Strategic and Tactical Scheduling of Logistics Assistants

Leveraging Flexibility in Shifts and Tasks

Using Column Generation

– An Opportunity for Hospital Logistics

Inaugural-Dissertation

zur Erlangung des Doktorgrades an der

Wirtschaftswissenschaftlichen Fakultät

der Universität Augsburg

vorgelegt von

Jonas Karl Christoph Volland

2017

Erstgutachter: Prof. Dr. Jens O. Brunner

Zweitgutachter: Prof. Dr. Robert Klein

Vorsitzender der mündlichen Prüfung: Prof. Dr. Axel Tuma

Tag der mündlichen Prüfung: 16. Januar 2017

Contents

List of Tables

List of Figures

1 Introduction

In the following chapter, we introduce healthcare operations management and frame the problem of our work before providing an outline of the structure of the thesis.

1.1 Healthcare operations management and problem definition

The cost of healthcare has outgrown economic growth in all industrialized countries in recent years (OECD Publishing 2015). This trend is expected to continue in the future, while an unfavorable demographic development might even increase healthcare cost as percent of the gross domestic product. On the one hand, an aging population leads to increasing demand for healthcare; on the other hand, the number of individuals financing the healthcare system is declining. Consequently, societies in general and healthcare providers in particular are faced with increasing economic pressure.

Across all member countries of the Organization for Economic Co-operation and Development (OECD), hospitals account for approximately one third of total healthcare expenditures (OECD Publishing 2013). The thesis at hand addresses this cost block, and in particular its two largest single cost items, personnel costs and logistics-related costs (Poulin 2003, Ross & Jayaraman 2009). It is supposed that costs can be reduced while maintaining or even improving the quality of patient care (Jarrett 1998).

Optimization of personnel costs in hospitals has been identified as one key cost containment lever. Academics and practitioners alike acknowledge the potential (Brucker et al. 2011). While sophisticated personnel scheduling techniques have been developed for a variety of industries, healthcare is clearly the largest field of application for personnel scheduling. Within healthcare, most attention has been paid to nurse scheduling (Van Den Bergh et al. 2013). Physician scheduling has received far less attention in the past, but the increasing number of publications in recent years signals growing interest from academia (Brunner et al. 2009, Erhard et al. 2016). In contrast, hospital logistics management has not been given high priority in the past, especially when considering the high relevance of logistics optimization in other industries, such as manufacturing. Nevertheless, research attention has grown considerably in recent years (see Chapter 2). One major research stream examines the application of sophisticated logistics concepts in hospitals.

The thesis at hand addresses the introduction of a sophisticated logistics concept in hospitals. The concept comprises the introduction of a new type of employee, referred to as *logistics assistant* to take over tasks unrelated to patient care and relieve medical staff of activities

outside their core business. We present integrated shift and task scheduling problems that may arise in the course of the introduction of the new employee type.

Shift scheduling typically comprises an iterative process that starts with demand modeling based on predefined task schedules or staffing level requirements that are initially derived from the patient mix and their needs in terms of care (see Chapter 4 and Ernst et al. 2004). Often, demand is determined based on historical data and fixed inputs. Previous research suggests that replacing fixed inputs with integrated, multiple decision-making holds promising optimization potential in personnel scheduling (Van Den Bergh et al. 2013).

In the work at hand, we integrate shift and task scheduling into one holistic model. We derive the resource supply in the shift scheduling part of our work and define the resource demand in the task scheduling part. In both areas, we assume a high degree of flexibility. On the resource supply side, we incorporate flexibility in the employees' shift schedules. Unlike traditional shift scheduling approaches, there are no explicit shift patterns to be scheduled. Instead, we rely on highly flexible implicit shift scheduling. On the demand side, we schedule tasks in order to minimize peaks and create demand patterns that fit well with the available resources. The work at hand addresses two optimization problems that occur in the course of the introduction of logistics assistants, namely the strategic workforce sizing problem and the tactical shift and task scheduling problem. For both optimization problems, we present optimal solution procedures, i.e., a column generation algorithm and a branch-and-price approach. Throughout this work, we aim to leverage the flexibility incorporated in both shift scheduling and task scheduling to find optimal solutions.

Our goal is to answer the following research questions with this thesis:

(1)   Which areas of hospital logistics management have been addressed by the previous literature and which areas are most promising for future cost optimization?

(2)   What is the optimal number of logistics assistants in hospitals when fully leveraging the flexibility incorporated in shift and task scheduling?

(3)   What is the impact of flexibility in shift and task scheduling on the optimal number of logistics assistants?

(4)   What are optimal shift and task schedules when workers are employed with different degrees of flexibility and at different costs?

1.2   Structure of the thesis

The thesis at hand comprises three major chapters, apart from the introduction (Chapter 1) and the conclusion (Chapter 5). Research question (1) is addressed in Chapter 2. It provides a literature review on material logistics in hospitals. We present and discuss publications along four research streams and work out future research directions. Particular focus is given to publications that apply quantitative methods. In Chapter 3, we address research questions (2) and (3). We introduce the integrated shift and task scheduling problem that arises when introducing logistics assistants. In particular, the chapter addresses the strategic workforce sizing problem. We present a column generation-based solution approach to solve the problem. We then address the subsequent tactical scheduling problem in Chapter 4 to address research question (4). There, we aim to define shift and task schedules for a given number of employees with varying degrees of flexibility and costs. A branch-and-price approach to solving the problem is presented. The final Chapter 5 provides the conclusion and presents overarching future research opportunities. In the following, we provide a more detailed overview of the three key chapters of this work.

1.2.1   Material logistics in hospitals: A literature review

Hospital logistics management has been identified as one key cost containment lever to cope with steadily increasing healthcare costs in industrialized countries. In the countries of the OECD, healthcare expenditures have grown at an average of 4% per year between 2000 and 2009 (OECD Publishing 2015). The purpose of Chapter 2 is to provide an overview of the state of the art of research on material logistics management in hospitals, whereby particular focus is given to publications that apply quantitative methods. The contribution of this chapter is threefold: First, we provide research guidance by categorizing the literature and identifying major research streams. Second, we discuss applied methodologies; and third, we identify future research directions. We take a systematic approach to identify the relevant literature from 1998 to 2014. Applicable publications are categorized thematically and methodologically, and future research opportunities are identified. In total, this work identifies and discusses 145 publications. The literature is categorized into four research streams, i.e.,

(1)   supply and procurement,

(2)   inventory management,

(3)   distribution and scheduling, and

(4)   holistic supply chain management.

The use of optimization techniques is steadily growing importance. The number of related publications has continually grown and has peaked over the last three years in scope. Optimization has been successfully applied in research streams (1), (2), and (3). Research stream (4), holistic supply chain management, comprises a rather qualitative research field of literature dealing with supply chain management issues.

This chapter is based on Volland, Fügener, Schoenfelder, et al. (2016)[1].

1.2.2   A column generation approach for strategic workforce sizing

In order to cope with steadily increasing healthcare costs, hospitals introduce a new type of employee, referred to as logistics assistants, to take over logistical tasks from specialized nurses. As described in the introduction, hospitals are faced with the task of dimensioning their number. We present a mixed-integer program that allows the optimal number of logistics assistants to be defined, given predefined task requirements. We combine flexible shift scheduling with a task scheduling problem, incorporating flexibility both in terms of shift and task scheduling in order to define the minimum number of workers. We present a column generation-based solution approach that finds optimal solutions, and compare decomposition approaches with one and two subproblems. Neither the general model nor the solution approach are limited to logistics assistants but can also be applied to other problem settings in the healthcare industry and beyond. The approach is tested with 48 problem instances in total and compared against benchmarks. As part of our solution approach, we present a new lower bound for staff minimization problems with an unknown number of available workers. We show that flexibility in shift and task scheduling can lead to a decrease of 40 to 49% of the required workforce, compared to the non-flexible case.

This chapter is based on Volland, Fügener & Brunner (2016)[2].

1.2.3   A branch-and-price approach for tactical shift and task scheduling

We present an integrated optimization model that simultaneously performs shift and task scheduling. Problems of this type occur, for example, in the healthcare industry when logistics assistants are introduced in hospitals to take over tasks unrelated to patient care from medical staff. In our model, we simultaneously schedule flexible shifts relying on an implicit formulation and tasks respecting start time windows and precedence constraints. The goal is to

---

[1] http://dx.doi.org/10.1016/j.omega.2016.08.004

[2] http://dx.doi.org/10.1016/j.ejor.2016.12.026

define shift and task schedules for a given number of employees. In this tactical optimization problem, we consider different worker categories that are employed under different shift parameters and work at different costs. We propose a branch-and-price solution approach that relies on two subproblem types, namely shift and task scheduling. In shift scheduling, we consider two worker categories, flexible and inflexible workers. As part of the presented solution approach, we introduce a new network flow formulation for the shift scheduling subproblem that is solved with a shortest path-labeling algorithm. We present a numerical study that demonstrates the superiority of the new subproblem formulation and illustrates the validity of the branch-and-price approach.

2 Material logistics in hospitals: A literature review

Chapter 2 introduces material logistics management in hospitals by providing a literature review on previous publications in this field of research.

2.1 Introduction

In the countries of the OECD, total healthcare expenditures have grown at an average of 4% per year from 2000 to 2009 (OECD Publishing 2015), with hospitals accounting for 29% of total healthcare expenditures (OECD Publishing 2013). Of hospital costs, more than 30% are linked to logistics activities (Nachtmann & Pohl 2009). This makes logistics costs the second largest cost block after personnel costs (Poulin 2003, Ross & Jayaraman 2009). Material management and logistics have not been given high priority in hospital management research in the past compared to other industries. Possible reasons are the high complexity of healthcare supply chains and their merely supporting role in the foremost objective of hospital management, i.e., effective treatment of patients (Beier 1995). However, in the last 15 to 20 years, logistics has been identified as one key lever in managing healthcare costs (Dacosta-Claro 2002, De Vries 2011). The research suggests that efficient logistics management can eliminate around half of the logistics-related costs in hospitals (Poulin 2003).

The potential of hospital logistics optimization within the healthcare sector is considered significant by academics and practitioners alike. The most obvious upside from optimizing material logistics is that cost reductions do not directly affect the quality of patient care (Jarrett 1998). Currently, logistics-related activities are often performed by medical staff, taking away time from patient care activities. In a recent survey among registered nurses in the U.S., time wasted on activities other than patient care, such as restocking supplies, was the major driver that negatively impacted nurses' time at the bedside (Jackson Healthcare 2014). Relieving nurses from activities not related to patient care can thus improve the quality of care.

The aim of this chapter is to present the state of the art of research on material logistics management in hospitals. In the discussion, we set a distinct focus on publications that apply quantitative methods. Relevant papers are discussed in detail, e.g., by providing tables with deep-dive analyses on the applied methodologies. Our contribution is threefold: First, we provide research guidance by categorizing the literature and identifying major research streams. Second, we discuss applied methodologies; and third, we identify future research directions. There exist rather general literature reviews on healthcare operations research and operations management (e.g., Rais & Viana 2011, Hulshof et al. 2012, Fakhimi & Probert 2013). Additionally, there are a number of reviews on supply chain management (SCM) in healthcare, e.g., De Vries & Huijsman (2011), that focus on the question of whether or not there are

parallels between the industrial sector and healthcare services. Dobrzykowski et al. (2014) thematically assess a more general scope than this work, as they include operations management topics like service management, planning, and scheduling. Furthermore, they limit their review to publications from only seven U.S. journals and review a different time period (1982 to 2011). Consequently, to the best of our knowledge, there is currently no comprehensive review of material logistics in hospitals with a focus on quantitative methods. This chapter fills that research gap.

The remainder of this chapter is structured as follows: Section 2.2 presents the methodology of the literature review and introduces a framework to cluster the relevant literature thematically. Section 2.3 provides a quantitative overview of topics, applied methodologies, and the regional coverage of assessed publications, as well as an overview of all publications. Sections 2.4 to 2.7 discuss the literature along this framework and point out future research potential. We present a conclusion and a summary of research opportunities in the final section of this chapter.

2.2   Methodology

Scope. This chapter reviews all relevant publications regarding the logistical activities involved in handling physical goods in hospitals. Physical goods comprise all items that are directly linked to patient care, such as pharmaceuticals, medical consumables, food, laundry, sterile items, laboratory samples, and waste, etc. Pharmaceuticals represent 70% to 80% of the supply costs, while medical-surgical materials account for 20% to 25% (Rego et al. 2013). Products unrelated to the care of patients, such as office supplies or mail, are excluded. Further, although partly included in logistics activities, flow of information is excluded. Due to its distinct characteristics, such as the irregularity of supply and the lack of comparability with the items mentioned above, blood products are out of scope of this review. Comprehensive reviews on SCM of blood products are available in the literature (Beliën & Forcé 2012, Rossetti et al. 2012). Considering the supply chain of goods from manufacturing to use, this review starts with the supply chain partners one step upstream of the hospital, i.e., typically the hospital-supplier interface. One exception is Subsection 2.4.4, where we shed light on the interface between drug manufacturers and wholesalers and the implications for hospital purchasing. Also, reverse logistics are not in particular scope of this chapter; however, we refer to Srivastava (2008) for designing a reverse logistics network. Logistics activities associated with outpatient treatment, like home delivery of meals or outpatient medication, are out of scope. As an example, Liu et al. (2014) present related work. Our restriction of the scope is in line with the existing literature, as hospital-internal logistics activities are the major source of competitiveness within healthcare material management (Rivard-Royer et al. 2002, Landry & Philippe 2004). Personnel planning

and scheduling that is not directly related to logistics activities, as well as bed and patient transportation, are out of scope of this chapter.

Identification of publications. In order to identify the relevant literature, we undertook a five-step approach. First, we searched Google Scholar and Science Direct for relevant keywords, e.g., "hospital" and "logistics". Second, we performed a forward and backward search of the most relevant publications. Third, we developed the categorization framework presented below and classified the literature accordingly. Fourth, we performed another Google Scholar and Science Direct search applying relevant key words within the respective category. Fifth, we concluded the search process with a final forward and backward search within those publications. We limit our research to English articles published in peer-reviewed journals. Books, theses, PhD dissertations, conference articles, and working papers are neglected. Our focus is on publications after the year 1998 until year-end 2014. Of the papers published earlier, only the most often cited ones are included. A review of previous work is included in Jarrett (1998).

Literature classification framework. The literature is thematically classified along the framework in Figure 2.1. We identify four major research topics in the literature. Categories (1) to (3) comprise the supply chain that material follows before being used in hospitals. (1) Supply and procurement contains the literature regarding the purchasing of material as well as all activities related to the hospital-supplier interaction, for example outsourcing and means of supplier collaboration. Furthermore, the literature on demand forecasting is presented in this section. (2) Inventory management includes the literature on inventory policy, location planning, as well as classification schemes and practice-oriented inventory publications. Drug inventory management and drug shortages are also discussed. (3) Distribution and scheduling covers all material-linked distribution activities within and outside the hospital. In research topic (3), we focus on the actual transportation or distribution rather than the location of the goods, as well as on the handling of sterile medical devices. (4) Holistic supply chain management takes a comprehensive and mostly qualitative approach to optimizing the supply chain.

All four research topics including their subtopics are presented and discussed. The focus is on areas where quantitative methods are applied; however, for the sake of completeness and to provide insights on related research directions, the remaining research areas are also presented. The literature search yields 145 publications that are categorized along the presented framework. Where multiple categories are addressed, the publication is assigned to the most relevant category. Six methodological categories are distinguished: optimization (containing an operations research (OR) model), simulation/scenario analysis, empirical research, literature

review, theory/conceptual (introducing or discussing a new theory or concept), and case study (findings from practical research projects, etc.).
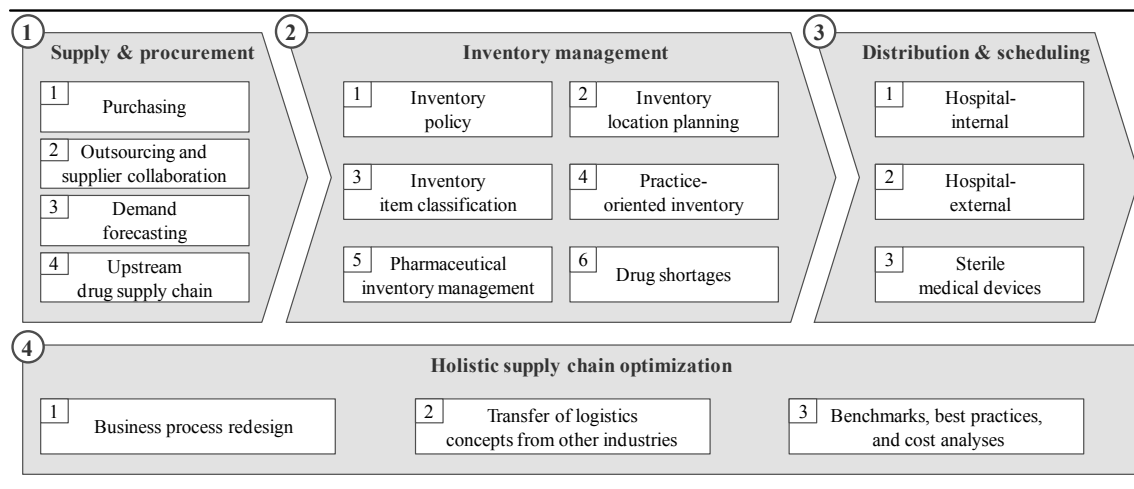
Figure 2.1: Literature classification framework

2.3   Publication meta-analyses

The subsequent section provides a quantitative assessment of the identified publications. Publications in this section are assessed from three perspectives: thematic, methodological, and regional. Thereafter, an overview is provided and research opportunities are identified.

Thematic categorization. We generally limit the scope to publications published from 1998 onwards. However, eight earlier publications are included due to their pivotal importance for the relevant literature streams. The development of publications over time and the thematic scope along our review framework is provided in Figure 2.2. The different shadings reflect the allocation to our four main categories.
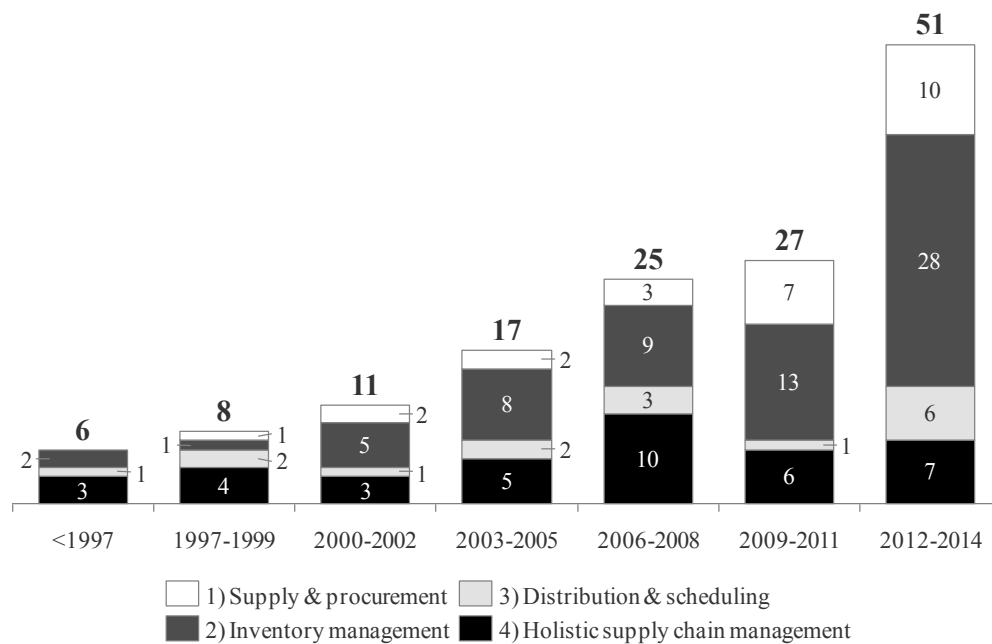
Figure 2.2: Thematic categorization

The number of publications has grown considerably over time. The increasing relevance of hospital material logistics in academia is indicated by the fact that the number of publications nearly doubled between 2009 to 2011 and 2012 to 2014 . The categories with the highest growth rates over the last years are (3) distribution and scheduling and (2) inventory management. These two categories combined account for an increase in publications from 14 between 2009 and 2011 to 34 between 2012 and 2014. Over the entire period, the majority of publications, 66 papers, were published in (2) inventory management.

Methodological categorization. The methodologies applied in the reviewed publications are presented in Figure 2.3. In the chart, the color of the shapes indicates the quantitative nature of the applied methodologies. Over the period, the largest number of papers were published in the field of case studies, with 47 publications, and theory/conceptual, with 32 publications. The category of optimization experienced a large increase in publications from 4 in 2009 to 2011 to 15 in 2012 to 2014. This indicates a further evolving interest in the field of operations research on hospital materials management. The second quantitatively-focused category, simulation/scenario analysis, also peaks in the latter time segment, underlining the importance of quantitative research in hospital materials management.
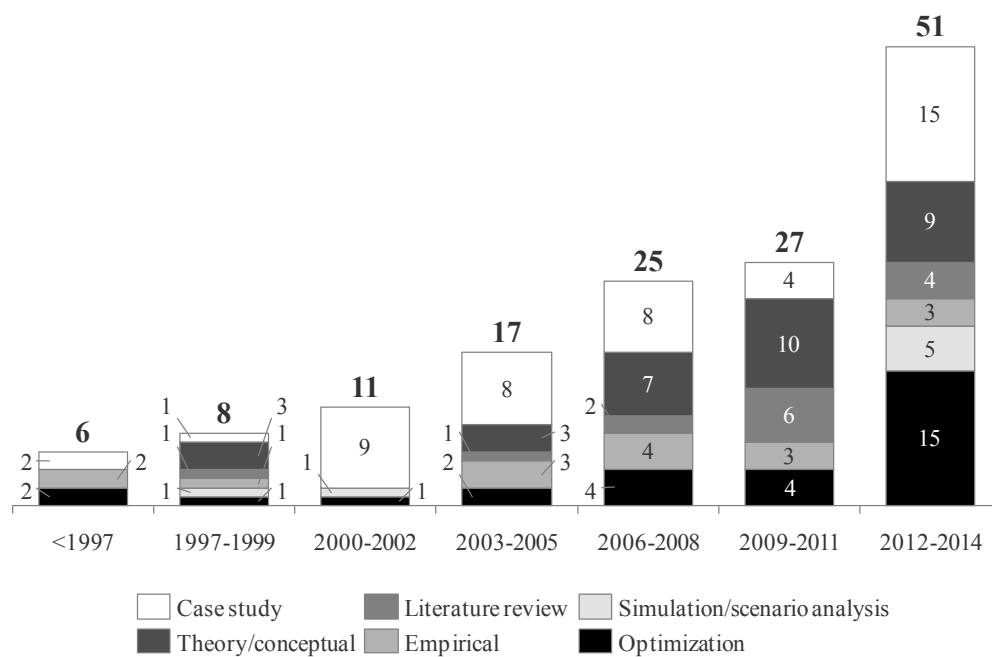
Figure 2.3: Methodological categorization

Regional categorization. This paragraph presents the regional scope of the examined publications. This is to be understood as the location the research relates to, e.g., where the relevant case hospital is located. However, the location of the research institution is not relevant for this overview. Figure 2.4 presents the regional scope for the entire time span. N/A indicates either that a publication has a global scope, e.g., literature reviews or methodology-focused publications, or that it was simply not possible to identify the location the publication relates to. The filling indicates the regional allocation of publications. North America and Europe are the continents with the most publications during the entire period with 65 and 37 articles, respectively. Combined, these account for more than two thirds of total publications. Within North America, the U.S. accounts for 54 publications, making it the country with the most publications worldwide. Within Europe, the U.K. and the Netherlands lead, with nine and six publications, respectively. Due to the strongly regulated nature of national healthcare systems and their significant regional differences, it is of pivotal importance to perform research with a clear regional focus, considering the specifics of national legislation. Our review considers only English publications, which might emphasize publications from English-speaking countries and thus slightly bias the displayed regional coverage.
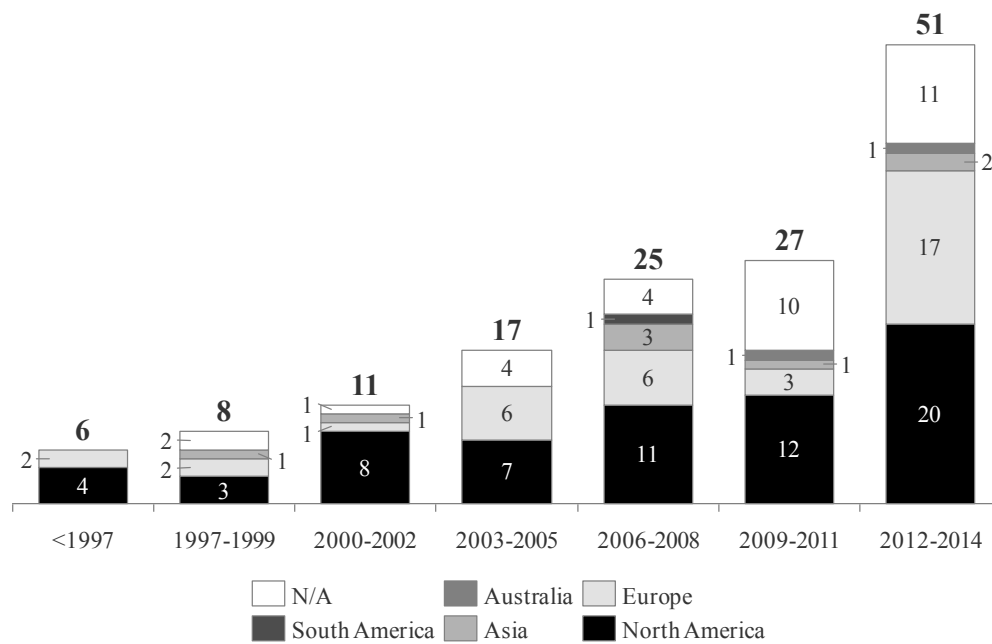
Figure 2.4: Regional coverage of publications

Overview of publications. A thematic and methodological overview of all publications is presented in Table 2.1. We determine that optimization techniques are mostly applied in (2) inventory management and (3) distribution and scheduling. Both research areas are characterized by rather well-defined problem settings that also occur in other industries, such as manufacturing. (1) Supply and procurement also offers possibilities to apply optimization techniques, but only five publications exist in this field of research. In (4) holistic supply chain management, however, no publications exist that apply optimization techniques. This may be due to the high complexity of integrated processes and the hard-to-define nature of holistic supply chain problems. Instead, the applied research is mostly qualitative research, namely theoretical/conceptual research or case studies. Another interesting insight is that most literature reviews are published within this field. This underlines the fact that this review fills a research gap, as it focuses on publications applying quantitative techniques. Another apparent research gap lies within the application of case studies and conceptual research in the field of (3) distribution and scheduling.

Table 2.1: Overview of publications

	(1) Supply & procurement	(2) Inventory management	(3) Distribution & scheduling	(4) Holistic supply chain management
Optimization	Hu & Schwarz (2011) Iacocca et al. (2013) Rego et al. (2013) Ross & Jayaraman (2009) Zhao et al. (2012)	Baboli et al. (2011) Bijvank & Vis (2012) Dellaert & Dan De Poel (1996) Guerrero et al. (2013) Kelle et al. (2012) Little & Coughlan (2008) Nicholson et al. (2004) Priyan & Uthayakumar (2014) Rosales et al. (2014) Rosales et al. (2015) Uthayakumar & Priyan (2013) Vila-Parrish et al. (2012)	Augusto & Xie (2009) Bailey et al. (2013) Banerjea-Brodeur et al. (1998) Kergosien et al. (2013) Lapierre & Ruiz (2007) Medaglia et al. (2008) Michelon et al. (1994) Ozturk et al. (2014) Shih & Chang (2001) Swaminathan (2003) Tlahig et al. (2013) Van De Klundert et al. (2007)	
Simulation/ scenario analysis	Azzi et al. (2013)	Gebicki et al. (2014) Pasin et al. (2002)	Dean et al. (1999) Di Mascolo & Gouin (2013)	Iannone et al. (2013) Iannone et al. (2014)
Empirical	Burns & Lee (2008) Nollet & Beaulieu (2003) Oumlil & Williams (2011) S. Kumar et al. (2008)	Baumer et al. (2004) Beier (1995) De Vries (2011) Fox & Tyler (2003) Hall et al. (2013) Huarng (1998) Kaakeh et al. (2011) Lee & Shim (2007) McBride et al. (2013)		Kafetzidakis & Mihiotis (2012) Kim & Schniederjans (1993) Zheng et al. (2006)
Literature review	Jack & Powers (2009)	Coustasse et al. (2013) Fosso Wamba et al. (2013) Irani et al. (2010) James et al. (2009)	Beliën et al. (2012)	Abukhousa et al. (2014) De Souza (2009) De Vries & Huijsman (2011) Ford & Scanlon (2007) Jarrett (2006) Jun et al. (1999) Mazzocato et al. (2010) Young et al. (2004)
Theory/ concept	Brennan (1998) Cruz & Marques (2013) Fein (2007) Van Donk (2003)	Almarsdóttir & Traulsen (2005) Alspach (2012) Danas et al. (2006) Fox et al. (2009) Gu et al. (2011) Gupta & Huang (2013) Huys & Simoens (2013) Johnson (2011) Le et al. (2011) Lee & Özer (2007) Mangan & Powers (2011) Mazer-Amirshahi et al. (2014) McLaughlin et al. (2013) McLaughlin et al. (2014) Meiller et al. (2011) Phillips & Berner (2004) Rider et al. (2013) Vaillancourt (2012) Ventola (2011)		Chandra (2008) Di Martinelly et al. (2009) Jarrett (1998) Kim et al. (2006) Kollberg et al. (2007) Landry & Beaulieu (2010) Meijboom et al. (2011) Whitson (1997) Young & McClean (2008)
Case study	Bendavid et al. (2010) Bendavid et al. (2012) Bhakoo & Chan (2011) Bhakoo et al. (2012) Guimarães et al. (2013) Haavik (2000) Matopoulos & Michailidou (2013) Mustaffa & Potter (2009) Rivard-Royer et al. (2002) Varghese et al. (2012)	Abijith & Fosso Wamba (2012) Beso et al. (2005) Çakıcı et al. (2011) Chan et al. (2012) Chang et al. (2012) Danas et al. (2002) Fitzpatrick et al. (2005) Franklin et al. (2008) Granlund & Wiktorsson (2013) Griffith et al. (2012) Gupta et al. (2007) Khurana et al. (2013) Koppel et al. (2008) Maviglia et al. (2007) Novek (2000) Patterson et al. (2002) Piccinini et al. (2013) Poley et al. (2004) Poon et al. (2006) Thomas et al. (2000)	Mühlich et al. (2003)	A. Kumar et al. (2008) Aptel & Pourjalali (2001) Born & Marino (1995) Coulson-Thomas (1997) Dacosta-Claro (2002) Ferretti et al. (2014) Fillingham (2007) Heinbuch (1995) Jayaraman et al. (2000) Kriegel et al. (2013) Landry & Philippe (2004) McKone-Sweet et al. (2005) Pan & Pokharel (2007) Trägårdh & Lindberg (2004) Venkateswaran et al. (2013) Yasin et al. (2003)

2.4 Literature on topic (1) supply and procurement

A natural approach to minimizing material costs is to reduce the actual purchasing costs. Supply chain costs constitute a major part of hospitals' operating expenses. For example, U.S. hospitals spent $27.7B on drugs alone in 2009 (Doloresco et al. 2011). Hospital material management literature focuses mainly on four areas. The first is bundling purchasing volumes to increase purchasing power. This stream of literature chiefly discusses hospital group purchasing organizations (GPOs). Hospitals strive to increase their purchasing power against suppliers by combining their respective purchasing volumes. The second stream of literature primarily discusses hospital inventory outsourcing approaches, e.g., stockless inventory systems or vendor-managed inventory (VMI), with regard to supplier integration. The third subsection offers an overview of demand forecasting, which is of high relevance for the hospital-supplier interface, while the last subsection sheds light on specifics of the upstream pharmaceutical supply chain and potential implications for hospital buyers. All four streams are presented below.

2.4.1 Purchasing

Hospitals are mostly organized in GPOs, i.e., voluntary alliances aggregating hospitals' purchasing volumes. In the U.S., 90% to 98% of hospitals are organized in purchasing alliances (Burns & Lee 2008). GPOs help to reduce material costs in two ways. First, they allow economies of scale to be leveraged due to purchasing volume bundling. Second, they enhance price transparency and create price ceilings through framework contracts in which agreed-upon price bands (Burns & Lee 2008). A qualitative comparison of advantages and disadvantages of GPOs is provided by Burns & Lee (2008) and Rego et al. (2013). In line with the remainder of this chapter, publications containing optimization models are displayed separately (see Table 2.2) and discussed in depth.

Rego et al. (2013) present a decision support tool helping hospital purchasing managers identify and assess alternative GPO forms. For a defined group of hospitals willing to cooperate, the tool presents the number, size and composition of GPOs, and a financial assessment. A metaheuristic comprised of a two-module hybrid variable neighborhood search (VNS) and tabu search is applied to solve the optimization problem. The tool allows alternative cooperative purchasing strategies to be evaluated and is applicable to a wide range of purchasing groups. Ross & Jayaraman (2009) focus on the single-hospital level. They assess how products should be bundled when placing orders at suppliers. They focus particularly on bundling new products with refurbished products, an option several U.S. healthcare providers have recently started to explore in order to reduce material costs. Examples of refurbished products include investment

goods such as medical devices or electric beds, which are bundled with (new) consumable products. The authors develop a mixed-integer program (MIP) aiming to minimize the total purchasing costs. They build a heuristic based on simulated annealing (SA) to find near-optimal purchasing strategies, i.e., which products to buy from which supplier, and decide whether to conduct a bundled or single-item purchase. Potential item surpluses in bundles exceeding the buyer's requirements are minimized in the objective function (apart from purchasing costs). Hu & Schwarz (2011) assess the general role of GPOs in the healthcare supply chain and their impact on pricing mechanisms with a Hotelling duopoly model. They find that GPOs indeed achieve lower prices for healthcare providers through increased competition among manufacturers. However, they also point out downsides of GPOs like reducing incentives for manufacturers to innovate and enhance their existing product portfolio.

Table 2.2: Purchasing publications containing optimization models

Publication	Problem description	Model characteristics	Objective function	Type of goods
Hu & Schwarz (2011)	Role of GPOs in healthcare supply chain and cost reduction potential	Hotelling duopoly model	Minimize costs	Not specified
Rego et al. (2013)	Decision support to set number of GPOs, size and composition for hospitals willing to cooperate	Metaheuristic (hybrid VNS/ tabu-search)	Minimize hospitals' shared supply chain costs	Not specified
Ross & Jayaraman (2009)	Purchasing strategy for bundled refurbished medical products	Mixed-integer problem solved with SA heuristic	Minimize total acquisition costs of purchasing plan	Durable items (new & refurbished)

Non-optimization-focused publications include the following: Oumlil & Williams (2011) discuss strategic purchasing alliances in the healthcare sector in terms of both organizational and personal factors. Organizational factors include the hospitals' type and size, while personal factors comprise, e.g., the education and experience levels of purchasing managers. They find that selected demographic characteristics of purchasing managers are linked to decisions on alliances. For instance, job experience and the success of the alliance are related. Burns & Lee (2008) provide an empirical study on the utilization, services, and performance of hospital purchasing alliances from the hospital material management's point of view. The authors find that purchasing alliances achieve cost reduction by lowering product purchasing prices, especially for commodity and pharmaceutical products. They stress that alliances further reduce transaction costs, as contracts are jointly negotiated. However, cost-benefit realization is hindered for service products or when physicians prefer certain items. Burns & Lee (2008) further present several literature streams – not necessarily related to operations management – in the field of purchasing alliances, such as pooling alliances and value chain alliances.

Another empirical study states that purchasing groups are subject to lifecycle stages (Nollet & Beaulieu 2003). The authors identify critical characteristics influencing the development of

purchasing groups. These include payers' intervention, the nature of benefits, procurement strategy, the nature of relationships with suppliers, structure, and resources. They further develop a lifecycle model to show the evolution of GPOs and the changing importance of these characteristics.

Regarding future research opportunities in terms of methodology, Ross & Jayaraman (2009) underline the combinatorial complexity of practical problems in healthcare logistics, marketing, and purchasing. They propose the development of heuristics in order to cope with large problem instances. Generally, there seems to be a bias in this research stream towards manufacturing industries; thus, healthcare in particular provides further research opportunities (Oumlil & Williams 2011). Other potential research fields include the assessment of performance determinants of GPOs to facilitate comparisons across purchasing alliance characteristics, e.g., in terms of size or management. Such research would furthermore allow identification of the potential to differentiate between GPOs (Burns & Lee 2008). Also, assessing outsourcing activities compared to GPOs seems worthwhile for future research. Past publications indicate that the attractiveness of outsourcing logistics is positively correlated to hospital size (Oumlil & Williams 2011).

2.4.2 Outsourcing and supplier collaboration

It is widely accepted that outsourcing logistics activities to third-party providers can generate significant efficiency advantages for both parties due to economies of scale and scope, reduction of fixed costs, and a focus on core competencies (Azzi et al. 2013, Iannone et al. 2014). There are many general studies on logistics outsourcing, but literature concerning healthcare is rather scarce, which is in line with the overall tendency in the healthcare sector to embrace new SCM practices only slowly (McKone-Sweet et al. 2005). However, outsourcing inventory decisions to healthcare providers has recently gained importance, especially in practice, where outsourcing concepts are widely applied (Nicholson et al. 2004). Kim (2005) stresses the potential of VMI in the healthcare sector. The author finds that hospitals can significantly reduce inventory stock. However, he states that supply chain integration might be hindered by the absence of standards for information sharing and a lack of participation of pharmaceutical manufacturers in collaborations.

In order to assess outsourcing opportunities, scenario modeling is applied in several publications. Azzi et al. (2013) consider different outsourcing options for a healthcare network in central Italy, comprising several hospitals and one centralized logistics hub. The authors evaluate three scenarios with varying outsourcing degrees both qualitatively and quantitatively: logistics self-management, partial logistics outsourcing, and total logistics outsourcing. The

qualitative assessment is mainly based on an extensive literature review, while the quantitative assessment of outsourcing options is performed using a system dynamics simulation. The authors state that logistics outsourcing is often the most economical option for different sets of distribution network layouts. Van Donk (2003) develops a tool to assess several potential supply chain designs between a hospital and its supplier of medical and non-medical gases. Nicholson et al. (2004) compare inventory costs of an in-house three-echelon distribution network versus an outsourced two-echelon distribution network (i.e., direct delivery to the care unit) for non-critical medical items. For a detailed analysis of this paper, see Subsection 2.5.1.

There is a large stream of literature presenting and discussing case studies without quantitative methods or simulation/scenario analyses. One example that is thematically linked to the previously discussed publication is the work by Rivard-Royer et al. (2002). They present a case study in a Canadian hospital that applies a hybrid stockless inventory management system. Hybrid means that suppliers have two options to deliver to the hospital: Either they supply goods to the hospital's central warehouse, which is the traditional approach, or they pack products according to the needs of the respective care unit and deliver direct. The authors find that the hybrid model may yield marginal benefits compared to the traditional approach. They also show that different forms of packaging are a significant source of cost savings. This packaging issue is analyzed in more detail in the publication of S. Kumar et al. (2008). The authors empirically assess whether package design plays a significant role in hospitals' purchasing decision-making processes. They find that packaging and environmentally friendly supplies currently do not play a pivotal role in purchasing decisions in the U.S. Further case studies regarding VMI concepts in the hospital setting are presented in Mustaffa & Potter (2009), Bhakoo et al. (2012), Guimarães et al. (2013), and Matopoulos & Michailidou (2013). All publications provide a good overview of the overall concept as well as its application in the hospital setting. Bhakoo et al. (2012) state that VMI has been widely ignored in the healthcare industry. They qualitatively assess different collaborative arrangements between hospitals and pharmaceutical suppliers, such as the "ward box", a variant of VMI where hospitals place direct orders for required items in a specific ward and the suppliers deliver to the ward without taking the detour to a central warehouse. Remarkably, they find that hospital material managers are more willing to undertake collaborative arrangements along the supply chain than their suppliers. Guimarães et al. (2013) present an assessment of VMI with regard to its benefits, risks, barriers, and enablers. They further conduct a case study of a multi-location hospital that aims to create transparency along its value chain. Matopoulos & Michailidou (2013) study the application of co-managed inventory (CMI), a form of VMI where hospitals remain partly responsible for inventory. The authors present a case study for a Greek hospital. Mustaffa & Potter (2009) assess a private hospital in Malaysia and its supplier relations and identify two

issues: urgent orders and stock availability at the wholesaler. Based on their findings, they propose the introduction of a VMI setup in order to cope with these difficulties. Bhakoo & Chan (2011) summarize complexity factors around pharmaceutical healthcare supply chains and present factors that hinder the implementation of e-business processes in the procurement area of healthcare supply chains: lack of consistency, poor data quality, and the global nature of supply.

Comparable to VMI approaches are consignment agreements in the hospital sector, where ownership of goods remains with the suppliers until they are consumed. This approach is mostly applied to expensive items, such as implants (Epstein & Dexter 2000). Compared to VMI approaches, recent literature on consignment agreements is rather scarce and focuses on case studies (Bendavid et al. 2010, Bendavid et al. 2012). The authors present an RFID-based traceability system for consignment and high-value products. Compared to other systems in the market, such as RFID-enabled cabinets or smart shelves, the system is rather simple and has lower technological requirements.

One obvious future research area is the extension of the outsourcing degree from VMI, where suppliers take over responsibility for hospitals' inventories, towards just-in-time (JIT) delivery. JIT means that suppliers provide goods to point-of-use locations in hospitals without intermediate buffer inventories. There seem to be some obstacles to the implementation of JIT concepts in hospital supply, however. Identifying the underlying reasons and providing ideas on how to overcome those difficulties could be an interesting future research field. For a continued discussion on the general applicability of JIT, please refer to Subsection 2.7.2. Also, as mentioned above, the availability of information across the supply chain might obstruct the applicability of more integrated supply chain concepts. Identifying ways to increase data transparency while respecting intellectual property rights and legal constraints might hold future research opportunities. Another potential research area could be the application of optimization techniques in outsourcing. So far, optimization models have not been applied in this research field. Potential questions include determining the characteristics of products that could be outsourced or defining the optimal degree of outsourcing, i.e., determining which product categories are appropriate for outsourcing.

