# A Math-Heuristic Approach for Two Echelon Vendor Managed Inventory Routing Problem

**Mehmet Can Aydın · Utku Yavuz · Yarkın Akçıl**
**Alp O. Özver · Çağla D. Özmen · Aslı Erol**
Department of Industrial Engineering
Özyegin University
Çekmeköy, İstanbul 34794, Turkey

## Abstract

This paper studies a two-echelon vendor managed inventory routing problem of a honey packager company that delivers packaged products from a single facility to multiple retailers and customers. The objective is creating a supply chain which minimizes the total distribution cost while satisfying customers' demand on time through retailers. The complex nature of the problem originates from connecting the inventory management and the routing process which makes getting an exact solution to the problem difficult. We propose a mathematical optimization model and develop a three-step clustering-based math-heuristic algorithm to solve the problem since commercial solvers fail to provide high-quality solutions within a given time limit. The performance of the algorithm is tested with randomly generated dimensions. The algorithm yields (on average) 16% improvement compared to objective value performances of the commercial solver.

## Keywords
Vendor managed inventory, inventory routing problem, math-heuristics, clustering

## 1. Introduction
The inventory routing problem (IRP) is a variant of the vehicle routing problem, where the priority is on the vehicle's routing as well as the inventory management. The inventory routing problem is defined as follows: A vehicle fleet of fixed capacities serve customers of fixed demand by the period from the central depot. Customers must be assigned vehicles and vehicles are routed so that the total route cost is minimized. These homogeneous vehicles deliver a single product to multiple customers depending on their consumption rate/orders. The route costs depend on the distance between customers or between customers and the depot (Malandraki and Daskin 1992). Determining and managing the optimum inventory amount are equally important as the routing process, where the inventory holding costs are significant and meeting customers' demands perfectly are of paramount importance (Moin et al. 2011). In addition to these objectives, Ekici et al. (2015) shows that, the vehicle routes, delivery amounts and the starting and ending points of the vehicles can be tracked while not exceeding the inventory capacity for both warehouse(s) and vehicle(s) and avoiding product shortage over the designated periods.

Due to the complex nature of the Inventory Routing Problem, several solution methods have also been introduced. The problem mentioned in Federgruen and Zipkin (1984) is a single day inventory routing problem with a limited amount of inventory. The solution proposed includes modified vehicle routing problem heuristics formulated as nonlinear integer programming which aims to minimize inventory holding, shortage and transportation costs while determining routes of each vehicle. Ghiani and Improta (2000) transforms the vehicle routing problem into a capacitated arc routing problem (CARP) and proposes an exact algorithm to solve the routing problem in and out of the clusters, treating the inside of the clusters as a generalized traveling salesman problem determining the least-cost throughout the circuit. Erdogan and Miller-Hooks (2012) states that the green vehicle routing problem (G-VRP) is formulated as a mixed integer linear program. The problem includes the modified Clarke and Wright's savings heuristic and the density-based clustering algorithm (DBCA). DBCA separates the vehicle routing problem into routing and clustering. Crainic et al. (2011) studied a two-echelon vehicle routing problem (2E-VRP), and they solved their problem with a clustering algorithm. Their aim was to minimize the total transportation cost. In this article, first-level vehicle routing is made between depot and satellites, and second-level vehicle routing is made between satellites and customers. Perboli et al. (2011) studies two-echelon vehicle routing problem which is an extension of a classic vehicle routing problem. Their transportation network has two levels. In the first level it is the connection from the depot to intermediate depots. In the continuation, the second level represents the connection between intermediate depots and the customers with infinite time horizons.

Crainic et. al. (2008) proposes clustering-based heuristics for the two-echelon vehicle routing problem. They work in two stages. The first phase attempts to provide a decent workable solution using a clustering algorithm assigns customers to its closest satellite in Euclidean distance, while the second phase aims to develop it. Dondo and Cerda (2007) studied multi-depot heterogeneous flat vehicle routing problems with time windows. This is accomplished by identifying a small number of feasible clusters, each enclosing several customer locations then calculating average travel distances and times between any two of them. Riberio and Lourenço (2003) proposes a model for an inventory system that corresponds to the vendor managed inventory system. The objective of the model is to minimize total inventory cost for the distributor while stock-out is allowed. The inventory routing problem emerges in the context of vendor managed inventory, in which a supplier decides on product completion for its customers. Walter et al. (1983) addresses integrated inventory management and vehicle scheduling in several variants of inventory routing problem model planning horizon that is finite. There are multiple solution approaches for inventory routing problem and genetic algorithm (GA) is one of them. Abdelmaguid and Dessouky (2006) stated that a depot has an adequate supply of inventory that can satisfy all demand throughout the finite time horizon inventory routing problem with multiple fleet sizes. Maximum Level (ML) policy is used as an inventory policy. To solve their problem, they used a genetic algorithm. Moin et al. (2011) studied an inventory routing problem that has a single assembly plant and multiple unique suppliers that supply distinct products. The study has been made based on a many-to-one structure. The demand is deterministic for the product, and time-varying. A hybrid genetic algorithm that considers both inventory and transportation costs has been proposed to solve the problem. Rohmer et al. (2019) claimed a two-echelon inventory routing problem (2E-IRP) which is formulated as a mixed-integer linear programming (MILP). It is considered as a single depot problem with a finite time horizon which is represented as *t*. Since the products have an expiration date, they are sorted in the inventory accordingly. Park et al. (2016) proposed an inventory routing problem with the vendor managed inventory system that produced products distributed from a single manufacturer to multiple retailers. To determine replenishment times and quantities and vehicle routes, they proposed a genetic algorithm to solve the problem. A mixed integer linear programming model is constructed for the VMIRPL. The objective of the model is to maximize supply chain profits for the distributor while lost sales are allowed.

