Simplifying Last Mile Delivery: A Simulation-Based Analysis for Warehouse Location Optimization in We-Doo's Operations

Yash Bhargava
Dept. of Computing
National College of Ireland
Dublin, Ireland
x2220@student.ncirl.ie

Abstract - A good warehouse selection is very important for minimizing operational costs and improving the satisfaction of the customer in logistics operations. In this research report, we have simulated to find the best and optimized warehouse location by implementing the simulation run for 20 days and 5 different warehouses and supporting the decision of the best warehouse location using the ANOVA test and investigating the delivery route used by cargo bikes driver to determine the best and optimal warehouse location. ANOVA was used to compare the average operational cost across the 5 different warehouse locations.

Index Terms: Warehouse, ANOVA, simulation, Monte Carlo optimization.

I. Introduction

Recently, there was an increasing growth in digitalization and shopping online, by which there was a huge number of parcels to be delivered to customers each day. China has become the largest number of express delivery parcels in the world [6]. Most logistics companies provide cargo transportation services using land, air, and sea to match the world for digitalization and growing economic patterns [1]. The most compatible delivery centre plays an important role by optimizing the delivery routes for the delivery boy to deliver the parcels to the customer's location. In the logistics business, there were very large numbers of parcels that needed to be delivered to the customer and it was also equally important for the logistics company to have a delivery centre at the best location so that the company could minimize their operational cost by following the best optimal route to deliver the parcels to the customers. In today's world, short trucks and big vans are mostly used for the transportation of and delivery of parcels due to this, there may be a huge amount of traffic on roads and very few cargo bikes are seen on the streets. The delivery of the parcels depends on the climate conditions. However, the complexity of the relationship between the operating cost of the warehouse and the delivery time makes it more difficult for the logistics company to deliver the parcels on time. A warehouse is where all the goods are manufactured and stored for transportation to delivery centres. Currently, supply chains compete in terms of delivery time and overall cost of the goods. Warehouse location selections influence capital investment, operational costs, and customer satisfaction. Warehousing has recently become one of the most critical contributors to a successful supply chain network. The warehouse's most suitable position helps supply chains succeed because of its low cost and increased revenue. The company must focus on selecting the most appropriate location among the options available for operating a warehouse. A warehouse was set up in a way that increases the overall effectiveness of the company's network, does not result in any problems in delivering goods, and increases the cost [5]. In this research report, we have evaluated the best delivery routes by simulating the run for 20 days for each of the 5 warehouses calculated the operating cost for each day for all 5 warehouses and conducted an ANOVA test to determine the best warehouse location for the delivery of the parcels. In the following report, the breakdown of this report is as follows: In the next section, we review some previous research work, in section 3 we go through the methodology of this research study, in section 4 we explore the findings of the research and interpretation of the results. In the final section, we have suggested some of the future work that needs to be done in future.

II. Literature Review

According to the research done by E. B. Tirkolaee et al [] in 2024 which was about the application of multimodal last-mile e-mobility concept by the modelling of travelling salesmen problem using drone and bicycle (TSP-DB) where they developed the linear programming model to minimize the travelling time. They have tried three vehicles truck, a drone, and a bicycle to serve the customers for last-mile delivery while the truck was used as the primary vehicle and the drone and bicycle were used as emergency vehicles. For investigating the complexity of this problem, validity, and applicability of the developed model where the dataset that consists of 64 different benchmarks was defined which resulted in accurately solving the problem for 50 customers in 685 seconds. Furthermore, the comparison was also made between the travel salesmen problem with drones and bicycles (TSP-D-B), the travel salesmen problem (TSP) and the travel salesmen problem with drones (TSP-D) in which TSP-D-B outperformed them. TSP-D-B have achieved the maximum time savings in most times.

