Fuzzy Clustering in Grouping Traditional Market Distribution and Genetic Algorithm Application in Routing of Packed Cooking Oil Distribution

Teja Primawati Utami[a], Syamsul Ma'arif[b], Yandra Arkeman[c], Liesbetini Hartoto[d]

[a]Student, Bogor 16160, Indonesia

[b,c,d]Lecturer, Bogor 16160, Indonesia

[a]Email: teja.p.utami@gmail.com

[b]Email: syamsul4958@gmail.com

[c]Email : yandra_ipb@yahoo.com

[d]Email : liesbetinihartoto@yahoo.com

Abstract

This paper presents the modeling of intelligent routing of transportation of packaging cooking oil from the center to traditional market in the cluster in Indonesia, especially in Jakarta. Indonesia is the nation who has many islands. Every island has different population of people. Every day the public go to traditional market to buy main consumption products as palm cooking oil etc. The price of palm cooking oil at the market, sensitively will increase, especially when it becomes lack, by means sustainability of recent palm cooking oil stock at the market is very important. Focus of this research is to demonstrate how to optimize of routing distribution from distribution center to markets in the cluster. Optimum route expected can guarantee the availability of product and stock in the market to maintain the price. The clustering is created by fuzzy clustering and the routing is created by Transportation Salesperson Problem (TSP) with Genetic Algorithm (GA) method. GA is a method for solving optimization problem based on evolutionary theory in biology.

----

* Corresponding author.

E-mail address: teja.p.utami@gmail.com

1. Introduction

The movement of finishing product to customers is a market distribution. In market distribution, the end customer represents the final destination [1]. The ability to deliver goods as customer ordered is part of service. It should be called logistic that are integrated production and distribution. The logistics components of a corporation consist of: (1) a number of manufacturing plants, (2) zero, one, or more distribution echelons with distribution centres, (3) the customers, (4) the suppliers of components and raw materials, (5) recycling centres for used products and returned packaging containers, and finally (6) the transportation channels that link all of the above components [2].

Indonesia is the nation who has many islands. Every island has different population of people. Every region has many traditional markets to serve daily public consumption. These traditional markets need distribution center to assure availability of consumption product in these traditional markets. How to optimum determine distribution center of each region may defined by clustering. The author in [3] show that clustering process by fuzzy (fuzzy clustering) give the result better that defined it by firm directly approach. In this paper we focus on Jakarta region.

Jakarta is capital city of Indonesia that has 5 regions. These regions are North Jakarta, South Jakarta, West Jakarta, East Jakarta and Center Jakarta (Figure 1). Traditional market in Jakarta is coordinated by PD Pasar Jaya. PD Pasar Jaya has 153 traditional markets (figure 2). Every traditional market uncontrollable for availability of goods and disparity of goods prices. This paper purpose to make distribution center to solve this problem.
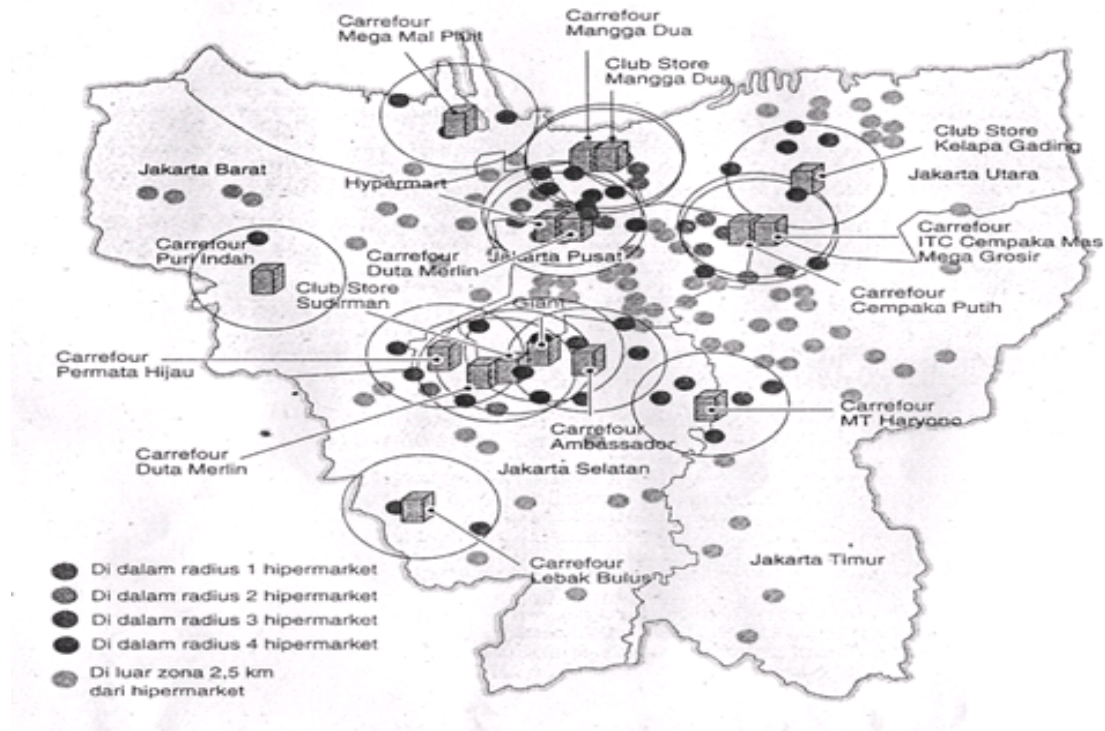

Figure 1: Region of Jakarta Map

Figure 2: Traditional markets at Jakarta

Distribution centre will be created by fuzzy clustering. Partition clustering essentially deals with the task of partitioning a set of entities into a several homogeneous clusters, about a suitable similarity measure. The main deference between the traditional hard clustering and fuzzy clustering can be stated as follow. While in hard clustering an entity belongs only to one cluster, in fuzzy clustering entities are allowed belong to many clusters with different degrees of membership.

