DEVELOPING AN INTELLIGENT LOGISTICS AND DISTRIBUTION SYSTEM FOR A LARGE NUMBER OF RETAIL OUTLETS: A BIG DATA ANALYTICS APPROACH

Zulkarnain1*, Komarudin1, Fauziah Arofah1, Irvanu Rahman1

1Department of Industrial Engineering Universitas Indonesia, Depok 16424, Indonesia

(Received: January 2020 / Revised: February 2020 / Accepted: May 2020)

ABSTRACT

The logistics and distribution system in the retail industry in Indonesia has its own complexity. The growth and productivity of the retail outlets in Indonesia have been increasing from year to year. Distribution activities in this study are related to the formation of salesman visit routes involving around 38,900 customer base retail outlets, which are quite numerous in number, calling for a challenging approach to find the optimized solution. Therefore, the case study in this research will be discussed on the concept of the Traveling Salesman Problem (TSP) and the Vehicle Routing Problem (VRP), considering the work balance and visit pattern constraints. The methods used in this research are the balanced K-means and the Minimum Span Tree – Kruskal’s Walk algorithm, which are proven to solve the problem with a shorter computation time and a more balanced daily route than the current conditions as the results.

Keyrords: Balanced K-Means; Intelligent Logistics; Optimization; Vehicle Routing Problem

1. INTRODUCTION

The retail industry in Indonesia is one of the most developed businesses. Nielsen Executive Director of Retail Measurement Services, Teguh Yunanto, said that Indonesia is a country with the second-largest retail outlet in the world after India, with the number of outlets reaching more than 2.5 million (Djumena, 2015). Productivity from retail outlets also grew from 2004 to 2008 and is expected to continue to grow. Modern and traditional retails are expected to grow up to 17% and up to 10%, respectively (Pandin, 2009).

Retail is the last axis of the goods distribution process. The retail industry can be divided into two categories: modern retail (modern trade) and traditional retail (general trade). The modern retailer is a large retailer, which generally has a sufficient number of stores with complete and modern facilities. In contrast, traditional retailers are more straightforward retailers with less capital than modern retailers (Soliha, 2008). On this occasion, the study focuses on the distribution of traditional retail outlets, comprising wholesalers, stores, and small shops selling personal needs products.

*Corresponding author’s email: zulkarnain@ie.ui.ac.id, Tel. +62-21-78888805
DOI: https://doi.org/10.32783/csid-jid.v3i1.106
The route is the sequence of nodes/customers to be visited (Hernandez, Gendreau, & Potvin, 2017). The optimal route is expected to minimize cost and time. Generating retail outlet visit routes in Indonesia is very complex. Generally, the route to visit retail outlets in Indonesia has a visit pattern that repeats every two weeks. When a retail outlet can contribute significantly to the turnover, it is usually categorized as a big wholesale, and it will be visited once a week on the same weekday.

On the other hand, the other retail outlet categories will be visited once every two weeks regularly. The pattern is designed by within a year, there are 52 weeks, weeks 1, 3, 5, and so on are called odd weeks, while the 2nd, 4th, 6th, and so on are called even weeks. A two-weekly retail outlet can be visited either on the even or the odd weeks, while a once-weekly retail outlet will be visited only on an odd and an even week alone. In one visit, the visited retail outlet is an outlet that is either adjacent to or affordable within a single trip, thereby saving the operational costs. Thus, it is necessary to establish a classified route based on the salesman allocation, the active salesman's day, and the route pattern of visits based on the Vehicle Routing Problem (VRP) concept.

Until now, the known largest number of the object used as research material for VRP is 20,000 data (Kytöjoki, Nuortio, Bräysy, & Gendreau, 2007). Previous research suggested a future research work by increasing the number of vertices and adding constraints based on real-world cases (Laporte, Gendreau, Potvin, & Semet, 2000; Mohammed et al., 2017). This is a consideration in the selection of case studies that will be examined. The case study consists of a cluster of big data outlet visits (38,900 customer base), which contains information about names of salespeople, schedule of regular visits outlet, outlet locations, as well as odd-even visit patterns influenced by outlet channel. All customer outlets involved in this study are situated in one of the big cities in the most densely populated island in Indonesia, namely Java island. There are not any outlets located across the different islands, and the distribution activities are operated by using road transport modes.

