UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Engineering and Information Technology

Optimal Task Scheduling and Flight Planning for Multi-Task Unmanned Aerial Vehicles

by

Bin Liu

A Thesis Submitted
in Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Sydney, Australia

2022

Certificate of Authorship/Originality

I, Bin Liu declare that this thesis, is submitted in fulfilment of the requirements for the award of doctor of philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with Nanjing University of Posts and Telecommunications.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 22/02/2022

ABSTRACT

Optimal Task Scheduling and Flight Planning for Multi-Task Unmanned Aerial Vehicles

by

Bin Liu

Unmanned aerial vehicles (UAVs), also known as drones, play an important role in various areas due to their agility and versatility. Integrated with many embedded components, the UAV is capable of conducting multiple tasks simultaneously. Coordinating different tasks to a multi-task UAV can be challenging. The reason is that tasks may require different levels of commitment and tolerate different latencies. Another reason is that multi-tasking can give rise to difficulties in the UAV's energy management, as many UAVs are battery-powered. In this thesis, we study the optimal flight planning, control, and routing for the multi-task UAV.

The main contributions of this thesis can be summarized as follows.

- This thesis presents a novel energy-efficient UAV flight planning framework, which integrates UAVs into intelligent transportation systems for energy-efficient, delay-sensitive delivery services. The UAV can dynamically choose actions from cruise speed, full speed, recharging at a roadside charging station, or hitchhiking and recharging on a collaborative vehicle. The objective is to minimize the energy consumption of the UAV and ensure timely delivery. We reveal the conditions under which the UAV's flight planning changes in terms of the remaining flight distance or the elapsed time. Consequently, the optimal flight planning can be instantly made by comparing with the thresholds.

- This thesis presents a new online control framework for multi-task UAVs, which allows a UAV to perform in-situ sensing while delivering goods. A new

finite-horizon Markov decision process (FH-MDP) problem is formulated to ensure timely delivery, minimize the UAV's energy consumption, and maximize its reward for in-situ sensing. We prove the monotonicity and subadditivity of the FH-MDP, such that the FH-MDP has an optimal, monotone deterministic Markovian policy. We find that the optimal policy consists of flight distance-related and time-related thresholds at which the optimal action of the UAV switches. As a result, the optimal actions of the UAV can be obtained by comparing its state with the thresholds at a linear complexity.

- This thesis presents a novel multi-task UAV routing framework, which aims to minimize the UAV's energy consumption, maximize its sensing reward, and ensure its timely arrival at the destination. We interpret possible flight waypoints as location-dependent tasks, hence accommodating the waypoints and in-situ sensing in a unified process of task selection. We construct a weighted time-task graph, and transform the optimal routing of the UAV to a weighted routing problem, which can be optimally solved by the celebrated Bellman-Ford algorithm.

Acknowledgements

I wish to thank all the people whose assistance was a milestone in the completion of this thesis.

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Ren Ping Liu and Prof. Hongbo Zhu, for all the opportunities, patience and help throughout my study. Their meticulous academic attitudes much impress me. I am grateful for their advice and guidance to grow. I would also like to convey my sincere thanks to my co-supervisor, Prof. Y. Jay Guo, for his invaluable advice and immense knowledge. It's really joyful to receive his appreciation and support.

I would like to express my profound appreciation to my co-supervisor, Prof. Wei Ni, who has the substance of a genius: he convincingly guided and encouraged me to be professional and do the right thing even when the road got tough. Without his persistent help, the completion of this thesis would have been impossible.

Finally, I would like to express my deep and sincere gratitude to my parents for their encouragement and unconditional love. They support me to pursue research and being myself. This work would not have been possible without their input.

Bin Liu

Sydney, Australia, 2022.

List of Publications

Journal Papers

J-1. **B. Liu**, W. Ni, R. Liu, Q. Zhu, Y. J. Guo, and H. Zhu. "Novel Integrated Framework of Unmanned Aerial Vehicle and Road Traffic for Energy-Efficient Delay-Sensitive Delivery," *IEEE Transactions on Intelligent Transportation Systems*, Early Access 2021. (Chapter 3).

J-2. **B. Liu**, W. Ni, R. Liu and H. Zhu. "Optimal Selection of Heterogeneous Network Interfaces for High-Speed Rail Communications," *IEEE Transaction on Vehicular Technology*, vol. 69, no.12, pp. 15005-15018. 2020.

J-3. **B. Liu**, W. Ni, R. Liu, Y. J. Guo, and H. Zhu. "Optimal Control of Multi-Task Drone for Delay-Aware Goods Delivery and In-Situ Sensing," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Submitted 2021. (Chapter 4).

J-4. **B. Liu**, W. Ni, R. Liu, Y. J. Guo, and H. Zhu. "Optimal Routing of Unmanned Aerial Vehicle for Joint Goods Delivery and In-situ Sensing," *IEEE Transactions on Intelligent Transportation Systems*, under major revision (revised and resubmitted), 2021. (Chapter 5).

J-5. **B. Liu**, Q. Zhu, and H. Zhu. "Trajectory optimization and resource allocation for UAV-assisted relaying communications," *Wireless Networks*, vol. 26, no. 1, pp. 1–11, Nov. 739-749, 2019.

J-6. **B. Liu**, Q. Zhu, and H. Zhu. "Rotman lens-based two-tier hybrid beamforming for wideband mmWave MIMO-OFDM system with beam squint," *EURASIP Journal on Wireless Communications and Networking*, vol.2018, no.1 pp.267, 2018.

Conference Papers

C-1. **B. Liu**, W. Tan, H. Hu, and H. Zhu. "Hybrid beamforming for mmWave MIMO-OFDM system with beam squint," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy, Sept. 2018, pp. 1422–1426.

C-2. **B. Liu**, Q. Zhu, and H. Zhu. "Delay-Aware LTE WLAN Aggregation for 5G Unlicensed Spectrum Usage," *IEEE 85rd Vehicular Technology Conference (VTC Spring)*, Sydney, Australia, Jun. 2017, pp. 1-6.

Contents

List of Figures

Abbreviation

3GPP: 3rd Generation Partnership Project

BS: Base Station

CAV: Connected and Automated Vehicle

DP: Dynamic Programming

FAA: Federal Aviation Administration

ILP: Integer Linear Programming

IoT: Internet of Things

ITS: Intelligent Transportation Systems

LoRaWAN: Long Range Wide Area Network

LoS: Line-of-Sight

MDP: Markov decision process

POI: Places of Interest

SWIPT: Simultaneous Wireless Information and Power Transfer

UAV: Unmanned Aerial Vehicle

VRP: Vehicle Routing Problem

WPT: Wireless Power Transfer

WSN: Wireless Sensor Network

Chapter 1

Introduction

1.1 Background

Unmanned aerial vehicles (UAVs), also known as drones, are playing an important role in various areas due to their agility and versatility [1–3]. Featured by on-board data processing, autonomy and boundless mobility, the UAV is empowered to provide close integration of the cyber and physical worlds by a synergy of computational power and physical processes [4, 5]. It can be used for various applications in the smart city, such as environmental sensing [6, 7], weather monitoring [8], cargo delivery [9], emergency search and rescue [10, 11], etc.

UAV optimization has been recently widely actively as an important use case in 5G advanced working groups, such as 3GPP Technical Specification Group Service and System Aspects (SA) 6, ITU. For example, in 3GPP release-17 [12], the UAV optimization is specified as the new usage area. With UAV optimization, timely delivery can be guaranteed, which can provide a delay-aware delivery service. As shown in this case in Fig. 1.1, the delivery approach can be customized with the delivery time requirement. In the paradigm of smart cities, UAVs can unify the physical and cyber world via ubiquitous and heterogeneous smart devices through 3D connectivity [13].

Equipped with embedded sensors, such as cameras, microphones, and thermometers, a drone is often capable of conducting multiple different tasks [14, 15]. As illustrated in Fig. 1.2, a multi-task drone equipped with sensors and dispatched to deliver goods can also carry out in-situ sensing along its delivery route.

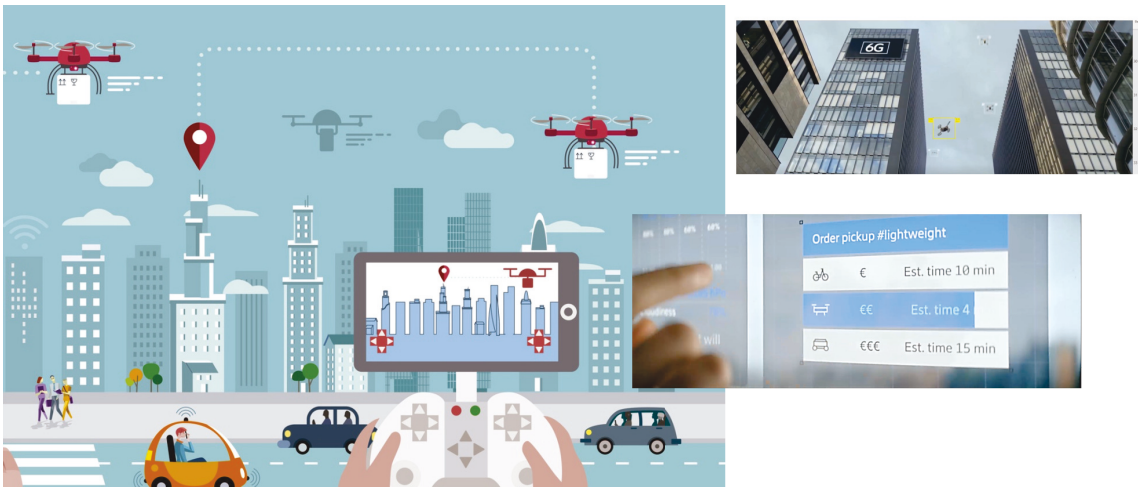


Figure 1.1 : Delay-aware delivery service: a timely aerial delivery can be guaranteed by the UAV optimization.

1.2 Research Challenges

Coordinating different tasks (e.g., goods delivery and in-situ sensing) to a multi-task UAV can be challenging. The reason is that roadside tasks may require different levels of commitment and tolerate different latencies. Another reason is that multi-tasking can give rise to difficulties in the energy management of a UAV, as many UAVs are powered by batteries. It is important to utilize the battery energy efficiently while completing multiple tasks by their deadlines. The key challenges for coordinating different tasks are listed as follows.

- **Challenge 1:** Battery limitation and energy management. Though the UAVs have great flexibility, the problem is that many UAVs are powered by batteries, limiting delivery range. Moreover, multi-tasking can give rise to difficulties in the energy management of the UAVs.

- **Challenge 2:** Muti-Tasking management. Different tasks may require different levels of commitment and tolerate different latencies. For example, the goods delivery is typically subject to strict deadlines, while in-situ sensing is

Figure 1.2 : Multi-task UAV for joint goods delivery and in-situ sensing.

a secondary task and only performed when possible.

- **Challenge 3:** Time management. The UAV needs to slow down and to perform tasks, which may prevent the timely arrival of the UAV.

- **Challenge 4:** Time-sensitive and location-dependent task request. The sensing task is often time-sensitive and location-dependent, as sensing platforms usually interest in information at a specified location and time. The UAV needs to arrive at the the place of interest within the requested time.

1.3 Research Objectives

In this thesis, we study the flight planning, optimal control, and optimal routing for the multi-task UAV in smart city with consideration of diverse service requirements. The aims of this thesis are to:

- **Objective 1:** establish a novel energy-efficient UAV flight planning frame-

Figure 1.3 : Research Objectives

work, which aims to minimize the energy consumption of the UAV and ensure timely arrival of the UAV.

- **Objective 2:** establish a novel multi-task UAV control framework, which aims to maximize its reward for sensing task conducted, minimize the drone's energy consumption, and ensure the ensuring timely goods delivery.

- **Objective 3:** establish a novel multi-task UAV routing framework, which aims to minimize the UAV's energy consumption, maximize its sensing reward, and ensure its timely arrival at the destination.

1.4 Research Contributions

This subsection provides a brief summary of the contribution this thesis.

1.4.1 Energy-Efficient Flight Planning of Multi-Task UAV

Chapter 3 presents a novel closely integrated framework of UAVs and road vehicles for energy-efficient delay-sensitive goods delivery, where a UAV can hitchhike and get recharged on willing vehicles. The UAV also has access to roadside charging stations. The different speeds of a UAV can be selected for energy saving [16, 17]. We optimize the UAV's selection of actions to minimize the energy consumption of the UAV over the flight distance while ensuring its timely arrival at the destination within a specified delay. The action selection is formulated as a finite-horizon Markov decision process (MDP) problem. A new dynamical programming (DP)-based algorithm is developed to obtain the optimal action policy. Another important aspect is that we analyze the conditions under which the optimal selections can be made. By exploiting the conditions of the optimal policy, we prove that the optimal policy is monotone to the remaining flight distance and the elapsed travel time. The optimal policy only changes when the remaining flight distance or the elapsed time meets the conditions. As a result, the optimal actions of the UAV can be instantly made by comparing the distance and time against the thresholds derived from the conditions, thereby eliminating the need of DP.

The key contributions of the chapter are as follows:

- We establish a novel UAV delivery framework, which closely integrates UAVs into the ITS to provide energy-efficient, delay-sensitive goods delivery services. The UAV can dynamically choose actions from cruise speed, full speed, recharging at a roadside charging station, or hitchhiking and recharging on a collaborative vehicle. The objective is to minimize the energy consumption of the UAV and ensure timely delivery.

- We formulate the action selection problem of the UAV as a finite-horizon discrete-time MDP problem, and solve the problem by developing a new DP

and backward induction-based algorithm.

- We unveil the necessary conditions under which the action of the UAV changes in the optimal policy. Accordingly, we derive the thresholds for the remaining flight distance and the elapsed travel time, at which an optimal action changes. As a result, the optimal actions can be selected instantly by thresholding.

1.4.2 Optimal Routing of Multi-Task UAV

Chapter 4 presents a new online control framework for multi-task UAVs, which allows a UAV to perform in-situ sensing while delivering goods. A new finite-horizon Markov decision process (FH-MDP) problem is formulated to ensure the timely delivery of goods, minimize the UAV's energy consumption, and maximize its reward for in-situ sensing. The key idea is that we prove the monotonicity and subadditivity of the FH-MDP, and hence the FH-MDP has the optimal, monotone deterministic Markovian policy. Another important finding is that the optimal policy consists of flight distance- and time-related thresholds at which the optimal action of the UAV switches. As a result, the multi-task UAV can select its optimal actions (i.e., flight speed, or in-situ sensing) online by comparing its state with the thresholds at a linear complexity. The optimality of the proposed approach is validated numerically by comparing with its computationally expensive, dynamic programming (DP)-based alternative. The main contributions of this chapter are listed in the following.

- We establish the new multi-task UAV control framework, where the UAV can dynamically select its action at each slot, including full speed, cruise speed, or performing in-situ sensing. A new FH-MDP problem is formulated to maximize its reward for sensing, minimize the UAV's energy consumption, and ensure the timely goods delivery.

- By proving the monotonicity and subadditivity of the FH-MDP, we prove the

existence of the monotone deterministic Markovian policy which is the optimal policy for the FH-MDP problem.

- We unveil the new threshold-based structure of the optimal policy, and derive the thresholds at which the action of the UAV optimally switches. The UAV can select its optimal action instantly at each time slot by comparing its current state with the thresholds at a linear complexity.

Extensive simulations corroborate that the new approach is able to guarantee the UAV's timely delivery, minimize its energy consumption, and maximize its sensing reward, as validated by comparison with the computationally expensive, standard dynamic programming (DP)-based solver for the FH-MDP problem. The simulations also show that our algorithm can save 24% of the energy of the UAV under our considered simulation setting, as compared to its alternative methods.

1.4.3 Optimal Routing of Multi-Task UAV

Chapter 5 optimizes the route of a UAV for joint goods delivery and in-situ sensing, so as to ensure timely goods delivery, minimize the UAV's energy consumption, and maximize its reward for in-situ sensing conducted. The key idea is that we interpret possible flight waypoints as location-dependent tasks, hence accommodating the waypoints and in-situ sensing in a unified process of task selection. Another crucial aspect is that we construct a weighted time-task graph, where each vertex captures both the location and time of a task. The feasible routes of the UAV are interpreted as a series of feasible vertexes which do not violate the delivery deadline and speed limit in the graph. By distributing the rewards of sensing tasks to the edges of the graph, we transform the optimal routing of the UAV to a weighted routing problem in the time-task graph and solve the problem using the Bellman-Ford algorithm. As validated by simulations, the new approach can outperform its

potential alternatives by at least 18% in terms of task reward in a real-world case study. The main contributions of this chapter are listed in the following.

- We establish a new UAV routing framework for joint goods delivery and in-situ sensing. A new routing problem is formulated to jointly ensure timely goods delivery, minimize the UAV's energy consumption, and maximize its reward for in-situ sensing conducted.
- We convert this joint aerial sensing and delivery routing problem to a task-time routing problem, where the task-time route refers to a series of differently located tasks selected over time. By interpreting the UAV's flight as virtual tasks, we reveal that the task-time route of the UAV has a unified graph representation.
- We derive the feasible task-time route conditions, under which a timely arrival of the UAV at the destination is guaranteed. By distributing the rewards of sensing tasks to the edges of the graph, we finally transform the routing of the UAV to a weighted routing problem in the time-task graph and optimally solve the problem using the celebrated Bellman-Ford algorithm.

1.5 Thesis Organization

This thesis focuses on flight planning, optimal control, and optimal routing for the multi-task UAV with consideration of diverse service requirements. This thesis is organized as follows:

Chapter 1: This chapter gives an overview of this research. It presents the research background of this thesis, followed by the main motivations and key research contributions.

Chapter 2: This chapter presents a comprehensive literature review of the related work on energy management, task scheduling, flight planning of the UAV.

Chapter 3: This chapter thesis presents a novel energy-efficient UAV flight planning framework, which integrates UAVs into intelligent transportation systems for energy-efficient, delay-sensitive goods delivery. Dynamic programming is first applied to minimize the energy consumption of a UAV and ensure its timely arrival at its destination, by optimizing the flight planning of the UAV.

Chapter 4: This chapter thesis present presents a new online control framework for multi-task UAVs, which allows a UAV to perform in-situ sensing while delivering goods. A new finite-horizon Markov decision process (FH-MDP) problem is formulated to ensure the timely delivery of goods, minimize the UAV's energy consumption, and maximize its reward for in-situ sensing.

Chapter 5: This chapter presents a novel multi-task UAV routing framework, which aims to minimize the UAV's energy consumption, maximize its sensing reward, and ensure its timely arrival at the destination. The optimal routing of the UAV is transformed to a weighted routing problem in the time-task graph and solved by the Bellman-Ford algorithm.

Chapter 6: This chapter summarizes the research of this thesis, and discusses future works.

Chapter 2

Literature Survey

Unmanned aerial vehicles play an important role in various areas due to their agility and versatility, which have attracted significant attention from academia and industry in recent years. The UAV can emerge as an aerial platform that integrates sensors, devices, networks, and/or applications. With the onboard communication, computation, sensing and control units, the UAV can build a closed loop from data perceiving, information exchanging, decision making to the final execution, thus providing a close integration of the cyber and physical worlds.

The rest of this chapter will review the related work on energy management, task scheduling, flight planning of the UAV and UAV-assisted sensing and relay communication.

2.1 Energy Management of the UAV

Many UAVs are powered by batteries, limiting delivery range. The delivery range can extended delivery range with optimal battery management [18] and trajectory optimization [19], but the delivery over a large distance is still challenging. Deploying charging stations along roads is an option to overcome the battery limitation [20–22]. The position and density of charging stations are critical to the delivery range. The minimum number of charging stations that satisfy a coverage requirement was presented in [23]. Nonetheless, it would take non-trivial effort, investment, and delay to deploy the charging stations densely.

Another existing option is to allow UAVs to collaborate with ground vehicles [24–

26], where a truck carries a set of packages and one or several UAVs. The UAVs can launch from the truck, deliver the packages, and return and dock onto the truck after completing the deliveries. A joint optimization and planning framework of ground truck and UAV goods delivery was proposed in [24, 25]. The framework decides the use of ground trucks, drones, or combined to serve customers through multi-stage stochastic integer programming. However, the uncertainty of customers' demand and travel time was not considered in the framework. Das *et al.* [26] formulated synchronized truck and drone routing as a vehicle routing problem (VRP), and applied the genetic algorithm to minimize the travel cost. The model is limited to a fixed pair of trucks and drones. Flexibilities were explored by using public passenger vehicles, such as buses and trams, instead of trucks [27, 28]. Huang *et al.* [27] presented drone routing problem under a stochastic public transportation model. A label setting algorithm was developed to solve a drone path problem to minimize travel time in [28].

While extending UAV delivery coverage, the collaboration of UAVs and trucks could incur high running costs because drivers are still needed. The maintenance, labor, and fuel costs of trucks are relatively high. With increasingly maturing automation technologies, connected and automated vehicles (CAVs) and shared mobility are revolutionizing intelligent transportation systems (ITS) [29–31]. Delivery logistics benefit from systems of shared vehicles [32]. In particular, UAVs can hitchhike on vehicles driving in the same direction to save energy. With docking mechanisms and the power charging connectors on the roof of vehicles [33, 34], UAVs can latch onto the vehicles and get recharged while hitchhiking.

2.2 Task Scheduling of the UAV

UAVs enjoy excellent flexibility and have been widely applied to provide environmental sensing [6, 35], weather monitoring [8, 36], goods delivery [9], and disaster

rescue [10,11,37]. Existing studies have been predominantly focused on the scheduling of single-task drones . In [38], the navigation of a drone was formulated as a generalized traveling salesman problem (TSP) for persistent monitoring missions. By determining the order of the targets to be visited and monitored, the longest interval elapsed between consecutive visits was minimized to optimize the route of the drone. Based on a vector field approach, a curved path-following method was proposed in [39] to control a drone conducting video tracking and surveillance. By exploiting the notion of input-to-state stability, the along-track error and the cross-track error were reduced for robust surveillance and tracking. Bartolini *et al.* [40] investigated the multi-trip task assignment of squads of drones for early target inspection. To maximize the target coverage, an NP-hard Integer Linear Programming (ILP) was formulated to schedule multiple drones. A greedy and prune (GaP) algorithm was developed to obtain 1/2-approximation of the optimality in polynomial time. Scheduling a truck-drone delivery system was studied in [41] and [25], where a truck carries parcels and acts as a moving dock for drones. The scheduling policy developed in [41] and [25] ensured a drone to take off from the truck, return to the truck after finishing the deliveries. The drone schedule problem was an MINLP problem [42].

While a decomposition approach was adopted, a high computational complexity was still undergone [26]. A distributed schedule framework was proposed for drone navigation and roadside charging in [43], and multi-agent deep reinforcement learning was used to minimize the energy consumption of the drone. Zhang *et al.* [44] considered a drone powered by both solar energy and roadside charging stations. By optimizing the drone's trajectory and schedule policy, sustainable communication services were maintained with energy outage prevention. However, the above existing works cannot be extended to schedule a multi-task drone running both goods delivery and in-situ sensing. Moreover, the existing works are computationally ex-

pensive despite their focus on single-task drones, due to the MINLP nature of the works.

2.3 Flight Planning of the UAV

Flight planning for joint goods delivery and in-situ sensing has yet to be studied in the literature. Existing studies have been on either delivery route planning or task selection. The UAV delivery routing problem was characterized as finding the optimal route for the UAV to deliver goods to the destination [21]. In [9], the navigation of a delivery drone was formulated as a generalized vehicle routing problem (VRP), which was a mixed-integer linear programming (MILP). The simulated annealing algorithm was applied to find the optimal route that minimizes both the delivery time and energy consumption. Francesco *et al.* [45] considered the drone delivery problem under varying wind conditions. The routing was carried out by the shortest path algorithm to maximize the percentage of delivery completion under a given battery budget. A synchronized truck and drone routing framework was developed in [24] and [26]. By jointly planning the routes of a ground truck and UAV for goods delivery, timely deliveries were guaranteed with the minimum travel costs.

In a different yet relevant context, UAVs were employed to collect data from ground sensors, where the velocity and data collection schedule of a UAV was optimized to minimize the packet loss of the sensors in [46] and [47]. The trajectory planning and ground node selection of a UAV was jointly optimized in [48] to minimize network data loss with incomplete network states knowledge. The trajectory was optimized in [49] to improve the throughput of cell-edge users. The UAV-enabled Internet of Things (IoT) networks were studied in [50], where the velocity control of UAV and the transmission scheduling of IoT nodes was jointly designed to minimize the data packet losses. However, these existing results [9, 21, 24, 26, 45–53]

cannot apply to the routing problem for multi-task UAVs, such as used for the joint goods delivery and in-situ sensing, because of the stringent delay requirement of goods delivery, as well as the time sensitivity and location dependence of in-situ sensing.

2.4 UAV-assisted In-Situ Sensing and Relay Communication

Wireless sensor networks have been largely deployed for various sensory solutions: agricultural and environmental monitoring or smart cities application, enabled by an enormous amount of sensor nodes (SNs) [54–60]. The sensory nodes typically send the collected data to the sink node. Constrained by the distance and channel character, the direct sink node transmission approach is nevertheless energy-consuming. Powered by a battery, these sensor nodes are challenging to be recharged [61, 62]. Hence, it is of great importance to design the energy-efficient data collection approach for prolonging the lifetime of WSNs [63,64]. The mobile sink was considered in [65] to enhance the network capacity in the WSNs.

Utilizing the UAV for WSNs has been deemed as a promising solution of energy-efficient data collection [66–68]. Featured by flexible mobility, the UAV can move sufficiently close to the SNs. The line-of-sight (communication links between the UAV and SNs can reduce transmission energy consumption. Different from conventional communication techniques, UAV's deployment and trajectory optimization issues are the new challenges for UAV-aided wireless communications. The performance is enhanced by the UAV placement optimization [69]. Many efforts have been devoted to the UAV's deployment optimization. Aiming at downlink coverage maximization, a three-dimensional (3D) UAV deployment optimization was considered in [70]. Moreover, 3D UAV deployment was also investigated in [71] to maximize the serving users. Apart from the placement optimization, the authors in [72] maximized throughput by power allocation together with the UAV's trajectory optimization.

The compressive sensing (CS) for WSN data gathering was studied in [73] to offer a transmission-efficient architecture. Using the UAV as LoRaWAN gateways, the energy-efficient surveillance scheme was investigated in [74] for the intelligent transportation systems.

As a promising solution for data collection and dissemination, UAV-aided WSNs has captured great attention [75]. With practical protocols and experiments, characterization of communication links of UAV-aided WSNs was given in [76, 77]. The multi-sensor uplink interference mitigation was proposed in [78] by leveraging the UAV beamforming. The SNs' wake-up policy and UAV's trajectory were jointly optimized in [63] to minimize the energy consumption of SNs. The UAV flight time for sensory data collection was minimized in [79]. The multi-agent reinforcement learning (MARL) framework was proposed in [80], where each UAV acts as the agent and automatically selects its communicating node, power levels and subchannels without any information exchange among UAVs. Leveraging proactive caching, reference [81] disseminated the popular data to a set of selected nodes that cooperatively cache all the files. Then, one SN retrieved the requested data from its local cache or from its nearest neighbour cached the file via device-to-device (D2D) communications, which could largely enhance the UAV endurance. The UAV was deployed in the WSN to gather environmental data for monitoring area [82]. The UAV also play an active role in optimizing the WSN topology. The authors in [83] further considered recharging out-of-battery ground devices with energy transfer from the UAV to extend the network lifetime. The system power consumption is minimized by deploying the UAV as a relay between the base station and sensors in [84].

It is worth noting that most of the existing works mentioned above focused on energy efficiency or throughput enhancement and only considered the UAV trajectory optimization [85], but overlooked the SNs transmission policy design. Most

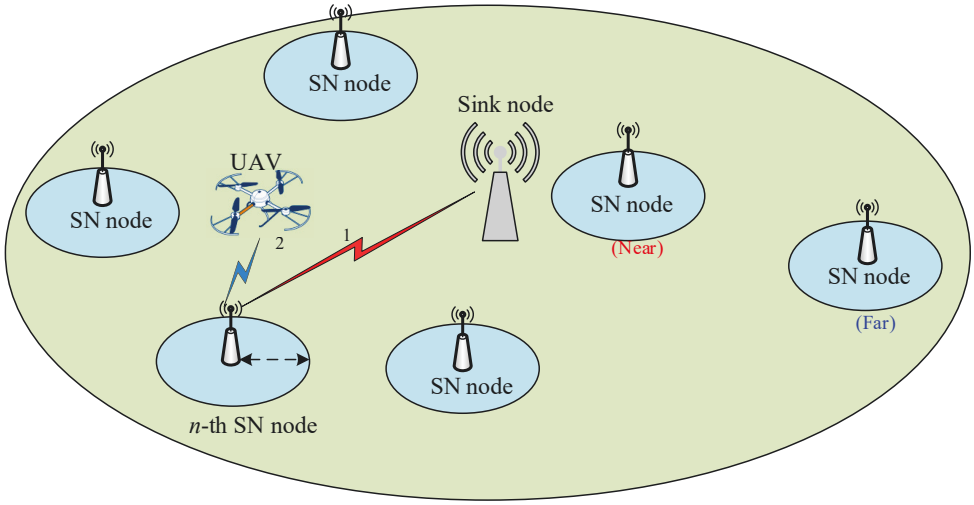


Figure 2.1 : UAV-assisted WSN

of the sensory data are delay-tolerant. This character helps to enhance further the UAV data collection performance in energy efficiency and throughput since the sensory data only needed to be uploaded to the data collector before the given time expires. Each SN needs to upload an amount of data to the data collector within a given time. At each slot, the SNs decide to upload data with optional transmission modes: i.e., 1) waiting, 2) directly transmitting data to the sink node, 3) uploading data to UAV transmission when available. On one side, the direct sink node transmission approach is always available. It costs more energy due to the distance and poor transmission link. On the other side, the UAV data gathering is an energy-effective transmission mode with the short-distance LoS link, which may not always be available. As the sensory data transmission is delay-tolerant, it is of the SN's interest to wait for the UAV collecting data, which save transmission energy of SNs. However, constrained by the buffer size and transmission time, the long wait leads to the unfinished transmission and incurs a buffer overflow and new sensory data loss.

A few works have aslo been seen for the UAV-assisted relaying communication

[86–91]. Two kinds of UAV relays are mainly investigated, i.e., static and mobile relays [92]. The static UAV relays are usually designed to enhance the wireless network's performance. By optimizing the static deployment, including altitude [93], horizontal placement [94], the static UAV relays harvest the line-of-sight link to improve the throughput or reliability of the network system. The optimum altitude of UAV to maximize the reliability, including power loss, outage probability, and bit error rate, was derived in [93]. It further clarified the optimum altitudes of UAV would be affected by the relay protocols, i.e., decode-and-forward (DF) and amplify-and-forward (AF). Fixed flying altitude, the horizontal locations of UAV were optimized in [94] to reap the higher throughput by deploying the UAV relay to the LoS position.

However, the static placement of UAVs cannot fully utilize the degree of freedom of UAV mobility and manoeuvrability features [72, 95–98]. The UAV trajectory optimization for data offloading in the edge area of multiple cells was proposed in [99] to enhance the sum rate of edge users. The UAV relays can achieve better performance to assist the wireless network by continuous position optimization. The throughput maximization with the consideration of the UAV mobility is studied in [72]. Constrained by the UAV flight duration, the trajectory optimization was investigated in [95] to minimize the transmission completion time. Aimed at rate-maximization and energy-minimization designs, the authors in [96] jointly optimized the initial/final locations, velocities, and minimum/maximum speed and acceleration of UAV. The hover duration of UAV together with trajectory was studied in [97, 100] to achieve better performance in terms of the average data rate. Forming the multiple UAVs as relays, both the single multi-hop link or multiple dual-hop links were evaluated in [101] under different path loss parameters and UAV positions. Trajectory optimization for multiple UAV relays was studied in [102]. Joint multiple UAVs trajectories and power allocation were optimized in [92] to obtain

the maximum end-to-end throughput. Moreover, the security issue in UAV-assisted relaying is critical [91, 103–107]. The significant throughput gain in UAV-assisted communication is due to its mobility and line-of-sight channel. However, on the other side, the LoS channels make UAV transmission easy to eavesdrop [108]. The UAV-assisted secure transmission for scalable videos in hyper-dense networks via caching was studied in [109].

