

# Genetic Algorithm Applied to the Capacitated Vehicle Routing Problem: An Analysis of the Influence of Different Encoding Schemes on the Population Behavior

Stanley Jefferson de Araujo Lima<sup>a\*</sup>, Sidnei Alves de Araújo<sup>b</sup>

<sup>a,b</sup>*Universidade Nove de Julho - UNINOVE, Informatics and Knowledge Management Graduate Program., Rua Vergueiro, 235/249 - Liberdade, Zip Code: 01504-001, São Paulo / SP, Brazil*

<sup>a</sup>*Email: stanleyjefferson@outlook.com, <sup>b</sup>Email: saraujo@uni9.pro.br*

## Abstract

Genetic Algorithm (GA) is an optimization method that has been widely used in the solution of NP-Hard (Non-deterministic Polynomial-time) problems, among which is the Vehicle Routing Problem (VRP), widely known in the literature due to its applications in the logistics and supply sectors, and which is considered in this work. However, finding solution for any optimization problem using GA presupposes the adoption of a solution encoding scheme that, according to the literature, impacts its performance. However, there is a lack of works in the literature exploring this theme. In this work we carry out an analysis of the main encoding schemes (binary and integer) employed in the GA for the solution of the capacitated VRP (CVRP), in order to evaluate the influence of each of them on the behavior of the GA population and, consequently, on the algorithm performance. To this end, we developed a computational tool that allows visualizing the GA individuals (solutions) mapped to a two-dimensional space. Based on the experiments conducted, we observed that, in general, integer vectors provide better conditions for GA individuals to explore the solution space, leading to better results. The results found, besides corroborating some assumptions in the literature, may justify the preference for integer encoding schemes to solve CVRP in recent literature works. In addition, this study can contribute to the choice and/or proposition of heuristics that allow GA to search for better quality solutions for the VRP with less computational effort.

**Keywords:** Genetic Algorithm; Encoding Scheme; GA Behavior; Capacitated Vehicle Routing Problem.

---

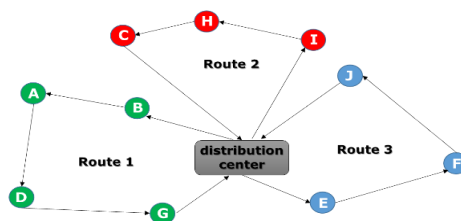
\* Corresponding author.

## 1. Introduction

The Genetic Algorithm (GA) is a metaheuristic method based on Charles Darwin's theory of natural evolution. According to that theory, the individuals of a population that best adapt to the environment in which they live are more likely to survive and reproduce [1]. The evolution process in GA occurs from a population of initial individuals, through the application of genetic operators for selection, crossover and mutation of individuals, which are the encoded solutions of the problem. The GA has been widely used in the solution of highly complex optimization problems, known in the literature as NP-Hard (Non-deterministic Polynomial-time), among which is the Vehicle Routing Problem (VRP), investigated in this work, which has been attracting the attention of many researchers in recent years due not only to the difficulty of its solution, but also to its presence in several practical situations in the logistics and supply sectors. The VRP basically consists in determining a set of routes so that a fleet of vehicles can serve a set of geographically dispersed customers, optimizing, for example, the total cost of routing and/or the total distance traveled, respecting the operational restrictions intrinsic to the considered environment [2,3]. In general, the VRP involves two main steps: grouping customers by route and defining the sequence of service sequence. There are several variants of the VRP, which take into account the vehicle's capacity (each vehicle has a specific capacity), time windows (customers must be served at a specific time), heterogeneous vehicle fleet (different vehicles), among others [4]. The most basic version is the Capacitated Vehicle Routing Problem (CVRP), considered in this work and described in detail in section 2.1, which considers a homogeneous set of limited capacity vehicles, initially located in the same depot and being the only restriction imposed the vehicle capacity [5,6]. The literature review conducted in this work showed that GA has been widely used in the resolution of VRP due to the good results it has shown for this problem. However, finding solutions to any combinatorial optimization problem through GAs presupposes the adoption of a solution encoding scheme, that is, how to represent the solution on the GA chromosome. Thus, several solution encoding schemes for the VRP have been proposed. However, according to the literature, the encoding scheme can impact on the performance of the GA [7-9]. The problem is that there is a lack of studies that explore how such schemes influence the behavior of the GA population and, consequently, its performance. In this work we carry out an analysis of the main solution encoding schemes employed in the GA to optimize the CVRP, in order to evaluate their influences on the behavior of the GA performance well as what are their advantages and disadvantages. For that, we developed and present in this work a computational tool that allows visualizing the GA individuals mapped to a two-dimensional space.

## 2. Theoretical background

### 2.1. Capacitated Vehicle Routing Problem (CVPR)



**Figure 1:** Example of vehicle routing with three routes starting from the distribution center (depot)

In the CVPR, illustrated in Figure 1, each customer has its demand previously defined, which must be fully delivered by only one vehicle. All vehicles have identical capacity, and depart from a unique distribution center [2,5,10].

