Job Assignment Problem on Online Transportation Order Using Hungarian Algorithm

Andreas Handojo
Industrial Engineering Department,
Sepuluh Nopember Institute of Technology
Informatics Department, Petra Christian University
Surabaya, Indonesia
handojo@petra.ac.id

Nyoman Pujawan
Industrial Engineering Department
Sepuluh Nopember Institute of Technology
Surabaya, Indonesia
0000-0002-9587-8152

Budi Santosa
Industrial Engineering Department
Sepuluh Nopember Institute of Technology
Surabaya, Indonesia
0000-0003-1753-9603

Moses Laksono Singgih
Industrial Engineering Department
Sepuluh Nopember Institute of Technology
Surabaya, Indonesia
0000-0002-6200-6411

***Abstract*— The more advanced a civilization is, the higher the mobility of its population. It takes a means of transportation that can answer the needs of the community. Currently, online transportation is an option in answering public transportation needs. The uniqueness of the online transportation business is that passengers can place orders anytime and anywhere. On the other hand, the driver can also be active/available anytime and anywhere. This causes problems in fulfilling orders, providers must consider various factors such as the time and distance of the driver to get to the pick-up location in determining which driver is assigned to fulfill the order. This study simulates order assignment using the Hungarian method, Minimum Value of Selected Factor in Data Collection, and random assignment. From the test results, it is known that the highest average pick-up duration and pick-up distance are obtained using the random method and the lowest average using the Hungarian method. The highest average time to determine the assignment is obtained by the Hungarian method, while the smallest is the random method. Based on these results, it can be concluded that the Hungarian method is better than the other two methods.**

Keywords*— *online transportation, Hungarian, job assignment, simulation, passenger

I. INTRODUCTION

Transportation is an important part of people's lives. In the last decades, online transportation has become one of the preferred choices for public transportation [1][2]. Passengers can order transportation (generally using cars and motorbikes) online from anywhere and anytime. The online transportation service provider company will then appoint available drivers to fulfill the order. The demand for online transportation is similar to transportation in general, for example, in the morning the demand is quite high (rush hour) appearing in suburban areas to go to the city center (going to work) while in the afternoon the demand appears on the contrary (going home from work).

One of the uniqueness of the online transportation business model is that drivers are not permanent employees of online transportation service providers. Drivers also do not have fixed working hours so drivers can work anytime and stop working at any time. This condition certainly causes uncertainty in meeting demand [3][4][5].

The location of the driver to the passenger's location is very influential on this transportation model because meeting demand with supply from other areas that are very far away is very impossible because the time and distance required by the supply/driver to reach the pick-up location can be very large. This, can increase transportation costs (due to the distance factor) and waiting time (due to the travel time factor). This can also increase dissatisfaction from both drivers and passengers which can also lead to order cancellations (from passengers) or cancellation of assignment matching (from drivers).

Therefore, the online transportation service provider requires the right mechanism in making assignments to drivers to fulfill an order. Where this mechanism can consider the time and distance factor for the driver to get to the pick-up location. This study will simulate online transportation business processes starting from requesting orders from passengers, available drivers, locations of passengers and drivers, possible cancellations from passengers and drivers. The distance and travel time from the driver to the pick-up location will use the google map API [6]. The assignment process from the driver will use Hungarian, Minimum Value of Selected Factor in Data Collection (MVSF) in this case will be used First Come First Serve (FCFS), and random assignment..

II. HUNGARIAN ALGORITHM

The Hungarian algorithm is one of the combinatorial optimization methods that aims to solve the problem of placement or installation of 2 types of objects by finding the optimal value of the variable that connects the 2 types of objects. Hungarian method is an algorithm for the optimal assignment problem by producing the installation of variables with the smallest resources [7].

Steps from the Hungarian method can be described as follow:

1.Find the minimum element from each row and subtract that value from all the elements of the row.

2. Find the minimum element from each column and subtract that value from all the elements of the column.

3. Let m=minimum number of lines required to cover all the zeroes in the table

4. while (m!=number of row/columns)

- Find the minimum element from the uncovered elements
- Subtract this element from all the other uncovered elements
- Add this element to the elements where the lines are intersecting
- Find new m

5. Use the zeroes to assign possible combinations- i.e. wherever there is a zero present, task can be assigned.

6. Find the minimum cost

7. End

III. NORMALIZATION

The value of each factor from the data such as duration, distance, total passengers who have been served will be normalized using the min max formula to generate a new value with a lower limit of 1 and an upper limit of 10 as shown on equation 1.

