Integrated demand management and vehicle routing problems in last-mile delivery

Modeling, analytical discussion, and solution approaches

Vienna Klein

Vollständiger Abdruck der von der Fakultät für Wirtschafts- und Organisationswissenschaften der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines

Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

genehmigten Dissertation.

Gutachter:

Univ.-Prof. Dr. Claudius Steinhardt

Univ.-Prof. Dr. Andreas Brieden

Die Dissertation wurde am 30.05.2022 bei der Universität der Bundeswehr München eingereicht und durch die Fakultät für Wirtschafts- und Organisationswissenschaften am 10.08.2022 angenommen. Die mündliche Prüfung fand am 23.09.2022 statt.

Integrated demand management and vehicle routing problems in last-mile logistics

Modeling, analytical discussion, and solution approaches

Vienna Klein

Abstract

The proliferation of e-commerce and the progress of communication technology has led to the emergence and establishment of new business models in last-mile delivery. In *attended home delivery*, e.g., the customer and the provider agree on a certain delivery time window for a certain day, in which the provider promises delivery. In *same-day delivery*, e.g., the customer expects to receive the delivery on short notice, i.e., within a few hours. These new business models have in common that they come along with very high customer expectations regarding offered services, delivery speed, and the accuracy of shipping notifications. As a consequence, managing last-mile delivery has evolved from optimizing fulfillment operations alone to, additionally, steering demand, i.e., to integrate demand management and vehicle routing. Therewith, providers are able to steer customer choices toward efficient delivery options and, at the same time, realize higher prices for some delivery options such that additional revenue can be generated.

However, despite its high relevance, neither in the scientific literature nor in the common industrial practice there exists a common understanding of such integrated demand management and vehicle routing problems (i-DMVRPs) or a respective modeling framework. Furthermore, there is no integrative and anticipatory solution approach for an i-DMVRP in a same-day delivery setting.

This dissertation is a comprehensive contribution to the research on i-DVMRPs by closing those research gaps. In particular, in this dissertation, a detailed but general definition of i-DMVRPs is derived from literature and practice and a respective modeling framework is developed and discussed analytically. Further, in this dissertation, the first integrative and anticipatory solution approach for an i-DMVRP in a same-day delivery context is developed, presented, and evaluated comprehensively.

Zusammenfassung

Der Wachstum des elektronischen Handels und die Fortschritte in der Kommunikationstechnologie haben zur Etablierung neuer Geschäftsmodelle in der Zustellung auf der letzten Meile geführt. Beim *Attended Home Delivery* z.B. vereinbaren der Kunde und der Anbieter ein bestimmtes Zeitfenster für einen bestimmten Tag, in dem die Lieferung garantiert wird. Beim *Same-Day Delivery* z.B. erwartet der Kunde die Lieferung innerhalb weniger Stunden am selben Tag. Diesen neuen Geschäftsmodellen ist gemein, dass sie mit sehr hohen Kundenerwartungen hinsichtlich der angebotenen Leistungen, der Liefergeschwindigkeit und der Genauigkeit des Lieferavis einhergehen. Infolgedessen hat sich die Optimierung der letzten Meile von der rein operativen Betrachtung der Belieferungsprozesse und der Tourenplanung dahin gehend entwickelt, dass Nachfragemanagement und Tourenplanung integriert betrachtet werden. Damit sind Anbieter in der Lage, die Wahl der Kunden in Richtung kostengünstiger Lieferoptionen zu lenken und gleichzeitig höhere Preise für einige Lieferoptionen zu erzielen.

Trotz der hohen Relevanz dieses Themas gibt es weder in der wissenschaftlichen Literatur noch in der Praxis ein einheitliches Verständnis der resultierenden integrierten Demand-Management- und Vehicle-Routing-Probleme (i-DMVRPs) oder einen entsprechend einheitlichen Modellierungsansatz. Darüber hinaus gibt es keinen integrativen und antizipativen Lösungsansatz für Same-Day Delivery Probleme.

Diese Forschungslücken werden mit der vorliegenden Dissertation geschlossen. Es wird eine detaillierte, aber allgemeine Definition von i-DMVRPs aus Literatur und Praxis abgeleitet und ein entsprechender Modellierungsrahmen entwickelt und analytisch diskutiert. Darüber hinaus wird der erste integrative und antizipative Lösungsansatz für ein Same-Day Delivery Problem entwickelt, vorgestellt und umfassend evaluiert.

Table of contents

List of tables 6

List of figures 8

List of procedures 9

List of abbreviations 11

I Introduction 13

II Theoretical foundations 21

1 Demand management **27**

1.1 Definition, history, and conceptual delimitation 27

1.2 Approaches and modeling 28

1.2.1 Quantity-based demand management 29

1.2.2 Price-based demand management 32

1.3 Relevance of opportunity cost and their properties 33

1.4 Customer choice modeling 34

2 Vehicle routing **37**

3 Stochastic dynamic problems **41**

3.1 Modeling stochastic dynamic problems 41

3.2 Approaching stochastic dynamic problems 43

III Integrated demand management and vehicle routing problems (i-DMVRPs) 47

4 Description of i-DMVRPs 53

4.1 i-DMVRPs with disjoint booking and service horizons 54

4.2 i-DMVRPs with overlapping booking and service horizons 57

5 Literature on i-DMVRPs in last-mile delivery 63

5.1 Surveys and general frameworks for i-DMVRPs in last-mile delivery 64

5.2 Solution approaches for i-DMVRPs with disjoint booking and service horizons 68

5.2.1 Anticipatory learning-based approaches 69

5.2.2 Anticipatory non-learning-based approaches 70

5.3 Solution approaches for i-DMVRPs with overlapping booking and service horizons 72

5.3.1 Anticipatory learning-based approaches 73

5.3.2 Anticipatory non-learning-based approaches 75

5.4 Tabular overview and identification of research gaps 78

6 Modeling i-DMVRPs 85

6.1 General notation 85

6.2 Markov decision process model for i-DMVRPs with disjoint booking and service horizons 90

6.3 Markov decision process model for i-DMVRPs with overlapping booking and service horizons 96

6.3.1 Natural Model 96

6.3.2 Modified Model 103

6.3.3 Model equivalency 104

IV Analytical discussion of opportunity cost for i-DMVRPs 109

7 Transferring the concept of opportunity cost to i-DMVRPs 113

7.1 A general definition of opportunity cost 113
7.2 Comparing opportunity cost in traditional applications and i-DMVRPs 117
7.3 Extension of the concept of opportunity cost for i-DMVRPs 121

8 Properties and analytical discussion of opportunity cost for i-DMVRPs 125
8.1 Decomposability of opportunity cost into marginal cost to serve (MCTS) and displacement cost (DPC) 125
8.2 Potential negativity of MCTS and DPC 128
8.3 Non-negativity of opportunity cost 131
8.4 Monotonicity of value functions 136

V Development of a novel dynamic demand-management and online tour-planning approach for same-day delivery 139

9 General solution framework and hierarchical demand-management decomposition 145
9.1 Problem setting 145
9.2 General solution framework 146
9.3 Hierarchical demand-management decomposition 148

10 Sample-scenario value approximation and tour planning with two-tier demand-management integration 151
10.1 Deterministic same-day delivery demand-management and tour-planning problem with first-tier demand management 151
10.1.1 Mixed-integer programming formulation 151
10.1.2 Heuristic solution approach 154
10.2 Value approximation and second-tier demand management 157
10.3 Anticipatory tour planning 159

11 Computational study 161
11.1 Setup 161
11.1.1 Setting-independent parameters 161

11.1.2 Setting-dependent parameters 164
11.1.3 Instance generation 165
11.2 Value of anticipation 165
11.2.1 Experimental design and performance metrics 165
11.2.2 Numerical results 167
11.2.3 Analysis and insights 171
11.3 Value of explicit pricing optimization 173
11.3.1 Experimental design 173
11.3.2 Numerical results 174
11.3.3 Analysis and insights 176

VI Conclusion **179**

Bibliography **187**

List of tables

5.1 Surveys that feature related problems 65

5.2 Anticipatory solution approaches for i-DMVRPs in last-mile delivery 79

5.3 Classification of Parts IV and V with regard to the analyzed literature 83

6.1 Exemplary offer sets 87

6.2 General notation for i-DVMRP modeling 88

7.1 Customer requests of the problem instance underlying the discussion of opportunity cost properties in i-DVMRPs 118

7.2 Available offer sets of the problem instance underlying the discussion of opportunity cost properties in i-DVMRPs 119

11.1 Setting-dependent parameters 165

11.2 Results for setting 1V_100 and different look-ahead periods, averaged over 300 simulation runs 169

11.3 Results for setting 1V_100 and different look-ahead periods, averaged over 300 simulation runs 172

11.4 Results for setting 2V_100 and different look-ahead periods, averaged over 300 simulation runs 172

11.5 Results for setting 3V_100 and different look-ahead periods, averaged over 300 simulation runs 172

11.6 Results for setting 1V_150 and different look-ahead periods, averaged over 300 simulation runs 172

11.7 Results for setting 2V_150 and different look-ahead periods, averaged over 300 simulation runs 172

11.8 Results for setting 3V_150 and different look-ahead periods, averaged over 300 simulation runs 172
11.9 Results for setting 1V_200 and different look-ahead periods, averaged over 300 simulation runs 173
11.10 Results for setting 2V_200 and different look-ahead periods, averaged over 300 simulation runs 173
11.11 Results for setting 3V_200 and different look-ahead periods, averaged over 300 simulation runs 173
11.12 Exploitation of optimization levers by solution approaches 177

List of figures

1 Share of total supply chain costs by type worldwide 2018, Capgemini (2019) 15
2 Retail e-commerce sales worldwide from 2014 to 2025, eMarketer (2022) 16

6.1 Overview of the MDP model of the i-DMVRP booking and fulfillment process 89
6.2 Schematic representation of the transitions in a DJP 92
6.3 Schematic representation of the transitions in an OP 99

7.1 Overview of the MDP model of the i-DMVRP booking and fulfillment process including the interim state 114
7.2 Customer locations of the problem instance underlying the discussion of opportunity cost properties in i-DVMRPs 117

8.1 Decision Tree - Example 7.1 129
8.2 Decision Tree - Example 7.2 131

9.1 General solution framework 147

11.1 Arrival rates 164
11.2 Value of anticipation by different look-ahead horizons 167
11.4 Mean contribution margins across 300 simulation runs: boxplots and 95%-confidence intervals 169
11.5 Value of anticipation by different look-ahead horizons for setting 1V_100, 5 samples, averaged over 300 simulation runs 170

11.7 Pricing benchmark - results averaged over 300 simulation runs . . . 176

List of Algorithms

1 Feasibility repair scenarios 156

2 Tour-planning decision and post-processing 160

List of abbreviations

3PL	third-party logistics provider
AC-BP	availability control with base prices
ADP	approximate dynamic programming
AHD	attended home delivery
BAM	basic attraction model
DP	dynamic program
DPC	displacement cost
DJP	i-DMVRP with disjoint booking and service horizons
GAM	generalized attraction model
i-DMVRP	integrated demand management and vehicle routing problem
LMD	last-mile delivery
MCTS	marginal cost to serve
MDP	Markov decision process
MIP	mixed integer program
MNL	multinomial logit model
MOD	mobility-on-demand
OCBP	opportunity cost based pricing
OP	i-DMVRP with overlapping booking and service horizons
OPA	original pricing approach
p-MTVRP	profitable multi-trip vehicle routing problem

p-VRPRDT	profitable multi-trip vehicle routing problem with release and due times
p-VRPTW	profitable single trip vehicle routing problem with time windows
RSB	revenue shopping baskets
RD	revenue deliveries
DC	delivery costs
CM	contribution margin
MUM	maximum utility model
NOD	number of deliveries
SDD	same-day delivery
SDD-DMTP	same-day delivery demand-management and tour-planning problem
TOP	team orienteering problem
TSP	traveling-salesman problem
VFA	value function approximation
VRP	vehicle routing problem

Part I

Introduction

"The last mile is always the hardest."

While originally certainly not intended to describe logistic processes, this old proverb is indeed literally true for last-mile delivery (LMD), which describes the final step of a supply chain and, thus, usually defines the delivery of goods to private customers in urban areas (Boysen et al. 2021).

In fact, with a share of 40% to 50% of the overall supply chain cost, LMD is responsible for the largest share of costs in the supply chain (see Figure 1). This share is even expected to further increase due to considerably rising fuel cost, mainly driven by the Russia-Ukraine conflict (Andrew Travis (2022), World Bank (2022). Thus, efficient LMD execution is absolutely essential in preserving a profitable business that involves home delivery.

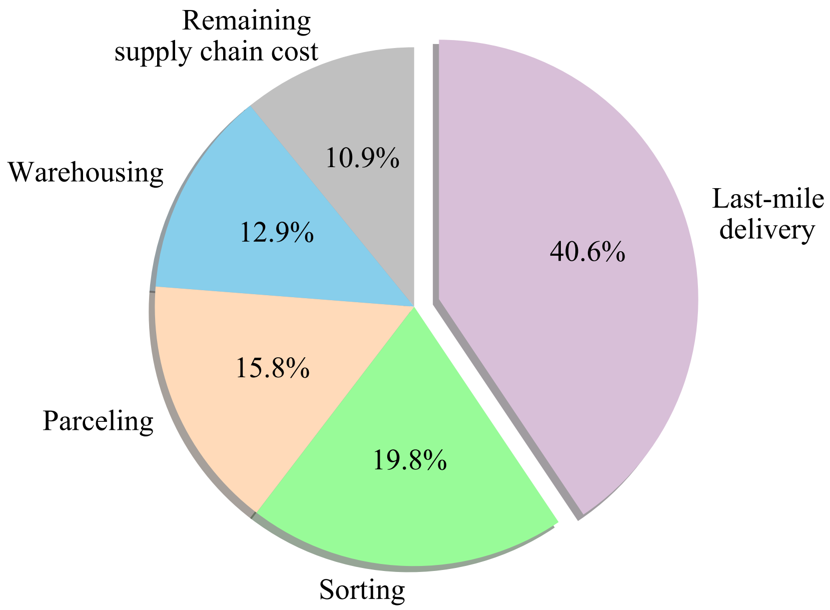


Figure 1: Share of total supply chain costs by type worldwide 2018, Capgemini (2019)

Most recently, in practice, the evolution of LMD has been substantially influenced by two major factors – the proliferation of e-commerce and the ongoing progress of communication technology:

Proliferation of e-commerce – From 2014 to 2019, global retail e-commerce sales nearly tripled, and they are forecast to nearly double over the next few years to reach an expected USD 7.4 trillion in sales volume by 2025 (see Figure 2). Therewith, the importance of LMD increases considerably, as LMD is a crucial part of e-commerce.

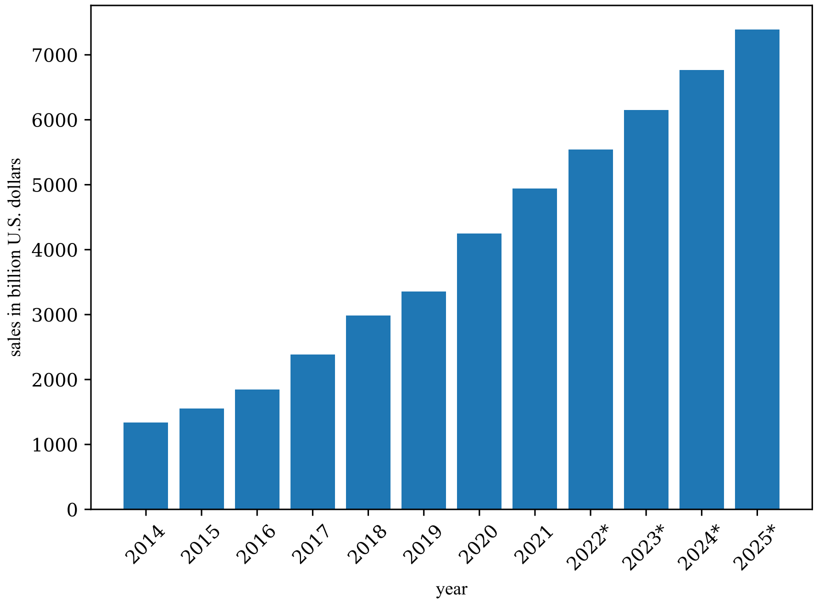


Figure 2: Retail e-commerce sales worldwide from 2014 to 2025, eMarketer (2022)

Progress of communication technology – The progress of communication technology has led to new possibilities regarding the interaction between providers and customers. This, in turn, accelerated the emergence and establishment of new business models such as *attended home delivery (AHD)* and *same-day delivery (SDD)* (Agatz et al. 2013). These business models bring opportunities and, at the same time, threats to existing LMD providers:

In *AHD*, the customer and the provider agree on a certain delivery time window for a certain day, in which the provider promises delivery (Koch and Klein (2020), Vinsensius et al. (2020)). This enables the expansion of e-commerce to goods, which have to be received in person by the customer, as for example pharmaceutical goods. Further, it helps reducing re-delivery cost substantially.

In *SDD*, the customer expects to receive the delivery on short notice, i.e., within a few hours (Voccia et al. (2019), Ulmer (2020a)). This enables the providers to expand their customer base to customers who seek instant gratification as with shopping in brick-and-mortar stores.

As a consequence, managing LMD has evolved from *optimizing fulfillment operations* alone to, additionally, steering demand towards efficient fulfillment operations, i.e., to incorporating *demand management* into LMD (Agatz et al. (2013), Yang and Strauss (2017), Klein et al. (2019), Fleckenstein et al. (2021)).

This is because, on the one hand, the described recent developments have led to very high customer expectations regarding offered services, delivery speed, and the accuracy of shipping notifications. On the other hand, in turn, they have led to differentiated, and therewith also higher willingness-to-pay from customer side (McKinsey and Company (2016), PwC (2018)). More precisely, providers are now able to steer customer choices and, at the same time, realize higher prices for some delivery options (Archetti and Bertazzi (2021), Agatz et al. (2021)). In summary, the resulting LMD processes can be planned to be more efficient and, further, additional revenue can be generated.

Due to the increasing importance of optimizing demand management and fulfillment operations in LMD practice, also a broad body of research emerged addressing the respective fields. This research mainly discusses the respective problems on an operational level (Waßmuth et al. 2022), as it holds high optimization potential, but with very complex optimization requirements. As shown by Figure 3, optimizing demand management and vehicle routing on the operational level has to be done in an integrative manner, i.e., under consideration of the effects each measure has on the other (Koch and Klein (2020), Fleckenstein et al. (2021)). This leads to the highly relevant, but exceptionally complex class of *integrated demand management and vehicle routing problems (i-DMVRPs)*. However, as will be shown in Chapter 5, in the broad body of research on i-DMVRPs, there is neither a common understanding of i-DMVRPs, nor a respective, general modeling framework. Instead, existing research mainly tackles either component individually and existing solution approaches do not involve anticipation for both components simultaneously.

This dissertation is a comprehensive contribution to the research on i-DMVRPs with regard to *problem definition*, *modeling*, *analytical discussion*, and *solution*

Decision making level	**Frequency**	**Long-term effect**	**State depen dency**	**Demand management**	**Fulfillment operations**
Strategic	low	high	static	Offering AHD/SDD? In which areas?	Operating own fleet or working with 3PL?
Tactical	medium	medium	static	Which delivery options/ delivery fees are generally offered?	How many vehicles? Which 3PL? Operating hours?
Operational	high	low	dynamic	What to offer when a certain request arrives?	How to route vehicles?

i-DMVRPs

Figure 3: Classification of i-DMVRPs into the different levels of decision making

approaches. In particular, in this dissertation, a detailed but general definition of i-DMVRPs is derived from literature and practice, which provides an overall taxonomy. This taxonomy forms building blocks to classify and analyze i-DMVRPs and their solution approaches, which is done within a complete literature discussion. Thereby, multiple research gaps are uncovered. The first two concern a unified modeling framework and an analytical discussion of the same. To fill these gaps, in this dissertation such unified modeling framework for i-DMVRPs is developed. Moreover, it is discussed analytically with a special focus on opportunity cost properties, which are essential for developing efficient solution approaches. Another major research gap identified is the non-existence of an integrative and anticipatory solution approach for i-DMVRPs in an SDD context. This research gap is also closed in this dissertation by the development of a respective solution approach. It includes a novel, specifically tailored anticipatory demand-management decomposition and an online tour-planning heuristic. In a comprehensive computational study it is shown, that this approach dominates benchmark approaches from the literature with regard to the overall contribution margin generated. Additionally, it is shown that, compared to a myopic benchmark approach, the contribution margin can be increased by up to 50%. Further, managerial insights are elaborated and summarized.

The remainder of the dissertation is structured as follows: After introducing the basic theory of related research fields, i.e., of demand management, vehicle routing problems (VRPs), and stochastic dynamic problems, in Part II, i-DMVRPs are comprehensively introduced in Part III. This includes the analysis of i-DMVRPs literature and the resulting identification of research gaps as well as the introduction of the unified modeling framework for i-DMVRPs. Afterwards, in Part IV the respective modeling framework is analyzed analytically and essential opportunity cost properties are derived and proven. Then, Part V is the heart of this dissertation. In this part, the newly developed approach for solving an i-DMVRP in an SDD problem setting is introduced.[1] Finally, Part VI concludes this dissertation by summarizing the most important findings from Part III to Part V.

[1]It has to be noted, that key insights from Parts III to V have already been made available as working papers named "On the concept of opportunity cost in integrated demand management and vehicle routing" and "Demand management and online routing for same-day delivery".

Part II

Theoretical foundations

As described in the Introduction, the widespread adoption of digital distribution channels enables logistical service providers to not only optimize and execute delivery fulfillment operations, but to intervene earlier in the overall process by actively managing the booking processes.

Incorporating demand management in addition to fulfillment optimization enables different optimization levers. Those can be classified according to their decision-making levels, depending on the long-term nature of the respective measures (Agatz et al. 2008). Figure 3 in the Introduction exemplarily shows which challenges can be addressed on which decision-making level. Decisions which are taken with low frequency and long-term effects are classified as decisions on the *strategic level*. Decisions which are taken with a medium frequency and a medium long-term effect are classified as decisions on the *tactical level*. Decisions which are taken with a high frequency and short-term effects are classified as decisions on the *operational level* (Boysen et al. 2021). Further, the decision-making level influences the required extent of the mutual integration of both components, demand management and fulfillment operations, within an optimization. Thereby, optimization on the operational level is the most complex, as it requires both components to be highly integrated, as the following discussion shows:

On the operational level, the *demand side* can be optimized, e.g., by dynamically deciding which customer requests to accept or reject, or which fulfillment options and/or prices to offer when a specific customer request arrives (see for example Bruck et al. (2018), Mackert (2019), Koch and Klein (2020)). Optimization potential on the *fulfillment operations* side can be leveraged by implementing sophisticated and anticipatory vehicle routing algorithms. Those can be applied to dynamically revise tour-plans according to the set of already confirmed customer orders and under consideration of expected incoming customer requests (see for example Voccia et al. (2019), Ulmer (2020a), Soeffker et al. (2021)). Thereby, predictions regarding incoming customer requests as well as the customer orders already confirmed strongly depend on the demand-management measures applied. At the same time, incorporating demand-management measures requires knowledge about the effect a potential demand-management decision has on the fulfillment operations in order to evaluate its efficiency.

Consequently, the i-DMVRPs considered in this dissertation share a common structure: A logistical service provider offers logistical services characterized by origin and destination in combination with other parameters, e.g., fees and time commitments. These services are offered throughout a booking horizon during which customer requests arrive *dynamically*. For every incoming customer request, the provider specifies

- the *availability of fulfillment options*,
- the *prices of fulfillment options*,
- or makes *accept/reject decisions*.

Subsequently, the customer makes a purchase choice, i.e., places an order, based on their *stochastic* individual preferences and the offered options. Fulfilling all customer orders takes place throughout the service horizon by means of a fixed number of vehicles. Capacities of other resources, as for example driver working hours, may also be limited. The booking and service horizons can be disjoint or overlapping.

Given the capacity restrictions as well as other operational constraints, such as potentially guaranteed service levels, the provider's objective is maximize profit by means of demand management and fulfillment optimization, i.e., routing optimization (Agatz et al. 2013). Hence, the operational planning of respective LMD applications is no longer limited to solving *VRPs*. Instead, providers integrate *demand management*. Thereby, generally, the demand-management component is a stochastic dynamic problem since customer requests arrive dynamically from random locations and, additionally, their choice behavior is unknown to the provider.

Therefore, in this part of the dissertation, brief introductions to the theory of *demand management*, *VRPs*, and *stochastic dynamic problems* is given. The goal is to provide a general foundation and mutual understanding of the taxonomies used in this dissertation. Further, it is the target to give the reader an idea of the considered fields of research and provide them with references to the relevant introductory literature.

Part II is structured as follows. In Chapters 1 and 2, the general foundation to *demand management* and *VRPs* is set. In Chapter 3, stochastic dynamic problems are introduced.

1. Demand management

1.1 Definition, history, and conceptual delimitation

The term *demand management* emerged from the broad body of literature that deals with revenue management. There is no consensus regarding the scope of the measures that are summarized under this term and how it relates to revenue management. Anderson and Carroll (2007) provide a very broad definition: "Demand management involves dynamically managing overall demand by optimising the use of distribution channels to reach target customer segments, leveraging and enhancing existing customer relationships, and taking effective RM [revenue management] actions." In their understanding, demand management expands "the tactical tools of RM to a more strategic level". Thus, they summarize every demand influencing measure from strategic marketing measures to operational price setting or availability control under the umbrella of demand management.

In contrast, other researchers understand demand management to be part of revenue management or even equate the two terms. Talluri and Van Ryzin (2006), Chapter 1, state: "RM [revenue management] is concerned with [...] demand management decisions and the methodology and systems required to make them." More recently, Agatz et al. (2013) conclude that "In fact, revenue management is demand management.".

Historically, revenue management originates from the airline industry in the 1970s. After the deregulation of the North American airline market, American Airlines introduced initial demand management measures. First, they introduced capacity controlled fares and shortly after, they implemented a comprehensive revenue management system in order to compete with new low-cost carriers (Gallego et al. (2019), Preface). Inspired by American Airlines' success story, revenue management found its way into related business models such as car rental, hotel industry, and cruise ships (Yeoman and McMahon-Beattie (2010), Chapter 1). Those traditional areas of revenue management application share the following characteristics: relatively fixed capacity, perishable inventory, variable but predictable demand, and a favorable cost structure, which is high fixed costs and low variable costs (Huefner (2015), Chapter 2). By this time, revenue management was also referred to as *yield management*, but finally, the term revenue management prevailed. It captures the typical cost structure of traditional areas of revenue management applications (Strauss et al. 2018): Variable cost can be neglected (Weatherford and Bodily 1992) and applying demand influencing measures are mainly targeted to improve revenue.

During the last decade, the research on revenue management expanded to new areas of applications such as manufacturing (Lohnert and Fischer 2019), last-mile delivery (Klein et al. 2020), and shared-mobility systems (Soppert et al. 2022). In those business models, variable fulfillment cost are not negligible and a shift from the term revenue management to the term demand management can be observed in the respective publications.

In this dissertation, the term demand management is used instead of the term revenue management, since i-DMVRPs address business models with non-negligible fulfillment cost. Nevertheless, the term summarizes the same measures and approaches from revenue management literature.

1.2 Approaches and modeling

Demand management approaches can be subdivided into *quantity-based* and *price-based* approaches. Both types are further described in the following. If not stated differently, the following description are based on the seminal book by Talluri and Van Ryzin (2006), Chapters 2.1 and 2.5, respectively.

1.2.1 Quantity-based demand management

Quantity-based demand-management approaches address the optimization of capacity allocation to demand originating from different customer segments, i.e., from customers with different willingness to pay. A simple problem setting is selling different fare classes of the same resource, e.g., selling comparable seats in the same compartment of a plane on the same flight for different fares and thereby addressing business customers with higher fares and leisure customers with lower fares. Quantity-based demand management is then concerned with dynamically controlling the availability and the number of available tickets of those different fare classes dynamically over the booking horizon.

Typical quantity-based types of demand-management approaches are determining *booking limits* or *protection levels*, or calculating *bid prices*, which are briefly outlined in the following. The presentation aims at giving the reader an initial idea of how to steer demand.

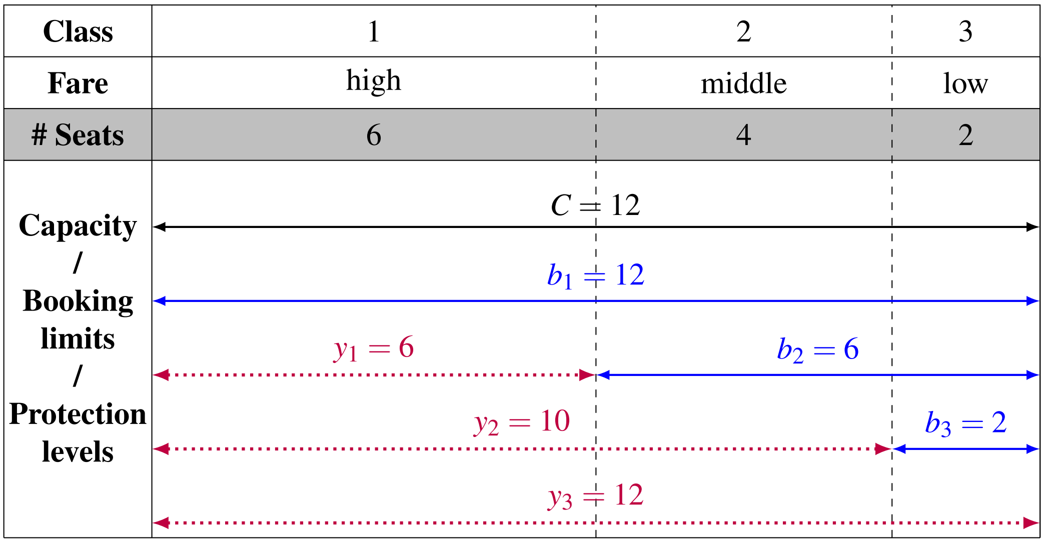


Figure 1.1: Relationship between booking limits, protection levels, and overall capacity C, cf. Talluri and Van Ryzin (2006), Chapter 2

Booking limits/Protection levels – The booking limit of a fare class defines the maximum capacity that is available for this and for lower fare classes. Protection levels can be understood as the complement; the protection level of a fare class defines the capacity that needs to be reserved for this and higher fare classes. Consequently, booking limits and protection levels have the following relationship: the booking limit of a fare class equals the overall capacity minus the sum of protection levels of all higher fare classes. This relationship is depicted in Figure 1.1 for an example with three fare classes, $i = \{1,2,3\}$, and an overall capacity of $C = 12$. The booking limit of a fare class i is denoted by b_i. The respective protection level is denoted by y_i.

Bid prices – Bid prices are threshold values defining whether an incoming request is accepted or rejected. More precisely, if a request's revenue exceeds a bid price, it is accepted. Otherwise, it is rejected. For optimal control, bid prices have to be updated after every sale and, if they depend on the remaining time of the booking horizon, with progress of time respectively. In this dissertation, bid price control is classified as a quantity based approach, in accordance with Talluri and Van Ryzin (2006). When used as threshold prices for accepting customer requests, the resulting number of customer request acceptances for certain fare classes equal those resulting from a respective control bybooking limits/protection level. However, bid prices can also be interpreted as price-based demand management as, of course, from bid prices, also dynamic prices to set for each incoming customer request can be derived.

Further, optimal demand-management decisions can be derived from dynamic programming formulations. In the following, two dynamic programming formulations for optimal, quantity-based demand management are presented. The first one models *availability control*, i.e., for every incoming customer request, it has to be decided whether it is accepted or rejected. The second one models *assortment optimization*, i.e., for every incoming customer request, it has to be decided which products to offer from a pre-defined set of alternatives. The following model for availability control stems from Talluri and Van Ryzin (2006), Chapter 2.5, the one for assortment optimization from Strauss et al. (2018):

Availability control – For modeling a basic availability control problem it is assumed that there are $t = 1, ..., T$ decision epochs in which customer requests from n

booking classes arrive. For every decision epoch t, a request of booking class j arrives with arrival probability $\lambda_j(t)$ such that the overall arrival rate equals $\sum_{j=1}^{n}(\lambda_j(t))$. Without loss of generality, it is assumed that decision epochs are set sufficiently small such that at most one customer request arrives per decision epoch. However, if a request of class j arrives and is accepted, a revenue r_j realizes for $j = 1,...,n$. For simplicity, a random variable $R(t)$ is introduced. It equals r_j if a customer request of class j arrives in decision epoch t. Otherwise $R(t) = 0$ holds. Further, decision variable u is introduced with $u = 1$ if there is a customer request and it is accepted. Otherwise $u = 0$ holds. The overall capacity is denoted by C and the remaining capacity is denoted by x. Then, the dynamic program (DP) for availability control can be represented by its value function $v_t(x)$ that equals the well-known Bellman equation:

$$\begin{aligned} v_t(x) &= \mathbb{E}[\max_{u\in\{0,1\}} \big(R(t)u + v_{t+1}(x-u)\big)] \\ &= v_{t+1}(x) + \mathbb{E}[\max_{u\in\{0,1\}} \big(\big(R(t) - \Delta v_{t+1}(x)\big)u\big)] \end{aligned} \tag{1.1}$$

with

$$\Delta v_{t+1}(x) = v_{t+1}(x) - v_{t+1}(x-1) \tag{1.2}$$

representing the expected marginal value of one unit of capacity in decision epoch $t+1$ and boundary conditions:

$$v_{T+1}(x) = 0 \qquad \forall x = 1,...,C, \tag{1.3}$$

and

$$v_t(0) = 0 \qquad \forall t = 1,...,T. \tag{1.4}$$

Assortment optimization – Assortment optimization models dynamically optimize the set of products offered to each incoming customer request. An offered set of products is called *offer set* and is denoted by g. The remaining capacity is denoted by a vector $\mathbf{x}$ and a product's i capacity consumption is represented by column vector $\mathbf{u}_i$. The set of offer sets that can feasibly be offered with remaining capacity $\mathbf{x}$ is denoted by $\mathscr{G}(\mathbf{x})$. The probability that an arriving customer chooses product i when an offer set g is offered is denoted by $P^i(g)$ and the corresponding revenue is denoted by r^i. Then, the assortment optimization problem can also be represented by its value function:

$$v_t(\mathbf{x}) = \max_{g \in \mathscr{G}(\mathbf{x})} [\sum_{i \in g} \lambda P^i(g)\big(r^i - \Delta_i v_{t+1}(\mathbf{x})\big)] \tag{1.5}$$

with

$$\Delta_i v_{t+1}(\mathbf{x}) = v_{t+1}(\mathbf{x}) - v_{t+1}(\mathbf{x} - \mathbf{u}_i) \tag{1.6}$$

representing the expected marginal value of the capacity consumption that is related with selling product i and boundary conditions (1.3) and (1.4).

For more details on quantity-based types of demand-management measures, the interested reader is referred to Talluri and Van Ryzin (2006).

1.2.2 Price-based demand management

Price-based demand management is closely related to quantity-based demand management in that the above described dynamic models (1.1) and (1.5) can also be applied in order to steer the availability of different fare classes or to manage different price lists (= combinations of discrete price points for different products) as offer sets, for example. In both approaches, customers *experience* the resulting demand-management control as dynamic pricing. However, there is a wide range of demand-management approaches that specifically aim at dynamic pricing, meaning that they explicitly optimize prices. The interested reader is referred to Talluri and Van Ryzin (2006), Chapter 5, for a detailed introduction into dynamic pricing. Further, the reader is referred to Koch and Klein (2020) for an example of how to incorporate a continuous pricing problem in an assortment optimization model.

1.3 Relevance of opportunity cost and their properties

Talluri and Van Ryzin (2006), Chapter 2, state that: "capacity should be allocated to a request if and only if its revenue is greater than the value of the capacity required to satisfy it." Thus, to evaluate potential demand-management decisions in general, a provider needs information about the marginal value of the resources required for fulfilling the resulting order. In the previously introduced models for availability control and assortment optimization, (1.1) and (1.5), this marginal value is derived from comparing two state values: the state value resulting from the conversion of the request into a certain order and the state value assuming no order is placed (c.f. Equations (1.2) and (1.6)). These differences are also referred to by the terms *(expected) displacement cost* or *opportunity cost* and in both models, Equations (1.1) and (1.5), the eminence of their role is obvious: the decision problem in both DP formulations essentially equals maximizing the expected difference of revenue and opportunity cost. (This can also be transferred to the DP formulations for price-based demand-management problems introduced in Talluri and Van Ryzin (2006) in Chapter 5.) Consequently, optimal policies can be derived from opportunity cost (as represented by Equations (1.2) and (1.6)). Further, with known opportunity cost, optimal policies can be replicated by the previously introduced types of demand-management approaches, i.e., booking limits/protection levels and bid prices, i.e., by setting bid prices equal to the given opportunity cost. In summary, demand-management problems can be decomposed into two steps: (1) calculating opportunity cost, (2) optimizing demand-management decisions with opportunity cost as input.

However, for most demand-management problems, the value function suffers from the "curses of dimensionality" (Koch and Klein (2020). and so does the calculation of opportunity cost. Thus, for determining optimal/good demand-management decisions, researchers typically rely on an approximation of opportunity cost (Gallego et al. (2019), Chapter 2). Thereby, exploiting known properties of the value function and the opportunity cost function, such as monotonicity or non-negativity, improves the accuracy of approximation approaches and subsequent demand-management decisions substantially. For the first task, i.e., the approximation of opportunity cost, approaches based on linear programming (Adelman 2007) as well as statistical learning (Koch 2017) are known to perform better if constraints are imposed that ensure that the resulting approximation also exhibits these properties. For the second task, i.e., solving the demand-management problem, the validity of certain opportunity cost properties yields direct insights of the resulting optimal policies or simplifies the approximation of good policies substantially (Gönsch and Steinhardt (2015), Maddah et al. (2010)).

