ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MODIFIED K-MEANS CLUSTERING MODEL IN MULTI STORE DELIVERY SERVICE

PURBA DARU KUSUMA
School of Electrical Engineering, Telkom University, Indonesia
E-mail: purbodaru@gmail.com

ABSTRACT

There are several delivery service problems in companies that have multiple stores in one city. These problems occur especially for companies that offer products that these products must be delivered to the customers' location by using their own delivery service. For several companies, they distribute their stock in a single main warehouse and in their stores. In the other side, their delivery service fleet is also distributed in their main warehouse and in every store. This condition triggers inefficiency in stock and the delivery fleet. In this work, we propose the centralized shared delivery service model. As a centralized model, the delivery service is handled by the central management so that coordination in delivery process among vehicles can be more efficient. As a shared system, the vehicle is not dedicated for single store only so that the vehicle can deliver products that come from more than one store in a single trip. In warehouse management, we use single warehouse concept so that all purchased products from all stores will be delivered from the main warehouse. In this work, we propose modified k-means clustering model in managing the delivery process. By using clustering mechanism, each vehicle will deliver products that their destination location is near to each other. In this work, we propose two variants of the k-means clustering model. In the first variant, we combine the k-means clustering method with the round robin method. In the second variant, we combine the k-means clustering method with sequential vehicle creation method. There are research findings after we have done tests. The increasing of the city size makes all observed variables increase. This condition occurs in all models. The increasing of the maximum delivery distance does not affect the total delivery distance but makes the number of vehicles decrease and in the other side makes the delivery distance per vehicle increase. The increasing of the number of stores does not affect the total delivery distance. In the first model, the increasing of the number of stores makes the number of vehicles increase and the delivery distance per vehicle decrease. In the other models, the increasing of the number of stores does not affect the number of vehicles and the delivery distance per vehicle. The increasing of the number of destinations makes all observed variables increase.

Keywords: *Delivery Service, K-Means Clustering, Round Robin, Single Warehouse Multi Store.*

1. INTRODUCTION

It is common for retail companies to have multiple physical stores in local, regional, national, or international coverage [1]. There are many retail companies that have multi store, such as: 7-Eleven [2], Starbuck [2], or IKEA [1]. By having multiple stores, they can increase the customer awareness [3], sales [3], and build better relationship with the customer [1]. By having more than one store, a company can applied customized pricing or uniform pricing to attract the customer in that area and to adapt with the local cost that may be different from one location to another [4].

By deploying more than one store, they can be as close as possible to the customer [1] because the customer can visit the store that its location is the nearest to his location. In the other hand, by having more than one store, company has better economics of scale in purchasing and lower marketing and distribution cost [1]. Having many stores in some regions also increases this company's market dominance [5]. For example, 7-Eleven Japan increases its stores from 5,500 stores in 21 prefectures to more than 10,000 stores in 32 prefectures between 1994 and 2003 [5].

Besides those advantages, having multiple stores needs more effort and increases complexity.

Warehouse and distribution centers are usually expensive to build, operate, and maintain [6]. Having lots of stores means the company must have and operate vehicles to transport goods between the warehouse and the shop [6]. The companies should distribute their stock into every store. Meanwhile, the stock quality changes during the time. It is common that for some products, their stock are enough in some stores but are not enough in other stores. In many cases, this condition makes the customers cancel the transaction because they cannot find their interesting products in the store that they visit.

Besides inefficiency in storage, there are problems in transporting product from supplier to retailer. This problem occurs in basic single-warehouse multi-retailer inventory system [14]. In this basic model, there are two transportation steps. The first step is transportation from supplier to the main warehouse [14]. The second step is transportation from main warehouse to retailers. Unfortunately, there is not any product transportation from one retailer to other retailers [14].

The problem becomes more complex when the characteristic of the products is products that need to be delivered to the destinations after being purchased, for example: home appliances, furniture, etc. In this case, company must have delivery service to deliver the purchased products because customer cannot pick their purchased products up by themselves. For many companies, their delivery vehicle is dedicated to single store. In the other side, some stores have more than one delivery vehicle because there are lots of daily transactions in these stores. If the basic single-warehouse multi-retailers model [14] is implemented in this situation, then there are three transportation steps: (1) supplier-warehouse, (2) warehouse-retailers, and (3) retailer-customers.

When the company implements the dedicated delivery service model, there are potentials in inefficiency. First, for stores that have high number of daily purchased products, sometimes the purchased products cannot be delivered in a single day because they need more vehicles and the related stores cannot utilize other idle stores' vehicles because this idle vehicle is dedicated to the other store even the other stores produce lower number of purchased products. In the other side, for lower utilized vehicle, it cannot deliver other stores' purchased product.

Based on these problems, some companies implement single main warehouse. Meanwhile, stocks in every store act as a display only. So, all purchased products will be delivered from the main warehouse. In this scenario, the challenge is planning the route. So, the research question in this work is how to create the route plan model so that the vehicle utilization and the delivery process can be increased.

Based on this research question, the research purpose in this work is developing the delivery service model that can optimize the company in three ways. The first is increasing the vehicle utilization and balancing the load among vehicle. The second is reducing the delivery distance because reducing delivery distance will reduce the delivery time and delivery cost. The third is reducing transportation step that usually occurs in basic single–warehouse multi-retailers system [14].

