International Journal of Applied Science and Engineering

Optimization of capacitated vehicle routing problem using initial route with same size K-means and greedy algorithm for vaccines distribution

Jessi Febria[1], Christine Dewi[1,2], Evangs Mailoa[1*]

[1] *Faculty of Information Technology, Satya Wacana Christian University, Central Java, Indonesia*

[2] *Department of Information Management, Chaoyang University of Technology, Taiwan, ROC*

ABSTRACT

Vaccines are the solution that is currently underway to tackle COVID-19. In this paper, vaccine distribution for hospitals in Central Java is developed. The problem case in this paper is classified as a Capacitated Vehicle Routing Problem (CVRP). The method proposed is using an initial route that follows the cluster-first route-second method (CFRS). The same size K-means is used for the clustering phase and the Greedy algorithm is used for the routing phase. The result of the initial route is a clustered route for each vehicle with a balanced capacity. Then, each cluster was re-optimized using metaheuristics Guided Local Search from Google OR-tools. Our experiment results have proven that using the initial route has the effect of reducing runtime by 97.37% - 99.17% when compared to without the initial route. This is because using initial routes with the same size K-means means breaking the problem into parts, then using the Greedy algorithm can reduce the number of possible routes. However, the total distance increased by 8.22% - 16.69% because no cluster member is allowed to move to another cluster.

Keywords: Capacitated vehicle routing problem, Vaccine distribution, K-means.

OPEN ACCESS

Received: April 20, 2021
Revised: August 10, 2021
Accepted: September 7, 2021

Corresponding Author:
Evangs Mailoa
evangs.mailoa@uksw.edu

Publisher:
Chaoyang University of Technology
ISSN: 1727-2394 (Print)
ISSN: 1727-7841 (Online)

1. INTRODUCTION

Coronavirus disease (COVID-19) carries terrible nightmares around the world. As of 22 February 2021, more than 110 million people have been infected with COVID-19. This is what has prompted many parties to compete to make vaccines against COVID-19. World Health Organization R&D Blue Print (2021) lists 84 vaccines in clinical development and 184 vaccines in pre-clinical development.

In Indonesia, vaccination has started, but it is not yet widespread. As of 22 February 2021, only 700 thousand people have been vaccinated out of 270 million Indonesians. In the future, vaccines must be distributed widely and massively. Weintraub et al. (2021) explained that the COVID-19 vaccine is urgent and needs proper vaccine distribution strategies to ensure that the distribution suppresses transmission.

Vaccine distribution for hospitals in Central Java is developed in this paper. Central Java is the third province with the most COVID-19 cases in Indonesia. Vaccines are distributed in 285 hospitals in Central Java where each hospital has different vaccine needs because it is based on the hospital's class. Distribution is carried out by several vehicles that had capacity limitations. The problem case in this paper is classified as a Capacitated Vehicle Routing Problem (CVRP) that is a variant of the Vehicle Routing Problem (VRP).

Furthermore, VRP has always been a concern in the field of Operations Research since its discovered by Dantzig and Ramser (1959). VRP has many constraints because it came from a real-world problem. Some of the famous constraints in VRP are

https://doi.org/10.6703/IJASE.202112_18(6).002

Capacitated VRP (CVRP) (Ibrahim et al., 2019), VRP with Time Windows (VRPTW) (Gil et al., 2020), VRP with pickups and deliveries (Ancele et al., 2021), and VRP with multiple depots (Jayarathna et al., 2020).

In this paper, Capacitated Vehicle Routing Problem (CVRP) is the problem case. CVRP is classified as an NP-hard problem and it can be exactly solved only for a small-sized problem, the heuristics approach yields the best results but does not guarantee optimality (Caric and Gold, 2008). Using metaheuristics is the common way to solve CVRP because the problem can be solved in a reasonable computational time although the result is not the best solution. Some metaheuristics that are widely used to solve CVRP are: Genetic Algorithms (Kramer, 2017), Guided Local Search (Dorigo and Socha, 2018), Ant Colony Optimization (Dorigo and Socha, 2018), Simulated Annealing (Delahaye et al., 2019), and Tabu Search (Sivaram et al., 2019).

One of the methods to solve CVRP is the cluster-first route-second method (CFRS). CFRS is a two-phase technique, the first are clustering, and then the second is routing where each cluster is treated independently as a Traveling Salesman Problem (TSP).

The following are the applications of the CFRS method in several papers to solve CVRP. Li et al. (2020) help the E-commerce product distribution problem with the K-means as the clustering method and Tabu Search as the routing method. Mostafa and Eltawil (2017) employed K-means as a clustering method and added valid inequalities as a routing method for customer supply chains.

Comert et al. (2018) help supermarket chains using K-means, K-medoids, and random clustering method than using branch and bound algorithm in each cluster. They compare the K-means, K-medoids and random clustering method and discovered that K-medoids provided better solution combined with branch and bound algorithm. Shalaby et al. (2020) handle the clustering phase using a modified demand weighted fuzzy c-means and used traditional optimization software for the routing phase.

