Information Sciences 323 (2015) 106–129

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Improving the vector generation strategy of Differential Evolution for large-scale optimization

Carlos Segura [a,*], Carlos A. Coello Coello [b], Alfredo G. Hernández-Díaz [c]

[a] Area of Computer Science, Centre for Research in Mathematics (CIMAT), Callejón Jalisco s/n, Mineral de Valenciana, Guanajuato, Guanajuato 36240, Mexico
[b] Evolutionary Computation Group, Department of Computer Science, Center of Research and Advanced Studies, National Polytechnic Institute, Mexico City 07300, Mexico
[c] Department of Economics, Quantitative Methods and Economic History, Pablo de Olavide University, Seville, Spain

ARTICLE INFO

Article history:
Received 29 September 2014
Revised 31 March 2015
Accepted 16 June 2015
Available online 25 June 2015

Keywords:
Differential evolution
Diversity preservation
Global numerical optimization
Large-scale optimization
Vector generation strategy

ABSTRACT

Differential Evolution is an efficient metaheuristic for continuous optimization that suffers from the curse of dimensionality. A large amount of experimentation has allowed researchers to find several potential weaknesses in Differential Evolution. Some of these weaknesses do not significantly affect its performance when dealing with low-dimensional problems, so the research community has not paid much attention to them. The aim of this paper is to provide a better insight into the reasons of the curse of dimensionality and to propose techniques to alleviate this problem. Two different weaknesses are revisited and schemes for dealing with them are devised. The schemes increase the diversity of trial vectors and improve on the exploration capabilities of Differential Evolution. Some important mathematical properties induced by our proposals are studied and compared against those of related schemes. Experimentation with a set of problems with up to 1000 dimensions and with several variants of Differential Evolution shows that the weaknesses analyzed significantly affect the performance of Differential Evolution when used on high-dimensional optimization problems. The gains of the proposals appear when highly exploitative schemes are used. Our proposals allow for high-quality solutions with small populations, meaning that the most significant advantages emerge when dealing with large-scale optimization problems, where the benefits of using small populations have previously been shown.

1. Introduction

Differential Evolution (DE) [43] is an efficient population-based metaheuristic initially designed for continuous optimization. From its inception, it has yielded remarkable results in several optimization competitions [12], such as in the *2005 IEEE Congress on Evolutionary Computation* (CEC) competition on real parameter optimization [39], or in the special issue on *Scalability of Evolutionary Algorithms and other Metaheuristics for Large Scale Continuous Optimization Problems*, recently organized for the *Soft Computing Journal* (SOCO) [21]. In addition, it has been successfully applied to demanding practical optimization problems [16,56].

In spite of the promising results obtained with different variants of DE, several potential weaknesses have been discovered. One of the first weaknesses studied is the large dependency between the DE parameters and the quality of the results. Several

* Corresponding author. Tel.: + 52 473 732 7155.
E-mail addresses: carlos.segura@cimat.mx, csegura@ull.es (C. Segura), ccoello@cs.cinvestav.mx (C.A. Coello Coello), agarher@upo.es (A.G. Hernández-Díaz).

http://dx.doi.org/10.1016/j.ins.2015.06.029
0020-0255/

studies have shown that DE is very sensitive to the setting of its control parameters [15,60]. Furthermore, several DE trial vector generation strategies have been proposed [17,37], hampering the choice of DE's parameters and components. In order to sidestep this drawback, some schemes consider the simultaneous use of several DE components and parameter values [34,55].

Some weaknesses involving the specific way in which new individuals are created in DE have also been identified [19,20]. DE borrows the idea from the Nelder and Mead method [28] of employing information from the vector population to alter the behavior of the variation scheme. Thus, the perturbation performed by the scheme is internally induced, i.e., the strength of the mutation depends on the contents of the population. The main associated problem is that DE has the potential to generate only a limited number of different trial solutions within one generation. In [19], it was shown that this can lead to a situation where DE may stop proceeding toward a global optimum even though the population has not even converged to a local optimum. This situation is called stagnation. The likelihood of stagnation occurring depends on the population size in question, and is more likely to occur when low population sizes are involved. An additional weakness of DE was reported in [20], where the authors generated a set of complex problems by applying genetic programming with the aim of detecting the drawbacks of different metaheuristics. It was shown that, in general, DE is less expansive than other metaheuristics. Thus, DE might be deceived into converging on the wrong peak, and once there, it could be impossible for this approach to escape.

