Genetic Algorithm For Solving The Uncapacitated Facility Location Problem

Rajesh Kumar.S*, Jerrin George Abraham**,Vineeth Mathew**,Jith John Francis**,
Mathews Oommen**, Nithin.P.Mathew**
*Assistant Professor, **Engineering Students
Department of Mechanical Engineering
Mar Baselios College of Engineering and Technology, Kerala University, Trivandrum, Kerala

Abstract

Facility location problem is to find locations for new facilities such that the conveying cost from facilities to customers is minimized. In practice, some factors such as demands, allocations, even locations of customers and facilities are usually changing. In an uncapacitated facility location problem, the customers are supplied by the nearest factory. However, in a capacitated problem, the customers may not be supplied by the nearest factory only. Facility location problem has been studied for half a century because of its widely practical application backgrounds. The problem in this paper consists of a company whose five adequate facilities have to be located considering the variable constraints using an operational solving tool known as genetic algorithm.

Keywords–Facility location, mixed integer programming models, Genetic algorithm.

1. Introduction

Decisions about the distribution system are a strategic issue for almost every company.The facilities we plan today help an organization achieve supply chain excellence. Supply chain management entails not only the movement of goods but also the decisions about, (i) where to produce, what to produce and how much to produce at each site, (ii). What quantity of goods to hold in inventory at each stage of the process, (iii) how to share the information among the parties in the process and finally, (iv) Where to locate plants and distribution centers.

Facilities are critical components of multilevel networks necessary for supply chain management. The major objectives of facility planning are to, (i) Improve customer satisfaction conforming to customer promises, (ii) increase return on assets, (iii) maximize speed for quick customer response, (iv) effectively use people, equipment, space and energy[5].

Location allocation models cover formulations which range in complexity from simple linear, single-stage, single-product, un capacitated, deterministic models to non-linear probabilistic models Location decisions may be the most critical and the most difficult of the decisions needed to realize an effective supply chain. Inefficient and inappropriate location for production and assembly plants as well as distribution centre's will result in excess costs being incurred throughout the lifetime of the facility.

Vehicle routing and inventory decisions are generally secondary to facility locations in the sense that facilities are expensive to construct and difficult to modify, while routing and inventory decisions can be modified periodically without difficult. In this case study we consider a company currently having two major plants in Kerala and five storage locations and 47 customer destinations all across India. The aim of this case study is to re-allocate the existing storage locations to new five locations across India so as to decrease the transportation cost from the plant to the warehouse and from the storage locations to the customer destinations considering into account several various variable factors involved in the placement of the facility locations.

The outline of this case study is as follows, first we review some of the models which has contributed to the current state of-the-art. Then we do a detail study

about the problem of allocating the facility and its constraints, and then apply a suitable methodology to allocate the facilities at the appropriate locations considering the various factors using an algorithm.

2. LITERATURE SURVEY

2.1 Types of models

Facility location models can be broadly classified as follows[2]:

1. The shape or topography of the set of potential plants yields models in the plane, network location models, and discrete location or mixed-integer programming models, respectively. For each of the subclasses distances are calculated using some metric.

2. Objectives may be either of the minsum or the minmax type. Minsum models are designed to minimize average distances while minmax models have to minimize maximum distances. Predominantly, minsum models embrace location problems of private companies while minmax models focus on location problems arising in the public sector.

3. Models without capacity constraints do not restrict demand allocation. If capacity constraints for the potential sites have to be obeyed demand has to be allocated carefully. In the latter case we have to examine whether single-sourcing or multiple-sourcing is essential.

4. Single-stage models focus on distribution systems covering only one stage explicitly. In multi-stage models the flow of goods comprising several hierarchical stages has to be examined.

5. Single-product models are characterized by the fact that demand, cost and capacity for several products can be aggregated to a single homogeneous product. If products are inhomogeneous their effect on the design of the distribution system has to be analyzed, viz. multi-product models have to be studied.

6. Frequently, location models base on the assumption that demand is inelastic, that is, demand is independent of spatial decisions. If demand is elastic the relationship between, e.g., distance and demand has to be taken into account explicitly. In the latter case cost minimization has to be replaced through, for example, revenue maximization.

7. Static models try to optimize system performance for one representative period. By contrast dynamic models reflect data (cost, demand, capacities, etc.) varying over time within a given planning horizon.

8. In practice model input is usually not known with certainty. Data are based on forecasts and, hence, are likely to be uncertain. As a consequence, we have either deterministic models if input is (assumed to be) known with certainty or probabilistic models if input is subject to uncertainty.

