Risk and Logistics Homework 1

Wei Li s2346729 Anqi He s2308848 Qiqi Xiao s2323497

February 15, 2023

1 Question 1 *Total Cost Function*

Total Cost is consists of 3 parts which are **setup**, **operating** and **transportation** costs.

- **Setup**

  In our case, $Y(j) = t$ means on *location j*, we build a warehouse in *t period*. If we do not build this warehouse in all periods on *location j*, then set $Y(j) = 11$. Therefore, we can just sum the correspond setup costs for those locations whose $Y(j)$ value is less than 11.

- **Operating**

  If we build a warehouse at *location j*, we need to pay every years' operating costs for this warehouse since the year it build to the end of the last year.

- **Transportation**

  As the function is to compute the total cost for certain warehouses built in certain periods(these are independent variables $Y(j)$ for this function), and our aim is to find the minimal costs to satisfy all demands. Under given Y_j, the operating and setup costs are 'fixed', so we need to find the minimal transportation cost. We find 2 methods to calculate transportation costs.

  The first is, for each periods, for each customers, for each product groups, calculate all distance from this customer to every built warehouse(s) then to every suppliers who can supply this kind of product, then finding the minimize distance. And we allocate this customer's this kind of product to the warehouse we find, and allocate this demand to the supplier we find.

  The second is, for each periods, for each customers, we search all distances from this customer to every built warehouse(s) and find the shortest path. Then we allocate this customer to the warehouse in this path. All warehouses can sum every kinds of products they need to satisfy allocated customers. Then, find the nearest suppliers in each product groups to warehouses.

  There is one line of code provided by tutor.

```
    ! The vector of supplier allocations
    Z: array(Locations,Products) of integer
```

The upper code means that we can only arrange 1 supplier to a warehouse for 1 kind of product. Therefore, we choose the second method, as the first 1 may lead to arrange more than 1 suppliers in 1 kind of product to 1 warehouse. And the second method likes realistic more, customers have demands, then they search for the nearest warehouses(supermarkets), warehouses need to satisfy those demand, then they search for the neareast suppliers.

- **Function Conclusion**
  $TotalCosts = Setup + Opeating + Transportation$

  $Setup = \sum SetupWarehouses(j) \quad \text{where } j \in J \setminus \{j : Y(j) = 11\}$
  $Operating = \sum(10 - 1 + Y(j)) \times OperatingWarehouses(j) \quad \text{where } j \in J \setminus \{j : Y(j) = 11\}$

  $Transportation = Trans\_SW(supplier- > warehous) + Trans\_CW(warehouse- > customer)$

  $Trans\_CW = \sum_{t\in P, i\in C} Cost(i,j) \times (\sum_{k\in G} Demand(i,k)) \quad j\ satisfy\ (min\ Distance(i,j))$
  $weight\_CW(i,k,j) = demand(i,k,j) \quad j\ satisfy\ (min\ Distance(i,j))$

  $weight\_SW(j,s) = \sum_{i\in C} weight\_CW(i,k,j)$ s satisfy (min $SupplierDistance(j,s)$) and provide corresponding product k.

  $Trans\_SW = \sum_{j\in J, s\in S} weight\_SW(j,s) \times CostSupplier(j,s)$

  where $P : Period\ \{1..10\}$, $C : Customer\ \{1..440\}$, $j \in J \setminus \{j : Y(j) \le p$, $G : ProductGroup\ \{1..4\}$, $s \in S : Supplier\{1..53\}$

- **Result**
  For file "CaseStudyData.txt", the result is 2346525.915.
  For file "CaseStudyDataAggregated.txt", the result is 2346727.605.

2 Question 2 *Construction Heuristic*

We choose greedy heuristic, because greedy heuristic has an omega rule to make the iteration stop when the total cost cannot be improved.

Our basic concept is first initialize all $Y_j = 11$ (j: Location number, Y(j): j location warehouse is built in Y period($Y(j) \le 10$ if j is built)).Than run the iteration, each iteration add 1 warehouse into our solution to find the minimize total cost.(i.e. in iteration 1, choose 1 warehouse to built in Y year; in iteration 2 which can minimize total cost, choose 2 warehouses to built in corresponding Y year, which can minimize total cost and the total cost is less than previous iteration(s) choice;...). In the end, we will choose i element from feasible region($Period \times Locations = 4400$) and get a minimal total cost by every iterations' comparations.

Besides, we set 1 constraint that model have to arrange 1 warehouse built in period 1 in each times' iteration to satisfy customers' demands.And if a warehouse is built in i year($i \leq 10$), then it cannot be built again in other years. Iterations will be stopped when it reaches omega rule(cannot find a better total cost).