This model will be developed by combining the round robin model and the k-means clustering model. The reason of combining these methods is as follows. Round robin model is a very famous model in load balancing work. This model is applied in many fields and is used in many researches, especially in computing and telecommunication. Mishra et al uses round robin in fair scheduling process for relief logistic [7]. In their work, the round robin method is used to allocate the limited number of distribution vehicles in disaster affected region [7]. Zongyu and Xingxuan use round robin scheduling algorithm in web cluster system [8]. Li and Chang use weighted round robin method for network load balancing in multiple link environment [9]. In their work, they stated that their method is efficient in maintaining load balance for multiple input and output network structure [9].

Meanwhile, k-means is a famous model in clustering work. This model is also studied in many researches in solving the clustering problems. Singh et al combined K-means, Heuristic K-means, and Fuzzy C-Means in document clustering and their work perform better result and more stable method [10]. Dehariya et al use K-means and Fuzzy K-means for clustering data in image processing work and this method gives better segmentation result [11].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

This work is the continuation of our previous work in logistic modeling [12,13]. In the first work, we use nearest distance method in developing collaborative delivery service model for merchants in the trade center area [12]. In the second work, we combine the round robin method with the least effort method in scheduled shipping service that is run by the local courier service [13].

This paper is organized as follows. In the first section, we explain the background, research motivation, research question, research purpose, and the paper organization. In the second section, we explain the existing problem in multi store delivery system. In the third section, we propose the model. In the fourth section, we make the test, discuss the test result, and describe the research findings. In the fifth section, we conclude the work and propose the future search potentials.

2. PROBLEMS IN EXISTING MULTI STORE DELIVERY SERVICE

There are classic logistic problems in a retail company that has multiple stores. The first problem is coordinating the stock in every store. In common case, company purchases goods from their suppliers, keeps them in its main or central warehouse, and distributes them into every store based on the request of the store. In this situation, stock in every store is not always the same to each others. The problem is when there is a customer who wants to purchase product in a store, sometimes the store cannot fulfill this request because its stock is not enough even this requested product is available in other stocks.

The illustration is as follows. Suppose that there is a company that has five stores in a city $\{s_1, s_2, s_3, s_4, s_5\}$. Meanwhile, each store sells product g_1. The number of products of g_1 in every store is follows {3, 4, 10, 5, 2} consecutively related to the store index. Then, there is a customer comes to store s3 and requests five units of product g_1. In this case, the store can fulfill this customer's request because its won stock of g_1 is more than or equal to the request. In other hand, there is another customer comes to store s_1 and requests 4 units of product g_1. In this case, the transaction will be fail because this store cannot fulfill the customer's request. Its stock is less than this customer's request. Unfortunately, if the stocks from several stores are combined, store s_1 will be able to fulfill this request.

Another problem is the delivery process. This problem occurs in company that sells product that must be delivered to the customer's location after it is purchased. In this case there are several common models that are adopted. The first model is the company adopts distributed warehouse and every store has dedicated vehicle or vehicles. The second model is the company adopts single warehouse so that all products will be delivered from the main warehouse. This model can be divided into two sub models. The first sub model is dedicated vehicle sub model. Similar to the first model, in this sub model, every store has its dedicated vehicle or vehicles. In the second sub model, the shared vehicle is adopted so that a vehicle can deliver products that the transactions come from multiple store as long as the destinations is near to each others.

The example of this problem is as follows. Suppose that there is a company that has three stores $\{s_1, s_2, s_3\}$. This company has single main warehouse w_1. Store s_1 must deliver products to customers $\{c_1, c_3, c_5\}$. Store s_2 must deliver products to customers $\{c_2, c_4, c_6\}$. Store s_3 must deliver products to customer $\{c_7\}$.The illustration is shown in Figure 1.

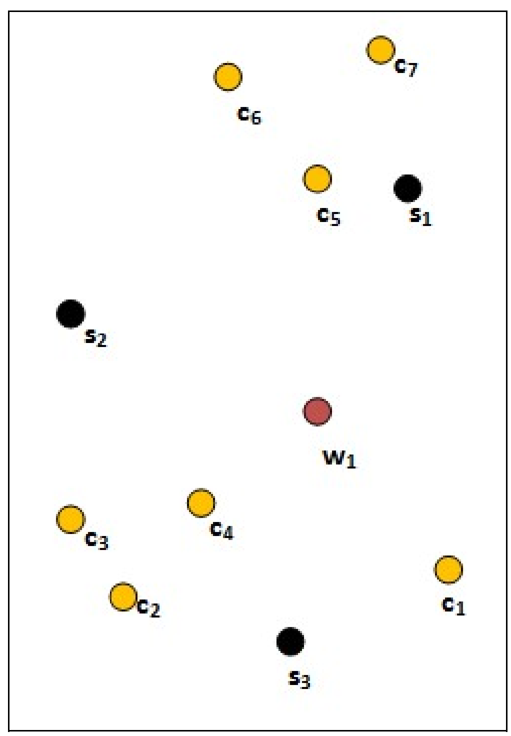

Figure 1. Multi Store Delivery Problem

When the first model is applied, the delivery route plan is as follows. Let us assume that there is three vehicles $\{v_1, v_2, v_3\}$ that are dedicated to each store consecutively. The nearest distance model is used as routing model. The route plan for v_1 is $s1 \rightarrow c_5 \rightarrow c_1 \rightarrow c_3$. The route plan for v_2 is $s_2 \rightarrow c_4 \rightarrow c_2 \rightarrow c_6$. The route plan for v3 is $s_3 \rightarrow c_7$.

