The Vehicle Rescheduling Problem

Remy Spliet[a*]

Adriana F. Gabor[a]

Rommert Dekker[a]

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

E12009-43

November 24, 2009

Abstract

The capacitated vehicle routing problem is to find a routing schedule describing the order in which geographically dispersed customers are visited to satisfy demand by supplying goods stored at the depot, such that the traveling costs are minimized. In many practical applications, a long term routing schedule has to be made for operational purposes, often based on average demand estimates. When demand substantially differs, constructing a new schedule is beneficial. The vehicle rescheduling problem is to find 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 two mathematical programming formulations of the rescheduling problem are presented as well as two heuristic methods, a two-phase heuristic and a modified savings heuristic. Computational and analytical results show that for sufficiently high deviation costs, the two-phase heuristic generates a schedule that is on average close to optimal or even guaranteed optimal, for all considered problem instances. The modified savings heuristic generates schedules of constant quality, however the two-phase heuristic produces schedules that are on average closer to the optimum.

Keywords: Vehicle routing, vehicle rescheduling problem, operational planning

1 Introduction

The capacitated vehicle routing problem (CVRP) is a classical problem in operations research. Consider a depot where goods are stored and a set of locations which have nonneg-

*Corresponding author: spliet@ese.eur.nl

ative demand for the goods. A set of 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 that is visited by every vehicle, in such a way that the total traveling costs are minimized. The CVRP is known to be an NP-hard problem.

Because the CVRP has many practical applications, a wide scala of solution methods can be found in the literature. The branch-and-cut scheme of Baldacci et al. (2004) seems presently to be the most successful at solving CVRP instances of up to 100 customer locations. Using this scheme, CVRP instances of up to 50 locations can consistently be solved to optimality while this can only be done for specific problem instances including more 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 (1995), Toth and Vigo (2002), Laporte (1992 and 2007) and Laporte et al. (2000) 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 situations of frequent periodic deliveries, it is beneficial for operational processes to determine the moment of delivery before the orders are placed. It is for instance very costly, if at all possible, to roster delivery handling personnel one day before they are needed. It is therefore very common to determine a long term schedule, henceforward *master schedule*, that serves as a schedule for every periodic delivery over a certain period of time in which multiple deliveries are made.

Such a master schedule is made before demand realizations become apparent. As a result the master schedule will not always be feasible as for instance high demand might cause the capacity of a vehicle to be insufficient to make deliveries to all locations on its route planned in the master schedule. In such cases the master schedule needs to be deviated from. Moreover, low demand may lead to inefficient use of vehicle capacity, such that lower traveling costs might be obtained by deviating from the master schedule. The construction of a new schedule when demand realizations become known, will be referred to as *rescheduling*.

In the literature, rescheduling is mainly considered in conjunction with designing a master schedule. A very popular method for rescheduling is the one proposed by Bertsimas (1992). In this method, the master schedule is used until a vehicle arrives at a location where its cargo

is depleted. Next it returns to the depot to refill. Finally it resumes the master schedule where it left off. When the demand distribution is known and when this rescheduling method is used, the expected costs of a master schedule can easily be calculated. Furthermore, this method allows for simple upper and lower bounds to be calculated for the expected costs of the master schedule. The main application area of the rescheduling method proposed by Bertsimas is when demand is not learned prior to the arrival of a vehicle at the depot. For situations where the demand realization is known prior to the dispatching of the vehicles, more efficient schedules can be obtained when using a different method of rescheduling. As the rescheduling problem itself also has many practical applications, the main focus of our paper will be to model and solve the rescheduling problem, disregarding the construction of a master schedule.

To minimize the traveling costs that are incurred after rescheduling, a CVRP can be solved using the demand realization to construct a new feasible routing schedule. This may, however, result in increased costs at the customer side of the network. For instance, when personnel is hired for handling deliveries, and deliveries arrive late, personnel is first idle followed by working overtime, increasing labor costs significantly. Furthermore, as also Bertsimas and Simchi-Levi (1996) recognize, stability in a distribution network is crucial; implementing new schedules may confuse drivers and regularity and personalization of service will not be guaranteed. The importance of rescheduling to practice, is also recognized by Li et al. (2007 and 2009). Given an assignment of vehicles to trips, they designed a model to reassign the vehicles to trips when one of the vehicles breaks down. Costs are incurred when trips are delayed. The main application of this model is in passenger transportation, where traveling costs and capacity constraints do not play an important role. In our paper the negative effects at the customer side of rescheduling are modeled by costs incurred when deviating from the master schedule. The vehicle rescheduling problem (VRSP) is to find the optimal trade-off between the traveling costs and the costs of deviation, while satisfying the capacity constraints. A new schedule has to be constructed that minimizes the total traveling costs as well as the total deviation costs.

Another option to compensate for fluctuating demand is stabilizing demand by keeping inventory. However, this may be impossible in practice by lack of storage space. Furthermore, as Haughton (1998) concludes, the benefits of demand stabilization strategies do not generally outweigh the costs.

In this paper a model is presented for the rescheduling problem where demand is known before vehicles are dispatched. We have encountered this situation in several retail chains. This allows for solution methods yielding more efficient schedules than methods based on

models where demand is assumed not to be known, as these models are typically used in the literature. Note that most literature on rescheduling is aimed at constructing the master schedule rather than designing a good model and finding accompanying solution methods for the rescheduling phase. Examples of this are the L-Shaped integer method to find the optimal master schedule by Laporte et al. (2002), a tabu search heuristic by Gendreau et al. (1996), a rollout algorithm by Secomandi (2001) and an evolutionary algorithm by Tan et al. (2007).

In the following section, the VRSP is presented. Two mixed integer linear formulations are included. They are extensions of mixed integer linear formulations for the CVRP, a closely related problem. In fact, the CVRP is a specific case of the VRSP and which implies that the VRSP is also an NP-hard problem. Heuristics will have to be used to find solutions for the VRSP in practical applications. Two heuristic algorithms are proposed: First the *two-phase heuristic* is introduced. This is a heuristic that makes use of specific properties of the VRSP model. An analytical bound on the difference between of the generated solution and the optimal solution is presented for this algorithm. Furthermore the *modified savings algorithm* is described. This is a basic extension to the VRSP setting of the savings algorithm proposed for the CVRP by Clarke and Wright (1964). The two-phase heuristic is mainly proposed for its potential to produce good solutions whereas the modified savings algorithm is presented as a comparative heuristic that is able to produce reasonably good solutions with great speed and that is easy to adapt to future extensions to the VRSP. Finally the sensitivity of the VRSP model to the value of certain parameters is investigated and the performance of the two heuristics is analyzed by comparing their solutions with each other and by comparing their solutions to the optimal solutions that are found using a direct implementation of the VRSP.

2 The vehicle rescheduling problem formulation

In this section VRSP is presented. First a description of this problem is given and the notation that will be used throughout the paper is introduced. This is followed by two alternative mixed integer linear formulations of the VRSP. These formulations are based on the most commonly used and most successful formulations of the capacitated vehicle routing problem (CVRP) in the literature.

2.1 Problem description

Consider a directed complete graph $G = (V, E)$. The set of nodes $V = \{0, 1, ..., n\}$ correspond to a single depot 0 and the customers $V' = \{1, ..., n\}$. 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 $Q \geq 0$ is at our disposal. Furthermore, for every location $i \in V'$ the demand q_i is given such that $Q \geq q_i > 0$. The vehicles will be used to supply demand. In this paper only none trivial problem instances will be discussed, it is assumed that $\sum_{j \in V'_M} c_{0j} q_j > 0$.

