

G. NILAY YÜCENUR, Ph.D.  
E-mail: nilayyucenur@beykent.edu.tr  
Beykent University  
Industrial Engineering Department  
Ayazağa-Maslak Campus, İstanbul, Turkey

Science in Traffic and Transport  
Original Scientific Paper  
Submitted: 8 Jan. 2019  
Accepted: 21 Nov. 2019

## TWO-STEP META-HEURISTIC APPROACH FOR A VEHICLE ASSIGNMENT PROBLEM – CASE FROM İSTANBUL/TURKEY

### ABSTRACT

*In this paper, a two-step meta-heuristic approach is proposed for vehicle assignment problem with geometric shape-based clustering and genetic algorithm. First, the geometric shape-based clustering method is used and then the solution of this method is given to the genetic algorithm as initial solution. The solution process is continued by genetic algorithm. There are 282 bus lines in İstanbul European side. Those buses should be assigned to six bus garages. The proposed method is used to determine the minimum distance between the bus lines and garages by assigning buses to garages. According to the computational results, the proposed algorithm has better clustering performance in terms of the distance from each bus-line start point to each bus garage in the cluster. The crossover rate changing method is also applied as a trial in order to improve the algorithm performance. Finally, the outputs that are generated by different crossover rates are compared with the results of the k-Nearest Neighbour algorithm to prove the effectiveness of the study.*

### KEY WORDS

*vehicle assignment problem; geometric shape-based clustering; genetic algorithm; crossover rate; the k-Nearest Neighbour algorithm;*

### 1. INTRODUCTION

Migration of rural population to urban areas is still continuing globally. Most major cities, especially in the developing countries cannot cope with the pace of population growth. This creates major problems that affect the city life quality. “Traffic” can be named one of the major problems for the cities. Increased public transport usage is one of the best ways to deal with the traffic problem.

Today, the increasing role of public transport has created a scientific research area of “total travel time optimization”. In this sense, minimization of the total travel time (transfer-waiting-in vehicle

included) is a major focus area. Both customers of public transport (shortened travel time), other residents (less traffic), and city economy (efficient use of funds) benefit from this optimization.

The planners usually use assignment and simulation tools to assess the number of performance measures and make decisions at the tactical, strategic, and planning levels [1]. The assignment problems are most often used for allocating workers to jobs and allocating jobs to machines. This can also be transformed into transportation problems. The purpose of the vehicle assignment models within the transportation problem is to find the lowest distance (optimum) between customers and vehicles by assigning the vehicle to the best route. Vehicle assignment problems appear in the literature as vehicle routing problems.

Besides, vehicle routing problems with multiple assignment options – such as multi depot and mixed fleet configurations – are used in many applications related to emergency assistance, transportation, robotics, production planning, health care, or maintenance [2]. These combinatorial optimization problems require that customers make decisions at two levels, both in determining the routing sequences and in assigning customers to some resources, such as warehouses or vehicle types.

In this paper, a two-step meta-heuristic approach is proposed for the vehicle assignment problem. This research focuses on minimizing the distances among buses and bus garages and trying to determine the number of buses to be assigned to each garage.

As an outline, Section 2 continues with the literature review, Section 3 includes the problem definition about a real vehicle assignment problem in İstanbul/Turkey. In Section 4, the problem mathematical model is introduced with its assumptions. Section 5 includes the solution methodology that

combines geometric shape-based clustering (GSBc) and genetic algorithm (GA) methods. In Section 6, the proposed algorithm and k-NN algorithm were performed, the results were found and the comparisons were made. Finally, the concluding remarks and suggestions for further studies are given in Section 7.

## 2. LITERATURE REVIEW

This paper is about a classical vehicle assignment problem. At this point, when the literature is examined, it is seen that the assignment problem model is converted to routing problem in general. In other words, the vehicle assignment problem consists of assigning, routing, or scheduling vehicles to lines or depots. In the literature, there are several kinds of assignment problems that are examined. For example, according to the constraints capacitated vehicle routing problem (VRP), distance constrained VRP, time window VRP, multi depot VRP, pick-up and delivery VRP, periodic VRP, stochastic VRP, time dependent VRP, according to the state of the roads symmetric or asymmetric VRP, according to routing situation open or close VRP and according to environment static and dynamic VRP [3-19]. Also, there are vehicle routing problems that contain more than one of these features. For example, Lin (2011) solved a VRP with pickup and delivery time windows and coordination of transportable resources; Afshar-Nadjafi and Afshar-Nadjafi (2017) solved a time-dependent multi depot VRP with time windows and heterogeneous fleet, and Zhou et al. (2018) examined multi depot two echelons VRP with delivery options and Behrouz [20-22].

On the other hand, there are lots of solution approaches to VRPs in the literature. At first, VRPs have been solved by linear programming models and branch-and-bound algorithms [23-27], but today, more realistic results are being obtained for the similar problems with metaheuristic algorithms. Anbuudayasankar et al. (2012) used a genetic algorithm; Yusoff et al. (2015) used particle swarm optimization; Jabir et al. (2017) used ant colonies; Wang et al. (2017) used taboo search and simulated annealing; Ng et al. (2017) used artificial bee colonies, and Yassen et al. (2017) used harmony search [5, 28-32].

In this study, by combining the techniques in the literature, this has been used in the problem of vehicle assignment which should be solved in the daily life and has been aimed to be an example to

different countries in solving similar problems. This paper does not only serve the proposed method to the academic community, but it also offers it for the use of the entire humanity.

Briefly, after all the literature review, there are two main objectives:

- Proposing a solution approach for the vehicle assignment problem with geometric shape-based clustering and genetic algorithm.
- Being the first study in the literature because of using geometric shape-based clustering and genetic algorithm in a new research area such as vehicle assignment problems.

The genetic algorithm that exemplifies the evolutionary mechanism in nature is a search method and it is used to find specific data from a data group. The method proposed by John Holland in the early 1970s is now used to find solutions to problems in many different areas. For example, Kowalski et al. (2018) used GA method for optimizing twin-screw food extrusion processing. With GA, Arthur et al. (2019) evaluated nuclear data, and Abreu et al. (2020) scheduled the open shops [33-35].