The dataset handled in this study is acknowledged as big data, since it involved 38,900 rows of customer data and more than four different attributes (columns of data) as customer nodes, to be optimally scheduled for visitation by 84 different salespeople. For a Traveling Salesman Problem (TSP) or a Vehicle Routing Problem (VRP), finding optimum routes that can connect this amount of nodes could be challenging, as previous studies usually have utilized 500-20,000 nodes, even most of them are less than that. Therefore, a study on completion of VRP for these visits with clustering and heuristic methods on a cluster of big data needs to be done. This study takes the form of symmetrical problems for the route of a visit by excluding outlet channels, time windows, and quantitative information.

2. LITERATURE STUDY

2.1. Travelling Salesman Problem (TSP) and Vehicle Routing Problem (VRP)

In 1800, an Irish mathematician, William Rowan Hamilton, and a British mathematician, Thomas Penyngton, introduced the concept of Traveling Salesman Problem (TSP). TSP is one of the classical optimization problems that have more than one combination. TSP concept is applied in the formation of a route, which has a depot with several destination nodes to be visited on a trip, where the destination node can only be visited once. These problems are generally represented in the graph to facilitate its completion. In the establishment of an optimal route, a strategy is needed to decide which route will be selected and used. This is because there is more than one solution that can be used to visit all the vertices in one trip. VRP is an extension of the TSP problem, which was introduced in 1959 by Dantzig and Ramzer (Besbes,
Developing an Intelligent Logistics and Distribution System
For A Large Number of Retail Outlets: A Big Data Analytics Approach

Dhouib, Wassan, & Marrekchi, 2019). The characteristic that distinguishes VRP from TSP is the introduction of capacity on the vehicle (Carlsson & Behroozi, 2017). VRP is a combinatorial optimization problem that is difficult to solve (Groër, Golden, & Wasil, 2010), and it is an optimization problem that can be applied to various fields of distribution, collection, logistics, and others (Pop et al., 2011). Therefore, VRP is an operational research issue, where the goal is to design the route at the minimum cost for a fleet of vehicles with the same capacity to serve customers who are geographically dispersed (Du & He, 2012).

VRP often applies to goods delivery activities or route visits. Components related to VRP are generally customers, depots, vehicles, and service vehicles. The primary purpose of VRP is to get an optimal route to minimize mileage, minimize cost, or balance routes. In contrast to TSP, VRPs need to generate at least one or more routes based on vehicle capacity or certain constraints. When a route cannot be facilitated by the capacity of the vehicle or other specified constraints, it will form the next route. Each route must start and finish at the same depot point, and all node points must be visited.

In order to solve a large-scale VRP, a good approach is required. When there is a small number of destination nodes, an approach by enumeration or branch and bound method can be used. However, when the number of destination nodes is enormous, the enumeration and branch and bound methods will be difficult to apply. Methods for solving large scale VRPs are usually considered into two main classes, namely: heuristics and metaheuristics (Laporte, Gendreau, Potvin, & Semet, 2000). Some examples of heuristic methods such as Nearest Neighbor, Saving Methods, and Minimum Spanning Tree (MST). Besides, for a vast number of points, it will be easier to use the clustering method before they are resolved by the heuristics method. This approach can increase the chance of finding the optimal solution with reasonably efficient computational time.

2.2. Route Generation Using Minimum Spanning Tree (MST) Approach

MST is a tree from a graph that has the minimum total weights. MST solution can be obtained by using Kruskal’s algorithm. Then, a route (TSP solution) can be constructed from the MST solution by replacing edges from nodes with degrees more than one with other edges. The replacement edges are selected so that the total weights are minimized, and the graph is connected.

Kruskal’s algorithm for MST starts by sequencing edges from the smallest weight to the largest weight. The next step is to select the most minimum edge as the initial sub tour. Then, it repeatedly adds a node with the smallest weights to the sub tour until no nodes left. Solving a problem using Kruskal's algorithm is illustrated in Figure 1, assuming point D is the center point.

![Figure 1. Completion of MST Using Kruskal's Algorithm](image-url)
In Figure 1, the second point is chosen, i.e., point A since it has the smallest distance with D. The next point to be selected is the one that has the smallest distance to point D or point A. Then, point B is selected and forms a D-B branch because B has the smallest distance from D if compared to the distance of B to A or the distance E from D. The procedure to select the next point follows the same procedure until all the points are connected.

