European Journal of Operational Research 216 (2012) 33–46

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Continuous Optimization

Differential Evolution algorithm with Separated Groups for multi-dimensional optimization problems

Adam P. Piotrowski *, Jaroslaw J. Napiorkowski, Adam Kiczko

Institute of Geophysics, Polish Academy of Sciences, Ksiecia Janusza 64, 01-452 Warsaw, Poland

ARTICLE INFO

Article history:
Received 15 March 2011
Accepted 22 July 2011
Available online 29 July 2011

Keywords:
Metaheuristics
Evolutionary Computation
Differential Evolution
Particle Swarm Optimization
Evolution Strategies

ABSTRACT

The classical Differential Evolution (DE) algorithm, one of population-based Evolutionary Computation methods, proved to be a successful approach for relatively simple problems, but does not perform well for difficult multi-dimensional non-convex functions. A number of significant modifications of DE have been proposed in recent years, including very few approaches referring to the idea of distributed Evolutionary Algorithms. The present paper presents a new algorithm to improve optimization performance, namely DE with Separated Groups (DE-SG), which distributes population into small groups, defines rules of exchange of information and individuals between the groups and uses two different strategies to keep balance between exploration and exploitation capabilities. The performance of DE-SG is compared to that of eight algorithms belonging to the class of Evolutionary Strategies (Covariance Matrix Adaptation ES), Particle Swarm Optimization (Comprehensive Learning PSO and Efficient Population Utilization Strategy PSO), Differential Evolution (Distributed DE with explorative-exploitative population families, Self-adaptive DE, DE with global and local neighbours and Grouping Differential Evolution) and multi-algorithms (AMALGAM). The comparison is carried out for a set of 10-, 30- and 50-dimensional rotated test problems of varying difficulty, including 10- and 30-dimensional composition functions from CEC2005. Although slow for simple functions, the proposed DE-SG algorithm achieves a great success rate for more difficult 30- and 50-dimensional problems.

1. Introduction

The global single-objective unconstrained optimization problems considered in this paper may be defined in the following form – for a real valued function $f(\boldsymbol{x})$, find the global optimum vector $\boldsymbol{x}^*$, such that:

$$f(\mathbf{x}^*) = \min_{\mathbf{x}\in\Omega} f(\mathbf{x}) \tag{1}$$

where $\boldsymbol{x}$ is a M-dimensional parameter vector, $\boldsymbol{x} = \{x^1,\ldots,x^M\}$, also called an individual, with domain $\Omega \subset R^M$.

Among a number of Evolutionary Computation methods proposed during the last twenty years, the Differential Evolution (DE) algorithm suggested by Storn and Price (1995, 1997) gained significant popularity. The DE method was successfully applied to some practical problems (Price et al., 2005; Rowiński and Piotrowski, 2008; Zhang et al., 2010; Cruz et al., 2010; Beynon et al., 2010; Piotrowski et al., 2011) and simple benchmark functions (discussion on difficulty of test problems may be found for example in Clerc, 2006). However, for a number of problems (Ilonen et al., 2003; Langdon and Poli, 2007; Mendes et al., 2009) DE performs poorly, especially when applied to rotated multimodal problems with higher dimensionality (Neri and Tirronen, 2010). Therefore, a number of modifications of DE have been proposed (Price et al., 2005; Qin et al., 2009; Neri and Tirronen, 2010; Das and Suganthan, 2011).

