Multi-Drone-Cell 3D Trajectory Planning and Resource Allocation for Drone-Assisted Radio Access Networks

by

Weisen Shi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Xianbin Wang Professor, Western University
Supervisor(s):	Xuemin (Sherman) Shen University Professor, University of Waterloo
Internal Member:	Kshirasagar Naik Professor, University of Waterloo
Internal Member:	Xiaodong Lin Adjunct Associate Professor, University of Waterloo
Internal-External Member:	Jun Liu Associate Professor, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Equipped with communication modules, drones can perform as drone-cells (DCs) that provide on-demand communication services to users in various scenarios, such as traffic monitoring, Internet of things (IoT) data collections, and temporal communication provisioning. As the aerial relay nodes between terrestrial users and base stations (BSs), DCs are leveraged to extend wireless connections for uncovered users of radio access networks (RAN), which forms the drone-assisted RAN (DA-RAN). In DA-RAN, the communication coverage, quality-of-service (QoS) performance and deployment flexibility can be improved due to the line-of-sight DC-to-ground (D2G) wireless links and the dynamic deployment capabilities of DCs. Considering the special mobility pattern, channel model, energy consumption, and other features of DCs, it is essential yet challenging to design the flying trajectories and resource allocation schemes for DA-RAN. In specific, given the emerging D2G communication models and dynamic deployment capability of DCs, new DC deployment strategies are required by DA-RAN. Moreover, to exploit the fully controlled mobility of DCs and promote the user fairness, the flying trajectories of DCs and the D2G communications must be jointly optimized. Further, to serve the high-mobility users (e.g. vehicular users) whose mobility patterns are hard to be modeled, both the trajectory planning and resource allocation schemes for DA-RAN should be re-designed to adapt to the variations of terrestrial traffic. To address the above challenges, in this thesis, we propose a DA-RAN architecture in which multiple DCs are leveraged to relay data between BSs and terrestrial users. Based on the theoretical analyses of the D2G communication, DC energy consumption, and DC mobility features, the deployment, trajectory planning and communication resource allocation of multiple DCs are jointly investigated for both quasi-static and high-mobility users.

We first analyze the communication coverage, drone-to-BS (D2B) backhaul link quality, and optimal flying height of the DC according to the state-of-the-art drone-to-user (D2U) and D2B channel models. We then formulate the multi-DC three-dimensional (3D) deployment problem with the objective of maximizing the ratio of effectively covered users while guaranteeing D2B link qualities. To solve the problem, a per-drone iterated particle swarm optimization (DI-PSO) algorithm is proposed, which prevents the large particle searching space and the high violating probability of constraints existing in the pure PSO based algorithm. Simulations show that the DI-PSO algorithm can achieve higher coverage ratio with less complexity comparing to the pure PSO based algorithm.

Secondly, to improve overall network performance and the fairness among edge and central users, we design 3D trajectories for multiple DCs in DA-RAN. The multi-DC 3D trajectory planning and scheduling is formulated as a mixed integer non-linear programming (MINLP) problem with the objective of maximizing the average D2U throughput. To

address the non-convexity and NP-hardness of the MINLP problem due to the 3D trajectory, we first decouple the MINLP problem into multiple integer linear programming and quasi-convex sub-problems in which user association, D2U communication scheduling, horizontal trajectories and flying heights of DBSs are respectively optimized. Then, we design a multi-DC 3D trajectory planning and scheduling algorithm to solve the sub-problems iteratively based on the block coordinate descent (BCD) method. A k-means-based initial trajectory generation scheme and a search-based start slot scheduling scheme are also designed to improve network performance and control mutual interference between DCs, respectively. Compared with the static DBS deployment, the proposed trajectory planning scheme can achieve much lower average value and standard deviation of D2U pathloss, which indicate the improvements of network throughput and user fairness.

Thirdly, considering the highly dynamic and uncertain environment composed by high-mobility users, we propose a hierarchical deep reinforcement learning (DRL) based multi-DC trajectory planning and resource allocation (HDRLTPRA) scheme for high-mobility users. The objective is to maximize the accumulative network throughput while satisfying user fairness, DC power consumption, and DC-to-ground link quality constraints. To address the high uncertainties of environment, we decouple the multi-DC TPRA problem into two hierarchical sub-problems, i.e., the higher-level global trajectory planning sub-problem and the lower-level local TPRA sub-problem. First, the global trajectory planning sub-problem is to address trajectory planning for multiple DCs in the RAN over a long time period. To solve the sub-problem, we propose a multi-agent DRL based global trajectory planning (MARL-GTP) algorithm in which the non-stationary state space caused by multi-DC environment is addressed by the multi-agent ngerprint technique. Second, based on the global trajectory planning results, the local TPRA (LTPRA) sub-problem is investigated independently for each DC to control the movement and transmit power allocation based on the real-time user trafc variations. A deep deterministic policy gradient based LTPRA (DDPG-LTPRA) algorithm is then proposed to solve the LTPRA sub-problem. With the two algorithms addressing both sub-problems at different decision granularities, the multi-DC TPRA problem can be resolved by the HDRLTPRA scheme. Simulation results show that 40% network throughput improvement can be achieved by the proposed HDRLTPRA scheme over the non-learning-based TPRA scheme.

In summary, we have investigated the multi-DC 3D deployment, trajectory planning and communication resource allocation in DA-RAN considering different user mobility patterns in this thesis. The proposed schemes and theoretical results should provide useful guidelines for future research in DC trajectory planning, resource allocation, as well as the real deployment of DCs in complex environments with diversified users.

Acknowledgements

First, I would like to express my deep and sincere gratitude to my supervisor, Professor Xuemin Shen, for his enormous support, endless help, and invaluable guidance during my Ph.D. study in University of Waterloo. It was Professor Shens insightful advice and continuous encouragement that help me accumulate knowledges, academic skills, and professional experiences, which pave a bright path for my future careers. I always expect to hear Professor Shen's comments after each weekly group meeting and subgroup meetings. From those comments, his enthusiasm and dedication to the career, comprehensive understanding of the research, conscientious and patient attitudes to every students and everything do inspire me a lot, and will be the precious treasure for me in future life. It is extremely lucky and great honor for me to be his student.

I would also like to express my gratitude to Professor Weihua Zhuang for her valuable guidance and great support in my Ph.D. study, especially in the Huawei SONAC project. The project experience did train me how to conduct in-depth and solid research works.

I would like to thank Professor Xiaodong Lin, Professor Kshirasagar Naik, Professor Jun Liu for serving my thesis examination committee. I would also like to thank Professor Xianbin Wang to serve as my external examiner. Their valuable suggestions and questions have significantly improved the quality of my thesis.

I would like to express my sincere appreciation to all the members of BBCR-SAG and BBCR-SDN subgroups. Without your valuable suggestions, continues encouragement and support, it is impossible for me to complete all my tasks and achievements. I would like to especially thank Professor Nan Cheng, Professor Qiang Ye, Professor Haibo Zhou, Professor Ning Zhang, Professor Shan Zhang, Professor Feng Lyu, Professor Peng Yang, Dr. Wenchao Xu, Dr. Xu Li, Professor Wei Quan, Junling Li, Huaqing Wu, Omar Alhussein, Jiayin Chen, Conghao Zhou, Dr. Phu Thinh Do, Kaige Qu, Si Yan, Dr. Haixia Peng, Mushu Li and Amr Salah Matar for their great help in both my research and life. I would also like to thank Dr. Nan Chen, Dongxiao Liu, Dr. Wen Wu, Liang Xue, Professor Kuan Zhang, Professor Jianbing Ni, Dr. Haohao Liao and all current and former BBCR members. I received precious friendship and support from all of you.

Finally, I would like to thank my parents for their understanding, support, and encouragement during my Ph.D. study.

Weisen Shi
May 11, 2020
Ottawa, Ontario, Canada

Dedication

This thesis is dedicated to my beloved parents, Ruiying Guo and Fengru Shi,

and in memory of my grandparents Yuzhen Song and Jianlin Guo.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
AI	Artificial Intelligence
AoI	Area of Interest
BCD	Block Coordinate Descent
BS	Base Station
BTS	Bursty Traffic Spot
C-RAN	Cloud-RAN
CDF	Cumulative Distribution Function
CH	Coverage Hole
D2B	Drone-to-Base-station
D2D	Drone-to-Drone
D2G	Drone-to-Ground
D2U	Drone-to-User
DA-RAN	Drone-Assisted Radio Access Networks
DC	Drone-Cell
DDPG	Deep Deterministic Policy Gradient
DI-PSO	per-Drone Iterated PSO
DQN	Deep Q-Network
DRL	Deep Reinforcement Learning
ESN	Echo State Networks
ETSI	European Telecommunications Standard Institute
FANET	Flying Ad Hoc Networks
GTP	Global Trajectory Planning
HAP	High Altitude Platform

HDRL	Hierarchical Deep Reinforcement Learning
HetNets	Heterogeneous Networks
ILP	Integer Linear Programming
IoT	Internet of Things
LAP	Low Altitude Platform
LEO	Low Earth Orbit Satellite
LoS	Line-of-Sight
LTE	Long-Term Evolution
LTPRA	Local Trajectory Planning and Resource Allocation
MARL	Multi-Agent Deep Reinforcement Learning
mmWAVE	Millimeter Wave
MINLP	Mixed Integer Non-Linear Programming
NLoS	Non-Line-of-Sight
NN	Neural Network
OFDMA	Orthogonal Frequency Division Multiple Access
PDF	Probability Distribution Function
PSO	Particle Swarm Optimization
QoS	Quality-of-Service
RAN	Radio Access Networks
RL	Reinforcement Learning
TDMA	Time Division Multiple Access
TPRA	Trajectory Planning and Resource Allocation
UAV	Unmanned Aerial Vehicle

Chapter 1

Introduction

Ubiquitous connectivity in anywhere at anytime with guaranteed quality-of-service (QoS) is expected by users of future radio access networks (RAN) [1]. However, the radio coverage holes (CHs) of terrestrial RAN caused by building blockage or lacking infrastructures, as well as the highly dynamic and uneven distribution of terrestrial data traffic pose great challenges to ensure the ubiquitous connectivity [2] [3] [4]. Although densely deploying massive small cells can be a feasible solution, it is inefficient and costly for RAN operators due to the high idle probability of small cells deployed for peak hours or remote areas [5]. To address the challenges in a cost-efficient way, drone, *a.k.a.* unmanned aerial vehicle (UAV), equipped with wireless communication modules is leveraged as drone-cell (DC) to assist future RAN, which forms the drone-assisted RAN (DA-RAN). The DC can benefit DA-RAN by its high-quality DC-to-ground (D2G) wireless links, the dynamic and controllable deployment, and the on-broad communication and computing capabilities. In this chapter, we first overview the drone communication, then introduce the DC and DA-RAN with their specific communication, mobility, and resource allocation features. Finally, we elaborate the three key problems investigated in this thesis.

1.1 Overview of Drone Communication

With the rapid development of automations and flying control technologies, drone is expected to perform essential roles in diversified scenarios of human life. As predicted by Tractica, the number of drones will continue to surge over the next several years, with the global annual unit shipment increasing more than tenfold from 6.4 million in 2015 to 67.7 million by 2021 [6]. Equipped with dedicated sensors or communication devices, drones

can undertake various services such as traffic surveillance [7], forest fire monitoring [8], homeland security [9], post-disaster rescue [10], logistics application (e.g. Amazon Prime Air [11], DHL Parcelcopter [12]) and communication assistance [13] [14].

As the frontier technology in the evolution to the Internet of things (IoT) era, employing drone to to support broadband wireless communication has attracted increasing research attentions from both academia and industry. Both Google and Facebook announced their drone communication projects (Project Loon [15], and Project Aquila [16]) respectively to enhance ground communication via drone's assistance. In 2016, Nokia proposed the UAV traffic management (UTM) architecture for connected UAVs, where the UTM unit provides functions of fleet management, automated UAV missions, three-dimensional (3D) navigation, and collision avoidance [17]. Intel and AT&T demonstrated the worlds first drone connecting to long-term evolution (LTE) networks at the 2016 Mobile World Congress [18]. Cooperating with China Mobile, Ericsson conducted the world's first 5G-enabled drone prototype field trial in Wuxi, China [19]. Qualcomm has published their Cellular Drone Communication solution to accelerate the development and optimization of 5G networks, and submitted their local drone collaboration and remote control use cases to 3GPP [20]. Huawei also announced the "Digital Sky Initiative" to spur development of drone applications and enable the low airspace digitized economy via enhanced low airspace network coverage [21]. As the response from standardization group, 3GPP released the work item "Enhanced LTE support for aerial vehicles" in 2018 [22].

Considering the different sizes, flying heights and utilities, drones with communication capabilities are usually divided into two main categories: the high altitude platform (HAP) drone that can work at an altitude of 10-22 km, and the low altitude platform (LAP) drone that can reach maximal hundreds meters heights. The HAP drones consist of large-size drones, such as the fix-wing airship [16] or balloons [15]. With the 10-20 km radio coverage radius, HAP drones are ideal platforms to support broadcast services, enhance communication services for rural areas, and relieve the network load on ground. Specific spectrum bands at 28GHz (27.5-28.35GHz for down-link, 31.0-31.3GHz for up-link), and 48GHz (47.9-48.2GHz, 47.2-47.5GHz for up-links and down-links) have been allocated to HAP drone to ground communications. The LAP drones generally include middle-size or small-size rotary-wing drones or helicopters, whose radio coverage radius can only reach hundreds meters. However, due to their flexibility and low-cost, multiple LAP drones can be dynamically deployed and form the drone swarm network to provide temporal communication services, e.g. post-disaster communication recovery, communication assistance for dense scenarios. Existing works usually use Wi-Fi technology to support LAP drone communication due to its availability [23]. Wi-Fi is also widely used in the communication between commercial drones and their controllers. Other research focusing

on the integration of drones and access networks mainly use LTE based communication technologies, especially for D2G communications [24]. Besides, the Zigbee [25], millimeter wave (mmWAVE) [26], and free-space-optics (FSO) [27] technologies are also implemented by preceding works in different scenarios. However, there is still no dedicated spectrum allocated to the LAP drone communications. Table 1.1 compares the difference between HAP and LAP drones.

Table 1.1: Difference between HAP and LAP Drones

Features	HAP drones	LAP drones
Size	Large-size	Middle-size, small-size
Flying Height	$17 - 22$ km [28]	Up to 4km, usually less than 200m [29]
Radio Coverage Radius	$20 - 30$ km [28]	Usually $100 - 300$m
Mobility	Stationary, periodically movement	Dynamic deployment
Duration	Months or years [28]	$0.5 - 2$ hours, up to 5+ hours [30] [31]
Deployment cost	High	Low
Maximal Payload Weight	$>$ 10kg	$3 - 5$kg [30]
Typical use cases	Broadcast services, rural area communication enhancement, ground network load offloading	Temporal communication assistance, post-disaster communication recovery.

The research and implementations of drone communication mainly focus on two aspects: the drone-to-drone (D2D) communication enabling the inter-drone data transmission in a swarm of drones; and the drone-to-ground (D2G) communication by which drones interacting with ground nodes (users or central controllers), as shown in Fig. 1.1.

Similar to the vehicle-to-vehicle (V2V) communication in the vehicular ad hoc networks (VANET), the LAP drones can communicate with each other in ad hoc manner without the assistance of any ground infrastructure, which is referred to as D2D communications. Multiple drones connected by D2D wireless links form a classic ad hoc network in the air, which is referred to as the flying ad hoc networks (FANET). Terrestrial users can access the FANET through any drones flying over them, while the interactions between Internet

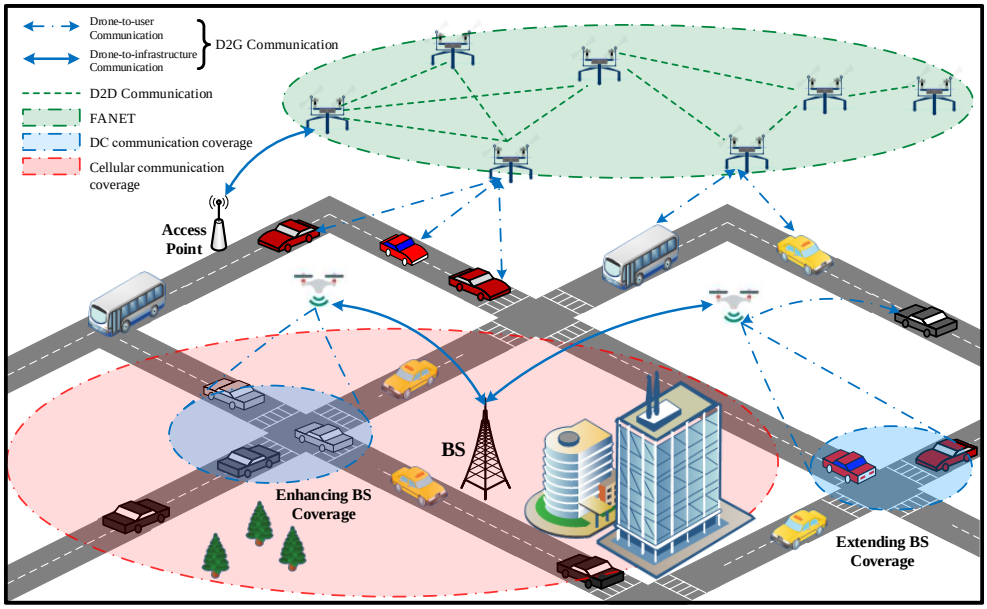


Figure 1.1: Overview of drone communication with D2D and D2G communications.

and the FANET are supported by some backbone drones (e.g. cluster heads of a swarm of drones) connecting to the terrestrial RAN or satellite networks. Focusing on the inter-drone communications, the FANET is mainly implemented in the infrastructure-less scenarios, such as battle field or post-disaster areas, to conduct data collection or communication recovery tasks. For instance, FANET has been leveraged during the 2011 Japan East great earthquake for monitoring the damaged reactors at the Fukushima Daiichi nuclear power plant [32]; In 2017, China mobile has successfully recovered LTE communication for over 30 km^2 post-earthquake area through UAV-based BS [10]. Existing D2D communication and FANET studies usually leverage the previous conclusions or achievements in VANET research due to their high similarity. However, two unique features in D2D communication have to be treated: 1) drones work in a 3D space while the vehicles are running on the two-dimensional (2D) plane, which introduces more complex wireless channel characters and antenna design requirements; and 2) the mobility of drones are usually centralized controlled by the controllers or monitors, which is highly predictable and controllable in network design.

Since the main purpose of drone communication is leveraging drone's specific features to improve the Internet access services for terrestrial users, it is inevitable to investigate the D2G communications. Compared with the terrestrial wireless links, the D2G communication links have their unique features which requires specific D2G channel models. For

instance, the D2G links have higher line-of-sight (LoS) probability than terrestrial wireless links due to the high flying altitude and dynamic deployment capability of drones; the large scale D2G pathloss is impacted by both horizontal distance and the drone flying height, etc. Besides the specific wireless link features, the throughputs and latency requirements for D2G communication are more strict than terrestrial wireless links because of the 3D mobility of drones.

1.2 Drone-Cell and Drone-Assisted Radio Access Networks

Providing ubiquitous connectivity for users and devices with diversified service requirements is regarded as one of the key performance metrics for future RAN [1]. However, coverage holes (CH) of terrestrial RAN prevail in both urban and rural scenarios due to the lack of infrastructures or blocking by obstacles [2]. Besides, the highly dynamic and uneven distribution of terrestrial data traffic poses great challenges to RAN in terms of guaranteeing diversified QoS requirements for users, especially in high-mobility scenarios [33]. To meet the ubiquitous connectivity requirements, considerable level of flexible deployment is expected by future RAN.

In current RAN, all BSs, small cells, and remote radio heads (RRHs) are deployed in certain geographical locations according to long-term traffic behaviors with little flexibility to be re-deployed. Such rigid RANs are reluctant to maintain ubiquitous connectivity for most 5G scenarios where dynamic data traffic occurs in both spatial and temporal domains [34]. Although densely deploying massive small cells is one intuitive way to improve the user coverage of RAN, the small cells deployed for peak hours or remote areas can have high idle probability. Such high expenditure and low efficiency are unacceptable for RAN operators [5]. Therefore, a new type of BSs or small cells with dynamic deployment capability is required by future RAN to address the challenges.

1.2.1 Overview of Drone-Cell

Equipped with specific wireless modules and controlled by corresponding controllers (e.g. edge servers on BSs), the flying drone can perform as DC to provide temporal and on-demand communication services for users in areas of interests [13]. Comparing with legacy BSs, the emerging DC technology is a promising solution to improve network flexibility due to four advantages:

1) Line-of-sight (LoS) connection: Constrained by the two-dimensional topology, traditional wireless links in RAN are obstructed by vehicles or buildings frequently. As a result, most messages are transmitted through non-line-of-sight (NLoS) wireless links which can seriously decrease the QoS of users [35]. On the contrary, DCs flying in the air have higher probability to connect ground users via LoS links, which facilitates highly reliable communications [36]. This advantage is further enhanced by DCs' mobility feature that allows 3D adjustments of their positions to avoid obstacles between DCs and users [37]. For the backhaul links connecting DCs and their corresponding BSs, since the flying height of DCs can be higher or close to the height of BS antennas, the backhaul connections are naturally LoS with little probability to be blocked by tall buildings [24]. Note that the higher LoS probability of D2G wireless link is only valid for outdoor users. For indoor users that naturally blocked by the wall of buildings, neither the DC nor the conventional BSs outside the room can provide LoS connections. Without loss of generality, in this thesis all the users served by the DC are outdoor users, e.g. mobile users, connected vehicles.

2) Dynamic deployment: Different from traditional BSs which are statically fixed on dedicated locations, DCs can be dynamically deployed according to the spatial and temporal changes of ground traffic, and allocated to different users or controllers on demands [38]. Comparing with deploying numerous static small-cells or RRHs, the dynamic deployment capability of DCs can guarantee the same QoS level with less cost.

3) Fully-controlled mobility: Different from connected vehicles in VANET whose mobility is controlled by drivers or autonomous vehicles themselves, the hovering positions and flying trajectories of DCs are fully controlled by the corresponding controllers. This fully-controlled mobility feature enables the dynamic deployment feature of DCs to assist ground infrastructures [39].

4) Computing & caching capabilities: As the flying access platform close to dedicated users, the DC holds the capability to conduct both computing and caching tasks through on-broad CPU and memory. Combining with the communication capability, DCs can serve as dynamic deployed edge devices to further improve the flexibility of ground networks. For instance, the computation tasks of high-mobility users can be offloaded to DCs that follows their mobility patterns, which enhances the hand-over performance and reduces the service delay.

Similar to the terrestrial heterogeneous networks (HetNets), two types of DCs are considered in existing works: the macro DCs and the small DCs. Marco DCs are embedded with full-functions of the ground BS to provide ubiquitous access services for ground users (e.g. enhancing radio coverage in rural areas; assisting core network in congestion scenarios, etc.). As the flying access point to core network, macro DCs have to ensure relatively large radio coverage range and continuously work for long time, which can only be sup-

ported by HAP drones. Theoretical analyses and field trials of using HAP drones to form macro DCs have been conducted by various preceding works [40] [41] [42].

Different from macro DCs, the small DCs are usually equipped with partial functions of the ground BS, such as the functions of a small-cell, to prevent large payload weight. Each small DC is controlled and connects to its corresponding BS. For each BS, multiple small DCs perform as its flying extensions that enhance its user coverage and flexibility. Consist of LAP drones, the radio coverage range and working time of small DCs are limited, while more flexibility and lower deployment cost are gained. So small DCs are mainly used for scenarios with dynamic and temporal communication requirements (e.g. hand-off enhancement for high mobility users; additional resource provision for temporary events; compensation for temporary coverage holes (CHs), etc.) Existing works investigating small DCs mainly focus on their 3D spatial deployment and mobility management issues [1].

1.2.2 Drone-Assisted Radio Access Networks Architecture

Inspired by the four advantages of DCs, the DA-RAN architecture is proposed by some pioneer works to enable flexible deployment capability for future RAN. In DA-RAN, single or multiple DCs are dynamically deployed to serve users in areas of interest (AoIs) using additional spectrum resource, and relay data between user and terrestrial BSs. AoIs include both the CHs of terrestrial RAN's communication coverage, as well as the bursty traffic spots (BTSs) where allocated terrestrial RAN resources are inadequate to support the dense traffic, e.g. congested roads, stadium with sports events, etc.

Different deployment architectures of DA-RAN are designed for different scenarios. In specific, various DA-RAN architectures in existing works can be categorized into two categories according to the DC types, i.e., the macro DC DA-RAN and the small DC DA-RAN. Fig. 1.2. illustrates the two architectures.

The Macro DC DA-RAN is referred to as the DA-RAN composed by macro DCs, where each DC is equipped with full functions of a ground BS. As an independent RAN flying over the terrestrial RAN, macro DC DA-RAN usually working on non-cellular band that provides additional spectrum resource to alleviate the spectrum scarcity and enhance the connectivity of ground users. The backhaul link is specifically defined for each DC, one DC can communicate with multiple ground access points (e.g. BSs), while multiple DCs can also access core network via a same central nodes, such as the satellite. The reason for using macro DC is to guarantee the large energy consumption required by full BS functions, and support large radio coverage area comparable to ground BS. Classic works in macro DC DA-RAN research mainly focus on the network design and DC deployment to

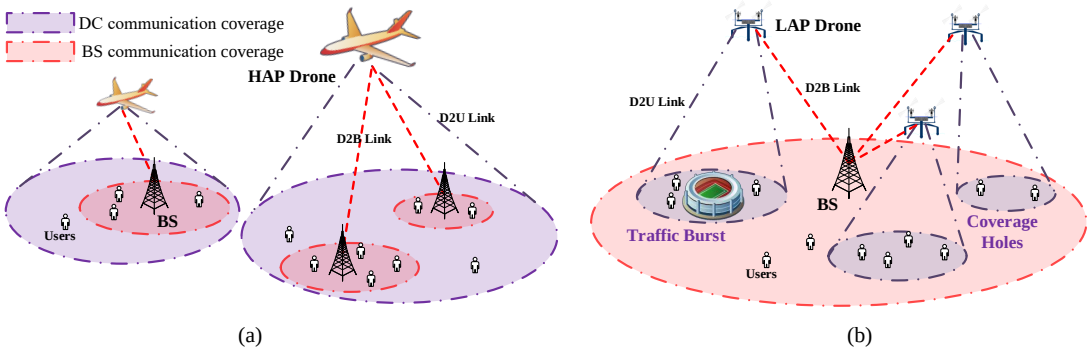


Figure 1.2: Two deployment architectures of DA-RAN: a) Macro DC DA-RAN, b) Small DC DA-RAN.

provide ubiquitous access services for users. In [40], Chandrasekharan *et al.* reported the detailed challenges in the design and implementation of a A-RAN consisting of LTE-A base stations. They show that the Helikite enabled aerial platform solutions and LTE-A can be proficiently used to provision Internet access to ground users. Dong *et al.* investigated the constellation design of DC deployment. Based on the radio coverage model of single DC, a cost-efficient optimization framework is designed to maximize network capacity under the QoS constraints [28]. Based on the multi-tire macro DC DA-RAN architecture proposed in [1], Alzenad *et al.* leveraged HAP and LAP drones to providing communication service by the help of FSO based air-to-ground links [27].

In small DC DA-RAN, all DCs only exchange data with corresponding central BSs. This architecture is similar to the cloud-RAN (C-RAN) architecture while the RRHs are replaced by DCs [43]. BSs perform as centralized controllers that manage the deployment and resource allocation for DCs. For each BS, it release multiple DCs over the AoIs within its communication coverage. Users in AoIs can directly access to DCs, then connect to the BS by the relay of DCs. Since each DC only performs as the flying extension of its corresponding BS, partial BS functions can be embedded on the DC to saving power consumption. To ensure fast re-deployments and eliminate the inter-drone interference, the small DC with relatively small radio coverage and high flexibility is the desirable platform to support small DC DA-RAN. Field experiments have been conducted to prove that LAP drone based DCs hold the capability of improving the signal strength for CHs under this architecture [44]. Emerging works in DA-RAN research mainly focus on the deployment and mobility design of DCs, which is the main topic of this proposal too. In this thesis, we focus on the small DC DA-RAN. Without specifications, all DCs discussed in the remainder

of this thesis are referred to as small DCs supported by LAP drones, and all DA-RAN in the following contents denote the small DC DA-RAN.

The main differences between DA-RAN and conventional RAN are compared in Table 1.2, where the wired links are denoted as —, the wireless links are denoted as $\cdots$.

Table 1.2: Comparison between DA-RAN and Conventional RAN

	DA-RAN	Conventional RAN
Components	Users in AoIs, DCs, macro BSs	Cellular users, small-cells, macro BSs
Data transmission	User $\cdots$ DC $\cdots$ BSs — core network	User $\cdots$ (small-cell —) BSs — core network
Spectrum resources	D2U: Unlicensed spectrum D2B: cellular spectrum, mmWAVE	Cellular spectrum
Channel models	Specific D2G channel models [45] [24]	Terrestrial wireless channel models (free-space model, Okumura-Hata model, etc.)

1.3 Motivations and Contributions

1.3.1 Challenges of Drone-Assisted Radio Access Networks

Although the DA-RAN has attracted increasing research attentions [46] and field tests [44], the design and implementation of DA-RAN still face four essential challenges.

1) D2G communication features: In specific, the D2G communication can be further classified into drone-to-user (D2U) and drone-to-BS (D2B) communications. *D2U communication:* The D2U communications focus on the data transmission between drones and their associated terrestrial users. For users in CHs, the D2U communication can directly use cellular bands to serve them, which realizes transparent network access for users. While in scenarios where DCs and ground BSs work cooperatively, to alleviate interference and bring additional resources, the D2U communication is expected to operate in different spectrum from the licensed cellular bands. Currently, the WiFi bands used by commercialized drone

products [25] and the TV White Space [47] are promising candidates for conducting D2U communications. *D2B Communication:* The D2B communications are conducted between drones and BSs. The high capacity and data rate requirements of D2B communication are essential challenges since each D2B link has to relay all the data between the BS and users served by the drones. Cross-layer optimization is required to handle the unexpected wireless fading and interference. For example, dedicated spectrum bands and high MAC priority can be allocated to D2B links; some authentications or routing discovery processes can be simplified or removed to reduce delay. One promising solution for this issue is the mmWAVE technology that can provide up to 20 Gbps data-rate transmission [48].

2) 3D deployment and trajectory designs for multiple DCs: Since the DA-RAN's capability to address terrestrial traffic variations is enabled by DC's dynamic deployment capability, the DC trajectory planning problem, which designs flying traces of DCs to serve terrestrial users in AoIs, is essential for DA-RAN research. On the other hand, the resource allocation decisions of each DC can be adapted to fit diversified user distribution status in different deploying locations. Therefore, the trajectory planning and resource allocation for the DC should be jointly investigated, which forms the DC trajectory planning and resource allocation (TPRA) problem. However, existing studies on DC TPRA research usually investigate the highly simplified and abstracted scenario where a single DC flying on a fixed height, which forms a 2D plane over the ground. To further improve the network performance, the 3D deployment and trajectory design, which fully exploits the dynamic deployment capability of DC, are expected in DA-RAN research. Besides, considering the large scenario with more than one BS and DC, the inter-DC interference must be addressed in the TPRA strategies for multiple DCs.

3) DC energy consumption constraint: Constrained by the limited battery capacity, the LAP drone based DC cannot keep working 24-hour long as the ground BS [49]. The energy constraint of DCs requires appropriate power consumption trade-off between flying control module and communication module. The power consumed by the flying control module is defined as propulsion power, which has been measured through experiments [50]. The results show that three types of flying modes (hovering, horizontal flying and vertical flying) consume same level of energy, while the weight of loads carried by drones impacts power consumption most. Currently, the most endurable DC can keep flying up to 5+ hours [30], which is adequate for temporary communication provision tasks. The power used by the communication module, *a.k.a.* communication power consumption, is usually hundred times smaller than the propulsion power consumption. For instance, a typical rotary-wing drone with 10 km/h flying speed can have 1158.919 Watt propulsion power, while only require around 1 Watt data transmit power, according to the power consumption models used in Chapter 5. Although the power consumption is dominated by

the propulsion power, the transmit power allocation can significantly impact the trajectory planning and deployment of one DC, which is directly related to propulsion power. The joint optimization of both communication and propulsion power consumption should be considered in DA-RAN design.

4) Support high-mobility users The DC TPRA problem has been studied by some pioneer works with the objectives of optimizing DA-RAN throughput or QoS of users [39] [51]. In those works, the trajectory of each DC is modeled as a closed curve composed by discrete 3D locations. The DC sequentially transverses each location and serves associated users to the locations according to scheduled time slots. Given this trajectory model, the terrestrial users or AoIs are assumed to be quasi-static nodes, which simplifies the problem to be solved by optimization methods. However, the assumptions of deterministic trajectory and the quasi-static user model is only applicable for static or low-mobility users with scheduled communications, e.g. data collection for massive Internet of things (IoT) devices [52]. To adapt to the non-static environment with high-mobility users, new DC TPRA approach is expected to make TPRA decisions for DCs according to environment variations [53].

Besides the technical challenges, most countries in the world have published various regulations in terms of the usage of commercial drones [37]. To ensure the legal use of DCs, all DCs in DA-RAN have to be maintained by the DA-RAN operators or governments directly, instead of controlled by the individual users. However, it is inevitable for the operators or governments to control massive DCs by autonomous algorithms or artificial intelligence (AI), such as the algorithms and schemes proposed in this thesis. The new regulations and laws concerning the AI-controlled or machine-controlled DCs are required for future DA-RAN. Nevertheless, in this thesis we focus on the technical challenges faced by DA-RAN.

1.3.2 Approaches and Contributions

To address the four challenges faced by current DA-RAN, in this thesis, we investigate the D2G communication and DC energy consumption in DA-RAN, and propose the multi-DC 3D trajectory planning and resource allocation schemes to serve both quasi-static and high-mobility terrestrial users. In specific, we focus on the following three research topics.

- To design the 3D deployment schemes of multiple DCs for quasi-static AoIs, we analyze the communication coverage, D2B link quality, as well as the optimal flying height of the DC, according to the state-of-the-art D2U and D2B channel models.

Then, a multi-DC 3D deployment problem is formulated, with the objective of maximizing the effectively covered users while maintaining D2U and D2B link qualities. To solve the problem, we propose a per-drone iterated particle swarm optimization (DI-PSO) algorithm. Compared with the pure particle swarm optimization (PSO) based algorithm derived by related works, the large particle searching space and the high violation ratio of D2G communication constraints are prevented by the DI-PSO algorithm. Simulations show that the DI-PSO algorithm can achieve higher user overage ratio with less computing complexity than that of the pure PSO based algorithm.

- To further improve the network performance and the fairness among edge and central users, we study the multi-DC 3D trajectory design in DA-RAN based on the static deployment results of DCs. We formulate the multi-DC 3D trajectory planning and scheduling problem as a mixed integer non-linear programming (MINLP) problem with the objective of maximizing the average D2U throughput. To address the non-convexity and NP-hardness of the MINLP problem caused by the 3D trajectories, we first decouple the MINLP problem into multiple integer linear programming (ILP) or quasi-convex sub-problems in which user association, D2U communication scheduling, horizontal trajectories and flying heights of DCs are respectively optimized. Then, a multi-DC 3D trajectory planning and scheduling algorithm is proposed to solve the sub-problems iteratively based on the block coordinate descent (BCD) method. We further design a k-means-based initial trajectory generation and a search-based start slot scheduling schemes to improve network performance and ensure inter-DC distance constraint, respectively. Compared with the static DBS deployment, the proposed trajectory planning can achieve 10-15 dB reduction on average D2U pathloss, and reduce the D2U pathloss standard deviation by 68%, which indicate the improvements of network throughput performance and user fairness.

- To design trajectories and allocated resource for multiple DCs in the dynamic environment composed by high-mobility users (e.g. vehicular users), we leverages the deep reinforcement learning (DRL) technique, and propose a hierarchical DRL based multi-DC trajectory planning and resource allocation (HDRLTPRA) scheme. To address the high uncertainties of high-mobility users, we decouple the multi-DC TPRA problem into two hierarchical sub-problems, i.e., the higher-level global trajectory planning (GTP) sub-problem and the lower-level local TPRA (LTPRA) sub-problem. In particular, the GTP sub-problem roughly plans trajectories for multiple DCs over a long time period in a large area. A multi-agent DRL based GTP (MARL-GTP) algorithm, which learns the joint trajectory planning policy for multiple DCs to maximize the accumulative number of users being served, is proposed to solve the

GTP sub-problem. Given the global trajectory planning results, the LTPRA sub-problem controls the real-time movement of single DC and resource allocation within its communication coverage. For the LTPRA sub-problem, we further design a deep deterministic policy gradient based LTPRA (DDPG-LTPRA) algorithm executed on each DC independently, which adjusts the DC movements and allocate transmit power over a continuous action space, in response to the real-time user traffic variations. Real-world scenario based simulations show the HDRLTPRA scheme can improve the total achieved network throughput by 40% when compared with the model-based TPRA scheme.

1.4 Thesis Outline

The remainder of the thesis is organized as follows: In Chapter 2, we present a comprehensive review of related works in terms of DC deployment, DC trajectory planning, and joint DC trajectory planning resource allocation in DA-RAN. In Chapter 3, we elaborate the analytical results of DC's communication features, and introduce the DI-PSO algorithm proposed for the multi-DC 3D deployment problem. In Chapter 4, the multi-DC 3D trajectory planning and scheduling problem is formulated and decoupled into sub-problems. Then, a 3D trajectory planning and scheduling algorithm is proposed to solve the sub-problem iteratively, which designs the 3D trajectories and communication schedule for multiple DCs. In Chapter 5, we formulate the multi-DC TPRA problem in highly dynamic and uncertain scenario with high-mobility users. Then, the problem is decoupled into two hierarchical sub-problems, and the MARL-GTP algorithm and the DDPG-LTPRA algorithm are introduced to solve the two sub-problems,respectively. Finally, we conclude the thesis and discuss future works in Chapter 6.

Chapter 2

Literature Review

This chapter aims to introduce the background and related works of DA-RAN research, including the D2G communication models, DC deployment, DC trajectory planning, as well as the joint trajectory planning and resource allocation for DCs.

2.1 D2G Communication Models

As the foundation of wireless communication and RAN research, the wireless channel model is an inevitable element defined in different scenarios or systems. Since the 3D working space and the dynamic deployment capability of drones bring new features to the wireless channel models of both D2D and D2G communications, it is unreasonable to using most of the classic channel models concluded from ground infrastructure based experiments. According to abundant field tests data in different scenarios, some pioneer researchers begin to derive dedicated channel models for drone communications, especially the D2G channel models.

In 2014, Al-Hourani *et al.* built an D2G pathloss model for low altitude platforms including DC [45], which sets the D2G channel model foundation for subsequent research on DA-RAN. In [45], a close-form expression of D2G pathloss model is proposed in which the probabilities of both LoS and NLoS D2G links in different scenarios are considered. Specifically, the LoS probability of D2G link is [45]:

$$P_{\text{LoS}}(r,h) = \frac{1}{1 + a\exp(-b(\arctan(\frac{h}{r}) - a))}, \tag{2.1}$$

where h is the DC flying altitude, r is the horizontal distance between the DC and the user. a and b are constant values determined by environment, such as urban, suburban, rural, etc. Neglecting the antenna heights of users and DCs, the average D2G pathloss can be calculated as follows [45]:

$$\begin{aligned} PL(r,h) = 20\log(\frac{4\pi f_c\sqrt{h^2+r^2}}{c}) \\ + P_{\text{LoS}}(r,h)\eta_{\text{LoS}} + (1-P_{\text{LoS}}(r,h))\eta_{\text{NLoS}}, \end{aligned} \tag{2.2}$$

where f_c (in Hz) is the carrier frequency, c (in m/s) is the speed of light. η_{LoS} and η_{NLoS} are average additional losses for LoS and NLoS links which are environment-dependent. The following Table 2.1 shows the values of parameters $(a, b, \eta_{\text{LoS}}, \eta_{NLoS})$ in different environments.

