PERSONALIZED SERVICE AND OPERATION OPTIMIZATION FOR EMERGING MOBILITY SOLUTIONS: FROM GROUND TO AIR

WANG YILUN

PhD

The Hong Kong Polytechnic University

2026

The Hong Kong Polytechnic University

Department of Industrial and Systems Engineering

Personalized Service and Operation Optimization for Emerging Mobility Solutions: From Ground to Air

Wang Yilun

A thesis submitted in partial fulfillment of the requirements

for

the degree of Doctor of Philosophy

September 2025

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

Signature: ________________________

Name of Student: Wang Yilun

Abstract

Urban mobility stands at a critical crossroads as rapid urbanization transforms how people and goods move through cities. By 2050, nearly 70% of the global population will live in urban areas, creating unprecedented challenges in meeting the mobility needs of growing populations. Various mobility solutions have been implemented to satisfy the increasing demand for on-demand transportation services, such as on-demand ride services for individuals (e.g., e-hailing), on-demand delivery services for parcels (e.g., same-day delivery), and the upcoming urban air mobility services. While these emerging ground and aerial services offer considerable potential to improve fleet utilization and reduce congestion, their successful implementation faces a fundamental challenge: balancing personalized service delivery with system-wide efficiency in dynamic and stochastic operational environments characterized by fluctuating demand, uncertain passengers behavior, and complex operational constraints. Current approaches often prioritize one at the expense of the other, leading to solutions that may face economic challenges or fail to meet passengers expectations. This critical gap demands innovative approaches that can effectively balance both priorities.

This thesis aims to address this challenge through a series of studies on advanced operations, particularly through personalized service and intelligent operating strategies for emerging mobility solutions across both ground and air domains to improve passenger satisfaction while ensuring economically viable urban mobility systems for the future.

Study I in Chapter 3 investigates the integration of parcel delivery and passenger ride services by a single fleet of mobility-on-demand companies. A dynamic confirmation, compensation, and routing problem is formulated to incentivize passengers to accept package delivery integration into existing passenger transport operations. We propose an innovative anticipatory policy underpinned by a novel regression-tree-based value function approximation to determine whether to serve newly arrived passengers, compensation amounts, and route decisions in a stochastic

and dynamic environment.

Study II in Chapter 4 extends this investigation to urban air mobility by designing and evaluating a novel urban air taxi service model that enables passenger pooling and flight transfers. This model explores the fundamental trade-offs between operational costs and individual passenger experience to establish the viability and potential of this advanced service modality. We formulate an integer programming model based on a hybrid time-space and time-space-battery graph and develop an adaptive arc generation algorithm to efficiently solve this problem.

Study III in Chapter 5 builds upon the foundations established in the previous studies to tackle the ultimate challenge of seamlessly integrating ground and air transportation into a unified, passenger-centric mobility service. To this end, we investigate a choice-aware air taxi service, where multiple multimodal trip options are offered to dynamically arriving passengers. We propose an efficient anticipatory policy supported by a value function approximation based on deep neural networks to optimize option offering, pricing, and dispatching decisions in a stochastic and dynamic environment.

Together, these three studies form a progressive and interconnected exploration of how intelligent personalized service mechanisms and routing can address the operational challenges of emerging mobility solutions from ground to air integrated systems. Managerial insights are derived from numerical experiments to support future mobility operations.

Publications Arising from the Thesis

Journal Publications

- **Yilun Wang**, Min Xu. (2025). Dynamic Confirmation, Compensation and Routing for Combined Urban Transportation of Passengers and Parcels. *Transportation Science*. (Accepted)
- **Yilun Wang**, Min Xu. (2025). Dynamic Pricing and Dispatching for Choice-aware Air Taxi Service. *Transportation Science*. (Under Review)
- **Yilun Wang**, Min Xu, Ya Xiong, Yadong Wang. (2026). Adaptive Arc Generation for Urban Air Mobility Service with Pooling and Flight Transfer. *Transportation Research Part B: Methodological*. (Submitted)

Conference Proceedings

- **Yilun Wang**, Min Xu. (2024). Dynamic Confirmation, Compensation and Routing for Combined Urban Transportation of Passengers and Parcels. *Presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)*. Crete, Greece.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Min Xu, for her tremendous support throughout my PhD journey. She patiently guided me through numerous challenges, significantly enhancing my analytical skills, logical thinking, and ability to consider problems from a comprehensive perspective. Her mentorship has broadened my academic horizons and enriched my doctoral experience immensely. I am also sincerely thankful to my co-supervisor, Prof. Carman Ka Man Lee, for her valuable advice on my research direction. Additionally, I extend my appreciation to the examination committee members, Prof. Andy Chow and Prof. Xiaoqian Sun, for their thorough review and invaluable feedback on my thesis.

I am profoundly grateful to The Hong Kong Polytechnic University for providing financial support for my doctoral studies, and I extend my thanks to the university staff for their assistance throughout my academic journey. My sincere thanks also go to my research colleagues, Dr. Da Lei, Dr. Haohan Xiao, Dr. Ting Wu, Dr. Xiaoyuan Yan, and Dr. Ziwei Yi, for their encouragement and support. I also wish to acknowledge the collaboration and help from my teammates: Dr. Shouguo Peng, Dr. Jiangyan Huang, Dr. Rumei Ji, Dr. Hang Su, Mr. Lu Yang, Mr. Qiaoqiao Ren, Mr. Tanghong Ran, Ms. Xiaohua Chen, , Ms. Lin Tu, Mr. Ya Xiong, Mr. Sen Du, and Ms. Ye Luo.

I am deeply indebted to my lifelong friends, Yiwen and Mingjie, for their unwavering companionship. Most importantly, I thank my parents, Wang Daguang and Ding Li, and my grandparents Ding Shitang and Wan Qiaoling for their endless love, understanding, and patience. They comforted me during moments of frustration, helped me maintain physical and mental well-being, and consistently encouraged me to move forward with clarity and courage whenever I felt lost.

This doctoral journey has been a long and challenging voyage. I am profoundly grateful to the lighthouses that guided me, the companions who sailed with me, and the harbor that always awaited my return. This thesis is dedicated to all of you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAG	Adaptive arc generation
ACCR	Anticipatory confirmation, compensation, and routing policy
ACCR2	ACCR that uses the regression-tree-based VFA with two attributes
A-DLT2	Anticipatory confirmation, compensation, and routing policy using dynamic look-up table with two common attributes
A-DLT4	Anticipatory confirmation, compensation, and routing policy using dynamic look-up table with four attributes
A-DQN	Anticipatory policy based on a deep Q network
APDP	Anticipatory pricing and dispatching policy
CBH-C	Cost-benefit policy with myopic compensation strategy
CBH-F	Cost-benefit policy with fixed compensation ratio
CUT	Combined urban transportation
DCR	Dynamic confirmation, compensation, and routing for CUT
DLT	Dynamic look-up table
DNN	Deep neural network
DQN	Deep Q network
eVTOL	Electric vertical takeoff and landing
E-SMP	Enhanced sparse master problem
MAG	Multi-arc generation
MDP	Markov decision process
MOB	Myopic pricing and dispatching policy
MYP	Myopic confirmation, compensation, and routing policy
OCP	Dynamic pricing strategy based on opportunity cost calculations
PG	Policy gradient
SAR	Share-a-ride problem
SDVRPs	Stochastic and dynamic vehicle routing problems
SMP	Sparse master problem
SP	Arc generation subproblem
SOB	Operator-designated policy
TP	Two-phase heuristic

TPDP	Two-phase anticipatory pricing and dispatching policy
UAM	Urban air mobility
VFA	Value function approximation

Chapter 1 Introduction

1.1 Research Background

Urban mobility stands at a critical crossroads as rapid urbanization transforms how people and goods move through cities. By 2050, nearly 70% of the global population will live in urban areas, creating unprecedented challenges in meeting the mobility needs of growing populations. Various mobility solutions have been implemented to satisfy the increasing demand for on-demand transportation services, such as on-demand ride services for individuals (e.g., e-hailing) or on-demand delivery services for parcels (e.g., same-day delivery, meal delivery). The mobility-on-demand market keeps expanding and is worth over $500 billion in 2021, with projections exceeding $1600 billion by 2031 (Allied Market Research, 2022).

However, as this market matures, its inherent limitations have become increasingly apparent, primarily in two dimensions. The first lies in operational inefficiency. The prevailing business model typically relies on separate operations, i.e., dedicated vehicles for passenger transport and others for goods delivery. This operational paradigm leads to significant inefficiencies: redundant vehicle deployments, increased vehicular miles traveled, exacerbation of urban congestion, and a larger environmental footprint due to parallel operations (Mourad et al., 2019). Interestingly, shared transportation of people and goods has been successfully adopted in long-haul transportation such as aviation, rail, and maritime shipping. However, urban mobility systems continue to operate under a separate paradigm, missing the opportunity to enhance the utilization of road space and infrastructure.

The second, more fundamental limitation is spatial constraints. As cities grow denser and more congested, ground transportation systems would face increasing pressure from physical road capacity, leading to traffic congestion and longer travel times. This ceiling on ground-based mobility fundamentally limits its ability to satisfy the rising demand for rapid and reliable services. These mounting pressures have prompted governments and companies to explore innovative transportation solutions beyond conventional ground-based systems.

These structural limitations have prompted governments and companies to explore innovative transportation solutions beyond conventional transportation systems in two complementary directions.

(1) Combined urban transportation of passengers and parcels

To tackle the operational inefficiency, combined urban transportation (CUT) emerges by integrating the transportation of passengers and parcels. The idea is to utilize a single fleet to cater to both types of requests (Cheng et al., 2023). By capturing the consolidation opportunity of the two types of requests, which often occur in the same region, this mobility solution has great potential to reduce vehicle miles traveled and operational costs (Mourad et al., 2019). A practical implementation of this concept can be seen in Uber Connect, a service that integrates package delivery into Uber's existing passenger transport operations (Uber, 2020). By allowing users to send packages to friends, family, or businesses through the same app used for ride-hailing, Uber Connect exemplifies how combining passenger and parcel services can enhance efficiency and meet diverse urban transportation needs. This service is particularly beneficial for various use cases, such as sending gifts, care packages, or important documents, offering a convenient alternative to traditional postal services (Sawers, 2020).

(2) Urban air mobility

Recent technological advances in electric vertical takeoff and landing (eVTOL) aircraft prompt the development of urban air mobility, which has great potential to revolutionize urban mobility by reducing congestion and travel times. The market of urban air mobility is expected to be worth \$1.5 trillion by 2040 (Jonas, 2019). These aircraft distinguish themselves from traditional helicopters through electric propulsion systems, resulting in reduced noise, pollution, operational costs, and energy consumption, making them particularly well-suited for urban environments (Pons-Prats et al., 2022). By leveraging these advantages, passengers can be transported in the airspace at lower altitudes within urban and suburban areas based on a network of vertiports. Vertiports are takeoff and landing infrastructure equipped with facilities for eVTOLs charging, and passengers embarking/disembarking. This service model can also accommodate passenger pooling for flights, offering a flexible and efficient mode of transportation between urban destinations (Rajendran and Srinivas, 2020). Companies such as Joby, EHang, Airbus, and Wisk are pioneering the development of eVTOL aircraft. Meanwhile, countries including the United States, China, Germany, and India are driving commercial urban air mobility readiness through government regulations, infrastructure development, and consumer acceptance initiatives. These coordinated efforts demonstrate its transition from concept to reality.

While CUT and UAM present promising potential to transform the urban transportation, their successful implementation faces two interconnected challenges regarding personalized service design and operational optimization.

(1) Personalized service design

The viability of both CUT and UAM depends on widespread passengers' adoption. However, passengers exhibit highly heterogeneous and often latent preferences across critical attributes such as travel time, cost, privacy, comfort, and willingness to share space or accept detours. This diversity creates a central tension: the pursuit of system-wide efficiency often conflicts with individual user expectations.

The implementation of CUT can induce detours required for parcel delivery during passenger trips. While some passengers may be highly time-sensitive and refuse such options, others might be price-sensitive and willingly trade a longer travel time for a reduced fare. A failure to account for this heterogeneous preference results in lost consolidation opportunities and suboptimal resource utilization, and potentially deters future use of the service.

As a novel and premium mode, passengers' preference towards UAM manifests differently. Critical adoption barriers include price sensitivity, safety perceptions, and travel time savings (Long et al., 2023). Business travelers typically prioritize fast travel and are willing to pay premium prices for services utilizing nearby vertiports, while price-sensitive passengers prefer more economical options despite potentially longer travel times. These interrelated factors present significant optimization challenges for air taxi operators seeking to establish economically viable and passenger-satisfying services in urban environments.

Therefore, it is imperative to develop personalized service mechanisms for those emerging mobility solutions to align individual preferences with system efficiency. Hence, we can transform the challenge of heterogeneous preferences into an opportunity for optimized market segmentation and resource allocation.

(2) Operational optimization

The implementation of CUT and UAM also introduces critical operational challenges of serving on-demand requests while minimizing operational costs. This creates a trade-off between the pursuit of higher efficiency (e.g., through consolidation or pooling) and service trade-offs (e.g., detours or long travel times) that may deter users. While these emerging mobility solutions share the same goal, the specific constraints and complexities differ significantly between ground and air.

The CUT operations focus on consolidating passengers and parcels by a single fleet. The decisions for on-demand CUT services should be responsive to the upcoming dynamic and stochastic requests in real-time operations. Determining whether to serve or reject an upcoming request, and planning the fleet's routes accordingly, are also important for practical implementation. Given a limited fleet size and service period, it is impractical to serve all requests. It is thus essential to balance the immediate and future profitability. This decision, referred to as a confirmation decision, is significant to profit management. Moreover, routing decisions will largely impact fleet utilization. The sequence of serving requests will not only affect travel expenses but also the capacity to serve future requests. In summary, a successful CUT operation will require multiple real-time decisions in a dynamic and stochastic environment, all of which are crucial to maximizing the total profits throughout the service period.

The UAM operations introduce a new layer of complexity due to their reliance on infrastructure and the three-dimensional nature of travel. Unlike traditional taxis that can pick up passengers virtually anywhere, the air mobility services are constrained to operating between designated takeoff and landing vertiports. This reliance on fixed vertiport locations limits service flexibility. Second, critical operational decisions regarding aircraft dispatching and charging have a significant impact on operational efficiency, not only affecting operating costs, but also the availability of serving future demand. The complexity intensifies when coordinating the personalized service with operational optimization. This requires decisions aligning with uncertain passengers' travel preferences. Finally, these challenges are mounting when operating in a dynamic and stochastic environment.

The optimization of CUT and UAM faces significant challenges that extend beyond traditional operational problems. These new mobility services should account for uncertain passenger behavior and individual preferences during the operation. This introduces a fundamental tension between system-wide cost efficiency and service quality, as these objectives often conflict in practice. Conventional models typically assume passengers comply with predetermined plans. However, only meeting these rigid service standards might not align with stochastic individual preference, which may hinder adoption and undermine the sustainable development of these emerging mobility service. Therefore, it is necessary to transform operational mechanisms from basic vehicle-task assignment systems into passenger-centric platforms that dynamically negotiate service value with passengers during the operation. This research aims to develop integrated personalized service and operational optimization framework for these emerging mobility service to convert the traditional trade-off between operational performance and passenger satisfaction into a new source of profitable service value. By providing services that truly align with user preferences, this transformation can support the successful implementation and long-term viability of emerging mobility solutions.

1.2 Research Objectives

The primary goal of this thesis is to develop intelligent frameworks that integrate personalized service and operational optimization for emerging mobility solutions. By balancing enhanced passenger experience with economic viability, these frameworks seek to accelerate market adoption and ensure the long-term operation of emerging transportation systems. The design of such frameworks crucially depends on understanding and incorporating the impact of operational decisions on passenger experience. To systematically address this, the thesis conducts three interconnected studies that progressively advance the framework's design by addressing different aspects of this challenge. This objective is pursued through three studies as follows:

(1) We investigate a novel compensation mechanism for CUT to capitalize on the consolidation opportunity of passengers and parcels in the urban area and mitigate the negative impact of detours required for consolidation on the passenger experience. Specifically, we explore the viability of offering compensations as incentives for detours required for delivery during passengers' trips in dynamic and stochastic scenarios, where service requests are received dynamically and the passengers' acceptance of detours remains uncertain. This research establishes a practical framework that incorporates dynamic confirmation, compensation, and routing decisions, which transform the typical tension between operational efficiency and passenger satisfaction into a mutually beneficial relationship by combining dynamic incentive systems with strategic operational decisions.

(2) To address the high operational costs that threaten the economic sustainability of UAM, we explore a novel on-demand shared air mobility service that incorporates passenger pooling and flight transfers to improve fleet utilization. Specifically, passengers are allowed to share the cabin and transfer at the intermediate vertiport. By formulating a model that synchronizes aircraft routing and charging, and passengers' itinerary planning decisions within a hybrid time-space-battery network, this study provides a unified framework to balance cost-effectiveness and passenger convenience in this emerging air mobility mode.

(3) Considering the stochastic passengers' preference towards air taxi service and their demand for door-to-door transportation, we further explore a choice-aware service framework to provide premium services that provide seamless multimodal transportation. Specifically, we investigate offering personalized multimodal trip options that cater to heterogeneous passenger preferences in a dynamic and stochastic environment. Each option represents a comprehensive door-to-door travel itinerary involving ground connection and aerial trip. Faced with the uncertainty of passengers' choice behavior, the operator must make multifaceted decisions involving

determining which set of trip options to offer, establishing pricing for each option, and executing real-time dispatching decisions once passengers reveal their choices. This research creates an operational framework that transforms standalone air taxi services into seamless door-to-door mobility experiences through choice-aware dynamic decision-making to enhance the total profit and service quality.

These three studies follow a coherent progression: from incentive-based approaches that guide passenger choices while allowing traditional alternatives, to system-wide optimization for those already using new mobility services, and finally to addressing the uncertainty in passenger choice behavior through multiple service options and dynamic pricing. The research evolves from ground transportation to aerial mobility and ultimately to integrated multimodal networks. Moreover, the research methodology advances from the first-layer incentive design to influence behavior, to the second-layer synchronized optimization of operational resources, and finally to the third-layer prescriptive frameworks that integrates behavioral uncertainty into the decision-making framework.

1.3 Thesis Organization

The thesis contains six chapters to address different challenges in emerging mobility solutions from ground to air, where three core chapters present studies of personalized service and operational optimization for CUT and UAM, respectively, as shown in Table 1.1.

Chapter 1 introduces the background of emerging mobility solutions and identifies implementation challenges associated with relevant business models. Research objective aiming at addressing those challenges are outlined to highlight the significance of the research topics.

Chapter 2 critically reviews the relevant research of CUT and UAM, and the corresponding solution methodologies. The research gaps are identified to emphasize the value of the study.

Chapter 3 investigates the viability of offering compensations as incentives for detours in a dynamic and stochastic environment. In this scenario, service requests are received dynamically, and the passengers' acceptance of detours remains uncertain. This problem is formulated as a Markov decision process (MDP) for sequential decision making. An anticipatory policy is proposed to integrate routing, confirmation, and compensation decisions in dynamic and stochastic scenarios. This is achieved by a novel regression-tree-based value function approximation (VFA) for learning future profitability given the fleet status. Numerical experiments demonstrate its excellent performance compared to benchmark policies.

Chapter 4 explores a novel UAM service problem that integrates passenger pooling and flight transfers, where passengers can share a flight and might change the aircraft during their trips, thereby potentially reducing operational costs while maintaining acceptable service levels for passengers. The problem is formulated as an integer programming model based on a hybrid time-space and time-space-battery graph. To solve this problem efficiently, we develop a multi-arc generation (MAG) algorithm as the fundamental solution framework, and then propose an adaptive arc generation (AAG) algorithm, which efficacy is validated through numerical experiments. Furthermore, an extensive sensitivity analysis examines the impact of key factors and derives managerial insights for UAM operators.

Chapter 5 investigates a choice-aware air taxi service, where operators offer multiple multimodal trip options to accommodate stochastic passenger preferences. We formulate this problem as an MDP and develop an anticipatory model that jointly optimizes option offering, pricing, and dispatching decisions, considering future profitability, while explicitly modeling uncertain passenger choice behavior. By exploiting the problem structure, we propose an efficient anticipatory pricing and dispatching policy for real-time decision-making, which is supported by a VFA based on deep neural networks (DNN). In numerical experiments, our approach significantly outperforms benchmark policies, leading to both higher profits and improved service quality during the operation. These findings provide practical guidance for air taxi operators in this emerging transportation sector.

Chapter 6 presents the overview and research contributions by this thesis and recommendations for future studies.

Table 1.1. Summary of thesis structure

Dimension	Study 1: CUT with compensation	Study 2: Shared UAM with transfer	Study 3: Choice-aware air taxi
Objective	Incentive-based mechanism that motivates passengers cooperation for parcel delivery	Coordination of aircraft routing, battery charging, and passenger itinerary planning with pooling and flight transfer	Integration of passenger choice uncertainty into pricing, option generation, and dispatching decisions
Problem formulation	Dynamic and stochastic programming	Integer programming	Dynamic and stochastic programming
Methodology	Approximate dynamic programming with value function approximation	Adaptive arc generation	Approximate dynamic programming with value function approximation
Progression	Incentive-based approaches to influence behavior	Synchronized optimization of resources and inconvenience	Prescriptive frameworks that integrates behavioral uncertainty into the decision-making framework

Chapter 2 Literature Review

In view of the focus of research, this chapter will critically review the relevant studies of integrated passenger-parcel transportation and urban air mobility operation. In addition, we will also discuss some of the most relevant research on stochastic and dynamic vehicle routing problems (SDVRPs).

2.1 Integrated Passenger-parcel Transportation

Research on integrated passenger-parcel transportation systems has received much attention in recent years, which utilizes the fixed infrastructure, e.g. metros or trams, or the free-floating fleet, e.g. taxis (Cheng et al., 2023). In light of the focus on this study, we mainly review the studies on the free-floating fleet-based CUT services. Li et al. (2014) first introduced a static share-a-ride problem (SAR), where passengers and parcels were transported by a single fleet of taxis. They formulated mixed-integer linear programming models for static cases with requests known a priori, while proposing a neighborhood search-based method for dynamic scenarios with real-time requests. The negative impact of consolidation on passengers was restricted by the maximum travel duration and the number of stops during the passenger trips. Two stochastic variants of SAR regarding stochastic travel time and stochastic delivery location were further investigated by Li et al. (2016). An adaptive large neighborhood search with sampling strategies was proposed to address the stochastic problems. Chen et al. (2016) explored a city-wide package delivery solution that exploited relays of taxis to help transport packages. To ensure the quality of passenger ride service, parcel delivery was prohibited during passenger trips. The authors developed a two-phase framework that employed historical taxi data to identify optimal delivery paths and an adaptive scheduling algorithm to guide real-time package routing. Chen et al. (2019) later proposed a package delivery scheme based on multi-hop ridesharing services and a prediction-based routing planning algorithm. The inconvenience to passengers caused by parcel delivery was not considered. Meinhardt et al. (2022) and Schlenther et al. (2020) explored simulation-based methods to execute freight tours with an on-demand mobility fleet.

Service priority was given to passengers by prohibiting parcel delivery in passenger trips. Moreover, in the study by Schlenther et al. (2020), vehicles were scheduled to serve freight requests only if a certain percentage of the vehicles in the fleet were idle. Bosse et al. (2023) considered a similar mechanism and proposed a dynamic priority rule to adjust the percentage of vehicles for serving passengers during the day. Fehn et al. (2022) investigated the integration of freight delivery in the passenger transportation of an on-demand ride-pooling service in an agent-based simulation framework. They considered both a moderate and a full integration scenario: in the moderate scenario, no stops for parcel pick-up and delivery were allowed during a passenger trip; whereas in the full integration scenario, detours for parcel delivery were allowed with passengers on board. Besides research on road networks, there are also studies on rail networks. Hörsting and Cleophas (2023) investigated the integrated planning of shared passenger and freight transport on a rail network and proposed a linear mixed-integer program to optimize the train schedule and cargo allocation.

The aforementioned studies on CUT either ignored the potential detour or restricted it by enforcing deterministic constraints. However, in practice, passengers' acceptance of detours could be uncertain. This ignorance may discourage passengers from using the CUT service in the long term, posing a threat to the operation's sustainability. Moreover, the solution algorithms proposed by most of the existing studies for the dynamic CUT problem were purely reactive, without adequately considering the stochastic information of future requests and the impact of current decisions on future profitability. To fill the research gaps, in this study, we consider offering compensation as incentives to passengers who are willing to accept detours during their rides in a dynamic and stochastic scenario, where the uncertainty of passengers' acceptance of detour options and the impact of future potential requests are explicitly considered.

2.2 Urban Air Mobility Operation

Research on urban air mobility has received significant attention in recent years (Rajendran and Srinivas, 2020). Strategic-level optimization focuses on infrastructure and fleet configurations. Chen et al. (2022a) investigated vertiport hub location optimization for air taxi operations in a metropolitan region, which was formulated as a discrete location problem. A mixed variable neighborhood search was proposed to solve this problem efficiently. Wang et al. (2022) studied a optimization problem for vertiport location and capacity considering interdependencies between strategic vertiport deployment, tactical operations, and passenger demand. Husemann et al. (2024) investigated the fleet configurations optimization for UAM operation involving vehicle types, battery capacity, and charging infrastructure. In addition to strategic-level optimization, operational challenges are also critical to UAM operation, involving passengers acceptance, service reliability, and coordination (Wandelt et al., 2023). In light of our research, we primarily review those who address operational-level problems, particularly involving vertiport assignment and aircraft routing/dispatching to fulfill on-demand requests under various operational constraints.

Espinoza et al. (2008a) pioneered work on the dial-a-flight problem, addressing the static scenario of routing a fleet of small jet planes to transport passengers from pre-specified origin airports to destination airports. An integer multicommodity network flow model was formulated to optimize the aircraft routing to minimize the operational costs. Espinoza et al. (2008b) further developed a large neighborhood search algorithm to solve this problem at a practical scale. Shihab et al. (2020) investigated dispatching eVTOLs to transport passengers and provide power grid service. A non-convex mixed integer quadratically constrained quadratic program was formulated to maximize the revenue, which was solved by Gurobi.

Beyond merely aerial transportation optimization, the integration of ground transportation has emerged as essential for providing seamless door-to-door air taxi services (Yan et al., 2024). Ale-Ahmad and Mahmassani (2021) investigated multimodal journey planning through coordinated vertiport assignment and shared flight arrangements, enabling passengers with diverse origins and destinations to utilize the same aircraft. A mixed integer programming was formulated to jointly optimize multimodal journey planning, pooling, and aircraft dispatching decisions, with a weighted multi-objective of maximizing revenue and minimizing the trip inconvenience.

While these operational advancements are critical, they must be complemented by passenger-centric considerations. To establish the air taxi as a competitive premium service, operators should also balance operational efficiency with tailored passenger mobility experience to ensure market viability and passenger satisfaction (Pons-Prats et al., 2022). Wu and Zhang (2020) developed a mixed integer program to jointly optimize aircraft charging and passenger pooling at individual vertiports with the objective of minimizing the total waiting times of passengers. Bennaceur et al. (2022) investigated air taxi operations featuring stratified service classes (regular and premium tiers), where premium passengers enjoyed prioritized scheduling and shorter waiting times. A two-phase framework was developed, which first optimized flight scheduling to minimize the overall weighted passengers waiting time, and then addressed aircraft routing using a variable neighborhood search algorithm to minimize the operational costs. Some studies have incorporated multimodal considerations and broader passenger preferences. Li et al. (2025) studied an urban air mobility service that accommodated passengers' multimodal trip preferences, such as maximum allowable transfers. A multi-stage solution framework was proposed to decouple the original problem into passengers' multimodal trip planning, available resource search, task assignment, and vehicle rebalancing. Munari and Alvarez (2019) investigated an aircraft routing problem that considered strategic flight upgrades to satisfy passenger-specific aircraft type demands. A mixed integer programming model was formulated to minimize the operational costs, which are solved by open-source optimization softwares.

The deficiency in flexible service design for UAM represents a significant barrier to system optimization. Current research on UAM optimizes routing and pooling for efficiency but remains limited by a point-to-point service paradigm. The omission of intermediate flight transfers might significantly restrict network flexibility, geographical coverage, and potential gains in fleet utilization, preventing air taxi systems from achieving their full operational and economic potential. Furthermore, current decision-making frameworks generally employ static and deterministic optimization methods, which is not adaptable to the inherently dynamic and stochastic nature of request arrivals in real-world operations. Furthermore, existing studies in air taxi service operations generally assume that passengers will accept operator-planned trip service. However, this assumption overlooks the heterogeneity and uncertainty of passengers' choice behavior. In reality, potential passengers may reject air taxi services based on individual preferences, which might be unknown and uncertain regarding factors such as price sensitivity and travel time requirements (Long et al., 2023). This oversight of stochastic passenger choice behavior creates a critical misalignment between service design and actual travel demand, which may compromise both passenger satisfaction and revenue potential.

Pricing decisions also play a crucial role in this business model by directly influencing both passenger choice behavior and operational decisions. However, few studies incorporate pricing decisions into the comprehensive decision-making framework for air taxi service operations. The limitations in current approaches necessitate the development of a comprehensive decision-making framework that integrates pricing optimization with operational decisions to align the service provision and actual travel demand within the dynamic and stochastic air taxi environment.

2.3 Challenges and Advances in Service-oriented SDVRPs

The inherently dynamic and stochastic nature of request arrivals in real-world operations presents significant challenges for both integrated passenger-parcel transportation and urban air mobility systems. This uncertainty of future demand is rather common in modern urban mobility services, where requests are heterogeneous regarding their arriving time, origin, destination, and service requirements. These challenges can be categorized as sub-problems within the SDVRPs, necessitating specialized solution approaches. Particularly, service-oriented SDVRPs must go beyond passive reaction to information changes, as it often leads to inefficiencies. For example, vehicles left idle in remote locations will be unable to respond quickly to new demand. Moreover, an overly tight routing schedule leaves little buffer to absorb surging demands. Instead, anticipatory planning is required, leveraging the vast amounts of historical data collected by companies to incorporate potential information changes into decision-making. This shift from descriptive and predictive analytics to prescriptive analytics enables more proactive and adaptive solutions, better suited to real-world dynamics. In this section, we review the main solution approaches for SDVRPs, examine the unique challenges posed by service-oriented extensions, and discuss recent advances in integrating service mechanisms with routing decisions.

The SDVRPs are typically modeled as sequential decision processes, which are addressed by a policy function that maps states into decisions at each decision epoch. Moreover, SDVRPs are featured with problem-specific variants of vehicle routing problems, where the decision space is formulated as a mixed-integer program based on the state at each decision epoch. The ultimate objective is to find an optimal policy maximizing the expected sum of rewards through request fulfillment. Solution approaches for SDVRPs can be widely classified into myopic approaches and anticipatory approaches. Myopic approaches implement a rolling horizon framework, which only utilizes the current problem state without involving anticipatory information about future requests (Angelelli et al., 2016; Arslan et al., 2019; Steever et al., 2019). In contrast, anticipatory approaches align with the prescriptive analytics paradigm by integrating predictive models to proactively optimize routing and resource allocation, thereby improving responsiveness and efficiency in urban mobility services. Given the large scale of most real-world problems, anticipatory approaches were widely explored to anticipate the impact of current decision making on the fleet's future performance. These approaches can be classified into two distinct methodological categories: VFA and PG methods (Powell, 2022). While the VFA methods focus on estimating the expected future reward-to-go value of potential decisions based on the current state to guide decision-making, the policy gradient methods directly learn a policy by optimizing the probability of selecting candidate decisions over the decision space, without explicit value estimation. Ulmer et al. (2015) proposed a cost-benefit heuristic, confirming the requests if the ratio of gain to cost for serving them exceeds a predefined threshold, which does not evaluate the reward-to-go value explicitly. Explicit anticipatory approaches for evaluation of reward-to-go value can be widely classified into online and offline approaches. Online anticipation is mainly achieved by sampling approaches, which takes into account of known requests and potential realizations of future periods via sampling (Azi et al., 2012; Bent and Van Hentenryck, 2004; Sungur et al., 2010; Srour et al., 2018; Voccia et al., 2019). This online anticipation requires significant computational effort within the execution phase, and is thus not capable for the problem under consideration. Offline methods do not require extensive computational effort during the implementation and achieve a decision policy during an offline learning phase. Offline methods aggregate the state spaces, often composed of a large number of dimensions, to a set of specific features, and then map the resulting vector of features to a value (Powell, 2007). The mapping depends on the approximate architecture used, such as lookup table, linear function, and neural nets. Schmid (2012) utilized a lookup-table-based VFA to estimate the future ambulance response time when making ambulance relocation and dispatching decisions. The classic look-up table might impede the solution quality due to a static partition of the state space. To address this, several methods have been proposed to extend the traditional VFA process. Ulmer et al. (2018) proposed a new method, which dynamically partitions the state space during the learning process, referred to as dynamic look-up table (DLT). The DLT-based VFA method has shown better performance on problems with routing components compared to the linear function, since they can capture the complex structure of the value function (Ulmer and Thomas, 2020). Müller et al. (2023) approximated the future profitability for vehicles by utilizing a similarity function to relevant data points. Mousavi et al. (2024) utilized the dual value to learn a value function for estimating the marginal cost of serving online orders at a later decision epoch. Chen et al. (2022b) proposed a deep Q-learning approach to estimate the expected number of customers served based on the state and decision.

Instead of learning the reward-to-go value to guide decision-making via VFA, the policy gradient methods directly learn a policy to make decisions without explicit value estimation. Typically, the policy is represented by a neural network that takes the current state as input and outputs the probability of selecting each candidate decision. Then, decisions are either determined by sampling according to the probability distribution in the output vector or by choosing the candidate decision with the highest probability. The policy can be learned through off-line simulation, where the policy's parameters are updated by the gradient based on the observed cumulative rewards (Powell, 2022). Policy gradient methods have been predominantly applied to SDVRPs with merely assignment or routing decisions (Nazari et al., 2018; Mao et al., 2020; Beirigo et al., 2022). The assignment-based SDVRPs focus on assigning requests to vehicles; the decision space is naturally discrete and manageable, which is amenable to either VFA or policy gradient approaches. The routing-based SDVRPs often require techniques of state aggregation and action restriction to render the state and decision space tractable for computational efficiency. However, the more complex SDVRPs that incorporate a service mechanism, involving additional decisions on pricing, incentives, or service options provision, pose significant challenges for both VFA and policy gradient methods. These challenges break a fundamental assumption in traditional approaches: that passengers passively accept whatever assignment or routing decisions the operator makes. In reality, passengers have their own preferences and choice behaviors that are stochastic and often unknown to the system, which introduces an additional layer of uncertainty beyond the typical stochasticity in demand patterns. When passengers can reject offered services based on their individual utility functions, this creates an interaction mechanism: the operational decisions influence passenger choices, which in turn affect system states and future decision-making. Moreover, this challenge also presents an opportunity. By explicitly modeling and responding to passenger preferences through service mechanisms like dynamic pricing, incentive design, or personalized service options, operators can potentially extract additional value from the system. These mechanisms allow for better demand management, improved resource utilization, and ultimately increased profitability. For instance, offering premium service options to passengers with higher willingness-to-pay or providing incentives to motivate passengers to choose services that benefit the system can help enhance efficiency and generate additional revenue.

Given these challenges, a related research lies in the field of SDVRPs with integrated pricing and routing decisions, aiming to maximize the reward generated by satisfying customer demand (Soeffker et al., 2022). For example, Figliozzi et al. (2007) extended the traveling-salesman problem by considering vehicle routing in a dynamic competitive auction environment, where contracts for shipments dynamically arrived and carriers competed for them in a sequence of one-shipment auctions. An approximate solution method with an online one-step look-ahead algorithm was proposed to determine the vehicle routing and auction pricing. Sayarshad and Chow (2015) investigated a dynamic dial-a-ride and pricing problem. Based on the price quoted by the system, a customer might choose to reject the service. The authors developed a non-myopic method based on an M/M/1 queue model to determine the trip price based on the routing decisions and approximation of the fleet's status. Luo and Saigal (2017) considered a dynamic pricing and commission setting problem for trip orders and drivers in a ride-sharing platform to maximize revenue. A continuous approach based on a network model was proposed to determine the price and commissions regarding the time point and location. Qiu et al. (2018) investigated a request-level dynamic pricing and routing problem in a shared mobility service. A dynamic programming framework was proposed to determine the price of exclusive and shared rides for arriving passengers and the vehicle routing for serving the passengers based on the selected ride options. Turan et al. (2020) further developed a deep reinforcement learning algorithm to address a joint routing, battery charging, and pricing problem of an on-demand mobility service.

Extensive research has been conducted to consider how operators can present these prices not as fixed values, but as part of a flexible choice menu for passengers. This broader perspective is captured in the concept of SDVRPs with service options. In this model, operators offer multiple options (e.g., with different prices, time windows, or service levels), allowing passengers to choose according to their preferences and make routing decisions based on passenger selections. These operations involve two key decision components: determining which options to offer and at what price, and routing vehicles efficiently based on resulting passenger choices. Existing research mainly involves four business models: attended home delivery, out-of-home delivery, same-day delivery, and dial-a-ride service (Fleckenstein et al., 2023). In the context of attended home delivery, the option refers to the time slot for attended home service. On the booking horizon, a set of feasible options is provided to passengers, while delivery services are performed on the delivery day after all passengers' selections are known. Strauss et al. (2021) studied dynamic pricing for attended home delivery. A linear programming formulation that integrates pricing decisions, routing costs implications, and stochastic passengers choice behaviors was formulated to maximize the expected total profits. Vinsensius et al. (2020) developed incentive mechanisms to guide passengers toward time slots that potentially lower operator delivery costs. An approximate dynamic programming approach was proposed to optimize these incentives, incorporating estimated routing costs and future orders. Akkerman et al. (2025) investigated out-of-home delivery options for e-commerce platforms, offering varying convenience levels ranging from door-to-door delivery to parcel lockers. They proposed a dynamic decision-making pipeline that first determines the set of options to offer, and then determines their prices by accounting for passenger selection behavior and estimated delivery costs. These cost estimates were derived using a VFA based on convolutional neural networks. In the context of same-day delivery services, the option refers to a delivery service with different deadlines. Based on passengers' selection, real-time routing decisions must be made to fulfill their demands. Ulmer (2020) addressed an interesting dynamic pricing and routing problem for same-day delivery services. Customers were offered multiple delivery options, each with a different deadline and customer-specific delivery price, and the customers would then be asked to select one delivery option based on their preferences. This problem was modeled as an MDP. An anticipatory pricing and routing policy with offline VFA was proposed to determine the price for each delivery deadline option instantly upon the arrival of a customer order and the corresponding vehicle routing for serving the customer by the selected deadline. This approach determines pricing strategies independently of passenger behavior modeling, focusing on operational feasibility and cost considerations. Azadeh et al. (2022) investigated dynamic pricing for options with different waiting times in dial-a-ride services. A mixed integer programming model was formulated to integrate routing and pricing decisions while accounting for passenger choice behavior, and was employed to make dynamic decisions in a rolling horizon approach.

Existing solution approaches for SDVRPs fail to address the complexities inherent in emerging mobility systems such as CUT and UAM. This inadequacy arises from three fundamental limitations in current methodologies. First, prior studies predominantly consider a single type of service request and employ one-time pricing decisions that operate independently without affecting previously committed service constraints. In contrast, the CUT considers two types of requests, i.e., passenger ride requests and parcel delivery requests. The compensation/pricing decision will be made to stimulate the consolidation between the two types of requests, which will affect the trip duration of some confirmed passenger orders in order to accommodate the services for new parcel requests. Moreover, a passenger could be asked for the acceptance of detour options multiple times before completing his or her trip, which would require several times of compensation/pricing decisions during the process. Second, prevailing decision-making frameworks for SDVRPs with service options rely on exhaustive enumeration of all feasible options rather than strategically selecting an optimal subset. Although preliminary work has begun exploring option selection (Akkerman et al., 2025; Azadeh et al., 2022), these approaches neglect to incorporate downstream routing costs and operational decisions into the construction of option sets or the pricing of individual options, resulting in suboptimal system performance. Third, existing research focuses primarily on temporal service differentiation, such as delivery time slots or waiting times. However, air taxi services present a fundamentally different challenge by offering multimodal trips that require consideration of additional factors beyond timing alone, including vertiport assignment and multimodal integration, which could exponentially expand the potential set of trip options, thereby necessitating a more strategic approach to option selection rather than exhaustive enumeration.

2.4 Research Gap Summary

The overall literature review reveals the development of emerging mobility systems, where critical gaps remain at the intersection of operational efficiency, passenger experience, and uncertainty management. This thesis identifies and addresses three interconnected research gaps:

1. The lack of adaptive personalization and proactive service management for CUT. Existing research on CUT often ignore the uncertainty of passenger acceptance for delivery-induced detours or use deterministic constraints not aligning with real-world conditions. Furthermore, solution approaches are predominantly reactive, failing to proactively anticipate future demand and the long-term impact of compensation decisions on system profitability and passenger satisfaction.

2. The deficiency in flexible service design for UAM. Current research on UAM optimizes routing and pooling for efficiency but remains limited by a point-to-point service paradigm. The omission of intermediate flight transfers might significantly restrict network flexibility, geographical coverage, and potential gains in fleet utilization, preventing air taxi systems from achieving their full operational and economic potential.

