Malaysian Journal of Mathematical Sciences 14(S) December : 15–29 (2020)
Special Issue: 2nd International Conference on Applied & Industrial Mathematics and Statistics 2019 (ICoAIMS 2019)

MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES

Journal homepage: http://einspem.upm.edu.my/journal

Routing Mail Delivery from a Single Depot with Multiple Delivery Agents

Rahman, A.*, Tan, H. I., Liew, W. and Shahruddin, N. S.

School of Mathematical Sciences, Universiti Sains Malaysia, Malaysia

E-mail: amirahr@usm.my
* *Corresponding author*

Received: 15 January 2020
Accepted: 30 July 2020

ABSTRACT

Optimising the distance covered during mail delivery by utilising multiple Travelling Salesman Problem (mTSP) models and techniques can simultaneously minimise operations cost and Greenhouse Gasses (GHG) emissions caused by the transportation part of the mail delivery process. The aim of this study is to solve a mail delivery problem with multiple delivery agents in an established mail delivery network and compare our optimised delivery route with the original delivery route. This study also aims to optimise the delivery route for different numbers of delivery agents to prepare for cases of employee shortage and vehicle breakdown. This study decomposes the mTSP into multiple separate TSPs by clustering nodes together using the k-Means Clustering method and assigning a delivery agent to each cluster, subsequently solving each cluster of TSP using Genetic Algorithm. The results show that our method is able to provide a better route for the mail delivery system compared to the original route, with a reduction of 11.25% in distance, which would no doubt reduce the transportation-related GHG emissions.

Keywords: Multiple Travelling Salesman Problem, Travelling Salesman Problem, Genetic Algorithm, k-Means Clustering.

1. Introduction

The mail delivery service is the final distribution process of delivering mail from a single distribution centre, or depot, to receivers in multiple locations. The sustainability aspect of delivery operations such as this cannot be disregarded as it could affect operations in the long term. The inability to properly plan and optimise delivery routes will not only lead to extra operational costs, but also contributes significantly to air pollution. Thus, route planning is significantly important in reducing fuel consumption which would then lead to lower operational costs and lesser harmful gas emissions. Consequently, this may result in a better brand reputation as stated by Perboli and Rosano (2019).

The mail delivery service frequently utilises multiple deliverymen or postmen as delivery agents. The delivery agents depart from the depot, deliver mail to different sets of receiver locations, and return to the depot when done. Each receiver can only accept mail from exactly one delivery agent. Thus, the problem is to optimise the routes of the delivery agents such that the distance travelled is minimised, which can be accomplished with the help of mathematical optimisation techniques.

A mail delivery problem with a single delivery agent can be modelled as a Travelling Salesman Problem (TSP). In general, TSP is a problem of finding the route with minimum distance in a network of nodes and edges, where all nodes need to be visited once except for the starting node which is visited twice. In a mail delivery problem, the nodes represent the single depot and multiple mail receiver locations, while the edges that connect a pair of nodes represents the road (or roads) connecting the respective locations.

Real-world mail delivery problems rarely involve only one delivery agent. A mail delivery problem with multiple delivery agents can be modelled as a Multiple Travelling Salesman Problem (mTSP). In general, mTSP is a problem of finding the route with minimum distance in a network of nodes and edges with multiple delivery agents, whereby each node must be visited exactly once by exactly one delivery agent, except for the starting node which is visited twice by all delivery agents.

Although,the mTSP is much more practical in real world applications, the study by Bektas (2006) reported that the mTSP is much less studied compared to the TSP. Both exact and approximation methods have been explored in solving mTSPs. In 1980, Laporte and Nobert (1980) used exact methods with some relaxations of the mTSP constraints. In 1986, Gavish and Srikanth (1986) attempted to solve first large-scale symmetric mTSP using the Branch-and-Bound method. In 1992, another exact method based on Quasi-Assignment (QA) relaxation was proposed by Gromicho et al. (1992) to solve the mTSP. The QA relaxation was obtained by relaxing the subtour elimination constraints.