Clustering has been around for many decades and located itself in a unique position as a fundamental conceptual and algorithmic landmark of data analysis. Almost since early of fuzzy sets, the role and potential of these information granules in revealing and describing structure in data was fully acknowledged and appreciated [4].

In the recent years clustering has undergone a substantial metamorphosis. From being an exclusively *data* driven pursuit, it has transformed itself into a vehicle which data centralized has been substantially augmented by the incorporation of domain knowledge thus giving rise to the next generation of *knowledge*-oriented and *collaborative* clustering. This fuzzy clustering is used.

After distribution centers are defined and need Travelling Salesperson Problem (TSP) to distribute the products to the markets. In the TSP, the goal is to find the shortest distance between *N* traveling points. The number of possible route for an *N* city tour requires *N*! additions. An exhaustive search through all possible paths is acceptable only when *N* is small. As *N* increases, several possible path grows geometrically. A 20-city tour involves 2.43 x 1018 additions. Even with 1 billion additions done in 1 second, this should take over 1852 years. Adding one more city would cause the number of additions to increase by a factor of 21. Obviously, exhaustive search becomes impractical.

To make it quicker and simpler, genetic algorithm is necessary used for saving the time purposes. Genetic Algorithm (GA) is a method for solving optimization problem that based on evolutionary theory in biology. This algorithm work with a population of candidate solutions named as chromosome that initially generated randomly from the area of the solution space of objective function. By using a mechanism of genetic operator i.e. crossover and mutation the population is evolved controlled by fitness function that directed to convergence condition [5].

This paper presents the application of GA approach in this cluster market of routing transportation problem called TSP. Although GA probably not led to the best solution, it could find a near optimal solution in a much less time (within several minutes).

2. Literature Review

2.1   *Fuzzy C-Mean (FCM)*

The author [6] noted that the fuzzy c-means (FCM) algorithm is one of the most widely used methods in fuzzy clustering. Data clustering is the process of dividing data elements into classes or clusters so that items in the same class are as similar as possible, and items in different classes are as dissimilar as possible. Depending on the nature of the data and the purpose for which clustering is being used, different measures of similarity may be used to place items into classes, where the similarity measure controls how the clusters are formed. Some examples of measures that can be used in clustering include distance, connectivity, and intensity.

There are two clustering known as hard and soft. In hard clustering, data is divided into distinct clusters, where each data element belongs to exactly one cluster. In fuzzy soft clustering, data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a process of assigning these membership levels, and then using them to assign data elements to one or more clusters [7].

The FCM algorithm attempts to partition a finite collection of n elements $X = \{x_1, ..., x_n\}$ into a collection of c fuzzy clusters with respect to somegiven criterion. Given a finite set of data, the algorithm returns a   list   of   c   cluster   centres $C = \{c_1, ..., c_c\}$ and   a   partition   matrix $U = u_{i,j} \in [0,1], \ i = 1,...,n, \ j = 1,...,c$, where each element u_{ij} tells the degree to which element x_i belongs to cluster c_j . Like the k-means algorithm, the FCM aims to minimize an objective function. The standard function is:

$$u_k(x) = \frac{1}{\sum_j \left(\frac{d(center_k,x)}{d(center_j,x)} \right)^{2/(m-1)}}.$$

Which differs from the k-means objective function by the addition of the membership values u_{ij} and the fuzzifier m. The fuzzifier m determines the level of cluster fuzziness. A large m results in smaller memberships u_{ij} and hence, fuzzier clusters. In the limit m = 1, the memberships u_{ij} converge to 0 or 1, that implies a crisp

partitioning. In the absence of experimentation or domain knowledge, m is commonly set to 2. The basic FCM Algorithm, given n data points (x1, . . ., xn) to be clustered, a number of c clusters with (c1, . . ., cc) the center of the clusters, and m the level of cluster fuzziness.

In fuzzy clustering, each point has a degree of belonging to clusters, as in fuzzy logic, rather than belonging completely to only one cluster. Thus, points on the edge of a cluster may be *in the cluster* to a lesser degree than points in the center of cluster. An overview and comparison of different fuzzy clustering algorithms is available.

Any point x has a set of coefficients giving the degree of being in the kth cluster $w_k(x)$. With fuzzy c-means, the centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster:

$$c_k = \frac{\sum_x w_k(x)x}{\sum_x w_k(x)}.$$

The degree of belonging, $w_k(x)$, is related inversely to the distance from x to the cluster center as calculated on the previous pass. It also depends on a parameter m that controls how much weight is given to the closest center. The fuzzy c-means algorithm is very similar to the k-means algorithm.

Choose a number of clusters. Assign randomly to each point coefficients for being in the clusters. Repeat until the algorithm has converged (that is, the coefficients' change between two iterations is no more than ε, the given sensitivity threshold).

Compute the centroid for each cluster, using the formula above. For each point, compute its coefficients of being in the clusters, using the formula above. The algorithm minimizes intra-cluster variance as well, but has the same problems as k-means; the minimum is a local minimum, and the results depend on the initial choice of weights.

The expectation-maximization algorithm is a more statistically formalized method which includes some of these ideas: partial membership in classes. Fuzzy c-means has been a very important tool for image processing in clustering objects in an image. In the 70's, mathematicians introduced the spatial term into the FCM algorithm to improve the accuracy of clustering under noise.

