Modelling, Simulation and Optimisation CA

1st Steven Travers - X23299525
NCIRL Modelling, Simulation & Optimisation CA
Postgraduate Diploma in Science in Data Analytics (PGDDA)

***Abstract*—This project aims to use various modelling and simulation techniques to identify the best possible warehouse location to minimise costs of last mile parcel delivery for We-Doo. Various models were generated and vast simulations were done with different test parameters to identify possible errors in the model. The Chinese Postman Problem algorithm was used to find the shortest possible delivery route where computer resources allowed the calculation. The final model should be the cheapest cost per parcel which satisfy the restraints laid out in the project description. Future work will outline possibilities to improve the models and various areas that were not accounted for in the modelling process, although these would be essential to providing We-Doo with the most accurate scenario and simulations possible.**

***Index Terms*—Traffic Modelling, Forecasting**

I. Introduction

A. We-Doo & Problem description

'We-Doo' is a start up company that operates a parcel delivery service to a local area. In recent months, there has been rapid expansion. This has led to the decision for the company to change its model and introduce a new warehouse for local areas. This new model is being piloted with one town. We-Doo have allocated a budget for this expansion project and the aim of this discussion is to identify which warehouse location, from a selection of 5 locations, is best placed to minimise costs and maximise efficiency for this expansion. To do this, extensive simulations will be ran on different models to test various ideas. These simulations will help to identify the warehouse location, and model, that is best suited for *We-Doo's* expansion plans. We-Doo have also outlined some key constraints to consider while generating the models. Some of constraints are as follows:

- Maximum cargo bike range and average speed
- Parcel preparation, handover and delivery times
- Operational costs and wages
- 10% of customers on average not being at home to accept the parcel

Throughout the project, models will be generated both following and not following these constraints. Where the models do not follow these constraints, the logic of the derivation of the model will be discussed.

II. Literature Review

In order to generate ideas for new models, several sources of information were examined. The Chinese Postman Problem was used as the main source of route finding algorithms for this project. The code needed for this algorithm was generated in lectures **??**. Two key last mile delivery operators were examined for this project and this examination led to the development of model 2 and 3 below.

A. Delivery Companies

Brief discussions with local delivery drivers for DPD & Amazon were had to try to understand their current model and potential areas to reduce bottlenecks in their workflow. There were some differences noted between these delivery companies. For example, with DPD, customers were contacted via email/phone the morning of their delivery with an ETA(estimated time of arrival). This is a multi purpose approach that benefits both the customer and the driver. Customers were benefit due to the fact that they now have a, in general, 2 hour slot for their delivery. They know to expect a knock on the door / phone call within that time. This then in turn benefits the driver as there is less of a wait for the driver to deliver the parcel. Amazons model differs slightly from this, with the Amazon app (and by email if required), a notification is sent when your parcel is out for delivery. Although there is not a specific time given, you have the option to track the delivery when the driver is several stops away. As before, this reduces the time the driver is waiting at the door for customer to answer.

B. Amazon Picture on Delivery

An area of interest which was identified is an Irish Times article that States that in 2019 just 7% of the Irish workforce usually worked from home. This has increased to 25% in 2022 [1]. The reason for highlighting this is, in the models produced for this project, it would be beneficial to show how the model compares with the scenario of reduced waiting time for driver at customer location. To do this, the idea of picture on delivery will be used. This was introduced by Amazon c2018, with the idea of the driver placing the delivery in a *secure* location and taking a picture of the delivery to send to the customer. This is the basis of model 2. The premise of this model is that the bottleneck for drivers waiting for the customer to answer the door is removed. The new model will show the driver placing the delivery, knocking on the door and taking a picture as they leave. This reduces the driver wait time significantly. It **MUST** be noted that we are not modelling for customer satisfaction and realistically this should be taken into account. This will be discussed in Future Work.

III. Methodology

A. Code Generation

The majority of the code used for the model generation was given throughout lectures. Variations in the code includes the generation of parameters for model 2 & 3. The other main difference is the map used and warehouse location generation. The parameters, which will be discussed in their relevant sections, were derived from the literature review. Jupiter Notebooks was used so the code could be separated for each of the models discussed.