2.4.3   Demand forecasting

One major obstacle for a better integration of hospitals and their suppliers is the unpredictable nature of hospital demand. Numerous researchers argue that the patient mix and the resulting demand for materials is very hard or impossible to predict (Haijema et al. 2007, Little & Coughlan 2008, Bailey et al. 2013, Cruz & Marques 2013, Hof et al. 2015). However, as

contracts with suppliers occasionally build on minimum purchase quantities, accurate demand forecasting is of high relevance for hospital purchasing managers. Brennan (1998) stresses the importance of regular demand forecasts based on clinical guidelines linking patient groupings' requirements with the resulting materials demand. To tackle unreliable resource demand predictability, Varghese et al. (2012) apply demand forecasting algorithms. Haavik (2000) stresses the importance of sharing hospital demand information with suppliers, e.g., by implementing VMI software able to forecast demand and placing orders with suppliers accordingly. Danas et al. (2002) reduce demand uncertainty by bundling several point-of-use inventory locations to one large virtual inventory. For further reading, we refer to Jack & Powers (2009), who provide a literature review on demand management and capacity management in healthcare services, and to Narayana et al. (2014), who investigate the redesign of the pharmaceuticals supply chain, not limited to hospitals. Improving forecasting mechanisms for hospital demand seems to hold worthwhile future research opportunities.

2.4.4   Upstream supply chain

In this subsection, we offer a brief overview of changes to the upstream drug supply chain, i.e., the interface between drug manufacturers and wholesalers. We specifically point out one aspect that has an impact on hospital pharmacies.

Starting in the year 2005, the payment and distribution scheme of the U.S. pharmaceutical supply chain went through a significant transition. Drug manufacturers and wholesalers changed their collaboration model from a buy-and-hold (BNH) to fee-for-service (FFS) system (Iacocca et al. 2013). In the BNH scheme, one of the wholesalers' major revenue sources was to speculate on drug price increases. When wholesalers held high stock levels and manufacturers increased their prices, wholesalers would pass the higher price on to their buyers, namely hospitals pharmacies. Apart from high stock levels, this scheme resulted in several other disadvantages, such as considerable fluctuation in wholesalers' order quantities, revenue losses for drug manufacturers, and unstable and unpredictable wholesaler revenues (Zhao et al. 2012). In the FFS scheme, however, the wholesalers agree to reduce or eliminate drug investment buying in return for fees paid by drug manufacturers to hold inventory and fulfill their distribution role (Iacocca et al. 2013). According to Fein (2007), the FFS scheme comes with two major threats for hospitals: First, the wholesalers' discount range is reduced as they share detailed order, inventory, and shipment data with drug manufacturers. This reduces their volume buying potential and, consequently, their discount range. Second, as inventory levels at wholesalers are reduced, the threat of drug shortages is significantly higher in the new scheme (see Subsection 2.5.6).

Table 2.3: Upstream supply chain publications containing optimization models

Publication	Problem description	Model characteristics	Objective function	Type of goods
Iacocca et al. (2013)	Comparison of effectiveness of buy-and-hold, fee-for-service, and direct-to-pharmacy contracts	Mathematical programming	Maximize profits	Pharmaceut.
Zhao et al. (2012)	Comparison of investment-buying and fee-for-service setup, definition of optimal policies and policy parameters	Multiperiod, stochasitc inventory problem	Maximize profits	Pharmaceut.

In this research area, two publications might be of interest for hospital pharmacy buyers (see Table 2.3). Zhao et al. (2012) investigate the design and benefits of FFS contracts and derive implications for inventory policies and their parameters for drug manufacturers and wholesalers. Iacocca et al. (2013) compare the differences between the two schemes and a third payment and distribution scheme, the direct-to-pharmacy (DTP) agreement, where wholesalers manage drug distribution and inventory for a fee, but the manufacturers remain the owner of the drug until it reaches the point of use. As the focus of the presented publications lies mostly on the manufacturer-wholesaler interface, we believe that future research should focus on the explicit implications for hospital buyers.

2.5 Literature on topic (2) inventory management

While the management of inventory systems has been widely discussed in the industrial context, healthcare managers have traditionally paid little attention to the management of inventories (Nicholson et al. 2004, De Vries 2011, Kelle et al. 2012, Rossetti et al. 2012, Guimarães et al. 2013). However, in recent years, the management of inventories has been identified as one key lever to realize efficiency improvements without negatively affecting the patient care. Scholars estimate that 10% to 18% of hospitals' net revenues are spent on inventory costs (Jarrett 1998, Nicholson et al. 2004). Hospitals in the U.S. and in France hold an average amount of $4,000 and $5,720 per bed, respectively, in medical supplies alone (Aptel & Pourjalali 2001).

In hospitals, the distribution of goods is typically designed as a multi-echelon inventory system. A central warehouse receives goods from suppliers. The central warehouse is commonly closely connected to the central pharmacy, being in charge of pharmaceuticals handling and the production of perishable drugs, e.g., intravenous fluids. The central warehouse regularly delivers to the point of use inventories that are typically located close to patient care locations (see Figure 2.5; similar figures may be found in Rivard-Royer et al. (2002) and Bijvank & Vis (2012)). Apart from this "traditional" method, two other goods distribution systems are typically applied in practice. In the semi-direct delivery, the suppliers skip the central warehouse and deliver directly to the point-of-use location. The third approach, direct delivery, is closest to JIT, meaning that the supplier takes responsibility for reacting to patient demand and refilling

supplies at the point-of-use locations (Aptel & Pourjalali 2001). Scholars distinguish between the hospital-external and hospital-internal supply chain. While external supply chain integration efforts receive most of the attention in the area of SCM, the internal supply chain remains the weak point of the chain as a whole (Landry & Philippe 2004).
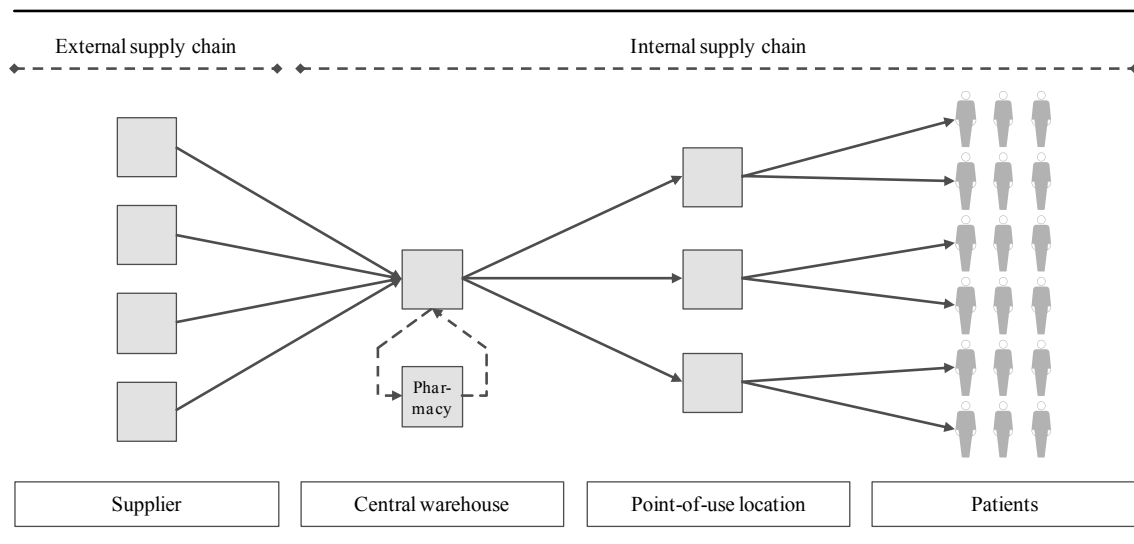
Figure 2.5: Illustrative supply chain

Regarding the setup of hospital inventory systems, several studies argue that hospital inventory management is to some extent comparable to that of other industries. Thus, proven concepts can be transferred to the healthcare industry. Due to the storage space constraints at the point of delivery, i.e., the care unit, laboratory, or operating theater, Little & Coughlan (2008) argue that the respective inventories are comparable to those in retail. Another retail inventory management aspect that could be incorporated in the healthcare environment is the application of "actual-use inventory management", meaning the use of point-of-use data in the upstream supply chain (Varghese et al. 2012). Danas et al. (2006) see strong similarities with the case of spare part inventories for production machines in industrial plants. Decision-makers are faced with a trade-off between the cost of production delays and the cost of safety stock. The literature on hospital inventory management is presented in the following section. The discussion starts with the most relevant field of literature, inventory policy, followed by publications in inventory location planning, inventory item classification, and practice-oriented inventory. We then present specifics and additional requirements for the management of drugs in pharmaceutical inventory management. The section concludes with a subsection focusing on drug shortages and strategies for avoiding them. Future research directions are provided at the end of each subsection.

2.5.1   Inventory policy

The following subsection presents the literature regarding inventory policy. The most widely discussed topic is the choice of a suitable inventory policy, which comprises the definition of the inventory review cycle (periodic or continuous) and parameter-setting for the reorder point, the reorder quantity, and/or the order-up-to level. All presented publications include optimization models. Aspects of hospital inventory management literature reviews may further be found in the following publications: Bijvank & Vis (2012) frame their problem with a brief review of replenishment policies for hospital inventory systems, De Vries (2011) gives an overall introduction of inventory management, and Rosales et al. (2014) introduce their research with a general review of inventory models and briefly discuss related quantitative models in the hospital setting.

The subsection starts with an overarching discussion of the inventory review logic. The publications are then clustered along the inventory locations that they address. It starts with multi-echelon, followed by the central inventory. The focus then shifts towards the patient, i.e., the point-of-use location. At the end of the subsection, we present future research opportunities.

Table 2.4: Inventory policy publications containing optimization models

Publication	Problem description	Model characteristics	Objective function	Type of goods	Review logic	Replenishment policy	Capacity const.	No. of products	Invent. type
Baboli et al. (2011)	Cost comparison of two-level pharmaceutical supply chain: Decentralized vs. centralized replenishment	Heuristic	Minimize total cost	Pharmaceut.	N/A	N/A	Yes	Multi	Multi echelon
Bijvank & Vis (2012)	Inventory policy (reorder points and order quantities) for point-of-use inventories	Heuristic	(1) Maximize service level (2) Minimize required capacity	Disposable items	Periodic	(T, S)	Yes	Multi	Point-of-use
Dellaert & Van De Poel (1996)	Easily applicable inventory policy for central inventory	Heuristic	Minimize ordering cost per cycle	Consumable items	Periodic	(T, s, c, S)	No	Multi	Central inventory
Guerrero et al. (2013)	Optimal inventory parameters for mutli-echelon inventory	Heuristic based on Markov DP	Minimize stock on hand of entire system (CU and CP)	Non-critical, e.g., infusion solutions	Periodic	(T, S)	Yes	Multi	Multi echelon
Kelle et al. (2012)	Easily applicable algorithm for inventory policy parameter setting in point-of-use inventory	Heuristic	(1) Minimize ordering & inventory cost (2) Minimize number of refills	Pharmaceut.	Contin.	(s, S)	Yes	Multi	Point-of-use
Little & Coughlan (2008)	Optimal stock levels under hospital space constraints	Heuristic (constraint programming)	(1) Maximize minimum service levels (2) Max. average service levels	Sterile items, consumables	Periodic	(T, S)	Yes	Multi	Point-of-use
Nicholson et al. (2004)	Costs and service level comparison for in-house three-echelon inventory vs. outsourced two-echelon distribution network	Heuristic (greedy-algorithm)	Minimize inventory cost (holding & backorder)	Non-critical inventory items	Periodic	(T, S)	No	Single	Multi echelon
Priyan & Uthayakumar (2014)	Extension of Uthayakumar & Priyan (2013): Fuzzyness, delivery quantity discrepancies, independent lead times	Fuzzy model defuzzied w. signed distance method	Minimize total cost	Pharmaceut.	Contin.	(s, Q)	Yes	Multi	Multi echelon
Rosales et al. (2014)	Hybrid inventory policy (periodic & continuous review) and inventory parameter definition	Heuristic based on simulation	Minimize long-run average cost per unit time	Pharmaceut., medical supplies	Periodic, continuous	(T, S, s, Q)	Yes	Single	Point-of-use
Rosales et al. (2015)	Two-bin replenishment system: (1) Parameter optimization for periodic review (2) RFID use to enable continuous review	SMDP, linear programming	Minimize order cost and stockout cost	High volume, low cost items	Periodic, continuous	N/A	Yes	Multi	Point-of-use
Uthayakumar & Priyan (2013)	Optimal inventory lot size, lead times and number of deliveries for hospital delivery	Langrangian multiplier algorithm	Minimize total cost	Pharmaceut.	Contin.	(s, Q)	Yes	Multi	Multi echelon
Vila-Parrish et al. (2012)	Inventory and production policy for pharmaceuticals with two stages: Raw material and highly-perishable finished good	Heuristic based on Markov decision process	Minimize expected inventory and production cost	Perishable pharmaceut., e.g., intrav.	Periodic	(T, S)	No	Single	Central inventory

Inventory review logic. As the notation in the literature varies, we introduce the following notation in order to make policies comparable: T (review cycle time), s (reorder point), Q (order quantity), S (order-up-to level), and c (can-order point). An overview of the basic inventory policies is provided in Figure 2.6. We distinguish between periodic review and continuous review. Periodic review comprises the *(T, S)* policy, which is also called "par level policy". It means that, after each review cycle T, orders are triggered so that the order-up-to level S would be in stock. Continuous review comprises two basic policies: The first is the *(s, Q)* policy, where a refill quantity of Q is triggered whenever inventory levels fall under reorder point s. The second continuous policy is the *(s, S)* policy, where instead of a fixed reorder quantity Q, orders are triggered so that the order-up-to level S would be in stock, as soon as storage falls below reorder point s. *N/A* indicates that the policy framework is not applicable. A performance comparison of periodic fixed order size replenishment policies and order-up-to policies is provided in Bijvank & Vis (2012). Regarding stock, we distinguish between safety stock, being i.e., the average buffer inventory, and the cycle stock, which is the average inventory above the safety stock (Kelle et al. 2012).

	When?	
	Every T periods (periodic)	When inventory level below reorder-point s (continuous)
Fixed order quantity Q	-	*(s, Q)*
Until order-up-to level S	*(T, S)*	*(s, S)*

How much?

Figure 2.6: General inventory policies

The relevant literature containing optimization models is presented in Table 2.4. Most publications apply a periodic inventory review policy, which is in line with historic and current practice in hospitals, especially at point-of-use inventory locations such as the wards (Nicholson et al. 2004). In light of the ongoing modernization of point-of-use technologies, like the introduction of advanced identification technologies such as barcodes or radio-frequency identification (RFID), researchers have recently started investigating new types of replenishment policies, such as hybrid policies (Rosales 2011, Bijvank & Vis 2012, Rosales et al. 2014). Rosales et al. (2014) develop such a hybrid replenishment policy. They generally perform a cost-efficient periodic review; however, in order to avoid costly stock-outs, they also permit continuous review. While PhD theses are not in the scope for this review, due to its

importance we refer to Rosales (2011), who elaborates on technology-enabled new inventory policies for hospitals. Kelle et al. (2012) study an automated ordering system, which allows for a continuous review. Uthayakumar & Priyan (2013) argue that periodic inventory review policies are not applicable in practical healthcare settings due to the uncertainty of patient arrivals and resulting demand. Most of the discussed publications incorporate capacity constraints in their inventory models, as space is a limiting factor in the hospital setting, especially at the point-of-use inventories.

When implementing the defined inventory parameters in practice, some typical obstacles exist. Inventory par levels often reflect the desired inventory levels of physicians and nurses rather than the calculated inventory levels. Thus, par levels tend to be experience- or policy-driven rather than data-driven (Prashant 1991). Furthermore, it appears to be common to keep "hidden" safety stock in several locations at care units due to difficulties in policy implementation, personal judgment, and silo-structured organizations (Guimarães et al. 2013).

Multi-echelon inventory. In the following publications, supplier inventories are discussed in parallel to hospital inventories. Baboli et al. (2011) provide a cost comparison for a joint optimization of a pharmaceutical supply chain from a retailer and hospital perspective. They take into account inventory and transportation costs and consider two cases where costs are optimized; a decentralized and a centralized case. In the decentralized case, companies independently optimize their costs, while in the centralized case, several participants in the supply chain are considered as a single organization. They focus on products with high demand and assume that demand for the respective products is deterministic. Uthayakumar & Priyan (2013) take an entire value chain perspective in the case of pharmaceuticals. In their inventory model, they include production and distribution of pharmaceuticals. They develop an algorithm to find the optimal inventory lot size, lead time, and number of deliveries with minimum costs under a continuous review policy. The algorithm is based on a Lagrange multiplier approach. In a second work, Priyan & Uthayakumar (2014) extend their model to cover a fuzzy-stochastic environment, discrepancies between ordered quantities and actually received quantities, and lead times consisting of mutually independent components. Based on the signed distance method, the environment is defuzzied and an optimal inventory policy is determined using the same Lagrange multiplier approach as in the first paper. Nicholson et al. (2004) assess the differences between an in-house three-echelon inventory system and an outsourced two-echelon distribution network, where the replenishment activities are performed by an outside agent who delivers directly to the point-of-use inventory locations in the hospital. The authors develop two optimization models to minimize the holding and backorder costs and apply a heuristic to solve

their problems. They find that outsourcing the distribution of non-critical items is a viable choice, enabling staff to concentrate on patient care activities.

Guerrero et al. (2013) develop a methodology to find near-optimal inventory policies for multi-echelon inventory networks, i.e., one central inventory and n point-of-use inventories. They aim to minimize the total stock on hand for the entire system and employ a Markov decision process. The reorder points at both echelons are derived by means of a probability calculation, while the optimal order-up-to level is one unit higher than the reorder point at the point-of-use inventories. The near-optimal order-up-to level at the central inventory is derived from a heuristic algorithm. Their approach is especially suitable for non-critical goods, such as infusion solutions. In summary, publications cover inventory parameter-setting as well as cost comparisons within and outside the hospital.

Central inventory. Dellaert & Van De Poel (1996) develop a simple and easily applicable inventory control rule for the hospital's central warehouse. The new policy, called "*(R, s, c, S)* policy*", is in our notation a (*T, s, c, S*) policy, which is an extension of the (*T, S*) policy that incorporates a can-order level c. Whenever inventory levels fall below c at the review, an order up to level S can be triggered; if inventory levels fall under s, an order must be triggered. For a periodic review with given review cycles, the inventory parameters are determined using a simple algorithm that minimizes ordering costs based on order bundling. A special case regarding the management of the central inventory is studied by Vila-Parrish et al. (2012). They discuss inpatient medication with two stages: raw materials and finished goods. This holds, for example, for intravenous fluids that are produced in the hospital pharmacy. All goods are perishable, but finished goods are of a more perishable nature. The authors model production and raw material ordering using a Markov decision process.

Point-of-use location. This part discusses the inventory closest to the patient, the point-of-use location. Bijvank & Vis (2012) determine the optimal inventory policy for hospital point-of-use inventories. The authors develop two exact models: a capacity model and a service model. In the capacity model, they maximize the service level subject to capacity restrictions, while in the service model the strategy is vice versa. They develop a simple heuristic inventory rule that can be easily applied by hospital staff for the capacity model. Little & Coughlan (2008) provide a constraint programming-based algorithm that finds optimal inventory parameters, which are service level, delivery frequency, and order-up-to amount for a periodic inventory policy (*T, S*). They especially stress space restrictions and criticality of items.

The two-bin replenishment system, a special replenishment system used in practice, provides for two equally-sized bins in the care units; one bin from which goods are taken and one reserve bin. Once one bin is empty, replenishment is triggered, mostly relying on Kanban logic. This

system is discussed in the following publications. Rosales et al. (2015)[3] study the two-bin replenishment inventory system in combination with RFID tags and assess the applicability of different replenishment policies. Generally, a periodic replenishment policy is applied. Rosales et al. (2015) assess two ways of optimizing the two-bin replenishment system: through parameter optimization for periodic review and through replenishment policy optimization shifting to a continuous replenishment policy. For parameter optimization, the periodic replenishment policy is modeled, and it is demonstrated that the average cost per unit time is quasi-convex, thus allowing for a simple search to find the optimal review cycle. The policy change-driven optimization is enabled through the incorporation of RFID tags, allowing for continuous replenishment. Using a semi-Markov decision process (SMDP), the optimal replenishment policy is modeled and heuristically determined. Landry & Beaulieu (2010) discuss the two-bin replenishment inventory system extensively and assess which lean concepts it addresses.

Automated inventory systems at the point of use are discussed by the following two publications: Kelle et al. (2012) provide the reorder point s and order-up-to level S for the point-of-use inventory of an automated ordering system in a continuous review setting. These inventory parameters are derived by means of a near-optimal allocation policy of safety stock and cycle stock. Parameters are derived using an iterative heuristic algorithm. Rosales et al. (2014) develop a hybrid inventory policy for point-of-use hospital inventories, called "(s, S, R, Q) policy". They combine a periodic (T, S) policy with a continuous (s, Q) policy. Consequently, in our notation this equals a (T, S, s, Q) policy. During the review cycles, reactive replenishments are allowed. This new policy is applicable in particular to automated dispense machines (ADMs) at point-of-use inventory locations. The authors develop a simulation-based heuristic to determine the parameter values for the reorder points, the order-up-to level, the order quantity, and the review cycle. They find that hybrid policies may provide substantial cost benefits versus purely periodic or purely continuous reviews.

Although a multitude of publications exist in the field of hospital inventory policy, this area remains promising for future research. Potential research includes the unpredictable nature of demand in hospitals and its implications on inventory policies. A majority of the presented publications focus on goods with high turnover and predictable demand. However, demand with low volume and lumpy characteristics and its potential effects on workload and policy-setting is hardly considered. This area of research is highlighted by Little & Coughlan (2008) and Kelle et al. (2012). Furthermore, most papers assume that demand for the different goods is independent

---

[3] The publication was available online in 2014 and therefore included in the review

(Nicholson et al. 2004, Kelle et al. 2012). Including dependencies in inventory models could be an interesting research field. Regarding the characteristics of inventory items, it could also be beneficial to incorporate expiration dates or special storage requirements such as cooling (Guerrero et al. 2013). Further, the effects of substitution products on service levels in case of stock-outs could be assessed, as proposed by Bijvank & Vis (2012), as could the fact that emergency deliveries from other care units would be possible, and at little cost for many goods. For multi-echelon inventory settings, a further research area could be to incorporate lead times in inventory models, especially between outside suppliers and the hospital, as proposed by Nicholson et al. (2004). A detailed assessment of the review cycle length at point-of-use inventories and lead times of respective suppliers could therefore be beneficial (Bijvank & Vis 2012). Emphasizing the hospitals' need for simplicity and ease of use could also be a potential future research area. Staff dealing with logistics activities in hospitals typically do not have the same technical background and knowledge as their counterparts in manufacturing industries. Consequently, implementing sophisticated inventory systems may be difficult in hospitals. Examples where the use could be facilitated include simple inventory policies for large-scale inventory systems (e.g., Rosales et al. 2014) or materials handling of ADMs.

2.5.2 Inventory location planning

Danas et al. (2002) provide a publication related to inventory layout planning. The authors introduce the concept of a virtual hospital pharmacy that bundles the inventories of several hospitals in a specific geographic region to allow for a more efficient use of storage capacity. Pasin et al. (2002) use a simulation tool to assess the impact of inventory pooling. They show that significant efficiency improvements can be generated when centralizing inventories of multiple hospitals. Thomas et al. (2000) assess placing an ADM in the point-of-use inventory for an operating room. The authors show that benefits can be realized through the reduction of medication preparation and setup time. For emergency medications, the preparation and setup time could be reduced from 15 to 5 minutes.

In the context of manufacturing industries, strategic planning of inventories like inventory location or layout planning is a large research area. Apart from defining inventory locations, the question of where goods should be stored in a multi-echelon inventory setting has been addressed (Cattani et al. 2011). However, in the healthcare context, inventory location or layout planning is a rather untouched research field. One potential justification is that, in the process of designing a hospital, planners focus on medical aspects, such as the location of operating theaters and wards. Logistics planning is often performed at a later stage, which leads to immature solutions that are not optimal from a materials management perspective (Dacosta-

Claro 2002). Future research opportunities could lie in the development of an integrated approach for hospital layout planning that better incorporates logistics aspects on a strategic level.

2.5.3   Inventory item classification

One lever to efficiently manage inventory is to categorize inventory items and establish individual inventory policies for these categories. This allows for standardized treatment of items within the same category, e.g., in terms of safety stock levels, required management attention, purchasing strategies, etc. In a case study by Beier (1995), 45% of U.S. hospital pharmacies were using a classification scheme to distinguish important items. Potential categorization methods include ABC analyses, meaning categorization along the items' monetary value and rate of consumption, and VED (vital, essential, and desirable) analyses, a classification scheme based on the criticality of items or combinations of the above. Khurana et al. (2013) develop a combination of ABC and VED classifications in order to define the management attention required for the different item categories. A case study for a combined ABC/VED classification is provided by Gupta et al. (2007). Danas et al. (2006) transfer the MASTA logic (multi-attribute spare tree analysis), a concept that was developed in the context of industrial spare parts, to the hospital inventory case. The idea is to classify each drug item along a classification tree in order to determine its stock and inventory strategy, and thus to ascertain whether that drug needs a safety stock within the respective clinic, hospital, or geographic region, or whether it can be supplied as a JIT item. Classification is performed along four dimensions: patient treatment criticality, supply characteristics, inventory problems, and usage rate. We further refer to Al-Qatawneh & Hafeez (2011), who present a multi-criteria inventory classification model based on criticality, cost, and usage value, acknowledging that conference proceedings are not in the scope of this review. Gebicki et al. (2014) incorporate drug characteristics in their inventory policy. They achieve higher patient safety and lower overall costs compared to traditional inventory management approaches. They evaluate the performance of several inventory policies with regards to total costs and service levels using event-driven simulation. The policies differ in the levels to which they incorporate information about the drugs, such as criticality or availability, cost components (e.g., whether stock-out or waste costs are included), and the application of sophisticated techniques, such as conditional demand forecasting.

Future research potential lies in the extension of the previously presented models in order to assess correlations between drug characteristics and the applied policy versus stock-out costs

and the individual cost components (Gebicki et al. 2014). Inventory item classification is further required for the use of innovative inventory systems, such as virtual pharmacies.

2.5.4   Practice-oriented inventory

Several practical case studies and empirical publications on inventory management exist. Huarng (1998) assesses materials management practices in Taiwan across several hospitals in an empirical study. Purchasing strategies, inventory turnover rates, and inventory fill rates are compared across the participating hospitals, and significant performance disparities are identified. An exploratory case study performed by De Vries (2011) underlines the complexity of inventory management in hospitals. The author indentifies and assesses the relevant stakeholders and their interests in the process of redesigning a hospital inventory system. Beier (1995) assesses inventory policies from a U.S. data sample and identifies cost improvement potential in inventory management and the collaboration with suppliers. Poley et al. (2004) present a case hospital containing two pharmacy inventory and distribution systems, i.e., a multi-echelon system and a patient-oriented ready-to-use distribution system. Both systems are systematically compared and differences in their cost structures are highlighted.

Future practical research on inventory management should be conducted in order to better understand the concrete differences between industrial settings and hospitals. Furthermore, reports on past inventory projects in hospitals would be highly beneficial in order to understand the dynamics and potential obstacles in the hospital setting (De Vries 2011).

2.5.5   Pharmaceutical inventory management

Pharmaceuticals impose high requirements on inventory management. According to Almarsdóttir & Traulsen (2005), inventory management for pharmaceuticals differs from other medical product categories based on its specific characteristics. While hospital inventory-related publications on pharmaceuticals were already discussed in detail above, we refer to Kelle et al. (2012), Uthayakumar & Priyan (2013), and Priyan & Uthayakumar (2014) who evaluate pharmaceutical supply chain specifics from the hospital's perspective. For a general introduction to hospital inventory management for pharmaceuticals, we refer to Vila-Parrish & Ivy (2013). Based on regulatory constraints, hospitals must make sure that information about the manufacturer, production lots and/or dates, shipping information, etc. must be registered and known (Çakici et al. 2011). In order to fulfill these identification requirements and prevent medication errors and costly return deliveries, hospitals rely on means of identification, namely barcodes and RFID. In the following subsection, both technologies and their application in

hospitals are presented. The subsection concludes with a brief overview of drug handling techniques.

The use of barcodes is the most widespread identification technology today. According to a cost-benefit analysis by Maviglia et al. (2007), hospitals can achieve significant savings when applying a barcode-based identification and dispensing system instead of a manual system. In their specific case, the break-even point for the upfront investment was reached within one year after the implementation of the new system. Poon et al. (2006) find that implementing a barcode-based hospital pharmacy system can significantly reduce the rate of dispensing errors. Pitfalls of such a system are presented by Phillips & Berner (2004). The work by Koppel et al. (2008) concentrates on workarounds that are performed by medical staff when barcode medication administration systems (BCMA) are in use. Another work by Patterson et al. (2002) presents a case study on implementation problems when using BCMA. The second and more technologically sophisticated identification technology is RFID. Key advantages include easier scanning and high product visibility along the supply chain, albeit at a high implementation cost. A general introduction to the supply chain implications of using RFID – not limited to healthcare – is provided by Lee & Özer (2007) and Irani et al. (2010). For RFID applications in the healthcare industry, we refer to Coustasse et al. (2013) and Fosso Wamba et al. (2013), who provide comprehensive literature reviews. The former find that, despite the rising penetration of RFID in healthcare, few empirical studies exist that assess the actual potential of RFID. An exception is the work by Abijith & Fosso Wamba (2012), who assess the financial impact of RFID-enabled transformation projects in the healthcare sector. Lee & Shim (2007) investigate the rationale behind introducing RFID in the healthcare industry. A highly practice-relevant work is provided by Chang et al. (2012), who elaborate on where to mount RFID tags on products from a material handling perspective. Potential future use cases for the information generated through the application of RFID are presented by Meiller et al. (2011) to further optimize material handling and reduce safety stock levels. A comparison of barcodes and RFID is provided by Çakici et al. (2011) and Chan et al. (2012). Regarding inventory policies, Çakici et al. (2011) find that continuous review is superior to periodic reviews whenever real-time information is available, which is the case for RFID-enabled inventories.

Regarding pharmaceuticals inventory design and drug distribution, hospitals employ either a traditional ward stock system or a unit dose drug inventory and distribution system. In the traditional system, inventory is held at the wards and can be divided into standard and patient-specific medication. In a unit dose system, drugs are picked in the central pharmacy according to the patients' actual needs (Piccinini et al. 2013). New material handling technologies are commonly first adopted in the central hospital pharmacy, where manual picking is replaced by

ADMs. There are case studies at hand examining their effects and learning from their implementation: The publications by Fitzpatrick et al. (2005), Franklin et al. (2008), and Piccinini et al. (2013) in particular evaluate ADMs in hospital pharmacies. Major effects include a significant reduction of dispensing errors, reduced picking times, increased staff satisfaction, and better use of storage capacity (Fitzpatrick et al. 2005, Franklin et al. 2008). Piccinini et al. (2013) present and analyze an automated picking workstation as part of an automated pharmacy distribution center for a group of hospitals. They focus on the actual picking step and assess how to pick very diverse and complex objects available on belts or in bins. Novek (2000) and Granlund & Wiktorsson (2013) more broadly assess the implementation and implications of automation in hospital-internal logistics. For a literature review on types and causes of dispensing errors, we refer to Beso et al. (2005) and James et al. (2009). Future research should further explore new handling technologies emerging in other industries with a focus on their applicability in hospitals.

2.5.6   Drug shortages

In recent years, drug shortages have occurred more and more often, having notable implications for hospital material management in general and inventory management in particular. Historically, the problem of drug shortages has been most common in either niche drug segments or developing countries. However, since the early 2000s, it has been reported that several drug groups are insufficiently supplied in the U.S. – a trend which experts argue is also prevalent in Europe, while exact data to prove this is missing (Fox & Tyler 2003, Le et al. 2011, Huys & Simoens 2013). Most of the existing literature focuses on drug shortages in the U.S. A number of publications assess the causes of drug shortages and their implications on the healthcare system. Reasons for shortages include the unavailability of raw material, production ramp-downs or manufacturing difficulties, mergers and acquisitions of drug manufacturers, voluntary recalls, regulatory issues, unexpected demand, natural disasters, and labor disruptions (Gu et al. 2011, Mangan & Powers 2011, Ventola 2011). Drug shortages have significant negative financial effects on the healthcare system as well as the quality of patient care (Baumer et al. 2004, Kaakeh et al. 2011, Alspach 2012). Nowadays, pharmacists and pharmacy technicians spend an average of eight to nine hours per week on drug shortage-related activities (Kaakeh et al. 2011).

Several authors present very hands-on suggestions on how to cope with drug shortages from a hospital logistics perspective. One key measure is to take proactive action, which includes purchasing strategies as well as the implementation of concrete action plans that ought to be developed before shortages occur. The plans include lists of substitute products and hospital

organizational issues, such as to define responsibilities in case shortages occur. The introduction of substitute products might have significant effects on logistics processes. For example, the IT and inventory systems must be able to cope with short-term changes to drug names, etc. (Mazer-Amirshahi et al. 2014). In inventory management, one lever for coping with drug shortages is changing the inventory policy and updating the order points and order quantities. Additionally, inventory sharing and pooling, as well as rationing and prioritizing policies should be considered (Johnson 2011). Having transparency on upcoming or expected shortages and the actual inventory level of respective drugs helps to be able to act proactively (Johnson 2011).

Apart from the rather practical action plans discussed in the previous publications, there are also publications that present general guidelines on how to prepare for (potential) drug shortages (Fox et al. 2009, Vaillancourt 2012, Gupta & Huang 2013, McLaughlin et al. 2013, Rider et al. 2013). We also refer to a number of studies that focus on the effects of shortages for certain drug groups (Griffith et al. 2012, Hall et al. 2013, McBride et al. 2013, McLaughlin et al. 2014). According to pharmacists, there is a lack of information needed to manage shortages, e.g., actual inventory data throughout the hospital (Kaakeh et al. 2011). Consequently, one future research area could be to further enhance data transparency and availability of up-to-date stock information.

2.6   Literature on topic (3) distribution and scheduling

This section discusses hospital-internal and hospital-external distribution and scheduling topics. Due to its distinct characteristics, the logistics of sterile items is presented in a third category. Hospital-internal distribution comprises mainly routing and scheduling problems of goods within the hospital, primarily from the central warehouse location to the respective care units. External distribution relates to inter-hospital transports as well as waste management. Sterile items handling comprises both transportation tasks and the actual sterilization process.

2.6.1   Hospital-internal distribution and scheduling

In the field of hospital-internal distribution and scheduling, four publications are identified that contain optimization models (see Table 2.5). "Traditional" pharmacy delivery is scarcely addressed in the literature (Augusto & Xie 2009). However, fairly specific issues are discussed, such as routing and scheduling problems of combined storage/delivery material management systems, e.g., mobile medicine delivery closets. Interestingly, due to the different delivery tools and setups, there are hardly any standards and common practices on how to transport materials in hospitals.

Michelon et al. (1994) and Augusto & Xie (2009) explore delivery problems of sophisticated inventory storage and delivery systems, i.e., medicine closets and twin trolleys. Augusto & Xie (2009) consider a hospital pharmacy delivery problem. In their study, pharmacy delivery is performed in a manner such that care units are equipped with mobile medicine closets. Periodically, these closets are collected and transported to the central pharmacy for inventory stock assessment and refill. The problem consists in creating a transportation and supply plan. The aim is to have a balanced workload for the two limiting human resource types, i.e., transporters and pharmacy assistants. The problem is formulated as a MIP, and a near-optimal schedule is determined using a standard solver. In a second step, a simulation model is applied to redesign the pharmacy delivery process in a case study. Michelon et al. (1994) compare two supply distribution systems. In their research case, supplies are delivered in a twofold way: First through so-called "twin trolleys" that contain most of the regularly required supplies, which are always doubled. One trolley is in use at the point of use and the "twin" is located at the central inventory location. Second, there are "non-twin trolleys" containing non-medical items, e.g., meals or cleaning products. Each of those trolleys is assigned to a number of point-of-use locations. In their publication, Michelon et al. (1994) assess whether it is beneficial to change the allocation of items to the twin or non-twin trolleys relying on a tabu search heuristic.

Table 2.5: Hospital-internal distribution and scheduling publications containing optimization models

Publication	Problem description	Model characteristics	Objective function	Type of goods
Augusto & Xie (2009)	Transportation and supply plan for mobile medicine closets located at care units; weekly replenishment in central pharmacy	Mixed-integer linear progamming, simulation	(1) Minimize number of routes (2) Minimize workload	Pharmaceut.
Banerjea-Brodeur et al. (1998)	Deliver quantity and schedule of regular linen delivery from central laundry to care units	PVRP solved with tabu search heuristic	Minimize total cost	Laundry (linen)
Lapierre & Ruiz (2007)	Multi-item inventory replenishment schedule under storage and manpower capacity constraints	Mixed-integer non-linear prob., meta-heuristic search	Minimize total inventory cost and minimize deviation of workload equilibrium	Not specified
Michelon et al. (1994)	Comparison of mobile inventory & distribution systems with varying amout of items kept locally in care units	Tabu search heuristic	Minimize number of tasks that cannot be performed by respective system	Medical, bed-related, meals, etc.

Linen delivery is modeled and optimized by Banerjea-Brodeur et al. (1998). The transportation system is reviewed based on shortfalls that regularly require emergency deliveries to the care units. The authors set up a periodic vehicle routing problem (VRP) in order to optimize delivery routing, scheduling, and quantities. To solve the VRP, a tabu search heuristic is applied.

Dean et al. (1999) focus on scheduling pharmacists who visit care units in order to trigger medication orders. In their model, drug prescriptions are added to the patient files, which are typically mounted to the bed of the respective patient. The study demonstrates that changing the time of day when the visit is performed affects the delay of medication arrivals.

The publication by Lapierre & Ruiz (2007) assesses scheduling activities and logistics optimization. The authors state that, in the context of hospital supply systems, basically two approaches exist to plan logistics activities. First, the inventory-oriented approach, where orders are placed in multi-echelon inventory settings whenever reorder points are met. In this predominant approach in the literature, the main focus lies on ensuring sufficient stock levels (see Section 2.5). However, according to the authors, this approach neglects further questions such as planning of scheduling activities and human resources. Consequently, the second approach focuses on scheduling and answering questions that include: When should employees work? How often should replenishments be performed? How often and when should supplies occur? Lapierre & Ruiz (2007) propose to schedule replenishments, purchasing activities, and supplier activities to avoid stock-outs and respect resource availability. The authors formulate a mixed-integer non-linear scheduling problem that balances employees' workloads. They develop a tabu search meta-heuristic algorithm for solving the problem.

Scheduling and questions around goods distribution of hospital-internal logistic activities appear to be a promising field for future research. Three potential research areas have been identified. First, the introduction of sophisticated inventory and delivery systems in hospitals raises optimization potential for associated activities. These systems include, for example, mobile medicine closets, twin trolleys, or the two-bin replenishment system. The motivation behind their introduction derives from hospital specifics that limit the applicability of standard solutions from other industries. Relevant characteristics include limited storage space at point-of-use locations or staff that are untrained in the use of the logistics system. Moreover, legislative constraints in drug handling make healthcare-specific solutions necessary. A corresponding use case is presented by Augusto & Xie (2009). The authors schedule pharmacists and transporters used when introducing mobile medicine closets. The case of twin trolleys is discussed by Michelon et al. (1994), while Dean et al. (1999) optimize the scheduling of pharmacist visits to wards. A topic that has not been touched so far is the two-bin replenishment system and associated scheduling requirements. Related scheduling activities, such as when to refill stock and in which overall schedule, have not yet been addressed. The work by Lapierre & Ruiz (2007) initially covers logistics-related scheduling tasks around hospital inventories. However, many potential areas for optimization remain. Second, as the majority of the presented publications apply heuristics to solve their models, the development of exact solution procedures could provide further interesting research areas. A third promising field is the extension of existing logistics-related scheduling activities to personnel planning and shift planning. Respective questions might appear when logistics activities are transferred to non-medical support staff.

2.6.2   Hospital-external distribution and scheduling

Hospital-external distribution and scheduling is hardly covered in the literature. In total, there are five relevant publications including optimization models (see Table 2.6). They handle inter-hospital transports, transports between suppliers and hospitals, as well as the collection and disposal of waste. In the context of this work, only hospital-related publications are discussed.

Bailey et al. (2013) investigate an alternative supply route for time-critical items to hospitals, which usually travel in conjunction with regular goods. They demonstrate that an unattended locker box can serve as an alternative delivery solution for urgent items, allowing those items to be separated from regular material flows. The authors use a hill-climbing optimization algorithm to identify the optimal size of a locker box to cover a certain service level in a typical hospital. Combined with results of staff interviews, they find that the introduction of unattended locker boxes would be beneficial in terms of speed of delivery and healthcare quality. Kergosien et al. (2013) address a two-level VRP with time windows, a heterogeneous fleet, multi-depot, multi-commodity, and split deliveries. The first level addresses fleet routing for collection and delivery of pharmaceuticals and hospital consumables. The second level addresses routing of employees between hospital unit buildings and sizing of warehouse employees. To solve the problem, two metaheuristic algorithms are presented; a genetic algorithm and a tabu search algorithm. Swaminathan (2003) discusses the allocation and distribution of scarce drugs to 150 hospitals in California. An optimization model is developed for that purpose that is solved based on an allocation heuristic.

