



# A hybrid approach based on genetic algorithm and nearest neighbor heuristic for solving the capacitated vehicle routing problem

Stanley Jefferson de Araujo Lima\*, Sidnei Alves de Araújo and Pedro Henrique Triguís Schimit

Programa de Pós-Graduação em Informática e Gestão do Conhecimento, Universidade Nove de Julho, Rua Vergueiro, 235/249, 01504-001, São Paulo, São Paulo, Brazil. \*Author for correspondence. E-mail: stanleyjefferson@outlook.com

**ABSTRACT.** This work presents a hybrid approach called GA-NN for solving the Capacitated Vehicle Routing Problem (CVRP) using Genetic Algorithms (GA) and Nearest Neighbor heuristic (NN). The first technique was applied to determine the groups of customers to be served by the vehicles while the second is responsible to build the route of each vehicle. In addition, the heuristics of Gillett & Miller (GM) and Downhill (DH) were used, respectively, to generate the initial population of GA and to refine the solutions provided by GA. In the results section, we firstly present experiments demonstrating the performance of the NN heuristic for solving the Shortest Path and Traveling Salesman problems. The results obtained in such experiments constitute the main motivation for proposing the GA-NN. The second experimental study shows that the proposed hybrid approach achieved good solutions for instances of CVRP widely known in the literature, with low computational cost. It also allowed us to evidence that the use of GM and DH helped the hybrid GA-NN to converge on promising points in the search space, with a small number of generations.

**Keywords:** capacitated vehicle routing problem; genetic algorithms; nearest neighbor; Gillett & Miller; downhill; complex networks.

## Uma abordagem híbrida baseada em algoritmo genético e heurística do vizinho mais próximo para solução do problema de roteamento de veículos capacitado

**RESUMO.** Neste trabalho é apresentada uma abordagem híbrida denominada AG-VMP para solução do Problema de Roteamento de Veículos Capacitado (PRVC) empregando Algoritmo Genético (AG) e heurística do Vizinho Mais Próximo (VMP). A primeira técnica foi empregada para determinar os grupos de clientes a serem atendidos pelos veículos, enquanto a segunda é responsável por construir a rota que cada veículo deve percorrer. Em adição, foram utilizados a heurística de Gillett & Miller (GM) e o algoritmo Descida de Encosta (DE), respectivamente, para geração da população inicial do AG e para o refinamento das soluções providas por ele. Na seção de resultados, primeiro apresentamos experimentos que demonstram o desempenho da heurística VMP na resolução dos problemas de Caminho Mínimo (CM) e Caixeiro Viajante (CV), os quais constituem a principal motivação para a proposição do AG-VMP. Os resultados dos experimentos com a abordagem híbrida demonstraram que ela é capaz de fornecer boas soluções para instâncias do PRVC amplamente conhecidas na literatura, com baixo custo computacional. Tais experimentos também permitiram evidenciar que o uso de GM e DE ajudou o AG-VMP convergir para pontos promissores no espaço de soluções, com um pequeno número de gerações.

**Palavras-chave:** problema de roteamento de veículos capacitado; algoritmo genético; heurística do vizinho mais próximo; Gillett & Miller; algoritmo descida de encosta; redes complexas.

## Introduction

The Vehicle Routing Problem (VRP) consists in defining the routes that a set of vehicles must follow to supply the demand of certain customers, respecting the operational restrictions imposed by the context in which the problem is inserted (Laporte, Gendreau, Potvin, & Semet, 2000).

In recent years, the VRP has attracted an increasing attention from researchers due to the great difficulty of its solution and its presence in

various practical situations (Cordeau, Laporte, Potvin, & Savelsbergh, 2007). In consequence, there has been great effort to develop more robust, agile and flexible algorithms that can be modeled according to the scenario that describes the problem.

The VRP has several variants, which take into account the capacity (each vehicle has a specific capacity), time windows (customers are taken care of in time), vehicle fleet heterogeneous (distinct vehicles), and multiple depots (vehicles can come

from different deposits (distribution centers) (Mandal, Pacciarelli, Lokketangen, & Hasle, 2015).

The Capacitated Vehicle Routing Problem (CVRP) is a variant of the VRP, and it consists, basically, in determining the routes to be followed by a fleet of homogeneous vehicles (in terms of capacity), to serve a given number of customers, without violating the capacities of the vehicles (Lee & Nazif, 2011).

The CVRP belongs to the class of NP-hard problems, which usually require non-polynomial-time algorithms for their solutions. In other words, the difficulty in finding the optimal solution for CVRP grows exponentially as the number of clients increases. Thus, the CVRP has been usually solved with the use of heuristic and metaheuristic algorithms such as Gillett & Miller (Gillett & Miller, 1974), Clarke and Wright (Laporte et al., 2000), Tabu Search, Particle Swarm Optimization, and Genetic Algorithms (Lu & Yu 2012).

