



The vehicle rescheduling problem



Remy Spliet*, Adriana F. Gabor, Rommert Dekker

Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

ARTICLE INFO

Available online 20 September 2013

Keywords:

Vehicle routing
Rescheduling
Operational planning

ABSTRACT

The capacitated vehicle routing problem (CVRP) is the problem of finding a routing schedule to satisfy demand by supplying goods stored at the depot, such that the traveling costs are minimized. For operational purposes, in many practical applications a long term routing schedule is made, often based on average demand. When demand substantially differs from the average, constructing a new schedule is beneficial. The vehicle rescheduling problem (VRSP) is the problem of finding a new schedule that not only minimizes the total traveling costs but also minimizes the costs of deviating from the original schedule. In this paper a mathematical programming formulation of the rescheduling problem is presented as well as a heuristic solution method referred to as the two-phase heuristic. We provide sufficiency conditions for which it produces the optimal solution. Finally, we perform computational experiments to study the performance of the two-phase heuristic.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Scheduling and rescheduling

The capacitated vehicle routing problem (CVRP) is a classical problem in operations research. Consider a depot where goods are stored and a set of customers that have nonnegative demand for these goods. A set of homogeneous vehicles of finite capacity is available to transport the goods from the depot to the customers. The vehicles start and end their routes at the depot. Costs are incurred for traveling from one location to another. The CVRP is to find a routing schedule that describes the sequence of locations visited by every vehicle that minimizes the total traveling costs, while the capacity constraints are satisfied. The CVRP is known to be an NP-hard problem.

Many solution methods can be found in the scientific literature for the CVRP. The branch-and-cut scheme of Baldacci et al. [1] seems presently to be one of the most successful at solving CVRP instances of up to 100 customer locations. For larger problem instances, many heuristic algorithms have also been developed that are able to find good solutions with greater speed. An overview of exact and heuristic algorithms can be found in Fisher [2], Toth and Vigo [3], Laporte [4,5] and Laporte et al. [6] amongst others.

In the classical CVRP, demand is deterministic and known. A situation that often occurs in practice is that demand only becomes apparent at a late moment. For example, in the retail industry it is very common that the orders of the individual stores

are placed only a few days, sometimes even just one day, before delivery. In these situations it is beneficial for operational processes to determine the delivery schedule before the orders are placed. It is for instance very costly, if at all possible, to make a roster for delivery handling personnel shortly before they are needed. A common solution to this problem is to determine a long term schedule, henceforward *master schedule*, that serves as a guiding schedule over a certain period of time in which multiple deliveries are made. For example, such a master schedule would describe the weekly or even daily deliveries for a period of six months. In practice, master schedules are usually constructed by solving deterministic CVRP instances based on average customer demand as predicted for the upcoming period. In this paper we assume that such a master schedule is given.

A master schedule is thus made before demand realizations become apparent. As a result, when the demand becomes known, the master schedule may not be optimal due to inefficient use of the vehicles, or may even be infeasible due to violation of the capacity constraints. In such cases the master schedule needs to be deviated from. This is often necessary in practice, for example when demand of customers is highly correlated, as typically is the case in a retail chain. The construction of a new schedule, when demand realizations become known, will be referred to as *rescheduling*.

1.2. Effects of rescheduling

After rescheduling, the new schedule will typically deviate from the master schedule. This can have negative effects on a distribution network. Locations are visited in a different order, by different trucks or by different truck drivers than initially planned. This may cause confusion among drivers and negatively affect the

* Corresponding author.

E-mail address: Spliet@ese.eur.nl (R. Spliet).

regularization and personalization of service, as is also recognized by Bertsimas and Simchi-Levi [7] and Li et al. [8,9]. Furthermore, consider the situation where personnel is hired only for handling deliveries and deliveries do not arrive at the agreed upon moment due to a deviation in the schedule. Here, labor costs increase due to the fact that personnel has to work overtime or has to be hired for another shift. When rescheduling is done by constructing a completely new schedule, many deviations may occur resulting in high additional costs.

Our experience with Dutch retail companies has shown that currently rescheduling is often done manually. Dispatchers typically operate under the notion that when a route needs to be deviated from, costs are lower when the first deviation occurs at a later stage in the route. There are several arguments that support this notion. Firstly, when deviating at a late stage in each route large portions of the master schedule remain intact, diminishing the above-mentioned negative effects of rescheduling. Secondly, the changes made in this manner are easily communicated through the distribution network. Finally, when changing the first locations of a route, the dispatching times of the truck will be altered. However, changing the working hours of a driver at a late moment is very expensive and often practically infeasible.

1.3. The vehicle rescheduling problem

In this paper, we propose a rescheduling model in which the negative effects of deviating from the master schedule are incorporated. We introduce deviation costs, which are incurred for each route that deviates from the master schedule. Furthermore, the height of the deviation cost per route is dependent on the customer at which the first deviation occurs and the position in the route it has. In this way, we are able to model the above described notion of dispatchers that deviations early in a route are more costly than deviations late in a route.

Given a master schedule and a demand realization, the goal is to find a new schedule that minimizes the total traveling and deviation costs, while satisfying the capacity constraints. This problem will be referred to as the vehicle rescheduling problem (VRSP). This model is of particular interest to for instance retail chains that control both the supply chain and the stores, as they not only incur the transportation costs, but also both the deviation costs at the supply side and at the customer side.

1.4. Rescheduling in current literature

In the literature, rescheduling is mainly considered in conjunction with designing a master schedule. Given a rescheduling method, the master schedule is designed before demand is known such that the expected costs incurred after rescheduling are minimized. The rescheduling method proposed by Bertsimas [10] is maybe the most popular method in the literature. In this method, the master schedule is used until a vehicle arrives at a location where its cargo is depleted. After it has returned to the depot to refill, the vehicle resumes the master schedule from the location where it left off. Under this rescheduling protocol, for specific demand distributions the expected costs of a master schedule can easily be calculated. For this reason, the rescheduling method proposed by Bertsimas has been incorporated in models used to design a master schedule with minimal expected costs. Examples of solution methods to solve these models are the L-Shaped integer method to find the optimal master schedule by Laporte et al. [11], a tabu search heuristic by Gendreau et al. [12], a rollout algorithm by Secomandi [13] and an evolutionary algorithm by Tan et al. [14].

