Large scale inventory routing problem with split delivery: a new model and Lagrangian relaxation approach

Yugang Yu, Haoxun Chen* and Feng Chu

Industrial System Optimization Group,
Institute of Computer Science and Engineering of Troyes,
University of Technology of Troyes,
12 rue Marie Curie, BP 2060,
Troyes 10010, France
E-mail: ygyums@gmail.com
E-mail: haoxun.chen@utt.fr
E-mail: feng.chu@utt.fr
*Corresponding author

Abstract: Inventory Routing Problem (IRP) integrates inventory planning with vehicle routing to minimise total logistics cost by coordinating inventory and transportation activities. Due to its complexity, an approximate model with new subtour elimination constraints is proposed for IRP with split delivery. Lagrangian Relaxation (LR) is used to decompose the model into subproblems that are solved by linear programming and Minimum Cost Flow (MCF) algorithms. A near-optimal solution of the model is constructed from the solution of the relaxed problem using a heuristic. The solution, which defines for each period the delivery volume for each customer, the number of times traversed by vehicles and the total quantity transported on each directed arc in the corresponding transportation network, is repaired to a feasible solution of the IRP by solving a series of assignment problems. Numerical experiments show that the proposed approach can find near-optimal solutions for the IRP with up to 200 customers in a reasonable computation time.

Keywords: logistics; Inventory Routing Problem (IRP); transportation; Lagrangian Relaxation (LR); services operations.

Reference to this paper should be made as follows: Yu, Y., Chen, H. and Chu, F. (2006) 'Large scale inventory routing problem with split delivery: a new model and Lagrangian relaxation approach', *Int. J. Services Operations and Informatics*, Vol. 1, No. 3, pp.304–320.

Biographical notes: Yugang Yu got his PhD in Information Management and Decision Science from the Business School of the University of Science and Technology of China in 2003. He is currently a post-doctoral research fellow at the Institute of Computer Science and Engineering, University of Technology of Troyes, France. His research interests are focused on supply chain management, operation research and game theory.

Haoxun Chen got his PhD in Systems Engineering from Xian Jiaotong University, China, in 1990. He is currently a Professor at the Institute of Computer Science and Engineering, University of Technology of Troyes, France. His main research interests are in supply chain management, production planning and scheduling and discrete event systems.

Feng Chu got her PhD in Computer Science from the University of Metz, France, in 1995. She is currently an Associate Professor at the Institute of Computer Science and Engineering, University of Technology of Troyes, France. She is mainly interested in modelling, analysis and optimisation of supply chains or production systems, transportation planning, production planning and scheduling.

1 Introduction

Inventory Routing Problem (IRP) arises when a supplier (e.g. a central depot) supplies a given commodity to a set of customers on a regular basis. For the supplier, the problem involves the daily specification of the number of deliveries, the route of each delivery and the delivery quantity for each customer in each route for inventory replenishment of the customers. In this paper, we consider the Inventory Routing Problem with Split Delivery (IRPSD) with the objective of minimising the total inventory and transportation cost. The problem is to determine the delivery quantity for each customer and a set of feasible vehicle routes for the deliveries in each period subject to vehicle capacity constraints and the inventory capacity constraint of each customer. The delivery volume for each customer in each period can be split and transported by multiple vehicles. It is assumed that the demand of each customer in each period must be satisfied on time.

The problem is of practical importance due to its logic for reducing the total cost by making a tradeoff between transportation and inventory costs across multiple periods. On the one hand, it allows customers' demands to be satisfied ahead of time by keeping local inventories. This makes the consolidation of small delivery volumes into larger ones to reduce the transportation cost possible. With the increased degree of freedom for transportation decisions, the transportation pressure in demand peak periods can also be alleviated. On the other hand, the consideration of inventory in the problem encourages the supplier's just-in-time deliveries to customers to reduce their inventory costs. Moreover, split delivery allows that the delivery of a customer's demand in any period can be satisfied by more than one vehicle, which makes customer's demand larger than a vehicle's capacity possible. The problem in each period is then a relaxation of widely discussed Capacitated Vehicle Routing Problem (CVRP). In CVRP, every customer must be serviced by one vehicle.

The literature relating to our present work includes Split Delivery Vehicle Routing Problem (SDVRP) and IRP. Research on the SDVRP is relatively recent although it widely exists in practice. Dror and Trudeau introduced the problem in 1989 and showed how this relaxation of the standard CVRP could lead to important savings, both in the total distance travelled and in the number of vehicles used. Despite this relaxation, the SDVRP remains NP-hard (see Dror and Trudeau, 1990). Dror et al. (1994) presented an integer programming formulation of the SDVRP and developed several classes of valid inequalities. They proposed algorithms for calculating the lower and the upper bound for the optimal objective value of the problem and reported the gap between the upper and lower bound from 0% to 9% for small problems. However, no exact algorithm was implemented.

Belenguer et al. (2000) studied the facial structure of the SDVRP polyhedron, and also identified several classes of valid inequalities. They considered instances up to 48 customers, and reported the gap between 0% and 12%. Once again, a feasible solution of the SDVRP cannot be obtained from the output of their algorithm in general.

The models developed in the above-mentioned papers do not offer an exact method to find an optimal solution of the SDVRP. All methods developed for the SDVRP are based on the introduction of variables for every vehicle although all vehicles involved are homogenous, such as in Ho and Haugland (2004) and Lee et al. (2004). In these papers, when a large number of customers are considered, their methods can not guarantee obtaining an optimal solution as the tabu search heuristic proposed by Ho and Haugland (2004). Exact methods can deal with only small problems with few customers, such as the shortest path approach proposed by Lee et al. (2004) where only seven customers are considered.