The aim of this paper is to design a two-echelon inventory routing system solution that is computationally scalable for the case study while satisfying all demand within an acceptable optimality gap. However, rather than a supply chain where the products are directly delivered to customers from the production center, a two-echelon logistics network is proposed where several intermediate retailers are utilized. The problem is inspired by a real-life honey packager company that aims to utilize their inventories and vehicles better and work with better routes to satisfy all demand of a specific product with a high seasonality effect. The considered problem revolves around a finite horizon with a single production facility and multiple retailers where a fleet of vehicles with finite capacities transport products from the single production facility to multiple customers through the usage of retailers, based on their forecasted yearly demand. Combined with the difficult-to-predict nature of the demand for the products, direct shipping is required to be implemented to satisfy several customers' demands based on their location and demand which also provides a model that is better suited for real-life cases. The problem definition, parameters and assumptions are taken from the real-life problem to ensure the model is easily implementable. The vendors incur costs such as warehouse and vehicle usage costs, inventory cost, distribution, and direct shipping costs. The introduced mathematical model output consists of a yearly transportation plan for the fleet of vehicles including which routes to take in each period, how much product to carry over these routes, the inventory levels of each warehouse and when to use direct shipping for which customer. Since the problem combines retailer inventory management with routing decisions, the commercial solvers fail to provide first-rate solutions within acceptable time horizons. Therefore, a three-step clustering based math-heuristic approach to reach high-quality solutions with a minimum loss from optimality is proposed.

## 2. Methods

We propose a mathematical formulation solution and a math-heuristic algorithm that includes a clustering method for solving the two-echelon vendor managed inventory routing problem (2E-VMIRP). First, we construct a mathematical model for solving the problem and later show that it is not possible to solve the two-echelon vendor managed inventory routing problem in an accepted time horizon due to the computational challenge of said problem. The complex nature of the problem originates from connecting the inventory management and the routing process which makes getting an exact solution to the problem difficult. Then we propose the math-heuristic algorithm consisting of three steps which are: clustering, second echelon routing and first echelon routing. The idea behind this approach is to divide the problem into three small parts and solve them separately with mathematical formulations that are connected via their inputs and outputs. In clustering we assign retailer(s) to customers to satisfy the demand and cluster the retailer and its customers. In the second step of the algorithm, we focus on the second echelon of the network and take the clusters from the previous step as input. In the final step of our algorithm, we focus on the first echelon routing taking the routing information from the previous step as an input.

## 2.1 Mathematical optimization model

The two-echelon vendor managed inventory routing problem in this study consists of a single production center supplying a single type of product to multiple retailers which then satisfies customer demands working in two echelons. The vendors plan inventory replenishment, ordering products from the production center according to the demand based on the forecasted amounts. In the first echelon, we consider that a product is shipped from the production center ($i$=0) to a set of possible retailers ($i$=1..., $Re$) over a discrete time horizon $T$ via first level vehicles ($v$=1..., $NV$) with routing. In the second echelon, the products delivered to the retailers are distributed to the customers ($j$=1..., $Cu$) via second level vehicles ($l$=1..., $NL$). A sample two-echelon logistics network and its routing decisions are shown in Figure 1.

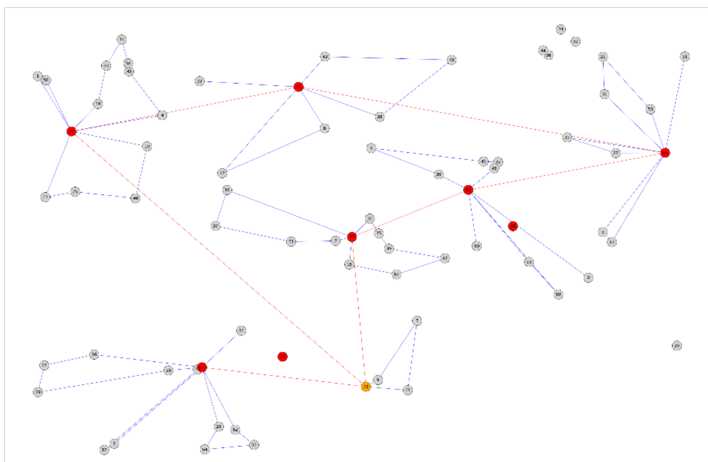


Figure 1: A sample two-echelon logistics network

Vehicles on the first level exit from the production facility and are routed to retailers, then route back to the production facility for further use. Customers can only be served by a designated retailer. It must be signified that backorder is not allowed by any means. Within the same period, separate vehicles exiting from retailers to carry demand over to the customers on the second level and route back to the retailer where they started. It is also possible for a vehicle to route between several customers before returning to the retailer, the same is also valid for vehicles exiting the production center and routing between retailers before routing back. To ensure demand is met, direct shipping methods are also implemented to the model to neutralize cases where truck capacity is exceeded or where retailers cannot satisfy customer demand from their own inventory. The direct shipping option acts as a new type of transportation where the demand of the customer is satisfied from the production center directly with a noticeably higher cost. The mathematical optimization model is formulated as follows and see Table 1 for declarations.

Table 1: Summary of notation for the original mathematical model

| **Variables** | |
|---|---|
| $R1_{0iv}^{t}$ | 1 if vehicle $v$ is utilized from the production facility to customer $j$ at time $t$;0 o.w. |
| $R3_{ikv}^{t}$ | 1 if vehicle $v$ is utilized from retailer $k$ to retailer $i$ in $t$; 0 o.w. |
| $R5_{i0v}^{t}$ | 1 if vehicle $v$ is utilized from retailer $k$ to the production facility at time $t$; 0 o.w. |
| $R0_{0jl}^{t}$ | 1 if vehicle $l$ is utilized from the production facility to customer $j$ at time $t$; 0 o.w. |
| $R2_{ijl}^{t}$ | 1 if vehicle $l$ is utilized from retailer $i$ to customer $j$ at time $t$; 0 o.w. |
| $R6_{mjl}^{t}$ | 1 if vehicle $l$ is utilized from customer $m$ to customer $j$ at time $t$; 0 o.w. |