9. In classical models the quality of demand allocation is measured on isolation for each pair of supply and demand points. Unfortunately, if demand is satisfied through delivery tours then, for instance, delivery cost cannot be calculated for each pair of supply and demand points separately. Combined location/ routing models elaborate on this interrelationship.

2.1.1 Continuous location models. Continuous location models (models in the plane) are characterized through two essential attributes: (a) the solution space is continuous, that is, it is feasible to locate facilities on every point in the plane. (b) Distance is measured with a suitable metric.. The objective of this model is to minimize the sum of distances between the facilities and m given demand points. The corresponding optimization problem can be solved efficiently by means of an iterative procedure.

2.1.2 Network location models. In network location models distances are computed as shortest paths in a graph. Nodes represent demand points and potential facility sites correspond to a subset of the nodes and to points on arcs. The network location model corresponding to the continuous multi-source Weber model is called p median problem. In the p-median problem p facilities have to be located on a graph such that the sum of distances between the nodes of the graph and the facility located nearest is minimized.

2.1.3 Set covering location problem. Here the objective is to locate the minimum number of facilities required to cover all of the demand nodes. This type of formulation assumes that the candidate

facility sites are located at the nodes of the network. A lower cost facility siting scheme might be possible if the facilities could be located along the arcs of the networks as well.

2.1.4 Maximal covering location problem. An under laying assumption of the set covering location problem is that all of the demand nodes must be covered in essence, there is no budget constraint. The maximal covering location problem was formulated to dress planning situations which have an upper limit on the number of facilities to be sited. The objective of the MCLP is to locate a predetermined number of facilities, p, in such a way to maximize the demand covered. Thus the MCLP assumes that there may not be enough facilities to cover all of the demand nodes.

2.1.5 p-center problem. The p-center problem addresses the problem of minimizing the maximum distance that the demand is from its closest facility given that we are siting a pre-determined number of facilities. There are several variable variations of the basic model. The vertex p-center problem restricts the set of candidate facility sites to the nodes of the network while absolute p-center problem permits the facilities to be anywhere along the arcs.

2.1.6 The p-dispersion problem. For all of the models discussed above the concern is with the distance between demand and new facilities. Also, an unspoken assumption is that being close to a facility is desirable. The p-dispersion problem differs from that problem in two ways. First, it is concerned only with the distance between new facilities. Second the objective is to maximize the minimum distance between any pair of facilities. Potential applications of PDP include the siting of military installations where separation makes them more difficult to attack or locating franchise outlets where separation reduces cannibalization among stores.

2.1.7 Mixed-integer programming models. Starting with a given set of potential facility sites many location problems can be modelled as mixed integer programming models. Apparently, network location models differ only gradually from mixed integer programming models because the former ones can be stated as discrete optimization models. A rough classification of discrete facility location models can be given as follows: (a) single- vs. Multistage models, (b) uncapacitated vs. capacitated models, (c) multiple- vs. single-sourcing, (d) single- vs. multi-product models, (e) static vs. dynamic models, and, last but not least, (f) models without and with routing options included.

3. Problem Definition and Mathematical Model

3.1 Problem Description

In India a company called English Indian clays limited, produces a product P which is used as one of the major raw materials for producing value added product by many other companies across India. The company is situated on the southern state of Kerala at Trivandrum as the above mentioned product raw material is cheaply and easily available in Trivandrum at nearby locations of the company. The company currently has two plants in the Trivandrum district which produces the same product. The two plants are located at veli and at thonnakal in the Trivandrum district. After the production they store the goods at their main warehouses located at Trivandrum itself, then it gets shifted to other warehouses located at various parts of India. Currently they have five warehouses situated in different parts of the country to meet their customers' demands. This paper aims to find the optimum location of the warehouses so that the cost for transporting goods from the plant to the warehouses and from the warehouses to the customer reduces. The existing warehouses are located at Trivandrum, kollam, Bangalore, Delhi, and Gujarat.

