Journal of Computational Science 5 (2014) 277–297

Contents lists available at ScienceDirect

Journal of Computational Science

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

A new population seeding technique for permutation-coded Genetic Algorithm: Service transfer approach

P. Victer Paul [a,*], A. Ramalingam [b], R. Baskaran [c], P. Dhavachelvan [a], K. Vivekanandan [d], R. Subramanian [a]

[a] Department of Computer Science, Pondicherry University, Puducherry, India
[b] Department of MCA, Sri Manakula Vinayagar Engineering College, Puducherry, India
[c] Department of Computer Science and Engineering, Anna University, Chennai, India
[d] Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry, India

ARTICLE INFO

Article history:
Received 31 December 2012
Received in revised form 16 May 2013
Accepted 26 May 2013
Available online 20 June 2013

Keywords:
Genetic Algorithm
Population seeding technique
Traveling Salesman Problem
Order distance vector
Combinatorial problem

ABSTRACT

Genetic Algorithm (GA) is a popular heuristic method for dealing complex problems with very large search space. Among various phases of GA, the initial phase of population seeding plays an important role in deciding the span of GA to achieve the best fit w.r.t. the time. In other words, the quality of individual solutions generated in the initial population phase plays a critical role in determining the quality of final optimal solution. The traditional GA with random population seeding technique is quite simple and of course efficient to some extent; however, the population may contain poor quality individuals which take long time to converge with optimal solution. On the other hand, the hybrid population seeding techniques which have the benefit of good quality individuals and fast convergence lacks in terms of randomness, individual diversity and ability to converge with global optimal solution. This motivates to design a population seeding technique with multifaceted features of randomness, individual diversity and good quality. In this paper, an efficient Ordered Distance Vector (ODV) based population seeding technique has been proposed for permutation-coded GA using an elitist service transfer approach. One of the famous combinatorial hard problems of Traveling Salesman Problem (TSP) is being chosen as the testbed and the experiments are performed on different sized benchmark TSP instances obtained from standard TSPLIB [54]. The experimental results advocate that the proposed technique outperforms the existing popular initialization methods in terms of convergence rate, error rate and convergence time.

1. Introduction

Genetic Algorithm (GA) is a well-known method for global optimization of complex problem very large search space based on the survival of the fittest concept of natural evolution [55]. The significant features of the GA, which makes it perform competently could be defined as follows: GA operates on a population of feasible solutions rather than on a single solution; the variety of genetic operators helps to explore unrevealed solutions in the large search space effectively; possibility to construct problem the specific genetic operators which can offer better solution search; population diversity helps to avoid the drawback of getting trapped in local optima and premature convergence. These flexible configurations encourage researchers to design novel GA with a modified operators and population seeding techniques to improve further its performance.

GA had been proven to be efficient at searching optimal solution among a large and complex search space in an adaptable way, controlled by the equivalent biological evolutionary mechanisms of reproduction, crossover, and mutation. Various phases of GA can be defined as population seeding (initial population), selection, reproduction, crossover, mutation and termination constraint, in which first step occurs once and the rest of the steps are repeated until the final condition is satisfied [16,52]. The first step of any GA is to generate a set of possible solutions randomly as an initial population or population seeding [34,37,41]. The quality of individual solutions in the initial population plays a critical role in determining the quality of the final solution that can be obtained using GA [32,40]. However, in traditional GA, population seeding is performed randomly which can be simple but, the whole population contains much of individuals with worst quality, infeasible solutions sometimes [36]. As a consequence, GA with a random population seeding technique requires longer search time to find

* Corresponding author. Tel.: +91 0413 2274430.
E-mail addresses: victerpaul@gmail.com, victerpaul.p@gmail.com (P. Victer Paul), a.ramalingam1972@gmail.com (A. Ramalingam), baaski@annauniv.edu (R. Baskaran), dhavachelvan@gmail.com (P. Dhavachelvan), k.vivekanandan@pec.edu (K. Vivekanandan), subbur@yahoo.com (R. Subramanian).

1877-7503/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jocs.2013.05.009

an optimal solution, number of generations required to evolve optimal solution increases, the search possibility for an optimal solution decreases and more importantly the convergence rate or quality of the optimal solution obtained is reduced. Thus, the requirement for a modified population seeding technique in GA is clear and in fact, several research works were proposed to support the dispute [61].

Lawrence and Amini [61] discussed about different GA configuration issues and claims that seeding the initial population with heuristics can improve the efficiency of the GA greatly. Togan and Daloglu [40] believes that performance and convergence ability of GA are critically influenced by the population seeding method and proposed two new self-adaptive member grouping strategies and a new strategy for population seeding. The large search space is collected into different groups and the list of cross section values are assigned as initial values to set the initial population automatically. This method is dedicatedly proposed for application in the area of structural engineering and tested in truss structures and transmission towers. Nearest Neighbor (NN) tour construction heuristic is one of the familiar alternatives for random population seeding in GA, particularly for TSP [3,16,42–46]. In NN technique, individuals in the population seeding are constructed with the city nearest to the current city and such good individuals can refine the subsequent search in the next generations [3]. Though NN works fine, it suffers with some critical factors: several cities are not included in the individuals created initially and have to be inserted at high costs in the end; neglecting several cities at the population seeding stage leads to severe errors in optimal solution construction and the diversity among the individuals created is very minimum.