3. The gap in choice-aware optimization for integrated ground-air mobility systems. The literature lacks frameworks that jointly optimize personalized service options, pricing strategies, and aircraft dispatching while accounting for uncertain passenger choice behavior. It is necessary to develop an intelligent decision-making framework that manages an integrated ground-air network by strategically influencing passenger choices through personalized, door-to-door mobility options.

By targeting these gaps, this thesis contributes a series of models and algorithms that advance the integration of personalized service mechanisms with routing decisions for emerging mobility services, bridging the divide between system efficiency and passenger-centric service.

Chapter 3 Dynamic Confirmation, Compensation, and Routing for Combined Urban Transportation

This chapter focuses on a dynamic confirmation, compensation, and routing problem for CUT, referred to as DCR problem. It aims to address the challenge of CUT implementation: the detours required for parcel deliveries could adversely affect the passenger experience. This chapter explores the viability of a dynamic compensation mechanism that offers compensations as incentives for detours in a dynamic and stochastic environment. In this scenario, service requests are received dynamically and the passenger's acceptance is unknown until their feedback is received. We formulate this dynamic and stochastic problem as an MDP for making sequential decisions. We further propose an anticipatory model and exploit the solution features. Based on the findings, the complexity of the original problem is reduced, and an effective anticipatory policy to make confirmation, compensation and routing decisions, referred to as the ACCR, is proposed to integrate multiple decisions in a unified framework while considering their impact on future profitability. This is achieved by a novel regression-tree-based VFA for learning future profitability given the fleet status. Extensive numerical experiments are conducted to demonstrate that the proposed policy significantly outperforms all the benchmark policies in terms of total profits.

The reminder of this chapter is organized as follows. Section 3.1 describes the DCR problem, and its MDP formulation. Section 3.2 elaborates on the solution method, i.e., the anticipatory policy integrating routing, confirmation, and compensation decisions. Section 3.3 introduces the regression-tree-based VFA. Numerical experiments are conducted in Section 3.4. Section 3.5 concludes this paper with future research directions.

3.1 Problem Description and Formulation

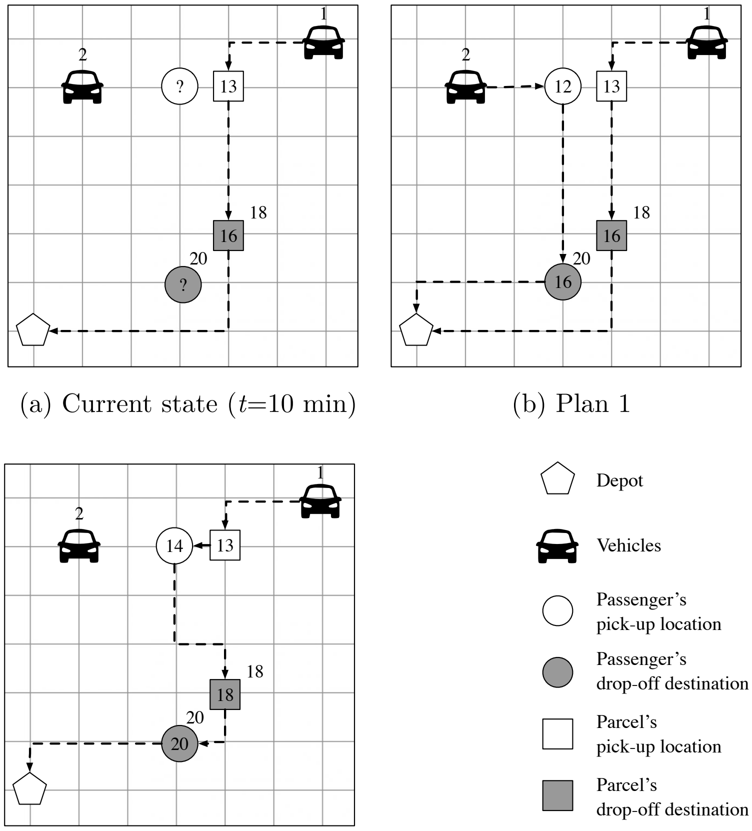
In this section, we first present problem description with an example illustration. Then, we formulate this stochastic and dynamic decision problem as an MDP for sequential decision making.

3.1.1 Problem description

For the DCR problem, we consider a service provider, who centrally operates a fleet of vehicles to serve both passenger ride requests and parcel delivery requests. The vehicles will depart from a depot in the beginning, and have to return to the depot by the end of the operation period. During the operation period, requests with specific pick-up and drop-off locations and a deadline to arrive at the drop-off location, arrive dynamically over time and space. Each request is associated with a known revenue. To maximize the total profit with a limited fleet size and service period, the service provider allows parcel delivery with passengers on board, but ride-pooling among passengers is not allowed. As a result, detours for parcel delivery could be induced in passengers' rides. To mitigate the negative effect on service quality, the service provider would notify the concerned passenger of the detour time limit with a certain amount of compensation as the incentive, referred to as the detour option. The passenger will be asked to indicate his or her acceptance of the detour option. The acceptance or rejection decision of passengers is unknown at the moment the detour option is offered. For ease of operation, the service operator will not consider passenger-parcel consolidation options that need approval from multiple passengers at the same time. However, parcels can still be transported during trips involving multiple passengers, as long as their acceptance doesn't require simultaneous approval. For example, a parcel can initially be accepted during one passenger's trip, and later, as a new passenger request is received, the parcel can continue to be transported during the new passenger's trip. Moreover, one passenger may be asked for detour acceptance multiple times during his/her trip. If the detour option is accepted, the service provider will consolidate the two requests within the detour time limit at the expense of compensation; otherwise, the service provider will respect the passenger's preference and seek other solutions.

Given the above information, the service provider would decide whether to serve each dynamically arriving request, which is termed as confirmation decision, and how to make real-time routing and compensation decisions that determine the visiting sequence of locations of each vehicle and the amount of compensation offered to concerned passengers so that all the confirmed requests are served by their respective deadlines within the detour time limits. For consistency, the vehicles are not allowed to be diverted to other locations in the midstream of traveling until arriving at the current heading location. After visiting all locations in the routing plan, the vehicle will wait at the last visited location, if there is a sufficient period of time to return to the depot before the end of the operation period. Otherwise, the vehicle will return to the depot immediately and terminate service. The objective of the DCR problem is to maximize the total profit of the service provider during the operation period, which is the total revenue of confirmed requests subtracting the travel cost and the amount of compensation accepted by passengers.

We illustrate the DCR problem using a simple example with two vehicles, one passenger ride request, and one parcel delivery request in Figure 5.1. This example is presented in a Manhattan-style grid network with 1-min drive for each arc segment. Figure 3.1(a) shows the initial system state at t=10 min, where Vehicle 1 plans to serve the delivery request, while Vehicle 2 stays still. The pick-up and drop-off locations of this request are represented by a white and dark square, respectively. The arrival times at the pick-up and drop-off locations are t=13 min and t=16 min, respectively, shown inside the square symbols. Simultaneously, a passenger ride request, of which the origin and destination are presented by a white and dark cycle, respectively, arrives with pick-up and drop-off times remaining to be determined. The deadlines of the two requests are displayed in the upper right corner of destination locations. Figures 3.1(b) and 3.1(c) represent two possible routing plans to serve it. Particularly, Figure 3.1(b) shows a routing plan with a direct ride, while Figure 3.1(c) shows a routing plan with a 2-min detour due to parcel drop-off during the passenger trip. If the service provider presumes that Plan 2 is better and decides the amount of compensation, a detour option, including the 2-min detour time limit and the compensation amount, will be forwarded to the passenger for approval. Note that the detour time excludes waiting at the origin, following industry practice where pickup delays are not compensated if arrival is within the agreed window.



(a) Current state (t=10 min) (b) Plan 1

(c) Plan 2

Figure 3.1. Illustration of an example of the DCR problem

If the passenger accepts it, the service provider will implement Plan 2; otherwise, Plan 1 will be executed. While Figure 3.1(c) focuses on a particular case about the integration of a new passenger order into a planned parcel trip, our study also considers the reverse scenario, where a newly arriving parcel order is integrated into a planned passenger trip, as part of the broader problem framework presented in Section 3.1.2.

3.1.2 MDP formulation

Let $\mathscr{O}$ denote the set of ride and delivery requests. The pick-up and drop-off locations of request $o \in \mathscr{O}$ are denoted by o^+ and o^-, respectively. Let e_o be the arrival time of request o, and l_o be the deadline of request o for arriving at location o^-. The revenue of request o is denoted by r_o. For each ride request o, we define detour sensitivity β_o to characterize the degree to which the passenger o would like to accept detour during the trip. It indicates the desired amount of compensation per minute of detour, which is unknown to the service provider. This means that, for a δ-min detour, the desired amount of compensation of ride request o is $\beta_o\delta$; for the offered amount of compensation η, the detour option will be accepted if $\eta \geq \beta_o\delta$ and rejected, otherwise.

Let $\mathscr{N}$ be the set of pick-up and drop-off locations of all requests, and 0 and $2|\mathscr{O}|+1$ denote the dummy nodes for the depot. The vehicles will initially start from node 0, and return to node $2|\mathscr{O}|+1$ at the end of operation period T. Let $\mathscr{A} = \{(i,j) \mid i,j \in \mathscr{N}\}$ be the arc set that includes the connections between any two locations in set $\mathscr{N}$. Let d_{ij} and $\bar{d}_{ij}$ denote the travel time and distance of arc (i,j), respectively. The DCR problem is defined on the graph $\mathscr{G} = (\mathscr{N}, \mathscr{A})$. This problem is a dynamic and stochastic problem, which can be formulated as an MDP for making sequential decisions. The MDP consists of five components: state variables, decision variables, exogenous information, transition function, and objective function. In what follows, we will elaborate on each component. The notations for the DCR problem and MDP formulation are summarized in Table 3.1.

State variables

Let $\mathscr{K}$ be the set of decision epochs. Each epoch $k \in \mathscr{K}$ corresponds to the time point t_k to make a decision triggered by the arrival of a new request o_k. The time point t_k is exactly the arrival time of the new request, i.e., e_{o_k}. The state variable $\mathscr{S}_k$ includes the information for decision making at epoch k, i.e., $\mathscr{S}_k = (t_k, o_k, \Theta_k)$, where Θ_k is the current routing plan of the fleet. Each element $\theta \in \Theta_k$ represents the path of an individual vehicle, which includes both spatial and temporal information. The spatial information contains the sequence of locations to visit. For the temporal information, the arriving times at the pick-up and drop-off locations of request o are denoted by $a^k_{o^+}$ and $a^k_{o^-}$, respectively. For the example in Figure 3.1(a), let o_1 represent the confirmed but not yet served parcel request. We can obtain its arrival times as $a^k_{o_1^+} = 13$, and $a^k_{o_1^-} = 16$. The initial state $\mathscr{S}_0$ is $(0, -, \Theta_0)$, where each path $\theta \in \Theta_0$ only contains dummy node 0 with arrival time 0.

Decision variables

The decision variables at epoch k are determined based on state variable $\mathscr{S}_k$

Table 3.1. Problem notations

Notations	Description
	Global Parameters
$\mathscr{O}$	Set of requests
$\mathscr{N}$	Set of locations
$\mathscr{A}$	Set of links
d_{ij}	Travel time of arc (i, j)
e_o	Arrival time of request o
o^+	Pick-up location of request o
o^-	Drop-off destination of request o
r_o	Revenue of request o
l_o	Deadline of dropping off request o
β_o	Detour sensitivity of passenger request o
$[0, T]$	Operation period
	State Variables
k	Decision epoch
t_k	Time at epoch k
$\mathscr{S}_k$	State variable at decision epoch k
o_k	New request arriving at epoch k
$\mathscr{O}_k^c$	Set of confirmed requests to serve at epoch k
$\mathscr{O}_k = \{o_k\} \cup \mathscr{O}_k^c$	Set of requests to consider at epoch k
Θ_k	Routing plan at epoch k
a_{o^-}	Planned arriving time at the drop-off destination of request o
a_{o^+}	Planned arriving time at the pick-up location of request o
	Decision Variables
ζ_k	Confirmation decision variables at epoch k
η_k	Compensation decision variables at epoch k
τ_k	Routing decision variables at epoch k
$\delta(o, \theta)$	Detour newly induced to o in θ
$\psi(\beta_o, \delta(o, \theta), \eta_{ok})$	Feedback from request o for detour option $(\delta(o, \theta), \eta_{ok})$
$\mathscr{X}(\mathscr{S}_k)$	Set of decision variables at epoch k
$\mathscr{X}^\pi(\mathscr{S}_k)$	Set of decision variables by policy π at epoch k
	Others
$\mathscr{W}_{k+1}(\mathscr{S}_k, x_k)$	Exogenous information
$\mathscr{F}(\mathscr{S}_k, x_k, \mathscr{W}_{k+1})$	Transition function
$\mathscr{R}(\mathscr{S}_k, x_k)$	Realized reward
$V(\mathscr{S}_k, x_k)$	Expected profits given $\mathscr{S}_k$ and x_k

and the feedback from passengers of ride requests at epoch k, which involves confirmation, compensation, and routing decisions. Let $x_k := (\zeta_k, \tau_k, \eta_k)$ denote the decision variables at epoch k, which include confirmation decision variable ζ_k, routing decision variables τ_k, and compensation decision variables η_k. The confirmation decision variable ζ_k is binary, which equals 1 if request o_k is served and 0 otherwise. Given the current routing plan Θ_k, let Θ_k^τ denote the set of potential routing paths for serving the requests in set $\mathscr{O}_k := \mathscr{O}_k^c \cup \{o_k\}$, where $\mathscr{O}_k^c$ is the set of requests confirmed to be served but not yet completely served by epoch k. Let τ_k be the vector representing the corresponding path-based routing decision variable at epoch k. Each entry of this vector is denoted by $\tau_{\theta k} \in \{0, 1\}$, where $\tau_{\theta k} = 1$ if path $\theta \in \Theta_k^\tau$ is chosen, and $\tau_{\theta k} = 0$ otherwise. Let ϕ_θ^o be the coefficient that equals 1 if request o is served on path θ, and 0 otherwise. Then we have the following constraints:

$$\begin{cases} \sum_{\theta \in \Theta_k^\tau} \phi_\theta^{o_k} \tau_{\theta k} = \zeta_k, \\ \sum_{\theta \in \Theta_k^\tau} \phi_\theta^{o} \tau_{\theta k} = 1, \ \forall o \in \mathscr{O}_k^c, \end{cases} \quad \forall k \in \mathscr{K}, \tag{3.1}$$

which ensures that the confirmed requests should be incorporated in the routing plan, and any rejected requests are excluded from consideration.

Let $\mathscr{C}(\theta)$ denote the set of passenger requests served along path $\theta \in \Theta_k^\tau$, and let $\delta(o, \theta)$ be the detour induced to the passenger request $o \in \mathscr{C}(\theta)$. Then, we can calculate the value of $\delta(o, \theta)$ by

$$\delta(o, \theta) = \begin{cases} (a_{o^-}^\theta - a_{o^+}^\theta) - (a_{o^-}^k - a_{o^+}^k), \ o \neq o_k, \\ (a_{o^-}^\theta - a_{o^+}^\theta) - d_{o^+ o^-}, \ o = o_k, \end{cases} \quad \forall o \in \mathscr{C}(\theta), \ \theta \in \Theta_k^\tau, \ k \in \mathscr{K}, \tag{3.2}$$

where $a_{o^+}^\theta$ and $a_{o^-}^\theta$ denote the arrival times at the pick-up and drop-off locations of request o along path θ. We can see that the above two sub-equations calculate the induced detour at epoch k for the previously confirmed ride requests and the new ride request, respectively. In the example illustrated in Figure 3.1(b) and 3.1(c), the new ride request o_k is considered for inclusion in two potential routing paths θ_1 and θ_2, respectively. For path θ_1, the arrival times at o_k's pick-up and drop-off locations are $a_{o_k^+}^{\theta_1} = 12$ and $a_{o_k^-}^{\theta_1} = 16$, respectively, resulting in no detour, i.e., $\delta(o_k, \theta_1) = 0$.

For path θ_2, the arrival times are $a^{\theta_2}_{o^+_k} = 14$ and $a^{\theta_2}_{o^-_k} = 20$, leading to a 2-min detour, i.e., $\delta(o_k, \theta_2) = 2$.

If $\delta(o, \theta) > 0$, compensation should be provided to ride request o along path θ. Let η_{ok} denote the amount of compensation offered to ride request o. Its decision space depends on the routing decision variables, which is formulated as follows:

$$0 \leq \eta_{ok} \leq M\delta(o,\theta)\tau_{\theta k},\ \forall o \in \mathscr{C}(\theta),\ \theta \in \Theta^{\tau}_{k},\ k \in \mathscr{K}, \tag{3.3}$$

where M is a sufficiently large number. These inequalities indicate that compensation is required only when the path inducing detours to passengers is selected.

Let $(\delta(o, \theta), \eta_{ok})$ be the detour option sent to passenger request o. We define a binary function $\psi(\beta_o, \delta(o, \theta), \eta_{ok})$ to model a passenger's acceptance of the detour option as follows:

$$\psi(\beta_o, \delta(o,\theta), \eta_{ok}) = \begin{cases} 1,\ \text{if } \eta_{ok} \geq \beta_o \delta(o,\theta), \\ 0,\ \text{otherwise}, \end{cases} \quad \forall o \in \mathscr{C}(\theta),\ \theta \in \Theta^{\tau}_{k},\ k \in \mathscr{K}. \tag{3.4}$$

We then have the following constraints for routing decision variables τ_k:

$$\tau_{\theta k} \leq \psi(\beta_o, \delta(o,\theta), \eta_{ok}),\ \forall o \in \mathscr{C}(\theta),\ \theta \in \Theta^{\tau}_{k},\ k \in \mathscr{K}, \tag{3.5}$$

which indicates that a path cannot be selected if the passenger rejects the detour option.

Exogenous information

The exogenous information $\mathscr{W}_{k+1}$ is the arrival of the next new request or the termination of the operation period. It triggers the subsequent epoch by revealing the next request o_{k+1} and its arrival time t_{k+1}. If no new requests appear by the end of the operation period T, the MDP will end, representing the termination of the service.

Transition function

The transition function $\mathscr{S}_{k+1} = \mathscr{F}(\mathscr{S}_k, x_k, \mathscr{W}_{k+1})$ transfers the current state $\mathscr{S}_k$

into $\mathscr{S}_{k+1}$ by decision x_k and exogenous information $\mathscr{W}_{k+1}$. The value t_{k+1} will be updated by the arrival of the new request o_{k+1} revealed by $\mathscr{W}_{k+1}$. Otherwise, the process will terminate, and $t_{k+1} = T$. The routing plan Θ_{k+1} will get updated from τ_k, excluding the locations visited before t_{k+1}.

Objective function

Let $\mathscr{R}(\mathscr{S}_k, x_k)$ denote the reward resulting from the decision in epoch k, which is calculated by the difference between the profits obtained by the decisions at the current epoch and the previous epoch as follows:

$$\mathscr{R}(\mathscr{S}_k, x_k) = r_{o_k}\zeta_k - (\sum_{\theta \in \Theta_k^\tau} c_\theta \tau_{\theta k} - \sum_{\theta \in \Theta_k} c_\theta) - \sum_{o \in \mathscr{O}_k} \eta_{ok}. \tag{3.6}$$

where c_θ denotes the travel cost of path θ. We can see that if request o_k is confirmed to be served, the realized reward will be its marginal profit, which is the request revenue minus extra travel cost (against the original routing plan Θ_k) and the amount of compensation accepted by passengers. Otherwise, the routing plan will not change, and the reward will be zero.

Let π be the policy that maps a state $\mathscr{S}_k$ to decision $x_k = \mathscr{X}^\pi(\mathscr{S}_k)$, where $\mathscr{X}^\pi(\mathscr{S}_k)$ is the decision rule derived from π, determining the decisions at each epoch k based on the current state $\mathscr{S}_k$. Let Π be the set of policies. We can formulate the objective function as finding the policy $\pi \in \Pi$ that maximizes the expected total reward conditional on the initial state as follows:

$$\max_{\pi \in \Pi} \mathbb{E} \left\{ \sum_{k=0}^{|\mathscr{K}|} \mathscr{R}(\mathscr{S}_k, \mathscr{X}^\pi(\mathscr{S}_k)) \middle| \mathscr{S}_0 \right\}. \tag{3.7}$$

3.2 Solution Method

The MDP can be solved by dynamic programming. According to Bellman's equation, the optimal policy can be found by addressing the following problem at each epoch k:

$$\max_{x_k} \mathscr{R}(\mathscr{S}_k, x_k) + V(\mathscr{S}_k, x_k), \tag{3.8}$$

where $V(\mathscr{S}_k, x_k) = \mathbb{E}\{\sum_{j=k+1}^{K} \mathscr{R}(\mathscr{S}_j, \mathscr{X}^\pi(\mathscr{S}_j))|\mathscr{S}_k, x_k\}$, which represents the expected profit that can be obtained based on state $\mathscr{S}_k$ and decision x_k. However, the dynamic programming method is intractable for large-scale problems due to the widely-known 'curse of dimensionality'. Therefore, we propose an anticipatory confirmation, compensation, and routing policy named ACCR based on approximate dynamic programming. Specifically, a routing heuristic is first proposed to generate potential feasible paths for serving the new request. We will then formulate an anticipatory model (AM) based on these new paths, and exploit the solution structure of a relaxed AM, referred to as R-AM. The findings and insights will be used to quickly find the optimal solution to AM in some cases. For the other cases, we further develop a compensation strategy to find good-quality solutions to AM. In addition, a new regression-tree-based VFA with a replay memory is developed to learn the reward-to-go value through simulation. The proposed policy ACCR is expected to generate good-quality solutions rapidly in dynamic scenarios.

In the following subsections, we first introduce the routing heuristic of ACCR, which is crucial to the formulation of AM. We will then present the AM and R-AM, exploit the structure of the model, and elaborate on the compensation strategy.

3.2.1 Routing heuristic

We design an insertion-based routing heuristic to rapidly identify potential routing plans for serving the new request. The algorithm will insert request o_k into any feasible locations along each path $\theta \in \Theta_k$ to generate a set of potential paths for serving request o_k. Upon determining a feasible insertion, the detour time induced for any affected passengers is then computed using Eq. (3.2). All newly generated paths are grouped in the set Θ_k^π, which is distinct from Θ_k^τ. While Θ_k^τ represents a general set of all potential routing plans, Θ_k^π specifically refers to the potential routing plans generated by the routing heuristic, which enforces specific feasibility conditions, making it a refined subset of Θ_k^τ. In addition to the request deadline constraint, those new paths will also satisfy ride-pooling and passenger-parcel consolidation constraints. Specifically, if request o_k is a passenger ride request, its origin and destination cannot be inserted into the trip of another ride request. If request o_k

is a delivery request, its pick-up and drop-off locations can not be inserted into the trips of two different ride requests. In this way, the insertion-based routing heuristic ensures that the insertion of request o_k into the current routing plan will not induce detours to multiple passenger requests.

If there are no feasible paths for serving the new request o_k, the request will be declined directly, and the current routing plan remains unchanged.

3.2.2 Anticipatory model

In this subsection, we formulate an anticipatory model (AM) to determine the approximate solution to problem (3.8) at each epoch k. The objective of the AM is to maximize the sum of the current reward $\mathscr{R}(\mathscr{S}_k, x_k)$ and the approximated future reward-to-go value $\hat{V}(\mathscr{S}_k, x_k)$ based on the newly generated paths in set Θ_k^π and the existing paths in set Θ_k. The value of reward-to-go $\hat{V}(\mathscr{S}_k, x_k)$ will be estimated by a VFA method elaborated in Section 5.3.

For ease of model building, we divide path set Θ_k^π into two subsets: one is $\Theta_k^{\pi 1}$, including all newly generated paths that do not induce detour to passengers, and the other is $\Theta_k^{\pi 2}$, including the rest of those paths that induce detour. Since each time the consolidation of passenger and parcel requests will affect one passenger at most, let p_θ denote the unique passenger request that is associated with a positive detour time $\delta_\theta > 0$ along path $\theta \in \Theta_k^{\pi 2}$. Let $\overline{\theta}$ denote the path in set Θ_k, from which the new path $\theta \in \Theta_k^\pi$ is generated by the insertion heuristic. Then the extra travel cost after the insertion of request o_k, denoted by $\Delta_\theta := c_\theta - c_{\overline{\theta}}$, $\forall \theta \in \Theta_k^\pi$, can be calculated instantly. Therefore, if a new path $\theta \in \Theta_k^\pi$ is selected to serve request o_k, while the other paths in set $\Theta_k \setminus \overline{\theta}$ remain unchanged, the realized reward $\mathscr{R}(\mathscr{S}_k, x_k)$ in Eq. (3.6) will be calculated by

$$z_\theta = \begin{cases} r_{o_k} - \Delta_\theta, \ \forall \theta \in \Theta_k^{\pi 1}, \\ r_{o_k} - \Delta_\theta - \eta_{p_\theta k}, \ \forall \theta \in \Theta_k^{\pi 2}. \end{cases} \tag{3.9}$$

In this modified reward calculation, the realized reward for serving request o_k is the revenue minus the marginal travel cost of the selected path, and is further reduced by the offered compensation if a detour path is chosen.

Given the above information, the AM at epoch k is formulated as follows:

[AM^k]

$$\max_{x_k} \sum_{\theta \in \Theta_k^\pi} z_\theta \tau_{\theta k} + \hat{V}(\mathscr{S}_k, x_k), \tag{3.10}$$

$$\text{s.t.} \sum_{\theta \in \Theta_k^\pi} \tau_{\theta k} = \zeta_k, \tag{3.11}$$

$$\sum_{\theta \in \Theta_k \cup \Theta_k^\pi} \phi_\theta^{o_k} \tau_{\theta k} = 1, \qquad \forall o \in \mathscr{O}_k^c, \tag{3.12}$$

$$\tau_{\theta k} \leq \psi(\beta_{p_\theta}, \delta_\theta, \eta_{p_\theta k}), \qquad \forall \theta \in \Theta_k^{\pi 2}, \tag{3.13}$$

$$\eta_{p_\theta k} \leq M \tau_{\theta k}, \qquad \forall \theta \in \Theta_k^{\pi 2}, \tag{3.14}$$

$$\eta_{p_\theta k} \geq 0, \qquad \forall \theta \in \Theta_k^{\pi 2}, \tag{3.15}$$

$$\zeta_k \in \{0, 1\}, \tag{3.16}$$

$$\tau_{\theta k} \in \{0, 1\}, \qquad \forall \theta \in \Theta_k^\pi. \tag{3.17}$$

The objective function in Eq. (3.10) is to select one potential routing plan to maximize the current reward and reward-to-go value. Constraint (3.11) connects the confirmation decision with the routing decision, which indicates that a newly generated path should be selected if the new request is confirmed to be served. Since the routing heuristic ensures that every $\theta \in \Theta_k^\pi$ serves the request o_k, the coefficient $\phi_\theta^{o_k}$ is always equal to 1 for these paths and can therefore be omitted for simplicity. Constraints (3.12) indicate that the confirmed requests should be served only once. Constraints (3.13) model the impact of the exogenous passenger's choice on the routing decision. If the passenger rejects the detour option, the corresponding routing plan can not be selected. Constraints (3.14) indicate that the compensation amount for the unselected path should be 0. Compared to Eq. (3.3), constraints (3.14) omits the coefficient $\delta(o, \theta)$ because only the set of paths $\Theta_k^{\pi 2}$ are considered. Since the routing heuristic ensures that each path in $\Theta_k^{\pi 2}$ have a detour component, and will not induce detours to multiple passenger requests, the coefficient $\delta(o, \theta)$ can be safely omitted without affecting the validity of the model. Constraints (3.15) to (3.17) define the domain of decision variables.

The uncertainty of passengers' responses in Eq. (3.13), which depends on the value of the compensation decision variable, makes it difficult to directly solve the model [AM^k]. Instead of addressing the problem directly, we will first solve a relaxed problem, R-AM, which ignores passenger acceptance constraints by assuming that the passengers would like to accept any detour regardless of the amount of compensation offered. Note that the purpose of formulating the relaxed model R-AM is to serves as a mediation to solve the original model AM. By exploring the relationship between solutions to AM and R-AM, as detailed in Proposition 1 and Observation 1, we can more easily identify the optimal solution to AM in certain cases. If not, we will further develop a compensation strategy to find good-quality solutions to AM.

By eliminating the passenger acceptance constraint, the R-AM at epoch k is formulated as follows:

[R-AM^k]

$$\max_{x_k}\ f(x_k) = \max_{x_k} \sum_{\theta \in \Theta_k^{\pi}} z_\theta \tau_{\theta k} + \hat{V}(\mathscr{S}_k, x_k) \tag{3.18}$$

$$\text{s.t. Eqs. } (3.11), (3.12), \text{and } (3.14) - (3.17).$$

We can see that, given the approximate reward-to-go value $\hat{V}(\mathscr{S}_k, x_k)$, the optimal solution to model [R-AM^k] will be chosen among the three groups of feasible solutions as follows:

- Group 1: Reject the new request and implement the original routing plan, i.e., $\zeta_k = 0,\ \tau_{\theta k} = 0,\ \forall \theta \in \Theta_k^{\pi},\ \tau_{\theta k} = 1,\ \forall \theta \in \Theta_k,\ \eta_k = 0.$
- Group 2: Serve the new request by a newly-generated path $\theta^* \in \Theta_k^{\pi 1}$ from its original path $\bar{\theta}^*$, without inducing a detour, and implement the original paths in set $\Theta_k \setminus \{\bar{\theta}^*\}$ for serving the other committed requests uncovered by this path, i.e., $\zeta_k = 1,\ \tau_{\theta^* k} = 1,\ \tau_{\theta k} = 0,\ \forall \theta \in \Theta_k^{\pi} \setminus \{\theta^*\} \cup \{\bar{\theta}^*\},\ \tau_{\theta k} = 1,\ \forall \theta \in \Theta_k \setminus \{\bar{\theta}^*\},\ \eta_k = 0.$
- Group 3: Serve the new request with a newly-generated path $\theta^* \in \Theta_k^{\pi 2}$ that

induces a detour, and implement the original paths in set $\Theta_k \setminus \{\bar{\theta}^*\}$ for serving the other committed requests uncovered by this path, i.e., $\zeta_k = 1$, $\tau_{\theta^* k} = 1$, $\tau_{\theta k} = 0,\ \forall \theta \in \Theta_k^\pi \setminus \{\theta^*\} \cup \{\bar{\theta}^*\}$, $\tau_{\theta k} = 1,\ \forall \theta \in \Theta_k \setminus \{\bar{\theta}^*\}$, $\eta_k \geq 0$.

Let Ω_1, Ω_2, and Ω_3 denote the corresponding feasible regions of the above three groups of solutions. For simplicity, we omit the index k in the notation, so we define $\Omega = (\Omega_1 \cup \Omega_2 \cup \Omega_3)$, which denotes the feasible regions of model [R-AM^k]. Correspondingly, we define the feasible region of model [AM^k] as $\hat{\Omega} = (\hat{\Omega}_1 \cup \hat{\Omega}_2 \cup \hat{\Omega}_3)$. Let $\hat{f}(x_k)$ and $f(x_k)$ denote the objective function values of models [AM^k] and [R-AM^k], respectively, and let x_k^* be the optimal solution to model [R-AM^k]. Then we have the following proposition:

Proposition 1: *If the optimal solution to model [R-AM^k] falls into Group 1 or Group 2, i.e., $x_k^* \in \Omega_1 \cup \Omega_2$, this solution will also be the optimal solution to model [AM^k], i.e., $x_k^* = arg\ max\ \hat{f}(x_k)$. If the optimal solution to model [R-AM^k] falls in Group 3, i.e., $x_k^* \in \Omega_3$, this solution will provide an upper bound for problem [AM^k], i.e., $\hat{f}(x_k) \leq \hat{f}(x_k^*), \forall x_k \in \hat{\Omega}$.*

Proof: Since the objective functions of models [AM^k] and [R-AM^k] are the same, we have

$$\hat{f}(x_k) = f(x_k),\ \forall x_k \in \Omega. \tag{3.19}$$

Because x_k^* is the optimal solution to model [R-AM^k], we have

$$f(x_k^*) \geq f(x_k),\ \forall x_k \in \Omega. \tag{3.20}$$

Therefore, the following relationship holds:

$$\hat{f}(x_k) = f(x_k) \leq f(x_k^*) = \hat{f}(x_k^*),\ \forall x_k \in \Omega. \tag{3.21}$$

Since model [R-AM^k] is a relaxation of model [AM^k] by ignoring Constraint (3.13), we have $\hat{\Omega}_1 = \Omega_1$, $\hat{\Omega}_2 = \Omega_2$, $\hat{\Omega}_3 \subset \Omega_3$, and accordingly, $\hat{\Omega} \subset \Omega$.

If $x_k^* \in \Omega_1 \cup \Omega_2$, it means that we have a feasible solution $x_k^* \in \hat{\Omega}$ for model [AM^k] such that

$$\hat{f}(x_k^*) \geq \hat{f}(x_k),\ \forall x_k \in \hat{\Omega}, \tag{3.22}$$

which indicates that the solution x_k^* is the optimal solution to model [AM^k].

If $x_k^* \in \Omega_3$, this solution may not be feasible for model [AM^k]. In other words, in case $x_k^* \in \Omega_3 \setminus \hat{\Omega}_3$, this solution provides an upper bound to model [AM^k]. $\square$

Based on this proposition, we will solve the model [R-AM^k] first and examine which group the optimal solution falls in. If it is in Group 1 or Group 2, we will implement the decision because this decision is optimal to model [AM^k]. Otherwise, we will further explore the solutions in Group 3 in the next subsection. Before proceeding to solve model [R-AM^k], we have the following observation:

Observation 1: *If the optimal solution to model [R-AM^k] falls in Group 3, i.e., $x_k^* \in \Omega_3$, the optimal compensation amount for model [R-AM^k] should be 0, i.e., $\eta_k^* = 0$.*

The above observation holds because the reward-to-go value $\hat{V}(\mathscr{S}_k, x_k)$ in objective function (3.18) will be affected only by the routing decision. Specifically, the reward-to-go can be rewritten as $\hat{V}(\mathscr{S}_k, \tau_{\theta k})$. If the compensation amount is positive, we can further improve the objective function value by reducing the compensation without violating any constraints. This means that we can eliminate the compensation decision without affecting the optimality of the original model. In other words, the model [R-AM^k] can be equivalently formulated as follows:
[R-AM-R^k]

$$\max_{(\zeta_k, \tau_{\theta k})} \sum_{\theta \in \Theta_k^\pi} (r_{o_k} - \Delta_\theta)\tau_{\theta k} + \hat{V}(\mathscr{S}_k, \tau_{\theta k}) \tag{3.23}$$

$$\text{s.t. Eqs. } (3.11), (3.12), (3.16), \text{and } (3.17).$$

According to the above finding, the optimal solution to model [R-AM^k] can thus be found by searching among all feasible path solutions in all three groups and identifying the one with the largest objective function value based on the following equation:

$$f^R(x_k) = \begin{cases} \hat{V}(\mathscr{S}_k, \tau_{\theta k}), \ \forall x_k \in \Omega_1, \\ r_{o_k} - \Delta_\theta + \hat{V}(\mathscr{S}_k, \tau_{\theta k}), \ \forall x_k \in \Omega_2 \cup \Omega_3, \end{cases} \quad \forall k \in \mathscr{K}. \tag{3.24}$$

Eq. (3.24) shows that the relaxed model [R-AM^k] computes the objective value as the reward-to-go value for rejecting o_k. For serving o_k, the objective value accounts for the immediate revenue, the reward-to-go value, and the extra travel cost of the selected path.

3.2.3 Compensation strategy

As aforementioned, if the optimal solution to model [R-AM^k] falls in Group 3, we need to further examine the solutions in Group 3 for model [AM^k]. In this case, the solutions in Groups 1 and 2 provide a lower bound. The best lower bound for model [AM^k] is $\underline{\hat{f}} = \max_{x_k \in \Omega_1 \cup \Omega_2} \hat{f}(x_k) = \max_{x_k \in \Omega_1 \cup \Omega_2} f^R(x_k)$, which can be readily obtained based on Eq. (3.24).

Since $\eta_{p_\theta k} \geq 0, \ \forall \theta \in \Theta_k^{\pi 2}$, we have $\hat{f}(x_k) \leq f^R(x_k), \ \forall x_k \in \Omega_3$, which means that $f^R(x_k)$ is the upper bound for each solution in Group 3, i.e., $x_k \in \Omega_3$. Therefore, if $f^R(x_k) \leq \underline{\hat{f}}$, the solution $x_k \in \Omega_3$ will not be better than the best solution in Groups 1 and 2, and we will not consider it anymore. Let $\Omega_3^{'}$ denote the subset of Ω_3 with solution x_k satisfying $f^R(x_k) > \underline{\hat{f}}$, i.e., $\Omega_3^{'} := \{x_k \in \Omega_3 | f^R(x_k) > \underline{\hat{f}}\}$. Hence, only a promising subset of newly generated paths in $\Theta_k^{\pi 2}$ associated with $\Omega_3^{'}$ is considered for compensation decision, which is denoted by $\Theta_k^{'\pi 2}$.

We can see from model [AM^k] that the compensation decision will influence both the current reward and the feasibility of a solution in terms of Constraint (3.13) via the value of $\psi(\beta_{p_\theta}, \delta_\theta, \eta_{p_\theta k})$. If a high compensation amount $\eta_{p_\theta k}$ is offered, the current reward will decrease while the detour option will be more likely to be accepted by passengers. If a low compensation amount $\eta_{p_\theta k}$ is offered, the current reward will increase but the detour option will be more likely to be rejected by passengers, rendering the solution infeasible. To balance the solution quality and feasibility, we define the probability of a passenger p_θ accepting the detour option $(\delta_\theta, \eta_{p_{\theta k}})$ as a function with respect to the compensation decision $\eta_{p_\theta k}$, i.e., $P(\beta_{p_\theta} \leq$

$\eta_{p_\theta k}/\delta_\theta)$. We will calculate the optimal compensation for each newly generated path in $\Theta_k^{\prime\pi 2}$, which is determined by maximizing the expected anticipatory profit as follows:

$$\max_{\eta_{p_\theta k}\geq 0} g(\eta_{p_\theta k}) = \max_{\eta_{p_\theta k}\geq 0} P(\beta_{p_\theta} \leq \eta_{p_\theta k}/\delta_\theta)(f^R(\tau_{\theta k}) - \eta_{p_\theta k}) + (1 - P(\beta_{p_\theta} \leq \eta_{p_\theta k}/\delta_\theta))\underline{\hat{f}},\ \forall \theta \in \Theta_k^{\prime\pi 2}. \tag{3.25}$$

Since the value of $f^R(x_k)$ is determined by $\tau_{\theta k}$ only, we rewrite $f^R(x_k)$ as $f^R(\tau_{\theta k})$ in the above equation for clarity. Note that $\underline{\hat{f}}$ is the objective function value of model [AM^k] of the baseline solution $\underline{x}_k^* = \arg\max_{x_k \in \Omega_1 \cup \Omega_2} \hat{f}(x_k)$, which is always feasible. The inclusion of $f^R(\tau_{\theta k})$ and $\underline{\hat{f}}$ in Eq. (3.25) enables the proposed compensation strategy to balance the trade-off between the expected profit from detour solutions (if accepted) and the fallback profit without detours (if rejected).

Let $\eta_{p_\theta k}^*$ denote the optimal solution to problem (3.25). Then we have the following proposition:

Proposition 2: *The optimal solution* $\eta_{p_\theta k}^*$ *falls in the interval* $[0, f^R(\tau_{\theta k}) - \underline{\hat{f}}]$.

Proof: When $\eta_{p_\theta k} \in [0, f^R(x_k) - \underline{\hat{f}}]$, we have $g(\eta_{p_\theta k}) \geq \underline{\hat{f}}$, which is the lower bound of problem (3.25).

Suppose that the optimal solution $\eta_{p_\theta k}^*$ does not fall in the interval $[0, f^R(x_k) - \underline{\hat{f}}]$, i.e., $\eta_{p_\theta k}^* > f^R(\tau_{\theta k}) - \underline{\hat{f}}$. Based on Eq. (3.25), we will have $g(\eta_{p_\theta k}^*) \leq \underline{\hat{f}}$. If $g(\eta_{p_\theta k}^*) < \underline{\hat{f}}$, it then contradicts the lower bound of the problem. If $g(\eta_{p_\theta k}^*) = \underline{\hat{f}}$, it shows no advantage over the objective of the baseline solution. Therefore, we have $\eta_{p_\theta k}^* \in [0, f^R(\tau_{\theta k}) - \underline{\hat{f}}]$. □

Based on this proposition, we employ a parameter $\lambda \in [0, 1]$ to calculate the compensation amount for each path considered in $\Theta_k^{\prime\pi 2}$ as follows:

$$\eta_{p_\theta k}^* = \lambda(f^R(\tau_{\theta k}) - \underline{\hat{f}}),\ \forall \theta \in \Theta_k^{\prime\pi 2},\ k \in \mathscr{K}. \tag{3.26}$$

This compensation strategy is applicable for any acceptance probability function $P(\beta_{p_\theta} \leq \eta_{p_\theta k}/\delta_\theta)$, which is adapted in generic scenarios without information of acceptance probability.

