Management Science Letters 7 (2017) 577–584

Contents lists available at GrowingScience

Management Science Letters

homepage: www.GrowingScience.com/msl

Clustering based hybrid approach for facility location problem

Ashish Sharma[a*], Ashish Sharma[b] and A. S. Jalal[b]

[a]*Department of Mathematics, GLA University, 17 km stone, NH#2, Mathura (UP), India*
[b]*Department of Computer Engineering & Applications, GLA University, 17 km stone, NH#2, Mathura (UP), India*

C H R O N I C L E

Article history:
Received: July 24, 2017
Received in revised format: August 6, 2017
Accepted: August 25, 2017
Available online:
August 25, 2017

Keywords:
Facility location problem
Fuzzy demand
Region search
Distance based FLP

A B S T R A C T

The main objective of facility location problem is the utilization of the facility by maximum number of possible customers so that the profit is maximized. For instance, in some services like wireless sensor networks, Wi-Fi, repeaters, etc., where the service area is limited, some specific equipment is installed in such a way that it could be used by maximum number of users. Here, the number of users for a particular facility is optimized with the help of clustering technique. The study develops a model for facility allocation problem. For the solution algorithm, a hybrid approach which is based on clustering and mixed integer linear programming (MILP) is proposed. The proposed method consists of two parts where in the first part, the K- means clustering technique is used and in the second part, for each cluster an MILP technique is implemented so that the facility which yields the maximum profit is obtained. Numerical examples for clustering and without clustering are presented. Analysis shows that due to clustering the average distance between facility and customer is significantly reduced.

1. Introduction

In some situations like wireless sensor networks, Wi-Fi, repeaters, signals transmission etc. the service area for a facility is limited. In other words, if the customer is in the range of a particular facility then only s/he can avail services. Another observation in day to day life is that customers generally prefer the facilities like automated telling machine (ATM), Petrol pumps, etc. which are near to them. In both the situations we can conclude that the distance between the facility and the customer plays an important role for the utilization of the facility. Thus in the installation of the facility one of the primary objectives is to keep the low distance between the customers and the facility so that the facility would be utilized by a large number of customers which resultantly increases the profit.

After going in the deep discussion with experts and published literatures of the clustering, we think that this objective can be met by using clustering techniques. By clustering we can club the customers which lie in a specific range. In the literature we have found that many authors considered clustering for FLP. Geetha et al. (2009) proposed a solution procedure for the capacitated clustering problem (CCP) by

* Corresponding author. Tel.: +91-5662-250826), Fax: +91-5662-241486
E-mail address: ashishsharma@gla.ac.in (A. Sharma)

doi: 10.5267/j.msl.2017.8.007

improving k-means clustering algorithm, by setting the priority as a measure which directs the search for better optimization. A new solution framework for Capacitated P-median problem (CPMP) is proposed by Sahraeian and Kaveh (2010). CPMP locates P facilities between the candidate points to satisfy the customer demand. Their approach is hybrid of K-means clustering algorithm. Saiedy et al. (2011) present a modeling of capacitated single allocation hub location problem with two indices.
A review of hierarchical facility location models is presented by Farahani et al. (2014). This review explains the facility layout problem (FLP) in two points of views. First one is based on cost minimization objective and the other one is based on service availability maximization objective. Another review paper on FLP given by Arabani and Farahani (2012) includes multiple dynamics of FLP with mathematical formulations. A review with different solution methods and the latest development for the three different categories of the FLP i.e. bi-objective, multi-objective and multi-attribute problems has provided by Farahani et al. (2010).

Some authors also used the integer programming techniques in order to obtain the best facility. The Capacitated Facility Location Problem (CFLP) is performed by Liao and Guo (2008) using mixed approach based on k-means and genetic algorithms. Kolisch and Dahlmann (2015) consider the dynamic assigning of replicas to the servers in their service based solution. To make more approximate solution Arya et al. (2004) uses the local search heuristic algorithm with k-means. Another type of facility location is Continuous Facility Location problem explored by (Meira & Miyazawa, 2008). Authors have presented algorithms with new approximation factors for Euclidean distances and squared Euclidean distances using k-means clustering algorithm for the small dimensions.