1.4 Customer choice modeling

According to Agatz et al. (2013) "revenue management aims to exploit market heterogeneities". This means that revenue management bases on the observation that different customers may have different utilities and preferences, denoted by u^i, for different products i, and on the idea to exploit this. The latter can be achieved, e.g., by varying prices or offer sets for different customers and therewith provoke a favorable customer behavior. To do so, the resulting customer behavior has to be anticipated. For this, typically, past transactions are analyzed and customers who have the same utilities and preferences u^i for the same products i are allocated to the same *customer segment*. Afterwards, a choice model is chosen to represent the typical behavior of the customers belonging to a certain segment and the choice model's parameters are estimated. Finally, the resulting customer choice probabilities $P^i(g)$ for choosing to buy product i, when offer set g is offered, can be used as input to solve a demand-management problem as, e.g., represented by Equations (1.1) or (1.5). Generally, it holds that $P^i(g) = P(u^i \geq \max\{u^i : i \in g\})$, which is calculated differently for different customer choice models. Further, different customer choice models differ in their capability of capturing substitution effects among the offered products and competitors' products available in the market (Gallego et al. (2019), Chapter 4). In the following, only the most common choice models found in recent i-DMVRP literature are described, based on the seminal book of Gallego et al. (2019), Chapter 4. For more customer choice models and insights about their structural properties, how to integrate them into different demand-management modeling approaches, and about how to estimate specific customer choice model parameters, the interested reader is referred to Gallego et al. (2019), Train (2009), and Strauss et al. (2018).

Maximum utility model

As the name suggests, the *maximum utility model (MUM)* assigns a purchase probability $P^i(g) = 1$ to the product i, which has the highest utility u^i among the offered products g and $P^{i'}(g) = 0$ for all other products in g with $u^{i'} < u^i$. If there is more than one product with equal utilities higher than all other utilities, the respective choice probability is assigned uniformly among them.

Basic attraction model

The *basic attraction model (BAM)* assigns purchase probabilities $P^i(g) \neq 0$ for all products i in the set of offered products g, with I offered products, that have a utility $u^i \neq 0$. Further, the no-purchase option is also assigned a utility u^0. Then, the purchase probability $P^i(g)$ for a product i in g is determined as follows:

$$P^i(g) = \frac{u^i}{u^0 + \sum_{i'=1}^{I} u^{i'}}. \tag{1.7}$$

The BAM assumes that the no-purchase option is independent of which and how many products are offered in g, i.e., substitution effects are underestimated. Thus, the generalized attraction model (GAM) is introduced.

Generalized attraction model

The GAM captures shadow attraction values w^j for every product j in the set of relevant products N which are not offered in g and could be purchased somewhere else in the market. Then, the no-purchase probability is increased by the sum of shadow attraction values of not offered alternatives. The resulting choice probabilities can be determined by:

$$P^i(g) = \frac{u^i}{u^0 + \sum_{j \in N \setminus g} w^j + \sum_{i'=1}^{I} u^{i'}}. \tag{1.8}$$

Multinomial logit model

The multinomial logit model (MNL) is a random utility model which differ from the previously described in that it is assumed that the utility a customer experiences when purchasing a certain product can typically be decomposed into an observable part and an unobservable part. Thus, if the observable part of the utility a customer experiences for product i is denoted by v^i, and the unobservable part by ε^i, the following holds: $u^i = v^i + \varepsilon^i$ (Train 2009). Thereby, v^i is assumed to be known, i.e., it depends on observable product specifications. The random variable ε^i is assumed to be unknown to the modeler and different random utility models differ regarding the assumptions underpinning the random component's distribution (Talluri and Van Ryzin 2006). The MNL bases on the assumption that the ε^i are independent and identically distributed random variables that follow a Gumble distribution. The resulting choice probabilities can be determined by:

$$P^i(g) = \frac{e^{\frac{u_i}{\mu}}}{\sum_{i' \in g} e^{\frac{u^{i'}}{\mu}}}, \tag{1.9}$$

with μ being a scale parameter of the Gumble distribution.

2. Vehicle routing

Generally, VRPs are a wide range of problems that concern the fulfillment of transportation requests with a given fleet of vehicles at minimum cost or maximum profit, subject to a number of constraints, such as capacity constraints. The first to introduce a VRP were Dantzig and Ramser (1959) who presented a generalization of the NP-hard traveling-salesman problem (TSP) and called it "The truck dispatching problem". Since then, a rich body of literature addressing VRPs, its variants, and related problems emerged. Further, due to the VRP's practical relevance, numerous commercial VRP solvers were developed (Golden et al. (2008), Preface). In the following, a brief introduction to VRPs and its variants is given. It is targeted to provide a general idea rather than to provide a comprehensive in-depth introduction, which would be out of scope for this dissertation. Thus, a discussion of how model and approach VRPs is omitted. Instead, the interested reader is referred to the seminal books of Toth and Vigo (2014) and Golden et al. (2008), the reviews of Pillac et al. (2013), Gendreau et al. (1996), and Cattaruzza et al. (2017), and the dissertation of Ulmer (2017) for a detailed discussion on how to model and approach VRPs and its variants. If not stated differently, the following discussion bases on the seminal book of Toth and Vigo (2014), Chapter 1.

Variants of the VRP and related problems can be classified along different dimensions. In the remainder of this section, the following dimensions are addressed: the evolution of information and uncertainty, environment and network characteristics, types of transportation requests, fleet characteristics, sources of uncertainty, constraints, and objectives.

Evolution of information and uncertainty

Pillac et al. (2013) give a comprehensive review of online tour-planning problems and introduce a corresponding two-dimensional classification scheme. The first dimension is named *information evolution* and classifies VRPs as *static* problems if all required information is known beforehand, and as *dynamic* problems if the information input changes over time. The second dimension is called *information quality* and it separates *deterministic* problems from *stochastic* problems. Thus, in their classification scheme, there are *static and deterministic* problems, *static and stochastic* problems, *dynamic and deterministic* problems as well as *dynamic and stochastic* problems, as shown by Figure 2.1. How the specific classification of a VRP along the here described dimensions substantially influences the classes of solution approaches that can be applied successfully. Chapters 2-4 of Toth and Vigo (2014) provide a review of how to tackle static and deterministic VRPs. For approaches of how to address static and stochastic VRPs, the reader is referred to Florio et al. (2020), Louveaux and Salazar-González (2018), Gauvin et al. (2014) and Lysgaard (2003). For solving dynamic and deterministic VRPs the same approaches as for static and deterministic VRPs are applied, with the difference, that frequent re-planning is conducted. Approaches of how to address dynamic stochastic VRPs are outlined in Section 5.3. Further, the interested reader is referred to Pillac et al. (2013) for a general outline of classes of solution approaches for dynamic VRPs.

		Information quality	
		Deterministic input	Stochastic input
Information evolution	Input known ex ante	*Static and deterministic*	*Static and stochastic*
	Input changes over time	*Dynamic and deterministic*	*Dynamic and stochastic*

Figure 2.1: Classifying VRPs with regard to the evolution of information and uncertainty, cf. Pillac et al. (2013)

Environment and network characteristics

Typically, VRPs address problems which are *node-based*, opposing to *arc-based* problems. This means that, in VRPs, transportation requests originate or terminate in certain points, i.e., nodes, of a transportation network which represent depots and customer locations. On the contrary, in arc-based problems the focus is not to reach certain points but to traverse certain links, i.e., arcs, of a transportation network. The latter represent streets for example. Then, VRPs can be classified according to whether the travel cost between nodes are *symmetric*, i.e., cost are independent of the direction travelled, or *asymmetric*, i.e., cost differ depending on the direction travelled. There are different modeling approaches that are valid for symmetric VRPs and cannot be applied to asymmetric ones and vice versa.

Types of transportation requests

Transportation requests that can be modeled in VRPs can be the distribution of goods from one or multiple depots to customers or, analogously, the collection of goods at customer locations and their delivery to one or multiple depots. The first are also known as *one-to-many VRPs*, the latter as *many-to-one VRPs*. Additionally, pick-up-and-delivery problems can be modeled as VRPs. Those are also known as *many-to-many VRPs*. Further, VRPs can be classified according to whether they address the transportation of goods, or the transportation of people, such as service technicians, medicals, pupils, or passengers.

Fleet characteristics

Another classification dimension regards the fleet involved in a VRP. Those typically differ in size and whether vehicles are homogeneous, i.e., share the same characteristics and capacity, or heterogeneous, i.e., vehicles differ in characteristics and capacity. Further, due to technological developments in recent years, many research articles now address VRPs that include multiple types of transportation vehicles, e.g., the delivery of goods to micro-depots by usual delivery trucks and subsequent fulfillment with electric cargo bikes, delivery robots, or drones (Boysen et al. (2021)).

Sources of uncertainty

In dynamic VRPs, there are various uncertain influences. The most considered are unknown *travel and service times*, unknown *requests*, and unknown *demand*. If there is exploitable information available on the distribution underlying the respective component, the VRP is considered stochastic.

Constraints

VRPs can be further classified according to the constraints considered. Those depend on its practical application and can for example range from *physical vehicle capacity*, *route lengths*, and *driver working hours* to specific *(un-)loading constraints* if for example a forklift is required to move the load. The problems under consideration in this dissertation are typically subject to *time constraints*, e.g., start times of delivery time windows and delivery deadlines. Additionally, it is relevant to consider whether vehicles can conduct *multiple consecutive tours* and whether transportation requests can be served in any arbitrary *delivery order* or if there are order constraints as in pick-up-and-delivery problems.

Objectives

Originally, the first introduced VRP was a *cost minimizing* problem. However, a vast variety of objectives of VRPs emerged in the literature, inspired by the underlying practical applications. There can be single and multiple objectives that range from finding feasible solutions to only selecting the most profitable customers. The most common objectives found in the literature relevant for this dissertation are either *minimizing cost*, *maximizing profit*, *maximizing the service area covered*, *maximizing fairness*, or *maximizing the number of accepted customers*.

Typically, in the literature, VRPs are implicitly classified among the relevant of the previously mentioned dimensions by speaking names such as *stochastic dynamic vehicle routing problem* (Ulmer 2017), *profitable multi-trip vehicle routing problem (p-MTVRP)* (Chbichib et al. 2012), *profitable single trip vehicle routing problem with time windows (p-VRPTW)* (Toth and Vigo 2014).

3. Stochastic dynamic problems

Demand-management problems as described above are stochastic dynamic problems by nature. Further, in i-DMVRPs as considered in this dissertation, also the tour-planning component can be a stochastic dynamic VRPs. Therefore, in this section, a brief introduction to stochastic dynamic problems is given. First, it is outlined how to model such problems. Then, a high-level introduction to approaching stochastic dynamic problems is given.

3.1 Modeling stochastic dynamic problems

Stochastic dynamic problems can be modeled as Markov decision processes (MDPs), which according to Puterman (2014), Chapter 1, are sequential decision making models, in which "the set of available actions, the rewards, and the transition probabilities depend only on the current state and action and not on states occupied and actions chosen in the past". In this section, the components of an MDP model, i.e., *decision epochs*, *state*, *action*, *transition*, and *objective* are briefly outlined. An in-depth discussion on the preliminaries for modeling a sequential decision process as an MDP, its history and characteristics, as well as its optimality criteria is out of scope for this dissertation. The interested reader is referred to the seminal books of Puterman (2014) for a comprehensive introduction to MDPs. If not stated differently, the following descriptions base on Puterman (2014), Chapter 2 with a slight adaption of notation to fit the remainder of this dissertation.

Decision epochs

The *decision epochs* t of an MDP define the points of time in which decisions have to be taken. They can either be continuous or discrete, but only the latter are considered in this dissertation. In this case, the decision epochs define the beginning of the stages of the MDP. Further, decision epochs can be subdivided according to whether they are time-based across (constant) time steps or event-based, e.g., across customer request arrivals. If time can be discretized, the latter equals time-based modeling with non-constant time steps. Further, the set of decision epochs can be infinite, corresponding to an infinite horizon problem, or finite, for problems that either have a natural end or are modeled as such. Then, T marks the last decision epoch of the MDP, i.e., $t \in \{1,2,...,T\}$. In this dissertation, only finite problems are considered.

State

The *state* s_t of an MDP maps the condition of the decision system at a certain decision epoch t and, thus, consists of all information that is known to that decision epoch and relevant for decision making. It can have different dimensions for different types of information. The set of all potential states is referred to as *state space* $\mathscr{S}$. It can be finite or infinite and can further depend on the decision epoch, i.e., vary with varying decision epochs, or be constant.

Action

In an MDP model, the *action* a_t in decision epoch t corresponds to the realization of a certain decision. The set of all potential actions is denoted *action space* $\mathscr{A}$ and can also be finite or infinite as well as dependent or independent of t. Further, it can be dependent on the state s_t, i.e., the set of actions to choose from can vary with varying states.

Rewards

With an action a_t in state s_t, a *reward* $r(s_t, a_t)$ realizes. This reward can be positive, i.e., equal revenue, or negative, i.e., equal cost. It can be dependent on the state $s_{t+1} \mid s_t, a_t$ to which the system transitions through action a_t when being in state s_t, or be independent of it. In the first case, it is modeled via an expectation by applying a *transition probability function* described in the following. However, it is required for optimal decision making that in decision epoch t either $r(s_t, a_t)$ or its expectation $\mathbb{E}(r(s_t, a_t))$ is known. In decision epoch $T+1$ no action is taken but still a reward r_{T+1} can realize. It is called *scrap value* or *salvage value* and can also be understood as a state value $v(s_{T+1})$ or as a function $r(s_{T+1})$ of the terminal state s_{T+1}.

Transition

The *transition* from one state s_t to a successor state s_{t+1} depends on the action taken in t, i.e., on a_t, and can be modeled as a transition probability function $p(s_{t+1} \mid s_t, a_t)$ for the transition to s_{t+1}, or $p_t(s_{t+1} \mid s_t, a_t)$, if it depends on decision epoch t. Usually, $\sum_{s_{t+1} \in \mathscr{S}} p(s_{t+1} \mid s_t, a_t) = 1$ holds.

Objective – Generally, the *objective* of an MDP is to maximize the sum of rewards, positive and negative, accrued over all decision epochs and including the salvage value:

$$\max \sum_{t=1}^{T} r(s_t, a_t) + r(s_{T+1}). \tag{3.1}$$

3.2 Approaching stochastic dynamic problems

It can be shown that the above described objective (3.1) can also be represented by the well-known Bellman equation, already described for the availability control problem (1.1 and the assortment optimization problem (1.5) in Section 1. Unfortunately, as mentioned already, it suffers from the curses of dimensionality (Powell et al. (2012), Chapter 2) such that it is not tractable for realistic sized instances. Thus, a rich variety of approaches to solve stochastic dynamic programs (at least heuristically) emerged from a wide range of different research communities. Powell (2019) summarize ten different research fields that address stochastic dynamic programs who all tackle similar problems with similar solution approaches but from different perspectives and, thus, based on different schemes of notation and taxonomies. The literature on tackling i-DMVRPs mainly emerged from the field of approximate dynamic programming (ADP), and only a few authors adapt notation, perspectives and taxonomies from the reinforcement learning community. However, those apply approaches that are also well known in ADP. Therefore, in the following, a short introduction to the ideas underlying *ADP* is given. It is the target to provide the reader with an idea of the general framework of how to tackle stochastic dynamic programs. A comprehensive introduction to the wide range and manifold challenges of specific solution approaches is out of scope of this dissertation. Instead, the interested reader is referred to the seminal book by Powell et al. (2012) for a comprehensive introduction to ADP, and to the seminal book by Sutton and Barto (2018) for a very accessible introduction to reinforcement learning. Further, for a broad overview of the different communities addressing stochastic dynamic problems the reader is referred to Powell (2019). The following brief introduction to the ideas of ADP bases on Powell et al. (2012), Chapter 4.

From Equation (1.1), it can be observed that the Bellman function is a recursive function that draws on the value v_{t+1} to calculate a value v_t. Thus, solving it to optimality is theoretically achieved by applying backwards recursion, i.e., by starting at the salvage value in T, passing that forward to $T-1$ to calculate v_{T-1}, which is then passed forward to calculate v_{T-2}, and so on. The key difference between this optimal dynamic programming and ADP is that in ADP the value function is calculated by stepping forward through the decision process rather than backward.

Thereby, a successor state's value v_{t+1} is not known for any decision epoch t and is thus *approximated* for making a decision a_t. After a decision is made, a reward r realizes and the system transitions to a successor state. Then, a value v_t is approximated drawing on the corresponding reward $r(s_t, a_t)$ and the approximated v_{t+1}, respectively. Consequently, the approximation of v_t, in the following denoted by $\hat{v}_t$, is based on a *sample realization* of the reward and the successor state.

These steps are replicated for every $t = 1, ..., T$, therewith following a *sample path*, i.e., a certain sequence of stochastic realizations within the decision process. After reaching the terminal decision epoch T, the whole process is repeated multiple times, updating the state value's estimate every time the respective state is visited. Thus, the state value's estimate is improved with every visit, until a termination criterion is reached.

Generally, this procedure is known as *value function approximation (VFA)* and existing approaches differ with regard to how the unknown v_{t+1} is approximated, and how the updates of state values are conducted. It is highly influenced by the underlying problems and applications and, therefore, forms the heart of ADP research. Examples of how the previously outlined, general framework of tackling stochastic dynamic problems can be applied to tackling i-DMVRPs, is the core of the literature discussed in Section 5.

Part III

Integrated demand management and vehicle routing problems (i-DMVRPs)

In Part III of this dissertation, i-DMVRPs are formally defined and the related literature is analyzed. Within this analysis, research gaps are identified, which motivate a deeper analytical investigation of i-DMVRPs as well as of their solution approaches.

Over the last decade, new areas of transportation and home delivery have emerged and have been successfully established with business models such as AHD, SDD, and mobility-on-demand (MOD) (Fleckenstein et al. 2021). Enabled by the ongoing evolution of communication technology, these business models have considerably changed the interaction between providers and customers in LMD services (Agatz et al. 2013). Thereby, customer expectations regarding offered services, delivery speed, and accuracy of shipping notifications have crucially increased (McKinsey and Company 2016). Additionally, the Covid-19-pandemic substantially accelerated the global proliferation and growth of home delivery services as it forced new customer groups to rely on such services who would otherwise have been hesitant to try them (Unnikrishnan and Figliozzi 2020).

Consequently, the fast growth of e-commerce and increasing customer expectations require the providers of LMD services to improve delivery efficiency and consider new measures to enable a profitable business operation. Hence, in practice, innovative transportation modes, such as crowd shipping and the delivery with drones, and *mechanisms to steer customer choice toward efficient* delivery operations, i.e., solving i-DMVRPs, are considered, and in parts successfully applied (Agatz et al. (2013), Agatz et al. (2021), Boysen et al. (2021), Archetti and Bertazzi (2021)).

In parallel, researchers began to address the same issues from a wide range of perspectives. Among engineering communities for instance, a vast body of literature emerged that considers the development and incorporation of delivery robots or drones, and technologies to deposit deliveries in car trunks were developed (Boysen et al. (2021), Chen et al. (2021)). Researchers of the operations research community who deal with tour-planning problems and VRPs shifted their research toward online optimization (Azi et al. (2012), Voccia et al. (2019)). The revenue management community began to explore transferring known revenue management instruments from, e.g., the airline industry to this area of transportation and home delivery (Klein et al. 2020). Overall, a vast body of literature emerged that addresses a wide range of existing business models with an even wider range of approaches to tackle the diversity of the related challenges (Fleckenstein et al. 2021).

As a consequence of different research communities considering these problems, the related literature bases on many different taxonomies, non-uniform classification schemes, and diverse terminology and modeling approaches. Hence, classifying and comparing business models, and selecting appropriate solution approaches for an existing problem among the approaches presented in the literature, is a challenge in itself.

Therefore, in Part III of the dissertation, an abstracted problem description of i-DMVRPs in LMD is provided and the existing literature that addresses demand management and online tour planning in i-DMVRPs is analyzed. Finally, a unified modeling framework to model i-DMVRPs is introduced. The respective contributions of this part are the following:

(1) A clear definition of i-DMVRPs in LMD that bases on a general taxonomy is introduced by abstracting from the literature and current practice. A unified terminology is derived and its components serve as building blocks according to which i-DMVRPs can be classified. Further, different types of i-DMVRPs are delineated and contrasted, which then allows to classify individual problem settings discussed in the literature, accordingly.

(2) A comprehensive literature overview is provided, in which different research streams are identified. Existing solution approaches are analyzed in depth and the results are provided in a comprehensive tabular overview. This supports the selection of an appropriate solution approach for new i-DMVRP applications.

(3) Essential research gaps are identified, which motivate a deeper investigation of i-DMVRPs, especially with regard to modeling, analytical discussion, the interpretation of opportunity cost, and anticipatory solution approaches.

(4) A unified modeling framework for i-DMVRPs is developed. It incorporates anticipatory demand-management and tour-planning decisions and explicitly clarifies

their temporal interdependencies. The proposed modeling approach is generic enough that a variety of problem settings can be modeled, but explicit enough that the reader has a clear understanding of the underlying business process.

This part of the dissertation is organized as follows: in Chapter 4, the addressed i-DMVRPs in LMD applications are defined and the taxonomy is introduced. Throughout the chapter, applications with disjoint booking and service horizons are delineated and contrasted from those with overlapping ones.

In Chapter 5, the related literature is reviewed. First, an overview of existing surveys that address i-DMVRPs either from a demand-management perspective, a LMD perspective, or integratively, is provided. Afterwards, the existing literature proposing anticipatory solution approaches is reviewed. Finally, a comprehensive tabular overview of the discussed solution approaches is provided and essential research gaps are identified.

In Chapter 6, the first of the research gaps identified in Chapter 5 is closed: a unified modeling framework for i-DMVRPs is developed.

4. Description of i-DMVRPs

In this chapter, different types of i-DMVRPs of LMD found in the literature and existing business models are abstracted with the target to derive a general description that follows a clear, specifically introduced terminology. This terminology bases on a decomposition of different influencing factors within the booking and fulfillment processes and serves as building blocks for future classifications. The derived building blocks are: *customer request arrivals, fulfillment options, offer sets, customer choice probabilities*, as well as *fulfillment operations* aspects when serving customer orders. Throughout the chapter, similarities of different types of i-DMVRPs are elaborated and decisive differences are highlighted. Later on, this allows to classify the settings discussed in the literature, i.e., AHD and SDD, accordingly.

First, a short overview of how the above mentioned influencing factors, i.e., the resulting building blocks, integrate into an i-DMVRP business process is given: as described in the introduction, i-DMVRPs comprise two integrated types of decisions, namely *demand-management decisions* and *tour-planning decisions*. Demand-management decisions have to be made for every *customer request arrival* and correspond to decisions on which *fulfillment options* to offer each particular customer at which delivery fees. The combination of a subset of fulfillment options with fixed prices is termed an *offer set* (Fleckenstein et al. 2021). Every offer set yields different customer *choice probabilities* according to which customers choose a fulfillment option. This either turns the customer request into a confirmed customer order, or the customer leaves the system without purchasing anything. In the latter case, the customer request is not considered any further. Finally, all confirmed customer orders have to be served by the provider's *fulfillment operations*.

At this point, different types of i-DMVRPs differ from each other. Figure 4.1 shows

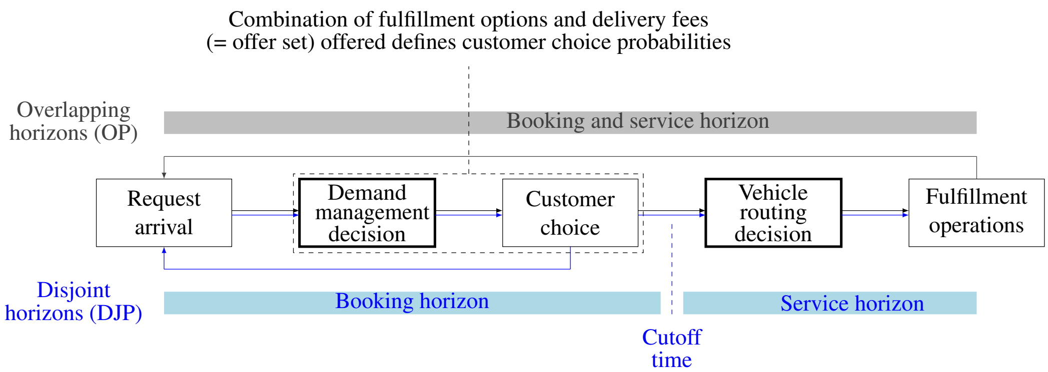


Figure 4.1: Overview i-DMVRP booking and fulfillment process

the corresponding temporal relationships of the previously introduced building blocks of different types of i-DMVRPs. In an *i-DMVRP with disjoint booking and service horizons* (DJP), all customer orders are served *after* a predefined cut-off time (Lang et al. (2021a), Koch and Klein (2020)). In an *i-DMVRP with overlapping booking and service horizons* (OP), the fulfillment operations run *in parallel* to incoming customer requests (Azi et al. (2012), Voccia et al. (2019)). However, in both types of i-DMVRPs tour-planning decisions correspond to solving an underlying VRP. In DJPs this is a static VRP that is only solved once after the booking horizon's end. In OPs, it corresponds to an online and dynamic VRP that has to be solved (Fleckenstein et al. 2021). Thus, in OPs, the provider continuously takes tour-planning decisions and executes them, while the booking horizon is still running. Below, DJPs and OPs are described in more detail along the previously introduced building blocks for deeper classification.

4.1 i-DMVRPs with disjoint booking and service horizons

In LMD, DJPs are better known as AHD problems. They concern business models in which customers request the delivery of perishable goods (e.g., groceries, flowers), personal/addressee-sensitive goods (e.g., pharmaceutical products, confidential documents), or bulky goods (e.g. home appliances or furniture), online or in-store (Campbell and Savelsbergh (2005), Lang et al. (2021a)). Those business models typically have in common that product delivery requires the presence of the customer because the products cannot be left in the mailbox. Since redelivery is costly, provider and customer consequently agree on a certain delivery time window in which the provider promises delivery (Agatz et al. (2008), Lang et al. (2021a), Fleckenstein et al. (2021)). Thereby, the customer and the provider follow diverging targets. The customer prefers a very narrow delivery time window in order to minimize the time in which they have to be at home. The provider, in contrast, prefers wide time windows in order to maximize flexibility in fulfillment operations (Köhler et al. (2019), Köhler et al. (2020)). Further, there are certain time windows, which are more frequently requested by customers than others (e.g., evening slots vs. morning slots) (Asdemir et al. (2009), Agatz et al. (2013)). Thus, the provider integrates demand-management measures into the process in order to steer customer choice behavior towards efficient fulfillment operations (Klein et al. 2019). In the following, the main components of such business models are described.

Customer request arrivals

In the considered problems, customer requests arrive sequentially at random times with either time-dependent or time-independent arrival rates, within a predefined booking horizon until a predefined cut-off time (Asdemir et al. (2009)). The arriving customers log in to the provider's website with registered profiles and fill their shopping basket or place delivery orders in-store. For every incoming customer request, the provider is assumed to know the corresponding location, as well as the shopping basket's potential revenue. The requesting customer expects to be offered a selection of fulfillment options with fixed delivery fees to choose from.

Fulfillment options

In DJPs, *fulfillment options* are typically predefined time windows (Agatz et al. 2008), either overlapping or non-overlapping, with either varying or fixed lengths, in which the provider commits to deliver (Fleckenstein et al. 2021). The set of fulfillment options could, for example, be composed of the following options: delivery the following day between 10am and noon, between 11am and 1pm, and between 10am and 1pm.

Offer sets and customer choice probabilities

For every incoming customer request, the provider decides on a subset of fulfillment options to offer. In doing so, the provider also selects a delivery fee for each fulfillment option, either from a predefined set of discrete price points or from a continuous (potentially limited) range. Thus, the set of all offer sets is either finite, if potential delivery fees originate from a finite set, or infinite, if potential delivery fees originate from a continuous range (Strauss et al. (2018), Klein et al. (2020), Fleckenstein et al. (2021)).

Consistent and logical offer-sets follow three guidelines:

(1) Within an offer set, each fulfillment option appears only once.

(2) If a customer can decide not to make a purchase, a fictive fulfillment option that represents a no-purchase option is included in every offer set. It is priced at zero.

(3) To ensure pricing consistency, the delivery fees of fulfillment options with longer time spans do not exceed the delivery fees of those with shorter spans.

Different offer sets yield different *choice probabilities* for different customers (Train (2009), Chapter 2, Gallego et al. (2019), Chapter 4).

Fulfillment operations

If a customer chooses an option other than the no-purchase option, their request turns into a confirmed *customer order* that needs to be served by the provider as promised. Therefore, every customer order is assigned a delivery time window in which they have to be served. Those time windows can either form hard constraints, i.e., customer orders have to be served within the exact time window, or soft constraints. If they form soft constraints, the customer orders could also be served before the beginning of the time window or after its end, but in acceptance of a corresponding penalty. All delivery time windows fall into a service horizon. It starts after a predefined cut-off time which marks the booking horizon's end (Asdemir et al. (2009), Agatz et al. (2013)). Thus, when the tour-planning decisions are taken, all customer orders, their locations, and their delivery time windows are known.

Generally speaking, in DJPs, tour-planning decisions correspond to solving a VRP. Its exact characteristics depend on the setting of the business model and the corresponding requirements (Agatz et al. (2008), Fleckenstein et al. (2021)). For example, some providers operate a fleet of a fixed number of homogeneous or non-homogeneous vehicles. Others collaborate with a large amount of sub-providers, such that the number of delivery vehicles is inexhaustible. For business models in which groceries are shipped, the number of orders that can be served with one delivery vehicle may not be limited by physical vehicle capacity. Opposingly, if bulky goods are shipped, the physical vehicle capacity may be a decisive factor. Other potential restrictions, among others, can be (Toth and Vigo (2014), Chapter 1):

- limited tour lengths (especially for electric vehicles),
- limited driver working hours,
- required (un-)loading times,
- access restrictions at customer sites,
- specific vehicle requirements for specific orders (e.g., for temperature-sensitive goods that require refrigeration or for bulky goods that require a forklift to be unloaded).

4.2 i-DMVRPs with overlapping booking and service horizons

In LMD, OPs are better known as SDD problems. They concern business models in which customers order the delivery of certain goods on short notice. More precisely, they expect to receive their order the very same day, within a few couple of hours, or even within the same hour. This concerns business models such as courier services, pharmaceutical product delivery, grocery and consumer goods delivery, meal delivery, and heating oil delivery (Fleckenstein et al. 2021). Some business models require the provider to load customer orders in a depot (Voccia et al. (2019), Chen et al. (2019), Ulmer (2020a)), as typical for grocery delivery. In other business models, all goods that can potentially be ordered are stocked in all delivery vehicles. In the latter case, a delivery vehicle does not have to return to the depot in order to load new customer orders but can insert a new customer order into a currently running tour (Ulmer et al. (2018), Ulmer et al. (2019), Ulmer (2020b)). A prominent example is heating oil delivery.

However, in both types of these business models, the provider starts their fulfillment operations immediately with, or shortly after, the first realized customer order (Archetti and Bertazzi 2021) to match the customers' expectations regarding fast delivery. While this brings instant gratification for the customer, it poses the following challenges to the provider, which are essential for profitable fulfillment operations and which can be tackled by integrating demand-management measures into the process (e.g. Klapp et al. (2018), Archetti and Bertazzi (2021)):

(1) Defining the start time of a tour and allocating orders to tours. If a tour starts early or a certain destination is visited early in the service horizon, the tour may not be profitable because there are not enough orders included to cover the resulting cost. If a tour starts late, the corresponding delivery vehicle may not yet be available when a more profitable customer request arrives. Further, it might be profitable to serve already known customer orders later in the service horizon with a different tour if more requests are expected from its vincinity.

(2) Covering peak times without having too many drivers and vehicles idle during low demand times.

In summary, the decisive difference between DJPs and OPs is that in DJPs the fulfillment operations start after the booking horizon's end, while in OPs, they run in parallel to the incoming customer requests. Hence, for OPs, tour-planning optimization has to be run in parallel to the demand management optimization as well. Consequently, while the demand-management related building blocks (customer request arrival, offer set definition, and calculating choice probabilities) structurally remain unchanged, OPs differ from DJPs by the delivery options offered to incoming customer requests and by the corresponding fulfillment operations. In the following, these distinguishing components are described in more detail.

Fulfillment options

In OPs, *fulfillment options* are typically predefined nested time spans in which the provider commits to deliver. The set of fulfillment options could, for example, comprise delivery within the next 90 minutes or within the next 300 minutes. It has to be noted that there is also business models for OPs in which fulfillment options are potentially overlapping time windows of equal or varying lengths as described for DJPs. Those business models for OPs are not discussed any further as they can be reduced to either the described DJPs or the described OPs (Waßmuth et al. 2022).

Fulfillment operations

If a customer chooses an option other than the no-purchase option, their request turns into a *customer order*. A customer order is assigned a *delivery deadline* that is calculated from the request time and the length of the chosen fulfillment option. As described before, in OPs, the service horizon starts with, or shortly after, the first realized customer order and ends when the last customer order of a day has been served. In the following, the fulfillment operations resulting from the two previously introduced different OP business models are delineated.

> *No depot returns required* – No depot return is required to load newly arrived customer orders if all goods that a customer can potentially order are stocked in the delivery vehicles. The provider continuously takes tour-planning decisions. Those correspond to decisions on whether a vehicle should idle at its current location, continue its trip (a pre-planned sequence of customer locations to be served), or which location to visit next.
>
> *Depot returns required* – Depot returns are required to load newly arrived customer orders if the delivery vehicles only stock goods for orders assigned

to the running tour. Thus, a newly arrived order can only be served either by a vehicle that visits a depot after the customer order's request time or by a vehicle that is idle in a depot. Again, the provider continuously takes tour-planning decisions. Those involve decisions on whether and when a vehicle should leave the depot/return to the depot and which orders will be assigned to outgoing vehicles. Once a vehicle left the depot, most providers fully execute the resulting tours as planned, without preemptive depot returns (e.g. Ulmer (2020a)). However, other providers allow running tours to be interrupted for depot returns in order to load new or different customer orders (e.g. Côté et al. (2021)).

Building block	DJPs	OPs
Customer request arrivals	sequentially within a predefined booking horizon, location/revenue are known	
Fulfillment options	predefined overlapping or non-overlapping time windows of varying or fixed length	predefined nested time spans of varying lengths
Offer sets	subset of fulfillment options offered with certain delivery fees	
Customer choice probabilities	probabilities with which a customer chooses a certain fulfillment option when being offered a certain offer set, calculated with a predefined choice model	
Fulfillment operations	start after the booking horizon's end, all customer orders and their location are known when tours are planned, no preemptive depot returns	run in parallel to booking horizon, not all customer orders are known when tours/trips are planned, preemptive depot returns, tours can be revised after their start

Table 4.1: Differences DJPs and OPs

For both types of OPs, tours and trips are either planned in such a way

- that no customer orders will be served later than their delivery deadline (Ulmer 2020a),
- that a predefined service level is matched (Klapp et al. 2018),
- or that penalty cost arising from late deliveries or the required outsourcing to a third-party logistics provider (3PL) are minimized (Voccia et al. 2019).

Travel and service times can thereby be assumed deterministic (Azi et al. (2012), Voccia et al. (2019), Côté et al. (2021)) or stochastic (Prokhorchuk et al. (2019)), depending on the specific problem setting. Moreover, the VRPs underlying specific fulfillment operations can be subject to further operational restrictions as already described for DJPs. Table 4.1 summarizes the differences between DJPs and OPs.

5. Literature on i-DMVRPs in last-mile delivery

Both, traditional demand management applications as described in Chapter 1 and LMD applications as described in Chapter 4, face the challenge of how to match fixed, scarce resources with heterogeneous demand. This similarity has prompted the establishment of vehicle routing as a new application for demand management (Agatz et al. 2013), which is also reflected by the rapid development of the respective literature. While there is some earlier work in the related field of stochastic and dynamic vehicle routing, the works by Campbell and Savelsbergh (2005) and Campbell and Savelsbergh (2006) can be viewed as the first contributions to integrating active demand management and vehicle routing. These publications initiated the literature stream on AHD demand-management problems, i.e., on DJPs in LMD (Yang et al. (2016), Koch and Klein (2020), and Vinsensius et al. (2020)).

On the contrary, Azi et al. (2012) present the first work on steering booking processes in parallel to fulfillment operations. Therewith, they initiate the literature stream on SDD demand-management problems, i.e., addressing OPs in LMD (Prokhorchuk et al. (2019), Ulmer (2020a)).

In this chapter, these two literature streams are examined. First, in Section 5.1, a short overview of the recent literature dealing with i-DMVRPs in LMD on a *general* level is given. This includes literature reviewing related research, elaborating different business concepts, and literature that features high-level modeling or solution frameworks. Afterwards, existing *solution approaches* for operational decision making in specific problem settings of such i-DMVRPs are analyzed. Thereby, the literature that addresses relevant DJPs (see Section 5.2) is distinguished from the literature that addresses respective OPs (see Section 5.3). As described earlier, the overlap of the booking and service horizons yields substantially different challenges since it incorporates an online tour-planning component, which is not required for DJPs. As a consequence, the literature on OPs mainly evolved from the literature on stochastic dynamic vehicle routing which is not the case for the literature on DJPs. The latter mainly evolved from the revenue management literature. However, for both problem settings the discussion is focused on anticipatory solution approaches. Further, the discussed literature is differentiated according to whether it involves *learning-based* anticipation or *non-learning-based* anticipation.

It has to be noted that there are other research streams related to i-DMVRPs in LMDs that consider a variety of research questions that are not subject to further discussion in the remainder of this dissertation. These consider, for example, innovative delivery modes such as crowd-sourced delivery (Dayarian and Savelsbergh 2020), delivery by autonomous vehicles or drones (Ulmer and Streng (2019), Ulmer and Thomas (2018)), or further operational aspects such as customer discrimination by delivery areas (Chen et al. 2020).

5.1 Surveys and general frameworks for i-DMVRPs in last-mile delivery

In this section, the existing literature considering i-DMVRPs in LMD in general is outlined. This literature mainly addresses recent research from two perspectives, either from a demand-management perspective or from an operational fulfillment/vehicle routing perspective. At first, *general surveys* of those main literature streams are outlined. Then, literature that is concerned with general modeling or solution *frameworks* is discussed.

Surveys

Table 5.1 gives an overview of the existing surveys on literature related to i-DMVRPs in LMD that are outlined in the following. It shows whether specific applications, i.e., DJPs, are considered or LMD concepts in general (G). There is no survey that specifically addresses OPs. Further, Table 5.1 summarizes whether the authors focus on a specific component, i.e., the demand management component (DM) or the operational fulfillment/VRP component (VRP), or consider i-DMVRPs. Additionally, Table 5.1 gives an overview of whether the respective survey addresses business concepts, mathematical models, and/or solution approaches. In the following, at first, surveys that consider i-DMVRPs from the demand-management perspective are outlined. Then, the respective literature considering i-DMVRPs from the operational fulfillment/VRP perspective. At last, literature that considers both perspectives in an integrative manner, i.e., literature that deals with i-DMVRPs comprehensively is reviewed.

