Distribution of goods and services has been important for ages. We still use the phrase “all roads lead to Rome” referring to the Roman construction of roads across the empire to efficiently distribute its armies and provisions. Until the 18th century, a general merchant was the key distributor, and most households were manufacturing their own clothes, candles, food, etc. During the industrial revolution, the production and transport of goods and services changed rapidly due to developments in manufacturing, transport and communication systems. Today, almost everything can be bought at shops or online and many services are outsourced, as for example repair, installation, and cleaning services. This leads to major distributional challenges of when to serve customers, how much to replenish, which route to take, etc.

Furthermore, we live in a fast-paced society in which time is valuable and everything should be arranged as efficient as possible. We want to indicate when we are available for service or delivery and we prefer narrow time windows to limit the waiting time and to better plan our daily activities. However, these high service level standards can increase the distribution cost. This leads to challenging routing and scheduling problems that on the one hand want to minimize the transportation cost and on the other hand need to keep the service levels high.

In this thesis, efficient routing solutions are developed that take preferences on service times of both the customer and the logistics service provider into account. We consider three of these routing problems in which service time windows play an important role. First, we investigate routing problems in which customers can indicate multiple time windows in which they are available for service. The logistics service provider has the freedom of selecting one of those time windows to service a customer. A heuristic and exact solution method are developed for this problem in Chapter 2 and 3, respectively. Secondly, a routing problem is solved in which the arrival times at the customers should be unpredictable. This problem is inspired by security considerations in a real-life application in cash transport. By using the theoretical insights from the first two chapters, a novel way for diversifying arrival times at customers is developed, and an
efficient solution approach is designed. Thirdly, travel and service times are uncertain in practice, which highlights the need to find a robust way to assign service time windows to customers such that the probability of being on time is as high as possible. Hence, a robust method is developed to simultaneously assign time windows to customers and determine a routing solution taking uncertainty in travel and service times into account.

In practice, a wide variety of routing problems can be observed, involving different objective functions and side constraints. The vehicle routing problem (VRP) is a classical optimization problem from the Operations Research literature that underpins the most fundamental distribution setting occurring in practice. As a result, it is the foundation on which a myriad of more applied distribution models and solvers have been designed. As the VRP is one of the most intensively studied optimization problems in the literature, it offers a competitive research field to develop and test new distribution models and solution approaches. All the chapters in this thesis consider variants of the VRP in which service time windows play an important role. Furthermore, we examine routing problems with deterministic and stochastic travel times and propose heuristic, exact, and robust solution methods.

This chapter offers a brief introduction of the VRP with time windows and its solution methods. Additionally, applications in the distribution of cash are discussed since this has been the motivation for a part of this research. Finally, the main research challenges are discussed, and an overview of the thesis is provided.

1.1 Vehicle routing problem with time windows

The total amount of inland freight transport in the EU was over 2,400 billion ton-kilometer of goods in 2016. Around three-quarters of this was transported by road (Eurostat [2018]). In the Netherlands, 70,000 million ton-kilometer of goods are transported by road freight transport and, furthermore, 350 million packages were sent by companies and consumers (Autoriteit Consument en Markt (ACM) [2017]). Since the online consumer market is still growing, the transportation cost and the number of kilometers driven will only increase. These numbers underline the importance of solving routing and scheduling problems.

Because of the practical relevance and academic challenges, the vehicle routing problem (VRP) is one of the most studied combinatorial optimization problems. In the VRP, a routing solution has to be determined that consist of at most $|K|$ routes that start and end at the depot and every customer is serviced exactly once by a single vehicle. The total demand serviced by a vehicle cannot exceed the vehicle capacity, $Q$. The goal of the VRP is to determine a set of routes with minimum cost. For encompassing reviews on the VRP see for example Golden et al. [2008] or Toth and Vigo [2014]. In this thesis, we focus on variants of the VRP with time windows (VRPTW) in which the service at every customer should start in a specific time window. The VRPTW is one of the most studied variants of the VRP as time windows arise naturally in many practical problems, for example, parcel deliveries, security patrol service, bank
1.1. Vehicle routing problem with time windows

deliveries, grocery delivery, and school bus routing. The time windows can be hard of soft. In the case of hard time windows, the vehicle has to wait until the start of the time window when the vehicle arrives too early at a customer. Generally arriving too early incurs no cost but arriving too late is not allowed. In the case of soft time windows, arriving too early or too late is both allowed at the expense of a penalty. In this thesis, we deal with both type of time windows, but in this chapter, we focus on the general VRPTW with hard time windows. Extensive reviews on the VRPTW can be found in Bräysy and Gendreau [2005b] and Baldacci et al. [2012].

