第 5 章

烟 花 算 法

燃放烟花是中国传统节日尤其是春节的一个重要活动。在这一天,成千上万的烟花在夜空中爆炸并产生美丽的图案。通常,不同价格和规格的烟花在黑夜中爆炸产生不同的效果。一般价格高昂的烟花产生的火花数量比较多,爆炸产生的火花分布的范围也较集中;价格低廉的烟花产生的火花数量比较少,爆炸产生的火花分布的范围也较分散。

烟花算法(FWA)针对燃放的烟花在空中爆炸这种行为建立数学模型,通过引入随机因素和选择策略形成一种并行的搜索方法。它结合了群体智能算法中的交互机制和进化算法中的选择机制,是一种独树一帜的智能优化算法。

5.1 基本烟花算法

基本 FWA 由四大部分组成,包括爆炸算子、变异算子、映射规则和选择策略。爆炸算子的作用是在烟花的周围产生一批新的火花。产生火花的个数以及爆炸的幅度都由爆炸算子确定。另外,通过变异算子产生的火花服从高斯分布。在两种算子的作用下,如果产生的火花不在可行域范围内,需要运用映射规则将新产生的火花映射至可行域范围内,再利用选择策略选择新的火花作为下一代烟花。

FWA 开始迭代,依次利用爆炸算子、变异算子、映射规则和选择策略,直到达到终止条件,即满足问题的精度要求或者达到最大函数评估次数。FWA 的实现包括如下几个步骤:

(1) 在特定的解空间随机产生一些烟花,每个烟花代表解空间的一个解。

(2) 根据适应度函数计算每个烟花的适应度值,并根据适应度值产生火花。火花的个数是基于免疫学中的免疫浓度的思想计算的,即适应度值越高的烟花产生火花的数目越多。

(3) 根据现实中的烟花属性并结合搜索问题的实际情况,在烟花的辐射空间内产生火花(某个烟花的爆炸幅度的大小由该烟花在函数上的适应度值决定,适应度值越大,爆炸幅度越小,反之亦然)。每个火花代表解空间的一个解。为了保证种群的多样性,需要对烟花进行适当变异,如高斯变异。

(4) 计算种群的最优解,判定是否满足要求,如果满足则停止搜索,没有满足则继续迭代。迭代的初始值为此次循环得到的最好的解和选择的其他的解。

下面逐个介绍基本 FWA 的各个算子。

5.1.1 爆炸算子

FWA 的初始化是随机生成 N 个烟花的过程。接着,需要对生成的这 N 个烟花应用爆炸算子,以便产生新的火花。爆炸算子是 FWA 的关键核心并起关键性的作用,包含爆炸强

度、爆炸幅度和位移操作。

1. 爆炸强度

爆炸强度是 FWA 中爆炸算子的核心,它模拟的是现实生活中烟花爆炸的方式。任何一个烟花爆炸时,这个烟花周围都会产生一片火花。FWA 首先需要确定每个烟花爆炸产生火花的个数,以及在什么幅度内产生这些火花。

通过观察一些典型优化函数的曲线图,可以直观地看出,最优点附近的优质点也相应较多较密。因此,通过爆炸强度让适应度函数值好的烟花,产生的火花个数较多。这样可以避免寻优时火花总是在最优值附近摆动,而无法精准地找到最优值。对于适应度函数值较差的烟花,由于产生适应度函数值好的火花的概率较小,为避免做过多的不必要的计算,通过爆炸强度使其产生火花数较少。这种适应度函数值差的烟花的作用是对其余空间做适度的探索,避免早熟。可根据各个烟花适应度值的大小,确定每一个烟花产生火花的数量,让适应度值好的烟花产生更多的火花,适应度值差的烟花产生更少的火花。

在 FWA 中,产生火花个数的公式如下:

$$S_i = m \cdot \frac{y_{\max} - f(x_i) + \varepsilon}{\sum_{i=1}^{N}(y_{\max} - f(x_i)) + \varepsilon}$$