More recently, Latah (2016) utilised the k-Means Clustering method with Ant Colony Optimization algorithm to solve the mTSP. In a similar vein, Lu et al. (2016) and Xu et al. (2018) presented the combination of the k-Means Clustering method with Genetic Algorithm to solve the mTSP. These three studies used the same framework of first clustering their networks into k smaller networks (where k is the number of salesmen), then solving each cluster independently using their stated metaheuristic methods. Lu et al. (2016) reported that using the k-Means Clustering method to group the nodes together before solving each cluster independently avoids the issue of path intersection among delivery agents and the usage of k-Means Clustering with Genetic Algorithm to solve mTSP leads to better result than solving it solely by Genetic Algorithm.

In this study, we attempt to solve the problem of delivering mail over an established mail delivery network with a single depot using multiple delivery agents. We model our mail delivery routing problem with multiple delivery agents as an mTSP, and solve the problem by combining the k-Means Clustering method with Genetic Algorithm similar to the approaches in Latah (2016), Lu et al. (2016) and Xu et al. (2018).

2. Problem Description

We study the mail delivery routing problem on the main campus of Universiti Sains Malaysia (USM) in Pulau Pinang, Malaysia. The distribution network is comprised of all buildings on campus including Schools, Centres, University Administration, Student Residence and the mail distribution centre (depot). Each node in the network represents multiple buildings grouped together due to their geographical proximity to one another. Google Maps is then used to obtain the shortest distance path between every pair of nodes. The resulting network, shown in Figure 1 has 25 nodes, where node 1 represents the depot.

The current mail delivery system in USM has 5 delivery agents, all of whom begin and end their routes at the depot. In the current system, the network is manually divided into 5 zones of nodes, one for each delivery agent. The current system does not set routes for their delivery agents. Instead, the delivery agents are given the task of deciding their own routes. While this allows for freedom for personal preference, it does not necessarily result in an optimised set of delivery routes on the whole. This information was obtained through personal communication.

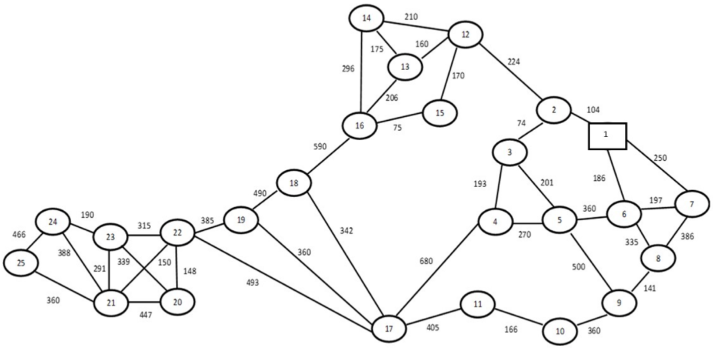

Figure 1: USM Mail Distribution Network.

In this study, we use a mathematical optimisation approach, and we model this mail delivery routing problem as a single depot mTSP with 5 salesmen. We compare our approach of solving this problem by combining the k-Means Clustering method with Genetic Algorithm with the manual approach utilised by the current system. In addition, we provide recommendations for situations where the number of delivery agents is reduced (separately) to 3 and 4 delivery agents due to factors such as employee shortage or vehicle breakdown.

3. Mathematical Formulation of the mTSP

We present an assignment-based Integer Programming formulation of the mTSP adapted from Bektas (2006) for our problem. The network is defined as a graph $G = (V, A)$, where V is the set of nodes and A is the set of arcs. The distance between each arc $(i, j) \in A$ is represented by d_{ij}, while m represents the number of delivery agents. The decision variables are defined as follows:

$$x_{ij} = \begin{cases} 1, & \text{if arc } (i,j) \text{ is used as part of a mail delivery route.} \\ 0, & \text{otherwise.} \end{cases}$$