Another research was done by Lingxiang Wei et al in 2024 which was about underground logistics systems (ULSs) that were low-carbon, energy-saving and ground-space-saving logistics solutions. The authors have designed a three-level last-mile delivery system (LDS) for the transfer of the parcels. The authors categorized the LDS into 2 modes i.e. self-pickup+ and home-entry+ and depending on the needs of the customers there were scenario-mode combinations. The authors analysed the five-scenario mode combinations and reduction of the fuel emission of LDS found out the solutions for last-mile delivery challenges and helped in planning and developing the ULSs.

Another research was done by Risald et al [] in 2018 which was about the best route selection to the nearest hospital using the Dijkstra and Floyd Warshall algorithm for victims of road accidents so that the number of deaths can be minimized. To find the best route to the nearest hospital the authors have combined Dijkstra's and Floyd Warshall's algorithms which means that the Dijkstra algorithm was used to find the fastest travel time and the other one was used to investigate the nearest distance to the hospital. The data of the hospital was used by Dijkstra's algorithm to calculate the fastest distance by the last traffic information and the Floyd Warshall algorithm was used to find the best route to the closest hospital. Finally, the author's aim for this research is to help the victims and the ambulance to find the best route to the closest hospital which improves the search time by reducing the time for decision making.

Another research was conducted by Florian W. et al [] in 2024 about reducing fuel emissions and improving last-mile delivery by utilizing rail freight. The questions about this research were how the hub located near the central station could reduce emissions and improve the shifts for rail freight as a pre-last mile. For the reduction of fuel emissions, the Stuttgart model was presented which was related to the Stuttgart 21 framework plan. The authors provided the solution for this problem by analysing different areas like trade, logistics and railway and based on that the role of the last mile should be decided to improve the logistics issue in Stuttgart. Another research was conducted by Yunjun Han et al which was about the optimal supply location and routing for the delivery of emergency material. In this research, the authors found it difficult to select the optimal supply location and the routing problem with warehouse selection and it was considered as an NP-Hard problem. The most important aspect of this research was the traffic jam, and this was solved by using the ILOG CPLEX system. After the testing of the numerical problem, it was clear that the cost of transportation and the demand were higher, and the number of commodities was higher. The complexity of the calculation made it more difficult with 60 nodes.

Finally, the research was conducted by Rajesh Kr. S. et al in 2018 was based in India where the auto components company wanted to set up their manufacturing unit in Iran. The author's purpose in this research study was to select the most optimal warehouse location in special economic zones and free trade zones. In this research, the authors used the Fuzzy AHP methodology to select the best possible four warehouse locations for an efficient and effective supply chain. In this study, the authors assigned each of the 4 locations with prioritised weights and based on that Salafchegan with the highest weight found to be the best location which was followed by Sirzan, Kish Island and the Port of Chabahar.

III. METHODOLOGY

In this section, we have explained the process of this simulation study, we have two parts in this simulation run i.e. Baseline Study and the second part of this methodology is the Simulation Study. For the development of the simulation run, we must first install the Simpy package and import it into our Jupyter Notebook to set up the environment for the simulation run. With the help of this package, we can use it for implementing real-world scenarios in our Notebook and can easily simulate it. As per the problem statement that was given to us, we must help the company find an optimal delivery centre location in the city to deliver parcels from the warehouse to the customer points such that the operating cost of the warehouse is optimized. The main components or classes of this simulation consist of "Recorder", "Parcel", "Customer", "Driver" and "Delivery Centre".

Firstly, we have a "Recorder" class which we are using for storing the data for each simulation run. There must be only one recorder file for each of the simulation runs of the warehouses which carries a reference to the "Recorder" class. In this class, I have defined a method "calculateDailyCost" in which I have assigned the constraints that were given in the problem such as operational_cost_per_km, pay_per_hour, and minimum_daily_pay are 0.08 cents, 30 euros, and 60 euros respectively. This method was used to calculate the daily operational cost for each of the candidate warehouses and store the operating cost of each warehouse in a data frame. This class records all the activities that take place during the simulation run i.e. delivery drivers working time, the length of the tour, and the number of parcels that were left over and calculated its statistics and plots based on that simulation. After the "Recorder" class, the next class was "Parcel" which was about the status of the parcels such that it follows the sequence of steps i.e. registration of the parcel in the warehouse, arrival of the parcel at the centre, out for delivery of parcel to the customer, whether the parcel was delivered to the customer or returned to the distribution centre.