In this paper, CFRS will be used as well for creating an initial route that will be re-optimized in each cluster. For creating CFRS, Clustering will be carried out with the same size K-means and Greedy algorithm for the routing phase. This paper implemented the same size K-means that will consider hospital demand so that each vehicle has the same vaccine capacity to distribute. The Greedy algorithm will solve the problem based on the shortest distance from the current node recursively. The first node used for the Greedy algorithm is the cluster's centroid. The reason for combining the same size K-means and Greedy algorithm for the CFRS method is because K-means and Greedy algorithm are well known for their speed and simplicity.

After the initial route is created, each cluster will be re-optimized using metaheuristic Guided Local Search. The purpose is to improve the performance of the Google Optimization Tools (OR-tools) (Perron, 2011) when solving CVRP using Guided Local Search. The rest of this paper is organized as follows. In section 2, the materials and methods are explained. Section 3 describes the result and discussion. Section 4 outlines the conclusion and possible future research direction.

2. MATERIALS AND METHODS

2.1 Data Collection and Processing

The dataset used in this paper is scraped from Central Java's Covid-19 Official Website (Tanggap COVID-19 Provinsi Jawa Tengah, 2021). This dataset contains 285 hospital data including the hospital's name, address, class, latitude, and longitude. Some hospitals do not have complete addresses, latitude, and longitude, so the data is completed using Geocoding API from Google Maps Platform.

Moreover, there are two ways of distance calculation in this paper. First, Euclidean distance used latitude and longitude data for processing the initial route. Second, distance matrix that used Distance Matrix API from Google Maps Platform. The size of the original distance matrix is 285 × 285, but it became a 286 × 286 distance matrix because one dummy data was added so that there was no CVRP depot. This distance matrix will be used in Google OR-tools.

2.2 Algorithm for Creating Initial Route

This paper will compare the result of Google OR-tools when solving CVRP without an initial route and with an initial route, especially using Guided Local Search as a metaheuristic. The initial route follows the CFRS with few modifications. There are four stages to create the initial route. The first stage, modify the number of points in each hospital based on hospital class. The second stage, use the traditional K-means algorithm to partition hospitals into several areas based on the number of vehicles. Create the same size K-means in the third stage. Last, modify each point again and use the Greedy algorithm for each cluster to get the initial route.

2.2.1 Modify the Number of Each Point Based on the Hospital Class

In Indonesia, hospitals are divided into several classes. Class shows the readiness of the hospital which in this case indicates readiness when dealing with COVID-19. There are three classes, namely class 1, 2, and 3 with class 1 hospitals being the most prepared. In this paper, it is assumed that class 1 requires 2000 vaccines, class 2 requires 1000 vaccines, and class 3 requires 500 vaccines. The lower limit which has been set is 500 vaccines, meaning that 1 point represents 500 vaccines. Moreover, hospitals with class 1 will become 4 points, hospitals with class 2 will become 2 points, and hospitals with class 3 will become 1 point. The purpose of this modification is to create a cluster that considers the number of needs.

https://doi.org/10.6703/IJASE.202112_18(6).002

2.2.2 K-means Clustering

The clustering analysis method is one of the main analytical methods. K-means is one of the simplest and widely used data clustering algorithms (Sinaga and Yang, 2020). K-means is often referred to as Lloyd's algorithm (Lloyd, 1982). The number of clusters must be defined in K-means. In this paper, each cluster represents one vehicle. For distance calculation, Euclidean distance is used. The Euclidean distance between data point a and b is given by formula (1) (Smith, 2011):

$$distance(a,b) = \sqrt{(a_x - b_x)^2 + (a_x + b_x)^2} \quad (1)$$

Given an initial set of K-means $m_1^{(1)}$, …, $m_k^{(1)}$, the K-means result is by proceeds these two steps (Mackay, 1995). First by assigning each observation to the cluster with the nearest mean as seen in formula (2). Second, with the update step, recalculate means (centroids) for observations assigned to each cluster as seen in formula (3).

$$S_i^{(t)} = \left\{x_p : \| x_p - m_i^{(t)} \|^2 \le \| x_p - m_j^{(t)} \|^2 \ \forall j, 1 \le j \le k\right\} \quad (2)$$

$$m_i^{(t+1)} = \frac{1}{\left|S_i^{(t)}\right|} \sum_{x_j \in S_i^{(t)}} x_j \quad (3)$$

2.2.3 Same Size K-means

The purpose of making the same size K-means is to make each vehicle has the same capacity to distribute. This is important because, with traditional K-means, the distribution capacity of each vehicle is not balanced.

The steps for creating the same size K-means used in this paper are described as follows.

1. Do clustering with traditional K-means.
2. Set the time limit.
3. If runtime is less than the time limit, iterate step 4 until step 7.
4. Set the first cluster as the current cluster.
5. If the current cluster is less than the equal number, move the shortest points in other clusters that exceed the equal number to the current cluster until the current cluster reaches the equal number.
6. If the current cluster exceeds the equal number, move the points in current clusters that have the shortest points to other clusters until the current cluster reaches the equal number.
7. Change the current cluster in order. Repeat steps 5 and 6 until all clusters are processed.
8. If all clusters are processed, go to step 3.

