Engineering Applications of Artificial Intelligence 123 (2023) 106351

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Multi-vehicle clustered traveling purchaser problem using a variable-length genetic algorithm

Arindam Roy [a,c], Samir Maity [b], Ilkyeong Moon [c,d,*]

[a] Department of Computer Science & Application, Prabhat Kumar College, Contai, 721404, India
[b] Department of Materials and Production, Aalborg University, Aalborg, 9220, Denmark
[c] Department of Industrial Engineering, Seoul National University, Seoul, 08826, Korea
[d] Institute of Engineering Research, Seoul National University, Seoul, 08826, Korea

ARTICLE INFO

Keywords:
Clustered TPP
k-means
Variable-length GA

ABSTRACT

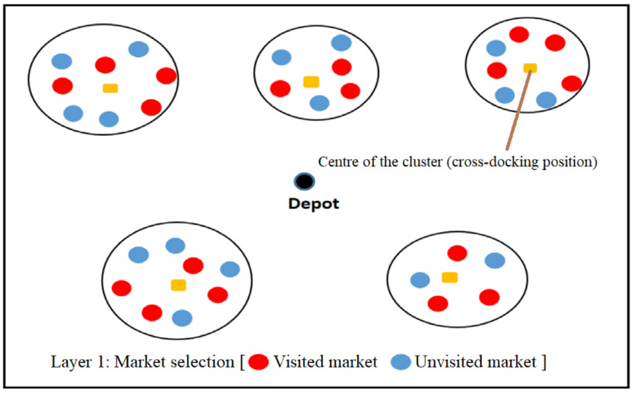
In this paper, we propose a multi-vehicle clustered traveling purchaser problem (MVCluTPP). Here, two types of procurement planning are proposed. In the first setup, the purchaser visits the markets clusterwise and, after satisfying the demand, returns to the depot with the purchased products, which are carried on the same path using a different vehicle. The other set up, in which products are purchased clusterwise and transported directly from the center of the cluster to the depot, but with the mandate that the purchaser visits the markets clusterwise. One of the multi-pronged aims of the model is to select the clusters and identify which markets to visit, determine the amount of procurement available in each market in a cluster, and develop an optimal routing plan in such a way that the overall system cost is minimized. The clusters are generated using k-means algorithm, and a variable-length chromosome genetic algorithm (VLC-GA) is proposed to optimize the cluster paths and to use a local heuristic to link the clusters to minimize the system cost. Furthermore, the superiority of the VLC-GA has been established through standard TPP and TSP instances, compared with exact methods, and some statistical tests are presented.

1. Introduction

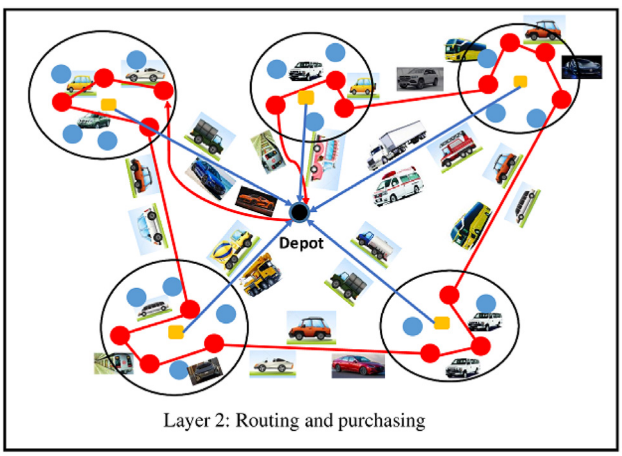
The objective of the traveling purchaser problem (TPP) is to find the optimal route that minimizes the traveling and purchasing costs (Ramesh, 1981). Nowadays, in the purchasing process, large bulk purchases normally involve separate vehicles, one to travel to the purchaser, and another to transport the goods (Nakano, 2019). Generally, purchasers are internal employees of the firm, whereas most of the transportation operations are outsourced to third-party transport providers. This type of evidence occurred in South Korea as well as in developing countries. Most markets are condensed within a region according to their product variability, such as in the case of vegetables, fish, groceries, cloth, and many other products. Typically, these regions are located side by side (clusterwise). Traditionally, in procurement planning, the purchaser visits markets and travels with one vehicle, and goods are transported in a different vehicle. The purchaser returns to the depot as soon as the demand is satisfied. Alternatively, in the new setup, the purchaser travels from market to market with bulk purchases; however, the goods that are purchased (clusterwise) are immediately sent to the depot. Fig. 1 shows the two components of the problem in a layered framework. Layer 1 illustrates the market selection, and layer 2 represents the act of traveling and purchasing and transporting goods to the depot. From these alternatives, the purchaser selects the strategy that incurs the minimum cost. Our research objective attempts to understand the cost of the impact of clusterwise disjoint traveling and transportation processes in comparison to the conventional purchasing process, in which the purchaser carries the purchased items. Transportation activity, including road transport in the logistics sector, is typically divided into two categories: the movement of people and the movement of goods. Also, the choice of the vehicle creates an opportunity for the decision maker to reduce traveling and transportation costs by choosing appropriate vehicle types. In many business environments, such as in raw materials and components purchasing, the selection of suppliers is a key procurement decision. However, these suppliers and sellers are typically sited in different geographical locations. Hence, to aid in the efficient purchasing of different products from them, the clustering concept is introduced. In both developed and developing countries, markets are located, according to their product categories and market concentrations, in different geographical locations, thereby forming a cluster, which is shown in a real-life snapshot in Fig. 2. Because of this, the purchasing processes for the different products may be facilitated by implementing a clusterwise purchasing plan. Such clusterwise procurement planning is still very

* Corresponding author at: Department of Industrial Engineering, Seoul National University, Seoul, 08826, Korea.
E-mail addresses: royarindamroy@yahoo.com (A. Roy), maitysamir13@gmail.com (S. Maity), ikmoon@snu.ac.kr (I. Moon).

https://doi.org/10.1016/j.engappai.2023.106351
Received 19 April 2022; Received in revised form 29 March 2023; Accepted 18 April 2023
Available online xxxx



(a). Market segmentation



(b). Multi-vehicle routing and purchasing

Fig. 1. Graphical view of multi-vehicle clustered TPP.

much uninvestigated by researchers. This article offers a novel slant on such procurement planning by proposing to design a clustered TPP (CluTPP) and different modes of clusterwise transportation. In this paper, we consider a clustered TPP with different types of vehicles for travel and transportation. The problem identifies which markets in a cluster to visit and procure the available products from; it also outlines how to transport those products as per the scenarios described. To solve the proposed MVCluTPP, a VLC-GA is developed. The proposed VLC-GA contains three steps: (i) introducing a k-means for effective clustering; (ii) determining how the VLC-GA works in each cluster; and (iii) designing a local heuristic to order the clusters optimally. The effectiveness of the proposed algorithm is established using standard TPP and TSPLIB instances by turning them into clustered TSPs and CluTPPs. The supremacy of the algorithms is shown through statistical analysis. The model is illustrated by numerical examples. This paper is organized as follows. A brief introduction and literature review are given in Section 1. In Section 2, our proposed mathematical model is discussed. We illustrate and describe in detail our developed VLC-GA in Section 3. In Section 4, computational results are presented in tabular and graphical form, and some managerial insights are discussed. Finally, we conclude the paper by discussing the importance of our research and the limitations and future scope of research in Section 5.

1.1. Literature survey

A concise literature survey has been done according to problem formulation and consequence methodology development, and this survey is encapsulated in Table 1. The literature survey of the TPP is divided into classical TPP, TPP with constraints, stochastic TPP, green TPP, dynamic TPP (DTPP), multi-vehicle TPP (MVTPP), MVTPP with constraints, bi-objective TPP, and bi-objective TPP with constraints, respectively; clustered TPP is still missing in the literature. Thereafter, categories of each variant of TPP are based on developed methodology, such as exact and heuristic. In Table 1, papers are listed in order of year of publication. Ramesh (1981) first introduced the TPP, which was a variant of the classical traveling salesman problem (TSP). Singh and van Oudheusden (1997) developed a branch and bound algorithm for solving the TPP. The main idea of this article was the selection of a subset of markets for the tour and the determination of an optimal tour of these markets that are successfully embedded into one. Laporte et al. (2003) developed a capacitated traveling purchaser problem (CTPP) for the first time. The author developed a mixed-integer linear programming model that was formulated and solved by exact methods. Different variants of TPP considering the constraints have been studied by several researchers. Mansini and Tocchella (2009b,a) developed a budget constraint TPP and a UTPP in which total traveling and purchasing costs are predefined and solved by a local heuristic with a variable neighborhood search (VNS). A procurement setting that explicitly incorporated both purchasing and transportation costs was studied by Mansini et al. (2012). This study found minimum purchasing and transportation costs by developing integer programming-based heuristics to solve the problem. A TPP with multiple stacks and deliveries was developed by Batista-Galván et al. (2013). In this model, they formulated a branch-and-cut algorithm and tested its performance on different instances. Gouveia et al. (2011) proposed a TPP that considered limiting the maximum number of markets to be visited and also included a limit on the number of items bought per market and formulated the setup in a dynamic programming approach. They used the Lagrangian-based method combined with a heuristic that attempts to transform relaxed solutions into feasible solutions. Several optimization approaches were applied for solving TPP by different researchers, and a concise survey on TPP research was conducted by Manerba et al. (2017). The TPP with fixed market visits using heuristics was done by Voß (1996). Choi and Lee (2011) developed a TPP formulation incorporating additional constraints, such as capacity, multi-purchaser, distance, and time restrictions. They introduced the integer linear programming formulation for multiple TPPs for maximizing system reliability under budget constraints. Kang and Ouyang (2011) formulated an extension of the TPP in which multiple types of products were sold from markets at different locations at random prices. Their formulation also considers uniform and normal distributions. Angelelli et al. (2016) analyzed a dynamic and stochastic variant of the TPP, in which the availability of products in each market was reduced over time, following a stochastic process. They introduced three variants of the heuristic approach using re-optimization to exploit new information. In the logistics field, each supplier offers a subset of products at different prices and has different availabilities. In reality, the price of a particular product also varies depending on availability, the season of purchase, the demand, and other factors. Also, the availability of a product may vary from time to time (Angelelli et al., 2016). As with the availability of a product in a market, the price of a product also varies from market to market. Under the following circumstances, these factors are considered to be random variables. In this paper, we consider the availability and price of products as random, and we follow an exponential distribution. Using multiple-vehicles in traveling makes the procurement system more practical. Riera-Ledesma and Salazar-González (2012, 2013) designed a simple school bus routing plan and a school bus routing with resource constraints as an MVTPP, and to solve these, they proposed a branch-and-cut approach and a

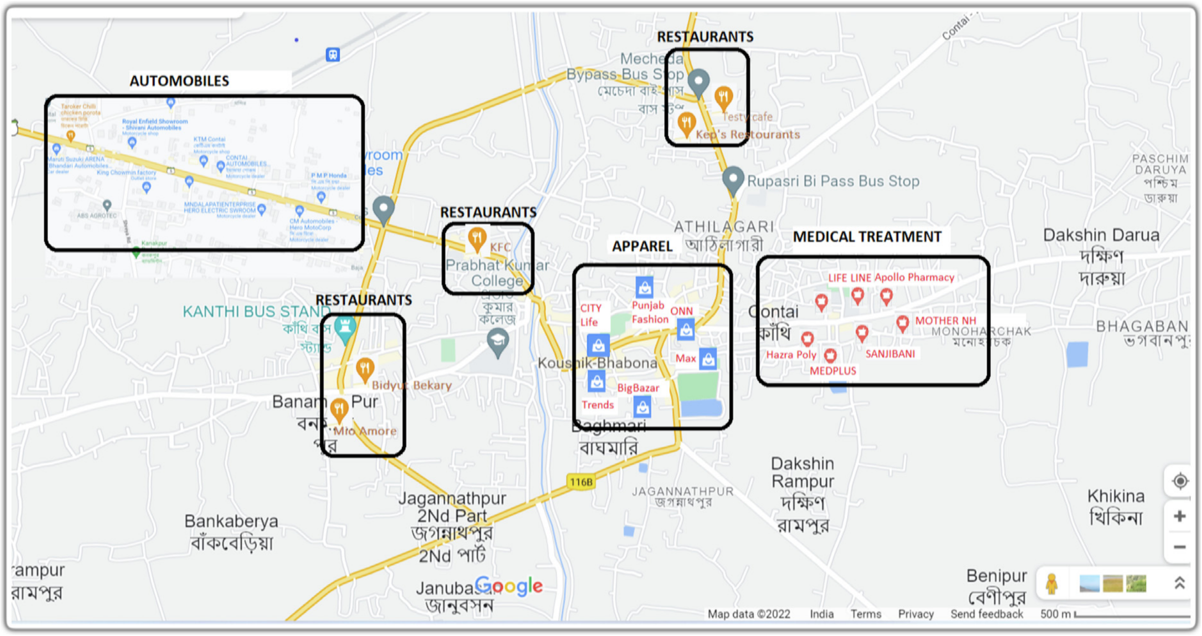


Fig. 2. Different market locations in the Contai subdivision in West Bengal, India.

column generation-based branch-and-price algorithm depending on a set partitioning formulation. Recently, Xu et al. (2022) presented a multi-parking lot and shelter heterogeneous vehicle routing problem with split pickup under emergencies. Bianchessi et al. (2014) developed a distance-constrained MVTPP, and an upper bound of the route length of each vehicle is considered. Manerba and Mansini (2015) developed an MVTPP with pairwise incompatible constraints and solved it by a branch-and-cut algorithm. Gendreau et al. (2016) considered an MVTPP with incompatible constraints. They considered a special case in which the demand for each product is unitary, and they proposed a column generation approach. They considered a multi-vehicle setup for the transportation of goods but did not consider a multi-vehicle setup for the purchasers traveling purposes. In our proposed model, we have considered both setups, which agree with reality. Pradhan et al. (2020) developed a traveling purchaser model in capacitated markets for minimizing system costs using multiple vehicles. In this setup, goods are transported from each market to the depot, which is not always feasible or cost-effective, and this research did not consider the clustering of markets. In our proposed multi-vehicle clustered traveling purchaser problem (MVCluTPP) model, we have allowed for the clustering of markets, and goods are transported from centers of clusters to the depot, which is more realistic. Very recently, Bianchessi et al. (2021) proposed a branch-price-and-cut algorithm for capacitated MVTPP under a unitary demand constraint.

Clustering analysis is a common tool used for market segmentation. The general concept behind clustering is to group similar items. The most popular partition-based clustering algorithm k-means was proposed by Hartigan and Wong (1979). Kuo et al. (2002) proposed a self-organizing feature map and k-means algorithm for market segmentation. They solved the proposed model by a two-stage method, using real-world data and simulated data. Liao and Guo (2008) developed a clustering-based location–allocation method and applied it to the capacitated facility location problem. This provided an approximately optimal solution to determine the location and coverage of a set of facilities to serve the demands of a large number of locations. Xie et al. (2019) developed an important k-means clustering with two variants of the firefly algorithm. Li et al. (2021) proposed a customer segmentation method based on the improved k-means algorithm and the adaptive particle swarm optimization (PSO) algorithm. Yu et al. (2018) implemented two improved k-means algorithms, where tri-level and bi-layer k-means algorithms were developed by the author. Cuellar-Usaquén et al. (2021) presented a GRASP-based methodology for the TPP based on three constructive procedures and two local search operators, which was strengthened with a path relinking strategy to improve the GRASP performance.

The TPP is an NP-hard problem since it generalizes three famous combinatorial optimization problems: the traveling salesman problem, the uncapacitated facility location problem, and the set covering problem (Mansini and Tocchella, 2009b, Manerba et al., 2017). In the literature, several solution methodologies are proposed, which are basically grouped into two categories: exact and heuristic solution approaches due to the complexity of the problem. Until now, CluTPP has been missing, which is formulated and illustrated in the present investigation. In this section, Table 1 presents different existing solution approaches for the TPP and its variants in the methodological order. It can be observed from Table 1 that although exact solution approaches have been utilized to solve the TPP, these approaches mostly fail to solve a large-size problem within a feasible time, due to the complexity of the problem. Therefore, several heuristic approaches have been developed to efficiently solve the TPP and its variants within an allowable computation time. Several researchers studied the TPP, such as Ochi et al. (2001), Boctor et al. (2003), Teeninga and Volgenant (2004), Riera-Ledesma and Salazar-González (2005b), Bernardino and Paias (2018), and others, and they solved it using heuristics and metaheuristics. Abreu et al. (2021) developed a new biased random key genetic algorithm with an integrated greedy local search procedure for open shop scheduling with capacitated vehicle routing. The proposed CluTPP demands intelligent clustering techniques to subdivide the markets. In artificial intelligence, k-means clustering is the most efficient technique, which is adopted here. Niknam et al. (2011) developed an efficient intelligent hybrid algorithm based on a modified imperialist competitive algorithm and k-means for data clustering. An approach to creating a time-balanced delivery within a city was developed by Menchaca-Méndez et al. (2022). Their study combined divisive clustering and simulated annealing considering real-world data. Adan et al. (2022) showed the influence of statistical feature normalization methods on k-nearest neighbors and k-means in the context of industry 4.0 in their article. An integrated k-means Laplacian (IKL) algorithm was developed by Rengasamy and Murugesan (2022). This study proposes three different ways of creating the normalized Laplacian matrix.