For IRP, in the early work, time horizon was often taken just a single day (see, Federgruen and Zipkin, 1984; Golden et al., 1984; etc.). The first effort to develop an approach that considers what happens beyond the next few days was made by Dror et al. (1985) and Dror and Ball (1987). A more comprehensive review of the IRP literature was given by Campbell et al. (2002) and Campbell and Savelsbergh (2004). Here we only review a few recent works on IRP that directly inspire our research. Chandra and Fisher (1994) considered, over a multiperiod horizon, a multiproduct production and distribution system, which integrates both the lot-sizing and the delivery routing decisions. An integrated optimisation model was proposed, based on a multistop routing problem formulation with the introduction of additional set-up constraints. A solution procedure was developed, which follows the typical decision process that most manufacturing firms adopt: the multiitem lot-sizing problem was first solved optimally, and the distribution plan was then given by using a heuristic algorithm. However, Yu et al. (2006) found that a solution of their model might not define a feasible solution for the IRP by tracing feasible routes according to the number of times that each arc is visited in the model's solution as they claimed. In fact their model offers only a lower bound for the optimal objective value of the IRP. Fumero and Vercellis (1999) developed a multiperiod optimisation model describing a single plant logistical system, in which multiple items are manufactured and delivered to customers and production, inventory and routing were considered in the same model. However in their paper, although the vehicles considered are homogeneous, they introduced decision variables for each individual vehicle, which significantly increases the number of decision variables.

In this paper, we focus on solving large scale IRPSD. In order to do so, we first propose an approximate integrated IRP model in which no variable related to a specific vehicle is introduced since the vehicles considered are homogeneous; this can significantly reduce the number of decision variables. New subtour elimination constraints for the IRPSD are introduced in the model. Lagrangian Relaxation (LR) is then used to decompose the model into two subproblems that can be solved by linear programming and Minimum Cost Flow (MCF) algorithms, respectively. Based on the solution of the Lagrangian relaxed problem, a MCF based heuristic is used to construct a feasible solution of the model. The solution is repaired to a near optimal solution of IRPSD by solving a series of assignment problems.

The main contributions of this paper are as follows:

1 We propose a new inventory routing model, in which vehicle specific decision variables are not required, this makes possible to formulate a larger problem with a much less number of decision variables compared with the models with vehicle specific variables. For example, a Vehicle Routing Problem (VRP) with 100 customers and 30 vehicles in each period in the model of Fumero and Vercellis, requires nearly 3×10^5 variables, whereas in our model the same number of variables can formulate the problem with more than 500 customers.

2 We propose new subtour elimination constraints for the vehicle routing problem with split delivery. The constraints are at least as tight as those of Chandra and Fisher (1994), but with much fewer number. If there are N customers considered, only N constraints are required in our model, whereas 2^N constraints are required in their model.

3 We developed some properties of the new model.

4 A Lagrangian relaxation approach is used to decompose the model into two subproblems, while preserving a global optimisation perspective.

5 A repair procedure is developed to repair a feasible solution of the approximate model to a feasible solution of IRPSD by solving a series of assignment problems.

The remainder of the paper is organised as follows: The problem description is given in Section 2, with the introduction of necessary notations. Section 3, presents the approximate model of IRPSD and its properties. Section 4, is dedicated to the LR approach, the most important part of the solution methodology of IRPSD. Based on the results of the LR approach, an approximate optimal solution of IRPSD is constructed in Section 5. Computational results of our algorithm are presented in Section 6. Concluding remarks are given in Section 7.

2 Problem description and notations

This paper is focused on an IRP consisting of a central depot, multiple customers and an unlimited number of homogenous capacitated vehicles. Each customer has a demand requirement of a product in each period which is satisfied by its supplier, the central depot, with a fleet of vehicles. The depot has to deliver the product to its customers, to ensure the complete fulfillment of their demands, which are considered deterministically known. Delivery takes place by means of a fleet of homogeneous vehicles with limited capacity. The delivery volume for each customer in each period can be split and transported by multiple vehicles. The objective is to minimise the sum of the inventory and transportation costs in a rolling horizon, while satisfying the demand requirements of each customer.

The inventory cost is related to the inventory level of each customer at the end of each period. The transportation cost considered includes not only fixed usage cost, which may be related to vehicle insurance, depreciation and drivers' rewards, but also a variable shipping cost, which depends both on transported quantities and travelled distances. This transportation cost structure, adopted by Fumero and Vercellis (1999), not only can approximately model purely distance proportional cost terms (such as costs for fuel) in classical VRP model but also is reasonable for linking the transportation cost to the delivery volume since the transportation cost charged may also be proportional to the shipped volume in the third-party logistics.

The notations required for modelling the problem are given as follows:

Indices

$t = 1,\ldots, T$ Period

$i, j = 0,1,\ldots,N$ Customer and depot, where $i, j = 1,\ldots,N$ are customer indexes, and 0 is the depot index

Parameters

C Vehicle capacity in volume

c_{ij} Variable shipping cost per unit of product along arc (i, j) where $c_{ij} = c_{ji}$ and the triangle inequality holds, that is, $c_{ij} + c_{jk} \geq c_{ik}$ for any i, j, k

c^b_{i0} Cost of travelling with an empty vehicle from customer i back to depot in each period

f_t Fixed vehicle cost per route in period t

h_{it} Holding cost per unit of product for customer i in period t

I_{i0} Initial inventory level of customer i at beginning of period 1

V_i Inventory capacity of customer i

r_{it} Demand for customer i in period t

Variables

d_{it} Delivery volume to customer i in period t

I_{it} Inventory level of customer i at the end of period t

q_{ijt} Quantity transported on directed arc (i, j) in period t

x_{ijt} Number of times traversed by vehicles on directed arc (i, j) in period t

3 Model

The integrated IRPSD model, denoted by P, can be formulated as

Model P:

$$Z = \min \sum_{t=1}^{T}\sum_{i=1}^{N} h_{it} I_{it} + \sum_{t=1}^{T}\sum_{\substack{j=0\\ j\neq i}}^{N}\sum_{i=0}^{N} c_{ij} q_{ijt} + \sum_{t=1}^{T}\sum_{i=1}^{N} f_t x_{i0t} + \sum_{t=1}^{T}\sum_{i=1}^{N} c^b_{i0} x_{i0t} \tag{1}$$

s.t.

