Heuristic approach for solving capacitated location-routing problems

H N Ahmad¹, T Bakhtiar¹, P T Supriyo¹

¹Division of Operations Research, Department of Mathematics, Bogor Agricultural University, Kampus IPB Dramaga, Bogor 16680, INDONESIA

Corresponding author: afiesnugraha1@gmail.com

Abstract. This paper deals with the standard capacitated location-routing problem (CLRP), where two types of interdependent decisions must be made: which facilities among a number of potential ones should be operated and which vehicle routes should be built to fulfil the demand of customers from the operated facilities using a given fleet. We propose a heuristic mechanism to address the problem. The algorithm consists of four stages, namely optimally clustering customers by $k$-means and assessed by Dunn index, selecting the nearest depot to be opened based on Euclidean distance calculation, customers-to-depot matching and searching the optimal distribution route. This heuristic algorithm is then implemented to a dataset comprises of 20 customers and 5 potential depot facility locations.

1. Introduction

The Vehicle Routing Problems (VRP) is a terminology which refers to a problem of searching routes for a fleet of vehicles of known capacities to service a number of customers with known locations and known demands for a certain commodity, given a set of constraints, possibly integrated with production scheduling [1]. While, Facility Location Problems (FLP) concerns with selecting the placement of a facility (often from a list of integer alternatives) to minimize transportation costs and to best meet the demanded constraints [2]. If VRP and FLP are mutually considered then we have the so-called location-routing problems (LRP). In fact, classical LRP integrates the two kinds of decisions, namely opening a subset of depots, assigning customers to them and determining vehicle routes, to minimize a total cost including the cost of open depots, the fixed costs of vehicles used, and the total cost of the routes [3]. Both exact and heuristic approaches can be employed in solving LRP.

Book chapter by Marinakis [4] provides the concise reference for the basic notion of LRP that includes the model and its variants as well as the solution method. While, book by Drezner and Hamacher [2] offers more comprehensive review. Prodhon and Prins [3] analyzes the recent literatures on the standard LRP and new extensions such as several distribution echelons, multiple objectives or uncertain data, including results of state-of-the-art metaheuristics method. A more recent survey paper on standard LRP is given by Schneider and Drexel [5] as it provides concise paper excerpts that convey the central ideas of each work, discuss recent developments in the field, provide a numerical comparison of the most successful heuristic algorithms, and list promising topics for further research. Farahani et al. [6] delivers a review on recent efforts and development in multi-criteria location problems in three categories including bi-objective, multi-objective and multi-attribute problems and their solution methods.

Throughout the paper, the following main characteristics are adopted [3]:
All relevant data are deterministic, i.e., are fully known in advance.
There is only one planning period, i.e., a static planning situation.
The number of potential locations for facilities is finite.
There is only one single, scalar objective function.
Each customer has a known demand that must be fulfilled by a delivery from exactly one of the potential facilities; there is no load transfers at intermediate locations allowed.
Each customer must be visited exactly once by one vehicle.
No inventory considerations apply, neither at facilities nor at customers.

2. Review on solution approach

From the perspective of solution method, a number of efficient algorithms and approaches have been introduced for addressing LRP, stimulated by the fact that it is an NP-hard combinatorial optimization problem that can be exactly solved only for limited instances of the problem. Belenguer et al. [7] is one proposing an exact approach based on branch-and-cut method to solve CLRP. While Tuzun and Burke [8] proposed a heuristic method based on two-phase tabu search. In the first phase, the best composition of opening depot was sought following by vehicle routing in the second phase. Prins et al. [9] utilized a greedy randomized adaptive search procedures (GRASP) equipped by learning process to decide depots should be opened and path relinking for routing. Prins et al. [10] suggested a solving approach by using lagrangean relaxation and granular tabu search. In the same effort, Duhamel et al. [11] also utilized GRASP and evolutionary local search (ESL) to solve the problem. Barreto et al. [12] proposed a clustering based analysis to address CLRP. Other heuristic approaches include the use of ant colony optimization [13, 14], hybrid genetic algorithm [15] and particle swarm optimization [16].

