2014 IEEE Congress on Evolutionary Computation (CEC)
July 6-11, 2014, Beijing, China

Analysis on Global Convergence and Time Complexity of Fireworks Algorithm

Jianhua Liu, Shaoqiu Zheng and Ying Tan

***Abstract*— Fireworks Algorithm (FWA) is a new proposed optimization technique based on swarm intelligence. In FWA, the algorithm generates the explosion sparks and Gaussian mutation sparks by the explosion operator and Gaussian mutation operator to search the global optimum in the problem space. FWA has been applied in various fields of practical optimization problems and gains great success. However, its convergence property has not been analyzed since it has been provided. Same as other swarm intelligence (SI) algorithms, the optimization process of FWA is able to be considered as a Markov process. In this paper, a Markov stochastic process on FWA has been defined, and is used to prove the global convergence of FWA while analyzing its time complexity. In addition, the computation of the approximation region of expected convergence time of FWA has also been given.**

I. Introduction

DEVELOPED by Tan Y. and Zhu Y. in 2010, Fireworks Algorithm (FWA) simulates the process of fireworks explosion in the night sky to search the optimal solution of optimization problem [1]. It is a new swarm intelligence (SI) algorithm based optimization techniques, which generates the explosion sparks and Gaussian mutation sparks by the explosion operator and Gaussian mutation operator to search the problem space. Compared to the other SI algorithms, FWA has distinctive advantages in solving optimization problems while presenting a different search manner.

Since its first introduction, FWA has shown its significance and superiority in dealing with the optimization problems and has been seen many improvements and applications with practical optimization problems [1]- [2]. Andreas and Tan used FWA to compute the initialization of non-negative matrix factorization and gains a little advantages compared to SPSO, FSS, GA [2], [3]. Pie et al. investigated the influences of approximation approaches on accelerating the FWA search by elite strategy [4]. In [4], they compared the approximation models, sampling methods, and sampling number on the FWA acceleration performance, and the random sampling method with two-degree polynomial model gains better performance on the benchmark functions. Gao Hongyuan and Ming Diao designed a cultural FWA which is used to search optimal value of filter design parameters with parallel search. Simulation results have shown that FIR and IIR digital filters based on the cultural FWA are superior to previous filters based on the other SI algorithms in terms of convergence speed and optimization results [5]. Zheng Yujun et.al proposed a hybrid multi objective fireworks optimization algorithm (MOFOA) for oil crop fertilization [6], which takes into consideration not only crop yield and quality but also energy consumption and environmental effects. Liu Jianhua et al. [7] provided a kind of new method to calculate the number of explosion sparks and amplitudes of fireworks explosion on FWA. The modified FWA improves the performance compared with original FWA. He Wenrui et al. [8] proposed a new framework that optimizes anti-spam model with heuristic swarm intelligence optimization algorithms. This framework could integrate various classifiers and feature extraction methods, which consider the spam detection problem as an optimization process which aims to achieve the lower error rate.

Though many researcher have developed the original algorithm and applications, FWA has never been analyzed about its convergence and time complexity since it has been provided. Same as the other SI algorithms, FWA can be considered as a kind of population-based Evolutionary Algorithms (EAs). The use of population has been regarded as one of the key features of EAs. EAs have been shown to be very effective in solving practical problems, yet many important theoretical issues of them are not clear. The expected first hitting time and Markov Stochastic Process are two important theoretical issues of evolutionary algorithms, since it implies the average computational time complexity. In the developments of convergence analysis of EAs, many researchers have developed this work. He Jun, and Xin Yao [9] used first hitting time and Markov Modal to compares (1 + 1) EAs to (N + N) EAs theoretically which is shown that a population can have a drastic impact on an EA's average computation time, changing an exponential time to a polynomial time (in the input size) in some cases. It is also shown that the first hitting probability can be improved by introducing a population. Yu Yang and Zhihua Zhou [10] established a bridge between the expected first hitting time and another important theoretical issue and proposed a new general approach to estimating the expected first hitting time. Based on this approach, they analyzed EAs with different configurations, including three mutation operators, with/without population, a recombination operator and a time variant mutation operator, on a hard problem. Huang Han et al [11] analyzed the ACO convergence time based on the absorbing Markov chain model. They presented a

Jianhua Liu is with School of Information Science and Engineering, Fujian University of Technology, Fuzhou, Fujian (email:jhliu@fjnu.edu.cn). Shaoqiu Zheng and Ying Tan (Corresponding author) are with Department of Machine Intelligence, School of EECS, Peking University Key Laboratory of Machine Perception (MOE), Peking University,Beijing,China

This work was supported by National Natural Science Foundation of China (NSFC) with Grant No. 61375119, 61170057 and 60875080 and by the Provincial Natural Science Foundation of Fujian with Grant No. 2012J01246. The author,Jianhua Liu, was supported by Chinese Scholarship Council(CSC) as visiting scholar in RMIT university, Melbourne,Australia.

Algorithm 1 Conventional Fireworks Algorithm.