Table 1
Existing solution approaches.

Reference	Problem	Category	Method
Ramesh (1981)	TPP	Exact	Lexicographic search
Singh and van Oudheusden (1997)	TPP	Exact	Branch-and-bound
Laporte et al. (2003)	TPP	Exact	Branch-and-cut
Riera-Ledesma and Salazar-González (2006)	TPP	Exact	Branch-and-cut
Golden et al. (1981)	TPP	Heuristic	GSH
Ong (1982)	TPP	Heuristic	TRH
Pearn (1991)	TPP	Heuristic	CAH
Pearn and Chien (1998)	TPP	Heuristic	GSH/TRH/CAH
Ochi et al. (2001)	TPP	Heuristic	GRASP/VNS
Boctor et al. (2003)	TPP/UTPP	Heuristic	Local search
Teeninga and Volgenant (2004)	TPP	Heuristic	GSH/TRH/CAH
Riera-Ledesma and Salazar-González (2005b)	TPP	Heuristic	Local search
Bernardino and Paias (2018)	TPP	Heuristic	Biased random key GA
Cuellar-Usaquén et al. (2021)	TPP	Heuristic	Grasp/path-relinking algorithm
Mansini and Tocchella (2009b,a)	TPP and UTPP with budget constraint	Heuristic	Local heuristic & VNS
Gouveia et al. (2011)	TPP with additional side constraints	Exact	Dynamic programming
Mansini et al. (2012)	Supplier selection problem with constraints	Heuristic	Integer programming based
Batista-Galván et al. (2013)	TPP with multiple stacks and deliveries	Exact	Branch-and-cut
Voß (1996)	TPP with fixed market visiting cost	Heuristic	TS/SA
Kucukoglu (2022)	TPP with first service option	Metaheuristic	ALNS
Kang and Ouyang (2011)	TPP with stochastic process	Exact	Dynamic programming
Hamdan et al. (2017)	Green TPP	Exact	Branch-and-cut
Cheaitou et al. (2020)	Sustainable TPP with speed optimization	Heuristic	GA
Pradhan et al. (2020)	Imprecise modified solid green TPP	Heuristic	Quantum-inspired GA
Roy et al. (2020)	Solid green TPP with uncertainty	Heuristic	Noble GA
Angelelli et al. (2017)	DTPP	Exact	Branch-and-cut
Angelelli et al. (2011)	DTPP	Heuristic	Greedy heuristic/Look-ahead
Bontoux and Feillet (2008)	UTPP	Heuristic	ACO
Riera-Ledesma and Salazar-González (2012, 2013)	MVTPP	Exact	Branch-and-cut/price
Bianchessi et al. (2014)	Distance constrained MVTPP	Exact	Branch-and-price
Manerba and Mansini (2015)	MVTPP with incompatible products	Exact	Branch-and-cut
Gendreau et al. (2016)	MVTPP with incompatible products	Exact	Branch-and-price
Bianchessi et al. (2021)	MVTPP with incompatible products	Exact	Branch-price-and-cut
Riera-Ledesma and Salazar-González (2005a)	Bi-objective TPP	Exact	Branch-and-cut
Almeida et al. (2012)	Bi-objective TPP	Heuristic	TA
Palomo-Martínez and Salazar-Aguilar (2019)	Bi-objective TPP with deliveries	Heuristic	Relinked VNS
Present study	Multi-vehicle Clustered TPP with stochastic parameters	Heuristic	VLC-GA

The present article adopts the k-means algorithm for the generation of the clustering of the markets. Introducing chromosome length into the solution of TPP creates challenges and is studied by researchers because it depends on the demand for products. That is why the length of every chromosome is not equal. For a particular TPP, we need to design a variable-length chromosome to get a feasible solution to the problem. In this study, for the first time, we proposed a variable-length GA to solve the TPP. Here, we developed a multi-parent comparison crossover-based variable-length chromosome genetic algorithm (VLC-GA) with a Boltzmann probability-based selection and a sigmoid random mutation considering variable-length chromosomes. Also, an innovative idea in the crossover operator and mutation operator in VLC-GA with three parents (two original and one a dummy) has been used to produce offspring. To the best of our knowledge, very few researchers outside of Pradhan et al. (2020) and Roy et al. (2020) have implemented in vitro fertilization (IVF) comparison crossover-based GA to solve the TPP. Kucukoglu (2022) implemented the TPP with a fast service option (TPP-FSO) and solved it by a metaheuristic process. The author developed an adaptive large neighborhood search (ALNS) algorithm for the TPP-FSO. Through the literature study, there is still no article found in the CluTPP. Our present investigation fills this lacuna.

This paper, therefore, contributes to a novel TPP model concept, along with expanding methodology in many ways: (i) it addresses a new multi-vehicle clustered TPP model; (ii) it includes two different scenarios for transporting goods from markets, which involve, in scenario 1, purchased goods transported directly to the depot for bulk purchasing, and, in scenario 2, goods transported with the purchaser in a separate vehicle; (iii) it considers random availability and price; (iv) it describes the development of the VLC-GA, including a multi-parent comparison crossover with a variable-length chromosome and a sigmoid random mutation technique; and (v) it provides policy-level insights for decision support systems.

2. Proposed multi-vehicle clustered traveling purchaser problem (MVCluTPP)

In this section, we have considered a multi-vehicle clustered traveling purchaser problem (MVCluTPP) in which a different number of markets are clustered. Two types of procurement planning are proposed (large-scale purchasing in scenario 1 and small-scale purchasing in scenario 2, presented in Pradhan et al., 2020). Due to the volatilities of the product availabilities and prices, we generate random instances, showing the effectiveness of the proposed model.

In Table 2, we present the notation and description of parameters and variables.

2.1. Problem description of clustered TPP (CluTPP)

The purchasing plan always follows a key decision process from a personal level to an organizational level. Many real-life parameters, such as market selection, are based on product availability, demand, price, location of the market, and other factors, all of which play a significant role. In most cases, there is no need to visit all the markets, but to visit them instead in a group-wise or clusterwise fashion, as in real-life situations. The market groups are identified based on their locations, products, and other factors. Keeping in mind this process, we have formulated a mathematical model termed "clustered TPP."

Classical TPP

In classical TPP, we consider a set of cities or markets and a set of products. Additionally, we consider c_{ij}, the cost of travel from market

Table 2
Notation and description of parameters variables.

Notation	Description
K	Set of clusters
M	Set of markets
M_p	Subset of M, $M_p \subseteq M$
P	Set of products
F	Set of vehicles for purchaser's travel
F'	Set of vehicles for transporting goods with the purchaser
TC	Set of transport companies
V	Set of markets, including the depot
V_k	Set of markets in cluster k
c_{ij}	Purchaser's travel cost between markets i and j (for single vehicle)
c_{ijf}	Purchaser's travel cost between markets i and j by vehicle $f \in F$
c'_{ij}	Unit transportation cost between markets i and j per unit of distance (for single vehicle)
$c'_{ijf'}$	Unit transportation cost between markets i and j per unit of distance by vehicle $f' \in F'$
r_{pi}	Unit price of product p at market i
R_{pi}	Maximum unit price of product p at market i
R'_{pi}	Minimum unit price of product p at market i
q_{pi}	Available quantity of product p in market i
Q_{pi}	Maximum available quantity of product p in market i
d_p	Total demand of product p
$\{0\}$	Depot
g_{k0}	Distance from center of cluster k to the depot
g'_{ij}	Distance between markets i and j
c_{k0}	Unit transportation cost from center of cluster k to the depot, per unit of distance by transport company (for single transport company)
$c^{(tc)}_{k0}$	Unit transportation cost from center of cluster k to the depot, per unit of distance by transport company tc, $tc \in TC$
C_k	Fixed cost of transporting goods from center of cluster k to the depot (for single transport company)
$C_{k(tc)}$	Fixed cost of transporting goods from center of cluster k to the depot by transport company tc, $tc \in TC$
C'	Fixed cost of transporting goods, when goods are transported with the purchaser in a separate vehicle (for single vehicle)
$C'_{f'}$	Fixed cost of transporting goods, when goods are transported with the purchaser by vehicle $f' \in F'$
Variable	**Description**
z_{pi}	Total purchased quantity of product p from market i
x_{ij}	1 if edge (i,j) is selected, $(i,j) \in E$, $i \neq j$; 0 otherwise
x_{ijf}	1 if edge (i,j) is selected by vehicle f, $(i,j) \in E$, $i \neq j$, $f \in F$; 0 otherwise
y_i	1 if market i is selected, $i \in M$; 0 otherwise
y_{if}	1 if market i is selected by vehicle f, $i \in M$, $f \in F$; 0 otherwise
$x'_{k0(tc)}$	1 if the arc $(k,0)$ between the center of cluster k and the depot is selected by the transport company $tc \in TC$; 0 otherwise
s_k	1 if the cluster $k \in K$ is visited by the purchaser; 0 otherwise

i to market j, and r_{pi}, the price per unit of product p at market i. A purchaser starts his or her journey from a domicile or depot, travels to a subset of n markets, purchases each of the products as per demand, and finally returns to his or her domicile or depot. The mathematical formulation of the classical TPP is given in Ramesh (1981).

Clustered TSP (CTSP)

Clustered TSP, proposed by Chisman (1975), is an extension of the classical TSP, which has several real-life applications. Here, cities or markets are grouped into clusters, and the cities or markets of each cluster must be visited contiguously. This problem can be defined as a complete weight graph, $G = (V, E)$, with a set of vertices (markets), $V = \{1, 2, \dots, n\}$, and a set of edges, $E = \{(i,j), i,j \in V, i \neq j\}$. The vertex set V is partitioned into disjoint clusters $V_1, V_2, \dots, V_m$ ($V_1 \cup V_2 \cup \dots \cup V_m = V$ & $V_1 \cap V_2 \cap \dots \cap V_m = \emptyset$. Let C be an $n \times n$ symmetric cost matrix on V such that c_{ij} is the traveling cost between city i and city j. The objective of the CTSP is to find a minimum cost Hamiltonian circuit over all vertices of the graph, G, where the vertices of each cluster must be visited consecutively. The salesman starts from a city or market and finally, after visiting all cities or markets, comes back to the starting city or market. The mathematical formulation of the clustered TSP is available in Chisman (1975).

Proposed clustered TPP (CluTPP)

There is a set of cities or markets (throughout this article, the terms market, node, vertices, or city are interchangeable), such as $M = \{1, 2, \dots, n\}, |M| = n$, in our proposed CluTPP. All markets are divided into some disjoint clusters. K is the set of clusters, such as $K = \{k_1, k_2, \dots, k_c\}$. A purchaser purchases a set of products, $P = \{1, 2, \dots, w\}$, from a subset of markets, M_p, in which, $M_p \subseteq M$, to fulfill the pre-specified demands d_p of each product, p. The purchaser starts a journey from a depot indexed at $\{0\}$ for traveling, and after fulfilling the demand, the purchaser returns to the depot. Let $G = (V, E)$ be an undirected graph in which $V = M \cup \{0\}$ is the set of nodes, and $E = \{(i,j) : i \neq j,\ i,j \in V\}$ is the set of edges. The market set $V' = V - \{0\}$ is partitioned into disjoint clusters $V_1, V_2, \dots, V_c (V_1 \cup V_2 \cup \dots \cup V_c = V'$ and $V_1 \cap V_2 \cap \dots \cap V_c = \emptyset$. V_i is the set of markets belonging to cluster k_i, and $k_i \in K$, $i = 1, 2, \dots, |K|$. The purchaser contiguously visits the clusters, with a subset of markets in a cluster, to satisfy the pre-specified demand, and returns to the depot. Let C be an $(n+1)\times(n+1)$ symmetric cost matrix on V such that c_{ij} is the traveling cost of the purchaser to travel between market i and market j. The notation r_{pi} is the purchasing price per unit of product, p, at market i. It is assumed that the availability of each product, p, in market i should be positive (i.e., $q_{pi} > 0$ and $\sum_{i \in M_p} q_{pi} \geq d_p$). The notation z_{pi} represents the quantity of product, p, purchased from market i. For any $U \subset V$, let $\delta^+(U)$ and $\delta^-(U)$ denote the sets $\{(i,j) \in E : i \in U, j \in V \backslash U\}$ and $\{(i,j) \in E : j \in U, i \in V \backslash U\}$, respectively.

2.2. Mathematical formulation of clustered TPP

The objective of the $CluTPP$ is to find the minimum system cost (traveling cost and purchasing cost) over the visited markets. Hence, the mathematical model of $CluTPP$ is as follows:

$$Minimize\ Z = \sum_{i \in V} \sum_{j \in V} c_{ij} x_{ij} + \sum_{p \in P} \sum_{i \in M_p} r_{pi} z_{pi} \tag{1}$$

$$subject\ to\ \sum_{k \in K} s_k \geq 1 \tag{2}$$

$$\sum_{i \in V} x_{i0} = \sum_{j \in V} x_{0j} = 1 \tag{3}$$

$$\sum_{j \in V} x_{ij} = \sum_{j \in V} x_{ji} = y_i,\ \forall i \in M \tag{4}$$

$$s_k = y_i\ (\exists\ i \in V_k),\ \forall k \in K \tag{5}$$

$$\sum_{i \in M_p} z_{pi} = d_p,\ \forall p = 1, 2, \dots, w \tag{6}$$

$$z_{pi} \leq q_{pi} y_i,\ p \in P,\ i \in M_p \tag{7}$$

$$u_i - u_j + (|V_k| - 1) x_{ij} \leq |V_k| - 2,\ \forall\ i,j \in V_k,\ k \in K \tag{8}$$

$$u_i \leq |V_k| - 2, \forall i \in V_k,\ k \in K \tag{9}$$

$$z_{pi} \geq 0,\ p \in P,\ i \in M_p \tag{10}$$

$$x_{ij} \in \{0,1\}, i \neq j,\ i,j \in V \tag{11}$$

$$s_k, y_i \in \{0,1\},\ i \in M,\ k \in K \tag{12}$$

$$u_i \geq 0,\ \forall i \in M \tag{13}$$

The objective function (1) minimizes the system cost (traveling and purchasing cost). Constraint (2) ensures that the purchaser visit at least one cluster. Constraint (3) guarantees that the tour starts and ends at the depot node. Constraint (4) shows that market i is visited by the purchaser and it ensures the continuity condition. Constraint (5) relates variable s_k to variable y_i. Constraint (6) ensures that the purchaser will purchase product, p, to satisfy its demand, d_p. Constraint (7) establishes that the purchaser cannot buy a product from a market if the market is not visited. The sub-tour elimination within a cluster is defined by the Constraints (8) and (9). Constraints (10)–(13) impose non-negative conditions on variables and binary variables.

2.3. Multi-vehicle clustered TPP (MVCluTPP)

A MVCluTPP is a TPP in which, at each market, some conveyances are available to travel to another market. Here, the purchaser travels to multiple markets and places orders that will be transported to the depot. Both the purchaser's travel and the transportation of goods may be arranged in multiple vehicle types. The mathematical formulation of the MVCluTPP is as follows.

$$Minimize\ S_1 = \sum_{i\in V}\sum_{j\in V}\sum_{f\in F} c_{ijf}x_{ijf} + \sum_{p\in P}\sum_{i\in M_p} r_{pi}z_{pi} \tag{14}$$

$$subject\ to\ \sum_{i\in V}\sum_{f\in F} x_{i0f} = \sum_{j\in V}\sum_{f\in F} x_{0jf} = 1 \tag{15}$$

$$\sum_{j\in V} x_{ijf} = \sum_{j\in V} x_{jif} = y_{if},\ \forall i\in M,\ f\in F \tag{16}$$

$$s_k = y_{if}\ (\exists\ i\in V_k,\ f\in F),\ \forall k\in K \tag{17}$$

$$z_{pi} \le q_{pi}y_{if},\ p\in P,\ i\in M_p,\ f\in F \tag{18}$$

$$\sum_{f\in F}(u_{if} - u_{jf} + (|V_k|-1)x_{ijf}) \le |V_k|-2,\ \forall\ i,j\in V_k,\ k\in K \tag{19}$$

$$\sum_{f\in F} u_{if} \le |V_k|-2, \forall i\in V_k, k\in K \tag{20}$$

$$x_{ijf}\in\{0,1\}, i\ne j,\ i,j\in V,\ f\in F \tag{21}$$

$$s_k, y_{if}\in\{0,1\},\ i\in M,\ f\in F,\ k\in K \tag{22}$$

$$u_{if}\ge 0, \forall i\in M, f\in F \tag{23}$$

with Constraints (2), (6), and (10). In this subsection, the above objective function and constraints have the same meaning as in Section 2.2 (Constraints (2)–(13)) by considering different types of vehicles.