B. Map Creation

There were 3 map variants generated for this project. They can be described by Table I where Customers is the count of customers and Nodes is the count of nodes.

Map	Seed	Customers	Nodes
SimpleData	0	5	20
TestData	9525	20	35
ProjectData	9525	100	35

TABLE I
VARIOUS MAP DETAILS

SimpleData was used to very quickly test ideas and ensure code was running correctly. Since simulations and path-finding algorithms are very resource intensive, TestData was used with the final map configuration and reduced customer count to reduce the run time of some ideas to be used. The ProjectData was used to ensure statistical tests could be used to see how beneficial models were and if there was statistical significance between models. The graph for the Project data map including customers can be seen in Fig. 1. Test and simple data can both be found in accompanying code but due to similarity of project and test data it is excluded from this report.

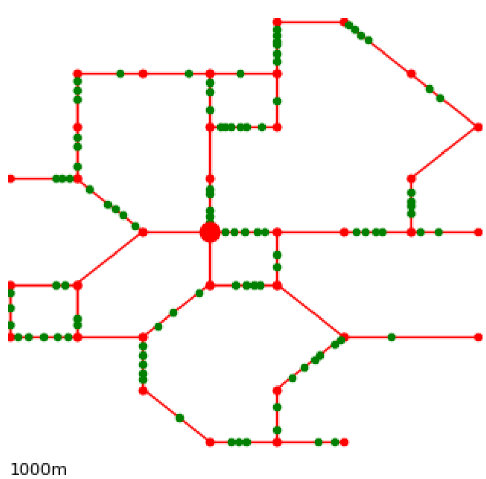

Fig. 1. Project Data map including customers and central warehouse location

C. Candidate Warehouse Locations

The code provided generates a set of N candidate warehouses. For this project, five potential warehouse locations were considered. The list of options can be seen in Fig. 2

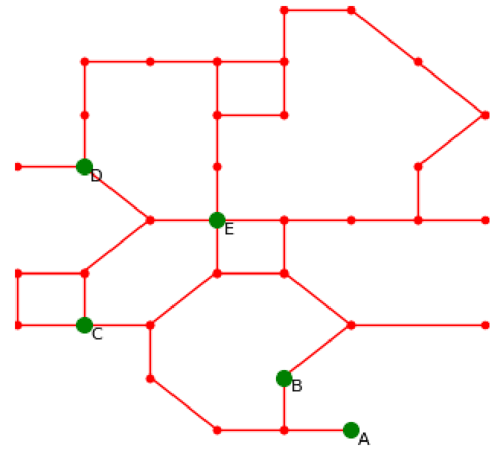

Fig. 2. Initial Candidate Warehouse Locations

Of the warehouse locations seen in Fig. 2 and based off initial inspection, it is assumed that warehouse in location E will be the cheapest. This hypothesis is based off the logic that this warehouse is the most central of all locations & in theory, should reach all customer locations in the smallest average distance. It must be noted that when testing with small sample size of customers, it is possible majority of customers are in the so called “Bottom Half” of the map. To reduce the risk of this occurring, it is imperative that the customer count is high enough to ensure best random spread of customers throughout the map.

D. Model 1 - Basic

The first model that will be tested is the standard model with no variations (excluding seed set up) in the initial code. This is used for a benchmark for which to aim to improve on. This model uses the Chinese Postman Problem (CPP) to find the shortest delivery route for each day.

Model 1a - One Driver. Using the CPP code for the project data, given no distance constraint and only one driver, there was a feasible solution. However, with one driver, covering the whole map of 100 customers is extremely difficult. Once the 40km constraint was added to the CPP code, the solution was infeasible. For this reason, and looking at the map, a conclusion was made to stick to at least two drivers for this project. This can be seen in block 70 of “Model1_Code.ipynb”.

Model 1b - Multiple Drivers.The new benchmark for this project will now be derived using the initial constraints in the code provided, with 2 drivers for CPP.

E. Model 2 - Picture on delivery

For the Picture on Delivery model, there were several parameters that needed to be updated in the code. The premise of this model is that the driver no longer stops to wait for the customer to come to the door after knocking as is the case with Fig. 3.