Reeves (1993) defines heuristic as a technique that although does not guarantee optimal solutions, it is able to find good solutions (sub-optimal in most cases) with feasible computational time. The metaheuristics, in turn, are characterized by guiding a set of heuristics and have been particularly useful in solving complex optimization problems. Heuristic and metaheuristic are more flexible than the exact algorithms to operate with more complicated and realistic objective functions and restrictions.

The choice of algorithm for solving VRP is a process of extreme importance, seeing that a well-defined algorithm can provide a better relation between the cost and benefit to define the routing. Thus, from all the existing techniques for solving the VRP, the Genetic Algorithms (GA) stand out by its versatility of construction, and the good results that it has been demonstrated in solving of complex problems, including VRP, as can be seen in Lee and Nazif, (2011); Tasan and Gen (2011); Ursani, Essam, Comforth, and Stocker (2011); Lu and Yu (2012); Kuo, Zulvia, and Suryadi (2012); Vidal, Crainic, Gendreau, Lahrichi, and Rei (2012); Reiter and Gutjahr (2012); Osaba, Diaz, and Oniera, (2014); Santicioli et al. (2015).

For Laporte et al. (2000), a research trend is the development of simplest, fast and robust metaheuristics that even with some prejudice to the quality of the solutions, allowing their incorporation into commercial packages.

In this sense, this paper presents a hybrid approach to solve CVRP using Genetic Algorithms (GA) and the following heuristics: Nearest

Neighbor (NN), Gillett & Miller (GM), and Downhill (DH). The NN heuristic consists of a greedy strategy that is coupled to GA for building the route of each vehicle in the solution. The second and third heuristics are employed, respectively, to: (i) generate solutions that are included in the initial population of GA, starting it with some feasible solutions, and (ii) refine the solutions generated by the GA-NN approach, after a certain number of generations without improvement.

In the experimental study, we demonstrated the efficiency of the NN for CVRP. It was applied it on the solution of Shortest Path Problem (SPP) and Traveling Salesman Problem (TSP). In these experiments, it was analyzed the influence of three complex networks coefficients (average clustering, average shortest path, network size) in the performance of NN, considering 2490 graphs. It was observed that NN performance improves as the average clustering coefficient of network (graph) increases. Thus, we could make sure that its use coupled to GA is a suitable alternative, since the all graphs of the instances from TSPLIB<sup>1</sup> considered in this work have clustering coefficients with maximum value.

The obtained results, compared with the best results found in the literature, indicate that proposed approach achieved good solutions for CVRP, with low computational cost. In addition, it was evidenced that the use of GM and DH improved the hybrid GA-NN to converge quickly on promising points in the search space.

## Material and methods

This section presents, firstly, a brief description on the considered problems (SPP, TSP, CVRP and Complex Networks) and techniques (GA, NN, GM and DH). After that, the methodology for conducting the experiments is presented.

### Shortest Path Problem (SPP)

The SPP consists in determining a path between a pair of points (origin and a destination) with shortest distance (Cormen, Leiserson, & Stein, 2001). In general, this problem is represented by a graph with several paths to be evaluated, which represents a computational difficulty. Thus, many researches have focused on the development of efficient algorithms to solve it. SPP is present in several areas such as computer networks, games, water and gas distribution, transport logistic, among others. In computer networks, for example, the SPP

<sup>1</sup> TSPLIB is a public library of sample instances for the TSP, CVRP and related problems from various sources and of various types. It is available in: <http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/>

consists of a routing task, in which information is transferred from a source to a destination following the shortest available path available (Cormen et al., 2001). There are many algorithms for SPP solution, many of them based on greedy strategies.

### Traveling Salesman Problem (TSP)

According to Helsgaun (2006), the TSP is one of the most well-known and intensely studied combinatorial optimization problem in history. It was one of the first types of routing problem to be studied. TSP consists of finding the smallest possible route, from the point of origin, in order to travel through a set of cities (points) and return to the origin, visiting each city exactly once. In other words, the objective of the TSP is to find in a graph  $G = (V, E)$ , in which  $V$  are the set of vertices and  $E$  the set of edges, the path with the shortest cost so that all the vertices are visited only once.

To problems of this type, the optimal solution could be found by enumerating all possible solutions, but even with the technological advance in the last decades, it becomes unfeasible as the number of vertices increases. In this sense, to solve practical problems, where there is a high number of vertices, heuristic and metaheuristics algorithms are normally used to obtain good solutions in an acceptable computational time. In this paper, for example, Nearest Neighbor (NN) was used to solve it.

### Capacitated Vehicle Routing Problem (CVRP)

According to Cunha (2000), routing problems can be seen as a problem of multiple traveling salesman with restriction of capacity (for example, each traveling salesman can only carry a limited capacity of objects). The CVRP is the classic version of VRP and consists of finding a set of routes, where each route is traveled by a vehicle, with the objective to minimize the total cost of determined routes, respecting the following restrictions: (1) Each route must start and finish at the distribution center; (2) Each customer must be visited just once; (3) The sum of the customers' demands included in a route cannot exceed the vehicle's capacity.