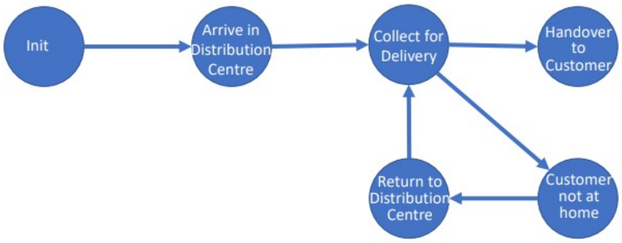


Fig. 1. Class Parcel

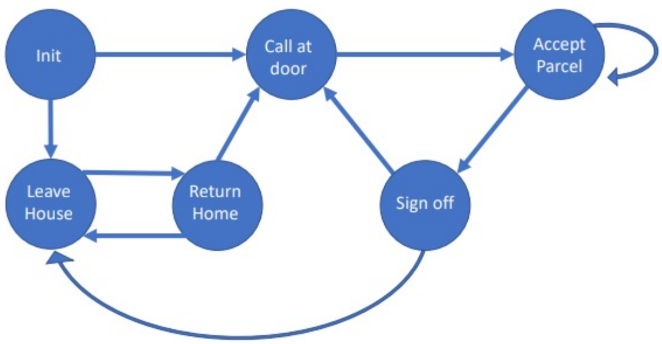


Fig. 2. Class Customers

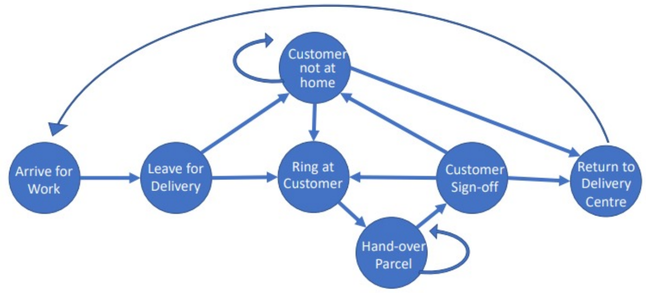


Fig. 3. Class Driver

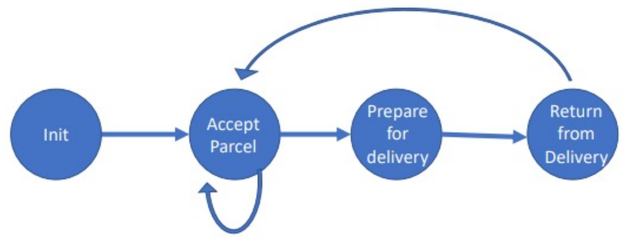


Fig. 4. Class Delivery Centre

The next class was "Customer" which was about the status of the customer in which we must first call the customer, whether the call was accepted or not, if accepted the parcel was received by the customer and if not received by the customer, then it should be signed off and the delivery boy leaves the house. The process followed in this class was about all the delivery steps from the delivery centre until the customer received the parcel. If not, it was returned to the delivery centre.

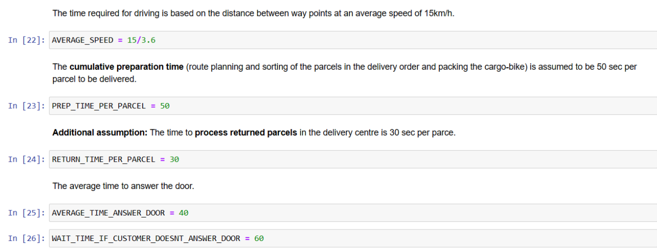
The next important class we used was "Driver" which was about the status of the driver such as the driver's arrival status for work, the driver left for delivery, delivery of the parcel status, return status of the parcel to the delivery centre and the arrival of the parcel back to the delivery centre. All this process was taken care of by the "Driver" class from leaving the parcel from the delivery centre by the driver to the customer or back to the return distribution centre.