The VRPTW is defined on a complete directed graph \( G = (V,A) \) with nodes \( V = \{0,1,\ldots,n,n+1\} \) and arcs \( A = \{(i,j) \in V \times V : i \neq j\} \). Node 0 and \( n + 1 \) represent the depot and nodes \( V' = \{1,\ldots,n\} \) correspond to the set of customers. A fleet of identical vehicles, is denotes by \( K \), with capacity \( Q \) is available to serve the customers. Each arc \((i,j) \in A\) is associated with a non-negative cost \( c_{ij} \) and travel time \( \tau_{ij} \). The service at every customer \( i \) must start within a given time window \([e_i,l_i]\). If the vehicle arrives before \( e_i \) it has to wait until the start of the time window. Arrivals after \( l_i \) are prohibited. The start time of servicing customer \( i \) is denoted by \( t_i \). Each node \( i \in V \) is associated with a demand \( q_i \) and a service time \( s_i \), where \( q_i \) and \( s_i \) are non-negative numbers. We assume \( s_0 = s_{n+1} = 0 \) and \( q_0 = q_{n+1} = 0 \).

Several mathematical formulations are proposed for the VRPTW. We present a mixed integer linear program (MILP) based on the binary flow variables \( x_{ij}^k \), which is one if vehicle \( k \) visits node \( j \) right after node \( i \) and zero otherwise. Using this decision variable, the VRPTW can be formulated by

\[
\begin{align*}
\min & \sum_{k \in K} \sum_{(i,j) \in A} c_{ij} x_{ij}^k \\quad & (1.1) \\
\text{s.t.} & \sum_{k \in K} \sum_{j \in V \setminus \{i\}} x_{ij}^k = 1, & \forall i \in V', \quad & (1.2) \\
& \sum_{j \in V \setminus \{0\}} x_{0j}^k = 1, & \forall k \in K, \quad & (1.3) \\
& \sum_{j \in V \setminus \{n+1\}} x_{j,n+1}^k = 1, & \forall k \in K, \quad & (1.4) \\
& \sum_{j \in V \setminus \{i\}} x_{ij}^k - \sum_{j \in V \setminus \{i\}} x_{ji}^k = 0, & \forall i \in V, k \in K, \quad & (1.5) \\
& \sum_{(i,j) \in A} q_{ij} x_{ij}^k \leq Q, & \forall k \in K, \quad & (1.6) \\
& t_i + s_i + \tau_{ij} \leq t_j + M(1 - \sum_{k \in K} x_{ij}^k), & \forall (i,j) \in A, i,j \in V', \quad & (1.7) \\
& e_i \leq t_i \leq l_i, & \forall i \in V', \quad & (1.8) \\
& x_{ij}^k \in \{0,1\}, & \forall (i,j) \in A, k \in K. \quad & (1.9)
\end{align*}
\]
Introduction

The objective function (1.1) is to minimize the total routing cost. Constraints (1.2) ensure that each customer is visited exactly once by a single vehicle. Constraints (1.3) and (1.4) ensure that every route starts and ends at the depot, respectively. The flow conservation constraints of the customers are given in (1.5). Constraints (1.6) ensure that the total demand of a vehicle route is not greater than the vehicle capacity. Constraints (1.7) prevent subtours and make the start times of servicing the customers consistent. Let \( M \) be a sufficiently large positive number, more precisely, we set \( M = \max_{(i,j) \in A} \{ l_i + s_i + \tau_{ij} - e_j \} \). Constraints (1.8) ensure that the service at a customer starts within the time window. Lastly, Constrains (1.9) impose binary conditions on the decision variables.

