函数 m , $l = \text{sort}(L, m)$,令

$$L[1]_d = L[l]_d = \infty$$

从 $i = 2$ 到 $l - 1$ 循环,

$$L[i]_d = L[i]_d + (L[i+1]_m - L[i-1]_m)$$

其中, $L[i]_m$ 表示集合 L 中的第 i 个个体对于第 m 个目标函数的值. $\text{sort}(L, m)$ 指对每个目标函数值对个体进行排序.

在NSGA-II算法中,合并父代和子代种群,接着快速非支配排序完成后,需要重新选择新父种群.这时,先填充排序级别低的种群到新父种群中,直到某一级 F_i 时,超过了种群 N 的数量,这时需要对 F_i 的所有个体进行拥挤度计算,选择拥挤度距离大的个体优先填充.

3 精英策略与算法改进

3.1 精英策略的改进

NSGA-II采用的精英策略是从父种群和子种群组合的二倍种群中挑选最优个体,即使子代种群是经过交叉变异的,其优劣性是未知的,但父种群中始终保持交叉变异前的最优个体,经过父子合并挑选最优个体,从而将优良种群保留到下一代.

尽管这种精英策略可以有效地保留精英个体到进化结束,避免丢失进化过程中取得的最优解,但是NSGA-II中采用的精英策略存在不足.假定一个种群的大部分非支配解都在级别为1的非支配曲面(线)上,但是这个曲面(线)可能远离真正的Pareto最优曲面(线).若采用上述精英策略,很可能会使得种群中选择的非支配解全是精英,这样就会造成种群的提前收敛或者收敛于局部最优解.

为了避免算法早熟或陷入局部最优解,确保搜索方向朝着真正的Pareto最优曲面(线),本文提出了一种改进的精英策略,如图2.

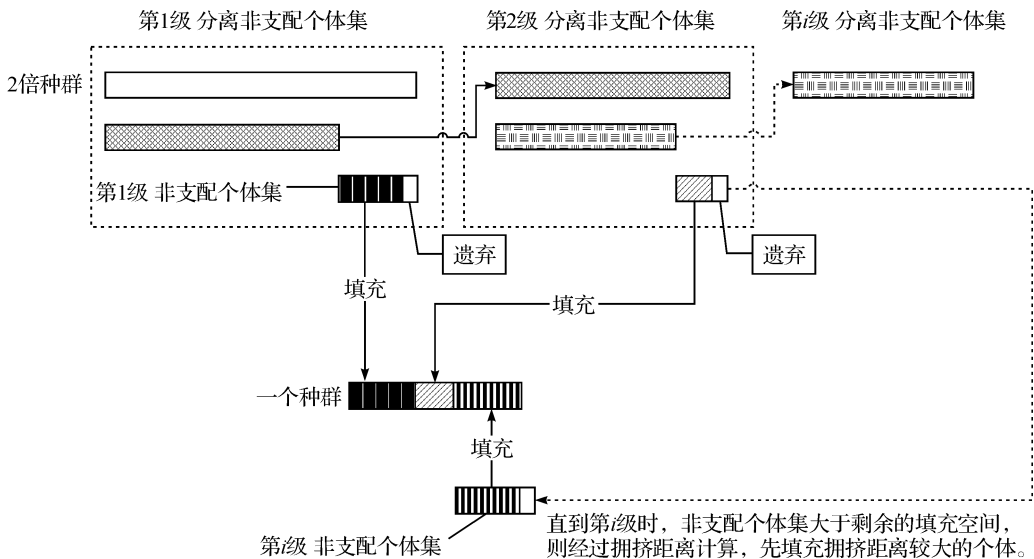


图2 精英策略改进图

Figure 2 Improved elitist strategy

精英策略体现在从父子种群(2popsize)挑选出下一代新父种群(popsize)的过程中. 从图中可以看出, 首先对二倍种群进行快速非支配排序, 从而得到第1级的非支配个体, 即 pareto 解, 这时不再对所有的 pareto 解填充到下一代种群中, 而是适当地“遗弃”个体. 本文采取将每一级的非支配个体集“遗弃”一个体的策略, 经过多级排序并进行填充后, 则一小部分精英被遗弃, 这样不但保证绝大部分精英个体留在下一代种群中, 也避免全部精英参与到下一代中, 从而可以防止种群提前收敛或陷入局部最优解.

