

# 云模型雾化特性及在进化算法中的应用

刘 禹<sup>1</sup>, 李德毅<sup>2</sup>, 张光卫<sup>1</sup>, 陈桂生<sup>2</sup>

(1. 北京航空航天大学软件开发环境国家重点实验室, 北京 100083; 2. 中国电子工程系统研究所, 北京 100840)

**摘 要:** 云模型在超熵变大时, 体现出雾化特性, 云滴离散程度变大但靠近概念核心的云滴不失数量优势. 雾化特性有利于表示进化算法中的遗传与变异, 在云进化算法(CBEA)中, 基于超熵变化的控制进化策略能够合理的调整选择压力, 从而决定进化方向. 实验表明, 算法在寻求函数最优解问题上表现出良好的效率和精度.

**关键词:** 进化算法; 云模型; 进化策略; 雾化特性

**中图分类号:** TP18 **文献标识码:** A **文章编号:** 0372-2112 (2009) 08-1651-08

## Atomized Feature in Cloud Based Evolutionary Algorithm

LIU Yu<sup>1</sup>, LI De-yi<sup>2</sup>, ZHANG Guang-wei<sup>1</sup>, CHEN Gui-sheng<sup>2</sup>

(1. State Key Lab of Software Development Environment, Beihang University, Beijing 100083, China;

2. China Institute of Electronics Engineering, Beijing 100840, China)

**Abstract:** The drops of the cloud spread around while the hyper entropy (He) is increasing. But many drops still stand in the central area of the cloud. Atomized feature of the cloud model can be used to adjust the strategies of the evolution. In cloud based evolutionary algorithm, changing He to affect the selection pressure and lead to different evolution result. By the experiments of nine typical test functions' optimization, the precision, stability and convergence rate of the algorithm are well proved.

**Key words:** genetic algorithm; cloud model; evolution strategies; atomized feature

### 1 引言

模糊性和随机性是不确定性的两个重要方面, 传统的研究方法往往单独从模糊性, 或者单独从随机性的角度研究不确定性. 如果能够寻找到一个模型既考虑随机性, 又考虑模糊性, 并且考虑二者之间的关联性, 那么对于不确定性的表达和分析应该更为全面和科学. 针对此问题, 李德毅教授提出了云模型<sup>[1]</sup>, 用一个统一的模型实现定性概念与定量描述之间的不确定转换, 并以此为基础发展了一系列的关键技术, 目前云模型已经发展成为一个新的不确定性处理和理论, 得到了广泛应用<sup>[2]</sup>.

#### 1.1 云模型定义

**定义 1** 设  $U$  是一个用精确数值表示的定量论域,  $C$  是  $U$  上的定性概念, 若定量值  $x \in U$ , 且  $x$  是定性概念  $C$  的一次随机实现,  $x$  对定性概念  $C$  的确定度  $\mu(x) \in [0, 1]$  是有稳定倾向的随机数:  $\mu: U \rightarrow [0, 1]$ ,  $\forall x \in U, x \rightarrow \mu(x)$ . 则  $x$  在论域上的分布称为云, 每一个  $x$  称为一个云滴<sup>[2]</sup>.

云模型的数字特征用期望  $Ex$ 、熵  $En$  和超熵  $He$  来

表征, 它们反映了定性概念  $C$  的整体特性.

#### 1.2 正态云发生器

正态云模型是以概率理论和模糊理论为基础, 在正态分布和钟形隶属函数的基础上发展起来的分析和处理不确定性问题的全新模型<sup>[1]</sup>. 正向正态云发生器是从定性到定量的映射.

### 2 云模型雾化性质

当  $He$  取值较小时, 云滴呈现出泛正态分布状态;

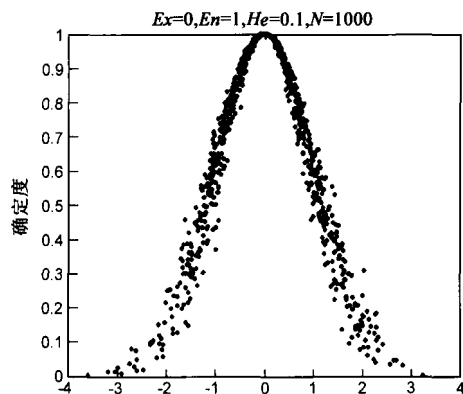


图1 泛正态云图

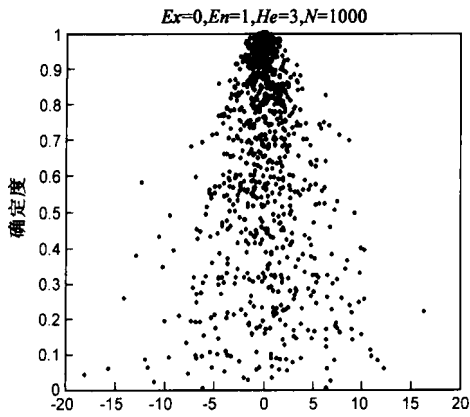


图2 雾化云图

当  $He$  取值较大时, 云滴所呈现的形状明显区别于正态分布, 外围云滴更加分散, 核心云滴出现明显的“抱团”, 将超熵取值较大时的云称之为“雾”. 随着  $He$  的变化, 云由一个极端(正态)到另一个极端(雾)的变化过程称之为雾化.