In a study by Groër et al. [15] they propose to reschedule in such a way that each customer is always visited by the same

driver and within the same time window. They apply this to a setting encountered in the small-package shipping industry in which a customer does not require service on all delivery days. In their paper they focus on generating a master schedule for large instances and do this using a local search heuristic. Similarly, Chen et al. [16] consider an arc-routing model for small-package delivery in which arcs that do not need service are skipped.

To the best of our knowledge, the literature on rescheduling strategies that take into account deviation costs is scarce. Li et al. [8,9] consider the problem of reassigning vehicles to trips, when one of the vehicles breaks down. In their model, costs are incurred when trips are delayed. The main application of this model is in passenger transportation, for situations where traveling costs and capacity constraints do not play an important role.

1.5. Contribution

In this paper, we introduce a novel model for the rescheduling problem which can be applied where demand is known shortly before vehicles are dispatched. We provide a mixed integer programming formulation based on a formulation of the CVRP by Baldacci et al. [1]. Using this formulation, the VRSP of moderate size can be solved by general purpose optimization software or slight modifications of existing algorithms for solving the CVRP.

Furthermore, we design a solution approach based on first removing the last locations of routes and rescheduling them. We will refer to this approach as the *two-phase heuristic*. We analyze the performance of this heuristic and derive sufficiency conditions on the value of the deviation costs for which the two-phase heuristic is guaranteed to give the optimal solution to the VRSP. Moreover, numerical experiments indicate that, in general, for low deviation costs the two-phase heuristic often provides optimal or near optimal solutions. Finally, we describe this algorithm in such a way that it can be implemented directly in existing commercial CVRP software available to many dispatchers in large distribution networks.

1.6. Outline

In the following section, the VRSP is described in detail and a mixed integer linear programming formulation is presented. In Section 3, the two-phase heuristic is presented. It is accompanied by an analysis of its behavior with respect to the deviation costs. Finally, in Section 4 the sensitivity of the solution to the VRSP with respect to the value of the deviation costs is investigated and the performance of the two-phase heuristic is studied by comparing its solutions with the optimal solutions of the VRSP.

2. The vehicle rescheduling problem formulation

In this section, first the vehicle rescheduling problem is defined. This is followed by a mixed integer programming formulation based on an existing model developed by Baldacci et al. [1]. In Section 4, we will use this MIP formulation in computational experiments.

2.1. Problem definition

Consider an undirected complete graph $G = (V, E)$. The set of nodes $V = \{0, 1, \dots, n+1\}$ corresponds to a starting depot 0, an ending depot $n+1$ and the set of customers $V' = \{1, \dots, n\}$. For every edge $(i, j) \in E$, traveling costs $c_{ij} \geq 0$ are given that satisfy the triangle inequality. We suppose that an unlimited number of vehicles of capacity Q are available for supplying goods to the

customers. Furthermore, for every location $i \in V'$ the demand q_i is given and satisfies $0 < q_i \leq Q$.

A route $r \in E$ is defined as a cycle in G including the depot. To every route $r = \{(0, i_1), \dots, (i_k, n+1)\}$ we associate the ordered set of vertices $\{i_1, \dots, i_k\}$. Throughout the paper we will use these representations interchangeably. It will be clear from the context whether the edge or the vertex representation is meant. A route r is called *feasible* when the total demand of the locations on r is less than or equal to the capacity of a vehicle, i.e., $\sum_{i \in r} q_i \leq Q$.

A routing schedule S is a collection of edge-disjoint routes such that all customers are included in exactly one route. Hence, $S = \bigcup_{i=1}^m r_i$, where for the routes r_1, \dots, r_m it holds that $r_i \cap r_j = \emptyset$ for $i \neq j$. A schedule S is called *feasible* when all routes in S are feasible. The set of all feasible schedules will be denoted by \mathbb{S} .

The classical CVRP, a closely related problem to the VRSP, can now be defined as finding a feasible schedule that minimizes the total traveling costs and can be formulated as

$$(\text{CVRP}) \quad \min_{S \in \mathbb{S}} \sum_{(i,j) \in S} c_{ij}.$$

Next we define the VRSP. Assume that a master schedule S_M is available. Note that this master schedule need not be feasible for all demand realizations, as capacity restrictions might be violated. The VRSP is to create a new feasible schedule S^* that minimizes both the traveling costs and the costs of deviating from the master schedule. Next we formally define a deviation and the accompanying deviation costs.

Consider a route $r_M = \{i_1, \dots, i_j, \dots, i_k\}$ in the master schedule S_M . When for a new schedule S there is a route $r \in S$ such that the first locations are $\{i_1, \dots, i_{j-1}\}$ and the following location, if any, is not i_j , then r *deviates from location i_j onwards in schedule S* . In other words, route r originates from route r_M and it has remained the same up to location i_j . Whenever a route deviates from location $i \in V'$ onwards, costs $u_i \geq 0$ are incurred. These costs are in practice often dependent not only on the location at which the master schedule is deviated from, but also on the position of that location in the route in the master schedule. However, as the master schedule is given and the position of each location in the route in the master schedule is fixed, introducing deviation costs per location is sufficient for our purposes. Thus, for a master schedule S_M and a schedule S , the *deviation costs* for a route from location i onwards are defined as

$$U(S_M, S, i) = \begin{cases} u_i & \text{if a route deviates from location } i \\ & \text{onwards in } S \text{ with respect to } S_M; \\ 0 & \text{otherwise.} \end{cases}$$

Throughout this paper we will assume that for each route in the master schedule that the deviation costs associated to the locations on that route are decreasing in the positions on that route.

For $r_M = \{i_1, \dots, i_j, \dots, i_k\}$ in the master schedule, if a route r deviates from location i_j onwards, we will refer to any i_v in route r_M for $j \leq v$ as a *rescheduled location*.

It is now possible to fully define the VRSP as finding a feasible schedule S^* such that it minimizes the total traveling and deviation costs for a given master schedule S_M

$$(\text{VRSP}) \quad \min_{S \in \mathbb{S}} \left[\sum_{(i,j) \in S} c_{ij} + \sum_{i \in V'} U(S_M, S, i) \right]$$

Note that if in the VRSP $u_i = 0 \forall i \in V'$ one obtains the classical CVRP. As the latter is NP-hard, the VRSP is also NP-hard.