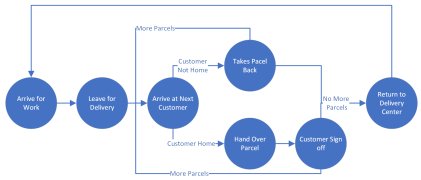

Fig. 3. Manually Created Event Graph for Delivery in Model 1

After knocking, the driver now takes a picture of the parcel in a secure location and leaves. This is a simplified model which drastically reduces driver wait time during deliveries. The event graph can be seen in Fig 4

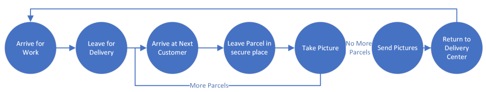

Fig. 4. Manually Created Streamlined Event Graph for Model 2

With the simplified event graph for parcel delivery, the parameters in Table II required updating. Some parameters were set to zero as they are no longer required, while others were increased such as Day end, to allow for additional time to send pictures manually on days end.

Parameter	Before	After
CUSTOMER_NOT_AT_HOME	10%	0%
AVE_TIME_ANSWER_DOOR	40	0
WAIT_TIME_IF_CUST_DOESNT_ANSWER_DOOR	50	0
AVG_HANDOVER_TIME	10	20
AVE_SIGN_OFF_TIME	10	0
DAY_END_PROCEDURE	600	1000

TABLE II
PARAMETER DEVIATIONS FOR MODEL 2

There are several expected outcomes from the updated simulation. Prior to running these simulations, they were noted. These outcomes include:

1) Delivery costs are reduced as driver bottleneck of waiting for customers is gone.
2) Route times are reduced for same reason.
3) Number of parcels left over per day should be significantly smaller. Potentially not Zero as there is still the 40km limit, thus making the system balanced.

F. Model 3 - Deliveries at Night Time

The final model which will be tested during this project is the possibility of doing deliveries at evening/ night time. The idea of pictures taken still applies, so the event graph doesn't not change significantly, except the driver does not knock if its too late. The reason for this simulation is to try to identify what has more of an impact on the price per delivery. This model aims to test the increase in average speed and the max bike distance due to efficiencies and reduced traffic. The changes to the parameters for this simulation can be seen in Table III

Parameter	Before	After
Shift Allowance Factor	NA	1.3x
Max_Distance	40km	45km
AVE_SPEED	15kmh	17.5kmh

TABLE III
SIMULATION MODEL CHANGES

To try to make the model more realistic, a shift allowance factor was applied, this was set at 1.3x. Another brief test that was done was with the reduced wait times, speed increase and increased distance, was one driver now feasible. This test failed as Plan returned infeasible solution. This can be seen in block [331] of model2_code.ipynb.

G. Analysis

Once the various models above were generated, the next stage was to run simulations. For this, there were a variety of parameter combinations tested. There are multiple different routes possible when generating new parameters. For example, it is possible to up the count of days, the number of drivers, parcels per day, customer counts, etc. For this project, it was decided that the following parameters would be tested:

- P - Average Parcels per customer per day.
- $Balance$ - The overlap for multiple drivers where applicable.
- N - Number of drivers.

For each of these parameters, a variety of value combinations were tested. Due to the resource and time constraints, more parameters could not be tested. These will be discussed in Future Work. A Python dictionary was utilised to loop through all the various combinations of parameters. This can be seen in "Model1_code" block 74.

The analysis stage was broken into 2 stages. The first stage involved running the simulations with a time limit of 200 seconds for the CPP to find the optimal route. If the optimal route was not found, a sub-optimal route may have been used. This is important as the methods of evaluating these models and parameters may have been hindered by a sub-optimal route. There are several factors that can influence this sub-optimal route, the main issue stems from resource constraints such as processing power & time constraints. Unfortunately. this is something that for this project could not be avoided. It would be possible to run all these simulations with infinite time, but due to constraints, a 200 second time limit was applied.

Code block 75 of "Model1_code" shows the simulations. As can be noted, the results of each of these simulations are stored in a dataframe which is later exported to excel. The purpose of this is to have a convenient way of visualising all of the different combinations quickly. When this relevant block was ran for each model, the results were briefly analysed in to sort by the smallest delivery cost. The exported excels can be found in the "datafolder".