Finally, the last important class was "Delivery Centre" which was about the status of the parcel when it was first received at the delivery centre by following the sequence such that accepting the parcel, status of the preparation of the parcel for delivery to the customer and if the parcel returns from the customer's location to the delivery centre again then the cycle continues, all this process was done using the methods that were defined in the delivery centre class.

A. Part 1: Baseline Simulation

To run a successful simulation, I have taken several constraints regarding the cargo bike such as the average speed of the cargo bike, preparation time per parcel, return time per parcel, the average time to answer the door and the wait time if the customer doesn't answer the door. To develop a simulation that can used for finding the optimal routes for the driver to deliver the parcels from different warehouse locations for the number of days (n=20). I have used my student ID (22220861) for this simulation as I have considered only my last four digits from the reverse side (1680) because for every different seed value the map generated for optimal delivery routes were different was different for every other seed value.

To find the best optimal route for the different number of days, the information that was provided in the document was the time required to cover the distance between at the average speed of 15 km/h, hence I set the AVERAGE_SPEED as 15/3.6, the preparation time was given as 50 seconds per parcel to be delivered, hence I have set the PREP_TIME_PER_PARCEL as 50, the time to process the per parcel when it was returned to the delivery centre, hence I have set the variable RETURN_TIME_PER_PARCEL to 30 and the average time required by the customer to answer the door was given us as 40 seconds, hence AVERAGE_TIME_ANSWER_DOOR as 40 and WAIT_TIME_IF_CUSTOMER_DOES NOT_ANSWER_THE_DOOR as 60.



The time required for driving is based on the distance between way points at an average speed of 15km/h.

```
In [22]: AVERAGE_SPEED = 15/3.6
```

The **cumulative preparation time** (route planning and sorting of the parcels in the delivery order and packing the cargo-bike) is assumed to be 50 sec per parcel to be delivered.

```
In [23]: PREP_TIME_PER_PARCEL = 50
```

Additional assumption: The time to **process returned parcels** in the delivery centre is 30 sec per parce.

```
In [24]: RETURN_TIME_PER_PARCEL = 30
```

The average time to answer the door.

```
In [25]: AVERAGE_TIME_ANSWER_DOOR = 40
```

```
In [26]: WAIT_TIME_IF_CUSTOMER_DOESNT_ANSWER_DOOR = 60
```

Fig. 5. Parameters from Specification

B. Simulation Study

In this part of the simulation study, the problem statement provided to us says to find the optimal best warehouse location among the various best warehouse candidate positions.

To do this, firstly I have chosen the number of warehouses I want to generate using the generateWarehouseLocation method as I have chosen the number of warehouses = 5. Then we simulated the delivery route for all the warehouses for the number of days = 20 and tried to find the operational cost of each warehouse for 20 days calculated the mean operational cost for each warehouse and performed the statistical test for the analysis to find the best warehouse location by comparing their operational cost and p-value and apply the Monte Carlo optimisation to find the best warehouse location among them.

IV. Result and Interpretation

In this section, after the simulation run was completed, we will explain the findings of the study and the decisions that we made for choosing the parameters that were used when we were running the simulation, we have calculated the daily working time vs density graph using the recorder class of each warehouse and their statistics such as mean and standard deviation which shows the histogram in which we can see the which shows the histogram was normally distributed for Warehouse 1 and the line chart for working time vs each day shows that as the number of days increasing the working time of the warehouse decreases as we can see the downward trend in the bar chart and as we can see all the other charts of the warehouse were following almost the same pattern as we can see in Fig

which also shows that there is not much difference among all 5 warehouse positions we have further calculated the operational cost for each warehouse for 20 days and visualized it on the line chart as we can see in Fig
This line chart shows the operational cost for each warehouse for 20 days we can see that the operational cost for each warehouse was found to be significantly low on Day 14 and Day 19 as there are no parcels were left on that day also the operational cost was very less as compared to other days which also shows that there is no significant difference between those warehouses and to support that we have done an ANOVA test and found the f-statistic= 0.65 and p-value= 0.62 which shows that there were no significant differences between the operational cost of the 5 warehouses, because the p-value of this test