$$I_{it} = I_{i,t-1} + d_{it} - r_{it} \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{2}$$

$$I_{i,t-1} + d_{it} \leq V_i \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{3}$$

$$\sum_{\substack{j=0\\j\neq i}}^{N} x_{ijt} = \sum_{\substack{j=0\\j\neq i}}^{N} x_{jit} \quad i = 0,\ldots,N,\ t = 1,\ldots,T \tag{4}$$

$$\sum_{\substack{j=0\\j\neq i}}^{N} q_{jit} - \sum_{\substack{j=0\\j\neq i}}^{N} q_{ijt} = d_{it} \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{5}$$

$$\sum_{i=1}^{N} q_{0it} = \sum_{i=1}^{N} d_{it} \quad t = 1,\ldots,T \tag{6}$$

$$q_{i0t} = 0 \quad i = 1,\ldots,N\ t = 1,\ldots,T \tag{7}$$

$$q_{ijt} \leq Cx_{ijt} \quad i = 0,\ldots,N,\ j = 1,\ldots,N\ i \neq j,\ t = 1,\ldots,T \tag{8}$$

$$I_{it} \geq 0 \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{9}$$

$$x_{ijt} \geq 0 \text{ and integer,} \quad i, j = 0,\ldots,N, i \neq j, t = 1,\ldots,T \tag{10}$$

$$\begin{aligned} &d_{it} \geq 0 \text{ and integer,} \quad i = 1,\ldots,N, q_{ijt} \geq 0 \text{ and integer,} \quad i = 0,\ldots,N,\\ &j = 1,\ldots,N,\ j \neq i, t = 1,\ldots,T \end{aligned} \tag{11}$$

The objective function (1) includes both inventory cost for each customer and transportation cost (fixed and variable transportation costs). Constraints (2) are the inventory balance constraints for each customer. Constraints (3) are the inventory capacity constraint of each customer. Constraints (4) ensure that the number of vehicles leaving from a customer or the depot is equal to that of arrival. Constraints (5) are the product flow conservation equations, assuring flow balance at each customer and eliminating all subtours. Constraints (6) assure the collection of the accumulative delivery volume at the depot. Constraints (7) guarantee that only empty vehicle returns to the depot. Constraints (8) formulate the vehicle capacity and logical relationship between q_{ijt} and x_{ijt}. Constraints (9) ensure that each customer's demand requirement is completely fulfilled without backorder.

If we consider only one period and assume that the delivery volume for each customer is given, the model *P* is reduced to a SDVRP. For the corresponding SDVRP, we have the following theorem 1. For readability, the subscript *t* in model *P* is dropped.

Theorem 1: *Consider the SDVRP in which q_{ij} is the quantity transported on directed arc (i, j), i, j = 0,1,…,N, i ≠ j. Let d_i denote a positive quantity that must be delivered to customer i, then the following constraints (12) and (13)*

$$\sum_{\substack{j=0\\j\neq i}}^{N} q_{ji} - \sum_{\substack{j=0\\j\neq i}}^{N} q_{ij} = d_i \tag{12}$$

$$q_{ij} \leq Cx_{ij} \quad i = 0,\ldots,N,\ j = 1,\ldots,N\ i \neq j \tag{13}$$

1 are at least as tight as the constraints

$$\sum_{j\in U_0\setminus S} \sum_{i\in S} x_{ji} \geq \sum_{i\in S} d_i / C \quad \forall S \subseteq \{1,\ldots,N\} \tag{14}$$

where $U_0 = \{0,1,\ldots,N\}$, *proposed by Chandra and Fisher (1994)*

2 *are a necessary condition for defining a feasible solution of the SDVRP*

3 *are a sufficient condition for subtour elimination.*

Proof:

1 For $\forall S \subseteq \{1,...,N\}$

$$\sum_{i\in S}\left(\sum_{\substack{j=0\\j\neq i}}^{N} q_{ji} - \sum_{\substack{j=0\\j\neq i}}^{N} q_{ij}\right) = \sum_{j\in U_0\setminus S}\sum_{i\in S} q_{ji} - \sum_{j\in U_0\setminus S}\sum_{i\in S} q_{ij} \le \sum_{j\in U_0\setminus S}\sum_{i\in S} q_{ji}$$

Considering constraints (12), we have

$$\sum_{j\in U_0\setminus S}\sum_{i\in S} q_{ji} - \sum_{j\in U_0\setminus S}\sum_{i\in S} q_{ij} = \sum_{i\in S} d_i$$

and

$$\sum_{j\in U_0\setminus S}\sum_{i\in S} q_{ji} \ge \sum_{i\in S} d_i$$

Considering constraints (13), since $q_{ij} \le Cx_{ij}$, we have

$$\sum_{j\in U_0\setminus S}\sum_{i\in S} q_{ji} \le C \sum_{j\in U_0\setminus S}\sum_{i\in S} x_{ji}$$

So

$$C \sum_{j\in U_0\setminus S}\sum_{i\in S} x_{ji} \ge \sum_{i\in S} d_i$$

that is,

$$\sum\nolimits_{j\in U_0\setminus S}\sum\nolimits_{i\in S} x_{ji} \ge \sum\nolimits_{i\in S} d_i / C$$

2 The result is easy to get and its proof is omitted here.

3 The reduction to absurdity is used to prove this sufficient condition.

Subtours are the tours that do not depart from and return to the depot. The subtours can be traced with $x_{ij} > 0$.

