Accelerating the Fireworks Algorithm with an Estimated Convergence Point

Yu, Jun
Graduate School of Design, Kyushu Unievrity

Takagi, Hideyuki
Faculty of Design, Kyushu University

Tan, Ying
School of Electronics Engineering and Computer Science, Peking University

Accelerating the Fireworks Algorithm with an Estimated Convergence Point

Jun YU[1], Hideyuki TAKAGI[2], Ying TAN [3]

[1]Graduate School of Design, Kyushu University, Fukuoka, Japan 815-8540

[2]Faculty of Design, Kyushu University, Fukuoka, Japan 815-8540

[3]School of Electronics Engineering and Computer Science, Peking University, Beijing, China 100871

Abstract

We propose an acceleration method for the fireworks algorithms which uses a convergence point for the population estimated from moving vectors between parent individuals and their sparks. To improve the accuracy of the estimated convergence point, we propose a new type of firework, the *synthetic firework*, to obtain the correct of the local/global optimum in its local areas fitness landscape. The synthetic firework is calculated by the weighting moving vectors between a firework and each of its sparks. Then, they are used to estimate a convergence point which may replace the worst firework individual in the next generation. We design a controlled experiment for evaluating the proposed strategy and apply it to 20 CEC2013 benchmark functions of 2-dimensions (2-D), 10-D and 30-D with 30 trial runs each. The experimental results and the Wilcoxon signed-rank test confirm that the proposed method can significantly improve the performance of the canonical firework algorithm.

Keywords— fireworks algorithm, estimated convergence point, synthetic individuals, acceleration

1   Introduction

The fireworks algorithm (FWA) [1], as a new member of swarm intelligence algorithms inspired by the explosion of real fireworks, has attracted much attentions in academia and industry. It simulates explosions repeatedly to implement local search points (sparks) around a specific point (firework) and evolves towards the optimal solution. Many improved versions of FWA have been proposed. The enhanced FWA (EFWA) [2] improves the corresponding operations of the original FWA and can achieve a better performance. Dynamic FWA (dynFWA) [3] uses a dynamic explosion amplitude for the current best firework to tune the search range more intelligently. An amplitude reduction strategy and local optima-based selection strategy [4] were also proposed to improve the performance of FWA obviously. Although many new ideas and mechanisms have been introduced to FWA to develop new variations, little attention was given to the generated sparks, which therefore offer a potential new direction for research.

Using gradient information has always been a very hot topic full of potential. Many practitioners have tried to build and use gradients to accelerate convergence. For example, [5] estimates the natural gradient for the exponential family based on regularized linear regression. In addition, [6] proposes an alternative way to compute search directions by exploiting neighborhood information. In this paper, we introduce a new type of firework, the synthetic firework. Using gradient information derived from the generated sparks, we can gain an understanding about the direction of local evolution on the fitness landscape. This local gradient information is then used to estimate a convergence point for the fireworks population.

The main objective of this paper is to use the estimated convergence point as an elite individual to accelerate FWA by substituting it for the worst firework individual in next generation if its fitness is better. The secondary one is to analyze the applicability of the proposed strategy, and introduce some topics which are open to discussion.

We introduce the framework of canonical FWA in Section 2.1 and a method for estimating the convergence point in Section 2.2. New types of fireworks are described in detail in Section 3. We evaluate them by comparing them with the original FWA using 20 benchmark functions of 3 different dimensions in Section 4. Finally, we discuss the experimental evaluations in Section 5 and conclude in Section 6.