In other words, the CVPR consists of finding a set of routes, each of which must be covered by a vehicle, with the aim of minimizing the total cost of the routes respecting the following restrictions: (a) each route must start and end in the same distribution center (depot); (b) each customer must be visited only once and by the same vehicle; (c) the sum of customer demands grouped on a route cannot exceed the vehicle's capacity. The mathematical formulation for the CVPR, adapted from [10] to the binary encoding scheme, can be expressed as follows:

$$\text{Minimize} \quad tc = \sum_{k=1}^K \sum_{j=1}^{nc} \sum_{i=0}^{nc} c_{ij} x_{ijk} \quad (1)$$

$$\text{Subject to:} \quad \sum_{k=1}^K \sum_{j=1}^{nc} x_{0jk} \leq K \quad (2)$$

$$\sum_{j=1}^{nc} x_{0jk} = \sum_{j=1}^{nc} x_{j0k} = 1, \quad k = 1, \dots, K \quad (3)$$

$$\sum_{k=1}^K \sum_{j=0}^{nc} x_{ijk} = 1, \quad i = 1, \dots, nc \quad (4)$$

$$\sum_{j=0}^{nc} x_{ijk} - \sum_{j=0}^{nc} x_{jik} = 0, \quad k = 1, \dots, K \quad i = 1, \dots, nc \quad (5)$$

$$\sum_{k=1}^K \sum_{i \in S} \sum_{j \in S} x_{ijk} \leq |S| - m(S), \quad \forall S \subseteq V \setminus \{0\}, |S| \geq 2 \quad (6)$$

$$\sum_{i=1}^{nc} d_i \sum_{j=0}^{nc} x_{ijk} \leq cv, \quad k = 1, \dots, K \quad (7)$$

$$x_{ijk} \in \{0,1\}, \quad i = 1, \dots, nc, j = 1, \dots, nc, k = 1, \dots, K \quad (8)$$

where:  $d_i$ : demand of customer  $i$ ,  $k$ : vehicle,  $K$ : set of vehicles,  $S$ : Set of customers,  $nc$ : number of customers,  $v(S)$ : minimum number of vehicles to attend  $S$ ,  $cv$ : capacity of vehicles,  $c_{ij}$ : cost of the path from the customer  $i$  to customer  $j$ ,  $tc$ : total cost of the composed routes,  $x_{ijk}$ : path from customer  $i$  to customer  $j$  with vehicle  $k$ .

Equation 2 ensures that  $K$  vehicles will be used starting from the distribution center, while the Equation 3 guarantees that each route has its beginning and ending at the distribution center. Equation 4 defines that customers must be attended exactly one time and the Equation 5 keeps the flow ensuring that the vehicle arrives at a customer and out of it, preventing the route ends prematurely. Equation 6 is attributed to the elimination of sub-routes, imposing on the model that routes that do not start and end at the depot are disregarded. In this restriction,  $v(S)$  represents the minimum number of vehicles required to attend a set of customers  $S$ . Equation 7 defines that the vehicle's capacity is not exceeded, Equation 8 defines that all elements of matrix  $x$  are binary numbers and, finally, Equation 9 below is used to evaluate the solutions generated by GA. It reflects the value of

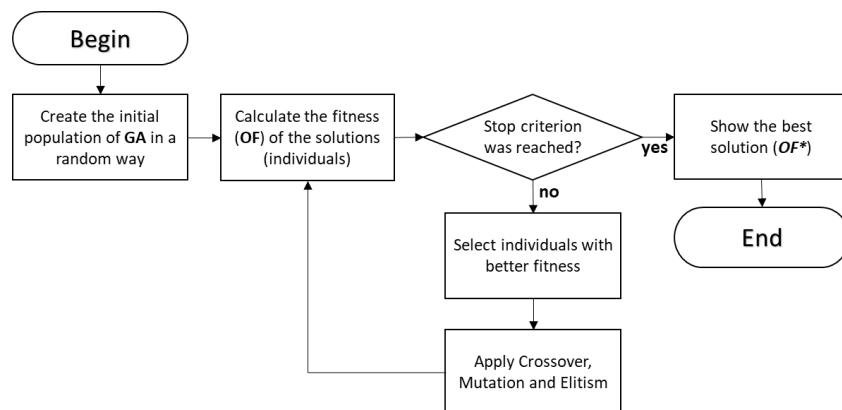
the objective function (*OF*) and combines the number of vehicles used in the solution, violated restrictions (Eqs. 2 to 7) and the total cost of routes (Equation 1).

$$OF = (KW_v) + (nrW_r) + tc \quad (9)$$

where:  $W_v$  is the weight assigned to the number of vehicles used in the solution;  $nr$  is the number of violated restrictions and  $W_r$  is the weight associated to the violated restrictions.

