Enhance Coppel's last mile delivery problem using a mean speed approach

Felipe Villaseñor Navarrete A01023976, Rodrigo Morales Aguayo A016324834, Rodrigo Hiroshi Argandar Nishizawa A01630005, Kaled Corona-Romero A01635604
Tecnologico de Monterrey, Escuela de Ingenieria y Ciencias, Campus Guadalajara, Ave. General Ramón Corona 2514, 45138 Zapopan, Jal., México

***Abstract*—Last year, e-commerce increased spectacularly due to the pandemic we are living in nowadays. Because of this, many stores did not have enough infrastructure or capacitation to satisfy the increasing demand on time. This article exposes the challenges of solving a Capacitated Vehicle Routing Problem and also an optimization of the delivery routes of the Mexican department store Coppel for a given day. The approach used to solve the problem consists of obtaining a traveling time matrix by dividing the distance matrix by the average speed of motor vehicles in the metropolitan area of Guadalajara.**

***Index Terms*—E-commerce, Coppel, CVRP, optimization, model distribution planning.**

I. Introduction

E-commerce arrived to Mexico and it came to stay. It earned a special place among Mexicans due to the comfort of being able to purchase an immense variety of products and services from any electronic device with internet connection. And if that was not enough, the lockdown resulting from the COVID-19 pandemic boosted e-commerce in Mexico at levels never seen before.

Such impact was reflected in an 81% increase in online sales during 2020, specially in food delivery, fashion, beauty and personal care products [1]. Online sales even reached 500% by mid April 2020 according to a study carried out by the consulting firm Kantar [2].

Consequently, the business had to adapt in order to satisfy the increasing demand. However, this isn't an easy task considering the vast amount of orders that must be shipped daily, the size and weight of each package and the capacity of the vehicles, the handling requirements of the products, the distance between the delivery points and the roads the vehicles may use.

There are indeed too many ways to deliver the orders of a given day. However, the deliveries must be done in the most efficient possible way. In the logistics world, this last part consisting of delivering the product to the final client is referred as the last mile delivery problem.

One of the most important points about delivery systems is the quality and effectiveness about it, primary on the process control. We take control as the sum of the systems and procedures for controlling the work flow and the utlization of capacity resources in order to meet specific performance standards.

The service provider continuously monitors, evaluates and refines the service delivery process in order to make it more effective, more cost-efficient and more customer-driven. Therefore, the control process continuously improves the service delivery and upgrades the quality of the service provided to customers [3].

Since they are inseparable, any improvement in the process will make a positive result on efectiveness and efficiency. Companies make certain structured questionnaires, giving it to customers and managers, each with different questions about the delivery system. Depending on the answers, they measure the variables to measure effectiveness, using differente scales.

One of the biggest challenges on delivery system is ensuring the quality of the delivery itself, first and probably the biggest challenge is the customer expectations, that´s why it´s important to talk to customers, to find out their expectations. After talking and knowing what they expect about a company, you have to exceed does expectations, customers often have new expectations after trying for the first time, that´s why companies have to step ahead of them, also to gain leverage over competitors.

Although, the consistency is probably the most difficult challenge, having great customer service all around each and every day, and each and every experience must be equally good on each sell. What affects customer satisfaction are cumulative experiences across multiple touchpoints and in multiple channels over time. Consistency can be achieved by having clear-cut policies on delivering great customer service. All employees must be awared of these policies and strict implementation of these policies must be attached to the long-term solution to achieve consistency across all business channels is to install a customer-first company culture.

Finally, a customer-driven company culture can be described as having a distinct appreciation for customer service, understanding the impact of their service and are willing to go out of their way to meet customer needs. In order to cultivate this culture, it must be emphasized during the hiring and training process. The employees must be able to see the company vision, fit into the culture and work with the company to achieve the company's goals.