Table 2.1: D2U Pathloss Parameters for Different Environments

Environment	$(a, b, \eta_{\text{LoS}}, \eta_{NLoS})$
Highrise Urban	$(27.23, 0.08, 2.3, 34)$
Dense Urban	$(12.08, 0.11, 1.6, 23)$
Urban	$(9.61, 0.16, 1, 20)$
Suburban	$(4.88, 0.43, 0.1, 21)$

Fig. 2.1 shows the average D2G pathloss versus the altitude of DC for different r and f_c. In Fig. 2.1, all pathloss curves decrease first then increase slowly with the increasing of altitude. This is because in low altitude space, D2G pathloss is mainly determined by the LoS probability, raising altitude leads to sharp increasing of $P_{\text{LoS}}(r,h)$ and decreases the pathloss level; while in high altitude space, $P_{\text{LoS}}(r,h)$ remains nearly constant for all altitude values, and the pathloss curves are dominated by free space attenuation instead of $P_{\text{LoS}}(r,h)$. Apart from the altitude, horizontal distance r and carrier frequency f_c also influence the D2G pathloss. Increasing r can raise all pathloss values on a curve and change the curve's shape. Increasing f_c impacts the first term in Eq. (2.2), and shifts the whole curve up by a constant value.

As the extension work, in 2017 Al-Hourani *et al.* formulated a D2B pathloss model for suburban scenario based on massive field experiments data [24]. The model indicates a trade-off in the channel performance as the vertical angle between the DC and the BS

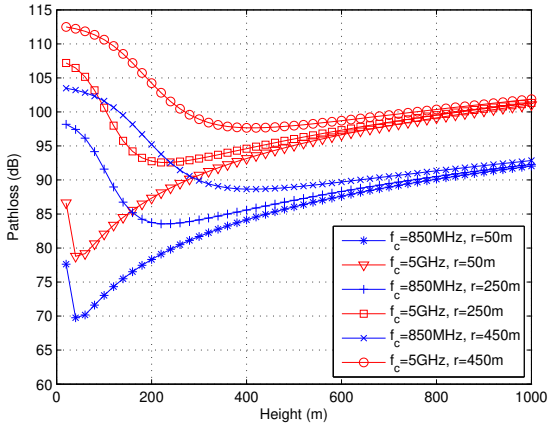


Figure 2.1: D2U pathloss model.

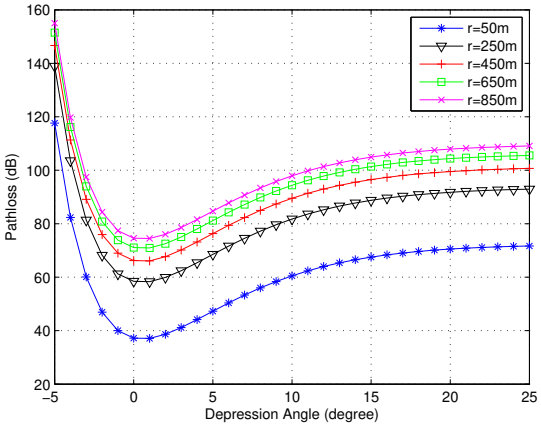


Figure 2.2: D2B pathloss model.

increases. Dominated by LoS links, there is no random factor in the model of average D2B pathloss [24]. The average D2B pathloss is calculated by the cellular-to-UAV pathloss model in [24], which can be expressed as follows:

$$PL(r_{\mathrm{DB}}, \theta) = 10\alpha \log(r_{\mathrm{DB}}) + A(\theta - \theta_0)\mathrm{e}^{(\frac{\theta_0 - \theta}{B})} + \eta_0, \tag{2.3}$$

where r_{DB} is the horizontal distance between the DC and the BS. θ indicates the vertical angle between the DC and the BS in degree. α, A, θ_0, B, and η_0 represent the terrestrial pathloss exponent, excess pathloss scaler, angle offset, angle scaler, and excess pathloss offset, respectively. Except r_{DB} and θ, all other parameters in Eq. (2.3) are constants depending on different environments. For the suburban environment investigated by [24], the values of parameter list $(\alpha, A, \theta_0, B, \eta_0)$ is $(3.04, -23.29, -3.61, 4.14, 20.7)$. Since all experiments and modeling conducted in [24] use 850MHz frequency band that falls into the widely used LTE bands ranging from 700MHz to 900MHz, the carrier frequency is not reflected as one parameter in Eq. (2.3).

Curves of average D2B pathloss versus DC to BS vertical angle under different r_{DB} are shown in Fig. 2.2. All pathloss curves decrease first then increase with the increasing of θ. The minimal pathloss values are achieved around 0° for all curves. Different from the D2G links modeled in [45], the D2B links can keep maintaining LoS condition due to less obstacles between them, so the dominate coefficient that impacts average pathloss is the spatial distance d_{DB} between DC and the BS. For any fixed r_{DB}, d_{DB} that minimizes the free space attenuation can be obtained when θ equals to 0°, which is reflected as the minimal value of the curve.

Note that for both the D2U and D2B pathloss models are large-scale pathloss models with the shadowing effects considered in the environment-based parameters. There is no widely accepted small-scale pathloss models for the D2G communication currently. Besides, the environment are all considered as ideal environment without unusual impactors, such as heavy snows or rains.

2.2 DC Deployment in DA-RAN

The main design objective of DC deployment is to optimize multiple DCs flying altitude, horizontal positions, and/or spatial density to achieve the maximum radio coverage in a given area without violate dedicated constraints, such as interference, D2G channel pathloss, etc. According to the DC deployment results, each BS release multiple DCs hovering over its CHs or dense traffic areas where additional communication resources are required, i.e. traffic congestion road, concerts and sports events, etc. Since the CHs and dense traffic areas are dynamically changing in both spatial and temporal domains, the DCs have to be re-deployed periodically based on their changes.

Recently there have been extensive research efforts to investigate the 3D DC deployment problem. Through field experiments, Dhekne *et al.* demonstrated DCs' capability of improving the signal strength in CHs when they perform as the aerial extensions of BSs [44]. In [1], multi-tier drone networks are introduced to complement terrestrial HetNets, the advancements and challenges related to the operation and deployment of DCs were investigated. To minimize the number of DCs needed to provide radio coverage for a group of distributed ground terminals (GTs), Lyu *et al.* designed a polynomial-time algorithm with successive DC placement, where the macro-BSs are placed sequentially starting on the area perimeter of the uncovered GTs along a spiral path towards the center, until all GTs are covered [54]. Considering the fly-hover-and-communicate scenario where the GTs are partitioned into disjoint clusters sequentially served by the DC as it hovers above the corresponding clusters, He *et al.* jointly optimized the DC's flying altitude and antenna beam-width for throughput optimization in DC downlink multicasting, downlink broadcasting, and uplink multiple access [55]. In [56], Zhang *et al.* studied the spectrum sharing of Drone-Small-Cells network modeled by the 3D Poisson point process, and found the optimal density of DSCs to maximize the network throughput while satisfying the cellular network efficiency constraint.

Leveraging the pathloss model in [45], some researchers focus on exploring the optimal DC deployment that maximizes specific performance metrics. In [51], Mozaffari *et al.* designed a clustering approach to find the optimal trajectories and locations of DCs that

maximize the information collection gain from ground IoT devices. Yang *et al.* proposed a holistic framework using DCs to assist 5G networks in flash crowd traffic scenarios, and design a "first-selfish and second-share" method for DC deployments [57]. In [56], Zhang *et al.* optimized the DC density in DC network to maximize the network throughput while satisfying the efficiency requirements of the cellular network. Bor-Yaliniz *et al.* formulated the 3D placement problem for single DC as a Mixed-Integer Nonlinear Programming (MINLP) problem, and solve it through bisection search algorithm [58]. An optimal DC placement algorithm maximizing the number of covered users with minimum power consumptions was designed by Alzenad *et al.*, in which the DC deployment problem is decoupled in the vertical and horizontal dimensions and solved respectively [59]. Zhou *et al.* studied the downlink coverage features of DC using Nakagami-m fading models, and calculated the optimal height and density of multiple DCs to achieve maximal coverage probability [60]. Considering delay-tolerant and delay-sensitive users, Kalantari *et al.* proposed an algorithm to find efficient DC deployment, user-DC associations and wireless backhaul bandwidth allocations to maximize the sum logarithmic rate of the users in a heterogeneous network. As the state-of-the-art work, even the Q-learning based approach was leveraged by Ghanavi *et al.* to find the optimum deployment position for the DC. Simulation results show this method provides an effective placement strategy which increases the QoS of wireless networks [61].

Various constraints have to be considered in the optimization of static DC deployment, i.e. Drone-to-User (D2U), Drone-to-BS (D2B) link qualities, battery capacity of DC, inter-DC interference, DC to terrestrial RAN interference, etc. Among them, the most essential one is the D2B link quality constraint, which is simplified or ignored by most works. Although the D2B link quality constraint is unnecessary under the assumption that the DC or DC swarm are isolated with terrestrial networks, in reality most services supported by DC still have to communicate with the core network enabled by terrestrial networks. Since the ultimate purpose of DSB is relaying data between terrestrial BSs and users in AoI, it is inevitable to ignore D2B communication constraints in the DC deployment research, especially for the DA-RAN scenario. Considering the D2B backhaul links, Kalantari *et al.* further explored the deployment of multiple DCs, and find the minimum number of DCs for dedicated user coverage constraint by using swarm intelligence based heuristic algorithm [62]. As the extension work of [62], the optimal 3D backhaul-aware placement of a DC in 2 different approaches, namely network-centric and user-centric, was found [63]. In [64] and [65], a swarm of DCs were regarded as backhaul/fronthaul hubs for small-cells via free-space-optics /mmWave links by Shah *et al.*, in which the DC deployment and association is jointly optimized through heuristic methods. In [66], Fouda *et al.* introduced DCs into the in-band integrated access and backhaul (IB-IAB) scenarios for 5G networks, and proposed the algorithm for DCs deployment. However, the D2B backhaul channel models used in

those works are either as same as the D2U pathloss model [65] or traditional terrestrial channel models [66]. In this thesis, we further implement the specific D2B channel model derived in [24] to highlight the D2B channel features.

2.3 DC Trajectory Planning in DA-RAN

Since the distribution of AoI changes dynamically in both spatial and temporal domains, the DCs have to be re-deployed periodically according to the dynamic distribution of AoI. In static DC deployment, the AoI located at the edge of the DC's radio coverage suffer relatively high pathloss comparing with the AoI located at the center of the DC's coverage. Although more strict D2U link quality constraint can be applied to minimize the pathloss difference between edge and center AoI, more DCs are required to ensure the same level of coverage ratio, which increases the deployment cost.

Comparing with the DC deployment in which DCs hovers as relatively stationary nodes, the DC trajectory planning is proposed in a mobile scenario where DCs keeps flying over multiple BSs or areas of interests periodically following the designed trajectory. To promote the fairness for all AoIs and maintain low deployment cost, specific DC can periodically flying over multiple AoIs that cannot be covered by the DC simultaneously, and provide communication servers for AoIs within its instantaneous radio coverage. The purpose of DC trajectory planning is to design an optimal trajectory for each DC which traverse all required AoIs to maximize QoS gains without violating any constraints. Comparing with statically deploying DCs and ignoring the uncovered AoIs, though some AoIs among the trajectory can suffer QoS lost in outage intervals where no data transmission occurs due to uncovered by DC, the overall QoS can be enhanced by leveraging trajectory planning.

The key of trajectory planning is to optimize two essential metrics: 1) the flying trajectory of the DC, and 2) the flying schedule of the DC within one period. Fig. 2.3 shows an scenario of the multi-DC trajectory planning problem. Within the radio coverage of one central BS, twenty AoIs are associated to four DCs, which is denoted as squares with different colors. The dotted-line with different colors represent the optimized trajectories for each DC respectively. As shown in Fig. 2.3, the DC trajectory usually exactly traverses over all AoIs to obtain optimum D2U channel conditions to each DA. However, noticing that the red and magenta trajectories both avoid traversing one AoI associated to it, this is caused by applying inter-DC interference constraints. Comparing with the DC deployment approach statically deploying each DC over the center of its radio coverage scope (denoted as dash-line), the multiple DC trajectories planning approach can periodically serve each AoI with optimum D2U channels, which improves overall QoS of all AoIs associated to

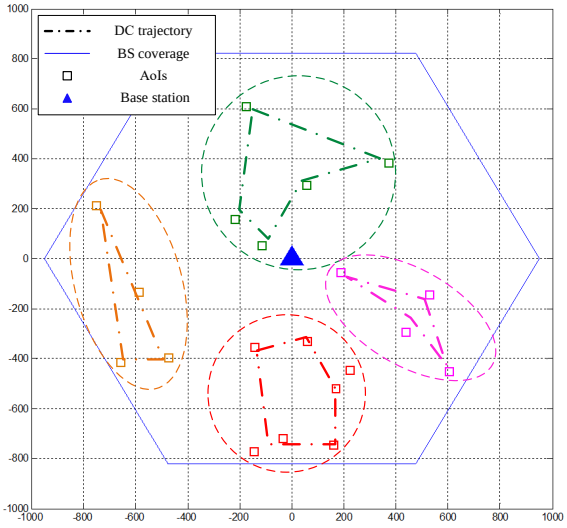


Figure 2.3: Overview of multi-DC trajectories planning problem.

each DC, and further enhances the performance of whole DA-RAN.

The trajectory of each DC is usually modeled as follows. Considering one arbitrary DC serves its AoIs via a periodic/cyclical TDMA manner with each period/cycle of duration denoted by T. Each period T is discretized into N equal-time slots, indexed by $n = 1, ..., N$. The elemental slot length $\delta_t = \frac{T}{N}$ is chosen to be sufficiently small to describe the minimal status-change for the DC. Based on this model, the DC trajectory over T can be approximated by the N two-dimensional sequences:

$$\mathrm{p}[n] = [x[n], y[n]]^T, \quad n = 1, ..., N \tag{2.4}$$

where $x[n]$ and $y[n]$ denote the horizontal coordinates of the DC at slot n. Two trajectory constraints are considered in our work: 1)Each DC needs to return to its initial location by the end of each period T such that users can be served periodically, which implies that the trajectory of each DC has to be a closed curve; 2) The instantaneous speed of DC cannot exceed the speed threshold $V_{\max}$. The two constraints can be represented as follows:

$$\mathrm{p}[1] = \mathrm{p}[N], \tag{2.5}$$

$$\|\mathrm{p}[n+1] - \mathrm{p}[n]\|^2 \leq \|V_{\max}\delta_t\|^2, \quad n = 1, ..., N-1. \tag{2.6}$$

According to 2.4, the distance from the DC to AoI i in time slot n can be expressed as:

$$d_i[n] = \sqrt{H^2 + \|\mathrm{p}[n] - \mathrm{w}_i\|^2}, \tag{2.7}$$

where H is the flying height of the DC, w_i represent the coordinates of AoI i.

Given the trajectory model, we can further schedule the N slots in T to serve different AoIs. For each DC, the scheduling process is modeled by defining a binary decision variable $s_i[n]$, which indicates AoI i is served by the DC in time slot n if $s_i[n]$ equals to 1; otherwise $s_i[n]$ equals to 0.

In most literatures, they consider the scenario where a dedicated DC periodically flies along the closed trajectory, therefore the time for one DC traverses the whole trajectory back to the original point is defined as one period of the trajectory. To improve the QoS gains, DC can have different status at specific points along the trajectory, for instance, large number of works allow DC hover dedicated time on some key points where the maximum D2G throughput can be obtained; while for the connecting paths between two key points where high D2G pathloss is suffered, the DC can increase flying speed to save the traversing time. Since that, the hovering and flying intervals corresponding to different DC status over one period have to be optimized, which is referred to as the flying schedule of DC trajectory. Existing work usually joint optimize both trajectory and flying schedule for one DC.

In [67], Li *et al.* proposed an cooperative relaying scheme in which multiple DCs relay data from terrestrial sensors to the BS using time division multiple access (TDMA). As a pioneer work, the UAVs trajectories are assumed to be pre-determined and not optimized, which simplifies the design to a DC-packet matching problem. Mozaffari *et al* studied both the static and mobile DC-enabled wireless networks underlaid with a device-to-device communication network [68]. Though the trajectory optimization is considered in this work, the D2U communications are only permitted at pre-defined stop points, which fails to exploit the impact of DC mobility feature on the network performance. Motivated by [68], Zeng *et al.* proposed a general framework for joint trajectory and communication optimization in D2U point-to-point communication scenario [69]. The trajectory designs in [70] and [69] can be considered as a generalization of the DCs trajectory planning problem, subject to practical constraints on the DCs mobility, such as its initial/final locations, maximum speed and acceleration, etc.

Besides the constraints from DC's mobility, some researchers further investigate the DC trajectory planning problem with general QoS constraints. For UAV-enabled multiuser communication networks, Lyu *et al.* proposed a novel cyclical multiple access scheme in [71], where the UAV periodically serves each of the ground users along its cyclical trajectory via TDMA. To maximize the minimum downlink throughput over ground users, Wu

et al. formulated a mixed integer non-convex optimization problem in which the multi-user communication scheduling and association are jointly optimized with DCs trajectories and power control schemes [39]. Block coordinate descent and successive convex optimization techniques are used for solving the problem. Considering the delay constraints, Wu *et al.* further studied a DC-enabled orthogonal frequency division multiple access (OFDMA) network where a DC is dispatched to serve a group of delay-sensitive users on the ground [72]. The DC trajectory and OFDMA resource allocation are jointly optimized to maximize the minimum average throughput of all users by leveraging an iterative parameter-assisted block coordinate descent method. In [73], Zhang *et al.* investigated the security issue in DC trajectory planning, which implements physical layer security technique into trajectory design and power control of DC to realize secure D2G communications. Even the proactive caching technique has been applied in DC trajectory planning. in [74], Xu *et al.* revealed the fundamental trade-off between the file caching cost, which is the total time required for the DC to transmit files to designated caching ground nodes, and the file retrieval cost, which is the average time required for serving one file request. A heuristic algorithm is proposed to jointly design the file catching policy, the DC trajectory and the communication scheduling, which characterizes the trade-off. To dynamically deploy multiple DCs while maintaining the connectivity among them, Zhao *et al.* proposed both the centralized and distributed DC motion control algorithms for scenarios with or without global information of users [75]. As the pioneer works of trajectory planning, [39] [72] set the foundation models for DC trajectory planning. However, in those works the flying height of all DC are treated as a pre-defined constant, and most of them idealize the D2B link quality constraints. In this paper, not only the D2B link quality constraint is introduced in problem formulation and optimization, but also the flying heights of each DC at every slots are jointly optimized with the horizontal trajectory.

2.4 Joint DC Trajectory Planning and Resource Allocation in DA-RAN

Given the trajectory planning results, the resource allocation decisions of each DC should be adapted to fit the user distribution status in different deploying locations. Therefore, the trajectory planning and resource allocation for the DC should be jointly investigated, which forms the DC trajectory planning and resource allocation (TPRA) problem. The resources can be allocated in DA-RAN includes the DC communication resource (bandwidth or physical resource blocks (RBs)), computing resource (CPU), caching resource (on-broad memory or cache), as well as the energy consumptions (transmit power, propulsion power of

DC), etc. Considering different user scenarios, two types of TPRA schemes are investigated by existing works, i.e., the model-based schemes that calculate the optimal TPRA decisions for quasi-static AoIs or users, and the learning-based schemes that dynamically determine the movements and resource allocations of DCs according to the variations of terrestrial user traffic.

2.4.1 Model-based Trajectory Planning and Resource Allocation

As the pioneer works of model-based DC TPRA research, [39] and [51] define the initial TPRA problem in which the DC periodically serves quasi-static terrestrial users through a discrete 3D trajectory. The model-based TPRA solution has been investigated in various scenarios including air-ground integrated communication assistance [34] and IoT data collections [76]. The model-based DC TPRA is generally jointly optimized with other impact factors such as DC power consumption [77], DC altitude and speed [78], and number of DCs [52], etc. Although the discrete trajectory and quasi-static user models simplify the optimization process, the accuracy of the TPRA solutions is inevitable decreased.

2.4.2 Learning-based Trajectory Planning and Resource Allocation for Mobile Users

To address the non-stationary environment with model-free methods, the learning-based approaches are proposed for the DC TPRA problem. Most learning-based TPRA works use the DRL framework.

Reinforcement learning (RL) is a learning process in which the learning agents sequentially interact with the environment, and automatically adjust their policies according to the feedback from the environment [79]. All RL processes can be modeled as the Markov decision process (MDP) [5]. At each time step, the current environment is observed by each agent as a certain state, and each agent selects a certain action according to its policy. Then, the environment transits into a new state jointly determined by the previous state and agents' actions. Each agent can received a reward from the environment at each step, which quantifies how well the action taken in that step given the specific state. Different to optimization and supervised learning approaches, RL does not rely on mathematical models of the environment or massive labeled data. In RL, the agents can automatically learn environment from their past experiences, and converge to stable policies.

One basic RL method widely used in learning based DC TPRA research is the Q-learning [80]. The agent in Q-learning selects actions based on the Q-table. Each Q-value

in the Q-table represents the estimated reward for the agent by executing the action at the state, under the current policy. The action associated with the largest Q-value is selected by the agent under each state. At each time step, the agent updates the Q-table based on the observed reward, therefore, the policy is also updated automatically. The policy will eventually converge to the optimal policy. Given an agent having A potential actions, and total S states in the environment, the Q-table contains $A \times S$ Q-values. Considering the high searching complexity in large action or state space, the Q-learning method is only applicable for scenarios composed by discrete action and state spaces with limited dimensions [81].

To address the high-dimensional state space, deep neural networks (NN) are utilized in Q-learning methods instead of traditional Q-table, which forms the DRL method [81]. At each time step, the agent inputs the feature vector representing the current state into the deep NN that estimates the Q-value for each action. The agent then selects the action with the largest estimated Q-value, and stores the experience including the state transition and reward into a replay buffer, which is used to train the deep NN for more accurate estimations. To implement deep reinforcement learning, the agent needs to store the deep NN and the replay buffer which stores its previous experiences. Generally, to obtain better results, the sizes of the deep NN and the replay buffer need to be large, which results in a high memory requirement. The training of the deep NN requires the gradients of estimated Q-values with respect to the parameters of the deep NN. This leads to a high computation complexity given a large DQN. Therefore, DRL methods require more memory and higher computational complexity than other RL approaches, such as Q-learning. Considering the relatively long training time, the DRL methods are usually trained in an off-line manner. For example, an agent can upload its experiences to a server to train the deep NN, then uses the well-trained NN for real-time data.

To extend the DRL methods into scenarios with high-dimensional action spaces, the actor-critic learning is proposed [82]. In actor-critic learning, the agent is consist of two roles, i.e. a critic and an actor. The actor select actions based on a probability distribution function (PDF) over the action space. The critic observes the states and rewards from the environment and evaluates the state value, i.e. the expected total reward that will be received in the future passing through the state, which can be considered as a state value function. The critic is used to improve the efficiency and stability for the training of the actor in term of optimal action selection. After each time step, the critic updates the state value based on the observed reward. Then, the actor updates its policy for the previous state towards the direction which maximizes the expected rewards. Compared with Q-learning methods, the actor-critic learning does not select action with the highest expected reward from the action space. Instead, the action is selected randomly following the PDF

learned by the actor, which decouples the complexity of action searching with the sizes of the action space. Therefore, the actor-critic learning can efficiently address large or continuous action spaces.

To leverage the DRL framework in DA-RAN, the DC takes TPRA actions according to the observed environment state, and then receives reward in each step. The policy of choosing TPRA actions is updated step by step to reach the convergence, with the objective of maximizing the long-term accumulative reward. In [43], Chen *et al.* studied the proactive deployment of DCs and content caching for optimizing the quality-of-experience (QoE) of wireless devices in the cloud RAN. Given the predicted behaviors of users, the proposed approach seeks to find the user-DC associations, the optimal DCs' locations, and the contents to cache at DCs. A novel algorithm based on the machine learning framework of echo state networks (ESN) is proposed to solve the problem. As the extension of [43], Chen *et al.* further customized the ESN based algorithm for Virtual Reality contents caching in DA-RAN [83], and propose a resource and cache management approaches for DA-RAN based on Liquid State Machine (LSM) learning [84]. In [85], the deterministic policy gradient (DPG) learning method is leveraged in single DC trajectory planning to maximize user throughput. An DRL based interference-aware trajectory planning scheme is proposed for single DC in [86], which achieves low D2U latency and high throughput. The authors in [87] design a relay scheme of single DC, which integrates both Q-learning and DRL to minimize both bit error rate of relayed signal and the DC power consumption. Considering multi-DC scenario, a decentralized DRL framework is proposed in [88] to solve multi-DC trajectory design problem based on the a sense-and-send protocol. Given the power consumption constraints of DC, the authors of [89] design a DRL-based energy-efficient trajectory planning method for fair communication coverage of multiple DCs. In [90], the DRL based computation tasks offloading is proposed for DC assisted IoT scenario. Considering the impact of user mobility, in [91], a DRL based DC deployment algorithm is proposed for multiple DCs serving mobile users, with the objective of maximizing the sum mean opinion score (MOS) of ground users. Although various impact factors are considered by existing DRL based TPRA research, the multi-DC environment, non-fixed user number and locations, as well as the joint trajectory planning and resource allocation of DC have never been considered simultaneously. In this work, we will propose the HDRLTPRA scheme to solve the multi-DC TPRA problem in high-mobility scenarios.

Chapter 3

Drone-Cell Communication Analysis and Deployment in DA-RAN

In this chapter, we propose a DA-RAN architecture in which DCs are leveraged to relay data between base stations and users. Based on the state-of-the-art D2U and D2B channel models, we first analyze the user coverage and the D2B backhaul connection features of DCs. We then formulate the 3D DC deployment problem with the objective of maximizing the user coverage while maintaining D2B link qualities, for a given number of DCs being deployed. To solve the problem, the PSO algorithm is leveraged for its low computational cost and unique features suiting the spatial deployment of DCs. We propose a DI-PSO algorithm that optimizes DC deployments for different DC numbers, and prevents the drawbacks of the pure PSO based algorithm derived from related works. Simulations show that the DI-PSO algorithm can achieve higher user coverage ratio with less complexity comparing to the pure PSO based algorithm.

3.1 Background and Motivations

Providing ubiquitous connectivity for users and devices with diversified service requirements is regarded as one of the key challenges in 5G networks [1]. To support the reliable and low-latency access of massive mobile users, considerable level of flexible deployment is required for future RAN. However, current BSs and RRHs are deployed in certain geographical locations according to long-term traffic behaviors with little flexibility to be re-deployed. Such rigid RANs are reluctant to maintain ubiquitous connectivity for most 5G scenarios where dynamic data traffic occurs in both spatial and temporal domains [34].

Though densely deploying BSs or RRHs is one intuitive way to improve the user coverage of RAN, the high expenditure and low efficiency brought by this method are unacceptable for RAN operators [1]. To enhance RAN's flexibility for supporting massive dynamic connections, the emerging DC communication technology is a promising solution.

Promoted by the well-developed flying control technologies and various commercial drone products, both academia and industry are devoting increasing number of efforts on the drone communication research [39]. Equipped with specific wireless modules and controlled by corresponding controllers (e.g. edge servers on BSs), flying drones can perform as DCs to provide temporal and on-demand communication services for areas of interests [13]. Comparing with legacy BSs, there are two advantages using DCs:

1) Line-of-Sight Connection: Compared with ground BSs, DCs flying in the air have higher probability to connect ground users via LoS links, which facilitates highly reliable communications [36]. This advantage is further enhanced by DC's mobility feature that allows 3D adjustments of the DC position to avoid obstacles between D2U links [37]. For the D2B links, since the flying height of DCs can be close to the height of BS antennas, the D2B connections are naturally LoS with little probability to be blocked by tall buildings [24].

2) Dynamical Deployment: Different from traditional BSs which are statically fixed on dedicated locations, drones can be dynamically deployed according to real-time requirements, and allocated to different users or controllers on demands. Two types of control methods are considered in DC deployments: 1) All DCs connect with corresponding central BSs taking charge of deployment control, which is similar to the C-RAN architecture while the RRHs functions are performed by flying DCs [43]; and 2) A swarm of DCs form a FANET and negotiate with each other to determine the deployment results. This method is mainly used in scenarios lacking of infrastructures, such as post-disaster communication recovery [92].

Although various approaches are proposed to optimize the deployment of DCs, the D2B communication is ignored or idealized in many works. Since the ultimate purpose of introducing DCs into RAN is enhancing users' accessibility to network services, it is inevitable to consider D2B communications and ensure their reliability in DC deployments. Therefore, the 3D DC deployment problem is not only affected by user distributions, but also constrained by the qualities of D2B links. In this chapter, we investigate the spatial deployment problem for multiple DCs, considering the D2B link constraint. The main contributions of this work are listed as follows.

- We propose a general framework of DA-RAN, in which the multi-DC 3D deployment problem is formulated to maximize the ratio of effectively covered users given a

dedicated number of DCs. The quality of D2B links, the maximal user number supported per DC, and the multi-DC interference to users are considered in the problem.

- Based on the D2U pathloss model in [45] and the D2B pathloss model in [24], we analysis the effective D2U communication coverage, D2B link quality, as well as the optimal flying height of the DC via stochastic method.
- To address the NP-hardness of the proposed multi-DC 3D deployment problem, we propose a pure PSO based algorithm by customizing heuristic algorithms in related works into the DA-RAN scenario. To further improve the heuristic's performance, a DI-PSO algorithm is designed to find the sub-optimal deployments corresponding to different numbers of DC respectively. Compared with the pure PSO based algorithm, the DI-PSO algorithm can achieve higher user coverage ratios with less iteration times.

The remainder of this chapter is organized as follows. In Section 3.2 the DA-RAN framework and the system model are presented, followed by the stochastic analyses of D2G communications in Section 3.3. The DC deployment problem is formulated in Section 3.4, with both pure PSO and DI-PSO algorithms being proposed in Section 3.5. Numerical results are presented in Section 3.6. The summary of the preliminary work is given in Section 3.7.

3.2 System Model

3.2.1 DA-RAN Network Model

The framework of DA-RAN is shown in Fig. 3.1. Similar to the C-RAN architecture, DCs perform as aerial RRHs that connect with their corresponding BSs. For each BS in DA-RAN, a swarm of DCs are deployed by it over the AoIs where users cannot have effective connections with the BS. Two typical types of AoI are TBSs and CHs. In TBSs, massive data traffic between users and the BS is generated at same time, which cannot be supported by the inadequate PRBs simultaneously. In CHs, effective user-to-BS links are blocked by obstacles, such as high buildings. Since appropriately deployed DCs can maintain reliable LoS D2U and D2B links, the unserved users in AoIs can communicate with the BS through accessing DCs which act as relays. According to the spatial and temporal variations of AoIs, the deployments of DCs can be adjusted by the BS in a flexible way, which enhances RAN's capability to cope with dynamic traffic.

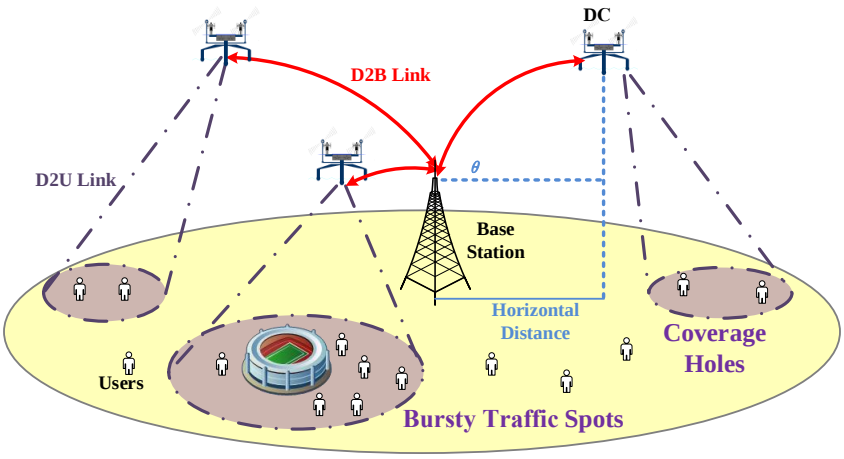


Figure 3.1: Drone assisted radio access networks.

The DA-RAN involves three types of links: U2B links, D2U links and D2B links.

U2B links: U2B links are classic Up/Down links between users and BSs without DC's involvement. In DA-RAN, U2B links co-exist with D2U and D2B links in AoIs without being interfered by them. There is no effective U2B link for users in CHs, while parts of users in TBSs can access BSs via U2B links.

D2U links: D2U links connect DCs and users in AoIs. To alleviate interference and bring additional resources for users in TBSs, the D2U links are expected to operate in different spectrum from the licensed U2B bands. Currently, the TV White Space [47], cognitive radio [93], and the WiFi bands used by commercialized drone products [25] are candidates for conducting D2U communications.

D2B links: DCs communicate with corresponding BSs through D2B links. Though the LoS feature of D2B links, due to the same level of DC flying heights and BS antenna heights, guarantees the reliability, the capacity of D2B links remains challenging since each D2B link has to relay all the data between the BS and users covered by the DC. One promising solution for the capacity issue is the mmWave technology that can provide up to 20Gbps data-rate transmission[48]. It is appealing and feasible to employ mmWave into D2B communications due to the following reasons. First, DCs are expected to hover on fixed position, and maintain the quasi-static status to corresponding BSs during the interval between adjacent re-deployments. The frequent re-directing of beams can be avoided, which results in more time for effective mmWave data transmissions [94]. Second, the ideal transmission environment of mmWave technology is the LoS links, which are naturally

supported by D2B communications [95]. Meanwhile, the MAC protocol of D2B links can be customized to promote performance [96]. For instance, the authentication process can be simplified to reduce latency [97].

Impacted by the flying height and mobility of DCs, both D2U and D2B links in DA-RAN have the unique channel features, which cannot be appropriately modeled by the common U2B pathloss model. In this chapter, state-of-the-art DC link pathloss models are leveraged to analyze D2U and D2B links. Specifically, the D2G pathloss model proposed in [45] is used to analyze D2U links, while the D2B pathloss model derived in [24] is leveraged to analyze D2B links.

3.2.2 User Distribution Model

Based on preceding DC pathloss models, we investigate the 3D deployment of multiple DCs corresponding to one BS. Given the 3D space, the BS is fixed at the origin point (coordinate $(0, 0, 0)$) with ground radio coverage radius R_{bs}. The ground radio coverage area of the BS on X-Y plane $|\mathcal{A}|_{\text{bs}}$ is modeled as a mesh that consists of multiple grids in square shape. The area of each grid is denoted as $|\mathcal{A}|_{\text{da}}$ with the side-length of $\sqrt{|\mathcal{A}|_{\text{da}}}$. By dividing $|\mathcal{A}|_{\text{bs}}$ into multiple grids densely, the side-length of each grid is far smaller than R_{bs} and the DC flying height h, so the average D2U pathloss of different users in one grid can be considered as equal. Without loss of generality, in this chapter, we normalize all D2U links within one grid to suffer the same D2U pathloss between the DC and the center of the grid. Specifically, each AoI is uniformed as a $20m \times 20m$ grid. Given the assumption that traffic is uniformly distributed in space and independent with each other, dedicated number of grids are randomly chosen as AoIs. For an arbitrary grid, the probability of being a AoI can be calculated as follows:

$$p_{\text{da}} = \frac{E_{\text{da}}}{N_{\text{g}}}, \tag{3.1}$$

where E_{da} is the average AoI number calculated through statistic, N_{g} is the total grid number within $|\mathcal{A}|_{\text{bs}}$. Based on (3.1), AoIs are uniformly distributed over the $|\mathcal{A}|_{\text{bs}}$. We assume that AoIs change their size and location with a low frequency, so each DC deployment can treat the random distributed AoIs as a quasi-static scenario where no spatial change occurs during the interval between adjacent re-deployments. Based on the current snapshot of AoIs distribution, the BS deploys DCs over AoIs to maximize user coverage, and re-deploys them when the AoIs distribution changes.

3.3 Theoretical Analyses of DC Deployment

In this part we analyze the user coverage and working zone of DCs in DA-RAN. The main notations used in the analyses are summarized in Table. 3.1.

Table 3.1: Summary of Main Notations

Notations	Descriptions
h	flying height of the DC
r	horizontal distance between one DC and one AoI
$\|\mathcal{A}\|_{\text{bs}}$	radio coverage area of the BS on X-Y plane
R_{bs}	radio coverage radius of the BS
$\|\mathcal{A}\|_{\text{da}}$	area of each grid
p_{da}	probability of being a AoI for an arbitrary grid
N_{g}	total grid number within the radio coverage of BS
λ	user density
u_{da}	number of user locating in an arbitrary AoI
$R_{\text{dc}}(h)$	effective radio coverage radius of one DC at height h
ξ_{LoS}	D2U LoS probability threshold
γ_{DU}	D2U free space pathloss threshold
$D_{\text{dc}}(h)$	average AoI number covered by one DC at height h
$U_{\text{dc}}(h)$	average user number covered by one DC at height h
$\mathbb{U}_{\text{Udc}}(h)$	upper bound of $U_{\text{dc}}(h)$
$\mathbb{L}_{\text{Udc}}(h)$	lower bound of $U_{\text{dc}}(h)$
H_{opt}	optimal DC height for maximizing $U_{\text{dc}}(h)$
r_{DB}	horizontal distance between DC and the BS
N_{dc}	number of DCs
N_{da}	number of AoIs
C_{dc}	Capacity of one DC
R	minimal data rate required by each user
γ_{DB}	D2B pathloss threshold
N_{Imax}	maximal multi-DC interference suffered by one AoI

3.3.1 Stochastic Analyses of DC Coverage

Without loss of generality, we assume that the user distribution follows a 2D Poisson Point Process (PPP) over the $|\mathcal{A}|_{\text{bs}}$ with a user density λ. Due to the independence of each PPP point, the number of user locating in an arbitrary AoI u_{da} follows a Poisson distribution:

$$P(u_{\text{da}}) = \frac{(\lambda|\mathcal{A}|_{\text{da}})^{u_{\text{da}}}}{u_{\text{da}}!}\text{e}^{-\lambda|\mathcal{A}|_{\text{da}}}. \tag{3.2}$$

Since NLoS D2U links are not capable of supporting effective data transmissions in reality, the effective radio coverage radius of the DC R_{dc} cannot be determined by directly using (2.2) with NLoS links. By constraining both LoS probability and free space pathloss, the refined DC radio coverage constraint is expressed as:

$$\begin{cases} P_{\text{LoS}}(r,h) > \xi_{\text{LoS}} \\ \frac{4\pi f_c\sqrt{h^2+r^2}}{c} < \gamma_{\text{DU}} \end{cases} \tag{3.3}$$

Since the $P_{\text{LoS}}(r,h)$ is a decreasing functions of r while the free space pathloss function is an increasing functions of r, the function with a minor return value determines R_{dc}. For a fixed height h, R_{dc} can be expressed as a function of h through rewriting (3.3) :

$$R_{\text{dc}}(h) = \min\big(\frac{h}{\tan(a - \frac{1}{b}\ln\frac{1-\xi_{\text{LoS}}}{a\xi_{\text{LoS}}})}, \sqrt{(\frac{c\gamma_{\text{DU}}}{4\pi f_c})^2 - h^2}\big), \tag{3.4}$$

where the min function returns the minor one of its two items, the former reflects the LoS probability constraint, while the later represents the free space pathloss constraint. Given $R_{\text{dc}}(h)$ and p_{da}, the average number of AoIs covered by one DC flying at height h can be calculated as:

$$D_{\text{dc}}(h) = p_{\text{da}}N_{\text{g}}\frac{|\mathcal{A}|_{\text{da}}}{|\mathcal{A}|_{\text{bs}}} = p_{\text{da}}N_{\text{g}}\frac{R_{\text{dc}}^2(h)}{R_{\text{bs}}^2}. \tag{3.5}$$

In following analyses, we use R_{dc} and D_{dc} to represent the returns of $R_{\text{dc}}(h)$ and $D_{\text{dc}}(h)$ for a given h, respectively.