2.2. Mixed integer programming formulation

Next we provide a mixed integer programming formulation of the VRSP which is a modification of an existing formulation of the CVRP.

As stated in Laporte [5], one of the most successfully used formulations of the CVRP is the two commodity flow formulation introduced by Baldacci et al. [1]. Due to the polynomial number of constraints and variables, a direct implementation of this formulation in general purpose mixed integer programming software is sufficient to find a solution to the CVRP for moderately sized instances. Moreover, any cutting-plane designed for the CVRP, like for instance generalized capacity constraints [1], can be applied here as well.

Next, we briefly discuss the parts inherited from the CVRP model of Baldacci et al. and refer the interested reader to their paper for more details. Furthermore, we elaborate on the addition of deviation costs.

For all $(i, j) \in E$, let ξ_{ij} indicate whether edge (i, j) is included in the new route. Next, let the variables $x_{ij} \in \mathbb{R}_+$, for $i, j \in V$, be flow variables. When traveling from i to j , x_{ij} might be interpreted as the load of a vehicle and x_{ji} the remaining capacity.

To model the deviation costs, the variable y_i is introduced for each location $i \in V'$, which indicates whether a route is deviating from location i onwards. Observe that for any route r , $\sum_{j \in r} u_j y_j = \sum_{j \in r} U(S_M, S, j)$.

In the following formulation we will use j_{\leq} to denote the set containing j and all locations which are visited prior to j on the same route in the master schedule. For a given master schedule S_M , the mixed integer programming formulation of the VRSP is

$$(\text{CF}) \min \sum_{(i,j) \in E} c_{ij} \xi_{ij} + \sum_{i \in V'} u_i y_i \quad (1)$$

$$\sum_{j \in V} (x_{ji} - x_{ij}) = 2q_i \quad \forall i \in V' \quad (2)$$

$$\sum_{i \in V'} x_{0i} = \sum_{i \in V'} q_i \quad (3)$$

$$\sum_{i \in V'} x_{i0} = Q \sum_{i \in V'} \xi_{0i} - \sum_{i \in V'} q_i \quad (4)$$

$$\sum_{i \in V'} x_{n+1i} = Q \sum_{i \in V'} \xi_{0i} \quad (5)$$

$$x_{ij} + x_{ji} = Q \xi_{ij} \quad \forall (i, j) \in E \quad (6)$$

$$\sum_{j \in V, i < j} \xi_{ij} + \sum_{j \in V, i > j} \xi_{ji} = 2 \quad \forall i \in V' \quad (7)$$

$$1 - \xi_{ij} \leq \sum_{k \in j_{\leq}} y_k \quad \forall (i, j) \in S_M, \quad i < j, j \in V' \quad (8)$$

$$1 - \xi_{ij} \leq \sum_{k \in i_{\leq}} y_k \quad \forall (j, i) \in S_M, \quad i < j, j \in V' \quad (9)$$

$$\xi_{ij} \in \{0, 1\} \quad \forall (i, j) \in E \quad (10)$$

$$x_{ij} \in \mathbb{R}_+ \quad \forall i, j \in V' \quad (11)$$

$$y_i \in \{0, 1\} \quad \forall i \in V' \quad (12)$$

The set of constraints (2)–(5) and (11) ensure that x provides a correct flow pattern from the source nodes 0 and $n+1$ to the sink nodes V' and 0. By constraints (2) the difference between inflow and outflow at a customer location of both the vehicle load and the remaining vehicle capacity is equal to the demand. Constraint (3) ensures that the outflow of the starting depot is equal to the total demand of all customer locations and constraint (5) ensures that the inflow at the ending depot is the total capacity of the used trucks. Note that $\sum_{j \in V'} \xi_{0j}$ represents the number of vehicles that are used. Constraint (4) sets the remaining capacity of trucks leaving the depot.

Constraints (6) ensure that there is either no flow through edge (i, j) when this edge does not belong to any route or that the total load and empty space defined for this edge is exactly Q otherwise. Constraints (7) ensure that exactly two edges incident to any

customer are used. The objective function $\sum_{(i,j) \in E} c_{ij} x_{ij}$ together with constraints (2)–(7) and (10)–(11) gives a correct formulation of the CVRP, as indicated by Baldacci et al. [1].

Finally, the VRSP formulation is completed as (8), (9) and (12) force y_i to take value 1 whenever a route deviates from location i onward. Moreover, note that as $u_i \geq 0$ for all $i \in V'$ and the formulation describes a minimization problem, for each route at most one location will have value $y_i = 1$.

Note that when using the above formulation, an optimal solution may exist including paths from 0 to 0 and from $n+1$ to $n+1$. The solution can in this case be transformed into a solution solely with paths from 0 to $n+1$, without an increase in costs. For paths from 0 to 0 or from $n+1$ to $n+1$ the corresponding flow variables x cannot be interpreted as the load of a vehicle or remaining capacity at an edge.

Note that the integrality of y can be relaxed without compromising the validity of the two commodity flow formulation.

3. Two-phase rescheduling heuristic

In this section we propose and analyze a two-phase heuristic. The main idea behind the two-phase heuristic is to start with the possibly infeasible master schedule S_M and modify it to make it feasible, resulting in a new schedule S^{TP} . In the first phase of the heuristic, a specific set of edges is removed from the master schedule and in the second phase new edges are added such that the obtained schedule is feasible and has low deviation costs. Next we describe the two phases in more detail.

3.1. Phase 1: removing edges

When choosing the edges that will be removed from S_M , the main criterion is to limit the total deviation costs that are incurred in the resulting schedule S^{TP} . For any route $r \in S_M$, the final edge is removed. Next, for each route the last edge is removed and this is repeated iteratively until the total demand of the remaining locations in r does not exceed Q . Denote by \hat{V} the set of resulting isolated locations, these will be the rescheduled locations in S^{TP} .

The result of Phase 1 is a rooted tree S_1 with root node 0 and vertex set $V \setminus \hat{V}$, representing the set of incomplete routes. The total demand of the locations on any path from the root to a leaf is at most Q .

Fig. 1 shows an example of a network of a single depot and several customers. The solid and dashed lines combined show the original schedule. The numbers next to the customers correspond to a realization of demand. The vehicles have a capacity of 10 units of demand. The dashed lines correspond to edges that are removed during the first phase.