The VRPTW is an NP-complete problem since it is a generalization of the TSP (Karp [1972]). In the past 50 years, many solution methods have been developed for the VRPTW. A brief description of these solution methods is presented in the next section.

1.2 Solution methods

The MILP formulated in the previous section can be solved using exact or heuristic solution methods. Exact algorithms are capable of finding optimal solutions, but are typically computational intensive. Heuristic algorithms are faster and can solve large instances, however, no guarantee of the solution quality can be given. Nowadays many state-of-the-art solution methods are based on a combination of heuristic and exact components. To support the main research challenges discussed in Section 1.3, we will briefly describe the exact and heuristic solution methods developed for the VRPTW.

1.2.1 Heuristic algorithms

Heuristic solution methods often find good-quality solution in relative short time. Most practical problems are solved by heuristic methods since they are flexible, fast, and capable of solving large instances. There are many different types of heuristic solution methods, for an overview see for example Bräysy and Gendreau [2005b] and Toth and Vigo [2014]. Almost all metaheuristic use the concept of local search. In a local search method, the current solution is iteratively improved by exploring neighboring solutions. If a neighboring solution improves the objective value it replaces the current solution; if no improving solution exists a local minimum has been found. A neighborhood is defined by a subset of solutions that can be derived by applying a simple modification to the current solution. An example of a simple neighborhood is the exchange neighborhood in which two customers exchange position. A large variety of neighborhoods are available for the VRPTW, an extensive review and accessible explanation of many neighborhoods can be found in Bräysy and Gendreau [2005a]. In most heuristic methods, multiple neighborhoods are used. There are two basic types of neighborhoods: intra-route and inter-route neighborhoods. In the first type, only a single route is changed compared to the current solution, while in the second type one or several customers are moved between two or more routes. In general, using more or
larger neighborhoods leads to a larger solution space and therefore holds the potential of obtaining a better solution. However, the evaluation of a larger solution space will be time consuming. Therefore, speed-up techniques can be used such as sequential search [Irnich 2008] or granular neighborhoods [Schneider et al. 2017, Toth and Vigo 2003]. In the first approach, many possible neighboring solutions do not have to be explored, since their objective value will not improve the current solution. In the second approach, the arc set is restricted such that the neighboring solutions only contains arcs that are likely to be part of high quality solutions.

When only neighboring solutions are accepted that improve the objective function, a local optimum will be found that can be very far from the optimal solution. To diversify the search and escape from a local minimum many different model concepts and acceptance criteria can be used, for example the concept of multi start, shaking, tabu search, or simulated annealing. In general, there are two widely used acceptance strategies; “first-accept” in which the first neighboring solution that satisfies the acceptance criteria is selected or “best-accept” in which all neighbors are explored, and the best neighboring solution is selected.

To create high-quality solutions, most state-of-the-art metaheuristics explore infeasible solutions during the search. Allowing infeasible solution makes it easier to reach different areas in the solution space. However, it increases the complexity since the solution space increases and the objective should include some terms that penalize infeasible solutions.

Traditionally, a hierarchical objective function is used in which first the number of vehicles is minimized and secondly, the travel time is minimized. In recent heuristic algorithms also a combination of the two components are used and various objectives are taken into account, as minimizing the total duration and the overtime (Goel and Kok 2012, Belhaiza et al. 2014, Taş et al. 2013). To minimize the number of vehicles, different approaches are applied. In many heuristic solution methods first the minimum number of vehicles is determined and then the travel cost is minimized. The most effective route minimization heuristic in the literature is proposed by Nagata and Braysy 2009.

1.2.2 Exact algorithms

Numerous exact solution methods have been proposed for the VRPTW. In contrast to heuristic solution methods, the objective consists often of one component, e.g., minimizing the total routing cost or the total travel time.

