Delft University of Technology

Project Operations Optimisation: Generalized Covering Traveling Salesman Problem

AE4441-16

González Álvarez, Marina - 5044030

Kloosterman, Luc - 4697006

January 11, 2020

Contents

1 Introduction

As stated in [1], the generalized covering traveling salesman problem is a recently introduced variant of the traveling salesman problem. Thus, given a set of cities, facilities and a depot, the salesman has the mission of visiting a subset of facilities -always starting from the depot- in order to cover some of the customers demand (customers live in the cities) so that a constraint function based on these clients coverage is satisfied.

Thereby, in this assignment the methodologies studied in the *AE4441-16 Operations Optimisation* subject for the resolution of MILP problems have been used and implemented with the aim of solving the generalised covering traveling salesman problem (GCTSP) presented in [1]. First, in Section 2, a detailed description of the model and all the constraints involved is presented. Moreover, a small example of the problem is solved and analysed so as to provide the reader with an easier and more visual representation of the constraints and the outcome they produce. Sencondly, in Section 3, a more detailed testing of the model is conducted. Then, a sensitivity analysis is done in Section 4 by varying certain parameters and analysing the effect of these modifications. Finally, some conclusions are extracted in Section 5, once the whole analysis has been completed.

2 Description of the model and testing

The aim of the traveling salesman problem (TSP) is to find a minimum length cycle of a given set of cities where each city must be visited exactly once. Sometimes, however, limitations on efficiency or resources may appear, so that it may not be feasible to visit each and every city: here is where the GCTSP comes into play by introducing the concept of *coverage*.

Thereby, the GCTSP works with a given set of n cities that include a depot, facilities and some customer cities. A coverage radius r_i is associated with each facility i and a demand d_j is associated with each customer j. The aim of the model is to construct a minimum length tour over a subset of the facilities so that the sum of all the customer demands covered by this subset is at least D. Moreover, a client is considered to be covered by a subset of facilities in case that it is within the coverage radius of one or more facilities. Thus, to solve this problem, one must find a subset of facilities that can satisfy the demand D while also minimizing the total distance traveled by the salesman.

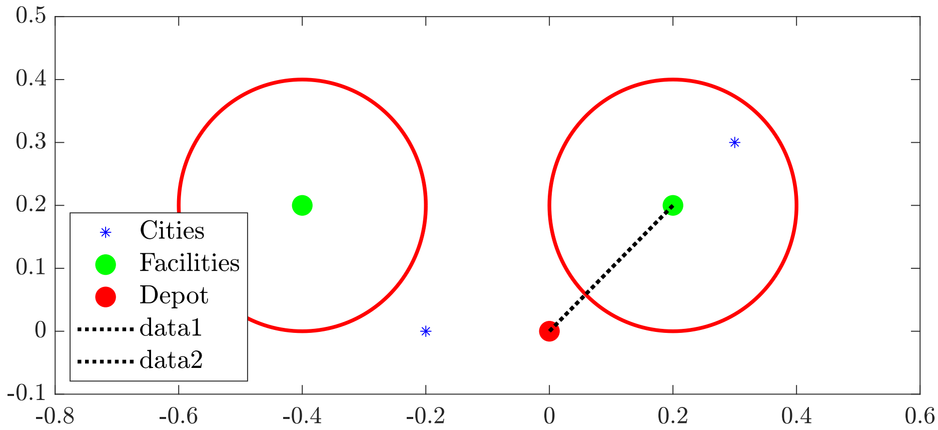


Figure 1: Simple example of a feasible solution with $C = 2$, $F = 2$ and $D = 1$.

Before starting with the actual explanation of the constraints, some mathematical notation is needed: $\{0\}$, F and C represent the depot, set of facilities and set of cities, respectively. Moreover, $E = \{(i, j) \,|\, i, j \in F \cup \{0\}\}$, which corresponds to the set of all edges. A cost (represented by a distance) t_{ij} is associated with each edge $(i, j) \in E$ and every customer i has a demand d_i (we will take $d_i = 1$ throughout the whole analysis) that will be satisfied by a tour only when it is within the coverage distance r_i of at least one facility on that tour. The facilities that are part of the tour are referred to as visited facilities, $F' \subseteq F$. As stated in [1], the GCTSP tries to find a tour of minimum length over $\{0\} \cup F$ that begins and ends at the depot, such that the sumtotal of the satisfied demands of the customers is at least D.

2.1 Mathematical formulation

In order to further explain the formulation of the objective function and constraints of the model, a simple example (see Figure 1) with two cities, two facilities and a depot will be analysed in conjunction with the theoretical purpose of each of the constraints, for clarity.