A *route* $r \subset E$ is defined as a cycle in G including the depot. For ease of notation we let every route r be accompanied by an ordered subset $r' \subseteq V'$. This allows us to represent a route $r = \{(0, i_1), ..., (i_k, 0)\}$ by $r' = \{i_1, ..., i_k\}$. A route is called feasible when the capacity of a vehicle allocated to drive this route is sufficient to contain the demand of all the locations on that route, $\sum_{i \in r'} q_i \leq Q$. The set including all feasible routes will be denoted by R.

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

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

$$\text{(CVRP)} \min_{S \in \mathbb{S}} \sum_{(i,j) \in S} c_{ij} \tag{1}$$

To define the VRSP we need the following additions to the CVRP. First of all, assume that a master schedule S_M is available. Note that this master schedule need not be feasible as capacity restrictions might be violated by demand realizations. The VRSP is to create a new feasible schedule S_R that not only minimizes the traveling costs but also takes costs of deviating from the master schedule into account. It is now left to define a deviation and the accompanying deviation costs.

A *deviation* is defined per location. We say that the new schedule does not deviate for location l when all locations visited prior to l on the route in both the master and the new schedule are the same and that it deviates otherwise. To be more precise, suppose location l is visited on route r_M in the master schedule S_M, r_M is accompanied by $r'_M = \{i_1^M, ..., i_v^M, l, ..., i_k^M\}$. In the new schedule S_R, l is visited on route r_R, which is accompanied by $r'_R = \{i_1^R, ..., i_w^R, l, ..., i_m^R\}$. When $v = w$ and $i_1^M = i_1^R, ..., i_v^M = i_w^R$, the new schedule does not deviate for location l; otherwise it does deviate. Therefore, if location l deviates, it immediately follows that all subsequent locations on the same route also deviate. As an example, suppose $r'_M = \{1, 2, 3, 4, 5, 6\}$ and $r'_R = \{1, 2, 4, 5, 7, 6\}$. The new schedule does not deviate for locations 1 and 2, but it does deviate for all locations 3 through 7 (we know that location 3 is moved to another route and 7 is moved from another route).

Whenever a new schedule deviates for location $i \in V'$, costs $u_i \geq 0$ are incurred. Let us therefore define the following function describing the incurred deviation costs for location i given an master and a new schedule, S_M and S_R respectively:

$$U(S_M, S_R, i) = \begin{cases} u_i, & \text{if } S_R \text{ deviates from } S_M \text{ for location } i; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

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

$$\text{(VRSP)} \quad \min_{S_R \in \mathbb{S}} [\sum_{(i,j) \in S_R} c_{ij} + \sum_{i \in V'} U(S_M, S_R, i)] \tag{3}$$

Note that the CVRP is a particular instance of the VRSP when $u_i = 0 \ \forall i \in V'$. As CVRP is NP-hard, so is VRSP.

2.2 Mixed integer linear formulations

To solve the VRSP, a mixed integer linear formulation can be used. Existing CVRP formulations can easily be extended to VRSP formulations. As sophisticated algorithms exist to solve the classical CVRP, they might be easily extended to incorporate the new features in a VRSP formulation to find a solution to this problem.

Perhaps the most commonly used formulation of the CVRP is the two index vehicle flow formulation. An extension of this yields an insightful version of the VRSP. A second formulation of the VRSP is based on the two commodity flow formulation of the CVRP. Although this formulation is less insightful, it is a powerful formulation of which direct implementation yields better results than using the the two index vehicle flow formulation of the VRSP. Computational results of using a direct implementation of this formulation are presented.

2.2.1 A two index vehicle flow formulation

The two index vehicle flow formulation of the CVRP was originally proposed by Laporte et al. (1985). Using this, solutions can be found by means of advanced branch-and-cut methods.

This formulation makes use of a function $b(W)$ for $W \subseteq V'$, denoting the minimum number of vehicles that are needed to satisfy the demand of all customers in W. Calculating $b(W)$ reduces to solving a bin-packing problem, however it can be shown that the CVRP formulation remains valid when replacing $b(W)$ by $\lceil \frac{\sum_{i \in W} q_i}{Q} \rceil$. Furthermore denote the complement of any set $W \subseteq V'$ by $\overline{W}$. For the two index vehicle flow formulation, let $\xi_{ij} \in \{0, 1\}$

indicate whether edge (i, j) is used or not.

A master schedule S_M is given. Introduce an indicator variable $y_i \in \{0, 1\}$ for $i \in V'$, which takes value 1 when location i deviates and 0 otherwise.

The two index vehicle flow formulation of the VRSP is given by:

$$(\mathbf{VF}) \ \min \sum_{(i,j)\in E} c_{ij}\xi_{ij} + \sum_{i\in V'} u_i y_i \tag{4}$$

$$\sum_{i\in V\setminus\{k\}} \xi_{ik} + \sum_{j\in V\setminus\{k\}} \xi_{kj} = 2 \ (k \in V') \tag{5}$$

$$\sum_{i\in W, j\in\overline{W}} \xi_{ij} \geq 2b(W) \ (\forall W \subseteq V') \tag{6}$$

$$1 - \xi_{ij} \leq y_j \ (\forall (i, j) \in S_M, j \in V') \tag{7}$$

$$y_j \leq y_k \ (\forall (i, j), (j, k) \in S_M, k \in V') \tag{8}$$

$$\xi_{ij} \in \{0, 1\} \ (i, j) \in E \tag{9}$$

$$y_i \in \{0, 1\} \ i \in V' \tag{10}$$

Constraints (5) ensure that every location is visited and departed from. Constraints (6) represent the generalized subtour constraints. They ensure that all created routes visit the depot and that the capacity constraints are not violated. A more detailed discussion of modeling the subtour elimination constraints can be found in Laporte (1986). The objective function $\sum_{(i,j)\in E} c_{ij}\xi_{ij}$ together with constraints (5),(6) and (9), give a correct formulation of the CVRP.

Constraints (7) are defined to force a variable y_j to take value 1 whenever edge (i, j) is used in the master schedule but not used in the solution of the VRSP. To complete the VRSP formulation, (8) ensure that whenever a location i is said deviates and therefore $y_i = 1$, then j also deviates and $y_j = 1$.

Finally note that in the literature it is common to assume symmetric traveling costs c_{ij}. If this is the case, G can be defined as an undirected graph and this effectively halves the number of decision variables in most mixed integer linear CVRP formulations. This can obviously also be done for VRSP formulations, however this yields a more cumbersome notation.

Lemma 1. *The integer program by (VF) is a correct formulation of the VRSP for a given master schedule S_M.*

Proof. To prove Lemma 1, first it will be shown that any solution to the VRSP can be transformed to a solution of (VF) and secondly the reverse will be shown.