Next, we present a compensation strategy for a special case when β_{p_θ} follows the normal distribution with a mean μ and a standard deviation σ. In this case, $P(\beta_{p_\theta} \leq \eta_{p_\theta k}/\delta_\theta)$ could be represented as the cumulative distribution function of standard normal distribution $\Phi(\beta_{p_\theta} \leq \frac{\eta_{p_\theta k}-\delta_\theta\mu}{\delta_\theta\sigma})$, which can be approximated by the following logistic function $L(\beta_{p_\theta})$ with a maximum absolute error of 0.0095 (Bowling et al., 2009):

$$\Phi(x) \approx L(x) = \frac{1}{1+e^{-1.702x}}. \tag{3.27}$$

Let $\tilde{g}(\eta_{p_\theta k})$ be the approximation of $g(\eta_{p_\theta k})$ by replacing $\Phi(x)$ with $L(x)$ in Eq. (3.25). We have the following proposition:

Proposition 3: *There exists a value $\epsilon = 0.0095(f^R(\tau_{\theta k})-\underline{\hat{f}})$ such that the maximum absolute error between $\tilde{g}(\eta_{p_\theta k})$ and $g(\eta_{p_\theta k})$ is less than ϵ.*

Proof: We have

$$|\tilde{g}(\eta_{p_\theta k}) - g(\eta_{p_\theta k})| = |\Phi(\frac{\eta_{p_\theta k}-\delta_\theta\mu}{\delta_\theta\sigma}) - L(\frac{\eta_{p_\theta k}-\delta_\theta\mu}{\delta_\theta\sigma})|(f^R(\tau_{\theta k}) - \underline{\hat{f}} - \eta_{p_\theta k}). \tag{3.28}$$

Since $|\Phi(x) - L(x)| \leq 0.0095$ and $\eta_{p_\theta k} \in [0, f^R(\tau_{\theta k}) - \underline{\hat{f}}]$, it follows that

$$|\tilde{g}(\eta_{p_\theta k}) - g(\eta_{p_\theta k})| \leq 0.0095(f^R(\tau_{\theta k}) - \underline{\hat{f}}). \tag{3.29}$$

Thus, we have $\epsilon = 0.0095(f^R(\tau_{\theta k}) - \underline{\hat{f}})$. $\square$

It indicates that $\tilde{g}(\eta_{p_\theta k})$ can approximate $g(\eta_{p_\theta k})$ with a low margin of error. Hence, we can solve the following problem with the objective function $\max \tilde{g}(\eta_{p_\theta k})$ to approximate the optimal compensation amount $\eta^*_{p_\theta k}$:

$$\max_{\eta_{p_\theta k} \in [0, f^R(\tau_{\theta k}) - \underline{\hat{f}}]} \tilde{g}(\eta_{p_\theta k}) = \frac{f^R(\tau_{\theta k}) - \underline{\hat{f}} - \eta_{p_\theta k}}{1 + e^{\frac{-1.702\eta_{p_\theta k}}{\delta_\theta \sigma} + \frac{1.702\mu}{\sigma}}} + \underline{\hat{f}}. \tag{3.30}$$

The function $\tilde{g}(\eta_{p_\theta k})$ is non-concave since its second derivative is not strictly negative as the values of parameters change. For example, suppose that $f^R(\tau_{\theta k}) - \underline{\hat{f}} = 10$, $\delta_\theta = 3$, and $\mu/\sigma = 2$, we have $\tilde{g}''(0) > 0$, and $\tilde{g}''(1) < 0$. Therefore, the conventional convex programming method can not be directly applied to problem (3.30). Instead, we can obtain its optimal solution according to the following proposition:

Proposition 4: *The optimal solution to problem* (3.30) *is* $\eta^*_{p_\theta k} = \max\{0, f^R(\tau_{\theta k}) - \underline{\hat{f}} - \frac{\delta_\theta \sigma}{1.702}(1 + W(e^g))\}$*, where* $g = \frac{1.702(f^R(\tau_{\theta k}) - \underline{\hat{f}} - \delta_\theta \mu)}{\delta_\theta \sigma} - 1$*, and* $W(\cdot)$ *is the Lambert W function, which could be found by Newton's method.*

Proof: For ease of deduction, we define parameters $w := f^R(\tau_{\theta k}) - \underline{\hat{f}}$, $c := \frac{-1.702}{\delta_\theta \sigma}$ and $d := \frac{1.702\mu}{\sigma}$, and rewrite the problem (3.25) as follows:

$$\max_{\eta_{p_\theta k} \in [0,w]} \tilde{g}(\eta_{p_\theta k}) = \frac{w - \eta_{p_\theta k}}{1 + e^{c\eta_{p_\theta k} + d}} + \underline{\hat{f}}. \tag{3.31}$$

The first derivative of $\tilde{g}(\eta_{p_\theta k})$ is given by

$$\tilde{g}'(\eta_{p_\theta k}) = \frac{h(\eta_{p_\theta k})}{(1 + e^{c\eta_{p_\theta k} + d})^2}, \tag{3.32}$$

where $h(\eta_{p_\theta k}) = e^{c\eta_{p_\theta k} + d}(c(\eta_{p_\theta k} - w) - 1) - 1$. The sign of $\tilde{g}'(\eta_{p_\theta k})$ represents the monotonicity of $\tilde{g}(\eta_{p_\theta k})$. Since $(1 + e^{c\eta_{p_\theta k} + d})^2 > 0$, the sign of $\tilde{g}'(\eta_{p_\theta k})$ depends on the sign of $h(\eta_{p_\theta k})$.

The first derivative of $h(\eta_{p_\theta k})$ is given by

$$h'(\eta_{p_\theta k}) = c^2(\eta_{p_\theta k} - w)e^{c\eta_{p_\theta k} + d}. \tag{3.33}$$

Since $\eta_{p_\theta k} \leq w$, it follows that

$$h'(\eta_{p_\theta k}) < 0, \forall \eta_{p_\theta k} \in [0, w], \tag{3.34}$$

which indicates that $h(\eta_{p_\theta k})$ is monotonously decreasing. Let $\hat{\eta}_{p_\theta k}$ be the value such that $h(\hat{\eta}_{p_\theta k}) = 0$. Since $h(w) = -e^{c\eta_{p_\theta k} + d} - 1 < 0$, we have $\hat{\eta}_{p_\theta k} < w$. Then, if $\hat{\eta}_{p_\theta k} > 0$ we have

$$\begin{aligned} h'(\eta_{p_\theta k}) &> 0,\ \forall \eta_{p_\theta k} \in (0, \hat{\eta}_{p_\theta k}), \\ h'(\eta_{p_\theta k}) &< 0,\ \forall \eta_{p_\theta k} \in (\hat{\eta}_{p_\theta k}, w). \end{aligned} \tag{3.35}$$

According to Eq. (3.32), we have

$$\begin{aligned} \tilde{g}'(\eta_{p_\theta k}) &> 0,\ \forall \eta_{p_\theta k} \in (0, \hat{\eta}_{p_\theta k}), \\ \tilde{g}'(\eta_{p_\theta k}) &< 0,\ \forall \eta_{p_\theta k} \in (\hat{\eta}_{p_\theta k}, w). \end{aligned} \tag{3.36}$$

It indicates that $\tilde{g}(\eta_{p_\theta k})$ increases in the region $[0, \hat{\eta}_{p_\theta k}]$ and decreases in the region $[\hat{\eta}_{p_\theta k}, w]$. Hence, we have

$$\hat{\eta}_{p_\theta k} = \arg \max_{\eta_{p_\theta k} \in (0,w)} \ \tilde{g}(\eta_{p_\theta k}). \tag{3.37}$$

It indicates that $\hat{\eta}_{p_\theta k}$ is the optimal solution to problem (3.30) such that $\eta^*_{p_\theta k} = \hat{\eta}_{p_\theta k}$.

If $\hat{\eta}_{p_\theta k} \leq 0$, we have

$$\tilde{g}'(\eta_{p_\theta k}) < 0, \ \forall \eta_{p_\theta k} \in [0, w], \tag{3.38}$$

which indicates that $\tilde{g}(\eta_{p_\theta k})$ is decreasing in the feasible region. Then, the optimal solution to problem (3.30) is $\eta^*_{p_\theta k} = 0$.

In summary,

$$\eta^*_{p_\theta k} = \max(0, \hat{\eta}_{p_\theta k}). \tag{3.39}$$

Next, our objective is to obtain the value of $\hat{\eta}_{p_\theta k}$ such that $h(\hat{\eta}_{p_\theta k}) = 0$, which is expanded as follows:

$$e^{c\hat{\eta}_{p_\theta k}+d}(c(\hat{\eta}_{p_\theta k} - w) - 1) - 1 = 0. \tag{3.40}$$

Let $u = c(\hat{\eta}_{\theta k} - w) - 1$; and we have

$$\hat{\eta}_{\theta k} = \frac{u + cw + 1}{c}. \tag{3.41}$$

By substituting Eq. (3.41) into Eq. (3.40), we obtain

$$e^u u = \frac{1}{e^{cw+d+1}}, \tag{3.42}$$

which could be rewritten in Lambert form as follows:

$$u = W(\frac{1}{e^{cw+d+1}}). \tag{3.43}$$

Note that $W(\cdot)$ is the Lambert W function, whose value could be obtained by Newton's method (Corless et al., 1996). By substituting the value of u into Eq.

(3.41), we can obtain the value of $\hat{\eta}_{\theta k}$ as follows:

$$\hat{\eta}_{\theta k} = \frac{W(\frac{1}{e^{cw+d+1}}) + cw + 1}{c}. \tag{3.44}$$

By substituting the parameters w, c, and d defined in the very beginning, we can conclude that $\eta^*_{p_\theta k} = \max\{0, f^R(\tau_{\theta k}) - \underline{\hat{f}} - \frac{\delta_\theta \sigma}{1.702}(1 + W(e^g))\}$, where $g = \frac{1.702(f^R(\tau_{\theta k}) - \underline{\hat{f}} - \delta_\theta \mu)}{\delta_\theta \sigma} - 1$. $\square$

Based on the compensation strategy, the compensation amount can be calculated for each path considered in $\Omega_3^{'}$. The corresponding objective function value of [AM^k] will be $\hat{f}(x_k) = f^R(\tau_{\theta k}) - \eta^*_{p_\theta k}$. Nevertheless, x_k may not be feasible due to passengers' denial of the detour option. To address it, we will check all promising solutions regarding $\psi(\beta_{p_\theta}, \delta_\theta, \eta_{p_\theta k})$ one by one in descending order of the objective function value of [AM^k] until one solution is accepted. Otherwise, the baseline solution will be implemented. For real-world implementation, time limits for each passenger's response and for the entire process of inquiry can be set to ensure a timely response to each request.

The overall procedure of ACCR is outlined in Algorithm 3.1. At each decision epoch k, the potential paths for serving o_k are first generated by the routing heuristic, denoted as a function RH($\mathscr{S}_k$) (see Line 1). If there are no feasible paths, o_k would be rejected (see Lines 2-4). Otherwise, if the optimal solution to [R-AM-R^k] does not induce detours, it is optimal to [AM^k] (see Lines 5-8). If x^*_k induces a detour, the compensation strategy will be activated to calculate the compensation amount for promising detour paths (see Lines 9-18). To be more specific, we will obtain the optimal non-detour solution as a baseline solution, where objective value serves as a threshold to identify promising detour paths (see Line 9). The set Ω'_3 is initialized as an empty set. We will calculate the compensation amount of those detour paths with objective function values greater than the baseline value (see Lines 11-13). The corresponding decision variables will be grouped in Ω'_3 (see Lines 14-16). The passenger's acceptance of the compensation offer is checked through offering the detour option based on the solutions in Ω'_3, starting with the solution having the highest objective value of [AM^k]. The accepted solution will be implemented. Conversely, if the passenger rejects the detour offer for a solution x^*_k,

Algorithm 3.1 Pseudocode of ACCR

Input: States $\mathscr{S}_k$, compensation parameter λ

Output: Decisions x_k^*

$\Theta_k^\pi \leftarrow \mathrm{RH}(\mathscr{S}_k)$ ▷ Generate potential paths to serve o_k
if $\Theta_k^\pi = \varnothing$ **then**
  Return $x_k^* \leftarrow (\zeta_k = 0, \tau_k = \mathbf{0}, \eta_k = \mathbf{0})$ ▷ Reject o_k
end if
$x_k^* \leftarrow$ Solve [R-AM-R^k]
if $x_k^* \in \Omega_1 \cup \Omega_2$ **then**
  Return x_k^* ▷ Return optimal solution
end if
$\underline{x}_k^* \leftarrow \underset{x_k \in \Omega_1 \cup \Omega_2}{\arg\max}\ f^R(x_k)$; $\underline{\hat{f}} \leftarrow f^R(\underline{x}_k^*)$ ▷ Baseline solution
$\Omega_3' \leftarrow \varnothing$
for $\theta \in \Theta_k^{\pi 2}$ **do**
  if $f^R(\tau_{\theta k}) > \underline{\hat{f}}$ **then**
    $\eta_{p_\theta k}^* \leftarrow \lambda(f^R(\tau_{\theta k}) - \underline{\hat{f}})$ ▷ Compensation strategy
    $\tau_k \leftarrow (\tau_{\theta k} = 1, \tau_{\theta' k} = 0,\ \forall \theta' \in \Theta_k^\pi \setminus \{\theta\} \cup \{\bar{\theta}\}, \tau_{\theta' k} = 1,\ \forall \theta' \in \Theta_k \setminus \{\bar{\theta}\})$
    $x_k \leftarrow (\zeta_k = 1, \tau_k, \eta_{p_\theta k} = \eta_{p_\theta k}^*)$
    $\hat{f}(x_k) \leftarrow f^R(\tau_{\theta k}) - \eta_{p_\theta k}^*$
    $\Omega_3' \leftarrow \Omega_3' \cup \{x_k\}$
  end if
end for
while $\Omega_3' \neq \varnothing$ **do**
  $x_k^* \leftarrow \underset{x_k \in \Omega_3'}{\arg\max}\ \hat{f}(x_k)$
  if $\psi(\beta_{p_\theta}, \delta_\theta, \eta_{p_\theta k}) = 1$ **then** ▷ Check passenger's acceptance
    Return x_k^* ▷ Return optimal solution
  else
    $\Omega_3' \leftarrow \Omega_3' \setminus \{x_k^*\}$ ▷ Remove infeasible solution
  end if
end while
Return $x_k^* \leftarrow \underline{x}_k^*$ ▷ Return baseline solution

this solution is removed from Ω_3', and the next-best solution is selected for evaluation. If, at the end of this process, no acceptable solution remains, the baseline solution will be adopted (see Lines 19-27).

3.3 Value Function Approximation

The proposed ACCR requires an estimate of expected profit $\hat{V}(\mathscr{S}_k, x_k)$. Since the compensation decision only affects the current reward, and does not affect the fleet's ability of serving future requests, we can estimate the value function by the state and routing decisions, i.e., $\hat{V}(\mathscr{S}_k, \tau_{\theta k})$. We design a VFA to learn the value function $\hat{V}(\cdot)$, which takes $(\mathscr{S}_k, \tau_{\theta k})$ as input vector and outputs the value of expected profit. This idea is to simulate the MDP to collect observations for updating $\hat{V}(\cdot)$ iteratively. By iterative offline simulation, a well-trained $\hat{V}(\cdot)$ can be employed for online implementation. The VFA involves a state aggregation to extract representative attributes from $(\mathscr{S}_k, \tau_{\theta k})$ and a regression tree model to train $\hat{V}(\cdot)$ based on observations stored in a replay memory. We first introduce the state aggregation, and then elaborate on the training process.

3.3.1 State aggregation

The high dimensionality of input vector to $\hat{V}(\cdot)$, i.e., $(\mathscr{S}_k, \tau_{\theta k})$, can negatively impact the computational efficiency. Instead of using the original input vector, we consider the state of each vehicle. Let $\bar{\Theta}_k^{\tau}$ denote the set of paths based on routing decisions. We define the reward-to-go value of the fleet $\hat{V}(\mathscr{S}_k, \tau_{\theta k})$ as the sum of the values of the individual vehicles as follows:

$$\hat{V}(\mathscr{S}_k, \tau_{\theta k}) := \sum_{\theta \in \bar{\Theta}_k^{\tau}} \hat{V}(\theta), \ \forall k \in \mathscr{K}, \tag{3.45}$$

where $\hat{V}(\theta)$ represents the profit a vehicle is expected to obtain in the future based on the path θ at epoch k.

Next, we propose four representative attributes to represent the input vector based on θ. Let $\chi(\theta) = \{m_1, \cdots, m_{|\chi(\theta)|}\}$ be the sequential visiting locations along θ, and $a_{m_i}^{\theta}$ represent the planned arrival time at location $m_i \in \chi(\theta)$. Note that

$m_{|\chi(\theta)|}$ is the dummy node of depot $2|\mathscr{O}|+1$, as the vehicle is required to return to the depot at the end of the service period. We extract four representative attributes to aggregate the sequential and temporal information of path θ. The first attribute is the arrival time on the heading location $a^{\theta}_{m_1}$, which represents the earliest available time of the vehicle to serve other requests and can be obtained directly from path θ. The second attribute is the left time after visiting all locations in $\chi(\theta)$, denoted by $b^{\theta}_{k} = T - a_{m_{|\chi(\theta)|}}$. A larger value of b^{θ}_{k} indicates that more requests can be served in the future.

The first two attributes are widely used in the dynamic and stochastic vehicle routing problem (Ulmer, 2017). In addition to those two attributes, we also propose two new attributes considering the available time a vehicle could utilize to serve future requests along the planned path without violating the deadline of already confirmed requests. We refer to this available time as 'flexibility'. Let l_{m_i} denote the deadline for visiting location $m_i \in |\chi(\theta)|$, which is calculated by

$$l_{m_i} = \begin{cases} l_o, \text{ if } m_i = o^-, \\ l_o - d_{o^+o^-}, \text{ if } m_i = o^+, \end{cases} \quad \forall o \in \mathscr{O}_k, \ k \in \mathscr{K}. \tag{3.46}$$

It indicates that if the location m_i is the drop-off location of the request, l_o will be the deadline of this location. Otherwise, the location deadline will be the request deadline minus the direct travel time from m_i to its drop-off location. Let $\bar{a}^{\theta}_{m_i}$ denote the latest arrival time at location m_i in path θ, which is calculated as follows:

$$\begin{cases} \bar{a}^{\theta}_{m_i} = T, \text{ if } i = |\chi(\theta)|, \\ \bar{a}^{\theta}_{m_i} = \min\{\bar{a}_{m_{i+1}} - d_{m_i m_{i+1}}, \ l_{m_i}\}, \text{ if } i \in \{|\chi(\theta)|-1, \cdots, 1\}, \end{cases} \quad \forall \ m_i \in \chi(\theta), \ \theta \in \bar{\Theta}^{\tau}_{k}, \ k \in \mathscr{K}. \tag{3.47}$$

The flexibility of each location m_i in path θ is defined as the difference between the latest arrival time $\bar{a}^{\theta}_{m_i}$ and planned arrival time $a^{\theta}_{m_i}$, which indicates the available time to serve other requests without violating the deadline of follow-up requests after vising this location. Based on the flexibility, we propose two new attributes: the minimum flexibility $\bar{f}^{\theta}_{k}$ and the average flexibility f^{θ}_{k}.

The minimum flexibility $\bar{f}_k^\theta$ is the smallest differences between the latest arrival time $\bar{a}_{m_i}^\theta$ and planned arrival time $a_{m_i}^\theta$ along path θ, which is calculated by

$$\bar{f}_k^\theta = \min_{m_i \in \chi(\theta)} \{\bar{a}_{m_i}^\theta - a_{m_i}^\theta\}, \ \forall\, \theta \in \bar{\Theta}_k^\tau,\ k \in \mathscr{K}. \tag{3.48}$$

This value represents the available time to serve future requests before visiting the location with the minimum flexibility in the path. To represent the available time after that, we use the average flexibility f_k^θ that averages the differences between $\bar{a}_{m_i}^\theta$ and $a_{m_i}^\theta$ as follows:

$$f_k^\theta = \frac{1}{|\chi(\theta)|} \sum_{m_i \in \chi(\theta)} (\bar{a}_{m_i}^\theta - a_{m_i}^\theta), \ \forall\, \theta \in \bar{\Theta}_k^\tau,\ k \in \mathscr{K}. \tag{3.49}$$

3.3.2 Regression-tree-based VFA

Through state aggregation, the $\hat{V}(\theta)$ can be expressed as $\hat{V}(a_{m_1}^\theta, b_k^\theta, \bar{f}_k^\theta, f_k^\theta)$ such that the dimension of the input vector is significantly reduced. In what follows, we present the iterative offline simulation for learning $\hat{V}(\cdot)$.

Let Γ be the set of sample scenarios for offline simulation. The offline simulation begins with the policy π^1 underpinned by an initial value function $\hat{V}^1(\cdot)$. Then, the MDP is simulated based on sample scenarios from Γ iteratively. Within the iteration n, for the sample scenario from Γ, a policy π^n is applied drawing on the value function $\hat{V}^n(\cdot)$ at each decision epoch. The realized states, decisions, rewards, and transited states are collected for updating $\hat{V}^{n+1}(\cdot)$, which supports the decision-making in the next iteration. Finally, a well-trained $\hat{V}^{|\Gamma|+1}(\cdot)$ can be obtained after this iterative offline simulation for online implementation. To learn $\hat{V}(\cdot)$, previous research has adopted dynamic lookup table (DLT), which could result in an exponential growth of the scale of subsets for storing observations as the dimension of the input vector increases (Ulmer et al., 2018; Ulmer, 2020). To avoid this issue, we propose a regression tree model for its controllable subset scale and optimized subset creation (Loh, 2011). Both of them learn $\hat{V}(\cdot)$ by creating subsets of the observation where the values are more similar to each other within each subset. Given the input vector, $\hat{V}(\cdot)$ finds the subset the input vector belongs to, and return the average value of observations in the subset. However, the DLT is essentially a 2^n-ary tree regarding the dimension n of input vector. This feature could result in exponential scale creation of subsets as the dimension increases, which might reduce the learning effect. To avoid this issue, we employ a regression tree model for its capability of controllable scale of subsets and optimized creation of subsets for more accuracy (Loh, 2011). We illustrate the process of subset creation in DLT and regression tree in Figure 3.2. Subsets are represented as rectangular with two-dimensional input vector, which are filled with color blue, red, and white to represent that they are generated in the first, second, and third iteration, respectively. Initially, there is only one subset storing all observation. As the accumulation of observations, the variance of values in the subset might increase beyond a threshold, which indicates the similarity of observations is inadequate. Then, the subset would be divided into children subset to reduce the variance. Through the process, the DLT generates 10 subsets by equally dividing the parent subsets into 4 children subsets with respect to two dimension each time. In comparison, the regression tree only generates 4 subsets altogether because each division splits the subset by one dimension regarding the variance reduction (Loh, 2011) . Let N be the number of split. Generally, the DLT could generate $|N2^n - N + 1|$ subsets after N splitting, which is exponential to the dimension of input vector n. In comparison, the regression tree only produces $N+1$ subsets, independent of the dimension of input vector, which is more capable of addressing high dimensionality. Furthermore, regression tree divides the subset by maximizing the reduction of variance, which could learn split conditions adaptively.

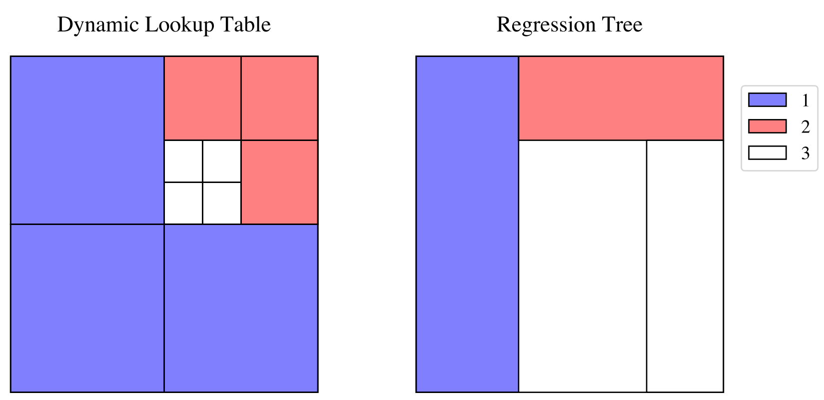


Figure 3.2. Comparision of dynamic lookup table and decision tree

To learn $\hat{V}(\cdot)$ by regression tree, a training dataset containing input vectors and realized profits is required. Given the sample scenario at iteration n, for each $\theta \in \Theta_k$ at epoch k, we calculate its realized profits backward at the end of MDP as follows:

$$v^n(a^\theta_{m_1}, b^\theta_k, \bar{f}^\theta_k, f^\theta_k) = \begin{cases} r_{o_k} - \Delta_{\theta'} - \eta_{p_{\theta' k}}, & \text{if } k = |\mathscr{K}^n|, \\ r_{o_k} - \Delta_{\theta'} - \eta_{p_{\theta' k}} + v^n(a^{\theta'}_{m_1}, b^{\theta'}_{k+1}, \bar{f}^{\theta'}_{k+1}, f^{\theta'}_{k+1}), & \text{otherwise,} \end{cases} \quad \forall k \in \mathscr{K}^n, \tag{3.50}$$

where $\theta' \in \bar{\Theta}^\tau_k$ is the selected potential routing path generated from θ. Considering that the policy is expected to improve through simulation, the accumulation of early poor observations might negatively affect the learning performance by the regression tree. To address this issue, we store all data points $((a^\theta_{m_1}, b^\theta_k, \bar{f}^\theta_k, f^\theta_k), v^n(a^\theta_{m_1}, b^\theta_k, \bar{f}^\theta_k, f^\theta_k))$ collected at iteration n in a replay memory P. The replay memory P is a first-in-first-out queue with capacity q, which discards early observations once its capacity is full. After each iteration, the regression tree will be updated based on the replay memory.

Algorithm 3.2 Training process of VFA

Input: Initialized value function $\hat{V}^1(\cdot)$
Output: Trained value function $\hat{V}^{|\Gamma|+1}(\cdot)$

$P \leftarrow \varnothing$
$n \leftarrow 1$
while $n \leq |\Gamma|$ **do**
  for $k = 1, \ldots, |\mathscr{K}^n|$ **do**
    $x_k \leftarrow \mathscr{X}^{\pi^n}(\mathscr{S}_k)$
    $\mathscr{S}_{k+1} = \mathscr{F}(\mathscr{S}_k, x_k, \mathscr{W}^n_{k+1})$
  end for
  Store the realized observations in P
  $\hat{V}^{n+1}(\cdot) \leftarrow$ Regression tree(P)
  $n \leftarrow n + 1$
end while
Return $\hat{V}^{|\Gamma|+1}(\cdot)$

The training process of $\hat{V}(\cdot)$ is outlined in Algorithm 3.2. The value function $\hat{V}^1(\cdot)$ is initialized as an empty regression tree that outputs 0 for any input. It indicates that ACCR is initialized as a myopic policy focusing only on the current reward, and gradually learns future profitability through simulation. At each iteration n, the MDP is simulated by applying policy π^n on one sample scenario. The realized profits are calculated based on the observations at the end of the MDP by Eq. (3.50), and the corresponding data points will be stored in the replay memory P. The $\hat{V}^{n+1}(\cdot)$ is obtained by training the regression tree model based on the data points in P, which is employed for the next iteration. After $|\Gamma|$ iterations, a well-trained $\hat{V}(\cdot)$ will be obtained for online implementation.

3.4 Numerical Experiments

This section presents the numerical experiments to evaluate the performance of the proposed ACCR. We first describe instance generation and parameter setting. Then, we will evaluate the performance of the ACCR against several benchmark policies, and evaluate the effectiveness of the proposed regression-tree-based VFA. Moreover, we will explore the advantages of the full integration in CUT.

3.4.1 Instance generation and parameter setting

In order to abstractly model a typical urban district for operational analysis, we consider a 7.5×7.5 km^2 square as the service area with the depot located in the center at location $(3.75, 3.75)$, which approximates the actual area of a core urban district like Shanghai's Xuhui, ensuring realistic spatial scale. The operation duration is set to be 360 min. We consider three different fleet sizes: 3, 5, and 7 vehicles. The travel cost is 2 units per km. The distance $\bar{d}_{ij}$ between locations i and j is computed as Euclidean distance. By assuming the constant speed 30 km/h, the travel time d_{ij} can be obtained readily. Regarding the number of requests, we consider three different settings: 300, 400, and 500 requests. The numbers of ride requests and delivery requests are the same, each accounting for half of the number of all requests. For their spatial distribution, we consider both uniform and normal distributions. Under uniform distribution, both pick-up and drop-off locations are generated randomly in the service area. Under normal distribution, the locations follow a normal distribution along the x and y coordinates with a mean of 3.75 and a standard deviation of 2. The arrival time of requests is uniformly distributed in the time window $[0, 320]$. The deadline l_o of ride request o is set to be $(e_o + d_{o^+o^-} + 20)$ min. The deadline of parcel request o is set to be $(e_o + d_{o^+o^-} + 60)$ min. The revenue of ride request o is calculated by $(15 + 10\bar{d}_{o^+o^-})$, and the revenue of delivery request o is calculated by $(5 + 2\bar{d}_{o^+o^-})$, both of which include a fixed component and a distance-dependent variable component. Regarding the individual detour sensitivity β_o, we assume that it follows a normal distribution with a mean $\mu = 1$ and a standard deviation $\sigma = 0.2$. Considering different combinations of fleet size, number of requests, and spatial distribution, a total of 18 instance settings are generated. For each instance setting, we evaluate the proposed ACCR by generating $1,000$ random sample scenarios. Regarding the parameters of the proposed VFA, the capacity q of replay memory is set to be 10,000; the minimum number of samples in a leaf node is set to be 200; and the minimum decrease of impurity is set to be 0.1.

3.4.2 Comparison results

In this subsection, we will conduct numerical experiments to evaluate the performance of the proposed ACCR. We will then assess the effectiveness of critical components in the proposed VFA, i.e., representative attributes and regression tree.

To evaluate the performance of the ACCR, we compare the results obtained by the ACCR with those obtained by five benchmark policies. The first is a myopic version of the ACCR, referred to as the MYP, considering the current reward only. The second is a cost-benefit policy with fixed compensation ratio (CBH-F), which represents a rule-based strategy that practitioners could implement. Specifically, it evaluates requests by comparing marginal profits to costs for each potential routing plan generated by the routing heuristic. Only requests with a profit-to-cost ratio above a set threshold are considered eligible for service. The compensation for detour routes is calculated by multiplying the detour by a fixed parameter. Then, the system selects the routing plan that yields the largest marginal profits to serve the request. The third policy, CBH-C, uses the same cost-benefit approach but applies our proposed compensation strategy without using VFA for future profitability. The details of the cost-benefit approach can be found in Ulmer et al. (2015). The fourth, named the A-DLT2, adopts the ACCR framework but with a different VFA method proposed in a previous study (Ulmer et al., 2018). It employs a dynamic look-up table (DLT) to learn $\hat{V}(\cdot)$, and uses the first two commonly used attributes $a^{\theta}_{m_1}$

and b_k^θ to aggregate the state of each path θ. The A-DLT2 serves as an alternative approximation within the ACCR framework, enabling a direct comparison between the proposed regression tree-based VFA and another approximation method. The fifth benchmark policy is also an anticipatory policy based on a deep Q network (DQN), denoted as the A-DQN. The DQN-based policy has shown efficiency in stochastic dynamic vehicle routing problems in the previous literature (Hildebrandt et al., 2023). For our problem, the A-DQN utilizes a DQN to estimate the Q-value of actions, which represents the expected reward from taking an action in a state. The action with the largest Q-value would be chosen. The DQN is essentially a deep neural network composed of one input layer, two hidden layers, and one output layer. The input vector to the DQN is composed of the current time, the representative attributes of each vehicle, and the marginal cost of serving the request by direct ride and detour ride. To avoid the explosion of the output layer of the DQN due to large action space, the action space is limited to four classes, including rejection, ride service without detour, ride service with a small detour, and ride service with a large detour. The ride service with a small detour refers to the case with the detour time less than 5 min, while the large detour refers to the case with the detour time exceeding 5 min. The amount of compensation for the detour ride is calculated by the proposed compensation strategy, given the Q-value of the detour ride and the baseline action with the largest Q-value. The parameters of the DQN are learned during the training process with 100,000 training epochs to ensure convergence.

Table 3.2 shows the comparison results for all instance settings, represented by the number of requests (#N), fleet size (#M), and the distribution type (SD) with the value of 0 for uniform distribution and value of 1 for normal distribution. The best objective values are highlighted in bold. The results demonstrate that the ACCR policy outperforms the others, achieving the best objective values in all instances. On average, the ACCR achieves a 55.0% higher profit compared to the MYP. When compared to the CBH-F and CBH-C, the ACCR yields 51.1% and 50.0% higher profits, respectively. Additionally, the ACCR outperforms the A-DLT2 with a 45.6% increase in profits.

We visualize the average profit each policy obtains in each fleet size setting as

Table 3.2. Comparison results of ACCR and benchmark policies

#N	#M	SD	ACCR	MYP	CBH-F	CBH-C	A-DLT2	A-DQN
300	3	0	**2,719**	1,674	1,735	1,757	1,814	1,586
300	5	0	**4,189**	3,556	3,581	3,617	3,608	3,043
300	7	0	**5,623**	5,606	5,601	5,614	5,449	5,518
400	3	0	**3,002**	1,431	1,493	1,521	1,618	2,046
400	5	0	**4,750**	3,128	3,190	3,200	3,320	2,620
400	7	0	**6,022**	5,240	5,284	5,318	5,295	5,332
500	3	0	**2,977**	1,290	1,359	1,361	1,494	2,676
500	5	0	**5,081**	2,780	2,845	2,834	3,047	3,240
500	7	0	**6,316**	4,789	4,867	4,863	5,008	5,097
300	3	1	**2,705**	1,695	1,755	1,793	1,788	1,450
300	5	1	**4,295**	3,370	3,411	3,447	3,420	3,867
300	7	1	**5,506**	5,187	5,169	5,225	5,074	4,528
400	3	1	**2,845**	1,493	1,556	1,579	1,625	2,532
400	5	1	**4,495**	3,065	3,111	3,130	3,191	3,090
400	7	1	**6,243**	4,888	4,924	4,947	4,934	5,177
500	3	1	**2,819**	1,340	1,415	1,424	1,509	1,911
500	5	1	**4,875**	2,786	2,848	2,849	2,991	3,429
500	7	1	**6,543**	4,547	4,600	4,606	4,705	4,169

the number of requests increases in Figure 3.3. Interestingly, as the number of requests grows under each fleet size setting, the profits obtained by MYP, CBH-F, and CBH-C decreased by around 9.7%, 9.4%, and 8.9%, respectively. A-DLT2 decreased by around 6.6% on average, and the profits obtained by A-DQN fluctuated with unstable performance. This indicates that these policies may not be able to maintain robust performance in an expanding demand market. In contrast, ACCR achieved an average profit increase of 6.4% under the same conditions, demonstrating its ability to effectively respond to an increased demand.

For ease of comparison, we present the relative gaps between the ACCR and benchmark policies, denoted by Λ-MYP, Λ-CBH-F, Λ-CBH-C, Λ-A-DLT2, Λ-A-DQN, respectively in Table 3.3. The gaps are calculated as $\frac{obj^{*}-obj}{obj} \times 100\%$, where obj^{*} represents the objective value obtained by ACCR, and obj represents the objective values obtained by the benchmark policies. We illustrate the variation of the average relative gaps with the increase of request number in Figure 3.4.

Table 3.3. Comparison of gaps between ACCR and benchmark policies

#N	#M	SD	Λ-MYP(%)	Λ-CBH-F(%)	Λ-CBH-C(%)	Λ-A-DLT2(%)	Λ-A-DQN(%)
300	3	0	62.4	56.8	54.8	49.9	71.5
300	5	0	17.8	17.0	15.8	16.1	37.7
300	7	0	0.3	0.4	0.2	3.2	1.9
400	3	0	109.8	101.1	97.5	85.5	46.8
400	5	0	51.9	48.9	48.4	43.1	81.3
400	7	0	14.9	14.0	13.2	13.7	12.9
500	3	0	130.7	119.1	118.8	99.2	11.3
500	5	0	82.8	78.6	79.3	66.7	56.8
500	7	0	31.9	29.8	29.9	26.1	23.9
300	3	1	59.6	54.1	50.9	51.2	86.6
300	5	1	27.4	25.9	24.6	25.6	11.1
300	7	1	6.2	6.5	5.4	8.5	21.6
400	3	1	90.5	82.9	80.2	75.1	12.4
400	5	1	46.7	44.5	43.6	40.8	45.5
400	7	1	27.7	26.8	26.2	26.5	20.6
500	3	1	110.3	99.3	98.0	86.8	47.5
500	5	1	75.0	71.2	71.1	63.0	42.2
500	7	1	43.9	42.3	42.1	39.1	57.0

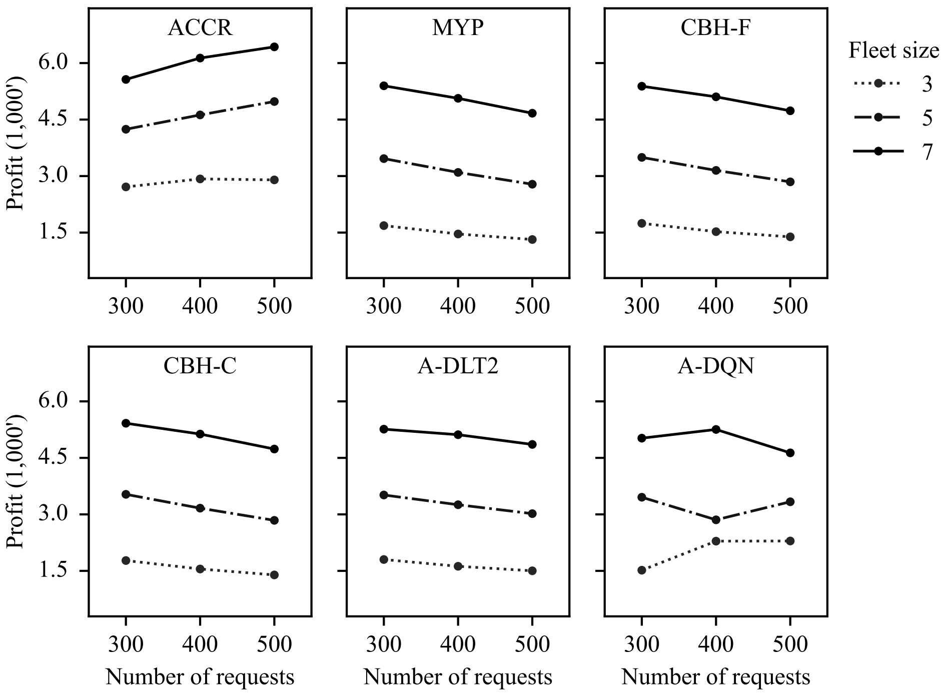


Figure 3.3. Impact of the increased number of requests on the policy's performance

It shows that as the number of requests increased from 300 to 500, the advantage regarding the solution quality of ACCR over MYP, CBH-F, CBH-C, and A-DLT2 became more apparent. Specifically, Λ-MYP rises from 29.0% to 79.1%,

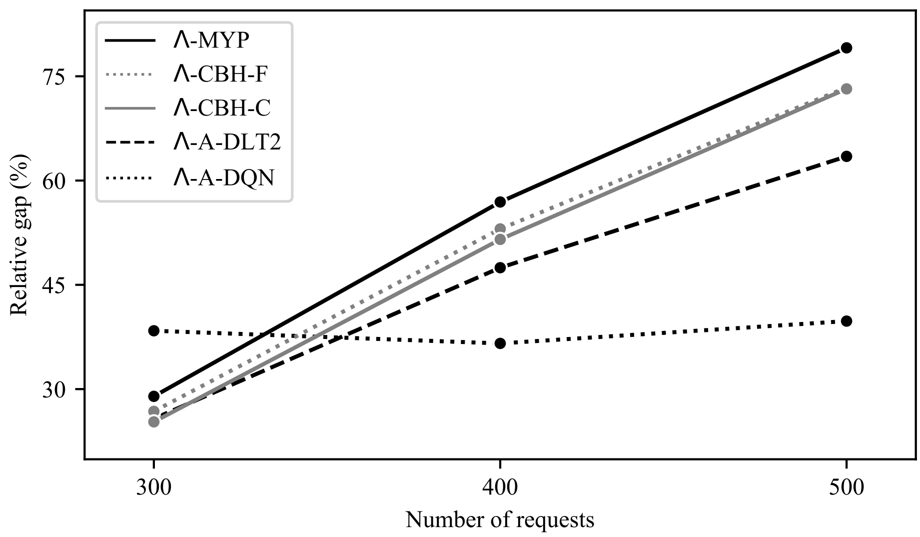


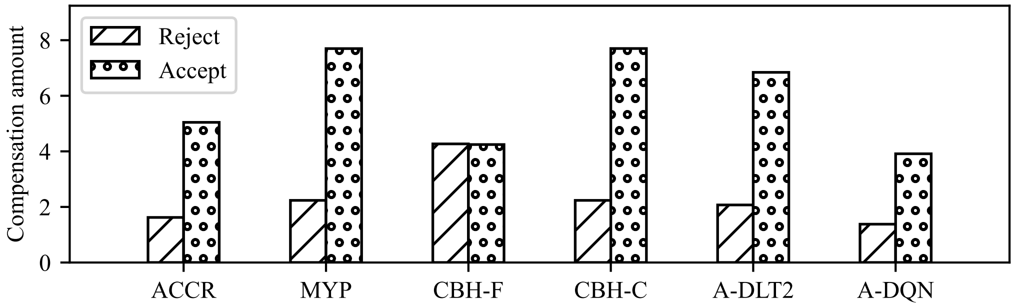
Figure 3.4. Relative gaps with the increase in the number of requests

Λ-CBH-F increases from 26.8% to 73.4%, Λ-CBH-C increases from 25.3% to 73.2%, and Λ-A-DLT2 grows from 25.7% to 63.5%. In contrast, the advantage of ACCR over A-DQN appears insensitive to the number of requests, and on average ACCR gains 38.2% more profits than A-DQN. The results reveal that MYP, the myopic version of ACCR, underperforms compared to the practical benchmarks CBH-F and CBH-C. However, when ACCR incorporates a DLT-based VFA designed for general SDVRPs, it significantly outperforms these practical benchmarks, demonstrating that even with a general-purpose VFA, the anticipatory policy is more effective. Furthermore, ACCR achieves its best performance when using our proposed regression-tree-based VFA, confirming the proposed VFA's effectiveness. Additionally, ACCR consistently outperforms A-DQN. These findings collectively demonstrate the effectiveness of both the anticipatory strategy and the proposed VFA.

