Vehicle Routing Problem with Time Window Constraint using KMeans Clustering to Obtain the Closest Customer

Nushwan Yousif B. Al-Nakash

Received: 3 December 2021 Accepted: 31 December 2021 Published: 14 January 2022

Abstract
In this paper, the problem statement is solving the Vehicle Routing Problem with Time Window constraint using the Ant Colony Algorithm with K-Means Clustering. In this problem, the vehicles must start at a common depot, pickup from various warehouses, deliver to the respective nodes within the time window defined by the customer and return back to the depot. The objectives defined are to reduce number of vehicles employed, the total logistics cost and to reduce carbon emissions. The mathematical model described in this paper has considered multiple pickup and multiple delivery points. The proposed solution of this paper aims to provide better and more efficient solution while minimizing areas of conflict so as to provide the best output on a large scale.

Index terms—ant colony optimization (ACO), k-means clustering, vehicle routing problem (VRP), time dependent vehicle routing problem (TDVRP).

1 I. Introduction
Transportation is one of the primary requisites of civilization and this fact continues to be true even today. In today’s world of quick and safe deliveries, there has been a need for better service, reduction of vehicles used, maximizing profit, reduction in travel time variation and reduction of overall travel cost. To define these problems together, the term Vehicle Routing Problems was coined. This problem deals with the supply chain of an organization. Transportation is the backbone of the logistics of any organization and it takes up about 40 to 50% of the total logistics cost, as stated in https://www.cogoport.com/blogs/transportcost (accessed on 11 October, 2021). This includes international and domestic transport, customs, all modes of transport such as air, water, land and so on. It can be inferred that transportation cost is a major and important factor in the supply chain of an organization, so its cost optimization becomes a necessity. The logistics branch of the organization must work on the management of transportation, deliver within customer provided time frames, competing with other organizations for better service and service rates effectively, handling unpredictable events and so on.

The world is witnessing the digital growth spurt and along with its influence on almost every sphere of life and nature. Integration of logistics and e-business will be a fruitful endeavor. This incorporation will lead to improvement in customer service, tracking, deliverance, time effectiveness as well as reduction in the overall cost.

Looking at the technical aspects of the Vehicle Routing Problem (VRP), there are initially p vehicles located at a depot that must deliver different amounts of supplies to q customers. Now, the VRP will aim to find the optimal route that a group of vehicles serve a group of users. This way a standard solution is obtained which contains all the routes that start and end at the depot, with the constraint that the goods are delivered within or before the time range set by the customer, capacity limit and the working time of the drivers are also considered.

This paper will discuss how the Ant Colony Optimization with KMeans Clustering (ACO-KMeans) has been employed to minimize costs when delivering goods from depot to the customer within or before the time frame constraint set. The mathematical model defined in this paper will tackle and solve the problems related to distribution, e-logistics, retail networks and so on.

Dantzig and Ramser [1] were the first ones to introduce the Vehicle Routing Problem in 1959. Their solution was based on Linear programming. It was a truck dispatching problem that dealt with the delivery of gasoline at gas stations. Later, [2] Clarke and Wright came up with the savings method and it was termed as the
II. Literature Survey

One of the heuristic solutions mentioned was provided by Hideki Hashimoto, Mutsunori Yagiura and Toshihide Ibaraki [8]. In their paper they generalized VRPTW by making travelling costs and duration to be time-dependent functions. They used local search algorithm to find the routes of vehicles and using that, evaluated a neighborhood solution. They proposed an algorithm that could efficiently pick optimal routes using data from previous dynamic programming recursion that were used to evaluate the present solution. They even included a filtering method that determines which spaces in the neighborhood are not to be searched so as to avoid dead ends in improving the solution. They finally conclude with a local search algorithm that combines all their modifications.

A metaheuristic solution was proposed by YiyoKuo [6]. In the research paper, the author has considered fuel consumption and carbon emission as the constraints to the Time-Dependent Vehicle Routing Problem (TDVRP). The paper has proposed an algorithm that determines a route that consumes less fuel and has the least carbon emissions. With this algorithm the author was able to provide an overall improvement of 22.69% in minimizing transportation distances and 24.61% improvement in fuel consumption.

