Determining the Number and Location of Warehouses to Minimize Logistics Costs of Business to Consumer (B2C) Distribution

I Nyoman Sutapa[1*], *Magdalena* Wullur[2], and *Tania* Nano Cahyono[1]

[1]Department of Industrial Engineering, Faculty of Industrial Technology, Petra Christian University, Jl. Siwalankerto No.121−131, Surabaya 60236, Indonesia
[2]Department of Management, Faculty of Economics and Business, Sam Ratulangi University, Bahu Link 8, Bahu, Manado 95115, Indonesia

Abstract. The article discusses a case study at a company that works on Fast Moving Consumer Goods (FMCG) section using B2B distribution system, and is conducting a market test for its new alternative tobacco products using B2C distribution system. The purpose of this study is to determine both number and location of warehouses that provides a minimum total logistics costs for the B2C process, where the company is currently still using company-owned B2B warehouses. Determination of the number and location of these warehouses is performed using Agglomerative Hierarchical Clustering where grouping is based on the shortest distance and is done using Evolutionary Solver. Centre of Gravity is also used to determine the location of the warehouses. Iteration will be carried out to obtain the number of warehouses that provide minimal total cost. The result of the modelling shows that the optimal number of warehouses is six and each warehouse will have their own market division.

Keywords: Agglomerative hierarchical clustering, center of gravity, fast moving consumer goods, optimal number of warehouses, tobacco alternative product.

1 Introduction

Distribution and logistics management is a critical company function. Professionals in this field play a key role in fulfilling customer demands, ordering and managing inventory, controlling inbound and outbound shipments, reducing costs, saving time, and meeting company objectives. Distribution is a main logistics activity that holds an important role in a business competitiveness. An efficient transfer of goods from the manufacture to the point of consumption in a cost-effective way while providing an acceptable service to the consumer is the main logistics concern [1].

Nowadays business to consumer (B2C) is one of the most common types of distribution. B2C is a new trend in the distribution of goods in the digital economy era. Currently, a company distribution policy observed in this case study, offers a new tobacco alternative product using that B2C e-commerce business process. Business to Consumer

* Corresponding author: mantapa@petra.ac.id

(B2C) is an activity where the company will directly serve the consumer [2]. According to a survey conducted by Pew Research Centre [3], online shopping is one of the fastest growing internet activities and many consumers prefer to do the online shopping. Therefore, the company is trying to adopt this new business model for its tobacco product.

The new tobacco alternative products are believed to reduce risk compared to the current tobacco products [4]. The company need to determine its number and location of warehouses with a minimum logistics cost for its new business process. The company is running a test market in Jakarta but still uses four available B2B warehouses. The company still does not know whether it is optimal or not in terms of the logistics costs. Based on an initial observation, there are four main variables that influence the total logistics costs, i.e. manpower, outbound and inbound transportation, and warehouse cost. The purpose of this study is to determine both number and location of warehouses for running the B2C process with a minimum logistics cost. This is important to determine, because in B2C, outbound logistics costs are very large, because end customers demand are generally very small so the frequency of delivery to them is very frequent [5].

Several previous studies have examined the problem of determining the number and location of warehouses to minimize logistics costs [6–8]. These models are able to produce optimal solutions, i.e. minimal outbound and inbound transportation costs, but have disadvantages because they do not consider other logistics costs, namely warehouse costs and manpower costs.

Brunaud et al. [6] proposed a mixed-integer linear programming model and Abo-Elnaga et al. [7] developed an active-set trust-region algorithm to determine the optimal number and location of warehouses. These models are able to produce the optimal solution but has the disadvantage of having to determine the number and location of the warehouse as the initial solution. The optimal solution obtained depends on the initial solution chosen, does not guarantee a global optimal solution.

The problem of distribution of goods in B2C is more complex than in B2B, because the number of items shipped is more frequent with smaller volumes, uncertain delivery times, irregular order arrival patterns, and limited time for order processing. Leung et al. [8] developing a B2C e-commerce intelligent system for re-engineering the e-order fulfilment process, but this system has not been able to determine the coverage area of customer demand that can be served by each warehouse.

To overcome some of the weaknesses of the existing model, the B2C distribution model developed is expected to determine the number and location of warehouses that minimize the total costs of inbound and outbound logistics, operational cost warehouses, and manpower costs, and each warehouse can reach consumers with area boundaries close and not overlapping.