### 2.2. Genetic Algorithm (GA)

The GA consists of a technique from evolutionary computation that simulates the mechanisms of natural selection, genetics and evolution. Its bias lies in the better an individual adapts to the environment, the greater the chances of surviving and generating descendants [1,11]. A GA individual is defined as a chromosome, consisting of genes, and represents a solution of the problem to be solved. The genes represent the variables of the problem, and each position of a gene is called allele that, in binary encoding scheme is represented by a bit. By means of competition, the chromosomes of a population with greater aptitude are selected and crossed each other, generating new chromosomes better than those ones of the previous population. So, at each generation the probability of one or more individual be a solution of the problem is increased [11]. The GA involves four main operations: evaluation, selection, crossover and mutation. In the evaluation operation fitness function is used to measure the aptitude of the individuals of the population, providing information such as the number of new individuals each one can generate according with its aptitude. The selection operation consists of the choice of the best individuals for reproduction. The crossover operation consists in recombination of genes from selected individuals, responsible to reproduce descendants more adapted to the next generation. Finally, a random change in a small number of alleles, with some small probability, is performed to preserve the diversity of the population. This operation, called mutation, is essential to avoid the premature convergence [1,11]. The GA operating scheme is illustrated in Figure 2.

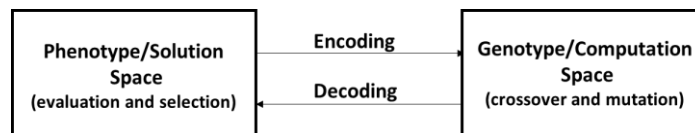


**Figure 2:** The GA operating scheme

It is important to highlight that a suitable convergence of the GA depends on some configuration parameters such as: solution encoding scheme, size of population, evaluation criteria, crossover and mutation rates, among

others. In general, these parameters are empirically defined. Other important concepts associated with GA, related to operations for solution's encoding and decoding are:

- **Genotype:** is related to the population in the computation space, in which the solutions are represented to be easily understood and manipulated by computers [12].
- **Phenotype:** is related to the population in the real-world solution space, in which the solutions are represented to be interpreted in real situations [8,12]. In other words, it is the structure built from the information of the genotype that allows chromosome decoding.
- **Encoding and Decoding:** in most cases, the phenotype and genotype spaces are different. Encoding is an operation that transforms a solution from the phenotype to genotype space, while decoding is responsible by transforming a solution from the genotype to the phenotype space (see Figure 3). The main encoding schemes are: Binary, Value (integer, float, string, etc), Permutation and Tree [8]. Since these operations are carried out repeatedly during the fitness value calculation (evaluation) in a GA, they need to be simple and fast [8].



**Figure 3:** Operations for encoding and decoding solutions generated by GA

### 2.3. Encoding schemes employed for solving VRP

For solving any problem using GA it is essential to define a scheme to represent a solution in the chromosome [8,9,12,13]. In the simplest cases, the binary alphabet {0,1} is used, which for many authors is the most natural representation for GA, although in practice the binary representation is not always the most appropriate. The encoding scheme may vary according to the problem and/or strategy adopted. However, it must allow the computational interpretation of the variables that define the solution of the problem in question. The literature review, considering the last decade, shows that there are several ways of encoding the VRP solutions in GA, as shown in Table 1. It is important to mention that although there are many works in the recent literature exploring the application of GA in the solution of the VRP, they usually focus on the quality of the solution found, presenting no discussion about the elements that can impact the behavior of the population of the AG and, consequently, in its functioning.

Based on Table 1 it is possible to notice that, in general, there is a strong tendency to use a solution encoding scheme for the VRP using vectors of integer numbers. Some authors such as [14], for example, justify the choice for the integer encoding scheme, explaining that it is a more direct representation of customers and/or vehicles, facilitating the decoding step. On the other hand, there are also authors who prefer to use binary coding because it is more natural for AG. Another trend observed is the use of mixed encoding, in which the chromosome is represented by multiple types of data, as the encoding described in [15]. However, there are no studies exploring how each scheme impacts the performance of the GA. The main CVRP solution encoding

schemes (binary and integer), which are considered in this work, are detailed in subsections 2.3.1 and 2.3.2.

**Table 1:** Encoding schemes employed in GA for VRP solution

Work	Data structure		Data type		
	Vector	Matrix	Binary	Integer	Real
Bermudez and his colleagues [16]	X			X	
Kansou e Yassine [17]	X			X	
Lau and his colleagues [18]	X			X	
Wang e Lu [19]	X			X	
Lee and Nazif [2]	X			X	
Ursani and his colleagues [20]	X			X	
Lu and Yu [14]	X			X	
Masum and his colleagues [21]		X		X	
Reiter and Gutjahr [22]	X			X	
Vidal and his colleagues [23]	X		X	X	
Nguyen and his colleagues [15]	X		X	X	
Cho and his colleagues [24]		X		X	
Liu and his colleagues [25]	X			X	
Osaba and his colleagues [26]		X			X
Khalili-Damghani and his colleagues [27]	X			X	
Lima and his colleagues [28]		X	X	X	
Mandal and his colleagues [4]	X			X	
Sánchez-Oro and his colleagues [3]	X			X	
Wang and his colleagues [29]		X		X	
Lima and Araújo [9]		X	X		
Lima and his colleagues [6]		X	X		
Hosseinabadi and his colleagues [30]	X			X	
Total	15	7	5	19	1

**2.3.1. Encoding scheme for CVRP solutions using a matrix of binary numbers**

The CVRP solution encoding scheme illustrated in Figure 4 is a generalization of the schemes employed, for example, in the works [6,10]. Such scheme is based on a three-dimensional matrix of binary numbers with  $M$  columns,  $N$  rows and  $Z$  depth, where  $M$  represents the number of customers to be served,  $N$  represents the customer service sequence and  $Z$  is defined by the number of vehicles needed to meet the total demand. In other words, the column indicates the customer to be visited, the row indicates the order of visiting and the depth indicates the vehicle assigned to serve the customer.