Variables

The variables that are used for the definition and subsequent solution of the model are:

- x_{ij}: variable used to indicate if an edge (i, j) is part of the tour (activated, $= 1$) or not (deactivated $= 0$). The total number of edges is: $\#x_{ij} = (\#facilities + depot) \cdot (\#facilities + depot - 1)$. $x_{ij} \in \{0, 1\}$.
- y_{ij}: to indicate whether the demand of customer $i \in C$ is satisfied by the facility $j \in F$ (activated, $= 1$) or not (deactivated $= 0$). $\#y_{ij} = \#cities \cdot \#facilities$. $y_{ij} \in \{0, 1\}$.
- z_i: to indicate if a facility $i \in F$ is visited by the tour (activated, $= 1$) or not (deactivated $= 0$). $\#z_i = \#facilties$. $z_i \in \{0, 1\}$.
- u_{ij}: auxiliary variable for the single loop constraint. $\#u_{ij} = (\#facilities + depot) \cdot (\#facilities + depot - 1)$. $u_{ij} \in \mathbb{Z}$.

Objective function

The objective function (see Equation 2.1) minimizes the total distance traveled by the salesman. In Table 1 all the values taken by the t_{ij} for the analysed example are shown. Thus, the function to minimize for this case (two cities $C = 2$, two facilities $F = 2$ and a depot) will have 6 terms: $Z = t_{01}x_{01} + t_{02}x_{02} + t_{10}x_{10} + t_{12}x_{12} + t_{20}x_{20} + t_{21}x_{21}$.

$$Minimize \sum_{(i,j)\in E} t_{ij}x_{ij} \tag{2.1}$$

Table 1: Example in Figure 1 - distances among depot and facilities.

t_{01}	t_{02}	t_{10}	t_{12}	t_{20}	t_{21}
0.2828	0.4472	0.2828	0.6000	0.4472	0.6000

Constraint #1: demand D satisfied by F'

The first constraint makes sure that the given total demand D is satisfied by a subset of facilities F' that are part of the tour.

$$\sum_{i \in C} \sum_{j \in F'} d_i y_{ij} \geq D \tag{2.2}$$

In order to actually implement this constraint, it has been divided in two different parts:

- Total demand D is satisfied (taking all facilities, F, into account): $\sum_{i \in C} \sum_{j \in F} d_i y_{ij} \geq D \rightarrow -\sum_{i \in C} \sum_{j \in F} d_i y_{ij} \leq -D$.
- Force those facilities j that cannot fulfill demand i to be zero.

Thus, Table 2 shows how this constraint would be shaped for the case of analysis.

Table 2: Example in Figure 1 - constraint #2.

y_{11}	y_{12}	y_{21}	y_{22}	b	Constraints
-1	-1	-1	-1	-1	$y_{11} + y_{12} + y_{21} + y_{22} \geq 1$
1	0	0	0	0	$y_{11} = 0$
0	1	0	0	0	$y_{12} = 0$
0	0	1	0	1	$0 \leq y_{21} \leq 1$
0	0	0	1	0	$y_{22} = 0$

Constraint #2: each customer allocated to at most one facility

The second constraint indicates that each customer must be allocated to at most one facility for satisfying the customer's demand, d_i. In this case, we have set $d_i = 1 \, \forall i \in C$, so that each of the cities represents one client that has a demand of 1. That is why, in other words, this constraint would also imply that each city can only be supplied by one facility maximum. For the particular case of analysis, this means that two additional constraints should be taken into account, as shown in Table 3.

$$\sum_{j \in F} y_{ij} \leq 1 \quad \forall i \in C \tag{2.3}$$

Table 3: Example in Figure 1 - constraint #2.

y_{11}	y_{12}	y_{21}	y_{22}	b	Constraints
1	1	0	0	1	$y_{11} + y_{12} \leq 1$
0	0	1	1	1	$y_{21} + y_{22} \leq 1$

Constraint #3: tour begins and ends at the depot

The third constraint basically assures that the path traveled by the salesman must both start and end at the depot. In other words, Equation 2.4 is making sure that only one edge departs from the depot ($\sum_{j\in F} x_{0j} = 1$) and only one edge arrives ($\sum_{j\in F} x_{j0} = 1$).

$$\sum_{j\in F} x_{0j} = 1 = \sum_{j\in F} x_{j0} \tag{2.4}$$

Table 4: Example in Figure 1 - constraint #3.

x_{01}	x_{02}	x_{10}	x_{12}	x_{20}	x_{21}	b	Constraints
1	1	0	0	0	0	1	$x_{01} + x_{02} = 1$
0	0	1	0	1	0	1	$x_{10} + x_{20} = 1$

