Global Journal of Mathematical Sciences: Theory and Practical.
ISSN 0974-3200 Volume 9, Number 1 (2017), pp. 65-79

http://www.irphouse.com

Improved Facility Location for Logistic Network Design using the proposed K-means - GA Fusion Model

Shaju Varghese [1] and Gladston Raj S [2]

[1] Research Scholar, Bharathiar University, Coimbatore, Tamilnadu, India.
[2]Head of Department of CS, Govt. College, Nedumangadu , Thiruvananthapuram, Kerala, India.

Abstract

Logistic Network design and management is the process of identifying and locating an optimal location for providing the facility. The location or the decision of where to locate the facility is important and it determines the efficiency and effectiveness of the logistic management. The distribution centers(DCs) should located in easy to access road ways and closer to the populated area to accelerate the delivery and also to reduce the overall transportation cost and corresponding time. Also, it is a tedious process to find the optimal location for a DC and to minimize the number of DCs and thus the overall cost.

In our previous work[19][20][21], we proposed the soft computing based models for facility location in logistic network design. In our previous work[19], we presented a genetic algorithm based model for facility location. In this work we propose a K-Means – GA fusion hybrid model for facility location in logistic network design.

Keywords: Logistic, Heuristic, Hybrid, Inbounded, Crossover, Mutation, Reproduction, Population, clustering, Optimal

I. INTRODUCTION

The basic concept of facility location analysis in logistic management is the K-center problem and P-median model. It is the process of identifying the facility location and generates a common model for finding the optimal locations. In this work it evaluated by two soft computing models known as k-means clustering and genetic algorithm.

Also, these model generate a hybrid model using these models to locate an optimal locations for facilities.

Facility Location Problem(FLP)

The facility location analysis and identification is a challenging and non-linear problem in the areas of resource procurement & management, production, supply, warehouse management. This problem is faced by many organizations in their operational stages. The requirement of location decision-making has led to the development of the location analysis and modeling as a part of the operations research[3]. This is the process of finding the optimal locations (one or more) for this problem in order to locate minimum service distance for maximum number of users. The facility can be identified as the bases, units, equipment, weapon systems, logistics, civil objects, etc.[3]. Various levels of researches have been conducted in this area and they proposed various approaches as the solution.

Location models are often difficult to solve, especially for large problem instances[3]. There are number of built in commercial tools are available that can solve the complexity of a location models. Besides location models are application dependent. Their objectives, constraints and variables are determined by particular problem under study. It is very difficult to develop an all-purpose model that can solve the common facility location problem that is optimal for existing applications[3].

The major solution techniques are

Exact Solution Techniques.

Heuristic Solution Techniques.

The heuristics solution methods are

- Bound on Optimal Solutions
- Worst Case Analysis
- Statistical Evaluation

Applications of FLP

The basic mathematical modeling concept and algorithms that are suitable for developing this type of mathematical models are the basic goal of this type of development. But the areas of innovative applications of this type of model are also important and also are considered at the time of development. The common applications include land use optimization for emergency facilities, transport route identification and optimization, retail shop outlet location selection, spatial objects optimization, performance optimization of urban modeling. When applied to telecommunication network design and capacity planning, the difficult task of designing and maintaining huge, reasonably optimal, telecommunication networks becomes more feasible. Many major telecommunication service providers are finding it imperative to upgrade or expand their facilities and services. The facility location models can make use of the design of various types of telecommunication networks. The capacitated optimization technique can play an important role for the improvement of the telecommunication network design. In this example, most of the analysis decisions have to be made concerning relation between concentrator location and cable expansion.

II. FACILITY LOCATIONS & LOGISTICS

The process of locating an optimal location for facilities with considering the attributes like facility construction costs, transportation costs, etc. This issue will affect the smooth transferring of the materials or products from one place to another in normal or emergency situations. Many studies and researches in this area contribute number solutions but most of the solutions are not common and purely concentrated on a particular type of problem.

The purpose of this work is to build a decision support system that helps the managers decide where to locate a facility.

Time complexity of problems of the above mentioned problem is a obvious one, so that linear programming based methods will take much time with respect to the depth of the problem and the assigned conditions. In this work, we will design a soft computing model for solving a simple facility location problem with minimum constraints.