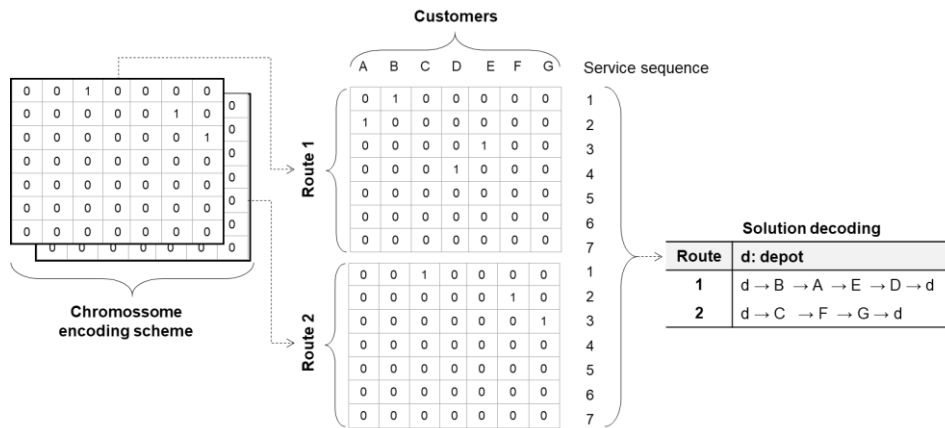


Figure 4: Operation of the encoding scheme using a matrix of binary numbers

2.3.2. Encoding scheme for CVRP solutions using a vector of integer numbers

Some authors such as [2,3,5,14,16,22] employed encoding schemes to represent CVRP solutions that can be generalized by the scheme illustrated in Figure 5. It consists of a vector of integers of length  $N$  representing the number of customers to be served, and each gene in the vector is a number that corresponds to a given customer. The sequence of genes in the vector determines the service order of customers, and the set of customers that make up each route is limited by the capacity of the vehicles, that is, each customer is assigned to a specific vehicle in the sequence in which it appears in the vector and, when the vehicle's capacity is exceeded, a new route is started.

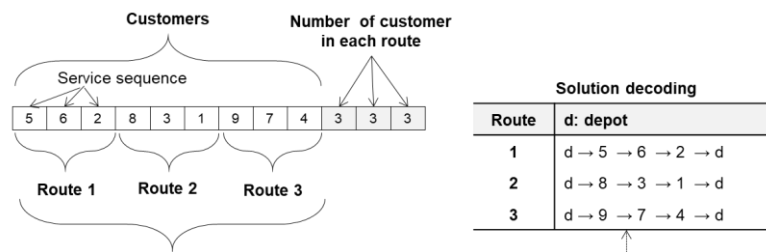


Figure 5: Operation of the encoding scheme using a vector of integer numbers

2.4. Computational tool for visualization and analysis of the GA population's behavior

The developed tool allows the visualization and analysis of individuals of the GA population mapping them from a  $n$ -dimensional space to a two-dimensional space, as illustrated in Figure 6. The results generated by the tool are recorded in a CSV (Comma Separated Values) text file, which can be read by the Microsoft Excel, and in video AVI (Audio Video Interleave) format, providing a more detailed analysis of the populations behavior.



**Figure 6:** Interface of the developed computational tool

In the panel on the right of the screen shown in Figure 6, where the individuals are mapped, the  $x$  axis is represented by the Euclidean distance and the  $y$  axis by the distance of Hamming. For that, it is considered that an individual  $\alpha$  with all the alleles null is represented by the point  $(x_0, y_0)$  and then, each individual  $i$  of a population  $P$  is mapped to a point  $(x_i, y_i)$ , based on the two distances between individuals  $i$  and  $\alpha$ . It is worth mentioning that these distances are widely used in the literature to measure the similarity between individuals from populations generated by metaheuristic algorithms [9]. In addition, the tool presents two graphs that are plotted at run time and allows obtaining the following performance evaluation measures of the GA:

- **Population Diversity (DIV):** this measure expresses, through a rate ranging from 0 to 1, the diversity of the population. Value of 0 for this rate indicates that all individuals in the population are similar, while value 1 indicates that all individuals are completely different.
- **GAP:** expresses how far the result obtained for a problem is from the best result reported in the literature for that problem. In this work it is calculated as follows:  $GAP = (OF - OF_{best}) / OF_{best}$ , being  $OF$  the value calculated by using Equation 9 and  $OF_{best}$  the  $OF$  value of the best solution found in the literature.
- **Explored Area (AE):** this measure reflects, through a rate ranging from 0 to 1, the percentage of the area of the solutions space explored during the execution of the GA. The higher the rate, the greater the area explored, indicating that the algorithm was more likely to explore more promising points in the solution space.
- **Dispersion of the population (DISP):** describes how dispersed the individuals of the population are, that is, the greater the average dispersion, the greater the area explored in the solution space.