Min max formula [8]:

$$v' = \frac{v - min}{max - min}(new\_max - new\_min) + new\_min \quad (1)$$

Where v' is the normalized data sample, v is the original data sample, min is the minimum data among any attribute of the original dataset, max is the maximum data among any attribute of the original dataset, new_min is the minimum of the normalized dataset, and new_max is the maximum of the normalized dataset.

IV. IMPLEMENTATION

A. Data Normalization

First, the simulation will randomize the appearance of orders from passengers and their pick-up locations. The simulation will also randomize the appearance and location of the driver. The simulation then makes a request to the Google Map API for the distance and travel time from the location of each driver to each pick-up location. In this way, the distance and travel time obtained will also consider traffic conditions on the path that must be passed in real time.

The data obtained (data of distance and travel time to the pick-up location, total number of passengers who have been served) is then normalized. before being processed by the method used, all the normalization results from the factors taken into consideration are added.

The following is an example of a simulation calculation with data for three passengers and three drivers. Data on pick-up duration, pick-up distance, and total passengers served can be seen in Tables 1, 3, and 5. While the normalization results can be seen in Tables 2, 4, and 6.

TABLE I. DURATION (IN SECONDS)

Passenger ID	Driver ID		
	1	*2*	*3*
1	276	488	220
2	943	180	339
3	614	252	677

TABLE II. DURATION NORMALIZATION

Passenger ID	Driver ID		
	1	*2*	*3*
1	2.122	4.626	1.437
2	10.000	1.000	2.854
3	6.078	1.815	6.825

TABLE III. DISTANCE (IN METERS)

Passenger ID	Driver ID		
	1	*2*	*3*
1	5394	4805	3105
2	1199	5293	2617
3	5320	4117	1930

TABLE IV. DISTANCE NORMALIZATION

Passenger ID	Driver ID		
	1	*2*	*3*
1	10.000	8.739	5.102
2	1.000	9.784	4.055
3	9.842	7.267	2.586

TABLE V. TOTAL CUSTOMER HAVE BEEN SERVED

Driver ID		
1	*2*	*3*
23	100	39

TABLE VI. TOTAL CUSTOMER HAVE BEEN SERVED

Driver ID		
1	*2*	*3*
1.000	10.000	2.870

The results of the normalization of the three factors are then added up (shown in Table 7).

TABLE VII. DURATION (IN SECONDS)

Passenger ID	Driver ID		
	1	*2*	*3*
1	13.123	21.000	18.790
2	14.365	20.783	11.950
3	7.540	16.911	12.291

B. *Assignment Process*

After adding all the factors, then the result will be calculated using the method specified by the user (Hungarian method, MVSF/FCFS, Random). From the data, the assignment results are obtained as shown in Table 8.

TABLE VIII. ASSIGNMEINT RESULT

Method		
Hungarian	***MVSF***	***Random***
Driver 1 → Passenger 3	Driver 1 → Passenger 3	Driver 1 → Passenger 2
Driver 2 → Passenger 1	Driver 2 → Passenger 2	Driver 2 → Passenger 3
Driver 3 → Passenger 2	Driver 3 → Passenger 1	Driver 3 → Passenger 1

The results of the assignment are then mapped on the Google Map as shown in Fig. 1.

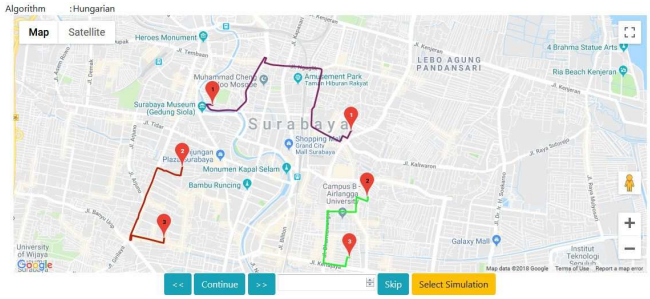

Fig. 1. Mapping results using the Hungarian method on Google Map

C. *Assignment Testing*

Testing is carried out on requests in the morning starting at 07.00 in the morning. This testing aims to see the results of the trial in conditions of traffic jams in the morning. Testing by adding the possibility of canceling orders from passengers and drivers by 25%. The test was carried out for 3641 seconds with a total of 364 batches. Test results (Table 9):

- Simulation with the Hungarian method gives the lowest average value of the duration and distance factor but produces the lowest total success of 133.
- The simulation with the MVSF method gives the average value of the duration and distance factor between Hungarian and random and produces the highest total success of 144.
- Simulation with random method gives the highest average value of the duration and distance factor and produces the highest total success of 144.