If there is a subtour that satisfies the constraints (12), we select all customers in this subtour to compose a set $S \subseteq \{1,\ldots,N\}$. Then for this subtour,

1 $x_{ji} = x_{ij} = q_{ji} = q_{ij} = 0,\ j \in U_0 \setminus S, i \in S$ and

2 $\exists x_{ji} > 0,\ i, j \in S.$

From the inequality $\sum_{j\in U_0\backslash S}\sum_{i\in S}x_{ji}\geq\sum_{i\in S}d_i/C$ derived in the proof of 1, we have $\sum_{i\in S}d_i=0$. Then, $d_i=0,\ \forall i\in S$ since all d_i are non-negative. In this case for the minimum problem of IRPSD, in its optimal solution, $x_{ij}=q_{ij}=0\ \ \forall i,j\in S$ according to constraints (13), which contradicts the existence of the subtour.

Note: Constraints (12) and (13) are neither a sufficient condition for defining a feasible solution of SDVRP nor a necessary condition for its subtour elimination. The model given here is only used to provide a lower bound of the optimal objective value of IRPSD.

The following property is proved by Dror and Trudeau (1990) in case of SDVRP with classical VRP cost structure. It can also be applied to the model *P* of IRPSD. Here we give the property directly without proof.

Theorem 2: *If model P is feasible, and* c_{ij} *i,j = 1,…, N satisfy the triangle inequality, then*

1. *it has an optimal solution where in each period no two routes with the same direction have more than one common customer*
2. $x_{ijt}\in\{0,1\}\ \ i,j=1,\ldots N.$

Note that the inventory decisions and VRP decisions of model *P* are interrelated only through the delivery qualities d_{it}. As soon as the qualities are given in each period, the corresponding SDVRP of the model in the period holds the same properties of classical SDVRP given by Theorem 2.

According to Theorem 2 and q_{i0t} = 0, *i* = 1,…, *N t* = 1,…, *T*, the model *P* can be transformed into the following equivalent model, denoted by *P1*.

Model *P1*:

$$Z=\min\sum_{t=1}^{T}\sum_{i=1}^{N}h_{it}I_{it}+\sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{i=0}^{N}c_{ij}q_{ijt}+\sum_{t=1}^{T}\sum_{i=1}^{N}f_t x_{i0t}+\sum_{t=1}^{T}\sum_{i=1}^{N}c_{i0}^{b}x_{i0t} \tag{15}$$

s.t. constraints (2)–(9), (11) and

$$x_{ijt}\in\{0,1\},x_{i0t},x_{0it}\ \text{are}\ integer\ i,j=1,\ldots,N\ \ i\neq j,\ \ t=1,\ldots,T \tag{16}$$

4 Lagrangian relaxation approach

IRPSD is more difficult than classical SDVRP because of its simultaneously considering inventory planning and vehicle routing. Since SDVRP is NP-hard, so is IRPSD. This inspires us to seek for an approximate approach to solve the problem. In this section we present a LR algorithm to compute a lower bound on the optimal objective value and to construct a near-optimal feasible solution of model *P1*. The constraints that complicate the resolution of this model are constraints (8), which couple the quantity transported on directed arc (*i*, *j*) in period *t*, q_{ijt}, with the number of times traversed by vehicles on directed arc (*i*,*j*) in period *t*, x_{ijt}. Here we relax the constraints by introducing a set of non-negative Lagrange multipliers $\lambda = (\lambda_{ijt})_{(N+1)\times N\times T}$. Then the Lagrangian relaxed problem is

$$\begin{aligned} Z_\lambda(I,d,q,x) &= \sum_{t=1}^{T}\sum_{i=1}^{N} h_{it} I_{it} + \sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{i=0}^{N} c_{ij} q_{ijt} + \sum_{t=1}^{T}\sum_{i=1}^{N} f_t x_{i0t} + \sum_{t=1}^{T}\sum_{i=1}^{N} c_{i0}^{b} x_{i0t} \\ &+ \sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{i=0}^{N} \lambda_{ijt}\left(q_{ijt} - Cx_{ijt}\right) = \sum_{t=1}^{T}\sum_{i=1}^{N} h_{it} I_{it} \\ &+ \sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{i=0}^{N}\left(\lambda_{ijt} + c_{ij}\right) q_{ijt} + \sum_{t=1}^{T}\sum_{i=1}^{N}\left(c_{i0}^{b} + f_t\right) x_{i0t} - C\sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{i=0}^{N} \lambda_{ijt} x_{ijt} \end{aligned} \tag{17}$$

s.t. (2)–(7), (9), (11), (16), and $\lambda \geq 0$.

The relaxed problem can be decomposed into the two independent subproblems: an inventory subproblem and a routing subproblem.

The inventory subproblem (denoted by INV), which determines the *I*, *d*, *q* values, is formulated as:

$$Z_\lambda^1(I,d,q) = \min \sum_{t=1}^{T}\sum_{i=1}^{N} h_{it} I_{it} + \sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{i=0}^{N}\left(\lambda_{ijt} + c_{ij}\right) q_{ijt} \tag{18}$$

s.t.

$$I_{it} = I_{i,t-1} + d_{it} - r_{it} \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{19}$$

$$I_{i,t-1} + d_{it} \leq V_i \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{20}$$

$$\sum_{\substack{j=0\\ j\neq i}}^{N} q_{jit} - \sum_{\substack{j=0\\ j\neq i}}^{N} q_{ijt} = d_{it} \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{21}$$

$$\sum_{i=1}^{N} q_{0it} = \sum_{i=1}^{N} d_{it} \quad t = 1,\ldots,T \tag{22}$$

$$q_{i0t} = 0 \quad i = 1,\ldots,N \ \ t = 1,\ldots,T \tag{23}$$

$$I_{it} \geq 0 \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{24}$$

$$d_{it} \geq 0 \ \ i = 1,\ldots,N,\ q_{ijt} \geq 0 \quad i = 0,\ldots,N \ \ j = 1,\ldots,N \ \ i \neq j \ \ t = 1,\ldots,T \tag{25}$$