| | |
|---|---|
| $R4_{jil}^t$ | 1 if vehicle *l* is utilized from customer *j* to retailer *i* at time *t*; 0 o.w. |
| $A_i$ | 1 if retailer *i* has inventory flow for any given period; 0 o.w. |
| $X_{0iv}^t$ | product flow on arc the production facility to retailer *i* by vehicle *v* at time *t*. |
| $Z_{ikv}^t$ | product flow on arc retailer *i* to retailer *k* by vehicle *v* at time *t*. |
| $Y_{ijl}^t$ | product flow on arc retailer *i* to county *j* by vehicle *l* at time *t*. |
| $W_{mjl}^t$ | product flow on arc customer *m* to customer *j* by vehicle *l* at time *t*. |
| $Q_{0jl}^t$ | product flow on arc the production facility to customer *j* by vehicle *l* at time *t*. |
| $DS_j^t$ | number of products delivered to customer *j* by using direct shipping at time *t*. |
| $I_i^t$ | inventory level of retailer *i* at time *t*. |
| **Parameters** | |
| $d_j^t$ | demand of customer *j* at time *t*. |
| $c_{0i}$ | cost of arc from the production center to retailer *i*. |
| $c_{ik}$ | cost of arc from retailer *i* to retailer *k*. |
| $c_{ij}$ | cost of arc from retailer *i* to customer *j*. |
| $c_{mj}$ | cost of arc from customer *m* to customer *j*. |
| $c_{0m}$ | cost of arc from the production center to customer *m*. |
| $FC$ | retailer usage cost. |
| $H$ | inventory holding cost for each product. |
| $CC$ | cost of carry for each product via routing. |
| $DC$ | cost of direct shipping per item. |
| $S$ | inventory capacity for each retailer. |
| $C_v$ | capacity of vehicles, as number of units for the first level. |
| $C_l$ | capacity of vehicles, as number of units for the second level. |
| $CR$ | retailers' periodic distribution capacity |
| $AR$ | number of different routes that retailers can do in a period. |
| $D_{0i}$ | distance between production center to retailer *i*. |
| $D_{ik}$ | distance between retailer *i* to retailer *k*. |
| $D_{ij}$ | distance between retailer *i* to customer *j*. |
| $D_{mj}$ | distance between customer *m* to customer *j*. |
| $MD1$ | the maximum distance a first level vehicle can travel. |
| $MD2$ | the maximum distance a second level vehicle can travel. |

**Mathematical Model:**

$$\begin{aligned}\min_{\mathbf{z}} \; & \sum_l \sum_t \sum_i \sum_j Y_{ij}^t CC + \sum_l \sum_t \sum_m R0_{0ml}^t c_{0m} + \sum_l \sum_t \sum_m R7_{m0l}^t c_{0m} \\ & + \sum_v \sum_t \sum_i R1_{0iv}^t c_{0i} + \sum_t \sum_{i,k} \sum_{i \neq k} \sum_v R3_{ikv}^t c_{ik} + \sum_v \sum_t \sum_i R5_{i0v}^t c_{i0} \\ & + \sum_l \sum_t \sum_i \sum_j R2_{ijl}^t c_{ij} + \sum_l \sum_t \sum_i \sum_j R4_{jil}^t c_{ij} + \sum_l \sum_t \sum_m \sum_j R6_{mjl}^t c_{mj} \\ & + \sum_t \sum_j DS_j^t DC + \sum_t \sum_i I_i^t H + \sum_i A_i FC \end{aligned} \tag{1}$$

s.t.

$$I_i^0 = 0 \tag{2}$$

$$\sum_v X_{0iv}^t + I_i^{t-1} + \sum_v \sum_k Z_{kiv}^t = \sum_l \sum_j Y_{ijl}^t + \sum_v \sum_k Z_{ikv}^t + I_i^t \quad \forall i, t \tag{3}$$

$$I_i^t \leq S \quad \forall i, t \tag{4}$$

$$I_i^T = 0 \quad \forall i \tag{5}$$

$$\sum_{i}\sum_{j} Y_{ijl}^{t} \le C_l \quad \forall l,t \tag{6}$$

$$\sum_{j} Q_{0jl}^{t} \le C_l \quad \forall l,t \tag{7}$$

$$\sum_{i} X_{0iv}^{t} \le C_v \quad \forall i,t \tag{8}$$

$$\sum_{i} R1_{0iv}^{t} \le 1 \quad \forall t,v \tag{9}$$

$$Q_{0jl}^{t} \le MR0_{0jl}^{t} \quad \forall j,t,l \tag{10}$$

$$Y_{ijl}^{t} \le MR2_{ijl}^{t} \quad \forall i,j,l,t \tag{11}$$

$$W_{mjl}^{t} \le MR6_{mjl}^{t} \quad \forall m,j,t,l \tag{12}$$

$$R1_{0iv}^{t} + \sum_{k} R3_{kiv}^{t} = \sum_{k} R3_{ikv}^{t} + R5_{i0v}^{t} \quad \forall i,t,v \tag{13}$$

$$\sum_{k}\sum_{v} R3_{kiv}^{t} + \sum_{v} R1_{0iv}^{t} \le 1 \quad \forall i,t \tag{14}$$

$$R0_{0jl}^{t} + \sum_{i} R2_{ijl}^{t} + \sum_{m} R6_{mjl}^{t} = \sum_{m} R6_{jml}^{t} + \sum_{i} R4_{jil}^{t} + R7_{j0l}^{t} \quad \forall l,t,j \ne m \tag{15}$$

$$\sum_{l} Q_{0jl}^{t} + \sum_{i}\sum_{l} Y_{ijl}^{t} + \sum_{m}\sum_{l} W_{mjl}^{t} - d_j^t + DS_j^t = \sum_{m}\sum_{l} W_{jml}^{t} \quad \forall j,t \tag{16}$$

$$\sum_{j}\sum_{l} Y_{ijl}^{t} \le CR \quad \forall t,i \tag{17}$$

$$\sum_{j}\sum_{l} Q_{0jl}^{t} \le CR \quad \forall t \tag{18}$$

$$\sum_{j}\sum_{l} R2_{ijl}^{t} = AR \quad \forall t,i \tag{19}$$

$$\sum_{j}\sum_{l} R0_{0jl}^{t} = AR \quad \forall t \tag{20}$$

$$\sum_{j} R2_{ijl}^{t} = \sum_{j} R4_{jil}^{t} \quad \forall t,i,l \tag{21}$$

$$\sum_{j} R0_{0jl}^{t} = \sum_{j} R7_{j0l}^{t} \quad \forall t,l \tag{22}$$

$$\sum_{t}\sum_{v} R1_{0iv}^{t} + \sum_{t}\sum_{k}\sum_{v} R3_{kiv}^{t} \le MA_i \quad \forall i \tag{23}$$

$$\sum_{i}\sum_{j} R2_{ijl}^{t} + \sum_{j} R0_{0jl}^{t} \le 1 \quad \forall l,t \tag{24}$$