Scenario 1: Goods are transported directly to the depot from each cluster

$$\left.\begin{array}{c} S_1 = \sum_{i\in V}\sum_{j\in V}\sum_{f\in F} c_{ijf}x_{ijf} + \sum_{p\in P}\sum_{i\in M_p} r_{pi}z_{pi} \\ S_2 = \sum_{k\in K}\left(C_{k(tc)} + \sum_{p\in P}\sum_{i\in M_p} c^{(tc)}_{k0}g_{k0}z_{pi}\right)s_k x'_{k0(tc)},\ tc\in TC \\ \text{Minimize } S = S_1 + S_2 \end{array}\right\} \tag{24}$$

subject to Constraints (15)–(23), along with $x'_{k0(tc)}\in\{0,1\}, k\in K, tc\in TC$. The objective function (24) contains two parts: S_1 stands for the purchaser's traveling cost and procurement cost, and S_2 stands for the transportation cost of the clusterwise goods. Goods are transported directly to the depot by the transport company, tc. In this scenario, constraints have the same meaning as in Section 2.2 (Constraints (2)–(13)) by considering different types of vehicles.

Scenario 2: Goods are transported with the purchaser in a separate vehicle

$$\left.\begin{array}{c} S_1 = \sum_{i\in V}\sum_{j\in V}\sum_{f\in F} c_{ijf}x_{ijf} + \sum_{p\in P}\sum_{i\in M_p} r_{pi}z_{pi} \\ S_2 = C'_{f'} + \sum_{p\in P}\sum_{i\in M_p}\sum_{j\in M_p} c'_{ijf'}g'_{ij}z_{pi}x_{ijf'},\ i\ne j\ f'\in F' \\ \text{Minimize } S = S_1 + S_2 \end{array}\right\} \tag{25}$$

subject to Constraints (15)–(23). The objective function (25) contains two parts: S_1 stands for the purchaser's traveling cost and procurement cost, and S_2 stands for the cost of transporting goods, where goods are transported with the purchaser in a separate vehicle, $f'\in F'$. In this scenario, constraints have the same meaning as in Section 2.2 (Constraints (2)–(13)) by considering different types of vehicles.

3. Proposed VLC-GA

In this section, we have proposed an algorithm with the combination of k-means (Algorithm 5), VLC-GA (Algorithm 2), and a local heuristic, or linking mechanism (Algorithm 3). We have clustered the markets based on k-means, and in each cluster, paths are optimized by using a novel VLC-GA. Finally, the clusters are linked to optimize the entire cycle using a developed heuristic. Because of the restriction of the TPP, we have designed variable-length chromosomes based on the GA. The VLC-GA also considered three operators: the Boltzmann probabilistic selection, the multi-parent comparison crossover (Algorithm 1), and the sigmoid random mutation operators for faster execution. The sequential steps of the methodology are described as follows:

3.1. Optimal path design

After generating the clustered market, the next aim is to find the optimum path in each cluster that satisfies the demand of the products. In the exact approach, it is not possible to determine the best path that considers traveling and purchasing costs in a feasible time, so the evolutionary computation is a good selection in this situation. Our developed VLC-GA is an innovative approach used to extend classical GA to solve our developed MVCluTPP when the chromosome length in each cluster is variable. Thus, we developed the GA with chromosomes of different lengths. The proposed VLC-GA is described in Section 3.3. In our developed model, we have used VLC-GA to generate the shortest path in each cluster. Each chromosome in each cluster constructs a TPP solution iteratively. Finally, all nodes are clustered into $|K|$ groups, and in each cluster, there is an optimal path. The centroid of each cluster will be used for the cluster connecting step.

3.2. Proposed variable-length chromosome genetic algorithm (VLC-GA)

Here, because of the demand for the products, different numbers of markets are selected that are treated as chromosomes of VLC-GA, as well as paths of the clusters that need to be optimized according to each cluster. The proposed VLC-GA worked identically in different clusters. Now the proposed VLC-GA is the combination of probabilistic selection, multi-parent comparison crossover, and sigmoid random mutation. Our developed VLC-GA is used in each cluster to create an optimum path. Here, we proposed a novel trait of VLC-GA using the probabilistic selection (the Boltzmann probability), the multi-parent comparison crossover, and a sigmoid random mutation, among a set of potential solutions, to get a new set of solutions. As usual, the sequence is continued until terminating conditions are encountered. The proposed VLC-GA and its procedures are presented below.

Representation:

Here, a complete path of $|V'_k|(V'_k\subseteq V_k)$ markets among $|V_k|$ markets represents a solution in the cluster k; $|V'_k|$ depends on the demand in cluster k. Hence, a $|V'_k|$ dimensional integer vector' $X_i = (x_{i1}, x_{i2}, \cdots, x_{i|V'_k|})$ is used to represent a solution in the cluster k, where $x_{i1}, x_{i2}, \cdots, x_{i|V'_k|}$ represented $|V'_k|$ consecutive markets in a path of the cluster k. Population size is the number of such solutions $X_i = x_{i1}, x_{i2}, \cdots, x_{i|V'_k|}), i = 1, 2, \cdots, N$, and is randomly generated by a random number generator. Here, N represents the number of chromosomes (i.e., solutions).

Probabilistic selection:

Here, we proposed a predefined value; for example, a probability of selection parameter (p_s). Each solution of f_i, randomly generates a number, r, from the range [0,1]. If $r < p_s$, the corresponding chromosome is stored to form the mating pool. For minimizing costs, a chromosome is selected in the neighborhood of the minimum solution of the entire solution space, so it propagates a higher convergence rate. From the initial population, the best fitted chromosome for TPP is chosen as having the least fitness (because TPP is a minimization problem) value (e.g., f_{min}). To form the mating pool, we use the Boltzmann-probability of each chromosome from the initial population.

Here, $p_B = e^{((f_{min}-f_i)/T)}$, T = T_0(1-a)^h, h = (1+100*(g/G)) where g = current generation number, G = maximum generation, T_0 = rand[5100], a =rand[0,1], f_i indicates the chromosome corresponding to X_i, and i = chromosome numbers. Procedures described by Maity et al. (2015) are followed in the present study.

Multi-parent comparison crossover

Better decision-making always depends on the information available, and more information is crucial for making better decisions. The most challenging factors in developing a VLC-GA are to adopt the genetic operations, to generate the proper solution, and to evaluate a fitness function value that satisfies constraints. The classical crossover techniques have a limitation on chromosomes of different lengths, but in the proposed technique, we overcome it. In this crossover, mating is performed among the parent nodes of different lengths. Here, we compared the costs (traveling and purchasing) until the demands were satisfied. In place of the two parents, we used multi-parent (three) and produced two offspring that differently replaced the parents (first two, any two, or worst two). In the proposed crossover method, we randomly chose three individuals (parents) to produce new individuals (offspring). To get the optimal result of a TPP, we took a tour from one node (market) to the next node (market) with minimum costs (traveling and purchasing costs), thereby maintaining the TPP conditions. Based on the above idea, we constructed the crossover mechanism in the following way: First, in order to select three individuals (parents) from the mating pool, we generated the random number, r, between [0, 1]. If $r < p_c$, we then selected that population for the first parent (say Pr_1). Similarly, we chose another parent (say Pr_2 and Pr_3). The multi-parent comparison crossover is done according to the following mechanism illustrated in Fig. 3. Here, three parents – parent 1, parent 2, and parent 3 – are taken with lengths 5, 3, and 4, and the demand of 150 units. After comparing the total cost (TC) between markets, offspring are generated. Here, offspring 1 and offspring 2, with lengths 4 and 3, respectively, satisfied the demand of 150. So, the above multi-parent comparison crossover procedure is as follows:

Algorithm 1: Multi-parent comparison crossover

Data: Set of parents' chromosomes
Result: A number of offspring
Initialization of the three parents, Pr_1, Pr_2, Pr_3, depends on the probability of crossover p_c.
Select a market (randomly) between 1 and the number of markets (say a_i).
Rearrange selected markets in the parent chromosomes by placing a_i in the first position of each parent.
In the first offspring, place a_i in the first position.
while *up to fulfill the demands* **do**
  Find the minimum total cost between a_i to each next node of the given parents.
  Place b_i (say) at the place i of the first offspring and update/rearrange each parent chromosome with b_i in the place i.

Sigmoid random mutation

In the sigmoid random mutation technique, in place of fixed mutation probability, (p_m), we dynamically updated the mutation probability by calculating it with the decrease of generation. In our developed VLC-GA, the sigmoid function returns a value between 0 and 1, depending on the parameter, α, and the current generation number. Following our expression of the sigmoid function, it will generate a value in [0, 1], which is used as the mutation probability. Increasing the generation number compels the value of p_m to decrease with a lower value, and vice versa. In the initial stage, in some iterations, a high value of p_m is maintained to explore the solution space, and gradually it stabilizes for convergence. The mutation process follows the following steps:

a. Generation–dependent mutation (p_m): We propose the probability of mutation (p_m) by

$$p_m = \alpha(1 + e^{-g}), \alpha \in [0, 0.5]$$

where g is the number of generations.

b. Selection for mutation: For each solution of $p(t)$, generate a random number, r, in the interval [0,1]. If $r < p_m$, then the solution is taken for mutation.

c. Mutation process: To mutate a solution $X = (x_1, x_2, \dots, x_{|V_k'|})$ of TPP, select two random integers, i, j, in the range [1, $|V_k'|$]. Then interchange x_i, x_j to get new solutions that replace the parent solution. Our proposed variable-length chromosome genetic algorithm (VLC-GA) is given in Algorithm 2.

Algorithm 2: Proposed variable-length chromosome genetic algorithm (VLC-GA)

Data: Max gen (S_0), population size (pop_size), problem data (cost, availability, demand, distance matrices, and other parameters)
Result: The optimum and near-optimum solutions in each cluster
Set initial generation t ← 0.
Randomly generate initial population $p(t)$, where f_i, i=1,2...,pop_$size$ are the chromosomes, $|V_k'|$ numbers of the node in each chromosome represent a solution or path in the cluster k of the TPP, and $|V_k'|$ depends on the demand.
Evaluate the fitness of each solution of the initial population $p(t)$.
while ($t \le S_0$) **do**
  Update the generation t ← t+1.
  Selection procedure:
  Determine the Boltzmann probability (p_B) of each chromosome of $p(t)$ according to the probabilistic selection in subsection 3.2.
  Create the mating pool based on p_s and p_B.
  Crossover procedure:
  while *depend on* p_c **do**
    Select the parents using p_c from the mating pool.
    Perform the crossover operations on selective chromosomes / solutions following the multi-parent comparison crossover mechanism in subsection 3.2.
    Generate offspring and replace the parents.
  Mutation procedure is done according to the sigmoid random mutation in subsection 3.2.
  Select the offspring for mutation based on p_m.
  Exchange the place of these nodes.
  Store the new offspring into offspring set.
  Compare the fitness and store the local optimum and near-optimum solutions.
Store the global optimum and near-optimum results in each cluster.

3.3. Cluster linking

After clusterwise purchasing, the purchaser comes back to the depot. A route needs to be designed here for visiting all the required markets (i.e., to link the visiting cluster design to a local heuristic). For this mechanism, we randomly selected a cluster (Fig. 4(a)). Next, we found the nearest cluster from the randomly selected cluster depending on the shortest distance (Euclidean distance, Fig. 4(a)) from all other centroids of the remaining clusters. To link these two clusters, we had to find two nodes or markets from these two clusters in such a way that one node from a cluster was nearest to the centroid of another cluster, and vice versa (Fig. 4(b) and Fig. 4(c)). Finally, we linked these two nodes and got a path by linking two end nodes. This process must be repeated until all clusters are linked (Fig. 4 and Fig. 5). The linking process is completed and depicted in Fig. 6. The procedure of the re-linking of the clusters is given in Algorithm 3 with the help of Figs. 4 to 6.

Algorithm of VLC-GA

Here, the proposed VLC-GA is described in Algorithm 4. The flowchart is given in Fig. 7.

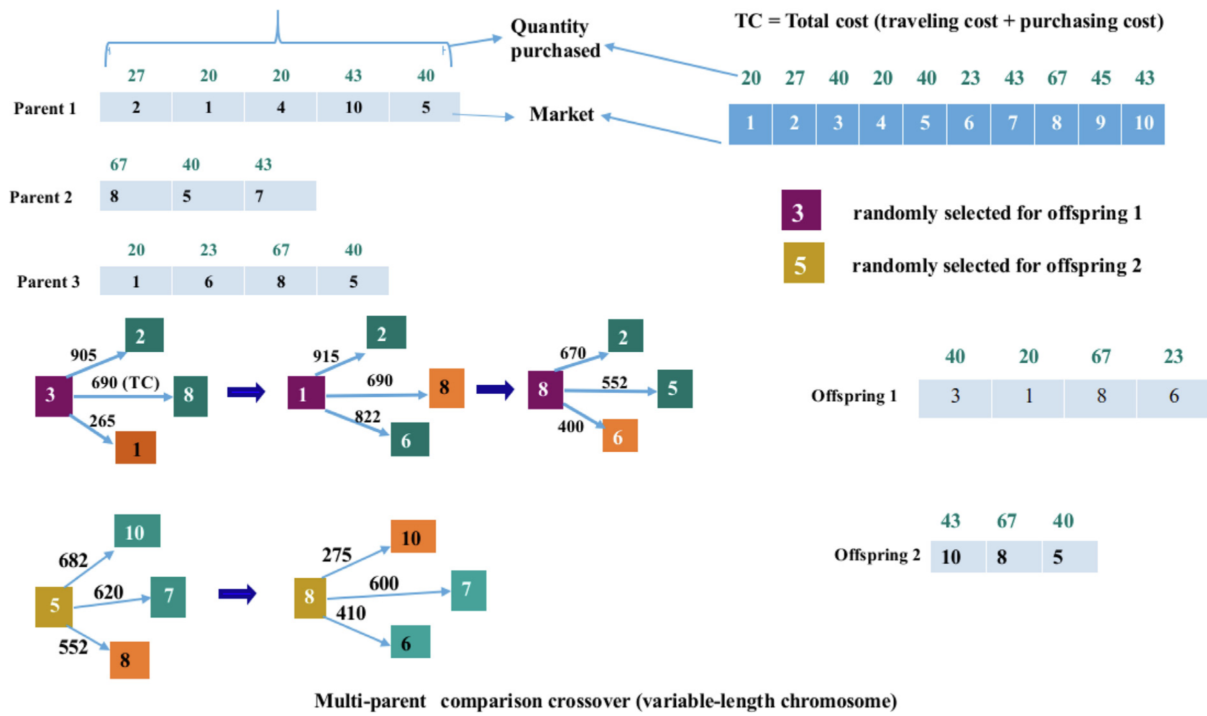


Fig. 3. Multi-parent comparison crossover.

Algorithm 3: Cluster linking

Data: Set of clusters with a unique optimized path in each cluster
Result: A complete tour among all clusters
Consider the clusters (with optimized path).
Randomly select a cluster.
while *(all clusters are connected)* **do**
  Find the nearest cluster from the randomly selected cluster depending on the minimum distance among all other centroids.
  Determine the distance from the centroid of the selected cluster to the nearest cluster and identify the shortest distance.
  Determine the distance from the centroid of the nearest cluster to the selected cluster and identify the shortest distance.
  Find the minimum distance measured in the previous two steps and connect corresponding markets.
  After linking these two clusters, generate a new path having visited two end markets.
  Find the nearest centroid among the remaining clusters from those two end points to find out the next cluster to re-link.
Generate a path connecting all the clusters considering the depot.

