

文章编号:1671-6833(2020)04-0023-05

基于混合策略的快速非支配排序算法 II

张茂清¹, 汪 镭¹, 崔志华², 郭为安³

(1.同济大学 电子与信息工程学院,上海 201804; 2.太原科技大学 计算机科学与技术学院,山西 太原 030024; 3.同济大学 中德工程学院,上海 201804)

摘 要:快速非支配排序算法 II(fast non-dominated sorting algorithm II, NSGA-II)是经典多目标优化算法。然而,其采用的锦标赛策略存在重复选择交叉个体的缺陷,导致后代个体多样性降低。为解决此问题,提出两种改进策略:第一,引入 Lévy 分布。Lévy 分布具有同时平衡局部搜索和全局搜索的能力。通过将 Lévy 分布引入到执行交叉操作的父代个体,可增加发现父代个体周围潜在较优个体的概率。第二,引入三交叉个体策略。一般的两个交叉个体存在来自同一个体的可能性,引入三交叉个体可以明显降低重复选择父代个体的现象。大量实验结果表明,所提策略可有效改进 NSGA-II 的整体性能。

关键词:NSGA-II;多目标优化;锦标赛策略;混合策略;种群多样性

中图分类号:TP391 **文献标志码:**A **doi:**10.13705/j.issn.1671-6833.2020.04.007

0 引言

在实际工程应用中,有许多目标函数为 2~3 个的优化问题^[1-3],且目标函数间具有彼此冲突的特点,此类问题一般被认为 NP-Hard 问题。传统的方法,如线性规划、梯度下降等,由于对问题特征具有较为严格的要求,导致解决此类问题时付出极为昂贵的时间代价,甚至在有限时间内无法获得满意解。进化算法的出现为此类问题提供了新思路。典型算法如 NSGA-II^[4]、SPEAII(improved strength pareto evolutionary algorithm)^[5],以及 NNIA(multi-objective immune algorithm with non-dominated neighbor-based selection)^[6]等在此类问题上均表现突出。其中,NSGA-II 作为多目标优化算法的典型代表,受到很多研究者的广泛关注。根据研究角度不同,可将近些年研究成果作如下分类。

就应用角度而言,NSGA-II 及其改进方法已经被应用到许多实际优化问题中。为解决服装调度生产中最大完成时间和最小延期交货时间的问题,陆金芳^[7]提出改进排序适应度和按需分层策略,极大提高了服装调度的效率。黄敏镁等^[8]为了最大化供应链环境下协商 Agent 自身效用和合

作企业建议相似度,采用了正整数和小数混合的实数编码方式,并在遗传操作中增加了约束限制,以剔除算法运行中产生不可行个体。在应急物流系统设计中,需要综合考虑系统总成本、总耗时以及道路安全性等问题。陈刚等^[9]针对此问题特点提出了改进个体交叉和变异方式,并进一步将精英策略融入 NSGA-II。在将 NSGA-II 应用于水污染修复管理模型时,NSGA-II 不能有效地收敛到真实 Pareto 前沿。为提升 NSGA-II 收敛性,宋健^[10]提出将 HCS(hill climber with step)融入 NSGA-II,极大地提升了算法收敛能力。

在理论研究方面,NSGA-II 也得到了极大的改进。为解决 NSGA-II 拥挤度距离机制无法有效区分多样性个体的缺陷,崔志华等^[11]提出了基于平均距离聚类的多样性评价指标,并进一步提出了基于平均距离聚类的 NSGA-II。受无免费午餐定理启发,陈辅斌等^[11]提出将免疫平衡原理引入 NSGA-II 选择策略。为解决 NSGA-II 中拥挤度距离机制无法衡量个体周围个体密度的缺陷,王祥^[12]提出将密度聚类思想融入到个体拥挤度评价机制,并进一步提出了个体邻域的构建方法。汪文文等^[13]将禁忌搜索的思想融入精英保留策略,有效地平衡了全局搜索和局部搜索。为了拓展 NSGA-II 在高纬多

目标优化问题上的性能,Yang 等^[14]提出利用基于网格支配的方法区分个体,并进一步利用网格支配产生后代。同样针对此类问题,Zhang 等^[15]提出将分解的思想引入到多目标优化问题中,可有效避免 Pareto 支配失效的问题。