3.2 交叉算子的选择

目前实数编码的多目标进化算法中, 最常用是算术交叉^[13]. 这里采用模拟二进制交叉算子(simulated binary crossover, 简称 SBX).

其定义为: 对两个父个体 x_1 和 x_2 , 按照以下方式生成两个子个体 c_1 和 c_2 :

$$\begin{cases} c_{1,i} = [(1 + \beta)x_{1,i} + (1 - \beta)x_{2,i}]/2 \\ c_{2,i} = [(1 - \beta)x_{1,i} + (1 + \beta)x_{2,i}]/2 \end{cases}, 1 \leq i \leq n \quad (2)$$

其中, β 为随机变量, 在每一维上都需要重新生成, 方法如下:

$$\beta = \begin{cases} (2\mu)^{\frac{1}{\eta}}, \mu \leq 0.5 \\ (2(1 - \mu))^{\frac{1}{\eta}}, \mu > 0.5 \end{cases} \quad (3)$$

其中, μ 是均匀分布于区间 $(0, 1)$ 上的随机数, η 是交叉参数, 为常数.

3.3 变异算子的选择

均匀变异操作取某一范围内均匀分布的随机数来替换原有基因值, 可使得个体在搜索空间内自由移动, 但不便于对重点区域进行局部搜索.

非均匀变异是将原有基因值作一随机扰动, 以扰动后的结果作为变异后的新基因值. 对每个基因座都以相同的概率进行变异运算后, 相当于整个解向量在解空间中作了一个轻微的变动. 随着遗传算法的运行, 非均匀变异可使得最优解的搜索过程更加集中在某一最有希望的重点区域中^[6].

在进行由 $X = x_1 x_2 \cdots x_k \cdots x_l$ 向 $X' = x_1 x_2 \cdots x'_k \cdots x_l$ 的非均匀变异操作时, 若变异点 x_k 处的基因值取值范围为 $[U_{\min}^k, U_{\max}^k]$, 则新的基因值 x'_k 由下式确定:

$$x'_k = \begin{cases} x_k + \Delta(t, y), \text{if random}(0, 1) = 0 \\ x_k - \Delta(t, y), \text{if random}(0, 1) = 1 \end{cases} \quad (4)$$

式中, $\Delta(t, y)$ (y 代表 $U_{\max}^k - \nu_k$ 和 $\nu_k - U_{\min}^k$) 表示 $[0, y]$ 范围内符合非均匀分布的一个随机数, 要求随着进化代数 t 的增加, $\Delta(t, y)$ 可按式定义:

$$\Delta(t, y) = y(1 - r^{(1-t/T)^b}) \quad (5)$$

式中: r 为 $[0, 1]$ 范围内符合均匀概率分布的一个随机数; T 为最大进化代数; b 是一个固定参数, 它决定了随机扰动对进化代数 t 的依赖程度.

4 算法 INSGA-II 的测试

以经典多目标函数为例, 对本文所提出的基于精英策略的改进非支配遗传算法 (INSGA-II) 进行测试, 下面分别给出算法流程, 以及算法的测试结果及比较分析.

4.1 INSGA-II 算法流程

算法步骤:

- 1) 随机初始化种群, 大小为 N ;
- 2) 将初始种群进行非支配排序, 并进行选择、交叉、变异操作^[14], 然后合并初始种群;
- 3) 将合并后的种群进行快速非支配排序, 并进行拥挤度计算;
- 4) 选取新父种群过程: 首先选取第一级非支配个体集 F_1 , 遗弃一个体后填充到新父种群中; 接着选取后面的集合, 当选取到 F_i 时, 新父种群中的个数超过 N 时, 选取拥挤度大的个体优先填充;
- 5) 对新父种群进行选择、交叉、变异操作, 判断是否满足停止准则, 若是, 退出算法, 输出结果, 否则转 6);
- 6) 将子代与父代进行合并, 转到 3).