However, most of these works do not consider Simultaneous Wireless Information and Power Transfer (SWIPT) to charge the UAV for data forwarding [110, 111, 111–115]. The source node sends data and energy instantaneously over different subcarriers (SCs) to the UAV relay, and then UAV relay forwards the data to the destination with the harvested energy. Several studies have integrated wireless power transfer (MPT) technologies into the UAV network. The authors in [116] deployed a UAV to deliver wireless energy to a set of receivers and maximize the amount of energy transferred by the UAV trajectory optimization. It was found that hovering location and duration of UAV can be leveraged to enhance the MPT efficiency. With UAV's trajectory prediction, the energy harvesting efficiency was improved in [117]. SWIPT was introduced into UAV-assisted relay [118], where file, power, and trajectory jointly optimized to maximize the throughput. However, most of these works did not fully consider the resource allocation problem for energy transfer, such as the power allocation and SC allocation. In the two-hop SWIPT OFDM relay systems, the SC allocation and power allocation are jointly optimized to maximize the capacity in [119, 120]. The joint subcarrier and power allocation for the full-duplex wireless powered OFDM communication networks was studied in [121]. The authors in [122] developed a dynamic SC allocation to improve the achievable rate, where the SCs are both used for information transmission and energy harvesting. However, these works failed to consider the dynamic position of relays or the transmission nodes in time.

2.5 Summary

Most of the current work considers deploying the UAV for a single kind of tasks, such as good delivery, traffic monitoring, and data gathering [123]. Facing the trend of digital transformation and Industrial Internet of Things (IIoT) , more vertical and horizontal integration is expected in the future [124–126]. The single-task optimization is not usually optimal in the system integration, for example, the energy consumption of UAV flight is typically much greater than the sensing and data collection [127]. It is not cost-effective to consume a large amount of propulsion energy to collect low-value, and delay-tolerant information [128].

With various onboard cameras, sensors and transmitters, UAVs can perform weather and air quality monitoring, scene monitoring and understanding, and various data-gathering tasks, such as meter reading. However, coordinate different tasks (e.g., goods delivery and in-situ sensing) to a multi-task UAV. The reason is that roadside tasks may require different levels of commitment and tolerate different latencies. Another reason is that multi-tasking can give rise to difficulties in the energy management of a UAV. Many UAVs are powered by batteries. It is important to utilize the battery energy efficiently while completing multiple tasks by their deadlines. This thesis studies flight planning, optimal control, and routing for the multi-task UAV in the smart city, considering diverse service requirements.

Chapter 3

Energy-Efficient Flight Planning for Multi-Task UAV

3.1 Introduction

Unmanned aerial vehicles have emerged as a promising platform for fast, energy- and cost-efficient package delivery [129,130]. More recently, UAVs have been an important tool during the COVID-19 pandemic, creating more resilient supply chains and contactless delivery services [131, 132]. The Federal Aviation Administration (FAA), for example, has issued numerous drone flight waivers to enable UAV-assisted pandemic response [133,134]. Amazon, Alphabet, and JD.com are rolling out UAV delivery for quicker and cheaper on-demand delivery services while avoiding unnecessary human contact [135].

Many UAVs are powered by batteries, limiting delivery range. The delivery range can extended delivery range with optimal battery management [18] and trajectory optimization [19], but the delivery over a large distance is still challenging. Deploying charging stations along roads is an option to overcome the battery limitation [20,21]. The position and density of charging stations are critical to the delivery range. The minimum number of charging stations that satisfy a coverage requirement was presented in [23]. Nonetheless, it would take non-trivial effort, investment, and delay to deploy the charging stations densely.

Another existing option is to allow UAVs to collaborate with ground vehicles [24–26], where a truck carries a set of packages and one or several UAVs. The UAVs can launch from the truck, deliver the packages, and return and dock onto the truck after

completing the deliveries. A joint optimization and planning framework of ground truck and UAV goods delivery was proposed in [24, 25]. The framework decides the use of ground trucks, drones, or combined to serve customers through multi-stage stochastic integer programming. However, the uncertainty of customers' demand and travel time was not considered in the framework. Das *et al.* [26] formulated synchronized truck and drone routing as a vehicle routing problem (VRP), and applied the genetic algorithm to minimize the travel cost. The model is limited to a fixed pair of trucks and drones. Flexibilities were explored by using public passenger vehicles, such as buses and trams, instead of trucks [27, 28]. Huang *et al.* [27] presented drone routing problem under a stochastic public transportation model. A label setting algorithm was developed to solve a drone path problem to minimize travel time in [28].

While extending UAV delivery coverage, the collaboration of UAVs and trucks could incur high running costs because drivers are still needed. The maintenance, labor, and fuel costs of trucks are relatively high. With increasingly maturing automation technologies, connected and automated vehicles (CAVs) and shared mobility are revolutionizing intelligent transportation systems (ITS) [29, 30]. Delivery logistics benefit from systems of shared vehicles. In particular, UAVs can hitchhike on vehicles driving in the same direction to save energy. With docking mechanisms and the power charging connectors on the roof of vehicles [33, 34], UAVs can latch onto the vehicles and get recharged while hitchhiking.

This chapter presents a novel closely integrated framework of UAVs and road vehicles for energy-efficient delay-sensitive goods delivery, where a UAV can hitchhike and get recharged on willing vehicles. The UAV also has access to roadside charging stations. The different speeds of a UAV can be selected for energy saving [16, 17]. We optimize the UAV's selection of actions to minimize the energy consumption of the UAV over the flight distance while ensuring its timely arrival

at the destination within a specified delay. The action selection is formulated as a finite-horizon Markov decision process (MDP) problem [136]. A new dynamical programming (DP)-based algorithm is developed to obtain the optimal action policy. Another important aspect is that we analyze the conditions under which the optimal selections can be made. By exploiting the conditions of the optimal policy, we prove that the optimal policy is monotone to the remaining flight distance and the elapsed travel time. The optimal policy only changes when the remaining flight distance or the elapsed time meets the conditions. As a result, the optimal actions of the UAV can be instantly made by comparing the distance and time against the thresholds derived from the conditions, thereby eliminating the need of DP.

The key contributions of the chapter are as follows:

- We establish a novel UAV delivery framework, which closely integrates UAVs into the ITS to provide energy-efficient, delay-sensitive goods delivery services. The UAV can dynamically choose actions from cruise speed, full speed, recharging at a roadside charging station, or hitchhiking and recharging on a collaborative vehicle. The objective is to minimize the energy consumption of the UAV and ensure timely delivery.

- We formulate the action selection problem of the UAV as a finite-horizon discrete-time MDP problem, and solve the problem by developing a new DP and backward induction-based algorithm.

- We unveil the necessary conditions under which the action of the UAV changes in the optimal policy. Accordingly, we derive the thresholds for the remaining flight distance and the elapsed travel time, at which an optimal action changes. As a result, the optimal actions can be selected instantly by thresholding.

The rest of this chapter is organized as follows. Section II describes the system

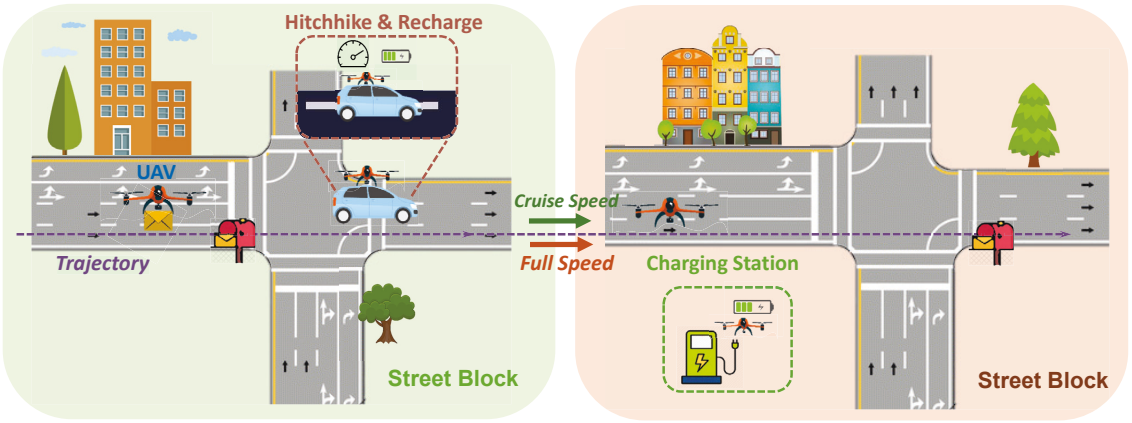


Figure 3.1 : A UAV flies along a route to deliver goods at different flight speeds. As shown in the left-hand side (LHS) of the figure, the UAV can hitchhike on collaborative vehicles and get recharged by the vehicle. As illustrated in the right-hand side (RHS) of the figure, the UAV can stop and recharge at roadside charging stations.

model. Section III formulates the problem of interest and solves the problem by using DP. In Section IV, we reveal that the optimal action selection can be translated to a thresholding process, and achieved with a substantially low, linear computational complexity. Simulation results are provided in Section V, followed by conclusions in Section VI.

3.2 System Model

Fig. 3.1 illustrates the proposed system, where a UAV is employed to collect goods at a number of locations and deliver the goods to the destination within a specified delay T. The UAV flies along roads to avoid buildings and foliage [137]. The pickup locations are also along the roads and accessible to the UAV. The UAV has access to roadside charging stations. The UAV flies along a route with the length of L (km) to visit the pickup locations sequentially, until it reaches the destination. In this chapter, we propose the UAV to collaborate with ground vehicles driving along

(some parts of) the flight path: The UAV can hitchhike on the vehicles driving in the same direction to save energy. With the docking mechanism and the power charging connector on the roof of the vehicles [33], the UAV can latch onto the vehicles and get recharged while hitchhiking.

A central coordinator (or coordinating platform) monitors road traffic and responds to the requests for goods delivery. The coordinator collects the routes of ground vehicles (and their willingness to collaborate with the UAV), e.g., via cellular systems or other wireless technologies. Accordingly, the coordinator plans the route and schedule of the UAV, so that the UAV can provide energy-efficient, delay-sensitive goods delivery services. The UAV and the collaborative vehicles are notified of the route and schedule, e.g., via cellular systems. This model has been considered in practice. For example, Mercedes-Benz has run a pilot project in Zurich to test an efficient van and drone-based cooperative system for on-demand delivery [138]. Amazon has planned to extend the service range of the Amazon Prime Air drone by integrating the autonomous robotaxi [139].

The objective is to minimize the energy consumption of the UAV, and ensure the UAV arrives at the destination within the time T. We discretize T into time slots with slot length τ, denoted by $t \in \mathcal{T} = \{0, \tau, 2\tau, \cdots, T/\tau\}$. We assume that the UAV can choose different speeds, or get recharged at a roadside charging station, or hitchhike on a collaborative vehicle and get recharged when hitchhiking. The action a specifies the UAV's decision at a decision epoch, and

1. *Cruise speed* (CS: $a = 0$): the UAV flies at the cruise speed (i.e., the eco mode), $v_0 = v_{\text{eco}}$;

2. *Full speed* (FS: $a = 1$): the UAV flies at its full speed, i.e., $v_1 = v_{\max}$;

3. *The UAV stops at a roadside charging station to get recharged* (CE: $a = 2$);

4. *The UAV hitchhikes on and gets recharged by a vehicle driving at the speed of* $v_3 = v_{\text{car}}$*, which depends on the speed limit of the road* (VE: $a = 3$).

The total flight length L is discretized with granularity υ, and collected in the set $l \in \mathcal{L} = \{0, \upsilon, 2\upsilon, \cdots, L/\upsilon\}$. We segment the UAV's route in accordance with street blocks. The indexes to the street blocks are $b \in \mathcal{B} = \{1, \cdots, B\}$, where B is the number of street blocks along the route of the UAV. L_b is the length of street block b. All street blocks are divided into four categories, according to their resource and service availability:

1. $\mathcal{B}^{(0)}$: the street blocks without service and resource available for the UAV;
2. $\mathcal{B}^{(1)}$: the street blocks with a roadside charging station;
3. $\mathcal{B}^{(2)}$: the street blocks with vehicles willing to provide hitchhiking and on-board charging services;
4. $\mathcal{B}^{(3)}$: the street blocks with both a roadside charging station and willing vehicles.

Note that a street block can be categorized differently at different time slots, depending on the presence of willing vehicles [140–142]. For example, a street block belonging to $\mathcal{B}^{(0)}$ at an earlier time slot can belong to $\mathcal{B}^{(2)}$ if a willing vehicle passes at a later time slot. By optimally selecting the UAV's action at each time slot, we minimize the energy consumption of the UAV over the flight distance L and ensure its timely arrival at the destination within the delay T.

3.3 Problem Formulation

In this section, we formulate the UAV's action selection problem, which minimizes the energy consumption of the UAV while ensuring its timely arrival at the

destination. The problem is shown to be a finite-horizon Markov decision process (FH-MDP) [136, 143]. We develop a new dynamical programming (DP)-based algorithm to obtain the optimal action policy.

Let $\boldsymbol{s} = (l, b)$ define the state of the UAV, where l is the remaining flight distance to the destination (in meters), and b indicates the street block which the UAV is currently in. Charging at a roadside station or hitchhiking on a vehicle can only take place in the street blocks with a charging station or willing vehicles. The different actions which can be possibly taken at different categories of street blocks are given by

$$\mathcal{A}^{(b)} = \begin{cases} \{0, 1\}, & \text{if } b \in \mathcal{B}^{(0)}; \\ \{0, 1, 2\}, & \text{if } b \in \mathcal{B}^{(1)}; \\ \{0, 1, 3\}, & \text{if } b \in \mathcal{B}^{(2)}; \\ \{0, 1, 2, 3\}, & \text{if } b \in \mathcal{B}^{(3)}. \end{cases} \tag{3.1}$$

In other words, actions $a = 0$ and $a = 1$ are always possible to the UAV, while the rest of the actions ($a = 2$ and 3) depend on the UAV's location.

For each action, the distance traveled per time slot (in meters) along the route is given by

$$Z(a, b) = \begin{cases} v_a \tau, & a \in \{0, 1\}, b \in \mathcal{B}; \\ v_0(\tau - \tau_c(b) - \tau_d(a, b)), & a = 2, b \in \{\mathcal{B}^{(1)}, \mathcal{B}^{(3)}\}; \\ v_{car}(\tau - \tau_d(a, b)) - v_0 \tau_d(a, b), & a = 3, b \in \{\mathcal{B}^{(2)}, \mathcal{B}^{(3)}\}. \end{cases} \tag{3.2}$$

where $\tau_c(b)$ is the charging time at a charging station. $\tau_d(a, b)$ is the time for the docking and undocking operations at a charging station ($a = 2$) in street block $b \in \{\mathcal{B}^{(1)}, \mathcal{B}^{(3)}\}$, or the latching and unlatching operations on a vehicle ($a = 3$) in street block $b \in \{\mathcal{B}^{(2)}, \mathcal{B}^{(3)}\}$. We assume that the UAV flies at the cruise speed ($v = v_0$) to get recharged or get onto a vehicle. Both the charging time $\tau_c(b)$ and the docking/latching time $\tau_d(a, b)$ can differ between the street blocks $b \in \mathcal{B}$. The total of the charging and docking/latching time is shorter than a time slot, i.e.,

$\tau_c(b) + \tau_d(a,b) \leq \tau$.

The energy consumption of the UAV depends on its speed and takeoff weight. The power consumption of the UAV (in Watts) is a function of its speed, i.e., $p(v)$, given the takeoff weight [144, 145]. The energy consumption (in Joules) caused by action a in street block b is given by

$$\begin{aligned} E_t(a, \boldsymbol{s}) &= E_t(a,b) \\ &= \begin{cases} p(v_a)\tau, & a \in \{0,1\}, b \in \mathcal{B}; \\ p(v_0)(\tau - \tau_c(a,b)), & a = 2, b \in \{\mathcal{B}^{(1)}, \mathcal{B}^{(3)}\}; \\ p(v_0)\tau_d(a,b), & a = 3, b \in \{\mathcal{B}^{(2)}, \mathcal{B}^{(3)}\}. \end{cases} \end{aligned} \tag{3.3}$$

When the UAV hitchhikes on a vehicle, the UAV does not consume energy, but the energy is consumed for latching/unlatching operations. When the UAV stops at a charging station, the UAV does not consume energy.

The UAV's power consumption depends on its speed and takeoff weight [19,146]. Given the takeoff weight, the cruise speed gives the maximum power efficiency [144, 147, 148]. The vehicle can accelerate or decelerate to allow the UAV to latch. The energy consumption of the UAV for acceleration and deceleration is assumed to be negligible.

The energy that the UAV gets charged (in Joules) depends on the charging efficiency η (in Watts) and the charging time:

$$G_t(a, \boldsymbol{s}) = G_t(a,b) = \begin{cases} \eta_a \tau_c(b) & a = 2, b \in \{\mathcal{B}^{(1)}, \mathcal{B}^{(3)}\}; \\ \eta_a(\tau - \tau_d(a,b)), & a = 3, b \in \{\mathcal{B}^{(2)}, \mathcal{B}^{(3)}\}; \\ 0 & \text{otherwise}, \end{cases} \tag{3.4}$$

where η_a denotes the charging efficiency of a charging station if $a = 2$, or the charging efficiency of a vehicle if $a = 3$.

The UAV needs to arrive at the destination within time T, and the penalty

occurs if it does not arrive by $T+\tau$. We define the penalty as

$$E_{T+\tau}(\boldsymbol{s}) = E_{T+\tau}(l) = \psi(l), \tag{3.5}$$

where $\psi(l)$ is a non-negative, non-decreasing convex function of l, and $\psi(0) = 0$. The total cost with the penalty considered can be written as

$$R_t(a, \boldsymbol{s}) = \begin{cases} E_t(a,b) - G_t(a,b), & t \in \mathcal{T}; \\ E_{T+\tau}(l), & t = T+\tau. \end{cases} \tag{3.6}$$

The transition probability of the state, i.e., the probability that the current state $\boldsymbol{s}$ transits to state $\boldsymbol{s}'$ at the next time slot if action a is taken, is expressed as

$$\mathbb{P}(\boldsymbol{s}'|\boldsymbol{s}, a) = \mathbb{P}\left((l', b')|(l, b), a\right) = \mathbb{P}\left(b'|b\right)\mathbb{P}\left(l'|(l, b), a\right), \tag{3.7}$$

where the conditional transition probability of the distance l becoming l' on the condition that action a is taken in street block b is given by

$$\mathbb{P}\left(l'|(l, b), a\right) = \begin{cases} 1, \text{ if } l' = [l - Z(a,b)]^+; \\ 0, \text{ otherwise}, \end{cases} \tag{3.8}$$

and $[x]^+ = \max\{0, x\}$. $\mathbb{P}(b'|b)$ is the conditional probability of the UAV flying from street block b to b', as given by

$$\mathbb{P}\left(b'|b\right) = \begin{cases} 1, \text{ if } Z(a,b) \geq L_b; \\ 0, \text{ otherwise}, \end{cases} \tag{3.9}$$

where the UAV flies from block street b to b', if the flight distance exceeds the length of the former street block, L_b, during the time slot. With the objective to minimize the energy consumption of the UAV and ensure its timely arrival at the destination, the problem of interest is formulated as

$$\begin{aligned} \underset{\boldsymbol{\pi}\in\Pi}{\text{minimize}} \quad & \mathbb{E}^{\boldsymbol{\pi}}_{\boldsymbol{s}_1}\left[\sum_{t=1}^{T} R_t\left(\boldsymbol{s}^{\boldsymbol{\pi}}_t, \pi_t\left(\boldsymbol{s}^{\boldsymbol{\pi}}_t\right)\right) + R_{T+1}\left(\boldsymbol{s}^{\boldsymbol{\pi}}_{T+1}\right)\right] \\ \text{s.t.} \quad & \pi_t \in \{0, 1, 2, 3\}; \\ & \boldsymbol{s}_0 = (L, b_0), \end{aligned} \tag{3.10}$$

where the action policy is a set of sequential actions, denoted by $\boldsymbol{\pi} = \{\pi_t(l,b), \forall l \in \mathcal{L}, b \in \mathcal{B}, t \in \mathcal{T}\}$. Here, $\pi_t(l,b)$ is the action taken under state $\boldsymbol{s} = (l,b)$ at time t, i.e., $\pi_t(l,b) : \mathcal{L} \times \mathcal{B} \to \mathcal{A}$. The feasible set of $\boldsymbol{\pi}$ is denoted by Π. Let $\boldsymbol{s}^{\boldsymbol{\pi}}$ denote the state after policy $\boldsymbol{\pi}$ is taken at the state $\boldsymbol{s}$. $\mathbb{E}^{\boldsymbol{\pi}}_{\boldsymbol{s}_1}[\cdot]$ takes the expectation over the probability distribution of the UAV's mobility trajectory. $\boldsymbol{s}_0 = (L, b_0)$ is the initial state. b_0 is the UAV's starting street block.

Problem (3.10) is an FH-MDP problem [136, 149], and can be solved by DP. We define the step value function $V_t(\boldsymbol{s})$ as the Bellman equation [150], which is the minimum value of the expected cost function $H_t(l,b,a)$, as given by

$$V_t(\boldsymbol{s}) = V_t(l,b) = \min_{a \in \mathcal{A}^{(b)}} \{H_t(l,b,a)\}, \tag{3.11}$$

where

$$\begin{aligned} & H_t(l,b,a) \\ &= E_t(a,b) - G_t(a,b) + \sum_{b' \in \mathcal{B}} \sum_{l' \in \mathcal{L}} p\left((l',b')|(l,b),a\right) V_{t+\tau}(l',b') && \text{(3.12a)} \\ &= \min\left\{l, Z(a,b)\right\} \cdot \xi_t(a,b) + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{t+\tau}\left(\left[l - Z_t(a,b)\right]^+, b'\right). && \text{(3.12b)} \end{aligned}$$

The expected cost function $H_t(l,b,a)$ consists of two parts: the energy consumption at time t, and the expected cost at the next time slot $t+\tau$. (3.12b) is obtained by substituting the energy consumption (3.3), charged energy (3.4) and transition probability (3.7) into (3.12a). $\xi_t(a,b)$ is the normalized energy consumption for a unit flight distance, as given by

$$\xi_t(a,b) = \frac{E_t(a,b) - G_t(a,b)}{Z_t(a,b)} = \begin{cases} \frac{p(v_a)}{v_a}, & a \in \{0,1\}, b \in \mathcal{B}; \\ \frac{p(v_0)(\tau - \tau_c(b)) - \eta_a \tau_c(b)}{v_0(\tau - \tau_c(b) - \tau_d(a,b))}, & a = 2, b \in \{\mathcal{B}^{(1)}, \mathcal{B}^{(3)}\}; \\ \frac{p(v_0)\tau_d(a,b) - \eta_a(\tau - \tau_d(a,b))}{v_{car}(\tau - \tau_d(a,b)) - v_0\tau_d(a,b)}, & a = 3, b \in \{\mathcal{B}^{(2)}, \mathcal{B}^{(3)}\}. \end{cases} \tag{3.13}$$

Based on the principle of optimality [136], the optimal policy $\boldsymbol{\pi}^* = \{\pi_t^*(l,b), \forall l \in \mathcal{L}, b \in \mathcal{B}, t \in \mathcal{T}\}$ collects the optimal action taken at each state $\boldsymbol{s}$, as given by

$$\pi_t^*(\boldsymbol{s}) = \pi_t^*(l,b) = \underset{a \in \mathcal{A}^{(b)}}{\operatorname{argmin}} H_t(l,b,a). \tag{3.14}$$

Algorithm 1 summarizes the proposed DP-based algorithm to achieve the optimal policy, where the optimal action $\pi_t^*(l,b)$ is specified at each time slot by using the backward induction. Specifically, $\pi_t^*(l,b)$ is obtained by evaluating the expected cost function $H_t(l,b,a)$ in (3.12), and then updating these actions recursively backward from time slot $t = T$ to $t = 1$. The computational complexity of Algorithm 1 is $\mathcal{O}(ATL/v)$ [151]. With the increase of the flight distance L, the time T, and the location B, the state space of the FH-MDP can be large, resulting in a prohibitive computational complexity.

3.4 Monotone Optimal Policy and Threshold Structure

In this section, we reduce the computation complexity of the proposed DP-based Algorithm 1 by unveiling and exploiting the conditions of the optimal policy. Specifically, the optimal policy is monotone to the remaining flight distance l and the elapsed travel time t. The policy only changes when either of two conditions about l and t is met at every time instant. As a result, the optimal actions of the UAV can be generated by comparing the system state $\boldsymbol{s}_t$ with the thresholds of l and t developed from the conditions, instead of searching the entire state space.

Algorithm 1 DP-Based Optimal Policy Algorithm

Input: Required flight distance L, time constraints T, the map $\mathcal{B}$ containing the locations of charging station by GPS and willing vehicles.
Set boundary condition: $R_{T+1}(a, \boldsymbol{s})$ using (3.6)
Set $t \leftarrow T$
while $t > 1$ **do**
    for $b \in \mathcal{B}$ **do**
        Set $l \leftarrow 0$
        while $l < L$ **do**
            Compute $H_t(l, b, a)$ using (3.12)
            Set $\pi_t^*(l, b)$ using (3.14)
            $V_t(l, b) \leftarrow H_t(l, b, \pi_t^*(l, b))$
            $l \leftarrow l + v$
        end while
    end for
    $t := t - 1$
end while
Output the optimal policy $\boldsymbol{\pi}^*$.

3.4.1 Monotone Properties of Optimal Policy

We first deduce the monotone properties of the optimal policy. Assume that the charging time τ_c and the docking/latching time τ_d are location-independent, i.e., for $\forall b \in \mathcal{B}$, $\tau_c := \tau_c(b)$, $\tau_d(2) := \tau_d(2, b)$, and $\tau_d(3) := \tau_d(3, b)$. For notation brevity, we denote the power consumption $p_0 = p(v_0)$ and $p_1 = p(v_1)$. The distance traveled per

time slot in (4.2) can be rewritten as

$$z_a = \begin{cases} v_a\tau, & \text{if } a \in \{0,1\}; \\ v_0(\tau - \tau_c - \tau_d(2)), & \text{if } a = 2; \\ v_{car}(\tau - \tau_d(3)) - v_0\tau_d(3), & \text{if } a = 3. \end{cases} \tag{3.15}$$

From (3.15), $z_1 > z_0 \geq z_2$ holds for the non-zero docking/latching time of action $a = 2$, i.e., $\tau_d(2) > 0$. The normalized energy consumption of the UAV, i.e., (3.13), is written as

$$\xi_a = \begin{cases} p_0/v_0, & \text{if } a = 0; \\ p_1/v_1, & \text{if } a = 1; \\ \frac{p_0(\tau-\tau_c)-\eta_2\tau_c}{v_0[\tau-\tau_c-\tau_d(2)]}, & \text{if } a = 2; \\ \frac{p_0\tau_d(3)-\eta_3[\tau-\tau_d(3)]}{v_{car}[\tau-\tau_d(3)-v_1\tau_d(3)]}, & \text{if } a = 3. \end{cases} \tag{3.16}$$

According to [19], the normalized energy consumption of the UAV is higher at the full speed (action $a = 1$) than it is at the cruise speed, i.e., $\xi_1 > \xi_0$. Then, the expected cost function (3.12) can be rewritten as

$$H_t(l, b, a) = \min\{l, z_a\} \cdot \xi_a + \sum_{b' \in \mathcal{B}} \mathbb{P}(b'|b) V_{t+\tau}\left([l - z_a]^+, b'\right). \tag{3.17}$$

With the above observations, we reveal the following property of the optimal policy of the UAV.

Lemma 3.1. *The overtime penalty function satisfies*

$$E_{T+1}(l) - E_{T+1}(l - z_a) \geq E_t(a, b). \tag{3.18}$$

Proof. $E_t(a, b) = \xi_a z_a$ grows linearly with the traveled distance z_a. From (3.5), we obtain $E_{T+\tau}(l) - E_{T+\tau}(l - z_a) = \psi(l) - \psi(l - z_a)$. Since the penalty function $\psi(l)$ is a non-decreasing convex function of l, the marginal penalty function with respect to z_a, i.e., $\psi(l) - \psi(l - z_a)$, grows at least quadratically with z_a. Therefore, the marginal penalty with respect to z_a is larger than the energy cost $E_t(a, b)$. □

Lemma 3.1 indicates the possible penalty for late arrival is much higher than the energy consumption cost. We proceed to prove the monotone property of the minimum cost $V_t(l,b)$ in Lemma 3.2.

Lemma 3.2. *The Bellman equation (i.e., the minimum cost function), $V_t(l,b)$, satisfies*

(i) $V_t(l,b)$ is non-decreasing in l, $\forall b \in \mathcal{B}$, $t \in \mathcal{T}$.

(ii) $V_t(l,b)$ is non-decreasing in t, $\forall b \in \mathcal{B}$, $l \in \mathcal{L}$.

In other words, the expected cost is higher, when a longer flight distance is ahead, or the deadline is closer.

Proof. (a) By mathematical induction, we prove the Lemma (3.2), as follows. From $t = T+\tau$, we have $V_{T+\tau}(l,b) = E_{T+\tau}(l) = \psi(l)$, which is non-decreasing in l. Assume that $V_{t+1}(l,b)$ is non-decreasing in l. Therefore, $H_t(l,b,a)$ is a non-decreasing function of l from (3.12). As the minimum function does not change the monotonicity of the function, we observe that $V_t(l,b)$ is a non-decreasing function of l from (3.11). We can conclude that $V_t(l,b)$ is a non-decreasing function of l for $\forall t \in \mathcal{T} \bigcup \{T+\tau\}$.

(b) Similarly, we use mathematical induction from $t = T + \tau$

$$\begin{aligned} V_T(l,b) &= \min_{a \in \mathcal{A}^{(b)}} \{H_T(l,b,a)\} \leq \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}(l,b') \\ &= E_{T+\tau}(l) = V_{T+\tau}(l,b'). \end{aligned} \tag{3.19}$$

Assume that $\forall l \in \mathcal{L}$, $\forall b \in \mathcal{B}$, $V_{t+1}(l,b)$ is non-decreasing in t. We have $H_t(l,b,a)$ is a non-decreasing function of t, as can be obtained from (3.12). As the minimum function does not change the monotonicity of the function, $V_t(l,b)$ is also is a non-decreasing function of t from (3.11). We conclude that $V_t(l,b)$ is a non-decreasing function of t, for $\forall t \in \mathcal{T} \bigcup \{T+\tau\}$, which completes the proof. □

With the monotone property given in Lemma 3.2, we can conclude that the action needs to be taken only when these parameters satisfy certain conditions, as stated in Lemma 3.3.

Lemma 3.3. *(i) If the traveled distance yields $z_3 \geq z_1$, i.e., the latching/unlatching time of action $a = 3$, $\tau_d(3)$, and the vehicle's speed, v_{car}, satisfy*

$$v_{car} \geq v_{\max}, \tag{3.20}$$

$$\tau_d(3) \leq \frac{v_{car} - v_{\max}}{v_{car} + v_{\text{eco}}}\tau, \tag{3.21}$$

then action $a = 3$ is optimal in street block $b \in \{\mathcal{B}^{(2)}, \mathcal{B}^{(3)}\}$, $\forall l \in \mathcal{L}$, $t \in \mathcal{T}$, i.e.,

$$\pi_t^*(l, b) = 3. \tag{3.22}$$

(ii) If the normalized energy consumption of the UAV yields $\xi_2 \geq \xi_0$, i.e., the docking/undocking time of action $a = 2$, $\tau_d(2)$, and the charging time, τ_c, satisfy

$$\tau_d(2) \geq \frac{\eta_2}{p_0}\tau_c, \tag{3.23}$$

then action $a = 2$ should not be selected and the action set (4.1) is written as

$$\tilde{\mathcal{A}}^{(b)} = \begin{cases} \{0, 1\}, & \text{if } b \in \{\mathcal{B}^{(0)}, \mathcal{B}^{(1)}\}; \\ \{0, 1, 3\}, & \text{if } b \in \{\mathcal{B}^{(2)}, \mathcal{B}^{(3)}\}; \end{cases} \tag{3.24}$$

Proof. (a) We prove that $a = 3$ is the optimal action by showing that it can achieve the minimum cost $H_t(l, b, a)$ among all actions. Since $z_3 \geq z_1$, we have $z_3 \geq z_1 >$

$z_0 \geq z_2$. Then,

$$
\begin{aligned}
& H_t(l,b,a) \\
&= \min\{l, z_a\}\,\xi_a + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}\left(\left[l - z_a\right]^+, b'\right) && \text{(3.25a)} \\
&\geq \min\{l, z_3\}\,\xi_3 + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}\left(\left[l - z_a\right]^+, b'\right) && \text{(3.25b)} \\
&\geq \min\{l, z_3\}\,\xi_3 + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}\left(\left[l - z_3\right]^+, b'\right) && \text{(3.25c)} \\
&= H_t(l,b,3), && \text{(3.25d)}
\end{aligned}
$$

where (3.25b) is obtained due to the fact that $V_t(l,b)$ is non-decreasing in l, as shown in Lemma 3.2. Since action $a = 3$ can achieve the minimum expected cost function, $a = 3$ is optimal under the above conditions.

