

# 一种非均匀分布问题分布性维护方法

李密青, 郑金华, 李 珂

(湘潭大学信息工程学院, 湖南湘潭 411105)

**摘 要:** 几乎所有多目标进化算法(multi-objective optimization evolutionary algorithm, MOEA)都是针对 Pareto 最优面为均匀分布问题而言. 然而现实中很多问题 Pareto 最优面是非均匀分布的, 决策者希望得到一个与 Pareto 最优面分布类似的解集. 现存算法并不能有效解决该问题. 对此, 提出一种针对于非均匀分布多目标优化问题的维护方法(non-uniformly diversity maintenance method, NUDMM). 该方法定义一个反映个体分布“规则”程度的指标——杂乱度, 并设计一种降低种群杂乱度的方法, 在未知 Pareto 最优面分布规律情况下有效剔除造成种群混乱的个体. 通过与 NSGA-II 和 SPEA2 在不同维数下 8 个非均匀函数上对比实验, 表明 NUDMM 在有效保持问题真实分布的同时, 具有良好的收敛性.

**关键词:** 多目标优化; 多目标进化算法; 非均匀分布; 分布性维护; 测试函数; 杂乱度

**中图分类号:** TP18 **文献标识码:** A **文章编号:** 0372-2112 (2011) 04-0946-07

## A Diversity Maintenance Method for Non-Uniform Distribution Problem

LI Mi-qing, ZHENG Jin-hua, LI Ke

(Institute of Information Engineering, Xiangtan University, Xiangtan, Hunan 411105, China)

**Abstract:** Almost all of the multi-objective optimization evolutionary algorithms(MOEAs) are designed for the Pareto optimal front which is distributed uniformly. But in real world optimizations, the Pareto optimal front usually has a non-uniform distribution. A similar solution set distribution with Pareto optimal front is expected to obtain by decision makers. However, the existing algorithms cannot solve such problems effectively. In this paper, a diversity maintenance method for non-uniformly distributed multi-objective optimization problem(NUDMM) is proposed. In the algorithm, an indicator reflecting ‘regular’ degree of distribution-Messy is defined. And a method to decrease Messy of population is designed, which eliminates disordered individual on the condition that the distribution of the Pareto optimal front is unknown. From an extensive comparative study with NSGA-II and SPEA2 on eight non-uniform distribution test problems, it is observed that the proposed method has a good performance in maintaining the real distribution and convergence.

**Key words:** multi-objective optimization; multi-objective evolutionary algorithms; diversity maintenance; non-uniform distribution; test problem, messy

## 1 引言

现实世界的很多问题通常由多个目标组成, 解决多目标优化问题(multi-objective optimization problem, MOP)是很困难的. 这些目标往往是互相冲突, 它不同于单目标优化问题(single objective optimization problem, SOP)那样只有一个最优解, 而是只能在相互冲突的目标间做出折衷, 得到一组非劣解集<sup>[1,2]</sup>.

近几十年来, 越来越多的研究者将进化算法应用到 MOP 中. 自 80 年代中期, David Schaffer 在他的博士论文中第一次将进化算法应用于解多目标优化问题, 提出了向量评估遗传算法 VEGA<sup>[3]</sup>以来, MOEA 的发展经历了

三个阶段, 出现了大量有效的算法<sup>[4,2]</sup>.

在 MOP 中, 许多问题真实 Pareto 最优边界(True Pareto optimal front, True-POF)可能很大, 也可能包含无穷多个解, 列出所有这些解是不现实的, 通常选择一组能理想代表 True-POF 的子集(Pareto optimal front, POF)作为研究对象. 在 MOEA 中, 特别是进化后期非支配解集通常要大于种群规模. 这时, 我们需要得到一组能代表整个目标空间的有限非支配子集, 供决策者选择. 为此, 研究者们提出了很多有效策略. 其中, Goldberg 等<sup>[5]</sup>的适应度共享技术是最早提出的分布性维护策略, Deb 等<sup>[6]</sup>的聚集距离, Come 等<sup>[7]</sup>的网格方法和 Zitzler 等<sup>[8]</sup>的密度估计为四种经典维护策略.

然而,以上这些方法都是针对 POF 为均匀分布问题而言的,都是把得到一组均匀分布的解集作为基本目标.事实上,现实中很多应用问题的 POF 是非均匀分布的,决策者对在目标空间中不同的区域的解有不同的偏好,这就要求算法能得到一个与 POF 分布类似的解集.对此学者们也进行了一些相关研究:Fonseca 等<sup>[9]</sup>于 1995 年首先提到了非均匀分布的概念;Deb 等<sup>[10]</sup>提出了构造非均匀测试函数的方法,并构造了两个著名的非均匀测试函数<sup>[10,11]</sup>;Pedersen 等<sup>[12]</sup>根据决策者的需要把解集分布分为均匀和非均匀两种情况,但只讨论了前者;Fieldsend 等<sup>[13]</sup>较详细的分析了非均匀分布情况,强调了决策者偏好的重要性.然而,这些研究未曾涉及解决非均匀问题 MOEA 的设计.