流程图, 如图 3.

4.2 算例测试与分析

INSGA-II 采用参数为: 种群规模为 100; 联赛选择, 联赛规模为 2; 模拟二进制交叉, 交叉概率为 0.8, 交叉参数为 5; 非均匀变异, 变异概率为 0.05.

采用 INSGA-II 对五个经典多目标函数进行测试, 并与 NSGA-II 算法的结果进行比较, 分别给出各个函数的 INSGA-II 及 NSGA-II 算法的 Pareto 优化曲线图, 如下.

- 1) 函数 $F1$:

$$\begin{cases} f_1(x) = 1 - \exp\left(-\sum_{i=1}^n \left(x_i - \frac{1}{\sqrt{n}}\right)^2\right) \\ f_2(x) = 1 - \exp\left(-\sum_{i=1}^n \left(x_i + \frac{1}{\sqrt{n}}\right)^2\right) \end{cases} \quad -4 \leq x_1, x_2, x_3 \leq 4 \quad (6)$$

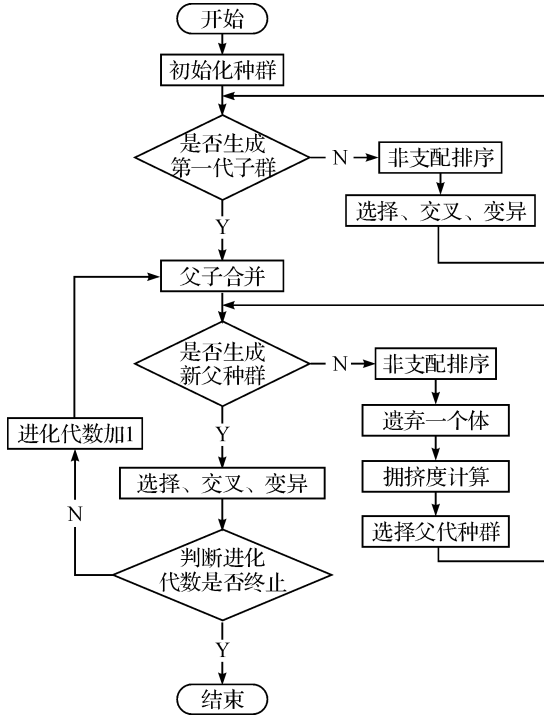


图 3 算法流程图

Figure 3 Flow chart of INSGA-II

函数 F_1 在两种算法下的 Pareto 优化曲线, 如图 4、图 5。

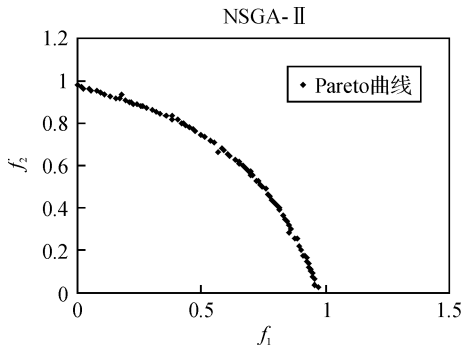


图 4 函数 F_1 采用 NSGA-II 得到的 Pareto 曲线图
Figure 4 Pareto front of function F_1 by NSGA-II

2) 函数 F_2 :

$$\begin{cases} f_1(x) = [1 + (A_1 - B_1)^2 + (A_2 - B_2)^2] \\ f_2(x) = [(x + 3)^2 + (y + 1)^2] \end{cases}$$

$$\begin{aligned} \text{其中 } A_1 &= 0.5\sin 1 - 2\cos 1 + \sin 2 - 1.5\cos 2 \\ A_2 &= 1.5\sin 1 - \cos 1 + 2\sin 2 - 0.5\cos 2 \\ B_1 &= 0.5\sin x - 2\cos x + \sin y - 1.5\cos y \\ B_2 &= 1.5\sin x - \cos x + 2\sin y - 0.5\cos y \end{aligned} \quad (7)$$