Both the selected Selected location and total cost from greedy heuristic from 'CaseStudyData.txt' and 'CaseStudyDataAggregated.txt' are shown below. From both dataset, we choose to build warehouse at location AB10 in Period 1. Besides, the total cost is 1401424.768.

```
Selected Locations: Y(AB10) = 1
Total cost of Greedy hueristic in 10 years: 1401424.768
time:   3923.075s
```

Figure 1: Selected location and total cost from greedy heuristic:CaseStudyData.txt

```
CaseStudyDataAggregated.txt
Selected Locations: Y(AB10) = 1
Total cost of Greedy hueristic in 10 years: 1401424.768
time:   35.854s

Process exited with code: 0
```

Figure 2: Selected location and total cost from greedy heuristic:CaseStudyDataAggregated.txt

3 Question 3 *The Interchange Heuristic*

The interchange heuristic with first improvement strategy was implemented to improve the results. For each selected location, the heuristic aims to find if there is a better location to replace it which can lead to a lower total cost. To simplify the problem, we consider the new locations will be set up in the same year as the original ones. In addition, we defined the neighborhoods only in terms of distance between two locations and in this case, only the locations that are within 200 far from the original ones can be selected to exchange. This threshold works to keep the number of neighborhoods within a reasonable limits and improve the efficiency of each iteration at the same time.

We go through each location in the starting solution and try to find a better location that is outside the solution to exchange in each iteration. If the total cost with the potential location is lower than that with the selected one, the results can therefore be improved by exchanging these two. Using the first improvement strategy, we will start the next iteration when a improving location is found in this iteration instead of finding the best location that can reduce total cost the most. The interchange heuristic will not change the number of locations and we assume that the set up period for each improving pair is the same so that the number of final selected locations set up in each year should be the same as the starting solutions.

1. The starting solution with the following selected locations

$$Y(DG1) = 1,\ Y(G12) = 3,\ Y(IV36) = 5,\ Y(ML8) = 7,\ Y(PH17) = 9$$

(a) After interchange heuristic with the data set "CaseStudyData.txt", the locations have turned into

$$Y(AB30) = 1,\ Y(G12) = 3,\ Y(IV36) = 5,\ Y(ML8) = 7,\ Y(PH17) = 9$$

where the location AB30 has replaced the original one DG1. The total cost have reduced from 3366110.121 to 3312345.279 so the solution has been improved.

```
After the interchange heuristic, the total cost is 3312345.279

Y(AB30) = 1
Y(G12) = 3
Y(IV36) = 5
Y(ML8) = 7
Y(PH17) = 9
time:   3248.261319
```

Figure 3: Final solutions for selected locations with data set CaseStudyData.txt

(b) Run the heuristic on the data set "CaseStudyDataAggregated.txt", we obtain the final set:

$$Y(DG1) = 1,\ Y(G12) = 3,\ Y(IV36) = 5,\ Y(ML8) = 7,\ Y(PH17) = 9$$

The heuristic fails to find any improving pair in this case and the total cost stays the same at 3366110.121.

```
The original total cost is 3366695.281
No improving pair found -> STOP
After the interchange heuristic, the total cost is 3366695.281
The final set of solutions is:
Y(DG1) = 1
Y(G12) = 3
Y(IV36) = 5
Y(ML8) = 7
Y(PH17) = 9
time:   31.82203846
```

Figure 4: Final solutions for selected locations with data set CaseStudyDataAggregated.txt

2. We will now use the solutions obtained from the last question as the starting solutions. The starting solution is as follows:

$$Y(AB10) = 1$$

(a) Run the heuristic on the data set "CaseStudyData.txt", we obtain the final set:

Y(AB11) = 1

The heuristic find an improving pair (AB10, AB11) in this case and the total cost reduce from 1401424.768 to 1378328.004.

```
The original total cost is 1401424.768
No improving pair found -> STOP
After the interchange heuristic, the total cost is 1378328.004
The final set of solutions is:
Y(AB11) = 1
time:   654.8174593
```

Figure 5: Final solutions for obtained locations with data set CaseStudyData.txt

(b) Run the heuristic on the data set "CaseStudyDataAggregated.txt", we obtain the final set:

Y(AB14) = 1

The heuristic find an improving pair (AB10, AB14) in this dataset and the total cost reduce from 1401424.768 to 1034212.927, experiencing a sharp decrease.

```
The original total cost is 1401424.768
No improving pair found -> STOP
After the interchange heuristic, the total cost is 1034212.927
The final set of solutions is:
Y(AB14) = 1
time:   5.490795209
```

Figure 6: Final solutions for obtained locations with data set CaseStudyDataAggregated.txt

Further Discussion:

This time, we tried to fix time and interchange different locations. (to interchange input solution, using the same building time but change where to locate to see if the total cost is decrease or not). And we set a limited distance to define the range of neighbours. The change of distance value will influence the efficiency of the model.

We also considered about some other methods. However, we do not have enough time to encode and run them this time, but we will try them after class. The methods are:

1. fixing the input locations and interchange periods.

2. Using another way to measure the "distance"(using in defining neighbours). We can use $\sqrt{weight_1 \times Distance^2 + weight_2 \times (period\_orginal - period\_interchange)^2}$ (here distance means distance between locations from given data). Setting 2 weights is to determine the importance of distance and periods' influence or adjust the difference of 2 variables(as the maximum value from 1 period to another is $11 - 1 = 10$, but distance usually higher than 100). By this way, we can change locations and periods at same time, and with the upper rule to define neighbours, the efficiency of the model will not be too low.