2.1 雾化的形成

超熵  $He = 0$  时, 云滴的分布为正态分布, 此时所有

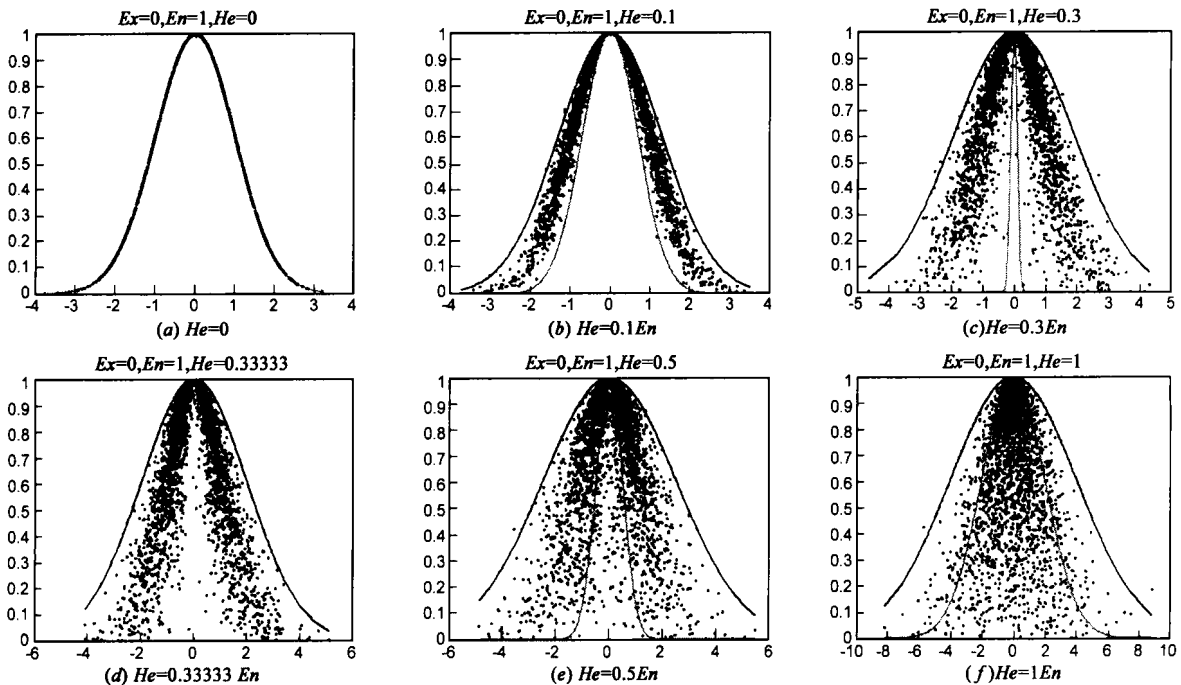


图3 雾化的形成

2.2 雾化特性的数学解析

若要研究正态云的雾化过程, 首先考虑正态云  $X$  的分布规律. 由正态云发生器算法<sup>[2]</sup>可知:

所有云滴  $x$  构成随机变量  $X$ .  $En'$  服从以  $En$  为期望值,  $He^2$  为方差的正态分布, 所以  $En'$  的概率密度函数为:

$$f_{En'}(x) = \frac{1}{\sqrt{2\pi He}} \exp\left\{-\frac{(x - En)^2}{2He^2}\right\}$$

当  $En'$  为定值时,  $X$  服从以  $Ex$  为期望、 $En'^2$  为方

差的正态分布, 如图 3(a) 所示, 正态点是云模型在图形上由正态曲线向云图转变的开始. 当  $He$  逐渐增大的时候, 云滴开始离散, 云滴的凝聚性变差了, 由于  $P\{En - 3He < En' < En + 3He\} = 0.997$ , 所以由正态云的实现算法可知, 当  $En - 3He > 0$ , 即  $He < En/3$  时, 99.7% 的云滴落在曲线  $y1 = e^{-(x - Ex)^2/2(En + 3He)^2}$  和  $y2 = e^{-(x - Ex)^2/2(En - 3He)^2}$  所围的区域内<sup>[3]</sup>, 如图 3(b)、(c) 所示.

当  $y2$  在  $He = En/3$ , 即  $En - 3He = 0$  时, 由于  $y2$  的指数趋向负无穷大, 函数值趋于 0, 如图 3(d) 所示.  $He > En/3$  时, 曲线  $y2$  的内径开始加宽, 部分云滴落在曲线  $y1$  与  $y2$  所围成的区域之外, 如图 3(e) 所示, 随着  $He$  的继续增大, 越来越多的云滴超出了两曲线所围区域, 如图 3(f). 而最终变化趋势为:  $\lim_{He \rightarrow \infty} y2 \rightarrow y1$ , 即两条函数曲线是趋向重合的, 此时所有云滴均逃离了两曲线所围范围. 可以将  $En/3$  称作正态云的雾化点, 当  $He < En/3$ , 云呈现泛正态状态;  $He > En/3$  时, 云呈现雾化状态.

差的正态分布, 此时  $X$  的概率密度函数为:

$$f_X(x | En') = \frac{1}{\sqrt{2\pi |En'|}} e^{-(x - Ex)^2/(2En'^2)}$$

由于  $En'$  是以  $Ex$  为期望,  $He^2$  为方差的正态随机数, 可能为负. 但是根据正态分布的概率密度公式  $f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x - Ex)^2/(2\sigma^2)}$  中的标准差  $\sigma$  不能为负<sup>[5]</sup>, 所以必须取绝对值.

由于  $En'$  是随机变量, 由条件概率密度公式, 可知  $X$  的概率函数为

$$f_X(x) = f_{En}(x) \times f_X(x|En)$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi He|y|} e^{[-(x-Ex)^2/(2\gamma^2) - (y-En)^2/(2He^2)]} dy$$

此函数是一个没有解析形式的概率密度函数,对任意的  $x$ ,可以通过数值积分计算得到相应的函数值.对于  $n$  个云滴的情况,可以利用 Paren 窗方法<sup>[6]</sup>估计  $X$  的概率密度函数.

