Management Science Letters 13 (2023) 235–240

Contents lists available at GrowingScience

Management Science Letters

homepage: www.GrowingScience.com/msl

K-means clustering for optimization of spare parts delivery

Kaushik D Ramgude[a*] and N.R. Rajhans[a]

[a]*Department of Manufacturing Engineering and Industrial Management, COEP Tech University, Pune 411005, MS, India*

C H R O N I C L E

Article history:
Received: March 1, 2023
Received in revised format:
March 29 2023
Accepted: June 22, 2023
Available online:
June 22, 2023

Keywords:
Transhipment
Travelling Salesman Problem
Spare Part Delivery
PMPML (Pune Mahanagar Parivahan Mahamandal Ltd.)

A B S T R A C T

Transhipment is an important logistics strategy that helps to improve supply chain efficiency and reduce transportation costs. It enables cargo to be transported to multiple destinations using different modes of transportation, such as ships, trains, trucks, and planes. This can help to reduce the overall transportation time and cost, as well as improve inventory management and distribution. In addition to its use in logistics and transportation, transhipment can also be used in other industries such as manufacturing, where it can be used to transfer raw materials or finished products between different facilities or production lines. This research paper examines the role of transhipment in improving the efficiency of spare part delivery systems to the PMPML depots from central workshop Swargate. PMPML has 12 depots in total (including central workshop). In many industries, the supply chain for spare parts is complex, with multiple suppliers, warehouses, and service centres involved. Transhipment, or the transfer of inventory between locations, can help to reduce lead times and improve inventory availability. In this paper, we analyze the impact of transhipment on key performance metrics such as order fulfilment, inventory turnover, and transportation costs. We also discuss the challenges associated with implementing transhipment in spare part delivery systems, including coordination between different parties, data sharing, and system integration.

© 2023 by the authors; licensee Growing Science, Canada

1. Introduction

Transhipment refers to the process of transferring goods or cargo from one mode of transportation to another, or from one shipment to another, while they are in transit (Orden, 1956). This is usually done at a transfer point such as a port, airport, or warehouse, where the cargo is temporarily stored before being transferred to the next mode of transportation or destination. Sales services for their products for replacement. Apart from OEMs, there are stores which keep spare parts for the equipment or products they use, which is required for repairing and maintenance purposes. These are further classified into central and distribution stores (Hong-Minh et al., 2000).

The data collected is from a bus transport facility in a large metro city. This facility maintains inventory for break down maintenance through its central workshop. The central workshop distributes the spare parts to 11 depots located in different parts of the metropolitan city. These depots forward their requirement to the central stores. The central store collectively receives and places orders for these spare parts and distributes it to the depots as per the requirements.

* Corresponding author.

E-mail address: ramgudekd19.prod@coeptech.ac.in (K. D Ramgude)

© 2023 by the authors; licensee Growing Science, Canada
doi: 10.5267/j.msl.2023.6.004

2. Literature Review

The first article for the Truck dispatching problem was published by Dantzig and Ramser who presented a larger truck dispatching problem, that is referred to as D&R problem and many more publications has been made (Dantzig & Ramser, 1959). The Clark and Wright algorithm is one of the most popular heuristic algorithms in the vehicle routing problem (Clarke & Wright, 1964). Transhipment can be modelled as a travelling salesman problem (TSP) when a company needs to deliver goods from a central warehouse to multiple locations (Murray & Chu, 2015). In this scenario, the TSP involves finding the shortest possible route that visits all the delivery locations and returns to the warehouse. The TSP is a classic optimization problem in which a salesman needs to visit a set of cities and return to the starting city, while minimizing the total distance traveled. The problem is known to be NP-hard, which means that finding the optimal solution for large sets of cities can be computationally expensive. Similarly, in the transhipment scenario, the warehouse is the starting point, and the delivery locations represent the cities. The goal is to find the optimal route that allows the company to deliver the goods to all locations while minimizing the total distance travelled.

To solve the transhipment problem as a TSP, various algorithms can be used, such as the nearest neighbour algorithm (Tsai et al., 2004), the Christofides algorithm (Biggs et al., 1986), or the 2-opt algorithm (Hasegawa et al., 1997). These algorithms aim to find the optimal route by iteratively selecting the best next location to visit based on a set of rules or heuristics. Overall, modelling transhipment as a TSP can help companies optimize their delivery routes, reduce transportation costs, and improve efficiency in their supply chain operations. The bus transport facility has a Central store which provides spare parts as and when required to their respective distribution stores. As the purchasing and procurement of spare parts is done at the central store only, the forecasting is done for the central store inventory.