Authors	Application	Perspective	Concepts	Models	Approaches
Agatz et al. (2008)	G	DM	✓	✓	✗
Agatz et al. (2013)	DJP	DM	✓	✗	✓
Archetti and Bertazzi (2021)	G	VRP	✓	✗	✓
Boysen et al. (2021)	G	VRP	✓	✗	✗
Fleckenstein et al. (2021)	G	i-DMVRPs	✓	✓	✓
Klein et al. (2020)	G	DM	✓	✓	✓
Snoeck et al. (2020)	DJP	VRP	✓	✗	✗
Soeffker et al. (2021)	G	VRP	✓	✓	✓
Waßmuth et al. (2022)	G	i-DMVRPs	✓	✗	✓

Table 5.1: Surveys that feature related problems

Demand-management perspective – Agatz et al. (2008) provide the first review on LMD concepts, more precisely on the distributional challenges in e-fulfillment, including initial ideas to connect demand management and LMD. The authors name two features of e-fulfillment systems that enable demand management. The first is an increased pricing flexibility compared to stationary retail, for example, and the second is an extensive availability of data concerning purchasing behavior. In those two features, the authors see the foundation for segment-specific pricing as well as promotion and conclude that they see "a shift from reactive forecasting to a much more active demand management in e-fulfillment". In a later review, Agatz et al. (2013) compare the demand-management-related processes of a large e-grocer with those prevalent in airline revenue management and elaborate similarities as well as decisive features of both concepts. As a result, they provide starting points for incorporating differentiated slotting/pricing, or dynamic slotting/pricing into business concepts of AHD and thereby focus on the demand-management side of i-DMVRPs. The same holds for Klein et al. (2020) who review recent generalizations and advances of revenue management techniques in traditional applications and new industry applications. They show how to transfer availability control to AHD problem settings and present the corresponding DP formulation. Further, they outline how the reviewed publications incorporate fulfillment cost and how the involved opportunity cost are approximated.

Operational fulfillment/VRP perspective – Archetti and Bertazzi (2021) consider i-DMVRPs from the other perspective, i.e., with a focus on operational fulfillment/VRP aspects. They review recent advancement and challenges of LMD systems especially in relation to e-commerce. They see pricing as a measure to balance demand among favorable and unfavorable delivery time windows and conclude: "We believe that the immediate challenge in time window assignment and management is related to defining proper pricing policies to influence customer requirements and favor a more balanced distribution of required delivery slots." They do not further elaborate demand-management measures in particular. The same is true for the survey by Snoeck et al. (2020). Although the authors specifically address revenue management in LMD with a focus on AHD problem settings, they do not discuss demand-management aspects but focus on the influences of potential extensions and future developments on the fulfillment operation level, such as incorporating a flexible crowdsourced fleet, or accounting for collection and delivery points. Further, Boysen et al. (2021) survey LMD research with a focus on newly emerged business concepts and Soeffker et al. (2021) discuss the related stochastic dynamic VRPs and embed them into a prescriptive analytics framework. Both consider pricing as an essential decision dimension in existing business concepts that concern LMD and stochastic dynamic vehicle routing respectively.

Integrated perspective – The most recent and comprehensive survey of literature on i-DMVRPs that integrates both perspectives, i.e., the demand-management and the last-mile-delivery perspective, is the survey by Fleckenstein et al. (2021). The authors provide a generalized problem definition and outline applications from LMD, i.e., AHD and SDD, as well as from the MOD sector. They propose a high-level, generic MDP modeling formulation and outline typically involved customer choice models. Further, they provide a comprehensive survey of general solution concepts and describe solution approaches for all involved subproblems, i.e., demand management-related subproblems and tour-planning-related subproblems. They conclude that the research area of i-DMVRPs can benefit from the development and proliferation of common model formulations and standardized solution approaches. In the following, recent research contributions in this directions are described in chronological order. Another survey that considers i-DMVRPs is a survey by Waßmuth et al. (2022). They survey recent literature dealing with i-DMVRPs on the strategic, tactical and operational level and specifically differentiate between two demand-management levers that are *offering* and *pricing*. Although they mention modeling aspects of the reviewed literature in their tabular overviews, they do not comprehensively discuss different modeling approaches and rather focus on brief outlines of solution approaches.

i-DMVRP frameworks

Ulmer et al. (2020) propose a generalized framework for modeling stochastic dynamic VRPs, which they call *route-based MDPs*. This framework can also be applied to model VRPs with stochastic customer requests. The authors aim at closing a gap between modeling stochastic dynamic VRPs and existing solution methods. They propose to include (preliminary) route plans into the state space. Further, they define immediate rewards for a certain action as the resulting myopic marginal changes in the value of those route plans. In parallel, Fleckenstein et al. (2021) propose a similar idea in their generalized modeling framework, namely, to include the vehicle/fulfillment state as part of the state definition. However, due to the specific objective function formulation Ulmer et al. (2020) propose, their MDP model framework is not intuitively transferable to i-DMVRPs. Presumably, this is the reason why many authors (see for example Klapp et al. (2020), Chen et al. (2019), Fleckenstein et al. (2021)) adopt their proposed state definition but not their proposed objective function when modeling i-DMVRPs.

Lang and Cleophas (2020) present a simulation-based benchmarking framework for prescriptive analytics based solution approaches tackling AHD optimization problems, denoted by SiLFul. It is an open source platform that provides problem settings and benchmark solution approaches to researchers. Therewith, the authors strive for more comparability and reproducibility of research results.

Hildebrandt et al. (2021) summarize solution frameworks for solving stochastic dynamic VRPs that originate from different research streams, namely *computer science* and *operations research*. They propose a high-level concept on how to combine those frameworks to build a reinforcement learning-based solution framework.

5.2 Solution approaches for i-DMVRPs with disjoint booking and service horizons

In this section, solution approaches for DJPs in LMD, i.e., for AHD problem settings, are discussed. Although tour-planning decisions in AHD problem settings are static ones that arise after the booking horizon has ended, involving a dynamic tour-planning component in the decision process during the booking horizon is valuable for two reasons: first, in order to check the feasibility of demand-management decisions, and second, to approximate future expected revenue and/or final fulfillment cost in order to evaluate the effect of demand-management decisions (Fleckenstein et al. 2021). The respective tour-planning component can be *myopic*, such that solely the already accepted customer orders are considered (Campbell and Savelsbergh (2006)), or *anticipatory*, such that future expected customer requests are taken into consideration as well. In the following, only anticipatory approaches are discussed as it has been shown that those outperform myopic ones (Fleckenstein et al. 2021). The elaboration of these approaches is structured as follows: at first, research that involves *learning-based anticipation* (see Section 5.2.1) is discussed. Afterwards, research that involves *non-learning-based anticipation* (see Section 5.2.2) is discussed.

Beyond the following discussion, the existing approaches can further be subdivided according to their individual objectives, e.g., maximization of the number of accepted customer requests, revenue or profit, or minimization of cost. They can also be subdivided by considering whether anticipation involves explicit tour planning and whether it is used to anticipate both, revenue and cost, or solely one of both. Further, the approaches can be subdivided by distinguishing different demand-management types such as assortment optimization for availability control or dynamic pricing, and/or by the use of different customer choice models. For a comprehensive and detailed classification on those differentiation dimensions, the interested reader is referred to the recent survey of Fleckenstein et al. (2021).

5.2.1 Anticipatory learning-based approaches

Learning-based approaches aim to learn accurate value function approximations and thus opportunity cost approximations either offline or online. Anticipatory learning-based approaches for optimizing demand-management decisions in AHD problem settings can be further subdivided according to whether they apply *look-up tables* or base on *parametric* or *non-parametric VFA*. In the following, the relevant literature is discussed in this order.

Lang et al. (2021b) compute a *look-up table* to solve an assortment optimization problem and determine which fulfillment options to offer to an incoming customer request by applying a backward dynamic programming approach. During the booking horizon, the look-up table returns values for different customer request characteristics. These are then linearly combined to an overall value as a basis for decision making. Therewith, the authors analyze the effect of different influencing factors and their weights in the objective function, e.g., revenue or coverage of the service area.

Ulmer and Thomas (2020) also apply a *look-up table*-based approach, which they combine with a *parametric VFA*. The authors aim at maximizing the overall revenue in an AHD problem setting by optimizing accept/reject decisions for every incoming customer request.

A different group of research papers addresses AHD demand-management problems more sophisticatedly by setting (continuous) prices or incentives dynamically for every incoming customer request. Thereby, Yang and Strauss (2017), Vinsensius et al. (2020), Koch and Klein (2020) and Lebedev et al. (2020) rely on pure *parametric VFAs* to calculate opportunity cost. In the first three of these, the authors implement linear regression models. Yang and Strauss (2017) and Vinsensius et al. (2020) base it on parameters that can be derived from the state of the system without calculating tentative route plans, i.e., without solving a VRP across existing and potentially anticipated customer orders . Koch and Klein (2020) require such tentative route plans to calculate the parameters' values before estimating opportunity cost. However, for all of these approaches, the resulting estimates are the input for a subsequent pricing problem, which in Yang and Strauss (2017) and Koch and Klein (2020) is solved with a continuous pricing algorithm and in Vinsensius et al. (2020) with an incentive-setting quadratic program as proposed by Campbell and Savelsbergh (2006). Lebedev et al. (2020) propose non-linear VFA models for profit-maximizing continuous pricing in an AHD problem setting and conduct a comprehensive sensitivity analysis on the trained parameters.

Finally, a third group of research papers examines the incorporation of neural networks for VFAs. Although generally relying on parameters, VFAs by neural networks is discussed to be *non-parametric* in some research papers (Fleckenstein et al. (2021)) or as a hybrid between parametric and non-parametric (Soeffker et al. (2021)). Therefore, here they are treated as a special group: Dumouchelle et al. (2021) rely on learned neural network models to incorporate them in a SARSA algorithm that directly returns accept/reject decisions. Lang et al. (2021a) train a neural network model to approximate cost and then solve an assortment optimization problem that bases on an MNL customer choice model.

5.2.2 Anticipatory non-learning-based approaches

The literature of non-learning-based solution approaches for demand-management problems in an AHD setting can be subdivided according to whether the anticipation within the tour-planning component is *sampling-based* or *seed-based*. In the following, *sampling-based* refers to solution approaches that sample future customer requests and incorporate those into tentative tour planning. *Seed-based* refers to solution approaches that agglomerate expected demand in pre-defined or dynamically adjusted locations as a basis for tentative tour planning.

Campbell and Savelsbergh (2005) are among the first to integrate demand management and vehicle routing. They consider a profit-maximizing AHD setting in which the provider decides on the acceptance or rejection of incoming customer requests. They rely on a *sampling-based* insertion heuristic that is performed from scratch for every decision epoch. It considers already accepted customer orders as well as sampled customer requests in order to approximate the monetary effects of accepting a customer request with regard to displaced expected revenue and resulting expected fulfillment cost. Yang et al. (2016) use the same insertion heuristic but amend it by additionally considering tentative tour plans from previous decision epochs. From the resulting tour plans, they derive an approximation of the expected increase of fulfillment cost for all potential fulfillment options and use this as an input to a continuous dynamic pricing approach that bases on an MNL customer choice model.

Angelelli et al. (2021) also draw on the idea to compare tentative tour plans including and not including the current customer request for a certain fulfillment option. The tentative tour plans are created with a team orienteering problem (TOP)-based tour-planning heuristic and involve accepted customer orders and expected customer requests. Then, the authors derive an approximation of opportunity cost, which is used to optimize accept/reject decisions in an one-vehicle AHD problem setting with regard to profit maximization.

Another *sampling-based* approach is the one proposed by Strauss et al. (2021). They approximate fulfillment cost as a linear function of the number of orders (accepted and expected) according to an approach introduced by Daganzo (1987). Thereby, the authors account for displaced demand and variable fulfillment cost. Generally, Strauss et al. (2021) consider a profit-maximizing dynamic pricing problem, which they tackle with a mixed integer program (MIP)-based approach that also relies on an MNL choice model as the approach by Yang et al. (2016).

The existing approaches that use *seed-based* anticipation within the tour-planning component incorporate it in deterministic linear programming approaches to solve the involved demand-management problem. Klein et al. (2018) use an MIP-based approximation of opportunity cost that involves seed-based tour planning. They incorporate the resulting estimate into a dynamic pricing problem. It is modeled like the MDP model of Yang et al. (2016). To solve the pricing problem, they solve an assortment optimization problem over dynamically re-optimized combinations of fulfillment options and price points. Mackert (2019) applies the same tour-planning approximation and uses the resulting opportunity cost estimate as input to a linearized assortment optimization problem arising under the assumption of a generalized attraction choice model.

Most recently, Giallombardo et al. (2020) agglomerate expected requests into seeds to incorporate them explicitly into an MIP-based solution approach. This optimizes tour-planning decisions and defines booking limits as a basis for solving the involved demand-management problem. Here, this approach is considered dynamic as the authors suggest to recompute the MIP's solution during the booking horizon in order to derive reoptimized booking limits that account for the current state of the system.

5.3 Solution approaches for i-DMVRPs with overlapping booking and service horizons

In this section, solution approaches for OPs in LMD are discussed. Due to the overlap of the booking and the service horizons, the respective fulfillment operations run in parallel to incoming customer orders, which, in turn, are dynamically incorporated into already running fulfillment operations. Thus, it is a stochastic dynamic problem. Consequently, only those approaches are discussed in the following, that are designed to take stochastic dynamic tour-planning into consideration.

The elaboration of the respective literature is structured as follows: as in Section 5.2, research that involves learning-based anticipation (see Section 5.3.1) is discussed first. Then, research that involves non-learning-based anticipation is addressed (see Section 5.3.2). The online tour-planning component decisively distinguishes OPs from DJPs with regard to the complexity that has to be accounted for by solution approaches, and thus, how solution approaches developed over time. To account for that, in each of the following Sections 5.3.1 and 5.3.2, a brief, high-level discussion of the most related research on pure online tour-planning approaches is provided in order to show how the discussed integrated approaches evolve from them. Thereby, pure online tour-planning approaches and integrated solution approaches for stochastic dynamic VRPs are distinguished following Fleckenstein et al. (2021). Accordingly, integrated approaches exceed basic feasibility control by (at least) allowing feasible customer requests to be rejected if the expected contribution to the objective is negative.

5.3.1 Anticipatory learning-based approaches

In this section, learning-based solution approaches for stochastic dynamic VRPs with stochastic requests are elaborated.

Pure online tour-planning

To solve stochastic dynamic VRPs with unknown requests, VFA approaches can be applied to derive tour-planning decisions. Recent publications on solving stochastic dynamic VRPs with stochastic customer requests as considered in typical i-DMVRPs are Ulmer (2017), Chapters 6-12, Ulmer et al. (2018), and Ulmer (2019). These publications present a variety of VFA approaches to take tour-planning decisions that will match as many customer requests as possible. In these approaches, a VFA is learned offline over a large number of simulation runs. The learned VFAs are then applied to assess post-decision values in an online decision period in order to take good tour-planning decisions.

Integrated approaches

Ulmer et al. (2019) combine an offline VFA with a simulation-based online rollout algorithm to solve a dynamic VRP with stochastic service requests for a single vehicle. They learn a look-up table offline, which is generated by approximate value iteration. It approximates state values based on temporal state information. The granularity of the look-up table adjusts dynamically during the learning process. For online decision making, this look-up table is combined with a simulation-based online rollout algorithm considering spatial information of potential post-decision states. Soeffker et al. (2017) also rely on such a dynamic look-up table and combine it with an approximate value iteration algorithm. Their overall target is to evaluate the effect of dynamic demand management on customer discrimination and to introduce a measure of fairness to be maximized. This measure aims on equalizing the chances of being accepted among the customer requests originating from different parts of the service area. However, look-up table based approaches cannot be implemented efficiently for realistic sized i-DMVRP applications. Considering multiple vehicles with multiple tours and offering multiple delivery options to incoming customer requests results in very large state and action spaces. Thus, even if a state space aggregation is applied, a respective look-up table is of intractable size.

In a different set of publications, researchers consider a pricing component within their integrated approaches. Ulmer (2020a) solves a dynamic tour-planning and pricing problem for an SDD problem setting by developing an anticipatory pricing and routing policy that is based on a sophisticated VFA approach and upstream policy learning. He is the first to present a VFA approach for a fleet of vehicles, which he does by separating the value function with regard to different vehicles. He includes the tour-plans of the vehicles in the state definition as introduced in Ulmer et al. (2020). To solve the pricing problem, the author relies on an opportunity cost estimate for different delivery options from comparing approximated state values. If the respective opportunity cost estimates are low, the corresponding delivery options are offered for *budget prices* derived from the upstream policy learning. Those prices represent the typical base prices of the delivery options. Only in cases where the approximated opportunity cost exceeds the budget price, the corresponding delivery options are priced differently. Then, the prices are set to equal the opportunity cost estimate. Therefore, this procedure ensures that only requests with a non-negative contribution to the overall objective are accepted. Tour-planning decisions are determined by a simple, non-anticipatory insertion heuristic.

In the same set of publications, Prokhorchuk et al. (2019) introduce a stochastic dynamic pricing and routing problem for an SDD problem setting with stochastic travel times. They also base decision making on approximating opportunity cost and amend the approach of Ulmer (2020a) by stochastic travel times, a different tour-planning heuristic that accounts for stochastic travel times, and by using standard VFA procedures.

A relatively new set of publications addresses integrated approaches by relying on offline reinforcement learning. More precisely, the authors apply deep-Q-learning algorithms to derive respective demand-management decisions. Chen et al. (2019) maximize the number of accepted customer requests in a setting where SDD fulfillment is conducted with delivery vehicles and drones. In Chen et al. (2020), fulfillment is only conducted by delivery vehicles, but the problem's setting is amended by the aspect of fairness within accepting/rejecting customer requests for SDD as in Soeffker et al. (2017).

5.3.2 Anticipatory non-learning-based approaches

In this section, non-learning-based solution approaches for stochastic dynamic VRPs with stochastic requests are discussed.

Pure online tour planning

Bent and Van Hentenryck (2004) introduce a multiple scenario approach to optimize tour-planning decisions in dynamic VRPs with stochastic customer requests. They aim at maximizing the number of accepted customer requests by constantly generating multiple tour plans based on sampled customer requests. From those tour plans, a distinguished tour plan is chosen by a so-called *consensus function* and frequently updated. It serves as input for taking decisions on which customers to serve next and by which vehicle. In Bent and Van Hentenryck (2007), the authors enhance the previous approach by including waiting and relocating strategies. With this approach, not only tour-planning decisions, e.g., a vehicle's next destination, but also dispatching decisions, i.e., which orders to allocate to which tours, are taken.

Integrated approaches

Among the considered integrated, non-learning-based approaches, the most relevant group of papers is based on the idea of the multiple scenario approach by Bent and Van Hentenryck (2004), as described above:

Azi et al. (2012) introduce an initial demand-management approach to a dynamic VRP with stochastic requests. They consider a profit maximization problem in determining which requests to accept and which to reject. To solve the tour-planning problem, they apply an adaptive large neighborhood search to scenarios that, like the ones in Bent and Van Hentenryck (2004), include already accepted customer orders and sampled customer requests. What is new about their approach is that they then compare scenario solutions with and without the current customer request and define the difference in solution quality as a scenario-specific opportunity value. If the sum of all scenario-specific opportunity values is positive, they accept the request. This approach delivers an estimate of whether or not the acceptance of a customer request yields a positive contribution to the overall objective, taking potential future developments into account.

Voccia et al. (2019) also adapt the ideas from Bent and Van Hentenryck (2004) and Bent and Van Hentenryck (2007). They aim at maximizing the number of feasibly inserted customer requests for a stochastic dynamic VRP with time windows and stochastic requests. The customer requests that are not inserted in a feasible solution are outsourced to a 3PL, which comes with a penalty cost per outsourced order. Their approach yields comprehensive tour-planning decisions including the set of orders allocated, vehicle assignment, as well as a schedule for each tour. Like Bent and Van Hentenryck (2007), they consider future, not yet realized customer requests by applying a sample-scenario approach. Thereby, they solve a multi-trip team orienteering problem with a standard implementation of a variable neighborhood search. Afterwards, the scenario solutions are used to construct anticipatory tour plans. Compared to Bent and Van Hentenryck (2007), they apply an enhanced consensus function that chooses partial plans according to their appearance frequency in the scenario solutions. Also, they include waiting strategies to improve the anticipatory quality of their solutions.

Côté et al. (2021) build on the approach by Voccia et al. (2019) and amend it by a regret heuristic, a different consensus function, and a specifically tailored branch-and-regret method. Further, they also consider settings in which pre-emptive depot returns are allowed.

Finally, there is another research stream dealing with OPs considering integrated approaches. Klapp et al. (2018) introduce a dynamic dispatch waves problem for an SDD problem setting in which decisions on which customer orders can be allocated to tours are taken while still unknown customer requests from known potential locations occur. They introduce an MIP based on the rolling-horizon solution approach, discuss policies calculated beforehand (called *a priori* policies), and enhance these policies by dynamically rolling them out during the joint booking and service horizon. With this approach, the authors aim at minimizing penalty costs for not-served orders. In Klapp et al. (2020), the authors extend this approach by introducing instant accept/reject decisions. Thereby, they explicitly model a dynamic programming model resembling the state definition in Ulmer et al. (2020), but maintaining tour plans that include anticipated customer requests in the state definition. Here, they take accept/reject decisions according to expected penalty costs derived from different policies. Both approaches aim at solving SDD problem settings with known customer locations served by a single vehicle only. Thus, the approaches cannot be transferred to realistic sized OPs in LMD without greater effort.

5.4 Tabular overview and identification of research gaps

In this section, the findings from analyzing the literature addressing i-DMVRPs in LMD are summarized and conclusions regarding essential research gaps are derived.

Tabular overview

Table 5.2 summarizes the literature that addresses anticipatory solution approaches for DJPs and OPs in LMD as described in Sections 5.2 and 5.3, as well as literature that deals with i-DMVRPs analytically. The second column shows for which applications, i.e., DJPs or OPs, a solution approach is designed. In the next two columns, it is indicated whether an approach involves anticipatory demand management (DM) (✓) and/or tour planning (TP) (✓) or not (✗). The fourth column shows whether the addressed anticipation is analytical, learning-based (✓), or non-learning-based (✗). The fifth column summarizes the objectives addressed by an approach. The observed objectives are the maximization of revenue (rev), profit (profit), customer request acceptances (accept), fairness (fair), coverage of the service area (serv), the minimization of cost (cost), or a (hierarchical) combination of those objectives. Approaches that aim at minimizing the number of rejected customer requests are counted as those maximizing customer request acceptances. The last column shows whether opportunity cost are considered explicitly and, if so, whether they are considered comprehensively, accounting for displaced acceptances and variable fulfillment cost (✓), or whether the displacement of expected revenue (DPC) or variable fulfillment cost (MCTS) are considered only. In the following, first, conclusions regarding the relevant literature on DJPs are summarized. Then, conclusions regarding the relevant literature on OPs are summarized.

Conclusion regarding the relevant literature on DJPs

Concerning the literature on DJPs in LMD, it becomes clear that solution ap-

[1] consider penalty cost but no fulfillment cost

Authors	Appli-cation	Anticipatory DM	Anticipatory TP	Learning based	Objective	OC
Asdemir et al. (2009)	DJP	✓	✗	analytical	rev	DPC
Lebedev et al. (2021)	DJP	✓	✗	analytical	prof	✓
Dumouchelle et al. (2021)	DJP	✓	✗	✓	profit	-
Koch and Klein (2020)	DJP	✓	✗	✓	profit	✓
Lang et al. (2021a)	DJP	✓	✗	✓	rev	DPC
Lang et al. (2021b)	DJP	✓	✗	✓	rev	DPC
Lebedev et al. (2020)	DJP	✓	✗	✓	profit	-
Ulmer and Thomas (2020)	DJP	✓	✗	✓	rev	-
Yang and Strauss (2017)	DJP	✓	✗	✓	profit	✓
Vinsensius et al. (2020)	DJP	✓	✗	✓	cost	MCTS
Angelelli et al. (2021)	DJP	✓	✗	✗	profit	✓
Campbell and Savelsbergh (2005)	DJP	✓	✗	✗	profit	✓
Giallombardo et al. (2020)	DJP	✓	✗	✗	profit	-
Klein et al. (2018)	DJP	✓	✗	✗	profit	✓
Mackert (2019)	DJP	✓	✗	✗	profit	✓
Strauss et al. (2021)	DJP	✓	✗	✗	profit	✓
Yang et al. (2016)	DJP	✓	✗	✗	profit	MCTS
Chen et al. (2019)	OP	✗	✓	✓	accept	-
Chen et al. (2020)	OP	✗	✓	✓	accept&fair	-
Ulmer (2020b)	OP	✗	✓	✓	accept	-
Ulmer et al. (2018)	OP	✗	✓	✓	accept	-
Ulmer et al. (2019)	OP	✗	✓	✓	accept	-
Azi et al. (2012)	OP	✗	✓	✗	profit	-
Côté et al. (2021)	OP	✗	✓	✗	accept&cost	-
Klapp et al. (2018)	OP	✗	✓	✗	cost&serv	-
Klapp et al. (2020)	OP	✗	✓	✗	profit	-
Voccia et al. (2019)	OP	✗	✓	✗	cost	-
Prokhorchuk et al. (2019)	OP	✓	✗	✓	rev&cost[1]	DPC
Soeffker et al. (2017)	OP	✓	✗	✓	accept&fair	-
Ulmer (2020a)	OP	✓	✗	✓	rev	DPC

Table 5.2: Anticipatory solution approaches for i-DMVRPs in LMD

proaches exist for a wide range of problem settings and that the underlying demand-management and tour-planning problems are tackled with a variety of different, learning-based and non-learning-based procedures. By definition, in DJPs, the tour-planning component is a static one, such that anticipation is only required in solving the demand-management subproblem. Most approaches thereby aim at optimizing the overall profit and follow the decomposed approach described in Section 1.3, for which a valid approximation of opportunity cost is essential.

However, only a limited number of works contains attempts to formally define and analyze opportunity cost in the context of i-DMVRPs. Furthermore, existing definitions are not consistent in terms of generality and scope. Some works provide definitions in the traditional sense. For example, the definitions by Campbell and Savelsbergh (2005) and Lang et al. (2021a) only incorporate the cost of displacements of future customers. On the contrary, some authors suggest that there is another component of opportunity cost in i-DMVRPs to consider besides displacement cost. For example, according to Klein et al. (2018), opportunity cost quantifies the "[. . .] 'consequences' concerning potential future requests and the resulting routing cost [. . .]". Further, Yang et al. (2016) and Koch and Klein (2020) state that the lost profits of potential future orders as well as final delivery costs have to be anticipated when approximating opportunity cost. Vinsensius et al. (2020) use the term 'opportunity cost' in the traditional sense, i.e., referring only to displacement cost, and introduce the broader term 'marginal fulfillment cost' for the sum of displacement cost and additional delivery cost, but then only consider the latter in their optimization approach. The most extensive, but still not unified definitions of opportunity cost for i-DMVRPs is provided by Yang and Strauss (2017), Mackert (2019), and Strauss et al. (2021). All explicitly highlight that opportunity cost comprises two components, one capturing displacement of future revenue and one capturing variable fulfillment cost.

Conclusion regarding the relevant literature on OPs

Concerning the literature on OPs in LMD, anticipation can be applied to either the demand-management problem, the tour-planning problem, or both. From Table 5.2 it can be conducted that there is no approach that combines anticipatory demand management with anticipatory tour planning. There are only approaches that address only one in an anticipatory manner and the other one myopically. Among the approaches that address anticipatory tour planning of which demand-management decisions are an implicit result, there exist learning-based and non-learning-based approaches. Among the approaches that apply explicit anticipatory demand management but base that on myopic tour planning, there are only learning-based approaches. Generally, most approaches do not aim at maximizing the overall profit. Instead, the existing approaches rather focus on maximizing the number of acceptances or revenue, neglecting fulfillment cost, or focus on cost-optimization, neglecting displacement of revenue.

Further research gaps arise regarding the consideration of opportunity cost. There is no approach that tackles the demand-management problem by explicitly considering an opportunity cost estimate that considers displacement of revenue and changes of fulfillment cost. Only the approaches by Ulmer (2020a) and Prokhorchuk et al. (2019) explicitly consider opportunity cost. They define it as a value function difference without further elaborating on its structure and then derive an opportunity cost estimate as a basis for decision making that only incorporates the displacement of expected revenue. Additionally, to best of the author's knowledge there exists no research tackling OPs in LMD analytically.

Resulting research gaps

From the findings from Section 5.1 and the previous conclusions, the following research gaps are identified:

- There is no explicit but unified MDP model for i-DMVRPs with disjoint and overlapping booking and service horizons.
- There is no MDP model for i-DMVRPs, neither for DJPs nor for OPs, that explicitly accounts for the mutual integration of demand-management and tour-planning decisions.
- Although it has been identified in the literature that the concept of opportunity cost for i-DMVRPs differs from the original understanding, there is no unified definition of the same. Further, for OPs, the potential of optimizing demand-management decisions based on opportunity cost estimates has not been exploited at all.
- In the related literature, there is no analytical discussion on i-DMVRP models that also address OPs.
- While there is a wide range of solution approaches that tackle DJPs, the existing solution approaches for OPs are not sufficiently holistic and evolved to integrate anticipation in demand management and tour planning at the same time. Additionally, literature proposing approaches that aim at profit

optimization and thereby consider revenue and cost at the same time, is scarce for OPs.

In the following, those research gaps are addressed as follows: first, in Chapter 6, a unified but explicit MDP modeling framework for both discussed types of i-DMVRPs, i.e., DJPs and OPs, is introduced. This modeling framework explicitly accounts for the integration of demand-management and tour-planning decisions. It resembles the approach from Ulmer et al. (2020), i.e., to incorporate (tentative) tour plans in the state definition, and the approach from Fleckenstein et al. (2021) to keep the objective function generic and incorporate both types of decisions. However, opposed to Fleckenstein et al. (2021) and Ulmer et al. (2020), the model introduced in Chapter 6 explicitly formalizes the two types of decisions, i.e., demand-management and tour-planning decisions and further emphasizes their mutual interdependency in a respective value function formulation.

Second, in Chapters 7 to 8 (Part IV) of this dissertation, the traditional concept of opportunity cost is amended to account for displacement of revenue and variable fulfillment cost at the same time. It is a comprehensive concept that is valid for DJPs and OPs. Then, four central properties of the newly derived concept of opportunity cost are analytically discussed and proven, and implications for i-DMVRP solution approaches are derived. Opposed to the discussion of Asdemir et al. (2009), the newly derived concept accounts for displaced revenue and variable fulfillment cost and, contrary to the discussion of Lebedev et al. (2021), the insights from Part IV are valid for DJPs and OPs. Moreover, Lebedev et al. (2021) only derive insights regarding the monotonicity of optimal prices but do not explicitly address the characteristics of the value function itself. This research gap is also closed in Part IV.

Third, in Chapters 9 to 11 (Part V) of this dissertation, a new solution approach for an SDD problem setting is presented that accounts for anticipation in the demand-management component and in the tour-planning component at the same time. It is the first solution approach for OPs that considers anticipation in both components. It incorporates non-learning-based anticipation and involves a newly developed demand-management decomposition to approximate opportunity cost with the comprehensive objective to maximize the overall profit of an SDD provider. Thus, it closes the last of the previously mentioned research gaps.

Table 5.3 classifies Part IV and V according to the literature classification scheme in Table 5.2.

	Appli-cation	Anticipatory DM	Anticipatory TP	Learning based	Objective	OC
Part IV	DJP & OP	✓	✓	analytical	p	✓
Part V	OP	✓	✓	✗	p	✓

Table 5.3: Classification of Parts IV and V with regard to the analyzed literature

6. Modeling i-DMVRPs

In this chapter, MDP models for DJPs and OPs are presented. Among all the coexisting modeling approaches in the literature, this work aims to provide a generic model that is valid for the majority of DJP and OP business models, but is explicit enough for the reader to immediately get an idea of the underlying process. Further, the presented OP model is a generalization of the DJP model such that one modeling approach covers both types of i-DMVRPs. Additionally, although not part of the MDP model itself, the corresponding value functions are presented as optimality equations/solution approaches (Powell et al. (2012), Chapter 3, Puterman (2014), Chapter 4.

First, some general notation of the previously discussed problem components is introduced. Then, an MDP model for DJPs is presented. Afterwards, two generalizations for OPs are presented. The first one is the most intuitive modeling approach for OPs, i.e., the corresponding value function strictly maps the underlying business process. Therefore it is referred to by *natural model*. However, as shown later in Section 8.4, important opportunity cost properties do not hold for this modeling approach. This is why the second modeling approach is presented, for which these properties hold. It is a modification of the first, and thus, in the following it is referred to by *modified model*. Then, the equivalency of the two models is proven, showing that, by model transformation, it is possible to exploit all opportunity cost properties discussed later.

6.1 General notation

In this section, the general notation for the i-DMVRP components described in Chapter 4 is introduced. At the end of this section, a tabular overview of the introduced notation is provided in Table 6.2.

Customer request arrivals

An incoming *customer request* and the *customer order* it potentially turns into is denoted by c. It is further described by its *location* $(x,y)_c$, its *revenue* r_c, and its *arrival rate* $\lambda_c(t)$. Without loss of generality, it is assumed that multiple customer requests can arrive from the same location with the same revenue, such that $\lambda_c(t)$ is independent of whether a customer request c has already realized before or not. The underlying assumption is that a request c only defines location, revenue and choice behavior, instead of referring to a particular person. Thus, c can refer to different persons from the same location, i.e., from the same house, street, or block, with the same socio-economic characteristics. To differ such requests from each other, every incoming request is assigned its *request time* t_c^{req} which further plays an important role in OPs. The *set of all potential customer requests*, more precisely, the set of all potential combinations of location, revenue and choice parameters, is denoted by C.

Fulfillment options

As described in Chapter 4, DJPs and OPs differ in the fulfillment options involved. The fulfillment options of DJPs are delivery time windows with a start time in the future. The fulfillment options of OPs are delivery time spans that immediately start after a customer request's conversion to a customer order.

> *DJPs* – Fulfillment options of DJPs are referred to by indices in ascending order according to their *start time*. The corresponding index set of fulfillment options is denoted by $\mathscr{I}$ with elements $0,1,...,I$ if there are I fulfillment options plus a fictive no-purchase option $i=0$. The start time of the fulfillment option with index i, denoted by $t^{beg}(i)$, is earlier than the start time of fulfillment option i', denoted by $t^{beg}(i')$, for $i<i'$. The *length* of a time window is denoted by $l(i)$. When a customer request c turns into a customer order by choosing a fulfillment option $i \neq 0$, it is assigned a corresponding *delivery start time*, $t_c^{beg}=t^{beg}(i)$, as well as a corresponding *delivery deadline*

$t_c^{due} = t_c^{beg} + l(i)$.

OPs – Fulfillment options of OPs are also referred to by indices in ascending order but according to their length. Thus, the length of the fulfillment option with index i, denoted by $l(i)$, is shorter than the length of fulfillment option i', denoted by $l(i')$, for $i < i'$. When a customer request c turns into a customer order by choosing a fulfillment option $i \neq 0$, its delivery time window starts instantaneously, $t_c^{beg} = t_c^{req}$ and its delivery deadline is set to $t_c^{due} = t_c^{req} + l(i)$.

Offer sets and customer choice probabilities

Offer sets comprise a (sub-)set of available fulfillment options with assigned delivery fees. The delivery fee that is assigned to delivery option i within a certain offer set is denoted by r^i. Among different offer sets, customers experience different utilities for the offered fulfillment options. This results in varying *choice probabilities* for a fulfillment option i depending on the offer set g, which are denoted by $P^i(g)$. Table 6.1 illustratively shows a set of offer sets $\mathcal{G}$ with $|\mathcal{G}| = 9$. Every $g \in \mathcal{G}$ is depicted in a row with exemplary purchase probabilities. In the example, there are two different fulfillment options $\{1,2\}$ with $l(1) < l(2)$, and two potential prices r^{i1}, r^{i2} each, with $r^{11} > r^{12} > r^{21} > r^{22}$. The no-purchase probability for a certain offer set g is denoted by $P^0(g)$ and equals $1 - \sum_{i=1}^{I} P^i(g)$.

	prices of fulfillment options		choice probabilities			
g	$i = 1$	$i = 2$	P^0	P^1	P^2	Σ
1	r^{12}	r^{22}	0.2	0.4	0.4	1
2	r^{12}	r^{21}	0.3	0.4	0.3	1
3	r^{12}	not offered	0.4	0.6	0.0	1
4	r^{11}	r^{22}	0.3	0.2	0.5	1
5	r^{11}	r^{21}	0.4	0.2	0.4	1
6	r^{11}	not offered	0.6	0.4	0.0	1
7	not offered	r^{22}	0.3	0.0	0.7	1
8	not offered	r^{21}	0.5	0.0	0.5	1
9	not offered	not offered	1	0.0	0.0	1

Table 6.1: Exemplary offer sets

Fulfillment operations

In the considered i-DMVRPs, during the service horizon, V homogeneous or non-homogeneous *delivery vehicles* serve the customer orders outgoing from one or more depots. The set of all delivery vehicles is denoted by $\mathscr{V}$. For DJPs and OPs that require depot returns, when a vehicle leaves a depot to serve customers, a *tour* is planned. A tour is denoted by θ^v for a vehicle $v \in \mathscr{V}$ and is defined by a start time t^{start} and a set of loaded customer orders $L = \{c_1, c_2, c_3, ...\}$. Further, to store the order in which a given tour will reach customer locations, a set of tuples is introduced, which assigns positions χ_{c_i} to customer orders $c_i \in L$. This set is denoted by $X = \{(c_1, \chi_{c_1}), (c_2, \chi_{c_2}), (c_3, \chi_{c_3}), ...\}$. Hence, $\theta^v = (t^{start}, L, X)$. Accordingly, the fields of the tuple of a given tour are referred to by $t^{start}(\theta^v), L(\theta^v)$ and $X(\theta^v)$.