3. Method

In this work we consider the capacitated location-routing problem (CLRP), the most basic and general variant of LRP by adding capacity constraint on both depots and vehicles. We formulate the problem in term of mixed integer linear programming (MILP) and approaches heuristically by k-means clustering for grouping customers. Dunn index is measured to determine the optimal number of clusters. Our procedure consists of four stages: optimally clustering customers by k-means, selecting the best depot, allocating customers to selected depot, and searching the optimal distribution route.

3.1. Dataset

For algorithm testing, we consider a hypothetical dataset presented in [10]. The dataset consists of 20 customers with known level of demand and 5 candidates of facility with homogeneous capacity of 140 units. The scatter plot of nodes for dataset is provided in Figure 1. In this plot, black diamonds denote the customers and squares represent the potential locations of facilities.

3.2. Customers clustering

Clustering of customers is the first step of our heuristic algorithm. We employ the well-known k-means clustering for grouping customers with \(2 \leq k \leq \lceil n/4 \rceil\), where \(n\) is the number of customers. The best number of clusters \(k\) is determined by means of Dunn index. Let \(C = \{C_1, C_2, ..., C_k\}\) be a set of clusters, let \(\delta: C \times C \to \mathbb{R}^+\) be a cluster-to-cluster distance measure and let \(\Delta: C \to \mathbb{R}^+\) be a cluster diameter measure. The Dunn index \(DI\) for the set \(C\) is defined as [18]:
Figure 1. The location of customers and potential locations of facilities in the third dataset

\[
D I(C) = \frac{\min_{i \neq j} \{\delta(C_i, C_j)\}}{\max_{i \in S} \{\Delta(C_i)\}},
\]

where \(\delta(C_i, C_j) = \min_{x \in C_i, y \in C_j} \{d(x, y)\}\) and \(\Delta(C_i) = \max_{x, y \in C_i} \{d(x, y)\}\) with \(d: C \times C \rightarrow \mathbb{R}^+\) is object-to-object distance measure like Manhattan distance or Euclidean distance. For a given assignment of clusters, a higher Dunn index indicates better clustering. The process of customers clustering is required to decide the number of depots should be operated.

3.3. Selecting the depot facilities

The mechanism for selecting the best depot facilities relies on the calculation of Euclidean distance between centroid of cluster and depot facility. Let \((a_i, b_i)\) denotes the centroid of cluster \(C_i\) and \((p_j, q_j)\) denotes the location of (potential) depot facility \(j\), then the depot facility \(j\) will be operated as long as its total distance \(D_j\) is minimum:

\[
D_j = \min_{1 \leq m \leq k} \sum_{i=1}^{k} \sqrt{(p_j - a_i)^2 + (q_j - b_i)^2},
\]

where \(m\) is the number of (potential) depot facilities. If the total demand of all customers cannot be supplied by depot \(j\), then the second best depot, i.e., depot facility with the next minimum total distance, should be opened. The process of opening next depots is continued until the total demand satisfied.

3.4. Customer-to-depot allocation

Customer-to-depot assignment is undertaken by measuring the Euclidean distance of each customer to each selected depot. A number of customers with minimum distance are then assigned to the nearest selected depot. The process of matching is continued until all customers assigned to depots. In this process, a customer reallocation might be needed regarding the capacity reachability of each selected depot.
3.5. Vehicle routing

The last step of the algorithm is searching the best routes from depot to customers for delivering products. A fleet of vehicles, each with known and fixed capacity, starts at a designated depot and returns to the same depot after visiting customers where service or product is demanded. The objective is either to minimize the total distance of all the routes or total cost.

Since the customers demand fulfilment by depot is guaranteed in the previous step, then in this paper we consider the most standard capacitated VRP. We here assume that a fleet of vehicles with known loading capacity is always available in each depot to deliver products. One may refer to Toth and Vigo [17] for a comprehensive account on this topic.