Assume that k AoIs occurs in the $|\mathcal{A}|_{\text{da}}$ of an arbitrary DC. Based on (3.2), u_{dc} users can be covered by the DC with the probability of:

$$P(u_{\text{dc}}) = \frac{(\sum_{i=0}^{k} p_i\lambda|\mathcal{A}|_{\text{da}})^{u_{\text{dc}}}}{u_{\text{dc}}!}\text{e}^{-\sum_{i=0}^{k} p_i\lambda|\mathcal{A}|_{\text{da}}}, \tag{3.6}$$

where p_i indicates the LoS link probability between AoI i and the DC, which is calculated by (2.1). $p_i\lambda$ represents the number of users effectively covered by the DC in AoI i.

According to the attributes of Poisson distribution, the average number of users covered by the DC $\mathrm{E}(u_{\mathrm{dc}})$ is:

$$\mathrm{E}(u_{\mathrm{dc}}) = \sum_{i=0}^{k} p_i \lambda |\mathcal{A}|_{\mathrm{da}}. \tag{3.7}$$

Since the values of p_i are affected by the horizontal and vertical distances between AoI i and the DC, even for a given flying height h and the corresponding R_{dc}, p_is of different AoIs vary. On the other hand, k in (3.7) is a random variable whose distribution is impacted by R_{dc}. Therefore, we have the following Proposition 1.

Proposition 1. *Given the DC flying at height h, the average number of effectively covered users U_{dc} can be calculated as follows:*

$$U_{\mathrm{dc}}(h) = \int_0^{R_{\mathrm{dc}}} \frac{2r\lambda |\mathcal{A}|_{\mathrm{da}} p_{\mathrm{da}} N_{\mathrm{g}}}{R_{\mathrm{bs}}^2 (1 + a\exp(-b(\arctan(\frac{h}{r}) - a)))} \,\mathrm{d}r. \tag{3.8}$$

Proof. Because the value of p_i is determined by the location of AoI i for given h, p_i turns to be an i.i.d random variable for all AoIs. Given the fact that p_i and k are independent, we have:

$$\begin{aligned} U_{\mathrm{dc}}(h) &= \mathrm{E}(u_{\mathrm{dc}}) = \mathrm{E}(p_i \lambda |\mathcal{A}|_{\mathrm{da}}) \mathrm{E}(k) \\ &= \mathrm{E}(p_i \lambda |\mathcal{A}|_{\mathrm{da}}) p_{\mathrm{da}} N_{\mathrm{g}} \frac{R_{\mathrm{dc}}^2}{R_{\mathrm{bs}}^2}. \end{aligned} \tag{3.9}$$

In (3.9), the expectation of k equals to the average number of AoIs covered by one DC D_{dc} calculated through (3.5). The remaining task is calculating $\mathrm{E}(p_i \lambda |\mathcal{A}|_{\mathrm{da}})$. For the fixed h, p_i can be treated as a function of r, and its PDF is dominated by r's distribution. Leveraging the geometric features of r, the PDF of r can be expressed as follows:

$$P(r) = \frac{\mathrm{d}C(r)}{\mathrm{d}r} = \frac{\mathrm{d}}{\mathrm{d}r}(\frac{\pi r^2}{\pi R_{\mathrm{dc}}^2}) = \frac{2r}{R_{\mathrm{dc}}^2} \qquad r \in [0, R_{\mathrm{dc}}], \tag{3.10}$$

where $C(r)$ is the cumulative distribution function (CDF) of r. Based on (3.10), the $\mathrm{E}(p_i \lambda |\mathcal{A}|_{\mathrm{da}})$ is calculated:

$$\begin{aligned} &\mathrm{E}(p_i \lambda |\mathcal{A}|_{\mathrm{da}}) = \lambda |\mathcal{A}|_{\mathrm{da}} \mathrm{E}(p_i(r)) \\ &= \lambda |\mathcal{A}|_{\mathrm{da}} \int_0^{R_{\mathrm{dc}}} \frac{1}{1 + a\exp(-b(\arctan(\frac{h}{r}) - a)} P(r) \,\mathrm{d}r \\ &= \int_0^{R_{\mathrm{dc}}} \frac{2\lambda |\mathcal{A}|_{\mathrm{da}} r}{R_{\mathrm{dc}}^2 (1 + a\exp(-b(\arctan(\frac{h}{r}) - a))} \,\mathrm{d}r. \end{aligned} \tag{3.11}$$

With both $\mathrm{E}(p_i\lambda|\mathcal{A}|_{\mathrm{da}})$ and $\mathrm{E}(k)$ being derived, the expression of $U_{\mathrm{dc}}(h)$ can be obtained through substituting (3.11) into (3.9):

$$\begin{aligned} U_{\mathrm{dc}}(h) &= p_{\mathrm{da}} N_{\mathrm{g}} \frac{R_{\mathrm{dc}}^2}{R_{\mathrm{bs}}^2} \times \\ &\int_0^{R_{\mathrm{dc}}} \frac{2\lambda|\mathcal{A}|_{\mathrm{da}} r}{R_{\mathrm{dc}}^2(1 + a\exp(-b(\arctan(\frac{h}{r}) - a)))}\,\mathrm{d}r \\ &= \int_0^{R_{\mathrm{dc}}} \frac{2r\lambda|\mathcal{A}|_{\mathrm{da}} p_{\mathrm{da}} N_{\mathrm{g}}}{R_{\mathrm{bs}}^2(1 + a\exp(-b(\arctan(\frac{h}{r}) - a)))}\,\mathrm{d}r. \end{aligned} \tag{3.12}$$

Proposition 1 has been proven. □

Though $U_{\mathrm{dc}}(h)$ can be expressed as a function of h with an integral of r included, the close-form of $U_{\mathrm{dc}}(h)$ is hard to be derived due to hyper-geometric functions are involved after calculating the integral. However, we can still analyze the upper and lower bounds of $U_{\mathrm{dc}}(h)$ through fixing the $P_{\mathrm{LoS}}(r,h)$ of each AoI as 1 and the smallest available $P_{\mathrm{LoS}}(r,h)$ which is achieved at the radio coverage boundary of R_{dc}, respectively. In this way, the integral of r can be removed from both upper and lower bound expressions of $U_{\mathrm{dc}}(h)$. Specifically, the upper bound of $U_{\mathrm{dc}}(h)$ is:

$$\mathbb{U}_{\mathrm{Udc}}(h) = \lambda|\mathcal{A}|_{\mathrm{da}} p_{\mathrm{da}} N_{\mathrm{g}} \frac{R_{\mathrm{dc}}^2}{R_{\mathrm{bs}}^2}, \tag{3.13}$$

and the lower bound is:

$$\begin{aligned} \mathbb{L}_{\mathrm{Udc}}(h) &= \frac{\lambda|\mathcal{A}|_{\mathrm{da}}}{1 + a\exp(-b(\arctan(\frac{h}{R_{\mathrm{dc}}}) - a))} p_{\mathrm{da}} N_{\mathrm{g}} \frac{R_{\mathrm{dc}}^2}{R_{\mathrm{bs}}^2} \\ &= \frac{\lambda|\mathcal{A}|_{\mathrm{da}} p_{\mathrm{da}} N_{\mathrm{g}} R_{\mathrm{dc}}^2}{R_{\mathrm{bs}}^2(1 + a\exp(-b(\arctan(\frac{h}{R_{\mathrm{dc}}}) - a)))}. \end{aligned} \tag{3.14}$$

According to (3.4), R_{dc} is determined by the minimal value of one increasing function of h and one decreasing function of h. The maximal R_{dc} can be obtained at the intersection point of two functions, let:

$$\frac{H_{\mathrm{opt}}}{\tan(a - \frac{1}{b}\ln\frac{1-\xi_{\mathrm{LoS}}}{a\xi_{\mathrm{LoS}}})} = \sqrt{(\frac{c\gamma_{\mathrm{DU}}}{4\pi f_c})^2 - H_{\mathrm{opt}}^2}, \tag{3.15}$$

the optimal height H_{opt} equals to:

$$H_{\mathrm{opt}} = \frac{\tan(a - \frac{1}{b}\ln\frac{1-\xi_{\mathrm{LoS}}}{a\xi_{\mathrm{LoS}}})\frac{c\gamma_{\mathrm{DU}}}{4\pi f_c}}{\sqrt{1 + \tan^2(a - \frac{1}{b}\ln\frac{1-\xi_{\mathrm{LoS}}}{a\xi_{\mathrm{LoS}}})}}. \tag{3.16}$$

Based on preceding analyses, we have the Corollary 1 of Proposition 1.

Corollary 1. *Given the DC flying height h, the upper bound of the average user number $U_{\mathrm{dc}}(h)$ efficiently covered by the DC is the function of h, and obtains its maximum when h equals to H_{opt}.*

3.3.2 DC Working Zone with D2B Link Constraints

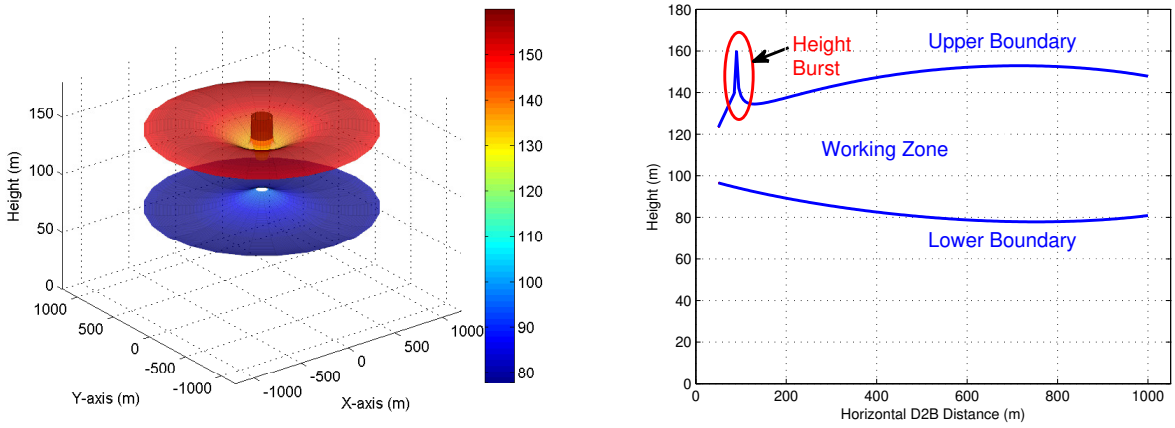
To maintain the qualities of D2B links, a working zone in which DC can maintain the D2B pathloss smaller than a given threshold γ_{DB} has to be identified. However, since the D2B pathloss model is a function of r_{DB} and θ, the available working zone of DC in the 3D space cannot be estimated intuitively. In this chapter, we transform θ into a function of h and r_{DB} to represent the working zone.

Fig. 3.2 shows an example of the DC working zone in both 3D and side views. The BS in Fig. 3.2 is configured with the height of $100m$ and $R_{\mathrm{bs}} = 1000m$, and γ_{DB} is set as $80dB$. According to Fig. 3.2, the DC working zone is constrained by a torus shape hovering around the top of BS horizontally. By increasing r_{DB}, the difference of the upper and lower height boundaries increases first, then decreases slowly. The maximal difference is achieved around $r_{\mathrm{DB}} = 700m$ with the available DC height ranging from $80m$ to $150m$. The torus shape of DC working zone can be explained as follows. According to Fig. 2.2, given a fixed pathloss threshold, the available θ range is reduced with the increasing of r_{DB}. The corresponding h_θ for each θ is calculated as:

$$h_\theta = r_{\mathrm{DB}} \tan(\theta) + H_{\mathrm{bs}}, \tag{3.17}$$

whose absolute value rises with r_{DB}'s increasing as long as θ is fallen into the available range. For a small r_{DB}, the reducing level of θ range is compensated by the increasing of h_θ, so the working zone height range are extending by increasing r_{DB}. While for a large r_{DB}, the increasing of h_θ cannot make up the serious reduction of θ range, and the working zone height range turns to shrink when r_{DB} keeps increasing.

One interesting effect observed from Fig. 3.2 is the sharp height burst of the working zone that is close to the BS. From Fig. 2.2, we notice that for a small r_{DB}, the pathloss curve can converge to a value less than the threshold when height increases to infinity, such as the curve corresponding to $r_{\mathrm{DB}} = 50m$. Without a θ to constraining the upper boundary, the available height at such r_{DB} can reach a large value until being limited by the large-scale pathloss term in (2.3). However, as long as r_{DB} rises to certain value whose corresponding pathloss curve reaches higher than the threshold, the upper boundary of height range is back to normal with the constraint of θ. Though this height burst provide



(a) 3D view.

(b) 2D projection on X-Z plane.

Figure 3.2: DC working zone.

additional space for DC deploying, its usage is still relatively limited in reality due to the narrow range.

3.4 3D Deployment Problem for Multiple DCs

Based on the preceding analyses, we formulate the 3D DC deployment problem in this section.

Considering a quasi-static AoIs scenario between adjacent re-deployments, N_{da} grids are selected as AoIs. A boolean parameter ϵ_i is defined to show the coverage state of the ith AoI ($0 \leq i \leq N_{\mathrm{da}}$):

$$\epsilon_i = \begin{cases} 1 & \text{if AoI } i \text{ is covered by a DC,} \\ 0 & \text{if AoI } i \text{ is uncovered.} \end{cases} \tag{3.18}$$

By defining ϵ_i, the purpose of enhancing user coverage can be translated to the objective of maximizing the summation of ϵ_i through deploying DCs in appropriate positions. Assuming the BS release N_{dc} number of DCs, the pathloss between DC j ($0 \leq j \leq N_{\mathrm{dc}}$) and AoI i is constrained by (3.3). Defining that AoI i can be effectively covered by DC j, as long as the D2U pathloss between them is less than threshold γ_{DU}. Since that, (3.18) can be

updated as follows:

$$\epsilon_{ij}(x_j, y_j, h_j) = \begin{cases} 1 & \text{if (3.3) holds,} \\ 0 & \text{else,} \end{cases} \tag{3.19}$$

where h_j is DC j's hovering height, r_{ij} is the horizontal distance between DC j and AoI j.

Three constraints are considered in the deployment problem. First, the maximum number of users served by one DC is defined as:

$$M_u = \lfloor \frac{C_{\text{dc}}}{R} \rfloor, \tag{3.20}$$

where R is minimal data rate required by each user, and C_{dc} is the capacity of one DC. To simplify the analysis, we assume that both R and C_{dc} are the same for arbitrary user and DC respectively. Leveraging the average number of effective users in one AoI calculated by (3.11), (3.20) can be transformed to express the maximal AoI number that can be supported by one DC:

$$C_{\text{da}} = \lfloor \frac{C_{\text{dc}}}{\text{E}(p_i \lambda |\mathcal{A}|_{\text{da}}) R} \rfloor, \tag{3.21}$$

The second constraint is maintaining the qualities for every D2B links. Defining a pathloss threshold γ_{DB}, we argue that any DCs being deployed must guarantee that their average D2B pathloss level $PL(r_{jb}, \theta_{jb})$ is less than γ_{DB}, where r_{jb} is the horizontal distance between DC j and the BS, and θ_{jb} represents the vertical angle between DC j and the antenna of BS. This constraint sets the working zone in which DCs choose their optimal deploying positions.

Multi-DC interference is considered as the third constraint. Noticing that all D2U communications are carried by the same spectrum band, one user can only be allocated to one DC. When one AoI is served by two or more DCs simultaneously, the users in such AoI can suffer serious interference from redundant DC, which is defined as multi-DC interference. To prevent this issue, we use I_i as the indicator of multi-DC interference suffered by AoI i:

$$I_i(x_j, y_j, h_j) = \max\{\sum_{j=1}^{N_{\text{dc}}} \epsilon_{ij}(x_j, y_j, h_j) - 1, 0\}. \tag{3.22}$$

When AoI i is uncovered or covered by only one DC, I_i equals to 0; otherwise I_i equals to the number of interfering DCs. By defining a maximal allowed I_i value, the overlapping of different DCs' user coverages can be minimized effectively.

Finally, the 3D DC deployment problem is formulated as the following optimization problem to maximize $\mathcal{U}$, i.e., the number of AoIs being covered:

$$\begin{aligned} \max_{x_j,y_j,h_j} \quad & \mathcal{U} = \sum_{j=1}^{N_{\mathrm{dc}}} \sum_{i=1}^{N_{\mathrm{da}}} \epsilon_{ij}(x_j, y_j, h_j) - \sum_{i=1}^{N_{\mathrm{da}}} I_i(x_j, y_j, h_j) \\ s.t. \quad & \sum_{i=1}^{N_{\mathrm{da}}} \epsilon_{ij}(x_j, y_j, h_j) \leq \lfloor \frac{C_{\mathrm{dc}}}{\mathrm{E}(p_i \lambda |\mathcal{A}|_{\mathrm{da}}) R} \rfloor \\ & PL(r_{jb}, \theta_{jb}) < \gamma_{\mathrm{DB}} \\ & I_i(x_j, y_j, h_j) \leq N_{\mathrm{Imax}}, \end{aligned} \tag{3.23}$$

where (x_j, y_j, h_j) is the coordinate of DC j being deployed at. N_{Imax} indicates the maximum allowed interference level suffered by one DA. Note that the constraints 1 and 2 are applying for each DC j, while the constraint 3 is examined for every AoI i.

3.5 Heuristic Solutions

Similar to the BS planning problem in [98], the DC deployment problem in (3.23) is also recognized as NP-hard [62]. Due to the heterogeneous pathloss models for D2U and D2B links, the optimal solution is hard to reach through mathematical deductions. So heuristic algorithms, especially the evolutionary heuristics such as genetic algorithm, PSO, etc., are considered as alternative choices to approaching the optima.

PSO algorithm is employed by both [62] and [98] to solve similar DC coverage problems. PSO has following advantages comparing with other heuristics: 1) fewer numerical parameters are required which simplifies the implementation; 2) lower computational cost and faster convergence speed [99]; 3) Unlike other algorithms (e.g., GA) that require a finite set of DC combinations [98], PSO can deal with an infinite set of DC combinations, which is suitable for DC deployment scenario where re-deployment speed matters.

3.5.1 Pure Particle Swarm Optimization based Algorithm

We customize the PSO algorithms proposed in [62] and [98] to form a pure PSO based algorithm, which is shown in Algorithm 1. Compared to the PSO algorithms in [62] and [98], the pure PSO algorithm involves the D2B link constraint specified in the DA-RAN.

In the pure PSO algorithm, the velocity vector of particle l at iteration t is defined as:

$$\begin{aligned}\mathbf{V}^{(l)}(t) = \phi\mathbf{V}^{(l)}(t-1) + c_1\phi_1(\mathbf{W}^{(l,local)}(t-1) - \mathbf{W}^{(l)}(t-1)) \\ + c_2\phi_2(\mathbf{W}^{(global)}(t-1) - \mathbf{W}^{(l)}(t-1)),\end{aligned} \tag{3.24}$$

where ϕ is the inertia weight that determines convergence speed. c_1 and c_2 are personal and global learning coefficients respectively. ϕ_1 and ϕ_2 are positive random variables. With the $\mathbf{V}^{(l)}(t)$ obtained from (3.24), $W^{(l)}(t)$ is updated as:

$$\mathbf{W}^{(l)}(t) = \mathbf{W}^{(l)}(t-1) + \mathbf{V}^{(l)}(t). \tag{3.25}$$

Algorithm 1 Pure PSO based algorithm for DC deployment in DA-RAN

1: Define a particle chosen space $\mathcal{S}$ as same as the DC working zone corresponding to γ_{DB}.
2: Generate L random particles in $\mathcal{S}$ as initial population. For each initial particle $\mathbf{W}^{(l)}(0)$,$l = 1, \ldots, L$, it is expressed as a random vector with size N_{dc}.
3: Re-generate the invalid elements in each $\mathbf{W}^{(l)}(0)$ vector until they meet constraint 1 and 3 in (3.23) simultaneously. Randomly generate initial velocity vector $\mathbf{V}^{(l)}(0)$ for each particle.
4: Calculate $U^{(l)}(0)$ through (3.23) for each $\mathbf{W}^{(l)}(0)$. Set $U^{(global)} = \max\{U^{(l)}(0), l = 1, \ldots, L\}$, $\mathbf{W}^{(global)} = \mathbf{W}^{(l)}(0)$ which achieves the $U^{(global)}$. Set $U^{(l,local)} = U^{(l)}(0)$, $\mathbf{W}^{(l,local)} = \mathbf{W}^{(l)}(0)$. Define maximum iteration number M_{Ite}.
5: **for** $t = 1, \ldots, M_{Ite}$ **do**
6: **for** $l = 1, \ldots, L$ **do**
7: Calculate $\mathbf{W}^{(l)}(t)$.
8: **while** $\mathbf{W}^{(l)}(t)$ exceeds any (3.23) constraints **do**
9: Re-generate the invalid elements in $\mathbf{W}^{(l)}(t)$.
10: **end while**
11: Calculate $U^{(l)}(t)$,$\mathbf{V}^{(l)}(t)$.
12: **if** $U^{(l)}(t) > U^{(l,local)}$ **then**
13: $U^{(l,local)} = U^{(l)}(t)$, $\mathbf{W}^{(l,local)} = \mathbf{W}^{(l)}(t)$.
14: **end if**
15: **if** $U^{(l,local)} > U^{(global)}$ **then**
16: $U^{(global)} = U^{(l,local)}$, $\mathbf{W}^{(global)} = \mathbf{W}^{(l,local)}$.
17: **end if**
18: **end for**
19: **end for**

The pure PSO based algorithm treats the combination of N_{dc} DC positions as one particle represented in vector form, which brings additional dependency between DCs. Since $\mathbf{W}^{(l)}(t)$'s elements (each DC position) are dependent with each other during iteration, the diversity of the particle searching space is seriously reduced. Though relatively less computational cost and shorter convergence time can be obtained in this way, the probability of falling into local optima increases. To compensate this, number pf particle and iteration times have to be added, which increases the complexity and total cost. Besides, the vector formed $\mathbf{W}^{(l)}(t)$ is hard to be re-generated to ensure that all elements fits (3.23)'s constraints simultaneously, especially the interference constraint which only allows the adjustments of partial elements. As a result, redundant re-generations must be executed before randomly generating a suitable particle.

3.5.2 Per-Drone Iterative PSO Algorithm

To alleviate the aforementioned issues of pure PSO based algorithm, we propose the DI-PSO algorithm which employs the PSO algorithm independently on each DC. The detail of DI-PSO is given in Algorithm 2. In DI-PSO, the particles are no longer vectors but individual positions. Due to the different particle structures, the updating functions of particle velocity and position are revised as follows:

$$\begin{aligned} V_d^{(l)}(t) = \phi V_d^{(l)}(t-1) &+ c_1\phi_1(W_d^{(l,local)}(t-1) - W_d^{(l)}(t-1)) \\ &+ c_2\phi_2(W_d^{(global)}(t-1) - W_d^{(l)}(t-1)), \end{aligned} \tag{3.26}$$

and

$$W_d^{(l)}(t) = W_d^{(l)}(t-1) + V_d^{(l)}(t). \tag{3.27}$$

For each DC d, PSO algorithm is used to calculate its best deployed position in which the maximal number of covered AoIs is achieved. With each DC being independently deployed, the spatial ergodicity of DI-PSO is increased which leads to higher probability of finding the global optima. meanwhile, the set of AoIs to be covered by DC d ($\mathbf{A}(d)$) is updated by each DC iteration with only uncovered AoIs remaining in it. This mechanism can reduce the times of re-generating $W_d^{(l)}(t)$ through prevent inter-DCs interference in a proactive way.

3.6 Simulations

Simulations are conducted to compare the performance of both proposed algorithms. Considering the single BS with 900m radio coverage radius in a suburban scenario, the sim-

Algorithm 2 DI-PSO algorithm for DC deployment in DA-RAN

Define a particle chosen space $\mathcal{S}$ as same as the DC working zone corresponding to γ_{DB}.
Initiate uncovered AoIs set $\mathbf{A}(1)$ contains all AoIs.
for $d = 1, \ldots, N_{\mathrm{dc}}$ **do**
    Generate L random particles $W_d^{(l)}(0)(l = 1, \ldots, L)$ as initial population of DC d within $\mathcal{S}$.
    Re-generate the invalid $W_d^{(l)}(0)$ until they meet constraint 1 and 3 in (3.23) simultaneously. Randomly generate initial velocity $V_d^{(l)}(0)$ for each $W_d^{(l)}(0)$.
    Calculate $U_d^{(l)}(0)$ with input $\mathbf{A}(d)$. Set $U_d^{(global)} = \max\{U_d^{(l)}(0), l = 1, \ldots, L\}$, $W_d^{(global)} = W_d^{(l)}(0)$ which achieves the $U_d^{(global)}$. Set $U_d^{(l,local)} = U_d^{(l)}(0)$, $W_d^{(l,local)} = W_d^{(l)}(0)$.
    for $t = 1, \ldots, M_{Ite}$ **do**
        for $l = 1, \ldots, L$ **do**
            Calculate $W_d^{(l)}(t)$.
            while $W_d^{(l)}(t)$ exceeds (3.23) constraints **do**
                Re-generate $W_d^{(l)}(t)$.
            end while
            Calculate $U_d^{(l)}(t)$,$V_d^{(l)}(t)$.
            if $U_d^{(l)}(t) > U_d^{(l,local)}$ **then**
                $U_d^{(l,local)} = U_d^{(l)}(t)$, $W_d^{(l,local)} = W_d^{(l)}(t)$.
            end if
            if $U_d^{(l,local)} > U_d^{(global)}$ **then**
                $U_d^{(global)} = U_d^{(l,local)}$,
                $W_d^{(global)} = W_d^{(l,local)}$.
            end if
        end for
    end for
    Calculate $\mathbf{A}_{new}(d)$ as the set of AoIs newly covered by DC d.
    Update $\mathbf{A}(d+1) = \mathbf{A}(d) - \mathbf{A}_{new}(d)$.
end for

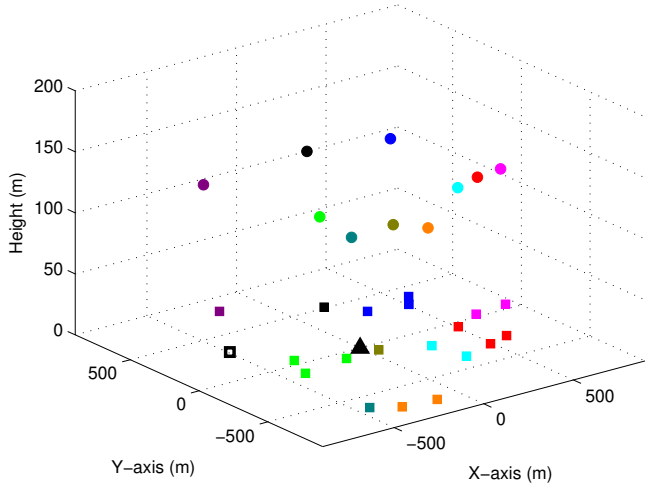
ulation parameters are listed in Table. 3.2. 2.4GHz WiFi band is chosen as the carrier f_c for its wide adoption in previous works and commercial drone products [100] [25]. The 850MHz LTE band are used to support D2B communications. To suit the scenario where the D2B pathloss model is built, propagation parameters for suburban are chosen. Allocating different bands to D2U and D2B links also prevent the interference between them. Both γ_{DU} and γ_{DB} are set within the range used by related works [58].

Table 3.2: Simulation Parameters for Multi-DC 3D Deployment

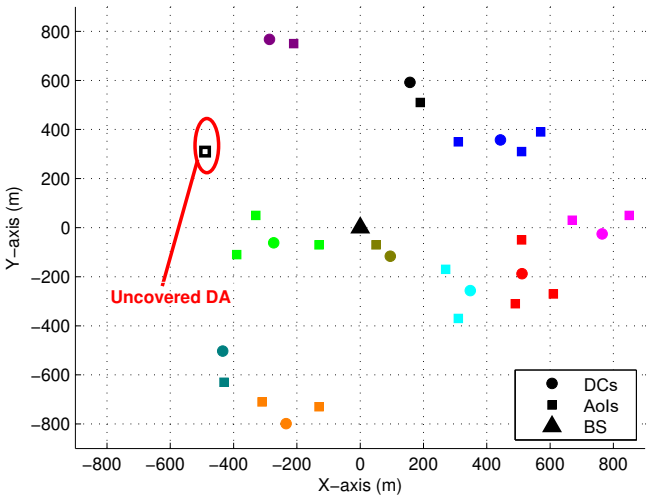
Simulation Parameters	Numerical Values
BS radio coverage radius R_{bs}	900m
AoI size	20m×20m
$(a, b, \eta_{\mathrm{LoS}}, \eta_{NLoS})$	(4.88,0.43,0.1,21)
(D2U f_c, D2B f_c)	(2.4GHz, 850MHz)
$(\gamma_{\mathrm{DU}}, \gamma_{\mathrm{DB}}$(normalized in dB))	(89dB, 80dB)
λ	0.025 person/m^2
C_{dc}	1Gbps
R	100Mbps
N_{Imax}	2
Particle population	400
Maximum number of iteration	50

Fig. 3.3 shows one example DC deployment generated by the DI-PSO algorithm. N_{da} is 20 and N_{dc} equals 10 in this scenario. Triangle dot locating at the origin of coordinate represents the BS. AoIs are indicated as square dots and DCs are illustrated as circle dots. AoIs covered by the same DC are with the same color as the DC's. In this example, 19 out of 20 AoIs are covered by DCs with only one uncovered AoI that locates near the edge of the BS radio coverage. Since all DCs have to guarantee D2U and D2B pathloss less than γ_{DU} and γ_{DB} simultaneously, the marginal location of the uncovered AoI implies no DC can be allocated for it without breaking pathloss constraints. The label of each DC in Fig. 3.3 represents its iteration order d in DI-PSO algorithm. The first deployed DC is allocated with 3 AoIs, while the fifth to tenth DCs can only occupy 1 AoI by each. This trend is caused by updating $\mathbf{A}(d)$ per-iteration with only uncovered AoIs being left, which can be regarded as one kind of greedy mechanism to find the corresponding maximal AoIs coverage for any given N_{dc}.

Fig. 3.4 shows the number of effectively covered AoIs versus the available DC number N_{dc} for both algorithms. Using the same AoIs scenario shown in Fig. 3.3, simulations are



(a) 3D view.



(b) 2D projection on X-Y plane.

Figure 3.3: DC deployment example generated by DI-PSO algorithm.

conducted 100 times for each algorithm under each DC number. Noting that the error-bars attached on each curve indicate the standard deviation level of their corresponding test points. As shown in Fig. 3.4, two algorithms are comparable when N_{dc} equals to 1 or 2.

By increasing N_{dc}, both curves keep increasing while the DI-PSO always maintains average 2 more AoIs being covered comparing with pure PSO. This poor performance of pure PSO algorithm is caused by the limited diversity of particle searching space. The increasing trends of two curves are all dented when N_{dc} reaches large values, while the pure PSO curve increases slower with average 3 covered AoIs being exceeded by DI-PSO curve when $N_{\mathrm{dc}} = 10$. Besides, little variance is achieved by DI-PSO algorithm for each N_{dc}, which indicates DI-PSO's stability of converging to optimized results.

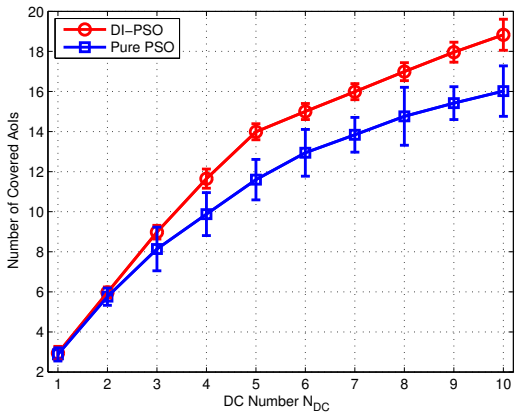


Figure 3.4: AoI coverage comparison between pure PSO and DI-PSO algorithms with different number of DCs.

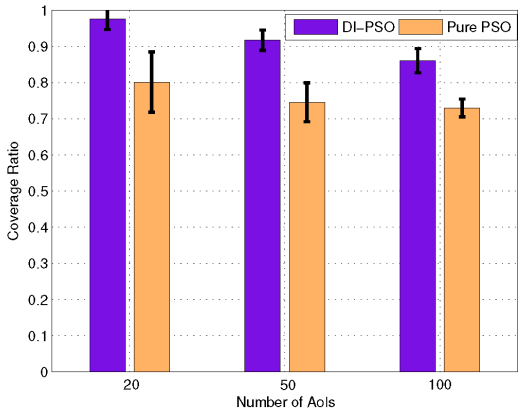


Figure 3.5: AoI coverage comparison between pure PSO and DI-PSO algorithms with different numbers of AoIs.

Fig. 3.5 compares the AoI coverage ratios of the two algorithms for the AoI number of 20, 50 and 100, respectively. The coverage ratio is defined as the number of covered AoIs over the total number of AoIs. The number of available DCs N_{dc} is fixed as 10. For each AoIs number, 100 snapshots of AoI distribution are generated to run Monte Carlo tests. As Fig. 3.5 indicates, AoI coverage ratios of both algorithms are decreased by increasing the number of AoIs. Since a larger number of AoIs leading to higher AoI density within the fixed BS radio coverage area, the serious inter-DC interference caused by the high AoI density is the main reason for the decrease of the AoI coverage ratio. Comparing with the pure PSO algorithm, the DI-PSO algorithm achieves higher AoI coverage ratios with less variances in all scenarios. Note that the performance of pure PSO based algorithm in both Fig. 3.4 and Fig. 3.5 is better than the similar simulation in [38]. This performance increase is caused by applying new D2B pathloss model to generate the working zone which constrains the particle searching space for the pure PSO based algorithm. The number of particle re-generations is also reduced by applying the working zone at the beginning of

the algorithm.

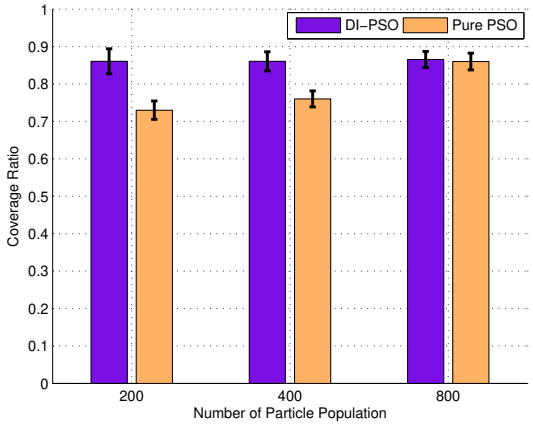


Figure 3.6: Performance comparison between pure PSO and DI-PSO algorithms with different particle populations.

To test the impacts of the limited particle searching space in pure PSO. We further compare the performance of two algorithms with the particle population set as 200, 400, and 800. The 100 snapshots of AoI distribution with 100 AoIs are used for testing. Fig. 3.10 shows the simulation results. For DI-PSO, its AoI coverage ratios under different particle populations remains no change, with only the standard deviation decreases by increasing the particle population. This result indicates that the DI-PSO can effectively converge to the optimal solution with a small particle population. For pure PSO based algorithm, its AoI coverage ratio keeps rising with the increasing particle population, and reaches the same level when the particle population equals to 800. However, the AoI coverage ratio achieved by pure PSO with 800 particle population size can be reached by DI-PSO with only 200 particle population size, which proves that the DI-PSO is more efficient than the pure PSO based algorithm in terms of computation costs.

3.7 Summary

In this chapter, we have theoretically analyzed the DC user coverage and feasible working zone by leveraging the emerging D2U and D2B pathloss models. The 3D deployment problem of DC in DA-RAN has been formulated to maximize the user coverage while maintaining D2B link qualities, which is solved by the DI-PSO heuristic solution. With

DCs' capabilities of providing LoS links and enabling dynamic deployments, the results from this research can shed light on the DA-RAN research. In addition, the DC user coverage and working zone analysis results offer general understandings to inspire future research concerning the DC radio coverage.

Chapter 4

Multi-Drone-Cell Trajectory Planning and Communication Scheduling

The static 3D deployment of DC leads to uneven communication performance between users located at the center and the edge of the DC's radio coverage. To improve user fairness and network performance, in this chapter, we design 3D trajectories of multiple DCs in the DA-RAN where DCs fly over associated AoIs and relay communications between the BS and users in AoIs. We formulate the multi-DC 3D trajectory planning and scheduling as a MINLP problem with the objective of minimizing the average D2U pathloss. The 3D trajectory variations in both horizontal and vertical directions, as well as the state-of-the-art DC-related channel models are considered in the formulation. To address the non-convexity and NP-hardness of the MINLP problem, we first decouple it into multiple ILP and quasi-convex sub-problems in which AoI association, D2U communication scheduling, horizontal trajectories and flying heights of DCs are respectively optimized. Then, we design a multi-DC 3D trajectory planning and scheduling algorithm to solve the sub-problems iteratively based on the BCD method. A k-means-based initial trajectory generation and a search-based start slot scheduling are considered in the proposed algorithm to improve trajectory design performance and ensure inter-DC distance constraint, respectively. Extensive simulations are conducted to investigate the impacts of DC quantity, horizontal speed and initial trajectory on the trajectory planning results. Compared with the static DC deployment, the proposed trajectory planning can achieve 10-15 dB reduction on average D2U pathloss, and reduce the D2U pathloss standard deviation by 68%, which indicate the improvements of network performance and user fairness.

4.1 Background and Motivations

To overcome the coverage and flexibility challenges faced by current RAN, the drone communication technology is proposed as a promising solution. Equipped with specific wireless transceivers, drones can communicate with both terrestrial users and cellular BSs using WiFi [92] or LTE [24] technologies. By integrating drone communication with terrestrial RAN, the DA-RAN, in which drones perform as DCs to relay data between users in AoIs and the associated terrestrial BS, has been proposed and verified by field experiments [44]. In DA-RAN, AoI includes both the CHs and the BTSs where the allocated RAN spectrum resources are temporarily inadequate, e.g. congested road, concerts and sports events, etc. Compared with the terrestrial RAN, DA-RAN advances in following four aspects: 1) The LoS probability for the D2G wireless link is higher than the terrestrial BS-to-user wireless link [101]. Experiments indicate that LoS links probability is the dominating factor to increase network performance [102]; 2) DCs can be dynamically deployed and dispatched to different controllers/users with respect to the spatial and temporal traffic variations [103]; 3) unlike connected vehicles whose mobility is controlled by drivers or autonomous driving controller, the trajectories of DCs can be fully controlled by system providers, which empowers DCs with the dynamic deployment feature [90] [104]; 4) DC are capable of executing computing tasks by equipping with CPU or caching modules [105] [43]. However, it is challenging to fully utilize the potential of DCs due to the following two reasons. First, the 3D mobility of DC poses great complexity on the DC spatial placement, especially in multi-DC scenarios [13]. Second, specific channel models are required to highlight the unique features of D2U and D2B channels[46].

