Homework 1 - Vehicle Routing Formulations

Sebastián Cancino Espinoza

Abstract

The Capacitated Vehicle Routing Problem is a NP-Hard Problem widely studied in the literature. Many formulations have been proposed in order to achieve better and faster solutions. This work studied four of these formulations, with polynomial number of subtour elimination constraints: one node-based formulation (MTZ-L) and three arc-based formulations (GG, BHM and MCF). The formulations were tested in 10 instances obtained from sets A and P of the CVRPLIB using CPLEX. The results show the superiority of the GG and BHM over the MTZ-L and MCF formulations.

1. Introduction

The Vehicle Routing Problem (VRP) was proposed by Dantzig and Ramser in 1959. It consists on, given a set of K vehicles, finding K routes, one for each vehicle that goes from the depot to a number n of costumers and then return to the starting point. Each costumer must be visited only one time, and every costumer should be visited.

The capacitated version of this problem, known as Capacitated Vehicle Routing Problem, introduces a parameter Q that represents the capacity of the vehicles and now each costumer has a demand that must be satisfied. Given these new conditions, the routes to be found must take in consideration that a vehicle cannot satisfy a total demand greater than its capacity. This is, the summation of the demands of the costumers visited by a vehicle must be less or equal that the capacity of that vehicle. The goal is to minimize the total travel time, cost or distance.

The CVRP is a combinatorial, NP-hard problem that has been widely studied in the literature. In this work we compare several, well-known CVRP formulations, just as in Aksen A., Öncan T. and Sadati M. (2018) work. This experimental process is not as extensive as the mentioned work, because the main goal is not to replicate the paper but to test and prove its results and conclusions.

2. Models

Given a complete directed graph $G = (I, A)$, where $I = \{1, \ldots, n\}$ is the set of nodes and A is the set of arcs, the general mathematical formulation for the CVRP studied in this work is the one that follows.

Sets

0: $node\ representing\ the\ depot$ I: $set\ of\ nodes\ for\ costumers\ (1, \dots, n)$

I_0: $I \cup 0$

Parameters

Q: vehicle capacity in demand units.

q_i: demand or load to be delivered or collected from customer i. $(i \in I)$

d_{ij}: distance from node i to node j, $(i, j \in I_0, d_{ij} \neq d_{ji})$

Variables

$X_{ij} = 1$ if arc (i, j) is traversed by a vehicle, 0 otherwise $(i, j) \in I_0$.

Model

$$minimize\ Z = \sum_{i \in I_0} \sum_{j \in I_{0, i \neq j}} d_{ij} X_{ij} \tag{1}$$

$subject\ to$

$$\sum_{j \in I_0, i \neq j} X_{ji} = 1, \qquad i \in \mathrm{I} \tag{2}$$

$$\sum_{j \in I_0, i \neq j} X_{ij} = 1, \qquad i \in \mathrm{I} \tag{3}$$

$$Subtour\ Elimination + Feasible\ Tour\ Construction\ Constraints \tag{4}$$

$$X_{ij} \in \{0, 1\} \qquad i, j \in \mathrm{I}_0 \tag{5}$$

The goal, as was stated, is to minimize the total cost or distance of the total travels, represented by Z. Constraints (2) and (3) ensures that only one vehicle enters and one vehicle get out from each costumer. Constraint (4) ensures that there is no subtour in the solution. Several approaches for this constraint have been proposed in the literature. The main difference in the formulations studied in this work is exactly the approach to this constraint. We consider the MTZ-L, GG, BHM and MCF formulations. Now these four formulations and a brief explanation of each one is presented.

2.1. A node-based formulation: The Lifted MTZ

The Lifted MTZ is modeled as a collection problem. It uses node-specific decision variables U_i that represents the load of a vehicle right after departing from customer i. This set of constraints replaces constraint (4) in the original model.

