Wireless Communications and Networking with Unmanned Aerial Vehicles: Fundamentals, Deployment, and Optimization

Mohammad Mozaffari

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

Walid Saad, Chair
Jeffrey H. Reed
Harpreet Dhillon
Mantu K. Hudait
Anil Kumar S. Vullikanti

May 16, 2018
Blacksburg, Virginia

Keywords: Unmanned Aerial Vehicle (UAV), 3D Wireless Networks, Drone, Optimization, Optimal Transport Theory, Performance Analysis

Wireless Communications and Networking with Unmanned Aerial Vehicles: Fundamentals, Deployment, and Optimization

Mohammad Mozaffari

ABSTRACT

The use of aerial platforms such as unmanned aerial vehicles (UAVs), popularly known as drones, has emerged as a promising solution for providing reliable and cost-effective wireless communications. In particular, UAVs can be quickly and efficiently deployed to support cellular networks and enhance their quality-of-service (QoS) by establishing line-of-sight communication links. With their inherent attributes such as mobility, flexibility, and adaptive altitude, UAVs admit several key potential applications in wireless systems. Remarkably, despite these inherent advantages of UAV-based communications, little work has analyzed the performance tradeoffs associated with using UAVs as aerial wireless platforms. The key goal of this dissertation is to develop the analytical foundations for deployment, performance analysis, and optimization of UAV-enabled wireless networks. This dissertation makes a number of fundamental contributions to various areas of UAV communications that include: 1) Efficient deployment of UAVs, 2) Performance evaluation and optimization, and 3) Design of new flying, three-dimensional (3D) wireless systems. For deployment, using tools from optimization theory, holistic frameworks are developed for the optimal 3D placement of UAV base stations in uplink and downlink scenarios. The results show that the proposed deployment approaches significantly improve the downlink coverage for ground users, and enable ultra-reliable and energy-efficient uplink communications in Internet of Things (IoT) applications. For performance optimization, a novel framework is developed for maximizing the performance of a UAV-based wireless system, in terms of data service, under UAVs' flight time constraints. To this end, using the mathematical framework of optimal transport theory, the optimal cell associations, that lead to a maximum data service to ground users within the limited UAVs' hover duration, are analytically derived. The results shed light on the tradeoff between hover time and quality-of-service in UAV-based wireless networks. For performance evaluation, this dissertation provides a comprehensive analysis on the performance of a UAV-based communication system in coexistence with a terrestrial network. In particular, a tractable analytical framework is proposed for analyzing the coverage and rate performance of a network with a UAV base station and device-to-device (D2D) users. The results reveal the fundamental tradeoffs in such a UAV-D2D network that allow adopting appropriate system design parameters. Then, this dissertation sheds light on the design of three new drone-enabled wireless systems. First, a novel framework for effective use of cache-enabled UAVs in wireless networks is developed. The results demonstrate how the users' quality of experience substantially improves by exploiting UAVs' mobility and user-centric information. Second, a new framework is proposed for deploying and operating a drone-based antenna array system that delivers wireless service to ground users within a minimum time. The results show significant enhancement in QoS, spectral and energy efficiency while levering the proposed drone antenna array system. Finally, to effectively incorporate various use cases of drones ranging from aerial users to base stations, the new concept of a fully-fledged 3D cellular network is introduced. For this new type of 3D wireless network, a unified framework for deployment, network planning, and performance optimization is developed that yields a maximum coverage and minimum latency in the network. In a nutshell, the analytical foundations and frameworks presented in this dissertation provide key guidelines for effective design and operation of UAV-based wireless communication systems.

Wireless Communications and Networking with Unmanned Aerial Vehicles: Fundamentals, Deployment, and Optimization

Mohammad Mozaffari

General Audience Abstract

Unmanned aerial vehicles (UAVs), commonly known as drones, have been the subject of concerted research over the past few years, owing to their autonomy, flexibility, and broad range of application domains. Indeed, UAVs have been considered as enablers of various applications that include military, surveillance and monitoring, telecommunications, delivery of medical supplies, and rescue operations. The unprecedented recent advances in drone technology has made it possible to widely deploy UAVs, such small aircrafts, balloons, and airships for wireless communication purposes. In particular, if properly deployed and operated, UAVs can provide reliable and cost-effective wireless communication solutions for a variety of real-world scenarios. On the one hand, drones can be used as aerial base stations that can deliver reliable, cost-effective, and on-demand wireless communications to desired areas. On the other hand, drones can function as aerial user equipments, known as cellular-connected UAVs, in coexistence with ground users.

Despite such promising opportunities for drones, one must address a number of technical challenges in order to effectively use them for each specific networking application. For instance, while using drone-BS, the key design considerations include performance characterization, optimal 3D deployment of drones, wireless and computational resource allocation, flight time and trajectory optimization, and network planning. Meanwhile, in the drone-UE scenario, handover management, channel modeling, low-latency control, 3D localization, and interference management are among the main challenges.

Therefore, this dissertation addresses the fundamental challenges in UAV-enabled wireless communications that allows providing broadband, wide-scale, cost-effective, and reliable wireless connectivity. To this end, various mathematical frameworks and efficient algorithms are proposed to design, optimize, deploy, and operate UAV-based communication systems. The results shows that, the proposed aerial communication system can deliver ultra-reliable, and cost-effective wireless services, thus providing ubiquitous high speed Internet connectivity for the whole world.

To my parents, my sister, and Shaghayegh

Acknowledgments

First and foremost, I owe my deepest gratitude to my advisor, Dr. Walid Saad, for his continuous support of my Ph.D study, his patience, motivation, and immense knowledge and experience. I would like to thank you for the amount of time and effort, ideas, and funding you have generously dedicated to make my Ph.D. experience productive and stimulating. I am also grateful for believing in me and giving me the freedom to pursue diverse, yet coherent research directions which has made my Ph.D. study a joyful and unforgettable journey. I am thankful for your priceless advice and great supervision that have helped me to grow as a research scientist and find the right path for my future career. I would like to thank the members of my Ph.D. advisory committee, Dr. Jeffrey H. Reed, Dr. Harpreet Dhillon, Dr. Mantu K. Hudait, and Dr. Anil Kumar S. Vullikanti, for their valuable comments which have helped me to substantially improve the quality of this dissertation.

I am grateful for having the opportunity to work with great researchers and scientists during my Ph.D. study to publish scholarly papers. Special thanks to my collaborators, Dr. Merouane Debbah, Dr. Mehdi Bennis, and Dr. Ismail Guvenc for their time and effort to provide me with their insightful comments to improve my work throughout my PhD.

I would like to thank the folks at Wireless@VT, especially my friends at NetSciWis lab for their help and support. In particular, thanks to Priyabrata for his great help during my PhD. Additional gratitude is offered to my amazing friends: Homa, Mehrnoosh, Rashidi, Eragh, Lotfian, Beshkar, and Eskandar.

I am greatly thankful to my family, Zahra, Ali, Saeed, MohammadReza, and Shayan for their sustained encouragement, support, and unconditional love.

Last, but not least, words cannot express how grateful I am to Mehrnaz. Indeed, without her support, I could have never been able to finish this dissertation.

Contents

List of Figures

List of Tables

Chapter 1

Motivation, Background, and Contributions

Recently, the use of unmanned aerial vehicles (UAVs) such as drones, small aircrafts, and airships has received significant attention for various applications that includes military, surveillance and monitoring, telecommunications, delivery of medical supplies, and rescue operations, among others [1–14]. For each application, one needs to use an appropriate type of UAV that can meet various requirements imposed by the application, environment, and Federal rules. In fact, to properly use UAVs for each specific application, several factors such as UAV capabilities and flying altitudes need to be taken into account. In general, UAVs can be categorized, based on their altitudes, into high altitude platforms (HAPs) and low altitude platform (LAPs). HAPs have altitudes above 17 km and are typically quasi-stationary [11, 15]. LAPs, on the other hand, can fly at altitudes of up to a few kilometers, can quickly move, and have a flexible deployment [15]. Compared to HAPs, the deployment of LAPs can be done more rapidly thus making them more appropriate for time-sensitive applications (e.g. emergency situations). However, HAPs have longer endurance and, hence, are more useful for long term (e.g. up to few months) operations. UAVs can also be categorized, based on type, into fixed-wing and rotary-wing UAVs. Compared to rotary-wing UAVs, fixed-wing UAVs such as small aircrafts have more weights, higher speed, and they need to move forward in order to remain aloft. In contrast, rotary-wing UAVs such drones and quadcopter drones, can hover and remain stationary over a given area [11]. In Figure 1.1, we provide an overview on the different types of UAVs, their functions, and capabilities.

From a networking perspective, if properly deployed and operated, UAVs can provide reliable and cost-effective wireless communication solutions for a variety of real-world scenarios. On the one hand, drones can be used as aerial base stations (BSs) that can deliver reliable, cost-effective, and on-demand wireless communications to desired areas [4–13, 15–20]. On the other hand, drones can function as aerial user equipments (UEs), known as cellular-connected UAVs, in coexistence with ground users. In particular, Compared to conventional, terrestrial base stations, UAV-based aerial base stations can provide on-demand communications. Furthermore, the adjustable altitude of UAVs enables them to effectively establish line-of-sight (LoS) communication links thus mitigating signal blockage and shadowing. Due to such advantages, UAVs admit many potential use

cases in wireless networks. For instance, UAV-based aerial base stations can be deployed to enhance the wireless capacity and coverage at temporary events or hotspots such as sport stadiums and outdoor events. Moreover, they can be used in public safety scenarios to support disaster relief activities and to enable communications when conventional terrestrial networks are damaged [6]. Another important application of UAVs is in the Internet of Things (IoT) in which the devices have small transmit power and may not be able to communicate over a long range. In such scenarios, a UAV can provide a means to collect the IoT data from one device and deliver it to the intended receiver [21] and [22]. UAVs can also be used for surveillance scenarios, a key use case for the IoT. Last but not least, in regions or countries in which building a complete cellular infrastructure is very expensive, deploying UAVs is highly beneficial as it removes the need for expensive towers and infrastructure deployment.

Key examples of recent projects on employing UAVs for wireless connectivity include Google Loon project and Facebook's Internet-delivery drone [23]. Within the scope of these practical deployments, UAVs are being used as HAPs to deliver Internet access to developing countries and provide airborne global Internet connectivity. Moreover, Qualcomm and AT&T are planning to employ LAP-UAVs for enabling widescale communications in the upcoming fifth generation (5G) wireless networks [24]. Despite the numerous advantages and critical applications that can use UAVs as flying base stations, one must overcome a number of technical challenges. These challenges include three-dimensional (3D) deployment, performance analysis, mobility, air-to-ground channel modeling, user association, and flight time optimization, among others [5, 7–13, 15–18, 25–33]. In this chapter, we first present the main benefits and applications of UAV-based communication systems. Then, we provide the challenges and the relevant existing literature, followed by our contributions.

1.1 Benefits and Applications of UAV Communications

In this section, we present the main advantages and potential applications of the UAVs in wireless networks.

1.1.1 UAVs as Base Stations for Coverage and Capacity Enhancement

Due to the extensive use of wireless devices in variety of applications, there is an ever-growing need for a worldwide coverage and a high speed wireless service. Nevertheless, due to the several limitations such as costs and geographical constraints (e.g. mountain and forest) there are many regions which do not have wireless service or suffer from poor connectivity and quality-of-service. Example of regions with poor wireless coverage includes rural areas and countries (e.g. in Africa) in which building a complete terrestrial infrastructure is very expensive or infeasible. Moreover, during major public events such as Olympic games in which a substantial demand for communication is observed, the existing cellular network infrastructure, capacity, and coverage will need to be

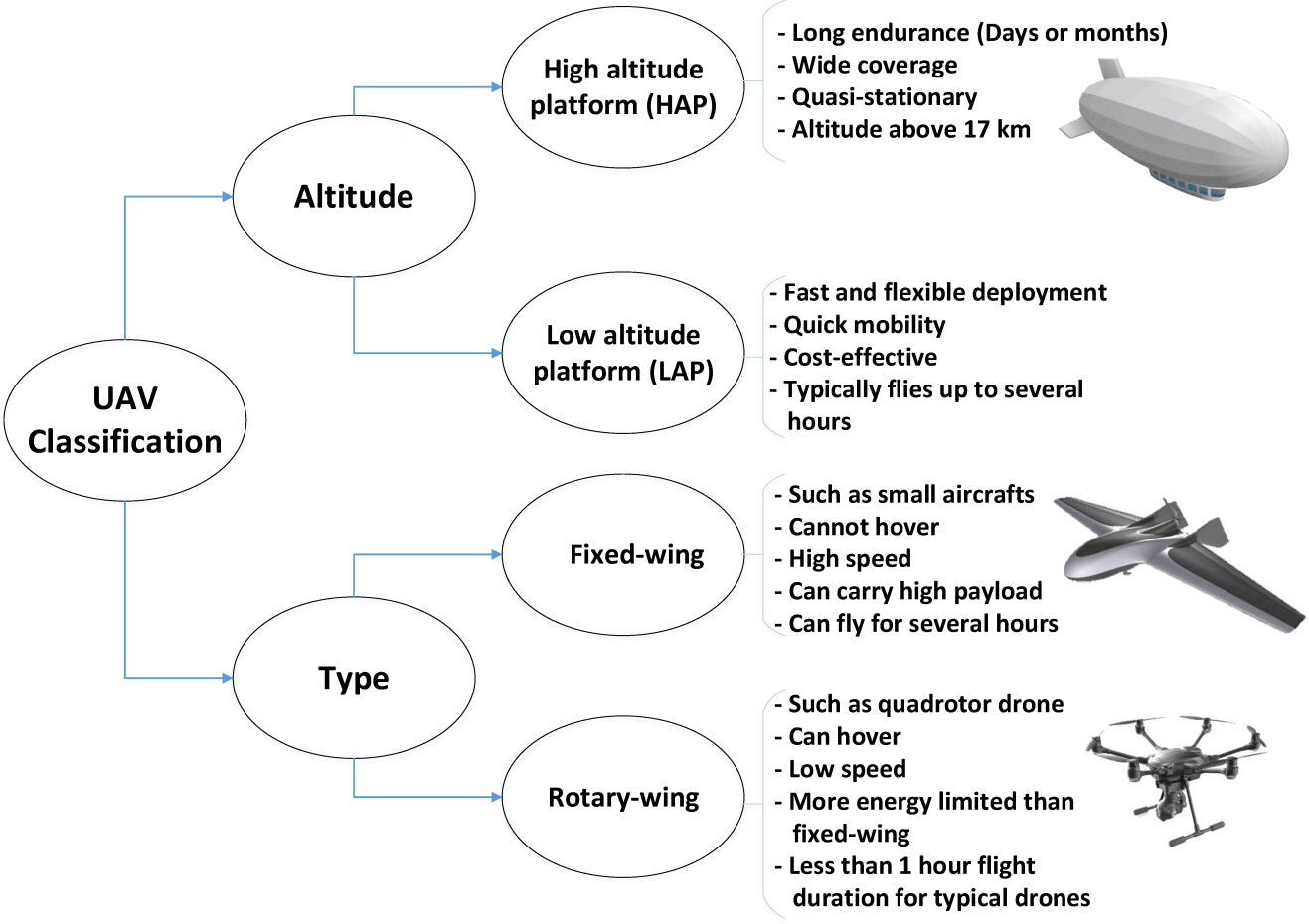


Figure 1.1: UAV Classification

rapidly boosted in order to cope with such a surge in demand. In such scenarios, UAV-based wireless communication provides an ideal solution to provide wireless services. For instance, AT&T and Verizon are planning to use flying drones to provide temporarily boosted Internet coverage for college football national championship and Super Bowl [34].

For the aforementioned scenarios, the use of UAVs is quite natural due to their agility, mobility, flexibility, and adaptive altitude which, in turn, enables them to boost the performance of existing ground wireless networks in terms of coverage, capacity, delay, and overall quality-of-service. This is due to the fact that, mobile UAVs can intelligently move and change their locations to provide on-demand service for specific areas which are not effectively covered by terrestrial networks. Another key feature of UAVs that can potentially assist in coverage and rate enhancement is having LoS connections towards ground users. In particular, due to the aerial nature of the UAVs and their high altitude, they can be effectively deployed to reduce the shadowing and blockage effects. In

addition, the coverage range of UAVs can be intelligently optimized by adjusting their altitude. In particular, UAVs can be deployed at optimal altitudes at which they can provide a maximum coverage or capacity for ground users. Therefore, by optimal deployment and intelligent mobility of UAVs according to demands, and establishing LoS communication links, UAVs can enhance coverage and capacity of wireless networks. Clearly, UAV-based wireless communications can effectively provide fast, reliable and cost-effective connectivity to areas which are either congested (e.g. hotspots) or poorly covered by terrestrial networks [7, 12, 29, 35].

1.1.2 UAVs in Public Safety Scenarios

Natural disasters such as floods, hurricanes, tornados, and severe snow storms are devastating for communities across the countries. During wide-scale natural disaster and unexpected events, the existing terrestrial communication networks can be damaged or even completely destroyed, and become significantly overloaded. In particular, cell phone towers, and ground communications infrastructure are often compromised during natural disasters. In such critical scenarios, there is a vital need for public safety communications between first responders and victims for search and rescue operations. Consequently, a robust, fast, and capable emergency communication system is needed to enable effective communications during public safety operations. To enable reliable communications during the natural disasters, the use of aerial networks such as UAVs, as shown in Figure 1.2, is considered as a promising solution [35]. Since UAVs do not need high infrastructure (e.g. cables), they can easily fly and dynamically change their positions to provide on-demand communications to ground users. In fact, due the unique features of UAVs such as mobility, flexible deployment, and rapid reconfiguration, they can effectively establish public safety communications. For instance, UAVs can be deployed as aerial base stations in order to deliver broadband connectivity to areas with damaged terrestrial wireless infrastructures. Moreover, flying UAVs can continuously move to provide full coverage to a given areas within a minimum possible time. Therefore, the use of UAV-mounted base stations can be a very appropriate solution for providing fast and ubiquitous connectivity in public safety scenarios.

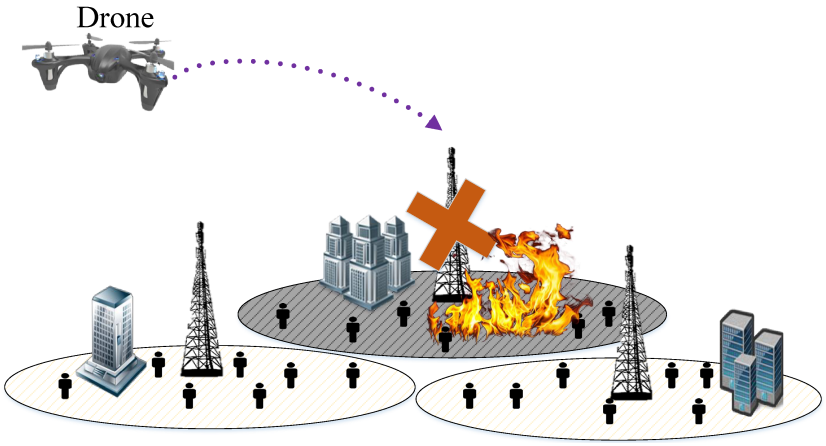


Figure 1.2: Drone in public safety scenarios.

1.1.3 Energy-Efficient IoT Communications

Wireless networking technologies are rapidly evolving into a massive IoT environment that must integrate a heterogeneous mix of devices ranging from conventional smartphones and tablets to vehicles, sensors, wearables, and naturally, drones. Realizing the much coveted applications of the IoT such as smart cities infrastructure management, healthcare, transportation, and energy management [21, 36, 37] requires effective wireless connectivity among a massive number of IoT devices that must reliably deliver their data, typically at high data rates or ultra low latency. The massive nature of the IoT requires a major rethinking to the way in which conventional wireless networks (e.g., cellular systems) operate.

For instance, in an IoT environment, energy efficiency, ultra low latency, reliability, and high-speed uplink communications become major challenges that are not typically as critical in conventional cellular network use cases [37]. In particular, IoT devices are highly battery limited and are typically unable to transmit over a long distance due to their energy constraints. For instance, in areas which experience an intermittent or poor coverage by terrestrial wireless networks, battery-limited IoT devices may not be able to transmit their data to distant base stations due to their power constraints. Furthermore, due to the various applications of IoT devices, they might be deployed in environments with no terrestrial wireless infrastructure such as mountains and desert areas.

In this regard, UAVs can also play a key role for IoT communications. In particular, in areas which experience an intermittent or poor coverage by terrestrial wireless networks, battery-limited IoT devices may not be able to transmit their data to distant base stations due to their power constraints. Furthermore, due to the various applications of IoT devices, they might be deployed in environments with no terrestrial wireless infrastructure such as mountains and desert areas. In such challenging scenarios, UAVs can be used as flying base stations to provide reliable and energy-efficient uplink IoT communications [4, 7, 38]. In fact, due to the aerial nature of the UAVs and

their high altitude, they can be effectively deployed to reduce the shadowing and blockage effects as the major cause of signal attenuation in wireless links. As a result of such efficient placement of UAVs, the communication channel between IoT devices and UAVs can be significantly improved. Subsequently, the battery-limited IoT devices will need a significantly lower power to transmit their data to UAVs. In other words, UAVs can be placed based on the locations of IoT devices enabling those devices to successfully connect to the network using a minimum transmit power. Moreover, UAVs can also serve large IoT networks by dynamically updating their locations based on the activation pattern of IoT devices. Clearly, by exploiting such unique features of UAVs, the connectivity of IoT networks can be significantly improved.

1.1.4 UAVs as Flying Backhaul for Terrestrial Networks

Wired backhauling is a common approach for connecting base stations to a core network in terrestrial networks. However, wired connections can be expensive and infeasible due to geographical constraints, especially when dealing with ultra dense cellular networks [39–41]. While wireless backhauling is a viable and cost-effective solution, it suffers from blockage and interference that degrade the performance of the radio access network [42]. In this case, UAVs can play a key role in enabling cost-effective, reliable, and high speed wireless backhaul connectivity for ground networks. In particular, UAVs can be optimally placed to avoid obstacles and establish LoS and reliable communication links. Moreover, the use of UAVs with mmW capabilities can establish high data rate wireless backhaul connections that are needed to cope with high traffic demands in congested areas. UAVs can also create a reconfigurable network in the sky and provide multi-hop LoS wireless backhauling opportunities. Clearly, such flexible UAV-based backhaul networks can significantly improve the reliability, capacity, and operation cost of backhauling in terrestrial networks.

In summary, we have presented the main advantages and potential applications of UAVs is wireless networks. In particular, we have discussed the effective use of UAVs as aerial base stations for coverage and capacity enhancement in cellular networks and public safety scenarios. Furthermore, in this section, the promising application of UAVs in IoT systems and aerial backhauling were discussed. Despite such benefits and applications of UAVs in wireless networks, there are several technical challenges that must be also taken into account. Next, we will discuss various UAV-related challenges in detail.

1.2 Challenges and Related Studies in UAV Communications

Beyond the numerous advantages of the UAVs, effectively utilizing UAVs requires meeting several key challenges. The first challenge in UAV-based communication systems is air-to-ground channel modeling. The air-to-ground channel characteristics significantly differ from classical ground communication channels. In particular, any movement and vibration of UAVs can affect the chan-

nel characteristics. Moreover, the air-to-ground channel is highly dependent on the altitude and type of the UAV, elevation angle, and type of the propagation environment. Therefore, finding a generic channel model for UAV-to-ground communications needs comprehensive simulations and measurements in various environments. In addition, the impact of UAV's altitude must be characterized for the channel modeling. Clearly, capturing such factors is challenging in air-to-ground channel modeling.

The optimal placement of UAVs is another challenging task as it depends on many parameters such as objective of the UAVs' deployment (e.g. coverage maximization, etc.), locations of ground users, and UAV-to-ground channel characteristics which itself is a function of the UAV's altitude. In addition, simultaneously deploying multiple UAVs becomes more challenging due to the impact of inter-cell interference on the system performance. In fact, the deployment of UAVs is significantly more challenging than deploying ground base stations, as is done in conventional cellular network planning. Unlike terrestrial base stations UAVs needs to be deployed in a continuous 3D space while considering the impact of altitude on the air-to-ground channel characteristics. Moreover, while deploying UAVs, their flight and energy constraints need to be also taken into account, as they directly impact the network performance.

Trajectory optimization of UAVs is also an important challenge in designing UAV-enabled communication systems. Naturally, optimizing the flight path of UAVs is challenging as it requires considering many physical constraints and parameters. For instance, while finding the trajectories of UAVs for a performance optimization, one needs to consider various key factors such as channel variation due to the mobility, energy consumption of UAVs, and flight constraints. Furthermore, solving a continuous UAVs' trajectory optimization problem is known to be analytically challenging as it involves finding an infinite number of optimization variables (i.e. UAV's locations) [11].

Furthermore, while designing UAV-based communication systems, a fundamental performance analysis needs to be done in order to evaluate the impact of design parameters on the overall system performance. Naturally, devising a fundamental analysis of the wireless performance of UAV-based system will substantially differ from conventional ground base stations due to the altitude and potential mobility of UAVs as well as their different channel characteristics. The stringent energy limitations of UAVs also introduce unique challenges. In fact, the limited available on-board energy of UAVs which leads to the short flight duration is a major factor impacting the performance of wireless communications using UAVs. For instance, UAVs must use their on-board energy for their transmission, processing, operation, and mobility purposes. Such energy constraints, in turn, lead to limited flight and hover time durations. Hence, while designing UAV communication systems, energy and flight constraints of UAVs need to be explicitly taken into account. Indeed, such limitations of UAVs present a unique design challenge for UAV-based communication systems by significantly affecting the system performance.

Other challenges for UAV communications include security and cyber-physical attacks, wireless backhaul connectivity, and interference management [11, 12, 35, 43–46]. Figure 1.3 summarizes the benefits, applications, and challenges of the UAV-based communications. Next, we present the existing works that have addressed some of the challenges pertaining to air-to-ground channel

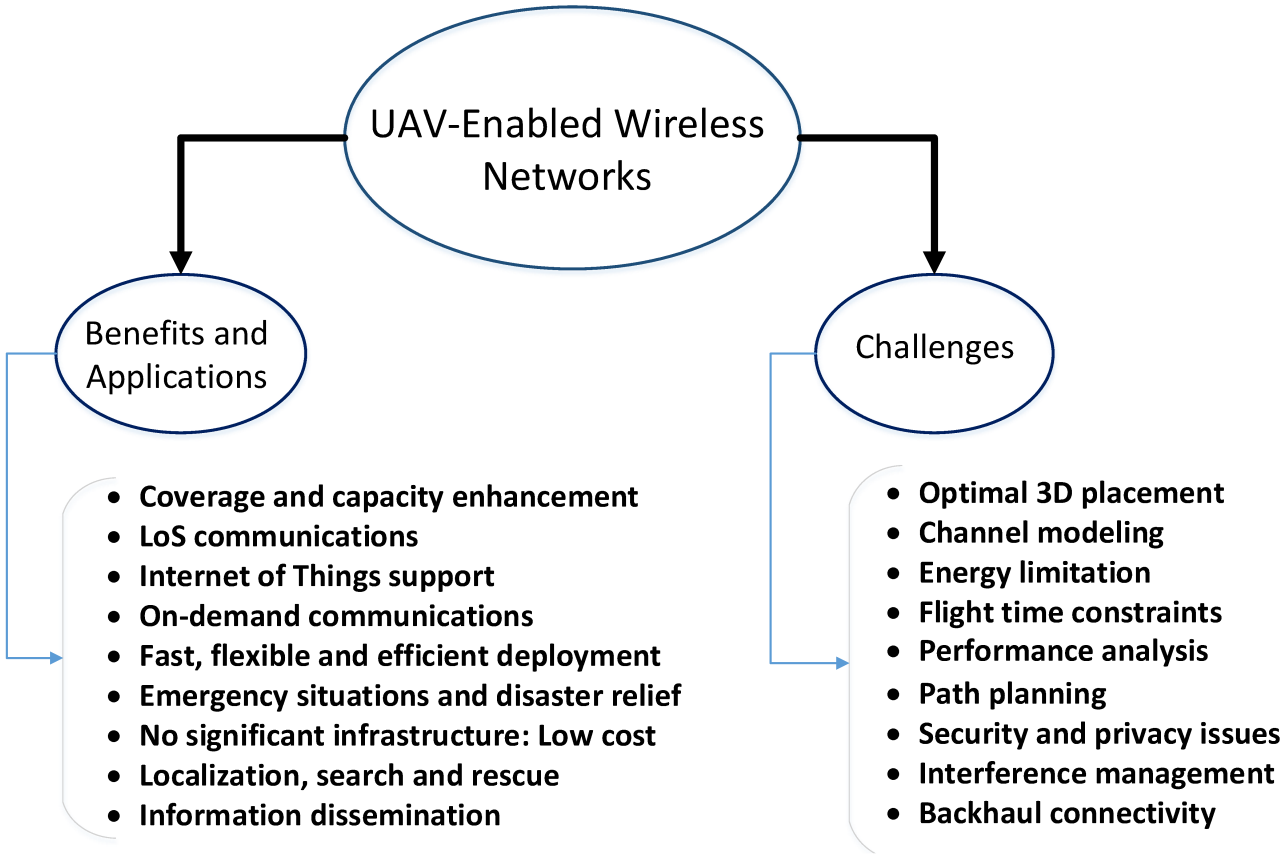


Figure 1.3: Opportunities, applications, and challenges of UAV-enabled wireless networks.

modeling, deployment, performance analysis, trajectory optimization, and energy efficiency.

1.2.1 Air-to-Ground Path Loss Modeling

Wireless signal propagation is naturally affected by the medium between the transmitter and the receiver. The air-to-ground (A2G) channel characteristics for UAV-to-ground transmission significantly differ from classical ground communication channels which, in turn, can determine the performance of UAV-based wireless communications in terms of coverage and capacity [15, 47–49]. Here, we discuss a number of recent suites on A2G channel modeling. In [47] and [48], the authors perform thorough path loss modeling for high altitude A2G communications. As discussed in [47], due to the different propagation environments and high elevation angle between transmitter and receiver in A2G, A2G links experience lower path loss and shadowing compared to the cellular network links. The authors in [49] presented a channel propagation model for high altitude platforms and ground users communications in an urban area. In this work, based on empirical re-

sults, the statistical characteristics of the channel are modeled as a function of the elevation angle. In particular, the authors in [49] considered LoS and NLoS links between the HAP and ground users and derived the probability of occurrence associated with each link. In [50], the likelihood of LoS links for A2G communication was derived as a function of elevation angle and average height of buildings in urban environments. In addition, there are some measurement-based studies on UAV-to-ground channel modeling such as [51–53] that identified some of the key channel characteristics. However, the results given in these studies are mainly valid for a specific system setup and environment. In fact, the works in [51–53] do not provide any generic channel model that can be used for designing UAV-based wireless communication systems. Nonetheless, these works provide some insights on the A2G channel characteristics that can be used to find a more generic channel model.