$$X_{0iv}^{t} \le MR1_{0iv}^{t} \quad \forall i,v,t \tag{25}$$

$$Z_{ikv}^{t} \le MR3_{ikv}^{t} \quad \forall i,k,v,t \tag{26}$$

$$\sum_{l} R6_{mjl}^{t} = 0 \quad \forall m = j,t \tag{27}$$

$$\sum_{i}\sum_{l} R2_{ijl}^{t} + \sum_{m}\sum_{l} R6_{mjl}^{t} + \sum_{l} R0_{0jl}^{t} \le 1 \quad \forall t,j \tag{28}$$

$$\sum_{v} R3_{ikv}^{t} = 0 \quad \forall i = k,t \tag{29}$$

$$\sum_{i} R1_{0iv}^{t} D_{0i} + \sum_{i}\sum_{k} R3_{ikv}^{t} D_{ik} + \sum_{i} R5_{i0v}^{t} D_{0i} \le MD_1 \quad \forall t,v \tag{30}$$

$$\sum_{j} R0_{0jl}^{t} D_{0j} + \sum_{j} R7_{j0l}^{t} D_{0j} + \sum_{i}\sum_{j} R2_{ijl}^{t} D_{ij} + \sum_{m}\sum_{j} R6_{mjl}^{t} D_{mj} + \sum_{i}\sum_{j} R4_{jil}^{t} D_{ij} \le MD_2 \quad \forall t,l \tag{31}$$

$$X_{0iv}^{t}, Y_{ijl}^{t}, Z_{ikv}^{t}, W_{mjl}^{t}, Q_{0jl}^{t}, I_{i}^{t}, DS_{j}^{t} \geq 0 \text{ and integer} \quad (32)$$
$$R0_{0jl}^{t}, R1_{0iv}^{t}, R2_{ijl}^{t}, R3_{ikv}^{t}, R4_{jil}^{t}, R5_{i0v}^{t}, R6_{mjl}^{t}, R7_{jil}^{t}, A_{i} \in 0,1 \quad (33)$$

The objective function (1) is the minimization of all the related costs which are cost of carry per product, cost of using the specific arc, cost of direct shipping, holding cost of inventory and retailer usage fixed cost. Constraint (2) sets the initial inventory to zero. Constraint (4) sets the inventory capacity for all the retailers. Constraint (5) ensures that no inventory is left at the end of the last period. Constraint (8) sets the vehicle capacity for the first level while ensuring that a delivery can be smaller or equal to the capacity of the vehicle. Constraints (6) and (7) set the vehicle capacity for the second period while ensuring that a delivery can be smaller or equal to the capacity of the vehicle. Constraint (9) ensures that a vehicle that comes out from the production center can visit one retailer at maximum in a particular period. Constraint (14) enables a retailer to be visited only by another retailer or the production center. Constraint (3) sets the inventory flow balance for all the retailers for the first level. Constraint (16) sets the second level product flow, and ensures that all demand is satisfied via available channels. Constraints (13) and (15) ensure that the number of entering and exiting vehicles and their indices are equal for the first and second level. Constraints (10), (11) and (12) force the flow to be present only if that second level arc is used by that second level vehicle. Constraints (25) and (26) force the flow to be present only if that first level arc is used by that first level vehicle. Constraint (17) makes sure that the number of products distributed by a retailer cannot exceed the retailer distribution capacity. Constraint (18) makes sure that the number of products distributed to customers by the production center cannot exceed its distribution capacity. Constraints (19) and (20) limit the possible routes a retailer or the production center can do at any period to the different available routes for every period. Constraint (29) computes that a retailer cannot tour back to itself. Constraint (27) computes that a customer cannot tour back to itself. Constraints (21) and (22) ensure that the vehicles return to their designated starting point whether it is the production center or a retailer. Constraint (24) ensures that only a single retailer can utilize a second level vehicle for any period t. Constraint (28) ensures that visiting is made at most once to satisfy demand of a customer via routes. Constraint (23) limits the retailer to be utilized only if it receives product from the production center or any other retailer. Constraints (30) and (31) ensure that the distances traveled by the first and second level vehicles does not exceed the maximum allowed distance for each of them. Constraints (32) and (33) enforce binary, integer and non-negativity conditions upon the variables.

## 2.2 Math-Heuristic approach

The complex nature of the problem combined with its dimensions forms a computationally challenging problem to be solved in commercial solvers. Therefore, a heuristic approach is mandatory for the problem to be solved efficiently. We propose a math-heuristic algorithm consisting of three steps; clustering, second echelon routing and first echelon routing. In the proposed solution we use different mathematical formulations for each of the steps. The first mathematical model aims to create clusters consisting of a single retailer and multiple customers. The customers within the cluster are served by their corresponding retailer. After the clustering is complete, the second mathematical model is solved for each cluster on the second echelon, using the output of the first model. Following after is the third model, which is solving the routing for the first echelon, yielding the routing between the production center and each retailer that is included in a cluster in the previous model. Over the improvement period of the algorithm construction process, we continually generated randomized data sets to observe the wellness of the algorithm compared to the result of the mathematical model.

### *2.2.1 Step 1 – Clustering*

In this step we aim to effectively cluster the retailers and the customers together. The key point is to select the best possible retailer-customer combinations with a method that does not steer the result too far from the optimal solution. A utility based mathematical model is introduced as the mentioned method. The objective of the model is to pick the retailer-customer pairings with the best demand over distance value, which is called the utility. Also, we incur an addition to the total utility value to consider the effect of the products distributed through used retailers on the objective function. Production center will also be used for distributing products to customers directly, meaning that customers close to the production center can be served via second level vehicles departed from the production center. The maximization of the utility needs an upper-bound to limit the model. Note that the last retailer node which is the dummy retailer corresponds to direct shipping. We introduced direct shipping as a retailer to compare the efficiency and fitness of direct shipping and comparing it to other possible retailers. Keep in mind that the dummy retailer has no capacity and does not involve any parameters while being available for usage every period. The dummy retailer is assumed to be at an equal distance to every customer even though this assumption is not possible in real life. See Table 2 for all declarations in clustering mathematical model.