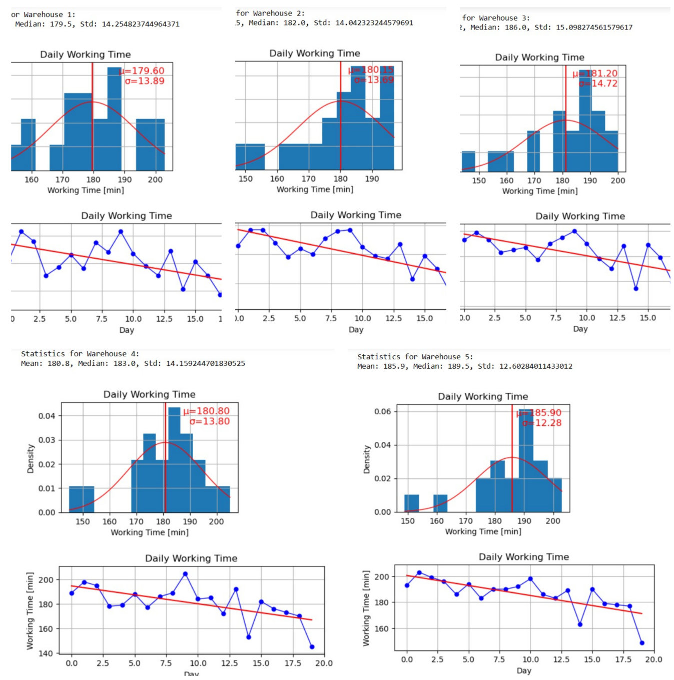


Fig. 6. Daily Working Time

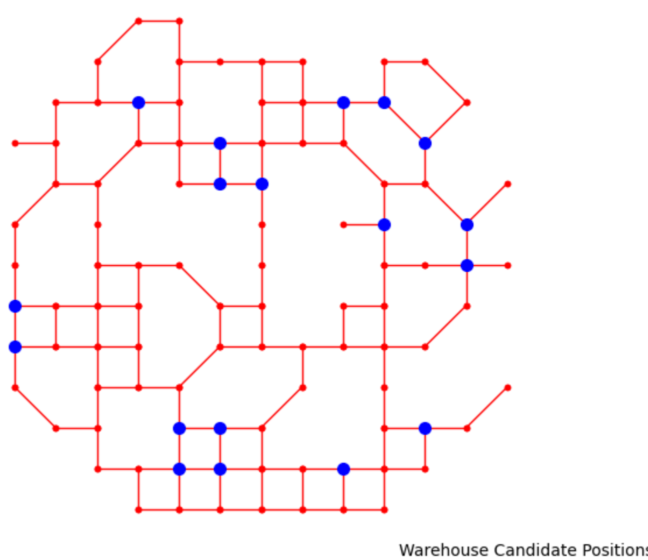


Fig. 7. Warehouse Candidate Position

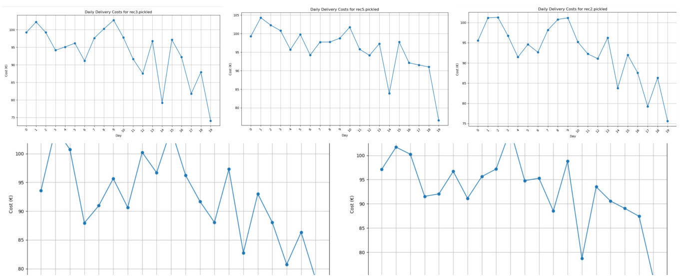


Fig. 8. Daily Delivery Costs

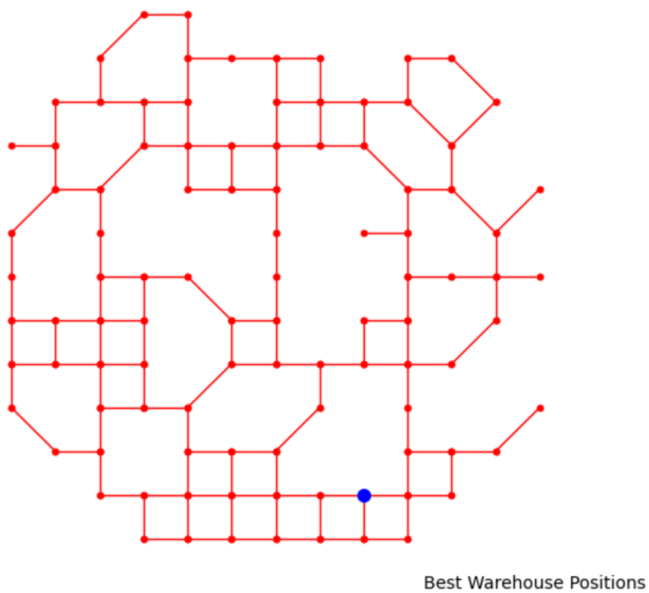


Fig. 9. Best Warehouse Position

was greater than the threshold value of 0.05. Finally, using the generateWarehouseLocation method, we have created a map of our simulations and plotted the warehouse candidate positions on the map then using the Monte Carlo optimisation algorithm we found the best Warehouse location among them as we can see in Fig and Fig.

V. Reflection and Future Work

Based on the research we have done so far; we have taken the number of warehouses = 5 and rum the simulation for 20 days for each warehouse. The simulation run was done to determine the best location for the warehouse so that the operational cost for the warehouse would be minimised. Hence, with this, we have the best warehouse with an average operational cost of €92.37 Warehouse 1 which is the best warehouse location. After all the simulation runs were complete then we conducted the ANOVA test on the operational cost for each of the warehouses and based on the results of the test we found that there was no significant difference between the 5 warehouse locations as the p-value was greater than the threshold i.e. 0.05 which shows that there was very less difference between the operational cost among the warehouses. Based on the above considerations, we have determined the best warehouse location for this problem which results in minimizing the operating cost of the warehouse for each day in future we can simulate different numbers of warehouses and can run simulations for a greater number of days so that we can further minimize the operational cost of the best warehouse.