Branch-and-bound frameworks have been used extensively to solve the VRP exactly. In a branch-and-bound framework, the state-space is searched systematically. At the root node, the linear relaxation is solved (full relaxed solution space) and the algorithm explores branches of the tree, which represent subsets of the solution set. Usually, solving the linear relaxation of an integer program provides a fractional lower bound on the value of the integer problem. For traditional VRP integer programs, branching is usually performed by setting a fractional variable to 0 for one branch and to 1 for
the other branch. Before enumerating the candidate solutions of a branch, the branch is checked against upper and lower estimated bounds on the optimal solution, and is discarded if it cannot produce a better solution than the best one found so far by the algorithm. The gap between the lower bound and the integer optimal value has an important impact on the size of the branching tree (more branching is needed when the gap is larger).

To improve the lower bound, cuts are added when solving the LP in each node of a branch-and-bound tree. This cutting plane method embedded in a branch-and-bound framework is called branch-and-cut. For general integer programs, cutting planes are valid inequalities that cut off a fractional solution of their LP relaxation without losing any of the feasible integer solutions. These cutting planes techniques are used when the number of linear constraints is large or valid inequalities are added to strengthen the linear relaxation (which are exponential in size). A valid inequality or cutting plane is not needed to solve the MIP but removes some fractional solutions. An example of cuts that are often used are the capacity $k$-path inequalities proposed by Kohl et al. [1999]. These cuts ensure that given subset of $S \subset N$ customers is visited by $v(S)$ vehicles with $v(S)$ the minimum number of vehicles required to serve all customers in $S$. Other valid inequalities for the VRPTW can be found in Toth and Vigo [2014].

Next to branch-and-cut, branch-and-price is one of the leading algorithms to exactly solve the VRPTW. In branch-and-price, column generation is used to solve the linear relaxations at every node in a branch-and-bound tree. Column generation is an exact method to solve a problem with many variables or columns. For the VRPTW, it is typically based on the set-partitioning model in which a minimum-cost set of feasible routes has to be determined. This is called the master problem. Since there are to many possible routes, only a subset of the feasible routing set is taken into account, resulting in a restricted master problem. New routes (columns) are generated in the pricing problem by using the dual variables of the solution of the master problem. Hence, the pricing problem searches for routes with negative reduced cost which should be included in the master problem. This process continues until no additional routes are found that can reduce the objective value. For the VRPTW, the subproblem is an elementary shortest path problem with resource constraints (ESPPRC), where the resources are the time windows and the capacity. In branch-and-price, the restricted master problem is solved in every node of the branch-and-bound tree and the corresponding dual solution is used in the pricing problem. The ESPPRC is usually solved by a labeling algorithm. The label represents the characteristics of the corresponding partial path from the depot to a customer. The labeling algorithm extends the labels by using extension functions. In this way, routes are generated by extending the paths all the way to the depot. Enumerating all possible feasible paths is time consuming, therefore dominance criteria are used to discard solutions that cannot result in an optimal solution. To speed up the algorithm, a bidirectional search can be used that extends the labels also backward from the end depot to the customers. For difficult instances (i.e., instances with extra routing constraints or many customers in one route), often a large part of the computational time is devoted to solving the ESPPRC subproblem. Therefore, heuristic solution
methods can be used to find routes with negative reduced cost in a faster way. When no route is found by the heuristic approach, the exact method is called. Another option to speed up the algorithm is to relax the subproblem by allowing the generation of paths containing cycles, e.g., the $ng$-path relaxation of Baldacci et al. [2012]. These problems are easier to solve but result in a weaker lower bound. Lastly, valid inequalities can be added to the branch-and-price algorithm to strengthen the lower bound. Many state-of-the-art algorithms are based on these branch-and-cut-and-price algorithm, which are a combination between branch-and-price and branch-and-cut.

