

An Alternative Approach using Genetic Algorithm Based Heuristics for Capacitated Maximal Covering Location Allocation Problem

S. Sarifah Radiah Shariff*

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia
E-mail: shari990@salam.uitm.edu.my

Noor Hasnah Moin

Institute of Mathematical Sciences, Faculty of Science, University of Malaya,
50603 Kuala Lumpur, Malaysia
E-mail: noor_hasnah@um.edu.my

Mohd Omar

Institute of Mathematical Sciences, Faculty of Science, University of Malaya,
50603 Kuala Lumpur, Malaysia
E-mail: mohd@um.edu.my

Abstract

The Maximal Covering Location Problem (MCLP) has the objective of maximizing the total demand volume covered by a facility within a maximum allowable travel distance, S . In this paper, three sets of data 30-node, 324-node and 818-node networks with no existing facilities are analyzed as the capacitated model of MCLP (CMCLP) and first solved by the commercial optimization software, CPLEX. As the software produces results which violate the capacity constraint when the constraints are tight, an alternative approach of Genetic Algorithm based heuristics is used to solve the problem for more competitive results. Combination of the best number of facilities to open and the random node order assignment is used to maximize the percentage of total demand covered. The approach is found to solve all the capacitated MCLP in shorter time and in more promising result compared to CPLEX. Finally, the result of the approach is presented and applied to analyze a "real-world" example on a selected area, Telok Panglima Garang, Selangor which is currently served by 5 public health care facilities. This is to explore the implications of the model as the area already has existing facilities and to provide insight on future decisions that can be made on the expansion and development of current facilities.

Keywords: *Maximal Covering Location Problem (MCLP), CMCLP, genetic algorithm, random permutation*

1. Introduction

The growth field of locational analysis has been fed with deriving optimal and near-optimal solutions to locations problems for over the past three decades

[see Daskin (1995), Drezner (1995) and Mirchandani and Francis (1990)]. However due to the complexity of the problem in hand, many existing mathematical techniques fail to find optimal or near optimal

*Corresponding Author

solutions within a reasonable amount of time. Many attempts have been made to devise heuristic solution procedures for the location and allocation problems and its variants which produce solutions of acceptable quality in reasonable amount of computer time (see for example Maniezzo et al (1998), Lorena and Senne (2004), Diaz and Fernandez (2006), Osman and Ahmadi (2002, 2004, 2005), Fleszar and Hindi (2008), Galvao et al (1996, 2000) and Lorena and Lopes (1994). In particular, genetic algorithm based on natural evolution has been developed to solve location allocation problems [see Jaramillo (2002), Beasley and Chu (1996), Correa et al (2004) and Alp and Erkut(2003)].

In this paper, we propose a genetic based algorithm to solve the Maximal Covering Location Problem (MCLP) with the capacity constraint (CMCLP). The objectives of this paper are to evaluate the effectiveness of the new approach in solving the MCLP and CMCLP compared to the result initiated by commercial optimization software CPLEX. To further realize the efficiency of the new approach, a locational analysis has been done to the “real data” to analyze the efficiency of past locations decisions and the effects of constraints on location decisions that are presents.

This paper starts with a brief review of well-known formulation of MCLP. Afterwards, the capacity constraint is incorporated into the model to become CMCLP. In Section 3, genetic algorithm based heuristics is introduced. In the initial study, the representation of number of facilities to open is designed and followed by the representation of random node order assignment. Section 4 summarized the data used in the study, three sets of adapted data and real world data. Computational results on both adapted data and real world data are summarized in Section 5, which highlighted the overall better solution resulted from the proposed approach, when compared to CPLEX as it is the package used to obtain the initial optimal solution. The conclusion and proposals for future work are presented in the last section.

2. Modelling and problem formulation

The objective function for the MCLP is adapted from Pirkul and Schilling (1991). The sets I and J represent

the clients and sites for facilities respectively. Variable x_{ij} is 1 if client i is assigned to facility j and y_j is 1 if a facility is sited at j . a_i is the demand volume at demand node i and d_{ij} is the distance between demand node i and facility j . C_j is the capacity of a facility j . c_{ij} is 1 if the demand volume is assigned to a facility within the coverage distance S , where S is the maximum service distance or time. The objective maximizes the total population assigned to a facility within the coverage distance S .

$$\text{Maximize} \quad \sum_{i \in I} \sum_{j \in J} c_{ij} a_i x_{ij} \quad (*)$$

Subject to

$$\sum_{j \in J} y_j \leq p, \quad \forall j \in J \quad (1)$$

$$\sum_{j \in J} x_{ij} = 1, \quad \forall i \in I \quad (2)$$

$$x_{ij} \leq y_j \quad \forall i \in I, j \in J \quad (3)$$

$$\sum_{i \in I} \sum_{j \in J} c_{ij} a_i x_{ij} \leq C_j \quad \forall j \in J \quad (4)$$

$$y_j = 1 \quad \text{for all existing facilities} \quad (5)$$

$$x_{ij}, y_j \in [0,1] \quad \forall i \in I, j \in J \quad (6)$$

where

- I = the index set of all demand nodes,
- J = the index set of all facility sites,
- a_i = the demand volume at node i ,
- p = the number of facilities to be sited,
- d_{ij} = the travel distance from demand node i and facility j ,

$$x_{ij} = \begin{cases} 1 & \text{if the demand node } i \text{ is served by facility } j \\ 0 & \text{otherwise} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if a facility is sited at } j \\ 0 & \text{otherwise} \end{cases}$$

$$c_{ij} = \begin{cases} 1 & \text{if } d_{ij} \leq S \\ 0 & \text{otherwise} \end{cases}$$

In this formulation, constraint (1) limits the total number of facilities to no more than p , while constraint (2) ensures that all demand nodes are assigned to a facility. The level of service provided to covered demand is obviously controlled by S ; however, an uncovered demand node could be assigned to any available facility, regardless of its proximity. Constraint (3) guarantees that a demand node is only allocated to an open facility and the capacity constraint on each facility is imposed by constraint (4). Constraints (5) and (6) fix the locations of the facilities that already exist and impose the integrality restriction respectively.

