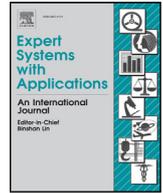




ELSEVIER

Contents lists available at ScienceDirect

Expert Systems With Applications

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

A population initialization method for evolutionary algorithms based on clustering and Cauchy deviates



Dražen Bajer^a, Goran Martinović^{a,*}, Janez Brest^b

^a Faculty of Electrical Engineering, J.J. Strossmayer University of Osijek, Kneza Trpimira 2b, 31000 Osijek, Croatia

^b Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova ul. 17, 2000 Maribor, Slovenia

ARTICLE INFO

Article history:

Received 24 February 2016

Revised 18 April 2016

Accepted 5 May 2016

Available online 10 May 2016

Keywords:

Cauchy deviates

Clustering

Differential evolution

Evolutionary algorithms

Initial population

Mutation

ABSTRACT

The initial population of an evolutionary algorithm is an important factor which affects the convergence rate and ultimately its ability to find high quality solutions or satisfactory solutions for that matter. If composed of good individuals it may bias the search towards promising regions of the search space right from the beginning. Although, if no knowledge about the problem at hand is available, the initial population is most often generated completely random, thus no such behavior can be expected. This paper proposes a method for initializing the population that attempts to identify i.e., to get close to promising parts of the search space and to generate (relatively) good solutions in their proximity. The method is based on clustering and a simple Cauchy mutation. The results obtained on a broad set of standard benchmark functions suggest that the proposed method succeeds in the aforementioned which is most noticeable as an increase in convergence rate compared to the usual initialization approach and a method from the literature. Also, insight into the usefulness of advanced initialization methods in higher-dimensional search spaces is provided, at least to some degree, by the results obtained on higher-dimensional problem instances—the proposed method is beneficial in such spaces as well. Moreover, results on several very high-dimensional problem instances suggest that the proposed method is able to provide a good starting position for the search.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Evolutionary algorithms (EAs) (Eiben & Smith, 2003) are population-based search and optimization methods. As such, the initial population plays a key role in terms of convergence rate and may affect the success of a EA in finding high quality or satisfactory solutions. Most often the initial population is generated completely random (Kazimipour, Li, & Qin, 2013; Maaranen, Miettinen, & Penttinen, 2007) (using uniform deviates). Such initial populations are as a rule very diverse. This is a desirable feature since it enables an extensive exploration of the search space and helps keeping the population from premature convergence and being trapped in local optima. On the other hand, such populations are generally composed of low quality individuals exclusively. Consequently, this requires a substantial amount of time for reaching promising regions of the search space and ultimately, requires more time in order to converge towards high quality solutions.

When the evaluation of the objective function is computationally expensive or the number of function evaluations is limited, an increased convergence speed towards promising regions of the search space is beneficial since it would reduce the time needed for finding satisfactory or high quality solutions. An increased convergence rate may be achieved by introducing good solutions/individuals into the initial population (Martinović & Bajer, 2012). Needless to say, this is not always easy.

When knowledge about the problem being solved is available, simple heuristics may be employed to obtain good solutions which can then be introduced into the initial population. The traveling salesman problem (TSP) is a prime example since there exists a multitude of heuristics. A modified nearest neighbor algorithm was proposed in Martinović and Bajer (2012) and utilized for the population initialization of a genetic algorithm (GA). Wang, Duan, and Zhang (2009) proposed an improved greedy algorithm which was used to initialize half of the population while the rest was initialized randomly. An extensive comparison of various population initialization methods used for the TSP can be found in e.g., Paul et al. (2015). A number of population initialization methods were proposed for many problems other than the TSP. For example, Guerrero, Berlanga, and Molina (2012) presented two distinct

* Corresponding author. Fax: +38531495402.

E-mail addresses: drazen.bajer@etfos.hr (D. Bajer), goran.martinovic@etfos.hr (G. Martinović), janez.brest@um.si (J. Brest).

population initialization methods for the segmentation problem and compared them to the usual (random) initialization approach. In order to solve the flexible job-shop scheduling, Zhang, Gao, and Shi (2011) proposed a new initialization procedure.

However, many real-world problems are like black boxes, hence no knowledge about their internals is available (black-box optimization (Audet, 2014; Cassioli & Schoen, 2013)). Generating an initial population that contains good individuals in such cases represents a significant problem. Thus, it is not surprising that the population is most often initialized completely random, since it is very hard to find good solutions in an efficient and simple manner. Nonetheless, a few attempts trying to overcome those problems can be found in the literature. Rahnamayan, Tizhoosh, and Salama (2008); (2007) proposed a method that employs opposition-based learning (OBL). This initialization approach was utilized by Dong et al. (2012) in a hybrid algorithm for the circle detection problem. Furthermore, a completely different approach to the initialization of the EA population was proposed by Ali, Pant, and Abraham (2013). The approach employs either a slightly modified quadratic interpolation or nonlinear simplex method. Both the aforementioned methods generate a new population from a starting uniformly random one. Also, both methods were tested for initializing the population of differential evolution (DE) which is a very competitive EA for numerical optimization.

This paper proposes a new method for initializing the population of EAs when there is no problem knowledge available. It is based on clustering and uses Cauchy deviates. Clustering is used in an attempt to identify i.e., to get close to promising regions of the search space which are then represented by cluster centers. A simple Cauchy mutation is then employed in order to generate new individuals around those centers while favoring the better ones. Finally, individuals created uniformly random are introduced in order to create a sufficiently diverse initial population.