- **Percentage of feasible solutions (PSF):** expressed through a rate ranging from 0 to 1, it reflects the capacity of the GA in converting non-feasible solutions in feasible ones.

### 3. Experimental setup

First, a bibliographic survey was carried out and from it were selected the 22 works presented in Table 1. From these works we identified the two main encoding schemes adopted in the CVRP solution (binary encoding scheme – BES and integer encoding scheme – IES), which were described in subsections 2.3.1 and 2.3.2, and are analyzed in this work using the computational tool presented in section 2.4. The BES was chosen because it is more natural for GA and because of its wide use in the literature for solving combinatorial optimization problems, while IES was chosen for being a more direct representation of customers and vehicles that facilitates the decoding operation, as stated in many works that employ this scheme. To analyze the influence of BES and IES encoding schemes on the GA population behavior, it was applied in 3 instances (data sets that describe CVRP scenarios) from the collection proposed by [31]. The instances E-n13-k4, E-n33-k4 and E-n76-k14, presented in Table 2, were chosen because they represent different degrees of optimization difficulty (easy, medium and difficult) and different sizes of scenarios (small, medium and large) taking into account the number of customers and vehicles [32]. For example, the instance E-n13-k4 considers a CVRP scenario with 13 customers and 4 vehicles, being classified as small (few vehicles and few customers). The optimal solutions for the three instances, used for calculating the GAP, were extracted from the work of [33].

**Table 2:** Scenarios described by the instances considered on this work

Data that describe the scenario	Instance		
	E-n13-k4	E-n33-k4	E-n76-k14
Number of costumers	13	33	76
Number of vehicles	4	4	14
Capacity of the vehicle	6000	8000	100
Optimal solution ( $OF_{best}$ )	247	835	1021

**Table 3:** GA parameter configuration

Parameter	Adopted values
Encoding scheme ( $eSch$ )	BES      IES
Population size ( $Pop$ )	200
Number of generations ( $nGer$ )	50
Crossover rate ( $cr$ )	0,90;    0,70;    0,50
Mutation rate ( $mr$ )	0,10;    0,05;    0,01
Elitism rate ( $er$ )	0,10
Selection method	roulette wheel
Number of feasible individuals included into the initial population	2

In order to describe the behavior of the GA population according to the encoding scheme adopted, computational experiments were performed. For this, an experiment planning (Design of Experiments – DoE) was also carried out considering different configurations of the main operators of the GA (crossover and mutation), which are presented in Table 3. This planning resulted in a set of 18 experiments whose results are described in Table 4. To make a fair comparison between the two encoding schemes, the same initial population, randomly generated with uniform distribution, was adopted for starting the GA. In addition, two individuals generated by the Gillett & Miller heuristic [34] were included in the initial population. In that way, the GA already started with feasible solutions.

**4. Results and discussion**

The analysis of the influences of BES and IES on the behavior of the GA population was based on the comparison of the following performance evaluation measures: average GAP ( $\overline{GAP}$ ), average diversity ( $\overline{DIV}$ ), average explored area ( $\overline{EA}$ ) and average computational cost ( $\overline{CC}$ ). These average values were obtained from the measurement values calculated for the three instances considered in this work. The results obtained in the experiments are consolidated in Table 4, in which the values highlighted in bold indicate the best performances of the GA, while the underlined values indicate the worst performances. To assist the understanding of these results the data of the Table 4 are shown graphically in the figures 7 to 9. The solution obtained for each considered instance was not presented here, since that the main focus of this work is not the solution obtained by the GA itself, but how the process for obtaining it was impacted by the encoding scheme. Regarding the average GAP ( $\overline{GAP}$ ), it can be seen in the graph of Figure 6 that, in general, IES provided better results (lower values of  $\overline{GAP}$ ) than those obtained by BES. With IES, experiments 10 and 13 provided the lowest  $\overline{GAP}$  (0,20). Employing BES, the lowest  $\overline{GAP}$  (0,30) was obtained in experiment 7. On the other hand, the worst performance of GA in relation to the  $\overline{GAP}$  were obtained in experiments 17 and 3, in which the values of 0,23 and 0,50 were obtained using IES and BES, respectively.

**Table 4:** Consolidation of the results obtained with BES and IES

Configuration		BES (Experiments 1 to 9)				IES (Experiments 10 to 18)			
cr	mr	Exp.	$\overline{GAP}$	$\overline{DIV}$	$\overline{EA}$	Exp.	$\overline{GAP}$	$\overline{DIV}$	$\overline{EA}$
0.5	0.01	1	0.42	0.24	<u>0.07</u>	10	<b>0.20</b>	0.70	0.20
0.5	0.05	2	0.48	0.24	0.09	11	0.21	0.72	0.21
0.5	0.10	3	<u>0.50</u>	<b>0.25</b>	0.09	12	0.21	<u>0.68</u>	0.24
0.7	0.01	4	0.31	0.24	0.09	13	<b>0.20</b>	0.70	0.22
0.7	0.05	5	0.37	<b>0.25</b>	<b>0.10</b>	14	0.22	<b>0.73</b>	0.24
0.7	0.10	6	0.38	<b>0.25</b>	<b>0.10</b>	15	0.21	0.69	<b>0.27</b>
0.9	0.01	7	<b>0.30</b>	<u>0.23</u>	<u>0.07</u>	16	0.22	0.70	<u>0.17</u>
0.9	0.05	8	0.36	<u>0.23</u>	0.08	17	<u>0.23</u>	<b>0.73</b>	0.18
0.9	0.10	9	0.38	<u>0.23</u>	0.08	18	0.22	<u>0.68</u>	0.21

