Enhance Differential Evolution with Random Walk

Zhi-hui Zhan and Jun Zhang* (Corresponding Author)
Department of Computer Science, Sun Yat-sen University, Guangzhou, P. R. China, 510275
Key Laboratory of Digital Life, Ministry of Education, P. R. China
Key Laboratory of Software Technology, Education Dept. of Guangdong Province, P. R. China
*junzhang@ieee.org

ABSTRACT

This paper proposes a novel differential evolution (DE) algorithm with random walk (DE-RW). Random walk is a famous phenomenon universally exists in nature and society. As random walk is an erratic movement that can go in any direction and go to any place, it is likely that this mechanism can be used in search algorithm to bring in diversity. We apply the random walk mechanism into conventional DE variants with different parameters. Experiments are conducted on a set of benchmark functions with different characteristics to demonstrate the advantages of random walk in avoiding local optima. Experimental results show that DE-RWs have general better performance than their corresponding conventional DE variants, especially on multimodal functions.

Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search – *Heuristic methods*; G.1.6 [**Numerical Analysis**]: Optimization – *Global optimization*.

General Terms

Algorithms, Performance, Reliability, Experimentation

Keywords

Differential evolution (DE); random walk; Brownian motion; multimodal.

1. INTRODUCTION

Differential evolution (DE) [1], as together with genetic algorithms [2], ant colony optimization [3], and particle swarm optimization [4], have become significant and promising global optimizers in recent years [5]. A main inefficiency of DE is that it may sometime be trapped by local optima, especially on complex multimodal functions. Among current studies that try to enhance DE via adaptive [6] or co-evolutionary [7] strategies, they may obtain performance improvements by scarifying the DE simplicity. In this paper, we try a very simple way to enhance DE by bringing the random walk (RW) into DE, term DE-RW. Random walk, also referred to as Brown motion, is a famous phenomenon universally exists in nature and society which is an erratic movement that can go in any direction and go to any place [8]. Therefore it is likely that this RW mechanism can be used in the DE search process to bring in diversity.

In DE, suppose the individual vector is $\boldsymbol{X_i}$=[x_{i1}, x_{i2}, …, x_{iD}], where i is the individual index and D is the problem dimension, then the mutant vector $\boldsymbol{V_i}$=[v_{i1}, v_{i2}, …, v_{iD}] is obtained as:

GECCO'12 Companion, July 7–11, 2012, Philadelphia, PA, USA.
ACM 978-1-4503-1178-6/12/07.

$$\boldsymbol{V_i} = \boldsymbol{X_{best}} + F\times(\boldsymbol{X_{r1}}-\boldsymbol{X_{r2}}) \quad (1)$$

where F is the '*amplification factor*' parameter in [0, 1], $r1$ and $r2$ are different individuals, and *best* is the global best individual index. The crossover vector $\boldsymbol{U_i}$=[u_{i1}, u_{i2}, …, u_{iD}] is obtained as:

$$u_{id} = \begin{cases} v_{id}, \text{if rand}(0,1) < CR \text{ or } d = k \\ x_{id}, \text{ otherwise} \end{cases} \quad (2)$$

where CR is the '*crossover rate*' parameter in [0, 1] and k is a randomly selected dimension by i subject to $k\in\{1, 2, …, D\}$. For DE-RW, it has the same algorithm framework as conventional DE except minor modification in the crossover operator. In DE-RW, the crossover operator on the one hand combines the individual vector and the mutant vector, and on the other hand lets some dimension perform random walk, to form the target vector, as:

$$u_{id} = \begin{cases} v_{id}, & \text{if rand}(0,1) < CR \text{ or } d = k \\ rand(L_d, H_d), & \text{else if rand}(0,1)<RW \\ x_{id}, & \text{otherwise} \end{cases} \quad (3)$$

where L_d and H_d are the low and high search boundaries of the d^{th} dimension, and parameter RW is used to control the influence of random walk. Studies in Brownian motion show that the random walk phenomenon is related to the temperature that the random walk is more intensive when the temperature is higher [8]. Emulating this, the parameter RW is set to be larger in the early evolutionary phase in order to bring more diversity, and gradually decreases to smaller in the later evolutionary phase when the algorithm converges. The parameter RW is controlled as:

$$RW = 0.1-0.099\times g/G \quad (4)$$

where g and G are the current generation number and the maximal number of generation, respectively.