$$U_i - U_j + QX_{ij} + (Q - q_i - q_j)X_{ij} \leq Q - q_j \qquad i,j \in \mathrm{I}, i \neq j \qquad (6)$$

$$U_i \geq q_j + \sum_{j \in I, j \neq i} q_j X_{ij} \qquad i \in \mathrm{I} \qquad (7)$$

$$U_i \leq Q - \sum_{j \in I, j \neq i} q_j X_{ij} \qquad i \in \mathrm{I} \qquad (8)$$

$$U_i \leq Q - (Q - q_i)X_{0i} \qquad i \in \mathrm{I} \qquad (9)$$

$$U_i \leq Q - (Q - \max_{j \in I, i \neq j}\{q_j\} - q_i)X_{0i} - \sum_{j \in I, j \neq i} q_j X_{ij} \qquad i \in \mathrm{I} \qquad (10)$$

$$U_i \geq 0 \qquad i \in \mathrm{I} \qquad (11)$$

These constraints ensure the validity of the definition of the variable U_i, and that the capacity of the vehicles is never exceeded. From now, this formulation is referred to as MTZ-L.

2.2. Single-commodity flow formulation

The single-commodity flow formulation is based on the model proposed in 1978 by Gavish and Graves for the Travelling Salesman Problem. It uses a decision variable F_{ij} $(i \in I, j \in I)$ that represents the amount of goods that flows from customer i to customer j, after collecting the load at customer i. Just like the previous formulation, this set of constraints replaces constraint (4).

$$\sum_{j \in I_0, i \neq j} F_{ij} + q_i = \sum_{j \in I_0, i \neq j} F_{ij} \qquad i \in \mathrm{I} \qquad (12)$$

$$F_{ij} \geq q_j X_{ij} \qquad i,j \in \mathrm{I}_0, i \neq j \qquad (13)$$

$$F_{ij} \leq (Q - q_j) X_{ij} \qquad i,j \in \mathrm{I}_0, i \neq j \qquad (14)$$

$$F_{ij} \geq 0 \qquad i,j \in \mathrm{I}_0 \qquad (15)$$

Constraint (12) is the flow conservation constraint at each customer node. (13) and (14) are lower and upper bounds for the new decision variable. This formulation is denoted as GG from now on.

2.3. Two-commodity flow formulation

For this formulation, a new set $I_0' = \{0, 1, \dots, n, n+1\}$ is defined. A dummy node, corresponding to a copy of the depot is added to the model, represented by the $(n+1)$ index. A new variable G_{ij} is created to represent a directed flow path from i to j.Hence, variables G_{ij} and G_{ji} represents two different flows in opposite directions, one that goes from node 0 to $(n+1)$, and the other that goes from node $(n+1)$ to node 0. Whenever $X_{ij} = 1$, $G_{ij} = Q - G_{ji}\ holds\ true.$

$$minimize\ Z' = \sum_{i \in I_0} \sum_{j \in I_0', i \neq j} d_{ij} X_{ij} + \sum_{i \in I} d_{i(n+1)} X_{i(n+1)} \qquad (16)$$

$subject\ to$

$$\sum_{j \in I_0, i \neq j} (G_{ji} - G_{ij}) = 2q_i \qquad i \in \mathrm{I} \qquad (17)$$

$$\sum_{j \in I} G_{0j} = \sum_{j \in I} q_j \qquad (18)$$

$$\sum_{j \in I} G_{j0} = KQ - \sum_{j \in I} q_j \qquad (19)$$

$$\sum_{j \in I} G_{(n+1)j} = KQ \qquad (20)$$

$$G_{ij} + G_{ji} = QX_{ij} \qquad i,j \in I_0', i \neq j, i \neq (n+1) \qquad (21)$$

$$\sum_{j \in I'_0, j>i} X_{ij} + \sum_{j \in I'_0, j<i} X_{ji} = 2 \qquad i \in \mathrm{I} \qquad (22)$$

$$G_{ij} \geq 0 \qquad i, j \in I'_0 \qquad (23)$$

$$X_{ij} \in \{0, 1\} \qquad i, j \in I'_0 \qquad (24)$$

Note that (16) is exactly twice the objective function of previous models (1). This model is referred to as BHM from now, in recognition to its original authors (Baldacci, Hadjiconstantinou and Mingozzi, 2004).