Another paper which has VRP with an added constraint where customers can request for delivery or pickup with the requirement that in every single delivery route, all pickups and deliveries to the customers are completed. This is known as Vehicle Routing Problem with the Time Window Constraint [12] that has been set by customers. Another modified VRP with the added constraint of using limited number of vehicles of varying holding capacity has been published as Mixed Fleet Vehicle Routing (MFRV) [13].

This paper has five sections in total. Section 1 deals with the introduction while section 2 deals with the literature survey. Section 3 handles the mathematical model of the proposed system [ACO using KMeans Clustering Algorithm], section 4 will explain the approach to the solution, section 5 will have the results and case studies, with section 6 concluding the paper.

2 Figure 1: General VRP solving method

Many works of solving the VRP with the Time Window Constraint were inherited from the travelling salesman problem. The method used by the salesman to find the best and optimal route to deliver the goods to the respective customers from one or more depot and also take the goods from the customer back to the respective depots within the constraints set, has been extensively used in VRP, with the inclusion of extra constraints. Similar VRP variants have been mentioned below:

- Vehicle Routing Problem with the Time Window Constraint [12] that has been set by customers.
- Another modified VRP with the added constraint of using limited number of vehicles of varying holding capacity has been published as Mixed Fleet Vehicle Routing (MVFR) [13].

Another paper which has VRP with an added constraint where customers can request for delivery or pickup with the requirement that in every single delivery route, all pickups and deliveries to the customers are completed. This is known as Vehicle Routing Problem with Backhauls (VRPB) [13].

Figure ?? gives a generalized view of how a VRP is solved.
VRP with soft time windows. According to this paper, the classical ant colony algorithm has been modified to
efficiently solve the local optimum problem. Their research has given proof that it can achieve a lower routing
cost at a high convergence rate than the classical ant colony (ACO) and the stimulated annealing ant colony
algorithms.

Looking into other heuristic strategies involved, [17] has the space-filling curve with optimal partitioning
as a solution while another has three-phase heuristics which has been developed by grouping a heuristic-based
clustering algorithm solving VRP [18]. Summary of other important state-of-art modern heuristics is available
in [19,20].

In this paper, we will be solving the Vehicle Routing Problem with Time Windows constraint using the
modified Ant Colony Optimization with KMeans Clustering. Ants use pheromones to leave behind a trail for
its comrades so as to use the optimal path fixed to reach the food source. There has been several researches
based on this behaviour of ants, such as [21], which was the first paper to be published on this topic. Papers
[22,23,24,25,26,27] have various hybrid versions of ACO in varied fields.

Using this behaviour of ants and with the help of previous research work based on a somewhat similar problem,
this paper aims to solve VRPTW using the KMeans Clustering algorithm to find the most optimal path to the
customer.

4 Mathematical Model of Proposed System

This part will use certain terms and elements from [28]. It is a case study based on VRPTW regarding fresh
distribution centres. There will be two subsets of service nodes: pickup set ?? and delivery set ?? ->.
The values of these terms are ?? = ?? and ?? = ?? respectively. Now, starting depot node is set to
0 and end depot is set to (?? + ?? + 1). A node will be replicated if it needs both delivery and pickup. Each
vehicle has its set capacity and operation cost. If there is an order between pickup node ?? and delivery node ??
then there will be a set ?? which contains pairs of (??, ??).

Looking at the objective function that minimizes total travelling cost, the equation is as follows
?? ?? ?? ?? ?? ??? = ? ??

Here, ?? refers to the number of clusters and ?? refers to the centroid of clusters.

The next equation makes sure that each node is served by at least one vehicle.

Equation (3) showcases the constraint where the same vehicle ?? must pick and order from node ?? and
deliver it to node ??.

Equation (4) is used to ensure that there is only one vehicle ?? that visits node ??.

A vehicle must pass starting and ending depots at least once and this is shown by equations ( ??) and (5).

If a vehicle reaches a node, it must leave as well. This is shown in equation [6].

Equations (7) and (8) have integrated time constraints, subtour elimination and load constraints.


5 Approach to the Solution

In this paper, the Vehicle Routing Problem with Time Window constraint has been resolved using a modified
version of the Ant Colony Optimization using KMeans Clustering. Marco Dorigo was the first person to introduce
Ant Colony Optimization, in the 90s, in his Ph.D. thesis. The solution algorithm is based on the behaviour of
ants, the way they live in colonies and search for food. While an ant goes around, searching for food, it leaves
behind pheromones that act as a beacon. It acts as a communication mechanism and each time the ant leaves a
pheromone trail, it tells the other ants about the quality and quantity of food the former ant had been carrying.
This way, there are several set paths that the ants use based on the number of pheromones released in a path.
The shortest and fastest route is chosen for maximum traffic.