Employee retention is also a key factor to success. Staff turnover not only affects the internal aspects of the business, it is also not good for the customers. Dealing with the same person develops trust and is a good way to build long lasting relationships.

In the last year, the Mexican department store Coppel S.A. of C.V., also had a notorious increase in online sales, increasing it 7 times. [4]. Coppel is the greatest chain store of department stores in Mexico, with 1,450 stores. It is an omnichannel store that has the objetive of giving the customers a great buying experience. To accomplish this, Coppel can deliver all the orders in time and maximize its profits with its current sales, route optimization is essential to minimize costs.

II. Problem description

The Last Mile Problem (LMP), is the final part of the delivery chain and the most important part because it's when an item gets transported to its final destination. Normally, customers want speed and a minuscule fee in this process, however, this problem happens to be the most complicated and expensive part of delivery logistics, which is quite the problem online retailers are trying to solve. The first step to solve this is to understand the Traveling Salesman Problem (TSP), which is the attempt to find the route that visits all the planned places (nodes), returns to the origin while it minimizes costs, whether it is time, carbon footprint or traveled distance.

To illustrate how deliveries can be optimized, we are going to take Coppel's deliveries from January 16th, 2021 to Guadalajara and Zapopan and we are going to plan its vehicles routes so they deliver all the orders traveling as little distance as possible.

In order to optimize Coppel's last mile delivery, a modification of the TSP called Capacitated Vehicle Routing Problem (CVPR) is required. This problem also minimizes the costs of visiting all nodes but considers having k available vehicles for the deliveries, each one with its one capacity, and the vehicles must return to the origin node after making the deliveries. In this case, the origin node is the Distribution Center (CD), also referred to as the node 0. Naturally, solving a CVPR entails a series of challenges that make it a complex problem.

One of the major challange of solving a CVRP, is the increase in the number of nodes someone has to encompass, in this case, e-commerce. The following arguments are why TSP is so challenging to solve:

- Traffic congestion
- Sudden change of routes
- Last minute order updates and requests
- Vehicle issues (low fuel, blew tire)
- Accurate delivery window timings

It is nearly impossible to solve CVRP with even the best manual efforts or algorithms, beacuse there is might be data that is not accurate and that can change in a matter of seconds [5]. But, with the rise of technology, new techniques have been implemented such as heuristics that simplifies the problem and helps to reach the primary objectives of CVRP, as follows:

- Optimized decisions for each vehicle and route
- Save fuel and labor costs
- Recognize the right addresses
- Making delivery turnaround time shorter
- Reducing cost per mile
- Improved productivity

In order to model the situation, the decision variables and the parameters to use must be defined. In this case, the decision variables are x_{ijk} and u_i, where x_{ijk} is a binary variable that takes the value 1 when the node j is visited after the node i (each node represents a delivery point) by the vehicle k and it is 0 otherwise. The variable u_i is an integer used to assure every vehicle starts and ends in the node 0.

On the other hand, the data provided by Coppel that is required to solve the problem is the demand of each client (the volume of its package) and the number of vehicles and their capacities. Any required missing data will lead into assumptions, like the capacity of the vehicles.

III. Methodology

In order to implement the model, some modifications were made to the original Coppel's database. First of all, the orders from days other than January 16 2021 were excluded. Furthermore, pick-ups and repairs were also deleted so that the database only includes deliveries. Finally, the repeated delivery addresses were also removed and its orders' volumes were added, resulting in 181 delivery locations.

The next step to solve the problem was to obtain the cost matrix. Assuming that Coppel desires to minimize traveling time, our cost matrix must be a traveling time matrix. However, since the traveling time varies tremendously depending on the departure time, the day of the week, the time of the year, whether there has been an accident in the route or not and the vehicle speed; the traveling time distance was obtained indirectly.

The Google Maps API was used to calculate the distance between all the nodes (by using the streets instead of calculating the euclidean distance) resulting in a distance matrix, and it was divided by a slightly higher speed than the last registered average speed of motor vehicles in the metropolitan area of Guadalajara (23.5 km/h in 2018) [6].