The most widely adopted A2G channel model for low altitude platforms is presented in [15] and thus, we explain it in more detail. As shown in [15], the path loss between the UAV and the ground node depends on the locations of the UAV and the ground node as well as the type of propagation environment (e.g., rural, suburban, urban, high-rise urban). In this case, depending on the environment, A2G communication links can be either LoS or NLoS. Note that, without any additional information about the exact locations, heights, and number of the obstacles, one must consider the randomness associated with the LoS and NLoS links. As a result, many of the existing literature on UAV communication (e.g. [5, 12, 18, 27, 28, 54–57]) adopted the probabilistic path loss model given in [8, 15]. As discussed in these works, the LoS and NLoS links can be considered separately with different probabilities of occurrence. The probability of occurrence is a function of environment, density and height of buildings, and elevation angle between the UAV and ground node. The common probabilistic LoS model is based on the general geometrical statistics of various environments provided by the International Telecommunication Union (ITU-R) [58]. In particular, for various types of environments, the ITU-R provides some environmental-dependent parameters to determine the density, number, and hight of the buildings (or obstacles). For instance, according to [58], the buildings' heights can be modeled using a Rayleigh distribution as:

$$f(h_B) = \frac{h_B}{\gamma^2} \exp\left(\frac{-h_B}{2\gamma^2}\right), \tag{1.1}$$

where h_B is the height of buildings in meters, and γ is a environmental-dependent parameter [8]. Clearly, due to the randomness (uncertainty) associated with the height of buildings (from a UAV perspective), one needs to consider a probabilistic LoS model while designing the UAV-based communication systems. Therefore, using the statistical parameters provided by ITU-R, other works such as [15] and [8] derived an expression for the LoS probability, which is given as follows [5, 15, 18, 27, 28, 54–57]:

$$P_{\text{LoS}} = \frac{1}{1 + C\exp(-B\left[\theta - C\right])}, \tag{1.2}$$

where C and B are constant values which depend on the environment (rural, urban, dense urban, or others) and θ is the elevation angle in degrees. Clearly, $\theta = \frac{180}{\pi} \times \sin^{-1}\left(\frac{h}{d}\right)$, with h being the

UAV's altitude, and d is the distance between the UAV and a given ground user. In this case, the NLoS probability will be $P_{\text{NLoS}} = 1 - P_{\text{LoS}}$.

Equation (8.2) shows that the probability of having LoS connection between the aerial base station and ground users is an increasing function of elevation angle. In other words, by increasing the elevation angle between the receiver and the transmitter, the blockage effect decreases and the communication link becomes more LoS.

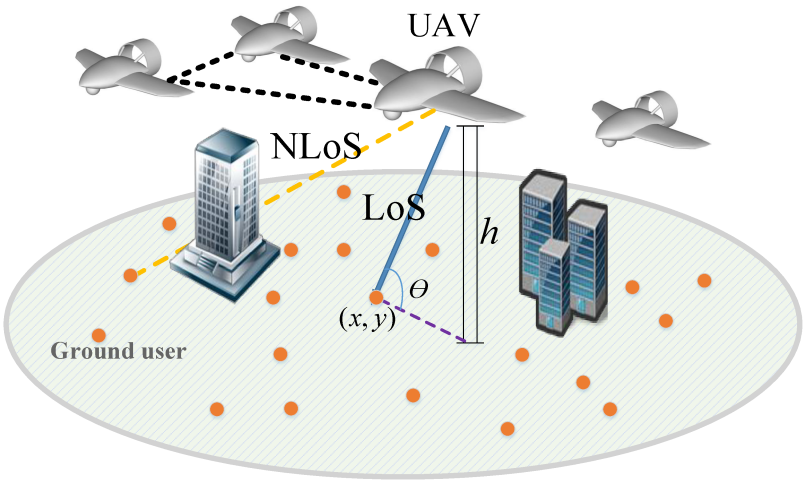


Figure 1.4: UAV communication.

In summary, UAV-to-ground communication channel depends on a number of key factors such as propagation environment, and elevation angle between the UAV and a given ground node. For path loss modeling in air-to-ground communications, a common approach is to use the probabilistic LoS/NLoS model. In this model, the LoS probability associated with a UAV-to-ground link is determined based on the elevation angle as well as the type of propagation environment.

1.2.2 Optimal Deployment

The three dimensional deployment of UAVs is one of the key challenges in UAV-based communications. In fact, the adjustable height of UAVs and their potential mobility provide additional degrees of freedom for an efficient deployment. As a result, optimal deployment of UAVs has received significant attention. In [8], the authors derived the optimal altitude enabling a single UAV to achieve a maximum coverage radius. In this work, the deterministic coverage range is determined by comparing the average path loss with a specified threshold. As shown in [8], for very low altitudes, due to the shadowing effect, the probability of LoS connections between transmitter and receiver decreases and, consequently, the coverage radius decreases. On the other hand, at very high altitudes, LoS links exist with a high probability. However, due to the large distance between transmitter and receiver, the path loss increases and consequently the coverage performance

decreases. Therefore, to find the optimal UAV's altitude, the impact of both distance and LoS probability should be considered simultaneously. In [9], we extended the results of [8] to the case of two UAVs while considering interference between the UAVs. In [10], we investigated the optimal 3D placement of multiple UAVs, that use directional antennas, to maximize total coverage area. Following our work, in [18], the authors studied the efficient deployment of aerial base stations to maximize the coverage performance. Furthermore, the authors in [18] determined the minimum number of drones needed for serving all the ground users within a given area. In [12], the authors studied the 3D placement of drones with the goal of maximizing the number of ground users which are covered by the drone. The work in [27] analyzed the impact of a UAV's altitude on the sum-rate maximization of a UAV-assisted terrestrial wireless network. In [29], the authors used evolutionary algorithms to find the optimal placement of LAPs and portable base stations for disaster relief scenarios. In this work, by deploying the UAVs at the optimal locations, the number of base stations required to completely cover the desired area was minimized. As discussed in [29], increasing the number of LAPs significantly reduces the number of required portable base stations. However, a high number of LAPs induces high interference which can degrade the coverage performance.

Moreover, the deployment of UAVs for supplementing existing cellular infrastructure was discussed in [30]. In particular, a general view of the potential integration of UAVs with cellular networks was presented. Integrating UAVs with the existing cellular network provides an enhanced coverage with low costs as the deployment of UAVs does not require significant infrastructure investments. In [59], the authors investigated the optimal deployment of a UAV that acts as a wireless relay between the transmitter and the receiver. The optimal location of the UAV was determined by maximizing the average rate while ensuring that the bit error rate will not exceed a specified threshold. As shown in [59], the UAV should be placed closer to the node (transmitter or receiver) which has a lower link quality to the UAV. The authors in [60] studied the use of UAVs to maintain the connectivity of the wireless networks. In this work, flying UAVs are deployed as a relay network to guarantee the massage delivery of sensors to destinations. The work in [61] investigates the use of multiple UAVs as wireless relays in order to provide service for ground sensors. In particular, this work addressed the tradeoff between connectivity among the UAVs and maximizing the area covered by the UAVs. However, the work in [61] does not consider the use of UAVs as aerial base stations and their mutual interference in downlink communications.

In general, despite being identified as a key challenge as discussed in [8–10, 18, 27, 29, 30, 59–61], the optimal 3D placement of UAVs acting as aerial base stations has not been thoroughly investigated in the literature. In particular, prior to our works in [9, 10], the optimal 3D deployment of multiple UAVs for various performance metrics optimization was not studied , except for some ad hoc network relaying scenarios [62, 63]. Furthermore, the previous studies did not analyze the deployment of UAVs for coverage maximization over bounded target areas.

1.2.3 Performance Analysis

A fundamental analysis of the performance of UAV-enabled wireless systems is required in order to evaluate the impact of each design parameter on the overall system performance. In particular, the performance of the UAV systems must analyzed in terms of the key quality-of-service (QoS) metrics such as coverage probability and the throughput. Such performance evaluations can also reveal the inherent tradeoffs that one faces when designing UAV-based systems.

Prior to our seminal work in this area in [7], most of the existing works focused on performance analysis of UAVs acting as relays, or in ad hoc networks [63–66]. For instance, the work in [64] evaluated the performance of a UAV ad hoc network in terms of achievable transmission rate and end-to-end delay. In [65], the authors studied the use of macro UAV relays to enhance the throughput of the cellular networks. The work in [63], derived the probability of successful connectivity among ground nodes in a UAV-assisted ad hoc network. In [66], the authors analyzed the performance of UAVs acting as relays for ground nodes in a wireless network. In particular, the authors derived closed-form expressions for signal-to-noise-ratio (SNR) distribution and ergodic capacity of UAV-ground nodes links. In contrast, in [7], we considered the use of UAVs as stand-alone aerial base stations. In particular, we investigated the downlink coverage and rate performance of a single UAV that co-exists with a device-to-device communication network. Following our work in [7], the authors in [17] derived an exact expression for downlink coverage probability for ground receivers which are served by multiple UAVs. In particular, using tools from stochastic geometry [67,68], the work in [17] provided the coverage analysis in a finite UAV network considering a Nakagami-m fading channel for UAV-to-user communications. In [55], the performance of a single drone-based communication system in terms of outage probability, bit error rate, and outage capacity is investigated. The work in [69], analyzed the coverage and throughput for a network with UAVs and underlaid traditional cellular networks. In this work, using the 3D and 2D Poisson point processes (PPP), the downlink coverage probability and rate expressions are derived. In [70], the authors evaluated the performance of using UAVs for overload and outage compensation in cellular networks. Clearly, such fundamental performance analysis is needed to provide various key design insights for UAV communication systems.

1.2.4 Energy Efficiency

Flying drones have a limited amount of on-board energy which must be used for transmission, mobility, control, data processing and payloads purposes [71]. Consequently, the flight duration of drones is typically short and insufficient for providing a long-term, continuous wireless coverage. The energy consumption of the UAV also depends on the role/mission of the UAV, weather conditions, and the traveling path. Recently, there has been some studies on the energy saving and power efficiency. For instance, the work in [72] provides an analytical framework for minimizing the energy consumption of a fixed-wing UAV by determining the optimal trajectory of the UAV. In [73], the authors proposed an energy-efficient scheduling framework for cooperative UAVs communications. In [74], the authors studied the energy efficiency of drones in target tracking scenarios

by adjusting the number of active drones. Energy harvesting from vibrations and solar sources for small UAVs was investigated in [75]. Power saving by exploiting beam alignment and focusing the transmit power in narrow beams was analyzed in [76]. Furthermore, the authors in [77] studied the use of antenna array on UAVs for improving the SNR and consequently for reducing the required transmit power. The work in [78] investigated an optimal resource allocation scheme for an energy harvesting flying access point.

In summary, the energy efficiency of UAVs requires careful consideration as it significantly impact the performance of UAV-communication systems. In fact, the limited on-board energy of UAVs is a key constraint for deployment and mobility of UAVs in various applications.

1.2.5 Trajectory Optimization

The optimal path planning for UAVs is another important challenge in the UAV-based communication systems. The trajectory of a UAV is significantly affected by different factors such as flight time, energy constraints, ground users' demands, and collision avoidance. In [26], the authors investigated the optimal trajectory of UAVs equipped with multiple antennas for maximizing sum-rate in uplink communications. The work in [79] maximizes the throughput of a relay-based UAV system by jointly optimizing the UAV's trajectory as well as the source/relay transmit power. The work in [80] jointly optimized user scheduling and UAV trajectory for maximizing the minimum average rate among ground users. In [81], a UAV path planning algorithm for photographic sensing of a given geographical area was proposed. The algorithm of [81] seeks to minimize the total energy consumption of the UAV while covering the entire survey area. To this end, the authors in [81] computed the optimal set of waypoints and the optimal speed of the UAV in the path between the waypoints. In [82], considering collision avoidance, no-fly zones, and altitude constraints, the optimal paths of UAVs that minimize the fuel consumption were computed using the mixed integer linear programming. The authors in [83] investigated the path planning for UAVs in the search and localization [84] applications using the camera measurements. In this work, the problem of path planning was addressed based on maximizing the likelihood of target detection. In [63], the authors investigated how to optimally move UAVs for improving connectivity of ad hoc networks assuming that the UAV have complete information about the location of nodes. The work in [85] jointly optimized user scheduling and UAV trajectory to maximize the minimum rate of ground users. In addition, there are some works that studied the UAV trajectory optimization for localization purposes. For instance, the work in [86] investigated path planning of multiple UAVs for localization of a passive emitter. In this work, using the angle of arrival (AOA) and time difference of arrival (TDOA) information, the set of waypoints which leads to a minimum localization error was determined. However, the work in [86] is limited to localization and does not directly address any wireless communication problem. Other works on UAV navigation and cooperative control are found in [87–91].

To summarize the related works UAV communications, Table 1.1 lists the challenges and their relevant references.

Table 1.1: Categories of the existing literature on UAV

Research on UAV	**Related references**
Channel modeling	[8, 15, 47–53, 58, 92]
Deployment	[8–10, 12, 13, 18, 28–30, 59, 60, 65]
Performance analysis	[7, 9, 17, 55, 63, 66, 69]
Energy efficiency	[71–78]
Trajectory optimization	[26, 63, 79–83, 85–91]

1.3 Limitations of Existing Works

As discussed in Section 1.1 and 1.2, in order to reap the benefits of UAV-enabled wireless systems, many technical challenges such as optimal deployment of UAVs, performance analysis, and flight time considerations need to be taken into account. While the previous studies presented in Section 1.2 have addressed some of these challenges such as channel modeling, and trajectory optimization, the literature lacks comprehensive studies on deployment, performance analysis and optimization of UAV communications, notably when UAVs are seen as flying aerial base stations. In particular, several UAV-related problems such as optimal 3D deployment of multiple UAVs as aerial base stations, coexistence of aerial and terrestrial networks, and UAV communications under explicit flight time constraints need to be thoroughly studied. Moreover, the existing prior art has not studied wireless-centric applications of UAVs, such as the use of UAVs within IoT networks or as cache-enabled flying base stations. In summary, the main limitations of the previous studies on UAV communications are as follows:

- Despite the notable number of works on positioning UAVs for ad hoc network relaying [63–66], those works cannot provide a full picture on how to optimally deploy UAVs that act as aerial base stations. In particular, the previous works on deployment of UAV relays do not address the coverage maximization problem which is one of the main goal of deploying UAV-mounted base station. Furthermore, the previous studies do not optimize the altitudes of UAVs that can be done by capturing specific characteristics of air-to-ground channels (e.g. elevation angle dependency).

- While the performance of UAVs, acting as relays, has been studied, these results cannot readily extend to cases in which UAVs are used as stand-alone base stations, as is envisioned for beyond 5G networks. This is due to the fact that the different applications of relays and base stations requires analyzing different performance metrics. For instance, in the UAV-mounted base station case, one can analyze specific wireless metrics such as the coverage probability in uplink or downlink, and system sum-rate considering various multiple access, scheduling, and resource allocation techniques.

- If future networks are to deploy UAVs as flying base stations, energy efficiency and flight time constraints become very critical challenge. For instance, even though some works such as [71–78] have studied energy efficiency, these studies remain largely limited in scope as they do not analyze the interplay between energy efficiency and wireless communication performance. In particular, one must optimize the performance of UAV-enabled wireless systems under energy and flight time constrains of UAVs.
- The advances in UAV capabilities naturally expanded their networking application domains, as compared to the past few years in which they were limited to military and surveillance applications. For instance, UAVs can be used to support IoT and machine-to-machine (M2M) networks by enabling the devices to successfully connect to the network using a minimum transmit power. Moreover, the use of cache-enabled UAVs is a promising solution for traffic offloading in cellular networks. Nevertheless, to exploit the benefits of UAVs in such applications, a holistic framework in needed for efficient deployment and mobility management of UAVs.
- Co-existence of aerial base stations with terrestrial cellular networks has not been studied in the literature. Clearly, while designing the UAV-assisted cellular networks, several fundamental tradeoffs needs to be considered. For instance, the overall network performance can be improved by intelligent mobility of UAVs, however, it is limited by the capabilities of the UAVs and can lead to a significant UAVs' energy consumption. Unlike a stand-alone aerial network, in such UAV-assisted cellular networks, the mobility and deployment of UAVs must be performed with respect to the terrestrial network. In particular, to maximize the overall system performance, one must address various challenges related to aerial and terrestrial base stations such as deployment, mobility, interference management, and resource allocation.
- Trajectory optimization has been substantially studied for UAVs, from a robotics/control perspective [86–91]. However, those works do not consider UAVs to be wireless equipped or, at least, do not analyze the coupling between mobility and wireless quality-of-service. In fact, to effectively use UAV base stations, the trajectory of UAVs needs to be optimized with respect to wireless metrics such as throughput and coverage as well as energy constraints of UAVs.

  In summary, despite the existence of a notable body of work on UAV-based communications, this existing literature is largely limited in scope and does not provide the necessary fundamentals needed to understand the full potential of wireless networking using UAV-mounted base stations.

1.4 Summary of Contributions

The main contribution of this dissertation is to develop analytical foundations for deployment, performance analysis, and optimization of UAV-enabled wireless networks. In particular, we propose

various frameworks for efficient deployment, mobility, performance analysis, and optimization of UAV-based wireless communication systems. By using the proposed frameworks, the performance of such UAV systems can be optimized in terms of coverage, capacity, reliability, and energy efficiency, while taking into consideration the unique features of UAVs, such as their flight time, mobility, and flexible altitude. To enable these contributions, this dissertation weaves together notions from optimization theory, learning, optimal transport theory, probability, and statistics. Indeed, using such advanced mathematical tools, this dissertation develops in-depth analytical foundations and efficient algorithms to design, optimize, deploy, and operate UAV-based communications. In summary, our contributions are given as follows:

1.4.1 Deployment and Mobility

Optimal 3D Placement of UAVs for Maximized Coverage

In Chapter 2, we investigate the optimal 3D deployment of multiple UAVs in order to maximize the downlink coverage performance with a minimum transmit power. In particular, given a target geographical area, the coverage requirements of the ground users and a number of UAVs that use directional antennas, we develop a novel framework to determine the optimal 3D locations of the UAVs. First, we analytically derive the downlink coverage as a function of the UAV's altitude and the antenna gain. Next, using circle packing theory [93], we propose an efficient deployment method which leads to the maximum coverage performance while ensuring that the coverage areas of UAVs do not overlap. Our results show that, considering the size of the desired area, the number of UAVs and the beamwidth of the antennas, the altitudes and locations of the UAVs can be properly adjusted for satisfying the coverage requirements.

Efficient Deployment and Mobility of UAVs for Energy-Efficient IoT Communications

As we discussed in Section 1.1.3, UAVs can also play a key role in IoT communications which is composed of small, battery-limited devices such as sensors, and health monitors. In Chapter 3, we investigate the efficient deployment and mobility of multiple UAVs, used as aerial base stations to collect data from ground IoT devices. In particular, to enable reliable uplink communications for IoT devices with a minimum total transmit power, we propose a novel framework to jointly optimize: a) 3D placement and mobility of the UAVs, b) device-UAV association, and c) uplink power control. First, given the locations of active IoT devices at each time instant, the optimal UAVs' locations and associations are determined. Next, to dynamically serve the IoT devices in a time-varying network, the optimal mobility patterns of the UAVs are analyzed. To this end, based on the activation process of the IoT devices, the time instances (update times) at which the UAVs must update their locations are derived. Moreover, the optimal 3D trajectory of each UAV is obtained in a way that the total energy used for the mobility of the UAVs is minimized while serving the IoT devices. The results reveal an inherent tradeoff between the number of update

times, the mobility of the UAVs, and the transmit power of the IoT devices. In essence, a higher number of updates can lead to lower transmit powers for the IoT devices at the cost of an increased mobility for the UAVs.
The results of this work constitute one of the first comprehensive studies on the joint optimal 3D deployment of aerial base stations, device association, and uplink power control in an IoT ecosystem.

1.4.2 Performance Analysis and Optimization

Optimized UAV Communications under Flight Time Considerations

The limited flight time duration of the UAVs presents a unique design challenge for UAV-based communication systems [94] and [95]. For instance, the performance of such systems significantly depends on the hover time of each UAV, which is defined as the flight duration within which the UAV must stay in the air over a given area for providing wireless service to ground users. In fact, with a higher hover time of the UAV, the users can receive wireless service for a longer period. Thus, by increasing the hover time, a UAV can meet higher load requirements and serve a larger area. However, the hover time of a UAV is naturally limited due to the highly constrained battery-provided, on-board energy, as well as flight regulations such as no-fly time/zone constraints. Hence, while analyzing the UAV-based communication systems, the hover time constraints must be also taken into account. In this case, there is a need for a framework to analyze and optimize the performance of UAV-based communications based on the hover time of UAVs. To this end, in Chapter 4, we develop a novel framework for optimized UAV-to-ground communications under explicit UAVs' hover time constraints. In particular, we consider a network in which multiple UAVs are deployed as aerial base stations to provide wireless service to ground users that are distributed over a geographical area based on an arbitrary spatial distribution. We investigate two key practical scenarios: *UAV communication under hover time constraints*, and *UAV communication under load constraints*. In the first scenario, based on the maximum possible hover times of UAVs, the average data service delivered to the users under a fair resource allocation scheme is maximized by finding the optimal cell partitions associated to the UAVs. Using the powerful mathematical framework of optimal transport theory [96], this cell partitioning problem is proved to be equivalent to a convex optimization problem. Subsequently, a gradient-based algorithm is proposed for optimally partitioning the geographical area based on the users' distribution, hover times, and locations of the UAVs. In the second scenario, given the load requirements of ground users, we derive the minimum average hover time that the UAVs need for completely servicing their ground users. To this end, first, we propose an optimal bandwidth allocation scheme for serving the users. Then, given this optimal bandwidth allocation, we derive the optimal cell partitions associated with the UAVs by exploiting optimal transport theory. The results show that our proposed cell partitioning approach leads to a significantly higher fairness among the users compared to the classical weighted Voronoi diagram. In addition, our results characterize a fundamental tradeoff between the hover time of UAVs and bandwidth efficiency while serving the ground users.

To our best knowledge, Chapter 4 is the first work that provides a framework to analyze and optimize the performance of UAV-based communications based on the hover time of UAVs.

Performance Analysis of UAV-assisted Wireless Networks in Coexistence with D2D

Performing a fundamental performance analysis is a key step in designing UAV-based wireless systems. For scenarios in which there is limited or no infrastructure support, beyond the use of UAVs, D2D communications can also lead to coverage and capacity enhancement. In Chapter 5, we provide coverage and rate analysis for a flying UAV base station used to provide the wireless communications to a given geographical. In particular, the coexistence between the UAV, that is transmitting data in the downlink, and an underlaid D2D communication network is considered. For this model, a tractable analytical framework for the coverage and rate analysis is derived. Two scenarios are considered: a static UAV and a mobile UAV. In the first scenario, the average coverage probability and the system sum-rate for the users in the area are derived as a function of the UAV altitude and the number of D2D users. In the second scenario, using the disk covering problem, the minimum number of stop points that the UAV needs to visit in order to completely cover the area is computed. Furthermore, considering multiple retransmissions for the UAV and D2D users, the overall outage probability of the D2D users is derived. The results reveal the fundamental tradeoffs in such UAV-D2D network that allow adopting appropriate system design parameters.
This chapter presents the first comprehensive fundamental analysis on the performance of UAV communication in the presence of underlaid D2D links.

1.4.3 Communications and Control for Wireless Drone-Based Antenna Array

One promising approach to provide high data rate and low service time is to utilize multiple drones within an antenna array system composed of multiple single-antenna drones [97]. Compared to conventional antenna array systems, a drone-based antenna array has the following advantages. First, the number of antenna elements (i.e., drones) is not limited by space constraints. Second, the gain of the drone-based antenna array can be increased by adjusting the array element spacing. Third, the mobility and flexibility of drones enable an effective mechanical beam-steering in any three-dimensional (3D) direction. Clearly, a high gain drone-based antenna array can provide high data rate wireless services to ground users thus reducing the service time. This chapter develops a new framework for deploying and operating a drone-based antenna array system that delivers wireless service to a number of ground users within a minimum time. In particular, we minimize the service time that includes both the transmission time and the control time needed to control the movement and orientation of the drones. To this end, we minimize the transmission time, by optimizing the drones' locations, as well as the control time that the drones need to move between these optimal locations. To minimize the transmission time, first, we determine the optimal drone spacing for which the array directivity is maximized. In this case, using perturbation theory [98],

we solve the drone spacing optimization problem by successively solving a number of perturbed convex optimization problems. Next, given the derived drone spacing, we optimally adjust the locations of the drones according to the position of each ground user. In order to serve different users, the drones must dynamically move between the derived optimal locations, during the control time period. To minimize the control time of quadrotor drones, we determine the optimal speeds of rotors such that the drones can update their positions and orientations within a minimum time. In this case, using *bang-bang* control theory [99], we derive a closed-form expression for the minimum control time as a function of external forces (e.g., wind and gravity), the drone's weight, and the destinations of drones. Our results show that the proposed drone antenna array approach can significantly reduce the service time and improve the spectral and energy efficiency of the network. In particular, our approach yields 48% improvement in spectral efficiency compared to a case in which the same number of drones are deployed separately. The results also reveal a tradeoff between the control time and transmission time while varying the number of drones.

1.4.4 Cell Association in UAV-assisted Cellular Networks

Another contribution of this dissertation is to introduce a framework for delay-optimal cell association in a cellular network in which both UAVs and terrestrial BSs co-exist. In particular, given the locations of the UAVs and terrestrial BSs as well as any general spatial distribution of users, we find the optimal cell association by exploiting the framework of *optimal transport theory* [100]. Within the framework of optimal transport theory, one can address cell association problems for any general spatial distribution of users. In fact, the main advantage of optimal transport theory is to provide tractable solutions for a variety of cell association problems in wireless networks. In our problem, we first prove the existence of the optimal solution to the cell association problem, and, then, we characterize the solution space. The results show that, our approach results in a significantly lower delay compared to a conventional signal strength-based association.

1.4.5 Proactive Deployment of Cache-Enabled UAVs

The use of cache-enabled UAVs is a promising solution for traffic offloading and information dissemination in wireless networks. Therefore, we develop a novel framework that leverages user-centric information, such as content request distribution and mobility patterns, to effectively deploy cache-enabled UAVs. Our proposed framework includes the prediction of the users' requests, and optimal deployment and movement of the UAVs based on the demands. Hence, cache-enabled UAVs can effectively service users that cannot be serviced by static, ground base stations. In this case, we first propose a conceptor-based echo state networks (ESNs) algorithm that can accurately predict the content request distribution and mobility patterns of each user. Then, using these predictions, we determine the optimal locations and content caching strategies for the UAVs. In particular, we analytically derive the optimal user-UAV association, the optimal locations of the UAVs as well as the content to cache at the UAVs. Our results show that, the proposed approach

yields a significant gain in terms of the percentage of the users with satisfied quality of experience (QoE) compared to a benchmark scenario without UAVs.
It should be noted that, this work is the first to investigate the use of caching at the level of UAVs by exploiting user-centric information.

1.4.6 Foundations of a 3D Cellular Network with Drones

In order to support the key roles of drones in wireless networking applications, there is a need for developing the novel concept of a *3D cellular network* that incorporates both drone-BSs and drone-users (drone-UEs). Therefore, we introduce the novel concept of a fully-fledged drone-based 3D cellular network while proposing a new framework for designing the two fundamental problems of deployment and 3D cell association in such a 3D cellular network. In particular, our proposed framework includes a tractable approach for three-dimensional placement of drone-BSs and a latency-minimal 3D cell association scheme for servicing drone-UEs. For deployment, we introduce an approach based on truncated octahedron cells that determines the locations of drone-BSs that can cover a 3D space with a minimum number of drone-BSs. For the latency-minimal 3D cell association, given the locations of drone-BSs and the distribution of drone-UEs, we find the optimal 3D cell association for which the total latency of serving drone-UEs is minimized. In this case, we analytically characterize the optimal 3D cell partitions by exploiting tools from optimal transport theory. Our results show that the proposed approach significantly reduces the latency of serving drone-UEs compared to classical signal-to-interference-plus-noise ratio (SINR)-based cell association approaches. This work presents the foundations for a new breed of cellular networks that can operate in three dimensions using drones.

1.4.7 Sum-Rate Analysis for HAP Drones with Tethered Balloon Relay

In addition, we investigate the use of interference alignment techniques for maximizing the sum-rate of a relay-aided HAP drones wireless system when the CSI is not available. In particular, to achieve the maximum sum-rate, we propose a DF scheme involving HAP drones, ground receivers, and one relay with multiple antennas. In this scenario, we show that it is possible to achieve the maximum possible sum-rate by exploiting the IA scheme. Moreover, we derive a closed-form analytical expression for the capacity of an X channel with tethered balloon relay. Simulation results verify our analytical results and show that a significant sum-rate gain can be achieved by using the proposed scheme.

1.5 List of Publications

As a byproduct of the above contributions, thus far, this dissertation has led to the following key publications:

1.5.1 Journal Publications

1. **M. Mozaffari**, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems," submitted, 2018.

2. **M. Mozaffari**, A. Taleb Zadeh Kasgari, W. Saad, M. Bennis, and M. Debbah, "Beyond 5G with UAVs: Foundations of a 3D Wireless Cellular Network," submitted for journal publication, 2018.

3. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: performance and tradeoffs," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3949-3963, June 2016.

4. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Mobile Unmanned Aerial Vehicles (UAVs) for Energy-Efficient Internet of Things Communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574-7589, Nov. 2017.

5. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Wireless Communication using Unmanned Aerial Vehicles (UAVs): Optimal Transport Theory for Hover Time Optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8052-8066, Dec. 2017.

6. 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 Journal on Selected Areas in Communications (JSAC), Special Issue on Human-In-The-Loop Mobile Networks*, vol. 35, no. 5, pp. 1046-1061, May 2017.

7. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Communications Letters*, vol. 20, no. 8, pp. 1647-1650, Aug. 2016.

8. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Optimal Transport Theory for Cell Association in UAV-Enabled Cellular Networks," *IEEE Communications Letters*, vol. 21, no. 9, pp. 2053-2056, Sept. 2017.

9. S. Pai, **M. Mozaffari**, M. Magarini, W. Saad, and P. Muthuchidambaranathan, "Sum-Rate Analysis for High Altitude Platform (HAP) Drones with Tethered Balloon Relay," *IEEE Communications Letters*, accepted and to appear, 2018.

10. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Communications and Control for Wireless Drone-Based Antenna Array," submitted for journal publication, 2018.

1.5.2 Conference Publications

1. **M. Mozaffari**, A. Taleb Zadeh Kasgari, W. Saad, M. Bennis, and M. Debbah, "3D Cellular Network Architecture with Drones for Beyond 5G," submitted for conference publication, 2018.

2. Q. Zhang, **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Machine Learning for Predictive On-Demand Deployment of UAVs for Wireless Communications," submitted for conference publication, 2018.

3. A. French, **M. Mozaffari**, A. Eldosouky, and W. Saad,"Environment-Aware Deployment of Wireless Drones Base Stations with Google Earth Simulator," submitted for conference publication, 2018.

4. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Drone-based Antenna Array for Service Time Minimization in Wireless Networks," in *Proc. of IEEE International Conference on Communications (ICC)*, Kansas City, USA, May 2018.

5. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Optimal Transport Theory for Effective UAV-Based Wireless Communications under Flight Time Constraints," extended abstract in *IEEE Communication Theory Workshop (CTW)*, May 2017.

6. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Performance Optimization for UAV-Enabled Wireless Communications under Flight Time Constraints," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Singapore, 2017.

7. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Mobile Internet of Things: Can UAVs provide an energy-efficient mobile architecture?," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, USA, Dec. 2016.

8. M. Naderisoorki, **M. Mozaffari**, H. Manshaei, and H. Saidi, "Resource Allocation for Machine-to-Machine Communications with Unmanned Aerial Vehicles," in *Proc. of the IEEE Global Communications Conference (GLOBECOM), Workshop on Wireless Networking, Control, and Positioning for Unmanned Autonomous Vehicles*, Washington, DC, USA, Dec. 2016.

9. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Optimal transport theory for power-efficient deployment of unmanned aerial vehicles," in *Proc. of IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, May 2016.

10. **M. Mozaffari**, W. Saad, M. Bennis, and M. Debbah, "Drone small cells in the clouds: Design, deployment and performance analysis," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, USA, Dec. 2015.

11. V. Namblar, E. Vattapparamban, A. I. Yurekli, I. Guvenc, **M. Mozaffari**, and W. Saad, “SDR Based Indoor Localization Using Ambient WiFi and GSM Signals," in *Proc. of the International Conference on Computing, Networking and Communication (ICNC), 3rd National Workshop for REU Research in Networking and Systems (REUNS)*, Silicon Valley, USA, Jan. 2017.

Chapter 2

Efficient Deployment of Multiple Unmanned Aerial Vehicles for Optimal Wireless Coverage

2.1 Background, Related Works, and Contributions

As we discussed in Chapter 1, the use of wireless aerial platforms is a promising approach to improve the performance of wireless communication networks. In particular, UAVs can act as flying base stations to enhance the coverage and rate performance of wireless networks in different scenarios such as temporary hotspots and emergency situations [9]. In addition, mobile UAVs can establish efficient communications with ground sensors for alarm messages delivery scenarios [101]. Indeed, using UAVs as aerial base stations provides several advantages compared to the terrestrial base stations. First, due to their higher altitude, aerial base stations have a higher chance of line-of-sight (LoS) links to ground users. Second, UAVs can easily move and have a flexible deployment, and hence, they can provide rapid, on-demand communications. Finally, using directional antennas, one can further enhance the UAV-based communications due to the possibility of using effective beamforming schemes [102].

Despite the numerous advantages for using UAVs as flying base stations, one must overcome a number of technical challenges. These challenges include the optimal 3D deployment of UAVs, energy limitations, interference management, and path planning [7, 9, 61, 101, 102]. In particular, the deployment problem is of paramount importance as it highly impacts the energy consumption as well as the interference generated by UAVs. However, only a limited number of existing literature have addressed the interplay between UAV deployment and wireless performance [7,9,29,61]. For instance, in [61], the use of multiple UAVs as wireless relays in order to provide service for ground sensors is investigated. This work addressed the tradeoff between connectivity among the UAVs and maximizing the area covered by the UAVs. However, the work in [61] does not consider

the use of UAVs as aerial base stations and their mutual interference in downlink communications. In [29], the authors used evolutionary algorithms in order to find the optimal placement of LAPs and portable base stations for disaster relief scenarios. However, the model of [29] assumes that overlapping LAPs' coverage areas is allowed by using inter-cell interference coordination (ICIC). However, ICIC requires further communications between LAPs.

In this chapter, we investigate the optimal 3D deployment of multiple UAVs in order to maximize the downlink coverage performance while using a minimum transmit power. Given a target geographical area, the coverage requirements of the ground users, and a number of UAVs that use directional antennas, we develop a novel framework to determine the optimal 3D locations of the UAVs. First, we derive the downlink coverage probability for a UAV as a function of the UAV's altitude and the antenna gain. Next, using circle packing theory [93], we propose an efficient deployment method which leads to the maximum coverage performance while ensuring that the coverage areas of UAVs do not overlap. Our results show that, considering the size of the desired area, the number of available UAVs and the gain (or beamwidth) of the directional antennas, the altitude and locations of the UAVs can be appropriately adjusted for satisfying the coverage requirements. In addition, our results reveal the minimum number of UAVs required to guarantee a target coverage for a given geographical area.

The rest of this chapter is organized as follows. In Section 2.2, we present the system model and describe the air-to-ground channel model. Section 2.3 presents the coverage analysis and the proposed deployment method. In Section 2.4, we provide the simulation results, and Section 2.5 concludes the chapter.

2.2 System Model

Consider a circular geographical area of radius R_c, as illustrated in Figure 2.1, within which M UAVs must be deployed to provide wireless coverage for ground users located within the area. In this model, we consider a stationary low altitude aerial platform such as quadrotor UAVs. The UAVs are assumed to be symmetric having the same transmit power and altitude. We denote the UAV's directional antenna half beamwidth by θ_B, and, thus, the antenna gain can be approximated by [103]:

$$G = \begin{cases} G_{\text{3dB},} & \frac{-\theta_B}{2} \leq \varphi \leq \frac{\theta_B}{2}, \\ g(\varphi), & \text{otherwise}, \end{cases} \tag{2.1}$$

where φ is the sector angle, $G_{\text{3dB}} \approx \frac{29000}{\theta_B^2}$ with θ_B in degrees, is the main lobe gain [104]. Also, $g(\varphi)$ is the antenna gain outside of the main lobe. For the air-to-ground channel modeling, a common approach is to consider the LoS and non-line-of-sight (NLoS) links between the UAV and the ground users separately [15]. Each link has a specific probability of occurrence which depends on the elevation angle, environment, and relative location of the UAV and the users. Clearly, for NLoS links the shadowing and blockage loss is higher than the LoS links. Therefore, the received signal power from UAV j at a user's location can be given by [15]:

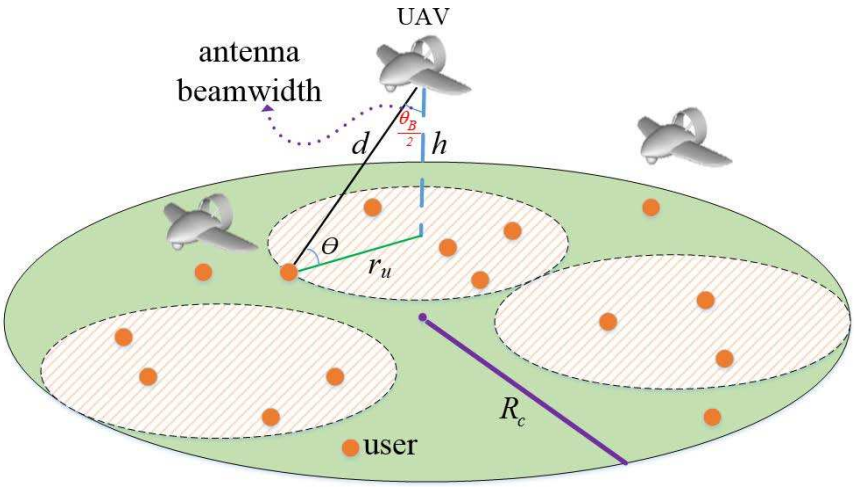


Figure 2.1: System model.

$$P_{r,j}(dB) = \begin{cases} P_t + G_{\text{3dB}} - L_{dB} - \psi_{\text{LoS}}, & \text{LoS link}, \\ P_t + G_{\text{3dB}} - L_{dB} - \psi_{\text{NLoS}}, & \text{NLoS link}, \end{cases} \tag{2.2}$$

where $P_{r,j}$ is the received signal power, P_t is the UAV's transmit power, and G_{3dB} is the UAV antenna gain in dB. Also, L_{dB} is the path loss which for the air-to-ground communication is:

$$L_{dB} = 10n\log\left(\frac{4\pi f_c d_j}{c}\right), \tag{2.3}$$

where f_c is the carrier frequency, c is the speed of light, d_j is the distance between UAV j and a ground user, and $n \geq 2$ is the path loss exponent. Also, $\psi_{\text{LoS}} \sim N(\mu_{\text{LoS}}, \sigma^2_{\text{LoS}})$ and $\psi_{\text{NLoS}} \sim N(\mu_{\text{NLoS}}, \sigma^2_{\text{NLoS}})$ are shadow fading with normal distribution in dB scale for LoS and NLoS links. The mean and variance of the shadow fading for LoS and NLoS links are $(\mu_{\text{LoS}}, \sigma^2_{\text{LoS}})$, and $(\mu_{\text{NLoS}}, \sigma^2_{\text{NLoS}})$. As shown in [15], the variance depends on the elevation angle and type of the environment as follows:

$$\sigma_{\text{LoS}}(\theta_j) = k_1 \exp(-k_2\theta_j), \tag{2.4}$$

$$\sigma_{\text{NLoS}}(\theta_j) = g_1 \exp(-g_2\theta_j), \tag{2.5}$$

where $\theta_j = \sin^{-1}(h/d_j)$ is the elevation angle between the UAV and the user, k_1, k_2, g_1, and g_2 are constant values which depend on environment. Finally the LoS probability is given by [15]:

$$P_{\text{LoS},j} = \alpha\left(\frac{180}{\pi}\theta_j - 15\right)^{\gamma}, \tag{2.6}$$

where α and γ are constant values reflecting the environment impact. Note that, the NLoS probability is $P_{\text{NLoS},j} = 1 - P_{\text{LoS},j}$. Next, using this air-to-ground channel model, we derive the downlink coverage probability and characterize an efficient scheme for deploying of multiple UAVs.

2.3 Optimal Multi-UAV deployment

First, we find the coverage radius of each UAV in the presence of interference from other UAVs. To this end, the coverage probability of a single UAV needs to be derived. Then, we propose an efficient deployment strategy for M UAVs that maximizes the total coverage performance while maximizing the coverage lifetime.

Theorem 1. The coverage probability for a ground user, located at a distance $r \leq h.\tan(\theta_{\rm B}/2)$ from the projection of the UAV j on the desired area, is given by:

$$P_{\rm cov} = P_{{\rm LoS},j} Q\left(\frac{P_{\rm min} + L_{dB} - P_t - G_{\rm 3dB} + \mu_{\rm LoS}}{\sigma_{\rm LoS}}\right) + P_{{\rm NLoS},j} Q\left(\frac{P_{\rm min} + L_{dB} - P_t - G_{\rm 3dB} + \mu_{\rm NLoS}}{\sigma_{\rm NLoS}}\right), \tag{2.7}$$

where $P_{\rm min} = 10\log\left(\beta N + \beta \bar{I}\right)$ is the minimum received power requirement (in dB) for a successful detection, N is the noise power, β is the signal-to-interference-plus-noise-ratio (SINR) threshold. $\bar{I}$ is the mean interference power received from the nearest UAV k which is given by: $\bar{I} \approx P_t g(\varphi_k)\left[10^{\frac{-\mu_{{\rm LoS},k}}{10}} P_{{\rm LoS},k} + 10^{\frac{-\mu_{{\rm NLoS},k}}{10}} P_{{\rm NLoS},k}\right]\left(\frac{4\pi f_c d_k}{c}\right)^{-n}$. Also, $Q(.)$ is the Q function.

Proof. The downlink coverage probability for a ground user considering the mean value of interference between UAVs can be written as:

$$\begin{aligned} P_{\rm cov} &= \mathbb{P}\left[\frac{P_{r,j}}{N+\bar{I}} \geq \beta\right] = \mathbb{P}\left[P_{r,j}(dB) \geq P_{\rm min}\right] \\ &= P_{{\rm LoS},j}\mathbb{P}\left[P_{r,j}({\rm LoS}) \geq P_{\rm min}\right] + P_{{\rm NLoS},j}\mathbb{P}\left[P_{r,j}({\rm NLoS}) \geq P_{\rm min}\right] \\ &\overset{(a)}{=} P_{{\rm LoS},j}\mathbb{P}\left[\psi_{\rm LoS} \leq P_t + G_{\rm 3dB} - P_{\rm min} - L_{dB}\right] \\ &+ P_{{\rm NLoS},j}\mathbb{P}\left[\psi_{\rm NLoS} \leq P_t + G_{\rm 3dB} - P_{\rm min} - L_{dB}\right] \\ &\overset{(b)}{=} P_{{\rm LoS},j} Q\left(\frac{P_{\rm min} + L_{dB} - P_t - G_{\rm 3dB} + \mu_{\rm LoS}}{\sigma_{\rm LoS}}\right) \\ &+ P_{{\rm NLoS},j} Q\left(\frac{P_{\rm min} + L_{dB} - P_t - G_{\rm 3dB} + \mu_{\rm NLoS}}{\sigma_{\rm NLoS}}\right), \end{aligned} \tag{2.8}$$

where $\mathbb{P}[.]$ is the probability notation, and $P_{\rm min} = 10\log\left(\beta N + \beta\bar{I}\right)$. Clearly, due to the use of directional antennas, interference received from the nearest UAV k is dominant. Hence, $\bar{I}$ can be written as:

$$\begin{aligned} \bar{I} &\approx P_{{\rm LoS},k}\mathbb{E}\left[P_{r,k}({\rm LoS})\right] + P_{{\rm NLoS},k}\mathbb{E}\left[P_{r,k}({\rm NLoS})\right] \\ &= P_t g(\varphi_k)\left[10^{\frac{-\mu_{\rm LoS}}{10}} P_{{\rm LoS},k} + 10^{\frac{-\mu_{\rm NLoS}}{10}} P_{{\rm NLoS},k}\right]\left(\frac{4\pi f_c d_k}{c}\right)^{-n}. \end{aligned}$$

where $\mathbb{E}[.]$ is the expected value over the received interference power. The mean interference

is a reasonable approximation for the interference and leads to a tractable coverage probability expression. In (2.8), (a) is a direct result of (8.1), and (b) comes from the complementary cumulative distribution distribution function (CCDF) of a Gaussian random variable. Furthermore, $r \leq h.\tan(\theta_{\rm B}/2)$ implies that a user can be covered only if it is in the coverage range of a directional antenna with beamwidth θ_B. This proves the theorem. □

Note that, Theorem 1 provides the coverage probability for users located at any arbitrary range r. From Theorem 1, it is observed that changing the UAV's altitude impacts the coverage by affecting several parameters including the distance between the UAV and users, LoS probability, and the feasible coverage radius ($r \leq h.\tan(\frac{\theta_{\rm B}}{2})$). For instance, increasing a UAV's altitude leads to a higher path loss and LoS probability, as well as a higher feasible coverage radius. Clearly, in the presence of interference, the UAVs need to increase their transmit power in order to meet the coverage requirements. Furthermore, as the number of UAVs increases, the distance between the UAVs decreases, and hence, the interference from the nearest UAV will increase. The coverage radius of a UAV, r_u, is the maximum range within which the probability that users are covered by the UAV is greater than a specified threshold (ϵ). Clearly, r_u depends on the transmit power, antenna beamwidth, ϵ, number of UAVs, and UAVs' locations. Thus, the coverage radius is given by:

$$r_u = \max\{r | P_{\rm cov}(r, P_t, \theta_B) \geq \varepsilon\}. \tag{2.9}$$

Now, consider the geographical area of interest which should be covered by multiple UAVs. The UAVs must be placed in a way to maximize the total coverage, and to avoid any overlapping in their coverage areas. Furthermore, while maximizing the total coverage, each UAV must use a minimum transmit power in order to maximize the coverage lifetime. The coverage lifetime is defined as the maximum time that the UAVs can provide coverage for the given area.Now, assuming a circular coverage area for each UAV, the problem can be formulated as follows:

$$\underset{i\in\{1,\ldots,M\}}{(\vec{r}_j^{*}, h^*, r_u^*)} = \arg\max M.r_u^2, \tag{2.10}$$

$$\text{st. } ||\vec{r}_j - \vec{r}_k|| \geq 2r_u, \;\; j \neq k \in \{1, ..., M\}, \tag{2.11}$$

$$||\vec{r}_j + r_u|| \leq R_c, \tag{2.12}$$

$$r_u \leq h.\tan(\theta_{\rm B}/2), \tag{2.13}$$

where M is the number of UAVs, R_c is the radius of the desired geographical area, $\vec{r}_j$ is the vector location of UAV j within the 2D plane of the desired area considering the center of the area as the origin, and r_u is the maximum coverage radius of each UAV. Considering the optimization problem in (2.10), constraint (2.11) ensures that no coverage overlap occurs, and (2.12) guarantees that UAVs do not cover outside of the desired area. As a result, the potential interference between UAVs will be avoided, and also, users located outside the given area (undesired area) will not be affected by UAVs' transmissions.

Solving (2.10) is challenging due to the high number of unknowns and the nonlinear constraints. We model this problem by exploiting the so-called *circle packing problem* [93]. In the circle packing problem, M circles should be arranged inside a given surface such that the packing density

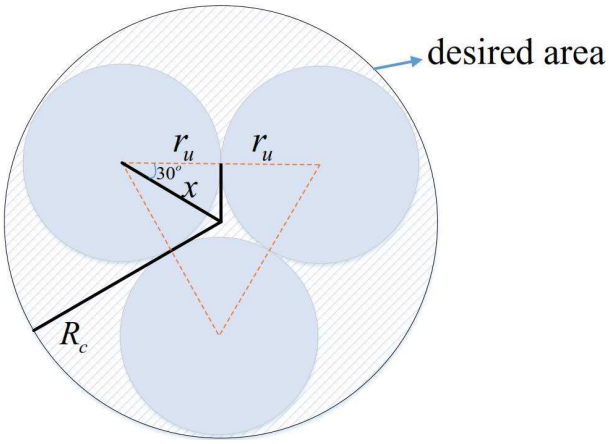


Figure 2.2: Packing problem in a circle with 3 circles.

is maximized and none of the circles overlap. As an illustrative example, Figure 2.2 shows the optimal packing of 3 equal circles inside a bigger circle. Also, Table 2.1 shows the radii of non-overlapping small circles which lead to the maximum packing density for a given circular area [93]. Clearly, the radius of each circle decreases as the number of circles increases. In Table 2.1, the total coverage represents the maximum portion of the desired area which can be covered by multiple circles. Note that, in general, the circle packing problem in a bounded area is known to be intractable [93]. In particular, it is not possible to find a general packing strategy that is optimal for any arbitrary M. Therefore, for each value of M a specific packing strategy needs to be provided. As an example, we derive the optimal packing method for $M = 3$. Consider a circular area with radius R_c. Clearly, the packing density is maximized if the maximum distance between the farthest circles is minimized. For $M = 3$, all the circles tangent with each other if their centers are placed on the vertices of a equilateral triangle. Hence, Figure 2.2 corresponds to the optimal placement. Then, from Figure 2.2, $x = \frac{r_u}{\cos(30^o)}$ and,
$R_c = r_u + x = r_u \left(1 + \frac{2}{\sqrt{3}}\right) \rightarrow r_u = \frac{\sqrt{3}R_c}{2+\sqrt{3}} \approx 0.464 R_c$.

In our model, each circle corresponds to the coverage region of each UAV, and maximizing packing density is related to maximizing the coverage area with non-overlapping smaller circles. Therefore, given the radius of the desired area and the number of symmetric UAVs, we can determine the required coverage radii of UAVs as well as their 3D locations which lead to the maximum coverage. Subsequently, based on the required coverage range of each UAV (r_u), the minimum transmit power of UAVs can be computed. Note that, the UAV's altitude should be adjusted based on the coverage radius and the antenna beamwidth by using $h = \frac{r_u}{\tan(\theta_B/2)}$.

Next, as a function of the number of UAVs, we derive an upper bound for the altitude which guarantees the non-overlapping condition between the UAVs' coverage regions.

Proposition 1. Given M UAVs, and R_c, the radius of the desired area, an upper bound for the maximum UAVs' altitude for which the coverage overlap does not occur, is given by:

Table 2.1: Covering a circular area with radius R_c using identical UAVs- the *circle packing in a circle* approach.

Number of UAVs	Coverage radius of each UAV	Maximum total coverage
1	R_c	1
2	$0.5R_c$	0.5
3	$0.464R_c$	0.646
4	$0.413R_c$	0.686
5	$0.370R_c$	0.685
6	$0.333R_c$	0.666
7	$0.333R_c$	0.778
8	$0.302R_c$	0.733
9	$0.275R_c$	0.689
10	$0.261R_c$	0.687

$$h \leq \frac{q_m R_c}{(2+q_m)\tan(\frac{\theta_B}{2})}, \tag{2.14}$$

where q_m is the maximum value of variable $q \in \mathbb{R}$ that satisfies the following inequality: $\frac{\pi}{\sin^{-1}(q/2)}\left(\frac{q\sqrt{3}+\sqrt{4-q^2}}{q}\right)+\sqrt{12}(1-M)\geq 0$.

Proof. The maximum packing density (D) that can be achieved for a circle, using M equal smaller circles, is upper bounded by [93]:

$$D \leq \frac{M q_m^2}{(2+q_m)^2}. \tag{2.15}$$

Also, clearly, $D = \frac{M r_u^2}{R_c^2}$, and hence, $\frac{M r_u^2}{R_c^2} \leq \frac{M q_m^2}{(2+q_m)^2}$.
Finally, considering $r_u \leq \frac{q_m R_c}{(2+q_m)}$, and $\tan(\frac{\theta_B}{2}) = \frac{r_u}{h}$, we have $h \leq \frac{q_m R_c}{(2+q_m)\tan(\frac{\theta_B}{2})}$. □

Proposition 1, provides a necessary condition on the UAVs' altitude for avoiding an overlapping coverage.

2.4 Simulation Results and Analysis

In our simulations, we consider the UAV-based communications over 2 GHz carrier frequency ($f_c = 2\,\text{GHz}$) in an urban environment with $\alpha = 0.6$, $\gamma = 0.11$, $k_1 = 10.39$, $k_2 = 0.05$, $g_1 = 29.06$, $g_2 = 0.03$, $\mu_{\text{LoS}} = 1\,\text{dB}$, $\mu_{\text{NLoS}} = 20\,\text{dB}$, and $n = 2.5$ [15]. Moreover, using Theorem 1 and (2.9), the coverage radius of each UAV is computed based on $\epsilon = 0.80$, $\beta = 5$, and $N = -120\,\text{dBm}$. Furthermore, here, the maximum coverage time duration of the UAVs (coverage lifetime) is inversely proportional to the transmit power of UAVs.

Figure 2.3 shows the total coverage and the coverage lifetime as a function of the number of UAVs (M) for $R_c = 5000\,\text{m}$, and $\theta_B = 80^o$. In this figure, the maximum achievable coverage while

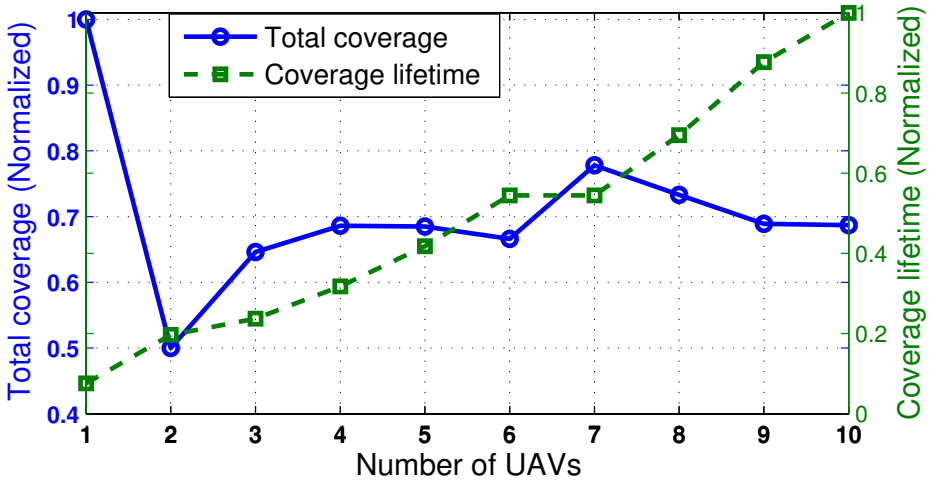


Figure 2.3: Total coverage and coverage lifetime versus number of UAVs for $R_c = 5000\,\text{m}$.

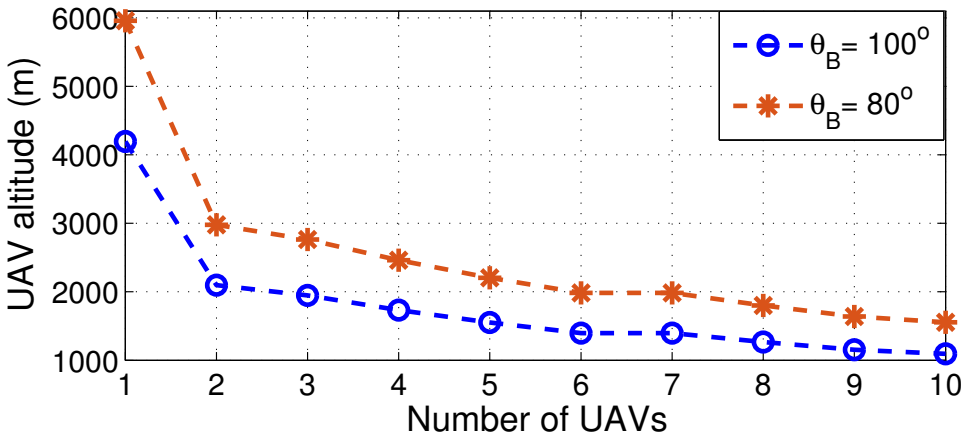


Figure 2.4: UAV's altitude versus number of UAVs.

maximizing the coverage lifetime is shown for different number of UAVs. Clearly, by increasing the number of UAVs, the coverage lifetime increases due to the decrease in the transmit power of each UAV. In Figure 2.3, for the given area with a radius 5000 m, a single UAV has the maximum coverage performance. However, in this case, the single UAV has a minimum coverage lifetime. Therefore, depending on the size of the area, coverage and coverage lifetime requirements, an appropriate number of UAVs needs to be deployed.

Figure 3.11 illustrates the optimal UAVs' altitude versus the number of UAVs. Clearly, the altitude of UAVs should be decreased as the number of UAVs increases. For higher number of UAVs, to avoid overlapping between the coverage regions of the UAVs, the coverage radius of UAVs must be decreased by reducing their height according to $h = \frac{r_u}{\tan(\theta_B/2)}$. As shown, by doubling the number of UAVs from 3 to 6, the optimal altitude is reduced from 2000 m to 1300 m. Furthermore, the optimal altitude is lower for higher antenna beamwidths.

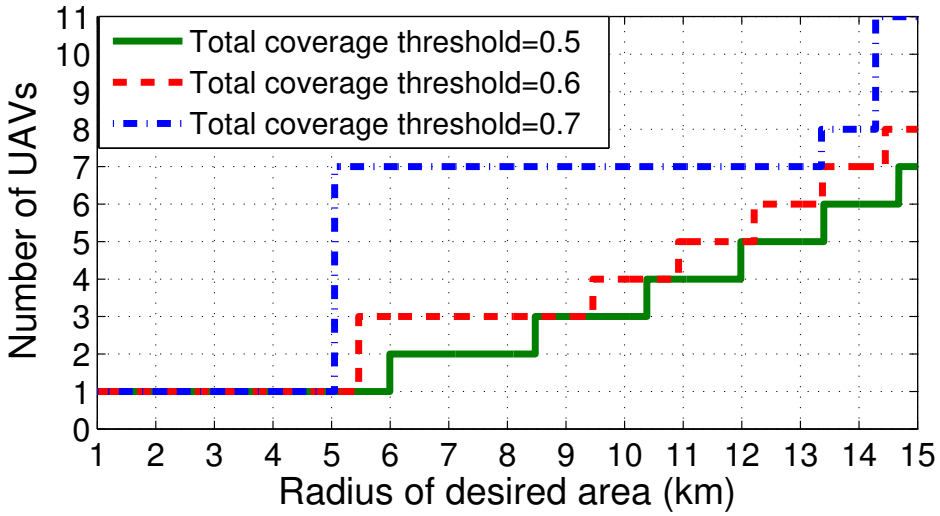


Figure 2.5: Number of required UAVs versus radius of desired area.

Figure 2.5 shows the minimum required number of UAVs in order satisfy the coverage requirement of the given geographical area. In this figure, the coverage threshold corresponds to the minimum portion of the given area which needs to be covered by the UAVs. This result is based on $P_t = 35\,\text{dBm}$, $\theta_B = 80^o$, and optimal altitudes subject to $h < 5000\,\text{m}$. Interestingly, to satisfy at least 0.7 coverage requirement with a maximum coverage lifetime, either one UAV or more than 6 UAVs are required. In other words, for $1 < M < 7$, the 0.7 coverage performance cannot be achieved. In general, as the size of the desired area increases, more UAVs are needed to meet the coverage requirement. Clearly, for $R_c < 5400\,\text{m}$, using a single UAV can satisfy a 0.6 coverage threshold. However, for a larger target area, more UAVs must be used to reach the coverage threshold. Therefore, the optimal number of UAVs for an efficient system design is significantly dependent on the coverage requirement, and the size of geographical area.

2.5 Summary

In this chapter, we have studied the optimal deployment of multiple UAVs equipped with directional antennas used as aerial base stations. First, the downlink coverage probability was derived based on the probabilistic LoS/NLoS links and considering the shadow fading. Next, given a desired geographical area which needs to be covered by multiple UAVs, an efficient deployment approach was proposed based on the circle packing theory that leads to a maximum coverage while each UAV uses a minimum transmit power. The results have shown that, the optimal altitude and location of the UAVs can be determined based on the number of available UAVs and the gain/beamwidth of the directional antennas.

Chapter 3

Mobile Unmanned Aerial Vehicles for Energy-Efficient Internet of Things Communications

3.1 Background, Related Works, and Contributions

As previously discussed in Section 1.1.3, compared to the terrestrial base stations, the advantage of using UAV-based aerial base stations is their ability to provide on-demand communications. Furthermore, the high altitude of UAVs enables them to effectively establish line-of-sight (LoS) communication links thus mitigating signal blockage and shadowing. Due to their adjustable altitude and mobility, UAVs can move towards potential ground users and establish reliable connections with a low transmit power. Hence, they can provide a cost-effective and energy-efficient solution to collect data from ground mobile users that are spread over a geographical area with limited terrestrial infrastructure.

Indeed, UAVs can play a key role in the Internet of Things which is composed of small, battery-limited devices such as sensors, and health monitors [21, 37]. These devices are typically unable to transmit over a long distance due to their energy constraints [21]. In such IoT scenarios, UAVs can dynamically move towards IoT devices, collect the IoT data, and transmit it to other devices which are out of the communication ranges of the transmitters [21]. In this case, the UAVs play the role of moving aggregators or base stations for IoT networks [4]. However, to effectively use UAVs for the IoT, several challenges must be addressed such as optimal deployment, mobility and energy-efficient use of UAVs as outlined in [5] and [7].