Building block	Notation	Description
Customer request arrivals	c	customer request/customer order
	$(x,y)_c$	location of customer request c
	r_c	revenue of customer request c
	$\lambda_c(t)$	(time-dependent) arrival rate of customer request c
	t_c^{req}	request time of customer request c
	C	set of all potential customer requests
	t_c^{beg}	delivery start time of customer order c
	t_c^{due}	delivery deadline of customer order c
Fulfillment options	i	index referring to a certain fulfillment option
	$t^{beg}(i)$	beginning of fulfillment option with index i
	$l(i)$	length of fulfillment option with index i
	I	number of different potential fulfillment options
	$\mathscr{I} = 0, ..., I$	index set to refer to fulfillment options
Offer sets	$\mathscr{G}$	set of all potential offer sets
	g	certain offer set
	$r^i(g)$	delivery fee of delivery option i in offer set g
	$P^i(g)$	probability that delivery option i is chosen when offer set g is presented
	$P^0(g)$	no-purchase probability when offer set g is presented
Fulfillment operations	V	number of delivery vehicles
	$\mathscr{V}$	set of delivery vehicles
	v	index referring to a certain delivery vehicle
	θ^v	tour of vehicle v
	$t^{start}(\theta^v)$	start time of tour θ^v
	$L(\theta^v)$	set of customer orders loaded to tour θ^v
	$X(\theta^v)$	set of positions of customer orders in tour θ^v

Table 6.2: General notation for i-DVMRP modeling

For OPs that do not require depot returns, tours do not have to be planned entirely as the decision which customer orders to serve with which vehicle can be revised continuously. Still, it is common to at least plan partial tours, also referred to by *trips*. Regarding the notation, trips are treated as tours with the difference, that they do not start in a depot. Instead, the customer order on the first position defines the start location.

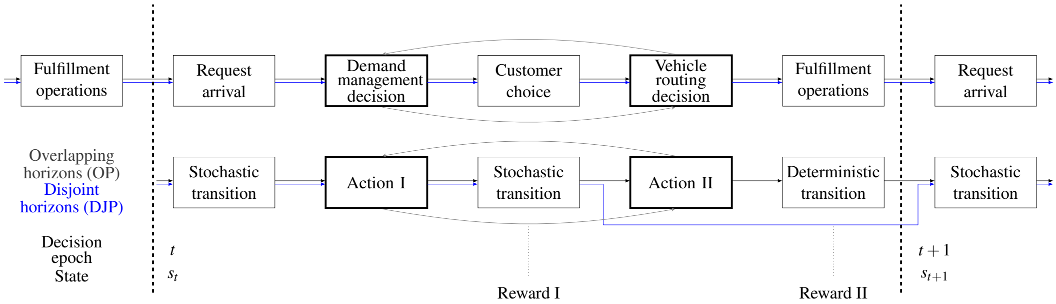


Figure 6.1: Overview of the MDP model of the i-DMVRP booking and fulfillment process

Generally, the MDP of an i-DMVRPs works as follows: in every *decision epoch* with a customer request arrival, the provider has to take an *action*, under consideration of the current *state* of the system. More precisely, to determine the feasibility of potential actions, the provider evaluates the customer orders already confirmed though not yet being delivered. For OPs, the provider additionally evaluates the current state of the delivery operations. The action then taken yields a *transition* as well as a *reward* in that a customer chooses a fulfillment option from the set of offered fulfillment options (including the no-purchase option). In OPs, this triggers the execution of the corresponding tour-planning decision. According to the transition, rewards follow: if a customer chooses to buy, they pay a delivery fee and the shopping basket's revenue realizes. Further, tour cost realize for every vehicle movement which happen in parallel to the booking process for OPs and after the booking process' end for DJPs. The system transitions to the next state, which differs from the previous one, potentially by the newly accepted customer order. In OPs it further differs by new tours/trips and the delivery execution's progress up to the next decision epoch. The objective that the provider seeks in their decision making is to maximize the total contribution margin accrued over all decision epochs. Figure 6.1

shows how these introduced MDP model components integrate into the booking and fulfillment process of DJPs and OPs.

In the following Sections 6.2 and 6.3, decision epoch, state, action, transition, rewards, objective, and additionally the corresponding value functions are described in detail.

6.2 MDP model for i-DMVRPs with disjoint booking and service horizons

In the following, the MDP model components of DJPs are described.

Decision epochs

In the considered problems, the stages of the MDP, i.e., the *decision epochs* are either modeled over (constant) time steps, or over potentially incoming customer requests. The latter equals MDP modeling with non-constant time steps (Puterman (2014), Chapter 2). However, time plays an important role in i-DMVRPs: requesting customers are offered delivery time windows or delivery time spans which have to be matched when solving the underlying tour-planning problem such that travel and service times have to be calculated. Thus, in this dissertation, an MDP model over constant time steps, denoted by $t = 1, ..., T$, is chosen. Without loss of generality and as standard in demand-management literature, these time steps are assumed to be sufficiently small that no more than one customer request arrives per decision epoch.

State

The state s_t of the beginning of decision epoch t stores information about all confirmed but not yet being served customer orders. The set of those orders is denoted by $\mathscr{C}_t$. For every order $c \in \mathscr{C}_t$, it contains all available information stored in a quadruple: $((x,y)_c, t_c^{req}, t_c^{beg}, t_c^{due})$. All possible combinations of customer requests $c \in C$, with all possible arrival times and potentially chosen fulfillment options define the state space $\mathscr{S}$, with $s_t \in \mathscr{S}$.

Action

In the DJPs under consideration, if a customer request arrives, action a_t corresponds to a demand-management decision. It corresponds to selecting which offer set to present to the current customer request, denoted by c_t. The offer set presented at decision epoch t is then denoted by g_t. If no customer request arrives, no action is taken. Thus, the action is formally defined as:

$$a_t = g_t \qquad \text{for } t = 1, ..., T, \text{ if there is a customer request } c_t \tag{6.1}$$

The corresponding action space depends on the current state of the system and the current customer request and is denoted by $\mathscr{A}(s_t, c_t)$. For DJPs, it corresponds to the offer sets which only contain feasible fulfillment options given state s_t and the current customer request c_t. This set of feasible offer sets is denoted by $\mathscr{G}(s_t, c_t)$. Thus, the action space of the MDP model for DJPs is defined as $\mathscr{A}(s_t, c_t) = \mathscr{G}(s_t, c_t)$ with $a_t \in \mathscr{A}(s_t, c_t)$. (It has to be noted that, when solving DJPs, theoretically, no feasibility check has to be made during the booking process. As will be seen later, in theory, the optimal solution is found by solving the recursive value function. Thereby, infeasibilities are implicitly considered in the boundary condition. Consequently, an infeasible VRP and the corresponding costs are passed through to the state where they arise and, thus, a corresponding decision is not taken. Consequently, the action space always equals the set of all offer sets for every state, i.e., $\forall s_t \in \mathscr{S}, c_t \in \boldsymbol{C} : \mathscr{G}(s_t, c_t) = \mathscr{G}$. However, in practice, this optimal solution is intractable such that heuristic solution approaches are applied. Then, feasibility checks can be useful for every demand-management decision and the action space reduces from all offer sets $\mathscr{G}$ to all feasible offer sets $\mathscr{G}(s_t, c_t)$.)

Transition

As Figure 6.1 shows, the transition model of DJPs only comprises (stochastic) demand-management-related transitions. Figure 6.2 is a schematic representation of such transitions involved in a DJP. It shows the temporal relation between two consecutive states with their respective transitions. Thereby, the dashed lines represent stochastic outcomes and the solid lines represent deterministic outcomes.

As Figure 6.2 shows, the stochastic event of whether there is a new customer request arriving or not can be observed at the beginning of a decision epoch t with state s_t. If there is a customer request c_t and a demand-management decision is made, a transition, namely the customer choice i', follows. This is depicted in the upper stream of Figure 6.2. The transition is stochastic, and potential outcomes i' can be observed with known probabilities $P^{i'}(g_t)$. This stochastic transition defines whether the state, more precisely, the state component $\mathscr{C}_t$, alters from one state s_t to a successor state s_{t+1} by adding a new customer order c_t. If no customer request is observed in decision epoch t, no decision needs to be made and no respective transition follows. This is depicted in the lower stream of Figure 6.2. The transitions of a DJP can be formalized as follows:

$$\mathscr{C}_{t+1} = \begin{cases} \mathscr{C}_t, & \text{if there is no customer request observed in } t, \\ & \text{or if the incoming request does not turn into} \\ & \text{a customer order with probability } P^0(g_t) \\ (\mathscr{C}_t \cup \{c_t\}), & \text{if there is a customer request in } t \text{ and it} \\ & \text{turns into a customer order with probability} \\ & \sum_{i=1}^{I} P^i(g_t) \end{cases} \tag{6.2}$$

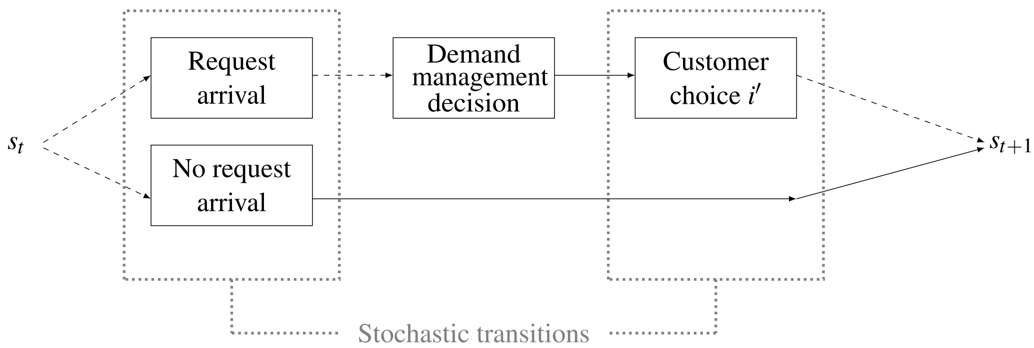


Figure 6.2: Schematic representation of the transitions in a DJP

Rewards

In DJPs, rewards are attributed to demand-management decisions, or more precisely, to the (stochastic) customer choice that results from a demand-management decision. With probability $\sum_{i=1}^{I} P^i(g_t)$ the customer request turns into a customer order and the resulting reward is positive. It is composed of the revenue r_{c_t} of the shopping basket of customer request c_t plus the delivery fee $r^i(g_t)$ resulting from a respective customer choice for a fulfillment option i, determined by the presented offer set g_t.

Objective

The objective of the considered DJPs is maximizing the contribution margin across all decision epochs. This equals maximizing the sum of the revenues of all confirmed customer orders minus the final fulfillment cost of the corresponding fulfillment operations that realize after the booking horizon's end in decision epoch $T+1$. The latter is referred to by logistics related rewards at decision epoch $T+1$, denoted by r^l_{T+1} Thus, the objective can be formalized as:

$$\max \sum_{t=1}^{T} (r_{c_t} + r^{i'}(g_t)) + r^l_{T+1} \tag{6.3}$$

Value function

The above described objective of DJPs can be represented by its value function that equals the well-known Bellman equation. As the name suggests, it captures the value of being in a given state and can be applied to find an optimal policy for an MDP model (Powell et al. (2012), Chapter 3). The value function for the described model of DJPs is defined as:

$$\begin{aligned} V(s_t) = \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \max_{g_t \in \mathscr{G}(s_t, c_t)} \sum_{i \in g_t} P^i(g_t) \cdot \left[r^i(g_t) + r^i_{c_t} + V(s_{t+1} \mid s_t, c_t, i) \right] \\ + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot V(s_{t+1} \mid s_t), \end{aligned} \tag{6.4}$$

with boundary condition:

$$V(s_{T+1}) = r^l_{T+1}. \tag{6.5}$$

The first line of equation (6.4) reflects the value and decision making in decision epoch t when a customer requests c_t arrives. The second line reflects the respective value when no customer request arrives. Thus, for a certain arriving customer request c_t, the provider derives a corresponding demand-management decision by solving:

$$\max_{g_t \in \mathscr{G}(s_t, c_t)} (\cdot). \tag{6.6}$$

Equation (6.5) defines the *salvage value* (Puterman (2014), Chapter 2) and equals the negative cost of the final tour-planning problem r^l_{T+1}. Thus, it equals the negative of the optimal objective value of a (capacitated) VRP defined on a complete graph with nodes for all customer orders in $\mathscr{C}_{T+1}$. This tour-planning problem is referred to by $VRP_{\mathscr{C}_{T+1}}$. The respective objective value is further referred to by the cost function $C(VRP_{\mathscr{C}_{T+1}})$. In case there is no feasible solution to the final tour-planning problem, the corresponding cost is set to infinity. The boundary condition can hence be defined more precisely as:

$$V(s_{T+1}) = r^l_{T+1} = \begin{cases} -\infty & \text{if } VRP_{\mathscr{C}_{T+1}} \text{ has no feasible solution} \\ -C(VRP_{\mathscr{C}_{T+1}}) & \text{else} \end{cases} \tag{6.7}$$

In the following, an MIP for a basic VRP of the fulfillment operations of a DJP is stated. For the sake of generality, it is kept as simple as possible. It therefore does not take into account any capacity or time constraints other than the typical delivery time window constraints but can be generalized accordingly if needed. Delivery time windows are assumed to be hard constraints. Consequently, if it is not possible to serve all customer orders within their individual delivery time windows, there is no feasible solution to the VRP. Further, it is assumed that a fleet of V homogeneous vehicles operates the fulfillment execution with one tour each, starting and ending at a single depot. Thereby, the time needed travelling from customer c to customer c' is denoted by $\tau_{cc'}$ and without loss of generality it is assumed that it contains the service time of serving customer c'. The related travel cost are represented by $\zeta_{cc'}$. A fictive node denoted by c_0 represents the depot. M denotes a sufficiently large number. For ease of readability, in the following model, index $T+1$ of $\mathscr{C}_{T+1}$ is omitted.

The model includes two types of decision variables. A binary one, denoted by $x_{cc'}^{v}$, defining the order in which vehicle v visits customer orders. A positive real valued one, denoted by $a_{c'}^{v}$, defining the time vehicle v serves customer c':

$$x_{cc'}^{v} = \begin{cases} 1 & \text{if customer } c' \text{ is served directly after customer } c \text{ by vehicle } v \\ 0 & \text{else} \end{cases} \qquad \forall c, c' \in \mathscr{C} \cup \{c_0\}, v \in \mathscr{V}$$

$$a_{c'}^{v} \in \mathbb{R}^{+} \qquad \forall c' \in \mathscr{C} \cup \{c_0\}, v \in \mathscr{V}$$

The VRP to determine the salvage value of the MDP model of DJPs can be formulated as follows (Toth and Vigo (2014), Chapter 1) and is briefly described below:

$$C(VRP_{\mathscr{C}}) = \min \sum_{v\in\mathscr{V}} \sum_{c\in\mathscr{C}\cup\{c_0\}} \sum_{c'\in\mathscr{C}\cup\{c_0\}} x_{cc'}^{v} \cdot \zeta_{cc'} \tag{6.8}$$

s.t.

$$\sum_{v\in\mathscr{V}} \sum_{c\in\mathscr{C}\cup\{c_0\}} x_{cc'}^{v} = 1 \qquad \forall c' \in \mathscr{C} \tag{6.9}$$

$$t_{c'}^{beg} \leq a_{c'}^{v} + (1 - \sum_{c\in\mathscr{C}\cup\{c_0\}} x_{cc'}^{v}) \cdot M \qquad \forall v \in \mathscr{V}, c' \in \mathscr{C} \tag{6.10}$$

$$a_{c'}^{v} \leq t_{c'}^{due} + (1 - \sum_{c\in\mathscr{C}\cup\{c_0\}} x_{cc'}^{v}) \cdot M \qquad \forall v \in \mathscr{V}, c' \in \mathscr{C} \tag{6.11}$$

$$a_{c}^{v} + x_{cc'}^{v} \cdot \tau_{cc'} \leq a_{c'}^{v} + (1 - x_{cc'}^{v}) \cdot M \qquad \forall v \in \mathscr{V}, c \in \mathscr{C} \cup \{c_0\}, c' \in \mathscr{C} \tag{6.12}$$

$$\sum_{c\in\mathscr{C}\cup\{c_0\}} x_{cc'}^{v} = \sum_{c\in\mathscr{C}\cup\{c_0\}} x_{cc'}^{v} \qquad \forall v \in \mathscr{V}, c' \in \mathscr{C} \tag{6.13}$$

$$\sum_{c'\in\mathscr{C}\cup\{c_0\}} x_{c_0c'}^{v} \leq 1 \qquad \forall v \in \mathscr{V} \tag{6.14}$$

$$\sum_{c'\in\mathscr{C}} \sum_{v\in\mathscr{V}} x_{c_0c'}^{v} \leq V \tag{6.15}$$

$$\sum_{c'\in\mathscr{C}} x_{c_0c'}^{v+1} \leq \sum_{c'\in\mathscr{C}} x_{c_0c'}^{v} \qquad \forall v \in \mathscr{V} \backslash \{V\} \tag{6.16}$$

The objective function (6.8) minimizes the overall travel cost. Constraints (6.9) ensure that all customer orders are being served. Constraints (6.10)-(6.12) are time restrictions. They ensure that customer orders are served within their allotted time window. Constraints (6.13) ensure flow conservation. Constraints (6.14) and (6.15) ensure that the number of available vehicles is not exceeded. Constraints (6.16) are symmetry breaking constraints.

6.3 MDP model for i-DMVRPs with overlapping booking and service horizons

In this section, the MDP model of DJPs introduced in Section 6.2 is generalized to an MDP model for OPs. In order for this section to be self-contained, there are some duplications with Section 6.2. However, the respective descriptions are kept short. Again, the description is accompanied by a schematic representation of the underlying stochastic dynamic decision process.

6.3.1 Natural Model

Decision epochs

As in the MDP model for DJPs, decision epochs are modeled over constant time steps $t = 1, ..., T$.

State

Due to the overlap of the booking and the service horizon in OPs and contrary to the DJP model, two state components are required in the MDP model of OPs. The first component is the set of confirmed and not yet being delivered customer orders as in the MDP model for DJPs. It is also denoted by $\mathscr{C}_t$. The second component is the overall tour plan at decision epoch t, denoted by ϕ_t (see modelling of route-based MDPs in Ulmer et al. (2020)). It contains the currently running tours θ_t^v for every vehicle $v \in \mathscr{V}$ as described in Chapter 6.1. If the vehicle v is idle in the depot it holds that $\theta_t^v = ()$. If the vehicle v is idle in a customer location it equals a tour that only consists of this customer order. Thus, the state is defined as:

$$s_t = (\mathscr{C}_t, \phi_t) \tag{6.17}$$

The corresponding state space needs to be adjusted as well, such that all possible combinations of customer requests from the registered customer pool, with all possible arrival and due times and with all possible tour plans define the state space $\mathscr{S}$, with $s_t \in \mathscr{S}$.

Action

While in DJPs actions are only taken when there is a customer request c_t, for OPs, actions have to be taken in any decision epoch. Thereby, actions in decision epochs in which a customer request arrives are differentiated from those in decision epochs in which no customer request arrives. In the former case, two types of decisions have to be made integratively, namely demand-management and tour-planning decisions. In the latter case, only tour-planning decisions have to be made:

> *Customer request* – If in a decision epoch t a customer request c_t arrives, a demand-management decision has to be made by selecting which offer set $g \in \mathscr{G}$ to present the requesting customer, as in the DJP model. Again, the offer set presented at decision epoch t is denoted by g_t. In OPs, tour-planning decisions have to be made additionally. More precisely, for every delivery option $i \in g_t$, potential tour-planning decisions are required. Such a tour-planning decision defines the subsequent state's possible overall tour plan which depends on the yet unknown customer choice for a delivery option i. Therefore, ϕ_t^i are introduced for $i \in g_t$ as the tour plans that will be executed if the customer chooses delivery option i. These tour plans are then included in the action definition as a second component.
>
> *No customer request* – If in a decision epoch t no customer request arrives, the corresponding action only comprises tour-planning decisions ϕ_{t+1}^0 without a new customer request.

Accordingly, the action a_t of decision epoch t has two distinct cases:

$$a_t = \begin{cases} (g_t, (\phi^i_{t+1})_{i \in g_t}) & \text{if there is a customer request at } t \\ \phi^0_{t+1} & \text{else} \end{cases} \tag{6.18}$$

Correspondingly, the action space when being in state s_t and observing customer request c_t or observing no customer request, denoted by $\mathscr{A}(s_t, c_t)$, or $\mathscr{A}(s_t)$ respectively, is also defined for the above-mentioned two distinct cases:

Customer request – For the first case, if there is a customer request in t, the action space comprises two components. The first component equals the action space of the MDP model for DJPs. It is denoted by $\mathscr{G}(s_t, c_t)$ and contains all feasible offer sets given state s_t and customer c_t. (It has to be noted that, unlike for DJPs, for OPs a feasibility check is required to that point. Tour-planning decisions that result form a demand-management decisions and the corresponding customer choice are executed during the decision process. Thus, infeasible tour-planning decisions cannot be defined and hence also demand-management decisions that necessarily yield infeasible tour-planning decisions cannot be defined either.) The other component, denoted by $(\Phi^i_{t+1}(s_t, c_t))_{i \in \mathscr{G}(s_t, c_t)}$, defines all potential tour plans that are feasible given the current state s_t, and assuming that the current customer request c_t turns into a customer order with a deadline according to delivery option $i \in \mathscr{G}(s_t, c_t)$. This could also comprise the decision that no new tour will start. Then, the tour plan does not change.

No customer request – If there is no customer request in t, the action space accordingly comprises all feasible tour plans for the set of confirmed customers $\mathscr{C}_t$. In this case, the set of all potential tour plans is denoted by $\Phi^0_{t+1}(s_t)$.

Consequently,

$$\mathscr{A}(s_t, c_t) = \begin{cases} \mathscr{G}(s_t, c_t), (\Phi^i_{t+1}(s_t, c_t))_{i \in \mathscr{G}(s_t, c_t)} & \text{if there is a customer request } c_t \\ \Phi^0_{t+1}(s_t) & \text{else.} \end{cases} \tag{6.19}$$

It has to be noted that the different types of OPs presented in Chapter 4.2 only differ in their corresponding action spaces. For example, for OPs that do not require depot returns, all tours can constantly be revised by inserting newly arrived customer orders. For OPs that require depot returns, the tour-planning component of the action space comprises all tours currently running at t and potentially new tours. New tours can be planned for vehicles that are standing idle at the depot, i.e., vehicles for which $\theta_t^v = ()$ holds, or vehicles that are returning to the depot during the decision epoch. If running tours are to be revised, a preemptive depot return has to be scheduled first.

Transition

The transition model of OPs comprises demand-management-related and tour-planning-related transitions, contrary to DJPs in which only demand management-related transitions have to be considered. Figure 6.3 is a schematic representation of the transitions occuring in OPs, with dashed lines representing stochastic outcomes and solid lines representing deterministic outcomes. Again, at the beginning of a decision epoch t with state s_t the stochastic event of whether there is a new customer request c_t arriving or not can be observed. The resulting transitions differ accordingly.

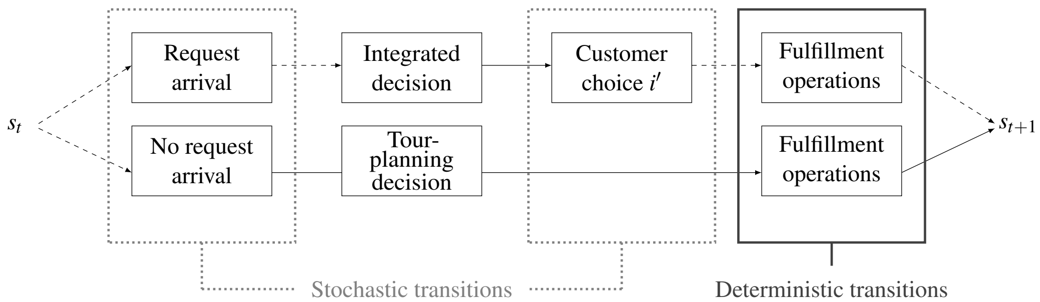


Figure 6.3: Schematic representation of the transitions in an OP

Customer request – If there is a request c_t, integrated demand-management and tour-planning decisions are made and a transition, namely the customer choice i' follows. As in DJPs, this is depicted in the upper stream of Figure 6.3. This transition is stochastic, and potential outcomes can be observed with known probability $P^{i'}(g_t)$. Again, it defines whether the first state component, namely the set of confirmed, yet still to be served customer orders $\mathscr{C}_t$, alters from one state s_t to a successor state s_{t+1} by adding a new customer order. Following this, in OPs, another transition realizes before the next state S_{t+1}. This transition concerns the execution of deliveries. It strictly follows the tour-planning decision $\phi^{i'}_{t+1}$ in a_t. Thereby, $i' = 0$ represents the case that the customer of customer request c_t has rejected all offered delivery options other than the fictive no-purchase option. If deterministic travel times are assumed as depicted in Figure 6.3, this transition is purely deterministic. Thus, in state s_{t+1}, ϕ_{t+1} is set to the respective $\phi^{i'}_{t+1}$ of action a_t. If travel times are stochastic, according probability distributions have to be considered when modeling the corresponding transitions with respect to delivery times. However, whether stochastic or deterministic, fulfillment execution-related transitions also influence the first state component. For example, for OPs that require depot returns, all customer orders from set $\mathscr{C}_t$ that are newly loaded onto a vehicle according to the new tour-plan ϕ_{t+1} are removed from $\mathscr{C}_t$. Therefore, the set $\Psi(\phi_{t+1} \mid \phi_t)$ is introduced, which contains all such customer orders. For OPs that do not require depot returns, $\Psi(\phi_{t+1} \mid \phi_t)$ is the set of customer orders that are newly allocated to tours and removed from $\mathscr{C}_t$.

No customer request – If no customer request is observed in state s_t, only tour-planning decisions are made. The corresponding deterministic or stochastic transition of the delivery execution alters the system from state s_t to the successor state s_{t+1}. This is depicted in the lower stream of Figure 6.3.

The transitions of the state components can be formalized as follows:

$$\phi_{t+1} \quad = \quad \phi^{i'}_{t+1} \tag{6.20}$$

$$\mathscr{C}_{t+1} = \begin{cases} \mathscr{C}_t \setminus \Psi(\phi_{t+1} \mid \phi_t), & \text{if there is no customer request } c_t\text{, or if the incoming request } c_t \text{ does not turn into a customer order with probability } P^0_{g_t} \\ (\mathscr{C}_t \cup \{c_t\}) \setminus \Psi(\phi_{t+1} \mid \phi_t), & \text{if there is a customer request } c_t \text{ that turns into a customer order with probability } \sum_{i=1}^{I} P^i_{g_t} \end{cases} \tag{6.21}$$

Rewards

In OPs, some rewards are attributed to demand-management decisions, others to tour-planning decisions. As in DJPs, the reward accrued with a demand-management decision is positive with probability $\sum_{i=1}^{I} P^i(g_t)$. It is composed of the shopping basket revenue r_{c_t} and the delivery fee $r^i(g_t)$ resulting from a respective customer choice for a fulfillment option i, determined by the presented offer set g_t. The reward accrued with a tour-planning decision is either zero or negative. Such rewards are called logistics-related rewards of a transition from s_t to s_{t+1}, given the decision $\phi^{i'}_{t+1}$ for the chosen delivery option i'. They are formally denoted by $r^l_{\phi^{i'}_{t+1}}$ and equal the sum of fulfillment cost that newly arises with a decision $\phi^{i'}_{t+1}$ and a respective customer choice i'. In OPs that do not require depot returns such cost arises, for example, from starting a trip to a certain customer location. In OPs that require depot returns such cost arises, for example, from starting new tours in $t+1$.

Objective

The objective of solving the OPs under consideration is also to maximize the overall profit. It can be formalized as follows:

$$\max \sum_{t=1}^{T} (r_{c_t} + r^{i'}(g_t) + r^l_{\phi^{i'}_{t+1}}) \tag{6.22}$$

Value function

The above described objective (6.22) of OPs can be represented by the corresponding value function that equals the well-known Bellman equation. It captures the value of being in a given state and can be applied to find an optimal policy for the MDP model (Powell et al. 2012). Specified to the described model of OPs, the value function explicitly models the temporal mutual interdependencies of the two integrated decisions, i.e., the demand-management decision and the tour-planning decision:

$$\begin{aligned} V(s_t) = \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \max_{g_t \in \mathscr{G}(s_t, c_t)} & \\ \Bigg(\sum_{i \in g_t} P^i(g_t) \cdot \big[r^i(g_t) + r^i_{c_t} + & \max_{\phi^i_{t+1} \in \Phi^i_{t+1}(s_t, c_t)_{i \in \mathscr{G}(s_t, c_t)}} \big(r^l_{\phi^i_{t+1}} + V(s_{t+1} \mid s_t, \phi^i_{t+1}) \big) \big] \Bigg) \\ + \big(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\big) \cdot & \max_{\phi^0_{t+1} \in \Phi^0_{t+1}(s_t)} \big(r^l_{\phi^0_{t+1}} + V(s_{t+1} \mid s_t, \phi^0_{t+1}) \big) \end{aligned} \tag{6.23}$$

With boundary condition:

$$V(s_{T+1}) = 0 \tag{6.24}$$

The first two lines of equation (6.23) reflect the value and decision making in decision epoch t, for when a customer request c_t arrives. The third line reflects the corresponding value and decision making for when no customer request arrives. For a certain arriving customer request c_t, the provider derives a corresponding demand-management decision by solving

$$\max_{g_t \in \mathscr{G}(s_t, c_t)} (\cdot). \tag{6.25}$$

To do so, the provider needs to consider the value of all delivery options (including the no-purchase option) i that the current customer might choose. This is obtained by solving

$$\max_{\phi^i_{t+1} \in \Phi^i_{t+1}(s_t, c_t)_{i \in \mathscr{G}(s_t, c_t)}} (\cdot). \tag{6.26}$$

If no customer request arrives, the tour-planning decisions equal solving equation (6.26) with $i = 0$.

6.3.2 Modified Model

As described above, for the OP model presented in Section 6.3.1, not all opportunity cost properties discussed later hold. Therefore, in the following, an alternate model for OPs is presented, for which these properties hold. This alternate model is a modification of the model introduced in Section 6.3.1 and it differs from the natural model in regard of *cost realization* and *cost modeling*. Thereby, cost realization concerns the point of time in which the cost are incurred in the real application. Cost modeling concerns the decision epoch in which the corresponding cost are taken into account within the MDP model. In the natural model, cost realization and cost modeling match. In the modified model, cost modeling is delayed. Therefore, the state of the respective MDP is augmented (compared to the natural OP model) and the corresponding transition, and the value function are adapted. All other model components equal those as described for the natural OP model.

State

For the modified model, a third state component denoted by $r_t^{l\ cum}$ is added. It captures the cumulative logistics-related rewards, i.e., the negative of the cumulated fulfillment cost that realized before decision epoch t. Thus, the state is defined as:

$$s_t = (\mathscr{C}_t, \phi_t, r_t^{l\ cum}) \tag{6.27}$$

The state space $\mathscr{S}$ comprises all combinations of possible customer requests and arrival times with potentially chosen fulfillment options and potential cumulative logistics-related rewards.

Transition

The transition of the additional state component $r_t^{l\ cum}$ equals:

$$r_{t+1}^{l\ cum} = r_t^{l\ cum} + r^l_{\phi_{t+1}^{i'}}. \tag{6.28}$$

Value function

For the modified model, cost modeling is delayed to decision epoch $T+1$. Consequently, during the decision epochs $t = 1, ..., T$, only rewards $r^i(g_t)$ and $r^i_{c_t}$ are considered in the value function, which is hence defined as:

$$\begin{aligned} V(s_t) = \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \max_{g_t \in \mathscr{G}(s_t, c_t)} \\ \left(\sum_{i \in g_t} P^i(g_t) \cdot \left[r^i(g_t) + r^i_{c_t} + \max_{\phi^i_{t+1} \in \Phi^i_{t+1}(s_t, c_t)_{i \in \mathscr{G}(s_t, c_t)}} V(s_{t+1} \mid s_t, \phi^i_{t+1}) \right] \right) \\ + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot \max_{\phi^0_{t+1} \in \Phi^0_{t+1}(s_t)} V(s_{t+1} \mid s_t, \phi^0_{t+1}) \end{aligned} \tag{6.29}$$

Rewards $r^l_{\phi^{i'}_{t+1}}$ are only considered in the boundary condition in that the salvage value equals the respective state component:

$$V(s_{T+1}) = r^{l\ cum}_{T+1} \tag{6.30}$$

When the value function of the natural model is denoted by $V(s_t)$ and the one of the modified model as $V'(s_t)$, then the following relationship holds: $V(s_t) = V'(s_t) - r^{l\ cum}_t$.

Table 6.3 summarizes the introduced modeling framework by combining and delineating the MDP model components of the introduced DJP model, the natural OP model, and the modified OP model, respectively.

6.3.3 Model equivalency

In this section, the equivalency of the natural and the modified model is proven. This shows that it is possible to exploit all opportunity cost properties that are discussed later on in Chapter 8, not only for DJPs but also for OPs. Before proving the model equivalency, the corresponding value functions of the natural and the modified model are reformulated for ease of presentation. Then, model equivalency is proven by induction.

Component	DJP	natural OP	modified OP
Decision epochs	t = 1,..., T		
State	$s_t = \mathscr{C}_t$	$s_t = (\mathscr{C}_t, \phi_t)$	$s_t = (\mathscr{C}_t, \phi_t, r_t^{l\ cum})$
Actions	$a_t = g_t$	$a_t = (g_t, (\phi^i_{t+1})_{i \in g_t})$	
Transitions	$\mathscr{C}_{t+1} = \begin{cases} \mathscr{C}_t, \\ (\mathscr{C}_t \cup \{c_t\}), \end{cases}$	$\phi_{t+1} = \phi^{i'}_{t+1}$ $\mathscr{C}_{t+1} = \begin{cases} \mathscr{C}_t \setminus \Psi(\phi_{t+1} \mid \phi_t) \\ (\mathscr{C}_t \cup \{c_t\}) \setminus \Psi(\phi_{t+1} \mid \phi_t) \end{cases}$	$r^{l\ cum}_{t+1} = r^{l\ cum}_t + r^-_t$ $\phi_{t+1} = \phi^{i'}_{t+1}$ $\mathscr{C}_{t+1} = \begin{cases} \mathscr{C}_t \setminus \Psi(\phi_{t+1} \mid \phi_t) \\ (\mathscr{C}_t \cup \{c_t\}) \setminus \Psi(\phi_{t+1} \mid \phi_t) \end{cases}$
Rewards	r_t^+, r_t^-		
Values	$V(s_t) = r_t^+ + V(s_t+1))$	$V(s_t) = r_t^+ + r_t^- + V(s_t+1))$	$V(s_t) = r_t^+ + V(s_t+1))$
Boundary conditions	$V(s_{T+1}) = r^l_{T+1}$	$V(s_{T+1}) = 0$	$V(s_{T+1}) = r^{l\ cum}$

Table 6.3: Differences of i-DMVRP models

The value functions $V(s_t)$ and $V'(s_t)$ of the natural and the modified model, respectively, are reformulated as follows: first, the maximization operators are replaced by the corresponding optimal decisions. The optimal demand-management decision, i.e., $\max_{g_t \in \mathscr{G}(s_t, c_t)}$, is represented by g_t^*. The optimal tour-planning decision, i.e., $\max_{\phi^i_{t+1} \in \Phi^i_{t+1}(s_t, c_t)_{i \in \mathscr{G}(s_t, c_t)}}$, is represented by $\phi^{*\ i}_{t+1}$. Further, expressions r_t^+ and r_t^- are introduced to replace expectations over positive rewards (revenues) and negative rewards (costs), respectively.

Value function of the natural model:

$$\begin{aligned}
V(s_t) &= \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \left(\sum_{i \in g_t^*} P^i(g_t^*) \cdot \left[r^i(g_t^*) + r^i_{c_t} + r^l_{\phi^{*\,i}_{t+1}} + V(s_{t+1} \mid s_t, \phi^{*\,i}_{t+1}) \right] \right) \\
&\quad + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot \left(r^l_{\phi^{*\,0}_{t+1}} + V(s_{t+1} \mid s_t, \phi^{*\,0}_{t+1}) \right) \\
&= \underbrace{\sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot \left(r^i(g_t^*) + r^i_{c_t} \right)}_{r_t^+} \\
&\quad + \underbrace{\sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot r^l_{\phi^{*\,i}_{t+1}} + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot r^l_{\phi^{*\,0}_{t+1}}}_{r_t^-} \\
&\quad + \underbrace{\sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot V(s_{t+1} \mid s_t, \phi^{*\,i}_{t+1}) + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot V(s_{t+1} \mid s_t, \phi^{*\,0}_{t+1})}_{V(s_{t+1})} \\
&= r_t^+ + r_t^- + V(s_{t+1})
\end{aligned} \tag{6.31}$$

Value function of the modified model:

$$\begin{aligned}
V'(s_t) &= \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \\
&\quad \left(\sum_{i \in g_t^*} P^i(g_t^*) \cdot \left[r^i(g_t^*) + r^i_{c_t} + V'(s_{t+1} \mid s_t, \phi^{i\,*}_{t+1}) \right] \right) \\
&\quad + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot V'(s_{t+1} \mid s_t, \phi^{0\,*}_{t+1}) \\
&= \underbrace{\sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot (r^i(g_t^*) + r^i_{c_t})}_{r_t^+} \\
&\quad + \underbrace{\sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot V'(s_{t+1} \mid s_t, \phi^{i\,*}_{t+1}) + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot V'(s_{t+1} \mid s_t, \phi^{0\,*}_{t+1})}_{V'(s_{t+1})} \\
&= r_t^+ + V'(s_{t+1})
\end{aligned} \tag{6.32}$$

Model equivalency is given, if $V(s_t) = V'(s_t) - r_t^{l\ cum}$ holds for every t, and $V(s_0) = V^{'}(s_0)$ holds as well. This is proven by induction, for which the following prepositions are used:

Preposition 6.1. $V_{s_{T+1}} = 0$

Preposition 6.2. $r_t^{l\ cum} = r_{t+1}^{l\ cum} - r_{\phi_{t+1}^{i'}}^{l}$

The proof starts with the result for the terminal state s_{T+1} to which the system transitioned via a certain sample path denoted by ω, i.e., a specific sequence of stochastic realizations throughout the decision epochs. Thus, for this realized sample path, the respective probabilities in r_t^{+}, r_t^{-} and $V(s_{t+1})$ of equations (6.32) and (6.31) equal 1 and the probabilities of other realizations $\omega' \neq \omega$ equal 0, which is formalized by an additional index ω. Then, independent of which sample path realizes, i.e., for every sample paths in the set of all potential sample paths, $\omega \in \Omega$, the following proof by induction can be conducted:

Proof. By induction:

Initial case:

$$V^{\omega}(s_{T+1}) = V'^{\omega}\ (s_{T+1}) - r_{T+1}^{\omega\ l\ cum} = 0 \tag{6.33}$$

Equation 6.33 holds by definition.