Let  $G = (V, E)$  be a graph in which  $V = 0, \dots, n$  is a set of vertices that represent the customers and  $E$  the set of edges that connects customers to each other and to distribution center. To each edge  $(v_i, v_j)$  is associated a cost  $c_{ij}$ . A set of  $K$  identical vehicles with capacity  $cv$  is allocated to the distribution center ( $v_0$ ). Each customer  $v_i$  has a demand  $d_i$  and the demand of distribution center is defined as  $d_0 = 0$ . Mathematically the CVRP can be expressed as follows Equation 1 to 8:

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

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

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

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

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

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

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

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

where:

$d_i$ : demand of customer  $v_i$ ;  $k$ : vehicle;  $K$ : set of vehicles;  $S$ : set of customers;  $nc$ : number of customers;  $m(S)$ : 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 routes;  $x_{ijk}$ : path from the 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 only once and the Equation 5 keeps the flow ensuring that the vehicle arrives at a customer and out of it, preventing the route ends prematurely. The Equation 6 prevents the formulation of routes which do not include the distribution center. In this restriction,  $m(S)$  represents the minimum number of vehicles required to attend a set of customers  $S$ .

To ensure that the number of vehicles used to attend the customers of set  $S$  is not less than  $m(S)$ , the restriction 6 establishes, indirectly, that the capacity of vehicle is not exceeded. However, to let it explicit, the Equation 7 is used to formulate the capacity restriction. The Equation 9 is used to evaluate the solutions generated by the proposed approach. It reflects the value of the objective function (OF) and involves the number of vehicles used in the solution, violated restrictions (Equation 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 given to the violated restrictions.

### Complex networks

Many situations may be mathematically modeled by a network (or graph). Nodes that are connected by edges express a wide range of problems. For instance: the network formed by authors (nodes) and the works they published together (edges); computers (nodes) and network computers (edges); individuals (nodes) and the contact they make with each other during a certain period (edges). However, regular networks seem not to represent the dynamics of such problems, with complex networks, which have random characteristics, being used more frequently (Barabasi, 2003).

The approach proposed by Erdős and Rényi (1959) started with random networks: given  $N$  nodes, a fraction  $p$  of the possible edges link the nodes. Then, considering a regular network with all nodes connected to  $m$  nearest neighbors, a fraction  $p$  of the connections is rewired, allowing long-distance nodes being connected, but maintaining average small distances between nodes. This process forms networks called small-world, proposed by Watts and Strogatz, in 1998 (Watts & Strogatz, 1998). Topological parameters of small-world networks are similar to real networks formed by individuals in a population (Boccaletti, Latorab, Chavezf, & Hwanga, 2006). Moreover, some real networks also present a power law distribution in the degree (number of connections) distribution, that is, highly connected nodes are more likely to receive new edges. This is a scale-free network property (Bollobás, Riordan, Spencer, & Tusnády, 2001), also present in Barabási-Albert networks, whose formation is based on richer get richer rule (Albert & Barabási, 2002).

Usually, the most important topological parameters to analyze a network are: average clustering coefficient, which is the average percentage of connections between neighbors of each node; average shortest path, which is the average shortest path between any two nodes of the network; degree distribution and network size (Newman, 2010).

### Genetic Algorithm (GA)

GA is a population-based optimization method on the Charles Darwin's theory of species evolution.

It is inspired by natural selection model and can be understood as a process where the most adapted individuals have a greater chance of surviving in an adverse environment (Goldberg, 1989). Each new generation is constituted from the most adapted individuals who pass on, consequently, their characteristics to their descendants, aiming at the preservation of acquired qualities. This process happens in a cyclical way so that each new generation is better than previous generation. Usually an objective function is used to calculate the aptitude (fitness) of an individual.

Thus, GA artificially applies the principles of the evolutionary process proposed by Charles Darwin, with the objective of optimizing a search technique in the search space, seeking a solution of good quality.

In a GA population, an individual (or chromosome), represents a possible solution to the investigated optimization problem. The chromosome is usually represented by a bit string in which the genes (chromosome substrings) are associated with the decision variables of the problem and each bit represents an allele of the chromosome. In the beginning, the population is randomly generated and then it evolves by means of applying GA operators (selection, crossover, and mutation). In the selection operation, a percentage of best candidates are selected to be crossed, generating a new population. In other words, the genetic material of best chromosomes is mixed to generate the individuals of the next population. Several selection methods have been proposed in the literature. The most common are roulette wheel, tournament and ranking. In the latter, for example, only a portion of chromosomes with better fitness are selected to generate a new population. The mutation is responsible for randomly changing a small portion of the alleles (usually defined by a percentage) aiming to preserve the diversity of the population, that is, to avoid that individuals of population become too similar to each other (Goldberg, 1989). In fact, the mutation is the key point of the genetic algorithm to avoid local minimum (or maximum) points.