First observe that any schedule S_R^* that is a solution to the VRSP can easily be transformed into a solution to (VF). Take $\xi_{ij} = 1$ when $(i,j) \in S_R^*$ and 0 otherwise. Furthermore, let $y_j = 1$ when location j deviates in the schedule S_R^* and 0 otherwise. Note that since (4), (5), (6) and (9) give a correct formulation of the CVRP, the chosen values of ξ satisfy them. Next take $(i,j),(j,k) \in S_M$. Consider constraint (7). When location k deviates, the constraint is trivially satisfied as $y_k = 1$. When location k does not deviate, edge (j,k) is used and $\xi_{jk} = 1$. which implies that (7) is again satisfied. Now consider constraint (8). When location k deviates, the constraint is trivially satisfied as $y_k = 1$. If location k does not deviate, this implies that not only edge $(j,k) \in S_R^*$ but also $(i,j) \in S_R^*$. Therefore $y_j = 0$ and (8) is also satisfied in this case. Inspection shows that for this choice of ξ and y it holds that $\sum_{(i,j)\in S_R^*} c_{ij} + \sum_{i\in V'} U(S_M, S_R^*, i) = \sum_{(i,j)\in E} c_{ij}\xi_{ij} + \sum_{i\in V'} u_i y_i$. Therefore an optimal solution of the VRSP can be transformed into a solution of (VF) with equal costs. The optimal value of the VRSP is thus at least the optimal value of (VF).

Let the optimal solution to (VF) be given by (ξ^*, y^*). This solution can be transformed into a solution of the VRSP as follows. Let $S_R = \{(i,j)|\xi_{ij}^* = 1\}$. As ξ^* satisfies (5), (6) and (9), it follows that $S_R \in \mathbb{S}$ and that $\sum_{(i,j)\in S_R} c_{ij} = \sum_{(i,j)\in E} c_{ij}\xi_{ij}^*$. It is now left to show that the total deviation costs of S_R, $\sum_{i\in V'} U(S_M, S_R, i)$, are equal to $\sum_{i\in V'} u_i y_i^*$. This is the case when for $u_k > 0$ it holds that $y_k^* = 1$ if and only if location k deviates in S_R.

First it will be proven that for $k \in V'$ with $u_k > 0$, if $y_k^* = 1$ then location k deviates in S_R. Suppose that there exists an $y_k^* = 1$ while location k does not deviate in S_R, for $(j,k) \in S_M$. Consider the solution (ξ^*, y') such that $y_i' = 0$ for all locations where $y_i^* = 1$ while location i does not deviate in S_R, and $y_i' = y_i^*$ otherwise. This solution does not violate any constraints. In particular the value of element y_k' does not cause violation of (7) and (8) as both $\xi_{jk}^* = 0$ and $y_j' = 0$. Moreover, it decreases the objective value as $u_k > 0$. Therefore a minimum is attained for the new solution (ξ^*, y'), contradicting the optimality of (ξ^*, y^*). We conclude that for $k \in V'$ with $u_k > 0$, when $y_k^* = 1$, location k deviates for the solution S_R.

Next it will be proved that for $k \in V'$ with $u_k > 0$, that $y_k^* = 1$ if location k deviates in S_R. Suppose that location k deviates for S_R while $y_k^* = 0$, for $(j,k) \in r \in S_M$. When $j = 0$, k is the first location on route r. This implies that $(0,k) \notin S_R$ and by construction it must hold that $\xi_{jk} = 0$ contradicting (7). For the first location k on any route, when location k deviates in S_R it follows that $y_i^* = 1$. When $j \in V'$, k is not the first location to be visited on

route r. As location j is visited prior to location k, observe that when j deviates in S_R then $y_j^* = 1$. If location k deviates for S_R this implies that either $(j,k) \notin S_R$ and by construction it must hold that $\xi_{jk} = 0$, contradicting (7), or there exists a location i deviating in S_R for $(i,j) \in S_M$ such that $y_i = 1$, contradicting (8). It follows that for $u_k > 0$, $y_k^* = 1$ if location k deviates for the solution S_R.

Summarizing, for $u_k > 0$ it holds that $y_k^* = 1$ if and only if location k deviates in S_R. It indeed follows that $\sum_{(i,j)\in S_R} c_{ij} + \sum_{i\in V'} U(S_M, S_R, i) = \sum_{(i,j)\in E} c_{ij} + \xi_{ij}^* \sum_{i\in V'} u_i y_i^*$. Therefore an optimal solution of the (VF) can be transformed into a solution of VRSP with equal costs. The optimal value of the VRSP is at most the optimal value of (VF). Hence, by the previous result, the optimal value of the VRSP is equal to the optimal value of the (VF). □

2.2.2 A two commodity flow formulation

As stated in Laporte (2007), the most successfully implemented formulation of the CVRP is the two commodity flow formulation, introduced by Baldacci et al. (2004). It lends itself to be solved using advanced branch-and-cut methods as is done in Baldacci et al. (2004). It can also be extended to a VRSP in a similar way as the two index vehicle flow formulation.

To simplify notation, assume symmetric costs c_{ij}. Let $n+1$ represent a copy of the depot and let $\hat{G} = (\hat{V}, \hat{E})$ be an the undirected graph where $\hat{V} = V \bigcup \{n+1\}$ and $\hat{E} = \{(i,j) | i,j \in \hat{V}, i < j\}$. A route can now be seen as an ordered subset $\tilde{r} \subset \hat{E}$ such that the induced subgraph $\hat{G}(\tilde{r})$ is a path from 0 to $n+1$. Note that as $\tilde{r}$ is an ordered subset, the edges are traversed in a particular order.

For all $(i,j) \in \hat{E}$, let ξ_{ij} take value 1 to indicate the use of edge (i,j) and 0 otherwise. Next define the variables $x_{ij} \in \mathbb{R}_+$ and $x_{ji} \in \mathbb{R}_+$for all $(i,j) \in \hat{E}$. These are flow variables that might be interpreted as the load of a vehicle x_{ij} and the remaining capacity x_{ji} when traversing edge (i,j).

For a given master schedule S_M, the two commodity flow formulation is:

$$\text{(CF)} \ \min \sum_{(i,j)\in E} c_{ij}\xi_{ij} + \sum_{i\in V'} u_i y_i \tag{11}$$

$$\sum_{j\in\hat{V}} (x_{ji} - x_{ij}) = 2q_i \ (\forall i \in V') \tag{12}$$

$$\sum_{j\in V'} x_{j0} = \sum_{i\in V'} q_i \tag{13}$$

$$\sum_{j \in V'} x_{j0} = Q \sum_{j \in V'} \xi_{0j} - \sum_{i \in V'} q_i \tag{14}$$

$$\sum_{j \in V'} x_{n+1j} = Q \sum_{j \in V'} \xi_{0j} \tag{15}$$

$$x_{ij} + x_{ji} = Q\xi_{ij} \ (\forall (i,j) \in \hat{E}) \tag{16}$$

$$\sum_{j \in \hat{V}, i<j} \xi_{ij} + \sum_{j \in \hat{V}, i>j} \xi_{ji} = 2 \ (\forall i \in V') \tag{17}$$