1: Select n positions for initial fireworks;
2: Calculate the number of sparks for each firework;
3: Calculate the amplitude of explosion for each firework;
4: Generate the sparks of explosion for each firework;
5: Generate the sparks of Gaussian mutation for each firework;
6: Select n locations for next generation fireworks
7: If condition does not meet, algorithm turns to 2
8: Output results

general result for the estimation of convergence time to reveal the relationship between convergence time and pheromone rate. Chen Tianshi et al [12] studied the computational time complexity of a simple EDA in order to gain more insight into EDAs complexity. They proved theoretically that the univariate marginal distribution algorithm(UMDA) with margins can solve the BVLeadingOnes problem efficiently. Yi Shengqiu et al [13] focused on the theoretical analysis of quantum-inspired evolutionary algorithms with $H\epsilon$ gate which is a modified version of the rotation gate. Applying the theory and analytical techniques in non-homogeneous Markov chains, they obtained the conclusion that quantum-inspired evolutionary algorithms converge in probability under some mild conditions. Ding Lixin and Jinghu Yu [14] introduced some techniques for the analysis of time complexity of evolutionary algorithms (EAs) based on a finite search space. The Markov property and the decomposition of a matrix are employed for the exact analytic expressions of the mean first hitting times that EAs reach the optimal solutions (FHT-OS). Huang Han et al [15] used an absorbing Markov process to analyze the time complexity of Evolutionary Programming and ant colony algorithm, which constructs some theory of analysis on swarm intelligence.

Same as other evolutionary algorithms, FWA can be considered as the Markov stochastic process to prove its global convergence and compute the expected time. This paper will define a Markov stochastic process for FWA and develop its theory model. Using the conceptions and theorems on FWA, we prove its global convergence and derive the approximate region of expected convergence time.

The paper is organized as follows: Section II introduces the FWA ; the Markov Modal of FWA is constructed in Section III. Section IV analyzes and proves the global convergence of FWA. The Section V gives the basic theory of time complexity. The time complexity of FWA is analyzed in Section VI. The conclusion is drawn in Section VI.

II. Fireworks Algorithm

For the model of Fireworks Algorithm, the sparks and fireworks are the potential solutions for an optimization problem. In term of the literatures [1] and [7], the procedure of conventional Firework Algorithm can be seen in Algorithm.1.

For a minimal optimization problem, according to the idea of FWA, a good firework denotes a position with the better fitness which means that the firework may be close to the optimal/local minimal point. Therefore, the good firework will generate bigger number of sparks in the smaller amplitude of explosion. On the contrast, the bad firework will generate smaller number of sparks in the larger amplitude of explosion. For each firework, the number of explosion sparks and amplitude of explosion needs to be calculated before it explodes, which formulas of each dimension is as follows:

- Calculating the Number of Sparks which is operated in step 2 of Algorithm 1:

$$S_i = m\frac{y_{\max} - f(x_i) + \tau}{\sum\limits_{i=1}^{n}(y_{\max} - f(x_i)) + \tau} \tag{1}$$

$$m_i = \begin{cases} round(a \cdot m) & if \;\; s_i < a \cdot m \\ round(b \cdot m) & if \;\; s_i > b \cdot m \\ round(S_i) & otherwise \end{cases} \tag{2}$$

  where m_i is the number of the sparks of the ith firework explosion, m is the total number of sparks generated by the n fireworks. y_{max} is the maximum value of the objective function among the n fireworks, and τ is a small constant which is utilized to avoid zero-division-error. The constant a and b are the const parameters.
- Calculating the Amplitude of Explosion which is operated in step 3 of Algorithm 1:

$$A_i = \tilde{A}\frac{f(x_i) - y_{\min} + \tau}{\sum\limits_{i=1}^{n}(f(x_i) - y_{\min}) + \tau} \tag{3}$$

  Where $\tilde{A}$ denotes the maximum explosion amplitude and y_{min} is the minimum value of the objective function among n fireworks. A_i is the exploding region of ith firework.
- In original FWA, it selects n location for next generation fireworks in each iteration. The best firework or spark needs to be selected as the first location of next generation fireworks, and the other locations of next generation fireworks are selected by the following Eq.(4) and Eq.(5), which are computed in step 6 of Algorithm 1:

$$R(x_i) = \sum_{j \in K} d(x_i, x_j) = \sum_{j \in K} ||x_i - x_j|| \tag{4}$$

$$p(x_i) = \frac{R(x_i)}{\sum_{j \in K} R(x_i)} \tag{5}$$

  where the x_i is the location of ith sparks or firework, $d(x_i, x_j)$ is the distance between two sparks or fireworks. K is the set of sparks and firework generated in current generation. The $p(x_i)$ is the probability which the ith firework or spark is selected as the firework of next generation.