Table 2: Summary of notation for the clustering mathematical model

| | |
|---|---|
| **Variables** | |
| $x_{ij}^t$ | 1 if customer *i* is assigned to retailer *j* at time *t*; 0 o.w. |
| $y_j^t$ | 1 if retailer *j* is utilized at time *t*; 0 o.w. |
| **Parameters** | |
| $U_{ij}^t$ | utility of customer *i* - retailer *j* pairing (demand/distance) at time *t*. |
| $MR$ | possible number of retailers can be used. |
| $MD$ | tolerated distance for each retailer-customer pair. |
| $DR$ | minimum demand required to be assigned to retailers for usage. |
| **Abbreviations** | |
| $pc$ | production center. |
| $dm$ | dummy retailer (direct shipment). |

**Clustering Mathematical Model:**

$$\max_{\mathbf{z}} \quad \frac{1}{\alpha}\sum_{i}\sum_{j\notin dm}\sum_{t} U_{ij}^t x_{ij}^t + \frac{1}{\beta}\sum_{i}\sum_{j\notin dm}\sum_{t} d_i^t x_{ij}^t \tag{1}$$

s.t.

$$\sum_{j} x_{ij}^t = 1 \quad \forall i,t \tag{2}$$

$$\sum_{j\notin\{pc,dm\}} y_j^t \leq MR \quad \forall t \tag{3}$$

$$x_{ij}^t \leq y_j^t \quad \forall i,j,t \tag{4}$$

$$\sum_{i} d_i^t x_{ij}^t \leq CR \quad \forall j,t \tag{5}$$

$$D_i^t \; x_{ij}^t \leq MD \quad \forall i,j,t \tag{6}$$

$$\sum_{i}\sum_{t} d_i^t x_{ij}^t \geq \sum_{t} DR \;\; y_j^t \quad \forall j \notin \{pc,dm\} \tag{7}$$

$$y_j^t = 1 \quad \forall t, j = \{pc,dm\} \tag{8}$$

$$y_j^t = y_j^{t+1} \quad \forall t, j \notin \{pc,dm\} \tag{9}$$

$$x_{ij}^t, y_j^t \in \{0,1\} \tag{10}$$

The objective function (1) maximizes the utility by also considering the amount of demand distributed via routing, the nature of the objective function is further explained in §*2.1.1.1*. Constraint (2) allows a customer to only receive products from one retailer. Constraint (3) limits the total number of retailers utilized to not exceed the maximum number of retailers. Constraint (4) allows a retailer can only be assigned to customers if it is utilized. Constraint (5) ensures that product flow from a retailer cannot exceed the retailer's total distribution capacity. Constraint (6) limits the distance traveled from a customer to a retailer to not exceed the maximum tolerated distance. Constraint (7) makes sure that if a retailer is utilized, it meets the minimum demand output requirement. Constraint (8) enables direct shipping/dummy retailers and production centers to be used and they are not included in the calculation of maximum retailer capacity. Constraint (9) makes sure that if a retailer is utilized at period t, it can be utilized for all other periods without incurring further costs. Constraint (10) sets the binary condition for the variables.

### *2.2.1.1 Determining $\alpha$ and $\beta$ values*

As seen in the objective function of the clustering mathematical model which maximizes utility, also to consider the amount of products delivered via routing it is added to the objective function. However, while comparing the average demand and utility parameters' values there is a significant difference between their magnitudes, and this difference affects the objective function critically. To reduce this effect, it is required to normalize both sides of the objective function with designated coefficients. To determine these coefficients, the objective function is separated into two parts and the model is run separately for each part. While the first execution maximizes only the utility function, the second execution aims to only maximize the amount of products delivered via routing. Then the optimal objective values from both models are designated as $\alpha$ and $\beta$ coefficients and implemented to the original Clustering Mathematical Model. The utility and demand functions are divided by $\alpha$ and $\beta$ values respectively. With this implementation featuring the optimal values $\alpha$ and $\beta$ of the two objective functions, the combined objective function can take values in the range of [0,2] and the resulting clustering mathematical model is normalized.

### *2.2.2 Step 2 – Second Echelon Routing*

The clustering mathematical model is giving clusters of retailer-customer assignments as an output which we will use as an input in this step. Recall that, the production center will behave like a retailer in this model to ensure customers assigned to the production center are served. The introduced mathematical model is a vehicle routing problem model that minimizes costs and determines the second level routes between the retailer and its assigned customers and between customers. Therefore, all decision variables and parameters related to the first level are removed. The second echelon routing includes the possibility of direct shipping that the model can prefer depending on the condition. Recall that direct shipping is sending the product/s via a different and direct transport with an increased cost to the customer in need without stopping somewhere else. Every cluster taken from the clustering step contains only one retailer and each customer must be served by at most one retailer, which creates the necessity of running the model for each cluster separately.

**Second Echelon Routing Mathematical Model:**

$$\min_{z} \sum_{l}\sum_{t}\sum_{i}\sum_{j} Y_{ij}^{t}CC + \sum_{t}\sum_{j} DS_{j}^{t}DC + \sum_{l}\sum_{t}\sum_{i}\sum_{j} R2_{ijl}^{t}c_{ij} + \sum_{l}\sum_{t}\sum_{i}\sum_{j} R4_{jil}^{t}c_{ij} + \sum_{l}\sum_{t}\sum_{m}\sum_{j} R6_{mjl}^{t}c_{mj} \quad (11)$$

**s.t.**

$$\sum_{l} R6_{mjl}^{t} = 0 \quad \forall t, m = j \quad (12)$$

$$Y_{ijl}^{t} \le MR2_{ijl}^{t} \quad \forall i, j, l, t \quad (13)$$

$$W_{mjl}^{t} \le MR6_{mjl}^{t} \quad \forall m, j, t, l \quad (14)$$

$$\sum_{i} R2_{ijl}^{t} + \sum_{m} R6_{mjl}^{t} = \sum_{m} R6_{jml}^{t} + \sum_{i} R4_{jil}^{t} \quad \forall l, t, j \ne m \quad (15)$$

$$\sum_{i} R2_{ijl}^{t} + \sum_{m} R6_{mjl}^{t} \le 1 \quad \forall t, j \quad (16)$$

$$\sum_{i}\sum_{j} Y_{ijl}^{t} = C_{l} \quad \forall l, t \quad (17)$$

$$\sum_{i}\sum_{l} Y_{ijl}^{t} + \sum_{m}\sum_{l} W_{mjl}^{t} - d_{j}^{t} + DS_{j}^{t} = \sum_{m}\sum_{l} W_{jml}^{t} \quad \forall j, t \quad (18)$$