2.2. Cluster Analysis

The author [8] noted that cluster analysis is the task of assigning a set of objects into groups (called clusters) so that the objects in the same cluster are more similar to each other more than to those in other clusters. Clustering is a main task of exploratory data mining, and a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics.

Cluster analysis itself is not of specific algorithm, but the general task to be solved. It can be achieved by

various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. Popular notions of clusters include groups with low distances among the cluster members, dense areas of the data space, intervals or particular statistical distributions. Clustering can therefore be formulated as a multi-objective optimization problem. The appropriate clustering algorithm and parameter settings (including values such as the distance function to use, a density threshold or the number of expected clusters) depend on the individual data set and intended use of the results. Cluster analysis as such is not an automatic task, but an iterative process of knowledge discovery or interactive multi-objective optimization that involves trial and failure. It will often be necessary to modify preprocessed and parameters until the result achieves the desired properties.

2.3. Center of cluster

The author [9] noted that the clusters are represented by a central vector, which may not necessarily be a member of the data set. When the number of clusters is fixed to k, *k*-means clustering gives a formal definition as an optimization problem: find the k cluster centers and assign the objects to the nearest cluster center, such that the squared distances from the clusters are minimized.

The optimization problem itself is known as NP-hard, and thus the common approach is to search only for approximate solutions. A particularly well known approximate method is Lloyd's algorithm, often actually referred to as "*k-means algorithm*". It does however only find a local optimum, and is commonly run multiple times with different random initializations. Variations of k-means often include such optimizations as choosing the best of multiple runs, but also restricting the centroids to members of the data set (k-medoids), choosing medians (k-medians clustering), choosing the initial centers less randomly (K-means++) or allowing a fuzzy cluster assignment (Fuzzy c-means).

Most k-means-type algorithms require the number of clusters - k - to be specified in advance, which is considered to be one of the biggest drawbacks of these algorithms. Furthermore, the algorithms prefer clusters of approximately similar size, as they will always assign an object to the nearest centroid. This often leads to incorrectly cut borders in between of clusters (which is not surprising, as the algorithm optimized cluster centers, not cluster borders).

2.4. Travelling Salesman Problem (TSP)

The idea of the travelling salesman problem (TSP) is to find a tour of a given number of cities, visiting each city exactly once and returning to the starting city where the length of this tour is minimized. The first instance of the travelling salesman problem was from Euler in 1759 whose problem was to move a knight to every position on a chess board exactly once [10].

Travelling Salesperson Problem (TSP) is one of the issues combinatorial optimization, if there are a number of cities (or place) and the cost of travel from one city to other cities. Description of the problem is how to find the cheaper route of visit all the cities, each the city is only visited once, and must back to the original departure

city. The combination of all the existing route is the factorial of the number of cities. Travel cost can be a distance, time, fuel, convenience, and so forth.

2.5. Genetic Algorithm

Genetic algorithms are search techniques and optimization which is inspired by the principles of genetics and natural selection (Darwin's theory of evolution). This algorithm is used to obtain the exact solution for the optimization problem of a single variable or multi variable.

The author [11] noted that Genetic Algorithm is a general purpose guided random search based on the natural selection principles of biological evolution to improve the potential solutions. GA includes random elements which help to prevent the search begin trapped in local minimum. These properties overcome some of the short comings of conventional optimization approaches in ill-structured problems.

Being inherently parallel, GA are performed over a population of solution candidates. The manipulation process uses genetic operators to produce a new population of individuals (offspring) by manipulation the solution candidates. The algorithms start working by evaluating thousands of scenarios automatically until they find an optimal answer. The genetic algorithms bias the selection of chromosomes so that those with the better fitness functions tend to reproduce more often than those with worse evaluations.

Given an optimisation problem, GA first encodes the parameters into solution candidates. In the initial phase, the population consists of randomly enervated heterogeneous solution candidates. After all chromosomes go through evaluation process, an initial population will improve as parents are replaced by better and better children. The best individual in the final population can be a highly evolved solution to the problem. The author [12] noted that generally the genetic algorithm process consists of the following steps: Encoding, Evaluation, Crossover, Mutation and Decoding**.**

3. Methods

Fuzzy clustering is one method that can capture the uncertainty situation of real data and it is well known that fuzzy clustering can obtain a robust result as compare with conventional hard clustering [4]. Following the emphasis on the general problem of data analysis, that is a solution able to analyze a huge amount of complex data, the merit of fuzzy clustering is then presented.

After cluster was constructed, next step is to design routing from center of cluster to the members. The members and the cluster are traditional market in Jakarta, Indonesia. Routing is designed by Transportation Salesperson Problem and Genetic Algorithm is used to make optimization. The methodology framework is bellow on figure 3.

4. Discussion

There is a great interest in clustering techniques because of the vast amount of data generated in every field

including business, health sciences, engineering and aerospace. It is essential to extract useful information from the data. Clustering techniques are widely used in pattern recognition and related applications. This research monograph presents the clusters for traditional market in Jakarta, which these have each distribution centre.
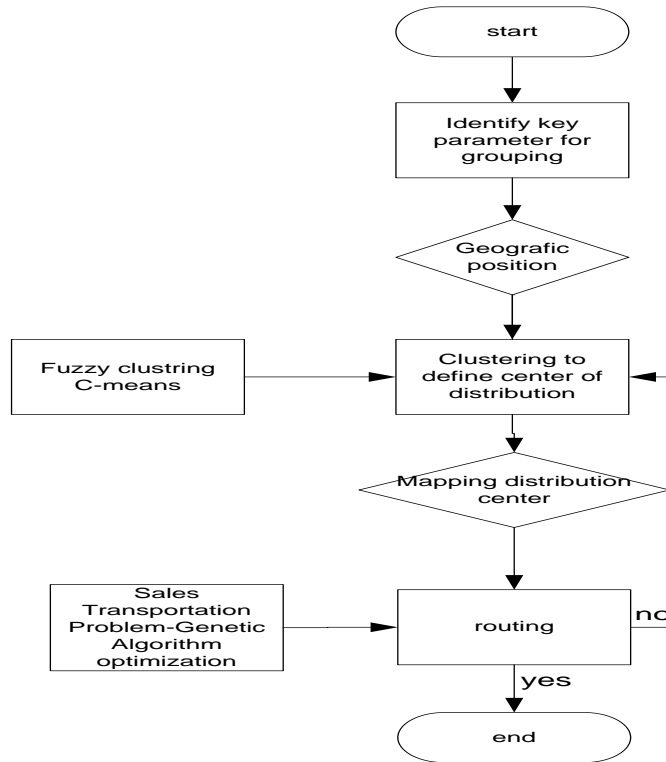
Figure 3: The Methodology Framework

4.1. Identify Parameter for Grouping

Clustering of traditional market in Jakarta is constructed by 4 parameters combining. These are latitude position, longitude position, and density of traders at the markets and accessibility of 153 transitional markets.

4.2. Clustering to define centre of distribution

Centres of distribution of traditional markets in Jakarta are defined by fuzzy clustering. we use MATHLAB to create the clustering. Fuzzy clustering with c-means is used for data analysis. The author [13] noted that the algorithm of fuzzy c-means (FCM) are below:

1.   Input data to be in the cluster is a matrix of n x m (n = number of data sample, m – attribute for each data).

X_{ij} = sample data to i (i = 1,2,…,n), attribute to-j (j=1,2,…,m).

Number of cluster (c)   = 15

Square (w)   = 2

Maximum iteration (maxIter)   = 100

Error (ξ)   = 10^{-5}

First objective function (P_0)   = 0

First iteration (t)   = 1

2.   Random number (μ_{ik}) generated,

i = 1,2,…,n; k = 1,2,…,c; with sequence below.

$$Q_j \sum_{k-1}^{c} \mu_{ik} \tag{1}$$

j = 1,2,…,m

which are,

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \tag{2}$$

3.   Center of cluster to-k ; V_{kj} with k = 1,2,…,c; and j = 1,2,…,m

$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^{n}(\mu_{ik})^w} \tag{3}$$

4.   Objective fungsion at iteration to-t, P_t :

$$P_t = \sum_{i-1}^{n} \sum_{k=1}^{c} \left(\left[\sum_{j-1}^{m} (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \tag{4}$$

5.   Partition matrix change

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(X_{ij}-V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c} \left[\sum_{j=1}^{m}(X_{ij}-V_{kj})^2 \right]^{\frac{-1}{w-1}}} \tag{5}$$

i = 1,2,…,n ; and k = 1,2,…,c

6.   Finish iteration

If : $(|P_t - P_{t-1}| < \varepsilon)$ or (t> $MaxIter$) so iteration is stoping.;   *(6)*

If not t = t+1, looping go to 3.

Clustering of numerical data forms the basis of many classifications and modeling algorithms system. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior.

Fuzzy Logic Toolbox tools allow finding clusters in input-output training data. It could use the cluster information to generate a Sugeno-type fuzzy inference system that best models the data behavior using a minimum number of rules. The rules partition themselves according to the fuzzy qualities associated with each of the data clusters.

Quasi-random two-dimensional data is used to illustrate how FCM clustering works. To load the data set and plot it, type the following commands:

load sheet1.dat

plot(sheet1 (:,1), sheet1 (:,2),'o')

Next, type the command-line function fcm to find two clusters in this data set until the objective function is no longer decreasing much at all. The syntax is

[center,U,objFcn] = fcm(sheet1sheet1,15);

Table 1: history of the objective function across the iterations

ITERATION	FCN
Iteration count = 1	obj. fcn = 0.041209
Iteration count = 2	obj. fcn = 0.030768
Iteration count = 3	obj. fcn = 0.030263
Iteration count = 4	obj. fcn = 0.029399
Iteration count = 5	obj. fcn = 0.027972
Iteration count = 6	obj. fcn = 0.026048
Iteration count = 7	obj. fcn = 0.024263
Iteration count = 8	obj. fcn = 0.022998
Iteration count = 9	obj. fcn = 0.021964
Iteration count = 10	obj. fcn = 0.021105
Iteration count = 11	obj. fcn = 0.020507
Iteration count = 12	obj. fcn = 0.020139
Iteration count = 13	obj. fcn = 0.019853
Iteration count = 14	obj. fcn = 0.019580
Iteration count = 15	obj. fcn = 0.019307
Iteration count = 16	obj. fcn = 0.019042
Iteration count = 17	obj. fcn = 0.018754
Iteration count = 18	obj. fcn = 0.018506
Iteration count = 19	obj. fcn = 0.018407
Iteration count = 20	obj. fcn = 0.018366
Iteration count = 21	obj. fcn = 0.018337
Iteration count = 22	obj. fcn = 0.018313
Iteration count = 23	obj. fcn = 0.018292
Iteration count = 24	obj. fcn = 0.018275
Iteration count = 25	obj. fcn = 0.018262
Iteration count = 26	obj. fcn = 0.018251
Iteration count = 27	obj. fcn = 0.018242

The fcm function is an iteration loop built on top of the following routines:

- initfcm — initializes the problem
- distfcm — performs Euclidean distance calculation
- stepfcm — performs one iteration of clustering

To view the progress of the clustering, plot the objective function by typing the following commands:

```
figure
plot(objFcn)
title('Objective Function Values')
xlabel('Iteration Count')
ylabel('Objective Function Value')
```
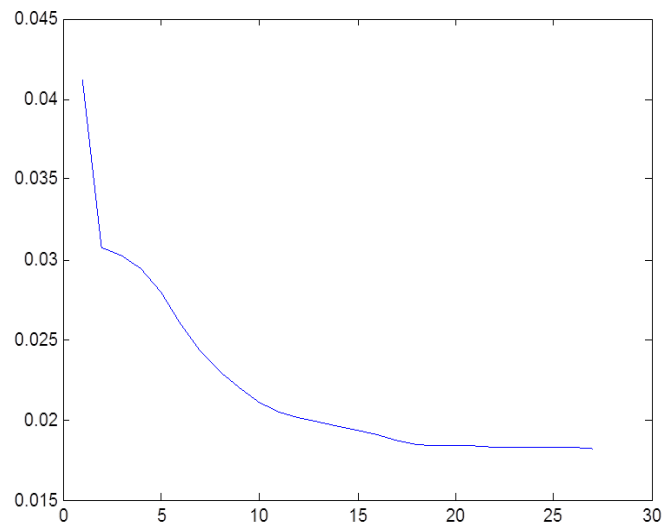
Figure of convergency is presented in figure 4.

Figure 4: Convergency Iteration

Finally, plot the fifteen cluster centers found by the fcm function using the following code:

```
maxU = max(U);
index1 = find(U(1, :) == maxU);
index2 = find(U(2, :) == maxU);
figure
line(fcmdata(index1, 1), fcmdata(index1, 2), 'linestyle',...
```

'none','marker', 'o','color','g');

line(fcmdata(index2,1),fcmdata(index2,2),'linestyle',...

'none','marker', 'x','color','r');

hold on

plot(center(1,1),center(1,2),'ko','markersize',15,'LineWidth',2)

plot(center(2,1),center(2,2),'kx','markersize',15,'LineWidth',2)

Coordinate geographic center of each cluster is presented below.

Table 2: coordinate geographic center of each cluster

NO	LATITUDE	LONGITUDE
1	6.166.857	1.065.258
2	6.117.264	1.065.397
3	6.097.971	1.065.185
4	6.113.407	1.064.485
5	6.132.045	1.065.233
6	608.422	1.064.826
7	6.198.298	1.065.389
8	6.177.212	1.065.032
9	6.112.826	106.504
10	614.756	1.064.774
11	6.090.052	1.064.985
12	6.111.806	1.064.832
13	6.075.793	1.065.701
14	6.151.379	1.065.046
15	6.085.015	1.064.342

4.3 Mapping distribution center by fuzzy clustering

These are 15 clusters traditional markets in Jakarta were defined. The centers of cluster are present with different color. One of them will be present in figure 5. These are cluster 1, 2 and 3. Members of the traditional markets in cluster 1 is: Tanjungpriok, Tanah tinggi, Sunterjaya, Sumurbatu, Senen, Rawasari, Kemayoran, Kartini, Kampungraya, Joharbaru, Gunung Sahari selatan, Galur, Cempaka baru, Cempaka putih barat and serdang. Distribution centers in cluster 1 is Senen traditional marketFurthermore, members of the traditional markets in cluster 2 is: Cipinang, Jati, Jatinegara Kaum, Kayu Putih, Klender, Malaka Jaya, Pondok kopi, Pulogadung, Pulogebang, Rawamangun and Utan Kayu Utara. Distribution centers in cluster 2 is Jatinegara the traditional market. While members of the traditional markets in cluster 3 are: Cawang, BatuAmpar, Cililitan,

Dukuh, Gedong, Kebon Pala, KramatJati, Makassar, Pinangranti, and Rambutan. Distribution centers in Cluster 3 is the traditional market Kramat Jati.
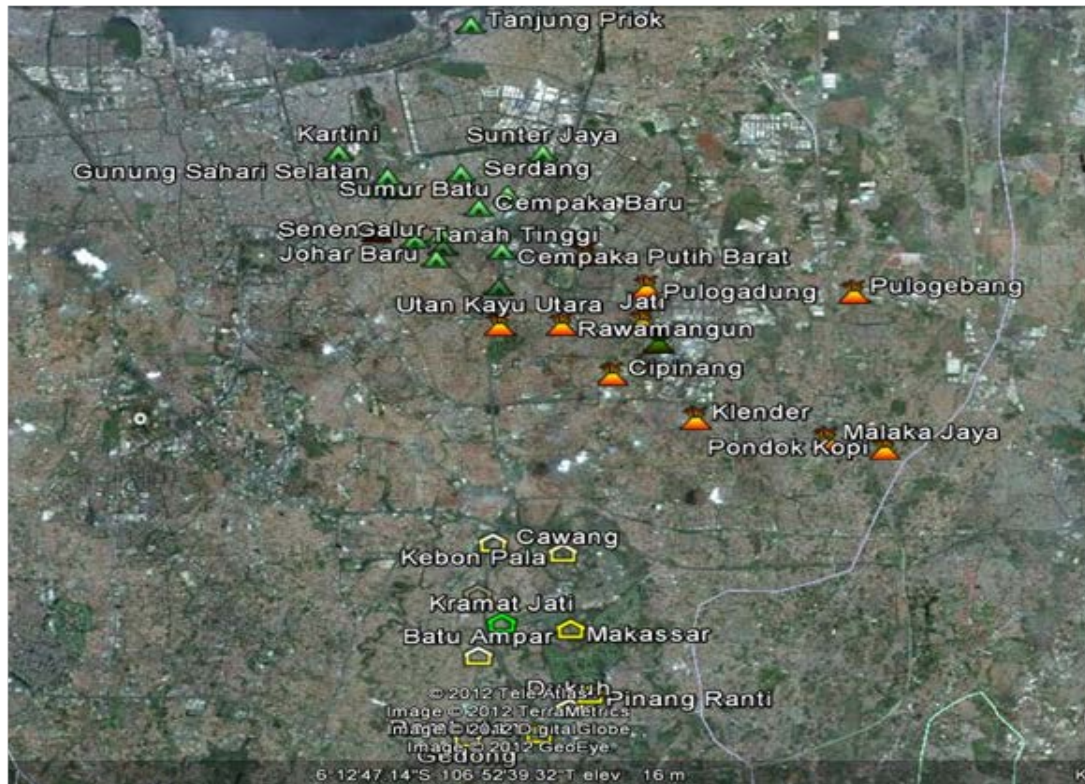

Figure 5: Cluster 1, 2 and 3 of Jakartas traditional markets clustering

4.4 Routing with Transportation Salesperson Problem-Genetic Algorithms.

The TSP is a standard problem in optimization. The objective in this paper is to minimize the travelling distance of *N* cities in a 10 km square radius from (0,0). Figure 3 for cluster 1 that color is yellow shows a 8-city tour starting from green dot (kramat jati) color. it is the centre to 7-others cities in the 10 km square radius, where the yellow dots are indicates the city needed to be traveled these are Cawang, BatuAmpar, Cililitan, Dukuh, Gedong, Kebon Pala, KramatJati, Makassar, Pinangranti, and Rambutan.

Figure 7 describes the flow of the optimization of TSP by GA. GA first encodes the travelling cities into chromosome. The population size is 1. After the chromosome goes through evaluation process, a fitness value is assigned to the chromosome. The child is then compared with the parent. If it is fitter than the parent, it will replace the parent, or it will not be used. Then the parent will reproduce a child through neighborhood mutation (which will be discussed in part v in this section). The process repeats until it reaches the maximum number of generations. The chromosome in the final population is a highly evolved solution to the problem.

4.4.1 Coding

In GA, the parameters to be optimized are encoded into chromosomes (Figure 9) and each chromosome is a

solution candidate. The encoding scheme depends on the nature of parameters to be optimized. In this problem, each city going to be visited is represented by an integer. The chromosome S, is a sequence of integers, can be formed by encoding the list of cities in the order they are visited. The length of chromosome equals to N.
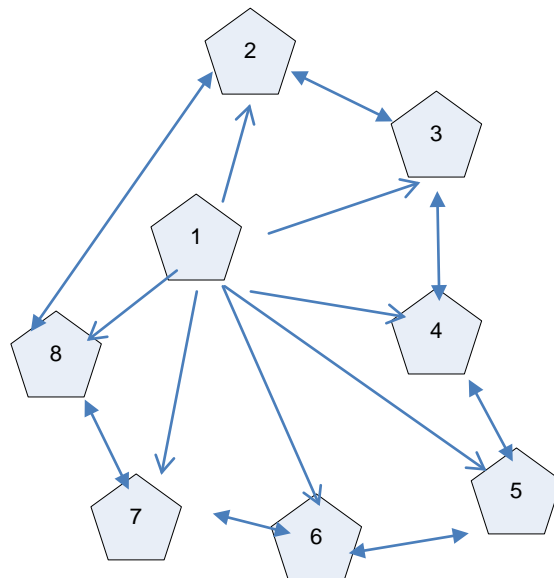

Figure 6: Cluster 3

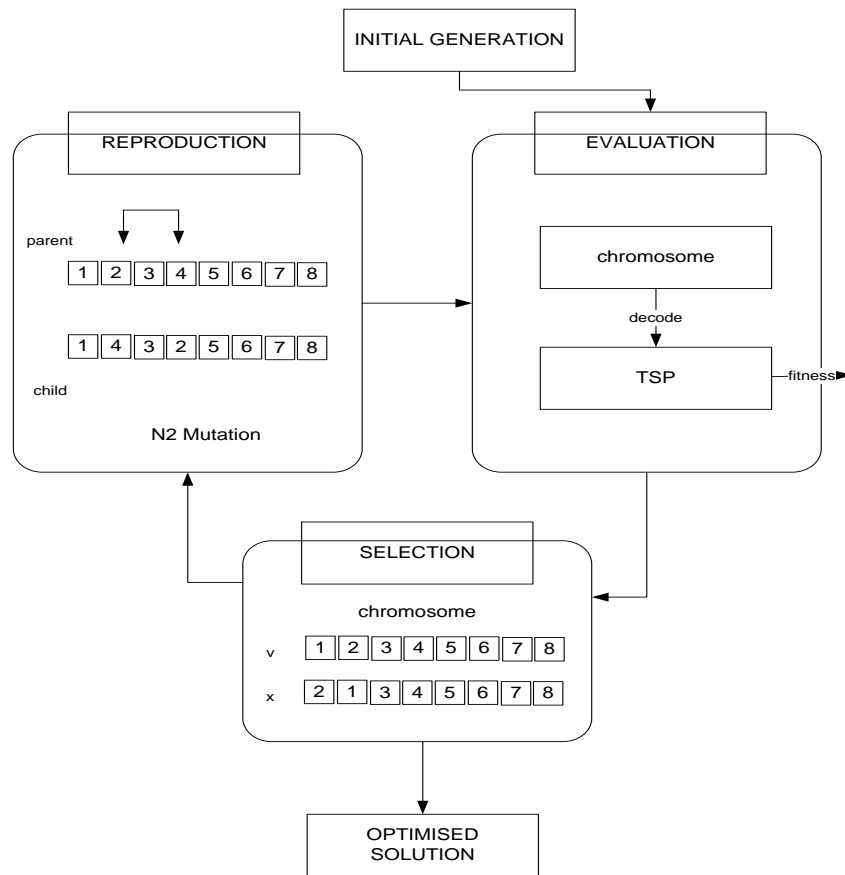**Figure 7:** Cluster one as a sample for routing TSP

Figure 8: Mechanisms of the proposed algorithms

4.4.2 Initialization

In this problem, we set the population size equals to 1 and the initial population is randomly generated.

4.4.3 Evaluation

In the evaluation module, each chromosome is coded with the integer of the cities to be travelled and the travelling time is calculated. The fitness value, calculated according to the fitness function, which is defined by the designer, is assigned to the chromosome.

4.4.4 Reproduction and Generation Selection

The reproduction module selects the alleles to be mutated. Then a new child chromosome is produced. The new chromosome is compared with parent chromosome. Elitism is used in the generation selection. If the new chromosome fitter than the parent, then it replaces the parent, else it will not be used. This avoids the loss of potential candidates by copying the best member into the succeeding generation.

4.4.5 Neighborhood Mutations

Conventional crossover and mutation are the most commonly used operations in GA to obtain offspring.

However, simple crossover and mutation may lead to violation of the constraint of TSP, as the city to be travelled may be missed or duplicated. As shown in Figure 9, the crossover operation will not work. Let's say, we have a 2nd crossover point. Every number in parent 1 before the crossover point is copied into the same position in child 1. Then, every number after the crossover point in parent 2 is put into child 1. The opposite is done for child 2.
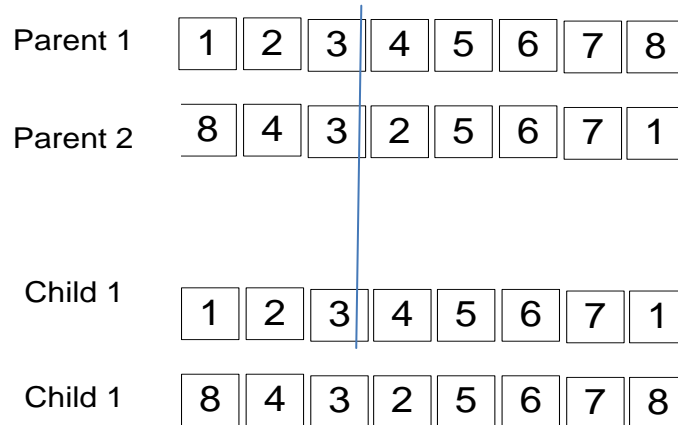
Parent 1 | 1 | 2 | 3 ‖ 4 | 5 | 6 | 7 | 8 |

Parent 2 | 8 | 4 | 3 ‖ 2 | 5 | 6 | 7 | 1 |

Child 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 |

Child 1 | 8 | 4 | 3 | 2 | 5 | 6 | 7 | 8 |

Figure 9: Crossover of chromosome

After the crossover operation, in Child 1, the city 1 is visited twice and city 8 is missed. The reproduction should preserve all the cities required in the chromosomes from the parents to the children. A different approach has therefore been adapted to the reproduction of chromosomes. The author [14] noted that a neighborhood is defined for the best chromosome in a generation and the chromosome only evolves to one of its neighbors.

The choice of chromosomes for the initial generation plays a vital role in the convergence toward the optimal solution. In order to smooth out this effect, 20tests have been carried out for each traffic condition with each neighborhood definition. In each test, the chromosomes of the initial generation are selected randomly from the set of possible sequences. The average of the minimum crossover the 20 tests is then calculated. All the simulation runs are performed on MATHLAB. Figure 10 summarizes the average travelling distance (of 20 tests) of 8 cities over 100 generations for different neighborhood mutations.

The following pseudo-code that is created for solve the above problems with the TSP using genetic algorithms

function Fitness (Kromosom[i])→integer
{calculate the fitness value of eachchromosome}

Declaration

Jum : integer
j : integer
Chromosome][] : array of integer of integer

Distance function (input A, B : integer) → integer

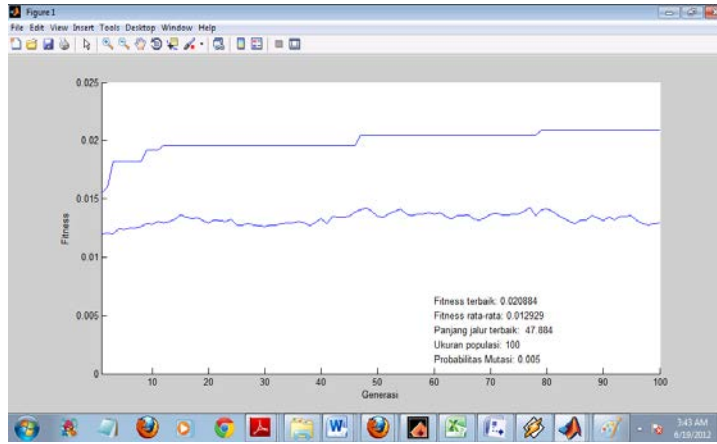*{generate the distance between two cities A and B }*

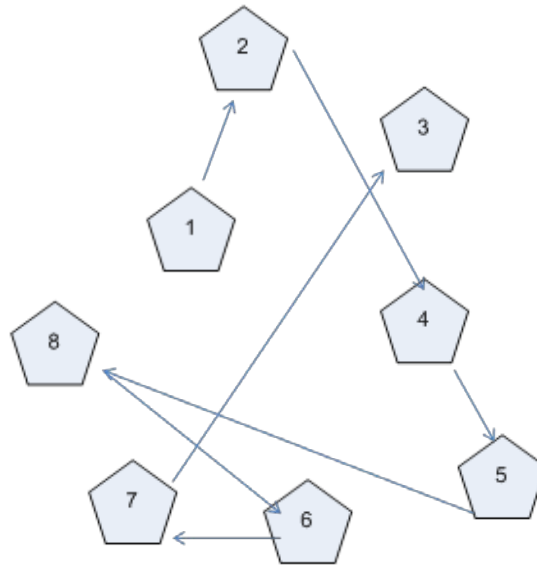**Figure 10:** The average travelling distance (of 20 tests) of 8 cities over 100 generations.