2.4. Multi-commodity flow formulation

The fourth and last model considered was proposed by Letchford and Salazar-Gonzalez in 2015 and uses two extra decision variables F_{ij}^k and G_{ij}^k. These are binary variables are defined for each costumer $k \in I$ and each arc $(i, j) \in A$. F_{ij}^k takes value 1 if and only if a vehicle traverses (i, j) on the way from the depot to k. F_{ij}^k takes value 1 if and only if a vehicle traverses (i, j) on the way from the depot to k. G_{ij}^k takes value 1 if and only if a vehicle traverses (i, j) on the way from k to the depot. This formulation can be written replacing (4) with the following set of constraints:

$$\sum_{j \in I} F_{0j}^k - \sum_{j \in I} F_{j0}^k = 1 \qquad k \in \mathrm{I} \qquad (25)$$

$$\sum_{j \in I_0, j \neq i} F_{ij}^k - \sum_{j \in I_0, j \neq i} F_{ji}^k = 1 \qquad i, k \in \mathrm{I}, \mathrm{i} \neq \mathrm{k} \qquad (26)$$

$$\sum_{j \in I} G_{j0}^k - \sum_{j \in I} G_{0j}^k = 1 \qquad k \in \mathrm{I} \qquad (27)$$

$$\sum_{j \in I_0, j \neq i} G_{ij}^k - \sum_{j \in I_0, j \neq i} G_{ji}^k = 1 \qquad i, k \in \mathrm{I}, \mathrm{i} \neq \mathrm{k} \qquad (28)$$

$$F_{ij}^k + G_{ij}^k \leq X_{ij} \qquad k \in I, (i, j) \in I'_0, i \neq j \qquad (29)$$

$$\sum_{k \in I \setminus \{i, j\}} q_k \left(F_{ij}^k + G_{ij}^k\right) \leq (Q - q_i - q_j) X_{ij} \qquad (i, j) \in I'_0, i \neq j \qquad (30)$$

$$F_{ij}^k,\ G_{ij}^k \geq 0 \qquad k \in I, (i,j) \in I'_0 \tag{31}$$

This formulation is referred to as MFC from now on.

3. Computational Result

Ten instances of the CVRP problem were solved using the four different formulations. The instances were obtained from the Capacitated Vehicle Routing Problem Library (CVRPLIB). In particular, five instances from the Set A (Augerat, 1995) and five instances from Set P (Augerat, 1995) were selected. These instances use two-dimensional Euclidean distances, rounded to the nearest integer.

The models were solved in CPLEX and programmed in Python 3.6.8 using the DOcplex interface, on an Intel Celeron N2920 1.86 GHz CPU with 8 GB of RAM. For all models and instances a time limit of 2000 seconds of clock time was set. At this time the best integer solution found by the model was compared with the best-known solution provided by the CVPLIB. These results are presented in Table 1.

Table 1. Computation Results

		MTZ-L		GG		BHM		MCF	
Instance	BKS	Z	T	Z	T	Z	T	Z	T
P-n16-k8	450	450*	17.8	450*	12.0	450*	10.2	450*	17.8
P-n19-k2	212	212	2000	212*	21.9	212*	9.0	219	2000
P-n20-k2	216	217	2000	216*	25.9	216*	4.9	221	2000
P-n23-k8	529	529	2000	529*	679.3	529*	255.6	529*	1064.4
A-n32-k5	784	784	2000	784	2000	784*	180.9	2176	2000
A-n37-k6	949	991	2000	949	2000	1012	2000	3122	2000
P-n40-k5	458	470	2000	458	2000	458*	139.5	1468	2000
A-n44-k6	937	963	2000	955	2000	1041	2000	3265	2000
A-n46-k7	914	944	2000	914	2000	914	2000	3518	2000
A-n62-k8	1288	1426	2000	1322	2000	2857	2000	7098	2000

Z: objective value for the best integer solution found. (*) indicates that the solver converged to optimality within the time limit.

T: total elapsed time of computation in clock seconds. 2000 is the time limit.

Note that, although several models found the optimal solution within the time limit, the solver didn't end exploring all the branches of the Branch & Bound tree in all cases. For this reason, optimality can't be ensured.

The MTZ-L formulation reached the time limit in almost all instances, but the best-found solution is relatively close to the optimal solution. In fact, it reached the optimal solution (not converged to the optimal solution) in four of the ten cases.