The Integer Linear Programming formulation of the mTSP is as follows:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{(i,j)\in A} d_{ij}x_{ij} && && (1)\\
\text{subject to} \quad & \sum_{j\in V\backslash\{1\}} x_{1j} = m, && && (2)\\
& \sum_{j\in V\backslash\{1\}} x_{j1} = m, && && (3)\\
& \sum_{i\in V} x_{ij} = 1, && \forall j \in V\backslash\{1\}, && (4)\\
& \sum_{j\in V} x_{ij} = 1, && \forall i \in V\backslash\{1\}, && (5)\\
& \sum_{i\in S}\sum_{j\in S} x_{ij} = |S| - 1, && \forall S \subseteq V\backslash\{1\},\ S \neq \emptyset, && (6)\\
& x_j \in \{0,1\}, \quad \forall (i,j) \in A. && && (7)
\end{aligned}
$$

The objective function (1) minimizes the total distance travelled over all mail delivery routes. Constraints (2) and (3) ensure that all delivery agents respectively depart from and return to node 1. Constraints (4) and (5) ensure that exactly one mail delivery route respectively enters and exists each node except for node 1. Constraint (6) prevents subtours from occurring, while (7) ensures that the decision variables only take on binary values.

4. Methodology

Our proposed methodology reduces the mTSP into multiple TSP instances to be solved, and exists in two parts: (1) Divide the network (except for the depot) into zones of nodes using the k-Means Clustering method, and (2) Include the depot in each zone and use Genetic Algorithm to solve a separate TSP for each zone. Figure 2 shows the process of converting a 5-salesman mTSP into 5 TSP instances using the k-Means Clustering method (k=5). The salesmen and TSP instances correspond to delivery agents and zones respectively.

4.1 k-Means Clustering

The method of k-Means Clustering is one of the most popular clustering methods due to its ability and efficiency compared to other clustering techniques in Gayathri et al. (2015). The k-Means Clustering is basically a simple partitioning process to find k non-overlapping clusters (see Forgy (1965)). The pre-processing of k-Means Clustering in mTSP can break down the problem into several small problems and it is more efficient than dealing with mTSP itself (see Lu et al. (2016)). In relation to this study, this has allowed us to break down our problem into k=5 small problems which correspond to the number of delivery agents available. The k-Means Clustering will divide all the nodes within the distribution zones into five smaller zones which will be assigned to each delivery agent. The steps of the algorithm are (seeJohnson and Wichern (2007)):

1. Partition the nodes into k initial clusters.
2. Proceed through the list and assign a node into a cluster whose centroid (mean) is nearest. Recalculate the mean of the cluster that receives new node and cluster that loses a node.
3. Repeat step (1) until step (3), until no more reassignments of node take place.

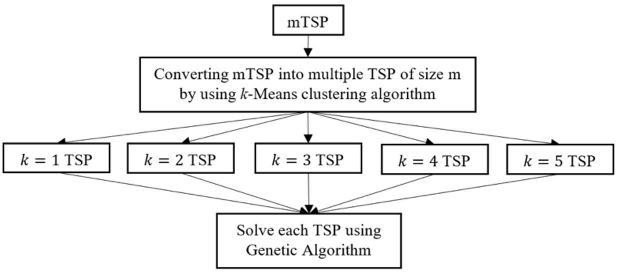

Figure 2: Converting a 5-salesman mTSP into 5 TSP instances.

4.2 Genetic Algorithm

Once the network is divided into zones of nodes, Genetic Algorithm (GA) is used to solve the resulting TSP for each zone. The selection, crossover and mutation operators used are those deemed suitable for a TSP. Tournament selection is the chosen selection operator as it is reported by Goldberg and Deb (1991) as being efficient as well as simple to implement. Self-crossover is the crossover operator chosen because the TSP requires each node to be visited exactly once, where typical crossover between two parent chromosomes may result in infeasible offspring that have repeating nodes. The swap mutation is the chosen mutation operator to act as a preventer from being trapped in a local minimum.