Many theoretical problems related to TSP were treated in the 1800s by the Irish mathematician Sir William Rowan Hamilton and by the British mathematician Thomas Penyngton Kirkman The general form of the TSP was first studied by mathematicians starting in the 1930s by Karl Menger in Vienna and Harvard (Wu & Yang, 2004). The problem was later promoted by Hassler Whitney and Merrill Flood at Princeton (Biggs et al., 1986). A detailed treatment of the connection between Menger and Whitney, and the growth of the TSP as a topic of study can be found in Alexander Schrijver's paper (Christofides, 1976).

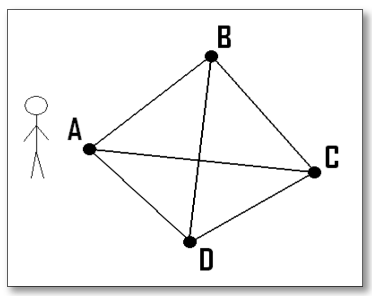

Fig. 1. An overview of the TSP problem

Given the distances between any two points, what is the shortest route a salesman can make from point A, visiting all points, and returning to point A? That could be one of the questions asked for the graph above. In any TSP the best solution would be to try all the ordered combinations to see which one is easiest and shortest. A lot of mathematicians use techniques of dynamic programming (Bellman, 1962; Bouman et al., 2018) to solve the problem exactly in time. Popular techniques are the big O notation which is demonstrated as $O(2^n)$. In 1954, the most important progress on the TSP was made by Rand researchers (Schrijver, 2005) when they developed a new method for solving the TSP, the Cutting Plane method, which became a prototype in linear programming (Hoffman & Wolfe, 1985; Christofides, 1976).

3. Methodology

3.1 Existing Spare Parts Delivery system

The Delivery problems faced were:

- PMPML's fleet has more than 800 active buses on road. A significant number of breakdowns is seen on a daily basis across all the 11 depots.
- For every bus breakdown there is a considerable number of losses that the organization have to incur.
- Spare parts are often required in order to repair/ maintain the different sub systems of a bus. The respective depots send their requirements of the required spare parts to the Central stores at Swargate. The Central Store then issues the parts to the respective depots as per the available stock or places an order for the same to the concerned spare parts supplier.

- A maintenance bus from each depot visited the central stores for receiving these parts. Thus, the total time consumed in the whole process starting from the paperwork until the actually dispatching of all the parts required by the 11 depots was reducing the office as well as depot's productivity.
- Also, the total distance travelled by all the maintenance buses was approximately 258 kilometres. This was accompanied with other costs such as fuel costs, overhead costs, salaries , etc.
- Table 1 represents the distances of respective depots from the central stores at Swargate. Some of the depots are located far away from the centre while some of them are in relatively closer vicinity.
- Calculating the total distance travelled (round trip) , it is approximately 258 km every day.

Table 1
Distances from the Central workshop.

Depot Name	Distance from central Workshop (Km)	Latitude	Longitude
Balewadi	17	18.5745	73.76544
Bhosari	19	18.62453	73.85133
Hadapsar	10	18.50054	73.93842
Katraj	5	18.45573	73.85671
Kothrud	9	18.5062	73.79496
Marketyard	4	18.48363	73.87108
N.T. Wadi	5	18.53384	73.84801
Nigdi	23	18.66304	73.76989
Pimpri	18	18.63543	73.81249
Pune Station	5	18.52672	73.87035
Shewalwadi	14	18.49544	73.9714

To reduce the number of buses plying between the central workshop and depot we need to cluster the depots. If we reduce the number of buses, it will also reduce the travel time, processing time, waiting time at respective depot levels. These problems can be solved by application of K-means clustering algorithm.