Constraint #4: indegree and outdegree of the visited facilities

Equation 2.5 satisfies the constraints of indegree and outdegree of the visited facilities. Thus, this means that in case that a certain facility is visited, then one edge must arrive and one edge must depart.

$$\sum_{(i,j)\in E} x_{ij} + \sum_{(j,i)\in E} x_{ji} = 2z_i \quad \forall i \in F \tag{2.5}$$

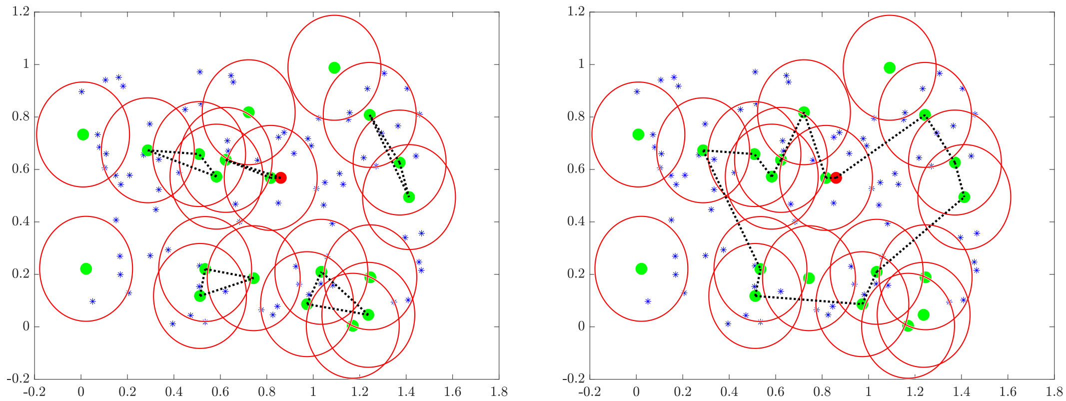
Table 5: Example in Figure 1 - constraint #4.

x_{01}	x_{02}	x_{10}	x_{12}	x_{20}	x_{21}	z_1	z_2	b	Constraints
1	0	1	1	0	1	-2	0	0	$x_{10} + x_{12} + x_{01} + x_{21} = 2z_1$
0	1	0	1	1	1	0	-2	0	$x_{20} + x_{21} + x_{02} + x_{12} = 2z_2$

Constraint #5: subtour elimination

Finally, Equation 2.6 represents the subtour elimination constraint. A graphical representation of this restriction can be observed in Figure 2. Concretely, for a problem with $C = 80$, $F = 20$ and $D = 50$ a solution with more than one subtour is represented in Figure 2a. Once we add constraint #5 the feasible solution corresponding to the minimum possible distance traveled by the salesman with only one subtour can be observed in Figure 2b.

$$\begin{aligned} u_i - u_j + nx_{ij} \le n - 1 \quad 2 < i \neq j \le n \\ 0 \le u_i \neq n - 1 \quad 2 \le i \le n \end{aligned} \tag{2.6}$$



(a) No single loop constraint

(b) Adding single loop constraint

Figure 2: Feasible solution example with $C = 80$, $F = 20$ and $D = 50$.

2.2 Mathematical formulation of the optimised model

During the implementation of this model it was found that the size of the problem increases tremendously when using an increasingly large data set. For example, a small data set which contains 2 facilities and 2 cities has 6 x_{ij} variables, 4 y_{ij} variables, 2 z_i variables and 6 u_{ij} variables, which gives a total of 18 variables. However, when the number of facilities and cities is doubled to 4 each the total number of variables will already be 60. If we do this one more time, so that we work with 8 cities and 8 facilities, we would have 184 variables.

Thus, an alternative approach is also proposed in order to decrease the size of the problem for larger data sets. This is done by modifying some of the variables, the objective function and constraints #1, #2 and #5. This section will expand on how the modification of the model works and will show the mathematical formulation after these changes are included.

Variables

As seen in Table 1, the weight (distance) assigned to x_{ij} is identical to that assigned to x_{ji} ($t_{ij} = t_{ji}$) and, therefore, the number of x_{ij} variables can be reduced to one half, by considering undirected links among the facilities. Secondly, all y_{ij} variables can be omitted when modifying constraint #1 and #2. Finally, by modifying constraint #5 all u_{ij} variables are also no longer needed. Thereby, for the case of the optimised model, the variables that are used for the definition and subsequent solution of the model are:

- x_{ij}: variable used to indicate if an edge (i, j) is part of the tour (activated, $= 1$) or not (deactivated $= 0$). The total number of edges is: $\#x_{ij} = (\#facilities + depot) \cdot (\#facilities + depot - 1)/2$. $x_{ij} \in \{0, 1\}$.