式中,S_i表示第 i 个烟花产生的火花个数,参数 i 的取值范围从 1 到 N。m 是一个常数,用来限制产生的火花总数。$y_{\max}$是当前种群中适应度值最差的个体的适应度值。$f(x_i)$表示个体 x_i的适应度值。最后一个参数 ε 取一个极小的常数,以避免出现分母为零的情况。

每个烟花爆炸火花的数量限制公式如下:

$$\hat{S}_i = \begin{cases} \text{round}(a \cdot m), & S_i < am \\ \text{round}(b \cdot m), & S_i > bm \\ \text{round}(a \cdot m), & \text{其他} \end{cases}$$

式中,a 和 b 是常数,$a<b<1$,S_i是火花数量的界限,round()是四舍五入函数。

2. 爆炸幅度

通过观察一些函数的曲线图,可以直观地看出,通常函数的最优值、局部极值附近的点的函数值通常也较优。因此在 FWA 中,通过控制爆炸幅度,让适应度函数值好的烟花爆炸幅度减小,这样才能更有效地收敛到各个极值,直至最终找到最优值。相反,适应度函数值较差的点,往往离最优值都较远,只有让这些适应度函数值差的烟花产生大幅度的变异,才能使其有效地到达最优值附近。这就是控制烟花爆炸幅度的基本思想。

烟花爆炸幅度范围的计算公式如下:

$$A_i = \hat{A} \cdot \frac{f(x_i) - y_{\min} + \varepsilon}{\sum_{i=1}^{N}(f(x_i) - y_{\min}) + \varepsilon}$$

式中,A_i表示第 i 个烟花的爆炸幅度范围,即爆炸的火花将在这个范围内随机产生位移,但不能超出这个范围。$\hat{A}$是一个常数,表示最大的爆炸幅度。参数$y_{\min}$是当前种群中适应度值最好的个体的适应度值。

3. 位移操作

在计算出爆炸幅度之后,需要确定烟花在爆炸幅度范围内的位移。这里用到的是随机

位移的方法，对烟花进行位移变异。这样，每个烟花都有自己特定的火花数目和爆炸幅度。在某个爆炸幅度内，能随机产生一个位移，生成新的火花，保证了种群多样性。通过爆炸算子，每个烟花都能生成一批新的火花，为寻找全局最优解提供了保障。

位移操作是对烟花的每一维进行位移，其公式如下：

$$\Delta x_i^k = x_i^k + \mathrm{rand}(0, A_i)$$

式中，$\mathrm{rand}(0, A_i)$表示在幅度 A_i 内生成的均匀随机数。

算法 5-1 给出了 FWA 产生火花的伪代码。

算法 5-1　产生火花

```
1. 初始化烟花，并计算出每个烟花的适应度值 f(x_i)；
2. 计算每个烟花生成的火花个数 S_i；
3. 计算每个烟花生成火花的爆炸幅度 A_i；
4. z = rand(1, dimension) //随机选择维度集合 z；
5. for k = 1→dimension do
6.    if k ∈ z then
7.          Δx_i^k = x_i^k + rand(0, A_i)；
8.    end if;
9. end for
```

评论：基本 FWA 通过两个计算爆炸火花数和爆炸幅度的公式来实现烟花之间的相互配合，即好的烟花在小范围内产生较多火花以进行开采，而差的烟花在大范围内产生较少火花以进行探索。尽管思路可取，但这两个具体公式有不合理之处。比如，最差的烟花的爆炸火花数和最好的烟花的爆炸幅度都很接近 0，这显然不符合算法设计的本意。尽管爆炸火花数通过阈值限制的方式进行了修补，但最好的烟花的爆炸幅度的问题是更致命的。这也引发了之后许多对基本 FWA 的探讨和改进。

5.1.2　变异算子

为进一步提高种群的多样性，在 FWA 中引入高斯变异。高斯变异火花产生的过程如下：在烟花种群中随机选择一个烟花，对选择得到的烟花随机选择一定数量的维度进行高斯变异。

高斯变异在选中的烟花和最好的烟花之间进行变异，产生新的火花。高斯变异可能产生超出可行“解空间”范围的火花。当火花在某一维度上超出边界时，将通过映射规则映射到一个新的位置。