$$\sum_{j} R2_{ijl}^{t} \le 1 \quad \forall t, i, l \quad (19)$$

$$\sum_{i}\sum_{j} R2_{ijl}^{t}D_{ij} + \sum_{m}\sum_{j} R6_{mjl}^{t}D_{mj} + \sum_{i}\sum_{j} R4_{jil}^{t}D_{ij} \le MD_{2} \quad \forall t, l \quad (20)$$

$$\sum_{j} d_{j}^{t} - \sum_{j} DS_{j}^{t} = N_{i}^{t} \quad \forall t, i \quad (21)$$

$$Y_{ijl}^{t}, W_{mjl}^{t}, DS_{j}^{t}, N_{i}^{t} \ge 0 \text{ and integer} \quad (22)$$

$$R2_{ijl}^{t}, R4_{jil}^{t}, R6_{mjl}^{t} \in \{0, 1\} \quad (23)$$

The objective function (11) is the minimization of all the related costs which are cost of carry per product, cost of using the specific arc and cost of direct shipping. Constraint (12) limits the customer to tour back to itself. Constraints (13) and (14) allow the flow to be present only if that arc is used by a second level vehicle. Constraint (15) ensures that the number of entering and exiting vehicles and their indices are equal. Constraint (16) allows the routing to be present only in the purpose of satisfying a demand. Constraint (17) is the vehicle capacity constraint for the second level. Constraint (18) is the second level flow constraint, and ensures that all demand is satisfied via available channels. Constraint (19) limits the route to only start from a retailer, arriving to a customer. Constraint (20) ensures that the distance traveled by a second level vehicle does not exceed the maximum allowed distance. Constraint (21) enables the dummy demand of the retailers to be equal to the demand of the customers that has been satisfied by that retailer. The Constraints (22) and (23) are binary, integer and non-negativity conditions for the variables.

$$\sum_{j} d_j^t - \sum_{j} DS_j^t = N_i^t \quad \forall t, i \tag{21}$$

The dummy demand $N_i^t$ is constructed in order to track the amount of products required to the retailer which is equal to the products distributed to the assigned customers from that retailer. This dummy demand is initialized in both echelons but with different agendas. In the second echelon the dummy demand is initialized as a decision variable in order to prevent limitation and allow the formulation to find the optimal amount. However, in the first echelon the dummy demand is initialized as a parameter and takes it's value from the second echelon routing as an input.

### *2.2.3 Step 3 – First Echelon Routing*

In this step we use some of the outputs from the previous mathematical model, the Customer Bubble Model. The retailers that are used in the previous model are all assigned with customer/s, as it is known that they must be opened for the optimal result to be achieved in our next model. It is assumed that those retailers are available and the other ones are not available, calculating the cost of using a retailer accordingly. The introduced mathematical model is an IRP model that minimizes costs and determines the routes between the production center and retailers and between retailers. In the second echelon routing there are several clusters that must be solved and the same model must be run for each of them; however, in this model all the retailers and the production center are in the same cluster, so running the model once for the entire cluster will cover all the possible routes.

**First Echelon Routing Mathematical Model:**

$$\min_{\mathbf{z}} \quad \sum_{t}\sum_{i}\sum_{v} R1_{0iv}^t c_{0i} + \sum_{t}\sum_{i,k}\sum_{i\neq k}\sum_{v} R3_{ikv}^t c_{ik} + \sum_{t}\sum_{i}\sum_{v} R5_{i0v}^t c_{0i} + \sum_{t}\sum_{i} I_i^t H \tag{24}$$

$$\text{s.t.} \tag{25}$$

$$I_i^0 = 0 \quad \forall i \tag{26}$$

$$I_i^T = 0 \quad \forall i \tag{27}$$

$$I_i^t \le S_i \quad \forall i, t \tag{28}$$

$$\sum_{v} X_{0iv}^t + I_i^{t-1} + \sum_{v}\sum_{k} Z_{kiv}^t = N_i^t + \sum_{v}\sum_{k} Z_{ikv}^t + I_i^t \quad \forall i, t \tag{29}$$

$$\sum_{i} X_{0iv}^t \le C_v \quad \forall v, t \tag{30}$$

$$\sum_{i} R1_{0iv}^t \le 1 \quad \forall v, t \tag{31}$$

$$R1_{0iv}^t + \sum_{k} R3_{kiv}^t = \sum_{k} R3_{ikv}^t + R5_{i0v}^t \quad \forall i, v, t \tag{32}$$

$$\sum_{v} R1_{0iv}^t + \sum_{k}\sum_{v} R3_{kiv}^t \le 1 \quad \forall i, t \tag{33}$$

$$X_{0iv}^t \le M R1_{0iv}^t \quad \forall i, v, t \tag{34}$$

$$Z_{ikv}^t \le M R3_{ikv}^t \quad \forall i, k, v, t \tag{35}$$

$$\sum_{v} R3_{ikv}^t = 0 \quad \forall i = k, t \tag{36}$$

$$\sum_{i} R1_{0iv}^t D_{0i} + \sum_{i}\sum_{k} R3_{ikv}^t D_{ik} + \sum_{i} R5_{i0v}^t D_{0i} \le MD_1 \quad \forall t, v \tag{37}$$

$$X_{0iv}^t, Z_{ikv}^t, I_i^t \ge 0 \text{ and integer} \tag{38}$$

$$R1_{0iv}^{t}, R3_{ikv}^{t}, R5_{i0v}^{t} \in \{0,1\} \quad (39)$$

The objective function (24) is the minimization of all the related costs which are the cost of using the specific arc and inventory holding cost. Constraint (25) sets the initial inventory of every retailer to zero. Constraint (26) makes sure that at the end of the last period there are no products in the inventory. Constraint (27) sets the inventory capacity of all the retailers. Constraint (28) is the inventory balance constraint for all the retailers, and ensures product needs to be distributed from each used retailer is satisfied by the production center. Constraints (29) and (30) set the vehicle capacity and make sure that the delivery amounts can be smaller than or equal to the first level vehicle capacity. Constraint (31) ensures that the number of entering and exiting vehicles and their indices are equal. Constraint (32) enables a retailer can be visited by another retailer or the production center. Constraints (33) and (34) allow the flow to be present only if that arc is used by a first level vehicle. Constraint (35) makes sure that a retailer cannot tour back to itself. Constraint (36) limits the distance traveled by a first level vehicle to be smaller than or equal to the maximum allowed distance. Constraints (37) and (38) are the binary, integer and non-negativity constraints for the variables. See Table 3 in order to understand complete working principle of the math-heuristic approach.

