Hybrid Algorithm for the Facility Location Problem based on Density based Clustering and Profit Maximization

Ashish Sharma[1], Ashish Sharma[2†] and Anand Singh Jalal[3]

[1]*Department of Mathematics, Institute of Applied Sciences and Humanities, GLA University, Mathura (UP)*
[2, 3]*Department of Computer Engineering & Applications, Institute of Engineering and Technology, GLA University, Mathura (UP)*
[1]*ashishsharma@gla.ac.in, [2]ashishs.sharma@gla.ac.in, [3]asjalal@gla.ac.in*

Abstract

The objective in the allocation of the facility depends up to the situation under consideration. For instance, in case of ATM, shopping malls, public utility services like schools, hospitals, etc. the facility is assigned, to such locations where the density of the users is more so that the facility will be utilized by as large as possible users. In such situation, the profit is at second priority as compared to utilization of facility. In this paper, we consider the same scenario and provide a hybrid algorithm for the solution of facility allocation problem. In the first-step, we use DBSCAN clustering technique, and after clustering, mixed integer linear programming technique is used in each cluster to get the best facility which will generate the maximum profit. Flowchart of the proposed algorithm and numerical example is presented.

1. Introduction

The facility location problem deals with the finding of the best location among the available one, which fulfills the objectives under consideration. The objective of the facility location problem depends upon the situation for example if we want to install a business outlet then the main objective will be the profit, on the other hand, if we want to install a medical facility then the main objective will be the utilization of the facility by as much as possible beneficiary. Similarly, bank ATM is also generally installed in a densely populated area. In such situations, generally the density-based clustering algorithms are used in order to get the information that which area contains the dense population. After getting the information about density, a facility is installed and if more than one option for opening the facilities are available then the preference will be given to those locations which will provide the maximum profit. For the density-based clustering, DBSCAN is one of the well-known clustering techniques, which is very much used by researchers.

In real-life applications, DBSCAN are used in many areas for instance [1] proposes a modification of DBSCAN clustering algorithm for identifying traffic accident-prone locations. [2] apply an ontological approach to the DBSCAN algorithm in the form of knowledge representation for constraint clustering. [3] applied two clustering techniques, k-means and DBSCAN, to an annotated Twitter dataset in order to evaluate the use of clustering for detecting different types of sentiment. They find that the results are very encouraging for DBSCAN as compared to k-means [4] proposed a modified generalized density-based clustering algorithm to deal with fuzziness in the values describing the population demographics which can be used for ATM location.

Through clustering, the determination of facility locations could be guided. [5] proposed an algorithm based on Sample Average Approximation that utilizes clustering techniques to update the sample sizes for the facility location problem dynamically. In real life, the facility location problem has been studied in various fields such as business management, medical diagnostics, and communication networks. Mixed Integer Programming is generally used in the computation of these models[6]. [7] proposes a mathematical model to aid banks in restructuring their branch locations by maintaining, closing or opening branches. The restructuring process is modeled as a nonlinear problem and is formulated as a mixed binary, integer linear model.

As stated by Melin and kauzo [8] the requirement of FLP is everywhere in real life, like establishing of various public services, emergency services, telecommunication networks services *etc*. This makes FLP is one of the interesting and complex problems for day to day life.[9] provide details of several other facility location models. [10, 11] Reza Zanjirani Farahani *et al* has contributed a review of hierarchical facility location models based on the objective of cost minimization and service availability maximization.[12] has given a review report concerning the characteristics of the dynamics of FLPs as well as the mathematical formulations that have been studied by the literature. [13] gives a clustering-based location-allocation method to the Capacitated Facility Location Problem. Thus to the best of our knowledge DBSCAN with Mixed integer linear programming technique is not used by any author for the solution of FLP.

The nature of FLP is so vast so that is not possible to design a single algorithm which we can apply in various FLP to find the solution. In the present scenario when the availability of service is essential to retain customers and to gain the profit also by providing all services is very challenging. The service quality can be maintained in the current scenario.

In this paper, we are assuming the position coordinated for 'M' customers and 'N' facilities. Over the set of 'M' customers and the 'N' facilities we are using the DBSCAN clustering technique for the formation of density based clusters. Each customer has certain demand and each facility have 'fixed cost' for the setting up the facility, 'variable cost' for the regular operation and the limited capacity or supply. We are also considered the transportation cost between facility and customer. After the cluster formation, mixed integer linear programming technique is used to optimize the profit function in the presence of capacity constraint. In the objective function, we are using the revenue earned from the customer, facility cost, variable cost, and transportation cost.