Ant Colony Optimization (ACO) algorithm is a probabilistic technique based on the above phenomenon to find
the optimal path. With the inclusion of KMeans Clustering, this modified approach has solved the constraints
of the MPMDVRPTWIF, which has resulted in shorter time consumption, delivery within the time window
and lower transportation cost along with the inclusion of multiple pickup and delivery nodes wherein a pickup
point might or might not have multiple delivery locations. The flowchart below showcases the solution setup. In
the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, each arc $(?, ?)$ has been assigned a variable called pheromone trail $??$. The probabilities of better solution is directly proportional to the pheromone intensity. This means that when an ant wants to go to another node from its current node, it will choose the one with the maximum pheromone intensity. To make this work, a fixed quantity of pheromone is allocated to every arc. To decide which node to proceed to, the form that will be used for finding the best solution from the set. When determining the optimal route for the best solution, pheromone update is used which includes pheromone deposition and pheromone evaporation. \+ ?? \+ 0 and ?? is a constant. In each iteration, ?? number of ants find??? ??? = ? ?? ??=1 ?? ??.

7 ??
of average total distance. Then pheromone is updated by the elitist and best ants. After the evaporation process, only the best and the elitist ants can update the pheromone deposits on the optimal path chosen, which is given by the equation?? ??? = ?? ??? + ? ?? ?? * + ?? ?? ?1 ??=1 ?? ?? ?? (16)

Where,?? ??? ?? = \{ ?? ?? ?? ?? ?? ?? ?? ?? ?? ??? ?? ?? ?? ??? ?? ?? ?? ?? ?? ?? ???? ?? ?? ??

Looking at equations 17 and 18, it can be concluded that there are two types of pheromone deposits that are deposited on the trails during the pheromone update process. First, if ?? elitist ants have travelled a path, that path will be updated as the best solution so far (????), in accordance with the ACO+KMeans Clustering algorithm. ??? ??? * denotes the pheromone update by the elitist ants. Second, out of the ?? ants available, only (?? ? 1) best ants, in the current iteration, can deposit pheromone on the path they have traversed. The term ??? ??? * is used to denote the pheromone quantity laid down on the trails that have been traversed by them and the amount of pheromone that have been deposited by the ants are determined by their solution quality ?? ?? and rank ?? and the value is equal to ???? ??? ?? . To summarize, the elitist ants need to increase the probability of the best-solution so far after each iteration as the values that are updated will act as reference values for the next iteration. The ranking methodology has been employed in \[17\] so as to reduce pheromone deposition on those routes that have relatively lesser favourability.

8 Case Study a) Dataset Used
The dataset has been obtained from consulting firms Horizon Consulting Inc. and EATEAM Inc. and their clients from students CPT and it has been preprocessed to Solomon-100 standard test set which have 20 problem cases. The pre-processed real world dataset from above firms includes x-y location coordinates, service time, demand by customers and ready time. This section will be in comparison with \[30\] as it has used the same data set. This comparison will help in proving that the proposed solution from this paper is the better method of solving the (MPMDVRPTWHF) as it gives better cost reduction with lesser percentage of carbon emissions, along with optimized fuel consumptions and lesser vehicles used.

9 b) Parameters Defined
The parameters defined in this paper are derived from \[30\] as this paper is in comparison with the latter. Similar to \[30\] the delivery vehicle used is a refrigerator car and the set of pre-defined parameters are given below.

10 Parameters
11 c) Result Analysis
The entire result section has used the Pareto optimal principle for obtaining the solution. The Pareto Principle states that 80 percent of a project’s benefit comes from 20 percent of the work. The optimal version of it makes
the sub objectives suppressed so as to efficiently solve the main objective. Due to this there is very little scope of conflict of objectives from the sub objectives and a noiseless solution id obtained.

Referring to [30], this paper the objectives chosen will be carbon emission reduction, total cost, time frame and customer satisfaction.

