Rich Routing Problems Arising in Supply Chain Management

Verena Schmid\textsuperscript{a,b}, Karl F. Doerner\textsuperscript{c}, Gilbert Laporte\textsuperscript{d}

\textsuperscript{a}Departamento de Ingeniería Industrial, Universidad de Los Andes, Bogotá, Colombia
\textsuperscript{b}Department of Business Administration, University of Vienna, Austria,
verena.schmid@univie.ac.at
\textsuperscript{c}Department of Production and Logistics Management, Johannes Kepler University Linz, Austria,
karl.doerner@jku.at
\textsuperscript{d}CIRRELT and Canada Research Chair in Distribution Management, HEC Montréal, Canada,
gilbert.laporte@cirrelt.ca

Abstract

The purpose of this paper is to provide basic models for highly relevant extensions of the classical vehicle routing problem in the context of supply chain management. The classical vehicle routing problem is extended in various ways. We will especially focus on extensions with respect to lotsizing, scheduling, packing, batching, inventory and intermodality. The proposed models allow for a more efficient use of resources, while explicitly taking into account interdependencies among the subproblems. The contribution of this survey is twofold: i) it provides an overview of recent and suitable literature for the interested scholar; ii) it presents six integrative models for the above mentioned extensions.

Keywords: Routing, Modeling, Survey, Supply Chain Management

1. Introduction

With the emergence of international markets and the growth of globalization, the management of supply chains has gained increased attention. The high complexity of the underlying procurement, production and distribution processes, as well as the increasing number of parties involved, create the necessity for efficient decision support systems. Traditionally, these three processes have mostly been solved as single problems with little interaction between them. They yield highly challenging combinatorial optimization problems which are mostly NP-hard. Because of the increase in computational power and the development of efficient metaheuristics these problems can now be solved in a more integrated fashion.

Real-world systems tend to be highly complex and intertwined. The focus on single subproblems neglects structural interdependencies and may lead to suboptimal
decisions. Nowadays an efficient and sustainable use of resources is becoming critical for the survival of organizations. This can only be achieved by considering the interdependencies of integrated supply chain services explicitly.

In this paper we discuss several rich routing problems arising in Supply Chain Management (SCM). These problems integrate several intertwined dimensions of SCM and typically contain complicated side constraints. We will present six highly relevant multi-level optimization problems, which extend the classical Vehicle Routing Problem (VPR) by integrating different down- and upstream operations. In every extension two relevant decision problems will be combined. Tactical and operational problems incorporating production (lotsizing, scheduling), warehousing (batching, packing) and inventory decisions will be covered. The collection of models is representative of various integrated SCM models which can and should be solved simultaneously. Our choice of models is not exhaustive, but should serve as an impulse for future research as there is need for efficient solution procedures capable of solving these highly complex optimization problems.

![Figure 1: Vehicle routing in the supply chain](image)

In order to tackle combined problems and eventually develop an efficient (hybrid) metaheuristic, we believe it is fundamental to deeply understand the subproblems
involved. The papers on which this study is based have been selected on the basis of their quality, readability, and on the authors’ preferences. In our opinion they provide a good starting point for further research.

Extensions related to concepts such as Vendor Managed Inventory (VMI) and Multi-echelon distribution systems have already been extensively studied since the mid-1980s (see Federgruen and Zipkin, 1984). The routing problem with container loading has been studied more recently; for example the first paper on packing and routing has been published in Gendreau et al. (2006). Following the publication of Geismar et al. (2008), production routing with scheduling and lotsizing aspects has attracted an increasing number of researchers. City logistics (see Perboli et al., 2011b) has gained a substantial amount of attention in the research community in recent years. To the best of your knowledge, combinations of warehousing and routing related aspects have not yet been considered.

The remainder of this paper is organized as follows. In Section 2 we present a model for a general routing problem with time windows, which forms the basis for all combined modeling approaches presented in this paper. Figure 1 depicts the considered subproblems and their interdependencies. Section 3 deals with a combined modeling approach integrating lotsizing with the downstream delivery process. From the lotsizing perspective, a solution with few setup operations is preferable. From the routing perspective, several smaller lots allow for more flexibility and smaller delivery costs. Section 4 addresses the integration of production schedules with the delivery process. The production process of all items is typically split into several tasks, which need to be performed on designated machines given possible precedence requirements. The scheduling problem focuses on assigning tasks to machines at an operational level. Again a similar tradeoff is encountered. The earlier the tasks associated with individual products are completed, the more flexibility is left for the routing problem.

Sections 5 and 6 focus on the integration of warehouse operations and freight space management with the distribution process. Order batching refers to the collection process of items in a warehouse using manual or automatic picking devices. An indoor routing problem may be solved for these picking devices. In the case of goods with specific handling and storage requirements, such as frozen, perishable or live products, the indoor and outdoor routing problems need to be tightly synchronized. The packing problem relates to the operational aspect of loading the goods in the vehicle. Besides the standard capacity and loading restrictions, the unloading efforts at the customer sites should be minimized.

In Section 7 we model the integration of the underlying replenishment process. Instead of reacting to orders needing to be fulfilled on any given day, an additional
degree of freedom is allowed for the distributor who can decide when to restock goods at the customer locations, while avoiding stockouts, and how to plan the vehicle routes. This gives rise to an Inventory-Routing Problem (IRP). The final extension presented in Section 8 considers a multi-echelon distribution network, where goods need to be transported over long distances and may be transshipped. Typically they are delivered to a local hub where transshipment occurs, and may then be sent to a second hub where they will be unloaded and then distributed. The problem is to synchronize these operations in order to reduce the system-wide cost. Conclusions follow in Section 9.

For each problem variants we provide a table of distinguished publications, which were selected due to the novelty of the proposed model formulation or because of the quality of their solution approach. The selection is illustrative rather than comprehensive.

2. Vehicle Routing Problems

The Vehicle Routing Problem with Time Windows (VRPTW) is defined on a network $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, where $\mathcal{V} = \{0, n + 1\} \cup \{1, \ldots, n\}$, 0 and $n + 1$ are two copies of the depot, $\mathcal{C} = \{1, \ldots, n\}$ is the customer set, and $\mathcal{A} = \{(i, j) : i, j \in \mathcal{V}, i \neq j\}$ is the arc set. With each arc $(i, j) \in \mathcal{A}$ is associated a travel cost $c_{ij}$ and a travel time $t_{ij}$. The cost and travel time matrices satisfy the triangle inequality. With each node $i \in \mathcal{V}$ is associated a demand $q_{ip}$ of product $p \in \mathcal{P}$ that needs to be satisfied by one vehicle, and a service time $s_i$. Furthermore a time window $[a_i, b_i]$ is associated with every node, within which service is supposed to start. Vehicles arriving early need to wait until the start of the time window and a unit penalty $c^{L}_i$ is incurred for starting service after $b_i$. For the depot nodes, we assume that $q_{0p} = q_{n+1,p} = s_0 = s_{n+1} = 0$. The time windows associated to the depot (where $a_0 = a_{n+1}, b_0 = b_{n+1}$) represent the earliest possible departure from the depot as well as the latest possible return time at the depot, respectively. A fleet of $m$ identical vehicles of capacity $Q$ (denoted by the set $\mathcal{K} = \{1, \ldots, m\}$) is based at the depot. Unless otherwise stated, each vehicle may execute at most one route, and split deliveries are not allowed.

The VRPTW consists of finding a set of $m$ routes such that

1. each route starts and ends at the depot,
2. each customer is visited by exactly one vehicle,
3. the total demand of customers assigned to a single vehicle does not exceed its loading capacity,
4. routes start and end within a predefined time window, and
the sum of the routing cost and of the penalties for time window violations is minimized.

For a survey of the VRPTW, the reader is referred to Cordeau et al. (2002).

2.1. Mathematical Model

The proposed model is a three-index vehicle flow formulation which uses binary flow variables \( x_{ijk} \), equal to 1 if and only if vehicle \( k \in K \) traverses arc \((i, j) \in A\). Similarly, binary variables \( y_{ik} \) are equal to 1 if and only if node \( i \) is visited by vehicle \( k \). Time variables \( w_{ik} \) define the start of service by vehicle \( k \) at node \( i \). Decision variables \( u_i \) model the delay with respect to the start of service of node \( i \).