用 x_i^k 表示第 i 个个体在第 k 维上的位置，此时高斯变异的计算方式如下：

$$x_i^k = x_i^k g$$

式中，g 是服从如下均值为 1、方差为 1 的高斯分布的随机数。

$$g = N(1, 1)$$

算法 5-2 给出了 FWA 中高斯变异的伪代码。

算法 5-2　高斯变异

```
1. 初始化烟花，并计算出每个烟花的适应度值 f(x_i)；
2. 计算高斯变异的系数 g = N(1,1)；
```

```
3. z=rand(1,dimension)        //随机选择 z 个维度
4. for k=1→dimension do
5.    if k∈z then
6.          x_i^k = x_i^k g
7.    end if
8. end
```

评论：基本 FWA 中的高斯变异火花是在烟花位置上乘以一个随机数产生的，因而会导致这种火花在原点附近较为集中。而对于一般的优化问题，其实不应假设原点有任何特殊性。这意味着基本 FWA 在面对最优点在原点和不在原点的两类函数时性能会有所不同。

5.1.3 映射规则

如果某一个烟花靠近可行域的边界，而爆炸幅度范围又覆盖到边界以外的区域，则可能在可行域范围外产生火花。这种火花是无用的，因此需要通过一种规则将其拉回可行域范围内。这里采用映射规则来应对这种情况。映射规则确保所有个体留在可行的空间。如果在边界附近又产生一些越界的火花，则它们将被映射到可行域的范围。

采用模运算的映射规则，其公式如下：

$$x_i^k = x_{\min}^k + |x_i^k| \% (x_{\max}^k - x_{\min}^k)$$

式中，x_i^k表示超出边界的第 i 个个体在第 k 维上的位置，$x_{\max}^k$和 $x_{\min}^k$分别表示第 k 维上的边界上下界。百分号表示模运算。

评论：映射规则是为了解决带边界约束的优化问题。但边界约束只是约束中的其中一种，除此之外还有其他类型的等式和不等式约束。关于如何处理带约束的优化问题，当前的研究还很不充分。另外，除了基于取模运算的映射规则之外，还有很多种其他可选的映射规则，例如映射到边界、随机映射、镜面映射等。

5.1.4 选择策略

运用爆炸算子和变异算子并保证产生的火花在可行域范围之后，需要从产生的火花中选择出一部分作为下一代的烟花。FWA 用到的是基于距离的选择策略。为了选择进入下一代的个体，选择策略采用的方式是每次都留下最优个体，再选择其他的 $N-1$ 个个体。为保证种群的多样性，采用 $N-1$ 个个体中和其他个体距离更远的个体有更多的机会被选中的选择策略。

在 FWA 中，采用欧氏距离用来度量任意两个个体之间的距离。$d(x_i, x_j)$表示任意两个个体 x_i和 x_j之间的欧氏距离。

$$R(x_i) = \sum_{j=1}^{K} d(x_i - x_j) = \sum_{j=1}^{K} ||x_i - x_j||$$

用 $R(x_i)$表示个体 x_i与其他个体的距离之和，$j \in K$ 是指第 j 个位置属于集合 K。其中，集合 K 是爆炸算子和高斯变异产生的火花的位置集合。个体选择采用轮盘赌方式，每个个体被选择的概率用 $p(x_i)$ 表示。

$$p(x_i)=\frac{R(x_i)}{\sum_{j\in K}R(x_i)}$$

离其他个体距离更远的个体具有更多的机会成为下一代个体。这种选择策略保证了FWA的种群多样性。

算法5-3是FWA的伪代码。

算法5-3 FWA的伪代码

```
1. 随机选择 n 个烟花的位置;
2. while 当前函数评估次数<最大函数评估次数 do
3.   对于 n 个烟花;
4.   for 所有烟花 x_i do
5.       计算每个烟花产生的火花个数 S_i;
6.       计算每个烟花产生火花的幅度 A_i;
7.   end for
8.   随机产生火花;
//  m̂是烟花高斯变异产生的火花数;
9.   for k=1→m̂ do
10.         随机选择一个烟花 x_i 并产生一个火花;
11.  end for
12.  依据映射规则对火花进行映射;
13.  依据选择策略选择最好的烟花以及其他烟花;
14. end while
```