Induction hypothesis:

$$V^{\omega}(s_{t+1}) = V'^{\omega}(s_{t+1}) - r_{t+1}^{\omega\ l\ cum} \tag{6.34}$$

Induction step:

From Equation (6.31) the following relationship can be derived: $V^{\omega}(s_t) = V^{\omega}(s_{t+1}) + r_t^{\omega\ +} + r_t^{\omega\ -}$. Then, the induction hypothesis (6.34) and Preposition 6.2 are substituted:

$$V^{\omega}(s_t) = (V'^{\omega}(s_{t+1}) - r_{t+1}^{l\ cum}) + r_t^{\omega\ +} + r_t^{\omega\ -} \tag{6.35}$$

$$= (V'^{\omega}(s_t) - r_t^{\omega\ +}) - (r_t^{l\ cum} + r_{\phi_{t+1}^{i'}}^{l}) + r_t^{\omega\ +} + r_t^{\omega\ -} \tag{6.36}$$

Since $r^{l}_{\phi^{i'}_{t+1}} = r^{\omega\,-}_{t}$ for i' that is observed in sample path ω and $r^{l}_{\phi^{i'}_{t+1}} = 0$ for all other i', this yields:

$$V^{\omega}(s_t) = V'^{\omega}(s_t) - r^{l\ cum}_{t} \tag{6.37}$$

for any $t = 0, ..., T$ and thus also for $t = 0$, i.e., $V_0 = V'_0 + r^{l\ cum}_{0}$. □

It holds by definition that $r^{l\ cum}_{0} = r^{\omega\,-}_{0} = 0$. Consequently, $V_0 = V'_0$ holds and the natural and the modified OP models are equivalent.

The most important insights from Part III can be summarized as follows: i-DMVRPs have been introduced and delineated. Further, a uniform taxonomy has derived to classify practical applications and new research problem settings. The related literature has been discussed with regard to different perspectives and with a special focus on modeling, the definition of opportunity cost, and solution approaches. Therewith, substantial research gaps have been elaborated and the first one has been closed. More precisely, a unified, explicit MDP modeling framework that captures temporal interdependencies of the two types of integrated decisions has been developed. The following Parts IV and V address the outstanding research gaps that have been identified in Chapter 5.

Part IV

Analytical discussion of opportunity cost for i-DMVRPs

In Part IV of this dissertation, a unified, specifically tailored definition of opportunity cost for i-DMVRPs is introduced.

As shown in Section 1.3, in traditional demand-management applications, like the airline or car rental industry (Klein et al. (2020)), solving the value function is intractable, such that many researchers approach demand-management problems by decomposing them into the two subproblems: (1) calculating opportunity cost, (2) optimizing demand-management decisions. This is why the analysis of opportunity cost and its properties has already become a standard tool for characterizing problem settings and for tailoring efficient approximation methods (Talluri and Van Ryzin (2006), Chapter 2).

Solving the value function of an i-DMVRPs is also intractable for realistic sized instances, but unfortunately, the results of the existing opportunity cost analyses from traditional demand-management applications cannot be transferred directly. This is due to a substantially more complex cost structure in i-DMVRPs compared with those of, e.g., the airline or car rental industry. More precisely, in i-DMVRPs, demand-management decisions influence subsequent tour-planning decisions and vice versa across the entire planning horizon (Agatz et al. (2013)). Thus, the variable cost of a certain demand-management decision depends on the set of finally served customer orders and can therefore only be determined ex post. This necessitates a specific definition and interpretation of opportunity cost as it substantially changes its composition.

However, despite the well-known relevance of opportunity cost for demand management, the discussion in Chapter 5 demonstrates that there is no such unified definition and interpretation for i-DMVRPs in the literature. Existing solution approaches base decision making on different opportunity cost definitions and thus incorporate different (sub-)components, respectively. As a consequence, optimization potential is left unexploited for individual problem settings. Further, the transferability of solution approaches among different problem settings is inhibited.

Therefore, a unified definition of opportunity cost that explicitly addresses the complex cost structure of i-DMVRPs is introduced. It is tailored to capture all cost components occurring in DJPs as well as in OPs. The respective contributions of this part are the following:

(1) The theoretical foundation of the concept of opportunity cost for i-DMVRPs is deepened by elaborating the decisive differences between opportunity cost in traditional demand-management applications and in i-DMVRPs.

(2) A unified opportunity cost definition, specifically tailored for i-DMVRPs, which is valid for DJPs and OPs at the same time, is developed and analyzed analytically.

(3) Central properties of opportunity cost for i-DMVRPs are derived and proven, which is essential for developing efficient approximation methods and the transferability of solution approaches among different problem settings.

This part of the dissertation is organized as follows: in Chapter 7, the concept of opportunity cost from traditional demand-management applications is analyzed with regard to crucial underlying assumptions which are then discussed for i-DMVRPs. Afterwards, in Chapter 8, four central properties of the newly derived opportunity cost components and also of the respective value functions are analytically elaborated and proven.

7. Transferring the concept of opportunity cost to i-DMVRPs

In this chapter, the concept of opportunity cost from traditional demand-management applications is transferred to i-DMVRPs. First, in Section 7.1, a general definition of opportunity cost is provided. For this, a so called interim state is introduced that separates the effects of a demand-management decision from those of the simultaneous tour-planning decision. Afterwards, in Section 7.2, underlying assumptions from traditional applications are investigated with regard to i-DMVRPs. As a conclusion, in Section 7.3, the traditional interpretation of opportunity cost is amended for i-DMVRPs.

7.1 A general definition of opportunity cost

Due to the well-known curses of dimensionality (Powell et al. (2012), Chapter 4), it is not possible to solve the i-DMVRP models presented in Chapter 6. Instead, as described in Section 1.3, it is common to decompose the demand-management problem. Thereby, the difference of two values of alternate states succeeding certain actions is determined as a change of future reward due to an action. This difference, called *opportunity cost*, is then used as input for finding the optimal demand-management decision.

In i-DMVRPs, in particular in OPs, an action can comprise two types of integrated decisions, namely, demand-management and tour-planning decisions. However, in this part of the dissertation, only the effects of a demand-management decision are of interest. Thus, it is the target to calculate opportunity cost from comparing state values that only reflect such effects separated from potential effects of tour-planning decisions of the same decision epoch. Hence, in the following, a new conceptual approach to separate the effects of two or more integrated decisions in one decision epoch of an MDP, as for example the demand-management and the tour-planning decisions in Equation (6.23), is developed:

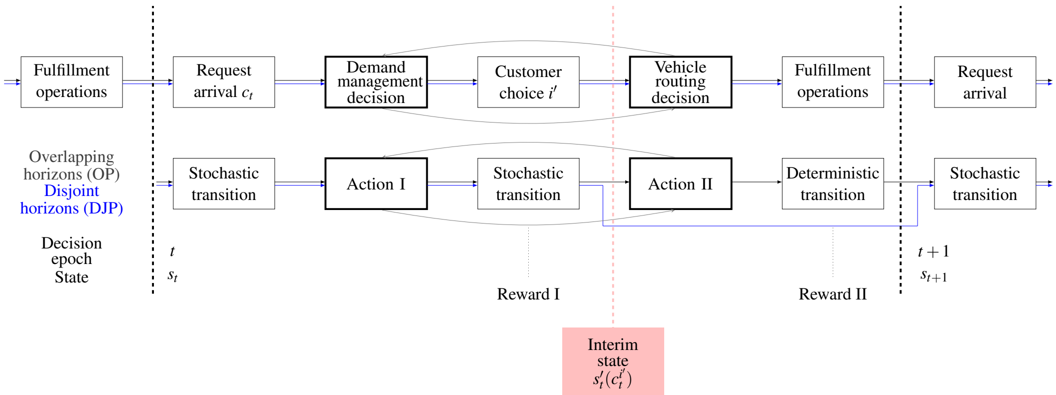


Figure 7.1: Overview of the MDP model of the i-DMVRP booking and fulfillment process including the interim state

A fictive state is introduced for each decision epoch $t = 1, ..., T$. It is referred to as *interim state* and denoted as $s'_t \mid s_t, c_t, i$. This interim state describes the state that is reached if a certain fulfillment option i is sold in state s_t with customer request c_t. ($s'_t \mid s_t$ denotes the respective interim state at decision epoch t, if no fulfillment option is sold.) The idea behind it is similar to the idea of the *post-decision state* introduced by Powell et al. (2012), Chapter 4, namely to isolate different effects of decisions and information on the state variable. More precisely, the post-decision state separates the deterministic effect of a decision from the stochastic effect of the *same* decision in order to ease decision making. Still, s'_t does not equal a post-decision state, as it does not separate stochastic and deterministic effects of the same decision. Instead, it separates the effects of two *different* decisions, i.e., the stochastic and deterministic effects of the demand-management decision from the deterministic effects of a tour-planning decision taken in the same decision epoch. Figure 7.1 is a modification of the overview of the MDP model of the i-DMVRP booking and fulfillment processes known from Chapter 6, but it includes the interim state in order to illustrate its integration into the presented MDP models.

Generally, the value of such an interim state $s'_t \mid s_t, c_t, i$ can be calculated as the sum of the logistics-related rewards $r^l_{\phi^{*\,i}_{t+1}}$ and the successor state's value, i.e., of state $s_{t+1} \mid s_t, \phi^{*\,i}_{t+1}$:

$$V(s'_t \mid s_t, c_t, i) = r^l_{\phi^{*\,i}_{t+1}} + V(s_{t+1} \mid s_t, \phi^{*\,i}_{t+1}). \tag{7.1}$$

Thus, also the state value $V(s_t)$ can be calculated based on an interim state value, thereby isolating the demand-demand management decision as follows:

$$\begin{aligned} V(s_t) = \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \max_{g_t \in \mathscr{G}(s_t, c_t)} \left(\sum_{i \in g_t} P^i(g_t) \cdot \left[r^i(g_t) + r^i_{c_t} + V(s'_t \mid s_t, c_t, i) \right] \right) \\ + \left(1 - \sum_{c_t \in C} \lambda_{c_t}(t)\right) \cdot V(s'_t \mid s_t, c_t, 0). \end{aligned} \tag{7.2}$$

In DJPs, the respective value of an interim state equals the value of the successor state, i.e., $V(s'_t \mid s_t, c_t, i) = V(s_{t+1} \mid s_t, c_t, i)$, since there is no tour-planning decision during the booking horizon. In the following, for ease of presentation, interim state $s'_t \mid s_t, c_t, i$ is denoted by $s'_t(c^i_t)$ and interim state $s'_t \mid s_t, c_t, 0$ is denoted by $s'_t(0)$.

Then, the following definition formalizes the concept of opportunity cost for solving the demand-management problem of i-DMVRPs:

Definition 7.1. *The opportunity cost $\Delta V(s_t, c_t, i)$ of a certain fulfillment option i when being in a certain state s_t and a customer request c_t arrives is calculated as the difference of the values of the following two interim states: (1) the interim state following a demand-management decision that results in the customer of c_t choosing the no-purchase option $i = 0$ and (2) the interim state following a demand-management decision that results in the customer of c_t choosing that particular fulfillment option i. Thus, it is defined as*

$$\Delta V(s_t, c_t, i) = V(s'_t(0)) - V(s'_t(c^i_t)). \tag{7.3}$$

This opportunity cost is then used as input to solve the demand-management problem, which can be illustrated by the following reformulation of the corresponding value functions (6.4) and (6.29), respectively. This reformulation is typical in revenue-management literature (e.g., Yang et al. (2016), Fleckenstein et al. (2021)) and yields:

$$V(s_t) = \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \underbrace{\max_{g_t \in \mathscr{G}(s_t, c_t)} \sum_{i \in g_t} P^i(g_t) \cdot \left[r^i(g_t) + r^i_{c_t} - \Delta V(s_t, c_t, i) \right]}_{\substack{\text{Demand management problem} \\ \text{here: assortment optimization}}} + V(s'_t(0)) \tag{7.4}$$

Since a demand-management decision only has to be taken when a certain customer request arrives, the probability $\sum_{c_t \in C} \lambda_{c_t}(t)$ is not relevant for decision making. Also, the second summand of equation (7.4), i.e., $V(s'_t(0))$, is not relevant as it is a constant and independent of the decision. Further, $P^i(g_t)$, $r^i(g_t)$ and $r^i_{c_t}$ are assumed to be known. Thus, if the opportunity cost of a fulfillment option i, $\Delta V(s_t, c_t, i)$, is known and the set $\mathscr{G}(s_t, c_t)$ is not large, it is possible to solve the demand-management problem by full enumeration. Indeed, for i-DMVRPs, the set $\mathscr{G}(s_t, c_t)$ is assumed to be of tractable size. However, calculating opportunity cost still involves solving a recursive function that is intractable for realistic sized instances (Fleckenstein et al. (2021)) and would further need to be calculated in real-time. This is why, for realistic problems, it is necessary to find accurate and efficient approximation approaches for opportunity cost and, thus, for the value function (Yang et al. 2016). This motivates a deeper understanding of opportunity cost and of its peculiarities and properties in i-DMVRPs. Consequently, in the following, opportunity cost in traditional demand-management problems and opportunity cost in i-DMVRPs are compared and decisive differences are discussed. As a conclusion, the definition of opportunity cost is augmented by a specifically tailored interpretation for i-DMVRPs. Then, in Chapter 8, four central properties of the value functions and the corresponding opportunity cost in i-DMVRPs are elaborated and proven analytically.

7.2 Comparing opportunity cost in traditional applications and i-DMVRPs

As described in Section 1.1, in traditional demand-management applications, the following two assumptions are made (Weatherford and Bodily (1992)):

Assumption 1 – Supply is inflexible, i.e., capacities are fixed.

Assumption 2 – Variable costs associated with the usage of capacity are either negligible or at least directly attributable to individual orders.

If those assumptions hold, opportunity cost is equivalent to displacement cost (DPC) and is defined as “the expected loss in future revenue from using the capacity now rather than reserving it for future use” (Talluri and Van Ryzin (2006), Chapter 2).

In the following, it is shown that this definition cannot be transferred to i-DMVRPs by investigating the underlying assumptions in the respective context. For illustrative reasons and in order to generate a general intuition of opportunity cost in i-DMVRPs and its decisive characteristics, it is investigated for a simple problem instance, which is introduced in the following. However, the respective results can be generalized to more complex problem instances.

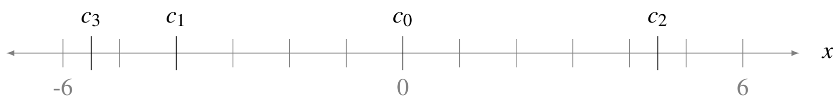


Figure 7.2: Customer locations of the problem instance underlying the discussion of opportunity cost properties in i-DVMRPs

The considered problem instance is a very basic instance of the DJP presented in Sections 4.1 and 6.2. It is assumed that there are only three potential customer requests, denoted as c_1, c_2, and c_3. As depicted in Figure 7.2, they are located on a line with one single, centrally located depot, denoted as c_0. There are three decision epochs in which customer requests arrive with time-dependent arrival rates $\lambda_{c_i}(t) = 0.5$, if $i = t$ and $\lambda_{c_i}(t) = 0$, else. The potential revenues associated with those customer requests are $10, 10$ and 20 *monetary units* (*MU*) for requests c_1, c_2, and c_3, respectively. The customer request characteristics are summarized in Table 7.1. Further, for every customer request that turns into a customer order, the same physical capacity consumption is assumed. It equals the size of one trunk. To serve customer orders, the provider has a single vehicle available, which can only load two trunks at a time and is not allowed to conduct multiple trips.

c_i	(x)	r_{c_i}	$\lambda_{c_i}(t)$
c_0	(0)	0	–
c_1	(-4)	10	$\lambda_{c_i}(t) = \begin{cases} 0.5 & \text{if } i = t \\ 0 & \text{else} \end{cases}$
c_2	(4.5)	10	
c_3	(-5.5)	20	

Table 7.1: Customer requests of the problem instance underlying the discussion of opportunity cost properties in i-DVMRPs

For each incoming customer request, the provider can only offer one fulfillment option ($i = 1$) and the fictive no-purchase option ($i = 0$). The provider does not charge a delivery fee, i.e., $r^1 = r^0 = 0MU$. Consequently, there are only two potential offer sets $g \in \mathcal{G} = \{1, 2\}$, as depicted in Table 7.2, among which the provider can choose one to present to an incoming customer request. It is assumed that a customer request turns into a customer order with probability $P^1 = 1$, if $i = 1$ is offered within an offer set and $P^0 = 1$, otherwise. Thus, presenting offer set $g = 1$ to an incoming customer request equals accepting it. Presenting offer set $g = 2$ to an incoming customer request equals rejecting it. Thus, for ease of readability, in the following, presenting offer set $g = 1$, i.e., action $a_t = 1$, is referred to as accepting the respective customer request. Presenting offer set $g = 2$, i.e., action $a_t = 2$, is referred to as rejecting it. It has to be noted that, without loss of generality, this can be transferred to problem settings with more than one fulfillment option. Then, "accepting a customer request" can be understood as accepting a customer for a certain fulfillment option or rejecting them. Fuel cost are assumed to equal 1MU per *unit length (UL)*. With this problem instance at hand, first, Assumption 1 of traditional demand-management applications is discussed:

Assumption 1 – In most i-DMVRPs, the first assumption of traditional demand-

	prices of fulfillment options		choice probabilities	
g	$i=0$	$i=1$	P^0	P^1
1	0	0	0	1
2	0	not offered	1	0

Table 7.2: Available offer sets of the problem instance underlying the discussion of opportunity cost properties in i-DVMRPs

management applications is valid. Depending on the problem at hand, either driver working times, fleet sizes, or loads represent resources with fixed capacities. As expected, such limited resources may cause a displacement of demand as the following example illustrates:

> Example 6.1: In this example, decision epoch $t = 2$ of the above described problem instance is investigated. It is assumed that a customer request c_1 arrived in the previous decision epoch and turned into a customer order. In $t = 2$, customer request c_2 realizes. Thus, $\mathscr{C}_2 = \{c_1\}$ and the provider has to decide whether to accept the current customer request c_2 with action $a_2 = 1$, or reject it by action $a_2 = 2$. Clearly, if the provider accepts c_2, it turns into a confirmed customer order. Then, it is not possible to also accept the customer request c_3, which realizes at the subsequent decision epoch with probability $\lambda_{c_3}(3) = 0.5$. Consequently, decision $a_2 = 1$ results in expected displacement cost that equals $\lambda_{c_3}(3) \cdot r_{c_3} MU = 0.5 \cdot 20MU = 10MU$. This means that an expected revenue of $10MU$ is displaced due to limited vehicle capacities if decision $a_2 = 1$ is taken.

Assumption 2 – The second assumption of traditional demand-management problems, negligible or directly attributable variable cost, does not hold in most i-DMVRPs. This can be shown by considering fuel cost as an example: within one tour of one vehicle that visits various customer locations, it is not possible to attribute a certain share of fuel consumption to the individual customer locations. The consumption depends on the specific combination of customer locations in the tour, and hence, the share of each customer location changes, when other customer locations are added to or removed from the tour. Thus, there is no way to calculate and attribute individual fuel cost (see for example Vinsensius et al. (2020)). Further, such costs are not negligible as the following example shows.

Example 6.2a: Again, the same decision epoch, with equal state and potential actions of the problem instance as described in Example 6.1, is considered. If fuel cost were to be neglected, DPC would equal opportunity cost, i.e., $\Delta V(s_2, c_2, 1) = 10MU$. Since the immediate contribution of action $a_2 = 1$ also equals $r_{c_2} = 10MU$, both decisions, $a_2 = 1$ or $a_2 = 2$, are equally good decisions for the provider.

Example 6.2b: However, Figure 7.2 shows clearly that the additional fulfillment cost in case that customer request c_2 turns into a confirmed customer order equals $9MU$ when optimal subsequent decisions are assumed. In turn, rejecting customer request c_2 by action $a_2 = 2$ and then accepting customer request c_3 instead, only leads to additional fulfillment cost of $3MU$. Since customer request c_3 realizes with probability $\lambda_{c_3}(3) = 0.5$, the expected increase in delivery cost caused by decision $a_2 = 1$ is calculated as $9MU - 0.5 \cdot 3MU = 7.5MU$. Considering this cost additionally to the previously calculated DPC, action $a_2 = 1$ causes an expected cost of $17.5MU$. Since the immediate contribution of accepting customer request c_2 is below this expected cost, the provider has to decide for $a_2 = 2$ in order to operate profitably.

Example 6.2a) and b) show that the exact same combination of state and customer request of the same problem instance yield different optimal demand-management decisions depending on whether fulfillment cost are taken into consideration or are being neglected. This shows that Assumption 2 does not hold in i-DMVRPs and, thus, that variable fulfillment cost cannot be neglected. Consequently, the traditional concept of opportunity cost, which equalizes opportunity cost and displacement cost (Talluri and Van Ryzin (2006), Chapter 2), has to be adapted for i-DMVRPs. Therefore, in the following Section 7.3 a novel definition for opportunity cost is introduced.

7.3 Extension of the concept of opportunity cost for i-DMVRPs

It has been shown that neglecting variable fulfillment cost leads to suboptimal demand-management decisions in i-DMVRPs and, hence, that it is required to take them into consideration for decision making. Further, it has been shown, that variable fulfillment cost cannot be attributed to individual customer requests in i-DMVRPs. Thus, it is not possible to consider them implicitly by optimizing demand-management decisions with regard to customer requests' individual contribution margins instead of revenues.

Instead, in the literature on i-DMVRPs, some authors explicitly consider the marginal increase of variable cost caused by the acceptance of a customer request as already exemplarily calculated in Example 6.2b. In the literature, this marginal increase is referred to as *(marginal cost to serve (MCTS))* (see for example Vinsensius et al. (2020), Strauss et al. (2021)). In a static context, a request's MCTS can be calculated by optimizing the tour plan for all accepted customer orders including the current request and comparing its fulfillment cost with the cost of the optimal tour plan without the current request. Calculating this value at a certain decision epoch of an i-DMVRP yields exact myopic MCTS (Fleckenstein et al. 2021). However, those are not sufficient for optimal decision making, as the following example shows:

Example 6.3: Again, the same problem instance as in the previous examples, with the same potential customer requests regarding locations, revenues, and arrival rates as depicted in Table 7.1 is considered. Fuel cost are again $1MU/UL$. This time, decision epoch $t = 1$ is examined. There are no confirmed customer orders yet, i.e., $\mathscr{C}_1 = \{\}$, and a request c_1 realizes. To calculate DPC, the sum of all expected revenues that can be accrued under optimal decision making in the subsequent decision epochs until the terminal decision epoch starting from interim state s'_1 has to be calculated. It equals $15MU$ as explained in the following. If customer request c_1 is rejected, it is still possible to accept c_2 and c_3 if they realize. Thus, it is possible to accrue the respective revenues

$10MU$ and $20MU$ with the respective arrival probabilities $\lambda_{c_2}(2) = \lambda_{c_3}(3) = 0.5$. From that, the corresponding sum of expected revenues under optimal decision making starting from interim state $s'_1(c_1)$, which equals $10MU$, has to be substracted. This yields $DPC = 15MU - 10MU = 5MU$. The myopic MCTS equal $8MU$. Consequently, if the provider bases decision making on DPC and myopic MCTS, the resulting optimal decision is to reject customer request c_1 by action $a_1 = 2$.

Nevertheless, the optimal decision resulting from solving the value function is accepting customer request c_1 by action $a_1 = 1$. This is also the intuitive decision when looking at Figure 7.2 and considering the vincinity to potential future customer request c_3, or more precisely, when considering the cost-related, anticipatory opportunity effect of accepting customer request c_1. Thus, for optimal decision making, opportunity cost for i-DMVRPs cannot only be revenue-related in form of DPC but also have to take cost-related effects on opportunity cost into account in form of expected MCTS. In the following, only the term MCTS is used and thereby refers to expected MCTS. If a different interpretation applies, it is explicitly stated. Correspondingly, the definition of opportunity cost is amended as follows:

Definition 7.2. *In i-DMVRPs with variable costs that are not directly attributable to customer requests, opportunity cost comprises two components: DPC as the difference of cumulated future expected revenues caused by accepting a customer request and MCTS as the difference of future expected fulfillment cost caused by accepting a customer request, both assuming optimal decision making in subsequent decision epochs.*

To define both terms formally, at first *cumulated future revenues* and *expected future fulfillment cost* are defined:

Definition 7.3. *The cumulated future revenues given interim state s'_t at decision epoch t and optimal decision making in all subsequent decision epochs until T is* *defined as:*

$$\begin{aligned}\mathbb{E}(R_t \mid s'_t) = & \underbrace{\sum_{c_{t+1} \in C} \lambda_{c_{t+1}}(t+1) \cdot \sum_{i \in g^*_{t+1}} P^i(g^*_{t+1}) \cdot \left(r^i(g^*_{t+1}) + r^i_{c_{t+1}}(g^*_{t+1})\right)}_{r^+_{t+1}} \\ & + \mathbb{E}(R_{t+1} \mid s'^{\,*}_{t+1}),\end{aligned} \tag{7.5}$$

with boundary condition:

$$\mathbb{E}(R_T \mid s'_T) = 0, \tag{7.6}$$

and with g^*_{t+1} denoting the optimal demand-management decision in $t+1$, and $s'^{\,*}_{t+1}$ denoting the resulting interim state, assuming optimal tour-planning decisions in t.

Definition 7.4. *The expected future fulfillment cost given interim state s'_t at decision epoch t and optimal decision making in all subsequent decision epochs until T is defined as:*

$$\begin{aligned}\mathbb{E}(C_t \mid s'_t) = & r^l_{\phi^*_{t+1} i'} + \sum_{c_{t+1} \in C} \lambda_{c_{t+1}}(t+1) \cdot \sum_{i \in g^*_{t+1}} P^i(g^*_{t+1}) \cdot \mathbb{E}(C_{t+1} \mid s'^{\,*}_{t+1}(c^i_{t+1})) \\ & + (1 - \sum_{c_{t+1} \in C} \lambda_{c_{t+1}}(t+1)) \cdot \mathbb{E}(C_{t+1} \mid s'^{\,*}_{t+1}(0)),\end{aligned} \tag{7.7}$$

with boundary condition:

$$\mathbb{E}(C_T \mid s'_T) = r^l_{\phi^*_{T+1} i'} \tag{7.8}$$

and with i' denoting the customer request c_t's choice realization ($i = 0$, if there is no customer request c_t), g^*_{t+1} denoting the optimal demand-management decision in $t+1$, and $s'^{\,*}_{t+1}$ denoting the resulting interim state. For DJPs holds $r^l_{\phi^*_T i'} = r^l_{T+1} = -C(VRP_{\mathscr{C}_{T+1}})$.

In the following, DPC and MCTS are formally defined:

Definition 7.5. *DPC of accepting customer request c_t for fulfillment option i in decision epoch t and state s_t is defined as:*

$$\Delta\mathbb{E}(R_t \mid s_t, c_t, i) = \mathbb{E}(R_t \mid s'_t(0)) - \mathbb{E}(R_t \mid s'_t(c^i_t)). \tag{7.9}$$

with

Calculating DPC depends on the state of the system and all consecutive decisions and transitions. Thus, it suffers from the curse of dimensionality in that the number of combinations of potential future acceptances/rejections is intractable.

Definition 7.6. *MCTS of accepting customer request c_t for fulfillment option i, in decision epoch t and state s_t is defined as:*

$$\Delta\mathbb{E}(C_t \mid s_t, c_t, i) = \mathbb{E}(C_t \mid s'_t(c^i_t)) - \mathbb{E}(C_t \mid s'_t(0)). \tag{7.10}$$

Like DPC, calculating MCTS also depends on the state of the system and all consecutive decisions and transitions. Thus, it also suffers from the curse of dimensionality in that the number of potential tour-planning decisions that have to be evaluated is intractable for realistic-sized instances. Additionally, it is well known that the underlying VRPs for every one of those tour-planning decisions are NP-hard (see for example Vinsensius et al. (2020)).

In the following chapter, Chapter 8, the newly derived definition of opportunity cost as well as the corresponding MCTS and DPC are investigated analytically.

8. Properties and analytical discussion of opportunity cost for i-DMVRPs

In Section 7.3, it has been shown that calculating opportunity cost or its components, DPC and MCTS, is intractable for realistic-sized i-DMVRPs. Consequently, solution approaches for i-DMVRPs require valid approximations of opportunity cost. Therefore, in the following, four central properties of the discussed i-DMVRPs, more precisely of the value functions (6.4), (6.23), and (6.29), and the derived opportunity cost values, are elaborated. Those are:

1. Decomposability into MCTS and DPC
2. Potential negativity of MCTS and DPC
3. Non-negativity of opportunity cost
4. Monotonicity of the value functions

8.1 Decomposability of opportunity cost into MCTS and DPC

To prove the decomposability of opportunity cost into MCTS and DPC, it is first shown that there is a valid decomposition for the value function of an interim state into two components. In other words, the value of an interim state, $V(s'_t)$, equals the difference of expected cumulated future revenues $\mathbb{E}(R_t \mid s'_t)$ and expected future routing cost $\mathbb{E}(C_t \mid s'_t)$, under optimal decisions. This leads to the following Lemma:

Lemma 8.1. *The value (function) of an interim state s_t' can be decomposed into two additive components, one capturing expected future revenues and one capturing expected future cost:*

$$V(s_t') = \mathbb{E}(R_t \mid s_t') + \mathbb{E}(C_t \mid s_t'). \tag{8.1}$$

Lemma 8.1 holds, if

$$\mathbb{E}(R_t \mid s_t') + \mathbb{E}(C_t \mid s_t') = r^l_{\phi^{*\,i}_{t+1}} + V(s_{t+1} \mid s_t, \phi^{*\,i}_{t+1}) \tag{8.2}$$

holds for every interim state s_t'.

Proof. By induction:

Initial case:
In the terminal decision epoch T, rewards are defined by the boundary conditions (6.24), (7.6) and (7.8). Substituted into (8.2), this leaves:

$$0 + r^l_{\phi^{*\,i}_{T+1}} = r^l_{\phi^{*\,i}_{T+1}} + 0, \tag{8.3}$$

showing that (8.2) holds for T.

Induction hypothesis: (8.2)

Induction step:

$$
\begin{aligned}
&\mathbb{E}(R_{t-1} \mid s'_{t-1}) + \mathbb{E}(C_{t-1} \mid s'_{t-1}) \\
=& \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot \left(r^i(g_t^*) + r^i_{c_t}(g_t^*)\right) + \mathbb{E}(R_t \mid s_t'^{\,*}) \\
&+ r^l_{\phi_t^{*}\, i'} + \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot \mathbb{E}(C_t \mid s_t'^{\,*}(c_t^i)) + (1 - \sum_{c_t \in C} \lambda_{c_t}(t)) \cdot \mathbb{E}(C_t \mid s_t'^{\,*}(0)) \\
=& r^l_{\phi_t^{*}\, i'} + \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \sum_{i \in g_t^*} P^i(g_t^*) \cdot \left(r^i(g_t^*) + r^i_{c_t}(g_t^*) + \mathbb{E}(C_t \mid s_t'^{\,*}(c_t^i))\right) \\
&+ (1 - \sum_{c_t \in C} \lambda_{c_t}(t)) \cdot \mathbb{E}(C_t \mid s_t'^{\,*}(0)) + \mathbb{E}(R_t \mid s_t'^{\,*})
\end{aligned}
\tag{8.4}
$$

Now, repeatedly substituting (7.1) and again the induction hypothesis (8.2) until T yields:

$$
r^l_{\phi_t^{*}\, i} + V(s_t \mid s_{t-1}, \phi_t^{*\, i}) \tag{8.5}
$$

□

With this in mind, the first property can be formally proven:

Property 8.2. *Opportunity cost can be decomposed into DPC and MCTS:*

$$
\Delta V(s_t, c_t, i) = \Delta\mathbb{E}(R_t \mid s_t, c_t, i) + \Delta\mathbb{E}(C_t \mid s_t, c_t, i). \tag{8.6}
$$

To prove Property 8.2, Lemma 8.1 is substituted into Equation (7.3). Further, Definitions 7.5 and 7.6 are substituted, which results in:

Proof.

$$
\begin{aligned}
\Delta V(s_t, c_t, i) &= V(s_t'(0)) - V(s_t'(c_t^i)) \\
&= \mathbb{E}(R_t \mid s_t'(0)) - \mathbb{E}(R_t \mid s_t'(c_t^i)) - \mathbb{E}(C_t \mid s_t'(0)) + \mathbb{E}(C_t \mid s_t'(c_t^i)) \\
&= \Delta\mathbb{E}(R_t \mid s_t, c_t, i) + \Delta\mathbb{E}(C_t \mid s_t, c_t, i)
\end{aligned}
\tag{8.7}
$$

□

8.2 Potential negativity of MCTS and DPC

In Example 6.2 of Chapter 7.2, a problem instance has been discussed, in which MCTS and DPC are both positive. However, contrary to traditional demand-management applications in which DPC can only be positive (Talluri and Van Ryzin (2006), Chapter 2), in i-DVMRPs, MCTS and DPC can be negative. This is the next property that will be discussed. Hence, in the following, constructed examples are presented in which either of the components is negative. First, a situation with negative MCTS is elaborated.

Example 7.1: The same problem instance as in Example 6.3 is considered, with the same potential customer requests regarding locations, revenues, and arrival rates as depicted in Table 7.1. Fuel cost are again assumed to equal $1MU/UL$ and decision epoch $t = 1$ is examined with $\mathscr{C}_1 = \{\}$. Customer request c_1 realizes. Decision $a_1 = 1$, i.e., accepting c_1, results in value $V(s'_1(c_1)) = 0.5MU$. Rejecting it by decision $a_1 = 2$, results in value $V(s'_1(0)) = 5MU$. Consequently, the corresponding opportunity cost of decision $a_1 = 1$ for customer request c_1 when being in the considered state equals $4.5MU$. DPC are calculated as in Example 6.3, thus, $\Delta\mathbb{E}(R_1 \mid s_1, c_1, 1) = 5MU$. Exploiting Property 8.2 yields $MCTS = \Delta\mathbb{E}(C_1 \mid s_1, c_1, 1) = \Delta V(s_1, c_1, 1) - \Delta\mathbb{E}(R_1 \mid s_1, c_1, 1) = 4.5MU - 5MU = -0.5MU$.

For illustrative reasons, Figure 8.1 shows the partial decision tree for this problem instance, originating in state s_1, assuming a customer request arrives. Random nodes are depicted as circles, which in Figure 8.1 represent whether there is a customer request or not. The outgoing upper arc always represents the arrival of a customer request, the outgoing lower arc represents the case that there is no such arrival. Decision nodes are depicted as rectangles and represent demand-management decisions. The upper arcs originating in such nodes represent accepting the respective customer request. The corresponding lower arcs represent rejecting the customer request. Optimal decisions in each decision epoch, derived from solving the corresponding value function, are depicted as solid arcs originating in the demand-management decision nodes.

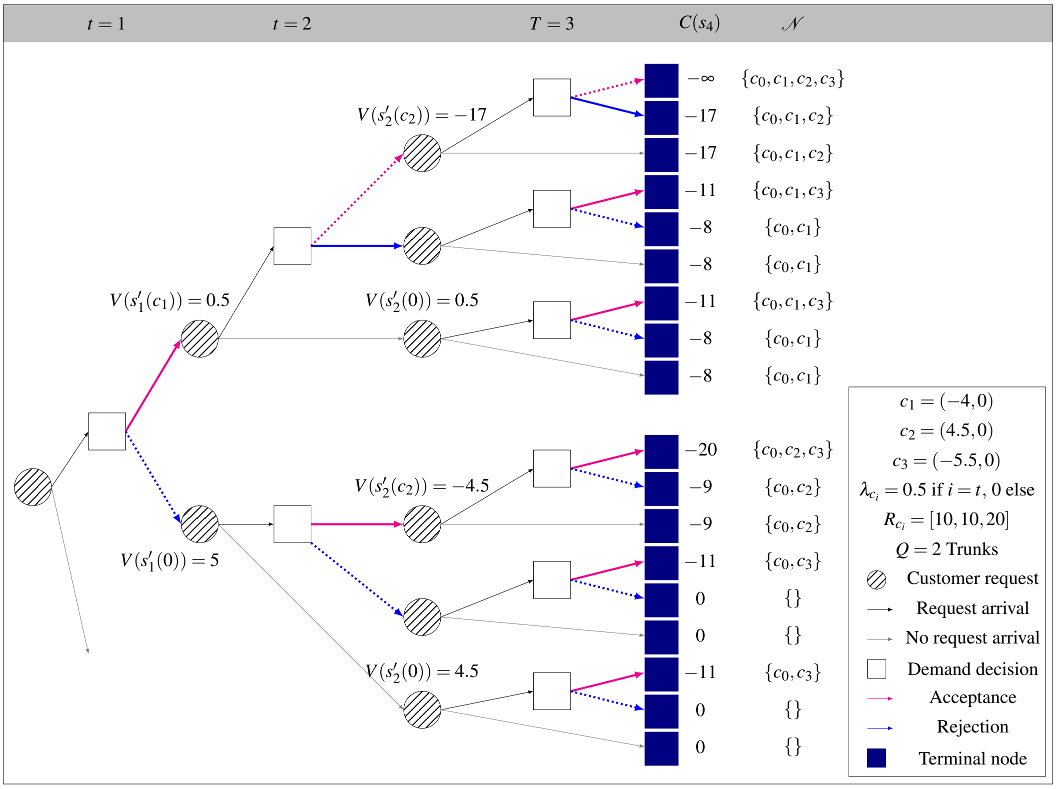


Figure 8.1: Decision Tree - Example 7.1

The intuition behind negative MCTS, as they occur in Example 7.1, is the following: accepting the corresponding customer request and following the subsequent optimal decisions leads to expected final routing cost that are lower than those generated by optimal decisions following the rejection of the same customer request. In other words, accepting a certain customer request inhibits the acceptance of one or more later customer requests, which would otherwise be accepted with optimal decisions and would lead to higher final fulfillment cost. In the example given, if customer request c_1 were to be rejected, it would be optimal to accept customer request c_2 in case it realizes. However, a final fulfillment tour that includes a customer order c_2 costs $9MU$ more than any final fulfillment tour that does not. Due to restricted capacities, by accepting customer request c_1, also accepting a customer request c_2 becomes suboptimal. Thus, the corresponding MCTS are negative and the acceptance of a request leads to an overall decrease of expected fulfillment costs.