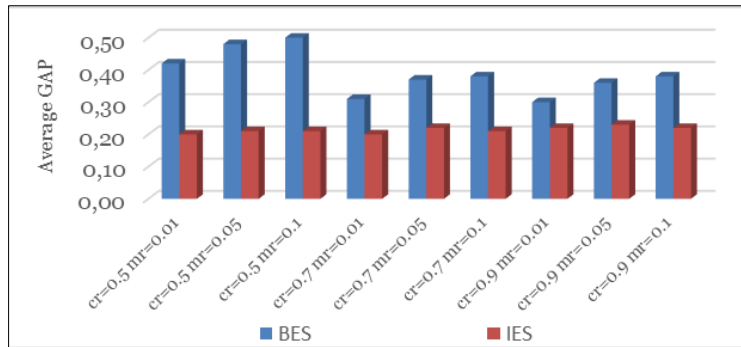


Figure 7: Average GAP

IES also privileged the performance of the GA in terms of diversity ( $\overline{DIV}$ ), as shown in the graph of Figure 8, in which it can be seen that in all IES experiments it provided a greater population diversity than that obtained using BES. IES provided the best GA performance in experiments 14 and 17 reaching  $\overline{DIV}$  of 0.73 and the worst performance in experiments 12 and 18, in which the value of  $\overline{DIV}$  was 0.68. The BES scheme showed the best performances in experiments 3, 5 and 6 reaching a  $\overline{DIV}$  value of 0.25, and the worst performances in experiments 7, 8 and 9 obtaining the value of 0.23 for  $\overline{DIV}$ .

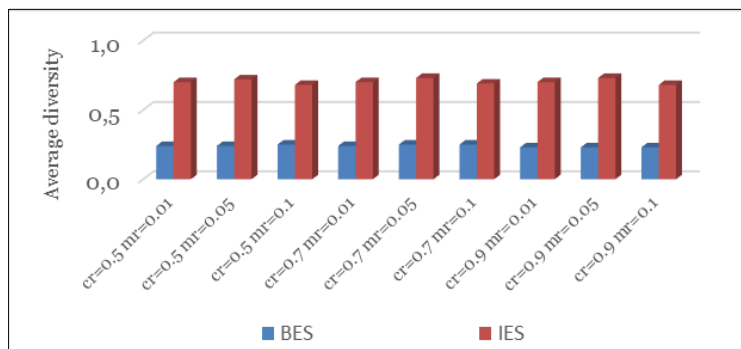


Figure 8: Average diversity

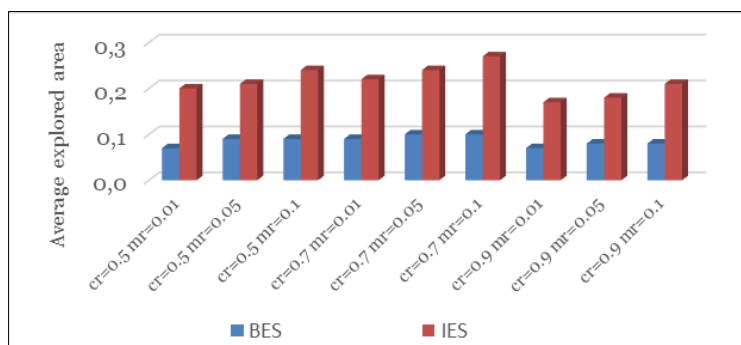


Figure 9: Average explored area

Regarding the average explored area ( $\overline{EA}$ ), it can be seen in the graph of Figure 9 that GA obtained the better performances using IES. This scheme showed its best performance in experiment 15 reaching

$\overline{EA}$  of 0.27. On the other hand, in experiment 16 IES presented its worst performance obtaining  $\overline{EA}$  of 0.17. BES provided the best performance in experiments 5 and 6 obtaining  $\overline{EA}$  of 0.10 and the worst performances in experiments 1 and 7 in which the value of  $\overline{EA}$  was 0.07.