2   Related Research

2.1   Fireworks Algorithm

Real fireworks are launched into the sky, and many sparks are generated around the fireworks. The explosion process of a firework can be viewed as a local search around a specific point. FWA simulates this explosion process

iteratively to find the optimal solution. Fig. 1 illustrates the process of FWA, which consists principally of three operations: explosion, mutation and selection [1].
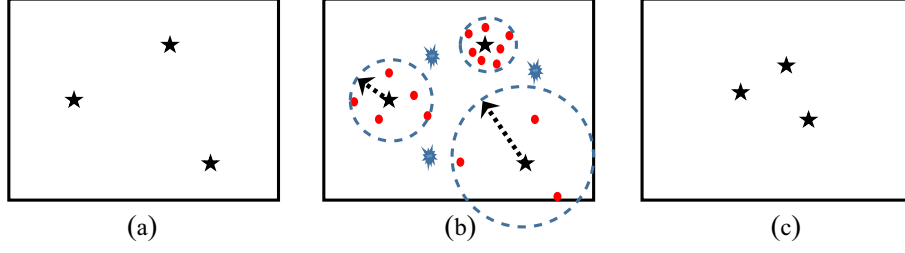

Figure 1: The search process of FWA. (a) fireworks are generated, (b) sparks are created around each firework, and mutation points are also generated, (c) new fireworks are created in the next generation using the sparks from (b). Steps (b) and (c) are iterated until the termination condition is satisfied.

Since there are some limitations in classic FWA and its performance is also not very prominent among all its subsequent variants, such as EFWA and dynFWA, we employ the more powerful EFWA [2] as our baseline algorithm and combine it with our proposed strategy. The EFWA introduces five major improvements into conventional FWA to improve its performance. For details on these improvements, refer to [2].

2.2 Method for Estimating the Convergence Point for a Population

The convergence point for the moving vectors between parent individuals and their offspring in the next EC search generation can be calculated mathematically [7, 8]. Let us begin by defining symbols. $\boldsymbol{a}_i$ and $\boldsymbol{c}_i$ in the Fig. 2 are the i-th parent individual and its offspring individual, respectively ($\boldsymbol{a}_i, \boldsymbol{c}_i \in \mathbb{R}^d$). The i-th moving vector is defined as the direction vector, $\boldsymbol{b}_i = \boldsymbol{c}_i - \boldsymbol{a}_i$. The unit direction vector of the $\boldsymbol{b}_i$ is given as $\boldsymbol{b}_{0i} = \boldsymbol{b}_i/\|\boldsymbol{b}_i\|$, i.e. $\boldsymbol{b}_{0i}^T \boldsymbol{b}_{0i} = 1$.
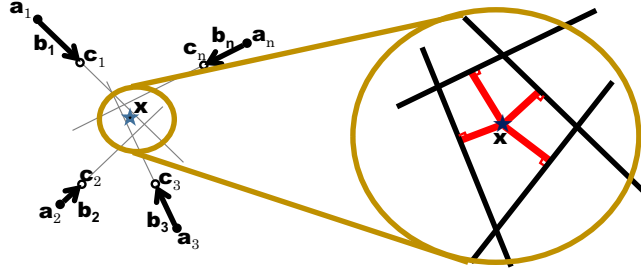
Figure 2: The moving vector $\boldsymbol{b}_i$ ($= \boldsymbol{c}_i - \boldsymbol{a}_i$) is calculated from a parent individual $\boldsymbol{a}_i$ and its offspring $\boldsymbol{c}_i$ in the d-dimensional searching space. The $\star$ mark is the convergence point for these moving vectors.

Let $\boldsymbol{x} \in \mathbb{R}^d$ be the point that is the nearest to the n extended directional line segments, $\boldsymbol{a}_i + t_i \boldsymbol{b}_i$ ($t_i \in R$). The *nearest*, means that the total distance from $\boldsymbol{x}$ to the n extended directional line segments, $\mathrm{J}(\boldsymbol{x}, \{t_i\})$ in Eq.(1), becomes the minimum. We may insert an orthogonality condition, Eq. (2), into Eq. (1) and thus remove t_i.

$$\mathrm{J}(\boldsymbol{x}, \{t_i\}) = \sum_{i=1}^{n} \|\boldsymbol{a}_i + t_i \boldsymbol{b}_i - \boldsymbol{x}\|^2 \tag{1}$$

$$\boldsymbol{b}_i^{\mathrm{T}}(\boldsymbol{a}_i + t_i \boldsymbol{b}_i - \boldsymbol{x}) = 0 \quad \text{(orthogonal condition)} \tag{2}$$