In the following, an example with negative DPC is elaborated.

Example 7.2: To show a situation with negative DPC, the same problem instance as in the previous examples is considered. Thereby, the same potential customer requests regarding locations and arrival rates, but with potential revenues $r_{c_1} = 10MU$, $r_{c_2} = 10MU$, and $r_{c_3} = 10.5MU$ are assumed. Furthermore, now it is assumed that the physical vehicle capacity is unrestricted, and instead, the maximum route length is constrained to $12UL$. Travelling one UL still costs $1MU$.

Again, decision epoch $t = 1$ is investigated, and again, it is assumed that there is no confirmed customer order yet, i.e., $\mathscr{C}_1 = \{\}$, and a customer request c_1 realizes. Now, $\mathbb{E}(R_1 \mid s'_1(0)) = 5MU$ as with rejecting c_1 by action $a_1 = 2$, the subsequent optimal decisions lead to a future revenue of $10MU$ with probability $\lambda_{c_2}(2) = 0.5$. If request c_2 does not realize, still, a request c_3 will not be accepted. In case the current customer request c_1 converts into a confirmed customer order, in turn, it is optimal to also accept a customer request c_3, if it realizes. Consequently, $\mathbb{E}(R_1 \mid s'_1(c_1)) = 5.25MU$ and, thus, $\Delta\mathbb{E}(R_1 \mid s_1, c_1, 1) = -0.25MU < 0MU$.

Figure 8.2 illustrates optimal decisions for all decision epochs, derived from solving the corresponding value function.

The intuition behind negative DPC, as they occur in Example 7.2, is the following: turning the corresponding customer request into a customer order enables accepting one or more expected future customer requests in its vincinity that otherwise would not be profitable with regard to their related fulfillment cost and revenues. Thus, without including MCTS in the consideration of opportunity cost, DPC cannot be negative. This is a decisive difference between the traditional concept of opportunity cost and the newly derived concept for i-DMVRPs. Thus, it is formalized in the following:

Property 8.3. *MCTS and DPC can both be negative:*

$$\exists\, MCTS < 0 \tag{8.8}$$

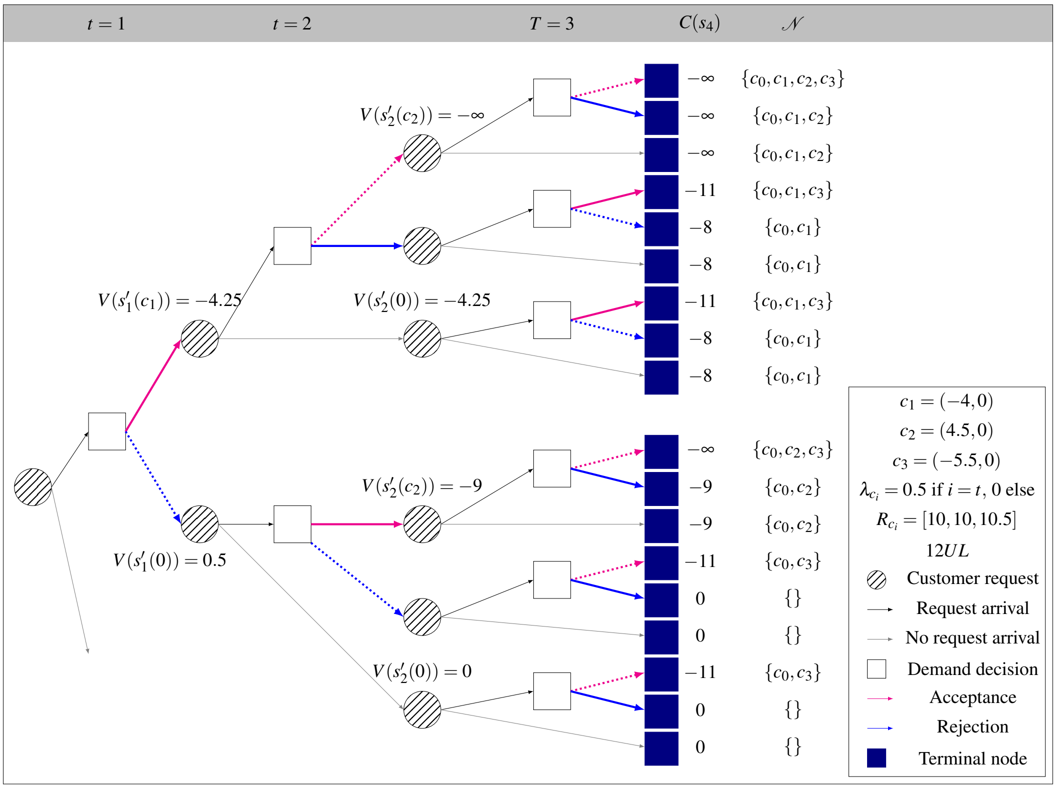


Figure 8.2: Decision Tree - Example 7.2

and

$$\exists\, DPC < 0. \tag{8.9}$$

Proof. By counterexamples 7.1 and 7.2. □

8.3 Non-negativity of opportunity cost

Despite the observation that both MCTS and DPC can be negative, it can be proven that, for value functions (6.4), (6.23), and (6.29) of the described MDP models, the resulting opportunity cost are always non-negative. To prove this property, it is shown that the value of the interim state following the acceptance of a customer request c_t by action $a_t = 1$ cannot be higher than the value of the interim state following a rejection of the same customer request c_t by action $a_t = 2$.

The corresponding proof builds on three lemmata, which formalize i-DMVRP model characteristics.

The first lemma concerns the stochastic transition probabilities, i.e., arrival rates and customer choice probabilities in a decision epoch t. Both are independent of the corresponding set of confirmed customer orders $\mathscr{C}_t$.

Lemma 8.4. *Stochastic transition probabilities are independent of the set of already confirmed customer orders:*

$$\forall t \in 1,...,T, i \in \mathscr{I}, c_i \in C, g_t \in \mathscr{G} : \lambda_{c_i}(t) \textit{ independent of } \mathscr{C}_t \tag{8.10}$$

and

$$P^i(g_t) \textit{ independent of } \mathscr{C}_t. \tag{8.11}$$

Proof. By definition. □

The second lemma concerns the relationship of the action spaces of any two states s_t and $\widehat{s_t}$, which only differ in that the latter contains the same confirmed customer orders as the first, but additionally, contains exactly one customer order, denoted by $\widehat{c}$, more, i.e., $\widehat{\mathscr{C}_t} = \mathscr{C}_t \cup \{\widehat{c}\}$. For those two states, the action space of the latter is a subset of the action space of the first.

Lemma 8.5. *The action space of any state $\widehat{s_t} = (\mathscr{C}_t \cup \{\widehat{c}\}, \phi_t)$ is a subset of the action space of a corresponding state $s_t = (\mathscr{C}_t, \phi_t)$:*

$$\forall t \in 1,...,T, c_t \in C, \widehat{c} \in C : \mathscr{A}(\widehat{s_t}, c_t) \subseteq \mathscr{A}(s_t, c_t). \tag{8.12}$$

Generally, the action space of a state comprises two components:

(1) the demand-management component $\mathscr{G}(s_t, c_t)$

(2) the tour-planning component $(\Phi^i_{t+1}(s_t, c_t))_{i \in \mathscr{G}(s_t, c_t)}$.

The demand-management component strongly depends on the tour-planning component in that it only comprises delivery options, for which a feasible tour plan can be found. Thus, to prove Lemma 8.5, it is sufficient to show that the tour-planning component of $\mathscr{A}(\widehat{s_t}, c_t)$ is a subset of the respective tour-planning component of $\mathscr{A}(s_t, c_t)$.

Proof. $(\Phi^i_{t+1}(s_t, c_t))_{i \in \mathscr{G}(s_t, c_t)}$ corresponds to the solution space of the constraint satisfaction variant of the underlying VRP. Enforcing the fulfillment of an additional customer order $\widehat{c}$ requires an additional constraint compared to fulfilling the set of customer orders $\mathscr{C}_t$, without $\widehat{c}$. This constraint is either redundant or further restricts the solution space, which proves Lemma 8.5 as long as the triangle inequality holds. □

The third lemma concerns the state space of an i-DMVRP MDP model. More precisely, it says that, for any state $\widehat{s_t} = (\widehat{\mathscr{C}_t}, \phi_t)$, there exists a state $s_t = (\mathscr{C}_t, \phi_t)$, which only differs in that it does not include a certain customer order $\widehat{c}$.

Lemma 8.6. *For every state $\widehat{s_t} = (\widehat{\mathscr{C}_t}, \phi_t)$, there exists a corresponding state $s_t = (\mathscr{C}_t, \phi_t)$ with $\mathscr{C}_t = \widehat{\mathscr{C}_t} \setminus \{\widehat{c}\}$:*

$$\forall \widehat{s_t} \text{ with } t \in 1, ..., T : \exists s_t : \mathscr{C}_t = \widehat{\mathscr{C}_t} \setminus \{\widehat{c_t}\}. \tag{8.13}$$

To prove Lemma 8.6, it has to be shown that, assuming the rejection of a certain customer request $\widehat{c}$, the same future decisions can be made as assuming the acceptance of $\widehat{c}$, and that those decisions result in the same transitions.

Proof. By Lemma 8.5, the same future decisions can be made assuming the rejection of a customer request $\widehat{c}$ as by assuming its acceptance. Further, by Lemma 8.4, those decisions result in the same subsequent transitions. □

Now, a certain decision sequence $\pi = (a_t, a_{t+1}, a_{t+2}, ..., a_T)$ is considered and applied to a certain sample path ω, which starts in an interim state s'_t. Given Lemmata 8.4 to 8.6, two general lemmata regarding the resulting *revenue*, denoted by $R^{\pi\ \omega}_t(s'_t)$, and regarding the resulting *fulfillment cost*, denoted by $C^{\pi\ \omega}_t(s'_t)$, can be derived. More precisely, the first of these lemmata says that, applying π to ω, assuming it starts in the interim state $s'_t(c^i_t)$, results in the same cummulative revenues as it does, assuming it starts in the corresponding interim state $s'_t(0)$. The second of these lemmata states that, applying π to ω, assuming it starts in the interim state $s'_t(c^i_t)$, results in lower or equal fulfillment cost as it does, assuming it starts in the corresponding interim state $s'_t(0)$.

Lemma 8.7. *Applying decision sequence π to sample path ω, assuming it starts in interim state $s'_t(c^i_t)$ results in the same sum of cummulated revenues as it does, assuming it starts in the corresponding interim state $s'_t(0)$:*

$$R_t^{\pi\ \omega}(s'_t(c^i_t)) = R_t^{\pi\ \omega}(s'_t(0)) \tag{8.14}$$

Proof. From Lemmata 8.5 and 8.6 follows that any π that can be applied feasibly to ω starting in interim state $s'_t(c^i_t)$ can also feasible applied to ω starting in interim state $s'_t(0)$. Then, 8.7 directly follows from Lemma 8.4 since, starting from both interim states, the same set of customer orders are received, when the same decisions are made. □

Lemma 8.8. *Applying decision sequence π to sample path ω, assuming it starts in the interim state $s'_t(c^i_t)$, results in higher or equal fulfillment cost as it does, assuming it starts in the corresponding interim state $s'_t(0)$:*

$$\mid C_t^{\pi\ \omega}(s'_t(c^i_t)) \mid \geq \mid C_t^{\pi\ \omega}(s'_t(0)) \mid \tag{8.15}$$

Proof. π and ω start with the same set of confirmed customer orders $\mathscr{C}_{t-1}$, irrespective of whether their start is assumed from interim state $s'_t(c^i_t)$ or $s'_t(0)$. Then, assuming π and ω start in interim state $s'_t(c^i_t)$, customer order c_t is added to the set of confirmed customer orders $\mathscr{C}_{t+1}$ whereas it is not added assuming π and ω start in interim state $s'_t(0)$. Afterward, starting from interim states $s'_t(c^i_t)$ and $s'_t(0)$, respectively and applying the same decision sequence π to ω, which is possible by Lemmata 8.5 and 8.6, again, the same customer orders are confirmed, which results from Lemma 8.4. Consequently, starting π and ω in interim state $s'_t(c^i_t)$ results in subsequent states $\widehat{s_{t'}} = (\widehat{\mathscr{C}_{t'}}, \phi_{t'})$ and starting in interim state $s'_t(0)$ results in subsequent states $s_{t'} = (\mathscr{C}_{t'}, \phi_{t'})$ for $t' = t+1, ..., T$, with $\widehat{\mathscr{C}_{t'}} = \mathscr{C}_{t'} \cup \{c_t^i\}$. This proves Lemma 8.8 analogously to the proof of Lemma 8.5, as the underlying VRP is more restricted and consequently the resulting fulfillment cost cannot be smaller as long as the triangle inequality holds (Asdemir et al. (2009)). □

Combining Lemmata 8.7 and 8.8 shows that applying a decision sequence π to sample path ω starting in interim state $s_t'(c_t^i)$ cannot result in a higher objective value than starting in interim state $s_t'(0)$. This is formalized in the following Lemma:

Lemma 8.9. *If the same decision sequence π is applied to sample path ω starting in interim state $s_t'(c_t^i)$, it cannot yield a higher value than it does when starting in interim state $s_t'(0)$:*

$$\forall \omega \in \Omega : V^{\pi\ \omega}(s_t'(c_t^i)) \leq V^{\pi\ \omega}(s_t'(0)). \tag{8.16}$$

With this in mind, the third crucial opportunity cost property can be formally proven by contradiction.

Property 8.10. *Opportunity cost are non-negative:*

$$\Delta V(s_t, c_t, i) \in \mathbb{R}_0^+. \tag{8.17}$$

Proof. It is assumed that for a sample path ω, starting in an interim state $s_t'(c_t^i)$, an optimal sequence of decisions, denoted by π^*, is found which results in a value $V^{\pi^*\ \omega}(s_t'(c_t^i))$ higher than any value that can be accrued on the same sample path starting in interim state $s_t'(0)$. With Lemmata 8.4 to 8.6, π^* can be applied to the sample path starting in interim state $s_t'(0)$ and, with Lemma 8.9, results in at least the same value. The original assumption is proven wrong and hence:

$$V^{\omega}(s_t'(c_t^i)) = V^{\pi^*\ \omega}(s_t'(c_t^i)) \leq V^{\omega}(s_t'(0)) = V^{\pi^*\ \omega}(s_t'(0)). \tag{8.18}$$

This proof by contradiction can be replicated for every sample path $\omega \in \Omega$. Thus:

$$V(s_t'(c_t^i)) = V^{\pi^*}(s_t'(c_t^i)) \leq V(s_t'(0)) = V^{\pi^*}(s_t'(0)) \tag{8.19}$$

and substituted in equations (7.3) it proves that $\Delta V(s_t, c_t, i) \geq 0$. □

8.4 Monotonicity of value functions

For the presented DJP model and the modified OP model, Property 8.10 directly implies that the value function is monotonically decreasing in the number of confirmed but not yet being served customer orders $|\mathscr{C}_{s_t}|$, i.e., the following holds: $\Delta V(s_t, c_t, i) = V(s'_t(0)) - V(s'_t(c^i_t)) \geq 0 \Leftrightarrow V(s'_t(0)) \geq V(s'_t(c^i_t))$. Analogously, it can be shown that the value functions (6.4) and (6.29) are monotonically decreasing in time, i.e., in *consecutive* states s_t and $s_{t'}$, with $t' > t$. For two states, s_t and $s_{t'}$, to be consecutive, there must exist an optimal decision sequence π^* with resulting stochastic transitions, that causes the system to transition from state s_t to state $s_{t'}$ in a finite number of decision epochs. Then, it can be shown that $V(s_t) \geq V(s'_t) \geq V(s_{t'})$. Therefore, it has to be shown that $V(s_t) \geq V(s'_t)$ and $V(s'_t) \geq V(s_{t+1})$ hold, as the latter directly implies that $\forall t' > t : V(s'_t) \geq V(s_{t'})$.

Property 8.11. *The value function is monotonically decreasing in the course of consecutive states:*

$$\forall t' > t : V(s_t) \geq V(s_{t'}). \tag{8.20}$$

Proof. $V(s_t) \geq V(s'_t)$ directly follows from Equation 7.4 with the following intuition: starting from a certain state s_t in decision epoch t means that there is one more customer request c_t potentially contributing to the state value than compared to the interim state of the same decision epoch resulting from any demand-management decision s'_t. If all potentially arriving requests $c_t \in C$ are not profitable based on s_t, the optimal demand-management decision is the rejection in any case, and $V(s_t) = V(s'_t)$ holds, because no revenue can be collected in t. Otherwise, if there is at least one potentially arriving request c_t that is profitable, it is optimal to accept this request and the associated expected revenue positively contributes to $V(s_t)$. Then, $V(s_t) > V(s'_t)$ holds. $V(s'_t) \geq V(s_{t+1})$ directly follows from Equation (7.1), since for the considered value functions $\forall t \in 1, ..., T, i \in \mathscr{I} : r^I_{\phi^{*}_{t+1} i} = 0$ holds by definition of the underlying MDP formulations. Thus $\forall t \in 1, ..., T$, it holds that $V(s'_t) = V(s_{t+1})$. □

It has to be noted, that Property 8.11 is the only of the discussed properties which cannot be transferred to the natural OP model. Due to the cost modeling during the booking horizon, the respective value function (6.23) is not monotone in time. This can have a destabilizing effect for learning value function approximations which is why the modified OP model with value function (6.29) is introduced in Chapter 6.3.2.

This concludes Part IV of this dissertation in which two of the identified research gaps have been closed. Namely, a unified definition of opportunity cost for i-DMVRPs has been introduced and the value functions of the underlying i-DMVRP MDP models as well as the derived opportunity cost properties have been discussed analytically.

Part V

Development of a novel dynamic demand-management and online tour-planning approach for same-day delivery

In Part V of this dissertation, a newly developed, integrated approach to demand management and tour planning for an OP, more precisely for an SDD problem setting, is presented. As shown in Chapter 5, heuristic optimization approaches for DJPs have been broadly discussed in the literature. Also, optimization approaches that address dynamic tour planning in OP settings and, correspondingly, initial demand management approaches that aim on optimized fulfillment operations, separately, are already discussed in the literature. However, comprehensive approaches that optimize fulfillment operations with regard to anticipatory decision making and aim on levering the potential of active demand-management measures at the same time form a gap in i-DMVRP literature. One potential reason for that gap is the increased problem complexity of OPs compared with DJPs. As the booking horizon and the service horizon overlap, demand management for OPs is substantially more difficult to optimize than for the broadly investigated DJPs. This can be ascribed to the necessity of incorporating an online tour-planning component. More precisely, a decision dimension is added to the demand-management problem, which itself is computationally intractable. With OPs, the decision on which fulfillment options to offer at which prices must be made online. In addition, the decision on which orders to allocate to which tours and when to start them, has to be made online and under consideration of potential future decisions.

In practice however, the need to optimize demand management and tour planning for same-day delivery business models comprehensibly and in an anticipatory manner has been demonstrated. Despite increasing demand for SDD and the customers' willingness to pay higher delivery fees for faster delivery, a large number of SDD providers went out of service or shifted their service portfolio toward more profitable business branches. Meanwhile, new businesses that promise delivery of groceries and common every day goods proliferate in urban last-mile-delivery markets.

Therefore, a new heuristic solution approach for an SDD problem setting is developed. It integratively optimizes the SDD demand-management and tour-planning components, both under anticipation of future customer requests and decisions. The approach is developed to make the concept of SDD profitable and improve provider services and, thus, customer satisfaction. It is tailored to exploit two demand-management levers, namely reserving more capacity for higher valued customers, and steering the stochastic customer choice toward efficient fulfillment options. Simultaneously, it improves online tour planning. This is achieved, by combining ideas of multiple scenario approaches for online tour planning with the ideas of state value approximation via sampled trajectories, such as those known from rollout algorithms/Monte Carlo methods (see for example Sutton and Barto (2018), Chapter 5). Further, a novel hierarchical demand-management decomposition is incorporated.

The contributions of this part are the following:

(1) A comprehensive anticipatory solution approach to the i-DMVRP is proposed. It links a pricing optimization problem to an anticipatory sample-scenario based value approximation method, relying on an explicit online tour-planning heuristic. It does not require extensive offline learning, and guarantees applicability and scalability to realistically sized problem instances.

(2) A novel hierarchical demand-management decomposition for solving stochastic scenarios is introduced, in which demand management is decomposed into two steps. The first step is to perform tentative tour planning that will decide which customer requests to accept and by what time they can be served. The second step is to re-integrate the provider's actual demand-management decisions and the customers' corresponding choices into the first step's solution. This allows to amend the solution of a customized MIP tour-planning model by retrospectively integrating realistic customer choice behavior, and to derive accurate values of the solutions.

(3) In a comprehensive computational study, a potential of up to 50% of contribution margin increase by anticipation in decision making within the newly developed approach is demonstrated. However, due to the online tour-planning component that distinguishes demand management of SDD problems from other demand-management problems in last-mile delivery research, anticipation in decision making does not always lead to substantial profitability improvement. Such cases are elaborated, and thus a differentiated insight into the problem is given. Further, the results are compared with those of other demand-management and pricing approaches, adopt-

ing ideas from the existing literature. It is shown how explicit price optimization increases the profit margin compared with these benchmarks and discussed how the different approaches affect the solution structure.

This part of the dissertation is organized as follows. In Chapter 9 the general solution framework and the idea behind the hierarchical demand-management decomposition are outlined. Then, in Chapter 10, the solution approach is presented with its technical details. Subsequently, in Chapter 11, the solution approach is evaluated comprehensively in a broad numerical study. Further, the corresponding results and insights that emerged from them are summarized and future research directions are pointed out.

9. General solution framework and hierarchical demand-management decomposition

In this chapter, the problem setting under consideration is introduced, the general solution framework (Chapter 9.2) is presented, and the newly developed hierarchical demand-management decomposition (Chapter 9.3) approach is outlined. The latter is one of the core ideas underlying the solution approach. As will be shown later, the approach relies on a sample-scenario value approximation and tour-planning approach, of which a detailed technical description is given in Chapter 10.

9.1 Problem setting

The proposed approach interacts with the SDD booking and service horizon as follows: a customer logs in to the website with information about their location and delivery preferences stored in the profile, chooses a shopping basket online while expecting a selection of narrow delivery time spans to be offered at affordable prices. This initiates a delivery request in response to which the provider has to make a demand-management decision. Therefore, simultaneously to the customer's login, the provider samples different customer request trajectories and conducts tentative tour-planning optimization, as well as value approximation with the help of the newly developed hierarchical demand-management decomposition. From the corresponding solutions, the provider derives anticipatory decisions on which delivery time spans to offer the current customer and at what prices. The customer then chooses one of the options offered or leaves the website without purchasing, following their own individual, stochastic preferences. If the customer decides to purchase, the delivery request becomes a confirmed customer order with a delivery deadline and the tour planning is updated on the basis of the previously sampled trajectories. To enable prompt delivery, the execution of deliveries might start/continue immediately, even though further customer requests could arrive. When a new customer request arrives, the whole process starts over again.

9.2 General solution framework

The proposed solution approach is based on the following consideration: if the optimal tour-planning decision and the associated value of the successor state, more precisely

$$V'(s^i_{t+1}) := \max_{\phi^i_{t+1}\in\Phi^i_{t+1}} (r^I_{\phi^i_{t+1}} + V(s_{t+1} \mid s_t, \phi^i_{t+1})) \tag{9.1}$$

were known, solving the value function (6.23) would be simplified tremendously. This is because with known $V'(s^i_{t+1})$ and if $|\mathscr{G}|$ is not large, Equation (6.23) could even be solved to optimality by total enumeration across all $g \in \mathscr{G}$ (Yang et al. (2016)). In the same-day delivery demand-management and tour-planning problem (SDD-DMTP), it is indeed assumed that $|\mathscr{G}|$ is of tractable size. However, $V'(s^i_{t+1})$ cannot be determined exactly. Thus, a problem-specific approximation of $V'(s^i_{t+1})$ is proposed. More specifically, every time a customer request arrives and the demand-management problem of Equation (6.23) has to be solved, a procedure to approximate $V'(s^i_{t+1})$ is carried out. This procedure simultaneously returns heuristic tour-planning decisions. The proposed approximation approach is a forward ADP approach (c.f. Powell et al. (2012), Chapter 4). It is based on a heuristic solution of the SDD-DMTP on sampled realizations of customer requests. Thus, it is referred to as *sample-scenario value approximation and tour-planning* approach. The full procedure is depicted schematically in Figure 9.1.

Every time a customer request arrives, the sample-scenario based approximation of $V'(s^i_{t+1})$ is carried out.

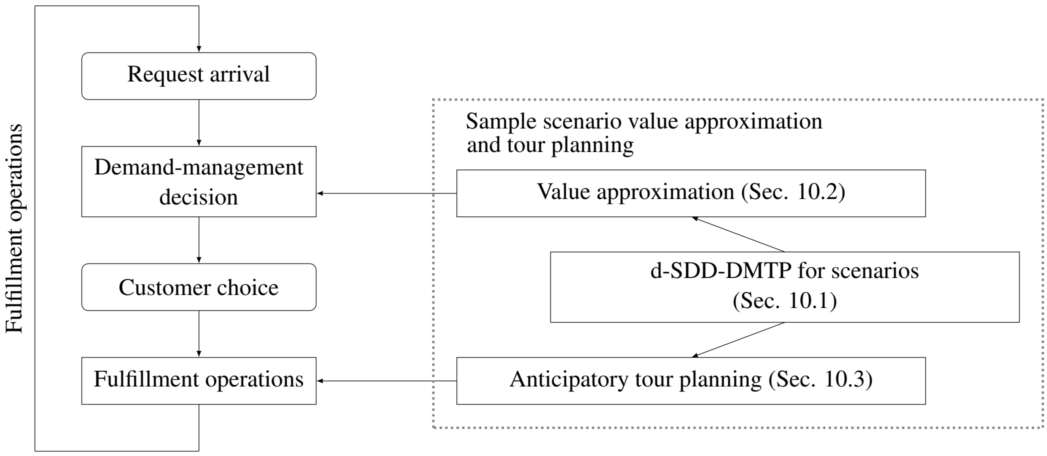


Figure 9.1: General solution framework

Therefore, different customer request realizations are sampled into the future to generate scenarios. A scenario $\omega \in \Omega_t^i$ at time t and for a certain fulfillment option i consists of three types of customers: first, the *confirmed and not yet being delivered* customer orders $\mathscr{C}_t$; second, the *current* customer request c_t with assigned deadline according to i, i.e., $t_{c_t}^{due} = t + l(i)$, if $i \neq 0$; third, a *sampled* realization of customer requests N^{ω}, sampled from t on until the end of a predefined sampling horizon length. Consequently, the scenarios are state-, time-, and customer-choice-specific. For all scenarios, a deterministic version of the SDD-DMTP (d-SDD-DMTP) is solved (see Chapter 10.1) which returns state-specific *scenario tour plans* that are then used for the following two purposes:

(1) *Value approximation*: Based on the scenario tour plans, a value for each scenario is approximated. Then, the average across all state-specific scenario values is used as the approximation of the corresponding state's value (see Chapter 10.2). This in turn, is the input to the provider's demand-management decision. All relevant successor state values are substituted in equation (6.23), and the provider can enumerate across all $g \in \mathscr{G}$ to decide which offer set to present to the current customer. Presenting an offer set then triggers a customer choice i' which the provider can observe.

(2) *Anticipatory tour planning*: According to the resulting customer choice i', the

provider finally has to make certain tour-planning decisions $\phi_{t+1}^{i'}$, which again, are constructed from the scenario tour plans (see Chapter 10.3).

It has to be noted that the tour-planning decisions are based on predictions into the future. This means they consider later customer requests and also time-steps in between future customer arrivals. Accordingly, it is only necessary to revise tour plans if a new customer request arrives. Only then, new explicit information about tour-optimization potential becomes known. Therefore, the respective delivery operations are executed until a new customer request arrives; only then will the whole decision-making procedure begin anew. Therefore, unlike the MDP model of OPs, the solution approach is not defined across all decision epochs t in the booking horizon. Instead, it is event driven, i.e., driven by customer request arrivals.

Regarding the literature discussed in Section 5.3, the solution approach falls in the class of non-learning approaches. For decision making, it uses an information model predictively (Soeffker et al. (2021)) and is conducted fully online.

9.3 Hierarchical demand-management decomposition

The sample-scenario value approximation and tour-planning approach adapts the online tour-planning ideas of Bent and Van Hentenryck (2004) as well as Voccia et al. (2019) and substantially extends them in order to include demand-management decisions. The basic idea is to solve the d-SDD-DMTP for different scenarios, and then to average the values of the scenario solutions to derive state values. Theoretically, due to the assumption of all customer orders to be known within an ex-post solution, the averaged scenario values have a tendency to overestimate the actual state value. However, as explained in more detail in Chapter 10.2, this has no major impact on the decision.

To manage the particular challenges that result from integrating demand management, a hierarchical demand-management decomposition is proposed. This procedure heuristically divides the demand management into two subsequent tiers, namely *accepting/rejecting customer requests* (first tier) and explicitly steering customer choices by *selecting the offer sets* to present (second tier):

(1) *Acceptance/rejection*: When the d-SDD-DMTP for a scenario ω is solved, in the optimization it is first assumed that a decision on accepting/rejecting a customer requests can be taken directly. This allows to formulate the first-tier problem as a deterministic, p-MTVRP. It is formalized as an MIP in Chapter 10.1.1 and can be solved by a tailored heuristic as presented in Chapter 10.1.2. This results in an anticipatory scenario solution, i.e., in a *scenario tour plan* $\phi^{\omega}(s^i_{t+1})$. It comprises tours that are already running, currently starting, or will start at any time from the decision moment on until the sample-horizon's end. Note that, a scenario tour plan can now comprise more than one tour per vehicle. The tours include all customer orders from $\mathscr{C}_t$ in state s_t and the current customer request c_t with a choice i'-specific delivery deadline. Further, they include an optimized subset of the sampled customer requests that are scheduled for delivery within the longest predefined fulfillment option.

(2) *Offer set selection*: From the first-tier solutions, it is possible to construct tour-planning decisions for the SDD-DMTP. However, it is not yet possible to derive a precise value approximation as input for demand-management decisions. This is due to neglecting the choice and no-choice probabilities of the sampled customer requests and their resulting delivery revenues. Consequently, in the second-tier demand management, these aspects are retrospectively captured and integrated into the scenario solutions. Therefore, the main target is to reconstruct demand-management decisions for a scenario's sampled customers in such a way that the same scenario tour plan results as in the first-tier problem's solution. More precisely, for every sampled customer request of a scenario solution, it has to be determined which offer sets provoke purchase choices with which the corresponding scenario solution is feasible. Across those offer sets, the expected contribution for every sampled customer is maximized and a scenario value $\tilde{V}^{\omega}$ can be determined. This procedure is described in Chapter 10.2.

10. Sample-scenario value approximation and tour planning with two-tier demand-management integration

In this chapter, a detailed presentation of the three components of the sample-scenario value approximation and tour-planning approach is given.

10.1 Deterministic SDD demand-management and tour-planning problem with first-tier demand management

At first, a MIP formulation for the d-SDD-DMTP is presented in Chapter 10.1.1 and then it is shown how to solve it heuristically in Chapter 10.1.2.

10.1.1 MIP formulation

The d-SDD-DMTP with first-tier demand management is a deterministic profitable multi-trip vehicle routing problem with release and due times (p-VRPRDT). It is defined across nodes for the already confirmed and not yet being delivered customer orders, a node representing the current customer request, and nodes for all sampled customers. Additionally, one node c_0 that represents a centrally located depot with coordinates $(x,y)_{c_0} = (0,0)$, is needed. Thus, the corresponding set of nodes $\mathscr{N}^{\omega}$ equals the the following union: $\mathscr{C}_t \cup \{c_t\} \cup N^{\omega} \cup \{c_0\}$. For every confirmed customer order and for the current customer request, this set contains information about the customer's location $(x,y)_c$ and their confirmed delivery deadline t_c^{due}. For every sampled customer request, the set contains information about the customer's location $(x,y)_c$, their request time t_c^{req}, the revenue of their requested shopping-basket r_c, and their utility u_c^i for fulfillment options $i \in \mathscr{I}$. Further, all sampled customers $c \in N^\omega$ are assigned a preliminary delivery deadline according to the longest available delivery span, $t_c^{due} = t_c^{req} + max\{l(i) \mid i \in \mathscr{I}\}$. The underlying idea is that if those customers are included in a scenario's solution, it is always possible to offer them at least one, namely the longest, fulfillment option when their request realizes. This ensures that it is always possible to feasibly reconstruct a first-tier solution with the second-tier demand management (see Chapter 10.1.2 and 10.2).

In the d-SDD-DMTP, V homogeneous vehicles operate chronologically ordered tours $k \in \mathscr{K} = 1,...,K$. $\zeta_{cc'}$ represents the costs of travelling from the location of customer order c to the location of customer order c'. $\rho_{c'}$ is a customer order individual penalty which equals the value of the shopping basket for all $c' \in N^\omega$ and equals a sufficiently high number M for all $c' \in \mathscr{C}_t \cup \{c_t\}$. Since $\mathscr{C}_t$ only contains customer orders for which a feasible solution (without delays and dropped visits) is available, these penalties ensure that no confirmed customer order is dropped when solving the model. The parameter t describes the current decision period. The following decision variables are included in the model:

$$x_{cc'}^{vk} = \begin{cases} 1 & \text{if customer } c' \text{ is served after customer } c \text{ on tour } k \text{ by vehicle } v \\ 0 & \text{else} \end{cases} \qquad \forall c,\, c' \in \mathscr{N}^\omega : c \neq c',\, v \in \mathscr{V},\, k \in \mathscr{K}$$

$$a_{c'}^{vk} \geq t \quad \forall c' \in \mathscr{N}^\omega,\, k \in \mathscr{K},\, v \in \mathscr{V} \qquad \text{Delivery time at customer location } c' \text{ on tour } k \text{ of vehicle } v$$

$$A^{vk} \geq t \quad \forall k \in \mathscr{K},\, v \in \mathscr{V} \qquad \text{Erliest departure time of tour } k \text{ of vehicle } v$$

$$B^{vk} \geq t \quad \forall k \in \mathscr{K},\, v \in \mathscr{V} \qquad \text{Time of finishing tour } k \text{ of vehicle } v \text{ in the depot}$$

The d-SDD-DMTP can be formulated as the following MIP, which is further explained below:

$$\min \sum_{v\in\mathscr{V}}\sum_{k\in\mathscr{K}}\sum_{c\in\mathscr{N}^\omega}\sum_{c'\in\mathscr{N}^\omega} x^{vk}_{cc'}\cdot\zeta_{cc'} + \sum_{c'\in\mathscr{N}^\omega\setminus\{0\}}(1-\sum_{v\in\mathscr{V}}\sum_{k\in\mathscr{K}}\sum_{c\in\mathscr{N}^\omega} x^{vk}_{cc'})\cdot\rho_{c'} \quad (10.1)$$

s.t.

$$\sum_{v\in\mathscr{V}}\sum_{k\in\mathscr{K}}\sum_{c\in\mathscr{N}^\omega} x^{vk}_{cc'} \leq 1 \qquad \forall c'\in\mathscr{N}^\omega\setminus\{c_0\} \quad (10.2)$$

$$A^{vk} \leq a^{vk}_0 \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K} \quad (10.3)$$

$$t^{req}_{c'}\cdot\sum_{c\in\mathscr{N}^\omega} x^{vk}_{cc'} \leq A^{vk} \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K},\, c'\in\mathscr{N}^\omega\setminus\{c_0\} \quad (10.4)$$

$$t^{due}_{c'} + (1-\sum_{c\in\mathscr{N}^\omega} x^{vk}_{cc'})\cdot M \geq a^{vk}_{c'} \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K},\, c'\in\mathscr{N}^\omega\setminus\{c_0\} \quad (10.5)$$

$$a^{vk}_{c'} + (1-x^{vk}_{cc'})\cdot M \geq a^{vk}_c + x^{vk}_{cc'}\cdot\tau_{cc'} \qquad \forall v\in\mathscr{V}, k\in\mathscr{K}, c\in\mathscr{N}^\omega,\ c'\in\mathscr{N}^\omega\setminus\{c_0\} \quad (10.6)$$

$$A^{vk} + \sum_{c\in\mathscr{N}^\omega}\sum_{c'\in\mathscr{N}^\omega\setminus\{c_0\}} x^{vk}_{cc'}\cdot\tau_{cc'} \leq B^{vk} \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K} \quad (10.7)$$

$$B^{vk} \leq A^{vk+1} \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K}\setminus\{K\} \quad (10.8)$$

$$\sum_{c\in\mathscr{N}^\omega} x^{vk}_{cc'} = \sum_{c\in\mathscr{N}^\omega} x^{vk}_{c'c} \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K},\, c'\in\mathscr{N}^\omega\setminus\{c_0\} \quad (10.9)$$

$$\sum_{c'\in\mathscr{N}^\omega\setminus\{c_0\}}\sum_{v\in\mathscr{V}} x^{v0}_{0c'} \leq V \quad (10.10)$$

$$\sum_{c'\in\mathscr{N}^\omega\setminus\{c_0\}}\sum_{k\in\mathscr{K}} x^{vk}_{0c'} \leq K \qquad \forall v\in\mathscr{V} \quad (10.11)$$

$$\sum_{c'\in\mathscr{N}^\omega} x^{vk}_{0c'} \leq 1 \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K} \quad (10.12)$$

$$\sum_{c'\in\mathscr{N}^\omega\setminus\{0\}} x^{v+1\ 0}_{c_0c'} \leq \sum_{c'\in\mathscr{N}^\omega\setminus\{c_0\}} x^{v0}_{0c'} \qquad \forall v\in\mathscr{V}\setminus\{V\} \quad (10.13)$$

$$\sum_{c'\in\mathscr{N}^\omega\setminus\{0\}} x^{vk+1}_{0c'} \leq \sum_{c'\in\mathscr{N}^\omega\setminus\{c_0\}} x^{vk}_{0c'} \qquad \forall v\in\mathscr{V},\, k\in\mathscr{K}\setminus\{K\} \quad (10.14)$$

The objective function (10.1) minimizes the overall travel costs and the sum of the penalties of all dropped visits. Dropping a sampled customer in the solution of the MIP means rejecting their request. Therefore, (10.1) balances the increase in travel costs for visiting a sampled customer and their shopping basket value – if marginal costs to serve and displacement costs are higher than a customer's shopping basket value, this customer request is rejected. Constraints (10.2) enable dropping visits/rejection of customer requests. Thus, in combination with the objective function, this represents the first-tier demand management. Constraints (10.3)-(10.8) are time restrictions, which ensure that a tour starts neither before t, nor before all allocated customer orders have realized, that all customer orders will be served on time, that the duration of a tour is the sum of all travel times of that tour, and that a vehicle can only start a new tour after having returned to the depot. Constraints (10.9) ensure flow conservation. Constraints (10.10) - (10.12) ensure that the number of available vehicles and the maximum number of tours are not exceeded. Constraints (10.13) - (10.14) are symmetry breaking constraints.