The VRPTW can be formally described as the following multicommodity network flow model with time windows and capacity constraints.

\[
\sum_{i \in V} \sum_{j \in V} c_{ij} \sum_{k \in V} x_{ijk} + \sum_{i \in V} c_i^L u_i \rightarrow \min \tag{2.1}
\]

\[
\sum_{i \in V} \sum_{p \in P} q_{ip} y_{ik} \leq Q \quad \forall k \in K \tag{2.2}
\]

\[
\sum_{k \in K} y_{ik} = 1 \quad \forall i \in C \tag{2.3}
\]

\[
\sum_{k \in K} y_{ik} = m \quad \forall i \in \{0, n+1\} \tag{2.4}
\]

\[
\sum_{i \in V} x_{ijk} = y_{jk} \quad \forall j \in V \setminus \{0\}, k \in K \tag{2.5}
\]

\[
\sum_{j \in V} x_{ijk} = y_{ik} \quad \forall i \in V \setminus \{n+1\}, k \in K \tag{2.6}
\]

\[
w_{ik} + s_i + t_{ij} \leq w_{jk} + M(1 - x_{ijk}) \quad \forall i, j \in V, k \in K \tag{2.7}
\]

\[
a_i \leq \sum_{k \in K} w_{ik} \leq b_i + u_i \quad \forall i \in C \tag{2.8}
\]

\[
a_i \leq w_{ik} \leq b_i + u_i \quad \forall k \in K, i \in \{0, n+1\} \tag{2.9}
\]

\[
x_{ijk} \in \{0, 1\} \quad \forall i, j \in V, k \in K \tag{2.10}
\]

\[
y_{ik} \in \{0, 1\} \quad \forall i \in V, k \in K \tag{2.11}
\]

\[
w_{ik} \geq 0 \quad \forall i \in V, k \in K \tag{2.12}
\]

\[
u_i \geq 0 \quad \forall i \in V \tag{2.13}
\]
The objective function (2.1) minimizes the total cost. Constraints (2.2) state that the capacity of the vehicles cannot be exceeded. Constraints (2.3) and (2.4) ensure that every customer is visited exactly once and that the depot is used by every vehicle \( k \). Flow conservation is guaranteed by Constraints (2.5) and (2.6). Constraints (2.7) ensure feasibility with respect to time. Constraints (2.8)–(2.9) enforce the time windows, whose violations are penalized in the objective function. Finally, conditions (2.10)–(2.13) impose conditions on the variables.

This model serves as a basis for a variety of real-world extensions. The most common feature includes a different objective function. For example, one could also optimize the number of vehicles in the solution. Time windows may be treated differently depending on the application at hand. The fleet of vehicles may also be heterogeneous. Allowing split deliveries, i.e. serving a single customer by more than one vehicle, may add an extra degree of flexibility to the underlying problem. Vehicles may be based at different depots. In the case of site-dependent VRPs only a subset of vehicles can serve any given customer. Further rich real-world constraints include labour restrictions, which limit the route durations and enforce working time regulations (see Kok et al., 2010). The model can be extended to include multiple product types, which may have different storage or transportation requirements. In the VRP with pickup and deliveries items will be both picked up and delivered to customers, and precedence requirements may apply (see Berbeglia et al., 2007 for a survey). Additional synchronization requirements may be enforced for the case of rendez-vous options, which allow vehicles to meet and exchange goods (see Parragh and Doerner, 2011).

2.2. Solution Approaches

Since the decision problem is typically NP-hard the use of exact algorithms is limited and metaheuristics are the privileged solution approach. In recent years several metaheuristic concepts have been proposed, some of which are now presented.

An early metaheuristic, proposed by Glover (1986), is Tabu Search (TS). The main idea of this method is to select the best allowed solution in the neighborhood of the current solution. To avoid cycling, some attributes of recently explored solutions are set tabu (i.e. they are prohibited) for a number of iterations. Two of the first TS implementations for VRPS were proposed by Taillard (1993) and Gendreau et al. (1994). In Taillard (1993) neighborhoods are defined in terms of moves and swaps. Gendreau et al. (1994) introduced Taburoute, where single customer relocations, but no swaps are considered. In this algorithm intermediate solutions may be infeasible. This algorithmic approach has further been extended by Taillard et al. (1997) to
VRPTWs with soft time windows. Granular TS was introduced by Toth and Vigo (2003). It consists of working on a sparse graph by removing edges that are unlikely to appear in an optimal solution. Cordeau et al. (2001, 2004) later proposed a streamlined TS implementation, applicable to several VRP variants.

Several other local search metaheuristics have been put forward. For example Variable Neighborhood Search (VNS), introduced by Mladenović and Hansen (1997), sequentially nested neighborhoods of increasing sizes are considered. The exploration of a neighborhood is done by only accepting improving moves. When this becomes impossible, the next neighborhood is considered. Bräysy (2003) have successfully applied this concept to the VRPTW. Other successful concepts based on VNS have been proposed in Prins (2004) for the multi-depot VRPTW and by Hemmelmayr et al. (2009a) for the periodic VRP.

Pisinger and Ropke (2007) have introduced the concept of Adaptive Large Neighborhood Search (ALNS) for the VRP. Solutions are partially destroyed by means of a ruin operator and are recreated by applying a constructive repair mechanism. The choice of operators is biased to favor those whose past performance has been better.

A completely different optimization concept is bioinspired. Evolutionary algorithms such as Genetic Algorithms (GA) and Ant Colony Optimization (ACO) mimic concepts coming from natural selection and populations respectively. The most promising algorithms for routing problems have been developed by Prins (2004) and Reimann et al. (2004).

We refer to Bräysy and Gendreau (2005a) and Bräysy and Gendreau (2005b) for recent surveys of heuristics in general. An overview of heuristics dedicated to problems from the domain of routing problems can be found in Laporte (2007).

Exact algorithms are of limited applicability for most routing problems. However, a number of successful exact methods have been applied to the basic VRP. Baldacci et al. (2008) and Baldacci and Mingozzi (2009) solve VRPs based on a set partitioning formulation using column generation. Furthermore, algorithms based on branch-and-cut have been introduced by Letchford et al. (2002) and Lysgaard et al. (2004). A combination of the previously mentioned concepts has been published in Fukasawa et al. (2006). Feillet et al. (2004) solved the VRPTW using column generation. An approach based on Lagrangian relaxation has been proposed in Kallehauge et al. (2006) and Kallehauge et al. (2007).

For recent surveys on exact algorithms for routing problems we refer to Baldacci et al. (2007, 2010); Kallehauge (2008); Laporte (2009).

When scanning the recent literature, we have observed an increased interest in so-called matheuristics (Doerner and Schmid, 2010). These methods combine exact and heuristic approaches and have proven successful for rich vehicle routing prob-
lems. In the following we summarize the most important hybridization schemes.
When solving rich VRPs three main algorithmic frameworks are commonly used: i) set-covering based approaches, where a subset of feasible routes is (heuristically) generated and the restricted problem is solved optimally; ii) local branching, which is a promising approach when it comes to solving large instances of a mixed-integer problem; the solution process is sped up by fixing and unfixing decision variables and hence reducing the time needed to solve the resulting subproblem; and iii) decomposition, a highly efficient method which partitions the problem into smaller subproblems by exploiting their structure and hierarchy. We believe that the latter idea is particularly promising for combined problems such as those presented in this paper. Once a problem has been decomposed, it can usually be solved efficiently by means of an available algorithm. The challenge remains to find effective information mechanisms that will guide the solution process of the subproblems.

3. Lotsizing Problems in Production Planning

Two central problem classes arise in the area of production planning. In the first, lotsizing decisions are made in order to determine a rough production plan taking into account machine resources and availabilities at a tactical and aggregate level. A more detailed schedule can then be designed, where resource capacities are explicitly taken into account, and feasibility with respect to their usage and the timing of production is considered. Traditionally these two subproblems have been solved independently. In this and the following section, we focus on their integration with distribution related aspects, both at a tactical level. The integrated view is especially important in situations where there only exist limited intermediate storage facilities, and manufactured products need to be distributed immediately, a situation frequently encountered in the case of highly perishable goods.

When solving lotsizing problems one typically tries to optimize the tradeoff between setup and inventory holding costs. Producing large batches of products at a time reduces the number of setup operations involved. However, the resulting inventory stock levels and holding costs become higher. In contrast, producing smaller quantities of products at a time leads to lower average inventory costs, but results in a setup costs increase. The lotsizing problem can be integrated within a downstream supply chain model. Customers may only be serviced once all required products are ready to be dispatched and shipped. Typical applications include the production and subsequent delivery of bakery products, or daily newspapers which are produced during a night shift and need to be distributed early on the following day.

The earlier customer orders are ready, the earlier vehicles may depart from the central production site and distribute the ordered items to customers. Inventory
holding costs at the production site are considered to be negligible, but a tradeoff still exists.

3.1. Literature Review

A survey of scheduling and lotsizing problems and their extensions can be found in Drexl and Kimms (1997). Another good overview on capacitated lotsizing problems is that of Quadt and Kuhn (2008). Buschkühl et al. (2010) survey dynamic capacitated lotsizing problems, whereas Stadtler (2000, 2003) presents an improved rolling schedule algorithm for the dynamic single and multi level lotsizing problem. Different problem formulations for the dynamic multi level lotsizing problem are discussed in Stadtler (1996), and three metaheuristics have been implemented by Tempelmeier and Derstroff (1996), Pitakaso et al. (2006) and Almder (2010). General models focusing on the integration of lotsizing and scheduling related aspects are discussed in Fleischmann and Meyr (1997). Algorithms for the integration of the two subproblems with sequence-dependent setup times have been proposed by Mateus et al. (2010) and by Meyr (2002). Lütke Entrup et al. (2005) have developed an integrated algorithm taking into account special requirements of perishable goods.