Table 3: Pseudocode of the proposed math-heuristic algorithm

| | |
|---|---|
| **Clustering** | |
| **Inputs:** | Utility $U_{ij}^{t}$, demand $d_i^t$, and the clustering model. |
| **Procedure:** | **Step 0:** Start with *j* available retailers (initially *j*=0), for every visit increase *j* by one.<br>**Step 1:** Execute the clustering model twice to maximize both side of the clustering model's objective function separately, keep optimal values as $\alpha$ and $\beta$.<br>**Step 2:** Using $\alpha$ and $\beta$ values, execute the initial clustering model to find best retailer-customer pairs.<br>**Step 3:** Keep the quantity of demand carried with direct shipping determined in **Step 2**, together with the number of retailers available.<br>**Step 4:** If *j* = 0, go to **Step 0**, and proceed to **Step 5** without visiting **Step 4**, else go to **Step 5**.<br>**Step 5:** Compare the difference in quantity of demand carried with direct shipping when available retailers equal *j* and $j-1$. If the difference is bigger than the parameter *DR*, go to **Step 0**, else break the procedure and keep all customer-retailer pairing results when number of available retailers is equal to $j-1$. |
| **Outputs:** | $x_{ij}^{t}$ values if customer *i* assigned to retailer *j* at time *t*, and $y_j^t$ values if retailer *j* is used. |
| **Second Echelon** | |
| **Inputs:** | Customer-retailer assignment list, used retailers list from clustering, and the second echelon routing mathematical model. |
| **Procedure:** | **Step 0:** Start with *k*<sup>th</sup> index in used retailers list (initially *k*=0), and keep the value in index as $i$.<br>**Step 1:** For each retailer, period *t* is 1 initially, if $t > t_{max}$, then increase *k* by one, and go to **Step 0**, else go to **Step 2**.<br>**Step 2:** Indicate which customer is assigned to retailer *i* at time *t* by using customer-retailer assignment list, and corresponding parameters.<br>**Step 3:** Execute the second echelon routing mathematical model with indicated customers and keep amount of product needs to be distributed via routing by retailer *i* at time *t* as initialized decision variable $N_i^t$. Increase *t* by one and go to **Step 1**. |
| **Outputs:** | Total cost as objective function value and routing output of the second echelon.<br>Amount of product needs to distribute via routing for every used retailer *i*, and every *t* as $N_i^t$. |
| **First Echelon** | |
| **Inputs:** | $N_i^t$ as parameter; for every used retailer and every t, the first echelon routing mathematical model. |
| **Procedure:** | Execute the first echelon routing mathematical model to satisfy product needed to used retailers for distribution. |
| **Outputs:** | Total cost as objective function value and routing output of the first echelon. It will be initialized the utilization cost of a retailer for each used retailer. |

# 3. Results and Discussion

Table 4: Computational result comparisons on randomly generated instances

| c | r | t | Model Result | Gap (%) | Algorithm Result | CPU (sec) | Improv. (%) |
|---|---|---|---|---|---|---|---|
| 30 | 4 | 2 | 211960 | 60.7 | **197182** | 52 | 7.9% |
| 30 | 4 | 2 | 170139 | 49.5 | **148720** | 12 | 14.8% |
| 40 | 5 | 2 | 198413 | 48.0 | **182240** | 72 | 8.1% |
| 40 | 5 | 2 | 223397 | 57.4 | **204596** | 726 | 10.8% |
| 50 | 7 | 3 | 366869 | 40.3 | **302869** | 1404 | 17.4% |
| 50 | 7 | 3 | 425680 | 37.6 | **410536** | 77 | 6.4% |
| 65 | 8 | 3 | 642517 | 49.3 | **507150** | 1135 | 29.8% |
| 65 | 8 | 3 | 517274 | 42.6 | **428679** | 281 | 30.3% |

The Table 4 provides the results of the mathematical model and the proposed algorithm, the optimality gap of the mathematical model after running time of four hours in CPLEX, the amount of time it takes for the algorithm to solve the problem instance as seconds and the improvement on results of the proposed algorithm compared to the mathematical model. Note that, the percentile for improvement being positive means that the algorithm performed better as the smaller result is preferred for the minimization objective for the problem. The n, r, and t stand for the number of customers, number of retailers and the time horizon for the specific instance, respectively. The locations of the production center, retailers, and the customers on three randomly generated test instances are shown in Figure 2. The amount of time that is required for the proposed algorithm to solve the problem instance is highly dependent on the number of customers that are connected to any retailer. Even though for two different instances sizes are the same, computational time to get a solution may differ due to size of clusters imported to the algorithm to be solved. The Improv. column is included to observe the performance between two different solutions introduced in this paper. Although the computational time of the proposed algorithm may be high in specific instances, it is still much better than the original mathematical model solved in CPLEX. The goal for the algorithm was to reduce the computational time with minimum loss from optimal results. As a result, we decreased the computational time severely and consistently achieved better results than the 4 hours of run time in CPLEX for the original mathematical model. The algorithm's routing results for the test instances in Figure 2 are available in Appendix A.