Cluster analysis or clustering is the problem of grouping a set of objects. In the problem, in order for the objects to be placed in the same set, they must be more similar to each other than the elements in other sets. In shape-based clustering, the objects are clustered within the same shape. In recent years, Meesrikamolkul et al. (2012) used shape-based clustering method for time series, and Xu et al. (2015) used the same method for analysing structural data [36-37].

The k-NN algorithm proposed by Cover and Hart in 1967 is used by using data from a given sample set. The distance of the new data to be included in the sample data set is calculated according to the existing data, the k-neighbourhood is examined and the nearest neighbour is included in the set. Although it is used for solving both classification and regression problems, it is mostly used in solving classification problems in industry. For example, Tharwat et al. (2018) recognized human activities in mobile crowd sensing environment with k-NN algorithm. Huan et al. (2019) classified the polarized switching and Bania and Halder (2020) grouped the medical data using the same technique [38-40].

## 3. PROBLEM DEFINITION

This research focuses on determining whether the number of buses to be assigned to each garage is determined by considering the distances between

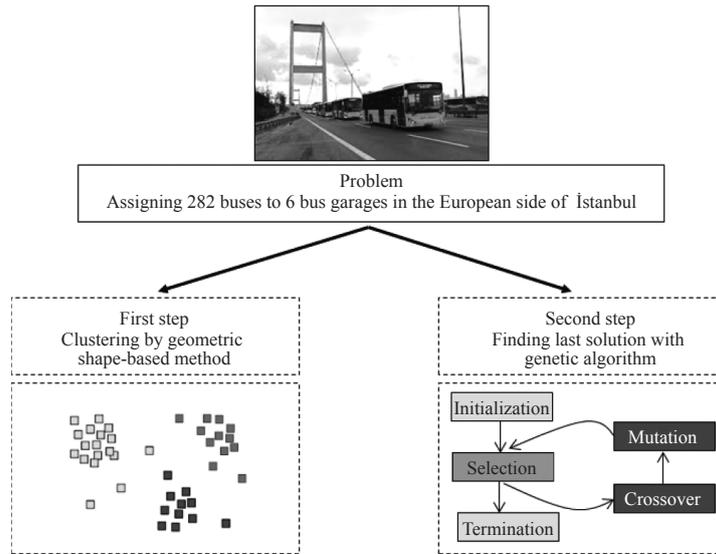


Figure 1 – Research model

the bus garages and the bus line start points. A two-step meta-heuristic approach has been proposed for solving the problem in this study, and the solution has been found by using clustering and genetic algorithm approaches together.

In this study, firstly, geometric shape-based clustering (GSBc) method is used and then the solution of this method is given to the genetic algorithm as an initial solution, and the solution process continues by genetic algorithm (GA).

Figure 1 shows the research model according to the main objectives.

#### 4. PROBLEM FORMULATION: MATHEMATICAL MODEL

In this paper, minimizing the total travelled distance between each bus garage and the bus-line start points is an objective. The buses which have different line start points try to be assigned to bus garages that are located in different places. For minimizing the travelled distance between each bus garage and the bus-line start points, the buses have to be assigned to the closest bus garage.

The buses which are parked in the same garage need to be clustered in a group. This is the main problem of this research. The data are acquired from the İstanbul Electric, Tramway and Tunnel Establishments (İETT) which is responsible for İstanbul's public transportation alternatives such as bus, mini-bus, metro, tram, metrobus, train...etc. In this study, the location data of 282 buses that serve in İstanbul European side and the six bus garages where these buses could park and the distance data from

each bus garage to each bus-line start point are used. The variables and parameters which are used in the mathematical model can be defined as follows:

$B$  -  $(1, \dots, b)$  set of buses;

$G$  -  $(1, \dots, g)$  set of bus garages;

$C$  -  $(1, \dots, c)$  set of circles;

$P$  -  $B \cup G$  set of all points;

$d_{i,j}$  - Euclidian distance between bus-line start point  $i$  and bus garage  $j$  where  $i \in B$  and  $j \in G$ ;

$L_c$  - set of buses in bus garage circle  $c$  ( $c \in C$ );

$X_{ij}$  - 1, if bus which starts serving from line start point  $i$  parks in bus garage  $j$  ( $i \in B$  and  $j \in G$ ); 0, otherwise.

After specifying the variables and the parameters, the problem can be formulated as below:

$$\min Z = \sum_{i=1}^B \sum_{j=1}^G d_{ij} X_{ij} \quad (1)$$

subject to

$$\sum_{j=1}^G X_{ij} = 1 \quad (i \in B, j \in G) \quad (2)$$

$$\sum_{i \in P} X_{ij} \leq 1 \quad (i \in B, j \in G \text{ and } B \in P, G \in P) \quad (3)$$

$$C = G \quad (4)$$

Minimizing the total distance between each bus garage (Equation 1) and the bus-line start points which are parked in the bus garage is the objective function. Constraint 2 allows each bus to be parked in a bus garage. At the end of the day, there should be no buses that have not returned to the garage. Each bus must be available in the garage to start the next day's tour. Constraint 3 requires that from each

bus-line start point a bus can leave at most once and *Constraint 4* represents that the circle number is equal to the bus garage number.

The bus garages and the bus-line start point number and the locations are known. In this research, according to the assumption that each bus garage has unlimited parking places, so that the garages are large enough to offer parking to all the buses. Each bus has to be assigned to only one bus garage and each bus has to start its route from the garage which it is assigned to.

In the current application the buses are assigned to the bus garage nearest to their last stop, but the next day they have to start their tour from their line starting point. Because of this reason, the assignment of the buses to the vicinity of their line starting points is a very critical problem. Finding more productive and more economic solutions is important for the municipal economy and resources.

### 5. SOLUTION METHODOLOGY

In this section, the proposed algorithm for the solution is described in two steps. In the first step, the geometric shape-based clustering method is mentioned, and in the second step, the genetic algorithm method which starts to work by accepting the clustering method results as the initial solution to itself is given. In Section 5.3 the k-NN algorithm is examined as a comparison technique for the proposed algorithm.

#### 5.1 Initial solution: Geometric Shape-Based clustering (GSBc)

The basis of the first step of the two-step meta-heuristic approach is based on Thangiah and Salhi's (2001) method about adaptive clustering. In this method, there are circles with different radii as the geometric shape [41]. In the adaptive clustering method the main goal is to minimize the routing cost with the circles which will be the chromosome to genetic algorithm. If the chromosomes have the ability to adapt the route to the circles, this method can be used effectively for routing the vehicles.