(b) We prove that action $a = 2$ should not be chosen by showing that the action gives the higher expected cost $H_t(l,b,a)$ than the action $a = 0$.

$$
\begin{aligned}
H_t(l,b,2) &= \min\{l, z_2\}\,\xi_2 + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}\left(\left[l - z_2\right]^+, b'\right) \\
&\geq \min\{l, z_2\}\,\xi_0 + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}\left(\left[l - z_2\right]^+, b'\right) \\
&\geq \min\{l, z_0\}\,\xi_0 + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}\left(\left[l - z_0\right]^+, b'\right) \\
&= H_t(l,b,0).
\end{aligned} \tag{3.26}
$$

Since action $a = 2$ always achieves a higher expected cost than action $a = 0$, it is always better to choose the action $a = 0$ instead. □

Lemma 3.3 indicates that action $a = 3$ is optimal, when the average driving speed of the vehicle hitchhiked by the UAV is faster than the UAV's maximum speed, and the latching/unlatching delay is negligible. The vehicle can slow down to allow the UAV to latch. Another observation is that the charging action $a = 2$ should not be chosen if the docking/undocking time is longer than the threshold provided in (3.23).

In other words, the threshold of taking action $a = 2$ depends on the charging time τ_c and the charging efficiency η_2 of a roadside charging station. It is concluded that the optimal action a changes as a function of the traveled distance z_a and the energy consumption ξ_a, while the willing vehicle's speed, the docking/undocking time, the latching/unlatching time, the charging time, and the charging efficiency serve as the parameters shaping the function. In this sense, we only need to consider z_a and ξ_a in the optimal action selection process in the following.

3.4.2 Threshold Structure of Optimal Policy

As revealed in Lemma 3.3, the optimal action is only chosen when a condition is met. We extract the conditions when the action can be possibly chosen and define the valid action set, denoted by $\tilde{\mathcal{A}}$, as stated Lemma 3.4.

Lemma 3.4. *(**Valid Action Set**) Consider two actions in the action set, $\hat{a}$, and $\check{a} \in \mathcal{A}$. If $z_{\hat{a}} \geq z_{\check{a}}$ and $\xi_{\hat{a}} \leq \xi_{\check{a}}$, then the action $\check{a}$ should not be chosen and needs to be precluded from the action set. The action set shrinks to $\tilde{\mathcal{A}} = \mathcal{A} \setminus \{\check{a}\}$.*

Proof. For action $\hat{a}$, and $\check{a} \in \mathcal{A}$, if $z_{\hat{a}} \geq z_{\check{a}}$ and $\xi_{\hat{a}} \leq \xi_{\check{a}}$,

$$\begin{aligned} H_t(l,b,\hat{a}) &= \min\{l, z_{\hat{a}}\}\,\xi_{\hat{a}} + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) V_{T+\tau}\left(\left[l - z_{\hat{a}}\right]^+, b'\right) \\ &= H_t(l,b,\hat{a}) + \left(\min\{l, z_{\hat{a}}\}\,\xi_{\hat{a}} - \min\{l, z_{\check{a}}\}\,\xi_{\check{a}}\right) \\ &\quad + \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) \left[V_{T+\tau}\left(\left[l - z_{\hat{a}}\right]^+, b'\right) - V_{T+\tau}\left(\left[l - z_{\check{a}}\right]^+, b'\right)\right] \\ &\leq H_t(l,b,\check{a}), \end{aligned} \tag{3.27}$$

where the inequality is based on Lemma 3.1. In other words, for action $\hat{a}$ and $\check{a}$ in the action set $\mathcal{A}$, if action $\check{a}$ provides in a larger expected cost $H_t(l,b,\hat{a})$ than action $\hat{a}$, it is always better to choose action $\hat{a}$. Therefore, action $\check{a}$ is not in the action set. □

With Lemma 3.4, we show the subadditivity and superadditivity of $H_t(l, b, a)$ in the valid action set with respect to the remaining flight distance l and the elapsed time t, as shown in Theorems 3.1 and 3.2. The definition and some properties of subadditivity and superadditivity are summarized in brief in Appendix A, with more details in [150, 152]. By comparing any two actions in the valid set, we first reveal that there exist the thresholds of the remaining flight distance and the elapsed time at which the optimal action changes in Theorems 3.1 and 3.2. Then, we derive the thresholds and propose the new threshold-based optimal action selection, as summarized in Theorem 3.3.

Theorem 3.1. *(**Distance threshold**) Consider any two actions in the action set, $\{a^+, a^-\} \in \mathcal{A}^{(b)}$, and the corresponding remaining distances l^+ and $l^- \in \mathcal{L}$, with $a^+ > a^-$ and $l^+ > l^-$:*

(i) If $z_{a^+} \geq z_{a^-}$ and $\xi_{a^+} \geq \xi_{a^-}$, $H_t(l, b, a)$ is superadditive in $\mathcal{L} \times \{a^-, a^+\}$ and $\pi_t^*(l, b)$ is a monotone increasing function of l. There exists a threshold of l, $\exists l^*(b, t) \geq 0$,

$$\pi_t^*(l, b) = \begin{cases} a^+, & \text{if } l \leq l^*(b, t); \\ a^-, & \text{otherwise.} \end{cases} \tag{3.28}$$

(ii) If $z_{a^+} \leq z_{a^-}$ and $\xi_{a^+} \leq \xi_{a^-}$, $H_t(l, b, a)$ is subadditive in $\mathcal{L} \times \{a^-, a^+\}$ and $\pi_t^*(l, b)$ is a monotone decreasing function of l. There exists a threshold of l, $\exists l^*(b, t) \geq 0$,

$$\pi_t^*(l, b) = \begin{cases} a^-, & \text{if } l \leq l^*(b, t); \\ a^+, & \text{otherwise.} \end{cases} \tag{3.29}$$

(iii) If $z_{a^+} \geq z_{a^-}$ and $\xi_{a^+} \leq \xi_{a^-}$, based on the valid action set in Lemma 3.4, a^- is precluded and

$$\pi_t{}^*(l, b) = a^+. \tag{3.30}$$

If $z_{a^+} \leq z_{a^-}$ *and* $\xi_{a^+} \geq \xi_{a^-}$, *action* a^+ *is precluded and*

$$\pi_t{}^*(l,b) = a^-. \tag{3.31}$$

Theorem 3.2. *(**Time threshold**) Consider any two actions in the action set,* $\{a^+, a^-\} \in \mathcal{A}^{(b)}$, *the time slot* t^+ *and* $t^- \in \mathcal{T}$, *where* $a^+ > a^-$ *and* $t^+ > t^-$*:*

(i) If $z_{a^+} \geq z_{a^-}$ *and* $\xi_{a^+} \geq \xi_{a^-}$, $H_t(l,b,a)$ *is superadditive in* $\mathcal{T} \times \{a^-, a^+\}$ *and* $\pi_t^*(l,b)$ *is a monotone increasing function of* t. *There exists a threshold of* t, $\exists t^*(b,l) \geq 0$,

$$\pi_t^*(l,b) = \begin{cases} a^-, & \text{if } t \leq t^*(b,l); \\ a^+, & \text{otherwise.} \end{cases} \tag{3.32}$$

(ii) If $z_{a^+} \leq z_{a^-}$ *and* $\xi_{a^+} \leq \xi_{a^-}$, $H_t(l,b,a)$ *is subadditive in* $\mathcal{T} \times \{a^-, a^+\}$ *and* $\pi_t^*(l,b)$ *is a monotone decreasing function of* t. *There exists a threshold of* t, $\exists t^*(b,l) \geq 0$,

$$\pi_t^*(l,b) = \begin{cases} a^+, & \text{if } t \leq t^*(b,l); \\ a^-, & \text{otherwise.} \end{cases} \tag{3.33}$$

(iii) If $z_{a^+} \geq z_{a^-}$ *and* $\xi_{a^+} \leq \xi_{a^-}$, *based on the valid action set in Lemma 3.4,* a^- *is precluded and*

$$\pi_t{}^*(l,b) = a^+. \tag{3.34}$$

If $z_{a^+} \leq z_{a^-}$ *and* $\xi_{a^+} \geq \xi_{a^-}$, *action* a^+ *is precluded and*

$$\pi_t{}^*(l,b) = a^-. \tag{3.35}$$

Proof. The proof of the threshold structure is based on the superadditivity or subadditivity of $H_t(l,b,a)$. First, we prove that $H_t(l,b,a)$ is a superadditive or subadditive function under different $\mathcal{L} \times \mathcal{A}^{(b)}$ conditions. For $a^+, a^- \in \mathcal{A}^{(b)}$ and $l^+, l^- \in \mathcal{L}$, with $a^+ > a^-$ and $l^+ > l^-$, we have

$$
\begin{aligned}
&H_t(l^+, b, a^+) - H_t(l^+, b, a^-) \\
&= z_{a^+}\xi_{a^+} - z_{a^-}\xi_{a^-} \\
&+ \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) \left(V_{T+\tau}\left(\left[l^+ - z_{a^+}\right]^+, b' \right) - V_{T+\tau}\left(\left[l^+ - z_{a^-}\right]^+, b' \right) \right).
\end{aligned} \tag{3.36}
$$

$$
\begin{aligned}
&H_t(l^-, b, a^+) - H_t(l^-, b, a^-) \\
&= z_{a^+}\xi_{a^+} - z_{a^-}\xi_{a^-} \\
&+ \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) \left(V_{T+\tau}\left(\left[l^- - z_{a^+}\right]^+, b' \right) - V_{T+\tau}\left(\left[l^- - z_{a^-}\right]^+, b' \right) \right).
\end{aligned} \tag{3.37}
$$

$$
\begin{aligned}
&[H_t(l^+, b, a^+) - H_t(l^+, b, a^-)] - [H_t(l^-, b, a^+) - H_t(l^-, b, a^-)] \\
&= \sum_{b' \in \mathcal{B}} \mathbb{P}\left(b'|b\right) \Big[\left(V_{T+\tau}\left(\left[l^+ - z_{a^+}\right]^+, b' \right) - V_{T+\tau}\left(\left[l^+ - z_{a^-}\right]^+, b' \right) \right) \\
&\quad - \left(V_{T+\tau}\left(\left[l^- - z_{a^+}\right]^+, b' \right) - V_{T+\tau}\left(\left[l^- - z_{a^-}\right]^+, b' \right) \right) \Big].
\end{aligned} \tag{3.38}
$$

With the incremental property, if $z_{a^+} \geq z_{a^-}$, we obtain

$$
H_t(l^+, b, a^+) - H_t(l^+, b, a^-) \geq H_t(l^-, b, a^+) - H_t(l^-, b, a^-). \tag{3.39}
$$

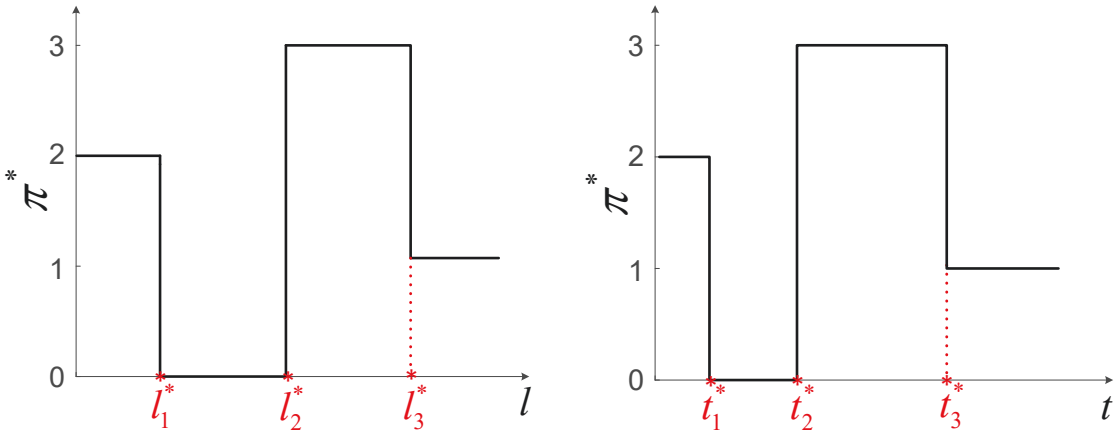
$H_t(l, b, a)$ is superadditive in $\mathcal{L} \times \{a^-, a^+\}$ and

$$
\pi_t^*(l, b) = \underset{a \in \{a^-, a^+\}}{\operatorname{argmin}} \; H_t(l, b, a). \tag{3.40}
$$

is monotone increasing in l. There exists such $l^*(b, t) \geq 0$ that: $\pi_t^*(l, b) = a^+$ when $l > l^*(b, t)$, and $\pi_t^*(l, b) = a^-$ when $l \leq l^*(b, t)$.

On the contrary, if $z_{a^+} < z_{a^-}$, then $H_t(l, b, a)$ is subadditive in $\mathcal{L} \times \{a^-, a^+\}$ and $\pi_t^*(l, b)$ is monotone decreasing in l. There exists such $l^*(b, t) \geq 0$ that: $\pi_t^*(l, b) = a^-$ when $l > l^*(b, t)$, and $\pi_t^*(l, b) = a^+$ when $l \leq l^*(b, t)$. This completes the proof of Theorem 3.1.

A similar approach can be applied to prove that $H_t(l, b, a)$ is superadditive in $\mathcal{T} \times \{a^-, a^+\}$ if $z_{a^+} \geq z_{a^-}$ and subadditive if $z_{a^+} < z_{a^-}$. We conclude that $\pi_t^*(l, b)$



(a) Threshold in l (b) Threshold in t

Figure 3.2 : Threshold illustration

is monotone increasing in t if $z_{a^+} \geq z_{a^-}$, and monotone decreasing in t if $z_{a^+} < z_{a^-}$. Then, we obtain the threshold in the elapsed time t:

1. if $z_{a^+} \geq z_{a^-}$, there exists such $t^*(b,l) \geq 0$ that: $\pi_t^*(l,b) = a^+$ when $t > t^*(b,l)$, and $\pi_t^*(l,b) = a^-$ when $t \leq t^*(b,l)$.

2. if $z_{a^+} < z_{a^-}$, there exists such $t^*(b,l) \geq 0$ that: $\pi_t^*(l,b) = a^-$ when $t > t^*(b,l)$, and $\pi_t^*(l,b) = a^+$ when $t \leq t^*(b,l)$,

which completes the proof of Theorem 3.2. By pairwise comparing the action in the valid action set as done in this example, we can obtain the threshold-based optimal policy, as summarized in Theorem 3.3.

□

For example, for $b \in \mathcal{B}^{(0)}$, we have the action set as $a \in \mathcal{A}^{(b)} = \{0,1\}$, $a^+ = 1$, and $a^- = 0$. Since $z_{a^+} > z_{a^-}$ and $\xi_{a^+} > \xi_{a^-}$, we conclude $H_t(l,b,a)$ is superadditive in $\mathcal{L} \times \{0,1\}$ and $\pi_t^*(l,b)$ is a monotone increasing function of l. In other words, there exists such $l^*(b,t) \geq 0$ that: $\pi_t^*(l,b) = 1$ when $l > l^*(b,t)$, and $\pi_t^*(l,b) = 0$ when $l \leq l^*(b,t)$.

3.4.3 Threshold-Based Optimal Policy Decision

Based on the thresholds established in Theorems 3.1 and 3.2, the optimal policy can be obtained by comparing the remaining flight distance and the elapsed time, as described in Theorem 3.3.

Theorem 3.3. *The optimal policy* $\boldsymbol{\pi}^* = \{\pi_t{}^*(l,b), \forall l \in \mathcal{L}, b \in \mathcal{B}, t \in \mathcal{T}\}$ *can be specified as*

1. *For* $b \in \mathcal{B}^{(0)}$ *and* $a \in \mathcal{A}^{(b)} = \{0,1\}$, *we have*

$$\pi_t{}^*(l,b) = \begin{cases} 0 \text{ (CS)}, & \text{if } l \le l_1^*(b,t); \\ 1 \text{ (FS)}, & \text{otherwise;} \end{cases} \tag{3.41}$$

$$\pi_t{}^*(l,b) = \begin{cases} 0 \text{ (CS)}, & \text{if } t \le t_1^*(b,l); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.42}$$

2. *For* $b \in \mathcal{B}^{(1)}$ *and* $a \in \mathcal{A}^{(b)} = \{0,1,2\}$, *we have*

$$\pi_t{}^*(l,b) = \begin{cases} 2 \text{ (CE)}, & \text{if } l \le l_1^*(b,t); \\ 0 \text{ (CS)}, & \text{if } l_1^*(b,t) < \ l \le l_2^*(b,t); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.43}$$

$$\pi_t{}^*(l,b) = \begin{cases} 2 \text{ (CE)}, & \text{if } t \le t_1^*(b,l); \\ 0 \text{ (CS)}, & \text{if } t_1^*(b,l) < \ t \le t_2^*(b,l); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.44}$$

3. *For* $b \in \mathcal{B}^{(2)}$ *and* $a \in \mathcal{A}^{(b)} = \{0,1,3\}$, *there are three cases:*

   (i) *Case 1: if* $z_0 > z_3$, *then*

$$\pi_t{}^*(l,b) = \begin{cases} 3 \text{ (VE)}, & \text{if } l \le l_1^*(b,t); \\ 0 \text{ (CS)}, & \text{if } l_1^*(b,t) < \ l \le l_2^*(b,t); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.45}$$

$$\pi_t{}^*(l,b)=\begin{cases} 3\ (\text{VE}), & \text{if } t \le t_1^*(b,l); \\ 0\ (\text{CS}), & \text{if } t_1^*(b,l) < \ t \le t_2^*(b,l); \\ 1\ (\text{FS}), & \text{otherwise.} \end{cases} \tag{3.46}$$

(ii) Case 2: if $z_1 > z_3 \ge z_0$, then

$$\pi_t{}^*(l,b)=\begin{cases} 3\ (\text{VE}), & \text{if } l \le l_1^*(b,t); \\ 1\ (\text{FS}), & \text{otherwise.} \end{cases} \tag{3.47}$$

$$\pi_t{}^*(l,b)=\begin{cases} 3\ (\text{VE}), & \text{if } t \le t_1^*(b,l); \\ 1\ (\text{FS}), & \text{otherwise.} \end{cases} \tag{3.48}$$

(iii) Case 3: if $z_3 \ge z_1$, then

$$\pi_t{}^*(l,b) = 3. \tag{3.49}$$

4. For $b \in \mathcal{B}^{(3)}$ and $a \in \mathcal{A}^{(b)} = \{0,1,2,3\}$, there are four cases:

(i) Case 1: if $z_0 > z_3 > z_2$, then

$$\pi_t{}^*(l,b)=\begin{cases} 2\ (\text{CE}), & \text{if } l \le l_1^*(b,t); \\ 3\ (\text{VE}), & \text{if } l_1^*(b,t) < \ l \le l_2^*(b,t); \\ 0\ (\text{CS}), & \text{if } l_2^*(b,t) < \ l \le l_3^*(b,t); \\ 1\ (\text{FS}), & \text{otherwise.} \end{cases} \tag{3.50}$$

$$\pi_t{}^*(l,b)=\begin{cases} 2\ (\text{CE}), & \text{if } t \le t_1^*(b,l); \\ 3\ (\text{VE}), & \text{if } t_1^*(b,l) < \ t \le t_2^*(b,l); \\ 0\ (\text{CS}), & \text{if } t_2^*(b,l) < \ t \le t_3^*(b,l); \\ 1\ (\text{FS}), & \text{otherwise.} \end{cases} \tag{3.51}$$

(ii) Case 2a: if $z_1 > z_3 \ge z_0 > z_2$ and $\xi_1 > \xi_0 \ge \xi_3 > \xi_2$, the action set reduces to the valid action set $\tilde{\mathcal{A}}^{(b)} = \{1,2,3\}$, then

$$\pi_t{}^*(l,b)=\begin{cases} 2\ (\text{CE}), & \text{if } l \le l_1^*(b,t); \\ 3\ (\text{VE}), & \text{if } l_1^*(b,t) < \ l \le l_2^*(b,t); \\ 1\ (\text{FS}), & \text{otherwise;} \end{cases} \tag{3.52}$$

$$\pi_t{}^*(l,b)=\begin{cases} 2 \text{ (CE)}, & \text{if } t \le t_1^*(b,l); \\ 3 \text{ (VE)}, & \text{if } t_1^*(b,l) < \ t \le t_2^*(b,l); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.53}$$

(iii) Case 2b: if $z_1 > z_3 \ge z_0 > z_2$ *and* $\xi_1 > \xi_3 > \xi_0 > \xi_2$*, then*

$$\pi_t{}^*(l,b)=\begin{cases} 2 \text{ (CE)}, & \text{if } l \le l_1^*(b,t); \\ 0 \text{ (CS)}, & \text{if } l_1^*(b,t) < \ l \le l_2^*(b,t); \\ 3 \text{ (VE)}, & \text{if } l_2^*(b,t) < \ l \le l_3^*(b,t); \\ 1 \text{ (FS)}, & \text{otherwise;} \end{cases} \tag{3.54}$$

$$\pi_t{}^*(l,b)=\begin{cases} 2 \text{ (CE)}, & \text{if } t \le t_1^*(b,l); \\ 0 \text{ (CS)}, & \text{if } t_1^*(b,l) \le \ t \le t_2^*(b,l); \\ 3 \text{ (VE)}, & \text{if } t_2^*(b,l) \le \ t \le t_3^*(b,l); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.55}$$

(iv) Case 3: if $z_3 \ge z_1$*, then*

$$\pi_t{}^*(l,b) = 3. \tag{3.56}$$

Here, $l_n^*(b,t)$ is the n-th threshold of l at time slot t and $n \in \{1,2,3\}$. $t_n^*(b,l)$ is the n-th threshold of t.

Theorem 3.3 establishes the threshold structure of the optimal policy to solve problem (3.10). The optimal actions are generated by comparing the system state $\boldsymbol{s}_t$ with the thresholds $l_n^*(b,t)$ and $t_n^*(b,l)$, instead of searching the state space. Fig. 3.2 illustrates the the optimal action function $\pi_t^*(l,b)$ with respect to l and t: For $b \in \mathcal{B}^{(3)}$, in Case 2b of Theorem 3.3-4), $z_1 > z_3 \ge z_0 > z_2$, $\xi_1 > \xi_3 > \xi_0 > \xi_2$, and the action set is $\mathcal{A}^{(b)} = \{0,1,2,3\}$. As shown in Fig. 3.2a, for a state $\boldsymbol{s}_t = (l,b)$, the optimal action is to choose 2 (CE), 0 (CS), 3 (VE), and 1 (FS), with the increase of the remaining flight distance l.

In practice, the latching/unlatching time of action $a = 3$ is relatively negligible, as compared to the docking/undocking time in action $a = 2$, i.e., $\tau_d(3) < \tau_d(2)$,

since a vehicle drives in the same direction as the UAV and the UAV does not need to reduce its speed to zero (as opposed to docking at a charging station). Thus, the normalized energy consumption yields $\xi_3 \leq \xi_2$ due to the short latching/unlatching time. As a result, the threshold-based optimal policy can be specified for street block $b \in \mathcal{B}^{(3)}$, where both CE ($a = 2$) and VE ($a = 3$) are available, as stated in Corollary 3.1.

Corollary 3.1. *If $\xi_3 \leq \xi_2$, then action $a = 2$ should not be chosen in any street block belonging to $\mathcal{B}^{(3)}$, and the valid action set is $\hat{\mathcal{A}}^{(3)} = \{0, 1, 3\}$. The threshold-based optimal policy specified in Theorem 3.2 can be simplified into three cases.*

1. *Case 1: If $z_1 > z_0 > z_3$, then*

$$\pi_t{}^*(l,b) = \begin{cases} 0 \text{ (CS)}, & \text{if } l \leq l_1^*(b,t); \\ 3 \text{ (VE)}, & \text{if } l_1^*(b,t) \leq l \leq l_2^*(b,t); \\ 1 \text{ (FS)}, & \text{otherwise;} \end{cases} \tag{3.57}$$

$$\pi_t{}^*(l,b) = \begin{cases} 0 \text{ (CS)}, & \text{if } t \leq t_1^*(b,l); \\ 3 \text{ (VE)}, & \text{if } t_1^*(b,l) \leq t \leq t_2^*(b,l); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.58}$$

2. *Case 2: If $z_1 > z_3 \geq z_0$, then*

$$\pi_t{}^*(l,b) = \begin{cases} 3 \text{ (VE)}, & \text{if } l \leq l_1^*(b,t); \\ 1 \text{ (FS)}, & \text{otherwise;} \end{cases} \tag{3.59}$$

$$\pi_t{}^*(l,b) = \begin{cases} 3 \text{ (VE)}, & \text{if } t \leq t_1^*(b,l); \\ 1 \text{ (FS)}, & \text{otherwise.} \end{cases} \tag{3.60}$$

3. *Case 3: If $z_3 \geq z_1$, then*

$$\pi_t{}^*(l,b) = 3. \tag{3.61}$$

3.4.4 Threshold-based Optimal Policy Algorithm

The thresholds can be evaluated efficiently by exploiting the property of the thresholds, as stated in Corollary 3.2.

Corollary 3.2. *The threshold $l_n^*(b,t)$ is a non-increasing function of t, i.e.,*

$$l_n^*(b,t) \geq l_n^*(b,t+\tau),\ \forall b \in \mathcal{B}. \tag{3.62}$$

The threshold $t_n^*(b,l)$ is a non-increasing function of l, i.e.,

$$t_n^*(b,l) \geq t_n^*(b,l+\upsilon),\ \forall b \in \mathcal{B}. \tag{3.63}$$

Proof. We prove from the corollary based on the definition of the threshold $l^*(b,t)$ in Theorems 3.1 and 3.2. We first prove Corollary 3.2 (i): In Case (i) of Theorem 3.1, for a given $b \in \mathcal{B}$ and $t \in \mathcal{T}$, based on the definition of the threshold $l^*(b,t)$ in (3.28), we have $\pi^*(l,b) = a^+$ if $l \geq l^*(b,t)$. From Lemma 4.1, we observe that the expected cost is higher when it is closer to the deadline. Thus, we have $\pi_{t+\tau}^*(l,b) = a^+$ if $l \geq l^*(b,t)$. Based on the definition of the threshold $l^*(b,t+\tau)$ at time $t+\tau$, we have $l^*(b,t+\tau) \geq l^*(b,t)$. This also holds in Case (ii) of Theorem 3.1, which completes the proof of (3.62).

Corollary 3.2 (ii) can be proved likewise, as follows: In Case (i) of Theorem 3.2, for a given $b \in \mathcal{B}$ and $l \in \mathcal{L}$, based on the definition of $t^*(b,l)$ in (3.32), we have $\pi_t^*(l,b) = a^+$ if $t \geq t^*(b,l)$. From Lemma 4.1, we observe that the expected cost is higher when a longer flight distance is ahead. Thus, $\pi_t^*(b,l+\upsilon) = a^+$ if $t \geq t^*(b,l)$. Based on the definition of the threshold in (3.32), we have $t^*(b,l) \geq t^*(b,l+\upsilon)$. The proof of Case (ii) can be similarly obtained, and (3.63) is proved. □

The optimal actions of the UAV can be generated by comparing the system state $\boldsymbol{s}_t$ with the thresholds of l and t developed from the conditions, instead of searching the entire state space. Based on Theorems 3.3, we summarize the proposed

threshold-based monotone optimal action selection policy in Algorithm 3, where the thresholds are generated by Algorithm 2.

In light of Corollary 3.2, the thresholds, $l_n^*(b,t)$ and $t_n^*(b,l)$, can be evaluated by first determining the number of thresholds at time slot t, i.e., $N_{\text{Thres}} = |\tilde{\mathcal{A}}^{(b)}| - 1$. Next, we arrange the action indexes in the ascending order of z_a in the set $\{1, \cdots, |\tilde{\mathcal{A}}^{(b)}|\}$, and calculate the i-th threshold $l_i^*(b,t)$ between the i-th and $(i+1)$-th action. Based on Corollary 3.2, we evaluate the distance threshold based on the threshold of the previous time slot, i.e., $l \leftarrow l_i^*(b, t+\tau)$ (see Line 6 in Algorithm 2). Finally, with the increase of l, the solution for $\text{argmin}_{a \in \mathcal{A}^{(b)}} H_t(l,b,a)$ switches from the i-th to the $(i+1)$-th action in the set. The value of l at which the switch occurs is the threshold $l_i^*(b,t)$. The time threshold $t_n^*(b,l)$ can be obtained in the same way. As revealed in Theorem 3, the optimal actions of the UAV ${\pi_t}^*(l,b)$ can be obtained by evaluating the thresholds $\{l^*, t^*\}$ with the state $\boldsymbol{s}_t = (l,b)$ (see Line 15 in Algorithm 3) retrospectively from $t = T$ to $t = 1$. The computational complexity of the optimal action selection drops from $\mathcal{O}(ALT/\upsilon)$ to around $\mathcal{O}(A \cdot \max\{L/\upsilon, T\})$ by using Algorithm 3 [153], as compared to Algorithm 1.

The proposed algorithms can be potentially extended for UAV swarms. In particular, the central coordinator monitors the road traffic and responds to the requests for goods delivery. Multiple UAVs can be assigned to deliver different goods. The coordinator plans the routes and schedules for the UAVs with the proposed algorithms, to minimize the total energy consumption of the UAVs, while ensuring each UAV's timely arrival at the destination. This can be done by running the proposed DP-based algorithm (Algorithm 1) or the proposed threshold-based algorithm (Algorithm 3). Specifically, the coordinator can generate the policy of the UAVs one after another by using the proposed algorithms (e.g., Algorithm 3). The policy of a UAV can impact on the policies of the others, given the shared resources, e.g., the available charging stations or willing vehicles at every time slot. Therefore, the

resources assigned to some of the UAVs are obsolete when scheduling the rest of the UAVs.

Algorithm 2 Threshold Generation Process

function Threshold Generation
  Obtain the number of threshold $N_{\text{Thres}} = |\tilde{\mathcal{A}}^{(b)}| - 1$
  Initialization: threshold set Ω
  Reorder the action index in the ascending order of z_a in the set $\{1, \cdots, |\tilde{\mathcal{A}}^{(b)}|\}$.
  for action $i \leftarrow \{1, \cdots, |\tilde{\mathcal{A}}^{(b)}|\}$ **do**
    Set $l \leftarrow l_i^*(b, t+\tau)$
    Set the threshold indicator $flag \leftarrow 0$
    while $l \leq L$ and $flag = 0$ **do**
      Compute $H_t(l, b, a)$ using (3.12);
      Set $\pi_t^*(l, b) \leftarrow \text{argmin}_{a \in \mathcal{A}^{(b)}} H_t(l, b, a)$
      Set $V_t(l, b) \leftarrow H_t(l, b, \pi_t^*(l, b))$
      if $\pi_t^*(l, b) = i$ **then**
        $l_i^*(b, t) \leftarrow l$
        $flag \leftarrow 1$;
      end if
      $l \leftarrow l + \upsilon$
    end while
  end for
end function

Algorithm 3 Threshold-based Optimal Policy Algorithm

Planning Stage: input L, T, $\mathcal{B}$

Set boundary condition: $R_{T+\tau}(a, \boldsymbol{s})$ using (4.7)

Set $t \leftarrow T$

while $t > 1$ **do**

    for $b \in \mathcal{B}$ **do**

        Call Threshold Generation Process

    end for

    $t \leftarrow t - \tau$

end while

Execution Stage: input threshold structure

Initialization: $t \leftarrow 1$, $l \leftarrow L$

while $t \leq T$ **do** and $l > 0$

    Get the block index of the UAV location at the time t: b

    Get the current state $\boldsymbol{s}_t = (l, b)$

    Decide the optimal action ${\pi_t}^*(l, b)$ by comparing the current state $\boldsymbol{s}_t$ with threshold structure $\{l^*, t^*\}$ with Theorem 3.3

    $a^* \leftarrow {\pi_t}^*(l, b)$

    Update $l \leftarrow [l - z_{a^*}]^+$, $t \leftarrow t + \tau$

end while

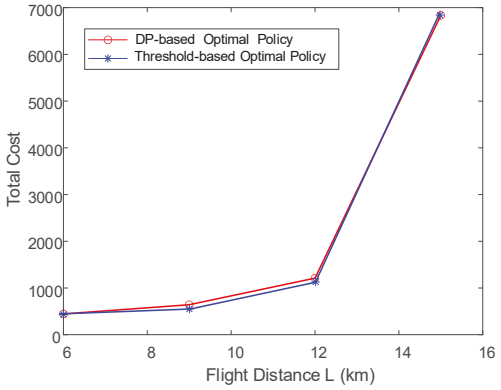
3.5 Numerical Results

In this section, we first verify the threshold structure of the monotone optimal policy, and show the optimality of the threshold-based action selection. Then, we evaluate the proposed schemes with practical octocopter flight parameters from DJI Agras MG-1 [154]. The energy consumption and the flight completion ratio [155] are evaluated in comparison with benchmarks. The maximum flight distance is also assessed with consideration of both the battery and arrival deadline T.