Table 2.6: Hospital-external distribution and scheduling publications containing optimization models

Publication	Problem description	Model characteristics	Objective function	Type of goods
Bailey et al. (2013)	Feasibility demonstration of unattended locker box delivery system; specification of locker box characteristics	Hill climbing optimization algorithm	Maximize number of orders to be stored within locker box	Urgent items
Kergosien et al. (2013)	Transportation flow design between hospital units; warehouse employee dimensioning	2-VRP sol. w. meta-heuristics (generic alg. & tabu serach)	Minimize the sum of delays and minimize required number of employees	Pharma-ceutilcals, consumables
Medaglia et al. (2008)	Optimal facility location for hospital waste treatment network	Biobjective facility loc. prob. (MIP), sol. with heuristic	Multiobjective: (1) Minimize transport. cost (2) Min. affected population	Waste
Shih & Chang (2001)	Route and schedule for periodic waste collection of hospital network	PVRP and MIP to assign routes to days of week	(1) Minimize transportation cost (2) Minimize daily travel mileage in a week	Waste
Swaminathan (2003)	Decision support for allocating scarce drugs to hospitals	Multiobjective optimization model, solved with heuristic	Minimize total value of drug budget left over / maximizing total value of allocated drugs	Pharma-ceuticals

The collection and disposal of waste is considered a separate field of research that mainly deals with VRPs. Due to its relevance for hospital logistics, publications that cover hospital waste disposal are briefly presented. For more detailed information on waste collection and waste

management, we refer to the literature review by Beliën et al. (2012). Medaglia et al. (2008) design a hospital waste disposal network in Columbia. They formulate the problem as a bi-objective obnoxious facility location problem (BOOFLP) that incorporates the trade-off between finding the cost-optimal facility locations and the negative effects on the population close to waste treatment facilities. They solve their model with a multi-objective evolutionary algorithm. Shih & Chang (2001) develop a periodic VRP to model a routing and scheduling problem for the collection of infectious hospital waste. They develop a MIP to assign routes to particular days of the week in a second step. An overview of infectious waste management in European hospitals is provided by Mühlich et al. (2003).

Three potential future research areas have been identified: First, it could be promising to assess the robustness of the presented VRPs and their modifications, e.g., through the application of discrete-event simulation as proposed by Kergosien et al. (2013). Second, referring to Section 2.5, where we identified hospital layout planning as one potential future research area, incorporating inter-hospital transportation issues into layout planning could be a promising future research field. Third, emergency deliveries within hospital networks could furthermore be assessed within this context.

2.6.3   Sterile medical devices

The handling of sterile medical items is a distinct field of research within hospital distribution and scheduling. Fineman & Kapadia (1978) were among the first to address this problem in the OR literature. For a brief introduction of sterilization logistics, see Di Mascolo & Gouin (2013). There are two kinds of sterile medical items: single-use and reusable medical items. We consider the latter because of the complexity of the repetitive sterilization process, which is not necessary for single-use items. Typically, reusable sterile items such as surgical instruments are sterilized after usage either in a hospital-internal sterilization department or by external service providers.

Table 2.7: Sterile device publications containing optimization models

Publication	Problem description	Model characteristics	Objective function	Type of goods
Ozturk et al. (2014)	Near optimal parallel job batch definition for washing step of sterilization process for sterile medical devices	Heuristic based on branch and bound	Minimize makespan of sterilization process	Sterile medical devices
Tlahig et al. (2013)	Comparison of decentralized in-house vs. central sterilization service of hospital network	MIP, solved via addition of approp. customized cuts	Minimize total cost (transportation, sterilization, resource transfer, acquisition)	Sterile medical devices
Van De Klundert et al. (2007)	Sterilization cost minimization (transportation and inventory costs) for outsourcing sterilization of medical devices	Dynamic programming	Minimize total cost	Sterile medical devices

An overview of publications in this research stream applying optimization models is presented in Table 2.7. Ozturk et al. (2014) assess the washing process itself, which they claim to be the bottleneck of the sterilization process. They provide a branch-and-bound-based heuristic in order to optimize the washing machine schedule. Van De Klundert et al. (2007) state that hospitals intensively attempt to outsource sterilization activities in order to save costs and free up space in the central sterilization service departments (CSSDs). However, outsourcing comes with downsides, for example longer transportation distances and potentially lower availability. In their work, the authors formulate an optimization problem aiming to minimize inventory and transportation costs as a lot sizing and transportation model, which is solved in polynomial time by dynamic programming. Further, they extend the model to a dynamic, non-deterministic setting addressing the value added by real-time information availability, e.g., when applying RFID. Additionally, they present a bundling problem regarding the composition of medical item nets. Tlahig et al. (2013) assess two different setups of sterilization services. They compare decentralized in-house sterilization against centralized sterilization services in a hospital network. In their model, they aim to find the general setup (centralized versus decentralized), the optimal location, and the optimal capacity. The problem is modeled as a MIP and solved based on the addition of appropriate cuts. Di Mascolo & Gouin (2013) also aim to improve sterilization services in hospitals. They assess the implications of changes to the processes and the organization. To do so, they develop a generic discrete-event simulation model, allowing the authors to represent and quantify any sterilization service in the respective health establishment in France.

Future research may focus on further assessing the performance of different sterilization services in hospitals (Di Mascolo & Gouin 2013), as well as different organizational setups. These include mixed forms, where some sterile items might be treated within the hospital, while others are sent to external service providers (Tlahig et al. 2013). Also, the incorporation of uncertainty in scheduling the washing process seems to be a worthwhile research field (Ozturk et al. 2014).

2.7   Literature on topic (4) holistic supply chain management

This section presents publications regarding the management of the entire supply chain. They do not contain optimization models, but are qualitative or conceptual. Consequently, all presented areas offer new perspectives on incorporating and developing optimization models. We classify the publications into three categories. The first, "business process redesign", covers all topics associated with the assessment and redesign of logistics processes and the organization of the hospitals' logistics function. The second, "transfer of logistics concepts from other industries", presents publications that assess whether logistics concepts that are

successfully implemented in other industries, such as lean, can be transferred to hospital logistics. The final part, "benchmarks, best practices, and cost analyses", discusses practice-related publications, mostly case studies that assess logistics cost and its components, as well as cost comparisons across countries or within hospital departments. The major discussion points and most relevant conclusions are presented.

2.7.1   Business process redesign

In this subsection, publications are presented that aim to improve hospital business processes. Several approaches according to the literature are shown and tools are presented. Generally, it is accepted that logistics processes in hospitals hold significant cost improvement potential. One relevant lever is to redesign logistics processes by implementing SCM practices (e.g., Haavik 2000, Poulin 2003, A. Kumar et al. 2008, De Vries 2011).

Landry & Philippe (2004) generally consider the role of logistics and show how it can serve healthcare and improve the quality of care. A number of publications focus on reengineering the hospital-internal logistics processes, which are the major weak point in hospital logistics (Born & Marino 1995, Coulson-Thomas 1997, Jayaraman et al. 2000). Chandra (2008) discusses trends, issues, and solution techniques for hospital SCM and presents a generic supply chain problem modeling methodology. Kriegel et al. (2013) evaluate what role external contract logistics service providers can play in the German hospital sector. Zheng et al. (2006) and Iannone et al. (2014) assess the potential of the supply chain integration through enhanced IT integration, e.g., data and information sharing. This enables a higher visibility of inventory data and a reduction of lead times and safety stock. In order to analyze the healthcare supply chain, several tools are at hand to support the decision-making process. The tools comprise process modeling techniques (A. Kumar et al. 2008, Di Martinelly et al. 2009, Iannone et al. 2013) and simulation techniques (Jun et al. 1999, Abukhousa et al. 2014).

2.7.2   Transfer of logistics concepts from other industries

Whether logistics concepts that have been successfully implemented in other industries, e.g., car manufacturing or retail, are transferable to the healthcare sector, is an intensively debated topic in the literature. Most publications conclude that, in general, these concepts are applicable in healthcare, but that there are major obstacles that need to be overcome. Young et al. (2004) very broadly discuss the applicability of industrial processes to healthcare, i.e., lean thinking, theory of constraints, Six Sigma, and scenario simulation. They conclude that all concepts are applicable to the healthcare sector. However, they state that they cannot be expected to deliver

improvements immediately, but will typically need to undergo an iterative implementation process in order to be successful.

Lean thinking emerged in operations in the early 1990s, in service operations management in the mid-to late 1990s, and in the healthcare sector in the early 2000s (Laursen et al. 2003). The applicability of lean thinking to healthcare, not necessarily related to material logistics, is conceptually discussed by Kim et al. (2006), Fillingham (2007), Kollberg et al. (2007), Young & McClean (2008), De Souza (2009), and Mazzocato et al. (2010). These authors find that lean thinking has been applied successfully in a wide range of healthcare applications, but while lean thinking usually takes a holistic approach to problems, healthcare often remains limited to narrower applications with limited organizational reach (Mazzocato et al. 2010). Although there seems to be a general agreement on the potential of lean healthcare, it remains challenging to quantify the potential and critically assess its impact. Compared to other industries, such as the automotive industry, healthcare lags behind regarding the implementation of lean concepts (De Souza 2009). Also, there is no clear definition of the term "value" in healthcare, hindering the reduction of non-value-adding activities, as is standard in industry operations (Young & McClean 2008). However, Kim et al. (2006) rather optimistically conclude that, in the healthcare sector, and especially in hospitals, lean thinking can provide significant process improvements and thus improve the quality and efficiency of patient care. A range of publications encompasses case studies where lean concepts have been implemented in healthcare. Trägårdh & Lindberg (2004) provide a study of a lean production-inspired transformation project in the healthcare sector in Sweden. Landry & Beaulieu (2010) present the case of a two-bin Kanban system for point-of-use inventories and discuss its implications for the inventory system. Venkateswaran et al. (2013) show that by applying the 5S (sort, set to order, shine, standardize, and sustain) methodology in hospital warehouses, significant increases in inventory turnover can be achieved. 5S represents activities that are required to create a desired work environment. Varghese et al. (2012) assess whether actual usage inventory management practices used in the retail industry are applicable in healthcare inventory systems. In particular, they evaluate whether ABC classification, demand characteristic classification, forecast-based demand planning, and inventory control policies are beneficial in the healthcare setting. They create a mathematical model that assesses the possibility of optimizing parameters for a (s, Q) inventory policy based on actual usage inventory management practices and real data. The authors conclude that, by applying those concepts, cost improvements may be achieved.

The applicability of JIT to healthcare logistics is assessed in several publications. Jarrett (1998) states that the healthcare industry had not implemented JIT at that time and provides examples from the literature to prove this point. Already very early, Kim & Schniederjans (1993)

demonstrate that JIT or stockless material management can significantly improve hospital operations. Heinbuch (1995) provides a case study for the successful implementation of JIT in the hospital sector and proves that significant cost improvements can be achieved. Whitson (1997) even argues that materials management in the hospital's pharmacy would be an ideal candidate for implementing JIT due to its manufacturing-like operations characteristics. Jarrett (2006) again underlines the general transferability of JIT concepts to the healthcare sector but claims that there is a research gap regarding the actual implementation of JIT in healthcare. Organizational modification to support the introduction of JIT is presented by Yasin et al. (2003). Contrary to the previously presented papers, Danas et al. (2002) argue that JIT can hardly be applied in healthcare due to the unpredictable nature of the patient mix and the resulting difficulty in forecasting product demand. S. Kumar et al. (2008) state that one reason why the healthcare sector has been slow to embrace SCM practices compared to other industries is the danger stock-out situations present to the health of patients.

More broadly, a variety of publications assesses whether SCM practices from other industries can be applied in the healthcare sector. De Vries & Huijsman (2011) identify five future research areas in healthcare SCM based on a review of the literature: first, the future role of information technology; second, the influence of different stakeholders on establishing SCM relationships within and between health service providers; third, understanding the strengths and weaknesses of management philosophies like lean/agile manufacturing, business process management, and lean Six Sigma; fourth, defining performance metrics of healthcare SCM; and fifth, applying SCM techniques not only to logistics, but also to patient flow. Ford & Scanlon (2007) discuss SCM performance measurements and supplier contracting principles including their applicability to healthcare. Meijboom et al. (2011) assess the applicability of SCM practices to patient care. McKone-Sweet et al. (2005) find that, while the importance of SCM in healthcare is widely recognized, there is only limited research on the unique challenges of healthcare SCM. Operational, organizational, and environmental barriers that hinder the implementation of SCM in healthcare are presented.

2.7.3 Benchmarks, best practices, and cost analyses

Several practice-related publications, mostly in the form of case studies, compare cost characteristics across hospital departments, different countries, etc. and provide benchmarks for the hospital logistics costs setup. Aptel & Pourjalali (2001) compare logistics costs and differences in hospital SCM of large hospitals between France and the U.S. Pan & Pokharel (2007) investigate logistics activities of hospitals in Singapore and specifically assess what kinds of activities are performed by the logistics departments. The same field of research is

covered by Dacosta-Claro (2002), who studies the tasks and management approaches of hospital materials managers. Ferretti et al. (2014) assess implications of reorganizing hospital materials processes and organization, while Kafetzidakis & Mihiotis (2012) more generally evaluate the awareness of logistics in hospitals in Greece. Potential future research could include a global benchmarking tool that allows for comparison of different logistics setups as well as knowledge transfer and sharing of lessons learned.

2.8 Conclusion

The healthcare sector in general and hospitals in particular face significant challenges due to steadily increasing healthcare costs. In hospitals, logistics-related costs are the second largest cost block after personnel costs. In order to reduce material-related logistics costs, healthcare academics and practitioners alike acknowledge the potential of applying quantitative methods. These methods have already proven their potential in other industries, such as manufacturing or service, but need to be modified to account for healthcare-specific problem settings. Moreover, the existence of several implementation difficulties is obvious due to the operational complexity of hospital logistics as well as organizational barriers. Further, staff entrusted with logistics activities in hospitals often have no logistics background, which makes implementation of sophisticated concepts difficult. Optimal solutions are overruled or tend to be policy- and experience-driven rather than data-driven.

The purpose of this chapter is to present the state of the art of research in hospital materials logistics with a specific focus on publications applying quantitative methods. A comprehensive literature review is conducted. Our contribution is threefold: First, we provide guidance for researchers by categorizing the literature and identifying major research streams; second, we methodologically discuss the publications; and third, we identify future research directions. Four major research fields are identified, of which three, i.e., (1) supply and procurement, (2) inventory management, and (3) distribution and scheduling, apply optimization techniques. The remaining identified research field, (4) holistic supply chain management, comprises a rather qualitative field of literature. In total, 145 publications are identified, categorized, and discussed thematically and methodologically. The largest thematic category in terms of number of publications is (2) inventory management (66 publications) over the entire time span, followed by (4) holistic supply chain management (38 publications), (1) supply and procurement (25 publications), and (3) distribution and scheduling (16 publications), respectively. The number of publications in the field of hospital logistics has been growing over the last years. For example, during the previous three years in scope, the total number of publications nearly doubled compared to the years before. Apart from their relevance for academics, the results of this

chapter and the overview it provides should also be of interest for practitioners in hospital materials management functions.

Hospital materials management is a steadily growing field of research in which further promising research opportunities exist. Opportunities are presented in detail in the respective sections. In summary, we identify five overarching research possibilities: First, when integrating the four identified major research streams with applied methodologies, it becomes apparent that the field of (4) holistic supply chain management offers further research potential with regard to the application of quantitative techniques. So far, integrated optimization across the entire supply chain has not yet been performed in the hospital logistics context. Second, answering the question of how to measure performance in hospital logistics is also a promising future research opportunity. Metrics from other industries, e.g., throughput time, are not directly applicable to hospital logistics, as they do not take into account the patient care specifics. Third, future research should continue to incorporate the healthcare and hospital view into operations management, and transfer established concepts from other industries into healthcare while accounting for industry specifics. In doing this, it is of pivotal importance to adjust research according to regional specifics due to the high importance of national legislation and strongly regulated nature of the healthcare industry. Fourth, it could be worthwhile to assess which enablers exist that could further push the implementation of sophisticated logistics concepts in hospitals. Potential enablers include consistent information technology systems and data standards across hospitals, clearly defined data interfaces between hospitals and their suppliers, or the introduction of uniform RFID technology. Fifth, in this context, it could be worthwhile to more specifically assess why healthcare has not yet reached the same professional level as other industries and to identify and evaluate potential implementation barriers. However, as healthcare is lagging behind in terms of the implementation of quantitative tools as well as SCM practices, other more successful industries should stand as an example for future research.

3 A column generation approach for strategic workforce sizing

In Chapter 3, we address the strategic workforce sizing problem for dimensioning the number of logistics assistants in hospitals' logistics divisions. A column generation-based solution approach is presented.

3.1   Introduction

Healthcare costs have outgrown GDP increases in all OECD countries in recent years (OECD Publishing 2015), with hospitals accounting for approximately 30% of healthcare expenditures (OECD Publishing 2013). In hospitals, the two single largest cost items are personnel costs and logistics-related costs, respectively (Poulin 2003, Ross & Jayaraman 2009). Consequently, one major challenge for hospital managers is containing costs, without negatively affecting the quality of patient care. A second major challenge results from the current and projected shortage of qualified medical staff. This holds, for example, for nurses in Germany and in the United States of America (Bundesagentur für Arbeit/German Federal Employment Agency 2014, U.S. Department of Labor 2013). Hospital managers found numerous levers for coping with the challenges mentioned above. Strong potential has been identified in scheduling workers more effectively and efficiently (Van Den Bergh et al. 2013). Additionally, logistics process improvements are of high relevance, in particular the application of logistics concepts that are already being successfully applied in more mature industries, such as manufacturing (De Vries & Huijsman 2011). Another promising cost containment lever lies within relieving medical staff of tasks not related to patient care. This allows physicians and nurses to focus on their core business, namely caring for patients (Jackson Healthcare 2014).

In light of the identified levers, hospitals introduce a new type of employee, referred to as logistics assistants to take over logistics tasks from medical staff. It is suggested that nurses spend approximately 10% of their time on logistics tasks (Landry & Philippe 2004). Relieving them of tasks such as stock replenishments and order placements, cleaning, laundry collection and distribution, or food preparation and disposal could increase nurses' efficiency in terms of net time for patient care, improve job satisfaction, and thus enhance the quality of care. The nature of the respective tasks is bivalent: Most tasks are rather fixed in time, because they are linked either to patient care schedules, such as food distribution, or to hospital-wide logistics processes. For example, this holds for certain material transport tasks that need to be performed within fixed distribution slots. Other tasks, however, such as cleaning or inventory replenishment, are more flexible in nature, as they merely need to be performed within one day, for example. Tasks may be linked by precedence relations, i.e., preceding tasks have to be finished before the next task may start.

When introducing logistics assistants, hospitals are faced with the problem of dimensioning their number. After defining which tasks ought to be reassigned, the number of required assistants needs to be defined. Our approach for this strategic optimization problem is twofold: First, we leverage the flexible nature of logistics tasks as stated above. Second, we incorporate flexible shift scheduling for logistics assistants, meaning that start times and shift lengths are arbitrary, while regulatory requirements are fulfilled. Our objective is to find the optimal, i.e., the minimum number of logistics assistants and provide a feasible schedule of all performed tasks. We do not explicitly assign tasks to logistics assistants, i.e., the task assignment problem is not addressed in our work.

The contribution of this chapter of the thesis is as follows: We develop a new problem formulation as a mixed-integer program (MIP) that integrates shift scheduling and task scheduling. We introduce a column generation-based algorithm to generate high-quality solutions to our problem in a reasonable time. As part of our approach, we present a lower bound for staff minimization problems, i.e., with an unknown number of available workers. We demonstrate the bound's superior behavior against standard column generation for our test instances. We apply our model to the use case of logistics assistants in hospitals. To do so, we use three real-world test instances in a hospital with 1,800 beds and approx. 5,300 employees. In total, we demonstrate the applicability of our solution approach based on 48 test instances that are based on real-world instances and additionally generated instances. Furthermore, we analyze the value of flexibility in both shift and task scheduling. Based on our real-world test instances, we demonstrate that efficiency gains of 40% to 49% are possible by incorporating flexibility compared to a case with reduced flexibility. While we illustrate the applicability of our approach based on a use case, the model is limited neither to logistics assistants nor to hospitals, but has a wide range of potential application areas in healthcare and beyond, e.g., in production support functions or in the service industry.

The remainder of this chapter is structured as follows: In the following Section 3.2 we frame the problem by reviewing the relevant literature. A detailed problem description and the formulation of the MIP are provided in Section 3.3. We present our solution approach in Section 3.4 and perform a numerical study in Section 3.5. Section 3.6 concludes this part of the thesis with a summary and provides a short outlook to future research opportunities.

3.2   Literature review

Hospital logistics management in general and scheduling problems in particular are growing fields of research holding promising research opportunities (see Chapter 2). In this thesis, we combine flexible shift scheduling and task scheduling. In order to give an overview of previous

work, we outline the relevant literature in the respective fields and provide insights on combined approaches.

Flexible shift scheduling. Comprehensive reviews exits on personnel scheduling in general (Ernst et al. 2004, Brucker et al. 2011, Van Den Bergh et al. 2013). Van Den Bergh et al. (2013) find that nurse scheduling and scheduling in the healthcare sector comprise the biggest application areas for staff scheduling problems. Moreover, there are dedicated reviews on nurse scheduling (Cheang et al. 2003, Burke et al. 2004), as well as a review on physician scheduling (Erhard et al. 2016). Our model relies on an implicit shift scheduling formulation, which is different from predefined shift types. Predefined shift types are characterized by fixed shifts that cover daily demand based on a set covering approach. The initial idea of implicit shift scheduling originates from Moondra (1976), who defines sets of variables for the amount of shifts that start and end in every time period. This idea is picked up by various authors, for example Bechtold & Jacobs (1990), who rely on an implicit model formulation to schedule lunch breaks. Thompson (1995) presents an implicit model that states that shifts starting in a specific time period have to end within one of the subsequent periods. Çezik et al. (2001) generate weekly shift rosters with two rest days. Time-based workload considerations are included in Burke et al. (2006). A further publication by Côté et al. (2011) employs new implicit modeling ideas based on context-free grammars to model shift constraints. In their work, a feasible shift is represented by a word. Closest to our approach is the scheduling literature building on flexible shift formulations (Brunner et al. 2009, 2010, Brunner & Edenharter 2011, Stolletz & Brunner 2012).

In terms of solution techniques, column generation-based approaches are widely used for staff scheduling problems. Bard & Purnomo (2005) rely on column generation to solve a nurse scheduling problem that incorporates individual preferences. Brunner & Edenharter (2011) employ a comparable column generation-based approach for physician scheduling with different experience levels. Beliën & Demeulemeester (2007) focus on decomposition approaches and compare staff-decomposed and task-decomposed column generation approaches. Building on column generation, branch-and-price approaches are furthermore employed in staff scheduling problems in healthcare: Jaumard et al. (1998) employ a branch-and-price approach for mid-term nurse scheduling, while Brunner et al. (2010) use a similar approach but deal with mid-term physician scheduling. Individual preferences are incorporated by Purnomo & Bard (2007), while Beliën & Demeulemeester (2006) employ a branch-and-price approach for scheduling hospital trainees.

Task scheduling. Regarding the task scheduling part of our model, there are strong parallels with resource-constrained project scheduling problems (RCPSP). Recent developments in this

area are covered by a book by Schwindt (ed.) & Zimmermann (ed.) (2015). In particular, some parallels exist between our work and RCPSPs with flexible resources or multi-skilled resources. For an introduction to this topic, we refer to Correia & Saldanha-da-Gama (2015). A general overview of variants and extensions of the RCPSP is provided by Hartmann & Briskorn (2010).

Integrated approaches. Only few publications exist that integrate shift and task scheduling. We present the existing literature in two steps. First, we discuss two previous papers in detail that are close to our approach and work out differences and similarities to this chapter of our work. Second, we extend the scope of the literature review and provide a broader overview of publications that apply integrated models both in the healthcare industry and beyond.

Table 3.1: Literature on integrated shift and task scheduling applying column generation

			Model characteristics					Solution approach						
Literature	Objective	Application	Staff scheduling			Task scheduling		Methodology	Decomp.		Shift SP		Task SP	
			Days -on/-off	Explicit shifts	Implicit shifts	Prece- dence	Time windows		1 SP	2 SP	IP	ShP	IP	ShP
Beliën & Demeulemeester (2008)	Minimize number of employees	Nurse and surgery scheduling	X	X		*n/a (surgery sched.)*		Branch & price		X		X	X	
Maenhout & Vanhoucke (2016)	Minimize staff costs	*n/a (methodological)*	X			X		Branch & price	X		X		*n/a (no T-SP)*	
This section of the thesis	Minimize number of employees	Sizing of logistics assistants			X	X	X	Column generation	X	X	X		X	

There are two publications that use column generation approaches to solve integrated shift and task scheduling problems (see Table 3.1). Beliën & Demeulemeester (2008) present an integrated model for surgery and nurse scheduling on a tactical level. In the task scheduling part of their model, they define surgery schedules, while in the staff scheduling part, the nurses' days-off are set in line with an explicit shift model. In surgery scheduling, they allocate resource blocks to surgeons subject to capacity constraints. They solve the integrated problem by relying on a branch-and-price algorithm. Maenhout & Vanhoucke (2016) present a branch-and-price algorithm for integrated project staffing and task scheduling to perform days-on and -off scheduling for a homogenous workforce. In a prior work by the same authors, different scheduling policies are evaluated for the presented problem type (Maenhout & Vanhoucke 2015). In staff scheduling, our work differs from the previous publications in that we rely on an implicit shift formulation. Regarding task scheduling, Maenhout & Vanhoucke (2016) only consider precedence constraints among tasks, while our model also incorporates time windows. The problem decomposition of the two previous works differs. Beliën & Demeulemeester (2008) decompose their problem into two subproblems, i.e., a shift scheduling subproblem and a task scheduling subproblem. The shift scheduling subproblem is solved by relying on a shortest path (ShP) formulation. Maenhout & Vanhoucke (2016) include the task scheduling part of their

work in the master problem and consequently have only one subproblem type. While our algorithm relies on a decomposition based on two subproblem types, we also compare both approaches in detail in a computational study. Next to our solution approach and the presented lower bound, we consider this a contribution of our work. In summary, this chapter of the thesis at hand is the first to perform implicit shift scheduling in an integrated model and to incorporate time windows in task scheduling. In addition, we present a new solution approach.

Beyond the presented publication by Beliën & Demeulemeester (2008), integrated surgery and staff scheduling has repeatedly been applied in the healthcare sector. Commonly, both tasks are performed in an iterative process starting with surgery scheduling, followed by nurse scheduling and assignment (Di Martinelly et al. 2014). However, only few attempts have been made to incorporate them into one holistic approach: Di Martinelly et al. (2014) integrate elements of nurse scheduling with planning and scheduling of surgeries on an operational level. Moreover, integrated approaches for surgery and physician scheduling have been presented (Van Huele & Vanhoucke 2014, 2015). None of the publications, however, incorporate precedence constraints and flexible shifts. Maenhout & Vanhoucke (2013) develop an integrated model that provides the required staff level of nurses for several wards. Also, they perform nurse staffing based on each ward's staffing policy. A branch-and-price approach is applied to solve their problem. In an early work, Venkataraman & Brusco (1996) develop a comparable integrated model in order to assess workforce management policies. A recent publication by Kim & Mehrotra (2015) applies a two-stage stochastic integer programming approach to schedule nurses and derive staffing decisions.

Beyond the healthcare sector, integrated approaches exist in the area of staff sizing and dimensioning. Beliën et al. (2013) integrate shift scheduling of an aircraft line maintenance company with staff dimensioning; however, they do not incorporate task scheduling. They formulate a MIP that is solved based on problem reformulation and a tabu search algorithm. The following publications integrate task scheduling with staff scheduling but do not include either task relations or flexible shifts. Moreover, none of the publications below is in the healthcare context. An early work in integrated planning and scheduling is the publication by Alfares & Bailey (1997), who perform task planning and manpower scheduling. They aim to minimize project durations and define the days-off, as well as to minimize staffing costs for construction projects. To do so, they present an integer programming-based and a heuristic solution procedure. In a related work, Alfares et al. (1999) incorporate multiple labor categories. Bailey et al. (1995) similarly integrate project task planning and manpower scheduling and develop a MIP that is solved heuristically. In the two latter publications, the integrated models outperform the iterative, two-step models. Wu & Sun (2006) develop a mixed-integer, non-linear program

for simultaneously planning project tasks for multiple projects and assigning workers to projects. They include learning effects of workers and present a genetic algorithm to solve their problem aiming to minimize outsourcing costs. The tasks are planned within time windows, but unlike in our work, precedence relations are neglected. Heimerl & Kolisch (2010) and Kolisch & Heimerl (2012) simultaneously schedule IT projects and assign human resources with multiple skills: Heimerl & Kolisch (2010) develop a MIP and show its superiority compared to heuristics used in practice. Kolisch & Heimerl (2012) introduce a metaheuristic based on a generalized minimum cost flow network that decomposes the problem into a binary scheduling problem and a continuous staffing problem. A variation of this problem is the simultaneous planning of project tasks and assignment to employees: An exemplary application in the chemical industry is performed by Bassett (2000). The output comprises a monthly assignment of employees to job roles. Both a mathematical programming model and a simple heuristic are developed to solve the problem. Valls et al. (2009) consider a comparable task planning and resource assignment problem, which they solve with a hybrid genetic algorithm that combines local search with genetic population management. Drezet & Billaut (2008) schedule a project and assign employees with different skill levels to the respective tasks. A MIP is formulated to which a greedy algorithm is applied to obtain a feasible solution, followed by a tabu search algorithm to improve the solution. A related problem is discussed in the work of Smet et al. (2014). They assign tasks to a defined number of multi-skilled workers, whose working time schedules have been determined beforehand. They introduce a two-phase matheuristic approach. A related problem – the combination of shift scheduling and machine job shop scheduling – is presented in the following two publications: Artigues et al. (2009) combine integer linear programming techniques and decomposition-based constraint programming to better balance production costs and employee satisfaction. Guyon et al. (2014) introduce two comparable solution procedures for a similar problem. They base their solution strategy first on decomposition and cut generation, and second on a hybridization of cut generation with a branch-and-bound approach.

In summary, a number of earlier publications integrate shift scheduling aspects with task scheduling. However, to the best of our knowledge, our approach is unique through the combination of a task scheduling formulation with implicit shift scheduling. Consequently, this chapter of the thesis fills a research gap both from a methodological perspective and through its application to healthcare.

3.3 Problem description and optimization model

Section 3.3 comprises three parts. We describe the problem in the first part, then describe the MIP in the second. In the last part of this section, we present potential model extensions to our generic model.

3.3.1 Problem description

The context of our problem in this chapter is strategic in nature, aiming to find the minimum number of workers required to fulfill all tasks. We therefore need to define a task schedule respecting time windows and precedence relations, as well as shift types and their occurrences. The work performed by logistics assistants is considered the resource supply, while the tasks represent resource demand. It has to be ensured that at any point in time, resource supply at least matches resource demand, meaning that all logistics tasks are performed (see Figure 3.1). By leveraging full flexibility in supply and demand, we can ensure that demand is always met. The resource supply in any time period constitutes all contracted workers on duty at that time. Compared to nurses, logistics assistants can be employed very flexibly. Shifts are flexible, meaning that the start time and the duration of each shift are flexible, as long as legislative constraints are fulfilled. The resource demand in each time period is the sum of required resources in that time period. The required resources derive from the tasks to be performed. Our model innovatively combines two existing modeling approaches: flexible shift scheduling (e.g., Brunner et al. 2010) for the supply and a task scheduling approach considering precedence relations and time windows, as in RCPSPs (e.g., Schwindt (ed.) & Zimmermann (ed.) 2015), for resource demand. Consequently, we innovatively apply demand management to shift scheduling problems. Please note, however, that we do not solve the operational problem of task assignment, i.e., assigning specific employees to tasks, see Ernst et al. (2004). The overall model logic is shown in Figure 3.1. We develop the model and our solution approach for the use case of logistics assistants in hospitals. However, there are numerous other application areas within and beyond the healthcare industry. Within the healthcare industry, we identify comparable problem structures when integrating master surgery scheduling and physician or nurse scheduling (see Section 3.2). Beyond the healthcare industry, comparable problems occur in production support functions. In line production, for example, we think of logistics tasks such as buffer stock refilling or other activities that are not directly linked to the assembly line cycle. Comparable activities in job shop production settings are also conceivable. Moreover, service operations comprise a potential application area beyond healthcare. In particular, we point to facility management of large office complexes, which includes cleaning, disposal, and maintenance activities.
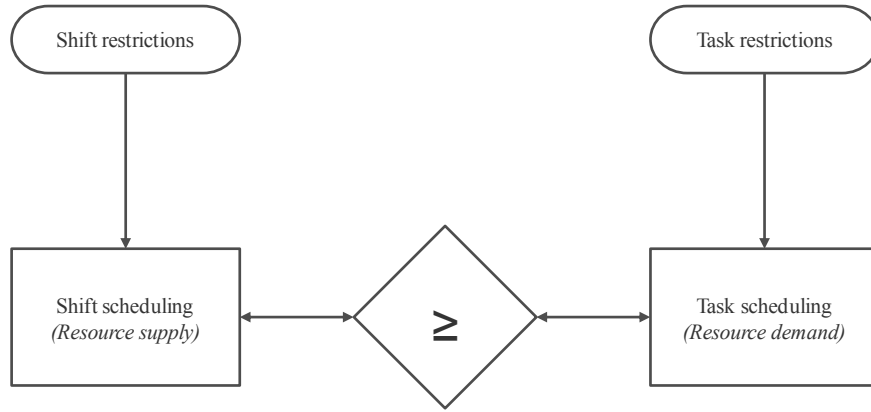
Figure 3.1: Model logic: Matching supply and demand

3.3.2 Mixed-Integer Program

We consider a cyclic problem with a planning horizon of one week. This is in line with existing approaches in the shift scheduling literature and also reflects the repetitive nature of logistics tasks by weeks. Most of the logistics tasks need to be performed daily and in rather fixed time windows, e.g., food distribution. Other tasks, however, such as material supply or cleaning, have larger time windows and do not need to be performed every day. $\mathcal{A}$ is the set of all activities/tasks $a \in \mathcal{A}$. $\mathcal{T}$ is the set of all time periods $t \in \mathcal{T}$. Possible start windows of each task a are depicted in $\mathcal{T}_a^{start}$. For each task, the set of possible start periods $\mathcal{T}_a^{start}$ is defined as the range from the earliest start time $T_a^{earliest}$ to the latest start time T_a^{latest}. Consequently, it holds that $\mathcal{T}_a^{start} = \{T_a^{earliest}, \dots, T_a^{latest}\} \subseteq \mathcal{T}$. The shifts need to be defined for workers $w \in \mathcal{W} = \{1, \dots, W\}$, where W is sufficiently large, i.e., reflecting the labor market for low-skilled logistics workers. When building the model, we limit the number of workers but have to select a sufficiently large number in order to rule out potential infeasibilities. We assume that all workers are equally qualified, which is reasonable for logistics assistants. As a consequence, all tasks can be performed by all workers. Shifts are flexibly scheduled, meaning that shift starts, durations, and consequently their end times are flexible. They are defined by the shift scheduling parameters, being the minimum shift length $T^{shiftMin}$ and the maximum shift length $T^{shiftMax}$. Moreover, a sequence of shifts is constrained by a maximum weekly working time $T^{workMax}$ and a minimum rest time between two consecutive shifts T^{rest}. All presented shift scheduling parameters are measured in time periods. In our model, we assume full flexibility of logistics assistants. Thus, we neglect any potential unwillingness to work in flexible shifts. However, unlike nurses, logistics assistants are rather low-skilled, which limits their bargaining power towards employers. An analysis of how shift flexibility impacts the

required number of logistics assistants is provided in Subsection 3.5.3. Resource demand is determined by all active tasks in a time period and their respective demand R_a. We further assume that the duration of each task L_a is deterministic and known, which is reasonable for the logistics tasks in scope. Moreover, the resource demand of each task is constant while it is performed. The set $\mathcal{A}_a^{pred}$ with $a \in \mathcal{A}_a^{pred} = \{1, \dots, A_a^{pred}\} \subseteq \mathcal{A}$ constitutes all tasks that are direct predecessors of task a.

To formulate the MIP, we introduce the following binary decision variables: x_{at}^{start} take the value 1 if task a starts in time period t and 0 otherwise. The variables y_{wt}^{start} are 1 when a shift of worker w starts in time period t and 0 otherwise. When worker w works in period t, y_{wt} is 1 and 0 otherwise. Decision variables y_w^{work} equal 1 if worker w is employed and 0 if not at costs C^{work}. In order to reflect the work demand in period t, the non-negative integer decision variables d_t are introduced. As a modeling device to achieve feasibility in early iterations, we introduce the auxiliary non-negative integer decision variables z_t, representing the number of demand undercoverage in period t. Their costs C^{out} are set sufficiently high compared to the costs of logistics assistants to ensure that no demand undercoverage exists in any optimal solution.

For the remainder of this section, we apply the following notation: Sets are indicated by calligraphic capital letters, their cardinality with capital letters. Indices are represented by lowercase letters and parameters by capital letters with indices. Decision variables are lowercase letters with indices.

Sets with indices

$\mathcal{T}$ set of time periods with index $t \in \mathcal{T} = \{1, \dots, T\}$

$\mathcal{T}_a^{start}$ set of possible start periods of task a with index $t \in \mathcal{T}_a^{start}$

$$= \{T_a^{earliest}, \dots, T_a^{latest}\} \subseteq \mathcal{T}$$

$\mathcal{A}$ set of tasks with index $a \in \mathcal{A} = \{1, \dots, A\}$

$\mathcal{A}_a^{pred}$ set of direct predecessors of task a with index $a \in \mathcal{A}_a^{pred} = \{1, \dots, A_a^{pred}\} \subseteq \mathcal{A}$

$\mathcal{W}$ set of workers with index $w \in \mathcal{W} = \{1, \dots, W\}$

Task scheduling parameters

R_a resource demand of task a for each time period (number of workers)

L_a length of task a in time periods

$T_a^{earliest}$ earliest start time of task a

T_a^{latest} latest start time of task a

Shift scheduling parameters

$T^{shiftMin}$ minimum shift length in time periods

$T^{shiftMax}$ maximum shift length in time periods

T^{rest} minimum rest time periods between two consecutive shifts

$T^{workMax}$ number of working periods per planning horizon, e.g., per week

Cost parameters

C^{work} cost of employing a logistics assistant

C^{out} cost of assigning an outside worker per period

Decision variables

x_{at}^{start} 1 if task a starts in period t, 0 otherwise

y_{wt}^{start} 1 if worker w starts to work in period t, 0 otherwise

y_{wt} 1 if worker w works in period t, 0 otherwise

y_w^{work} 1 if worker w is employed, 0 otherwise

d_t work demand in period t

z_t amount of demand undercoverage in period t

The objective function is displayed in (1). While the goal of this chapter of the thesis is to find the minimum number of employees, we show the objective function for the more general case of cost minimization. In order to obtain the minimum number of employees as the objective, C^{work} has to be set to 1. The first term represents the cost of employing logistics assistants, each with a cost of C^{work}. The second term represents the cost of assigning demand undercoverage at a cost of C^{out} per period of an outside assignment.

Objective function

$$\text{Min } z^{MIP} = \sum_{w \in \mathcal{W}} C^{work} y_w^{work} + \sum_{t \in \mathcal{T}} C^{out} z_t \tag{1}$$

The constraints can be divided into three blocks, i.e., the shift scheduling constraints (2) to (12), the task scheduling constraints (13) to (17), and the joint constraints (18) and (19).

Shift scheduling constraints

$$\sum_{t \in \mathcal{T}} y_{wt} \leq T^{workMax} y_w^{work} \qquad \forall w \in \mathcal{W} \qquad (2)$$

$$y_{wt}^{start} = y_{wt}\left(1 - y_{w(t-1)}\right) \qquad \forall w \in \mathcal{W}, t \in \mathcal{T} \setminus \{1\} \qquad (3)$$

$$y_{wt}^{start} = y_{wt} \qquad \forall w \in \mathcal{W}, t = 1 \qquad (4)$$

$$\sum_{\tau=t}^{t+T^{shiftMin}-1} y_{w\tau} \geq T^{shiftMin} y_{wt}^{start} \qquad \begin{array}{l} \forall w \in \mathcal{W}, \\[4pt] t \in \left\{ \begin{array}{c} 1, \dots, \\ T - T^{shiftMin} + 1 \end{array} \right\} \end{array} \qquad (5)$$

$$\sum_{\tau=t}^{T} y_{w\tau} + \sum_{\tau=1}^{T^{shiftMin}-T+t-1} y_{w\tau} \geq T^{shiftMin} y_{wt}^{start} \qquad \begin{array}{l} \forall w \in \mathcal{W}, \\[4pt] t \in \left\{ \begin{array}{c} T - T^{shiftMin} + 2 \\ , \dots, T \end{array} \right\} \end{array} \qquad (6)$$

$$1 - y_{w\left(t+T^{shiftMax}\right)} \geq y_{wt}^{start} \qquad \begin{array}{l} \forall w \in \mathcal{W}, \\[4pt] t \in \left\{ \begin{array}{c} 1, \dots, \\ T - T^{shiftMax} \end{array} \right\} \end{array} \qquad (7)$$

$$1 - y_{w\left(t+T^{shiftMax}-T\right)} \geq y_{wt}^{start} \qquad \begin{array}{l} \forall w \in \mathcal{W}, \\[4pt] t \in \left\{ \begin{array}{c} T - T^{shiftMax} \\ +1, \dots, T \end{array} \right\} \end{array} \qquad (8)$$

$$\sum_{\tau=1}^{t-1} (1 - y_{w\tau}) + \sum_{\tau=T-T^{rest}+t}^{T} (1 - y_{w\tau}) \geq T^{rest} y_{wt}^{start} \qquad \begin{array}{l} \forall w \in \mathcal{W}, \\[4pt] t \in \{1, \dots, T^{rest}\} \end{array} \qquad (9)$$

$$\sum_{\tau=t-T^{rest}}^{t-1} (1 - y_{w\tau}) \geq T^{rest} y_{wt}^{start} \qquad \begin{array}{l} \forall w \in \mathcal{W}, \\[4pt] t \in \{T^{rest} + 1, \dots, T\} \end{array} \qquad (10)$$

$$y_{wt}, y_{wt}^{start} \in \{0,1\} \qquad \forall w \in \mathcal{W}, t \in \mathcal{T} \qquad (11)$$

$$y_w^{work} \in \{0,1\} \qquad \forall w \in \mathcal{W} \qquad (12)$$

The shift scheduling constraints (2) to (12) define the shift scheduling problem and are similar to those of previous work (e.g., Brunner & Edenharter 2011). Constraints (2) ensure that the maximum working time is not exceeded by limiting the sum of all working periods in the planning horizon for each worker. Constraints (3) link y_{wt}^{start} and y_{wt}, forcing y_{wt}^{start} to be 1 if y_{wt} changes from 0 to 1, i.e., a new shift begins. Constraints (4) comparably link y_{wt}^{start} and y_{wt} for the start of the planning horizon. Non-linear constraints (3) can easily be linearized, as shown in the appendix. Constraints (5) and (6) set the minimum shift length. Constraints (6) are required to cover the end of the planning horizon. Comparably, constraints (7) and (8) set the

maximum shift length during and at the end of the planning horizon. Constraints (9) and (10) ensure that a minimum rest time of T^{rest} is adhered to between two consecutive shifts. Constraints (9) are necessary to cover the beginning of the planning horizon when the current time period t is smaller than or equal to the minimum rest time T^{rest}. Constraints (6), (8), and (9) make the schedules cyclic. Constraints (11) and (12) are the binary constraints.