It is worth emphasizing that the great challenge in the use of GA is to choose well its configuration parameters, which include the selection method and the mutation rate. In addition, despite its good performance in solving optimization problems, there may be some difficulty for GA to find optimal or sub-optimal solutions, especially in more complex problems such as CVRP. In these cases, a

good alternative is the use of a local search algorithm to refine the solutions found by the GA.

#### Nearest Neighbor heuristic (NN)

According to Cormen et al. (2001), the constructive heuristic of Nearest Neighbor works basically starting the solution from the origin vertex and adding to the route the vertex closest to the last vertex added, finalizing the algorithm when all the vertices are visited (in the case of TSP) or the destination vertex is found [in the case of (SPP)].

The NN algorithm is easy to implement and run quickly, but sometimes it may lose shorter routes, which are easily noticed by human vision because of its 'greedy' nature. As a general guide, if the last steps of the route are comparable in length to that of the first steps, the course is reasonable; but if they are much larger, it means that there are much better routes.

#### Downhill Heuristic (DH)

The Downhill algorithm is one of the most basic local search techniques (uses a heuristic function for making locally optimal choices). At each step of the algorithm, the solution generated by the current state is replaced by the solution represented by its best neighbor. If we consider, for example, a two-dimensional solution space, the neighbors of a solution  $(x, y)$  could be  $(x+1, y)$ ,  $(x-1, y)$ ,  $(x, y+1)$  and  $(x, y-1)$ .

The algorithm closes when it does not find a neighbor better than that represented by the current state (Russel & Norvig, 1995). Due to its characteristics of local search, DH is widely used to refine solutions generated by metaheuristic algorithms.

The use of a local search algorithm such as DH presupposes, in addition to an evaluation function, a methodology for generating the neighbors of a solution, which is important for the success of the approach. A methodology that generates many neighbors can lead to an excessive increase of the computational cost of the algorithm.

#### Gillett & Miller heuristic (GM)

The GM algorithm (Gillett & Miller, 1974) is based on the notion of economies and is a two-stage heuristic (Cluster first, Route second) widely used for solving Vehicle Routing Problems. GM solves the VRP as follows: in the first stage, the vertices (customers) are clustered using some criterion of proximity and performing the savings calculations.

Then, the sequence of visitation for each group of vertices is obtained by solving the TSP using, for example, the NN algorithm.

#### Methodology for conducting the experiments

In this work, two experiments were conducted. The first set of experiments consist of analyzing the influence of three complex networks coefficients (average clustering, average shortest path, network size) on the performance of NN applied to solve SPP and TSP, considering a total of 2,490 complex networks (graphs) with different topologies (Small World, Barabasi Albert, Scale Free and Erdős & Rényi). We have considered the following intervals for the coefficients: average clustering (from 0.1 to 1.0), average shortest path (from 1.0 to 2.5) and network size (from 5 to 10). In order to allow the comparison of the results obtained by NN for SPP and TSP, the optimal solutions were obtained for each graph using, respectively, Dijkstra and exhaustive search algorithms. In addition, it was considered that a solution would be suboptimal when it had a gap of up to 10% in relation to the optimal solution.

The second set of experiments was conducted to evaluate the performance of proposed GA-NN approach for solving the CVRP, considering a set of 16 instances from Christofides<sup>2</sup> and TSPLIB<sup>3</sup> libraries, with up to 30 customers. The results obtained by GA-NN were compared with the best results found in the literature.

For the development of GA-NN approach, the programming language C/C++ and GAlib<sup>4</sup> library was employed. The GAlib is a free library widely used in solving combinatorial optimization problems. In all experiments we used a desktop computer with the following configuration: Intel Celeron 2955U–1.40 GHz processor; 4 GB of RAM; Windows 7 Ultimate 32-bits operating system.

#### Results and discussion

##### Analysis of the influence of complex network coefficients on the performance of NN

Table 1 presents the results obtained in preliminary experiments conducted to demonstrate the influence of average clustering, average shortest path and network size coefficients in the performance of NN when it is applied to solve SPP and TSP.

<sup>2</sup>Available in: <http://vrp.atd-lab.inf.puc-rio.br/index.php/en/>

<sup>3</sup>Available in: <http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/vrp/>

<sup>4</sup>Available in: <http://lancet.mit.edu/ga/>

**Table 1.** Percentage of optimal and sub-optimal solutions obtained by NN with relation to network coefficients.

Coefficient	Interval	Optimal and sub-optimal solutions	
		SPP	TSP
Average clustering	0.00 – 0.33	52.24%	5.77%
	0.34 – 0.66	56.23%	18.02%
	0.67 – 1.00	61.57 %	43.53%
Average shortest path	1 – 1.46	61.26%	39.56%
	1.47 – 1.93	55.10%	8.85%
	1.94 – 2.5	47.11%	-
Network size	5 – 6	67.73%	42.86%
	7 – 8	61.00%	4.16%
	9 – 10	53.56%	4.16%

As we can be seen in Table 1, the performance of NN heuristic for both problems (SPP and TSP) improves as the clustering coefficient increases. Thus, it is possible to consider that the average clustering coefficient directly influences the performance of NN. This information was essential because we could make sure that the use of NN coupled to GA is a good alternative for solving CVRP, since the all graphs from Christofides and TSPLIB instances have average clustering coefficients equal to 1. Concerning to average shortest path and network size coefficients, the observed behavior indicates that the performance of NN is better in small graphs and/or small paths.