Many authors used the clustering techniques for the solution of FLP. (Chaves & Dahlmann, 2010) has proposed the Capacitated Centered Problem. It consists partitioning a collection of 'n' points into 'p' disjoint clusters with a noted capability. Wena and Iwamura (2008) considere the facility location problem and vehicle routing problem both simultaneously. In another paper the multi-level un-capacitated facility location problem is considered by (Kaya & Urek, 2016). They have considered mixed integer linear programming (MILP) as a solution and validated. Shishebori and Ghaderi (2015) consider the combined facility location/network design problem with regard to transportation link disruptions and develops a mixed integer linear programming formulation to model it. An efficient hybrid algorithm based on LP relaxation and variable neighborhood search meta-heuristic is developed in order to solve the mathematical model.

To the best of our knowledge we did not found any work which used the clustering and the MILP for the solution of the FLP. Thus to meet this limitation, in this paper, we first defined the mathematical model of the FLP which contains the objective functions and the constraints based on capacity and demand. For the solution of the FLP we used the mixed algorithm of clustering and the MILP. In which in the first step we apply K-means clustering techniques over a set of customers and facilities so that we will get the required number of clusters. Now in each clusters, the possibilities are:

1. Each cluster contains some customers and some facilities.
2. Some clusters contain customers as well as facilities and some clusters contain only customers.
3. Some clusters contain customers as well as facilities and some clusters contain only facility.

Thus in the presence of all such possibilities we prepare an algorithm for the solution. After clustering we apply the MILP on each cluster to get the best facility which provides the maximum profit. Thus by summing up the profit of individual cluster we will get the overall optimum profit. Initially we start this algorithm for 2 clusters and then we gradually increase the clusters. At some stage we will get the maximum profit and if we further increase the clusters then the profit will start decreasing. Thus with the help of this algorithm we can also obtain the number optimum clusters for a particular FLP.

The basic research question we addressed in this paper are;

(1) Is clustering is helpful to obtain the basic objective of minimizing the distance between the customer and facility?
(2) What will be the optimum number of clusters, if number of opening of facilities is not restricted?

2. The proposed method

2.1 Notations, Definitions and Properties

The following notations are used in this paper:
i= Index of facilities, $1 \le i \le n$
j =Index of customers $1 \le j \le m$
(A_j, B_j) =location of the customer j, $1 \le j \le m$.
(A_i, B_i) = location of the facility i, $1 \le i \le n$
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 = The Opening and operating facility fixed cost of a facility i
V_i = The Operating cost, i.e. a real number and unit variable of the facility i
x_i = Decision variable, i.e. binary, the value 1, if facility i is open, and zero otherwise
y_{ij} = The supplied quantity to from i to j
t_{ij} =The cost of transportation from i to 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. Clustering is based upon the euclidian distance of the customer and the facility.
4. Transportation cost of per unit per unit distance for all customers is same.

2.2. Modeling formulation

After finding the Euclidian distance between the facility and the customer we apply k-means algorithm for the formation of clusters so that customers and the facilities are divided into different clusters. To gain the maximum profit we have to identify the facility location point for which the profit is maximum that can be done by profit function that we have defined in our next section. The facility for which the objective function is maximum, is achieved by mixed integer linear programming. This point becomes the facility location point in the cluster. In this problem, our main aim is to maximize the profit function in the presence of the constraint related to supply, capacity, demand and distance. In the objective function, we subtract the unit operating cost and the unit transportation cost from the unit price charged to customer, so that we get the unit revenue earned and then we will multiply it by the quantity supplied so that, we get the total revenue earned. Now, for the profit function we subtract the operating cost of the facility from the total revenue earned. Thus, the objective function is

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

We restricted that the demand for a particular customer must be greater than or equals to the total supply from all the facilities. In notations it is defined as

$$\sum_{j=1}^{m} y_{ij} \le D_j, j = 1, 2, \ldots m \tag{2}$$

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

$$\sum_{j=1}^{m} y_{ij} \le s_i x_i, i = 1, 2, \ldots n \tag{3}$$

Thus the final facility allocation problem will be

$$\max\left\{\sum_{i=1}^{n}\sum_{j=1}^{m}[(r_j - V_i - t_{ij})y_{ij}] - \sum_{i=1}^{n} c_i x_i\right\} \tag{4}$$