评论:选择机制在群体算法中非常少见,是FWA独有的设计。基于密度的选择机制有利于保持选出的烟花的多样性和代表性。但也有学者指出可以采用其他类型的选择机制,以适应不同的需要。

5.1.5 基本烟花算法特点分析

1. 爆发性

在FWA的一次迭代开始后,烟花在辐射范围内爆炸,会产生其他火花。等本次迭代结束后,FWA选择 N 个火花作为下一代的烟花,恢复了烟花的数目,并为下次迭代的爆发做好准备。每次迭代,烟花都会爆发,说明FWA具有爆发性。

2. 瞬时性

当一次迭代计算开始后,各个烟花依据适应度值的不同,产生不同的火花个数和爆炸幅度。接着,FWA将在爆炸算子和变异算子的作用下产生火花。最后,首先选出最优个体,再依据距离选择其他的 $N-1$ 个个体。这些选择出来的 N 个个体将作为下一代爆炸的烟花,其余的火花不再保留。不保留的火花或烟花将消亡,说明FWA具有瞬时性。

3. 简单性

和群体智能算法一样,每个烟花只需要感知自身周围的信息,遵循简单的规则,完成自身的使命。总体来看,FWA本身并不复杂,由简单个体组成,说明FWA具有简单性。

4. 局部性

在FWA中,所有的烟花都会在相应的爆炸幅度内产生火花。除非超出可行域,产生的

火花都局限在一定的范围内。FWA 的局部性特点体现了 FWA 强大的局部搜索能力，可以用在算法运算的后期更加精细地搜索最优解。因此，FWA 具有局部性。

5. 涌现性

烟花之间通过竞争与协作，群体之间表现出简单个体不具有的高度智能性。烟花之间相互作用，比单个个体的行为要复杂得多，因此 FWA 具有涌现性。

6. 分布并行性

在 FWA 的每次迭代过程中，各个烟花在不同坐标范围内依次爆炸，即对不同的坐标区间进行一次搜索，在最后会将所有的火花和烟花综合起来，进行下一代烟花的选择。在一次迭代中，算法实质上是并行搜索，表现出 FWA 的分布并行性。

7. 多样性

种群多样性是影响群体优化算法性能的关键。群体多样性的保持，可以保证算法跳出局部极值点，从而可以收敛到全局最优点，这正是群体优化算法与一般优化算法的显著区别。群体多样性越高，算法中的个体分布越广，找到最优值的可能性越大，同时还不会明显影响算法的收敛能力。因此，种群多样性是 FWA 的一个重要特点。FWA 的多样性主要体现在以下三方面。

（1）火花个数和爆炸幅度的多样性：在爆炸算子的作用下，依据各个烟花的适应度值，其产生不同个数的火花和不同的爆炸幅度。适应度值高的烟花产生更多的火花，爆炸幅度相对较小，而适应度值低的烟花产生更少的火花，爆炸幅度相对较大。因此，保证了火花个数和爆炸幅度的多样性。

（2）位移操作和高斯变异的多样性：FWA 有两种算子，第一种是爆炸算子，第二种是变异算子。在爆炸算子的位移操作中，对计算出来的幅度范围，随机产生一个位移，将在选中的烟花加上这个随机位移。在变异算子的作用下，选中的烟花需要乘以一个满足高斯分布的随机数。爆炸算子与烟花的适应度值有关，变异算子与烟花本身的坐标值有关。两种算子本质上是不同的，都保证了爆炸的多样性。

（3）烟花的多样性：通过一定的选择机制，保留下来的烟花坐标值各不相同，从而保证了 FWA 的多样性特点。另外，在选择策略中，距离其他火花距离更远的火花更容易被选中，也体现出 FWA 中烟花的多样性特点。