Algorithm 4: VLC-GA-Cluster linking

Data: Problem data, GA parameters, number of clusters ($|K|$)
Result: The optimum tour of the clustered TPP
Initialize the set of markets and the number of clusters ($|K|$).
Execute k-means algorithm according to Algorithm 5 in subsection 4.1.
Start VLC-GA.
Initialize the total number of products, availability, demand, and price of each product in each market.
Distribute the total demand among clusters.
Start VLC-GA for each cluster separately to find the best path to fulfill the pre-defined demand of each product in each cluster.
Execute VLC-GA according to Algorithm 2 in subsection 3.2.
Execute cluster linking among all clusters to find the complete optimum tour to satisfy the demand according to Algorithm 3.

4. Empirical test

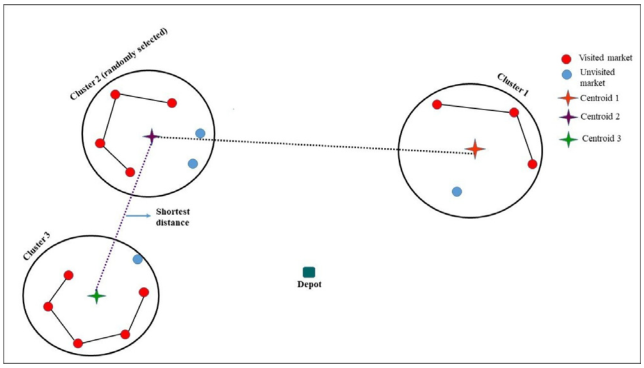
Before illustrating the computational results, we first describe the market clustering using k-means. Expected values of availability and cost of the products are described here.

4.1. Cluster computation

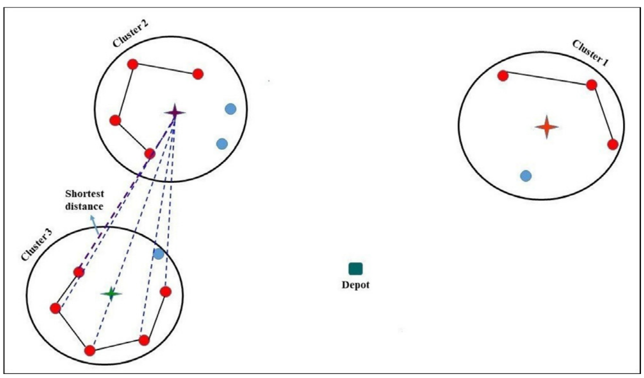
The markets are distributed in distinct geographical locations, with different types of products, availability, prices, and other factors. Here, the clusters are generated based on the markets' positions. Thus, optimal clustered design is a challenging task for the proposed MVCluTPP. For simplicity, we used the k-means algorithm to group the markets. Thus, the proposed methodology is based on sub-categories according to market clustering, optimal path finding, and cluster connecting. This section uses the clustering technique to improve the performance of the evolutionary computation algorithm (i.e., GA in the CluTPP). This section describes the following steps (i.e., market clustering using k-means).

Market clustering using k-means

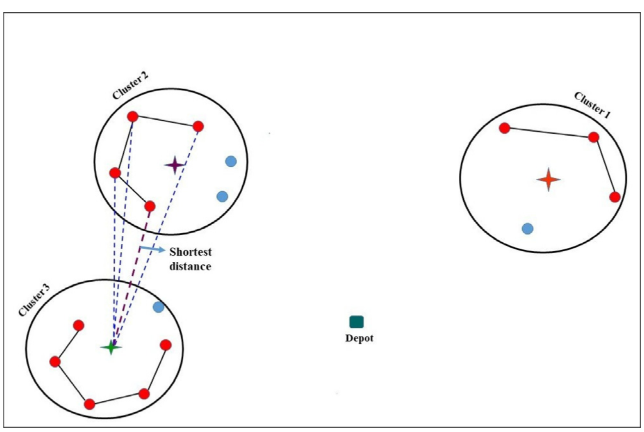
Here, markets are clustered using k-means algorithms. It is a distance-based clustering technique, from which centroids of each cluster are calculated. This algorithm performs an iterative alternating fitting process to form a specific number of clusters. Initially, $|K|$ points are randomly selected to be a first guess of the centroids of the clusters.



(a) Distance from randomly selected cluster to other cluster



(b) Distance between centroid 2 and all visited markets in cluster 3



(c) Distance between centroid 3 and all visited markets in cluster 2

Fig. 4. Randomly selecting a cluster and the shortest distance between clusters.

Each node or market is assigned to the nearest centroid to form temporary clusters. In the next step, the centroids are replaced by the cluster means, and the nodes, or markets and centroids, are reassigned. This process continues until no further changes occur in the clusters. The k-means is one of the partition-based clustering techniques proposed by Hartigan and Wong (1979). The classical k-means is described here in Algorithm 5. In this proposed MVCluTPP model, the whole market set is grouped or clustered using k-means. The number of clusters is selected by the elbow method (Syakur et al., 2018), the distance based on Euclidean types, and, finally, sum square errors (SSE) examine its consistency. Fig. 8 represents the elbow method to find the optimum number of clusters for k-means clustering. The k-means clustering algorithm is given below:

Algorithm 5: k-means clustering algorithm

Data: Set of given markets, M, and the number of clusters, $|K|$
Result: Unique set of markets in each cluster; centroid of each cluster
Initialize a positive integer, c (number of clusters, $c = |K|$) using elbow method.
Generate markets of each cluster (initial set) randomly.
Determine the centroid of each cluster.
while *the centroids no longer move* **do**
    Set each market to the cluster for which it is nearest to the centroid.
    Update the centroid of each cluster.

4.2. Generation of products availability and prices

In this procurement and planning process, the availability and price of products are considered as random. The random variables corresponding to the availability and price of the products have the following p.d.f as

$$f(\hat{q}_{pi}) = \begin{cases} \frac{\lambda_1 . e^{-\lambda_1 . q_{pi}}}{1 - e^{-\lambda_1 . Q_{pi}}}, & 0 \leq q_{pi} \leq Q_{pi}, \quad \lambda_1 \geq 0, \ p = 1, 2, \ldots, w, \ i = 1, 2, \ldots, n \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

$$f(\hat{r}_{pi}) = \begin{cases} \frac{\lambda_2 . e^{-\lambda_2 . r_{pi}}}{e^{-\lambda_2 . R'_{pi}} - e^{-\lambda_2 . R_{pi}}}, & R'_{pi} \leq r_{pi} \leq R_{pi}, \quad \lambda_2 \geq 0, p = 1, 2, \ldots, w, \ i = 1, 2, \ldots, n \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

So, the expected values of these random variables are

$$E(\hat{q}_{pi}) = \frac{1}{\lambda_1} - \frac{Q_{pi} e^{-\lambda_1 . Q_{pi}}}{(1 - e^{-\lambda_1 . Q_{pi}})}, \quad p = 1, 2, \ldots, w, \ i = 1, 2, \ldots, n \tag{28}$$

$$E(\hat{r}_{pi}) = \frac{1}{\lambda_2} + \frac{(R'_{pi} e^{-\lambda_2 . R'_{pi}} - R_{pi} e^{-\lambda_2 . R_{pi}})}{(e^{-\lambda_2 . R'_{pi}} - e^{-\lambda_2 . R_{pi}})}, \quad p = 1, 2, \ldots, w, \ i = 1, 2, \ldots, n \tag{29}$$

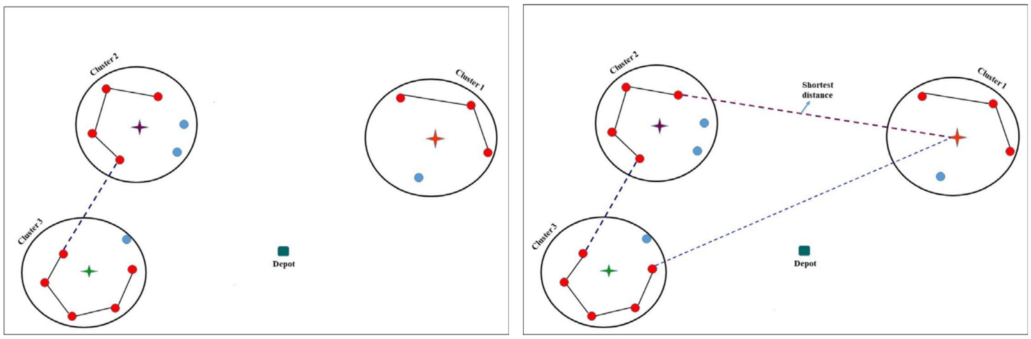
In the computational scheme, q_{pi}, and r_{pi} are replaced by $E(\hat{q}_{pi})$, and $E(\hat{r}_{pi})$.

4.3. Computational experiments

Our proposed algorithm is created in Python on a PC with an Intel Core i5 processor running at 3.2 GHz and 8 GB of RAM. The input data and code are available on the given link (https://github.com/Samir-Maity/Clustered-Travelling-Purchaser-Problem). In this study, the proposed VLC-GA is our developed algorithm. The goal of our study is to fulfill the main objective of minimizing the total cost (which includes the cost of traveling, purchasing, and transportation).

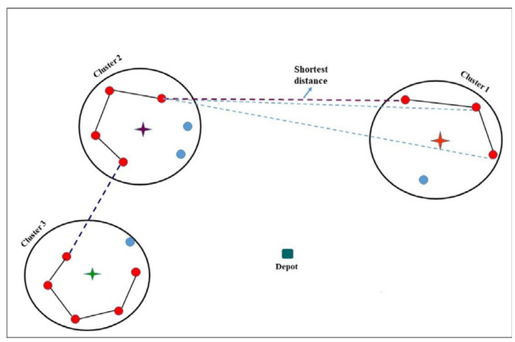
We chose the bays29, a standard TSPLIB (Reinelt, 1995) instance of 29 nodes, as a dataset to numerically illustrate the proposed MVCluTPP using a developed VLC-GA. In bays29, each node represents a market, except for node 0. Node 0 represents the depot, and the other 28 nodes are considered as 28 different markets in our problem. Here, we have considered the bays29 distance matrix as the standard distance matrix for the numerical illustration of the problem.

Table 3 shows the performance of our proposed VLC-GA and exact solver CPLEX based on standard TPP instances considered by Angelelli et al. (2017). The selected nine instances are between 17 nodes and 42 nodes. Market planning for all selected instances is prepared for



(a) Linking between center 2 and cluster 3 (nearest points)

(b) Distance from two end points to cluster 1



(c) Distance from end point of cluster 2 to all visited markets in cluster 1

Fig. 5. Cluster linking and finding the nearest cluster.

Table 3
Standard TPP instances (Angelelli et al., 2017) using VLC-GA and CPLEX.

Instances	Total market	No of visited markets	Purchased amount	Routing cost	Purchasing cost	Total cost	Mean	SD	computation time (s)	Total cost (CPLEX)	Comp. time sec. (CPLEX)
gr17p01_s02	17	8	562	1249	1702	2951	2983.67	23.75	43.71	2951	250
gr21p01_s00	21	10	847	1997	2397	4394	4451.67	93.84	44.91	3935	180
ulysses22p01_s00	22	12	1131	2974	4581	7555	7690.56	157.15	45.11	7478	2685
gr24p01_s02	24	9	778	888	2037	2925	3009.44	46.25	43.48	2544	3600
fri26p01_s00	26	11	948	592	2668	3260	3740.89	279.92	45.85	3260	3600
bays29p01_s00	29	10	901	968	2400	3368	3748.44	259.18	86.43	2943	3600
bayg29p01_s00	29	9	813	942	2151	3093	3274.22	221.06	85.31	2831	3600
dantzig42p01_s00	42	15	1493	704	4139	4843	4988.44	289.54	85.54	—	—
swiss42p01_s01	42	16	1387	938	3782	4720	4895.56	138.68	87.73	—	—

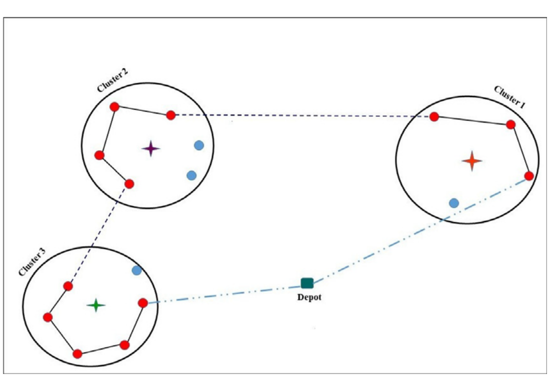


Fig. 6. Optimal path connecting with the depot.

single item purchasing, and instance-wise availability, price, and total demand are considered according to Angelelli et al. (2017). Table 3 represents the number of visited markets, purchasing cost, routing cost, and statistical measures with mean and standard deviation (SD) based on the total cost to fulfill the instance-wise demand. Also, the total cost obtained using CPLEX within one-hour run time is presented in Table 3 (dantzig42p01_s00 and swiss42p01_s01 did not give any results within one-hour time). It has been observed that, in all cases, CPLEX gives better results than VLC-GA.

In another experiment, markets are randomly generated (following Riera-Ledesma and Salazar-González, 2005b, 2006) by generating integer coordinate vertices in the [0, 1000]×[0, 1000] square according to a uniform distribution and by calculating routing costs by Euclidean distances. Each product, k, was associated with $|M_k|$ randomly selected markets, where $|M_k|$ was randomly generated from 15 to 70. The number of products is generated from 2 to 7 for various types of demands ranging from 700 to 7000. For these instances, results are depicted in Fig. 12 (small instances) and Fig. 13 (large instances).

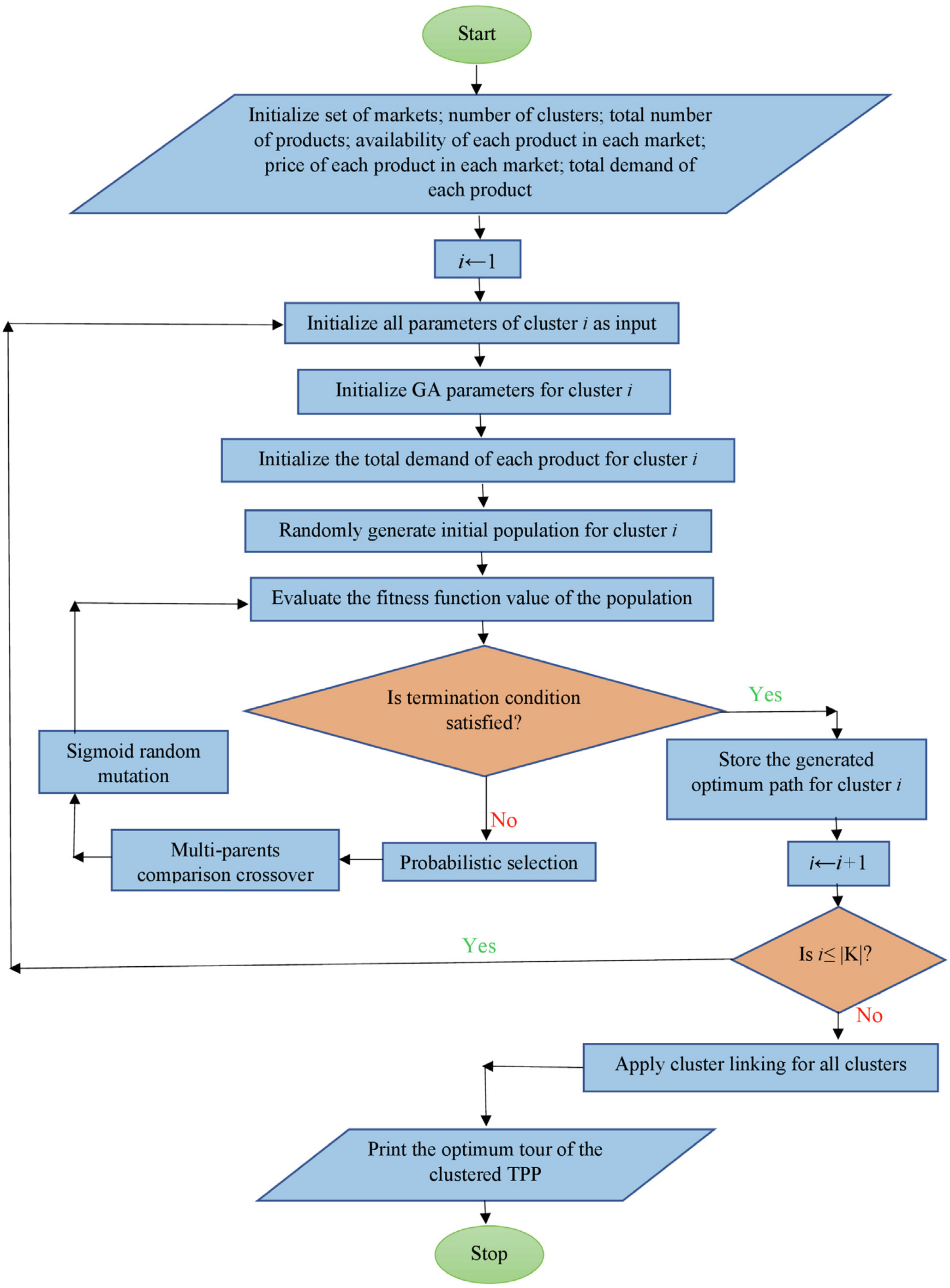


Fig. 7. Flowchart of VLC-GA.