In this paper, the GSBc method is used with circles as the first step of the meta-heuristic algorithm. The two-step meta-heuristic algorithm is different from the Thangiah and Salhi's method. Because the researchers' method uses only the first step of the solution as a clustering method, the proposed algorithm's first step is not associated with the ob-

jective function of the vehicle assignment problem. Besides, in the proposed algorithm the circles with a random radius are drawn from the bus garages. This application is different from the researchers' method. Unlike the study of Thangiah and Salhi (2001) clustering method, this algorithm is used only for the clustering step of the assignment problem. Furthermore, in this algorithm the circles are drawn from the bus garages with random radii. The bus garages are the origins of the circles and the number of garages is equal to the number of circles.

In the first step of the two-step meta-heuristic algorithm, the circles with different radii are drawn from the bus garages and these circles do not overlap. The radii of the circles are randomly determined with the integer numbers ranging from 1 to 10. When assigning the buses to be clustered within a circle, there are three different conditions that need to be considered. The bus location can be outside the circumference of the drawn circle from the centre of the bus garage, it can be on the circumference of the drawn circle from the bus garage centre or inside a drawn circle from the bus garage centre.

Figure 2 presents an example which is obtained from two bus garages, G1 and G2, and three buses such as bus1, bus2, and bus3. The assignment of the buses to the bus garages is similar with the Thangiah and Salhi's method [41]. If one bus location is inside a circle which is drawn from the bus garage, this bus will be assigned to the respective circle as shown in the example in Figure 2. For example, bus1 will be assigned to circle G1. If one bus location is on one circle's circumference which is drawn from the bus garage, this bus will be assigned to the respective circle as shown in the example in Figure 2. For example, bus2 will be assigned to circle G1. If a bus is not inside or on the circumference of any circle, then this bus distance to all circle circumferences are calculated. The bus is assigned to the bus garage

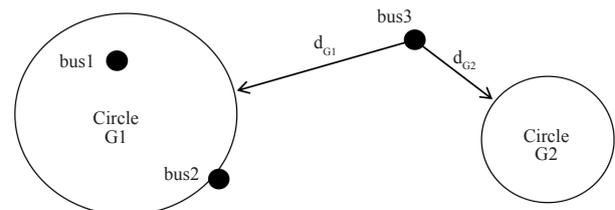


Figure 2 – Schematic representation of the process of assigning buses to bus garages

whose Euclidean distance is the closest between the bus and all circles. For example, bus3 is outside of G1 and G2, as seen in *Figure 2*.

If the distance between bus3 and circle G1 circumference which is drawn with a random radius from bus garage 1 location is  $d_{G1}$  and similarly the distance between bus3 and the circumference of circle G2 which is drawn with a random radius from bus garage 2 location is  $d_{G2}$ , then bus3 will be assigned to G1 if  $d_{G1} < d_{G2}$ , or to G2 if  $d_{G2} < d_{G1}$ .

The algorithms are developed on Python 2.7 and simulated on Intel core i7-7500 CPU @ 2.70GHz. *Algorithm 1* shows the pseudo-code of the first solution step of this paper about geometric shape-based clustering.

*Algorithm 1 – The pseudo-code of the first solution step*

```

START
  READ Distance data
  SET lineGarageDistance
  SET garageIDs
  INIT Random garage diameters
  FOR line IN BusLines DO:
    SET Distance = garageSemiDiameter - lineGarageDistance
    SET lineGarageID = min(Distance)
  RETURN lineGarageIDs
END
    
```

## 5.2 Genetic Algorithm (GA)

In the two-step meta-heuristic approach, GA is used for grouping the buses to the bus garages after the GSBC process. After the initial assignment of all the buses to garages using GSBC method, the genetic algorithm process begins. The buses are allocated to bus garages by drawn circles in the first step based on the geometric shape-based clustering. This initial assignment that needs improvement is the initial solution for GA.

If the number of bus garages is supposed to be  $m$ , there are  $2^m$  different line-up combinations in the problem and these combinations generate the population pool of the genetic algorithm. The fitness value is calculated for each population in order to obtain the objective function that minimizes the total distance between bus garages and bus-line starts. After calculating the fitness value, the GA operators are executed respectively for all populations. First of all, after the selection of the chromosomes by the weighted roulette wheel technique from the binary-coded string population, the crossover operator starts to run by single point version. Secondly, the mutation operator is applied to the offspring which have been obtained from the crossover operator by changing one gene of them from 0 to 1 or from 1

to 0. The fitness value is recalculated after the application of the crossover and mutation operators. The genetic algorithm process is stopped when it reaches the maximum produced generation number, because this number is the stop rule for the algorithm. Otherwise, the algorithm is repeated with all its steps.

By subjecting individuals to the fitness function, the fitness value that determines how close the solution is to the optimal solution is determined. The initial population-generated genetic algorithm works with three evolution operators. These are: selection, crossover and mutation operators.

*Representation of the chromosomes:* The chromosomes are coded into binary strings. In this type of coding, a binary string which changes with a value of 0 or 1 identifies each chromosome. *Figure 3* shows this type of coding.

```

Chromosome A:  1101010001110100
Chromosome B:  0101001111011101
    
```

*Figure 3 – Binary string coding*

*Generating an initial population:* In two-step meta-heuristic approach, the initial population is generated by geometric shape-based clustering circles with random-length radii.

*Selection of chromosomes for crossover operator:* The selection process is performed by the weighted roulette wheel technique. In the weighted roulette wheel technique the chromosome fitness scores are calculated and these chromosomes are weighted according to their fitness scores. The selection is made according to these weighted scores. Thus, this selection is made between chromosomes with higher fitness scores, and the chances of producing better generations will increase.

The selection process starts with the spinning of the roulette wheel. The roulette wheel performs up to the number of the population size; in each spinning for a new population a single chromosome is selected by its weighted fitness score.

This is a stochastic method. Individuals are placed on a roulette wheel so that the area of each slice is proportional to the conformity value, with successive lengths proportional to the conformity values. The individual who owns the range of the random number generated is selected. The process is repeated until the desired number of individuals is reached.