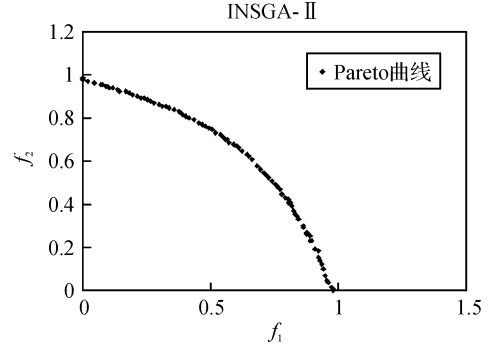


图 5 函数 F_1 采用 INSGA-II 得到的 Pareto 曲线图
Figure 5 Pareto front of function F_1 by INSGA-II

函数 F_2 在两种算法下的 Pareto 优化曲线, 如图 6 和图 7。

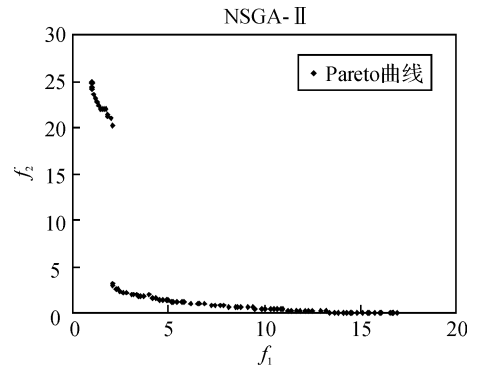


图 6 函数 F_2 采用 NSGA-II 得到的 Pareto 曲线图
Figure 6 Pareto front of function F_2 by NSGA-II

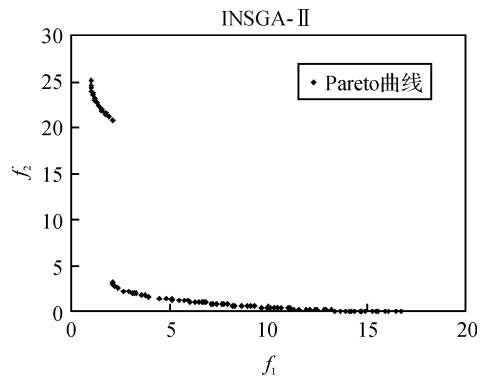


图 7 函数 F_2 采用 INSGA-II 得到的 Pareto 曲线图
Figure 7 Pareto front of function F_2 by INSGA-II

3) 函数 F3:

$$\begin{cases} f_1(x) = \sum_{i=1}^{n-1} (-10 \exp(-0.2 \sqrt{x_i^2 + x_{i+1}^2})) \\ f_2(x) = \sum_{i=1}^n (|x_i|^{0.8} + 5 \sin(x_i)^3) \\ -5 \leq x_1, x_2, x_3 \leq 5 \end{cases} \quad (8)$$

函数 F3 在两种算法下的 Pareto 优化曲线, 如图 8 和图 9.

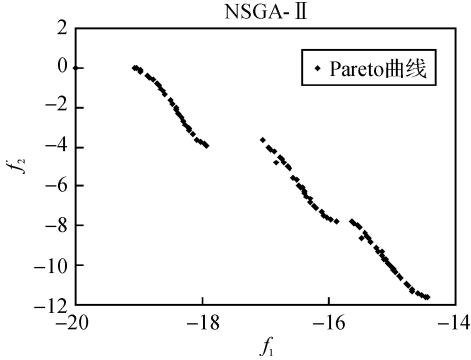


图 8 函数 F3 采用 NSGA-II 得到的 Pareto 曲线图
Figure 8 Pareto front of function F3 by NSGA-II