### Proposed GA-NN

In the proposed approach, the chromosome of GA (Figure 1) represents just the set of customers that must be attended by each vehicle, while the sequence or order of visits to customers (route) is solved by the NN algorithm.

Customers	A	B	C	D	E	F	G
Vehicle 1	1	1	0	1	0	0	1
Vehicle 2	0	0	1	0	1	1	0

**Figure 1.** Chromosome representation of the GA.

As illustrated in Figure 1, the chromosome consists of a binary matrix of  $N$  rows by  $M$  columns, where  $n$  represents the number of vehicles and  $M$  the number of costumes. So, alleles (array elements) with a value of 1 in each line indicate the customers that will be attended by a vehicle.

The configuration parameters of GA were defined empirically as follows: Chromosome encoding = Binary Matrix; Population Size = 1200; Population subset for refinement ( $R$ ) = 1/3 of the population; Number of generations without improvement ( $gw$ ) = 30; Number of Generations (used as stop criterion) = 5000; Population rate of replacement = 80%; Elitism (%) = 20; Crossover rate (%) = 80; Selection Method = Roulette; Rate

of Mutation (%) = 1%; Type of Mutation = Flip-bit. Concerning the mutation, 1% of the alleles from population, randomly selected, will have their values changed. The flip-bit operation means that if the original allele value is 1, it will be switched to zero, and vice versa.

In the refinement process by DH algorithm, the neighborhood of a current state (solution) is generated by shifting the elements of chromosome's matrix in a circular manner along the dimension  $M$  (left and right). Thus, at each iteration of DH algorithm, only two neighbor solutions are examined.

The sequence of techniques of proposed GA-NN approach is illustrated in Figure 2.

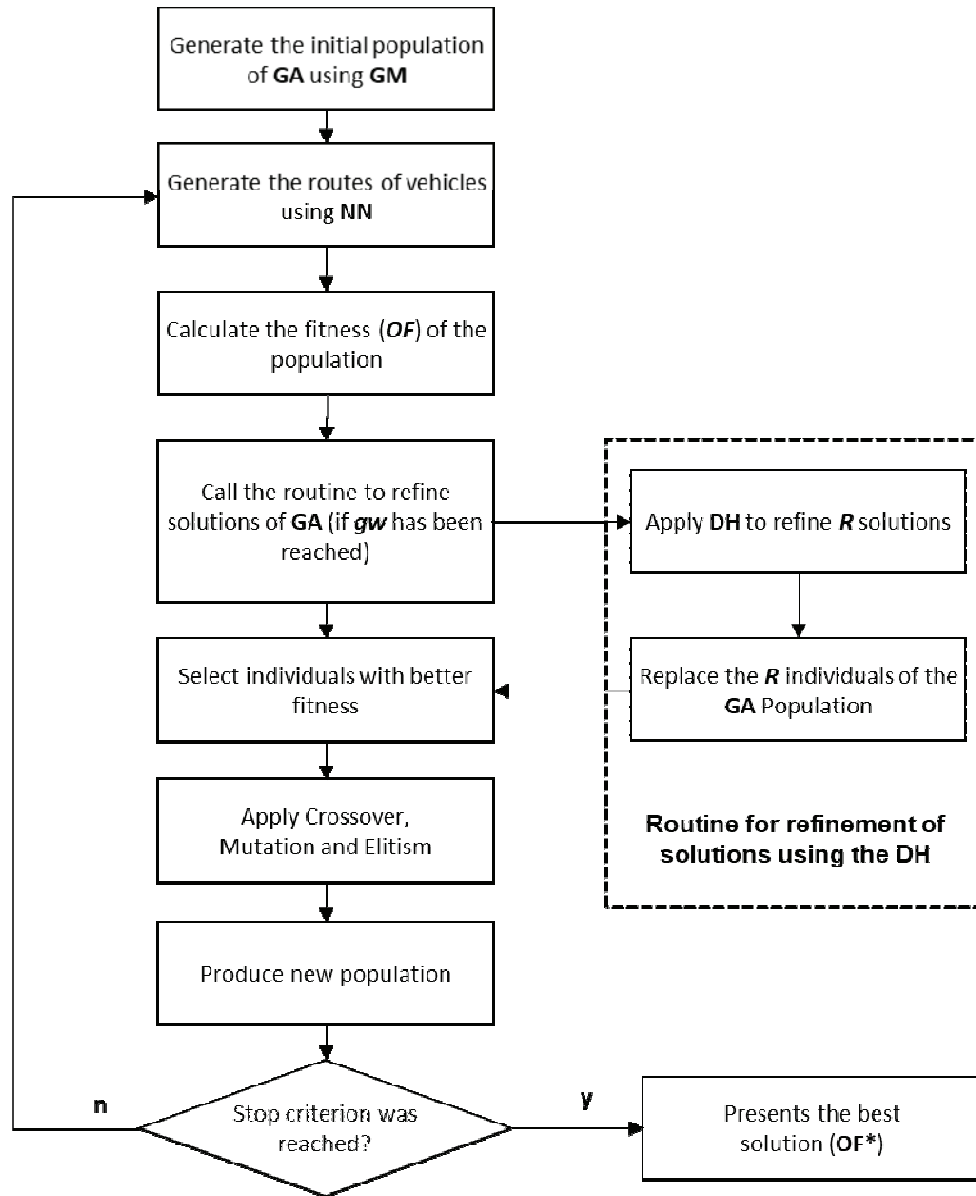
For evaluating the proposed approach, ten tests for each instance were conducted and the results obtained were compared with the best solutions found in the literature.