When the second model with dedicated vehicle sub model is applied, the delivery route plan is as follows. The route plan for v_1 is $w_1 \rightarrow c_1 \rightarrow c_3 \rightarrow c_5$. The route plan for v_2 is $w_1 \rightarrow c_4 \rightarrow c_2 \rightarrow c_6$. The route plan for v3 is $w_1 \rightarrow c_7$.

When the second model with shared vehicle sub model is applied then the delivery route plan is as follows. Suppose that the nearest distance model is used and the maximum number or customers is three customers. The route plan for the v_1 is $w_1 \rightarrow c_4 \rightarrow c_2 \rightarrow c_3$. The route plan for the v_2 is $w_1 \rightarrow c_1 \rightarrow c_5 \rightarrow c_6$. The route plan for the v_3 is $w_1 \rightarrow c_7$.

As it is shown in Figure 1, basically the customers' location can be clustered into three clusters. Cluster one contains c_2, c_3, and c_4. Cluster two contains c_5, c_6, and c_7. Each cluster will be executed by one vehicle. In this scenario, the shared vehicle model is used. By using this clustering method, the delivery process will be more efficient. Based on this opportunity, in this work, we use k-means clustering method as a basis to propose new delivery model.

Unfortunately, basic clustering method cannot be implemented directly to solve this problem. In the delivery system, one key parameter is the vehicle maximum capacity. It means that the number of products that can be picked up in a single trip is limited by the vehicle maximum capacity. Meanwhile, in the basic k-means clustering method, the maximum number of members in one cluster is not concerned. The other problem is that in the delivery service, the maximum number of destinations that can be visited in a single trip is also limited by the maximum travel distance. This problem is also not concerned in the basic k-means clustering method. Based on this condition, the k-means clustering should be modified so that it can be implemented to solve these problems.

As it is shown in Figure 1, the utilization among vehicles is not fair. These two vehicles visit three destinations. In the other side, the last vehicle visits one destination. In the real world, drivers are usually paid daily or monthly. The number of destinations that are visited in a day is ignored. So, it will be unfair for the two drivers who visit three locations because they receive same amount of wage with the driver who visits only one location.

The other problem is the delivery vehicle is usually bought by loan. Company must pay fix monthly payment no matter this vehicle utilization is high or low. Based on this problem, load balancing among vehicles must be implemented.

3. PROPOSED MODEL

Based on problems that are explained in section two, we propose new delivery service model in this work. This model is developed based on clustering method so that this proposed model is different with models in our previous work [12]. In this work, the model is divided into two steps. The first step is clustering the destinations. The second step is dispatching the destinations in the cluster into the vehicles and balancing the load among vehicles.

In this work, we propose two models. Both models use k-means clustering method to process the first step. The difference between two models is in the second step. The first model uses round robin method in the second step. The second model uses the nearest destination method in the second step.

In the first step, our proposed models use k-means clustering method. In the k-means clustering method, the number of clusters is defined previously and each cluster contains single centroid. The main algorithm of k-means clustering is shown in Figure 2.

```
begin
 initialize_centroid_location()
 status ← "run"
 while status = "run"
 begin
  dispatch_destination()
  update_centroid_location()
  status ← check_status()
 end
end
```

Figure 2. K-means Clustering Main Algorithm

The first process in k-means clustering is initializing the centroids' location. In Figure 2, this process is executed by using the initialize_centroid_location procedure. The initial centroid location is determined randomly and this random number follows uniform distribution. This process is shown in Equation 1 and Equation 2. In

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Equation 1 and 2, e denotes the centroid. Variable x denotes the horizontal position and variable y denotes the vertical position. Variable i denote the cluster index.

$$e_{x,i,0} = random(0, width) \quad (1)$$

$$e_{y,i,0} = random(0, length) \quad (2)$$

After the location of the centroids has been determined, the next process is dispatching the destinations into the selected cluster. In Figure 2, this process is executed by using the dispatch_destination procedure. By using k-means clustering, the selected cluster (l_{sel}) for the destination is cluster that its centroid location is the nearest to the destination location. This process is formulized by using Equation 3 and Equation 4. In Equation 4, the distance between centroid and destination is the Euclidean distance between them.

$$l_{sel,j} = l \mid \min(d(e, c_j)) \wedge l \in C \quad (3)$$

$$d(e, c) = \|e - c\| \quad (4)$$

After all destinations have been dispatched, then the next step is updating the centroids' location. This process is executed by using the update_centroid_location procedure. In this process, there are two options. If the cluster has members then the new location of the centroid will be the average location of this centroid's members. Else, this centroid will stay at its previous location. This process is determined by using Equation 5 and Equation 6.