III. THE STOCHASTIC MODAL OF FIREWORKS ALGORITHM

It is assumed that FWA search is undertaken for the essential infimum which is defined as Eq.(6).

$$\psi = inf(t : \nu[n \in S|f(z) < t] > 0) \tag{6}$$

where $\nu[A]$ is the Lebesgue measure on the set A. The above equation means that there must be more than one point in a subset of search space yielding function values arbitrarily close to ψ, so that ψ is the infimum of the function values from nonzero Lebesgue measurable set, and then the stochastic process of Fireworks Algorithm is established as follows.

Definition 1: $\{\xi(t)\}_{t=0}^{\infty}$ is named as the stochastic process of Fireworks Algorithm, where $\xi(t) = \{F(t), T(t)\}$ which $F(t) = \{F_1(t), F_2(t), \ldots, F_n(t)\}$ denotes the position of n fireworks in the problem space in the step t; And $T(t) = \{A(t), S(t)\}, A(t) = \{A_1(t), A_2(t), \ldots, A_n(t)\}$ denotes the explosion amplitude of n fireworks, $S(t) = \{s_1(t), s_2(t), \ldots, s_n(t)\}$ denotes the explosion sparks number of n fireworks.

Now, an optimality region can be defined as follow.

Definition 2: $R_\varepsilon = \{x \in S|f(x) - f(x^*) < \varepsilon, \varepsilon > 0\}$ is named as the optimal region of function $f(x)$, where x^* represents the optimal solution of function $f(x)$ in the problem space.

In term of the Definition 2, if the algorithm finds a point in the optimal region, an acceptable approximation to the global minimum of the function has been acquired by the algorithm. According to the above definition of infimun,ψ, the Lebesgue measure of optimal solution space must be no zero, which means that $v(R_\varepsilon) > 0$ is held.

Definition 3: The optimal state of FWA is defined as $\xi^*(t) = \{F^*(t), T(t)\}$, where there exist $F_i(t) \in R_\varepsilon$ and $F_i(t) \in F^*(t), i \in 1, 2, \ldots, n$.

The definition 3 means that the best firework in the optimal state $\xi^*(t)$ of FWA is in the optimal region R_ε. So here exist $F_i(t) \in R$ and $|f(F_i(t)) - f(x^*)| < \varepsilon$, $x^* \in R_\varepsilon$.

Lemma 1: The stochastic process of FWA, v$\{\xi(t)\}_{t=0}^{\infty}$, is Markov stochastic process.

proof: $\{\xi(t)\}_{t=0}^{\infty}0$ is the stochastic process of discrete times, because the state $\xi(t) = \{F(t), T(t)\}$ is decided by the $\{F(t-1), T(t-1)\}$, so the probability $P\{\xi(t+1)|\xi(1), \xi(2), \ldots, \xi(t)\} = P\{\xi(t+1)|\xi(t)\}$, which means that the probability of $(t+1)$th state occurring is not related to the probability of tth state occurring. Therefore, the $\{\xi(t)\}_{t=0}^{\infty}$ is Markov stochastic process.

The proof is completed.

Definition 4: (optimal state space) Given Y is represented as the state space of FWA's state $\xi(t)$ and $Y^* \subset Y$. Y^* is named as the optimal state space if there exists a solution $s^* \in F^*$ such that $s^* \in R_\varepsilon$ for any state $\xi(t)^* = \{F^*, T\} \in Y$.

In the term of the above definition, it means that $|f(s^*) - f(x^*)| < \varepsilon$ for any $x^* \in F^*$. If the state of the Fireworks Algorithm can arrive at the optimal state, there exists a firework in the fireworks which stays in the optimal region R_ε and the optimal solution of problem has been acquired by the Fireworks Algorithm. After this time, the optimal solution must be in the optimal region forever.

Definition 5: Given a Markov stochastic process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$, if $\{\xi(t)\}_{t=0}^{\infty}$ s.t. $P\{\xi(t+1) \notin Y^*|\xi(t) \in Y^*\} = 0, \{\xi(t)\}_{t=0}^{\infty}$ is named as an absorbing Markov process.

Lemma 2: The stochastic process of FWA,$\{\xi(t)\}_{t=0}^{\infty}$,is an absorbing Markov stochastic process.

proof: According to the lemma 1, the stochastic process of FWA, $\{\xi(t)\}_{t=0}^{\infty}$, is a Markov stochastic process. If $F_1(t) \in F(t)$ stays in the optimal solution space R_ε, the state $\xi(t) = \{F(t), T(t)\}$ must belong to the optimal state space Y^*. Because $F_1(t)$ is the best location in all fireworks of FWA, $f(F_1(t+1))$ is not worse than $f(F_1(t))$ in term of the step 6 of FWA. So, the state $\xi(t+1)$ must belong to the optimal state space Y^*. Therefore, $P\{\xi(t+1) \notin Y^*|\xi(t) \in Y^*\} = 0$, the stochastic process of FWA, $\{\xi(t)\}_{t=0}^{\infty}$, is an absorbing Markov process.

The proof is completed.

IV. THE GLOBAL CONVERGENCE OF FIREWORKS ALGORITHM

In this section, the definition of convergence is given, which is used to analyze the convergence of Fireworks Algorithm.