3. Genetic algorithm-based heuristics

Genetic algorithms are implemented as a computer simulation in which a population of abstract representations (called chromosomes or the genotype of the genome) of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and in each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been attained, or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached.

3.1 Initial study

In the initial study we solve the MCLP using the commercial optimization software CPLEX and found very interesting result. The result is explained in the next section. In order to compare the results obtained, we design a Genetic Algorithm to determine the number of facilities to open and the algorithm is written in MATLAB. The natural representation scheme for the location allocation problem is a binary vector of length n , where n is the number of facilities and element i in the chromosome is given by (Jaramillo, 2002)

$$i = \begin{cases} 1 & \text{if the facility is open} \\ 0 & \text{otherwise} \end{cases}$$

3.2 Capacitated Facility Problem

From the initial study, we expand the model to include the capacity constraints imposed at each facility. Due to the tight capacity constraint imposed on some of the facilities, it is observed that the solutions produced by CPLEX violate some of the facility constraints and the CPU time taken is fairly large. We propose a Genetic algorithm based heuristics to overcome this problem since the results from the initial study of un-capacitated MCLP case are very promising and are compatible with those found using the CPLEX.

We propose a representation that combines the representation employed in the un-capacitated case (the representation of number of facilities to open) and the random permutation of all the nodes that are to be assigned to the open facilities. The new

Figure 1. An example of a chromosome where only two facilities A and D are open. Points 1,2,3,4 and 5 are the demand points.

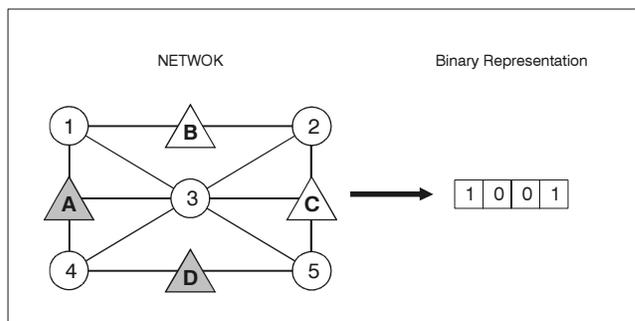
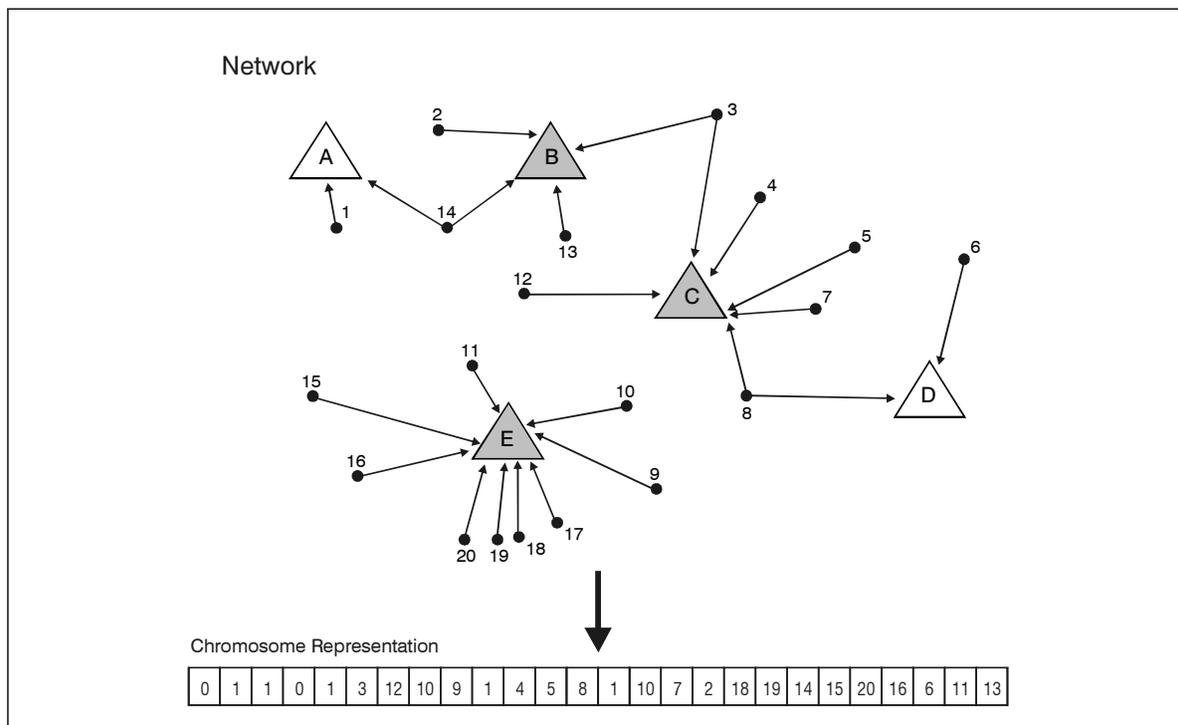


Figure 2. An example of a chromosome where only three facilities B, C and E are open. Points 1,2,3,4, 5 20 are the demand points.



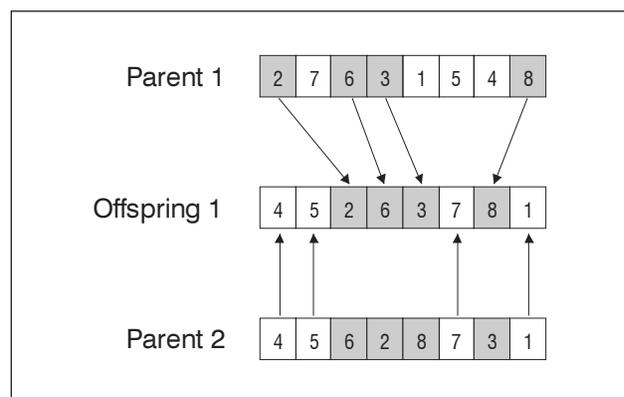
representation is a vector of length n where n and m are the number of facilities and the number of demand points respectively. The new representation comprises of a binary vector concatenated with a random permutation of integer vectors. We note that the first part of the chromosome identifies which facility or facilities that are open and the second part determines the way the allocation to the open facilities is traversed. This implies that the assignment of demand to the facilities is carried out according to the way the demand points appear in the chromosome. An example of 5 facilities and 20 demand nodes are given in Figure 2. In the example, demand point 3 is assigned first, followed by demand points 12, 10, 9, 1 and so on. Demand point 13 is considered last. It is interesting to observe that the way the assignment is carried out has substantial influence on the objective value obtained.