The aforementioned weaknesses are closely related to the high selection pressure introduced by DE's survivor selection scheme and to its premature loss of diversity [2]. In the field of evolutionary computation, several schemes for dealing with such drawbacks have been proposed, some of which have been considered with the aim of improving the capabilities of DE. In fact, since its inception, the hybridization of DE with annealing procedures was studied to reduce the selection pressure [37]. Some more recent schemes include the introduction of stochasticity into the selection process [2] or the use of generational replacement [3].

The papers listed above enumerated some of the potential weaknesses of DE. However, most state-of-the-art DE schemes do not introduce any special mechanisms to address these weaknesses, probably because in most cases their effects are not very significant. In fact, in many cases the main effects of these weaknesses can be bypassed by increasing the population size [19]. Thus, it seems that with the traditional parameterizations of DE, the effect of these potential weaknesses is not very cumbersome. The result is that not much attention has been paid to these weaknesses. On the other hand, it is well-known that DE suffers from the curse of dimensionality, which refers to various phenomena that arise when dealing with high-dimensional spaces that do not occur in low-dimensional settings. In these cases, the results can generally be improved by applying micro-DE [30], i.e., schemes with low population sizes, meaning that inducing additional intensification is useful in these cases. However, such a reduction in the population size can lead to problems in terms of reliability [60].

The main aim of this paper is to show that when dealing with high-dimensional problems, the weaknesses mentioned above might arise and greatly deteriorate the performance of DE. Thus, in this paper, two techniques to address these drawbacks are proposed. Instead of considering general proposals from the field of evolutionary computation, our modifications take into account the particular variation method of DE so as to preserve its basic principles. Then, we show the utility of these schemes when faced with two well-known sets of high dimensional optimization benchmark problems. This is tested with several different non-hybrid variants of DE. Experimental evaluation shows that the benefits are more significant when exploitative configurations of DE are used, such as when considering low population sizes. This indicates that the effects of the weaknesses analyzed herein have a higher likelihood of appearing when large-scale complex problems are involved because of the required balance toward intensification. Note that it is beyond the scope of this paper to design a complete state-of-the art optimization scheme, which usually involves hybrid methods that combine several optimization methodologies. Instead, we want to better explore the reasons for the sub-optimal behavior of DE when applied to large-scale problems, which is why simple baseline and non-hybrid state-of-the-art algorithms are used.

The rest of the paper is organized as follows. A brief summary of DE and a discussion of the relevant background are given in Section 2. Section 3 is devoted to justifying the development of the new scheme by analyzing certain mathematical properties of several related schemes. The new proposal is described in Section 4. Then, our experimental validation is presented in Section 5. Finally, our conclusions and some lines of future work are given in Section 6.

2. State of the art in DE

2.1. Fundamentals

DE was initially proposed as a direct search method for single-objective continuous optimization problems [44]. In continuous optimization, the variables governing the system to be optimized are given by a vector $\vec{X} = [x_1, x_2, x_3, \ldots, x_D]$, where each variable x_i is a real number. The number of variables (D) defines the dimensionality of the optimization problem. Finally, the objective function $f(\vec{X})(f : \Omega \subseteq \mathbf{R}^D \rightarrow \mathbf{R})$ measures the quality of each set of variables. The aim of the optimization—considering a minimization problem—is to find a vector $\vec{X*} \in \Omega$ in which $f(\vec{X*}) \leq f(\vec{X})$ holds for all $\vec{X} \in \Omega$. The problems most typically addressed with DE are box-constrained optimization problems. In these cases, the region Ω is specified with the lower (a_j) and upper (b_j) bounds of each variable (j) in the problem.

DE is a population-based stochastic algorithm that belongs to the broad class of Evolutionary Algorithms (EAs). As with other EAs, it randomly initializes a population (P) with NP individuals ($P = \{\vec{X}_1, \ldots, \vec{X}_{NP}\}$). Each individual is a vector with D real numbers. The value of the jth variable of individual X_i is denoted by $x_{i,j}$. Then, the population evolves over successive iterations to explore the search space. In DE, the term *vector*, instead of *individual*, is commonly used. At each DE iteration, the following steps