- z_i: to indicate if a facility $i \in F$ is visited by the tour (activated, $= 1$) or not (deactivated $= 0$). $\#z_i = \#facilties$. $z_i \in \{0, 1\}$.

Modification of the objective function

After introducing the modification in our approach based on the use of undirected links, instead of directed links among the facilities, the length of the objective function is reduced to only 3 terms. Thus, the shape of the objective function for the same problem considered in the previous section is now $Z = t_{01}x_{01} + t_{02}x_{02} + t_{12}x_{12}$. The distances between the facilities for the problem shown in Figure 1 are stated in Table 6.

Table 6: Example in Figure 1 - distances among depot and facilities after modification.

t_{01}	t_{02}	t_{12}
0.2828	0.4472	0.6000

Modification of constraint #1 and #2

Constraints #1 and #2 can be modified to be one single constraint as shown in Equation 2.7. This constraint adds a weight d_i (demand of each city) to z_i. Thus, the meaning of this constraint can be interpreted as: the sum of all the demand of the cities that are visited in the tour (thereby, the demand that is indeed satisfied) must be greater or equal than a total demand D fixed at the begininning of each simulation. This constraint is shown in Table 7 for the problem of Figure 1.

$$\sum_{i=1}^{F} d_i z_i \geq D \tag{2.7}$$

Table 7: Example in Figure 1 - modified demand constraint

z_1	z_2	b	Constraints
1	1	1	$1 \cdot z_1 + 1 \cdot z_2 \geq 1$

Modification of constraint #5

It is obvious that when we choose to omit the u_{ij} variables, an alternative way of forcing a single loop in the solution has to be obtained. In the original model this is done by setting a constraint on all possible subtours. When increasing the data set not only the numbers of variables is increased but also the number of possible subtours increases substantially. Therefore, constraint #5 is modified to be a lazy constraint so that, in the case that more than one loop is detected, it will add some additional constraints making the particular loops discovered infeasible. The mathematical definition is given by equation 2.8. Table

8 shows such a constraint for a problem with $F = 4$ where a loop exists between facilities 1, 2 and 4.

$$\sum x_{ij} \leq n - 1 \tag{2.8}$$

Table 8: Example in Figure 1 - example of single loop constraint.

x_{01}	x_{02}	x_{03}	x_{04}	x_{12}	x_{13}	x_{14}	x_{23}	x_{24}	x_{34}	b	Constraints
0	0	0	0	1	0	1	0	1	0	2	$x_{12} + x_{14} + x_{24} = 3$

3 Further verification

Although the validity of the model has already been tested in Section 2 while explaining how each of the constraints involved in the definition of the model works, some additional testing will be conducted in the current section. Thus, a simple example with six cities ($C = 6$) and three facilities ($F = 3$) will be analysed for different values of the minimum total demand that has to be satisfied (D). In this case, each city has again an inherent demand of $d_i = 1$ so that, for example, $D = 2$ means that the demand of at least two different cities should be satisfied.

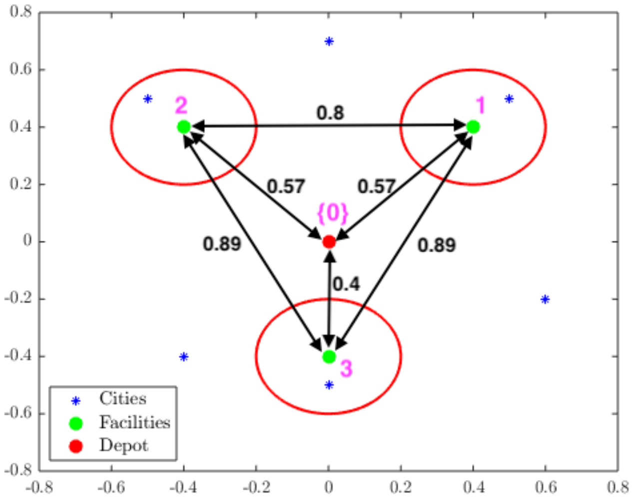


Figure 3: Representation of distances for a simple example with $C = 6$ and $F = 3$.

Figure 3 shows the distances among the different facilities and the depot. In Figures 4, 5 and 6 the results obtained for the cases of $D = 1$, $D = 2$ and $D = 3$, respectively, are graphically displayed.