3.2 Assumptions and modeling

1. Each red dot represents a demand point.
2. Each green node represents a warehouse.

3. Each node represents a possible location of a facility warehouse.
4. Only one facility may be located per node.
5. The travel cost are symmetric i.e. cost of each arc or link i.e. cost for transporting goods from one facility to another facility is directly proportional to its distance.
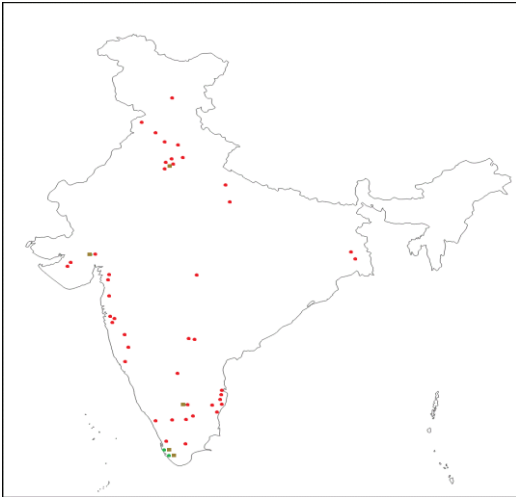

Figure 1. Locations of Customers in India

The model is formulated as mixed integer programming model. The model developed here minimizes the overall cost, which includes facility operation transportation, maintenance and storage and it is subjected to a set of various constraints and non negativity constraints.

3.3 Indices

I= set of location of plant, indexed as i.
J=Set of (possible locations) candidate facility location, indexed as j.
K=Set of customer locations indexed by k.

3.4 Model parameters

U_j = total available capacity of warehouse j $\in$ J

$d_k = demand$ at customer locations k $k \in K$

C_{ij} = Unit cost of shipping between plant site $i \in I$, and candidate facility $j \in J$.

v_j = Variable cost of locating facility at a candidate site $j \in J$

C_{jk} = Unit cost of shipping between candidate facility, $k \in K$ and candidate site $j \in J$.

f_j = Fixed cost of locating a facility at candidate site $j \in J$.

3.5 Decision variables

$$Z_j = \begin{bmatrix} 1 & \text{if we locate at candidate site } j \in J \\ 0 & \text{if not} \end{bmatrix}$$

X_{ij} = Quantity of commodity shipped between supply site $i \in I$ and candidate facility $j \in J$

X_{jk} = Quantity of commodity shipped between candidate facility $j \in J$ and customer location $k \in K$

3.6 Mathematical model

$$\text{Minimize } \sum_{i \in I}\sum_{j \in J} C_{ij}.X_{ij} + \sum_{j \in J}\sum_{k \in K} C_{jk}.X_{jk} + \sum_{j \in J} f_j.Z_j + \sum_{j \in J} v_j.Z_j$$

Subjected to the following constraints:

$$\sum_{j \in J}\sum_{k \in K} X_{jk} = d_k, \forall \ k \in K$$

$$\sum_{i \in I}\sum_{j \in J} X_{ij} \leq U_j.Z_j$$

$$X_{ij}, X_{jk} \geq 0 \quad \forall \ i \in I; \forall \ j \in J, \forall \ k \in K.$$

4. Solution methodologies

This network design problem is an extension of the traditional uncapacitated facility location problem

(UFLP). UFLP was shown to be NP-complete by Krarup, which implies that the problem that we are studying also belongs to the NP-complete class of problems. Therefore, we propose an efficient heuristic algorithm for solving this kind of problem using LP-based Genetic Algorithm.

4.1 Genetic Algorithm

The proposed mathematical model is a NP hard problem. Thus genetic algorithm-based heuristic is used in order to obtain good solutions. Figure demonstrates the steps of this approach[4]. The potential locations of plant, warehouse are summarized in table 3 and monthly demands of each distribution centre locations are shown in table 4. For simplicity Euclidean distance is used in measuring travel distance.