Several studies optimizing the multi-DC spatial placements to support terrestrial users emerges in recent year, which can be divided into two categories, i.e., static DC deployment and DC trajectory planning. The static DC deployment research focus on optimizing the hovering positions of DCs to maximize terrestrial users QoS. However, the static deployment fails to guarantee the fairness for users, in which the users located at the edge of the DC's radio coverage suffer relatively higher pathloss compared with the users located at the center of the DC's coverage. In addition, most existing DC deployment works focus on optimizing the D2U communication, while ignore or idealize the D2B link quality constraints. To promote the fairness for all users and maintain low deployment cost, some researchers further propose the DC trajectory planning approach that allows DCs fly over and serve AoIs periodically according to designed trajectories. The purpose of DC trajectory planning is optimizing AoIs association and trajectory for each DC to maximize user QoS [39] [72]. However, three issues remain unsolved in current works. First, to reduce the complexity of optimization problems, most existing works assume that all DCs fly at a pre-defined constant height, which shrinks the 3D trajectory planning into a 2D horizontal

trajectory planning, and fails to realize the performance improvements by adjusting DC flying heights. Second, the commonly used assumption of free-space propagation model cannot reflect the unique D2G channel features. Third, the D2B link quality constraint is also omitted by most DC trajectory planning works.

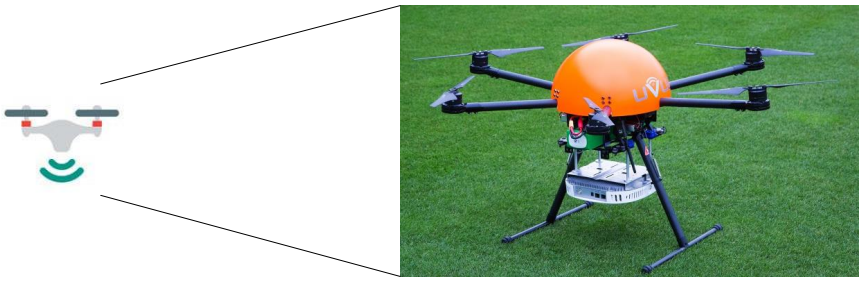
To address those issues, in this section, we investigate the 3D trajectory planning and scheduling for multiple DCs in the DA-RAN. Considering the state-of-the-art D2U [45] and D2B [24] channel models, and constraints of D2B link qualities, the multi-DC 3D trajectory planning problem is formulated as a MINLP problem which aims at minimizing the average D2U pathloss for all users within one trajectory period. By decoupling the MINLP problem into multiple quasi-convex or ILP sub-problems, we can separately optimize the AoI association, D2U communication scheduling, DC horizontal trajectories and flying heights in each sub-problem, respectively. In essence, we adopt the BCD mechanism to devise a multi-DC 3D trajectory planning and scheduling algorithm, in which the sub-problems are iteratively optimized and converge to the optima. The main contributions of this work are listed as follows:

- We investigate the 3D trajectory planning of multiple DCs in which both the flying heights and horizontal trajectories of DC are optimized together instead of optimizing horizontal trajectories on a 2D plane. As far as we know, this is the first work considering the real 3D trajectory in which the flying height of any DC can be adjusted at different slots on its trajectory.

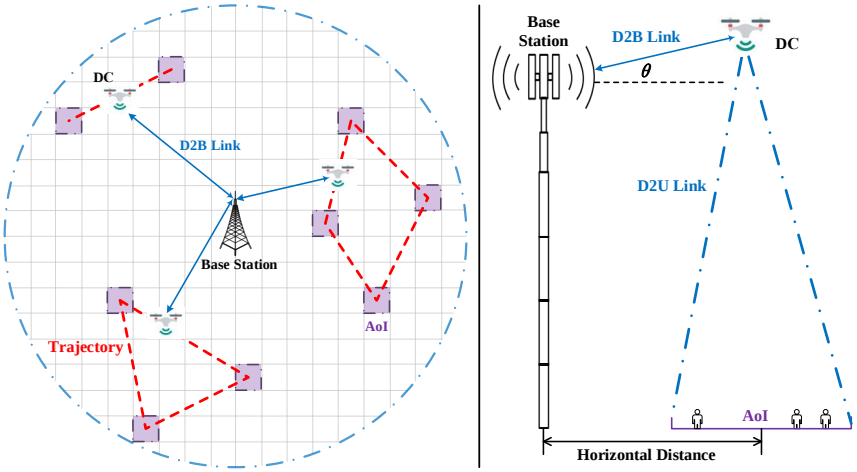
- To make the system model more practical, we employ the state-of-the-art D2U and D2B pathloss models rather than the traditional pathloss models (e.g., Friis equation) in the system model. We formulate the multi-DC trajectory planning problem, which turn to be an MINLP, and decouple it into multiple sub-problems to resolve the non-convexity. A protect distance constraint between any two DCs at every time slots is considered in the problem formulation to suppress the physical collision and mutual interference of DCs. Instead of modifying the 3D trajectories, we ensure the protect distance constraint by scheduling the start slot of each trajectory to avoid introducing non-convex constraints in trajectory-related sub-problems.

- A BCD based algorithm is proposed to separately optimize AoI association, D2U communication scheduling, horizontal trajectories and flying heights of DCs in different sub-problems, respectively. Besides, a k-means-based scheme is devised to generate the DC initial trajectories for further improvements on performance.

- We conduct extensive simulations and results demonstrate that the proposed 3D trajectory planning and scheduling algorithm can reduce the average D2U pathloss



(a) DC photograph [106]



(b) System model

Figure 4.1: Multi-DC 3D trajectory planning and scheduling.

by $15-20$ dB, and lower the D2U pathloss standard deviation by 68%, in comparison with the static DC deployment algorithm based on PSO.

The remainder of this chapter is organized as follows. In Section 4.2 the system model for DC trajectory planning and scheduling in DA-RAN is introduced. Then the multi-DC 3D trajectory planning and scheduling problem is formulated in Section 4.3. In Section 4.4 the MINLP problem in Section IV is decoupled into sub-problems and the BCD based algorithm is proposed to solve it. Simulation and numerical results are carried out in Section 4.5, and the conclusion is given in Section 4.6.

4.2 System Model

In this section, we introduce the DA-RAN scenario, D2U and D2B channel models, as well as the DC trajectory model used for further analyses and problem formulation.

4.2.1 Drone Assisted Radio Access Networks

Fig. 4.1 shows the system model and the DC photograph for the multi-DC 3D trajectory planning and scheduling. The DC is supported by the quad-copter or multi-rotor drones, which is regarded as LAP with the low flying height [13], limited communication coverage [13] and static hovering capability. Based on the DA-RAN architecture, we investigate the scenario in which multiple DCs are controlled by a single BS to support users in AoI through the state-of-the-art wireless relay techniques [107] [108]. The radio coverage area of the BS $|\mathcal{S}|_{\text{bs}}$ is a circle with radius r_{BS}, and is divided into a mesh consisting of multiple grids on X-Y plane. The side-length of each grid is denoted as L_{aoi}. Without loss of generality, the average D2U pathloss for any users in one grid can be treated as equal since L_{aoi} is far smaller than r_{BS}. Assuming that users are uniformly distributed over $|\mathcal{S}|_{\text{bs}}$, and each grid can be chosen as AoI with same probability. Therefore, the user association is equal to AoI association in this work. Both users in AoIs and DCs are considered as identical devices with identical transmit power and uplink/downlink bandwidth. Considering the fact that AoIs change their distribution in a relatively low frequency, the dynamic distribution of AoIs can be treated as a quasi-static scenario between successive trajectory planning. Based on the current snapshot of AoIs distribution, the BS running trajectory planning and scheduling algorithm to calculate optimal trajectories for all DCs, and update them to the DCs via D2B links. When the BS senses significant changes of AoIs distribution, re-planning process is triggered to design new trajectories for DCs, otherwise, the trajectory planning result keeps constant. The set of AoIs to be served and the set of DCs to be deployed are denoted as $\mathcal{U}$ and $\mathcal{D}$, respectively. Their cardinalities, $|\mathcal{U}|$ and $|\mathcal{D}|$, represent the number of AoIs and DCs, respectively.

4.2.2 D2U and D2B Channel Models

Based on the state-of-the-art D2G channel research [45] [24], both the D2U and D2B links are modeled in our work, respectively. For D2U links, the LoS probability is calculated as [45]

$$P_{\text{LoS}}(r_{\text{DU}}, h) = \frac{1}{1 + a\exp(-b(\arctan(\frac{h}{r_{\text{DU}}}) - a))} \tag{4.1}$$

where r_{DU} is the horizontal distance between DC and the AoI, h represents the flying heights of the DC. a and b are environment-based constant values. The average D2U pathloss can be derived based on (4.1) [45]:

$$\begin{aligned} PL(r_{\mathrm{DU}}, h) = 20\log(\frac{4\pi f_c\sqrt{h^2 + r_{\mathrm{DU}}^2}}{c}) \\ + P_{\mathrm{LoS}}(r_{\mathrm{DU}}, h)\eta_{\mathrm{LoS}} + (1 - P_{\mathrm{LoS}}(r_{\mathrm{DU}}, h))\eta_{\mathrm{NLoS}} \end{aligned} \tag{4.2}$$

where f_c (Hz) and c (m/s) are carrier frequency and speed of light, respectively. η_{LoS} and η_{NLoS} are additional losses for LoS and NLoS links obtained through field test data, which involves the impacts of shadowing components. a, b, η_{LoS} and η_{NLoS} are all environment-based parameters.

D2B links are designed to provide high-reliability data transmission between DCs and their corresponding BS. The average D2B pathloss is calculated as follow by implementing the D2B channel model in [24]:

$$PL(r_{\mathrm{DB}}, \theta) = 10\alpha\log(r_{\mathrm{DB}}) + A(\theta - \theta_0)\mathrm{e}^{(\frac{\theta_0 - \theta}{B})} + \eta_0 \tag{4.3}$$

where r_{DB} and θ denote the horizontal distance and the vertical angle between the DC and the BS antenna, respectively. α, A, θ_0, B, and η_0 are the terrestrial pathloss exponent, excess pathloss scalar, angle offset, angle scalar, and excess pathloss offset, respectively. All of them are environment-based parameters and involving impacts of shadowing components. Except r_{DB} and θ, all other parameters in (4.3) are environment-based constants. Since the D2B channel model in [24] use 850 MHz LTE bands, (4.3) contains no parameter representing carrier frequency.

Both the D2U [45] and D2B [24] pathloss models are large-scale pathloss models. For the small scale fading and multipath effects, currently there is no specific model for the drone-to-ground communication links. Moreover, since the objective of our multi-DC trajectory planning is to minimize the mean D2U pathloss of the system, the small-scale pathloss can be average out at zero or a constant offset during the analysis. Therefore, based on the assumptions in D2U and D2B channel models, in this work we do not focus on the small scale fading and multipath effects in D2U and D2B links.

4.2.3 DC Trajectory Model

For an arbitrary DC $d \in \mathcal{D}$, we design its trajectory such that the DC serves the associated AoI set $\mathcal{A}_d \subseteq \mathcal{U}$ periodically. Within one period T, d flies over all its associated AoI and serves them sequentially according to the scheduling result. Since the continuous time can

introduce infinite number of position variables to describe the DC trajectory, we discrete the period T into N equal-time slots to simplify the formulation. The length of each slot $\delta_t = \frac{T}{N}$ can be set as small as possible to approximate the continuous optimal trajectory. Based on this model, the trajectory of DC d within each T can be modeled as a N-length sequence composed by three-dimensional vectors:

$$\mathbf{W}_d[n] = [x_d[n], y_d[n], h_d[n]], \quad n = 1, ..., N \tag{4.4}$$

where $x_d[n]$, $y_d[n]$ and $h_d[n]$ denote the 3D coordinates of DC d at slot n. DC d is considered to follow the same trajectory $\mathbf{W}_d[n]$ over consecutive periods until the re-planning process is triggered. For multiple DCs working simultaneously, they share the same trajectory period length T to simplify the trajectory planning and scheduling.

Several trajectory constraints are considered in our work: 1) Each DC needs to return to its initial location by the end of each period T, which implies that the trajectory of each DC is a closed curve in 3D space. 2) Within any slot n, the horizontal and vertical shifts of any DC cannot exceed the maximal horizontal distance $V_{\max}\delta_t$, and the maximal height difference $H_{\max}\delta_t$, respectively. $V_{\max}$ and $H_{\max}$ are maximum allowed horizontal and vertical speeds. 3) For any slot n, the 3D distance between any two DCs cannot be smaller than a pre-defined protect distance $Z_{\min}$, which prevents the physical collision, disturbance and mutual interference among DCs. Since calculating interference from non-associated DCs to any AoI based on D2U pathloss model is highly complex and makes the trajectory planning problem unsolvable, in this work, we assume that the mutual interference (to AoIs) among DCs can be effectively avoided by ensuring the protect distance constraint.

4.2.4 AoI Association and D2U Communication Scheduling

In this work, the DC-AoI association is denoted by the binary variable $a_{d,u}$. $a_{d,u} = 1$ when AoI $u \in \mathcal{U}$ is associated to DC $d \in \mathcal{D}$, and otherwise $a_{d,u} = 0$. The D2U communication scheduling is denoted by the binary variable $k_{d,u}[n]$ for $\forall d \in \mathcal{D}, u \in \mathcal{U}, n \in \mathcal{N}$. If AoI u is severed by DC d in slot n, $k_{d,u}[n]$ is set as 1; otherwise, $k_{d,u}[n] = 0$. For each DC with pre-defined trajectory planning and AoI association results, a D2U communication scheduling scheme is designed to allocate each slot to the corresponding AoI, and guarantee the fairness among all associated AoIs. Several constraints are considered in the AoI association and scheduling model: 1) One DC can serve maximal $|\mathcal{A}_d|_{\max}$ number of AoIs. 2) In any slot n, one DC d can serve at most one AoI $u \in \mathcal{A}_d$; in all slots, one AoI u can only be associated to one DC. 3) For any given T and $\mathcal{A}_d$, the total N slots are uniformly scheduled to each $u \in \mathcal{A}_d$ to ensure fairness. 4) The slot amount scheduled to every AoIs cannot be smaller than a pre-defined threshold $S_{\min}$, which indicates the minimal user service time

constraint. 5) To prevent the overloads and delay caused by frequent switching between associated AoIs, all slots scheduled to one u within T have to be consecutive.

4.3 Problem Formulation

In this section we formulate the multi-DC 3D trajectory planning problem based on the aforementioned system model.

According to (4.4), the 3D distance from the DC d to AoI u in time slot n can be expressed as

$$\begin{aligned} m_{d,u}[n] &= \sqrt{h_d[n]^2 + \|\mathbf{l}_d[n] - \mathbf{l}_u\|^2} \\ &= \sqrt{h_d[n]^2 + r_{d,u}[n]^2} \end{aligned} \tag{4.5}$$

where $\mathbf{l}_u = [x_u, y_u]$ is the 3D coordinate of AoI u. $\mathbf{l}_d[n]$ is d's 2D projection on X-Y plane $\mathbf{l}_d[n] = [x_d[n], y_d[n]]$. $r_{d,u}[n]$ denotes the horizontal distance between DC d and AoI u. Without loss of generality, we set the BS at the original point of coordinate system. By substituting $m_{d,u}[n]$ for the D2U and D2B distances in (4.2) and (4.3), we can calculate the D2U pathloss between DC d and AoI u in slot n

$$\begin{aligned} P_{d,u}[n] = &\, 20\log(\frac{4\pi f_c m_{d,u}[n]}{c}) \\ &+ P_{\text{LoS}}(r_{d,u}[n], h_d[n])\eta_{\text{LoS}} \\ &+ (1 - P_{\text{LoS}}(r_{d,u}[n], h_d[n]))\eta_{\text{NLoS}} \end{aligned} \tag{4.6}$$

as well as the D2B pathloss between the BS and DC d in slot n

$$\begin{aligned} P_{d,B}[n] = &\, 10\alpha\log(\|\mathbf{l}_d[n]\|) \\ &+ A(\theta_{d,B}[n] - \theta_0)\mathrm{e}^{(\frac{\theta_0 - \theta_{d,B}[n]}{B})} + \eta_0 \end{aligned} \tag{4.7}$$

where $\theta_{d,B}[n] = \arctan(h_d[n]/\|\mathbf{l}_d[n]\|)$ in degree.

Since all users and DCs are identical devices with fixed transmission power and transmission bandwidth in each period, the achievable D2U data rate between DC d and AoI u is negative correlated with the D2U pathloss. Therefore, the aim of the multi-DC 3D trajectory planning and scheduling problem is minimizing the average D2U pathloss of the network over one period T.

Define $\mathbf{A} = \{a_{d,u},\ \forall d, u\}$, $\mathbf{K} = \{k_{d,u}[n],\ \forall d, u, n\}$ and $\mathbf{W} = \{\mathbf{W}_d[n],\ \forall d, n\}$, the trajectory planning and scheduling problem can be formulated as

$$\min_{\mathbf{A},\mathbf{K},\mathbf{W}} \quad \frac{1}{N|\mathcal{U}|}\sum_{u=1}^{|\mathcal{U}|}\sum_{d=1}^{|\mathcal{D}|} a_{d,u}\Big(\sum_{n=1}^{N} k_{d,u}[n]P_{d,u}[n]\Big) \tag{4.8}$$

$$s.t. \quad \sum\nolimits_{u=1}^{|\mathcal{U}|} a_{d,u} \le |\mathcal{A}_d|_{\max}, \quad \forall d, \tag{4.8a}$$

$$\sum\nolimits_{d=1}^{|\mathcal{D}|} a_{d,u} = 1, \quad \forall u, \tag{4.8b}$$

$$\sum\nolimits_{u=1}^{|\mathcal{U}|} k_{d,u}[n] = 1, \quad \forall d, n, \tag{4.8c}$$

$$\sum\nolimits_{d=1}^{|\mathcal{D}|} k_{d,u}[n] = 1, \quad \forall u, n, \tag{4.8d}$$

$$\sum\nolimits_{n=1}^{N} k_{d,u}[n] = \frac{N}{|\mathcal{A}_d|}, \quad \forall d, u, \tag{4.8e}$$

$$\sum\nolimits_{n=1}^{N} k_{d,u}[n] \ge S_{\min}, \quad \forall d, u, \tag{4.8f}$$

$$\sum\nolimits_{o}^{\frac{N}{|\mathcal{A}_d|}} k_{d,u}[(n+o) \bmod N] \le \frac{N}{|\mathcal{A}_d|}, \quad \forall d, u, n, \tag{4.8g}$$

$$a_{d,u}, k_{d,u}[n] \in \{0, 1\}, \quad \forall d, u, n, \tag{4.8h}$$

$$\mathbf{W}_d[1] = \mathbf{W}_d[N+1], \quad \forall d, \tag{4.8i}$$

$$\|\mathbf{l}_d[n+1] - \mathbf{l}_d[n]\| \le V_{\max}\delta_t, \quad \forall d, n, \tag{4.8j}$$

$$|h_d[n+1] - h_d[n]| \le H_{\max}\delta_t, \quad \forall d, n, \tag{4.8k}$$

$$\|\mathbf{W}_i[n] - \mathbf{W}_j[n]\| \ge Z_{\min}, \quad \forall n, i, j \ne i, \tag{4.8l}$$

$$P_{d,B}[n] \le P_{\mathrm{DB}}, \quad \forall d, u, n. \tag{4.8m}$$

In (4.8), $|\mathcal{A}_d| = N/\sum_{u=1}^{|\mathcal{U}|} a_{d,u}$ is the number of AoIs associated to DC d. $\|\mathbf{W}_i[n] - \mathbf{W}_j[n]\|$ represents the 3D distance between DC i and j at slot n. $a \bmod b$ is the modulo operation between a and b. P_{DB} is the pathloss threshold for D2B communication. (4.8a)-(4.8h) are AoI association and D2U communication scheduling constraints, in which (4.8a) is constraint 1); (4.8b)-(4.8d) represent the constraint 2); (4.8e) and (4.8f) corresponds to constraint 3) and 4), respectively; (4.8g) indicates constraint 5). (4.8i)-(4.8l) correspond to DC trajectory constraints 1), 2) and 3). (4.8m) is the D2B pathloss constraint.

Due to the quadratic and exponential terms in (4.8) and constraints, as well as the binary variable $a_{d,u}$, $k_{d,u}[n]$, problem (4.8) is a MINLP problem [58]. Besides, the optimization objective (4.8) and constraints are non-convex for DC trajectory $\mathbf{W}$, which is difficult to solve directly.

4.4 Multi-DC 3D Trajectory Planning and Scheduling Algorithm

Although the objective and constraints in problem (4.8) are non-convex or non-linear for the decision variables, the problem can still be transformed into solvable forms (e.g. quasi-convex or ILP) by setting parts of the decision variables as constants. Then, the MINLP problem can be decoupled into multiple sub-problems which are solvable for parts of the decision variables. Specifically, for the multi-DC 3D trajectory planning and scheduling problem, we divide the decision variable set into four blocks (i.e. $\mathbf{A}$, $\mathbf{K}$, $\mathbf{L} = \{\mathbf{l}_d[n],\ \forall d, n\}$ and $\mathbf{H} = \{h_d[n],\ \forall d, n\}$), and propose multiple sub-problems in which all blocks or their sub-blocks are optimized, respectively. However, the problem (4.8) remains non-convex to DC trajectory variable $\mathbf{W}$ even with given $\mathbf{A}$ and $\mathbf{K}$. Therefore, we further divide $\mathbf{W}$ into two independent blocks, i.e. the horizontal DC trajectory $\mathbf{L}$ and the DC flying height $\mathbf{H}$.

4.4.1 AoI Association Optimization

Given the constant $\mathbf{K}$, $\mathbf{L}$ and $\mathbf{H}$, which indicate the pre-defined trajectories of multiple DCs, the AoI association sub-problem can be written as an ILP problem:

$$\begin{aligned} \min_{\mathbf{A}} \quad & \frac{1}{N|\mathcal{U}|} \sum_{u=1}^{|\mathcal{U}|} \sum_{d=1}^{|\mathcal{D}|} a_{d,u} \left(\sum_{n=1}^{N} k_{d,u}[n] P_{d,u}[n]\right) \\ s.t. \quad & (4.8a), (4.8b), \quad a_{d,u} \in \{0,1\} \quad \forall d, u. \end{aligned} \tag{4.9}$$

Since exact $\mathbf{K}$ can only be determined with given $\mathcal{A}_d$, an initial D2U communication scheduling $\mathbf{K}_0$, in which $k_{d,u}[n] = 1,\ \forall d, u, n$, is defined for the first AoI association optimization. The branch and bound method supported by various solvers (e.g. Gurobi [109]) can be used to solve problem (4.9) efficiently.

4.4.2 D2U Communication Scheduling Optimization

Based on the optimized $\mathbf{A}$, as well as the constant $\mathbf{L}$ and $\mathbf{H}$, the D2U communication scheduling sub-problem is an ILP problem too:

$$\begin{aligned} \min_{\mathbf{K}} \quad & \frac{1}{N|\mathcal{U}|} \sum_{u=1}^{|\mathcal{U}|} \sum_{d=1}^{|\mathcal{D}|} a_{d,u} \left(\sum_{n=1}^{N} k_{d,u}[n] P_{d,u}[n]\right) \\ s.t. \quad & (4.8a), (4.8b), (4.8c), (4.8d), (4.8e), (4.8f), (4.8g), \\ & k_{d,u}[n] \in \{0,1\} \quad \forall d, u, n. \end{aligned} \tag{4.10}$$

It is worth noting that constraint (4.8e) and (4.8g) turn to be linear constraint to $\mathbf{K}$ given constant $\mathbf{A}$. Same as problem (4.9), problem (4.10) can be efficiently solved by the branch and bound method.

4.4.3 DC Horizontal Trajectory Optimization

The sub-problem to optimize $\mathbf{L}$ with constant $\mathbf{A}$, $\mathbf{K}$ and $\mathbf{H}$ can be expressed as

$$\begin{aligned} \min_{\mathbf{L}} \quad & \frac{1}{N|\mathcal{U}|}\sum_{u=1}^{|\mathcal{U}|}\sum_{d=1}^{|\mathcal{D}|} a_{d,u}(\sum_{n=1}^{N} k_{d,u}[n]P_{d,u}[n]) \\ s.t. \quad & (4.8j), (4.8m), \quad \mathbf{l}_d[1] = \mathbf{l}_d[N+1] \quad \forall d. \end{aligned} \tag{4.11}$$

According to (4.6), (4.11) is non-convex for $\mathbf{L}$. Instead of jointly optimizing $\mathbf{L}$, we further divide the block into its element variable $\mathbf{l}_d[n]$ and revise (4.11) as

$$\begin{aligned} \min_{\mathbf{l}_d[n]} \quad & \frac{1}{N|\mathcal{U}|}a_{d,u}[n]k_{d,u}[n]P_{d,u}[n] + \\ & \frac{1}{N|\mathcal{U}|}\sum_{u=1}^{|\mathcal{U}|}\sum_{d=1}^{|\mathcal{D}|} a_{d,u}(\sum_{\bar{n}=1,\bar{n}\neq n}^{N} k_{d,u}[\bar{n}]P_{d,u}[\bar{n}]) \\ s.t. \quad & (4.8j), (4.8m), \quad \mathbf{l}_d[1] = \mathbf{l}_d[N+1] \quad \forall d. \end{aligned} \tag{4.12}$$

Keeping other $\mathbf{l}_d[\bar{n}],\ \forall d, u, \bar{n} \neq n$ fixed, the second part of (4.12) turns to be constant. With the given $\mathbf{H}$, we can prove that $P_{d,u}[n]$ is a quasi-convex and non-decreasing function to D2U horizontal distance $r_{d,u}[n], \forall d, u, n$. Therefore, minimizing the objective function in (4.12) equals minimizing $r_{u,d}[n]^2 = \|\mathbf{l}_d[n] - \mathbf{l}_u\|^2$, which is a quadratic convex optimization problem for $\mathbf{l}_d[n]$:

$$\begin{aligned} \min_{\mathbf{l}_d[n]} \quad & \|\mathbf{l}_d[n] - \mathbf{l}_u\|^2 \\ s.t. \quad & (4.8j), (4.8m), \quad \mathbf{l}_d[1] = \mathbf{l}_d[N+1], \quad \forall d. \end{aligned} \tag{4.13}$$

It is worth noting that the feasible region of $\mathbf{l}_d[n]$ constrained by (4.8m) can form a convex set in any X-Y plane by ignoring the working-zone burst close to the BS antenna [23].

4.4.4 DC Flying Height Optimization

Similar to problem (4.11), the sub-problem optimizing $\mathbf{H}$ is also non-convex with given $\mathbf{A}$, $\mathbf{K}$ and $\mathbf{L}$. Further decoupling $\mathbf{H}$ into $h_d[n]$ $\forall d, n$, the sub-problem to optimize each $h_{u,d}[n]$

is

$$\begin{aligned}\min_{h_d[n]} \quad & \frac{1}{N|\mathcal{U}|}a_{d,u}[n]k_{d,u}[n]P_{d,u}[n]+ \\ & \frac{1}{N|\mathcal{U}|}\sum_{u=1}^{|\mathcal{U}|}\sum_{d=1}^{|\mathcal{D}|}a_{d,u}(\sum_{\bar{n}=1,\bar{n}\neq n}^{N}k_{d,u}[\bar{n}]P_{d,u}[\bar{n}]) \\ s.t. \quad & (4.8k),(4.8m), \quad h_d[1]=h_d[N+1] \quad \forall d.\end{aligned} \tag{4.14}$$

To solve (4.14), we first transform (4.6) as the summation of one function of $r_{d,u}[n]$ and one function of $\theta_{d,u}[n]=\arctan(h_d[n]/r_{d,u}[n])$:

$$P_{d,u}[n]=20\log(\frac{4\pi f_c}{c}r_{d,u}[n])+\eta_{\text{NLoS}}+F(\theta_{d,u}[n]) \tag{4.15}$$

where

$$\begin{aligned}F(\theta_{d,u}[n]) &= 20\log(\sec(\theta_{d,u}[n]) \\ &\quad + \frac{\eta_{\text{LoS}}-\eta_{\text{NLoS}}}{1+a\exp(-b(\theta_{d,u}[n]-a))}.\end{aligned} \tag{4.16}$$

Since $r_{d,u}[n]$ is pre-defined with given $\mathbf{L}$, (4.15) is constant and optimizing $h_d[n]$ equals optimizing $\theta_{d,u}[n]$ in (4.14) by substituting $r_{d,u}[n]\tan(\theta_{d,u}[n])$ for corresponding $h_d[n]$. Fig. 4.2 shows the curves of $P_{d,u}[n]$ versus $\theta_{d,u}[n]$ under different $r_{d,u}[n]$. For (4.15), we can prove the following Proposition 2:

Proposition 2. *The (4.15) is a quasi-convex function to $\theta_{d,u}[n]$ with only one global minimum.*

Proof. Since the first two parts of (4.15) are constant or the function of $r_{d,u}[n]$, analyzing the convexity of (4.15) is equal to analyzing the convexity of (4.16). (4.16) can be regarded as the summation of two functions:

$$\begin{aligned}F(\theta_{d,u}[n]) &= F_1(\theta_{d,u}[n])+F_2(\theta_{d,u}[n]) \\ &= 20\log(\sec(\theta_{d,u}[n])+F_2(\theta_{d,u}[n]) \\ &= 20\log(\sec(\theta_{d,u}[n])+\frac{\eta_{\text{LoS}}-\eta_{\text{NLoS}}}{1+a\exp(-b(\theta_{d,u}[n]-a))}.\end{aligned} \tag{4.17}$$

The first-order derivations of $F_1(\theta_{d,u}[n])$ and $F_2(\theta_{d,u}[n])$ are

$$F_1'(\theta_{d,u}[n])=\frac{20}{\ln(10)}\tan(\theta_{d,u}[n]), \tag{4.18a}$$

$$F_2'(\theta_{d,u}[n])=\frac{ab(\eta_{\text{LoS}}-\eta_{\text{NLoS}})\exp(-b(\theta_{d,u}[n]-a))}{(1+a\exp(-b(\theta_{d,u}[n]-a)))^2}. \tag{4.18b}$$

We further calculate the second-order derivations of $F_1(\theta_{d,u}[n])$ and $F_2(\theta_{d,u}[n])$ as

$$F_1''(\theta_{d,u}[n]) = \frac{20}{\ln(10)} \sec^2(\theta_{d,u}[n]), \tag{4.19a}$$

$$F_2''(\theta_{d,u}[n]) = \frac{ab^2(\eta_{\text{NLoS}} - \eta_{\text{LoS}}) \exp(-b(\theta_{d,u}[n] - a))}{(1 + a\exp(-b(\theta_{d,u}[n] - a)))^2} - \frac{2a^2b^2(\eta_{\text{NLoS}} - \eta_{\text{LoS}}) \exp(-2b(\theta_{d,u}[n] - a))}{(1 + a\exp(-b(\theta_{d,u}[n] - a)))^3}. \tag{4.19b}$$

Note that (4.19a) is always larger than zero for all $\theta_{d,u}[n] \in [0°, 90°]$, so (4.18a) is proved to be non-decreasing function. Let (4.19b) equals zero, we can calculate the only root of (4.19b) $\theta_{d,u}[n]_{\text{root}} = a + \ln(a)/b$ at which (4.18b) achieves its global minimum.

Given the suburban scenario parameters $(\eta_{\text{LoS}}, \eta_{\text{NLoS}}, a, b) = (0.1, 21, 4.88, 0.43)$, we can calculate that:

$$F_1'(90°) + F_2'(90°) \geq 0, \tag{4.20a}$$

$$F_1'(\theta_{d,u}[n]_{\text{root}}) + F_2'(\theta_{d,u}[n]_{\text{root}}) \leq 0, \tag{4.20b}$$

$$F_1'(0°) + F_2'(0°) \leq 0. \tag{4.20c}$$

which indicates $F_1'(\theta_{d,u}[n]) + F_2'(\theta_{d,u}[n])$, i.e., (4.22), has only one root between $\theta_{d,u}[n]_{\text{root}}$ and 90°. Therefore, we can conclude that (4.16) is a uni-modal function with only one global minimum for all $\theta_{d,u}[n] \in [0°, 90°]$.

Define the $\theta_{d,u}[n]$ achieves the global minimum of (4.16) as $\theta_{d,u}[n]_{\text{opt}}$, $a \in [0°, 90°]$, $b \in [0°, 90°]$ and $t = \lambda a + (1 - \lambda)b \ \forall \lambda \in [0, 1]$. If $a \leq b \leq \theta_{d,u}[n]_{\text{opt}}$, (4.16) is non-increasing function to t and $F(b) \leq F(t) \leq F(a)$. If $\theta_{d,u}[n]_{\text{opt}} \leq a \leq b$, (4.16) is non-decreasing function to t and $F(a) \leq F(t) \leq F(b)$. If $a \leq \theta_{d,u}[n]_{\text{opt}} \leq b$, (4.16) ensures that $F(t) \leq \max\{F(a), F(b)\}$. Since that, we can argue that for any $a \in [0°, 90°]$ and $b \in [0°, 90°]$:

$$F(\lambda a + (1 - \lambda)b) \leq \max\{F(a), F(b)\} \quad \forall \lambda \in [0, 1] \tag{4.21}$$

which corresponds to the definition of quasi-convex function. Therefore, (4.15) is a quasi-convex function with only one global minimum. □

To obtain the optimal $\theta_{d,u}[n]_{\text{opt}}$ at which $P_{d,u}[n]$ reaches the global minimum, we let the first-order derivation of $P_{d,u}[n]$ to $\theta_{d,u}[n]$ equals zero:

$$\frac{\partial P_{d,u}[n]}{\partial \theta_{d,u}[n]} = \frac{20}{\ln(10)} \tan(\theta_{d,u}[n]) + \frac{ab(\eta_{\text{LoS}} - \eta_{\text{NLoS}}) \exp(-b(\theta_{d,u}[n] - a))}{(1 + a\exp(-b(\theta_{d,u}[n] - a)))^2} = 0. \tag{4.22}$$

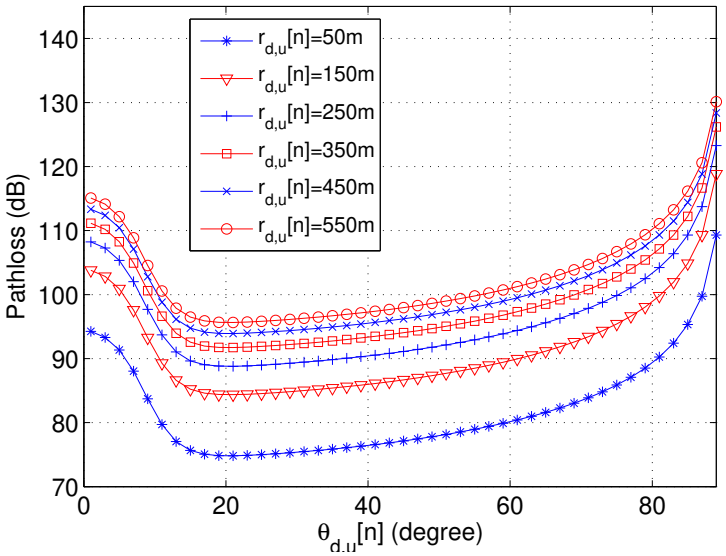


Figure 4.2: $P_{d,u}[n]$ versus $\theta_{d,u}[n]$.

(4.22) is a transcendental equation without closed-form solution. However, considering the fact that (4.15) has only one global minimum which is the single solution for (4.22), we can further calculate the second-order derivation of $P_{d,u}[n]$ to $\theta_{d,u}[n]$:

$$\begin{aligned}\frac{\partial^2 P_{d,u}[n]}{\partial \theta_{d,u}[n]^2} &= \frac{20}{\ln(10)} \sec^2(\theta_{d,u}[n]) \\ &+ \frac{2a^2b^2(\eta_{\text{LoS}} - \eta_{\text{NLoS}}) \exp(-2b(\theta_{d,u}[n] - a))}{(1 + a\exp(-b(\theta_{d,u}[n] - a)))^3} \\ &- \frac{ab^2(\eta_{\text{LoS}} - \eta_{\text{NLoS}}) \exp(-b(\theta_{d,u}[n] - a))}{(1 + a\exp(-b(\theta_{d,u}[n] - a)))^2}.\end{aligned} \tag{4.23}$$

Then, the $\theta_{d,u}[n]_{\text{opt}}$ can be calculated through the Newton-Raphson method:

$$\theta_{d,u}[n]_{i+1} = \theta_{d,u}[n]_i - \frac{P_{d,u}[n]'(\theta_{d,u}[n]_i)}{P_{d,u}[n]''(\theta_{d,u}[n]_i)} \tag{4.24}$$

where the iteration stops when $\theta_{d,u}[n]_{i+1} - \theta_{d,u}[n]_i \leq \epsilon$ and $\theta_{d,u}[n]_{\text{opt}} = \theta_{d,u}[n]_{i+1}$. The calculation of $\theta_{d,u}[n]_{\text{opt}}$ is constrained by $h_d[n]_{\text{d}} \leq r_{d,u}[n]\tan(\theta_{d,u}[n]) \leq h_d[n]_{\text{u}}$ where $h_d[n]_{\text{d}}$ and $h_d[n]_{\text{u}}$ are upper and lower bounds of $h_d[n]$ due to D2B link quality constraint. After

obtaining $\theta_{d,u}[n]_{\text{opt}}$ for each $\theta_{d,u}[n]$, the optimal $h_d[n]_{\text{opt}}$ can be calculated as

$$h_d[n]_{\text{opt}} = \begin{cases} h_d[n]_{\text{u}}, & \theta_{d,u}[n]_{\text{opt}} \geq \arctan(\frac{h_d[n]_{\text{up}}}{r_{d,u}[n]}) \\ h_d[n]_{\text{d}}, & \theta_{d,u}[n]_{\text{opt}} \leq \arctan(\frac{h_d[n]_{\text{down}}}{r_{d,u}[n]}) \\ r_{d,u}[n]\tan(\theta_{d,u}[n]_{\text{opt}}), & \text{otherwise.} \end{cases} \tag{4.25}$$

4.4.5 Protect Distance Constraint

Given the assumption that all trajectories are closed curves in 3D space, the start slot of each trajectory can be any slots on the trajectory. Because all trajectories are assumed to have same length of period N, for any trajectory, different start slots can lead to different inter-DC distances at all following slots. Since that, it is efficient to schedule the start slot of each DC to prevent violating protect distance constraint (4.8l) in following slots. We address protect distance constraint in the start slot scheduling process due to two reasons: First, the feasible set of constraint (4.8l) is non-convex for trajectory related variables, i.e., $\mathbf{L}$ and $\mathbf{H}$, optimizing them in horizontal trajectory or height optimization problems can significantly increase the problem complexity. Besides, by ensuring the protect distance constraint through start slot scheduling, the DC trajectory planning result can be maintained, which achieves better average D2U pathloss than modifying those optimized trajectories. Since the start slot scheduling is not an optimization problem, we can accept any start slots set as long as it ensures constraint (4.8l). In this work, we apply a greedy-based searching algorithm, as shown in Algorithm 3, to iteratively schedule the start slots of all DCs $d_i\ \forall i = 1, 2, \ldots, |\mathcal{D}|$. In each iteration, d_i sequentially sets its start slot as $n_j\ \forall j = 1, 2, \ldots, N$ and calculate the 3D distances between d_i and previous scheduled $d_k\ \forall k \leq i$ at every slots. If any start slot n_j ensures protect distance constraint at every slots, the start slot of d_i is temporally scheduled to n_j and break to the d_{i+1} iteration. If all $n_j\ \forall j = 1, 2, \ldots, N$ on d_i's trajectory cannot ensure protect distance constraint, the algorithm abandons the current and all previous scheduled DCs and re-run the first iteration of d_1 with updated start slot $n_1 = n_1 + 1$. Algorithm 3 stops until all d_i are scheduled with feasible start slots.