之所以至今不存在解决非均匀分布问题的有效算法,我们认为主要存在两方面困难:(I)非均匀问题 POF 分布的规律性难以确定;(II)非均匀分布解集对搜索存在一定的“阻碍”作用.另外,值得一提的是文献[14]采用一种带网格的小生境技术来保持问题真实分布,然而,这种技术需要根据问题 POF 分布人为设置疏密区间.事实上问题 POF 通常是不可知的,设计非均匀问题维护方法最主要的困难就是 POF 分布的规律性难以掌握,不同问题解集非均匀分布情况是不同的.

在本文中,我们首先分析经典 MOEAs 的分布性保持策略不适用非均匀分布 MOP 的原因,然后定义一个反映种群分布规则程度的指标——杂乱度,并提出一种在未知 Pareto 最优面分布规律情况下用杂乱度进行分布性维护的方法.最后选取并构造一些具有非均匀分布的测试函数进行实验.

## 2 分布性保持策略分析

分布性保持是 MOEA 设计中一个非常重要的方面.下面,我们简要地介绍几个典型的分布性保持策略(适应度共享,网格,聚集距离,密度估计),并说明它们不适应非均匀分布问题的原因.

适应度共享技术<sup>[5]</sup>是一种基于小生境的分布性保持策略.它根据小生境内个体数目连续地更新适应度共享函数,以获得非支配边界上良好的分布结果.网格方法<sup>[7]</sup>通过等分每一个坐标轴将解空间均匀划分为若干超立方体网格,进而剔除拥有个体最多的网格中的个体,达到保持分布度的目的.聚集距离方法是由 Deb 等于 2002 年提出的<sup>[6]</sup>.个体的聚集距离定义为与其在每个子目标上相邻的两个个体的距离差之和,通过计算种群中每个个体的聚集距离,构造一个偏序集,选择聚集距离大的  $N$  个个体进入下一代种群.基于密度的维护方法是由 Zitzler 等提出的<sup>[8]</sup>,该方法采用剪切技术对种群进行维护,依次选择距离最近的个体进行删除.

事实上,以上方法都是以维护非支配边界分布均匀性为目的,都是针对 POF 为均匀分布的测试问题而言的.现在我们用 Deb 的方法<sup>[10]</sup>构造一个简单的最小化问题,图 1 为规模 8 时其 POF 分布情况.

$$\begin{cases} f_1(\mathbf{X}) = x_1^{1.8} \\ f_2(\mathbf{X}) = g(\mathbf{X}_2)(1 - f_1(\mathbf{X})/g(\mathbf{X}_2)) \\ g(\mathbf{X}_2) = 1 + x_2 \end{cases} \quad (1)$$

容易发现,该 MOP 真实 Pareto 最优边界分布是不均匀的,个体在  $f_1$  上的值越接近 1 越稀疏.设图 2 为种群进化到一定时期的非支配个体的分布情况,下面我们分别用小生境,网格,聚集距离,基于密度的方法对其进行维护.

图 2 中待维护个体数目为 10,设种群规模为 8.图 3(a)~(d)为 4 种算法的种群维护结果.

图 3(a),(b)中的维护方法需要参数的设置,参数的不同,会导致不同的分布结果.图 3(a),(c)子图的维护方

法剔除了个体 B,E;(b),(d)子图的维护方法剔除了个体 B,H. 容易发现,虽然 4 种维护方法得到的结果有所不同,但都是试图剔除密集程度大的个体,“努力”使种群分布均匀.然而正是这种“均匀”破坏了问题的真实分布.其实通过观察,我们发现通常非均匀问题最优边界的分布并非杂乱无章,而是隐含一定的“规则”.譬如,个体密集程度在目