To evaluate the quality of the obtained solutions, we employed a measure called GAP, widely used in the literature to express how far the result obtained for a problem is from the best result reported in the literature for that problem. In our case,  $GAP = (OF^* - OF_{Best}) / OF_{Best}$ , being  $OF^*$  the best solution obtained by GA-NN and  $OF_{Best}$  the best solution reported in the literature. In addition, the average of the objective function ( $OF$ ) values obtained in the 10 tests ( $OF_{Med}$ ), the standard deviation of  $OF$  values ( $OF_{\sigma}$ ), the computation time and the best solution ( $OF^*$ ) obtained by GA-NN, for each instance, are also presented in Table 2.

From Table 2, we can observe that the approach enabled promising results. The maximum GAP has not exceeded 18 and 62.5% of the instances showed a GAP less than 5%. Regarding the stability, in 62.5% of instances the GA-NN presented the maximum standard deviation ( $OF_{\sigma}$ ) up to 10, indicating its good performance.

With respect to computational cost, as can be seen in Table 2, the processing time is ranging from 27 s for smaller instance (Eil7) to 140 s for larger instance (E-Eil30 n30-k3).

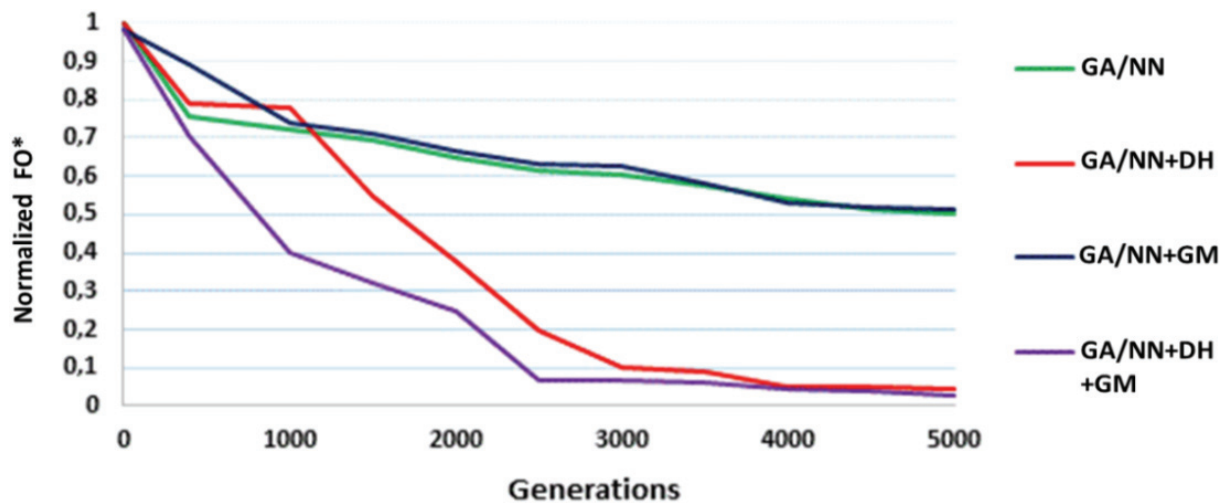
Experiments considering GA-NN with and without the use of heuristics GM and DH were also performed as follows: GA-NN = Only Genetic Algorithm/Nearest Neighbor (without the use of heuristics); GA-NN/DH = Genetic Algorithm/Nearest Neighbor + Downhill; GA-NN/GM = Genetic Algorithm/Nearest Neighbor + Gillett & Miller; GM = only Gillett & Miller; GA-NN/DH/GM = Genetic Algorithm/Nearest Neighbor + Gillett & Miller + Downhill; These results are presented in Table 3 and Figure 2.