$$1 - \xi_{ij} \leq y_j \ (\forall (i,j) \in S_M, i < j, j \in V') \tag{18}$$

$$1 - \xi_{ij} \leq y_i \ (\forall (j,i) \in S_M, i < j, j \in V') \tag{19}$$

$$y_j \leq y_k \ (\forall (i,j),(j,k) \in S_M, k \in V') \tag{20}$$

$$\xi_{ij} \in \{0,1\} \ (i,j) \in \hat{E} \tag{21}$$

$$x_{ij} \in \mathbb{R}_+ \ (i,j) \in \hat{E} \tag{22}$$

$$x_{ji} \in \mathbb{R}_+ \ (i,j) \in \hat{E} \tag{23}$$

$$y_i \in \{0,1\} \ i \in V' \tag{24}$$

The set of constraints (12)-(15), (22) and (23) ensure that any solution x defines a correct flow pattern. By (12) the difference of both the vehicle load and the remaining capacity between inflow and outflow at a customer location is equal to the demand. Constraints (13)-(15) make sure that the in and outflow of the depot and its copy are matching when taking into account the amount of goods that are delivered per customer. Note that $\sum_{j \in V'} \xi_{0j}$ is the number of vehicles that are used. Constraints (16) ensure that there is either no flow through edge (i,j) when this edge does not belong to any route and that the total load and empty space defined for this edge is exactly Q otherwise. That exactly two edges incident to any customer are used, is ensured by constraints (17). The objective function $\sum_{(i,j) \in E} c_{ij}\xi_{ij}$

together with constraints (12)-(17) and (21)-(23), gives a correct formulation of the CVRP. Finally the VRSP formulation is completed as (18)-(20) and (24) force y_i to take value 1 whenever location i is interpreted to deviate and 0 otherwise.

As indicated in Baldacci et al. (2004), the set of solutions to the CVRP corresponds to the set of solutions of the two commodity flow formulation of the CVRP. A solution that describes a set of disjoint paths from node 0 to node $n+1$ can straightforwardly be interpreted as a routing schedule. It is possible however that an optimal solution describes paths from 0 to 0, or from $n+1$ to $n+1$. In this case a simple transformation can be made without increasing the objective value to find paths from 0 to $n+1$. For paths from 0 to 0 or from $n+1$ to $n+1$ the corresponding flow variables x can not be interpreted as the load of a vehicle or remaining capacity at an edge. These observations also hold in the two commodity flow formulation of the VRSP.

Lemma 2. *The integer program (CF) is a correct formulation of the VRSP for a given master schedule S_M.*

The proof of lemma 2 is analogue to that of lemma 1 and is therefore omitted. Observe that although the graph on which the problem is defined is in this case undirected, applying induction on the locations in an arbitrary route as is done in the proof of lemma 1 gives the desired result nonetheless.

3 Solution methods

The mixed integer linear formulations can not easily be solved using standard mixed integer programming software like CPLEX. Only small instances of the problem can be solved directly within reasonable computation times. This is already apparent as the VRSP can be considered to be an extension of the CVRP to which this comment also applies, as remarked in the introduction. For practical purposes, larger problem instances need to be considered. Therefore, two heuristics are presented in this section. They will be able to find a good feasible solution to large problem instances of the VRSP with limited computational effort.

3.1 Two-phase rescheduling 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. In the first phase of the heuristic, a set of edges is removed from the master schedule such that a set of locations $V'_R \subset V'$ no longer has any incident edges. The complement of V'_R, denoted by $V'_M = V' \backslash V'_R$, is the set of locations that still have edges incident to them. Next, what remains of the master schedule will be

completed again by adding edges such that all locations in V'_R are visited and the resulting schedule S_R^{TP} is feasible.

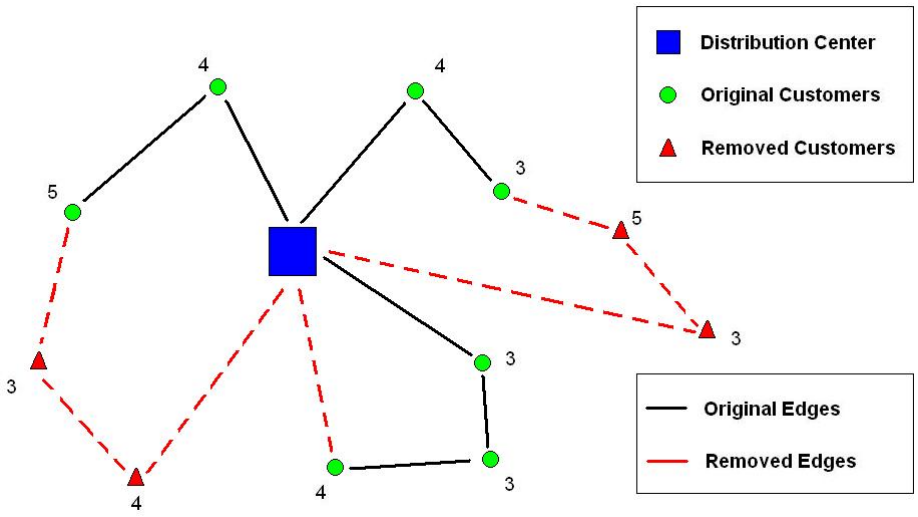
3.1.1 Phase 1: removing edges

When removing edges from S_M the main criterion is to limit the total deviation costs that are incurred in the resulting schedule S_R^{TP} of the heuristic. For any route $r \in S_M$, edges are removed in reverse order of traversal until the total demand of the locations that still have incident edges to them in r does not exceed Q. Note that the edge between the last location on a route and the depot is always removed.

The result of this procedure is a set of edges S'_R representing an incomplete schedule as a rooted tree with root node 0. The total demand of the locations on any path from the root to a leaf is less than or equal to Q. All locations in the set V'_M have incident edges in S'_R and the other locations in $V' \backslash V'_M$ do not.

Figure 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. After execution of the edge removal procedure, the dashed lines correspond to edges that are removed.

Figure 1:



3.1.2 Phase 2: Adding edges

In this phase, edges are added to the incomplete schedule S'_R such that it becomes a feasible schedule S_R^{TP}. This is done at minimal additional traveling costs. The problem that needs

to be solved can therefore be defined as:

$$S_R^{TP} = \arg \min_{S \in \mathbb{S} | S'_R \subset S} \sum_{(i,j) \in S} c_{ij} \tag{25}$$

In principal this is a restricted version of a CVRP. It is restricted in the sense that all edges $(i, j) \in S'_R$ need to be used. This can be solved using any CVRP algorithm that allows for certain edges to be included beforehand. For instance a method based on the two index vehicle flow formulation of the CVRP can be used by replacing ξ_{ij} with 1 for all $(i, j) \in S'_R$.

For implementation of problem (25) in standard CVRP software edges may not be predefined to be used in a solution. The nature of the heuristic or the software package used might prevent this. Note that (25) can be reformulated as an unrestricted CVRP using artificial customer locations, as follows. Contract all paths from root to leaf in S'_R into nodes. Let V'_{CR} be the set of these contracted nodes. The costs of using edges connecting any two vertices in $V'_R \bigcup \{0\}$ 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 $j \in V'_R \bigcup \{0\}$ 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 $c_{ij} = \infty$.

Let q_i for $i \in V'_{CR}$ 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 V'_R does not change.

Consider a solution to the CVRP problem defined on the complete graph with the set of customer locations $V'_{CR} \bigcup V'_R$ with demand and costs as defined above. As the costs of using an edge between any location in V'_R and any vertex in V'_{CR} is 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 it can be reformulated into an unrestricted CVRP. In this reformulation, there are as many artificial nodes defined as there were routes in the master schedule and there are locations that need to deviate. Fortunately, in most practical cases we encountered only a small subset of the customer locations need to deviate limiting the size of the CVRP in the second phase. For solving the CVRP any specialized exact algorithm or heuristic can be used.