In this trial, the MVSF and random methods both produced the highest total success with the same value of 144 compared to the Hungarian algorithm with a value of 133. However, the Hungarian method was still superior in the efficiency of the average value of the lowest duration and distance factor.

TABLE IX. TESTING RESULT

	Hungarian	***MVSF***	***Random***
Total Batch	364		
Calculation Formula	Duration + Distance + Total Service		
Running Time	3641 seconds		
Range Driver Wait	100 – 1000 Seconds		
Range Passenger Wait	100 – 1000 Seconds		
Driver Cancel Probability	25%		
Passenger Cancel Probability	25%		
Assignment Success	133	144	144
Cancel by Driver	17	12	9
Cancel by Passenger	29	22	24

The smallest and largest values displayed in the boxplot form of the testing data can be seen in Fig. 2, 3, and 4.

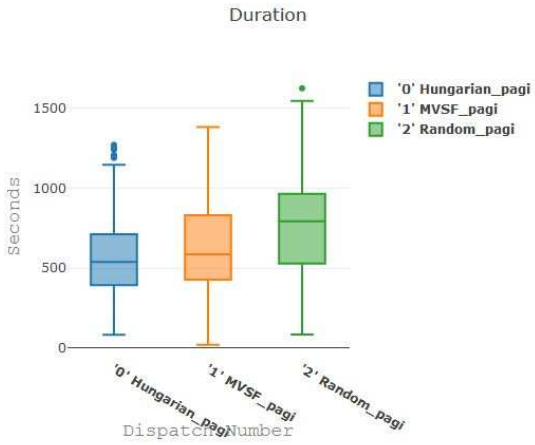

Fig. 2. Boxplot Duration

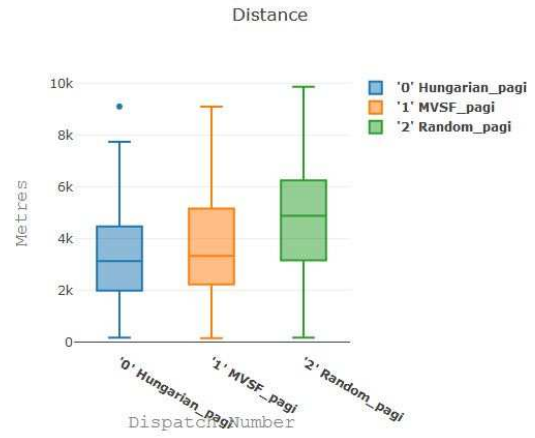

Fig. 3. Boxplot Distance

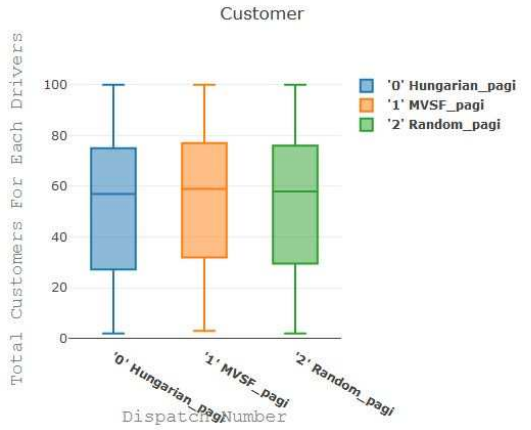

Fig. 4. Boxplot Total Customer Served

From the test results (as shown on Fig.5 to 8) with legend blue color as Hungarian, orange color as MVSF, and green color as Random method, it can be seen that the highest average pick-up duration (Fig. 5) and pick-up distance (Fig. 6) was obtained using the random method. While the lowest average is by using the Hungarian method. Meanwhile, the average number of passengers served by the three methods in each batch is roughly the same (Fig. 7). The highest average time required to determine the assignment was obtained by the Hungarian method, while the smallest was the random method (Fig. 8).