The implementation of crossover has to take into account the different types of representation that is encoded in the chromosome. In this study, the parent chromosomes are split into two separate vectors before the process of crossover takes place. The binary vector of the selected chromosome will

undergo a uniform crossover whilst an order based crossover proposed by Syswerda (1990) is implemented for the integer permutation vector. The implementation of the process of crossover is illustrated in Figure 3.

It is interesting to note that the same procedure is applied to the mutation process. The binary vector will undergo the classical mutation operator of

Figure 3. An example of order based crossover operator (OBX) when there are only 8 demand points



flipping from zero to one and vice versa, whilst the integer representation will undergo a simple insertion operator. We implement, for all the GAs, a probability of crossover of 0.7 and a mutation rate of 0.1. We chose incremental replacement method in which the average fitness of the population will improve if the child solutions have better fitness values than those of the solutions replaced. This method was chosen over the other method, generational replacement which generates a new population of children and replaces the whole parent population (Beasley and Chu, 1996). Using this method, the best solutions are always in the population and the newly created solutions are immediately available for selection and reproduction. Note that when replacing a solution, care must be taken to prevent excessive copies of a solution from entering the population. Allowing too many duplicate solutions to exist in the population may be undesirable because a population could come to consist of identical solutions, causing premature convergence and thus severely limiting the GAs ability to explore new search space.

The whole process above can be summarized as follows:

Step 1: Generate initial population. Each chromosome comprises of a binary vector of length n and an integer permutations of length N .

Step2: Perform selection based on the fitness value of the chromosome using the Stochastic Universal Sampling.

Step 3: Perform crossover by first splitting the parent chromosomes into two separate vectors: the binary vector and the integer representation. Perform crossover separately for each vector. Recombine both vectors to form a legal offspring.

Step 4: Perform mutation on the offspring.

Step 5: Repeat until maximum number of generations is attained.

The objective function for the MCLP is given by equations (*). However, when the constraints are tight, there is a tendency for the capacity constraints to be violated, or some of the demand nodes are not assigned. In view of this, we modified the present objective function to incorporate the penalty term

when there exists some demand nodes that are not assigned or when the capacity constraint of the facilities are violated. The new objective function is as follows:

$$\sum_{j \in J} \sum_{i \in I} c_{ij} a_{ij} x_{ij} + \beta \sum_{j \in J} \sum_{i \in I} c_{ij} \quad (**)$$

where β is the penalty factor for the MCLP.

4. Data

In this paper we used three data sets, 30-nodes, 324-nodes and 818-nodes network, which were adapted from Correa, Chaves and Lorena (2008). The 30-nodes network data was provided by Marianov and Serra (1998) and used together with the respective two others, network data obtained from a geographical data base of Sao Jose dos Campos, Brazil by Correa et al. In order to measure the applicability of the approach the following parameter values have been calculated for the three sets and summarized in the table below. The number of facilities to be sited, p , is chosen based on result produced by Correa et al. Since each demand node is a potential location of the facility, the total number of demand nodes is randomized to determine which demand node would also be the facility site. In each data set, every facility is assumed to have the same capacity and the capacity volume per facility is set to be 70 percent above the average demand volume per facility. The 70 percent is chosen as one demand node can have more than average demand volume. For example, in Set III node 814 has demand volume of 992 compared to the average demand volume per facility for Set III is only 583. The average distance between demand nodes and the facilities are also calculated in order to determine the maximum allowable distance S .

The real data set is taken from the network location of public healthcare facilities in Telok Panglima Garang, a district in the state of Selangor, Malaysia. It is located about 27 km southeast of Shah Alam (the capital of the state of Selangor), and comprises a small town centre, a few medium-sized residential areas and villages as well as small to medium workshops and factories. At present, the area is served by 5 units of facilities, is 8071 hectares in size, with total potential visitors of 30458. The 5

Table 1. Profile for 3 sets of network

| Description | Set I | Set II | Set III |
|--|-------|--------|---------|
| Number Nodes | 30 | 324 | 818 |
| Number of facilities to site - p | 5 | 20 | 50 |
| Total demand volume | 5470 | 12152 | 29168 |
| Average demand per facility | 1094 | 608 | 583 |
| Capacity volume per facility | 1900 | 1000 | 1000 |
| Average distance (between demand nodes and facilities) | 1.52 | 1.23 | 2.59 |

facilities are presented as SD, SL, KM, KB and TPG. The population in the study area is sparsely distributed including an approximately 18 hectares or 24 percent of the area which is unpopulated, as there are chicken (poultry) and palm oil farms situated in this area. Each clinic is assigned a target population to service per year. Table 2 summarizes the target population breakdown for the five clinics under study. It also describes the breakdown of target service volume for every clinic.

In this study it is assumed that the total population of 30458 is distributed uniformly within the area of study. The whole area is divided into 179 smaller sub regions, with their demand nodes located centrally and are approximately 1 km apart from each other. As the area comprises of five service boundaries with different target population volume, the assignment of demand nodes are done in two ways:

1. The population demand is uniformly distributed *within its own service boundary*.
2. The population demand is uniformly distributed *within the whole area of study*

We note that the demand volume for each node differs within the different service boundary regions when the population demand is uniformly

Table 2. Targets for consultations and health provision for Year 2007 (Telok Panglima Garang Health Clinics)

| | TPG | KB | SD | KM | SL | Total |
|---|-------|-------|------|-------|-------|-------|
| Total Population | 22722 | 11128 | 5696 | 14109 | 12585 | 66240 |
| Potential Visitors (children 0-6 years old and women 15-44 years old) | 10085 | 5114 | 2617 | 6485 | 6157 | 30458 |

Table 3. Demand nodes distribution before and after the revised assignment

| Service Area | Number of Demand Node | Volume per Demand Node |
|---------------------------------------|-----------------------|------------------------|
| Original 179 nodes | | |
| SD | 20 | 131 |
| SL | 32 | 192 |
| KM | 26 | 249 |
| KB | 46 | 111 |
| TPG | 55 | 183 |
| Total | 179 | |
| After Revised assignment to 221 nodes | | |
| SD | 20 | 131 |
| SL | 32 | 192 |
| KM | 45 | 144 |
| KB | 46 | 111 |
| TPG | 78 | 129 |
| Total | 221 | |

distributed *within its own service boundary*. It is also noted that the nodes, located in KM, are assigned very high demand volume at more than 240 per node. In order to reduce the high demand volume, the regions are further divided into smaller regions which are within approximately 0.5 km of each other and this results in an increase of the demand nodes from 179 to 221 nodes. Table 3 displays the demand nodes and their respective volume for each service area before and after the revised assignment.