Before presenting a bound on the difference between the solution generated by the two-phase heuristic and the optimal solution, some properties of its solutions are presented.

3.1.3 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:

$$U^* = \min_{S \in \$} \sum_{i \in V'} U(S_M, S, i) \tag{26}$$

In the next lemma it is shown that the deviation costs of the schedule obtained by the two phase heuristic are equal to U^*. This implies that in order to find U^*, one can apply the procedure described in the first phase of the heuristic, which does not require the construction of a schedule.

Lemma 3. *The two-phase heuristic produces a schedule S_R^{TP} such that the total deviation costs are minimized:*

$$\sum_{i \in V'} U(S_M, S_R^{TP}, i) = U^* \tag{27}$$

Proof. Since S_R^{TP} is a feasible solution for finding the schedule with the minimum deviation costs, $U^* \leq \sum_{i \in V'} U(S_M, S_R^{TP}, i)$. As any location $i \in V'_M$ keeps its place in the master route in the new schedule S_R^{TP} it holds that location i does not deviate in S_R^{TP} and therefore for all $i \in V'_M$, $U(S_M, S_R^{TP}, i) = 0$. Observe that as $V' = V'_M \bigcup V'_R$ we have $\sum_{i \in V'} U(S_M, S_R^{TP}, i) = \sum_{i \in V'_M} U(S_M, S_R^{TP}, i) + \sum_{i \in V'_R} U(S_M, S_R^{TP}, i) \leq \sum_{i \in V'_R} u_i$. This provides us with an upper bound on the costs of using schedule S_R^{TP}.

Next it will be proven that $\sum_{i \in V'_R} u_i \leq U^*$. Denote by S_{U^*} the schedule for which U^* is attained and let V_{U^*} be the set of locations that deviate in this schedule. It will be shown that $V'_R \subseteq V_{U^*}$. Consider a location $j \in V'_R$ and the route $r = \{i_1, ..., i_l, j, ..., i_k\} \in S_{U^*}$ and denote $r_M = \{i_1, ..., i_l\}$. If j does not deviate in S_{U^*}, none of the locations in r'_M deviate and it must hold that $\sum_{i \in r'_M} q_i + q_j \leq Q$. However, this contradicts the construction of V'_R, hence we can conclude that $\sum_{i \in V'_R} u_i \leq U^*$. This yields the desired result. □

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 costs of traveling. In such a case in particular the two-phase heuristic produces the optimal schedule for the VRSP.

Lemma 4. *Let $u_{\min} = \min_{i \in V'} u_i$, $c_{\min} = \min_{(i,j) \in E} c_{ij}$ and let S_R^{TP} be the schedule produced by the two-phase heuristic. If $u_{min} \geq \sum_{i \in V'} [c_{0i} + c_{i0}] - (n + \lceil \frac{\sum_{i \in V'} q_i}{Q} \rceil) c_{\min}$ then the schedule S_R^{TP} is optimal.*

Proof. For any schedule S_R that is a solution to the VRSP, let the costs be decomposed into

two parts, the traveling costs $Z_R = \sum_{(i,j)\in S_R} c_{ij}$ and the deviation costs $U_R = \sum_{i\in V'} U(S_M, S_R, i)$ respectively. Note that the costs of using the schedule S_R^{TP} can also be decomposed into Z_R^{TP} and U^*.

Observe that $Z_R^{TP} \leq \sum_{i\in V'}[c_{0i}+c_{i0}]$ and, as $\lceil \frac{\sum_{i\in V'} q_i}{Q} \rceil$ is a lower bound on the number of vehicles that are needed, $Z_R \geq (n + \lceil \frac{\sum_{i\in V'} q_i}{Q} \rceil)c_{\min}$ where $c_{\min} = \min_{(i,j)\in E} c_{ij}$. Therefore:

$$\sum_{i\in V'}[c_{0i} + c_{i0}] - (n + \lceil \frac{\sum_{i\in V'} q_i}{Q} \rceil)c_{\min} \geq Z_R^{TP} - Z_R \tag{28}$$

Note that for any schedule S_R with at least one more deviation than S_R^{TP} it holds that $U_R \geq U^* + u_{\min}$. From this it follows for $u_{min} \geq \sum_{i\in V'}[c_{0i} + c_{i0}] - (n + \lceil \frac{\sum_{i\in V'} q_i}{Q} \rceil)c_{\min}$, that:

$$U_R \geq U^* + Z_R^{TP} - Z_R \tag{29}$$

Therefore for all S_R with at least one more deviation than S_R^{TP} it holds that:

$$Z_R^{TP} + U^* \leq Z_R + U_R \tag{30}$$

Finally, note that for any schedule with an equal number of deviations as S_R^{TP}, (30) is always satisfied. This proves the optimality of S_R^{TP} for all $u_{min} \geq \sum_{i\in V'}[c_{0i} + c_{i0}] - (n + \lceil \frac{\sum_{i\in V'} q_i}{Q} \rceil)c_{\min}$.

□

Tight Example:

To show that the bound on $u_{\min}$ cannot be improved, consider the following tight example. The example is an instance with two locations, all distances are symmetric and equal, $c_{01} = c_{02} = c_{12} = c$. The master schedule consists of a return trip to location one and a separate return trip to location two. The demand realizations are such that $q_1 + q_2 \leq Q$ and the two routes might feasibly be merged. Furthermore, let $u_2 = u_{\min}$. Three feasible schedules exist. The master schedule can be used as a solution to the rescheduling problem, this schedule will be referred to as S_1 and yields total costs $4c$. A second option is to have only one route, first visiting location 1 and then location 2, this schedule will be referred to as S_2 and yields the total costs of $3c + u_{min}$. The last option is the reverse of S_2 and can be disregarded in this discussion as the costs of using this schedule will never be lower than the costs of using S_2. The two-phase heuristic will produce schedule S_1. This schedule will be optimal if and only if $u_{min} \geq c = \sum_{i\in V'}[c_{0i} + c_{i0}] - (n + \lceil \frac{\sum_{i\in V'} q_i}{Q} \rceil)c_{\min}$.

For specific problem instances the relative difference between the traveling costs and the deviation costs need not be high for the two-phase heuristic to produce the optimal solution. However, it is not possible to provide a general guarantee for small relative differences. For the problem instances used for the computational results, numeric values are presented to give insight in the necessary magnitude of the difference in costs for the two-phase heuristic to produce optimal schedules.

The difference of the costs when applying the schedule produced by the two-phased heuristic as opposed to using the optimal solution in the worst case scenario, can offer insight in the performance of the heuristic. In the next lemma an upperbound is given on the relative difference between the costs of the optimal solution and those of using the two-phase heuristic.

Lemma 5. *The costs of using the routing schedule produced by the two-phased heuristic* S_R^{TP} *is at most* $\min\{\frac{Q}{q_{\min}}, \frac{2Qc_{\max}}{(Q+\bar{q})c_{\min}}+1\}$ *times the costs of the optimal schedule* S_R^* *for the VRSP, where* $q_{\min} = \min_{j\in V'} q_j$, $\bar{q} = \frac{\sum_{i\in V'} q_i}{n}$, $c_{\max} = \max_{(i,j)\in E} c_{ij}$ *and* $c_{\min} = \min_{(i,j)\in E} c_{ij}$.

Proof. Note that for ease of notation we use symmetric traveling costs c_{ij} in this analysis. An analogue proof can be given when traveling costs are not symmetric.