This remaining paper is organized as follows. All the notations and the explanation of FLP are provided in Section 2. Section 3 contains the analysis of the FLP. Particular cases are discussed in section 4. Section 5 contains the algorithm for the solution of FLP. The numerical example is presented in section 6. Section 7 concludes the paper.

2. Proposed Model

In general, service-based facility location models are based on the radius. Therefore, such facilities require the approach which fulfills their service in a radius.
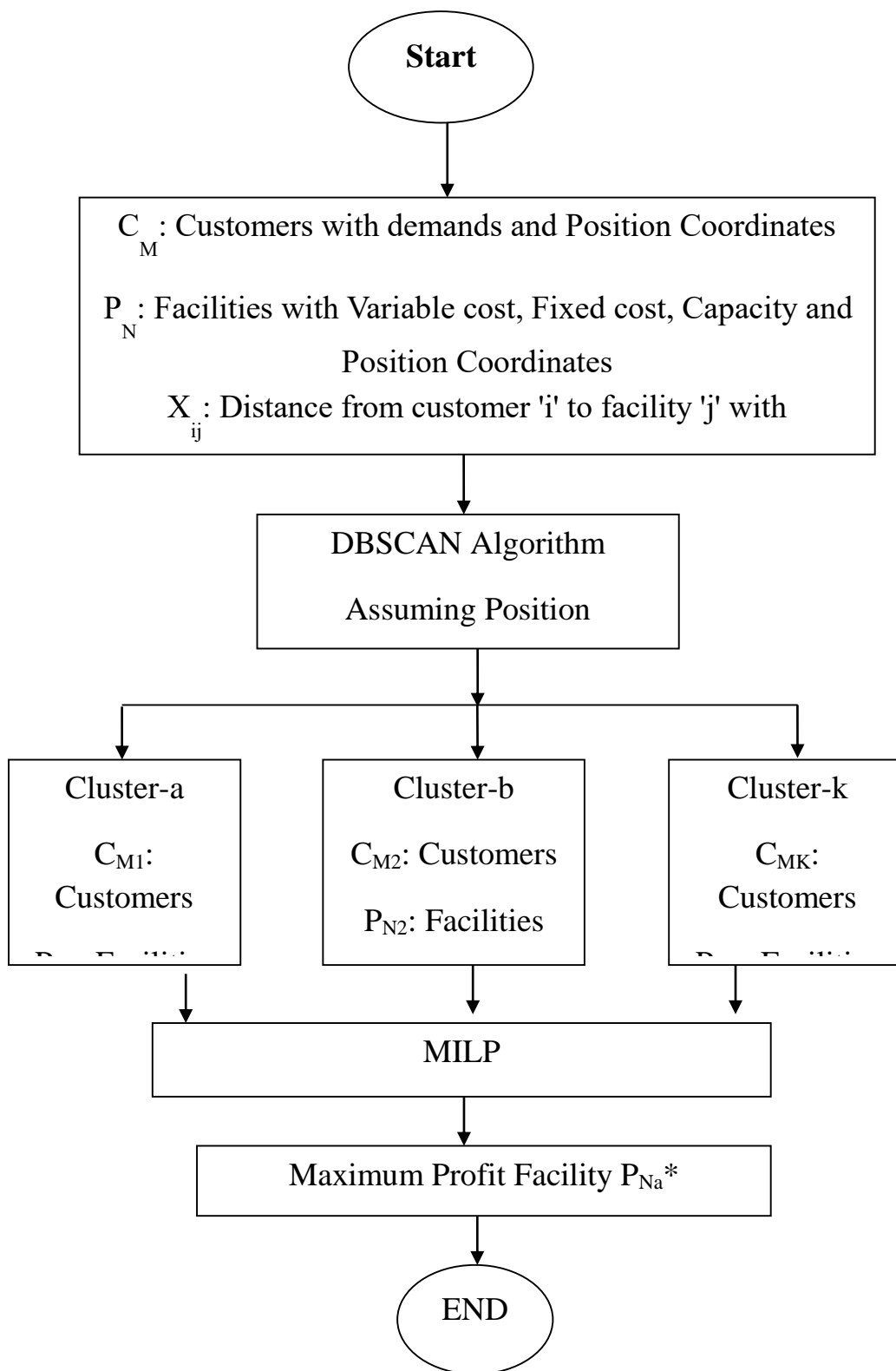
Figure 1. Solution Procedure Flow Chart

1. Notations, Definitions and Properties

The following notations are used in this paper:

i= Index of facilities, $1 \leq i \leq n$

j =Index of customers $1 \leq j \leq m$

(A_j, B_j) =location of the customer j, $1 \leq j \leq m$.