From (19), we have $I_{it} = I_{i,0} + \sum_{\tau=1}^{t}(d_{i\tau} - r_{i\tau})$. The subproblem can then be reformulated as

$$\begin{aligned} Z_\lambda^1(d,q) &= \min \sum_{t=1}^{T}\sum_{i=1}^{N} h_{it}\left(I_{i,0} + \sum_{\tau=1}^{t}\left(d_{i\tau} - r_{i\tau}\right)\right) \\ &+ \sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{i=0}^{N}\left(\lambda_{ijt} + c_{ij}\right) q_{ijt} \end{aligned} \tag{26}$$

s.t.

$$I_{i,0} + \sum_{\tau=1}^{t} d_{i\tau} - \sum_{\tau=1}^{t-1} r_{i\tau} \le V_i \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{27}$$

$$\sum_{\substack{j=0 \\ j\neq i}}^{N} q_{jit} - \sum_{\substack{j=0 \\ j\neq i}}^{N} q_{ijt} = d_{it} \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{28}$$

$$\sum_{i=1}^{N} q_{0it} = \sum_{i=1}^{N} d_{it} \quad t = 1,\ldots,T \tag{29}$$

$$q_{i0t} = 0 \quad i = 1,\ldots,N \ \ t = 1,\ldots,T \tag{30}$$

$$I_{i,0} + \sum_{\tau=1}^{t} \left(d_{i\tau} - r_{i\tau} \right) \ge 0 \quad i = 1,\ldots,N,\ t = 1,\ldots,T \tag{31}$$

$$d_{it} \ge 0 \ \ i = 1,\ldots,N,\ q_{ijt} \ge 0 \quad i = 0,\ldots,N \ \ j = 1,\ldots,N \ \ i \neq j \ \ t = 1,\ldots,T \tag{32}$$

The routing subproblem (denoted by *ROU*), which determines the x values, is formulated as:

$$Z_\lambda^2(x) = \min \sum_{t=1}^{T} \sum_{i=1}^{N} \left(c_{i0}^b + f_t \right) x_{i0t} - C \sum_{t=1}^{T} \sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{i=0}^{N} \lambda_{ijt} x_{ijt} \tag{33}$$

s.t.

$$\sum_{\substack{j=0 \\ j\neq i}}^{N} x_{ijt} = \sum_{\substack{j=0 \\ j\neq i}}^{N} x_{jit} \quad i = 0,\ldots,N, \quad t = 1,\ldots,T \tag{34}$$

$$x_{ijt} \in \{0,1\} \quad j \neq i, \quad x_{i0t}, x_{0jt} \text{ are } \textit{integer}, \quad i, j = 1,\ldots,N, \quad t = 1,\ldots,T \tag{35}$$

For *ROU*, it can be further decomposed into T subproblems, one for each period, given by:

$$Z_\lambda^{2t}\left(x_{(t)} \right) = \min \sum_{i=1}^{N} \left(c_{i0}^b + f_t \right) x_{i0t} - C \sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{i=0}^{N} \lambda_{ijt} x_{ijt} \tag{36}$$

s.t.

$$\sum_{\substack{j=0 \\ j\neq i}}^{N} x_{ijt} = \sum_{\substack{j=0 \\ j\neq i}}^{N} x_{jit} \quad i = 0,\ldots,N \tag{37}$$

$$x_{ijt} \in \{0,1\} \quad j \neq i, \quad x_{i0t}, x_{0jt} \text{ are } \textit{integer} \ i, j = 1,\ldots,N \tag{38}$$

For each given $\{\lambda_{ijt}\}_{(N+1) \times N \times T} > 0$, we have:

$$Z_\lambda(I, B, d, q, x) = Z_\lambda^1(d, q) + Z_\lambda^2(x) = Z_\lambda^1(d, q) + \sum_{t \in T} Z_\lambda^{2t}\left(x_{(t)} \right)$$

and

$$Z_\lambda(I,d,q,x) \le Z^*$$

where Z^* is the optimal objective value of model *P1*.

The subproblem *INV* is a linear programming and can be easily solved by using a commercial software such as Lingo.

The subproblem *ROU* can be transformed into a MCF problem that can be solved by using the out-of-kilter algorithm or Klein, Jewell, Busacker & Gowan's method (see e.g. Wolsey, 1998). These algorithms run in polynomial time, and have a computational complexity lower than the fastest linear programming algorithms. For our *ROU* subproblem, it is solved by using a MCF algorithm after relaxing $x_{ijt} \in \{0,1\}$ to $0 \le x_{ijt} \le 1$. The optimal solution of the relaxed subproblem is still integral since it is a network flow problem and hence has the totally unimodular structure.

The LR approach maximises the objective of the corresponding dual problem by using the subgradient method. We use an adaptive step sizing strategy to set the step size of the method in each iteration. The algorithm procedure of the approach is given as follows.

Step 0 Initiate $\lambda^0 = 0$, $\theta^0 = 1$ and $k = 0$.

Step 1 Solve the inventory subproblem and the routing subproblem.

Step 2 Set step size s^k in iteration k by

$$s^k = \beta \frac{Z^* - Z^k}{\left\| g^k \right\|^2}$$

where β is a parameter with $0 < \beta < 1$. Z^k is the current dual value (lower bound); Z^* is estimated by $(1+\omega/\theta^\rho)Z^{[k]}$, where $Z^{[k]}$ is the best dual value obtained prior to iteration k, parameters $\omega \in [0.1, 1.0]$, $\rho \in [1.1, 1.5]$ and the value of θ in iteration $k + 1$ is given by $\theta^{k+1} = \max(1, \theta^k - 1)$ if $Z^k > Z^{[k]}$, otherwise $\theta^{k+1} = \theta^k + 1$;

$$\left\| g^k \right\|^2 = \sum_{t=1}^{T} \sum_{\substack{j=1 \\ j \ne i}}^{N} \sum_{i=0}^{N} \left(q_{ijt}^k - Cx_{ijt}^k \right)^2$$

Step 3 Update the Lagrange multipliers in iteration $k + 1$

$$\lambda_{ijt}^{k+1} = \max\left\{ \lambda_{ijt}^k + s^k \left(q_{ijt}^k - Cx_{ijt}^k \right), 0 \right\}$$

Step 4 Check the stopping criterion.