2.2.4 Modify Each Point and Use the Greedy Algorithm for Each Cluster

After getting the results from the same size K-means, the points need to be modified again to produce the same number as the number of hospitals. Cluster determination at each point is selected from the dominant cluster. In the last step, use the Greedy algorithm to connect all points in each cluster and the initial route is ready to use. The step for using the Greedy algorithm in each cluster is described as follows.

1. Select the centroid for starting point.
2. Find for unvisited nearest point and go there.
3. Repeat step 2 until all point is visited.

2.3 Google OR-tools

Google Optimization Tools (Google OR-tools) is an open-source, fast, and portable software suite for solving combinatorial optimization problems. Google OR-tools provides algorithms for solving Travelling Salesman Problem and Vehicle Routing Problem.

In our work, the results of CVRP when using the initial route and without the initial route are compared. Both using libraries from Google OR-tools with Guided Local Search as the metaheuristics. Guided Local Search is a high-level strategy that interacts with the local improvement procedure using an efficient penalty-based technique. This interaction results in a process that can escape local optima, improving the efficiency and robustness of the underlying local search algorithms (Tsang et al., 2016).

For distance calculation, a distance matrix is used. The distance matrix is used for the final stage, either with the initial route or without the initial route. The use of a distance matrix is to make the route solution to be in accordance with real-life situations.

When solving with the initial route, each cluster is run separately and treated as Travelling Salesman Problem. When solving without an initial route, the problem is treated as Vehicle Routing Problem. The parameters used in Google OR-tools are shown in the table below (Google developers, 2020).

3. RESULTS AND DISCUSSION

3.1 Initial Route Results

Our experiment employed data from 285 hospitals in Central Java which have an area of 32,801 km². The experiment is assumed with the number of vehicles 3 (low), 5 (medium), and 7 (high). The number of vehicles represents the number of clusters created which is the k value in K-means. Furthermore, by using the algorithm described in this paper, an initial route is created.

Fig. 1 shows the process of creating an initial route for 3 vehicles. Fig. 1(a) was the k-means results with modified points that make 285 points become 380 points. The points increase due to capacity adjustment as described in Algorithm 1. In Fig. 1 (b), the same size K-means was created. The last step is to modify the points again back to 285 points, then the points are connected with the Greedy algorithm, see Fig. 1(c). The result of Fig. 1(c). will be used as the initial route in Google OR-Tools. The same step is shown for 5 vehicles in Fig. 2, and Fig. 3 for 7 vehicles.

Further, Same size K-means result as seen in Fig. 1(b) and Fig. 2(b), no point in the cluster has deviated too far. But in Fig. 3(c), some points have deviated too far and it should be more effective if those points move to another cluster.

Table 1. Google OR-tools parameters

Parameter	Value	Description
FirstSolutionStrategy	PATH_CHEAPEST_ARC	Starting from a route "start" node, connect it to the node which produces. The cheapest route segment then extends the route by iterating on the last node added to the route. FirstSolutionStrategy is used when the program runs without an initial route.
LocalSearchMetaheuristic	GUIDED_LOCAL_SEARCH	Uses guided local search to escape local minima. This is generally the most efficient metaheuristic for vehicle routing.
ReadAssignmentFromRoutes	initial_route	Specify a set of initial routes for a CVRP. The initial route given is based on the algorithm described in this paper.
Time_Limit	Between 100 to 1000	The time limit is in milliseconds to the search.