2 Research methods

Centre of Gravity (CoG) is a technique used to find the best location at a distribution point so as to minimize the costs incurred for distribution [9]. The main purpose of this method is to find distance that is efficient in terms of the cost of moving goods or services from existing location. The first step needed to find CoG is to know the total demand in each region. CoG calculation will be performed for each cluster result. Agglomerative Hierarchical Clustering (AHC) is one of the cluster methods that is based on hierarchical approach, where it is started with a single cluster which then repeatedly combines the two closest clusters to cover all existing clusters [10]. AHC has three techniques which are single linkage (*min*), complete linkage (*max*) and group average. Single linkage is chosen in this study because it is based on the proximity of two clusters from the minimum distance between two different clusters. Single linkage Formulation (1):

$$d_{uv} = min\{d_{uv}\} \tag{1}$$

where $\{d_{uv}\}$ is the distance between u and v from each cluster u and v.

Clustering is done using Evolutionary Solver from Microsoft Excel. This cluster method is based on the shortest distance between an area and another region, where the regional approach is based on the sub-districts in Jakarta. Evolutionary Solver is based on the idea of survival of the fittest from the theory of evolution [11]. This evolutionary solver can be used for optimization problems [12]. Evolutionary Solver has constraints called population, and the function or destination cell will be evaluated based on existing constraints. Changes in constraint will increase the likelihood that new populations will be located near old populations that have good values. In principle, cluster is a closest market grouping method, after that an analysis of the location of potential facilities is carried out through CoG.

Grouping starts with a warehouse in each demand or market side, then calculate the total cost generated by this solution. The total cost will be the highest total logistics cost because it is obtained with the maximum number of facilities. Then, the number of warehouses will be reduced one by one by clustering and becoming one potential location. This potential location is obtained using CoG. Finally, calculate the total logistics cost after reducing the number of warehouses and repeat all of the steps until there is one warehouse left.

3 Result and discussion

The new tobacco alternative is an import product and the goods will be produced every two months. The determination of the number and location for warehouses using data from Nielsen RTA X2 which is a market survey data based on sales of secondary markets and not sales volumes in Jakarta. The data used for this research is 10 % of SPM (white cigarette machine) sales data in 2017. The data obtained is the SPM sales data for each part of Jakarta (South, West, North, Central and East).

3.1 Determine demand for each sub-district

Demand from each Jakarta regions will be used to determine the demand for each sub-district. The approach used to determine the location of the warehouse starts from sub-districts because the company's business processes are currently based on sub-districts. Determining the number of demand for each sub-district is done based on the proportion of the male population from Indonesia's population census data in 2010.

3.2 Determine the number of orders and workers

The number of orders is depending on the demand of each sub-district, and it also affects the number of delivery. The number of delivery will influence the outbound cost. Outbound cost is the total cost of shipping from warehouse to end-customer using Last Mile Delivery which is Go-Send.

The number of workers is also affect the manpower cost. Each worker has eight hours of work a day. The average order processing time based on current test market is 8.19 min. This time includes picking and packing activities. Therefore, each worker can proceed 2 512 order for 2 mo.

3.3 Clustering and CoG

Clustering is done using Evolutionary Solver from Microsoft Excel by grouping sub-districts that have the closest distance. Distance measurement is done using Google Maps between each sub-districts (using their Latitude and Longitude from Google Maps) and it will be done again after clustering. The initial clustering from East Jakarta can be seen in Fig. 1.
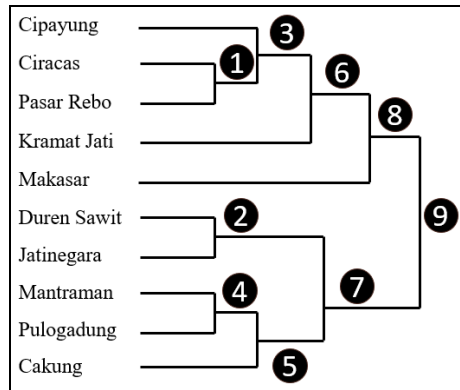
Fig. 1. Initial clustering for East Jakarta Region

Ciracas and Pasar Rebo will be clustered first because of the close distance. The next step is to calculate the Centre of Gravity. CoG will be calculated for each cluster result to determine the warehouse location based on their Longitude and Longitude and demand.

3.4 Iteration on each cluster

Iteration will be carried out based on each cluster results. Iteration is done by counting the total logistics costs of four variables. The first iteration will have the highest logistics costs because each sub-districts will have their own warehouses. Manpower cost will be influenced by the worker's capacity and also the demand from each region. Early stages of clustering will minimize manpower costs, but final stages will consider outbound costs. These two variables have a greater influence according to the current test-market conditions.