$$e_{x,i,n+1} = \begin{cases} e_{x,i,n}, n(l_i) = 0 \\ \dfrac{\sum c_x \mid c \in l_i}{n(l_i)} else \end{cases} \quad (5)$$

$$e_{y,i,n+1} = \begin{cases} e_{x,i,n}, n(l_i) = 0 \\ \dfrac{\sum c_y \mid c \in l_i}{n(l_i)} else \end{cases} \quad (6)$$

After the new centroid is being updated, the next process is checking the status whether it should iterate or stop. In this work, the iteration will stop if the total destination to centroid distance is minimized. This process is executed by using the check_status function. This process id determined by using Equation 7 to Equation 9. In Equation 7, the iteration still continues if the current total distance is lower than the previous total distance. In Equation 8, the total distance is the summation of distance of all destinations with their centroid. In Equation 9, this destination between the destination and its centroid is the Euclidean distance between them.

$$status = \begin{cases} run, d_{tot,n} < d_{tot,n-1} \\ stop, else \end{cases} \quad (7)$$

$$d_{tot} = \sum_{j=1}^{n_d} d(l_{sel,j}, c_j) \quad (8)$$

$$d(l_{sel,j}, c_j) = \|l_{sel,j} - c_j\| \quad (9)$$

After the first step is done, the second process will be executed. Based on the clustering process in the first step, clustering process produces several numbers of clusters which each cluster has its own members. The number of members in each cluster may be different among clusters. Besides that, because the total distance of cluster members to its centroid is different among clusters, the travel distance to deliver products in each cluster may be different too.

In the first balancing model, we use round robin method for load balancing. The process of the round robin balancing process is as follows. At the beginning, all destinations' status is set 0 which means this destination has not been executed. Then the round robin process begins. At the first time, each vehicle is located in main warehouse. Then, the rotating process runs for every vehicle to find its destination within its cluster. This process is determined by using nearest destination method so that the vehicle will take its nearest available destination. Each time the vehicle finds its destination then this destination will be allocated to this vehicle and the vehicle position is set as this current destination position. After a destination has been allocated then its status is set 1. If the vehicle reaches its maximum delivery distance then this vehicle stops to find new destination and new vehicle is generated to replace this vehicle. This new vehicle initial position is in the main warehouse. When a vehicle does not meet any available destination in its own cluster then this vehicle will search its nearest available destination in other clusters. This process will stop after all destinations have been allocated. This round robin based balancing process main algorithm is shown in Figure 3.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

```
Begin
 token ← 0
 n_availc ← n_c
 while n_availc > 0 do
 begin
  c_sel ← nearest_dest(v_token)
  c_sel,status ← 1
  v_token,dist ← v_token,dist+d(c_sel,v_token)
  v_token,x ← c_sel,x
  v_token,y ← c_sel,y
  if v_token,dist ≥ d_max then
  begin
   v_token,status ← 0
   create_new_vehicle(token)
  end
  n_availc ← n_availc - 1
    if n_availc > 0 then
   token ← next_turn(token)
 end
end
```

Figure 3. Round Robin Process Main Algorithm

The explanation of this main algorithm is as follows. At the beginning the token is set 0 so that the vehicle that represents cluster 0 get the first turn. Variable n_{availc} denotes the number of available destinations in the system. At the beginning, this variable's value is the total number of destinations. The iteration runs as long as the number of available destinations is more than zero. The first process in the iteration is finding the nearest destination by using the find_nearest_dest function. Variable c_{sel} denote the selected destination. Then, the vehicle's delivery distance ($v_{token,dist}$) and position ($v_{token,x}$;$v_{token,y}$) are updated. If this vehicle's delivery distance exceeds the maximum delivery distance (d_{max}) then this current vehicle will be inactivated and new vehicle in this cluster will be created by using the create_new_vehicle function. If there is any available destination then the token is passed to the next vehicle by using the next_turn function.

The second balancing sub model is called sequential model. In this model, the destinations are allocated sequentially based on the cluster index. The allocation process in the next cluster will runs after the allocation process in the current cluster ends. Similar to the round robin based model, the nearest destination method is used in this sequential model. If a vehicle reaches the maximum delivery distance then this vehicle is deactivated and is replaced by new vehicle. If a vehicle cannot find any available destinations in its own cluster then this vehicle will find available destination in other cluster. Similar to the round robin model, iteration in this second model will run until all destinations have been allocated. This second model main algorithm is shown in Figure 4.

In this algorithm, at the beginning, new vehicle is created. In this step, the token that indicates the active cluster has not been determined. Then, the active cluster is determined by using the nearest_cluster function. In this function, the selected cluster is the cluster that still has at least one available destination and its centroid is the nearest the active vehicle's current position. Then the iteration begins. The iteration will stop if there is not any available destination in the system.

```
Begin
 create_new_vehicle()
 token ← nearest_cluster()
 n_availc ← n_c

while n_availc > 0 do
 begin
  if n_avail(token) = 0 then
   deactivated_cluster(token)
   token ← nearest_cluster()
  else
  begin
   c_sel←nearest_dest(v_ac,token)
   c_sel,status ← 1
   v_ac,dist ← v_active,dist + d(c_sel,v_ac)
   v_ac,x ← c_sel,x
   v_ac,y ← c_sel,y

   if v_ac,dist ≥ d_max then
   begin
    v_ac,status ← 0
    create_new_vehicle()
   end
   n_availc ← n_availc - 1
  end
 end
end
```

Figure 4. Sequential Process Main Algorithm

At the beginning of the iteration, the first process is checking whether the current active cluster still has at least one available destination. If there is not any available destination in this cluster, then this current cluster will be deactivated and the next active cluster will be activated by using the same nearest_cluster function. Else, the nearest available destination in this cluster will be dispatched to this current active vehicle. The next processes following this process are updating the selected destination status, updating this vehicle's delivery distance, and updating the vehicle's current position. Similar to the first sub model, the vehicle's delivery distance will be checked whether its value has exceeded the maximum delivery distance. If it is true then this vehicle will be deactivated and the new vehicle will be created.