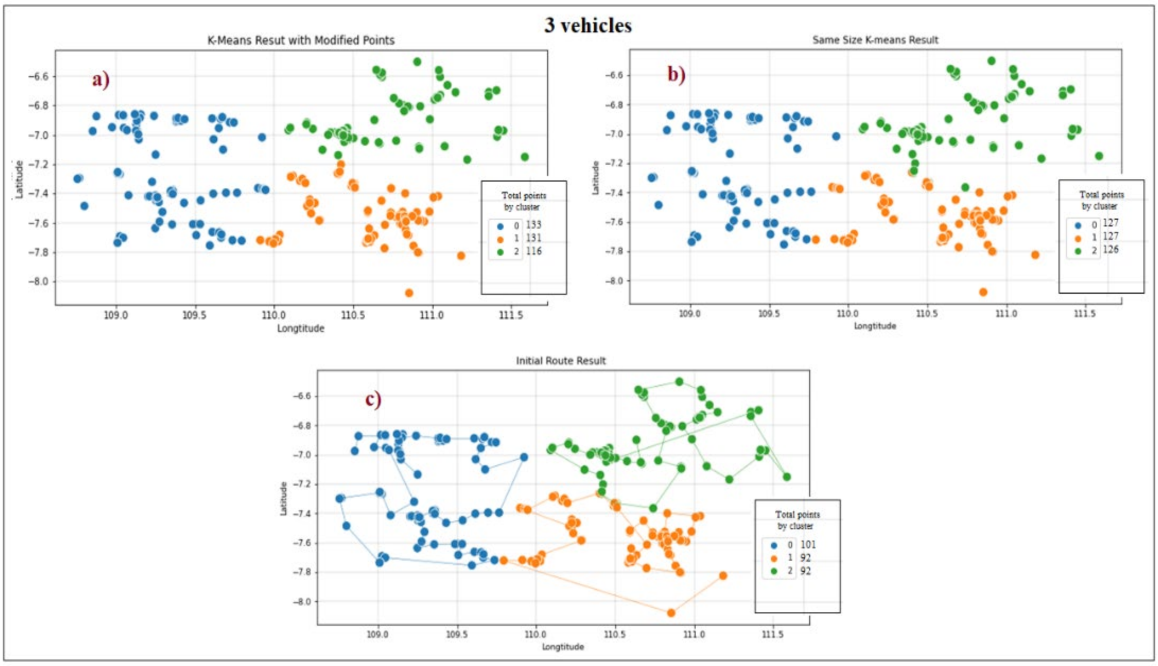

Fig. 1. Initial route for 3 vehicles. (a) K-Means result with modified points. (b) Same size K-means result. (c) Initial route result using Greedy algorithm

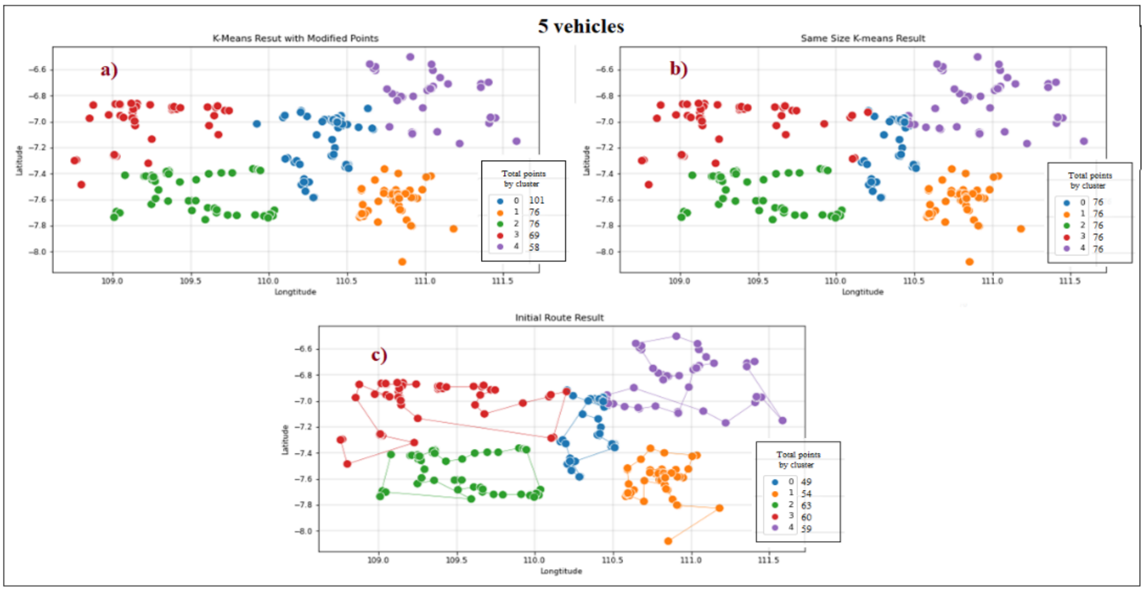

Fig. 2. Initial route for 5 vehicles. (a) K-Means result with modified points. (b) Same size K-means result. (c) Initial route result using Greedy algorithm