We note that each node is assigned the same demand volume of 170 when the population demand is uniformly distributed *within the whole area of study*. It is observed that there are 47 nodes that fall into the unpopulated regions situated within the service areas of KB and TPG. For this, an analysis is carried out to examine the effect of assigning a demand volume of zero for all the 47 nodes that fall into the unpopulated region. Consequently, this results in a demand volume of 231 for the remaining nodes.

5. Results and discussions

The model is solved using CPLEX 5.2 to obtain optimal solution and the GA is written in MATLAB. Both algorithms are executed on a PC COMPAQ Presario V3000, equipped with 80 MB of RAM and an Intel Core Duo microprocessor running at 200 MHz.

5.1 Initial Result of Maximum Coverage Location Problem (MCLP)

The first part of the study examines the uncapacitated MCLP with the objective function to maximize the percentage of population covered within some maximum allowable distance, S . The optimal solutions are obtained using CPLEX.

Table 4 summarizes the result for Set I, the 30 node network. Both CPLEX and GA resulted in 95.2 percent of total demand volume covered with only 4 facilities opened. It is not possible to achieve 100% coverage for $S=3$ because there are three nodes of total demand volume 260 (or 4.8 percent) which are not within the coverage distance. Both results are obtained in less than 2 seconds. However, it can be seen that for CPLEX, all the demand volume are cumulated in the first assigned facilities which end up with Facility A having 4640. Similarly, for Set II, the 324 node network, both CPLEX and GA resulted in 100 percent coverage as seen in Table 5. The optimal solution requires 8 facilities whilst GA requires 9 facilities to be open respectively. However, as in data Set I, the total demand volume is cumulated mostly in Facility A, the earliest assigned facility. The result for Set III, the 818 node network in Table 6 is even more interesting with CPLEX only requires 7 facilities to be open compared to 20 by GA and Facility A has accumulated almost 13000 demand volume using CPLEX.

5.2 Results of Capacitated Maximum Coverage Location Problem (CMCLP)

The results for CMCLP are summarized in Table 4, 5 and 6. For Set I, both CPLEX and GA produce compatible result with the GA solutions requiring only 4 facilities to be open for both MCLP and CMCLP. We note that the solution for Set II did not converge and the best feasible solution found after 2 and a half hours (or almost 9000 seconds!) of running time. In the CMCLP case, there are two facilities which violate the capacity constraint of 1000 per facility and the total volume covered is only 93.8 compared to 100 percent by GA requiring 15 facilities to be open. We note that, in GA the capacity constraints are treated as hard constraints thus will not have any solution that may violate the capacity

constraint. GA requires an average time of 228 seconds. In Set III with even larger number of demand nodes, CPLEX only produces one feasible solution with 52.4 percent of total demand volume covered. GA achieves only 96.6 percent coverage with 33 open facilities. There is one node with high volume of 992 (or 3.4 percent) which could not be allocated within the capacity limit of 1000. The average running time taken is 1396 seconds.

5.3 Result of the real data

Ministry of Health Malaysia (MOH) has laid down in its planning policy that the capacity for each Rural Clinic (RC) is to serve 4000 population only. The statistics from the ministry showed that 88.5% of population lives within 5 km of a health facility and 81% within 3 km (NST, 2007). Based on this information, the analysis is carried out on the two maximum allowable distance values of $S=3$ km and 5 km respectively.

Table 7 and 8 present the results for the two assumptions mentioned in Section 4. For the first case where the population demands are distributed uniformly within the service boundary area, it is observed that the choice of facilities to be opened is exactly the same for the 179 and 221 nodes (refer to Table 7). However the results differ slightly in the percentage of the population covered when only two facilities are open. The percentage coverage is slightly higher for the 221 nodes and this is due to the demand volume being more scattered at the facility area. It is interesting to note that 100 percent coverage is achieved for $S=3$ km when only 3 facilities are open for both numbers of nodes. When S is reduced to 5 km, a full coverage is only achievable when all the 5 facilities are open. It is interesting to observe that if only one facility is open, 100 percent coverage can only be realized when $S=5$ km.

Table 8 tabulates the results when the demand nodes are uniformly distributed over the whole service area. It is found that the highest coverage achieved is only 99.4 percent and it is noted that the nodes that are not covered are those (47 nodes) that are located in the unpopulated areas. Consequently, 100 percent coverage is achieved when the demand volumes for these nodes are reduced to 0. This implies that the percentage of coverage is still satisfactory (more than 95 percent) if there is a new development in the currently unpopulated area.