**1.3 Cash transport and research challenges**

Because of the development of credit cards, debit cards, and contactless payments, the demand for cash is declining rapidly. In the Netherlands, the number of withdrawals and deposits at bank offices and ATMs declines with 6% to 10% per year. Yet by law, citizens should have access to an ATM within 5 km from their home address. In 2008 there were more than 9,500 ATMs in the Netherlands while in 2018 there are only 7,800 left (Trouw [2018]). Since replenishing ATMs and bank branches is a costly task due to safety requirements, banks are joining forces to efficiently replenish ATMs. Efficiently organizing the Dutch cash supply chain is the key task of Geldmaat (former Geld Service Nederland), a joint venture of the three largest Dutch banks. Convinced of the necessity to explore novel ways to manage the cash supply chain, Geldmaat joined forces with the Vrije Universiteit Amsterdam and ORTEC to secure support from the Netherlands Organisation for Scientific Research (NWO) for three PhD trajectories. This PhD thesis focuses on two research challenges in cash transport. First, because of security reasons, unpredictable routing solutions are developed while keeping the transportation cost low. Secondly, robust routing solutions and time window assignments are presented to replenish the bank branches, while keeping uncertainty in travel and service time into account. Below these research challenges are described in more detail.

**Unpredictable routing**

The transport of cash is performed by cash-in-transit (CIT) companies that transfer banknotes from cash centers to ATMs or banks. These CITs have to protect themselves against attacks by using armored trucks, vehicle tracking, special safes, etc. For security reasons and due to legal regulations, CITs have to use varying routes to serve their customers. There are two streams in the literature, each focusing on a different aspect of route unpredictability; (i) order diversification, in which the sequence in which customers are serviced changes over time (see Talarico et al. [2015a], Talarico et al. [2015b]), and (ii) arrival time diversification, in which the arrival times at customers fluctuates over time (see Michallet et al. [2014]). Interviews with safety managers of CIT companies in the Netherlands indicate that the moment of arrival at the customer, and the periods when the vehicle is stationary, are the most vulnerable parts of the
transportation journey. Therefore, we developed a model to alternate the arrival times at each customer while still minimizing transportation costs. To the best of our knowledge, there are only three studies on diversifying the arrival times at customers in the literature (see Calvo and Cordone [2003], Yan et al. [2012] and Michallet et al. [2014]). The models proposed in these papers are computationally intensive, so we proposed a more efficient a powerful solution method to diversify the arrival times. To do this, we extended the research on the VRP with multiple time windows and used the theoretical results to solve this problem.

Multiple time windows

In the vehicle routing problem with multiple time windows (VRPMTW), every customer has multiple time windows in which it is available for service. The service at every customer has to start in one of its time windows and if a vehicle arrives before or between time windows, it waits until the start of the next time window. The objective of the problem is to minimize the duration of all routes in the solution while ensuring that the capacity and time window constraints are satisfied. Since different departure times from the depot can lead to different time window selections and therefore different waiting times at the customer, the route duration depends on the departure time from the depot. Compared to the VRP with a single time window per customer, complexity increases significantly. Therefore, all existing algorithms are based on heuristic solution methods (see Favaretto et al. [2007], Belhaiza et al. [2014], Belhaiza et al. [2017]). There exist exact algorithms to minimize the duration of a given route, however, these algorithms are computationally intensive, therefore the route duration is often approximated in local search (see Tricoire et al. [2010] and Belhaiza et al. [2014]). We developed an efficient method to exactly handle route duration in the local search to improve the solution quality of heuristic solution methods. Furthermore, we proposed the first exact solution method for the VRPMTW to gain insight in the problem structure and to support the development of future (hybrid) algorithms.

Time window assignment

Service time windows at banks and shops are often imposed by the CIT or are based on an agreement between the customer and the CIT. When determining these time windows, the uncertainty in travel times should be taken into account. If the CIT would arrive too early at a bank branch, then it may have to wait because the bank employee is not available or the safe is still closed by a time lock. Since CIT trucks are most vulnerable when standing still, a robust routing solution and time window assignment have to be determined that avoids waiting time by taking uncertainty in travel and service time into account.