This MIP formulation is a generalization of a p-MTVRP, which additionally considers time restrictions. It is an adaption of the p-MTVRP formulation of Chbichib et al. (2012) and of a multi-trip team orienteering problem with time windows formulation by Voccia et al. (2019). Further, it is closely related to a multi-trip vehicle routing problem with time windows and release dates introduced by Cattaruzza et al. (2016) who do not state a formal MIP.

10.1.2 Heuristic solution approach

p-MTVRPs belong to the class of NP-hard problems (Chbichib et al. (2012)). Thus, the MIP given in Chapter 10.1.1 cannot be solved for all sampled scenarios in reasonable time. Instead, a heuristic approach which consists of the following three steps is proposed:

Relaxation – First, the explicit consideration of depot returns is relaxed in the d-SDD-DMTP. The resulting problem is a p-VRPTW (Toth and Vigo (2014), Chapter 1). A customer request's arrival time now forms the start of their delivery time window, while the delivery deadline remains unchanged. The trick is that all vehicles can now start only one tour, but can serve customer requests that have not yet realized at the time the tour starts.

Solving the relaxed problem – Next, the resulting p-VRPTW is solved heuristically by means of a standard tour-planning software (e.g. Google OR Tools, `https://developers.google.com/optimization`). The result is a tour plan with one tour per vehicle, including confirmed and sampled customer orders.

Feasibility repair – When the tours start, not all sampled customer orders have already realized, which is why depot returns have to be added to the tours. Thus, for feasibility, the respective tours are repaired as follows. To generate feasible tour plans, a vehicle's tour is interrupted for a depot return each time a sampled customer order has to be served whose request had not yet arrived when the tour started in the depot. For this vehicle, a new tour is planned to serve the original tour's remaining customers in the same order, until it has to be interrupted for another depot return. If a depot return causes a late delivery for a sampled customer, the customer is removed from the tour; yet, if the depot return causes a late delivery for a confirmed customer, the latest sampled customer is removed from the tour and, according to vehicle availability, the departure time is updated to an earlier time. This procedure is repeated until all late deliveries have been removed. If the algorithm does not find a feasible solution without late deliveries, at the end of the algorithm, an empty scenario tour plan ϕ^{ω} and a scenario value $\tilde{V}^{\omega}(s^{i}_{t+1}) = -\infty$ is returned. For the original decision problem (6.23), this results in not offering the corresponding fulfillment option i to the current customer request c_t. It has to be noted that, since a tour plan can now comprise more than just one tour per vehicle, an index k is added to the tour notation, i.e., θ^{vk} denotes the k^{th} tour of vehicle v.

The procedure is more formally presented in the following Algorithm 1.

Algorithm 1 Feasibility repair scenarios

$\phi \leftarrow$ Heuristic solution of P-VRPTW
$\theta^{v}(\phi) \leftarrow$ Tour of vehicle v according to solution ϕ
$a_c^v \leftarrow$ Delivery time of customer c with vehicle v according to solution ϕ
for v in $\mathscr{V}$ **do**
    Initialize first tour θ^{v1} by adding depot $c = 0$ and customer order c with smallest a_c^v according to $\theta^{v}(\phi)$
    Calculate current departure time: $A^{v\ next} \leftarrow a_c^v - \tau_{0c}$
    $\theta^{v\ next} \leftarrow \theta^{v1}$
    $\phi^{\ v} \leftarrow \theta^{v\ next}$
    repeat
        $\theta^{v\ curr} \leftarrow \theta^{v\ next}$
        $A^{v\ curr} \leftarrow A^{v\ next}$
        $L(\theta^{v\ curr}) \leftarrow \{\}$
        $X(\theta^{v\ curr}) \leftarrow \{\}$
        Add not yet planned customer orders c of $\theta^{v}(\phi)$ to $L(\theta^{v\ curr})$ with increasing a_c^v until $t_c^{req} \geq A^{v\ curr}$ and amend $X(\theta^{v\ curr})$ accordingly
        Cut tour by adding depot return and calculate return time $B^{v\ prev}$
        $\theta^{v\ curr} \leftarrow (A^{v\ curr}, L(\theta^{v\ curr}), X(\theta^{v\ curr}))$
        Append $\theta^{v\ curr}$ to ϕ^{v}
        Initialize next tour $\theta^{v\ next}$ by adding depot $c = 0$
        Add not yet planned customer order c with next smallest a_c^v according to $\theta^{v}(\phi)$
        Calculate latest departure time: $A^{v\ next} \leftarrow max\{a_c^v - \tau_{0c}, B^{v\ prev}\}$
    until all customer orders c in $\theta^{v}(\phi)$ are planned to tours
    for *tour* in ϕ^{v} **do**
        Update all $a_c^{v\ tour}$ according to $A^{v\ tour}$ and travel times $\tau_{cc'}$
        if $a_c^{v\ tour} \geq t_c^{due}$ for any sampled customer order c in $\theta^{v\ tour}$ **then**
            Remove c from $\theta^{v\ tour}$ and update all left $a_c^{v\ tour}$ according to $A^{v\ tour}$ and travel times $\tau_{cc'}$
        if $a_c^{v\ tour} \geq t_c^{due}$ for any confirmed customer order c in $\theta^{v\ tour}$ **then**
            repeat
                Remove sampled customer order c with highest t_c^{req}
                Update $A^{v\ tour}$ according to vehicle availability
                Update all left $a_c^{v\ tour}$ according to $A^{v\ tour}$ and travel times $\tau_{cc'}$
            until there are no longer any late deliveries
$\phi \leftarrow \{\phi^{v} : v \in \mathscr{V}\}$

10.2 Value approximation and second-tier demand management

The heuristic presented in Chapter 10.1.2 is used to solve the d-SDD-DMTP for scenarios $\omega \in \Omega_t^i$. In this way, scenario-specific tour plans ϕ^ω are generated. Those are anticipatory in the sense that they anticipate future customer requests, but only under the consideration of the first-tier demand management. Thus, as explained in Chapter 9.3, before determining the scenario value $\tilde{V}^\omega(s_{t+1}^i)$, the second-tier demand management has to be re-integrated into the solution. Thereby, in order to derive the best possible estimate, it is the target to imitate the original demand management of the SDD-DMTP as the value function (6.23) solved it as closely as possible. Imagining solving Equation (6.23) by hand: in a first step the feasible action space would intuitively be defined by excluding all infeasible decisions from the consideration. For the demand-management decision this means determining which fulfillment options can be feasibly offered to the current customer, i.e., defining the set of feasible offer sets. In a next step, the offer set with the highest expected sum of immediate reward and successor state value is offered to the requesting customer. This last step includes making tour-planning decisions.

In re-integrating the second-tier demand management into a scenario's solution ϕ^ω, this procedure is mimicked with two modifications:

> *Modification 1* – When identifying the offer sets for the accepted, sampled customer requests $c \in N^\omega \cap \{L(\theta^{vk}) : \theta^{vk} \in \phi^\omega\}$ that are feasible with respect to the first-tier demand-management solution, all tour-planning decisions have already been determined. Thus, the specific delivery times for customer orders a_c^{vk} are already defined. Consequently, fulfillment options are only feasible, if a_c^{vk} can be matched within the fulfillment option.
>
> *Modification 2* – When selecting which offer set to offer, only the expectation regarding the immediate rewards is considered. DPC and MCTS can be neglected.

The second modification can be made without sacrificing accuracy because the scenario solution, i.e., the acceptance and delivery times of all requesting customers in the scenario under consideration, has already been decided. Thus, it does not matter whether the currently considered customer chooses one of the offered fulfillment options or the no-purchase option. The value that might be incurred with subsequent customer requests will not change for this scenario. Consequently, in the second-tier demand management, neither DPC nor MCTS are relevant for optimizing offer set selection. Other approaches that approximate values/costs via heuristically solving scenarios ex-post in order to derive tour-planning decisions are for example Azi et al. (2012), Campbell and Savelsbergh (2005), and Angelelli et al. (2021).

More formally, re-integrating second-tier demand management into a solution ϕ^{ω} can be described as follows. For every $c \in N^{\omega} \cap \{L(\theta^{vk}) : \theta^{vk} \in \phi^{\omega}\}$, the procedure determines which fulfillment options $i \in \mathscr{I}$ can feasibly be offered according to their planned delivery time a_c^{vk} when following ϕ^{ω}. Next, for each of those customers, a subset $\mathscr{G}_c'(\phi^{\omega}) \subset \mathscr{G}$ defines all offer sets that include only the valid fulfillment options i. To approximate the sampled customer's contribution $r_{c\ \phi^{\omega}}$ to a scenario's value $\tilde{V}^{\omega}(s_{t+1}^i)$, the expected reward across all $g \in \mathscr{G}_c'(\phi^{\omega})$ is maximized:

$$r_{c\ \phi^{\omega}} = \underset{g \in \mathscr{G}_c'(\phi^{\omega})}{\arg\max} \sum_{i \in g} P^i(g) \cdot (r_c^i + r^i), \tag{10.15}$$

if a customer order c is being accepted in the first-tier solution, otherwise $r_{c\ \phi^{\omega}} = 0$. A scenario's value is then defined as $\tilde{V}^{\omega}(s_{t+1}^i) = \sum_{c \in N^{\omega}} r_{c\ \phi^{\omega}} - r_{\phi^{\omega}}^l$. Following this, $V'(s_{t+1}^i)$ is approximated by

$$\widehat{V}'(s_{t+1}^i) = \frac{\sum_{\omega \in \Omega_t^i} \tilde{V}^{\omega}(s_{t+1}^i)}{\mid \Omega_t^i \mid}. \tag{10.16}$$

Finally, the SDD-DMTP's demand-management decision is taken by substituting (10.16) in the value function. That yields the following demand-management decision policy for when a customer request arrives:

$$g^* = \arg\max_{g \in \mathcal{G}} \left(\sum_{i \in g} P^i(g) \cdot [r^i(g) + r^i_{c_t} + \widehat{V}'(s^i_{t+1})] \right). \tag{10.17}$$

It has to be noted that the value approximation described above relies on solving scenarios ex-post, under the assumption that all customer arrivals were known. This could lead to a systematic over-estimation of the actual value of a state. However, for deciding on which offer set to present to an incoming customer, this over-estimation is not a major issue for the reason that when solving Equation 10.17, not the absolute level of the values $\widehat{V}'(s^i_{t+1})$ for $i \in g$ is decision-relevant, but the differences between them, i.e., the opportunity cost of the fulfillment options. As the potential over-estimation is systematic, it applies similarly to all those values.

10.3 Anticipatory tour planning

Having described how to approximate values to derive demand-management decisions based on scenario tour plans, now it is explained how to take tour-planning decisions. For every potential customer choice i and the corresponding successor state s^i_{t+1}, a set of scenario tour plans $\phi^\omega \in \Omega^i_t$ with values $\tilde{V}^\omega(s^i_{t+1})$ is available from the scenarios' solutions. This can be used to derive tour-planning decisions. Typically, in multiple-scenario approaches, at this point, a consensus function measures the robustness of partials of those tour-plans and then constructs a robust overall tour plan, called a distinguished plan (Bent and Van Hentenryck (2004), Voccia et al. (2019)). Due to the large number of stochastic influences in the SDD problem setting under consideration, i.e., customer location, request arrival time, and customer choice, the scenario solutions exhibit high variability. This is why typical consensus functions proved not to perform well in pre-tests. Therefore, tour-planning decisions are derived from the sampled tour plan ϕ^ω that has the highest value $\tilde{V}^\omega(s^{i'}_{t+1})$ of all the tour plans in $\Omega^{i'}_t$, while fully aware of and accepting that the derived tour plan's performance might naturally be lower in entirely different realizations. It is selected as distinguished plan ϕ^* and comprises planned tours θ^{vk} for all $v \in \mathscr{V}$. The tours of one vehicle v still contain sampled and confirmed customer orders. Thus, in line with the literature on multiple-scenario approaches, all sampled customer orders are removed from those tours and the delivery times a_c^{vk} of all remaining confirmed customer orders c, as well as the return times to the depot, are updated according to A^{vk} and relevant $\tau_{cc'}$. Finally, the tours of one vehicle v start sequentially at given start times A^{vk}. This procedure is more formally described in Algorithm 2.

An executable tour at state s_{t+1}^i, derived from the tour-planning decision ϕ^* for vehicle $v \in \mathscr{V}$ is denoted as θ^{*vk}. All θ^{*vk} for $v \in \mathscr{V}$, $k \in \mathscr{K}$ of ϕ^* form the tour-planning decision ϕ_{t+1} in t and for all subsequent t' until a new customer request arrives. If a new customer request arrives, the full decision-making procedure as presented in Figure 9.1 starts all over again. For the tour-planning decisions that means all tours in ϕ_{t+1} that have not already started by the time of the new customer request, can be revised.

Algorithm 2 Tour-planning decision and post-processing

```
i ← Customer choice
s^i_{t+1} ← Regarding successor state
φ* ← argmax_{φ^ω : ω ∈ Ω^i_t} Ṽ^ω(s^i_{t+1})

for v ∈ 𝒱 do
    for k ∈ 𝒦 do
        Remove all sampled customers c ∈ θ^{vk}
        for remaining customers c ∈ θ^k do
            Update a^{vk}_c according to A^{vk} and travel and service times τ_{cc'}
        θ^{*vk} ← θ^{vk}
φ* ← {θ^{*vk} : v ∈ 𝒱, k ∈ 𝒦}
φ_{t+1} ← φ*
```

11. Computational study

In this chapter, a computational study on a variety of parameter settings is presented in which the proposed solution approach is applied in different variants, e.g., with different lengths of the sampling horizon. Additionally, benchmark approaches are presented and the respective results are compared with each other. In particular, the effectiveness of the proposed approach is assessed and the value of anticipation, as well as that of an explicit price optimization are evaluated. In Section 11.1, first, the parameters of the settings under consideration are described and it is explained how instances are generated. In Sections 11.2 and 11.3, the extensive computational experiments' results on the value of anticipation and of explicit price optimization are discussed.

11.1 Setup

The computational study is based on a number of different *settings* that are examined in a stochastic simulation, by applying and comparing different anticipation and pricing approaches. In Section 11.1.1, the parameters that are commonly used throughout all considered settings are specified. In Section 11.1.2, the parameters that may vary across settings are introduced. In Section 11.1.3, it is described how instances are generated for each setting within the stochastic simulation.

11.1.1 Setting-independent parameters

The following parameters are defined identically for all settings considered in the computational study.

Time horizon and fulfillment options

The considered time horizon corresponds to the booking and service course of one day. It is represented by 900 episodes, which could be thought of as representing 900 minutes from 7am to 10pm. The booking horizon consists of 600 minutes, i.e., it starts at 7am and ends at 5pm. Thus, $T = 600$. The service horizon starts with the first accepted customer order and ends at 10pm, latest. In all settings, offer sets can be generated based on two possible fulfillment options, i.e., delivery within 90 minutes or within 300 minutes.

Customer segments

Customers are defined by a segment affiliation, their location, their arrival times, and arrival rates. A customer's segment affiliation defines the potential contribution margins of selected shopping baskets. More precisely, it indicates a probability distribution across the potential contribution margins in connection with a purchase decision. Further, it defines their utility for different fulfillment options with different prices. In the computational study, it is assumed that there are two segments, distinguishing between *segment-one* customers and *segment-two* customers. The contribution margin of a segment-one customer is drawn from a discrete uniform distribution over $[75,85,100]$ monetary units (MU). The contribution margin of a segment-two customer is drawn from a discrete uniform distribution over $[20,35,40]$ MU. Additionally, segment-one customers have a higher observable utility for shorter fulfillment options than segment-two customers. The basic observable utilities before pricing u^i_{basic} of segment-one customers are 22 and 14, and those of segment-two customers are 13 and 10.5 for the short and the long fulfillment options. To calculate the observable utility for a fulfillment option with a certain price u^i, the corresponding basic utility u^i_{basic} is reduced by the offered price r^i, but it cannot be negative, i.e., $u^i = max\{u^i_{basic} - r^i, 0\}$. Also, the no-purchase option has a utility for customers from both segments. For segment-one customers, this utility equals 2, while for segment-two customers it equals 3. This reflects that segment-two customers are more likely to purchase via a traditional, non-SDD fulfillment option or in a brick-and-mortar store.

The purchase probabilities for different fulfillment options within the offer sets are modeled according to a basic attraction model (Gallego et al. (2019), Chapter 4). Therefore, the purchase probabilities for fulfillment options i in an offer set $g \in \mathcal{G}$ can be calculated by solving

$$P^i(g) = \frac{u^i}{\sum_{i \in g} u^i}. \tag{11.1}$$

Service area and customer locations

The service area is simulated on a squared grid with a width of 120 distance units (DU), with a centrally located depot. On this grid, 200 customer locations are generated in advance, drawn from a discrete uniform distribution. These locations will be used later on in instance generation. Travelling one DU equals one minute in the simulation run and costs 0.3 MU. Thus, all potential customer locations on this grid can be visited within 120 minutes. Thus, if vehicle capacity allows, every customer can at least be offered the longest fulfillment option.

Arrival rates

The customer arrival process is modeled with customer segment s-specific time dependent arrival rates λ_t^s (Lebedev et al. 2021). Further, two peaks in the arrival rates are assumed which show common online shopping behavior, namely customers placing orders during their lunch break or after returning home from work. For the lower valued segment-two customers, lower and wider peaks are assumed than for the higher valued segment-one customers in order to reflect more flexible working conditions with lower income. The distribution of arrival rates is illustratively depicted in Figure 11.1.

Pricing approach

For the dynamic pricing component, two price points per fulfillment option are assumed: 8 or 10 MU for guaranteed delivery within 90 minutes and 5 or 7 MU for guaranteed delivery within 300 minutes. Further, it is assumed that no fulfillment option other than the no-purchase option needs to be offered. These pricing parameters

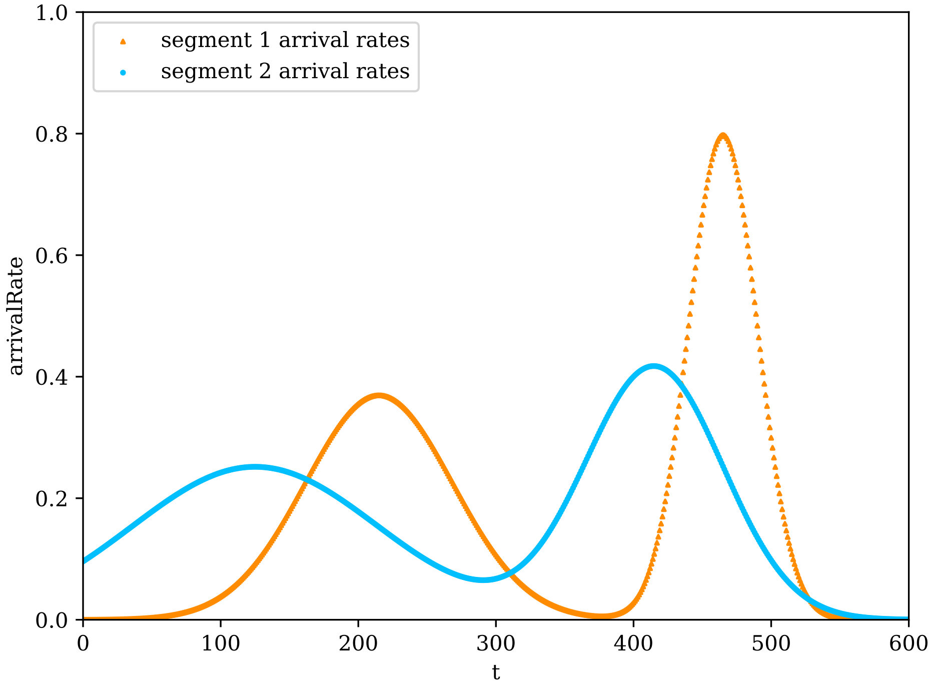


Figure 11.1: Arrival rates

result in nine potential price lists from which the provider can select one to offer to an incoming request.

11.1.2 Setting-dependent parameters

The considered settings differ in terms of the expected number of incoming customer requests and the number of delivery vehicles. More specifically, settings resulting from each possible combination of 100, 150, and 200 expected customer requests with one, two, and three delivery vehicles are considered. The corresponding settings are shown in Table 11.1, which states the settings' names.

		customer requests		
		100	150	200
	1	1V_100	1V_150	1V_200
vehicles	2	2V_100	2V_150	2V_200
	3	3V_100	3V_150	3V_200

Table 11.1: Setting-dependent parameters

11.1.3 Instance generation

To ensure comparability, the proposed approach and the benchmark approaches are tested on the same set of registered customers, which is referred to as the *customer base*. More precisely, based on the customer segments' and customer locations' characteristics described in Section 11.1.2, a customer base of 3000 different customer requests is generated initially. 30% of these customer requests are segment-one customer requests. Then, for each setting, instances represent particular demand streams that are obtained by event-based discrete simulation based on the arrival rates and according to the setting's expected number of customer requests. Customer requests' characteristics are obtained by sampling from the customer base. 300 instances are generated for each setting. It has to be noted that, again to ensure comparability, the same 300 instances are used for settings that differ only in the number of delivery vehicles.

11.2 Value of anticipation

In the following, the value of anticipation for the SDD-DMTP is discussed with respect to the developed approach as presented in Chapters 9 and 10.

11.2.1 Experimental design and performance metrics

In studying the impact of different levels of anticipation, different variants of the proposed approach are applied. They differ as to the length of the sample horizon used for approximating the scenario values and route planning (see Section 9.2).

Different sample horizon lengths of 30, 60, 90, and 120 minutes are evaluated. Here, the decision making is based on the anticipation of a total of 15 scenarios, a number that led to good decisions in performed pre-tests. Scenarios are sampled by drawing new customer requests from the customer base each time a decision has to be taken. Further, the anticipatory approach is benchmarked against myopic decision making. Myopic decisions are taken in exactly the same way as in the anticipatory approach, except that all potential successor state values in Equation (6.23) are set to 0. Further, the tour-planning decisions are taken without anticipated customer requests. Thus, in this approach the demand-management decision is based only on myopic MCTS of a request.

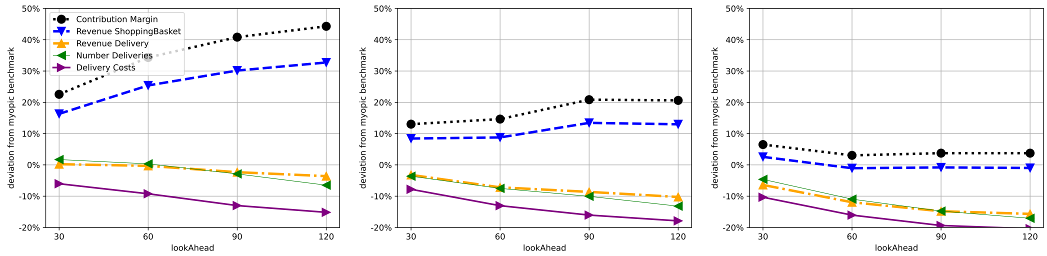
To measure performance, for each setting and each length of the sample horizon, the deviation from the myopic benchmark is evaluated with respect to the following metrics:

Metric		Description
Revenue shopping Baskets	(RSB)	sum of contribution margins of all shopping baskets sold in one instance
Revenue Deliveries	(RD)	sum of delivery fees accrued by selling fulfill-ment options throughout one instance
Delivery Costs	(DC)	overall cost of delivery operations, i.e., all exe-cuted delivery tours in one instance
Contribution Margin	(CM)	RSB + RD - DC
Number Of Deliveries	(NOD)	number of accepted customer requests that turned into orders and are being served in the course of one instance

The deviation of a given metric from the myopic benchmark for a given setting with a given sample horizon length is determined as follows: the results of the metric across the 300 test instances of the setting under consideration is averaged and compared with the corresponding averaged values resulting from solving the same 300 instances with the myopic benchmark approach. For example, the deviation of the CM with a sample horizon length of 30 minutes from the myopic results is calculated by $\frac{CM^{30}}{CM^{myopic}} - 1$.

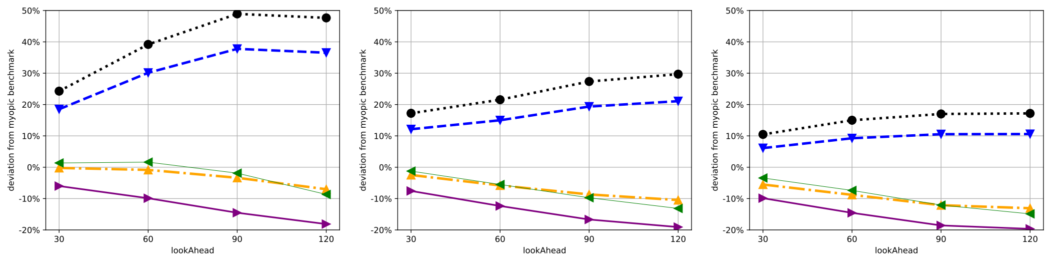
11.2.2 Numerical results

The obtained results are shown in Figure 11.2. On the tested settings, it is possible



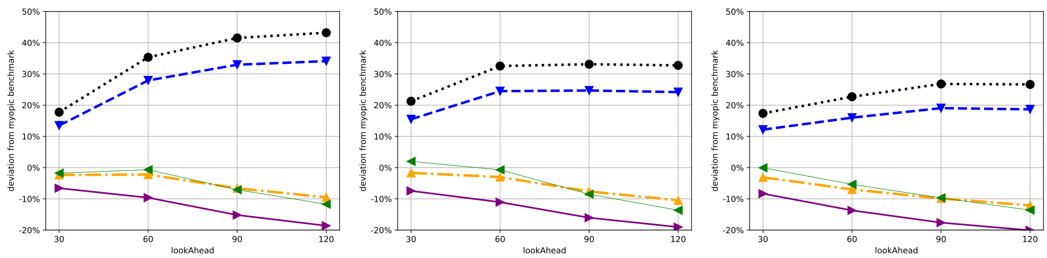
(a1) 1 vehicle **(a2)** 2 vehicles **(a3)** 3 vehicles

(a) 100 customers



(b1) 1 vehicle **(b2)** 2 vehicles **(b3)** 3 vehicles

(b) 150 customers

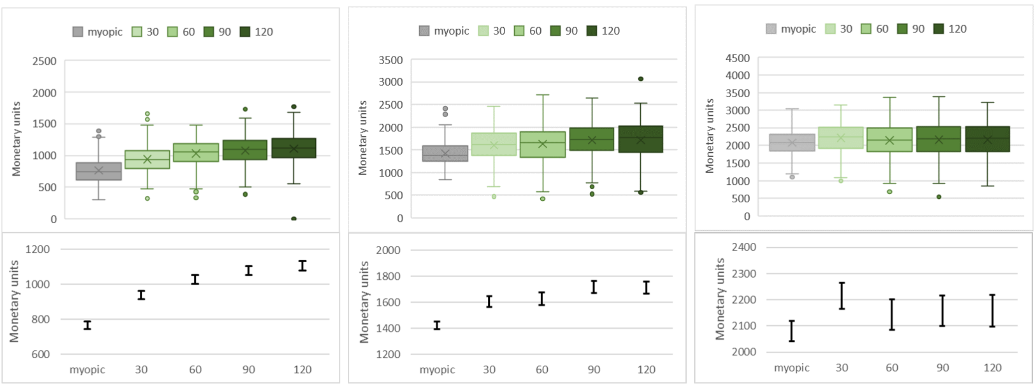


(c1) 1 vehicle **(c2)** 2 vehicles **(c3)** 3 vehicles

(c) 200 customers

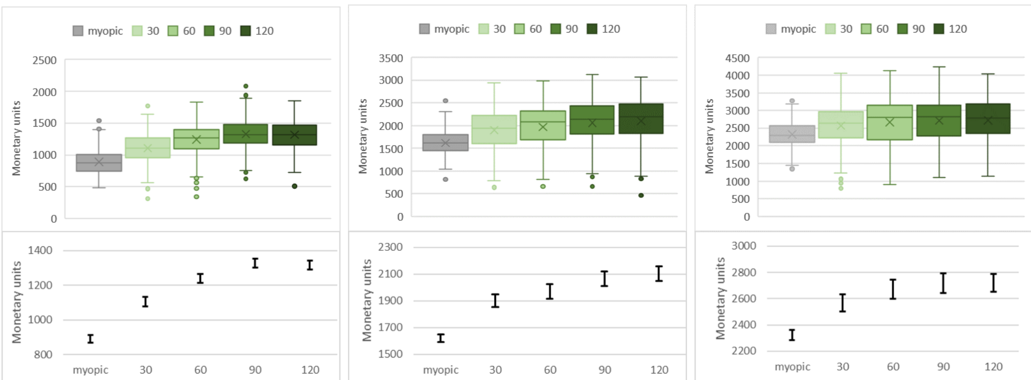
Figure 11.2: Value of anticipation as deviation from myopic benchmark of CM, RSB, RD, ND, and DC by different look-ahead horizons (30, 60, 90, 120), averaged over 300 simulation runs

to achieve an increase in CM of 15% to 50%. First, the increase grows degressively as the sample horizon length increases, until it reaches a peak at a sample horizon length of 90 or 120 minutes for most settings. It then slowly decreases for longer sample horizon lengths, which is displayed in more detail in Figure 11.4), where the absolute values of the mean CM across all 300 instances, as well as the corresponding 95%-confidence intervals are depicted. For almost all settings, these intervals of the myopic approach and the anticipatory approaches do not overlap.



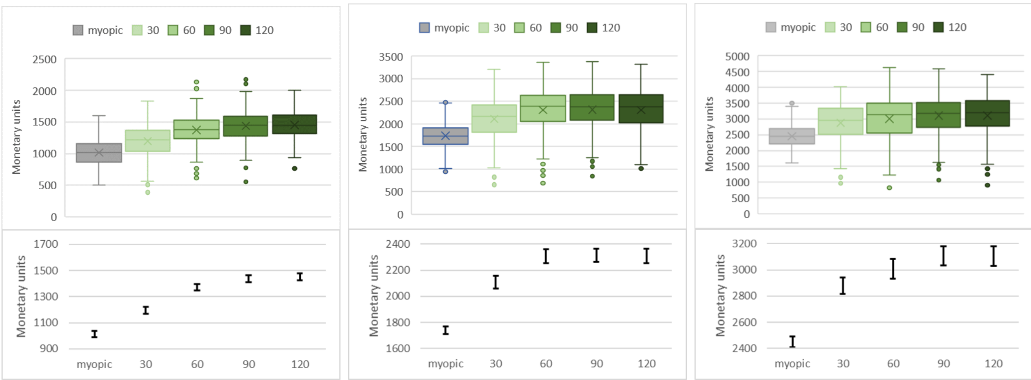
(a1) 1 vehicle **(a2)** 2 vehicles **(a3)** 3 vehicles

(a) 100 customers



(b1) 1 vehicle **(b2)** 2 vehicles **(b3)** 3 vehicles

(b) 150 customers



(c1) 1 vehicle

(c2) 2 vehicles

(c3) 3 vehicles

(c) 200 customers

Figure 11.4: Mean contribution margins across 300 simulation runs: boxplots and 95%-confidence intervals of the myopic approach, and the anticipatory approach with look-ahead horizons 30, 60, 90, and 120

For the respective settings, this implies with a confidence of 95%, that the increase in CM results from the anticipation approach. The only setting in which the increase in CM is smaller than 10% and where 95%-confidence intervals overlap, is the setting with low resource scarcity, in which the myopic approach also yields good results. Regarding the degressive course of the CM increase with increasing sample horizon length, pre-tests have shown that using fewer samples flattens the growth and shifts the peak to a shorter sample horizon length. This is shown illustratively for setting 1V_100 in Figure 11.5. Increasing the sample size does not substantially shift the peak to a longer sample horizon length.

		look-ahead			
	myopic	30	60	90	120
# segment 1 customers	5.01	7.14	8.71	9.85	10.67
# segment 2 customers	14.16	12.36	10.53	8.77	7.27
# 90 minutes choice	1.25	1.60	1.44	1.21	1.67
# 300 minute choice	17.92	17.90	17.80	17.41	16.26
average price 90 minutes	8.00	8.49	8.55	8.69	8.78
average price 300 minutes	5.00	5.36	5.45	5.52	5.53

Table 11.2: Results for setting 1V_100 and different look-ahead periods, averaged over 300 simulation runs

Table 11.2 shows further numerical results for the 1V_100 setting, namely the

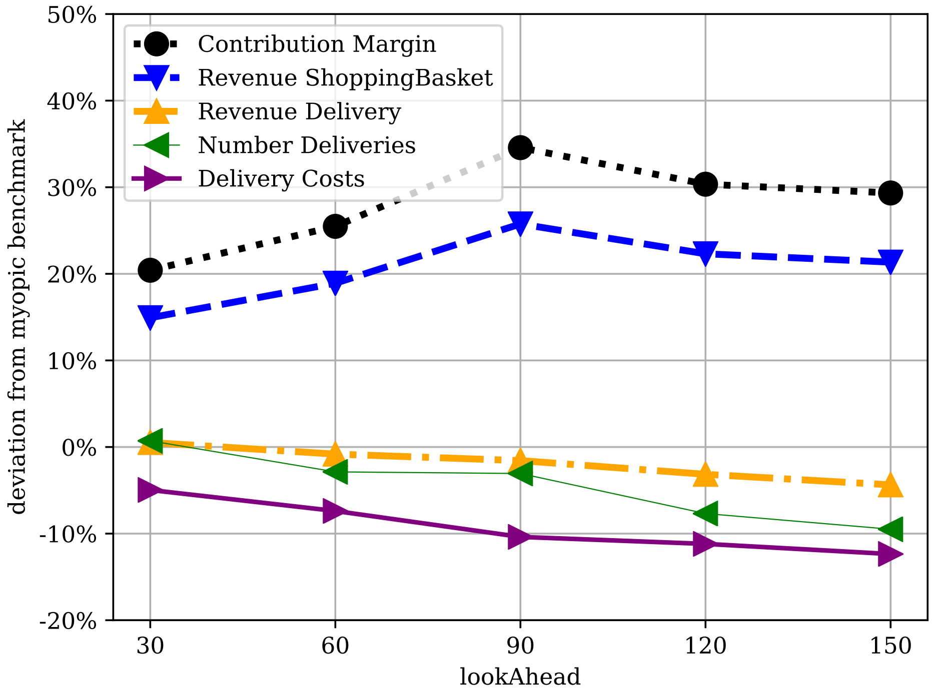


Figure 11.5: Value of anticipation by different look-ahead horizons for setting 1V_100, 5 samples, averaged over 300 simulation runs

average absolute values of customer choices, the segments of customer orders, and the average prices paid for fulfillment options per instance. Here, it can be observed that, as the length of the sampling horizon increases, the average number of highly valued customer orders accepted in an instance increases, and correspondingly, the average number of low-value customer orders accepted, decreases. Another trend observed as the length of the sample horizon increases, is the increase in the average prices paid for the delivery spans. The average number of customer decisions for the different delivery intervals shows no obvious pattern. All of these observations are representative of the results in the other instances, as can be seen in Tables 11.3 to 11.11.

11.2.3 Analysis and insights

According to the previously described observations, the contribution margin that can be achieved with anticipation is always higher than the contribution margin of the myopic benchmark. This is mainly due to the fact that the revenues generated by selling shopping baskets increase and the delivery costs decrease disproportionately to the decrease in delivery orders. Combined with Tables 11.2 to 11.11, this shows that anticipation indeed allows preserving capacity for high-value customer orders, and also to generally steer customer choice with respect to a favorable spatial structure. Thus, compared with myopic decision making, through anticipation delivery efficiency can be improved.