Table 4. 30 node network (Set I)

| Facilities | MCLP (S=1.52) | | | | Capacity | CMCLP (S=1.52) | | | |
|-------------------------------|---------------|---------------|---------------|---------------|-------------|----------------|---------------|---------------|---------------|
| | CPLEX | | GA | | | CPLEX | | GA | |
| | No of nodes | Demand Volume | No of nodes | Demand Volume | | No of nodes | Demand Volume | No of nodes | Demand Volume |
| A | 21 | 4640 | 7 | 2520 | 1900 | 8 | 1500 | 0 | 0 |
| B | 4 | 370 | 5 | 400 | 1900 | 9 | 1330 | 8 | 1770 |
| C | 3 | 280 | 5 | 460 | 1900 | 4 | 360 | 5 | 460 |
| D | 2 | 180 | 10 | 1830 | 1900 | 6 | 1260 | 9 | 1520 |
| E | 0 | 0 | 0 | 0 | 1900 | 3 | 1020 | 5 | 1460 |
| TOTAL | 30 | 5470 | 27 | 5210 | 9500 | 30 | 5470 | 27 | 5210 |
| Population Volume Covered | 5210 | | 5210 | | | 5210 | | 5210 | |
| Percentage of Coverage | 95.20% | | 95.20% | | | 95.20% | | 95.20% | |

Table 5. 324 node network (Set II)

| Facilities | MCLP (S=1.23) | | | | Capacity | CMCLP (S=1.23) | | | |
|-------------------------------|----------------|---------------|----------------|---------------|--------------|----------------|---------------|----------------|---------------|
| | CPLEX | | GA | | | CPLEX | | GA | |
| | No of nodes | Demand Volume | No of nodes | Demand Volume | | No of nodes | Demand Volume | No of nodes | Demand Volume |
| A | 168 | 4619 | 0 | 0 | 1000 | 60 | 1588 | 17 | 342 |
| B | 21 | 484 | 0 | 0 | 1000 | 16 | 486 | 28 | 755 |
| C | 70 | 3313 | 29 | 1158 | 1000 | 13 | 641 | 27 | 983 |
| D | 3 | 87 | 0 | 0 | 1000 | 13 | 618 | 28 | 1000 |
| E | 44 | 2049 | 64 | 1521 | 1000 | 16 | 956 | 44 | 1000 |
| F | 0 | 0 | 17 | 932 | 1000 | 16 | 1056 | 17 | 943 |
| G | 10 | 771 | 0 | 0 | 1000 | 20 | 835 | 0 | 0 |
| H | 0 | 0 | 0 | 0 | 1000 | 13 | 369 | 0 | 0 |
| I | 0 | 0 | 0 | 0 | 1000 | 21 | 571 | 30 | 998 |
| J | 4 | 223 | 11 | 249 | 1000 | 13 | 337 | 20 | 501 |
| K | 4 | 606 | 0 | 0 | 1000 | 10 | 421 | 11 | 994 |
| L | 0 | 0 | 0 | 0 | 1000 | 15 | 508 | 0 | 0 |
| M | 0 | 0 | 14 | 304 | 1000 | 13 | 411 | 0 | 0 |
| N | 0 | 0 | 73 | 1828 | 1000 | 13 | 337 | 10 | 243 |
| O | 0 | 0 | 33 | 2943 | 1000 | 10 | 725 | 12 | 1000 |
| P | 0 | 0 | 41 | 1426 | 1000 | 15 | 538 | 16 | 998 |
| Q | 0 | 0 | 0 | 0 | 1000 | 12 | 315 | 30 | 770 |
| R | 0 | 0 | 0 | 0 | 1000 | 8 | 350 | 20 | 996 |
| S | 0 | 0 | 0 | 0 | 1000 | 18 | 643 | 14 | 629 |
| T | 0 | 0 | 42 | 1791 | 1000 | 9 | 447 | 0 | 0 |
| TOTAL | 324 | 12152 | 324 | 12152 | 20000 | 324 | 12152 | 324 | 12152 |
| Population Volume Covered | 121152 | | 12152 | | | 11398 | | 12152 | |
| Percentage of Coverage | 100.00% | | 100.00% | | | 93.80% | | 100.00% | |

Table 6. 818 node network (Set III)