Definition 6 :(convergence) Given an absorbing Markov process $\{\xi(t)\}_{t=0}^{\infty} = \{F(t), T(t)\}$ and an optimal state space $Y^* \subset Y, \lambda(t) = P\{\xi(t) \in Y^*\}$ denotes the probability that the stochastic state arrives at the optimal state in step t, if $lim_{t\to\infty}\lambda(t) = 1$, $\{\xi(t)\}_{t=0}^{\infty}$ is convergence.

In term of above definition, the convergence of Markov stochastic process depends on the probability of $P\{\xi(t) \in Y^*\}$. If it converges to 1 with the time t, the Markov process $\{\xi(t)\}_{t=0}^{\infty}$ will be convergence.

Theorem 1: Given an absorbing Markov process $\{\xi(t)\}_{t=0}^{\infty}$ of FWA and optimal state space $Y^* \subset Y$. If $P\{\xi(t) \in Y^*|\xi(t-1) \notin Y^*\} \geq d \geq 0$ for any t and $P\{\xi(t) \in Y^*|\xi(t-1) \in Y^*\} = 1$, then $P\{\xi(t) \in Y^*\} \geq 1 - (1-d)^t$.

proof: Let $t = 1, P\{\xi(1) \in Y^*\}$

$$= P\{\xi(1) \in Y^*|\xi(0) \in Y^*\} \cdot P\{\xi(0) \in Y*\}$$

$$+P\{\xi(1) \in Y^*|\xi(0) \notin Y^*\} \cdot P\{\xi(0) \notin Y^*\}$$

$$\geq P\{\xi(0) \in Y^*\} + d \cdot P\{\xi(0) \notin Y^*\}$$

$$= P\{\xi(0) \in Y^*\} + d \cdot (1 - P\{\xi(0) \in Y^*\})$$

$$= d + (1-d) \cdot P\{\xi(0) \in Y^*\}$$

Because $(1-d) \geq 0$,so $d + (1-d) \cdot P\{\xi(0) \in Y^*\} \geq d$, then $P\{\xi(1) \in Y^*\} \geq d = 1 - (1-d)^1$;

Now, it is assumed that the $P\{\xi(t) \in Y^*\} \geq 1 - (1-d)^t$ is held for any $t < k-1$, and then for $t = k$,

$$P\{\xi(k) \in Y^*\} = P\{\xi(k) \in Y^*|\xi(k-1) \in Y^*\}$$

$$\cdot P\{\xi(k-1) \in Y^*\} + P\{\xi(k) \in Y*|\xi(k-1) \notin Y^*\}$$

$$\cdot P\{\xi(k-1) \notin Y^*\}$$

$$= P\{\xi(k-1) \in Y^*\} + P\{\xi(k) \in Y^* | \xi(k-1) \notin Y^*\}$$

$$\cdot P\{\xi(k-1) \notin Y^*\}$$

$$\geq P\{\xi(k-1) \in Y^*\} + d \cdot (1 - P\{\xi(k-1) \in Y^*\})$$

$$= d + (1-d) \cdot P\{\xi(k-1) \in Y*\}$$

$$\geq d + (1-d) \cdot (1 - (1-d)^{k-1}) = 1 - (1-d)^k$$

Consequently, $P\{\xi(t) \in Y^*\} \geq 1 - (1-d)^t$ is true for any $t \geq 1$.

The proof is completed

According to the step 5 of Algorithm 1 on FWA, FWA has the operation of mutation. It is assumed for simplicity that the operation is the stochastic mutation.

Theorem 2: Given FWA an absorb state Markov process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \in Y$, then $lim_{t\to\infty}\lambda(t) = 1$ which means that $\xi(t)_{t=0}^{\infty}$ will converge to the optimal state Y^*.

Proof: In term of the step 5 of FWA in Algorithm 1, FWA can provide the mutation operator, so the probability that the firework of FWA arriving optimal region R_ε from non optimal region by mutation operator is denoted as $P_{mu}(t)$. It is expressed as follow:

$$P_{mu} = \frac{\nu(R_\varepsilon) \cdot n}{\nu(S)}.$$

where $\nu(S)$ is the Lebegue measure value of the problem space S, n is the number of fireworks.

Because $\nu(R_\varepsilon) > 0$, so $P_{mu} > 0$.

In term of the stochastic Markov process $\{\xi(t)\}_{t=0}^{\infty}$ of FWA, it holds that

$$\lambda(t) = P\{\xi(t) \in Y^* | \xi(t-1) \notin Y^*\} = P_{mu}(t) + P_{ex}(t).$$

where $P_{ex}(t)$ denotes the probability of the fireworks of FWA arriving the optimal region R_ε by the firework's explosion in FWA.

So, $P\{\xi(t) \in Y^* | \xi(t-1) \in Y^*\} \geq P(mu) > 0$.