特殊地,当  $He = 0$  时,  $X$  的概率密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi En}} e^{-(x-Ex)^2/(2En^2)}$$

因为所有的云滴  $x$  均来自期望为  $Ex$  的正态随机变量,所以  $X$  的期望  $EX = Ex$ ;

根据无需确定度的逆向云算法<sup>[4]</sup>对方差  $DX$  的计算方法的分析,可知:方差

$$DX = En^2 + He^2$$

该式表明正态云算法产生的云滴,是一个期望为  $Ex$ ,方差为  $En^2 + He^2$  的随机变量,呈现泛正态分布.参数  $He$  是偏离正态分布程度的度量,可以用来反映在论域空间影响概念的诸因素的不够均匀或者不够独立的现象.

### 2.3 雾化特性的统计特征

为研究正态云云滴整体的离散过程,考虑如下统计特征:

**定义 2.1 云论域区间  $C_d$**  云论域区间分割了正态云所表示的概念的论域范围,若记云  $C(X)$  中所有云滴的集合为  $Drops = \{x | x \in C(X)\}$ ,则论域区间  $C_d$  为云滴在  $X$  轴上的投影(投影点与距离  $Ex$  的距离不大于  $d$ )所构成的区域.对于二维正态云模型,如图 5 所示,论域区间可以看做以概念核心为中心,以变量  $d$  为半径的圆形.此时称  $d$  为论域区间的宽度.可以近似的将最大论域区间定义为云滴最小值  $\min(Drops)$  到云滴的最大值  $\max(Drops)$  之间的线段,最大论域宽度则为此线段长度的一半.

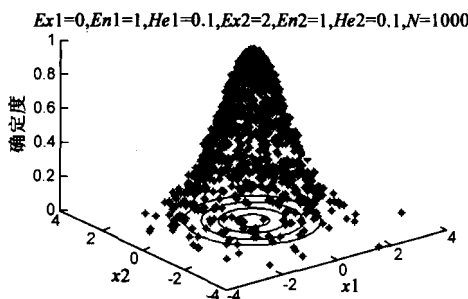


图4 二维正态云的论域区间

**定义 2.2 云密度** 用来表示投影在论域区间单位宽度长度上的云滴数,一维正态云记为  $\rho = \Delta cnt / \Delta d$ ,二维正态云记为  $\rho = \Delta cnt / \Delta s$ ,云密度反映论域某个区间

上云滴的密集程度.

由于云滴本身是离散的点,对一维正态云模型,统计  $X$  的某一区间  $A(A \in C_d)$  上的密度更有意义,可以近似的用投影在区间  $A$  上的云滴个数与区间长度的比值来表示区间云密度,记为  $\rho_A$ ,而将云滴整体的平均密度记为  $\rho_T$ .

### 实验 1

(1)取数字特征  $Ex = 0, En = 1, He = h$ ,通过正态云发生器生成云滴( $N = 1000, h$  初值为 0);

(2)计算云近似最大论域区间  $C_T$  与云平均密度  $\rho_T, i$  初值为 0;

(3)重复步骤(1)~(2),  $i = i + 1, 1000$  次,得到正态云模型在  $He$  为  $h$  时的平均密度  $C_{dh}$  与  $\rho_T$ ;

(4)变化参数  $h, h = h + 0.01 En$ ,重复步骤(1)~(3),分别计算出  $He$  在区间  $[0, 10En]$  上的近似最大论域区间和云密度.

### 实验 1 结果

随着超熵的变大,定性概念的论域范围越来越宽广,如图 5 所示.而云团的整体密度呈下降趋势,下降过程中,当  $He$  取值在  $[0, 5En]$  范围内时,云密度下降明显;当  $He$  超过  $5En$  时,下降趋势趋于缓和;当  $He$  超过  $7En$  时继续扩大时,云团已经充分离散,如图 6 所示.

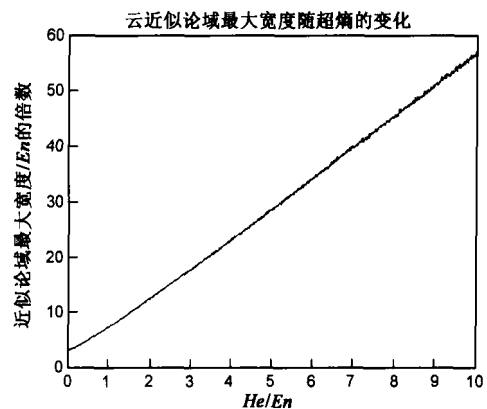


图5 云近似最大论域区间宽度变化

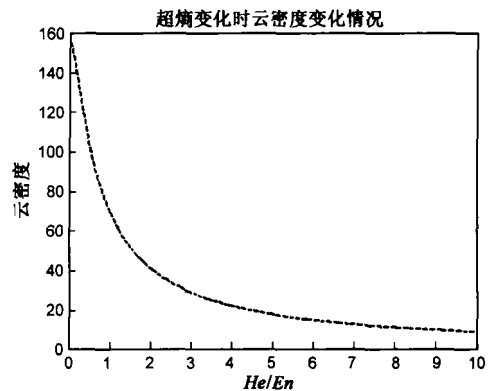


图6 云整体密度变化

### 2.3.1 云滴离散的区域趋势

本节讨论各个典型论域区间内的云密度变化情况:

骨干区间(记为 A)  $[Ex - 0.67En, Ex + 0.67En]$ ;

