ROUTE FIRST-CLUSTER SECOND METHOD FOR PERSONAL SERVICE ROUTING PROBLEM

MELİKE KÜBRA EKİZ[1*], MUHAMMET BOZDEMİR[2], BURCU ÖZCAN TÜRKKAN[1]

[1]*Departman of Industrial Engineering, University of Kocaeli, Kocaeli, 41001, Turkey*

[2]*Department of Industrial Engineering, Konya Technical University, Konya, 42250, Turkey*

Abstract: The Vehicle Routing Problem (VRP), which has many sub-branches, is a difficult problem that cannot be solved using classical methods. This study includes a case study for Service Routing Problem, which is one of the sub-branches of VRP. The case study is a problem of determining service routes for staffs of a company. In this context, we first assigned the employees to the stations, and then we reached a solution using the route first-cluster second heuristic method. We used the Genetic Algorithm (GA) to improve the route and compared the results by creating different scenarios in clustering methods.

Keywords: vehicle routing problem, clustering methods, service routing problem

1. INTRODUCTION

Logistic term provide that is deliver to customers when a product or service is produced in companies. Nowadays, the term of logistics has become more important because of the development of technology, the increase of competition and the expansion of e-commerce services. With the increase of technological activities, the position of the products at that moment, the times of access to the customer or supplier etc. situations can be monitored more easily. The main reason for the increase in the competition between companies and firms is that the logistics costs constitute a large part of the companies' budget. Therefore, the objectives such as minimizing the duration, distance or cost of the routes used in the transportation of the product to the customer or supplier, and maximizing the occupancy rates of the vehicles used arise. Also e-commerce services are preferred nowadays, therefore increasing the usage rates of e-commerce companies increases the importance of logistics concept day by day.

In this context, the problem of routing, which is a part of the term of logistics, has been discussed. First of all, the problem that was introduced to the literature as a Traveling Salesman Problem (TSP) has as the only tool assumption. The problem of dealing with a large number of vehicles and single warehouses is included in the literature as a Multiple Travelling Salesman Problem (mTSP) [1].

The Vehicle Routing Problem (VRP), which was first discovered by Dantzig and Ranser [2], is the problem of minimizing the travel distance or duration of vehicles with a large number of capacities. For VRP, optimal results were obtained with mathematical models proposed in the literature. With the addition of new constraints to the problem over time, the size of it increased and became a combinatorial and turned into an np-hard. Also, with the addition of these constraints, different VRPs emerged. For example; VRP with time windows, pick-up and delivery, simultaneous pick-up and delivery etc.

---

[*] Corresponding author, email melike.ekiz@kocaeli.edu.tr

Mathematical models for VRP, which contain many constraints in real life problems such as personnel or school service, garbage, medical waste collection, bread and milk distribution, are not sufficient and an optimal or suitable solution cannot be obtained. Therefore, artificial intelligence applications and heuristic – meta-heuristic methods are frequently used in order to obtain suitable solutions for VRP in the np-hard class. In our study, the personnel service problem, which has become an important problem for the companies, has been discussed.

In this study is proceeding as follows. In chapter 2, we mention the VRP and its subgroups which that have different constraints. In chapter 3, we mention literature review of VRP with time windows, service and school service routing problems which are subgroups of VRP related to our problem. In chapter 4, the Genetic Algorithm (GA), which is one of the meta-heuristic methods used in the solution of our problem, is included. In addition, we mention clustering algorithms used to the route obtained in this chapter. In chapter 5, we have implemented the case study and steps. In Chapter 6, we mention the case study and lastly we report the conclusions in chapter 6.

2. VEHICLE ROUTING PROBLEM (VRP)

VRP is the problem of determining the routes required to return to the warehouse by determining the objective function in the form of the minimization of total transportation distance or transportation time, provided that the fleet of vehicles in a central warehouse has a certain homogeneous or heterogeneous capacity. The assumptions of the classical VRP model are as follows:
- Routes start in the warehouse and end there;
- The capacity of the vehicles is homogeneous;
- Demands are deterministic;
- Every demand point must be met once and the entire demand must be met (demands cannot be divided).