| Facilities | MCLP (S=25.9) | | | | Capacity | CMCLP (S=25.9) | | | |
|---------------------------|---------------|---------------|-------------|---------------|----------|----------------|---------------|-------------|---------------|
| | CPLEX | | GA | | | CPLEX | | GA | |
| | No of nodes | Demand Volume | No of nodes | Demand Volume | | No of nodes | Demand Volume | No of nodes | Demand Volume |
| 1 | 296 | 12427 | 21 | 111 | 1000 | 13 | 270 | 0 | 0 |
| 2 | 64 | 1638 | 0 | 0 | 1000 | 16 | 375 | 18 | 783 |
| 3 | 0 | 0 | 0 | 0 | 1000 | 15 | 239 | 0 | 0 |
| 4 | 312 | 9324 | 54 | 4655 | 1000 | 18 | 552 | 17 | 996 |
| 5 | 6 | 95 | 0 | 0 | 1000 | 12 | 533 | 26 | 999 |
| 6 | 15 | 276 | 16 | 290 | 1000 | 13 | 1097 | 26 | 890 |
| 7 | 0 | 0 | 31 | 606 | 1000 | 19 | 606 | 21 | 958 |
| 8 | 0 | 0 | 0 | 0 | 1000 | 20 | 968 | 15 | 1000 |
| 9 | 89 | 4343 | 0 | 0 | 1000 | 16 | 888 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 1000 | 12 | 378 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 1000 | 14 | 255 | 27 | 993 |
| 12 | 36 | 1065 | 0 | 0 | 1000 | 19 | 576 | 38 | 901 |
| 13 | 0 | 0 | 0 | 0 | 1000 | 13 | 503 | 12 | 998 |
| 14 | 0 | 0 | 0 | 0 | 1000 | 21 | 659 | 14 | 147 |
| 15 | 0 | 0 | 34 | 746 | 1000 | 21 | 2056 | 35 | 991 |
| 16 | 0 | 0 | 0 | 0 | 1000 | 19 | 606 | 15 | 987 |
| 17 | 0 | 0 | 0 | 0 | 1000 | 16 | 471 | 23 | 1000 |
| 18 | 0 | 0 | 47 | 2418 | 1000 | 22 | 498 | 27 | 999 |
| 19 | 0 | 0 | 27 | 583 | 1000 | 16 | 792 | 28 | 984 |
| 20 | 0 | 0 | 0 | 0 | 1000 | 19 | 574 | 27 | 987 |
| 21 | 0 | 0 | 0 | 0 | 1000 | 15 | 743 | 0 | 0 |
| 22 | 0 | 0 | 0 | 0 | 1000 | 17 | 625 | 24 | 996 |
| 23 | 0 | 0 | 18 | 941 | 1000 | 21 | 417 | 0 | 0 |
| 24 | 0 | 0 | 0 | 0 | 1000 | 20 | 612 | 15 | 1000 |
| 25 | 0 | 0 | 35 | 509 | 1000 | 11 | 394 | 36 | 330 |
| 26 | 0 | 0 | 0 | 0 | 1000 | 20 | 832 | 0 | 0 |
| 27 | 0 | 0 | 20 | 631 | 1000 | 15 | 220 | 0 | 0 |
| 28 | 0 | 0 | 0 | 0 | 1000 | 10 | 147 | 32 | 995 |
| 29 | 0 | 0 | 0 | 0 | 1000 | 9 | 137 | 5 | 196 |
| 30 | 0 | 0 | 0 | 0 | 1000 | 20 | 612 | 16 | 200 |
| 31 | 0 | 0 | 0 | 0 | 1000 | 11 | 720 | 30 | 924 |
| 32 | 0 | 0 | 49 | 1158 | 1000 | 19 | 580 | 0 | 0 |
| 33 | 0 | 0 | 0 | 0 | 1000 | 15 | 550 | 0 | 0 |
| 34 | 0 | 0 | 73 | 3931 | 1000 | 10 | 244 | 29 | 951 |
| 35 | 0 | 0 | 94 | 2143 | 1000 | 13 | 314 | 0 | 0 |
| 36 | 0 | 0 | 0 | 0 | 1000 | 21 | 741 | 14 | 1000 |
| 37 | 0 | 0 | 0 | 0 | 1000 | 19 | 1011 | 0 | 0 |
| 38 | 0 | 0 | 22 | 239 | 1000 | 16 | 364 | 8 | 109 |
| 39 | 0 | 0 | 0 | 0 | 1000 | 19 | 671 | 0 | 0 |
| 40 | 0 | 0 | 0 | 0 | 1000 | 14 | 994 | 29 | 987 |
| 41 | 0 | 0 | 84 | 3563 | 1000 | 23 | 518 | 23 | 987 |
| 42 | 0 | 0 | 7 | 145 | 1000 | 12 | 267 | 10 | 891 |
| 43 | 0 | 0 | 54 | 3387 | 1000 | 21 | 1047 | 0 | 0 |
| 44 | 0 | 0 | 0 | 0 | 1000 | 20 | 1026 | 0 | 0 |
| 45 | 0 | 0 | 35 | 1001 | 1000 | 17 | 517 | 36 | 1000 |
| 46 | 0 | 0 | 0 | 0 | 1000 | 17 | 395 | 0 | 0 |
| 47 | 0 | 0 | 55 | 1016 | 1000 | 16 | 470 | 47 | 997 |
| 48 | 0 | 0 | 42 | 1095 | 1000 | 11 | 325 | 56 | 1000 |
| 49 | 0 | 0 | 0 | 0 | 1000 | 24 | 619 | 0 | 0 |
| 50 | 0 | 0 | 0 | 0 | 1000 | 8 | 160 | 38 | 1000 |
| TOTAL | 818 | 29168 | 818 | 29168 | 50000 | 818 | 29168 | 817 | 28176 |
| Population Volume Covered | 29168 | | 29168 | | | 15283 | | 28176 | |
| Percentage of Coverage | 100.00% | | 100.00% | | | 52.40% | | 96.60% | |

Table 7. Coverage Percentage when demand nodes are uniformly distributed within service boundary only (S=5km)

| Number of Nodes | Number of Facilities | Objective Function Value | Coverage Percentage | Open Facility |
|-----------------|----------------------|--------------------------|---------------------|-----------------|
| 179 | 1 | 26963 | 88.5 | KM |
| | 2 | 30085 | 98.8 | SL,KB |
| | 3 | 30458 | 100 | SD,KM,TPG |
| | 4 | 30458 | 100 | SD,SL,KM, TPG |
| | 5 | 30458 | 100 | SD,SL,KM,KB,TPG |
| 221 | 1 | 26587 | 87.3 | KM |
| | 2 | 30270 | 99.4 | SL,KB |
| | 3 | 30458 | 100 | SD,KM,TPG |
| | 4 | 30458 | 100 | SD,SL,KM, TPG |
| | 5 | 30458 | 100 | SD,SL,KM,KB,TPG |

Table 8. Coverage percentage when demands are uniformly distributed over the whole study area (S=5km)

| Number of Nodes | Number of Facilities | Objective Function Value | Coverage Percentage | Open Facility |
|-----------------|----------------------|--------------------------|---------------------|-----------------|
| 179 | 1 | 22788 | 74.8 | KM |
| | 2 | 28925 | 95 | SD,KM |
| | 3 | 30287 | 99.4 | SD,KM,TPG |
| | 4 | 30287 | 99.4 | SD,SL,KM, TPG |
| | 5 | 30287 | 99.4 | SD,SL,KM,KB,TPG |
| 132 | 1 | 25157 | 82.6 | KM |
| | 2 | 29997 | 98.5 | SD,KM |
| | 3 | 30458 | 100 | SD,KM,TPG |
| | 4 | 30458 | 100 | SD,SL,KM, TPG |
| | 5 | 30458 | 100 | SD,SL,KM,KB,TPG |

Further analysis shows that all the uncovered nodes fall in the unpopulated areas. Consequently, full coverage is obtained when the nodes in the unpopulated area are assigned a demand of zero. As in the previous model, full coverage is only achieved when only one facility is open is when the maximum allowable distance *S* is 7 km.

We note that, in practice, every facility in the area is assigned a different demand volume to serve in proportion to its nominal capacity. It can be seen from Table 2, in which facilities SD and KB are assigned the demand volume 2617 and 5114, respectively. However, applying the MCLP models has reassigned the demand nodes and its volume which are very different from the actual practice, as shown in Table 9. This is an interesting result to observe as this gives indication on the ideal capacity of each facility in order to achieve a maximum coverage. This result may also assist future decisions on the expansion of the current facilities.

5.4 Result for the Capacitated Maximum Covering Location Problem (CMCLP)

Referring to Table 2, some of the facilities, such as SL, KM and KB are operating above their capacities (of 4000 each) by 54%, 62% and 28% respectively, whilst the other facilities, SD and TPG are operating below their respective capacities. They are operating within 35% and 50% of their capacities only. In view of this, the CMCLP is applied and results are compared to the existing practice.