Let the traveling costs and deviation costs of S_R^{TP} be given by Z_R^{TP} and U^* respectively. Similarly, let the traveling and deviation costs of S_R^* be given by Z_R^* and U_R^*. Furthermore let $Z^* = \min_{S\in\$} \sum_{(i,j)\in S} c_{ij}$. To prove the theorem, it is first shown that $\frac{Z_R^{TP}+U^*}{Z_R^*+U_R^*} \leq \frac{Q}{q_{\min}}$ and secondly that $\frac{Z_R^{TP}+U^*}{Z_R^*+U_R^*} \leq \frac{2Qc_{\max}}{(Q+\bar{q})c_{\min}} + 1$.

In Simchi-Levi et al. (1997) it is proven that for a CVRP with unequal demands it holds that $2\sum_{j\in V'} c_{0j}q_j \leq QZ^*$. Now observe that:

$$Z^{TP} \leq 2\sum_{i\in V'} c_{0i} \leq \frac{2}{q_{\min}} \sum_{i\in V'} c_{0i}q_i \leq \frac{Q}{q_{\min}} Z^* \tag{31}$$

Which implies that:

$$\frac{Z_R^{TP}+U^*}{Z_R^*+U_R^*} \leq \frac{\frac{Q}{q_{\min}}(Z^*+U^*)}{Z^*+U^*} \leq \frac{Q}{q_{\min}} \tag{32}$$

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

$$\frac{Z_R^{TP}+U^*}{Z_R^*+U_R^*} \leq \frac{Z_R^{TP}+U^*}{Z^*+U^*} < \frac{2nc_{\max}}{(n+\lceil \frac{\sum_{i\in V'} q_i}{Q} \rceil)c_{\min}} + 1 \leq \frac{2nc_{\max}}{(n+\frac{\sum_{i\in V'} q_i}{Q})c_{\min}} + 1 = \frac{2Qc_{\max}}{(Q+\bar{q})c_{\min}} + 1 \tag{33}$$

Here the strict inequality follows from $\frac{a+b}{c+b} < \frac{a}{c} + 1$ for $a, b, c > 0$. This concludes the

proof.

□

Tight Example:

Note that this bound can not be improved upon, by considering the following tight example. 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 $\frac{Z_R^{TP}+U^*}{Z_R^*+U_R^*} = 1 = \frac{Q}{q_{\min}} = \min\{\frac{Q}{q_{\min}}, \frac{2Qc_{\max}}{(Q+\bar{q})c_{\min}} + 1\}$.

3.2 Modified savings heuristic

The savings algorithm due to Clarke and Wright (1964) is a very insightful algorithm that has been widely used to solve the CVRP problem. This algorithm has great flexibility in incorporating features like time window constraints which play a crucial role in many practical applications. Moreover this algorithm can easily be adapted to incorporate the features of the VRSP. In the modified savings algorithm, the schedule is created in a constructive manner as opposed to using S_M as a basis to remove edges from like in the two-phase heuristic. This allows for greater freedom and flexibility in designing a schedule.

The savings algorithm starts by assigning a vehicle to serve every individual customer. The initial schedule consists of n routes from the depot to a customer and back. The costs of this schedule are then of course $2\sum_{j=1}^{n} c_{0j}$. Routes will be merged to improve these costs until no further improvement can be found using merges or until any further merges will render the schedule infeasible (for instance because capacity constraints are violated). If the last location on a route is i and the first on another route is j, a merge between these routes is obtained by traveling directly from i to j instead of returning and starting from the depot. The savings of such a merge is $s_{ij} = c_{i0} + c_{0j} - c_{ij}$.

Let S^m be the schedule after m merges. Furthermore let S_{ij}^m be the schedule resulting from merging the routes ending at location i and starting at location j when the current schedule is S^m. For the modified savings algorithm the rescheduling costs need to be incorporated. These should be calculated and added to s_{ij}. This has to be done for the initial values as well as after every merge. Define the rescheduling savings in iteration m as:

$$u_{ij}^m = \sum_{i \in V'} U(S_M, S_{ij}^m, i) - \sum_{i \in V'} U(S_M, S^m, i) \tag{34}$$

Denote A as the set of all pairs (i, j) such that the route ending with customer i can feasibly be merged with customer j. It must hold for all pairs (i, j) in A that: i and j are on different routes; Customers i and j are the last and first location on a route respectively;

A merge between the route ending with i and the route starting with j must not violate the capacity constraint. The algorithm is given by:

Step 1: Construct S^0 in which every customer is visited by an individual vehicle. Initialize $A = E'$.

Step 2: For all $(i,j) \in E'$ calculate s_{ij} and u_{ij}^m, let the total savings be $S_{ij} = s_{ij} + u_{ij}^m$ and let m=0.

While $\max_{(i,j)\in A} S_{ij} \geq 0$

Step 3: Merge the route ending with i and the route starting with j for $(i,j) = \arg\max_{(i,j)} S_{ij}$ resulting in schedule S^{m+1}, set $m = m + 1$.

Step 4: Update A.

Step 5: Calculate u_{ij}^m for all $(i,j) \in A$.

Step 6: Calculate $S_{ij} = s_{ij} + u_{ij}^m$ for all $(i,j) \in A$.

Observe that the algorithm finds a feasible solution in polynomial time.

4 Computational results

In this section, the results of numerical experiments are presented. They will provide insight into the sensitivity of the model to different values of the deviation costs. Furthermore, the performance of the proposed heuristics is evaluated empirically, by applying them to a range of test cases and comparing the results.

In this section, randomly generated test cases are used. For every problem instance, first an master schedule S_M is generated based on some demand realization. Next a new demand realization is generated and rescheduling will yield a new feasible schedule. Unless stated otherwise, the following settings are used for the generation of individual problem instances:

- n customer locations are randomly generated, uniformly distributed over a square geographical area with sides of length 20 units and a depot in the center.
- The costs of traveling from one location to another is equal to the Euclidean distance between both.
- All vehicles have a capacity of 30 units.
- Demand for the master problem is normally distributed with mean 5 and standard deviation 1.5, truncated from below to 1 and from above to 30.
- Demand for the new schedule is normally distributed with standard deviation 1.5. The demand average is equal to 1.5 times the realization for the master schedule.

The deviation costs will be specified for every individual experiment.

These numbers are inspired by a practical case in a retail chain with recurrent sales actions. The specific demand structure ensures the necessity to deviate from the original schedule. The algorithms are tested on cases where deviation will indeed be necessary as demand has typically increased. The original schedule is generated using a direct implementation of the CVRP using a two commodity flow formulation (Baldacci et al. 2004) yielding an optimal schedule.

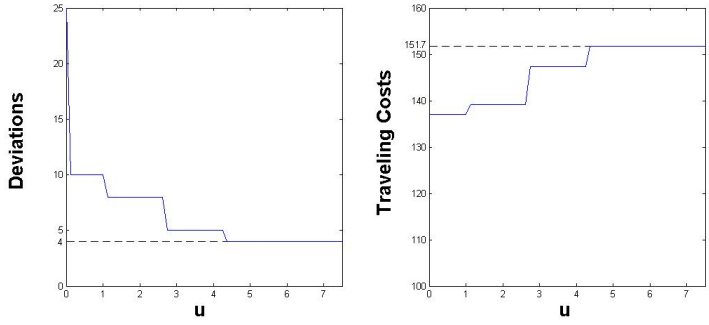
All tests are performed on an Intel(R) Core(TM)2 Duo CPU E8400, 2.99 GHz, 3.21 GB of RAM. For all solutions that are found by direct implementation of a mixed integer programming formulation, ILOG CPLEX 10.1 is used.