虽然 NSGA-II 在理论和应用方面已经取得了很大的进步,但是采用的锦标赛策略会导致重复父代个体的产生,并由此导致后代多样性受到影响。为解决此问题,笔者从以下两方面进行改进,第一,引入 Lévy 分布,增加发现父代个体周围潜在较优个体的能力;第二,提出三交叉父代策略,进一步降低重复父代个体对后代多样性的影响。最后提出基于混合策略的 NSGA-II (fast non-dominated sorting genetic algorithm II based on hybrid strategies, HSNSGA-II)。

1 基本概念以及锦标赛选择策略

1.1 基本概念

一个典型多目标优化问题可形式化表示如下^[1]:

$$\begin{aligned} \min F(\mathbf{X}) &= [f_1(\mathbf{X}), f_2(\mathbf{X}), \dots, f_M(\mathbf{X})], \quad (1) \\ \text{s.t. } \mathbf{X} &\in \Omega, \end{aligned}$$

式中: $\mathbf{X} = (x_1, x_2, x_3, \dots, x_D)$ 是一个 D 维决策向量; $\Omega \subseteq \mathbf{R}^n$ 为决策空间; M 为目标函数数量。由于多目标优化问题中不同目标间具有相互冲突的特点,导致不存在最优个体,而是最优解集。以下为多目标优化问题中的主要概念。

定义 1: 若变量 \mathbf{X} 目标函数 $f(\mathbf{X}) = (f_1(\mathbf{X}), f_2(\mathbf{X}), \dots, f_M(\mathbf{X}))^T$, 变量 \mathbf{X}' 目标函数 $f(\mathbf{X}') = (f_1(\mathbf{X}'), f_2(\mathbf{X}'), \dots, f_M(\mathbf{X}'))^T$, 当且仅当对于 $\forall i \in \{1, 2, \dots, M\}, f_i(\mathbf{X}) \leq f_i(\mathbf{X}')$ 成立, 且存在 $k \in \{1, 2, \dots, M\}$, 使得 $f_k(\mathbf{X}) < f_k(\mathbf{X}')$ 严格成立, 则称 \mathbf{X} 支配 \mathbf{X}' , 记作: $\mathbf{X} < \mathbf{X}'$ 。

定义 2: 决策空间上所有 Pareto 最优解构成的集合称为 Pareto 最优解集。

定义 3: Pareto 最优解集在目标空间上对应解集合称 Pareto 前沿面。

1.2 基本 NSGA-II 框架以及缺陷分析

作为多目标优化领域里的典型代表算法, NSGA-II 有两个核心策略, 非支配排序和拥挤度距离。非支配排序用于强化种群个体间的选择压力, 拥挤度距离用于评价个体的多样性, 通过上述两种策略达到种群个体不断进化。种群更新主要过程可简述如下。

设 P 和 Q 分别为具有 N 个个体的父代种群

和对应的子代种群。首先, 将两种种群合并为 $C = P \cup Q$ 。为从合并的种群 C 中选择出 N 个后代个体, 利用非支配排序策略对种群 C 进行操作, 可得到多个 Pareto 前沿面。然后, 从第 1 层前沿面开始, 累积计算个体数量, 直到第 l 层个体数量首次超过 N 。为从第 l 层选择出较优个体, 利用拥挤度距离计算第 l 层个体中每个个体的拥挤度, 选择拥挤度距离大的个体作为下一代个体。

NSGA-II 中交叉操作和变异操作为常见操作, 在此不再赘述。需要指出的是, NSGA-II 中选择操作所用策略为锦标赛策略, 描述如下:

- (1) 确定每次选择个体数量, 在此为 2 个。
- (2) 从种群中随机选择个体, 选择适应度值较优个体作为后代个体。
- (3) 重复上述步骤(2), 直到达到预定个体数量。

从上述步骤可以看出, 若个体 I 为第 1 层 Pareto 前沿面上个体, 且具有较优多样性指标, 则其在下次被选中的概率依然很大。

图 1 展示了采用锦标赛策略选择父代个体重复个体数量统计数据。采用测试函数为 ZDT2, 种群数量为 50, 关于测试函数详细信息在后续实验部分详细阐述。从上述统计数据中可以看出, 每代都有平均 15 个个体的重复量, 最高甚至达到 32 个重复个体。这些重复个体虽然携带了较优的基因, 可以为后代个体提供较好的搜索方向, 但是也造成了后代多样性较差的缺陷。