It was decided to use an average speed higher than the average (35 km/h) because the average speed was calculated considering a displacement of 24 km in average [6] instead of the traveled distance. It was also decided not to use a higher speed because of the following reasons: The last registered data of the average speed is outdated and it might have changed, to have the worst-case scenario covered (like a day with a lot of traffic) and considering that delivery vehicles might be slower than the average vehicle. The speed was used in km/min to have the objective function in minutes.

The model also considers the service time in each node in function of its order volume. Considering that the packages are only delivered at the entrance and sometimes they are brought inside the building (they are not brought to another floor nor they are assembled or installed) and using the biggest packages as reference (a sofa and a mattress), the service time was estimated and the next function was built:

$$Service\ time_i = \sqrt{20d_i} + 4$$

where d_i is the order volume of the node i in m^3

The delivery locations for the given day (January 16th, 2021) are shown in the Fig.1.

IV. Model

$$Min\ Z = \sum_i \sum_j \sum_k c_{ijk} x_{ijk} \tag{1}$$

s.t.

$$\sum_k \sum_j x_{ijk} = 1 \quad \forall i, \tag{2}$$

$$\sum_k \sum_i x_{ijk} = 1 \quad \forall j, \tag{3}$$

$$u_i - u_j + n\, x_{ijk} \leq n - 1 \quad \forall i, \forall j, \forall i \neq j, \forall k \tag{4}$$

$$\sum_j x_{ljk} = \sum_i x_{ilk} \quad \forall k, \forall l, \tag{5}$$

$$\sum_i \sum_j d_i x_{ijk} \leq Q_k \quad \forall k, \tag{6}$$

$$\sum_i \sum_j (c_{ijk} + (\sqrt{20d_i} + 4)) x_{ijk} \leq 480 \quad \forall k, \tag{7}$$

$$x_{ikj} \in \{0, 1\}, \quad u_i \in \mathbb{N}. \tag{8}$$

Restrictions in the equations 2 and 3 make reference to the binary variable, where we make sure that every node is visited once.
Restriction 4 assures that there won´t be any sub-tours and that every delivery vehicle departs and arrives to the DC.
Restriction 5 makes sure the vehicle that arrives to a node is the same one that leaves it.
Restriction 6 makes sure that the delivery vehicles won´t exceed its capacity (volume capacity). Meanwhile d_{ij} is the distance (by using the streets) from the node i to the node j and Q_k is the volume capacity of the vehicle k.
Restriction 7 makes that every driver doesn't work more than 8 hours (480 minutes) considering driving and service time.

Finally, the equations in 8 are the technical restrictions of the model. $x_{i,j,k}$ is a binary variable and u_i must be a natural number.

V. Solution

The model was solved using Python's library vrpy. Furthermore, some heuristics were used to reduce the problem's complexity; a greedy randomized heuristic and a pricing strategy called 'Hyper'.

The greedy randomized heuristic 'generates a path starting at the Source node and then randomly selects an edge among the γ outgoing edges of least reduced cost that do not close a cycle and that meet operational constraints (γ is a parameter). This is repeated until the Sink node is reached . The same procedure is applied backwards, starting from the Sink and ending at the Source, and is run 20 times. All paths with negative reduced cost are added to the pool of columns.' [7]

On the other hand, the 'Hyper' pricing strategy is a feature in the vrpy library that uses Hyper-Heuristics for the dynamic selection of the other 3 available pricing strategies [7].

The problem was solved by visiting all the 181 delivery locations in 1146 minutes, considering traveling and service time using 5 vehicles. One of the routes is shown in the following map:

On the other hand, the problem was solved with the same approach but grouping the delivery locations according to the predefined Coppel's routes (referred as the divided problem). This way, the nodes were visited in 1595 minutes using 10 vehicles. Our solution to the CVRP takes about 28% less time using half the vehicles compared with the solution obtained by dividing the nodes according to Coppel's routes.