In [26], the authors investigated the optimal trajectory of UAVs equipped with multiple antennas for maximizing sum-rate in uplink communications. The work in [79] maximizes the throughput of a relay-based UAV system by jointly optimizing the UAV's trajectory as well as the source/relay transmit power. However these works considered a single UAV in their models. In [7], we inves-

tigated the optimal deployment and movement of a single UAV for supporting downlink wireless communications. However, this work was restricted to a single UAV and focused on the downlink without accounting for IoT devices. The work in [105] provided a comprehensive downlink coverage analysis for a network in which a finite number of UAVs serve the ground users. In [38], the authors used UAVs to efficiently collect data and recharge the clusters' head in a wireless sensor network which is partitioned into multiple clusters. However, this work is limited to a static sensor network, and does not investigate the optimal deployment of the UAVs. While the energy efficiency of uplink data transmission in a machine-to-machine (M2M) communication network was investigated in [106] and [107], the presence of UAVs was not considered. In fact, none of the prior studies in [4, 5, 7, 8, 10, 11, 21, 26, 37, 38, 79, 105–107], addressed the problem of jointly optimizing the deployment and mobility of UAVs, device association, and uplink power control for enabling reliable and energy-efficient communications for IoT devices. In this chapter, we provide a *comprehensive analysis on the joint optimal 3D deployment of aerial base stations, device association, and uplink power control in an IoT ecosystem*.

In this chapter, we introduce a novel framework for optimized deployment and mobility of multiple UAVs for the purpose of energy-efficient uplink data collection from ground IoT devices. In particular, we consider an IoT network in which the IoT devices can be active at different time instances. To minimize the total transmit power of these IoT devices, given device-specific signal-to-interference-plus-noise-ratio (SINR) constraints, we propose an efficient approach to jointly and dynamically find the UAVs' locations, the association of devices to UAVs, and the optimal uplink transmit power. Our proposed framework is composed of two key steps. First, given the locations of the IoT devices, we propose a solution for optimizing the deployment and association of the UAVs. In this case, we solve the formulated problem by decomposing it into two subproblems which are solved iteratively. In the first subproblem, given the fixed UAVs' locations, we find the jointly optimal device-UAV association and the devices' transmit power. In the second subproblem, given the fixed device association, we determine the joint 3D UAVs' locations. For this subproblem, we transform the non-convex continuous location optimization problem to a convex form and provide tractable solutions. Next, following our proposed algorithm, the results of solving the second subproblem are used as inputs to the first subproblem for the next iteration. Here, we show that our proposed approach leads to an efficient solution with a reasonable accuracy compared to the global optimal solution that requires significant overhead. Clearly, the UAVs' locations and the device association that we obtain in this first step will depend on the locations of active IoT devices.

In the second step, we analyze the IoT network over a time period during which the set of active devices changes. In this case, we present a framework for optimizing the UAVs' mobility by allowing them dynamically update their locations depending on the time-varying devices' activation process. First, we derive the closed-form expressions for the time instances (update times) at which the UAVs must move according to the activation process of the devices. Next, using the update time results, we derive the optimal 3D UAVs' trajectory such that the total movement of the UAVs while updating their locations is minimized. Our simulation results show that, using the proposed approach, the total transmit power of the IoT devices can be significantly reduced compared to a

case in which stationary aerial base stations are deployed. The results also verify our analytical derivations for the update times and reveal an inherent tradeoff between the number of updates, the mobility of the UAVs, and transmit power of the IoT devices. In particular, it is shown that a higher number of updates leads to lower transmit powers for the IoT devices at the cost of higher UAVs' energy consumptions.

The rest of this chapter is organized as follows. In Section 3.2, we present the system model and problem formulation. Section 3.3 presents the optimal deployment of UAVs and device association. In Section 3.4, we address the mobility and update time of the UAVs. In Section 3.5 we provide the simulation and analytical results, and Section 3.6 draws some conclusions.

3.2 System Model and Problem Formulation

Consider an IoT system consisting of a set $\mathcal{L} = \{1, 2, ..., L\}$ of L IoT devices. Examples of such devices include various types of sensors used for environmental monitoring, smart traffic control, and smart parking devices. In this system, a set $\mathcal{K} = \{1, 2, ..., K\}$ of K rotary wing UAVs must be deployed to collect the data from the ground IoT devices. These UAVs can dynamically move, when needed, to effectively serve the IoT devices using uplink communication links. Here, the term *served* by a UAV implies that the uplink SINR is above the threshold and, thus, the UAV can successfully collect data from the ground IoT device. In our model, we assume that the devices transmit their data to the UAVs in the uplink using frequency division multiple access (FDMA) over R orthogonal channels. Let $E_{\max}$ be the maximum energy that each UAV can spend on its movement. The locations of device $i \in \mathcal{L}$ and UAV $j \in \mathcal{K}$ are, respectively, given by (x_i, y_i) and $\boldsymbol{v}_j = (x_j^{\text{uav}}, y_j^{\text{uav}}, h_j)$ as shown in Fig. 3.1. In our model, we consider a centralized network in which the locations of the devices and UAVs are known to a control center located at a central *cloud* server. The cloud server will determine the UAVs' locations, device association, and the transmit power of each IoT device.

We analyze the IoT network within a time interval $[0, T]$ during which the IoT devices can be active at different time instances and must be served by the UAVs at some pre-defined time slots. At the beginning of each slot, the positions of the UAVs as well as the device-UAV association are updated based on the locations of currently active devices that are assumed to be known to the cloud center[1]. Hereinafter, the time instance at which the UAVs' locations and associations are *jointly* updated, is referred to as the *update time*. The update times are denoted by t_n, $1 \leq n \leq N$, with N being the number of updates. At each update time t_n, based on the location of active devices, the optimal UAVs' locations and the corresponding association must be determined for effectively serving the ground devices. Here, the IoT devices that become active during $[t_{n-1}, t_n)$ are served by the UAVs during the time period $[t_n, t_{n+1})$. Note that, during $[t_{n-1}, t_n)$, the UAVs' locations and their device association do not change until the next update time, t_n. Clearly, since at different update times, a different subset of devices might be active, the locations of the UAVs

[1]We consider static IoT devices in delay-tolerant applications which their fixed locations and activation patterns are known to the cloud center.

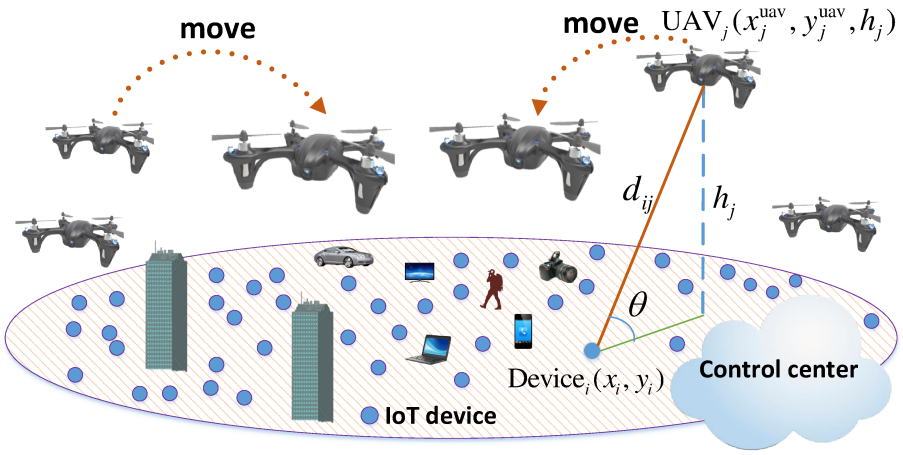


Figure 3.1: System model.

must dynamically change at each update time. Therefore, each UAV's *trajectory* will consist of N stop locations at which the UAV serves the ground devices. Note that, in our model, the UAVs' locations are not necessarily updated once the set of active devices changes. Instead, we consider some specific time instances (update times) at which the UAVs locations device associations, and devices' transmit power are optimized. In particular, considering the fact that the set of active devices may continuously change, continuously updating the UAVs' locations, the devices transmit powers, and the device-UAV associations may not be feasible as it can lead to low reliability, high UAVs' energy consumption, and a need to solve complex real-time optimization processes. In our model, the update times are design parameters that depend on the activity of the devices, and the energy of UAVs. Given this model, our objective is to find the optimal joint UAVs' locations and device association at each update time t_n so as to minimize the total transmit power of the active devices while meeting each device's SINR requirement. Moreover, we need to develop a framework for determining the update times as well as the UAVs' mobility to handle dynamic changes in the activity of the devices. To this end, first, we present the ground-to-air channel model and the activation models for the IoT devices.

3.2.1 Ground-to-Air Path Loss Model

In our model, while optimizing the locations of the UAVs, the information available includes the ground devices' locations, and the type of environment (e.g. rural, suburban, urban, highrise urban, etc.). Note that, in such practical scenarios, one will not have any additional information about the exact locations, heights, and number of the obstacles. Therefore, one must consider the randomness associated with the LoS and NLoS links while designing the UAV-based communication system. Therefore, for ground-to-air communications, each device will typically have a LoS view towards a specific UAV with a given probability. This LoS probability depends on the environment, location of the device and the UAV as well as the elevation angle [8]. One suitable expression for the LoS

probability is given by [5, 7, 8]:

$$P_{\text{LoS}}^{ij} = \frac{1}{1 + \psi \exp(-\beta \left[\theta_{ij} - \psi\right])}, \tag{3.1}$$

where ψ and β are constant values which depend on the carrier frequency and type of environment such as rural, urban, or dense urban, and θ_{ij} is the elevation angle. Clearly, $\theta = \frac{180}{\pi} \times \sin^{-1}\left(\frac{h_j}{d_{ij}}\right)$, where $d_{ij} = \sqrt{(x_i - x_j^{\text{uav}})^2 + (y_i - y_j^{\text{uav}})^2 + h_j^2}$ is the distance between device i and UAV j.

From (8.2), we can see that by increasing the elevation angle or increasing the UAV altitude, the LoS probability increases. The path loss model for LoS and non-line-of-sight (NLoS) links between device i and UAV j is given by [5] and [8]:

$$L_{ij} = \begin{cases} \eta_1 \left(\frac{4\pi f_c d_{ij}}{c}\right)^{\alpha}, & \text{LoS link}, \\ \eta_2 \left(\frac{4\pi f_c d_{ij}}{c}\right)^{\alpha}, & \text{NLoS link}, \end{cases} \tag{3.2}$$

where f_c is the carrier frequency, α is the path loss exponent, η_1 and η_2 ($\eta_2 > \eta_1 > 1$) are the excessive path loss coefficients in LoS and NLoS cases, and c is the speed of light. Note that, the NLoS probability is $P_{\text{NLoS}}^{ij} = 1 - P_{\text{LoS}}^{ij}$. Typically, given only the locations of the UAVs and devices, it is not possible to exactly determine which path loss type (LoS/NLoS) is experienced by the device-UAV link. In this case, the path loss average considering both LoS and NLoS links can be used for the device-UAV communications [5] and [8]. Now, using (8.2) and (3.2), the average path loss between device i and UAV j can be expressed as:

$$\bar{L}_{ij} = P_{\text{LoS}}^{ij} \eta_1 \left(\frac{4\pi f_c d_{ij}}{c}\right)^{\alpha} + P_{\text{NLoS}}^{ij} \eta_2 \left(\frac{4\pi f_c d_{ij}}{c}\right)^{\alpha} = \left[P_{\text{LoS}}^{ij} \eta_1 + P_{\text{NLoS}}^{ij} \eta_2\right] \left(K_o d_{ij}\right)^{\alpha}, \tag{3.3}$$

where $K_o = \frac{4\pi f_c}{c}$. Clearly, the average channel gain between the UAV and the device is $\bar{g}_{ij} = \frac{1}{\bar{L}_{ij}}$. Note that, by using the average channel gain, there is no need to account for LoS and NLoS links separately, and, hence, the SINR expressions become more tractable. Therefore, we use the average channel gain to model the interference and desired links for all device-UAV communications while computing the SINRs.

3.2.2 IoT Device Activation Model

Indeed, the activation of IoT devices depends on the services that they are supporting. For instance, in some applications such as weather monitoring, smart grids, or home automation, the IoT devices need to report their data periodically. However, the IoT devices can have random activations in health monitoring, or smart traffic control applications. Therefore, the UAVs must be properly deployed to collect the IoT devices data while dynamically adapting to the activity patterns of these devices. Naturally, the optimal locations of the UAVs and their update times depend on the activation process of the IoT devices. Here, we consider two activation models. In the first model,

the IoT devices are randomly activated, as in smart traffic control applications. In this case, the concurrent transmissions of a massive number of devices within a short time duration can lead to a bursty traffic as pointed out in [108] and [109]. In fact, when massive IoT devices attempt to transmit within a short time period, the arrival patterns become more bursty [110]. Thus, 3GPP suggests a *beta* distribution to capture this traffic characteristic of IoT devices [111]. In this case, each IoT device will be active at time $t \in [0, T]$ following the beta distribution with parameters κ and ω [109–111]:

$$f(t) = \frac{t^{\kappa-1}(T-t)^{\omega-1}}{T^{\kappa+\omega-1}B(\kappa,\omega)}, \tag{3.4}$$

where $[0, T]$ is the time interval within which the IoT devices can be active, and $B(\kappa, \omega) = \int_0^1 t^{\kappa-1}(1-t)^{\omega-1}\mathrm{d}t$ is the beta function with parameters κ and ω [112].

In addition, IoT devices such as smart meters typically report their data periodically rather than randomly. For such devices, the activation process is deterministic and assumed to be known in advance. In such case, we assume that device i becomes active each τ_i seconds during $[0, T]$ time duration. Clearly, the number of activations for a device i during $[0, T]$ is $\left\lfloor \frac{T}{\tau_i} \right\rfloor$.

3.2.3 Channel Assignment Strategy

Here, given only the devices' locations, a practical channel assignment approach is to assign different channels to devices which are located in proximity of each other. This approach significantly mitigates the possibility of having strong interference between two closely located devices. For the channel assignment problem, we have adopted a constrained *K-mean clustering* strategy [113], which is an efficient distance-based clustering approach in which a set of given points are grouped into K clusters based on their proximity. In this case, given the number of active devices, L_n, and the number of orthogonal channels, $R \leq L_n$, we group the devices based on proximity, and assign different channels to devices which are in the same group.

Now, we present our optimization problem to find the UAVs' locations, device association, and transmit power of IoT devices at each update time t_n during $[0, T]$:

(OP):

$$\min_{\boldsymbol{v}_j, \boldsymbol{c}, \boldsymbol{P}} \sum_{i=1}^{L_n} P_i\,, \quad \forall i \in \mathcal{L}_n,\ \forall j \in \mathcal{K}, \tag{3.5}$$

$$\text{s.t. } \frac{P_i \bar{g}_{ic_i}(\boldsymbol{v}_{c_i})}{\sum\limits_{k \in \mathcal{Z}_i} P_k \bar{g}_{kc_i}(\boldsymbol{v}_{c_i}) + \sigma^2} \geq \gamma, \tag{3.6}$$

$$0 < P_i \leq P_{\max}, \tag{3.7}$$

where L_n is the total number of active devices at update time t_n, and $\mathcal{L}_n$ is the set of devices' index. $\boldsymbol{P}$ is the transmit power vector with each element P_i being the transmit power of device i. Also, $\boldsymbol{v}_j$ is the 3D location of UAV j, and $\boldsymbol{c}$ is the device association vector with each element c_i being the index of the UAV that is assigned to device i. $P_{\max}$ is the maximum transmit power

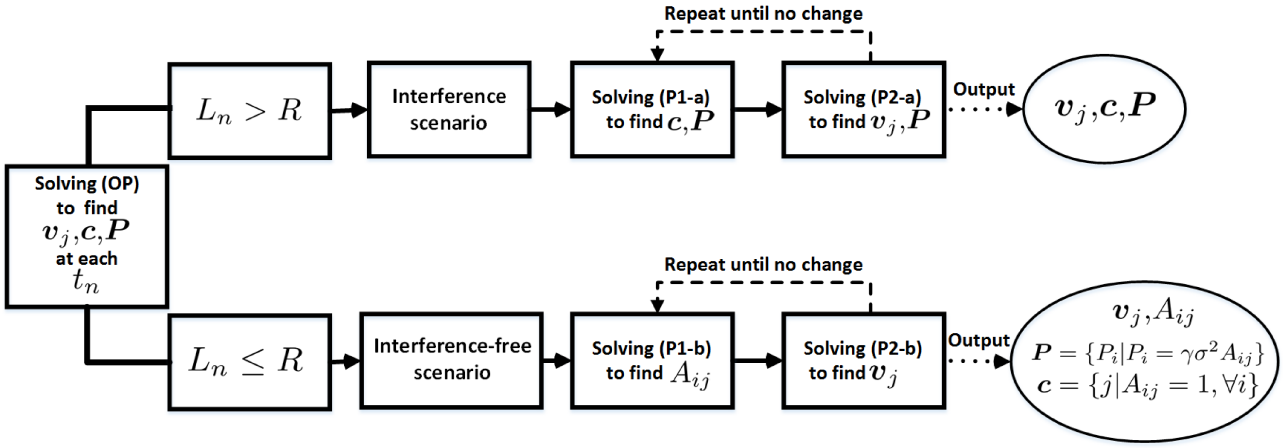


Figure 3.2: Block diagram for the proposed solution.

of each IoT device, and σ^2 is the noise power. Furthermore, $\bar{g}_{ic_i}(\boldsymbol{v}_{c_i})$ is the average channel gain between device i and UAV c_i which is a function of the UAV's location. Also, $\bar{g}_{kc_i}(\boldsymbol{v}_{c_i})$ is the average channel gain between interfering device k and UAV c_i. In (3.6), $\mathcal{Z}_i$ is the set of all other devices that use the same channel as device i and create interference. γ is the SINR target which must be achieved by all the devices, (3.6) represents the SINR requirement, and (3.7) shows the maximum transmit power constraint. Hereinafter, we call **(OP)** the *original problem*.

Note that, in (3.5), the transmit power of the IoT devices, the 3D locations of the UAVs, and the UAV-device associations are unknowns. Clearly, the locations of the UAVs impact the channel gain between the devices and UAVs, and, hence, they affect the transmit power of each device, P_i. Furthermore, given (3.6), due to the mutual interference between the devices, the transmit power of each device depends also on the transmit power of the interfering devices as well as the device-UAV associations. In addition, the device-UAV associations depend on the UAVs' locations which are also unknowns. Therefore, there is a mutual dependency between all the optimization variables in **(OP)**. Moreover, considering (8.2) and constraint (3.6), we can see that, this optimization problem is highly non-linear and non-convex. Indeed, solving (3.5) is significantly challenging due to the mutual dependency of the optimization variables, non-linearity, and non-convexity of the problem. Next, we propose a framework for solving this optimization problem.

In essence, our proposed framework for solving **(OP)** proceeds as follows. At each update time t_n, given the fixed UAVs' locations, we find the optimal device-UAV association and the transmit power of the devices. Next, given the fixed UAV association from the previous step, we determine the optimal locations of the UAVs and update the transmit power the devices accordingly. This procedure is done iteratively until the 3D UAVs' locations, device association, and the transmit power of the devices are found. Clearly, at each step, the total transmit power of the devices decreases, and, hence, the proposed algorithm converges. Fig. 3.2 shows a block diagram that summarizes the main steps for solving **(OP)**. Next, we discuss, in detail, each block of the proposed solution in Fig. 3.2.

3.3 UAV Deployment and Device Association with Power control

Here, given the locations of active IoT devices, we minimize the total transmit power of the devices by solving (3.5). Clearly, the UAVs' locations and the device association are mutually dependent. In particular, to find the device association, the locations of the UAVs must be known. Moreover, the UAVs' locations cannot be optimized without knowing the device association.

Therefore, we decompose **(OP)** into two subproblems that will be solved iteratively. In the first subproblem, given the locations of the UAVs, we find the optimal device association and the transmit power of the devices such that the uplink SINR requirements of all active devices are satisfied with a minimum total transmit power. In the second subproblem, given the device association resulting from the first subproblem, we determine the optimal locations of the UAVs for which the transmit power of the devices is minimized. Note that, this is an iterative process in which the results of each subproblem are used in the other subproblem for the next iteration. These computations are performed by the control center until the 3D UAVs' locations, device association, and transmit power of the devices are obtained.

Note that, given the limited number of available orthogonal channels, the interference between the devices will depend on the number of active devices at each update time. Clearly, there is no interference when the number of active devices at time t_n is less than the number of orthogonal channels, or equivalently $L_n \leq R$. Given that, in the interference-free scenario, one can provide a more tractable analysis, here, for the deployment and association steps. Therefore, we will investigate the interference and interference-free scenarios, separately.

3.3.1 Device Association and Power Control

Here, given initial locations of the UAVs, we aim to find the optimal device association as well as the transmit power of each IoT device such that the total transmit power used for successful uplink communications is minimized.

Interference scenario

In the presence of uplink interference when $L_n > R$, the power minimization problem at update time t_n will be given by:

(P1-a):

$$\min_{\boldsymbol{c},\boldsymbol{P}} \sum_{i=1}^{L_n} P_i\,, \quad \forall i \in \mathcal{L}_n,\ \forall j \in \mathcal{K}, \tag{3.8}$$

$$\text{s.t. } \frac{P_i \bar{g}_{ic_i}}{\sum\limits_{k \in \mathcal{Z}_i} P_k \bar{g}_{kc_i} + \sigma^2} \geq \gamma, \tag{3.9}$$

$$0 < P_i \leq P_{\max}. \tag{3.10}$$

To solve **(P1-a)**, we need to jointly find the optimal device association and the transmit power of all active devices under the SINR constraints for the given UAVs' locations. Clearly, given the fixed UAVs' locations, optimization variables are the device association and the transmit power of the devices. Note that, satisfying the SINR requirement of each device significantly depends on the distance and altitude of its serving UAV. Therefore, the feasibility of the optimization problem in (3.8) depends on the locations of the UAVs. Next, we derive an upper bound and a lower bound for the altitude of serving UAV j as a function of its distance from device i.

Proposition 2. The lower and upper bounds for the altitude of a UAV j needed to serve a device i (meeting its SINR requirement), are given by:

$$d_{ij} \sin\left(\frac{1}{\beta} \ln\left(\frac{\psi Q}{1-Q}\right) + \psi\right) \leq h_j \leq \left(\frac{P_{\max}}{\gamma K_o^{\alpha} \sigma^2 \eta_1}\right)^{1/\alpha}, \tag{3.11}$$

where d_{ij} is the distance between UAV j and device i, and $Q = \frac{P_{\max}}{\gamma d_{ij}^{\alpha} {K_o}^{\alpha} \sigma^2 (\eta_1 - \eta_2)} - \frac{\eta_2}{\eta_1 - \eta_2}$.

Proof. Let I_i be the cumulative interference from interfering devices on device i, then:

$$\begin{aligned}
&\text{SINR}_i = \frac{P_i \bar{g}_{ij}}{I_i + \sigma^2} \geq \gamma, \\
&d_{ij}^{\alpha} \leq \frac{P_i}{\gamma K_o^{\alpha} \left(I_i + \sigma^2\right) \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 P_{\text{NLoS}}^{ij}\right)} \leq \frac{P_{\max}}{\gamma K_o^{\alpha} \sigma^2 \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 (1 - P_{\text{LoS}}^{ij})\right)}, \\
&P_{\text{LoS}}^{ij} \geq \frac{P_{\max}}{\gamma d_{ij}^{\alpha} K_o^{\alpha} \sigma^2 \left(\eta_1 - \eta_2\right)} - \frac{\eta_2}{\eta_1 - \eta_2}, \\
&\text{considering } Q = \frac{P_{\max}}{\gamma d_{ij}^{\alpha} K_o^{\alpha} \sigma^2 \left(\eta_1 - \eta_2\right)} - \frac{\eta_2}{\eta_1 - \eta_2}, \text{and using equation (8.2)}, \\
&\theta_{ij} \overset{(a)}{\geq} \frac{1}{\beta} \ln\left(\frac{\psi Q}{1-Q}\right) + \psi, \\
&h_j \geq d_{ij} \sin\left(\frac{1}{\beta} \ln\left(\frac{\psi Q}{1-Q}\right) + \psi\right),
\end{aligned} \tag{3.12}$$

where (a) stems from (8.2). Also, we have:

$$d_{ij}^{\alpha} \leq \frac{P_{\max}}{\gamma K_o^{\alpha} \sigma^2 \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 (1 - P_{\text{LoS}}^{ij})\right)} \overset{(b)}{\leq} \frac{P_{\max}}{\gamma K_o^{\alpha} \sigma^2 \eta_1}, \tag{3.13}$$

where in (b) we consider $P_{\text{LoS}} = 1$ which is equivalent to $h_j = d_{ij}$. Finally,

$$h_j \leq \left(\frac{P_{\max}}{\gamma K_o^{\alpha} \sigma^2 \eta_1}\right)^{1/\alpha}. \tag{3.14}$$

Clearly, (3.12) and (3.14) prove the proposition. □

Proposition 2 provides the necessary conditions for the UAV's altitude needed in order to be able to serve the IoT device. From (3.11), the minimum altitude must increase as the distance increases. In other words, the UAV's altitude needs to be adjusted based on the distance such that the elevation angle between the device and the UAV exceeds $\frac{1}{\beta}\ln\left(\frac{\psi Q}{1-Q}\right)+\psi$. Furthermore, as expected, the maximum altitude of the UAVs significantly depends on the maximum transmit power of the devices as given in (3.14).

Now, given the fixed UAVs' locations, problem **(P1-a)** corresponds to the problem of joint user association and uplink power control in the terrestrial base station scenario. The algorithm presented in [114] and [115] leads to the global optimal solution to the joint user association and uplink power control under the SINR and maximum transmit power constraints. As a result, the optimal transmit power of users and the base station association for which the total uplink transmit power is globally minimized, is determined. In problem **(P1-a)**, the IoT devices correspond to the users, and fixed positioned UAVs correspond to the terrestrial base stations. For our case, this algorithm, as given in Algorithm 1, will proceed as follows. We start with an initial value for transmit power of all active devices in step 3. Then, in step 6 we compute $\rho_{ij}^{(t)}$ at iteration t. In this case, $\rho_{ij}^{(t)}$ represents the minimum required transmit power of device i to reach an SINR of 1 while connecting to UAV j, given the fixed transmit power of other devices. In step 7, we find the minimum transmit power of device i if it connects to the best UAV. Then, the index of the best UAV which is assigned to device i is given in step 8. In step 7 we update the transmit power of device i in order to achieve an SINR of γ. Steps 6 to 7 must be repeated for all devices to obtain the optimal transmit power and the device association vectors.

Algorithm 1 Iterative algorithm for joint power control and device-UAV association

1: **Inputs:** Locations of UAVs and IoT devices
2: **Outputs:** Device association vector ($\boldsymbol{c}$), and transmit power of all devices ($\boldsymbol{P}$).
3: Set $t=0$, and initialize $\boldsymbol{P}^{(0)}=\left(P_1^{(0)},...,P_K^{(0)}\right)$.
4: Define $\rho_{ij}^{(t)}=\frac{\sigma^2+\sum\limits_{k\in\mathcal{Z}_i}P_k^{(t)}\bar{g}_{kj}}{\bar{g}_{ij}}$.
5: Compute $S_i(\boldsymbol{P}^{(t)})=\min\limits_{j\in\mathcal{K}}\rho_{ij}^{(t)}$.
6: Find $c_i(\boldsymbol{P}^{(t)})=\arg\min\limits_{j\in\mathcal{K}}\rho_{ij}^{(t)}$.
7: Update $P_i^{(t+1)}=\min\left\{\gamma S_i(\boldsymbol{P}^{(t)}),P_{\max}\right\},\forall i\in\mathcal{L}_n$.
8: Repeat steps 6 to 7 for all devices until $\boldsymbol{P}^{(t)}$ converges.
9: $\boldsymbol{P}=\boldsymbol{P}^{(t)}$, $\boldsymbol{c}=\left[c_i(\boldsymbol{P}^{(t)})\right],\forall i\in\mathcal{L}_n$.

As shown in [114], after several iterations this algorithm quickly converges to the global optimal solution if the SINR of each device is equal to γ. Hence, by solving **(P1-a)**, we are able to find the optimal transmit power of the devices and the device association for any given fixed locations of the UAVs. Then, the device association and transmit power of the devices will be used as inputs for solving the second subproblem in which the UAVs' locations need to be optimized. Next, we

investigate the device association and power control problem in an interference-free scenario.

Interference-free scenario

At each update time t_n, if the number of active devices is lower than the number of orthogonal channels or equivalently $L_n \leq R$, there will be no interference between the devices. Unlike in the interference scenario, here, the transmit power of each device can be computed only based on the channel gain between the device and its serving UAV. Therefore, considering (3.3), and (3.6) without interference, the minimum transmit power of device i in order to connect to UAV j is $P_i = \gamma\sigma^2\bar{L}_{ij}$. In this case, given the locations of the UAVs (fixed for all $\boldsymbol{v}_j$), $\bar{L}_{ij}$ is known for all devices and problem **(P1-a)** can be simplified. Hence, the optimal association problem under minimum power in the interference-free scenario will be:

(P1-b):

$$\min_{A_{ij}} \sum_{j=1}^{K} \sum_{i=1}^{L_n} A_{ij}\bar{L}_{ij}, \tag{3.15}$$

$$\text{s.t.} \sum_{j=1}^{K} A_{ij} = 1, \ \ \forall i \in \mathcal{L}_n, \tag{3.16}$$

$$A_{ij}\bar{L}_{ij} \leq \frac{P_{\max}}{\gamma\sigma^2}, \ A_{ij} \in \{0,1\}, \ \forall i \in \mathcal{L}_n, j \in \mathcal{K}, \tag{3.17}$$

where $\bar{L}_{ij}$ is the average path loss between device i and UAV j, which is known, give the locations of the UAV and the device. A_{ij} is equal to 1 if device i is assigned to UAV j, otherwise A_{ij} will be equal to 0. Clearly, the optimization problem in (3.15) is an integer linear programming (ILP). In general, this problem can be solved by using standard ILP solution methods such as the cutting plane. However, these solutions might not be efficient as the size of the problem grows. In particular, due to the potentially high number of IoT devices, a more efficient technique for solving (3.15) is needed. Here, we transform problem (3.15) to a standard assignment problem [116] which can be solved in polynomial time. In the assignment problem, the objective is to find the optimal one-to-one assignment between two sets of nodes with a minimum cost. In our problem, the devices and the UAVs can be considered as two sets of nodes that need to be assigned to each other with an assignment cost of L_{ij} between nodes i and j. However, compared to the classical assignment problem, **(P1-b)** has an additional constraint in (3.17) which results from the transmit maximum power constraint. This constraint indicates that device i cannot be assigned to UAV j if $\bar{L}_{ij} > \frac{P_{\max}}{\gamma\sigma^2}$. Therefore, in the assignment problem we can consider $L_{ij} = +\infty$ to avoid assigning device i to UAV j when $\bar{L}_{ij} > \frac{P_{\max}}{\gamma\sigma^2}$ that implies the constraint in (3.17) is violated. Subsequently, using the updated assignment costs, L_{ij}, problem **(P1-b)** will be transformed to the classical assignment problem which can be solved using the Hungarian method with a time complexity of $O((L_nK)^3)$ [117]. We note that, in absence of interference, problems **(P1-a)** and **(P1-b)** have the same solution. Next, we present the second subproblem of the original optimization problem (3.5) in order to optimize the UAVs' locations.