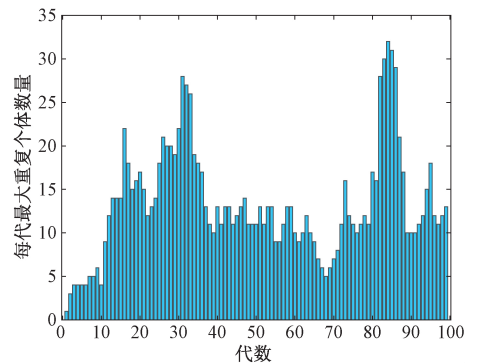


图 1 重复个体数量统计

Figure 1 Statistics of repeated individuals

2 基于混合策略的 NSGA-II

针对 NSGA-II 上述缺陷, 笔者提出引入 Lévy 分布策略和三交叉父代策略。下面简要介绍 Lévy 分布策略, 然后详细介绍本文改进算法。

Lévy 分布概率密度函数可形式化表示如下:

$$L(\delta) \sim u \cdot t^{-1-\delta}, \quad 0 < \delta < 2, \quad (2)$$

其简化形式可表示如下:

$$L(\delta) \sim \phi \cdot u / |v|^{1/\delta}, \quad (3)$$

其中, u 和 v 是符合高斯分布的随机数, δ 设置为 $1.5^{[13]}$, ϕ 定义如下:

$$\phi = \left(\frac{\Gamma(1+\delta) \cdot \sin(\pi \cdot \delta/2)}{\Gamma\left(\frac{1+\delta}{2}\right) \cdot \delta \cdot 2^{(\delta-1)/2}} \right)^{1/\delta}. \quad (4)$$

图 2 展示了 Lévy 分布所生成的 100 个采样点取值。统计一下 Lévy 分布取值范围, 可发现, 93% 的取值落在 $[-1.8, 1.8]$ 内。从图 2 可以看出, Lévy 分布不仅可以使算法搜索聚焦于个体局部区域, 也可使搜索跳出局部范围, 增加后代多样性, 避免陷入局部最优。

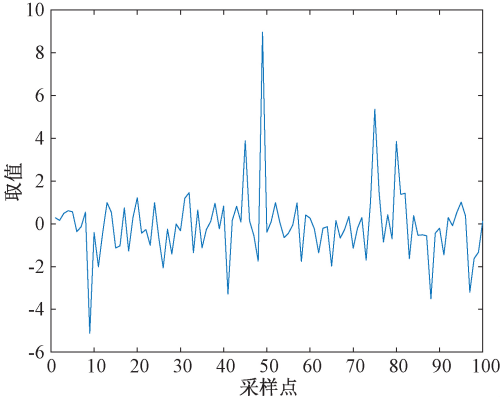


图 2 Lévy 分布采样点

Figure 2 Sampling points with Lévy distribution

设 P_1 、 P_2 、 P_3 分别为锦标赛选择策略所选出父代个体。原交叉算子可定义如下:

$$Q_1 = \frac{(P_1 + P_2)}{2} + \text{beta} \cdot \frac{(P_1 - P_2)}{2}, \quad (5)$$

其中, beta 是 NSGA-II 中定义参数, 请参考文献 [4]。

改进后交叉算子可表示如下:

$$Q_1 = \frac{(P_1 + P_2 + P_3)}{3} + L \cdot \text{beta} \cdot \frac{(P_1 - P_2 - P_3)}{3}, \quad (6)$$

式中: L 即为上文所述 L 分布函数。从上式可以看出, Lévy 起到扰动作用, 可以极大增加发现父代个体周围潜在较优个体的概率, 同时引入 P_3 则可进一步减小父代个体为同一个个体的概率。