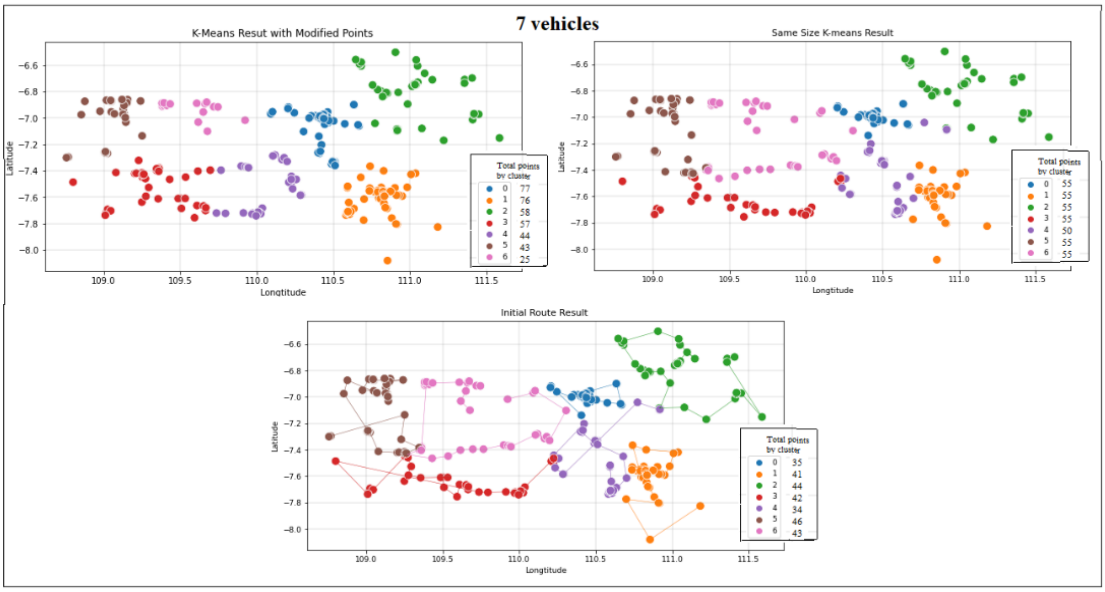

Fig. 3. Initial route for 7 vehicles. (a) K-Means result with modified points. (b) Same size K-means result. (c) Initial route result using Greedy algorithm

3.2 Result of CVRP Plotted in Central Java Map

The final route after being processed with Google OR-Tools is shown in Fig. 4 for 3 vehicles, Fig. 5 for 5 vehicles, and Fig. 6 for 7 vehicles. All results are plotted on a map of Central Java. The result with no initial route, as seen in Fig. 4(a), Fig. 5(a), and Fig. 6(a) represent that each vehicle has many scattered points. These points are scattered far away, but the road is not passed many times. Hence, the result with the initial route is shown in Fig. 4(b), Fig. 5(b), Fig. 6(b), respectively. Each vehicle has adjacent points, but the road is passed many times and looked less effective.

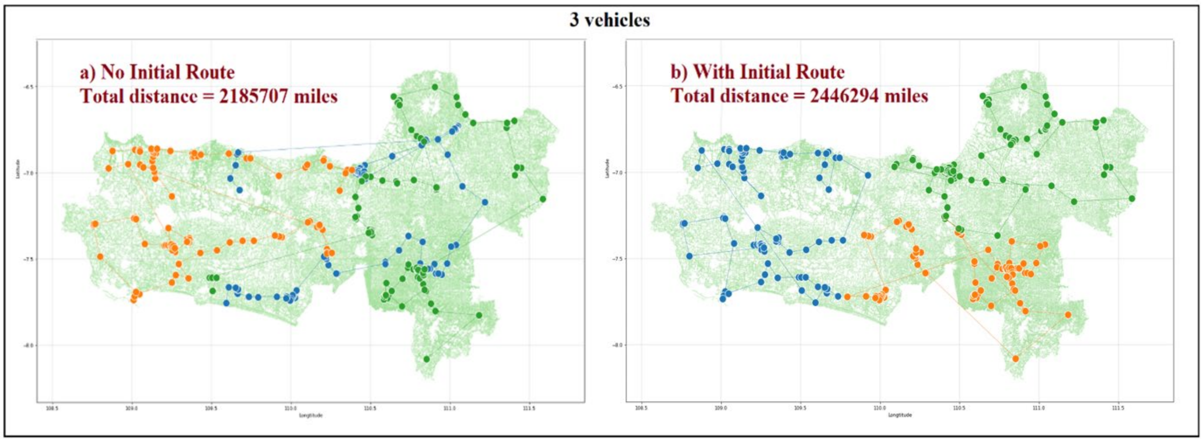

Fig. 4. Route plotted for 3 vehicles. (a) No initial route results. (b) With initial route result

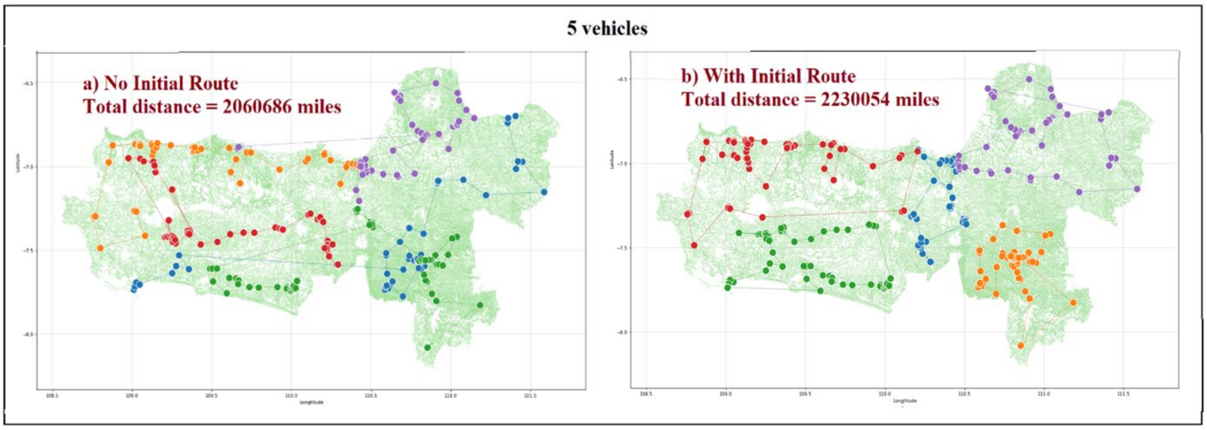

Fig. 5. Route plotted for 5 vehicles. (a) No initial route results. (b) With initial route result