3.3.2 Optimal Locations of the UAVs

In this section, given the optimal device association, our goal is to find the optimal locations of the UAVs for which the total transmit power of the devices is minimized. In other words, considering the mobile nature of the UAVs, we intelligently update the location of each UAV based on the location of its associated IoT devices.

Interference scenario

In this scenario, given the UAV-device associations, the optimization problem to find the 3D locations of the UAVs and the transmit power of the devices will be:

(P2-a):

$$\min_{\boldsymbol{v}_j, \boldsymbol{P}} \sum_{i=1}^{L_n} P_i\,, \quad \forall i \in \mathcal{L}_n,\ \forall j \in \mathcal{K}, \tag{3.18}$$

$$\text{s.t. } \frac{P_i \bar{g}_{ij}(\boldsymbol{v}_j)}{\sum\limits_{k \in \mathcal{Z}_i} P_k \bar{g}_{kj}(\boldsymbol{v}_j) + \sigma^2} \geq \gamma, \tag{3.19}$$

$$0 < P_i \leq P_{\max}, \tag{3.20}$$

where $\boldsymbol{v_j} = (x_j^{\text{uav}}, y_j^{\text{uav}}, h_j)$ indicates the 3D location of UAV j. Clearly, the channel gains used in (3.19) depend on the locations of the UAVs. Note that, according to (8.2) and (3.3), $\bar{g}_{ij}(\boldsymbol{v}_j)$ is a non-convex function of $\boldsymbol{v}_j$. Consequently, constraint (3.19) is also non-linear and non-convex. Furthermore, the transmit power of the devices and the UAVs' locations are mutually dependent. On the one hand, the location of each UAV must be determined such that its associated devices can connect to it with a minimum transmit power. On the other hand, the UAV's location will impact the amount of interference received from other interfering devices. Indeed, solving the optimization problem in **(P2-a)** is challenging as the problem is highly non-linear and non-convex. In particular, the complexity of this problem stems from the mutual dependence between the transmit power of the devices and the locations of the UAVs.

Our proposed approach to solve **(P2-a)** is based on optimizing the location of each UAV separately. Note that, using the results of **(P1-a)**, for each UAV, the associated and non-associated devices and their transmit power, P_i^*, are known. Our proposed solution proceeds as follows. The cloud starts by considering a single UAV and then optimizing its location given the fixed transmit power for the non-associated devices. Then, the cloud updates the transmit power of the associated devices according to the new location of their serving UAV. Hence, at each step, the location of a UAV and the transmit power of its associated devices are updated. At each iteration, after finding P_i^*, we set $P_{\max} = P_i^*$ for the next iteration. This ensures that the transmit power of the devices does not increase during the iterative process. The entire process is repeated by the cloud for all UAVs one-by-one, until the transmit power of the devices cannot be further reduced by changing the UAVs' locations. Note that, at each step, one must determine the optimal location of each UAV such that the total transmit power of its associated devices is minimized.

Now, let $\mathcal{C}_j$ be the set of devices' index associated to UAV j. Given (3.3), (3.18), and (3.19), the optimal location of UAV j can be determined by solving the following problem:

$$\min_{\boldsymbol{v_j}} \sum_{i \in \mathcal{C}_j} F_i(\boldsymbol{v}_j), \tag{3.21}$$

$$\text{s.t.} \quad F_i(\boldsymbol{v}_j) = \gamma \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 P_{\text{NLoS}}^{ij}\right) (K_o d_{ij})^{\alpha} \left[\sum_{k \in Z_i} \frac{P_k}{\left(\eta_1 P_{\text{LoS}}^{kj} + \eta_2 P_{\text{NLoS}}^{kj}\right) (K_o d_{kj})^{\alpha}} + \sigma^2\right], \tag{3.22}$$

$$F_i(\boldsymbol{v}_j) \leq P_i^*, \ \ \forall i \in \mathcal{C}_j, \tag{3.23}$$

Note that, P_{LoS}^{ij}, P_{LoS}^{kj}, d_{kj}, and d_{ij} depend on the locations of UAVs ($\boldsymbol{v_j}$). Also, (3.23) guarantees that the transmit power of each device is reduced by updating the location of serving UAV.

Clearly, (3.21), (3.22), and (3.23) are non-linear and non-convex. Considering the fact that the objective function and constraints are twice differentiable, we convert (3.21) to a quadratic form which can be solved using efficient techniques. In particular, we adopt the sequential quadratic programming (SQP) method as one of the most powerful algorithms for solving large scale and constrained differentiable non-linear optimization problems [118]. Clearly, considering the high non-linearity of (3.22) as well as the large number of constraints, the SQP is a suitable method for solving our optimization problem. In the SQP method, the objective function is approximated by a quadratic function, and the constraints are linearized. Subsequently, the optimization problem is solved by solving multiple quadratic subproblems. In our optimization problem, to find the optimal location of UAV j, $\boldsymbol{v}_{j,k}$, we start with an initial point $\boldsymbol{v}_{j,k}$ (starting with $k = 0$). Then, we use the first order necessary optimality or Karush-Kuhn-Tucker (KKT) conditions to find the Lagrangian variables. In particular, we use:

$$\nabla L(\boldsymbol{v}_{j,k}, \boldsymbol{\lambda}_k) = \sum_{i \in \mathcal{C}_j} \nabla F_i(\boldsymbol{v}_{j,k}) + \nabla \boldsymbol{w}_i(\boldsymbol{v}_{j,k}) \boldsymbol{\lambda}_k = 0, \tag{3.24}$$

where $L(\boldsymbol{v}_{j,k}, \boldsymbol{\lambda}_k) = \sum_{i \in \mathcal{C}_j} F_i(\boldsymbol{v}_{j,k}) + \boldsymbol{\lambda}^T \boldsymbol{w}(\boldsymbol{v}_{j,k})$ is the Lagrangian function, $\boldsymbol{\lambda}_k$ is the vector of Lagrangian variables, and $\boldsymbol{w}(\boldsymbol{v}_{j,k})$ is a vector of functions with each element being $\boldsymbol{w}_i(\boldsymbol{v}_{j,k}) = (F_i(\boldsymbol{v}_{j,k}) - P_i^*)$. Then, given $\boldsymbol{v}_{j,k}$, we determine the Lagrange variables by [118]:

$$\boldsymbol{\lambda}_k = - \left[\boldsymbol{w}_i(\boldsymbol{v}_{j,k})^T \nabla \boldsymbol{w}_i(\boldsymbol{v}_{j,k})\right]^{-1} \nabla \boldsymbol{w}_i(\boldsymbol{v}_{j,k})^T \sum_{i \in \mathcal{C}_j} \nabla F_i(\boldsymbol{v}_{j,k}). \tag{3.25}$$

In the next step, we update $\boldsymbol{v}_{j,k+1} = \boldsymbol{v}_{j,k} + \boldsymbol{d}_k$, where $\boldsymbol{d}_k$ is the solution to the following quadratic programming problem:

$$\boldsymbol{d}_k = \arg\min_{\boldsymbol{d}} \sum_{i \in \mathcal{C}_j} \nabla F_i(\boldsymbol{v}_{j,k})^T \boldsymbol{d} + \frac{1}{2} \boldsymbol{d}^{\boldsymbol{T}} \nabla^2 \left[L(\boldsymbol{v}_{j,k}, \boldsymbol{\lambda}_k)\right] \boldsymbol{d}, \tag{3.26}$$

$$\text{s.t. } F_i(\boldsymbol{v}_{j,k}) + \nabla F_i(\boldsymbol{v}_{j,k})^T \boldsymbol{d} - P_i^* \leq 0, \quad \forall i \in \mathcal{C}_j, \tag{3.27}$$

where, ∇ and ∇^2 indicate the gradient and Hessian operations. Clearly, (3.26) is an inequality constrained quadratic programming. Moreover, it can be shown that the Hessian matrix, $\nabla^2 \left[L(\boldsymbol{v}, \boldsymbol{\lambda}_k)\right]$, is not positive semidefinite, and, hence, (3.26) is non-convex in general. In this case, the two possible solution approaches are the active set, and the interior point methods. Typically, the active set method is preferred when the Hessian matrix is moderate/small and dense. The interior point, however, is a suitable approach when the Hessian matrix is large and sparse [119]. In our problem, due to the potential possible high number of active devices, the number of constraints can be high. Therefore, the Hessian matrix, $\nabla^2 \left[L(\boldsymbol{v}_{j,k}, \boldsymbol{\lambda}_k)\right]$, is large and sparse, and, hence, the interior point method is used.

Finally, based on (3.24)-(3.27), the optimal location of each UAV ($\boldsymbol{v}_j$), given the fixed device association, will be determined. Next, we address the UAVs' location optimization in an interference-free scenario.

Interference-free scenario

In the absence of interference, we are able to provide tractable analysis on the UAVs' locations optimization. Considering $\alpha = 2$ for LoS ground-to-air propagation [8], the optimal locations of the UAVs will be given by:

(P2-b):

$$\min_{\boldsymbol{v_j}} \sum_{i \in \mathcal{C}_j} K_o^2 \sigma^2 \gamma \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 P_{\text{NLoS}}^{ij}\right) {d_{ij}}^2, \tag{3.28}$$

$$\text{s.t. } \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 P_{\text{NLoS}}^{ij}\right) d_{ij}^2 \leq \frac{P_{\max}}{K_o^2 \sigma^2 \gamma}, \ \ \forall i \in \mathcal{C}_j. \tag{3.29}$$

This optimization problem is non-convex over $\boldsymbol{v_j} = (x_j^{\text{uav}}, y_j^{\text{uav}}, h_j)$. However, given any altitude h_j, we can provide a tractable solution to this problem. First, given h_j, we consider the following function that is used in (3.28):

$$q(d_{ij}) = K_o^2 \sigma^2 \gamma \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 P_{\text{NLoS}}^{ij}\right) d_{ij}^2. \tag{3.30}$$

Clearly, considering the fact that $0 \leq P_{\text{LoS}}^{ij} \leq 1$, and $P_{\text{NLoS}}^{ij} = 1 - P_{\text{LoS}}^{ij}$, we have:

$$K_o^2 \sigma^2 \gamma \eta_1 d_{ij}^2 \leq q(d_{ij}) \leq K_o^2 \sigma^2 \gamma \eta_2 d_{ij}^2. \tag{3.31}$$

From (9.54), we can see that $q(d_{ij})$ is bounded between two quadratic functions that each is linearly proportional to d_{ij}^2. Now, using the least square estimation method, we find the coefficients α_1 and α_2 such that, given any h_j, $q(d_{ij})$ is approximated by the following convex quadratic function:

$$q(d_{ij}) \approx \alpha_1 d_{ij}^2 + \alpha_2. \tag{3.32}$$

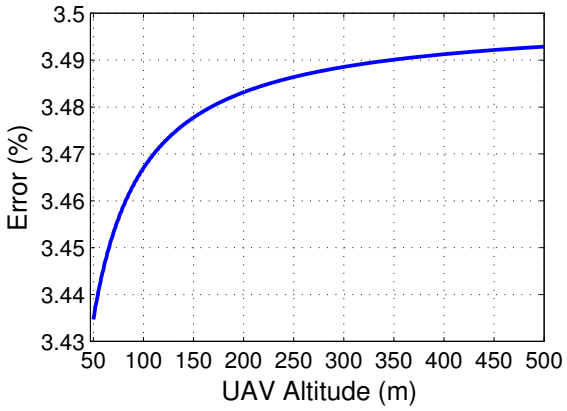


Figure 3.3: Error in the objective function approximation.

where α_1 and α_2 are altitude dependent coefficients. Note that, using the quadratic approximation, the solution of (3.28) becomes more tractable.

Fig. 3.3 shows the error in the objective function (3.28) due the quadratic approximation. As we can see from Fig. 3.3 which is obtained based on the parameters in Table 3.1, the error is less than 4% for different UAVs' altitudes.

Now, in constraint (3.29), we consider $D = \left(\eta_1 P_{\text{LoS}}^{ij} + \eta_2 P_{\text{NLoS}}^{ij}\right) d_{ij}^2$. Clearly, D is an increasing function of d_{ij} since $\eta_1 - \eta_2 < 0$, and for a fixed altitude, LoS probability is a decreasing function of distance. Therefore, using ${d_{ij}}^2 = (x_i - x_j^{\text{uav}})^2 + (y_i - y_j^{\text{uav}})^2 + h_j^2$, and (3.32), for any given h_j we can write the optimization problem (3.28) as:

$$\min_{x_j^{\text{uav}}, y_j^{\text{uav}}} \sum_{i \in \mathcal{C}_j} \left(x_j^{\text{uav}} - x_i\right)^2 + \left(y_j^{\text{uav}} - y_i\right)^2 + {h_j}^2, \tag{3.33}$$

$$\text{s.t. } \left(x_j^{\text{uav}} - x_i\right)^2 + \left(y_j^{\text{uav}} - y_i\right)^2 + h_j^2 - \epsilon^2 \leq 0, \ \ \forall \ i \in \mathcal{C}_j, \text{ and } j \in \mathcal{K}. \tag{3.34}$$

where $\varepsilon = \{d | K_o^2 \sigma^2 \gamma \left(\eta_1 P_{\text{LoS}} + \eta_2 P_{\text{NLoS}}\right) d^2 = P_{\max}\}$. Next, we derive the solution to problem (3.33) that corresponds to finding the optimal UAVs' locations.

Theorem 2. The solution to (3.33) is given by $\boldsymbol{s}^* = (x_j^{\text{uav}*}, y_j^{\text{uav}*}) = -\boldsymbol{P}(\boldsymbol{\lambda})^{-1}\boldsymbol{Q}(\boldsymbol{\lambda})$, with the vector $\boldsymbol{\lambda}$ that maximizes the following concave function:

$$\max_{\boldsymbol{\lambda}} \ \frac{1}{2}\boldsymbol{Q}(\boldsymbol{\lambda})^T \boldsymbol{P}(\boldsymbol{\lambda})^{-1}\boldsymbol{Q}(\boldsymbol{\lambda}) + r(\boldsymbol{\lambda}), \tag{3.35}$$

$$\text{s.t. } \ \boldsymbol{\lambda} \geq 0, \tag{3.36}$$

where $\boldsymbol{P}(\boldsymbol{\lambda}) = \boldsymbol{P}_o + \sum_{i=1}^{|\mathcal{C}_j|} \lambda_i P_i$, $\boldsymbol{Q}(\boldsymbol{\lambda}) = \boldsymbol{Q}_o + \sum_{i=1}^{|\mathcal{C}_j|} \lambda_i Q_i$ and $r(\boldsymbol{\lambda}) = r_o + \sum_{i=1}^{|\mathcal{C}_j|} \lambda_i r_i$, with $\boldsymbol{P}_o$, $\boldsymbol{Q}_o$, r_o, $\boldsymbol{P}_i$, $\boldsymbol{Q}_i$, and r_i given in the proof.

Proof. As we can see from (3.33), the optimization problem is a quadratically constrained quadratic program (QCQP) whose general form is given by [120]:

$$\min_{\boldsymbol{s}} \ \frac{1}{2}\boldsymbol{s}^T\boldsymbol{P}_o\boldsymbol{s} + \boldsymbol{Q}_o^T\boldsymbol{s} + r_o, \tag{3.37}$$

$$\text{s.t. } \frac{1}{2}\boldsymbol{s}^T\boldsymbol{P}_i\boldsymbol{s} + \boldsymbol{Q}_i^T\boldsymbol{s} + r_i, \ \ i \in \mathcal{C}_j. \tag{3.38}$$

Given (3.33) and (3.34), we have:

$$\boldsymbol{P}_o = \begin{bmatrix} 2|\mathcal{C}_j| & 0 \\ 0 & 2|\mathcal{C}_j| \end{bmatrix}, \boldsymbol{P}_i = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \boldsymbol{Q}_o = \begin{bmatrix} -2\sum\limits_{i=1}^{|\mathcal{C}_j|} x_i & -2\sum\limits_{i=1}^{|\mathcal{C}_j|} y_i \end{bmatrix}^T, \boldsymbol{Q}_i = \begin{bmatrix} -2x_i & -2y_i \end{bmatrix}^T.$$

Also, $r_o = \sum\limits_{i=1}^{|\mathcal{C}_j|} x_i^2 + \sum\limits_{i=1}^{|\mathcal{C}_j|} y_i^2$, and $r_i = x_i^2 + y_i^2 + h_j^2 - \epsilon^2$ with $\varepsilon = \{d | K_o^2\sigma^2\gamma\left(\eta_1 P_{\text{LoS}} + \eta_2 P_{\text{NLoS}}\right) d^2 = P_{\max}\}$. Note that, $\boldsymbol{P}_o$ and $\boldsymbol{P}_i$ are positive semidefinite matrices, and, hence, the QCQP problem in (3.37) is convex. Now, we write the Lagrange dual function as:

$$\begin{aligned} f(\lambda) &= \inf_{\boldsymbol{s}} \left[\frac{1}{2}\boldsymbol{s}^T\boldsymbol{P}_o\boldsymbol{s} + \boldsymbol{Q}_o^T\boldsymbol{s} + r_o + \sum_i \lambda_i \left(\frac{1}{2}\boldsymbol{s}^T\boldsymbol{P}_i\boldsymbol{s} + \boldsymbol{Q}_i^T\boldsymbol{s} + r_i \right) \right] \\ &= \inf_{\boldsymbol{s}} \left[\frac{1}{2}\boldsymbol{s}^T\boldsymbol{P}(\boldsymbol{\lambda})\boldsymbol{s} + \boldsymbol{Q}(\boldsymbol{\lambda})^T\boldsymbol{s} + r(\boldsymbol{\lambda}) \right]. \end{aligned} \tag{3.39}$$

Clearly, by taking the gradient of the function inside the infimum with respect to s, we find $\boldsymbol{s}^* = -\boldsymbol{P}(\boldsymbol{\lambda})^{-1}\boldsymbol{Q}(\boldsymbol{\lambda})$. As a result, using $\boldsymbol{s}^*$, $f(\boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{Q}(\boldsymbol{\lambda})^T\boldsymbol{P}(\boldsymbol{\lambda})^{-1}\boldsymbol{Q}(\boldsymbol{\lambda}) + r(\boldsymbol{\lambda})$. Finally, the dual of problem (3.37) or (3.33) will be:

$$\max \ f(\boldsymbol{\lambda}), \ \text{s.t. } \boldsymbol{\lambda} \geq 0, \tag{3.40}$$

which proves Theorem 2. □

Using Theorem 2, for a fixed altitude, we find the optimal 2D coordinates of the UAV, $\boldsymbol{s}^* = (x_j^{\text{uav}*}, y_j^{\text{uav}*})$. Then, the optimal UAV's altitude is the argument that minimizes the following one-dimensional function as:

$$h_j^* = \arg\min_{h_j} \left[\alpha_1 \left(h_j^2 + \|\boldsymbol{s}^*\|^2 \right) + \alpha_2 \right]. \tag{3.41}$$

where α_1 and α_2 are the altitude dependent coefficients given in (3.32). Note that, given (3.41), the optimal altitude of the UAV is obtained via one dimensional search over a feasible range of altitudes. Consequently, we can determine the optimal 3D location of each UAV.

Note that, the device association (presented in Section III), and UAVs' locations optimization (in Section IV) are applied iteratively until there is no change in the location update step. Clearly, at each iteration, the total transmit power of the devices is reduced and the objective function is monotonically decreasing. Hence, the solution converges after several iterations. Note that,

our proposed approach provides a suboptimal solution to the original problem. Nevertheless, our solution has a reasonable accuracy but significantly fast compared to the global optimal solution that can be achieved by the brute-force search, as will be further corroborated in the simulations.

Thus far, we considered the IoT network at one snapshot in the time duration $[0, T]$. Next, we analyze the IoT network considering the entire time duration $[0, T]$ in which the set of active devices changes. In this case, to maintain the power-efficient and reliable uplink communications of the devices, the UAVs must update their locations at different update times t_n.

3.4 Update Times and Mobility of UAVs

Here, we find the optimal update time and trajectory of the UAVs to guarantee the reliable uplink transmissions of the IoT devices. Clearly, the trajectory of the UAVs, as well their update time depend on the activation process of the IoT devices. Furthermore, to move along the optimal trajectories, the UAVs must spend a minimum total energy on mobility so as to remain operational for a longer time. In the considered ground IoT network, the set of active IoT devices changes over time. Consequently, the UAVs must frequently update their locations accordingly. Note that, the UAVs do not continuously move as they must stop, serve the devices, and then update their locations. Furthermore, the mobility of the UAVs is also limited due to their energy constraints. Hence, the UAVs update their locations only at some specific times. In this case, during time interval $[0, T]$, we need to find update times t_n, $1 \leq n \leq N$ with N updates, and a framework for optimizing the mobility of the UAVs at different update times. Here, for tractability, we assume that the devices are synchronized at $t = 0$. In this case, the synchronization process needs to be done only once during the entire activation period $[0, T]$. Note that, our optimization problems for jointly finding the optimal UAVs' locations, the device association, and devices' transmit power at each update time do not depend on this synchronization assumption.

3.4.1 Update Time Analysis

First, we propose a framework to find the update times of the UAVs. As discussed in Section II, each UAV's trajectory consists of multiple stop locations (determined in update times) at which each UAV serves its associated ground devices. Clearly, the update times depend on the activation of the IoT devices during the given time period $[0, T]$. Indeed, the number of update times, N, impacts the optimal location and trajectory of the UAVs as well as the power consumption of the IoT devices. A higher number of updates leads to a shorter time interval between the consecutive updates. Hence, a lower number of devices will be active during the shorter time interval. In such a case, the active devices experience lower interference from each other while transmitting their data to the UAVs. Therefore, the IoT devices can use lower transmit power to meet their SINR constraint. However, a higher number of updates requires more mobility and higher energy consumption for the UAVs. Next, we provide insightful analysis on the update time based on the probabilistic and periodic activation models of the IoT devices.

Periodic IoT activation

In some applications such as weather monitoring, smart grids (e.g. smart meters), and home automation, the IoT devices can report their data periodically. Therefore, the devices are activated periodically. Let τ_i be the activation period of device i during $[0, T]$. Without loss of generality, assume $\tau_1 \leq \tau_2 \leq ... \leq \tau_L$. Due to the periodic nature of devices' activation, we can find the exact number of active devices at each update time t_n.

Proposition 3. The exact number of active IoT devices at update time t_n is given by:

$$b_n = \sum_{i=1}^{L} \mathbb{1}\left(\left\lfloor \frac{{t_n}^-}{\tau_i} \right\rfloor > \left\lfloor \frac{t_{n-1}}{\tau_i} \right\rfloor \right), \quad n > 1, \tag{3.42}$$

$$b_1 = \arg\max_i \left\{ t_1 > \tau_i \right\}, \tag{3.43}$$

where $\mathbb{1}(.)$ is the indicator function which can only be equal to 1 or 0, and ${t_n}^- = \lim_{\varepsilon \to 0^+} (t_n - \varepsilon)$.

Proof. User i becomes active during $[t_{n-1}, t_n)$ if there exists $q \in \mathbb{N}$ such that $t_{n-1} \leq q\tau_i < t_n$. Thus, the number of activations of device i before t_n must be greater than the one until t_{n-1}. Considering the fact that the number of activations before t_n is $\left\lfloor \frac{{t_n}^-}{\tau_i} \right\rfloor$, and until t_{n-1} is $\left\lfloor \frac{t_{n-1}}{\tau_i} \right\rfloor$, we must have:

$$\left\lfloor \frac{{t_n}^-}{\tau_i} \right\rfloor > \left\lfloor \frac{t_{n-1}}{\tau_i} \right\rfloor. \tag{3.44}$$

Hence, the total number of active devices which need to be served at t_n is equal to:

$$b_n = \sum_{i=1}^{L} \mathbb{1}\left(\left\lfloor \frac{{t_n}^-}{\tau_i} \right\rfloor > \left\lfloor \frac{t_{n-1}}{\tau_i} \right\rfloor \right). \tag{3.45}$$

Finally, considering $t_0 = 0$, we can write b_1 as: $b_1 = \arg\max_i \left\{ t_1 > \tau_i \right\}$. □

Proposition 3 gives the exact number of devices that must be served by UAVs at each update time. In this case, the update times can be adjusted according to the number of devices that can be served by the UAVs. Indeed, knowing the exact number of active devices enables us to determine the update times in a deterministic and efficient way based on system requirements.

Probabilistic IoT activation

Certain IoT devices can have probabilistic activations in applications such as health monitoring, and smart traffic control. In this case, each IoT device becomes active at time $t \in [0, T]$ following the beta distribution as given in (3.4). For this scenario, we will next derive the specific update times as a function of the average number of active devices.

Theorem 3. The update times during which, on the average, a total of a_n devices must be served by the UAVs, are given by:

$$t_n = T \times I^{-1}\left(\frac{a_n}{L} + I_{\frac{t_{n-1}}{T}}\left(\kappa,\omega\right),\kappa,\omega\right),\ n>1, \tag{3.46}$$

$$t_1 = T \times I^{-1}\left(\frac{a_1}{L},\kappa,\omega\right), \tag{3.47}$$

where $I_x(.)$ is the regularized incomplete beta function and $I^{-1}(.)$ is the inverse of this function. L is the total number of IoT devices, and $[0,T]$ is the time interval during which IoT devices can be active.

Proof. First, we find the probability that each device becomes active in order to send its data to a UAV at update time t_n. As discussed in the system model, a device needs to transmit at time t_n if it becomes active during time $[t_{n-1},t_n)$. Thus, the probability that each device needs to be served at t_n is:

$$\begin{aligned} p_n &= \int_{t_{n-1}}^{t_n} \frac{t^{\kappa-1}(T-t)^{\omega-1}}{T^{\kappa+\omega-1}B(\kappa,\omega)}\mathrm{d}t = \int_{\frac{t_{n-1}}{T}}^{\frac{t_n}{T}} \frac{t^{\kappa-1}(1-t)^{\omega-1}}{B(\kappa,\omega)}\mathrm{d}t, \\ &= \frac{B_{\frac{t_n}{T}}(\kappa,\omega) - B_{\frac{t_{n-1}}{T}}(\kappa,\omega)}{B(\kappa,\omega)} = I_{\frac{t_n}{T}}(\kappa,\omega) - I_{\frac{t_{n-1}}{T}}(\kappa,\omega), \end{aligned} \tag{3.48}$$

where $B_x(\kappa,\omega) = \int_0^x y^{\kappa-1}(1-y)^{\omega-1}\mathrm{d}y$ is the incomplete beta function with parameters κ and ω, and $I_x(.)$ is the regularized incomplete beta function.
Now, the average number of active devices at t_n is given by:

$$\begin{aligned} a_n &= \sum_{k=1}^{L}\binom{L}{k}{p_n}^k(1-p_n)^{L-k} = Lp_n\sum_{k=1}^{L}\frac{(L-1)!}{(k-1)!\,(L-k)!}{p_n}^{k-1}(1-p_n)^{L-k}, \\ &= \sum_{k'=0}^{L'}\frac{(L')!}{(k')!\,(L'-k')!}{p_n}^{k'-1}(1-p_n)^{L'-k'} = Lp_n, \end{aligned} \tag{3.49}$$

where in (a), we used $L' = L-1$ and $k' = k-1$. Note that, (3.49) corresponds to the mean of a binomial distribution. Then we have:

$$L\left[I_{\frac{t_n}{T}}(\kappa,\omega) - I_{\frac{t_{n-1}}{T}}(\kappa,\omega)\right] = a_n. \tag{3.50}$$

This leads to:

$$t_n = T \times I^{-1}\left(\frac{a_n}{L} + I_{\frac{t_{n-1}}{T}}\left(\kappa,\omega\right),\kappa,\omega\right). \tag{3.51}$$

Finally, considering $I_0(.) = 0$, we find $t_1 = T \times I^{-1}\left(\frac{a_1}{L},\kappa,\omega\right)$. □

Clearly, the update times need to be determined based on the IoT devices' activation patterns. In fact, t_n depends on the number of IoT devices, and their activation distribution. Furthermore,

according to (3.46), each t_n depends also on the previous update time, t_{n-1}. This is due to the fact that, the number of active devices that need to be served at t_n, depends on the update time difference $t_n - t_{n-1}$. Using Theorem 3, the update times of the UAVs can be set based on the average number of active devices. Typically, at each update time, the number of devices which need to be served by the UAVs should not be high in order to avoid high interference. However, considering the number of available resources (orthogonal channels and UAVs), it is preferable to serve a maximum number of active devices at each update time. Hence, in this case, the number of active devices at each update time must not be relatively low.Therefore, considering system requirements and different parameters such as mutual interference between devices, acceptable delay for serving the devices, and number of available channels, an appropriate t_n must be adopted. For instance, using Theorem 3, the update times can be set such that the average number of active devices be lower than the number of channels, R, to avoid interference between the devices. Next, we investigate the UAVs' mobility during the update times.

3.4.2 UAVs' Mobility

Thus far, we have determined the update times as well as the stop locations at each update time. Here, we investigate how the UAVs should move between the stop locations at different update times. In this case, considering the energy limitation of the UAVs, $E_{\max}$, we find the optimal trajectory of each UAV to guarantee reliable and energy-efficient uplink transmissions of active IoT devices. The UAVs update their locations according to the activity of the IoT devices. Therefore, the UAVs move from their initial locations at t_{n-1} to a new optimal locations at t_n. This mobility should be done in such a way that the UAVs spend a minimum total energy on the mobility so as to remain operational for a longer time. In fact, given the optimal sets of UAVs' locations at t_{n-1} and t_n obtained from Section III, we determine how to move the UAVs between the initial and the new sets of locations in order to minimize total mobility of the UAVs.