The criterion may be given by

1 The dual value Z^k is not improved for a given number of iterations, or
2 A given maximal total iteration number is reached.

If the criterion is met, stop and output all required results. Otherwise, set $k = k + 1$ and go to Step 1.

The solution of the LR problem not only provides a lower bound for model *P1*, but also can be used to construct a near optimal feasible solution of *P1*.

5 Construction of feasible solution

5.1 Construction of feasible solution of model P1

Based on the *d*, *q* values obtained by solving the Lagrangian relaxed problem, a feasible solution of model *P1* can be constructed by solving the following problem, denoted by *P2*.

Model *P2*:

$$Z_\lambda(x) = \min \sum_{t=1}^{T} \sum_{i=1}^{N} \left(c_{i0}^{b} + f_t \right) x_{i0t} \tag{39}$$

s.t.

$$\sum_{\substack{j=0 \\ j \neq i}}^{N} x_{ijt} = \sum_{\substack{j=0 \\ j \neq i}}^{N} x_{jit} \quad i = 0,\dots,N \quad t = 1,\dots,T \tag{40}$$

$$\left\lceil \frac{q_{ijt}}{C} \right\rceil \leq x_{ijt} \quad i = 0,\dots,N, \quad j = 1,\dots,N \quad i \neq j \quad t = 1,\dots,T \tag{41}$$

$$x_{ijt} \in \{0,1\} \quad j \neq i, \quad x_{i0t}, x_{0jt} \text{ are } \textit{integer } i, j = 1,\dots,N \quad t = 1,\dots,T \tag{42}$$

where q_{ijt} $i = 0,\dots,N,$ $j = 1,\dots,N$ $i \neq j$ are obtained from the solution of the relaxed problem. Problem *P2* can be reformulated again as a minimal cost flow problem as follows.

$$Z_\lambda(x) = \min \sum_{t=1}^{T} \sum_{i=1}^{N} \left(c_{i0}^{b} + f_t \right) x_{i0t} \tag{43}$$

s.t.

$$\sum_{\substack{j=0 \\ j \neq i}}^{N} x_{ijt} = \sum_{\substack{j=0 \\ j \neq i}}^{N} x_{jit} \quad i = 0,\dots,N \quad t = 1,\dots,T \tag{44}$$

$$\left\lceil \frac{q_{ijt}}{C} \right\rceil \leq x_{ijt} \leq 1 \quad i = 1,\dots,N, \quad j = 1,\dots,N \quad i \neq j \quad t = 1,\dots,T \tag{45}$$

$$x_{0jt} \geq \left\lceil \frac{q_{0jt}}{C} \right\rceil, \quad j = 1,\dots,N \quad t = 1,\dots,T \tag{46}$$

$$x_{ijt} \geq 0 \quad j \neq i, \quad i, j = 0,\dots,N \quad t = 1,\dots,T \tag{47}$$

The problem can be decomposed into *T* subproblems, one for each period. Each subproblem can be solved by using an MCF algorithm.

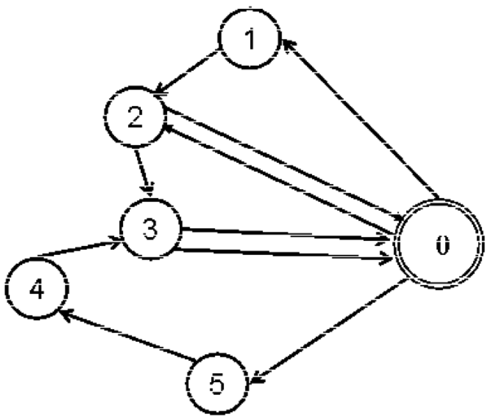
In order to obtain a better upper bound of model *P1*, the LR approach constructs a feasible solution in last several iterations or in every iteration before the stopping criterion is met, and the best feasible solution is selected as the final solution of model *P1*.

5.2 Repair P1's solution to a feasible solution of IRPSD

In Section 5.1, a feasible solution, *d*, *q*, *x*, of *P1*, is obtained, but this solution may not define a feasible solution of the original problem IRPSD, in the sense that from *x*, *q*, we cannot construct a set of feasible vehicle routes for each period so that the number of times traversed and the total quantity transported by the vehicles on each directed arc that links any two customers or the central depot and a customer are exactly given by *x*, *q*. In order to obtain a feasible solution of IRPSD, a method is proposed to trace a set of feasible routes in every period based on the solution of model *P1*. Because the method is the same for every period, for simplification, *we omit the subscript t* of the corresponding variables and parameters in the following discussion.

With x_{ij} $i, j = 0,\ldots,N$ $i \neq j$ of the feasible solution of model *P1*, a directed *transportation network*, as illustrated in Figure 1, can be defined in each period t where two customer nodes (or the depot node and a customer node) *i*, *j* are connected with a directed arc (*i*, *j*) by x_{ij} times if $x_{ij} \geq 1$. The directed arcs associated with $\{x_{ji} \mid x_{ji} \geq 1, j = 0,\ldots,N\}$ are called *incoming* arcs of customer node *i*, and the directed arcs associated with $\{x_{ij} \mid x_{ij} \geq 1, j = 0,\ldots,N\}$ are called *outgoing* arcs of customer node *i*. For customer node *i*, $\{q_{ji} \mid x_{ji} \geq 1, j = 0,\ldots,N\}$ form its *inflows*, and $\{q_{ij} \mid x_{ij} \geq 1, j = 0,\ldots,N\}$ form its *outflows*.