Suppose $\mathbb{C}$ denotes the set of all customers and $\mathbb{V}$ denotes the set of all vehicles available in the depot. If the number of customers is $n$, and thus $\mathbb{C} = \{1,2,\ldots,n\}$, then the set of all nodes, i.e., depot and customers, is denoted by $\mathbb{N} = \{0\} \cup \mathbb{C} = \{0,1,2,\ldots,n\}$, where 0 is index for the only depot. The following parameters are used: $c_{ij}$ is the traveling cost between node $i$ and $j$, $W$ and $Q$ are capacities of depot and vehicle, respectively, and $d_j$ is the demand level of customer $j$. For decision variable, $x_{ijk} = 1$ if vehicle $k$ travel from node $i$ to node $j$, $x_{ijk} = 0$ otherwise. VRP can be formulated such that the following total cost is minimized:

$$
\min_{x_{ijk}} Z := \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} \sum_{k \in \mathbb{V}} c_{ij} x_{ijk},
$$

subject to the constraints

$$
\sum_{j \in \mathbb{N}} \sum_{k \in \mathbb{V}} x_{ijk} = 1, \quad \forall j \in \mathbb{C},
$$

$$
\sum_{i \in \mathbb{N}} \sum_{k \in \mathbb{V}} x_{ijk} = 1, \quad \forall i \in \mathbb{C},
$$

$$
\sum_{j \in \mathbb{C}} \sum_{i \in \mathbb{N}} \sum_{k \in \mathbb{V}} d_j x_{ijk} \leq Q, \quad \forall k \in \mathbb{V},
$$

$$
\sum_{j \in \mathbb{C}} d_j \leq W,
$$

$$
\sum_{j \in \mathbb{C}} x_{ijk} - \sum_{j \in \mathbb{C}} x_{jik} = 0, \quad \forall i \in \mathbb{N}, \forall k \in \mathbb{V},
$$

$$
\sum_{j \in \mathbb{C}} x_{ok} \leq 1, \quad \forall k \in \mathbb{V},
$$

$$
\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{C}} x_{ijk} \leq |\mathbb{S}| - 1, \quad \forall \mathbb{S} \subseteq \mathbb{C}, \forall k \in \mathbb{V}.
$$

Constraints (4) and (5) ensure each customer is visited and exited once. Constraint (6) guarantees that the total delivered product is not exceeding the capacity of vehicle. While condition (7) requires the total demand of all customers is not exceeding the capacity of the depot. Constraint (8) enforces the route continuity, meaning that as soon as a vehicle reaches a customer to deliver products, it should leave that place at once. The requirement that all trips must be started at depot is given by (9), while (10) eliminates sub-tour.

4. Results and discussion

We implement our formulation and algorithm to a dataset that consists of 20 customers and 5 depot candidates. It is assumed that the amount of products demanded by customers vary but the capacity of depots are homogeneous. Each depot has a sufficient number of vehicles to deliver the products, however the loading capacity of vehicles are assumed to be the same, which is 70 unit/vehicle. Table 1
provides the location of all nodes as well as their demand or capacity levels, while Figure 1 depicts the scatter plot.

Based on $k$-means clustering analysis on customers location as well as their respecting Dunn index, it is suggested that customers can be best clustered into four clusters as indicated by its highest Dunn index (see Table 2). Under this 4-group clustering, it is known that clusters $C_1$ and $C_3$ consist of 7 customers, cluster $C_2$ has 4 customers and cluster $C_4$ has only 2 members (see Table 3).

| Node   | Coordinate $x$ | Coordinate $y$ | Demand (Capacity) |
|--------|----------------|----------------|-------------------|
| Customer 1 | 20  | 35  | 17  |
| Customer 2 | 8   | 31  | 18  |
| Customer 3 | 29  | 43  | 13  |
| Customer 4 | 18  | 39  | 19  |
| Customer 5 | 19  | 47  | 12  |
| Customer 6 | 31  | 24  | 18  |
| Customer 7 | 38  | 50  | 13  |
| Customer 8 | 33  | 21  | 13  |
| Customer 9 | 2   | 27  | 17  |
| Customer 10 | 1   | 12  | 20  |
| Customer 11 | 26  | 20  | 16  |
| Customer 12 | 20  | 33  | 18  |
| Customer 13 | 15  | 46  | 15  |
| Customer 14 | 20  | 26  | 11  |
| Customer 15 | 17  | 19  | 18  |
| Customer 16 | 15  | 12  | 16  |
| Customer 17 | 5   | 30  | 15  |
| Customer 18 | 13  | 40  | 15  |
| Customer 19 | 38  | 5   | 15  |
| Customer 20 | 9   | 40  | 16  |
| **Total** | | | **315** |

| Node   | Coordinate $x$ | Coordinate $y$ | Demand (Capacity) |
|--------|----------------|----------------|-------------------|
| Depot 1 | 6   | 7   | 140 |
| Depot 2 | 19  | 44  | 140 |
| Depot 3 | 37  | 23  | 140 |
| Depot 4 | 35  | 6   | 140 |
| Depot 5 | 5   | 8   | 140 |
| **Total** | | | **700** |

| Number of clusters | 2  | 3  | 4  | 5  |
|---------------------|----|----|----|----|
| Dunn index          | 0.166 | 0.155 | 0.253 | 0.181 |