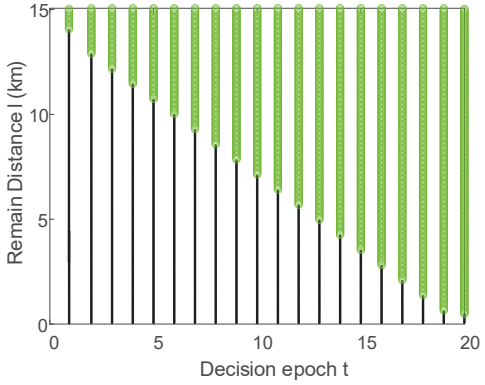
3.5.1 Threshold Structure Verification

We verify the threshold structure of the monotone optimal policy and its optimality in Fig. 3.3. First, we compare the proposed DP-based policy and threshold-based policy obtained by Algorithms 1 and 3, respectively. The total cost is plotted against the required flight distance L in Fig. 3.3a, where $T = 20$ min, $\langle z_0, z_1, z_2, z_3 \rangle = \langle 380, 720, 180, 600 \rangle$, and $\xi_1 > \xi_0 > \xi_2 > \xi_3$. It is observed that the threshold-based optimal policy achieves the same performance with the DP-based method. The threshold-based method achieves the optimality with a much lower computation complexity. Here, the total cost includes the energy consumption and the overtime penalty. As shown in Fig. 3(a), the total cost is dominated by the penalty when the flight distance $L > 15$ km. This is because the flight cannot be completed within 20 min.

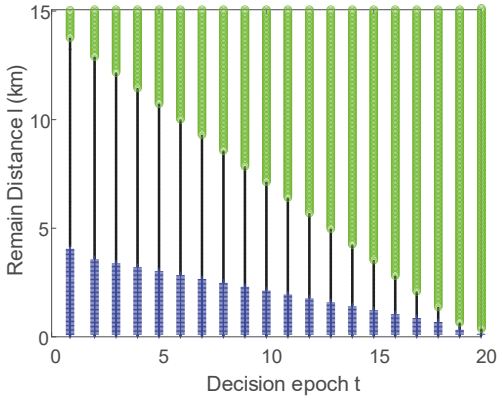
In Figs. 3.3b-3.3f, we illustrate the threshold structure of the optimal policy. As shown in Fig. 3.3b, for $b \in \mathcal{B}^{(0)}$, the valid set is $\mathcal{A}^{(0)} = \{0, 1\}$. $\pi_t^*(l, b)$ increases monotonically with the remaining distance l (i.e., from 0 to 1), when $z_1 > z_0$ and $\xi_1 > \xi_0$. There exists such $l^*(b, t) \geq 0$ that $\pi_t^*(l, b) = 1$ when $l > l^*(b, t)$, and $\pi_t^*(l, b) = 0$ when $l \leq l^*(b, t)$. The value of $l^*(b, t)$ is the $x-$axis coordinate of the switching point from action $a = 0$ (dark dot $\cdot$) to action $a = 1$ (green circle $\circ$) in Fig. 3.3b. This verifies the threshold structure in street blocks of the 0-th category,



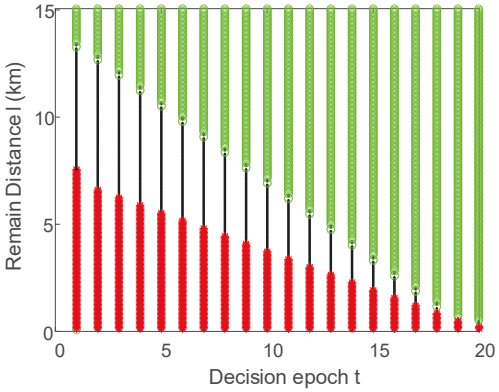
(a) Total cost versus required flight distance L for the given time $T = 20$ min



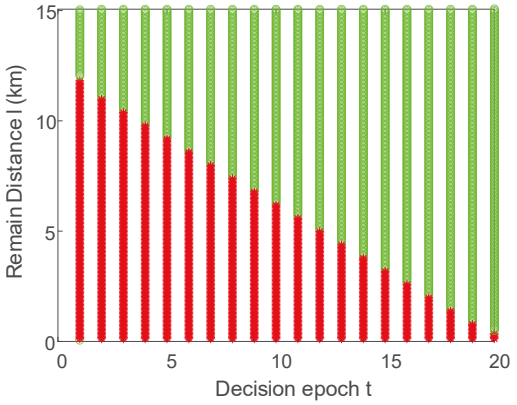
(b) $b \in \mathcal{B}^{(0)}$,$\langle z_0, z_1 \rangle = \langle 480, 720 \rangle$ m and $\xi_1 > \xi_0$



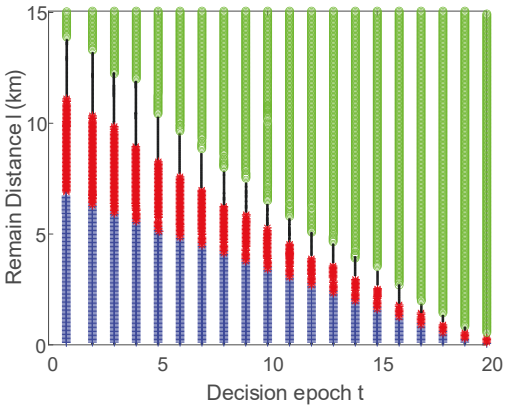
(c) $b \in \mathcal{B}^{(1)}$,$\langle z_0, z_1, z_2 \rangle = \langle 480, 720, 180 \rangle$ m and $\xi_1 > \xi_0 > \xi_2$



(d) $b \in \mathcal{B}^{(2)}$,$\langle z_0, z_1, z_3 \rangle = \langle 480, 720, 360 \rangle$ m and $\xi_1 > \xi_0 > \xi_3$



(e) $b \in \mathcal{B}^{(3)}$, $z_1 > z_3 > z_0 > z_2$, $\langle z_0, z_1, z_2, z_3 \rangle = \langle 480, 720, 180, 600 \rangle$ m and $\xi_1 > \xi_0 > \xi_2 > \xi_3$



(f) $b \in \mathcal{B}^{(3)}$, $z_1 > z_0 > z_3 > z_2$ $\langle z_0, z_1, z_2, z_3 \rangle = \langle 480, 720, 180, 360 \rangle$ m and $\xi_1 > \xi_0 > \xi_3 > \xi_2$

Figure 3.3 : Threshold structure verification: (b)-(f) are obtained under $L = 15$ km, and $T = 20$ min; dark dot $\cdot$, green circle $\circ$, blue plus $+$, and red star $*$ represent the CS ($a = 0$), FS ($a = 1$), CE ($a = 2$) and VE ($a = 3$) action, respectively.

i.e., $b \in \mathcal{B}^{(0)}$, as stated in Theorem 4.3-1). The other cases stated in Theorem 3 are validated in the same way in Figs. 3.3c-3.3f.

As shown in Fig. 3.3e, for $b \in \mathcal{B}^{(3)}$, $a \in \mathcal{A}^{(b)} = \{0, 1, 2, 3\}$ $z_1 > z_3 > z_0 > z_2$, and $\xi_1 > \xi_0 > \xi_2 > \xi_3$, and the valid set reduces to $\tilde{\mathcal{A}}^{(b)} = \{1, 3\}$, which is consistent with Case 2 in Corollary 3.1. In the case where $b \in \mathcal{B}^{(3)}$, $z_1 > z_0 > z_3 > z_2$, and $\xi_1 > \xi_0 > \xi_3 > \xi_2$, the threshold structure is shown in Fig. 3.3f, validating Case 1 of Theorem 4.3-4).

3.5.2 Performance Comparisons

We evaluate the schemes with practical octocopter flight parameters from DJI Agras MG-1 [154], where the UAV takeoff weight is 21 kg, the power consumption is $p_0 = 3250$ W under the cruise velocity of $v_0 = 8$ m/s, and the power consumption is $p_1 = 6300$ W under the full speed of $v_1 = 12$ m/s. The UAV is equipped with two battery packs (DJI MG-12000 [156]), each with the capacity of 532 Wh. Each battery pack has two charging modes: slow charging at 600 W and fast charging at 1200 W. We assume that the charging stations operate in the fast charging mode at $\eta_2 = 1200$ W, and the willing vehicles are only equipped with slow charging modules $\eta_3 = 600$ W.

The vehicle's speed is $v_{car} = 15$ m/s. The density of different street block categories is $\rho(\mathcal{B}) = [0.6, 0.1, 0.2, 0.1]$. The time slot length is $\tau = 1$ min. The average docking/undocking time is $\bar{\tau}_d(2) = 7.5$ s. The average latching/unlatching time is $\bar{\tau}_d(3) = 6$ s. The average charging time at a roadside charging station is $\bar{\tau}_c = 30$ s.

The proposed method is applied to select the action of the UAV in the presence and absence of collaborative willing vehicles. In particular,

1. Delay-sensitive action selection with vehicles collaboration (DVC): The UAV

selects the optimal actions with the proposed method, where all actions including different speeds, recharging at a roadside charging station, and hitchhiking are considered.

2. Delay-sensitive action selection with roadside charging station (DRC): The UAV selects the optimal actions obtained by the proposed method in the absence of willing vehicles ($a = 3$).

For comparison purpose, we also simulate the following benchmark schemes:

- On-the-spot action selection schemes: the UAV chooses to get recharged at a roadside charging station, or hitchhike on a willing vehicle whenever available. Two actions are chosen with an equal probability in the overlapping area ($\mathcal{B}^3$) [157]. The UAV's speed can be set up in the following three different ways as three benchmarks:

    1. On-the-spot with low-speed priority (OLP): The UAV chooses the cruise speed, when there is no charging station or willing vehicle.

    2. On-the-spot with equal speed probability (OEP): The UAV chooses the cruise speed and the full speed with equal probability, when there is no charging station or willing vehicle.

    3. On-the-spot with high-speed priority (OHP): The UAV chooses the full speed, when there is no charging station or willing vehicle.

- Speed selection only: The UAV only chooses the speed and does not get recharged at a roadside charging station or willing vehicles.

    1. Cruise-speed only (CSO): The UAV flies to the destination at the cruise speed.

    2. Full-speed only (FSO): The UAV flies to the destination at the full speed.

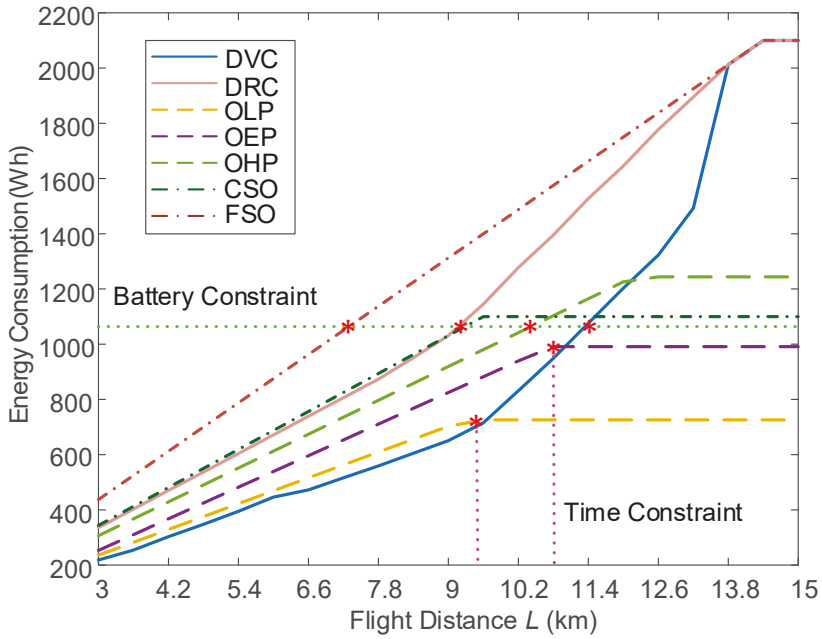


Figure 3.4 : Energy consumption under different flight distances under delay requirement $T = 20$ min

First, we compare the energy consumption under different flight distances under a given delay requirement T. When $T = 20$ min, the energy consumptions are shown in Fig. 3.4, where the star (*) marks the farthest flight distance that each scheme can achieve within the given time T or battery constraint. In the FSO scheme, the UAV always flies at the full speed, hence giving the steepest growth of energy consumption. Thus, FSO is the first to meet the battery constraints with the growing flight distance. It gives the minimum flight distance of 7.2 km among all schemes. Flying at a single speed, FSO and CSO only allow the UAV to rely on the battery. Their flight distances are shorter than the other schemes'. We observe that DRC and CSO have similar performances. This is because the allowed flight delay is relatively short, as compared to the large flight distance. The UAV using the DRC scheme is more likely to fly at the cruise or full speed to the destination to avoid the overtime penalty, even though the charging stations are available. However, a

slightly lower energy consumption observed in the DRC scheme can be observed when the overall distance is $L < 7$ km, as compared with CSO. The reason is that recharging at some roadside stations is still preferable in the DRC scheme when the overall flight distance is short.

By allowing hitchhiking and recharging at willing vehicles, OLP, OEP OHP, and DVC can fly long distances. The maximum flight distances of OLP and OEP depend on the flight delay T, rather than the battery capacity of the UAV. The on-the-spot action selection schemes, i.e., OLP, OEP and OHP, are opportunistic and do not optimize against T and other flight requirements. In the OLP scheme, the UAV cruises for energy saving when there is no charging station available, and gets recharged whenever possible without consideration of potential overtime. In this way, more energy is obtained but the arrival is delayed, as shown in Fig. 3.7. Being aware of trajectory and arrival deadline, the proposed methods, DVC and DRC, dynamically choose actions depending on the remaining flight time and flight distance, and can achieve the lowest energy consumption while guaranteeing timely arrival.

Fig. 3.5 compares the energy consumption of the considered schemes under $T =$ 30 min. The same observations of FSO, CSO, and DRC can be obtained as in Fig. 3.4. With a longer allowed flight delay, the maximum flight distances of the OLP, OEP, and OHP schemes depend primarily on the battery capacity of the UAV. Moreover, with the longer flight time, the UAV can get more opportunities to be recharged. This gives a gentle growth of the energy consumption, and extends the maximum flight distance of the OLP, OEP, OHP, and DVC schemes.

Fig. 3.6 shows the flight completion rate over the flight distance $L = 9$ km, as T grows. We see that the proposed method (DVC) is the only one that completes the flight within 14 min. The maximum flight duration of FSO is around 10 min, and

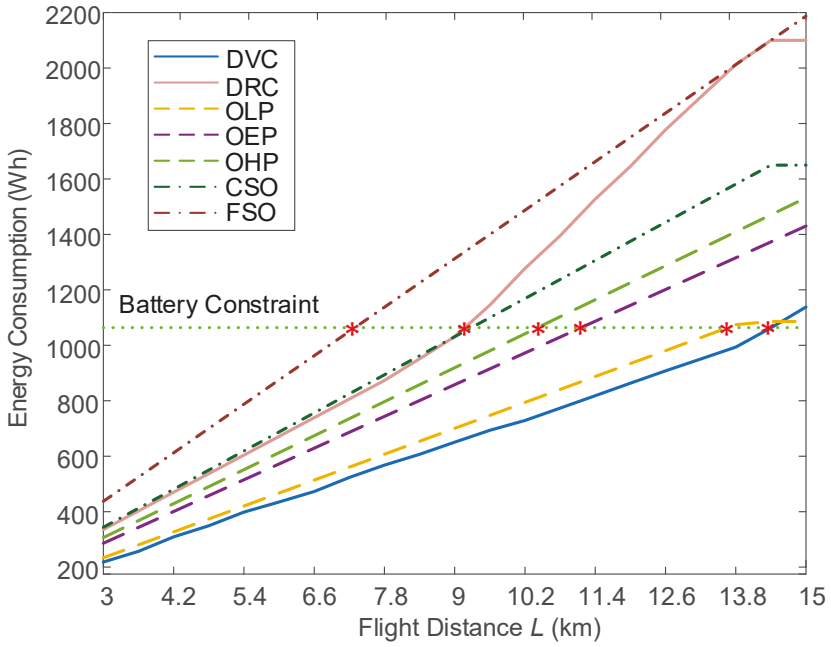


Figure 3.5 : Energy consumption under different flight distances under delay requirement $T = 30$ min

the maximum flight distance is 7.2 km. FSO cannot complete the flight, even with the growth of T.

Fig. 3.7 shows that the energy consumptions of the considered algorithms under different allowed flight delays T. The triangle ($\mathbf{\Delta}$) marks the earliest arrival time for the flight distance $L = 9$km. It is observed that the proposed method (DVC) completes the flight within 14 min at the cost of high energy consumption. The energy consumption continuously decreases by choosing to recharge and hitchhike. 39% and 10% energy reductions can be achieved, as compared with CSO and OLP schemes, respectively. In this sense, the proposed method DVC strikes a good balance between energy consumption and flight delay.

Fig. 3.8 plots the energy of the UAV obtained from the roadside charging stations and willing vehicles under different allowed flight delay T and the distance $L = 7.2$ km. The longer allowed flight time provides the UAV with more opportunities

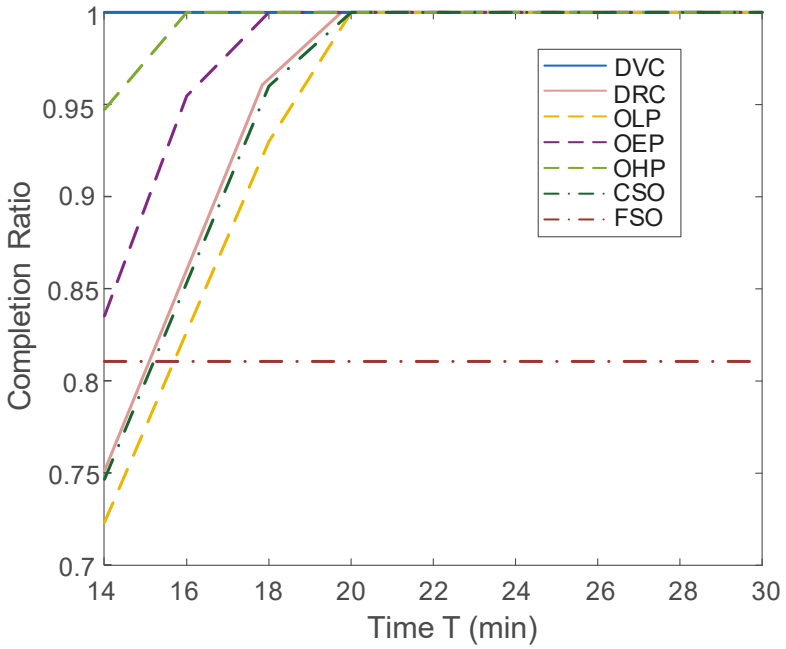


Figure 3.6 : Flight completion rate for the flight distance $L = 9$ km with the increase of T

to pass the street blocks with charging stations and/or willing vehicles. Without vehicles for hitchhiking and recharging, DRC obtains the energy from the charging station and the recharged energy of the UAV can increase with the allowed flight time. It is observed that the recharged energy obtained by the on-the-spot schemes keep the same under different allowed flight delays. This is because the on-the-spot schemes are unaware of the allowed flight delays, and cannot benefit from the roadside charging stations and willing vehicles. Moreover, the on-the-spot schemes choose actions without comparing their costs, and cannot provide the optimal action policy. The proposed methods, DVC and DRC, schedule actions with full awareness of the allowed flight delay and distance. Therefore, they can make the best use of the charging stations and willing vehicles, and achieve timely arrivals at the destination.

Fig. 3.9 evaluates the proposed method in terms of its capability to extend the flight distance, where flight completions are plotted under different flight delays.

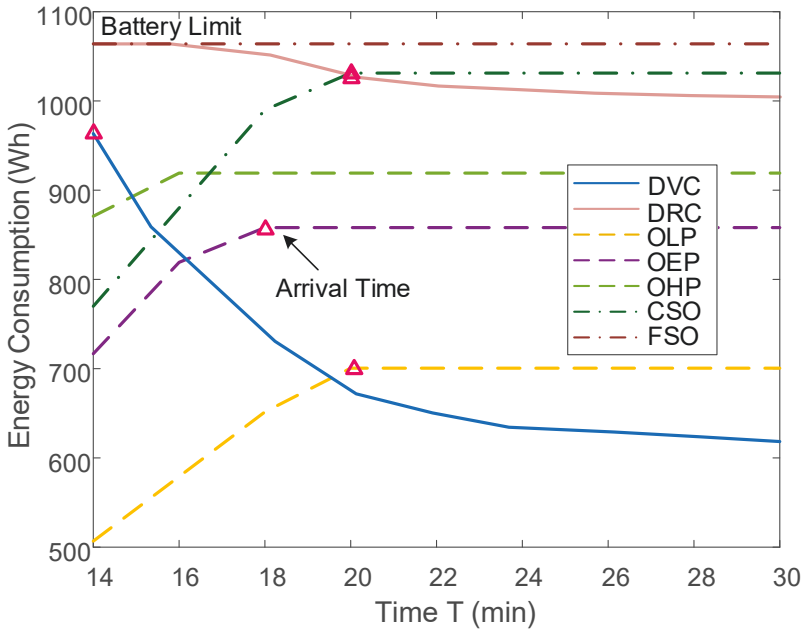


Figure 3.7 : Energy consumptions for the flight distance $L = 9$ km with the increase of T

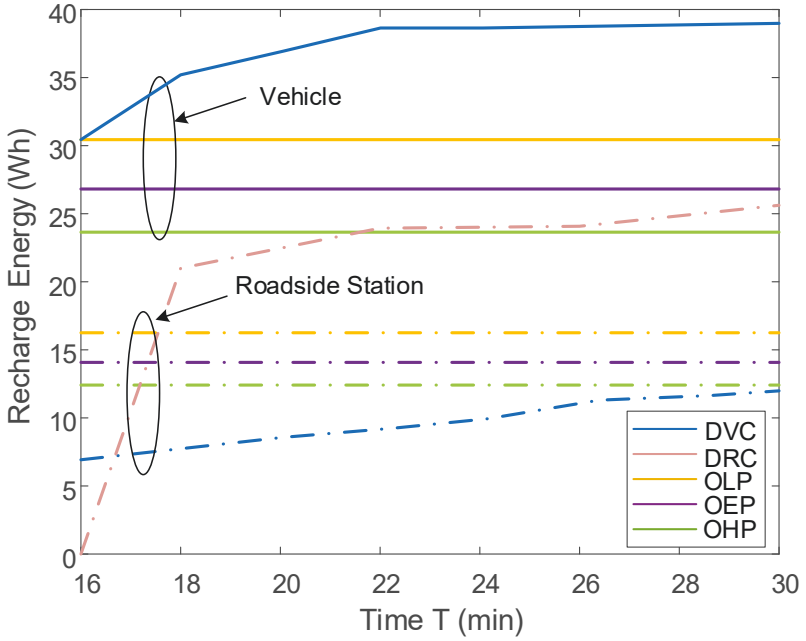


Figure 3.8 : Recharge energy under different allowed flight delay T for $L = 7.2$ km

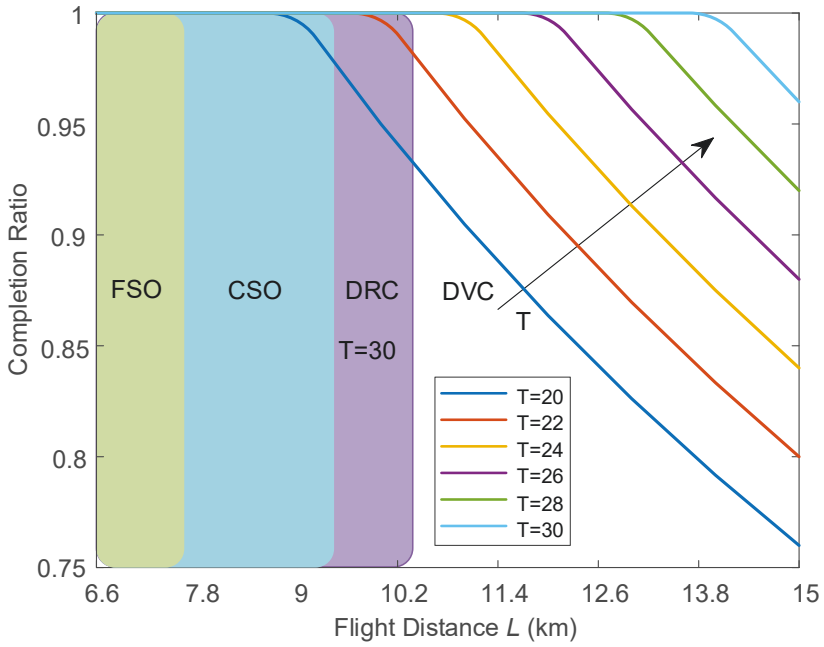


Figure 3.9 : Flight completions of the proposed method (DVC) under different flight delays

It is observed that the maximum flight distance increases with the allowed flight delay T. The maximum flight range can reach as far as 13.8 km. Even with no delay requirement, the maximum flight distances of FSO and CSO are around 7.2 km and 9.3 km, respectively. To this end, the proposed methods improve the flight distance by at least 48%. In particular, by getting recharged at the roadside charging stations during the flight alone, the UAV flight range of the proposed method DVC can approach to 10.3 km under $T = 30$ min. By leveraging hitchhiking and on-vehicle charging, 34% extension in the flight range can be achieved.

3.6 Conclusion

In this chapter, we integrated closely UAVs and ground vehicles to deliver delay-sensitive goods delivery services, where a UAV can hitchhike on collaborative vehicles for energy saving and recharging. We minimized the energy consumption of the

UAV and ensured its timely arrival at the destination, by optimizing its actions at different points of its planned route, including its flight mode (i.e., full speed or cruise speed), hitchhiking (and recharging on collaborative ground vehicles), or stopping and recharging at roadside charging stations. A new DP-based algorithm was first developed to obtain the optimal action of the UAV. Then, we revealed that the optimal action of the UAV only changes when the remaining flight distance or the elapsed time meets certain conditions. Accordingly, thresholds were derived for the distance and time. By comparing the remaining flight distance and the elapsed time with the thresholds, the optimal actions of the UAV can be instantly made. Simulation results showed that the proposed algorithms can improve the flight distance by 48%, as compared with existing alternatives. Our future work will include joint optimization of path planning and action selection to further improve the efficiency of goods delivery services.

Chapter 4

Optimal Control of Multi-Task UAV

4.1 Introduction

Equipped with embedded sensors, such as cameras, microphones, and thermometers, a UAV is often capable of conducting multiple different tasks [14, 15]. As illustrated in Fig. 4.1, a multi-task UAV equipped with sensors and dispatched to deliver goods can also carry out in-situ sensing along its delivery route.

Coordinating different tasks (e.g., goods delivery and in-situ sensing) is challenging to a multi-task UAV. One reason is that roadside tasks may require different levels of commitment and tolerate different latencies [158, 159]. In the given example of the delivery UAV which also performs in-situ sensing in Fig. 4.1, goods delivery is typically subject to strict deadlines, while in-situ sensing is a secondary task and only performed when possible. Another reason is that multi-tasking can give rise to difficulties in the energy management of a UAV [160]. It is important to utilize the battery energy efficiently while completing multiple tasks by their deadlines [21, 161]. The control of the multi-task delivery UAV would involve the joint optimization of the UAV's flying speed and task schedule along the trajectory. The energy consumption of UAVs is a nonlinear function of the speed, leading to non-convexity in both objectives and constraints [17]. The decisions consist of both continuous and binary variables, i.e., the UAV's speed and tasks selection [162]. Such a UAV scheduling problem is typically a mixed integer non-linear program (MINLP) [163, 164], requires a sequential decision process, and incurs a prohibitive computational complexity, especially when there is a large state space [165].

Existing studies have been predominantly focused on single-task UAVs. In [38], the navigation of a UAV was formulated as a generalized traveling salesman problem (TSP) for persistent monitoring missions. By determining the order of the targets to be visited and monitored, the longest interval elapsed between consecutive visits was minimized to optimize the route of the UAV. Based on a vector field approach, a curved path-following method was proposed in [39] to control a UAV conducting video tracking and surveillance. By exploiting the notion of input-to-state stability, the along-track error and the cross-track error were reduced for robust surveillance and tracking. Bartolini *et al.* [40] investigated the multi-trip task assignment of squads of UAVs for early target inspection. To maximize the target coverage, an NP-hard Integer Linear Programming (ILP) was formulated to schedule multiple UAVs. A greedy and prune (GaP) algorithm was developed to obtain 1/2-approximation of the optimality in polynomial time. Scheduling a truck-UAV delivery system was studied in [41] and [25], where a truck carries parcels and acts as a moving dock for UAVs. The scheduling policy developed in [41] and [25] ensured a UAV to take off from the truck, return to the truck after finishing the deliveries.

While a decomposition approach was adopted, a high computational complexity was still undergone [26]. A distributed schedule framework was proposed for UAV navigation and roadside charging in [43], and multi-agent deep reinforcement learning was used to minimize the energy consumption of the UAV. Zhang *et al.* [44] considered a UAV powered by both solar energy and roadside charging stations. By optimizing the UAV's trajectory and schedule policy, sustainable communication services were maintained with energy outage prevention. However, the above existing works cannot be extended to schedule a multi-task UAV running both goods delivery and in-situ sensing. Moreover, the existing works are computationally expensive despite their focus on single-task UAVs, due to the MINLP nature of the works [162].

This chapter presents a novel and optimal control policy for a multi-task UAV which performs delay-bounded goods delivery and in-situ sensing whenever possible. A new finite-horizon Markov decision process (FH-MDP) problem is formulated to ensure timely goods delivery, minimize the UAV's energy consumption, and maximize its reward for in-situ sensing. By proving the monotonicity and subadditivity of the FH-MDP, we confirm the existence of a monotone deterministic Markovian policy for the optimal action selection of the UAV. The optimal action (i.e., the speed and the decision of performing in-situ sensing) is monotone regarding the remaining trip and elapsed time. We also reveal the optimal action only changes when either of two thresholds regarding the remaining distance or elapsed time is met. By comparing its state with the thresholds, the UAV can optimally decide its action at each time slot. This is conducive to the online optimization of the delivery UAV's schedule, adapting to arising in-situ sensing requests.

The main contributions of this chapter are listed in the following.

- We establish the new multi-task UAV control framework, where the UAV can dynamically select its action at each slot, including full speed, cruise speed, or performing in-situ sensing. A new FH-MDP problem is formulated to maximize its reward for sensing, minimize the UAV's energy consumption, and ensure the ensuring timely goods delivery.

- By proving the monotonicity and subadditivity of the FH-MDP, we prove the existence of the monotone deterministic Markovian policy which is the optimal policy for the FH-MDP problem.

- We unveil the new threshold-based structure of the optimal policy, and derive the thresholds at which the action of the UAV optimally switches. The UAV can select its optimal action instantly at each time slot by comparing its current state with the thresholds at a linear complexity.

Figure 4.1 : Illustration of a multi-task UAV performing goods delivery and in-situ sensing at places of interest (POIs).

Extensive simulations corroborate that the new approach is able to guarantee the UAV's timely delivery, minimize its energy consumption, and maximize its sensing reward, as validated by comparison with the computationally expensive, standard dynamic programming (DP)-based solver for the FH-MDP problem. The simulations also show that our algorithm can save 24% of the energy of the UAV under our considered simulation setting, as compared to its alternative methods.

This chapter is arranged as follows. Section II depicts the system configuration. Section III casts the multi-task UAV control problem. In Section IV, we prove the threshold-based structure of the optimal policy, which can be obtained online with a considerably low complexity. Numerical and simulation results are provided based on practical flight parameters in Section V. Finally, Section VI concludes the chapter.

4.2 System Model

Fig. 4.1 illustrates the proposed multi-task UAV system, where a multi-task UAV delivers the parcels to the destination and performs in-situ sensing along the delivery route. The UAV is expected to reach the destination by no later than a specified time T, and the preplanned route length is M (km). The UAV can perform in-situ sensing at the places of interest (POIs) along the delivery route and report the sensing results to get rewards. The UAV needs to slow down and hover over POIs to perform sensing. Hovering over the POI takes extra time and may prevent the timely delivery. Since the UAVs are typically powered by on-board batteries, it is critical to meticulously control the speed of the UAV and sensing activities for timely goods delivery and maximum sensing rewards.