The importance of the interdependence of the two subproblems has already been mentioned in several surveys (see for example Ehrengüç et al., 1999 and Mula et al., 2010) but few integrated algorithms exist. Chandra and Fisher (1994) have demonstrated on small instances that gains can be obtained by integrating the two aspects of the problem within the same algorithm. This is especially true for industries characterized by relatively high distribution costs compared to setup related costs (e.g. bottled and packaged goods and wood products). Another heuristic approach for a single product in a multiple period horizon is presented in Boudia et al. (2008).

Integrating lotsizing and distribution planning is particularly important for products with a limited shelf life, such as newspapers and food. Different solution methods for problems coming from this area of application have been proposed in Hurter and Van Buer (1996), Van Buer et al. (1999), Chen and Vairaktarakis (2005), Cunha and Mutarelli (2007), Ahumada and Villalobos (2011) and Bilgen and Günther (2010). The paper of Geismar et al. (2008) presents a heuristic for an integrated production and transportation problem for products with short life span. Additional extensions taking also into account inventory related aspects have been proposed in Fumero and Vercellis (1999), Lei et al. (2006), Vidyarthi et al. (2007), Bard and Nananukul (2009) and Bard and Nananukul (2010). More recently Adulyasak et al. (2012) suggested an optimization-based adaptive large neighborhood search heuristic, which is able to produce high quality solutions in short computing times.
An overview on selected publications in the domain of lotsizing and routing is shown in Table 1.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Problem characteristics</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chandra and Fisher (1994)</td>
<td>basic combined lotsizing and routing model</td>
<td>1-step (MIP) and 2-step approach (MIP for production decision, simple heuristics for subsequent routing decision)</td>
</tr>
<tr>
<td>Geismar et al. (2008)</td>
<td>integrated production and transportation scheduling problem, no inventory, perishable goods</td>
<td>2-step heuristic (GA/MA for routing, heuristic for scheduling)</td>
</tr>
<tr>
<td>Hurter and Van Buer (1996)</td>
<td>perishable goods (newspapers)</td>
<td>2-step heuristic (RFCS heuristic for VRP, sorted list of routes implies schedule)</td>
</tr>
<tr>
<td>Bilgen and Günther (2010)</td>
<td>block planning for production, FTL &amp; LTL routing</td>
<td>1-step approach (MIP)</td>
</tr>
<tr>
<td>Bard and Nananukul (2009)</td>
<td>integrated production, inventory &amp; distribution routing</td>
<td>2-step approach (solve lot-sizing decision with surrogate for routing costs using MIP, resulting VRP using TS)</td>
</tr>
<tr>
<td>Adulyasak et al. (2012)</td>
<td>integrated production, inventory &amp; distribution routing</td>
<td>2-step approach (different setup solutions using MIP, resulting VRP using ALNS and network flow)</td>
</tr>
</tbody>
</table>


3.2. Mathematical Model

Let $\mathcal{T} = \{1, \ldots, |\mathcal{T}|\}$ denote the set of time periods per day available for production, and let $\mathcal{T}' = \mathcal{T}\setminus\{|\mathcal{T}|\}$. The total production capacity in $t$ is denoted by $C_t$. Each setup operation of product $p$ is penalized by a factor $c_p^S$ in the objective function.

Let $v_{pt}$ be the amount of product $p$ produced in time slot $t$. The binary decision variable $v_{pt}^S$ is equal to 1 if and only if product $p$ is made in period $t$. The stock level of product $p$ at the start of period $t$ is $h_{pt}$. The amount of goods of type $p$ to be loaded into vehicle $k$ in time period $t$ is denoted by $o_{pkt}$. The binary decision
variables $z_{kt}$ link the production and the routing decision: they are equal to 1 if and only if vehicle $k$ leaves the depot at time $t$.

The underlying mathematical program is as follows:

$$\sum_{p \in P} \sum_{t \in T} c_p^S v^S_{pt} + \sum_{i \in V} \sum_{j \in V} c_{ij} \sum_{k \in K} x_{ijk} + \sum_{i \in V} c_i^L u_i \rightarrow \min$$  \hfill (3.1)

$$\sum_{p \in P} v_{pt} \leq C_t \quad \forall t \in T$$  \hfill (3.2)

$$v_{pt} \leq M v^S_{pt} \quad \forall p \in P, t \in T$$  \hfill (3.3)

$$h_{p1} = 0 \quad \forall p \in P$$  \hfill (3.4)

$$h_{pt} + v_{pt} - \sum_{k \in K} o_{pkt} = h_{p,t+1} \quad \forall p \in P, t \in T'$$  \hfill (3.5)

$$\sum_{t \in T} z_{kt} \leq 1 \quad \forall k \in K$$  \hfill (3.6)

$$w_{0k} = \sum_{t \in T} t z_{kt} \quad \forall k \in K$$  \hfill (3.7)

$$o_{pkt} \leq \sum_{i \in V} q_{ip} y_{ik} + M (1 - z_{kt}) \quad \forall p \in P, k \in K, t \in T$$  \hfill (3.8)

$$o_{pkt} \geq \sum_{i \in V} q_{ip} y_{ik} - M (1 - z_{kt}) \quad \forall p \in P, k \in K, t \in T$$  \hfill (3.9)

$$v_{pt}, h_{pt} \geq 0 \quad \forall p \in P, t \in T$$  \hfill (3.10)

$$v^S_{pt} \in \{0, 1\} \quad \forall p \in P, t \in T$$  \hfill (3.11)

$$o_{pkt} \geq 0 \quad \forall p \in P, k \in K, t \in T$$  \hfill (3.12)

$$z_{kt} \in \{0, 1\} \quad \forall k \in K, t \in T.$$.  \hfill (3.13)

The objective function (3.1) minimizes costs related to routing and setup times. In addition to (2.2)–(2.13) the following constraints are needed. Constraints (3.2) ensure that the available capacity per time period is not exceeded. Constraints (3.3) mean that setup costs accrue if product $p$ is made in period $t$. Once items have been produced they will temporarily be stored at the local inventory. Constraints (3.4) and (3.5) model the inventory balance equation per product and time period. Constraints are imposed in order to link the routing model with the lotsizing model and to ensure its feasibility with respect to the availability of products. By Constraints (3.6), a vehicle may leave the depot at most once. Constraints (3.7) link the actual start of the
tour of vehicle \( k \) with the corresponding binary indicator \( u_{t}^{k} \). Constraints (3.8) and (3.9) define the load of vehicle \( k \) upon leaving the production site. Constraints (3.10)–(3.13) impose bounds on the decision variables.

4. Machine Scheduling Problems in Production Planning

In the previous section, the VRP was extended with respect to lotsizing, a problem arising at the tactical level. The resulting operational decision problem is the combination of machine scheduling and routing. Scheduling is a well-studied problem in operations management. Depending on the characteristics of the underlying production process one can differentiate between job shop, flow shop and the more general open job problems. Scheduling lies at the central core of operational production planning, where jobs need to be assigned to resources (machines), and schedules must fulfill a number of precedence or other technical requirements. The outcome of the planning process itself is much more detailed, compared to lotsizing problems, since it explicitly considers the exact timing and assignments decisions. Once a good has been produced, it may be delivered to the next layer of the supply chain, and transportation may only start after the last task of the production process is completed. Scheduling decisions interfere with decisions for the underlying routing problem and vice versa. By combining scheduling and routing decisions we aim for better decisions on a broader perspective.

4.1. Literature Review

Different hybrid and metaheuristic approaches (see for instance Blum, 2005, Guéret et al., 2000, Hall et al., 2001 and Stecco et al., 2008, 2009) have been proposed for different variants of the scheduling problem. Typical applications include just-in-time production processes, where orders become known at short notice and the production process, as well as the subsequent distribution process, needs to be tightly managed on a daily basis. Outside production related contexts, scheduling problems also arise in the field of health care. In hospitals, main resources such as operation theaters need to be assigned to patients. The resulting optimization problem can be formally modeled as a job shop scheduling problem, where surgeries are scheduled in operating rooms. For models related to this context, we refer the reader to Cardoen et al. (2009, 2010) and to Van Oostrum et al. (2008). An integrated approach for both scheduling and routing related aspects in this domain has been developed in Schmid and Doerner (2011).

An integrated model for production and routing problems has been proposed by Xu and Chiu (2001) for the routing of service technicians. Technicians have
different skill levels, and tasks must be assigned (i.e. scheduled) according to their skill levels. Once tasks have been assigned, a routing problem with time windows and other constraints must be solved. Chang and Lee (2004) focus on the scheduling problem, while considering the consequences in terms of capacity requirements for the delivery problem. More recent papers take into account the underlying routing problem explicitly and solve the combined problem simultaneously (see Kovacs et al., 2011 and Bredström and Rönqvist, 2008).