The $\hat{\boldsymbol{x}}$ that minimizes the total distance in the Eq. (1) is obtained by partially differentiating each element of $\boldsymbol{x}$ and setting them equal 0. Finally, the convergence point $\hat{\boldsymbol{x}}$ is given by Eq. (3), where $\boldsymbol{I}_d$ is the unit matrix.

$$\hat{\boldsymbol{x}} = \left\{ \sum_{i=1}^{n} \left(\boldsymbol{I}_d - \boldsymbol{b}_{0i} \boldsymbol{b}_{0i}^{\mathrm{T}} \right) \right\}^{-1} \left\{ \sum_{i=1}^{n} \left(\boldsymbol{I}_d - \boldsymbol{b}_{0i} \boldsymbol{b}_{0i}^{\mathrm{T}} \right) \boldsymbol{a}_i \right\} \tag{3}$$

3 Proposed Method

We introduce a new kind of firework, named the synthetic firework, to make full use of the many generated sparks, which are otherwise only involved in the selection operation and then destroyed. The synthetic fireworks and fireworks of the current generation form many moving vectors which can be used to estimate a convergence point

that is expected to locate near the global optimum. The estimated point is regarded as an elite individual and replaces the worst individual from the next generation if its fitness is better.

The method for calculating the synthetic fireworks is as follows. Each firework and its generated sparks form a subgroup, and we can construct many vectors between the firework and its generated sparks. If the firework is worse than one of the generated sparks, this vectors direction is considered to be promising. Otherwise, its antipode is used to calculate a synthetic firework. There are many methods to evaluate the potential of these directions. In this paper, we simply use the fitness difference between the endpoint and the start point of a vector to evaluate it. Thus, the larger the fitness difference is, the higher will be the weight of the vector. In order to not increase the number of fitness evaluations, we only calculate the antipode for a firework which is lacking a fitness evaluation if the antipodal direction is to be used. The fitness difference of the original vector is roughly used to evaluate the used antipodal direction. Finally, a synthetic firework can be roughly calculated by weighting those vectors with Eq. 4 in each firework group. Fig. 3 illustrates how a synthetic firework is thus formed.
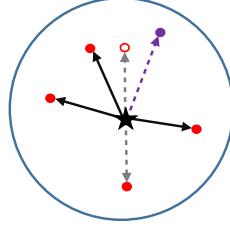

Figure 3: A synthetic firework is generated from a firework and its generated sparks. The black five-pointed star and the red solid points represent the firework and its generated sparks, respectively. The presence of a red hollow circle means that the antipode has been used. The purple solid point is the synthetic firework obtained by weighting these vectors.

$$v^i = \sum_{j=1}^{m} \frac{f(x^i) - f(s_j^i)}{\sum_{i=1}^{n} ||(f(x^i) - f(s_j^i))||} * (s_j^i - x^i) + x^i \tag{4}$$

Where x^i and s_j^i represent the i-th firework and its j-th generated spark or antipodal point. v^i is the i-th synthetic firework of the i-th firework group; m is the number of generated sparks of the i-th firework; $f()$ is a fitness function.

We can obtain new synthetic fireworks up to the population of the the current firework generation. Since we do not increase the number of fitness evaluations and a new synthetic firework is expected to be better than the firework belonging to its subgroup, we will not evaluate the synthetic fireworks. A moving vector is calculated from the current firework to the newly generated synthetic firework in each subgroup, and the convergence point is estimated using these moving vectors with the estimation method described in the Section 2.2. Algorithm 1 outlines the flow of EFWA using our proposed strategy.

Algorithm 1 The framework for the fireworks algorithm using our proposed strategy. Steps 11 to 16 are from our proposal.