The transportation cost consist of two variables which are outbound and inbound cost. Outbound cost calculates the cost of using Go-Send as Last Mile Delivery which is IDR 2 500 km^{-1} where the distance is calculated from the warehouse to the market (end-customer) and it is obtained using Google Maps. Each sub-district coordinate is the location of the market itself. The minimum cost of Go-Send is IDR 20 000 and will be borne by consumers. If the distance is greater than eight kilometres, then the company will have borne that cost. Inbound cost calculates the cost of using third-party logistics (3-PL) where it is also based on the distance between airport and warehouses. The new alternative tobacco product is an import product so airport will be its arrival place.

Warehouse costs will be influenced by the cost of renting a place and the components used in the warehouse. Warehouse rental costs will be affected by the total area of the warehouse. Warehouse components such as computers, tables and chairs will follow the number of workers because each worker will need one of those things to carry out the order fulfilment process. Iteration will be carried out for each region in Jakarta.

The result of 10 iterations from a simulation of the East Jakarta region, the number of warehouses that provide the minimum total logistics costs is two warehouses with

IDR 84 685 346 for every 2 mon. The two warehouses of the East Jakarta region are located in the Cakung sub-district (-6.20532; 106.91262) and Ciracas sub-district (-6.30618; 106.87547). Based on eight of iterations for a simulation of the West Jakarta region, two warehouses will give the minimum total logistics costs which is IDR 66 011 466 for every 2 mo. The location of the two warehouses are in Gambir sub-district (-6.16340; 106.79987) and Cengkareng sub-district (-6.16000, 106.735192). One warehouse will provide the minimum logistics costs for the Central Jakarta region. This condition is due to the location of sub-districts in Central Jakarta that are close to each other so there are no outbound costs in each iteration. The total logistics costs are IDR 60 082 194 for every 2 mo and it is located in Senen sub-district (-6.17833; 106.84029). One warehouse will give the minimum total logistics costs for the North Jakarta region, even if the outbound costs is bigger than the usage of two warehouses. This condition happens because the manpower costs for one warehouse is smaller than the difference in outbound costs. The cost of using one warehouse in North Jakarta is IDR 57 898 477 and it is located in Koja sub-district (-6.12465; 106.90001). The South Jakarta region will also use one warehouse. The minimum total logistics costs are IDR 80 870 648 and it is located in Cilandak sub-district (-6.26712; 106.81127). This warehouse will fulfil requests from 10 sub-districts in the South Jakarta region. Based on the results of the simulation of the five Jakarta region above, it was found that there were 7 warehouses, where East and West Jakarta will each have two warehouses while Central, North and South Jakarta will each have one warehouse. The total logistics cost for seven warehouses are IDR 349 548 131 for every 2 mo.
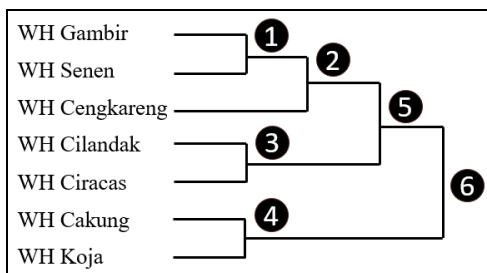

Fig. 2. Clustering for seven warehouses (WH)

From the seven warehouse clusters, iterations will be carried out again to see the possibility of lower logistics costs. Clustering is based on the shortest distance from the seven warehouses above. Calculation of COG will be carried out based on the result of clustering and will also pay attention to demand from the market area (sub-district) served by warehouses. Iterations were carried out again based on the cluster in Fig. 2. Utilization of six warehouses will provide a smaller logistics costs than the use of six warehouses (Fig. 3). The third iteration and so on will give higher total logistics costs. Therefore, six warehouses are the best decision.
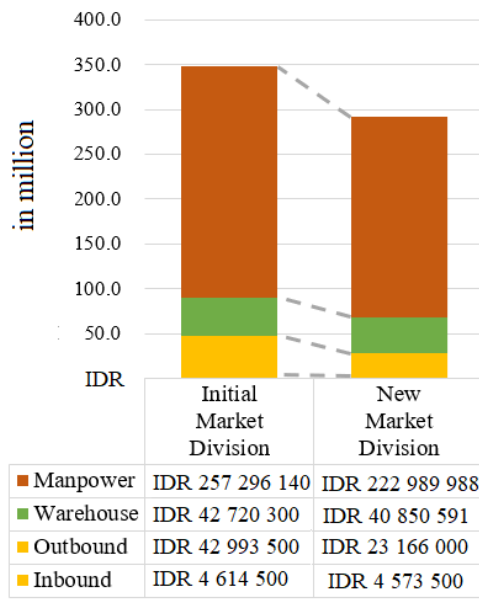
3.5 Determine the market area