Therefore, because the Markov process $\{\xi(t)\}_{t=0}^{\infty}$ of FWA is an absorbing Markov process while the condition of the theorem 1 is held, the following equation can be got.

$$P\{\xi(t) \in Y^*\} = 1 - (1 - P_{mu}(t))^t.$$

So, $limt_{t\to\infty} P\{\xi(t) \in Y^*\} = 1$.

Consequently, the Markov process $\{\xi(t)\}_{t=0}^{\infty}$ of FWA will converge to the optimal state.

The proof is completed.

V. The Basic Theory of Time Complexity on Fireworks Algorithm

On the analysis of evolution computation based on Marko Model, Huang Han and Hao Zhifeng have done for the evolutionary programming [15] and ant colony optimization [16]. In this section, the definitions of some conception and theorems on FWA are as follows which are referred to the literatures [15] and [16].

Definition 6: (expected convergence time) Given FWA an absorbing state Markov process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$, if γ is a stochastic nonnegative value such that: if $t \geq \gamma$, $P\{\xi(t+1) \in Y^*\} = 1$; if $0 \leq t \leq \gamma$, $P\{\xi(t+1) \notin Y^*\} < 1$, then the γ is named as the convergence time of FWA. The expected value Eγ is named as the expected convergence time of FWA.

The expected convergence time describes the expected time of arriving the global optimal solution with probability 1 at the first time. The smaller the expected value Eγ is, the faster the convergence of FWA is and the more effective FWA is. However,it can also use the expected first hitting time(EFHT) as a index of convergence time which is given as follows [9].

Definition 7:(expected first hitting time) Given FWA an absorbing state Markov process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$; μ is a stochastic value such that: if $t = \mu$, $\xi(t) \notin Y^*$; if $0 \leq t \leq \mu$, $\xi(t) \notin Y^*$. The expected value Eμ is named as Expected First Hitting Time.

The following theorem give the method to compute the Expected First Hitting Time,Eλ.

Theorem 3:Given FWA an absorbing state Markov process $\xi(t)_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$. If $\lambda(t) = P\{\xi(t) \in Y^*\}$ and $limt_{t\to\infty}\lambda(t) = 1$, the Expected Convergence Time is $E\gamma = \Sigma_{t=0}^{\infty}(1 - \lambda(t))$.

Proof:

$$\lambda(t) = P\{\xi(t) \in Y^*\} = P\{\mu \leq t\}$$
$$\Rightarrow \lambda(t) - \lambda(t-1) = P\{\mu \leq t\} - P\{\mu \leq t-1\}$$
$$\Rightarrow P\{\mu = t\} = \lambda(t) - \lambda(t-1),$$

then

$$E\mu = 0 \cdot P\{\mu = 0\} + \Sigma_{t=0}^{\infty} t \cdot P\{\mu = t\},$$
$$E\mu = \Sigma_{t=0}^{\infty} t \cdot (\lambda(t) - \lambda(t-1))$$
$$= \lambda(1) - \lambda(0)) + 2 \cdot (\lambda(2) - \lambda(1)) + \ldots$$
$$+ t \cdot (\lambda(t) - \lambda(t-1)) + \ldots$$
$$= \Sigma_{i=1}^{\infty}(\lambda(t) - \lambda(t-1)) + \Sigma_{i=2}^{\infty}(\lambda(t) - \lambda(t-1))$$
$$+ \ldots + \Sigma_{i=t}^{\infty}(\lambda(t) - \lambda(t-1)) + \ldots$$
$$= (lim_{t\to\infty}\lambda(t) - \lambda(1)) + (lim_{t\to\infty}\lambda(t) - \lambda(2))$$
$$+ \ldots + (lim_{t\to\infty}\lambda(t) - \lambda(t-1)) + \ldots$$
$$= \Sigma_{i=1}^{\infty}(limt_{t\to\infty}\lambda(t) - \lambda(t-1)) = \Sigma_{i=1}^{\infty}(1 - \lambda(t-1))$$
$$= \Sigma_{i=1}^{\infty}(1 - \lambda(t))$$

So,

$$E\gamma = E\mu = \Sigma_{i=1}^{\infty}(1-\lambda(t)).$$

The proof is completed

According to the theorem 1, it is difficult to compute the expected convergence time Eγ becouse it is hard to acquire the value of $\lambda(t)$,. So, its estimation is given as follows. The proof of following lemmas theorems can been referred to [16].

Lemma 3: Given two stochastic nonnegative variable, μ and ν, and $D_u(.)$ and $D_v(.)$ denote the distribution functions of μ and ν,respectively. The expected value of μ and ν can hold $E\mu < E\nu$ if $D_u(.) \geq D_v(.)$ for $t = 0, 1, 2, \ldots$.