4.4.6 Proposed Algorithm

By decoupling the decision variable set into multiple blocks, i.e., $\mathbf{A}_t$, $\mathbf{K}_t$, $\mathbf{l}_d[n]\ \forall d, n$, $h_d[n]\ \forall d, n$, each block's sub-problem can be optimized respectively with other blocks keeping constant. Therefore, the problem (4.8) can be solved through iteratively optimizing

Algorithm 3 Start slots scheduling algorithm

Generate start slots set $\mathcal{S} = \{n_1 = 1, n_2 = 1, \ldots, n_{|\mathcal{D}|} = 1\}$ for DC $d_1, d_2, \ldots, d_{|\mathcal{D}|}$.
for $i = 1, 2, \ldots, |\mathcal{D}|$ **do**
    for $j = 1, 2, \ldots, N$ **do**
        Set d_i's start slot as n_j.
        Calculate distance between d_i and d_k $\forall k \leq i$ at all slots with d_i starts at n_j.
        Break if all distances are larger than $Z_{\min}$
    end for
    if $Z_{\min}$ is violated for all $n_j \in N$ **then**
        Set $i = 1,\ n_1 = n_1 + 1$.
    end if
end for

those sub-problems until the results converge, which yields to the classic BCD method. Based on the BCD method, we propose the algorithm to solve the multi-DC 3D trajectory planning and scheduling problem, which shows in Algorithm 4. $\mathbf{A}_t$, $\mathbf{K}_t$, $\mathbf{W}_t$ denote the AoI association, D2U communication scheduling and DC trajectories after each iteration t, respectively. $\mathbf{W}_t$ is composed by $\mathbf{L}_t$ and $\mathbf{H}_t$. According to the BCD method, the proposed algorithm ensures convergence since the global optimal results of all sub-problem are accurately achieved [39] [110].

Initial trajectories are required for the first iteration of AoI association. Without loss of generality, we set the same initial height $h_0 \in [h_d[n]_d, h_d[n]_u]$ for all DC. For each DC, we apply a circle initial trajectory with radius $r_0 = 1$ m. Given the assumption in subsection A that the initial D2U communication scheduling $k_{d,u}[n]$ equals one for $\forall d, u, n$, it is better to deploy the center of each circle trajectory to the position where the summation of D2U pathloss between its adjacent AoIs (will be associated to the DC with high probability) is minimized. Therefore, the classic k-means algorithm can be effectively applied to determine the initial trajectory center of each DC by substituting D2U pathloss for the geometric distance in original algorithm. To reduce the convergence time and improve the result quality, we apply the k-means ++ algorithm which prefers centroid seeds with large mutual distances [111].

4.5 Numerical Results

We conduct extensive simulations to verify the performance of our proposed algorithm in minimizing average D2U pathloss of the network. The simulations are link level without

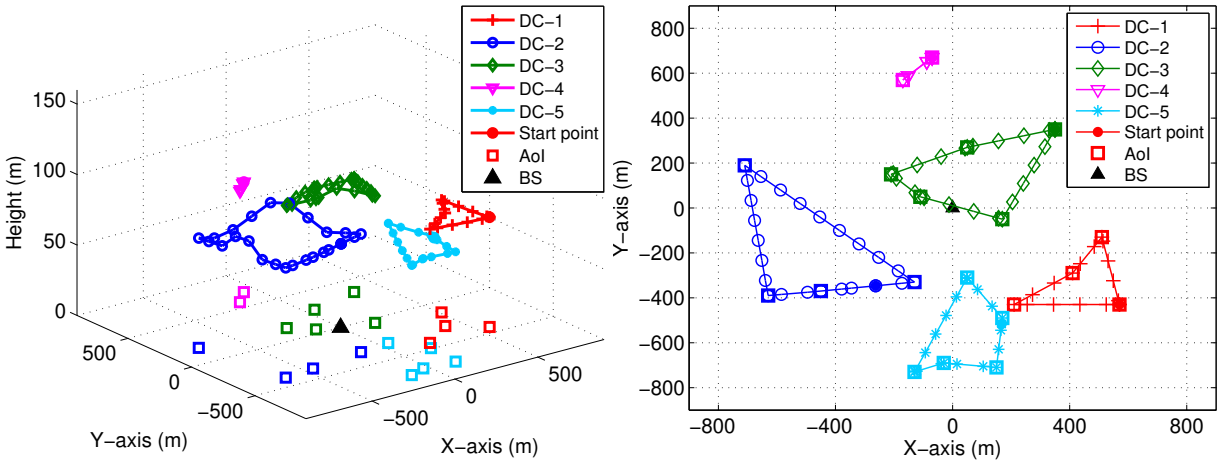
Algorithm 4 Multi-DC 3D trajectory planning and scheduling algorithm

Initiate initial D2U communication scheduling $\mathbf{K}_0$, initial height h_0.
Calculate initial horizontal trajectory $\mathbf{L}_0$ through k-means ++ algorithm.
Set $t = 1$, $\Delta W = \infty$.
while $\Delta W \geq \epsilon$ **do**
    Solve problem (4.9) to obtain $\mathbf{A}_t$ by treating $\mathbf{K}_{t-1}$, $\mathbf{L}_{t-1}$ and $\mathbf{H}_{t-1}$ as constants.
    Solve problem (4.10) to obtain $\mathbf{K}_t$ by treating $\mathbf{A}_t$, $\mathbf{L}_{t-1}$ and $\mathbf{H}_{t-1}$ as constants.
    for $d \in \mathcal{D}, n = 1, 2, \ldots, N$ **do**
        Solve problem (4.12) to obtain $\mathbf{L}_d[n]_{\text{opt}}$ by treating $\mathbf{A}_t$, $\mathbf{K}_t$, $\mathbf{H}_{t-1}$ and $\mathbf{L}_d[\bar{n}]$ $\forall d, \bar{n} \neq n$ as constants.
        Update $\mathbf{L}_t$ with $\mathbf{l}_d[n]_{\text{opt}}$.
    end for
    for $d \in \mathcal{D}, n = 1, 2, \ldots, N$ **do**
        Solve problem (4.14) to obtain $h_d[n]_{\text{opt}}$ by treating $\mathbf{A}_t$, $\mathbf{K}_t$, $\mathbf{L}_t$ and $h_d[\bar{n}]$ $\forall d, \bar{n} \neq n$ as constants.
        Update $\mathbf{H}_t$ with $h_d[n]_{\text{opt}}$.
    end for
    Update $\mathbf{W}_t$ with $\mathbf{L}_t$ and $\mathbf{H}_t$.
    $t = t + 1$.
    $\Delta W = \mathbf{W}_t - \mathbf{W}_{t-1}$.
end while
Run Algorithm 3 to ensure protect distance constraint.

simulating specific MAC or upper layers protocols. The BS is located at the origin point (coordinate $(0, 0, 0)$) and the side-length of grid is set to 20 m. Both D2U and D2B pathloss models are configured in suburban scenario. To provide additional spectrum resources for DA-RAN and reduce the interference to terrestrial RAN users, the frequency band of D2U communication f_c is expected to be different from the licensed cellular band. Like most commercial drone products [29] [112] and DC related works [113] [44] [51] [75], we use the 2.4 GHz unlicensed band as the carrier for D2U communications. D2B communications use the 850 MHz LTE band according to the D2B pathloss model [24]. By allocating different carrier frequencies, the interference between D2U and D2B communication can be prevented. Initial height h_0 is set to 80 m within the working-zone of DC over the whole BS radio coverage area [23]. We treat δ_t as the minimal time unit to calculate related variables including $V_{\max}$, $H_{\max}$, etc. There is no need to assign specific value for δ_t in the simulation, however, according to the general specifications of commercial drones ($50 - 70$ km/h for horizontal speed, $3 - 5$ m/s for ascent/descent speed) [50], the approximate value of δ_t is around 10 s. Table. 4.1 shows detail simulation parameters.

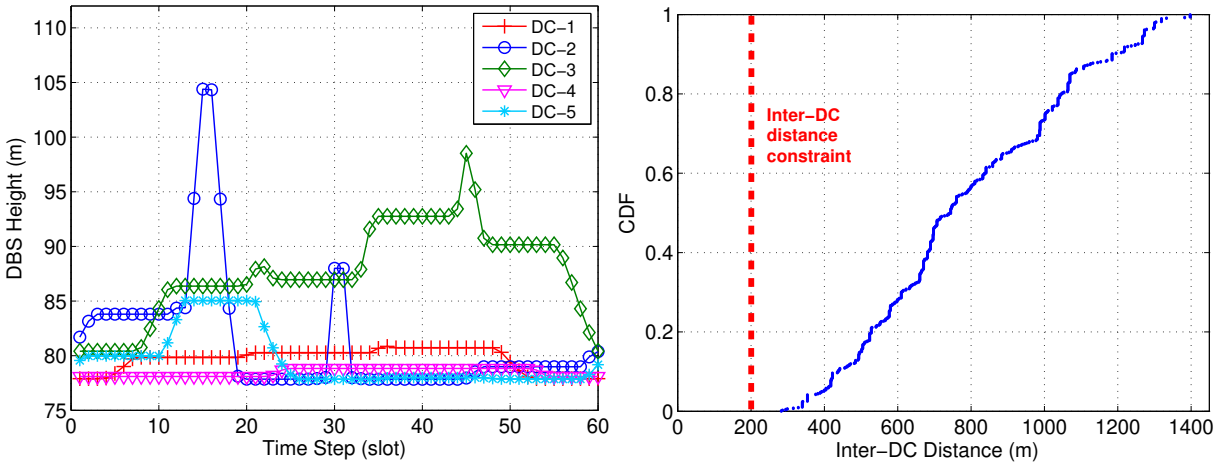
Table 4.1: Simulation Parameters for Multi-DC 3D Trajectory Planning and Communication Scheduling

Parameters	Numerical Values
BS radio coverage radius r_{BS}	900 m
AoI number $\lvert\mathcal{U}\rvert$	20
D2U parameters $(\eta_{\text{LoS}}, \eta_{\text{NLoS}}, a, b)$	(0.1,21,4.88,0.43)
D2B parameters $(\alpha, A, \theta_0, B, \eta_0)$	(3.04,-23.29,-3.61,4.14,20.7)
Carrier frequencies (D2U, D2B)	(2.4 GHz, 850 MHz)
Slots amount in one period N	60 slot
D2B pathloss constraint P_{DB}	80 dB
Minimal per-AoI slot number $S_{\min}$	10 slot
Maximal per-DC AoI number $\lvert\mathcal{A}_d\rvert_{\max}$	6
Maximal horizontal speed $V_{\max}$	$30, 50, 70, 90, 110$ m/slot
Maximal vertical speed $H_{\max}$	10 m/slot
Protect Distance $Z_{\min}$	200 m
Trajectory difference ϵ	0.1 m for each slot



(a) 3D view

(b) 2D projection on X-Y plane



(c) Heights

(d) Protect distance guarantee

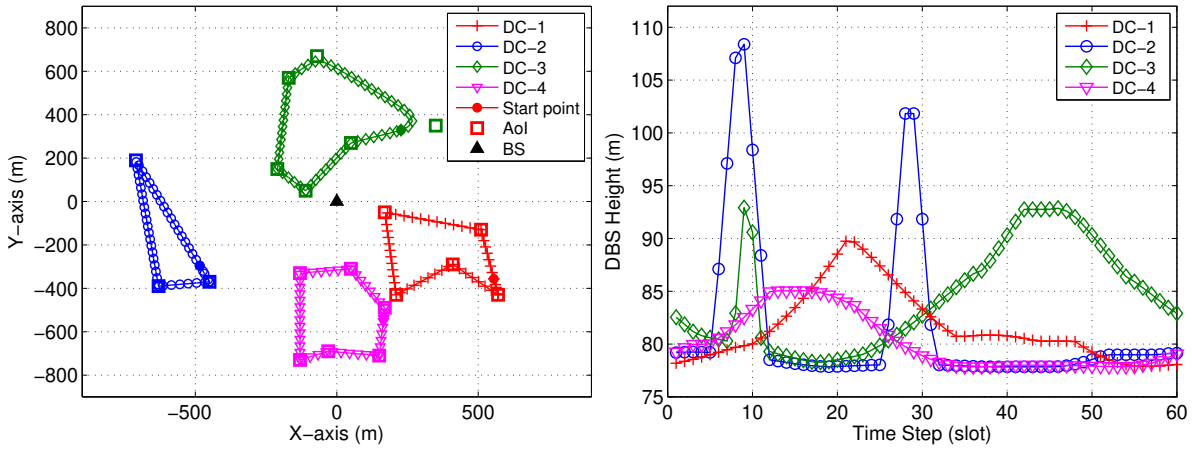
Figure 4.3: Trajectory planning results of 5 DCs serving 20 AoIs.

4.5.1 3D Trajectory Planning for Multiple DCs

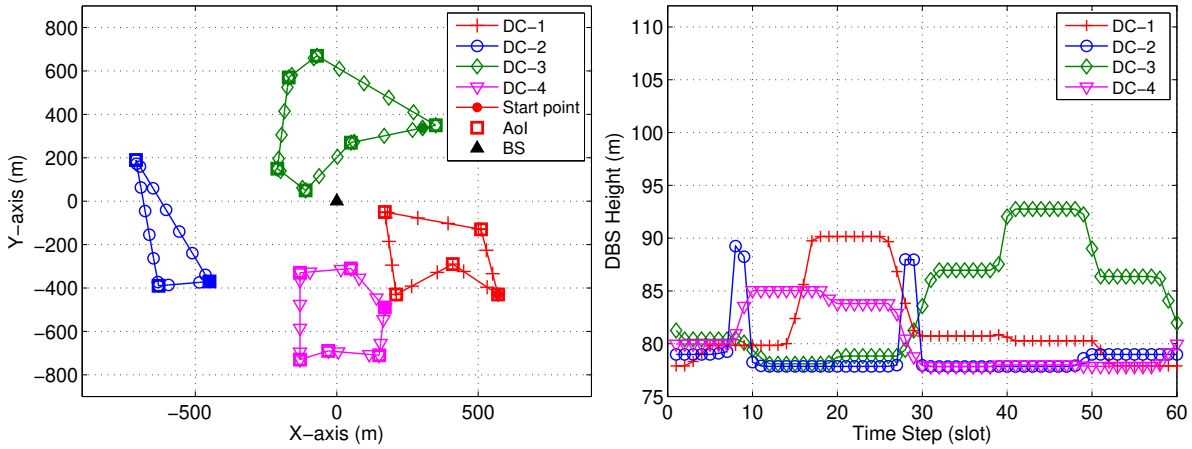
Fig. 4.3 shows the scenario where the trajectories of five DCs are optimized to serve twenty AoIs with $V_{\max} = 90$ m/slot. The closed curves dotted by different markers in Fig. 4.3(a) and Fig. 4.3(b) denote different DC trajectories; the squares on the X-Y plane represent AoIs. AoIs are associated to corresponding DCs with same colors. Fig. 4.3(c) illustrates the changes of flying height within one period. As shown in Fig. 4.3(a) and 4.3(b), for each DC, the optimized trajectory can fly over all its associated AoIs and form a closed curve in 3D space. In Fig. 4.3(c), all DC flying height curves are lower bounded around 78 m, which is the lower bound of $h_d[n]_d$ due to the D2B pathloss constraint.

Note that for each trajectory in Fig. 4.3(b), the summation of dots on the section between two AoIs is less than 50% of the total slots number N. Besides, the height curves in Fig. 4.3(c) show the trend to maintain fixed values for consecutive slots. Combining Figs. 4.3(b) and 4.3(c), we can justify that the remaining dots in Fig. 4.3(b) are overlapped above the associated AoIs, and those overlapped dots corresponds to the consecutive slots have fixed heights in Fig. 4.3(c). In other word, the proposed algorithm prefers hovering DCs above the associated AoIs, while leave minimal slots for the travelling process between adjacent hovering positions. Such a "hovering effect" can be explained as follows. For any slot n, the proposed algorithm is prone to small $r_{d,u}[n]$ which minimizes the D2U pathloss with given $h_d[n]$. On the contrary, the smaller the $r_{d,u}[n]$ is, the higher the probability that $r_{d,u}[n]\theta_{d,u}[n]_{\text{opt}} \leq h_d[n]_{\text{d}}$. Based on (4.25), if $r_{d,u}[n]\theta_{d,u}[n]_{\text{opt}} \leq h_d[n]_{\text{d}}$, the optimal height equals the lower bound of flying height $h_d[n]_{\text{d}}$ at current position. Since the minimal $r_{d,u}[n]$ equals zero, the minimal average D2U pathloss can be achieved by the trajectory with most slots hovering above the AoIs. From Fig. 4.3(c), we can see that several height bursts occur when each DC is flying between two AoI with a long inter-AoI distance. At those slots, the DC is relatively far from the scheduled AoI and $r_{d,u}[n]\tan(\theta_{d,u}[n]_{\text{opt}})$ can fall in the feasible height range between $h_d[n]_d$ and $h_d[n]_u$ constrained by D2B pathloss threshold. Therefore, the optimal heights prefer to approach the value of $r_{d,u}[n]\tan(\theta_{d,u}[n]_{\text{opt}})$ in those slots, which leads to the height bursts. Fig. 4.3(d) shows the CDF of inter-DC distance within one period T. The red dotted line is the protect distance constraint. It can be seen that all inter-DC distances of the final trajectory planning result are larger than the protect distance threshold $Z_{\min}$, which indicates the effectiveness of the start slots scheduling algorithm to ensure protect distance constraint.

Fig. 4.4 presents two groups of trajectory planning results with $V_{\max} = 30$ m/slot and $V_{\max} = 110$ m/slot, respectively. The available DC number $|\mathcal{D}|$ equals four for both groups. As shown in Fig. 4.4, the AoI associations are same under different $V_{\max}$. The trajectories in Fig. 4.4(a) cannot fly over every associated AoI since the maximal horizontal speed is too small to ensure the DCs to approach every associated AoIs within one period. In

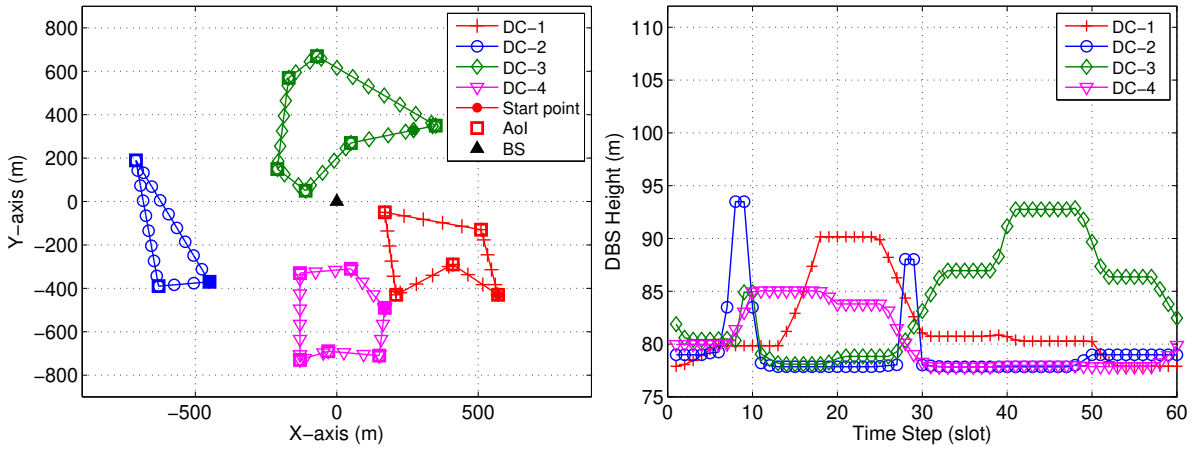


(a) $V_{\max} = 30$ m/slot (b) Heights of $V_{\max} = 30$ m/slot

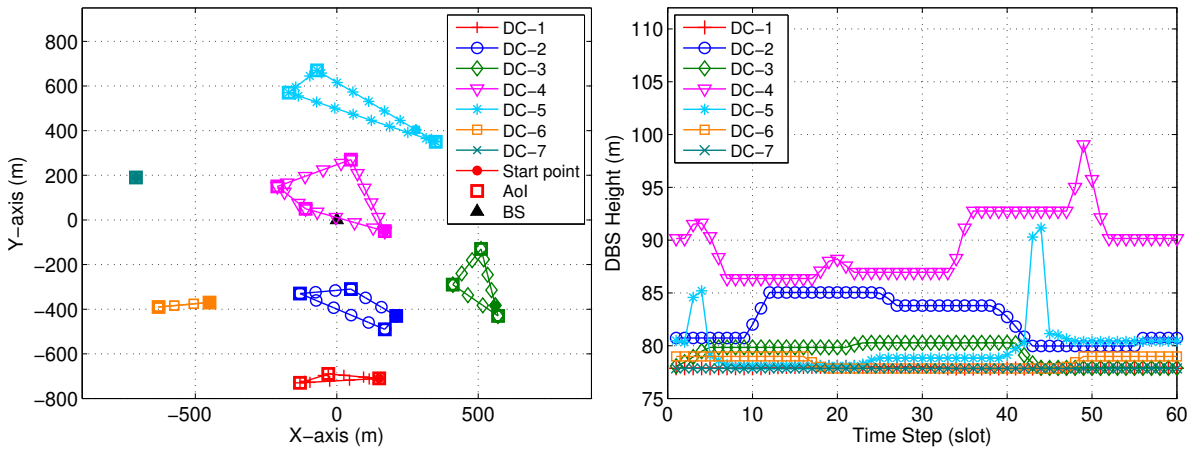


(c) $V_{\max} = 110$ m/slot (d) Heights of $V_{\max} = 110$ m/slot

Figure 4.4: Trajectory planning results impacted by horizontal speeds.



(a) $|\mathcal{D}| = 4$ (b) Heights of $|\mathcal{D}| = 4$



(c) $|\mathcal{D}| = 7$ (d) Heights of $|\mathcal{D}| = 7$

Figure 4.5: Trajectory planning results impacted by DC number.

Fig. 4.4(c) where $V_{\max}$ is high enough, the DCs can even hovering on each associated AoI for few slots since the flying interval between two hovering positions requires less slots. Comparing Fig. 4.4(b) and 4.4(d), we can see that the variation of flying height with $V_{\max} = 30$ m/slot is larger than the height variation with $V_{\max} = 110$ m/slot. Because in low $V_{\max}$ scenario, the traveling process between two AoIs requires more slots than that in the high $V_{\max}$ scenario, the optimal height $r_{d,u}[n]\tan(\theta_{d,u}[n]_{\text{opt}})$ in small $V_{\max}$ has higher probability to fall into the feasible flying height range. Considering the four DCs trajectory planning scenario, the CDF of D2U pathloss under different $V_{\max}$ are compared in Fig. 4.6. Given any pathloss threshold, We can see that the probability of D2U pathloss less than the threshold raises as the $V_{\max}$ increases.

Fig. 4.5 compares the trajectory planning results when $|\mathcal{D}| = 4$ and $|\mathcal{D}| = 7$. The horizontal speed is set as $V_{\max} = 70$ m/slot for both scenarios. As the number of available DC increases, the average number of AoI associated to one DC is reduced, some trajectories can even degenerate to one static deployment position when the corresponding DC is associated with only one AoI. On the other hand, since the average $r_{d,u}[n]$ length is also reduced with the decreasing of associated AoI number for each DC, the variation of flying height can be reduced with the increasing of $|\mathcal{D}|$. Fig. 4.7 shows the CDF of D2U pathloss under different $|\mathcal{D}|$. Similar to Fig. 4.6, the probability of D2U pathloss less than any given threshold increases as more numbers of DC are provided.

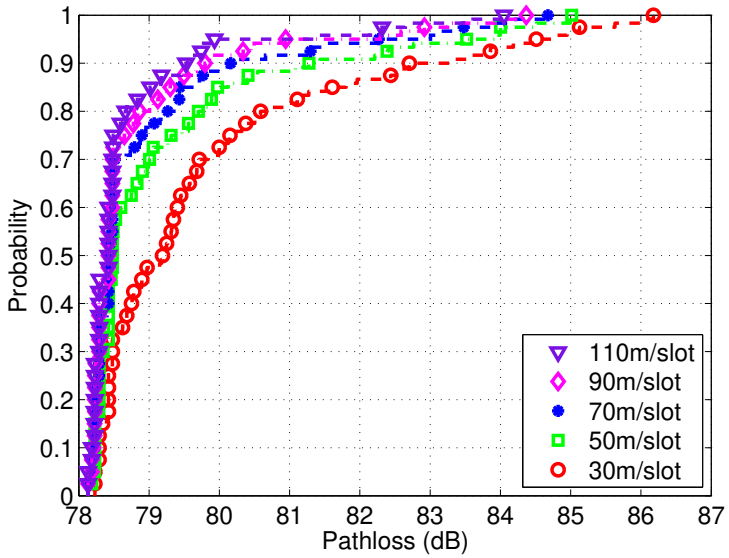


Figure 4.6: CDF of D2U pathloss under different $V_{\max}$.

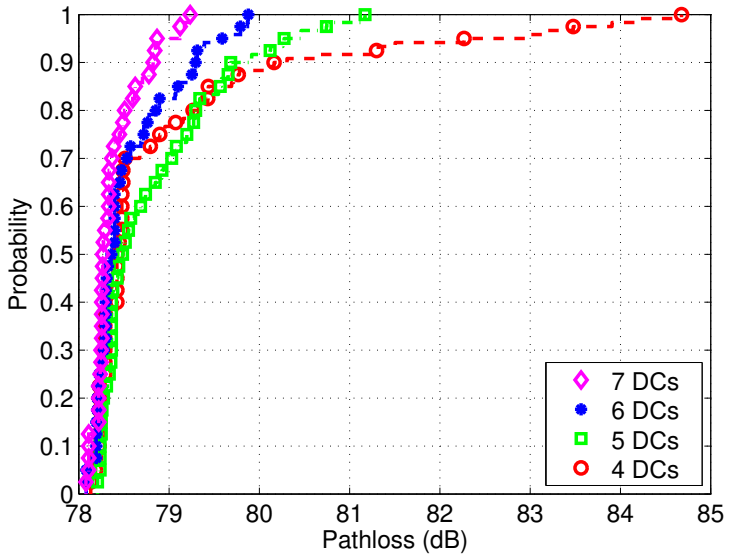


Figure 4.7: CDF of D2U pathloss under different $|\mathcal{U}|$.

Fig. 4.6 and Fig. 4.7 indicate that the D2U pathloss performance can be influenced by both $V_{\max}$ and $|\mathcal{D}|$. In Fig. 4.8, we further investigate the average pathloss performance with different $V_{\max}$ and $|\mathcal{D}|$. From preceding analyses, we know that both the higher $V_{\max}$ and larger $|\mathcal{D}|$ can lead to smaller average $r_{d,u}$, which eventually reduces the average pathloss level. According to Fig. 4.8, given the same number of DC, the average pathloss level decreases slightly as the $V_{\max}$ increases; while significant average pathloss level reduction occurs as the $|\mathcal{D}|$ increases under fixed maximal horizontal speed. Therefore, we can conclude that both raising the horizontal speed and increasing the number of available DC can promote the average D2U pathloss performance of DC trajectory planning, while increasing the number of available DCs is proved to be more efficient than raising horizontal flying speed. Note that this conclusion is valid for the average D2U pathloss performance of the whole network only. Since the standard deviations of pathloss (error-bars) plotted in Fig. 4.8 are highly overlapped with each other, the D2U pathloss performance of specific DC-to-AoI pair can vary a lot. Nevertheless, as the horizontal speed and available DC number increase, the standard deviation is reduced, which indicates that the user fairness can also be promoted by raising $V_{\max}$ and $|\mathcal{D}|$.

We further analyze the impacts of different initial trajectories (ITs) to the achieved average D2U pathloss performance. We compare four types of ITs, i.e., 1) circle IT with the center location determined by k-means ++; 2) point IT (where the trajectory shrinks to one hovering point) with the point location determined by k-means ++; 3) circle IT with

the center location uniformly distributed over $|\mathcal{S}|_{\text{bs}}$; 4) point IT with the point location uniformly distributed over $|\mathcal{S}|_{\text{bs}}$. Fig. 4.9 shows the comparison result. We can see that the k-means-based circle IT used in the proposed algorithm achieves minimal average D2U pathloss and pathloss standard deviation. Comparing the performance gaps between ITs, we note that applying k-means ++ algorithm has more significant impact to both average D2U pathloss and pathloss standard deviation than using circle-shaped IT.

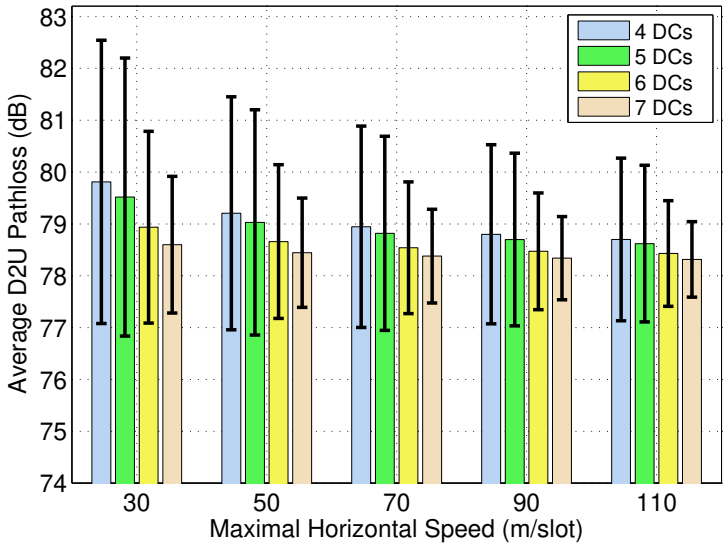


Figure 4.8: Average D2U pathloss with different horizontal speeds and DC numbers.

4.5.2 Performance Comparison

To highlights the efficiency of the proposed multi-DC 3D trajectory planning and scheduling, we compare the average D2U pathloss performance achieved by our proposed trajectory planning algorithm, as well as the static DC deployment scheme in Fig. 4.10 and Table 4.2. For static DC deployment algorithm, we use the DI-PSO algorithm proposed in [23]. Without loss of generality, we use all data achieved by $V_{\max} = 30, 50, 70, 90, 110$ m/slot to calculate the average D2U pathloss of the proposed trajectory planning algorithm. From Fig. 4.10, we can see that the average D2U pathloss of both algorithms are reduced as the available DC number increases. However, the D2U pathloss performance achieved by our trajectory planning algorithm maintains $10 - 15$ dB smaller than that achieved by the DI-PSO algorithm.

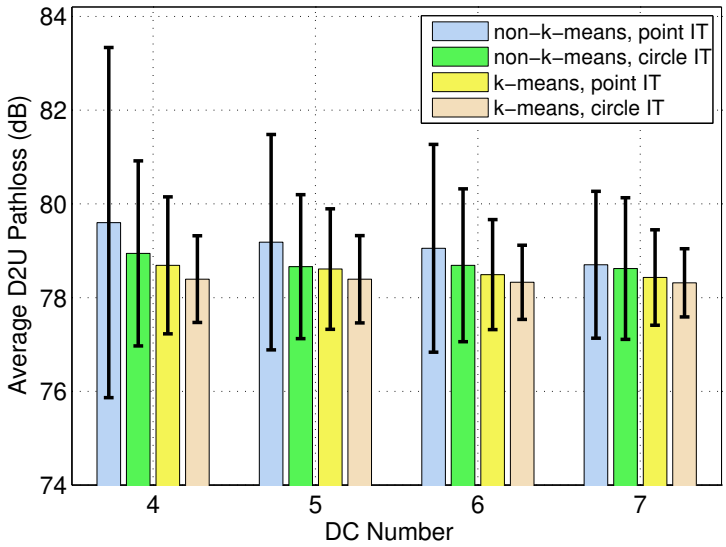


Figure 4.9: Initial trajectory comparison.

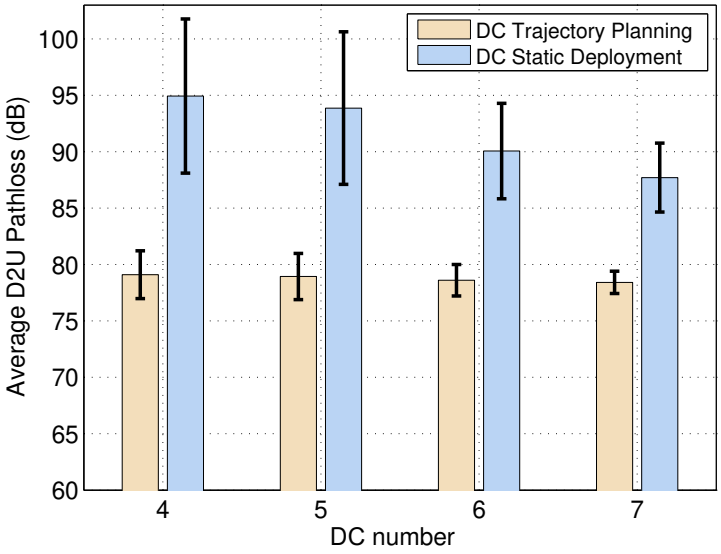


Figure 4.10: Average D2U pathloss comparison between trajectory planning and static deployment of multiple DCs.

The user fairness promotion provided by the proposed trajectory planning algorithm is indicated by the error-bars in Fig. 4.10 and the D2U pathloss standard deviation comparison in Table 4.2. σ_t and σ_s are D2U pathloss standard deviations for DC trajectory planning and static DC deployment, respectively. In Fig. 4.10, we can see that the standard deviation of D2U pathloss achieved by both algorithms are reduced as the available DC number increases. While the σ_t maintains less than half of the σ_s, there is no overlap between the error-bars achieved by two algorithms. From Table 4.2, we can calculate that the DC trajectory planning can lower the D2U pathloss standard deviation by 68.34% on average compared with the static DC deployment.

To highlight the cost-efficiency of the trajectory planning algorithm, and provide a guideline to determine the number of required DC, we compare the minimal required DC number under different D2U pathloss thresholds in Table 4.3. The D2U pathloss threshold is a strict constraint, which means that for any DC at any slot, the D2U pathloss to any AoI cannot exceed it. As shown in Table 4.3, under the same D2U pathloss threshold, the minimal required DC amount is decreased as the maximal horizontal speed of DC increases. Besides, given any threshold levels, the static DC deployment always requires two to four more DCs than the DC trajectory planning algorithm, which implies that the DC trajectory planning algorithm is more economical than the static DC deployment.

Table 4.2: D2U Pathloss Standard Deviation Comparison

DC number	4	5	6	7
σ_t	2.1203	2.0476	1.3923	0.9887
σ_s	6.8313	6.7562	4.2329	3.0530
$(\sigma_s - \sigma_t)/\sigma_s$	68.96%	69.69%	67.11%	67.61%

Table 4.3: Minimal DC Amounts Comparison

D2U pathloss threshold (dB)	98	95	92	89	86
30m/slot	3	4	5	6	8
30m/slot	3	4	5	6	8
70m/slot	3	4	4	6	7
90m/slot	3	4	4	5	7
110m/slot	3	3	4	4	6
Static deployment	6	7	8	10	10

4.6 Summary

In this chapter, we have studied the 3D trajectory planning and scheduling for multiple DCs in DA-RAN with the state-of-the-art D2U and D2B pathloss models considered. We have formulated the MINLP problem to minimize the average D2U pathloss achieved by multi-DC trajectory planning and scheduling. To solve the MINLP problem, we have decoupled the MINLP problem into multiple solvable sub-problems, and devised a BCD-based multi-DC 3D trajectory planning and scheduling algorithm in which the AoI association, D2U communication scheduling, horizontal trajectories, and flying altitudes of DCs are iteratively optimized. A start slot scheduling algorithm and a k-means-based circle IT have been proposed to ensure the protect distance constraint and generate initial DC trajectories. We have investigated the impacts of available DC number, horizontal speed and different IT on the achieved average D2U pathloss. Simulation results have shown that the proposed algorithm can achieve $10-15$ dB average D2U pathloss reduction, and promote pathloss standard deviation by 68% when compared with the static DC deployment algorithm.

Chapter 5

Hierarchical DRL-based Multi-Drone-Cell Trajectory Planning and Resource Allocation for High-Mobility Users

To address the high dynamics and uncertainties of high-mobility users, in this chapter, we propose a hierarchical DRL based multi-DC trajectory planning and resource allocation (HDRLTPRA) scheme. The objective of the multi-DC TPRA problem is maximizing the accumulative network throughput when satisfying user fairness, DC power consumption and DC-to-ground link quality constraints. We first decouple the multi-DC TPRA problem into two hierarchical sub-problems, i.e., the higher-level global trajectory planning sub-problem and the lower-level local TPRA sub-problem. In particular, the GTP sub-problem roughly plans trajectories for multiple DCs over a long time period in a large area, while the LTPRA sub-problem controls the real-time movement of single DC and resource allocation within its communication coverage. Then the HDRLTPRA scheme is designed, which consists of two customized DRL algorithms for the two sub-problems respectively. For the GTP sub-problem, we propose a multi-agent DRL based MARL-GTP algorithm which learns the joint trajectory planning policy for multiple DCs to maximize the accumulative number of users being served. The non-stationary state space caused by multi-DC environment is addressed by the multi-agent fingerprint technique. Given the global trajectory planning results, we further design a DDPG based DDPG-LTPRA algorithm for the LTPRA sub-problem. The DDPG-LTPRA algorithm is executed on each DC independently to adjust its movements and allocate transmit power over a continuous

action space, in response to the real-time user traffic variations. Simulations based on real-world scenario show the performance improvements provided by the proposed HDRLTPRA scheme over the model-based TPRA scheme.

5.1 Background and Motivations

In DA-RAN, DCs are dynamically deployed to serve users using additional spectrum resource, and relay data between served user and terrestrial BSs via D2B links. Since the DA-RAN's capability to address terrestrial traffic variations is enabled by DC's dynamic deployment feature, the DC trajectory planning problem, which designs flying traces of DCs to serve terrestrial users in specific areas, is essential for DA-RAN research. The specific areas include both the blind spots of terrestrial RAN's communication coverage, as well as the BTSs where allocated terrestrial RAN resources are inadequate to support the dense traffic, e.g. congested roads, stadium with sports events, etc. On the other hand, the resource allocation decisions of each DC can be adapted to fit the user distribution status in different deploying locations. Therefore, the trajectory planning and resource allocation for the DC should be jointly investigated, which forms the DC trajectory planning and resource allocation (TPRA) problem.

The DC TPRA problem has been studied by some pioneer works with the objectives of improving given performance metrics [39] [51], e.g. network throughput, QoS of users, etc. In these non-learning-based works, the terrestrial users or uncovered areas are modeled as static nodes for each trajectory planning and resource allocation process. If the user locations or traffic patterns change significantly, re-planning and re-allocation are executed. The trajectory of each DC is modeled as a closed curve composed by discrete 3D locations, the DC sequentially transverses each location and serves the associated users (covered areas) at scheduled time slots. Given the fixed user locations and known traffic model, a deterministic TPRA decision for each DC, which cannot adapt itself to traffic variations, is calculated through optimization methods. However, the deterministic trajectory and static user model is only applicable for fixed or low-mobility users with scheduled communications, e.g. data collection for massive Internet of things (IoT) devices [52]. To adapt to the non-static environment, the learning-based TPRA research leverages machine learning techniques to make TPRA decisions for DCs according to environment variations [53]. However, most of the existing works consider scenarios with single DC, fixed number of users, single BS communication coverage, or only designing the trajectory without resource allocation [85] [86] [87] [88] [114]. When focusing on a large RAN area with non-static users, the number of users is spatio-temporal variant due to the high user mobility. Besides, multiple DCs are required to cooperatively serve the ground users, which involves

potential mutual interference among DCs. Therefore, it is still challenging to propose a learning-based TPRA scheme in general environment with multiple DCs, non-fixed number of high-mobility users over large RAN areas.