Task scheduling constraints

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} x_{at}^{start} = 1 \qquad \forall a \in \mathcal{A} \qquad (13)$$

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} t x_{at}^{start} - \sum_{t=T_b^{earliest}}^{T_b^{latest}} t x_{bt}^{start} \geq L_b \qquad \begin{array}{l} \forall a \in \mathcal{A}, \\ b \in \mathcal{A}_a^{pred} \neq \emptyset \end{array} \qquad (14)$$

$$\sum_{a \in \mathcal{A}} R_a \left(\sum_{\tau=\max\{T_a^{earliest};t-L_a+1\}}^{\min\{t;T_a^{latest}\}} x_{a\tau}^{start} \right) = d_t \qquad \forall t \in \mathcal{T} \qquad (15)$$

$$x_{at}^{start} \in \{0,1\} \qquad \forall a \in \mathcal{A}, t \in \mathcal{T}_a^{start} \qquad (16)$$

$$d_t \in \mathbb{N}_0 \qquad \forall t \in \mathcal{T} \qquad (17)$$

Constraints (13) to (17) represent the task scheduling constraints. They are derived from standard literature (e.g., Schwindt (ed.) & Zimmermann (ed.) 2015). Constraints (13) ensure that each task is performed. Constraints (14) make sure that, for successive tasks, all predecessors are finished before the successor begins. Constraints (15) add up the demand of all tasks a at time period t, which is represented by auxiliary variables d_t. Constraints (16) ensure that x_{at} are binary, while d_t are non-negative and integer according to constraints (17). However, the integrality condition can be relaxed because d_t will always be integer, with R_a being integer and $x_{a\tau}^{start}$ being binary according to (15).

Joint constraints

$$\sum_{w \in \mathcal{W}} y_{wt} + z_t \geq d_t \qquad \forall t \in \mathcal{T} \qquad (18)$$

$$z_t \in \mathbb{N}_0 \qquad \forall t \in \mathcal{T} \qquad (19)$$

Constraints (18) ensure that for each time period t, the resource supply at least matches resource demand. These constraints link the shift scheduling part of our model to the task scheduling model part. Constraints (19) define that variables z_t are non-negative and integer, while in line with d_t, the integrality condition can be dropped.

3.3.3 Model extensions

The model described above is generic and can easily be extended. In the following, we present some possible model extensions known from the shift scheduling literature. Regarding the nomenclature, we add lowercase letters to the original reference if additional constraints are added, while replaced constraints are named according to their original reference and "*".

Breaks. Shifts exceeding certain durations usually contain breaks. During breaks, the available resource capacities are reduced. Thus, modeling breaks creates more detailed (and more realistic) shift rosters, at the price of increased complexity. Breaks could be integrated into our model as in Brunner et al. (2009).

Additional decision variables

y_{wt}^{break} 1 if worker w is on break in period t, 0 otherwise

Additional shift scheduling parameters

T^{pre} minimum number of working periods before a break

T^{post} minimum number of working periods after a break

Constraints

$$\sum_{\tau=t+T^{pre}}^{\min\{t+T^{shiftMax}-T^{post}-1;T\}} y_{w\tau}^{break} \geq y_{wt}^{start} \qquad \begin{aligned} &\forall w \in \mathcal{W}, \\ &t \in \{1, \dots, T-T^{pre}\} \end{aligned} \qquad (10a)$$

$$\sum_{\tau=\max\{0;t-T^{shiftMax}+T^{pre}\}}^{\max\{0;t-T^{post}-1\}} y_{w\tau}^{break} \geq y_{w(t-1)} - y_{wt} \qquad \begin{aligned} &\forall w \in \mathcal{W}, \\ &t \in \left\{\begin{matrix} 2, \dots, \\ T-T^{pre}-1 \end{matrix}\right\} \end{aligned} \qquad (10b)$$

$$\sum_{t\in\mathcal{T}} y_{wt}^{break} = \sum_{t\in\mathcal{T}} y_{wt}^{start} \qquad \forall w \in \mathcal{W} \qquad (10c)$$

$$y_{wt}^{break} \in \{0,1\} \qquad \forall w \in \mathcal{W}, t \in \mathcal{T} \qquad (11a)$$

The additional constraints ensure that a break is assigned to each shift, but not before T^{pre} periods after the start of a shift (10a) and not after T^{post} periods before the end of a shift (10b). Constraints (10c) ensure that the number of breaks is equal to the number of shift starts, i.e., in combination with (10a) that each shift contains one break. The additional decision variables y_{wt}^{break} are binary decision variables and are introduced in (11a). The joint constraints (18) have to be modified accordingly in order to reflect that workers who are on a break do not provide resource supply. The updated constraints are displayed in (18*).

$$\sum_{w \in \mathcal{W}} y_{wt} - \sum_{w \in \mathcal{W}} y_{wt}^{break} + z_t \geq d_t \qquad \qquad \forall t \in \mathcal{T} \qquad \qquad (18^*)$$

Different qualification levels. Both workers and tasks could inhibit different qualification levels. For example, there could be some tasks $\mathcal{A}^{exp}$ that can only be performed by experienced workers $\mathcal{W}^{exp}$, although experienced workers can also perform regular tasks. We present a straightforward model extension.

Additional sets with indices

$\mathcal{A}^{exp}$ set of tasks that require an experienced worker with index $a \in \mathcal{A}^{exp}$

$$= \{1, \dots, A^{exp}\} \subseteq \mathcal{A}$$

$\mathcal{W}^{exp}$ set of experienced workers with index $w \in \mathcal{W}^{exp} = \{1, \dots, W^{exp}\} \subseteq \mathcal{W}$

Additional decision variables

d_t^{exp} work demand for experienced workers in period t

Constraints

$$\sum_{a \in \mathcal{A}^{exp}} R_a \left(\sum_{\tau = \max\{T_a^{earliest}; t - L_a + 1\}}^{\min\{t; T_a^{latest}\}} x_{a\tau}^{start} \right) = d_t^{exp} \qquad \qquad \forall t \in \mathcal{T} \qquad \qquad (15a)$$

$$d_t^{exp} \in \mathbb{N}_0 \qquad \qquad \forall t \in \mathcal{T} \qquad \qquad (17a)$$

$$\sum_{w \in \mathcal{W}^{exp}} y_{wt} + z_t \geq d_t^{exp} \qquad \qquad \forall t \in \mathcal{T} \qquad \qquad (18a)$$

The demand for experienced workers is defined in (15a). Constraints (18a) ensure that a sufficient number of experienced workers are scheduled. The work demand for experienced workers is represented by the non-negative integer decision variables d_t^{exp}, which are introduced in (17a).

Full-time and part-time workers. Our base model assumes that all workers work the same (maximum) number of periods and are associated with the same costs. However, there might be full-time or part-time employees. In an extreme case, each worker could have an individual (maximum) number of working periods associated with individual costs. Moreover, the shift parameters could be worker-specific, for example due to employment contracts with a varying number of maximum working periods. In that case, the cost parameters and shift scheduling parameters would be worker-specific, which would be reflected by an additional index w for the respective parameter.

Linking tasks to workers. Our base model ensures that the required number of workers is available in each time period. We do allow task splitting, e.g., one logistics assistant starts serving food for 30 minutes and another logistics assistant takes over afterwards. Although this is current practice in our partner hospital, there might be tasks that should not be split, or where a dedicated assignment of workers to tasks is required. Obviously, this extension leads to a more detailed level of planning and to increased computational complexity. We introduce the binary decision variables z_{aw}^{link} that are 1 if worker w is assigned to task a and 0 otherwise.

Additional decision variables

z_{aw}^{link} 1 if worker w is assigned to task a, 0 otherwise

Constraints

$$x_{at}^{start}L_a - \left(1 - z_{aw}^{link}\right)L_a \leq \sum_{\tau=t}^{t+L_a-1} y_{w\tau} \qquad \forall a \in \mathcal{A}, t \in \mathcal{T}_a^{start}, w \in \mathcal{W} \qquad (18b)$$

$$\sum_{w\in\mathcal{W}} z_{aw}^{link} = R_a \qquad \forall a \in \mathcal{A} \qquad (18c)$$

$$\sum_{\tau=\max\{t-L_a;0\}}^{t} x_{a\tau}^{start} + \sum_{\tau=\max\{t-L_b;0\}}^{t} x_{b\tau}^{start} \leq 3 - z_{aw}^{link} - z_{bw}^{link} \qquad \begin{aligned} &\forall a \neq b \in \mathcal{A}, w \in \mathcal{W}, \\ &t \in \left\{ \begin{array}{l} \min\{T_a^{earliest}; T_b^{earliest}\}; \\ \max\left\{\begin{array}{l} T_a^{latest} + L_a; \\ T_b^{latest} + L_b \end{array}\right\} - 1 \end{array} \right\} \end{aligned} \qquad (18d)$$

$$z_{aw}^{link} \in \{0,1\} \qquad \forall a \in \mathcal{A}, w \in \mathcal{W} \qquad (19a)$$

Constraints (18b) ensure that if a task is started and a worker is assigned to that task, the worker is assigned to a shift during all periods the task is active. Due to (18c), enough workers are assigned to each task, and (18d) forbid any overlap between two activities assigned to the same worker. Constraints (19a) define that z_{aw}^{link} are binary decision variables.

Controlling days-off. Our generic model limits the maximum number of working periods, but not the maximum number of working days in the planning horizon. One might wish to ensure that a minimum number of days-off is adhered. In the formulation presented below, a day-off is defined as a day without a working period. Please note that other definitions are possible, for example a day without a shift start. In this case, one would need to replace y_{wt} with y_{wt}^{start}.

Additional sets with indices

$\mathcal{D}$ set of working days with index $d \in \mathcal{D} = \{1, \dots, D\}$

Additional parameters

R^{min} minimum number of days-off in the planning horizon

Additional decision variables

y_{wd}^{off} 1 if day d is a day-off for worker w, 0 otherwise

Constraints

$$y_{wt} \leq 1 - y_{wd}^{off} \qquad \forall w \in \mathcal{W}, d \in \mathcal{D}, t \in \left\{ \frac{(d-1)T}{D} + 1, \dots, dT/D \right\} \qquad (10\text{d})$$

$$\sum_{t=(d-1)T/D+1}^{dT/D} y_{wt} \geq 1 - y_{wd}^{off} \qquad \forall w \in \mathcal{W}, d \in \mathcal{D} \qquad (10\text{e})$$

$$\sum_{d \in \mathcal{D}} y_{wd}^{off} \geq R^{min} \qquad \forall w \in \mathcal{W} \qquad (10\text{f})$$

$$y_{wd}^{off} \in \{0,1\} \qquad \forall w \in \mathcal{W}, d \in \mathcal{D} \qquad (11\text{b})$$

We introduce the set $\mathcal{D}$, which represents all days of the planning horizon. The newly introduced binary decision variables y_{wd}^{off} are 1 if day d is a day-off for worker w and 0 otherwise (11b). Ensuring that a minimum number of days-off is adhered can be incorporated by adding additional constraints. Constraints (10d) ensure that, whenever there are active shifts during a day, this day must not be a day-off. In parallel, constraints (10e) ensure that when there are no working periods during one day, that day is a day-off. We ensure that a minimum number of days-off R^{min} is adhered in constraints (10f).

3.4 Solution approach

In Section 3.4, we present our approach to solve the integrated strategic shift and task scheduling problem. The presented MIP cannot be solved with a standard solver within an acceptable time (see Section 3.5). The model described in Subsection 3.3.2 is a combination of a flexible shift scheduling problem and a task scheduling problem. By fixing the supply vector, i.e., eliminating shift flexibility, our model can be reduced to a RCPSP. As the RCPSP is known to be NP-hard in the strong sense (Blazewicz et al. 1983), our integrated problem is NP-hard in the strong sense as well and can consequently only be solved efficiently for small problem instances.

We propose an exact algorithm for our problem that comprises three major steps: First, we develop a tight lower bound for our problem. In order to do so, we apply a Dantzig-Wolfe

decomposition and use a column generation approach to solve the reformulation, which we terminate once the best lower bound has been proven. In the second step, we heuristically find a feasible start solution for our optimization problem, leveraging the columns generated in the first step and additionally generated columns using a sophisticated heuristic procedure. Third, our problem is solved as a MIP in the formulation as presented in Subsection 3.3.2 with a standard solver, albeit benefiting from the information gathered in steps one and two. We provide the lower bound as a hard constraint and use the start solution to warm-start the MIP (see Figure 3.2).

In the following three Subsections 3.4.1 to 3.4.3, we present the problem reformulation and the resulting master problem and subproblems. In Subsection 3.4.4, we introduce our lower bound to terminate the column generation. We then present the method to find a valid start solution (see Subsection 3.4.5), after which we summarize the solution algorithm. Additionally, we present an alternative decomposition approach.

1 Find lower bound		**2** Find good start solution
▪ Column-generation approach: Master problem coordinates two subproblems ▪ Early termination once best lower bound has been found		▪ Choose high quality task schedules ▪ Generate well fitting shift patterns ▪ Solve master problem as integer program relying on columns from (1) and additionally generated columns

3 Solve MIP building on (1) and (2)
▪ Set lower bound from (1) as hard constraint
▪ Leverage start solution from (2) to warm-start solution procedure
▪ Solve enhanced MIP to optimality

Figure 3.2: Overview of the solution approach

3.4.1 Master Problem (MP)

We reformulate the MIP based on a Dantzig-Wolfe decomposition. The problem decomposes by shift schedules and task schedules, which yields two subproblems (SPs) and one master problem (MP). The MP coordinates the SPs and contains the joint constraints (18) and (19). S-SP, i.e., the shift SP, contains the shift scheduling constraints (2) to (12) and generates feasible shift schedules. T-SP, i.e., the task SP, generates task schedules and contains the task scheduling constraints (13) to (17). For the remainder of this chapter, we apply the following nomenclature: A shift schedule is a sequence of working periods, indicated by 1, and off-periods, indicated by 0. The resource supply is defined by the sum of all active shift schedules multiplied with the number of logistics assistants working in the respective schedule. Shift schedules are represented by the set $\mathcal{J}$, with $j \in \mathcal{J}$ representing one element thereof. In order to indicate

whether a period is a working period or an off-period in schedule j, we introduce the binary parameter Y_{jt} that describes the parameter vector $\vec{Y}$ for all t. A task schedule constitutes all start periods of every task a in the planning horizon and is represented by set $\mathcal{P}$. As a consequence, task schedule $p \in \mathcal{P}$ defines the resulting resource demand vector $\vec{D}$ that describes d_t for all t. In the MP, the respective demand in period t of task schedule p is represented by the parameter D_{pt}. We further introduce additional decision variables. The binary variables ρ_p indicate whether task schedule p is selected, while the integer variables λ_j represent the number of employees working in shift schedule j.

Sets with indices

$\mathcal{P}$ set of task schedules with index $p \in \mathcal{P} = \{1, \dots, P\}$

$\mathcal{J}$ set of shift schedules with index $j \in \mathcal{J} = \{1, \dots, J\}$

Parameters

Y_{jt} 1 if shift schedule j is active in period t, 0 otherwise

D_{pt} work demand of task schedule p in period t

Decision variables

ρ_p 1 if task schedule p is selected, 0 otherwise

λ_j number of workers in shift schedule j

In the following, we state the MP.

Master problem

$$\text{Min } z^{MP} = \sum_{j \in \mathcal{J}} C^{work} \lambda_j + \sum_{t \in \mathcal{T}} C^{out} z_t \tag{20}$$

subject to

$$\sum_{j \in \mathcal{J}} Y_{jt} \lambda_j - \sum_{p \in \mathcal{P}} D_{pt} \rho_p + z_t \geq 0 \qquad\qquad \forall t \in \mathcal{T} \tag{21}$$

$$\sum_{p \in \mathcal{P}} \rho_p \geq 1 \tag{22}$$

$$\lambda_j \in \mathbb{N}_0 \qquad\qquad \forall j \in \mathcal{J} \tag{23}$$

$$z_t \in \mathbb{N}_0 \qquad\qquad \forall t \in \mathcal{T} \tag{24}$$

$$\rho_p \in \{0,1\} \qquad\qquad \forall p \in \mathcal{P} \tag{25}$$

The objective function (20) minimizes the total costs and is equal to the objective function of the (original) MIP. The total costs constitute the number of employees working in one specific shift schedule λ_j multiplied with the cost per shift plus the cost for demand undercoverage. Constraints (21) are the key joint constraints that ensure that in any period t, the total resource supply at least equals resource demand. They equal constraints (18) of our MIP formulation. Constraint (22) forces that one task schedule is chosen. For modeling reasons, we do not claim equality (= 1) but use a greater-than-or-equal-to relationship. As selecting more than one task schedule negatively impacts our objective, only one schedule is chosen in all optimal solutions. Constraints (23) and (25) set the newly introduced decision variables λ_j non-negative integer and ρ_p binary, respectively. If all potential shift and task columns are included in MP, it is equivalent to the original MIP formulation as presented in Subsection 3.3.2 with a huge number of columns.

We add columns to our MP by solving the two subproblems iteratively. During column generation, the search for new columns is guided by the dual problem of the linearized MP. In this restricted MP, the integrality and binary conditions of constraints (23), (24), and (25) are dropped and λ_j and ρ_p are treated as continuous variables. As the solution space of both subproblem types is bounded, their solutions mark extreme points of the Dantzig-Wolfe reformulation, i.e., the MP-LP model. Please note that we apply the following nomenclature: We refer to the restricted problem as the linear programming master problem MP-LP. In contrast, we refer to MP-IP when the integrality conditions of constraints (23), (24), and (25) are adhered to. The corresponding dual problem of MP-LP is provided below.

Dual master problem

$$\text{Max } z^{DMP} = \mu \tag{26}$$

subject to

$$C^{work} - \sum_{t \in \mathcal{T}} Y_{jt}\delta_t \geq 0 \qquad\qquad \forall j \in \mathcal{J} \tag{27}$$

$$\sum_{t \in \mathcal{T}} D_{pt}\delta_t - \mu \geq 0 \qquad\qquad \forall p \in \mathcal{P} \tag{28}$$

$$\delta_t \leq C^{out} \qquad\qquad \forall t \in \mathcal{T} \tag{29}$$

$$\delta_t \geq 0 \qquad\qquad \forall t \in \mathcal{T} \tag{30}$$

$$\mu \geq 0 \tag{31}$$

For the dual MP-LP, we introduce additional decision variables, namely δ_t, which are the dual variables of constraints (21) and μ being the dual variable value of constraint (22). The objective function and the constraints are derived from duality.

3.4.2   Shift scheduling subproblem (S-SP)

Once the primal MP-LP is solved to optimality with the existing subset of columns, we use S-SP to generate new shift schedules. New shift schedules are shift columns with negative reduced costs, i.e., we are looking for dual infeasibility cuts. The objective function of S-SP, depicted by $\bar{C}_j$, is the negative reduced cost of a new shift schedule and is derived from the dual of MP-LP. With $\delta_t \geq 0$ being the dual value of constraints (21) of MP-LP, the objective of S-SP is depicted in (32). It minimizes the reduced cost of a feasible shift schedule. They are independent from worker w, thus we drop index w and decision variables y_w^{work} in our S-SP formulation. Additionally, we allow y_t and y_t^{start} to be manually set to 0 for periods where no tasks are performed, such as during nights. Apart from these changes, the constraints of S-SP are the same as the shift scheduling constraints of our MIP, namely constraints (2) to (12).

$$\text{Min } \bar{C}_j = C^{work} - \sum_{t \in \mathcal{T}} \delta_t y_t \tag{32}$$

subject to

$$\sum_{t \in \mathcal{T}} y_t \leq T^{workMax} \tag{33}$$

$$y_t^{start} = y_t(1 - y_{t-1}) \qquad\qquad \forall t \in \mathcal{T} \setminus \{1\} \tag{34}$$

$$y_t^{start} = y_t \qquad\qquad t = 1 \tag{35}$$

$$\sum_{\tau=t}^{t+T^{shiftMin}-1} y_\tau \geq T^{shiftMin} y_t^{start} \qquad\qquad \forall t \in \left\{ \begin{matrix} 1, \dots, \\ T - T^{shiftMin} + 1 \end{matrix} \right\} \tag{36}$$

$$\sum_{\tau=t}^{T} y_\tau + \sum_{\tau=1}^{T^{shiftMin}-T+t-1} y_\tau \geq T^{shiftMin} y_t^{start} \qquad\qquad \forall t \in \left\{ \begin{matrix} T - T^{shiftMin} + 2 \\ , \dots, T \end{matrix} \right\} \tag{37}$$

$$1 - y_{t+T^{shiftMax}} \geq y_t^{start} \qquad\qquad \forall t \in \left\{ \begin{matrix} 1, \dots, \\ T - T^{shiftMax} \end{matrix} \right\} \tag{38}$$

$$1 - y_{t+T^{shiftMax}-T} \geq y_t^{start} \qquad\qquad \forall t \in \left\{ \begin{matrix} T - T^{shiftMax} \\ +1, \dots, T \end{matrix} \right\} \tag{39}$$

$$\sum_{\tau=1}^{t-1}(1-y_\tau) + \sum_{\tau=T-T^{rest}+t}^{T}(1-y_\tau) \geq T^{rest}y_t^{start} \qquad \forall t \in \{1, \dots, T^{rest}\} \qquad (40)$$

$$\sum_{\tau=t-T^{rest}}^{t-1}(1-y_\tau) \geq T^{rest}y_t^{start} \qquad \forall t \in \{T^{rest}+1, \dots, T\} \qquad (41)$$

$$y_t, y_t^{start} \in \{0,1\} \qquad \forall t \in \mathcal{T} \qquad (42)$$

As long as there are feasible solutions of S-SP with $\bar{C}_j < 0$, LP optimality of MP-LP has not been proven, i.e., the identified column is dual infeasible and can consequently be added to MP-LP. Once S-SP is solved to optimality and negative, we add the hereby obtained column to MP-LP. The new column takes the following shape:

$$\begin{bmatrix} C^{work} \\ \vec{Y} \\ 0 \end{bmatrix}$$

C^{work} and $\vec{Y}$ are parameters that characterize the new shift schedule. The parameter $\vec{Y}$ comprises a vector of length T of on- and off-periods 1 and 0, which denotes the new shift schedule, where elements Y_t are equal to 1 if in the solution of the subproblem, y_t equals 1 and 0 otherwise for all $t \in \mathcal{T}$.

3.4.3 Task scheduling subproblem (T-SP)

Comparable to S-SP, where we generate shift schedules, we use T-SP to generate task schedules. The objective function of T-SP derives from the dual problem of the MP-LP, with $\mu > 0$ being the dual value of constraint (22) of the MP-LP. New, promising columns have negative reduced costs. The objective is to minimize the costs of new task schedules. The constraints (44) to (48) of T-SP are the same as the task scheduling constraints of the MIP presented in Subsection 3.3.2. In line with the shift scheduling subproblem, we allow x_{at}^{start} to be manually set to 0 for periods where no tasks are performed. The T-SP objective function is presented in (43).

$$\text{Min } \bar{C}_p = \sum_{t \in \mathcal{T}} \delta_t d_t - \mu \qquad (43)$$

subject to

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} x_{at}^{start} = 1 \qquad \forall a \in \mathcal{A} \qquad (44)$$

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} tx_{at}^{start} - \sum_{t=T_b^{earliest}}^{T_b^{latest}} tx_{bt}^{start} \geq L_b \qquad \begin{array}{l} \forall a \in \mathcal{A}, \\ b \in \mathcal{A}_a^{pred} \neq \emptyset \end{array} \qquad (45)$$

$$\sum_{a \in \mathcal{A}} R_a \left(\sum_{\tau=\max\{T_a^{earliest};t-L_a+1\}}^{\min\{t;T_a^{latest}\}} x_{a\tau}^{start} \right) = d_t \qquad \forall t \in \mathcal{T} \qquad (46)$$

$$x_{at}^{start} \in \{0,1\} \qquad \forall a \in \mathcal{A}, t \in \mathcal{T}_a^{start} \qquad (47)$$

$$d_t \in \mathbb{N}_0 \qquad \forall t \in \mathcal{T} \qquad (48)$$

In line with S-SP, we call T-SP as long as feasible solutions with $\bar{C}_p < 0$ exist. We solve T-SP to optimality and add the resulting column to MP-LP. This column contains the demand profile of the respective task schedule, which is represented by parameter $\vec{D}$ of length T. Elements $D_t \in \mathbb{N}_0$ represent the demand resulting from the T-SP solution d_t for all $t \in \mathcal{T}$. The task column is represented by:

$$\begin{bmatrix} 0 \\ \vec{D} \\ 1 \end{bmatrix}$$

We call S-SP and T-SP iteratively one after another. When neither of the subproblems yields a valid solution, i.e., no further columns with negative reduced costs are obtained, the solution of the MP-LP is proven to be optimal. We have found a valid lower bound of MP-IP and consequently of our original MIP (see Subsection 3.3.2).

3.4.4 Lower bounds

When running the presented column generation, a large number of iterations is necessary until optimality of MP-LP is proven. Optimality is achieved once the optimally solved subproblems no longer price out, i.e., no dual infeasibility cuts are found. The slow convergence of the MP-LP is known as the tailing-off effect, which was first observed by Gilmore & Gomory (1963). As our goal of employing the column generation is to find a good integer lower bound of our problem, we can terminate column generation before achieving optimality of MP-LP.

To find a valid lower bound, we leverage information we obtain from the column generation. The basic idea is to build on information from MP-LP as well as the two subproblems. Our goal is to terminate column generation at iteration i when our current lower bound z_i^{LB} is larger than or equal to our current objective value of MP-LP z_i, i.e., $z_i^{LB} \geq z_i$. We present the lower bound in (49). It is valid at any main iteration of column generation, where j^* and p^* represent the

column indexes of optimal solutions generated by the S-SP and T-SP, respectively. Consequently, $\bar{C}_{j*}$ and $\bar{C}_{p*}$ represent the objective values of the optimal solutions of S-SP and T-SP. We initialize z_0^{LB} with $-\infty$ and set $z_i^{LB} = \max\{z_{i-1}^{LB}; z_i^{LB}\}$.

$$z_i^{LB} = \left\lceil z_i + \bar{C}_{j*} \sum_{j \in \mathcal{J}} \lambda_j + \bar{C}_{p*} \right\rceil \tag{49}$$

In the less general case where the cost of assigning a worker to one shift schedule is fixed to 1, i.e., $C^{work} = 1$ and $z^{MP} = \sum_{j \in \mathcal{J}} \lambda_j$, which holds for our test instances, Lübbecke & Desrosiers (2005) present a comparable lower bound. For all our real-world test instances, both lower bounds yield the same results. Vanderbeck & Wolsey (1996) likewise leverage dual information in order to prune nodes in their branch-and-price approach. If the number of employees to be scheduled is fixed, lower bounds have been presented by Wolsey (1998) and Brunner & Stolletz (2014), which are well known as the Lagrangian bound. Comparably, if there is an upper bound for the number of workers to be scheduled, lower bounds applied to the bin-packing problem as presented by Pisinger & Sigurd (2007) could be used. We provide a proof of the lower bound below.

Proposition. The following lower bound is valid at any main iteration of column generation, where j^* and p^* represent the column index of optimal solutions generated by the S-SP and T-SP, respectively. We initialize z_0^{LB} with $-\infty$ and set $z_i^{LB} = \max\{z_{i-1}^{LB}; z_i^{LB}\}$.

$$z_i^{LB} = \left\lceil z_i + \bar{C}_{j*} \sum_{j \in \mathcal{J}} \lambda_j + \bar{C}_{p*} \right\rceil \tag{50}$$

Proof. As presented previously, $\bar{C}_j$ and $\bar{C}_p$ represent the reduced cost of a shift/task schedule column that have been added to MP, respectively. It holds for all shifts:

$$\bar{C}_j = C^{work} - \sum_{t \in \mathcal{T}} Y_{jt} \delta_t \geq C^{work} - \sum_{t \in \mathcal{T}} Y_{jt}^* \delta_t = \bar{C}_{j*} \qquad \forall j \in \mathcal{J} \tag{51}$$

Comparably, the following holds for all task schedules:

$$\bar{C}_p = \sum_{t \in \mathcal{T}} D_{pt} \delta_t - \mu \geq \sum_{t \in \mathcal{T}} D_{pt}^* \delta_t - \mu = \bar{C}_{p*} \qquad \forall p \in \mathcal{P} \tag{52}$$

$\bar{C}_{j*}$ and $\bar{C}_{p*}$ represent the objective values of the optimal solutions of S-SP and T-SP, respectively. We now can compute z_i^{LB} as shown below.

$$z^{MP} = \sum_{j \in \mathcal{J}} C^{work} \lambda_j + \sum_{t \in \mathcal{T}} C^{out} z_t$$

$$\geq \sum_{j \in \mathcal{J}} C^{work} \lambda_j + \sum_{t \in \mathcal{T}} \delta_t z_t = \sum_{j \in \mathcal{J}} \left(\bar{C}_j + \sum_{t \in \mathcal{T}} Y_{jt} \delta_t \right) \lambda_j + \sum_{t \in \mathcal{T}} \delta_t z_t$$

$$= \sum_{j \in \mathcal{J}} \bar{C}_j \lambda_j + \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} Y_{jt} \delta_t \lambda_j + \sum_{t \in \mathcal{T}} \delta_t z_t$$

$$\geq \sum_{j \in \mathcal{J}} \bar{C}_{j^*} \lambda_j + \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} Y_{jt} \delta_t \lambda_j + \sum_{t \in \mathcal{T}} \delta_t z_t$$

$$\geq \bar{C}_{j^*} \sum_{j \in \mathcal{J}} \lambda_j + \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} Y_{jt} \delta_t \lambda_j + \sum_{t \in \mathcal{T}} \delta_t \left(\sum_{p \in \mathcal{P}} D_{pt} \rho_p - \sum_{j \in \mathcal{J}} Y_{jt} \lambda_j \right)$$

$$= \bar{C}_{j^*} \sum_{j \in \mathcal{J}} \lambda_j + \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} Y_{jt} \delta_t \lambda_j + \sum_{t \in \mathcal{T}} \sum_{p \in \mathcal{P}} D_{pt} \delta_t \rho_p - \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} Y_{jt} \delta_t \lambda_j$$

$$= \bar{C}_{j^*} \sum_{j \in \mathcal{J}} \lambda_j + \sum_{p \in \mathcal{P}} (\bar{C}_p + \mu) \rho_p$$

$$\geq \bar{C}_{j^*} \sum_{j \in \mathcal{J}} \lambda_j + \sum_{p \in \mathcal{P}} (\bar{C}_{p^*} + \mu) \rho_p = \bar{C}_{j^*} \sum_{j \in \mathcal{J}} \lambda_j + (\bar{C}_{p^*} + \mu) \sum_{p \in \mathcal{P}} \rho_p$$

$$\geq \bar{C}_{j^*} \sum_{j \in \mathcal{J}} \lambda_j + \bar{C}_{p^*} + \mu = z_i + \bar{C}_{j^*} \sum_{j \in \mathcal{J}} \lambda_j + \bar{C}_{p^*}$$

The first inequality derives from constraints (29) of the dual restricted MP. The following reformulation as well as the second inequality are based on equations (51), with $\sum_{j \in \mathcal{J}} \lambda_j \geq 0$. The third reformulation builds on constraints (21) from the restricted MP, with $\delta_t \geq 0$. The reformulation after simplifying the term is based on expression (52), with $\rho_p \geq 0$. The second to last inequality is again based on (52). The last inequality follows from constraint (22) of the MP-LP. As we are in a minimization context with an integer objective value, we round up z_i^{LB} to the next integer number (see (50)). □

3.4.5 Start solution and improvement procedure

After computing the lower bound, the second step of our solution approach is to identify a good, feasible start solution. If the solution of MP-LP is integer incidentally and equals our lower bound, we have found the optimal solution and skip the following step. If not, we generate a start solution by solving MP-IP. The number of task schedule columns strongly impacts the

solution time of MP-IP, so we only select a subset of all generated task schedule columns. In order to identify high-quality columns, we search schedules with flat demand. In this context, "flat" indicates schedules whose maximum demand in one time period D_t is small. This method turned out to be effective in preliminary testing, as peaks strongly impact the workforce size. By avoiding schedules with peak demands, we realize a good match to shift schedules and consequently achieve a low number of workers. For each task schedule p, we define the maximum demand $D_p^{max} = \max_{t \in \mathcal{T}} D_{pt}$. We finally select a portion of task schedules with the smallest D_p^{max}.

As a preliminary testing result, 10% and a total maximum of 15 task columns proved to be an appropriate portion. High numbers of task columns negatively affect the solution time. The testing showed that by tripling the share, the solution time of MP-IP was 15 times higher than with a share of 10%, not taking into account the additional time needed to generate the additional shift columns. The objective value of MP-IP, however, remained the same. Including too few task columns in MP-IP had a negative effect on the solution of MP-IP. The impact of shift columns on the solution quality of MP-IP is different from that of task schedule columns. While the negative impact on the solution time is significantly smaller, the solution quality increases with the number of shift schedules. Consequently, we add additional shift columns to the MP-IP. In order to find high-quality shift schedules that fit well with the selected task schedules, we employ a simple heuristic, which we present in the following. The basic idea is to generate well-fitting shift columns for each of the selected task schedules. We do so by modifying the objective function (32) of S-SP. The new objective is given in (53).

$$\text{Max } \hat{C}_j = \sum_{t \in \mathcal{T}} D_t^p y_t \tag{53}$$

We generate shift schedules that maximize the coverage of the selected demand profile resulting from a specific task schedule. We call the updated shift problem iteratively and update D_t^p in each iteration. We update by subtracting the chosen shift schedule y_t for all t in T from our demand profile, which results in the new demand vector $D_t^p \leftarrow D_t^p - y_t$. We terminate for task schedule p once $\sum_{t \in \mathcal{T}} D_t^p = 0$ and continue with the next task schedule until all selected task schedules have been considered. The constraints of the shift SP remain unchanged. While we acknowledge that this procedure already provides valid solutions to our problem, we further improve the obtained solutions in the subsequent steps.

In the next step, we solve MP-IP. We rely exclusively on the selected high-quality task schedules. Regarding shift schedules, we include all schedules generated to find the lower bound, as well as the columns additionally generated by our heuristic. Preliminary testing has

shown that the resulting solution represents a good, feasible start solution and upper bound to the original MIP (1) to (19).

In the last step of our solution procedure, we optimize the start solution by applying the MIP presented in Subsection 3.3.2. We provide the MIP with the start solution that allows a warm start of the search procedure. In addition to the start solution, we can further speed up the solution by limiting the size of the set of workers $\mathcal{W}$ to the smallest required employee number of a feasible solution, i.e., the best known the upper bound. Furthermore, we provide the MIP with the lower bound and include it as an additional hard constraint, which allows us to cut-off the search tree in the branch-and-cut procedure of our standard solver. Compared to our benchmark, the MIP is significantly accelerated (see computational study in Subsection 3.5.2).
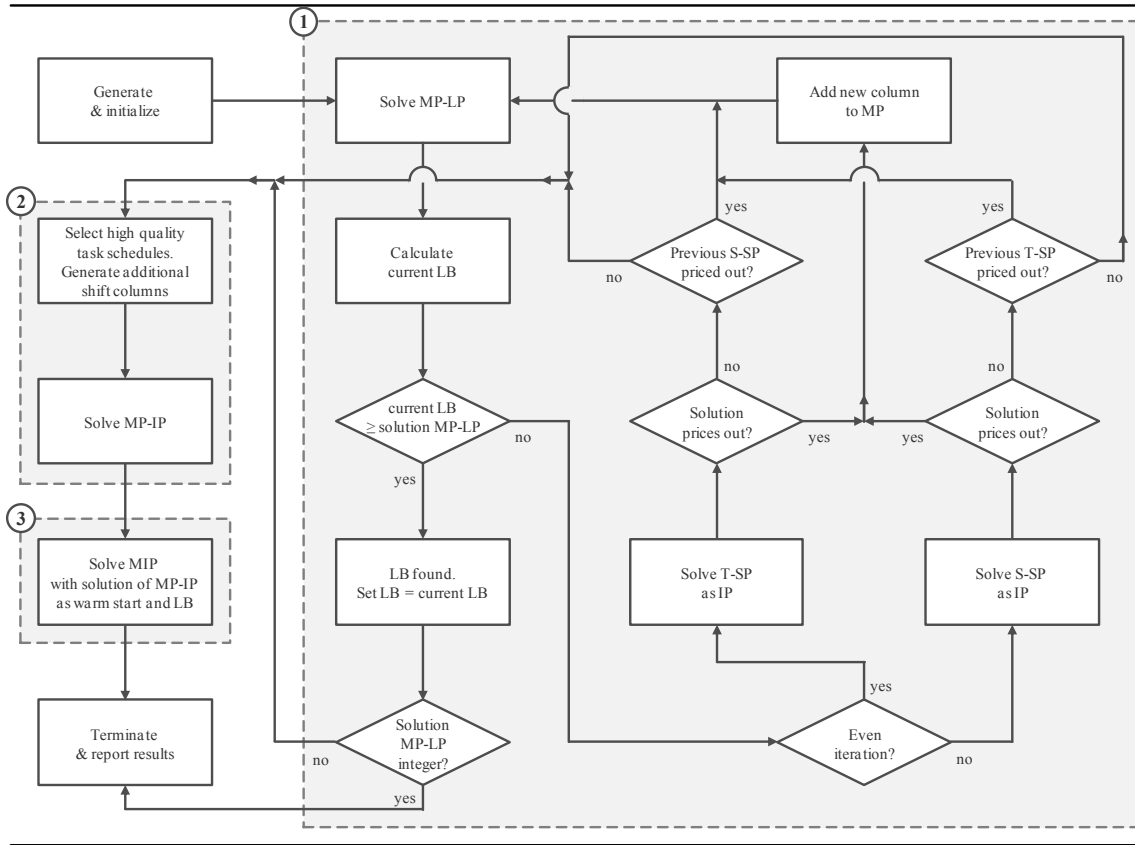
Figure 3.3: Illustrative flow chart of the solution approach

To summarize our solution approach, we provide an overview in Figure 3.3. The three major steps of the algorithm are indicated by the areas highlighted in gray. Step 1 comprises the column generation to find the lower bound. In each iteration, we check whether the lower bound has already been found. The objective values of the two subproblems are initialized with negative infinity, so that we can calculate the lower bound from the first iteration onwards. If the lower bound has been found, we terminate column generation early and proceed with step 2, i.e., the search for a good, feasible start solution. If not, we iteratively call the two subproblems

one after another. We call T-SP in every even iteration and S-SP in every odd iteration. When the respective subproblem prices out, we add the resulting column to MP and solve MP-LP. If both T-SP and S-SP no longer price out, the standard stopping criterion of the column generation is fulfilled. In this case, we proceed with step 2 of our algorithm and select the best lower bound until that point. In this step, we solve MP-IP with the selected high-quality task columns and all generated shift columns. In the final step 3, we improve our start solution by calling the MIP with our lower bound. A pseudo-code of our solution approach is shown in the appendix.

3.4.6 Alternative problem decomposition

Instead of decomposing the problem into MP, T-SP, and S-SP, which we call 2SP, one could also include the task scheduling in MP and only work with one subproblem, S-SP, called 1SP. This approach has been suggested in a previous publication by Maenhout & Vanhoucke (2016). The performance of the different decomposition approaches is compared in Subsection 3.5.4. Below, we present MP-IP for 1SP using the notation introduced previously.

$$\text{Min } z^{MP} = \sum_{j \in \mathcal{J}} C^{work} \lambda_j + \sum_{t \in \mathcal{T}} C^{out} z_t \tag{54}$$

subject to

$$\sum_{j \in \mathcal{J}} Y_{jt} \lambda_j - d_t + z_t \geq 0 \qquad\qquad \forall t \in \mathcal{T} \tag{55}$$

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} x_{at}^{start} = 1 \qquad\qquad \forall a \in \mathcal{A} \tag{56}$$

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} t x_{at}^{start} - \sum_{t=T_b^{earliest}}^{T_b^{latest}} t x_{bt}^{start} \geq L_b \qquad\qquad \begin{aligned} &\forall a \in \mathcal{A}, \\ &b \in \mathcal{A}_a^{pred} \neq \emptyset \end{aligned} \tag{57}$$

$$\sum_{a \in \mathcal{A}} R_a \left(\sum_{\tau = \max\{T_a^{earliest}; t-L_a+1\}}^{\min\{t; T_a^{latest}\}} x_{a\tau}^{start} \right) = d_t \qquad\qquad \forall t \in \mathcal{T} \tag{58}$$

$$x_{at}^{start} \in \{0,1\} \qquad\qquad \forall a \in \mathcal{A}, t \in \mathcal{T}_a^{start} \tag{59}$$

$$\lambda_j \in \mathbb{N}_0 \qquad\qquad \forall j \in \mathcal{J} \tag{60}$$

$$z_t \in \mathbb{N}_0 \qquad\qquad \forall t \in \mathcal{T} \tag{61}$$

$$d_t \in \mathbb{N}_0 \qquad\qquad\qquad \forall t \in \mathcal{T} \qquad\qquad (62)$$

3.5 Numerical study

In this section of the thesis, we present and discuss numerical results for the strategic workforce sizing problem. It is divided into four parts: First, we demonstrate the convergence behavior of our lower bound. Second, we illustrate the performance of our solution approach for real-world problem instances. To do so, we rely on three scenarios, i.e., small, medium, and large, which are based on real data and which differ in the number of tasks that need to be planned. We discuss the solution quality and solution times compared to solving the benchmark MIP described in Subsection 3.3.2 with a standard solver. Third, we assess the impact of flexibility on the required number of employees and runtimes. To do so, we rely on the three initial real-world instances and vary the degree of flexibility so that we obtain a total of 12 instances. Fourth, we investigate our solution approach based on a computational study with 36 additionally generated test instances and compare our problem decomposition into two subproblems with the alternative problem decomposition with one subproblem as presented in Subsection 3.4.6. To solve the linear and integer programs, we use standard solver IBM ILOG CPLEX Optimization Studio Version 12.6.1 on an i5-4300 bi-core CPU with 1.90 GHz, 64-bit, and 8 GB RAM, running on a Windows 7 operating system.