We divide T evenly into time slots with the duration of each slot being τ (in seconds). Let $t \in \mathcal{T} = \{1, 2, \cdots, T/\tau\}$ be the index to the time slots. The UAV can choose A number of different speeds, or hover and sense over POI at a time slot. Let v_a denote the UAV's speed, $a \in \mathcal{A} = \{1, \cdots, A\}$. $v_1 = v_{\max}$ is the full speed. $v_A = v_{\text{eco}}$ is the cruise speed. $a = A + 1$ indicates the UAV performs in-situ sensing. The total flight length M is discretized into segments with length λ per segment. Let $m \in \mathcal{M} = \{0, \lambda, 2\lambda \cdots, M\}$ be the index to the segments. We also divide the UAV's path in line with street blocks. The index to a street block is $l \in \mathcal{L} = \{1, \cdots, L\}$, where L denotes the number of street blocks along the flight path from start to destination. Each street block has a length of L_0. The street blocks are categorized into two types:

1. $\mathcal{L}^{(0)}$: The street blocks that have no POI;

2. $\mathcal{L}^{(1)}$: The street blocks that have POIs and require in-situ sensing.

We define $\boldsymbol{s} = (m, l)$ to be the state of the UAV. m is the distance ahead towards

the final destination. l is the UAV's current street block. Since in-situ sensing is only performed in the POIs, the different actions that the UAV can possibly take at different street blocks are expressed as

$$\mathcal{A}^{(l)} = \begin{cases} \mathcal{A}, & \text{if } l \in \mathcal{L}^{(0)}; \\ \mathcal{A} \bigcup \{A+1\}, & \text{if } l \in \mathcal{L}^{(1)}. \end{cases} \tag{4.1}$$

In other words, the action of selecting different speeds $a \in \mathcal{A}$ are always possible, while the action of in-situ sensing ($a = A+1$) heavily depends on the current location of the UAV.

4.2.1 Mobility Model of the UAV

The traveled distance of the UAV over a time slot when the UAV takes action a is given by

$$D_t(l,a) = D(l,a) = \begin{cases} v_a \tau, & \text{if } a \in \mathcal{A},\ l \in \mathcal{L} = \mathcal{L}^{(0)} \bigcup \mathcal{L}^{(1)}; \\ v_A(\tau - \tau_s(l)) + v_a \tau_s(l), & \text{if } a = A+1,\ l \in \mathcal{L}^{(1)}. \end{cases} \tag{4.2}$$

where $\tau_s(l)$ is the sensing duration at street block l. After sensing, the UAV proceeds at the cruise speed v_A for the rest of the slot.

4.2.2 Energy Consumption of the UAV

The UAV's energy consumption is primarily dominated by its propulsion energy and the in-situ sensing related energy [144,145]. The propulsion power consumption of the UAV (in Watts) depends on its flying speed and is usually a function of the flying speed, i.e., $p(v_a)$ [16]. Consider a rotary-wing UAV, and the propulsion power of the UAV at the speed v_a is given by [19]

$$\begin{aligned} p(v_a) = & \underbrace{P_0 \left(1 + \frac{3v_a^2}{Q_{\text{tip}}^2}\right)}_{\text{blade profile}} + \underbrace{P_i \left(\sqrt{1 + \frac{v_a^4}{4V_0^4}} - \frac{v_a^2}{2v_0^2}\right)^{1/2}}_{\text{induced}} \\ & + \underbrace{\frac{1}{2}\chi f_0 g W v_a^3}_{\text{parasite}}, \end{aligned} \tag{4.3}$$

where P_0 and P_i are the blade profile power and induced hovering power, respectively; Q_{tip} is the tip speed of the rotor blade; V_0 is the mean rotor induced hovering speed; f_0 and g are the fuselage drag ratio and rotor solidity, respectively; and χ and W are the air density and rotor disc area, respectively.

The energy consumption (in Joules) when action a is chosen at street block l can be written as

$$\begin{aligned} & E_t(l,a) = E_t(l,a) \\ & = \begin{cases} p(v_a)\tau, & \text{if } a \in \mathcal{A} \text{ and } l \in \mathcal{L}; \\ p(v_1)[\tau - \tau_s(l)] + [p(v_a) + p_s]\tau_s(l), & \text{if } a = A+1 \text{ and } l \in \mathcal{L}^{(1)}, \end{cases} \end{aligned} \tag{4.4}$$

where p_s is the normalized power consumption for collecting and reporting information at block l. When the UAV performs in-situ sensing, the UAV also consumes power to hover over the POIs. The energy used for in-situ sensing is usually much lower than the UAV's propulsion [19].

4.2.3 Reward for In-situ Sensing

A reward can be provided to the UAV for the in-situ sensing activities conducted at the POIs, as given by

$$R_t(l,a) = R(l,a) = \begin{cases} \mu\tau_s(l), & \text{if } a = A+1 \text{ and } l \in \mathcal{L}^{(1)}; \\ 0, & \text{otherwise,} \end{cases} \tag{4.5}$$

where μ is the unit (monetary) sensing reward per second.

4.2.4 Penalty for Late Arrival

The UAV is expected to reach the destination by time T. A (monetary) penalty is issued if the UAV fails to do so, i.e., the UAV arrives at or after time $T+1$. We define the penalty as

$$E_{T+1}(\boldsymbol{s}) = E_{T+1}(m) = \Phi(m), \tag{4.6}$$

where $\Phi(m)$ is non-decreasing and convex in m, and $\Phi(m) = 0$, $\forall m \leq 0$. The cost function concerning the considered penalty is expressed as

$$U_t(\boldsymbol{s}, a) = \begin{cases} E_t(l, a) - R_t(l, a), & \text{if } t \in \mathcal{T}; \\ E_{T+1}(m), & \text{if } t = T + 1. \end{cases} \tag{4.7}$$

4.3 Problem Statement

In this section, we select the actions of the UAV to minimize the energy usage of the UAV, maximize the sensing reward, and complete the goods delivery in time. With the definition of state $s = (m, l)$, the state transition probability is the probability of state $\boldsymbol{s}$ transiting to state $\boldsymbol{s}'$ if action a is performed, as given by

$$\begin{aligned} \Pr(\boldsymbol{s}'|\boldsymbol{s}, a) &= \Pr\left((m', l')|(m, l), a\right) \\ &= \Pr\left(l'|l\right) \Pr\left(m'|(m, l), a\right), \end{aligned} \tag{4.8}$$

where $\Pr\left(m'|(m, l), a\right)$ gives the distance transition probability, i.e., the conditional probability of the remaining distance transiting from m to m' if the UAV takes action a at street block l, and can be written as

$$\Pr\left(m'|(m, l), a\right) = \begin{cases} 1, \text{ if } m' = [m - D(l, a)]^+; \\ 0, \text{ otherwise}, \end{cases} \tag{4.9}$$

and $[m]^+ = \max\{0, m\}$. $\Pr(l'|l)$ is the location transition probability of the UAV when the UAV's location changes from street block l to street block l', as given by

$$\Pr\left(l'|l\right) = \begin{cases} 1, \text{ if } D(l, a) \geq L_0; \\ 0, \text{ otherwise}, \end{cases} \tag{4.10}$$

where the UAV flies from street block l to street block l' on the condition that the flight distance over the time slot outreaches a street block.

To minimize the energy consumption of the UAV, maximize the sensing reward,

and deliver the goods in time, the multi-task UAV control problem is formulated as

$$\begin{aligned} \underset{\boldsymbol{\pi}\in\Pi}{\text{minimize}} \quad & \mathbb{E}_{\boldsymbol{s}}^{\boldsymbol{\pi}}\left[\sum_{t=1}^{T} U_t\left(\boldsymbol{s}_t^{\boldsymbol{\pi}}, \pi_t\left(\boldsymbol{s}_t^{\boldsymbol{\pi}}\right)\right) + U_{T+1}\left(\boldsymbol{s}_{T+1}^{\boldsymbol{\pi}}\right)\right] \\ \text{s.t.} \quad & \pi_t \in \mathcal{A} \cup \{A+1\}; \\ & \boldsymbol{s}_0 = (M, l_0), \end{aligned} \tag{4.11}$$

where the control policy of the UAV, denoted by $\boldsymbol{\pi}$, is composed of sequential actions during the flight. $\boldsymbol{\pi} = \{\pi_t(m,l), \forall m \in \mathcal{M}, l \in \mathcal{L}, t \in \mathcal{T}\}$, where $\pi_t(m,l)$ denotes the action chosen at state $\boldsymbol{s} = (m,l)$ for $t \in \mathcal{T}$, i.e., $\pi_t(m,l)$ provides the mapping: $\pi_t(m,l) : \mathcal{M} \times \mathcal{L} \rightarrow \mathcal{A} \cup \{A+1\}$. All the possible action policies of the UAV constitute the feasible set, denoted by Π. $\boldsymbol{s}^{\boldsymbol{\pi}}$ represents the state after applying policy $\boldsymbol{\pi}$ under state $\boldsymbol{s}$. $\mathbb{E}_{\boldsymbol{s}}^{\boldsymbol{\pi}}[\cdot]$ denotes the expectation with respect to the probability mobility distribution of the UAV. $\boldsymbol{s}_0 = (M, l_0)$ is the initial state of the UAV. l_0 is the starting street block of the UAV.

Problem (4.11) is an FH-MDP problem. Conventionally, the optimal control policy can be obtained offline by using dynamic programming (DP) [136]. The minimum cost $V_t(\boldsymbol{s})$ is the optimal Bellman equation [150], as given by

$$V_t(\boldsymbol{s}) = V_t(m,l) = \min_{a \in \mathcal{A}^{(l)}} \{K_t(m,l,a)\}, \tag{4.12}$$

where $K_t(m,l,a)$ is the expected cost function and

$$\begin{aligned} K_t(m,l,a) &= E_t(l,a) - R_t(l,a) \\ &\quad + \sum_{l' \in \mathcal{L}} \sum_{m' \in \mathcal{M}} \Pr\left((m',l')|(m,l),a\right) V_{t+\tau}(m',l') \qquad (4.13a) \\ &= \min\left\{m, D_t(l,a)\right\} \cdot \zeta_t(l,a) \\ &\quad + \sum_{l' \in \mathcal{L}} \Pr\left(l'|l\right) V_{t+\tau}\left(\left[m - D_t(l,a)\right]^+, l'\right). \qquad (4.13b) \end{aligned}$$

The expected cost function $K_t(m,l,a)$ is composed of the immediate cost at slot t, and the estimated cost in the remaining slots. (4.13b) is achieved by substituting the energy consumption (4.4), the in-situ sensing reward (4.5), and the transition

probability (4.8) into (4.13a). $\zeta(l,a)$ is the normalized cost regarding the traveled distance per time slot, i.e.,

$$\zeta_t(l,a)=\frac{U_t(l,a)}{D_t(l,a)}=\begin{cases}\frac{p(v_a)}{v_a}, & \text{if } a\in\mathcal{A}, l\in\mathcal{L};\\ \frac{p(v_A)(\tau-\tau_s(l))+[p(v_a)+p_s-\mu]\tau_s(l)}{v_A(\tau-\tau_s(l))+v_a\tau_s(l)}, & \text{if } a=A+1, l\in\mathcal{L}^{(1)}.\end{cases} \tag{4.14}$$

From the principle of optimality [136], $\boldsymbol{\pi}^*=\{\pi_t^*(m,l),\forall m\in\mathcal{M},l\in\mathcal{L},t\in\mathcal{T}\}$ is the optimal policy, when the action selected at state $\boldsymbol{s}$ satisfies

$$\pi_t^*(\boldsymbol{s})=\pi_t^*(m,l)=\underset{a\in\mathcal{A}^{(l)}}{\operatorname{argmin}} K_t(m,l,a). \tag{4.15}$$

The optimal action $\pi_t^*(m,l)$ can be achieved by using the backward induction of DP. However, the computational complexity of DP is $\mathcal{O}((A+1)TM/\lambda)$ [166]. As the growth of the trip distance M, allowed time T, and speed mode number A, the state space of the multi-task UAV control problem, i.e., problem (4.11), can be large, incurring an excessive complexity.

4.4 Optimal Monotone Policy and Threshold Structure

If the Bellman equation of an FH-MDP satisfies two conditions of monotonicity and subadditivity, a monotone deterministic Markovian policy can be built as the optimal policy for the FH-MDP [167]. The monotone deterministic Markovian policy can be obtained by comparing the state of the FH-MDP with thresholds pertinent to the monotonicity and subadditivity conditions, hence reducing the computational complexity significantly.

In this section, we develop the monotone deterministic Markovian policy for the action selection of the multi-task UAV. Specifically, we first verify that the optimal Bellman equation of the FH-MDP, i.e., $V_t(m,l)$, is non-decreasing in both the remaining distance m and the elapsed time t. In other words, the Bellman

equation is monotone. We also prove the subadditivity of the expected cost function of the FH-MDP, $K_t(m, l, a)$. Then, we establish the optimal monotone deterministic Markovian policy of problem (4.11), based on the subadditivity, i.e., the optimal action $\pi_t^*(m, l)$ in (4.15) is monotone in both m and t. The optimal policy exhibits a threshold-based structure, where the switch of the optimal action is only triggered when either of two thresholds about m and t is met. By evaluating the UAV's state against the thresholds, the action of the UAV can be optimally selected at each time slot.

4.4.1 Monotonicity of Bellman Equation

We first reveal the monotonicity of the optimal Bellman equation $V_t(m, l)$ in (4.12). Assume that the sensing time τ_s is location-independent, i.e., $\tau_s := \tau_s(l)$, $\forall l \in \mathcal{L}$. The traveled distance of the UAV over a slot in (4.2) can be updated as

$$d_a = \begin{cases} v_a\tau, & \text{if } a \in \mathcal{A} \text{ and } l \in \mathcal{L}; \\ v_A(\tau - \tau_s) + v_{A+1}\tau_s, & \text{if } a = A+1 \text{ and } l \in \mathcal{L}^{(1)}. \end{cases} \tag{4.16}$$

From (4.16), $d_1 > \cdots > d_A > d_{A+1}$ holds for $v_{\max} > v_{\text{eco}} > v_{\text{hov}}$ and non-zero sensing time, i.e., $\tau_s > 0$. We normalize the cost function of the UAV, i.e., (4.14), with the unit traveled distance per time slot, as given by

$$\zeta_a = \begin{cases} p(v_a)/v_a, & \text{if } a \in \mathcal{A}; \\ \dfrac{p(v_A)(\tau - \tau_s) + [p(v_{A+1}) + p_s - \mu]\tau_s}{v_A(\tau - \tau_s) + v_{A+1}\tau_s}, & \text{if } a = A+1, l \in \mathcal{L}^{(1)}. \end{cases} \tag{4.17}$$

Hence, the expected cost function (4.13) is reformed as

$$K_t(m, l, a) = \min\{m, d_a\} \cdot \zeta_a + \sum_{l' \in \mathcal{L}} \mathbb{P}(l'|l) V_{t+1}\left([m - d_a]^+, l'\right). \tag{4.18}$$

The following lemma shows that the overtime penalty function for not completing the delivery in time is much higher than the rewards for executing the in-situ sensing, or the energy saving of choosing a lower speed.

Lemma 4.1. *The cost function* $U_t(l,a)$ *in* (4.7) *satisfies*

$$U_{T+1}(m) - U_{T+1}(m - d_a) \geq U_t(l,a). \tag{4.19}$$

Namely, the incurred overtime penalty is higher than the energy consumption for the same flying distance.

Proof. $U_t(l,a) = \zeta_a d_a$ is a linearly increasing function of the traveled distance d_a. With (4.6), for $t = T+1$, the cost function $U_{T+1}(m) - U_{T+1}(m - d_a) = \Phi(m) - \Phi(m - d_a)$. The penalty function $\Phi(m)$ gives a non-decreasing convex function of m, and $\Phi(m) - \Phi(m - d_a)$ rises more than quadratically with d_a. The penalty of not accomplishing a flight distance of d_a in time is larger than the cost of flying the distance. □

Based on Lemma 4.1, we establish the monotonicity of $V_t(m,l)$ in (4.12), as follows.

Theorem 4.1. ***(Monotonicity)*** $\forall l \in \mathcal{L}$, *the optimal Bellman equation* $V_t(m,l)$ *satisfies*

(i) $V_t(m,l)$ *does not decrease in* m, $\forall t \in \mathcal{T}$.

(ii) $V_t(m,l)$ *does not decrease in* t, $\forall m \in \mathcal{M}$.

Namely, the higher expected cost incurs when a longer flight distance or less flight time remains before the UAV can reach the destination in time. The Bellman equation (4.12) exhibits monotonicity.

Proof. (a) The proof of Theorem 4.1 is built on mathematical induction, as follows. From $t = T+1$, we get $V_{T+1}(m,l) = U_{T+1}(m) = \Phi(m)$, which is non-decreasing in m. Under the assumption that $V_{t+1}(m,l)$ increases with m, so is $K_t(m,l,a)$ based on

(4.13). Since the minimum function does not violate monotonicity, $V_t(m,l)$ is non-decreasing in m. In conclusion, $V_t(m,l)$ does not decrease in m, $\forall t \in \mathcal{T} \bigcup \{T+1\}$.

(b) The mathematical induction is adopted again. First, for $t = T$, we have

$$\begin{aligned} V_T(m,l) &= \min_{a \in \mathcal{A}^{(l)}} \{K_T(m,l,a)\} \leq \sum_{l' \in \mathcal{L}} \Pr(l'|l) V_{T+1}(m,l') \\ &= U_{T+1}(m) = V_{T+1}(m,l'). \end{aligned} \tag{4.20}$$

Under the assumption that $\forall l \in \mathcal{L}$, $\forall b \in \mathcal{L}$, we find that $V_{t+1}(m,l)$ does not decrease in t. Based on (4.13), $K_t(m,l,a)$ is non-decreasing in t. Since the minimum function preserves the monotonicity, $V_t(m,l)$ does not decrease with the increase of t, $\forall t \in \mathcal{T} \bigcup \{T+1\}$. □

4.4.2 Subadditivity and Optimal Policy

As revealed in Theorem 4.1, we obtain the monotonicity of the minimum cost function, which is the prerequisite for the optimal monotone deterministic Markovian policy of problem (4.11). In this subsection, we reveal the optimal policy, and show its threshold-based structure. In what follows, we first derive the irreducible set of candidate actions at any street block l, denoted by $\widetilde{\mathcal{A}}^{(l)}$. Then, we prove the subadditivity of the expected cost function in (4.13), i.e., $K_t(m,l,a)$, in the action set. Based on the subadditivity, we reveal the threshold-based structure of the optimal policy.

Lemma 4.2. *Suppose that there are two candidate actions, i.e., a^+, $a^- \in \mathcal{A}^{(l)}$. If*

$$d_{a^+} \geq d_{a^-} \text{ and } \zeta_{a^+} \leq \zeta_{a^-}, \tag{4.21}$$

then action a^- would never be selected in the optimal policy and the candidate action set can be reduced to $\widetilde{\mathcal{A}}^{(l)} = \mathcal{A}^{(l)} \setminus \{a^-\}$. In other words, the actions satisfying (4.21) *would consume more energy over the same flight distance and can be precluded for the selection of the optimal action.*

Proof. For actions a^+, $a^- \in \mathcal{A}$, if $d_{a^+} \geq d_{a^-}$ and $\zeta_{a^+} \leq \zeta_{a^-}$,

$$\begin{aligned}
&K_t(m, l, a^+) \\
&= \min\{m, d_{a^+}\}\zeta_{a^+} + \sum_{l' \in \mathcal{L}} \Pr(l'|l) V_{t+1}\left([m - d_{a^+}]^+, l'\right) \\
&= K_t(m, l, a^+) + (\min\{m, d_{a^+}\}\zeta_{a^+} - \min\{m, d_{a^-}\}\zeta_{a^-}) \\
&+ \sum_{l' \in \mathcal{L}} \Pr(l'|l)\left[V_{t+1}\left([m - d_{a^+}]^+, l'\right) - V_{t+1}\left([m - d_{a^-}]^+, l'\right)\right] \\
&\leq K_t(m, l, a^-),
\end{aligned} \tag{4.22}$$

where the inequality is obtained according to Lemma 4.1, namely, for actions a^+, $a^- \in \mathcal{A}$, if action a^- expects a higher cost $K_t(m, l, a^+)$ than action a^+, it is more cost-effective to select action a^+ instead. As a result, action a^- must not be included in the irreducible candidate action set. □

Based on Lemma 4.2, we can derive that the sensing action ($a = A+1$) can only be selected in the optimal policy when the following condition specified in Lemma 4.3 is satisfied.

Lemma 4.3. *If the distance traveled per time satisfies $d_A > d_{A+1}$, the sufficient condition of choosing the sensing action in the optimal policy is given by*

$$\zeta_{A+1} \leq \zeta_A. \tag{4.23}$$

There is also a low bound of the sensing reward per POI, denoted by μ_{Low}, which is achieved if $\tau_s = 1$ and satisfies

$$\mu_{\text{Low}} > p(v_{A+1}) + p_s - \frac{v_{\text{hov}}}{v_{\text{eco}}} p(v_A). \tag{4.24}$$

Proof. Lemma 4.3 is proved by showing that action $a = A+1$ should not be selected

if $\zeta_{A+1} > \zeta_A$, as it would lead to a higher expected cost $K_t(m, l, a)$,

$$
\begin{aligned}
& K_t(m, l, A+1) \\
&= \min\{m, d_{A+1}\}\,\zeta_{A+1} + \sum_{l' \in \mathcal{L}} \Pr(l'|l) V_{t+1}\left([m - d_{A+1}]^+, l'\right) && \text{(4.25a)} \\
&\geq \min\{m, d_{A+1}\}\,\zeta_{A} + \sum_{l' \in \mathcal{L}} \Pr(l'|l) V_{t+1}\left([m - d_{A+1}]^+, l'\right) && \text{(4.25b)} \\
&\geq \min\{m, d_{A}\}\,\zeta_{A} + \sum_{l' \in \mathcal{L}} \Pr(l'|l) V_{t+1}\left([m - d_{A}]^+, l'\right) && \text{(4.25c)} \\
&= K_t(m, l, A), && \text{(4.25d)}
\end{aligned}
$$

where (4.25b) is obtained by the condition $\zeta_{A+1} > \zeta_A$, (4.25c) is established, since $V_t(m, l)$ is a non-decreasing function of m, as stated in Theorem 4.1. Due to the fact that action $a = A+1$ results in a higher cost than action $a = A$ in any circumstance, action $a = A$ is more cost-effective and can be selected. As a result, action $a = A+1$ can only be chosen when $\zeta_{A+1} \leq \zeta_A$. □

Lemma 4.3 suggests that the in-situ sensing action $a = A + 1$ should not be chosen if the sensing reward is not larger than the threshold specified in (4.24), as the reward cannot compensate for the energy and time consumed to complete the sensing task. The threshold of selecting action $a = A + 1$ is dependent on the unit reward μ, flight speed, and power consumption, which are salient parameters of the optimal policy and thresholds.

Next, we introduce the definitions and important properties of subadditivity in Definition 4.1 [136, 150].

Definition 4.1. *Function $g(i, j, k)$ is subadditive in $\mathcal{I} \times \mathcal{J}$, for given k, if $\forall i^+, i^- \in \mathcal{I}$, $i^+ > i^-$, and $\forall j^+, j^- \in \mathcal{J}$, $j^+ > j^-$.*

$$
g(i^+, j^+, k) - g(i^+, j^-, k) \leq g(i^-, j^+, k) - g(i^-, j^-, k). \tag{4.26}
$$

If $g(i,j,k)$ is subadditive in $\mathcal{I}\times\mathcal{J}$ for a given k, then

$$f(i,k)=\arg\min_{j\in\mathcal{J}} g(i,j,k) \tag{4.27}$$

decreases monotonically with the increase of i.

Following Definition 4.1, we can show the subadditivity of $K_t(m,l,a)$, and the optimal action $\pi_t^*(m,l)$ in (4.15) is monotone in regards of the remaining distance m and the elapsed time t. By pairwise comparing the actions in the action set, we find two thresholds regarding m and t, at which a switch of the UAV's optimal actions takes place.

Theorem 4.2. *Suppose two candidate actions, a^+, $a^-\in\widetilde{\mathcal{A}}^{(l)}$ with $a^+>a^-$:*

*(i) (**Distance threshold**) If $d_{a^+}\le d_{a^-}$ and $\zeta_{a^+}\le\zeta_{a^-}$, then $K_t(m,l,a)$ yields subadditivity in $\mathcal{M}\times\{a^-,a^+\}$ and $\pi_t^*(m,l)$ decreases monotonically with m. The following threshold of m exists: $\exists m^*(l,t)\ge 0$,*

$$\pi_t^*(m,l)=\begin{cases} a^+, \text{ if } m\le m^*(l,t);\\ a^-, \text{ otherwise.}\end{cases} \tag{4.28}$$

*(ii) (**Time threshold**) If $d_{a^+}\le d_{a^-}$ and $\zeta_{a^+}\le\zeta_{a^-}$, then $K_t(m,l,a)$ yields subadditivity in $\mathcal{T}\times\{a^-,a^+\}$ and $\pi_t^*(m,l)$ decreases monotonically with t. The following threshold of t exists: $\exists t^*(l,m)\ge 0$,*

$$\pi_t^*(m,l)=\begin{cases} a^+, \text{ if } t\le t^*(l,m);\\ a^-, \text{ otherwise.}\end{cases} \tag{4.29}$$

Proof. The threshold-based structure can be revealed by exploiting the subadditivity of $K_t(m,l,a)$. First, we deduce that the subadditivity of $K_t(m,l,a)$ in $\mathcal{M}\times\mathcal{A}^{(l)}$. For $a^+,a^-\in\mathcal{A}^{(l)}$ with $a^+>a^-$ and $m^+,m^-\in\mathcal{M}$ with $m^+>m^-$,

$$\begin{aligned} &K_t(m^+,l,a^+)-K_t(m^+,l,a^-)\\ &=d_{a^+}\zeta_{a^+}-d_{a^-}\zeta_{a^-}\\ &+\sum_{l'\in\mathcal{L}}\Pr(l'|l)\left(V_{t+1}\left(\left[m^+-d_{a^+}\right]^+,l'\right)-V_{t+1}\left(\left[m^+-d_{a^-}\right]^+,l'\right)\right), \end{aligned} \tag{4.30}$$

$$
\begin{aligned}
&K_t(m^-, l, a^+) - K_t(m^-, l, a^-) \\
&= d_{a^+}\zeta_{a^+} - d_{a^-}\zeta_{a^-} \\
&+ \sum_{l' \in \mathcal{L}} \Pr(l'|l) \left(V_{t+1}\left(\left[m^- - d_{a^+}\right]^+, l' \right) - V_{t+1}\left(\left[m^- - d_{a^-}\right]^+, l' \right) \right),
\end{aligned} \tag{4.31}
$$

$$
\begin{aligned}
&[K_t(m^+, l, a^+) - K_t(m^+, l, a^-)] - [K_t(m^-, l, a^+) - K_t(m^-, l, a^-)] \\
&= \sum_{l' \in \mathcal{L}} \Pr(l'|l) \Bigg[\left(V_{t+1}\left(\left[m^+ - d_{a^+}\right]^+, l' \right) - V_{t+1}\left(\left[m^+ - d_{a^-}\right]^+, l' \right) \right) \\
&\quad - \left(V_{t+1}\left(\left[m^- - d_{a^+}\right]^+, l' \right) - V_{t+1}\left(\left[m^- - d_{a^-}\right]^+, l' \right) \right) \Bigg].
\end{aligned} \tag{4.32}
$$

If $d_{a^+} \leq d_{a^-}$, we obtain

$$
K_t(m^+, l, a^+) - K_t(m^+, l, a^-) \leq K_t(m^-, l, a^+) - K_t(m^-, l, a^-) \tag{4.33}
$$

$K_t(m, l, a)$ is subadditive in $\mathcal{M} \times \{a^-, a^+\}$. Based on Definition 4.1,

$$
\pi_t^*(m, l) = \underset{a \in \{a^-, a^+\}}{\operatorname{argmin}} K_t(m, l, a), \tag{4.34}
$$

monotonically decreases with the growth of m. We can obtain such $m^*(l, t) \geq 0$ that: $\pi_t^*(m, l) = a^+$ if $m \leq m^*(l, t)$ and $\pi_t^*(m, l) = a^-$ if $m > m^*(l, t)$.

Likewise, we can obtain that $K_t(m, l, a)$ is subadditive in $\mathcal{T} \times \{a^-, a^+\}$ if $d_{a^+} < d_{a^-}$, by taking the same approach. We conclude that $\pi_t^*(m, l)$ decreases monotonically in t if $d_{a^+} < d_{a^-}$. The threshold can be established regarding the elapsed time t: if $d_{a^+} < d_{a^-}$, we have $t^*(l, m) \geq 0$, and $\pi_t^*(m, l) = a^+$ if $t \leq t^*(l, m)$ or $\pi_t^*(m, l) = a^-$ if $t > t^*(l, m)$. This proves Theorem 4.2. By comparing the actions in the candidate action set, the threshold-based optimal monotone deterministic Markovian policy is established.

□

With the thresholds discovered in Theorem 4.2, the optimal monotone deterministic Markovian policy of the UAV's action selection can be achieved by evaluating m and t with the thresholds, as stated in Theorem 4.3.

Theorem 4.3. *Suppose that the UAV can take the actions of full speed, cruise speed, and in-situ sensing for illustration convenience (i.e., $A = 2$). The optimal policy $\boldsymbol{\pi}^* = \{{\pi_t}^*(m,l), \forall m \in \mathcal{M}, l \in \mathcal{L}, t \in \mathcal{T}\}$ can be given as*

1. *For $l \in \mathcal{L}^{(0)}$ and $a \in \mathcal{A}^{(l)} = \{1,2\}$, we have*

$${\pi_t}^*(m,l) = \begin{cases} 2 & \text{if } m \le m_1^*(l,t); \\ 1 & \text{otherwise;} \end{cases} \tag{4.35}$$

$${\pi_t}^*(m,l) = \begin{cases} 2 & \text{if } t \le t_1^*(l,m); \\ 1 & \text{otherwise.} \end{cases} \tag{4.36}$$

2. *For $l \in \mathcal{L}^{(1)}$ and $a \in \mathcal{A}^{(l)} = \{1,2,3\}$, we have*

$${\pi_t}^*(m,l) = \begin{cases} 3 & \text{if } m \le m_1^*(l,t); \\ 2 & \text{if } m_1^*(l,t) < \ m \le m_2^*(l,t); \\ 1 & \text{otherwise.} \end{cases} \tag{4.37}$$

$${\pi_t}^*(m,l) = \begin{cases} 3 & \text{if } t \le t_1^*(l,m); \\ 2 & \text{if } t_1^*(l,m) < \ t \le t_2^*(l,m); \\ 1 & \text{otherwise.} \end{cases} \tag{4.38}$$

Here, $m_n^*(l,t)$ represents the n-th threshold of the remaining flight distance m at slot t. $t_n^*(l,m)$ is the n-th threshold of the elapsed time t at the remaining distance m. $n \in \{1,2\}$. This theorem can be readily generalized to the case of $A > 2$.

Theorem 4.3 reveals the optimal policy of problem (4.11), which can be produced by making a comparison between the state $\boldsymbol{s}_t$ and the two thresholds $m_n^*(l,t)$ and $t_n^*(l,m)$, rather than searching the state space. Algorithm 4 describes the optimal policy with the thresholds generated using Algorithm 5 (as described below). The computational complexity of the optimal policy can be reduced from $\mathcal{O}((A+1)MT/\lambda)$ using the standard DP method to $\mathcal{O}\left((A+1)\cdot\max\{M/\lambda, T\}\right)$ using Algorithm 4 [168].

4.4.3 Threshold-based Action Selection

Theorem 4.2 exhibits the threshold-based structure of the optimal policy, which is monotone in m and t, and remains unchanged as long as neither of the two thresholds about m and t is satisfied. To derive the thresholds, we first present conditions that the thresholds satisfy, as specified in Corollary 4.1.