Yingzi et al. [32] proposed a Greedy GA (GGA), in which the population seeding is performed using Gene Bank (GB). The GB is built by assembling the permutation of '*n*' cities based on their distance. In GGA method, the population of individuals is generated from the GB such that the individuals are of above-average fitness and short defining length. In GGA, with the increase in number of cities leads to augmented problem complexity and performance degradation, and large collection of GB individuals enlarges the cost of computation at each generation. The improved performance of GGA is justified using TSP with maximum of hundred cities and its performance deteriorates with large number of cities. In [38], Fuyan et al. proposed K-means algorithm, based on the work reported in [46], which is considered to generate much infeasible solutions, to obtain the initial population in which '*N*' number of individuals are partitioned and assigned to one of '*K*' clusters. As a result, using K-means algorithm, generating infeasible solutions in the population seeding stage is avoided. Performance evaluation is performed with a maximum of 10 cities with only time based analysis and be deficient to validate the proposed technique for large number of cities and convergence capability. Yugay et al. [36] proposes a modified GA with sorted initial population method based on theory of better parents produce better offsprings. In this approach, a large initial pool of population is generated and ranked in accordance to their fitness values and at last, a certain number of individuals with bad fitness are omitted. This approach also suffers with the issues discussed with NN tour construction heuristic technique such as premature convergence, reduced search space exploration and minimum population diversity. Hence, the traditional GA does not provide effective performance when applied to some of the combinatorial problems like TSP [33], so each stage in the traditional GA has been modified in order to achieve a better output and thus resulted in hybrid GAs [28–32,34,49,50].

Recently, many researchers proposed modified versions of GA, particularly for solving TSP using random population seeding techniques [3,16,30–32,34,39–41,48,64–66]. Though several modified population seeding techniques for GA have been proposed, many researchers still continue to work with a random population seeding technique because of the complexity nature of algorithm, which is difficult to understand and implement, problem specific modifications are required to apply and problem such as premature convergence, ineffective search space exploration and less population diversity. This implies that the researchers are interested in random population seeding technique to accomplish a better search space exploration and finding best optimal solution at the cost of high convergence time. In [64], Xing et al. proposed a hybrid approach combining an improved Genetic Algorithm and optimization strategies using random population seeding technique. An efficient hybrid mutation Genetic Algorithm has been proposed using a random population seeding technique in [34]. Chang and Ramakrishna [39] proposed a GA for shortest path routing problem, in which author reveals the reason for preferring random population seeding method than heuristic initialization. Although the mean fitness of the individuals generated using heuristic initialization are high so that it may help the GA to obtain the better solutions faster, but it ends up in exploring a small part of the search space and never find the global optimal solutions because of the lack of diversity in the initial population generated [47]. In some of the works, the authors used hybrid population seeding technique which combines random and any of the modified population seeding technique [33,61]. In [62], Qu and Sun proposed a synergetic approach to GA by adding some new randomly generated individuals into the population after each generation in order to prevent premature convergence and to obtain final optimum.

To summarize, the modified population seeding techniques has the advantages of good quality or generating potential sequence individuals at an early stage and the ability to find near optimal solutions at few generations; however they lacks in randomness, diversity of individuals generated, ability to explore more search space and finding the global solution. On the other hand, the random population seeding technique has the advantages of individual diversity, can explore search space efficiently and finding optimal solution; however they have the disadvantages of individuals with worst potential sequence and requires longer search time to converge an optimal solution. The controversies between these two categories of population seeding techniques motivate to propose an efficient population seeding technique with characteristics of randomness, individual diversity and potential sequence. Thus in this paper, an efficient Ordered Distance Vector (ODV) based population seeding technique with three different varieties has been proposed for permutation-coded GA. The popular combinatorial optimization problem of Traveling Salesman Problem (TSP) is being chosen as the testbed to validate and claim the efficacy of the proposed population seeding technique. Experiments were performed over the different sized benchmark TSP instances obtained form the TSPLIB [54]. The organization of the paper is as follows: Section 2 offers sufficient background information over GA and TSP to improve the understanding of this paper. Section 3 describes the proposed technique and its variants along with the corresponding algorithms. The different phases of the experiments are summarized in Section 4. This section also reports experimental results with corresponding analyses. And finally, Section 5 presents the conclusive remarks of the work reported in this paper.

2. Background information

As stated earlier, this section offers a brief introduction on Genetic Algorithm (GA) and Traveling Salesman Problem (TSP) to improve the understandability of this paper.

2.1. Genetic Algorithm

Genetic Algorithms (GAs), a subclass of evolutionary algorithms, is a stochastic optimization technique based on the principles