3.5.1 Early termination of column generation

In this subsection, we provide a comparison between our introduced lower bound stopping criterion and the convergence of the MP-LP objective value. When using our lower bound, we are able to terminate column generation early once the lower bound value z_i^{LB} exceeds the current objective value z_i of MP-LP. Without this lower bound, the column generation process terminates whenever the subproblems no longer price out.

We illustrate the convergence behavior for our medium-sized problem instance of 200 tasks (see Subsection 3.5.2). The convergence behavior is depicted in Figure 3.4. The horizontal axis represents the number of column generation iterations, while the vertical axis represents the MP-LP objective value z_i and the current value of the lower bound at iteration i. Applying the lower bound stopping criterion, the column generation can be terminated significantly earlier than when the regular stopping criterion is used. In this particular instance, we terminate the column generation after 109 iterations (2 minutes) compared to 620 iterations (22 minutes); this equals 18% in terms of iterations or 9% in terms of time compared to the regular stopping criterion. A comparable behavior is observed for the other instances: The small problem instance terminates after 120 iterations instead of 130 (92%), while the large instance terminates

after 257 iterations compared to 448 iterations (57%) of performing the entire column generation procedure. As later iterations consume more solution time, the positive effect is even higher in terms of time required to solve the problem than in terms of iterations.
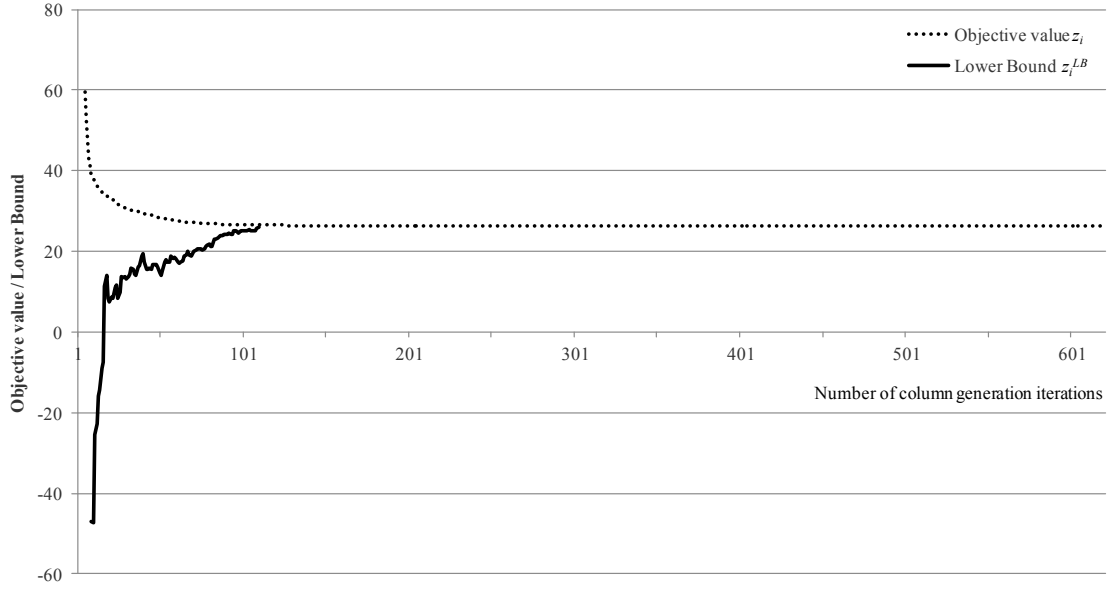
Figure 3.4: Convergence of objective value and lower bound

3.5.2 Investigation of real-world cases

In the following, we provide a comparison of the solution quality and time between our solution approach and a benchmark. The benchmark is the MIP formulation as presented in Subsection 3.3.2, which is solved with the standard solver CPLEX.

Parameters. The planning horizon of our model is one week, which is in line with the logistics task scheduling horizon of our real-world data. One time period represents 15 minutes, resulting in a planning horizon of 672 time periods. We use three problem instances, which are based on real data from the pediatric clinic of a large German hospital with 200 beds and 8 care units. Based on the available data, we generated three test instances, namely small, medium, and large. The task number has been scaled in order to represent different problem size settings. The small instance reflects the size of the original pediatric clinic in terms of required employees. The medium-sized test instance, in contrast, could be an average hospital, while the large instance reflects a large hospital.

When translating the real data into our test instances, we identified two task types. First, we find *day-long tasks* that last eight hours, i.e., a typical working day. These could include, for example, background tasks like refilling stock in the central warehouse. The second task type

represents *flexible tasks*, which have rather short start windows, such as food preparation or disposal. Such tasks are linked to other logistics processes within the hospital. This category also comprises tasks that can be performed over a longer timeframe, for example cleaning or refilling stock at wards. Consequently, tasks vary in their duration and in the size of their start windows. While our model allows more than one resource to be allocated to each task a, we set the resource demand of each task $R_a = 1$. An overview of the tasks is provided in Table 3.2.

In our base case, we set the shift parameters as follows: The minimum shift length $T^{shiftMin}$ is set to 16, which equals 4 hours, while the maximum shift length $T^{shiftMax}$ is 40, i.e., 10 hours. The rest time between two consecutive shifts T^{rest} is 48, which accounts for 12 hours. Over the entire planning horizon, we limit the maximum number of working periods $T^{workMax}$ to 160, which equals 40 hours and is in line with German labor contracts. The cost parameters, i.e., the cost of employing logistics assistants and the cost of demand undercoverage, are set such that it is always preferable to employ logistics assistants. We set the relative cost of employing logistics assistants C^{work} to 1, while the cost of demand undercoverage C^{out} is set to 1,000.

Table 3.2: Problem instances

Instance		Small	Medium	Large
Day-long tasks				
Tasks	*[#]*	67	119	140
Number of precedence relations	*[#]*	57	102	120
Average length	*[Hours]*	8.0	8.0	8.0
Flexible tasks				
Tasks	*[#]*	63	81	360
Number of precedence relations	*[#]*	80	112	437
Average length	*[Hours]*	0.8	0.9	1.0
Total				
Tasks	*[#]*	130	200	500
Number of precedence relations	*[#]*	137	214	557
Average length	*[Hours]*	4.5	5.1	2.9

The parameters of our benchmark instances are the same. However, unlike in our solution approach, we need to hand the number of available workers to the MIP. The cardinal number W of the set of workers heavily influences the solution time of our benchmark, as it impacts the number of variables. When choosing W, we aim to provide the MIP with a realistic number of available workers to avoid making the benchmark too slow. Additionally, we may not build on information that we do not have. Consequently, we select W as follows: For all tasks, we multiply their resource demand and length and divide the resulting number by the minimum shift length multiplied with a reasonable number of working days (5), which ensures that a sufficient number of workers are available. Summarizing, we calculate the initial number of workers $W = \sum_{a \in \mathcal{A}} L_a R_a / (5 * T^{shiftMin})$. This results in 30 workers for the small instance, 52

for the medium instance, and 74 workers for the large problem instance. Please note that there might be potential problem instances where this approach is not an upper bound.

Performance. We present the performance of our solution approach in the following. We compare the results and the runtime of our solution procedure with the MIP as presented in Subsection 3.3.2, solved with the standard solver. We limit the computation time for the benchmark instances to 10 hours. Once CPLEX is unable to find the optimal solution within this time limit, we provide the best solution found until then.

Table 3.3: Computational results

Instance		Small	Medium	Large
Our solution appraoch				
Runtime to find lower bound	*[Min.]*	2	2	12
Runtime to generate additional shift columns	*[Min.]*	1	2	7
Runtime to solve MP-IP	*[Min.]*	0	0	1
Runtime of "warm start" MIP	*[Min.]*	1	2	39
Total runtime	***[Min.]***	**4**	**6**	**59**
Lower bound: Column generation	*[# workers]*	16	27	37
Upper bound: Start solution of MIP	*[# workers]*	17	29	47
Final solution: MIP solved to optimality	***[# workers]***	**16**	**27**	**37**
Benchmark				
Runtime (interrupted after 600 minutes)	*[Min.]*	600	600	600
Best solution in runtime	*[# workers]*	17	29	36*

(*) plus 2,441 time periods of under coverage

The performance of our solution approach is presented in Table 3.3. For all instances, we compare the runtime and the objective value, and provide information on the intermediate steps. The total runtime of our solution approach comprises the three individual steps presented above. First, we present the runtime required to find the lower bound. For the small and medium problem instances, we find the lower bound within 2 minutes, while the large instance requires 12 minutes. We generate a total of 869, 777, and 1,810 columns until we find the lower bound for the small, medium, and large instance, respectively. Creating additional shift schedules based on the updated S-SP takes another 1 to 7 minutes, depending on the problem size. We heuristically generate an additional 1,419, 2,187, and 8,770 shift columns. The number of task columns is a major driver of the complexity of solving the MP-IP. As we reduce their number to 10% of all generated columns (max. 15), MP-IP is solved especially quickly and is performed within a maximum of one minute. Solving the warm-started MIP to optimality for the small and medium-sized instances also runs within just a few minutes, while it takes 39 minutes for the large instance. Note that for nights, we force shifts y_t and tasks x_t to zero between 6:15 p.m. and 6:00 a.m., where no logistics tasks need to be performed. Further, we only allow a maximum of five working days per week and one shift per day. Summarizing the performance

of our solution approach, all problem instances are solved to optimality within one hour. In all three examples, the integer solution equals the lower bound that we obtained from column generation, which demonstrates the strength of this lower bound. However, this observation might not be valid in general. As we are discussing a strategic planning problem, the presented solution times can be reasonably tolerated. Furthermore, the small and medium-sized instances are even solved within four or six minutes, respectively. In the final result, the average/minimum/maximum number of working hours per employee is as follows: small instance: 38.9/35.0/40.0, medium instance: 39.0/33.5/40.0, and large instance: 39.9/39.0/40.0. In order to schedule only 40-hour shifts, one could assign additional on-periods to employees without negatively impacting the objective value.

Comparing our solution procedure to the benchmark MIP, we demonstrate that our solution is superior in terms of both solution quality and runtime. Within the time limit of 10 hours, none of our test instances is solved to optimality. However, the solution of the small and medium instance is only one or two workers larger than the optimum. The solution of the large problem instance includes hypothetical demand undercoverage: The total objective value is 2,441,036, which comprises 36 scheduled logistics assistants and 2,441 time periods of uncovered demand. The MIP gap is 13.4%, 11.8%, and 100% for the small, medium, and large instance, respectively.

3.5.3   The value of flexibility in shifts and tasks

In this subsection, we assess the value of flexibility in our model. Flexibility is incorporated in our model in a twofold manner. First, tasks can be scheduled flexibly, provided they are started within given start time windows and fulfill the precedence relations. Second, we allow a high degree of flexibility in shifts through flexible start times and shift durations. In order to evaluate how flexibility impacts the number of required workers, we vary the degree of flexibility in both tasks and shifts. We build our analysis on the three instances developed in Subsection 3.5.2. In the base case presented there, both task schedules and shift schedules are flexible. In the upcoming analysis, we reduce the degree of flexibility for shifts only, for tasks only, and for both.

In order to reduce the degree of flexibility of shifts, we allow only two shift types of the same length. In particular, we fix the shift length to eight hours. Furthermore, we allow only two start times each day, namely 6:00 a.m. and 10:15 a.m. These two shift types allow demand to be covered at any point in time for our test instances. Having two shift types – one morning shift and one late shift – is in line with what we see in practice. For the task side, we reduce the degree of flexibility by significantly limiting the task start time window. In particular, we leave

$T_a^{earliest}$ unchanged for all tasks, but set T_a^{latest} at half the length of the original start window. Consequently, the start windows are reduced by 50%.

A summary is shown in Figure 3.5. We provide the number of required workers for all three problem instances (large, medium, and small) when we solve them to optimality, as well as the utilization. Utilization is defined as the sum of all periods of the selected task schedule demand divided by the sum of all periods of the selected shift schedules, i.e., the resource supply. For both tasks and shifts, we compare the flexible case to the non-flexible or fixed case, which results in four quadrants of the presented matrix. Of the four cases, the upper right FULL-FLEX reflects our base case. Here, we see the lowest number of required workers across all three test instances. In the bottom left quadrant of our matrix, we show the NO-FLEX, which yields the worst results.
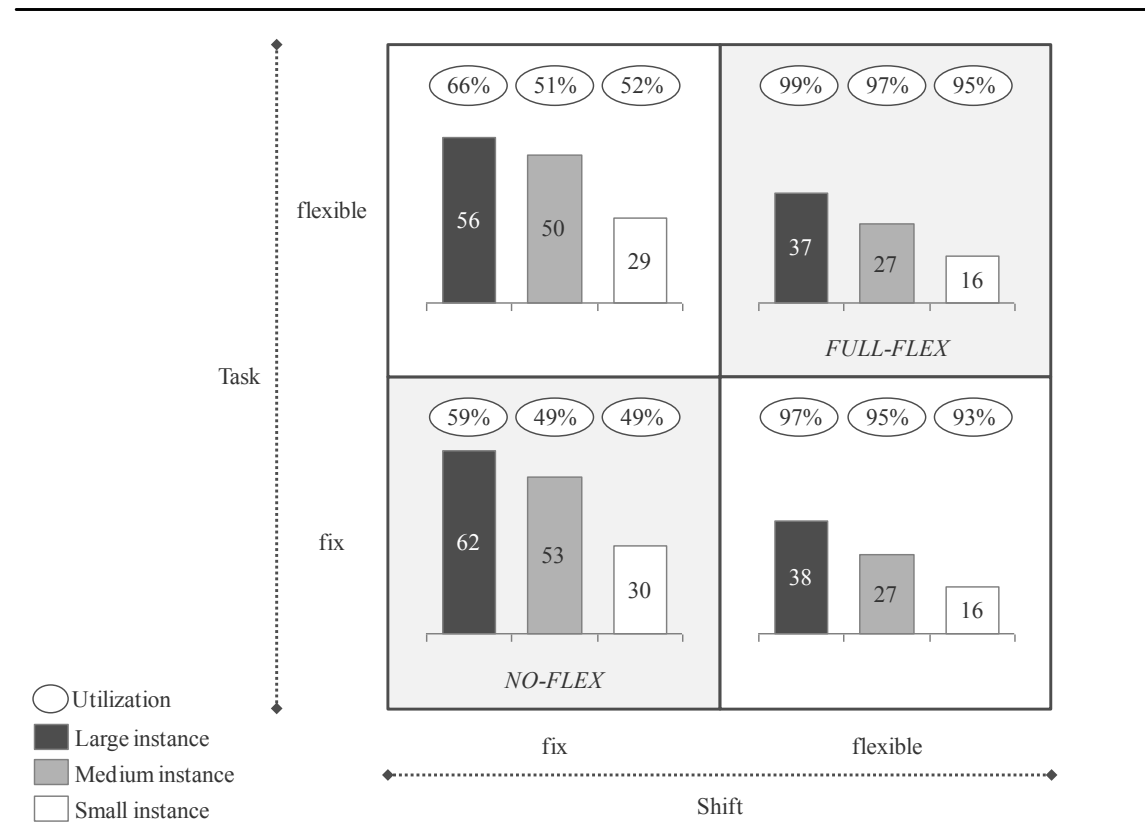
Figure 3.5: Number of required workers with different levels of flexibility

We first discuss the results of the large instance. In NO-FLEX, we need 62 workers to cover our demand. Increasing the level of flexibility in tasks and shifts reduces the required employees to 56 (-10%) and 38 (-39%), respectively. In FULL-FLEX, the optimum is 37 workers, which equals a 40% reduction compared to NO-FLEX. In the medium instance, we see the same overall direction. In the no-flexibility case, 53 workers are required, compared to 50 (-6%) in the case of flexible tasks, or 27 (-49%) in the case of flexible shifts. In FULL-FLEX, 27

workers are needed as well. The results of the small instance are similar, with a starting value of 30 workers in the non flexible case. By increasing task flexibility, this number can be decreased to 29 (-3%), or 16 (-47%) when increasing shift flexibility. The joint minimum is 16, which equals a total reduction potential of 47%. Increasing flexibility lowers the required number of workers but leads to a higher utilization of employees, which results in higher workloads and shorter overall rest times.

A summary of the computational complexity of the different flexibility scenarios is provided in Table 3.4. We draw two major conclusions from the analysis. First, we underline that our solution approach is superior to the benchmark CPLEX in all problem instances. Especially scenarios with a high degree of flexibility are not solvable in a straightforward approach. Second, we observe that flexibility increases the complexity of our model and leads to longer runtimes. For scenarios with a low degree of flexibility, however, CPLEX can be an option to find optimal solutions within the time limit of 10 hours.

From a managerial perspective, there is significant cost reduction potential in increasing the flexibility of shifts and tasks. In our case, the major cost reduction lever is an increase in shift flexibility. However, we have to note that in the no-flexibility shift case, the day-long tasks need two shifts. Consequently, one could argue that, in practice, different shift schedules would be applied. Allowing full flexibility in shifts comes with some downsides. A major drawback is the negative effect on employee satisfaction, as shift patterns come with a high degree of variance in terms of both shift lengths and daily start times. Additionally, employee preferences for certain shift types or working periods are not reflected. A high variability in shifts also leads to higher organizational flexibility and a stronger need to provide a high level of transparency on shift schedules. Also, in the flexible shift framework it is possible that not all employees work the same number of hours. Measures need to be defined to reintroduce the element of fairness.

Table 3.4: Runtimes of different flexibility scenarios

Instance		NO-FLEX SMA	MED	LAR	TASK-FLEX / SHIFT-FIX SMA	MED	LAR	TASK-FIX / SHIFT-FLEX SMA	MED	LAR	FULL-FLEX SMA	MED	LAR
Our solution appraoch													
Runtime (total)	[m.]	1	1	5	1	7	23	21	16	62	4	6	59
Best solution in runtime	[#w]	30	53	62	29	50	56	16	27	38	16	27	37
Benchmark													
Runtime (int. aft. 600 m.)	[m.]	32	301	283	26	600	243	600	600	600	600	600	600
Best solution in runtime	[#w]	30	53	62	29	51	56	17	28	1,399,036	17	29	2,441,036

In summary, we see significant cost reduction potential both from increasing task flexibility and shift flexibility. If we assume that NO-FLEX represents the status quo, there is significant improvement potential. In order to overcome disadvantages in terms of employee satisfaction and organizational complexity, one could reduce the level of flexibility by allowing only

predefined shift patterns for certain employees (or employee groups). Our approach may be used to identify promising shift patterns.

3.5.4   Investigation of the solution approach

In the following subsection, we compare the solution approach based on decomposing the initial problem into two subproblems, T-SP and S-SP, with the decomposition with only one subproblem type S-SP (see Subsection 3.4.6). We demonstrate the applicability of our solution approach by solving 36 additional test instances. We start by elaborating on problem instance generation and discuss the results of the numerical study thereafter.

Generating test instances. The real-world instances presented in Subsection 3.5.2 are the starting point for the generation of additional test instances. In line with the real-world instances, we distinguish between three problem sizes, namely small, medium, and large. In order to distinguish the real-world instances from the additionally generated instances, we call the latter SMA, MED, and LAR. They differ in the man-hours that are required to fulfill all tasks, not respecting shift patterns. The problem sizes largely reflect the real-world problem instances and account for:

- SMA: 600 man-hours per week,
- MED: 1,000 man-hours per week,
- LAR: 1,400 man-hours per week.

Next, we allocate the man-hours to working days. The work is distributed equally from Monday to Friday, while we allocate only 70% of a regular week-day demand to Saturdays and Sundays, respectively. As a result, we obtain the man-hours per day. We then allocate the man-hours per day to task types. Unlike the real-world instances, where we distinguish between two task types, we distinguish between three task types in the newly generated test instances. This approach helps us to achieve transparency on the effects of time window and precedence constraints on the performance of the algorithm. The task types are *day-long tasks*, *peak tasks*, and *precedence tasks*. The characteristics of each task type are shown in Table 3.5.

Table 3.5: Task type characteristics

Task type		Earliest start time	Latest start time	Minimum length	Maximum length	Resource demand	Precedence constraints
(1) Day-long task		6h	10h	6h	8h	1	NO
(2) Peak task	-early	6h	8h	1h	2h	1	NO
	-mid day	10h	12h	1h	2h	1	NO
	-late	14h	16h	1h	2h	1	NO
(3) Precedence task		Mon, 6h	Fri, 13h	2h	4h	1	YES

We generate three task distribution scenarios S1, S2, and S3 to allocate the man-hours per day to task types. S1 is the base case with 2% precedence tasks, 83% day-long tasks, and 15% peak tasks. This largely reflects the averages of our real-world data instances. In S2, the share of precedence tasks is tripled, while in S3, the share of peak tasks is tripled. The man-hours allocated to peak tasks are distributed equally to three time windows during the day (see Table 3.5). In the subsequent step, we replace the generated man-hours with concrete tasks. Their length is randomly generated based on a uniform distribution within their minimum and maximum length. Once the defined man-hours are filled with concrete tasks, we stop. With this approach, we slightly overshoot the defined working hours for our scenarios. Precedence tasks have a large start window but are subject to precedence constraints. In line with the real-world instances, they are assumed to be parallel-serial task chains of varying sizes between two and five. The described approach for generating test instances results in nine task problem types. We double their number by reducing the start time windows by 50% with fixed earliest start times. The original start windows are referred to as large window (LW) instances, while the instances with smaller start time windows are named SW. Moreover, we assess flexible shifts (FL) and fixed shifts (FX). In fixed shifts, we allow only one shift length, namely eight hours, and two start periods, one morning and one late shift. Summarizing, we generate 36 additional test instances that resemble real-world behavior.

Table 3.6: Results of the computational study

Instance	2SP								1SP						
	t_{lb} [m.]	t_{asc} [m.]	t_{ip} [m.]	t_{mip} [m.]	t_{total} [m.]	LB [#w]	UB [#w]	Fin sol [#w]	t_{lb} [m.]	t_{ip} [m.]	t_{mip} [m.]	t_{total} [m.]	LB [#w]	UB [#w]	Fin sol [#w]
SMA_S1_LW_FL	2	1	0	1	4	18	20	18	1	0	-	1	18	18	18
SMA_S1_LW_FX	16	1	0	0	18	18	24	18	0	0	-	0	18	18	18
SMA_S1_SW_FL	1	1	0	-	1	20	20	20	1	0	-	1	20	20	20
SMA_S1_SW_FX	1	0	0	0	1	20	26	20	0	0	-	0	20	20	20
SMA_S2_LW_FL	3	1	0	2	6	17	19	17	1	0	-	1	17	17	17
SMA_S2_LW_FX	30	2	0	0	32	17	23	18	0	0	-	0	18	18	18
SMA_S2_SW_FL	1	1	0	1	2	19	20	19	1	0	-	1	19	19	19
SMA_S2_SW_FX	4	2	0	0	6	19	23	19	0	0	-	0	19	19	19
SMA_S3_LW_FL	9	2	0	3	15	17	21	17	1	1	2	4	17	18	17
SMA_S3_LW_FX	30	2	0	0	32	18	27	19	0	0	0	0	18	19	19
SMA_S3_SW_FL	1	1	0	1	2	22	24	22	1	0	-	1	22	22	22
SMA_S3_SW_FX	8	2	0	0	10	24	32	24	0	0	-	0	24	24	24
MED_S1_LW_FL	2	1	0	177	180	28	32	29	1	0	179	180	28	29	29
MED_S1_LW_FX	30	2	0	0	32	28	37	29	0	0	-	0	29	29	29
MED_S1_SW_FL	1	1	0	1	2	32	33	32	1	0	-	1	32	32	32
MED_S1_SW_FX	3	2	0	0	6	33	39	33	0	0	-	0	33	33	33
MED_S2_LW_FL	5	1	0	4	10	27	30	27	1	1	6	8	27	28	27
MED_S2_LW_FX	30	2	0	0	32	27	35	28	0	0	-	0	28	28	28
MED_S2_SW_FL	1	1	0	2	5	29	31	29	1	0	-	1	29	29	29
MED_S2_SW_FX	9	2	0	0	11	30	36	30	0	0	-	0	30	30	30
MED_S3_LW_FL	11	4	1	37	53	27	35	27	1	1	6	8	27	28	27
MED_S3_LW_FX	30	3	0	0	33	29	48	31	0	0	180	180	30	31	31
MED_S3_SW_FL	2	1	0	1	3	34	37	34	1	0	-	1	34	34	34
MED_S3_SW_FX	4	3	0	0	7	40	54	40	0	0	-	0	40	40	40
LAR_S1_LW_FL	3	2	0	175	180	38	44	39	1	0	179	180	38	39	39
LAR_S1_LW_FX	30	7	0	0	37	38	49	40	0	0	0	1	39	40	40
LAR_S1_SW_FL	1	1	0	2	4	43	44	43	1	0	-	1	43	43	43
LAR_S1_SW_FX	5	3	0	0	8	44	53	44	0	0	-	0	44	44	44
LAR_S2_LW_FL	1	1	0	3	5	58	63	58	0	0	-	0	58	58	58
LAR_S2_LW_FX	2	2	0	0	5	58	68	58	0	0	-	0	58	58	58
LAR_S2_SW_FL	1	0	0	-	2	65	65	65	0	0	-	0	65	65	65
LAR_S2_SW_FX	1	0	0	0	1	65	82	65	0	0	-	0	65	65	65
LAR_S3_LW_FL	14	8	1	24	46	37	48	37	1	1	178	180	37	38	38
LAR_S3_LW_FX	30	6	0	0	36	38	65	41	1	0	179	180	40	41	41
LAR_S3_SW_FL	2	1	0	2	5	49	51	49	1	0	-	2	49	49	49
LAR_S3_SW_FX	6	5	0	0	11	54	74	54	1	0	-	1	54	54	54

Interpretation of results. The results of the computational analysis are depicted in Table 3.6. We compare the time to find the lower bound t_{lb}, the time to generate additional shift columns t_{asc}, the time to solve MP-IP t_{ip}, and the time to solve the warm-started MIP t_{mip}. We distinguish between the lower bound (LB), the upper bound (UB) from solving MP-IP and the final solution. Cases in which the final step of the solution approach, i.e., solving the MIP, is not required because an optimal solution has already been found are indicated by "-" in the table. Solution times of "0" indicate runtimes of less than 30 seconds. In our study, we limit the total solution time to three hours and compare the best objective value that we obtained within this time. The column generation time is limited to 30 minutes.
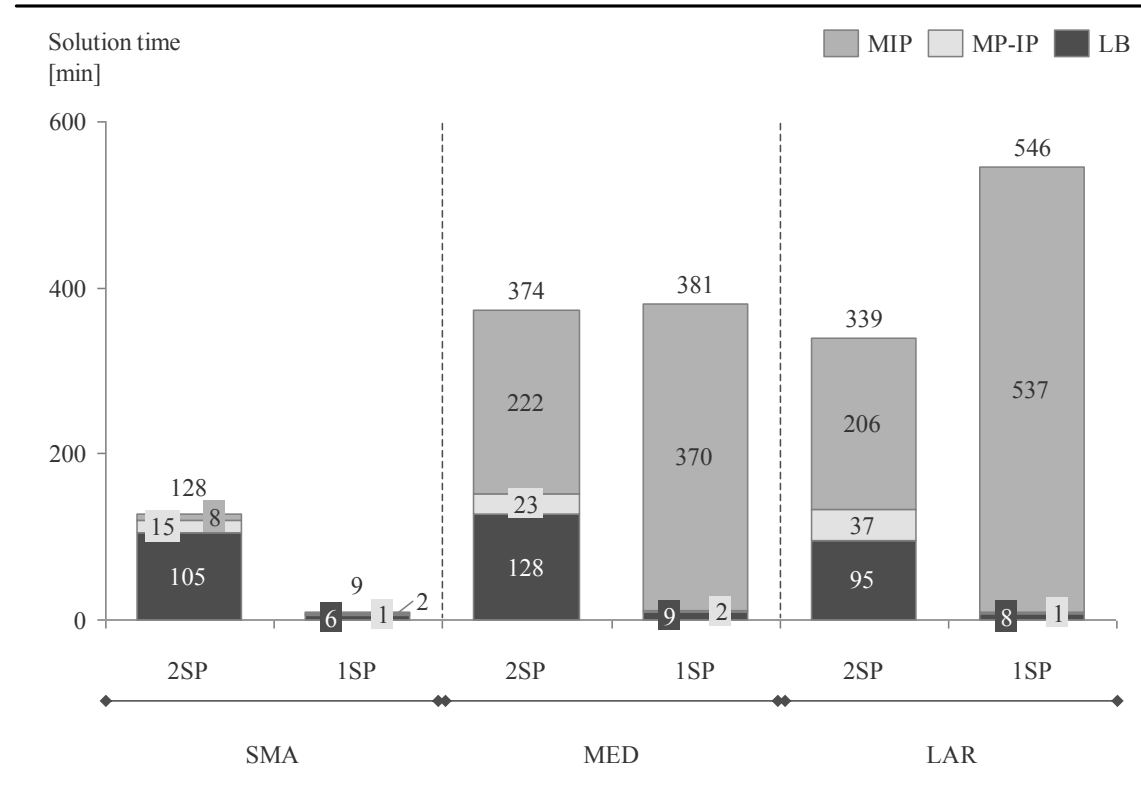
Figure 3.6: Runtimes

2SP and 1SP achieve the same final solutions for the small and medium-sized test instances. For the large cases, however, one instance cannot be solved to optimality within three hours by 1SP, while 2SP achieves the optimal solution. Although in theory 2SP should achieve the same or better lower bounds than 1SP, 1SP yields superior lower bounds in six cases. This can be explained by the time limit of 30 minutes, which 2SP exceeds in those instances. All cases where column generation of 2SP is stopped due to the time limit belong to the "LW_FX" instances, namely large time windows in T-SP and fixed shifts. In 1SP, column generation stops in most cases as all columns are priced out, while 2SP most often terminates due to the lower bound stopping criterion as introduced in Subsection 3.4.4. Solving MP-IP results in the optimal solution for the majority of S1 instances, so the step of solving the warm-started MIP can be skipped. This behavior is observed in 10 of 12 small instances, 8 of 12 medium-sized instances, and 8 of 12 large instances.

Runtimes for the different test instance sizes for 2SP and 1SP are displayed in Figure 3.6. In the first part of the algorithm, i.e., finding lower bounds, 1SP outperforms 2SP. In the second step, i.e., solving MP-IP and adding new shift columns for 2SP, we see slight advantages for 1SP as well. However, due to the long runtimes of the warm-started MIP, improving the solution quality and proving optimality takes very long in 1SP. Consequently, for large problem instances, 2SP outperforms 1SP in terms of solution runtimes.

3.6 Conclusion

In this chapter, we introduce a new model that combines shift scheduling with task scheduling. Our goal is to determine the optimal, i.e., minimum number of workers required for given task requirements. To do so, we integrate flexible shift scheduling with task scheduling. In order to solve our model efficiently, we present an exact column generation-based solution approach. As part of this approach, we develop a lower bound for staff minimization problems with an unknown number of workers to terminate column generation early. A focus of our model is the incorporation of flexibility in terms of both shift scheduling and task scheduling. We test our solution approach for scheduling logistics assistants in a large case hospital and generate artificial problem instances based on real-world data. All real-world test instances are solved to optimality within a reasonable amount of time, which clearly outperforms a benchmark solved with CPLEX. Based on the real-world test instances, we demonstrate that by incorporating flexibility, efficiency gains of 40 to 49% are possible, compared to a case with reduced flexibility. Also, we randomly generate additional test instances and demonstrate the validity of our solution approach. While we apply our model to a specific use case in the healthcare industry, we believe that our model and algorithmic approach are limited neither to this type of employee nor to this industry, but have a wide range of potential application areas, e.g., in production support functions or in the service industry (see Subsection 3.3.1).

This chapter of the thesis at hand offers several opportunities for future research in terms of both model extensions and methodology. Regarding model extensions, one option is to better incorporate demand undercoverage, which for now is a modeling device only. We set cost parameters in such a way that, in an optimal solution, demand undercoverage does not appear. To extend the model, apart from the extensions discussed in Subsection 3.3.3, one could, however, include a second type of employee, e.g., nurses that can step in when the number of logistics assistants is not sufficient to perform all tasks. This would further strengthen the practical applicability of our model and would allow us to include one more source of flexibility. Another potential model extension is the inclusion of stochasticity in task scheduling. Currently, we assume that all tasks are fixed in length; however, in some applications, their length can vary. Apart from extending our model, it could also be worthwhile to reformulate the subproblems so that they are solvable even faster, allowing more sophisticated solution approaches such as branch-and-price algorithms (see Chapter 4). In this case, it might be possible that the decomposition into two subproblems is faster than when only one subproblem is considered, as once shifts have been fixed through branching, the resulting master problem incorporates a RCPSP.

Summarizing our findings, we conclude that there is considerable cost savings potential in simultaneously scheduling shifts and tasks of logistics assistants in hospitals and leveraging the incorporated flexibility.

4 A branch-and-price approach for tactical shift and task scheduling

After addressing the strategic workforce sizing problem in Chapter 3, we move to the tactical shift and task scheduling problem in this chapter. The problem is solved with a branch-and-price algorithm.

4.1 Introduction

Shift rostering is usually considered an iterative process. It starts with *demand modeling,* which comprises the deduction of workforce demand based on predicted incident patterns. In the next step, *days-off scheduling*, days-off and days-on are allocated to different lines of work. *Shift scheduling* comprises the definition of the type and number of shifts that are required to fulfill work demand in each time period of the planning horizon. In a traditional, explicit approach, shifts include, for example, morning shifts, late shifts, and night shifts. In the last step, *line of work construction,* individual workable rosters are created for each employee that cover the entire planning horizon, which mostly comprises between one and up to several consecutive weeks (Ernst et al. 2004). The process is laid out in Figure 4.1.

Shift rostering steps	Outcome
1 Demand modeling	Demand profile: Number of staff needed in each time period of the planning horizon
2 Days-off scheduling	Distribution of rest days and working days for different lines of work
3 Shift scheduling	Number of shift (types) required to fulfill the demand
4 Line of work construction	Shift roster for each staff member
...	
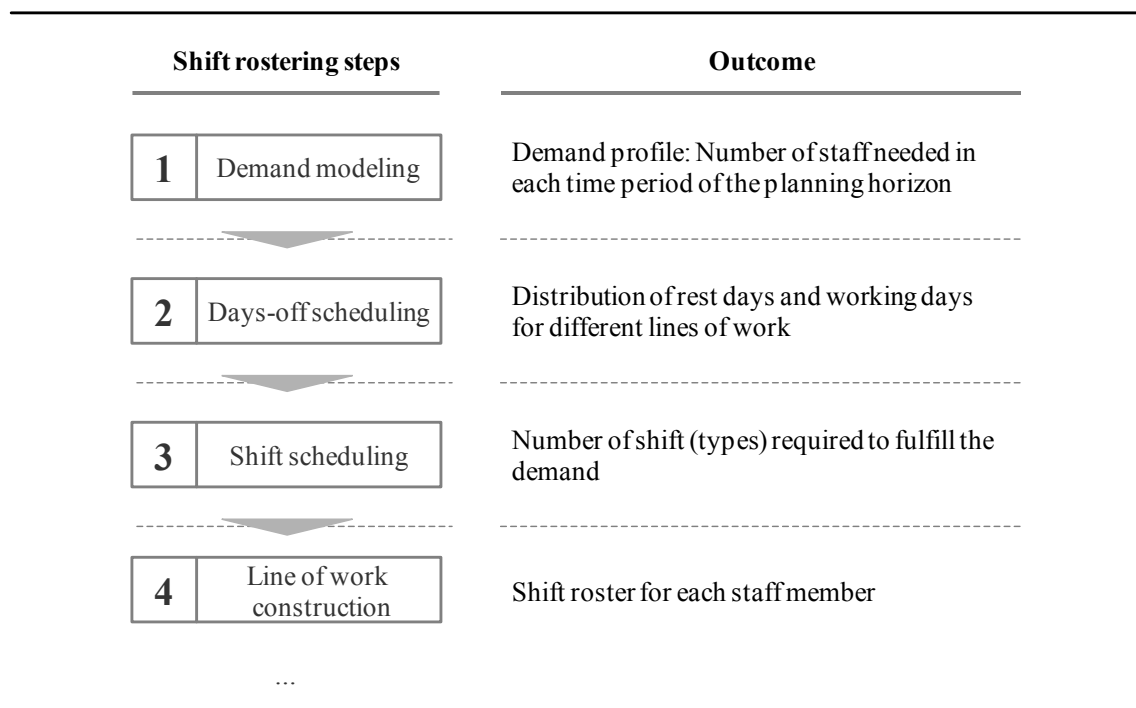
Figure 4.1: Shift rostering according to Ernst et al. (2004)

In demand modeling, the literature suggests three alternative incident categories that result in demand, namely *task-based demand*, *flexible demand*, and *shift-based demand*. In the first category, task-based demand, the respective demand results from predefined tasks that have to be completed. The tasks usually have some inherent flexibility, for example, as task start times

are subject to certain start time windows. Demand is forecast based on the collection and use of historical data in order to define the staffing levels required to fulfill the task requirements. Flexible demand forecasting is applied in environments where demand is subject to a high degree of stochasticity. Consequently, the likelihood of incidents is incorporated. In the latter demand category, shift-based demand, demand is derived from the number of staff that has to be on duty in certain time periods, for example hospital nurses that have to be present in care units (Ernst et al. 2004).

This work addresses a task-based demand setting, in which demand results from the tasks that have to be performed in a certain time period. However, the setting of this work differs from the presented standard hierarchical scheduling process. Instead of assuming an iterative process in order to develop individual lines of work, we perform all four steps in an integrated manner. By doing so, we make use of the flexibility incorporated in shift and task scheduling. Our aim is to reduce inefficiencies, i.e., oversupply or undersupply of work at certain time periods that might occur once iterative planning is performed. In line with Chapter 3, we were confronted with this problem while introducing logistics assistants in our partner hospital. In the previous chapter, we introduce the basic model and a column generation-based solution procedure for solving the strategic workforce sizing problem in an integrated manner. In this chapter, we move from strategic sizing to tactical scheduling, which comprises the definition of shift and task schedules for a given number of employees. In this tactical problem, we allow varying degrees of shift flexibility for workers with different shift parameters and different costs. The contribution of this chapter is the introduction of a branch-and-price approach to optimally solve the tactical shift and task scheduling problem. In the introduction of the algorithm, we present a new network flow (NF) formulation of the flexible shift scheduling problem that is solved with a shortest path (ShP)-labeling algorithm.

The rest of this chapter is structured as follows: In Section 4.2, we provide a brief overview of the literature on branch-and-price approaches for integrated staff and task scheduling problems. We provide a detailed description of the problem in Section 4.3, followed by a presentation of the problem formulation and the solution approach in Section 4.4. A numerical study in Section 4.5 evaluates the solution procedure. We conclude with a brief discussion and point out future research opportunities in Section 4.6.

4.2 Literature review

This work integrates two research streams, namely flexible shift scheduling and task scheduling, and presents a branch-and-price solution algorithm. In order to provide an overview of the most relevant literature, we limit the review to publications applying a branch-and-price approach and

follow a two-step process. First, we present the literature that relies on branch-and-price approaches to solve shift scheduling problems in the healthcare industry. Second, we discuss integrated approaches for shift and task scheduling applying a branch-and-price approach. In the latter segment, we extend the scope to publications beyond healthcare. For a broader literature overview on the combination of shift and task scheduling, as well as the individual research areas, we refer to Section 3.2.

Shift scheduling in healthcare relying on branch-and-price approaches. Branch-and-price-based shift scheduling is applied to three employee groups in the healthcare industry: nurses, physicians, and hospital trainees. Branch-and-price-based nurse scheduling was initially performed by Jaumard et al. (1998), who generate shift schedules with a pricing subproblem (SP) that is formulated as a resource-constrained ShP problem. In their master problem (MP), they ensure that work demand is met while costs are minimized and nurses' preferences are maximized considering team balancing. A comparable approach is undertaken by Purnomo & Bard (2007), who balance individual preferences while minimizing personnel costs. They focus on creating cyclic schedules and evaluate several branching strategies. The authors apply a well-performing rounding heuristic to find integer solutions in early stages. Maenhout & Vanhoucke (2010) focus on the incorporation of multiple objectives and assess several branching and pruning strategies for nurse scheduling problems. In a further work by the same authors, the solution approach is extended to incorporate nurse staffing and allocation to several hospital departments (Maenhout & Vanhoucke 2013). A branch-and-price approach in physician scheduling is applied by Brunner et al. (2010). The authors investigate two different branching strategies: master variable branching and subvariable branching. Beliën & Demeulemeester (2006, 2007) study the scheduling of hospital trainees in two closely related publications. In contrast to the other approaches, the authors decompose their problem along activities that have to be fulfilled rather than staff members. In the latter publication, a comparison between staff-decomposed and activity-decomposed solution approaches is provided.

Branch-and-price approaches for integrated shift and task scheduling. The literature on integrated shift and task scheduling is scarce. Only two publications are identified that present integrated models and apply branch-and-price approaches. As the two publications have already been presented in Section 3.2, we focus on their solution approach and work out similarities to and differences from our branch-and-price algorithm.

