NAZARBAYEV UNIVERSITY
SCHOOL OF ENGINEERING AND DIGITAL SCIENCES

EMEM 537 - Engineering Decision Tools

Term Project: Alem TAT - Application of Design optimization and Travelling Salesman problem to moving company

Instructor: Professor Prashant Jamwal

Prepared by: Zhanar Ibrayeva
Nurtas Rash
Zhanybek Suleiman
Shynggys Toremurat
Sabit Shaikholla

Link: https://github.com/sabit-shaiholla/EMEM-537-EDT-project.git

1. Abstract

Our group wants to consider a local delivery company named Alem TAT. The Alem TAT company has existed in the logistics services market for 18 years. The main activity of the company is cargo transportation and express delivery. Alem TAT offers a full range of courier services on the principle of "hand to hand", "door to door" in the shortest possible transit time, in integrity and safety. They are facing the problem of non-optimal loading of packages to the trucks and non-optimal delivery routes. So, our project will consist of two parts. In the first part our aim will be to optimize packaging of several products by applying design constraints so that the truck is loaded at maximum. In the second part we are going to solve the Travelling Salesman Problem. The company trucks need to deliver packages to 12 delivery points within Nur-Sultan. We want to optimize the route so that the distance travelled is minimized. The specific constraint in this case is that the route should start and finish at the warehouse.

2. Introduction

In order to optimize the distance and capacity, we will use the power of Microsoft Solver and Python. Our goal, technically speaking, is to reduce distance and boost the capacity of the trucks. Minimizing the distance, thus, spending less time on delivery and increasing capacity, thus, increasing the potential profit of the company are classic problems in transportation companies. The developed algorithm takes into account the capacity and distance and establishes for them specific constraints. For instance, for the capacity problem constraints are a design of the goods. For the distance, as a constraint is that the route should start and finish at the warehouse. In the results and recommendation sections, the algorithm and the solution given by the MS Excel Solver's general reduced gradient and application of Python in optimizing the distance between the routes.

3. Literature review

The literature review was mainly conducted in order to consider such options as design optimization, vehicle routing problem, different approaches of design optimization, classification of algorithms for the VRP.

The design optimization area is an area where project performance can be significantly higher than in the initial implementation (Papalambros et al, 2017). There are the following optimization steps: variables (description of design options), objective (choosing to maximize or minimize the function), constraints (compiling equities and inequalities of variables), feasibility (defining a value for variables that satisfy all constraints and objectives).

There are several approaches of design optimization. For example, the computer-based approach considered by Parkinson et al. (2013) requires several qualifications. To begin, it is needed to create a quantitative model to calculate the answers of interest. For example, if the task is to maximize heat transfer, it is necessary to calculate heat transfer for various design configurations. In case of minimizing costs, there is needed to first calculate the cost. Often, creating such an accurate quantitative model is not an easy task. Thus, creating an accurate and reliable quantitative model is the most important step in optimization. Parkinson et al. (2013) reports that it often takes about 90% of design optimization efforts to develop and test a quantitative model.

Vehicle Routing problem (VRP) solves combinatorial optimization and integer programming tasks. It answers the question "What is the optimal set of routes for a fleet of vehicles that must be crossed in order to deliver them to a given set of customers?" It summarizes the travelling salesman problem (TSP), which first appeared in an article by George Danzig and John Ramser in 1959 (Dantzig et al, 1959), in which the first algorithmic approach was created and applied to the supply of gasoline. Often, the context is to deliver goods in a central warehouse to customers who have placed orders for such goods.

Grigorios et al. (2020) in their paper provide classification of algorithms for the VRP applied in freight transportation cases and companies. Algorithms for solving VRP variants are divided by three main groups which are exact, heuristic, and metaheuristic methods. Grigorios et al. (2020) describe 16 applied algorithms in VRP variants (capacitated VRP, VRP with time windows, VRP with pickups and deliveries, heterogeneous fleet VRP, VRP with multiple depots and collaborative VRP, green VRP, open VRP, dynamic and stochastic VRP, multi-trip VRP, multi-echelon VRP, time-dependent VRP, two dimensional and three dimensional VRP, consistent VRP, split delivery VRP, periodic VRP, truck and trailer VRP) which shows the variety and complexity of routing problems.

Moreover, huge research efforts are aimed at developing efficient algorithms in order to develop optimal solutions for voluminous problems. For example, in an article Xuesong et al. (2018) have developed a program based on time dependent and state-dependent path searching framework. The authors provide in the paper an open source lightweight VRP with pickup and delivery with time windows (VRPPDTW) to provide a high-quality and computationally efficient solution mechanism for on-demand transportation applications (Xuesong et al., 2018).

The results of literature review demonstrate that the optimization of design and transport problems is an actual task. Thus, they require a lot of research work including the development of complex and accurate algorithms and models.

4. Methodology

4.1. Knapsack optimization method

Imagine that there is a backpack that needs to be packed, however there are two constraints: there is a weight limit, and all the items have monetary value. The goal is to maximize the value of the selected items and at the same time not going over the weight limit. The knapsack optimization method can be a good solution to the resource allocation problems. The knapsack problem was developed early in 1897 by a mathematician Tobias Dantzig. A study conducted by Skiena (1999) showed that the knapsack problem was the third most needed algorithmic problems out of 75. There are three basic types of a knapsack problem (Chu and Beasley, 1998):

- **0-1 knapsack problem:** the number of copies of each type is restricted to 0 or 1.
  Maximize $\sum_{i=1}^{n} v_i x_i$
  Subject to $\sum_{i=1}^{n} w_i x_i \ \le W;\ x_i \in \{0,1\}$
  Here,
  n – number of items from 0 to n
  w_i – weight
  v_i – value
  W – maximum weight capacity
- **The bounded knapsack problem:** there is no restriction that there is only one item, instead number of copies are restricted with the maximum non-negative integer.
  Maximize $\sum_{i=1}^{n} v_i x_i$
  Subject to $\sum_{i=1}^{n} w_i x_i \ \le W\ ;\ x_i \in \{0,1,2,3,\dots,c\}$
- **The unbounded knapsack problem:** the only restriction is that number of copies need to be non-negative integer by eliminating upper bound.
  Maximize $\sum_{i=1}^{n} v_i x_i$
  Subject to $\sum_{i=1}^{n} w_i x_i \ \le W;\ x_i \ge 0;\ x_i \in Z$

In our case, to solve the problem there are two constraints: weight and volume limits.

4.2. The vehicle routing problem

The logistics companies must handle hundreds of delivery points every day with fixed number of fleets. The goal is to minimize costs and at the same time maximize efficiency. This is exactly the question of vehicle optimization problem. This is one of the difficult math problems to solve when multiple constraints need to be considered. For example, the number of vehicles, capacity, time window of stores, types of goods, human resources, traffic, and road condition (Golden, Raghavan and Wasil, 2008). Once the company manages to solve VRP, vehicle utilization can be improved by reducing travelling time and saving a lot on logistics cost. In other words, it will help to find the optimal route for the delivery fleet. The vehicle routing problem is the generalization of Travelling Salesman Problem. Modelling of VRPs has three different approaches:

- Vehicle flow formulations – cannot be used in practical applications
- Commodity flow formulations – used to find exact solution
- Set partitioning problem – used for general route cost

The basic vehicle routing problem can be handled by genetic algorithm where customer demand is known, and delivery is realized from the single depot (Baker and Ayechew, 2003). The constraints are the vehicles' weight limit and limit on the distance travelled. The aim is to calculate the set of delivery routes satisfying the constraints with minimum cost, which can be related to the minimum distance travelled.

In our case, there would be two constraints:

- The start and finish of the routes at one place
- The number of trucks is 10

The problem is solved by using Python.

5. Results & Recommendations

5.1. Design Optimization Problem

In order to maximize the capacity utilization of the delivery truck fleet of the company the multi-variable knapsack optimization problem was constructed. The packaging and truck load problems are intrinsic to movers' companies and closely tied to their performance. Because the timely delivery of packages to customers as well as optimum truckload of the shipments are the key success factors for such firms.

Alem TAT owns a fleet of ten box trucks dedicated for carrying a maximum of five tons of cargo. The transportation service of the company is performed as shown in Figure 1. First, a courier is forwarded to the customer's site to receive goods, which are packaged in standard shipping boxes. Then received packages are loaded into the truck and delivered to the destination point.

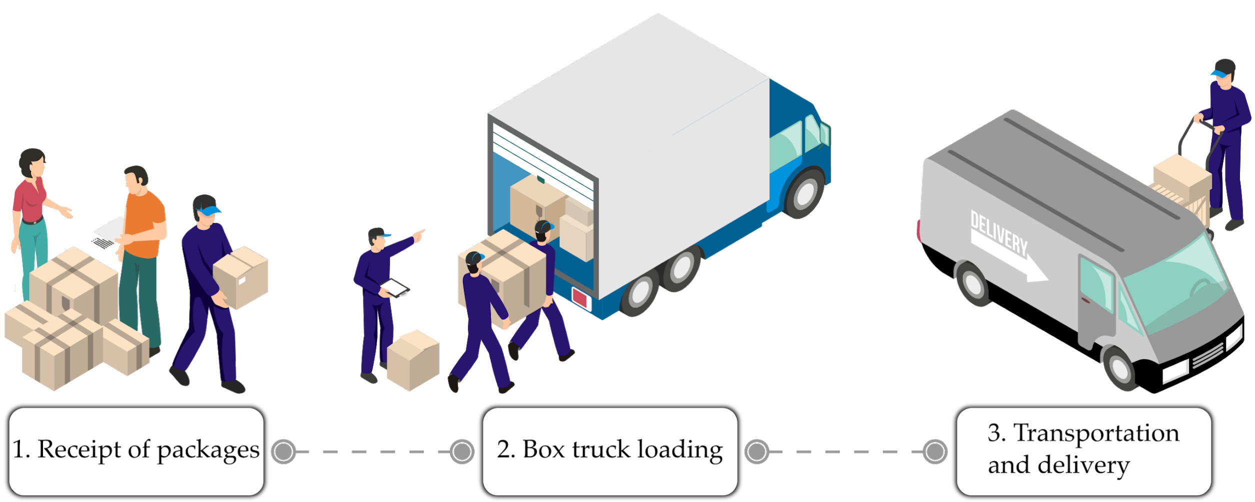


Figure 1. Shipping procedure of Alem TAT

In this process the optimum loading of a truck is one of the critical issues that impacts the number of hauls of each truck on a daily basis and thereby the company's cost efficiency in the long run. To formulate the objective function of the problem the standard dimensions, volume, maximum load weight, and price of shipping boxes as can be seen from Table 1 were retrieved from the company's staff.

Table 1. Basic shipping box data

Box	Length (m)	Width (m)	Height (m)	Volume (m3)	Weight (kg)	Price/box (tg)
Premium postal box	0.092	0.383	0.295	0.0104	0.3000	763
Rectangular shipping box	0.068	0.177	0.081	0.0010	0.3600	410
Kraft box	0.064	0.223	0.15	0.0021	0.2400	320
Pizza box	0.037	0.26	0.26	0.0025	0.3000	174
Square shipping box	0.084	0.3	0.3	0.0076	1.0000	584
Shipping box with lid	0.233	0.233	0.233	0.0126	1.2000	1980
Small removal box	0.34	0.375	0.25	0.0319	2.5000	438
Medium-sized cardboard box	0.17	0.5	0.375	0.0319	3.0000	460
Large cardboard box	0.34	0.5	0.375	0.0638	5.0000	611
XXL cardboard box	0.34	0.76	0.375	0.0969	8.0000	780

The main goal of the company in each turn of the truck loading is to maximize the occupancy of the cabin. Since the trucks have a difference in cabin volume and load capacities these are the main constraints of the problem. Other constraints as well as the objective function are provided in the equation Figure 2. To find the solution of this optimization problem the Microsoft Excel Solver was utilized as presented in Figure 3.

Maximize

$$f(x) = 0.0104x_1 + 0.001x_2 + 0.0021x_3 + 0.0025x_4 + 0.0076x_5 + 0.0126x_6 + 0.0319x_7 + 0.0319x_8 + 0.0638x_9 + 0.0969x_{10}$$

Subject to

$$0.0104x_1 + 0.001x_2 + 0.0021x_3 + 0.0025x_4 + 0.0076x_5 + 0.0126x_6 + 0.0319x_7 + 0.0319x_8 + 0.0638x_9 + 0.0969x_{10} \leq 20$$

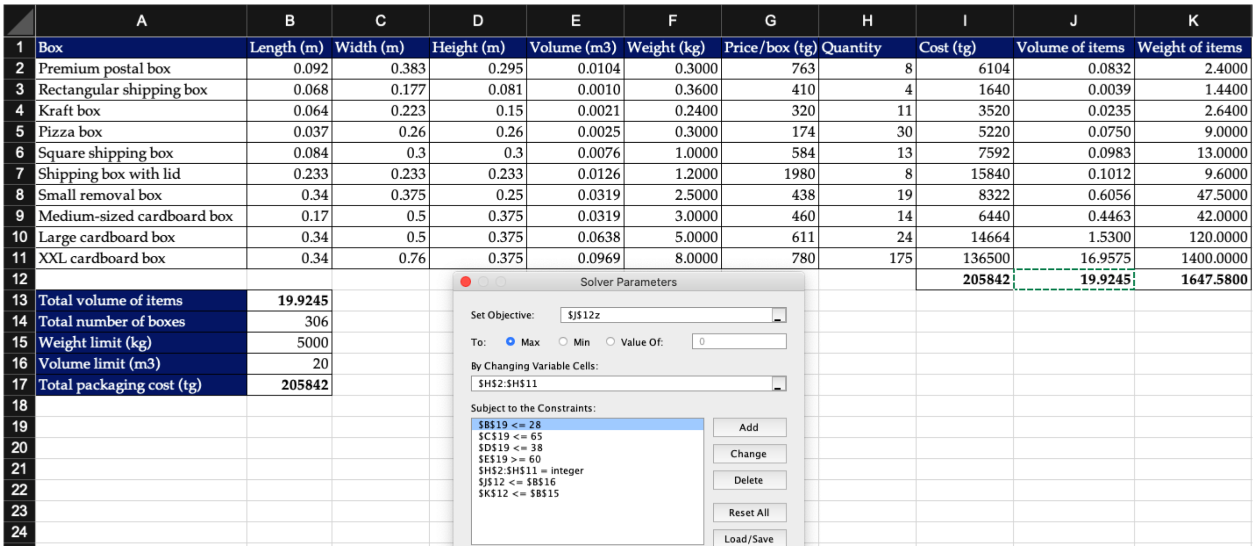
$$0.3x_1 + 0.36x_2 + 0.24x_3 + 0.3x_4 + x_5 + 1.2x_6 + 2.5x_7 + 3x_8 + 5x_9 + 8x_{10} \leq 5000$$

$$x_1 + x_2 + x_3 \leq 28$$

$$x_4 + x_5 + x_6 \leq 65$$

$$x_9 + x_{10} \geq 60$$

Figure 2. The objective function and constraints of the knapsack optimization problem



Box	Length (m)	Width (m)	Height (m)	Volume (m3)	Weight (kg)	Price/box (tg)	Quantity	Cost (tg)	Volume of items	Weight of items
Premium postal box	0.092	0.383	0.295	0.0104	0.3000	763	8	6104	0.0832	2.4000
Rectangular shipping box	0.068	0.177	0.081	0.0010	0.3600	410	4	1640	0.0039	1.4400
Kraft box	0.064	0.223	0.15	0.0021	0.2400	320	11	3520	0.0235	2.6400
Pizza box	0.037	0.26	0.26	0.0025	0.3000	174	30	5220	0.0750	9.0000
Square shipping box	0.084	0.3	0.3	0.0076	1.0000	584	13	7592	0.0983	13.0000
Shipping box with lid	0.233	0.233	0.233	0.0126	1.2000	1980	8	15840	0.1012	9.6000
Small removal box	0.34	0.375	0.25	0.0319	2.5000	438	19	8322	0.6056	47.5000
Medium-sized cardboard box	0.17	0.5	0.375	0.0319	3.0000	460	14	6440	0.4463	42.0000
Large cardboard box	0.34	0.5	0.375	0.0638	5.0000	611	24	14664	1.5300	120.0000
XXL cardboard box	0.34	0.76	0.375	0.0969	8.0000	780	175	136500	16.9575	1400.0000
								205842	**19.9245**	**1647.5800**

Total volume of items	**19.9245**
Total number of boxes	306
Weight limit (kg)	5000
Volume limit (m3)	20
Total packaging cost (tg)	**205842**

Figure 3. Solution using Microsoft Excel Solver

Based on the results of this solution the company should fill the cabinets of trucks with 306 boxes in total giving the priority to XXL cardboard boxes. Owing to this decision Alem TAT would reach the full load of its trucks without exceeding both volume and weight limits. Finally, the cost of packaging for each truck is estimated to be KZT 205,842.

5.2. Vehicle Routing Problem for moving company

After considering the process of optimum loading of tracks, another issue to consider is the so-called vehicle routing problem which is an integer programming problem identifying the optimal set of routes for vehicles to deliver products and goods to a given set of customers. As we have seen in the previous part, the proper design optimization of vehicle loading with a set of products available with a set of constraints was provided. In the same manner, the analysis of vehicle routing problem is provided further. Let's consider the Alem TAT company: it's a moving company which has a goal to deliver the products and goods to several locations provided the customer demand and vehicles' overall capacity. The items to be delivered have their own attributes such as quantity, weight and cost of packaging. The main problem is to deliver the items to designated locations with the minimum cost possible to achieve.

Based on the thorough literature and analysis of the state-of-the-art solutions available to solve the Vehicle Routing Problem (VPR), it was determined that integer optimization problem with graph theory applied will be most suitable as it is one of the efficient and effective solutions. In addition, it can be said that VPR is NP hard problem, meaning that the computational effort required to solve this problem increases exponentially with the problem size. It also should be noted that the Traveling Salesman problem is a particular case of Vehicle Routing Problem with the limitation of vehicle (in that case – it is a salesman who travels and sales items to customers).

To start with, the definition of graph network is to be assigned. The designated locations can be considered as the nodes of the graph and including our departure depot (company's warehouse). The decision variable is assigned as a binary variable - whether the vehicle took the route (1) or not (0). The main objective function is to minimize the sum of traveled distances by all vehicles. The definitions of the network can be seen as follows:

$$G = (L, R) - a\ graph\ showing\ the\ location\ and\ route\ to\ the\ customer$$

$$N = \{0,1, \dots n\} - a\ set\ of\ nodes\ (customer's\ location\ and\ depot - warehouse)$$

$$V = \{0,1, \dots k\} - a\ set\ of\ vehicles$$

$$D_i \geq 0 - demand\ of\ the\ respective\ customer$$

$$Q_j \geq 0 - capacity\ of\ each\ vehicle\ (same\ for\ all)$$

$$C_{ij} - distance\ covered\ (cost)\ by\ vehicle\ (distance\ between\ i\ and\ j)$$

$$x_{ij}^k = \begin{cases} 1 \\ 0 \end{cases} - whether\ the\ vehicle\ took\ the\ route\ or\ not$$

$$Minimize \sum_k \sum_{i,j} C_{ij} * x_{ij}^k - main\ objective\ function$$

Following that, the set of constraints should be defined accordingly:

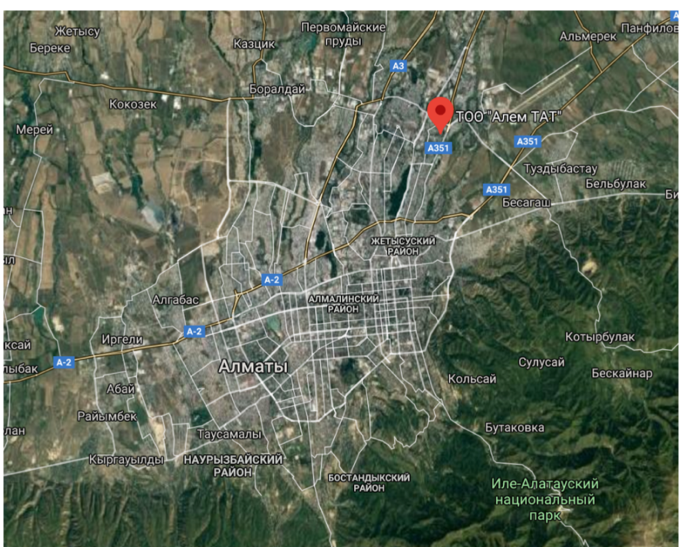
1. Only 1 particular vehicle will be allowed to deliver items to particular customer
2. All the vehicles will depart from one point which is depot (warehouse)
3. Number of entering and number of leaving vehicles on edge nodes are the same (1 vehicle per customer)
4. The delivery capacity of the vehicles cannot exceed the loading limit which is the same for all of them (Q)
5. The subtours will be removed for each vehicle (Ulrich and Stanêk 2017).
6. Decision variable constraint – either vehicle will take or not the route (binary)

1) $\sum_k \sum_{i,j} x_{ij}^k = 1 - only\ 1\ vehicle\ to\ deliver\ to\ particular\ customer$
2) $\sum_{i,j} x_{0j}^k = 1 - all\ vehicles\ will\ be\ departed\ from\ the\ depot$
3) $\sum_{i,j} x_{ji}^k - \sum_{i,j} x_{ij}^k = 0 -$
   $number\ of\ entering\ and\ number\ of\ leaving\ vehicles\ are\ the\ same$
4) $\sum_{i,j} q_{ij} * x_{ij}^k \leq Q -$
   $the\ delivery\ capacity\ of\ each\ vehicle\ should\ exceed\ the\ loading\ limit$
5) $\sum_k \sum_{i,j} x_{ij}^k \leq |S| - 1: constraint\ for\ removing\ subtours\ for\ the\ vehicles$
6) $x_{ij}^k = \begin{cases} 1 \\ 0 \end{cases} - decision\ variable\ constraint$

After considering all the constraints and defining main objective function, it was decided to write a program based on Python language as it has variety of great packages and tools for visualization and solving mathematically expressed equations. The main idea was to apply the concepts of dynamic programming so that the developed application could be used universally.

To start with, libraries and packages that will be utilized are as follows: Numpy, MatPlotLib, Pandas, Gmaps, Itertools are used to operate with numerical values and visualizations, whereas the Pulp library is used to solve mathematically expressed equations. The main aim was to provide the visualization of the graph network on the street view map using Google Maps Application Programming Interface (API). However, as there were introduced some limitations from Google Cloud Platform, the fullest potential of proper usage of this tool couldn't be provide as expected. Nevertheless, this API is used in other places to calculate the distances between locations based on their latitude and longitude. The proper API key should be invoked in Google Cloud Platform in order to use this API.