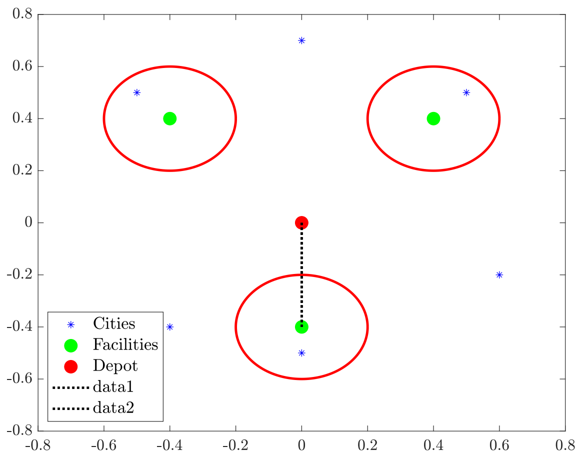


Figure 4: Test case #1 with $D = 1$

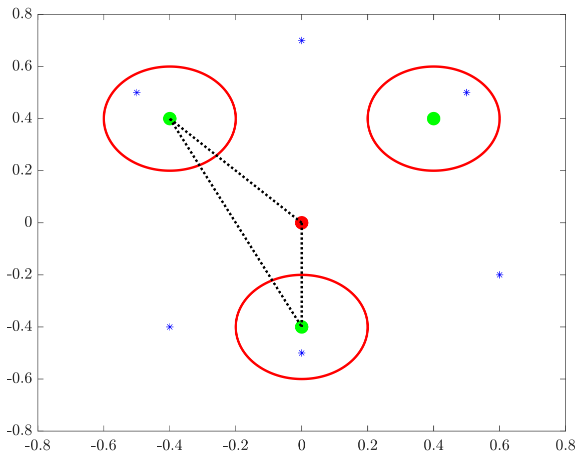


Figure 5: Test case #1 with $D = 2$

The reason why these simple examples have been chosen is because it is almost trivial to check whether the feasible solution obtained once the model has been implemented and fed to the solver is correct or not.

For the case of $D = 1$ the traveling salesman has the option of visiting any of the three facilities and the customer demand would already be satisfied. However, in order to minimize the distance, he/she should choose to visit the facility that is closest to the depot. Thus, in Figure 4 we see that, indeed, the solution obtained with our model would be to travel from the depot to facility #3 and back, as expected (this corresponds to the shortest path according to the scheme shown in Figure 3). The same reasoning applies to the cases of $D = 2$ and $D = 3$ as shown in Figures 5 and 6.

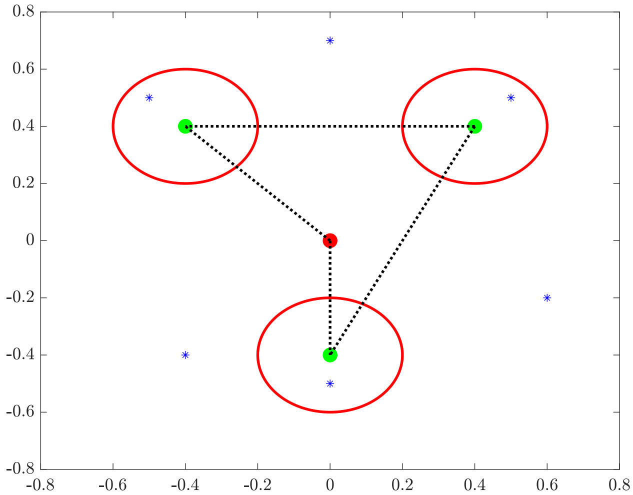


Figure 6: Test case #1 with $D = 3$

4 Sensitivity Analysis

In this section, we will analyse the effect of modifying several key parameters on the outcome of the model. In order to perform this analysis, we define the baseline problem shown in Figure 7, with 150 cities, a coverage radious of 0.15 and a total demand of 50. The locations of both the cities and the facilities have been randomly generated. The solution to this problem can also be observed in Figure 7. Concretely, the variables which will be modified throughout the sensitivity analysis are the magnitude of the coverage radius (r_i), the number of facilities (F) and the demand (D). After conducting these modifications, their effect on the total distance traveled by the salesman in the optimal solution and the computation time will be studied. Finally, it is worth to mention that all the computations have been done on a computer with a 2.70 GHz Intel Core-i7 processor and 8 GB RAM.