*Crossover:* In this paper, the single point crossover method is used for crossover operator. The purpose of crossover process is to produce offspring by altering the chromosome genes and thus the offspring can be produced with higher fitness values than the parent chromosomes with high fitness values. In a single point crossover, the chromosomes match randomly each other from the population pool. For each chromosome pair selected, a random location is selected from the intervening genes except for the first and the last gene. This point shows the crossover point and the crossover is performed from this point. The genes that come after this point are mutually displaced in both chromosomes. The offspring are obtained when the genes are crossed by a single point crossover. Chromosomes must be of the same length for this operation. *Figure 4* shows a single point crossover example.

*Mutation:* This operator has the ability for the generation of new chromosomes from the existing population pool. The mutation process increases the genetic diversity and prevents the problem from converging to any local optimum. In this paper, mutation is performed by changing one gene of the chromosomes randomly. The selected chromosome one gene is changed from 0 to 1 or from 1 to 0 with mutation. *Figure 5* shows mutation example which is made by changing one gene.

*Algorithm 2* shows the second solution step pseudo-code which is about the genetic algorithm.

Chromosome A (parent A):	<b>110101</b>	<b>0001110100</b>
Chromosome B (parent B):	010100	1111011101
Child 1:	<b>110101</b>	1111011101
Child 2:	010100	<b>0001110100</b>

*Figure 4 – Single point crossover example*

*Algorithm 2 – Second solution step pseudo-code about two-step meta-heuristic approach*

```

START
SET InitialChromosome = array(lineGarageID)
INIT Population = combination(InitialChromosome)
FOR person IN Population DO:
    SET fitnessValue(person) = sumproduct(lineGarageID, lineGarageDistance)
SET roulette = weighted(Population)
CHOOSE p Parents FROM population
SET RANDOM(mutationFactor)
IF RANDOM(mutationFactor) > 0.50 THEN DO:
    Population = mutate(Population)
SET crossoverRate
FOR p IN size(Parents)/2 DO:
    SET parent\1 = RANDOM SELECT(Parents)
    DEL parent\1 FROM Parents
    SET parent\2 = RANDOM SELECT(Parents)
    DEL parent\2 FROM Parents SET child\1, child\2 = crossover(parent\1, parent\2)
    APPEND child\1 AND child\2 TO Children
    RETURN Children
FOR ALL children IN Children DO:
    SET fitnessValue(child) = sumproduct(lineGarageID, lineGarageDistance)
    REPLACE person WITH child IN Population
    IF fitnessValue(person) < fitnessValue(child)
RETURN min(fitnessValue(person))
END
    
```

Chromosome (parent): 1101010001110100  
 Child: 1101011001110100

*Figure 5 – Mutation example with the changing one gene*

### 5.3 K Nearest Neighbour (k-NN) algorithm

The k-NN algorithm is used extensively in the classification because of ease of implementation, its simplicity, and a powerful and useful learning process. In addition, the k-NN algorithm is particularly preferred for classification applications due to the lack of training, analytically traceable, adaptable to local knowledge, parallel execution, and resistance to noisy training data.

According to this algorithm used in the classification, the characteristics of the new individual to be categorized should be taken into consideration from k priorities of the previous individuals. The distance criterion, the number of neighbours (k) and the weighting method are effective and important parameters in the performance of the k-NN algorithm.

The Euclidean distance is the most frequently used distance measure in the classification and clustering algorithms. In this study, the distances between the bus-line start points and the bus garages were calculated with the Euclidean distances with the coordinate information of the bus-line start points.

Another parameter used in the k-NN algorithm is the neighbour number (k). In the algorithm, the classification is made based on the value of the neighbouring number (k) parameter. In the classification process, for k=1, only the closest neighbour is assigned, all the data in the dataset are considered as the number of sample approaches to the number of samples (N), and the selection is made according to the order.

The last parameter in the algorithm is the weighting method. It is intended that the adjoining neighbouring specimens, that are closer to the classifier being weighted by assigning weight values to the neighbours, will make more contributions on the basis of the majority. According to the most commonly used weight assignment method, each neighbour's weight is  $1/d$  or  $1/d^2$ ; where  $d$  is the distance between neighbours.

## 6. COMPUTATIONAL EXPERIMENTS

This paper is about a vehicle assignment problem which involves more than one bus and bus garage. The question is which bus should be assigned to whichever bus garage according to the distances between bus garages and the bus-line starts. Minimizing the total distance between the bus-line start points and bus garages is an objective of this research.

In order to measure the proposed algorithm effectiveness, after the establishment of the mathematical model, the vehicle assignment problem is solved by the k-NN algorithm firstly and then, secondly, the same problem is solved by the two-step meta-heuristic approach. The research results obtained from the two methods are presented in this section.

The proposed two-step meta-heuristic approach was applied to the data which were taken from the İstanbul Electric, Tramway and Tunnel Establishments (İETT) which is responsible for İstanbul public transportation alternatives. In these data, the distances between the starting points of 282 bus lines which serve the European side of İstanbul and the six bus garages where they can park is given in kilometres.

### 6.1 Results

In this paper, the two-step meta-heuristic approach is used for the assignment problem. In the first step clustering is performed by geometric shape-based approach and the results of this step are used for an initial solution of the genetic algorithm. In this problem, according to the known distances between bus garages and bus-line start points, it is determined which buses should park in which garages.

This is the first study which proposes a two-step meta-heuristic approach for the assignment problems based on a GSBc and GA. The

computational time of the proposed approach was obtained for high crossover rate as 5.32 min and for low crossover rate as 3.42 min.

### GSBc

In the two-step meta-heuristic approach, the initial population is generated by GSBc and then the GA is performed with all its parameters. Figure 6 shows the bus-line start points and bus garage locations.

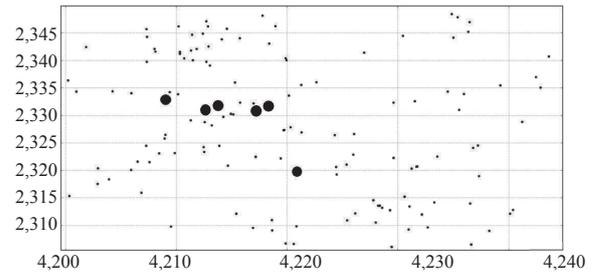


Figure 6 – Bus-line start points and bus garage locations

After the determination of the location of the bus-line start points and the bus garages, in the first part of the study, circles were drawn with random radii ranging from 1 to 10 from bus garages based on geometric shape-based clustering method. In Figure 7, circles with random values have a user interface and the first step final solution of the proposed approach. Figure 8 shows a detailed representation of the distribution.