The classic VRP model is given below. Here, i and j represent nodes and also 0 and k represents the warehouse and vehicle, respectively. M, N and Q are number of vehicles, number of customers and vehicle capacity, respectively. Also parameters of VRP, c_{ij} is transportation cost from node i to node j and q_i is demand of node i (customer).

Decision variables:

$$x_{ijk} = \begin{cases} 1 & \text{if vehicle of k visits from node i to node} \\ 0 & \text{otherwise} \end{cases}$$

$y_i =$ sub-tour qualifying decision variable

$$\text{Min } Z = \sum_{i=0}^{N} \sum_{j=0}^{N} \left(c_{ij} \sum_{k=1}^{M} x_{ijk} \right) \tag{1}$$

$$\sum_{i=0}^{N} \sum_{k=1}^{M} x_{ijk} = 1, \qquad j=1,...,N \tag{2}$$

$$\sum_{i=0}^{N} x_{ipk} - \sum_{j=0}^{N} x_{pjk} = 0, \qquad k=1,...,M; p=0,...,N \tag{3}$$

$$\sum_{i=1}^{N} \left(q_i \sum_{j=0}^{N} x_{ijk} \right) \leq Q, \qquad k=1,...,M \tag{4}$$

$$\sum_{j=1}^{N} x_{0jk} = 1, \qquad k=1,...,M \tag{5}$$

$$y_i - y_j + N \sum_{k=1}^{M} x_{ijk} \leq N-1, \qquad i \neq j=1,...,N \tag{6}$$

$$x_{ijk} \in \{0,1\}, \qquad i=j=1,...,N; k=1,...,M \tag{7}$$

$$y_i \text{ free} \tag{8}$$

where the objective function defined in equation (1) is to minimize the cost. equation (2) assures that each node *i* is assigned to at least node *j* with any *k*. equation (3) ensured that a tour between the nodes. Equation (4) is the capacity constraint and equation (5) allows each vehicle to leave the warehouse once. Equation (6) provides sub-round qualifying and finally, we give integrality constraints which define the nature of the decision variables.

Different varieties are formed and divided into sub-groups by adding different constraints to VRP over time. Some of these subgroups are listed below:
- VRP with mixed vehicle capacity;
- VRP with multiple warehouse;
- VRP with indeterminate demand;
- VRP with pick-up and delivery;
- VRP with simultaneous pick-up and delivery;
- Time-dependent VRP;
- School bus routing problem;
- VRP with time window.

3. LITERATURE REVIEW

In this section of the study, we discussed the study of VRP with capacity constrained, time window and school service routing problem, which are the subgroups of VRP models which are the closest to the problem we discussed. In the personnel transportation problems, VRP with capacity restricted model can be preferred due to the fact that the vehicle has a certain capacity [3]. In addition, due to the time window in the personnel transportation problem, the VRP with time windows has also been utilized. The VRP with time windows is divided into two subgroups: VRP with hard and flexible time windows. In the case of the VRP with a hard time window, vehicles arriving at the node before the earliest start-up time must remain on that node until the specified time window, and vehicles that arrive after the latest start time cannot start the service. In the VRP with flexible time window, there is a penalty cost for vehicles that cannot reach the service within the specified time window.

Koç and Karaoğlan [4] developed a mathematical model for multi-use and time-window VRP and applied to existing test problems in the literature. Unlike VRP, the VRP with time window includes the concepts of early start time (a_i) and late start time (b_i) of vehicles. The node with the [a_i, b_i] time window should be serviced in this range. The difference between VRP and multi-use VRP is that a vehicle can take a new route after completing the route. Ünlü, Uçar, Akkuş and Şen [5] conducted a multi-vehicle dynamic routing application with time window for cargo companies.