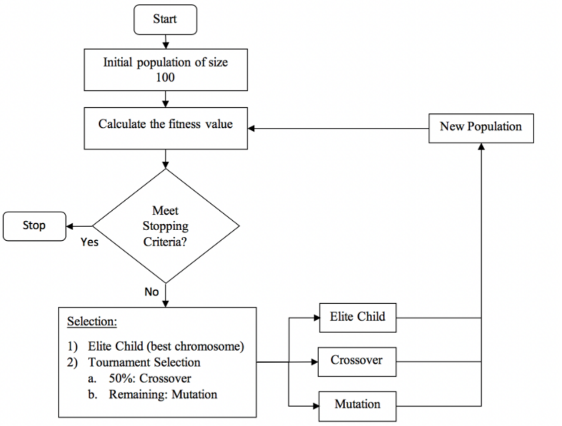

Figure 3: Flowchart for Genetic Algorithm.

The Genetic Algorithm used is based on the one implemented by the MATLAB Genetic Algorithm Solver in the Global Optimization Toolbox (see mat). We use a population size of 100 chromosomes and tournament size of four. There are three types of children produced in the reproduction process: elite child, crossover children and mutation children. Per generation, there is only one elite child, the crossover rate is 50% to produce a half of 99 chromosomes of crossover children, and the remaining chromosome will be used to produce mutation children. The evaluation of the fitness function of each chromosome is defined as the total distance travelled by all delivery agents. The stopping criteria used is either reaching a maximum number of 500 iterations or the average relative change of the fitness value over the course of 200 iterations is less than the default function tolerance, which equal to $1.0000e^{-6}$. The flowchart of GA is as shown in Figure 3.

5. Results and Discussion

The k-Means Clustering method was implemented in Minitab 16.1, while Genetic Algorithm was implemented in Matlab R2012b.

5.1 Current System Vs Our Approach

To date, the current delivery system only divides the distribution network into 5 zones without using any clustering methods. The delivery routes have not been optimised as the routes taken by the delivery agents are decided by the delivery agents themselves. Thus, to compare the performance of our approach versus the current system, we use Genetic Algorithm to solve the routing problem as TSPs in each zone of the current system to obtain an optimised set of routes. Table 1 shows the distribution of nodes to zones for the current system (manually determined) and our approach (determined using the k-Means Clustering method). Notice that in our approach, the k-Means Clustering Method determined that Zone 5 has 2 more nodes (nodes 17 and 18) than in the current system. We believe this is due to the network layout itself, where nodes 17 and 18 are closer to nodes 19 to 25 than the other nodes in the network.

Table 1: Zones for the Current System and Our Approach.

	Nodes	
Zone	Current System	Our Approach
1	2, 3	2, 3, 4, 5
2	4, 5, 6, 7, 8, 9, 10	6, 7, 8
3	11, 17, 18	9, 10, 11
4	12, 13, 14, 15, 16	12, 13, 14, 15, 16
5	19, 20, 21, 22, 23, 24, 25	17, 18, 19, 20, 21, 22, 23, 24, 25

Figures 4 and 5 show the routes obtained for each zone for the current system and our approach respectively. The routes in both figures are represented by paths consisting of nodes linked by solid and dashed lines. Two nodes linked by a solid line are adjacent to one another on the network. Two nodes linked by a dashed line are not adjacent to one another on the network. The dashed line indicates that the shortest path between the pair of nodes passes by other nodes on the network.

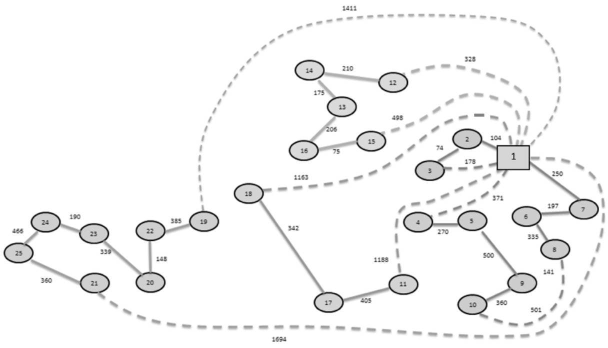

Figure 4: Delivery Routes for Current System.