Logistic Management

Logistics is the management of a transfer of things from one place to another. Transportation is the mode of transfer in this process. This process involves the pack movement, warehousing, packing, production, material handling, integration of information flow, security etc.

History of Logistics

The concept of logistics was introduced late 19th century. Another definition for logistics is that it is the detailed co-ordination of a complex operation consisting of many people, warehouses, distribution etc. In a logistic management process the planning, implementing and controlling procedures for the systematic and time bounded transfer from one place to another.

The logistics activities can be classified as

1. Inbounded logistics – Arranging the internal movement of the materials.
2. Out bounded logistics – The products are moving from the production unit to the end user.

The major access fields are procurement, distribution, after sales, disposal, global/domestic, emergency, production, construction, asset control, reverse logistics.

Applications

1. Military Logistics

In this case it is very important to maintain the security. Also, the logistics management failures cause critical problems. An Integrated Logistics Support (ILS) is a discipline tool used to maintain the best service with lowest cost.

2. Business Logistics

   The major goal of the business logistics is that distribution of the products in the actual quality at the exact time at the correct place for the right price in the perfect condition. Since 1960s only the concept of business logistics came to the business environment. This is the basic concept of the shipping process of worldwide supply chain management.

Facility Location in Logistics Network

In Bansal [2], a Public Logistics Network (PLN) for the continental U.S. was designed. This research uses the same design approach that was developed by Bansal. Bansal's design can be explained as four step process that includes generation of the Underlying Road Network (URN), developing the network of public Distribution centers (DCs), estimation of average package delivery time, and finding public DC locations that minimize average package delivery time.

In this work, the PLN will be designed using a similar process with some modifications and simplification in the steps. In this work we use a simplified version of that design so that, instead of using the “average package delivery time” as a metric for optimization, we used simple distance as the metric in the fitness function of the soft computing model. This approach was used to minimize the optimization time. Since the distance is directly proportional to package delivery time, we believe that this approach also will lead to equal results, logically with in lesser time.

- Generation of the Underlying Road Network (URN)
- Developing the network of Distribution centers (DCs),
- Finding public DC locations that minimize the distance between the DCs and the User locations.

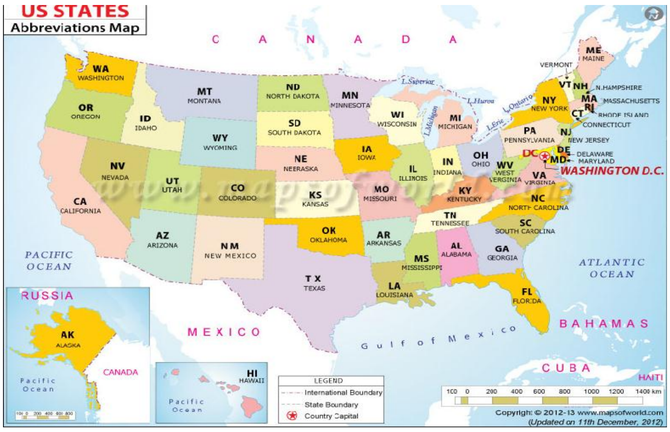


Figure 1: US states Map and their names

Generation of the Road Network of USA

The following map shows the road network that we created from the US census data set.

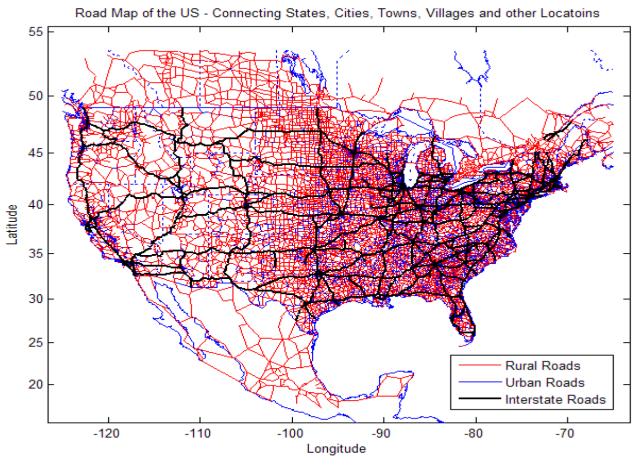


Figure 2. Road Network of USA

Generation of the Underlying Road Network of Regional Distribution Centers.