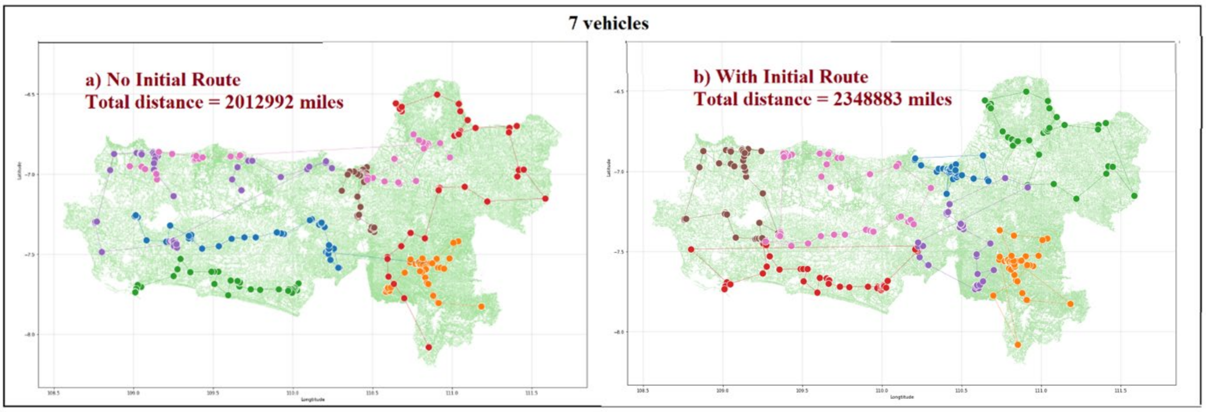

Fig. 6. Route plotted for 7 vehicles. (a) No initial route results. (b) With initial route result

3.3 Total Distance Comparison

As seen in Fig. 7, Guided Local Search in Google OR-tools with initial route performs worse in total distance. This is because points can't move to another cluster, make the solution stuck on local optima. In Table 2, the detailed results are shown with the better results are shown as bold text for each vehicle. Each problem was repeated 10 times with a time limit change from 100 milliseconds to 1000 milliseconds.

From Table 2, the worst increase is in Vehicle 7, this is following the results of the same size K-means with Vehicle 7 that is not good enough. The results of the same size K-means greatly affect the total distance.

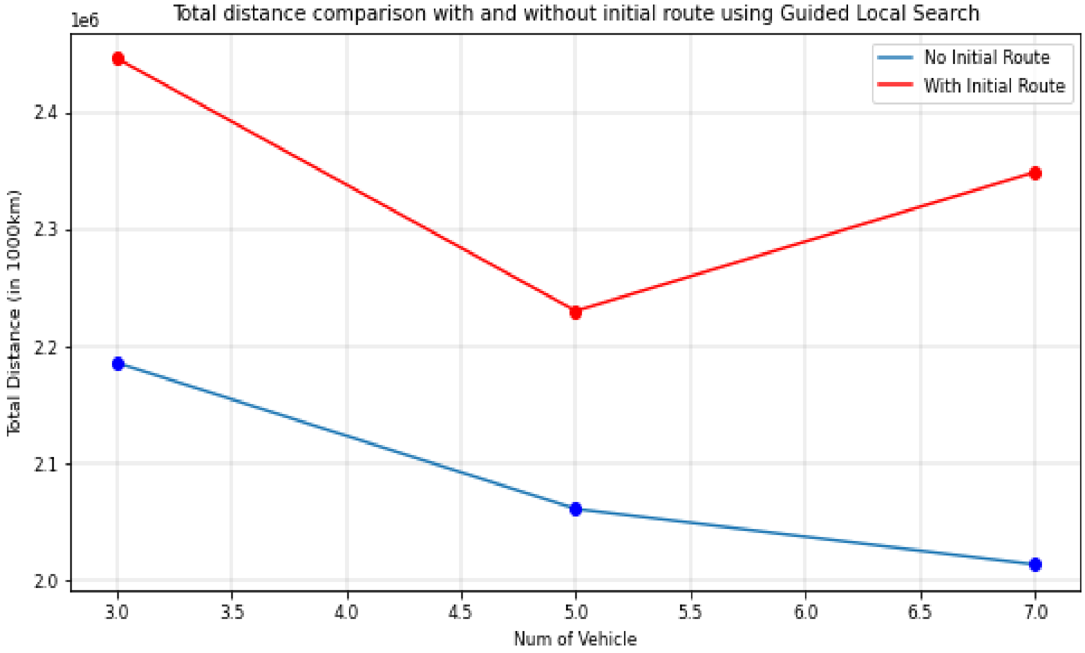

Fig. 7. Total distance comparison

Table 2. Total distance comparison of each solution

Vehicle Number	Without Initial Route (miles)	With Initial Route (miles)	Initial Route Effect
Vehicle: 3	**2185707**	2446294	+11.92%
Vehicle: 5	**2060686**	2230054	+8.22%
Vehicle: 7	**2012992**	2348883	+16.69%