Atmaca [6] carried out by using the data of a cargo company, the VRP model was simultaneously studied. The results were compared with the current status of the cargo company in terms of the number of vehicles, vehicle occupancy rate, and route lengths. In the study conducted by Hezer and Kara [7], VRP with simultaneous pick-up and delivery was applied to sample test problems by using Bacterial Foraging Optimization Algorithm. The results of the proposed model are compared with insertion based heuristic algorithm. Çetin and Gencer [8] proposed a mathematical model by combining the VRP model with the exact time window and the simultaneous pick-up and delivery. The objective function of many VRPs in the literature has been taken as the minimization of the received path or cost while the objective function is taken as the minimization of the delays. Cetin and Gencer [9] proposed a VRP model which is taking into account the situations where vehicle capacities are not equal or the vehicle types are not the same and with simultaneous pick-up and delivery with exact time windows. Koç and Karaoğlan [10] suggested a mathematical model for the problem, which is called time-dependent VRP, when the speed of the vehicles varied between the nodes.

Bozyer, Alkan, and Fığlalı [11] mentioned that mathematical models would be insufficient if the size of the VRP problem increased and they proposed a heuristic method which is based on before clustering after then routing. Şahin and Eroğlu [12] have addressed the problem of limited capacity VRP and have studied with the heuristic and meta-heuristic methods that are available in the literature. Düzakın and Demircioğlu [13] examined the models and heuristic methods developed by VRP in their studies. Güvez, Dege and Eren [14] determined the positions of health institutions on the digital map and calculated their distance to each other and worked on the collection of medical wastes with the lowest cost using the integer programming model. Onder [15] handled the multi-warehouse VRP and implemented an application on bread distribution.

For school service routing problems, nodes (stops) are usually determined for students using a clustering algorithm. In school service problems, the number of schools and accordingly the ringing time is important. If the vehicles serve more than one school, the ring time of each school is important. Gündoğar and Akıl [16] examined routing problems and considered the problem of routing of service vehicles as a scheduling problem. Uzun and Tezel [17] used the heuristic method for each individual to benefit from the service in his / her time window in the use of rehabilitation service of disabled individuals. Uzumer and Eren [18] has optimized the school service routing problem with the linear program model. Ünsal and Yiğit [19] mentioned that both the fuel and the cost will be contributed by the optimal route of service vehicles, and the time spent in the service and the walking distance will be reduced by the personnel who use the vehicles. In order to benefit from the developing technology and to facilitate the routing problem, data were obtained from GSP, GIS and firstly nodes were determined by using K-means clustering algorithm. In determining the most suitable route for service vehicles, GA which is one of the artificial intelligence techniques is used.

4. DESCRIPTION OF THE PROBLEM AND METHODS

In our study, we determined a problem which is a own service vehicles to be able to provide the most appropriate service to their staff in the morning and in the evening. In this context, the purpose of the company is to minimize the factors such as route and cost. It also wants the vehicles to reach the company in time and to minimize the time spent by the personnel in the service. The request of the company in the routing problem is the smallest walking distance to the designated nodes (stops) by clustering.

In our study, we used route first—cluster second method for the service routing problem. In this context, the problem was first thought to be TSP and the first route was created with the nearest neighbor algorithm. The first route using the GA was improved and the nodes were assigned to the vehicles by controlling their capacity using 3 different clustering methods. Below, we will first talk about the GA used to improve TSP and then 3 different clustering methods.

4.1. Genetic Algorithm (GA)
Individuals in nature struggle with each other to have resources such as food, shelter and water. The lives of the victors in the struggles are longer and their genes are transferred to the next generations. Thus, weak individuals are eliminated. The good characteristics from different individuals will merge over time and better individuals will emerge. The GA based on this basic logic is a heuristic method that does not guarantee us the best solution, but in a short time it gives good solution which is near to best solution. Steps of GA are as follows:
-A random group of all possible solutions in the search space is selected and coded as a series of solutions. Then, a random process is performed to create the starting population.
-The fit value is calculated for each set of solutions. The calculated fit value indicates the quality of the solution.
-A group is selected by means of a selection mechanism. Selected group are subjected to crossing and mutation processes and new generations are created.
-The above procedures are continued until the stop criterion is established. The most appropriate individual is chosen as the solution.