Further, a degression in the increase of contribution margin with an increasing sample horizon length can be observed. Such degression is explained by the lengths of the sample horizon becoming longer, and as this happens, the proportion of uncertainty in decision making increases. Thus, these results indicate that the solutions' quality decreases if the sample horizon is too long or if too few samples are used. This can be ascribed to the following observation: for every decision, increasing the sampling horizon length also increases the number of sampled, and hence uncertain requests, while the number of certain orders does not increase. Additionally, due to the tight delivery spans that distinguish SDD from other LMD services, all certain orders in the scenarios will be served shortly after the time when the sampling starts. Hence, sampling into the future too far leads to decision making based on tours that include only uncertain orders. This distorts the precision of the value approximation.

look-ahead	myopic	30	60	90	120
# segment 1 customers	5.01	7.14	8.71	9.85	10.67
# segment 2 customers	14.16	12.36	10.53	8.77	7.27
# 90 minutes choice	1.25	1.60	1.44	1.21	1.67
# 300 minute choice	17.92	17.90	17.80	17.41	16.26
average price 90 minutes	8.00	8.49	8.55	8.69	8.78
average price 300 minutes	5.00	5.36	5.45	5.52	5.53

Table 11.3: 1 vehicle – 100 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	9.93	12.76	13.65	15.47	16.13
# segment 2 customers	25.19	21.08	18.82	16.12	14.36
# 90minutes choice	3.33	4.13	3.35	2.75	2.95
# 300 minute choice	31.78	29.71	29.13	28.84	27.53
average price 90 minutes	8.00	8.31	8.34	8.42	8.50
average price 300 minutes	5.00	5.24	5.31	5.34	5.37

Table 11.4: 2 vehicle – 100 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	15.61	17.68	17.98	19.05	19.61
# segment 2 customers	33.94	29.58	26.14	23.18	21.49
# 90minutes choice	6.11	6.93	5.16	3.50	3.94
# 300 minute choice	43.44	40.32	38.96	38.74	37.16
average price 90 minutes	8.00	8.32	8.31	8.40	8.44
average price 300 minutes	5.00	5.17	5.21	5.22	5.27

Table 11.5: 3 vehicle – 100 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	5.69	8.63	10.54	12.39	13.06
# segment 2 customers	15.82	13.17	11.33	8.72	6.59
# 90minutes choice	1.06	1.52	1.49	1.34	1.56
# 300 minute choice	20.45	20.28	20.38	19.77	18.08
average price 90 minutes	8.00	8.43	8.58	9.00	8.90
average price 300 minutes	5.00	5.39	5.55	5.61	5.61

Table 11.6: 1 vehicle – 150 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	10.82	14.73	16.49	18.82	20.24
# segment 2 customers	28.55	24.15	20.72	16.72	13.95
# 90minutes choice	3.21	4.17	3.63	2.82	3.17
# 300 minute choice	36.16	34.71	33.57	32.72	31.02
average price 90 minutes	8.00	8.36	8.43	8.59	8.54
average price 300 minutes	5.00	5.31	5.40	5.46	5.46

Table 11.7: 2 vehicle – 150 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	15.87	19.45	22.07	24.26	25.23
# segment 2 customers	40.17	34.67	29.83	25.02	22.47
# 90minutes choice	5.84	6.92	5.22	3.65	4.36
# 300 minute choice	50.20	47.19	46.68	45.62	43.34
average price 90 minutes	8.00	8.29	8.38	8.42	8.48
average price 300 minutes	5.00	5.23	5.29	5.32	5.34

Table 11.8: 3 vehicle – 150 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	6.46	9.15	11.74	13.58	14.58
# segment 2 customers	17.25	14.14	11.82	8.46	6.35
# 90minutes choice	1.02	1.58	1.69	1.14	1.60
# 300 minute choice	22.70	21.72	21.87	20.89	19.34
average price 90 minutes	8.00	8.53	8.78	9.11	9.00
average price 300 minutes	5.00	5.43	5.61	5.68	5.70

Table 11.9: 1 vehicle – 200 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	11.34	15.85	19.48	21.37	22.64
# segment 2 customers	30.47	26.78	22.02	16.87	13.46
# 90minutes choice	3.11	4.05	3.69	2.69	3.23
# 300 minute choice	38.70	38.58	37.82	35.56	32.87
average price 90 minutes	8.00	8.36	8.47	8.62	8.65
average price 300 minutes	5.00	5.37	5.49	5.54	5.53

Table 11.10: 2 vehicle – 200 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	15.73	21.32	24.79	27.91	29.12
# segment 2 customers	44.19	38.56	31.93	26.21	22.64
# 90minutes choice	5.91	6.84	5.45	4.16	4.54
# 300 minute choice	54.01	53.04	51.27	49.96	47.22
average price 90 minutes	8.00	8.33	8.40	8.50	8.53
average price 300 minutes	5.00	5.28	5.34	5.40	5.41

Table 11.11: 3 vehicle – 200 customers

11.3 Value of explicit pricing optimization

Here, the value of the explicit pricing approach as described in Section 11.1.1 is elaborated and compared with two benchmark pricing approaches.

11.3.1 Experimental design

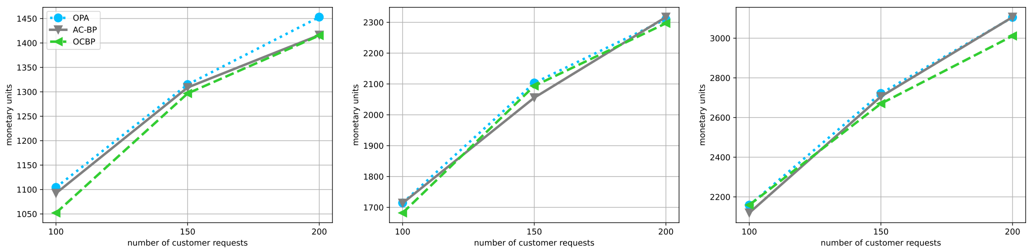
To determine the value of (explicitly) using a pricing optimization model within the proposed approach, two benchmark variants are elaborated. The first pricing benchmark reflects pure availability control, in which the provider can only decide whether to offer certain fulfillment options or not. Thereby, all prices are set to the corresponding lower prices from the explicit pricing approach described in Section 11.1.1. The second pricing benchmark replaces solving an explicit pricing optimization problem in the developed approach by a simple pricing rule based on opportunity cost estimation, which mimicks an idea followed by Ulmer (2020a). If a fulfillment option's calculated opportunity costs are low, its base price is set the lower price point used in the explicit pricing optimization. If the opportunity cost of an option exceed this base price, the price is set to the opportunity costs. For calculating those, Ulmer (2020a) follows a definition by Yang et al. (2016).

They define opportunity cost as the difference between the values of the states that result from rejecting a customer and those from accepting the customer (for a certain fulfillment option). In the conducted benchmark study, this definition is also followed and opportunity cost are calculated accordingly, based on state values resulting from the newly developed approximation approach (see Section 10.2). The first benchmark is referred to as 'AC-BP' (for 'availability control with base prices') and the second as 'OCBP' (for 'opportunity cost based pricing'). Further, the developed explicit pricing approach is referred to as 'OPA' (for 'original pricing approach').

The study is conducted on the same 300 instances for each setting as in Section 11.2.1. State values, and therewith opportunity cost, are approximated by averaging the values of 15 samples across a sample horizon of 120 minutes length. Based on the analysis in Section 11.2.2, this has proven to be the best combination for the considered settings. In this way, the effects of bad opportunity cost estimation by sub-optimal sampling horizon lengths/number of samples is minimized. Again, performance is measured by evaluating the average of the contribution margins, the number of accepted customer orders, the sum of revenues from shopping baskets and from selling fulfillment options, as well as of the delivery costs.

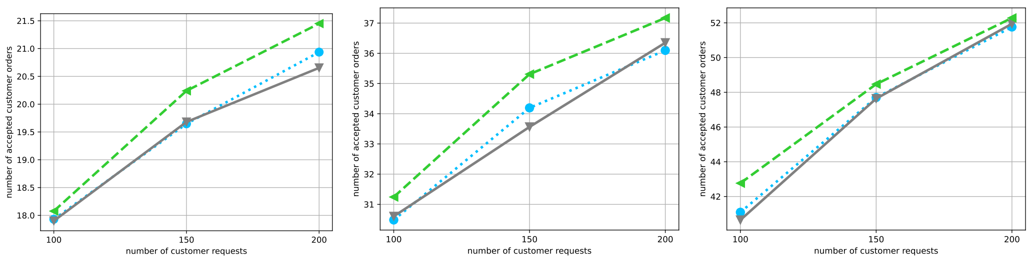
11.3.2 Numerical results

The obtained results are given in Figure 11.7. Although the results of the average contribution margins are close, the OPA yields better results than the benchmark approaches in nearly all settings. The AC-BP only yields a higher averaged CM in the settings with 200 customers, with two as well as with three vehicles; however, the results of the OPA are exceeded by less than 0.5% and 0.005%, respectively. In the setting with 100 customers and three vehicles, the OCBP yields a less than 0.05% higher CM than the OPA (see Figure 11.6a). The OCBP, on average, accepts the most customer requests of all settings (see Figure 11.6b), but in most settings its average RSB falls below the other approaches' RSB. Also, it yields a substantially higher DC for all settings and yields the highest RD in only three settings, where it



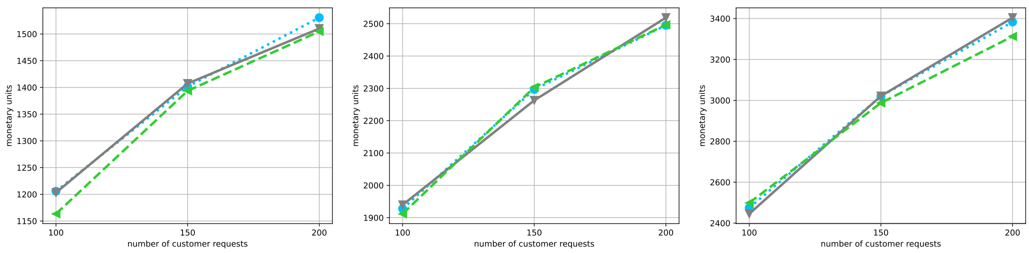
(a1) 1 vehicle **(a2)** 2 vehicles **(a3)** 3 vehicles

(a) Average contribution margin



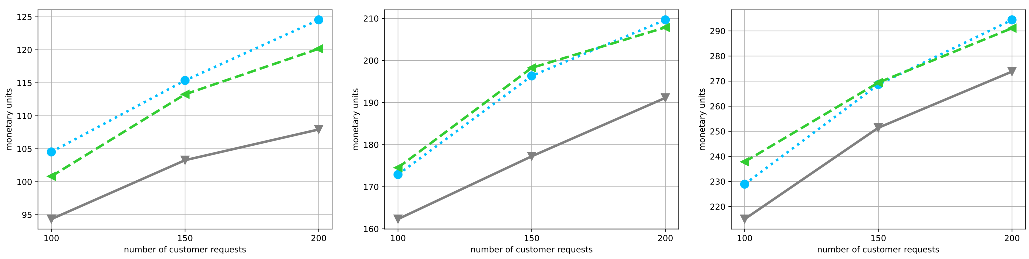
(b1) 1 vehicle **(b2)** 2 vehicles **(b3)** 3 vehicles

(b) Average number of accepted customer orders



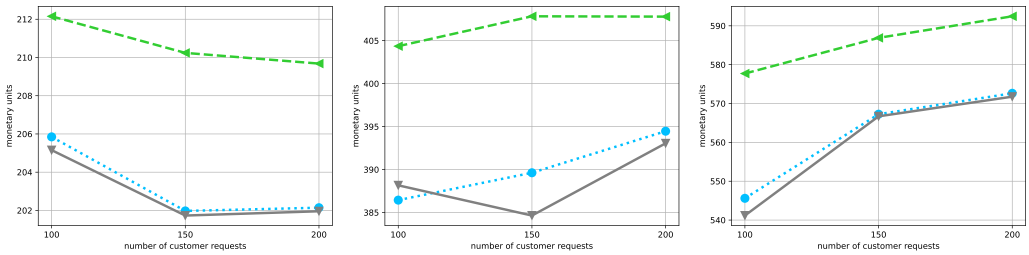
(c1) 1 vehicle **(c2)** 2 vehicles **(c3)** 3 vehicles

(c) Average sum of revenue from shopping baskets



(d1) 1 vehicle **(d2)** 2 vehicles **(d3)** 3 vehicles

(d) Average sum of revenue from selling fulfillment options



(e1) 1 vehicle **(e2)** 2 vehicles **(e3)** 3 vehicles

(e) Average delivery costs

Figure 11.7: Pricing benchmark in monetary units over number of customer requests 100, 150, and 200 - results averaged over 300 simulation runs

does not substantially exceed the RD of the OPA. In most instances the AC-BP accepts the lowest number of customer requests, also with substantially lower RD than the other approaches, but it still accrues a comparably high RSB. It even exceeds the other approaches' RSB in four settings. Also, the AC-BP yields the lowest DC of all instances except one.

11.3.3 Analysis and insights

The results in Section 11.3.2 show that the different pricing approaches rely on three different levers to increase the contribution margin, and that each of the various approaches exploits those levers to a different extent. The levers observed are *increasing the overall revenue by setting higher prices* (L1) if possible (mainly observed for the OCBP), *increasing the overall revenue by preserving capacity for high-value customer orders* (L2) (mainly observed for the BP and AC-BP), and *reduce overall delivery costs by steering customer choices* (L3) toward the most efficient fulfillment options and rejecting those requests that negatively affect routing efficiency (mainly observed for the AC-BP and OPA). Table 11.12 summarizes the exploitation of the different optimization levers by the evaluated solution approaches.

The OCBP has the highest pricing flexibility, as prices originate from a continuous range instead of being chosen from a predefined, finite set of price points. Therefore, this approach can exploit L1 the most (see Figure 11.6d) and hence can also accept the most customer requests. Still, regarding the CM, for most settings the OCBP

		Levers		
		L1 Pricing flexibility	L2 Preserving capacity	L3 Reducing delivery cost
Solution approach	OPA	✓	✓	✓
	AC-BP	✗	✓	✓
	OCBP	✓	✗	✗

Table 11.12: Exploitation of optimization levers by solution approaches

performs worse than the other approaches due to exploiting L2 and L3 less effectively. This can be derived from the lower or under-proportionally higher RSB (see Figure 11.6c), and from the over-proportionally higher DC (11.7e).

In contrast, the AC-BP has the lowest pricing flexibility and cannot exploit L1 as the much lower RD (11.6d) shows. Still, the AC-BP is a performs well regarding the exploitation of L2 and L3. The same can be observed for the OPA. Further, it can be recognized that the OPA also exploits L1. In addition to exploiting L1, the OPA enables enlarging the provider's service provision, as the OPA can offer delivery of customer requests that the AC-BP would deny and the customers can themselves decide whether to accept or reject the corresponding offer. This helps to improve customer goodwill and long-term customer loyalty.

Part VI

Conclusion

Despite the increasing practical relevance of i-DMVRPs and the vast body of literature addressing respective problems, a large number of research gaps concerning practically relevant questions still existed, e.g., a general taxonomy, a unified modeling framework for i-DMVRPs. Therefore, existing analytical discussions, as well as existing solution approaches could not be transferred between different types of i-DMVRPs.

After introducing the problem and its relevance in Part I and laying a theoretical foundation in Part II, this dissertation contributes to the research on i-DMVRPs as follows: first, by comprehensively analyzing the broad body of related literature in Part III and, based on this, by deriving a highly explicit but still sufficiently general problem definition that includes a mutual taxonomy for further classification. Moreover, this part contributes by introducing a modeling framework, which is then deeply analyzed analytically in Part IV. Additionally, a specifically tailored solution approach that combines anticipation in both components, i.e., demand management and tour planning, in an integrative manner is developed and presented in Part V. In the following, for each Part III to V, a conclusion, managerial insights, and an outlook for future research directions are given.

Part III: Integrated demand management and vehicle routing problems

In Part III of this dissertation, i-DMVRPs were formally defined and different types of i-DMVRPs were delineated following a newly introduced, uniform taxonomy that was derived from practical applications as well as existing research. The related literature was discussed with regard to different perspectives and with a special focus on modeling, the definition of opportunity cost, and solution approaches. Therewith, substantial research gaps were identified and elaborated. Those are the lack of:

- an explicit but unified MDP model for i-DMVRPs with disjoint and overlapping booking and service horizons,
- a unified definition of opportunity cost for i-DMVRPs,
- an analytical discussion on i-DMVRP models that also address OPs,

- and a solution approach for OPs that integrates anticipation in demand management and tour planning at the same time.

Then, a unified modeling framework was introduced, in which an MDP model for OPs is the generalization of a respective model for DJPs. Additionally, a modified model was introduced for OPs in order to preserve a monotone value function. Mathematical proof shows the equivalency of this modified model and the original model with regard to the objective function value. Therewith, different types of i-DMVRPs can be modeled within the proposed framework, and the modeling approaches are transferable to a wide range of underlying problem settings. Further, the same problem setting can be modeled by different but equivalent approaches such that it is possible to choose the model which better suites a given solution approach.

Part IV: Analytical discussion of opportunity cost for i-DMVRPs

In Part IV of this dissertation, the MDP models introduced in Chapter 6 were investigated analytically. Thereby, it was shown, that the traditional interpretation of opportunity cost cannot be transferred to i-DMVRPs, and therefore, the respective definition of opportunity cost was generalized. Further, central opportunity cost properties and the monotonicity of the introduced value functions were investigated.

The following conclusions and managerial insights can be derived: first, the discussion shows that neglecting variable fulfillment cost in the estimation of opportunity cost can lead to sub-optimal decisions. Thus, approximating solely DPC to derive demand-management decisions is only valid for some specific problem settings and can lead to expensive suboptimal decisions. Thereby, generally, the following relationship can be derived: (1) The tighter the physical vehicle capacity, the more important is the influence of DPC in decision making. (2) The more spatially dispersed the demand, the more important is the influence of MCTS in decision making. This has to be taken into account when selecting or developing a solution approach for an i-DMVRP setting, as not all solution approaches can approximate DPC and MCTS at the same time. Further, it has to be taken into account that the approximation of DPC and MCTS can be variably complex for different problem settings and that, at the same time, the required approximation precision can deviate. A complex demand-management setting, e.g., tends to require more precision in the approximation of opportunity cost than a simple demand-management setting.

Second, from the observation that DPC and MCTS both can be negative, the following insights can be derived and exploited for heuristic decision making:

> *Negative DPC* – If the respective customer request is accepted, further customer requests from its vincinity are expected and should also be accepted if they realize.
>
> *Negative MCTS* – If the respective customer request is accepted, fulfillment cost can be saved in that the acceptance causes other customer requests to be displaced, which come from locations that are less profitable with regard to fulfillment cost.

Third, the exploitation of the investigated opportunity cost properties and the monotonicity of the value functions can improve solution approaches substantially (c.f. Koch (2017), Adelman (2007)).

Further, the previous introduction and analysis of opportunity cost for i-DMVRPs is a starting point for deeper analyses of the impact intensity and scope of the monotonicity of value functions and the discussed opportunity cost properties on different solution approaches. Specific matters of interest for future studies are for example the impact of the monotonicity of the value function on the stability of learning-based solution approaches or the respective impact of the delay of the routing rewards in the modified OP model. A different, promising research direction is the investigation of how to preserve the previously described properties for heuristic MDP models that base on state space aggregation or incorporate heuristic tour-planning approaches.

Part V: Development of a novel dynamic demand-management and online tour-planning approach for same-day delivery

In Part V of this dissertation, the SDD-DMTP was investigated, with special attention to explicitly incorporating two types of required decisions, namely demand-management decisions and tour-planning decisions. The problem under consideration is characterized by overlapping booking and service horizons. This adds an online tour-planning component to the demand-management problem, which itself is computationally intractable. Thus, it makes the overall problem substantially more difficult to optimize than the related DJPs, to which a variety of solution approaches is dealt with in the literature as discussed in Section 5.2.

A non-learning based solution approach has been developed which provides integrated decision making for the two types of decisions and does not require extensive offline learning. In this approach, both decisions are anticipatory and based on the combination of two central ideas – multiple scenario approaches for online tour-planning and approximation of state values – which is done by averaging across sampled trajectories, such as those known from rollout algorithms. Further, a hierarchical demand-management decomposition has been developed.

Moreover, an extensive numerical study was conducted, that also shows the superiority of the newly developed approach, first, with regard to incorporating anticipation, and second, with regard to different pricing benchmarks, in two parts:

In the first part of the study, the performance of the approach regarding different levels of anticipation was assessed. The assessment demonstrated that anticipation can increase the contribution margin with up to 10-50% in the settings under consideration, especially if delivery resources are scarce. By incorporating anticipation through sampling, it was found that appropriately limiting the length of the sample-horizon can substantially improve decision making. The main reason for this is that as the length of the sample horizon increases, decisions are made with increasing uncertainty. This is especially relevant for practical settings where booking and service horizons overlap, as in the SDD case. If the sampling horizon is too long, anticipatory decisions are based on tours that contain only sampled orders and no confirmed ones.

In the second part of the study, three pricing approaches have been compared: pure availability control, the proposed explicit pricing approach, and a simple pricing rule based on opportunity cost. Comparing the different approaches with each other, it was revealed that as price flexibility increases (from fixed prices to a limited number of possible price points to possible prices from an unbounded continuous space), the quality of the resulting tours decreases. This demonstrates that the integrated state value approximation and decision-making approach does indeed allow to steer customer choice toward efficient fulfillment options, while at the same time preserving capacity for high-value customer orders. Compared with the other two approaches, this one resulted in the best ratios of number of customer requests accepted to corresponding sum of revenues from shopping baskets, and delivery efficiency. Further, it was shown that the approach that accepts the most customer requests is not necessarily the best in terms of contribution margin, as it yields the highest delivery costs. In practice, when choosing a pricing approach, one has to examine closely which is more relevant for long-term success – losing a customer's goodwill due to being rejected or due to higher delivery cost.

The conducted study's results provide starting points for future efforts in several directions. The first direction concerns anticipation in solving i-DMVRPs with overlapping booking and service horizons. In future studies, it could be useful to examine hybrid anticipation approaches that combine learning based and non-learning based decision making. Thus, a good starting point would be to explore whether adding a previously learned end-of-horizon valuation to the presented approach would improve its performance. The second direction concerns the pricing component of the newly developed approach and the different variations compared. The observed results show that an increase in price flexibility leads to a decrease in cost efficiency, which is a very intriguing direction for deeper analyses, especially when dealing with continuous explicit price optimization and more complex customer choice models. The third direction concerns an entirely different, more revenue management oriented view. It would be very elucidating to further investigate the hierarchical demand management decomposition approach that has been developed. Particularly, it could be studied how this approach performs in different environments and for different problems, e.g., with more complex pricing and choice models, and whether it would then still be possible to apply it in online algorithms.

Overall, the research underlying this dissertation shows that i-DMVRPs form a highly practice-relevant, but extraordinary complex class of optimization problems, which have not yet been discussed to completeness in the literature. Therefore, the results presented in this dissertation contribute to both practice as well as research. First, the formally introduced problem definition and the derived taxonomy helps practitioners to classify their practical problems and, thus, supports an efficient, targeted search for solution approaches. At the same time, it helps researchers to efficiently gain an overview of existing approaches and, thus, substantially supports the delineation of future research from existing works. Analogously, the introduced unified modeling framework enables transferring existing solution approaches among different problem settings and its analytical discussion sets a crucial foundation for the acceleration of transferring approaches from different fields of research to the research on i-DMVRPs. Finally, with the newly developed solution approach, practitioners are provided with a tool to comprehensively tackle SDD problems and, therewith, operate SDD services profitably for the first time.

Bibliography

Adelman, D. (2007). Dynamic bid prices in revenue management. *Operations Research*, 55 (4), 647–661.

Agatz, N., Campbell, A. M., Fleischmann, M., Van Nunen, J., and Savelsbergh, M. (2013). Revenue management opportunities for internet retailers. *Journal of Revenue and Pricing Management*, 12(2), 128–138.

Agatz, N., Fan, Y., and Stam, D. (2021). The impact of green labels on time slot choice and operational sustainability. *Production and Operations Management*, 30(7), 2285–2303.

Agatz, N. A., Fleischmann, M., and Van Nunen, J. A. (2008). E-fulfillment and multi-channel distribution–a review. *European Journal of Operational Research*, 187(2), 339–356.

Anderson, C. K. and Carroll, B. (2007). Demand management: Beyond revenue management. *Journal of Revenue and Pricing Management*, 6(4), 260–263.

Andrew Travis. (2022) Pain at the pump: How will rising gas prices affect last-mile delivery?. URL `https://www.foodlogistics.com/transportation/last-mile/article/22172109/onfleet-pain-at-the-pump-how-will-rising-gas-prices-affect-lastmile-delivery`. Accessed May 10, 2022.

Angelelli, E., Archetti, C., Filippi, C., and Vindigni, M. (2021). A dynamic and probabilistic orienteering problem. *Computers & Operations Research*, 136, 105454.

Archetti, C. and Bertazzi, L. (2021). Recent challenges in routing and inventory routing: E-commerce and last-mile delivery. *Networks*, 77(2), 255–268.

Asdemir, K., Jacob, V. S., and Krishnan, R. (2009). Dynamic pricing of multiple home delivery options. *European Journal of Operational Research*, 196(1), 246–257.

Azi, N., Gendreau, M., and Potvin, J.-Y. (2012). A dynamic vehicle routing problem with multiple delivery routes. *Annals of Operations Research*, 199(1), 103–112.

Bent, R. and Van Hentenryck, P. (2007). Waiting and relocation strategies in online stochastic vehicle routing. *International Joint Conferences on Artificial Intelligence Organization*, 7, 1816–1821.

Bent, R. W. and Van Hentenryck, P. (2004). Scenario-based planning for partially dynamic vehicle routing with stochastic customers. *Operations Research*, 52(6), 977–987.

Boysen, N., Fedtke, S., and Schwerdfeger, S. (2021). Last-mile delivery concepts: A survey from an operational research perspective. *OR Spectrum*, 43(1), 1–58.

Bruck, B. P., Cordeau, J.-F., and Iori, M. (2018). A practical time slot management and routing problem for attended home services. *Omega*, 81, 208–219.

Campbell, A. M. and Savelsbergh, M. (2006). Incentive schemes for attended home delivery services. *Transportation Science*, 40(3), 327–341.

Campbell, A. M. and Savelsbergh, M. W. (2005). Decision support for consumer direct grocery initiatives. *Transportation Science*, 39(3), 313–327.

Capgemini. (2019). Share of total supply costs worldwide in 2018, by type of cost. URL `https://www.statista.com/statistics/1043253/share-of-total-supply-chain-costs-by-type-worldwide/`. Accessed Januar 5, 2022.

Cattaruzza, D., Absi, N., and Feillet, D. (2016). The multi-trip vehicle routing problem with time windows and release dates. *Transportation Science*, 50(2), 676–693.

Cattaruzza, D., Absi, N., Feillet, D., and González-Feliu, J. (2017). Vehicle routing problems for city logistics. *EURO Journal on Transportation and Logistics*, 6(1), 51–79.

Chbichib, A., Mellouli, R., and Chabchoub, H. (2012). Profitable vehicle routing problem with multiple trips: Modeling and variable neighborhood descent algorithm. *American Journal of Operational Research*, 2(6), 104–119.

Chen, C., Demir, E., Huang, Y., and Qiu, R. (2021). The adoption of self-driving delivery robots in last mile logistics. *Transportation Research Part E: Logistics and Transportation Review*, 146, 102214.

Chen, X., Ulmer, M. W., and Thomas, B. W. (2019). Deep q-learning for same-day delivery with a heterogeneous fleet of vehicles and drones. *arXiv preprint arXiv:1910.11901*.

Chen, X., Wang, T., Thomas, B. W., and Ulmer, M. W. (2020). Same-day delivery with fairness. *arXiv preprint arXiv:2007.09541*.

Côté, J.-F., de Queiroz, T. A., Gallesi, F., and Iori, M. (2021). *Dynamic Optimization Algorithms for Same-day Delivery Problems*. Bureau de Montreal, Université de Montreal.

Daganzo, C. F. (1987). Modeling distribution problems with time windows: Part I. *Transportation Science*, 21(3), 171–179.

Dantzig, G. B. and Ramser, J. H. (1959). The truck dispatching problem. *Management Science*, 6(1), 80–91.

Dayarian, I. and Savelsbergh, M. (2020). Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders. *Production and Operations Management*, 29(9), 2153–2174.

Dumouchelle, J., Frejinger, E., and Lodi, A. (2021). Can machine learning help in solving cargo capacity management booking control problems? *arXiv preprint arXiv:2102.00092*.

eMarketer. (2022). Retail e-commerce sales worldwide from 2014 to 2025. URL `https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/`. Accessed May 13, 2022.

Fleckenstein, D., Klein, R., and Steinhardt, C. (2021). Recent advances in integrating demand management and vehicle routing: A methodological review. *SSRN*.

Florio, A. M., Hartl, R. F., and Minner, S. (2020). Optimal a priori tour and restocking policy for the single-vehicle routing problem with stochastic demands. *European Journal of Operational Research*, 285(1), 172–182.

Gallego, G., Topaloglu, H., et al. (2019). *Revenue management and pricing analytics*. Springer.

Gauvin, C., Desaulniers, G., and Gendreau, M. (2014). A branch-cut-and-price algorithm for the vehicle routing problem with stochastic demands. *Computers & Operations Research*, 50, 141–153.

Gendreau, M., Laporte, G., and Séguin, R. (1996). Stochastic vehicle routing. *European Journal of Operational Research*, 88(1), 3–12.

Giallombardo, G., Guerriero, F., and Miglionico, G. (2020). Profit maximization via capacity control for distribution logistics problems. *arXiv preprint arXiv:2008.03216*.

Golden, B. L., Raghavan, S., Wasil, E. A., et al. (2008). *The vehicle routing problem: Latest advances and new challenges*. Springer.

Gönsch, J. and Steinhardt, C. (2015). On the incorporation of upgrades into airline network revenue management. *Review of Managerial Science*, 9(4), 635–660.

Hildebrandt, F. D., Thomas, B., and Ulmer, M. W. (2021). Where the action is: Let's make reinforcement learning for stochastic dynamic vehicle routing problems work! *arXiv preprint arXiv:2103.00507*.

Huefner, R. (2015). *Revenue Management: A path to increased profits*. Business Expert Press.

Klapp, M. A., Erera, A. L., and Toriello, A. (2018). The dynamic dispatch waves problem for same-day delivery. *European Journal of Operational Research*, 271(2), 519–534.

Klapp, M. A., Erera, A. L., and Toriello, A. (2020). Request acceptance in same-day delivery. *Transportation Research Part E: Logistics and Transportation Review*, 143, 102083.

Klein, R., Mackert, J., Neugebauer, M., and Steinhardt, C. (2018). A model-based approximation of opportunity cost for dynamic pricing in attended home delivery. *OR Spectrum*, 40(4), 969–996.

Klein, R., Neugebauer, M., Ratkovitch, D., and Steinhardt, C. (2019). Differentiated time slot pricing under routing considerations in attended home delivery. *Transportation Science*, 53(1), 236–255.

Klein, R., Koch, S., Steinhardt, C., and Strauss, A. K. (2020). A review of revenue management: Recent generalizations and advances in industry applications. *European Journal of Operational Research*, 284(2), 397–412.

Koch, S. (2017). Least squares approximate policy iteration for learning bid prices in choice-based revenue management. *Computers & Operations Research*, 77, 240–253.

Koch, S. and Klein, R. (2020). Route-based approximate dynamic programming for dynamic pricing in attended home delivery. *European Journal of Operational Research*, 287(2), 633–652.

Köhler, C., Ehmke, J. F., Champbell, A. M., and Cleophas, C. (2019). Flexible dynamic time window pricing for attended home deliveries. *Working Paper Series*.

Köhler, C., Ehmke, J. F., and Campbell, A. M. (2020). Flexible time window management for attended home deliveries. *Omega*, 91, 102023.

Lang, M. and Cleophas, C. (2020). Establishing an extendable benchmarking framework for e-fulfillment. *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 1589–1598.

Lang, M. A., Cleophas, C., and Ehmke, J. F. (2021). Anticipative dynamic slotting for attended home deliveries. *Operations Research Forum*, 2(4), 1–39.

Lang, M. A., Cleophas, C., and Ehmke, J. F. (2021). Multi-criteria decision making in dynamic slotting for attended home deliveries. *Omega*, 102, 102305.

Lebedev, D., Margellos, K., and Goulart, P. (2020). Approximate dynamic programming for delivery time slot pricing: A sensitivity analysis. *arXiv preprint arXiv:2008.00780*.

Lebedev, D., Goulart, P., and Margellos, K. (2021). A dynamic programming framework for optimal delivery time slot pricing. *European Journal of Operational Research*, 292(2), 456–468.

Lohnert, N. and Fischer, K. (2019). Booking limit based revenue management approaches for customer-value oriented make-to-order production. *Logistics Management*, 268–282.

Louveaux, F. V. and Salazar-González, J.-J. (2018). Exact approach for the vehicle routing problem with stochastic demands and preventive returns. *Transportation Science*, 52 (6), 1463–1478.

Lysgaard, J. (2003). Cvrpsep: A package of separation routines for the capacitated vehicle routing problem.

Mackert, J. (2019). Choice-based dynamic time slot management in attended home delivery. *Computers & Industrial Engineering*, 129, 333–345.

Maddah, B., Moussawi-Haidar, L., El-Taha, M., and Rida, H. (2010). Dynamic cruise ship revenue management. *European Journal of Operational Research*, 207(1), 445–455.

McKinsey and Company. (2016). Parcel delivery the future of last mile. URL `https://www.mckinsey.com/~/media/mckinsey/industries/travel%20transport%20and%20logistics/our%20insights/how%20customer%20demands%20are%20reshaping%20last%20mile%20delivery/parcel_delivery_the_future_of_last_mile.ashx`. Accessed December 12, 2017.

Pillac, V., Gendreau, M., Guéret, C., and Medaglia, A. L. (2013). A review of dynamic vehicle routing problems. *European Journal of Operational Research*, 225(1), 1–11.

Powell, W. B. (2019). A unified framework for stochastic optimization. *European Journal of Operational Research*, 275(3), 795–821.

Powell, W. B., Simao, H. P., and Bouzaiene-Ayari, B. (2012). Approximate dynamic programming in transportation and logistics: A unified framework. *EURO Journal on Transportation and Logistics*, 1(3), 237–284.

Prokhorchuk, A., Dauwels, J., and Jaillet, P. (2019). Stochastic dynamic pricing for same-day delivery routing. *arXiv preprint arXiv:1912.02946*.

Puterman, M. L. (2014). *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons.

PwC. (2018). Global consumer insights survey 2018. URL `https://www.pwc.com/gx/en/retail-consumer/assets/delivery-expectations-global-consumer-insights-survey.pdf`. Accessed September 1, 2019.

Snoeck, A., Merchan, D., and Winkenbach, M. (2020). Revenue management in last-mile delivery: State-of-the-art and future research directions. *Transportation Research Procedia*, 46, 109–116.

Soeffker, N., Ulmer, M. W., and Mattfeld, D. (2017). On fairness aspects of customer acceptance mechanisms in dynamic vehicle routing. *Proceedings of Logistikmanagement 2017*, 17–24.

Soeffker, N., Ulmer, M. W., and Mattfeld, D. C. (2021). Stochastic dynamic vehicle routing in the light of prescriptive analytics: A review. *European Journal of Operational Research*, 298(3), 801–820.

Soppert, M., Steinhardt, C., Müller, C., and Gönsch, J. (2022). Differentiated pricing of shared mobility systems considering network effects. *Transportation Science*, 56(5), 1279–1303.

Strauss, A. K., Klein, R., and Steinhardt, C. (2018). A review of choice-based revenue management: Theory and methods. *European Journal of Operational Research*, 271 (2), 375–387.

Strauss, A. K., Gülpınar, N., and Zheng, Y. (2021). Dynamic pricing of flexible time slots for attended home delivery. *European Journal of Operational Research*, 294(3), 1022–1041.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Talluri, K. T. and Van Ryzin, G. J. (2006). *The theory and practice of revenue management*. Springer Science & Business Media.

Toth, P. and Vigo, D. (2014). *Vehicle routing: Problems, methods, and applications*. SIAM.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Ulmer, M. W. (2019). Anticipation versus reactive reoptimization for dynamic vehicle routing with stochastic requests. *Networks*, 73(3), 277–291.

Ulmer, M. W. (2020). Dynamic pricing and routing for same-day delivery. *Transportation Science*, 54(4), 1016–1033.

Ulmer, M. W. (2020). Horizontal combinations of online and offline approximate dynamic programming for stochastic dynamic vehicle routing. *Central European Journal of Operations Research*, 28(1), 279–308.

Ulmer, M. W. and Streng, S. (2019). Same-day delivery with pickup stations and autonomous vehicles. *Computers & Operations Research*, 108, 1–19.

Ulmer, M. W. and Thomas, B. W. (2018). Same-day delivery with heterogeneous fleets of drones and vehicles. *Networks*, 72(4), 475–505.

Ulmer, M. W. and Thomas, B. W. (2020). Meso-parametric value function approximation for dynamic customer acceptances in delivery routing. *European Journal of Operational Research*, 285(1), 183–195.

Ulmer, M. W., Mattfeld, D. C., and Köster, F. (2018). Budgeting time for dynamic vehicle routing with stochastic customer requests. *Transportation Science*, 52(1), 20–37.

Ulmer, M. W., Goodson, J. C., Mattfeld, D. C., and Hennig, M. (2019). Offline–online approximate dynamic programming for dynamic vehicle routing with stochastic requests. *Transportation Science*, 53(1), 185–202.

Ulmer, M. W., Goodson, J. C., Mattfeld, D. C., and Thomas, B. W. (2020). On modeling stochastic dynamic vehicle routing problems. *EURO Journal on Transportation and Logistics*, 9(2), 100008.

Ulmer, M. W. (2017). *Approximate dynamic programming for dynamic vehicle routing*. Springer.

Unnikrishnan, A. and Figliozzi, M. A. (2020). A study of the impact of covid-19 on home delivery purchases and expenditures.

Vinsensius, A., Wang, Y., Chew, E. P., and Lee, L. H. (2020). Dynamic incentive mechanism for delivery slot management in e-commerce attended home delivery. *Transportation Science*, 54(3), 567–587.

Voccia, S. A., Campbell, A. M., and Thomas, B. W. (2019). The same-day delivery problem for online purchases. *Transportation Science*, 53(1), 167–184.

Waßmuth, K., Köhler, C., Agatz, N., and Fleischmann, M. (2022). Demand management for attended home delivery–a literature review. *ERIM Report Series Reference Forthcoming*.

Weatherford, L. R. and Bodily, S. E. (1992). A taxonomy and research overview of perishable-asset revenue management: Yield management, overbooking, and pricing. *Operations Research*, 40(5), 831–844.

World Bank. (2022). Natural gas commodity prices in europe and the united states from 1980 to 2021, with a forecast for 2022. URL `https://www.statista.com/statistics/252791/natural-gas-prices/#:~:text=In%202022%2C%20the%20price%20of,further%20constraints%20resulting%20from%20the`. Accessed May 22, 2022.

Yang, X. and Strauss, A. K. (2017). An approximate dynamic programming approach to attended home delivery management. *European Journal of Operational Research*, 263 (3), 935–945.

Yang, X., Strauss, A. K., Currie, C. S., and Eglese, R. (2016). Choice-based demand management and vehicle routing in e-fulfillment. *Transportation Science*, 50(2), 473–488.

Yeoman, I. and McMahon-Beattie, U. (2010). *Revenue management: A practical pricing perspective*. Springer.