Corollary 4.1. *The distance threshold $m_n^*(l,t)$ does not increase with the elapsed time t, i.e.,*

$$m_n^*(l,t) \geq m_n^*(l,t+1),\ \forall l \in \mathcal{L}. \tag{4.39}$$

The time threshold $t_n^*(l,m)$ does not increase with the remaining flight distance m, i.e.,

$$t_n^*(l,m) \geq t_n^*(l,m+\lambda),\ \forall l \in \mathcal{L}. \tag{4.40}$$

Proof. Corollary 4.1 is proved by leveraging the threshold structure $m^*(l,t)$ in Theorem 4.2. The proof of Corollary 4.1 (i) is first given as follows. In Theorem 4.2, for given $l \in \mathcal{L}$ and $t \in \mathcal{T}$, by referring to the threshold structure $m^*(l,t)$ in (4.28), we draw $\pi^*(m,l) = a^-$ when $m \geq m^*(l,t)$. As shown in Lemma 4.1, it is observed that the estimated cost is higher if the remaining delivery time is less. Therefore, $\pi_{t+1}^*(m,l) = a^-$ if $m \geq m^*(l,t)$. According to the definition of the threshold $m^*(l,t+1)$ at time $t+1$, we have $m^*(l,t) \geq m^*(l,t+1)$.

Corollary 4.1 (ii) is proved likewise: For given $l \in \mathcal{L}$ and $m \in \mathcal{M}$, by leveraging the definition of $t^*(l,m)$ in (4.29), $\pi_t^*(m,l) = a^-$ when $t \geq t^*(l,m)$. According to Lemma 4.1, a longer remaining flight distance can result in a higher expected cost. $\pi_t^*(l,m+\lambda) = a^-$ if $t \geq t^*(l,m)$. With reference to the threshold structure defined in (4.38), we have $t^*(l,m) \geq t^*(l,m+\lambda)$. □

From Corollary 4.1, the thresholds, $m_n^*(l,t)$ and $t_n^*(l,m)$, are generated with the following steps. We start by determining the irreducible set, $\tilde{\mathcal{A}}$, based on

Algorithm 4 Threshold-based Action Selection Algorithm

1: *Planning Stage*: M, T, and $\mathcal{L}$.
2: Run Generation of the thresholding
3: *Selection Stage*:
4: Set $t \leftarrow 1$, $m \leftarrow M$
5: **while** $t \leq T$ and $m > 0$ **do**
6: Collect the current location of the UAV and obtain the current state $\boldsymbol{s}_t = (m, l)$
7: Obtain ${\pi_t}^*(m, l)$ by comparing $\boldsymbol{s}_t$ and the thresholds $\{m^*, t^*\}$
8: $a^* \leftarrow {\pi_t}^*(m, l)$
9: Update $m \leftarrow [m - d_{a^*}]^+$, $t \leftarrow t + 1$
10: **end while**

Lemma 4.2, and decide the number of thresholds needed per slot t, $N_{\text{Thres}} = |\tilde{\mathcal{A}}^{(l)}| - 1$. The indexes to the actions are reordered in the descending order of d_a in the set $\{1, \cdots, |\tilde{\mathcal{A}}^{(l)}|\}$. Then, we assess the n-th threshold, $m_n^*(l, t)$, between the n-th and the $(n+1)$-th action. Referring to Corollary 4.1, the UAV can start evaluating the distance threshold on the basis of the threshold at slot $t+1$, i.e., $m \leftarrow m_n^*(l, t+1)$. Finally, with the growth of m, the action that satisfies $\text{argmin}_{a \in \mathcal{A}^{(l)}} K_t(m, l, a)$ changes from the n-th to the $(n+1)$-th in the action set $\tilde{\mathcal{A}}^{(l)}$. The value of m when the UAV changes its action is the threshold $m_n^*(l, t)$. The same approach can apply to produce the time threshold $t_n^*(l, m)$.

Remark: It may not be possible to acquire all sensing requests beforehand in practice [169]. For example, a POI-free street block can become a POI street block when a new demand arises for in-situ sensing in the street block. Algorithm 4 suits online applications, as it provides the optimal UAV schedule based on in-situ requests arisen. Specifically, when a new request of in-situ sensing is generated along

Algorithm 5 Generation of the thresholds

Calculate the total number of thresholds $N_{\text{Thres}} = |\mathcal{A}^{(l)}| - 1$
for action $n \leftarrow \{1, \cdots, |\tilde{\mathcal{A}}^{(l)}|\}$ **do**
    Set $m \leftarrow m_n^*(l, t+1)$
    Set the threshold indicator $flag \leftarrow 0$
    while $m \leq M$ and $flag = 0$ **do**
        Compute $K_t(m, l, a)$ using (4.13);
        Set $\pi_t^*(m, l) \leftarrow \text{argmin}_{a \in \mathcal{A}^{(l)}} K_t(m, l, a)$
        Set $V_t(m, l) \leftarrow K_t(m, l, \pi_t^*(m, l))$
        if $\pi_t^*(m, l) = n + 1$ **then**
            $m_n^*(l, t) \leftarrow m$
            $flag \leftarrow 1$;
        end if
        $m \leftarrow m + \lambda$
    end while
end for

the UAV flight, we first compare the unit reward of the new in-situ sensing request with the low bound of the unit reward μ_{Low} given in (4.24). If the unit reward of the new in-situ sensing request is lower than μ_{Low}, the in-situ sensing is not performed and the UAV maintains the current schedule. If the unit reward of the in-situ sensing request is higher than μ_{Low} and the sufficient condition in Lemma 4.3 is satisfied, Algorithm 2 is executed to update the thresholds of the optimal policy, followed by Algorithm 1 to compare the UAV's current state $\boldsymbol{s}_t$ with the thresholds to decide if the new in-situ sensing task can be performed without jeopardizing the timely delivery. In this way, the algorithms can run online to update the UAV's schedule in response to new in-situ sensing requests arisen.

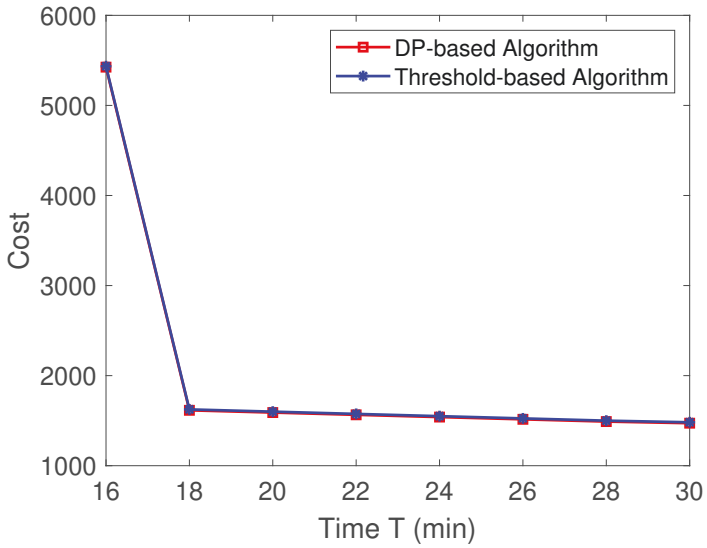


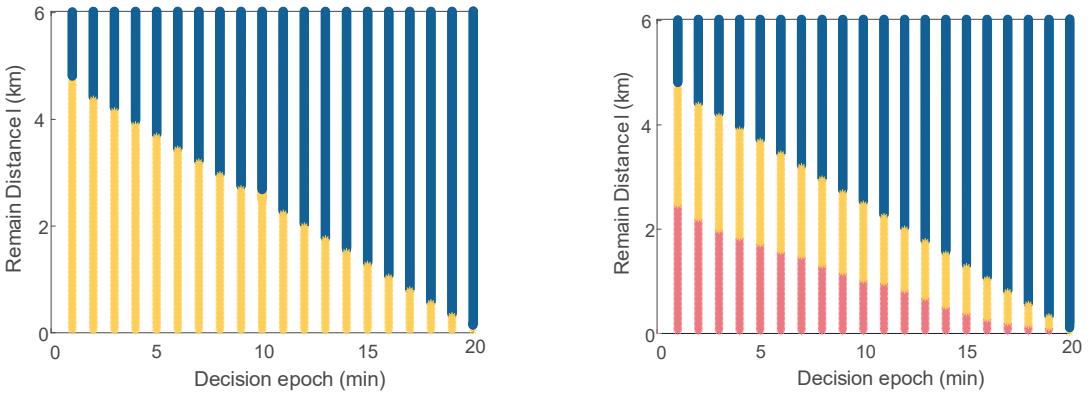
Figure 4.2 : Total cost with the increase of required arrival time T for $M = 10.2$ km

4.5 Numerical and Simulation Result

In this section, we evaluate the policy based on real-life UAV flight parameters from DJI Agras [154]. Comparison studies with identified benchmarks are given in terms of the UAV's energy consumption, in-situ sensing reward, and timely flight completion ratio [155].

4.5.1 Threshold Structure Verification

Fig. 4.2 verifies the new threshold-based optimal policy, i.e., the proposed Algorithm 1 by comparing with the DP-based optimal policy described at the end of previous section. Fig. 4.2 plots the total cost of the flight under the different values of the required arrival T when the distance $M = 10.2$ km. The total cost of the flight is defined in (4.7), consisting of the energy usage and overtime penalty. Fig. 4.2 confirms that the new threshold-based policy coincides with the conventional optimal DP-based approach, with a significantly lower computation complexity. As also observed in Fig. 4.2, the penalty of late arrival dominates the total cost when the



(a) $l \in \mathcal{L}^{(0)}$, $d_1 > d_2$ $\zeta_1 > \zeta_2$ (b) $l \in \mathcal{L}^{(1)}$, $d_1 > d_2 > d_3$, $\zeta_1 > \zeta_2 > \zeta_3$

Figure 4.3 : Validation of the new threshold-based optimal policy, where are achieved when $M = 6$ km, and $T = 20$ min; blue circle $\circ$, yellow star $*$, and red stars $*$ stand for actions $a = 1$ (full speed), $a = 2$ (cruise speed), and $a = 3$ (in-situ sensing), respectively.

required delay time is $T < 18$ min, since the flight cannot finish in time.

Figs. 4.3a and 4.3b exhibit the threshold-based structure of the optimal monotone deterministic Markovian policy. Fig. 4.3a shows that, for $l \in \mathcal{L}^{(0)}$, $\mathcal{A}^{(0)} = \{1, 2\}$. $\pi_t^*(m, l)$ decreases as the remaining distance m grows (i.e., from 1 to 2), when $d_1 > d_2$ and $\zeta_1 > \zeta_2$. We can see such $m^*(l, t) \geq 0$ that $\pi_t^*(m, l) = 2$ when $m \leq m^*(l, t)$, and $\pi_t^*(m, l) = 1$ when $m > m^*(l, t)$. $m^*(l, t)$ provides the x-coordinate value of the switching point where the optimal switches from action $a = 2$ (yellow star $*$) to action $a = 1$ (blue circle $\circ$), as the growth of the remaining flight distance m. This validates the new threshold-based structure in a POI-free street block, i.e., $l \in \mathcal{L}^{(0)}$, as stated in Theorem 4.3-1). It is also observed that the value of the threshold $m^*(l, t)$ is monotonic to the elapsed time t, which characterizes the monotonic property of the threshold, as described in Corollary 3.2. The case of a POI street block described in Theorem 4.3-2) can be verified in the same way in Fig. 4.3b, and is suppressed for brevity.

4.5.2 Performance Comparison

We assess the proposed algorithms based on real-life UAV flight parameters from DJI Agras [154]. The full speed is $v_1 = 10$ m/s, the cruise speed is $v_2 = 6$ m/s, and the hover speed is $v_3 = 1$ m/s. The power consumption is $p_1 = 5700$ W and $p_2 = 3200$ W for flying the full speed and the cruise speed, respectively. The total power consumption of hovering for in-situ sensing is $p_3 + p_s = 3300$ W. The time slot length is $\tau = 1$ min. The unit sensing reward per second is $\mu = 3000$ W, and the average sensing time is $\bar{\tau}_s = 48$ s. The density of street blocks with sensing requirements is $\rho(\mathcal{L}^{(1)}) = 0.4$; unless specified otherwise.

For comparison, we also evaluate the benchmarks based on on-the-spot action policy [170]:

- Benchmark 1: The UAV performs in-situ sensing when required, and chooses the full speed for the rest of the flight.
- Benchmark 2: The UAV performs in-situ sensing when required, and chooses the cruise speed for the rest of the flight.
- Benchmark 3: The UAV chooses randomly 50% of the POI street blocks to conduct in-situ sensing, and flies at the full speed for the rest of the flight.
- Benchmark 4: The UAV chooses randomly 50% of the POI street blocks to conduct in-situ sensing, and cruises for the rest of the flight.

Fig. 4.4 plots the energy usage under different required delivery distances, when the specified delay T is 30 min. The star ($*$) represents the outmost flight distance of the UAV that each scheme can reach by T. When the UAV always chooses the full speed (namely, Benchmarks 1 and 3), the energy consumption of the UAV undergoes the steepest growth. When the UAV chooses to cruise for energy saving

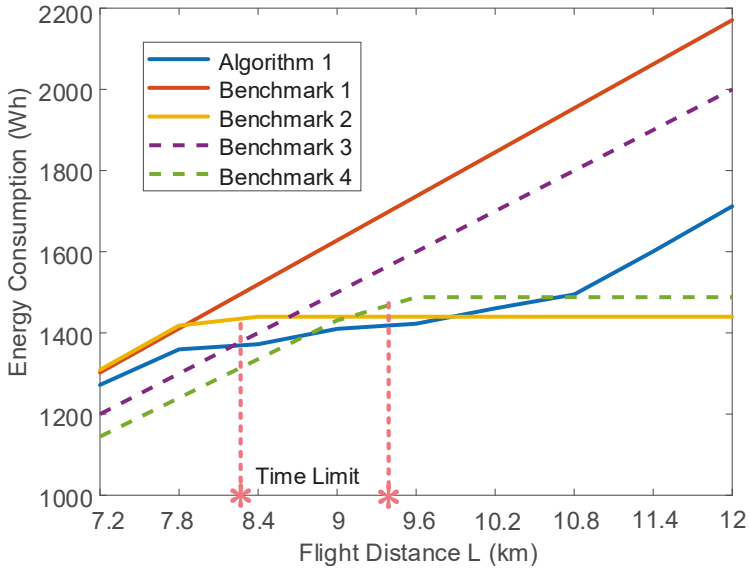


Figure 4.4 : The energy usage against the required delivery distances, when the specified delay T is 30 min.

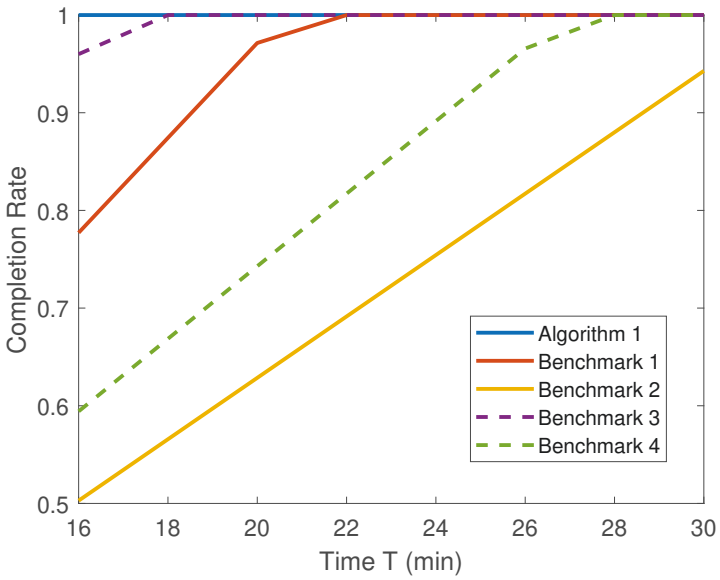


Figure 4.5 : The timely flight completion ratio of the distance $M = 8.4$ km, as the allowed flight time T increases.

(namely, Benchmarks 2 and 4), more energy is saved but the flight distance is limited. In particular, in Benchmark 2, the UAV performs in-situ sensing at all

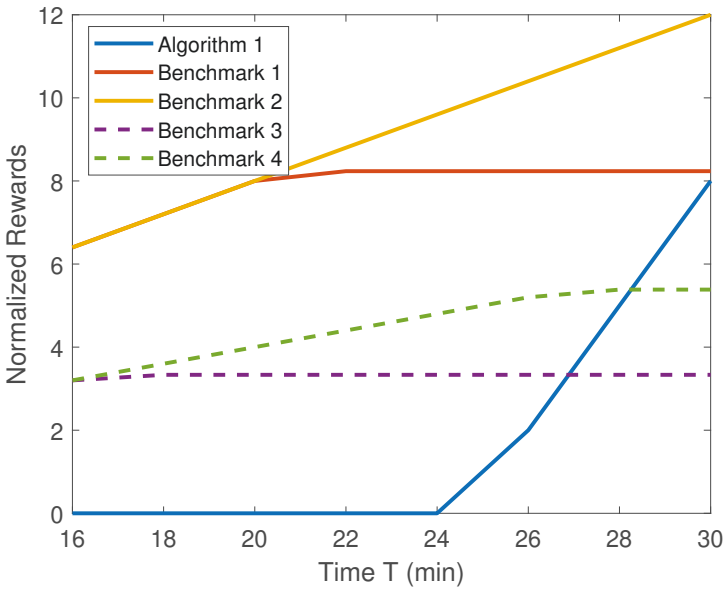


Figure 4.6 : The normalized in-situ sensing rewards for the flight distance $M = 8.4$ km, as the allowed flight time T increases.

POI street blocks, at the cost of timely completion of the goods delivery. As a consequence, Benchmark 2 is the first to be restrained by the delivery deadline as the required flight distance increases, resulting in the shortest flight distance of 7.9 km among all schemes. In contrast, Algorithm 1 dynamically chooses the actions of the delivery UAV by assessing the remaining distance and elapsed time, while ensuring the delivery is in time with the minimum energy consumption. Fig. 4.4 demonstrates that the proposed Algorithm 1 can save 18% and 24% of the energy of the UAV, as compared to Benchmarks 1 and 3.

Figs. 4.5 and 4.6 plot the flight completion rate and the normalized rewards over the flight distance $M = 8.4$ km, respectively, as T grows. It is observed in Fig. 4.5 that only Algorithm 1 can accomplish the flight within 18 min. Being aware of the delivery deadline, Algorithm 1 dynamically chooses actions for the UAV in accordance with the remaining distance and elapsed time, and the UAV performs in-situ sensing when it does not compromise the timely delivery. As shown

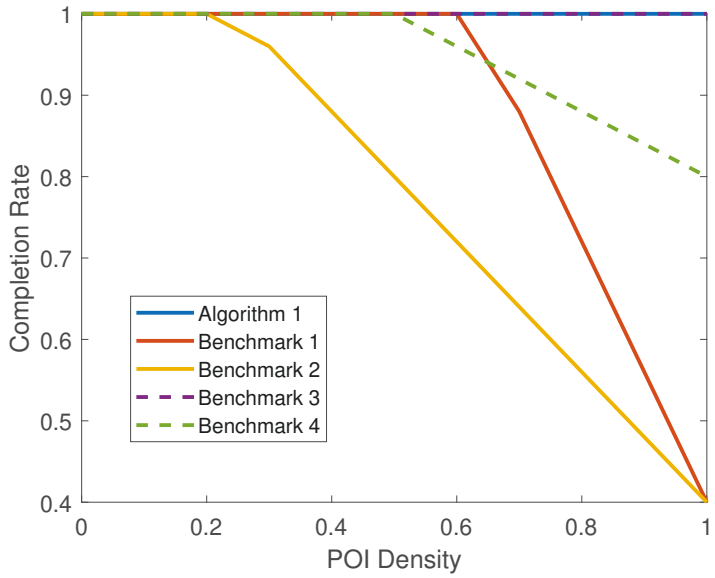


Figure 4.7 : The timely flight completion ratio under different POI densities $\rho(\mathcal{L}^{(1)})$

in Fig. 4.6, the proposed Algorithm 1 only performs in-situ sensing when $T > 24$ min. Benchmarks 1 and 2 perform sensing for rewards whenever possible. While more reward is returned, the arrival is delayed. Here the rewards are normalized over the unit sensing reward μ. On the other hand, Benchmarks 1 and 3 require the UAV to fly persistently at the full speed. This leads to a higher completion rate in exchange for energy consumption, as shown in Fig. 4.4.

Fig. 4.7 plots the completion rate of UAV flight under different POI densities $\rho(\mathcal{L}^{(1)})$, where the required flight distance is $M = 8.4$ km and the allowed flight time is $T = 28$ min. We see that the proposed Algorithm 1 can dynamically choose to perform in-situ sensing and hence complete the flight within the allowed delay T. The flight completion rates of Benchmarks 1 and 2 drop rapidly, as the POI density grows. This is because performing in-situ sensing takes time and may prevent the timely delivery.

Fig. 4.8 evaluates the normalized sensing rewards of the UAV obtained under different flight distances, with the growth of the allowed flight time T. The longer

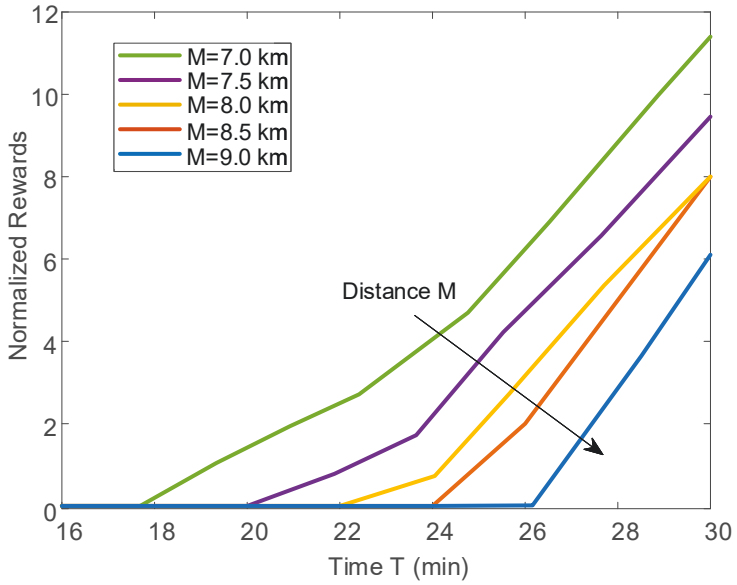


Figure 4.8 : The normalized in-situ sensing rewards under different flight distances, as the allowed flight time T increases.

flight time T grants the UAV more flexibility to choose in-situ sensing. With the same T, a shorter flight distance allows the UAV to obtain a higher reward through performing in-situ sensing. In the proposed Algorithm 1, the UAV starts to perform in-situ sensing, when the allowed flight time is $T > 18$ min under the flight distance $M = 7$ km. No in-situ sensing is performed when the allowed flight time is shorter than $T = 26$ min under the flight distance $M = 9$ km.

4.6 Conclusion

In this chapter, we established a new multi-task UAV control framework, where the UAV can dynamically choose its action from different speeds or in-situ sensing. We reveal the optimal control framework is a monotone deterministic Markovian policy with a simple threshold-based structure of the optimal policy. An optimal switch of the UAV's flight speed or in-situ sensing is only activated when the thresholds concerning the flight distance or time is met. By comparing its state with the

thresholds, the UAV can decide its action optimally in response to in-situ sensing requests arising. Extensive simulations demonstrated that the proposed algorithm can guarantee the UAV's timely delivery, minimize its energy consumption, and maximize its sensing reward, as validated by comparison with the computationally expensive, DP-based alternative; and substantially reduce the energy consumption of the UAV, as compared to existing approaches.

Chapter 5

Optimal Routing of Multi-Task UAV

5.1 Introduction

UAVs are increasingly widely employed to provide fast and cost-effective last-mile goods delivery, thanks to their excellent flexibility and maneuverability [171–173]. Equipped with many embedded sensors, e.g., cameras, microphones and thermometers, UAVs have also been widely considered to provide aerial sensing functionalities [174]. There is a great opportunity of integrating sensing capabilities into delivery UAVs. A delivery UAV can also carry out in-situ sensing along its delivery route, as illustrated in Fig. 5.1.

In this chapter, we optimize the flight route for UAV-based goods delivery and in-situ sensing. Goods delivery and in-situ sensing often have different commitment and timing requirements [175]. In particular, goods delivery has stringent deadlines, while in-situ sensing can be only performed when it is possible and does not violate the delay requirement of the goods delivery. A sensing task is often time-sensitive and location-dependent [176, 177]. On the other hand, it is important for the drone to utilize the battery energy efficiently while completing the delivery and sensing tasks, since commercially available drones are primarily powered by onboard batteries [178]. In this sense, it is crucial to meticulously design the UAV flight routes by taking the delivery requirement, in-situ sensing, and energy consumption into consideration.

To the best of our knowledge, UAV routing for joint goods delivery and in-situ sensing has yet to be studied in the literature. Existing studies have been on either

delivery route planning or task selection. The UAV delivery routing problem was characterized as finding the optimal route for the UAV to deliver goods to the destination [21]. In [9], the navigation of a delivery drone was formulated as a generalized vehicle routing problem (VRP), which was a mixed-integer linear programming (MILP). The simulated annealing algorithm was applied to find the optimal route that minimizes both the delivery time and energy consumption. Francesco *et al.* [45] considered the drone delivery problem under varying wind conditions. The routing was carried out by the shortest path algorithm to maximize the percentage of delivery completion under a given battery budget. A synchronized truck and drone routing framework was developed in [24] and [26]. By jointly planning the routes of a ground truck and UAV for goods delivery, timely deliveries were guaranteed with the minimum travel costs. However, these existing results [9, 21, 24, 26, 45–50] cannot apply to the considered routing problem for joint goods delivery and in-situ sensing, because of the stringent delay requirement of goods delivery, as well as the time sensitivity and location dependence of in-situ sensing.

This chapter presents a new and efficient approach to optimizing the route of a UAV for joint goods delivery and in-situ sensing. In-situ sensing is time-sensitive and location-dependent. A new routing problem is formulated to jointly ensure timely goods delivery, minimize the UAV's energy consumption, and maximize its reward for in-situ sensing conducted. To solve the problem, we first interpret the flight waypoints as location-dependent tasks, hence accommodating the waypoints and in-situ sensing in a unified process of task selection. Then, we construct a weighted time-task graph in which each vertex captures both the location and time of a task. Accordingly, the feasible routes of the UAV can be interpreted as a series of feasible vertexes which do not violate the delivery deadline and the UAV's speed limit in the graph. By distributing the rewards of sensing tasks to the edges of the graph, we finally transform the routing of the UAV to a weighted routing problem in

the time-task graph and optimally solve the problem using the celebrated Bellman-Ford algorithm. As validated by simulations, the new approach can outperform its potential alternatives by at least 18% in terms of task reward in a real-world case study.

The rest of this chapter is organized as follows. Section II describes the system model. Section III formulates the considered problem and elaborates on the proposed method. Numerical results are provided based on practical flight parameters in Section IV. Finally, Section V concludes the chapter.

5.2 System Model

Fig. 5.1 illustrates the proposed joint aerial sensing and delivery system, where a UAV delivers goods to the destination and performs in-situ sensing on the way. The UAV is expected to reach the destination within a specified delay T. The UAV flies along roads to avoid buildings and foliage and can conduct in-situ sensing tasks at the road intersections. The UAV gets rewarded for the in-situ sensing tasks conducted. We consider a time-sensitive and location-dependent sensing task model [176]. To incentivize the UAV to perform the sensing tasks, the UAV is offered a (different) reward for each sensing task. Given the time sensitivity and location dependence of the in-situ sensing tasks, the UAV performs different tasks and gets different rewards by taking different routes. To this end, the UAV's trajectory needs to be carefully designed to maximize the rewards, minimize the energy consumption of the UAV, and ensure its timely goods delivery.

We discretize T into time slots with slot duration $\Delta\tau = 1$. Let $\mathcal{T} = \{1, 2 \cdots, T\}$ collect the indexes to the time slots. The UAV needs to arrive at the destination by no later than $t = T$ for timely goods delivery. The considered area is segmented into street blocks, and indexes to the to intersections of the street blocks are collected in the set $\mathcal{B} = \{1, \cdots, B\}$, where B is the total number of intersections. The initial

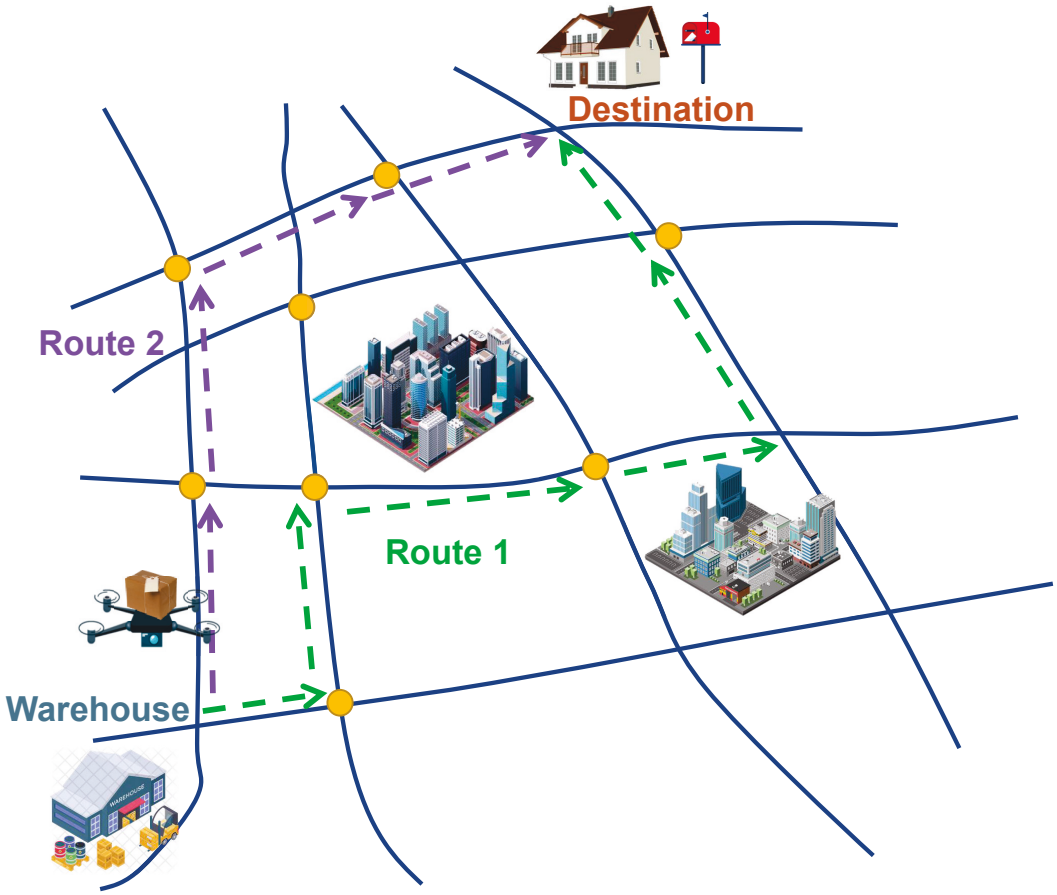


Figure 5.1 : Illustration of joint aerial in-situ sensing and delivery framework: A UAV is dispatched to deliver goods from the warehouse to the destination, and can also carry out in-situ sensing at the road intersections when required (yellow circle).

position and the delivery destination are b_I and b_D, respectively. b_I, $b_D \in \mathcal{B}$.

The total number of tasks is N. The index to the sensing tasks is $n \in \mathcal{N} = \{1, \cdots, N\}$. Each task n is featured by its location $b[n]$, starting time $t[n]$, and task reward $\mu[n]$. The duration of each task is $\Delta\tau$.

- Task location $b[n] \in \mathcal{B}$: The place (road intersection) of interest, where the sensing task is needed.

- Starting time of task $t[n] \in \mathcal{T}$: We assume that each task cannot be executed

before time $t[n]$, and it takes one time slot to execute the task.

- Task reward $\mu[n]$: The reward $\mu[n] > 0$.

The spatio-temporal map of requested in-situ sensing tasks is described as a function of location and time:

$$\Lambda(b,t) = \begin{cases} 1, & \text{if } t = t[n] \text{ and } b = b[n], \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

We assume that the UAV flies from road intersection b to intersection b' at a speed of v_f, and the power consumption is p_f. The energy consumption of the UAV moving from road intersection b to b' is given by

$$E_f[b,b'] = \tau[b,b']p_f, \tag{5.2}$$

where $\tau[b,b'] = l[b,b']/v_f$ is the flight time from road intersection b to b', and $l[b,b']$ is the distance between intersection b and b'. When the UAV performs a sensing task, the UAV consumes power to hover over the area at a speed of v_h and collect sensing data. The propulsion power consumption of hovering is $p_h = p(v_h)$. The power consumption for sensing is p_s. The energy consumption of the UAV performing the n-th sensing task is given by

$$E_h[n] = (p_s + p_h)\Delta\tau = p_s + p_h. \tag{5.3}$$

The propulsion power for hovering is much lower than that for moving, i.e., $p_h < p_f$. The energy consumption of the in-situ sensing is usually much smaller than that of the UAV's propulsion [19].