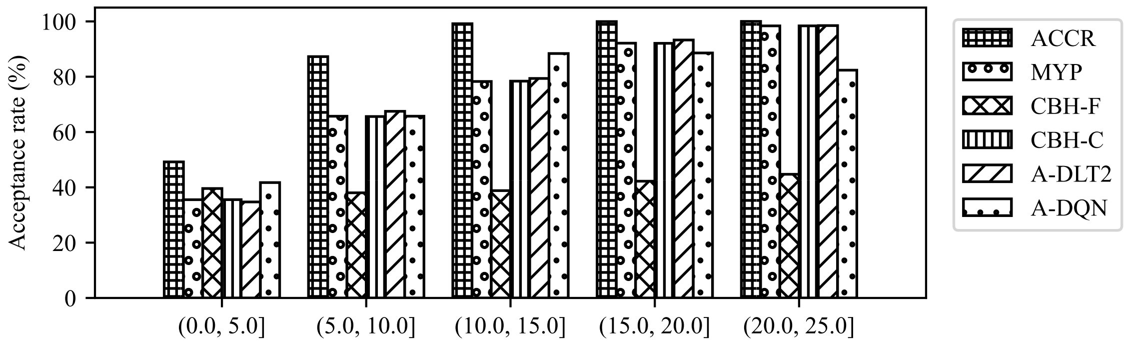
In the following, we analyze the results in detail. First, we show the value of ACCR by analyzing its effectiveness of capturing consolidation opportunities, compared with benchmark polices. Then, we evaluate the efficacy of the proposed VFA on the performance of the ACCR. Finally, we conduct sensitivity analysis of passengers' detour sensitivity parameters on the system's performance.

3.4.3 Value of ACCR

In this part, we analyze ACCR's effectiveness of capturing consolidation opportunities, compared with benchmark policies. We first analyze the relationship between offered compensation amount and acceptance rate through different polices in Figure 3.5. The acceptance rate is defined as the proportion of detour options accepted by passengers out of all detour offers made. Since different policies could offer different detour options, the observed acceptance rate reflects the ability of each policy to catch consolidation opportunities by offering favorable detour options. Figure 3.5(a) shows that the average accepted compensation amount of polices MYP and CBH-C are the highest. The reason might be that the overlook of the reward-to-go value could overestimate the benefit of consolidation, leading to offering high compensation amount by the proposed compensation strategy. With the improved anticipation of reward-to-go value, the average accepted compensation amount reduces. However, the average accepted compensation amount under ACCR is not the highest or the lowest compared to other benchmark polices, which is 1.8 to 2.7 lower, compared to the policies MYP, CBH-C, and A-DLT2, and is 0.8 to 1.1 higher, compared to the polices CBH-F and A-DQN. Nevertheless, the ACCR achieves the highest acceptance rate of detour options within all compensation amount intervals, as shown in Figure 3.5(b). Specifically, within the interval of (0.0, 5.0], the ACCR achieves around 50% acceptance rate, while others achieve at most 42%. The acceptance rate under ACCR improves to a level of 87% within the interval of (5.0, 10.0], while others achieve at most 67%. With the further increasing of compensation amount, the acceptance rate under ACCR achieves almost 100%. Even though other polices, such as MYP, CBH-C, and A-DLT2, offer higher compensation mount, they do not achieve higher acceptance rate compared to the ACCR. Low compensation amount would naturally lead to rejection of detour options, but the ACCR can offer appropriate compensation amount through a well estimation of the benefits of consolidation, leading to higher acceptance rate of detour options.



(a) Average accepted and rejected compensation amount



(b) Acceptance ratio within different compensation amount intervals

Figure 3.5. Compensation amount and acceptance rate under the ACCR and benchmark policies

Figure 3.6 illustrates the onboard rate, defined as the ratio of parcels delivered with passengers on board to the total number of delivered parcels, as the number of requests increases under the ACCR and benchmark policies. The ACCR achieves the highest onboard rate among all benchmark polices. Specifically, with the number of requests increase from 300 to 500, the onboard rate under ACCR increases from around 62% to 70%, while the rate under other benchmark polices even decreases from around 31% to 22%. It indicates that the ACCR has a better ability of capturing the consolidation opportunities during the operation, and is scalable to an expanding market.

3.4.4 Effectiveness of regression-tree-based VFA

To evaluate the efficacy of the regression tree, we compared the performance of using ACCR and the policy that uses DLT-based VFA within the same policy framework, employing either two commonly used attributes, $a^{\theta}_{n_1}$ and b^{θ}_{k}, or all four

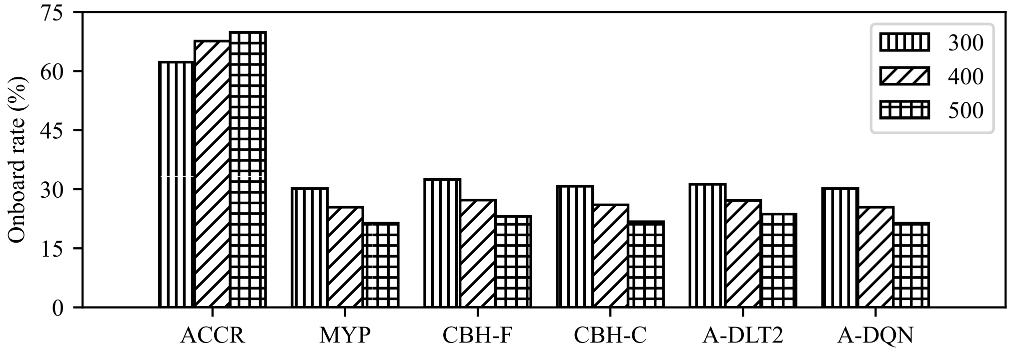


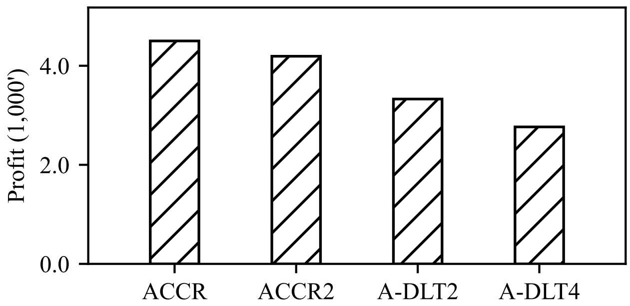
Figure 3.6. Onboard rate with increased number of requests under the ACCR and benchmark polices

attributes. We refer to the policy that uses the regression-tree-based VFA with two attributes as ACCR2. We refer to the policy that uses the DLT-based VFA with four attributes as A-DLT4.

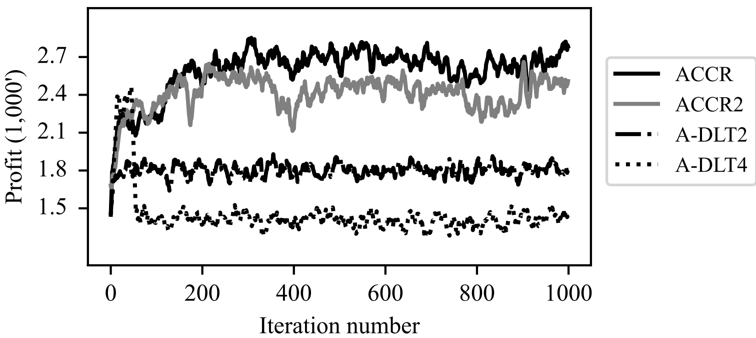
Table 3.4. Comparison results of ACCR and ACCR2

#N	#M	SD	ACCR	ACCR2	Λ-ACCR2 (%)
300	3	0	**2,719**	2,655	2.4
300	5	0	**4,189**	4,151	0.9
300	7	0	**5,623**	5,101	10.2
400	3	0	**3,002**	2,868	4.7
400	5	0	**4,750**	4,576	3.8
400	7	0	**6,022**	5,669	6.2
500	3	0	**2,977**	2,851	4.4
500	5	0	**5,081**	4,744	7.1
500	7	0	**6,316**	6,112	3.3
300	3	1	**2,705**	2,465	9.8
300	5	1	**4,295**	4,019	6.9
300	7	1	**5,506**	5,153	6.8
400	3	1	**2,845**	2,600	9.4
400	5	1	**4,495**	4,249	5.8
400	7	1	**6,243**	5,835	7.0
500	3	1	**2,819**	2,439	15.6
500	5	1	**4,875**	4,234	15.1
500	7	1	**6,543**	5,724	14.3

The comparison results between ACCR and ACCR2 are tabulated in Table 3.4. The relative gap Λ-ACCR2 is also computed. The comparison results are illustrated in Figure 4.1. Figure 3.7(a) shows the average profits obtained by ACCR, A-DLT4, ACCR2, and A-DLT2, under all instance settings. The performance of A-DLT4 is significantly inferior to ACCR, gaining 38.6% less profits on average, and is even worse than A-DLT2, with 16.9% less profits. Moreover, even with just two attributes, ACCR2 outperforms A-DLT2, yielding 26.2% more profits. It indicates the efficacy of the proposed regression-tree-based in enhancing policy performance.



(a) Profits obtained by ACCR, ACCR2, A-DLT2, and A-DLT4



(b) Training process of ACCR, ACCR2, A-DLT2, and A-DLT4

Figure 3.7. Comparison of ACCR, ACCR2, A-DLT2, and A-DLT4

We further illustrate the training process of the four policies in the instance setting with 300 requests and 3 vehicles under the uniform spatial distribution in Figure 3.7(b) (The observed fluctuations are attributed to the variability introduced by different instance realizations). It shows that the performance of A-DLT4 improves initially, but then deteriorates drastically, converging at 46.5% and 21.1% below ACCR and A-DLT2, respectively. In contrast, the performance of ACCR and ACCR2 is enhanced consistently to a significantly higher level. In summary, the proposed regression-tree-based VFA successfully handles the increased dimension of representative attributes, contributing to an excellent policy performance. Conversely, the DLT-based VFA, as used in A-DLT4, suffers from the increased dimensionality, resulting in significantly worse policy performance.

By employing four attributes, ACCR outperforms ACCR2 achieving 7.4% more profits on average, as shown in Figure 3.7(a). We further visualize the variation of relative gap Λ-ACCR2 with the increasing numbers of requests in Figure 3.8. It shows that compared to ACCR2, ACCR obtains 6.0% more profits when the numbers of requests are 300 and 400, and 10.0% more profits when the number of requests increases to 500. The results validate the efficacy of the proposed two new attributes for the VFA, which could improve the performance of the proposed ACCR.

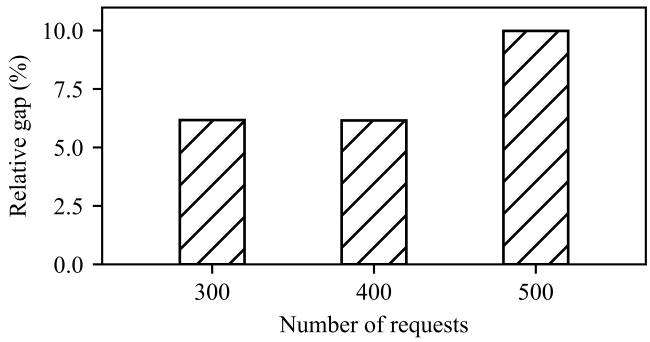


Figure 3.8. Relative gaps between ACCR and ACCR2 with the increase in the number of requests

3.4.5 Analysis of detour sensitivity

To explore the effect of detour sensitivity on the combined transportation system, we conducted a sensitivity analysis of the parameter β. Specifically, we vary β across a range of values to reflect scenarios where passengers are more averse (higher β) or more inclined (lower β) to accept detours. The analysis examines the impact of the parameter μ, which represents the mean value of the random variable β in the normal distribution for individual passengers, on system performance, from the perspectives of compensation amount, the extent of combined transportation, and profits. The μ ranges from 1 to 10, with higher values indicating passengers being more averse to the detour.

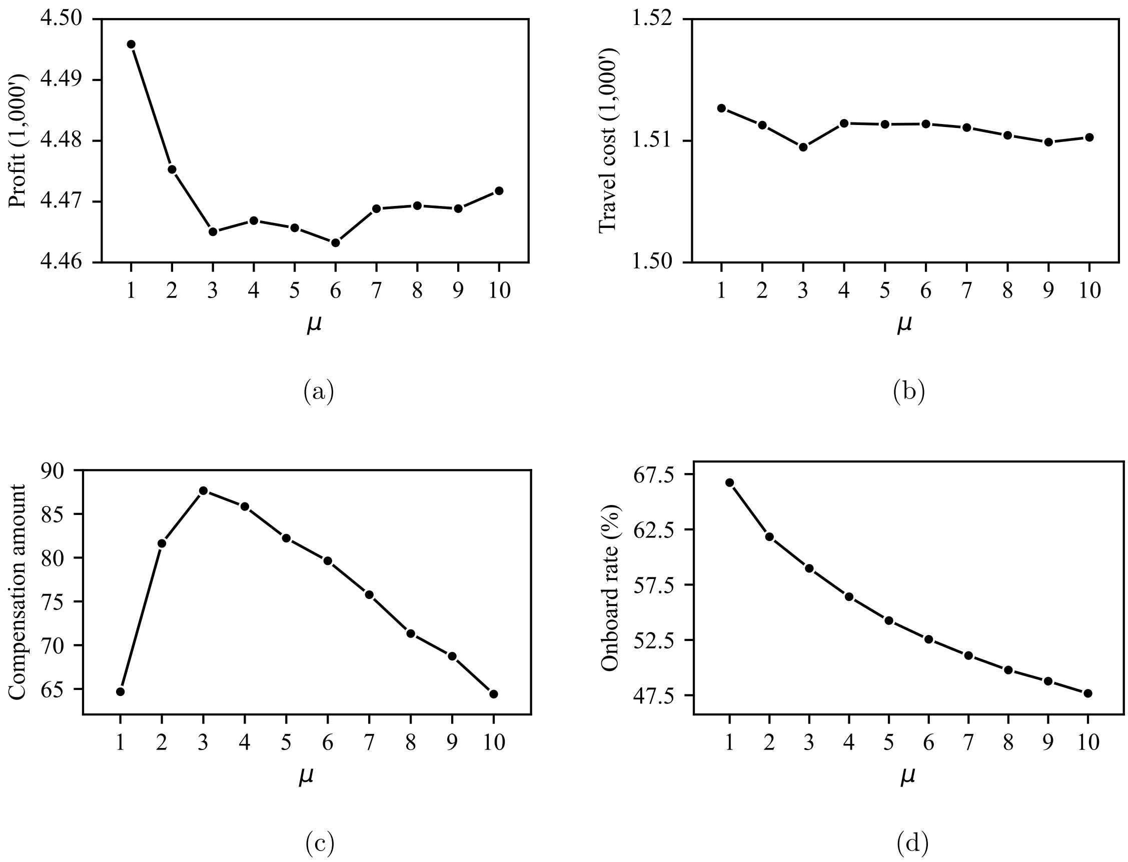


Figure 3.9. Sensitivity analysis of μ on (a) profits, (b) travel costs, (b) compensation amount, and (d) onboard rate

Figure 3.9(a) shows that when passenger sensitivity to detours increases from 1 to 3, system profits decrease by about 1%. However, as sensitivity continues to increase, system profits remain stable without significant further decline. Additionally, travel costs remain constant as passengers' inclination to detours changes, as illustrated in Figure 3.9(b). It indicates that under ACCR, the system could adaptively regulate the progress of combined transportation, ensuring stable system performance even with varying levels of detour sensitivity. Figure 3.9(c) displays that when μ increases from 1 to 3, the system's compensation expenditure rises by 33.0%. However, as μ continues to increase, the total compensation amount gradually decreases and returns to the original level. Meanwhile, the onboard rate decreases progressively with the increase in detour sensitivity, dropping from 67.5% to 47.5%, illustrated in Figure 3.9(d). This indicates that when passengers start to be more sensitive to detours, the system increases compensation expenditures to capture opportunities for consolidation. However, as passengers are more adverse to the detour, certain consolidation opportunities could be abandoned strategically. This leads to more parcels being delivered without passengers on board and causes the system to operate in a more moderate setting.

The key finding indicates that under ACCR, the system could adaptively regulate the progress of combined transportation, ensuring stable system performance even with varying levels of detour sensitivity. As detour sensitivity initially increases, the system raises compensation expenditures to capture consolidation opportunities. However, as passengers become more adverse to detours, some consolidation opportunities are strategically abandoned, causing the system to operate in a more moderate setting.

We conduct an extended analysis to explore how incorporating the direct travel time into the acceptance criterion influences system performance, where passengers tend to be more resistant to detours during shorter trips, while they are more willing to accept detours as travel time increases. To model this relationship, we used an exponential decay function:

$$\mu_o = N_0 + (N_1 - N_0)e^{-\alpha d_{o^+o^-}}, \tag{3.51}$$

which indicates that average passengers' detour sensitivity is initially highest at N_1 and gradually decreases to N_0 as the direct travel time from its origin to destination increases. This model is influenced by the parameter α, where a large value of α indicates that an increase in travel time has a stronger effect on reducing passengers detour sensitivity. We selected α values of $\ln(2)/10$, $\ln(2)/5$, $\ln(2)/2$ to represent the low, medium, and high levels of the effect of increases in travel time on reducing passengers' detour sensitivity. Figure 3.10(a) illustrates that as the effect shifts from low to medium, system profits initially stabilize without significant increases. It suggests that the operator should balance the cost of incentives for passengers to accept detours against the diminishing returns of profits. When the effect becomes high, profits increases by 0.5%, indicating that the operator should prioritize making detours more attractive to passengers on long trips to further improve profits. Meanwhile, the travel cost remains constant as passengers' inclination to detours regarding the travel time changes, illustrated in Figure 3.10(b). Figure 3.10(c) shows that the transition from low to medium levels of effect results in an 8% increase in the system's compensation expenditure. However, when passengers on long-distance trips are more inclined to accept detours, total compensation expenditure decreases by 16%, and the onboard rate progressively increases from 54% to 66%, as shown in Figure 3.10(d). These results suggest that initially, the system increases compensation to attract more profitable opportunities. As passengers on longer trips become more willing to accept detours, the required compensation amount decreases, leading to lower total expenditures and a more efficiently combined operation. In summary, the operator should balance incentive costs against diminishing profit returns and prioritize strategies that encourage detours among passengers on long trips to boost overall system profitability.

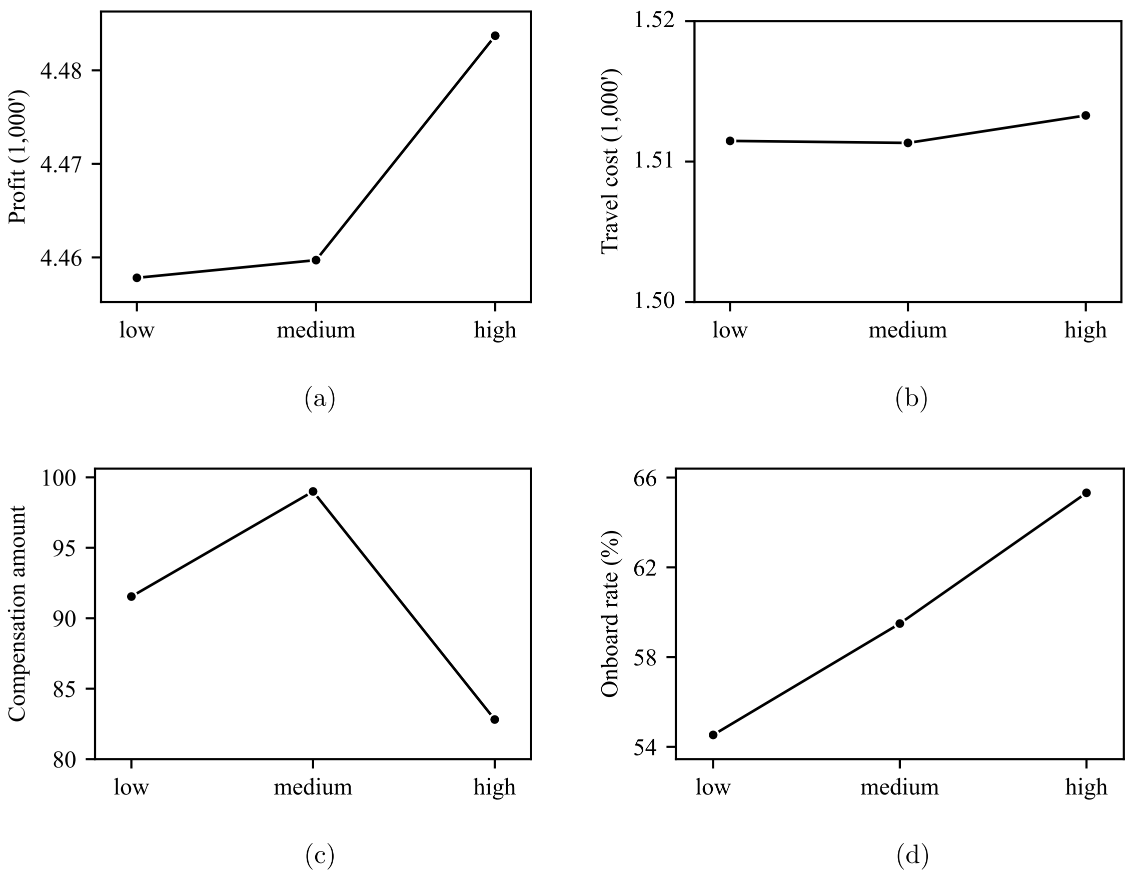
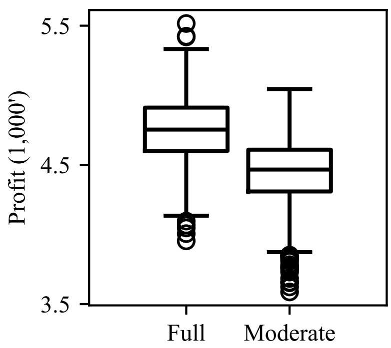


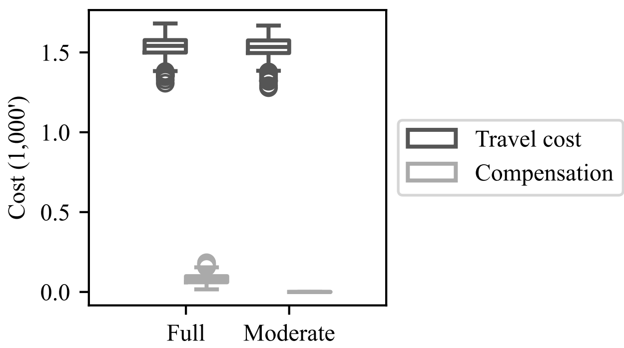
Figure 3.10. Sensitivity analysis of μ regarding the direct travel time on (a) profits, (b) travel costs, (c) compensation amount, and (d) onboard rate

3.4.6 Advantages of full integration

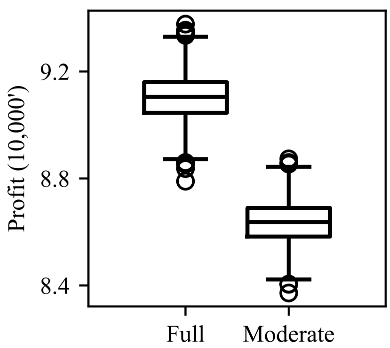
As mentioned earlier, the CUT can be implemented in two modes, either full integration investigated in this paper or moderate integration that does not allow parcel delivery when passengers are on board, which does not require compensation decisions. In this subsection, we evaluate the benefit of the full integration by comparing it to the moderate integration in small-scale (400 requests and 5 vehicles) and large-scale (4,000 requests and 50 vehicles) instances under uniform spatial distribution. We employ ACCR to make confirmation and routing decisions in the moderate integration while excluding the compensation strategy (ignoring lines 10-26 in Algorithm 3.1), which adopts the baseline solution at each decision epoch.



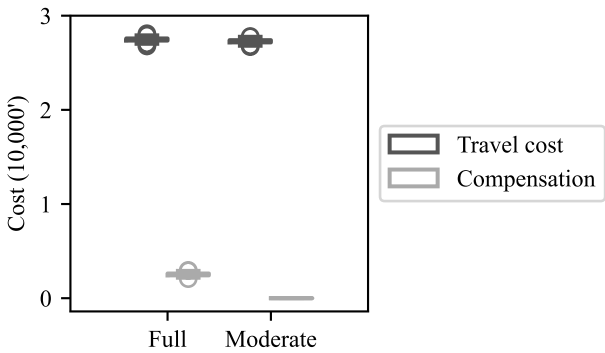
(a) Profits in small-scale instances



(b) Travel costs and compensation in small-scale instances



(c) Profits in large-scale instances



(d) Travel costs and compensation in large-scale instances

Figure 3.11. Comparison of CUT in the full and moderete integration

The comparison results of the total profits, travel cost, and compensation amount in each mode are visualized in Figure 3.11. In small-scale instances, the profits obtained in the full integration are on average 6.7% higher than those obtained in the moderate integration, as shown in Figure 3.11(a). The average travel costs under the full and moderate integration are nearly the same (1,536 vs. 1,531), as shown in Figure 3.11(b). The advantage of full integration over moderate integration is also obvious in the large-scale instance setting, where the full integration could obtain on average 5.4% more profits with similar travel costs compared to the moderate integration, as shown in Figures 3.11(c) and 3.11(d). The results indicate that in the full integration, vehicles might better utilize consolidation opportunities along their routes, with only marginal increases in travel costs. Furthermore, the compensation costs incurred in the full integration were quite modest, amounting to only 5% of the total travel costs. This suggests that the additional costs associated with the compensation strategy in the full integration are relatively minor. The service provider is thus recommended to implement the full-integrated CUT to capture more consolidation opportunities with improved profitability at marginal additional costs.

3.5 Concluding Remarks

In this study, we investigated a dynamic confirmation, compensation, and routing problem for CUT, where compensation was offered as an incentive to motivate passengers to accept detours for parcel delivery. This dynamic and stochastic problem was formulated as an MDP for making sequential decisions to maximize the total profits. We further proposed an anticipatory model to incorporate multiple decisions while considering their impact on future profitability. By exploiting the structure of the anticipatory model, we propose an effective anticipatory policy ACCR to make real-time decisions of confirmation, compensation, and routing. This policy was supported by a novel regression-tree-based VFA method for learning the future profitability given the current fleet status. The numerical results have validated the effectiveness of our policy and the significance of the proposed regression-tree-based VFA on the performance of the policy. We also find that compared to moderate integration, which does not allow parcel delivery during passengers' trips, by motivating passengers to accept parcel delivery, full integration can capture more consolidation opportunities with marginal costs, leading to improved profitability.

Chapter 4 Adaptive Arc Generation for Urban Air Mobility Service with Pooling and Flight Transfer

This chapter focuses on a novel UAM service with pooling and flight transfer, referred to as the UAM-PF problem. The concept of transit transportation has been widely adopted in established urban transportation systems such as ride-sharing, mass transit railways, and logistics (Wang et al., 2023; Owais et al., 2021; Cortés et al., 2010; Azcuy et al., 2021; Chen et al., 2019). This sharing model holds great potential to improve the utilization of fleets and service level (Singh et al., 2021; Teubner and Flath, 2015). In the UAM, the synergy of transfer and pooling transforms a system of many independent, inefficient point-to-point routes into an interconnected network. This approach can consolidate dispersed demand, which significantly reduces empty legs and enhances city-wide connectivity, ultimately enabling a more flexible, efficient urban air transportation system.

In this problem, passengers place trip orders specifying their origin and destination vertiports, and the earliest departure time at the origin and deadlines for arriving at the destination. The operator must arrange appropriate flights for each passenger, implementing passenger pooling and flight transfer strategies for a seamless multiple segment journey. Additionally, the operator must manage aircraft operations by dispatching aircraft to transport passengers according to the planned routes and making charging plans for aircraft at vertiports to ensure continuous service availability. The objective of this research is to minimize the total operational costs while maintaining high service quality, specifically balancing direct operational costs with penalties associated with passenger inconvenience, thereby addressing the fundamental trade-off between cost efficiency for the operator and time efficiency for passengers in on-demand air mobility services. We formulate an integer programming based on a mixed time-space and time-space-battery graph, which can effectively integrate passenger flight planning, aircraft dispatching, and charging decisions in a unified framework. Given the complexity of the problem, solving the original integer programming model directly using commercial solvers is computationally challenging. To this end, we propose two arc generation algorithms that construct sparse graphs with a restricted subset of nodes and arcs to enhance the computational efficiency while maintaining solution quality. We first propose a MAG algorithm to provide the foundational framework to build the sparse graph iteratively. Then, we develop an AAG algorithm by adaptively adjusting the construction of a sparse graph based on feedback from previous iterations, which enables more intelligent exploration of the solution space. Numerical experiments demonstrate the excellent performance of the AAG algorithm compared to benchmark algorithms regarding solution quality and computational efficiency.

The remainder of this chapter is organized as follows. Section 4.1 presents the problem statement and formulation. Section 4.2 elaborates on the solution method for MAG and AAG. Numerical experiments are conducted in Section 4.3. Section 4.4 presents the concluding remarks of this chapter.

4.1 Problem Statement and Formulation

In this section, we present the formulation of the UAM-PF problem. We begin with a general problem statement that outlines the global notations, which are summarized in Table 4.1 for reference. Subsequently, we introduce a discretization approach for time periods and battery levels to make the problem computationally tractable. Finally, we formulate this problem as an integer programming model.

Table 4.1. Problem notations

Notations	Descriptions
	Indices and sets
N_v	Set of vertiports
$\{0, \lvert N_v\rvert{+}1\}$	Set of dummy nodes
N	Set of all nodes, $N = N_v \cup \{0, \lvert N_v\rvert{+}1\}$
A	Set of arcs of nodes i and $j \in N$
G_v	Physical network, $G_v = (N, A)$
(i, j)	Index of arcs
K	Set of trip orders
$\mathscr{T}$	Set of discrete time points

Continued on next page

Table 4.1 – continued from previous page

Notations	Descriptions
$N_{\mathscr{T}}$	Set of time-space nodes
$\tilde{A}_{\mathscr{T}}$	Set of traveling time-space arcs
$H_{\mathscr{T}}$	Set of waiting time-space arcs
$A_{\mathscr{T}}$	Set of time-space arcs, $A_{\mathscr{T}} = \tilde{A}_{\mathscr{T}} \cup H_{\mathscr{T}}$
(i, t, j, t')	Index of time-space arcs
$G_{\mathscr{T}}$	Time-space graph, $G_{\mathscr{T}} = (N_{\mathscr{T}}, A_{\mathscr{T}})$
B	Set of discrete charge points
$N_{\mathscr{T}B}$	Set of time-space-battery nodes
$\tilde{A}_{\mathscr{T}B}$	Set of traveling time-space-battery arcs
$H_{\mathscr{T}B}$	Set of waiting time-space-battery arcs
$A_{\mathscr{T}B}$	Set of time-space-battery arcs, $A_{\mathscr{T}B} = \tilde{A}_{\mathscr{T}B} \cup H_{\mathscr{T}B}$
$(i, t, \beta, j, t', \beta')$	Index of time-space-battery arcs
$G_{\mathscr{T}B}$	Time-space-battery graph,$G_{\mathscr{T}B} = (N_{\mathscr{T}B}, A_{\mathscr{T}B})$
	Parameters
T	Service period
Q	Vertiport capacity
M	Fleet size
d_{ij}	Travel time of arc (i, j)
f_{ij}	Flight cost of arc (i, j)
v_{ij}	Battery consumption of arc (i, j)
ϕ	Charging rate
q	Load capacity of aircraft
k^+	Origin vertiport of order k
k^-	Destination vertiport of order k
e_k	Earliest departure time of order k
l_k	Deadline of arriving at destination of order k
p_k	Number of passengers in order k
Δ	Time interval
Δ_b	Battery level interval
ρ	Time inefficiency penalty
	Decision variables

Continued on next page

Table 4.1 – continued from previous page

Notations	Descriptions
$x^k_{itjt'}$	Binary variable indicating whether order k travels through arc (i, t, j, t')
$z^{\beta\beta'}_{itjt'}$	Integer variable indicating the number of aircraft traveling through arc $(i, t, \beta, j, t', \beta')$
$y^{\beta\beta'}_{itjt'}$	Integer variable indicating the number of aircraft carrying passengers through arc $(i, t, \beta, j, t', \beta')$
$w^{\beta\beta'}_{itjt'}$	Integer variable indicating the number of aircraft without carrying passengers through arc $(i, t, \beta, j, t', \beta')$

4.1.1 Problem statement

Let the graph $G_v = (N, A)$ represent the network of vertiports. The node set is defined as $N = N_v \cup \{0, |N_v|+1\}$, where N_v represents the set of vertiports, and nodes 0 and $|N_v|+1$ are dummy nodes representing starting and ending depots for all aircraft, respectively. Each vertiport has a capacity limit of Q aircraft, which constrains the total number of aircraft simultaneously occupy the vertiport. Each arc $(i, j) \in A$ represents a flight connection between nodes i and $j \in N$. Each flight arc $(i, j) \in A$ is characterized by three parameters: travel time d_{ij}, operating cost f_{ij}, and battery consumption v_{ij}. The travel time d_{ij} encompasses the actual flight duration plus any necessary service times. The fleet consists of M homogeneous aircraft, each with a passenger capacity of q seats and battery state-of-charge (SoC) $\in [0, 100]\%$. Aircraft can charge at any vertiport at a constant rate of ϕ percent per minute, and partial charging is permitted during waiting periods. Let K represent the set of all trip orders to be served within the service period $[0, T]$. Each order $k \in K$ is defined by four attributes: origin vertiport $k^+ \in N$, destination vertiport $k^- \in N$, earliest departure time e_k, and latest arrival time l_k.

To address the operational challenge of coordinating passenger pooling (multiple trips sharing compatible flights) and transfers at intermediate vertiports to improve service efficiency, the UAM-PF problem optimizes passenger itinerary planning (including routing and transfers) and aircraft routing (including flight scheduling and charging), where passengers are served by available aircraft at each flight segment. The objective is to minimize operating costs and time-related penalties while ensuring all trips are completed by the travel requirement.

4.1.2 Discretization of time period and battery level

The continuous nature of time and battery levels presents computational challenges for solving the UAM-PF problem. To develop a tractable integer programming formulation, we employ a discretization approach that transforms the continuous problem into a discrete network flow problem. This discretization enables us to track aircraft movements and battery states through time while maintaining computational efficiency. We first discretize the time horizon to construct time-space graphs for modeling time-dependent trajectories, then extend this framework to include the SoC, creating a three-dimensional time-space-battery graph that captures all relevant aircraft states.

Time discretization

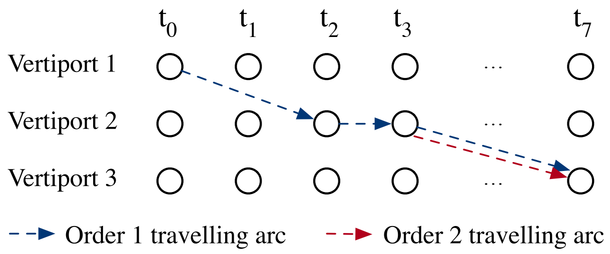
We partition the service period $[0, T]$ into uniform intervals of length Δ, yielding the discrete time set $\mathscr{T} = \{0, \Delta, 2\Delta, \ldots, n\Delta\}$, where $n = T/\Delta$ represents the number of time intervals. For notational convenience, we assume Δ divides T evenly. This discretization transforms the physical vertiport network into a time-space network. Each physical vertiport $i \in N_v$ is replicated at each time point $t \in \mathscr{T}$, creating time-space nodes (i, t). The complete node set becomes $N_{\mathscr{T}} = \{(i, t) | i \in N_v, t \in \mathscr{T}\} \cup \{(0, 0), (|N_v|+1, n\Delta)\}$, $(0, 0)$ and $(|N_v|+1, n\Delta)$ represent the dummy nodes at the beginning and end of the service period.

The time-space arc set consists of two types of arcs: (i) flight arcs $\tilde{A}_{\mathscr{T}} = \{(i, t, j, t') | i \neq j, (i, t), (j, t') \in N_{\mathscr{T}}\}$, where each arc (i, t, j, t') represents departing from the vertiport i at time t and arriving at the vertiport j at time $t' = \left\lceil \frac{t+d_{ij}}{\Delta} \right\rceil \Delta)$, and (ii) waiting arcs $H_{\mathscr{T}} = \{(i, t, i, t') | t' = t + \Delta, t < n\Delta, (i, t) \in N_{\mathscr{T}}\}$, where each arc represents waiting at the vertiport i from time for one time interval. Thus, we extend the physical graph G_v into the time-space graph $G_{\mathscr{T}} = (N_{\mathscr{T}}, A_{\mathscr{T}})$. To ensure feasibility after time discretization, we adjust the timing requirement of trip orders. For each trip order $k \in K$, we map its earliest possible departure time e_k to $\lceil \frac{e_k}{\Delta} \rceil \Delta$, and its deadline l_k to $\lfloor \frac{l_k}{\Delta} \rfloor \Delta$. This approach ensures that all constraints in the original problem remain satisfied in the discretized version.

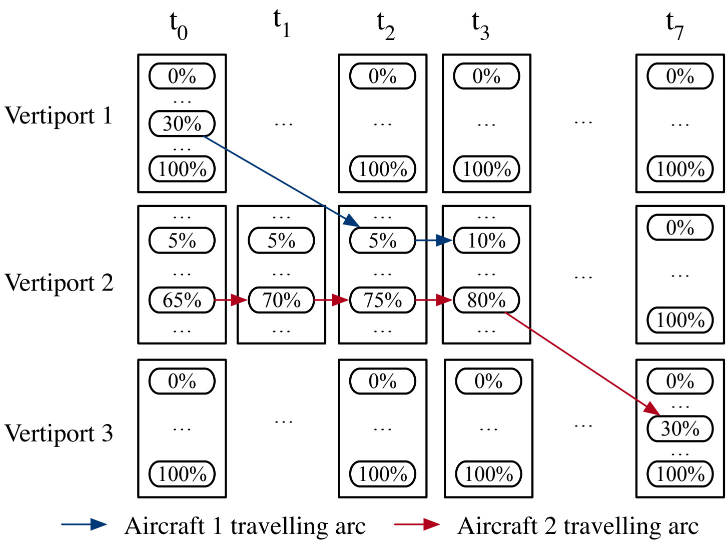
Battery level discretization

To track aircraft' battery level, we extend the time-space graph by discretizing the SoC into $B = \{0, \Delta_b, \ldots, n_b\Delta_b\}$, where the Δ_b represents the SoC granularity and $n_b = \left\lfloor \frac{100\%}{\Delta_b} \right\rfloor$. For notational convenience, we assume Δ_b divides 100% evenly, so that $n_b\Delta_b = 100\%$. This yields $n_b + 1$ discrete battery levels. Each time-space node (i, t) is further expanded into multiple time-space-battery nodes (i, t, β) for each $\beta \in B$. Let $N_{\mathscr{T}B} = \{(i, t, \beta) | (i, t) \in N_{\mathscr{T}}, \beta \in B\}$ represent the set of all time-space-battery nodes. The time-space-battery arc set $A_{\mathscr{T}B}$ contains two types of arcs: (i) flight arcs with battery consumption $\tilde{A}_{\mathscr{T}B} = \{(i, t, \beta, j, t', \beta') | \beta > v_{ij}, \beta \in B, (i, t, j, t') \in \tilde{A}_{\mathscr{T}}\}$, where $\beta' = \left\lfloor \frac{\beta - v_{ij}}{\Delta_b} \right\rfloor \Delta_b$ represents the SoC after consuming v_{ij} during flight, and (ii) waiting arcs with charging $H_{\mathscr{T}B} = \{(i, t, \beta, i, t', \beta') | \beta \in B, (i, t, i, t') \in H_{\mathscr{T}}\}$, where $\beta' = \min\{\left\lfloor \frac{\beta + \phi\Delta}{\Delta_b} \right\rfloor \Delta_b, n_b\Delta_b\}$ represents charging at rate ϕ for duration Δ unless the battery is fully charged. We denote the time-space-battery graph by $G_{\mathscr{T}B} = (N_{\mathscr{T}B}, A_{\mathscr{T}B})$.