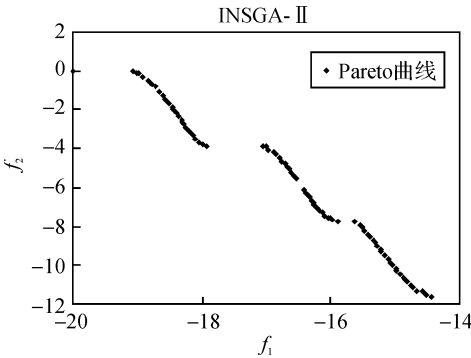


图 9 函数 F3 采用 INSGA-II 得到的 Pareto 曲线图
Figure 9 Pareto front of function F3 by INSGA-II

4) 函数 F4:

$$\begin{cases} f_1(x) = x_1 & 0 \leq x_1 \leq 1 \\ f_2(x) = g \left(1 - \sqrt{\frac{x_1}{g}} \right) & -5 \leq x_2, \dots, x_{10} \leq 5 \end{cases}$$

其中

$$g(x) = 91 + \sum_{i=2}^{10} (x_i^2 - 10 \cos(4\pi x_i)) \quad (9)$$

函数 F4 在两种算法下的 Pareto 优化曲线, 如图 10 和图 11.

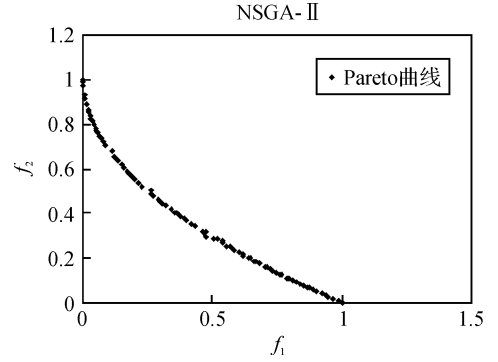


图 10 函数 F4 采用 NSGA-II 得到的 Pareto 曲线图
Figure 10 Pareto front of function F4 by NSGA-II

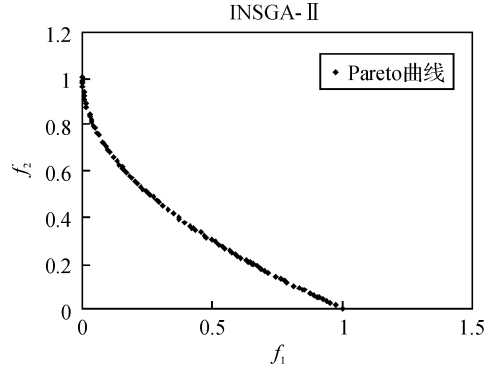


图 11 函数 F4 采用 INSGA-II 得到的 Pareto 曲线图
Figure 11 Pareto front of function F4 by INSGA-II

5) 函数 F5:

$$\begin{cases} f_1(x) = 1 - \exp(-4x_1) \sin^6(6\pi x_1) \\ f_2(x) = g(1 - (f_1/g)^2) \\ 0 \leq x_i \leq 1 \quad i = 1, \dots, 10 \end{cases}$$

其中
$$g(x) = 1 + 9 \left(\sum_{i=2}^{10} x_i/9 \right)^{0.25} \quad (10)$$

函数 F5 在两种算法下的 Pareto 优化曲线, 如图 12 和图 13.

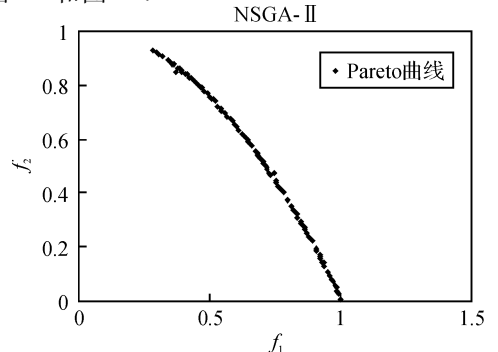


图 12 函数 F5 采用 NSGA-II 得到的 Pareto 曲线图
Figure 12 Pareto front of function F5 by NSGA-II