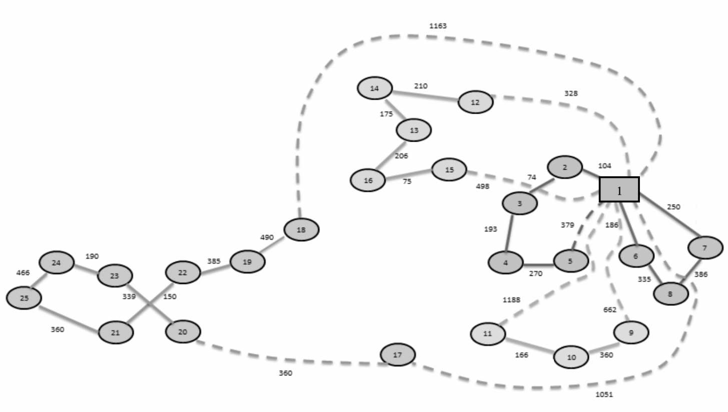

Figure 5: Delivery Routes for Our Approach.

Table 2 shows the routes obtained for each zone and the separate as well as the total distances after solving the TSPs in all 5 zones for the current system and our approach.

Table 2: Route Optimisation for Current System and Our Approach.

	Current System		Our Approach	
Zone	Route	Distance (m)	Route	Distance (m)
1	1-2-3-1	356	1-2-3-4-5-1	1020
2	1-4-5-9-10-8-6-7-1	2771	1-7-8-6-1	1157
3	1-11-17-18-1	3098	1-9-10-11-1	2376
4	1-12-14-13-16-15-1	1492	1-12-14-13-16-15-1	1492
5	1-19-22-20-23-24-25-21-1	4993	1-17-20-23-24-25-21-22-19-18-1	5235
TOTAL		12710		11280

Based on Table 2, it is observed that our approach yields a set of routes that results in an 11.25% reduction of 1430 meters in distance travelled by the delivery agents over the current system. Recall that the distance given by the current system is one that has been optimised using Genetic Algorithm, as opposed to the actual practice which allows delivery agents to make their own routes without any regard for efficiency. Less distance travelled means more cost-efficiency related to transportation and lesser emission of harmful gasses such as GHG. Moreover, less distance travelled also means less time needed to travel to all locations, which could indirectly increase customer satisfaction regarding the timeliness of the delivery service.

As for the paths and distances travelled separately by each delivery agent, notice that the k-Means Clustering method has distributed the zones such that, with the exception of zone 5, the total distance travelled in each zone does not vary by much. This was due to the partitioning of all nodes near the cluster centroid and as a result, our approach has reduced the distance travelled in Zone 1 to Zone 4 by 1672 meters. Zone 5 remains the zone with the longest distance as the nodes in this zone are located the farthest from the depot, where the distance from the depot to zone 5 and vice versa is 2214 meters which is 42.29% of the total distance travelled. It can be concluded that our approach improves upon the current system not only in terms of total distance travelled, but also in fairness of job distribution i.e. the distance travelled by each delivery agent. The fairness of job distribution may well lead to better employee satisfaction which would no doubt result in a healthier work environment.

5.2 Reduction in Number of Delivery Agents

As presented in the previous section that our approach yields better routes compared to the current system, we can confidently make recommendations to prepare for cases such as employee shortage and vehicle breakdown which would result in the reduction of the number of delivery agents. In these cases, the management will have to make smart decisions ensuring smooth delivery without much delay. The first recommendation is to recalculate the route given the available number of delivery agents or the available number of delivery vehicles. This can be done on site as the algorithm only takes a few seconds to recalculate the routes. However, prevention is better than cure. Thus, being prepared is better in handling such problems, and quick solution is readily available for implementation. A well-prepared organization will certainly be able to handle difficult situations efficiently, which would indirectly improve the organization's reputation.