Table 9. Comparing the number of demand nodes and volume assigned to each facility

| Number of facilities | No model (Existing assignment) | | MCLP Within Its Own Service Boundary | | MCLP Within The Whole Study Area | |
|----------------------|-----------------------------------|--------------|---|--------------------------|-------------------------------------|--------------------------|
| | Number of nodes assigned | Volume | Volume | Number of nodes assigned | Volume | Number of nodes assigned |
| SD | 20 | 2617 | 105 | 10620 | 105 | 14060 |
| SL | 32 | 6157 | 39 | 10626 | 39 | 8546 |
| KM | 26 | 6485 | 27 | 6406 | 27 | 6006 |
| KB | 46 | 5114 | 6 | 2060 | 6 | 1386 |
| TPG | 55 | 10085 | 2 | 746 | 2 | 460 |
| TOTAL | 179 | 30458 | 179 | 30458 | 179 | 30458 |

Table 10 summarizes the results for the CMCLP when the total demand volume is **uniformly distributed within its own service boundary**. It is observed that for the capacity that is based on the government planning policy, 4 facilities are assigned demand volume more than its nominal capacity by CPLEX when the maximum distance S=3 km. The coverage is also as low as 38.1 percent only. GA produces better results where all the capacities are not violated and the coverage increases to 90.6 percent. It is worth noting that when S=5 km, the coverage increases to 91.6 percent for CPLEX with GA achieving 100 percent coverage. It is also noted that CPLEX produces result that violate the capacity constraint of facility SL.

The results for the population demand volume being **uniformly distributed within the whole study area** are illustrated in Table 11. When the capacity follows the government policy, the results using

CPLEX violate the capacity constraints for all the facilities except TPG. As mentioned earlier CPLEX will always assign all nodes regardless whether the capacity or the maximum coverage constraints are violated as opposed to GA where all constraints are treated as hard constraints. We note that the results produced by GA do not manage to assign all the demand nodes. It is observed that almost 25 percent (51) nodes have not been assigned for S=3 km and this is reduced to 3 percent (5) nodes when S=5 km. However the coverage produced by GA is higher in both cases.

The approach with GA produces overall better results compared to CPLEX. The overall percentage of coverage for this case is also a bit worst compared to when demand volume is distributed within its own service boundary. This is due to the fact that the density of demand volume actually varies between different service boundaries.

Table 10. Comparison of CPLEX versus GA results for CMCLP (demand volume uniform within its own service boundary)

| Facilities | Capacity Based on Govt Policy | CMCLP (S=5km) | | | | CMCLP (S=3km) | | | |
|-------------------------------|-------------------------------------|-----------------|---------------|-----------------|---------------|-----------------|---------------|-----------------|---------------|
| | | CPLEX | | G A | | CPLEX | | G A | |
| | | Number of nodes | Demand Volume |
| SD | 4000 | 61 | 2870 | 39 | 3893 | 73 | 5747 | 31 | 3966 |
| SL | 4000 | 22 | 4213 | 22 | 3981 | 25 | 5625 | 19 | 3996 |
| KM | 4000 | 19 | 3828 | 37 | 3956 | 23 | 4867 | 31 | 3952 |
| KB | 4000 | 13 | 2724 | 18 | 3967 | 27 | 6309 | 20 | 3975 |
| TPG | 20000 | 64 | 16823 | 62 | 14661 | 31 | 7910 | 42 | 11713 |
| TOTAL | 36000 | 179 | 30458 | 179 | 30458 | 179 | 30458 | 143 | 27602 |
| Population Volume Covered | | 27980 | | 30458 | | 11618 | | 27602 | |
| Percentage of Coverage | | 91.9 | | 100 | | 38.1 | | 90.6 | |

Table 11. Comparison of CPLEX versus GA results for CMCLP (demand volume uniform within the whole service area)

| Facilities | Capacity Based on Govt Policy | CMCLP (S=5km) | | | | CMCLP (S=3km) | | | |
|-------------------------------|-------------------------------------|-----------------|---------------|-----------------|---------------|-----------------|---------------|-----------------|---------------|
| | | CPLEX | | G A | | CPLEX | | G A | |
| | | Number of nodes | Demand Volume |
| SD | 4000 | 63 | 4841 | 30 | 3918 | 71 | 6460 | 21 | 3921 |
| SL | 4000 | 17 | 3691 | 18 | 3919 | 29 | 6464 | 17 | 3923 |
| KM | 4000 | 16 | 3228 | 36 | 3923 | 24 | 5305 | 31 | 3927 |
| KB | 4000 | 4 | 923 | 17 | 3927 | 31 | 6694 | 17 | 3927 |
| TPG | 20000 | 79 | 17775 | 73 | 13850 | 24 | 5535 | 42 | 8081 |
| TOTAL | 36000 | 179 | 30458 | 174 | 29537 | 179 | 30458 | 128 | 23779 |
| Population Volume Covered | | 29308 | | 29308 | | 29537 | | 9465 | |
| Percentage of Coverage | | 96.2 | | 96.2 | | 97 | | 31.1 | |

6. Conclusions

In this paper, we proposed a new approach in solving the capacitated maximal covering location problem (CMCLP). The approach which is based on GA is applied in the MCLP case to assess the effectiveness of the algorithm and extended to solve the CMCLP. The solution is represented as a combination of binary and integer vector and the algorithm attempts to maximize coverage by simultaneously determining the number of facilities to be open and the order in which the demand nodes are to be assigned. This results in overall comparably better solutions than those found by CPLEX in terms of objective function values as well as computational times when tested on 3 randomly generated data sets.

The algorithm is then applied to the real data and an analysis on the past location decision in the study area, assuming that the demand volume is uniformly distributed is carried out. The initial solution was solved by using the commercial optimization software CPLEX in which each facility can serve as many demand volume. The location decision was further analyzed by adding the capacity constraint. The commercial optimization software such as CPLEX has produced incompatible result when it violates the capacity constraint in certain cases. The proposed algorithm based on GA obtained better results and in much shorter running processing time.