1: Initialize n fireworks randomly.
2: Evaluate the fitness of each firework.
3: **while** the termination condition is not satisfied **do**
4:    Generate explosion sparks around each firework.
5:    Use Gaussian mutation to obtain Gauss sparks.
6:    **if** sparks are generated outside the search area **then**
7:       Use a mapping rule to bring them back into the area.
8:    **end if**
9:    Evaluate the fitness of each generated spark.
10:    Select n fireworks for the next generation from the generated sparks and the current fireworks.
11:    Calculate the synthetic fireworks for each subgroup.
12:    Obtain moving vectors using the synthetic fireworks and the current fireworks.
13:    Estimate a convergence point.
14:    **if** the estimated convergence point is better than the worst firework in the next generation **then**
15:       Replace the worst firework with the estimated point.
16:    **end if**
17: **end while**
18: end of program.

Note that our proposed strategy does not change the structure of the original FWA when it is combined with other fireworks algorithms. It simply uses the fireworks and the generated sparks to build local gradient information, then uses this to estimate a convergence point to accelerate convergence.

4 Experimental Evaluations

We use 20 benchmark functions from the CEC2013 benchmark test suite [9] in our evaluations, which is designed for real parameter single-objective optimization. Table 1 shows their types, characteristics, variable ranges, and optimum fitness values. These landscape characteristics include shifted, rotated, global on bounds, unimodal and multi-modal. We test them with 3 dimensional settings: $D = 2$, 10 and 30. We select EFWA [2] as our test baseline and combine it with our proposal for this experiment using parameters as described in table 2, where the definition of the symbols can be found in the original literature [1] [2].

For fair evaluations, we evaluate convergence against the number of fitness calls rather than generations. We test each benchmark function with 30 trial runs in 3 different dimensional spaces. We apply the Wilcoxon signed-rank test on the fitness values at the stop condition, i.e. the maximum number of fitness calculations, and compare EFWA with (EFWA + our proposed method). Table 3 shows the result of these statistical tests.

Table 1: Benchmark Function: Uni=unimodal, Multi=multimodal.

No.	Types	Characteristics	Ranges	Optimum fitness value
F_1		Sphere function		-1400
F_2		Rotated high conditioned elliptic function		-1300
F_3	Uni	rotated Bent Cigar function	$[-100, 100]$	-1200
F_4		Rotated discus function		-1100
F_5		different powers function		-1000
F_6		Rotated Rosenbrock's function		-900
F_7		Rotated Schaffers function		-800
F_8		Rotated Ackley's function		-700
F_9		Rotated Weierstrass function		-600
F_{10}		Rotated Griewank's function		-500
F_{11}		Rastrigin's function		-400
F_{12}	Multi	Rotated Rastrigin's function	$[-100, 100]$	-300
F_{13}		Non-continuous rotated Rastrigin's function		-200
F_{14}		Schwefel's function		-100
F_{15}		Rotated Schwefel's function		100
F_{16}		Rotated Katsuura function		200
F_{17}		Lunacek BiRastrigin function		300
F_{18}		Rotated Lunacek BiRastrigin function		400
F_{19}		Expanded Griewank's plus Rosenbrock's function		500
F_{20}		Expanded Scaffer's F_6 function		600

Table 2: Parameter setting of EFWA.

Parameters	Values
# of fireworks for 2-D, 10-D and 30-D search	5
# of sparks m	50
# of Gauss mutation sparks,	5
constant parameters	$a = 0.04\ b = 0.8$
Maximum amplitude A_{max}	40
stop condition; MAX_{NFC}, for 2-D, 10-D, and 30-D search	1,000, 10,000, 40,000
dimensions of benchmark functions, D	2, 10, and 30
# of trial runs	30

Table 3: Statistical test results of the Wilcoxon signed-rank test for average fitness values of 30 trial runs of the proposal (EFWA + our proposed method) and conventional method (EFWA) at the stop condition, MAX_{NFC}. $A \gg B$ and $A > B$ mean that A is significant better than B with significant levels of 1% and 5%, respectively. $A \approx B$ means that although A is better than B, there is no significant difference between them.