subject to

$$\sum_{j=1}^{m} y_{ij} \le D_j, j = 1, 2, \ldots m \tag{5}$$

$$\sum_{j=1}^{m} y_{ij} \le s_i x_i, \quad i = 1, 2, \ldots n \tag{6}$$

$$x_i \in \{0, 1\}, i = 1, 2, \ldots n \tag{7}$$

$$y_{ij} \ge 0, i = 1, 2, \ldots n; j = 1, 2, \ldots m \tag{8}$$

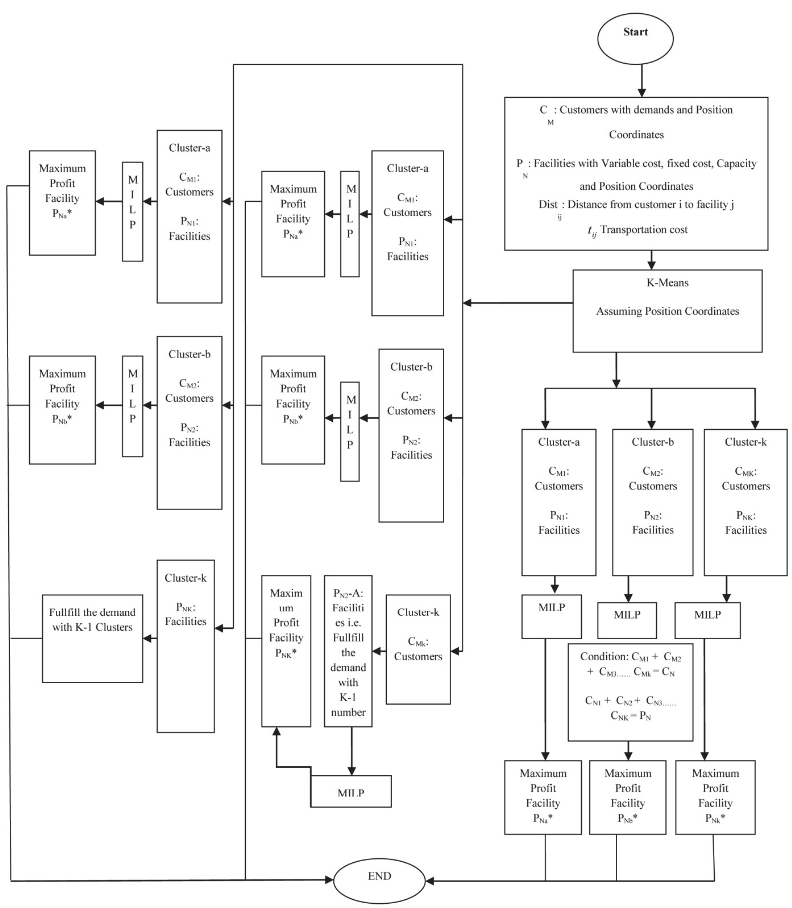

Fig. 1. shoes the flow chart of the proposed approach for FLP

3. Numerical Examples

We now present the numerical examples to illustrate the solution procedure.
Example1: 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 demand for all the 50 customers (C1 to C50).Transportation Cost 200/unit/unit distance and Price charged to client j is 150000 per Unit. We have taken the data sets of (Lin, 2014).

Table 1
Fixed Cost ($\times 10^7$), Capacity and Variable Cost ($\times 10^6$)

Locations	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Fixed cost	4.385	1.056	5.312	1.018	2.930	1.1372	4.455	4.412	4.185	5.16
Capacity	4000	1500	1800	1300	1000	3400	3700	1800	4000	4200
Variable cost	1.0	1.1	1.1	1.1	1.1	1.0	1.0	1.1	1.0	1.0

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

Results: Table 3(a) to Table 3 (d) shows the result of the numerical example 1 for the cases without clustering and clustering for number of clusters = 2-5 when facility and variable cost are different.

Table 3(a)
Number of clusters = 2

	Without Clustering		With Clustering	
Locations	P1	P6	P10	P6
Total distance of a customer from a particular facility	43.31	39.11	22.69	21.18
Total distance between all customer & facilities	82.42		43.87	
No. of customers whose demand was satisfied	20	30	26	24
Average Customer Distance from a particular facility	2.17	1.31	0.88	0.89
Profit for each cluster	……………..		36220030	47025410
Total Profit	82933210		83245440	

Table 3(b)
Number of clusters = 3

	Without Clustering			With Clustering		
Locations	P1	P6	P8	P10	P6	P4
Total distance of a customer from a facility	43.53	36.93	3.65	19.29	28.38	25.97
Total distance between all customer & facilities	84.11			73.64		
No. of customers whose demand was satisfied	19	29	2	16	21	13
Average Customer Distance from facility	2.29	1.27	1.82	1.21	1.35	1.99
Profit for each cluster	……………….			27689400	37980460	1.9E+07
Total Profit	65480730			84235350		