The population in RDC is represented by total 925 U.S. census blocks that are plotted on the map of RDC. A sub-graph of the road network was generated that is then followed by the removal of two-degree nodes from the network. Each point in this network is a potential location for a DC

The following graph/map shows the road network of Alabama(AL), USA that will be the example of a sub graph we created and used to create the regional distribution centers (RDC) that we are interested in.

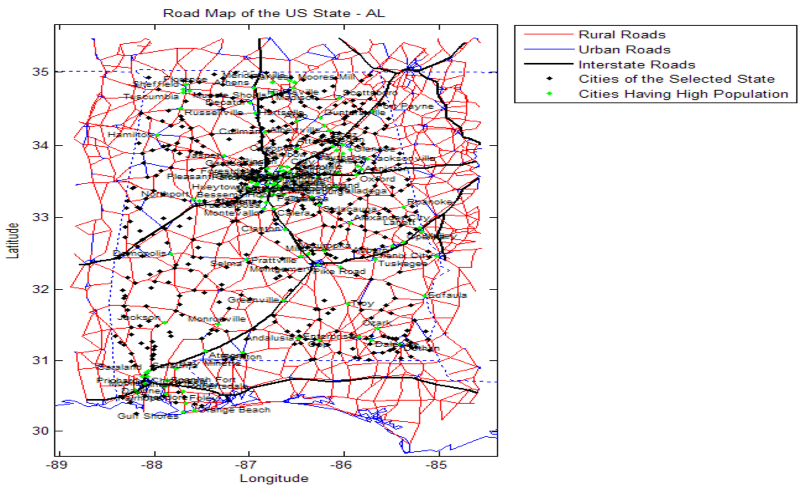


Figure 3. Underlying Road Network of RDCs

Network of DCs

If needed, we may also create the network of DCs as follows.

The arcs connecting census blocks to the URN are added to the network and the shortest time paths and distances between each pair of points are calculated using Dijkstra's algorithm.. DCs will be located at some of the key points and then connected to each other using Delaunay Triangulation [22] to form a network of public DCs. The shortest time paths between all pairs of DCs is found and those paths and distances are then used to calculate the percent flow of the packages, wij from DCi to DCj using order based proximity factors developed by Kay and Parlikad [4]

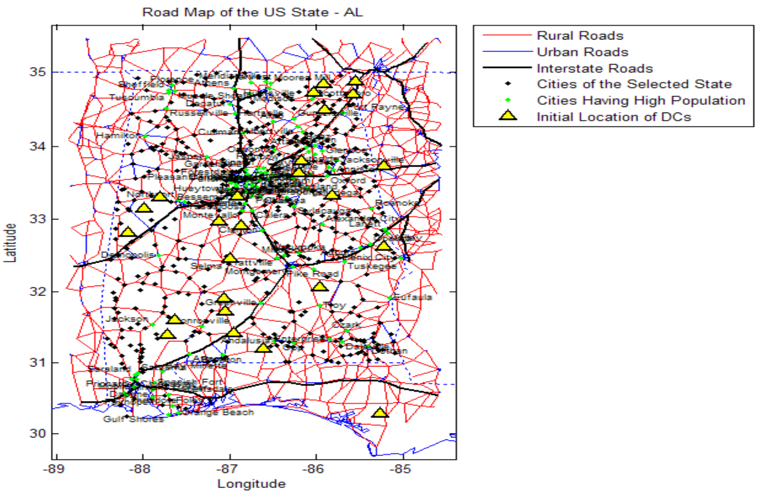


Figure 4: Randomly placed facilities

K-Means Clustering

Simply put, k-Means Clustering is an algorithm among several that attempt to find groups in the data. In pseudo code, it follow this procedure:

The vector **m** contains a reference to the sample mean of each cluster. **x** refers to each of our examples, and **b** contains our "estimated class labels.