In both encoding schemes, for increasing the explored area and consequently increase the diversity of the population it is suggested to apply diversification heuristics to ensure that a sufficiently large number of regions in the solution space are explored in a balanced way. On the other hand, if the GA presents high diversity and high GAP, it is recommended to use intensification heuristics, in order to concentrate searches in promising regions in the solution space. These information are useful to support the choice and/or proposition of heuristics that allow GA to find better quality solutions for the VRP with less computational effort. With respect to the computational cost, considering the processing time for the three instances, it was possible to observe that BES has an average computational cost of 815 s, which can be understood as high, if compared to the computational cost demanded by IES (56 s). In this sense, even though the use of BES suggests greater simplicity in the operations carried out, it demands a higher computational effort because it is a sparse representation. The results reported in this work indicate that the adopted encoding scheme significantly impacts the behavior of the population of GA and, consequently, in its performance. Thus, the choice of the solution encoding scheme as well as the configuration of the genetic operators of the GA must be done carefully. It is worth mentioning that genetic operators also impact GA's behavior. The lack of this analysis is a limitation of this work that we intend to overcome in future works. The experiments conducted pointed out that even though the GA being an effective, robust, and flexible metaheuristic, it depends on other heuristic mechanisms to better explore the solution space, especially when applied to NP-Hard problems. The problem is to know at what point in the GA evolution cycle a heuristic should be applied and which heuristic is most appropriate. It is in this context that the computational tool described in this work is inserted, since it allows investigating, through the available performance evaluation measures, what type of heuristic should be adopted and at what time it should be applied.

## 5. Conclusions

From the literature review conducted in this work we could verify the adoption of different solution encoding schemes for solving CVRP. Most of the works, especially the most recent, uses integer vectors perhaps because they simplify the decoding step since it is a more direct representation of customers and vehicles. The results of the experiments conducted showed that the encoding scheme impacts significantly in the behavior of the GA. The results also showed that the BES allows an intense exploration in subareas of the solution space, since the measures of diversity and explored area indicate that this scheme keeps the population concentrated in a determined subarea. On the other hand, IES provided better performance of the GA in all analyzed aspects (GAP, diversity, explored area, and computational cost). Perhaps this is the justification for the widespread use of this encoding scheme in more recent works, although this explanation is not provided in such works. This study is important because it can assist in proposing heuristics that allow GA to search for better quality solutions with less computational effort. In future work it is intended to carry out a more detailed analysis of the impact not only of the encoding schemes, but also of the main configuration parameters (crossover and mutation), on the performance of the GA. In addition, it is intended to propose a set of inference rules that can

be used to reconfigure the AG during the evolution cycle, in order to improve its performance in the solution of the CVRP and correlated combinatorial optimization problems.

### **Acknowledgements**

The authors would like to thank the CNPq–Brazilian National Research Council by the research scholarship granted to S. A. Araújo (Proc. 313765/2019-7).

### **References**

- [1]. Holland, J. H. *Adaptation in natural and artificial systems*. Cambridge: MIT Press, 1992, pp. 228.
- [2]. Lee, S. L., Nazif, H. Optimised crossover genetic algorithm for capacitated vehicle routing problem. *Applied Mathematical Modelling*, v. 36, 2011, pp. 2110-2117.
- [3]. Sánchez-Oro, J., López-Sánchez, A. D., Colmenar, J. M. A general variable neighborhood search for solving the multi-objective open vehicle routing problem. *Journal of Heuristics*, 2017, pp. 1-30.
- [4]. Mandal, S. K., Pacciarelli, D., Lokketangen, A., Hasle, G. A memetic NSGA-II for the bi-objective mixed capacitated general routing problem. *J. Heuristics*, v.21, 2015, pp. 359-390.
- [5]. Azad, T., Hasin, M. Ahsan Akhtar. Capacitated vehicle routing problem using genetic algorithm: a case of cement distribution. *International Journal of Logistics Systems and Management*, v. 32, n. 1, 2019, pp. 132-146.
- [6]. Lima, S. J. A., Araújo, S. A., Schimit, P. T. H. A hybrid approach based on genetic algorithm and nearest neighbor heuristic for solving the capacitated vehicle routing problem. *Acta Scientiarum Technology*, v. 40, 2018, pp. e36708.
- [7]. Cheng, X., AN, L., Zhang, Z. Integer Encoding Genetic Algorithm for Optimizing Redundancy Allocation of Series-parallel Systems. *Journal of Engineering Science & Technology Review*, v. 12, n. 1, 2019, pp. 126-136.
- [8]. Kumar, A. Encoding schemes in genetic algorithm. *International Journal of Advanced Research in IT and Engineering*. v. 2, n. 3, 2013, pp. 1-7.
- [9]. Lima, S. J. A., Araújo, S. A. A new binary encoding scheme in genetic algorithm for solving the capacitated vehicle routing problem. In: *International Conference on Bioinspired Methods and Their Applications*. Springer, Cham, 2018, pp. 174-184.
- [10]. Vieira, H. P. *Metaheuristics for solving vehicle routing problems with time window*. 2013, 108f. Dissertation (master's degree in applied mathematics) - Institute of Mathematics, Statistics and Scientific Computing, State University of Campinas, Campinas, Brazil.
- [11]. Goldberg, D. E. *Genetic Algorithms in search, optimization and machine learning*. U.S.A., Addison-Wesley Publishing Company, 1989.
- [12]. Brooker, R. J. *Concepts of genetics*. New York: McGraw-Hill, 2012.
- [13]. Michalewicz, Z, Hartley, S. J. Genetic algorithms + data structures = evolution programs. *Mathematical Intelligencer*, v. 18, n. 3, 1996, pp. 71.
- [14]. Lu, C., Yu, V. F. Data envelopment analysis for evaluating the efficiency of genetic algorithms on solving the vehicle routing problem with soft time windows. *Computers & Industrial Engineering*. v. 63,

2012, pp. 520-529.