Fig. 9 shows the market planning for different standards of TPP. Fig. 10 depicts a clusterwise purchasing plan. According to the proposed MVCluTPP, here the goods are transported from the center of a cluster to the depot. A comparison between routing, purchasing costs, and demand is shown in Fig. 11.

Different markets' clusterwise visits with product availability and demand are shown in Fig. 12. The demand variations generate a different number of clusters. An increase in market concentration in the procurement settings between transportation distances may decline. Similar observations are found in Fig. 13 for large-scale business processes.

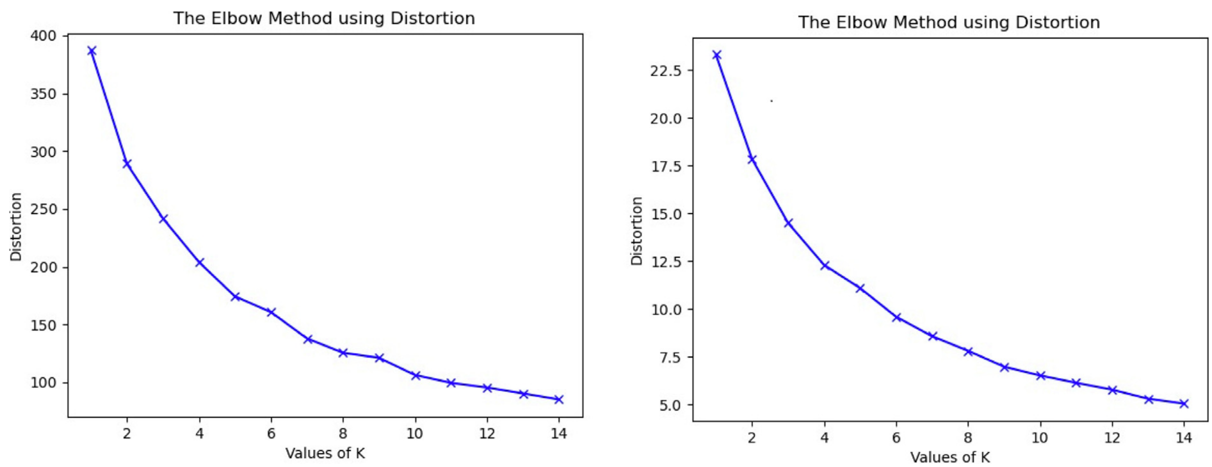
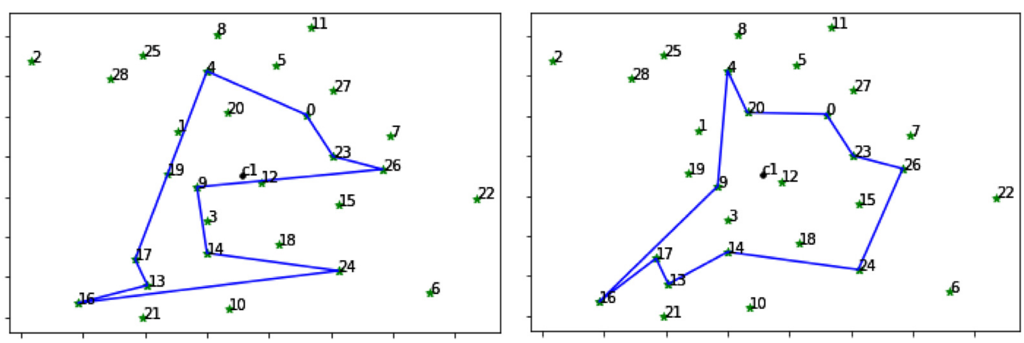
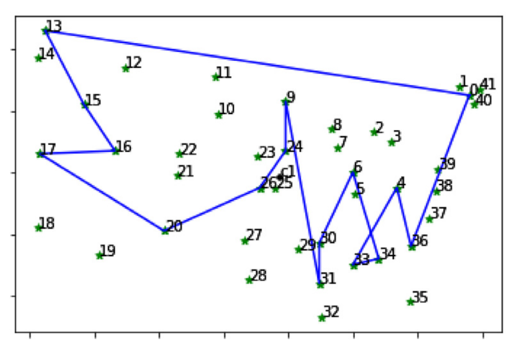


Fig. 8. Elbow method to find the optimal number of clusters.



(a) Market planning for bayg29p01 (b) Market planning for bays29p01



(c) Market planning for dantzig42p01

Fig. 9. Market planning to fulfill demands 1855, 2342, and 3561 for bayg29p01, bays29p01, and dantzig42p01, respectively.

4.4. CluTPP with scenario-wise goods transportation for a single item in deterministic parametric values

In this subsection for the empirical test, we took into account the CluTPP with scenario-wise goods transportation considering a single product and a single vehicle type. The following numerical data are used to illustrate the proposed model: Unit travel cost for the purchaser = \$1.5, c'_{ij} = \$1.0, and c_{k0} is randomly generated as being between \$1.3 and \$1.4. Then r_{pi} is randomly generated as being between \$10 and \$20; q_{pi} is randomly generated as being between 100 to 150; and C_k = \$150 if the transported goods exceed 300 units; otherwise \$100, C' = \$200.

Fig. 14 illustrates demand versus different costs (including the traveling cost, purchasing cost, transportation cost, and total system cost) scenario-wise. The transportation cost increases sharply when the demand is shifted from 1500 units to 2000 units in scenario 2. It also increases when the demand increases to 2500 and 3000 units. This observation agrees with reality, because the transportation cost massively increases when goods are carried with the purchaser in a different vehicle. Hence, the decision maker can select the mode of

Table 4
Performance Test of VLC-GA on CTSP Using TSPLIB Instances.

Instances	Optimum number of clusters	Best objective function value	Mean objective function value	SD	% of SD	Computation time (mean value in sec)
bays29	3	2160.00	2278.40	95.27	4.18	503.19
bayg29	3	1700.00	1848.80	99.79	5.40	510.12
dantzig42	3	795.00	849.60	41.15	4.84	511.76
eil51	3	768.00	809.00	23.25	2.87	544.22
berlin52	5	9410.00	9781.90	249.16	2.55	574.26
st70	4	1498.00	1571.90	46.62	2.97	645.26
pr76	5	133692.00	144030.20	6381.31	4.43	696.38
gr96	4	73519.00	78514.10	3410.24	4.34	815.29
rat99	3	2396.00	2591.50	166.24	6.41	816.25
rd100	4	10396.00	11052.20	534.90	4.84	820.17

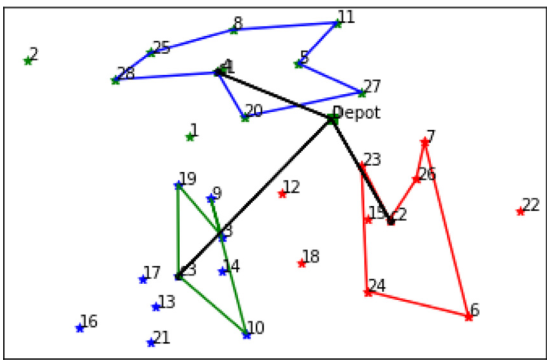


Fig. 10. Clusterwise routing and transporting of goods from center of cluster to the depot.

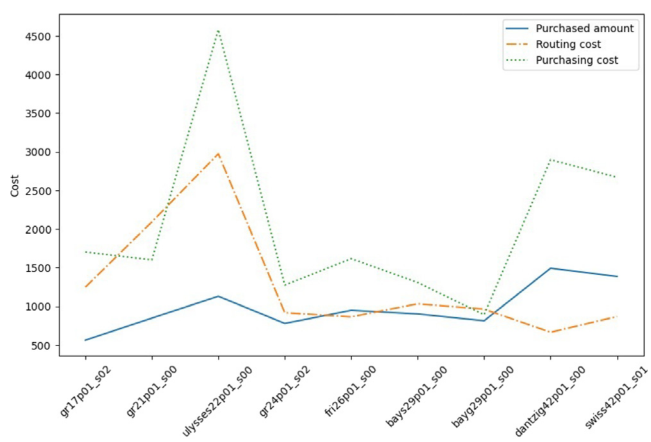


Fig. 11. Instance-wise routing and purchasing cost.

transportation (from the center of clusters to the depot or transported with the purchaser in a separate vehicle).

4.5. MVCluTPP for multi-item setup in deterministic parametric values

The following numerical data are used to illustrate the proposed model.

The unit travel cost for the purchaser = (\$1.5, \$1.3, \$1.6), $c'_{ijf'}$ = (\$1.0, \$1.2), $c^{(tc)}_{k0}$ = (\$1.3, \$1.4); r_{1i} is randomly generated as being between \$10 and \$20; r_{2i} is randomly generated as being between \$12 and \$22; r_{3i} is randomly generated as being between \$12 and \$20; q_{1i} is randomly generated as being between 30 and 50; q_{2i} is randomly generated as being between 40 and 70; q_{3i} is randomly generated as being between 50 and 70; and $C_{k(tc)}$ = (\$140, \$150, \$160) if the goods being transported exceed 300 units; otherwise (\$95, \$100, \$105), $C'_{f'}$ = (\$200, \$210).

Here, Fig. 15 also represents a similar observation to the one made in Fig. 14 (Section 4.4). This figure presents demand versus different costs scenario-wise. The transportation cost increases sharply with the demand when it is shifted from 1500 units to 2000 units and above in scenario 2, which agrees with reality. Hence, the decision maker can select the mode of transportation according to their demand.

4.6. MVCluTPP in a random environment

Here, $\lambda_1 = \frac{1}{100}$, $\lambda_2 = \frac{1}{25}$, R_{pi} is randomly generated as being between \$25 and \$35; R'_{pi} is randomly generated as being between \$5 and \$10; Q_{pi} is randomly generated as being between 120 and 150; and other numerical data have the same values as in Section 4.5.

Here, Fig. 16 represents similar observations of demand versus different costs, scenario-wise. Decision makers can choose the appropriate scenario for transporting goods according to their requirements.

4.7. Performance test of VLC-GA

In this subsection, we have shown the superiority of our developed VLC-GA with a local heuristic for cluster linking by considering some TSPLIB instances and instances generated in Section 4.3.

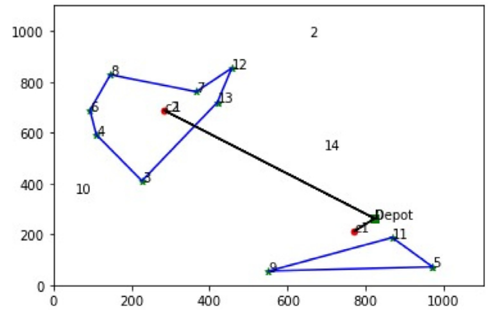
Performance test on CTSP using TSPLIB instances

To judge the effectiveness and feasibility of the proposed intelligent hybrid algorithm (VLC-GA) with a local heuristic for cluster linking, we applied it to 10 standard benchmarks using TSPLIB Reinelt (1995). Here, TSPLIB instances are solved by converting them into CTSP considering the optimum number of clusters (using the elbow method). Table 4 shows the results of 10 standard TSPLIB instances using VLC-GA with local heuristics by converting them into CTSP. The results are presented in terms of the total costs. Under 25 independent runs that include best-found results and average results with a percentage of SDs are presented here. In all cases, SD values given by the VLC-GA are very small, which indicates that this method is stable and that results in each run do not differ much from the mean. We also obtained the lowest percentage of SDs in most cases. It can be seen from Table 4 that for different TSPLIB instances, we received a better mean value, which implies that our proposed algorithm generates results that are closer to the best objective function value. Also, most of the SDs of the standard problems are convincible. Therefore, our developed VLC-GA algorithm is quite efficient and effective.

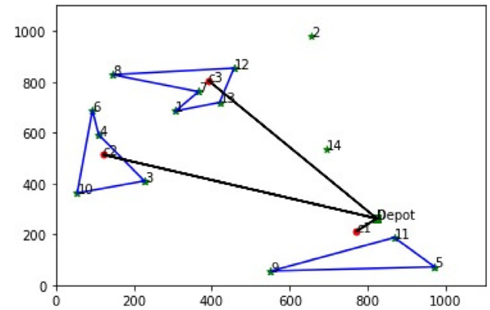
Performance test on CluTPP using TSPLIB instances

Here, we designed the classical TSP problems from TSPLIB and converted them into TPP and CluTPP by considering the availability and demand for the optimum number of clusters, using the elbow method. Here, for the numerical presentation, we consider the demand to be 7000 units and the price of the product to be randomly generated from between \$10 and \$20.

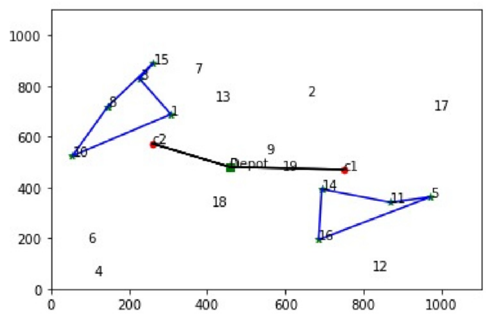
This subsection establishes the effectiveness of our developed VLC-GA meta-heuristics by calculating the mean, SD, and percentage of



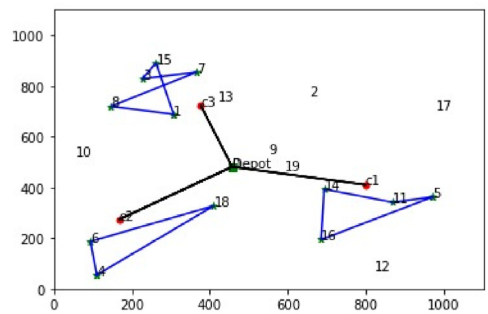
(a) Node: 15; Product: 02; Demand: 700



(b) Node: 15; Product: 03; Demand: 1,000



(c) Node: 20; Product: 02; Demand: 700



(d) Node: 20; Product: 03; Demand: 1,000

Fig. 12. Market planning to fulfill different demands of different products for randomly generated markets for small instances.

Table 5
Performance Test of VLC-GA on CluTPP Using TSPLIB Instances.