Figure 4.1 illustrates the discretization approach through an example involving two aircraft serving two trip orders across a three-vertiport network. Figure 4.1(a) shows the time-space representation where each physical vertiport is replicated across discrete time points (some nodes omitted for visual clarity). Order 1 requests transportation from vertiport 1 to vertiport 3, while order 2 requests service from vertiport 2 to vertiport 3. The operator designs the service plan, where order 1 is routed through vertiport 2 as a transfer point, and order 2 waits until t_3 to pool with order 1 for the shared flight to vertiport 3, arriving at t_7. Figure 4.1(b) illustrates the aircraft operation on the time-space-battery graph, where aircraft 1 is dispatched to pickup order 1 at vertiport 1 with 30% SoC, sufficient only for the first leg. Considering this limitation, the operator pre-positions aircraft 2 at vertiport 2 and schedules it to charge from t_0 to t_3. At the designated transfer time t_3, the operator arranges for passengers from both orders to board aircraft 2, which now has adequate SoC for the flight to vertiport 3.



(a) Time-space graph for trip orders



(b) Time-space-battery graph for aircraft

Figure 4.1. Example of mixed time-space-battery graph

4.1.3 Mathematical formulation

We formulate the UAM-PF problem using a multi-layer graph approach that combines a time-space graph $G_{\mathcal{T}}$ for passenger flows and a time-space-battery graph $G_{\mathcal{T}B}$ for aircraft operations. Let $x^k_{itjt'}$ represent the binary flow decision, indicating whether trip order k travels on arc $(i, t, j, t') \in A_{\mathcal{T}}$. Let $z^{\beta\beta'}_{itjt'}$ denote the number of aircraft traveling on arc $(i, t, \beta, j, t', \beta') \in A_{\mathcal{T}B}$. We minimize the total system cost, comprising operating costs and a penalty for passenger time inefficiency. The penalty parameter ρ weights the extra travel time relative to direct flight time for each order.

Based on the graph $G = (G_{\mathcal{T}}, G_{\mathcal{T}B})$, the integer programming model [M(G)] is formulated as follows:

[M(G)]

$$\min \sum_{(i,t,\beta,j,t,\beta')\in A_{\mathscr{T}B}} f_{ij} z_{itjt'}^{\beta\beta'} + \sum_{k\in K} \frac{\rho}{d_{k^+k^-}}((\sum_{(i,t,k^-,t')\in \tilde{A}_{\mathscr{T}}} t' x_{itk^-t'}^{k}) - e_k - d_{k^+k^-}) \tag{4.1}$$

$$\sum_{(j,t',i,t)\in A_{\mathscr{T}}} x_{jt'it}^{k} - \sum_{(i,t,j,t')\in A_{\mathscr{T}}} x_{itjt'}^{k} = \begin{cases} -1, \ i = k^+, t = \left\lceil \frac{e_k}{\Delta} \right\rceil \Delta, \\ 1, \ i = k^-, t = \left\lfloor \frac{l_k}{\Delta} \right\rfloor \Delta, \\ 0, \ \text{otherwise}, \end{cases} \quad \forall k \in K, (i,t) \in N_{\mathscr{T}} \tag{4.2}$$

$$\sum_{(j,t',\beta',i,t,\beta)\in A_{\mathscr{T}B}} z_{jt'it}^{\beta'\beta} - \sum_{(i,t,\beta,j,t',\beta')\in A_{\mathscr{T}B}} z_{itjt'}^{\beta\beta'} \begin{cases} \geq -M, \ i = 0, t = 0, \\ \leq M, \ i = |N_v|+1, t = T, \\ = 0, \ \text{otherwise}, \end{cases} \quad \forall (i,t,\beta) \in N_{\mathscr{T}B} \tag{4.3}$$

$$\sum_{k\in K} p_k x_{itjt'}^{k} \leq \sum_{(i,t,\beta,j,t',\beta')\in A_{\mathscr{T}B}} q z_{itjt'}^{\beta\beta'}, \ \forall (i,t,j,t') \in \tilde{A}_{\mathscr{T}} \tag{4.4}$$

$$\sum_{(i,t,\beta,j,t',\beta')\in A_{\mathscr{T}B}} z_{itjt'}^{\beta\beta'} \leq Q, \ \forall (i,t) \in N_{\mathscr{T}} \tag{4.5}$$

$$x_{itjt'}^{k} \in \{0,1\}, \ \forall k \in K, (i,t,j,t') \in A_{\mathscr{T}} \tag{4.6}$$

$$z_{itjt'}^{\beta\beta'} \in Z^*, \ \forall (i,t,\beta,j,t',\beta') \in A_{\mathscr{T}B} \tag{4.7}$$

The objective function (4.1) minimizes the sum of the total operating cost and the penalties for passenger time inefficiency. The penalty term measures the relative increase in travel time compared to direct flights. Constraints (4.2) ensure flow balance at each time-space node for each order such that orders depart from their origin at or after their earliest departure time and must reach their destination by the deadline. Constraints (4.3) ensure flow balance at the time-space-battery nodes for aircraft such that at most M aircraft can depart from the depot, the same fleet must return by the end of service, and flow is conserved at all intermediate nodes. Constraints (4.4) link passenger and aircraft flows, ensuring sufficient seat capacity on each flight. Constraints (4.5) restrict the number of aircraft simultaneously present at each vertiport. Constraints (4.6) to (4.7) define the decision variable domains.

4.2 Solution Method

The complexity of [M(G)] largely depends on the scale of the graph G, i.e., the number of nodes and arcs. A graph with all possible nodes and arcs, referred to as the full graph, could generate a prohibitively large number of decision variables, making the resulting model computationally intractable for commercial solvers.

While column generation is a standard approach for large-scale optimization problems, applying it directly to our model presents unique challenges. The core issue lies in the non-separable structure of the objective function, which violates the independence assumption of standard Dantzig-Wolfe decomposition. Unlike problems amenable to standard Dantzig-Wolfe decomposition, our objective function combines two interdependent components: aircraft operating costs (dependent on z variables) with passenger time penalties (dependent on x variables). If each column represents a single aircraft route, its operating costs can be calculated independently. However, the passenger time penalties depend on complete itineraries that typically span multiple flight legs across the fleet. Consequently, the total cost of a solution cannot be calculated as the simple sum of individual column costs, preventing the direct calculation of reduced costs in the pricing subproblem. To address this computational challenge, we propose alternative approaches based on arc generation. Rather than working with a full graph, we construct sparse graphs with a strategic subset of nodes and arcs, which significantly enhances the computational efficiency while generating high-quality solutions.

4.2.1 MAG algorithm

The MAG algorithm is first proposed by Jacquillat et al. (2022) for relay logistics problems, which provides a framework for optimizing routing decisions through iterative graph construction. The procedure of the MAG algorithm is illustrated in Figure 4.2. It begins with an initialized sparse graph $\bar{G}$ and iteratively solves a sparse master problem (SMP) and a subproblem. The SMP is developed based on the current sparse graph to obtain dual multipliers. The subproblem uses the dual multipliers to identify promising arcs that could improve the objective value. The sparse graph $\bar{G}$ is updated by adding the identified promising arcs, and the process repeats until no further beneficial arcs are found. The final solution is obtained by solving the integer program on the finalized sparse graph.

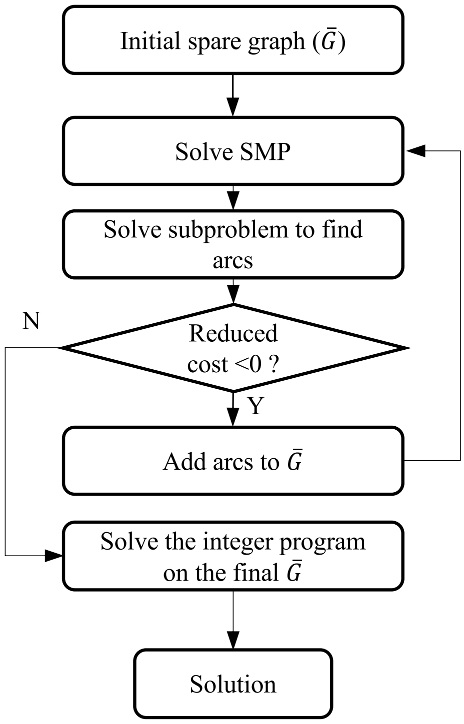


Figure 4.2. Procedure of MAG algorithm

Sparse master problem

The SMP is a linear relaxation of the original problem [M(G)] restricted to a sparse graph $\bar{G}$. While the time-space-battery network for aircraft remains complete to ensure feasibility, we restrict the time-space network by maintaining only a subset of arcs for each trip order.

To enable order-specific arc generation, we decompose the passenger flow network by trip order. For each order k, let $N^k_{\mathcal{T}} \subseteq N_{\mathcal{T}}$ be the set of order-specific nodes and $A^k_{\mathcal{T}} \subseteq A_{\mathcal{T}}$ be the set of order-specific arcs. This yields an order-specific time-space graph $G^k_{\mathcal{T}} = \{N^k_{\mathcal{T}}, A^k_{\mathcal{T}}\}$. The complete graph can thus be expressed as $G = (\{G^k_{\mathcal{T}} | k \in K\}, G_{\mathcal{T}B})$, comprising $|K|$ order-specific time-space graphs and one time-space-battery graph for aircraft operations.

Let $\bar{G} = (\{\bar{G}^k_{\mathscr{T}} | k \in K\}, G_{\mathscr{T}B})$ represent the sparse graph, where $\bar{G}^k_{\mathscr{T}} = (\bar{N}^k_{\mathscr{T}}, \bar{A}^k_{\mathscr{T}})$ with $\bar{A}^k_{\mathscr{T}} \subseteq A^k_{\mathscr{T}}$ representing the currently available arcs for order k. The sparse master problem (SMP) is formulated as follows:

$$\min \sum_{(i,t,\beta,j,t,\beta') \in A_{\mathscr{T}B}} f_{ij} z^{\beta\beta'}_{itjt'} + \sum_{k \in K} \frac{\rho}{d_{k^+k^-}} \Big(\Big(\sum_{(i,t,k^-,t') \in \bar{A}^k_{\mathscr{T}}} t' x^k_{itk^-t'} \Big) - e_k - d_{k^+k^-} \Big) \tag{4.8}$$

$$\text{s.t.} \sum_{(j,t',i,t) \in \bar{A}^{kn}_{\mathscr{T}}} x^k_{jt'it} - \sum_{(i,t,j,t') \in \bar{A}^{kn}_{\mathscr{T}}} x^k_{itjt'} = \begin{cases} -1, & i = k^+, t = \left\lceil \frac{e_k}{\Delta} \right\rceil \Delta, \\ 1, & i = k^-, t = \left\lfloor \frac{l_k}{\Delta} \right\rfloor \Delta, \\ 0, & \text{otherwise,} \end{cases} \quad \forall k \in K, (i,t) \in \bar{N}^k_{\mathscr{T}} \tag{4.9}$$

Eqs. (4.3) to (4.5)

$$x^k_{itjt'} \in [0,1], (i,t,j,t') \in \bar{A}^{kn}_{\mathscr{T}}, \ \forall k \in K \tag{4.10}$$

$$z^{\beta\beta'}_{itjt'} \geq 0, \ \forall (i,t,\beta,j,t',\beta') \in A_{\mathscr{T}B} \tag{4.11}$$

The SMP maintains the same objective structure as [M(G)], minimizing operating costs and time inefficiency penalties, but with two critical modifications. First, the passenger flow variables $x^k_{itjt'}$ and aircraft flow variables $z^{\beta\beta'}_{itjt'}$ are relaxed, transforming the problem into a linear program. Second, passenger flow variables are restricted by the sparse graph $\bar{G}^k_{\mathscr{T}}$ instead of the full graph $G^k_{\mathscr{T}}$. This restriction appears in both the objective function (4.8) and constraints (4.9). The remaining constraints for aircraft flow balance, capacity linking, and vertiport capacity remain unchanged, with the aircraft network remaining complete to ensure feasibility. This formulation reduces computational complexity while preserving the problem structure needed for iterative improvement.

Arc generation subproblem

The arc generation subproblem identifies promising arcs to add to the sparse graph by exploiting dual information from the SMP. Using the optimal dual multipliers, we compute reduced costs of all possible arcs and select those with negative reduced costs, indicating potential for objective improvement.

From the optimal solution of SMP($\bar{G}^n$), we obtain λ_{it}^k as dual multipliers for constraints (4.9), and $\pi_{itjt'}$ as dual multipliers for constraints (4.4). For each arc $(i,t,j,t') \in A_{\mathscr{T}}^k$ and order $k \in K$, the reduced cost is calculated by

$$c_{itjt'}^k = \begin{cases} \lambda_{it}^k - \lambda_{k^-t'}^k + \pi_{itk^-t'} + \frac{\rho t'}{d_{k^+k^-}}, \ (i,t,k^-,t') \in \tilde{A}_{\mathscr{T}}, \\ \lambda_{it}^k - \lambda_{jt'}^k + \pi_{itjt'}, \ (i,t,j,t') \in \tilde{A}_{\mathscr{T}}, \\ \lambda_{it}^k - \lambda_{jt'}^k, \ (i,t,j,t') \in H_{\mathscr{T}}, \end{cases} \quad \forall (i,t,j,t') \in A_{\mathscr{T}}^k, k \in K. \tag{4.12}$$

Let $\bar{x}_{itjt'}^k$ denote the solution value for each arc $(i,t,j,t') \in \bar{A}_{\mathscr{T}}^k$. The arc generation subproblem (SP) is then formulated as follows:

$$\min \sum_{k \in K} \sum_{(i,t,j,t') \in A_{\mathscr{T}}^k} c_{itjt'}^k \delta_{itjt'}^k \tag{4.13}$$

$$\sum_{(j,t',i,t) \in A_{\mathscr{T}}} (\bar{x}_{jt'it}^k + \delta_{jt'it}^k) - \sum_{(i,t,j,t') \in A_{\mathscr{T}}} (\bar{x}_{itjt'}^k + \delta_{itjt'}^k) = \begin{cases} -1, \ i = k^+, t = \left\lceil \frac{e_k}{\Delta} \right\rceil \Delta \\ 1, \ i = k^-, t = \left\lfloor \frac{l_k}{\Delta} \right\rfloor \Delta \\ 0, \ \text{otherwise} \end{cases}, \ \forall k \in K, (i,t) \in N_{\mathscr{T}}$$

$$\bar{x}_{itjt'}^k + \delta_{itjt'}^k \geq 0, \ \forall k \in K, (i,t,j,t') \in A_{\mathscr{T}} \tag{4.14}$$

The arc generation subproblem aims to find a vector $\boldsymbol{\delta}$ that indicates a direction in which costs can be reduced while maintaining flow balance feasibility. Since order flows for different orders are independent, the SP can be decomposed into $|K|$ independent shortest path problems. Specifically, for each order $k \in K$, the problem is to find a path on the full graph $G_{\mathscr{T}}^k$ from the origin $(k^+, \left\lceil \frac{e_k}{\Delta} \right\rceil \Delta)$ to the destination $(k^-, \left\lfloor \frac{l_k}{\Delta} \right\rfloor \Delta)$, such that the sum of the reduced costs of constituent arcs is minimized. If the shortest path has a negative total cost, add all its constituent nodes and arcs to update the sparse graph from $\bar{G}^n$ to $\bar{G}^{n+1}$.

The procedure stops when no negative reduced cost paths exist. This indicates that no additional arcs can improve the SMP objective. Since the time-space graphs are directed and acyclic (time only moves forward), each shortest path problem can be solved efficiently using dynamic programming algorithms (Righini and Salani, 2006).

4.2.2 AAG algorithm

The MAG algorithm provides a systematic approach to sparse graph construction. However it treats flight costs as fixed values regardless of passenger loads, potentially missing opportunities to identify cost-efficient flight arrangements. The AAG algorithm addresses this limitation by introducing adaptive individual flight costs that represent the cost per passenger on each time-space arc. These costs are dynamically updated based on actual passenger flows at each iteration, enabling the arc generation process to better capture the economic benefits of consolidation.

Enhanced sparse master problem

Instead of using fixed flight costs f_{ij}, we introduce adaptive per-passenger costs $\bar{f}^n_{itjt'}$ that reflect actual utilization patterns at iteration n, which is calculated as follows:

$$\bar{f}^{n+1}_{itjt'} = \begin{cases} \frac{f_{ij} m_{itjt'}}{\sum_{k \in K} p_k \bar{x}^k_{itjt'}}, & m_{itjt'} > 0 \\ \frac{1}{n} \sum_{n'=0}^{n} \bar{f}^{n'}_{itjt'}, & m_{itjt'} = 0 \end{cases} \tag{4.15}$$

where $m_{itjt'} = \lceil \sum_{k \in K} \bar{x}^k_{itjt'} / q \rceil$ is the number of aircraft required. For active arcs ($m_{itjt'} > 0$), this formula distributes the total flight cost evenly among all passengers. This ensures that when multiple passengers share flights, each pays a proportionally lower cost, naturally capturing pooling benefits. For inactive arcs ($m_{itjt'} = 0$), the formula calculates the average of all historical cost estimates for that arc. This prevents information loss and provides reasonable cost estimates for arcs that may become active in future iterations, ensuring continuity in the cost structure. The initial value $\bar{f}^0_{itjt'}$ for each arc (i, t, j, t') is set to f_{ij}.

Let $y_{it\beta jt'\beta'}$ and $w_{it\beta jt'\beta'}$ denote the number of aircraft carrying passengers and not carrying passengers through the arc $(i, t, \beta, j, t', \beta')$, respectively. By replacing the fixed cost term with adaptive costs, we formulate the enhanced SMP as follows:

[E-SMP]

$$
\begin{aligned}
\min \sum_{k\in K} \sum_{(i,t,j,t')\in \bar{A}^k_{\mathscr{T}}} \bar{f}^n_{itjt'} x^k_{itjt'} + \sum_{(i,t,\beta,j,t,\beta')\in A_{\mathscr{T}B}} f_{ij} w^{\beta\beta'}_{itjt'} \\
+ \sum_{k\in K} \frac{\rho}{d_{k^+k^-}}\Big(\Big(\sum_{(i,t,k^-,t')\in \bar{A}^k} t' x^k_{itk^-t'}\Big) - e_k - d_{k^+k^-}\Big)
\end{aligned} \tag{4.16}
$$

s.t. Eqs. (4.2) to (4.3) and (4.5)

$$\sum_{k\in K} x^k_{itjt'} \leq \sum_{(i,t,\beta,j,t',\beta')\in A_{\mathscr{T}B}} q y^{\beta\beta'}_{itjt'},\ \forall (i,t,j,t') \in \bar{A}^{kn}_{\mathscr{T}} \tag{4.17}$$

$$z^{\beta\beta'}_{itjt'} = y^{\beta\beta'}_{itjt'} + w^{\beta\beta'}_{itjt'},\ \forall (i,t,\beta,j,t',\beta') \in A_{\mathscr{T}B} \tag{4.18}$$

$$x^k_{itjt'} \in [0,1], (i,t,j,t') \in \bar{A}^{kn}_{\mathscr{T}},\ \forall k \in K \tag{4.19}$$

$$y^{\beta\beta'}_{itjt'} \geq 0,\ \forall (i,t,\beta,j,t',\beta') \in A_{\mathscr{T}B} \tag{4.20}$$

$$w^{\beta\beta'}_{itjt'} \geq 0,\ \forall (i,t,\beta,j,t',\beta') \in A_{\mathscr{T}B} \tag{4.21}$$

Unlike the original objective in Eq. (4.8), Eq. (4.16) employs adaptive per-passenger costs instead of fixed flight costs. This key distinction allows the model to capture actual utilization patterns at each iteration by calculating flight costs based on the product of individual passenger costs and passenger flow variables. Constraints (4.17) ensure adequate aircraft dispatch for each flight. Constraints (4.18) decompose total aircraft flow into passenger-carrying and empty repositioning components. Constraints (4.19) restrict order flow variables to the sparse graph structure. Constraints (4.20) to (4.21) define the region of aircraft flow variables.

Subproblem

While the subproblem structure in AAG follows the same decomposition principle as MAG, the key distinction lies in the computation of reduced costs. The adaptive cost structure $\bar{f}^n_{itjt'}$ provides a crucial advantage for the subproblem: it enables the passenger flow identification process to account for varying costs across different time-space arcs, encouraging the discovery of arcs with high pooling potential. This mechanism leads to more effective sparse graph construction, which ultimately enhances computational efficiency in solving the overall problem.

Based on the dual multipliers and the updated individual cost, the reduced cost for each arc $(i, t, j, t') \in A^k_{\mathscr{T}}$ and order $k \in K$ is calculated as follows:

$$c^k_{itjt'} = \begin{cases} \bar{f}^{n+1}_{itk^-t'} p_k + \lambda^k_{it} - \lambda^k_{k^-t'} + \pi_{itk^-t'} p_k + \frac{\rho t'}{d_{k^+k^-}}, \ (i, t, k^-, t') \in \tilde{A}_{\mathscr{T}}, \\ \bar{f}^{n+1}_{itjt'} p_k + \lambda^k_{it} - \lambda^k_{jt'} + \pi_{itjt'} p_k, \ (i, t, j, t') \in \tilde{A}_{\mathscr{T}}, \\ \lambda^k_{it} - \lambda^k_{jt'}, \ (i, t, j, t') \in H_{\mathscr{T}}, \end{cases} \quad \forall (i, t, j, t') \in A^k_{\mathscr{T}}, \ k \in K. \tag{4.22}$$

Given these modified reduced costs, the AAG subproblem follows the same solution logic as MAG.

4.2.3 Solution procedure

Both MAG and AAG follow a three-phase solution framework: (i) an initialization phase that constructs an initial sparse graph, (ii) an iterative arc generation process that expands the sparse graph, and (iii) a final optimization phase that solves for an optimal integer solution on the finalized sparse graph. While the overall structure is identical for both algorithms, AAG incorporates adaptive cost updates within the iterative phase.

The sparse graph $\bar{G}^0$ is initialized using a heuristic algorithm. Let γ-shortest path for order k represent the travel time of the path that does not exceed the direct flight time $d_{k^+k^-}$ multiplying γ. For each order k, we identify multiple γ-shortest paths from its origin to destination on the full time-space graph $G^k_{\mathscr{T}}$. The arcs from γ-shortest paths are used to initialize sparse time-space graphs in $\bar{G}^0$.

The solution procedure is outlined in Algorithm 4.3. The algorithm begins by initializing the sparse graph $\bar{G}^0$ and the adaptive individual flight cost (see Lines 1-7). At each iteration, to obtain the dual variables, if the algorithm is AAG, the E-SMP is solved; otherwise, the SMP is solved (see Lines 10–14). After obtaining the dual variables, the reduced costs are updated for all arcs in the time-space

Algorithm 4.3 Solution procedure

1: $n \leftarrow 0$; Initialize $\bar{G}^n$
2: **if** algorithm is AAG **then**
3:     **for** $(i,t,j,t') \in \tilde{A}_{\mathscr{T}}$ **do**
4:         $\bar{f}^n_{itjt'} \leftarrow f_{ij}$
5:     **end for**
6:     **for** $(i,t,j,t') \in H_{\mathscr{T}}$ **do**
7:         $\bar{f}^n_{itjt'} \leftarrow 0$
8:     **end for**
9: **end if**
10: **if** algorithm is AAG **then**
11:     $\boldsymbol{\lambda}, \boldsymbol{\pi} \leftarrow$ Solve E-SMP;
12: **else**
13:     $\boldsymbol{\lambda}, \boldsymbol{\pi} \leftarrow$ Solve SMP;
14: **end if**
15: **for** $(i,t,j,t') \in A_{\mathscr{T}}$ **do**
16:     **if** algorithm is AAG **then**
17:         Update $\bar{f}^{n+1}_{itjt'}, c_{itjt'}$
18:     **else**
19:         Update $c_{itjt'}$
20:     **end if**
21: **end for**
22: $I_{\mathscr{T}} \leftarrow$ Solve SP;
23: **if** $I_{\mathscr{T}} = \varnothing$ **then**
24:     Terminate
25: **else**
26:     Construct $\bar{G}^{n+1}$ by augmenting $\bar{G}^n$ with $I_{\mathscr{T}}$
27:     $n \leftarrow n+1$
28:     Go to Line 10
29: **end if**
30: Solve $M(\bar{G}^n)$

network. Additionally, if the algorithm is AAG, the adaptive individual flight costs are also updated (see Lines 15-21). Then, the subproblem is solved to obtain the set of potential arcs $I_{\mathscr{T}}$ (see Lines 22). If $I_{\mathscr{T}}$ is empty, indicating no more beneficial arcs can be found, the process is terminated; otherwise, the $\bar{G}^{n+1}$ is constructed by augmenting $\bar{G}^{n}$ with $I_{\mathscr{T}}$, and the algorithm continues with the next iteration (see Lines 23-29). Finally, an optimal integer solution could be obtained by solving [M($\bar{G}^{n}$)] with the finalized sparse graph (see Line 30).

4.3 Numerical Experiment

We use Gurobi 8.1 as the linear programming solver for the procedure of the AAG algorithm, and the MIP solver to obtain an optimal integer solution based on the finalized sparse graph. In this section, we first describe the instance setting. Then, we evaluate the performance of the AAG algorithm. Next, we conduct a sensitivity analysis to assess the impact of an increased number of orders and vertiport capacity on the operational performance. Finally, we will explore the impact of the time inefficiency penalty on the trade-off between operational costs and the delay time of passengers.

4.3.1 Instance and parameter setting

We consider a 75 km $\times$ 75 km square region. The operation duration is set to 480 min. We consider three different vertiport numbers: 7, 8, and 9. The vertiports are randomly distributed within the region, with a minimum distance of 10 km between any two vertiports. Each vertiport can accommodate up to 8 aircraft. The aircraft are uniform, with a maximum flight distance of 144 km, a flight speed of 120 km/h, a charging speed of 2% per min, and a load capacity of 5 passengers (Zhou and Tan, 2024). We consider five different settings for the number of trip orders ranging from 300 to 500. For each order, its origin and destination vertiports are assigned randomly. The earliest departure times of trip orders are uniformly distributed within the time window $[0, 360]$. The deadline l_k for request k is set to be $e_k + d_{k^+k^-} + 40$ min. Each trip order has 1 to 3 passengers. The penalty for time inefficiency ρ is set to be 10. Given the combination of vertiport number and order number, 15 instance settings are randomly generated. Regarding the algorithm-related parameters, we set the time interval Δ to 5 min, the interval of battery level Δ_b to 2%, and γ to 0.02.

4.3.2 Performance evaluation

To evaluate the performance of the AAG algorithm, we will compare its results with those obtained by using Gurobi for solving the $[M(G)]$ problem, the MAG algorithm, and a heuristic algorithm. The heuristic algorithm, referred to as the TP algorithm, decomposes the decisions for order flow and aircraft flow into two phases. In the first phase, the order flow is determined independently of the aircraft flow. The sequence of trip orders is shuffled, and the shortest path for each order in the shuffled sequence is obtained on the full graph G. Additionally, the impact of order flow on the aircraft flow is considered. During the process, if the number of required aircraft at any time-space node equals the vertiport capacity Q, the corresponding time-space traveling arcs are removed from G in the following subsequent determination of the shortest path for orders. In the second phase, the aircraft flow is determined by solving $[M(G)]$ with the order flow fixed from the first phase. A limit of 2 hr is imposed for solving each of these instances. In other words, for a specific instance, all the algorithms will stop if either an optimal solution is obtained within 2 hr or the elapsed time exceeds 2 hr.

The comparison results are presented in Table 4.2, where instance settings are represented by order number #K and vertiport number #N. The objective value (Obj) and computation time (Time) of each method are reported. The best objective values are highlighted in bold. For ease of comparison, the relative gaps, Gap1, Gap2, and Gap3, are computed by $\frac{obj^{*}-obj1}{obj1} * 100\%$, $\frac{obj^{*}-obj2}{obj2} * 100\%$, and $\frac{obj^{*}-obj3}{obj3} * 100\%$ respectively, where $obj^{*}, obj1$, $obj2$, and $obj3$ represent the objective values obtained by AAG, Gurobi, MAG, and TP algorithms. The results demonstrate that the AAG algorithm outperforms the others, achieving the best objective values in all instances. Compared to Gurobi, the AAG algorithm achieves an average of 10.2% less objective value and is around 9 times faster. When the number of orders exceeds 300 for instances with 7 vertiports, and 400 for instances with 8 and 9 vertiports, the AAG algorithm can rapidly find good-quality solutions, whereas the Gurobi can not obtain a feasible solution within the time limit. The MAG and TP algorithms also face this drawback as the number of orders increases. Compared to the MAG algorithm, the AAG algorithm achieves an average of 26.9% less objective value, while compared to the TP algorithm, it achieves an average of 48.8% less objective value. The results demonstrate that the AAG algorithm effectively addresses the challenge of maintaining computational efficiency as the order number increases, whereas the benchmark algorithms fail to do so.

To better understand the mechanism of the AAG algorithm, Figure 4.3 illustrates the number of order flow variables in the original full graph and the finalized sparse graphs obtained by the AAG and MAG algorithms. It indicates that the AAG algorithm can obtain a sparse graph with a reduced number of variables, representing only 7.1% of the full graph, and 36.29% of the sparse graph obtained by the MAG algorithm. The effectiveness of the AAG algorithm may be attributed to its ability of achieving a higher level of sparsity, which reduce the search space effectively, while maintaining the quality of the solution. This balance between sparsity and solution quality could be a key factor in its excellent performance. In contrast, while the MAG algorithm also reduces the number of variables, it may not be as effective in preserving the most relevant arcs, failing to maintain the computational efficiency as the order number increases.

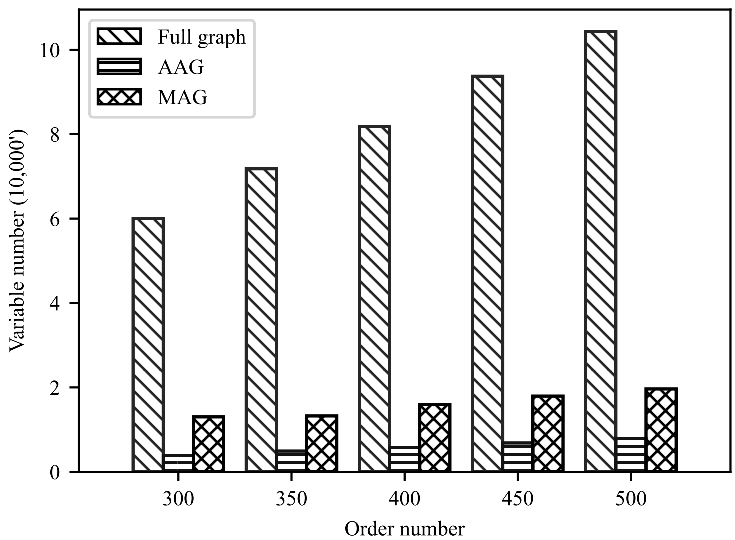


Figure 4.3. Comparison of order flow variables

Table 4.2. Comparison results

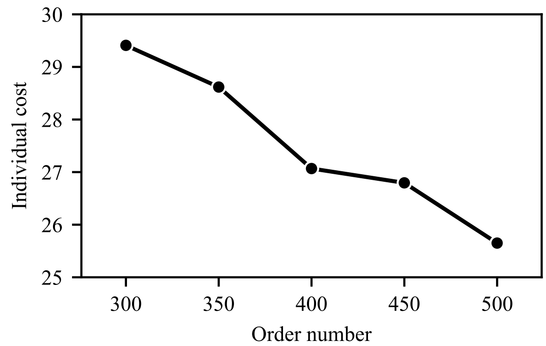
		Gurobi		AAG		MAG		TP		Gap1	Gap2	Gap3
#K	#N	Obj	Time(s)	Obj	Time(s)	Obj	Time(s)	Obj	Time(s)			
300	7	9,029	7,200	**8,705**	300	12,650	5,042	-	7,200	-3.6%	-31.2%	-
350	7	-	7,200	**9,794**	503	-	7,200	-	7,200	-	-	-
400	7	-	7,200	**10,557**	717	-	7,200	-	7,200	-	-	-
450	7	-	7,200	**11,680**	1,005	-	7,200	-	7,200	-	-	-
500	7	-	7,200	**12,997**	1,246	-	7,200	-	7,200	-	-	-
300	8	11,477	7,200	**9,404**	341	11,813	3,776	18,285	4,844	-18.1%	-20.4%	-48.6%
350	8	11,620	7,200	**10,762**	562	15,453	3,814	-	7,200	-7.4%	-30.4%	-
400	8	13,656	7,200	**12,172**	903	16,521	6,587	-	7,200	-10.9%	-	-
450	8	-	7,200	**13,024**	1,101	-	7,200	-	7,200	-	-	-
500	8	-	7,200	**13,947**	1,643	-	7,200	-	7,200	-	-	-
300	9	10,013	7,200	**8,981**	421	11,833	2,301	17,365	980	-10.3%	-24.1%	-48.3%
350	9	11,495	7,200	**10,081**	613	14,223	3,704	19,984	1,076	-12.3%	-29.1%	-49.6%
400	9	12,116	7,200	**11,037**	700	15,118	3,950	-	7,200	-8.9%	-	-
450	9	-	7,200	**12,029**	1,046	-	7,200	-	7,200	-	-	-
500	9	-	7,200	**13,910**	1,120	-	7,200	-	7,200	-	-	-
AVG.		11,343.7	7,200.0	**11,272.1**	814.6	13,943.0	5,784.9	18,544.3	6,220.0	-10.2%	-26.9%	-48.8%

4.3.3 Sensitivity analysis

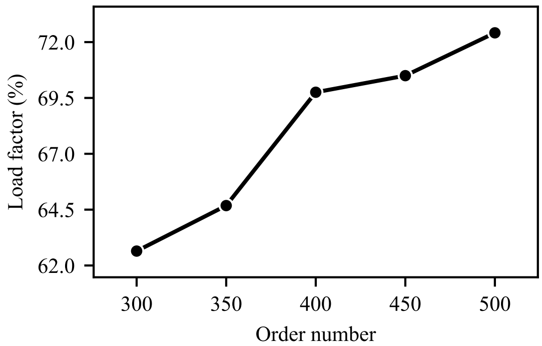
In this subsection, we conduct a sensitivity analysis to evaluate the impact of external factors and operator-controllable parameters on operational performance. Specifically, regarding the external factors, we examine the influence of the number of orders and vertiport capacity on operational performance. Subsequently, we investigate the role of the time inefficiency penalty ρ in balancing operational costs against passenger delay time.

Analysis of the impact of increased number of orders

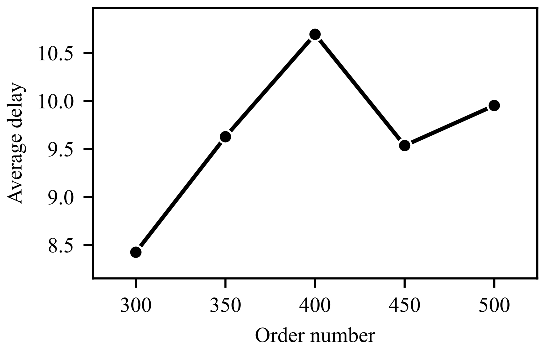
We evaluate the impact of an increased number of orders on three metrics: individual cost, which is the total operational cost divided by the number of trip orders; load factor, which is the average percentage of occupied seats on an aircraft in relation to the total available seats; and average delay time, which is the average difference between the actual arrival time at the destination and the earliest possible arrival time. Figure 4.4(a) shows that individual cost decreases by 12.8% as the order number increases from 300 to 500, which could be attributed to the increased load factor from around 62.0% to 72.0%, as shown in Figure 4.4(b). It indicates that as the market continues to expand, the operational cost can be spread over a larger number of trips, leading to lower per-order costs. Figure 4.4(c) illustrates that the delay time increases from around 8.5 min to 10.5 min as the order number increases from 300 to 400, then decreases to around 9.5 min with 450 orders, and increases again to 10 min with 500 orders. The variation in delay times may be attributed to the system's shifting priority between cost reduction and time efficiency as the number of orders increases. When the system is under moderate load (e.g., 300 to 400 orders), the increased delay time may be due to the focus on increasing aircraft utilization to reduce costs, which in turn increases delay time. However, as more orders are added, the system might adapt by optimizing the flight plan, thereby reducing delays temporarily. The subsequent increase in delay time at 500 orders might indicate another round of shifting priority. These results demonstrate that the system can effectively balance operational costs and time efficiency in an expanding market, such that as the number of orders increases, the individual cost decreases, while average delay time does not increase linearly.



(a) Individual cost with increased order number



(b) Load factor with increased order number

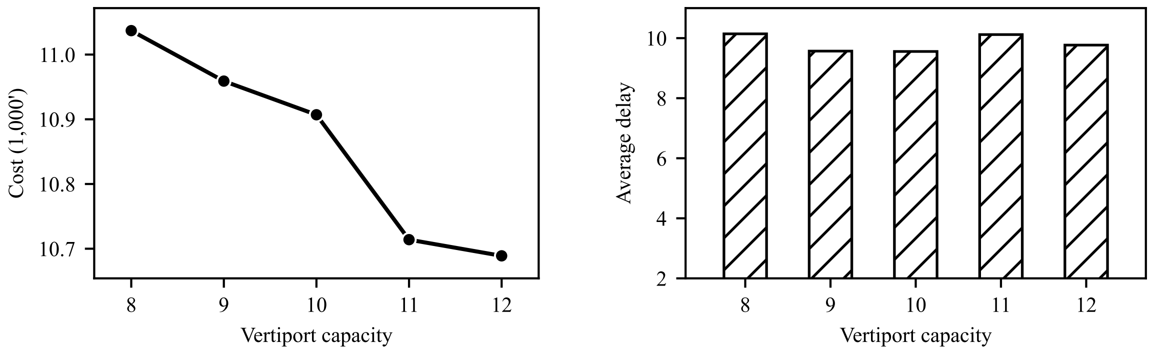


(c) Average delay with increased order number

Figure 4.4. Impact of increased number of orders

Analysis of the impact of increased vertiport capacity

We evaluate the impact of increased vertiport capacity on total operational costs and average delay time, which results are obtained on instances with 400 orders, and 8 vertiports with capacity ranging from 8 to 12. Figure 4.5(a) shows a 3% decrease in operational costs as vertiport capacity increases from 8 to 12. The primary reason for this cost reduction may be that with increased vertiport capacity, more aircraft can park at the vertiport to charge or wait for future orders rather than relocating to another vertiport due to limited parking capacity. With more available parking, the system might optimize aircraft assignment, minimizing unnecessary flights, and thereby reducing costs. However, this cost-saving benefit diminishes as the vertiport capacity increases from 11 to 12. The results suggest that during the planning stage, the service operator should balance the construction cost of increasing vertiport capacity against its benefit in the reduction of operational costs. However, simply increasing vertiport capacity does not necessarily lead to greater time efficiency. As shown in Figure 4.5(b), the average delay time fluctuates between 9 and 10 min. Therefore, in the following, we explore the parameter ρ, which can directly affect the system's focus on the time efficiency.



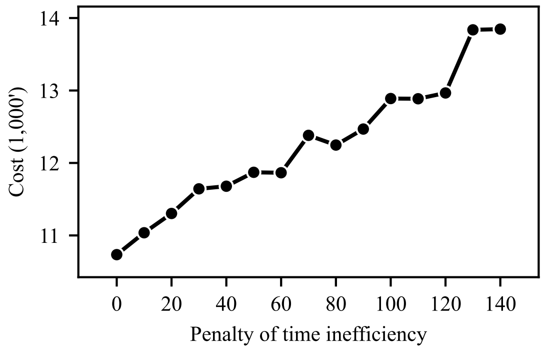
(a) Operational cost with increased vertiport capacity

(b) Average delay time with increased vertiport capacity

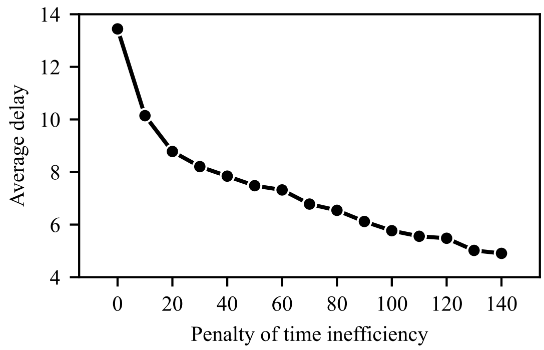
Figure 4.5. Impact of increased vertiport capacity

Analysis of the impact of penalty of time inefficiency

To evaluate the impact of the time inefficiency penalty ρ on the trade-off between the operational cost and time efficiency, we analyze the results by setting ρ ranging from 0 to 140 in instances with 400 orders and 8 vertiports. Figure 4.6(a) shows that as the ρ increases from 0 to 140, the operational cost rises by around 29.0%, while the average delay time decreases by 64.6%. Additionally, the initial increase in ρ has a more significant impact on reducing delay time. As ρ increases from 0 to 10, the average delay time decreases by around 20.5%, while operational costs increase by just 0.8%. However, this effect diminishes gradually as ρ continues to increase. Notably, when ρ increases from 130 to 140, the average delay time decreases by only 2.2%. The primary reason for this diminishing effect may be that the system initially has enough flexibility to make adjustments to the flight plan to reduce delays without significant cost increases. However, as ρ grows, the system may reach a point where further improvements in time efficiency require disproportionately higher costs. In summary, the time inefficiency penalty ρ enables service operators to balance the operational cost and the time efficiency within the system, with small values of ρ being particularly effective in reducing delays at marginal additional costs. Operators should actively recalibrate their trade-off strategies to prevent cost surges due to excessive focus on reducing delay time.