4. IMPLEMENTATION AND DISCUSSION

These two proposed models then are implemented into single warehouse multi store delivery simulation application. This application is a web based application and it is developed by using PHP language. The implementation is used to observe and to evaluate performances of these proposed models. Besides comparing to each other, in the implementation and testing, these proposed models are also compared with the nearest distance model that has been developed in our previous work [12]. In this test, the first model represents the round robin-k means clustering model. The second model represents the sequential-k means clustering model. The third model represents the nearest driver model.

The environment of the application is a virtual city. The shape of this city is square so its width and length is equal. In our previous work, the city size is static. In this work, the city size ranges from the minimum value to the maximum value. In this paper, the city size represents the city width or length.

When the simulation runs, some stores and destinations are built. These stores and destinations location is generated randomly and it follows uniform distribution. A main warehouse is built in the center of the city.

After these nodes are built then these destinations are dispatched to certain vehicles. In these proposed models, the dispatching process is started with the clustering process. After the clustering is done and all destinations are grouped into selected cluster then the next process is allocating vehicle to deliver products to the destinations. In this step, the round robin sub model is different to the sequential sub model.

In the nearest distance model, the vehicle is generated sequentially. All vehicles' initial position is in the warehouse. Then, the vehicle will get the available destination that its position is the nearest the vehicle current position. This selected destination location then will be the vehicle's next position. This process will iterate until this vehicle's delivery distance reaches the maximum delivery distance. After that, this vehicle will be deactivated and new vehicle will be activated. This dispatching process will iterate until all destinations has been dispatched.

Four test groups are done to observe and to evaluate the performance of the model. In this test, there are several adjusted variables and observed variables. The adjusted variables are city size, maximum delivery distance, number of stores, and number of destinations. The observed variables are total distance, number of vehicle, and average distance per vehicle. There is default value of the observed variables. This default value is shown in Table 1.

Table 1. Variables Default Value

Variable	Default Value
City size	20 kilometer
Maximum delivery distance	75 kilometer
Number of store	10 units
Number of destinations	150 nodes

In the first test, we observe and evaluate the relationship between the city size and the observed variables. In this test, the city size ranges from 15 kilometer to 25 kilometer with the step size is 1 kilometer. Other adjusted variables are set at their default value. There are five simulation sessions in every step. The result is shown in Table 2.

Journal of Theoretical and Applied Information Technology
15th January 2019. Vol.97. No 1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2. Test Result for All Models with Various City Sizes

size (km)	Round Robin-K Means Model			Sequential-K Means Model			Nearest Distance Model		
	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)
15	686.6	13.8	49.8	702.2	9.6	73.2	708.6	9.6	74.0
16	771.3	14.5	53.3	815.0	11.0	74.0	789.5	10.5	75.3
17	804.2	15.0	53.4	820.6	11.0	74.6	824.0	11.0	75.0
18	863.0	16.0	54.0	892.8	11.6	77.2	878.6	11.4	77.2
19	917.0	17.4	52.8	947.6	12.6	75.2	928.8	12.4	74.8
20	998.6	17.6	56.8	1,035.8	13.2	78.4	967.2	12.8	75.6
21	1,006.0	18.8	53.8	1,040.8	13.3	78.5	1,001.0	13.0	77.0
22	1,075.0	19.2	56.2	1,103.4	14.2	77.8	1,075.8	14.0	76.8
23	1,111.4	19.0	58.8	1,151.6	14.8	77.8	1,113.8	14.2	78.4
24	1,164.6	19.8	59.2	1,194.6	15.2	78.6	1,158.0	15.0	77.4
25	1,174.6	19.2	61.4	1,252.0	16.0	78.2	1,159.8	15.2	76.2

Based on Table 2, it is shown that when the city size increases, the total delivery distance increases too. This condition occurs in all models. By comparing among all models, the total delivery distance among all models is very small. The total delivery distance gap between the maximum value and the minimum value is shown in Table 3. The third column of Table 3 denotes the percentage between the gap and the average total delivery distance among models. In Table 3, it is shown that the percentage between the gap and the average value is less than 10 percent with the minimum percentage is 2.42 percent and the maximum percentage is 7.71 percent.

Table 3. Gap in Total Delivery Distance in the First Test

Size (km)	Gap (km)	Percentage (%)
15	22.0	3.15
16	43.8	5.52
17	19.8	2.43
18	29.8	3.39
19	30.6	3.29
20	68.6	6.86
21	39.8	3.91
22	28.4	2.62
23	40.2	3.57
24	36.6	3.12
25	92.2	7.71

In the number of vehicles aspect, when the city size increases, the number of vehicles increases too. This condition occurs in all models. By comparing among models, the round-robin-k means clustering model produces the highest number of vehicles. The nearest distance model produces the lowest number of vehicles. Although the sequential-k means clustering model produces moderate value, its position is closer to the nearest distance model rather than to the round robin-k means clustering model.

In delivery distance per vehicle aspect, when the city size increases, the delivery distance per vehicle trend is different among models. In the first model, the delivery distance per vehicle tends to increase and its value is the lowest among other models. In the second model, the delivery distance per vehicle tends to increase but with very low gradient. Meanwhile, its value is much higher than the first model. In the third model, the delivery distance per vehicle tends to fluctuate with low amplitude. Its value is competitive compared with the second model.

In the second test, we observe and evaluate the relationship between the maximum delivery distance and the observed variables. In this test, the maximum delivery distance ranges from 50 kilometer to 100 kilometer with the step size is 5 kilometer. Other adjusted variables are set at their default value. There are five simulations sessions in every step. The result is shown in Table 4.

Based on data in table 4, it is shown that when the maximum delivery distance increases, the total delivery distance tends to fluctuate. This condition occurs in all models. The gap between the maximum and the minimum value of the total delivery distance among models is very small. Although the gap is very small, the first model produces the lowest average total delivery distance. The second model produces the highest average total delivery distance.

Table 4. Test Result for All Models with Various Maximum Delivery Distance

d_{max} (km)	Round Robin-K Means Model			Sequential-K Means Model			Nearest Distance Model		
	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)
50	992.8	23.6	42.2	1,015.4	18.8	54.0	958.4	18.2	52.8
55	958.6	20.2	47.6	991.0	16.8	59.2	954.0	16.4	58.2
60	976.6	20.2	48.2	1,019.8	16.2	63.2	1,002.0	15.8	63.4
65	896.2	18.2	49.2	934.8	13.8	67.8	918.2	13.8	66.2
70	952.0	18.0	52.8	982.4	13.6	72.2	962.8	13.4	71.8
75	963.2	17.8	54.2	995.4	13.2	75.4	959.6	12.6	76.0
80	938.6	16.8	56.0	957.4	11.8	81.0	964.4	12.0	80.4
85	976.8	16.0	61.2	1,009.0	11.6	87.2	998.2	11.8	84.6
90	1,005.6	16.0	63.0	1,039.4	11.4	91.4	1,002.6	11.2	89.4
95	979.0	14.8	66.4	1,005.2	10.6	95.0	976.2	10.2	95.8
100	969.4	14.0	69.6	1,012.2	10.2	99.4	1,003.4	10.0	100.4

In the number of vehicles aspect, when the maximum delivery distance increases, the number of vehicles decreases. This condition occurs in all models. Comparing among model, the first model produces the highest number of vehicles. In the other side, the number of vehicles in the second model tends to similar to the number of vehicles in the third model.

In the delivery distance per vehicle aspect, when the maximum delivery distance increases, the delivery distance per vehicle increases too. This condition occurs in all models. The first model produces the lowest delivery distance per model. The delivery distance per vehicle in the second model is similar to the delivery distance in the third model. Meanwhile the increasing delivery distance per vehicle in the second model and in the third model is faster than the increasing delivery distance per vehicle in the first model.

In the third test, we observe and evaluate the relationship between the number of stores and the observed variables. In this test, the number of stores ranges from 5 units to 15 units with the step size is 1 unit. Other adjusted variables are set at their default value. There are five simulations sessions in every step. The result is shown in Table 5.

Table 5. Test Result for All Models with Various Numbers of Stores

n_s (unit)	Round Robin-K Means Model			Sequential-K Means Model			Nearest Distance Model		
	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)
5	963.2	15.0	64.2	988.0	13.0	76.0	970.8	12.6	77.0
6	990.4	15.8	62.8	1,023.4	13.2	77.6	976.6	12.8	76.2
7	953.6	15.6	61.2	980.0	12.8	76.6	1,029.2	13.6	75.8
8	959.4	16.4	58.6	992.4	13.0	76.4	990.8	13.2	75.0
9	995.2	17.4	57.2	1,010.2	13.2	76.6	983.6	13.0	75.6
10	1,010.2	17.4	58.0	1,034.4	13.6	76.2	1,012.6	13.2	76.8
11	1,000.0	18.4	54.2	1,028.8	13.2	78.0	1,025.6	13.2	77.8
12	1,045.8	19.4	54.2	1,098.8	14.0	78.6	1,030.0	13.4	77.0
13	959.8	19.4	49.4	996.4	13.0	76.6	956.0	12.4	77.0
14	973.6	19.2	50.8	1,006.6	13.0	77.6	995.8	12.8	78.0
15	958.2	20.2	47.6	1,017.6	13.4	76.2	939.6	12.2	77.0

Based on Table 5, it is shown that the increasing of the number of stores does not affect the total delivery distance. The total delivery distance tends to fluctuates during the increasing of the number of stores. Although the total delivery distance tends to fluctuate, the first model produces the lowest total delivery distance. The second model produces the highest total delivery distance.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In the number of vehicles aspect, there is different condition among models in the increasing of the number of stores. In the first model, when the number of stores increases, the number of vehicles increases too. In the other side, in the second model and in the third model, the number of vehicles tends to fluctuate during the increasing of the number of stores. So, in the second model and in the third model, the number of stores does not affect the number of vehicles. By comparing the number of vehicles among models, the first model produces the highest number of vehicles. The number of vehicles in the second model tends to similar to the number of vehicles in the third model. Meanwhile, the average number of vehicles in the second model is a little bit higher than in the third model.

In the delivery distance per vehicle aspect, there is different condition among models. In the first model, when the number of stores increases, the delivery distance per vehicle tends to decrease. In the second model and in the third model, the delivery distance per vehicle tends to fluctuate during the increasing of the number of stores. So, in the second model and in the third model, the number of stores does not affect the delivery distance per vehicle. By comparing among models, the first model produces the lowest delivery distance per vehicle. The second model produces similar value of the delivery distance per vehicle with the third model does.

In the fourth test, we observe and evaluate the relationship between the number of destinations and the observed variables. In this test, the number of destinations ranges from 100 nodes to 200 nodes with the step size is 10 nodes. Other adjusted variables are set at their default value. There are five simulations sessions in every step. The result is shown in Table 6.

Table 6. Test Result for All Models with Various Numbers of Destinations

n_d (node)	Round Robin-K Means Model			Sequential-K Means Model			Nearest Distance Model		
	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)	d_{tot} (km)	n_v (unit)	d_{av} (km)
100	688.0	13.4	51.6	710.0	9.4	75.4	655.4	8.8	74.4
110	728.4	14.0	52.2	757.2	10.0	75.8	781.4	10.4	75.2
120	793.4	14.8	54.0	827.2	11.0	75.2	776.0	10.2	76.2
130	849.8	15.6	54.6	868.2	11.2	77.6	823.8	10.6	77.6
140	928.6	17.2	54.0	946.2	12.2	77.6	930.2	12.0	77.6
150	962.2	17.6	54.8	983.4	12.8	76.8	990.6	13.2	75.0
160	1,061.8	18.6	57.2	1,091.0	14.2	76.8	1,090.4	14.0	77.8
170	1,067.4	18.4	58.2	1,113.4	14.6	76.2	1,111.0	14.6	76.2
180	1,172.8	20.0	58.6	1,226.6	16.0	76.8	1,157.8	14.8	78.2
190	1,174.0	19.8	59.6	1,212.6	15.8	76.6	1,186.2	15.6	75.8
200	1,294.6	22.0	59.0	1,322.8	17.0	77.8	1,279.8	16.4	78.0

Based on Table 6, it is shown that the increasing of the number of destinations affect the total delivery distance. This condition occurs in all models. When the number of destinations increases, the total delivery distance increases too. By comparing among models the gap between the highest value and the lowest value is small. This condition is shown in Table 7. As it is shown in the third column in Table 7, the percentage between the gap and the average value is less than 10 percent and it indicates that the gap is small.

Table 7. Total Delivery Distance Gap in The Fourth Test

n_d (node)	Gap (km)	Percentage (%)
100	54.6	7.98
110	53.0	7.01
120	51.2	6.41
130	44.4	5.24
140	17.6	1.88
150	28.4	2.90
160	29.2	2.70
170	46.0	4.19
180	68.8	5.80
190	38.6	3.24
200	43.0	3.31

In the number of vehicles aspect, when the number of destinations increases, the number of vehicles increases too. This condition occurs in all

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

models. By comparing among models, the first model produces the highest number of vehicles while the third model produces the lowest number of vehicles. Although the second model is in the middle, its value is nearer the third model rather than the first model.

In the delivery distance per vehicle aspect, the increasing of the number of destinations makes the delivery distance per vehicle increase too but with low gradient. This condition occurs in all models. The first model produces the lowest delivery distance per vehicle. The second model produces similar delivery distance per vehicle with the third model does.

There are limitations in this work. First, in this work, the product quantity and size for every destination are assumed equal. Unfortunately, in the real world, the product size and quantity for every destination or delivery order may be different. Second, the number of destinations that can be executed in single trip is limited only by the maximum delivery distance. In this scenario, vehicle's capacity is ignored. In the real world, the relation between the vehicle's capacity in one side and order quantity and size of orders in another side is critical. Third, in this work, the product is assumed homogenous. In the real world, there are various types of products in a vehicle. For example, a furniture store may deliver sofa, bed, and cupboard in a single trip. Another example is electronic store may deliver television, refrigerator, and washing machine in a single trip. Fourth, the vehicle speed is assumed static. That is why in this work, the delivery process is limited by the maximum delivery distance. In the real world, the vehicle speed may change during the time and place. The vehicle speed will be slower during the busy hour. In city central, vehicle speed is usually slower too.

There are improvements in this work compared with the previous works [12,14]. Compared with the previous work [12], the improvement is as follows. Basically, there is situation similarity between the coordinated trade center shipping delivery system [12] with single-warehouse multi-store delivery system. The merchants in trade center can be seen as stores in this work. In the previous work [12], the delivery model is developed by using nearest distance method. In this work, we use k-means method that is combined with round robin method or nearest driver method. Based on the simulation result, it is shown that each method has advantages and disadvantages compared with other methods.