In this chapter, we investigate the multi-DC TPRA problem to serve high-mobility users (e.g. vehicular users) over large RAN areas with multiple BSs. Considering the user fairness, DC energy consumption, D2U and D2B communication constraints, the multi-DC TPRA problem in high-mobility scenario is formulated as a constrained Markov decision process (CMDP) which aims at maximizing the long-term accumulative network throughput over the large area. However, the multi-DC TPRA problem is intractable for conventional DRL-based algorithms in complex environment due to the high spatio-temporal network dynamics and the inter-DC interference. Therefore, we propose a hierarchical DRL based HDRLTPRA scheme to decouple the highly complicated problem into two hierarchical sub-problems. The objective of higher-level global trajectory planning sub-problem is planning global trajectories for multiple DCs by the RAN controller to maximize the accumulated number of served users over a large area and a long time. The global trajectory determines the sequence of areas (e.g. BTSs) served by the DC. The GTP decision step (i.e., planning interval) for the global trajectory is relatively long to contain multiple fine-grained TPRA decisions within it. Based on the global trajectory decision at each GTP step, the lower-level local trajectory planning and resource allocation sub-problem is addressed by each DC independently to control real-time movement and allocate resources within the pre-determined area by GTP sub-problem. To solve the two sub-problems in HDRLTPRA, a multi-agent DRL based GTP (MARL-GTP) algorithm and a deep deterministic policy gradient (DDPG) based LTPRA (DDPG-LTPRA) algorithm are respectively designed. The main contributions of this work are three-folded:

- We propose an effective multi-DC HDRLTPRA scheme for the high-mobility scenario. In HDRLTPRA, the higher-level MARL-GTP algorithm addresses the complexities caused by multiple DCs and long-term variations of user distributions. The lower-level DDPG-LTPRA algorithm addresses the real-time variations of served user numbers and locations, with the state space constrained by the output of higher-level MARL-GTP algorithm. This hierarchical DRL framework allows the HDRLTPRA to converge to sub-optimal TPRA solutions with high probability.

- To generate the global trajectories, we design the MARL-GTP algorithm which implements fully cooperative multi-agent DRL to fit the multi-DC environment. In specific, the multi-agent fingerprints and prioritized experience replay (PER) methods are applied to design the hyper-parameters and NNs in MARL-GTP algorithm to address the non-stationary environment and sparse rewards.

- In response to the real-time user mobility, we design the DDPG-LTPRA algorithm executed by each DC independently to adjust the real-time DC flying control and resource allocation (transmit power). In DDPG-LTPRA algorithm, the DDPG enables the trajectory planning and resource allocation over continuous spaces. Besides, the complexity of DDPG-LTPRA's input are reduced by mathematical analyses of D2U communication, which further improves the convergence performance.

The remainder of the chapter is organized as follows. In Section 5.2, we introduce the system model in which high-mobility users and multiple DCs are involved. Then the multi-DC TPRA problem is formulated and decoupled into two hierarchical sub-problems in Section 5.3. In Section 5.4, the HDRLTPRA scheme is proposed with the higher-level MARL-GTP algorithm and the lower-level DDPG-TPRA algorithm. Real-world scenario based simulations are carried out in Section 5.5, followed by conclusions in Section 5.6.

5.2 System Model

5.2.1 DA-RAN Scenario with High-Mobility Users

Fig. 5.1 shows the DA-RAN scenario where two DCs embedded on rotary-wing drones are released by the DA-RAN to serve high-mobility users (e.g., vehicular users) over a large area. In this work, we consider downlink transmissions in DA-RAN where DCs relay data from BS to their served users using additional spectrum resources. Define the DA-RAN scenario as $\mathcal{G}$ whose area is large enough to contain multiple BSs, such as a university campus with its affiliated regions. In the DA-RAN scenario $\mathcal{G}$, denote $\mathcal{B} = \{b_1, b_2, \ldots, b_{|\mathcal{B}|}\}$ as the set of BSs with cardinality $|\mathcal{B}|$, and $\mathcal{D} = \{d_1, d_2, \ldots, d_{|\mathcal{D}|}\}$ as the set of DCs with cardinality $|\mathcal{D}|$. The communication coverage of each BS $b \in \mathcal{B}$ is modelled as a hexagon area with a circumscribed circle radius R_{b}. The blue and yellow areas in the right side of Fig. 5.1 represent the coverage of two BSs, respectively. The communication coverage of each DC $d \in \mathcal{D}$ is also a hexagon area with a circumscribed circle radius R_{d}, as shown in the left side of Fig. 5.1. To simplify the environment for DC trajectory planning, we evenly divided the whole scenario into a hexagon mesh where the size of each unit hexagon (unit) equals one DC's communication coverage. Denote each unit as g, we can redefine $\mathcal{G}$ as a set $\mathcal{G} = \{g_1, g_2, \ldots, g_{|\mathcal{G}|}\}$ with the cardinality $|\mathcal{G}|$. The communication coverage of BS b is further modelled as a set $\mathcal{G}_b = \{g_b \ldots\}$ containing multiple units covered by BS b. Note that $\mathcal{G}_b \cap \mathcal{G}_{\bar{b}} \geq \emptyset, \forall b \neq \bar{b}$ due to the coverage overlapping of adjacent BSs.

Because of the high-mobility feature, the total numbers and locations of users in the whole scenario $\mathcal{G}$ are random variables changing over time. Define the decision time step

t with length δ_t, during which both DCs and high-mobility users can only move within sufficiently small ranges with negligible effects on the network performance. The locations of each DC d and each user u at step t are represented by $z_d(t) = (x_d(t), y_d(t), h_d(t))$ and $(x_u(t), y_u(t))$, respectively, where x_d, y_d, h_d and x_u, y_u are 3D and 2D Cartesian coordinates, respectively. All users served by DC d at step t form a set $\mathcal{U}_d(t)$ with the cardinality (number of users) $|\mathcal{U}_d(t)|$, and all users in unit g at step t is represented by the set $\mathcal{U}_g(t)$ with the cardinality $|\mathcal{U}_g(t)|$. Since we mainly address users' high-mobility issue in this work, we assume all users are running the same type of services with homogeneous downlink data traffic patterns and bandwidth requirements.

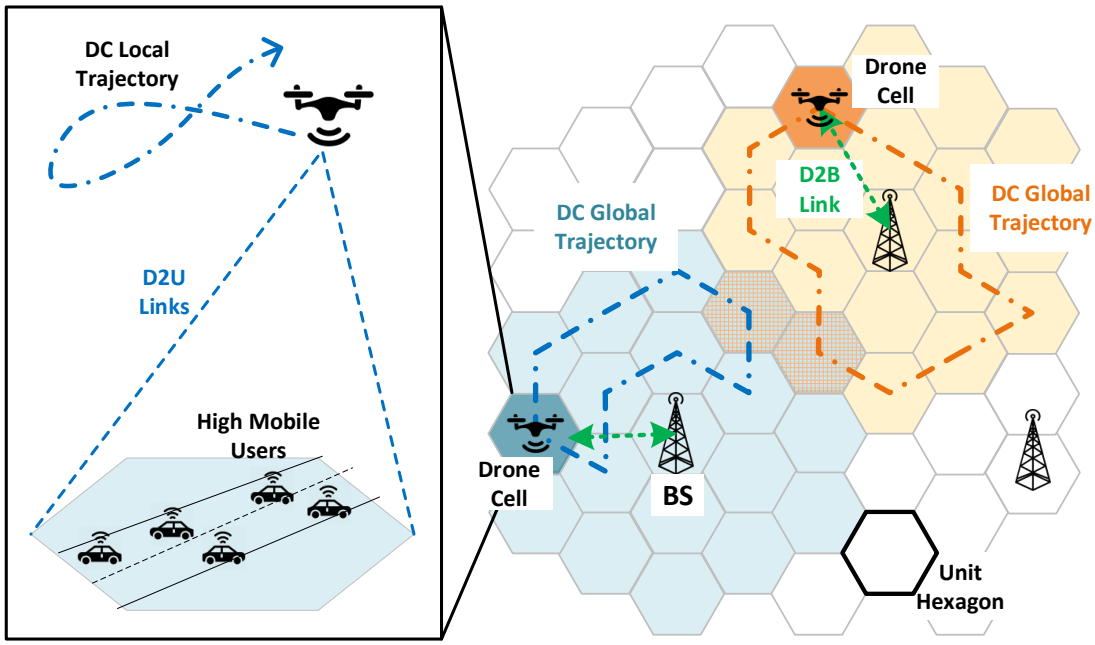


Figure 5.1: System model.

5.2.2 DC Communication Model

The state-of-the-art D2U and D2B channel models are used to indicate the high LoS probabilities of both D2U and D2B links. According to [45], the average D2U pathloss

(dB) is calculated by:

$$P_{du}^{L}\big(r_{du}(t),h_d(t)\big) = 20\log\Big(\frac{4\pi f_c}{c}\sqrt{h_d(t)^2+r_{du}(t)^2}\Big) + Pr_{\text{los}}\eta_{\text{los}} + \big(1-Pr_{\text{los}}\big)\eta_{\text{nlos}} \tag{5.1}$$

where $r_{du}(t)$ (m), $h_d(t)$ (m), f_c (Hz), and c (m/s) are D2U horizontal distance, DC flying height, carrier frequency, and speed of light, respectively. Denoted by Pr_{los}, the LoS probability is calculated by:

$$Pr_{\text{los}} = \frac{1}{1+a\exp(-b\arctan(\frac{h_d(t)}{r_{du}(t)})+ab)}. \tag{5.2}$$

In (5.1) and (5.2), a, b, η_{los}, and η_{nlos} are all environment-based constants.

The D2B channel is naturally modeled as LoS channel with environment-based offsets [24]. The D2B pathloss (dB) is calculated by:

$$P_{db}^{L}\big(r_{db}(t),h_d(t)\big) = 10\alpha\log\big(r_{db}(t)\big)+\eta_0 + A\Big(\arctan(\frac{h_d(t)}{r_{db}(t)})-\theta_0\Big)\mathrm{e}^{\big(\frac{\theta_0-\arctan(\frac{h_d(t)}{r_{db}(t)})}{B}\big)} \tag{5.3}$$

where $r_{db}(t)$, A, α, θ_0, η_0, and B are D2B horizontal distance, excess pathloss scalar, terrestrial pathloss exponent, angle offset, excess pathloss offset, and angle scalar, respectively. The 850 MHz LTE bands is used in (5.3) [24]. Both (5.1) and (5.3) are large-scale pathloss models. Since this work mainly investigates the impact of user mobility on the environment complexity, the small-scale channel shadowing and fading are not involved for analysis simplicity.

To prevent interference between DCs, we assume all DCs are assigned with orthogonal spectrum resources with the same total bandwidth B_{D}. Since the users are homogeneous in terms of their bandwidth requirement, to ensure user fairness, the B_{D} of each DC is evenly shared by all its associated users with a maximal per user bandwidth constraint b_{U}. Denote DC d's downlink transmit power to user u at step t by $P_{du}(t)$, the downlink throughput of user u at step t is calculated by:

$$c_{du}(t) = \begin{cases} b_{\text{U}}\log_2\big(1+\frac{P_{du}(t)\beta_{du}^{-1}(r_{du}(t),h_d(t))}{\sigma_0 b_{\text{U}}}\big), & b_{\text{U}} \le \frac{B_{\text{D}}}{|\mathcal{U}_d(t)|} \\ \frac{B_{\text{D}}}{|\mathcal{U}_d(t)|}\log_2\big(1+\frac{P_{du}(t)\beta_{du}^{-1}(r_{du}(t),h_d(t))}{\sigma_0 B_{\text{D}}/|\mathcal{U}_d(t)|}\big), & otherwise \end{cases} \tag{5.4}$$

where $\beta_{du}(r_{du}(t), h_d(t)) = 10^{P^L_{du}(r_{du}(t),h_d(t))/10}$ is the power attenuation coefficient, σ_0 is the spectral density of noise power.

Although the bandwidth are assumed to be evenly shared by users of the DC, the achieved throughput of users are still uneven due to different D2U 3D distances and different transmit power levels allocated by the TPRA strategies. We apply Jains fairness index [115] to measure the throughput fairness between all users served by one DC d at step t:

$$m_d(t) = \frac{\left(\sum_{u\in\mathcal{U}_d(t)} c_{du}(t)\right)^2}{|\mathcal{U}_d(t)| \sum_{u\in\mathcal{U}_d(t)} c_{du}(t)^2} \tag{5.5}$$

where $m_d(t) \in [1/n, 1]$. The fairness level between users increases as the fairness index $m_d(t)$ increases.

5.2.3 DC Energy Consumption Model

The limited total battery energy of each DC E_d is mainly consumed by three parts, i.e., the computation energy, data transmission energy and propulsion energy. The computation energy enables the signal processing and computation functions on DC, which is relatively much smaller than the communication and propulsion energy. Without loss of generality, we ignore the computation energy consumption in this work. Denote $P_d(t)$ as the total transmit power of DC d at step t:

$$P_d(t) = \sum_{u\in\mathcal{U}_d(t)} P_{du}(t). \tag{5.6}$$

The propulsion power energy is used to keep the DC aloft and adjust the movements. For a rotary-wing DC flying with speed $v_d(t)$ at time step t, the propulsion power consumption can be modeled as [116]:

$$P_d^{\text{prop}}(t) = P_{\text{b}}(1 + \frac{3v_d(t)^2}{{V_{\text{tip}}}^2}) + \frac{P_{\text{i}}V_{\text{h}}}{v_d(t)} + \frac{D_0 S_0 \rho A_0 v_d(t)^3}{2}. \tag{5.7}$$

where P_{b} and P_{i} are DC's blade profile power and induced power in hovering state, respectively, V_{tip} denotes the tip speed of the rotor blade, V_{h} is the mean rotor induced velocity in hovering state, D_0, S_0, ρ, and A_0 are the fuselage drag ratio, rotor solidity, air density, and rotor disc area, respectively.

Define the service endurance T_d as the DC's continuously flying time from fully charged state to energy depletion, we have:

$$\sum_{t=1}^{T_d} \big(P_d(t) + P_d^{\text{prop}}(t)\big)\delta_t \leq E_d. \tag{5.8}$$

For TPRA task conducted over long time period $T > T_d$, it is impossible for the DC to keep active with its limited battery capacity. In this chapter, we assume all DCs can fly back to the associated BSs to charge their batteries, with the same charging speed denoted by p_{crg} Joule per step t. The DCs in charging state cannot serve any users. In the following sections, we use $e_d(t)$ to denote the remaining battery energy of DC d at step t:

$$\begin{aligned} e_d(t) = E_d + \\ \sum_{\tau \leq t}^{t} \Big[p_{\text{crg}}\zeta(\tau) - \big(P_d(\tau) + P_d^{\text{prop}}(\tau)\big)\delta_t(1-\zeta(\tau))\Big] \end{aligned} \tag{5.9}$$

where $\zeta(\tau) = 1$ when DC d is charging at step τ, otherwise $\zeta(\tau) = 0$.

5.3 Problem Formulation

In this section, the multi-DC TPRA problem is formulated first, then decoupled into the higher-level GTP sub-problem and the lower-level LTPRA sub-problem according to the HDRL framework.

5.3.1 Multi-DC TPRA Problem

The objective of the multi-DC TPRA problem is maximizing accumulative network throughput over time period T, by choosing the appropriate TPRA decisions for each DC d at each step t to serve high-mobility users in DA-RAN scenario $\mathcal{G}$. Since the highly dynamic and uncertain user distributions over the scenario $\mathcal{G}$ can be assumed to evolve in an ergodic way, we can model the multi-DC TPRA problem as a MDP.

A typical MDP is denoted by a tuple $(\mathscr{S}, \mathcal{A}, R, P)$, in which $\mathscr{S}$ is the state space, $\mathcal{A}$ is action space, $R := \mathscr{S} \times \mathcal{A} \to \mathbb{R}$ is the reward function, and $P := \mathscr{S} \times \mathcal{A} \times \mathscr{S} \to \mathbb{R}$ is state transition probability. Denote the user distribution status in $\mathcal{G}$ at step t by $U(t)$, we define the system state at step t as $\mathbf{S}(t) = [U(t), \mathbf{Z}(t), \mathbf{E}(t)]$. $\mathbf{Z}(t)$ is the set of all

DCs' locations at step t, and $\mathbf{E}(t)$) is the set of all DCs' remaining energy at step t. We denote all DCs' trajectory planning decisions and resource allocation decisions at each step t by $\mathbf{A}_z(t)$ and $\mathbf{A}_r(t)$, respectively. The system action at step t can be represented by $\mathbf{A}(t) = (\mathbf{A}_z(t), \mathbf{A}_r(t))$. The state transitions from step t to $t+1$ are updated by three components. The first component is the user distribution status change $U(t) \to U(t+1)$, which depends on the highly dynamic and uncertain environment. The second component is all DCs' trajectories updates $\mathbf{Z}(t+1) = \mathbf{Z}(t) + \mathbf{A}_z(t)$ by taking action $\mathbf{A}_z(t)$. The third component is the remaining energy updates depend on both trajectory planning and resource allocation actions.

Define the time-invariant stationary policy mapping from any state $\mathbf{S} \in \mathscr{S}$ to any action $\mathbf{A} \in \mathscr{A}$ as $\mathbf{\Pi}(\mathbf{S}, \mathbf{A})$. We aim to find the optimal policy $\mathbf{\Pi}^*$ to maximize the long-term expectation of average network throughput. Therefore, we formulate the multi-DC TPRA problem, which can be regard as a CMDP, as follows:

$$\max_{\mathbf{\Pi}} \mathbb{E}\left[\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{d=1}^{D} \sum_{u}^{\mathcal{U}_d(t)} c_{du}\Big(\mathbf{S}(t)\Big) \middle| \mathbf{\Pi}\right] \tag{5.10}$$

$$s.t. \quad e_d(t) \geq 0, \qquad \forall d, t, \tag{5.10a}$$

$$m_d(t) = 1, \qquad \forall d, t, \tag{5.10b}$$

$$P_{\mathrm{DB}}^{L}\big(r_{db}(t), h_d(t)\big) \leq \gamma_{\mathrm{DB}}, \qquad \forall d, t, \tag{5.10c}$$

$$\mathcal{U}_d(t) \cap \mathcal{U}_{\bar{d}}(t) = \emptyset, \qquad \forall d \neq \bar{d}, t, \tag{5.10d}$$

$$v_d(t) < V_{\max}, \qquad \forall d, t. \tag{5.10e}$$

In (5.10), γ_{DB} denotes the maximal allowed D2B large-scale pathloss which ensures the D2B link quality. (5.10a) is the remaining energy constraint of each DC. (5.10b)-(5.10c) are user fairness constraint and D2B link quality constraint for each DC at each time step, respectively. (5.10d) ensures multiple DCs do not have coverage overlap to improve DC utilization efficiency. (5.10e) is the DCs' maximal flying speed constraint.

Considering massive high-mobility users over the large area $\mathcal{G}$, the size of state space $\mathcal{S}$ for $\mathbf{\Pi}(\mathbf{S}, \mathbf{A})$ is tremendous, which is infeasible for simple DRL based algorithm to solve. To address the complexity, the hierarchical DRL framework is leveraged which decouples the problem into multiple sub-problems with smaller state space $\mathscr{S} \subset \mathcal{S}$, then solves the whole problem by solving all sub-problems iteratively. In this chapter, we decouple the multi-DC TPRA problem into two hierarchical sub-problems, i.e., the multi-DC GTP sub-problem, and the single DC LTPRA sub-problem.

5.3.2 Multi-DC GTP Sub-problem

Note the fact that although the real-time user distribution varies dynamically over different time steps, the number of users within each unit $|\mathcal{U}_g(t)|$ changes little between adjacent steps, since most users' moving range within one step t cannot exceed the unit area. On the other hand, the $|\mathcal{U}_g(t)|$ varies smoothly over a long time period T (e.g. number of vehicles within one road segment at peak hours and normal hours over one day time). This long-term trends of different units must be learned by the trajectory planning algorithm to determine global trajectories for DCs.

In this work, we denote each DC d's global trajectory by $\mathbf{z}_d^{\mathrm{r}}(t_{\mathrm{r}})$. The global trajectory $\mathbf{z}_d^{\mathrm{r}}(t_{\mathrm{r}})$ determines the unit g at which the DC is designed to fly over and serve users at GTP step t_{r}. The length of each GTP step δ_{r} is longer than δ_t since the user number statistic in each unit varies slowly with time step. Define the average number of users in unit $g \in \mathcal{G}$ over GTP step t_{r} as $\overline{U_g}(t_{\mathrm{r}})$, the system state for multi-DC GTP sub-problem is represented by $\mathbf{s}^{\mathrm{r}}(t_{\mathrm{r}}) = [\overline{\mathcal{U}}_g(t_{\mathrm{r}}), \mathbf{z}^{\mathrm{r}}(t_{\mathrm{r}}), \mathbf{E}(t_{\mathrm{r}})]$. $\overline{\mathcal{U}}_g(t_{\mathrm{r}})$ is the set of all $\overline{U_g}(t_{\mathrm{r}})$, $\mathbf{z}^{\mathrm{r}}(t_{\mathrm{r}})$ is the set of all DCs' global trajectory locations (assigned units) at GTP step t_{r}, $\mathbf{E}(t_{\mathrm{r}})$ is the set of all DCs' remaining battery energy at GTP step t_{r}. Although the statistic data $\overline{U_g}(t_{\mathrm{r}})$ average out the random bursts of user traffic over short step t, the long-term uncertainties and dynamics of the user traffic are kept in $^{\mathrm{r}}$, which still requires the DRL based algorithm to address. In next time step $(t_{\mathrm{r}}+1)$, we define that the DC can only move to one of the six neighbor units or remain in its current unit. Therefore, the system action $\mathbf{a}^{\mathrm{r}}(t_{\mathrm{r}}) \in {}^{\mathrm{r}}$ is composed by D units selected from the seven potential units of each DC, respectively. The state transitions are updated by the mean user number variations $\overline{\mathcal{U}}_g(t_{\mathrm{r}}) \rightarrow \overline{\mathcal{U}}_g(t_{\mathrm{r}}+1)$, the DC location update $\mathbf{z}^{\mathrm{r}}(t_{\mathrm{r}}+1) = \mathbf{z}^{\mathrm{r}}(t_{\mathrm{r}}) + \mathbf{a}^{\mathrm{r}}(t_{\mathrm{r}})$, as well as the remaining energy update by reducing E_δ^{r} Joule at each t_{r}. The E_δ^{r} is calculated by $\frac{E_d \delta_{\mathrm{r}}}{T_d}$, which constrains the total energy consumption (consumed by local TPRA actions) of each DC within one GTP step period. Since we assume all users are homogeneous with a constant data traffic rate, the objective of (5.10), is equivalent to maximizing the long-term expectation of average number of served users. Denote the average number of users served by DC d at GTP step t_{r} by $\overline{U_d}(t_{\mathrm{r}})$, the multi-DC GTP sub-problem can be formulated as a CMDP:

$$\max_{\pi^{\mathrm{r}}} \mathbb{E}\left[\lim_{T\to\infty} \frac{1}{T} \sum_{t_{\mathrm{r}}=1}^{T} \sum_{d=1}^{D} \overline{U_d}(t_{\mathrm{r}}) \middle| \pi^{\mathrm{r}} \right] \tag{5.11}$$

$$s.t. \quad e_d^{\mathrm{r}}(t_{\mathrm{r}}) \geq 0, \qquad \forall d, t_{\mathrm{r}}, \tag{5.11a}$$

$$g(d, t_{\mathrm{r}}) \in \mathcal{G}_b, \qquad \forall d, t_{\mathrm{r}}, \tag{5.11b}$$

$$g(d, t_{\mathrm{r}}) \neq g(\bar{d}, t_{\mathrm{r}}), \qquad \forall g_k, d \neq \bar{d}, t_{\mathrm{r}}, \tag{5.11c}$$

where π^{r} is the stationary policy mapping from $\mathbf{s}^{\mathrm{r}} \in {}^{\mathrm{r}}$ to $\mathbf{a}^{\mathrm{r}} \in {}^{\mathrm{r}}$. $g(d, t_{\mathrm{r}})$ denotes the

unit g served by DC d at GTP step t_{r}. (5.11a) indicates the non-negative remaining battery energy. For $T \geq T_d$, this constraint can be guaranteed by charging the DC at its associated BS periodically. (5.11b) constrains the DC can only serve units within the communication of its associated BS. (5.11c) constrains each unit cannot be served by multiple DC simultaneously, which prevents the potential collisions and interference.

Note that the DC is able to adapt its antenna or beam directions to keep covering all users in $g_k(t_{\mathrm{r}})$ during each GTP step t_{r}. Given the fixed communication coverage (one unit) at each GTP step t_{r}, the optimal DC flying height $h_d^{\mathrm{opt}}(t_{\mathrm{r}})$ maximizing D2U throughput can be calculated according to [23]. Therefore, we assume that the optimal flying height of DC $h_d^{\mathrm{opt}}(t_{\mathrm{r}})$ within each GTP step t_{r} is known a priori, which simplifies the state and action space for the single DC LTPRA sub-problem.

5.3.3 Single DC LTPRA Sub-problem

Given the global trajectory planning result $g(d, t_{\mathrm{r}})$ at each GTP step t_{r}, each DC d independently solves the single DC TPRA sub-problem within its communication coverage $g(d, t_{\mathrm{r}})$ at each step t. Although the decision step is t to capture the real-time user mobility patterns, the complexity of state space ${}^{\mathrm{l}}_{d}$ and action space ${}^{\mathrm{l}}_{d}$ are both reduced due to the small serving area (one unit) and single DC decision (no joint action space of multiple DCs). For each DC d, the state at step t is represented by $\mathbf{s}_d^{\mathrm{l}}(t) = [\mathcal{U}_d(t), \mathbf{z}_d^{\mathrm{l}}(t), e_d^{\mathrm{l}}(t)]$, where $\mathcal{U}_d(t)$ is the status of served users, $\mathbf{z}_d^{\mathrm{l}}(t)$ is the DC location, $e_d^{\mathrm{l}}(t) \in [0, E_\delta^{\mathrm{r}}]$ is the remaining energy. The action at step t $\mathbf{a}_d^{\mathrm{l}}(t)$ is composed by trajectory planning action and resource allocation action. Therefore, the state transitions from t to $t+1$ are determined by the user distribution variations, trajectory planning actions and energy consumption by TPRA actions. With the objective of maximizing the expectation of average DC throughput, we formulate the single DC LTPRA sub-problem for each DC d as follows:

$$\max_{\pi_d^{\mathrm{l}}} \mathbb{E}\left[\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{u}^{\mathcal{U}_d(t)} c_{du}\Big(\mathbf{s}_d^{\mathrm{l}}(t), \mathbf{a}_d^{\mathrm{l}}(t)\Big) \middle| \pi_d^{\mathrm{l}} \right] \tag{5.12}$$

$$s.t. \quad (5.10a), (5.10b), (5.10c), (5.10e), \tag{5.12a}$$

where the LTPRA stationary policy π_d^{l} maps $\mathbf{s}_d^{\mathrm{l}} \in {}^{\mathrm{l}}_{d}$ to $\mathbf{a}_d^{\mathrm{l}} \in {}^{\mathrm{l}}_{d}$.

5.4 Hierarchical DRL based Trajectory Planning and Resource Allocation Scheme

In this section, we first propose the HDRLTPRA scheme to decouple the multi-DC TPRA problem in a hierarchical way. Then the MARL-GTP algorithm for higher-level sub-problem, and the DDPG-LTPRA algorithm for lower-level sub-problem are introduced, respectively.

5.4.1 HDRLTPRA Scheme

Considering the unknown transition probabilities, and the large state spaces due to highly dynamic and uncertain user mobility, DRL based algorithms are required to solve the two sub-problems. To integrate DRL algorithms for different sub-problems to solve the multi-DC TPRA problem, we apply the hierarchical DRL framework [117] and propose the HDRLTPRA scheme, as shown in Fig. 5.2.

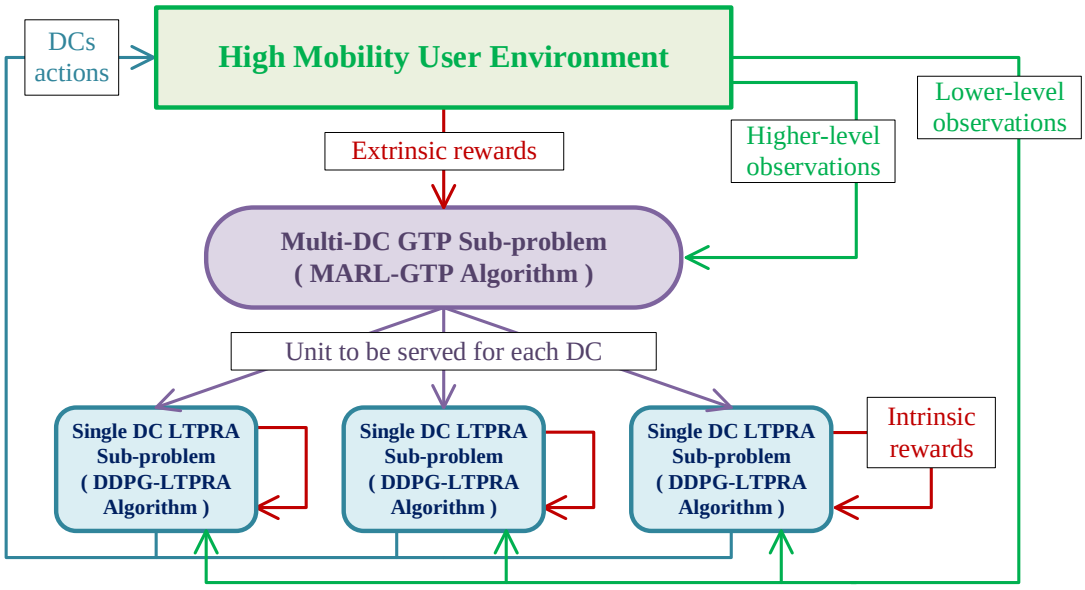


Figure 5.2: Architecture of HDRLTPRA scheme.

The proposed HDRLTPRA scheme has two levels. The MARL-GTP algorithm is executed on higher-level which addresses the multi-DC GTP sub-problem. The higher-level

agent interacts with the high-mobility environment in a low-frequency rate, and observes the statistic user data, i.e., average user number within each unit in GTP step t_{r}, as the input state. The extrinsic reward received by higher-level agent is the average number of users served by all DCs in current GTP step. The output actions of higher-level agent are units to be served by every DC in next GTP step.

The lower-level of HDRLTPRA is composed by multiple agents executing the DDPG-LTPRA algorithms to address the single DC LTPRA sub-problem. Each DC acts as one agent and executes the DDPG-LTPRA algorithm independently. The global trajectory planning result received from higher-level agent constrains the movement and serving range of the lower-level agent within one unit, which allows the DDPG-LTPRA algorithm to determine TPRA decisions in response to the real-time environment variations in the local unit. The intrinsic reward is obtained by each DC independently from interacting with the environment, which is defined as the total achieved throughput by the DC over the currently served unit.

The higher-level and lower-level agents can be trained in both interacting or separate ways given the online or off-line environment data, respectively. The training phase can only be stopped after all agents in higher and lower levels reach their convergence. In real implementation, the higher-level agent is implemented in the central controller of the DA-RAN to collect global data over the whole scenario and interact with multiple DCs; each lower-level agent is implemented on individual DC and receives higher-level global trajectory planning results via D2B communications.

5.4.2 MARL-GTP Algorithm

Given the discrete action spaces of the higher-level GTP sub-problem, the deep Q-network (DQN) can be used to solve the sub-problem with fast convergence. However, since the GTP sub-problem considers multiple DCs, the dimension (size) of the joint action space $|^{\mathrm{r}}| = 7^D$ increases exponentially as DC number D increases. To prevent the "curse of dimensionality" for action space in multi-DC environment, the multi-agent DRL is used in the MARL-GTP algorithm. MARL has been widely applied to solve the high-dimensional joint action space issue in existing works [118]. By implementing separate DRL algorithms in each agent, the global convergence is approximated through each agent's learning based on its own observation and the inter-communications between agents. In the MARL-GTP algorithm, we apply the fully cooperative MARL that all agents jointly maximize the global accumulative reward. Each DC is implemented with an identical DQN as the learning agent. Detail designs of the MARL-GTP algorithm are listed as follows.

Action Design

The joint action space of multiple DCs is decoupled into multiple local action spaces for different agents. The local action for each agent is $\mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}})$, which is an integer between $[0, 6]$. $\mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}}) = 0$ indicates the DC hovers above current unit in next step, otherwise the DC is assigned to the corresponding neighbor units. Since the flying speeds and transmit powers of DCs are not controlled by global trajectory planning, $\mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}})$ contains no power control actions. The $e_d^{\mathrm{r}}(t_{\mathrm{r}})$ is updated by directly reducing E_δ^{r} at each GTP step. For the charging process, we assume that the DC keeps charged during the whole charging GTP step and the step period δ_{r} is long enough for the DC to be fully charged.

State Design

According to Section 5.3.2, the set of average user numbers in all $g \in \mathcal{G}$ at GTP step t_{r}, $\overline{\mathcal{U}}_g(t_{\mathrm{r}})$, is used as the state component to represent user distribution changes. This design is realistic since the DA-RAN's central controller, in which the MARL-GTP algorithm is implemented, can collect the data over the whole scenario. For each DC associated with BS b, we use the subset $\overline{\mathcal{U}}_g^b(d, t_{\mathrm{r}}) \subset \overline{\mathcal{U}}_g(t_{\mathrm{r}})$, which is the set of average user numbers in all $g \in \mathcal{G}_b$, as the state component to reduce complexity due to constraint (5.11b). The ID of currently served unit $g(d, t_{\mathrm{r}})$ by the DC is involved in the state to indicate current DC location. The ID of unit where the associated BS located g_b and the remaining battery energy level $e_d^{\mathrm{r}}(t_{\mathrm{r}})$ are also involved in the state for DC charging behavior.

By involving all above designs, the state vector for one agent in MARL-GTP algorithm is represent as:

$$\mathbf{s}_d^{\mathrm{r}} = [\overline{\mathcal{U}}_g^b(d, t_{\mathrm{r}}), g(d, t_{\mathrm{r}}), g_b, e_d^{\mathrm{r}}(t_{\mathrm{r}})]. \tag{5.13}$$

In case where the dimension of $\overline{\mathcal{U}}_g^b(d, t_{\mathrm{r}})$ is too large, the "dimension spread" technique is applied to the last three elements in $\mathbf{s}_d^{\mathrm{r}}$ by duplicated them for multiple times, which balances the input weights of different factors.

Reward Design

At each step, the summation of $N_k(d, t_{\mathrm{r}})$ from all DCs is feedback to every DCs as step reward $W^{\mathrm{r}}(t_{\mathrm{r}})$. On the other hand, to promote the low-power DC flying back to BS for charging, a punishment reward is applied for low-power DC whose value is proportional to

the horizontal D2B distance. The general expression of step reward is denoted by:

$$W^{\mathrm{r}}(t_{\mathrm{r}}) = \begin{cases} \sum_{d=1}^{D} \overline{U_d}(t_{\mathrm{r}}), & \frac{e_d^{\mathrm{r}}(t_{\mathrm{r}})}{E_\delta^{\mathrm{r}}} > |r_{db}(t_{\mathrm{r}})| \\ \sum_{d=1}^{D} \overline{U_d}(t_{\mathrm{r}}) - w_{\mathrm{pun}}^{\mathrm{r}} |r_{db}(t_{\mathrm{r}})|, & \frac{e_d^{\mathrm{r}}(t_{\mathrm{r}})}{E_\delta^{\mathrm{r}}} \leq |r_{db}(t_{\mathrm{r}})| \\ W_{\mathrm{charge}}^{\mathrm{r}}, & e_d^{\mathrm{r}}(t_{\mathrm{r}}) = |r_{db}(t_{\mathrm{r}})| = 0 \\ -w_{\mathrm{pun}}^{\mathrm{r}} |r_{db}(t_{\mathrm{r}})|, & e_d^{\mathrm{r}}(t_{\mathrm{r}}) = 0, |r_{db}(t_{\mathrm{r}})| \neq 0 \end{cases} \tag{5.14}$$

where $-w_{\mathrm{pun}}^{\mathrm{r}}$ is the punishment reward constant, $|r_{db}(t_{\mathrm{r}})|$ is the normalized D2B horizontal distance equals to the minimal number of units to be traversed from the DC to the BS. $W_{\mathrm{charge}}^{\mathrm{r}}$ is the charging reward granted by reaching the BS for charging when $e_d^{\mathrm{r}}(t_{\mathrm{r}}) = 0$.

Priority Experience Replay

The punishment reward in (5.14) can be regarded as the "reward sharing" [119] technique to promote the agent leaning sparse behaviors with high reward (e.g. $W_{\mathrm{charge}}^{\mathrm{r}}$). To further increase the probability of those high-reward sparse behavior being learned, the priority experience replay (PER) technique [119] is applied into MARL-GTP algorithm. Instead of randomly choosing experienced state-actions pairs in the buffer, the PER stores the past experience in a priority tree, where the state-actions pairs with higher step rewards have higher probability to be chosen for learning.

Multi-agent Fingerprint

In MARL, the convergence is hard to be ensured since the movements of other agents break the environment of the each agent into non-stationary [118]. In [118], the non-stationary environment issue in MARL is solved by the multi-agent fingerprint technique, which involves the abstracted state information of other agents into the each agent's state design.

Apart from the step IDs of other agents suggested by [118], we also involve a neighbor DC indicator vector $\mathbf{o}_d = [o_1, o_2, \ldots, o_6]$ into state design. $\mathbf{o}_d$ contains six elements representing the neighboring DCs' status in the six neighbor units of the DC. Each element in $\mathbf{o}_d$ equals one when its represented unit is occupied by other DC in next time step, otherwise the element equals zero. Compared with directly listing all DCs' locations in the state,

the $\mathbf{o}_d$ limits the state non-stationary caused by other DCs' movements within six binary dimensions, which highly simplifies the state space complexity.

To use $\mathbf{o}_d$, an inter-agent information exchange process must be executed by all agents in each step, as shown in Fig. 5.3. In the controller, each agent sequentially takes global trajectory planning actions, then the remaining agents keep updating their vector $\mathbf{o}_d$ according to previous agents' actions, and use the latest one in their turns to determine the action. Note that the agent 1 in Fig. 5.3 has no constraint from other DC in its action chosen, while the last agent is constrained by all previous DCs. Such unfairness between state and action spaces of different agents can lead to the "lazy agent problem" [118], i.e., the agent with the most constraints becomes inactive to prevent affecting other agents performance. To overcome the "lazy agent problem", a simple round robin rule is applied to iteratively select the head agent for the inter-agent information exchange process at each step.

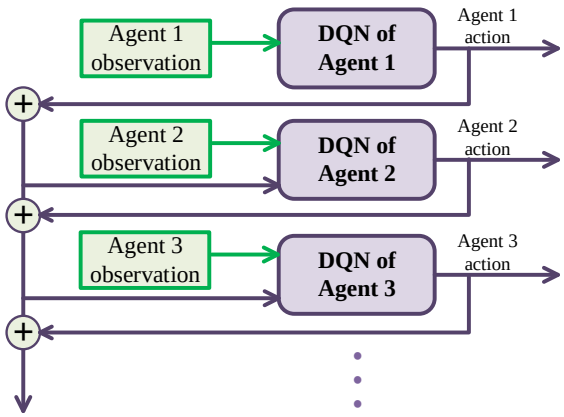


Figure 5.3: Inter-agent information exchange.