```
ANOVA Test Results:
F-Statistic: 0.6537628196938253, P-Value: 0.6256440911153279
There is no significant difference between the warehouses.
```

Fig. 10. Annove Test Results

References

[1] E. Babaee Tirkolaee, E. Cakmak, and S. Karadayi-Usta, "Traveling salesman problem with drone and bicycle: multimodal last-mile e-mobility," International Transactions in Operational Research, Mar. 2024, Published, doi: 10.1111/itor.13452.

[2]] L. Wei, Y. Chen, D. Guo, J. Ji, Z. Chen, and C. Zhuo, "A last-mile delivery system for underground logistics with 'self-pickup +' and 'home-entry +' modes," Tunnelling and Underground Space Technology, vol. 147, p. 105678, May 2024, doi: 10.1016/j.tust.2024.105678.

[3] "Best routes selection using Dijkstra and Floyd-Warshall algorithm," IEEE Conference Publication — IEEE Xplore, Oct. 01, 2017.

[4] F. Wondratschek, V. Gangadharan, and N. Wilson, "How might a City Hub at Stuttgart Central Station reduce Emissions and improve the Pre-last Mile Delivery by Rail Freight?" Jan. 09, 2024. https://www.scienceopen.com/hosted-document?doi=10.14293/PR2199.000632.v1

[5] "Optimal Supply Location Selection and Routing for Emergency Material Delivery," IEEE Conference Publication — IEEE Xplore, Sep. 01, 2007. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=4341789

[6] R. K. Singh, N. Chaudhary, and N. Saxena, "Selection of warehouse location for a global supply chain: A case study," IIMB Management Review, vol. 30, no. 4, pp. 343–356, Dec. 2018, doi: 10.1016/j.iimb.2018.08.009.