V. CONCLUSION

Online transportation (generally using cars and motorbikes) is growing and is becoming one of the choices of public transportation. One of the uniqueness of the online transportation business is that passengers can place orders anytime and anywhere. On the other hand, the driver is not an employee of a service provider company. The driver also does not have working hours so he can be active or inactive at any time. Similar to passengers, drivers can also be active/available from any location. This causes problems in fulfilling orders, providers must consider factors such as which drivers are assigned to fulfill existing orders. The system must also consider the time and distance of the driver to get to the pick-up location. Because the higher the factor, the longer the passengers have to wait and the greater the costs incurred by the driver. Therefore, providers must be able to incorporate these factors as a data assignment process. In this study, a simulation is carried out that compares the assignment by considering the factors above using Hungarian, Minimum Value of Selected Factor in Data Collection (MVSF) in this case will be used First Come First Serve (FCFS), and random assignment.

From the test results, it can be seen that the highest average pick-up duration and pick-up distance are obtained using the random method. While the lowest average is by using the Hungarian method. Meanwhile, the average number of passengers served by the three methods in each batch is approximately the same. The highest average time needed to determine the assignment is obtained by the Hungarian method, while the smallest is the random method. Based on these results, it can be concluded that the Hungarian method is better than the other two methods. For future research development, testing can also add other factors such as driver rating, passenger/driver activity level, driver waiting time. Future research can also make comparisons with other methods such as taboo search, branch and bound, goal programming. So that the best method for doing assignments is obtained.

ACKNOWLEDGMENT

We thank our colleagues Vincent Christian, Henry Novianus Palit, Tanti Octavia, Justinus Andjarwirawan for their valuable input, discussion and feedback.

REFERENCES

[1] A.A. Damaini, G.S. Nugroho, and S. Suyoto, “Fraud Crime Mitigation of Mobile Application Users for Online Transportation. International Journal of Interactive Mobile Technologies,” 2018, 12 (3):153, https://doi.org/10.3991/ijim.v12i3.8070.

[2] M. Abdullah, N. Ali, M.A. Javid, C. Dias, and T. Campisi, “Public transport versus solo travel mode choices during the COVID-19 pandemic: Self-reported evidence from a developing country,” Transportation Engineering, Volume 5, 2021, 100078, ISSN 2666-691X, https://doi.org/10.1016/j.treng.2021.100078

[3] Q. Luo, and R. Saigal, “Dynamic Pricing for On-Demand Ride-Sharing : A Continuous Approach Dynamic Pricing for On-Demand Ride-Sharing : A Continuous Approach,” SSRN Electronic Journal, 2017, https://doi.org/10.2139/ssrn.3056498

[4] F. Miao, S. Han, A.M. Hendawi, M.E. Khalefa, J.A. Stankovic, and G.J. Pappas, Data-driven distributionally robust vehicle balancing using dynamic region partitions. Proceedings of the 8th International Conference on Cyber-Physical Systems – 2017, pp. 261–271. https://doi.org/10.1145/3055004.3055024

[5] A. Handojo, N. Pujawan, B. Santosa, and M.L. Singgih, “Online Transportation Order Matching Considering Passenger Priority Using Recency Frequency Monetary Method,” 2022. Asia Pacific Conference on Manufacturing Systems and The International Manufacturing Engineering Conference

[6] Google LLC. 2022. GoogleMaps platform documentation. Retrieved May 18, 2022 from https://developers.google.com/maps/documentation/

[7] Z. Pi-Xu, W. Hang-Yu, S. Zhang-Song, X. Wen-Kai, and W. Zhong-Hong, “Plot Association for Multi-Sensor Based on Hungary Algorithm,” Procedia Engineering, vol 24, 2011, Pages 17-21, ISSN 1877-7058, https://doi.org/10.1016/j.proeng.2011.11.2594

[8] P. J. Muhammad Ali, “Investigating the Impact of Min-Max Data Normalization on the Regression Performance of K-Nearest Neighbor with Different Similarity Measurements”, ARO, vol. 10, no. 1, 2022, pp. 85-91, http://dx.doi.org/10.14500/aro.10955