Leveraging the multi-agent fingerprint technique, we revise the state design for MARL-GTP as:

$$\mathbf{s}_d^{\mathrm{r}} = [\overline{\mathcal{U}}_g^b(d, t_{\mathrm{r}}), \mathbf{o}_d, g(d, t_{\mathrm{r}}), g_b, e_d^{\mathrm{r}}(t_{\mathrm{r}}), t_{\mathrm{r}}]. \tag{5.15}$$

In (5.15), we can directly use the currently step ID t_{r} as fingerprint because the MARL-GTP algorithm defines no break behaviors over the episode T, therefore, all DCs can complete each step simultaneously.

Algorithm 5 shows the details of MARL-GTP algorithm within one episode.

Algorithm 5 MARL-GTP Algorithm

Initialize the replay buffer $M^{\mathrm{r}}(d)$ with size $|M^{\mathrm{r}}(d)|$ for each DC.
Initialize the evaluation DQN $Q_d^e(\mathbf{s}_d^{\mathrm{r}}, \mathbf{a}_d^{\mathrm{r}}|\theta_d^e)$ with random parameters θ_d^e for each DC.
Initialize the target DQN $Q_d^t(\mathbf{s}_d^{\mathrm{r}}, \mathbf{a}_d^{\mathrm{r}}|\theta_d^t)$ with random parameters θ_d^t for each DC.
Initialize $\epsilon(d) = 1$ (fully random), $t_{\mathrm{r}} = 0$, and the ϵ_{decay}
for each episode **do**
    for each t_{r} **do**
        for $d \in \mathcal{D}$ **do**
            Update $\mathbf{o}_d$ with other DCs' actions.
            Choose $\mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}})$ from $\mathbf{s}_d^{\mathrm{r}}(t_{\mathrm{r}})$ using ϵ-greedy.
            if take $\mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}})$ violating (5.12a) **then**
                $\mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}}) = 0$.
            end if
            Take $\mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}})$, obtain reward $W^{\mathrm{r}}(t_{\mathrm{r}})$ according to (5.14), observe next state $\mathbf{s}_d^{\mathrm{r}}(t_{\mathrm{r}}+1)$.
            Store the transition $\big(\mathbf{s}_d^{\mathrm{r}}(t_{\mathrm{r}}), \mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}}), W_d^{\mathrm{r}}(t_{\mathrm{r}}), \mathbf{s}_d^{\mathrm{r}}(t_{\mathrm{r}}+1)\big)$ in $M^{\mathrm{r}}(d)$.
            Sample a training-batch from $M^{\mathrm{r}}(d)$.
            Calculate $\big(\mathbf{s}_d^{\mathrm{r}}(t_{\mathrm{r}}), \mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}})\big)$'s target Q value:
                $y_d^{\mathrm{r}}(t_{\mathrm{r}}) = W_d^{\mathrm{r}}(t_{\mathrm{r}}) + \gamma \max_{\mathbf{a}_d^{\mathrm{r}}} Q_d^t\big(\mathbf{s}_d^{\mathrm{r}}(t+1), \mathbf{a}_d^{\mathrm{r}}\big|\theta_d^t\big)$.
            Perform gradient decent on θ_d^e to minimize:
                $\big[y_d^{\mathrm{r}}(t_{\mathrm{r}}) - Q_d^e\big(\mathbf{s}_d^{\mathrm{r}}(t_{\mathrm{r}}), \mathbf{a}_d^{\mathrm{r}}(t_{\mathrm{r}})\big|\theta_d^e\big)\big]^2$
            if $\mod(t_{\mathrm{r}}, N) = 0$ **then**
                $\theta_d^t = \theta_d^e$.
                $\epsilon(d) = \epsilon(d) - \epsilon_{decay}$
            end if
        end for
        Round robin the order of DCs in inter-agent information exchange.
    end for
end for

5.4.3 DDPG-LTPRA Algorithm

Given the unit $g_k(d, t_\mathrm{r})$ decided by the higher-level MARL-GTP algorithm at GTP step t_r, We propose the DDPG-LTPRA algorithm for each DC to serve all users within $g_k(d, t_\mathrm{r})$ over one GTP step period δ_r. To promote the accuracy of local TPRA results, we allow the DC to choose TPRA actions from a continuous action space, which is enabled by the DDPG DRL technique. Based on the typical DDPG NN architecture, we specify the DDPG-LTPRA algorithm with the following designs.

Action Design

According to Section 5.3, the DC has two types of actions to increase the total achieved throughput, i.e. the transmit power control action, and the trajectory planning action. Considering the high mobility of users, the number of users within the served unit can change with each step t. Since the variable action and state dimensions are not supported by any DRL frameworks, it is hard and inefficient to directly allocate transmit power at per-user level by the DDPG-TPRA algorithm. Therefore, in this work, we define the amount of total transmit power $P_d(t)$ at each step t as the transmit power control action. Given the $P_d(t)$, we apply a simple fair power allocation mechanism for all served users in step t, which ensures strict per-user throughput fairness with $m_d(t) = 1$. Specifically, the power allocated to each user u is calculated by:

$$P_{du}(t) = \frac{P_d(t)\beta_{du}(t)}{\sum_{u \in \mathcal{U}_d(t)} \beta_{du}(t)}. \tag{5.16}$$

The choice of $P_d(t)$ is dependent on the current number of users being served, as well as the total available energy within one GTP step period δ_r.

Given the fixed flying height $h_d^{\mathrm{opt}}(t_\mathrm{r})$ defined in Section 5.3.2, the trajectory planning action of the DC is determining its horizontal flying speed $v_d(t)$. We represent $v_d(t)$ by two components $(v_d^x(t), v_d^y(t))$ along x and y-axis in the action space. Define the maximal speed of both components as $V_{\max}^{\mathrm{c}} = V_{\max}/\sqrt{2}$, the values of $v_d^x(t)$ and $v_d^y(t)$ are selected between $[-V_{\max}^{\mathrm{c}}, V_{\max}^{\mathrm{c}}]$. Therefore, the trajectory action space of a DC is defined as a horizontal 2D square centered at DC's current location with diagonal length $2V_{\max}$. In general, the output action of the DDPG-LTPRA algorithm is represent as:

$$\mathbf{a}_d^{\mathrm{l}}(t) = [v_d^x(t), v_d^y(t), P_d(t)]. \tag{5.17}$$

Algorithm 6 DDPG-TPRA Algorithm

Initialize the replay buffer M^{l} with size $|M^{\mathrm{l}}|$.
Initialize the actor network $\Phi_a(\mathbf{s}_d^{\mathrm{l}}|\varphi_a)$ with random parameters φ_a, and target actor network $\Phi_a^t(\mathbf{s}_d^{\mathrm{l}}|\varphi_a^t)$ with parameters $\varphi_a^t = \varphi_a$.
Initialize the critic network $\Phi_c(\mathbf{s}_d^{\mathrm{l}}, \mathbf{a}_d^{\mathrm{l}}|\varphi_c)$ with random parameters φ_c, and target critic network $\Phi_c^t(\mathbf{s}_d^{\mathrm{l}}, \mathbf{a}_d^{\mathrm{l}}|\varphi_c^t)$ with parameters $\varphi_c^t = \varphi_c$.
Initialize soft update rate τ.
for each episode **do**
    for each step t **do**
        Observe current state $\mathbf{s}_d^{\mathrm{l}}(t)$.
        Choose $\mathbf{a}_d^{\mathrm{l}}(t) = \Phi_a(\mathbf{s}_d^{\mathrm{l}}(t)|\varphi_a) + \mathcal{N}$ where $\mathcal{N}$ is exploration factor.
        if take $\mathbf{a}_d^{\mathrm{l}}(t)$ violating (5.11b), (5.11c) **then**
            Refine the action $\mathbf{a}_d^{\mathrm{l}}(t)$ within the constraints.
        end if
        Take $\mathbf{a}_d^{\mathrm{l}}(t)$, obtain reward $W^{\mathrm{l}}(t)$ according to (5.22), observe next state $\mathbf{s}_d^{\mathrm{l}}(t+1)$.
        Store the transition $\big(\mathbf{s}_d^{\mathrm{l}}(t), \mathbf{a}_d^{\mathrm{l}}(t), W^{\mathrm{l}}(t), \mathbf{s}_d^{\mathrm{l}}(t+1)\big)$ in M^{l}.
    end for
    Sample a training-batch with N_{bat} transitions from M^{l}.
    Calculate the critic target by:
    $y_t(t) = W^{\mathrm{l}}(t) - \gamma\Phi_c^t\big(\mathbf{s}_d^{\mathrm{l}}(t+1), \Phi_a^t(\mathbf{s}_d^{\mathrm{l}}(t+1)|\varphi_a^t)\big|\varphi_c^t\big)$
    Update critic network parameter φ_c by minimizing critic loss function:
    $L(\Phi_c) = \frac{1}{N_{\mathrm{bat}}}\sum_t \big[y_t(t) - \varphi_c\big(\mathbf{s}_d^{\mathrm{l}}(t), \mathbf{a}_d^{\mathrm{l}}(t)\big|\varphi_c\big)\big]^2$.
    Update actor network parameter φ_a by minimizing actor loss function:
    $L(\Phi_a) = -\frac{1}{N_{\mathrm{bat}}}\sum_t \varphi_c\big(\mathbf{s}_d^{\mathrm{l}}(t), \Phi_a(\mathbf{s}_d^{\mathrm{l}}(t)|\varphi_a)\big|\varphi_c\big)$.
    Soft update target networks:
    $\varphi_c^t = \tau\varphi_c + (1-\tau)\varphi_c^t$;
    $\varphi_a^t = \tau\varphi_a + (1-\tau)\varphi_a^t$
end for

State Design

Considering the variant numbers of users at each step, the state $\mathbf{s}_d^{\mathrm{l}}(t)$ has to be designed with fixed number of dimensions instead of using all users' locations. To find a properer state which indicates the feature of user distribution with fixed size, the geometric center $c_g(t)$ of all users in the served unit g at step t are selected. We choose $c_g(t)$ as the state component to represent user distribution feature due to the following corollary:

Corollary 2. *Given $P_d(t)$ and the per-user power allocation mechanism in (5.16), the lower bound of DC d's achieved total throughput at step t can be maximized by hovering above the geometric center $c_g(t)$ of all users in $\mathcal{U}_d(t)$.*

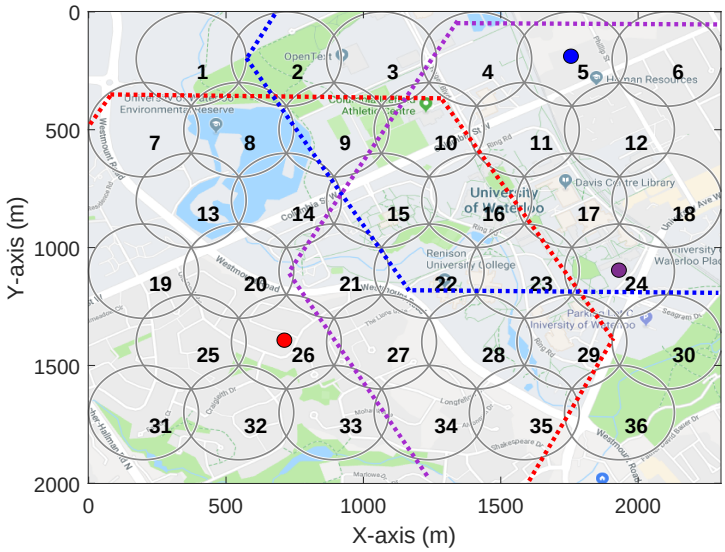


Figure 5.4: Real-world based simulation scenario.

Detail proof of the Corollary 2 is shown as follows.

Proof. According to the power allocation mechanism (5.16) for each user $u \in \mathcal{U}_d(t)$, the achieved total throughput for DC d is:

$$\begin{aligned} C_d(t) &= \sum_{u}^{\mathcal{U}_d(t)} b_{\mathrm{U}} \log_2 \left[1 + \frac{P_d(t)\beta_{du}(t)}{\sum_{u}^{\mathcal{U}_d(t)} \beta_{du}(t)} \times \frac{\beta_{du}^{-1}(t)}{\sigma_0 b_{\mathrm{U}}}\right] \\ &= \sum_{u}^{\mathcal{U}_d(t)} b_{\mathrm{U}} \log_2 \left[1 + \frac{P_d(t)}{\sigma_0 b_{\mathrm{U}} \sum_{u \in \mathcal{U}_d(t)} \beta_{du}(t)}\right]. \end{aligned} \tag{5.18}$$

For each step t, given $P_d(t)$, $\sigma_0 b_{\mathrm{U}}$, b_{U}, $h_d(t) = h_d^{\mathrm{opt}}(t)$ as constants, the maximal $C_d(t)$ is achieved by minimize $\sum_{u\in\mathcal{U}_d(t)} \beta_{du}(t)$, where

$$\begin{aligned}\beta_{du}(t) &= 10^{\frac{P^L_{du}(r_{du}(t),h_d(t))}{10}} \\ &= \big(\frac{4\pi f_c}{c}\big)^2 (h_d(t)^2 + r_{du}(t)^2) \times 10^{\frac{Pr_{\mathrm{los}}\eta_{\mathrm{los}} + \big(1-Pr_{\mathrm{los}}\big)\eta_{\mathrm{nlos}}}{10}}.\end{aligned} \tag{5.19}$$

Combine with (5.2), it is easy to prove that the component $\beta_{\mathrm{los}} = 10^{\frac{Pr_{\mathrm{los}}\eta_{\mathrm{los}} + \big(1-Pr_{\mathrm{los}}\big)\eta_{\mathrm{nlos}}}{10}}$ is a monotonic increasing function of D2U horizontal distance $r_{du}(t)$. Given the system model, the maximal D2U horizontal distance can be the diameter of one unit's circumscribed circle $2 \times R_{\mathrm{d}}$, which maximizes the $\max(\beta_{\mathrm{los}}) = \beta_{\mathrm{los}}^{\max}$. By applying the maximal β_{los} to all users' pathloss calculation, the upper bound of all users' pathloss summation at step t is calculated by:

$$\begin{aligned}UB\big(\sum_u^{\mathcal{U}_d(t)} \beta_{du}(t)\big) &= C \sum_u^{\mathcal{U}_d(t)} (h_d^{\mathrm{opt}}(t)^2 + r_{du}(t)^2) \\ &= C \sum_u^{\mathcal{U}_d(t)} r_{du}(t)^2 + C|\mathcal{U}_d(t)| h_d^{\mathrm{opt}}(t)^2\end{aligned} \tag{5.20}$$

where $C = \beta_{\mathrm{los}}^{\max}\big(\frac{16\pi^2 f_c^2}{c^2}\big)$. Since the second component in (5.20) $C|\mathcal{U}_d(t)| h_d^{\mathrm{opt}}(t)^2$ is constant, the minimal upper bound can be calculated by minimize $\sum_u^{\mathcal{U}_d(t)} r_{du}(t)^2$, which is achieved by the geometric center of all users. Note that the achieved total throughput $C_d(t)$ is lower bounded by using the upper bound defined in (5.20), therefore, we have proven that the lower bound of achieved total throughput $C_d(t)$ of DC d at step t can be maximized when the DC hovering above $c_g(t)$. □

According to Corollary 2, given different user distributions, flying towards $c_g(t)$ can always be the optimal action to maximize the minimal guaranteed total throughput. Since the purpose of the DDPG-TPRA algorithm is to learn the optimal deterministic action for each state, using $c_g(t)$ in state design can not only simplify the state space, but also provide guidance information for the DC's trajectory planning at each step. Note that the maximal total throughput at each step t might not be achieved by hovering above $c_g(t)$, $c_g(t)$ only provides a search direction with high probability to find the optimal point nearby. The exact optimal point maximizing total throughput is found by the DDPG-LTPRA algorithm. Together with the DC location component $z_d(t)$, as well as the remaining energy of DC within δ_{r}, the state for the DDPG-TPRA algorithm can be represented by:

$$\mathbf{s}_d^{\mathrm{l}}(t) = [z_d(t), c_g(t), e_d^{\mathrm{l}}(t)] \tag{5.21}$$

where $e_d^{\mathrm{l}}(t) \in [0, E_\delta^{\mathrm{r}}]$.

Reward Design

The total throughput achieved by the DC is set as step reward for DDPG-TPRA algorithm, together with a negative reward $-w_{\mathrm{pun}}^{\mathrm{l}}$ for states with $E_d^{\mathrm{r}}(t) = 0$:

$$W^{\mathrm{l}}(t) = \begin{cases} \sum_{u \in \mathcal{U}_d(t)} c_{du}(t), & e_d^{\mathrm{l}}(t) \neq 0 \\ -w_{\mathrm{pun}}^{\mathrm{l}}, & e_d^{\mathrm{l}}(t) = 0. \end{cases} \tag{5.22}$$

Details of the DDPG-TPRA algorithm for one DC are shown in Algorithm 6.

5.5 Simulations

To validate the performance of the proposed HDRLTPRA scheme, we build a real-world based simulation scenario as shown in Fig. 5.4. The scenario contains all roads in the campus region of University of Waterloo with a size of 2300m × 2000m. Three BSs are included in the scenario, which are represented by solid-circles with their communication coverage denoted by dotted lines in Fig. 5.4. The scenario is divided into 36 units, each hexagon unit is represented by its circumscribed circle. In the simulation, vehicles are considered as the high-mobility users. We build the scenario in traffic simulator VISSIM to generate highly authentic vehicle traffic by jointly considering the impacts of traffic signals, driver behaviors, traffic conditions, etc [120]. The total TPRA task period T is set to six hours from 9:00-am to 3:00-pm. Without loss of generality, we set the GTP step length $\delta_{\mathrm{r}} = 15$ minutes, and one step t length $\delta_t = 10$ seconds. The service endurance for each DC T_d is set to three hours, which is reasonable according to the existing industry products with 5+ endurance time[30]. We compare the performance of our HDRLTPRA scheme with the non-learning-based baseline scheme in [52]. Detail simulation parameters are shown in Table 5.1 and Table 5.2. All simulations are running on the server with Intel Xeon Gold 6128, 3.4GHz, 4 processors CPUs and 128G RAM.

In Table 5.2, the hyper-parameters of NNs are tuned by experiments. All those parameters can jointly impacts the NN outputs and convergence speeds. The NN output results are sensitive to the architecture parameters of NNs, such as the number of layers, nodes and batch size; while the convergence speed are more sensitive to the learning rate and decay speed for reward and ϵ (randomness level). One interesting observation is that the

Table 5.1: Simulation Parameters for HDRLTPRA Scheme

Parameters	Numerical Values
BS radio coverage radius R_{b}	1200 m
DC radio coverage radius R_{d}	200 m
D2U parameters $(\eta_{\mathrm{LoS}}, \eta_{\mathrm{NLoS}}, a, b)$	(0.1,21,4.88,0.43)
D2B parameters $(\alpha, A, \theta_0, B, \eta_0)$	(3.04,-23.29,-3.61,4.14,20.7)
Carrier frequencies (D2U, D2B)	$(2.4\ \mathrm{GHz}, 850\ \mathrm{MHz})$
D2B pathloss constraint γ_{DB}	80 dB
DC maximal speed $V_{\max}$	10 m/s
P_d^{prop} parameter $(P_{\mathrm{b}}, P_{\mathrm{i}}, V_{\mathrm{tip}}, V_{\mathrm{h}}$ $, D_0, S_0, \rho, A_0)$	$(577, 793, 200, 7.2$ $, 0.3, 0.05, 1.225, 0.785)$
Full battery Energy E_d	300 Wh
Charging speed p_{crg}	1200 W/s
Noise spectral density σ_0	-174 dBm/Hz
Total bandwidth B_{D}	20 MHz
Maximal per user bandwidth constraint b_{U}	1 MHz

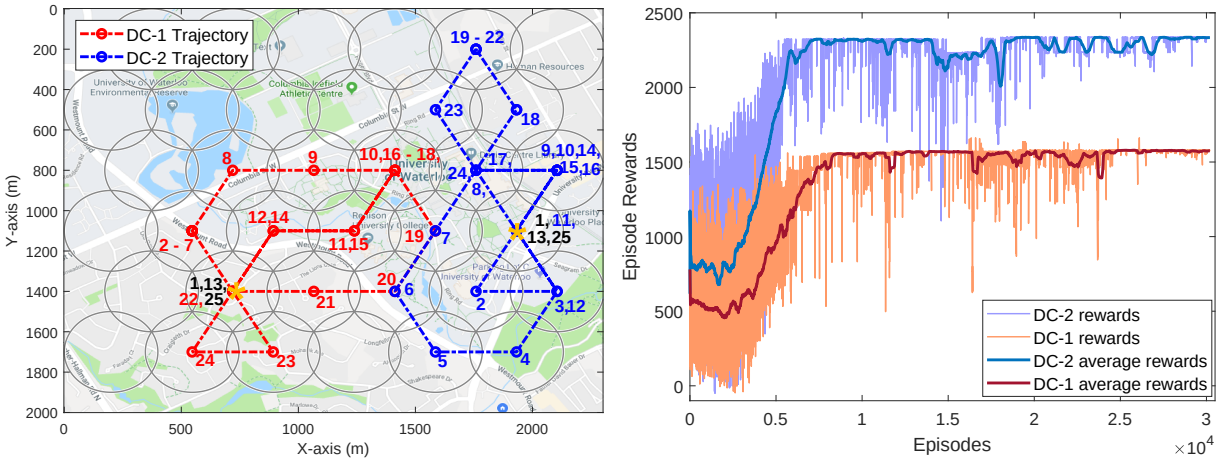
form of active function impacts both output results and convergence performance significantly. There for the most essential task for parameter tunning of HDRLTPRA scheme is active function selection.

Figs. 5.5 and 5.6 show the trajectory planning results and performance of the MARL-GTP algorithm with two and three DCs, respectively. Each BS is associated by no more than one DC in both scenarios. Figs. 5.5(a) and 5.6(a) plot the global trajectory planning results of the last three episodes out of 30000 episodes for both scenarios, in which the global trajectories of different DCs are denoted by dash-lines with different colors. Each number x associated to the units traversed by the trajectories indicates that the DC is planned to serve the unit at the x-th GTP step. Since the HARL-GTP algorithm has reached convergence, the trajectory planning results of the last three episodes are identical in both scenarios. Note that the MARL-GTP algorithm tends to reduce overlapping between different DCs' trajectories. In both scenarios, all DCs are assigned to the associated BS units at steps 13 and 25 (marked with black color). Given the three hours DC service endurance, step 13 and 25 are exactly the steps at which the DC battery is used up. This effect indicates that the charging behavior is successfully learned by the MARL-GTP algorithm.

Figs. 5.5(b) and 5.6(b) show the convergence performance for training the MARL-

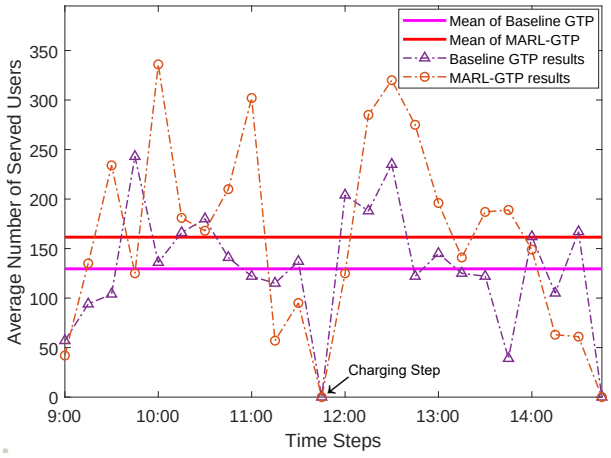
Table 5.2: Neural Network Parameters for Simulations

Parameters	Numerical Values
MARL-GTP Parameters	
Number of layers for DQN	3 (except output layer)
Number of nodes for each layer	$(128, 64, 64)$
Active function	Relu
Minimal ϵ, ϵ decay	$(0.001, 0.0002)$
Learning rate, reward decay γ	$(0.005, 0.9)$
Replay buffer size, batch size	$(20000, 128)$
Maximal episode, steps per episode	$(30000, 24)$
DDPG-LTPRA Parameters	
Number of layers for actor networks	3
Number of layers for critic networks	2
Number of nodes for actor layers	$(30, 30, 3)$
Number of nodes for critic layers	$(30, 1)$
Actor networks active function	(Relu, Relu, tanh)
Critic networks active function	(Relu, Relu)
Learning rates for actor, critic networks	$(0.001, 0.002)$
Reward decay γ	0.9
Soft update rate τ	0.01
Replay buffer size, batch size	$(100000, 512)$
Maximal episode, steps per episode	$(20000, 90)$



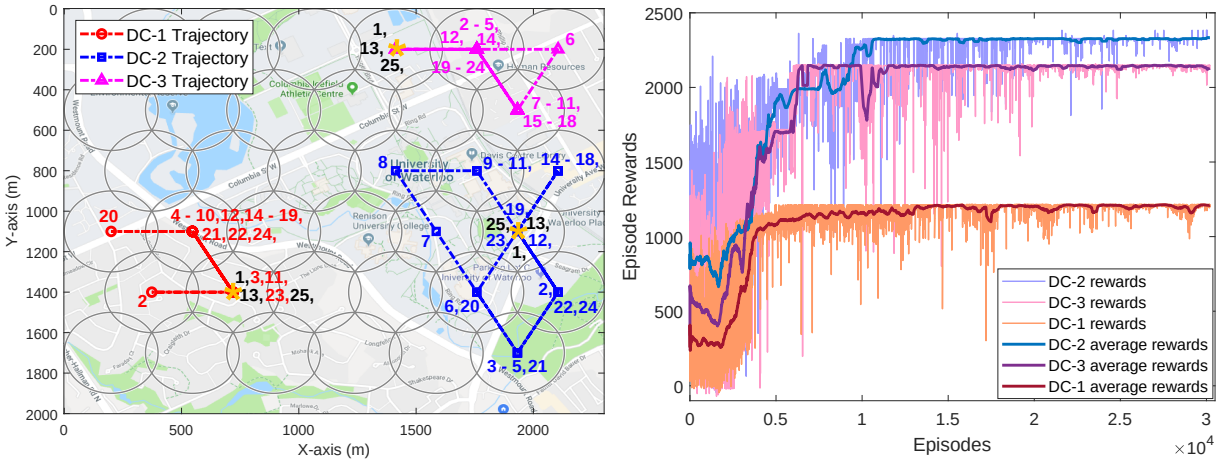
(a) Trajectory planning results

(b) Convergence performance



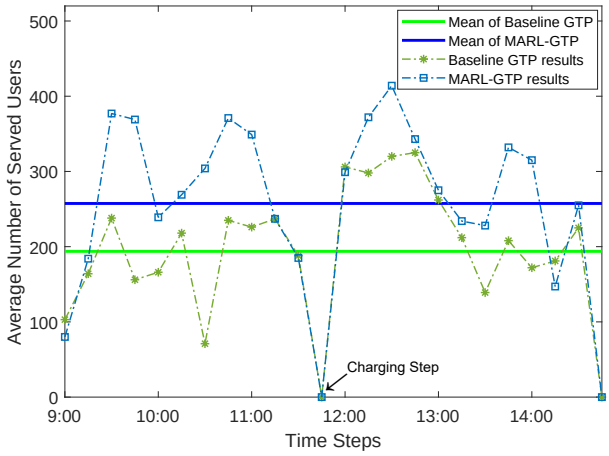
(c) User coverage performance

Figure 5.5: Global trajectory planning of 2 DCs by MARL-GTP algorithm.



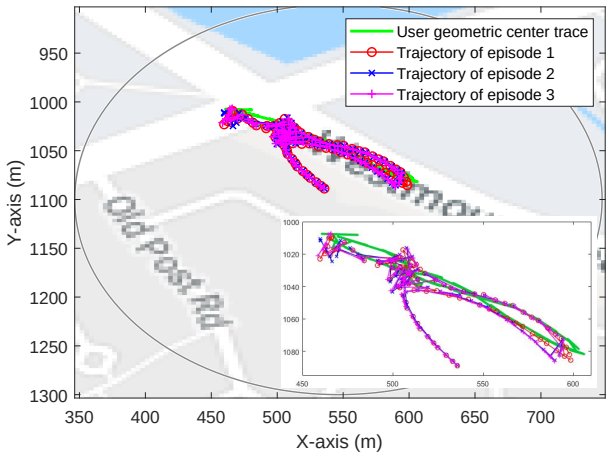
(a) Trajectory planning results

(b) Convergence performance

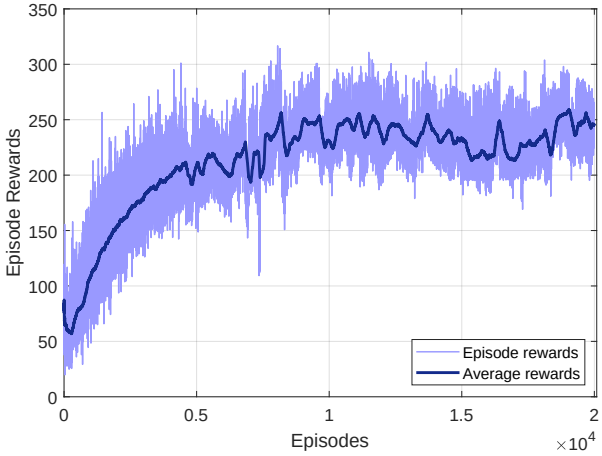


(c) User coverage performance

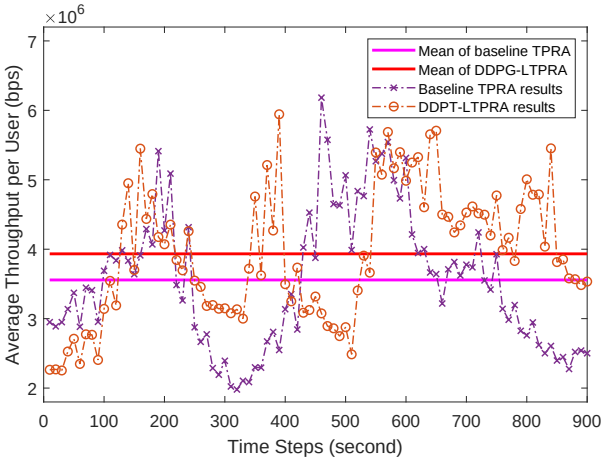
Figure 5.6: Global trajectory planning of 3 DCs by MARL-GTP algorithm.



(a) Trajectory Planning Results



(b) DDPG-LTPRA algorithm convergence performance



(c) Throughput performance

Figure 5.7: Local TPRA of single DC by DDPG-LTPRA algorithm.

GTP algorithm in both scenarios. We can observe that all DCs' episode rewards reach convergence after 25000 episodes training. In reality, the times of the algorithm to reach convergence in both scenarios are around 30 minutes. The performance comparisons of the MARL-GTP algorithm and the baseline GTP algorithm in terms of the average number of served users per GTP step are shown in Figs. 5.5(c) and 5.6(c), with their mean values denoted by horizontal lines. Note that the MARL-GTP algorithm can serve more users almost at all GTP steps, except for the few steps before the charging step. This effect is caused by the punishment reward design in the MARL-GTP algorithm for learning the charging behavior. In terms of the average number of served users over time period T, around $20\% - 25\%$ improvements are achieved by the MARL-GTP algorithm when compared with the baseline scheme.

Fig. 5.7 shows the TPRA results and performance of the DDPG-TPRA algorithm in unit 20. The trajectory of the last three episodes over 20000 episodes training is shown in Fig. 5.7(a). Note that the three trajectories are different in terms of each step's action choice, but all have the trend to follow the variations of vehicular users' geometric center, as shown in Fig. 5.7(a). This phenomenon indicates the DDPG-TPRA algorithm's capability to tracing the weight center of all users for maximizing total throughput. Fig. 5.7(b) shows the convergence performance of the DDPG-TPRA algorithm. We can observe that the mean value of the episode rewards approximates to 240 after 10000 episodes training, which takes around 20 minutes running on the server. However, each episode reward still varies between 200 to 280 after their mean value reaching convergence. This indicates that the DDPG-LTPRA algorithm can successfully approximate a sub-optimal long-term TPRA tread to maximize the accumulative throughput, but still requires longer training time to converge to an exact optimal trajectory. The achieved per-user average throughput at each step (within one episode) of the DDPG-TPRA and the baseline algorithms are shown in Fig. 5.7(c). Since all users in the unit are served by the DC at any steps, the trajectory designed by the non-learning-based baseline algorithm can converge to one point given the statistic vehicle traffic over the whole episode [52]. As shown in Fig. 5.7(c), the TPRA policy from the DDPG-TPRA algorithm achieves higher per-user average throughput than that of the baseline algorithm at most steps, and the DDPG-TPRA algorithm overhauls the baseline algorithm by 10% in terms of the achieved mean throughput over all steps.

By combining the MARL-GTP and DDPG-TPRA algorithm, the performance of the HDRLTPRA scheme in terms of total throughput at each steps over T is shown in Fig. 5.8. This simulation is conducted in the two DCs scenario. We can observe that the achieved total throughput is mainly dominated by the number of served users, which is maximized by the MARL-GTP algorithm. Compared with the non-learning-based solution, the HDRLTPRA scheme can achieve 40% performance promotion in terms of the

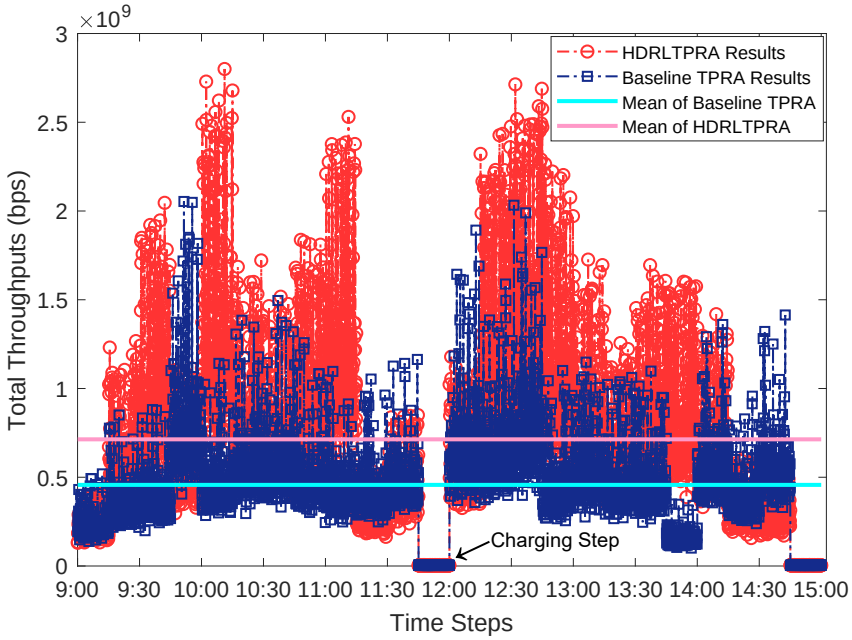


Figure 5.8: Accumulated throughput performance of HDRLTPRA scheme.

accumulative network throughput.

5.6 Summary

In this chapter, we have proposed the HDRLTPRA scheme for multiple DCs to serve high-mobility users over the large area for a long time. Based on the hierarchical DRL framework, we have decoupled the multi-DC TPRA problem into two hierarchical sub-problems to reduce the environment complexity. For the higher-level multi-DC GTP sub-problem, the MARL-GTP algorithm has been proposed, in which the MARL and multi-agent fingerprint techniques are applied to promote convergence in complex environment. Given the global trajectory planning results from the MARL-GTP algorithm, the DDPG-LTPRA algorithm for lower-level LTPRA sub-problem has been designed to adapt DC movement and allocate transmit power according to the real-time user distributions. Simulations show the HDRLTPRA scheme can enhance the total network throughput by 40% when compared with the model-based TPRA scheme.

Chapter 6

Conclusions and Future Works

In this chapter, we summarize the main contributions of this thesis, and discuss future research directions.

6.1 Main Research Contributions

In this thesis, we investigate the multi-DC 3D trajectory planning and resource allocation issues in DA-RAN for both quasi-static and high-mobility terrestrial users. In specific, three DC trajectory planning and resource allocation schemes are proposed for different DA-RAN scenarios: i.e., the DI-PSO scheme designs quasi-static 3D deployment of multiple DCs; the model-based multi-DC 3D trajectory planning and communication scheduling scheme determines stationary trajectories and communication schedules; as well as the DRL-based HDRLTPRA scheme generates dynamic TPRA decisions in response to the high uncertainties of high-mobility users. Features of D2G communications, DC energy consumptions, DC movements, and inter-DC interference are considered by the proposed schemes. The proposed schemes can provide guidance for RAN operators in terms of the implementation of DA-RAN and the design of their DC trajectory planning and resource allocation strategies. The main research contributions of this thesis are summarized as follows.

1. The general framework of DA-RAN was proposed, in which multiple DCs relay data between the associated BSs and their covered users in specific AoIs. Based on the state-of-the-art D2U and D2B channel pathloss models, the effective D2U communication coverage, D2B link quality, as well as the optimal flying height of the DC were

theoretically analyzed. Then, we formulated the multi-DC 3D deployment problem with the objective of maximizing the ratio of effectively covered users given a dedicated number of DCs. A DI-PSO algorithm was designed to find the sub-optimal DC deployment results corresponding to different numbers of DCs, which can achieve higher user coverage ratios with less iteration times when compared with the pure PSO based algorithm.

2. We investigated the 3D trajectory planning of multiple DCs in which both the flying heights and horizontal trajectories of DC are optimized together instead of optimizing horizontal trajectories on a 2D plane. Considering the state-of-the-art D2U and D2B pathloss models and the mutual interference of DCs, we formulated the multi-DC trajectory planning problem as an MINLP, and decoupled it into multiple sub-problems to resolve the non-convexity. Instead of modifying the 3D trajectories, we ensure the inter-DC interference constraint by scheduling the start slot of each trajectory to avoid introducing non-convex constraints in trajectory-related sub-problems. A BCD based algorithm was proposed to separately optimize AoI association, D2U communication scheduling, horizontal trajectories and flying heights of DCs in different sub-problems, respectively. Besides, a k-means-based scheme are devised to generate the DC initial trajectories for further improvements on performance. Compared with the quasi-static DC deployment algorithm based on PSO, the proposed 3D trajectory planning and scheduling algorithm can significantly improve the network performance and user fairness.

3. We proposed an effective multi-DC HDRLTPRA scheme for the scenario with high-mobility users. In HDRLTPRA, the higher-level MARL-GTP algorithm addresses the complexities caused by multiple DCs and long-term variations of user distributions, while the lower-level DDPG-LTPRA algorithm addresses the real-time variations of served user numbers and locations. This hierarchical DRL framework allows the HDRLTPRA to converge to sub-optimal TPRA solutions with high probability. To generate the global trajectories, we designed the MARL-GTP algorithm which implements fully cooperative multi-agent DRL to fit the multi-DC environment. In specific, the multi-agent fingerprints and PER methods are applied to design the hyper-parameters and NNs in MARL-GTP algorithm to address the non-stationary environment and sparse rewards. We further designed the DDPG-LTPRA algorithm executed by each DC independently to adjust the real-time DC movements and resource allocations over continuous spaces, which addresses the real-time user mobility. The complexity of DDPG-LTPRA's input were reduced by mathematical analyses of D2U communication to improve the convergence performance. The proposed HDRLTPRA scheme can improve the total network throughput by 40% when

compared with the model-based TPRA scheme.

6.2 Future Works

For the future research, there are some interesting related topics as follows.