Instances	Optimum number of clusters	Best objective function value	Mean objective function value	SD	% of SD	Computation time (mean value in sec)
kroB100	4	116 734.00	119 935.25	2169.88	1.81	822.34
kroC100	4	138 206.00	145 229.63	5393.57	3.71	802.54
kroD100	4	123 798.00	127 125.13	2490.66	1.96	820.55
kroE100	4	136 585.00	141 627.63	2704.65	1.91	822.86

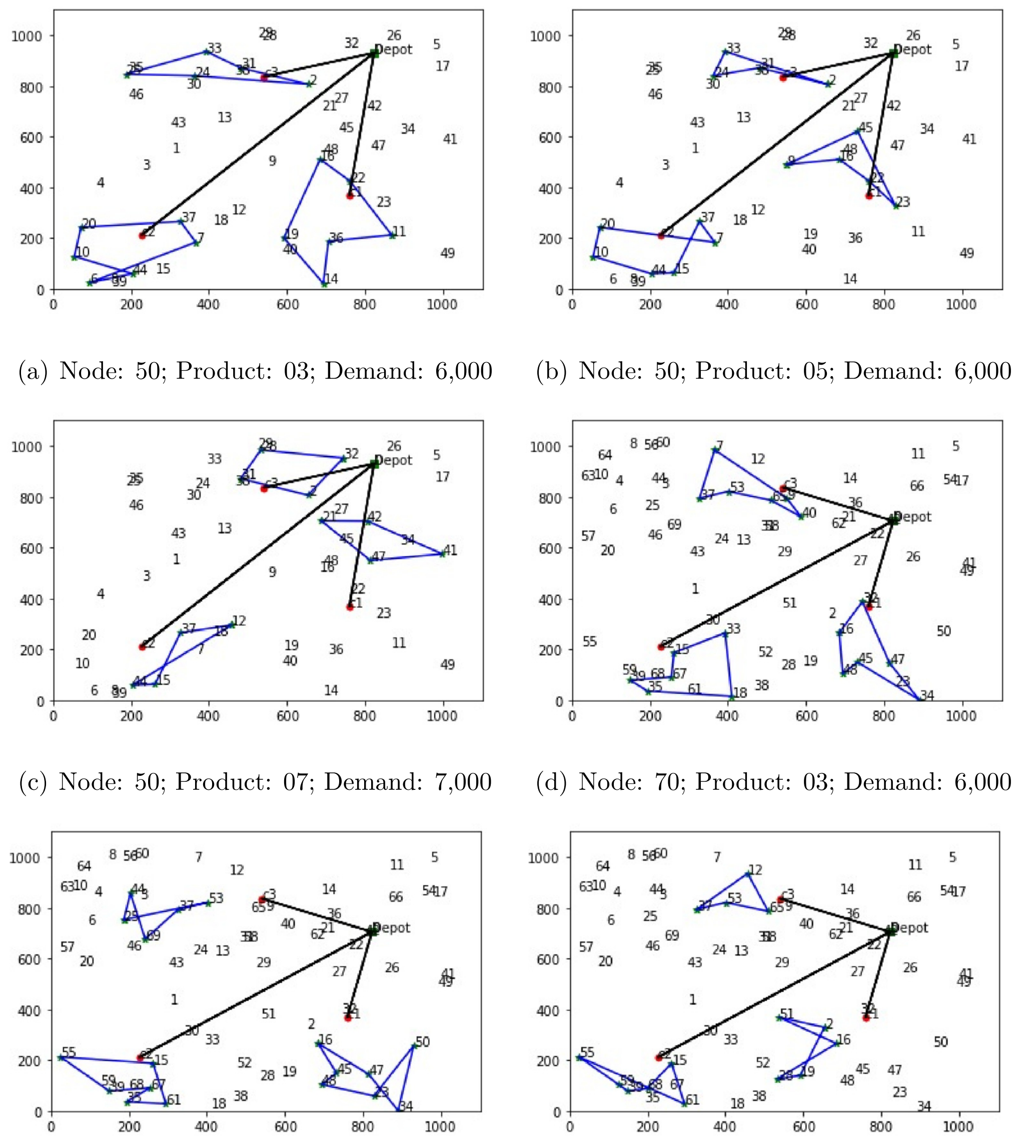
Table 6
Comparison of Exact Algorithm and VLC-GA in CluTPP, MVCluTPP (Multi-item), and MVCluTPP (Multi-item with random).

Model	Demand	Total cost (Exact algorithm)				Total cost (VLC-GA)				Deviation			
		Scenario 1	Runtime (s)	Scenario 2	Runtime (s)	Scenario 1	Runtime (s)	Scenario 2	Runtime (s)	Total cost Scenario 1	Scenario 2	Runtime (s) Scenario 1	Scenario 2
CluTPP (single item)	1000	19 565.97	10 800	16 838.78	10 800	15 852.87	620.22	14 441.40	658.48	−18.98%	−14.24%	−94.26%	−93.90%
	1500	30 925.09	10 800	27 056.41	10 800	24 271.60	686.11	22 896.07	725.93	−21.51%	−15.38%	−93.65%	−93.28%
	2000	35 936.26	10 800	42 211.94	10 800	33 472.71	742.89	33 978.90	780.20	−06.86%	−19.50%	−93.12%	−92.78%
	2500	44 730.81	10 800	56 396.16	10 800	44 345.77	804.22	47 141.59	842.26	−0.86%	−16.41%	−92.55%	−92.20%
	3000	56 594.01	10 800	73 297.23	10 800	53 908.52	865.98	57 325.87	905.29	−04.75%	−21.79%	−91.98%	−91.62%
MVCluTPP (multi-item)	1000	21 522.24	10 800	19 237.36	10 800	19 163.53	630.29	17 587.78	662.20	−10.96%	−08.57%	−94.16%	−93.87%
	1500	34 242.24	10 800	31 660.23	10 800	31 156.23	698.26	29 443.20	739.58	−09.01%	−07.00%	−93.53%	−93.15%
	2000	38 005.20	10 800	41 738.39	10 800	35 520.74	756.25	38 682.46	797.25	−06.54%	−07.32%	−93.00%	−92.62%
	2500	51 923.36	10 800	55 002.61	10 800	48 314.38	825.02	51 554.43	863.92	−06.95%	−06.27%	−92.36%	−92.00%
	3000	58 299.22	10 800	68 237.27	10 800	56 012.26	972.22	60 756.19	1032.66	−03.92%	−10.96%	−90.99%	−90.44%
MVCluTPP (multi-item) with random	1000	20 899.52	10 800	18 932.66	10 800	18 379.01	640.21	16 974.80	670.23	−12.06%	−10.34%	−94.07%	−93.79%
	1500	32 600.22	10 800	28 437.36	10 800	28 397.77	704.23	26 753.54	747.22	−12.89%	−05.92%	−93.48%	−93.08%
	2000	37 005.52	10 800	40 245.98	10 800	33 944.22	751.48	37 305.50	797.62	−08.27%	−07.31%	−93.04%	−92.61%
	2500	51 237.84	10 800	53 365.09	10 800	47 232.08	988.86	50 452.66	1024.42	−07.81%	−05.46%	−90.84%	−90.52%
	3000	60 038.29	10 800	65 092.88	10 800	54 648.86	1042.22	59 006.28	1089.29	−08.98%	−09.35%	−90.35%	−89.91%

SD of the 10 best near-optimal solutions. The performance of the proposed meta-heuristics was statistically tested by running it 25 times independently and calculating the average value, SD, and percentage of SD by considering the 10 best solutions against 25 solutions. Results are given in Table 5, and for all cases, we received better results. Our given mean objective function value for all TSPLIB instances is very close to their corresponding best objective function value with minimum percentage SD, which proves the superiority of our developed VLC-GA.

Parametric analysis

Fig. 17 shows that the performance of the VLC-GA depends on the size of the population (pop_size). The size of the population affects the total cost, such as the total cost decreases while population size increases, and the converging point is 65. The same observations happen in the number of generations. The total cost decreases while the number of generations increases, and the converging point is near 500. A study on exact approaches with a comparison of proposed algorithms is shown in Table 6, where the instances have been taken as per the description in Section 4.3 and the exact approaches perform within



(a) Node: 50; Product: 03; Demand: 6,000 (b) Node: 50; Product: 05; Demand: 6,000

(c) Node: 50; Product: 07; Demand: 7,000 (d) Node: 70; Product: 03; Demand: 6,000

(e) Node: 70; Product: 05; Demand: 6,000 (f) Node: 70; Product: 07; Demand: 7,000

Fig. 13. Market planning to fulfill different demands of different products for randomly generated markets for large instances.

three hours limitation. The performance regarding the running time according to Table 6 shows that the proposed VLC-GA is ahead in each case of the scenarios and the different demand settings.

Comparison of Exact algorithm and VLC-GA

Regarding solution quality (total cost) and running times with a variation of different demand settings, Table 6 shows that the supremacy of the proposed VLC-GA is effective in finding solutions within a reasonable time considering the standard TSPLIB instance, bays29. An exact algorithm has been implemented with optimization programming language (OPL) and solved with the CPLEX solver using the exact branch-and-cut algorithm. VLC-GA is compared with the exact algorithm using the same parameters. Detailed optimal procurement planning (sequence of markets, amount of purchased products) with different scenarios under varying demand settings using VLC-GA available is shown in Tables 7, 8, and 9 (Appendix A).

Comparisons among the proposed VLC-GA, Classical GA, and ACO

Here, Fig. 18 is depicted with different demand settings, system cost variations with the proposed VLC-GA, classical GA (roulette wheel selection, cyclic crossover, and random mutation), and ACO considering the standard TSPLIB instance, bays29. In most of the cases proposed, VLC-GA is ahead of other used algorithms. The study on the proposed CluTPP based on demand variations through classical GA, ACO, and proposed VLC-GA shows the supremacy of the algorithm. In some cases, scenario 1 gives better results, but when demand increases, scenario 2 gives better results ahead of other scenarios. Thus, a trade-off between family shopping versus bulk purchase planning (business shopping) emerged and is depicted in Fig. 18.

4.8. Discussion and managerial insights

In each case, we could find some meaningful results scenario-wise. For all cases, the optimal number of clusters is 3 (for bays29). In the CluTPP, for a single item in a deterministic environment, market selection with the purchased amount, different costs, and the total cost for different demand values are all presented in Table 7. In the MV-CluTPP, for multi-item goods in a deterministic environment, market selection with the purchased amount, different costs, and the total cost for different demand values are all presented in Table 8, and Table 9

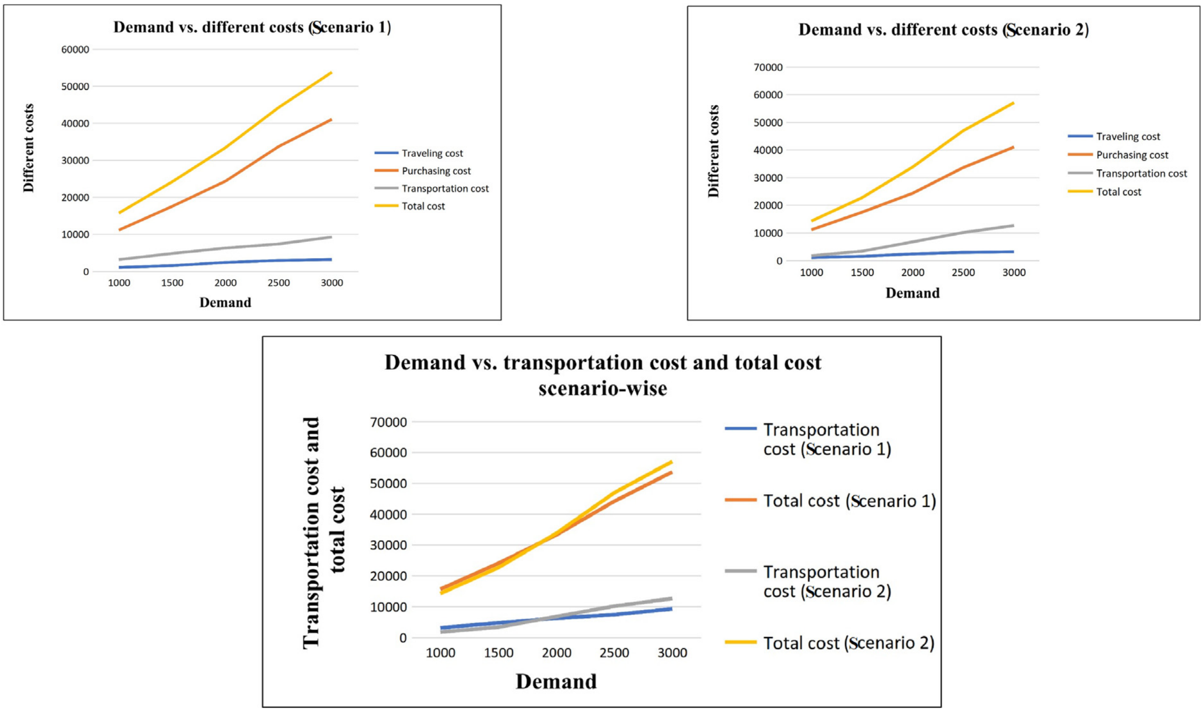


Fig. 14. Demand vs. different costs, scenario-wise, for a single item.

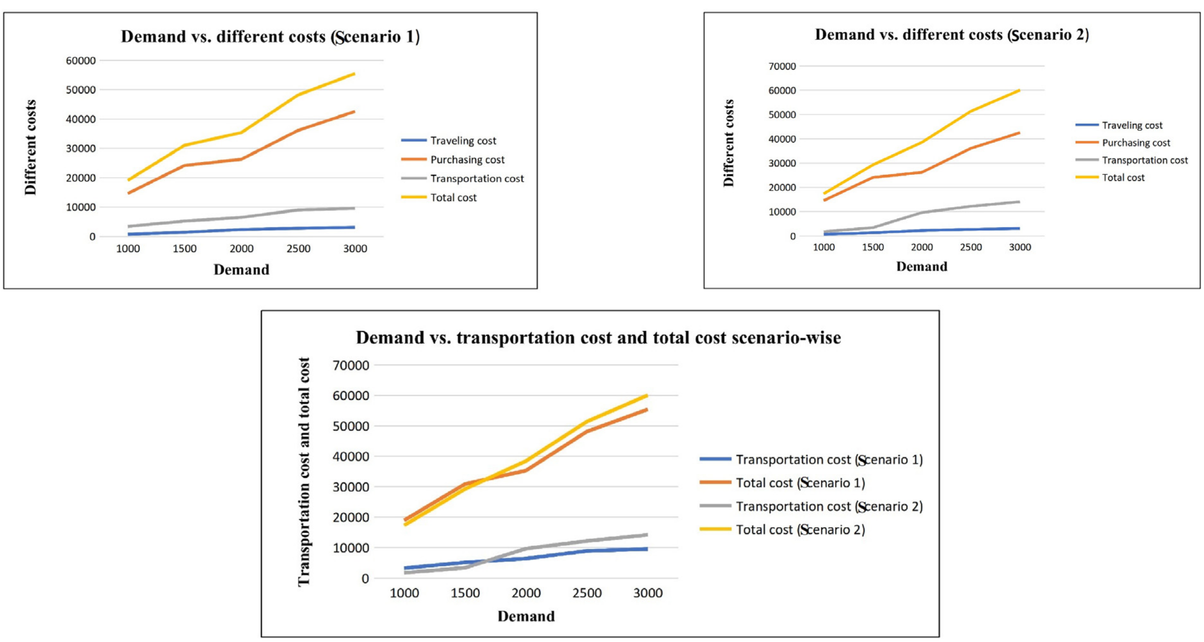


Fig. 15. Demand vs. different costs, scenario-wise, for multi-item goods.

represents corresponding values for the MVCluTPP model for multi-item goods in a random environment. We found that the travel cost is minimized for the MVCluTPP model compared to the CluTPP model, which agrees with reality, because of the choice of conveyances. Our developed model can reduce costs because of the suitable choice of the purchaser's vehicles, as well as the proper choice of transport vehicles. Experimental results illustrate the developed MVCluTPP model, product-wise, in a random environment. The MVCluTPP model can be applied in situations in which different types of products are available in different markets with varying prices, adhering to the quality of randomness, which agrees with reality. Market clustering is one of the most important strategies in the marketing process. Market clustering is a technique generally used for clustering similar sub-markets. This technique is used to enhance the decision-making abilities of consumers and retailers. Market clustering can separate a group of markets into smaller subgroups whose objects are more similar according to a set of criteria than are objects assigned to another cluster. The above-proposed idea can segment customer databases to identify different types of usage, revenue, cost, and profitability segments. Market clustering does not need to be complicated; it just needs to be useful. This is a powerful technique used to improve business understanding of the market. The clustering method can be beneficial, and it needs to be made relevant

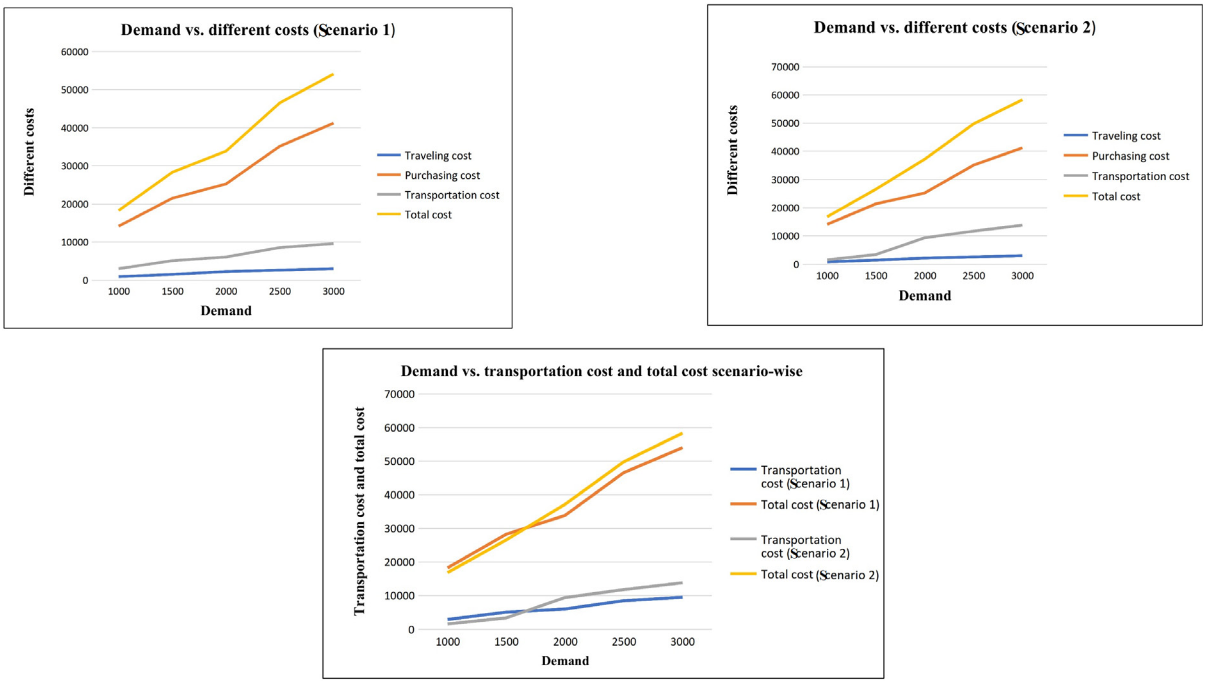
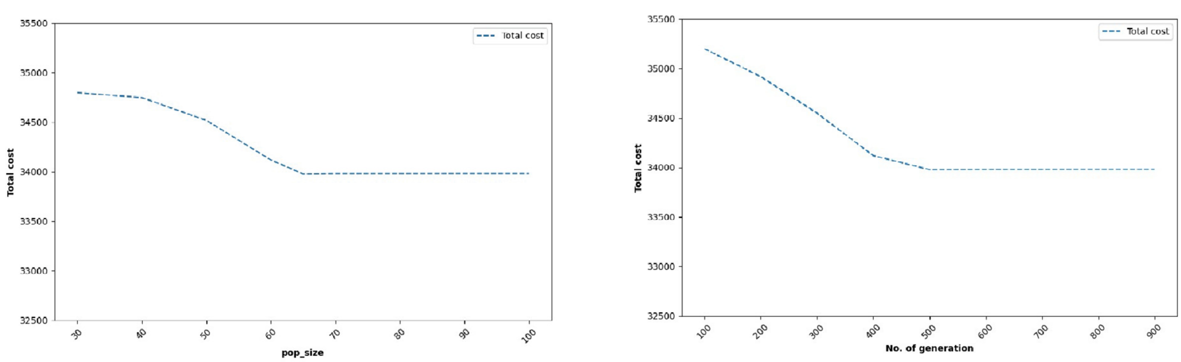


Fig. 16. Demand vs. different costs, scenario-wise, in a random environment.