基于混合策略的 NSGA-II 伪代码如下。

- 1: 初始化种群以及相关参数
- 2: While (是否满足结束条件) do
- 3: 快速非支配排序
- 4: 采用锦标赛策略, 从种群中选择较优个体
- 5: 采用式 (6) 对父代个体执行交叉操作
- 6: 对个体执行变异操作
- 7: 环境选择, 更新种群
- 8: End while
- 9: 输出种群

3 实验结果及分析

3.1 参数设置

为综合测试笔者所提算法性能, 将 HSNSGA-II 传统算法 SPEA2^[5]、PEASII^[16]、NNIA^[6] 以及最近提出的算法 MOEAIGDNS (multi-objective evolutionary algorithm based on enhanced IGD)^[17] 和 ADCNSGA-II (NSGA-II with average distance clustering)^[1] 进行对比。注意, 所有参数设置均按照原始参考文献设置, 有兴趣读者请参阅文献 [5-6, 16-17]。实验所用计算机为 Inter Core i 5-2400 3.10 GHz CPU, 6.00 GB 内存, Windows7 操作系统, 运行环境为 MATLAB7.9。每个算法独立运行 20 次, 对 ZDT 测试函数集最大迭代 100 次; 对 DTLZ1 函数最大迭代 700 次; 对 DTLZ2、DTLZ4 和 DTLZ5 函数迭代 250 次; 对 DTLZ3 函数迭代 1 000 次; 种群个体为 50。

采用 ZDT 测试函数集^[18], 这些函数具有凸、凹、连续、非连续和具有多重局部最优等特点; 评价指标采用 $IGD^{[19]}$ 。 IGD 是一个综合指标, 可同时评价算法收敛性和多样性, 其值越小表示算法性能越优, 定义为:

$$IGD(P, P^*) = \frac{\sum_{v \in P^*} \text{dist}(v, P)}{|P^*|}, \quad (7)$$

式中: P^* 为目标空间真实 Pareto 前沿面上均匀分布点的集合; P 为算法所求解集; $\text{dist}(v, P)$ 为点 v 和集合 P 中点的最小欧几里得距离。

3.2 算法对比及分析

表 1 列出了笔者所提算法和对比算法在测试函数上的实验结果, 每个结果为算法运行 20 次的 IGD 的均值与方差, 最优结果用黑体显示。从上述实验结果可以看出, 与传统算法相比, HSNSGA-II 在 ZDT1、ZDT2、ZDT3、ZDT4、ZDT6、DTLZ2、DTLZ5 上表现均较为突出, 仅在 DTLZ1、DTLZ3、DTLZ4 上表现较差。相比于最近提出的算法

MOEAIGDNS 和 ADCNSGA-II,可以看出 HSNSGA-II 仍然具有较大的优势。因此,综合上述实验结果,可以看出笔者所提策略明显地提高了 NSGA-II 的性能。