3.4 Runtime Comparison

The runtime between both methods is very different, as seen in Fig. 8. The summary based on Table 3 is as follows:

1. Using the initial route can reduce runtime by up to 98.89% in Google OR-tools.
2. The average runtime when using the initial route is 5

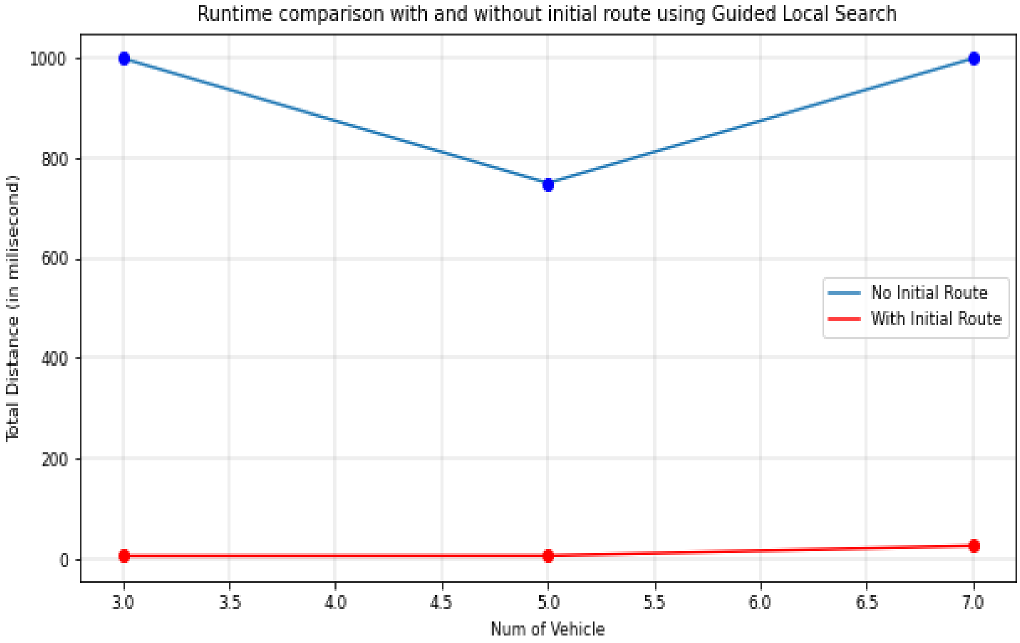

Fig. 8. Runtime comparison

Table 3. Runtime comparison of each solution

Vehicle Number	Without Initial Route (milliseconds)	With Initial Route (milliseconds)	Initial Route Effect
Vehicle: 3	1000.16614	**5.85955**	-99.41%
Vehicle: 5	750.00998	**6.25637**	-99.166%
Vehicle: 7	1000.00953	**26.30466**	-97.37%

- 6 milliseconds. However, it depends on how fast the program creates the same size K-means. In vehicle 7, it takes a longer time to create the same size K-means.

3. While using the initial route the runtime increases with the number of vehicles.

4. CONCLUSION

In this paper, the results of using the initial route that follow the cluster-first route-second method (CFRS) in solving Capacitated Vehicle Routing Problem (CVRP) are presented. For creating an initial route, the same size K-means is used for the clustering phase and the Greedy algorithm is used for the routing phase. Re-optimization of each cluster was also done by using the metaheuristics method Guided Local Search from Google OR-tools.

As shown in Table 2, the initial route has the effect of reducing runtime by 97.37% - 99.166% when compared to without the initial route. This is because using initial routes with the same size K-means means breaking the problem into parts, then using the Greedy algorithm can reduce the number of possible routes. Nevertheless, when it is run with Google OR-tools using metaheuristics Guided Local Search, there is a reduction in runtime which is very significant.

However, the total distance when using the initial route increased by 8.22% - 16.69% as seen in Table 3. This is because no cluster member is allowed to move to another cluster, and make the route stuck on local optima. Hence, there is no relationship between clusters because each cluster is run separately.

In conclusion, using the initial route with the same size K-means and Greedy algorithm can solve CVRP in a more acceptable runtime, especially in problems that must be solved quickly such as vaccine distribution. However, this method does not produce a better total distance. For future work, the testing will be carried out on a larger dataset. Also, the addition of constraints such as Time Window should be considered.

ACKNOWLEDGMENT

This research would like to thank the support of Satya Wacana Christian University and Chaoyang University of Technology, Taiwan.

REFERENCES

Ancele, Y., Hà, M.H., Lersteau, C., Matellini, D.B., Nguyena, T.T. 2021. Toward a more flexible VRP with pickup and delivery allowing consolidations, Transportation Research Part C: Emerging Technologies, 128, p. 103077. DOI: https://doi.org/10.1016/j.trc.2021.103077

Caric, T., Gold, H. 2008. Vehicle routing problem. Edited by T. Caric and H. Gol. InTech. DOI: 10.5772/64.