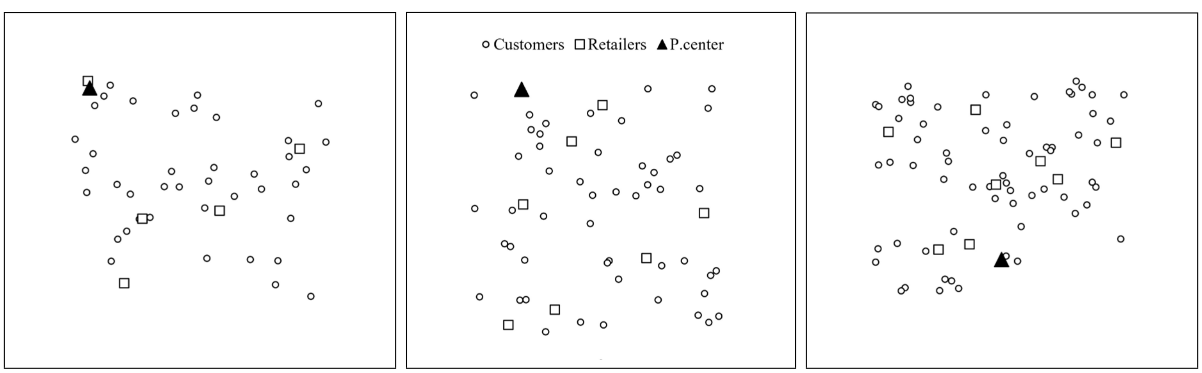



Figure 2: Randomly generated instances with 40, 50, 65 customers respectively

# 4. Conclusion

The two-echelon vendor managed inventory routing problem is discussed in this paper, and the clustering-based math-heuristic algorithm consists of three parts is proposed. The first phase constructs clusters including a single retailer and multiple customers in each. The customers are paired with their corresponding retailers using a utility-based function. The second step solves the routing problem in each cluster independently and the third step solves the routing problem between each cluster's retailer and the production center. The proposed algorithm achieves an on average 16% improvement compared to mathematical model's four hours performance in commercial solver CPLEX on all instances run, while decreasing computational time to minutes. The proposed math-heuristic algorithm is compared with the initial model in several instances to yield a proper estimate about the performance of the algorithm.

## References

Tamer F Abdelmaguid and Maged M Dessouky. A genetic algorithm approach to the integrated inventory-distribution problem. *International Journal of Production Research*, 44(21):4445– 4464, 2006.

Walter J Bell, Louis M Dalberto, Marshall L Fisher, Arnold J Greenfield, Ramchandran Jaikumar, Pradeep Kedia, Robert G Mack, and Paul J Prutzman. Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer. *Interfaces*, 13(6):4–23, 1983.

Teodor Gabriel Crainic, Simona Mancini, Guido Perboli, and Roberto Tadei. *Clustering-based heuristics for the two-echelon vehicle routing problem*, volume 46. CIRRELT Montréal, 2008.

Teodor Gabriel Crainic, Simona Mancini, Guido Perboli, and Roberto Tadei. Multi-start heuristics for the two-echelon vehicle routing problem. In *European Conference on Evolutionary Computation in Combinatorial Optimization*, pages 179–190. Springer, 2011.

Rodolfo Dondo and Jaime Cerdá. A cluster-based optimization approach for the multi-depot heterogeneous fleet vehicle routing problem with time windows. *European journal of operational research*, 176(3):1478–1507, 2007.

Ali Ekici, Okan Orsan Özener, and Gültekin Kuyzu. Cyclic delivery schedules for an inventory¨ routing problem. *Transportation Science*, 49(4):817–829, 2015.

Sevgi Erdoğan and Elise Miller-Hooks. A green vehicle routing problem. *Transportation research part E: logistics and transportation review*, 48(1):100–114, 2012.

Awi Federgruen and Paul Zipkin. A combined vehicle routing and inventory allocation problem. *Operations Research*, 32(5):1019–1037, 1984.

Gianpaolo Ghiani and Gennaro Improta. An efficient transformation of the generalized vehicle routing problem. *European Journal of Operational Research*, 122(1):11–17, 2000.

Chryssi Malandraki and Mark S Daskin. Time dependent vehicle routing problems: Formulations, properties and heuristic algorithms. *Transportation science*, 26(3):185–200, 1992.

Noor Hasnah Moin, Said Salhi, and NAB Aziz. An efficient hybrid genetic algorithm for the multi-product multi-period inventory routing problem. *International Journal of Production Economics*, 133(1):334–343, 2011.

Yang-Byung Park, Jun-Su Yoo, and Hae-Soo Park. A genetic algorithm for the vendor-managed inventory routing problem with lost sales. *Expert Systems with Applications*, 53:149– 159, 2016.

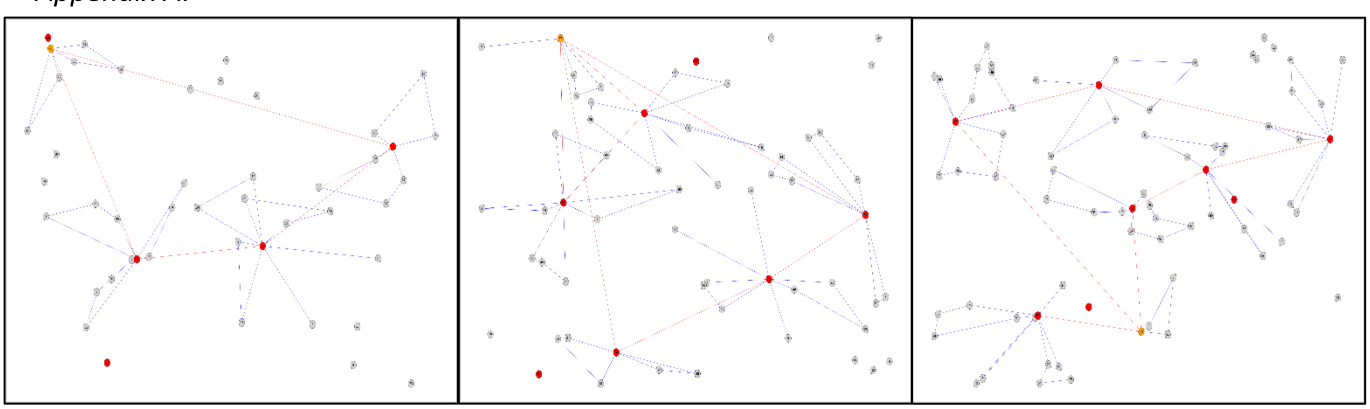
Guido Perboli, Roberto Tadei, and Daniele Vigo. The two-echelon capacitated vehicle routing problem: Models and math-based heuristics. *Transportation Science*, 45(3):364–380, 2011.

Rita Ribeiro and Helena Ramalhinho-Lourenço. Multi-period vendor managed inventory systems. 2003.

SUK Rohmer, GDH Claassen, and Gilbert Laporte. A two-echelon inventory routing problem for perishable products. *Computers & Operations Research*, 107:156–172, 2019.

## Appendix

*Appendix A:*



Appendix A: Algorithm results of the instances in Figure 2