(A_i, B_i) = location of the facility i, $1 \leq i \leq n$

D_j = demand of customer j

r_j = price charged to client j, *i.e.* per Unit

s_i = capacity of facility i

c_i = fixed cost of a facility i

V_i = variable cost of facility i

x_i = binary decision variable, *i.e.,* the value 1, if facility 'i' is open, and 0 otherwise

y_{ij} = supplied quantity to from facility i to customer j

t_{ij} =The cost of transportation from facility i to customer j

d_j = demand of the customer j

We use the following assumptions for the model
1. The capacity, facility cost and the variable cost of the facility are fixed.
2. The demand of the individual customer is fixed.
3. A customer can use only one facility.
3. Clustering is based on the density of the customers.
4. Transportation cost of per unit per unit distance for all customers is same.

3. Facility Location Problem

In this model, the objective function is defined by

$$\max\left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} [(r_j - V_i - t_{ij})y_{ij}] - \sum_{i=1}^{n} c_i x_i \right\} \tag{1}$$

which contains the unit price charged to the customer (r_j), operating cost (V_i), transportation cost (t_{ij}) and the facility opening cost (c_i). Here the price charges to the customer is subtracted by the operating cost, transportation cost and the facility opening cost to get the revenue.

The capacity for a particular facility must be greater than or equal to the supply. In terms of mathematical inequalities, it is defined as

$$\sum_{j=1}^{m} y_{ij} \leq s_i x_i, i = 1, 2, \ldots n \tag{2}$$

Thus the final facility allocation problem in will as

$$\max\left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} [(r_j - V_i - t_{ij})y_{ij}] - \sum_{i=1}^{n} c_i x_i \right\} \tag{3}$$

Subject to

$$\sum_{j=1}^{m} y_{ij} \leq s_i x_i, \quad i = 1, 2, \ldots n \tag{4}$$

$$x_i \in \{0, 1\}, i = 1, 2, \ldots n \tag{5}$$

$$y_{ij} \geq 0, i = 1, 2, \ldots n; \quad j = 1, 2, \ldots m \tag{6}$$

Algorithm
Input: C_M : Customers with their demand
P_N : Facilities with their attributes
X_{ij} : distance from customer i to facility j
Output: Return Best facility for each cluster
The algorithm consists of three steps :
Clusters containing nearby facilities and customers
Find out the best facility for each cluster which provides Maximum Profit.
End

4. Numerical Examples

We now present the numerical examples to illustrate the solution procedure.

Example1: Consider the case when parameters of the FLP are: Table 1 shows the values of the Fixed Cost, Capacity and Variable Cost for all the 10 locations (P1 to P10). Table 2 shows the distance matrix, that is, the demand of 50 customers (C1 to C50) from 10 locations (P1 to P10)., Transportation Cost 1600/unit/unit distance and Price charged to client j is 150000 per Unit.

Table 1. Fixed Cost, Capacity and Variable Cost

Locations	Fixed Cost	Capacity	Variable cost
P1	43850000	4000	100000
P2	105620000	1500	110000
P3	53120000	1800	110000
P4	101800000	1300	110000
P5	29300000	1000	110000
P6	113720000	3400	100000
P7	44550000	3700	100000
P8	44120000	1800	110000
P9	41850000	4000	100000
P10	51600000	4200	100000

Table 2. Distance Matrix

Table 2: Demand for 50 customers (C1 to C50).

Customers	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
Demand	12	14	121	12	134	5	114	21	132	14
Customers	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
Demand	15	14	21	32	7	15	14	21	32	11
Customers	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30
Demand	15	14	21	32	74	15	24	21	22	64
Customers	C31	C32	C33	C34	C35	C36	C37	C38	C39	C40
Demand	24	15	11	22	12	14	25	4	21	32
Customers	C41	C42	C43	C44	C45	C46	C47	C48	C49	C50
Demand	72	15	14	21	7	65	54	21	32	18

Total Customer = 50, Total Demand = 2062 units

5. Results

Table 3 shows the result of the numerical example. Here we presented the results for both the clustering and without clustering.