基本区间(记为 B)  $[Ex - En, Ex + En]$ ;

外围区间(记为 C)  $[Ex - 2En, Ex + 2En]$ ;

弱外围区间(记为 D)  $[Ex - 3En, Ex + 3En]$ ;

为验证各区间内云滴的离散程度,制定如下实验方案:

#### 实验 2

(1)取数字特征  $Ex = 0, En = 1, He = h$ ,通过正态云发生器生成云滴 ( $N = 1000, h$  初值为 0);

(2)计算云滴  $X$  投影于骨干区间内的云滴数,得到骨干区间内的云密度  $\rho_{Ai}, i$  初始 0;

(3)计算云滴  $X$  投影于基本区间内的云滴数,得到基本区间内的云密度  $\rho_{Bi}$ ;

(4)计算云滴  $X$  投影于外围区间内的云滴数,得到外围区间内的云密度  $\rho_{Ci}$ ;

(5)计算云滴  $X$  投影于弱外围区间内的云滴数,得到弱外围区间内的云密度  $\rho_{Di}$ ;

(6)重复步骤 1-5,  $i = i + 1, 1000$  次,得到正态云模型在  $He$  为  $h$  时各区间密度的均值  $\rho_{Ah}, \rho_{Bh}, \rho_{Ch}, \rho_{Dh}$ ;

(7)变化参数  $h, h = h + 0.01En$ ,重复步骤 1-6,分别计算出  $He$  在区间  $[0, 10En]$  上的各区间的区间密度.

#### 实验 2 结果

表 1 中给出了实验 2 的具有典型代表性的部分结果,从实验结果中可以看出:

对于骨干区间  $[Ex - 0.67En, Ex + 0.67En]$ ,当  $He < En$  时,随着  $He$  的增大,骨干区间密度  $\rho_A$  没有减小反而增大;当  $He > En$  时,  $\rho_A$  呈现出减小的趋势,此时才与云滴整体密度的趋势相符合.对于基本区间  $[Ex - En, Ex + En]$ ,变化规律与骨干区间相类似,不过  $He$  在  $0.4En$  附近时,  $\rho_B$  达到最大值,随后  $\rho_B$  开始减小,也就是在  $He$  的变化过程中,  $\rho_B$  的峰值的出现早于  $\rho_A$ .

表 1 区间密度变化表

$\rho$	$\rho_A$	$\rho_B$	$\rho_C$	$\rho_D$
$0.1En$	746.6179	684.7030	476.4225	332.1780
$0.2En$	759.6194	689.8260	473.6000	331.4220
$0.4En$	805.6433	700.8140	464.9240	328.6190
$0.6En$	835.8358	700.2820	454.9200	324.7180
$0.8En$	839.7015	690.7040	445.0510	320.3093
$1En$	822.0836	673.6640	435.0830	315.6290
$2En$	664.6358	555.9010	379.4250	287.7340
$3En$	539.5642	461.0640	329.9595	259.0500
$5En$	394.8224	344.7770	260.3735	213.0607
$8En$	285.3493	253.7970	199.3790	168.6613

### 2.3.2 云滴偏好依附率

偏好依附特性是指在复杂网络中,各个节点之间的连接并不是有序的、均匀的,拥有较多连接的核心节点,对于其它节点的具有较大的吸引力.考虑正态云模型,根据定义可知位于区间  $[Ex - 0.67En, Ex + 0.67En]$  之内的点,占全部定量值的 22.33%,它们对定性概念的贡献占总体贡献的 50%.这一性质体现了正态云模型的偏好依附特性,而偏好依附的唯一核心即为云的期望.作为定性概念的核心,云的期望  $Ex$  具有最大的吸引力,越靠近  $Ex$ ,云滴越密集,对概念的贡献也越大.

**定义 2.3 区间 R 的偏好依附率:**指定区间  $R, R$  以  $Ex$  为对称点,  $R$  长度不大于  $C_d, R$  的区间密度  $\rho_R$  与云滴整体密度  $\rho_T$  之比称之为区间  $R$  内的云滴对  $Ex$  的偏好依附率,简称区间  $R$  的偏好依附率,记为  $CPA$  (Cloud Preferential Attachment),  $CPA_R = \rho_R / \rho_C$ . 区间  $R$  的偏好依附率体现了  $R$  区间内的云滴的离散速率,  $CPA$  越高,表示  $R$  区间内的云滴离散速度越慢,也就是越能保持“抱团”特性.

#### 实验 3

(1)取数字特征  $Ex = 0, En = 1, He = h$ ,通过正态云发生器生成云滴 ( $N = 1000, h$  初值为 0);

(2)计算云平均密度  $\rho_T, i$  初始为 0;

(3)计算云滴  $X$  投影于骨干区间内的云滴数,从而得到骨干区间内的云密度  $\rho_{Ai}$ ,再而得到  $CPA_{Ai}$ ;

(4)计算云滴  $X$  投影于基本区间内的云滴数,从而得到基本区间内的云密度  $\rho_{Bi}$ ,再而得到  $CPA_{Bi}$ ;

(5)计算云滴  $X$  投影于外围区间内的云滴数,从而得到外围区间内的云密度  $\rho_{Ci}$ ,再而得到  $CPA_{Ci}$ ;

(6)计算云滴  $X$  投影于弱外围区间内的云滴数,从而得到弱外围区间内的云密度  $\rho_{Di}$ ,再而得到  $CPA_{Di}$ ;

(7)重复步骤(1)~(6),  $i = i + 1, 1000$  次,得到正态云模型在  $He$  为  $h$  时各区间偏好依附率的均值  $CPA_{Ah}, CPA_{Bh}, CPA_{Ch}, CPA_{Dh}$ ;

(8)变化参数  $h, h = h + 0.01En$ ,重复步骤 1-7,分别计算出  $He$  在区间  $[0, 10En]$  上的各区间的偏好依附

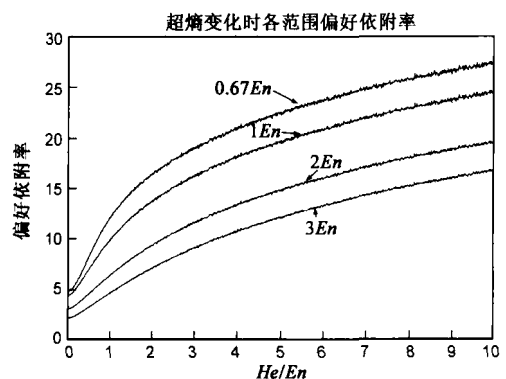


图 7 各区间偏好依附率的变化

率.

**实验 3 结果**

纵向比较表 2 中的数据,四个区间的偏好依附率都呈现明显的上升趋势,表明相对云整体的离散程度,各区间内的离散较慢,反向可以说明云整体宽度随着  $He$  的增大,呈现明显的增大趋势;横向比较数据可知,由 A 至 D,偏好依附率逐渐降低,这说明越靠近概念核心,云滴呈现出的抱团特性越明显,云滴整体扩散的过程中,云期望附近的云滴扩散的最慢.

表 2 区间偏好依附率变化表

CPA \ He	A	B	C	D
0.1En	4.9907	4.5777	3.1856	2.2211
0.2En	5.5239	5.0163	3.4437	2.4099
0.4En	7.1688	6.2360	4.1372	2.9243
0.6En	9.0225	7.5602	4.9109	3.5054
0.8En	10.7147	8.8137	5.6789	4.0876
1En	12.1066	9.9216	6.4074	4.6480
2En	16.6001	13.8830	9.4747	7.1843
3En	19.0275	16.2565	11.6330	9.1344
5En	22.5991	19.7375	14.9025	12.1940
8En	25.6922	22.8531	17.9528	15.1883

**3 云进化**

近代进化理论认为,进化的单位不是个体而是种群.使用云模型研究进化算法需要解决云模型对进化的表示和对进化的控制两个问题.

**3.1 进化知识的表示**

借鉴遗传学和生物学理论,给出本文进化算法中采用的若干概念:

**定义 3.1 基因 (Gene)** 遗传学中基因是一个 DNA 片段,在本文的进化算法中,基因指待优化对象的一个待优化的属性,记为  $g$ .对以上无约束优化问题  $g_i = x_i$ ,基因的取值  $|g_i|$  称为基因值,  $i$  称为基因编号,基因组 (Genome) 在遗传学上代表每一条染色体上的全部基因,本文中指个体基因的有序向量,  $G = (g_1, g_2, \dots, g_{DIM})$ ,  $DIM$  为基因的个数,也就是求解问题的维数,称为基因数,记为  $|G|$ .算法中种群个体只有一条染色体.基于基因组的概念,个体的编码可以看成基因组的一组特定取值  $Individual = (|g_1|, |g_2|, \dots, |g_{DIM}|)$ .

**定义 3.2 种群基因云 (Population Gene Cloud)**

用云模型建模一个种群中所有个体的同一个基因的基因值的分布特征得到的云,称为该种群的一个基因云.基因云是一个种群中个体基因值的定性表示,反映了种群基因的整体统计特征.第  $k$  个种群的第  $j$  个基因的基因云记为  $C_{k,j} = (Ex, En, He)$ ,  $(Ex, En, He)$  也称为种群基因的特征向量.

通过联合控制  $En$  和  $He$ ,可以调整进化过程中种群的选择压力,用  $En$  控制遗传的粒度,用  $He$  控制子代个体的凝聚离散程度和搜索半径的广度,结合第二部分的统计数据,得到如下结论,如表 3 所示.

表 3  $En$  和  $He$  对遗传变异的定性控制分析

En	He	描述
较大	较大	凝聚程度较小,离散程度较大,搜索半径较大,小比例后代个体围绕在父体周围,适合大范围搜索
较大	较小	遗传粒度较大,凝聚程度较大,离散程度较小,搜索半径较小,大比例后代个体围绕在父体周围,适合在新位置搜索初期采用
较小	较大	遗传粒度较小,凝聚程度较小,离散程度较大,搜索半径较大,小比例后代个体围绕在父体周围,搜索过程中适合跳出小局部
较小	较小	遗传粒度较小,凝聚程度较大,离散程度较小,搜索半径较小,大比例后代个体围绕在父体周围,搜索过程中适合局部求精

**3.2 遗传与变异的调整策略**

**定义 3.3 进化代 (Evolving Generation)** 进化过程中新群落的一次产生称为一个进化代.

**定义 3.4 适应度 (Fitness)** 群落中各个个体对环境的适应程度,是衡量个体优劣的标准,根据适应度的大小,决定某个个体是被保留还是被淘汰.适应度通常是费用、盈利、方差等目标的表达式,适应度函数的构造可以参考遗传算法中适应度评估函数的构造方法.

**定义 3.5 精英个体 (Elite Individual)** 指进化过程中发现的适应能力较好的个体,分为当代精英和跨代精英,出现跨代精英的进化代称为非平凡进化代;没有出现跨代精英的进化代称为平凡进化代;两个跨代精英个体之间相隔的进化代数称为连续平凡代数,是连续没有出现跨代精英个体的进化代数.连续出现跨代精英的进化代数称为连续非平凡代数.

**定义 3.6 定性进化策略 (Qualitative Evolving Strategy)** 指进化过程中对种群进化操作的定性控制策略,亦即通过调整种群基因云的参数  $Ex$ 、 $En$  和  $He$  来优化子代种群产生的策略.通过制定进化策略解决两方面问题:

(1)局部求精,当出现了跨代精英个体时,算法可能找到了新的极值邻域,或更加逼近了老的极值邻域,按照“当前优秀个体周围往往存在更加优秀的个体”原则,此时需要求精操作.一方面通过减小  $En$  来降低子代产生的遗传粒度,另一方面减小  $He$  以增大子代个体产生的凝聚程度和降低搜索半径,从而加大搜索的精度和稳定度以达到快速局部求精的目的.实际操作中,  $En$  可以按照某种函数关系下降,比如  $En$  线性降为原来的  $1/K (K > 1)$ ,  $K$  称为求精系数.  $He$  的变化则服从

根据经验编制的局部求精超熵控制表,如表 4 所示,表中超熵  $He$  取熵  $En$  的倍数,具体应用中需要根据连续非平凡代数的大小查表确定倍数  $\gamma$ .

(2)局部求变(进化式变异) 当在若干进化代均没有发现新的跨代精英,即连续平凡代数达到一定的阈值  $\theta_{local}$  时,算法可能陷入了一个局部最优邻域,说明当前位置周围没有更好的发现,此时需要跳出这个小局部,并在该局部附近尝试寻找新的局部最优.局部求变的方法是提高  $En$  和  $He$ ,实际应用中, $En$  的变化可以按照某种函数关系上升,比如  $En$  线性提高为原来的  $L$  倍, $L$  称为求变系数,一般情况下  $L \leq K$ ,可取  $L = \lceil \sqrt{K} \rceil$ .  $He$  的变化则服从根据经验编制的局部求变超熵控制表,如表 5 所示,表中超熵  $He$  取熵  $En$  的倍数,具体应用中需要根据连续非平凡代数的大小查表确定倍数  $\lambda$ .如果原  $He$  的值大于查表所得值,则以原  $He$  为准,否则以查表所得值为准.对函数的局部最优值非常邻近的剧烈波动的多峰函数,进化式变异有利于在众多邻近的局部最优值中寻找全局最优.当连续进化代数大于  $\theta + 6$  时,取  $\lambda = 20$ ,即  $He$  固定取  $En$  的 20 倍,表明算法在此后的进化代将以该位置为中心进行大范围搜索,但实验表明当连续进化代数大于  $\theta + 6$  时种群往

往已经被淘汰.

表 4 局部求精超熵控制表

连续非平凡进化代数	$\gamma$ ( $He = \gamma En$ )	云区间密度 ( $\delta = 1 En$ )	云区间密度 ( $\delta = 3 En$ )	论域区间最大宽度 ( $1 C_T 1$ ) ( $En$ 的倍数)	偏好依附率 (CPA) ( $\delta = 1 En$ )
1	10	21.703	44.573	113.62	12.334
2	8	25.373	50.505	90.420	11.467
3	4	39.372	70.231	45.830	9.0203
4	2	55.598	86.383	24.650	6.8530
5	1	67.459	94.714	14.722	4.9653
6	0.618	70.080	97.268	10.915	3.8243

表 5 局部求变超熵控制表

连续平凡进化代数	$\lambda$ ( $He = \lambda En$ )	云区间密度 ( $\delta = 1 En$ )	云区间密度 ( $\delta = 3 En$ )	论域区间最大宽度 ( $1 C_T 1$ ) ( $En$ 的倍数)	偏好依附率 (CPA) ( $\delta = 1 En$ )
$\theta_{local} + 1$	0.618	70.080	97.268	10.915	3.8243
$\theta_{local} + 2$	1	67.459	94.714	14.722	4.9653
$\theta_{local} + 3$	2	55.598	86.383	24.650	6.8530
$\theta_{local} + 4$	4	39.372	70.231	45.830	9.0203
$\theta_{local} + 5$	8	25.373	50.505	90.420	11.467
$\theta_{local} + 6$	10	21.703	44.573	113.62	12.334

### 3.3 云进化的具体实现

云进化算法 (Cloud Based Evolutionary Algorithm,

CBEA)遵循进化计算技术的一般过程:

(1)系统初始化为一组随机解,即随机初始化群落

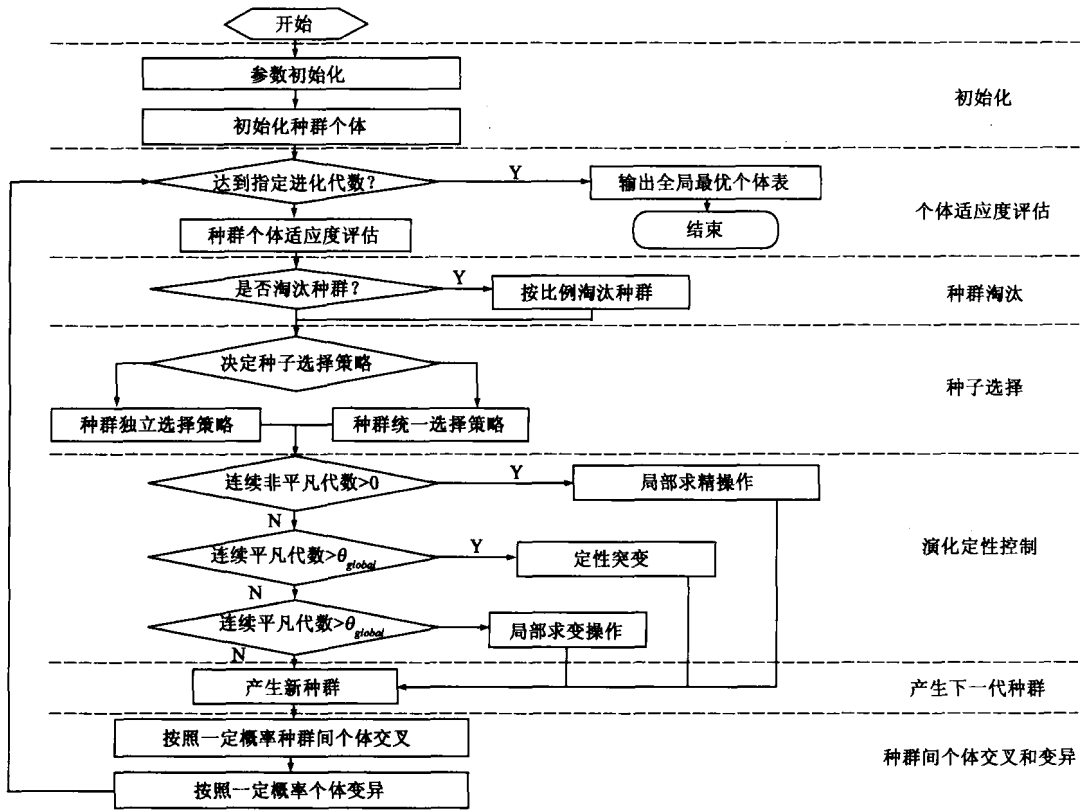


图8 定性进化算法流程图

中个体的值.

(2)适应度计算:计算所有种群中个体的适应度

(3)种群淘汰:根据条件淘汰适应性不好的种群

(4)种子选择:根据适应度选择种群的种子个体

(5)种群进化定性控制:根据进化情况对各个种群进行定性控制

(6)生成新种群:根据种子个体产生新种群,形成新群落

(7)种群间个体的遗传操作:交叉操作、变异

(8)如果达到进化代数则算法停止,最优秀的个体即为最优解,否则转步骤 2.

进化算法的详细结构如图 8 所示.

如果达到了进化代数,算法结束,此时全局最优个体即为所求最优解,此个体的适应度就是最优解的函数值.

#### 4 实验及结果分析

为了测试算法性能,选择如下非约束类高维测试函数进行测试,所选函数均为最小值问题,直接以函数值衡量函数适应度:

F01 Ackley 函数:

$$f(x) = 20 + e - 20e^{-(1/5)\sqrt{(1/n)\sum_{i=1}^n x_i^2 - e^{-(1/n)\sum_{i=1}^n \cos(2\pi x_i)}}}, -32.768 < x_i < 32.768, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0), f(x^*) = 0$$

F02 Griewank 函数:

$$f(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1, -32.768 < x_i < 32.768, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0), f(x^*) = 0$$

F03 Rastrigin 函数:

$$f(x) = 10n + \sum_{i=1}^n [x_i^2 - 10\cos(2\pi x_i)], -5.12 < x_i < 5.12, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0), f(x^*) = 0$$

F04 De Jong's first function 是连续、单峰的凸函数.

$$f(x) = \sum_{i=1}^n x_i^2 - 5.12 < x_i < 5.12, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0); f(x^*) = 0$$

F05 是一个阶梯函数,它具有一个最小值且不连续

$$f(x) = \sum_{i=1}^n (|x_i + 0.5|)^2, -100 < x_i < 100, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0); f(x^*) = 0$$

F06 Rotated hyper-ellipsoid 函数,连续单峰凸函数,是旋转的超椭圆体.

$$f(x) = \sum_{i=1}^n \left( \sum_{j=1}^i x_j \right)^2, -100 < x_i < 100, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0); f(x^*) = 0$$

F07 多维单峰函数.

$$f(x) = \sum_{i=1}^n |x_i| + \prod_{i=1}^n |x_i|, -10 < x_i < 10, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0); f(x^*) = 0$$

F08 多维单峰函数.

$$f(x) = \max\{|x_i|, 1 \leq i \leq n\}, -100 < x_i < 100, i = 1, 2, \dots, n, n = 30, x^* = (0, 0, \dots, 0); f(x^*) = 0$$

F09 Rosenbrock's valley 又称香蕉函数,或第二 De Jong 函数,是一个连续,单峰,全局单极值函数,全局极值落在一个狭长的山谷中,收敛到全局极值十分困难.

$$f(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2], -30 < x_i < 30, i = 1, 2, \dots, n, n = 30, x^* = (1, 1, \dots, 1); f(x^*) = 0$$

算法对每个函数独立运行 50 次求得最优解的平均值和方差如表 6 所示.为了进行对比,选择五种代表性的进化算法进行比较. OGA/Q (Orthogonal Genetic Algorithm with Quantization)<sup>[5]</sup> 是一种结合正交实验设计方法的遗传算法,是对 OGA 方法的改进,算法在减少计算量的同时有效的改进了 GA<sup>[6]</sup> 的全局搜索能力,效果很好. StGA (Stochastic Genetic Algorithm)<sup>[7]</sup> 是一种优秀的改进遗传算法,它通过改进编码策略来形成并动态调整搜索区域,算法用随机采样产生每一个区域点,并用其中最优秀的个体代表该区域,如果一个区域中找到了优秀的个体,那么该区域中点的采样变化度就降低,否则就升高采样变化度,该算法的思想和本文算法有一定的相似. OEA (Organizational Evolutionary Algorithm)<sup>[8]</sup> 是一种基于组织协同进化的数值优化算法.表 6 为算法计算