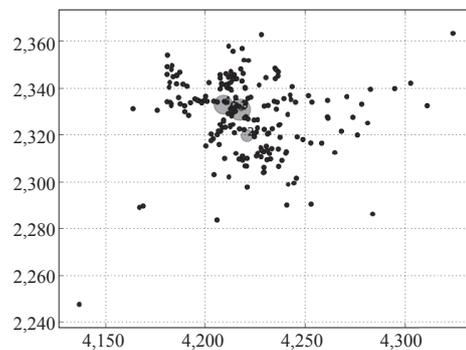


Figure 7 – Drawn circles with random radii from bus garages

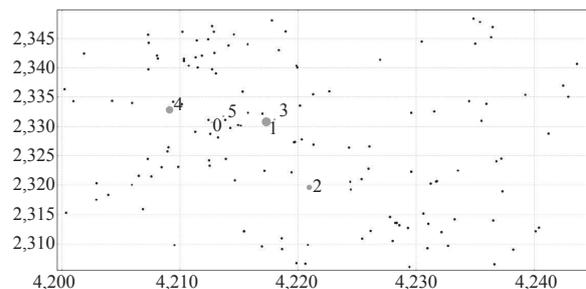


Figure 8 – Detailed representation of the distribution of the buses to bus garages

GA

After the first step of the proposed approach was completed, the second step, the GA, was performed after the initial solution was obtained from the GSBc method.

In the genetic pool of a population size of 1,000, as a control operator, the crossover rate and the mutation rate were left random. The results obtained from the proposed algorithm are shown in Figure 9.

As seen in Figure 9, with random crossover rate, the best solution was obtained in the 566<sup>th</sup> iteration.

In order to prove the effectiveness of the crossover operator in the proposed two-step meta-heuristic approach, the problem has been also computed with a crossover rate set to 60% and to 3%, and the obtained results are shown in Figures 10 and 11.

As seen in Figure 10, with 60% crossover rate, the best solution was obtained in the 609<sup>th</sup> iteration.

As seen in Figure 11, with 3% crossover rate, the best solution was obtained in the 225<sup>th</sup> iteration.

K-NN algorithm results

In this paper, the distances between the bus-line start points and the garages are calculated to solve the problem of assigning the buses to the bus garages, the  $k$  values are taken 15, the line starts are not weighted and the buses are assigned to the nearest garage under these conditions. The pseudo-code of the k-NN algorithm and the results of the algorithm can be seen in Algorithm 3 and Figure 12, respectively.

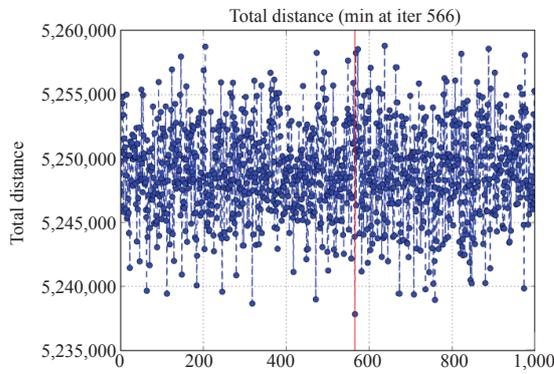


Figure 9 – Results with random crossover rate

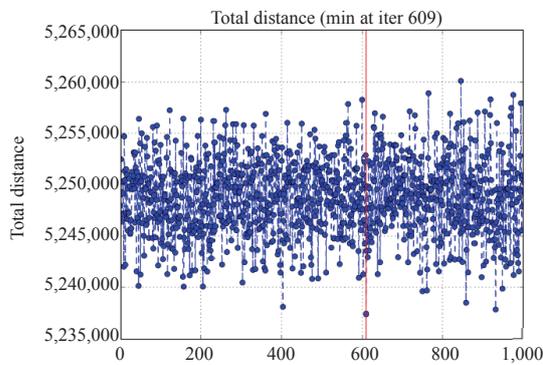
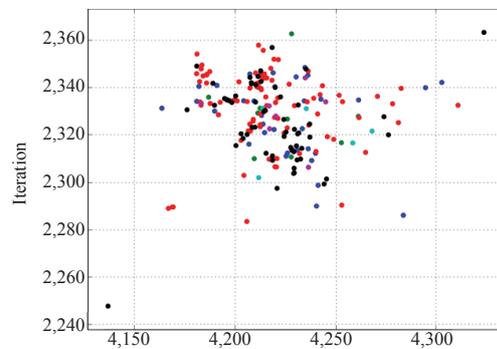


Figure 10 – Results with 60% crossover rate

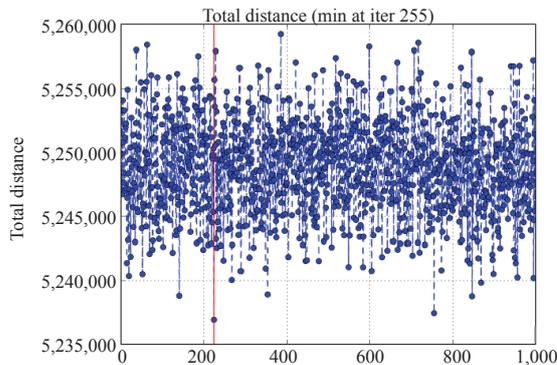
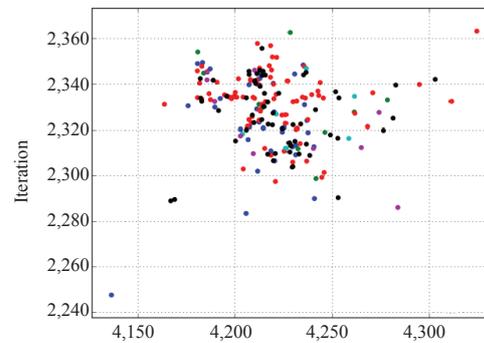