Table 3(c)
Number of clusters = 4

Locations	P1	P3	P6	P8	P7	P8	P1	P6
Total distance of a customer from a facility	43.05	8.55	34.68	3.65	30.74	19.93	56.17	18.75
Total distance between all customer & facilities	89.96				125.59			
No. of customers whose demand was satisfied	19	2	27	2	15	12	1	22
Average Customer Distance from facility	2.27	4.2	1.29	1.83	2.05	1.66	56.17	0.853
Profit for each cluster	…………….				23842850	21986210	741830	34446350
Total Profit	86897460				81017240			

Table 3(d)

Number of clusters = 5

Locations	P1	P3	P4	P6	P8	P4	P10	P3	P1	P7
Total distance of a customer from a facility	42.68	22.88	34.82	3.65	0	18.75	12.69	10.13	56.17	28.83
Total distance between all customer & facilities			104.03					126.57		
No. of customers whose demand was satisfied	17	2	28	2	0	22	9	5	2	12
Average Customer Distance from facility	2.52	11.44	1.25	1.83	0	0.85	1.41	2.03	28.09	2.4
Profit for each cluster			………………..			34446370	17467210	7959210	741830	20187890
Total Profit			86887280					80802510		

Thus the maximum profit 8,42,35,350 and the minimum distance 73.64 units, corresponds to the number of clusters = 3.

Example 2: Now for the further insights of the model, we consider the case when the variable cost and fixed cost are same for all the facilities. We considered fixed cost = 4,00,000 and Variable cost = 1,00,000for all the locations. All other parameters like demand of customers, transportation cost and the price charged to client j are same as considered in example 1. Tables 4(a-d) report the result when facility and variable cost are different.

Table 4(a)

Number of clusters = 2

Approach	Without Clustering		With Clustering	
Locations	P4	P10	P10	P4
Total distance of a customer from a particular facility	18.2	23.82	22.69	21.18
Total distance between all customer & facilities	42.02		43.87	
No. of customers whose demand was satisfied	25	25	26	24
Average Customer Distance from a particular facility	0.728	0.9528	0.87269231	0.8825
Profit for each cluster	…….		38199330	37757270
Total Profit	73707850		75956600	

Table 4(b)

Number of clusters = 3

Locations	P4	P6	P10	P10	P6	P4
Total distance of a customer from a particular facility	15.42	17.76	21.52	19.29	28.38	25.97
Total distance between all customer & facilities		54.7			73.64	
No. of customers whose demand was satisfied	21	9	20	16	21	13
Average Customer Distance from a particular facility	0.74	1.98	1.08	1.21	1.36	1.99
Profit for each cluster		……..		28668700	29475170	18094250
Total Profit		76012940			76238120	

Table 4(c)

Number of clusters = 4

Locations	P4	P6	P7	P10	P7	P8	P1	P6
Total distance of a customer from a particular facility	12.21	15.42	17.13	19.78	30.74	19.93	56.17	18.75
Total distance between all customer & facilities		64.54				125.59		
No. of customers whose demand was satisfied	6	21	8	15	15	12	1	22
Average Customer Distance from a particular facility	2.04	0.74	2.15	1.32	2.05	1.67	56.17	0.86
Profit for each cluster		…….			23818410	21979330	545680	28346810
Total Profit		76655960				74690230		

From the Table 4(a) to Table 4 (d) we can observe that in case of clustering, when the number of cluster are 2, the total distance travelled by customers is less as compare to without clustering, which results the reduced transportation cost hence the more profit. Now if we further increase the number of clusters the total distance in case of clustering starts increasing, which results the more transportation cost and less profit. Thus from Table 4, for the three clusters we will get the maximum profit Rs. 7,62,381,20and the minimum distance 73.64 units. This shows that the clustering to a certain limit is useful as it decrease the distance between the facility and the customers. The proposed algorithm also provides the number of optimum clusters so that we will get the maximum profit.

Table 4(d)
Number of clusters = 5

Locations	P1	P3	P4	P6	P8	P4	P10	P3	P1	P7
Total distance of a customer from a particular facility	42.68	22.88	34.82	3.65	0	18.75	12.69	10.13	56.17	28.83
Total distance between all customer & facilities			104.03					126.57		
No. of customers whose demand was satisfied	17	2	28	2	0	22	9	5	2	12
Average Customer Distance from a particular facility	2.52	11.44	1.25	1.83	0	0.85	1.41	2.02	28.08	2.4
Profit for each cluster			…….			28346710	17446510	7972330	545680	20163450
Total Profit			77254310					74474680		

From the Table 3(a to d) & 4(a to d) we can observe that in case of clustering, when the number of cluster are up to 3, the total distance travelled by customers is less as compare to without clustering, which results the reduced transportation cost hence the more profit. Now if we further increased the number of clusters the total distance in case of clustering starts increasing, which results the more transportation cost and less profit. Thus from Table 3, for the three clusters we will get the maximum profit Rs8, 42, 35,350 and the minimum distance 73.64 units. Thus we can conclude that, to a certain limit, clustering is useful, which decrease the distance between the facility and the customers and once the profit starts decreasing as we increase the number of clusters it continuously decreases. This is very helpful result which can be used to stop the solution procedure. This is also explained in the Fig. 2 which is plotted for the number of clusters against the profit for example 1 and 2. The proposed algorithm also provides the number of optimum clusters =3 which provides the maximum profit.