表 1 实验结果对比
Table 1 Comparison of experimental results

函数	类别	HSNSGA-II	NSGA-II	NNIA	PESA-II	SPEA2	MOEAIGDNS	ADCNSGA-II
ZDT1	均值	1.745 7e-1	7.100 5e-1	6.174 6e+0	1.037 0e+0	1.995 2e+0	1.710 2e+0	1.985 3e-1
	方差	1.13e-1	5.10e-1	3.49e+0	6.90e-1	1.11e+0	6.59e-1	3.41e-1
ZDT2	均值	6.014 3e-1	1.050 3e+0	1.274 6e+1	9.532 1e-1	1.642 4e+0	1.678 0e+0	6.047 5e-1
	方差	5.97e-2	2.50e-1	5.05e+0	9.14e-2	7.13e-1	5.52e-1	6.21e-2
ZDT3	均值	1.992 5e-1	5.095 5e-1	6.636 5e+0	1.071 5e+0	7.253 1e-1	1.058 7e+0	1.992 8e-1
	方差	1.12e-1	1.41e-1	3.02e+0	4.24e-1	3.27e-1	3.76e-1	2.18e-1
ZDT4	均值	2.276 0e-1	4.809 5e-1	1.026 0e+0	8.995 2e-1	7.858 6e-1	8.859 0e-1	2.841 1e-1
	方差	1.42e-1	1.60e-1	6.80e-1	2.58e-1	3.99e-1	3.85e-1	1.04e-1
ZDT6	均值	6.995 9e-2	2.184 0e-1	3.942 7e-1	1.228 1e-1	1.837 6e-1	6.391 1e-1	7.124 2e-1
	方差	8.16e-2	4.65e-2	1.26e-1	1.35e-2	1.32e-1	1.21e-1	5.17e-2
DTLZ1	均值	2.077 5e-1	4.533 5e-3	4.720 1e-3	9.309 3e-3	3.914 6e-3	3.973 1e-3	3.864 1e-3
	方差	1.86e-1	3.20e-5	1.40e-4	1.83e-3	2.01e-4	3.41e-4	1.94e-1
DTLZ2	均值	5.201 1e-3	1.021 8e-2	1.069 4e-2	1.882 7e-2	8.568 2e-3	8.170 7e-3	4.864 2e-3
	方差	2.84e-4	3.97e-4	8.68e-4	4.30e-3	1.61e-4	1.12e-4	5.21e-4
DTLZ3	均值	8.686 3e+0	1.143 5e-2	1.259 5e-2	2.185 4e-2	9.032 0e-3	1.193 6e-2	8.147 2e-1
	方差	5.00e+0	1.08e-3	3.71e-4	9.84e-4	3.47e-4	1.12e-4	4.14e-1
DTLZ4	均值	1.524 5e-1	1.159 0e-2	1.039 4e-2	3.803 0e-1	3.753 7e-1	7.420 9e-1	1.072 4e-2
	方差	3.30e-1	2.05e-4	9.05e-6	5.12e-1	5.19e-1	6.89e-8	3.01e-1
DTLZ5	均值	5.071 6e-3	1.117 9e-2	1.098 0e-2	1.723 8e-2	8.437 5e-3	8.151 2e-3	5.284 6e-3
	方差	1.90e-4	1.02e-3	5.34e-4	8.56e-4	8.32e-6	1.19e-5	2.01e-4

图 3 展示了 HSNSGA-II 和 NSGA-II 在 ZDT3 测试函数上的实验结果对比。从图 3 可以看出, HSNSGA-II 可以搜索到右下角区域,而标准 NSGA-II 却无法有效搜索到该区域,说明所提混合策略可以提高后代个体的多样性并达到了预期,进一步验证了笔者所提策略的有效性。

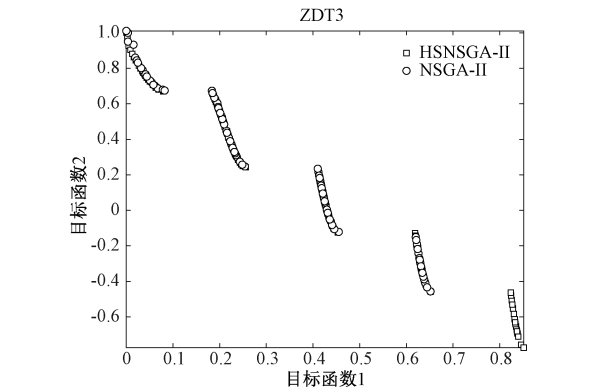


图 3 在 ZDT3 函数上实验结果对比
Figure 3 Comparison of experimental results on ZDT3

4 结论

NSGA-II 是多目标优化领域较为经典算法,然而其采用的锦标赛选择策略可导致重复选择父代个体,进而导致后代多样性受到影响。为解决此问

题,笔者提出融合 Lévy 分布和三交叉父代的策略,第一个策略可有效强化搜索父代周围潜在个体的能力,第二个策略可进一步降低所选父代为同一个体的现象。通过实验对比,验证了笔者所提算法在多个测试集上的有效性。

参考文献:

[1] 崔志华,张茂清,常宇,等.基于平均距离聚类的 NSGA-II[J/OL].自动化学报:1-12(2019-04-11)[2019-12-20]. <https://doi.org/10.16383/j.aas.c180540>.
[2] 蒋佩华,华冰,黄宇,等.基于遗传算法的变质量航天器姿态控制方法[J].郑州大学学报(工学版),2019,40(4):1-7.
[3] 龙志伟,肖松毅,王晖,等.基于粒子群算法的水资源需求预测[J].郑州大学学报(工学版),2019,40(4):32-35,47.
[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] ZITZLER E, LAUMANN S M, THIELE L. SPEA2: improving the strength pareto evolutionary algorithm[C]//Proceedings of Conferences on Evolutionary Methods for Design Optimization and Control with Ap-