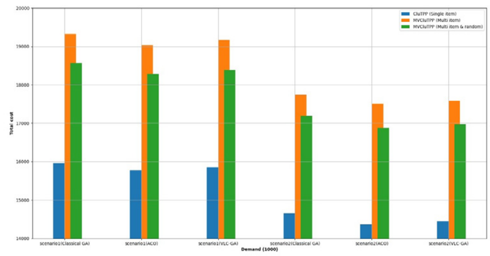
(a) Population size vs. total cost (b) Number of generations vs. total cost

Fig. 17. Population size and number of generations vs. total cost.

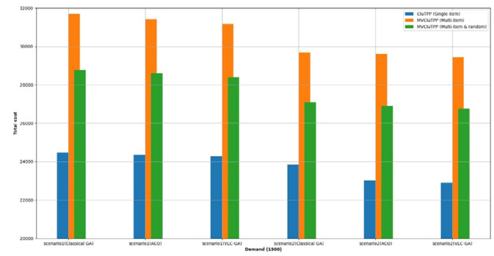
to procurement planning. To be more specific, whatever the particular need is, the more likely it is that clustering will meet that need better than will any other random procurement strategy. Universal clustering often yields specific segments that offer some real business-relevant insights, but a well-designed clustering strategy will result in better procurement planning and will help decision makers choose the best path for marketing to customers and planning for purchases. Also, our proposed k-means clustering technique was able to quickly cluster large datasets. Here, researchers can define the number of clusters through a suitable technique to perform an actual study. This method is most useful for testing different types of models with a varying number of clusters. Our method can also be used to identify homogeneous groups of buyers. The buying behavior of a group of like-minded customers can be studied by measuring favorite stores, brand loyalty, frequency of purchases, and price-point-to-purchase decisions. Our proposed model can be directly involved with many real-life applications, such as production scheduling and manufacturing systems. The TPP can be designed for many network applications, such as rail and subway lines, irrigation networks, and other networks. Applications of TPP can also be specified in telecommunication networks, industrial design, and generic communication networks. Such infrastructures consist of several local access networks (LANs) that collect traffic of user nodes at the switching centers and of a backbone network that routes high volume traffic among switching centers.

5. Conclusions

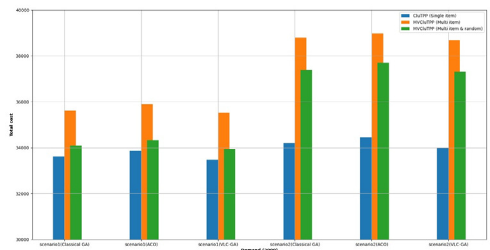
In this study, for the first time, we illustrated, formulated, and represented an MVCluTPP in a deterministic and random environment. The classical TPP was rendered more realistic by introducing setups such as the following: Different types of conveyances were made available at market locations for the purchaser's traveling purposes, for transporting goods alongside the purchaser, and also for transporting goods from the center of clusters, scenario-wise. Two alternatives for transporting goods to the depot (i.e., transporting goods from the market to the depot directly and transporting goods along with the purchaser for the entire route) were considered, and the better choice was selected. Our study compared the conventional purchasing process against disjoint travel scenarios and transportation plans. Suggested conveyance types for travel and transportation include those that offer more flexibility



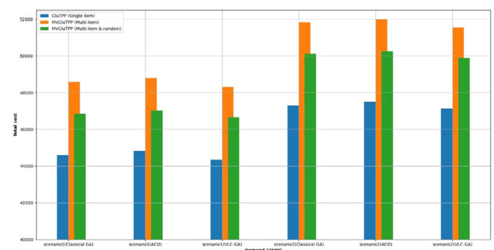
(a) Scenario-wise total cost for demand 1,000



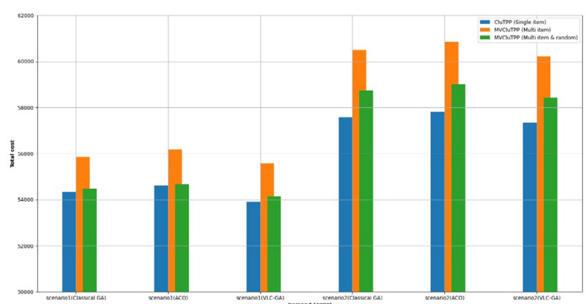
(b) Scenario-wise total cost for demand 1,500



(c) Scenario-wise total cost for demand 2,000



(d) Scenario-wise total cost for demand 2,500



(e) Scenario-wise total cost for demand 3,000

Fig. 18. Scenario-wise total cost for Classical GA, ACO, and VLC-GA for different demands.

to the decision maker by allowing them to make proper choices to reduce costs. The developed MVCluTPP was formulated in deterministic and random cost parameters and was solved using the proposed VLC-GA with k-means clustering. Here, we successfully established the superiority of the VLC-GA with probabilistic selection, multi-parent comparison crossover, and sigmoid random mutation, compared with exact methods (CPLEX solver). Our investigation is directly connected to real-life purchasing and distribution problems. These results can be used in other optimization areas, like multi-vehicle transportation problems, production policy studies, large-scale integration vehicle routing strategies, and other fields. This research work gives a new variant of TPP, termed clustered TPP, which is a new term in the literature. Although we attempted to include certain practical features in our model, we could not accommodate all eventualities, and remaining limitations can be studied further. For example, in our developed MVCluTPP, one central depot is considered. It might be more applicable and interesting to extend the model to a multi-depot problem. Also, such a study might consider multiple paths between two markets, and along each path, multiple conveyances could be available. Thus, expanding our model to explore these possibilities might be an interesting extension area that researchers could investigate in the future.

CRediT authorship contribution statement

Arindam Roy: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Samir Maity:** Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Validation. **Ilkyeong Moon:** Writing – review & editing, Validation, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgments

This research is supported by the Brain Pool Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, and Future Planning [Grant No. NRF-2020H1D3A2A01085443]. The authors are grateful for the valuable comments from the associate editor and three anonymous reviewers.

Appendix A

According to Table 7, three clusters are taken, and for the demand of 1000 units, the optimal path becomes 10 →9 →3 →23 →7 →8 →11 →5 →27. The traveling cost is $1215.00; the purchasing cost is $11,298.00; and scenario-wise, transportation costs are $3339.87 (scenario 1) and $1928.40 (scenario 2). The overall system cost in scenario 1 is greater, at $15,852.87, than it is in scenario 2 (where it is $14,441.40). Similarly, for the demand of 1500 units, the same observations can be found in such a way that the optimal path becomes 3 →9 →10 →21 →24 →23 →7 →20 →11 →5 →27. The traveling cost increases to $1683.00, because more markets are visited, and the overall system cost also is greater in scenario 1, at $24,271.60, than it is in scenario 2 (where it is $22,896.07). When demand has increased to 2000 units, the optimal path becomes 10 →21 →16 →9 →19 →3 →24 →6 →23 →7 →20 →8 →11 →25 →5 →27. The traveling cost increases to $2545.50, because more markets are visited to fulfill more demand, and the transportation cost in scenario 1 ($6435.21) is less than it is in scenario 2 (where it is $6941.40). Furthermore, the overall system cost in scenario 1 ($33,472.71) is minimized, compared to scenario 2 (where it is $33,978.90). Thus, a demand for 2000 units of goods transported directly from the markets after being purchased is beneficial, cost-wise, for the procurement planner, when compared to the planner's carrying the same unit of goods in a different vehicle. A similar observation can be found when the demand increases to 2500 units and 3000 units, which can be seen in Table 7.

Section 4.5 considers the MVCluTPP with multi-products or items in a deterministic environment. Here, we observed similar results to those in Section 4.4. Table 8 represents clusterwise market selection and the sequences of those selections with the amount of purchased quantity, item-wise, from different markets. Also, the optimal path of travel (which includes the traveling cost, the purchasing cost, the transportation cost, and the total cost) is presented scenario-wise. From the presented results in Table 8, we can see that the travel cost is minimized compared to its iteration in Table 7, because of the choice by the purchaser of using multiple vehicles. Also, in this model, when the demand is for 2000 units or more, the transportation cost, as well as the total cost, increases in scenario 2 (where goods are transported with the purchaser in a separate vehicle).

Section 4.6 considers the MVCluTPP with multi-products or items in a random environment. In reality, the availability and price of products are not always fixed. Here, we have considered the availability and

Table 7
Optimal path and different costs of CluTPP for a single item (No. of clusters=3).

Demand	Cluster	Market	Visited market	Optimal path	Traveling	Purchasing	Transportation cost		Total cost	
			(purchased quantity)		cost	cost	Scenario 1	Scenario 2	Scenario 1	Scenario 2
	1	6, 7, 12, 15, 22, 23, 24, 26	23 (146), 7(59)							
1000	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	27(26), 5(146), 11(144), 8(120)	10 →9 →3 →23 →7 →	1215.00	11 298.00	3339.87	1928.40	15 852.87	14 441.40
	3	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	9(89), 3(138), 10(132)	8 →11 →5 →27						
	1	6, 7, 12, 15, 22, 23, 24, 26	7(147), 23(138), 24(137)							
1500	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	20(134), 27(138), 5(147), 11(127)	3 →9 →10 →21 →24 →23 →	1683.00	17 651.00	4937.60	3562.07	24 271.60	22 896.07
	3	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	3(138), 9(147), 10(132), 21(115)	7 →20 →11 →5 →27						
	1	6, 7, 12, 15, 22, 23, 24, 26	24(132), 6(146), 23(138), 7(147)							
2000	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	11(146), 8(144), 20(134), 27(138), 5(147), 25(18)	10 →21 →16 →9 →19 →3 →24 →6 →	2545.50	24 492.00	6435.21	6941.40	33 472.71	33 978.90
	3	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	9(147), 19(124), 3(132), 10(138), 21(144), 16(25)	23 →7 →20 →8 →11 →25 →5 →27						
	1	6, 7, 12, 15, 22, 23, 24, 26	23(92), 24(143), 6(146), 7(147), 15(122)	3 →9 →10 →21 →13 →16 →24 →						
2500	2	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	3(132), 9(147), 10(138), 21(144), 13(124), 16(30)	6 →7 →15 →23 →1 →20 →5 →	3084.00	33 700.00	7561.77	10 357.59	44 345.77	47 141.59
	3	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	27(45), 8(144), 28(123), 2(135), 25(120), 11(146), 5(147), 20(134), 1(141)	11 →25 →2 →28 →8 →27						
	1	6, 7, 12, 15, 22, 23, 24, 26	26(146), 15(143), 24(122), 23(138), 12(133), 7(147), 6(15)	16 →13 →14 →21 →10 →9 →19 →3 →						
3000	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	25(118), 28(123), 8(144), 5(147), 11(146), 27(138), 20(134), 1(141)	12 →23 →24→15 →26 →6 →7 →	3288.00	41 193.00	9427.52	12 844.87	53 908.52	57 325.87
	3	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	16(132), 13(147), 14(124), 21(121), 10(122), 9(124), 19(144), 3(138)	11 →5 →8 →28 →25→1 →20 →27						

Table 8
Optimal path and different costs of MVCTPP for multi-item goods (No. of clusters=3).

Demand	Cluster	Market	Visited market	Optimal path	Traveling	Purchasing	Transportation cost		Total cost	
			(purchased quantity item-wise)		cost	cost	Scenario 1	Scenario 2	Scenario 1	Scenario 2
	1	3, 9, 10, 13, 14, 16, 17, 19, 21	13(44, 59, 66), 3(34, 43, 62), 14(34, 43, 62)							
1000	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	20(4, 56, 67), 5(4, 64, 65), 25(4, 62, 70), 28 (40, 62, 70)	13 →14 →3 →23 →	916.50	14 751.00	3496.03	1920.28	19 163.53	17 587.78
	3	6, 7, 12, 15, 18, 22, 23, 24, 26	23(48, 42, 55)	20 →28 →25 →5						

(*continued on next page*)

Table 8 (*continued*).

Demand	Cluster	Market	Visited market (purchased quantity item-wise)	Optimal path	Traveling cost	Purchasing cost	Transportation cost Scenario 1	Transportation cost Scenario 2	Total cost Scenario 1	Total cost Scenario 2
1500	1	6, 7, 12, 15, 22, 23, 24, 26	22(46, 49, 52), 6(46, 49, 52), 7(37, 58, 67)							
	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	27(35, 45, 55), 1 (44, 56, 67), 20(47, 64, 65), 5(44, 56, 67)	21 →14 →17 →3 →6 →22	1508.00	24 309.00	5339.23	3626.20	31 156.23	29 443.20
	3	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	21(34, 43, 62), 14(43, 62, 65), 17(37, 57, 52), 3(37, 57, 52)	7 →5 →20 →1 →27						
2000	1	6, 7, 12, 15, 22, 23, 24, 26	23(48, 42, 55), 24(43, 54, 50), 12(33, 66, 59), 15(33, 66, 59)							
	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	8(34, 46, 70), 25(40, 62, 70), 20(44, 56, 67), 5(44, 56, 67)	21 →10 →13 →14 →9 →12→	2459.20	26 401.00	6660.54	9822.26	35 520.74	38 682.46
	3	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	21(34, 43, 62), 9(47, 66, 62), 14(43, 62, 65), 13(44, 59, 66), 10(44, 59, 66)	15 →23 →24 →25 →8 →5 →20						
2500	1	6, 7, 12, 15, 22, 23, 24, 26	15(43, 54, 50), 12(32, 53, 65), 7(33, 66, 59), 26(37, 58, 67)	17 →19 →13 →16 →21 →14 →						
	2	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	19(44, 49, 52), 13, 16, 21(34, 43, 52), 14, 3, 18, 17(37, 57, 52)	3 →18 →12 →15 →26 →7 →	2926.60	36 254.00	9133.78	12 373.83	48 314.38	51 554.43
	3	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	1(44, 56, 67), 20(44, 56, 67), 8(34, 46, 70), 11(46, 55, 53), 27(38, 69, 52)	11 →8 →20 →1 →27						
3000	1	6, 7, 12, 15, 22, 23, 24, 26	12(46, 49, 52), 23, 6(48, 42, 55), 15, 24, 22(43, 54, 50)	22 →24 →15 →6 →23 →12 →16 →						
	2	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	20(44, 56, 67), 5, 4, 25(40, 62, 70), 27, 8, 1(35, 45, 55)	10 →17 →19 →9 →14 →18 →1 →	3299.00	42 965.00	9748.26	14 492.19	56 012.26	60 756.19
	3	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	16(38, 52, 63), 10, 17, 19(44, 49, 52), 9, 14, 18(43, 62, 65)	8 →27 →25 →4 →5 →20						

price of products as random, and follow an exponential distribution. Also, we observed similar results, as in Section 4.4. Table 9 represents clusterwise market selection and the sequences of these selections with the amount of purchased quantity, item-wise, from different markets in a random environment. Also, the optimal path of travel (which includes the traveling cost, the purchasing cost, the transportation cost, and the total cost) is presented scenario-wise. From the results in Table 9, we can see that the travel cost is minimized, compared to its iteration in Table 7, because of the choice by the purchaser of using multiple vehicles. Also, this subsection depicted similar observations of the transportation cost for different demands, scenario-wise.