Theorem 4: Given FWA an absorbing state Markov process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$. If $\lambda(t) = P\{\xi(t) \in Y^*\}$ such that $0 \leq D_l(t) \leq \lambda(t) \leq D_h(t) \leq 1 (\forall t = 0, 1, 2, \ldots)$ and $limt_{t\to\infty}\lambda(t) = 1$, then:

$$\Sigma_{i=1}^{\infty}(1 - D_l(t)) \leq E\gamma \leq \Sigma_{i=1}^{\infty}(1 - D_t(t)).$$

Theorem 5: Given FWA an absorbing state Markov process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$; if $\lambda(t) = P\{\xi(t) \in Y^*\}$ and $0 \leq a(t) \leq \lambda(t) \leq b(t)$, $\Sigma_{l=1}^{\infty}[(1 - \lambda(0))\Pi_{i=0}^{\infty}(1 - a(t))] \leq E\gamma \leq \Sigma_{t=0}^{\infty}[(1 - \lambda(0))\Pi_{i=1}^{\infty}(1 - a(t))]$.

Corollary 1: Given FWA an absorbing state Markov process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$; and $\lambda(t) = P\{\xi(t) \in Y^*\}$. If $a \leq P\{\xi(t+1) \in Y^* | \xi(t+1) \notin Y^*\} \leq b(a, b > 0)$ and $lim_{t\to\infty}\lambda(t) = 1$, then the expected convergence time $E\gamma$ of FWA is such that:

$$b^{-1}[1 - \lambda(0)] \leq E\gamma \leq a^{-1}[1 - \lambda(0)].$$

The above corollary and theorems indicate that the formula $P\{\xi(t) \in Y^* | \xi(t-1) \notin Y^*\}$ can give description of the probability of arriving at the optimal state from the non optimal state. The estimation of value range of Eλ is able to be computed by the range of value of $P\{\xi(t) \in Y^* | \xi(t-1) \notin Y^*\}$.

VI. The analysis of time complexity on Fireworks Algorithm

The time complexity of FWA is used to define the expected convergence time Eλ. In term of the *Corollary* 1 in previous section, it is mainly related to the probability of the FWA state arriving optimal region R_ε from non optimal region which is the formula, $P\{\xi(t+1) \in Y^* | \xi(t-1) \notin Y^*\}$. In the section, we will further analyze the formula to get the time complexity of FWA. FWA includes three operations: explosion, mutation and selection, but the operations which directly make the Markov state of FWA get to the optimal region are explosion and mutation, so the following theorem is given.

Theorem 6: Given FWA an absorbing state Markov process $\{\xi(t)\}_{t=0}^{\infty}$ and optimal state space $Y^* \subset Y$, then FWA is such that:

$$\begin{aligned} \frac{\nu(R_\varepsilon) \times n}{\nu(S)} &\leq P\{(\xi(t+1)) \in Y^* | \xi(t) \notin Y^*\} \\ &\leq \nu(R_\varepsilon)\left(\frac{n}{\nu(S)} + \sum_{i=1}^{n} \frac{m_i}{\nu(\mathrm{A}_i)}\right). \end{aligned} \tag{7}$$

where $\nu(R_\varepsilon)$ is the Lebegue measure value of the optimal regionR_ε. The $\nu(S)$ is the Lebegue measure value of the problem search region S, $v(A_i)$ is the Lebegue measure value of the explosion region A_i of ith firework.

Proof: In term of the step of FWA, the FWA includes two operations to generate the sparks: explosion operator and mutation operator. So the following equation is got:

$$P(\xi(t+1) \in Y^* | \xi(t) \notin Y^*) = P_{mu} + P_{ex}. \tag{8}$$

where P_{mu} denotes the probability that the fireworks of FWA arrive optimal region R_ε from non optimal region by mutation operator. P_{ex} is the probability which explosion operator of n fireworks make some sparks stay optimal region R_ε.

For P_{mu}, it is assumed that the mutation operator operation is randomness with uniform distribution. The probability of one firework being mutated to the optimal region $R_\varepsilon$is:

$$\frac{\nu(\mathrm{R}_\varepsilon)}{\nu(\mathrm{S})}.$$

So the probability of n fireworks to be randomly mutated to the optimal region R_ε,$P_{mu}(t)$, is:

$$\frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})}.$$

That means:

$$P_{mu} = \frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})}.$$

For the P_{ex} , the probability that ith firework explodes and make sparks stay in the optimal range R can be derived as follows:

$$\frac{\nu(A_i \cap R_\varepsilon) \times m_i}{\nu(A_i)}.$$

So, the probability which explosion of n fireworks make some sparks stay optimal region R_ε,P_{ex},is:

$$P_{ex} = \sum_{i=1}^{n} \frac{\nu(A_i \cap R_\varepsilon) \times m_i}{\nu(A_i)}.$$

where A_i denotes the search space in which the ith firework explodes; m_i is the number of spark which the ith firework generates.