Explained perhaps more simply in words, the algorithm roughly follows this approach:

1) Choose some manner in which to initialize the m_i to be the mean of each group (or cluster), and do it.
2) For each example in your set, assign it to the closest group (represented by m_i).
3) For each m_i, recalculate it based on the examples that are currently assigned to it.
4) Repeat steps 2-3 until m_i converge.

Now that we have some rudimentary understanding of what k-means

Initialize $\mathbf{m}_i$, $i = 1,\ldots,k$, for example, to k random $\mathbf{x}^t$
Repeat
    For all $\mathbf{x}^t$ in X
    $b_i^t \leftarrow 1$ if $\|\mathbf{x}^t - \mathbf{m}_i\| = \min_j \|\mathbf{x}^t - \mathbf{m}_j\|$
        $b_i^t \leftarrow 0$ otherwise
    For all $\mathbf{m}_i$, $i = 1,\ldots,k$
        $\mathbf{m}_i \leftarrow$ sum over t $(b_i^t \mathbf{x}^t)$ / sum over t (b_i^t)
Until $\mathbf{m}_i$ converge

Genetic Algorithm

Genetic algorithm is a *population-based search* method. Genetic algorithms are acknowledged as good solvers for tough problems. It is an iterative procedure maintaining a population of structures that are candidate solutions to specific domain challenges. During each temporal increment (called a generation), the structures in the current population are rated for their effectiveness as domain solutions, and on the basis of these evaluations, a new population of candidate solutions is formed using specific genetic operators such as reproduction, crossover, and mutation.

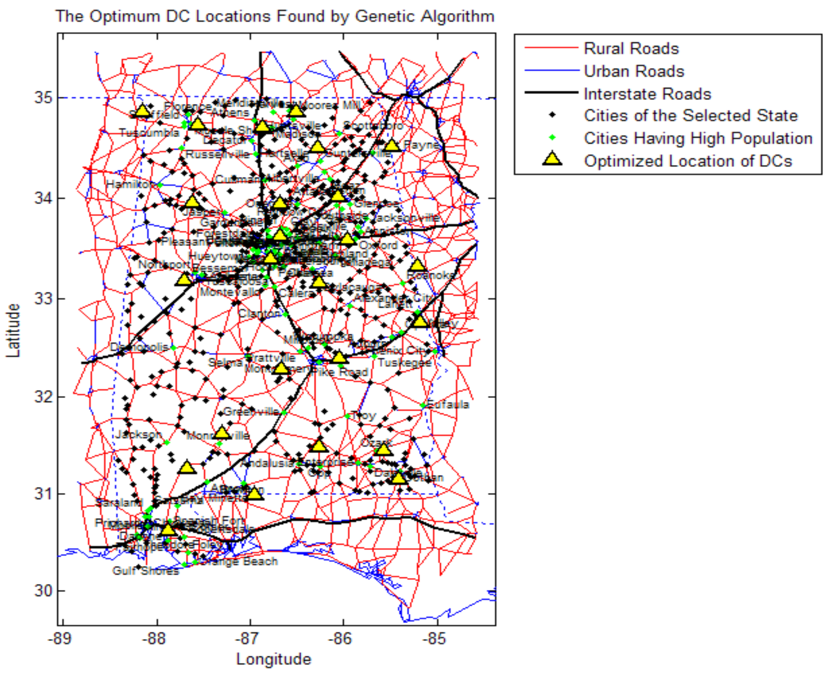


Figure 5: Optimum Locations by Genetic Algorithm

The K-Means-GA Fusion Method

We used linear Fusion approach in design. In this algorithm, instead of initializing the population with random candidate locations, we first use the k-means algorithm for doing the initial estimate on better starting locations.