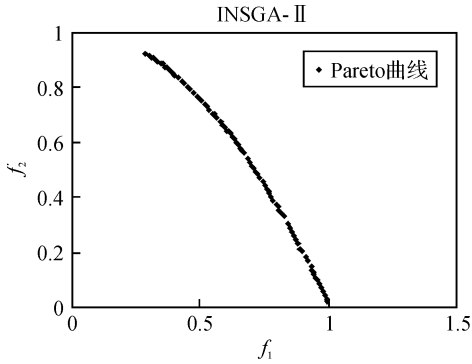


图13 函数F5采用INSGA-II得到的Pareto曲线图
Figure 13 Pareto front of function F5 by INSGA-II

以上5组图是将五个函数分别采用两种算法得到的,可以明显看出,采用INSGA-II算法得到的pareto曲线比起NSGA-II得到的pareto曲线的分布更加均匀。

5 结 语

本文针对NSGA-II算法做出三方面改进:改进精英策略,在选取精英过程中,遗弃一小部分精英个体,可避免算法因精英的快速集中陷入局部最优解;使用模拟二进制交叉算子,从而提高了算法的搜索性能;变异算子采取了非均匀分布变异,从而更好地维持了种群的多样性.将改进后的INSGA-II算法对五个典型的多目标优化进行测试实验,并与NSGA-II算法的结果进行比照,结果表明改进后的算法的pareto曲线分布均匀,算法更加稳定.因此,INSGA-II是一种有效的非支配排序遗传算法。

【参 考 文 献】

[1] 周 明, 孙树栋. 遗传算法原理及应用[M]. 北京: 国防工业

出版社, 1999: 45-54.

- [2] 张火明, 陆慧娟, 卫 伟. 混合离散变量模拟退火方法及其应用[J]. 中国计量学院学报, 2006, 17(1): 44-49.
- [3] 高明正, 金尚忠, 张火明. 改进的混合粒子群优化算法[J]. 中国计量学院学报, 2008, 19(3): 260-264.
- [4] MURUGAN P, KANNAN S, BASKAR S. NSGA-II algorithm for multi-objective generation expansion planning problem[J]. Electric Power Systems Research, 2009, 79: 622-628.
- [5] 谢 涛, 陈火旺. 多目标优化与决策问题的演化算法[J]. 中国工程科学, 2002(2): 59-68.
- [6] SRINIVAS N, KALYANMOY D. Multiobjective optimization using nondominated sorting in genetic algorithms[J]. Evolutionary Computation, 1994, 2(3): 221-248.
- [7] KALYANMOY D, AGRAWAL S, PRATAB A, et al. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. KanGAL Report 200001[R]. Indian: Indian Institute of Technology, 2000.
- [8] SUN Y J, SHEN G Z. Improved NSGA-II multi-objective genetic algorithm based on hybridization-encouraged mechanism[J]. Chinese Journal of Aeronautics, 2008, 21: 540-549.
- [9] 关志华, 寇纪淞, 李敏强. 一种改进的非支配排序遗传算法 INSGA[J]. 天津大学学报, 2002, 35(4): 429-434.
- [10] DAVID A, VAN V, GARY B L. Multi-objective evolutionary algorithms: analyzing the state-of-the-Art[J]. Evolutionary Computation, 2000, 8(2): 125-147.
- [11] DEB K, AGRAWAL R B. Simulated binary crossover for continuous search space[J]. Complex Systems, 1995, 9(4): 115-148.
- [12] BÄCK T, SCHWEFEL H P. An overview of evolutionary algorithms for parameter optimization[J]. Evolutionary Computation, 1993, 1(1): 1-23.
- [13] 张 敏, 罗文坚, 王煦法. 一种基于正态分布交叉的 ϵ -MOEA[J]. 软件学报, 2009, 20(2): 305-314.
- [14] 张火明, 高明正, 张晓菲. 基于交叉变异的连续蚁群算法[J]. 中国计量学院学报, 2009, 20(3): 259-262.