- plications to Industrial Problems. Barcelona: CIMNE, 2001: 5–10.
- [6] GONG M G, JIAO L C, DU H F, et al. Multiobjective immune algorithm with nondominated neighbor-based selection[J]. Evolutionary computation, 2008, 16(2): 225–255.
- [7] 陆金芳.改进的 NSGAII 算法在服装企业生产调度中的应用研究[D].广州:暨南大学,2018.
- [8] 黄敏镁,袁际军,曹亮.基于 NSGAII 的协同产品开发项目自动协商决策[J].运筹与管理,2017, 26(3):86–91,99.
- [9] 陈刚,付江月.基于 NSGAII 的应急物流多目标 LRP 研究[J].软科学,2016,30(4):135–139.
- [10] 宋健.多目标遗传算法的改进及其在地下水污染修复管理中的应用[D].南京:南京大学,2017.
- [11] 陈辅斌,李忠学,杨喜娟.基于改进 NSGA2 算法的多目标柔性作业车间调度[J].工业工程,2018, 21(2):55–61.
- [12] 王祥.基于改进 NSGA-II 算法的应急物资模块化调度问题建模与求解[D].合肥:合肥工业大学,2018.
- [13] 汪文文,方玺,何朗,等.NSGA-II 算法的改进及其在应急管理中的应用[J].计算机工程与应用, 2018, 54(16):241–247.
- [14] YANG S X, LI M Q, LIU X H, et al. A grid-based evolutionary algorithm for many-objective optimization[J]. IEEE transactions on evolutionary computation, 2013, 17(5):721–736.
- [15] ZHANG Q F, LI H. MOEA/D: a multiobjective evolutionary algorithm based on decomposition[J]. IEEE transactions on evolutionary computation, 2007, 11(6):712–731.
- [16] CORNE D, JERRAM N, KNOWLES J, et al. PESA-II: region-based selection in evolutionary multiobjective optimization[C]//Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001). San Francisco: Morgan Kaufman, 2001:283–290.
- [17] TIAN Y, ZHANG X Y, CHENG R, et al. A multi-objective evolutionary algorithm based on an enhanced inverted generational distance metric[C]//2016 IEEE Congress on Evolutionary Computation (CEC). Vancouver: IEEE, 2016: 24–29.
- [18] ZITZLER E, DEB K, THIELE L. Comparison of multiobjective evolutionary algorithms: empirical results[J]. Evolutionary computation, 2000, 8(2):173–195.
- [19] COELLO C A C, CORTES N C. Solving multiobjective optimization problems using an artificial immune system[J]. Genetic programming and evolvable machines, 2005, 6(2):163–190.

Fast Non-dominated Sorting Genetic Algorithm II Based on Hybrid Strategies

ZHANG Maoqing¹, WANG Lei¹, CUI Zhihua², GUO Weian³

(1.School of Electronics and Information, Tongji University, Shanghai 201804, China; 2.School of Computer Science and Technology, Taiyuan University of Science and Technology, Shanxi 030024, China; 3.Sino-Germany College of Applied Sciences, Tongji University, Shanghai 201804, China)

Abstract: Fast Non-dominated Sorting Algorithm II (NSGA-II) was an classical multi-objective optimization algorithm, However, the tournament selection strategy in NSGA-II had the drawback of repeated selection of the same individuals, resulting in the low diversity of offspring population. To tackle this problem, this paper proposed two strategies. The first one was to incorporate Lévy distribution. Lévy distribution had the ability of balancing the local and global search. Incorporating Lévy distribution to parent individuals to do crossover operator could increase the probability of discovering potential better individuals around patent individuals. The second one was to introduce tri-crossing crossover strategy. In general, two parent individuals had the possibility of coming from the same individual. The introduction of tri-crossing individuals could obviously reduce the phenomenon of repeated selection of parent individuals. Extensive experiments demonstrated that the proposed method could efficiently improve the overall performance of NSGA-II.

Key words: NSGA-II; multi-objective optimization; tournament selection strategy; hybrid strategies; population diversity