3.2. Phase 2: adding edges

In this phase, edges are added to the incomplete schedule S_1 such that it becomes a feasible schedule S^{TP} . This is done at minimal additional traveling costs. The problem that needs to be solved can therefore be defined as

$$S^{TP} = \arg \min_{S \in \mathbb{S}_{S_1 \cup \hat{V}}} \sum_{(i,j) \in S} c_{ij} \quad (13)$$

This is an instance of the CVRP in which certain edges are fixed. In some standard CVRP software, it may not be possible to prescribe the use of certain edges in the generated solution. In these cases, (13) can be reformulated as a CVRP without fixing edges by using artificial customer locations as follows. Contract each path in S_1 from root to a leaf into a node. Let V_{CR} be the set of these contracted nodes. The costs of using edges connecting

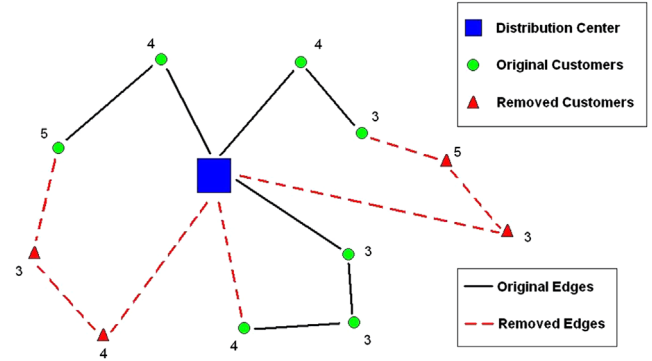


Fig. 1. Example phase 1: removed edges from the master schedule.

any two vertices in $\hat{V} \cup \{0, n+1\}$ remain unchanged. For $i \in V_{CR}$ let c_{0i} be equal to the costs of traversing the edge starting at the depot and ending at the first location on the path contracted into i . Similarly, for $j \in V_2 \cup \{0\}$ let c_{ij} be the costs of traversing the edge starting at the last location on the path contracted into i and ending at location j . Furthermore, for $i \neq 0$ and $j \in V_{CR}$ let $c_{ij} = \infty$.

For each $i \in V_{CR}$, let q_i be the total demand of the locations on the path contracted into i . Clearly, after the first phase of the heuristic, the demand q_i for $i \in V_{CR}$ does not exceed the vehicle capacity Q . The demand for the locations in \hat{V} does not change.

Consider a solution to the CVRP problem defined on the complete graph with the set of customer locations $V_{CR} \cup \hat{V}$ with demand and costs as defined above. As the costs of using an edge between any location in \hat{V} and any vertex in V_{CR} are infinite, in an optimal schedule any node in V_{CR} will be preceded only by the depot. A feasible schedule to the VRSP is now found by expanding back all the contracted nodes.

The problem that has to be solved in the second phase of the heuristic is obviously an NP-hard problem as the CVRP can be reduced to it. Fortunately, in most practical cases, the size of this CVRP is small. This is due to the fact that the number of nodes is equal to the number of routes in the master schedule (these are the artificial nodes) plus the number of isolated nodes, which is relatively low.

Next we present some properties of the solution obtained by the two phase heuristic.

3.3. Properties of the two-phase heuristic

Consider the problem of finding a feasible schedule that minimizes the total deviation costs when a master schedule S_M is given as follows:

$$U^* = \min_{S \in \mathbb{S}_{S_1 \cup \hat{V}}} U(S_M, S, i) \quad (14)$$

In the next proposition it is shown that when the deviation costs are decreasing on each route, i.e. $u_i \geq u_j$ for $(i,j) \in S_M$, the deviation costs of the schedule obtained by the two-phase heuristic are equal to U^* and that the number of rescheduled locations in this schedule is minimal.

Proposition 1. *If the deviation costs are decreasing on each route, the two-phase heuristic produces a feasible schedule S^{TP} such that the number of rescheduled locations and the total deviation costs are minimized.*

Proof. For a feasible schedule S , denote by V_S the set of rescheduled locations. We will next show that the minimum number of rescheduled locations is $|\hat{V}|$, by proving that $\hat{V} \subseteq V_S$ for any feasible

schedule S . Consider a location $j \in \hat{V}$ and the route $r = \{i_1, \dots, i_j, \dots, i_k\}$ in S . If $i_j \notin V_S$, none of the locations i_1, \dots, i_j are rescheduled and it must hold that $\sum_{l=1}^j q_{i_l} \leq Q$. However, this contradicts the construction of \hat{V} , hence $\hat{V} \subseteq V_S$. As \hat{V} is the set of rescheduled locations in S^{TP} , $|\hat{V}|$ is the minimum number of rescheduled locations.

Since $\hat{V} \subseteq V_S$, any schedule S deviates from the same locations onward as the schedule S^{TP} or from earlier locations. As the deviation costs are decreasing on each route, the two-phase heuristic provides a schedule such that the deviation costs are minimized. \square

Note that in order to determine the minimum number of rescheduled locations and the minimum deviation costs, one can apply the procedure described in the first phase of the heuristic, which does not require the construction of a schedule.

For certain values of the parameters, the schedule that minimizes the total deviation costs achieves the optimal value for the VRSP. This is in particular the case when the costs of deviating are very large relative to the traveling costs. In such a case the two-phase heuristic produces the optimal schedule for the VRSP, as stated in the next proposition. For notational convenience, let $u_{n+1} = 0$ and $u_i^\delta = u_i - u_j$ for all (i, j) in the master schedule.

Proposition 2. Let $u_{\min} = \min_{i \in V'} u_i^\delta$, $c_{\min} = \min_{(i,j) \in E} c_{ij}$ and let S^{TP} be the schedule produced by the two-phase heuristic. If

$$u_{\min} \geq \sum_{i \in V'} [c_{0i} + c_{i,n+1}] - \left(n + \left\lceil \frac{\sum_{i \in V'} q_i}{Q} \right\rceil \right) c_{\min},$$

then the schedule S^{TP} is optimal.

Proof. For any schedule S that is a feasible solution to the VRSP, let $Z_S = \sum_{(i,j) \in S} c_{ij}$ denote the traveling costs and $U_S = \sum_{i \in V'} U(S_M, S, i)$ the deviation costs.