Table 1. Nodes location, demand and capacity

Table 2. Dunn index
Table 3. Result of 4-group clustering

| Cluster | Centroid x | y | Customer          |
|---------|------------|---|-------------------|
| C₁      | 16.286     | 40| 1, 4, 5, 12, 13, 18, 20 |
| C₂      | 4          | 25| 2, 9, 10, 17      |
| C₃      | 25.714     | 18.143| 6, 8, 11, 14, 15, 16, 19 |
| C₄      | 33.500     | 46.500| 3, 7              |

Table 4. Distance between centroid and depot

| Cluster | Depot 1 | Depot 2 | Depot 3 | Depot 4 | Depot 5 |
|---------|---------|---------|---------|---------|---------|
| C₁      | 34.566 | 4.834   | 26.797  | 38.810  | 33.931  |
| C₂      | 18.111 | 24.207  | 33.060  | 36.359  | 17.029  |
| C₃      | 22.645 | 26.715  | 12.286  | 15.286  | 23.064  |
| C₄      | 48.130 | 14.714  | 23.759  | 40.528  | 47.901  |
| Total   | 123.452| 70.470  | 95.903  | 130.984 | 121.926 |

Table 4 provides the distance between centroid of each cluster and depot candidates. Depot 2 has the smallest total distance of 70.470. Thus, Depot 2 with capacity 140 units is selected for opening. However, according to Table 1 the total demand is 315 units. It means that in addition to Depot 2, operating two more depot is required such that the total capacity of depots is 420 units. Depots 3 and 5 which have the next smallest total distances should be opened. Table 5 gives the distance between each customer and Depot 2, 3 and 5, from which we can decide the customer-depot matching based on the smallest distance. At this stage, the total demand for Depot 2 is 171 units, which is exceeding its capacity. While those for Depots 3 and 5 are 73 and 71 units, respectively. It is imperative to reallocate a few number of customers of Depot 2 to either Depot 3 or Depot 5. The last two columns of Table 5 calculates additional distance incurred by customer reallocation from Depot 2 to Depots 3 and 5. Thus, it is suggested to reallocate Customers 17 and 2 to Depot 5, respectively. Consequently, total demand should be fulfilled by Depot 2 becomes 171 − 15 − 18 = 138 units and total demand for Depot 5 becomes 71 + 15 + 18 = 104 units.

Figure 2 illustrates the solution of VRP over 3 depots and 20 customers. It is known that Depot 2 should dispatch 3 vehicles to deliver products to 9 customers with total distance of 88.991, Depot 3 uses 2 vehicles to serve 5 customers with total distance of 71.930 and Depot 5 operates 2 vehicles to visit 6 customers with total distance of 85.619. In Figure 2, red squares indicate the selected depots.

5. Conclusion

A heuristic algorithm based on k-means clustering and Euclidean distance has been introduced to approach the CLRP. The algorithm comprises of customers clustering assessed by Dunn index, selecting the depot to be operated among a number of potential facilities, customer-depot matching and vehicle routing. The algorithm has successfully been implemented to a dataset consists of 20 customers and 5 potential depot facilities.

Future research direction in this area includes the application of hierarchical methods and use of route length formulae; extension to dynamic and stochastic problems; development of integrated methods in logistics, e.g., solving the location-routing-inventory and the location-routing-packing problems; and multiple objective LRP.
Acknowledgments

Authors gratefully acknowledge financial support provided by Ministry of Research, Technology and Higher Education of the Republic of Indonesia under grant No. 011/SP2H/LT/DRPM/IV/2017 through the PUPT Scheme of Bogor Agricultural University.