The following pseudo code outlines the design of K-Means-GA Fusion for improving Facility Location.

```
Function k-mean-GA_FLP
begin
 kmeans :  initialize random centroids
           run k-means for n iterations to find
           N optimum initial candidate locations
 INITIALIZE population with N candidate
solutions provided by k-means algorithm;
   (Each candidate solution will represent N
locations of the facilities)
   EVALUATE each candidate;
      (find fitness of each candidate using the
fitness function)
   repeat
      SELECT parents;
         (Select two candidate having best
fitness value)
      RECOMBINE pairs of parents;
         (use single point crossover on the
         selected candidates and generate new
         population – this includes the original
         parents)
      MUTATE the resulting children;
         (use gaussian mutation on entire
population)
      EVALUATE children;
         (find fitness of all new candidates of the
population)
   until TERMINATION-CONDITION is
satisfied
end
Place the facilities on the locations optimized
by GA
```

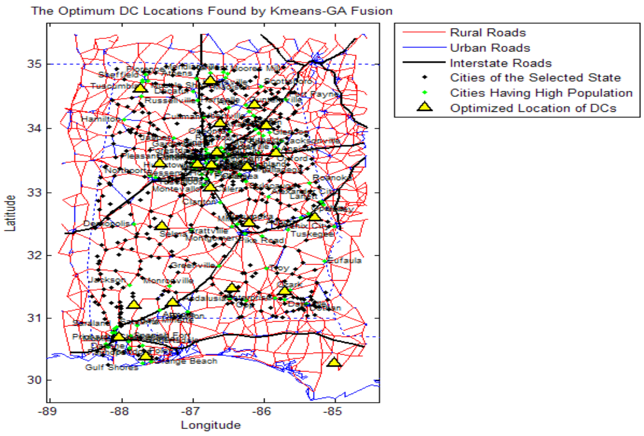


Figure 6: Facility Location by ***k-means-GA Fusion Method***

The Fitness Function

The following function is used to find the fitness at the set of facility location XY_F= (X_i, Y_j), where i,j = 1 to n. The set of points which has the lowest fitness vaue will be the optimum location for placing facilities. XY_U is the locations of all the customers(cities, towns, villages)

Function d= EuclideanDist(XY_U, XY_F**)**

Begin

// Compute the Euclidean distance with each coordinate

[R,C]=size(**XY_U**);

//sum squared data - save re-calculating repeatedly later

XYsq=repmat(sum(**XY_U**.^2,2),1,NoDCs);

// The distance Function d^2 = (x-c)^2 = x^2 + c^2 -2xc

Dist = XYsq + repmat (sum ((XY_F.^2)' ,1), R, 1) - 2.*(**XY_U** *(**XY_U** '));

//label points

[**d**,Classes]=min(Dist,[],2);

d =sqrt(sum(**d**));

return (d)