An overview on selected publications on machine scheduling and routing can be found in Table 2.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Problem characteristics</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kovacs et al. (2011)</td>
<td>job assignment, skill levels, routing</td>
<td>1-step, ALNS</td>
</tr>
<tr>
<td>Bredström and Rönqvist (2008)</td>
<td>job assignment, skill levels, routing, precedence, synchronization</td>
<td>1-step, MIP-based heuristic</td>
</tr>
</tbody>
</table>

MIP: mixed integer program, ALNS: adaptive large neighborhood search.

4.2. Mathematical Model

The job shop scheduling problem is defined as follows. A set of $P = \{1, \ldots, P\}$ of products must be processed on machines. The production process of each individual product $p$ consists of a set of tasks $h \in B_p$ that need to be performed on specific machines $r \in R$. The processing time of each task $h$ is denoted by $d_h$. Once the operation of a task has started it cannot be interrupted. The set of all tasks is denoted by $B = \bigcup_{p \in P} B_p$, the first (last) task associated with product $p$ is denoted by $e_p$ ($f_p$). Tasks must be scheduled on machines in such a way that the sequence of tasks associated with the same final product $p$ satisfies some precedence requirements. The goal is to determine a feasible schedule, such that each machine may only process one task at a time and the average cycle time (i.e. the time for processing all tasks) per product is minimized. The demand of customer $i$ for product $p$ is denoted by $q_{ip}$. Let $v_h$ be the starting time of the operation associated with task $h \in B$. Binary decision variables $z_{hl}$ are defined for all pairs of tasks $h, l \in B$, processed on the same machine $r \in R$; they are equal to 1 if and only if task $h$ is executed before task $l$. Let $g_p$ denote the earliest starting date of the production process of product $p$. The formulation is then:

$$\sum_{p \in P} c^C_p (v_{f_p} + d_{f_p} - g_p) + \sum_{i \in V} \sum_{j \in V} c_{ij} \sum_{k \in K} x_{ijk} + \sum_{i \in V} c^L_i u_i \rightarrow \min \quad (4.1)$$
The objective function (4.1) combines routing and scheduling costs, namely the minimization of cycle times. The resulting cycle time of product \(p\) is penalized by parameter \(c_p^C\). Constraints (4.2) and (4.3) are disjunctive constraints which ensure that machines can only execute at most one task at a time. Precedence requirements are expressed by Constraints (4.4). The start of production or product \(p\) may not start before the associated release date \(g_p\) (4.5). A vehicle servicing customer \(c\) may only leave from the depot once the production process of its demand is completed (4.6).

5. Order Batching in Warehouse Logistics

After having focused on production related extensions, we now look at inventory related issues. Warehouse and distribution operations are two important problems in logistics. The goal consists of efficiently picking several orders consisting of several items in a warehouse. These items are typically dispersed within the warehouse. Picking operations are performed by fully automated guided vehicles or, more traditionally, by man-to-order picking systems. In order to avoid inefficient batching operations, items belonging to the same order have to be picked up in the same picker route. The two resulting subproblems are classified as order batching and order picking problem. The first problem consists of combining different orders on a single picker route. The aim of the second problem is to identify good routes for all batched items.

After orders have been batched and picked up they need to be distributed to their destinations. This is typically achieved by solving a VRPTW. Traditionally batching-routing operations and vehicle routing have been tackled independently, but ideally this should not be the case since the solutions to these two problems are interrelated.
The external distribution process, from the warehouse to the customers, can be modeled as a standard VRPTW. The indoor picking process is an adaptation of a VRP model, but the number of tours per picking device is not limited to one per planning period, typically a day. Furthermore no explicit time windows or due dates have to be satisfied. The goal is to efficiently batch customers into picking orders, given set of customer orders and the storage location of all items, so that the total length of all picking tours required is minimized. Since customer orders are not allowed to be split we assume that the capacity of the largest customer order still fits into any (indoor) picking device and (outdoor) vehicle.

We assume that the inventory level in the warehouse is large enough to cover all demands and each item is associated with a unique storage location. Each order item is associated with a unique storage location in the warehouse.

5.1. Literature Review

The combined batching and picking problem is a well-known optimization problem. An overview of traditional heuristics for the basic problem can be found in De Koster et al. (1999). Two approaches based on ant colony optimization and iterated local search have been proposed in Henn et al. (2010). The design decisions of the proposed methods are based on problem-specific heuristic decision rules. Different metaheuristic rules for solving the underlying routing problem have been proposed in Ho and Tseng (2006). In this paper the solution method for the embedded routing problem is randomised by borrowing some concepts of simulated annealing. An exact branch-and-price algorithm has been developed in Gademann and Velde (2005).

These problems have been further extended. Petersen and Schmenner (1999) take storage policies into account (i.e. where to store different items). In Theys et al. (2010) more emphasis is put on the development of efficient algorithms for the underlying routing problem. The problem has been further extended to a dynamic real-time setting with rolling horizon by Le-Anh et al. (2010).

A step towards the direction of integrating batching-routing operations and vehicle routing has been made by Tsai et al. (2008) who considers due dates explicitly the the first subproblem. In this paper we propose an integrated model formulation.

An overview on selected publications in the domain of order batching and routing can be found in Table 3.
Table 3: Selected literature on order batching and routing

<table>
<thead>
<tr>
<th>Reference</th>
<th>Problem characteristics</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Koster et al. (1999)</td>
<td>survey</td>
<td>ILS and savings-based ants for batching, Largest-Gap &amp; S-Shape heuristics</td>
</tr>
<tr>
<td>Henn et al. (2010)</td>
<td>order batching and intra-warehouse routing</td>
<td></td>
</tr>
<tr>
<td>Tsai et al. (2008)</td>
<td>minimizing downstream impact by penalizing laten-</td>
<td>2-step GA for batching and intra-warehouse routing</td>
</tr>
<tr>
<td></td>
<td>t of deliveries</td>
<td></td>
</tr>
</tbody>
</table>

ILS: iterated local search, GA: genetic algorithm

5.2. Mathematical Model

Mathematically the problem can be described as follows. Let \( \mathcal{P} = \{1, \ldots, n^I\} \) denote the set of order items and let \( q_p^I \) be the size of item \( p \). The set of ordered items of customer \( i \) is denoted by \( \mathcal{P}_i \). The order picking of batched routes within the warehouse can be modeled as a special type of VRP. In order to improve readability, the corresponding parameters for the indoor-related routing aspect are highlighted by the superscript \( I \). A fleet \( \mathcal{K}^I = \{1, \ldots, m^I\} \) identical picking devices, each with capacity \( Q^I \), is available at the central depot within the warehouse. The set of tours executed by each picking device is denoted by \( \mathcal{R} = \{1, \ldots, |\mathcal{R}|\} \). The VRP with Order Batching is defined on a network \( \mathcal{G}^I = (\mathcal{V}^I, \mathcal{A}^I) \), where \( \mathcal{V}^I \) denotes the set of nodes, which consists of the set of storage locations associated with order items \( p \in \mathcal{P} \) and two copies 0 and \( n^I + 1 \) of the depot. Non-negative costs \( c_{hl}^I \) and travel times \( t_{hl}^I \) are associated with every arc \( (h, l) \in \mathcal{A}^I \). The total demand of customer \( i \) is given by \( q_i \), where \( \sum_{h \in \mathcal{P}} q_h^I = q_i \).

The model extends that of the VRPTW defined in Section 2, which will cover the transportation part from the warehouse to the customers. Additional decision variables have to be defined in order to model the intra-warehouse picking process. Binary flow variables \( x_{hklr}^I \) take on value one if and only if item \( h \) is picked up immediately before item \( l \) by vehicle \( k \) on its \( r \)-th tour (where \( h, l \in \mathcal{V}^I, k \in \mathcal{K}^I, r \in \mathcal{R} \)). Binary indicator variables \( y_{hkr}^I \) \((h \in \mathcal{V}^I, k \in \mathcal{K}^I, r \in \mathcal{R})\) are equal to 1 if and only if picking device \( k \) collects item \( h \) on its \( r \)-th tour. Similarly \( z_{ikr}^I \) are equal to 1 if and only if picking device \( k \) collects all order items of customer \( i \) on route \( r \). The arrival time of picking device \( k \) on route \( r \) at the location \( h \) is modeled in terms of \( w_{hkr}^I \).