Not only CITs or retail shops are dealing with time window assignment (TWA) problems, but also in many attended home deliveries and repair and maintenance services time windows have to be assigned to customers. When determining time windows, there are often conflicts of interest between the service provider and the customer.
On the one hand, the service provider typically prefers wide time windows for route flexibility and to lower the risk of being too early or too late. The customer, on the other hand, prefers narrow time windows to limit the waiting time and better plan his daily activities. To improve the service to the customers, service providers want to assign small time windows to the customers (possibly within larger initial time windows) in which they can still service the customers with high probability. In the vehicle routing problem with time window assignments (VRP-TWA), the routes and time window assignments have to be determined simultaneously prior to departure. In the literature on VRP-TWA, only solution methods for the stochastic variant are proposed in which the distribution of the travel times of the arcs is completely known (see Jabali et al. [2015] and Vareias et al. [2017]). In practice, it can be hard to estimate the distribution of the travel times of all roads. Therefore, we are the first to develop a robust solution method assuming that only some characteristics of the distributions are known.

1.4 Thesis outline

In this thesis, three routing problems are considered in which service time windows play an important role. The content of this thesis is organized in five chapters. Chapter 2 deals with the vehicle routing problem with multiple time windows (VRPMTW), in which the service at each customer should start in one of its multiple time windows. This research aims to find a set of routes that minimizes the total route duration. Because of the presence of multiple time windows, the duration of a route does not only depend on the customer sequence but also on the departure time from the depot. Different departure times from the depot corresponds to a different selection of time windows, impacting vehicle waiting times. Therefore, the problem complexity of the VRPMTW is significantly higher than the complexity of the VRP with a single time window. In this chapter, an exact algorithm to minimize the duration of a given route is presented that has lower worst-case and average computational time than existing algorithms. It is also the first algorithm that can efficiently recalculate the minimal route duration in local search operations. Furthermore, the benefit of using an exact evaluation method in a heuristic solution method for the VRPMTW is experimentally shown.

Because of the high complexity of the VRPMTW, all existing algorithms are based on heuristic solution methods. In Chapter 3, the first exact solution method for the VRPMTW is proposed. Our approach is based on a branch-and-cut-and-price approach in which a tailored labeling algorithm solves the pricing problem. In this labeling algorithm, start time intervals are used that presents the start time of servicing a customer such that the preceding customers can feasibly be serviced. Based on these intervals, new dominance criteria are proposed to discard non-promising solutions. The computational experiments show that the algorithm is capable of solving instances up to 100 customers.

In Chapter 4, the models and insights of the VRPMTW are used to find unpredictable
Introduction

routing solutions. In this chapter, unpredictable routes are generated by varying the arrival times at each customer over successive visits. By deleting the arrival intervals of previous days, the problem is transformed into a VRPMTW that is solved in a rolling horizon setting of one day. Since this chapter is inspired by the transportation of cash, waiting times are not allowed because an armored truck is most vulnerable when it is stationary. Therefore, the algorithms of Chapter 2 are adjusted to efficiently determine whether routes or local search operations are time window feasible. To allow infeasible solutions in the heuristic search, four different penalty functions are proposed and tested. Our new solution approach is able to improve all benchmark instances from the literature at significantly lower computational times. Furthermore, real-world data is used to examine the trade-off between unpredictability and transportation cost.

In Chapter 3 we do not assume that the time windows are part of the input, but as in many practical applications, the assignment of time windows is part of the optimization problem. When determining routes and assigning time windows to customers prior to departure, logistics service providers are faced with uncertainty in travel and service times. The objective of the proposed vehicle routing problem with time window assignments (VRP-TWA) is to simultaneously determine routes and time window assignments such that the risk of violating the time windows is minimized. We are the first to solve the robust variant in which the probability distributions of the uncertain parameters are not completely known, but some statistics as the mean, minimum and maximum travel time can be estimated. For the subproblem of minimizing the risk of a given set of routes, we show that this problem is convex and derive the subgradient. The robust VRP-TWA is solved by iteratively generating cuts from the subproblem which are added in a branch-and-cut fashion. To test the quality and robustness of the robust VRP-TWA model, it is compared to a stochastic variant of the VRP-TWA in which the travel time distributions are known. We are the first to propose an exact solution method for the stochastic VRP-TWA and we show that the proposed robust method is close to the optimal solution. Furthermore, the trade-off between travel time and risk of violating a time window is examined.

Finally, Chapter 6 presents the conclusions and directions for future research.