Beliën & Demeulemeester (2008) propose an integrated model for master surgery scheduling and nurse scheduling. Their basic idea resembles our approach, namely that work demand is not fixed but can be altered in order to achieve demand profiles that have a good fit with the generated shift schedules. In their work, resource demand for nurses is the result of the master

surgery schedule. In line with the decomposition applied in this work, the authors decompose their problem into two SPs. The surgery schedules are determined in one SP which is formulated as an integer program (IP) and solved with a standard solver. In this surgery scheduling SP, resource blocks are assigned to surgeons subject to capacity constraints. This differs from our work, as we derive demand from task scheduling subject to start time windows and precedence constraints. The second SP generates shift schedules and is formulated as a dynamic program. Shift scheduling relies on days-on and days-off scheduling based on explicit shifts. In our work, we schedule shifts implicitly, which allows us to incorporate a higher degree of flexibility. In branching, Beliën & Demeulemeester (2008) only consider driving the surgery scheduling part to integrality, while driving shift schedule solutions to integrality is not considered. In the subsequent chapter, we apply a branching logic that addresses both shift and task scheduling.

Maenhout & Vanhoucke (2016) present a project scheduling approach that they combine with days-on and days-off shift scheduling. Their general approach resembles the ideas of the previously presented work and our approach in that resource demand as a result of task scheduling is not fixed but can be integrated with shift scheduling. The shift scheduling part is modeled as an IP and is limited to days-on and days-off scheduling. In this chapter, we present a NF formulation that solves the shift scheduling SP to optimality in a significantly shorter time than the IP formulation. Moreover, our model allows a very high degree of flexibility by relying on implicit shift formulations. In task scheduling, Maenhout & Vanhoucke (2016) consider precedence relations between tasks only, while our approach also incorporates task start time windows. The authors present a branch-and-price solution approach to solve the problem. Our problem decomposition differs from that applied by Maenhout & Vanhoucke (2016); they decompose the problem into one MP and one SP generating shift schedules, while our decomposition considers one SP (type) for shift scheduling and another SP for task scheduling. A comparison between the two problem decompositions is provided in Subsection 3.5.4. The authors present and assess different branching strategies for both shift and task branching.

We draw two major conclusions from the literature review: First, we acknowledge that shift scheduling problems in the healthcare industry have repeatedly been solved relying on branch-and-price approaches. Second, we show that integrated approaches that extend shift scheduling with task scheduling have received some attention in recent years. Branch-and-price approaches appear to be a suitable technique to find solutions to this type of problem. To the best of our knowledge, this thesis is the first to propose an optimal solution procedure for the integrated shift and task scheduling problem based on two SPs. Moreover, the integrated approach relying

on a shift scheduling SP based on an implicit shift formulation solved with a NF ShP approach as presented in this chapter is novel.

4.3 Problem description

The problem comprises an integrated shift and task scheduling problem for an equally skilled workforce, allowing different levels of flexibility in the shift patterns. Varying degrees of flexibility come at different costs. In line with Chapter 3, hospital managers are confronted with this problem type when introducing logistics assistants in hospitals. While the case of logistics assistants was the starting point of our research, we believe that the presented problem setting is not limited to this employee group, but has numerous application areas. Potential fields of application are characterized by a homogenous, mostly low-qualified workforce and a set of tasks with well-projectable occurrence and a known duration. Prerequisites include that all employees are able to fulfill all tasks and that there is neither a ramp-up or handover time for performing the tasks nor a required break between the distinct tasks. If ramp-ups or handover times are to be incorporated in the model, the tasks need to be provided with a buffer covering the additional time requirements. Apart from hospitals, promising application areas include production support functions and service operations management (see Section 3.3).

The strategic workforce sizing question is addressed in Chapter 3. The problem addressed in this chapter is of a tactical nature. The goal is to define shift and task schedules for a given number of employees. In this tactical optimization problem, we consider different worker categories who are employed under different shift parameters and work at different costs. In order to find optimal solutions, we perform two optimization activities in an integrated manner: The basic idea of our approach is to ensure that resource supply covers resource demand in every time period. The basic modeling assumptions for resource supply and resource demand are the same as in Chapter 3, but shift scheduling, i.e., resource supply, allows for a higher granularity by reflecting different levels of shift flexibility. In the following, we elaborate in detail on resource supply and resource demand.

4.3.1 Resource supply

Resource supply comprises the sum of all contracted employees working in a specific time period. Our model relies on an implicit shift formulation. This means that we do not provide predefined, explicit shift types. Instead, sequences of on- and off-periods are modeled implicitly reflecting collaborative agreements, legislative requirements, and labor contracts. We apply the following nomenclature for the remainder of this chapter: A shift schedule is a sequence of working periods/on-periods 1, and rest periods/off-periods 0. It is active in a time period when it

is an on-period. For a detailed literature review on implicit shift scheduling, we refer to Section 3.2. We assume that there are two categories of workers that differ in terms of how flexibly they can be employed. A higher degree of flexibility results in higher hourly cost of employing workers. While neither of the two worker categories is completely inflexible, we refer to *inflexible* or *fixed workers* for the less flexible category and *flexible workers* for the more flexible category. Inflexible workers' shift patterns typically comprise a working day of a fixed length, e.g., eight hours per day. As the name suggests, the second worker category, flexible workers, is employed very flexibly. We assume that shift start times, durations, and consequently their finish times are arbitrary as long as they comply with the constraints stated above. The total number of on-periods is limited by the allowed number of working periods in the planning horizon. The number of consecutive on-periods has to be within the minimum and maximum shift length. Between two on-periods, we request that a minimum number of off-periods is adhered.

4.3.2   Resource demand

The resource demand is a result of the tasks being performed in a certain time period and their respective resource needs. We assume that the duration of each task is deterministic and known, and that the number of resources required to perform a task is known and constant during the execution of the respective task. Task scheduling is subject to task start time windows and precedence relations between tasks.

4.4   Solution approach

In Subsection 3.3.2, we present the compact formulation of our basic model. While it does not reflect different worker categories, a model extension is straightforward. We show that real-world problem instances in the compact formulation are not solvable in an acceptable time by a standard solver. In this section, we present a branch-and-price approach to solve the tactical integrated shift and task scheduling problem. Branch-and-price algorithms are occasionally referred to as IP column generation. In what follows, we present a Dantzig-Wolfe decomposition of our original problem. It decomposes by shifts and tasks and consequently relies on one MP and two pricing SP types for shift and task scheduling.

The branch-and-price algorithm comprises two major steps. In the first step, column generation is applied to find valid lower bounds (LBs) for each node of the search tree. To do so, we consider the restricted linear MP, which only contains a subset of columns of all possible shift and task schedules and is referred to as the restricted MP or MP-LP. Columns are generated by solving the pricing SPs. Promising columns are iteratively added to MP-LP, which is

subsequently re-optimized. As we are in a minimization context, we search SP solutions with negative reduced costs. Adding columns to MP-LP improves its objective value. The task scheduling SP (T-SP) generates task schedules, while shift schedules are generated in the shift scheduling SP (S-SP). Please note that there are two S-SP types, reflecting two worker categories: flexible and inflexible workers. The column generation procedure is repeated until the SPs no longer price out, i.e., there are no SP solutions with negative reduced costs. This means that MP-LP is solved to optimality and that we found a LB of our initial problem. In the second step of the branch-and-price algorithm, branching is performed in order to drive non-integer solutions of MP-LP to integrality. For more details on the branch-and-price approach, we refer to Vanderbeck & Wolsey (1996), Barnhart et al. (1998), and Vanderbeck (2000).

The next section is structured as follows: We present the MP in Subsection 4.4.1, followed by S-SP including the newly developed NF formulation in Subsection 4.4.2. T-SP is displayed in Subsection 4.4.3, and heuristics to find upper bounds (UBs) in Subsection 4.4.4. We conclude by elaborating on the branching logic (Subsection 4.4.5) and the overall algorithm (Subsection 4.4.6).

4.4.1 Master problem (MP)

The MP fulfills two functions: First, it coordinates the pricing SPs and second, it ensures that resource supply fully covers resource demand in every time period. The basic idea is comparable to the MP of Subsection 3.4.1; however, the objective function is different and the model presented below allows shift scheduling for different worker categories. In the model, we consider a one-week planning horizon in one-hour increments. This approach is in line with the shift scheduling literature and supports the repetitive nature of logistics tasks in weeks. The set $\mathcal{T}$ represents the time periods $t \in \mathcal{T}$, while the set of workers to be scheduled is indicated by $\mathcal{W}$, with individual workers $w \in \mathcal{W}$. We distinguish between subsets $\mathcal{W}^{fix}$ for inflexible workers and $\mathcal{W}^{flex}$ for flexible workers.

A shift schedule for worker w is depicted by $j \in \mathcal{J}^w$. The resource supply in time period t comprises the sum of all active shift schedules in that period. The binary parameter Y_{jwt} indicates whether shift schedule j for worker w is active in time period t. It describes the binary parameter $\vec{Y}$ of the respective shift schedule. The cost per period for employing worker w is depicted by C_{jw}. The number of workers w working in shift pattern j is depicted by decision variables λ_{jw}. The sum of all active periods in shift pattern j is represented by H_j.

The demand structure is slightly different, as selecting one task schedule defines the entire demand for every time period t. The task schedules are represented by the set $\mathcal{P}$ with individual schedules $p \in \mathcal{P}$. The decision variables ρ_p are the task schedule selection variables. They are 1 if schedule p is selected and 0 otherwise. The resulting resource demand of that respective task schedule p in time period t is displayed in D_{pt}, which represents the integer demand vector $\vec{D}$ of the respective task schedule.

As a modeling device, we allow demand undercoverage to already achieve feasibility at the beginning of column generation. The amount of demand undercoverage in period t is depicted by decision variables z_t. The cost of demand undercoverage in one time period C^{out} is set so high that no undercoverage occurs in an optimal solution, i.e., it holds that z_t equals zero in all time periods.

Similar to Chapter 3, we apply the following nomenclature: Capital calligraphic letters represent sets and capital letters stand for their cardinality. Lowercase letters represent indices, while parameters are indicated by capital letters with indices. Decision variables are lowercase letters with indices.

Sets with indices

$\mathcal{W}$ set of workers with index $w \in \mathcal{W} = \{1, \dots, W\}$, with flexible workers $w \in \mathcal{W}^{flex}$ and inflexible workers $w \in \mathcal{W}^{fix}$ and $\mathcal{W}^{fix} \cup \mathcal{W}^{flex} = \mathcal{W}$, $\mathcal{W}^{fix} \cap \mathcal{W}^{flex} = \emptyset$

$\mathcal{J}^w$ set of shift schedules of worker w with index $j \in \mathcal{J}^w = \{1, \dots, J^w\}$

$\mathcal{P}$ set of task schedules with index $p \in \mathcal{P} = \{1, \dots, P\}$

$\mathcal{T}$ set of time periods with index $t \in \mathcal{T} = \{1, \dots, T\}$

Parameters

Y_{jwt} 1 if shift schedule j of worker w is active in period t, 0 otherwise

C_{jw} cost of employing worker w per period

C^{out} cost of assigning an outside worker per period

H_j number of working periods in shift schedule j

D_{pt} work demand of task schedule p in period t

Decision variables

λ_{jw} number of workers w in shift schedule j

ρ_p 1 if task schedule p is selected, 0 otherwise

z_t amount of demand undercoverage in period t

Master problem

The integer MP is presented below.

$$\text{Min } z^{MP} = \sum_{w \in \mathcal{W}} \sum_{j \in \mathcal{J}^w} C_{jw} H_j \lambda_{jw} + \sum_{t \in \mathcal{T}} C^{out} z_t \tag{63}$$

subject to

$$\sum_{w \in \mathcal{W}} \sum_{j \in \mathcal{J}^w} Y_{jwt} \lambda_{jw} - \sum_{p \in \mathcal{P}} D_{pt} \rho_p + z_t \geq 0 \qquad \forall t \in \mathcal{T} \tag{64}$$

$$\sum_{p \in \mathcal{P}} \rho_p = 1 \tag{65}$$

$$\sum_{j \in \mathcal{J}^w} \lambda_{jw} = 1 \qquad \forall w \in \mathcal{W} \tag{66}$$

$$\lambda_{jw} \in \{0,1\} \qquad \forall w \in \mathcal{W}, j \in \mathcal{J}^w \tag{67}$$

$$\rho_p \in \{0,1\} \qquad \forall p \in \mathcal{P} \tag{68}$$

$$z_t \in \mathbb{N}_0 \qquad \forall t \in \mathcal{T} \tag{69}$$

The objective function (63) minimizes the total hourly cost of employing workers and the cost of demand undercoverage. In constraints (64), we make sure that resource demand, which derives from the selected task schedule, is sufficiently covered by resource supply as a result of the number of workers in active shifts and demand undercoverage. In constraint (65), we ensure that exactly one plan p is selected. Comparably, we ensure in constraints (66) that, for every worker w, one shift schedule j is selected. Constraints (67) and (68) define the binary variables, and constraints (69) define that demand undercoverage z_t are non-negative integer variables.

It might be necessary to limit the share of the workload covered by flexible workers. In order to incorporate that limitation into our model, one would simply add constraint (70), where parameter $S^{flex,max}$ represents the maximum share of the total workload of flexible workers in relation to the total workload of inflexible workers.

$$S^{flex,max} \sum_{w \in \mathcal{W}^{fix}} \sum_{j \in \mathcal{J}^w} H_j \lambda_{jw} - \sum_{w \in \mathcal{W}^{flex}} \sum_{j \in \mathcal{J}^w} H_j \lambda_{jw} \geq 0 \tag{70}$$

MP columns are added iteratively by solving the SPs. The search for new columns is guided by the dual version of MP-LP, which is depicted below. In line with Chapter 3, we apply the

following nomenclature: When integrality conditions of constraints (67) to (69) are adhered, we refer to the integer programming MP or MP-IP. As stated before, the restricted linear MP relying on only a subset of shift and task columns is referred to as MP-LP. In MP-LP, the integrality constraints of variables λ_{jw}, ρ_p, and z_t are dropped and the variables are treated as continuous variables. We introduce the variables $\delta_t \geq 0$ representing the duals of constraints (64), variable $\mu \in \mathbb{R}$ representing the dual of constraint (65), and variables $\pi_w \in \mathbb{R}$ representing the duals of constraints (66). The objective function and the constraints derive from duality.

Dual master problem

$$\text{Max } z^{DMP} = \mu + \sum_{w \in \mathcal{W}} \pi_w \tag{71}$$

subject to

$$C_{jw} H_j - \sum_{t \in \mathcal{T}} Y_{jwt} \delta_t - \pi_w \geq 0 \qquad \forall j \in \mathcal{J}^w, w \in \mathcal{W} \tag{72}$$

$$\sum_{t \in \mathcal{T}} D_{pt} \delta_t - \mu \geq 0 \qquad \forall p \in \mathcal{P} \tag{73}$$

$$\delta_t \leq C^{out} \qquad \forall t \in \mathcal{T} \tag{74}$$

$$\delta_t \geq 0 \qquad \forall t \in \mathcal{T} \tag{75}$$

$$\mu \in \mathbb{R} \tag{76}$$

$$\pi_w \in \mathbb{R} \qquad \forall w \in \mathcal{W} \tag{77}$$

4.4.2   Shift subproblem (S-SP)

After solving MP-LP to optimality relying on the available shift and task columns, we generate new shift columns with S-SP. As presented in Subsection 3.4.2, the shift scheduling pricing problem can be modeled implicitly relying on an IP formulation. In that case, the SP objective function, depicted by $\bar{C}_{jw}$, equals the negative reduced cost of a new shift schedule. It derives from the dual problem of MP-LP and is depicted in (78). The constraints of the S-SP IP formulation are shown in Subsection 3.4.2.

$$\text{Min } \bar{C}_{jw} = C_{jw} \sum_{t \in \mathcal{T}} y_t - \sum_{t \in \mathcal{T}} \delta_t y_t - \pi_w \tag{78}$$

In the following, we present a NF representation of the S-SP and a ShP-labeling algorithm to efficiently solve the problem to optimality. By way of illustration, we provide a graphical extract and a brief example of the network at the end of this subsection. The basic idea of the

new NF formulation was initially presented by Brunner (2015). Other publications that include ideas of NF formulations for shift scheduling are at hand (see, e.g., Jaumard et al. 1998, Caprara et al. 2003, Maenhout & Vanhoucke 2010, and Brunner et al. 2013). We display the shift scheduling parameters below, as they are required for the NF formulation. They are the same as in Chapter 3. The graph representation of the shifts is novel.

Shift parameters

$T^{shiftMin}$ minimum shift length in time periods

$T^{shiftMax}$ maximum shift length in time periods

T^{rest} minimum rest time periods between two consecutive shifts

$T^{workMax}$ maximum number of working periods per planning horizon, e.g., per week

The number of consecutive working periods is constrained by the minimum shift length $T^{shiftMin}$ and the maximum shift length $T^{shiftMax}$. The total number of working periods in the planning horizon must not exceed a maximum working time $T^{workMax}$. Moreover, a minimum rest time between two consecutive shifts T^{rest} has to be adhered.

Network definition. Consider a directed and loopless graph $G(V, E)$ with a set of vertices V and a set of edges E. Each vertex $v \in V$ represents a state that is defined by a tuple (t, t^{on}). The current time period is depicted by $t \in \mathcal{T} = \{1, \dots, T\}$, while $t^{on} \in \{0, T^{shiftMin}, \dots, T^{workMax}\}$ represents the number of working periods since the beginning of the planning horizon. An edge $e_{ij} \in E = \left((t_i, t_i^{on}), (t_j, t_j^{on})\right)$ represents a sequence of working periods starting in t_i, followed by a sequence of rest periods ending in $t_j - 1$. Consequently, a path through the network is a representation of a shift schedule, i.e., sequences of working periods and rest periods. The formal definition of all vertices $v \in V$ is given in (79) to (82).

$$V = V' \cup V'' \cup V''' \tag{79}$$

$$V' = \left\{ (t, t^{on}) \middle| \begin{array}{l} t \in \mathcal{T}, t^{on} = 0 \lor T^{shiftMin} \leq t^{on} \leq T^{workMax} \\ \land\, t^{on} \leq t - T^{rest} - 1 \end{array} \right\} \tag{80}$$

$$V'' = \left\{ (t, t^{on}) \middle| \begin{array}{l} t = T + 1, t^{on} = 0 \lor T^{shiftMin} \leq t^{on} \leq T^{workMax} \\ \land\, t^{on} \leq t - T^{rest} - 1 \end{array} \right\} \tag{81}$$

$$V''' = (T + 2, -1) \tag{82}$$

There are three types of vertices, represented by sets V', V'', and V'''. Vertex set V' represents all vertices within the planning horizon. A state constitutes the current time period $t_i \in \mathcal{T}$ as well as the number of all possible working periods t_i^{on} until time period t_i. As shown in (80), the number of working periods t_i^{on} can be 0 in case that there has not yet been a working

period. If there have already been working periods before time period t_i, t_i^{on} is greater than or equal to $T^{shiftMin}$ and can take a maximum value of $T^{workMax}$. Moreover, it is implicitly clear that t_i^{on} must be smaller than $t_i - T^{rest} - 1$, as otherwise the respective vertex could not have been reached by a valid sequence of working periods and rest periods. Vertices V'' as displayed in (81) represent vertices at the end of the planning horizon. As edges $e_{ij} \in E = \left((t_i, t_i^{on}), (t_j, t_j^{on})\right)$ mark a sequence of working periods and rest periods that end in $t_j - 1$, V'' is required in order to represent working and rest sequences that last until the end of the planning horizon. Vertex V''' is the final vertex. In order to indicate the final vertex, we artificially set t_j^{on} to -1. All valid paths end in the final vertex V'''. Note that an ID is assigned to all vertices. The IDs are non-negative integer numbers with zero as the ID of the start vertex.

We present the formal definition of all edges $e_{ij} \in E$ in (83) to (88).

$$E = E_1 \cup E_1^{rest} \cup E_2 \cup E_2^{rest} \cup E_3 \tag{83}$$

$$E_1 = \left\{ \left((t_i, t_i^{on}), \begin{pmatrix} t_i + n + T^{rest}, \\ t_i^{on} + n \end{pmatrix}\right) \middle| \begin{array}{l} (t_i, t_i^{on}) \in V', n \in \{T^{shiftMin}, \dots, T^{shiftMax}\}, \\ t_i + n + T^{rest} \le T, t_i^{on} + n \le T^{workMax} \end{array} \right\} \tag{84}$$

$$E_1^{rest} = \left\{ \left((t_i, t_i^{on}), (t_i + 1, t_i^{on})\right) \middle| (t_i, t_i^{on}) \in V', t_i + 1 \le T \right\} \tag{85}$$

$$E_2 = \left\{ \left((t_i, t_i^{on}), (t_i + n, t_i^{on} + n)\right) \middle| \begin{array}{l} (t_i, t_i^{on}) \in V', n \in \{T^{shiftMin}, \dots, T^{shiftMax}\}, \\ t_i + n = T + 1, t_i^{on} + n \le T^{workMax} \end{array} \right\} \tag{86}$$

$$E_2^{rest} = \left\{ \left((t_i, t_i^{on}), (t_i + 1, t_i^{on})\right) \middle| (t_i, t_i^{on}) \in V', t_i = T \right\} \tag{87}$$

$$E_3 = \left\{ \left((t_i, t_i^{on}), (t_i + 1, -1)\right) \middle| (t_i = T + 1, t_i^{on}) \in V'' \right\} \tag{88}$$

We distinguish between five different edge types, which are represented by edge sets $E_1, E_1^{rest}, E_2, E_2^{rest}$, and E_3. The union of the five edge sets forms edge set E. In each vertex v, a sequence of working periods or a rest period can be started. The start of a rest period is represented by E_1^{rest} or E_2^{rest}. If one of these two edges is selected, the sequence of rest periods is extended by one time period. E_1^{rest} as presented in (85) represents rest edges that originate from vertices $v \in V'$. At the end of the planning horizon, rest periods are indicated by E_2^{rest}, as shown in (87). The start of a working period during the planning horizon is represented by E_1 (see (84)). Then, we make sure that we start a shift with a minimum of $T^{shiftMin}$ working periods and T^{rest} rest periods. If we are at the end of the planning horizon, we drop the prerequisite of having rest periods at the end of an on/off-pattern. Consequently, we only demand that the minimum number of working periods $T^{shiftMin}$ is adhered. Practically speaking, this means that shift schedules are not necessarily cyclic and a shift schedule may end

with a sequence of working periods. This case is represented by edge set E_2, as shown in (86). The edge set E_3 (see (88)) represents edges that connect all paths to the final vertex, so that all paths start in the source vertex $(0,0)$ and end in sink vertex $(T+2, -1)$.

Please note that the number of edges leaving each vertex $v \in V'$ comprises the rest edge plus the edges representing all possible shift lengths. In order to make that structure clear, we provide an extract of the layout of our graph in Figure 4.2. Let us assume $v \in V'$. Edge (A) represents the rest edge. This means that no working period is started in time period $t = t_i$, but the sequence of rest periods is extended by one time period. Consequently, in the succeeding vertex, the time period t_i is augmented by 1, while the number of total working periods within the planning horizon t_i^{on} remains unchanged. Edges (B), (C), and (D) represent the start of a working period in time period t_i^{on}. Edge (B) represents the start of a working period with minimum length, i.e., $T^{shiftMin}$. Consequently, the number of working periods in the planning horizon of the successor t_j^{on} is equal to $t_i^{on} + T^{shiftMin}$. The respective time period t_j is the time period of the predecessor t_i plus shift length $T^{shiftMin}$ plus the minimum number of rest periods T^{rest}. Edge (C) is comparable to edge (B) with the only difference that the sequence of working periods is one time period longer. Edge (D), in contrast, represents the start of the longest possible shift length in vertex (t_i, t_i^{on}), $T^{shiftMax}$.
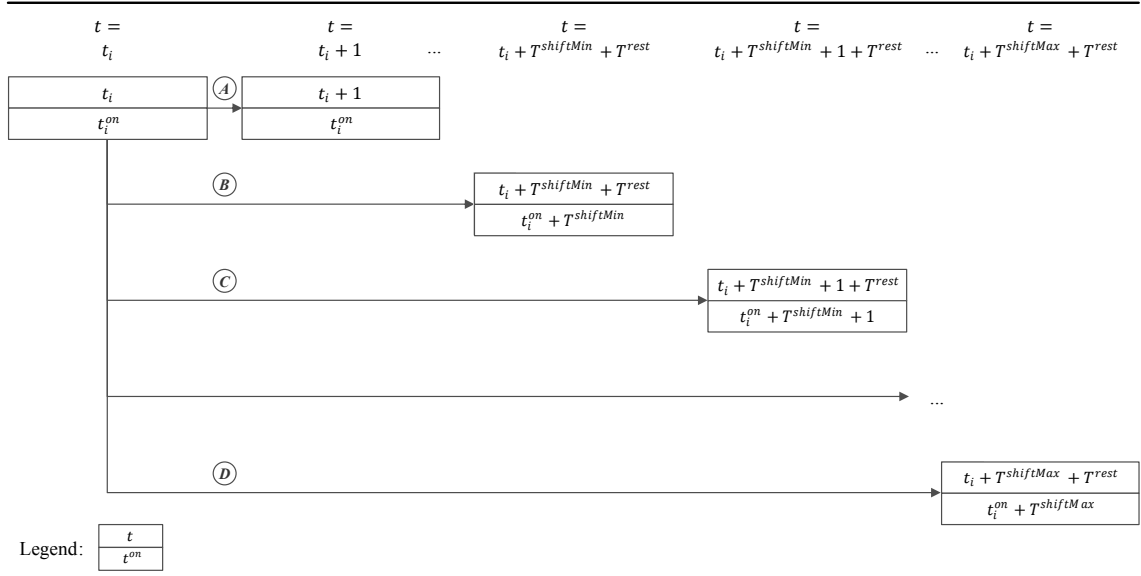

Figure 4.2: Extract of the shift graph

An illustrative example of the graph for a small problem instance with a time horizon of $t \in \mathcal{T} = \{1, \dots, 6\}$ is shown in Figure 4.3. The horizontal axis represents the time periods t, while the vertical axis represents the number of working periods in the planning horizon. The

start of a rest period is represented by horizontal edges. In our illustrative example, the path through the graph $\big((1,0), (2,0), (6,3), (7,3), (8,-1)\big)$ equals a shift sequence of $(0,1,1,1,0,0)$.
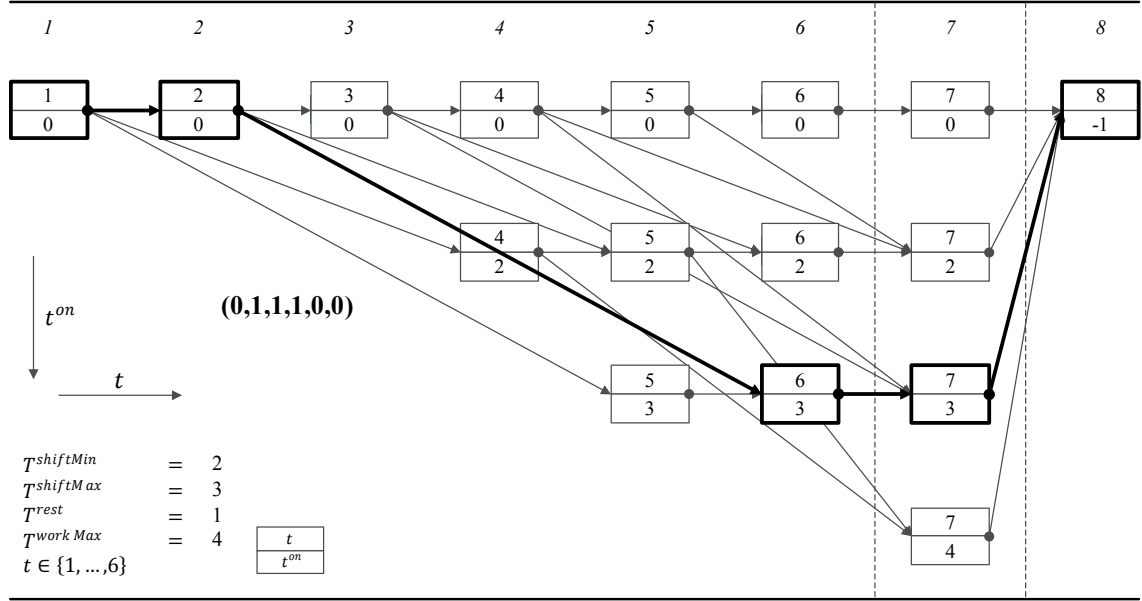

Figure 4.3: Illustrative example of a shift graph

Edge weights. We present the edge weights in the following. Horizontal edges, which represent time periods where no working periods are started, have an edge weight of 0. This holds for E_1^{rest} and E_2^{rest}. Edges which stand for the start of a working period, i.e., E_1 and E_2, are weighted corresponding to the negative sum of dual values δ_t associated with the number of working periods of the shift that begins. The weights of edges E_3, namely the edges that flow into the final vertex, correspond to the cost per hour C_{jw} multiplied with the total number of working hours during the entire planning horizon $H_j = \sum_{t\in\mathcal{T}} y_t$. In our model, the cost per hour is constant over the planning horizon, so we can add these costs at the end of each path. This is beneficial from an implementation perspective, as we only need to set the respective edge weights once when initializing the graph. In later iterations, no update is necessary. If the cost per working period were time-dependent, one would need to add the respective costs to the edge weights at the start of working periods.

Labeling algorithm. The ShP through the network represents the optimal solution of S-SP. In the following, we briefly describe a simple labeling algorithm that we use to solve S-SP. A pseudo-code of the algorithm is presented in Figure 4.4. The algorithm comprises two major steps, i.e., a forward labeling step and a backward readout step. The basic idea is to assign a label to each vertex $v \in V$ that contains the cost of the ShP from the start vertex to the respective vertex and the ID of the previous vertex. The ShP is read in the backward readout step.

In the labeling step, the algorithm iterates through the vertices of the network, starting with the start vertex and moving forward for each time period $t \in \mathcal{T}$ from the smallest to the largest ID. When going through the network, we label each succeeding vertex with the cost of the ShP to the start vertex $c_{j,new}^{path} = c_{i,current}^{path} + c_{ij}$, with c_{ij} being the weight of edge e_{ij}. Moreover, we save the ID of the previous vertex. If the respective succeeding vertex has already been labeled, we compare the current lowest cost $c_{j,current}^{path}$ with the new lowest cost $c_{j,new}^{path}$ and update the label if a lower-cost path has been found.

Once the labeling step has iterated through all edges, the backward readout step follows. It starts with the final vertex and reads the cost of the path. The path cost assigned to the final vertex equals the objective value of S-SP without considering the dual values π_w. The objective is read and saved. In order to find the ShP, we go backwards through the path, reading out the IDs of the best previous vertex. The backward iteration stops when we arrive at the start vertex. All vertices that are visited and that consequently lie between the final vertex and the start vertex comprise the shortest path. In the end, we translate the path into a feasible shift schedule.

Labeling algorithm

STEP 1: Labeling

1:	reset all vertex labels
2:	for all t
3:	for all edges *(i, j)* with $t_i = t$
4:	calculate $c_{j,new}^{path} = c_{i,current}^{path} + c_{ij}$
5:	if $\left(c_{j,new}^{path} < c_{j,current}^{path} \right)$
6:	update label $\left(ID_i, c_{j,new}^{path} \right)$
7:	end if
8:	end for
9:	end for

STEP 2: Readout

10:	select final vertex
11:	add final vertex to path
12:	read label: objective
13:	read label: previous vertex
14:	while (previous vertex ≠ start vertex)
15:	add previous vertex to path
16:	previous vertex = read label predecessor
17:	end while
18:	add start vertex to path

Figure 4.4: Labeling algorithm

Once S-SP is solved and it holds that $\bar{C}_{jw} < 0$, i.e., the new column is dual-infeasible, we add the resulting shift column to MP-LP. As long as there are feasible solutions of S-SP with negative reduced costs, adding columns improves the MP-LP objective value z^{MP}. To solve

MP-LP to optimality, we do not need to solve S-SP optimally. Nevertheless, the presented approach solves S-SP to optimality. The shift column we add to MP-LP takes the following shape:

$$
\begin{bmatrix}
C_{jw} \sum_{t \in \mathcal{T}} y_t \\
\vec{Y} \\
0 \\
\vec{I}
\end{bmatrix}
$$

The cost of the new shift column is depicted by $C_{jw} \sum_{t \in \mathcal{T}} y_t$. The parameter $\vec{Y}$ of length T represents the new shift schedule as sequences of on-periods 1 and off-periods 0 for each time period $t \in \mathcal{T}$. $\vec{I}$ is a binary parameter vector of length W that indicates for which worker $w \in \mathcal{W}$ the shift schedule has been generated.

4.4.3   Task subproblem (T-SP)

New task schedules for MP-LP are generated by solving T-SP. Again, we search for dual infeasibility cuts. Solutions of T-SP with negative reduced costs comprise promising columns to add to MP-LP. The T-SP presented below resembles the T-SP in Subsection 3.4.3. For reasons of readability, we state the T-SP again. We start with the notation.

Sets with indices

$\mathcal{A}$ set of tasks with index $a \in \mathcal{A} = \{1, \dots, A\}$

$\mathcal{A}_a^{pred}$ set of direct predecessors of task a with index $a \in \mathcal{A}_a^{pred} = \{1, \dots, A_a^{pred}\} \subseteq \mathcal{A}$

$\mathcal{T}_a^{start}$ set of possible start periods of task a with index $t \in \mathcal{T}_a^{start}$

             $= \{T_a^{earliest}, \dots, T_a^{latest}\} \subseteq \mathcal{T}$

Parameters

R_a resource demand of task a for each time period (number of workers)

L_a length of task a in time periods

$T_a^{earliest}$ earliest start time of task a

T_a^{latest} latest start time of task a

Decision variables

x_{at}^{start} 1 if task a starts in period t, 0 otherwise

d_t work demand in period t

We rely on the following sets: The set of all tasks a is represented by $\mathcal{A}$, while the subsets $\mathcal{A}_a^{pred}$ represent the sets of all direct predecessors of task a. $\mathcal{T}_a^{start}$ represents the potential start periods of tasks a. Each task is subject to four parameters. R_a is the number of workers required to perform task a. It is constant during the length of the task L_a. The task start time window is depicted by the earliest start time $T_a^{earliest}$ and the latest start time T_a^{latest}. The binary decision variables x_{at}^{start} indicate the start of task a. They are 1 if task a starts in period t and 0 otherwise. The non-negative integer decision variables d_t represent the demand in time period $t \in \mathcal{T}$.

The objective function of T-SP derives from constraint (73) of the dual problem of MP-LP. The T-SP is displayed below.

$$\text{Min } \bar{C}_p = \sum_{t \in \mathcal{T}} \delta_t d_t - \mu \tag{89}$$

subject to

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} x_{at}^{start} = 1 \qquad\qquad \forall a \in \mathcal{A} \tag{90}$$

$$\sum_{t=T_a^{earliest}}^{T_a^{latest}} t x_{at}^{start} - \sum_{t=T_b^{earliest}}^{T_b^{latest}} t x_{bt}^{start} \geq L_b \qquad\qquad \forall a \in \mathcal{A}, b \in \mathcal{A}_a^{pred} \neq \emptyset \tag{91}$$

$$\sum_{a \in \mathcal{A}} R_a \left(\sum_{\tau=\max\{T_a^{earliest};t-L_a+1\}}^{\min\{t;T_a^{latest}\}} x_{a\tau}^{start} \right) = d_t \qquad\qquad \forall t \in \mathcal{T} \tag{92}$$

$$x_{at}^{start} \in \{0,1\} \qquad\qquad \forall a \in \mathcal{A}, t \in \mathcal{T}_a^{start} \tag{93}$$

$$d_t \in \mathbb{N}_0 \qquad\qquad \forall t \in \mathcal{T} \tag{94}$$

The objective (89) is to minimize the cost of a new task schedule. In constraints (90), we make sure that each task is performed. Constraints (91) make sure that precedence relations between tasks are respected. They state that a succeeding task may only start once all its predecessors have been finished. In constraints (92), we add up the respective resource demand d_t for each time period which is required in the objective function. Constraints (93) and (94) set the variables x_{at}^{start} binary and d_t non-negative integer, respectively.

A slightly modified formulation was assessed in a pretesting phase. In the alternative formulation, we dropped constraints (92) and variables d_t. Furthermore, we replaced the objective function (89) with the following objective function:

$$\text{Min } \bar{C}_p = \sum_{a \in \mathcal{A}} \sum_{t \in \mathcal{T}_a^{start}} R_a c_{at} x_{at}^{start} - \mu \tag{95}$$

We introduce the matrix c_{at} for the alternative formulation. It contains the sum of the dual values associated with starting task a in time period t. The matrix is updated every time T-SP is called. Unfortunately, preliminary testing showed that by modifying T-SP accordingly, the solution procedure could not be accelerated.

Once the T-SP solution is negative, i.e., $\bar{C}_p < 0$, we add the resulting column to MP-LP. The column takes the following shape:

$$\begin{bmatrix} 0 \\ \vec{D} \\ 1 \\ \vec{0} \end{bmatrix}$$

The new task column contains the demand profile that results from the T-SP solution. It is depicted by vector $\vec{D}$, with elements $D_t \in \mathbb{N}_0$ representing the demand. $\vec{0}$ is a zero vector with elements 0 and a length of W. Please note that, in line with S-SP, T-SP does not need to be optimally solved to improve the MP-LP objective value z^{MP}. Once a T-SP solution with negative reduced costs is found, the resulting column can be added to MP-LP to improve the objective value z^{MP}. In this case, we found a column with dual infeasibility cuts. Preliminary testing suggests that selecting the first T-SP solution with negative reduced costs is superior to solving T-SP to optimality, as we achieve better overall runtimes to solve MP-LP to optimality.

4.4.4 Upper bound heuristics

Generating high-quality integer solutions during the enumeration of the search tree can significantly speed up the solution procedure. Nodes with LBs higher than or equal to the best/lowest UB do not need to be further investigated and can be fathomed. This means that the respective nodes are disregarded, as they will not yield better UBs than the current best known UB. In order to find high-quality UBs, we rely on two heuristics and an improvement heuristic, which we present in the following. The first global UB is obtained in the root node and subsequently updated whenever a better UB is found.

Finding upper bounds. In each iteration, we employ a simple rounding heuristic that works as follows: When MP-LP is solved to optimality, we select the task schedule with the highest ρ_p in

the MP-LP solution and set its value to 1 and all remaining ρ_p to 0. We then select the highest λ_{jw} for each worker w, set it to 1 and all other λ_{jw} of the respective worker to 0. If the selected supply and demand profiles result in demand undercoverage, we add the cost of the respective undercoverage. Obviously, the rounding heuristic does not necessarily provide high-quality solutions, as undercoverage occurs frequently. Nevertheless, the computational complexity of the heuristic is very low. In order to find high-quality UBs, we solve MP-IP to obtain integer solutions every 50 nodes. As the number of task columns in MP-IP has a negative impact on the runtime of MP-IP, we limit their number. In line with the approach presented in Subsection 3.4.5, we select task columns with flat demand patterns. We restrict the task column number to 5, which proved to be efficient in preliminary testing. Please note that in Subsection 3.4.5 we limit their number to the minimum of 15 and 10% of all generated task columns. There, we solve MP-IP only once, so longer runtimes can be accepted. In the presented branch-and-price approach, however, we solve MP-IP every 50 nodes, so runtimes must be kept low. We consider all generated shift columns when solving MP-IP.

Improving upper bounds. Once either of the two presented heuristics finds an improved UB, we hand the new solution to a simple improvement heuristic. The basic idea is to shorten working periods when they are not needed to cover the resource demand. The heuristic works as follows: It leaves the selected task schedule unchanged. Based on the task schedule and the selected shifts, the heuristic identifies all time periods in the planning horizon where resource supply exceeds resource demand. Subsequently, the heuristic aims to reduce this resource oversupply. It does so by checking each shift schedule that is active in the respective oversupply period if it marks the start or end of a working period. If this is the case and the length of the working periods exceeds the minimum number of consecutive on-periods $T^{shiftMin}$, we set the shift in the respective time period to 0. The procedure terminates when neither of the shift schedules can be further shortened.

4.4.5 Branching

Once column generation terminates and a LB for the respective node of the search tree is found, branching is necessary if the MP-LP solution is fractional. We employ the same general branching logic to drive the task scheduling variables ρ_p and the shift scheduling variables λ_{jw} to integrality. We start branching by selecting the most fractional MP-LP variable, regardless of whether it is λ_{jw} or ρ_p. Depending on which variable is most fractional, we employ task branching or shift branching. The branching scheme we apply is a subvariable or constraint branching scheme that was first presented by Ryan & Foster (1981). For a comparable problem

to ours, the branching scheme is successfully applied by Beliën & Demeulemeester (2008) for task branching.

Task branching. We branch on the demand per period. Let us assume that the most fractional MP-LP variable is ρ_{p^*}. We randomly select another $\rho_p > 0$, say $\rho_{p^{**}}$. We compare the respective demand vectors of the two task columns, D_{p^*t} and $D_{p^{**}t}$. Starting with $t = 0$, we select the first period t' as our branching periods where it holds that $D_{p^*t'} \neq D_{p^{**}t'}$. The case that both demand vectors are equal does not occur, as duplicate columns are not added to MP-LP. Branching takes place by generating two child nodes in the search tree. In the left node, we impose that, if $D_{p^*t'} < D_{p^{**}t'}$, $D_{t'} \leq D_{p^*t'}$, while in the right node, we force $D_{t'} \geq D_{p^*t'} + 1$. In case $D_{p^*t'} > D_{p^{**}t'}$, the left node imposes $D_{t'} \leq D_{p^{**}t'}$, while the right node requests $D_{t'} \geq D_{p^{**}t'} + 1$. Consequently, we impose a maximum demand value in the left branch and a minimum demand value in the right branch for the selected branching period and branch on decision variables d_t in T-SP. In the subsequent node, the branching constraint is incorporated into T-SP. Moreover, we exclude task columns from MP-LP that do not fulfill the new branching constraints. To do so, we check in each node at the beginning of column generation which of the generated task columns fulfills the branching constraints. Moreover, we remove task columns that were not in the basis of the MP-LP solution in the last 50 nodes. Columns that are excluded from MP-LP are saved and potentially re-included in subsequent nodes.