End

The Fitness Function

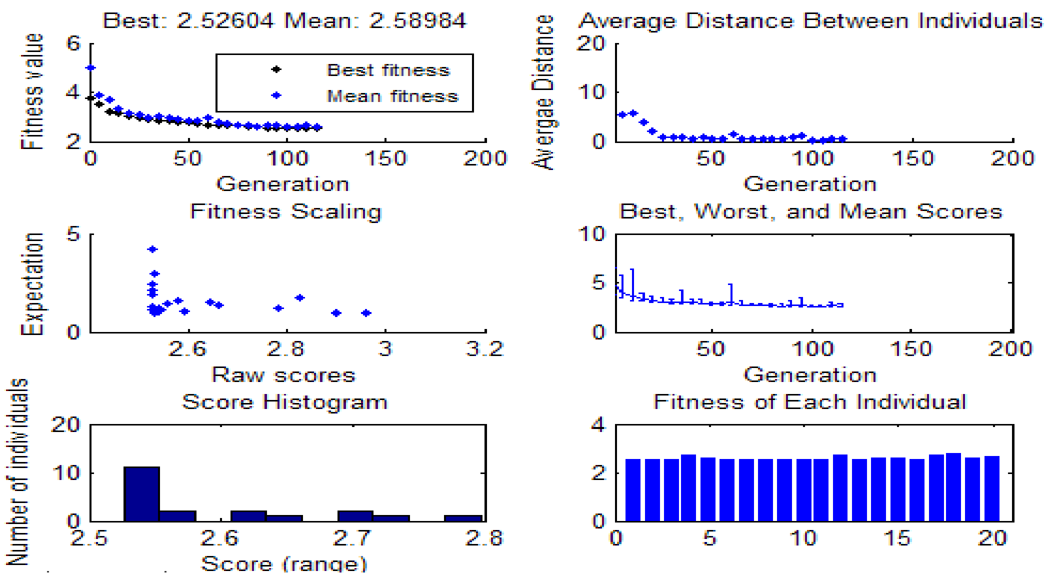
III. RESULTS AND DISCUSSION

We have implemented the proposed soft computing based models for facility location in logistics analysis using Mat lab software version R2012s. We used some of the functions of Logistics Engineering Toolbox “Mat log Version 16” in this research. We tried to use almost equal input parameters for each and every evaluated method. We used the USA census data and map data which is much suitable for this kind of research. We decided to use USA data because, it is the only data refereed in some of the previous works and there seems no such detailed data available for any other country for validating the methods of facility location and logistics analysis.

The Parameters of the Soft Computing Models

Genetic Algorithm options

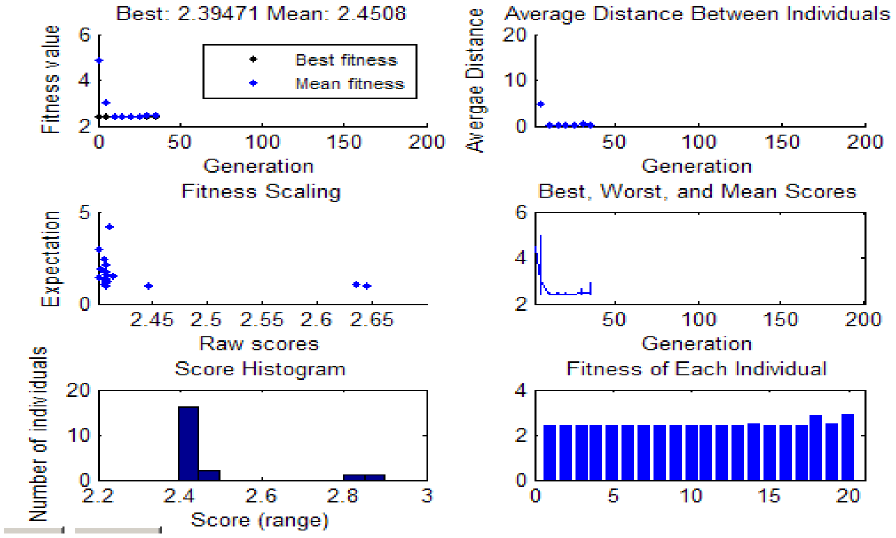
StallGenLimit	:	20
Generations	:	200
PopulationSize	:	20
CrossoverFcn	:	Two Point Crossover
MutationFcn	:	Gaussian k-means Clusterign
MaxRepetitions	:	10
MaxIter	:	20



The Time Taken for Finding Optimum Locations of DCs : 2.81 sec

The Average Distance between DCs and Customer Locations: 2.53Units

Figure : 7 - Genetic Algorithm Working Model



The Time Taken for Finding Optimum Locations of DCs : 1.57 sec

The Average Distance between DCs and Customer Locations: 2.39 Units

Figure 8: K-Means – GA Fusion Working Model

The following table shows the overall results of this work. Since the performance of a soft computing model will depend up on several factors, and some random conditions, we run each algorithm several times and only selected the values which are minimum.

Table 1: The Overall Performance for Locating 25 Facilities

Facility Location Method	Avg. Distance	Time Consumed
Random	6.16	-
GA	2.53	2.81
k-means-GA Fusion	2.39	1.57

The following graph shows the performance of the algorithms in terms of the minimum average distance achieved. The average distance is the average of distance between all DCs and the Customer locations. Each Customer is bound with a nearest DC and the distance between each customer to that corresponding DC is calculated and then the averages of all such distances were calculated. In this graph, the first column shows the reference distance which is nothing but the initial average distance of DCs that are randomly placed on the map.

With respect to the average distance, the k-means-GA fusion model performed well. It means, k-means-GA found the optimum facility locations better than the Normal GA methods.