![Figure 2. Completion of MST Preorder Walk from Kruskal's Algorithm](image)

Figure 2. Completion of MST Preorder Walk from Kruskal's Algorithm

The solution formed from Figure 1 is a tree and it is not necessarily a route solution because some nodes may need to be passed more than once to get to the depot. The resulted solution is D-A-D-B-C-B-E-B-D, with a total distance = 52. So, a preorder walk algorithm is needed to fix it, as illustrated in Figure 2. After performing the preorder walk, the resulted route is D-A-B-C-E-D, with total distance = 44. The steps performed as in Figure 1 and Figure 2 generate a solution to TSP, whereas can be modified to obtain a VRP solution by adding constraints.

### 2.3. Clustering Method

The clustering method is a method for categorizing data based on data similarity. In its application, this method is usually used to identify a group that does not necessarily have members with identical characteristics but has members with similar characteristics. If the data has information about the geographical location of a point, then this method can classify it based on proximity to the location of other near points. In the route formation case, clustering data based on point position can make the route generation efficient. It happens because this method can avoid the situation when a subsequent search point is feasible but located very far, which impact to extend more routes to check. The clustering method can produce a faster computation time (Mohammed et al., 2017).

### 2.4. Balanced K-means

When using the K-means method, the existence of outliers can generate poor results. This is a common problem that happens almost in every method of data modeling (Agusta, 2007). The outlier is an observation that is too different from most observations (Cousineau & Chartier, 2010). In the K-means method itself, the outlier becomes a critical factor for the final cluster result. In this study, the outlier data are characterized by the position of the outlets.

The K-means method assigns data to be part of a cluster. The characteristics of a cluster can change as data are assigned to a more suitable cluster. No mechanism prevents K-means to generate unbalanced clusters. Therefore, it needs some modifications to the K-means method so that the cluster to become more balanced. The incorporation of a balancing method does not necessarily result in the same number of customers of each cluster being formed. Nevertheless, the resulted number of customers in each cluster would not be too different.

The purpose of cluster balancing is to assign the same workload amongst salespeople, i.e., on the number of outlets handled by each salesperson. The size of the workload given to workers will affect the performance of employees (Shah et al., 2011). The excessive workload can cause stress and suboptimal employees at work, while too little workload can also cause losses to the company. The workload can be used to determine the number of salespeople required to
Developing an Intelligent Logistics and Distribution System
For A Large Number of Retail Outlets: A Big Data Analytics Approach

The topic of balancing workloads in distribution and routes is interesting to be developed. Workload balancing can minimize travel costs (Naderi & Kilic, 2016), minimize vehicle mileage and balance the vehicle time utilization (Nikolakopoulou, Kortesis, Synefaki, & Kalfakakou, 2004), maintain a balance between transportation cost and service level (Zhou, Min, & Gen, 2002), and produce fair schedules (Decerle, Grunder, Hajjam El Hassani, & Barakat, 2017; Razzazi & Esmaeeli, 2014).

The balanced K-means method is a modification of the basic method of K-means. This method will balance the number of members in each cluster. The proposed balanced K-means algorithm is as follows:

1. Determine the desired number of clusters.
2. Determine the amount of capacity of each cluster member formed by dividing a total number of data (nodes) by the desired number of clusters ± to 10%. A tolerance of 10% is set in this study to facilitate the results and speed up the calculation process.
3. Select the centroid or the center point at random.
4. Allocate each data to the nearest centroid while the requirement of the capacity of cluster members of the centroid is still adequate.
5. Calculate the average of data in each cluster and make the average as a new centroid or center point.
6. Allocate each data to the nearest centroid or new center point if cluster members' capacity of the centroid is still adequate.
7. If there are still data to move the cluster or there are changes in the value of centroid compared to previous results, then go back to step 5. Otherwise, the algorithm will stop the mechanism due to each cluster has reached a balanced or homogenous number of data (nodes) to be handled.