Shift branching. In shift branching, we apply the same general approach, i.e., we branch on active or inactive shifts per period and worker. If the variable with the most fractional value in MP-LP is $\lambda_{j^*w'}$, we randomly select another $\lambda_{jw'} > 0$ of the same worker w', say $\lambda_{j^{**}w'}$. In line with task branching, our branching period is the first time period t' where it holds that $Y_{j^*w't'} \neq Y_{j^{**}w't'}$. In the left node, we impose that $Y_{w't'} \leq 0$, and in the right node, we force that $Y_{w't'} \geq 1$. Consequently, we branch on y_t in S-SP for one specific worker; we impose in the branching period that the shift schedule is inactive in the left branch and force the shift schedule to be active in the right branch. In order to reflect the branching constraints in our NF representation, we remove edges from the graph that represent forbidden shift patterns in the branching period. In line with task branching, we exclude shift columns from MP-LP that do not fulfill the branching constraints when resuming column generation in the subsequent node. Also, we remove shift columns that were not in the basis of MP-LP in the last 100 nodes. Shift columns that do not match the branching constraints are stored and potentially reused in subsequent nodes.

4.4.6 Algorithm

In the following, we present the branch-and-price algorithm. An illustrative flow chart of the solution approach is provided in Figure 4.5. The two major steps – first, defining a LB for each node of the search tree by solving MP-LP with means of column generation, and second, finding high-quality UBs – are indicated by areas highlighted in gray.
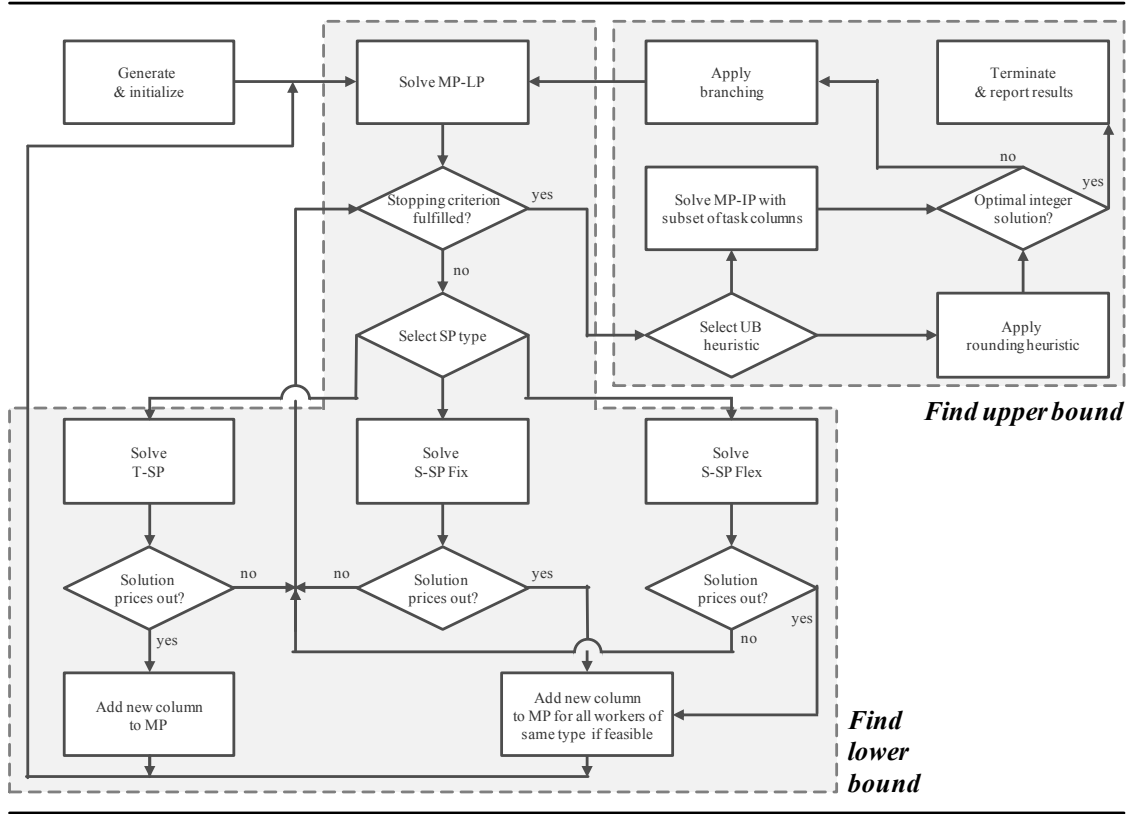

Figure 4.5: Illustrative flow chart of the branch-and-price approach

MP-LP initialization. We initialize MP-LP by adding artificial supercolumns for tasks and shifts. Adding an artificial task schedule is necessary to trigger the creation of new shifts. The task schedule supercolumn has a demand vector $\vec{D}$ with identical elements D_t at each time period of the planning horizon $t \in \mathcal{T}$ that are larger than the maximum peaks of the real demand profiles. Shift supercolumns need to be added to ensure MP-LP feasibility at the start of column generation. Remember that one shift schedule per worker has to be chosen in MP-LP. Consequently, we add one shift supercolumn per worker at high costs. Parameters are set so that in an optimal solution, MP-LP does not rely on the supercolumns. After adding the supercolumns, we start the column generation procedure to find the optimum of MP-LP.

SP selection. When MP-LP is solved to optimality with the existing columns, we apply the following logic for selecting SPs and generating new columns for MP-LP: We iteratively call T-

SP and S-SP. When calling S-SP, we iterate between flexible and inflexible workers. If we have more workers of the same category, we iteratively call the S-SPs of all workers of the respective category. Whenever a SP is called and does not find a solution with negative reduced costs, we go directly to the next SP without re-calling MP-LP. Once a new column is added to MP-LP, we re-optimize MP-LP with the new set of columns.

Column generation stopping criteria. Column generation for solving MP-LP to optimality terminates when at least one of two stopping criteria is fulfilled. The first stopping criterion comprises the column generation standard stopping criterion, in which column generation terminates when neither of the SPs finds new solutions with negative reduced costs. In that case, MP-LP is solved to optimality and adding additional columns would not improve the MP-LP objective value. We return the MP-LP objective value rounded to the next integer number $\lceil z_i \rceil$ as the node's LB. Please note that this stopping criterion is only valid when neither of the SPs based on the identical dual values find new solutions with negative reduced costs. Consequently, T-SP and all S-SPs need to be called and checked. This stopping criterion leads to a comparably high number of required column generation iterations before proving MP-LP optimality. In the root node, however, it is sufficient to check the solution of one S-SP type (flexible or inflexible) due to the symmetry of the S-SPs before branching. The second stopping criterion considers the current MP-LP objective value z_i. When z_i falls below the integer LB of its parent node, we stop the column generation procedure and return the integer LB of the parent node, which is equal to $\lceil z_i \rceil$. When we add branching constraints to the child node, the child node's solution space is more restricted than that of its parent node. Consequently, an integer LB of the child node must be equal to or larger than the parent node's LB. Due to rounding of z_i, however, it is possible that the current node's MP-LP objective value z_i falls below the parent node's (rounded) LB. In that case, the objective value of MP-LP z_i cannot be reduced as much, as this results in a better LB. We therefore terminate column generation once the current node's MP-LP objective value falls below its parent node's LB.

Shift column management. It is possible to have multiple workers per worker category. Each worker is assigned to one S-SP. However, for several workers of the same category, e.g., flexible workers, the shift parameters are identical. This results in a high degree of symmetry among the S-SPs of the same worker category. The symmetry is reduced through branching, as different branching rules are applied to the individual S-SPs. We apply the following logic to reduce the required S-SP iterations: Whenever we find a new shift column, e.g., for a flexible worker, we check whether the generated shift schedule is a feasible solution for all other workers of the same category. In this case, we add the generated shift column to MP-LP for all flexible workers for which the generated column is a feasible solution.

4.5 Numerical study

In this section, we present and discuss numerical results. It is divided into two parts: First, we compare the runtimes of the presented S-SP formulations, namely the IP and the NF formulation. Second, we show the behavior of our branch-and-price approach. All calculations are performed on an i5-4300 bi-core CPU with 1.90 GHz, 64-bit, and 8 GB RAM, running on a Windows 7 operating system. We coded the algorithms in Java and used CPLEX Version 12.6.1 to solve the mixed integer and linear programs applying the CPLEX standard settings.

4.5.1 Shift subproblems

In the following study, we compare the different SP formulations, namely the IP formulation solved with CPLEX and the NF formulation solved with the labeling algorithm presented in Subsection 4.4.2. As mentioned above, S-SP does not need to be optimally solved in order to find a LB at each node of the search tree, i.e., to solve MP-LP to optimality. Whenever a feasible solution is found, the resulting column can be added to MP-LP to improve its objective value. Consequently, we compare the results of solving S-SP IP to optimality with solving S-SP IP without requiring the solution to be optimal. Note that S-SP NF always provides optimal solutions. In the study, we do not require T-SP to be solved to optimality but accept the first feasible solution. We terminate the column generation procedure once MP-LP in the root node is solved to optimality, i.e., the root node LB has been found.

Generating test instances. We generate 18 test instances building on the instances presented in Subsection 3.5.2. Start time windows and task lengths are modified to account for the different planning granularity. In Chapter 3, we schedule either only flexible (FL) or only inflexible/fixed (FX) workers, which is reasonable for strategic planning. When switching from the strategic planning level of Chapter 3 to the tactical planning level of this chapter, we incorporate different shift flexibilities into each test instance. We distinguish between flexible and fixed workers whose number is defined prior to solving the optimization problem. We determine the number of inflexible workers W^{fix} by allocating 50% of the required man-hour demand over the entire planning horizon to this worker category and divide the resulting man-hours by the maximum number of working periods in the planning horizon for fixed workers. The number of flexible workers W^{flex} is set so that, when we apply our UB heuristics at the root node, no demand undercoverage occurs. The shift parameters of inflexible and flexible workers are as follows: Inflexible shifts have a fixed shift length of eight hours, i.e., $T^{shiftMin} = T^{shiftMax} = 8$. The maximum number of working periods in the planning horizon $T^{workMax}$ is 40. The number of working periods of flexible shifts is between four and six hours, i.e., $T^{shiftMin} = 4$

and $T^{shiftMax} = 6$. Flexible workers can work a maximum of 20 hours in the planning horizon. For both worker categories, we set the minimum required number of rest periods between two consecutive shifts T^{rest} to 12. The hourly costs per working period C_{jw} are two for inflexible and three for flexible workers.

Table 4.1: Results of the S-SP study

Instance	Parameters		S-SP IP (optimal)				S-SP IP (feasible)				S-SP NF			
	W_{flex} [#w]	W_{fix} [#w]	LB [mu]	#itMP [#]	$t_{s\text{-}sp}$ [s.]	t_{total} [s.]	LB [mu]	#itMP [#]	$t_{s\text{-}sp}$ [s.]	t_{total} [s.]	LB [mu]	#itMP [#]	$t_{s\text{-}sp}$ [s.]	t_{total} [s.]
SMA_S1_LW	29	8	1,663	218	14.363	36.524	1,663	265	3.394	16.231	1,663	167	0.201	9.984
SMA_S1_SW	29	8	1,663	384	21.318	40.925	1,663	462	7.117	31.827	1,663	399	0.386	23.569
SMA_S2_LW	33	8	1,672	211	9.410	25.234	1,672	227	2.850	19.320	1,672	155	0.153	12.923
SMA_S2_SW	33	8	1,672	211	8.739	22.905	1,672	227	2.829	19.097	1,672	155	0.192	13.498
SMA_S3_LW	36	8	1,777	160	8.161	19.304	1,777	258	3.102	20.515	1,777	182	0.193	14.273
SMA_S3_SW	36	8	1,904	360	21.322	47.308	1,904	562	10.105	58.006	1,904	381	0.387	32.289
SMA_SUBTOTAL	-	-	-	1,544	83.313	192.200	-	2,001	29.397	164.996	-	1,439	1.512	106.536
MED_S1_LW	47	14	2,698	225	9.814	28.589	2,698	237	2.984	21.914	2,698	158	0.222	14.791
MED_S1_SW	47	14	2,698	357	22.036	57.922	2,698	501	7.670	56.486	2,698	364	0.439	36.915
MED_S2_LW	46	14	2,683	142	7.483	22.447	2,683	244	3.002	30.339	2,683	148	0.157	18.629
MED_S2_SW	46	14	2,683	293	14.270	48.267	2,683	470	6.643	65.761	2,683	328	0.291	42.883
MED_S3_LW	57	14	2,815	192	7.798	27.219	2,815	284	3.376	36.132	2,815	177	0.181	23.673
MED_S3_SW	57	14	3,073	288	26.065	68.414	3,073	522	8.944	81.637	3,073	370	0.304	51.973
MED_SUBTOTAL	-	-	-	1,497	87.466	252.858	-	2,258	32.619	292.269	-	1,545	1.594	188.864
LAR_S1_LW	68	19	3,794	232	8.954	33.097	3,794	254	3.041	31.120	3,794	182	0.164	23.368
LAR_S1_SW	68	19	3,794	337	20.808	59.209	3,794	511	7.808	82.521	3,794	356	0.236	51.090
LAR_S2_LW	70	19	3,791	218	10.001	44.289	3,791	266	3.139	49.285	3,791	203	0.270	36.716
LAR_S2_SW	70	19	3,791	274	13.968	61.857	3,791	452	6.169	91.626	3,791	325	0.239	63.641
LAR_S3_LW	80	19	3,878	231	10.016	52.589	3,878	327	4.023	64.902	3,878	270	0.265	54.426
LAR_S3_SW	80	19	4,192	360	22.697	86.705	4,192	458	8.624	104.037	4,192	367	0.465	76.711
LAR_SUBTOTAL	-	-	-	1,652	86.444	337.746	-	2,268	32.804	423.491	-	1,703	1.639	305.952
TOTAL	**-**	**-**	**-**	**4,693**	**257.223**	**782.804**	**-**	**6,527**	**94.820**	**880.756**	**-**	**4,687**	**4.745**	**601.352**

Interpretation of results. The results of the numerical study are shown in Table 4.1. In our analysis, we compare the root node LB in monetary units (mu), the number of MP iterations required to terminate column generation in the root node (#itMP), and the total time required to solve the S-SP $t_{s\text{-}sp}$. This includes updating the edge weights of the graph in S-SP NF. Furthermore, we display the total time required to terminate column generation in the root node t_{total}.

Across all instances, solving S-SP takes longest for S-SP IP optimally solved and shortest for S-SP NF. When accepting feasible solutions for S-SP IP, runtimes can be reduced by 63% compared to S-SP IP optimally solved. From a runtime perspective, S-SP NF provides the best results, with a runtime reduction of 98% compared to the runtime of S-SP IP optimally solved. Interestingly, when comparing S-SP IP optimally solved with S-SP IP feasibly solved, we find that the runtime advantages are overcompensated by an increased number of MP iterations that are necessary to prove MP-LP optimality. Across all instances, the number of MP iterations increases by 39%, from 4,693 to 6,527 iterations. As a result, the higher number of required iterations offsets the S-SP runtime gains. Consequently, dropping the optimality requirement for S-SP IP results in an increased total runtime of 13% compared to solving S-SP IP to optimality. The best results are achieved by S-SP NF, with a nearly equal number of required MP iterations as solving S-SP IP to optimality and a runtime reduction of 23%. An overview of the solution times of solving S-SP and proving MP-LP optimality in the root node is provided in Figure 4.6.
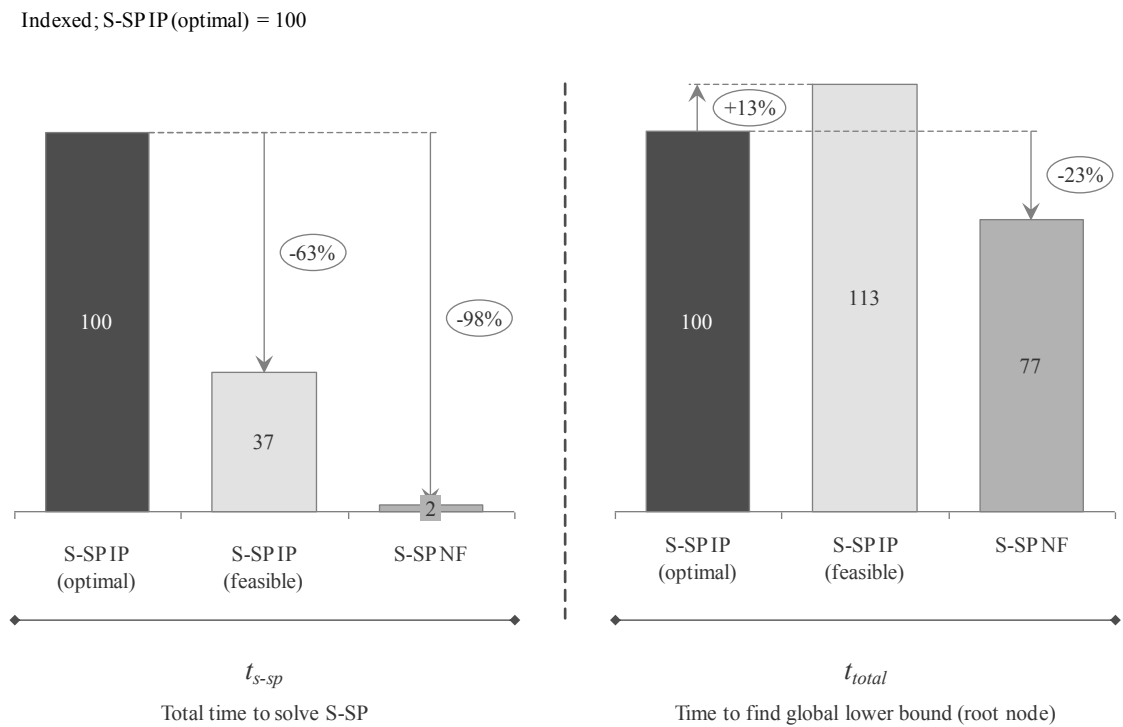

Figure 4.6: S-SP runtimes

4.5.2 Branch-and-price approach

In the following study, we demonstrate the behavior of the presented branch-and-price approach. When running the column generation part of the algorithm, we do not request T-SP to be solved to optimality but accept the first feasible solution that is found. We add the resulting task column to MP-LP. To solve S-SP, we rely on the NF formulation solved with the presented labeling algorithm. We stop the branch-and-price procedure at a node limit of 200 nodes and present the best obtained result.

Generating test instances. We generate 12 test instances applying the same overall logic as in Subsection 4.5.1. In line with the previous data sets, we rely on three task categories: day-long tasks, peak tasks, and precedence tasks. In order to generate small test instances for one week, we assign one day-long task to each day as well as three peak tasks for the morning, noon, and evening of each day. This results in four tasks per day or 28 tasks per week. We add two precedence tasks that can be performed from Monday through Friday, which results in 30 tasks per week in the base case. In the base demand profile D1, we set the length of each task randomly within the given range. This results in a total weekly demand of 68 man-hours. We create an additional demand profile by increasing the task lengths respecting the maximum task duration, resulting in 86 man-hours, namely D2. The two generated base case demand profiles have long task start windows (LW). We create additional demand patterns with reduced start windows (SW) in line with our approach applied in Chapter 3. Additional instances are generated by deleting the precedence tasks (NO_PREC) and the peak tasks, so that only the day-long tasks remain (DAY_ONLY). The approach provides 12 demand profiles that we use in our computational study. The shift parameters of inflexible and flexible workers are the same as in Subsection 4.5.1. The number of scheduled workers and the total demand in man-hours per instance are shown in Table 4.2. In total, we schedule between two and six workers.

Table 4.2: Number of workers

Instance	D1_LW	D1_LW_NO_PREC	D1_LW_DAY_ONLY	D1_SW	D1_SW_NO_PREC	D1_SW_DAY_ONLY	D2_LW	D2_LW_NO_PREC	D2_LW_DAY_ONLY	D2_SW	D2_SW_NO_PREC	D2_SW_DAY_ONLY
W^{fix} *[#]*	1	1	1	1	1	1	1	1	1	1	1	1
W^{flex} *[#]*	4	4	1	4	4	1	5	5	2	5	5	2
Demand *[man hours]*	68	66	45	68	66	45	86	78	50	86	78	50

Interpretation of results. The results of the computational study are presented in Table 4.4. For the root node and the end node, we present the best LB, the best UB, and the gap that is defined

as the difference between UB and LB divided by the UB. LB and UB are displayed in monetary units (mu). Also, we show the number of explored nodes and the ID of the node where the best incumbent solution was found. Regarding the solution times, we compare the time to solve the MP-LP t_{mp}, the time to solve T-SP $t_{t\text{-}sp}$, and the time to solve S-SP in total $t_{s\text{-}sp}$ and per SP, as well as the time to find the UBs heuristically t_{ub}. In addition, we display the total solution time t_{total} and the total solution time per explored node. All times are in seconds.

When terminating the solution procedure at the node limit of 200 nodes, none of the instances is solved to optimality. We illustrate that our algorithm finds optimal solutions based on one instance for which we omit the node limit. For the remaining 11 instances, a gap between the best integer solution and the global LB remains. Nevertheless, the gap in the end node is significantly reduced compared to the root node in all instances. While we interrupt the solution procedure at 200 nodes, preliminary testing suggests that significantly enlarging the number of explored nodes hardly results in better UBs or a proof of optimality. A complete enumeration of the search tree requires a huge number of nodes to be explored based on the applied branching scheme. Moreover, a problem of solving small instances becomes apparent in the study. As the degree of flexibility is limited with small instances, it may be that due to the minimum number of working periods in a shift, an oversupply in resource demand occurs that can be neither reduced nor "filled" with resource demand. This leads to rather large optimality gaps in the small test instances.

Based on the optimally solved problem instance D2_SW_DAY_ONLY, we illustrate that the proposed algorithm finds optimal solutions. In the root node, the column generation procedure finds a LB of 110. Solving MP-IP with a subset of task columns and all generated shift columns results in an initial feasible solution of 155, which is improved to 152 by the improvement heuristic. Consequently, in the root node, there is an optimality gap of 42 or 27.6%. In node 33, the rounding heuristic finds an improved feasible solution of 122, which equals an optimality gap of 9.8%. In node 100, the MP-IP heuristic is applied, which results in a new UB of 116 and an optimality gap of 5.2%. To prove that 116 is the optimal solution, the branch-and-price procedure continues until node 253. In node 196, the best LB is increased from 110 to 112, as the best LB of all active nodes equals 112. The optimality gap after increasing the LB equals 3.4%. In the final node 253, the set of active nodes is reduced to zero and consequently, all nodes have been explored and fathomed, as their LBs exceed or are equal to the global UB. Consequently, we have found the optimal solution and verified its optimality.

Table 4.3: MP-IP and improvement heuristics

Instance	Heuristics			
	MP-IP UB *[mu]*	Improv. UB *[mu]*	Improvement *[mu]*	*[%]*
D1_LW	296	284	12	4.1
D1_LW_NO_PREC	305	269	36	11.8
D1_SW	251	242	9	3.6
D2_LW	332	308	24	7.2
TOTAL	**1,184**	**1,103**	**81**	**6.8**

The proposed heuristic procedures deliver suitable UBs. Of the presented 12 test instances, UBs of five instances were generated by the MP-IP heuristic that considers a subset of all generated task columns and all generated shift columns. In four instances, the improvement heuristic can further enhance the obtained solution. This behavior is displayed in Table 4.3, where we show all instances where the MP-IP heuristic provides the final UB and an improvement is successful. We compare the UBs of the MP-IP heuristic and the improved UBs and outline the absolute and relative improvement. We see that a relative improvement between 3.6% and 11.8% is achieved. On average, the UBs are improved by 6.8%.

Exploring all nodes until the node limit of 200 takes between 20 and 458 seconds. The solution time for solving S-SP is the largest block of the different parts of the algorithm. However, when considering the number of S-SPs and dividing the total S-SP solution time by the number of S-SPs, solving T-SP clearly dominates the total solution time. For all D1 instances, solving T-SP requires 30% of the total solution time, while for all D2 instances, the share is 18%. The average is 22% across all instances. Solving S-SP NF per SP only takes 5% of the total solution time, which is significantly less than the time needed to solve T-SP. The time dedicated to heuristically generating UBs accounts for 4% of the total solution time, which seems to be a reasonable share given the potential advantages of cutting off nodes whose LBs exceed the best UB and thus reducing the size of the search tree. The runtime per node is within a range of 0.101 seconds as a minimum and 2.288 seconds as a maximum. While the lower end seems acceptable, runtimes of over two seconds per node result in very long total runtimes for proving optimality given the large number of nodes to be explored when completely enumerating the search tree.

Table 4.4: Results of the branch-and-price study

Instance	Root node			End node					Solution times						
	Best LB [mu]	Best UB [mu]	Gap [%]	Best LB [mu]	Best UB [mu]	Gap [%]	Expl. nodes [#]	Solut. node id [#]	t_{mp} [s.]	$t_{t\text{-}sp}$ [s.]	$t_{s\text{-}sp}$ [s.]	$t_{s\text{-}sp}/SP$ [s.]	t_{ub} [s.]	t_{total} [s.]	$t_{total}/node$ [s.]
D1_LW	164.0	2293.0	92.8	164.0	284.0	42.3	200	50	50.002	79.877	40.658	8.132	5.100	213.097	1.065
D1_LW_NO_PREC	160.0	1269.0	87.4	160.0	269.0	40.5	200	50	12.098	9.336	5.153	1.031	2.387	59.056	0.295
D1_LW_DAY_ONLY	101.0	3134.0	96.8	101.0	116.0	12.9	200	24	3.259	7.368	4.253	2.127	2.374	34.164	0.171
D1_SW	167.0	268.0	37.7	167.0	242.0	31.0	200	50	11.774	26.903	16.213	3.243	2.484	87.370	0.437
D1_SW_NO_PREC	167.0	269.0	37.9	167.0	212.0	21.2	200	76	7.124	25.659	29.428	5.886	1.889	93.572	0.468
D1_SW_DAY_ONLY	103.0	2134.0	95.2	103.0	116.0	11.2	200	22	1.130	3.128	2.053	1.027	1.267	20.163	0.101
D1_SUBTOTAL									85.387	152.271	97.758	21.443	15.501	507.422	0.423
D2_LW	218.0	1341.0	83.7	218.0	308.0	29.2	200	200	29.589	22.063	10.107	1.685	9.722	106.824	0.534
D2_LW_NO_PREC	194.0	1326.0	85.4	194.0	245.0	20.8	200	156	40.886	38.844	202.761	33.794	6.198	457.635	2.288
D2_LW_DAY_ONLY	110.0	182.0	39.6	110.0	122.0	9.8	200	38	3.840	6.929	4.886	1.629	2.453	39.703	0.199
D2_SW	225.0	305.0	26.2	225.0	248.0	9.3	200	190	29.436	60.377	38.521	6.420	9.837	181.786	0.909
D2_SW_NO_PREC	206.0	289.0	28.7	206.0	221.0	6.8	200	141	10.686	33.341	48.269	8.045	3.672	132.936	0.665
D2_SW_DAY_ONLY	110.0	152.0	27.6	116.0	116.0	0.0	253	100	5.961	15.198	11.895	3.965	7.008	66.611	0.263
D2_SUBTOTAL									120.398	176.752	316.439	55.537	38.890	985.495	0.787
TOTAL									205.785	329.023	414.197	76.980	54.391	1,492.917	0.609

4.6   Conclusion

In Chapter 4, we present a branch-and-price algorithm to solve the tactical integrated shift and task scheduling problem. We aim to find shift and task schedules for a given number of workers and a predefined set of tasks that have to be fulfilled. By scheduling shifts and tasks in an integrated manner, we exploit the flexibility in both areas. In shift scheduling, we consider two different worker categories – flexible and inflexible workers – that differ in their shift parameters and in their costs. An increased degree of shift flexibility comes at higher costs. We present a branch-and-price approach that relies on subvariable or constraint branching for driving solutions to integrality. To find upper bounds, we present two heuristics and one improvement heuristic. As part of our solution approach, we introduce a new NF formulation for the S-SP that is solved to optimality with a ShP-labeling algorithm. We demonstrate the superiority of the S-SP NF formulation in a computational study. Also, we compare runtimes of S-SP IP solved optimally versus S-SP IP solved feasibly. Interestingly, we find that runtime advantages of solving S-SP IP feasibly are overcompensated by an increased number of MP iterations that are necessary to prove MP-LP optimality. Based on 18 test instances building on real-world data, we show that the runtimes to solve the S-SP are reduced by 98% when using the NF formulation compared to solving the S-SP IP to optimality with standard solver CPLEX. Moreover, we demonstrate that the presented branch-and-price approach finds optimal solutions for small test instances.

This chapter of the thesis offers several future research opportunities. The presented model relies on a rather large number of different SPs. There are two SP types for task and shift scheduling. Moreover, there is one S-SP for each worker. Coordinating, updating, and incorporating SP solutions consume a lot of effort in the solution procedure. Consequently, future research should focus on how to reduce the handling effort. One possibility is to incorporate task scheduling in the MP. Standing alone, the T-SP consumes a high share of the solution time. Consequently, speeding up the T-SP could significantly reduce the computation time of the entire algorithm. Potential levers include a different T-SP formulation. While we find acceptable solutions for small problem instances in early stages of the branch-and-price procedure, enumerating the entire search tree with the proposed branching scheme requires a large number of nodes to be checked. Consequently, investigating alternative branching schemes could also improve the solution procedure.

In summary, we show that our newly introduced NF representation of the S-SP provides optimal solutions in a very short time. Moreover, we illustrate that the presented branch-and-price approach finds optimal solutions for small problem instances.

5 Conclusion

The thesis at hand aims to answer four research questions in the context of healthcare operations management in hospitals. In Chapter 1, we raised the following research questions:

(1)   Which areas of hospital logistics management have been addressed by the previous literature and which areas are most promising for future cost optimization?

(2)   What is the optimal number of logistics assistants in hospitals when fully leveraging the flexibility incorporated in shift and task scheduling?

(3)   What is the impact of flexibility in shift and task scheduling on the optimal number of logistics assistants?

(4)   What are optimal shift and task schedules when workers are employed with different degrees of flexibility and at different costs?

The presented research questions are addressed in the three major chapters of this work. Research question (1) is addressed in Chapter 2. We introduce hospital logistics management by providing a comprehensive literature review with a particular focus on publications that apply quantitative methods. We categorize publications along four major research streams, discuss the applied methodologies, and work out future research opportunities in each research stream. In addition, we present overarching research opportunities. The four identified research streams are (1) supply and procurement, (2) inventory management, (3) distribution and scheduling, and (4) holistic supply chain management. Research stream (2) has received the most attention in the past, considering the number of previous publications. In total, we present and assess 145 publications. The significant increase in the number of publications in recent years indicates the growing importance of hospital logistics. For example, during the period 2012 to 2014, the number of publications nearly doubled compared to 2009 to 2011. We further identify a growing relevance of applying quantitative methods in hospital materials management.

The subsequent Chapter 3 addresses research questions (2) and (3). We introduce the integrated shift and task scheduling problem and address the strategic workforce sizing question when introducing logistics assistants in hospitals. We present a column generation solution approach to define the minimum number of employees leveraging flexibility in shifts and tasks. In the course of introducing our solution approach, we present a lower bound for staff minimization problems with an unknown number of available workers. This approach allows us to terminate column generation early compared to the column generation standard stopping criterion. The algorithm and its performance are tested with 48 problem instances that are based on real data and compared to benchmarks. We show that the proposed algorithm clearly outperforms a

benchmark solved with standard solver CPLEX. Moreover, we compare two different decomposition approaches of the original problem. We show that fully leveraging flexibility in shift and task scheduling can lead to a decrease of 40 to 49% of the required workforce, compared to the non-flexible case.

Research question (4) is finally addressed in Chapter 4, where we introduce a branch-and-price algorithm to solve the subsequent tactical shift and task scheduling problem. The goal is to define shift and task schedules for a predefined number of employees. We apply a subvariable or constraint branching scheme in order to drive solutions to integrality. Two heuristics and one improvement heuristic are used to generate good upper bounds. The model relies on two worker categories that differ in their degree of flexibility and in their costs. As part of the solution approach, we introduce a shortest path network flow formulation for the shift subproblem that is solved to optimality with a labeling algorithm. Based on a numerical study, we demonstrate the superiority of the new formulation compared to an integer program subproblem solved with standard solver CPLEX. We show that the runtime for solving the shift pricing problem can be reduced by up to 98%. Also, we show that the proposed algorithm is appropriate for efficiently solving small-sized test instances.

Chapters 3 and 4 build on comparable assumptions for shift and task scheduling but address two different hierarchical planning levels. Chapter 3 addresses strategic workforce sizing, while Chapter 4 addresses the subsequent tactical shift and task scheduling problem for workers with different degrees of flexibility and different costs. In both chapters, we present optimal solution procedures that are based on column generation. Chapter 3 presents an approach where we use information from column generation and a sophisticated heuristic procedure to warm-start the initial mixed-integer program. In Chapter 4, we present a branch-and-price algorithm.

Our integrated scheduling approach works particularly well when shift and task scheduling adhere to a high degree of flexibility (see Chapter 3). Flexibility in shift scheduling is incorporated by relying on an implicit shift formulation, i.e., there are no predefined, explicit shift types. Instead, shifts are modeled implicitly so that collaborative agreements, legislative requirements, and labor contracts are fulfilled. Task scheduling incorporates flexibility because task start times may be set very flexibly as long as precedence constraints and start time windows are adhered. In task scheduling, the high degree of flexibility seems reasonable for the logistics tasks in scope. The task occurrence can be predicted accurately and their length and resource demand is deterministic and known. In shift scheduling, however, one might argue that the assumed high degree of flexibility is overly optimistic due to the negative implications of flexible shifts on employees. Working in shifts that are implicitly scheduled may result in shift patterns that are characterized by a high variance. For example, days-on and days-off may

alternate week by week and shift start times and durations might be different every day. This might negatively impact job satisfaction. Moreover, implementing implicitly scheduled shifts may be subject to high organizational burdens. Under normal conditions, the optimal solution of our model should be close to a feasible solution in practice. However, we acknowledge that our solution provides a lower bound rather than a solution with large buffers. Due to operational constraints, it might consequently become necessary to employ more workers than suggested by our model.

The thesis at hand offers several opportunities for future research. In order to overcome the acceptance and implementation issues outlined above, our model could be extended to incorporate employee preferences into shift scheduling. It would be worthwhile if workers could state their preferences regarding days-off, shift lengths, and shift start times. This could increase acceptance among workers and facilitate the implementation of the implicitly modeled shift schedules. In task scheduling, future research could incorporate stochasticity. In our model, we assume that tasks are well projectable and that their lengths and resource demands are deterministic. While the presented assumptions are valid for our use case, incorporating stochasticity could further extend the scope of application of our approach to settings characterized by stochastic task occurrences. In terms of the proposed methodology, future research should focus on ways to speed-up task scheduling. While we propose an efficient and well-performing methodology for shift scheduling, task scheduling relies on a rather slow IP formulation solved with a standard solver. Speeding-up task scheduling holds high potential to shorten the overall solution times.

In summary, we believe that integrating shift and task scheduling can help to further optimize staff sizing and scheduling decision-making in hospitals. Consequently, it can help to reduce healthcare costs and tackle the challenges presented in the introduction. We demonstrate the applicability of our integrated approach in the healthcare industry but believe that integrating shift and task scheduling can also be applied in other industries beyond healthcare, for example manufacturing or service industries.

Appendix I: Linearization of constraints (3)

According to Brunner & Edenharter (2011), constraints (3) can be replaced with the three linear constraints (3*), (3**), and (3***).

$$y_{wt}^{start} = y_{wt}\left(1 - y_{w(t-1)}\right) \qquad \forall w \in \mathcal{W}, t \in \mathcal{T} \setminus \{1\} \qquad (3)$$

Linearization

$$-y_{wt} + y_{wt}^{start} \leq 0 \qquad \forall w \in \mathcal{W}, t \in \mathcal{T} \setminus \{1\} \qquad (3^*)$$

$$y_{w(t-1)} + y_{wt}^{start} \leq 1 \qquad \forall w \in \mathcal{W}, t \in \mathcal{T} \setminus \{1\} \qquad (3^{**})$$

$$y_{wt} - y_{w(t-1)} - y_{wt}^{start} \leq 0 \qquad \forall w \in \mathcal{W}, t \in \mathcal{T} \setminus \{1\} \qquad (3^{***})$$

Appendix II: Algorithm details for strategic workforce sizing

Algorithm
STEP 1: Find lower bound
1:   initialize MP-LP, T-SP, S-SP
2:   set $it = 0$
3:   while $\left(z_i^{LB} < z_i\right)$
4:   solve MP-LP to obtain z_i
5: $it = it + 1$
6:   calculate z_i^{LB}
7:   if $\left(z_i^{LB} \geq z_i\right)$
8:   return z_i^{LB}
9:   terminate while and go to *STEP 2*
10:   else
11:   if (*it* even)
12:   solve T-SP
13:   else
14:   solve S-SP
15:   end if
16:   if (new column prices out)
17:   add column to MP
18:   else if (S-SP infeasible AND T-SP infeasible)
19:   return $\lceil z_i \rceil$
20:   terminate while and go to *STEP 2*
21:   end if
22:   end if
23:   end while
STEP 2: Find start solution
24:   for all generated task columns
25:   calculate $D_p^{max} = \max_{t \in \mathcal{T}} D_{pt}$
26:   end for
27:   select 10% of plans (max. 15) with minimum D_p^{max}
28:   for all selected task columns
29:   select demand D_t^p
30:   while $\left(\sum_{t \in \mathcal{T}} D_t^p > 0\right)$
31:   generate shift columns heuristically with updated S-SP
32: $D_t^p \leftarrow D_t^p - y_t$
33:   end while
34:   end for
35:   solve MP-IP with selected task columns and all generated shift columns
STEP 3: Improve start solution
36:   use MIP in compact formulation (1) - (8)
37:   add lower bound z_i^{LB} from *STEP 1* as hard constraint
38:   use feasible solution from *STEP 2* to warm start
39:   terminate and report results

Appendix III: Abbreviations

ADM	Automated Dispense Machine
BCMA	Barcode Medication Administration System
BNH	Buy-and-Hold
BOOFLP	Bi-objective Obnoxious Facility Location Problem
CMI	Co-managed Inventory
CP	Central Pharmacy
CSSD	Central Sterilization Service Department
CU	Care Unit
DTP	Direct-to-Pharmacy
FFS	Fee-for-Service
GPO	Group Purchasing Organization
IP	Integer Program
JIT	Just-In-Time
LB	Lower Bound
LP	Linear Program
MASTA	Multi Attribute Spare Tree Analysis
MIP	Mixed-Integer Program
MP	Master Problem
MP-IP	Master Problem – Integer Program
MP-LP	Master Problem – Linear Program
NF	Network Flow
OECD	Organization for Economic Co-operation and Development
OR	Operations Research
PVRP	Periodic Vehicle Routing Problem
RCPSP	Resource-Constrained Project Scheduling Problem
RFID	Radio Frequency Identification

SA	Simulated Annealing
SCM	Supply Chain Management
ShP	Shortest Path
SMDP	Semi Markov Decision Process
SP	Subproblem
S-SP	Shift-Subproblem
T-SP	Task-Subproblem
UB	Upper Bound
VED	Vital, Essential, Desirable
VMI	Vendor-managed Inventory
VNS	Variable Neighborhood Search
VRP	Vehicle Routing Problem
5S	Sort, Set to order, Shine, Standardize, and Sustain

Bibliography

Abijith, A. & Fosso Wamba, S., 2012. Business value of RFID-enabled healthcare transformation projects. *Business Process Management Journal*, 19(1), pp.1–28.

Abukhousa, E., Al-Jaroodi, J., Lazarova-Molnar, S. & Mohamed, N., 2014. Simulation and modeling efforts to support decision making in healthcare supply chain management. *The Scientific World Journal*, (Article ID 354246), pp.1–16.

Al-Qatawneh, L. & Hafeez, K., 2011. Healthcare logistics cost optimization using a multi-criteria inventory classification. In *Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management, January 22-24, 2011.* Kuala Lumpur, Malaysia, pp. 506–512.

Alfares, H.K. & Bailey, J.E., 1997. Integrated project task and manpower scheduling. *IIE Transactions*, 29(9), pp.711–717.

Alfares, H.K., Bailey, J.E. & Lin, W.Y., 1999. Integrated project operations and personnel scheduling with multiple labour classes. *Production Planning & Control*, 10(6), pp.570–578.

Almarsdóttir, A.B. & Traulsen, J.M., 2005. Cost-containment as part of pharmaceutical policy. *Pharmacy World and Science*, 27(3), pp.144–148.

Alspach, J.G., 2012. Is the drug shortage affecting patient care in your critical care unit? *Critical Care Nurse*, 32(1), pp.8–13.

Aptel, O. & Pourjalali, H., 2001. Improving activities and decreasing costs of logistics in hospitals: A comparison of U.S. and French hospitals. *The International Journal of Accounting*, 36(1), pp.65–90.

Artigues, C., Gendreau, M., Rousseau, L.M. & Vergnaud, A., 2009. Solving an integrated employee timetabling and job-shop scheduling problem via hybrid branch-and-bound. *Computers and Operations Research*, 36(8), pp.2330–2340.

Augusto, V. & Xie, X., 2009. Redesigning pharmacy delivery processes of a health care complex. *Health Care Management Science*, 12(2), pp.166–178.

Azzi, A., Persona, A., Sgarbossa, F. & Bonin, M., 2013. Drug inventory management and distribution: Outsourcing logistics to third-party providers. *Strategic Outsourcing: An International Journal*, 6(1), pp.48–64.

Baboli, A., Fondrevelle, J., Tavakkoli-Moghaddam, R. & Mehrabi, A., 2011. A replenishment

policy based on joint optimization in a downstream pharmaceutical supply chain: Centralized vs. decentralized replenishment. *The International Journal of Advanced Manufacturing Technology*, 57(1–4), pp.367–378.

Bailey, G., Cherrett, T., Waterson, B. & Long, R., 2013. Can locker box logistics enable more human-centric medical supply chains? *International Journal of Logistics Research and Applications*, 16(6), pp.447–460.

Bailey, J., Alfares, H. & Win Yuan Lin, 1995. Optimization and heuristic models to integrate project task and manpower scheduling. *Computers & Industrial Engineering*, 29(1), pp.473–476.