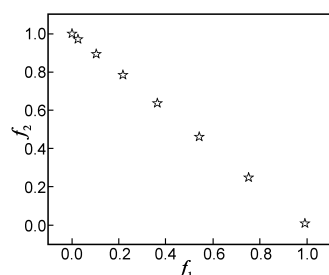


图1 规模为8时POF分布情况

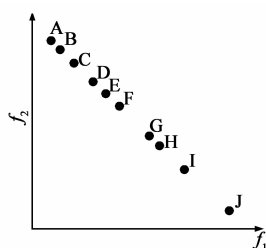


图2 种群分布示例

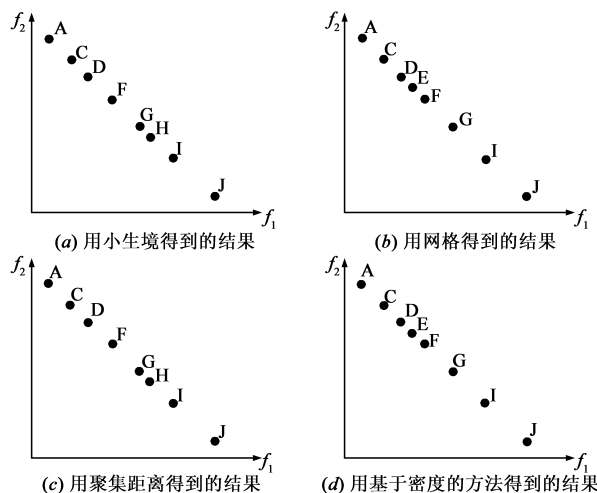


图3 4种算法的种群维护结果

标空间慢慢变化(如式(1)的测试问题),或者在一部分区域内密集,在另一部分区域内稀疏等等.下面,我们就按这种隐含的“规则”对非均匀分布种群进行维护.

### 3 解决非均匀问题的分布性维护方法

对于有一定分布规律的非均匀问题,在没有种群维护方法的情况下,其解集分布通常会或多或少的表现其偏好,如解集可能更多的分布在 Pareto 面中密集的区域.我们的方法就是利用问题的这种性质,选择群体中与周围个体差异最大的个体进行维护.下面,我们定义一个反映个体与其周围个体分布状况的指标——杂乱度.

#### 3.1 杂乱度

**定义 1** 杂乱度.对种群  $P$  生成一棵欧氏最小生成树(Euclidean minimum spanning tree, EMST),对于  $P$  的任一个体  $i$ ,定义  $i$  杂乱度  $messy_i$  为:

$$messy_i = l_{i\_max} / l_{i\_min} - 1 / d_i \quad (2)$$

EMST 为权重为点之间的欧氏距离的最小生成树.式(2)中  $d_i$  为个体  $i$  在 EMST 中的度数;  $l_{i\_max}$ ,  $l_{i\_min}$  分别为 EMST 中连接  $i$  的最长边和最短边.个体杂乱度由边长之比和个体度数两部分组成,边长之比越大,度数越高,杂乱度越大.

由 EMST 的贪婪性和连通性可知,EMST 中的边实质是连接不同聚类之间的最短距离(单个个体也可以看作为一个聚类),即种群由 EMST 中任意一边划分的两个个体集(聚类)之间,不存在比该边更短的距离.这样,EMST 中具有两个以上边的个体(即度数大于等于 2 的个体)可以看作连接不同聚类的中间个体,边的长度可以看作个体与聚类之间的距离.自然地,个体最大边与最小边的比反映了个体与不同聚类之间距离的最大差异,大的值表明了个体与周围不同聚类的“联系”参差不齐,相对混乱.另外,对于 EMST 中度数为 1 的个体,它们的最大边与最小边相同.这些个体只与一个聚类连接,没有反映个体与不同聚类之间联系的差异,我们赋予它较小的杂乱度.此外,个体在 EMST 中的度数也一定程度上反映了个体位置关系,通常边界个体具有较低的度数,这样在边长之比相近但分布位置不同的个体之间,边界个体具有更低的杂乱度.下面我们用个体杂乱度对种群进行维护.

#### 3.2 种群维护

杂乱度是一个反映个体规则程度的相对概念.容易发现,对杂乱度大的个体的淘汰不一定会使种群分布更有规律,相反很大程度上会使周边个体杂乱度有所增加.这里采用淘汰与最大杂乱度个体距离最近个体的方法降低种群杂乱度.算法伪码如下:

#### 算法 1

参数设置.  $Q$ :待维护非支配集;  $N$ :种群规模;  $a[i]$ :标示  $i$  是否已进入 EMST,  $a[i] = 1$  标示已进入,  $a[i] = -1$  标示未进入;  $edge\_max[i]$ :连接个体  $i$  最大边的边长;  $edge\_min[i]$ :连接个体  $i$  最小边的边长;  $minind\_edge[i]$ :与  $i$  组成连接个体  $i$  最小边的个体;  $messy[i]$ :个体  $i$  杂乱度;  $sel\_ind$ :种群中拥有最大杂乱度的个体.

while  $|Q| > N$  do

// 计算个体之间欧氏距离

for each  $j$  from 1 to  $N$

for each  $k$  from 1 to  $N$

if  $j \neq k$

$C[j][k] = \text{Euclidean}(j, k)$

end if

end for

end for

// 初始化赋值

for each  $i$  from 1 to  $N$

$a[i] = -1$

$edge\_max[i] = 0$

$edge\_min[i] = \text{INF}$

$degree[i] = 0$

$minind\_edge[i] = \text{NULL}$

end for

// 记录最大边、最小边以及个体度数

$a[1] = 1$

findmin() /\* find  $h$ , where  $C[1][h]$  is minimum  
&&  $a[h] = -1$  \*/

$a[h] = 1$

renew(1) /\* renew  $edge\_max[1]$ ,  $edge\_min[1]$ ,  
 $degree[1]$  and  $minind\_edge[1]$  \*/

renew( $h$ )

for each  $i$  from 3 to  $N$

findmin() /\* find  $j$  and  $k$ , where  $C[j][k]$  is the  
minimum &&  $a[j] = 1$ ,  $a[k] = -1$   
\*/