Now, let $\mathcal{I}_{n-1}$ and $\mathcal{I}_n$ be two sets comprising the UAVs' locations at two consecutive update times t_{n-1} and t_n. Our goal is to find the optimal mapping between these two sets in a way that the energy used for transportations (between two sets) is minimized. Not that, in our model, the total energy which each UAV can use for the mobility during $[0, T]$ is limited to $E_{\max}$. Clearly, in the multiple updates (mobilities) during $[0, T]$, the maximum energy consumption of each UAV at each update is equal to the remaining energy of the UAV. Let $\mathit{\Gamma}_{n,k}$ be the remaining energy of the UAV at the location having index $k \in \mathcal{I}_{n-1}$ at time t_n. Then, we can write the following UAVs' mobility optimization problem:

$$\min_{\boldsymbol{Z}} \sum_{l\in\mathcal{I}_n} \sum_{k\in\mathcal{I}_{n-1}} E_{kl} Z_{kl}, \tag{3.52}$$

$$\text{s.t.} \sum_{l\in\mathcal{I}_n} Z_{kl} = 1, \quad \sum_{k\in\mathcal{I}_{n-1}} Z_{kl} = 1, \tag{3.53}$$

$$E_{lk} \leq \mathit{\Gamma}_{n,k}\,, \;\; Z_{kl} \in \{0, 1\}, \; \forall k \in \mathcal{I}_{n-1}, \forall l \in \mathcal{I}_n, \tag{3.54}$$

where $\mathcal{I}_{n-1}$ and $\mathcal{I}_n$, are the initial and new sets of UAVs' locations at times t_{n-1} and t_n. $\boldsymbol{Z}$ is

the $|\mathcal{I}_n| \times |\mathcal{I}_n|$ assignment matrix with each element Z_{kl} being 1 if UAV k is assigned to location l, and 0 otherwise. E_{kl} is the energy used for moving a UAV from its initial location with index $k \in \mathcal{I}_{n-1}$ to a new location with index $l \in \mathcal{I}_n$. Also, $\Gamma_{n,k}$ is the remaining energy for the UAVs at time t_n. Note that, (3.54) guarantees that UAVs remain operational until the end of the period T. The total energy consumption of the rotary wing UAV while moving between two stop locations can be computed as done in [121]:

$$E = \frac{D}{v}\left(P_V + P_H\right), \tag{3.55}$$

where D is the distance between two stop locations, D/v is the flight duration, P_V is the power consumption for vertical movement, and P_H is the power consumption for horizontal movement. Clearly, if the altitude difference between two stop locations is Δh, the effective vertical and horizontal velocities will be $v_v = v\sin\phi$ and $v_h = v\cos\phi$, with $\phi = \sin^{-1}\left(\frac{\Delta h}{D}\right)$. According to [122] and [123], P_H is composed of parasitic power and induced power needed for overcoming the parasitic drag and the lift-induced drag. The parasitic power, based on [123, equations (13.32), (13.27), and (11.3)], can be given by:

$$P_P = \frac{1}{2}\rho C_{D_o} A_e v_h^3 + \frac{\pi}{4} N_b c_b \rho C_{D_o} \omega^3 R^4 \left(1 + 3\Big(\frac{v_h}{\omega R}\Big)^2\right), \tag{3.56}$$

where v_h is the effective horizontal velocity, C_{D_o} is the drag coefficient, ρ is the air density, c_b is the blade chord, N_b is the number of blades, and A_e is the reference area (frontal area of the UAV) [122] and [123]. We note that the second term in (3.56) represents the blade power profile.

Using [123, equations (13.19), (13.13), and (12.2)], the induced power (assuming zero tilt angle) can be computed by:

$$P_I = \omega R W \times \lambda, \tag{3.57}$$

where R is the rotor disk radius, W is the weight of the UAV, and ω is the angular velocity. Also, given [123, (13.18), (13.13), and (12.1)], we can find λ by solving the following equation:

$$f(\lambda) = 2\rho\pi\omega^2 R^4 \lambda \sqrt{\frac{v_h^2}{\omega^2 R^2} + \lambda^2} - W = 0. \tag{3.58}$$

The power consumption due to the vertical climbing and descending (assuming rapid descent) can be given by [123, equations (12.35), (12.47), (12.50)]:

$$P_V = \begin{cases} \frac{W}{2} v_v + \frac{W}{2}\sqrt{{v_v}^2 + \frac{2W}{\rho\pi R^2}}, & \text{climbing,} \\ \frac{W}{2} v_v - \frac{W}{2}\sqrt{{v_v}^2 - \frac{2W}{\rho\pi R^2}}, & \text{descending (in windmill state),} \end{cases} \tag{3.59}$$

where v_v is the effective vertical velocity. Finally, the total mobility energy consumption is computed using (3.55)-(3.59).

Clearly, the optimization problem in (3.52) is an integer linear programming (ILP). Following the similar approach we used for solving (3.15), we transform problem (3.52) to a standard assignment

Table 3.1: Simulation parameters.

Parameter	Description	Value
$P_{\max}$	Maximum transmit power of each device	200 mW
α	Path loss exponent for LoS links	2
σ^2	Noise power	-130 dBm
γ	SINR threshold	5 dB
L	Total number of IoT devices	500
η_1	Additional path loss to free space for LoS	3 dB
η_2	Additional path loss to free space for NLoS	23 dB

problem which can be solved using the Hungarian method in a polynomial time with a complexity of $O(|\mathcal{I}_n|^3)$. To this end, we need to remove constraint (3.54) by considering $E_{lk} = +\infty$ when the constraint is not satisfied. To determine when (3.54) is not satisfied, we use $\mathcal{I}_{n-1}$ and $\mathcal{I}_n$ to compute E_{lk}, and compare it with the remaining energy of the UAVs, $\Gamma_{n,k}$. Then, in the objective function (3.52), we replace each E_{lk} corresponding to the unsatisfied constraint with $E_{lk} = +\infty$. Consequently, (3.52) is transformed into a standard assignment problem. The result of solving (3.52) will be the assignment matrix, $\boldsymbol{Z}$, that optimally assigns the UAVs to the destinations. Therefore, the locations of the UAVs are updated according to the new destinations. Then, having the destinations of each UAV at different update times, we can find the optimal trajectory of the UAVs.

3.5 Simulation Results and Analysis

For our simulations, the IoT devices are deployed within a geographical area of size 1 km $\times$ 1 km. In this case, we consider a total number of 500 IoT devices which are uniformly distributed on the area. Furthermore, we consider UAV-based communications in an urban environment with $\psi = 11.95$ and $\beta = 0.14$ at 2 GHz carrier frequency [8]. Table 3.1 lists the simulation parameters. Here, we analyze the transmit power of the IoT devices, the energy consumption of UAVs on their mobility, and the update times. In our update time analysis, unless otherwise stated, we consider the probabilistic activation model for the IoT devices with the beta distribution parameters $\kappa = 3$, and $\omega = 4$ [111]. When applicable, we compare our results with pre-deployed stationary aerial base stations (UAVs) scenario while adopting the optimal device association and power control technique of Subsection 3-1. In the stationary case, the locations of UAVs are assumed to be fixed over the target area and they are not updated according to the devices' locations. All statistical results are averaged over a large number of independent runs.

Note that, in the given IoT network, serving all the active devices may not be possible due to the limitations on the number of UAVs and the maximum transmit power of the devices. Thus, in Fig. 3.4, we show the achieved system *reliability* which, here, is defined as the probability that all the active devices can be served by the UAVs. Clearly, the reliability depends on the locations and transmit powers of the devices as well as the number of UAVs.

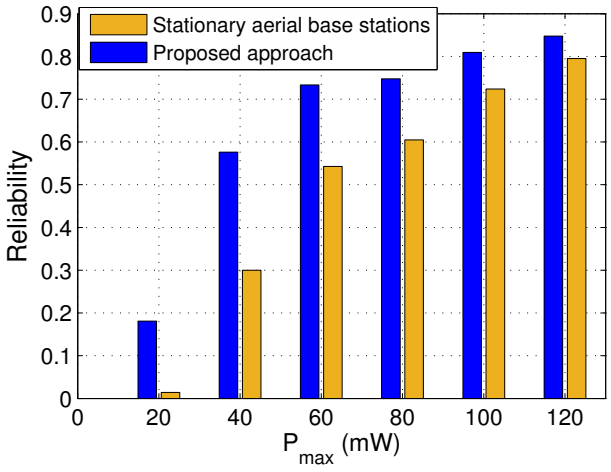


Figure 3.4: Reliability comparison between the proposed approach and stationary aerial base stations using 5 UAVs.

Fig. 3.4 shows the reliability as the maximum transmit power of the devices, $P_{\max}$, varies. In this case, 5 UAVs are deployed to serve 100 active IoT devices. Clearly, as $P_{\max}$ increases, the reliability also increases. In fact, for higher $P_{\max}$ values, the devices have higher a chance to successfully connect to UAVs. From Fig. 3.4, we can see that, our proposed approach leads to a significantly improved reliability compared to the case in which stationary aerial base stations are used. In particular, the difference between the reliability of the stationary case and our proposed approach is significant for lower $P_{\max}$. Indeed, a higher reliability is achieved by dynamically optimizing the UAVs' locations based on the locations of the IoT devices. As shown in Fig. 3.4, by increasing $P_{\max}$ from 40 mW to 100 W, the reliability increases from 0.3 to 0.72 for the stationary case, while it increases from 0.58 to 0.82 in our proposed approach. Furthermore, the proposed approach yields a maximum of 28% improvement in the system reliability.

Fig. 3.5 shows a snapshot of the UAVs' locations and their associated IoT devices (indicated by the same color) resulting from the proposed approach. In this figure, 5 UAVs are efficiently deployed to serve 100 active IoT devices which are uniformly distributed on the area. In this case, all the devices are able to send their data to the associated UAVs by using a minimum total transmit power. Here, the 3D locations of the UAVs as well as the device association are determined based on the locations of the ground IoT devices and their transmit power.

In Fig. 3.6, we show the total transmit power needed by the IoT devices for reliable uplink communications, versus the number of UAVs in the interference scenario. Clearly, the total transmit power of the IoT devices can be reduced by deploying more UAVs. For instance, considering 100 active devices and 20 available channels, using our proposed approach, the total transmit power decreases from 2.4 W to 0.2 W by increasing the number UAVs from 5 to 10. Furthermore, using the proposed approach, the total transmit power of the devices decreases by 45% (on the average) compared to the stationary case. Clearly, for a lower number of UAVs, the proposed approach leads to higher power reduction compare to the stationary case. In other words, intelligently opti-

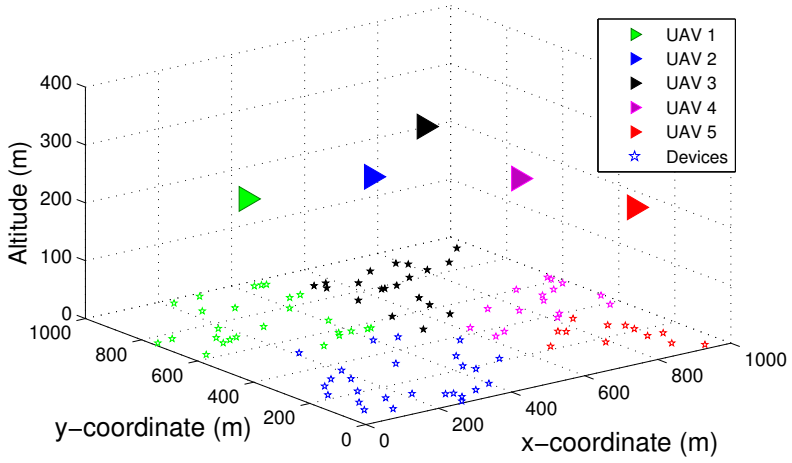


Figure 3.5: UAVs' locations and associations for one illustrative snapshot.

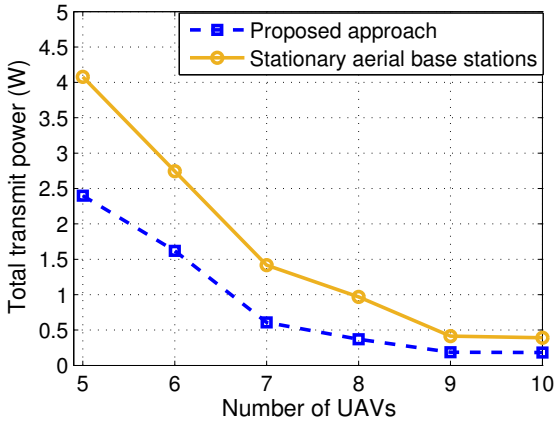


Figure 3.6: Total transmit power of devices vs. number of UAVs in the presence of interference.

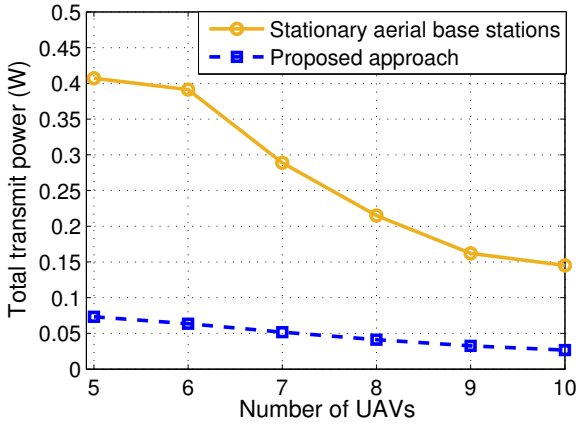


Figure 3.7: Total transmit power of devices vs. number of UAVs in the interference-free scenario.

mizing the locations of UAVs provides more power reduction gains when the number of UAVs is low. In fact, for very dense networks with a high number of UAVs, updating the UAVs' locations is obviously no longer necessary compared to a case with a low number of UAVs. For instance, as we can see from Fig. 3.6, the power reduction gain achieved by deploying 5 UAVs is around 7 times larger than the case with 10 UAVs.

Fig. 3.7 shows the total transmit power of the IoT devices as a function of the number of UAVs in an interference-free scenario. Compared to the interference scenario, the devices can obviously use a lower transmit power for sending their data to the UAVs. For instance, by efficiently deploying only 5 UAVs, the devices can establish reliable uplink communications with a total transmit power of 70 mW. Furthermore, Fig. 3.7 shows that, our proposed approach leads to an average of 80%

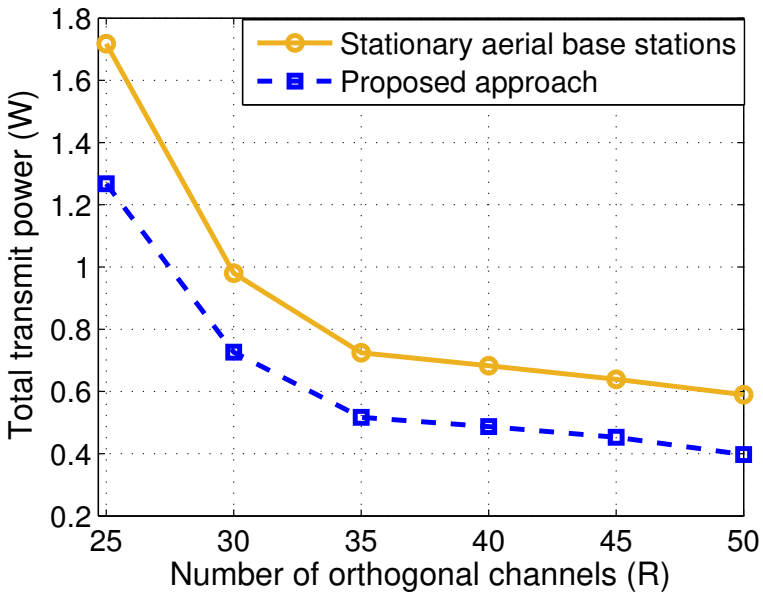


Figure 3.8: Total transmit power of devices vs. number of orthogonal channels.

power reduction compared to the stationary case.

Fig. 3.8 shows the total transmit power of devices used for meeting the SINR requirement as the number of available channels varies. The result in Fig. 3.8 corresponds to a case with 100 active devices which are served by 5 UAVs. Clearly, the total transmit power decreases as the number of channels increases. This is due to the fact that, when more orthogonal resources are available, the interference between the devices will decrease. As a result, each device can reduce its transmit power while connecting to the serving UAV. From Fig. 3.8, we can see that, by increasing the number of channels from 25 to 50, the total transmit power of devices can be reduced by 68% in the proposed approach. In fact, the average number of interfering devices decreases from 4 to 2 when we increase the number of channels from 25 to 50. Consequently, less interference is generated by the devices while transmitting to the UAVs.

In Fig. 3.9, we show the average number of active devices that must be served by UAVs at different update times t_n which are normalized by T. Clearly, the number of active devices at each update time depends the activation process of the devices and the number of update times that indicates how frequently the UAVs serve the devices. In Fig. 3.9, due to the beta distribution-based activation pattern of the IoT devices, the number of active devices decreases when t_n exceeds 0.5 for $N = 10$. From Fig. 3.9, we can see that, for a higher number of update times or equivalently shorter time period between consecutive updates, the average number of devices that need to transmit their data decreases. For instance, considering $t_n = 0.6$, the average number of active devices decreases from 180 to 80 when the number of updates increases from 5 to 10. We also note that, while a lower number of active devices leads to a lower interference between the devices, it requires more updates and mobility for the UAVs. Fig. 3.9 also verifies that the analytical results in Theorem

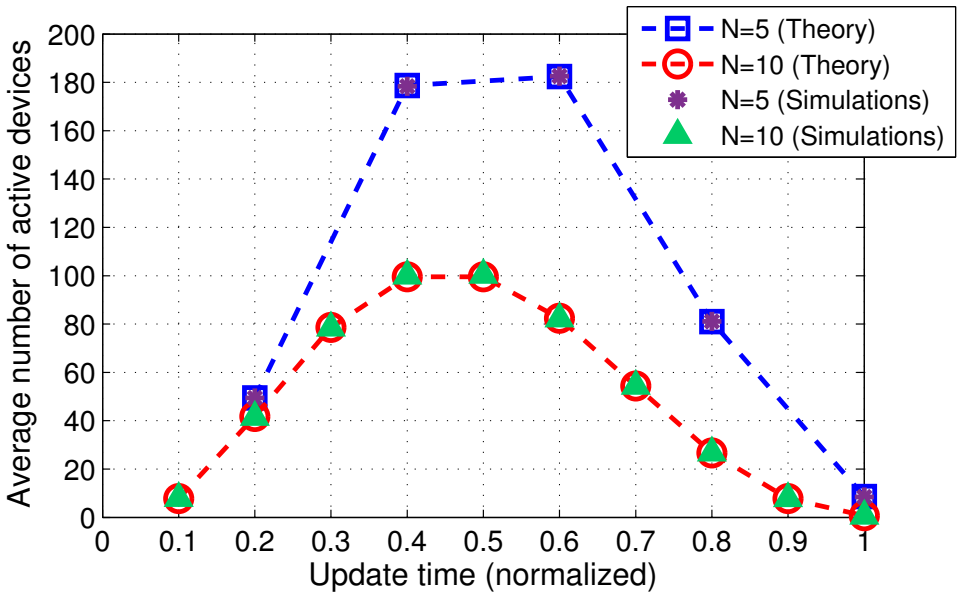


Figure 3.9: Average number of active devices at update times for the probabilistic activation.

2 match the simulations. Furthermore, in Fig. 3.10, we show the exact number of active devices for the *periodic activation case* obtained from Proposition 2. In this case, each device becomes active with a certain activation period, τ_i. As expected, for a higher number of updates, a lower number of active devices will need to be served by the UAVs. For instance, by increasing the number of updates from 10 to 30, on the average, the number of active devices decreases by 58%. Moreover, Fig. 3.10 shows that the maximum number of active devices for 10 updates is about two times larger than the case with 30 updates. Therefore, in order to avoid the interference between the devices, the number of orthogonal channels must be increased by a two-fold factor when the number of updates decreases from 30 to 10.

Fig. 3.11 presents a direct result of Theorem 2 that computes the update times based on the average number of active devices. Fig. 3.11 shows how to set update times in order to ensure that the number of devices (which needs to be served) at each update time does not exceed a specified number, a. As we can see from Fig. 3.11, to achieve a lower value of a, updates must occur more frequently to reduce the time interval between the consecutive updates. For example, as can be seen from this figure, to meet $a = 100$, 75, and 50, the 5th update must occur at $t_n = 0.41$, 0.55, and 1. Moreover, Fig. 3.11 shows that, the number of updates increases as a decreases. For example, in this case, to reduce a from 100 to 50, the number of updates needs to be doubled.

Fig. 3.12 shows the impact of the number of updates on the amount of energy that the UAVs use to move. For our simulations, we have considered $v = 10\,\text{m/s}$, $\rho = 1.225\,\text{kg/m}^{-3}$, $\omega = 20\,\text{rad/s}$, $R = 0.5\,\text{m}$, $c_b = 10\,\text{cm}$, $N_b = 4$, and $W = 50\,\text{N}$ [123]. Intuitively, a higher number of updates requires more mobility of the UAVs. Therefore, by increasing the number of updates, the total energy consumption of the UAVs will also increase. As we can see from Fig. 3.12, by increasing

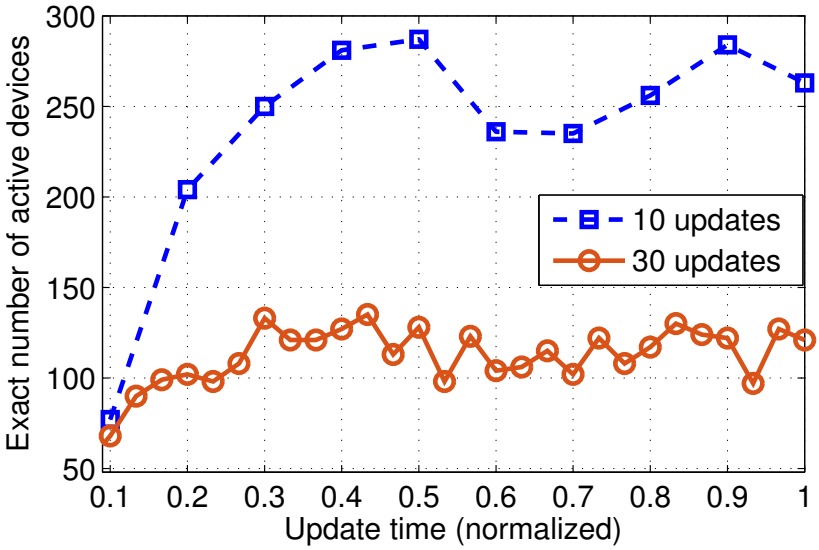


Figure 3.10: Exact number of active devices at different update times for the periodic activation.

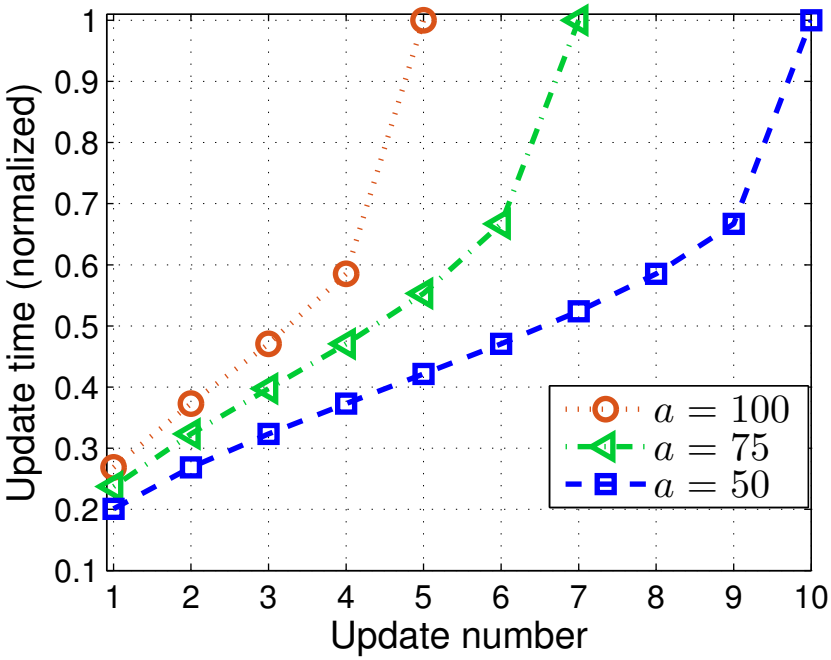


Figure 3.11: Update times for different average number of active devices.

the number of updates from 3 to 6, the energy consumption of UAVs increases by a factor of 2.1 when the target area size is $1\,\text{km} \times 1\,\text{km}$. Note that, the mobility of the UAVs also depends on the size of geographical area in which the devices are distributed. Clearly, on average, the UAVs need

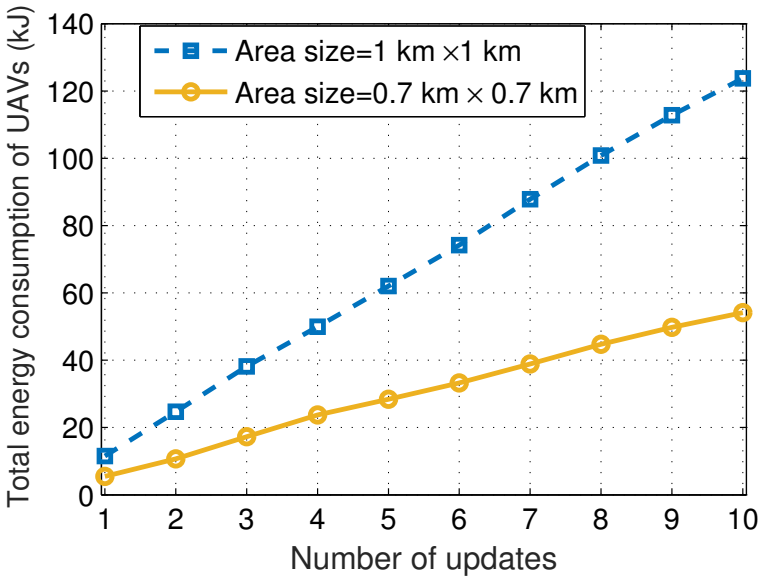


Figure 3.12: Total UAV energy consumption vs. number of updates.

to move further for covering a larger area.

Interestingly, there is an inherent tradeoff between the number of updates, mobility of the UAVs, and transmit power of the IoT devices. In fact, considering Fig. 3.12, a higher number of updates leads to a higher energy consumption of the UAVs due to the higher mobility. In addition, as shown in Fig. 3.9, as the number of updates increases, a lower number of the IoT devices will be active at each update time, and, hence, there will be lower interference between the devices. As a result, the transmit power of the devices that is needed for satisfying the SINR requirement, can be reduced. Note that, as we showed in Fig. 3.8, the devices' transmit power decreases as the interference decreases (by increasing the number of orthogonal channels). Hence, while a higher number of updates leads to a lower devices' transmit power, it requires more UAVs' mobility.

Fig. 3.13 shows the overall convergence of the proposed power minimization algorithm that is used for solving the original problem (3.5) considering 5 UAVs. As we can see from the figure, in this case, the total transmit power of the IoT devices converges after 5 iterations. In Fig. 3.13, each iteration corresponds to a joint solution to the device association and UAVs' locations optimization problems. Clearly, after several iterations, updating the device association and UAVs' locations will no longer improve the solution.

In Fig. 3.14, we show an example to compare the accuracy and time complexity of our proposed approach with the optimal solution obtained by an exhaustive search. Here, to perform an exhaustive search over the continuous space, we have discretized the space with a resolution of 0.1 m. In this case, two UAVs are deployed to serve the devices. Clearly, the average gap between the proposed solution and the optimal solution is around 11%. However, in this example, the proposed solution is around 500 times, on the average, faster than the optimal solution.

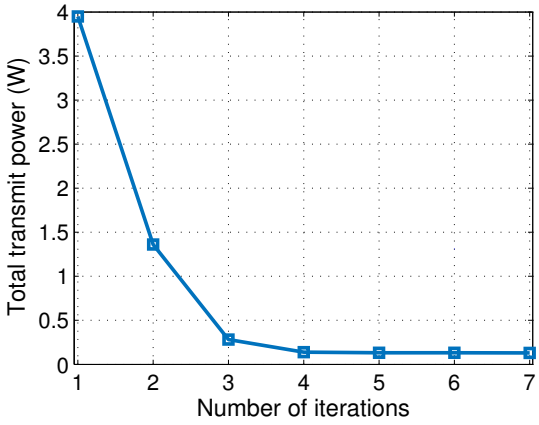


Figure 3.13: Overall convergence of the algorithm.

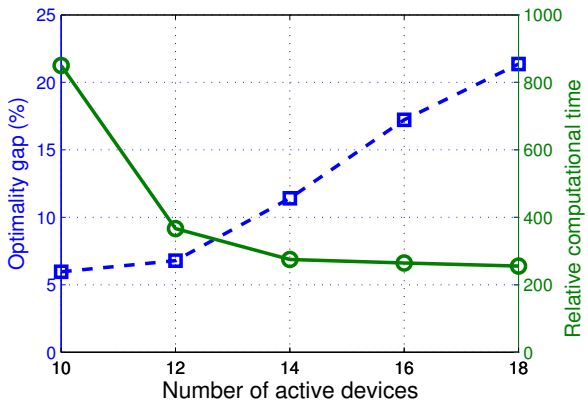


Figure 3.14: Proposed approach vs. optimal solution.

3.6 Summary

In this chapter, we have proposed a novel framework for efficiently deploying and moving UAVs to collect data in the uplink from ground IoT devices. In particular, we have determined the jointly optimal UAVs' locations, device association, and uplink power control of the IoT devices such that the total transmit power of the devices under their SINR constraints is minimized. In addition, we have investigated the effective movement of the UAVs to collect the IoT data in a time-varying IoT network. For this case, based on the devices' activation process, we have derived the update time instances at which the UAVs must update their locations. Furthermore, we have obtained the optimal trajectories that are used by the UAVs to dynamically serve the IoT devices with a minimum energy consumption. The results have shown that by intelligently moving and deploying the UAVs, the total transmit power of the devices significantly decreases compared to the case with pre-deployed stationary aerial base stations. Moreover, there is a fundamental tradeoff between the number of updates, the UAVs' mobility, and the devices' transmit power.

Chapter 4

Hover Time Optimization in UAV-Enabled Wireless Networks

4.1 Background, Related Works, and Contributions

With their inherent attributes such as mobility, flexibility, and adaptive altitude, UAVs have several key potential applications in wireless systems, as pointed out in Chapter 1. Despite the several benefits and practical applications of using UAVs as aerial base stations, one must address many technical challenges such as performance analysis, deployment, air-to-ground channel modeling, user association, and flight time optimization [11] and [12].

In [15], the authors performed air-to-ground channel modeling for UAV-based communications in various propagation environments. In [27] and [18], the authors studied the efficient deployment of aerial base stations to maximize the coverage and rate performance of wireless networks. In [124], the authors investigated the energy-efficient path planning of a UAV-mounted cloudlet which is used to provide offloading opportunities to ground devices. The work in [26] studied the optimal trajectory and heading of UAVs for sum-rate maximization in uplink communications. An analytical framework for trajectory optimization of a fixed-wing UAV for energy-efficient communications was presented in [72]. The work in [80] jointly optimized user scheduling and UAV trajectory for maximizing the minimum average rate among ground users. The authors in [17] derived an exact expression for downlink coverage probability for ground receivers which are served by multiple UAVs.

Another important challenge in UAV-based communications is user (or cell) association. The work in [16] investigated the area-to-UAV assignment for capacity enhancement of heterogeneous wireless networks. However, this work is limited to the case with a uniform spatial distribution of ground users. Moreover, the work in [16] does not consider any fairness criteria that can be affected by the network congestion in a non-uniform users' distribution case. In addition, this work ignores the UAVs' flight time constraints while determining the cell partitions associated

with the UAVs. In [125] the optimal cell partitions associated to UAVs were determined with the goal of minimizing the UAVs' transmit power while satisfying the users' rate requirements. However, in [125], the impact of flight time constraints on the performance of UAV-to-ground communications is not taken into account.