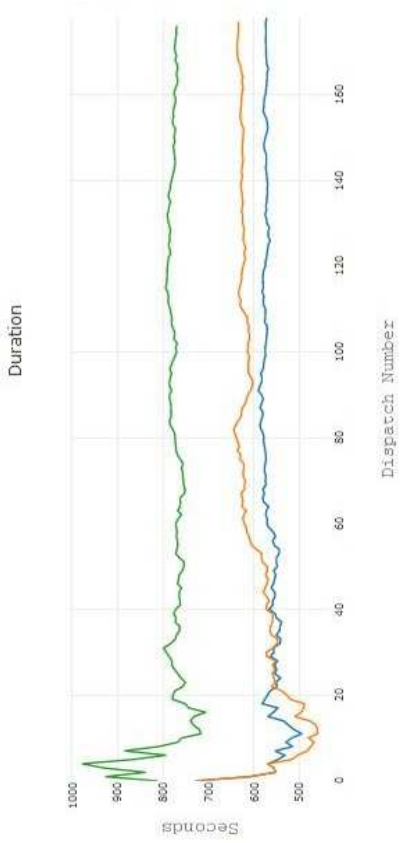

Fig. 5. Average Duration Result

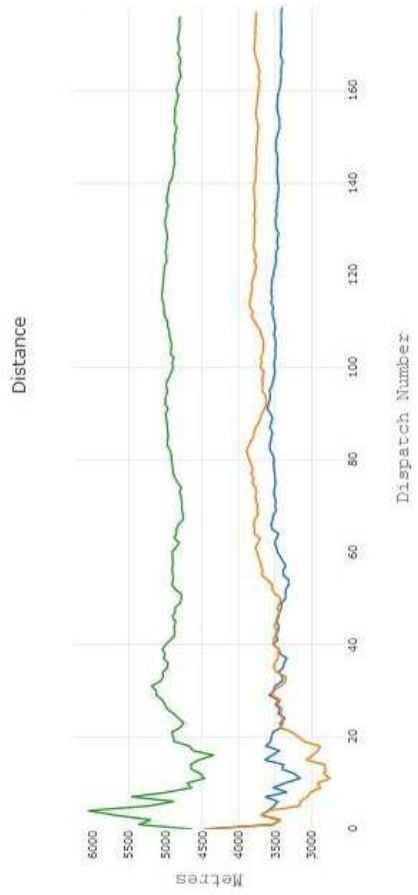

Fig. 6. Average Distance Result

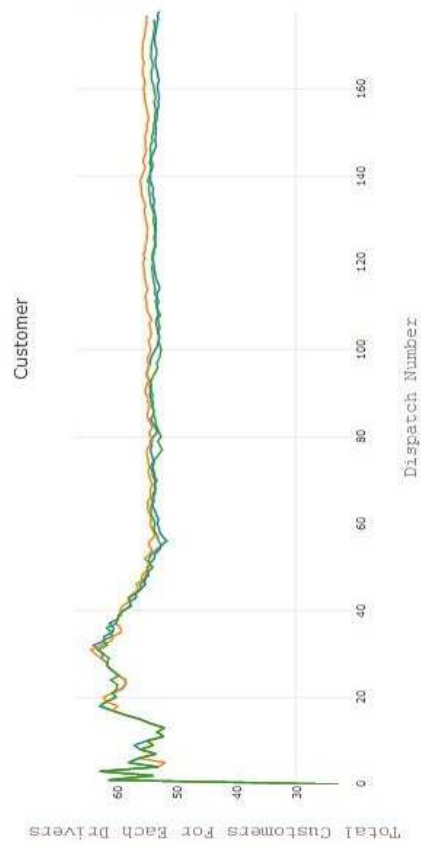

Fig. 7. Average Passenger Served Result

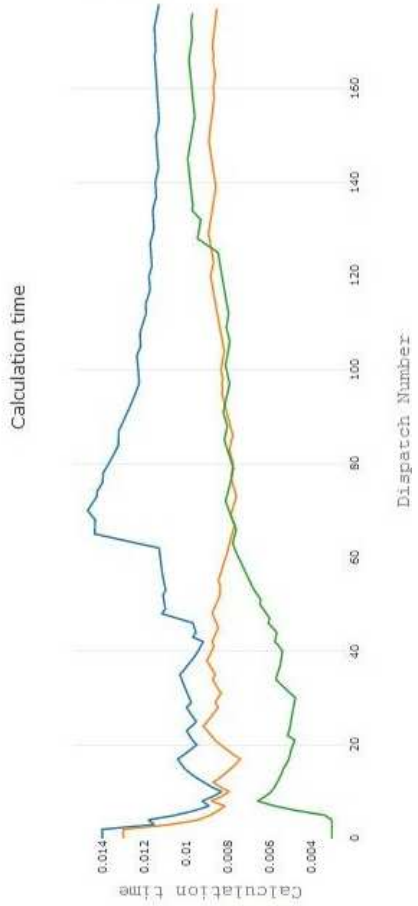

Fig. 8. Average Calculation Time Result