Note that $Z_{S^{TP}} \leq \sum_{i \in V'} [c_{0i} + c_{i,n+1}]$. Since $\lceil \sum_{i \in V'} q_i / Q \rceil$ is a lower bound on the number of vehicles that are needed, for every feasible solution S to the VRSP it holds that $Z_S \geq (n + \lceil \sum_{i \in V'} q_i / Q \rceil) c_{\min}$. Therefore,

$$Z_{S^{TP}} \leq Z_S + \sum_{i \in V'} [c_{0i} + c_{i,n+1}] - \left(n + \left\lceil \frac{\sum_{i \in V'} q_i}{Q} \right\rceil \right) c_{\min}.$$

Furthermore, $U_{S^{TP}} = U^*$. If S has at least one more rescheduled location than \hat{V} , then $U^* + u_{\min} \leq U_S$. Hence, for $u_{\min} \geq \sum_{i \in V'} [c_{0i} + c_{i,n+1}] - (n + \lceil \sum_{i \in V'} q_i / Q \rceil) c_{\min}$, it follows that

$$Z_{S^{TP}} + U^* \leq Z_S + U_S. \quad (15)$$

If S and S^{TP} deviate from the same locations onward, (15) is always satisfied. This proves the optimality of S^{TP} for all instances with $u_{\min} \geq \sum_{i \in V'} [c_{0i} + c_{i,n+1}] - (n + \lceil \sum_{i \in V'} q_i / Q \rceil) c_{\min}$. \square

Tight example. To show that the bound on u_{\min} cannot be improved, consider the following example. Let $V = \{0, 1, 2\}$ and $c_{ij} = c$, for all $i, j \in V$. The master schedule S_M consists of a return trip to location 1 and a separate return trip to location 2. The demand realizations are such that $q_1 + q_2 \leq Q$ and therefore the two routes might feasibly be merged. Furthermore, let $u_1 = u_2 = u_{\min}$. There are only three feasible solutions to this VRSP. The master schedule can be used as a solution to the rescheduling problem and yields a total cost of $4c$. The other two solutions visit nodes 1 and 2 on one route and both have costs $3c + u_{\min}$. The two-phase heuristic will produce S_M . The resulting schedule will be optimal if and only if $u_{\min} \geq c = \sum_{i \in V'} [c_{0i} + c_{i,n+1}] - (n + \lceil \sum_{i \in V'} q_i / Q \rceil) c_{\min}$.

For specific problem instances the relative difference between the traveling and the deviation costs need not be high for the two-phase heuristic to produce the optimal solution. However, we

are not able to provide a general guarantee for small relative differences. In Section 4, we analyze the impact of the numerical values of the deviation costs on the optimality of the solution obtained by the two-phase heuristic.

In the next proposition a worst case bound is provided on the ratio of the solution value of the solution provided by the two-phase heuristic and the optimum.

Proposition 3. The costs of using the routing schedule S^{TP} produced by the two-phased heuristic are at most $\min\{Q/q_{\min}, 2Qc_{\max}/(Q + \bar{q})c_{\min} + 1\}$ times the costs of the optimal schedule S^* for the VRSP, where $q_{\min} = \min_{j \in V} q_j$, $\bar{q} = \sum_{i \in V} q_i / n$, $c_{\max} = \max_{(i,j) \in E} c_{ij}$ and $c_{\min} = \min_{(i,j) \in E} c_{ij}$.

Proof. Let the traveling costs and deviation costs of S^{TP} be given by $Z_{S^{TP}}$ and U^* respectively. Similarly, let the traveling and deviation costs of S^* be given by Z_{S^*} and U_{S^*} respectively. Furthermore let $Z^* = \min_{S \in \mathcal{S}} \sum_{(i,j) \in S} c_{ij}$. To prove the proposition, it is first shown that $(Z_{S^{TP}} + U^*) / (Z_{S^*} + U_{S^*}) \leq Q/q_{\min}$ and second that $(Z_{S^{TP}} + U^*) / (Z_{S^*} + U_{S^*}) \leq 2Qc_{\max}/(Q + \bar{q})c_{\min} + 1$. For ease of notation, assume $c_{0i} = c_{i,n+1}$.

In Simchi-Levi et al. [17, p. 220] it is proven that for the CVRP it holds that $2 \sum_{j \in V} c_{0j} q_j \leq QZ^*$. Now observe that

$$Z_{S^{TP}} \leq 2 \sum_{i \in V'} c_{0i} \leq \frac{2}{q_{\min}} \sum_{i \in V'} c_{0i} q_i \leq \frac{Q}{q_{\min}} Z^*,$$

which implies that

$$\frac{Z_{S^{TP}} + U^*}{Z_{S^*} + U_{S^*}} \leq \frac{Q}{q_{\min}} \frac{(Z^* + U^*)}{Z^* + U^*} \leq \frac{Q}{q_{\min}}.$$

Next, as $\lceil \sum_{i \in V'} q_i / Q \rceil$ is a lower bound on the number of vehicles that are used, it follows that

$$\begin{aligned} \frac{Z_{S^{TP}} + U^*}{Z_{S^*} + U_{S^*}} &\leq \frac{Z_{S^{TP}} + U^*}{Z^* + U^*} < \frac{2nc_{\max}}{\left(n + \left\lceil \frac{\sum_{i \in V'} q_i}{Q} \right\rceil \right) c_{\min}} + 1 \leq \frac{2nc_{\max}}{\left(n + \frac{\sum_{i \in V'} q_i}{Q} \right) c_{\min}} \\ &+ 1 = \frac{2Qc_{\max}}{(Q + \bar{q})c_{\min}} + 1 \end{aligned}$$

Here the strict inequality follows from $(a+b)/(c+b) < a/c + 1$ for $a, b, c > 0$. This concludes the proof. \square

Tight example. In this example we show that the bound provided in Proposition 3 cannot be improved upon. Consider a problem instance of n locations and let an arbitrary master schedule be given. Now let demand be given by $q_i = Q$ for all $i \in V'$. Obviously there is only one feasible schedule, hence $(Z_{S^{TP}} + U^*) / (Z_{S^*} + U_{S^*}) = 1 = Q/q_{\min} = \min\{Q/q_{\min}, 2Qc_{\max}/(Q + \bar{q})c_{\min} + 1\}$.

4. Numerical experiments

In this section, results of numerical experiments are presented to provide insight into the sensitivity of the model with respect to different values of the deviation costs. Furthermore, the performance of the two-phase heuristic is evaluated empirically by applying it to several test cases.

The following settings are used for the generation of individual problem instances:

- n customer locations are randomly generated according to a uniform distribution over a square with sides of length 20 units. The depot is situated in the center of the square.
- The traveling costs between two locations are equal to the Euclidean distance between them.

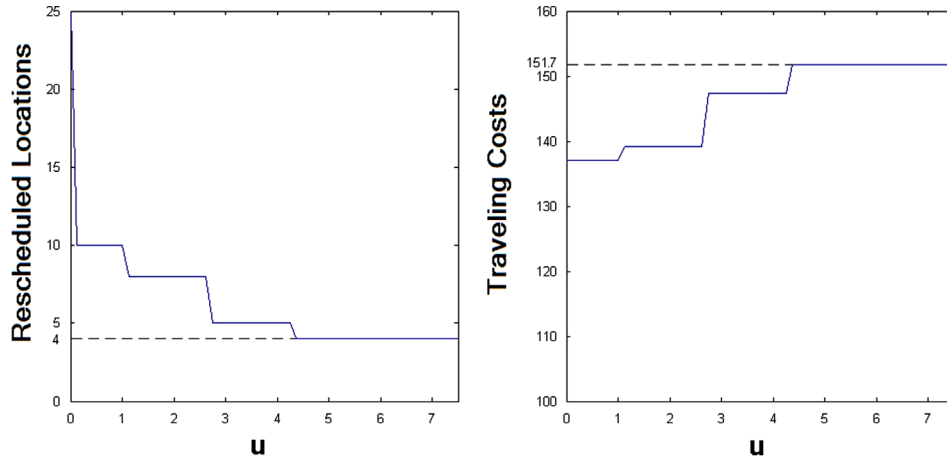


Fig. 2. Number of rescheduled locations and traveling costs.

- All vehicles have a capacity of 60 units.
- Presumed demand is normally distributed with mean 5 and standard deviation 1.5, truncated from below to 1 and from above to 60.
- Actual demand per location is normally distributed with standard deviation 1.5 and the demand average equal to 1.5 times the realization of the presumed demand, also truncated from below to 1 and from above to 60.

For each experiment, we indicate the corresponding deviation costs. For every problem instance, first a master schedule S_M is generated by solving a CVRP using presumed demand. This is done either exactly or heuristically depending on the experiment at hand. Next a demand realization is generated to represent actual demand. The deviation costs will be specified for every individual experiment. As actual demand will typically be higher than presumed demand in our experiments, deviations from the master schedule will most often be necessary. These instances are inspired by a practical case in a retail chain with recurrent sales actions.

We have implemented the branch-and-cut algorithm by Baldacci et al. [1] to solve the CVRP to optimality. It uses their two-commodity flow formulation to find lower bounds. These are strengthened by separating capacity cuts using a greedy randomized algorithm. This algorithm is used to generate the master schedule in some instances and to solve the CVRP in the second phase of the two-phase heuristic. We use the same algorithm to solve the VRSP to optimality by adding constraints (8), (9) and (12) to the formulation. For the instances where the master schedule is found by solving a CVRP heuristically, we use the savings algorithm by Clarke and Wright [18].

All experiments are performed on a Pentium(R) Dual-Core CPU, E5800, 3.2 GHz with 4.00 GB of RAM. The branch-and-cut algorithm makes use of ILOG CPLEX 12.3 to solve the linear programming relaxations.

4.1. Impact of deviation costs

Recall that, by Proposition 2, for large values of u_{\min} with respect to the traveling costs, the two phase heuristic gives the optimal solution. It is, thus, interesting to study whether optimality is also obtained for lower values of u_{\min} . We will refer to the lowest value of u_{\min} for which the two-phase heuristic produces the optimal schedule as the *critical level* and we will denote it by u_{critical} .

Next we will argue that in order to analyze u_{critical} it is sufficient to look at the number of rescheduled locations in an optimal

schedule. Note that the minimal number of rescheduled locations can easily be determined by applying the first phase of the rescheduling heuristic and, by Proposition 1, the solution of the two phase heuristic has the minimal number of rescheduled locations.

If an optimal solution has a minimal number of rescheduled locations, the locations that deviate in both the optimal solution and the solution provided by the two-phase heuristic are identical. Since the two-phase heuristic inserts the deviating locations such that the traveling costs are minimized, the schedule produced by the two-phase heuristic must be optimal. Hence, in order to assess the optimality of the schedule generated by the two-phase heuristic, it is sufficient to look at the number of deviations in the optimal schedule.

For the numerical experiments in this paragraph we use the following deviation costs. For each route r in the master schedule and costs u , we assign deviation costs $(|r|+1-i)u$ to the i th location on r , where $|r|$ indicates the number of locations visited by r . Hence, $u_{\min} = u$. We generate an instance with $u=0$ and solve it. Next, we repeatedly modify the deviation costs by increasing the value of u by 0.125, and solve the modified instance. We repeat this until the two-phase heuristic provides the optimal solution for a modified instance.

Let us first look at the value of u_{critical} for an example. Consider a single randomly generated instance of 25 customer locations. For this example, the upper bound on u_{critical} given in Proposition 2 is equal to 349.65. As remarked in Section 3.3, the minimum number of rescheduled locations can be found beforehand by applying the first phase of the two-phase heuristic. The optimal schedules are found using a direct implementation of the two commodity flow formulation of the VRSP.

In Fig. 2 the number of rescheduled locations and traveling costs in the optimal solution is depicted, for different values of u . Notice that when $u=0$, all locations are rescheduled. However, as u grows slightly above 0, a new schedule is found with less rescheduled locations but with equal traveling costs.

Fig. 2 illustrates the fact that the number of rescheduled locations in the optimal schedule decreases as u grows. In this particular instance, the minimum number of rescheduled locations is 4 and the critical value is 4.25, a much lower value than the theoretical upper bound. However, the value of u_{critical} is meaningless unless related to the traveling costs. Let \bar{c}_M be the average of the traveling costs over the edges used in the master schedule. For the example depicted in Fig. 2, $\bar{c}_M = 4.13$, which can be considered very close to the critical level u of 4.25.

We have repeated this experiment for 100 randomly generated instances. In 29 of the cases, u_{critical} lies below $0.5\bar{c}_M$, in 61 cases

Table 1Deviation costs average $0.25\bar{c}_M$.

n	BL2	BL3	CD_{TP}	T_{opt}	T_{TP}	OPT not found
10	141.4 (20.5)	2022 (1230)	2.9 (4.4)	0.13	0.01	0
20	294.2 (28.0)	2795 (1413)	2.4 (2.7)	7.92	0.05	0
30	440.5 (28.9)	3835 (1724)	3.0 (2.8)	256.83	0.37	2
40	587.5 (39.6)	3960 (1637)	4.9 (3.9)	1483.44	1.84	37

Table 2Deviation costs average $0.75\bar{c}_M$.

n	BL2	BL3	CD_{TP}	T_{opt}	T_{TP}	OPT not found
10	141.4 (20.5)	2022 (1230)	0.9 (2.1)	0.235	0.01	0
20	294.2 (28.0)	2795 (1413)	1.0 (1.6)	17.636	0.06	0
30	440.5 (28.9)	3835 (1724)	1.2 (1.5)	481.763	0.43	4
40	587.5 (39.6)	3960 (1637)	1.0 (0.7)	2436.45	1.16	45

below \bar{c}_M and in 82 cases below $1.5\bar{c}_M$. Observe that 53 of the critical levels do not differ more than 50% of the value of \bar{c}_M . When we calculate the bound derived in Proposition 2, the two-phase heuristic could only have been guaranteed to generate the optimal schedule for $u \geq 102.4\bar{c}_M$ on average.

Finally, we discuss the tradeoff between transportation costs and the number of rescheduled locations. The schedules with minimal traveling costs in the first example have a traveling cost equal to 137.1. Among these schedules the best in terms of number of rescheduled locations is a solution with 10 rescheduled locations. The schedule with minimal number of rescheduled locations, 4, has a traveling costs equal to 151.7. Observe that in this case, the number of rescheduled locations decreases by 6 while the traveling costs increase by 10.6%. For the 100 generated cases, the average increase in traveling costs between the schedule with minimal traveling costs and the schedule with minimal number of rescheduled locations is 10.8%, with standard deviation 5.1.

4.2. Algorithm performance

The performance of the two-phase heuristic is evaluated on multiple test instances. For these cases, deviation costs decreasing in locations per route are obtained by generating a positive cost decrease u_i^p for each customer $i \in V$. We use a normal distribution with mean equal to either $0.25\bar{c}_M$ or $0.75\bar{c}_M$ and a standard deviation of $0.5\bar{c}_M$. The cost decreases are truncated from below at 0. These parameters were chosen such that it is unlikely that the generated instances either revert to standard CVRP because all u are near or equal to 0, or that they are sufficiently high so that the two-phase heuristic is guaranteed optimal.

The performance of the two-phase heuristic is compared to solving the VRSP to optimality using the branch-and-cut algorithm. For each instance a time limit of 1 h is maintained for both the heuristic and the exact algorithm.

In Tables 1 and 2 the results of computational experiments for instances of different sizes are presented. The master schedule is generated by solving a CVRP to optimality. For each value of n , representing the number of customer locations, 50 instances were generated. The instances used for Table 1 have an average deviation cost decrease equal to $0.25\bar{c}_M$ and the ones used for Table 2 have an average cost decrease equal to $0.75\bar{c}_M$.

Column BL2 shows the average value of the theoretical bound on u_{\min} in Proposition 2, and the standard deviation in between brackets. Column BL3 presents the average worst case bound as described in Proposition 3, expressed in percentages, and the standard deviation in between brackets. Column CD_{TP} shows, in

Table 3Deviation costs average $0.25\bar{c}_M$, heuristic master schedule.

n	BL2	BL3	CD_{TP}	T_{opt}	T_{TP}	OPT not found
10	137.3 (18.5)	1863 (745)	3.4 (6.3)	0.142	0.011	0
20	290.5 (27.4)	2645 (1143)	2.7 (3.4)	2.686	0.056	0
30	443.1 (32.7)	3925 (1778)	4.5 (3.5)	232.103	0.425	0
40	598.1 (42.5)	4135 (1807)	4.6 (3.3)	939.925	3.567	22

percentages, the average difference between the cost of the schedule produced with the two-phase heuristic and the cost of the optimal schedule, and the standard deviation in between brackets. Note that these costs include both the traveling costs and the deviation costs. Finally, the values in T_{opt} and T_{TP} represent average running times in seconds of the exact algorithm and the two-phase heuristic respectively. As a master schedule was assumed to be given, the time needed to generate it is not incorporated. The column OPT not found indicates the number of instances out of 50, for which the optimal solution was not found within a 1 h time limit. These instances were not considered in the presented averages.

In these experiments, the total costs of the schedules generated by the two-phase heuristic are on average not more than 2.9% above the optimum in Table 1 and not more than 1.0% in Table 2. As can be seen in column BL3, this differs significantly from the theoretical performance bound in Proposition 3. As expected, the solutions produced by the two-phase heuristic are on average closer to the optimum for the instances with average cost decreases of $0.75\bar{c}_M$ than for the instances with average cost decreases of $0.25\bar{c}_M$.

From the columns indicating running times it can be concluded that using the two-phase heuristic reduces the solution time significantly with respect to solving it to optimality using the branch-and-cut algorithm.

Out of the 39 instances with low deviation costs for which the optimal solution is not found, a feasible solution was found for 14 instances using the branch-and-cut algorithm. The average difference in solution value with respect to the two-phase heuristic is 4.4%. The two-phase heuristic solved these instances using an average computation time of 262.17 s. Out of the 49 instances with high deviation costs for which the optimal solution is not found, a feasible solution was found for 23 instances using the branch-and-cut algorithm. The average difference in solution value with respect to the two-phase heuristic is -0.7% . The two-phase heuristic solved these instances using an average computation time of 3.32 s.

4.3. Impact of the master schedule

Using an inefficient master schedule with respect to traveling costs, might affect the performance of the two-phase heuristic. When rescheduling using an inefficient master schedule, the traveling costs might be considerably reduced at the expense of deviating early in a route. In such cases, the two-phase heuristic will not perform well as it never generates unnecessarily early deviations. To investigate the effect of the quality of the master schedule on the performance of the two-phase heuristic, the experiment is repeated using instances where the master schedule is obtained by solving a CVRP heuristically using the savings algorithm by Clarke and Wright [18]. Tables 3 and 4 show the results of these experiments.