3. METHODS AND EXPERIMENTS

In this study, the data of 39,066 customer base (retail outlets) of a company producing and distributing personal needs are used. The customer base consists of several customers who will be visited within two weeks, and it contains complete information about customers (retail outlets). However, the information that will be used in this study focus on location (Latitude and Longitude), customer ID, traffic patterns, and salesman ID.

The first step taken is an initial analysis that aims to get an overview of the current situation. In this step, the customer base data with a coordinate point (0,0) are eliminated because it does not represent correct data. Secondly, a customer base that has a double data record with two different salespeople are also eliminated, so that 38,900 customer base data is obtained, and 84 salespeople are identified. Among those selected customers, 38,537 outlets have a two-weekly visit pattern, and the other 253 outlets have a once-a-week visit pattern. The data then are used as a basis for generating optimal route schedule visits using the clustering and heuristic methods.

This study generates the route of every salesperson for each day by considering the existing visit pattern. Java programming is used as a computational tool. The existence of the visit pattern makes the scoring based on the consumer ID needs to be done. Thus, the clustering process for outlets with a visit pattern once a week can be visited on the same day. The study performs scoring with the removal of duplicate data by consumer’s ID and provides a score, as shown in Table 1. After the scoring, clustering outlets based on geographical location using the balanced K-means is conducted. The clustering is performed in three steps.
First, clustering is to classify the outlet point in which an area of each salesperson is assigned. At this stage, the clustering is done with a value of K=84, according to the number of salespeople registered in the customer base. The second clustering is done to clustering outlet point by salesperson working days (Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday) so that the value of K = 6. The second clustering is done within each result of the first clustering. The third clustering is done in second clustering results aimed at visits (odd and even) patterns.

In contrast to the first clustering and the second one, the third clustering is done in two stages. The first stage is to separate consumers ID with scores equals to 2, and consumers ID with a score of 1. Consumer ID with a score of 2 will automatically become a member of the cluster of odd and even, while for consumer ID with score 1 will be clustered in the same way as the first and the second clustering, with K = 2 for odd and even. After that, the even cluster on consumer ID with score 2 will be combined with the even cluster result on consumer ID with score 1, as well as an odd cluster.

Table 1. Scoring process based on consumer id on the customer base

| ID Customer | ID Customer | Score |
|-------------|-------------|-------|
| 600000643   | 600007829   | 1     |
| 600000660   | 600007830   | 1     |
| 600000661   | 600000053   | 1     |
| 600007822   | 600007830   | 2     |
| 600007822   | 600007831   | 1     |
| 600009062   | 600007831   | 2     |
| 600007827   | 600000059   | 2     |
| 600007827   | 600007832   | 2     |
| 600007820   | 600007832   | 1     |

The next step after clustering is the route forming. The method used is the MST method with Kruskal’s walk preorder algorithm. The MST method is a fast and high-quality heuristic method in the total Euclidean distance. In this study, 1,008 daily routes for two weeks for 84 salespeople are obtained. From those routes, the distance of each route and the total distance are calculated, which aims to compare with the existing route at the current customer base condition.

4. RESULTS AND DISCUSSION

In this study, the developed program has been run and performed ten times for the same case study data. It aims to provide an overview of the results of the running program. Table 2 shows the running program results for a total distance route formed along with time calculated in seconds. There are ten different results for every running program. This is due to the random initiation of K-means clustering method within the program. K-means is a standard method used as a clustering tool.

Random initiation on K-means can indeed produce several different cluster models on each calculation. However, the K-means initiation process randomly tends to produce better results and independent compared with setting an initial of K-means (Peña, Lozano, & Larrañaga, 1999). The process of finding a solution on K-means will stop when the centroid and members of the iteration cluster are running the same as the centroid and members in the previous iteration cluster. Random initiation will significantly affect the K-means scenario to find a
solution. This is the reason why the total distance and time calculations in Table 2 for each running program may vary.

Table 2. Total route distance and program calculation time

| Running Results | Route Distance | Calculation Time (s) |
|-----------------|----------------|---------------------|
| 1               | 415.09         | 97.24               |
| 2               | 412.73         | 138.00              |
| 3               | 411.40         | 127.36              |
| 4               | 413.12         | 154.59              |
| 5               | 412.33         | 128.61              |
| 6               | 410.14         | 109.51              |
| 7               | 411.85         | 232.59              |
| 8               | 413.99         | 6.10                |
| 9               | 410.49         | 136.80              |
| 10              | 415.20         | 120.62              |

If the result of the total distance in Table 2 is compared, the total distance is formed not too much different on every run. It indicates that the result of the established program has relatively homogenous (the standard deviation of the total distance generated is 1.76). The smallest total distance in Table 2 occurs in the sixth experiment, which is 410.14. Therefore, the 6th program run will be taken as the best result and used as a comparison with the existing route.