Banerjea-Brodeur, M., Cordeau, J.-F., Laporte, G. & Lasry, A., 1998. Scheduling linen deliveries in a large hospital. *The Journal of the Operational Research Society*, 49(8), pp.777–780.

Bard, J.F. & Purnomo, H.W., 2005. Preference scheduling for nurses using column generation. *European Journal of Operational Research*, 164(2), pp.510–534.

Barnhart, C., Johnson, E.L., Nemhauser, G.L., Savelsbergh, M.W.P. & Vance, P.H., 1998. Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, 46(3), pp.316–329.

Bassett, M., 2000. Assigning projects to optimize the utilization of employees' time and expertise. *Computers and Chemical Engineering*, 24(2), pp.1013–1021.

Baumer, A.M., Clark, A.M., Witmer, D.R., Geize, S.B., Vermeulen, L.C. & Deffenbaugh, J.H., 2004. National survey of the impact of drug shortages. *American Journal of Health-System Pharmacy*, 61, pp.2015–2023.

Bechtold, S.E. & Jacobs, L.W., 1990. Implicit modeling of flexible break assignments in optimal shift scheduling. *Management Science*, 36(11), pp.1339–1351.

Beier, F.J., 1995. The management of the supply chain for hospital pharmacies: A focus on inventory management practices. *Journal of Business Logistics*, 16(2), pp.153–174.

Beliën, J., De Boeck, L. & Van Ackere, J., 2012. Municipal solid waste collection and management problems: A literature review. *Transportation Science*, 48(1), pp.78–102.

Beliën, J. & Demeulemeester, E., 2008. A branch-and-price approach for integrating nurse and surgery scheduling. *European Journal of Operational Research*, 189(3), pp.652–668.

Beliën, J. & Demeulemeester, E., 2007. On the trade-off between staff-decomposed and activity-decomposed column generation for a staff scheduling problem. *Annals of*

Operations Research, 155(1), pp.143–166.

Beliën, J. & Demeulemeester, E., 2006. Scheduling trainees at a hospital department using a branch-and-price approach. *European Journal of Operational Research*, 175(1), pp.258–278.

Beliën, J., Demeulemeester, E., De Bruecker, P., Van Den Bergh, J. & Cardoen, B., 2013. Integrated staffing and scheduling for an aircraft line maintenance problem. *Computers and Operations Research*, 40(4), pp.1023–1033.

Beliën, J. & Forcé, H., 2012. Supply chain management of blood products: A literature review. *European Journal of Operational Research*, 217(1), pp.1–16.

Bendavid, Y., Boeck, H. & Philippe, R., 2010. Redesigning the replenishment process of medical supplies in hospitals with RFID. *Business Process Management Journal*, 16, pp.991–1013.

Bendavid, Y., Boeck, H. & Philippe, R., 2012. RFID-enabled traceability system for consignment and high value products: A case study in the healthcare sector. *Journal of Medical Systems*, 36(6), pp.3473–3489.

Van Den Bergh, J., Beliën, J., De Bruecker, P., Demeulemeester, E. & De Boeck, L., 2013. Personnel scheduling: A literature review. *European Journal of Operational Research*, 226(3), pp.367–385.

Beso, A., Franklin, B.D. & Barber, N., 2005. The frequency and potencial causes of dispensing erros in a hospital pharmacy. *Pharmacy World and Science*, 27(3), pp.182–190.

Bhakoo, V. & Chan, C., 2011. Collaborative implementation of e-business processes within the health-care supply chain: The Monash Pharmacy Project. *Supply Chain Management: An International Journal*, 16(3), pp.184–193.

Bhakoo, V., Singh, P. & Sohal, A., 2012. Collaborative management of inventory in Australian hospital supply chains: Practices and issues. *Supply Chain Management: An International Journal*, 17(2), pp.217–230.

Bijvank, M. & Vis, I.F.A., 2012. Inventory control for point-of-use locations in hospitals. *Journal of the Operational Research Society*, 63(4), pp.497–510.

Blazewicz, J., Lenstra, J.K. & Rinnooy Kan, A.H.G., 1983. Scheduling subject to resource constraints: Classification and complexity. *Discrete Applied Mathematics*, 5(1), pp.11–24.

Born, C.M. & Marino, D., 1995. Improving materials management through re-engineering. *Healthcare Financial Management*, 49(9), pp.30–37.

Brennan, C.D., 1998. Integrating the healthcare supply chain. *Healthcare Financial Management*, 52(1), pp.31–34.

Brucker, P., Qu, R. & Burke, E., 2011. Personnel scheduling: Models and complexity. *European Journal of Operational Research*, 210(3), pp.467–473.

Brunner, J.O., 2015. Flexible personnel scheduling using branch and price. *8th International Congress on Industrial and Applied Mathematics, August 11, 2015, Beijing, China*.

Brunner, J.O., Bard, J.F. & Köhler, J.M., 2013. Bounded flexibility in days-on and days-off scheduling. *Naval Research Logistics*, 60(8), pp.678–701.

Brunner, J.O., Bard, J.F. & Kolisch, R., 2009. Flexible shift scheduling of physicians. *Health Care Management Science*, 12(3), pp.285–305.

Brunner, J.O., Bard, J.F. & Kolisch, R., 2010. Midterm scheduling of physicians with flexible shifts using branch and price. *IIE Transactions*, 43(2), pp.84–109.

Brunner, J.O. & Edenharter, G.M., 2011. Long term staff scheduling of physicians with different experience levels in hospitals using column generation. *Health Care Management Science*, 14(2), pp.189–202.

Brunner, J.O. & Stolletz, R., 2014. Stabilized branch and price with dynamic parameter updating for discontinuous tour scheduling. *Computers and Operations Research*, 44, pp.137–145.

Bundesagentur für Arbeit/German Federal Employment Agency, 2014. *Der Arbeitsmarkt in Deutschland - Fachkräfteengpassanalyse*, Nürnberg.

Burke, E.K., De Causmaecker, P. & Vanden Berghe, G., 2004. The state of the art of nurse scheduling. *Journal of Scheduling*, 7(6), pp.441–499.

Burke, E.K., De Causmaecker, P., Petrovic, S. & Vanden Berghe, G., 2006. Metaheuristics for handling time interval coverage constraints in nurse scheduling. *Applied Artificial Intelligence*, 20(9), pp.743–766.

Burns, L.R. & Lee, J.A., 2008. Hospital purchasing alliances: Utilization, services, and performance. *Health Care Management Review*, 33(3), pp.203–215.

Çakici, Ö.E., Groenevelt, H. & Seidmann, A., 2011. Using RFID for the management of pharmaceutical inventory-system optimization and shrinkage control. *Decision Support Systems*, 51(4), pp.842–852.

Caprara, A., Monaci, M. & Toth, P., 2003. Models and algorithms for a staff scheduling

problem. *Mathematical Programming*, 98(1–3), pp.445–476.

Cattani, K.D., Jacobs, F.R. & Schoenfelder, J., 2011. Common inventory modeling assumptions that fall short: Arborescent networks, Poisson demand, and single-echelon approximations. *Journal of Operations Management*, 29(5), pp.488–499.

Çezik, T., Günlük, O. & Luss, H., 2001. An integer programming model for the weekly tour scheduling problem. *Naval Research Logistics*, 48(7), pp.607–624.

Chan, H.L., Choi, T.M. & Hui, C.L., 2012. RFID versus bar-coding systems: Transactions errors in health care apparel inventory control. *Decision Support Systems*, 54(1), pp.803–811.

Chandra, C., 2008. The case for healthcare supply chain management: Insights from problem-solving approaches. *International Journal of Procurement Management*, 1(3), pp.261–279.

Chang, C.H., Lai, Y.L. & Chen, C.C., 2012. Implement the RFID Position Based System of Automatic Tablets Packaging Machine for Patient Safety. *Journal of Medical Systems*, 36(6), pp.3463–3471.

Cheang, B., Li, H., Lim, A. & Rodrigues, B., 2003. Nurse rostering problems - A bibliographic survey. *European Journal of Operational Research*, 151(3), pp.447–460.

Correia, I. & Saldanha-da-Gama, F., 2015. A modeling framework for project staffing and scheduling problems. In C. Schwindt & J. Zimmermann, eds. *Handbook on Project Management and Scheduling Vol. 1*. Springer International Publishing Cham, pp. 547–564.

Côté, M.-C., Gendron, B. & Rousseau, L.-M., 2011. Grammar-based integer programming models for multiactivity shift scheduling. *Management Science*, 57(1), pp.151–163.

Coulson-Thomas, C.J., 1997. Process management in a hospital and healthcare context: Evidence from the HOCAPRIT project. *Business Process Management Journal*, 3(2), pp.118–132.

Coustasse, A., Tomblin, S. & Slack, C., 2013. Impact of radio frequency identification (RFID) technologies on the hospital supply chain: A literature review. *Perspectives in Health Information Management*, 10(Fall), pp.1–17.

Cruz, C.O. & Marques, R.C., 2013. Flexible contracts to cope with uncertainty in public-private partnerships. *International Journal of Project Management*, 31(3), pp.473–483.

Dacosta-Claro, I., 2002. The performance of material management in health care organizations. *International Journal of Health Planning and Management*, 17(1), pp.69–85.

Danas, K., Ketikidis, P. & Roudsari, A., 2002. A virtual hospital pharmacy inventory: An approach to support unexpected demand. *Journal of Medical Marketing*, 2(2), pp.125–129.

Danas, K., Roudsari, A. & Ketikidis, P.H., 2006. The applicability of a multi-attribute classification framework in the healthcare industry. *Journal of Manufacturing Technology Management*, 17(6), pp.772–785.

Dean, B., Van Ackere, A., Gallivan, S. & Barber, N., 1999. When should pharmacists visit their wards? An application of simulation to planning hospital pharmacy services. *Health Care Management Science*, 2(1), pp.35–42.

Dellaert, N. & Van De Poel, E., 1996. Global inventory control in an academic hospital. *International Journal of Production Economics*, 46–47, pp.277–284.

Dobrzykowski, D., Saboori Deilami, V., Hong, P. & Kim, S.-C., 2014. A structured analysis of operations and supply chain management research in healthcare (1982–2011). *International Journal of Production Economics*, 147, pp.514–530.

Doloresco, F., Fominaya, C., Schumock, G.T., Vermeulen, L.C., Matusiak, L., Hunkler, R.J., Shah, N.D. & Hoffman, J.M., 2011. Projecting future drug expenditures - 2011. *American Journal of Health-System Pharmacy*, 68(10), pp.921–932.

Van Donk, D.P., 2003. Redesigning the supply of gasses in a hospital. *Journal of Purchasing and Supply Management*, 9(5–6), pp.225–233.

Drezet, L.E. & Billaut, J.C., 2008. A project scheduling problem with labour constraints and time-dependent activities requirements. *International Journal of Production Economics*, 112(1), pp.217–225.

Epstein, R.H. & Dexter, F., 2000. Economic analysis of linking operating room scheduling and hospital material management information systems for just-in-time inventory control. *Anesthesia & Analgesia*, 91(2), pp.337–343.

Erhard, M.J.P., Schoenfelder, J., Fügener, A. & Brunner, J.O., 2016. State of the art in physician scheduling. *Working paper, Faculty of Business and Economics, University of Augsburg*.

Ernst, A.T., Jiang, H., Krishnamoorthy, M. & Sier, D., 2004. Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*, 153(1), pp.3–27.

Fakhimi, M. & Probert, J., 2013. Operations research within UK healthcare: A review. *Journal of Enterprise Information Management*, 26(1/2), pp.21–49.

Fein, A.J., 2007. Managing wholesaler agreements in the fee-for-service era. *American Journal*

of Health-System Pharmacy, 64(11), pp.1153–1153.

Ferretti, M., Favalli, F. & Zangrandi, A., 2014. Impact of a logistic improvement in an hospital pharmacy: Effects on the economics of a healthcare organization. *International Journal of Engineering, Science and Technology*, 6(3), pp.85–95.

Fillingham, D., 2007. Can lean save lives? *Leadership in Health Services*, 20(4), pp.231–241.

Fineman, S.J. & Kapadia, A.S., 1978. An analysis of the logistics of supplying and processing sterilized items in hospitals. *Computers & Operations Research*, 5(1), pp.47–54.

Fitzpatrick, R., Cooke, P., Southall, C., Kauldhar, K. & Waters, P., 2005. Evaluation of an automated dispensing system in a hospital pharmacy dispensary. *The Pharmaceutical Journal*, 274, pp.763–765.

Ford, E.W. & Scanlon, D.P., 2007. Promise and problems with supply chain management approaches to health care purchasing. *Health Care Management Review*, 32(3), pp.192–202.

Fosso Wamba, S., Anand, A. & Carter, L., 2013. A literature review of RFID-enabled healthcare applications and issues. *International Journal of Information Management*, 33(5), pp.875–891.

Fox, E.R., Birt, A., James, K.B., Kokko, H., Salverson, S., Soflin, D.L. & Hawkins, B., 2009. ASHP guidelines on managing drug product shortages in hospitals and health systems. *American Journal of Health-System Pharmacy*, 66(15), pp.1399–1406.

Fox, E.R. & Tyler, L.S., 2003. Managing drug shortages: Seven years' experience at one health system. *American Journal of Health-System Pharmacy*, 60(3), pp.245–253.

Franklin, B.D., O'Grady, K., Voncina, L., Popoola, J. & Jacklin, A., 2008. An evaluation of two automated dispensing machines in UK hospital pharmacy. *International Journal of Pharmacy Practice*, 16(1), pp.47–53.

Gebicki, M., Mooney, E., Chen, S.-J.G. & Mazur, L.M., 2014. Evaluation of hospital medication inventory policies. *Health Care Management Science*, 17(3), pp.215–229.

Gilmore, P.C. & Gomory, R.E., 1963. A linear programming approach to the cutting stock problem - part II. *Operations Research*, 11(6), pp.863–888.

Granlund, A. & Wiktorsson, M., 2013. Automation in healthcare internal logistics: A case study on practice and potential. *International Journal of Innovation and Technology Management*, 10(3), pp.1–20.

Griffith, M.M., Gross, A.E., Sutton, S.H., Bolon, M.K., Esterly, J.S., Patel, J.A., Postelnick, M.J., Zembower, T.R. & Scheetz, M.H., 2012. The impact of anti-infective drug shortages on hospitals in the United States: Trends and causes. *Clinical Infectious Diseases*, 54(5), pp.684–691.

Gu, A., Wertheimer, I.A., Brown, B. & Syaya, T.F., 2011. Drug shortages in the US - Causes, impact, and strategies. *Innovations in Pharmacy*, 2(4), pp.1–8.

Guerrero, W.J., Yeung, T.G. & Guéret, C., 2013. Joint-optimization of inventory policies on a multi-product multi-echelon pharmaceutical system with batching and ordering constraints. *European Journal of Operational Research*, 231(1), pp.98–108.

Guimarães, C.M., De Carvalho, J.C. & Maia, A., 2013. Vendor Managed Inventory (VMI): Evidences from lean deployment in healthcare. *Strategic Outsourcing: An International Journal*, 6(1), pp.8–24.

Gupta, D.K. & Huang, S.-M., 2013. Drug shortages in the United States: A critical evaluation of root causes and the need for action. *Clinical Pharmacology and Therapeutics*, 93(2), pp.133–135.

Gupta, R., Gupta, K.K., Jain, B.R. & Garg, R.K., 2007. ABC and VED analysis in medical stores inventory control. *Medical Journal Armed Forces India*, 63(4), pp.325–327.

Guyon, O., Lemaire, P., Pinson, É. & Rivreau, D., 2014. Solving an integrated job-shop problem with human resource constraints. *Annals of Operations Research*, 213(1), pp.147–171.

Haavik, S., 2000. Building a demand-driven, vendor-managed supply chain. *Healthcare Financial Management*, 54(2), pp.56–61.

Haijema, R., Van Der Wal, J. & Van Dijk, N.M., 2007. Blood platelet production: Optimization by dynamic programming and simulation. *Computers & Operations Research*, 34(3), pp.760–779.

Hall, R., Bryson, G.L., Flowerdew, G., Neilipovitz, D., Grabowski-Comeau, A. & Turgeon, A.F., 2013. Drug shortages in Canadian anesthesia: A national survey. *Canadian Journal of Anesthesia*, 60(6), pp.539–551.

Hartmann, S. & Briskorn, D., 2010. A survey of variants and extensions of the resource-constrained project scheduling problem. *European Journal of Operational Research*, 207(1), pp.1–14.

Heimerl, C. & Kolisch, R., 2010. Scheduling and staffing multiple projects with a multi-skilled

workforce. *OR Spectrum*, 32(2), pp.343–368.

Heinbuch, S.E., 1995. A case of successful technology transfer to health care just-in-time. *Journal of Management in Medicine*, 9(2), pp.48–56.

Hof, S., Fügener, A., Schoenfelder, J. & Brunner, J.O., 2015. Case mix planning in hospitals: A review and future agenda. *Health Care Management Science*, pp.1–14.

Hu, Q.J. & Schwarz, L.B., 2011. Controversial role of GPOs in healthcare-product supply chains. *Production and Operations Management*, 20(1), pp.1–15.

Huarng, F., 1998. Hospital material management in Taiwan: A survey. *Hospital Materiel Management Quarterly*, 19(4), pp.71–81.

Van Huele, C. & Vanhoucke, M., 2014. Analysis of the integration of the physician rostering problem and the surgery scheduling problem topical collection on systems-level quality improvement. *Journal of Medical Systems*, 38(6), pp.1–16.

Van Huele, C. & Vanhoucke, M., 2015. Decomposition-based heuristics for the integrated physician rostering and surgery scheduling problem. *Health Systems*, 4(3), pp.159–175.

Hulshof, P.J.H., Kortbeek, N., Boucherie, R.J., Hans, E.W. & Bakker, P.J.M., 2012. Taxonomic classification of planning decisions in health care: A structured review of the state of the art in OR/MS. *Health Systems*, 1(2), pp.129–175.

Huys, I. & Simoens, S., 2013. European drug shortages: A call for action! *International Journal of Pharmacy Practice*, 21(1), pp.1–2.

Iacocca, K., Zhao, Y. & Fein, A., 2013. Resell versus direct models in brand drug distribution. *International Journal of Pharmaceutical and Healthcare Marketing*, 7(4), pp.324–340.

Iannone, R., Lambiase, A., Miranda, S., Riemma, S. & Sarno, D., 2013. Modelling hospital materials management processes. *International Journal of Engineering Business Management*, 5(15), pp.1–12.

Iannone, R., Lambiase, A., Miranda, S., Riemma, S. & Sarno, D., 2014. Pulling drugs along the supply chain: Centralization of hospitals' inventory. *International Journal of Engineering Business Management*, 6(21), pp.1–11.

Irani, Z., Gunasekaran, A. & Dwivedi, Y.K., 2010. Radio frequency identification (RFID): Research trends and framework. *International Journal of Production Research*, 48(9), pp.2485–2511.

Jack, E.P. & Powers, T.L., 2009. A review and synthesis of demand management, capacity

management and performance in health-care services. *International Journal of Management Reviews*, 11(2), pp.149–174.

Jackson Healthcare, 2014. Practice trends and time at bedside. How certain foctors are impacting the satisfaction of registered nurses. *https://www.jacksonhealthcare.com/media/208952/rntrendsreport_ebook0114.pdf (accessed: August 11, 2015)*.

James, K.L., Barlow, D., McArtney, R., Hiom, S., Roberts, D. & Whittlesea, C., 2009. Incidence, type and causes of dispensing errors: A review of the literature. *The International Journal of Pharmacy Practice*, 17(1), pp.9–30.

Jarrett, P.G., 2006. An analysis of international health care logistics: The benefits and implications of implementing just-in-time systems in the health care industry. *Leadership in Health Services*, 19(1), pp.1–10.

Jarrett, P.G., 1998. Logistics in the health care industry. *International Journal of Physical Distribution & Logistics Management*, 28(9/10), pp.741–772.

Jaumard, B., Semet, F. & Vovor, T., 1998. A generalized linear programming model for nurse scheduling. *European Journal of Operational Research*, 107(1), pp.1–18.

Jayaraman, V., Burnett, C. & Frank, D., 2000. Separating inventory flows in the materials management department of Hancock Medical Center. *Interfaces*, 30(4), pp.56–64.

Johnson, P.E., 2011. Drug shortages: Impact and strategies. *JNCCN Journal of the National Comprehensive Cancer Network*, 9(8), pp.815–819.

Jun, J.B., Jacobson, S.H. & Swisher, J.R., 1999. Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society*, 50(2), pp.109–123.

Kaakeh, R., Sweet, B.V., Reilly, C., Bush, C., DeLoach, S., Higgins, B., Clark, A.M. & Stevenson, J., 2011. Impact of drug shortages on U.S. health systems. *American Journal of Health-System Pharmacy*, 68(19), pp.1811–1819.

Kafetzidakis, I. & Mihiotis, A., 2012. Logistics in the health care system: The case of Greek hospitals. *International Journal of Business Administration*, 3(5), pp.23–32.

Kelle, P., Woosley, J. & Schneider, H., 2012. Pharmaceutical supply chain specifics and inventory solutions for a hospital case. *Operations Research for Health Care*, 1(2), pp.54–63.

Kergosien, Y., Lenté, C., Billaut, J.-C. & Perrin, S., 2013. Metaheuristic algorithms for solving

two interconnected vehicle routing problems in a hospital complex. *Computers & Operations Research*, 40(10), pp.2508–2518.

Khurana, S., Chhillar, N. & Gautam, V.K.S., 2013. Inventory control techniques in medical stores of a tertiary care neuropsychiatry hospital in Delhi. *Health*, 5(1), pp.8–13.

Kim, C.S., Spahlinger, D.A., Kin, J.M. & Billi, J.E., 2006. Lean health care: What can hospitals learn from a world-class automaker? *Journal of Hospital Medicine*, 1(3), pp.191–199.

Kim, D., 2005. An integrated supply chain management system: A case study in healthcare sector. In K. Bauknecht et. al., ed. *E-Commerce and Web Technologies*. Springer-Verlag Berlin Heidelberg, pp. 218–227.

Kim, G.C. & Schniederjans, M.J., 1993. Empirical comparison of just-in-time and stockless materiel. *Hospital Materiel Management Quarterly*, 14(4), pp.65–74.

Kim, K. & Mehrotra, S., 2015. A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. *Operations Research*, 63(6), pp.1431–1451.

Van De Klundert, J., Muls, P. & Schadd, M., 2007. Optimizing sterilization logistics in hospitals. *Health Care Management Science*, 11(1), pp.23–33.

Kolisch, R. & Heimerl, C., 2012. An efficient metaheuristic for integrated scheduling and staffing IT projects based on a generalized minimum cost flow network. *Naval Research Logistics*, 59(2), pp.111–127.

Kollberg, B., Dahlgaard, J.J. & Brehmer, P.-O., 2007. Measuring lean initiatives in health care services: Issues and findings. *International Journal of Productivity and Performance Management*, 56(1), pp.7–24.

Koppel, R., Wetterneck, T., Telles, J.L. & Ben-Tzion, K., 2008. Workarounds to barcode medication administration systems: Their occurence, cause and threat to patient safety. *Journal of the American Informatic Association*, 15(4), pp.408–423.

Kriegel, J., Jehle, F., Dieck, M. & Mallory, P., 2013. Advanced services in hospital logistics in the German health service sector. *Logistics Research*, 6(2–3), pp.47–56.

Kumar, A., Ozdamar, L. & Zhang, C.N., 2008. Supply chain redesign in the healthcare industry of Singapore. *Supply Chain Management: An International Journal*, 13(2), pp.95–103.

Kumar, S., DeGroot, R.A. & Choe, D., 2008. Rx for smart hospital purchasing decisions: The impact of package design within US hospital supply chain. *International Journal of Physical Distribution & Logistics Management*, 38(8), pp.601–615.

Landry, S. & Beaulieu, M., 2010. Achieving lean healthcare by combining the two-bin kanban replenishment system with RFID technology. *International Journal of Health Management and Information*, 1(1), pp.85–98.

Landry, S. & Philippe, R., 2004. How logistics can service healthcare. *Supply Chain Forum*, 5(2), pp.24–30.

Lapierre, S.D. & Ruiz, A.B., 2007. Scheduling logistic activities to improve hospital supply systems. *Computers & Operations Research*, 34(3), pp.624–641.

Laursen, M.L., Gertsen, F. & Johansen, J., 2003. Applying lean thinking in hospitals - Exploring implementation difficulties. In *Proceedings of the 3rd International Conference on the Management of Healthcare and Medical Technology, September 7-9, 2003*. Warwick, UK, pp. 1–15.

Le, P., Seoane-Vazquez, E., Rodriguez-Monguio, R., Fox, E.R., Szeinbach, S.L., Dunehew, A.R. & Montagne, M., 2011. The prevalence of pharmaceutical shortages in the United States. *Journal of Generic Medicines*, 8(4), pp.210–218.

Lee, C.-P. & Shim, J.P., 2007. An exploratory study of radio frequency identification (RFID) adoption in the healthcare industry. *European Journal of Information Systems*, 16(6), pp.712–724.

Lee, H. & Özer, Ö., 2007. Unlocking the value of RFID. *Production and Operations Management*, 16(1), pp.40–64.

Little, J. & Coughlan, B., 2008. Optimal inventory policy within hospital space constraints. *Health Care Management Science*, 11(2), pp.177–183.

Liu, R., Xie, X. & Garaix, T., 2014. Hybridization of tabu search with feasible and infeasible local searches for periodic home health care logistics. *Omega*, 47, pp.17–32.

Lübbecke, M.E. & Desrosiers, J., 2005. Selected topics in column generation. *Operations Research*, 53(6), pp.1007–1023.

Maenhout, B. & Vanhoucke, M., 2015. A resource type analysis of the integrated project scheduling and personnel staffing problem. *Annals of Operations Research*.

Maenhout, B. & Vanhoucke, M., 2016. An exact algorithm for an integrated project staffing problem with a homogeneous workforce. *Journal of Scheduling*, 19(2), pp.107–133.

Maenhout, B. & Vanhoucke, M., 2013. An integrated nurse staffing and scheduling analysis for longer-term nursing staff allocation problems. *Omega*, 41(2), pp.485–499.

Maenhout, B. & Vanhoucke, M., 2010. Branching strategies in a branch-and-price approach for a multiple objective nurse scheduling problem. *Journal of Scheduling*, 13(1), pp.77–93.

Mangan, M.N. & Powers, M.F., 2011. Drug shortages and the role of the pharmacy technician: A Review. *Journal of Pharmacy Technology*, 27(6), pp.247–250.

Di Martinelly, C., Baptiste, P. & Maknoon, M.Y., 2014. An assessment of the integration of nurse timetable changes with operating room planning and scheduling. *International Journal of Production Research*, 52(24), pp.7239–7250.

Di Martinelly, C., Riane, F. & Guinet, A., 2009. A Porter-SCOR modelling approach for the hospital supply chain. *International Journal of Logistics Systems and Management*, 5(3–4), pp.436–456.

Di Mascolo, M. & Gouin, A., 2013. A generic simulation model to assess the performance of sterilization services in health establishments. *Health Care Management Science*, 16(1), pp.45–61.

Matopoulos, A. & Michailidou, L., 2013. Healthcare supply chains: A case study of hospital-vendor collaborative practices. *International Journal of Logistics Systems and Management*, 15(2), pp.288–303.

Maviglia, S.M., Yoo, J.Y., Franz, C., Featherstone, E., Churchill, W., Bates, D.W., Gandhi, T.K. & Poon, E.G., 2007. Cost-benefit analysis of a hospital pharmacy bar code solution. *Archives of Internal Medicine*, 167(8), pp.788–794.

Mazer-Amirshahi, M., Pourmand, A., Singer, S., Pines, J.M. & Van Den Anker, J., 2014. Critical drug shortages: Implications for emergency medicine. *Academic Emergency Medicine*, 21(6), pp.704–711.

Mazzocato, P., Savage, C., Brommels, M., Aronsson, H. & Thor, J., 2010. Lean thinking in healthcare: A realist review of the literature. *Quality and Safety in Health Care*, 19(5), pp.376–382.

McBride, A., Holle, L.M., Westendorf, C., Sidebottom, M., Griffith, N., Muller, R.J. & Hoffman, J.M., 2013. National survey on the effect of oncology drug shortages on cancer care. *American Journal of Health-System Pharmacy*, 70(7), pp.609–617.

McKone-Sweet, K.E., Hamilton, P. & Willis, S.B., 2005. The ailing healthcare supply chain: A prescription for change. *The Journal of Supply Chain Management*, 41(1), pp.4–17.

McLaughlin, M., Kotis, D., Thomson, K., Harrison, M., Fennessy, G., Postelnick, M. & Scheetz, M., 2013. Empty shelves, full of frustration: Consequences of drug shortages and

the need for action. *Hospital Pharmacy*, 48(8), pp.617–618.

McLaughlin, M.M., Pentoney, Z., Skoglund, E. & Scheetz, M.H., 2014. Projections for antiinfective drug shortages and time to actual resolution. *American Journal of Health-System Pharmacy*, 71(23), pp.2074–2078.

Medaglia, A.L., Villegas, J.G. & Rodríguez-Coca, D.M., 2008. Hybrid biobjective evolutionary algorithms for the design of a hospital waste management network. *Journal of Heuristics*, 15(2), pp.153–176.

Meijboom, B., Schmidt-Bakx, S. & Westert, G., 2011. Supply chain management practices for improving patient-oriented care. *Supply Chain Management: An International Journal*, 16(3), pp.166–175.

Meiller, Y., Bureau, S., Zhou, W. & Piramuthu, S., 2011. Adaptive knowledge-based system for health care applications with RFID-generated information. *Decision Support Systems*, 51(1), pp.198–207.

Michelon, P., Cruz, M.D. & Gascon, V., 1994. Using the tabu search method for the distribution of supplies in a hospital. *Annals of Operations Research*, 50(1), pp.427–435.

Moondra, S.L., 1976. An LP model for work force scheduling for banks. *Journal of Bank Research*, 7(4), pp.299–301.

Mühlich, M., Scherrer, M. & Daschner, F.D., 2003. Comparison of infectious waste management in European hospitals. *Journal of Hospital Infection*, 55(4), pp.260–268.

Mustaffa, N.H. & Potter, A., 2009. Healthcare supply chain management in Malaysia: A case study. *Supply Chain Management: An International Journal*, 14(3), pp.234–243.

Nachtmann, H. & Pohl, E., 2009. The state of healthcare logistics: Cost and quality improvement opportunities. *Center for Innovation in Healthcare Logistics, Arkansas*.

Narayana, S.A., Kumar Pati, R. & Vrat, P., 2014. Managerial research on the pharmaceutical supply chain - A critical review and some insights for future directions. *Journal of Purchasing and Supply Management*, 20(1), pp.18–40.

Nicholson, L., Vakharia, A.J. & Selcuk Erenguc, S., 2004. Outsourcing inventory management decisions in healthcare: Models and application. *European Journal of Operational Research*, 154(1), pp.271–290.

Nollet, J. & Beaulieu, M., 2003. The development of group purchasing: An empirical study in the healthcare sector. *Journal of Purchasing and Supply Management*, 9(1), pp.3–10.

Novek, J., 2000. Hospital pharmacy automation: Collective mobility or collective control? *Social Science & Medicine*, 51(4), pp.491–503.

OECD Publishing, 2015. *Fiscal sustainability of health systems: Bridging health and finance perspectives*, Paris.

OECD Publishing, 2013. *Health at a glance 2013*, Paris.

Oumlil, A. Ben & Williams, A.J., 2011. Strategic alliances and organisational buying: An empirical study of the healthcare industry. *International Journal of Procurement Management*, 4(6), pp.610–626.

Ozturk, O., Begen, M.A. & Zaric, G.S., 2014. A branch and bound based heuristic for makespan minimization of washing operations in hospital sterilization services. *European Journal of Operational Research*, 239(1), pp.214–226.

Pan, Z.X.T. & Pokharel, S., 2007. Logistics in hospitals: A case study of some Singapore hospitals. *Leadership in Health Services*, 20(3), pp.195–207.

Pasin, F., Jobin, M. & Cordeau, J.-F., 2002. An application of simulation to analyse resource sharing among health-care organisations. *International Journal of Operations & Production Management*, 22(4), pp.381–393.

Patterson, E.S., Cook, R.I. & Render, M.L., 2002. Improving patient safety by identifying side effects from introducing bar coding in medication administration. *Journal of the American Medical Informatics Association*, 9(5), pp.540–553.

Phillips, M.T. & Berner, E.S., 2004. Beating the system - Pitfalls of bar code medication administration. *Journal of Healthcare Information Management*, 18(4), pp.16–18.

Piccinini, P., Gamberini, R., Prati, A., Rimini, B. & Cucchiara, R., 2013. An automated picking workstation for healthcare applications. *Computers & Industrial Engineering*, 64(2), pp.653–668.

Pisinger, D. & Sigurd, M., 2007. Using decomposition techniques and constraint programming for solving the two-dimensional bin-packing problem. *INFORMS Journal on Computing*, 19(1), pp.36–51.

Poley, M.J., Bouwmans, C.A.M., Hanff, L.M., Roos, P.J. & Van Ineveld, B.M., 2004. Efficiency of different systems for medication distribution in an academic children's hospital in the Netherlands. *Pharmacy World and Science*, 26(2), pp.83–89.

Poon, E.G., Cina, J.L., Churchill, W., Patel, N., Featherstone, E., Rothschild, J.M., Keohane, C.A., Whittemore, A.D., Bates, D.W. & Gandhi, T.K., 2006. Medication dispensing errors

and potential adverse drug events before and after implementing bar code technology in the pharmacy. *Annals of Internal Medicine*, 145(6), pp.426–434.

Poulin, É., 2003. Benchmarking the hospital logistics process: A potential cure for the ailing health care sector. *CMA Management*, 77(1), pp.20–23.

Prashant, N.D., 1991. A systematic approach to optimization of inventory management functions. *Hospital Material Management Quarterly*, 12(4), pp.34–38.

Priyan, S. & Uthayakumar, R., 2014. Optimal inventory management strategies for pharmaceutical company and hospital supply chain in a fuzzy-stochastic environment. *Operations Research for Health Care*, 3(4), pp.177–190.

Purnomo, H.W. & Bard, J.F., 2007. Cyclic preference scheduling for nurses using branch and price. *Naval Research Logistics*, 54(2), pp.200–220.

Rais, A. & Viana, A., 2011. Operations research in healthcare: A survey. *International Transactions in Operational Research*, 18(1), pp.1–31.

Rego, N., Claro, J. & De Sousa, J.P., 2013. A hybrid approach for integrated healthcare cooperative purchasing and supply chain configuration. *Health Care Management Science*, 17(4), pp.303–320.

Rider, A.E., Templet, D.J., Daley, M.J., Shuman, C. & Smith, L.V., 2013. Clinical dilemmas and a review of strategies to manage drug shortages. *Journal of Pharmacy Practice*, 26(3), pp.183–191.

Rivard-Royer, H., Landry, S. & Beaulieu, M., 2002. Hybrid stockless: A case study - Lessons for health-care supply chain integration. *International Journal of Operations & Production Management*, 22(4), pp.412–424.

Rosales, C.R., 2011. Technology enabled new inventory control policies in hospitals. *Department of Operations and Business Analytics of the College of Business, University of Cincinnati*.

Rosales, C.R., Magazine, M. & Rao, U., 2014. Point-of-use hybrid inventory policy for hospitals. *Decision Sciences*, 45(5), pp.913–937.

Rosales, C.R., Magazine, M. & Rao, U., 2015. The 2Bin system for controlling medical supplies at point-of-use. *European Journal of Operational Research*, 243(1), pp.271–280.

Ross, A.D. & Jayaraman, V., 2009. Strategic purchases of bundled products in a health care supply chain environment. *Decision Sciences*, 40(2), pp.269–293.

Rossetti, M.D., Buyurgan, N., Pohl, E. & Converge, A., 2012. Medical supply logistics. In R. Hall, ed. *Handbook of Healthcare System Scheduling*. Springer New York Heidelberg Dordrecht London, pp. 245–280.

Ryan, D.M. & Foster, B.A., 1981. An integer programming approach to scheduling. In A. Wren, ed. *Computer Scheduling of Public Transport Urban Passenger Vehicle and Crew Scheduling*. North-Holland Publishing Company, pp. 269–280.

Schwindt (ed.), C. & Zimmermann (ed.), J., 2015. *Handbook on project management and scheduling vol. 1*, Springer International Publishing Cham.

Shih, L.-H. & Chang, H.-C., 2001. A routing and scheduling system for infectious waste collection. *Environmental Modeling and Assessment*, 6(4), pp.261–269.

Smet, P., Wauters, T., Mihaylov, M. & Vanden Berghe, G., 2014. The shift minimisation personnel task scheduling problem: A new hybrid approach and computational insights. *Omega*, 46, pp.64–73.

De Souza, L.B., 2009. Trends and approaches in lean healthcare. *Leadership in Health Services*, 22(2), pp.121–139.

Srivastava, S., 2008. Network design for reverse logistics. *Omega*, 36(4), pp.535–548.

Stolletz, R. & Brunner, J.O., 2012. Fair optimization of fortnightly physician schedules with flexible shifts. *European Journal of Operational Research*, 219(3), pp.622–629.

Swaminathan, J.M., 2003. Decision support for allocating scarce drugs. *Interfaces*, 33(2), pp.1–11.

Thomas, J.A., Martin, V. & Frank, S., 2000. Improving pharmacy supply-chain management in the operating room. *Healthcare Financial Management*, 54(12), pp.58–61.

Thompson, G.M., 1995. Improved implicit optimal modeling of the labor shift scheduling problem. *Management Science*, 41(4), pp.595–607.

Tlahig, H., Jebali, A., Bouchriha, H. & Ladet, P., 2013. Centralized versus distributed sterilization service: A location–allocation decision model. *Operations Research for Health Care*, 2(4), pp.75–85.

Trägårdh, B. & Lindberg, K., 2004. Curing a meagre health care system by lean methods-translating "chains of care" in the Swedish health care sector. *The International Journal of Health Planning and Management*, 19(4), pp.383–398.

U.S. Department of Labor, 2013. Economic news release: Employment projections 2012-2022

(USDL-13-2393).

Uthayakumar, R. & Priyan, S., 2013. Pharmaceutical supply chain and inventory management strategies: Optimization for a pharmaceutical company in a hospital. *Operations Research for Health Care*, 2(3), pp.52–64.

Vaillancourt, R., 2012. Drug shortages: What can hospital pharmacists do? *The Canadian Journal of Hospital Pharmacy*, 65(3), pp.175–179.

Valls, V., Pérez, Á. & Quintanilla, S., 2009. Skilled workforce scheduling in service centres. *European Journal of Operational Research*, 193(3), pp.791–804.

Vanderbeck, F., 2000. On Dantzig-Wolfe decomposition in integer programming and ways to perform branching in a branch-and-price algorithm. *Operations Research*, 48(1), pp.111–128.

Vanderbeck, F. & Wolsey, L.A., 1996. An exact algorithm for IP column generation. *Operations Research Letters*, 19(4), pp.151–159.

Varghese, V., Rossetti, M., Pohl, E., Apras, S. & Marek, D., 2012. Applying actual usage inventory management best practice in a health care supply chain. *International Journal of Supply Chain Management*, 1(2), pp.1–10.

Venkataraman, R. & Brusco, M.J., 1996. An integrated analysis of nurse staffing and scheduling policies. *Omega*, 24(1), pp.57–71.

Venkateswaran, S., Nahmens, I. & Ikuma, L., 2013. Improving healthcare warehouse operations through 5S. *IIE Transactions on Healthcare Systems Engineering*, 3(4), pp.240–253.

Ventola, C.L., 2011. The drug shortage crisis in the United States - Causes, impact, and management strategies. *Pharmacy and Therapeutics*, 36(11), pp.740–757.

Vila-Parrish, A.R. & Ivy, J.S., 2013. Managing supply critical to patient care: An intorduction to hospital inventory management for pharmaceuticals. In B. T. Denton, ed. *Handbook of Healthcare Operations Management: Methods and Applications*. Springer Science + Business Media New York, pp. 447–463.

Vila-Parrish, A.R., Ivy, J.S., King, R.E. & Abel, S.R., 2012. Patient-based pharmaceutical inventory management: A two-stage inventory and production model for perishable products with Markovian demand. *Health Systems*, 1(1), pp.69–83.

Volland, J., Fügener, A. & Brunner, J.O., 2016. A column generation approach for the integrated shift and task scheduling problem of logistics assistants in hospitals. *European Journal of Operational Research*, (forthcoming).

Volland, J., Fügener, A., Schoenfelder, J. & Brunner, J.O., 2016. Material logistics in hospitals: A literature review. *Omega*, (forthcoming).

De Vries, J., 2011. The shaping of inventory systems in health services: A stakeholder analysis. *International Journal of Production Economics*, 133(1), pp.60–69.

De Vries, J. & Huijsman, R., 2011. Supply chain management in health services: An overview. *Supply Chain Management: An International Journal*, 16(3), pp.159–165.

Whitson, D., 1997. Applying just in time systems in health care. *IIE Solutions*, 29(8), pp.32–37.

Wolsey, L.A., 1998. *Integer programming*, Wiley New York.

Wu, M.-C. & Sun, S.-H., 2006. A project scheduling and staff assignment model considering learning effect. *The International Journal of Advanced Manufacturing Technology*, 28(11–12), pp.1190–1195.

Yasin, M.M., Small, M.H. & Wafa, M.A., 2003. Organizational modifications to support JIT implementation in manufacturing and service operations. *Omega*, 31(3), pp.213–226.

Young, T., Brailsford, S., Connell, C., Davies, R., Harper, P. & Klein, J.H., 2004. Using industrial processes to improve patient care. *BMJ*, 328(7432), pp.162–164.

Young, T.P. & McClean, S.I., 2008. A critical look at lean thinking in healthcare. *Quality & Safety in Health Care*, 17(5), pp.382–386.

Zhao, H., Xiong, C., Gavirneni, S. & Fein, A., 2012. Fee-for-service contracts in pharmaceutical distribution supply chains: Design, analysis, and management. *Manufacturing & Service Operations Management*, 14(4), pp.685–699.

Zheng, J., Bakker, E., Knight, L., Gilhespy, H., Harland, C. & Walker, H., 2006. A strategic case for e-adoption in healthcare supply chains. *International Journal of Information Management*, 26(4), pp.290–301.