Figure 1 Directed transportation network



In this network, if the number of incoming arcs and the number of outgoing arcs of every customer node are both equal to 1, that is, each customer's delivery is realised by a single vehicle, then a set of feasible routes can be naturally traced from the feasible solution *d*, *q*, *x* of *P1*. Thus the solution is also a feasible solution of IRPSD. Otherwise, there exists at least one customer node for which the number of its incoming arcs or the number of its outgoing arcs is greater than 1. In this case, some customers are common customers of multiple vehicle routes and two conditions must hold for tracing feasible routes:

1 any incoming arc of each customer must be matched with one of its outgoing arc and

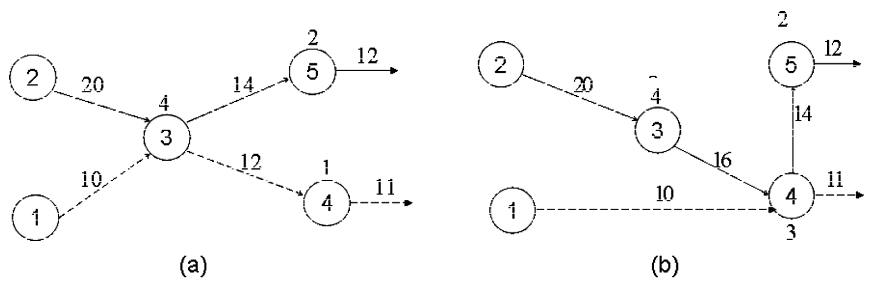
2 for each pair of matched arcs, the flow of the incoming arc must be no less than the flow of the outgoing arc.

If condition 1 or 2 cannot be satisfied, the feasible solution of *P2* will not be a feasible solution of the original IRPSD problem.

To determine for each incoming arc of a customer node its matched outgoing arc, an assignment problem needs to be solved. Because the number of incoming arcs of each customer is equal to that of its outgoing arcs according to constraints (44), condition 1 is satisfied naturally by the solution of the assignment problem. However, condition 2 may be violated by the solution. Thus, the objective of the assignment problem is to minimise the number of the incoming-outgoing arc matches that violate condition 2 by penalising them. When the optimal objective value of the assignment problem for a customer node is not equal to zero, there are matches that don't satisfy condition 2, and thus the values, x q, related to the customer node need to be *adjusted* so that the tracing of feasible routes can continue.

The adjustment is performed in the following way: an infeasible match of incoming arc and outgoing arc with minimum violation of condition 2 is first selected. Suppose that the incoming arc and the outgoing arc are (j, i) and (i, k), respectively, and that another outgoing arc of customer node i is (i, k'). The arcs (j, i) and (i, k') are then replaced by arcs (j, k) and (k, k'). The flows of arcs (j, k), (k, k') and (i, k) are adjusted accordingly to adapt to the arc replacement. For example, for the network in Figure 2(a), no matter how the incoming arcs and the outgoing arcs of customer node 3 are matched, it is impossible to make condition 2 satisfied. In this case, the value adjustment of x, q based on the solution of the assignment problem is required. The solution of the corresponding assignment problem for customer node 3 suggests that the incoming arc (1, 3) is matched with outgoing arc (3, 4) to trace a route, and another outgoing arc of the customer node is arc (3, 5). In this case, the values x, q in Figure 2(a) are adjusted to those in Figure 2(b). The value adjustment is important because it makes the matching of incoming arcs and outgoing arcs of customer node 3 feasible, and then *the procedure of tracing feasible routes can proceed* to customer node 4 – an immediate successor of customer node 3.

Figure 2 An example of the value adjustment



Note: The numbers associated with each directed arc (i, j) and node i are q_{ij} and d_j respectively.

When the matching procedure and the value adjustment for a customer node are finished, they will be repeated on the immediate successors of the customer node until all customer nodes have been examined.

For each ending customer node i, whose unique immediate successor is node 0, the number of its outgoing arcs is x_{i0} and the flows of all its outgoing arcs are zero ($q_{i0} = 0$), the objective value of any feasible solution of the corresponding assignment problem must be zero. That is, any match of incoming and outgoing arcs of the customer node is a feasible match. In this case, the above-mentioned value adjustment is not required and all feasible routes are then traced. As a result, a feasible solution of IRPSD is obtained.

In order to improve the quality of the solution of IRPSD, a local search is further adopted by merging two non-fully loaded vehicle routes if the merging is feasible and reduces the total cost.

6 Numerical examples

In this section, we evaluate the performance of our proposed algorithm using randomly generated problem instances. The parameters of the instances are generated in the following way: the length of the time horizon is taken as $T = 5$, which corresponds to five working days in each week; C, f_t, h_{it}, I_{i0}, V_i and r_{it} are randomly and uniformly generated from the intervals [100, 300], [400, 700], [0.5, 2], [400, 800], [50, 400], respectively. For c_{ij}, to ensure that the triangle inequality holds, we first generate the coordinates of all customers and the central depot from a 10 × 10 square, and then calculate c_{ij} as the physical distance between customers i and j. c_{i0}^b is set to $20 \times c_{i0}$.

Our algorithm was coded in C++ using callable library of Lingo 6.0. The algorithm constructs a feasible solution of model *P1* based on the solution of its relaxed problem in every iteration. The final feasible solution of IRPSD is obtained by repairing the best feasible solution of *P1*. The numerical test was performed on a Pentium IV 1.7GHz PC with 256 MB RAM and the termination condition for each instance for the LR approach is 150 iterations. The notations used for presenting the results of the instances are given in Table 1.