**Figure 2.** Sequence of techniques employed in the proposed approach.

**Table 2.** Results of experiments with the proposed approach.

Source	Instance	nc	K	OF <sub>best</sub>	Proposed approach				Time (s)
					OF*	OF <sub>Med</sub>	GAP	OF <sub>σ</sub>	
TSPLIB	Eil7	7	2	114	114	114	0%	0	27
	Eil13	13	4	290	290	296	0%	5	40
	Eil22	22	4	375	375	429	0%	24	90
	Eil23	23	5	875	886	886	1%	0	100
	Eil30	30	3	545	562	582	4%	10	140
	P-n16-k8	16	8	450	450	468	0%	21	50
	P-n19-k2	19	2	212	220	224	4%	2	60
Christofides	P-n20-k2	20	2	216	225	230	5%	4	70
	P-n21-k2	21	2	211	223	223	6%	1	80
	P-n202-k2	22	2	216	222	232	2%	4	90
	P-n22-k8	22	8	590	642	686	9%	29	90
	P-n23-k8	23	8	529	538	565	2%	0	100
	E-n13-k4	13	4	247	290	295	18%	4	40
	E-n22-k4	22	4	375	398	453	7%	19	90
	E-n23-k3	23	3	569	569	583	0%	21	100
	E-n30-k3	30	3	534	574	617	8%	25	140



**Figure 3.** Convergence of GA throughout the generations for solving the Eil30 instance, with and without the use of heuristics.

**Table 3.** Results of the experiments with and without the use of heuristics GM and DH.

Source	Instance	Fbest	GM	GA-NN	GA-NN/ GM	GA-NN/ DH	GA-NN/ GM/DH
TSPLIB	Eil7	114	114	114	114	114	114
	Eil13	290	332	340	308	334	290
	Eil22	375	573	624	472	385	375
	Eil23	875	1039	1074	1012	984	886
	Eil30	545	795	765	750	676	562
	P-n16-k8	450	474	532	462	450	450
	P-n19-k2	212	296	288	253	245	220
	P-n20-k2	216	273	282	255	242	225
Christofides	P-n21-k2	211	271	249	249	230	223
	P-n22-k2	216	271	352	268	244	222
	P-n22-k8	590	635	722	618	630	642
	P-n23-k8	529	553	675	542	542	538
	E-n13-k4	247	332	302	302	290	290
	E-n22-k4	375	522	502	472	460	398
	E-n23-k3	569	762	822	690	569	569
	E-n30-k3	534	792	884	687	574	574
Average GAP (%)			26.6	33.0	17.0	10.0	3.9

These results indicate that, in general, the use of the proposed approach shows better results when employing heuristics GM and DH coupled to GA-NN. These results indicate that the utilization of these heuristics assists the GA-NN to converge on promising points in the search space, with a smaller number of generations, as illustrated in Figure 2.

In summary, it can be observed in Figure 2 that when GA-NN is using DH its convergence is much faster, which leads it to a decrease the computational cost. However, it should be noted that the neighbor structure used in DH plays a crucial role in the local search. In our experiments, to explore the neighborhood of a solution, we just shift the elements of chromosome. This operation, although fast, can generate a high number of infeasible solutions. Thus, further investigations using, for example using k-opt operators, may bring improvements to GA-NN.

## Conclusion

The first contribution of this study consists of the experiments with NN for solving SPP and TSP, which demonstrates that NN heuristic can produce great solutions for CVRP in networks presenting high clustering coefficients. Another important contribution was the second set of experiments, in which we demonstrated that the proposed GA-NN obtained good results for the instances used, considering the quality of solutions and computational time. Through these experiments it was also possible to demonstrate that Gillett & Miller and Downhill heuristics helped the GA to converge on promising points in the search space with a small number of generations.

In this sense, the proposed GA-NN represents a great alternative to be employed in applications that can be used in daily business operations. As future work, we intend to investigate other techniques to obtain feasible solutions for composing the initial



population of the GA, such as Clarke & Wright algorithm; to investigate other types of neighborhood structures for the refinement of solutions, for example the 2-Opt and 3-Opt operators and, finally, to apply the GA-NN in other instances from the literature, well as in real scenarios, aiming to evaluate its applicability in a commercial software.

### Acknowledgements

S. J. de A. Lima would like to thank University Nove de Julho (Uninove) and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (Capes) for his scholarship. S. A. de Araújo would like to thank the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for his research scholarship (Process 311971/2015-6). P.H.T. Schimit is partially supported by grants #303743/2016-6 and #402874/2016-1 of CNPq and grant #2017/12671-8 of São Paulo Research Foundation (Fapesp).