Table 2. Supply from Plant

PLANT	SUPPLY
P1	4000
P2	12000

TABLE 2: Supply from warehouses

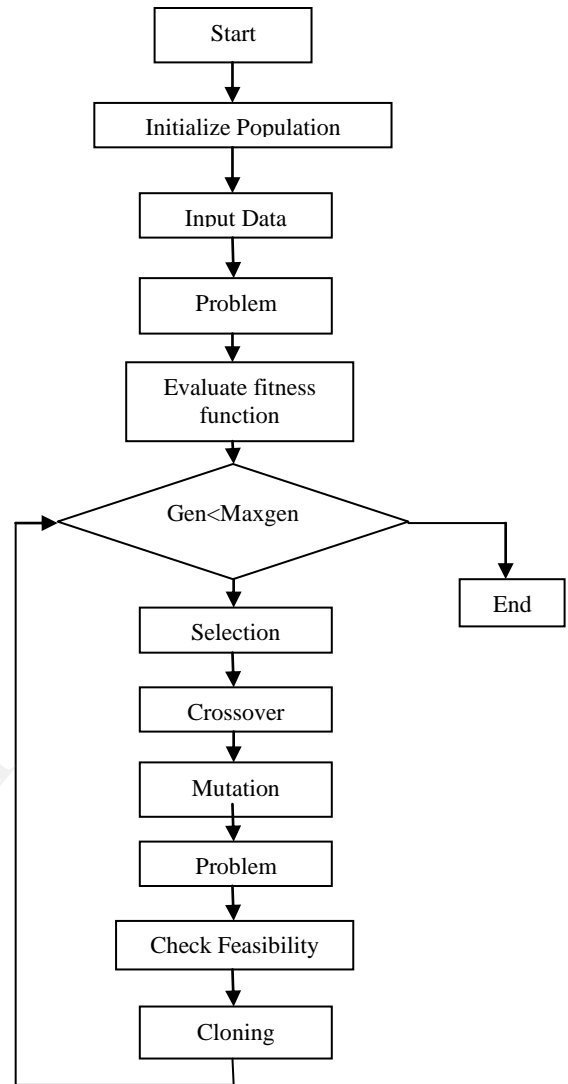| WAREHOUSES	SUPPLY
W1	2000
W2	2500
W3	2000
W4	10000
W5	3000

Figure 2. Flowchart representation of Genetic Algorithm

Table 3. Potential sites of warehouses and Plants

Plants	Site Coordinates		Warehouses	Site Coordinates	
	X	Y		X	Y
P1	1140	210	W1	1040	2470
			W2	1165	705
P2	1145	240	W3	1125	260
			W4	1160	220
			W5	630	1900

Table 4. Demands and coordinates of customers

Customers	Demands	Site Coordinates	
		X	Y
D1	15	1530	635
D2	187	1300	1250
D3	127	675	1460
D4	525	2260	1830
D5	100	1370	1660
D6	2420	690	1590
D7	15	845	1170
D8	1625	675	1700
D9	188	675	1420
D10	1505	1100	2470
D11	467	1290	1245
D12	704	995	2530
D13	355	800	1280
D14	622	1505	770
D15	859	630	1900
D16	608	650	1410
D17	105	795	1030
D18	43	1245	1230
D19	16	1015	2320
D20	15	1415	770
D21	490	1515	1120
D22	48	1425	1205
D23	774	1165	705
D24	150	785	990
D25	75	915	2250
D26	309	1060	370
D27	180	460	1800
D28	250	1125	960
D29	48	650	1650
D30	153	1220	715
D31	61	1470	730
D32	29	1370	2210
D33	63	1000	2910
D34	9	990	2515
D35	14	900	2738
D36	32	1330	2380
D37	113	445	1775
D38	528	1090	2560
D39	60	1450	725
D40	16	1275	600
D41	385	1040	2470
D42	32	1020	2475
D43	16	1180	560
D44	22	2250	1800
D45	16	1215	350
D46	10	980	510
D47	150	1085	2440

4.2 Genetic operators

4.2.1 Representation Scheme. To design a GA for a particular problem, we first need to devise a suitable representation scheme that shows the solution characteristics. In solving the location problem using a GA previous researchers usually defined the chromosomes as being the potential sites. In this paper, we developed a representation scheme that consisted of 2-dimensional binary strings as the chromosome structure. The 2-dimensional binary representation of an individual's chromosome is illustrated in Figure.

50	21	98	142	13

Figure 3. Representation

4.2.2 Fitness Function and Parents Selection. The fitness function is a measure of goodness of a solution. In this paper, we use the fitness function as its objective function. For the crossover operation, we select two parent solutions. The various selection operators have one principle in common, in that the probabilistically superior solution is the one to be chosen. In this study, we used the simplest and most frequently used roulette wheel method.

4.2.3 Crossover Operator. After the reproduction phase is over, the population is enriched with better individuals. Reproduction makes clones of goods strings, but does not create new ones. Cross over

operator is applied to the mating pool with the hope that it would create a better string. The aim of the crossover operator is to search the parameter space. In addition, search is to be made in a way that information stored in the present string is maximally preserved because these parent strings are instances of good strings selected during reproduction.

Parent 1	31	57	43	189	100
Parent 2	15	78	20	42	111

Offspring 1	31	57	20	42	111
Offspring 2	15	78	43	189	100

Figure 4: Example for crossover operation

4.2.4 Mutation Operator. After cross over, the strings are subjected to mutation. Mutation of a bit involves flipping it, changing 0 to 1 and vice versa with a small mutation probability P. The bit-wise mutation is performed bit-by-bit by flipping a coin with a probability P. Flipping a coin with a probability of P is simulated as follows. A simple genetic algorithm treats the mutation only a secondary operator with the role of restoring lost genetic materials.