8. 可扩充性

FWA 中烟花和火花的数量不确定，可以依据问题的复杂度确定。烟花和火花的数目可多可少，增加和减少个体都能有效地求解问题，因此 FWA 具有可扩充性。

9. 适应性

FWA 求解问题时，不要求问题具有显式表达，只要计算适应度值就能求解问题。同时，FWA 对问题的要求低，也能求解显式表达的问题。因此 FWA 具有适应性。

5.2 烟花算法的收敛性与稳定性分析

本节讨论 FWA 的收敛性和时间复杂度等理论，内容包括 FWA 的随机模型、全局收敛性、时间复杂度的基本理论和时间复杂度分析。

5.2.1 随机模型

假设 FWA 的随机模型采用基本下确界(essential infimum),并定义如下:

$$\varphi = \inf(t : v[n \in S \mid f(z) < t] > 0)$$

式中,$v[A]$是在集合 A 上的勒贝格测度(Lebesgue measure)。上述公式意味着在搜索空间的子集中,存在多个点,使得函数值趋近于 φ,使得在勒贝格可度量的非零集合中,φ 为函数值的下确界。

下面建立 FWA 的随机过程。

定义 5-1 $\{\xi(t)\}_{t=0}^{\infty}$为 FWA 的随机过程,其中 $\xi(t)=\{F(t),T(t)\}$,而 $F(t)=\{F_1(t),F_2(t),\cdots,F_n(t)\}$表示在 t 时刻 N 个烟花在解空间中的位置。$T(t)=\{A(t),S(t)\}$,其中 $A(t)=\{A_1(t),A_2(t),\cdots,A_n(t)\}$表示 N 个烟花爆炸的幅度,而 $S(t)=\{s_1(t),s_2(t),\cdots,s_n(t)\}$表示 N 个烟花爆炸的数目。

接下来定义最优区域。

定义 5-2 $R_\varepsilon=\{x \in S \mid f(x)-f(x^*)<\varepsilon,\varepsilon>0\}$是函数 $f(x)$的最优区域,x^* 表示函数 $f(x)$在解空间的最优解。

根据定义 5-2,如果算法找到一个位于最优区域的点,视为算法找到了函数的接近全局最优的一个可接受的解。根据定义 5-1,最优解空间的勒贝格测度必须不为零,这意味着 $v(R_\varepsilon)>0$。

定义 5-3 定义 FWA 的最优状态为$\xi^*(t)=\{F^*(t),T(t)\}$,同时,存在$F_i(t)\in R_\varepsilon$和$F_i(t)\in F^*(t)$,$i\in 1,2,\cdots,n$。

定义 5-3 说明在 FWA 的最优状态$\xi^*(t)$下,最好的烟花在最优区域R_ε中。所以这里存在$F_i(t)\in R$ 和$|f(F_i(t))-f(x^*)|<\varepsilon$,$x^*\in R_\varepsilon$。

引理 5-1 FWA 的随机过程$\{\xi(t)\}_{t=0}^{\infty}$是一个马尔可夫(Markov)过程。

证明略。

定义 5-4 (最优状态空间)用 Y 表示 FWA 的状态 $\xi(t)$的状态空间,$Y^*\subset Y$。只要存在一个解$s^*\in F^*$使得$s^*\in R_\varepsilon$在任意状态$\xi^*(t)=\{F^*,T\}\in Y$ 下成立,那么Y^*就是最优状态空间。

定义 5-4 指出,$|f(s^*)-f(x^*)|<\varepsilon$ 对任意 $x^*\in F^*$都成立。如果 FWA 可以达到最优状态,则必定有一个火花达到了最优区域R_ε,且 FWA 得到了最优解。此后,最优解一定在最优区域内。

定义 5-5 给定一个马尔可夫随机过程$\{\xi(t)\}_{t=0}^{\infty}$和优化状态空间$Y^*\subset Y$,$\{\xi(t)\}_{t=0}^{\infty}$如果满足 $P\{\xi(t+1)\notin Y^* \mid \xi(t)\in Y^*\}=0$,则被命名为吸收马尔可夫过程。

引理 5-2 FWA 的随机过程$\{\xi(t)\}_{t=0}^{\infty}$是一个吸收马尔可夫过程。

证明略。

5.2.2 全局收敛性

定义 5-6 (收敛性)给定一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}=\{F(t),T(t)\}$和一个优化状态空间$Y^*\subset Y$,$\lambda(t)=P\{\xi(t)\in Y^*\}$表示在 t 时刻,随机状态达到最优状态的概率。如果