Figure 11 – Results with 3% crossover rate

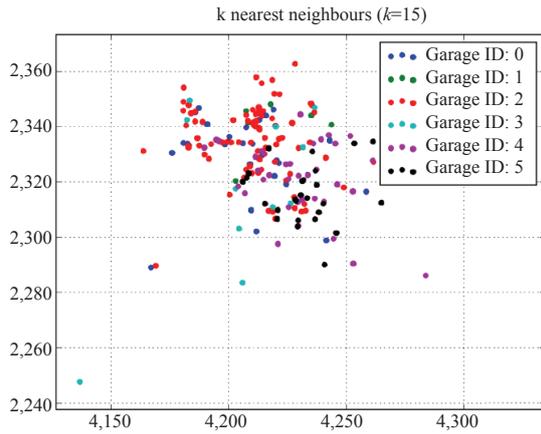


Figure 12 – Results of the k-NN algorithm for an assignment problem of the buses to the bus garages

Algorithm 3 – Pseudo-code of k-NN

```

START
  READ lineGarageDistance
  READ lineGarageIDs
  SET trainSize
  SET trainData = sample(lineGarageIDs, trainSize)
  SET testData = LIST(lineGarageIDs IF line NOT IN trainData)
  SET kNearestNeighbor
  FOR line IN testData DO:
    SET neighborCount = count(neighbour, garageID)
    SET lineGarageID = max(neighborCount, garageID)
  RETURN testData
END
    
```

In this paper, the nearest k-NN algorithm is used for comparison. The main reason is its simplicity that makes it easy to understand. The implementation of this algorithm on numerical data is easier than the implementation on categorical data. However, storing the data to be used in the algorithm takes up a lot of memory. Another disadvantage of k-NN algorithm is that it classifies a new object while the other objects distance calculation and comparison take a lot of time.

6.2 Comparison and analysis

All distances between the bus-line start points and bus garages are calculated in k-NN algorithm and the number of *k* that determines how many

closest neighbours will be searched is set to 15. The k-NN algorithm results are shown in the last column of Figure 13.

In the proposed two-step meta-heuristic algorithm genetic algorithm step, another comparison study was made by changing the crossover rates. With the changing of the crossover rates, it has been found that better results can be achieved with smaller distances between bus garages and bus line start points. The proposed algorithm results are shown in column 1, column 2 and column 3 of Figure 13 with different crossover rates.

In this paper the k-NN algorithm is used for comparison with the proposed two-step meta-heuristic algorithm. In the proposed algorithm, the crossover rate is set to a random value, 60% and 3% for reaching similar but unexplored regions of search space and the comparison results are presented in Figure 13.

In all genetic algorithm steps of the study, the genetic operators such as the population size was set at 1,000, the mutation operator was performed by changing one gene of the chromosomes randomly and the band gap was set at 60. On the other hand, the crossover rate was set at 60%, 3% or a random value.

As seen in Figure 13, the best solution for this assignment problem was obtained by the proposed two-step meta-heuristic algorithm with 3% crossover rate. The result obtained by the proposed algorithm with 3% crossover rate is 12.90% better than the result obtained by k-NN algorithm.

Moreover, the result obtained with 3% crossover rate in the proposed algorithm is 3.56% better than the result obtained with 60% crossover, and 2.39% better with the result obtained with random crossover.

As a conclusion, the proposed two-step meta-heuristic method provided a positive contribution to the use of the solution by giving very successful results according to the criteria quality and calculation time, code simplicity, flexibility and analyticity.

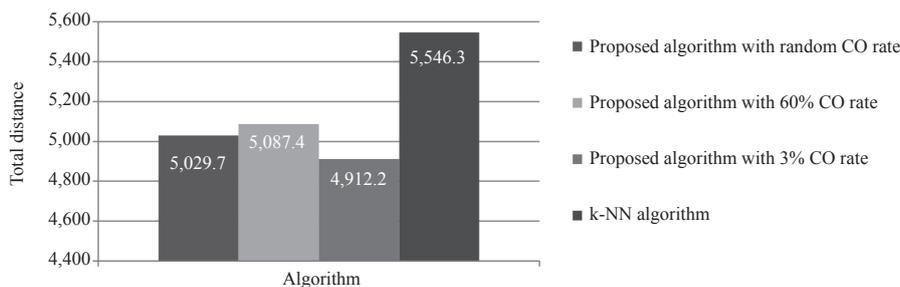


Figure 13 – Comparison results of algorithms

## 7. CONCLUSION

In this paper, the two-step meta-heuristic approach was proposed for a vehicle assignment problem with geometric shape-based clustering and genetic algorithm. Firstly, the geometric shape-based clustering method was used and then the solution of this method was given to the genetic algorithm as an initial solution, and the solution process was continued by the genetic algorithm.

The proposed method was used for determining that 282 bus lines serving Istanbul European side should be assigned to whichever six bus garages and trying to minimize the total distances from the bus garages to the starting points of the lines. For comparison, the problem was solved with k-NN algorithm and the proposed algorithm with different crossover rates.

The crossover operator is considered to be the engine of the genetic algorithm. In other words, the crossover process is the changing of two parts between two parents' chromosomes. The chromosome pairs are selected for crossing with the probability of P. Higher selection of the crossover probability leads to further degradation of chromosomes of the previous generation. This deterioration results in the disappearance of chromosomes with very good fitness values and the disappearance of good solutions. For this reason, the results were searched with both high and low crossover rates in the genetic algorithm step of this study and the better results were obtained with low crossover rate.

Finally, it is seen that the obtained results from two-step meta-heuristic algorithm with the low crossover rate were better than the nearest neighbour algorithm, which proved the efficiency of the study.

In the future, the proposed two-step meta-heuristic algorithm can be applied to a similar assignment problem by adding garage capacity constraints. Thus, a much more realistic work can be done. Furthermore, the proposed algorithm can also be developed for different types of assignment and routing problems. This research can also be expanded with the buses assigned to the garages, as well as in the garage layouts according to the bus sizes and service starting times.