The formulation is as follows:

\[
\sum_{h \in \mathcal{V}^I} \sum_{l \in \mathcal{V}^I} \sum_{k \in \mathcal{K}^I} \sum_{r \in \mathcal{R}} c_{hl}^I x_{hklr}^I + \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} c_{ij} \sum_{k \in \mathcal{K}} x_{ijk} + \sum_{i \in \mathcal{V}} c_{i}^I u_i \rightarrow \min \] (5.1)
\[
\sum_{h \in \mathcal{P}} q_h^I y_{hkr} \leq Q^I \quad \forall k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.2)
\]
\[
\sum_{k \in \mathcal{K}^I} \sum_{r \in \mathcal{R}} y_{hkr} = 1 \quad \forall h \in \mathcal{P} \quad (5.3)
\]
\[
\sum_{k \in \mathcal{K}^I} y_{hkr} = m_I^h \quad \forall h \in \{0, n^I + 1\}, r \in \mathcal{R} \quad (5.4)
\]
\[
\sum_{h \in \mathcal{V}^I} x_{hlkr} = y_{hkr} \quad \forall l \in \mathcal{V}^I \setminus \{0\}, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.5)
\]
\[
\sum_{l \in \mathcal{V}^I} x_{hlkr} = y_{hkr} \quad \forall h \in \mathcal{V}^I \setminus \{n^I + 1\}, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.6)
\]
\[
w_{hkr}^I + s_{h}^I + t_{hl}^I \leq w_{lkr}^I + M(1 - x_{hlkr}^I) \quad \forall h, l \in \mathcal{V}^I, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.7)
\]
\[
w_{n^I + 1, k, r - 1}^I \leq w_{lkr}^I \quad \forall k \in \mathcal{K}^I, r \in \mathcal{R} \setminus \{1\} \quad (5.8)
\]
\[
\sum_{k \in \mathcal{K}^I} \sum_{r \in \mathcal{R}} z_{ikr}^I = 1 \quad \forall i \in \mathcal{C} \quad (5.9)
\]
\[
y_{hkr}^I = z_{ikr}^I \quad \forall i \in \mathcal{C}, h \in \mathcal{P}_i, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.10)
\]
\[
w_{0k} \geq w_{n^I + 1, k, r}^I + M(2 - z_{i,k,l,r}^I - y_{ik}) \quad \forall i \in \mathcal{C}, k \in \mathcal{K}, k^I \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.11)
\]
\[
x_{hlkr}^I \in \{0, 1\} \quad \forall h \in \mathcal{V}^I, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.12)
\]
\[
y_{hkr}^I \in \{0, 1\} \quad \forall h \in \mathcal{V}^I, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.13)
\]
\[
w_{hkr}^I \geq 0 \quad \forall h \in \mathcal{V}^I, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.14)
\]
\[
z_{ikr}^I \in \{0, 1\} \quad \forall i \in \mathcal{C}, k \in \mathcal{K}^I, r \in \mathcal{R} \quad (5.15)
\]

The objective (5.1) is to minimize the total costs for transportation in the warehouse and to customers, including a penalty for violating their time windows. Constraints (5.2)–(5.7) are a straightforward extension of the corresponding VRPTW constraints, adapted to the intra-warehouse picking process for several identical picking devices. Constraints (5.8) mean that the picking device routes are feasible and do not overlap. Constraints (5.9) and (5.10) ensure that items ordered by any customer all are picked up by the same picking device on the same tour. Finally constraints (5.11) link the indoor and outdoor routing problems and ensure that vehicles may only depart from the warehouse to finally service the customers, once all ordered items have been collected and dispatched.
6. Container and Pallet Loading in Warehouse Logistics

Once items have been picked in a warehouse they need to be loaded into vehicles. The one-dimensional capacity restriction of the VRP ensures that the total capacity of the vehicle is not exceeded. In real life, however, the items are characterized by several dimensions. In this section we present a model for a two-dimensional container loading and routing problem.

As such, loading is already a challenging combinatorial optimization problem. The goal is to optimally cut or pack a set of items into a set of bins of predefined dimensions. Several versions of the problem exist, depending on the dimension of items (one-, two- or multi-dimensional) and their characteristics (guillotine, fragility, stability, etc.). In practical applications, items coming from a warehouse usually need to be loaded into a vehicle before being delivered to customers. The sequence of customers in the vehicle routes should be designed to avoid unnecessary unloading and repacking operations. Solving these packing and routing separately may lead to suboptimal decisions. Hence we will present an integrated model formulation capturing their interdependencies.

6.1. Literature Review

An excellent typology on general packing problems can be found in Wäscher et al. (2007). The two-dimensional bin packing problem has gained attention in the recent years. A survey is presented in Lodi et al. (2002b). An approximation algorithm based on TS for the two-dimensional bin packing problem can be found in Lodi et al. (1999a,b) and Baumgartner et al. (2011). An exact approach is presented in Martello and Vigo (1998) and Fekete et al. (2007).

A natural extension comes by adding a third dimension. The underlying problem is highly generic, as it can also be applied to more general container loading problems. A TS heuristic was first proposed by Lodi et al. (2002a), and a guided local search heuristic was presented in Faroe et al. (2003). These algorithms were outperformed in Crainic et al. (2008) and Perboli et al. (2011a), where two- and three-dimensional problems are tackled using a sophisticated constructive heuristic. Exact algorithms are described in Martello et al. (2000) and in den Boef et al. (2005). A complex variant of the container loading problem, arising from a real-world industrial application including additional features is presented in Ceschia and Schaerf (2010) and Christensen and Rousoe (2009).

A survey on applications of combined packing and routing problems can be found in Iori and Martello (2010). In its simplest version the consequences with respect to routing costs have been considered by means of Last-In-First-Out (LIFO) or First-In-First-Out (FIFO) constraints, which have been proposed in Carrabs et al. (2007),
Cordeau et al. (2010a), Cordeau et al. (2010b) and Erdoğan et al. (2009). Several policies allowing unloading and reloading along a vehicle route have been investigated by Battarra et al. (2010).

Advanced models considering multiple compartments within vehicles have been proposed by Chajakis and Guignard (2003), Oppen and Løkketangen (2008), Derigs et al. (2011) and El Fallahi et al. (2008). Extensions based on multiple loading stacks, where also rear-loading considerations need to be enforced have been presented in Côté et al. (2009), Doerner et al. (2007), Tricoire et al. (2011), Petersen and Madsen (2009) and Petersen et al. (2010).

More realistic extensions based on the integration of the two-dimensional vehicle loading and routing problem (2L-CVRP) have been proposed by Iori et al. (2007), who solved the underlying problem exactly. Gendreau et al. (2008) have developed a TS heuristic for this problem. The problem has recently been tackled by Duhamel et al. (2011), Fuellerer et al. (2009), Leung et al. (2011) and Zachariadis et al. (2009). The three-dimensional vehicle loading and routing problem (3L-CVRP) was first solved by Gendreau et al. (2006) using TS. Further approaches include those of Fuellerer et al. (2010), Moura and Oliveira (2009) and Tarantilis et al. (2009).

In this section we focus on a combination of the classical two-dimensional bin packing problem, and on the downstream routing problem. Items to be packed are mostly three-dimensional by nature. In practice, however, only two dimensions are typically considered because customer orders tend to be loaded on pallets which cannot be stacked on top of each other.

An overview on selected publications on combined container loading and routing problems is depicted in Table 4.
### Table 4: Selected literature on container loading and routing

<table>
<thead>
<tr>
<th>Reference</th>
<th>Problem characteristics</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>routing &amp; 1d packing</strong></td>
<td>1d, LIFO loading, P&amp;D, TSP</td>
<td>1-step, VNS</td>
</tr>
<tr>
<td>Carrabs et al. (2007)</td>
<td>1d, handling cost for rearrangements at customer locations, routing costs dominate, P&amp;D, TSP</td>
<td>2-step heuristic: solve routing &amp; packing problem sequentially to optimality, B&amp;C for optimizing routing &amp; packing simultaneously, restricted handling decisions</td>
</tr>
<tr>
<td>Battarra et al. (2010)</td>
<td>2-step heuristic: solve routing &amp; packing problem sequentially to optimality, B&amp;C for optimizing routing &amp; packing simultaneously, restricted handling decisions</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>routing &amp; 1.5d packing</strong></th>
<th>1.5d, VRP, compartments</th>
<th>1-step, LS, LNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derigs et al. (2011)</td>
<td>1.5d, TSP, multiple compartments, LIFO</td>
<td>LNS, embedded hierarchically, subproblem heuristically</td>
</tr>
<tr>
<td>Côté et al. (2009)</td>
<td>1.5d, 2-TSP, multiple compartments, fixed size, LIFO</td>
<td>SA, TS, ILS, LNS, embedded hierarchically, subproblem heuristically</td>
</tr>
<tr>
<td>Petersen and Madsen (2009)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>routing &amp; 2d packing</strong></th>
<th>2d, VRP, with(out) rearrangement</th>
<th>B&amp;C, embedded hierarchically, subproblem 2d single bin packing per route/vehicle: B&amp;B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iori et al. (2007)</td>
<td>2d, VRP, with(out) rearrangement, with(out) orientation, with(out) loading restrictions</td>
<td>ACO, embedded hierarchically, subproblem heuristically (bottom left/touching parameter) or B&amp;B guided TS, embedded hierarchically, subproblem heuristically</td>
</tr>
<tr>
<td>Fuellerer et al. (2009)</td>
<td>2d, VRP, sequential, unrestricted</td>
<td></td>
</tr>
<tr>
<td>Zachariadis et al. (2009)</td>
<td>3d, VRP, fragility, supporting areas, LIFO</td>
<td>TS, embedded hierarchically (TS), subproblem heuristically (TS)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>routing &amp; 3d packing</strong></th>
<th>3d, VRP, fragility, supporting areas, LIFO</th>
<th>TS, embedded hierarchically (TS), subproblem heuristically (TS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gendreau et al. (2006)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


### 6.2. Mathematical Model

The demand of customer $i \in C$ is typically characterized by a set of items $P_i$ of total weight $q_i$, each characterized by its width $f_h$ and height $e_h$ respectively, where $h \in P$ and $P = \bigcup_{i \in C} P_i$. The fleet of vehicles is characterized by a rectangular loading space of width $f^K$ and height $e^K$. We assume that items have a fixed orientation,
that they have to be packed parallel to the edges of the loading surface and cannot be rotated. Furthermore, since shifting items around when unloading at customers is often unpractical, we additionally impose that items should be packed in a way that unloading can be done easily by just sliding them toward the rear of the vehicle.