4.1 Model parameter sensitivity

For large deviation costs $u_{\min}$ with respect to the traveling costs, by lemma 4, the optimal schedule is a schedule with the minimum number of deviating locations. It is very interesting to see what the value of $u_{\min}$ must at least be in situations that are not necessarily worst case, to ensure the optimal schedule to have the minimal number of deviations. Not only does this give insight into the cost structure of deviating, also observe that from this value of $u_{\min}$ and higher, the two-phase heuristic is guaranteed to produce the optimal solution. For these experiments let all $u_i = u$.

As an example, consider a single randomly generated case of 25 customer locations, $n = 25$. This case exhibits behavior that is representative for that of all simulated cases. Figure 2 depicts properties of optimal schedules of the VRSP for this case using different values u for the deviation costs. These optimal schedules are found using a direct implementation of the two commodity flow formulation of the VRSP. The graph on the left shows the number of deviations and the graph on the right shows the total traveling costs of the optimal schedule. Notice that when $u = 0$, all locations deviate, however as u grows slightly above 0, a new schedule is found with less deviations but with equal traveling costs.

Figure 2:



The figure indeed supports the notion that large u forces the optimal schedule to have a decreased number of deviations. The figure shows that in this particular case the minimum number of deviations is 4 and is obtained for u greater or equal to 4.25. Obviously this value of 4.25 is meaningless unless related to the traveling costs.

To give a more general impression, let us relate the critical level u, the level from which onward the optimal solution of the VRSP is the schedule with the minimum number of deviations, with the average traveling costs of all edges used in the original schedule $\bar{c}_M$. For the example depicted in Figure 2 $\bar{c}_M = 4.13$, which can be considered very close to the critical level u of 4.25. For the 100 generated cases 40% of the critical u values lie below $0.5\bar{c}_M$, 71% of the critical values even lie below $\bar{c}_M$ and 87% lies below $1.5\bar{c}_M$. Observe that 47% of the critical levels do not differ more than 50% of the value of $\bar{c}_M$.

The schedules with minimal traveling costs in the example case, have traveling costs 137.1. Among these schedules the best in terms of number of deviations, is a solution with 10 deviations. The schedule with minimal number of deviations, 4, has traveling costs 151.7. Comparing these schedules, observe that in this case, 6 less locations need to deviate at the expense of an increase in traveling costs of 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 deviations is 9.7%, with standard deviation 5.3 percent points. It might therefore be argued that when the benefits of having a schedule with minimal number of deviations outweighs the costs of traveling close to an additional 9.7% units distance, it may be a good idea to minimize the number of deviations. In this case the VRSP can be solved using the two-phase heuristic.

4.2 Algorithm performance

To evaluate the performance of the two-phase heuristic and the modified savings heuristic, they are used to find solutions for multiple test cases. For these cases, the deviation costs are randomly generated using a normal distribution with mean equal to 0.75 times the average length of the edges used in the original schedule, the standard deviation is equal to 0.5 times this average length. The deviation costs 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 that the two-phase heuristic is guaranteed optimal.

In implementing the two-phase heuristic, the second phase is performed using a two commodity flow formulation of the CVRP. This is solved to optimality by a direct implementation of the model in. Obviously, the running time of the two-phase heuristic can be improved by

using sophisticated heuristics in this phase.

To obtain some insight in the quality of the solutions obtained by using the heuristics, they are used on small test cases. In these small cases, the optimal solutions can be found using a direct implementation of the VRSP in. The optimal solutions are then used as a benchmark.

Table 1 shows the results of computational experiments for instances of different size. For each value n representing the number of customer locations generated in a test case, 100 simulations were performed. For each individual generated instance of the VRSP, an optimal solution was found as well as solutions using the two-phase heuristic and the modified savings, respectively. The column CD_{TP} shows the average cost difference in percentages of the schedule produced with the two-phase heuristic with respect to the optimal schedule over all 100 instances, the standard deviation of the difference is included in between brackets. Likewise, CD_{MS} shows the average cost difference and standard deviation for the schedules produced with the modified savings heuristic. Note that these costs include both the traveling costs and the deviation costs. Next, the column $C_{TP<MS}$ shows the number of instances out of the 100, in which the costs of two-phase heuristic provided a schedule with lower costs than the modified savings heuristic and the column $C_{TP>MS}$ shows it the other way round. Finally, T_{opt}, T_{TP} and T_{MS} shows the average time in seconds it took the algorithms to finish all 100 cases. The standard deviation of the time to find the optimal solution varies between 1 and 3 times the average, for the two phased heuristic it is between 1 and 2 times the average and for the modified savings algorithm the standard deviation is 0. Note that for every individual case, solving it to optimality took the most amount of time and using the modified savings algorithm took the least amount of time.

Table 1:

n	CD_{TP}	CD_{MS}	$C_{TP<MS}$	$C_{TP>MS}$	T_{opt}	T_{TP}	T_{MS}
10	0.7(2.1)	16.2(16.0)	75	5	0.034	0.002	0.000
15	1.1(2.2)	15.9(12.3)	87	7	0.455	0.009	0.000
20	1.1(1.7)	17.1(11.2)	91	5	1.580	0.025	0.000
25	1.8(2.4)	15.8(8.5)	98	2	9.327	0.134	0.000
30	1.4(1.9)	16.9(8.4)	98	1	95.987	0.283	0.000

Although the difference in costs between the two-phase heuristic and the optimum seems to be increasing, the costs are on average not more than 2% above the optimum for instances of every considered size. This differs significantly from the performance bound that has been derived earlier. Furthermore, the costs of the modified savings algorithm are on average more than 15% above the optimum for instances of every considered size. Observe from columns $C_{TP<MS}$ and $C_{TP>MS}$, that the two-phase heuristic more often produces schedules with lower costs than the schedules of the modified savings algorithm than the other way around.

Also, as the size of the instances increases, the number of times in which the two-phase heuristic performs better increases as well. We see that for these choices in generating u, the two-phase heuristic performs better than the modified savings heuristic in terms of costs.

From the columns indicating running times it can obviously be concluded that using heuristic methods reduces the solution time significantly with respect to solving it to optimality using a direct implementation of the VRSP. Furthermore, the running time of the two-phase heuristic significantly increases as the size of the instances grows. This was to be expected as in the second phase an NP-hard problem of increasing size is solved to optimality. The modified savings algorithm always terminated within a millisecond.

Solving the VRSP to optimality using a direct implementation has not been done for problem instances including more than 30 customer locations. The computation time increases exponentially as is also supported by table 1. As practical problems most often include more than 30 customer locations, heuristics are typically used to find feasible solutions. It is therefore interesting to see how the results of using the two proposed heuristics compare on larger instances. First of all it is of interest whether or not the relative difference in costs of the generated schedules gets smaller for larger instances. Similarly it is of interest to see if the number of times that the two-phase heuristic outperforms the modified savings heuristic does not decrease.