For each model, only the simulation which provided the cheapest mean cost per parcel was used for future analysis. This is where there is potentially issues with sub-optimal routes being ignored. As one route could be sub-optimal but would benefit from a longer run time and could potentially have a lower mean cost per parcel. An example of the evaluation metrics produced in the table include:

- Plan Key, Warehouse location, Number of drivers, balance, P
- Mean cost per parcel, Mean daily cost and mean working time.

Once each of these simulations were ran and the top simulation for each model was identified, these were used to proceed for some further analysis. The next stage of the analysis included a more in depth simulation run. This is the last code block in each file which increases the simulation time to ensure that the optimal route was found. Brief results are found in Table IV The main difference is the length of the time limit to finding the optimal route and the number of days. For the first set of simulations, the days parameter was set to 30. For the in depth simulation, this was increased to 365 to represent one full year.

Model	Warehouse	Ave Cost PP	Ave. Working Time
1	E	€6.68	302.5
2	E	€6.05	275
3	D	€6.98	239.2

TABLE IV
MODEL PERFORMANCE COMPARISON

Interesting points from Table IV is the reduction in working time for both model 2 & 3, it should be noted that if these are possible models, there is a clear reduction in working time which could lead to driver satisfaction. If this is something *We-Doo* want to take into consideration.

The final stage of the analysis is ensuring the model that is chosen is balanced and analysing various graphs to ensure that there are no obvious issues with the model. This also ensures that the initial assumption of warehouses is correct / the logic was flawed or proven incorrect.

IV. RESULTS AND INTERPRETATION

The initial hypothesis for this project when viewing the candidate warehouses is that the warehouse that will be the cheapest is Warehouse E. As explained previously, this was because this warehouse is most central. After running the vast simulations for each model and warehouse on the smaller time limit, there were some interesting results. Warehouse A, B & C were by far the most expensive. This would be expected. However, an area of surprise was how competitive Warehouse D was for average cost per parcel. For Model 1 & 2, there was €0.10 difference in warehouse D & E. For Model 3, which is the nighttime deliveries, Warehouse D was the cheapest.

Although this is not a complete shock, since both warehouses are both close by, it was an interesting find. From Table IV, it can be noted that the cheapest simulation was for model 2 and warehouse E. The driver routes for this model can be seen in Fig.6 and Fig. for each respective driver.

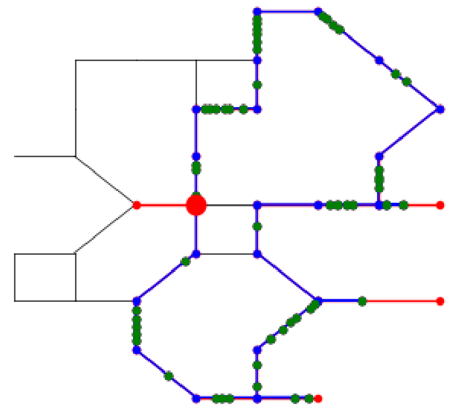

Fig. 5. CPP route for Driver 0 - Model 2

It can be noted that for both figures, the maximum distance is less than 40km. This might lead to the hypothesis that this is not a key factor in the model. It would be nice to further examine this with more time.

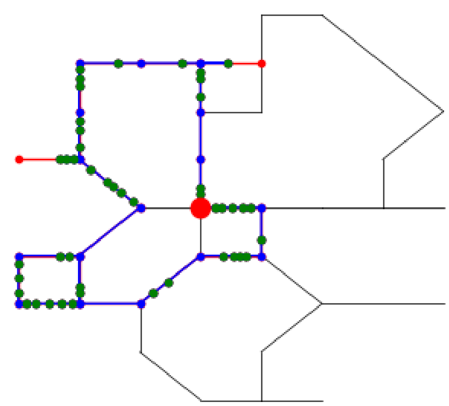

Fig. 6. CPP route for Driver 1 - Model 2

It may also be noted that customers are fairly evenly split between both drivers. With Driver 1 having just 8 more customers. The next important aspect of this model is to ensure that it is balanced. This means that the model can handle the parcels over a long period of time. To do this, there are 3 key graphs to consider, the parcels received, delivered and the Arrival/Delivery plots.