Although populations generated completely at random may span the whole search space, they become extremely sparse as the dimensionality of that space increases (Rahnamayan & Wang, 2009). Due to this fact, it is very unlikely that solutions in promising regions are generated. Hence, the approach aimed at generating a population that is not necessarily diverse, but one that contains at least a few relatively good solutions is adopted here.

The rest of the paper is organized as follows. Section 2 gives a brief and concise introduction to clustering and DE, since clustering is a key part of the proposed method, while DE is employed as a representative EA for the numerical experiments. The proposed method for initializing the EA population is described and analyzed in detail in Section 3. Setup of the conducted experimental analysis and the obtained results are reported and discussed in Section 4. Finally, the drawn conclusions and suggestions for future work are presented in Section 5.

2. Preliminaries

Generally, a global optimization problem can be represented with the pair (S, f) , where $S \subseteq \mathbb{R}^d$ is the search space, and $f : S \rightarrow \mathbb{R}$ a real-valued objective function. Solving this problem requires that a d -dimensional point $\mathbf{x}^* \in S$ is found such that

$$\forall \mathbf{x} \in S : f(\mathbf{x}^*) \leq f(\mathbf{x}) . \tag{1}$$

Global or more specifically, numerical optimization (Nocedal & Wright, 2006) problems may be subject to a multitude of linear and/or nonlinear constraints. The focus here is on unconstrained optimization problems which are subject to no constraints, although bound-constraints (box-constraints) typically apply.

Next, the clustering problem with focus on partitional clustering is described, followed by a brief review of some previous application of clustering in EAs. Also, a description of differential evolution

is given. These topics are introduced since clustering is an essential part of the proposed method, whereas DE is the EA that has been used for testing purposes.

2.1. Clustering

For a given set $\mathcal{A} = \{\mathbf{a}^j : j = 1, \dots, n\} \subset [\alpha, \beta] \subset \mathbb{R}^d$, where $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$, $\beta = (\beta_1, \dots, \beta_d) \in \mathbb{R}^d$, clustering (Theodoridis & Koutroumbas, 2008; Xu & Wunsch, 2009) requires the grouping/division of that set into $1 < k < n$ subsets π_1, \dots, π_k called clusters such that

$$\bigcup_{j=1}^k \pi_j = \mathcal{A}, \quad \pi_r \cap \pi_s = \emptyset, \quad r \neq s, \quad |\pi_j| \geq 1, \quad j = 1, \dots, k . \tag{2}$$

The grouping of the set \mathcal{A} into k subsets that satisfy (2) is a hard (or crisp) partition (or k -partition) of \mathcal{A} , and can be represented by the set $\Pi(\mathcal{A}) = \{\pi_1, \dots, \pi_k\}$.

Furthermore, by introducing a proximity measure (Teboulle, 2007; Theodoridis & Koutroumbas, 2008) between the data points, each cluster π_j can be assigned a representative in form of a center \mathbf{z}^j . Usually the Euclidean distance or norm ($\|\cdot\|$) is used. Each cluster π_j can be then represented by the center

$$\mathbf{z}^j = \frac{1}{|\pi_j|} \sum_{\mathbf{a} \in \pi_j} \mathbf{a} . \tag{3}$$

Conversely, a partition can be calculated based on provided centers $\mathcal{Z} = \{\mathbf{z}^1, \dots, \mathbf{z}^k\}$ according to the minimum distance principle

$$\pi_j = \{\mathbf{a} \in \mathcal{A} : \|\mathbf{a} - \mathbf{z}^j\| \leq \|\mathbf{a} - \mathbf{z}^r\|, \forall r = 1, \dots, k\}, \quad j = 1, \dots, k . \tag{4}$$

Since the number of possible partitions of the set \mathcal{A} is very large, a criterion for evaluating the different partitions must be introduced. The problem of finding an optimal partition may be defined as the global optimization problem (Teboulle, 2007; Xu & Wunsch, 2009)

$$\min \mathcal{F}(\mathcal{Z}, \Pi) = \sum_{j=1}^k \sum_{\mathbf{a} \in \pi_j} \|\mathbf{a} - \mathbf{z}^j\|^2 . \tag{5}$$

Different weights $w_j > 0$ can be assigned to each data point $\mathbf{a}^j \in \mathcal{A}$ in order to emphasize their importance, thus determining their contribution in the calculation of the centers. Accordingly, (3) becomes

$$\mathbf{z}^j = \frac{1}{W_j} \sum_{\mathbf{a}^r \in \pi_j} \mathbf{a}^r, \quad W_j = \sum_{\mathbf{a}^r \in \pi_j} w_r . \tag{6}$$

One of the most popular and widely used algorithms that searches for a locally optimal partition of \mathcal{A} (in terms of (5)) is the k -means (Theodoridis & Koutroumbas, 2008; Wu et al., 2007; Xu & Wunsch, 2009) algorithm. It is a simple alternating optimization procedure that calculates a partition based on fixed centers which is then used to correct/update those centers. The procedure is repeated until a termination criterion is met. Despite having some drawbacks, its simplicity and speed (as stated in Duda, Hart, and Stork (2000), in practice only a relatively low number of iterations is needed) make it a powerful and useful data mining tool (see e.g., Ayech and Ziou (2015); Öztürk, Cavusoglu, and Zengin (2015)). The k -means algorithm for clustering data with assigned weights is outlined in Algorithm 1.

2.2. Applications of clustering in EAs

Clustering has found applications in numerous research areas that range from astronomy to signal processing. The partitioning

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات