J. Appl. Res. Ind. Eng. Vol. 6, No. 2 (2019) 108–124

JARIE Journal of Applied Research on Industrial Engineering

Journal of Applied Research on Industrial Engineering
www.journal-aprie.com

The Extension Analysis of Natural Gas Network Location-Routing Design through the Feasibility Study

Filscha Nurprihatin*[*]*, Agnes Octa, Tasya Regina, Tony Wijaya, Julliete Luin, Hendy Tannady

Department of Industrial Engineering, Faculty of Design and Technology, Universitas Bunda Mulia, Jakarta, Indonesia.

PAPER INFO

Chronicle:
Received: 06 January 2019
Revised: 18 May 2019
Accepted: 02 June 2019

Keywords:
City Gas Network.
K-Means Clustering.
Heuristics Algorithm.
Feasibility Study.
Benefit-Cost Ratio.

ABSTRACT

From 2001 to the present, natural gas production in Indonesia dominates petroleum production. Most of the natural gas is used for export until 2013. From 2014 until now, most of the natural gas production is being utilized for domestic with an increasing trend. Domestic gas usage for households is still far below industry and commercial sectors. Domestic gas usage in the household can be done in two ways, namely city gas and Liquefied Petroleum Gas cylinder. The use of city gas is better in terms of price, mitigation, and gas emission. The government plans to build a new city gas network for 4,000 households. This study aim is to propose the design of city gas network, so the construction and operational costs become minimal. This research uses three stages, namely division of region by using the clustering algorithm, the gas network route determination using heuristics algorithm, and determination of the feasibility using the Benefit-Cost Ratio. We have successfully calculated the maximum region's capacity by making use of Weymouth Formula. The iterative clustering algorithm is done to make sure the location of Distribution Centers is well-defined. We modified the distance measurement by preferring driving distance rather than Euclidean in the interest of precision. In the end, we also discuss the feasibility study of the project. Based on the calculation, we have obtained that the gas network development project is feasible to run.

1. Introduction

The exploration, production, and utilization of petroleum (oil and gas) in Indonesia has been more than 100 years, meaning the industry is well-worn [1]. The development of oil and gas production in the country is described in Fig. 1. Fig. 1 tells that since 1966, oil production has always been above natural gas although it tends to decline since 1995. The year 2001 marked the dominance of natural gas production to date due to the tiring oil's exploration area. In 2014, natural gas production was about 8,147 million standard cubic feet per day [1].

The government is expected to increase the percentage of domestic usage along with the minister of energy and mineral resources regulation No. 3/2010 on the allocation and utilization of natural gas for domestic needs. Fig. 2 shows the fulfillment of domestic and foreign demand. Fig. 2 shows that in 2013

[*] Corresponding author
E-mail address: fnurprihatin@bundamulia.ac.id
10.22105/jarie.2019.174164.1082

53% of natural gas is used for domestic needs and increase to 57% by 2014. In the end of 2019, it will reach 64% [1].

The utilization of domestic gas can be divided into industrial sectors, commercial, household, transportation, and other sectors, as shown in Fig. 3. From 2013 to 2015 the usage of domestic gas had been dominated by other sectors, industries, and commercials. Meanwhile, household and transportation sectors occupy the last two sequences. This shows that domestic gas usage for households is still far below industrial and commercial sectors [2].

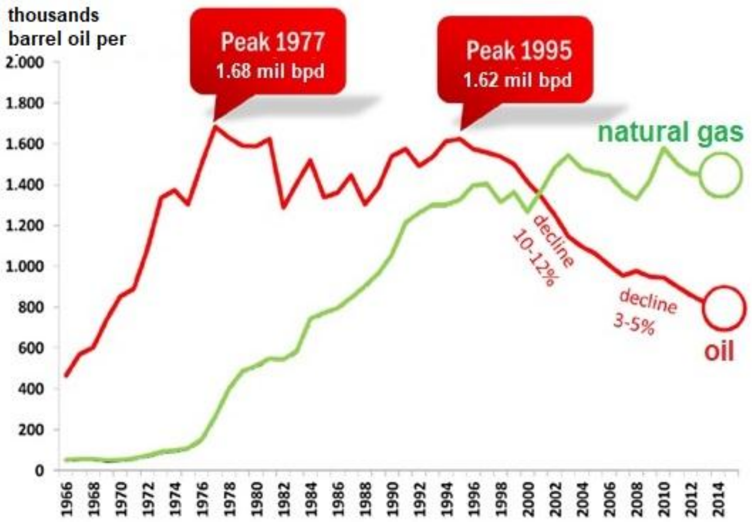

Fig. 1. *Development of oil and gas production [1].*

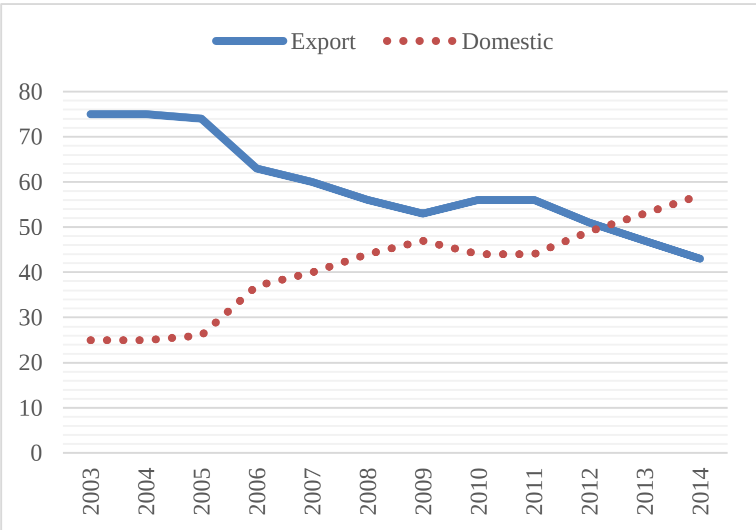

Fig. 2. *The fulfillment of domestic and foreign demand (%) [1].*

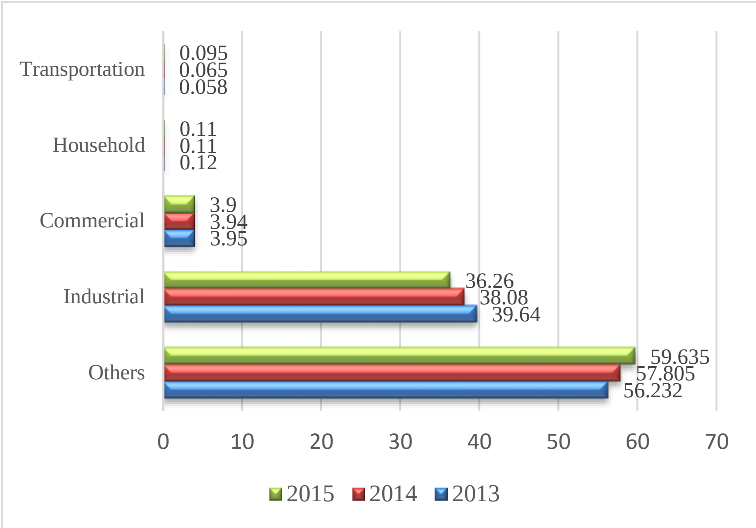

Fig. 3. *The utilization of natural gas in various sectors (%) [2].*

Domestic gas usage in the household can be utilized in two ways: city gas and Liquefied Petroleum Gas (LPG) cylinder. The difference between city gas and LPG cylinder lies with the main ingredient of gas ingredients and packaging. The main ingredients of city gas are methane and ethane, whereas in the LPG cylinder are propane and butane. In terms of packaging, LPG is more practical to use because it is packed in the form of gas cylinders, while the city gas can only be channeled directly using pipelines similar to water supply network. However, the city gas is relatively cheaper and greener compared to LPG cylinder for power generation, manufacturers, and others [3-5]. City gas network does not need to charge the transportation cost due to the natural gas can be channeled directly to households using pipelines. In addition, it is easy to handle and overcome gas leakage under city gas rather than LPG cylinders. Therefore, the usage of city gas is better than the LPG cylinder in terms of price and handling the leakage. Currently, Indonesian government wishes to increase the use of city gas in the household sector by making network development programs in some major cities.

The construction of the city gas network in Indonesia began in 2008 on the qualification documents stage such as Front-End Engineering Design (FEED) and Detail Engineering Design for Construction (DEDC). In 2009, city gas network construction was built in the city of Palembang and Surabaya. In 2009-2010, as many as 19,377 households have been installed gas network and continue to grow so that in the end of 2014, a total of 89,460 households have been able to enjoy the city gas network throughout the country [1]. By the end of 2019, this number is targeted at 1.23 million households [1].

The government plans to build new city gas network to continue the enormous development plan which physical construction was initiated in 2010. This new pipeline network will be installed for 4,000 households. A regional state-owned enterprise will handle the project and needs to minimize the length of the gas pipeline. In order to achieve it, this paper considers the location of the gas storage, the Distribution Center (DC) to accommodate gas from the storage, the distance from one node to another, and the number of households which will be supplied. All of these factors will form a supply chain.

In this problem, the gas distribution network is divided into several regions. Each region is represented by one DC and should not exceed the capacity of the number of households which will be supplied. From here, gas pipeline branches are created and flows directly to each household. The regional grouping is used to restrict each region capacity by utilizing K-Means clustering.

K-Means clustering has been utilized to overcome problems in wide areas such as medical and engineering. In medical field, the research has been conducted by making use of K-Means clustering along with fuzzy C-mean clustering to segregate brain tumor of normal and pathological tissues [6]. Another research successfully segmented automatically the lumen and vessel wall in Magnetic Resonance (MR) images of carotid arteries, a crucial step towards the evaluation of cerebrovascular diseases [7]. Other study tried to classify the Corpus Callosum (CC), an important brain structure and its volume and variations in shape which correlated with diseases like Alzheimer, schizophrenia, dyslexia, epilepsy, and multiple sclerosis [8]. Previous research is conducted in engineering field utilizing K-Means clustering to determine the location of the region with the highest percentage of water pollution in several cities in China based on the clustering of water pollution points [9]. In this research, we make use K-Means clustering to prevent the exceeding capacity from each region iteratively. If there is the exceeding capacity, we go to another round of iteration by adding the number of DCs. Unlike another research, we predefine the capacity according to Weymouth formula since the problem is related to gas flow rate.

Our aim is to determine the location of DCs as the representative of each region through K-Means clustering. Greater population of a region is the greater demand [10]. Our work selects the number and locations of DCs based on population at each node. Since the population in the region is greater than the number of households, so we make the logical proportion of it.

Traveling Salesman Problem (TSP) is usually used to solve the problem of determining the distribution line with Hamiltonian cycle in order to obtain the shortest distribution path and lowest construction cost. The TSP approach can be used to simulate logistics, job planning, vehicle scheduling, and so on [11]. This study aims to determine the city gas pipeline network using TSP approach. In minimizing the length of the city gas pipeline network, our work requires stages to identify and determine the distribution network with the algorithm. The algorithm is a set of steps or procedures that can be used to solve the problem that is arranged clearly and logically. Heuristics algorithm has been used to minimize the travel distance in third-party warehouse while order-picking under batch system [12]. Another study utilized the heuristics algorithm in distributing finished goods from the warehouse to the final customer [13] under Multiple-Depot Vehicle Routing Problem (MDVRP) with stochastic travel times [14]. Open Vehicle Routing Problem (OVRP) has been solved by modifying Clarke and Wright´s Savings Algorithm (SA) [15] to schedule office bus, and allocate bus stops that will use to pick up and drop employees according to their residence [16]. Sequential SA has been developed to improve profit by 116.84% for a freight transportation service provider in Thailand [17], while this paper utilizes SA to shorten the length of city gas network.

The designed city gas distribution network based on the TSP generates total construction costs. Therefore, we conduct a feasibility analysis of the project using Benefit Cost Ratio (BCR) method. The previous research utilized BCR to analyze the feasibility of a 5-year biannual Foot and Mouth Disease (FMD) vaccination program to avoid losses for farmers and to contribute to a decrease in the poverty rate [18]. In addition, other earlier study used BCR to analyze the cost and benefit perspectives gained by waging war [19]. The cost and benefit perspective is derived from the death toll due to war and the economic movement has an impact on the decision to start the war [19]. In this paper, we analyze benefits and costs based on natural gas selling prices and construction costs.

K-Means clustering and SA utilization in this study is aimed to determine DC and city gas network while minimizes the length of city gas pipeline. By minimizing the length of the pipeline, the operational costs of the gas pipelines network are assumed to be minimal as well. This study supports the government's program to increase gas usage by 18% in 2013 to 22% in 2025, and 25% in 2050 [1]. This study also contributes to provide a minimum length of city gas network for 4,000 households of the 306,000 households target by 2018 throughout the country [1]. The final goal of this research is to know the feasibility of development project of city gas network based on the cost-benefit ratio value.

2. Literature Review

2.1. Supply Chain and Logistics

Logistics is a dynamic and diverse function related to raw material and distribution management that must be able to adapt changes in demand and other limitations [20]. Furthermore, the scope of the supply chain is an integration between suppliers, logistics, and customers that not only focus on material movement but also information [20]. The main factors affecting supply chain performance consist of [21]:

- Location of facility.
- Inventory.
- Transportation.
- Information.
- Procurement of goods.
- Pricing.

2.2. Transportation

Transportation addresses the issue of distributing a commodity or product from a number of sources to a number of destinations to minimize freight costs [21]. The specific characteristics of transportation issues are as follows:

- There are a number of specific sources and objectives.
- The quantity of commodities distributed from each source and requested by each destination, the magnitude varies.
- Commodities sent or transported from a source to a destination, the magnitude corresponding to the demand, and capacity of the source.
- The cost of transporting a commodity from a source to a destination of magnitude varies.

There are 7 (seven) modes of transportation that can be used in the supply chain, such as [21]:

- Air,
- Package carriers,
- Truck,
- Rail,
- Water,
- Pipeline,
- Intermodal.

2.3. Pipeline

Pipeline as a transportation mode is limited by service coverage and the capability which can only be utilized for liquid products and gas [22]. In addition, the utilization of pipelines as a transportation mode requires substantial capital costs, such as pipeline construction costs, and related infrastructure. Transportation uses pipeline is suitable for large and stable demand characteristics that gives utilities 80 to 90 percent [21]. Pipelines are suitable for the long-term usage due to the costs incurred consist of development costs and maintenance costs.

2.4. Weymouth Formula

The Weymouth equation is used for high pressure, high flow rate, and large-diameter gas collection system. The formula is shown in Eq. (1) to calculate the gas flow rate through a pipeline based on the gravity value, compressibility, inlet and outlet pressure, pipeline diameter, and pipeline length [23].

$$Q = 433.5E\left(\frac{T_b}{P_b}\right)\left(\frac{P_1^2 - e^s P_2^2}{GT_f L_e Z}\right)^{0.5} D^{2.667}, \quad (1)$$

where:
Q = Flow rate volume, standard ft^3/day (scfd).
E = Pipe efficiency, decimal value less than or equal to 1.
P_b = Initial pressure, psia.
T_b = Initial temperature, °R (460 + °F).
P_1 = Upstream pressure, psia.
P_2 = Downstream pressure, psia.
G = Specificity gravity.
T_f = Average temperature of gas flow, °R (460 + °F).
L_e = Equivalent pipe length, mi.
Z = Gas compressibility factor.
D = Inner diameter of pipe, inches.

2.5. K-Means Clustering

Clustering is the process of grouping data items into a small number of groups so that each group has something of an essential equation [24]. There are two types of clustering algorithms used in DM, namely hierarchical clustering and non-hierarchical clustering [25].

Hierarchical clustering algorithms may be conducted either agglomerative (bottom-up) or divisive (top-down) [26]. Agglomerative algorithms start with each object being a separate cluster itself and successively merge groups according to some criterion. Divisive algorithms follow the opposite strategy. The cluster will form a kind of tree where there is a clear hierarchy between objects, from the most similar to the least similar. Logically, all objects will ultimately form a group.

The non-hierarchical clustering algorithm also known as Representative-Based Algorithm [24] begins first by determining the desired number of groups. Once the number of groups is known, then the grouping process is done without following the hierarchy process. K-Means clustering is a grouping technique for large data with similar characteristics and can be used after the number of groups is determined [7]. Non-hierarchical procedures do not involve a hierarchical construction process, but need to select a group center, and all objects within a certain range are included in the resulting group [25]. K-Means clustering algorithm is based on the Euclidean distance that divides data into a number of clusters and this algorithm works only on numerical attributes [24]. The steps of clustering with K-Means clustering algorithm are as follows [27, 28]:

- Select the number of clusters *k*.
- Many initializing techniques have been proposed, from simple methods, such as choosing the first k data points, forgy initialization (randomly choosing *k* data points in the dataset), and random partitions (dividing the data points randomly into *k* subsets), to more sophisticated methods, such as density-based initialization, intelligent initialization, furthest-first initialization (FF for short, it works by picking first center point randomly and then adding more center points which are furthest from existing ones), and subset furthest-first (SFF) initialization.
- Initialization of this cluster center k can be done in various ways. But the most often done is by random. Cluster centers are scored with random numbers.
- Allocate all data/objects to the nearest cluster. The proximity of two objects is determined by the distance of the two objects. Likewise, the proximity of a data to a particular cluster is determined by the distance between the data and the cluster center. In this stage, it is necessary to calculate the distance of each node to each cluster center. The distance between one data and one particular cluster will determine the cluster membership. To calculate the distance of all data to each cluster central point can use Euclidean distance formulated in Eq. (2). We improve the accuracy in calculating the distance of all data to each cluster central point; Google Maps application calculates distance based on driving distance, not by Euclidean distance.

$$d_{ij} = \sqrt{\sum_{k=1}^{n} \left(x_{ki} - x_{kj}\right)^2}, \quad (2)$$

where:

d_{ij} = Distance of data to *i* to center cluster *j*.

X_{ki} = Data to *i* on attribute data to *k*.

X_{kj} = Center point to *j* at attribute to *k*.

- Recalculate the cluster center with the current cluster membership. The cluster center is the average of all data/objects in a particular cluster. If desired, it can also use the median of the cluster. So, the mean is not the only size that can be used.
- Reassign each object using the new cluster center. If the cluster center does not change again then the clustering process is complete. Alternatively, return to step number 3 until the center of the cluster does not change anymore.

2.6. Traveling Salesman Problem (TSP)

The concept of Traveling Salesman Problem (TSP) is the determination of the salesman's travel route from a starting point to another point until all points are visited and ends to the starting point again, given that each point can only be passed once. It is used to find travel routes from selected cities (visit each city exactly once) where travel distance is minimized. TSP is defined as the task of finding the shortest cycle or path in the entire graph of nodes. This is a classic example of the NP-hard problem because of the list of cities and the distance from each city pair; the task of the TSP model is to find the shortest route possible that visits each city only once. Thus, the method of finding the optimal solution involves searching in distance solutions that grow or increase exponentially with the number of cities [29].

The TSP mathematical model is shown in the following equations [29]:

Objective function:

$$\text{Minimize:} \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}.x_{ij} \ , i \neq j \tag{3}$$

Subject to:

$$\sum_{j=1; j\neq i}^{n} x_{ij} = 1, \quad i = 1,2,\dots,n \tag{4}$$

$$\sum_{i=1; i\neq j}^{n} x_{ij} = 1, \quad j = 1,2,\dots,n \tag{5}$$

$$x_{ij} \in \{0,1\}; i,j = 1,2,\dots,n\ . \tag{6}$$

In the model above, Eq. (3) is aim to minimize the total distance or total travel costs. The constraint (4) guarantees the journey of a point only once, while the constraint (5) indicates that a point is visited only once. Meanwhile, the constraint (6) specifies that x_{ij} is a member of a binary number, 0 if point *j* is not visited from point *i*, and is 1 if point *j* is visited by point *i*.

2.7. Heuristics Algorithm

Determination of schedules and delivery routes from a location to multiple destinations is one of the most important operational decisions in the management of dissemination. The objective in the problem of determining the vehicle route is to minimize transportation costs approached by route length [30]. One of the heuristic methods in solving the problem of route determination and distribution schedule is by using the SA developed by Clarke and Wright in 1964 [30]. Although the SA does not need to provide optimum quality solutions because the optimality may be impossible, but the solutions provided are good [31].

SA calculation steps are as follows [13, 16, 17]:

– Determining the distance matrix

  In this step, we need to know the distance between the DCs and each node and the distance between nodes. The distance is obtained using the Google Maps app.

– Determining the Savings Matrix

  Savings matrix represents savings that can be realized by merging of two or more nodes in a single route and one vehicle. The savings can be distance and time, or cost. If each node x and y node are visited separately then the distance traveled is the distance from the DC to node x and back to the DC plus the distance from the DC to node y and then back to the DC.

  S(x,y) is denoted as a savings of distance when a merge of visits into one route is from the DC to node x and node x to node y then back to the DC. The savings formula is shown in Eq. (7).

$$S(x, y) = J(G, x) + J(G, y) - J(x, y). \quad (7)$$

- Allocating nodes to the route

The purpose of determining the vehicle route to the point is to maximize savings. Savings can be maximized by doing multiple iterations. The first step of this iteration procedure is the merger of two routes with the highest savings being one viable route. This procedure is done continuously until no more suitable combinations are found.

- Sorting destination point in a predefined route

There are three methods for sorting points in a known route from the iteration process, including:

- Farthest insert

This method is accomplished by the addition of nodes in a travel route, started with those with the greatest or furthest distance increase. This procedure will continue until all nodes enter the route.

- Nearest insert

This method is the opposite of the farthest insert method where this procedure starts from determining the route of the vehicle to the node that has the closest distance. Then this procedure will continue to repeat until all nodes enter the route of travel.

- Nearest neighborhood

The procedure of sorting the visit of the consumer with the approach of nearest neighbors starts from the DC then makes the addition of the node closest to the DC. At each stage, the existing route is built by adding node closer to the last visited node.

3. Research Methodology

This study begins by calculating the gas flow rate through a pipe using Weymouth Formula, then adjusting it as a fluid using the Ideal Gas Law. K-Means clustering was utilized to determine the maximum proportion of family heads and the maximum number of sectors (clusters). After that, we determined the Distribution Center (DC) from several locations with the largest proportion, then each location with the largest proportion is assumed to be a cluster. In determining members on each cluster, it is based on the smallest distance between each location with predetermined DC points. A review of each DC point is carried out using the Center of Gravity (COG) method so that each DC point is in the middle of the cluster. The COG method is repeated until there are no clusters that exceed proportion capacity and no cluster members move.

The route of the gas pipeline network is inefficiencies using Clarke and Wright's Savings Algorithm is done by two methods, namely Nearest Neighbor and Savings Matrix. After obtaining the route and total distance, a distance is observed on each cluster route, so that the method with the smallest distance is chosen to be realized. Finally, a calculation of Benefit Cost Ratio (BCR) is carried out to see the feasibility of the city's gas pipeline project.

4. Result and Discussion

4.1. Capacity of Each Region

The capacity of each region is determined based on the gas flow equation, using Weymouth Formula. The gas flow rate can be calculated using Eq. (8).

$$Q = 433.5E\left(\frac{T_b}{P_b}\right)\left(\frac{P_1^2 - e^S P_2^2}{GT_f L_e Z}\right)^{0.5} D^{2.667}\ . \quad (8)$$

The following is the calculation of each region's capacity:

$$Q = \frac{433.5(0.95)\left(\frac{519.67}{14.6959}\right)\left(\frac{16.1463^2 - 15.4211^2}{(0.8)(539.676)(1.5534)(0.9999)}\right)^{0.5}(2.48)^{2.667}}{1000}.$$

$$Q = 30.33\ mscfd.$$

$$Capacity\ per\ region = 30.33\ x\ 28.3128\ .$$

$$Capacity\ per\ region = 858.7\ sm^3/day\ .$$

$$Capacity\ per\ region = \frac{858.7}{\frac{(0.10\ x\ 14.5038 + 14.6959)}{14.6959}\ x\ \frac{(60 + 459.67)}{(1.8\ x\ 26.67) + 32 + 459.67}}.$$

$$Capacity\ per\ region\ = 752.59\ m^3/day$$

$$Maximum\ Number\ of\ household = \frac{752.59}{\frac{15}{30}}\ m^3/day\ .$$

$$Maximum\ Number\ of\ household = 1{,}505.18\ \approx 1{,}505\ households.$$

The mmaximum capacity in one region is calculated based on the value of gas flow rate (Q), where the value of Q obtained will be converted to m^3/day. The converted Q value is then divided by assuming average household consumption per month (15 m^3) divided by 30 days, results in a maximum number of the household for each region that is 1,505 households.

4.2. Proportion Number of Households

The proportion of data for the connection needs of each point is shown in Table 1. The proportion of data is done to determine the need for connection capacity at each node. Data on the number of households obtained in each node is then proportioned based on the density of the household in each node. The proportion value of each node obtained is then multiplied by the number of households to be supplied (4,000 households) and to obtain the number of households at each node.

Table 1. *Proportion of number of households per node in the area.*

Node	Number of Households	Proportion	Number of Households per Node
1	*539*	*0,0626671*	*251*
2	*283*	*0,0329032*	*132*
3	*356*	*0,0413905*	*166*
4	*342*	*0,0397628*	*159*
5	*199*	*0,0231368*	*93*
6	*635*	*0,0738286*	*295*
7	*452*	*0,052552*	*210*
8	*278*	*0,0323218*	*129*
9	*502*	*0,0583653*	*233*
10	*497*	*0,057784*	*231*
11	*464*	*0,0539472*	*216*
12	*277*	*0,0322056*	*129*
13	*439*	*0,0510406*	*204*
14	*575*	*0,0668527*	*267*
15	*439*	*0,0510406*	*204*
16	*795*	*0,0924311*	*370*
17	*480*	*0,0558075*	*223*
18	*286*	*0,0332519*	*133*
19	*374*	*0,0434833*	*174*
20	*389*	*0,0452273*	*181*
Total	*8601*	*1*	*4000*

4.3. K-Means Clustering and DC Determination

In this paper, we make the use of K-Means clustering method which has two variables, namely the coordinate variables of each point and the number of households in the form of proportion. The region's capacity is obtained from the number of households in each cluster and shall not exceed the maximum region's capacity. It is calculated based on the gas flow rate equation, which is 1,505 households per region. At the beginning of the assumption, we grouped the nodes into several clusters, so we can determine the number of clusters. The number of clusters is obtained by dividing the total households to be supplied with the maximum region's capacity. At this stage, we obtain 4000/1505 = 2.65 clusters rounded up into 3 clusters.

The initial DC coordinate points under 3 clusters scenario are determined by 3 largest households' values from the twenty nodes. At this stage, we obtained 3 DC coordinate points based on the largest number of households as presented in Table 1. Therefore, we chose the location 16, location 6, and location 14. Furthermore, the grouping of each node under three clusters scenario is based on the smallest distance. The smallest distance is obtained from the ratio of the distance of each node to DC1, DC2, and DC3. After all the nodes are grouped into 3 clusters based on the smallest distance, then we calculated the DC coordinates again based on the average value that has been grouped.

After obtaining the new DC coordinates, iterations continue to see if there are any changes to the membership of each cluster. The new coordinates of each DC are then averaged. Then, we explore the change in the distance of each point with each DC. If there is a change of membership under the three clusters scenario, then iteration will continue as the steps in iteration 1 until there is no change of membership under the three clusters scenario. The end result of clustering as well as the membership of the points under the three clusters scenario show the total number of households at least at one cluster surpass the specified maximum households that can be supplied from one DC. Therefore, clustering and re-grouping of all node's spots by K-Means clustering method as clustering steps must be done until no cluster outstrips the maximum region's capacity. This step is done by adding the number of clusters.

After doing the third iteration in clustering process under 5 cluster scenarios, we obtain DC coordinates from each cluster as shown in Table 2. Table 2 also presents the final membership of each cluster. Because of no cluster traverse the maximum region's capacity, no additional cluster scenario is needed.

Table 2. *Data coordination of each cluster under 3 DCs scenario (iteration 1).*

	x_i	y_i	Membership
DC1	*-6,3876020*	*106,8219400*	*10*
DC2	*-6,3824291*	*106,8279736*	*1, 6, 12, 15, 16, 18*
DC3	*-6,3743809*	*106,8375068*	*2, 3, 4, 5, 19, 20*
DC4	*-6,3799174*	*106,8340668*	*11, 14*
DC5	*-6,3885759*	*106,8274409*	*7, 8, 9, 13, 17*

4.4. Heuristics Algorithm Calculation

In this study, the heuristics algorithm is used to determine the necessary city gas pipeline route. We make the use of the two methods as a comparison to find the best route in terms of minimum distance, such as Nearest Neighbor (NN) and Clarke and Wright´s Savings Algorithm (SA). In calculating the total distance, we utilize the Google Maps application to gain the driving distance. Furthermore, the total driving distance generated from each route needs to be visually reviewed because the principle of the pipeline is different from the driving distance. Pipelines do not need to be installed twice from node A to node B just to make sure the gas flow from inverse order if they are in one flow.

Route determination is done on each cluster by using two heuristics algorithm to see the method with the least route results to be applied. In cluster 1 and cluster 4, the NN and SA methods produce the same total distance, as shown in Table 3. The NN method generates the minimum total distance in cluster 2 and cluster 5. Meanwhile, SA is the method that produces the minimum total distance on cluster 3. Fig. 4 until Fig. 8 shows the visualization the minimal length of city gas network for cluster 1 until cluster 5, respectively.

Table 3. *Total route and total distance.*

Cluster	Heuristics Methods	Route	Total Distance before Review	Total Distance after Review
1	*NN*	*DC-10-DC*	*0*	*0*
	SA	*DC-10-DC*	*0*	*0*
2	*NN*	*DC-6-16-12-15-18-1-DC*	*9777*	*2659*
	SA	*DC-1-18-15-6-16-12-DC*	*8304*	*3776*
3	*NN*	*DC-4-19-3-20-5-2-DC*	*9070*	*1878*
	SA	*DC-3-2-4-19-5-20-DC*	*8722*	*1529*
4	*NN*	*DC-14-11-DC*	*735*	*421*
	SA	*DC-11-14-DC*	*735*	*421*
5	*NN*	*DC-8-9-17-7-13-DC*	*3616*	*1804*
	SA	*DC-7-13-9-17-8-DC*	*3660*	*1949*

Fig. 4. *The performance of NN and SA for cluster 1.*

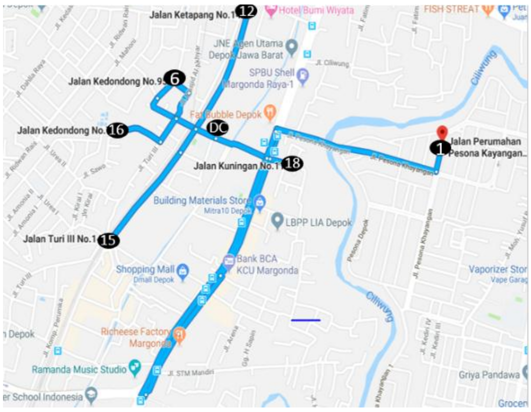

Fig. 5. *The performance of NN for cluster 2.*

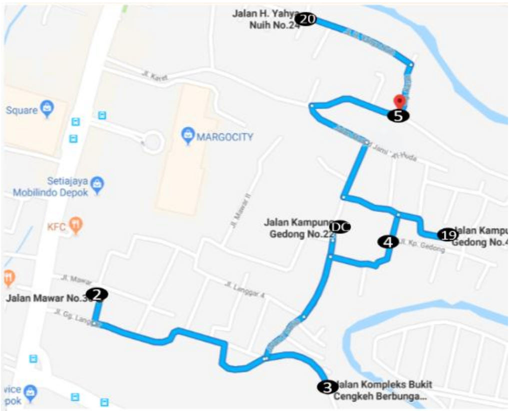

Fig. 6. *The performance of SA for cluster 3.*

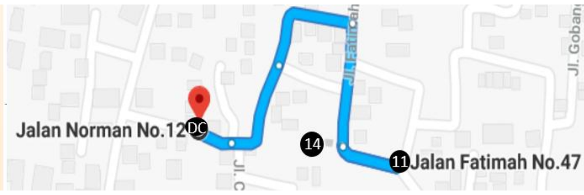

Fig. 7. *The performance of NN and SA for cluster 4.*

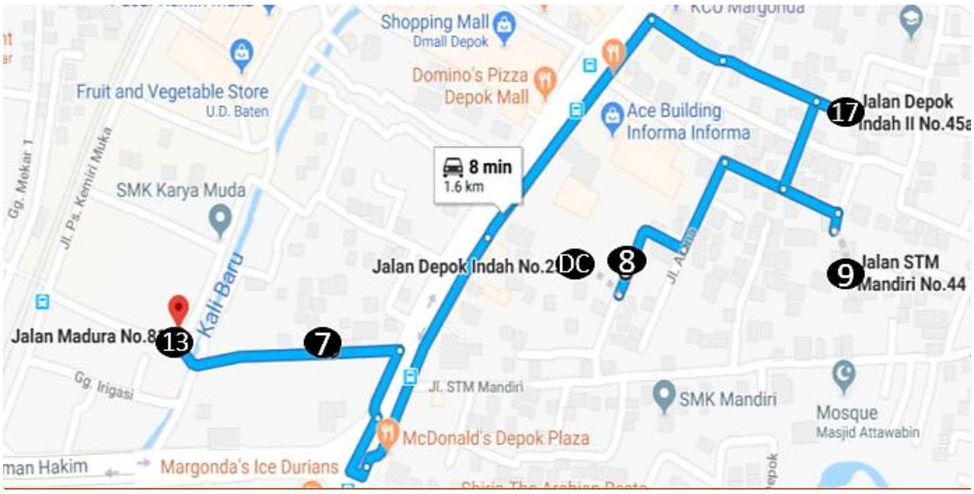

Fig. 8. *The performance of NN and SA for cluster 5.*

4.5. Benefit-cost Ratio (BCR)

Benefit-Cost Ratio (BCR) is used to analyze the feasibility of a project; the use of the equation used to determine the value of BCR depends on the use of Present Worth (PW) as well as the Annual Worth (AW). The selection of annual worth in BCR calculations is based on the data of each variable collected that is annual. The value of benefits is assumed as the government's annual revenue from natural gas sales for households. In addition, the investment value of a project that is assumed to be the cost of building the gas network is depreciated annually based on the project feasibility age. Table 4 explains the total initial investment on all city gas construction sites. The operational and maintenance costs of the project also appear annually, so based on all considerations of the data collected, the annual worth is selected to perform the BCR calculations.

Table 4. *The initial investment.*

No	Item	Size	Total	Unit	Price per Unit (IDR)	Total Price (IDR)
I	*Sector Regulator*	*Typical*	*1*	*Unit*		
1	*Regulator*		*5*	*pcs*	*65.000.000*	*325.000.000*
2	*Ball Valve*	*2 inches*	*10*	*pcs*	*5.000.000*	*50.000.000*
3	*Relief Valve*		*5*	*pcs*	*2.000.000*	*10.000.000*
4	*Shelter+Pipe+Instrument*		*5*	*set*	*20.000.000*	*100.000.000*
5	*Installation and Construction Work*		*5*	*LS*	*7.000.000*	*35.000.000*
	Total I	*Typical*	*I*	*Unit*		*520.000.000*
II	*Pipeline PE (MDPE 80 SDR 11)*			*Unit*		
1	*Pipeline 125 mm Diameter*		*6812*	*m*	*80.000*	*544.960.000*
2	*Pipeline 90 mm Diameter*		*6413*	*m*	*114.000*	*731.082.000*
	Total II					*1.276.042.000*
III	*Household*					
	Number of Households		*4000*	*Unit*	*2.535.000*	*10.140.000.000*
	Total III					*10.140.000.000*
Grand Total (I+II+III)						*11.936.042.000*

Calculation of depreciation value:

$$D_{SL} = \frac{11.936.042.000 - 0}{25} = 477.441.680 .$$

Calculation of annual worth investment:

$$AW_1 = \left(P\left(\frac{A}{P}, 12\%, 25\right)\right) - D_{SL} .$$

$$AW_1 = 11.936.042.000(0{,}1275) - 477.441.680 = 1.044.403.675.$$

Annual worth O & M calculation:

Annual operational and maintenance costs are calculated to be a maximum of 5%.

$$AW_{0\&M} = AW_1 \times 5\%.$$

$$AW_{O\&M} = 1.044.403.675 \times 5\% = 52.220.183{,}75.$$

Annual calculation is worth the benefit:

$$AW_{Benefit} = 2790 \times (15 \times 4000 \times 12) = 2.008.800.000.$$

$$BCR = \frac{2.008.800.000 - 52.220.183{,}75}{1.044.403.675 - 0} = 1{,}87 .$$

Assume the minimum gas price at break even point (BCR = 1).

$$BCR = \frac{AW_{Benefit - AW_{O\&M}}}{AW_1 - AW_{SL}} .$$

$$1 = \frac{X - 52.220.183{,}75}{1.044.403.675 - 0} .$$

$$X = 1.044.403.675 + 52.220.183{,}75 = 1.096.623.859 .$$

$$X = selling\ price \times (15 \times 4000 \times 12) .$$

$$Selling\ price = \frac{1.096.623.859}{(15 \times 4000 \times 12)} = 1.523/m^3 .$$

The calculation of BCR based on natural gas selling price determined by the government for Depok area is IDR 2,790/m^3. We assume the average use of natural gas per household is 15 m^3/month, the interest rate factor value at I = 12% and n = 25. The calculation results show that BCR> 1 means that the project is feasible to be realized.

5. Conclusions

Based on calculations using Weymouth Formula, the maximum region's capacity for each DC was 1,505 households. In other words, 1 (one) DC was only able to accommodate 1,505 households. At the stage of the first clustering under 3 (three) clusters scenario, there was at least one cluster transcended the region's capacity. After the third clustering under 5 clusters scenario, there was no cluster surpassed the region's capacity, meaning these DC points will guide the location to construct the facilities. In calculating the total distance generated using Google Maps application, we preferred and gained the driving distance rather than Euclidean distance in view of precision. The total distance generated from each route needs to be visually reviewed because the principle of the pipeline is different from the driving distance. Pipelines do not need to be installed twice from node A to node B just to make sure the gas flow from inverse order if they are in one flow. The total length of the gas pipeline is built in Depok city, which is 6,812 meters for pipes with a diameter of 125 mm and 6,413 meters for pipes with a diameter of 90 mm. The main purpose of minimizing the distance by having the closest route is to save the cost resulting from the construction of the city gas network. Further calculation shows the study feasibility of the project based on the incurred costs, as well as the earned income, is highly worthy to run.

References

[1] Ministry of Energy and Mineral Resources (2015). *Strategic plan of the ministry of energy and mineral resources 2015-2019*. Retrieved July 06, 2019 from http://www.apbi-icma.org//uploads/files/old/2016/01/Renstra-KESDM-2015-2019.pdf

[2] Ministry of Energy and Mineral Resources. (2016). *Handbook of energy and economic statistics of Indonesia 2016*. Retrieved July 06, 2019 from https://www.esdm.go.id/assets/media/content/content-handbook-of-energy-economic-statistics-of-indonesia-2016-lvekpnc.pdf

[3] Ahmad, H. (2017). Impact of slashing oil prices on the natural gas market. In S. K. Kar., & A. Gupta (Eds.), *Natural gas markets in India* (pp. 33-42). Springer, Singapore.

[4] Paliwal, P. (2017). Natural gas pricing. In S. K. Kar., & A. Gupta (Eds.), *Natural gas markets in india: opportunities and challenges* (pp. 75-93). Springer, Singapore.

[5] Kar, S. K., Sinha, P. K., & Dholakia, B. (2017). Building and sustaining natural gas business in India. In S. K. Kar., & A. Gupta (Eds.), *Natural gas markets in India: opportunities and challenges* (pp.167-196). Singapore: Springer Nature Singapore.

[6] Shanker, R., & Bhattacharya, M. (2017). Brain tumor segmentation of normal and pathological tissues using K-mean clustering with fuzzy C-mean clustering. In J. M. R. S. Tavares., & R. M. N. Jorge (Eds.), *European congress on computational methods in applied sciences and engineering* (pp. 286-296). Springer, Cham.

[7] Jodas, D. S., Pereira, A. S., & Tavares, J. M. R. (2017). Automatic segmentation of the lumen in magnetic resonance images of the carotid artery. In J. M. R. S. Tavares., & R. M. N. Jorge (Eds.), *European congress on computational methods in applied sciences and engineering* (pp. 92-101). Springer, Cham.

[8] Herrera, W. G., Cover, G. S., & Rittner, L. (2017). Pixel-based classification method for corpus callosum segmentation on diffusion-MRI. In J. M. R. S. Tavares., & R. M. N. Jorge (Eds.), *European congress on computational methods in applied sciences and engineering* (pp. 217-224). Springer, Cham.

[9] Li, C., Sun, L., Jia, J., Cai, Y., & Wang, X. (2016). Risk assessment of water pollution sources based on an integrated k-means clustering and set pair analysis method in the region of Shiyan, China. *Science of the total environment*, *557*, 307-316.

[10] Tartavulea, R. I. (2015). Model for determining the optimum location for performance improvement in supply-chain strategies. *European journal of interdisciplinary studies*, *7*(1), 39.

[11] Liu, D., Hou, M., & Qu, H. (2013). A simple model for the multiple traveling salesmen problem with single depot and multiple sink. *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*, *32*(2), 556-574.

[12] Lam, C. H., Choy, K. L., Ho, G. T., & Lee, C. K. M. (2014). An order-picking operations system for managing the batching activities in a warehouse. *International journal of systems science*, *45*(6), 1283-1295.

[13] Jeřábek, K., Majercak, P., Kliestik, T., & Valaskova, K. (2016). Application of clark and wright´ s savings algorithm model to solve routing problem in supply logistics [Special issue]. *NAŠE MORE: znanstveno-stručni časopis za more i pomorstvo*, *63*(3), 115-119.

[14] Sathyanarayanan, S., Joseph, K. S., & Jayakumar, S. K. V. (2015). A hybrid population seeding technique based genetic algorithm for stochastic multiple depot vehicle routing problem. *2015 international conference on computing and communications technologies (ICCCT)* (pp. 119-127). IEEE.

[15] Pichpibul, T., & Kawtummachai, R. (2013). A heuristic approach based on clarke-wright algorithm for open vehicle routing problem. *The scientific world journal*. http://dx.doi.org/10.1155/2013/874349
[16] Hashi, E. K., Hasan, M. R., & Zaman, M. S. U. (2015, November). A heuristic solution of the Vehicle Routing Problem to optimize the office bus routing and scheduling using Clarke & Wright's savings algorithm. *2015 international conference on computer and information engineering (ICCIE)* (pp. 13-16). IEEE.
[17] Charoenwong, C., & Pathomsiri, S. (2015, May). Vehicle routing for improving financial performance: A case study of a freight transportation service provider in Thailand. *4th international conference on advanced logistics and transport (ICALT)* (pp. 263-268). IEEE.
[18] Young, J. R., Suon, S., Rast, L., Nampanya, S., Windsor, P. A., & Bush, R. D. (2016). Benefit-cost analysis of foot and mouth disease control in large ruminants in cambodia. *Transboundary and emerging diseases*, *63*(5), 508-522.
[19] Hausken, K. (2016). Cost benefit analysis of war. *International journal of conflict management*, *27*(4), 454-469.
[20] Rushton, A., Croucher, P., & Baker, P. (2014). *The handbook of logistics and distribution management: Understanding the supply chain*. Kogan Page Publishers.
[21] Meindl, S. C. P. (2016). *Supply chain management--strategy, planning and operation*. Tsinghua University Press. wheat soybean others land for no use.
[22] Panagakos, G., Psaraftis, H. N., & Holte, E. A. (2015). Green corridors and their possible impact on the European supply chain. In C.-Y. Lee., & Q. Meng (Eds). *Handbook of ocean container transport logistics* (pp. 521-550). Springer, Cham.
[23] Amani, H., Kariminezhad, H., & Kazemzadeh, H. (2016). Development of natural gas flow rate in pipeline networks based on unsteady state Weymouth equation. *Journal of natural gas science and engineering*, *33*, 427-437.
[24] Aggarwal, C. C. (2015). *Data mining: the textbook*. Springer.
[25] Badiru, A. B. (2005). Cluster analysis: a tool for industrial engineers. In *Handbook of industrial and systems engineering* (pp. 393-402). CRC Press.
[26] Andritsos, P., & Tsaparas, P. (2017). Categorical data clustering. In C. Sammut., & G. I. Webb (Eds.), *Encyclopedia of machine learning and data mining* (pp. 188-193). Springer, Boston, MA
[27] Zafar, M. H., & Ilyas, M. (2015). A clustering based study of classification algorithms. *International journal of database theory and application, 8*(1), 11-22.
[28] Jin, X., & Han, J. (2017). K-Means clustering. In C. Sammut., G. I. Webb (Eds.), Encyclopedia of machine learning and data mining (pp. 695-697). Springer, Boston, MA.
[29] Diaby, M., & Karwan, M. H. (2016). *Advances in combinatorial optimization*. Singapore: World Scientific Publishing, 2016.
[30] Bramel, J., & Simchi-Levi, D. (1997). *The logic of logistics: theory, algorithms, and applications for logistics management* (pp. 175-240). New York: Springer.
[31] Nahmias, S., & Cheng, Y. (2005). *Production and operations analysis* (Vol. 6). New York: McGraw-hill.