Outbound cost gives a big influence on the total costs incurred, so sub-districts will be geographically grouped again to provide more minimal costs. This market grouping is based on the distance of each sub-districts to six warehouses. Each sub-district has the possibility to be served by those six warehouses, but sub-districts that are closer to a warehouse will provide smaller shipping distance and will reduce the outbound costs. The smallest distance from warehouse to each sub-district (market location) will be selected.
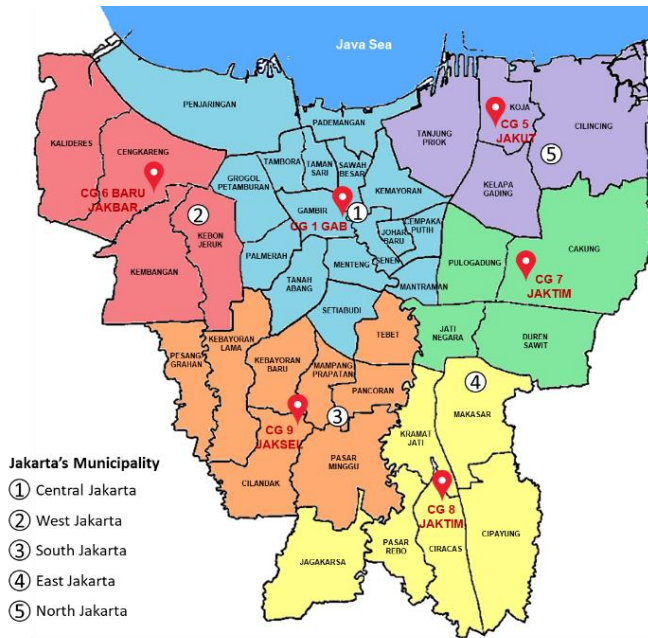
Fig. 3. Market division for six warehouses

Calculation of logistics costs for six warehouses with their new market division will be carried out again (Fig. 4). The division of market areas causes a decrease in the total costs. The total logistics costs from six warehouses with their own market division are IDR 291 580 079 for every 2 mo. There is 16 % decrease in the total logistics costs between the initial market area and the new market division of six warehouses.

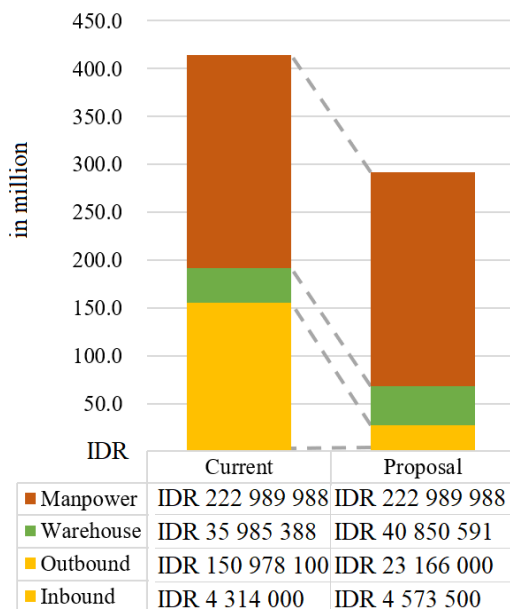|	Initial Market Division	New Market Division
■ Manpower	IDR 257 296 140	IDR 222 989 988
■ Warehouse	IDR 42 720 300	IDR 40 850 591
■ Outbound	IDR 42 993 500	IDR 23 166 000
■ Inbound	IDR 4 614 500	IDR 4 573 500

Fig. 4. Comparison of market areas

All of the four variables costs will decrease, especially the cost of manpower and outbound. The number of workers will decrease from 15 persons to 13 persons because the market division is no longer limited to sub-districts. Therefore, it can maximize the manpower's capacity. Outbound costs incurred also decreased due to the division of market areas that have been done based on the shortest distance.

3.6 Comparison between warehouse's current condition and proposal

Current test market of the new tobacco alternative product uses fur warehouses with their own market areas in Jakarta. The total logistics costs are IDR 414 267 476 for every 2 mo, while the new proposal of number and location of warehouses cost IDR 291 580 079 for every 2 mo. Thus, there is 29.6 % difference between current and proposed conditions.