For the same reason the VRSP is not solved to optimality, the master schedule will not be generated by solving a CVRP instance to optimality. The master schedule will be generated using the standard savings algorithm. Note that 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 some unnecessary deviations. In such cases, the two-phase heuristic will not perform well as it never generates unnecessary deviations. Therefore, for the test cases using the saving algorithm to generate the original schedule, also test cases of 20 and 30 locations are generated to compare to the test cases where the master is generated by solving a CVRP to optimality. Table 2 shows the results of the test cases where the savings algorithm is used. Columns TC_{TP} and TC_{MS} show the average costs and standard deviation over the same 100 instances for the two-phase heuristic and the modified savings algorithm. The interpretation of the last four columns are analogue to that of table 1.

For the test cases where the master schedule was generated by solving a CVRP to optimality, the modified savings algorithm constructed schedules with costs that were on average less than 15% above the costs of the schedules produced by the two-phase heuristic. When the master schedule is generated by applying the standard savings algorithm, this difference

Table 2:

n	TC_{TP}	TC_{MS}	$C_{TP<MS}$	$C_{TP>MS}$	T_{TP}	T_{MS}
20	128.9(17.0)	154.0(17.2)	97	3	0.032	0.000
30	175.6(18.3)	209.5(18.4)	100	0	0.324	0.000
40	220.8(18.0)	263.4(15.7)	100	0	1.652	0.000
50	263.1(20.3)	309.7(19.5)	100	0	14.042	0.000

is on average roughly 19%. However there is no clear indication that this difference varies as the size of the test cases vary. For larger test cases the two-phase heuristic outperforms the modified savings heuristic for every individual case. Therefore we again conclude that for this configuration in generating u, the two-phase heuristic performs better than the modified savings heuristic in terms of costs.

With respect to the running times, again data supports that the running time of the two-phase heuristic increases exponentially as the number of customer locations increase. The CVRP that has to be solved in the second phase on average includes more locations. Also for the larger problem instances the modified savings algorithm always terminated within a millisecond.

5 Concluding remarks

In this paper, the negative effects of deviating from a master schedule have been incorporated in the scheduling process. Insight has been obtained on the behavior of the optimal solution of the VRSP for different values of the deviation costs relative to the traveling costs.

Two mixed integer linear formulations have been presented. Using these formulations the VRSP can be solved to optimality using direct implementation. Computational results have shown that this can consistently be done for problem instances of up to 30 locations.

Two heuristics have been presented to solve instances of the VRSP including up to 50 customer locations, that find good solutions within a small amount of computation time. The two-phase heuristic makes use of specific properties of the VRSP to find a solution. For sufficiently high deviation costs, this algorithm generates a schedule that is on average close to optimal or even guaranteed optimal, for all considered problem instances. For general problem instances, an analytical bound on the difference between the solution generated by the two phase heuristic and the optimum is presented. However, numerical results indicate that this analytical upper bound is extremely far from the actual difference. As in the second phase of the 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 minimum number of locations that need to deviate. For larger problem instances, the second phase can be performed using a heuristic to solve the CVRP.

The second heuristic that is presented is the modified savings heuristic. Although computation times are extremely low, the performance in terms of total costs is worse than that of the two-phase heuristic. The modified savings algorithm can easily be extended to incorporate additional features like time-window constraints that play a great role in practical applications.

References

- Baldacci, R., Hadjiconstantinou, E. and Mingozzi, A. 2004, 'An Exact Algorithm for the Capacitated Vehicle Routing Problem Based on a Two-Commodity Network Flow Formulation', *Operations Research*, Vol. 52, No. 5, pp. 723-738.
- Bertsimas, D.J. 1992, 'A vehicle routing problem with stochastic demand', *Operations Research*, Vol. 40, No. 3, pp. 574-585.
- Bertsimas, D.J. and Levi, D.S. 1996, 'A New Generation of Vehicle Routing Research: Robust Algorithms, Addressing Uncertainty', *Operations Research*, Vol. 44, No. 2, pp. 286-304.
- Clarke, G. and Wright, J.W. 1964, 'Scheduling of Vehicles from a Central Depot to a Number of Delivery Points', *Operations Research*, Vol. 12, No. 4, pp. 568-581.
- Fisher, M. 1995, 'Vehicle Routing', *the handbooks in operations research*, ed. J. Desrosiers, Y. Dumas, M.M. Solomon, F. Soumis, W.B. Powell, P. Jaillet, A. Odoni, A. Federgruen, D. Simchi-Levi, A.A. Assad, B.L. Golden, M. Florian, D. Hearn, M. Labbé, D. Peeters, J.F. Thisse, R.H. Möhring, D. Wagner, F. Wagner, W.W. Sharkey, Elsevier Science B.V..
- Gendreau, M., Laporte, G. and Séguin, R. 1996, 'A tabu search heuristic for the vehicle routing problem with stochastic demands and customers', *Operations Research*, Vol. 44, No. 3, pp. 469-477.
- Haughton, M.A. 1998, 'The performance of route modification and demand stabilization strategies in stochastic vehicle routing', *Transportation Research B*, Vol. 32, No. 8, pp. 551-566.
- Laporte, G. 1986, 'Generalized Subtour Elimination Constraints and Connectivity Constraints', *The Journal of the Operational Research Society*, Vol. 37, No. 5, pp. 509-514.
- Laporte, G. 1992, 'The Vehicle Routing Problem: An overview of exact and approximate algorithms', *European Journal of Operational Research*, 59, pp. 345-358.
- Laporte, G. 2007, 'What you should know about the vehicle routing problem', *Naval research Logistics*, Vol. 54, No. 8, pp. 811-819.

- Laporte, G., Gendreau, M., Potvin, J. and Semet, F. 2000, 'Classical and modern heuristics for the vehicle routing problem', *International Transactions in Operational Research*, 7, pp. 285-300.
- Laporte, G., Louveaux, F.V. and van Hamme, L. 2002, 'An integer L-Shaped algorithm for the capacitated vehicle routing problem with stochastic demands', *Operations Research*, Vol. 50, No. 3, pp. 415-423.
- Laporte, G., Nobert, Y. and Desrochers, M. 1985, 'Optimal Routing under Capacity and Distance Restrictions', *Operations Research*, Vol. 33, No. 5, pp. 1050-1073.
- Li, J., Mirchandani, P.B. and Borenstein, D. 2007, 'The Vehicle Rescheduling Problem: Model and Algorithms', *Networks*, Vol. 50, No. 3, pp. 211-229.
- Li, J., Mirchandani, P.B. and Borenstein, D. 2009, 'A Lagrangian heuristic for the real-time vehicle rescheduling problem', *Transportation Research Part E*, 45, pp. 419-433.
- Secomandi, N. 2001, 'A Rollout Policy for the Vehicle Routing Problem with Stochastic Demands', *Operations Research*, Vol. 49, No. 5, pp. 796-802.
- Simchi-Levi, D., Chen, X., Bramel, J. 1995, *The Logic of Logistics. Theory, Algorithms, and Applications for Logistics and Supply Chain Management*, Springer Series in Operations Research, second edition,Springer Science+Business Media, Inc., New York.
- Tan, K.C., Cheong, C.Y. and Goh, C.K. 2007, 'Solving multiobjective vehicle routing problem with stochastic demand via evolutionary computation', *European Journal of Operational Research*, 177, pp. 813-839.
- Toth, P. and Vigo, D. 2002, 'Models, relaxations and exact approaches for the capacitated vehicle routing problem', *Discrete Applied Mathematics*, 123, pp. 487-512.