$a[k] = 1$

renew( $j$ )

renew( $k$ )

end for

// 计算并找出杂乱度最大个体

$max\_messy = 0$

for each  $i$  from 1 to  $N$

if  $edge\_min[i] \neq 0$

```
messy[i] = edge_max[i]/edge_min
[i] - 1/degree[i]
else
    messy[i] = INF
end if
if max_messy < messy[i]
    max_messy = messy[i]
    sel_ind = i
end if
end for
// 删除与最大杂乱度距离最近个体
delete(minind_edge[sel_ind])
|Q| = |Q| - 1
end while
```

算法 1 中算子 `renew()` 是根据新进入 EMST 的边调整个体在 EMST 中属性的函数. 另外, 由 EMST 贪心性可知, 最近个体组成的边都在 EMST 中, 因此在找出与  $i$  最近个体 (即调整 `minind_edge[i]`) 时, 只需比较 `edge_min[i]` 与新加入边的边长. 在计算个体杂乱度时, `edge_min[i] = 0` 表明个体  $i$  为重复个体, 赋予最大杂乱度, 算法 1 将首先淘汰与其重复的个体. 下面我们用法 1 对上节中的例子进行维护.

图 4(a) ~ (c) 为种群维护实例. 图 4(a) 中, 对种群生成 EMST 并找出最大杂乱度个体 G (`messy[G] = |FG|/|GH| - 1/2`). 因此, 与个体 G 最近的个体 H 被淘汰. 重复以上过程, 继续找出最大杂乱度个体 F, 则个体 E 被淘汰. 图 4(c) 为最终维护结果, 可以发现由左上至右下密度逐渐变小, 个体分布具有很好的规则性. 此外, 我们发现在三维或更高维数问题中, 位于非支配集内部度数较高的个体常拥有较大的杂乱度, 这可能是由于这些个体和周围联系比较紧密, 对周围不同密度个体集的连接起着中间桥梁的作用, 这样的个体是非常重要的. 而距离他们最近的个体通常度数较低, 与周围其它个体联系并不紧密, 一定程度上造成了种群分布的不规则, 算法 1 将首先剔除这类个体.