The analysis carried out on the data when the demand nodes are distributed uniformly in the whole service area provides an insight into future facilities that may be required if the currently unpopulated areas are to be developed. It also offers information on the future expansion of the current facilities.

References

- Ahmadi, S. and Osman, I.H. (2004). Density Based Problem Space Search for the Capacitated Clustering p -Median Problem. *Annals of Operations Research* 131 (1-14), pp. 21-43.
- Ahmadi, S. and Osman, I.H. (2005). Greedy Random Adaptive Memory Programming Search for the Capacitated Clustering Problem. *European Journal of Operational Research* 162(1), pp. 30-44.
- Alp, O, Erkut E. and Drezner Z (2003). An Efficient Genetic Algorithm for the p -Median Problem. *Annals of Operations Research* 122 (1-4), pp. 21-42, DOI: 10.1023/A: 1026130003508. At: <http://www.springerlink.com/content/qu06n670l285522x/>
- Beasley, J.E. and Chu P.C. (1996). A Genetic Algorithm for the Set Covering Location Problem. *European Journal of Operational Research* 94, pp. 392-404.
- Corrêa, E.S., Steiner, M.T.A., Freitas, A.A. and Carnieri, C. (2004). A Genetic Algorithm for Solving Capacitated p -Median Problem. *Numerical Algorithms* 35, pp. 373-388.
- Corrêa F.A., Chaves A.A. and Lorena L.A.N. (2008). Hybrid Heuristics for the Probabilistic Maximal Covering Location-Allocation Problem. *Operational Research: An International Journal* 7 (3), pp. 323-344.
- Daskin M.S. (1995). *Network and discrete location: models, algorithms, and applications*, Wiley, New York.
- Díaz, J.A. and Fernández, E. (2006). Scatter search and path re-linking for the capacitated p -median problem. *European Journal of Operational Research* 169 (2), pp. 570-585.
- Drezner Z., editor (1995). *Facility location: a survey of applications and methods*, Springer, New York.
- Fleszar K. and Hindi K. (2008). An Effective VNS for the Capacitated p -Median Problem. *European Journal of Operational Research* 191, pp. 612-622.
- Galvão R.D. and ReVelle C. (1996). A Lagrangean Heuristic For The Maximal Covering Location Problem. *European Journal of Operational Research* 88, pp. 114-123.
- Galvão, R.D., Espejo. L.G.A. and Boffey B. (2000). A Comparison Of Lagrangean And Surrogate Relaxations For The Maximal Covering Location Problem. *European Journal of Operational Research* 124, pp. 377-389.
- Jaramillo J.H., Bhadury J. and Batta R. (2004). On The Use Of Genetic Algorithm To Solve Location Problems. *Computers and Operations Research* 29, pp. 761-779.
- Lorena L. and Lopes F.B. (1994). A Surrogate Heuristics For Set Covering Problem. *European Journal of Operational Research* 79, pp. 138-150.
- Lorena L. and Senne E. (2004). A Column Generation Approach To Capacitated p -Median Model. *Computers and Operation Research* 31, pp. 863-876.
- Maniezzo V., Mingozzi, A. and Baldacci R. (1998). A Bionomic Approach To The Capacitated p -Median Model. *Journal of Heuristics* 4, pp. 263-280.
- Marianov V. and Serra D. (1998). Probabilistic Maximal Covering Location-Allocation Models For Congested Systems. *Journal of Regional Science* 38(3), pp. 401-424.
- Ministry of Health Annual Report 2005, 2005. Putrajaya, Selangor, Malaysia

- Mirchandani P.B. and Francis R.L. editors (1990). *Discrete location theory*, Wiley, New York.
- Osman I.H., Ahmadi, S. (2002). Guided Construction Search For The Capacitated p -Median Problem. Working paper, School of Business, American University of Beirut, Lebanon
- Pirkul, H., Schilling D. (1991). The Maximal Covering Location Problem With Capacities On Total Workload. *Management Science* 37(2), pp. 233-248.
- Syswerda G. (1990). Schedule Optimization Using Genetic Algorithms, in: L.Davis (Ed.) *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York.
- The New Strait Times* (2007). Doctors Urged To Treat, Not Refer To Hospitals: Director General of Health Tan Sri Dr Ismail Merican, Tuesday Feb 13, 8.
-

S.Sarifah Radiah Shariff is a doctoral student in the field of Operational Research at Institute of Mathematical Science, University of Malaya, Kuala Lumpur, and also a senior lecturer in Centre of Quantitative Methods, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Malaysia. She worked as a logistics executive doing inventory and transportation planning, warehouse management as well as demand forecasting in the international companies before joining UiTM as a statistics lecturer. She holds a Master's degree in Information Technology from UiTM and an undergraduate degree in Statistics and Mathematics from Purdue University, West Lafayette, IN, USA. Ms Shariff is currently focusing on the field of facility location decision as her scope of research.

Noor Hasnah Moin is an associate professor at the Institute of Mathematical Sciences, University of Malaya, Malaysia. She holds a doctoral degree in Applied and Computational Mathematics from University of Sheffield, United Kingdom and a master's degree in Operational Research, from the University of Sussex, also in the United Kingdom, and an undergraduate degree in Mathematics from the University of Malaya. Dr Noor Hasnah has been a reviewer for international journals such as International Journal of Production Economics, INFORMS Journal of Computing, European Journal of Operational Research, IEEE Transactions on Systems, Man and Cybernetics and local academic journals and has also served as reviewers for international and local conferences. Dr Noor Hasnah has published in academic journals, conference proceedings and working papers. Dr. Noor Hasnah is currently working in areas of vehicle routing, inventory routing and related problems.

Mohd Omar is a research assistant professor in the University of Malaya at Institute of Mathematical Sciences. He has conducted research in production and inventory planning, location allocation problem and yield management. He holds a doctoral degree in Operational Research from University of Exeter, United Kingdom and a master's degree in Management Science from University of Hull, United Kingdom. Dr M. Omar has been a frequent reviewer in academic journals and has also served in numerous professional organizations including the East Asia SIAM and International Society Inventory Research. Dr M. Omar has made numerous publications in academic journals, conference proceedings, technical reports and working papers.