2. EXPERIMENTAL STUDIES

The experiments are on six 30 dimensional functions as listed in Table 1 (f_1-f_6) [9][10] where f_1-f_3 are unimodal and f_4-f_6 are multimodal. By taking the parameters into considerations, both DE and DE-RW are tested on (F=0.1, CR=0.1), (F=0.5, CR=0.1), (F=0.5, CR=0.9), and (F=0.9, CR=0.9). Therefore, totally 4 couples DE and DE-RW variants are tested. All the algorithms use the same population size of 50 and the same maximal number of fitness evaluations (FEs) of 3.0×10^5 for each test function. Each function is simulated 30 times and the average results are compared in Table 2, with the better results with **boldface**. The symbols '1', '0', and '−1' are used to indicate that DE-RW performs significantly better than, similar to, and significantly worse than the corresponding DE, respectively, according to the Wilcoxon's rank sum tests with significant level α=0.05.

The comparisons in Table 2 show that the performance of DEs is significantly affected by the parameters settings.

Table 1. Test functions for comparisons

Test function	*Range*	Test function	*Range*	Test function	*Range*
$f_1(x)=\sum_{i=1}^{D} x_i^2$	$[-100,100]^{30}$	$f_3(x)=\sum_{i=1}^{D} ix_i^4 + random[0,1)$	$[-1.28,1.28]^{30}$	$f_5(x)=\sum_{i=1}^{D} -x_i \sin(\sqrt{\lvert x_i \rvert})+418.9829\times D$	$[-500,500]^{30}$
$f_2(x)=\max_i(\lvert x_i \rvert, 1\le i\le D)$	$[-100,100]^{30}$	$f_4(x)=\sum_{i=1}^{D-1}[100(x_{i+1}-x_i^2)^2+(x_i-1)^2]$	$[-10,10]^{30}$	$f_6(x)=\sum_{i=1}^{D}[x_i^2-10\cos(2\pi x_i)+10]$	$[-5.12,5.12]^{30}$

Table 2. Experimental results on DE and DE-RW with different parameters

Functions		DE	DE-RW		DE	DE-RW		DE	DE-RW		DE	DE-RW	
		F=0.1, *CR*=0.1			*F*=0.5, *CR*=0.1			*F*=0.5, *CR*=0.9			*F*=0.9, *CR*=0.9		
f_1	Mean	9.92E+02	**8.86E-03**	1	**4.73E-142**	3.78E-48	−1	**1.49E-279**	8.80E-264	−1	**2.36E-54**	3.48E-47	−1
	Std	7.66E+02	**6.73E-03**		**1.07E-141**	8.47E-48		**0.00E+00**	**0.00E+00**		**4.84E-54**	6.38E-47	
f_2	Mean	5.43E+01	**9.56E-01**	1	1.51E+01	**2.83E-01**	1	1.32E+01	**5.64E-01**	1	**7.15E-02**	1.32E-01	−1
	Std	6.47E+00	**1.30E-01**		4.52E+00	**8.40E-02**		5.91E+00	**3.16E-01**		1.35E-01	**9.91E-02**	
f_3	Mean	3.67E-01	**9.81E-03**	1	**2.55E-03**	7.06E-03	−1	5.09E-02	**1.59E-02**	1	**9.69E-03**	1.06E-02	0
	Std	2.73E-01	**2.74E-03**		**7.71E-04**	1.68E-03		4.53E-02	**6.11E-03**		3.73E-03	**3.67E-03**	
f_4	Mean	5.88E+03	**6.23E+01**	1	2.64E+01	**2.45E+01**	1	1.06E+00	**9.98E-17**	1	1.20E+00	**2.96E-05**	0
	Std	4.41E+03	**3.52E+01**		1.34E+01	**1.71E+00**		1.79E+00	**4.91E-16**		1.86E+00	**6.21E-05**	
f_5	Mean	1.31E+03	**4.88E-02**	1	3.51E+02	**3.82E-04**	1	4.05E+03	**3.82E-04**	1	1.66E+03	**3.82E-04**	1
	Std	3.68E+02	**4.21E-02**		1.95E+02	**8.48E-13**		7.68E+02	**3.36E-10**		4.79E+02	**1.42E-11**	
f_6	Mean	3.98E+01	**6.31E-03**	1	8.62E-01	**0.00E+00**	1	8.42E+01	**4.45E-05**	1	5.82E+01	**7.38E-04**	1
	Std	6.36E+00	**5.83E-03**		7.27E-01	**0.00E+00**		2.05E+01	**2.44E-04**		1.60E+01	**3.36E-03**	

'1', '0', and '−1' indicate that DE-RW performs significantly better than, similar to, and significantly worse than its corresponding DE, respectively, according to the Wilcoxon's rank sum tests

For example, DE (*F*=0.5&*CR*=0.9) performs best on the simple unimodal function (Sphere function f_1) while DE (*F*=0.5&*CR*=0.1) seems to be good at most multimodal functions, except that DE (*F*=0.5&*CR*=0.9) is better on the Rosenbrock function f_4. This may be due to that the variables of f_4 are dependent and therefore require large *CR* value.