Comert, S.E., Yazgan, H.R., Kır, S., Yener, F. 2018. A cluster first-route second approach for a capacitated vehicle routing problem: A case study, International Journal of Procurement Management, 11, 399–419. DOI: 10.1504/IJPM.2018.092766.

Dantzig, G.B., Ramser, J.H. 1959. The truck dispatching problem, Management Science, 6, 80–91. DOI: https://doi.org/10.1287/mnsc.6.1.80

Delahaye, D., Chaimatanan, S., Mongeau, M. 2019. Simulated annealing: From basics to applications, International Series in Operations Research and Management Science, 272, 1–35. DOI: 10.1007/978-3-319-91086-4_1.

Dorigo, M., Socha, K. 2018. An introduction to ant colony optimization, Handbook of Approximation Algorithms and Metaheuristics, Second Edition, 395–408. DOI: 10.1201/9781351236423-23.

Gil, A.F., Sánchez, M.G., Lalla-Ruiz, E., Castro, C. 2020. Cumulative VRP with time windows: A trade-off analysis, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12433 LNCS, 277–291. DOI: https://doi.org/10.1007/978-3-030-59747-4_18

Google developers, 2020. Google OR-Tools reference. Available at: https://developers.google.com/optimization/reference (Accessed: March 31, 2021).

Ibrahim, A.A., Lo, N., Abdulaziz, R.O., Ishaya, J.A. 2019. Capacitated vehicle routing problem, International Journal of Research-Granthaalayah, 7, 310–327. doi: https://doi.org/10.29121/granthaalayah.v7.i3.2019.976

Jayarathna, N., Lanel, J., Juman, Z.A.M.S. 2020. Five years of multi-depot vehicle routing problems, Journal of Sustainable Development of Transport and Logistics, 5, 109–123. DOI: 10.14254/JSDTL.2020.5-2.10.

Kramer, O. 2017. Genetic algorithms, Studies in Computational Intelligence, 679, 11–19. DOI: 10.1007/978-3-319-52156-5_2.

Li, L., Chen, Y., Meng, J. 2020. Improved Tabu search algorithm for solving the vehicle routing problem with soft time windows in B2C environment, in Proceedings of the 32nd Chinese Control and Decision Conference, CCDC 2020. Institute of Electrical and Electronics

Engineers Inc., 3629–3633. DOI: 10.1109/CCDC49329.2020.9164444.

Lloyd, S. 1982. Least squares quantization in PCM, IEEE Transactions on Information Theory, 28, 129–137.

Mackay, D.J.C. 1995. Information theory, Inference, and Learning Algorithms. Available at: http://www.inference.phy.cam.ac.uk/mackay/itila/ (Accessed: March 31, 2021).

Mostafa, N., Eltawil, A. 2017. Solving the heterogeneous capacitated vehicle routing problem using K-means clustering and valid inequalities, Proceedings of the International Conference on Industrial Engineering and Operations Management Rabat, Morocco, 11–13.

Perron, L. 2011. Operations research and constraint programming at google, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, p. 2. DOI: 10.1007/978-3-642-23786-7_2.

Shalaby, M.A.W., Mohammed, A.R., Kassem, S. 2020. Modified fuzzy c-means clustering approach to solve the capacitated vehicle routing problem, Proceedings - 2020 21st International Arab Conference on Information Technology, ACIT. DOI: 10.1109/ACIT50332.2020.9300082.

Sinaga, K.P., Yang, M.S. 2020. Unsupervised K-means clustering algorithm, IEEE Access, 8, 80716–80727. DOI: 10.1109/ACCESS.2020.2988796.

Sivaram, M., Batri, K., Mohammed, A.S., Porkodi, V. 2019. Exploiting the local optima in genetic algorithm using Tabu search, Indian Journal of Science and Technology. DOI: 10.17485/just/2019/v12i1/139577.

Smith, K.J. 2011. Precalculus: A functional approach to graphing and problem solving. Jones\& Bartlett Publishers.

Tanggap COVID-19 Provinsi Jawa Tengah, 2021. Rumah sakit rujukan COVID-19 Di jawa tengah. Available at: https://corona.jatengprov.go.id/rumah-sakit (Accessed: February 23, 2021).

Tsang, E.P.K. 2016. Market microstructure view project guided local search. doi: 10.1007/978-3-319-07153-4_2-1.

Weintraub, R.L., Subramanian, L., Karlage, A., Ahmad, I., Rosenberg, J. 2021. Covid-19 vaccine to vaccination: Why leaders must invest in delivery strategies now, Health Affairs, 40, 33–41. DOI: https://doi.org/10.1377/hlthaff.2020.01523

World Health Organization R&D Blue Print, 2021. Draft landscape and tracker of COVID-19 candidate vaccines. Available at: https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines (Accessed: March 31, 2021).