To model the underlying problem mathematically the following additional decision variables have to be defined. Variable $v_{hk}$ is equal to 1 if and only if item $h$ is placed in vehicle $k$. Binary variables $z_{hl}$ are equal to 1 if and only if item $h$ and $l$ are placed into the same vehicle. To avoid items from overlapping the binary variable $o_{L_{hl}}$ takes the value 1 if and only if item $l$ is located left of item $h$, hence these two items do not overlap horizontally. The binary variable $o_{U_{hl}}$ is equal to 1 if and only if item $l$ is located under item $h$, i.e. these items do not overlap vertically. Finally $c^X_h$ ($c^Y_h$) are used to model the $x$ ($y$) coordinates of the bottom left corner of item $h$.

The model can formally be described as follows:

$$\sum_{k \in K} c^F_k (1 - x_{0,n+1,k}) + \sum_{i \in V} \sum_{j \in V} c_{ij} \sum_{k \in K} x_{ijk} + \sum_{i \in V} c^L_i u_i \rightarrow \min \quad (6.1)$$

$$\sum_{k \in K} v_{hk} = 1 \quad \forall h \in \mathcal{P} \quad (6.2)$$

$$y_{ik} = v_{hk} \quad \forall k \in K, i \in C, h \in \mathcal{P}_i \quad (6.3)$$

$$c^X_h + f^P_h \leq f^K \quad \forall h \in \mathcal{P} \quad (6.4)$$

$$c^Y_h + e^P_h \leq e^K \quad \forall h \in \mathcal{P} \quad (6.5)$$

$$c^X_h + f^P_h \leq c^X_i + f^K (2 - z_{hl} - o^L_{hl}) \quad \forall h, l \in \mathcal{P} \quad (6.6)$$

$$c^Y_h + f^P_h \leq c^X_i + f^K (2 - z_{hl} - o^L_{hl}) \quad \forall h, l \in \mathcal{P} \quad (6.7)$$

$$c^Y_h + e^P_h \leq c^X_i + e^K (2 - z_{hl} - o^U_{hl}) \quad \forall h, l \in \mathcal{P} \quad (6.8)$$

$$c^Y_h + e^P_h \leq c^X_i + e^K (2 - z_{hl} - o^U_{hl}) \quad \forall h, l \in \mathcal{P} \quad (6.9)$$

$$o^L_{hl} + o^L_{lh} + o^U_{hl} + o^U_{lh} \geq 1 \quad \forall h, l \in \mathcal{P} \quad (6.10)$$

$$\sum_{i \in C} y_{ik} \leq M(1 - x_{0,n+1,k}) \quad \forall k \in K \quad (6.11)$$

$$c^X_h, c^Y_h \geq 0 \quad \forall h \in \mathcal{P} \quad (6.12)$$

$$z_{hl}, o^L_{hl}, o^U_{hl} \in \{0, 1\} \quad \forall h, l \in \mathcal{P} \quad (6.13)$$

$$v_{hk} \in \{0, 1\} \quad \forall h \in \mathcal{P}, k \in K, \quad (6.14)$$

and (2.2)–(2.13).

The objective (6.1) combines the usual routing considerations (travel costs and penalties for not meeting predefined time windows) as well as the number of vehicles.
Parameter $c_k^F$ is the cost per vehicle. Constraints (6.2) and (6.3) in combination with (2.3) impose that all items must be packed in vehicles and that all items $P_i$ belonging to customer $i$ must be packed in the same vehicle $k$. To ensure that items do not overhang the loading area of the vehicle Constraints (6.4) and (6.5) have to hold. Constraints (6.6)–(6.9) ensure that no two items overlap, if placed in the same vehicle $k$. Constraints (6.11) ensure that items may only be packed in vehicles that are actually used. For the routing part Constraints (2.2)–(2.13) described in Section 2 must hold.

The combined loading and routing problem may be extended in several ways. An example is the packing of three-dimensional items which can be stacked on top of each other. Due to stability issues, it may be necessary to guarantee a sufficient supporting area. Depending on the weight and stability of individual items, their fragility and stackability may also have to be taken into account. Depending on the nature of the items, they may have to be rotated before being packed in the vehicle.

7. Inventory Management and Vendor Managed Inventory

In this and the next sections we will focus on decision problems that emerge further downstream in the supply chain. More specifically, we will give an overview of the recent literature on IRPs. This problem combines storage at customer locations and routing decisions. Concepts such as Vendor Managed Inventory (VMI) have been known for several years. Nevertheless we have decided to include it in our survey in order to provide a comprehensive overview.

Both customers and the supplier can benefit from VMI. Customers no longer need to monitor their inventory levels and place orders. On the other hand an additional degree of freedom is introduced from the supplier’s point of view who becomes responsible for keeping customer inventories above a certain level depending on the contractual agreements with respect to stockout probabilities. This allows for more flexibility when it comes to upstream processes such as production and routing. IRPs also embed the routing problem solved by the supplier. The exact delivery days are chosen by the supplier, with the objective of minimizing routing costs, holding costs and stockouts costs.

7.1. Literature Review

We refer to Andersson et al. (2010) for a recent survey on industrial aspects emerging in the field of inventory-routing. The basic inventory problem is to determine an adequate order policy, ignoring potential consequences in terms of the routing aspect. A survey on inventory models can be found in Minner (2003) and Minner
and Transchel (2010). Further extensions include models based on the concepts of VMI, where order policies may be chosen by the distributor, rather than by the customers. VMI models with deterministic and stochastic demand have been investigated by Hemmelmayr et al. (2009b, 2010). Models in which routing is expected according to a full-truckload policy have been studied in Campbell and Savelsbergh (2004b) and in Cornillier et al. (2008a,b, 2009).

Less-than-truckload problems with periodic customer visits can be modeled as Periodic Vehicle Routing Problems (PVPRP). This case allows for more flexibility from the distributor's point of view, since visit patterns may be chosen from a candidate set. The problem has been studied in Cordeau et al. (1997) and further investigated by Francis et al. (2006) and Wen et al. (2010).

A realistic model formulation for IRP can be found in Archetti et al. (2007). For this problem, heuristics have been developed by Bertazzi et al. (2002) and Bertazzi et al. (2005). An MIP-based decomposition approach for a related problem was investigated by Campbell and Savelsbergh (2004a). Several variants of the inventory-routing problem along with rich real-world extensions have been tackled by Gaur and Fisher (2004), Oppen et al. (2010), Raa and Aghezzaf (2009), Savelsbergh and Song (2008), Abdelmaguid and Dessouky (2006), Archetti et al. (2011) and Aghezzaf et al. (2006). A stochastic extension of the IRP has been analyzed by Berman and Larson (2001), Kleywegt et al. (2002) and Kleywegt et al. (2004).