Following that, the user is asked to enter the number of customers and number of vehicles available (as this is provided for moving company). The capacity of each vehicle is set to constant (it might be changed accordingly; however it was decided to set it to constant as the vehicles' capacity is a constant value). Next, it is asked to enter the latitude and longitude of the depot (warehouse) from where the vehicles will start their travel. The probable location of Alem TAT's warehouse in Almaty, Kazakhstan, can be viewed on Figure 4A.



	latitude	longitude	demand
0	43.346226	76.978129	0
1	43.763506	76.356953	11
2	43.280156	76.815274	12
3	43.104267	76.534262	19
4	43.361023	77.036700	14
5	43.724881	77.287979	16
6	43.430443	76.815545	12
7	43.270634	76.911575	10
8	43.136565	76.843449	16
9	43.216793	77.083314	12

Figure 4. A) Alem TAT's warehouse location (depot) **B)** Locations of customers and their demand.

As for the simplicity of demonstration, the locations of customers (their latitude and longitude) are randomly calculated using normal (Gaussian) distribution with the mean of depot's location and standard deviation as 0.25 which is approximately equal to 40 km in distance (distance from one end of Almaty city to another). Also, the demands of those customers are also calculated using normal distribution with mean 10 and standard deviation of 20, for simplicity of calculations. Otherwise, it will be a tedious work to paste location of each item and it's corresponding demand. However, within the context of this research analysis, the random samples of normal distributions will be used instead. The probable data frame of locations of customers with their corresponding demands can be seen on Figure 4B. Note that 0^{th} location is depot with 0 demand (indexing in Python starts from 0).

Next, the distances between node edges are calculated using Google Maps API. Then, by using Pulp library's function of linear problem solver with attribute "CVRP (Capacitated Vehicle Routing Problem)" is used. Following that, all the constraints described in mathematical form are inserted. As a result, the number of vehicles required, and total sum of distances covered by those vehicles are calculated.

Lastly, using MatPlotLib library the routing scheme was presented in a plane coordinate system, where depot is presented in blue circle and node edges (customers) are represented in green circles. The black lines depicting the vehicle route taken. The example of routing scheme can be seen on Figure 5A. The particular case of Traveling Salesman Problem can be seen on the Figure 5B, where the vehicle (salesman) capacity constraint is much exceeding the sum of all demands of customers.

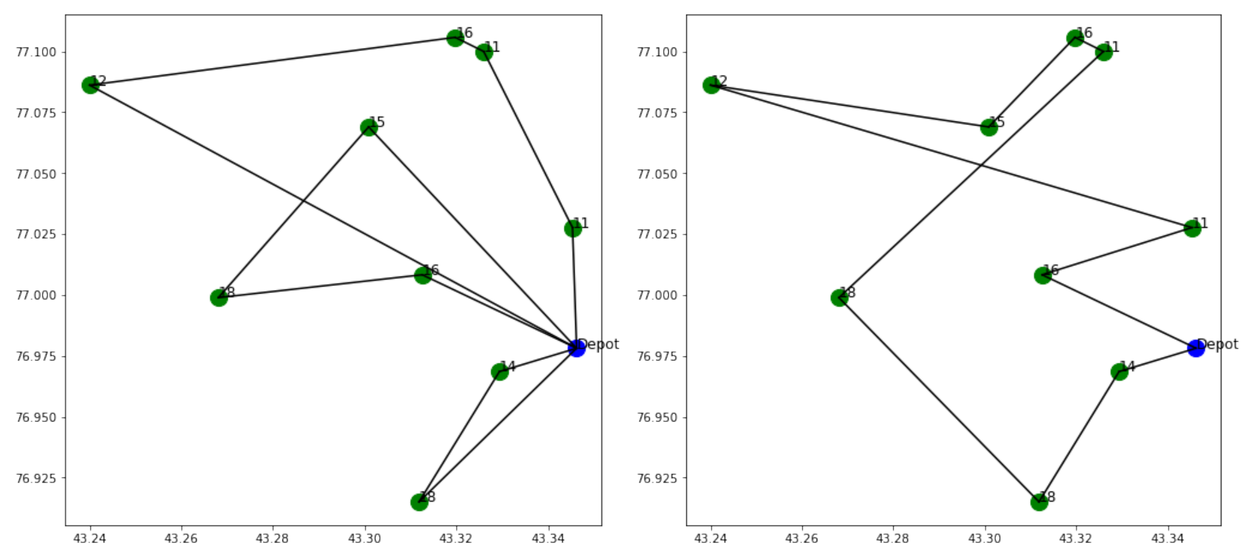


Figure 5. A) The routing scheme with 9 customers and 3 vehicles **B)** Case of Traveling Salesman Problem