Therefore, it is derived that:

$$\begin{aligned} P(\xi(t+1) \in Y^* | \xi(t) \notin Y^*) &= \frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})} \\ &+ \sum_{i=1}^{n} \frac{\nu(A_i \cap R_\varepsilon) \times m_i}{\nu(A_i)}. \end{aligned} \tag{9}$$

Because $0 \leq \nu(\mathrm{A}_i \cap R_\varepsilon) \leq \nu(R_\varepsilon)$,

$$0 \leq P_{ex} = \sum_{i=1}^{n} \frac{\nu(\mathrm{A}_i \cap R_\varepsilon) \times m_i}{\nu(\mathrm{A}_i)} \leq \sum_{i=1}^{n} \frac{\nu(R_\varepsilon) \times m_i}{\nu(\mathrm{A}_i)} = \nu(R_\varepsilon) \sum_{i=1}^{n} \frac{m_i}{\nu(\mathrm{A}_i)}.$$

And then,

$$\frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})} \leq P(\xi(t+1) \in Y^* | \xi(t) \notin Y^*) \leq \frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})} + \nu(R_\varepsilon) \sum_{i=1}^{n} \frac{m_i}{\nu(\mathrm{A}_i)} = \nu(R_\varepsilon) \left(\frac{n}{\nu(\mathrm{S})} + \sum_{i=1}^{n} \frac{m_i}{\nu(\mathrm{A}_i)} \right).$$

It is got that

$$\frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})} \leq P(\xi(t+1) \in Y^* | \xi(t) \notin Y^*) \leq \nu(R_\varepsilon) \left(\frac{n}{\nu(\mathrm{S})} + \sum_{i=1}^{n} \frac{m_i}{\nu(\mathrm{A}_i)} \right).$$

The proof is completed.

The above theorem gives the rude result because the right formula of the Eq.(7) is difficult to be confirmed and be computed. It is complex for FWA to run and hard to compute the probability about it. In order to realize the probability exactly, the Eq.(8) need to be further investigated which is given as follows.

$$P_{ex} = \sum_{i=1}^{n} \frac{\nu(S_i \cap R_\varepsilon) \times m_i}{\nu(S_i)}.$$

As it can be known, the formulas $\nu(S_i \cap R_\varepsilon)$ and m_i in the above equation play key role to the P_{ex} because the two formulas is dynamically changed with the algorithm running.

The formula $\nu(S_i \cap R_\varepsilon)$ is related to the firework location Fi. In term of Eq.(4) and Eq.(5), the distance between two of fireworks selected as the next generation is as far as possible, so it can be assumed that just one firework can stay in the optimal region R_ε in the same time. In other hand, it is further assumed that there is the highest probability for the best firework to get into the optimal region R_ε.

According to the above idea of Fireworks Algorithm, $\nu(A_i) \geq \nu(A_{best})$ and $m_i \leq m_{best}$,$i \in (1, 2, \dots n)$, where A_{best} and m_{best} is the exploding region and generating sparks number of the firework which fitness is best in the all fireworks, respectively. So, it can be derived as the follows:

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

It can be considered that$(A_i \cap R_\varepsilon) \cap (A_{best} \cap R_\varepsilon) = \phi$ for $i \in (1, 2, \dots, n)$ and $i \neq best$, especially in the early running time, so the following equation can be derived.

$$P(exp) = \sum_{i=1}^{n} \frac{\nu(S_i \cap R_\varepsilon) \times m_i}{\nu(S_i)} < \frac{\nu(S_{best} \cap R_\varepsilon) \times m_{best}}{\nu(S_{best})} < \frac{\nu(R_\varepsilon) \times m_{best}}{\nu(S_{best})}. \tag{10}$$

So the Eq.(7) can be changed to the follows:

$$\frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})} \leq P(\xi(t+1) \in Y^* | \xi(t) \notin Y^*) \leq \nu(R_\varepsilon) \left(\frac{n}{\nu(\mathrm{S})} + \frac{m_{best}}{\nu(S_{best})} \right). \tag{11}$$

The above Eq.(11) is more meaningful than the Eq.(7), which tells that the best firework is more important. In term of the Eq.(11) and Corollary 1, let $a = \frac{\nu(R_\varepsilon) \times n}{\nu(\mathrm{S})}$ and $b = \nu(R_\varepsilon) \left(\frac{n}{\nu(\mathrm{S})} + \frac{m_{best}}{\nu(S_{best})} \right)$, so the following equation is derived:

$$\frac{\nu(S) \times \nu(S_{best})}{\nu(R_\varepsilon) \times (n \times \nu(S_{best}) + m_{best} \times \nu(S))} \times (1 - \lambda(0)) \leq E\gamma \leq \frac{\nu(S)}{\nu(R_\varepsilon) \times n} \times (1 - \lambda(0)).$$

where $\lambda(t) = P\{\xi(t) \in Y^*\}$.