An interesting observation that supports the advantages of random walk is that under different parameter settings, DE-RWs are more promising than their corresponding DEs on most of the functions, especially on multimodal functions. DE-RW (*F*=0.1&*CR*=0.1) performs better than its DE variant on all the 6 functions. All the other three DE-RW variants obtain better results on all the 3 multimodal functions than their corresponding DE variants.

It should be noticed that DE-RW may result in slight worse solutions to simple unimodal functions. This is not surprising for that DE-RW may sacrifice a certain convergence speed on simple unimodal functions as it uses random walk to keep algorithm diversity for avoiding local optima. However, such sacrifice does not degenerate the algorithm performance much. DE-RWs can still obtain comparable results with their corresponding DEs. Moreover, DE-RW (*F*=0.5&*CR*=0.1) can obtain significantly better results than its DE on f_2 and DE-RW (*F*=0.5&*CR*=0.9) significantly outperforms its corresponding DE on both f_2 and f_3.

3. CONCLUSIONS

A novel DE variant using random walk mechanism is proposed in this paper. The DE-RW algorithm framework is as simple as conventional DE framework except that DE-RW uses a probability to control the random walk to bring random noises into the new solution vector. This way, the DE-RW algorithm can not only keep the simple algorithm implementation and the fast convergence characteristic of DE, but also keep diversity to avoid being trapped into local optima. This is important for the DE algorithm for optimizing both unimodal and multimodal functions. We have tested the DE-RW variants with different parameters configurations on a set of benchmark functions, by comparing them with their corresponding DE variants. The experimental studies demonstrate the intuition that DE-RW has stronger global search ability than conventional DE and therefore can obtain much better solutions to multimodal functions.

4. ACKNOWLEDGMENTS

This work was supported in part by the National Science Fund for Distinguished Young Scholars No.61125205, National Natural Science Foundation of China (NSFC) No.61070004, NSFC Joint Fund with Guangdong under Key Project U0835002, and the Scholarship Award for Excellent Doctoral Student Granted by Ministry of Education, China. The authors are with the Key Lab of Digital Life, MoE, China, and the Key Lab. of Softw. Tech, Edu. Depart of Guangdong Province. The corresponding author is Jun Zhang, email: junzhang@ieee.org.

5. REFERENCES

[1] Z. H. Zhan and J. Zhang, "Self-adaptive differential evolution based on PSO learning strategy," in *Proc. Genetic Evol. Comput. Conf.*, Jul. 2010, pp. 39-46.

[2] J. Zhang, H. Chung, and W. L. Lo, "Clustering-based adaptive crossover and mutation probabilities for genetic algorithms," *IEEE Trans. Evol. Comput.*, vol. 11, no. 3, pp. 326-335, Jun. 2007.

[3] W. N. Chen and J. Zhang, "Ant colony optimization approach to grid workflow scheduling problem with various QoS requirements," *IEEE Trans. Syst., Man, and Cybern. C.*, vol. 31, no. 1, pp. 29-43, Jan. 2009.

[4] W. N. Chen, J. Zhang, H. Chung, W. L. Zhong, W. G. Wu, and Y. H. Shi, "A novel set-based particle swarm optimization method for discrete optimization problems," *IEEE Trans. Evol. Comput.*, vol. 14, no. 2, pp. 278-300, April 2010.

[5] J. Zhang, Z. H. Zhan, Y. Lin, N. Chen, Y. J. Gong, J. H. Zhong, H. Chung, Y. Li, and Y. H. Shi, "Evolutionary computation meets machine learning: A survey," *IEEE Comput. Intell. Mag.*, vol. 6, no. 4, pp. 68-75, Nov. 2011.

[6] J. H. Zhong and J. Zhang, "Adaptive multi-objective differential evolution with stochastic coding strategy," in *Proc. Genetic Evol. Comput. Conf.*, Jul., 2011, pp. 665-672.

[7] Z. H. Zhan and J. Zhang, "Co-evolutionary differential evolution with dynamic population size and adaptive migration strategy," in *Proc. Genetic Evol. Comput. Conf.*, Jul., 2011, pp. 211-212.

[8] Wikipedia, *Random Walk*, Online: http://en.wikipedia.org/wiki/Random_walk, Access: March 2012.

[9] Z. H. Zhan, J. Zhang, Y. Li, and H. Chung, "Adaptive particle swarm optimization," *IEEE Trans. Syst., Man, and Cybern., B*, vol. 39, no. 6, pp. 1362-1381, Dec. 2009.

[10] Z. H. Zhan, J. Zhang, Y. Li, and Y. H. Shi, "Orthogonal learning particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 15, no. 6, pp. 832-847, Dec. 2011.