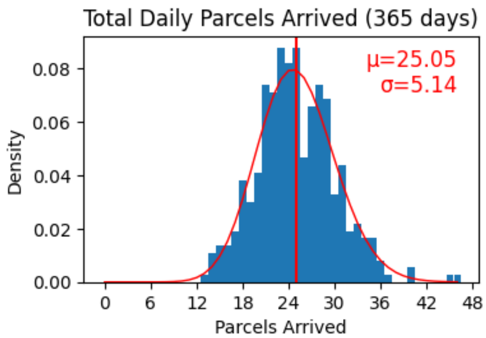

Fig. 7. Histogram of Parcels Arrived to Delivery Center

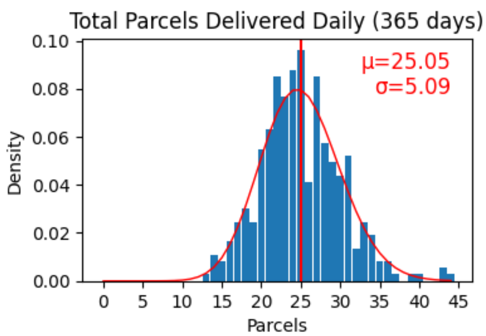

Fig. 8. Histogram of Parcels Delivered

Fig. 7 and 8 represent the parcels arriving to the delivery center and also the parcels delivered daily. The point to note is that both means are the same, meaning the average arrival and delivery rate daily is the same. This would make sense as in model 2, no / very little parcels are undelivered. This is the benefit of the "Picture on Delivery" Model.

Another key area for consideration when discussing model options with *We-Doo* is the working time. With a 3 hour constraint, it would be beneficial to the client to minimise the amount of days in which drivers are doing less than 2 hours work. This is due to the fact that there is a 2 hour flat rate. To visualise the average working time, a histogram is used. The results of which can be seen in Fig. 9

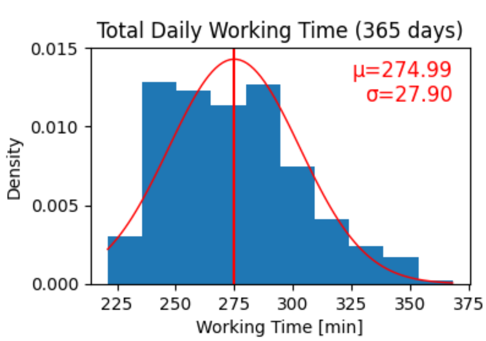

Fig. 9. Average Working Time Model 2

The mean working time for 2 drivers in model 2 for the given simulation is c280 minutes. This is c 140 minutes / above 2 hours for each driver which is excellent for *We-Doo*.

Concluding on the analysis, it can be noted that model 2 provides the lowest cost per parcel when using warehouse E. This is both a balanced model that meets *most* of the constraints. If there were discussions with the client, this model would be a viable option if they were willing to alter their model.

The biggest bottleneck in reducing costs for last mile delivery that's possible to change comes from the time waiting at doors etc for the driver. If the client was to move to "Picture on Delivery" model, this reduces the average cost per parcel by nearly 65c.

However, when explaining the results to *We-Doo*, it should be stressed that the above models mainly focus on reducing the cost. However, there should be many other factors taken into consideration for this project. An example is customer satisfaction. In the above models, the event of leaving a delivery at the door was tested. Although this is done in practice, there should be an option to exclude this from parcels say over a certain threshold. This reduces the possibility of theft. To keep this model simple, this scenario was ignored. However, with more time and programming knowledge, this should be explored prior to approaching *We-Doo*.

V. Reflections and Future Work

A. *Reflections*

Reflecting on this project, it was interesting to learn about several different route finding algorithms and how to implement them in python. Having the event graphs at the beginning helped to make understanding the code easier. Although the scope of this project is not completely code focused, with an increased ability in python, some of the possibilities to enhance this code, as will be discussed in Future work would have been possible.

One of the most interesting pieces of this project is the ability to change parameters in the code and see the impact that this has on working hours and costs. Slightly changing some parameters and rerunning the same code can give quite different results.