In the following part of the study, the tournament selection method and linear crossing and placement mutation method are mentioned which we used in the case study.

Tournament Selection Method: Two sequences are randomly selected from the whole generation. As a result of the struggle between these two series, whoever is better wins and gets a point. All individuals in the generation have the right to participate in the tournament twice. Those who succeed in both tournaments will print their name to the list.

Linear Crossing: Firstly, a random gene site is determined, and the genes of the mother and father are assigned to two children according to the location. Remaining genes in children are completed in a way not to repeat by mother and father.

Placement Mutation: A randomly determined gene from the population is placed in another random position within the same population.

After multiplication, crossing and mutation processes, an infinite number of generations are formed. The stop criterion is determined to stop an infinite number of these generations. The number of iterations is generally used as a stop criterion.

4.2. Clustering Methods and Distance Functions
Clustering methods divide assets into different groups according to their similarities. It is desired that these groups have the most similarity among themselves, while the similarity between the other groups is desired to have the least similarity [20]. K-Means, K-Medoids and K-Modes clustering algorithm were used in the study. The important issue in clustering methods is the distance functions. In this context, Euclidean, Manhattan and Square Euclidean distance functions are included in this study.

4.2.1. K-Means Clustering Algorithm
The K-Means clustering algorithm performs grouping by the number of sets determined by looking at the similarity of each asset in the data set. The stages of K-Means clustering algorithm are as follows [21]:
-Determining the number of clusters (determining the number of k) and separating the data into k sets;
-Assignment of the assets in the cluster to the closest cluster in terms of value (the value is calculated with the distance function);
-Step 2 is repeated until no asset is out.

4.2.2. K-Medoids Clustering Algorithm
The K-Medoids clustering algorithm, which has many variants in the literature, has been developed due to the fact that K-Means clustering algorithm. K-Medoids is better than K-Means clustering algorithm in terms of non sensitive to outliers, execution time and reduces noise [22]. The stages of K-Medoids clustering algorithm are as follows [23]:
-Random k elements are selected and these elements are determined as cluster center;
-New entities are added to the cluster and those that contribute the most to the cluster development are selected and are assigned as a central;
-Step 2 is repeated until no asset is out.

4.2.3. K-Modes Clustering Algorithm
The stages of the K-Modes clustering algorithm, a variant of the K-Means clustering algorithm, are as follows:
-Random k beginning mode is selected;
-Objects in the cluster are displaced if the mode of assets is closest to which cluster center;
-Cluster modes are updated after each displacement;
-Steps 2 and 3 are repeated until the displacement of the objects in the cluster is finished.

5. CASE STUDY

In order to solve the problem of personnel service, the method was applied to route first-cluster second. In order to improve the route, GA was chosen from artificial intelligence algorithms. We have developed a software for improving the route for GA. First of all, 20 initial solutions were created with the nearest addition heuristic. Tournament Selection Method was applied in the selection of the initial solutions. Generally, those who have 1 point in the tournament selection method should be eliminated. However, because of the diversity in the cluster, we have included even get 1 point fields in the generation set. We preferred the linear crossing method to multiply the individuals in the generation set.

A total of 2000 crossings were performed with 1000 iterations. The mutation operator was used to ensure genetic diversity. A random decimal number of 0 to 1 was generated for each mutation operator to apply the mutation operator. If the random decimal number is between 0 and 0.1, a mutation is performed. We checked the fit values within the generation clusters that were formed as a result of the selection, crossing and mutation process. The population cluster with the lowest fit value is considered a solution space. After determining the solution space, different scenarios were created according to the strategies formed according to clustering methods and distance functions. We have developed a software for applying the method developed for the solution of the personnel service problem to the existing data. This software uses Microsoft Access 2016 database program, C # programming language on Visual Studio.Net 2015 platform and GMap.NET library.