Using several test cases of 25, 50, 75 and 100 customers in three different scenarios, the proposed ACO algorithm with KMeans clustering provides a better solution in comparison. The results are arranged in the Pareto optimal solution format. [??] test table is used here for obtaining the most optimal path with better results of the constraints set. This solution has used the Pareto optimal approach and figure 3 has shown the comparison between [3] and this paper. It is clearly visible from the graph that the proposed algorithm of ACO+KMeans [PS_KP S O] clustering has better output in terms of carbon emission, customer satisfaction and total transportation cost. This part has used the c101(50) test table. 5 trucks have been employed with respective paths (0, 43, 42, 41, 40, 44, 46, 45, 48, 50, 49, 47, 0), (0, 5, 3, 7, 8, 10, 11, 9, 6, 4, 2, 1, 75, 0), (0, 20, 24, 25, 27, 29, 30, 28, 26, 23, 22, 21, 0), (0, 32, 33, 31, 35, 37, 38, 39, 36, 34, 0) and (0, 13, 17, 18, 19, 15, 16, 14, 12, ??). The end are 33.43 for carbon emissions, 5942.72 cost and 100 percent customer satisfaction. Figure 5 shows the comparison between [30] and this paper results while figure 6 displays the routes taken by the 5 trucks. The c101(75) dataset has been used in this part. The number of vehicles used is 8 with the most optimal paths chosen respectively: (0, 43, 42, 41, 40, 44, 46, 45, 48, 51, 50, 52, 49, 47, 0), (0, 5, 3, 7, 8, 10, 11, 9, 6, 4, 2, 1, 75, 0), (0, 32, 33, 31, 35, 37, 38, 39, 36, 34, 0), (0, 67, 65, 63, 62, 74, 72, 61, 64, 68, 66, 69, 0), (0, 20, 24, 25, 27, 29, 30, 28, 26, 23, 22, 21, ??), (0, 57, 55, 54, 53, 56, 58, 60, 59, 0), (0, 13, 17, 18, 19, 15, 16, 14, 12, ??) and (0, 71, 70, 73, 0).

The final results of carbon emissions, total cost and customer satisfaction are 54.96, 10639.71 and 100 percent respectively. Figures 7 and 8 showcase the comparison between [30] and this paper and the route distribution of the vehicles.

12 Global Journal of Computer Science and Technology

Volume XXII Issue I Version I This section has used the c101 (100) dataset. Now looking [30], there are better results in terms of carbon emission, cost and customer satisfaction (69.03, 13561.41 and 100 percent). Instead of 23 vehicles, 10 vehicles have been employed and the most optimal paths are chosen: (0, 5, 3, 7, 8, 10, 11, 9, 6, 4, 2, 1, 75, 0), (0, 43, 42, 41, 40, 44, 46, 45, 48, 51, 50, 52, 49, 47, 0), (0, 20, 24, 25, 27, 29, 30, 28, 26, 23, 22, 21, 0), (0, 90, 87, 86, 83, 82, 84, 85, 88, 89, 91, 0), (0, 57, 55, 54, 53, 56, 58, 60, 59, 0), (0, 13, 17, 18, 19, 15, 16, 14, 12, ??) and (0, 71, 70, 73, 0).

13 Global Journal of Computer Science and Technology

Volume XXII Issue I Version I In the Solomon-100 dataset, there are three formats of destination grouping. One is a cluster format (C), one is a random format (R) and one is a randomclustered format (RC). These three formats have been used for 25, 50, 75 and 100 customers. So other than C101, there are C201, R211, R201 and RC201. The comparison between the proposed algorithm (ACO+KMeans algorithm) and modified Ant Colony algorithm [30] have been given in Table 4. The data from Table 4 helps in evaluating the effectiveness of the proposed algorithm. Even with increase in the number of customers, be it clustered, random or both, there is barely any increase in the number of vehicles employed. With an average of 2.625 vehicles per case, this greatly affects the total travel, storage, damage and fuel costs while reducing the carbon footprint by a great extent, ultimately helping not only the economy of the organisation but also trying to improve the environmental condition of the Earth. It can be assumed from the results data that there is a high probability of increase in number of customers. As the number of vehicles employed is less, there is scope of increasing customer reach and maybe there is a chance of increasing the speed of delivery. With the new electronic vehicle usage, there will be even more cuts in the carbon footprint value and better customer coverage.