The performance of the DE algorithm frequently depends on the control parameter setting. Moreover, DE has only a limited flexibility to adapt to specific problems and frequently prematurely converges to local optimum. To overcome these problems many modifications of DE aims at self-adaptation of DE parameters (Liu and Lampinen, 2005; Brest et al., 2006; Salman et al., 2007; Qin et al., 2009), development of new mutation and crossover strategies (Price et al., 2005; Das et al., 2009) or hybridization of DE with other algorithms (Zhang and Xie, 2003; Omran et al., 2009). Distributed DE methods have recently been proposed based on distributed Evolutionary Algorithms and Island Models (Tanese, 1989; Gustafson and Burke, 2006), including Distributed DE algorithm with Explorative-Exploitative Families (DDE, Weber et al., 2009) and Grouping DE algorithm (GDE, Piotrowski and Napiorkowski, 2010). These algorithms distribute the population into a few sub-populations (or groups) and define simple communication rules between them. However, the idea of distributing a population into many small sub-populations and defining various contact mechanisms between them has not been developed within a DE

* Corresponding author. Tel.: +48 22 6915858; fax: +48 22 6915915.
E-mail address: adampp@igf.edu.pl (A.P. Piotrowski).

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.07.038

framework. Similar to the self-adaptation of control parameters or the use of different mutation and crossover strategies, the distribution of a population into small groups with various contact mechanisms between them should improve adaptation capability of the algorithm to various types of problems improving optimization performance.

The present paper proposes a new algorithm pertaining to distributed DE methods, called DE with Separated Groups (DE-SG). The DE-SG is based on the GDE algorithm proposed by the authors in Piotrowski and Napiorkowski (2010), but some its features were inspired by other distributed and self-adaptive DE methods, as well as Island Models. However, DE-SG differs from the other algorithms in its structure and its mechanisms aimed at sorting and exchanging of information.

Unlike in DDE and GDE, the population of individuals in DE-SG is divided into halves (rules of migration of individuals are different in each half) and then each half is further divided into small groups (or sub-populations) that operate independently. Because the exchange of information within a small group is quicker, small groups are able to speed up exploitation. To facilitate exploration, communication between sub-populations is introduced in a few ways. Communication is understood here as gaining information about the location of individuals belonging to different groups and the migration of individuals between groups. In DE-SG, like in GDE, if a particular group has no improvement, the group's members are allowed to communicate with the whole population for some time; doing so allows them to get the necessary knowledge needed to escape from local optimum. One elite group has constant access to information collected by the whole population. However, as the groups are small, relying only on information exchange turns out to be insufficient; hence, the mechanism of migration of individuals between groups is introduced. Rules governing the migration of individuals between groups within each half differ. Within one half, the best individuals migrate relatively quickly to an elite group, while within the other, the best individuals migrate slowly and are distributed more widely among various groups.

On the basis of the Self-adaptive DE (SaDE, Qin et al., 2009) and GDE, in DE-SG an offspring may be produced by one of two strategies of different nature. The first one is expected to perform better exploration, the second exploitation. Although DE-SG introduces a set of new parameter values, only the speed of communication and exchange of individuals between groups may require some tuning by the user, depending on the number of function calls allowed. All results presented in this paper were obtained with the fixed recommended parameter values.

The proposed method is compared to eight state of the art EC algorithms based on 19 rotated 10-to 50-dimensional test problems. The paper is organized as follows: Section 2 describes the DE algorithm and the overview of methods developed to improve its performance; Section 3 describes the DE-SG algorithm; Section 4 provides benchmark results and includes a discussion of the performance of all considered algorithms; Section 5 concludes the paper.

2. Differential Evolution algorithm (DE)

To find the optimal solution $\boldsymbol{x}^*$ in the M-dimensional space, the original DE algorithm (Storn and Price, 1995) randomly initializes population P with individuals $\boldsymbol{x}_i$, $i = 1, \ldots, K$ within Ω.

At each iteration, for every individual $\boldsymbol{x}_i$, $(i = 1, \ldots, K)$, three other distinct parameter vectors $\boldsymbol{x}_a$, $\boldsymbol{x}_b$, and $\boldsymbol{x}_c$ are randomly chosen from the population P. Then a new vector $\boldsymbol{u}_i$ is generated by an operation called mutation:

$$\mathbf{u}_i = \mathbf{x_a} + F \cdot (\mathbf{x_b} - \mathbf{x_c}) \tag{2}$$

where F is a predefined scaling parameter. The mutated vector $\boldsymbol{u}_i$ and the target (initial) vector $\boldsymbol{x}_i$ form two parents. In order to form the trial vector (offspring) $\boldsymbol{y}_i$, the crossover between $\boldsymbol{u}_i$ and $\boldsymbol{x}_i$ is defined component-wise as:

$$y_i^j = \begin{cases} u_i^j & \text{if } rand(0,1)_j \leqslant Cr \text{ or } j = j_{rand} \\ x_i^j & \text{otherwise} \end{cases} \tag{3}$$

where Cr is a predefined crossover parameter, j_{rand} is a random integer variable from $[1, M]$ and $rand(0,1)_j$ is a uniform random real variable from range $[0,1]$. Note that as a number of mutation and crossover strategies were proposed since 1995 (see Price et al., 2005; Mezura-Montes et al., 2006; Mishra, 2006; Qin et al., 2009; Das et al., 2009), the initial version defined above was named DE/rand/1/bin. Then, if

$$f(\mathbf{y_i}) \leqslant f(\mathbf{x_i}) \tag{4}$$

$\boldsymbol{y}_i$ replaces $\boldsymbol{x}_i$ – this operation is called selection. The entire algorithm terminates when one of the stopping criteria is met. In basic DE, three control parameters, namely scaling factor F, crossover rate Cr and population size K must be specified by the user.

Population size is probably the most problem-dependent control parameter. Storn and Price (1997) suggested that population size K should equal from $5M$ up to $10M$. Different opinions about proper population size may be found in the literature, ranging from $K < M$ to $K > 10M$ (see review in Das et al., 2009; Qin et al., 2009; Weber et al., 2009 or Das and Suganthan, 2011 for detailed discussion). Frequently K is left to be decided by the user.

The value of Cr also heavily depends on the problem (Price et al., 2005), while desirable value differs for separable (lower Cr) and non-separable problems (higher Cr). The value of F was suggested to equal 0.5 by Storn and Price (1997).

Looking for the improvement of original DE, many researchers considered various control parameter values (Gamperle et al., 2002; Rokkonen et al., 2005; Chakraborty et al., 2006; Kaelo and Ali, 2006), for example suggesting different values of F, generally between 0.4 and 1. More recently a number of algorithms with self-adaptive parameters were proposed (Zaharie and Pectu, 2003; Liu and Lampinen, 2005; Omran et al., 2005; Qin and Suganthan, 2005; Qin et al., 2009; Brest et al., 2006; Al-Anzi and Allahverdi, 2007; Zhang and Sanderson, 2009). The results of several self-adaptive DE algorithms were compared in Brest et al. (2007) and Qin et al. (2009)).

Detailed discussion of DE-based algorithms developed during last fifteen years is beyond the scope of the present paper, but good review may be found in Das et al. (2009), Qin et al. (2009), Weber et al. (2009), Neri and Tirronen (2010) and Das and Suganthan (2011). Most improvements aimed at an increase in the algorithm's ability to adapt to specific problems and avoid premature convergence. To achieve it, different methods combining DE with different neighborhood topologies (Tomassini, 2005), used for example in Particle Swarm Optimization (PSO, Eberhart and Kennedy, 1995) were proposed to slow down information propagation (Das et al., 2009; Omran et al., 2009). Alternatively, a concept of cooperative coevolution – developed to decompose high-dimensional problems into smaller, simpler to solve components (Yang et al., 2008) – was introduced into the DE approach. Also, an idea of merging advantages of multiple EC methods, including DE, has been developed into an adaptive multi-algorithm AMALGAM (Vrugt et al., 2009).

Another idea was adopted from Island or Species based models (Tanese, 1989; Holland, 2000; Liu et al., 2000). A distribution of individuals into sub-populations that occasionally communicate allows a better exploration of search space and usually slows down the algorithm's convergence to local optimum. Among a few distributed DE algorithms (Tasoulis et al., 2004; Falco et al., 2007; Apolloni et al., 2008) two become more successful (Weber et al., 2009 and Piotrowski and Napiorkowski, 2010). However, when