According to the Step1 of FWA, the initialization of n fireworks is generated in random. The following results can be acquired. Because $\lambda(0) = P\{\xi(0) \in Y^*\} << 1$, $1 - \lambda(0) = 1$; then

$$\frac{\nu(S) \times \nu(S_{best})}{\nu(R_\varepsilon) \times (n \times \nu(S_{best}) + m_{best} \times \nu(S))} \leq E\gamma \leq \frac{\nu(S)}{\nu(R_\varepsilon) \times n}. \tag{12}$$

Corollary 2: Fireworks Algorithm's expected convergence time Eλ such that:

$$\frac{\nu(S) \times \nu(S_{best})}{\nu(R_\varepsilon) \times (n \times \nu(S_{best}) + m_{best} \times \nu(S))} \leq E\gamma \leq \frac{\nu(S)}{\nu(R_\varepsilon) \times n}. \tag{13}$$

From the Eq.(12), the more lager value of R_ε and the smaller value of $\nu(S)$ are beneficia to the efficiency of FWA, but the two values are related to the search problem.The Eq.(12) indicates that the $\nu(S_{best})$ and m_{best} is very important for FWA's expected convergence time. But above some results are got under the condition of some assumptions. The more exact analysis need to be further done through considering the detail of some equations about originate FWA.

VII. Conclusions

Fireworks Algorithm is a novel swarm intelligence algorithm which generates the explosion sparks and Gaussian mutation sparks by the explosion operator and Gaussian mutation operator to search the problem space thus can be applied to solve practical optimization problems. However, little theoretical analysis or work on FWA has been done. Same as the other swarm intelligence algorithms, the optimization process of Fireworks Algorithm is considered as a Markov process. The time complexity of FWA can be analyzed using an absorbing Markov process. This paper has presented some conceptions on Markov stochastic process of FWA and has proven its global convergence. Moreover, we also present the approximate region of expected convergence time of FWA. Although the results given in the paper are incomplete and naive, its theorem analysis can provide a direction on the theory of Fireworks Algorithm.

REFERENCES

[1] Tan Y. and Zhu Y. C.,Fireworks Algorithms for Optimization, *Proceeding of International Conference on Swarm Intelligence (ICSI2010), Part II*,LNCS 6145, pp.355-364,2010.

[2] Janecek Andreas and Ying Tan," Using population based algorithms for initializing nonnegative matrix factorization," *Advances in Swarm Intelligence*, Springer Berlin Heidelberg, 307-316,2011.

[3] Janecek Andreas and Ying Tan, "Iterative improvement of the multiplicative update nmf algorithm using nature-inspired optimization,"*Seventh International IEEE Conference on Natural Computation (ICNC)*, Vol.3, 2011.

[4] Pei Yan et al, "An empirical study on influence of approximation approaches on enhancing fireworks algorithm," *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2012.

[5] Gao Hongyuan and Ming Diao, "Cultural firework algorithm and its application for digital filters design," *International Journal of Modelling, Identification and Control*,Vol.14 , pp. 324-331, 2011.

[6] Zheng Yu-Jun, Qin Song and Sheng-Yong Chen, "Multiobjective fireworks optimization for variable-rate fertilization in oil crop production," *Applied Soft Computing*, Vol.13,pp. 4253-4263,2013.

[7] Liu Jianhua, Shaoqiu Zheng and Ying Tan. "The Improvement on Controlling Exploration and Exploitation of Firework Algorithm," *Advances in Swarm Intelligence. Springer Berlin Heidelberg*, pp. 11-23, 2013.

[8] He Wenrui, Guyue Mi and Ying Tan, "Parameter Optimization of Local-Concentration Model for Spam Detection by Using Fireworks Algorithm," *Advances in Swarm Intelligence. Springer Berlin Heidelberg*,pp.439-450,2013.

[9] He Jun and Xin Yao, "From an individual to a population: An analysis of the first hitting time of population-based evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, Vol 6, pp.495-511,2002.

[10] Yu Yang, and Zhi-Hua Zhou. "A new approach to estimating the expected first hitting time of evolutionary algorithms." Artificial Intelligence 172.15 (2008): 1809-1832.

[11] Huang Han, Chun-Guo Wu and Zhi-Feng Hao, "A pheromone-rate-based analysis on the convergence time of ACO algorithm," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol.39,pp. 910-923, 2009.

[12] Chen Tianshi et al, "Analysis of computational time of simple estimation of distribution algorithms,"*IEEE Transactions on Evolutionary Computation*, Vol.14,pp.1-22, 2010.

[13]]Yi Shengqiu, Ming Chen and Zhigao Zeng, "Convergence analysis on a class of quantum-inspired evolutionary algorithms," *Seventh IEEE International Conference on Natural Computation (ICNC)*, Vol. 2., 2011.

[14] Ding Lixin and Jinghu Yu, "Some techniques for analyzing time complexity of evolutionary algorithms," *Transactions of the Institute of Measurement and Control*, Vol.34 pp.755-766,2012.

[15] Huang Han, Hao Zhifeng and Qin Yong, "Time Complexity of Evolutionary Programming", *Journal of Computer Research and Development*, Vol.45(11),pp.1850-1857,2008.(In Chinese).

[16] Huang Han, Hao Zhifeng,Wu ChunGuo and Qin Yong, "Time Convergence Speed of Ant Colony Optimization," *Chinese Journal of Computer*, Vol.30(8),pp.1344-1353,2007.(In Chinese).