3.2 K-means Clustering

Application of K- means clustering would drastically reduce the distances travelled as well as the time required to deliver the spare parts to respective depots. In this case we had only one centre for clustering that was Central Stores (Swargate) and the depots to be served were 11. In order to calculate the number of clusters required we used the elbow method. The elbow method is a technique used to determine the optimal number of clusters in a data set for unsupervised learning algorithms such as K-means clustering. The elbow method is named after the point at which the plot of the within-cluster sum of squares (WCSS) against the number of clusters resembles an elbow shape (K-means clustering, 2023). The elbow method involves the following steps:

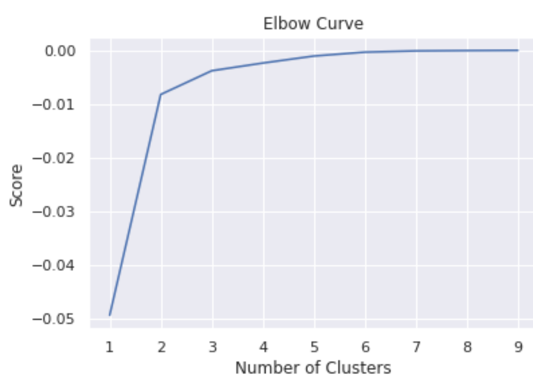

Fig. 2. Elbow Curve for optimal number of Clusters

1. Compute the sum of squared distances between each data point and its closest centroid (WCSS) for different values of k (number of clusters).
2. Plot the WCSS values against the number of clusters (k).
3. Look for the point where the decrease in WCSS starts to level off and the graph forms an elbow shape. This point is the optimal number of clusters.

The intuition behind the elbow method is that adding more clusters will reduce the WCSS, but at some point, the improvement in WCSS will be marginal, and adding more clusters will not lead to significant improvement in clustering quality. This is reflected in the elbow shape of the plot, where the WCSS decreases rapidly initially and then levels off (Elbow method clustering, 2023). Table 1 indicates the latitude and longitudinal coordinates of the respective depots. Using these coordinates, we calculate the WCSS score which is then plotted in figure 2.From figure 2 it is evident that the SSE score has negligible variations after k=3 (three clusters). Hence, we conclude that the optimum number of clusters required is 3 in this case.

4. Using Python for K-means clustering

In Python, we can use the scikit-learn library to perform K-means clustering on a dataset. We used python as it has an inbuilt library and user-friendly GUI. Accordingly, we generated a table indicating the clustering as well as the graph showing the colour coded depots that belonged to the same group. In python we can get the output in the form of both scatter plots as well as a table, the scatter plot is a comparatively better representation of the results. After getting the output we had to scale the data using the Scale (Minmax) function in Python. We have used Google Collab for running the Code and generating the outputs. Table 2 is the output generated which groups the depots according to the least WCSS from the central depot (Swargate.) Fig. 3 is a scatter plot of latitude against longitude which shows the newly formed clusters.

Table 2

Clusters generated using Python.

Sr. No	Latitude	Longitude	Depot	Cluster
0	18.50002	73.86122	Central workshop	1
1	18.52672	73.87035	Pune station	1
2	18.45573	73.85671	Katraj	1
3	18.50054	73.93842	Hadapsar	2
4	18.49544	73.9714	Shewalwadi Bus Stand	2
5	18.5745	73.76544	Balewadi Depot	0
6	18.63543	73.81249	Pimpri bus depot	0
7	18.5062	73.79496	Kothrud Depot	1
8	18.62453	73.85133	Bhosari Bus Depot	0
9	18.48363	73.87108	MARKET YARD BUS DEPO	1
10	18.53384	73.84801	P.M.P. Bus Depot	1
11	18.66304	73.76989	Nigdi Bus Depot	0

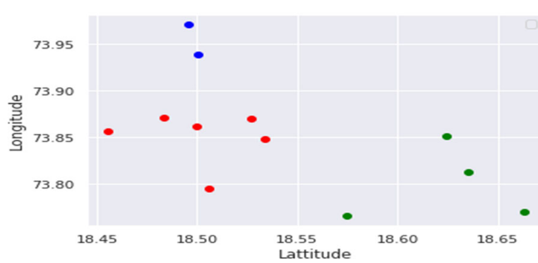

Fig. 3. Plot generated from Python with colour coded clusters.

5. Conclusion

After calculating the optimum number of clusters needed and generating the new clusters, we were able to calculate the tangible savings that PMPML could earn if this model was applied in their spare parts distribution system. We calculated the Before clustering distance covered by the trucks for spare parts delivery and the after clustering distance covered for the same data. We used the items issued data from PMPML's CargoFL system. The data used to calculate the savings was from January 2022 to October 2022.