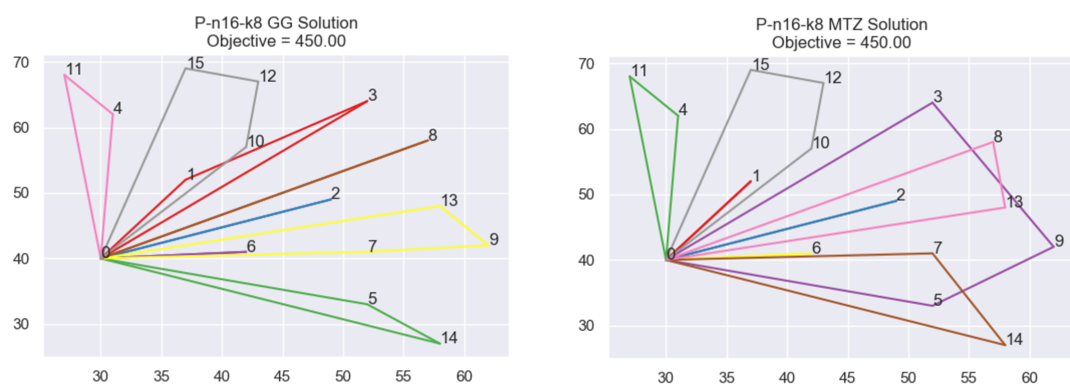
The GG formulation found the optimal solution in 8 instances (4 of them converged within the time limit). All instances with more than 30 nodes reached the time limit with a best-found solution relatively close to the optimal value.

The BHM formulation outperformed all other formulations, finding an optimal value for 6 of the 10 instances. It has the lower computation times for the instances that converged. Around 40 nodes (customers) the formulation reached the time limit for the instances considered.

The MCF formulation seems to work well only for greater values of K. It only converged to optimal solution for instances with $K = 8$ and, for these instances, the total time of computation is significantly bigger than the elapsed time for BHM and GG formulations. For the other instances, the best-found solutions are fare away from the optimal values.

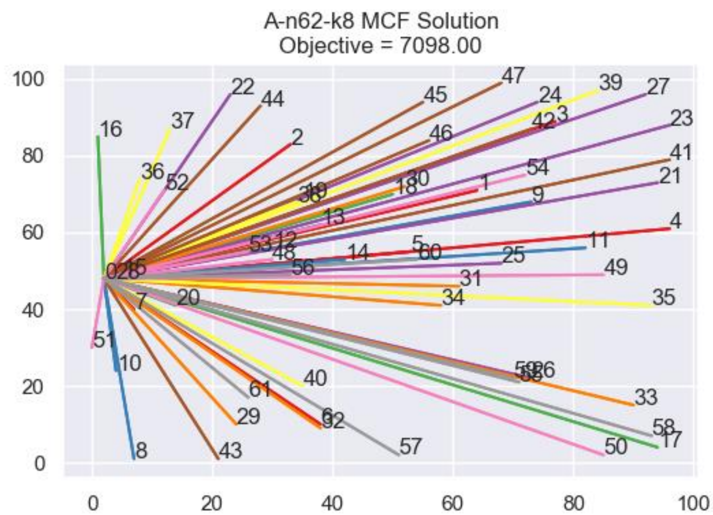
Some interesting findings are that for some instances, more than one optimal route was found. Figure 1 presents two different paths found by the MTZ-L and GG formulations. Although the routes are different, the values for the objective function are the same in both cases.

Figure 1: Plot of the MTZ-L and GG solutions for P-n16-k8 instance



On the other hand, looking at the plots for the MCF solutions, we can see that in all instances that reached the time limit, the formulation failed to found a feasible solution for the problem. Figure 2 present an example of this for the biggest instance evaluated. In this solution, all nodes have their own routes from the depot. This gives a total of 61 routes, while the instance only allows $K = 8$ routes.

Figure 2: MCF Solution for A-n62-k8



Figures for all solutions can be found in the attached output folder.

4. Conclusions

The results of the original authors are replicated in this work. The superiority of GG and BHM formulations is clear. While BHM takes less time on small instances, the GG formulation gives better solutions for bigger instances. The MTZ-L and MCF performs far worse than GG and BHM for almost all instances. This leads to that, using The BHM formulation outperformed all other formulations, finding an optimal value for 6 of the 10 instances. Using a cheap CPU as in this case, the BHM, and GG formulations can still reach relatively good solutions in a reasonable time, that are usable in many simple real-world applications of this problem.