We created the total number of people who will use the personnel service is 93. The stops were determined according to the proximity status of the regions where these 93 people live. We have established a route with 39 passenger stops and a destination company so a total of 40 stops. There are two different types of vehicles with capacities 27 and 16. The results obtained with the scenarios created by setting the number of clusters to $k= 4$ are given in Table 1.

Table 1. Results of scenarios created with clustering methods and distance functions.

Clustering Method	Distance Function	Result (km)
K-means	Euclidean	178.6593
K-Medoids	Euclidean	178.6593
K-Modes	Euclidean	178.6593
K-means	Manhattan	227.3776
K-Medoids	Manhattan	227.3776
K-Modes	Manhattan	178.6593
K-means	Square Euclidean	178.6593
K-Medoids	Square Euclidean	178.6593
K-Modes	Square Euclidean	178.6593

As a result of the scenarios, there were 2 different results, the best of which was 178.6593 km. The Euclidean and Square Euclidean distance function produced good results, but the Manhattan distance function did not produce good results in K-means and K-medoids methods. According to the best results, it has been assigned 3 service units with a capacity of 27 and 1 service units with a capacity of 16 people. As one of the best solutions, route lengths, number of staff and occupancy rate of services obtained by K-medoids clustering method and Euclidean distance function are given in Table 2. The average occupancy rate of services is 96.3%. This ratio can be said to be quite good.

Table 2. Results of K-medoids clustering method and Euclidean distance function.

Cluster	Route Length	Number of Staff	Occupancy Rate of Service
0	42.6223	26	96.3%
1	64.176	27	100%
2	15.6218	24	88.88%
3	56.2392	16	100%

6. CONCLUSIONS

In this study, first of all, it was followed route first-cluster second heuristic method for service routing problem. We created scenarios using 3 different clustering algorithms; K-means, K-medoids and K-modes and 3 different distance functions; Euclidean, Manhattan, Square Euclidean. According to these scenarios, the route length is determined by the developed software and the total route length is calculated. Different scenarios produce the best distance according to the clustering method and distance function.

 The best result has 178.6593 km total route lengths and 96.3% average occupancy rate of services. According to the best results, it has been assigned 3 service units with a capacity of 27 and 1 service units with a capacity of 16 people. We chose as one of the best solution which formed K-medoids clustering algorithm and Euclidean distance function. In this solution, 4 clusters were created and a service was assigned to each cluster. The occupancy rates of the services are 96.3, 100, 88.88, 100% respectively.

As a result of this study, it is not possible to comment that a specific clustering method and distance function give the best results. However, in general, the best results in clustering methods were obtained with the Square Euclidean and Euclidean distance functions.

REFERENCES

[1] Rostami, A.S., Mohanna, F, Keshavarz, H, Hosseinabadi, A.A.R., Solving multiple traveling salesman problem using the gravitational emulation local search algorithm, Applied Mathematics & Information Sciences, vol. 9, no. 3, 2015, p. 1-11.

[2] Dantzig, G.B., Ramser, J.H., The truck dispatching problem, Management Science, vol. 6, no. 1, 1959, p. 80-91.

[3] Dinçerler, V.A., Güven, N.E., Tanrikulu, M.M, Temel, M., Yitmen, M., Yaman, H., Bilkent Üniversitesi personel taşima sistemi için etkin ve ekonomik çözüm, Endüstri Mühendisliği Dergisi, vol. 15, no. 2, 2004, p. 2-14.

[4] Koç, K., Karaoğlan, İ., Çok kullanımlı ve zaman pencereli araç rotalama problemi için bir matematiksel model, Journal of the Faculty of Engineering and Arcchitecture of Gazi University, vol. 27, no. 3, 2012, p. 69-576.

[5] Ünlü, N., Uçar, E., Akkuş, G.B., Şen, B., Multi-vehicle dynamic routing with time- windows in parcel delivery, vol. 3, no. 2, 2017, p. 105-113.

[6] Atmaca, E., Bir kargo şirketinde araç rotalama problemi ve uygulaması, Tübav Bilim Dergisi, vol. 5, no. 2, 2012, p. 12-27.