Indeed, the *flight time duration of the UAVs* presents a unique design challenge for UAV-based communication systems [94] and [95]. For instance, the performance of such systems significantly depends on the *hover time* of each UAV, which is defined as the flight time during which the UAV must stay in the air over a given area for providing wireless service to ground users. In fact, with a higher hover time of the UAV, the users can receive wireless service for a longer period. Thus, by increasing the hover time, a UAV can meet higher load requirements and serve a larger area. However, the hover time of a UAV is naturally limited due to the highly constrained battery-provided, on-board energy, as well as flight regulations such as no-fly time/zone constraints [43]. Hence, while analyzing the UAV-based communication systems, the hover time constraints must be also taken into account. In this case, there is a need for a framework to analyze and optimize the performance of UAV-based communications based on the hover time of UAVs. In fact, to our best knowledge, none of the previous UAV studies such as [7,11,12,15–18,23,26,27,35,61,72,80,124], considered the hover time constraints in their analysis.

In this chapter, we develop a novel framework for optimized UAV-to-ground communications under explicit UAVs' hover time constraints. In particular, we consider a network in which multiple UAVs are deployed as aerial base stations to provide wireless service to ground users that are distributed over a geographical area based on an arbitrary spatial distribution. We investigate two key practical scenarios: *UAV communication under hover time constraints*, and *UAV communication under load constraints*. In the first scenario, given the maximum possible hover time of UAVs that is imposed by the limited on-board energy of UAVs and flight regulations, we maximize the average number of bits (data service) that is transmitted to the users under a fair resource allocation scheme. To this end, given the hover times and the spatial distribution of users, we find the optimal cell partitions associated to the UAVs. In this case, using the powerful mathematical framework of *optimal transport theory* [100], we propose a gradient-based algorithm that optimally partitions the geographical area based on the users' distribution as well as the UAVs' hover times and locations. In the second scenario, given the load requirements of ground users, we minimize the average hover time needed for completely serving the users. To this end, we introduce an optimal bandwidth allocation scheme as well as optimal cell partitions for which the average hover time of UAVs is minimized. Our results for the first scenario show that our proposed cell partitioning approach leads to a significantly higher fairness among the users compared to the classical weighted Voronoi approach. For the second scenario, the results show that the average hover time can be reduced by 64% by adopting the proposed optimal bandwidth allocation and cell partitioning approach. Furthermore, our results reveal an inherent tradeoff between the hover time of UAVs and bandwidth efficiency while servicing the ground users.

The rest of this chapter is organized as follows. In Section 4.2, we present the system model. In Section 4.3 we investigate Scenario 1 in order to maximize data service. Section 4.4 presents Scenario 2 for minimizing the hover time of UAVs. Simulation results are presented in Section 4.5

and conclusions are drawn in Section 6.5.

4.2 System Model

Consider a geographical area $\mathcal{D} \subset \mathbb{R}^2$ within which a number of wireless users are located according to a given distribution $f(x, y)$ in the two-dimensional plane[1]. Here, we consider a fixed number of users, N, which are independently distributed on the area. In this area, a set $\mathcal{M}$ of M UAVs are used as aerial base stations to provide wireless service for the ground users[2]. Let $\boldsymbol{s}_i = (x_i, y_i, h_i)$ be the three-dimensional (3D) coordinate of each UAV $i \in \mathcal{M}$ with h_i being the altitude of UAV i. We consider a downlink scenario in which each UAV adopts a frequency division multiple access (FDMA) technique to provide service for the ground users. Let P_i and B be, respectively, the maximum transmit power and the total available bandwidth [3] for UAV i. We use $\mathcal{A}_i$, as shown in Fig. 4.1, to denote the partition of the geographical area which is served by UAV i. In this case, all users located in cell partition $\mathcal{A}_i$ will be connected to UAV i. Hence, the geographical area is divided into M disjoint partitions each of which is associated with one of the UAVs. Let τ_i be the *hover time* of UAV i, defined as the time duration that a UAV uses to hover (stop) over the corresponding cell partition to service the ground users. During the hover time, the UAV must initiate connections to the ground users, perform required computations, and transmit data to the users. Let T_i be the effective data transmission period during which a UAV services the users. In general, the effective data transmission time is less than the total hover time. Consequently, we consider a *control time* as $g_i(.)$, a function of the number of users in $\mathcal{A}_i$, to represent the portion of the hover time that is not used for the effective data transmission. This control time naturally captures the total time that a UAV i needs to spend for computations, setting up connections, and control signaling. Intuitively, the control time will increase when the number of users in the corresponding cell partition increases.

In our model, we use the term *data service* to represent the amount of data (in bits) that each UAV transmits to a given user. Clearly, the data service depends on several factors such as the effective data transmission time (which is directly related to flight time), and the transmission bandwidth. Therefore, here, the effective data transmission times and bandwidth of the UAVs are considered as *resources* which are used for servicing the users. Given this model, to provide service for the ground users using UAVs, we consider two scenarios. The first scenario, **Scenario 1**, can be referred to as *UAV communications under hover time constraints*. In this case, given the maximum possible hover times (imposed by the energy and flight limitations of each UAV), we maximize the average data service to the users under a fair resource allocation policy by optimal cell partitioning of the area. Here, we optimally partition the geographical area based on the hover times and the spatial distribution of users. In Scenario 1, given the maximum possible hover time of each

[1] In this model, the locations of users can be modeled as homogeneous or inhomogeneous point process.

[2] For wireless backhauling of aerial networks, satellite and WiFi are considered as the two feasible candidates [43].

[3] While this model considers a frequency reuse factor of one, our model can accommodate any arbitrary frequency reuse factor.

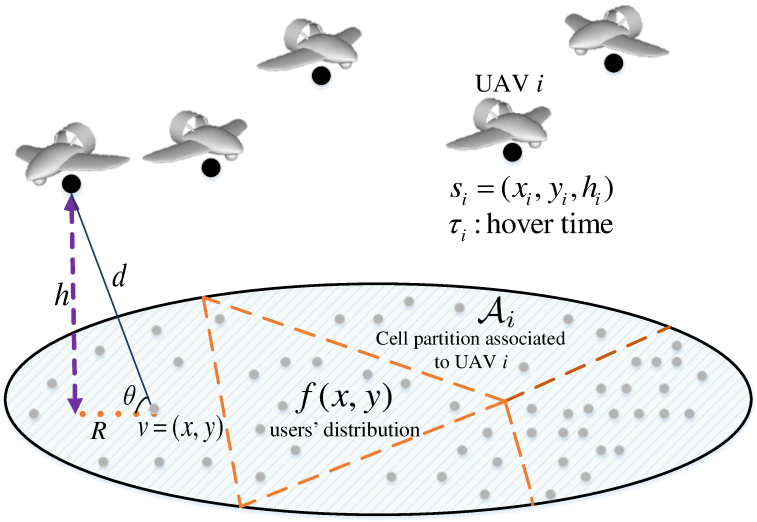


Figure 4.1: System model.

UAV, the total data service under user fairness considerations is maximized. In fact, Scenario 1 corresponds to resource-limited communication scenarios in which the amount of resources (e.g. hover times and bandwidth) is not sufficient to completely meet the demands. An example of such scenario is when battery-limited UAVs are deployed in hotspots with high number of users and demands. The second scenario, **Scenario 2**, is referred to as *UAV communication under load constraints*. In this case, our goal is to completely meet the demands of ground users by properly adjusting the hover time of the UAVs. In particular, given the load requirement of each user at a given location, we minimize the average hover time needed for completely serving the ground users. As a result, the load requirement of the ground users will be satisfied with a minimum average hover time of the UAVs. In this case, by minimizing the hover time, one can minimize the energy consumption of the UAVs as well as the time needed to completely serve the ground users. Such analyses in Scenario 2 are primarily useful is emergency situations in which all users need to be quickly served by the UAVs.

We note that, beyond the optimal cell partitioning, the optimal deployment and mobility of the UAVs based on the ground users' locations can provide performance gains. In the considered model, we assume that the locations of the UAVs are given as an input to the cell association problem. In other words, given any deployment configuration for the UAVs, our approach can be used to determine the optimal cell association.

4.2.1 Air-to-ground path loss model

The air-to-ground signal propagation is affected by the obstacles and buildings in the environment. In this case, depending on the propagation environment, air-to-ground communication links can be either LoS or non-line-of-sight (NLoS). In general, while designing a UAV-based communication system, a complete information about the exact locations, heights, and the number of obstacles may

not be available [15]. In such case, one must consider the randomness associated with the LoS and NLoS links [18], [15]. Clearly, the probability of having LoS communication links depends on the locations, heights, and the number of obstacles, as well as the elevation angle between a given UAV and it's served ground user. In our model, we consider a widely used probabilistic path loss model provided by International Telecommunication Union (ITU-R) [58], and the work in [15]. In this case, the path loss between UAV i and a given user at location (x, y) can be given by [18], and [15]:

$$\Lambda_i(x,y) = \begin{cases} \left(\frac{4\pi f_c d_o}{c}\right)^2 \left(d_i(x,y)/d_o\right)^2 \mu_{\text{LoS}}, & \text{LoS link,} \\ \left(\frac{4\pi f_c d_o}{c}\right)^2 \left(d_i(x,y)/d_o\right)^2 \mu_{\text{NLoS}}, & \text{NLoS link,} \end{cases} \tag{4.1}$$

where μ_{LoS} and μ_{NLoS} are different attenuation factors considered for LoS and NLoS links. Here, f_c is the carrier frequency, c is the speed of light, and d_o is the free-space reference distance. $d_i(x,y) = \sqrt{(x-x_i)^2 + (y-y_i)^2 + h_i^2}$ is the distance between UAV i and an arbitrary ground user located at (x, y). For the UAV-user link, the LoS probability is given by [15]:

$$P_{\text{LoS},i} = b_1 \left(\frac{180}{\pi}\theta_i - 15\right)^{b_2}, \tag{4.2}$$

where $\theta_i = \sin^{-1}(\frac{h_i}{d_i(x,y)})$ is the elevation angle (in radians) between the UAV and the user. Also, b_1 and b_2 are constant values reflecting the environment impact. Note that, the NLoS probability is $P_{\text{NLoS},i} = 1 - P_{\text{LoS},i}$. Clearly, considering $d_o = 1$ m, and $K_o = \left(\frac{4\pi f_c}{c}\right)^2$ the average path loss is $K_o {d_i}^2(x,y)[P_{\text{LoS},i}\mu_{\text{LoS}} + P_{\text{NLoS},i}\mu_{\text{NLoS}}]$. Hence, the received signal power from UAV i will be:

$$\bar{P}_{r,i}(x,y) = \frac{P_i/\bar{N}_i}{K_o {d_i}^2(x,y)\left[P_{\text{LoS},i}\mu_{\text{LoS}} + P_{\text{NLoS},i}\mu_{\text{NLoS}}\right]}, \tag{4.3}$$

where P_i is the total transmit power of UAV i which is equally divided among its associated users, and $\bar{N}_i = N \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y$ represents the average number of users in partition $\mathcal{A}_i$. Now, considering the fact that the noise and interference powers will linearly decrease with the number of users associated to UAV i, the received SINR for a user located at (x, y) and served by UAV i can be written as:

$$\gamma_i(x,y) = \frac{\bar{P}_{r,i}(x,y)}{I_i(x,y) + \sigma^2}, \tag{4.4}$$

where $I_i(x,y) = \beta \sum_{j \neq i} \bar{P}_{r,j}(x,y)$ is the received interference at location (x, y) stemming from all UAVs except UAV i. We also consider a weight factor $0 \leq \beta \leq 1$ to adjust the amount of interference and capture the impact of any interference mitigation technique. Naturally, $\beta = 1$ and $\beta = 0$, respectively, correspond to the full interference and interference-free scenarios.

Clearly, the throughput of a user located at (x, y) if it connects to UAV i is:

$$C_i(x,y) = W(x,y)\log_2\left(1 + \gamma_i(x,y)\right), \tag{4.5}$$

where $W(x, y)$ is the bandwidth allocated to the user at (x, y).

Subsequently, the total data service for the user provided by the UAV will be:

$$L_i(x, y) = T_i C_i(x, y), \tag{4.6}$$

where T_i is the effective transmission time of UAV i. Also, $L_i(x, y)$ represents the total number of bits transmitted to the user located at (x, y). Note that, the data service offered to each ground user depends on a number of key parameters such as the location of the user and the serving UAV, the bandwidth allocated to the user, and the effective data transmission time of the UAV, T_i. Here, we consider the available bandwidth and effective data transmission times as the *resources* used by the UAVs to service the ground users. Clearly, the amount of resources that each user can receive depends on several parameters such as the total number of users, cell partitions as well as bandwidth and hover times of the UAVs. Given this model, we next analyze Scenario 1.

4.3 Scenario 1: Optimal Cell Partitioning for Data Service Maximization under Fairness Constraints

In this section, our goal is to find the optimal cell partitions that maximize the average data service to the ground users based on the UAVs' hover times and the spatial distribution of the ground users. In this case, each cell partition is assigned to one UAV, and the users within the cell partition must be serviced by the corresponding UAV. We note that our model is based on the statistical spatial distribution of the ground users (that indicates the likelihood with which a user will be located at location (x, y)) without using any information about the exact or deterministic locations of the users. In other words, in order to find the optimal cell association, the only information that the network uses is the locations of UAVs as well as the spatial distribution of users within a given geographical area. Given such information, the two classical approaches for finding the cell association are distance-based association (Voronoi diagram), and SINR-based association (weighted Voronoi diagram) [126–128]. In classical cell partitioning approaches such as Voronoi and weighted Voronoi diagrams, the spatial distribution of users is not taken into account. As a result, some partitions can be highly congested with users and, hence, each user will receive significantly lower amount of resources than those in less congested partitions. Thus, such classical cell partitioning approaches can lead to a highly unfair data service for the users. In our cell partitioning approach, however, while maximizing the total data service, we can load balance users across cells by controlling the amount of resources that must be allocated to each user. Hence, our approach avoids creating unbalanced cell partitions and, thus, it leads to a higher level of fairness compared to the classical Voronoi approach.

Furthermore, there are other cell association approaches such as the ones that consider load balancing or fairness [129–131] that can be applied to uneven distributions of users. However, these methods cannot be applied to our model as they need the exact information about the configuration (i.e. locations) of the users. In other words, such approaches (e.g. in [129–131]) can only be used

when a realization of the users' distribution is given, and, thus, they are unable to capture a general probabilistic spatial distribution of users in cell association as considered here. Consequently, next, we present the details of our cell partitioning approach based on the UAVs' hover times and the spatial distribution of users.

Let τ_i be the hover time of UAV i during which it provides service for the users located in the corresponding cell partition, $\mathcal{A}_i$. The hover time is composed of the effective data transmission time and the control time. To ensure a fair resource allocation policy, we consider the following fairness criterion:

$$T_i = \tau_i - g_i \left(\int_{\mathcal{A}_i} f(x,y) \mathrm{d}x\mathrm{d}y \right), \ \forall i \in \mathcal{M}, \tag{4.7}$$

where g_i is the control time which depends on the number of the users located in $\mathcal{A}_i$. Note that, the average number of users within each cell partition $\mathcal{A}_i$ is linearly proportional to $\int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y$. In other words, given the spatial distribution of users, $f(x,y)$, and the total number of users, N, the average number of users in partition $\mathcal{A}_i$ is equal to $N \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y$ [127]. From (4.5) and (4.6) which are used to compute the amount of data service, we can see that the value $T_i B$ can be considered as the resources that UAV i uses to service users in $\mathcal{A}_i$. In this case, we consider the following resource allocation scheme:

$$\frac{\alpha_i T_i B}{N \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y} = \frac{\alpha_j T_j B}{N \int_{\mathcal{A}_j} f(x,y)\mathrm{d}x\mathrm{d}y}, \ \forall i \neq j \in \mathcal{M},^{4} \tag{4.8}$$

where α_i and α_j are resource allocation factors that can be adjusted to control the amount of resources allocated to users in partitions $\mathcal{A}_i$ and $\mathcal{A}_j$. Constraint (4.8) ensures that a UAV with more resources (bandwidth and hover time) or a higher resource allocation factor will serve a higher number of users.

Now, using (4.8) and considering the fact that $\int_{\mathcal{D}} f(x,y)\mathrm{d}x\mathrm{d}y = \sum_{k=1}^{M} \int_{\mathcal{A}_k} f(x,y)\mathrm{d}x\mathrm{d}y = 1$, we have the following constraint on the number of users (i.e. load) in each partition:

$$\int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y = \frac{\alpha_i T_i}{\sum_{k=1}^{M} \alpha_k T_k}, \ \forall i \in \mathcal{M}. \tag{4.9}$$

As we can see from (4.9), the number of users in each generated optimal partition will depend on the UAVs' resources. Clearly, when the UAVs have the same hover times, bandwidths, and resource allocation factors, (4.7)-(4.9) lead to $\int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y = \frac{1}{M}, \forall i \in \mathcal{M}$. This case implies that the identical UAVs will service equally-loaded cell partitions. Now, given (4.5), (4.6), and (4.9), we can write the average data service at location $(x,y) \in \mathcal{A}_i$ as:

$$L_i(x,y) = \frac{T_i B}{N \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y} \log_2\left(1+\gamma_i(x,y)\right) = \left(\frac{B}{N\alpha_i} \sum_{k=1}^{M} T_k \right) \log_2\left(1+\gamma_i(x,y)\right). \tag{4.10}$$

[4]Given hover times of the UAVs, τ_i, $\forall i \in \mathcal{M}$, we can compute T_i by solving the system of equations in (4.7) and (4.8).

Now, we formulate an optimization problem for maximizing the average data service by optimal partitioning of the target area. In this case, by defining $\lambda_i = \frac{B}{N\alpha_i}\sum_{k=1}^{M} T_k$, and $\omega_i = \frac{\alpha_i T_i}{\sum_{k=1}^{M}\alpha_k T_k}$, the data service maximization problem can be given by the following minimization problem:

$$\min_{\mathcal{A}_i, i\in\mathcal{M}} \sum_{i=1}^{M} \int_{\mathcal{A}_i} -\lambda_i \log_2\left(1+\gamma_i(x,y)\right) f(x,y)\mathrm{d}x\mathrm{d}y, \tag{4.11}$$

$$\text{s.t.} \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y = \omega_i, \ \ \forall i \in \mathcal{M}, \tag{4.12}$$

$$\mathcal{A}_l \cap \mathcal{A}_m = \emptyset, \ \ \forall l \neq m \in \mathcal{M}, \tag{4.13}$$

$$\bigcup_{i\in\mathcal{M}} \mathcal{A}_i = \mathcal{D}. \tag{4.14}$$

where (4.12) captures the constraint on the load of each cell partition. Also, (9.7) and (9.8) ensure that the cell partitions are disjoint and their union covers the entire target area $\mathcal{D}$.

Solving the optimization problem in (4.11) is challenging due to various reasons. First, the optimization variables $\mathcal{A}_i$, $\forall i \in \mathcal{M}$, are sets of continuous partitions (as we have a continuous area) which are mutually dependent. Second, to perfectly capture the spatial distribution of users, $f(x,y)$ is considered to be a generic function of x and y and, this leads to the complexity of the given two-fold integrations. In addition, due to the constraints given in (4.12), finding $\mathcal{A}_i$ becomes more challenging. To solve the optimization problem in (4.11), next, we model the problem by exploiting *optimal transport theory* [100].

4.3.1 Optimal Transport Theory: Preliminaries

Here, we present some primary results from optimal transport theory which will be used in the next subsection to derive the optimal cell partitions. Optimal transport theory goes back to the Monge's problem in 1781 which is stated as follows [100]. Given piles of sands and holes with the same volume, what is the best move (transport map) to entirely fill up the holes with the minimum total transportation cost. In general, this theory aims to find the optimal matching between two sets of points that minimizes the costs associated with the matching between the sets. These sets can be either discrete or continuous, with arbitrary distributions (weights). Mathematically, the Monge optimal transport problem can be written as follows. Given two probability distributions f_1 on $X \subset \mathbb{R}^n$, and f_2 on $Y \subset \mathbb{R}^n$, find the optimal transport map T from f_1 to f_2 that minimizes the following problem:

$$\min_{T} \int_{X} c\left(\boldsymbol{x}, T(\boldsymbol{x})\right) f_1(\boldsymbol{x})\mathrm{d}\boldsymbol{x}; \ T: X \to Y, \tag{4.15}$$

where $c(\boldsymbol{x}, T(\boldsymbol{x}))$ denotes the cost of transporting a unit mass from a location coordinate $\boldsymbol{x} \in X$ to a location $\boldsymbol{y} = T(\boldsymbol{x}) \in Y$. Also, as shown in Fig. 4.2, f_1 and f_2 are the source and destination probability distributions.

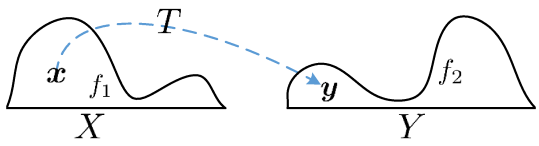


Figure 4.2: Transport map between two probability distributions.

Solving the Monge's problem is challenging due its high non-linear structure [100], and the fact that it does not necessarily admit a solution as each point of the source distribution must be mapped to only one location at the destination. However, Kantorovich relaxed this problem by using transport plans instead of maps, in which one point can go to multiple destination points. The relaxed Monge's problem is called Monge-Kantorovich problem which is written as [100]:

$$\min_{\pi} \int_{X\times Y} c\left(\boldsymbol{x},\boldsymbol{y}\right)\mathrm{d}\pi(\boldsymbol{x},\boldsymbol{y}), \tag{4.16}$$

$$\text{s.t.} \int_{X} \mathrm{d}\pi(\boldsymbol{x},\boldsymbol{y}) = f_1(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \int_{Y} \mathrm{d}\pi(\boldsymbol{x},\boldsymbol{y}) = f_2(\boldsymbol{y})\mathrm{d}\boldsymbol{y}, \tag{4.17}$$

where π represents the transport plan which is the probability distribution on $X \times Y$ whose marginals are f_1 and f_2.

The Monge-Kantorovich problem has two main advantages compared to the Monge's problem. First, it admits a solution for any semi-continuous cost function. Second, there is a dual formulation for the Monge-Kantorovich problem that can lead to a tractable solution. The duality theorem is stated as [100] and [132]:

Kantrovich Duality Theorem: Given the Monge-Kantorovich problem in (4.16) with two probability measures f_1 on $X \subset \mathbb{R}^n$, and f_2 on $Y \subset \mathbb{R}^n$, and any lower semi-continuous cost function $c(\boldsymbol{x},\boldsymbol{y})$, the following equality holds:

$$\min_{\pi} \int_{X\times Y} c\left(\boldsymbol{x},\boldsymbol{y}\right)\mathrm{d}\pi(\boldsymbol{x},\boldsymbol{y}) \tag{4.18}$$

$$= \max_{\varphi,\psi} \left\{ \int_{X} \varphi(\boldsymbol{x}) f_1(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_{Y} \psi(\boldsymbol{y}) f_2(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\,;\, \varphi(\boldsymbol{x}) + \psi(\boldsymbol{y}) \le c(\boldsymbol{x},\boldsymbol{y}),\, \forall(\boldsymbol{x},\boldsymbol{y}) \in X\times Y \right\}, \tag{4.19}$$

where $\varphi(\boldsymbol{x})$ and $\psi(\boldsymbol{y})$ are Kantorovich potential functions. As discussed in [100], this duality theorem provides a tractable framework for solving the optimal transport problems. In particular, we will use this theorem to tackle our optimization problem in (4.11).

We note that, in general, the solutions for the Monge-Kantorovich problem do not coincide with the Monge's problem. Nevertheless, when the source distribution, f_1, and the cost function are continuous, these two problems are equivalent [133]. In addition, the optimal transport map, $T : \boldsymbol{x} \to \boldsymbol{y}$, is linked with the optimal Kantorovich potential functions by:

$$T(\boldsymbol{x}) = \left\{\boldsymbol{y} | \varphi^*(\boldsymbol{x}) + \psi^*(\boldsymbol{y}) = c(\boldsymbol{x},\boldsymbol{y})\right\}, \tag{4.20}$$

where $\varphi^*(\boldsymbol{x})$ and $\psi^*(\boldsymbol{y})$ are the optimal potential functions corresponding dual formulation of the Monge-Kantorovich problem.

Given this optimal transport framework, we can solve our optimization problem in (4.11). In particular, we model this problem as a semi-discrete optimal transport problem in which the source measure (users' distribution) is continuous while the destination (UAVs' distribution) is discrete.

4.3.2 Optimal Cell Partitioning

Using optimal transport theory, we can find the optimal cell partitions, $\mathcal{A}_i$, for which the average total data service is maximized. In our model, users have a continuous distribution, and the locations of the UAVs can be considered as discrete points. Then, the optimal cell partitions are obtained by optimally mapping the users to the UAVs. In fact, given (4.11), the cell partitions are related to the transport map by [134]:

$$\left\{ T(\boldsymbol{v}) = \sum_{i \in \mathcal{M}} \boldsymbol{s}_i \mathbb{1}_{\mathcal{A}_i}(\boldsymbol{v}); \int_{\mathcal{A}_i} f(x,y) \mathrm{d}x \mathrm{d}y = \omega_i \right\}, \tag{4.21}$$

where $\omega_i = \frac{\alpha_i B T_i}{B \sum\limits_{k=1}^{M} \alpha_k T_k}$, as given in (4.12), is directly related to the hover time and the bandwidth of the UAVs. Also, $\mathbb{1}_{\mathcal{A}_i}(\boldsymbol{v})$ is the indicator function which is equal to 1 if $\boldsymbol{v} \in \mathcal{A}_i$, and 0 otherwise.

Therefore, the optimization problem in (4.11) can be cast within the optimal transport framework as follows. Given a continuous probability measure f of users, and a discrete probability measure $\Gamma = \sum\limits_{i \in \mathcal{M}} \omega_i \delta_{\boldsymbol{s}_i}$ corresponding to the UAVs, we must find the optimal transport map for which $\int_{\mathcal{D}} J\left(\boldsymbol{v}, T(\boldsymbol{v})\right) f(x,y) \mathrm{d}x \mathrm{d}y$ is minimized. In this case, $\delta_{\boldsymbol{s}_i}$ is the Dirac function, and J is the transportation cost function which is used in (4.11) and is given by:

$$J(\boldsymbol{v}, \boldsymbol{s}_i) = J(x, y, \boldsymbol{s}_i) = -\lambda_i \log_2\left(1 + \gamma_i(x,y)\right). \tag{4.22}$$

Clearly, the cost function, J, and the source distribution, f, are continuous. As a result, the Monge's problem coincides with the Monge-Kantorovich problem. Next, we propose a solution to (4.11) by exploiting the dual formulation of the Monge-Kantorovich problem.

Theorem 4. The optimization problem in (4.11) is equivalent to the following unconstrained maximization problem:

$$\max_{\psi_i, i \in \mathcal{M}} \left\{ F(\boldsymbol{\psi}^T) = \sum_{i=1}^{M} \psi_i \omega_i + \int_{\mathcal{D}} \psi^c(x,y) f(x,y) \mathrm{d}x \mathrm{d}y \right\}, \tag{4.23}$$

where $\boldsymbol{\psi}^T$ is a vector of variables $\psi_i, \forall i \in M$, and $\psi^c(x,y) = \inf\limits_i J(x,y,\boldsymbol{s}_i) - \psi_i$.

Proof. We use the Kantorovich duality theorem (4.18) in which $f(x,y)$ and $\Gamma = \sum_{i\in\mathcal{M}} \omega_i \delta_{\boldsymbol{s}_i}$ are two probability measures, and $J(\boldsymbol{v},\boldsymbol{s})$ is the cost function. Clearly, due to the continuity of $f(x,y)$ and $J(\boldsymbol{v},\boldsymbol{s}_i)$, the Monge's problem is equivalent to the Monge-Kantorovich problem.

$$\min_T \int_{\mathcal{D}} J\left(\boldsymbol{v}, T(\boldsymbol{v})\right) f(x,y)\mathrm{d}x\mathrm{d}y \tag{4.24}$$

$$= \max_{\varphi,\psi} \left\{ \int_{\mathcal{D}} \varphi(\boldsymbol{v}) f(x,y)\mathrm{d}x\mathrm{d}y + \int_S \psi(s) \sum_{i\in\mathcal{M}} \omega_i \delta_{s-s_i}\mathrm{d}s \,;\, \varphi(\boldsymbol{v}) + \psi(\boldsymbol{s}) \le J(\boldsymbol{v},\boldsymbol{s}) \right\} \tag{4.25}$$

$$= \max_{\varphi,\psi} \left\{ \int_{\mathcal{D}} \varphi(x,y) f(x,y)\mathrm{d}x\mathrm{d}y + \sum_{i=1}^{M} \psi(s_i)\omega_i \,;\, \varphi(x,y) + \psi(s_i) \le J(x,y,s_i),\, \forall i \in \mathcal{M} \right\}. \tag{4.26}$$

Note that, to maximize (4.26) given any ψ, we need to choose a maximum value for φ. Considering the fact that $\varphi(x,y) + \psi(\boldsymbol{s}_i) \le J(x,y,\boldsymbol{s}_i)$ must be satisfied for all $(x,y) \in \mathcal{D}$ and $\boldsymbol{s}_i \in S$, the maximum allowable value of φ is given by:

$$\varphi(x,y) = \psi^c(x,y) = \inf_{\boldsymbol{s}_i} J(x,y,\boldsymbol{s}_i) - \psi(\boldsymbol{s}_i), \tag{4.27}$$

where ψ^c is called the c-transform of ψ. Now, considering $\psi_i = \psi(\boldsymbol{s}_i)$, (4.26) and (4.27) lead to:

$$\min_T \int_{\mathcal{D}} J\left(\boldsymbol{v}, T(\boldsymbol{v})\right) f(x,y)\mathrm{d}x\mathrm{d}y = \max_{\psi_i, i\in\mathcal{M}} \left\{ F(\boldsymbol{\psi}^T) = \sum_{i=1}^{M} \psi_i\omega_i + \int_{\mathcal{D}} \psi^c(x,y) f(x,y)\mathrm{d}x\mathrm{d}y \right\}, \tag{4.28}$$

$$\psi^c(x,y) = \inf_i J(x,y,\boldsymbol{s}_i) - \psi_i. \tag{4.29}$$

As a result, the optimization problem in (4.11) is reduced to (4.23) with a set of M optimization variables, $\psi_i, \forall i \in \mathcal{M}$. This proves the theorem. □

Theorem 4 shows that the complex optimal cell partitioning problem in (4.11) can be transformed to a tractable optimization problem with M variables. In other words, by solving (4.23), one can use the optimal values of $\psi_i, \forall i \in \mathcal{M}$ to find the optimal cell partitions. Using Theorem 4, we can further proceed to solve (4.11) by presenting the following theorem:

Theorem 5. Given (4.23), F is a concave function of variables ψ_i, $i \in \mathcal{M}$. Also, we have:

$$\frac{\partial F}{\partial \psi_i} = \omega_i - \int_{\mathcal{D}_i} f(x,y)\mathrm{d}x\mathrm{d}y, \tag{4.30}$$

where $\mathcal{D}_i = \{(x,y) | J(x,y,\boldsymbol{s}_i) - \psi_i \le J(x,y,\boldsymbol{s}_j) - \psi_j, \forall j \ne i\}$.