### 4 实验设计与结果

#### 4.1 实验参数设置

为检验 NUDMM 对非均匀问题的有效性, 我们与两个著名 MOEAs (SPEA2, NSGA-II) 进行比较实验. 我们选取 8 个非均匀分

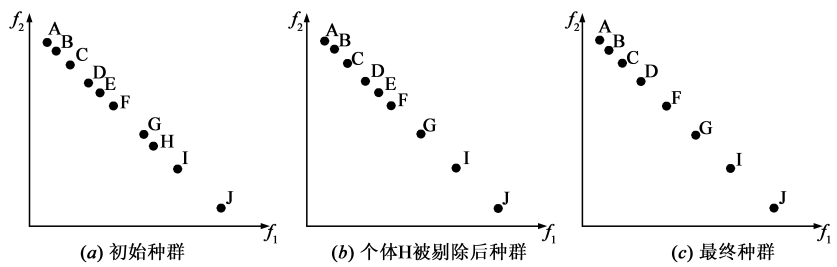


图4 算法1对图2种群维护过程

表 1 实验测试函数

测试问题	目标函数	约束条件及特征
NUF1	$f_1(X) = 1 - \exp(-4x_1) \sin^6(6\pi x_1)$	$m = 10, 0 \leq x_i \leq 1$ POF 凹 连续 较难收敛
	$f_2(X) = g(1 - (f_1/g)^2)$	
	$g(X) = 1 + 9((\sum_{i=2}^m x_i / (m-1))^{0.25})$	
NUF2	$f_1(X) = \sum_{i=1}^{n-1} (-10e^{(-0.2)} \sqrt{x_i^2 + x_{i+1}^2}),$ $f_2(X) = \sum_{i=1}^n ( x_i ^{0.8} + 5\sin(x_i)^3)$	$-5 \leq x_i \leq 5;$ $i = 1, 2, 3$ POF 非连续 较易收敛
NUF3	$f_1(X) = 1 - \exp(-4x_1) \sin^6(4\pi x_1)$ $f_2(X) = g(X)(1 - (f_1/g(X))^4)$ $g(X) = 1 + 10x_2$	$m = 2, 0 \leq x_i \leq 1$ POF 凹 连续 较易收敛
NUF4	$f_1(X) = x_1^2 + x_2^2$ $f_2(X) = g(X)(1 - (f_1/g(X))^{0.5})$ $g(X) = 1 + 10x_2$	$m = 3, 0 \leq x_i \leq 1$ POF 凸 连续 易收敛
NUF5	$f_1(X) = 1 - \exp(-2x_i) \sin^4(2\pi x_1)$ $f_2(X) = g(1 - (f_1/g)^{0.5} - (f_1/g) \sin(10\pi f_1))$ $g(X) = 1 + 9(\sum_{i=2}^m x_i / (m-1))$	$m = 30, 0 \leq x_i \leq 1$ POF 非连续 较难收敛
NUF6	$f_1(X) = 1 - x_1^2$ $f_2(X) = 1 - x_1^4$ $f_3(X) = g(X)(1 - (f_1/g(X))^4 - (f_2/g(X))^2)$ $g(X) = 1 + 10x_2$	$m = 2, 0 \leq x_i \leq 1$ POF 凹 连续 较易收敛
NUF7	$f_1(X) = 1 - x_1^3$ $f_2(X) = 1 - x_2^3$ $f_3(X) = g(1 - (f_1/g)^{0.5} - (f_2/g)^2)$ $g(X) = 1 + 9(\sum_{i=3}^m x_i / (m-1))$	$m = 30, 0 \leq x_i \leq 1$ POF 连续 难收敛
NUF8	$f_1(X) = (1 + g(X)) \cos(x_1^2 \pi / 2) \cdots \cos(x_{M-2}^2 \pi / 2) \cos(x_{M-1}^2 \pi / 2)$ $f_2(X) = (1 + g(X)) \cos(x_1^2 \pi / 2) \cdots \cos(x_{M-2}^2 \pi / 2) \sin(x_{M-1}^2 \pi / 2)$ ..... $f_M(X) = (1 + g(X)) \sin(x_1^2 \pi / 2)$ $g(X) = \sum_{x_i \in X_M} (x_i - 0.5)^2$	$n = 10, M = 3, \alpha = 10, 0 \leq x_i \leq 1$ POF 凹 较易收敛

布测试函数.如表 1 所示,NUF1 至 NUF5 为二目标测试函数. NUF1 为 Deb 提出的 ZDT6<sup>[10]</sup>,NUF2 为 Van Veldhuizen 的 MOP4<sup>[1]</sup>,NUF3 至 NUF5 为我们根据文献[10]设计的非均匀测试函数,其中  $f_1$  用于控制 POF 中解向量均匀性. NUF6 至 NUF8 为三目标测试函数,其中 NUF6,NUF7 为我们设计的非均匀测试函数,其中  $f_1, f_2$  用于控制 POF 中解向量均匀性. NUF8 为 Deb 提出的 DTLZ4<sup>[11]</sup>.

解集性能评价是 MOEA 中重要的一部分,分布性和收敛性是性能评价中两个基本指标<sup>[15]</sup>.然而,目前大多数分布性评价方法都是把解集分布均匀性作为测试标准<sup>[16]</sup>,它们并不适合非均匀问题.在此,我们选择两种借助于问题真实 Pareto 面的分布性评价方法:暴露程度(Exposure Degree, ED)<sup>[17]</sup>和改进的最大分散度(Improved Maximum Spread, MS')<sup>[18]</sup>.在收敛性方面,我们采用 Generational Distance(GD)<sup>[4]</sup>来估计算法的最终边界

与 POF 的趋近程度.此外,我们采用当前流行的 Hyper-volume(HV)方法<sup>[1]</sup>评价算法综合性能<sup>[19]</sup>,HV 把解集和空间参考点围成的超体积作为评价标准,超体积越大越好.在实验中,需要选择空间参考点,NUF1、NUF3 参考点设置为(2.0,2.0),NUF2 参考点为(-14.0,1.0),NUF4 参考点为(3.0,2.0),NUF5 参考点为(1.0,1.0)以及 NUF6 至 NUF8 参考点设置为(2.0,2.0,2.0).

在实验中,三种算法都采用实数编码,交叉概率 0.85,变异概率  $1/nreal$ ,  $nreal$  为决策变量维数,2、3 维测试函数种群规模分别为 100、200. SPEA2 内部种群与归档规模相同.每个算法对各个测试函数独立运行 30 次.

4.2 分布性评价

表 2、3 分别给出了三种算法对所有测试问题在分布性评价指标 ED,MS' 上的评价结果.在每个单元格中,第 1 行为均值,第 2 行为标准差.