1. **Joint allocation of multiple resource for learning-based TPRA of DCs:** Except the communication resources (transmission power, bandwidth) considered in this thesis, the computing and caching resources embedded by the DC can also be allocated based on the trajectory planning results, especially for the learning-based DC TPRA research. Existing learning-based TPRA works usually focus more on the trajectory planning policy and only consider one type of resources. Therefore, it is essential to investigate the joint allocation policies for multi-type resources in learning-based TPRA research. However, both the action and state spaces can be enlarged by involving more types of resources to be allocated, which further increases the environment complexity. To ensure the convergence performance, special designs of actions, states, and DRL frameworks must be applied in the DC TPRA schemes jointly allocating multi-type resources.

2. **Integrating with low-altitude satellite networks:** To serve massive terrestrial users without satellite communication terminals, the terrestrial-satellite access point (TSAP) is implemented to relay communications between terrestrial users and low earth orbit satellites (LEOs). Specifically, the TSAP can communicate with terrestrial users through various methods, such as C-band, WiFi, LTE-U and 5G-NR, etc.; the TSAP-to-satellite (T2S) communication is carried by Ka-band for its high throughput performance. The size and communication coverage of one TAST is comparable to a typical 5G small cell, which allows the TSAP to be implemented on the DC for temporal communication supports over demanding areas. Leveraging the DC TPRA research achievements, the dynamic deployment and trajectory design problem for mobile TSAPs can be a promising research directions.

   Since the mobile TSAP deployment problem integrates the DC TPRA and low-altitude satellite networks, some specific features must be considered in the research. First, compared with traditional backhaul links supported by wired networks, the T2S backhaul link for TSAP is a typical satellite-to-ground wireless link whose capacity and achieved data rate are limited with relatively large variations. Besides, since the effective serving time for one LEO to a dedicated terrestrial location is less than 10 minutes, the frequent hand-over and inter-LEO data exchanges can further

impact the T2S backhaul performance. To determine the dynamic deployment solution for multiple TSAPs, a set of comprehensive T2S backhaul constraints must be investigated with T2S channel models, inter-LEO hand-over and inter-LEO link qualities being considered. Second, due to the long distance of T2S wireless links and the limited data processing capability of LEOs, the capacity and delay performance for user-to-LEO communications relayed by TSAP is too constrained to support services with high-throughput or low-latency requirements. On the other hand, for some delay-tolerant services requiring long-distance (cross-continent) data transmission, the LEO network can provide better performance than terrestrial Internet due to less switching nodes. Therefore, when determining the serving areas for mobile TSAPs, we note only consider the spectrum scarcity, but also consider the user and traffic distributions of services suitable for LEO network. In this way, the mobile TSAPs are deployed to offload appropriate services traffic through LEO network, and release terrestrial spectrum resources for local or high-QoS required services. Third, leveraging the large coverage area of LEO network, and the direct user-to-satellite (U2S) link between terrestrial user and LEO for data collection, it is possible to conduct global optimization on LEO network for the deployment of all TSAPs within its coverage. Compared with traditional DC deployment over large area, the inter-BS or inter-DC coordination is substituted by LEO central control, which reduces the communication overhead and response speed for traffic variations.

References

[1] I. Bor-Yaliniz and H. Yanikomeroglu, "The new frontier in RAN heterogeneity: Multi-tier drone-cells," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 48–55, 2016.

[2] S. Zhang, W. Quan, J. Li, W. Shi, P. Yang, and X. Shen, "Air-ground integrated vehicular network slicing with content pushing and caching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2114–2127, 2018.

[3] M. Mozaffari, A. Taleb Zadeh Kasgari, W. Saad, M. Bennis, and M. Debbah, "Beyond 5g with UAVs: Foundations of a 3d wireless cellular network," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 357–372, 2019.

[4] X. Shen, J. Gao, W. Wu, K. Lyu, M. Li, W. Zhuang, X. Li, and J. Rao, "AI-assisted network-slicing based next-generation wireless networks," *IEEE Open J. Veh. Technol.*, vol. 1, no. 1, pp. 45–66, 2020.

[5] 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.

[6] L. Tractica, "Small unmanned aerial vehicles for aerial imaging, recreation, and aerial games: Global market analysis and forecasts," 2016. [Online]. Available: https://www.tractica.com/research/consumer-drones/.

[7] D. Wu, D. I. Arkhipov, M. Kim, C. L. Talcott, A. C. Regan, J. A. McCann, and N. Venkatasubramanian, "Addsen: Adaptive data processing and dissemination for drone swarms in urban sensing," *IEEE Trans. Comput.*, vol. 66, no. 2, pp. 183–198, 2017.

[8] W. Staff, "Fighting forest fires before they get bigwith drones," 2015. [Online]. Available: http://www.wired.com/2015/06/fighting-forest-fires-get-big-drones.

[9] R. L. Finn and D. Wright, "Unmanned aircraft systems: Surveillance, ethics and privacy in civil applications," *Computer Law & Security Review*, vol. 28, no. 2, pp. 184–194, 2012.

[10] Y. Dong, "Mobile high altitude base station, the first rise of the earthquake area, while providing services for 1200 mobile phones," 2017. [Online]. Available: http://www.fonow.com/view/191374.html.

[11] R. D'Andrea, "Guest editorial can drones deliver?" *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 3, pp. 647–648, 2014.

[12] DHL, "DHL parcelcopter 3.0," 2018. [Online]. Available: http://www.dpdhl.com/en/media_relations/specials/parcelcopter.html.

[13] N. H. Motlagh, T. Taleb, and O. Arouk, "Low-altitude unmanned aerial vehicles-based internet of things services: Comprehensive survey and future perspectives," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 899–922, 2016.

[14] W. Shi, H. Zhou, J. Li, W. Xu, N. Zhang, and X. Shen, "Drone assisted vehicular networks: architecture, challenges and opportunities," *IEEE Netw.*, vol. 32, no. 3, pp. 130–137, 2018.

[15] Google, "Project loon," 2018. [Online]. Available: https://x.company/projects/loon/.

[16] Facebook, "Connecting the world," 2018. [Online]. Available: https://info.internet.org/en/.

[17] Nokia, "UTM infrastructure and connected society," 2016. [Online]. Available: http://rpas-civops.com/wp-content/uploads/2016/11/S7.2_Nokia_DE_V1.pdf.

[18] Intel, "AT&T and intel showcase lte connected drones," 2016. [Online]. Available: https://www.intel.com/content/www/us/en/experience/drones/worlds-first-lte-connected-drone.html.

[19] Ericsson, "Ericsson and china mobile conduct worlds first 5G drone prototype field trial," 2016. [Online]. Available: https://www.ericsson.com/en/news/2016/8/ericsson-and-china-mobile-conduct-worlds-first-5g-drone-prototype-field-trial.

[20] Qualcomm, "Qualcomm technologies releases LTE drone trial results," 2017. [Online]. Available: https://www.qualcomm.com/news/onq/2017/05/03/qualcomm-technologies-releases-lte-drone-trial-results.

[21] Huawei, "Digital sky initiative," 2017. [Online]. Available: http://www.huawei.com/en/industry-insights/innovation/xlabs/industry-views/digital-sky-initiative.

[22] 3GPP TR 36.777, "Study on enhanced LTE support for aerial vehicles," V15.0.0, 2018.

[23] W. Shi, J. Li, W. Xu, H. Zhou, N. Zhang, S. Zhang, and X. Shen, "Multiple drone-cell deployment analyses and optimization in drone assisted radio access networks," *IEEE Access*, vol. 6, pp. 12 518–12 529, 2018.

[24] A. Al-Hourani and K. Gomez, "Modeling cellular-to-UAV path-loss for suburban environments," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 82–85, 2018.

[25] Y. Zhou, N. Cheng, N. Lu, and X. Shen, "Multi-UAV-aided networks: Aerial-ground cooperative vehicular networking architecture," *IEEE Veh. Technol. Mag.*, vol. 10, no. 4, pp. 36–44, 2015.

[26] Z. Xiao, P. Xia, and X.-G. Xia, "Enabling UAV cellular with millimeter-wave communication: Potentials and approaches," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 66–73, 2016.

[27] M. Alzenad, M. Z. Shakir, H. Yanikomeroglu, and M. Alouini, "FSO-based vertical backhaul/fronthaul framework for 5g+ wireless networks," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 218–224, 2018.

[28] F. Dong, H. Han, X. Gong, J. Wang, and H. Li, "A constellation design methodology based on qos and user demand in high-altitude platform broadband networks," *IEEE Trans. Multim.*, vol. 18, no. 12, pp. 2384–2397, 2016.

[29] DJI, "Mavic pro platinumspecs," 2018. [Online]. Available: https://www.dji.com/mavic-pro-platinum/info#specs.

[30] "Perimeter 8." [Online]. Available: https://skyfront.com/perimeter-8.

[31] MMC, "MMC hydrogen fuel cell drone now in its second generation: Introducing HyDrone 1800," 2017. [Online]. Available: http://www.mmcuav.com/2017/02/21/mmcs-hydrogen-fuel-cell-drone-now-in-its-second-generation-introducing-hydrone-1800/.

[32] J. Siminski, "Fukushima plants radiation levels monitored with an UAV," 2014. [Online]. Available: http://theaviationist.com/2014/01/29/fukushima-japan-uav.

[33] Q. Ye, J. Li, K. Qu, W. Zhuang, X. Shen, and X. Li, "End-to-end quality of service in 5G networks: Examining the effectiveness of a network slicing framework," *IEEE Veh. Technol. Mag.*, vol. 13, no. 2, pp. 65–74, 2018.

[34] N. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. Shen, "Software defined space-air-ground integrated vehicular networks: Challenges and solutions," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 101–109, 2017.

[35] F. Lv, H. Zhu, H. Xue, Y. Zhu, S. Chang, M. Dong, and M. Li, "An empirical study on urban ieee 802.11p vehicle-to-vehicle communication," in *Proc. IEEE SECON*, London, UK, Jun. 2016, pp. 1–9.

[36] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, 2016.

[37] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, 2016.

[38] W. Shi, J. Li, W. Xu, H. Zhou, N. Zhang, and X. Shen, "3D drone-cell deployment optimization for drone assisted radio access networks," in *Proc. IEEE/CIC ICCC*, Qingdao, China, Oct. 2017, pp. 1–6.

[39] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2109–2121, 2018.

[40] S. Chandrasekharan, K. Gomez, A. Al-Hourani, S. Kandeepan, T. Rasheed, L. Goratti, L. Reynaud, D. Grace, I. Bucaille, T. Wirth *et al.*, "Designing and implementing future aerial communication networks," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 26–34, 2016.

[41] E. Falletti, M. Laddomada, M. Mondin, and F. Sellone, "Integrated services from high-altitude platforms: a flexible communication system," *IEEE Commun. Mag.*, vol. 44, no. 2, pp. 85–94, 2006.

[42] W.-Q. Wang and D. Jiang, "Integrated wireless sensor systems via near-space and satellite platforms: A review," *IEEE Sensors J.*, vol. 14, no. 11, pp. 3903–3914, 2014.

[43] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1046–1061, 2017.

[44] A. Dhekne, M. Gowda, and R. R. Choudhury, "Cell tower extension through drones: Poster," in *Proc. ACM MobiCom*, New York, NY, Oct. 2016, pp. 456–457.

[45] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal lap altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, 2014.

[46] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, 2018.

[47] H. Zhou, N. Zhang, Y. Bi, Q. Yu, X. Shen, D. Shan, and F. Bai, "Tv white space enabled connected vehicle networks: Challenges and solutions," *IEEE Netw.*, vol. 31, no. 3, pp. 6–13, 2017.

[48] S. Zhang, G. Wang, and C. L. I, "Is mmwave ready for cellular deployment?" *IEEE Access*, vol. 5, pp. 14 369–14 379, 2017.

[49] N. Chen, M. Wang, N. Zhang, and X. Shen, "Energy and information management of electric vehicular network: A survey," *IEEE Commun. Surveys Tuts.*, to appear, 2020.

[50] C. Tseng, C. Chau, K. M. Elbassioni, and M. Khonji, "Autonomous recharging and flight mission planning for battery operated autonomous drones," 2017.

[51] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7574–7589, 2017.

[52] W. Shi, J. Li, N. Cheng, F. Lyu, S. Zhang, H. Zhou, and X. Shen, "Multi-drone 3D trajectory planning and scheduling in drone-assisted radio access networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8145–8158, 2019.

[53] N. Kato, Z. M. Fadlullah, F. Tang, B. Mao, S. Tani, A. Okamura, and J. Liu, "Optimizing space-air-ground integrated networks by artificial intelligence," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 140–147, 2019.

[54] 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.

[55] H. He, S. Zhang, Y. Zeng, and R. Zhang, "Joint altitude and beamwidth optimization for UAV-enabled multiuser communications," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 344–347, 2018.

[56] C. Zhang and W. Zhang, "Spectrum sharing for drone networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 136–144, 2017.

[57] P. Yang, X. Cao, C. Yin, Z. Xiao, X. Xi, and D. Wu, "Proactive drone-cell deployment: Overload relief for a cellular network under flash crowd traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2877–2892, 2017.

[58] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3D placement of an aerial base station in next generation cellular networks," in *Proc. IEEE ICC*, Kuala Lumpur, Malaysia, May. 2016, pp. 1–5.

[59] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, "3D 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, 2017.

[60] L. Zhou, Z. Yang, S. Zhou, and W. Zhang, "Coverage probability analysis of UAV cellular networks in urban environments," in *Proc. IEEE ICC Workshops*, Kansas City, MO, May. 2018, pp. 1–6.

[61] R. Ghanavi, E. Kalantari, M. Sabbaghian, H. Yanikomeroglu, and A. Yongacoglu, "Efficient 3d aerial base station placement considering users mobility by reinforcement learning," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, Barcelona, Spain, Apr. 2018, pp. 1–6.

[62] E. Kalantari, H. Yanikomeroglu, and A. Yongacoglu, "On the number and 3D placement of drone base stations in wireless cellular networks," in *Proc. IEEE VTC-Fall*, Montreal, QC, Sept. 2016, pp. 1–6.

[63] E. Kalantari, M. Z. Shakir, H. Yanikomeroglu, and A. Yongacoglu, "Backhaul-aware robust 3D drone placement in 5g+ wireless networks," in *Proc. IEEE ICC Workshops*, Paris, France, May 2017, pp. 109–114.

[64] S. A. W. Shah, T. Khattab, M. Z. Shakir, M. G. Khafagy, and M. O. Hasna, "Small cell association with networked flying platforms: Novel algorithms and performance bounds," 2018.

[65] S. A. W. Shah, T. Khattab, M. Z. Shakir, and M. O. Hasna, "A distributed approach for networked flying platform association with small cells in 5g+ networks," in *Proc. IEEE GLOBECOM*, Singapore, Singapore, Dec. 2017, pp. 1–7.

[66] A. Fouda, A. S. Ibrahim, I. Guvenc, and M. Ghosh, "UAV-based in-band integrated access and backhaul for 5G communications," in *Proc. IEEE VTC-Fall*, Chicago, IL, Apr. 2018, pp. 1–5.

[67] 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, 2016.

[68] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3949–3963, 2016.

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

[70] 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, 2016.

[71] J. Lyu, Y. Zeng, and R. Zhang, "Cyclical multiple access in UAV-aided communications: A throughput-delay tradeoff," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 600–603, 2016.

[72] Q. Wu and R. Zhang, "Common throughput maximization in UAV-enabled OFDMA systems with delay consideration," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6614–6627, 2018.

[73] G. Zhang, Q. Wu, M. Cui, and R. Zhang, "Securing UAV communications via joint trajectory and power control," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1376–1389, 2019.

[74] 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, 2018.

[75] H. Zhao, H. Wang, W. Wu, and J. Wei, "Deployment algorithms for uav airborne networks toward on-demand coverage," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2015–2031, 2018.

[76] Q. Zhang, M. Jiang, Z. Feng, W. Li, W. Zhang, and M. Pan, "IoT enabled UAV: Network architecture and routing algorithm," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3727–3742, 2019.

[77] M. Li, N. Cheng, J. Gao, Y. Wang, L. Zhao, and X. Shen, "Energy-efficient UAV-assisted mobile edge computing: Resource allocation and trajectory optimization," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3424–3438, 2020.

[78] H. Wang, H. Zhao, W. Wu, J. Xiong, D. Ma, and J. Wei, "Deployment algorithms of flying base stations: 5G and beyond with UAVs," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10 009–10 027, 2019.

[79] R. S. Sutton, A. G. Barto *et al.*, *Introduction to reinforcement learning*. MIT press Cambridge, 1998, vol. 135.

[80] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. AAAI*, 2016.

[81] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

[82] R. Lowe, Y. Wu, A. Tamar, J. Harb, O. P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *Advances in neural information processing systems*, 2017, pp. 6379–6390.

[83] M. Chen, W. Saad, and C. Yin, "Echo state learning for wireless virtual reality resource allocation in UAV-enabled LTE-U networks," in *Proc. IEEE ICC*, Kansas City, MO, May 2018, pp. 1–6.

[84] ——, "Liquid state machine learning for resource and cache management in LTE-U unmanned aerial vehicle (UAV) networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1504–1517, 2019.

[85] S. Yin, S. Zhao, Y. Zhao, and F. R. Yu, "Intelligent trajectory design in UAV-aided communications with reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8227–8231, 2019.

[86] U. Challita, W. Saad, and C. Bettstetter, "Cellular-connected UAVs over 5G: Deep reinforcement learning for interference management," *arXiv preprint arXiv:1801.05500*, 2018.

[87] X. Lu, L. Xiao, and C. Dai, "UAV-aided 5G communications with deep reinforcement learning against jamming," *arXiv preprint arXiv:1805.06628*, 2019.

[88] J. Hu, H. Zhang, and L. Song, "Reinforcement learning for decentralized trajectory design in cellular UAV networks with sense-and-send protocol," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6177–6189, 2019.

[89] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2059–2070, 2018.

[90] N. Cheng, F. Lyu, W. Quan, C. Zhou, H. He, W. Shi, and X. Shen, "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, 2019.

[91] X. Liu, Y. Liu, and Y. Chen, "Reinforcement learning in multiple-UAV networks: Deployment and movement design," *IEEE Trans. on Veh. Technol.*, vol. 68, no. 8, pp. 8036–8049, 2019.

[92] Z. M. Fadlullah, D. Takaishi, H. Nishiyama, N. Kato, and R. Miura, "A dynamic trajectory control algorithm for improving the communication throughput and delay in UAV-aided networks," *IEEE Netw.*, vol. 30, no. 1, pp. 100–105, 2016.

[93] N. Zhang, S. Zhang, J. Zheng, X. Fang, J. W. Mark, and X. Shen, "Qoe driven decentralized spectrum sharing in 5g networks: Potential game approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 7797–7808, 2017.

[94] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: Mm-wave beam steering without in-band measurement," in *Proc. IEEE INFOCOM*, Kowloon, Hong Kong, Apr. 2015, pp. 2416–2424.

[95] D. J. Deng, S. Y. Lien, J. Lee, and K. C. Chen, "On quality-of-service provisioning in ieee 802.11ax wlans," *IEEE Access*, vol. 4, pp. 6086–6104, 2016.

[96] Q. Ye and W. Zhuang, "Distributed and adaptive medium access control for internet-of-things-enabled mobile networks," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 446–460, 2017.

[97] W. Xu, H. A. Omar, W. Zhuang, and X. Shen, "Delay analysis of in-vehicle internet access via on-road wifi access points," *IEEE Access*, vol. 5, pp. 2736–2746, 2017.

[98] H. Ghazzai, E. Yaacoub, M. S. Alouini, Z. Dawy, and A. Abu-Dayya, "Optimized lte cell planning with varying spatial and temporal user densities," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1575–1589, 2016.

[99] M. Clerc and J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, no. 1, pp. 58–73, 2002.

[100] M. Garca-Fernndez, Y. lvarez Lpez, A. Arboleya, B. Gonzlez-Valds, Y. Rodrguez-Vaqueiro, M. E. D. C. Gmez, and F. L.-H. Andrs, "Antenna diagnostics and characterization using unmanned aerial vehicles," *IEEE Access*, vol. 5, pp. 23 563–23 575, 2017.

[101] M. Mozaffari, W. Saad, M. Bennis, Y. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2334–2360, 2019.

[102] F. Lyu, H. Zhu, H. Zhou, L. Qian, W. Xu, M. Li, and X. Shen, "Momac: Mobility-aware and collision-avoidance mac for safety applications in vanets," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10 590–10 602, 2018.

[103] Z. Xue, J. Wang, G. Ding, Q. Wu, Y. Lin, and T. A. Tsiftsis, "Device-to-device communications underlying UAV-supported social networking," *IEEE Access*, vol. 6, pp. 34 488–34 502, 2018.

[104] D. Takaishi, Y. Kawamoto, H. Nishiyama, N. Kato, F. Ono, and R. Miura, "Virtual cell based resource allocation for efficient frequency utilization in unmanned aircraft systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3495–3504, 2018.

[105] S. Jeong, O. Simeone, and J. Kang, "Mobile edge computing via a uav-mounted cloudlet: Optimization of bit allocation and path planning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2049–2063, 2018.

[106] N. Summers, "EE looks to drones and big balloons to tackle 4G 'notspots'," 2017. [Online]. Available: https://www.engadget.com/2017/02/21/ee-drone-balloon-truck-4g-coverage.

[107] Q. Li, S. Feng, X. Ge, G. Mao, and L. Hanzo, "On the performance of full-duplex multi-relay channels with df relays," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9550–9554, 2017.

[108] Q. Li, M. Yu, A. Pandharipande, X. Ge, J. Zhang, and J. Zhang, "Performance of virtual full-duplex relaying on cooperative multi-path relay channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3628–3642, 2016.

[109] "Gurobi optimizer 8.1." [Online]. Available: http://www.gurobi.com.

[110] D. P. Bertsekas, *Nonlinear programming*. Belmont, MA, USA: Athena scientific, 1999.

[111] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proc. ACM-SIAM*, Philadelphia, PA, Jan. 2007, pp. 1027–1035.

[112] N. Cheng, W. Xu, W. Shi, Y. Zhou, N. Lu, H. Zhou, and X. Shen, "Air-ground integrated mobile edge networks: Architecture, challenges, and opportunities," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 26–32, 2018.

[113] I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Spatial configuration of agile wireless networks with drone-bss and user-in-the-loop," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 753–768, 2019.

[114] T. Yu, X. Wang, and A. Shami, "UAV-enabled spatial data sampling in large-scale IoT systems using denoising autoencoder neural network," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1856–1865, 2019.

[115] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination," *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA*, 1984.

[116] 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.

[117] A. S. T. D. Kulkarni, K. Narasimhan and J. Tenenbaum, "Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation," in *Advances in Neural Information Processing Systems*, 2016, p. 36753683.

[118] J. Foerster, N. Nardelli, G. Farquhar, T. Afouras, P. H. Torr, P. Kohli, and S. Whiteson, "Stabilising experience replay for deep multi-agent reinforcement learning," in *Proc. ICML*, vol. 70, 2017, pp. 1146–1155.

[119] A. Y. Ng, D. Harada, and S. Russell, "Policy invariance under reward transformations: Theory and application to reward shaping," in *Proc. ICML*, vol. 99, 1999, pp. 278–287.

[120] N. Cheng, W. Quan, W. Shi, H. Wu, Q. Ye, H. Zhou, W. Zhuang, X. Shen, and B. Bai, "A comprehensive simulation platform for space-air-ground integrated network," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 178–185, 2020.

[121] J. Li, Y. Zhou, and L. Lamont, "Communication architectures and protocols for networking unmanned aerial vehicles," in *Proc. IEEE GLOBECOM Workshop*, Atlanta, GA, Dec. 2013, pp. 1415–1420.

[122] 3GPP, "Enhanced LTE support for aerial vehicles," 2017. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3231.

[123] M. Asadpour, B. V. den Bergh, D. Giustiniano, K. A. Hummel, S. Pollin, and B. Plattner, "Micro aerial vehicle networks: An experimental analysis of challenges and opportunities," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 141–149, 2014.

[124] H. Shokri-Ghadikolaei, C. Fischione, P. Popovski, and M. Zorzi, "Design aspects of short-range millimeter-wave networks: A mac layer perspective," *IEEE Netw.*, vol. 30, no. 3, pp. 88–96, 2016.

[125] S. Andreev, V. Petrov, M. Dohler, and H. Yanikomeroglu, "Future of ultra-dense networks beyond 5g: Harnessing heterogeneous moving cells," *arXiv preprint arXiv:1706.05197*, 2017.

[126] G. Ding, Q. Wu, L. Zhang, T. A. Tsiftsis, and Y.-D. Yao, "An amateur drone surveillance system based on cognitive internet of things," *IEEE Commun. Mag.*, to appear, 2018.

[127] S. Liu and M. St-Hilaire, "A genetic algorithm for the global planning problem of umts networks," in *Proc. IEEE GLOBECOM*, Miami, FL, Dec. 2010, pp. 1–5.

[128] A. Thornburg, T. Bai, and R. W. Heath Jr, "Performance analysis of outdoor mmwave ad hoc networks." *IEEE Trans. Signal Processing*, vol. 64, no. 15, pp. 4065–4079, 2016.

[129] L. Qualcomm, "Cellular drone communication," 2017. [Online]. Available: https://www.qualcomm.com/invention/technologies/lte/advanced-pro/cellular-drone-communication.

[130] C. Liang and F. R. Yu, "Wireless network virtualization: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 358–380, 2015.

[131] M. Richart, J. Baliosian, J. Serrat, and J.-L. Gorricho, "Resource slicing in virtual wireless networks: A survey," *IEEE Trans. Netw. Service Manag.*, vol. 13, no. 3, pp. 462–476, 2016.

[132] R. Kokku, R. Mahindra, H. Zhang, and S. Rangarajan, "Nvs: A substrate for virtualizing wireless resources in cellular networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 5, pp. 1333–1346, 2012.

[133] J. Sachs and S. Baucke, "Virtual radio: a framework for configurable radio networks," in *Proc. ICST*, 2008, p. 61.

[134] Y. Zaki, L. Zhao, C. Goerg, and A. Timm-Giel, "Lte mobile network virtualization," *Mobile Netw. Appl.*, vol. 16, no. 4, pp. 424–432, 2011.

[135] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," in *ACM SIGOPS*, vol. 37, no. 5, 2003, pp. 164–177.

[136] M. Li, F. R. Yu, P. Si, E. Sun, Y. Zhang, and H. Yao, "Random access and virtual resource allocation in software-defined cellular networks with machine-to-machine communications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6399–6414, 2017.

[137] ETSI, "Mobile-edge computing introductory technical white paper," 2014. [Online]. Available: https://portal.etsi.org/portals/0/tbpages/mec/docs/mobile-edge_computing_-_introductory_technical_white_paper_v1.

[138] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proc. MCC workshop on Mobile cloud computing*, 2012, pp. 13–16.

[139] K. Zhang, Y. Mao, S. Leng, S. Maharjan, and Y. Zhang, "Optimal delay constrained offloading for vehicular edge computing networks," in *Proc. IEEE ICC*, 2017, pp. 1–6.

[140] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 2017.

[141] O. Munoz, A. Pascual-Iserte, and J. Vidal, "Optimization of radio and computational resources for energy efficiency in latency-constrained application offloading," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4738–4755, 2015.

[142] W. Wang, B. Liang, and B. Li, "Multi-resource fair allocation in heterogeneous cloud computing systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 10, pp. 2822–2835, 2015.

[143] M. Chowdhury, Z. Liu, A. Ghodsi, and I. Stoica, "Hug: Multi-resource fairness for correlated and elastic demands." in *Proc. NSDI*, 2016, pp. 407–424.

[144] J. Khamse-Ashari, I. Lambadaris, G. Kesidis, B. Urgaonkar, and Y. Zhao, "Per-server dominant-share fairness (ps-dsf): A multi-resource fair allocation mechanism for heterogeneous servers," in *Proc. IEEE ICC*, Paris, France, May. 2017, pp. 1–7.

[145] M.-H. Chen, B. Liang, and M. Dong, "Joint offloading and resource allocation for computation and communication in mobile cloud with computing access point," in *Proc. IEEE INFOCOM*, Atlanta, GA, May. 2017, pp. 1–9.

[146] N. Lu, N. Zhang, N. Cheng, X. Shen, J. W. Mark, and F. Bai, "Vehicles meet infrastructure: Toward capacity-cost tradeoffs for vehicular access networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1266–1277, 2013.

[147] H. Wang, J. Wang, G. Ding, J. Chen, Y. Li, and Z. Han, "Spectrum sharing planning for full-duplex UAV relaying systems with underlaid d2d communications," *IEEE J. Sel. Areas Commun.*, to appear.

[148] V. V. Chetlur and H. S. Dhillon, "Downlink coverage analysis for a finite 3D wireless network of unmanned aerial vehicles," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4543–4558, 2017.

[149] Q. Ye, W. Zhuang, X. Li, and J. Rao, "End-to-end delay modeling for embedded vnf chains in 5G core networks," *IEEE Internet Things J.*, p. early access, 2018.

[150] I. Bisio, C. Garibotto, F. Lavagetto, A. Sciarrone, and S. Zappatore, "Unauthorized amateur UAV detection based on wifi statistical fingerprint analysis," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 106–111, 2018.

[151] X. Cao, P. Yang, M. Alzenad, X. Xi, D. Wu, and H. Yanikomeroglu, "Airborne communication networks: A survey," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1907–1926, 2018.

[152] S. Sra, S. Nowozin, and S. J. Wright, *Optimization for machine learning*. Mit Press, 2012.

[153] Z. Xue, J. Wang, G. Ding, H. Zhou, and Q. Wu, "Maximization of data dissemination in UAV-supported internet of things," *IEEE Wireless Commun. Lett.*, early access, 2018.

[154] M. E. Soussi, P. Zand, F. Pasveer, and G. Dolmans, "Evaluating the performance of eMTC and NB-IoT for smart city applications," in *Proc. IEEE ICC*, Kansas City, MO, May. 2018, pp. 1–7.

[155] S. Kumar, S. Suman, and S. De, "Backhaul and delay-aware placement of UAV-enabled base station," in *Proc. IEEE INFOCOM Workshop)*, Honolulu, HI, Apr. 2018, pp. 634–639.

[156] M. Tao, K. Ota, and M. Dong, "Locating compromised data sources in IoT-enabled smart cities: A great-alternative-region-based approach," *IEEE Trans Ind. Informat.*, vol. 14, no. 6, pp. 2579–2587, 2018.

[157] J. Li, N. Zhang, Q. Ye, W. Shi, W. Zhuang, and X. Shen, "Joint resource allocation and online virtual network embedding for 5g networks," in *Proc. IEEE GLOBECOM*, Singapore, Singapore, Dec. 2017, pp. 1–6.

[158] J. Liu and N. Kato, "A markovian analysis for explicit probabilistic stopping-based information propagation in postdisaster ad hoc mobile networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 81–90, 2016.

[159] M. T. Dabiri, S. M. S. Sadough, and M. A. Khalighi, "Channel modeling and parameter optimization for hovering UAV-based free-space optical links," *IEEE J. Sel. Areas Commun.*, early access, 2018.

[160] M. Moradi, K. Sundaresan, E. Chai, S. Rangarajan, and Z. M. Mao, "Skycore: Moving core to the edge for untethered and reliable UAV-based LTE networks," in *Proc. ACM MOBICOM*, New Delhi, India, Oct. 2018, pp. 35–49.

[161] W. Shi, J. Li, N. Cheng, F. Lyu, Y. Dai, H. Zhou, and X. Shen, "3d multi-drone-cell trajectory design for efficient iot data collection," in *Proc. IEEE ICC*, 2019. submitted.

[162] P. Yang, X. Cao, X. Xi, Z. Xiao, and D. Wu, "Three-dimensional drone-cell deployment for congestion mitigation in cellular networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9867–9881, 2018.

[163] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2059–2070, 2018.

[164] D. Grace, J. Thornton, , G. P. White, and T. C. Tozer, "Improving the system capacity of broadband services using multiple high-altitude platforms," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 700–709, 2005.

[165] G. Yang, X. Lin, Y. Li, H. Cui, M. Xu, D. Wu, H. Ryden, and S. B. Redhwan, "A telecom perspective on the internet of drones: From LTE-advanced to 5G," 2018. [Online]. Available: http://arxiv.org/abs/1803.11048.

[166] H. Wu, Z. Wei, Y. Hou, N. Zhang, and X. Tao, "Cell-edge user offloading via flying UAV in non-uniform heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, early access, 2020.

[167] M. Chen, W. Saad, and C. Yin, "Echo-liquid state deep learning for 360 content transmission and caching in wireless VR networks with cellular-connected UAVs," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6386–6400, 2019.

[168] C. Zhou, H. He, P. Yang, F. Lyu, W. Wu, N. Cheng, and X. Shen, "Deep RL-based trajectory planning for AoI minimization in UAV-assisted IoT," in *Proc. IEEE WCSP*, Xi'an, China, Oct. 2019, pp. 1–6.

[169] 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, 2019.

[170] 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.

List of Publications

Journal Papers

1. **W. Shi**, H. Zhou, J. Li, W. Xu, N. Zhang, and X. Shen, "Drone assisted vehicular networks: Architecture, challenges and opportunities", *IEEE Netw.*, vol. 32, no. 3, pp. 130-137, 2018.

2. **W. Shi**, J. Li, W. Xu, H. Zhou, N. Zhang, S. Zhang, and X. Shen, "Multiple drone-cell deployment analyses and optimization in drone assisted radio access networks", *IEEE Access*, vol. 6, pp. 12518-12529, 2018.

3. **W. Shi**, J. Li, N. Cheng, F. Lyu, S. Zhang, H. Zhou, and X. Shen, "Multi-drone 3D trajectory planning and scheduling in drone-assisted radio access networks", *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8145-8158, 2019.

4. **W. Shi**, J. Li, H. Wu, C. Zhou, N. Cheng, and X. Shen, "Drone-cell trajectory planning and resource allocation for highly mobile networks: A hierarchical DRL approach", *IEEE Internet Things J.*, under review.

5. **W. Shi**, J. Li, P. Yang, Q. Ye, X. Shen, X. Li, and J. Rao, "Two-level soft RAN slicing for customized service provisioning", under revision.

6. J. Li, **W. Shi**, P. Yang, Q. Ye, X. Shen, X. Li, and J. Rao, "A Hierarchical Soft RAN Slicing Framework for Differentiated Service Provisioning," *IEEE Wireless Commun.*, accepted, 2020.

7. J. Li, **W. Shi**, N. Zhang, and X. Shen, "Delay-Aware VNF Scheduling: A Reinforcement Learning Approach with Variable Action Set," *IEEE Trans. Cogn. Commun. Netw.*, accepted, 2020.

8. J. Li, **W. Shi**, N. Zhang, Q. Ye, S. Zhang, W. Zhuang, and X. Shen, "Joint Virtual Network Topology Design and Embedding for Service-Oriented 5G Core Networks," *IEEE Trans. Commun.*, under review, 2020.

9. N. Cheng, W. Quan, **W. Shi**, H. Wu, Q. Ye, H. Zhou, W. Zhuang, X. Shen, and B. Bai, "A comprehensive simulation platform for space-air-ground integrated network", *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 178185, 2020.

10. S. Zhang, B. Luo, J. Li, **W. Shi**, and X. Shen, "Hierarchical Soft Slicing to Meet Multi-Dimensional QoS Demand in Cache-Enabled Vehicular Networks", *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2150-2162, 2020.

11. O. Alhussein, P. T. Do, Q. Ye, J. Li, **W. Shi**, W. Zhuang, X. Shen, X. Li, and J. Rao, "A Virtual Network Customization Framework for Multicast Services in 5G Core Networks", *IEEE J. Sel. Areas Commun.*, Accepted, 2020.

12. N Cheng, F Lyu, W Quan, C Zhou, H He, **W. Shi**, and X. Shen, "Space/Aerial-Assisted Computing Offloading for IoT Applications: A Learning-Based Approach", *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117-1129, 2019.

13. N Cheng, W Xu, **W. Shi**, Y Zhou, N Lu, H Zhou, and X. Shen, "Air-ground Integrated Mobile Edge Networks: Architecture, Challenges, and Opportunities", *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 26-32, 2018.

14. S Zhang, J Chen, F Lyu, N Cheng, **W. Shi**, and X. Shen, "Vehicular Communication Networks in The Automated Driving Era", *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 26-32, 2018.

15. S. Zhang, W. Quan, J. Li, **W. Shi**, P. Yang, and X. Shen, "Air-Ground Integrated Vehicular Network Slicing with Content Pushing and Caching", *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2114-2127, Oct. 2018.

16. W Xu, H Zhou, N Cheng, F Lyu, **W. Shi**, J Chen, and X. Shen, "Internet of Vehicles in Big Data Era", *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 1, pp. 19-35, 2017.

17. J. Li, **W. Shi**, N. Zhang, Q. Ye, and X. Shen, "Multi-SFC embedding with end-to-end delay-guarantee: A game theoretical approach", manuscript.

18. J. Li, **W. Shi**, P. Yang, and X. Shen, "Cost-aware dynamic service function chain provisioning in SDN-NFV enabled networks", manuscript.

Conference Papers

1. **W. Shi**, J. Li, W. Xu, H. Zhou, N. Zhang, and X. Shen, "3D drone-cell deployment optimization for drone assisted radio access networks", *Proc. IEEE/CIC ICCC'17*, Oct. 2017, pp. 1-6.

2. **W. Shi**, J. Li, N. Cheng, F. Lyu, Y. Dai, H. Zhou, and X. Shen, "3D multi-drone-cell trajectory design for efficient IoT data collection", *Proc. IEEE ICC'19*, May. 2019, pp. 1-6.

3. F. Lyu, P. Yang, **W. Shi**, H. Wu, W. Wu, N. Cheng and X. Shen, "Online UAV Scheduling Towards Throughput QoS Guarantee for Dynamic IoVs", *Proc. IEEE ICC'19*, May. 2019, pp. 1-6.

4. J. Li, **W. Shi**, P. Yang, and X. Shen, "On Dynamic mapping and scheduling of service function chains in SDN/NFV-enabled networks", *IEEE GLOBECOM'19*, Dec. 2019, pp. 1-6.

5. J. Li, **W. Shi**, N. Zhang, and X. Shen, "Reinforcement learning based VNF scheduling with end-to-end delay guarantee", in *Proc. IEEE/CIC ICCC'19*, Aug. 2019, pp. 572-577.

6. J. Li, **W. Shi**, Q. Ye, W. Zhuang, X. Shen, and X. Li, "Online joint VNF chain composition and embedding for 5G networks", in *Proc. IEEE GLOBECOM'18*, Dec. 2018, pp. 1-6.

7. O. Alhussein, T.P. Do, J. Li, Q. Ye, **W. Shi**, W. Zhuang, X. Shen, X. Li, and J. Rao, "Joint VNF Placement and Multicast Traffic Routing in 5G Core Networks," *Proc. IEEE GLOBECOM'18*, Dec. 2018, pp. 1-6.

8. W Xu, H Wu, J Chen, **W. Shi**, H Zhou, N Cheng, and X. Shen, "ViFi: Vehicle-to-Vehicle Assisted Traffic Offloading via Roadside WiFi Networks", *Proc. IEEE GLOBECOM'18*, Dec. 2018, pp. 1-6.

9. J. Li, N. Zhang, Q. Ye, **W. Shi**, W. Zhuang, and X. Shen, "Joint resource allocation and online virtual network embedding for 5G networks", in *Proc. IEEE GLOBECOM'17*, Dec. 2017, pp. 1-6.

10. W Xu, H Zhou, **W. Shi**, F Lyu, and X. Shen, "Throughput Analysis of In-vehicle Internet Access via On-road WiFi Access Points", *Proc. IEEE VTC-Fall'17*, Sept. 2017, pp. 1-6.

Book Chapter

1. **W. Shi**, J. Li, and N. Zhang, "Resource Allocation in UAV-Aided Wireless Networks," *Encyclopedia of Wireless Networks, Springer*, 2019.