We explore separate situations where the number of delivery agents are reduced to 4 and 3. The distributions of nodes to zones for the different number of delivery agents are presented in Table 3 and Table 4. These distributions of nodes to zones are determined using the k-Means Clustering method. Note that the average number of nodes per zone increases as the number of zones reduces.

Table 3: Zones for 4 Delivery Agents.

Zone	Nodes
1	2, 3, 4, 5, 6, 7
2	8, 9, 10, 11
3	12, 13, 14, 15, 16
4	17, 18, 19, 20, 21, 22, 23, 24, 25

Table 4: Zones for 3 Delivery Agents.

Zone	Nodes
1	2, 3, 4, 12, 13, 14, 15, 16
2	5, 6, 7, 8, 9, 10, 11
3	17, 18, 19, 20, 21, 22, 23, 24, 25

Figures 6 and 7 show the routes obtained for each zone for 4 and 3 delivery agents respectively. As per Figures 4 and 5 the solid lines represent a direct link between nodes, while a dashed line represents a non-direct link between nodes. A non-direct link means that the path between the two nodes passes through other nodes on the network.

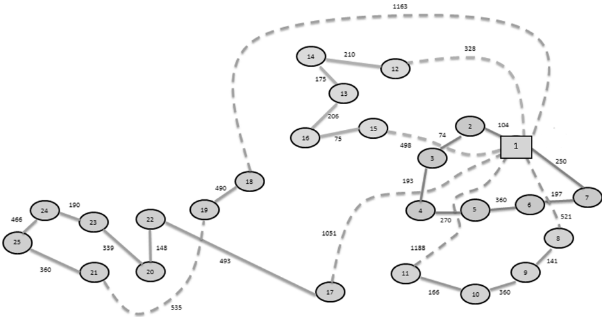

Figure 6: Delivery Routes for 4 Delivery Agents.

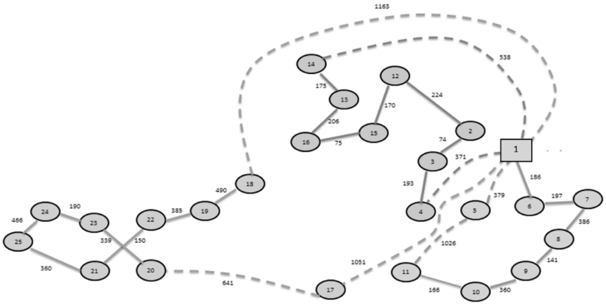

Figure 7: Delivery Routes for 3 Delivery Agents.

Table 5 and Table 6 respectively show the routes obtained for each zone and the separate and total distances after solving the TSPs in all zones for both the 4 and 3 delivery agents problems respectively.

Table 5: Route Optimisation for 4 Delivery Agents.

Zone	Route	Distance (m)
1	1-2-3-4-5-6-7-1	1448
2	1-8-9-10-11-1	2376
3	1-12-14-13-16-15-1	1492
4	1-17-22-20-23-24-25-21-19-18-1	5235
TOTAL		10551

Table 6: Route Optimisation for 3 Delivery Agents.

Zone	Route	Distance (m)
1	1-4-3-2-12-15-16-13-14-1	2026
2	1-6-7-8-9-10-11-5-1	2841
3	1-17-20-23-24-25-21-22-19-18-1	5235
TOTAL		10102

Notice that as the number of delivery agents reduces, the total distance also reduces (which is expected) as the number of times the depot node is visited reduces. With regards to the distribution of nodes to zones, note that the last zone for every problem solved by our approach (3, 4 and 5 delivery agents) has the exact same set of nodes. All other nodes are distributed fairly among other zones depending on the number of delivery agents. This shows that the k-Means Clustering method seemingly has the effect of distributing the nodes as evenly as possible into every zone.