- [15]. Nguyen, T.T., Yang, S., Branke, J. Evolutionary dynamic optimization: A survey of the state of the art. *Swarm and Evolutionary Computation*, 6, 2012, pp. 1-24.
- [16]. Bermudez, C., Graglia, P., Stark, N., Salto, C., Alfonso, H. Comparison of recombination operators in panmictic and cellular gas to solve a vehicle routing problem. *Inteligencia Artificial*, v.46, 2010, pp. 34-44.
- [17]. Kansou, A., Yassine, A. New upper bounds for the multi-depot capacitated arc routing problem. *International Journal of Metaheuristics*, v. 1, n. 1, 2010, pp. 81-95.
- [18]. Lau, H. C. W, Chan, T. M., Tsui, W. T., Pang, W. K. Application of genetic algorithms to solve the multidepot vehicle routing problem. *IEEE Transactions on Automation Science and Engineering*, v. 7, n. 2, 2010, pp. 383-392.
- [19]. Wang, C., Lu, J. An effective evolutionary algorithm for the practical capacitated vehicle routing problems. *Journal of Intelligent Manufacturing*, v. 21, n. 4, 2010, pp. 363-375.
- [20]. Ursani, Z., Essam, D., Cornforth, D., Stocker, R. Localized genetic algorithm for vehicle routing problem with time windows. *Applied Soft Computing*, v. 11, n. 8, 2011, pp. 5375-5390.
- [21]. Masum, A. K. M., Shahjalal, M., Faruque, Md. F., Sarker, Md. I. H. Solving the vehicle routing problem using genetic algorithm. *Internacional Journal of Advanced Computer Science and Applications*, v. 2, n.7, 2012, pp. 126-131.
- [22]. Reiter, P., Gutjahr, W. J. Exact hybrid algorithms for solving a bi-objective vehicle routing problem. *Central European Journal of Operations Research*, v. 20, n. 1, 2012, pp. 19-43.
- [23]. Vidal, T., Crainic, T. G., Gendreau, M., Lahrichi, N., Rei, W. A hybrid genetic algorithm for multi-depot and periodic vehicle routing problems. *Operations Research*, v. 60, n. 3, 2012, pp. 611-624.
- [24]. Cho, D. W., Lee, Y. H., Lee, T. Y., Gen, M. An adaptive genetic algorithm for the time dependent inventory routing problem. *Journal of Intelligent Manufacturing*, v. 25, n. 5, 2014, pp. 1025-1042.
- [25]. Liu, R., Jiang, Z., Geng, N. A hybrid genetic algorithm for the multi-depot open vehicle routing problem. *OR spectrum*, v. 36, n. 2, 2014, pp. 401-421.
- [26]. Osaba, E., Diaz, F., Oniera, E. Golden Ball: a novel meta-heuristic to solve combinatorial optimization problems based on soccer concepts. *Applications Intelligence*. v.41, 2014, pp. 145-166.
- [27]. Khalili-Damghani, K., Abtahi, A., Ghasemi, A. A new bi-objective location-routing problem for distribution of perishable products: evolutionary computation approach. *Journal of Mathematical Modelling and Algorithms in Operations Research*, v. 14, n. 3, 2015, pp. 287-312.
- [28]. Lima, S. J. A., Santos, R. A. R., Araújo, S. A. Optimization of the Capacitated Vehicle Routing Problem Using Genetic Algorithms and the Heuristics Downhill and Gillett & Miller. In: *Proc. of XXXV National Meeting of Production Engineering*. Rio de Janeiro: ABEPRO, v. 1. 2015, pp. 1-15.
- [29]. Wang, K., Lan, S, Zhao, Y. A genetic-algorithm-based approach to the two-echelon capacitated vehicle routing problem with stochastic demands in logistics service. *Journal of the Operational Research Society*, v. 68, n. 11, 2017, pp. 1409-1421.
- [30]. Hosseinabadi, A. A. R, Slowik, A., Sadeghilalimi, M., Farokhzad, M., Sangaiah, A. K. An Ameliorative Hybrid Algorithm for Solving the Capacitated Vehicle Routing Problem. *IEEE Access*, v. 7, 2019, pp. 175454-175465.

- [31]. Christofides, N. Vehicle Routing. In: *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*, Lawer, EL., Lenstra, J.K, Kan, A.H.G.R., Shmoys, D.B (Eds), John Wiley & Sons, 1985.
- [32]. Kalatzantonakis, P., Sifaleras, A., Samaras, N. Cooperative versus non-cooperative parallel variable neighborhood search strategies: a case study on the capacitated vehicle routing problem. *Journal of Global Optimization*, 2019, pp. 1-22.
- [33]. Reinelt, G., Wenger, M. K. Maximally violated mod-p cuts for the capacitated vehicle-routing problem. *INFORMS Journal on Computing*, v. 18, 2006, pp.466-476.
- [34]. Gillett, B., Miller, L. R. A heuristic algorithm for the vehicle dispatch problem. *Operations Research*, v. 22, 1974, pp. 340-349.