表 6 30 维函数优化结果对比

F	CBEA	OGA/Q	OEA	StGA	CLPSO	FEP
F01	2.8312E-015	4.44E-16	8.678E-14	3.52E-8	0	1.8E-2
	2.5632E-016	3.989E-17	1.001E-13	3.51E-9	0	2.1E-3
F02	0	0	2.767E-2	2.44E-17	4.56E-3	1.6E-2
	0	0	2.433E-2	4.54E-17	4.81E-3	2.2E-2
F03	0	0	2.7E-15	4.42E-13	4.85E-10	4.6E-2
	0	0	1.051E-14	1.14E-13	3.63E-10	1.2E-2
F04	1.1696E-166	0	4.21E-32	2.45E-15	4.46E-14	5.7E-4
	0	0	1.735E-31	5.25E-16	1.73E-14	1.3E-4
F05	0	/	0	0	/	0
	0	/	0	0	/	0
F06	5.0934E-119	0	1.238E-4	9.93E-29	/	1.6E-2
	1.6107E-118	0	8.287E-4	6.9E-29	/	1.4E-2
F07	5.2927E-097	0	4.169E-12	2.03E-7	/	8.1E-3
	7.6024E-097	0	2.289E-11	2.95E-8	/	7.7E-4
F08	1.1993E-065	0	7.361E-2	2.01E-8	/	3E-1
	2.2381E-065	0	2.476E-2	3.42E-9	/	5E-1
F09	2.6178E1	7.52E-1	4.07E-1	4.435E-2	2.1E1	5.06
	1.4174E-1	1.14E-1	1.392	0	2.98E0	5.87

结果的比较. FEP(Evolutionary programming made faster)<sup>[9]</sup>为一种改进的快速遗传规划算法.

根据表7的对比结果, CBEA的优化效果没有 OGA/Q好,但比其它几种算法有明显提高. OGA/Q算法本身带有一定的非计算复杂性,而且对不同的优化问题需要设计不同的正交表等,不便于应用, CBEA则较简单.可见本文算法的效率还是比较令人满意.在高维函数优化上 CBEA采用种群独立的种子选择策略,进化代数较大,收敛速度较低维函数慢些.

表7 30维函数优化性能对比

函数	CBEA	FEP
F01	150,000	200,000
F02	150,000	200,000
F03	200,000	500,000
F04	100,000	150,000
F05	150,000	150,000
F06	200,000	500,000
F07	150,000	200,000
F08	200,000	500,000
F09	200,000	2000,000

## 5 结论与展望

正态云模型在超熵变大的过程中,云滴整体区域离散,但靠近中心( $E_x$ )的云滴数量并不失数量优势,这一特征能够很好的控制进化过程中的求精和求变操作.利用正态云模型超熵的调整,可以模拟出不同的选择压力,从而控制进化的基因频率,实验证明,云进化算法具有优秀的执行效果,为进化算法的发展和云模型的实际应用提供了新思路.

正态云模型的另外一个特征是关于定性概念的确定度,在当前的云进化算法中,并没有考虑到确定度对种群的影响,此处可以作为未来研究的方向,以期更高的效率与精度.

### 参考文献:

- [1] 李德毅,刘常昱,杜鹃,韩旭.不确定性人工智能[J].软件学报,2004,15(11):1-13.  
Li De-yi, Liu Chang-yu, Du Yi, Han Xu. Artificial intelligence with uncertainty[J]. Journal of Software, 2004, 15(11): 1-13. (in Chinese)
- [2] 李德毅,杜鹃.不确定性人工智能[M].北京:国防工业出版社,2004. 137-166, 242-307.  
Li De-yi, Du Yi. Artificial Intelligence with Uncertainty[M]. Beijing: National Defense Industry Press, 2004. 137-166, 242-307. (in Chinese)

- [3] 刘常昱.不确定性知识表示研究[D].南京:解放军理工大学,2005. 67-84.
- [4] 李德毅,孟海军,史雪梅.隶属云和隶属云发生器[J].计算机研究与发展,1995,32(6):15-20.  
Li De-yi, Meng Hai-jun, Shi Xue-mei. Membership clouds and membership cloud generators[J]. Journal of Computer Research and Development, 1995, 32(6): 15-20. (in Chinese)
- [5] Y W Leung, F Wang. An orthogonal genetic algorithm with quantization for global numerical optimization[J]. IEEE Transactions on Evolutionary Computation, 2001, 5(1): 41-53.
- [6] 王正志,薄涛.进化计算[M].长沙:国防科技大学出版社,2000.
- [7] Zhenguo Tu, Yong Lu. A robust stochastic genetic algorithm (StGA) for global numerical optimization[J]. IEEE Transactions On Evolutionary Computation, 2004, 8(5): 456-470.
- [8] 刘静.协同进化算法及其应用研究[D].西安:西安电子科技大学,2004.  
Liu Jing. Coevolutionary Algorithms and Their Application [D]. Xi'an: Xidian University, 2004. (in Chinese)
- [9] Xin Yao, Yong Liu, Guangming Lin. Evolutionary programming made faster[J]. IEEE Transactions on Evolutionary Computation. 1999, 3(2): 82-102.

### 作者简介:



刘禹男,1980年生于河北辛集.北京航空航天大学计算机学院博士研究生.研究方向为不确定性人工智能与模式识别.  
E-mail: liuyu8014@163.com



李德毅男,1944年生于江苏泰县.中国工程院院士,主要研究方向为人工智能,复杂网络.  
E-mail: ziqinli@public2.bta.net.cn