6. Conclusion

Transportation is often linked with cost efficiency and GHG emission, as well as mail delivery, particularly because it relies heavily on transportation. Optimizing the mail delivery process and obtaining its minimal distance route benefits the service provider, the customer and the environment as less distance means less operational cost, greater customer satisfaction and less GHG emission. Our approach to solving the mTSP uses the combination of the k-Means Clustering method with Genetic Algorithm, which reduces the mTSP into multiple TSP instances, making the problem simpler to solve. Furthermore, it also solves the problem of path intersection with other delivery agents.

The results show that our approach improves upon the current system in terms of total distance travelled. The usage of the k-Means Clustering method in partitioning the distribution network is shown to result in better optimization due to its partitioning near the cluster centroid. We can also infer that the k-Means Clustering method distributes the nodes as evenly as possible among the delivery agents. In addition, we were also able to provide routes for different numbers of delivery agents as preparation for employee shortage or vehicle breakdown situations.

This study can be extended by exploring the usage of different types clustering methods and analysing their performance in terms of solution values and computational time.A possible direction of study could involve investigating other different numbers of (increased or decreased) delivery agents, or for the case where there is only 1 delivery agent. Another avenue of future work could involve exploring using software such as jsp to solve the mail delivery problem. A comparison could be made between our approach and jsprit (or equivalent) in terms of the performance of the solution quality as well as the computation time.

Acknowledgement

This study is supported by the Kementerian Pendidikan Malaysia FRGS grant 203/PMATHS/6711571.

References

GitHub page for jsprit. `https://github.com/graphhopper/jsprit`. accessed November 2020.

How the Genetic Algorithm Works. `https://se.mathworks.com/help/gads/how-the-genetic-algorithm-works.html`. accessed April 2019.

Bektas, T. (2006). The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega*, 34(3):209–219.

Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, 21:768–769.

Gavish, B. and Srikanth, K. (1986). An optimal solution method for large-scale multiple traveling salesmen problems. *Operations Research*, 34(5):698–717.

Gayathri, R., Cauveri, A., Kanagapriya, R., Nivetha, V., Tamizhselvi, P., and Kumar, K. P. (2015). A novel approach for clustering based on bayesian network. In *Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015)*, page 60. New York: Association for Computing Machinery.

Goldberg, D. E. and Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. *Foundations of Genetic Algorithms*, 1:69–93.

Gromicho, J., Paixão, J., and Bronco, I. (1992). Exact solution of multiple traveling salesman problems. In: Akgul M., Hamacher H. W., Tufekci S. (eds), *Combinatorial Optimization. NATO ASI Series (Series F: Computer and Systems Sciences)*, vol 82, pp. 291–292. Springer, Berlin.

Johnson, R. A. and Wichern, D. W. (2007). Applied multivariate statistical analysis. 6th. *New Jersey, US: Pearson Prentice Hall.*

Laporte, G. and Nobert, Y. (1980). A cutting planes algorithm for the m-salesmen problem. *Journal of the Operational Research Society*, 31(11):1017–1023.

Latah, M. (2016). Solving multiple TSP problem by K-means and crossover based modified ACO algorithm. *International Journal of Engineering Research and Technology*, 5(2):430–434.

Lu, Z., Zhang, K., He, J., and Niu, Y. (2016). Applying k-means clustering and genetic algorithm for solving MTSP. In: Gong M., Pan L., Song T., Zhang G. (eds). *Bio-inspired Computing - Theories and Applications. BIC-TA 2016. Communications in Computer and Information Science*, vol 682. pp. 278–284. Springer, Singapore.

Perboli, G. and Rosano, M. (2019). Parcel delivery in urban areas: Opportunities and threats for the mix of traditional and green business models. *Transportation Research Part C: Emerging Technologies*, 99:19–36.

Xu, X., Yuan, H., Liptrott, M., and Trovati, M. (2018). Two phase heuristic algorithm for the multiple-travelling salesman problem. *Soft Computing*, 22(19):6567–6581.