Table 5 provides an overview on selected publications on problems arising in inventory management and vendor managed inventory with routing related aspects.
Table 5: Selected literature on inventory and routing related problems

<table>
<thead>
<tr>
<th>Reference</th>
<th>Problem characteristics</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemmelmayr et al. (2009b)</td>
<td>periodic VRP, deterministic demand</td>
<td>1-step, VNS</td>
</tr>
<tr>
<td>Hemmelmayr et al. (2010)</td>
<td>periodic VRP, stochastic demand, evaluation of recourse actions</td>
<td>1-step, VNS</td>
</tr>
<tr>
<td>Campbell and Savelsbergh (2004b)</td>
<td>optimize volume delivered given fixed routing sequence</td>
<td>optimal policy</td>
</tr>
<tr>
<td>Cornillier et al. (2008b)</td>
<td>IRP, multiple compartments</td>
<td>iterative 2-step heuristic (determine visits heuristically, routing &amp; delivery volume exactly)</td>
</tr>
<tr>
<td>Archetti et al. (2007)</td>
<td>multi-period VRP, order-up-to level policy</td>
<td>1-step, B&amp;C</td>
</tr>
<tr>
<td>Campbell and Savelsbergh (2004a)</td>
<td>IRP</td>
<td>2-step heuristic (determine visits exactly, routing &amp; delivery volume heuristically)</td>
</tr>
<tr>
<td>Raa and Aghezzaf (2009)</td>
<td>cyclic VMI</td>
<td>2-step heuristic (heuristic for assignment &amp; delivery volume, insertion heuristic for resulting TSP)</td>
</tr>
<tr>
<td>Archetti et al. (2011)</td>
<td>high demand forces route duration to exceed duration of single period</td>
<td>2-step heuristic (MIP for inventory, LNS for routing), hybrid approach using SP</td>
</tr>
</tbody>
</table>


7.2. Mathematical Model

The following is a variation of the classical VRPTW formulation. Whereas the VRPTW focuses on a single period, the Inventory-Routing Problem (IRP) considers a multi-period time horizon, typically measured in terms of days. The model we present here deals with the repeated distribution of multiple products $p \in P$ from a single depot to a set of customers $c \in C$. Rather than by placing specific orders each customer $i$ is characterized by a given rate $q_{ip}$ by which product $p$ is consumed. At the beginning of the planning horizon each customer $i$ has an initial inventory level of $g_{ip}$ of product $p$. The inventory holding cost per customer, product and unit of time is given by $c_{ip}^H$. We assume that all products are available for distribution at the central depot in sufficient quantities.

To formulate the underlying IRP mathematically the following additional decision variables have to be defined. The quantity of product $p$ to be delivered to customer $i$ in time period $t$ by vehicle $k$ is equal to $v_{ipkt}$. Let $h_{ipt}$ be the stock level of product.
at customer \(i\) at the beginning of time period \(t\). The decision variables of the VRPTW are extended by an additional dimension \(t\), in order to model the time period when certain decisions (i.e. allocation \((y_{ikt})\), routing \((x_{ijkt})\), delay \((u_{it})\)) are implemented. The model itself can be described as follows:

\[
\sum_{i \in \mathcal{C}} \sum_{p \in \mathcal{P}} c_{ip}^{H} \sum_{t \in \mathcal{T}} h_{ipt} + \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} c_{ij} \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} x_{ijkt} + \sum_{i \in \mathcal{V}} \sum_{t \in \mathcal{T}} c_{i}^{L} \sum_{t \in \mathcal{T}} u_{it} \rightarrow \min \quad (7.1)
\]

- \(h_{ip1} = g_{ip} \quad \forall i \in \mathcal{C}, p \in \mathcal{P} \quad (7.2)\)
- \(h_{ipt} + \sum_{k \in \mathcal{K}} v_{ipkt} - q_{ip} = h_{i,p,t+1} \quad \forall i \in \mathcal{C}, p \in \mathcal{P}, t \in \mathcal{T}' \quad (7.3)\)
- \(h_{ipt} + \sum_{k \in \mathcal{K}} v_{ipkt} - q_{ip} = g_{ip} \quad \forall i \in \mathcal{C}, p \in \mathcal{P}, \text{ where } t = |\mathcal{T}| \quad (7.4)\)
- \(\sum_{i \in \mathcal{C}} \sum_{p \in \mathcal{P}} v_{ipkt} \leq Qy_{ikt} \quad \forall k \in \mathcal{K}, t \in \mathcal{T} \quad (7.5)\)
- \(\sum_{k \in \mathcal{K}} y_{ikt} \leq 1 \quad \forall i \in \mathcal{C}, t \in \mathcal{T} \quad (7.6)\)
- \(\sum_{k \in \mathcal{K}} y_{ikt} = m \quad \forall i \in \mathcal{C}, t \in \mathcal{T} \quad (7.7)\)
- \(\sum_{i \in \mathcal{V}} x_{ijkt} = y_{jkt} \quad \forall j \in \mathcal{V} \setminus \{0\}, k \in \mathcal{K}, t \in \mathcal{T} \quad (7.8)\)
- \(\sum_{j \in \mathcal{V}} x_{ijkt} = y_{ikt} \quad \forall i \in \mathcal{V} \setminus \{n + 1\}, k \in \mathcal{K}, t \in \mathcal{T} \quad (7.9)\)
- \(w_{ikt} + s_{i} + t_{ij} \leq w_{jkt} + M(1 - x_{ijkt}) \quad \forall i, j \in \mathcal{V}, k \in \mathcal{K}, t \in \mathcal{T} \quad (7.10)\)
- \(a_{i} \leq \sum_{k \in \mathcal{K}} w_{ikt} \leq b_{i} + u_{it} \quad \forall c \in \mathcal{C}, t \in \mathcal{T} \quad (7.11)\)
- \(a_{i} \leq w_{ikt} \leq b_{i} + u_{it} \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, i \in \{0, n + 1\} \quad (7.12)\)
- \(x_{ijkt} \in \{0, 1\} \quad \forall i, j \in \mathcal{V}, k \in \mathcal{K}, t \in \mathcal{T} \quad (7.13)\)
- \(y_{ikt} \in \{0, 1\} \quad \forall i \in \mathcal{V}, k \in \mathcal{K}, t \in \mathcal{T} \quad (7.14)\)
- \(w_{ikt} \geq 0 \quad \forall i \in \mathcal{V}, k \in \mathcal{K}, t \in \mathcal{T} \quad (7.15)\)
- \(u_{it} \geq 0 \quad \forall i \in \mathcal{V}, t \in \mathcal{T} \quad (7.16)\)
- \(h_{ipt} \geq 0 \quad \forall i \in \mathcal{C}, p \in \mathcal{P}, t \in \mathcal{T} \quad (7.17)\)
- \(v_{ipkt} \geq 0 \quad \forall i \in \mathcal{C}, p \in \mathcal{P}, k \in \mathcal{K}, t \in \mathcal{T}. \quad (7.18)\)
The objective (7.1) is to minimize distribution costs, as well as inventory holding costs, while avoiding any stockouts. Constraints (7.2) and (7.3) model the inventory balance equations for every product and customer. Constraints (7.4) ensure that the initial stock level is attained by the end of the planning horizon. Constraints (7.5) ensure that the vehicle capacities are not exceeded on any day \( t \) during the planning horizon. Constraints (7.7)–(7.12) are an adaptation of (2.4)–(2.9) to account for the multi-periodic nature of the underlying routing problem.

8. Multi-Echelon Distribution Logistics

In this section, dedicated to multi-echelon distribution systems, we take into account the fact that a delivery from a supplier to a customer may involve different stages and possibly different transportation modes. The underlying decision problem covers both short-haul and long-haul freight transportation. Items are typically consolidated at hub locations by relatively small vehicles. Long-distance freight transportation between designated hubs are then performed overnight. Once the destination hub is reached, items are unloaded and short-haul distribution takes place. Typically long-haul transportation operates on a fixed schedule. Given this schedule, short-haul transportation is planned both for collection and distribution. This approach, however, may lead to suboptimal decision since the three embedded sub-problems (short-haul collection, long-haul transportation, short-haul distribution) are interdependent. Postponing the departure time of long-haul freighters leads to more flexibility for the short-haul collection, but to less flexibility for the short-haul distribution from the destination hub.

Items are typically picked up in the afternoon and the long-haul transportation takes place overnight. As soon as the long-haul vehicle has reached its final destination, items are distributed. Customers are usually delivered during the morning. The same fleet of vehicles is used for performing pickup and delivery services. We assume that vehicles are available for performing their pickup tour in the afternoon after having finished their morning tours (delivering items that arrived in the early morning). When shipping items to or from customers the transshipment takes place via the closest hub. Such problems can arise in city logistics (Crainic et al., 2009).

8.1. Literature Review

More recent work related to our proposed model formulation can be found in Perboli et al. (2011b). These authors introduce the family of multi-echelon vehicle routing problems, where delivery from depots to customers is organized by routing
and consolidating freight through intermediate hubs. Heuristics have been proposed by Crainic et al. (2011). Gendron and Semet (2009) have further extended this problem. In addition to routing-related issues, location decisions also come into play. Further applications includes features such as cross-docking (Wen et al., 2009). In this case no long-haul transportation is executed, but vehicle movements need to be synchronized at transshipment facilities.

An overview on selected publications in the domain of multi-echelon distribution logistics is shown in Table 6.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Problem characteristics</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perboli et al. (2011b)</td>
<td>introduce problem, consolidate freight through intermediate hubs</td>
<td>2-step, linear relaxation, variable fixing</td>
</tr>
<tr>
<td>Crainic et al. (2011)</td>
<td>propose metaheuristic</td>
<td>2-step, multi-start LS for assignment, variable fixing</td>
</tr>
</tbody>
</table>

8.2. Mathematical Model

We now present a model formulation for the combined routing problem, considering both long-haul and short-haul routing operations simultaneously.