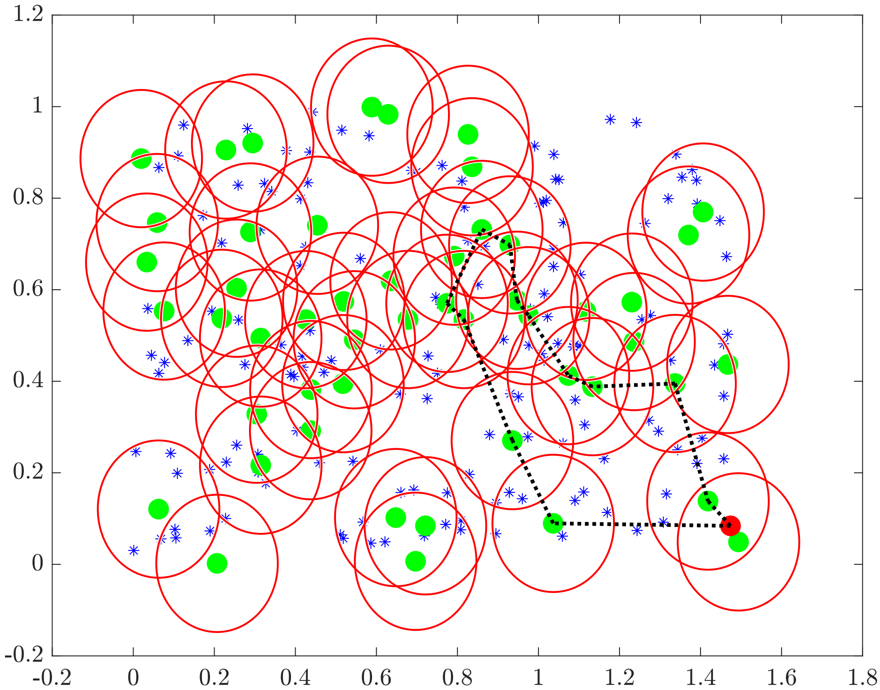


Figure 7: Solution to baseline problem

4.1 Modifying the coverage radius

When modifying the value of the coverage radius, the number of potential cities that can be supplied by a certain facility inherently changes. For the case of analysis (see Figure 7), the smallest r_i which still provides us with a feasible solution is 0.07 and, therefore, it will be the taken as a lower bound. On the other hand, a radius 0.21 will be chosen as an upper bound, as from this point on the change in the total distance and computation time is negligible. A *for loop* is implemented so that almost all possible values between the upper and lower bounds are tested.