Fig. 2. Number of clusters against the profit for examples 1 and 2

4. Conclusion

In this paper, a new clustering and mixed integer linear programming based hybrid approach for FLP is developed. The proposed approach efficiently minimizes the distance between the customer and the facility as well as it also maximizes the profit. This is verified and substantiated by the numerical analysis. With the help of proposed approach the user find whether the clustering is beneficial for him or not and if it is beneficial he can also obtain the optimal number of clusters which will maximize the profit. Numerically we proved that once the profit starts decreasing as the number of cluster increases, it continuously decreases this result help us to stop the solution procedure. For the future research, the practitioners can extend the model for more complex and realistic situations, like consideration of random distance.

References

Arabani, A. B., & Farahani, R. Z. (2012). Facility location dynamics: An overview of classifications and applications. *Computers & Industrial Engineering*, *62*(1), 408-420.

Arya, V., Garg, N., Khandekar, R., Meyerson, A., Munagala, K., & Pandit, V. (2004). Local search heuristics for k-median and facility location problems. *SIAM Journal on computing*, *33*(3), 544-562.

Chaves, A. A., & Lorena, L. A. N. (2010). Clustering search algorithm for the capacitated centered clustering problem. *Computers & Operations Research*, *37*(3), 552-558.

Farahani, R. Z., SteadieSeifi, M., & Asgari, N. (2010). Multiple criteria facility location problems: A survey. *Applied Mathematical Modelling*, *34*(7), 1689-1709.

Farahani, R. Z., Hekmatfar, M., Fahimnia, B., & Kazemzadeh, N. (2014). Hierarchical facility location problem: Models, classifications, techniques, and applications. *Computers & Industrial Engineering*, *68*, 104-117.

Geetha, S., Poonthalir, G., & Vanathi, P. T. (2009). Improved k-means algorithm for capacitated clustering problem. *INFOCOMP Journal of Computer Science*, *8*(4), 52-59.

Kaya, O., & Urek, B. (2016). A mixed integer nonlinear programming model and heuristic solutions for location, inventory and pricing decisions in a closed loop supply chain. *Computers & Operations Research*, *65*, 93-103.

Kolisch, R., & Dahlmann, A. (2015). The dynamic replica placement problem with service levels in content delivery networks: a model and a simulated annealing heuristic. *OR spectrum*, *37*(1), 217-242.

Liao, K., & Guo, D. (2008). A Clustering-Based Approach to the Capacitated Facility Location Problem. *Transactions in GIS*, *12*(3), 323-339.

Lin, P. C., Watada, J., & Wu, B. (2014). A parametric assessment approach to solving facility-location problems with fuzzy demands. *IEEJ Transactions on Electrical and Electronic Engineering*, *9*(5), 484-493.

Meira, L. A., & Miyazawa, F. K. (2008, March). A continuous facility location problem and its application to a clustering problem. In *Proceedings of the 2008 ACM symposium on Applied computing* (pp. 1826-1831). ACM.

Wen, M., & Iwamura, K. (2008). Facility location–allocation problem in random fuzzy environment: Using (α, β)-cost minimization model under the Hurewicz criterion. *Computers & Mathematics with Applications*, *55*(4), 704-713.

Sahraeian, R., & Kaveh, P. (2010). Solving Capacitated P-Median Problem by Hybrid K-Means Clustering and Fixed Neighborhood Search algorithm. In *Proceedings of the 2010 International Conference on Industrial Engineering and Operation Management,* 1-6.

Saiedy, H., Moezi, D. S., & Noruzi, M. (2011). Modeling of capacitated of single allocation hub location problems with n hub centre, *Journal of Scientific and Industrial Research*, *70* (1) 20-24.

Shishebori, D., & Ghaderi, A. (2015). An integrated approach for reliable facility location/network design problem with link disruption. *International Journal of Supply and Operations Management*, *2*(1), 640-661.