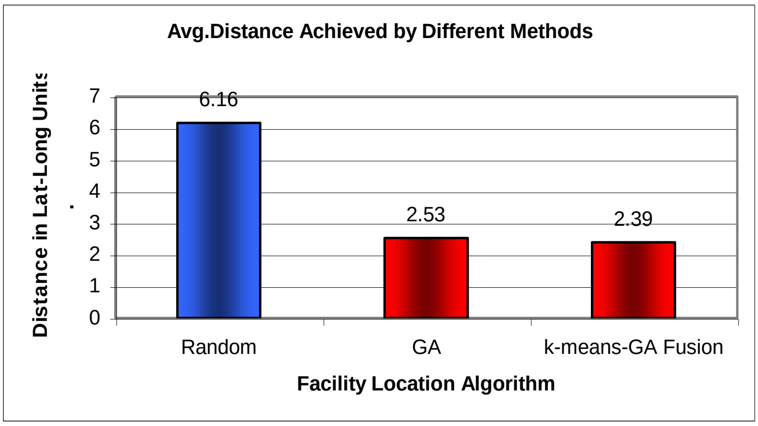


Figure 9. The Performance in Terms of distance

The following graph shows the performance in terms of cpu time. Unexpectedly, the k-means-GA model consumed less time. Even though the GA parameters were same, the fusion model consumed lesser time. Even there was no increase in time for the initial k-means step. Since the k-means step already found a good initial starting location, the search process of GA becomes very focussed so that, the individual steps of GA converge very fast. This may be the main reason for this low runtime of GA.

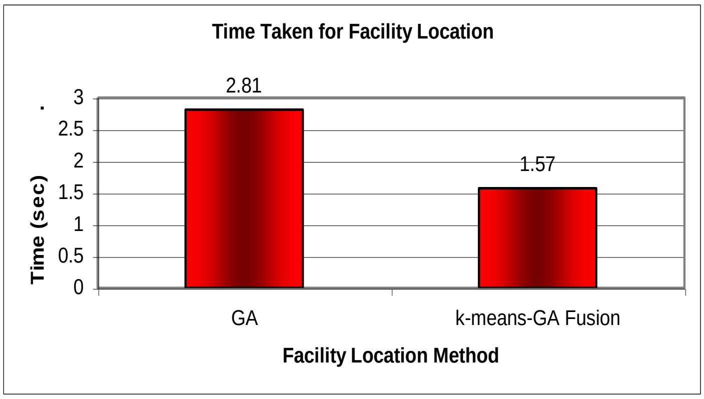


Figure 10. The Performance in Terms of CPU Time

IV. CONCLUSION

Facility location for logistics is a wide area for research. In this work we addressed the possibilities of using soft computing based models for facility location for logistics. We used soft computing based clustering approach for Facility Location Problem & Logistic Analysis.

As per the results, the hybrid, soft computing based optimization models successfully found optimum locations of facilities in considerably meaningful time limit.

In this work, we used a simple Euclidean distance function as a fitness function in the design of soft computing based location optimization model. Also the attributes like distance, road types, and city population are included in the model. But, there are much more constraints and parameters in a practical logistics problem that can be included in the design of the fitness function such as (1)loading unloading time at DCs, (2) different modes of travel times such as air travel time. This kind of more constraints and parameters can be included in future design of soft computing based optimization models. Our future works will address these issues.

We have designed the proposed models as minimum number objective problem. But there are facility location and logistics situations where there may be more objective during optimization. Future works may address the design of soft computing based optimization models for multi objective optimization scenarios.

REFERENCES

[1] Erlenkotter, D., "A Dual-Based Procedure for the Uncapacitated Facility Location." Operations Research, Vol 26(6), pg. 992-1009, 1978.

[2] Francis, R.L., McGinnis, F.L., Jr., White, J.A., "Facility Layout and Location: An Analytical Approach." Prentice Hall, 2nd Ed., 1974.

[3] Juliana Karakaneva, "A LOCATION PROBLEM MODELING AND SOLVING", Trakia Journal of Sciences, Vol 1, No 4, pp 1-7, 2003,ISSN 1312-1723, Copyright © 2003 Trakia University

[4] HK Smith, G Laporte and PR Harper, "Locational analysis: highlights of growth to maturity", http://eprints.soton.ac.uk/68930/1/Locational_Analysis_-_Smith,_Laporte_and_Harper.doc

[5] Michael J. Bucci, Michael G. Kay, Donald P. Warsing†, Jeffrey A. Joines, "Metaheuristics for Facility Location Problems with Economies of Scale", IIE Transactions

[6] Michael J. Bucci, Ryan Woolard, Jeffrey Joines, Kristin Thoney, Russell E. King, "An Application of Heuristics Incorporating Economies of Scale to Facility Location Problems in Carpet Recycling"

[7] Biehl, M., Prater, E., Realff, M.J., 2007, Assessing performance and uncertainty in developing carpet reverse logistics systems, Computers and Operations Research, 34, 443-463.