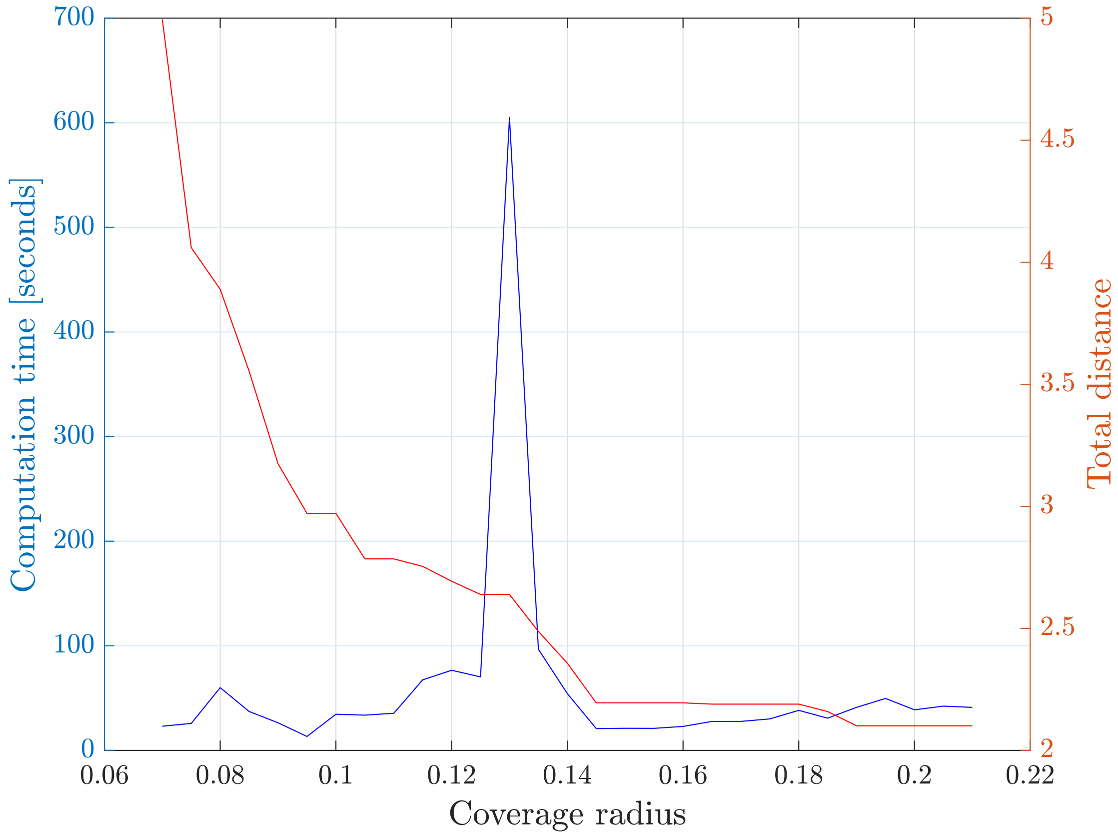


Figure 8: Total distance and computation time for varying coverage radii

The results obtained are shown in Figure 8. As it can be observed from the graph, the total distance decreases to an asymptotic value of approx. 2 when the coverage radius is increased. This clear decrease in the total distance traveled by the salesman as the value of the coverage radius increases is what would be expected from the model, as the possible ways in which the different cities can be assigned to the different facilities drastically increases. Thus, the length of the route traveled by the salesman remarkably diminishes. Moreover, the computation time does not show a clear relation to the variation of coverage radius. However, only with a coverage radius of 0.13 a substantial increase appears.

4.2 Modifying the number of facilities

When increasing the number of facilities the number of variables increases exponentially. Figure 9 shows the total distance and computation time for varying number of facilities. A small number of facilities of 10 is chosen as the lower bound and due to limitations in computation time the upper bound is set to 60. Just as the number of variables increases, the computation time also increases exponentially with increasing number of facilities. However, the total distance is lowest when the number of facilities is between 25 and 50.

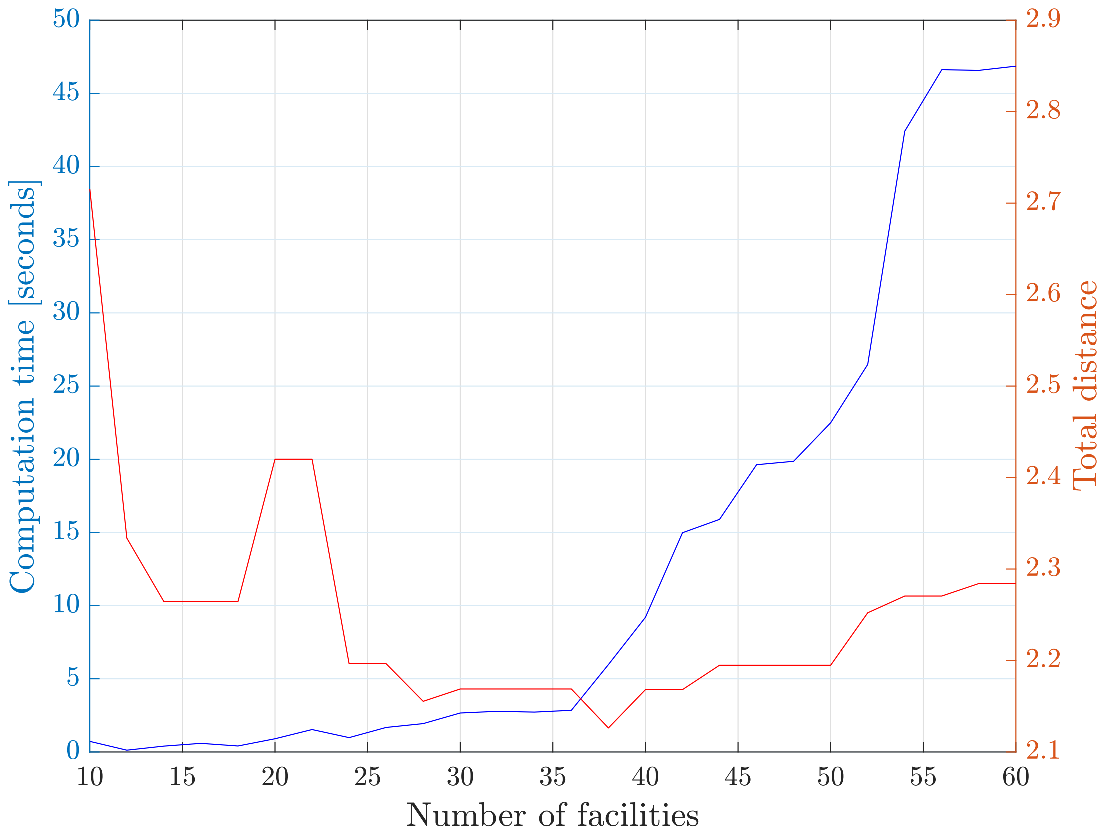


Figure 9: Total distance and computation time for varying number of facilities

4.3 Modifying the demand

When increasing the demand the number of cities which have to be supplied also increases. Figure 10 shows the total distance and computation time for varying number of facilities. A small demand of 10 is chosen as the lower bound and the upper bound is set to 140. As opposed to the case where the number of facilities is modified, the total distance is directly related to the demand. As it can be seen in the graph, the total distance increases nearly in a linear fashion with the increasing demand, as expected. The graph also shows that the computation time is highest for mid range demand values.