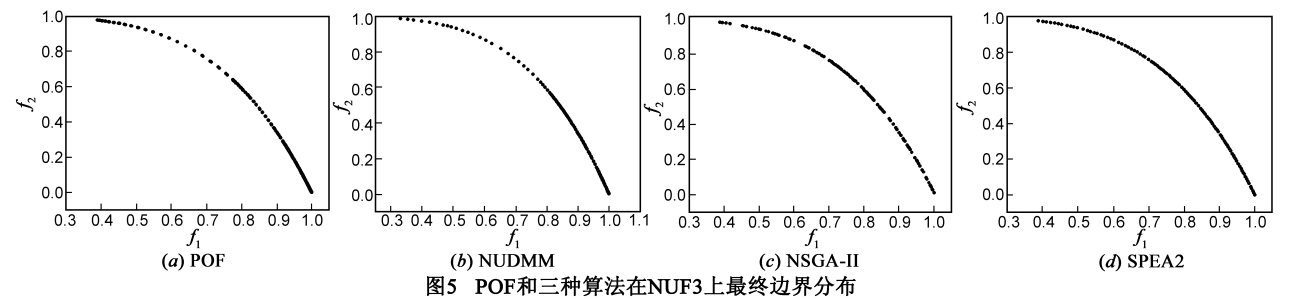


图5 POF和三种算法在NUF3上最终边界分布

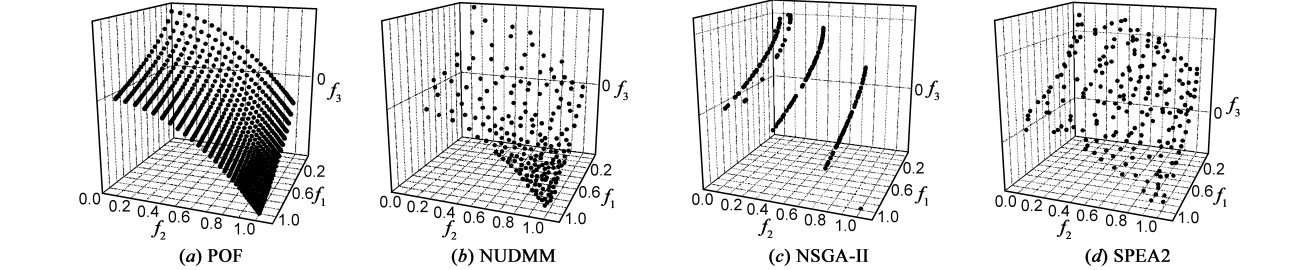


图6 POF和三种算法在NUF7上最终边界分布

表 2 分布性指标 ED 计算结果

Algorithms	NUF1	NUF2	NUF3	NUF4	NUF5	NUF6	NUF7	NUF8
NUDMM	0.005894	0.008405	0.002405	0.009321	0.039265	0.001861	0.028279	0.047941
	0.003002	0.006954	0.001574	0.007451	0.012423	0.001241	0.015417	0.032147
NSGA-II	0.008817	0.012471	0.007412	0.016242	0.045644	0.005412	0.106547	0.123654
	0.008651	0.009874	0.001542	0.005412	0.016571	0.002147	0.116471	0.132546
SPEA2	0.006471	0.008741	0.004115	0.009451	0.040212	0.002941	0.029654	0.066040
	0.001964	0.006878	0.001642	0.006327	0.010457	0.001365	0.023541	0.025324

可以发现在评价解集与真实 Pareto 面相似程度的指标 ED 上,NUDMM 始终具有最优结果,特别是在 NUF3,NUF6 和 NUF8 上体现得较为明显.对于广泛性评价指标 MS',三种算法结果相近,NUDMM 在 NUF3、NUF6、NUF7 上达到最优,而 NSGA-II 和 SPEA2 分别在 NUF2、NUF4 和 NUF1、NUF5、NUF8 上得到最优值.注意

到在 3 目标问题 NUF7 和 NUF8 上,NSGA-II 解集的分布范围要明显差于其它两种算法.这是由于 NSGA-II 解集收敛到局部 Pareto 面造成的.图 5、6 分别给出了三种算法在 NUF3 和 NUF7 上最终解集的分布情况.容易发现,相对于其它两种算法,NUDMM 的最终解集更接近于问题真实的分布.

表 3 广泛性指标 MS(计算结果)

Algorithms	NUF1	NUF2	NUF3	NUF4	NUF5	NUF6	NUF7	NUF8
NUDMM	0.99158	0.99942	0.99946	0.99901	0.98557	0.99926	0.97226	0.99326
	0.00083	0.00017	0.00011	0.00034	0.00264	0.00016	0.01034	0.00814
NSGA-II	0.99334	0.99944	0.99872	0.99906	0.98764	0.99862	0.84354	0.97486
	0.00086	0.00009	0.00032	0.00042	0.00186	0.00046	0.38592	0.02647
SPEA2	0.99387	0.99867	0.99241	0.99871	0.99101	0.99759	0.95641	0.99347
	0.00067	0.00016	0.00084	0.00031	0.00129	0.00033	0.05249	0.00947