Compared with basic single-warehouse multi-retailer system [14], the improvement is as follows. As it is mentioned above, in the basic single-warehouse multi-retailer system, there is transportation step from warehouse to retailer [14]. This step occurs regularly based on the retailer request or main warehouse initiative [14]. In our proposed model, this step is eliminated. So, there are only two transportation steps: (1) supplier-warehouse and (2) warehouse-customer.

5. CONCLUSION AND FUTURE WORK

Based on the explanation above, the k-means clustering based delivery model has been developed and has been implemented into the delivery process simulation application. This proposed model has reduced the number of transportation step by eliminating the periodic product transportation from warehouse to retailer because the product from warehouse is delivered directly to customer. This work has also contributed in enriching the previous method by using k-means method that is combined with round robin or nearest distance method.

There are research findings after we have done the tests. The increasing of the city size makes all observed variables increase. This condition occurs in all models. The increasing of the maximum delivery distance does not affect the total delivery distance but makes the number of vehicles decrease and in the other side makes the delivery distance per vehicle increase. The increasing of the number of stores does not affect the total delivery distance. In the first model, the increasing of the number of stores makes the number of vehicles increase and the delivery distance per vehicle decrease. In the other models, the increasing of the number of stores does not affect the number of vehicles and the delivery distance per vehicle. The increasing of the number of destinations makes all observed variables increase.

There are some future research potentials in logistic sector, especially in warehouse process and delivery process in retail industry. The change in customer behavior makes many companies must adapt and change their business process. The rapid growth in e-commerce business makes the logistic management and model become significant aspect. In one side, customer demands easy access, fast respond, and low cost delivery. In the other side, company needs more efficient logistic system so they can provide the customer's needs.

The other future research potential is implementing this delivery model with more realistic situation that has been simplified in the limitation of this work. First, the model is improved into delivery service situation where the product quantity and size may be different among delivery orders. Second, the model is improved into delivery service situation where the vehicle capacity is concerned and the capacity may be different among vehicles. Third, the model is improved into delivery service situation where the product that must be picked up is heterogeneous and each product type has specific caring mechanism. Fourth, the environment is an area with various characteristic so that the speed in one area may be different with the speed in other area.

REFERENCES:

[1] M.C.W. Janseen, V.A. Karamychev, P. van Reeven, "Multi-Store Competition: Market Segmentation or Interlacing?", Faculty of Economics, Erasmus University Rottterdam, 2004, dicsussion paper.

[2] A. Dasci, G. Laporte, "A Continuous Model for Multi-Store Competitive Location", *Operations Research*, vol.53(2), 2005.

[3] N.L. Reyhle, "Strength in Numbers: The beneftis of Having Multiple Storefronts", Brightpearl, 23 October, 2017, https://blog.brightpearl.com/the-benefits-of-having-multiple-storefronts.

[4] P.W. Dobson, M. Waterson, "Chain-Store Competition: Customized vs Uniform Pricing", Dept. of Economic, University of Warwick.

[5] S. Chopra, "Seven-Eleven Japan Co", Kellog Graduate School of Management, Northwestern University, 2000.

[6] J. Fernie, L. Sparks, *Logistics & Retail Management: Emerging Issues and New Challenges in the Retail Supply Chain*, Kogan Page, 2009.

[7] B.K. Mishra, K. Dahal, Z. Pervez, "Dual-mode Round-robin Greedy Search with Fair Factor Algorithm for Relief Logistics Scheduling", *Proceeding of The 4th International Conference on Information and Communication Technologies for Disaster management*, 11 December, 2017, Munster.

[8] X. Zongyu, W. Xingxuan, "A Predictive Modified Round Robin Scheduling for Web Server Clusters", Proceeding of 34th Chinese Control Conference, 28 July, 2015, Hangzhou.

[9] D.C. Li, F.M. Chang, "An In-Out Combined Dynamic Weighted Roudn Robin Method for Network Load Balancing", The Computer Journal, vol.50(5), 2007.

[10] V.K. Singh, N. Tiwari, S. Garg, "Document Clustering Using K-Means, Heuristic K-Means and Fuzzy C-Means", *Proceeding of International Conferenceon Computational Intelligence and Communication Networks*, 7 October, 2011, Gwalior.

[11] V.K. Dehariya, S.K. Shrivastava, R.C. Jain, "Clustering of Image Data Set Using K-Means and Fuzzy K-Means Algorithms", *Proceeding of International Conference on Computational Intelligence and Communication Networks*, 26 November, 2010, Bhopal.

[12] P.D. Kusuma, "Nearest Destination Based Coordinated Delivery Service in Trade Center", *Journal of Engineering and Applied Sciences*, 2018, accepted.

[13] P.D. Kusuma and A.B. Osmond, "Round Robin-Least Cost Combined Model in City Courier Scheduled Shipping Service", *Journal of Engineering and Applied Sciences*, vol.13(15), 2018, pp.6181-6187.

[14] Y. Jin, "Single-Warehouse Multi-Retailer Inventory Systems with Full Truckload shipments", Dept. of Mechanical and Industrial Engineering, University of Massachusetts, 2005.