Table 5. Customer-depot matching

| Customer | Distance | Assigned Depot | Total Demand | Additional distance due to reallocation |
|----------|----------|----------------|--------------|----------------------------------------|
| 1        | 9.055    | 20.809         | 30.887       | 2                                      | 11.753 | 21.831 |
| 2        | 17.029   | 30.083         | 23.195       | 2                                      | 13.054 | 6.165  |
| 3        | 10.050   | 21.541         | 42.438       | 2                                      | 11.491 | 32.388 |
| 4        | 5.099    | 24.840         | 33.615       | 2                                      | 19.740 | 28.516 |
| 5        | 3        | 30             | 41.437       | 2                                      | 27     | 38.437 |
| 7        | 19.925   | 27.018         | 53.413       | 2                                      | 171    | 33.489 |
| 12       | 11.045   | 19.723         | 29.155       | 2                                      | 8.678  | 18.109 |
| 13       | 4.472    | 31.828         | 39.294       | 2                                      | 27.355 | 34.822 |
| 17       | 19.799   | 32.757         | 22           | 2                                      | 12.958 | 2.201  |
| 18       | 7.211    | 29.411         | 32.985       | 2                                      | 22.200 | 25.774 |
| 20       | 10.770   | 32.757         | 32.249       | 2                                      | 21.986 | 21.479 |
| 6        | 23.324   | 6.083          | 30.529       | 3                                      |        |        |
| 8        | 26.926   | 4.472          | 30.871       | 3                                      |        |        |
| 11       | 25       | 11.402         | 24.187       | 3                                      | 73     |        |
| 14       | 18.028   | 17.262         | 23.431       | 3                                      |        |        |
| 19       | 43.382   | 18.028         | 33.136       | 3                                      |        |        |
| 9        | 24.042   | 35.228         | 19.235       | 5                                      |        |        |
| 10       | 36.715   | 37.643         | 5.657        | 5                                      | 71     |        |
| 15       | 25.080   | 20.396         | 16.279       | 5                                      |        |        |
| 16       | 32.249   | 24.597         | 10.770       | 5                                      |        |        |

Figure 2. The optimal routes.
References

[1] Moons S, Ramaekers K, Caris A and Arda Y 2017 Comp. & Indust. Eng. 104 224–245.
[2] Drezner Z and Hamacher HW 2002 Facility Location: Applications and Theory (Berlin: Springer-Verlag).
[3] Prodhon C and Prins C 2014 Euro. J. Oper. Res. 238 1–17.
[4] Marinakis Y 2009 Encyclopedia of Optimization Second Edition, ed CA Floudas and PM Pardalos (New York: Springer) p 1919–1925.
[5] Schneider M and Drezel M 2017 Ann. Oper. Res. 259 389–414.
[6] Farahani RZ, SteadieSeifi M and Asgaria N 2010 App. Math. Model. 34 1689–1709.
[7] Belenguer JM, Benavent E, Prins C, Prodhon C and Calvo RW 2011 Comp. Oper. Res. 38(6) 931–941.
[8] Tuzun D and Burke LI 1999 Euro. J. Oper. Res. 116(1) 87–99
[9] Prins C, Prodhon C and Calvo RW 2006 4OR 4(3) 221–238.
[10] Prins C, Prodhon C, Ruiz A, Soriano P and Calvo RW 2007 Transport Science 41(4) 470–483
[11] Duhamel C, Lacomme P, Prins C and Prodhon C 2010 Comp. Oper. Res. 37(1) 1912–1923.
[12] Barreto SS 2004 Análise e Modelização de Problemas de Localização-Distribuição PhD Thesis (Portugal: University of Aveiro)
[13] Nadizadeh A, Sahraeian R, Zadeh AS and Homayouni SM 2011 Afric. J. Buss. Managm., 5(17) 7499–7506.
[14] Ting CJ and Chen CH 2013 Int. J. Prod. Econ. 141(1) 34 – 44.
[15] Lopes RB, Ferreira C, Santos BS 2016 Comp. Scien. Res. 70(1) 155–162.
[16] Peng Z, Manier H and Manier A 2017 Int. Feder. Automat. Contr. 50(1) 14668–14673.
[17] Toth P and Vigo D 2014 Vehicle routing: problems, methods, and applications (Philadelphia: SIAM-MOS).
[18] Abonyi J and Feil H 2007 Cluster analysis for data mining and system identification (Berlin: Birkhäuser Verlag AG).