Dr. G. NİLAY YÜCENUR

E-mail: nilayyucenur@beykent.edu.tr

Beykent Üniversitesi, Endüstri Mühendisliği Bölümü

Ayazağa, İstanbul, Türkiye

## BİR ARAÇ ATAMA PROBLEMİ İÇİN İKİ ADIMLI META-SEZGİSEL YAKLAŞIM: İSTANBUL / TÜRKİYE ÖRNEĞİ

### ÖZET

*Bu makalede, bir araç rotalama problemi için geometrik şekil tabanlı kümeleme ve genetik algoritma ile iki aşamalı meta-sezgisel bir yaklaşım önerilmiştir. İlk önce geometrik şekil tabanlı kümeleme yöntemi kullanılmış ve daha sonra bu yöntemin çözümü başlangıç çözüm olarak genetik algoritmaya verilmiştir. Çözüm süreci genetik algoritma ile devam etmiştir. İstanbul'un Avrupa yakasında 282 otobüs hattı bulunmaktadır. Bu otobüsler 6 otobüs garajına atanmalıdır. Önerilen yöntem, garajlara otobüslerin atanmasında otobüs hatları ile garajlar arasındaki minimum mesafeyi belirlemek için kullanılmaktadır. Hesaplanan sonuçlara göre, önerilen algoritma, her bir otobüs hattı başlangıç noktasından her bir garaja olan mesafe açısından daha iyi kümeleme performansına sahiptir. Algoritma performansını artırmak için çaprazlama oranı değiştirme yöntemi de bir deneme olarak uygulanmaktadır. Son olarak, farklı çaprazlama oranları tarafından üretilen çıktılar, çalışmanın etkinliğini kanıtlamak için en yakın komşu algoritmasının sonuçlarıyla karşılaştırılır.*

### ANAHTAR KELİMELELER

*Araç atama problemi; geometrik şekil tabanlı kümeleme; genetik algoritma; çaprazlama oranı; en yakın komşu algoritması;*

### REFERENCES

- [1] Verbas Ö, Mahmassani HS, Hyland MF. Gap-based transit assignment algorithm with vehicle capacity constraints: Simulation-based implementation and large-scale application. *Transportation Research Part B*. 2016;93: 1-16. Available from: doi:10.1016/j.trb.2016.07.002
- [2] Vidal T, Crainic TG, Gendreau M, Prins C. Implicit depot assignments and rotations in vehicle routing heuristics. *European Journal of Operational Research*. 2014;237: 15-28. Available from: doi:10.1016/j.ejor.2013.12.044
- [3] Zhao Y, Leng L, Qian Z, Wang W. A Discrete hybrid invasive weed optimization algorithm for the capacitated vehicle routing problem. *Procedia Computer Science*. 2016;91: 978-987. Available from: doi:10.1016/j.procs.2016.07.127
- [4] Tlili T, Faiz S, Krichen S. A Hybrid metaheuristic for the distance-constrained capacitated vehicle routing problem. *Procedia – Social and Behavioral Sciences*. 2014;109: 779-783. Available from: doi:10.1016/j.sbspro.2013.12.543
- [5] Yassen ET, Ayob M, Nazri MZA, Sabar NR. An adaptive hybrid algorithm for vehicle routing problems with time windows. *Computers & Industrial Engineering*. 2017;113: 382-391. Available from: doi:10.1016/j.cie.2017.09.034
- [6] Damleijer K, Spliet R. A branch-and-cut algorithm for

- the time window assignment vehicle routing problem. *Computers & Operations Research*. 2018;89: 140-152. Available from: doi:10.1016/j.cor.2017.08.015
- [7] Oliveira FB, Enayatifar R, Sadaei HJ, Guimaraes FG, Potvin JY. A cooperative coevolutionary algorithm for the multi-depot vehicle routing problem. *Expert Systems with Applications*. 2016;43: 117-130. Available from: doi:10.1016/j.eswa.2015.08.030
- [8] Du J, Li X, Yu L, Dan R, Zhou J. Multi-depot vehicle routing problem for hazardous materials transportation: A fuzzy bilevel programming. *Information Sciences*. 2017;399: 201-218. Available from: doi:10.1016/j.ins.2017.02.011
- [9] Zachariadis EE, Tarantilis CD, Kiranoudis CT. The vehicle routing problem with simultaneous pick-ups and deliveries and two-dimensional loading constraints. *European Journal of Operational Research*. 2016;251(2): 369-386. Available from: doi:10.1016/j.ejor.2015.11.018
- [10] Kartal Z, Hasgul S, Ernst AT. Single allocation p-hub median location and routing problem with simultaneous pick-up and delivery. *Transportation Research Part E: Logistics and Transportation Review*. 2017;108: 141-159. Available from: doi:10.1016/j.tre.2017.10.004
- [11] Norouzi N, Amalnick MS, Alinaghiyan M. Evaluating of the particle swarm optimization in a periodic vehicle routing problem. *Measurement*. 2015;62: 162-169. Available from: doi:10.1016/j.measurement.2014.10.024
- [12] Shahparvari S, Abbasi B. Robust stochastic vehicle routing and scheduling for bushfire emergency evacuation: An Australian case study. *Transportation Research Part A: Policy and Practice*. 2017;104: 32-49. Available from: doi:10.1016/j.tra.2017.04.036
- [13] Huang Y, Zhao L, Woensel TV, Gross JP. Time-dependent vehicle routing problem with path flexibility. *Transportation Research Part B: Methodological*. 2017;95: 169-195. Available from: doi:10.1016/j.trb.2016.10.013
- [14] Aringhieri R, Bruglieri M, Malucelli F, Nonato M. An asymmetric vehicle routing problem arising in the collection and disposal of special waste. *Electronic Notes in Discrete Mathematics*. 2004;17: 41-47. Available from: doi:10.1016/j.endm.2004.03.011
- [15] Liu R, Jiang Z. The close-open mixed vehicle routing problem. *European Journal of Operational Research*. 2012;220(2): 349-360. Available from: doi:10.1016/j.ejor.2012.01.061
- [16] Yu VF, Jewpanya P, Redi AANP. Open vehicle routing problem with cross-docking. *Computers & Industrial Engineering*. 2016;94: 6-17. Available from: doi:10.1016/j.cie.2016.01.018
- [17] Atefi R, Salari M, Coelho LC, Renaud J. The open vehicle routing problem with decoupling points. *European Journal of Operational Research*. 2018;265(1): 316-327. Available from: doi:10.1016/j.ejor.2017.07.033
- [18] Okulewicz M, Mandziuk J. The impact of particular components of the PSO-based algorithm solving the dynamic vehicle routing problem. *Applied Soft Computing*. 2017;58: 586-604. Available from: doi:10.1016/j.asoc.2017.04.070
- [19] Abdallah AMFM, Essam DL, Sarker RA. On solving periodic re-optimization dynamic vehicle routing problems. *Applied Soft Computing*. 2017;55: 1-12. Available from: doi:10.1016/j.asoc.2017.01.047
- [20] Lin CKY. A vehicle routing problem with pickup and delivery time windows, and coordination of transportable resources. *Computers & Operations Research*. 2011;38(11): 1596-1609. Available from: doi:10.1016/j.cor.2011.01.021
- [21] Afshar-Nadjafi B, Afshar-Nadjafi A. A constructive heuristic for time-dependent multi-depot vehicle routing problem with time-windows and heterogeneous fleet. *Journal of King Saud University – Engineering Sciences*. 2017;29(1): 29-34. Available from: doi:10.1016/j.jksues.2014.04.007
- [22] Zhou, L, Baldacci R, Vigo D, Wang X. A multi-depot two-echelon vehicle routing problem with delivery options arising in the last mile distribution. *European Journal of Operational Research*. 2018;265(2): 765-778. Available from: doi:10.1016/j.ejor.2017.08.011
- [23] Kulkarni RV, Bhavne PR. Integer programming formulations of vehicle routing problems. *European Journal of Operational Research*. 1985;20(1): 58-67. Available from: doi:10.1016/0377-2217(85)90284-X
- [24] Park YB, Koelling CP. A solution of vehicle routing problems in a multiple objective environment. *Engineering Costs and Production Economics*. 1986;10(2): 121-132. Available from: doi:10.1016/0167-188X(86)90006-6
- [25] Radharamanan R, Choi LI. A branch and bound algorithm for the travelling salesman and the transportation routing problems. *Computers & Industrial Engineering*. 1986;11(1-4): 236-240. Available from: doi:10.1016/0360-8352(86)90085-9
- [26] Laporte G, Nobert Y, Taillefer S. A branch-and-bound algorithm for the asymmetrical distance-constrained vehicle routing problem. *Mathematical Modelling*. 1987;9(12): 857-868. Available from: doi:10.1016/0270-0255(87)90004-2
- [27] Brodei GR, Waters CDJ. Integer linear programming formulation for vehicle routing problems. *European Journal of Operational Research*. 1988;34(3): 403-404. Available from: doi:10.1016/0377-2217(88)90162-2
- [28] Anbuudayasankar SP, Ganesh K, Koh SCL, Ducq Y. Modified savings heuristics and genetic algorithm for bi-objective vehicle routing problem with forced backhauls. *Expert Systems with Applications*. 2012;39(3): 2296-2305. Available from: doi:10.1016/j.eswa.2011.08.009
- [29] Yusoff M, Ariffin J, Mohamed A. DPSO based on a min-max approach and clamping strategy for the evacuation vehicle assignment problem. *Neurocomputing*. 2015;148: 30-38. Available from: doi:10.1016/j.neucom.2012.12.083
- [30] Jabir E, Panicker VV, Sridharan R. Design and development of a hybrid ant colony-variable neighbourhood search algorithm for a multi-depot green vehicle routing problem. *Transportation Research Part D: Transport and Environment*. 2017;57: 422-457. Available from: doi:10.1016/j.trd.2017.09.003
- [31] Wang J, Jagannathan AKR, Zuo X, Murray CC. Two-layer simulated annealing and tabu search heuristics for a vehicle routing problem with cross docks and split deliveries. *Computers & Industrial Engineering*. 2017;112: 84-98. Available from: doi:10.1016/j.cie.2017.07.031
- [32] Ng KKH, Lee CKM, Zhang SZ, Wu K, Ho W. A multiple