5.3 Joint Aerial Sensing and Delivery Routing

In this section, we optimize UAV's route to minimize the UAV's energy consumption, maximize its sensing reward, and ensure its timely arrival at the destination.

We convert this joint aerial sensing and delivery routing problem to a task-time routing problem, where the task-time route refers to a series of differently located tasks selected over time. By interpreting the UAV's flight as virtual tasks, we reveal that the task-time route of the UAV has a unified graph representation. We derive the feasible task-time route conditions, under which a timely arrival of the UAV at the destination is guaranteed. The problem of interest is thus converted to search the set of feasible task-time routes for the route that minimizes the UAV's energy consumption and maximizes its total sensing reward. By aggregating the rewards and the costs as the edge weight of the unified graph, the task-time routing turns to be a weighted routing problem which can be optimally solved using the celebrated Bellman-Ford algorithm.

First, we unify the UAV's trajectory planning and in-situ sensing task selection by defining the UAV's movement as a set of virtual tasks $\mathcal{N}^{\text{vir}}$. The virtual task set should satisfy the condition that the UAV departs from the initial location b_I at starting time $t = 0$ and arrives at the destination location b_D by time $t = T$. In other words, the set of virtual tasks should contain the initial task n_I and the final task n_D, i.e., $\{n_I, n_D\} \subseteq \mathcal{N}^{\text{vir}}$, where the locations and times of the initial task n_I and the final task n_D satisfy $b[n_I] = b_I$, $t[n_I] = 0$, $b[n_D] = b_D$, and $t[n_D] = T$. The reward for any virtual task n is zero, i.e., $\mu[n] = 0$ for $n \in \mathcal{N}^{\text{vir}}$. By this means, the routing and the in-situ sensing decisions are unified in one task set, as given by

$$\mathcal{S} = \mathcal{N} \cup \mathcal{N}^{\text{vir}}. \tag{5.4}$$

As a result, each task can be represented by a task-time point (s, t), $\forall s \in \mathcal{S}$, $t \in \mathcal{T}$, as illustrated in Fig. 2, where the solid circles denote the sensing tasks and the hollow circles denote the virtual tasks. There is a total of $K = |\mathcal{S}| = BT$ tasks in the unified task set. $|\cdot|$ stands for cardinality.

We define the step cost of switching from task s to task s' as $c[s, s']$, which

incorporates both the energy consumption of the UAV's flight and sensing, as given by

$$c[s, s'] = \begin{cases} E_f[s, s'] + E_h[s'], & \text{if } s \neq s', \\ 0, & \text{otherwise.} \end{cases} \tag{5.5}$$

where $E_f[s, s'] = E_f\big[b[s], b[s']\big]$ is the energy consumption for the UAV to fly from the location of task s to the location of task s'. No cost occurs when $s = s'$, i.e., $c[s, s] = 0$. According to (5.2), the energy consumption $E_f[s, s']$ depends linearly on the shortest flight distance $l\big[b[s], b[s']\big]$ from the location of task s to the location of task s', i.e., from $b[s]$ to $b[s']$. The route may not be unique, as the UAV can take different paths between the road intersections $b[s]$ and $b[s']$. In this case, we can apply the Dijkstra's algorithm [179] to find the shortest flight route between task s and task s'.

Next, we present the UAV's task-time route with an acyclic directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ denote the sets of vertexes and edges, respectively. $\mathcal{V}$ collects all the task-time points, as given by

$$\mathcal{V} = \{(s, t) : s \in \mathcal{S}, t \in \mathcal{T}\}. \tag{5.6}$$

Each edge links two different task-time points. A set of directionally connected, acyclic edges indicate a possible route for the UAV, i.e.,

$$\mathcal{E}(\boldsymbol{r}) = \Big\{\big((s^k, t^k), (s^{k+1}, t^{k+1})\big) : k = 1, \ldots, K-1\Big\}, \tag{5.7}$$

where $\boldsymbol{r}$ denotes a UAV potential route, which connects a series of task-time points visited sequentially by the UAV, k indicates the k-th visited task/location, and K is the total number of tasks completed, $t^K \leq T$.

Let $\mathcal{V}(\boldsymbol{r})$ and $\mathcal{E}(\boldsymbol{r})$ denote the set of vertexes and edges traversed by a potential route $\boldsymbol{r}$. Each vertex is associated with an expected reward $\mu[(s, t)]$. Each edge is associated with a cost $c[s, s']$, as defined in (5.5). The considered routing problem

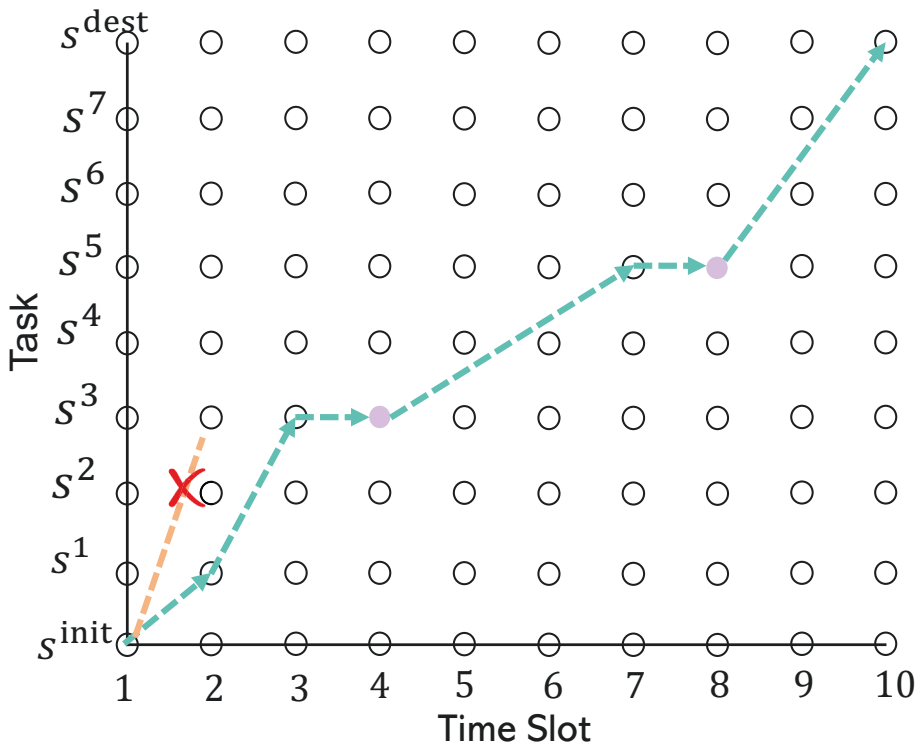


Figure 5.2 : Illustration of a feasible task-time route and an infeasible task-time route, where the solid circles denote sensing tasks and hollow circles represent virtual tasks (i.e., the UAV's flight).

can be formulated to maximize the net gain of the UAV, i.e., the total reward subtracted by the total cost, as given by

$$\max_{\boldsymbol{r}\in\mathcal{R}} U(\boldsymbol{r}) = \sum_{(s,t)\in\mathcal{V}(\boldsymbol{r})} \mu[(s,t)] - \sum_{((s,t),(s',t'))\in\mathcal{E}(\boldsymbol{r})} c[(s,t),(s',t')] \tag{5.8a}$$

$$\text{s.t.} \quad b\left[s^1\right] = b_I,\ b\left[s^K\right] = b_D, \tag{5.8b}$$

$$0 = t^1 < t^2 < \cdots < t^K \leq T, \tag{5.8c}$$

$$t^{k+1} - t^k = \tau\left[b\left[s^k\right], b\left[s^{k+1}\right]\right]. \tag{5.8d}$$

Here, all the possible routes are collected in set $\mathcal{R}$. Constraint (5.8b) specifies the starting point and the destination of the UAV. Constraint (5.8c) specifies the causality, i.e., the time elapses with the expansion of the route, and $t^K \leq T$ also

ensures the timely arrival of UAV at destination. Constraint (5.8d) represents the UAV mobility constraint: the time interval between the UAV performing task s^k and performing task s^{k+1} should allow the UAV to fly between the locations of two tasks.

Because of the constraints, the standard shortest path algorithm, such as the Dijkstra's algorithm and Bellman-Ford algorithm, cannot readily solve problem (5.8). Moreover, with the increases of the time T and the number of road intersections B, the task space and hence the scale of graph can be large, resulting in a prohibitive computational complexity. Nevertheless, we find that constraints (5.8c) and (5.8d) can be readily accounted for by assessing the feasibility of the individual edges in graph $\mathcal{G}$. Specifically, the graph $\mathcal{G}$ is a directed graph. A feasible directed edge exists from vertex (s,t) to vertex (s',t') with $t<t'$, not the other way around, due to the causality in (5.8c). The feasible edge between the two tasks s and s' must satisfies $t[s']-t[s]=l[b[s'],b[s]]/v_f$, and the weight of the edge is $c[s,s']$, according to the mobility constraint (5.8d). As illustrated in Fig. 5.2, the route from s_I to s^3 is not feasible as the UAV cannot reach the site of s^3 from s_I within a slot.

By connecting the feasible edges and disconnecting any infeasible edges which would violate causality and mobility, any possible routes from b_I to b_D in the resulting graph $\mathcal{G}$ inherently satisfy constraints (5.8c) and (5.8d). Problem (5.8) becomes selecting one of the feasible routes that maximizes the objective (5.8a), i.e.,

$$\begin{aligned}\max_{\boldsymbol{r}\in\mathcal{R}} U(\boldsymbol{r}) = &\sum_{(s,t)\in\mathcal{V}(\boldsymbol{r})} \mu[(s,t)] - \sum_{((s,t),(s',t'))\in\mathcal{E}(\boldsymbol{r})} c[(s,t),(s',t')] \\ \text{s.t.} \quad & b\left[s^1\right]=b_I,\ b\left[s^K\right]=b_D.\end{aligned} \tag{5.9}$$

To allow the standard shortest path algorithm [180] to apply, we further update the graph $\mathcal{G}$ by distributing the reward of each vertex s' into the cost at any feasible edge $[s,s']$, $\forall s\in\mathcal{S}$, as illustrated in Fig. 5.3. This transformation is equivalent

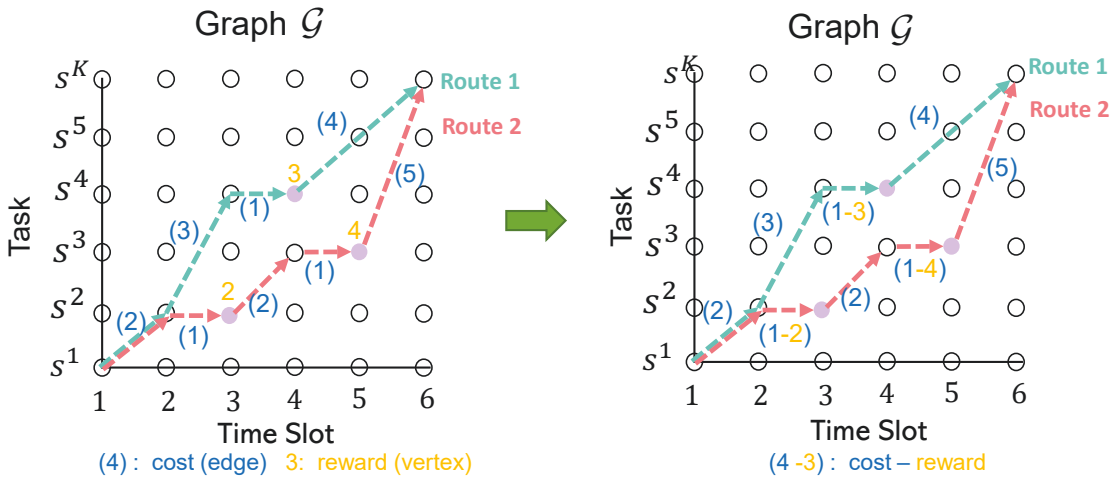


Figure 5.3 : Graph Updating: graph $\mathcal{G}$ is updated by replicating and distributing the reward of each vertex s' to the weights of any edge $[s, s']$, $\forall s \in \mathcal{S}$.

because each of the (directed) edges points to a single vertex. The reward of a vertex can be replicated and distributed to every edge pointing to the vertex. As a result, the weight of an edge of $\mathcal{G}$ provides a unified cost, as given by

$$h[(s,t),(s',t')] = c[(s,t),(s',t')] - \mu[(s,t)]. \tag{5.10}$$

Finally, problem (5.8) can be rewritten to find the route with the minimum sum edge weight in graph $\mathcal{G}^*$, as given by

$$\begin{aligned} \boldsymbol{r}^* = \arg\min_{\boldsymbol{r}} & \sum_{((s,t),(s',t'))\in\mathcal{E}(\boldsymbol{r})} h[(s,t),(s',t')] \\ \text{s.t.} \quad & b\left[s^1\right] = b_I,\ b\left[s^K\right] = b_D. \end{aligned} \tag{5.11}$$

which can be readily solved using the Bellman-Ford Algorithm. The proposed algorithm is summarized in Algorithm 6.

Algorithm 6 Joint Delivery and In-situ Sensing Routing

Initialization:

1: Obtain the initial location of the UAV (depot) b_I, the location of delivery b_D, and the time requirement T.

2: Sensing task collection: Collect the task descriptions from the sensing platform, including the location $b[n]$, time $t[n]$, and the reward $\mu[n]$ of the tasks, $\forall n \in \mathcal{N}$, $t \in \mathcal{T}$.

Route Planning:

3: Calculate the shortest flight route between the locations of any two tasks, e.g., by applying the Dijkstra's algorithm.

4: Construct the graph $\mathcal{G}^*$, including the vertex set $\mathcal{V}$ with (5.6) and the edge set $\mathcal{E}$ with (5.7), and preclude any edges violating constraints (5.8c) and (5.8d).

5: Compute the edge weights $h[(s,t),(s',t')]$, with (5.10)

6: Run the Bellman-Ford Algorithm in graph $\mathcal{G}$ from source to destination

$$\boldsymbol{r}^* = \arg\min_{\boldsymbol{r}} \sum_{((s,t),(s',t'))\in\mathcal{E}(\boldsymbol{r})} h[(s,t),(s',t')]$$

7: Output the optimal route $\boldsymbol{r}^*$.

5.4 Numerical Results

In this section, we evaluate the proposed routing method in comparison with two benchmark schemes. We also present a case study on a practical application with a realistic map and goods delivery requests.

5.4.1 Performance Comparison

We consider practical octocopter flight parameters from DJI Agras MG-1 [154], where the UAV takeoff weight is 19.8 kg. The flying speed is $v_f = 10$ m/s, and the hover speed is $v_h = 1$ m/s. The power consumption of the UAV flying at the

speed of $v_f = 10$ m/s is $p_f = 5700$ W. The total power consumption of hovering and performing in-situ sensing is $p_h + p_s = 2350$ W. The duration of a time slot is $\Delta\tau = 1$ min. In the simulation, we consider an area of 2.4 km $\times$ 2.4 km. As shown in Fig. **??**, there are nine street blocks with $B = 16$ road intersections in the area, and the length of each street block is around 800 m. The tasks are randomly generated and distributed at road intersections with probability $\rho = 0.4$. The reward of in-situ sensing tasks is in the form of energy (Wh). Three levels of reward are considered, i.e., $\mu = [145, 165, 195]$ Wh, which are equally likely assigned to all tasks in the simulations. We repeat the simulations for 1000 times with randomly generated UAV's delivery destination each time. At each time, the locations and time of in-situ sensing tasks are also randomly generated.

For comparison purposes, we also simulate the following two benchmarks based on [45]:

- On-the-spot sensing (OTS) scheme: The UAV flies directly to the delivery destination with the shortest path and performs sensing tasks the in-situ sensing tasks encountered along the shortest path.

- Greedy sensing routing (GSR) scheme: The UAV minimizes the cost for each time slot. More specifically, for each time slot, the GSR selects the next optimal intersection that minimizes the cost using a greedy method. In other words, the UAV finds the task with the highest reward. The problem of the GSR scheme is given by

$$\min_{s'} h[(s,t),(s',t')] \quad \text{s.t.} \quad (s',t') \in \mathcal{E}(\boldsymbol{r}) \tag{5.12}$$

First, we compare the averaged rewards offered by different routing schemes in Fig 5.4. The averaged rewards are obtained under different delivery deadlines through 1000 independent tests with randomized depot, destination, and in-situ

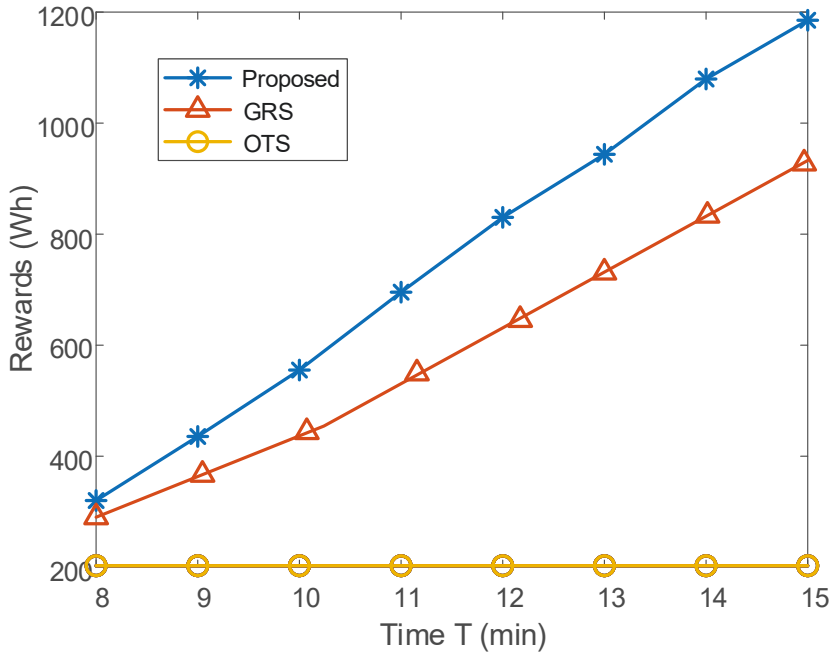


Figure 5.4 : Averaged rewards under different delivery deadline T when the delivery distance (straight line distance on map) is 3 km.

sensing tasks. The delivery distance (straight line distance on map) is set as 3 km. As shown in Fig 5.4, the rewards increase with the delivery deadline. The UAV can perform more sensing tasks and obtain more rewards with a longer allowed delivery time. The proposed algorithm can considerably outperform GSR, since the latter only considers the UAV's route one time slot ahead and finds the task that maximizes the reward of the slot. In contrast, the proposed algorithm can choose the route that maximizes its sensing rewards and minimizes its energy consumption cost, and thus achieve higher averaged rewards than the GSR scheme. The OTS scheme only allows the UAV to perform the sensing tasks opportunistically. The sensing tasks are only taken from those which are along the straight flight route and happen to be required at the time when the UAV passes by. In this way, the rewards of the OTS scheme can barely increase even given a later delivery deadline.

The energy consumption of different routing schemes is compared in Fig 5.5,

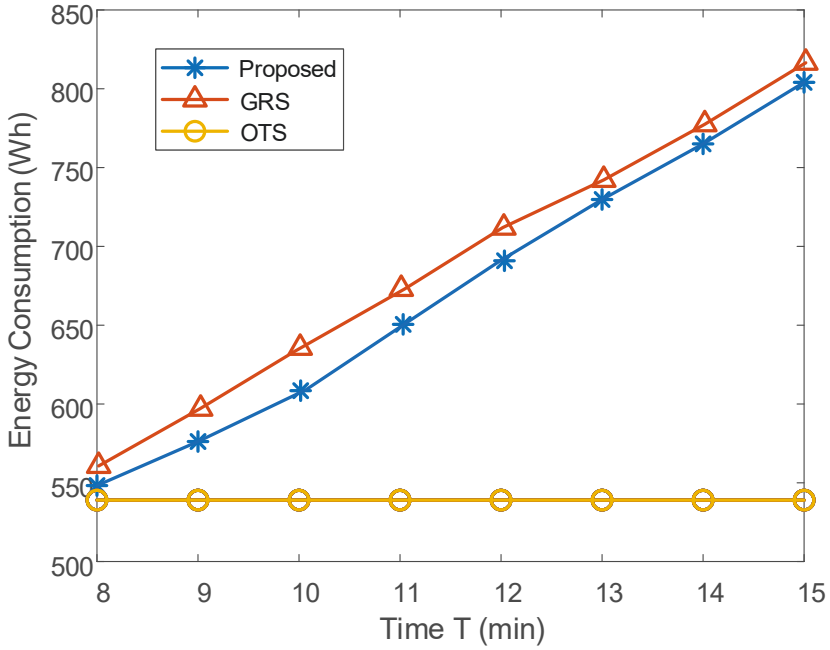


Figure 5.5 : Energy consumption comparison under different delivery deadline T when the delivery distance (straight line distance on map) is 3 km.

where the UAV delivers goods to a randomized destination 3 km away from the depot. Every result is the average of over 1000 independent tests. As shown in the figure, the energy consumption of both the proposed and GSR schemes increase with the delivery deadline T. With a longer delivery time, the UAV is allowed to perform more sensing tasks and increases energy consumption. Nevertheless, the proposed algorithm has a lower energy consumption than the GRS scheme. The GSR scheme has a higher energy consumption because the UAV deviates to achieve the highest reward for the next time slot. The OTS scheme guides the UAV directly to the destination, but it is at the cost of in-situ sensing opportunities and subsequently sensing rewards.

Fig. 5.6 plots the net gains of the considered schemes, as defined in (5.8), where different task demand densities are taken. The task demand density here refers to the ratio of the sensing tasks and all task-time points. The net gain is evaluated when

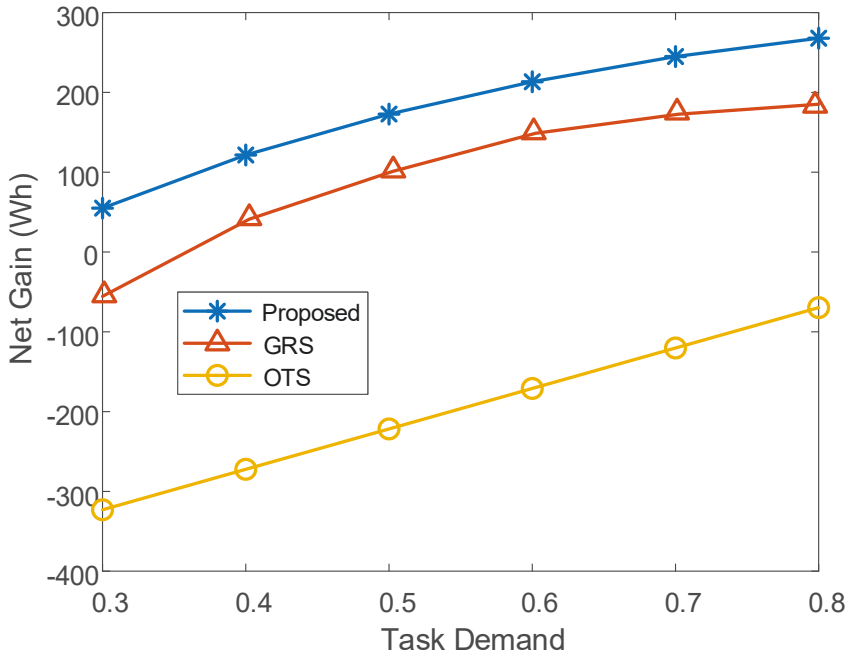


Figure 5.6 : Net gain under different task demand densities when the UAV is required to complete the 3 km delivery within 12 min.

the UAV is required to complete goods delivery over 3 km (straight line distance on map) within 12 min. It is observed that the net gains of all schemes grow with task demands. On the one hand, with large task demands, the UAV has more opportunities to perform sensing and obtain rewards. On the other hand, the required tasks are more likely to be along the shortest path. In other words, the UAV deviate less often from the shortest path to perform in-situ sensing. The energy consumption can be reduced. It is worth pointing out that a negative net gain indicates lower rewards obtained than the energy consumption of the flight; see (5.8).

5.4.2 Case Study

We study a practical application of the proposed routing and two benchmark schemes on our university campus. As shown in Fig 5.7, the UAV needs to deliver

Figure 5.7 : The map of our on-campus case study, where the blue, yellow, and orange lines indicate the proposed algorithm, OTS, and GRS, respectively.

goods to the destination within 15 min. The straight line distance between the UAV's initial position and destination is 1.5 km. In-situ tasks distributed with specified time requirements and rewards, as shown in Fig 5.7. It is found that the proposed algorithm achieves the maximum reward of 320 Wh, while the rewards of the OTS and GRS are 120 Wh and 270 Wh, respectively. The proposed algorithm can outperform OTS and GRS by 166% and 18%, respectively.

As illustrated in Fig 5.7, the GRS scheme finds the highest-reward nearsightedly for only one time slot ahead, and chooses the task-A instead of task-B. However, task-B and task-C are close in distance and can be performed sequentially with little extra energy consumption. Consequently, task-B and task-C can collectively provide higher reward and demand lower energy consumption than task-A alone. On the other hand, the flight distances of the proposed algorithm, OTS, and GRS

are 1.59 km, 1.50 km, and 1.79 km. The OTS scheme gives the straight line distance between the initial position and destination. The proposed algorithm only increases the 90 m flight distance than the OTS, while achieving more rewards. The proposed algorithm can achieve 166% task reward of the OTS with only 6% increase in flight distance. In summary, the proposed routing algorithm enables the UAV to perform more sensing tasks along the delivery route than its existing alternatives, while guaranteeing the time arrival of the delivery.

5.5 Conclusion

In this chapter, we developed a new routing framework for UAV-based joint goods delivery and in-situ sensing, which allows the UAV to minimize its energy consumption, maximize its sensing reward, and ensure its timely arrival at the destination. This was achieved by transforming the UAV routing of joint goods delivery and sensing to a weighted routing problem in a time-task graph, and efficiently solving the problem using the Bellman-Ford algorithm. Extensive simulations demonstrated that the proposed routing algorithm enables the UAV to perform more sensing tasks along the delivery route than its existing alternatives, while guaranteeing its timely goods delivery. Validated by a real-world case study, the new approach can outperform its alternatives by at least 18% in terms of task reward.

Chapter 6

Conclusion and Future Work

In summary, this thesis studies flight planning, optimal control, and routing for the multi-task UAV under diverse service requirements. The detailed research work and contributions of this thesis are summarized as follows.

- In Chapter 3, We establish a novel UAV delivery framework, which closely integrates UAVs into the ITS to provide energy-efficient, delay-sensitive goods delivery services. We minimized the energy consumption of the UAV and ensured its timely arrival at the destination, by optimizing its actions at different points of its planned route, including its flight mode (i.e., full speed or cruise speed), hitchhiking (and recharging on collaborative ground vehicles), or stopping and recharging at roadside charging stations. A new DP-based algorithm was first developed to obtain the optimal action of the UAV. Then, we revealed that the optimal action of the UAV only changes when the remaining flight distance or the elapsed time meets certain conditions. Accordingly, thresholds were derived for the distance and time. By comparing the remaining flight distance and the elapsed time with the thresholds, the optimal actions of the UAV can be instantly made. Simulations show that the proposed algorithms can improve the flight distance by 48%, as compared with existing alternatives.

- In Chapter 4, we established a new multi-task UAV control framework, where the UAV can dynamically choose its action from different speeds or in-situ sensing. We reveal the optimal control framework is a monotone deterministic Markovian policy with a simple threshold-based structure of the optimal

policy. An optimal switch of the UAV's flight speed or in-situ sensing is only activated when the thresholds concerning the flight distance or time is met. By comparing its state with the thresholds, the UAV can decide its action optimally in response to in-situ sensing requests arising. Extensive simulations demonstrated that the proposed algorithm can guarantee the UAV's timely delivery, minimize its energy consumption, and maximize its sensing reward, as validated by comparison with the computationally expensive, DP-based alternative; and

- In Chapter 5, we developed a new routing framework for UAV-based joint goods delivery and in-situ sensing, which allows the UAV to minimize its energy consumption, maximize its sensing reward, and ensure its timely arrival at the destination. This was achieved by transforming the UAV routing of joint goods delivery and sensing to a weighted routing problem in a time-task graph, and efficiently solving the problem using the Bellman-Ford algorithm. Extensive simulations demonstrated that the proposed routing algorithm enables the UAV to perform more sensing tasks along the delivery route than its existing alternatives, while guaranteeing its timely goods delivery.

The theoretical results obtained in this thesis provide new insights into the intelligent applications of UAVs in the future. The optimal scheduling and flight planning of the multi-task UAV in this thesis try to address the challenges in the service integration. However, due to time constraints, the relevant research the content involved in this thesis has certain limitations. In the future, the follow-up research can be carried out from the following aspects:

1. The UAV control and scheduling is a highly complex task. This thesis targets the upper layer control of the UAV flight. It would be interesting to incorporate the attitude and articulation control and aircraft flight dynamics into

consideration, such as altitude, airspeed, and dive angle control, to refine and improve the performance.

2. As an important part of non-terrestrial networks in the 6G, the UAV complement terrestrial connectivity in critical scenarios by leveraging its agility and versatility. For these real-time applications, it would be interesting to jointly optimize the task scheduling and flight planning of the UAV by applying online learning techniques.

Bibliography

[1] H. Wang, H. Zhao, J. Zhang, D. Ma, J. Li, and J. Wei, "Survey on unmanned aerial vehicle networks: A cyber physical system perspective," *IEEE Commun. Surv. Tutor.*, vol. 22, no. 2, pp. 1027–1070, 2019.

[2] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5g and beyond: Recent advances and future trends," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2241–2263, 2018.

[3] J. Hu, H. Zhang, L. Song, Z. Han, and H. V. Poor, "Reinforcement learning for a cellular internet of UAVs: Protocol design, trajectory control, and resource management," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 116–123, 2020.

[4] J. Tavčar and I. Horvath, "A review of the principles of designing smart cyber-physical systems for run-time adaptation: Learned lessons and open issues," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 49, no. 1, pp. 145–158, 2018.

[5] I. Makhdoom, I. Zhou, M. Abolhasan, J. Lipman, and W. Ni, "Privysharing: A blockchain-based framework for privacy-preserving and secure data sharing in smart cities," *Computers & Security*, vol. 88, p. 101653, 2020.

[6] D. Cavaliere, V. Loia, A. Saggese, S. Senatore, and M. Vento, "Semantically enhanced UAVs to increase the aerial scene understanding," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 49, no. 3, pp. 555–567, 2019.

[7] R. S. De Moraes and E. P. De Freitas, "Multi-UAV based crowd monitoring system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1332–1345,

2021.

[8] Z. Hu, Z. Bai, Y. Yang, Z. Zheng, K. Bian, and L. Song, "UAV aided aerial-ground IoT for air quality sensing in smart city: Architecture, technologies, and implementation," *IEEE Netw.*, vol. 33, no. 2, pp. 14–22, 2019.

[9] K. Dorling, J. Heinrichs, G. G. Messier, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 47, no. 1, pp. 70–85, 2017.

[10] A. V. Savkin and H. Huang, "Range-based reactive deployment of autonomous drones for optimal coverage in disaster areas," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 51, no. 7, pp. 4606 – 4610, 2021.

[11] H. Huang, A. V. Savkin, and C. Huang, "Decentralised autonomous navigation of a UAV network for road traffic monitoring," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 4, pp. 2558–2564, 2021.

[12] 3GPP, "Study on application layer support for unmanned aerial system (UAS)," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 23.755, 01 2021, version 17.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications /SpecificationDetails.aspx? specificationId=3588

[13] F. Qi, X. Zhu, G. Mang, M. Kadoch, and W. Li, "UAV network and IoT in the sky for future smart cities," *IEEE Netw.*, vol. 33, no. 2, pp. 96–101, 2019.

[14] Z. Sun, G. G. Yen, J. Wu, H. Ren, H. An, and J. Yang, "Mission planning for energy-efficient passive UAV radar imaging system based on substage division collaborative search," *IEEE Trans. Cybern.*, 2021 (Early access).

[15] A.-Y. Lu and G.-H. Yang, "Resilient observer-based control for cyber-physical

systems with multiple transmission channels under denial-of-service," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4796–4807, 2019.

[16] X. Li, J. Tan, A. Liu, and et al., "A novel UAV-enabled data collection scheme for intelligent transportation system through UAV speed control," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–11, 2020.

[17] A. Al-Hilo, M. Samir, C. Assi, and et al., "UAV-assisted content delivery in intelligent transportation systems-joint trajectory planning and cache management," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–13, 2020.

[18] Y. Wang, F. Xu, and et al., "Adaptive online power management for more electric aircraft with hybrid energy storage systems," *IEEE Trans. Transport. Electrific.*, vol. 6, no. 4, pp. 1780–1790, 2020.