As shown in Fig. 5, there is a significant reduction in outbound costs. Warehouse costs will increase slightly compared to the initial condition because there are two additional warehouses. There is no difference between manpower costs. Inbound costs are also slightly higher.

	Current	Proposal
■ Manpower	IDR 222 989 988	IDR 222 989 988
■ Warehouse	IDR 35 985 388	IDR 40 850 591
■ Outbound	IDR 150 978 100	IDR 23 166 000
■ Inbound	IDR 4 314 000	IDR 4 573 500

Fig. 5. Comparison between current and proposed conditions

4 Conclusions

It has been discussed the determination of the optimal number and location of warehouses using the Centre of Gravity, where consumer groupings are based on Agglomerative Hierarchical Clustering to get the shortest total distance from each warehouse to the consumer group. As result, the use of six warehouses will provide lower logistics costs compared to its current condition, four warehouses. The location of the six warehouses will be in Cengkareng Timur, Gambir, Rawabadak, Cakung, Ciracas and Cipete areas.

The use of warehouse that is close to consumers will reduce logistics costs. The determination of the warehouse should also be based on population density or the number of demand that exists because both things will give an impact on the outbound transportation costs incurred. Too few warehouses will increase the outbound transportation costs, while too many warehouses will increase labour costs, so the use of six warehouses will provide minimal costs. The results of this study are in line with the last recent studies related to optimal determination of warehouse location and number of warehouses, depending on the costs of inbound and outbound logistics.

This research still has various weaknesses. Future research will have expected to be able to utilize the big data captured on social media, which in more detail describes the real condition, so that the results obtained are more accurate. Authors should use the forms shown in Table 3 in the final reference list.

References

1. A. Rusthon, P. Croucher, P. Baker, *The Handbook of Logistics and Distribution Management*, India: Replika Press (2010). https://www.koganpage.com/product/the-handbook-of-logistics-and-distribution-management-9780749476779
2. S. Tsur, S. Abiteboul, R. Agrawal, U. Daval, J. Klein, G. Weikum, VLDB **1**:614−617(2001). http://www.vldb.org/conf/2001/P614.pdf
3. K. Purcell, *Search and email still top the list of most popular online activities*. [Research report], DC, USA: Pew Research Center (2011). P. 1−15. https://www.pewresearch.org/internet/wp-content/uploads/sites/9/media/Files/Reports/2011/PIP_Search-and-Email.pdf
4. J-P. Schaller, D. Keller, L. Poget, P. Pratte, E. Kaelin, D. McHugh, G. Cudazzo, et al., Regulatory Toxicology and Pharmacology, **81**,2:S27−S47(2016). https://www.sciencedirect.com/science/article/pii/S0273230016302902
5. B. Sainathuni, P.J. Parikh, X. Zhang, N. Kong, European Journal of Operational Research, **237**:690–700(2014). https://www.sciencedirect.com/science/article/pii/S0377221714001222
6. B. Brunaud, M.H. Bassett, A. Agarwal, J.M. Wassick, I.E. Grossmanna, Computers and Chemical Engineering, **111**:311–323(2018). https://www.sciencedirect.com/science/article/pii/S0098135417302107
7. Y. Abo-Elnaga, B. El-Sobky, L. Al-Naser, Journal of Taibah University for Science, **11**:353–358(2017). https://www.tandfonline.com/doi/abs/10.1016/j.jtusci.2016.04.003
8. K.H. Leung, K.L. Choy, P.K.Y. Siu, G.T.S. Ho, H.Y. Lam, C.K.M. Lee, Expert Systems with Applications, **91**:386–401(2018). https://www.sciencedirect.com/science/article/pii/S0957417417306280
9. X. Zhao, Based on gravity method of logistics distribution center location strategy research, *International Conference on Logistics Engineering, Management and Computer Science (LEMCS 2014)* Shenyang, China 2014. p. 584–587. https://www.atlantis-press.com/proceedings/lemcs-14/12283
10. Tan P.N., M. Steinbach, A. Karpatne, V. Kumar, *Introduction to Data Mining*, London: Pearson (2019). https://www-users.cs.umn.edu/~kumar001/dmbook/index.php
11. W.L. Winston, *Microsoft Excel 2019 Data Analysis and Business Modeling*, London: Pearson Education. https://www.pearson.com/us/higher-education/program/Winston-Microsoft-Excel-2019-Data-Analysis-and-Business-Modeling-6th-Edition/PGM2142991.html
12. X. Yu, M. Gen, *Introduction to Evolutionary Algorithms*, London: Springer (2010). https://www.google.com/books?hl=en&lr=&id=rHQf_2Dx2ucC&oi