4.3 收敛性、综合性评价

表 4 为三种算法在 8 个非均匀测试函数上收敛性评价结果.可以发现,除了在 NUF1 上,SPEA2 要优于 NUDMM 外.对其余非均匀测试函数,NUDMM 都拥有最优收敛度,特别是对较难收敛的 NUF7,NUDMM 要明显优于其它两种算法.表 5 为三种算法在 8 个非均匀测试函数 HV 评价结果.NUF8 括号内的数值表示算法在 30 次运行中成功的次数.可以发现,NUDMM 在 HV 上的优

势没有在 GD 上明显.NUDMM 在 NUF2,NUF3,NUF6,NUF7,NUF8 上拥有最优值,SPEA2 在 NUF1、NUF4,NSGA-II 在 NUF5 上达到最优的结果.这可能是由于 HV 是一种综合性能指标,它是由算法的收敛性、分布均匀性、分布广泛性决定的<sup>[19]</sup>.SPEA2,NSGA-II 在这些函数上通常能得到均匀分布的解集,因而能达到较大的 HV 值,虽然这并不能正确地反映非均匀问题的分布情况<sup>[13]</sup>.

表 4 收敛性指标 GD 计算结果

Algorithms	NUF1	NUF2	NUF3	NUF4	NUF5	NUF6	NUF7	NUF8
NUDMM	0.0159176	0.0024192	0.0006727	0.0018298	0.0008269	0.0001328	0.0049869	0.0002354
	0.0070097	3.953E-005	6.571E-005	0.0001614	0.0001159	2.027E-005	0.0009272	0.0001532
NSGA-II	0.0617145	0.0028161	0.0025321	0.0036463	0.0011891	0.0002051	0.0120774	0.0002412
	0.0135976	9.847E-005	0.0025383	0.0003516	9.674E-005	8.097E-006	0.0012363	0.0001051
SPEA2	0.0005604	0.0028876	0.0013903	0.0031531	0.0016485	0.0004208	0.0195618	0.0002861
	1.085E-005	6.954E-005	0.0015306	0.0010122	0.0001135	8.346E-005	0.0010531	0.0001931

表 5 综合性能评价指标 HV 计算结果

Algorithms	NUF1	NUF2	NUF3	NUF4	NUF5	NUF6	NUF7	NUF8
NUDMM	3.02310	37.0794	2.80754	6.28417	0.84564	7.33749	9.85338	7.10654(23)
	0.001546	0.005471	4.63E-005	0.000574	0.000219	0.000647	0.033214	0.471253
NSGA-II	3.02167	37.0775	2.80652	6.28291	0.84571	7.32762	9.56179	6.63796(11)
	0.002219	0.008126	8.85E-005	0.001027	0.000331	0.001351	0.045688	1.523476
SPEA2	3.02317	37.0751	2.80712	6.28617	0.84527	7.33253	9.77568	6.84031(15)
	0.001771	0.031549	2.25E-005	0.000147	0.000399	0.001736	0.034043	0.641257

另外,对于 NUF8,三种 MOEAs 并不是每次都能逼近问题的整个 Pareto 最优面.在 30 次独立运行中,NUDMM 成功 23 次,而 NSGA-II,SPEA2 分别为 11,15 次.可见,在求解容易收敛到部分 Pareto 最优面的 MOP 上,NUDMM 也具有一定的优势.

通过以上 8 个实验的分布性和收敛性、综合性能评价可以发现,NUDMM 是一种能有效解决非均匀问题的算法,它在很好的趋近 POF 最优边界的同时,较好的保持了问题真实分布情况.

5 结论

本文提出一种解决非均匀分布问题的多目标进化算法.利用 EMST 中邻近个体之间的联系,定义一个反映种群分布规则程度的指标,并设计一种降低种群杂乱度的维护方法,有效的剔除了那些造成种群混乱的个体,实验结果表明该算法能有效的解决非均匀问题,

在未知 POF 最优面分布规律的情况下得到与其分布类似的解集,为决策者提供了更有实用价值的方案.

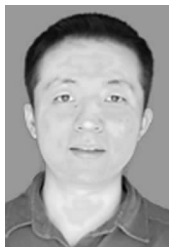
最后,需要指出的是,我们的方法是利用非均匀问题解集分布规律一定程度上反映其偏好这一特性进行维护,由于事先并不知道问题 Pareto 面的真实分布,得出的结果可能与真实的分布存在差异.如对于 Pareto 面分布为密集到稀疏再到密集的问题,我们得出的解集可能只是由稀疏到密集,我们的实验也证明了这点(如对测试问题 1 的结果).因此,如何更准确估计真实 Pareto 面的分布规律,甚至对真实 Pareto 面杂乱无章的非均匀问题进行有效地维护,是需要进一步研究的问题,这是我们将来的研究重点.