$\lim\limits_{t\to\infty}\lambda(t)=1$，则$\{\xi(t)\}_{t=0}^{\infty}$收敛。

根据上面的定义，马尔可夫随机过程的收敛取决于$P\{\xi(t)\in Y^*\}$的概率。如果在t时刻马尔可夫随机过程的收敛概率为1，可以认为马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$收敛。

定理 5-1 给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^*\subset Y$，如果对于任意的$t$，$P\{\xi(t)\in Y^*|\xi(t-1)\notin Y^*\}\geqslant d\geqslant 0$，且$P\{\xi(t)\in Y^*|\xi(t-1)\in Y^*\}=1$成立，则$P\{\xi(t)\in Y^*\}\geqslant 1-(1-d)^t$。

证明略。

FWA 包含高斯变异。为简化问题，假定该变异是一种随机变异。

定理 5-2 给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^*\subset Y$，则$\lim\limits_{t\to\infty}\lambda(t)=1$意味着$\{\xi(t)\}_{t=0}^{\infty}$能收敛到最优状态$Y^*$。

证明略。

以上的定义和定理证明了 FWA 的马尔可夫过程将收敛到最优状态。

5.2.3 时间复杂度的基本理论

定义 5-7 （期望收敛时间）给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^*\subset Y$，如果$\gamma$是一个随机非负值使得如果$t\geqslant\gamma$，有$P\{\xi(t+1)\in Y^*\}=1$；如果$0\leqslant t\leqslant\gamma$，有$P\{\xi(t+1)\notin Y^*\}<1$。那么$\gamma$就是 FWA 的收敛时间。FWA 的期望收敛时间用E_γ表示。

期望收敛时间E_γ描述的是 FWA 以 1 的概率初次得到全局最优解的时间。期望值E_γ越小，FWA 的收敛就越快，FWA 也就更加有效。但是，也可以用首次最优解期望时间(Expected First Hitting Time，EFHT)作为收敛时间的一个标志。

定义 5-8 （首次最优解期望时间）给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^*\subset Y$；$\mu$是一个随机值，使得如果$t=\mu$，则$\xi(t)\notin Y^*$；如果$0\leqslant t\leqslant\mu$，则$\xi(t)\notin Y^*$。期望值$E_\mu$称为首次最优解期望时间。

下面的定理给出了计算E_γ的方法。

定理 5-3 给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^*\subset Y$。如果$\lambda(t)=P\{\xi(t)\in Y^*\}$并且$\lim\limits_{t\to\infty}\lambda(t)=1$，则期望收敛时间$E_\gamma=\sum\limits_{t=0}^{\infty}(1-\lambda(t))$。

证明略。

因为很难得到$\lambda(t)$的值，所以很难计算出期望收敛时间E_γ。因此，只能给出估计的时间。

定理 5-4 给定两个随机非负变量u和v，并用$D_u(\cdot)$和$D_v(\cdot)$分别表示u和v的分布函数。如果$D_u(t)\geqslant D_v(t)$（$\forall t=0,1,2,\cdots$），则u和v的期望值$E_u<E_v$。

证明略。

定理 5-5 给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^*\subset Y$，如果$\lambda(t)=P\{\xi(t)\in Y^*\}$使得$0\leqslant D_l(t)\leqslant\lambda(t)\leqslant D_h(t)\leqslant 1$（$\forall t=0,1,2,\cdots$）且$\lim\limits_{t\to\infty}\lambda(t)=1$，那么$\sum\limits_{t=1}^{\infty}(1-D_h(t))\leqslant E_\gamma\leqslant\sum\limits_{t=1}^{\infty}(1-D_l(t))$。

证明略。

定理 5-6　给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^* \subset Y$，如果$\lambda(t)=P\{\xi(t)\in Y^*\}$且$0\leqslant a(t)\leqslant\lambda(t)\leqslant b(t)$，则$\sum_{t=1}^{\infty}\left[(1-\lambda(0))\prod_{t=1}^{\infty}(1-b(t))\right]\leqslant E_\gamma\leqslant\sum_{t=1}^{\infty}\left[(1-\lambda(0))\prod_{t=1}^{\infty}(1-a(t))\right]$。