(a) Operational cost with increased penalty of time inefficiency



(b) Average delay time with increased penalty of time inefficiency

Figure 4.6. Impact of penalty of time inefficiency

4.4 Concluding Remarks

In this research, we investigate a novel UAM-PF problem. We formulate this problem as an integer programming model based on a mixed time-space and time-space-battery network. We propose an AAG algorithm to solve this problem efficiently. The numerical results demonstrate the effectiveness of proposed algorithm, which significantly reduces the number of decision variables, and thus increases the computational efficiency. Further sensitivity analysis demonstrates that the system can effectively enhance the cost efficiency with stable time efficiency in an expanding market. Besides, the analysis of the impact of vertiport capacity suggests that during the planning stage, the service operator should balance the construction cost of increasing vertiport capacity against its benefit on the reduction of operational costs. Moreover, we explore the role of time inefficiency penalty in the operational performance, suggesting that operators should actively adjust their trade-off strategies to prevent cost surges due to excessive focus on reducing delay time.

Chapter 5 Dynamic Pricing and Dispatching for Choice-aware Air Taxi Service

This chapter focuses on a dynamic pricing and dispatching problem for choice-aware air taxi services, where operators must simultaneously optimize multimodal trip options, pricing strategies, and real-time fleet dispatching in a dynamic and stochastic environment. Passengers' requests arrive dynamically. For each newly arriving passenger, the operator offers a portfolio of trip options. Each option represents a comprehensive door-to-door travel itinerary combining multiple transportation modes with specific timing information. Passengers' preference towards each option is uncertain and unknown to the operator. The operator faces a complex sequential decision problem: determining which portfolio of trip options to offer, establishing dynamic prices for each option, and executing real-time dispatching decisions once passengers reveal their choices. The objective of the operator is to maximize total profit, calculated as passenger revenue minus operational costs. We formulate this problem as an MDP and develop an anticipatory optimization framework that jointly determines option offerings, pricing, and dispatching decisions while explicitly modeling how these decisions influence passenger choice behavior and future profitability. By exploiting the problem structure, we propose an anticipatory pricing and dispatching policy (APDP) that makes real-time decisions for each incoming request by balancing immediate revenue with future profitability. This policy leverages a DNN-based VFA, which efficiently learns to estimate future profitability given the fleet status. Our computational experiments demonstrate that our proposed policy significantly outperforms all benchmark polices and yields important managerial insights for air taxi service operations.

This chapter is organized as follows. Section 5.1 presents the problem description and formulation. Section 5.2 outlines the proposed anticipatory model and policy. Section 5.3 introduces the DNN-based VFA. In Section 5.4, we present numerical experiments. Section 5.5 presents the concluding remarks of this chapter.

5.1 Problem Description and Formulation

In this section, we first present problem description with an example illustration. Then, we formulate this stochastic and dynamic decision problem as an MDP for sequential decision making.

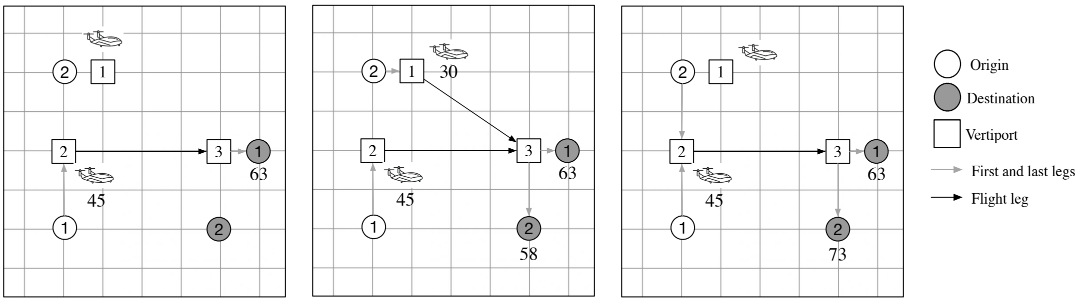
5.1.1 Problem description

We consider an air taxi service operating a fleet of eVTOL aircraft over a defined time horizon T. The service relies on a network of vertiports located throughout the urban area, which serve as dedicated infrastructure for takeoff, landing, and battery charging operations. Each aircraft has limited battery capacity, requiring periodic recharging at the vertiport at a constant rate, and the operator can flexibly decide to partially charge aircraft based on the flight requirements.

During the operational period, passengers submit real-time trip requests for the air taxi service, specifying their origin, destination, and arrival deadline. The operator designs a set of multimodal trip options, where each option consists of the ground transportation from origin to departure vertiport (first leg), air travel between vertiports (flight leg), and ground transportation from arrival vertiport to destination (last leg). These options vary in the multimodal travel route and timing. The operator also determines the price for each option. The passenger either selects a trip option or reject all options based on their individual preferences, which is unknown to the operator. After the passenger reveals his/her choice, the operator dispatches the aircraft and coordinates the complete multimodal journey. The first and last legs are outsourced to ride-hailing platforms that manage their own operations, incurring costs based on ground travel distance and duration. To enhance efficiency, passengers with compatible itineraries, specifically those sharing similar departure and arrival vertiports and time windows, may share flight legs within aircraft seating capacity. As trip requests arrive dynamically, the operator determines three key decisions: the set of multimodal trip options to offer, the price for each option, and the aircraft dispatching including charging schedules. Moreover, these decisions must ensure synchronization between ground transportation and flight departures and arrivals to provide a seamless journey. The objective of this problem is to maximize the expected profits, calculated as total revenue from selected options minus aircraft operating costs and outsourcing fees for ground transportation.

In what follows, we provide an example to illustrate trip options offering at a specific decision epoch of the problem. We consider a Manhattan-style grid, where the ground vehicle travels along street segment with a travel time of 10 min per segment, while the aircraft can travel directly between nodes with an aerial travel time of 2 min per segment. Figure 5.1(a) depicts a scenario with three vertiports at time $t = 20$ min, and the system already has a confirmed passenger, referred to as passenger 1, whose origin and destination are represented by circles with origin in light color and destination in dark color. Based on this passenger's selection before, the system has scheduled a flight departing from vertiport 2 at 45 min, with the passenger expected to reach their final destination at 63 min via ground transportation after the flight. This flight will be operated by an aircraft currently stationed at vertiport 2. Additionally, another aircraft is idle at vertiport 1. All aircraft have sufficient battery charge and available seating capacity to accommodate the new passenger.

At this decision epoch, a new trip request of passenger 2 has just arrived. The passenger's deadline for reaching the destination is 80 min. For the newly arrived trip request, the service operator offers two distinct trip options. Figure 5.1(b) displays the first option with a multimodal journey utilizing a flight departuring at 30 min from vertiport 1 to vertiport 3. Under this option, the arrival time at the destination is 58 minutes, representing the fastest possible route for the passenger. The operator prices this premium option at 300 revenue units. Figure 5.1(c) displays the second option with a multimodal journey utilizing the current flight departing at 45 min from vertiport 2 to vertiport 3. Under this option, the passenger would share the flight with the previously confirmed passenger, arriving at the destination at 73 min. While this option leads to longer travel time to the passenger, it incurs a lower marginal service cost for the operator by utilizing an already scheduled flight. This economy option may be offered at a reduced price of 200 revenue units. Suppose that the passenger selects the second option based on individual utility considering the travel time and costs. Based on this selection, the service operator arranges the



(a) Current state (t=20 min) (b) Trip option 1 (c) Trip option 2

Figure 5.1. Example of a decision epoch

corresponding multimodal journey, with the aerial portion fulfilled by a previously scheduled flight. The aircraft follows this dispatching plan until the next request arrives.

5.1.2 MDP formulation

The service network consists of a set of vertiports N_1. Let N_2 represent the set of origins and destinations of trip requests. Let $N = N_1 \cup N_2$ represent the set of all nodes. Let $A = \{(n, n')|n, n' \in N\}$ be the set of arcs between nodes, including the set of arcs for the flight leg, denoted as $A_1 = \{(n, n')|n \neq n', n, n' \in N_1\}$, and the set of arcs for the first and last leg, denoted as $A_2 = \{(n, n')|(n \in N_1 \wedge n' \in N_2) \vee (n \in N_2 \wedge n' \in N_1)\}$. Each arc $(n, n') \in A_1$ is associated with a flight travel time $\tau^f(n, n')$. Each arc $(n, n') \in A_2$ is associated with a ground travel time $\tau^g(n, n')$. The notations for the problem and MDP formulation are summarized in Table 5.1.

Table 5.1. Problem notations

Notations	Descriptions
	Global notations
T	Operational period
N_1	Set of vertiports
N_2	Set of origin and destination of trip requests
N	Set of all nodes, $N = N_1 \cup N_2$
A_1	Set of arcs for the flight leg

Continued on next page

Table 5.1 – continued from previous page

Notations	Descriptions
A_2	Set of arcs for the first and last leg
A	Set of all arcs, $A = A_1 \cup A_2$
$\tau^f(n, n')$	Flight travel time of arc $(n, n') \in A_1$
$\tau^g(n, n')$	Ground travel time of arc $(n, n') \in A_2$
	State variables
K	Set of decision epochs
t_k	Time point at epoch k
r_k	New request arriving at epoch $k \in K$
r_k^+	Origin of request r_k
r_k^-	Destination of request r_k
b_{r_k}	Deadline for arrival at new request r_k's destination
C_k	Set of previously accepted requests
n_r^+	Departure vertiport of request $r \in C_k$
n_r^-	Landing vertiport of request $r \in C_k$
$\bar{a}_{r^+}$	Promised departure time from the origin of request $r \in C_k$
$\bar{b}_{r^-}$	Promised arrival time at the destination of request $r \in C_k$
Θ_k	Set of current dispatching plans at epoch k, where each $\theta \in \Theta_k$ represents a path for individual aircraft
c_θ^f	Cost of implementing path θ
$c_{nn'}^g$	Ground travel cost of arc $(n, n') \in A_2$
χ_θ	Scheduled visiting vertiports in path θ
$\bar{\chi}_\theta$	Scheduled flights in path θ
$\gamma_{r\theta}^{nn'}$	Assignment variables indicating whether flight leg (n, n') in path θ serves request r
$a_{r\theta}^n$	Planned departure time of request r from vertiport n under path θ
S_k	State variables at epoch k
	Decision variables
O_k	Set of all possible trip options at epoch k
n_o^+	Departure vertiport of option $o \in O_k$
n_o^-	Landing vertiport of option $o \in O_k$

Continued on next page

Table 5.1 – continued from previous page

Notations	Descriptions
d_{ok}	Total travel time under option $o \in O_k$
Θ_{ok}	Set of feasible dispatching plans for option $o \in O_k$
z_k	Option inclusion variables
p_k	Pricing variables
y_k	Dispatching variables
x_k	Set of all decision variables at epoch k, i.e., $x_k = (z_k, p_k, y_k)$
$\mathscr{X}^\pi(S_k)$	Set of decision variables by policy π at epoch k
O_k^x	Set of offered trip options
$\bar{O}_k^x$	Complete option set $\bar{O}_k^x = O_k^x \cup \{o_0\}$, where o_0 is the outside option of rejecting all trip options
	Others
$W_{k+1}(S_k, x_k)$	Exogeneous information
u_o	Utility of option o
β_0	Base utility of the air taxi service
β_1	Coefficient of travel time sensitivity
β_2	Coefficient of price sensitivity
ϵ_{ok}	Unknown factors affecting the passenger's utility of option o for request r_k
$R(S_k, x_k, W_{k+1})$	Reward at epoch k
$S^M(S_k, x_k, W_{k+1})$	Transition function
$V(S_k, x_k)$	Reward-to-go value given S_k and x_k

State variables

Let K be the set of decision epochs. Each epoch $k \in K$ begins when new request r_k arrives at the time point t_k. The request r_k is characterized by its placement time t_k, origin r_k^+, destination r_k^-, and the arrival deadline b_{r_k}. Let C_k denote the set of previously accepted requests that have not yet reached their destinations at time t_k. For each request $r \in C_k$, let n_r^+ and n_r^- represent the departure and landing vertiport of the option selected by the passenger, respectively. Let $\bar{a}_{r^+}$ and

$\bar{b}_{r^-}$ represent the promised departure time from its origin and arrival time at its destination, respectively.

Let Θ_k represent the set of current dispatching plans for the fleet of aircraft to serve the requests in C_k. Each path $\theta \in \Theta_k$ corresponds to a operational schedule for a single aircraft, comprising a sequence of flight operations to serve assigned passengers and charging activities at vertiports. Let $\chi_\theta = \{n_1, \ldots, n_{|\chi_\theta|}\}$ represent the ordered sequence of vertiports visited by the aircraft following path θ. Each path $\theta \in \Theta_k$ contains detailed assignment and trip timing information for requests $r \in C_k$. Let the binary variable $\gamma_{r\theta}^{nn'}$ indicate whether flight leg (n, n') in path θ serves request r. The trip timing variables are denoted by $a_{r\theta}^n$, which represents the scheduled departure time of request r from vertiport n along path θ. Each path is constructed to satisfy all path-based electric vehicle routing constraints, including the aircraft seating capacity, passenger trip timing requirements, and battery energy limitations. Moreover, for any given flight in the path, all passengers assigned to that flight should have the same departure time.

Let S_k denote the state of the system at decision epoch k, which is characterized by the above components, i.e., $S_k = (t_k, r_k, C_k, \Theta_k)$. The state S_0 is initialized by $(0, -, \varnothing, \varnothing)$.

Decision variables

Decision variables x_k at epoch k are determined based on the state S_k. These variables encompass three interdependent decision types of determining a set of trip options to offer, pricing, and aircraft dispatching. Before defining these decision variables, we first introduce the concept of trip options.

Let o denote a trip option for request r_k, which specifies the departure and landing vertiport $n_o^+, n_o^- \in N_1$, respectively. Each trip option o represents a complete multimodal journey for request r_k, comprising three sequential legs: the first leg from origin to departure vertiport, (r_k^+, n_o^+), the flight leg between vertiports, (n_o^+, n_o^-), and the last leg from landing vertiport to destination, (n_o^-, r_k^-). For request r_k, the total travel time d_{ok} for option o is calculated by:

$$d_{ok} = \tau^g(r_k^+, n_o^+) + \tau^f(n_o^+, n_o^-) + \tau^g(n_o^-, r_k^-), \tag{5.1}$$

which represents the sum of ground transportation time to the departure vertiport, aerial travel time between vertiports, and ground transportation time from the landing vertiport to the final destination. For clarity and operational efficiency, we constrain our model to provide at most one trip option for each unique pair of departure and landing vertiports. Let $O_k = \{o|(n_o^+, n_o^-) \in A_1, \text{and } (n_o^+, n_o^-) \neq (n_{o'}^+, n_{o'}^-), \forall o' \neq o\}$ denote the set of all trip options that could potentially be offered for the request r_k at epoch k. Whether these options are actually offered to the passenger depends on the platform's offering decisions and operational feasibility.

The decision variables x_k integrate three elements: option inclusion variables z_k, pricing variables p_k, and dispatching variables y_k. The option inclusion variables z_k consist of binary entries z_{ok}, where $z_{ok} = 1$ indicates that trip option $o \in O_k$ is offered to the passenger, and $z_{ok} = 0$ otherwise. The pricing variables p_k contain entries p_{ok} that specify the fare for each offered option $o \in O_k$. To ensure operational feasibility, each trip option must be supported by a compatible aircraft dispatching plan. Let Θ_{ok} represent the set of feasible dispatching plans for each trip option o. The dispatching variables y_k include entries y_θ for each path $\theta \in \Theta_{ok}$, where $y_\theta = 1$ denotes that the path θ is selected to fulfill the service requirement of option o and existing commitments. Thus, the decision variables can be represented as $x_k = (z_k, p_k, y_k)$. For these decision variables to constitute a feasible solution, they must satisfy a set of operational constraints that ensure service continuity, temporal feasibility, and logical consistency between offering an option and having the operational capability to fulfill it. We formulate the feasible space of decision variables for each trip option o as follows:

$$\sum_{\theta \in \Theta_{ok}} \gamma_{r_k \theta}^{n_o^+ n_o^-} y_\theta = z_{ok} \tag{5.2}$$

$$\sum_{\theta \in \Theta_{ok}} \gamma_{r\theta}^{n_r^+ n_r^-} y_\theta = 1, \ \forall r \in C_k \tag{5.3}$$

$$t_k + \tau^g(r_k^+, n_o^+) - M(1 - y_\theta) \leq a_{r_k \theta}^{n_o^+}, \ \forall \theta \in \Theta_{ok} \tag{5.4}$$

$$\bar{a}_{r^+} + \tau^g(r^+, n_r^+) - M(1 - y_\theta) \le a_{r\theta}^{n_r^+}, \ \forall r \in C_k \tag{5.5}$$

$$a_{r_k\theta}^{n_o^+} + \tau^f(n_o^+, n_o^-) + \tau^g(n_o^-, r_k^-) - M(1 - y_\theta) \le b_{r_k}, \ \forall \theta \in \Theta_{ok} \tag{5.6}$$

$$a_{r\theta}^{n_r^+} + \tau^f(n_r^+, n_r^-) + \tau^g(n_r^-, r^-) - M(1 - y_\theta) \le \bar{b}_{r^-}, \ \forall r \in C_k \tag{5.7}$$

$$z_{ok} = \{0, 1\}, \ \forall o \in O_k \tag{5.8}$$

$$p_{ok} > 0, \ \forall o \in O_k \tag{5.9}$$

$$y_\theta = \{0, 1\}, \ \forall \theta \in \Theta_{ok} \tag{5.10}$$

Eq. (5.2) establishes the logical connection between offering a trip option and selecting a compatible dispatching plan: if option o is offered to request r_k, then exactly one path that serves this request with the specified flight leg must be selected. Eq. (5.3) ensures that all not-yet-served requests are fulfilled according to its previously assigned vertiport pairs to maintain service continuity for existing commitments. Eqs. (5.4) and (5.5) enforce the temporal feasibility of passenger arrivals at departure vertiports, which guarantees that passengers of the request r_k and not-yet-served requests can catch their flight on time. Eqs. (5.6) and (5.7) enforce the service requirements for arrival deadlines. Eq. (5.6) ensures that the request r_k reach the destination before its deadline, accounting for flight time and ground transportation after landing. Eq. (5.7) maintains the promised arrival times for existing requests. In Eqs. (5.4) to (5.7), M represents a sufficiently large constant that activates the constraints only when the corresponding path is selected. Eqs. (5.8) to (5.10) define the domains of the decision variables, where y_θ and z_{ok} are binary variables, and prices p_{ok} must be positive.

Exogenous information

The exogenous information $W_{k+1}(S_k, x_k)$ depends on the current states and decisions, including two stochastic elements: the passenger's choice among offered trip options, and the arrival of next request (or the termination of the service period).

The passenger's choice among trip options is determined by the perceived utility of each alternative, which primarily which for air taxi services is primarily influenced by two key factors: travel time and trip price (Long et al., 2023). Based on these factors, we formulate the passenger's utility of the option o using a discrete choice utility model (Train, 2009) as follows :

$$u_o = \beta_0 + \beta_1 d_{ok} + \beta_2 p_{ok} + \epsilon_{ok}. \tag{5.11}$$

Eq. (5.11) consists of observed utility components (the first three terms) and an unobserved utility component (the last term). The observed utility component captures the systematic factors influencing passenger choice. Specifically, the parameter β_0 represents the base utility for the air taxi service. The coefficient β_1 quantifies the disutility associated with travel time, and is negative indicating passenger's preference for shorter trips. Similarly, β_2 measures price sensitivity, and is negative indicating that higher fares reduce the attractiveness of an option. The unobserved utility component ϵ_{ok} accounts for all unknown factors affecting choice of option o, which is an independent and identically distributed (iid) random variable.

After establishing the utility model, we now characterize the passenger's choice behavior and the corresponding reward function. Let o_0 represent the outside option of rejecting all trip options. Based on the decision variable x_k, the set of offered trip options O_k^x is defined as $O_k^x = \{o \in O_k | z_{ok} = 1\}$. The complete choice set $\bar{O}_k^x$ is $\bar{O}_k^x = O_k^x \cup \{o_0\}$. Following utility maximization principle, the passenger of request r_k will select the option o^* that maximizes their utility:

$$o^* = \arg\max_{o \in \bar{O}_k^x} \ u_o. \tag{5.12}$$

This choice directly impacts the system's operational costs and revenue. Let c_θ^f represent the cost of implementing path θ, and let $c_{nn'}^g$ denote the ground travel cost of arc $(n, n') \in A_2$. Based on the selected option o^*, a reward $R(S_k, x_k, W_{k+1})$ is realized, which is calculated as:

$$R(S_k, x_k, W_{k+1}) = \begin{cases} p_{o^*k} - (c^g_{r_k^+ n_{o^*}^+} + c^g_{n_{o^*}^- r_k^-} + \sum_{\theta \in \Theta_{o^*k}} c_\theta^f y_\theta - \sum_{\theta \in \Theta_k} c_\theta^f), \text{ if } o^* \neq o_0, \\ 0, \text{ otherwise.} \end{cases} \tag{5.13}$$

This formulation indicates that when a passenger accepts a trip option, the reward equals the fare revenue minus the marginal service costs. These costs include ground transportation expenses for first and last legs, and the marginal flight operation costs for updating the dispatching plan to serve the new request. When a passenger rejects all offered options, no revenue is generated and no additional costs are incurred, resulting in a zero reward.

The exogenous information also reveals the information of the next request r_{k+1}. If no further requests arrive before the end of the operational period T, the service terminates, concluding the decision process.

Transition function

Based on the state S_k, decision variables x_k, and exogenous information $W_{k+1}(S_k, x_k)$, the transition function determines the next state $S_{k+1} = S^M(S_k, x_k, W_{k+1})$. The system advances to time t_{k+1} upon the arrival of new request r_{k+1}. All the requests in the set C_k with unfulfilled trips before t_{k+1} are carried over to C_{k+1}. If the passenger of r_k selects a trip option that will remain unfulfilled at t_{k+1}, then r_k is also added into C_{k+1} with updated promised departure time $\bar{a}_{r_k^+}$ from its origin and promised arrival time $\bar{b}_{r_k^-}$ at the destination. The dispatching plan is updated according to the passenger's choice. If the passenger selects option o^*, all potential dispatching plans $\theta \in \Theta_{o^*k}$ with $y_\theta = 1$ are executed, and the dispatching plan is updated to Θ_{k+1} at t_{k+1}. Alternatively, if the passenger rejects the service, the fleet continues to follow the original plan Θ_k, which then becomes Θ_{k+1} at time t_{k+1}.

Objective function

Let π be the policy that maps a state S_k to decision $x_k = \mathscr{X}^\pi(S_k)$. We can formulate the objective function as finding the policy $\pi \in \Pi$ that maximizes the expected total reward conditional on the initial state as follows:

$$\max_{\pi \in \Pi} \mathbb{E} \left\{ \sum_{k=0}^{|K|} R(S_k, \mathscr{X}^\pi(S_k)) \middle| S_0 \right\}. \tag{5.14}$$

5.2 Solution Method

In this section, we present the solution method. Theoretically, the sequential decision process model can be solved by dynamic programming based on the Bellman equation:

$$\max_{x_k \in X(S_k)} \mathbb{E}\{R(S_k, x_k) + V(S_k, x_k)\}, \tag{5.15}$$

where the value function $V(S_k, x_k) = \mathbb{E}\{\sum_{j=k+1}^{K} R(S_j, \mathscr{X}^{\pi}(S_j))|S_k, x_k\}$ represents the expected future profits the fleet could obtain based on the current state and decisions. However, this dynamic programming method becomes computationally intractable due to the exponential growth of the state space and the complex action space involving joint decisions on option inclusion, pricing, and dispatching. To address this challenge, we propose an anticipatory policy based on approximate dynamic programming. Our approach consists of three key components. First, we develop a dispatching heuristic to efficiently generate feasible trip options, ensuring that each option can be served by at least one feasible dispatching plan. Second, based on these trip options and dispatching plans, we formulate an anticipatory model (AM) that integrates future profitability estimates and stochastic passenger choice behavior. This allows us to optimize current decisions while accounting for their future system impacts. Third, by exploiting the structural properties of the model, we derive an effective dispatching and pricing strategy. Together, these components form a policy that generates computationally efficient solutions at each decision epoch.

In the following subsections, we first introduce the dispatching heuristic, which serves as a foundation for the AM. We then elaborate on how we exploit the model's structure to develop our dispatching and pricing strategy.

5.2.1 Dispatching heuristic

We propose a dispatching heuristic to construct the potential dispatching plan for each option, by assigning existing compatible flight or creating new flights when necessary.

For ease of presentation, we represent the scheduled flights in path θ as $\bar{\chi}_{\theta}$, where each flight $f_{\theta i} \in \bar{\chi}_\theta$ is featured by the departure vertiport $v^+_{\theta i}$, the departure time $e^+_{\theta i}$, the landing vertiport $v^-_{\theta i}$, the landing time $e^-_{\theta i}$, and the number of available seats $q_{\theta i}$. The dispatching heuristic is presented in Algorithm 5.4. For each option $o \in O_k$, the algorithm first attempts to find compatible existing flights along each path $\theta \in \Theta_k$ that can accommodate the flight leg (n^+_o, n^-_o) for request r_k (see Lines 2–7). A compatible flight must have vacant seats and satisfy the temporal constraints specified in Eqs. (5.4) – (5.7). In this case, the path to fulfill option o can be generated by assigning an available vacant seat to the passenger without modifying the exisiting flight schedule. In addition, the algorithm explores the possibility of creating a new flight (see Lines 8–13). For option $o \in O_k$, the algorithms evaluates all feasible insertions of a new flight serving the flight leg (n^+_o, n^-_o) into each path $\theta \in \Theta_k$. Finding feasible insertion for a new flight is a variant of the node insertion problem for electric vehicle routing, which can be solved efficiently using the insertion method in Keskin and Çatay (2016). Specifically, the algorithm first verifies that the temporal constraints in Eqs. (5.4) – (5.7) are satisfied. After verifying temporal constraints, the algorithm evaluates the feasibility of the insertion through the following steps. First, it checks battery sufficiency for all flights in the modified path. If battery levels are insufficient, it attempts to adjust the charging durations at vertiports within battery capacity limits and temporal constraints. If no feasible charging plan exists that maintains adequate battery levels throughout the path, the insertion is infeasible.

All newly generated dispatching plans for option o are collected in the set Θ^π_{ok}, which is distinct from Θ_{ok} defined in Section 5.1. While Θ_{ok} represents all theoretically possible dispatching plans, Θ^π_{ok} specifically refers to the potential dispatching plans generated by our dispatching heuristic. For each path $\theta \in \Theta^\pi_{ok}$, let $\bar{\theta} \in \Theta_k$ denote the original path that was modified to create path θ by the dispatching heuristic. If the algorithm can not find any feasible dispatching plans for option o, i.e., $\Theta^\pi_{ok} = \varnothing$, that option is deemed as infeasible. Let $O^\pi_k = \{o \in O_k | \Theta^\pi_{ok} \neq \varnothing\}$ be the set of all feasible trip options at epoch k, which represents all trip options for which at least one feasible dispatching plan exists. For each option $o \in O^\pi_k$ served by the path $\theta \in \Theta_{ok}^{\pi}$, the extra operating cost Δ_{θ}^{o} can be calculated as follows:

$$\Delta_{\theta}^{o} = c_{r_k^+ n_o^+}^{g} + c_{n_o^- r_k^-}^{g} + c_{\theta}^{f} - c_{\bar{\theta}}^{f}, \tag{5.16}$$

which comprises the ground traveling costs between request locations and vertiports, plus any extra costs incurred by the flight. Notably, when option o is served by an existing compatible flight, the marginal flight cost is zero.

Algorithm 5.4 Dispatching heuristic

```
Input:  Θ_k, option o
Output: Θ^π_ok
Θ^π_ok ← ∅
for θ ∈ Θ_k do
    for i = {1, ..., |χ̄_θ|} do
        if Compatible(f_θi, n_o^+, n_o^-, r_k) then
            θ' ← θ; γ_{r_k θ'}^{v^+_{θ'i} v^-_{θ'i}} ← 1; a_{r_k θ}^{v^+_{θ'i}} ← e^+_{θi}; q_{θ'i} ← q_{θ'i} − 1
            Θ^π_ok ← Θ^π_ok ∪ {θ'}
        end if
    end for
    for i = {1, ..., |χ̄_θ|+1} do
        if FeasibleInsertion(χ_θ, i, n_o^+, n_o^-, r_k) then
            θ' ← Insert(χ_θ, i, n_o^+, n_o^-, r_k)
            Θ^π_ok ← Θ^π_ok ∪ {θ'}
        end if
    end for
end for
Return Θ^π_ok
```

5.2.2 Anticipatory model

In this subsection, we formulate an anticipatory model (AM) at each epoch k. The AM employs a multinomial logit (MNL) choice model to explicitly account for stochastic passenger choice behavior. The model is formulated based on the newly generated paths in set Θ_{ok}^{π} for each option o and the existing paths in set Θ_k, aiming to maximize expected profits at each epoch while considering future impacts.

To facilitate computational efficiency, we discretize the pricing space. Let L_o represent the set of discrete price levels for option o, where each level $l \in L_o$ corresponds to a specific price point. We introduce binary pricing variables h_k, where each entry h_{olk} indicates whether price level l is selected for option o. We have

$$\sum_{l \in L_o} h_{olk} = z_{ok}, \forall o \in O_k^{\pi}, \tag{5.17}$$

which ensures that one price level is selected if and only if the option is offered.

With discrete price levels defined, we can now specify the deterministic utility for each option-price combination. For request r_k, the deterministic utility m_{olk} of option o at price level l is:

$$m_{olk} = \beta_0 + \beta_1 d_{ok} + \beta_2 l. \tag{5.18}$$

To capture the uncertainty in passenger choice behavior, we assume that unobserved utility component ϵ_{ok} follows the standard Gumbel distribution with location parameter $\mu = 0$ and scale parameter $\sigma = 1$. Accordingly, we calculate the probability $P(o, h_k)$ that a passenger selects a particular option o from the set of available options:

$$P(o, h_k) = \frac{\sum_{l \in L_o} e^{m_{olk}} h_{olk}}{\sum_{o \in O_k^{\pi}} \sum_{l \in L_o} e^{m_{olk}} h_{olk} + e^{m_0}}, \tag{5.19}$$

where the term m_0 represents the deterministic utility of the outside option. This formulation indicates that the selection probability for any given option depends not only on its own utility but also on the utilities of all other offered options, which are influenced by the pricing decision variable h_k.

Since each option might utilize different paths within dispatching plan Θ_k, we define Θ_k^o as an identical copy of the current dispatching plan Θ_k for each option o. The dispatching variables for option o are defined as $y_k^o = \{y_\theta | \theta \in \Theta_{ok}^{\pi} \cup \Theta_k^o\}$. This allows independent path selection for each option during optimization. Since the fleet's future profitability depends on current dispatching decisions, we denote the value function as $V(S_k, y_k^o)$, which will be elaborated in Section 5.3. Based on these dispatching plans and value function, the objective function of the AM captures the expected profits of two scenarios: (i) the passenger selects an option o from the available set O_k^{π}, and (ii) the passenger rejects all options. If option $o \in O_k^{\pi}$ is selected, its revenue is its price $p_{ok} = \sum_{l \in L_o} l h_{olk}$, and the reward consists of the immediate marginal profit $p_{ok} - \sum_{\theta\in\Theta^{\pi}_{ok}} \Delta^o_\theta y_\theta$, and the future value $V(S_k, y^o_k)$. If the passenger rejects all trip options, the reward will be the current reward-to-go value $V(S_k)$. Weighted by their selection probabilities, the expected profits are calculated as follows:

$$\begin{aligned}
&\sum_{o\in O^\pi_k} P(o,h_k)(p_{ok} - \sum_{\theta\in\Theta^\pi_o} \Delta^o_\theta y_\theta + V(S_k, y^o_k)) + P(o_0, h_k)V(S_k) \\
=&\sum_{o\in O^\pi_k} P(o,h_k)(p_{ok} - \sum_{\theta\in\Theta^\pi_o} \Delta^o_\theta y_\theta + V(S_k, y^o_k)) + (1 - \sum_{o\in O^\pi_k} P(o,h_k))V(S_k) \\
=&\sum_{o\in O^\pi_k} P(o,h_k)(p_{ok} - \sum_{\theta\in\Theta^\pi_o} \Delta^o_\theta y_\theta + V(S_k, y^o_k) - V(S_k)) + V(S_k). \qquad (5.20)
\end{aligned}$$

Based on this expected profit formulation, the anticipatory model [AM$_k$] at epoch k is formulated as follows:

[AM$_k$]

$$\max \sum_{o\in O^\pi_k} P(o,h_k)(p_{ok} - \sum_{\theta\in\Theta^\pi_o} \Delta^o_\theta y_\theta + V(S_k, y^o_k) - V(S_k)) \qquad (5.21)$$

$$\text{s.t. } Eq.(5.17)$$

$$\sum_{l\in L_o} l h_{olk} - p_{ok} = 0,\ \forall o \in O^\pi_k \qquad (5.22)$$

$$\sum_{\theta\in\Theta^\pi_{ok}\cup\Theta^o_k} \gamma^{n^+_r n^-_r}_{r\theta} y_\theta = 1,\ \forall r \in C_k, \forall o \in O^\pi_k \qquad (5.23)$$

$$\sum_{\theta\in\Theta^\pi_{ok}} y_\theta - z_{ok} = 0,\ \forall o \in O^\pi_k \qquad (5.24)$$

$$h_{olk} = \{0,1\},\ \forall l \in L_o, o \in O^\pi_k \qquad (5.25)$$

$$y_\theta = \{0,1\},\ \forall \theta \in \Theta^\pi_{ok} \cup \Theta^o_k, o \in O^\pi_k \qquad (5.26)$$

The objective function in Eq. (5.21) maximizes the expected profit among all potential trip options, where the term $V(S_k)$ from the full expected profit formulation in Eq (5.20) can be excluded as it is independent of the decision variables. For each option o, Eq. (5.22) defines the price p_{ok} as a function of the selected price level variables h_{olk}. Eq. (5.23) ensures continuity of service for previously accepted requests in set C_k, such that each request is served according to its previously chosen option. Eq. (5.24) guarantees that exactly one potential path is selected for serving the request r_k under option o when that option is offered. Since all paths in Θ^{π}_{ok} generated by the dispatching heuristic already satisfy the temporal constraints (5.4) to (5.7), these constraints are implicitly enforced and thus excluded from the model [AM$_k$]. Finally, constraints (5.25) to (5.26) define the range of decision variables.

5.2.3 Anticipatory policy

The original model [AM$_k$] is a nonlinear integer-programming problem, which is computationally intractable for real-time decision-making. To address this challenge, we exploit the problem's inherent structure and propose an efficient anticipatory policy that determines the option offering, pricing and aircraft dispatching.

The model [AM$_k$] contains interdependent binary variables z_k, h_k, and y_k, where z_k controls which options to offer and directly affects both the pricing decisions h_k and dispatching decisions y_k, creating computational challenges. To address this, we seek to eliminate z_k by exploiting the structure of the cost function. For ease of presentation, let $\eta_o(y^o_k) = \sum_{\theta\in\Theta^{\pi}_{ok}} \Delta^o_\theta y_\theta - V(S_k, y^o_k) + V(S_k)$, representing the cost of serving the option o, which incorporates both the immediate operational costs and the opportunity cost of future value function changes. Since this cost function depends on the dispatching decisions y^o_k, we characterize when these decisions are feasible. The decision space of y^o_k is determined by two constraint sets. Let Y^o_0 denote the set satisfying Eqs. (5.23), and $Y^o(z_{ok})$ denote the set satisfying Eq. (5.24), so that $y^o_k \in Y^o_0 \cap Y^o(z_{ok})$. Note that $Y^o(z_{ok})$ depends on the inclusion variable z_{ok}. When $z_{ok} = 1$, $Y^o(z_{ok})$ requires $\sum_{\theta\in\Theta^{\pi}_{ok}} y_\theta = 1$, denoted as Y^o_1. When $z_{ok} = 0$, it requires $\sum_{\theta\in\Theta^{\pi}_{ok}} y_\theta = 0$, denoted as Y^o_2.

The key insight is that the offering decision z_{ok} only affects the feasibility of dispatching decisions y^o_k, not their costs. This allows us to disconnect the decision spaces of y_k and z_k and incorporate the offering choice within the cost structure. We formalize this in the following proposition:

Proposition 1: *For any $y^o_k \in Y^o_0 \cap Y^o(z_{ok})$, there exists an $\bar{y}^o_k \in Y^o_0 \cap Y^o_1$ such that $\eta_o(y^o_k) = \eta_o(\bar{y}^o_k) \sum_{l\in L_o} h_{olk}$.*

Proof: We consider two cases based on the value of z_{ok}.

Case 1: If $z_{ok} = 1$, then $Y^o(z_{ok}) = Y_1^o$. Therefore, for any $y_k^o \in Y_0^o \cap Y_1^o$, setting $\bar{y}_k^o = y_k^o$ yields $\eta_o(y_k^o) = \eta_o(\bar{y}_k^o) \cdot 1 = \eta_o(\bar{y}_k^o) z_{ok}$ holds.

Case 2: If $z_{ok} = 0$, then $Y^o(z_{ok}) = Y_2^o$. By the definition of Y_2^o, for all $\theta \in \Theta_{ok}^\pi$, we have $y_\theta = 0$, which implies $\sum_{\theta \in \Theta_{ok}^\pi} \Delta_\theta^o y_\theta = 0$, and $V(S_k, y_k^o) - V(S_k) = 0$ (no dispatching changes). Therefore, when $z_{ok} = 0$, we have $\eta_o(y_k^o) = 0$. Let $\bar{y}_k^o$ be any element in $Y_0^o \cap Y_1^o$. Since $z_{ok} = 0$, we have $\eta_o(y_k^o) = 0 = \eta_o(\bar{y}_k^o) z_{ok}$. Thus, in both cases, there exists an $\bar{y}_k^o \in Y_0^o \cap Y_1^o$ such that $\eta_o(y_k^o) = \eta_o(\bar{y}_k^o) z_{ok}$. By substituting Eq. (5.17), we have $\eta_o(y_k^o) = \eta_o(\bar{y}_k^o) \sum_{l \in L_o} h_{olk}$. $\square$

Based on this propostion, we transform the model [AM$_k$] into a more computationally efficient formulation [AM2$_k$].

[AM2$_k$]

$$\max \sum_{o \in O_k^\pi} P(o, h_k)\Big(\sum_{l \in L_o} l h_{olk} - \eta_o(\bar{y}_k^o) \sum_{l \in L_o} h_{olk}\Big) \tag{5.27}$$

$$\text{s.t. Eq. (5.25)}$$

$$\sum_{l \in L_o} h_{olk} \leq 1, \ \forall o \in O_k^\pi \tag{5.28}$$

$$\bar{y}_k^o \in Y_0^o \cap Y_1^o, \ \forall o \in O_k^\pi \tag{5.29}$$

This transformation also eliminates the variables p_k by substituting $\sum_{l \in L_o} l h_{olk}$ for p_{ok}. Moreover, this transformation simplifies the constraint structure by replacing Eq. (5.17) with Eq. (5.28).

To analyze the structure of model [AM2$_k$], Let $\bar{y}_k$ represent the vector of dispatching variables with each entry $\bar{y}_k^o$ satisfying constraint (5.29). The transformed model [AM2$_k$] still requires jointly optimizing pricing decisions h_k and dispatching decisions $\bar{y}_k$. We observe that for any fixed pricing decision h_k, the contribution of each option o depends on its cost $\eta_o(\bar{y}_k^o)$. This cost structure suggests that minimizing each option's cost independently could lead to global optimality, which motivates the following proposition.