4.1. Clustering Results

The existing clustering conditions are found to be unbalanced. The balance of customers number in each cluster is essential in maintaining the performance of salespeople. The balanced K-means will improve the current condition. The proposed method can make the number of customers in each cluster is more balanced when compared to current conditions. In current conditions, there are 66 variations of the total number of customers or outlets (x-axis) for each salesperson to handle, comprising a range between 241 and 546 customers (Figure 3a). In contrast to the new conditions, every salesperson tends to handle a more evenly distributed number of customers, i.e., 463 outlets for every 82 salespeople (Figure 3c). Only two salespeople who serve 464 and 470 customers consecutively. Y-axis in both Figure 3a and Figure 3c shows the number of salespeople who are responsible for handling a number of outlets (x-axis).

The clustering of the customer base for each salesperson also significantly affects the number of outlets to be visited daily by a salesperson. A considerable variation (unbalanced) number of customers whom a salesperson visits daily present in the current condition, ranging from 0 to 63 customers visited daily (Figure 3b). In contrast, the proposed clustering results make the number of outlets to be visited daily be more balanced for each salesperson, i.e., around 36 to 41 customers visited daily (Figure 3d). Y-axis in Figure 3b and Figure 3d shows the number of events when any number of customers are visited daily by a salesperson (x-axis). Figure 3 shows the comparison of the work balance in terms of the total number of customers handled by each salesperson (Figure 3a and 3c) and the number of outlets to be visited daily by each salesperson (Figure 3b and 3d), between the existing and the proposed conditions.

4.2. Routing Results

The route forming is conducted after the clustering solution is obtained. In current conditions, the formation of these salespeople visit has not been done, so that with the adoption of the MST Kruskal’s Walk, 1,008 solutions are obtained. A scatter diagram is intended to describe the outlier routes and the existing outlets. Figure 4a shows a scatter diagram for the current condition, while Figure 4b is a scatter diagram for the routes formed by the proposed algorithm.

The number of destination nodes as an object formed in this research makes these hard to be seen in Figures 4a and 4b. However, there is a point where several suspected outliers are found in the customer base.
a) unbalanced total customers handled (x-axis) by each salesperson (existing); b) unbalanced number of customers visited daily (x-axis) by each salesperson (existing); c) more balanced number of customers handled (x-axis); d) more balanced customers visited daily (x-axis)

Figure. 3. Salesperson workload distribution
An outlier point on the customer base can extend the distance that will increase the transportation costs, so the existence of the outlier point becomes essential to be of awareness. When the outlet can provide a significant turnover and cover the operating costs, then the outlet can be maintained. Otherwise, it should be redirected to another distributor closer to its position.

From Figure 4a and Figure 4b, there is one outlier point that can be visible to the human eye (red box) with a location at the point (5.58; 108.08). The removal of a one-point outlier is also simulated in this study to compare the calculation of the total distance of the generated routes (Figure 4c and Figure 4d).

**4.3. Total Distance Results**

Based on the Euclidean distance formula, the total distance from the route formed at the current condition is 297.81, while the total distance from the new route (the 6th run) is 410.14. The total distance in the current state is much shorter than the proposed solution. This condition happens because the clustering that exists in the current conditions do not consider the balance of worker capacity on a daily basis. In Figure 3b, there are several significant differences in the number of visits per day. It affects the total distance difference for each daily route formed, which may be less likely to be too far away. The standard deviation value of the total daily route distance under current conditions is 5.42, whereas the standard deviation value of the total daily distance of the newly formed route is 0.23. If the working hours of each salesperson are relatively similar for each day, it can improve the performance of a salesperson. It also will affect the achievement of a salesperson in gaining incentives. Therefore, the establishment of a more balanced route is vital to consider.