证明略。

推论 5-1　给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^* \subset Y$和$\lambda(t)=P\{\xi(t)\in Y^*\}$，如果$a\leqslant P\{\xi(t+1)\in Y^* \mid \xi(t+1)\notin Y^*\}\leqslant b(a,b>0)$且$\lim\limits_{t\to\infty}\lambda(t)=1$，则 FWA 的期望收敛时间$E_\gamma$满足下列不等式：

$$b^{-1}[1-\lambda(0)]\leqslant E_\gamma\leqslant a^{-1}[1-\lambda(0)]$$

证明略。

上述推论和定理表明公式$P\{\xi(t)\in Y^* \mid \xi(t-1)\notin Y^*\}$可以描述 FWA 的烟花从非最优状态到最优状态的概率。E_γ值的估计范围可以通过$P\{\xi(t)\in Y^* \mid \xi(t-1)\notin Y^*\}$的值来计算。

5.2.4　时间复杂度分析

FWA 的时间复杂度需要计算期望收敛时间E_γ。依据推论 5-1，FWA 的时间复杂度主要和 FWA 的烟花从非最优区域到最优区域R_ε的概率相关，即$P\{\xi(t+1)\in Y^* \mid \xi(t-1)\notin Y^*\}$。这里，将进一步分析此公式来得到 FWA 的时间复杂度。FWA 包含爆炸算子、变异算子、映射规则和选择策略，但与 FWA 的马尔可夫状态到达最优区域直接相关的是爆炸算子和变异算子。因此，有下面的定理。

定理 5-7　给定 FWA 的一个吸收马尔可夫过程$\{\xi(t)\}_{t=0}^{\infty}$和一个优化状态空间$Y^* \subset Y$，则有

$$\frac{v(R_\varepsilon)\times n}{v(S)}\leqslant P\{\xi(t+1)\in Y^* \mid \xi(t)\notin Y^*\}\leqslant v(R_\varepsilon)\left(\frac{n}{v(S)}+\sum_{i=1}^{n}\frac{m_i}{v(A_i)}\right)$$

式中，$v(R_\varepsilon)$是最优区域R_ε的勒贝格测度值，$v(S)$是问题搜索区域S的勒贝格测度值，$v(A_i)$是第i个烟花的爆炸幅度A_i的勒贝格测度值。

上面的定理给出了非常粗糙的结果，因为实际的公式很难进行确定性的计算。FWA 很难准确计算出火花落在最优区域R_ε的概率。为了准确地实现，需要做如下变换：

$$P(\exp)=\sum_{i=1}^{n}\frac{v(S_i\cap R_\varepsilon)\times m_i}{v(S_i)}$$

式中，$v(S_i\cap R_\varepsilon)$和m_i随着算法的运行在动态改变，所以它们非常重要。$v(S_i\cap R_\varepsilon)$和烟花的位置F_i相关。FWA 的选择策略使得位置距离大的个体有更高的概率被选中，所以可以假定每次只有一个烟花处于最优区域R_ε，进一步假设适应度值最高的烟花进入最优区域R_ε的概率最高。依据上述假设，$v(A_i)>v(A_{\text{best}})$且$m_i>m_{\text{best}}$，$i\in(0,1,2,\cdots)$。其中$A_{\text{best}}$和$m_{\text{best}}$分别是适应度值最高的烟花的爆炸区域和生成火花的数目。由此得到

$$\frac{v(A_i\cap R_\varepsilon)\times m_i}{v(A_i)}<\frac{v(A_{\text{best}}\cap R_\varepsilon)\times m_{\text{best}}}{v(A_{\text{best}})}$$

考虑在算法运行初期$(A_i \cap R_\varepsilon) \cap (A_{\text{best}} \cap R_\varepsilon) = \varnothing$，其中$i \in (0,1,2,\cdots)$且$i \neq \text{best}$，可得下面的公式：

$$P(\exp) = \sum_{i=1}^{n} \frac{v(S_i \cap R_\varepsilon) \times m_i}{v(S_i)} < \frac{v(S_{\text{best}} \cap R_\varepsilon) \times m_{\text{best}}}{v(S_{\text{best}})} < \frac{v(R_\varepsilon) \times m_{\text{best}}}{v(S_{\text{best}})}$$