[7] Hezer, S., Kara, Y., Eşzamanlı dağıtımlı ve toplamalı araç rotalama problemlerinin çözümü için bakteriyel besin arama optimizasyonu tabanlı bir algoritma, Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 28, no. 2, 2013, p. 373-382.

[8] Çetin, S., Gencer, C., Kesin zaman pencereli - eş zamanlı dağıtım toplamalı araç rotalama problemi: matematiksel model, Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 25, no. 3, 2010, p. 579-585.

[9] Çetin, S., Gencer, C., Heterojen araç filolu zaman pencereli eş zamanlı dağıtım-toplamalı araç rotalama problemleri: matematiksel model, International Journal of Research and Development, vol. 3, no. 1, 2011, p. 19-26.

[10] Koç, K., Karaoğlan, İ., Zaman bağımlı araç rotalama problemi için bir matematiksel model, Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 29, no. 3, 2014, p. 549-558.

[11] Bozyer, Z., Alkan, A., Fığlalı, A., Kapasite kısıtlı araç rotalama probleminin çözümü için önce grupla sonra rotala merkezli sezgisel algoritma önerisi, Bilişim Teknolojileri Dergisi, vol. 7, no. 2, 2014, p. 29-37.

[12] Şahin, Y., Eroğlu, A., Metaheuristic methods for capacitated vehicle routing problem: literature review, Suleyman Demirel University, The Journal of Faculty of Economics and Administrative Sciences, vol. 19, no. 4, 2014, p. 337-355.

[13] Düzakin, E., Demircioğlu, M., Vehicle routing problems and solution methods, Çukurova Üniversitesi İİBF Dergisi, vol. 13, no. 1, 2009, p. 68-87.

[14] Güvez, H., Dege, M., Eren, T., Kırıkkale'de araç rotalama problemi ile tıbbi atıkların toplanması, International Journal of Engineering Research and Development, vol. 4, no. 1, 2012, p. 41-45.

[15] Önder, E., Optimization of multi-depot vehicle routing problem of Istanbul Halk Ekmek A,Ş. (IHE) by using meta-heuristic methods, Business Economy Institute Journal of Management, vol. 70, no. 1, 2011, p. 74-92.

[16] Gündogar, E., Akil, S., Servis araçları rotalama- çizelgeleme problemleri ve çözüm yaklaşımları, Sakarya University Journal of Science, 1998, p. 25-30.

[17] https://www.researchgate.net/publication/312320724 (19.05.2019).

[18] Uzumer, E., Eren, T., Okul servisi rotalama problemi: bir uygulama, International Journal of Engineering Research and Development, vol. 4, no. 2, 2012, p. 26-29.

[19] Ünsal, Ö., Yiğit, T., Yapay zeka ve kümeleme teknikleri kullanılarak geliştirilen yöntem ile okul servisi rotalama probleminin optimizasyonu, Journal of Engineering Sciences and Design, vol. 6, no. 1, 2018, p. 7-20.

[20] Ercan, C.S., Yazgan, H.R., Sertvuran, İ., Şengül, H., Sıkı zaman pencereli araç rotalama probleminin çözümü için yeni bir yöntem önerisi ve bir süpermarket zincirinde uygulanması, Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, vol. 22, no. 2, 2018, p. 685 – 694.

[21] Çakmak, Z., Uzgören, N., Keçek, G., Kümeleme analizi teknikleri ile illerin kültürel yapılarına göre sınıflandırılması ve değişimlerin incelenmesi, Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, vol. 1, no. 12, 2005, p. 1-21.

[22] Arora, P., Deepali, D., Varshney, S., Analysis of k-means and k-medoids algorithm for big data, Procedia Computer Science, vol. 78, 2016, p. 507-512.

[23] http://eacharya.inflibnet.ac.in/data-server/eacharya-documents/53e0c6cbe413016f23443704_INFIEP_33/93/LM/33-93-LM-V1-S1__kmedoids.pdf (25.06.2019).