参考文献

[1] Deb K. Multi-objective optimization using evolutionary algorithms [M].Chichester,UK:John Wiley & Sons,2001.

- [2] 公茂果,焦李成,杨咚咚,马文萍.进化多目标优化算法研究[J].软件学报,2009,20(2):271-289.  
Gong Maoguo, Jiao Licheng, Yang Dongdong, Ma Wenping. Research on evolutionary multi-objective optimization algorithms [J]. Journal of Software, 2009, 20(2): 271-289. (in Chinese)
- [3] Schaffer J D. Some Experiments in Machine Learning Using Vector Evaluated Genetic Algorithms [D]. Nashville: Vanderbilt University, PhD thesis, 1984.
- [4] 郑金华.多目标进化算法及其应用[M].北京:科学出版社,2007.
- [5] Goldberg, D E., Richardson J. Genetic algorithms with sharing for multimodal function optimization [A]. Proceedings of the Second International Conference on Genetic Algorithms [C]. Massachusetts, US; L Erlbaum Associates Inc, 1987, 41-49.
- [6] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.
- [7] Corne D W, Jerram N R, Knowles J D, Oates M J, PESA-II: region-based selection in evolutionary multiobjective optimization [A]. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO' 2001) [C]. San Francisco, California; Morgan Kaufmann Publishers, 2001. 283-290.
- [8] Zitzler E, Laumanns M, Thiele L. SPEA2: Improving the Strength Pareto Evolutionary Algorithm [R]. Switzerland: TIK-Report 103, ETH, 2001.
- [9] Fonseca, C M Fleming, P J. An overview of evolutionary algorithms in multi-objective optimization [J]. Evolutionary Computation, 1995, 3(1): 1-16.
- [10] Deb K. Multi-objective genetic algorithms: problem difficulties and construction of test problems [J]. Evolutionary Computation, 1999, 7(3): 205-230.
- [11] Deb K, Thiele D, Laumanns M, Zitzler E. Scalable test problems for evolutionary multiobjective optimization [A]. Evolutionary Multiobjective Optimization: Theoretical Advances and Applications [C]. London: Springer-Verlag, 2005. 105-145.
- [12] Pedersen G, Goldberg D E. Dynamic uniform scaling for multiobjective genetic algorithms [A]. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO' 2004) [C]. Washington, USA: Springer-Verlag, 2004. 11-23.
- [13] Fieldsend J E, Singh S. Optimizing forecast model complexity using multi-objective evolutionary algorithms [A]. Applications of Multi-Objective Evolutionary Algorithms [C]. Singapore: World Scientific, 2004. 675-700.
- [14] 薛娟.多目标进化算法中非均匀问题的研究[D].湘潭大学信息工程学院,2005.
- [15] 雷德明,严新平,吴智铭.多目标混沌进化算法[J].电子学报,2006,34(6):1142-1145.  
Lei Deming, Yan Xinping, Wu Zhiming. Multi-objective chaotic evolutionary algorithm [J]. Acta Electronica Sinica, 2006, 34(6): 1142-1145. (in Chinese)
- [16] 李密青,郑金华,谢炯亮,杨平,李晶.一种 MOEA 分布度的逐步评价方法[J].电子学报,2008,36(10):1986-1991.  
Li Miqing, Zheng Jinhua, Xie Jiongliang, Yang Ping, Li Jing. A stepwise diversity metric for MOEA [J]. Acta Electronica Sinica, 2008, 36(10): 1986-1991. (in Chinese)
- [17] 曾三友,蔡振华,张青,康立山.一种评估近似 Pareto 前沿多样性的方法[J].软件学报,2008,19(6):1301-1308.  
Zeng Sanyou, Cai Zhenhua, Zhang Qing, Kang Lishan. A novel method for assessing the diversity of approximation Pareto fronts [J]. Journal of Software, 2008, 19(6): 1301-1308. (in Chinese)
- [18] Goh C K, Tan K C. An investigation on noisy environments in evolutionary multiobjective optimization [J]. IEEE Transactions on Evolutionary Computation, 2007, 11(3): 354-381.
- [19] Miqing Li, Jinhua Zheng. Spread assessment for evolutionary multi-objective optimization [A]. Evolutionary Multi-Criterion Optimization, 5th International Conference (EMO 2009) [C]. Nantes, France; Springer, 2009. 216-230.

#### 作者简介



**李密青** 男,1981 年生于湖南津市,现为湘潭大学信息工程学院实验师.主要研究方向为多目标进化算法.

E-mail: limit1008@126.com



**郑金华(通讯作者)** 男,1963 年生于湖南邵阳,现为湘潭大学信息工程学院教授,博士生导师,CCF 高级会员.主要研究方向为进化计算、智能科学.

E-mail: jhzheng@xtu.edu.cn