Figure 11: Sample optimum routing of 8 cities from distribution center (1) to seven other cities

Algorithm

Jum ←Jarak(A,Kromosom[i][1])

for j ←2 to 4 do

Jum →Jum +Distance (chromosome[i][j-1], chromosome [i][j])

endfor

Jum ← sum + Jarak(Kromosom[i][4],A)→ Jum

Crossover procedure (input populasi: integer, ρc:real)

{parent selection on the cross over }

Declaration

k : integer

R[] : array of integer

function random (input a-b :

integer) → nteger

{generates random numbers from a number to b }

Algorithm

k= 0

While k <= populasi do

R[k] ← random(0-1)

if R[k] < ρc then

pilih Kromosom[k][]sebagai induk

endif

k←k+1

endwhile

function of Number mutations (input JumGen,

JumlahKromosom: integer, ρm: real) →integer

{count the number ofthemutations }

Declaration

TotalGen : integer

JumMutasi : integer

Algorithm

TotalGen ←JumGen * JumlahKromosom

ρm ←0.2

JumMutasi ←0.2*TotalGen→JumMutasi

5. Conclusion

Traditional market existing is very important to help Peoples life in Jakarta. Every day they go to traditional market to buy many things for basic need consumption. Availability of goods and stability of price are important to consider. This paper gives a solution by present the distribution centers to facilitate all of traditional markets. Jakarta has 153 traditional markets. Traditional markets in Jakarta should be distributed to 15 clusters. Each cluster has one center, it could be distribution center. MATLAB is used fuzzy clustering to mapping clusters of

traditional markets in Jakarta. The iteration to convergence is 27 iterations. Every distribution center is nearly optimum to distribute the cooking palms oils to all traditional markets in the cluster.

Genetic algorithms appear to find good solutions for the travelling salesman problem, however it depends much on the way the problem is encoded and which crossover and mutation methods are used. It seems the methods that use heuristic information or encode the edges of the tour (such as the matrix representation and crossover) perform the best and give good indications for future work in this area.

Overall, genetic algorithms have proved suitable for solving the travelling salesperson problem. It seems that the biggest problem with the genetic algorithms devised for the travelling salesperson problem is the difficulty to maintain structure from the parent chromosomes and still end up with a legal tour in the child chromosomes. Perhaps a better crossover or mutation routine that retains structure from the parent chromosomes would give a better solution than we have already found for some travelling salesman problems.

References

[1] Bowersox, Closs, Cooper. Supply Chain Logistic Management. Mc Graw Hill. Michigan State University, New York, 2002, pp. 155-156

[2] Goetschalckx, M, Vidal, C.J and Dogan, K. Modeling and design of global logistics systems: A review of integrated strategic and tactical models and design algorithms. European Journal of Operational Research. Vol 43, pp. 1–18, 2002.

[3] Hamzah, A, Pengenalan Poladengan Fuzzy Clustering. ACADEMIA ISTA. Lembaga Penelitian, Institut Sainsdan Teknologi AKPRIND,Yogayakarta. Vol.4.No.1, pp : 47-58, 2001.

[4] Sato, M and Jain, L. C. Innovation in Fuzzy Clustering, Theory and Applications. Springer Berlin Heindelberg. New York, 2006, pp. 9-15.

[5] Widyastuti, N and Hamzah, A. Application of genetic Algorithm to Enhance the Performance of Clustering Fuzzy in Pattern Recognition**.** Jurnal Berkala MIPA, 17(2), pp. 36-50, 2007.

[6] Bezdek, J. C., Ehrlich, R., and Full, W. FCM: The Fuzzy C-Means Clustering Algorithm, Journal Computers & Geosciences, Vol. 10 No. 2-1, pp. 191-203, 1984.

[7] Chattopadyay, S., Pratihar, P. K., and De Sarbar, S. C. A Comparative Study of Fuzzy C-Means Algorithm And Entropy – Based Fuzzy Clustering Algorithms, Journal Computing and Informatics, Vol 30, pp. 701-720, 2011.

[8] Everitt, B. S., Landau, S., and Leese, M. Cluster Analysis. Fourth Editions, Arnold, 2001, pp. 6-11.

[9] Krotha, R. S and Merugula, S. A Brief Survey on Document Clustering Technique Using MATHLAB. International Journal of Computer & Organization Trends, Vol 3 Issue 1, pp. 1-6, 2013.

[10] Michalewicz, Z. Genetic Algorithms + Data Structures = Evolution Programs Traveling Salesman Problem Using Genetic Algorithms. Springer, Verlag, 2nd edition 1994, pp : 112-125.

[11] Chan, P.T. and Rad, A.B. and Ho, T.K. *Optimisation of a class oftransportation problems using genetic algorithms.* In: Traffic And Transportation Studies. Proceedings of the Third International Conference on Traffic and Transportation Studies, Guilin, China, 23-25 July 2002, pp: 276-289.

[12] Bryant, K, Benjamin A. Advisor Genetic Algorithms and the Traveling Salesman Problem. Departement of Mathematic. Harvey Mudd College, 2000, pp. 11-18.

[13] Kusumadewi. Aplikasi logika Fuzzy untuk Pendukung Keputusan (ID): Graha Ilmu. 2004.pp: 34-40

[14] Gerard R. The Traveling Salesman: Computational Solutions for TSP Applications.. Springer-Verlag,. 1994, pp. 97-102.