[19] Y. Zeng, J. Xu, and R. Zhang, "Energy minimization for wireless communication with rotary-wing UAV," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2329–2345, 2019.

[20] V. Hassija and e. a. Chamola, "A distributed framework for energy trading between UAVs and charging stations for critical applications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5391–5402, 2020.

[21] R. G. Ribeiro and et al., "Unmanned aerial vehicle location routing problem with charging stations for belt conveyor inspection system in the mining industry," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 10, pp. 4186–4195, 2019.

[22] S. Hu, X. Chen, W. Ni, X. Wang, and E. Hossain, "Modeling and analysis of energy harvesting and smart grid-powered wireless communication networks: A contemporary survey," *IEEE Trans.Trans. Green Commun. Netw.*, vol. 4, no. 2, pp. 461–496, 2020.

[23] H. Huang and A. V. Savkin, "A method of optimized deployment of charging stations for drone delivery," *IEEE Trans. Transport. Electrific.*, vol. 6, no. 2, pp. 510–518, 2020.

[24] S. Sawadsitang, D. Niyato, P.-S. Tan, and P. Wang, "Joint ground and aerial package delivery services: A stochastic optimization approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2241–2254, 2018.

[25] Y. Liu, Z. Liu, J. Shi, G. Wu, and W. Pedrycz, "Two-echelon routing problem for parcel delivery by cooperated truck and drone," *IEEE Trans. Syst., Man, Cybern. Syst*, pp. 1–16, 2020.

[26] D. N. Das, R. Sewani, J. Wang, and M. K. Tiwari, "Synchronized truck and drone routing in package delivery logistics," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5772–5782, 2020.

[27] H. Huang, A. V. Savkin, and C. Huang, "Reliable path planning for drone delivery using a stochastic time-dependent public transportation network," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–10, 2020.

[28] H. Huang, A. V. Savkin, and C. Huang, "Drone routing in a time-dependent network: Toward low-cost and large-range parcel delivery," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 1526–1534, 2021.

[29] L. Zhao and A. Malikopoulos, "Enhanced mobility with connectivity and automation: A review of shared autonomous vehicle systems," *IEEE Intell. Transp. Syst. Mag.*, 2020 (Early access).

[30] J. Wan, X. Li, H.-N. Dai, A. Kusiak, M. Martínez-García, and D. Li, "Artificial-intelligence-driven customized manufacturing factory: Key technologies, applications, and challenges," *Proc. IEEE*, 2020.

[31] A. A. Khan, M. Abolhasan, W. Ni, J. Lipman, and A. Jamalipour, "A hybrid-fuzzy logic guided genetic algorithm (H-FLGA) approach for resource optimization in 5g VANET's," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6964–6974, 2019.

[32] ——, "An end-to-end (e2e) network slicing framework for 5g vehicular ad-hoc networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7103–7112, 2021.

[33] J. A. McClintock, D. Buchmueller, V. et al. J. M. Johansson, B. W. Porter, and A. J. Roths, "Systems, devices and methods delivering energy using an uncrewed autonomous vehicle," 2017, US Patent No. 9,778,653.

[34] L. Qian, S. Graham, and H. H.-T. Liu, "Guidance and control law design for a slung payload in autonomous landing: A drone delivery case study," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 4, pp. 1773–1782, 2020.

[35] H. Huang, A. V. Savkin, and W. Ni, "Navigation of a UAV team for collaborative eavesdropping on multiple ground transmitters," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10 450–10 460, 2021.

[36] K. Li, R. C. Voicu, S. S. Kanhere, W. Ni, and E. Tovar, "Energy efficient legitimate wireless surveillance of UAV communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2283–2293, 2019.

[37] H. Huang, A. V. Savkin, and W. Ni, "Decentralized navigation of a UAV team for collaborative covert eavesdropping on a group of mobile ground nodes," *IEEE Trans. Autom. Sci. Eng.*, 2022.

[38] S. K. K. Hari, S. Rathinam, S. Darbha, K. Kalyanam, S. G. Manyam, and D. Casbeer, "Optimal UAV route planning for persistent monitoring missions," *IEEE Trans. Robot.*, vol. 37, no. 2, pp. 550–566, 2020.

[39] S. Zhao, X. Wang, Z. Lin, D. Zhang, and L. Shen, "Integrating vector field approach and input-to-state stability curved path following for unmanned aerial vehicles," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 50, no. 8, pp. 2897–2904, 2018.

[40] N. Bartolini, A. Coletta, G. Maselli *et al.*, "A multi-trip task assignment for early target inspection in squads of aerial drones," *IEEE Trans. Mobile Comput.*, early access, 2020.

[41] S. Kim and I. Moon, "Traveling salesman problem with a drone station," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 49, no. 1, pp. 42–52, 2019.

[42] Y. Emami, B. Wei, K. Li, W. Ni, and E. Tovar, "Joint communication scheduling and velocity control in multi-UAV-assisted sensor networks: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10 986–10 998, 2021.

[43] C. H. Liu, Z. Dai, Y. Zhao, J. Crowcroft, D. Wu, and K. K. Leung, "Distributed and energy-efficient mobile crowdsensing with charging stations by deep reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 20, no. 1, pp. 130–146, 2021.

[44] L. Zhang, A. Celik, S. Dang, and B. Shihada, "Energy-efficient trajectory optimization for UAV-assisted IoT networks," *IEEE Trans. Mobile Comput.*, 2021 (Early access).

[45] B. Sorbelli *et al.*, "Energy-constrained delivery of goods with drones under varying wind conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 6048–6060, 2021.

[46] K. Li, W. Ni, and F. Dressler, "Continuous maneuver control and data capture scheduling of autonomous drone in wireless sensor networks," *IEEE Trans.*

Mobile Comput., early access,2021.

[47] K. Li *et al.*, "Online velocity control and data capture of drones for the internet of things: An onboard deep reinforcement learning approach," *IEEE Veh. Technol. Mag.*, vol. 16, no. 1, pp. 49–56, 2021.

[48] ——, "Joint flight cruise control and data collection in UAV-aided internet of things: An onboard deep reinforcement learning approach," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9787–9799, 2021.

[49] Y. Tang *et al.*, "Trajectory design for UAV assisted wireless networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2019, pp. 1–6.

[50] K. Li *et al.*, "On-board deep Q-network for UAV-assisted online power transfer and data collection," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 12 215–12 226, 2019.

[51] S. Hu, W. Ni, X. Wang, A. Jamalipour, and D. Ta, "Joint optimization of trajectory, propulsion, and thrust powers for covert UAV-on-UAV video tracking and surveillance," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1959–1972, 2020.

[52] C. Sun, W. Ni, and X. Wang, "Joint computation offloading and trajectory planning for uav-assisted edge computing," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5343–5358, 2021.

[53] K. Li, W. Ni, X. Wang, R. P. Liu, S. S. Kanhere, and S. Jha, "Energy-efficient cooperative relaying for unmanned aerial vehicles," *IEEE Trans. Mobile Comput.*, vol. 15, no. 6, pp. 1377–1386, 2015.

[54] G. Liu, "Data collection in mi-assisted wireless powered underground sensor networks: Directions, recent advances, and challenges," *IEEE Commun. Mag.*, vol. 59, no. 4, pp. 132–138, 2021.

[55] C. Di, F. Li, and S. Li, "Sensor deployment for wireless sensor networks: A conjugate learning automata-based energy-efficient approach," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 80–87, 2020.

[56] G. Han, X. Yang, L. Liu, S. Chan, and W. Zhang, "A coverage-aware hierarchical charging algorithm in wireless rechargeable sensor networks," *IEEE Netw.*, vol. 33, no. 4, pp. 201–207, 2019.

[57] F. Al-Turjman and A. Radwan, "Data delivery in wireless multimedia sensor networks: Challenging and defying in the IoT era," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 126–131, 2017.

[58] F. Xiao, X. Xie, Z. Li, Q. Deng, A. Liu, and L. Sun, "Wireless network optimization via physical layer information for smart cities," *IEEE Netw.*, vol. 32, no. 4, pp. 88–93, 2018.

[59] H. Huang, H. Yin, G. Min, X. Zhang, W. Zhu, and Y. Wu, "Coordinate-assisted routing approach to bypass routing holes in wireless sensor networks," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 180–185, 2017.

[60] K. W. Choi, L. Ginting, A. A. Aziz, D. Setiawan, J. H. Park, S. I. Hwang, D. S. Kang, M. Y. Chung, and D. I. Kim, "Toward realization of long-range wireless-powered sensor networks," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 184–192, 2019.

[61] D. Wu, H. Jing, H. Wang, C. Wang, and R. Wang, "A hierarchical packet forwarding mechanism for energy harvesting wireless sensor networks," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 92–98, Aug. 2015.

[62] J. Zhang, X. Hu, Z. Ning, E. C.-H. Ngai, L. Zhou, J. Wei, J. Cheng, and B. Hu, "Energy-latency tradeoff for energy-aware offloading in mobile edge

computing networks," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2633–2645, 2017.

[63] C. Zhan, Y. Zeng, and R. Zhang, "Energy-efficient data collection in UAV enabled wireless sensor network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 328–331, Jun. 2018.

[64] J. Zhang, Q. Cui, X. Zhang, W. Ni, X. Lyu, M. Pan, and X. Tao, "Online optimization of energy-efficient user association and workload offloading for mobile edge computing," *IEEE Trans. Veh. Technol.*, 2021.

[65] N. Gharaei, K. A. Bakar, S. Z. M. Hashim, A. H. Pourasl, and S. A. Butt, "Collaborative mobile sink sojourn time optimization scheme for cluster-based wireless sensor networks," *IEEE Sensors Journal*, vol. 18, no. 16, pp. 6669–6676, Jun. 2018.

[66] J. Ueyama, H. Freitas, B. S. Faical, G. P. Filho, P. Fini, G. Pessin, P. H. Gomes, and L. A. Villas, "Exploiting the use of unmanned aerial vehicles to provide resilience in wireless sensor networks," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 81–87, 2014.

[67] A. T. Erman, L. v. Hoesel, P. Havinga, and J. Wu, "Enabling mobility in heterogeneous wireless sensor networks cooperating with UAVs for mission-critical management," *IEEE Wireless Commun.*, vol. 15, no. 6, pp. 38–46, 2008.

[68] R. Ma, R. Wang, G. Liu, H.-H. Chen, and Z. Qin, "UAV-assisted data collection for ocean monitoring networks," *IEEE Netw.*, vol. 34, no. 6, pp. 250–258, 2020.

[69] J. Zhang, Z. Yong, and Z. Rui, "UAV-enabled radio access network: Multi-mode communication and trajectory design," *IEEE Trans. Signal Process*,

vol. 66, no. 20, pp. 5269–5284, Aug. 2018.

[70] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 434–437, Aug. 2017.

[71] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 604–607, 2017.

[72] Y. Zeng, R. Zhang, and T. J. Lim, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 4983–4996, Dec. 2016.

[73] K.-C. Lan and M.-Z. Wei, "A compressibility-based clustering algorithm for hierarchical compressive data gathering," *IEEE Sensors Journal*, vol. 17, no. 8, pp. 2550–2562, Feb. 2017.

[74] V. Sharma, I. You, G. Pau, M. Collotta, J. Lim, and J. Kim, "Lorawan-based energy-efficient surveillance by drones for intelligent transportation systems," *Energies*, vol. 11, no. 3, p. 573, Jan. 2018.

[75] K. Li, W. Ni, and F. Dressler, "Lstm-characterized deep reinforcement learning for continuous flight control and resource allocation in UAV-assisted sensor network," *IEEE Internet Things J.*, 2021.

[76] N. Ahmed, S. S. Kanhere, and S. Jha, "On the importance of link characterization for aerial wireless sensor networks," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 52–57, May 2016.

[77] X. Yuan, Z. Feng, W. Ni, Z. Wei, R. P. Liu, and C. Xu, "Connectivity of UAV

swarms in 3d spherical spaces under (un) intentional ground interference," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8792–8804, 2020.

[78] F. Jiang and A. L. Swindlehurst, "Optimization of UAV heading for the ground-to-air uplink," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 5, pp. 993–1005, Jun. 2012.

[79] J. Gong, T. H. Chang, C. Shen, and X. Chen, "Flight time minimization of UAV for data collection over wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1942–1954, Sept. 2018.

[80] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 729–743, 2019.

[81] X. Xu, Y. Zeng, Y. L. Guan, and R. Zhang, "Overcoming endurance issue: UAV-enabled communications with proactive caching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1231–1244, Jun. 2018.

[82] D. Popescu, C. Dragana, F. Stoican, L. Ichim, and G. Stamatescu, "A collaborative UAV-WSN network for monitoring large areas," *Sensors*, vol. 18, no. 12, p. 4202, Nov. 2018.

[83] S. Arabi, E. Sabir, H. Elbiaze, and M. Sadik, "Data gathering and energy transfer dilemma in UAV-assisted flying access network for IoT," *Sensors*, vol. 18, no. 5, p. 1519, May 2018.

[84] S. Fu, L. Zhao, Z. Su, and X. Jian, "UAV based relay for wireless sensor networks in 5G systems," *Sensors*, vol. 18, no. 8, p. 2413, Jul. 2018.

[85] S. Hu, X. Chen, W. Ni, X. Wang, and E. Hossain, "Modeling and analysis of energy harvesting and smart grid-powered wireless communication networks:

A contemporary survey," *IEEE Trans.Trans. Green Commun. Netw.*, vol. 4, no. 2, pp. 461–496, 2020.

[86] K. Li, W. Ni, M. Abolhasan, and E. Tovar, "Reinforcement learning for scheduling wireless powered sensor communications," *IEEE Trans.Trans. Green Commun. Netw.*, vol. 3, no. 2, pp. 264–274, 2018.

[87] N. Zhao, W. Lu, M. Sheng, Y. Chen, J. Tang, F. R. Yu, and K.-K. Wong, "UAV-assisted emergency networks in disasters," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 45–51, 2019.

[88] Y. Li and L. Cai, "UAV-assisted dynamic coverage in a heterogeneous cellular system," *IEEE Netw.*, vol. 31, no. 4, pp. 56–61, 2017.

[89] W. Y. B. Lim, S. Garg, Z. Xiong, Y. Zhang, D. Niyato, C. Leung, and C. Miao, "UAV-assisted communication efficient federated learning in the era of the artificial intelligence of things," *IEEE Netw.*, vol. 35, no. 5, pp. 188–195, 2021.

[90] Z. Feng, L. Ji, Q. Zhang, and W. Li, "Spectrum management for mmwave enabled UAV swarm networks: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 57, no. 1, pp. 146–153, 2019.

[91] Y. Wang, Z. Su, N. Zhang, and D. Fang, "Disaster relief wireless networks: Challenges and solutions," *IEEE Wireless Commun.*, vol. 28, no. 5, pp. 148–155, 2021.

[92] G. Zhang, H. Yan, Y. Zeng, M. Cui, and Y. Liu, "Trajectory optimization and power allocation for multi-hop UAV relaying communications," *IEEE Access*, vol. 6, pp. 48 566–48 576, Aug. 2018.

[93] Y. Chen, W. Feng, and G. Zheng, "Optimum placement of UAV as relays," *IEEE Communications Letters*, vol. 22, no. 2, pp. 248–251, Feb. 2018.

[94] J. Chen and D. Gesbert, "Optimal positioning of flying relays for wireless networks: A LoS map approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, Jul. 2017, pp. 1–6.

[95] Y. Zeng, X. Xu, and R. Zhang, "Trajectory design for completion time minimization in UAV-enabled multicasting," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2233–2246, Apri. 2018.

[96] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3747–3760, Jun. 2017.

[97] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Wireless communication using unmanned aerial vehicles (UAVs): Optimal transport theory for hover time optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8052–8066, Dec. 2017.

[98] Y. Wang, S. Li, W. Ni, M. Zhao, A. Jamalipour, and B. Wu, "Cooperative three-dimensional position mapping based on received signal strength measurements: Algorithm design and field test," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10 541–10 552, 2021.

[99] F. Cheng, S. Zhang, Z. Li, Y. Chen, N. Zhao, F. R. Yu, and V. C. Leung, "UAV trajectory optimization for data offloading at the edge of multiple cells," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6732–6736, Jul. 2018.

[100] X. Yuan, Z. Feng, W. Xu, W. Ni, J. A. Zhang, Z. Wei, and R. P. Liu, "Capacity analysis of UAV communications: Cases of random trajectories," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7564–7576, 2018.

[101] Y. Chen, N. Zhao, Z. Ding, and M.-S. Alouini, "Multiple UAVs as relays: Multi-hop single link versus multiple dual-hop links," *IEEE Trans. Wireless*

Commun., vol. 17, no. 9, pp. 6348–6359, 2018.

[102] A. Alsharoa, H. Ghazzai, M. Yuksel, A. Kadri, and A. E. Kamal, "Trajectory optimization for multiple UAVs acting as wireless relays," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Kansas City, USA, May 2018, pp. 1–6.

[103] X. Jiang, X. Chen, J. Tang, N. Zhao, X. Y. Zhang, D. Niyato, and K.-K. Wong, "Covert communication in UAV-assisted air-ground networks," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 190–197, 2021.

[104] M. Li, F. R. Yu, P. Si, R. Yang, Z. Wang, and Y. Zhang, "UAV-assisted data transmission in blockchain-enabled M2M communications with mobile edge computing," *IEEE Netw.*, vol. 34, no. 6, pp. 242–249, 2020.

[105] R. I. Ansari, N. Ashraf, S. A. Hassan, D. G. C., H. Pervaiz, and C. Politis, "Spectrum on demand: A competitive open market model for spectrum sharing for UAV-assisted communications," *IEEE Netw.*, vol. 34, no. 6, pp. 318–324, 2020.

[106] A. S. Abdalla, K. Powell, V. Marojevic, and G. Geraci, "UAV-assisted attack prevention, detection, and recovery of 5g networks," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 40–47, 2020.

[107] B. Shang, L. Liu, J. Ma, and P. Fan, "Unmanned aerial vehicle meets vehicle-to-everything in secure communications," *IEEE Commun. Mag.*, vol. 57, no. 10, pp. 98–103, 2019.

[108] F. Cheng, G. Gui, N. Zhao, Y. Chen, J. Tang, and H. Sari, "UAV-relaying-assisted secure transmission with caching," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3140–3153, Jan. 2019.

[109] N. Zhao, F. Cheng, F. R. Yu, J. Tang, Y. Chen, G. Gui, and H. Sari, "Caching UAV assisted secure transmission in hyper-dense networks based on interference alignment," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2281–2294, Jan. 2018.

[110] X. Sun, W. Yang, Y. Cai, and M. Wang, "Secure mmwave UAV-enabled swipt networks based on random frequency diverse arrays," *IEEE Internet Things J.*, vol. 8, no. 1, pp. 528–540, 2021.

[111] X. Sun, W. Yang, and Y. Cai, "Secure communication in noma-assisted millimeter-wave swipt UAV networks," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1884–1897, 2020.

[112] M. T. Mamaghani and Y. Hong, "Improving phy-security of UAV-enabled transmission with wireless energy harvesting: Robust trajectory design and communications resource allocation," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8586–8600, 2020.

[113] D. N. K. Jayakody, T. D. P. Perera, A. Ghrayeb, and M. O. Hasna, "Self-energized UAV-assisted scheme for cooperative wireless relay networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 578–592, 2020.

[114] S. Yin, Y. Zhao, L. Li, and F. R. Yu, "UAV-assisted cooperative communications with time-sharing information and power transfer," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1554–1567, 2020.

[115] C. Jeong and S. H. Chae, "Simultaneous wireless information and power transfer for multiuser UAV-enabled IoT networks," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8044–8055, 2021.

[116] X. Jie, Z. Yong, and Z. Rui, "UAV-enabled wireless power transfer: Trajectory design and energy optimization," *IEEE Trans. Wireless Commun.*, vol. 17,

no. 8, pp. 5092–5106, Aug. 2018.

[117] S. Jeong, J. Bito, and M. M. Tentzeris, "Design of a novel wireless power system using machine learning techniques for drone applications," in *Wireless Power Transfer Conference (WPTC)*, Kansas City, USA, May 2017, pp. 1–4.

[118] S. Yin, Y. Zhao, and L. Li, "UAV-assisted cooperative communications with time-sharing SWIPT," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kansas City, USA, May 2018, pp. 1–6.

[119] X. Zhou, C. K. Ho, and R. Zhang, "Wireless power meets energy harvesting: A joint energy allocation approach in OFDM-based system," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3481–3491, May 2016.

[120] X. Yuan, Z. Feng, J. A. Zhang, W. Ni, R. P. Liu, Z. Wei, and C. Xu, "Spatio-temporal power optimization for MIMO joint communication and radio sensing systems with training overhead," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 514–528, 2020.

[121] H. Kim, H. Lee, M. Ahn, H.-B. Kong, and I. Lee, "Joint subcarrier and power allocation methods in full duplex wireless powered communication networks for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4745–4753, Jul. 2016.

[122] K. Xiong, P. Fan, C. Zhang, and K. B. Letaief, "Wireless information and energy transfer for two-hop non-regenerative MIMO-OFDM relay networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1595–1611, Aug. 2015.

[123] M. O. Demir, A. E. Pusane, G. Dartmann, G. Ascheid, and G. K. Kurt, "A garden of cyber physical systems: Requirements, challenges, and implementation aspects," *IEEE Internet Things Mag.*, vol. 3, no. 3, pp. 84–89, 2020.

[124] H. X. Nguyen, R. Trestian, D. To, and M. Tatipamula, "Digital twin for 5g and beyond," *IEEE Commun. Mag.*, vol. 59, no. 2, pp. 10–15, 2021.

[125] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial internet of things: Challenges, opportunities, and directions," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4724–4734, 2018.

[126] Y. Liao, E. d. F. R. Loures, and F. Deschamps, "Industrial internet of things: A systematic literature review and insights," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4515–4525, 2018.

[127] J. Zhang, Y. Wu, G. Min, F. Hao, and L. Cui, "Balancing energy consumption and reputation gain of UAV scheduling in edge computing," *IEEE Trans.Trans. Cogn. Commun. Netw.*, vol. 6, no. 4, pp. 1204–1217, 2020.

[128] D. Liu, Y. Xu, J. Wang, J. Chen, K. Yao, Q. Wu, and A. Anpalagan, "Opportunistic UAV utilization in wireless networks: Motivations, applications, and challenges," *IEEE Commun. Mag.*, vol. 58, no. 5, pp. 62–68, 2020.

[129] P. A. Catherwood, M. Little, D. Finlay, and J. M. OBE, "Recovery of incapacitated commercial delivery drones using lpwan technology," *IEEE Intell. Transp. Syst. Mag.*, vol. 12, no. 2, pp. 6–19, 2019.

[130] B. Sorbelli and et al., "Energy-constrained delivery of goods with drones under varying wind conditions," *IEEE Trans. Intell. Transp. Syst.*, 2020 (Early Access).

[131] R. Gupta, A. Kumari, S. Tanwar, and N. Kumar, "Blockchain-envisioned softwarized multi-swarming uavs to tackle COVID-19 situations," *IEEE Netw.*, vol. 35, no. 2, pp. 160–167, 2020.

[132] K. Gai, Y. Wu, L. Zhu, K.-K. R. Choo, and B. Xiao, "Blockchain-enabled

trustworthy group communications in uav networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4118–4130, 2020.

[133] FAA, "Certificate of waiver no. 107w-2020-01564." [Online]. Available: https://www.faa.gov/uas/commercial_operators/part_107_waivers

[134] M. S. Hossain, G. Muhammad, and et al., "Explainable ai and mass surveillance system-based healthcare framework to combat COVID-19 like pandemics," *IEEE Netw.*, vol. 34, no. 4, pp. 126–132, 2020.

[135] S. H. Kim, "Choice model based analysis of consumer preference for drone delivery service," *J. Air Transp. Manag.*, vol. 84, p. 101785, 2020.

[136] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY USA: John Wiley & Sons, 2014.

[137] Y.-H. Hsu and R.-H. Gau, "Reinforcement learning-based collision avoidance and optimal trajectory planning in UAV communication networks," *IEEE Trans. Mobile Comput.*, 2020.

[138] Mercedes-Benz, "Pilot project for on-demand delivery," https://www.mercedes-benz.com/en/vehicles/transporter/vans-drones-in-zurich, accessed: 2010-09-30.

[139] Amazon , "Amazon acquires autonomous vehicle developer zoox," https://www.mhlnews.com/technology-automation/article/21135764/amazon-acquires-autonomous-vehicle-developer-zoox, accessed: 2010-09-30.

[140] S. C.-K. Chau, S. Shen, and Y. Zhou, "Decentralized ride-sharing and vehicle-pooling based on fair cost-sharing mechanisms," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 1936–1946, 2022.

[141] Z. Wang and et al., "Relationship between gaze behavior and steering performance for driver–automation shared control: a driving simulator study," *IEEE Trans. Intell. Veh.*, vol. 4, no. 1, pp. 154–166, 2018.

[142] E. Bulut, M. C. Kisacikoglu, and K. Akkaya, "Spatio-temporal non-intrusive direct V2V charge sharing coordination," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9385–9398, 2019.

[143] N. Ichihara, "Convergence of value functions for finite horizon markov decision processes with constraints," *Appl. Math. Optim.*, pp. 1–44, 2020.

[144] J. K. Stolaroff, C. Samaras, E. R. OâNeill, A. Lubers, A. S. Mitchell, and D. Ceperley, "Energy use and life cycle greenhouse gas emissions of drones for commercial package delivery," *Nature Communications*, vol. 9, no. 1, pp. 1–13, 2018.

[145] D.-H. Tran, T. X. Vu, S. Chatzinotas, S. ShahbazPanahi, and B. Ottersten, "Coarse trajectory design for energy minimization in UAV-enabled," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 9483–9496, 2020.

[146] T. Zhang, G. Liu, H. Zhang, W. Kang, G. K. Karagiannidis, and A. Nallanathan, "Energy-efficient resource allocation and trajectory design for UAV relaying systems," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6483–6498, 2020.

[147] M. Perreault and K. Behdinan, "Delivery drone driving cycle," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1146–1156, 2021.

[148] S. Sekander, H. Tabassum, and E. Hossain, "Statistical performance modeling of solar and wind-powered UAV communications," *IEEE Trans. Mobile Comput.*, vol. 20, no. 8, pp. 2686–2700, 2020.

[149] M. H. Cheung and J. Huang, "DAWN: Delay-aware Wi-Fi offloading and network selection," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1214–1223, Jun. 2015.

[150] V. Krishnamurthy, *Partially Observed Markov Decision Processes*. Cambridge University Press, 2016.

[151] J. Zhang and H. Jin, "Optimized calculation of the economic speed profile for slope driving: Based on iterative dynamic programming," *IEEE Trans. Intell. Transp. Syst.*, 2020 (Early access).

[152] D. Liu, L. Khoukhi, and A. Hafid, "Prediction-based mobile data offloading in mobile cloud computing," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4660–4673, 2018.

[153] C. Paquette and K. Scheinberg, "A stochastic line search method with expected complexity analysis," *SIAM Journal on Optimization*, vol. 30, no. 1, pp. 349–376, 2020.

[154] DJI, "AGRAS MG-1." [Online]. Available: https://www.dji.com/be/mg-1

[155] H. Wang, J. Wang, G. Ding, J. Chen, and J. Yang, "Completion time minimization for turning angle-constrained UAV-to-UAV communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4569–4574, 2020.

[156] DJI. MG-12000P Flight Battery User Guide v1.2. [Online]. Available: https://www.dji.com/be/mg-1p/info

[157] J. Liu, C. Bondiombouy, L. Mo, and P. Valduriez, "Two-phase scheduling for efficient vehicle sharing," *IEEE Trans. Intell. Transp. Syst.*, 2020.

[158] F. Tao and Q. Qi, "New IT driven service-oriented smart manufacturing: framework and characteristics," *IEEE Trans. Syst., Man, Cybern. Syst*,

vol. 49, no. 1, pp. 81–91, 2019.

[159] H. Sun, C. Peng, D. Yue, Y. L. Wang, and T. Zhang, "Resilient load frequency control of cyber-physical power systems under QoS-dependent event-triggered communication," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 51, no. 4, pp. 2113–2122, 2020.

[160] M. Kishk, A. Bader, and M.-S. Alouini, "Aerial base station deployment in 6G cellular networks using tethered drones: The mobility and endurance tradeoff," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 103–111, 2020.

[161] R. G. Ribeiro, L. P. Cota, T. A. Euzébio, J. A. Ramírez, and F. G. Guimarães, "Unmanned-aerial-vehicle routing problem with mobile charging stations for assisting search and rescue missions in postdisaster scenarios," *IEEE Trans. Syst., Man, Cybern. Syst*, 2021 (Early access).

[162] Y. Qu, H. Dai, H. Wang, C. Dong, F. Wu, S. Guo, and Q. Wu, "Service provisioning for UAV-enabled mobile edge computing," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3287–3305, 2021.

[163] F. Santoso, M. A. Garratt, and S. G. Anavatti, "Hybrid pd-fuzzy and pd controllers for trajectory tracking of a quadrotor unmanned aerial vehicle: Autopilot designs and real-time flight tests," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 51, no. 3, pp. 1817 – 1829, 2021.

[164] G. Qi, X. Li, and Z. Chen, "Problems of extended state observer and proposal of compensation function observer for unknown model and application in UAV," *IEEE Trans. Syst., Man, Cybern. Syst*, 2021 (Early access).

[165] F. Santoso, M. A. Garratt, S. G. Anavatti, and I. Petersen, "Robust hybrid nonlinear control systems for the dynamics of a quadcopter drone," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 50, no. 8, pp. 3059–3071, 2018.

[166] D. Liu, S. Xue, B. Zhao, B. Luo, and Q. Wei, "Adaptive dynamic programming for control: A survey and recent advances," *IEEE Trans. Syst., Man, Cybern. Syst*, vol. 51, no. 1, pp. 142–160, 2021.

[167] D. P. Bertsekas, "Affine monotonic and risk-sensitive models in dynamic programming," *IEEE Trans. Trans. Automat. Contr.*, vol. 64, no. 8, pp. 3117–3128, 2019.

[168] N. Bauerle and D. Lange, "Optimal control of partially observable piecewise deterministic Markov processes," *SIAM Journal on Control and Optimization*, vol. 56, no. 2, pp. 1441–1462, 2018.

[169] Y. Yang, W. Liu, E. Wang, and J. Wu, "A prediction-based user selection framework for heterogeneous mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 18, no. 11, pp. 2460–2473, 2019.

[170] M. H. Cheung, F. Hou, J. Huang, and R. Southwell, "Congestion-aware dns for integrated cellular and Wi-Fi networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1269–1281, 2017.

[171] A. P. Cohen, S. A. Shaheen, and E. M. Farrar, "Urban air mobility: History, ecosystem, market potential, and challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 6074–6087, 2021.

[172] J. Shi and et. al, "Flying social networks: Architecture, challenges and open issues," *IEEE Netw.*, 2021 (Early access).

[173] N. Cherif, W. Jaafar, H. Yanikomeroglu, and A. Yongacoglu, "3D aerial highway: The key enabler of the retail industry transformation," *IEEE Commun. Mag.*, vol. 59, no. 9, pp. 65–71, 2021.

[174] Z. Su *et al.*, "UAV enabled content distribution for internet of connected vehicles in 5G heterogeneous networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5091–5102, 2021.

[175] M. S. Rahman, I. Khalil, and M. Atiquzzaman, "Blockchain-powered policy enforcement for ensuring flight compliance in drone-based service systems," *IEEE Netw.*, vol. 35, no. 1, pp. 116–123, 2021.

[176] M. H. Cheung, F. Hou, J. Huang, and R. Southwell, "Distributed time-sensitive task selection in mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 20, no. 6, 2021.

[177] X. Lyu, H. Tian, W. Ni, Y. Zhang, P. Zhang, and R. P. Liu, "Energy-efficient admission of delay-sensitive tasks for mobile edge computing," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2603–2616, 2018.

[178] H. Y. Jeong *et al.*, "The flying warehouse delivery system: A quantitative approach for the optimal operation policy of airborne fulfillment center," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 12, pp. 7521–7530, 2020.

[179] B. Golden, "Shortest-path algorithms: A comparison," *Operations Research*, vol. 24, no. 6, pp. 1164–1168, 1976.

[180] H. Guo, X. Hou, Z. Cao, and J. Zhang, "GP3: Gaussian process path planning for reliable shortest path in transportation networks," *IEEE Trans. Intell. Transp. Syst.*, 2021 (Early access).