Table 9
Optimal path and different costs of MVCluTPP for multi-item goods in a random environment (No. of clusters=3).

Demand	Cluster	Market	Visited market (purchased quantity item-wise)	Optimal path	Traveling cost	Purchasing cost	Transportation cost Scenario 1	Transportation cost Scenario 2	Total cost Scenario 1	Total cost Scenario 2
1000	1	6, 7, 12, 15, 18, 22, 23, 24, 26	24(49, 58, 60), 18(0, 55, 59), 15(0, 54, 22)	24 →15 →18 →9 →						
	2	3, 9, 10, 13, 14, 16, 17, 19, 21	9(54, 64, 50)	5 →8 →25 →4	1001.60	14 272.76	3104.65	1700.44	18 379.01	16 974.80
	3	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	5(12, 45, 50), 4(12, 51, 64), 25(12, 50, 49), 8(52, 51, 64)							
1500	1	6, 7, 12, 15, 22, 23, 24, 26	12(47, 58, 45), 15(47, 58, 45), 26(47, 58, 60)							
	2	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	3(41, 53, 64), 9(55, 60, 57), 17(49, 62, 45), 14(55, 60, 57)	26 →15 →12 →14 →17 →9 →	1615.50	21 598.79	5183.48	3539.25	28 397.77	26 753.54
	3	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	11(56, 51, 52), 20(56, 51, 52), 27(41, 46, 58), 8(56, 51, 52)	3 →11 →8 →27 →20						
2000	1	6, 7, 12, 15, 18, 22, 23, 24, 26	23(53, 12, 47), 26(52, 58, 49), 6(45, 53, 55), 15(54, 0, 12)							
	2	3, 9, 10, 13, 14, 16, 17, 19, 21	16(65, 32, 60), 17(65, 32, 60), 14(65, 61, 60)	14 →17 →16 →15 →6 →26 →	2385.60	25 375.10	6183.52	9544.80	33 944.22	37 305.50
	3	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	25(52, 61, 61), 11, 27(48, 64, 52), 4, 1, 5(61, 65, 47)	23 →27 →11 →25 →5 →1 →4						
2500	1	6, 7, 12, 15, 22, 23, 24, 26	12(47, 58, 60), 15(47, 58, 60), 6(62, 67, 56), 7(59, 58, 57)							
	2	3, 9, 10, 13, 14, 16, 17, 18, 19, 21	17(51, 55, 34), 9, 14, 10(41, 53, 64), 19, 16(58, 56, 46)	17 →16 →19 →10 →14 →9 →12 →27	2790.62	35 551.96	8889.50	12 110.08	47 232.08	50 452.66
	3	1, 2, 4, 5, 8, 11, 20, 25, 27, 28	1(56, 65, 61), 4, 11, 27(41, 46, 58), 20, 28(62, 59, 34)	15 →6 →7 →11 →4 →1 →28 →20						

(*continued on next page*)

Table 9 (*continued*).

Demand	Cluster	Market	Visited market (purchased quantity item-wise)	Optimal path	Traveling cost	Purchasing cost	Transportation cost Scenario 1	Transportation cost Scenario 2	Total cost Scenario 1	Total cost Scenario 2
3000	1	6, 7, 12, 15, 22, 23, 26, 27	23(58, 62, 54), 7, 15(58, 57, 55), 22, 12(57, 59, 56)	21 →14 →10 →17 →3 →13 →						
	2	3, 9, 10, 13, 14, 16, 17, 18, 21, 24	10(59, 58, 53), 14, 21, 24(60, 51, 46), 13, 3, 17(48, 60, 37)	24 →12 →22 →15 →7 →23 →	3157.86	41 650.18	9840.82	14 198.24	54 648.86	59 006.28
	3	1, 2, 4, 5, 8, 11, 19, 20, 25, 28	28(49, 52, 57), 2, 20, 5(46, 59, 52), 4, 8, 1(56, 57, 56)	2 →28 →1 →8 →4 →5 →20						

References

Abreu, L.R., Tavares-Neto, R.F., Nagano, M.S., 2021. A new efficient biased random key genetic algorithm for open shop scheduling with routing by capacitated single vehicle and makespan minimization. Eng. Appl. Artif. Intell. 104, 104373.

Adan, I., Landa-Torres, I., Portillo, E., Manjarres, D., 2022. Influence of statistical feature normalisation methods on K-nearest neighbours and K-means in the context of industry 4.0. Eng. Appl. Artif. Intell. 111, 104807. http://dx.doi.org/10.1016/j.engappai.2022.104807.

Almeida, C.P., Gonçalves, R.A., Goldbarg, E.F., Goldbarg, M.C., Delgado, M.R., 2012. An experimental analysis of evolutionary heuristics for the biobjective traveling purchaser problem. Ann. Oper. Res. 199 (1), 305–341.

Angelelli, E., Gendreau, M., Mansini, R., Vindigni, M., 2017. The traveling purchaser problem with time-dependent quantities. Comput. Oper. Res. 82, 15–26.

Angelelli, E., Mansini, R., Vindigni, M., 2011. Look-ahead heuristics for the dynamic traveling purchaser problem. Comput. Oper. Res. 38 (12), 1867–1876.

Angelelli, E., Mansini, R., Vindigni, M., 2016. The stochastic and dynamic traveling purchaser problem. Transp. Sci. 50 (2), 642–658. http://dx.doi.org/10.1287/trsc.2015.0627.

Batista-Galván, M., Riera-Ledesma, J., Salazar-González, J.J., 2013. The traveling purchaser problem, with multiple stacks and deliveries: A branch-and-cut approach. Comput. Oper. Res. 40 (8), 2103–2115. http://dx.doi.org/10.1016/j.cor.2013.02.007.

Bernardino, R., Paias, A., 2018. Metaheuristics based on decision hierarchies for the traveling purchaser problem. Int. Trans. Oper. Res. 25 (4), 1269–1295.

Bianchessi, N., Irnich, S., Tilk, C., 2021. A branch-price-and-cut algorithm for the capacitated multiple vehicle traveling purchaser problem with unitary demand. Discrete Appl. Math. 288, 152–170.

Bianchessi, N., Mansini, R., Speranza, M.G., 2014. The distance constrained multiple vehicle traveling purchaser problem. European J. Oper. Res. 235 (1), 73–87.

Boctor, F.F., Laporte, G., Renaud, J., 2003. Heuristics for the traveling purchaser problem. Comput. Oper. Res. 30 (4), 491–504.

Bontoux, B., Feillet, D., 2008. Ant colony optimization for the traveling purchaser problem. Comput. Oper. Res. 35 (2), 628–637.

Cheaitou, A., Hamdan, S., Larbi, R., Alsyouf, I., 2020. Sustainable traveling purchaser problem with speed optimization. Int. J. Sustain. Transp. 1–20.

Chisman, J.A., 1975. The clustered traveling salesman problem. Comput. Oper. Res. 2 (2), 115–119. http://dx.doi.org/10.1016/0305-0548(75)90015-5.

Choi, M.J., Lee, S.H., 2011. The multiple traveling purchaser problem for maximizing system's reliability with budget constraints. Expert Syst. Appl. 38 (8), 9848–9853. http://dx.doi.org/10.1016/j.eswa.2011.02.018.

Cuellar-Usaquén, D., Gomez, C., Álvarez-Martínez, D., 2021. A GRASP/Path-Relinking algorithm for the traveling purchaser problem. Int. Trans. Oper. Res..

Gendreau, M., Manerba, D., Mansini, R., 2016. The multi-vehicle traveling purchaser problem with pairwise incompatibility constraints and unitary demands: A branch-and-price approach. European J. Oper. Res. 248 (1), 59–71. http://dx.doi.org/10.1016/j.ejor.2015.06.073.

Golden, B., Levy, L., Dahl, R., 1981. Two generalizations of the traveling salesman problem. Omega 9 (4), 439–441.

Gouveia, L., Paias, A., Voß, S., 2011. Models for a traveling purchaser problem with additional side-constraints. Comput. Oper. Res. 38 (2), 550–558. http://dx.doi.org/10.1016/j.cor.2010.07.016.

Hamdan, S., Larbi, R., Cheaitou, A., Alsyouf, I., 2017. Green traveling purchaser problem model: A bi-objective optimization approach. In: 2017 7th International Conference on Modeling, Simulation, and Applied Optimization. ICMSAO, IEEE, pp. 1–6.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A k-means clustering algorithm. J. R. Stat. Soc. Ser. C. Appl. Stat. 28 (1), 100–108. http://dx.doi.org/10.2307/2346830.

Kang, S., Ouyang, Y., 2011. The traveling purchaser problem with stochastic prices: Exact and approximate algorithms. European J. Oper. Res. 209 (3), 265–272. http://dx.doi.org/10.1016/j.ejor.2010.09.012.

Kucukoglu, I., 2022. The traveling purchaser problem with fast service option. Comput. Oper. Res. 105700.

Kuo, R.J., Ho, L.M., Hu, C.M., 2002. Integration of self-organizing feature map and K-means algorithm for market segmentation. Comput. Oper. Res. 29 (11), 1475–1493. http://dx.doi.org/10.1016/S0305-0548(01)00043-0.

Laporte, G., Riera-Ledesma, J., Salazar-González, J.-J., 2003. A branch-and-cut algorithm for the undirected traveling purchaser problem. Oper. Res. 51 (6), 940–951.

Li, Y., Chu, X., Tian, D., Feng, J., Mu, W., 2021. Customer segmentation using K-means clustering and the adaptive particle swarm optimization algorithm. Appl. Soft Comput. 113, 107924.

Liao, K., Guo, D., 2008. A clustering-based approach to the capacitated facility location problem. Trans. GIS 12 (3), 323–339. http://dx.doi.org/10.1111/j.1467-9671.2008.01105.x.

Maity, S., Roy, A., Maiti, M., 2015. A modified genetic algorithm for solving uncertain constrained solid travelling salesman problems. Comput. Ind. Eng. 83, 273–296. http://dx.doi.org/10.1016/j.cie.2015.02.023.

Manerba, D., Mansini, R., 2015. A branch-and-cut algorithm for the multi-vehicle traveling purchaser problem with pairwise incompatibility constraints. Networks 65 (2), 139–154. http://dx.doi.org/10.1002/net.21588.

Manerba, D., Mansini, R., Riera-Ledesma, J., 2017. The traveling purchaser problem and its variants. European J. Oper. Res. 259 (1), 1–18. http://dx.doi.org/10.1016/j.ejor.2016.12.017.

Mansini, R., Savelsbergh, M.W.P., Tocchella, B., 2012. The supplier selection problem with quantity discounts and truckload shipping. Omega 40 (4), 445–455. http://dx.doi.org/10.1016/j.omega.2011.09.001.

Mansini, R., Tocchella, B., 2009a. Effective algorithms for a bounded version of the uncapacitated TPP. In: Innovations in Distribution Logistics. Springer, pp. 267–281. http://dx.doi.org/10.1007/978-3-540-92944-4_14.

Mansini, R., Tocchella, B., 2009b. The traveling purchaser problem with budget constraint. Comput. Oper. Res. 36 (7), 2263–2274. http://dx.doi.org/10.1016/j.cor.2008.09.001.

Menchaca-Méndez, A., Montero, E., Flores-Garrido, M., Miguel-Antonio, L., 2022. An algorithm to compute time-balanced clusters for the delivery logistics problem. Eng. Appl. Artif. Intell. 111, 104795.

Nakano, M., 2019. Supply Chain Management: Strategy and Organization. Springer.

Niknam, T., Fard, E.T., Pourjafarian, N., Rousta, A., 2011. An efficient hybrid algorithm based on modified imperialist competitive algorithm and K-means for data clustering. Eng. Appl. Artif. Intell. 24 (2), 306–317.

Ochi, L.S., Silva, M.B., Drummond, L., 2001. Metaheuristics based on GRASP and VNS for solving traveling purchaser problem. In: Metaheuristics International Conference. Citeseer, pp. 489–494.

Ong, H.L., 1982. Approximate algorithms for the travelling purchaser problem. Oper. Res. Lett. 1 (5), 201–205.

Palomo-Martínez, P.J., Salazar-Aguilar, M.A., 2019. The bi-objective traveling purchaser problem with deliveries. European J. Oper. Res. 273 (2), 608–622.

Pearn, W., 1991. On the traveling purchaser problem. Working Paper 91–01, National Chiao Tung University.

Pearn, W.L., Chien, R., 1998. Improved solutions for the traveling purchaser problem. Comput. Oper. Res. 25 (11), 879–885.

Pradhan, K., Basu, S., Thakur, K., Maity, S., Maiti, M., 2020. Imprecise modified solid green traveling purchaser problem for substitute items using quantum-inspired genetic algorithm. Comput. Ind. Eng. 147, 106578. http://dx.doi.org/10.1016/j.cie.2020.106578.

Ramesh, T., 1981. Traveling purchaser problem. Opsearch 18 (1–3), 78–91.

Reinelt, G., 1995. TSPLIB. http://www.iwr.uni-heidelberg.de/groups/comopt/software, TSPLIB95.

Rengasamy, S., Murugesan, P., 2022. K-means–Laplacian clustering revisited. Eng. Appl. Artif. Intell. 107, 104535.

Riera-Ledesma, J., Salazar-González, J.J., 2005a. The biobjective travelling purchaser problem. European J. Oper. Res. 160 (3), 599–613. http://dx.doi.org/10.1016/j.ejor.2003.10.003.

Riera-Ledesma, J., Salazar-González, J.J., 2005b. A heuristic approach for the travelling purchaser problem. European J. Oper. Res. 162 (1), 142–152.

Riera-Ledesma, J., Salazar-González, J.-J., 2006. Solving the asymmetric traveling purchaser problem. Ann. Oper. Res. 144 (1), 83–97. http://dx.doi.org/10.1007/s10479-006-0014-y.

Riera-Ledesma, J., Salazar-González, J.-J., 2012. Solving school bus routing using the multiple vehicle traveling purchaser problem: A branch-and-cut approach. Comput. Oper. Res. 39 (2), 391–404.

Riera-Ledesma, J., Salazar-González, J.J., 2013. A column generation approach for a school bus routing problem with resource constraints. Comput. Oper. Res. 40 (2), 566–583.

Roy, A., Gao, R., Jia, L., Maity, S., Kar, S., 2020. A noble genetic algorithm to solve a solid green traveling purchaser problem with uncertain cost parameters. Amer. J. Math. Management Sci. 40 (1), 17–31.

Singh, K.N., van Oudheusden, D.L., 1997. A branch and bound algorithm for the traveling purchaser problem. European J. Oper. Res. 97 (3), 571–579.

Syakur, M.A., Khotimah, B.K., Rochman, E.M.S., Satoto, B.D., 2018. Integration K-means clustering method and elbow method for identification of the best customer profile cluster. IOP Conf. Ser.: Mater. Sci. Eng. 336, 012017. http://dx.doi.org/10.1088/1757-899x/336/1/012017.

Teeninga, A., Volgenant, A., 2004. Improved heuristics for the traveling purchaser problem. Comput. Oper. Res. 31 (1), 139–150.

Voß, S., 1996. Dynamic tabu search strategies for the traveling purchaser problem. Ann. Oper. Res. 63 (2), 253–275.

Xie, H., Zhang, L., Lim, C.P., Yu, Y., Liu, C., Liu, H., Walters, J., 2019. Improving K-means clustering with enhanced firefly algorithms. Appl. Soft Comput. 84, 105763.

Xu, L., Wang, Z., Chen, X., Lin, Z., 2022. Multi-parking lot and shelter heterogeneous vehicle routing problem with split pickup under emergencies. IEEE Access 10, 36073–36090.

Yu, S.-S., Chu, S.-W., Wang, C.-M., Chan, Y.-K., Chang, T.-C., 2018. Two improved k-means algorithms. Appl. Soft Comput. 68, 747–755.