It has been previously discussed that a removal scenario of one outlier is simulated. Hence its effect on the total distance of the route is also examined. After the outlier removal, the total distance formed is 294.65 in the current condition. The total distance after the outlier removal is not very significantly changed, as it only differs at 1%, which is quite reasonable because there is only one point omitted in the calculation. In contrast, the new standard deviation (after the outlier removal) is 4.49, which is decreased by 17%. However, on the proposed route, after
running the data without outliers as much as ten times, a minimum value of the total distance of 324.65 can be obtained. The total distance is decreased by 20.84% compared to the total distance that has been previously formed. The standard deviation is also decreased by 37% into 0.15. It proves that the outlier can also affect the balance of the formed route. However, once again, the result depends very much on the value of random initiation on K-means in the program. Therefore, it is necessary to develop a method for K-means initiation in the future.

The total distance from the route can also be affected by the center. In this study, there are 38,900 points of destination with one central point (center of gravity). As indicated by the number of destination points, the addition of a central point and making it as a decentralized type of distribution will likely shorten the total distance and minimize the standard deviation value. This finding can be used as a material for further research.

5. CONCLUSION

From this study, several remarks can be concluded. The study proves that the combination of the K-means and the MST-Walk balanced methods can generate more balanced routes with faster computing time for large amounts of data. The computation time is shorter since the developed algorithm will be efficiently limiting the search area for finding optimum solutions in a generated cluster, instead of checking the combination of all nodes as solution candidates one by one. From ten runs of the program, the minimum total Euclidean distance generated from the proposed route is 410.14 (with the standard deviation of 1.76 for those ten runs). The total route distance at the current condition is 297.81, shorter than the proposed total route distance. However, this happens because it has not taken the balanced amount of cluster members for each route into account. The current standard deviation for the daily route is quite high, i.e., 5.42, while the newly proposed standard deviation of the daily route is very much lower at 0.23. The proposed condition has a more balanced optimal route than the existing condition. Another insight from this study is the finding that several aspects could affect the total distance of the route, i.e., the position of center and the position of goal, as well as the capacity of each cluster to be formed.

For further research, an initial solution method can be developed to obtain more optimal K-means results. Finding solutions using metaheuristic methods could also be interesting to be studied. Adding the other constraints that are appropriate to the actual circumstances (e.g., quantitative information and time windows) to improve the optimal route solution is worth it to put on the next agenda. Following up the research with a decentralized distribution system would also be a good idea.

Finally, this study might not provide the best solution to the generated routes in terms of their total distance. However, the main contribution of the paper is to find near-optimum solutions, without violating the workload balance constraint amongst the salespeople, with reasonably efficient computation time, even though a large number of customer nodes are involved as the object of the study.

REFERENCES

Agusta, Y. (2007). K-Means-Penerapan, Permasalahan dan Metode Terkait. Jurnal Sistem Dan Informatika, 3(1), 47–60.
Besbes, W., Dhouib, D., Wassan, N., & Marrekchi, E. (2019). Solving Transport Problems: Towards Green Logistics. John Wiley & Sons.
Carlsson, J. G., & Behroozi, M. (2017). Worst-case demand distributions in vehicle routing. European Journal of Operational Research, 256(2), 462–472. https://doi.org/10.1016/j.ejor.2016.03.047
Developing an Intelligent Logistics and Distribution System for a Large Number of Retail Outlets: A Big Data Analytics Approach

Cousineau, D., & Chartier, S. (2010). Outliers detection and treatment: a review. *International Journal of Psychological Research, 3*(1), 58–67. https://doi.org/10.21500/20112084.844

Decerle, J., Grunder, O., Hajjam El Hassani, A., & Barakat, O. (2017). A general model for the home health care routing and scheduling problem with route balancing. *IFAC-PapersOnLine, 50*(1), 14662–14667. https://doi.org/10.1016/j.ifacol.2017.08.1907

Djumena, E. (2015). Jumlah Toko di RI Terbesar Kedua Dunia. Retrieved from https://entertainment.kompas.com/read/2011/03/15/12333392/Jumlah.Toko.di.RI.Terbesar.Kedua.Dunia.