$Func.$	2-D	10-D	30-D
f_1	proposal $\gg$ EFWA	proposal $\gg$ EFWA	proposal $\gg$ EFWA
f_2	proposal $\approx$ EFWA	proposal $\gg$ EFWA	proposal $\gg$ EFWA
f_3	proposal $\approx$ EFWA	proposal $>$ EFWA	proposal $>$ EFWA
f_4	EFWA $\approx$ proposal	proposal $\approx$ EFWA	proposal $\approx$ EFWA
f_5	proposal $\gg$ EFWA	proposal $\gg$ EFWA	proposal $\gg$ EFWA
f_6	proposal $\approx$ EFWA	EFWA $\approx$ proposal	proposal $\approx$ EFWA
f_7	proposal $>$ EFWA	EFWA $\approx$ proposal	proposal $\approx$ EFWA
f_8	proposal $\approx$ EFWA	EFWA $\approx$ proposal	proposal $\approx$ EFWA
f_9	EFWA $\approx$ proposal	EFWA $\approx$ proposal	EFWA $\approx$ proposal
f_{10}	proposal $>$ EFWA	proposal $\gg$ EFWA	proposal $\gg$ EFWA
f_{11}	proposal $\approx$ EFWA	proposal $\gg$ EFWA	proposal $\gg$ EFWA
f_{12}	EFWA $\approx$ proposal	proposal $\approx$ EFWA	proposal $\approx$ EFWA
f_{13}	proposal $\approx$ EFWA	proposal $\approx$ EFWA	proposal $\approx$ EFWA
f_{14}	proposal $\approx$ EFWA	proposal $>$ EFWA	proposal $\gg$ EFWA
f_{15}	proposal $\approx$ EFWA	proposal $\approx$ EFWA	EFWA $\approx$ proposal
f_{16}	EFWA $\approx$ proposal	proposal $\approx$ EFWA	proposal $\approx$ EFWA
f_{17}	proposal $\gg$ EFWA	proposal $\gg$ EFWA	proposal $\gg$ EFWA
f_{18}	proposal $>$ EFWA	proposal $\approx$ EFWA	EFWA $\approx$ proposal
f_{19}	proposal $\approx$ EFWA	proposal $\approx$ EFWA	proposal $\gg$ EFWA
f_{20}	proposal $\approx$ EFWA	proposal $\approx$ EFWA	EFWA $\approx$ proposal

5   Discussions

Most fireworks algorithm variants mainly use their computational resources for generating sparks, but the information from these sparks is not fully used. In our experimental evaluations, the total number of generated sparks was 10 times of that of the fireworks. It is clearly productive to consider how these many sparks can be used efficiently. We introduced a new type of firework, called the synthetic firework, to explore local gradient information on the fitness landscape. Thanks to the use of multiple vectors in each subgroup, the synthetic firework also has a certain anti-noise property, as its calculation cancels noise from the directions of the various moving vectors. This can help to improve the precision of the estimated convergence point. In any case, the proposed method increases each generations fitness calculations by only one - so we can say that it is a low risk, high return strategy.

What potential still remains for our proposed firework, the synthetic firework? Although we have used only the fitness difference between the two endpoints of a moving vector to evaluate it, we think that not only these fitness differences but also their lengths should be considered to understand the local gradient information more accurately, yielding further improvements in the estimate. Additionally, there are many other ways to weight moving vectors and increase the precision of the estimated convergence point. As an example, the fitness value at the beginning point or the end point of a moving vector can be used to evaluate it, which means that the lower the distance from the optimal area, the higher the weight given. A precise way of obtaining reasonable weights for the vectors is also a potential discussion topic.

We would like to point out that the new type of firework introduced can be used to speed up convergence. In this paper, we used synthetic fireworks to estimate a convergence point without evaluating their fitness. They have the potential to act as a new guide for individuals, helping move them toward a preferable evolutionary direction rather than random exploration. The new synthetic fireworks can be introduced into a population to improve the diversity and reduce selection pressure. How to use them reasonably is also a potential discussion topic.

We also performed an extra experiment to investigate the fitness of synthetic fireworks. We compared the synthetic firework with the firework individual belonging to its same subgroup. The experimental results show that in the early stages, synthetic fireworks are better than fireworks individuals, while the probability of a better synthetic firework decreases as the convergence progresses. For optimization problems with different characteristics, it seems reasonable to use a different method for assigning weights when creating the synthetic fireworks. Perhaps different optimization stages could use different weighting methods to obtain better synthetic fireworks. Summarizing the relationship between weighting method and optimization problem is thus also a potential topic for study.