Table 3. Result when Facility and Variable Cost are Different

Approach	Without Clustering			With Clustering	
Number of Clusters = 2					
Locations	P1 and P6			P1	P6
Total distance of a customer from a particular facility	18.2		23.82	22.69	21.18
Total distance between all customer & facilities	90.02			75.87	
No. of customers whose demand was satisfied	19		31	48	2
Average Customer Distance from a particular facility	4.72		2.9	1.5	10.59
Profit for each cluster	…….			53116210	2330782
Total Profit	48317120			55446992	
Number of Clusters = 3					
Locations	P1, P6 and P8			P8	P6	P5
Total distance of a customer from a particular facility	23.42	37.76	21.52	19.29	28.38	25.97
Total distance between all customer & facilities	84.7			73.64		
No. of customers whose demand was satisfied	21	9	20	12	29	9
Average Customer Distance from a particular facility	2.6	6.08	2.7	6.1	2.5	8.1

Locations	……..				10889400	32209620	5685117
Total distance of a customer from a particular facility	48378250				76238120		

Number of Clusters = 4							
Locations	P4, P6, P7 and P10				P6	P8	P1	P6
Total distance of a customer from a particular facility	12.21	15.42	17.13	19.78	10.74	9.93	32.17	28.75
Total distance between all customer & facilities	74.54				68.59			
No. of customers whose demand was satisfied	6	21	8	15	10	31	4	5
Average Customer Distance from a particular facility	2.04	0.74	2.15	1.32	12.5	4.051	31.39	25.11
Profit for each cluster	…….				31316610	9942294	1884606	61102020
Total Profit	48400020				49253712			

From the table it is clearly seen that if we group the customer on the basis the density then the profit will increase since the facility is used by as much as large customers.

6. Conclusion

The facility location problem with the fixed customer demand and the fixed capacity are considered. This paper basically contributes to the following aspects: (a) hybrid algorithm with DBSCAN clustering technique and mixed integer linear programming are considered. (b) Utilisation of facility is prior to the profit. (c) The proposed algorithm can also find the optimum number of clusters.The effectiveness of the solution is measured after applying the model on a real-life example. The computational results present the effectiveness and accuracy of the parameters to gain the maximum profit.

Reference

[1]   C. Qiu, H. Xu and Y. Bao, "Modified-DBSCAN Clustering for Identifying Traffic Accident Prone Locations", Proceedings Springer International Publishing of 17th International Conference, Yangzhou, China, (**2016**), pp. 99-105.

[2]   Q. Du, Z. Dong, C. Huang and F. Ren, "Density-Based Clustering with Geographical Background Constraints Using a Semantic Expression Model", ISPRS International Journal of Geo-Information, vol. 5 no. 5, **(2016)**, pp. 72.

[3]   Kumari, S. Surya, and G. Anjan Babu, "Sentiment on social interactions using linear and non-linear clustering", Proceeding of 2nd International Conference on, IEEE, Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), India, **(2016)**.

[4]   Kisore, N. Raghu and C.H. B. Koteswaraiah, "Improving ATM coverage area using density based clustering algorithm and voronoi diagrams", Information Sciences, vol. 376, **(2017)**, pp. 1-20.

[5]   E. Adindu, C. Sudipta, M. Mohammad, B. Linkan and E. Burak, "An enhanced sample average approximation method for stochastic optimization." International Journal of Production Economics, vol. 182, **(2016)**, pp. 230-252.

[6]   A.Y. Chen and T.-Y. Yu, "Network based temporary facility location for the Emergency Medical Services considering the disaster induced demand and the transportation infrastructure in disaster response", Transportation Research Part B: Methodological, vol. 91, **(2016)**, pp. 408-423.

[7]   Monteiro, M. S. Rodrigues and D. BMM Fontes, "Locating and sizing bank-branches by opening, closing or maintaining facilities", Operations Research Proceedings, Springer Berlin Heidelberg, **(2006)**, pp. 303-308.

[8]   MeilinWena and KakuzoIwamurab, "Facility location–allocation problem in random fuzzy environment: Using cost minimization model under the Hurewicz criterion", An International Journal of Computers and Mathematics with Applications, **(2008)**, pp. 704–713.

[9]   A.T. Murray, "Advances in location modeling: GIS linkages and contributions", Journal of Geographical Systems, vol. 12 no. 3, pp. 335-354.

[10] R. Z. Farahania, M. Hekmatfarb, B. Fahimniac and N. Kazemzadehd, "Hierarchical facility location problem: Models, classifications, techniques, and applications", Computers & Industrial Engineering, vol. 68, **(2014)**, pp.104–117.

[11] R. Z. Farahani, M. S. Seifi and N. Asgari, "Multiple criteria facility location problems: A survey", Applied Mathematical Modeling, vol. 34, **(2010)**, pp. 1689–1709.

[12] A.B. Arabani and R. Z. Farahani, "Facility location dynamics: An overview of classifications and applications", Computers & Industrial Engineering, vol. 62, **(2012)**, pp. 408–420.

[13] K. Liao and D. Guo, "A Clustering-Based Approach to the Capacitated Facility Location Problem", Transactions in GIS, vol. 12 no. 3, **(2008)**, pp. 323–339.