Du, L., & He, R. (2012). Combining Nearest Neighbor Search with Tabu Search for Large-Scale Vehicle Routing Problem. *Physics Procedia, 25*, 1536–1546. https://doi.org/10.1016/j.phpro.2012.03.273

Groër, C., Golden, B., & Wasil, E. (2010). A library of local search heuristics for the vehicle routing problem. *Mathematical Programming Computation, 2*(2), 79–101. https://doi.org/10.1007/s12532-010-0013-5

Hernandez, F., Gendreau, M., & Potvin, J. Y. (2017). Heuristics for tactical time slot management: a periodic vehicle routing problem view. *International Transactions in Operational Research, 24*(6), 1233–1252. https://doi.org/10.1111/itor.12403

Kytöjoki, J., Nuortio, T., Bräysy, O., & Gendreau, M. (2007). An efficient variable neighborhood search heuristic for very large scale vehicle routing problems. *Computers and Operations Research, 34*(9), 2743–2757. https://doi.org/10.1016/j.cor.2005.10.010

Laporte, G., Gendreau, M., Potvin, J. Y., & Semet, F. (2000). Classical and modern heuristics for the vehicle routing problem. *International Transactions in Operational Research, 7*(4–5), 285–300. https://doi.org/10.1111/j.1475-3995.2000.tb00200.x

Mohammed, M. A., Abd Ghani, M. K., Hamed, R. I., Mostafa, S. A., Ahmad, M. S., & Ibrahim, D. A. (2017). Solving vehicle routing problem by using improved genetic algorithm for optimal solution. *Journal of Computational Science, 21*, 255–262. https://doi.org/10.1016/j.jocs.2017.04.003

Naderi, S., & Kilic, K. (2016). A Decision Support System for Staff Workload Balancing. *Procedia Computer Science, 67–73*. https://doi.org/10.1016/j.procs.2016.09.371

Nikolakopoulou, G., Kortesis, S., Synefaki, A., & Kalfakakou, R. (2004). Solving a vehicle routing problem by balancing the vehicles time utilization. *European Journal of Operational Research, 152*, 520–527. https://doi.org/10.1016/S0377-2217(03)00042-0

O’Donnell, C. R., & Eggemeier, F. T. (1986). Workload Assessment Methodology: chapter 42. In Boff, K. R., Kaufman, L., Thomas, J. P. (Eds.), *Handbook of Perception and Human Performance, Volume 2* (pp. 1–49).

Pandin, M. L. (2009). Potret Bisnis Ritel di Indonesia: Pasar Modern. *Economic Review, 215*, 1–12. https://doi.org/10.1016/0029-554X(79)90243-X

Pavone, M., Arsie, A., Frazzoli, E., & Bullo, F. (2011). Distributed algorithms for environment partitioning in mobile robotic networks. *IEEE Transactions on Automatic Control, 5*(8), 1834–1848. https://doi.org/10.1109/TAC.2011.2112410

Peña, J. M., Lozano, J. A., & Larrañaga, P. (1999). An empirical comparison of four initialization methods for the K-Means algorithm. *Pattern Recognition Letters, 20*(10), 1027–1040. https://doi.org/10.1016/S0167-8655(99)00069-0

Pop, P. C., Zelina, I., Lupse, V., Sitar, C. P., & Chira, C. (2011). Heuristic algorithms for solving the generalized vehicle routing problem. *International Journal of Computers, Communications and Control, 6*(1), 58–165. https://doi.org/10.15837/ijccc.2011.1.2210

Razzazi, M., & Esmaeeli, A. (2014). Balanced allocation mechanism: An optimal mechanism for multiple keywords sponsored search auctions. *Information Sciences, 262*, 190–214. https://doi.org/10.1016/j.ins.2013.08.033

Shah, S. S. H., Jaffari, A. R., Aziz, J., Ejaz, W., Ul-Haq, I., & Raza, S. N. (2011). Workload
and Performance of Employees. *Interdisciplinary Journal of Contemporary Research In Business, 3*(5), 256–267.

Soliha, E. (2008). Analisis Industri Ritel di Indonesia. *Jurnal Bisnis Dan Ekonomi, 15*(2), 24251.

Zhou, G., Min, H., & Gen, M. (2002). The balanced allocation of customers to multiple distribution centers in the supply chain network: A genetic algorithm approach. *Computers and Industrial Engineering, 43*(1–2), 251–261. https://doi.org/10.1016/S0360-8352(02)00067-0