14 Conclusion

This paper discusses the vehicle routing problem with time window constraint (VRPTW) along with added constraints of number of vehicles, logistics cost, overall carbon emission rate along with multiple pickup and delivery points faced by firms EATEAM and Horizon Consulting Inc. in their logistical operations. A meta heuristic Ant Colony algorithm with KMeans Clustering was employed to solve the problem statement. Looking at the literature survey in this paper, it is observable that Vehicle Routing Problem has had several approaches
with varying results, which in turn leads to the fact that VRP with added constraints is a difficult problem to solve.

The solution provided in this paper has been compared with [30], which has a similar problem statement, and the results of the proposed Ant colony Algorithm with KMeans Clustering has performed far better and has provided very less scope of improvement in the discussed problem areas.

In future researches on similar topics, it’s a hope that this paper will be a good leverage for the researchers and this solution can be further modified for more improvements.

Figure 1: Figure 2:
Figure 2: (}
Figure 3: Figure 3:

Figure 4: Figure 4:
Figure 5: Figure 5:

Figure 6: Figure 6:
Figure 7: Figure 7:

---

Figure 8: Figure 8:
Figure 9:

<table>
<thead>
<tr>
<th></th>
<th>ACO</th>
<th>ACOMO</th>
<th>PS_KPSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>98.58</td>
<td>99.6</td>
<td>100</td>
</tr>
<tr>
<td>CE</td>
<td>135.4</td>
<td>129.87</td>
<td>69.03832969</td>
</tr>
<tr>
<td>C (CNY)</td>
<td>32103.39</td>
<td>31202.27</td>
<td>13561.40714</td>
</tr>
</tbody>
</table>

Figure 10: Figure 9:
Figure 11: Figure 10:

Solution for Case: c201 with 75 Customers

- Center
- Customer Location
Figure 12: Solution for Case: c201 with 100 Customers

Figure 13:
Figure 13: Solution for Case: c201 with 25 Customers

Figure 14: Solution for Case: c201 with 25 Customers
Solution for Case: c201 with 50 Customers

Figure 14: Figure 11:
Figure 15: Solution for Case: r201 with 25 Customers

Figure 17: [Additional figure or content not provided]
Figure 16: Solution for Case: r201 with 50 Customers

Figure 19:
Figure 17: Solution for Case: r201 with 75 Customers

- *Center*
- *Customer Location*
Solution for Case: r211 with 25 Customers

Figure 19: Figure 25:
Pheromone update is used to elevate the pheromone values that are found on good solution paths and decrease those that are on bad solution paths. In pheromone deposition and evaporation, pheromone values either increase or decrease at a constant rate [29]. The pheromone evaporation equation is given as such,

\[
\text{pheromone deposition and evaporation, pheromone} = \text{?}(??, ?) \text{(15)}
\]

Where trail persistence 1 ? ?? ? 0 of the evaporation factor 1 ?

\[
?? ?
\]

Figure 21:

<table>
<thead>
<tr>
<th>PID</th>
<th>NUM_CUST (CNY)</th>
<th>PS_KPSO</th>
<th>ACOMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>C101_25</td>
<td>25</td>
<td>3481.31</td>
<td>98.4</td>
</tr>
<tr>
<td>C101_50</td>
<td>50</td>
<td>9583.7</td>
<td>99.2</td>
</tr>
<tr>
<td>C101_75</td>
<td>75</td>
<td>20195.27</td>
<td>99.47</td>
</tr>
<tr>
<td>C101_100</td>
<td>100</td>
<td>31202.27</td>
<td>99.6</td>
</tr>
</tbody>
</table>

b. Results with other test cases

Figure 22: Table 2:
<table>
<thead>
<tr>
<th>PID</th>
<th>NUM_CUST</th>
<th>TOT_VEH</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_201_25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>C_201_50</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>C_201_75</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>C_201_100</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>r_201_25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>r_201_50</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>r_201_75</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>r_201_100</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>r_211_25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>r_211_50</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>r_211_75</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>r_211_100</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>r_c201_25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>r_c201_50</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>r_c201_75</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>r_c201_100</td>
<td>100</td>
<td>25</td>
</tr>
</tbody>
</table>

Figure 23: Table 4:


[Figliozzi M A] ‘The time dependent vehicle routing problem with time windows: benchmark problems, an efficient solution algorithm, and solution characteristics’. Figliozzi M A. *Transportation*