- colonies artificial bee colony algorithm for a capacitated vehicle routing problem and re-routing strategies under time-dependent traffic congestion. *Computers & Industrial Engineering*. 2017;109: 151-168. Available from: doi:10.1016/j.cie.2017.05.004
- [33] Kowalski RJ, Li C, Ganjyal GM. Optimizing twin-screw food extrusion processing through regression modeling and genetic algorithms. *Journal of Food Engineering*. 2018;234: 50-56. Available from: doi:10.1016/j.jfoodeng.2018.04.004
- [34] Arthur J, Bahrn R, Hutchinson J, Pozzi SA. Genetic algorithm for nuclear data evaluation applied to subcritical neutron multiplication inference benchmark experiments. *Annals of Nuclear Energy*. 2019; 133: 853–862. Available from: doi:10.1016/j.anucene.2019.07.024
- [35] Abreu LR, Cunha JO, Prata BA, Framinan JM. A genetic algorithm for scheduling open shops with sequence-dependent setup times. *Computers & Operations Research*. 2020;113: 104793. Available from: doi:10.1016/j.cor.2019.104793
- [36] Meesrikamolkul W, Niennattrakul V, Ratanamahatana CA. Shape-Based Clustering for Time Series Data. *Advances in Knowledge Discovery and Data Mining: Proceedings of the 16<sup>th</sup> Pacific-Asia Conference, PAKDD, Part I, Kuala Lumpur/Malaysia*; 2012. p. 530-541. Available from: doi:10.1007/978-3-642-30217-6\_44
- [37] Xu S, Zou S, Wang L. A Geometric Clustering Algorithm with Applications to Structural Data. *Journal of Computational Biology*. 2015;22(5): 436-450. Part I, 10.1089/cmb.2014.0162
- [38] Tharwat A, Hahdi H, Elhoseny M, Hassanien AE. Recognizing human activity in mobile crowdsensing environment using optimized  $k$ -NN algorithm. *Expert Systems with Applications*. 2018;107: 32-44. Available from: doi:10.1016/j.eswa.2018.04.017
- [39] Huan Y, Wei T, Wang Z, Lei C, Chen F, Wang X. Polarization switching and rotation in KNN-based lead-free piezoelectric ceramics near the polymorphic phase boundary. *Journal of the European Ceramic Society*. 2019;39(4): 1002-1010. Available from: doi:10.1016/j.jeurceramsoc.2018.11.001
- [40] Bania RK, Halder A. R-Ensembler: A greedy rough set based ensemble attribute selection algorithm with  $k$ NN imputation for classification of medical data. *Computer Methods and Programs in Biomedicine*. 2020;184: 105122. Available from: doi:10.1016/j.cmpb.2019.105122
- [41] Thangiah SR, Salhi S. Genetic clustering: An adaptive heuristic for the multi-depot vehicle routing problem. *Applied Artificial Intelligence*. 2001;15:361-383. Available from: doi:10.1080/08839510151087293