所以

$$\frac{v(R_\varepsilon) \times n}{v(S)} \leqslant P\{\xi(t+1) \in Y^* \mid \xi(t) \notin Y^*\} \leqslant v(R_\varepsilon)\left(\frac{n}{v(S)} + \frac{m_{\text{best}}}{v(S_{\text{best}})}\right)$$

设$a = \dfrac{v(R_\varepsilon) \times n}{v(S)}$，$b = v(R_\varepsilon)\left(\dfrac{n}{v(S)} + \dfrac{m_{\text{best}}}{v(S_{\text{best}})}\right)$，那么可以得到下面的公式：

$$\frac{v(S) \times v(S_{\text{best}})}{v(R_\varepsilon) \times (n \times v(S_{\text{best}}) + m_{\text{best}} \times v(S))} \times (1-\lambda(0)) \leqslant E_\gamma \leqslant \frac{v(S)}{v(R_\varepsilon) \times n} \times (1-\lambda(0))$$

FWA 初始种群中的n个烟花是随机生成的，因此可以得出$\lambda(t) = P\{\xi(t) \in Y^*\}$。由于$\lambda(0) = P\{\xi(0) \in Y^*\} \ll 1$，$1-\lambda(0) = 1$，因此

$$\frac{v(S) \times v(S_{\text{best}})}{v(R_\varepsilon) \times (n \times v(S_{\text{best}}) + m_{\text{best}} \times v(S))} \leqslant E\gamma \leqslant \frac{v(S)}{v(R_\varepsilon) \times n}$$

由此可以看出，R_ε的值越大，并且$v(S)$的值越小，将提高 FWA 的效率。但这两个变量都和搜索问题相关。$v(S_{\text{best}})$和m_{best}对于 FWA 的期望收敛时间非常重要。但上述结论是在一些假设条件下成立的。更精确的分析需要进一步考虑 FWA 公式的细节。

5.3 增强烟花算法和动态搜索烟花算法

在 FWA 被提出时，FWA 就表现出极为优异的性能。Zheng 等对 FWA 的算子进行了详细的分析，针对算法存在的缺陷进行了有效的改进并提出了增强烟花算法(EFWA)。EFWA 表现出比 FWA 更稳定可靠的优化性能。

在 EFWA 中，最小爆炸半径检测算子使得 EFWA 中适应度最优的烟花能够发挥其强大的搜索能力。然而，这种简单的仅仅根据当前的适应度值评估次数和最大适应度值评估次数确定的最小爆炸半径检测策略(MEACS)过于人为干预设定，并没有考虑到算法优化过程中的动态优化的信息。基于此，本章针对烟花种群中适应度值最优的烟花提出了依据算法优化过程中种群是否寻找到更优的解而动态地调整适应度值最优的烟花的爆炸半径的动态适应策略，并称其为动态搜索烟花算法(dynFWA)。

FWA 的搜索能力主要取决于烟花的爆炸算子的作用。在爆炸半径范围内，一个烟花能够同时产生一定数量的火花从而可以对烟花周围区域进行细致的搜索。在 FWA 和 EFWA 中，爆炸半径是用于调整烟花种群的局部搜索能力和全局搜索能力的关键参数。算法根据各个烟花当前位置的适应度值来计算其爆炸半径和爆炸火花数目。其主要思想是，一个烟花的适应度值越好(对最小化问题而言就是越小)，它生成的火花数量就越多，而范围就越小；反之，一个烟花的适应度值越坏(越大)，它生成的火花数量就越少，而范围就越大。从而，处于较好位置的烟花将在当前位置的较小区域周围进行局部搜索，而适应度值大的烟花将在更大范围进行全局搜索。相对于基本 FWA，EFWA 针对 FWA 存在的缺陷提出了多方面的改进。下面我们主要以 EFWA 为基准介绍 FWA，并分析其最小爆炸半径检测算