[8] Brimberg, J.,Hansen, P., Mladenovic, N.,Taillard, E.D. ,2000, Improvements and Comparison of Heuristics for Solving the Uncapacitated Multisource Weber Problem, Operations Research, 48, 444-460.

[9] Cooper, L., 1963, Location-allocation problems, Operations Research, 11, 331-343.

[10] CARE (Carpet America Recovery Effort) annual report, 2007, http://www.carpetrecovery.org/pdf/annual_report/07_CARE-annual-rpt.pdf

[11] CARE (Carpet America Recovery Effort) network website,2009, http://www.carpetrecovery.org/pdf/reclamation_centers/Carpet_Reclamation_ Center s.pdf

[12] Daskin, M.S., 1995, Network and discrete location: models, algorithms, and applications, John Wiley and Sons, New York.

[13] De Brito, M.P., Dekker, R., Flapper, S.D.P., 2003, Reverse Logistics - a review of case studies, ERIM Report Series.

[14] Fleishchmann, M., Krikke, H.R., Dekker, R., Flapper, S.D.P., 2000, A characterization of logistics networks for product recovery, Omega, 28, 653-666.

[15] Louwers, D., Kip, B.J., Peters, E., Souren, F., Flapper, S.D.P., 1999, A facility location allocation model for reusing carpet, Computers and Industrial Engineering,

[16] Mirchandani, P.B., Francis, R.L., 1990, Discrete Location Theory, Wiley, NewYork.

[17] Realff, M.J., Ammons, J.C., Newton, D., 1999, Carpet Recycling: Determining the 38:3, 547-567.

[18] Realff, M., Systems Planning for Carpet Recycling. 2006, Recycling in Textiles, Editor Youjiang Wang, CRC Press.

[19] Shaju Varghese, Gladston Raj S, "A Genetic Algorithm Based Optimization Model for Facility Location in Logistic Network Design", International Journal of Applied Engineering Research, Vol 10, No. 69, pp-338-344, 2005, ISSN 0973-4562.

[20] Shaju Varghese, Gladston Raj S, "Simulated Annealing and Direct Search Based Optimization Models for Facility Location in Logistic Network Design" International Journal of Computer Applications, Vol 132, pp- 31-37, December 2015, ISSN: 0975-8887.

[21] Shaju Varghese, Gladston Raj S, " Clustering Based Model For Facility Location In Logistic Network Using K-Means", International Journal of Scientific Inventions and Innovations, Volume 1 Issue 1, July 2016,

AUTHORS PROFILE

Mr. Shaju Varghese received his M.Sc. (Maths), M.C.A., and M.Phil. in computer Science. Now working as Head of the Department of Computer Applications at Baselios Poulose II Catholicos (B. P. C) College, Piravom, Kerala, India. He was the Principal Investigator of the Minor Research Project "Computerized Facility Location Analysis In Rural Area Using Clustering", 2010, funded by Universities Grant Commission, India. His research interest includes Data Mining, Facility Location Problem, and Cyber Criminology, and has presented papers in International and National Seminars. He has got three international journal publications. Currently he is pursuing Ph.D. in Computer Science at Bharathiar University, Tamilnadu, India.

Dr. Gladston Raj S. received his M.Sc (CS), M.Tech (Image Computing) and PhD in Computer Science from University of Kerala and Completed UGC-NET from University of Kerala and PGDCH (Computer hardware) from MicroCode, He is Now working as Head of the Department of Computer Science at Govt. College Nedumangad, Kerala, India. His area of interest includes Image Processing, Signal Processing, Data mining. He is providing research guidance for Ph.D scholars from different areas of research and has presented several invited talks in this areas of research.