5. Numerical analysis

For illustrative purpose, the problem described in section 3 was solved by using the proposed GA. For that, the code for the proposed GA was created in a programming language known as C++ and was run with certain parameters. These parameters are; population size =50, maximum number of generations =40; crossover rate = 80%; mutation rate = 10%. The figure below shows the best fitness values of each generation as a function of the number of generations. The GA solution procedure was executed on a desktop computer equipped with a Pentium D processor with a speed of 2.66GHz, and 2.00 GB of memory.
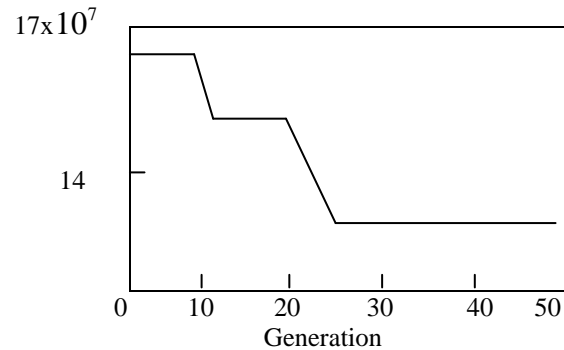

Figure 5. Evaluation curve of Genetic Algorithm

Table 5. The best solution for the selection of Warehouses

WAREHOUSES	NEW SITE COORDINATES	
	X	Y
W1	1350	2850
W2	900	1650
W3	1200	1200
W4	1350	900
W5	1650	600

6. Results and Discussions

The existing transportation cost from the plant to the warehouse and from the warehouses to the customers is found to be 14.65 crores. From GA the new transportation cost from the plant to the re allocated warehouses and from the warehouses to the customers is obtained as 13.7 crores. Hence there is a saving off of 95 lakhs per month. The new proposed locations are in Satara(Maharashtra), Kanpur(UP), Madurai(TN), Hyderabad, and Tirichirapally(TN) .

Thus this paper shows how proper facility location planning can improve customer satisfaction, increase return on assets, maximize speed for quick customer response, and effectively use people, equipment, space and energy to improve the profit of the company.
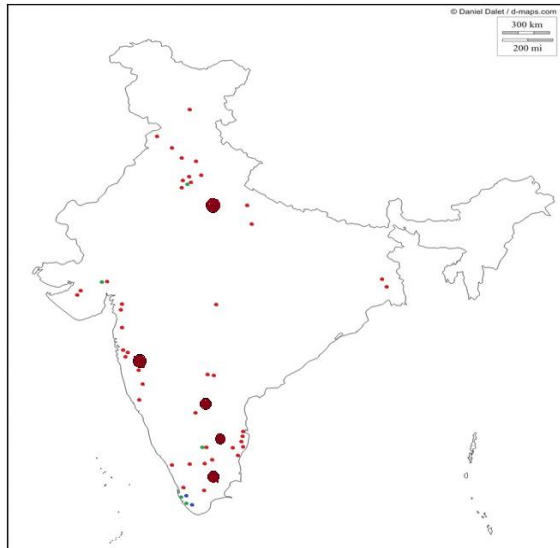

Figure 6. New Location of Warehouses in India

7. CONCLUSION

From this paper we can see that facility location planning plays a vital role in the proper placement of warehouse locations to reduce the transportation costs. This paper proposes a mathematical model and GA which aim to provide a minimum cost solution for the facility location problem. Found out the best solution for the location and allocations of the facilities to minimize the transportation costs. For the future research, the work can be extended to include selling price of finished goods, and consider multi mode transportation system etc.

REFERENCES

[1] Zvi Drezner , Horst W.Hamacher "facility location applications and theory", springer 1st edition.

[2] Andreas Klose, Andreas Drexl "Facility location models for distribution system" European journal of operational research 162(2005)4-29

[3] Eoksu Sim, Sungwon Jung,Haejoong Kim, and Jinwoo Park "A Generic Network Design for a Closed-Loop Supply Chain Using Genetic Algorithm.

[4] S.Rajasekaran, G.A.vijaylakshmi Pai "Neural Networks,Fuzzy Logic, and Genetic Algorithms synthesis and applications.

[5] Mark S.Daskin, Lawrence V. Snyder, Rosemary T. Berger, "Facility Location In supply Chain Design", Logistics Systems: Design and Optimization, Chapter 2.