It should be noted that the problem set considered in this research analysis is simplified and doesn't fully reflect the real-life case scenario which might include other types constraints. With the increase of edge nodes (number of customers), the computational complexity of the problem increases significantly (NP hard problem). Therefore, one probable future research might be to develop a sensitivity system to analyze how the constraints be properly reflected in finding optimal value of objective function.

The code is available through the link: https://github.com/sabit-shaiholla/EMEM-537-EDT-project.git

6. Conclusion

The report has discussed the optimization of capacity and distance between routes. It clearly shows the application of a multi-variable knapsack optimization problem and of solving it. By using the power of MS Excel Solver we obtained that the optimal capacity of trucks' cabinets can allow 306 boxes where XXL cardboard boxes have the highest priority. Also, it can be seen that with this capacity, the company can increase their costs by not exceeding both volume and weight limits. Apart from that, the report has shown the solution to the vehicle routing problem by using the integer optimization problem with graph theory. Along with this, the application of programming on Python with a variety of helpful packages was involved. So inserting all constraints the program can calculate the number of vehicles required, and the total sum of distances covered by those vehicles. The good thing here is that using the concepts of dynamic programming, this application can be used universally.

To demonstrate how the application works the locations of customers and their demands are calculated randomly with the particular mean and standard deviation. Since these constraints were selected randomly, as further research, the development of a sensor system to analyze how constraints such as locations and demands can affect finding the optimal solution might take place. Along with this, there might be other types of constraints, such as congestion on the road which also should be considered if they appear.

Reference list

Baker, B., & Ayechew, M. (2003). *A genetic algorithm for the vehicle routing problem. Computers & Operations Research*, *30*(5), 787-800. doi: 10.1016/s0305-0548(02)00051-5

Chu, P., Beasley, J. A. (1998). *Genetic Algorithm for the Multidimensional Knapsack Problem.* Journal of Heuristics 4, 63–86 . doi.org/10.1023/A:1009642405419

Dantzig, George B., Ramser, J.H. (1959). *The Truck Dispatching Problem. Management Science.* doi.org/10.1287/mnsc.6.1.80

Golden, B. L., Raghavan, S., & Wasil, E. A. (Eds.). (2008). *The vehicle routing problem: latest advances and new challenges* (Vol. 43). Springer Science & Business Media.

Grigorios D. K., Sotiris P. G., Evripidis P. K. (2020). *Vehicle routing problem and related algorithms for logistics distribution: a literature review and classification.* Available at: https://link.springer.com/article/10.1007/s12351-020-00600-7

Skiena, S. (1999). *Who is interested in algorithms and why?* ACM SIGACT News, 30(3), 65-74. doi: 10.1145/333623.333627

Papalambros, Panos Y., Wilde, Douglass J. (2017). *Principles of Optimal Design: Modeling and Computation. Cambridge University Press*

Parkinson A. R., Balling R. J., Hedengren J.D. (2013). *Optimization Methods for Engineering Design*. Brigham Young University

Pferschy, Ulrich, and Rostislav Staněk (2017). *Generating subtour elimination constraints for the TSP from pure integer solutions.* Central European journal of operations research 25, no. 1 (2017): 231-260.

Xuesong Zh., Lu T., Monirehalsadat M., Lijuan Zh., Yu Y., Yongxiang Zh., Pan Sh., Jiangtao L., Tie Sh. *(2018). Open-source VRPLite Package for Vehicle Routing with Pickup and Delivery: A Path Finding Engine for Scheduled Transportation Systems.* Urban Rail Transit V. 4, p. 68–85(2018)