From the results of the statistical tests, we find that the proposed method is beneficial for unimodal optimization problems ($f_1 - f_5$), while the performance on low-dimensional multimodal optimization problems is not obvious. This may be because the basic estimation method, which is clearly effective for unimodal optimization problems, is not always valid for multimodal problems where the moving vectors go toward different local optima. Further, the number of moving vectors is small (in this case, the number is 5), and even on some multimodal optimization problems, it is less than the number of peaks. Regardless, the proposed strategy does not show any deleterious effect. [10] confirmed the effectiveness of using an extra individual pool to preserve outstanding individuals from past generations and using this pool, instead of the current generation, to estimate convergence points. For the next stage, using past searching individuals to increase the number of moving vectors, and combining it with the clustering method may allow us to extend our proposal to multimodal optimization problems.

6   Conclusion

We propose a new kind of fireworks which uses the generated sparks to efficiently estimate a convergence point which can act as an elite individual to accelerate the fireworks algorithm. The controlled experiments confirm that the proposed strategy can significantly improve the performance of conventional EFWA, and the higher the dimension, the more obvious the effect.

In future work, we will further study the proposed synthetic fireworks and use them to beneficially guide the evolution of the population. Additionally, it is suggested that we can further improve the accuracy of the estimated point by using historical information to better understand the fitness landscape.

Acknowledgment

References

[1] Tan, Y. and Zhu, Y.C., "Fireworks algorithm for optimization," Advances in Swarm Intelligence, vol. 6145 of the series Lecture Notes in Computer Science, pp. 355–364 (2010).

[2] Zheng, S.Q., Janecek, A., and Tan, Y., "Enhanced Fireworks Algorithm," IEEE Congress on Evolutionary Computation, Cancun, Mexico, pp. 2069–2077 (June, 2013).

[3] Zheng, S.Q., Janecek, A., Li J.Z., and Tan, Y., "Dynamic Search in Fireworks Algoritm, " IEEE Congress on Evolutionary Computation, Beijing, China, pp. 3222–3229 (July, 2014).

[4] Yu, J. and Takagi, H., "Acceleration for Fireworks Algorithm Based on Amplitude Reduction Strategy and Local Optima-based Selection Strategy, " Int. Conf. on Swarm Intelligence, Fukuoka, Japan, pp. 477– 484 (July, 2017).

[5] Luigi M., and Matteo M., "Robust Estimation of Natural Gradient in Optimization by Regularized Linear Regression, " 1st International SEE Conference on Geometric Science of Information, Paris, France, pp. 861–867 (Aug. 2013).

[6] Schtze, O., Alvarado, S., Segura, C., and Landa, R., "Gradient subspace approximation: a direct search method for memetic computing, " Soft Computing, volume 21, issue 21, pp. 6331–6350 (Nov. 2017).

[7] Murata, N., Nishii, R., Takagi, H., and Pei, Y. "Estimation Methods of the Convergence Point of Moving Vectors Between Generations," Japanese Society for Evolutionary Computation Symposium 2014, Hatsukaichi, Japan, pp.210–215 (Dec., 2014). (in Japanese).

[8] Murata, N., Nishii, R., Takagi, H., and Pei, Y. "Analytical Estimation of the Convergence Point of Populations," 2015 IEEE Congress on Evolutionary Computation, Sendai, Japan, pp. 2619–2624 (May, 2015).

[9] Liang, J.J., Qu, B.Y., Suganthan, P.N., and Alfredo, G H. "Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization". http://al-roomi.org/multimedia/CEC_Database/CEC2013/RealParameterOptimization/CEC2013_RealParameterOptimization_TechnicalReport.pdf (2013).

[10] Yu, J. and Takagi, H., "Estimation of the Convergence Points of a Population Using an Individual Pool," 10th International Workshop on Computational Intelligence and Applications, Hiroshima, Japan, pp. 67–72 (Nov., 2017).