Proposition 2: *Let $G(h_k, \bar{y}_k)$ denote the objective function of model [AM2$_k$]. If there exists a decision vector $\bar{y}_k^*$ with each component $\bar{y}_k^{o*}$ satisfying $\eta_o(\bar{y}_k^{o*}) \leq \eta_o(\bar{y}_k^o)$ for all $\bar{y}_k^o \in Y_0^o \cap Y_1^o$. Then, for any feasible variable h_k to model [AM2$_k$], we have $G(h_k, \bar{y}_k^*) \geq G(h_k, \bar{y}_k)$, for all feasible $\bar{y}_k$.*

Proof: We have the difference between the objective function values as follows:

$$
\begin{aligned}
&G(h_k, \bar{y}_k^*) - G(h_k, \bar{y}_k) \\
&= \sum_{o \in O_k^\pi} P(o, h_k)((\eta_o(\bar{y}_k^o) - \eta_o(\bar{y}_k^{o*})) \sum_{l \in L_o} h_{olk})
\end{aligned} \tag{5.30}
$$

By the hypothesis of the lemma, we have $\eta_o(\bar{y}_k^o) - \eta_o(\bar{y}_k^{o*}) \geq 0$ for all $o \in O_k^\pi$. Given that $P(o, h_k) \geq 0$ and $\sum_{l \in L_o} h_{olk} \geq 0$ for all $o \in O_k^\pi$, each term in the summation is non-negative:

$$P(o, h_k)((\eta_o(\bar{y}_k^o) - \eta_o(\bar{y}_k^{o*})) \sum_{l \in L_o} h_{olk}) \geq 0, \tag{5.31}$$

which implies $G(h_k, \bar{y}_k^*) - G(h_k, \bar{y}_k) \geq 0$. Hence, we have $G(h_k, \bar{y}_k^*) \geq G(h_k, \bar{y}_k)$ for all feasible $\bar{y}_k$. □

This proposition reveals that the optimal dispatching decisions can be determined independently of the pricing decisions without affecting the optimality of [AM2$_k$]. Hence, we design our dispatching strategy, which obtains $\bar{y}_k^*$ by solving the following problem for each option $o \in O_k^\pi$ independently.

$$\min_{\bar{y}_k^o \in Y_0^o \cap Y_1^o} \sum_{\theta \in \Theta_{ok}^\pi} \Delta_\theta^o y_\theta - V(S_k, \bar{y}_k^o) + V(S_k) \tag{5.32}$$

This problem determines $\bar{y}_k^{o*}$ as the lowest cost solution for serving option o, which enables us to solve the dispatching problem as a preprocessing step, significantly reducing the computational complexity of the overall model.

Proposition 3: *Given the optimal dispatching decisions $\bar{y}_k^*$ obtained from our dispatching strategy, the problem of determining the set of options to offer and their optimal prices at each epoch k can be reformulated as the tractable model [AM3$_k$],* *as follows:*

$$\max \frac{\sum_{o\in O_k^\pi}\sum_{l\in L_o} e^{m_{olk}}(l-\eta_o(\bar{y}_k^{o*}))h_{olk}}{\sum_{o\in O_k^\pi}\sum_{l\in L_o} e^{m_{olk}}h_{olk}+e^{m_0}} \tag{5.33}$$

$$\text{s.t. Eqs. (5.25) and (5.28)}$$

Proof: Substituting $\bar{y}_k^*$ and Eq. (5.19) into the objective function Eq. (5.27), we have

$$\frac{\sum_{o\in O_k^\pi}(\sum_{l\in L_o} e^{m_{olk}}h_{olk})(\sum_{l\in L_o}(l-\eta_o(\bar{y}_k^{o*}))h_{olk})}{\sum_{o\in O_k^\pi}\sum_{l\in L_o} e^{m_{olk}}h_{olk}+e^{m_0}} \tag{5.34}$$

For each $o \in O_k^\pi$, we expand the the product of sums inside the numerator as follows:

$$\begin{aligned}&\sum_{l\in L_o}\sum_{l'\in L_o} e^{m_{olk}}(l'-\eta_o(\bar{y}_k^{o*}))h_{olk}h_{ol'k})\\ =&\sum_{l\in L_o} e^{m_{olk}}(l-\eta_o(\bar{y}_k^{o*}))h_{olk}^2)+\sum_{l\in L_o}\sum_{l'\in L_o,l'\neq l} e^{m_{olk}}(l'-\eta_o(\bar{y}_k^{o*}))h_{olk}h_{ol'k}).\end{aligned} \tag{5.35}$$

Since Eq. (5.28) ensures that $h_{olk}h_{ol'k}=0$, for all $l\neq l'$, the second term in Eq.(5.35) vanishes. For the first term, because h_{olk} is binary, we have $h_{olk}^2=h_{olk}$. Thus, the expression simplifies to:

$$\sum_{l\in L_o} e^{m_{olk}}(l-\eta_o(\bar{y}_k^{o*}))h_{olk}) \tag{5.36}$$

Replacing the original product of sums in Eq. (5.34) with Eq. (5.36), we thus obtain the objective function in the proposition consequently. Since $\bar{y}_k^*$ obtained by our dispatching strategy satisfies Eq. (5.29), Eq. (5.29) can be omitted safely in [AM2$_k$] after substitution of $\bar{y}_k^*$. Thus, [AM2$_k$] is equivalent to the model in the proposition. □

While Proposition 3 provides a tractable formulation, the pricing problem still requires searching over all discrete price levels in L_o for each option. To further improve computational efficiency, we analyze the structure of optimal pricing decisions.

Proposition 4: *The optimal pricing variable h_k^* satisfies that for each option o, either (i) $h_{olk}^*=0$, $\forall l\in L_o$, or (ii) $\exists l\in L_o$, $h_{olk}^*=1$ with $l\geq\eta_o(\bar{y}_k^{o*})$, and $h_{ol'k}^*=0$ for all $l'\neq l$ (exactly one price offered, at or above the threshold).* .

Proof: For the decision variable h_k to achieve optimality, it must satisfy several conditions, which we introduce sequentially.

Condition 1: Any optimal solution must satisfy $G(h_k) >= 0$. We prove this by contradiction. Suppose there exists an optimal h_k^* such that $G(h_k^*) < 0$. We can construct a feasible solution h_k' by either setting $h_{olk}' = 0, \forall o \in O_k^\pi, l \in L_o$ or setting $h_{olk}' = 1$ if and only if $l = \eta_o(\bar{y}_k^{o*})$, otherwise setting $h_{olk}' = 0$. In both cases, we have $G(h_k') = 0 > G(h_k^*)$ so that h_k' dominates h_k^*, contradicting the assumption that h_k^* is optimal. Therefore, any optimal solution must satisfy $G(h_k) >= 0$.

Let $h_{ok} = \{h_{olk} | l \in L_o\}$ be the decision variables for option o, and we have $h_k = \{h_{ok} | o \in O_k^\pi\}$. Let $E(h_{ok}) = \sum_{l \in L_o} e^{m_{olk}}(l - \eta_o(\bar{y}_k^{o*}))h_{olk}$ be the numerator of the objective function $G(h_k)$, and $D(h_k) = \sum_{o \in O_k^\pi} \sum_{l \in L_o} e^{m_{olk}} h_{olk} + e^{m_0}$ be the denominator. Then, we can reformulate $G(h_k)$ as

$$G(h_k) = \frac{\sum_{o \in O_k^\pi} E(h_{ok})}{D(h_k)}. \tag{5.37}$$

Condition 2: For each option $o \in O_k^\pi$, any optimal solution satisfies $E(h_{ok}) \geq 0$. We prove it by contradiction. Suppose there exists an optimal h_k^* and a subset $O_1 \subset O_k^\pi$ such that $E(h_{ok}^*) < 0$ for all $o \in O_1$. Since $G(h_k^*) \geq 0$ by Condition 1, we must have $\sum_{o \in O_2} E(h_{ok}^*) \geq |\sum_{o \in O_1} E(h_{ok}^*)|$ where $O_2 = O_k^\pi \setminus O_1$. Then we can construct a feasible solution h_k' as follows:

- For $o \in O_1$: Set $h_{olk}' = 1$ if and only if $l = \eta_o(\bar{y}_k^{o*})$, otherwise setting $h_{olk}' = 0$.
- For $o \in O_2$: Keep $h_{ok}' = h_{ok}^*$.

Then we have $E(h_{ok}') = 0, \forall o \in O_1$, and $E(h_{ok}') = E(h_{ok}^*) \geq 0, \forall o \in O_2$. For each $o \in O_1$, let l_o^* denote the price level where $h_{ol_o^*k} = 1$ in the original solution h_{ok}^*, and let l_o' denote the price level where $h_{ol_o'k} = 1$ in the constructed solution h_{ok}'. Since, for each option $o \in O_1$, $E(h_{ok}^*) < 0$, we must have $l_o^* < \eta_o(\bar{y}_k^{o*})$. Since $e^{m_{olk}}$ is decreasing in l and $l_o^* < \eta_o(\bar{y}_k^{o*}) = l_o', \forall o \in O_1$, and $l_o^* = l_o', \forall o \in O_2$ we have

$$D(h_k') - D(h_k^*)$$

$$= \sum_{o \in O_1} (e^{m_{ol'_o k}} - e^{m_{ol^*_o k}}) + \sum_{o \in O_2} (e^{m_{ol'_o k}} - e^{m_{ol^*_o k}})$$

$$= \sum_{o \in O_1} e^{m_{ol'_o k}} - e^{m_{ol^*_o k}} < 0 \tag{5.38}$$

Then we compare the objective function:

$$G(h'_k) - G(h^*_k) = \frac{\sum_{o \in O_1} (E(h'_{ok}) D(h^*_k) - E(h^*_{ok}) D(h'_k)) + \sum_{o \in O_2} E(h^*_{ok})(D(h^*_k) - D(h'_k))}{D(h'_k) D(h^*_k)}. \tag{5.39}$$

Since

- $E(h'_{ok}) = 0, E(h^*_{ok}) < 0, \forall o \in O_1$, then the first sum is greater than 0.
- Since $E(h^*_{ok}) > 0, \forall o \in O_2$, and $D(h^*_k) - D(h'_k) > 0$, then the second sum $\sum_{o \in O_2} E(h^*_{ok})(D(h^*_k) - D(h'_k)) \geq 0$.
- $D(h^*_k), D(h'_k) > 0$

Therefore, $G(h'_k) > G(h^*_k)$, contradicting optimality of h^*_k

From Condition 2, for each option o, we need $E(h_{ok}) \geq 0$. Given the definition $E(h_{ok}) = \sum_{l \in L_o} e^{m_{olk}} (l - \eta_o(\bar{y}^{o*}_k)) h_{olk}$, this is satisfied if and only if (i) either $h_{olk} = 0$ for all $l \in L_o$ (not option offered), or (ii) there exists $l \in L_o$ with $h_{olk} = 1$ and $l \geq \eta_o(\bar{y}^{o*}_k)$ (price at or above the threshold). $\square$

This proposition reveals that options should only be offered if they can be priced profitably above their cost. However, excessively high prices reduce passenger acceptance. To balance profitability and acceptance rate, our pricing strategy restricts the price range of L_o to $[\eta_o(\bar{y}^{o*}_k), \bar{l}_o]$ for each option, where the lower bound $\eta_o(\bar{y}^{o*}_k)$ ensures profitability, and the upper bound $\bar{l}_o$ ensures practical viability. The upper bound $\bar{l}_o$ is calculated as

$$\bar{l}_o = \sup\{l | \frac{e^{m_{olk}}}{e^{m_{olk}} + e^{m_0}} \geq \lambda\}, \tag{5.40}$$

where $\lambda \in (0, 1)$ is the minimum acceptable choice probability. Eq. (5.40) indicates that $\bar{l}_o$ is the highest price at which option o maintains at least λ probability of being selected when compared only to the outside option (rejecting the service).

This parameter λ controls the pricing flexibility, where higher value of λ lead to lower price bounds, while lower values allow more aggressive pricing. By restricting the searching space to this range, we can also reduce computational complexity while ensuring that offered prices remain both profitable and acceptable to passengers.

Based on the restricted price search space, we now present the computational framework of the pricing policy. Let $\bar{L}_o$ denote the restricted price level for option o. By replacing L_o with $\bar{L}_o$ in model [AM3$_k$], we obtain the restricted model [R-AM$_k$]. The [R-AM$_k$] model exhibits two important structural properties: (i) the objective function is both strictly quasi-convex and strictly quasi-concave, and (ii) the coefficient matrix is completely unimodular. As shown by Chen and Hausman (2000), these properties guarantee that the optimal solution to the linear relaxation of the problem is integral.

While these properties ensure integrality, [R-AM$_k$] remains a fractional optimization problem due to the choice probability structure. To this end, we apply the Charnes–Cooper transformation (Cooper and Charnes, 1962) to convert this fractional program into an equivalent linear program. To this end, we introduce a normalization variable $\iota := \frac{1}{\sum_{o\in O_k^\pi}\sum_{l\in\bar{L}_o} e^{m_{olk}}h_{olk}+e^{u_0}}$ that normalizes the denominator of the original objective function to 1 as follows:

$$\iota(\sum_{o\in O_k^\pi}\sum_{l\in\bar{L}_o} e^{m_{olk}}h_{olk} + e^{u_0}) = 1. \tag{5.41}$$

This normalization can transform the fractional objective into a linear one. For each binary variable h_{olk}, we define substitution variables $\upsilon_{olk} := \frac{h_{olk}}{\sum_{o\in O_k^\pi}\sum_{l\in\bar{L}_o} e^{m_{olk}}h_{olk}+e^{m_0}}$. With these transformations, the fractional program [AM2$_k$] can be reformulated as the following linear program:

$$\max \sum_{o\in O_k^\pi}\sum_{l\in\bar{L}_o} e^{m_{olk}}(l-\eta_o(\bar{y}_k^{o*}))\upsilon_{olk} \tag{5.42}$$

$$\text{s.t.} \sum_{l\in\bar{L}_o}\upsilon_{olk} \le \iota,\ \forall o\in O_k^\pi \tag{5.43}$$

$$\sum_{o\in O_k^\pi}\sum_{l\in\bar{L}_o} e^{m_{olk}}\upsilon_{olk} + e^{u_0}\iota = 1 \tag{5.44}$$

Eq. (5.42) represents the transformed objective function from Eq. (5.27), now expressed as a linear function of the new variables. Constraints (5.43) is the transformation of constraints (5.28). Constraint (5.44) is the normalization constraint from Eq. (5.41), where h_{olk} is replaced by the substitution variable $\frac{\upsilon_{olk}}{\iota}$. After solving this linear program, the optimal solution for the original variables can be calculated by $h^*_{olk} = \frac{\upsilon^*_{olk}}{\iota^*}$, for all $o \in O^\pi_k$ and $l \in L_o$. Due to the completely unimodular property of the constraint matrix in the original problem [AM2$_k$], these recovered values will be binary despite solving a continuous relaxation.

5.3 Value Function Approximation

The proposed anticipatory model requires the value function $V(S_k, y^o_k)$ to evaluate future profitability based on the fleet's state and dispatching decisions. However, exact computation of this value function is intractable due to the curse of dimensionality, we propose a DNN-based VFA than learns the value function through offline simulation. Our approach first conducts offline training, where we simulate the operation and use the observed states, decisions, and rewards to train the neural network. After the offline training is complete, the well-trained network can be deployed to provide real-time value estimates for APDP.

5.3.1 State aggregation

The original state and decision spaces exhibit high dimensionality, which significantly increases computational complexity and hinder the learning process. To address this challenge, we propose a state aggregation method that transforms the complex fleet-level value evaluation into more manageable aircraft-level evaluation.

Specifically, given the dispatching decision y^o_k, let $\bar{\Theta}^\pi_{ok} = \{\theta \in \Theta^\pi_{ok} \cup \Theta^o_k | y_\theta = 1\}$ represent the set of selected dispatching plans. We define the reward-to-go value of the fleet $V(S_k, y^o_k)$ as the sum of the values of each individual aircraft as follows:

$$V(S_k, y^o_k) = \sum_{\theta \in \bar{\Theta}^\pi_{ok}} V(\theta), \tag{5.45}$$

where $V(\theta)$ represents the expected cumulative reward profit an aircraft can obtain by following path θ.

To effectively approximate $V(\theta)$, we capture the critical factors that determine an aircraft's future profitability by characterizing the path θ using five representative attributes. The first three are classical attributes commonly used in SDVRPs (Soeffker et al., 2022). We first define them in our problem context.

Classical attributes

Recall that $e^+_{\theta i}$ is the scheduled departure time of flight $f_{\theta i}$ in path θ. To evaluate an aircraft's future profitability, we need to track when each aircraft becomes available for future flights, since earlier available time allows aircraft to serve more demand in the future. We define the first attribute, earliest available time $\mathbb{d}_\theta$, as follows:

$$\mathbb{d}_\theta = \begin{cases} t_k, \ \text{if } e^+_{\theta 1} > t_k, \\ e^-_{\theta 1}, \ \text{otherwise}, \end{cases} \tag{5.46}$$

where $e^+_{\theta 1}$ and $e^-_{\theta 1}$ are the departure and landing times of the first flight in path θ. If the first flight has not yet departed, the aircraft is available at the current time t_k; otherwise, it becomes available after completing its first flight at time $e^-_{\theta 1}$.

The second attribute is the remaining operational time denoted by $\mathbb{b}_\theta = T - e^-_{|\bar{\chi}_\theta|}$, which represents the time left after completing all flight legs in $\bar{\chi}_\theta$, where $e^-_{|\bar{\chi}_\theta|}$ denotes the landing time of the last flight in path θ. A larger value of $\mathbb{b}_\theta$ indicates greater potential for the aircraft to serve additional requests in the future.

Let $\bar{\chi}'_\theta = \{f_{\theta i} \in \bar{\chi}_\theta | e^+_{\theta i} > t_k\}$ denote the set of flight legs that have not yet departed at the current decision epoch. We capture the revenue potential from available seats along the path with the third attribute $\mathbb{m}_\theta$, defined as follows:

$$\mathbb{m}_\theta = \sum_{f_{\theta i} \in \bar{\chi}'_\theta} \frac{q_{\theta i}}{Q}. \tag{5.47}$$

Recall that $q_{\theta i}$ represents the number of available seats on flight f_i. This attribute quantifies the proportion of available seats for future passengers.

In addition to the three aforementioned attributes, we introduce two new attributes tailored to our problem.

Time efficiency

We introduce the fourth attribute, the 'trip time efficiency' of a given path, which captures competitive advantage that drives demand and enables premium pricing, as research has shown that potential time savings compared with ground transportation is a critical factor for attracting passengers (Long et al., 2023). We develop this attribute by first defining 'time efficiency' of each individual flight-request pairs, then aggregating these efficiencies along the path to measure the overall 'time efficiency' of the path. Specifically, let $\delta_{\theta ir}$ denote the 'time efficiency' of flight $f_{\theta i}$ with respect to request r, calculated by:

$$\delta_{\theta ir} = 1 - \frac{\tau^g(r^+, v_{\theta i}^+) + \tau^f(v_{\theta i}^+, v_{\theta i}^-) + \tau^g(v_{\theta i}^-, r^-)}{\tau^g(r^+, r^-)}, \tag{5.48}$$

where the numerator represents the total travel time of the multimodal journey, and the denominator represents the direct ground transportation time between origin and destination. This ratio captures the proportion of time saved when the passenger of request r takes a multimodal journey via flight $f_{\theta i}$ compared to using ground transportation alone.

Based on this individual 'time efficiency' $\delta_{\theta ir}$, we now define the 'time efficiency' of a flight leg. While $\delta_{\theta ir}$ captures time efficiency from the perspective of directness, a comprehensive evaluation must also incorporate temporal considerations. The efficiency of a flight leg becomes irrelevant if the flight's timing does not accommodate passengers' temporal requirements. Let $D(t, t')$ represent the set of expected compatible requests within the time interval $[t, t']$. This set contains requests that satisfies all temporal constraints specified in Eqs. (5.4)–(5.7) and can be obtained through offline simulation based on historical operational data. Let ζ_1 denote a predefined time window before departure. The popularity of the flight leg $f_{\theta i}$, denoted as $TE(f_{\theta i})$, is calculated as follows:

$$TE(f_{\theta i}) = \frac{1}{|D(e_{\theta i}^+ - \zeta_1, e_{\theta i}^+)|} \sum_{r \in D(e_{\theta i}^+ - \zeta_1, e_{\theta i}^+)} \delta_{\theta ir}, \tag{5.49}$$

which represents the average 'time efficiency' across all requests expected to arrive within ζ_1 before the departure of the flight. A higher value of $TE(f_{\theta i})$ indicates that the flight leg offers greater time savings for potential passengers, potentially attracting more demand and increasing profits.

Finally, we compute the overall 'time efficiency' of path θ by combining the 'time efficiency' value across all flight legs as follows:

$$\mathbb{p}_\theta = \frac{1}{|\chi'_\theta|} \sum_{f_{\theta i} \in \chi'_\theta} TE(f_{\theta i}). \tag{5.50}$$

This attribute captures the overall attractiveness of the path by averaging the 'time efficiency' value of all its future flight legs. The inclusion of this metric enables the value function to account for the potential demand patterns when evaluating different paths.

Coverage overlap

The fifth attribute, which we term 'coverage overlap', denoted as $\mathbb{s}_\theta$ for plan θ, reflects the effective coverage area of the aircraft flight plan and its similarity to others in the fleet. This metric serves as a crucial indicator of path coordination: excessive overlap may lead to redundant coverage and inefficient use of aircraft, while insufficient overlap may reduce on-demand service availability in low-coverage areas. We develop this concept through a hierarchical approach, starting from vertiport-level overlap and building up to fleet-level overlap.

We first introduce the coverage overlap between two vertiports. We model each vertiport's service area as a circle, representing the typical range within which passengers can conveniently access the vertiport. Let $s_1(n, n')$ denote the area of intersection between these circular service areas for vertiports n and $n' \in N$. A larger value of $s_1(n, n')$ indicates that the vertiports are located closer to each other, resulting in greater overlap in their service areas.

Building on this concept, we can measure coverage overlap between flight legs. Since only flights operating in similar time periods would compete for the same passenger demand, we incorporate a temporal dimension in our analysis. Let ζ_3

represent a time threshold that defines when two flights are considered temporally proximate enough to potentially serve the same passenger segments. For each flight leg $f_{\theta i}$ in path θ, we define its coverage overlap with respect to another flight $f_{\theta' j}$ in path θ' as follows:

$$s_2(f_{\theta i}, f_{\theta' j}) = \begin{cases} s_1(v^+_{\theta i}, v^+_{\theta' j}) + s_1(v^-_{\theta i}, v^-_{\theta' j}), \text{ if } e^+_{\theta i} - \zeta_3 \leq e^+_{\theta' j} \leq e^+_{\theta i} + \zeta_3, \\ 0, \text{ otherwise.} \end{cases} \tag{5.51}$$

This equation calculates the sum of the intersection area between their departure vertiports ($v^+_{\theta i}$ and $v^+_{\theta' j}$) and landing vertiports ($v^-_{\theta i}$ and $v^-_{\theta' j}$), only if their departure time are within ζ_3 time units. If the flights operate at significantly different times, their coverage overlap is considered zero regardless of spatial proximity, as they would serve temporally distinct passenger markets. A larger value of $s_2(f_{\theta i}, f_{\theta' j})$ indicates higher coverage between two flight legs, while a lower value indicates two paths are complementary by serving different passengers segments.

Extending this concept to the fleet level, we calculate the coverage overlap between two aircraft following θ and θ', respectively as follows:

$$s_3(\theta, \theta') = \sum_{f_i \in \bar{\mathcal{X}}'_\theta} \sum_{f_j \in \bar{\mathcal{X}}'_{\theta'}} s_2(f_{\theta i}, f_{\theta' j}), \tag{5.52}$$

which aggregates the overlap between the two aircraft's planned paths by summing the similarities between all pairs of their respective flight legs. Finally, we compute the average coverage overlap of the aircraft following path θ with all other aircraft in the fleet:

$$\mathbb{s}_\theta = \frac{1}{|\bar{\Theta}^\pi_{ok}| - 1} \sum_{\theta' \in \bar{\Theta}^\pi_{ok} \setminus \{\theta\}} s_3(\theta, \theta'). \tag{5.53}$$

To ensure comparability, all the proposed five attributes are normalized to a standard scale of $[0, 1]$. Based on the five attributes, we can now represent any plan θ as a feature vector $\mathbb{A}_\theta = (\mathbb{d}_\theta, \mathbb{b}_\theta, \mathbb{m}_\theta, \mathbb{p}_\theta, \mathbb{s}_\theta)$. The feature vector $\mathbb{A}_\theta$ serves as the input to our value function $V(\cdot)$, enabling efficient evaluation.

5.3.2 DNN-based VFA

Traditional VFA often relies on linear architectures to approximate value functions, which struggles to capture complex nonlinear relationships in the SDVRPs as the dimensionality increases (Soeffker et al., 2022). To address this limitation, we employ DNNs, which can effectively capture complex non-linear relationships through multiple hidden layers, when estimating the reward-to-go values from our representative attributes. Our DNN-based VFA learns the value function through offline simulation, which allows us to leverage historical training dataset, enabling computational efficiency for online implementation.

Our approach uses two neural networks: a main neural network $\hat{V}_1^{\xi}$ that approximates the value function, and a target network $\hat{V}_2^{\xi}$ that provides stable target values during training. This dual-network architecture, adapted from Deep Q-Learning (Mnih et al., 2015), prevents the instability that arises when the same network generates both predictions and targets. The target network parameters ξ' are periodically updated by copying from the main network, while the main network parameter ξ are updated continuously through gradient descent.

We generate training data through offline simulation using a set of scenarios Ω, where each scenario $\omega \in \Omega$ contains a sequence of randomly generated requests over the operation period. During simulation, we apply our proposed policy supported by $\hat{V}_1^{\xi}$ to make decisions at every epoch. For each epoch k under scenario ω, we collect aircraft-level experiences rather than full fleet states, based on our state aggregation approach. Specifically, when the passenger selects option o^* for request r_k, we record the experience tuple $(\theta^*, \psi, \tilde{\theta}^*)$, where θ^* is the path for serving the request r_k, ψ is the observed reward $R(S_k, x_k, W_{k+1})$, and $\tilde{\theta}^*$ is the updated path after transition at epoch $k+1$. These tuples are stored in an experience replay memory $\mathscr{M}$, from which we randomly sample batches for training to improve learning stability and efficiency (Fedus et al., 2020).

The main network is trained to minimize the mean squared error between predicted and target values. For a bach $\mathscr{B} \subset \mathscr{M}$, the loss function is:

$$\mathscr{L}(\xi) = \frac{1}{|\mathscr{B}|} \sum_{(\theta,\psi,\tilde{\theta})\in\mathscr{B}} (\psi + \alpha \hat{V}_2^{\xi'}(\mathbb{A}_{\tilde{\theta}}) - \hat{V}_1^{\xi}(\mathbb{A}_{\theta}))^2, \tag{5.54}$$

where the parameter α serves as a discount factor scaling the contribution of the future value. The target value $\psi+\alpha\hat{V}_2^{\xi'}(\mathbb{A}_{\tilde{\theta}})$ combines the immediate reward with the discounted future value estimated by the target network. Parameters are updated via stochastic gradient descent:

$$\xi \leftarrow \xi - \rho \nabla_{\xi} \mathscr{L}(\xi), \tag{5.55}$$

where the parameter ρ is the learning rate.

Algorithm 5.5 outlines the training process of our DNN-based VFA. The algorithm begins with an initialized value function $\hat{V}_1^{\xi}(\cdot)$, and a set of sampled scenarios Ω. The experience replay memory is initialized as an empty set (Line 1). The offline simulation proceeds for J iterations, with each iteration $j \in J$ using a scenario ω^j sampled from the scenario set Ω. (Lines 2-4). Within each scenario, decisions at each epoch k are made based on the current state and policy π^j supported by $\hat{V}_1^{\xi}(\cdot)$. At each epoch, the observed states, rewards, and transitions under policy π^j are stored in the experience replay memory (Lines 5-8). The main network parameters ξ are updated using experiences sampled from this memory based on Eq. (5.55), while the target network parameters ξ' are synchronizing periodically (Line 9). After processing all J scenarios, the trained value function $\hat{V}_1^{\xi}(\cdot)$ is returned (Line 13), which can be used for online deployment.

5.4 Numerical Experiments

This section presents the numerical experiments to evaluate the performance of our proposed APDP. We first describe the instance generation and parameter settings. Next, we evaluate the performance of the APDP against several benchmark policies. We then conduct a detailed analysis to explore the specific advantages of our approach regarding profitability and passenger experience. Finally, we perform sensitivity analyses on critical parameters to assess the robustness of the APDP.

Algorithm 5.5 Training process of VFA

Input: Initialized value function $\hat{V}_1^{\xi}(\cdot)$, $\hat{V}_2^{\xi'}(\cdot)$, sampled scenarios Ω
Output: Trained value function $\hat{V}_1^{\xi}(\cdot)$

$\mathscr{M} \leftarrow \varnothing$
$j \leftarrow 1$
while $j \leq J$ **do**
  Sample $\omega^j \leftarrow \Omega$
  for $k = 1, \dots, |K^j|$ **do**
    $x_k \leftarrow \mathscr{X}^{\pi^j}(S_k)$
    $S_{k+1} = S^M(S_k, x_k, W_{k+1}^j)$
    $\mathscr{M} \leftarrow \text{UpdateReplay}(\mathscr{M}, (S_k, R(S_k, x_k, W_{k+1}^j), S_{k+1}))$
    $\xi, \xi' \leftarrow \text{UpdateParameter}(\xi, \xi', \mathscr{M})$
  end for
  $j \leftarrow j + 1$
end while
Return $\hat{V}_1^{\xi}(\cdot)$

5.4.1 Instance generation and parameter settings

We consider a 60×60 km^2 square as the service area, with 10 vertiports distributed across the region. The operation duration is set to 360 min. We consider three different fleet sizes: 4, 6, and 8 aircraft. Based on existing eVTOL aircraft prototypes developed for city commuting, the following parameters are adopted: each aircraft can carry up to 5 passengers, has a cruising speed of 200 km/h, a maximum flight duration of 60 min, and requires 25 min for a full charge (Zhou and Tan, 2024). The aircraft operational cost is assumed to be 3 units per km. The speed of the ground transportation used for the first and last legs is 30 km/h. To approximate the real-world travel distance through road networks, the ground distance between two points is calculated as the Euclidean distance multiplied by a factor of 1.4 (Boscoe et al., 2012). The ground travel cost is 0.175 units per km.

For passenger demand, we consider three different total number of requests: 100, 200, and 300. The placement time of requests are uniformly generated within the interval [0, 300]. The origin and destination of each request are randomly generated within the area, with a minimum distance of 20 km between them, as those travelers would realistically consider using air taxi services (Al Haddad et al., 2020). The deadline for each request r_k is set as $t_k + \tau^g(r_k^+, r_k^-) + 30$, which is the placement time plus the direct ground travel time and 30 min. Regarding individual choice behavior, we assume β_0=3, β_1=2, and $\beta_2 = -0.01$.

Considering all combinations of the number of requests and fleet size, we have a total of 9 types of instance. For each type of instance, we evaluate the proposed APDP and benchmark policies by generating 1,000 random sampled scenarios. The algorithmic parameters are described as below. For pricing strategy, the parameter λ is set to 0.3. In the state aggregation, the time window parameters are set to $\zeta_1 = 30$, and $\zeta_2 = 30$. The time threshold ζ_3 that defines when two flights are considered temporally proximate is set to 15. The service range of each vertiport is set to 10 km. For the parameters of the DNN-based VFA, both the main network and target networks consist of four layers, including two hidden layers with a dimension of 128 units each. The learning rate is initialized at 1×10^{-3}, with a decay rate of 0.95 to a minimum value of 1×10^{-5}. The discount factor α is set to 0.9. The experience replay memory has a capacity of 1×10^5. We update the target network every 100 training steps.

5.4.2 Benchmark policies

To evaluate the performance of our proposed APDP, we implement four benchmark polices representing different approaches to trip option offering, pricing, and dispatching strategies. We first consider a traditional operator-designated policy, denoted as SOB, where the service operator takes full control of trip planning by assigning exactly one trip option to each passenger (Bennaceur et al., 2022). The SOB offers the passenger the trip option with the minimum marginal cost derived from the proposed dispatching heuristic. In case of a tie, the option with the shortest travel time for the passenger is selected. The SOB applies the basic pricing strategy. The basic price p_o^b for option o offered to the passenger r_k is calculated by:

$$p_o^b = p_f + \phi_1 \times (\tau^g(r_k^+, n_o^+) + \tau^g(n_o^-, r_k^-)) + \phi_2 \times \tau^f(n_o^+, n_o^-), \tag{5.56}$$

where p_f is a fixed price component, while ϕ_1 and ϕ_2 are pricing factors for ground connection and flight leg distances, respectively. These parameters are determined by using sampled average approximation, selecting the optimal configuration based on their performance on the training dataset.

The other three benchmark policies all offer multiple trip options for passengers, but differ either in their option offering, pricing, or dispatching strategies. The first policy adopts a two-phase framework, which has shown efficiency in some SDVRPs (Bennaceur et al., 2022; Akkerman et al., 2025). We denote this policy as TPDP. We design this benchmark policy to highlight the benefit of joint optimization of option offering, pricing and dispatching in our proposed policy. In TPDP, this two-phase framework first determines the set of options to offer by selecting time-efficient options from the set of all possible options O_k^π. This is achieved by calculating the time efficiency of each option $o \in O_k^\pi$ relative to request r_k:

$$\delta_{ok} = 1 - \frac{\tau^g(r_k^+, n_o^+) + \tau^f(n_o^+, n_o^-) + \tau^g(n_o^-, r_k^-)}{\tau^g(r_k^+, r_k^-)} \tag{5.57}$$

Only options with δ_{ok} greater than a predetermined threshold are considered. We calibrate this threshold by testing values in the range of $[0, 1]$ with intervals of 0.1, and select the value that performs best on the training dataset. Based on the set of selected options, the TPDP applies our proposed framework to determine the price of each selected option and make dispatching decisions.

The second policy applies a dynamic pricing strategy based on opportunity cost calculations (OCP), which has proven efficient for dynamic and stochastic vehicle dispatching problem with multiple choices (Ulmer, 2020). As an extension to the original framework to adapt to our problem setting, OCP employs the same option selection mechanism as TPDP. For each selected option o, the opportunity cost η_o is obtained by solving the problem (5.32), with the reward-to-go value evaluated by the proposed VFA. Moreover, the OCP uses a basic price to establish a minimum price threshold. The price for each trip option o is determined as $p_o = \max\,\{p_o^b, \eta_o\}$, ensuring the price accounts for future profitability when the opportunity cost exceeds the basic price p_o^b. The dispatching decision for each trip option follows Eq. (5.32). The OCP evaluates the impact of opportunity cost-based pricing without incorporating passenger choice behavior, enabling us to quantify the value of integrating choice models in dynamic pricing decisions and highlighting APDP's advantage of accounting for passenger behavior in optimization.

The last policy is the myopic version of the APDP, denoted as MPDP, which applies the same pricing and dispatching framework of APDP but neglects future profitability. This can be achieved simply by setting the reward-to-go value to 0 for all inputs in the proposed framework. This benchmark highlights the specific value of considering future impacts in the decision-making process while maintaining the choice-aware characteristic of the full APDP approach.

5.4.3 Performance evaluation

In what follows, we present the performance evaluation results of APDP and benchmark polices. The results of each type of instance are summarized in the Appendix **??**. Our experimental results demonstrate that the proposed APDP consistently outperforms all benchmark policies across all types of instance.

Table 5.2. Comparison results of APDP and benchmark policies

#N	#M	APDP	SOB	TPDP	OCP	MPDP
100	4	**23,954**	14,422	23,023	20,747	19,439
100	6	**24,786**	14,874	23,527	21,707	20,469
100	8	**24,487**	14,829	23,676	21,666	20,717
200	4	**47,635**	30,667	42,822	40,975	34,658
200	6	**50,160**	31,121	48,090	43,574	41,103
200	8	**51,168**	31,142	48,842	43,692	42,770
300	4	**68,626**	47,416	59,390	58,144	48,639
300	6	**76,892**	47,799	71,085	64,847	58,779
300	8	**78,071**	47,672	74,206	64,630	64,230

The detailed comparison results between APDP and benchmark polices are presented in Table 5.2, where the best results are highlighted in bold. It shows that APDP outperforms all benchmark polices regarding the profits across all scenario types. Figure 5.2 illustrates the comparison results of average profits obtained by APDP and benchmark polices among all instance. Specifically, APDP achieves 59.2% higher profits compared to traditional SOB. This demonstrates the fundamental advantage of offering multiple trip options with intelligent decisions. Compared to two-phase framework, APDP outperforms TPDP by 7.5%, which demonstrates the benefits of joint optimization of the proposed framework. Moreover, APDP

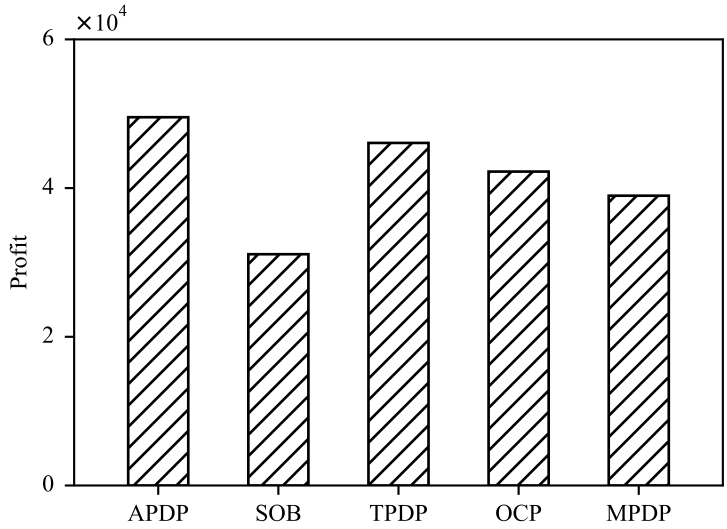


Figure 5.2. Profits obtained by APDP and benchmark polices

achieves 17.3% profits improvement over the OCP, which underscores the importance of explicitly considering passenger choice behavior in the decision-making process. As for MPDP, APDP obtains 27.1% more profits on average, which demonstrates the importance of considering future profitability rather than focusing solely on immediate gains.

For ease of comparison, we represent the relative gaps between the APDP and benchmark policies, denoted by Λ-SOB, Λ-TPDP, Λ-OCP, Λ-MPDP. The gaps are calculated as $\frac{obj^{*}-obj}{obj} \times 100\%$, where obj^{*} represents the objective value obtained by APDP, and obj represents the objective value obtained by benchmark polices. We illustrate the variation of the average relative gaps with increasing number of requests in Figure 5.3. While the relative advantage of APDP over SOB decreases from 66.0% to 56.5% as requests increase from 100 to 300, it maintains a substantial improvement of 56.5% even under high-demand scenarios. More notably, the relative gaps between APDP and more sophisticated benchmark policies widen with increasing number of requests. Specifically, Λ-MPDP increases from 20.8% to 30.3%, Λ-OCP increases from 14.2% to 19.2%, Λ-TPDP increases from 4.3% to 9.2%. These results demonstrate several crucial insights APDP's superior performance. First, the value of considering future profitability (the advantage over MPDP) is amplified under higher request volumes. When the system are loaded with high demand, offering poor trip options can negatively impact the operational efficiency if passen-

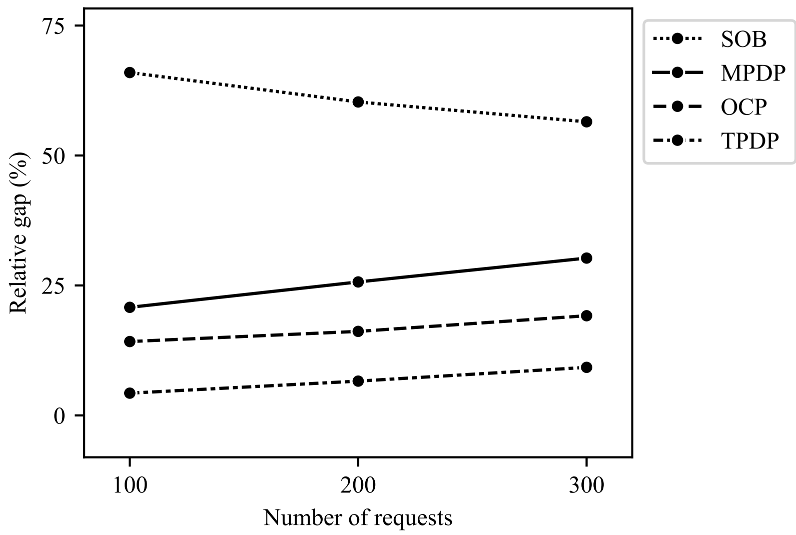


Figure 5.3. Relative gaps between APDP and benchmark polices

gers choose them, as these options utilize aircraft that could be deployed on flights where higher fares could be justified, thus reducing overall profits. Second, the importance of modeling passenger choice behavior (the advantage over OCP) becomes more apparent in busier scenarios, which allows for more strategic option offering and pricing decisions, thus improving operational efficiency. Third, the value of joint optimization (the advantage over TPDP) becomes increasingly critical as the system faces higher demand. It indicates that interdependencies between option offering, pricing and dispatching decisions have significant impact on operational efficiency in high demand scenario. Joint optimization enables APDP to make coherent decisions that account for these complex interdependencies, while TPDP's two-phase approach creates misalignments that increasingly reduce potential profits as demand increases. Together, these mechanisms allow APDP to enhance operational efficiency as demand increases, demonstrating its robustness for real-world air taxi operations.