USING CNG TRUCK	Before Clustering	After Clustering	Total savings
Total Distance covered in 166 days (Km)	42828	4757	38071
Total Fuel cost (₹)	386780	42819	₹ 3,43,961.00

Fig. 4. Indicative Savings using CNG Truck

USING DIESEL BUS	Before Clustering	After Clustering	Total savings
Total Distance covered in 166 days (Km)	42828	4757	38071
Total Fuel cost (₹)	1327668	147467	₹ 11,80,201.00

Fig. 5. Indicative Savings using Diesel Bus

A proposal of using CNG Trucks was made which projected a good amount of savings as indicated in Fig. 4. Whereas if the existing Diesel Buses are used we get more savings as indicated in the Fig. 5.

PMPML can save up to 10 million rupees per annum just by using this model of k-means clustering.

Apart from this , there are several tangible and intangible benefits such as

- Reduced Paperwork
- Less time needed in processing requests from each depot.
- Managing inventories at the depot level will be easier.
- Waiting time of vehicles will be reduced drastically at the central stores.
- Improved Productivity.
- Overhead costs of all vehicles will be reduced.
- Trucks/Buses will be loaded with maximum load capacities.
- Less number of trucks/ buses will be required.

Acknowledgements

We would like to thank the Department of Manufacturing Engineering and Industrial Management and COEP Tech University for providing us with the opportunity to work on this project. We would also like to thank Pune Mahanagar Parivahan Mahamandal Ltd. (PMPML) for providing us the data for our study.

References

Bellman, R. (1962). Dynamic programming treatment of the travelling salesman problem. *Journal of the ACM (JACM)*, *9*(1), 61-63.

Bouman, P., Agatz, N., & Schmidt, M. (2018). Dynamic programming approaches for the traveling salesman problem with drone. *Networks*, *72*(4), 528-542.

Biggs, N., Lloyd, E. K., & Wilson, R. J. (1986). *Graph Theory*, 1736-1936. Clarendon Press, USA.

Christofides, N. (1976). Worst-case analysis of a new heuristic for the travelling salesman problem. Carnegie-Mellon Univ Pittsburgh Pa Management Sciences Research Group.

Clarke, G., & Wright, J.R. (1964). Scheduling of Vehicle Routing Problem from a Central Depot to a Number of Delivery Points. *Operations Research, 12*, 568-581.http://dx.doi.org/10.1287/opre.12.4.568

Dantzig, G. B., & Ramser, J. H. (1959). The Truck Dispatching Problem. *Management Science, 6*(1), 80–91. http://www.jstor.org/stable/2627477

Elbow method (clustering) (2023) Wikipedia. Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/Elbow_method_(clustering) (Accessed: April 15, 2023).

Hasegawa, M., Ikeguchi, T., & Aihara, K. (1997). Combination of chaotic neurodynamics with the 2-opt algorithm to solve traveling salesman problems. *Physical Review Letters*, *79*(12), 2344.

Hoffman, K. L., Padberg, M., & Rinaldi, G. (2013). Traveling salesman problem. In S.I Gass & M.C. Fu (Eds.), *Encyclopedia of operations research and management science* (pp. 1573–1578). Boston: Springer. http://dx.doi.org/10.1007/978-1-4419-1153-7_1068

Hong-Minh, S. M., Disney, S. M., & Naim, M. M. (2000). The dynamics of emergency transhipment supply chains. *International Journal of Physical Distribution & Logistics Management*, *30*(9), 788-816.

K-means clustering (2023) Wikipedia. Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/K-means_clustering (Accessed: April 15, 2023).

Murray, C. C., & Chu, A. G. (2015). The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery. *Transportation Research Part C: Emerging Technologies*, *54*, 86-109.

Orden, A. (1956). The transhipment problem. *Management science*, *2*(3), 276-285.

Schrijver, A. (2005). On the History of Combinatorial Optimization (Till 1960).

Steinley, D. (2006). K-means clustering: a half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 1-34. https://doi.org/10.1348/000711005X48266

Tsai, C. F., Tsai, C. W., & Tseng, C. C. (2004). A new hybrid heuristic approach for solving large traveling salesman problem. *Information Sciences*, *166*(1-4), 67-81.

Wu, H., & Yang, Y. (2004, December). Application of continuous hopfield network to solve the TSP. In *ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, 2004.* (Vol. 3, pp. 2258-2263). IEEE.

© 2023 by the authors; licensee Growing Science, Canada. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).