Figures 2 and 3 are examples of one of the routes in our solution and one of the routes in the divided problem respectively. The arcs are shown as straight lines for better visualization, however, the distances were calculated going through the streets. While our route passes through 45 nodes in a time of 182 minutes, the other route passes through 14 nodes in a time of 83 minutes. Since our optimized route visits an average of 0.25 nodes per minute, the other one visits 0.17 nodes per minute, meaning our route is more efficient.

VI. Limitations

A. Google API

While it is true that Google API was an essential tool in the solution of the problem, using it also has its limitations. By using the mean speed approach, that is, dividing the distance matrix by the average speed of motor vehicles in the metropolitan area of Guadalajara, the bias caused by the fact that the Google API calculates the traveling time by setting the departure time as the same time the request was made to the Google service is reduced significantly. Nevertheless, we acknowledge that our approach should also have some bias considering that the routes used to calculate the distance matrix might change depending also on the time the request was made to the Google API.

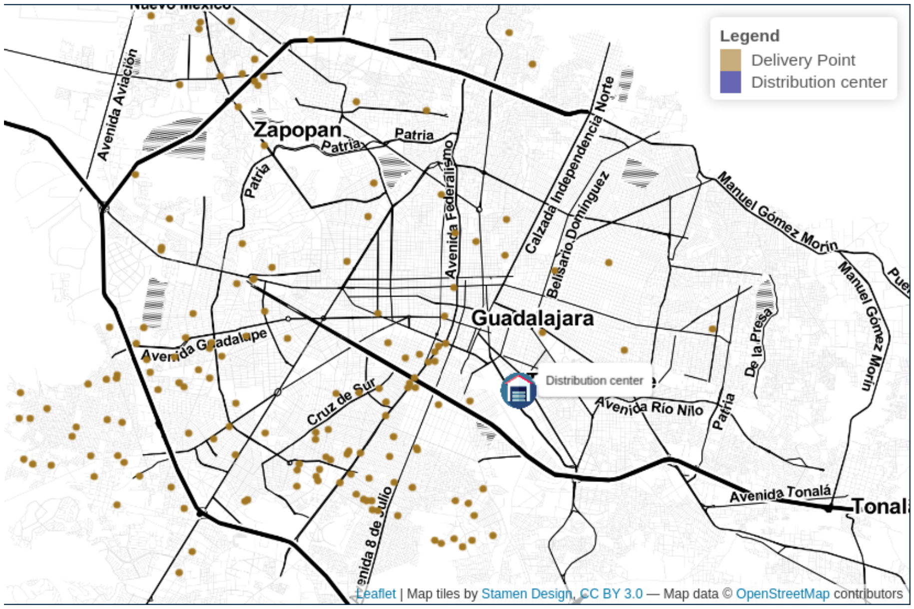

Fig. 1. Map of Guadalajara and Zapopan with the customer's locations and the DC location