To evaluate how different policies affect passenger experience, we analyze two key metrics: (i) the the offered ratio, which is the percentage of passengers who receive at least one trip option, and (ii) the acceptance ratio, which is the percentage of passengers who select an option. Figure 5.4 presents these metrics for APDP and four benchmark policies. APDP demonstrates superior performance with the highest offered ratio of 97% and acceptance ratio of 81%, which indicates its effectiveness in both reaching passengers and providing appealing options. When compared to traditional single-option, fixed-pricing approach (SOB), APDP achieves 5% higher offered ratio and 7% higher acceptance ratio. This result demonstrates that providing multiple dynamically-priced options enables the system to serve more passengers by offering alternatives that fit different preferences. As for TPDP, while both achieve the same acceptance ratio, APDP achieved 9% higher offered ratio. It indicates that APDP's joint optimization approach enables the system to offer options to more passengers without compromising service quality. In constrast, TPDP's two-phase approach, which separates option offering from pricing decisions, potentially limits the range of options that can be provided. Compared to OCP, APDP achieves 7% higher offered ratio and 5% higher acceptance ratio. This improvement demonstrates the value of understanding passenger choice behavior. Without consideration of passengers choice preferences, OCP offers options that may seem operational efficiency but fail to align with passenger preferences. APDP's higher acceptance rate confirms that the incorporation of passenger choice behavior into the decision process creates more appealing options. Compared to MPDP, APDP achieves 12% higher offered ratio and 3% higher acceptance ratio, which highlights the importance of considering future impacts. Since MPDP focuses only on immediate profits, it offers options without considering how these offerings affect the system's ability to provide options to future passengers in high-demand areas. This myopic decision-making causes that fewer passengers can be offered options later. The modest improvement in acceptance ratio further suggests that APDP's future-aware approach not only enables broader service coverage but also creates slightly better option quality. These results demonstrate that APDP's comprehensive approach addresses the specific limitations of each benchmark policy, resulting in superior service coverage and passenger satisfaction.

We conduct a detailed sensitivity analysis examining both internal and external factors affecting system performance: the critical components of the proposed policy that we can optimize, and the diverse passenger segments that represent the inherent variability in real-world operations. First, we examine the impact of state aggregation. Then, we investigate the impact of pricing thresholds in our proposed policy on system performance. Finally, we explore how different passenger segments

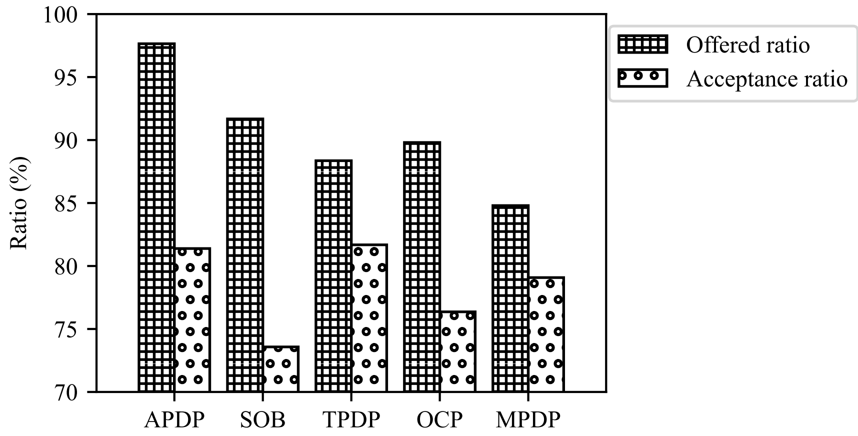


Figure 5.4. Offered and acceptance ratio under APDP and benchmark policies

with varying sensitivities to price and travel time influence system performance.

5.4.4 Impact of state aggregation

To evaluate the effectiveness of our proposed new attributes, we compare the performance of different feature combinations within our framework regarding both profits and passenger experiences. Specifically, we compared APDP with three variants with different state aggregation: APDP3, which uses only three traditional attributes (earliest available time, remaining time, available seats), APDP4 adding time efficiency to capture scheduled flight attractiveness, and APDP4s adding coverage overlap to measure service area redundancy.

Figure 5.5(a) presents the relative gaps between APDP and three variants, which reveals how each attribute contributes to system performance. Compared to APDP3, APDP achieves 14% higher profits, which demonstrates the critical limitations of relying solely on traditional attributes of basic temporal and capacity features. The value of each new attribute becomes evident through incremental comparisons. With coverage overlap, APDP outperforms APDP4 by 6% in profits. This attribute enables the system to prevent service redundancies, leading to improved operational efficiency. With time efficiency, APDP surpasses APDP4s by 5% in profits. This attribute empowers the system to differentiate between high and low-value flight opportunities, actively prioritizing options that align with passenger demand patterns and maximize revenue potential.

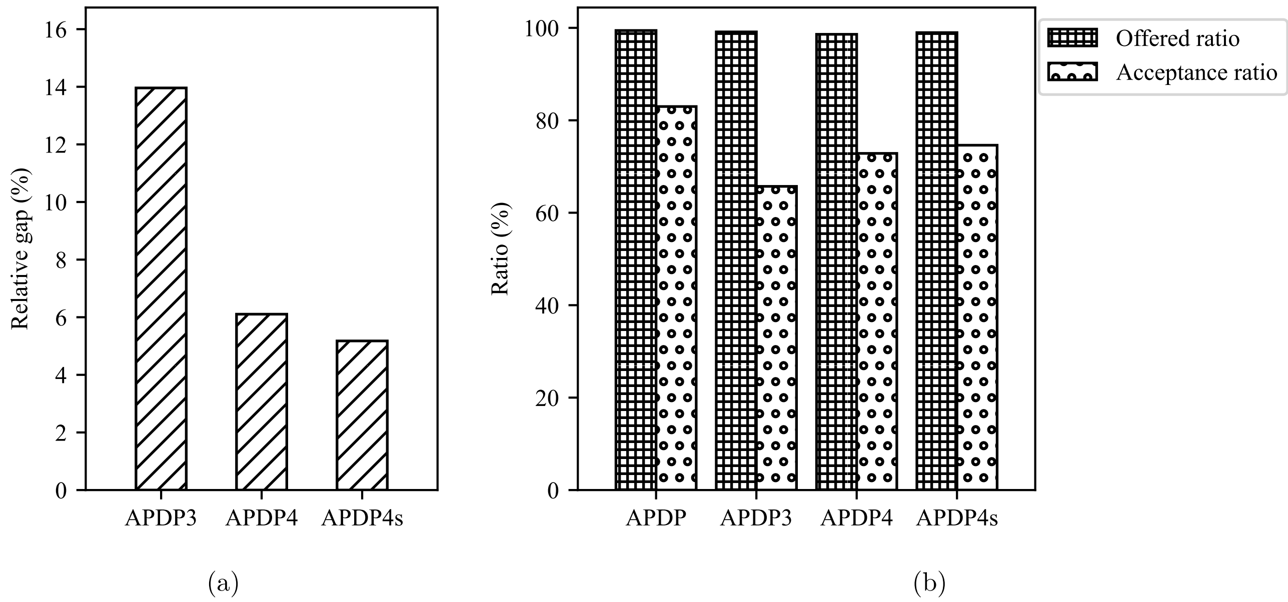


Figure 5.5. Impact of attributes on (a) profits, (b) offered and acceptance ratio

The impact on passenger experience further validates these attributes' importance. While all variants achieve similar offered ratios of approximately 99%, the acceptance ratios shows significant difference: APDP (81%), APDP4s (78%), APDP4 (70%), and APDP3 (62%). This progression clearly shows that each additional attribute enables the system to create more appealing options. Time efficiency helps identify options that better match passenger preferences, while coverage overlap enables the system to maintain service quality across the network, ensuring sustainable operation.

These results confirm that both proposed attributes are essential and complementary. Time efficiency enables demand-aware prioritization while coverage overlap ensures supply-side coordination. The APDP's superior performance demonstrates that effective on-demand systems require comprehensive state representations that capture market dynamics and operational efficiency simultaneously.

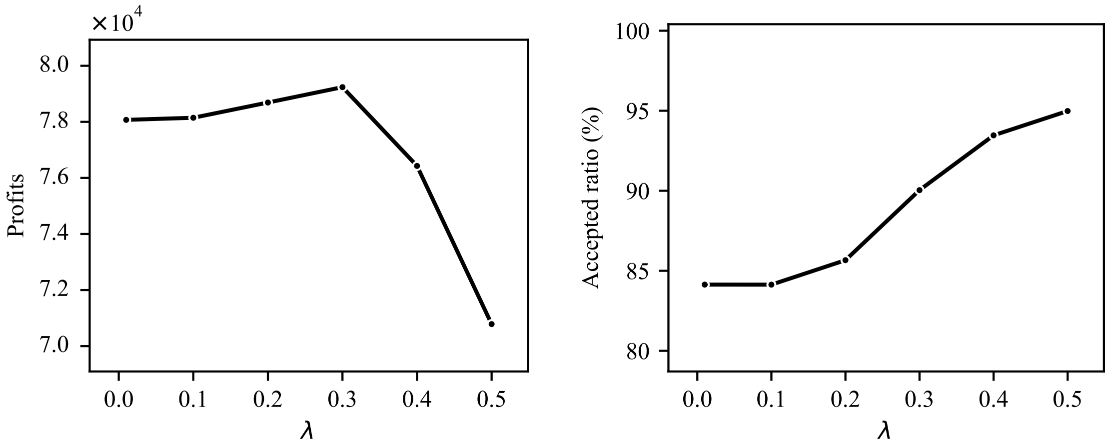
5.4.5 Impact of pricing flexibility

In what follows, we explore the role of pricing flexibility in balancing the trade-off between profitability and service accessibility. This pricing flexibility in our framework is controlled by the parameter λ, where higher value of λ imposes greater restrictions on dynamic pricing, resulting in a narrower price range. To understand this trade-off, we implement six different price flexibility settings with λ ranging from 0 to 0.5, analyzing their impact on profits and acceptance ratio as shown in Figure 5.6. Figure 5.6(a) reveals that when λ initially increases from 0 to 0.3, profits increase by 1.4%, and the acceptance ratio improves from 85% to approximately 90%, as shown in Figure 5.6(b). This suggests that moderate restriction of pricing flexibility can actually benefit the system by serving more passengers while maintaining profitability. The underlying reason is that unlimited price range ($\lambda = 0$) might generate options with excessively high price, which deter potential passengers demand. By slightly limiting the price space, the system avoids providing options with extreme prices, which enables to obtain higher profits and capture more demand. However, as λ increases further from 0.3 to 0.5, profits decline by 11% even though the acceptance ratio continues to rise from 90% to 95%. This result indicates that excessive restrictions of pricing flexibility significantly deteriorate profit potential. Even though, this conservative pricing strategy is beneficial for market penetration, it causes the system to lose opportunities to capture higher willingness-to-pay from certain passengers.

These findings reveal that the pricing flexibility plays a critical role in balancing profitability and service accessibility. The optimal strategy appears to involve moderate restriction of λ to prevent extreme pricing for better system performance. For practical implementation, the operator can adjust λ based on market condition: imposing restriction of pricing flexibility during off-peak period to maintain affordable prices that attract price-sensitive customers and improve fleet utilization, and easing restriction during peak periods to capture higher willingness-to-pay from urgent travelers.

5.4.6 Impact of different target passenger groups

For effective business planning, the comprehension of how various passenger segments respond to air taxi services serves as a fundamental element in developing targeted marketing strategies and optimizing operational policies. In this analysis, we evaluate how the price sensitivity and travel time sensitivity affect system per-



(a) (b)

Figure 5.6. Impact of price flexibility on (a) profits (b) acceptance ratio

formance. We first evaluate three passenger groups regarding low, medium and high price sensitivity, where the utility parameter for price β_1 is set to -0.010, -0.011, and -0.012, respectively. As shown in Figure 5.7(a), targeting high price-sensitive passengers results in a 20% decline in total profits, compared to low price-sensitive segments. This suggests operators should carefully balance their passenger segments to maintain financial sustainability. The operators could initially target less price-sensitive segments to establish profitability before expanding to broader markets. Despite this significant profit difference, similar acceptance ratio are maintained across all groups, as illustrated in Figure 5.7(b). The results reveal a critical advantage of providing multiple options by APDP, which enable the choice-aware service to extract available value from each passenger segments and prevent passenger loss that would occur under traditional single-price models, avoiding the typical trade-off between market coverage and profitability.

We also evaluate three passenger groups with low, medium and high travel time sensitivity, where the utility parameter for the travel time β_2 is set to 1.5, 2.0, and 2.5, respectively. As shown in Figure 5.8(a), targeting on passengers with high travel time sensitivity yields 6% higher profitability, and a 4% increase in option acceptance rates compared to those with low time sensitivity, as illustrated in Figure 5.8(b). This suggests that this choice-aware service can effectively provide suitable options for time-sensitive passengers while capturing the premium they are willing to pay

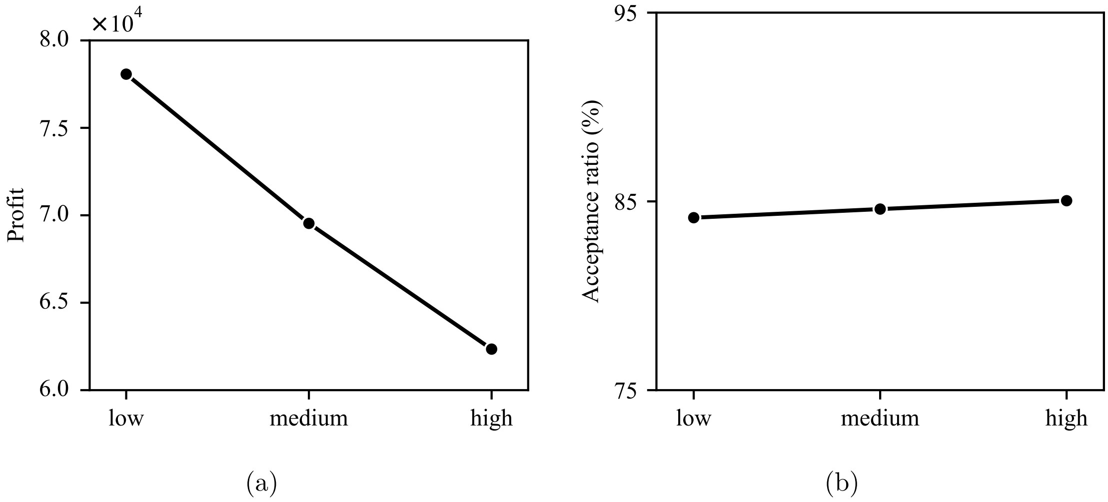


Figure 5.7. Impact of passengers with different price sensitivity on (a) profits, and (b) offered and acceptance ratio

for time savings.

The results of price-sensitive and time-sensitive segments reveal important strategic implications. Operators should strategically design their option portfolios to emphasize time-saving features rather than focusing solely on price differentiation. This strategy leverages air taxis' inherent speed advantage to attract time-conscious customers who demonstrate both higher profitability and greater service acceptance, positioning air taxi services as a premium time-saving solution rather than a cost-competitive alternative to ground transportation.

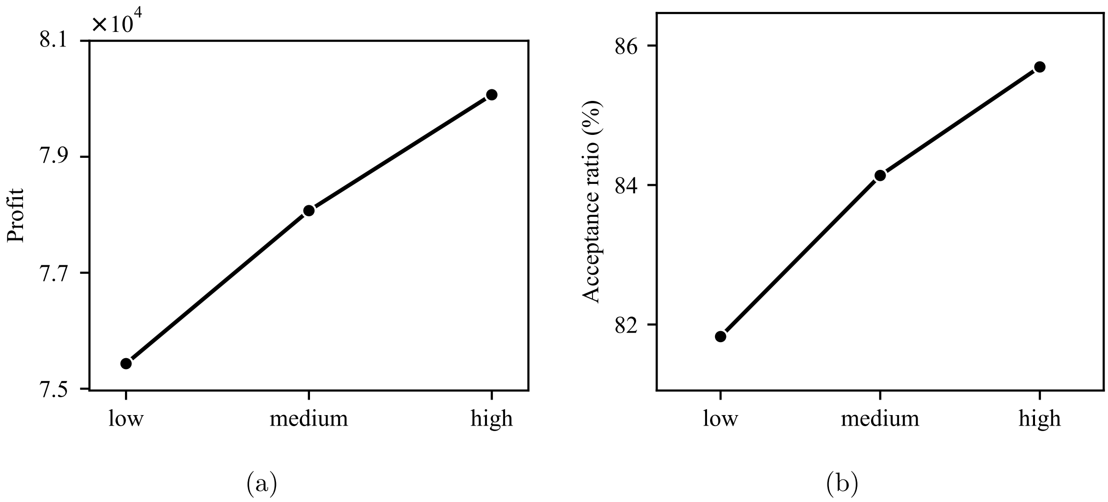


Figure 5.8. Impact of passengers with different travel time sensitivity on (a) profits, and (b) offered and acceptance ratio

5.5 Concluding Remarks

In this chapter, we investigates a choice-aware air taxi service, where the operator offers multiple multimodal trip options to accommodate stochastic passenger preferences. To make sequential decisions, this problem is formulated as a MDP. We propose an anticipatory model to incorporate decisions of option offering, pricing, and aircraft dispatching in a unified framework while explicitly accounting for both passenger's choice behavior and future profitability. By exploiting the problem structure, we propose the APDP for online deployment, which is supported by a DNN-based VFA. Our comprehensive numerical experiments demonstrate that APDP significantly outperforms benchmark policies in terms of both profitability and service levels. The detailed analysis of our results yields several important managerial insights for air taxi operators. First, offering multiple trip options with APDP substantially improves both profitability and service quality compared to other benchmark policies. Second, pricing strategies should balance immediate revenue against passenger acceptance rates, with moderate price thresholds providing optimal results. Finally, strategic implications of passengers segments are derived to support practical guidance of air taxi service.

Chapter 6 Conclusion

This chapter provides a comprehensive summary of the outcomes presented in this thesis. Then, the recommendations for future studies are presented.

6.1 Overview and Research Contributions

This thesis conducts a progressive series of studies on the integration of personalized service mechanisms with operational routing decisions for ground and air mobility platforms: (i) establishing adaptive incentive mechanism and routing decision-making framework for ground-level service consolidation, (ii) designing efficient operational framework for flexible air mobility service with pooling and transfer, and (iii) formulating choice-aware decision-making framework of pricing and dispatching in multimodal air taxi service. Together, they form a cohesive contribution to intelligent mobility management in the future.

Chapter 3 studies a dynamic confirmation, compensation, and routing for combined urban transportation of passengers and parcels, where compensation is adopted as incentive to motivate passengers to accept detour for delivery in the CUT considering the uncertainty of passengers' acceptance. We formulate this dynamic and stochastic problem as an MDP for making sequential decisions. We further propose an anticipatory model and exploit the solution features. Based on the findings, the complexity of the original problem is reduced, and an effective anticipatory policy to make confirmation, compensation and routing decisions, referred to as the ACCR, is proposed to integrate multiple decisions in a unified framework while considering their impact on future profitability. This is achieved by a novel regression-tree-based VFA for learning future profitability given the fleet status. We conduct extensive numerical experiments in a series of instances with varying request numbers, fleet sizes, and spatial distribution of requests. The numerical results have validated the effectiveness of our policy and the significance of the proposed regression-tree-based VFA on the performance of the policy. We also find that compared to moderate integration, which does not allow parcel delivery during passengers' trips, by motivating passengers to accept parcel delivery, full integration can capture more consolidation opportunities with marginal costs, leading to improved profitability.

Chapter 3 demonstrated the critical role of personalized service mechanisms in the successful integration with ground transportation networks. To explore the broader applicability of this concept, we extend the investigation within UAM, by exploring a novel flexible air mobility with passenger pooling and flight transfers in Chapter 4. This problem involves trip planning for passengers and aircraft dispatching and charging decisions to balance the trade-off between the operational costs and individual passenger experience, helping us understand the feasibility of such aerial systems. We formulate an integer programming based on a mixed time-space and time-space-battery graph, which can effectively integrates passenger flight planning, aircraft dispatching, and charging decisions in a unified framework. Given the complexity of the problem, solving the original integer programming model directly using commercial solvers is computationally challenging. To this end, we propose two arc generation algorithms that construct sparse graphs with restricted subset of nodes and arcs to enhance the computational efficiency while maintaining solution quality. We first propose a MAG algorithm to provide the foundational framework to build the sparse graph iteratively. Then, we develop a AAG algorithm by adjusting the construction of sparse graph process based on feedback from previous iterations, which enables more intelligent exploration of the solution space. Numerical experiments demonstrate the excellent performance of AAG algorithm compared to benchmark algorithms regarding solution quality and computational efficiency. Through sensitivity analysis, we find that urban air mobility with pooling and flight transfer is scalable, which has better utilization of available aircraft as more passengers use this service. Furthermore, this system could well balance the cost-effectiveness and convenient service with increased number of orders, which allows decision-makers to control the balance after carefully weighing the benefits in delay reduction against the associated operating costs.

After the investigation within the air mobility network in Chapter 4, Chapter 5 explores a more sophisticated personalized service mechanism within the UAM operation. It focuses on a dynamic pricing and dispatching problem for choice-aware air taxi services, where operators must simultaneously optimize multimodal trip options, pricing strategies, and real-time fleet dispatching in a dynamic and stochastic environment. We formulate this problem as an MDP and develop an anticipatory optimization framework that jointly determines option offerings, pricing, and dispatching decisions while explicitly modeling how these decisions influence passenger choice behavior and future profitability. By exploiting the problem structure, we propose a APDP that makes real-time decisions for each incoming request by balancing immediate revenue with future profitability. This policy leverages a DNN-based VFA, which efficiently learns to estimate future profitability given the fleet status.

Our computational experiments yield important managerial insights for air taxi service operations. First, our choice-aware service framework shows that the coordination of offering multiple dynamically priced options with aircraft dispatching can significantly enhance profits and serve more passengers compared to benchmark policies, and this advantage becomes even more pronounced under high-demand conditions, underscoring the framework's scalability. In addition, we demonstrate the importance of the proposed VFA for air taxi services. By capturing the alignment between scheduled flights and demand patterns, and the strategic redundancy across service areas, the VFA enables the policy to generate options that not only cater to passenger preferences but also maximize revenue potential. Furthermore, the choice-aware service enables operators to extract value from diverse passenger segments, achieving broader market penetration. This capability reveals strategic opportunities for targeted market selection and expansion.

6.2 Recommendations for Future Research Direction

Future research directions on personalized service mechanism and routing for emerging mobility solutions are various regarding the understanding of passenger choice behavior mechanism, strategic service integration, and strategic operational mechanism design.

(1) Understanding of passenger choice behavior mechanisms

The personalized service mechanism will largely depend on passengers' choice behavior. More empirical studies to quantitatively measure passenger sensitivity to detour time, price, waiting time, and shared riding with strangers, are thus necessary for a better understanding of this key parameter. Move beyond static parameters. Research should focus on machine learning models that can predict individual passenger preferences in real-time based on historical choices, context (e.g., different trip purpose), and even real-time mood.

(2) Strategic service integration and partnership models

The value of new mobility solutions can be further expanded when integrated seamlessly into the broader transportation ecosystem. For CUT, integration with public transit can be conducted to explore the role of CUT as a feeder or replacement for low-frequency bus lines, enhancing public transit's efficiency and reach by providing flexible first/last-mile connections and replacing underused routes.

For UAM, vertiport planning can be integrated with our framework to enhance network accessibility and multimodal connectivity. At the operational level, the model could be extended to incorporate strategic relocation of aircraft during off-peak periods to better position the fleet in anticipation of future demand. From a partnership perspective, the integration of dynamic contracts with ground transportation providers in our framework presents opportunities to create more seamless multimodal journeys. Furthermore, research on seamless Mobility-as-a-Service operation can be conducted to enhance the efficiency of both the ground and air mobility systems. This includes optimizing not just the aircraft dispatching but the routing of ground vehicles, considering multifaceted realistic factors involving real-time congestion, stochastic travel time, and low-altitude airspace capacity variations.

(3) Strategic and operational mechanism design

To ensure the practical implementation of the proposed optimization framework, it is necessary to further explore the coordination between optimization methodology and strategic mechanism. The optimization framework's practical applicability could be further enhanced by incorporating considerations of resilience, autonomy, infrastructure, regulation, acceptance, coordination, and equity (Sun et al., 2025). Moreover, the operational efficiency can be enhanced with extra strategic decisions, such as repositioning. The integration of predictive repositioning models that use forecasting based on contextual information such as events and weather can proactively relocate vehicles/aircraft during off-peak periods to reduce idle time and increase availability for future requests. More sophisticated algorithms can be investigated to evaluate the trade-off between waiting for potential consolidation/pooling opportunities and immediately servicing a request to balance the trade-off between passenger experience and system-wide operating costs.

References

Akkerman, Fabian, Dieter, Peter, and Mes, Martijn (2025). Learning dynamic selection and pricing of out-of-home deliveries. *Transportation Science* 59(2), pp. 250–278.

Al Haddad, Christelle, Chaniotakis, Emmanouil, Straubinger, Anna, Plötner, Kay, and Antoniou, Constantinos (2020). Factors affecting the adoption and use of urban air mobility. *Transportation research part A: policy and practice* 132, pp. 696–712.

Ale-Ahmad, Haleh and Mahmassani, Hani S (2021). Capacitated location-allocation-routing problem with time windows for on-demand urban air taxi operation. *Transportation Research Record* 2675(10), pp. 1092–1114.

Allied Market Research (2022). Mobility on Demand Market Size, Share, Competitive Landscape and Trend Analysis Report. Accessed October 28, 2024, https://www.alliedmarketresearch.com/mobility-on-demand-market.

Angelelli, Enrico, Mansini, Renata, and Vindigni, Michele (2016). The stochastic and dynamic traveling purchaser problem. *Transportation Science* 50(2), pp. 642–658.

Arslan, Alp M, Agatz, Niels, Kroon, Leo, and Zuidwijk, Rob (2019). Crowdsourced delivery—a dynamic pickup and delivery problem with ad hoc drivers. *Transportation Science* 53(1), pp. 222–235.

Azadeh, Sh Sharif, Atasoy, Bilge, Ben-Akiva, Moshe E, Bierlaire, M, and Maknoon, MY (2022). Choice-driven dial-a-ride problem for demand responsive mobility service. *Transportation Research Part B: Methodological* 161, pp. 128–149.

Azcuy, Irecis, Agatz, Niels, and Giesen, Ricardo (2021). Designing integrated urban delivery systems using public transport. *Transportation Research Part E: Logistics and Transportation Review* 156, p. 102525.

Azi, Nabila, Gendreau, Michel, and Potvin, Jean-Yves (2012). A dynamic vehicle routing problem with multiple delivery routes. *Annals of Operations Research* 199(1), pp. 103–112.

Beirigo, Breno A, Schulte, Frederik, and Negenborn, Rudy R (2022). A learning-based optimization approach for autonomous ridesharing platforms with service-level contracts and on-demand hiring of idle vehicles. *Transportation Science* 56(3), pp. 677–703.

Bennaceur, Mehdi, Delmas, Rémi, and Hamadi, Youssef (2022). Passenger-centric Urban Air Mobility: Fairness trade-offs and operational efficiency. *Transportation Research Part C: Emerging Technologies* 136, p. 103519.

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

Boscoe, Francis P, Henry, Kevin A, and Zdeb, Michael S (2012). A nationwide comparison of driving distance versus straight-line distance to hospitals. *The Professional Geographer* 64(2), pp. 188–196.

Bosse, Alexander, Ulmer, Marlin W, Manni, Emanuele, and Mattfeld, Dirk C (2023). Dynamic priority rules for combining on-demand passenger transportation and transportation of goods. *European Journal of Operational Research* 309(1), pp. 399–408.

Bowling, Shannon R, Khasawneh, Mohammad T, Kaewkuekool, Sittichai, and Cho, Byung Rae (2009). A logistic approximation to the cumulative normal distribution. *Journal of Industrial Engineering and Management* 2(1), pp. 114–127.

Chen, Chao, Zhang, Daqing, Ma, Xiaojuan, Guo, Bin, Wang, Leye, Wang, Yasha, and Sha, Edwin (2016). Crowddeliver: Planning city-wide package delivery paths leveraging the crowd of taxis. *IEEE Transactions on Intelligent Transportation Systems* 18(6), pp. 1478–1496.

Chen, Kyle D and Hausman, Warren H (2000). Mathematical properties of the optimal product line selection problem using choice-based conjoint analysis. *Management Science* 46(2), pp. 327–332.

Chen, Liting, Wandelt, Sebastian, Dai, Weibin, and Sun, Xiaoqian (2022a). Scalable vertiport hub location selection for air taxi operations in a metropolitan region. *INFORMS Journal on Computing* 34(2), pp. 834–856.

Chen, Xinwei, Ulmer, Marlin W, and Thomas, Barrett W (2022b). Deep Q-learning for same-day delivery with vehicles and drones. *European Journal of Operational Research* 298(3), pp. 939–952.

Chen, Yueyue, Guo, Deke, Xu, Ming, Tang, Guoming, Zhou, Tongqing, and Ren, Bangbang (2019). PPtaxi: Non-stop package delivery via multi-hop ridesharing. *IEEE Transactions on Mobile Computing* 19(11), pp. 2684–2698.

Cheng, Rong, Jiang, Yu, and Nielsen, Otto Anker (2023). Integrated people-and-goods transportation systems: from a literature review to a general framework for future research. *Transport Reviews*, pp. 1–24.

Cooper, A Charnes WW and Charnes, W (1962). Programming with linear fractional functionals. *Naval Research logistics quarterly* 9(3), pp. 181–186.

Corless, Robert M, Gonnet, Gaston H, Hare, David EG, Jeffrey, David J, and Knuth, Donald E (1996). On the Lambert W function. *Advances in Computational Mathematics* 5, pp. 329–359.

Cortés, Cristián E, Matamala, Martín, and Contardo, Claudio (2010). The pickup and delivery problem with transfers: Formulation and a branch-and-cut solution method. *European Journal of Operational Research* 200(3), pp. 711–724.

Espinoza, Daniel, Garcia, Renan, Goycoolea, Marcos, Nemhauser, George L, and Savelsbergh, Martin WP (2008a). Per-seat, on-demand air transportation part I: Problem description and an integer multicommodity flow model. *Transportation Science* 42(3), pp. 263–278.

Espinoza, Daniel, Garcia, Renan, Goycoolea, Marcos, Nemhauser, George L, and Savelsbergh, Martin WP (2008b). Per-seat, on-demand air transportation part II: Parallel local search. *Transportation Science* 42(3), pp. 279–291.

Fedus, William, Ramachandran, Prajit, Agarwal, Rishabh, Bengio, Yoshua, Larochelle, Hugo, Rowland, Mark, and Dabney, Will (2020). "Revisiting fundamentals of experience replay". *International conference on machine learning*. PMLR, pp. 3061–3071.

Fehn, Fabian, Engelhardt, Roman, Dandl, Florian, Bogenberger, Klaus, and Busch, Fritz (2022). Integrating Parcel Deliveries into a Ride-Pooling Service–An Agent-Based Simulation Study. *arXiv preprint arXiv:2205.04718*.

Figliozzi, Miguel Andres, Mahmassani, Hani S, and Jaillet, Patrick (2007). Pricing in dynamic vehicle routing problems. *Transportation Science* 41(3), pp. 302–318.

Fleckenstein, David, Klein, Robert, and Steinhardt, Claudius (2023). Recent advances in integrating demand management and vehicle routing: A methodological review. *European Journal of Operational Research* 306(2), pp. 499–518.

Hildebrandt, Florentin D, Thomas, Barrett W, and Ulmer, Marlin W (2023). Opportunities for reinforcement learning in stochastic dynamic vehicle routing. *Computers & Operations Research* 150, p. 106071.

Hörsting, Lena and Cleophas, Catherine (2023). Scheduling shared passenger and freight transport on a fixed infrastructure. *European Journal of Operational Research* 306(3), pp. 1158–1169.

Husemann, Michael, Kirste, Ansgar, and Stumpf, Eike (2024). Analysis of cost-efficient urban air mobility systems: Optimization of operational and configurational fleet decisions. *European Journal of Operational Research* 317(3), pp. 678–695.

Jacquillat, Alex, Schmid, Alexandria, and Wang, Kai (2022). Relay logistics: A multi-variable generation approach. *Available at SSRN 4241031*.

Jonas, Adam (Jan. 2019). Are flying cars preparing for takeoff? URL: https://www.morganstanley.com/ideas/autonomous-aircraft.

Keskin, Merve and Çatay, Bülent (2016). Partial recharge strategies for the electric vehicle routing problem with time windows. *Transportation Research Part C: Emerging Technologies* 65, pp. 111–127.

Li, Baoxiang, Krushinsky, Dmitry, Reijers, Hajo A, and Van Woensel, Tom (2014). The share-a-ride problem: People and parcels sharing taxis. *European Journal of Operational Research* 238(1), pp. 31–40.

Li, Baoxiang, Krushinsky, Dmitry, Van Woensel, Tom, and Reijers, Hajo A (2016). The share-a-ride problem with stochastic travel times and stochastic delivery locations. *Transportation Research Part C: Emerging Technologies* 67, pp. 95–108.

Li, Shanghan, Zhang, Tengfei, Xiao, Yiyong, and Li, Daqing (2025). On-demand ridesharing based on dynamic scheduling in urban air mobility. *Transportation Research Part C: Emerging Technologies* 175, p. 105111.

Loh, Wei-Yin (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1(1), pp. 14–23.

Long, Qi, Ma, Jun, Jiang, Feifeng, and Webster, Christopher John (2023). Demand analysis in urban air mobility: A literature review. *Journal of Air Transport Management* 112, p. 102436.

Luo, Qi and Saigal, Romesh (2017). Dynamic pricing for on-demand ride-sharing: A continuous approach. *Available at SSRN 3056498.*

Mao, Chao, Liu, Yulin, and Shen, Zuo-Jun Max (2020). Dispatch of autonomous vehicles for taxi services: A deep reinforcement learning approach. *Transportation Research Part C: Emerging Technologies* 115, p. 102626.

Meinhardt, Simon, Schlenther, Tilmann, Martins-Turner, Kai, and Maciejewski, Michal (2022). Simulation of on-demand vehicles that serve both person and freight transport. *Procedia Computer Science* 201, pp. 398–405.

Mnih, Volodymyr et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518(7540), pp. 529–533.

Mourad, Abood, Puchinger, Jakob, and Chu, Chengbin (2019). A survey of models and algorithms for optimizing shared mobility. *Transportation Research Part B: Methodological* 123, pp. 323–346.

Mousavi, Kianoush, Bodur, Merve, Cevik, Mucahit, and Roorda, Matthew J (2024). Approximate dynamic programming for pickup and delivery problem with crowdshipping. *Transportation Research Part B: Methodological* 187, p. 103027.

Müller, Christian, Gönsch, Jochen, Soppert, Matthias, and Steinhardt, Claudius (2023). Customer-centric dynamic pricing for free-floating vehicle sharing systems. *Transportation Science* 57(6), pp. 1406–1432.

Munari, Pedro and Alvarez, Aldair (2019). Aircraft routing for on-demand air transportation with service upgrade and maintenance events: Compact model and case study. *Journal of Air Transport Management* 75, pp. 75–84.

Nazari, Mohammadreza, Oroojlooy, Afshin, Snyder, Lawrence, and Takác, Martin (2018). Reinforcement learning for solving the vehicle routing problem. *Advances in Neural Information Processing Systems* 31.

Owais, Mahmoud, Ahmed, Abdou S, Moussa, Ghada S, and Khalil, Ahmed A (2021). Design scheme of multiple-subway lines for minimizing passengers transfers in mega-cities transit networks. *International Journal of Rail Transportation* 9(6), pp. 540–563.

Pons-Prats, Jordi, Živojinović, Tanja, and Kuljanin, Jovana (2022). On the understanding of the current status of urban air mobility development and its future prospects: Commuting in a flying vehicle as a new paradigm. *Transportation Research Part E: Logistics and Transportation Review* 166, p. 102868.

Powell, Warren B (2007). Approximate Dynamic Programming: Solving the curses of dimensionality. John Wiley & Sons.

Powell, Warren B. (2022). Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions. John Wiley & Sons,Hoboken, NJ.

Qiu, Han, Li, Ruimin, and Zhao, Jinhua (2018). Dynamic pricing in shared mobility on demand service. *arXiv preprint arXiv:1802.03559*.

Rajendran, Suchithra and Srinivas, Sharan (2020). Air taxi service for urban mobility: A critical review of recent developments, future challenges, and opportunities. *Transportation Research Part E: Logistics and Transportation Review* 143, p. 102090.

Righini, Giovanni and Salani, Matteo (2006). Symmetry helps: Bounded bi-directional dynamic programming for the elementary shortest path problem with resource constraints. *Discrete Optimization* 3(3), pp. 255–273.

Sawers, Paul (2020). Uber Connect lets you deliver things to friends and family. Accessed Jan. 3, 2025, https://venturebeat.com/business/uber-connect-lets-you-deliver-things-to-friends-and-family.

Sayarshad, Hamid R and Chow, Joseph YJ (2015). A scalable non-myopic dynamic dial-a-ride and pricing problem. *Transportation Research Part B: Methodological* 81, pp. 539–554.

Schlenther, Tilmann, Martins-Turner, Kai, Bischoff, Joschka Felix, and Nagel, Kai (2020). Potential of private autonomous vehicles for parcel delivery. *Transportation Research Record* 2674(11), pp. 520–531.

Schmid, Verena (2012). Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. *European Journal of Operational Research* 219(3), pp. 611–621.

Shihab, Syed Arbab Mohd, Wei, Peng, Shi, Jie, and Yu, Nanpeng (2020). "Optimal evtol fleet dispatch for urban air mobility and power grid services". *AIAA Aviation 2020 Forum*, p. 2906.

Singh, Ashutosh, Al-Abbasi, Abubakr O, and Aggarwal, Vaneet (2021). A distributed model-free algorithm for multi-hop ride-sharing using deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems* 23(7), pp. 8595–8605.

Soeffker, Ninja, Ulmer, Marlin W, and Mattfeld, Dirk C (2022). Stochastic dynamic vehicle routing in the light of prescriptive analytics: A review. *European Journal of Operational Research* 298(3), pp. 801–820.

Srour, F Jordan, Agatz, Niels, and Oppen, Johan (2018). Strategies for handling temporal uncertainty in pickup and delivery problems with time windows. *Transportation Science* 52(1), pp. 3–19.

Steever, Zachary, Karwan, Mark, and Murray, Chase (2019). Dynamic courier routing for a food delivery service. *Computers & Operations Research* 107, pp. 173–188.

Strauss, Arne, Gülpınar, Nalan, and Zheng, Yijun (2021). Dynamic pricing of flexible time slots for attended home delivery. *European Journal of Operational Research* 294(3), pp. 1022–1041.

Sun, Xiaoqian, Wang, Shuang, Zhang, Xuejun, and Wandelt, Sebastian (2025). LAERACE: Taking the policy fast-track towards low-altitude economy. *Journal of the Air Transport Research Society*, p. 100058.

Sungur, Ilgaz, Ren, Yingtao, Ordóñez, Fernando, Dessouky, Maged, and Zhong, Hongsheng (2010). A model and algorithm for the courier delivery problem with uncertainty. *Transportation Science* 44(2), pp. 193–205.

Teubner, Timm and Flath, Christoph M (2015). The economics of multi-hop ride sharing: Creating new mobility networks through IS. *Business & Information Systems Engineering* 57(5), pp. 311–324.

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

Turan, Berkay, Pedarsani, Ramtin, and Alizadeh, Mahnoosh (2020). Dynamic pricing and fleet management for electric autonomous mobility on demand systems. *Transportation Research Part C: Emerging Technologies* 121, p. 102829.

Uber (2020). Uber Connect - easier than ever to send packages to others. Accessed Jan. 3, 2025, https://www.uber.com/blog/uber-connect-ondemand-package-delivery/.

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

Ulmer, Marlin W, Brinkmann, Jan, and Mattfeld, Dirk C (2015). “Anticipatory planning for courier, express and parcel services”. *Logistics Management: Products, Actors, Technology-Proceedings of the German Academic Association for Business Research, Bremen, 2013*. Springer, pp. 313–324.

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

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

Ulmer, Marlin Wolf (2017). Approximate dynamic programming for dynamic vehicle routing. Springer International Publishing, Cham, Switzerland.

Vinsensius, Albert, Wang, Yuan, Chew, Ek Peng, and Lee, Loo Hay (2020). Dynamic incentive mechanism for delivery slot management in e-commerce attended home delivery. *Transportation Science* 54(3), pp. 567–587.

Voccia, Stacy A, Campbell, Ann Melissa, and Thomas, Barrett W (2019). The same-day delivery problem for online purchases. *Transportation Science* 53(1), pp. 167–184.

Wandelt, Sebastian, Wang, Shuang, Zheng, Changhong, and Sun, Xiaoqian (2023). AERIAL: A meta review and discussion of challenges toward unmanned aerial vehicle operations in logistics, mobility, and monitoring. *IEEE Transactions on Intelligent Transportation Systems* 25(7), pp. 6276–6289.

Wang, Dujuan, Wang, Qi, Yin, Yunqiang, and Cheng, TCE (2023). Optimization of ride-sharing with passenger transfer via deep reinforcement learning. *Transportation Research Part E: Logistics and Transportation Review* 172, p. 103080.

Wang, Kai, Jacquillat, Alexandre, and Vaze, Vikrant (2022). Vertiport planning for urban aerial mobility: An adaptive discretization approach. *Manufacturing & Service Operations Management* 24(6), pp. 3215–3235.

Wu, Zhiqiang and Zhang, Yu (2020). “Optimal eVTOL charging and passenger serving scheduling for on-demand urban air mobility”. *AIAA Aviation 2020 Forum*, p. 3253.

Yan, Yiping, Wang, Kai, and Qu, Xiaobo (2024). Urban air mobility (UAM) and ground transportation integration: A survey. *Frontiers of Engineering Management*, pp. 1–25.

Zhou, Qiang and Tan, Fu (2024). Avionics of Electric Vertical Take-off and Landing in the Urban Air Mobility: A Review. *IEEE Aerospace and Electronic Systems Magazine*.