From the instances used in Tables 3 and 4 more are solved within the 1 h time limit than from the instances in Tables 1 and 2. Moreover, the average computation time of the exact method is lower. Nevertheless, the computation time of the two-phase

Table 4Deviation costs average $0.75\bar{c}_M$, heuristic master schedule.

n	BL2	BL3	CD_{TP}	T_{opt}	T_{TP}	OPT not found
10	137.3 (18.5)	1863 (745)	1.7 (4.5)	0.198	0.011	0
20	290.5 (27.4)	2645 (1143)	1.3 (2.3)	3.768	0.049	0
30	443.1 (32.7)	3925 (1778)	1.8 (2.4)	493.469	0.489	3
40	598.1 (42.5)	4135 (1807)	1.5 (1.5)	1187.79	2.814	30

Table 5

Minimum number of deviating locations.

n	Exact CVRP master schedule	Heuristic CVRP master schedule
10	2.4 (1.0)	2.0 (1.1)
20	4.6 (1.4)	5.0 (1.4)
30	8.8 (2.0)	8.2 (2.0)
40	14.7 (2.4)	12.0 (3.1)

heuristic is slightly higher for the instances where the master schedule is generated heuristically. The difference between the optimal solution value and the value of the solution produced by the two-phase heuristic is not significantly different for the instances where the master schedule is generated heuristically.

4.4. Number of deviating locations

The minimum number of deviating locations is an important factor for the speed of the two-phase heuristic. In the second phase a CVRP with the number of customers roughly equal to the minimum number of deviating locations should be solved. It depends on the instance at hand what the minimum number of deviations is. The average minimum number of deviating locations of the instances used in Section 4.2 can be found in Table 5, as well as the standard deviation in between brackets.

In our implementation of the two-phase heuristic, we use an exact method to solve the reduced CVRP in the second phase. Many heuristic algorithms exist that can be used to deal with for instance large scale CVRP instances. When the size of the second phase CVRP grows too large for exact algorithms to be useful, such a heuristic can be employed. As the performance of existing exact algorithms and heuristics directly translate to our setting, we refer the interested reader to Laporte [5]. For a software library of heuristic methods for the CVRP that can be used in the second phase of the heuristic see for instance Groër et al. [19].

5. Conclusion

In this paper, the negative effects of deviating from a master schedule have been incorporated in a vehicle routing model, hence introducing the VRSP. Insight has been obtained into the behavior of the optimal solution of the VRSP for different values of the deviation costs relative to the traveling costs. We have formulated this problem that allows existing CVRP algorithms to be used to solve the VRSP after slight modifications.

Furthermore, we propose a two-phase heuristic that is capable of finding good solutions within a small amount of computation time. We have proven that this algorithm generates an optimal schedule when deviation costs are sufficiently high. Even when the deviation costs are not as high as required by our proposition, numerical experiments show that solutions of the two-phase heuristic are on average close to optimal. Moreover, for general problem instances, an analytical bound on the difference between the solution generated by the two phase heuristic and the optimum is presented. Numerical results indicate, however, that this analytical upper bound is extremely far from the actual difference. As in the second phase of this heuristic an instance of the CVRP, an NP-hard problem, has to be solved for the locations that need to be rescheduled, the computation time heavily depends on the number of isolated vertices after the first phase of the heuristic.

References

- [1] Baldacci R, Hadjiconstantinou E, Mingozzi A. An exact algorithm for the capacitated vehicle routing problem based on a two-commodity network flow formulation. *Operations Research* 2004;52(5):723–38.
- [2] Fisher M. Vehicle routing. In: Ball MO, Magnanti TL, Monma CL, Nemhauser GL, editors. *The handbooks in operations research and management science*, vol. 8. Elsevier Science; 1995.
- [3] Toth P, Vigo D. Models, relaxations and exact approaches for the capacitated vehicle routing problem. *Discrete Applied Mathematics* 2002;123:487–512.
- [4] Laporte G. The vehicle routing problem: an overview of exact and approximate algorithms. *European Journal of Operational Research* 1992;59:345–58.
- [5] Laporte G. What you should know about the vehicle routing problem. *Naval Research Logistics* 2007;54(8):811–9.
- [6] Laporte G, Gendreau M, Potvin J, Semet F. Classical and modern heuristics for the vehicle routing problem. *International Transactions in Operational Research* 2000;7:285–300.
- [7] Bertsimas DJ, Levi DS. A new generation of vehicle routing research: robust algorithms, addressing uncertainty. *Operations Research* 1996;44(2):286–304.
- [8] Li J, Mirchandani PB, Borenstein D. The vehicle rescheduling problem: model and algorithms. *Networks* 2007;50(3):211–29.
- [9] Li J, Mirchandani PB, Borenstein D. A Lagrangian heuristic for the real-time vehicle rescheduling problem. *Transportation Research Part E* 2009;45:419–33.
- [10] Bertsimas DJ. A vehicle routing problem with stochastic demand. *Operations Research* 1992;40(3):574–85.
- [11] Laporte G, Louveaux FV, van Hamme L. An integer L-Shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research* 2002;50(3):415–23.
- [12] Gendreau M, Laporte G, Séguin R. A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. *Operations Research* 1996;44(3):469–77.
- [13] Secomandi N. A rollout policy for the vehicle routing problem with stochastic demands. *Operations Research* 2001;49(5):796–802.
- [14] Tan KC, Cheong CY, Goh CK. Solving multiobjective vehicle routing problem with stochastic demand via evolutionary computation. *European Journal of Operational Research* 2007;177:813–39.
- [15] Groër C, Golden B, Wasil E. The consistent vehicle routing problem. *Manufacturing & Service Operations Management* 2009;11(4):630–43.
- [16] Chen S, Golden B, Wong R, Zhong H. Arc-routing models for small-package local routing. *Transportation Science* 2009;43(1):43–55.
- [17] Simchi-Levi D, Chen X, Bramel J. *The logic of logistics, theory, algorithms, and applications for logistics and supply chain management*. Springer series in operations research. 2nd ed. New York: Springer Science+Business Media, Inc; 2005.
- [18] Clarke G, Wright JW. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research* 1964;12(4):568–81.
- [19] Groër C, Golden B, Wasil E. A library of local search heuristics for the vehicle routing problem. *Mathematical Programming Computation* 2010;2(2):19–101.