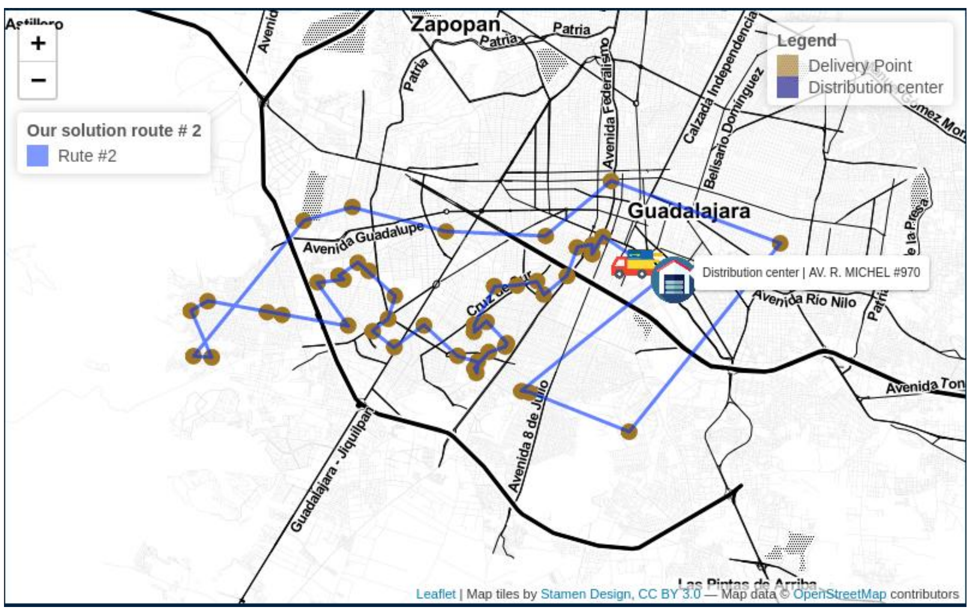

Fig. 2. Optimized route number 2 in our solution.

B. Assumptions and other limitations

In order to obtain a solution to the problem, some assumptions were made. These are that we are considering an average speed of 35 km/m for the vehicles. We are also assuming that every vehicle has a capacity of 20 m^3 which might also not been true, that the drivers work 8 hours at most, and that every delivery vehicle departs a arrives at the DC and not at any store. Other limitations of the solution of the problem include that we are only considering the volume of the packages, and not their weights and specific dimensions and we are also not considering if clients want their deliveries in a specific time span. Moreover, our main goal is to minimize the traveling time, which may differ from the store's goal (minimizing ecological footprint, time, costs, etc.).

VII. Conclusion

Now, more than ever, the availability of the internet is enhancing all the ways to comfort our lives, but behind this, there is a great effort from the logistic of an enterprise to improve the challenges that the clients demand. As seen in this report, nowadays people are becoming extremely demanding, not only for products but for service and quality of the companies, creating new ways for technology to take over for improving in every way possible, and, generating new products that take life to a new level.

Deliveries and achieving quality expectations are the biggest challenges on a delivery system chain. The Traveling Salesman

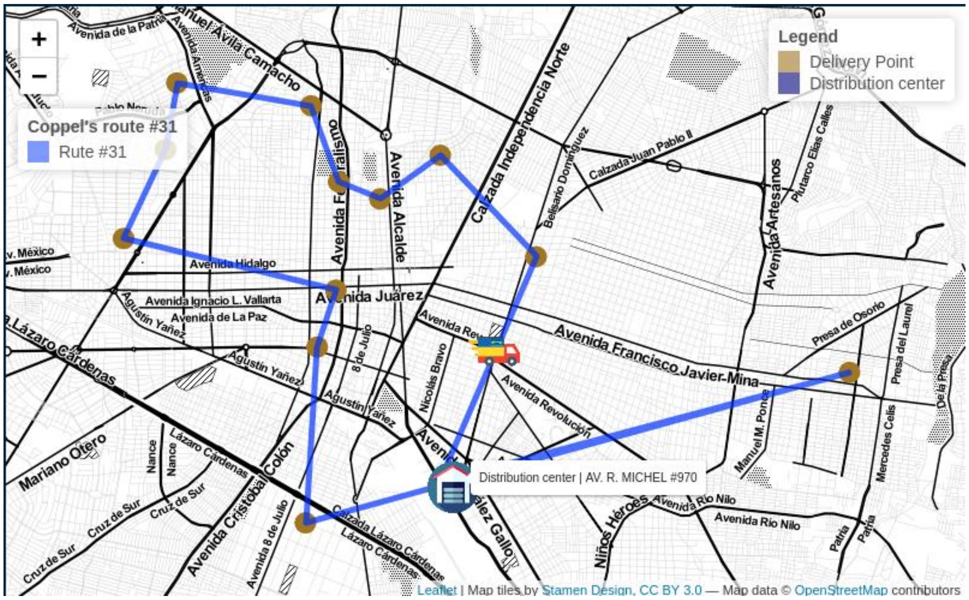

Fig. 3. Coppel's 31st route.