Reflecting on the possibility of including other parameters, a parameter that would have been interesting is a daily cost for the warehouse. An assumption that was made on this project were that all candidate warehouses are of equal cost. However, if we assume that the map generated using the seed is a town, it is very likely that a warehouse in the middle of the town will be far more expensive to rent than a warehouse on the outskirt. This would make selecting the relevant warehouse more complex.

With some more time, another aspect that would be worth adding is how finding the equation for how many parcels the client should deliver for each relevant warehouse to be most optimal, taking into account the cost of the warehouse.

In hindsight, this project would have benefit drastically by adding some more parameter changes, like increasing the customer count and the p parameter. It was interesting that the length of time the code took to run increased drastically and in some cases the simulations took upwards of 20 minutes. This restraint made complex simulations difficult.

B. Future Work

There are several possibilities for future work with this project, some of this includes adding more parameters to the code and testing several different models. An area that would be of potential benefit to *We-Doo* is trying to future proof these models.

As mentioned in an article on CapitalOneShopping, "From 2017 to 2022, the number of packages the average American received in a year increased by 73%." [2]. This is interesting as there would be merit in including some sort of forecasts of the value for the parameter P, the average number of parcels per day, based on these statistics. The growing presence of Amazon and websites such as Temu with cheap delivery it amplifying the growth of home deliveries throughout the world. It would be interesting to see what potential models parameters work best when the value for P increases significantly, which might help to plan for *We-Doo's* future. Another key parameter that would be worth future proofing is the max possible daily distance travelled. With improvements in Battery powered vehicles, how will this impact the parcel deliveries and costs with increased range. Another area to consider when future proofing is the change in cost of charging the bike and wages. How can *We-Doo* maximise the efficiency of the cargo bikes?

Another area that could be explored with some more time to research is customer behaviour. In this project, the main focus was to minimise the cost for the client. Although in an article published by McKinsey, c30% of people would pay an increased delivery charge for faster delivery [3]. The remaining 70% are content with having the cheapest form of delivery [3]. Following on from this, it would be useful to have another option for customers of *We-Doo* for them to select their method of delivery. If they would like faster delivery these get prioritised daily. Due to the time constraints of this project, it was not possible to add such logic into the code. It would be interesting to test several models using the above information such as is there a benefit to hiring a driver purely for next day / faster deliveries? These could also take the parcels that were not delivered on the previous day too potentially.

With some further programming knowledge, it would be interesting to see if it is possible to find an optimal group of parameters that minimise the cost for deliveries. This could be done by implementing some further choice of parameters and utilising machine learning methods. This information could provide useful analysis to *We-Doo* such as if they reduce driver wait time by x%, this will lead to a reduction of €y per delivery. Another parameter worth looking at is the driver salary. The histogram plots in the code show that there are some days in which the drivers work less than 2 hours. This is sub-optimal as the minimum salary per day is based off 2 hours work. Although some of the models segregate the map into areas, how would this change if at first, the maps are segregated, but when a driver is finished their route, can they help the other driver?

These are all possible factors to take into consideration before providing a definitive solution to *We-doo*. Although the model that was agreed in the results section is the best model that was found, there is nearly a guarantee that there are better models to be found. This must be stressed to the client when providing the conclusion of this paper as small changes in some parameters may lead to a better more optimal solution.

VI. Acknowledgement

A quick thanks to Christian Horn for the help in creation of code for this assignment [4]. Code was easy to follow and adapt to generate our own models.

VII. Other

A link to the github file to show own code adaptations can be found here: https://github.com/StevenTravers-97/Modelling–Simulation-and-Optimisation

References

[1] C. Gleeson, "Remote working: A quarter of ireland's workers now operate from home most of the time," Aug 2023. [Online]. Available: https://www.irishtimes.com/business/work/2023/08/08/ireland-adopting-hybrid-working-at-greatest-rate-in-eu/

[2] Feb 2024. [Online]. Available: https://capitaloneshopping.com/research/package-delivery-statistics/

[3] M. Joerss, F. Neuhaus, and J. Schrouml;der, "How customer demands are reshaping last-mile delivery," Oct 2016. [Online]. Available: https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/how-customer-demands-are-reshaping-last-mile-delivery

[4] C. Horn, "Christian horn mso class notes." [Online]. Available: https://moodle2023.ncirl.ie/course/view.php?id=1999