The underlying network differs slightly from that of the VRPTW defined in Section 2. The multi-echelon distribution model is defined on a network \( G = (\mathcal{V}, \mathcal{A}) \), where \( \mathcal{V} \) denotes the set of nodes associated with the set of \( H \) hubs \( \mathcal{H} = \{1, \ldots, n^H\} \), where the transshipment takes place, and a set \( \mathcal{C} = \{(n^H + 1), \ldots, (n^H + n)\} \) of customers. Dummy nodes \( \mathcal{H}^* = \{(n^H + n + 1), \ldots, (2(n^H) + n)\} \) for the hubs complete the set of nodes. Let \( \mathcal{C}^p_h \) denote the set of customers having items to be picked up and transshipped through hub \( h \). Similarly, let \( \mathcal{C}^d_h \) denote the set of customers receiving items that have originally been transshipped through hub \( h \). A fleet of \( m_h \) identical vehicles is located at hub \( h \in \mathcal{H} \). The set of vehicles located at hub \( h \) is denoted by \( \mathcal{K}_h = \{1, \ldots, m_h\} \). All transportation requests have to be performed on the very same day. Hence long-haul vehicles departing from hub \( h \) to \( l \) may only leave \( h \) once all items that have to be transported to hub \( l \) have been dropped off at hub \( h \). An additional service time of length \( s^p_h \) (\( s^d_h \)) has to be taken into account for the loading process from short- to long-haul vehicles (and vice versa). Customers may only be visited once for delivery and pickup. The distribution of items by means of vehicle \( k \in \mathcal{K}_l \) the next day may only start, once all items for all customers visited by vehicle \( k \) (probably coming from different hubs) have arrived.

There may be several transportation requests of items originating from the same
customer but not necessarily with the same destination or intermediate hub. Similarly, customers may receive several items. The total amount of transportation requests originating from (to be shipped to) customer $i$ is denoted by $q_i^P (q_i^D)$.

The model is an extension of the three-index vehicle flow formulation presented in Section 2. Basically, two VRPTWs starting from any hub have to be considered simultaneously, which model the short-haul transportation from customers to hubs and vice versa. Additionally, the timing of long-haul transportation on links between hubs needs to be considered. Binary variables $x_{ijk}^P (x_{ijk}^D)$ are equal to 1 if and only if vehicle $k$ traverses arc $(i, j)$ when picking up (delivering) items. Binary variables $y_{ik}^P (y_{ik}^D)$ are equal to 1 if and only if node $i$ is visited by vehicle $k$ during their afternoon (morning) tour. Similarly, time variables $w_{ik}^P (w_{ik}^D)$ and $u_i^P (u_i^D)$ are used to model the arrival time at nodes, as well as the delay at customer locations on the pickup or delivery routes. Additional decision variables $v_{hl}$ are used to model the time at which long-haul vehicles with destination hub $l$ leave hub $h$. Binary variables $z_{hk}^P (z_{hk}^D)$ are equal to 1 if and only if vehicle $k$ has picked up items that need to be transshipped to or from hub $h$.

The model can be formally described as follows:

$$\sum_{i,j \in V} \sum_{k \in K} c_{ij}^P x_{ijk}^P + \sum_{i \in V} c_i^L P u_i^P + \sum_{i,j \in V} \sum_{k \in K} c_{ij}^D x_{ijk}^D + \sum_{i \in V} c_i^L D u_i^D$$  \hspace{1cm} (8.1)

$$\sum_{i \in C} q_i^P y_{ik}^P \leq Q \hspace{1cm} \forall k \in K$$  \hspace{1cm} (8.2)

$$\sum_{k \in K} y_{ik}^P = 1 \hspace{1cm} \forall i \in C, \text{ where } q_i^P > 0$$  \hspace{1cm} (8.3)

$$\sum_{k \in K} y_{ik}^P = m_h \hspace{1cm} \forall h \in \{H, H^*\}$$  \hspace{1cm} (8.4)

$$\sum_{i \in V} x_{ijk}^P = y_{jk}^P \hspace{1cm} \forall j \in V \setminus H, k \in K$$  \hspace{1cm} (8.5)

$$\sum_{j \in V} x_{ijk}^P = y_{ik}^P \hspace{1cm} \forall i \in V \setminus H^*, k \in K$$  \hspace{1cm} (8.6)

$$w_{ik}^P + s_i^P + t_{ij} \leq w_{jk}^P + M(1 - x_{ijk}^P) \hspace{1cm} \forall i, j \in V, k \in K$$  \hspace{1cm} (8.7)

$$a_i^P \leq \sum_{k \in K} w_{ik}^P \leq b_i^P + u_i^P \hspace{1cm} \forall i \in C$$  \hspace{1cm} (8.8)

$$\sum_{i \in C} q_i^D y_{ik}^D \leq Q \hspace{1cm} \forall k \in K$$  \hspace{1cm} (8.9)
\[ \sum_{k \in K} y_{ik} = 1 \quad \forall i \in \mathcal{C}, \text{ where } q_i^D > 0 \quad (8.10) \]
\[ \sum_{k \in K} y_{hk} = m_h \quad \forall h \in \{ \mathcal{H}, \mathcal{H}^* \} \quad (8.11) \]
\[ \sum_{i \in V} x_{ijk} = y_{jk} \quad \forall j \in \mathcal{V} \setminus \mathcal{H}, k \in K \quad (8.12) \]
\[ \sum_{j \in V} x_{ijk} = y_{ik} \quad \forall i \in \mathcal{V} \setminus \mathcal{H}^*, k \in K \quad (8.13) \]
\[ w_{ik}^D + s_i^D + t_{ij} \leq w_{jk}^D + M(1 - x_{ijk}) \quad \forall i, j \in \mathcal{V}, k \in K \quad (8.14) \]
\[ a_i^D \leq \sum_{k \in K} w_{ik} \leq b_i^D + u_i \quad \forall i \in \mathcal{C} \quad (8.15) \]
\[ w_{hk}^P + d_h^P \leq v_{hl} + M(1 - z_{lk}^P) \quad \forall h \in \mathcal{H}^*, l \in \mathcal{H}, k \in \mathcal{K}_h \quad (8.16) \]
\[ v_{hl} + t_{hl} + d_l^D \leq w_{lk}^D + M(1 - z_{hk}^D) \quad \forall h \in \mathcal{H}^*, l \in \mathcal{H}, k \in \mathcal{K}_l \quad (8.17) \]
\[ \sum_{i \in \mathcal{C}^P} y_{ik} \leq Mz_{hk}^P \quad \forall k \in \mathcal{K}, h \in \mathcal{H} \quad (8.18) \]
\[ \sum_{i \in \mathcal{C}_h^P} y_{ik} \leq Mz_{hk}^D \quad \forall k \in \mathcal{K}, h \in \mathcal{H} \quad (8.19) \]
\[ v_{hl} \geq 0 \quad \forall h \in \mathcal{H}^*, l \in \mathcal{H} \quad (8.20) \]
\[ x_{ijk}, x_{ijk}^D \in \{0, 1\} \quad \forall i, j \in \mathcal{V}, k \in K \quad (8.21) \]
\[ y_{ik}, y_{ik}^D \in \{0, 1\} \quad \forall i \in \mathcal{V}, k \in K \quad (8.22) \]
\[ w_{ik}, w_{ik}^D \geq 0 \quad \forall i \in \mathcal{V}, k \in K \quad (8.23) \]
\[ u_i^P, u_i^D \geq 0 \quad \forall i \in \mathcal{V} \quad (8.24) \]
\[ z_{hk}^P, z_{hk}^D \in \{0, 1\} \quad \forall h \in \mathcal{H}, k \in \mathcal{K}_h \quad (8.25) \]

The objective (8.1) is to minimize total routing costs, including penalties for not respecting time windows. Constraints (8.2)–(8.8) model the VRPTW for the pickup routes. Note that vehicles may only leave from their home hub (8.4). Customer \( i \) can only be visited if there exists a corresponding transportation requests originating at \( i \) (8.3). Constraints (8.9)–(8.15) model the corresponding VRPTW for delivery routes. Constraints (8.16) ensure that long-haul vehicles may only leave hub \( h \) for destination \( l \) once all items to be transshipped via hub \( l \) have arrived and have been loaded. Similarly, Constraints (8.17) mean that short-haul vehicles may only depart from hub \( l \) once all items requested by customers on their route have arrived.
Constraints (8.18) and (8.19) link the corresponding binary indicators $z$ with the customers visited.

9. Conclusions

To the best of our knowledge this paper is the first to propose an overview on various combinatorial optimization problems arising in Supply Chain Management. Its objective was to propose a new family of integrative decision problems. Pure routing and transportation related models were extended to also capture other features of the supply chain, such as scheduling, lotsizing, batching, packing, inventory and intermodality. The embedded subproblems have traditionally been solved independently. However, because of their interdependent nature, a combined modeling and solution approach should be adopted to exploit their optimization potential. Due to the combinatorial complexity of the proposed decision problems there is a clear need for efficient metaheuristics to tackle them.

Acknowledgements

This research was partly supported by the Canadian Natural Sciences and Engineering Research Council under grant 39682-10. This support is gratefully acknowledged. Thanks are due to the referees for their valuable comments.
References