Problem, and a more specific version, the Capacitated Vehicle Routing Problems gives an approach to solve this likely to an impossible situation. The approach taken in this article reduces the bias caused by the fact that the API used to generate the cost matrix, sets the departure time as the same time the request is made.
Coppel would benefit from using the mean speed approach, this is, taking a distance matrix and dividing it by the average speed, because it results in an average time matrix that does not depend on the time the request was made to the API.

VIII. Acknowledgments

We thank Dr. Araceli Zavala Martínez, Ing. Martha Ivette Estrada Suarez, and Dr. Abelardo Ernesto Damy Solís for giving us assistancce in the elaboration of this work.

Rodrigo Morales Aguayo studies at Tecnológico de Monterrey, in pursue of the Data Science and Mathematics Engineering degree, expecting to graduate in January 2023. Currently, works at Roque-Gardea, which is a finances consulting group for start-up businesses. He wants to focus his studies in the finances and optimization areas. Loves sports, specifically: basketball, soccer and paddle tennis.

Kaled Corona-Romero is pursuing a Bachelor degree in Data Science and Mathematics at Tecnologico de Monterrey. He works as a researcher assistent with the researchers group Nanobiotechnology labs in the area of machine learning and image analysis. He also works as a co-research in the area of cybersecurity where he wants to develop new technology to enhance privacy. Besides academic activities, he spends its time writing a sci-fi novel, playing violin, and running.

Felipe Villaseñor Navarrete is also pursuing a B.S. in Data Science and Mathematics Engineering at Tecnologico de Monterrey, and he is also expected to graduate in June 2023. He is interested in the finance and marketing areas. His hobbies include training calisthenics, playing soccer, and playing video games.

Rodrigo Hiroshi Argandar Nishizawa is currently studiyng for his Data Science and Mathematics degree at Tecnologico de Monterrey, expected to graduate in June 2023. Currently working as a fitness coach at Commando Studio with three certifications involving fitness and basic first aids.

REFERENCES

[1] R. Noguez, "Confinamiento impulsa ventas online en México: crecen 81% en 2020," 2021. [Online]. Available: https://www.forbes.com.mx/negocios-ventas-online-mexico-2020/

[2] S. Sánchez, "México creció 500% sus ventas en línea por el confinamiento en casa: Kantar," 2020. [Online]. Available: https://www.forbes.com.mx/negocios-mexico-crecio-500-sus-ventas-en-linea-por-el-confinamiento-en-casa-kantar/

[3] G. Kostopoulos, "The Impact of Service Delivery System Effectiveness on Service Quality," 2012. [Online]. Available: https://core.ac.uk/download/pdf/209085732.pdf

[4] C. Martínez, "Coppel acata cierre de tiendas y vende online," 2020. [Online]. Available: https://www.eluniversal.com.mx/cartera/coppel-acata-cierre-de-tiendas-y-vende-online

[5] L. D., "Travelling Salesman Problem: It's more than cost-efficiency," 2020. [Online]. Available: https://blog.locus.sh/travelling-salesman-problem-and-how-can-tech-solve-it/

[6] G. de Jalisco, "Velocidad promedio de desplazamiento de los vehículos de motor en el Área Metropolitana de Guadalajara." [Online]. Available: https://seplan.app.jalisco.gob.mx/mide/panelCiudadano/detalleIndicador/397

[7] vrpy, "Solving Options." [Online]. Available: https://vrpy.readthedocs.io/en/latest/solving$_{o}ptions.html$