### References

- Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1), 47-97. doi 10.1103/RevModPhys.74.47
- Barabasi, A. L. (2003). *Linked: how everything is connected to everything else and what it means*. New York, NY: Plume Books.
- Boccaletti, S., Latorab, V., Chavezf, M., & Hwanga, D. U. (2006). Complex networks: structure and dynamics. *Physics Reports*, 424(4), 175-308. doi 10.1016/j.physrep.2005.10.009
- Bollobás, B., Riordan, O., Spencer, J., & Tusnády, G. (2001). The degree sequence of a scale-free random graph process. *Journal of Random Structures & Algorithms*, 18(3), 279-290. doi 10.1002/rsa.1009
- Cordeau, J. F., Laporte, G., Potvin, J. Y., & Savelsbergh, M. W. P. (2007). Transportation on demand. *Transportation, Handbooks in Operations Research and Management Science*, 14, 429-466. doi 10.1016/S0927-0507(06)14007-4
- Cormen, T. H., Leiserson, C. E. L. R., & Stein, R. C. (2001). *Introduction to algorithms* (2nd ed.). Cambridge, MA: The MIT Press.
- Cunha, C. B. (2000). Aspectos práticos da aplicação de modelos de roteirização de veículos a problemas reais. *Transportes*, 8(2), 51-74. doi 10.14295/transportes.v8i2.188
- Erdős, P., & Rényi, A. (1959). On random graphs. I. *Publicationes Mathematicae*, 6, 290-297.
- Gillett, B., & Miller, L. R. (1974). A heuristic algorithm for the vehicle-dispatch problem. *Operations Research*, 22(2), 340-349. doi 10.1287/opre.22.2.340
- Goldberg, D. E. (1989). *Genetic Algorithms in search, optimization and machine learning*. Boston, MA: Addison-Wesley Publishing Company.
- Helsgaun, K. (2006). An effective implementation of K-opt moves for the Lin-Kernighan TSP heuristic. *Writings on Computer Science*, 109, 1-99.
- Kuo, R. J., Zulvia, F. E., & Suryadi, K. (2012). Hybrid particle swarm optimization with genetic algorithm for solving capacitated vehicle routing problem with fuzzy demand—A case study on garbage collection system. *Applied Mathematics and Computation*, 219(5), 2574-2588. doi 10.1016/j.amc.2012.08.092
- Laporte, G., Gendreau, M., Potvin, J.-Y., & Semet, F. (2000). Classical and modern heuristics for the vehicle routing problem. *International Transactions in Operational Research*, 7(4-5), 285-300. doi 10.1016/S0969-6016(00)00003-4
- Lee, S. L., & Nazif, H. (2011). Optimized crossover genetic algorithm for capacitated vehicle routing problem. *Applied Mathematical Modelling*, 36(5), 2110-2117. doi 10.1016/j.apm.2011.08.010
- Lu, C., & Yu, V. F. (2012). Data envelopment analysis for evaluating the efficiency of genetic algorithms on solving the vehicle routing problem with soft time windows. *Computers & Industrial Engineering*, 63(2), 520-529. doi 10.1016/j.cie.2012.04.005
- Mandal, S. K., Pacciarelli, D., Lokketangen, A., & Hasle, G. (2015). A memetic NSGA-II for the bi-objective mixed capacitated general routing problem. *Journal Heuristics*, 21(3), 359-390. doi 10.1007/s10732-015-9280-7
- Newman, M. (2010). *Networks: an introduction*. New York, NY: Oxford University Press.
- Osaba, E., Diaz, F., & Onicera, E. (2014). Golden Ball: a novel meta-heuristic to solve combinatorial optimization problems based on soccer concepts. *Applications Intelligence*, 41(1), 145-166. doi 10.1007/s10489-013-0512-y
- Reeves, C. R. (1993). *Modern heuristic techniques for combinatorial problems*. New York, NY: John Wiley & Sons Inc.
- Reiter, P., & Gutjahr, W. J. (2012). Exact Hybrid Algorithms for solving a bi-objective vehicle routing problem. *Central European Journal of Operations Research*, 20(1), 19-43. doi 10.1007/s10100-010-0158-3
- Russel, S., & Norvig, P. (1995). *Artificial intelligence a modern approach*. Trenton, NJ: Prentice Hall.
- Santiciolli, F. M., Costa, S. E., Eckert, J. J., Heron, J. D., Silva, L. C. D. A., Carvalho, M. R., ... Pinho, A. F. (2015). Development of an adaptive genetic algorithm for simulation optimization. *Acta Scientiarum. Technology*, 37(3), 321-328. doi 10.4025/actascitechnol.v37i3.25986
- Tasan, A. S., & Gen, M. (2011). A genetic algorithm based approach to vehicle routing problem with simultaneous pick-up and deliveries. *Computers &*

- Industrial Engineering*, 62(3), 755-761. doi 10.1016/j.cie.2011.11.025
- Ursani, Z., Essam, D., Comforth, D., & Stocker, R. (2011). Localized genetic algorithm for vehicle routing problem with time windows. *Applied Soft Computing*, 11(8), 5375-5390. doi 10.1016/j.asoc.2011.05.021
- Vidal, T., Crainic, T. G., Gendreau, M., Lahrichi, N., & Rei, W. (2012). A hybrid genetic algorithm for multidepot and periodic vehicle routing problems. *Operations Research*, 60(3), 611-624. doi 10.1287/opre.1120.1048
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393, 440-442. doi 10.1038/30918
- Received on April 15, 2017.*  
*Accepted on September 26, 2017.*
- License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.