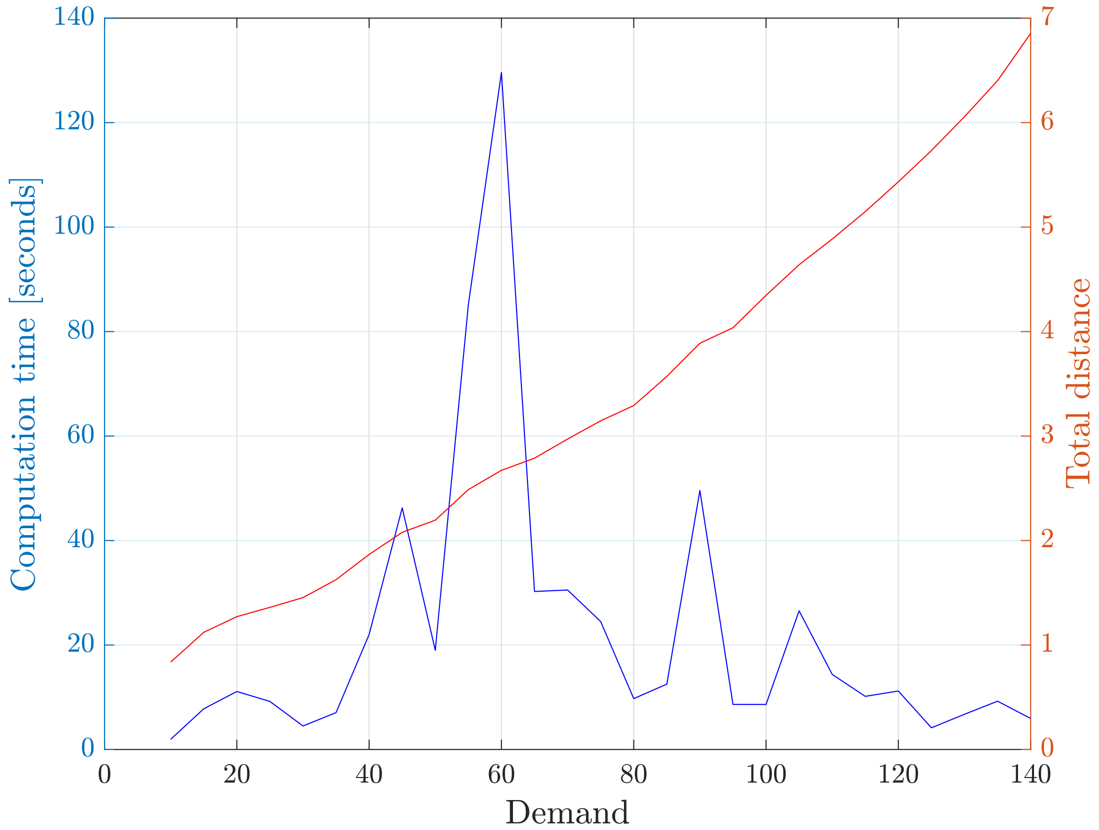


Figure 10: Total distance and computation time for varying demand

5 Conclusions and limitations

Throughout this project, we implemented a model to solve the generalised covering traveling salesman problem (GCTSP) as proposed in [1] by using the studied tools for the resolution of MILP problems as seen in class (see [2]). Thus, given a set of cities, facilities and a depot, we successfully implemented a set of constraints so that the salesman is able to satisfy a certain quantity of the total customers demand in a single tour of minimal length. Moreover, we checked the validity of our implementation by analysing the outcome of a simple case example. Finally, we performed a sensitivity analysis where we confirmed the expected behaviour of the model in terms of the total distance traveled by the salesman: increasing the coverage radius and the number of facilities decreases the total length of the tour, whereas increasing the demand, increases the total distance. We also analysed how the computation time varies with each of these modifications.

However, our work could be continued from this point on in order to generalise the model even more to cover a certain number of limitations that are present right now. For example, now we are only considering the possibility of having one only depot, but a more general version could include one or more starting points. Moreover, now we are considering a one-only-vehicle situation. Also, the fact that the variables are one-way instead of having two directed arcs linking each pair of facilities could also be improved. Finally, if a city belongs to the intersection between two facilities, it is automatically being assigned to its closest facility instead of considering the shortest-path option.

References

[1] Venkatesh, P., Singh, A., *An artificial bee colony algorithm with variable degree of perturbation for the generalized covering traveling salesman problem.* Applied Soft Computing Journal, 2019.

[2] Frederick S. Hillier and Gerald L. Lieberman *Introduction to Operations Research.* Mc Graw-Hill International Edition, 2014