Table 1 Notations used in numerical results

N_0	Sum of the number of customers and the depot
CT_{avg}	Average computational time (minutes: seconds)
CT_{min}	Minimal computational time (minutes: seconds)
CT_{max}	Maximal computational time (minutes: seconds)
Gap_{avg}	Average value of (Upper Bound-Lower Bound)/Upper Bound × 100%
Gap_{min}	Minimal value of (Upper Bound-Lower Bound)/Upper Bound × 100%
Gap_{max}	Maximal value of (Upper Bound-Lower Bound)/Upper Bound × 100%

In order to evaluate the performance of our method for problems with different sizes, four scenarios with N_0 = 50, 100, 150 and 200, respectively, are tested. The results for these scenarios are summarised in Table 2. For each scenario, 30 instances were randomly generated and tested.

Table 2 Numerical results

N_0	CT_{min}	CT_{max}	CT_{avg}	$Gap_{min}(\%)$	$Gap_{max}(\%)$	$Gap_{avg}(\%)$
50	4:09	5:58	4:46	0.62	4.97	2.93
100	20:20	24:02	21:55	1.38	5.87	3.47
150	36:55	48:12	43:29	1.47	6.71	3.88
200	75:22	89:59	82:42	2.56	6.47	4.31

From the results of the four scenarios, several important observations can be obtained as follows:

1. From Table 2, the average gaps between the upper bound and the lower bound for all scenarios are less than 5% with the largest and the lowest gaps being 6.71% and 0.62%, respectively, which shows that our algorithm can not only obtain a near-optimal solution, but also is insensitive to the change of the problem size N_0.

2. All instances are solved in a reasonable time, with the average computation time being only 82 min and 42 sec for the largest instances with $N_0 = 200$ on an ordinary computer. Each instance with $N_0 = 200$ contains 399,990 variables where 199,000 variables are integral. Note that we tested some instances with only 10 customers by directly solving them using LINGO 6.0 software, but no result was obtained for each of the instances after 24 hr of computation!

3. With the increase of the problem size, the increase of the average computation time is approximately linear with respect to the increase of the number of decision variables. For example, the average computation time for the instances with $N_0 = 50$ is 4 min and 46 sec, while the average computation time for the instances with $N_0 = 200$ in Table 2 is 82 min and 42 sec. With the increase of N_0 from 50 to 200, the number of the decision variables is increased by about 16 times from 24,990 to 399,990. At the same time, the computation time increase is about 17 times from 4 min and 46 sec to 82 min and 42 sec.

7 Conclusions

In this paper, in order to make large scale IRPSD tractable, we propose a new approximate model for it in which no variable is related to a specific vehicle since all vehicles involved are homogeneous. A Lagrangian relaxation approach is used to decompose the model into subproblems that can be solved by linear programming and MCF algorithms. Based on the solution of the Lagrangian relaxed problem, a MCF based heuristic is used to construct a feasible solution of the model which is further repaired to a near optimal solution of the IRPSD. The numerical experiments demonstrate that our proposed approach can obtain high quality solutions with the average optimality gap less than 5% for randomly generated instances with up to 200 customers in a reasonable computation time.

Acknowledgement

This work is supported by a post-doctoral fellowship granted by the Ministry of Research of France.

References

Belenguer, J.M., Martinez, M.C. and Mota, E. (2000) 'A lower bound for the split delivery vehicles routing problem', *Operations Research*, Vol. 48, No. 5, pp.801–810.

Campbell, A. and Savelsbergh, M.W.P. (2004) 'A decomposition approach for the inventory-routing problem', *Transportation Science*, Vol. 38, pp.488–502.

Campbell, A., Clarke, L. and Savelsbergh, M.W.P. (2002) 'Inventory routing in practice. The vehicle routing problem', in P. Toth and D. Viego (Eds). *SIAM Monographs on Discrete Mathematics and Applications*, pp.309–330.

Chandra, P. and Fisher, M.L. (1994) 'Coordination of production and distribution planning', *European Journal of Operational Research*, Vol. 72, pp.503–517.

Dror, M. and Ball, M. (1987) 'Inventory/routing: reduction from an annual to a short period problem', *Naval Research Logistics Quarterly*, Vol. 34, No. 6, pp.891–905.

Dror, M. and Trudeau, P. (1990) 'Split delivery routing', *Naval Research Logistics*, Vol. 37, pp.383–402.

Dror, M., Ball, M. and Golden, B. (1985) 'Computational comparison of algorithms for the inventory routing problem', *Annals of Operations Research*, Vol. 4, Nos. 1–4, pp.3–23.

Dror, M., Laporte, G. and Trudeau, P. (1994) 'Vehicle routing with split deliveries', *Discrete Applied Mathematics*, Vol. 50, pp.239–254.

Federgruen, A. and Zipkin, P. (1984) 'A combined vehicle routing and inventory allocation problem', *Operation Research*, Vol. 32, pp.1019–1036.

Fumero, F. and Vercellis, C. (1999) 'Synchronized development of production, inventory, and distribution schedules', *Transportation Science*, Vol. 33, No. 3, pp.330–340.

Golden, B.L., Assad, A.A. and Dahl, R. (1984) 'Analysis of a large scale vehicle routing problem with an inventory component', *Large Scale Systems*, Vol. 7, pp.181–190.

Ho, S.C. and Haugland, D. (2004) 'A tabu search heuristic for the vehicle routing problem with time windows and split deliveries', *Computers and Operations Research*, Vol. 31, pp.1947–1964.

Lee, C.G., Epelman, M.A., White III, C.C., et al. (2004) 'A shortest path approach to the multiple-vehicle routing problem with split pick-ups', *Working Paper*, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada.

Wolsey, L.A. (1998) *Integer Programming*, New York: Wiley.

Yu, Y.G., Chu, F. and Chen, H.X. (2006) 'A note on coordination of production and distribution planning', *European Journal of Operational Research,* article in press and available at: www.sciencedirect.com