Proof. Clearly, $\sum_{i=1}^{M} \psi_i \omega_i$ is a linear function of ψ_i. Also, given any $i \in \mathcal{M}$, $J(x, y, \boldsymbol{s}_i) - \psi_i$ is a linear function of ψ_i. Let $z(\boldsymbol{\psi}^T) = \inf_i J(x, y, \boldsymbol{s}_i) - \psi_i$ with $\boldsymbol{\psi}^T$ being a vector of all variables ψ_i, $i \in \mathcal{M}$. Then, we can observe that the hypograph of $z(\boldsymbol{\psi}^T)$, a set of points below $z(\boldsymbol{\psi}^T)$, is a convex set. Subsequently, considering the fact that a function is concave if and only if its hypograph is convex, we prove the concavity of $z(\boldsymbol{\psi}^T)$. Finally, since multiplying $z(\boldsymbol{\psi}^T)$ by a positive probability density function $f(x, y)$, and taking integration over (x, y) does not violate the concavity, F is also a concave function of $\boldsymbol{\psi}^T$.

To find the derivative of F with respect to ψ_i, we first compute $\frac{\partial \psi^c}{\partial \psi_i}$. Clearly, based on (4.29), we have:

$$\frac{\partial \psi^c}{\partial \psi_i} = \begin{cases} -1, & \text{if } \ J(x, y, \boldsymbol{s}_i) - \psi_i \leq J(x, y, \boldsymbol{s}_j) - \psi_j, \ \forall j \neq i, \\ 0, & \text{otherwise.} \end{cases} \tag{4.31}$$

Then, by defining $\mathcal{D}_i = \{(x, y) | J(x, y, \boldsymbol{s}_i) - \psi_i \leq J(x, y, \boldsymbol{s}_j) - \psi_j, \forall j \neq i\}$, the derivative of F, given in (4.23), will be:

$$\frac{\partial F}{\partial \psi_i} = \omega_i - \int_{\mathcal{D}_i} f(x, y) \mathrm{d}x \mathrm{d}y \, . \tag{4.32}$$

This proves the theorem. □

Theorem 5 shows the concavity of F as a function of $\boldsymbol{\psi}^T$. Thus, the optimal values for variables ψ_i, $\forall i \in \mathcal{M}$, can be obtained by maximizing F. Then, given the optimal ψ_i, $\forall i \in \mathcal{M}$, equations (4.20), and (4.21) are used to determine the optimal cell partitions corresponding to the optimization problem in (4.11). In this case, using the first derivative of F provided in (4.30), we propose a gradient-based method to determine the optimal vector $\boldsymbol{\psi}^T$ that leads to the optimal cell partitions. Here, using the gradient descent method is simple in terms of implementation and does not require computing the Hessian matrix of F which is needed in the Newton methods. In fact, given the intractable expression of F in (4.23), finding its second derivative is challenging and, thus, we adopt the gradient-based approach.

The proposed algorithm for finding the optimal cell partitions is shown as Algorithm 1 and proceeds as follows. The inputs are the distribution of users, hover times, locations of the UAVs, and $\rho > 0$ which is the threshold based on which the algorithm stops. In Algorithm 4, we first initialize vector $\boldsymbol{\psi}_t^T$ with t being the iteration number. Next, using (4.30), we compute $\nabla F(\boldsymbol{\psi}_t^T)$. In Step 6, we update $\boldsymbol{\psi}_t^T$ using step size ϵ_k. The appropriate step size at each iteration is determined through Steps 7 to 21. Here, we find the step size ϵ_k based on Armijo's rule which is a common approach for finding an appropriate step size in the gradient decent method [135]. In particular, Armijo's rule improves the convergence of the gradient descent method by adaptively adjusting the step size. In this case, the algorithm stops when the condition in 4 is not satisfied. Clearly, due to the concavity of F, the optimal solution to (4.23) is attained. Finally, based on the optimal vector $\boldsymbol{\psi}^T$, the optimal cell partitions are determined using Steps 24 and 25.

In summary, we proposed a framework for maximizing the average data service to ground users while considering some level fairness between the users. To this end, we used tools from optimal

Algorithm 2 Gradient method for optimal cell partitioning

Inputs: $f(x,y)$, ρ, τ_i, $\boldsymbol{s}_i$, $\forall i \in \mathcal{M}$.
Outputs: ψ_i^*, $\mathcal{A}_i$, $\forall i \in \mathcal{M}$.
Set initial values for $\boldsymbol{\psi}_t^T$, $(t=1)$.
while $\left\|\nabla F(\boldsymbol{\psi}_t^T)\right\|_2 > \rho$ **do**
    Set $k=1$, $\epsilon_1 = 1$.
    Update $\boldsymbol{\psi}_{t+1}^T = \boldsymbol{\psi}_t^T + \varepsilon_k \nabla F(\boldsymbol{\psi}_t^T)$, $k,t \in \mathbb{N}$.
    if $F(\boldsymbol{\psi}_t^T) < F(\boldsymbol{\psi}_{t+1}^T)$ **then**
        Go to Step 12 .
    else
        Go to Step 17 .
    end if
    while $F(\boldsymbol{\psi}_t^T) < F(\boldsymbol{\psi}_{t+1}^T)$ **do**
        $k \rightarrow k+1$.
        $\epsilon_k = 2^{k-1}\epsilon_1$.
        Update $\boldsymbol{\psi}_{t+1}^T = \boldsymbol{\psi}_t^T + \varepsilon_k \nabla F(\boldsymbol{\psi}_t^T)$, $k,t \in \mathbb{N}$.
    end while
    while $F(\boldsymbol{\psi}_t^T) > F(\boldsymbol{\psi}_{t+1}^T)$ **do**
        $k \rightarrow k+1$.
        $\epsilon_k = 2^{-k+1}\epsilon_1$.
        Update $\boldsymbol{\psi}_{t+1}^T = \boldsymbol{\psi}_t^T + \varepsilon_k \nabla F(\boldsymbol{\psi}_t^T)$, $k,t \in \mathbb{N}$.
    end while
    $t \rightarrow t+1$.
end while
$\psi_i^* = \boldsymbol{\psi}_t^T(i)$, $\forall i \in \mathcal{M}$.
$\mathcal{A}_i = \left\{(x,y) | J(x,y,\boldsymbol{s}_i) - \psi_i^* \leq J(x,y,\boldsymbol{s}_j) - \psi_j^*, \forall j \neq i\right\}$, $\forall i \in \mathcal{M}$.

transport theory to determine the optimal cell partition associated to each UAV that services the users in the cell partition.

4.3.3 Cell Partitioning in Uplink Case

For cell association problems, in general, analyzing the uplink can be intractable due to the challenges stemming from interference modeling. In particular, uplink interference depends on two key information: channel assignment strategy and cell partitions. Without such information (as in our case), one cannot write an exact expression for the interference. To analyze the uplink case, we can consider interference-free and interference scenarios, separately, as follows.

Interference-free scenario

In this case, we assume that the UAVs operate over different frequency bands which are divided among their associated users and, thus, there is no interference. The uplink SNR for a given user

located at (x, y) that connects to UAV i will be:

$$\gamma_i^{\text{UL}}(x,y) = \frac{P_u}{K_o {d_i}^2(x,y)\left[P_{\text{LoS},i}\mu_{\text{LoS}} + P_{\text{NLoS},i}\mu_{\text{NLoS}}\right]\sigma^2}, \tag{4.33}$$

where P_u is the transmit power of the user. Then, the uplink cell partitioning problem will be given by:

$$\min_{\mathcal{A}_i, i\in\mathcal{M}} \sum_{i=1}^{M} \int_{\mathcal{A}_i} -\left(\lambda_i \log_2\left(1+\gamma_i^{\text{UL}}(x,y)\right)\right) f(x,y)\mathrm{d}x\mathrm{d}y, \tag{4.34}$$

$$\text{s.t.} \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y = \omega_i, \ \ \forall i \in \mathcal{M}, \tag{4.35}$$

$$\mathcal{A}_l \cap \mathcal{A}_m = \emptyset, \ \ \forall l \neq m \in \mathcal{M}, \ \bigcup_{i\in\mathcal{M}} \mathcal{A}_i = \mathcal{D}. \tag{4.36}$$

where $\lambda_i = \frac{B}{N\alpha_i}\sum_{k=1}^{M} T_k$, and $\omega_i = \frac{\alpha_i T_i}{\sum_{k=1}^{M}\alpha_k T_k}$. Clearly, similar to the downlink case, the uplink cell association problem in (4.34) can be solved using Algorithm 1.

Interference case

In this scenario, UAVs share the same frequency band which is divided among the users associated to each UAV. Here, since it is not possible to compute the exact value of the interference power in the uplink, we provide an approximate expression for the average interference. To this end, we first determine cell partitions based on a weighted Voronoi diagram and, then, we compute the average interference power over the given cell partitions.

Clearly, for a user that connects to UAV i, interference stems from users located in all other cells, $\forall \mathcal{A}_k, k \neq i$. Now, the spatial probability density function of an interfering user located at partition $\mathcal{A}_k, k \neq i$ can be given by:

$$f_{\mathcal{A}_k}(x,y) = f(x,y|\,(x,y)\in\mathcal{A}_k) = \frac{f(x,y)}{P\left[(x,y)\in\mathcal{A}_k\right]} = \frac{f(x,y)}{\int_{\mathcal{A}_k} f(x,y)\mathrm{d}x\mathrm{d}y}. \tag{4.37}$$

Subsequently, the average interference power at UAV i from a user in partition $\mathcal{A}_k$ is:

$$I_{\mathcal{A}_k,i} = \int_{\mathcal{A}_k} P_u G_i(x,y) f_{\mathcal{A}_k}(x,y)\mathrm{d}x\mathrm{d}y, \tag{4.38}$$

where $G_i(x,y) = K_o d_i^2(x,y)\left[P_{\text{LoS},i}\mu_{\text{LoS},i} + P_{\text{NLoS},i}\mu_{\text{NLoS},i}\right]$ is the channel gain between the interfering user and UAV i. Hence, the average interference power at UAV i from all interfering users can be approximately given by:

$$I_i^{\mathrm{UL}} \approx \sum_{k \neq i} I_{\mathcal{A}_k, i}. \tag{4.39}$$

Note that, our approach for finding the average interference is analytically tractable. Moreover, this approach is not computationally demanding as it only needs to be done once, before solving our cell association problem. Clearly, the uplink SINR for a given user located at (x, y) that connects to UAV i will be:

$$\gamma_i^{\mathrm{UL}}(x, y) = \frac{P_u G_i(x, y)}{I_i^{\mathrm{UL}} + \sigma^2}. \tag{4.40}$$

Finally, (4.40) is used in our uplink cell association problem in (4.34) which can be solved using Algorithm 1. A more elaborate analysis on the uplink cell association problem can be subject of future work. In the next section, we investigate Scenario 2 in which the minimum average hover times of UAVs needed for completely servicing the users are determined.

4.4 Scenario 2: Minimum Hover Time For meeting Load Requirements

Our goal is to meet the load requirements of the ground users while minimizing the average hover times of the UAVs. In particular, given the demand of each user, we find the minimum required average hover time of the UAVs to completely serve the users. The hover time of each UAV depends on the distribution and load of the users, bandwidth allocation between users, and the cell partition that is assigned to the UAV. Next, we first derive an expression for the average hover time needed to serve any arbitrary cell partition under an optimal bandwidth allocation between the users. Then, we determine the optimal cell partitions for which the average hover times required for completely servicing the entire target area is minimized. We note that, given a specified partition $\mathcal{A}_i$, the hover time of the UAV i needed for serving the users depends on the bandwidth allocation strategy. Hence, we first derive the minimum average hover time of each UAV that can be attained by optimal bandwidth allocation to the users in the given cell partition.

Proposition 4. Let u be the load (in bits) of a user located at (x, y). The minimum average hover time of UAV i for serving partition $\mathcal{A}_i$ that can be achieved by optimally allocating the bandwidth to the users, is given by:

$$\tau_i = \int_{\mathcal{A}_i} \frac{Nu}{C_i^B(x, y)} f(x, y) \mathrm{d}x\mathrm{d}y + g_i \left(\int_{\mathcal{A}_i} f(x, y) \mathrm{d}x\mathrm{d}y \right), \tag{4.41}$$

where N is the total number of users, $C_i^B = B\log_2\left(1 + \gamma_i(x, y)\right)$, and $g_i(.)$ is the additional control time which is a function of the number of users in the cell partition.

Proof. Let $\mathcal{Y}$ be an arbitrary set of Y users in a cell partition $\mathcal{A}_i$. We denote the load and bandwidth allocated to user r by u_r and W_r. Then, the time needed to serve user r is given by:

$$t_r = \frac{u_r}{W_r E_r}, \tag{4.42}$$

where E_r is the spectral efficiency (bit/s/Hz) at the user's location. Clearly, to completely serve all the users in $\mathcal{A}_i$, the hover time of UAV i must be $\max t_r + g_{\mathcal{Y}}$, with $g_{\mathcal{Y}}$ being the additional control time. The hover time can be minimized by an optimal bandwidth allocation as follows:

$$\min_{W_r, r=1,\ldots,Y} \max t_r + g_{\mathcal{Y}}, \tag{4.43}$$

$$\text{s.t.} \sum_{r=1}^{Y} W_r = B. \tag{4.44}$$

The min-max problem in (4.43) can be transformed to:

$$\min Z + g_{\mathcal{Y}}, \tag{4.45}$$

$$\text{s.t. } Z \geq \frac{u_r}{W_r E_r}, \ \forall r \in \mathcal{Y}, \tag{4.46}$$

$$\sum_{r=1}^{Y} W_r = B. \tag{4.47}$$

Now, using (9.55) and (9.56), we have $Z \geq \frac{1}{B} \sum_{r=1}^{Y} \frac{u_r}{E_r}$. Hence, the minimum hover time under an optimal bandwidth allocation will be $\sum_{r=1}^{Y} \frac{u_r}{BE_r} + g_{\mathcal{Y}}$. Furthermore, it can be shown that $W_r = \frac{Bu_r}{E_r} / \sum_{k=1}^{Y} \frac{u_k}{E_k}$ is an optimal bandwidth allocation to user r. Clearly, $\sum_{r=1}^{Y} \frac{u_r}{BE_r}$ is also equal to the total time needed for sequentially serving the users using the entire bandwidth B. Therefore, given a cell partition $\mathcal{A}_i$ and the users' distribution, $f(x, y)$, the minimum average hover time of UAV i can be given by:

$$\tau_i = \int_{\mathcal{A}_i} \frac{Nu}{B\log_2\left(1 + \gamma_i(x, y)\right)} f(x, y) \mathrm{d}x\mathrm{d}y + g_i \left(\int_{\mathcal{A}_i} f(x, y) \mathrm{d}x\mathrm{d}y \right). \tag{4.48}$$

Finally, considering $C_i^B = B\log_2\left(1 + \gamma_i(x, y)\right)$, this proposition is proved. □

From (4.41), we can see that the hover time increases as the number of users increases. In fact, for a higher number users, both the total data transmission time and the additional control time increase. Moreover, the hover time increases as the load of the users increases. In particular, for an equal load of users, the hover time increases sublinearly by increasing the load as the control time does not depend on the load here. From (4.41), we can see that the hover time can be reduced by increasing the transmission rate. In addition, (4.41) implies that the UAV must hover for a longer time over a partition with higher users' density (i.e. $f(x, y)$).

Next, we minimize the total average hover time by solving the following optimization problem:

$$\min_{\mathcal{A}_i, i\in\mathcal{M}} \sum_{i=1}^{M} \int_{\mathcal{A}_i} \frac{Nu}{C_i^B(x,y)} f(x,y)\mathrm{d}x\mathrm{d}y + g_i\left(\int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y\right), \tag{4.49}$$

$$\text{s.t. } \mathcal{A}_l \cap \mathcal{A}_m = \emptyset, \ \ \forall l \neq m \in \mathcal{M}, \tag{4.50}$$

$$\bigcup_{i\in\mathcal{M}} \mathcal{A}_i = \mathcal{D}, \tag{4.51}$$

where the objective function in (4.49) represents the total average hover time needed for providing service for the target area.

Solving (4.49) is challenging as the optimization variables $\mathcal{A}_i$, $\forall i \in \mathcal{M}$, are sets of continuous partitions which are mutually dependent. Moreover, since g_i is a generic function of $\mathcal{A}_i$ and $f(x,y)$, this problem is intractable. Next, we use optimal transport theory to completely characterize the optimal solution. To this end, we first prove the existence of the solution to (4.49) for any semi-continuous function g_i, $\forall i \in \mathcal{M}$. Note that, in general, (4.49) does not necessary admit an optimal solution when the semi-continuity of g_i does not hold.

Proposition 5. The optimization problem in (4.49) generally admits an optimal solution.

Proof. Let $a_i = \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y$, then we also define a unit simplex as follows:

$$E = \left\{ \boldsymbol{a} = (a_1, a_2, ..., a_M) \in \mathbb{R}^M; \ \sum_{k=1}^{M} a_i = 1, a_i \geq 0, \forall i \in \mathcal{M} \right\}. \tag{4.52}$$

Clearly, given any vector $\boldsymbol{a}$, problem (4.49) can be considered as a classical semi-discrete optimal transport problem. In particular, considering $f(\boldsymbol{v}) = f(x,y)$, and $c(\boldsymbol{v}, \boldsymbol{s}_i) = \frac{u}{C_i^B(x,y)}$, (4.49) can be transformed to:

$$\min_{T} \int_{\mathcal{D}} c(\boldsymbol{v}, \boldsymbol{s}) f(\boldsymbol{v})\mathrm{d}\boldsymbol{v}, \ \ \boldsymbol{s} = T(\boldsymbol{v}), \tag{4.53}$$

where T is the transport map which is associated to cell partitions $\mathcal{A}_i$ by :

$$\left\{ T(\boldsymbol{v}) = \sum_{i=1}^{M} \boldsymbol{s}_i \mathbb{1}_{\mathcal{A}_i}(\boldsymbol{v}); \int_{\mathcal{A}_i} f(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = a_i \right\}. \tag{4.54}$$

Clearly, as discussed in Section 4.3, the optimal transport problem in (9.48) admits a solution. Hence, for any $\boldsymbol{a} \in E$, the problem in (4.49) has an optimal solution. Since E is a unit simplex in $\mathbb{R}^M$ which is a non-empty and compact set, the problem admits an optimal solution over the entire E. Thus, the proposition is proved. □

Next, we completely characterize the solution space of problem (4.49) which allows finding the optimal cell partitioning and the average hover time of each UAV.

Theorem 6. The optimal hover time of UAV i required to completely service the target area is given by:

$$\tau_i^* = \int_{\mathcal{A}_i^*} \frac{Nu}{C_i^B(x,y)} f(x,y)\mathrm{d}x\mathrm{d}y + g_i \left(\int_{\mathcal{A}_i^*} f(x,y)\mathrm{d}x\mathrm{d}y \right), \tag{4.55}$$

where $\mathcal{A}_i^*$ is the optimal cell partition given by:

$$\mathcal{A}_i^* = \left\{ (x,y) | \frac{Nu}{C_i^B(x,y)} + g_i'(a_i) \leq \frac{Nu}{C_j^B(x,y)} + g_j'(a_j), \forall j \neq i \in \mathcal{M} \right\}, \tag{4.56}$$

where $a_i = \int_{\mathcal{A}_i} f(x,y)\mathrm{d}x\mathrm{d}y$, and N is the total number of users.

Proof. See Appendix A.1. □

Using Theorem 6, we can find the optimal cell partitions as well as the minimum hover time needed to completely service the users. In fact, the target area is optimally partitioned in a way that the average hover time that the UAVs use to serve their users is minimized. Note that, for the special case where $g_i' = 0$, the result in (4.56) corresponds to the classical weighted Voronoi diagram. In this case, the users are assigned to the UAVs based on the maximum received signal strength criterion. Consequently, the users can be served with a maximum rate and, hence, the total required hover time is minimized. However, in general, the classical weighted Voronoi diagram is not optimal [127] as the effect of control time is ignored while generating cell partitions.

From (4.56), we can see that there is a mutual dependency between a_i and $\mathcal{A}_i, \forall i \in \mathcal{M}$. Therefore, solving (4.56) does not have an explicit form and, hence, an iterative-based approach is needed to find a solution to (4.56). Next, given the results of Theorem 6, we present an iterative algorithm based on [134] and shown in Algorithm 2, that solves (4.56) and finds the optimal cell partitions and the average hover time of each UAV. This algorithm guarantees the convergence to the optimal solution within a reasonable number of iterations [134]. In addition, this algorithm is practical to implement as its complexity grows linearly with the size of the area $\mathcal{D}$.

Algorithm 2 for finding the optimal cell partitions as well as the average hover times proceeds as follows. The inputs are load and distribution of the users, locations of the UAVs, control time function, and the number of iterations, L. Here, we use t to represent the iteration number. First, we generate initial cell partitions $\mathcal{A}_i^{(t)}$, and set $\phi_i^{(t)}(x,y) = 0$, $\forall i \in \mathcal{M}$, with $\phi_i^{(t)}(x,y)$ being a pre-defined parameter that will be used to update the cell partitions. Next, we update $\phi_i^{(t+1)}(x,y)$, and compute a_i in step 6. Then, in step 8, we update cell partitions by using (4.56). Finally, at the end of the iteration, the optimal cell partitions and the minimum average hover time of the UAVs are determined.

Algorithm 3 Iterative algorithm for optimal cell partitions and hover times

Inputs: $f(x,y)$, u, Z, g_i, $\boldsymbol{s}_i$, $\forall i \in \mathcal{M}$.
Outputs: $\mathcal{A}_i^*$, τ_i^*, $\forall i \in \mathcal{M}$.
Set $t=1$, generate an initial cell partitions $\mathcal{A}_i^{(t)}$, and set $\phi_i^{(t)}(x,y)=0$, $\forall i \in \mathcal{M}$.
while $t < Z$ **do**
    Compute $\phi_i^{(t+1)}(x,y) = \begin{cases} (1-1/t)\,\phi_i^{(t)}(x,y), & \text{if } (x,y) \in \mathcal{A}_i{}^{(t)}, \\ 1-(1-1/t)\left(1-\phi_i^{(t)}(x,y)\right), & \text{otherwise.} \end{cases}$
    Compute $a_i = \int_{\mathcal{D}} \left(1-\phi_i^{(t+1)}(x,y)\right) f(x,y)\mathrm{d}x\mathrm{d}y$, $\forall i \in \mathcal{M}$.
    $t \rightarrow t+1$.
    Update cell partitions using (4.56).
end while
$\mathcal{A}_i^* = \mathcal{A}_i^{(t)}$,
Compute τ_i^* using (4.41) based on $\mathcal{A}_i^*$, $\forall i \in \mathcal{M}$.

4.5 Simulation Results and Analysis

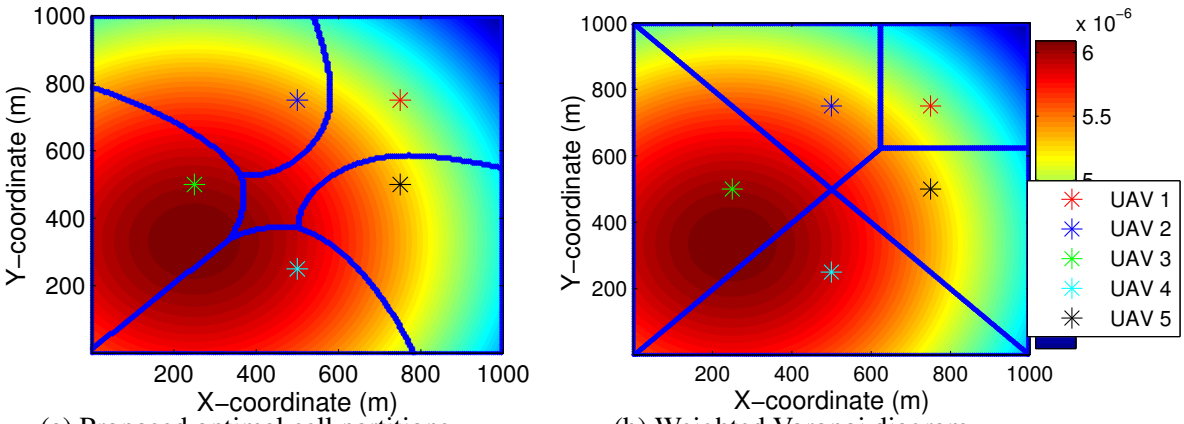
For our simulations, we consider a rectangular area of size $1000\,\mathrm{m} \times 1000\,\mathrm{m}$ in which the ground users are distributed according to a two-dimensional truncated Gaussian distribution which is suitable to model a hotspot area and is given by [128]:

$$f(x,y) = \frac{1}{\eta}\exp\Big[-\Big(\frac{x-\mu_x}{\sqrt{2}\sigma_x}\Big)^2\Big]\exp\Big[-\Big(\frac{y-\mu_y}{\sqrt{2}\sigma_y}\Big)^2\Big], \tag{4.57}$$

where $\eta = 2\pi\sigma_x\sigma_y \mathrm{erf}\left(\frac{L_x-\mu_x}{\sqrt{2}\sigma_x}\right)\mathrm{erf}\left(\frac{L_y-\mu_y}{\sqrt{2}\sigma_y}\right)$, and the size of the area is $L_x \times L_y$. Also, μ_x, σ_x, μ_y, and σ_y are the mean and standard deviation values of the x and y coordinates, and $\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}\mathrm{d}t$. In this case, ($\mu_x$, μ_y) represents the center of the hotspot, and the density of the users around the center is inversely proportional to the values σ_x and σ_y. In our simulations, we consider $\sigma_x = \sigma_y = \sigma_o$. Note that, although we consider the truncated Gaussian distribution of users, our analysis can also accommodate any other arbitrary distribution. Moreover, we deploy the UAVs based on a grid-based deployment with an altitude of 200 m. Unless stated otherwise, we consider a full interference scenario with an interference factor $\beta = 1$. Furthermore, for Scenario 1, we consider equal resource allocation factors (i.e. $\alpha_i = \alpha_j$, $\forall i,j \in \mathcal{M}$) . For the control time function, we consider $g_i(Na_i) = \alpha(Na_i)^2$, with α being an arbitrary constant factor. This function is a reasonable choice in our model as it is a superlinear function of the number of users and its value can be adjusted by factor α. However, any arbitrary Lipschitz continuous control function can also be considered in our model. The simulation parameters are listed in Table I. We compare our results, obtained based on the proposed optimal cell partitioning approach, with the classical weighted Voronoi diagram baseline. Note that, all statistical results are averaged over a large number of independent runs. Next, we present the results corresponding to Scenario 1 and Scenario 2, separately.

Table 4.1: Simulation parameters.

Parameter	Description	Value
f_c	Carrier frequency	2 GHz
P_i	UAV transmit power	0.5 W
N_o	Noise power spectral	-170 dBm/Hz
N	Number of ground users	300
μ_{LoS}	Additional path loss to free space for LoS	3 dB
μ_{NLoS}	Additional path loss to free space for NLoS	23 dB
B	Bandwidth	1 MHz
α	Control time factor	0.01
h	UAV's altitude	200 m
u	Load per user	10 Mb
μ_x, μ_y	Mean of the truncated Gaussian distribution	250 m, 330 m
b_1, b_2	Environmental parameters (dense urban)	0.36, 0.21 [15]



(a) Proposed optimal cell partitions. (b) Weighted Voronoi diagram.

Figure 4.3: Cell partitions associated to UAVs given the non-uniform spatial distribution of users.

4.5.1 Results for Scenario 1

Fig. 8.3 shows the proposed optimal cell partitions and the classical weighted Voronoi diagram. In this case, we consider 5 UAVs that provide service for the non-uniformly distributed ground users (truncated Gaussian distribution with $\sigma_o = 1000$ m). Moreover, in Scenario 1, we assume that the maximum hover time of each UAV is 30 minutes which corresponds to the typical hover time for quadcopter UAVs [136]. In Fig. 8.3, areas shown by a darker color have a higher population density. As we can see from Fig. 8.3b, the cell partitions associated with UAVs 3 and 4 have significantly more users than cell partition 1. Therefore, given the limited hover times, users located at cell partitions 3 and 4 cannot be fairly served by UAVs. However, in the proposed optimal cell partitioning case (obtained by Algorithm 1), the cell partitions change such that the average data service under a fair resource allocation constraint is maximized. For instance, as

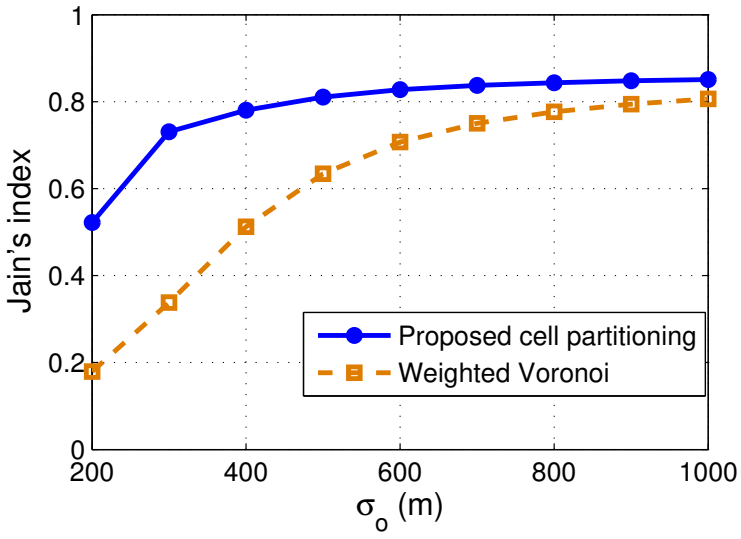


Figure 4.4: Jain's fairness index for average data service to users.

shown in Fig. 8.3a, the size of cell partitions 3 and 4 decreases compared to the weighted Voronoi diagram. As a result, the proposed cell partitions lead to a higher level of fairness among the users than the weighted Voronoi case.

To show how fairly the users can be served in different cases of cell partitioning, we use the Jain's fairness index. This metric can be applied to any performance metric such as rate or service load and it is maximized when all users receive an equal service. Here, we compute the Jain's index based on the data service that is offered to each user. The Jain's index is given by [137]:

$$F_{\text{Jain}}(l_1, l_2, ..., l_N) = \left(\sum_{i=1}^{N} l_i\right)^2 \times \left(N \sum_{i=1}^{N} {l_i}^2\right)^{-1}, \tag{4.58}$$

where N is the number of users, and l_i is the data service to user i. Clearly, $1/N \leq F_{\text{Jain}} \leq 1$, with $F_{\text{Jain}} = 1/N$ and $F_{\text{Jain}} = 1$ indicating the lowest and highest level of fairness.

Fig. 4.4 shows the Jain's fairness index for different values of σ_o which is given in (4.57). In this figure, as σ_o increases the spatial distribution of users becomes closer to a uniform distribution. As we can see from this figure, the minimum Jain's index corresponding to the proposed cell partitioning method is above 0.5. However, in the weighted Voronoi case, it can decrease to 0.18 for a highly non-uniform distribution of users with $\sigma_o = 200\,\text{m}$. This is due to the fact that, in the Voronoi case, users located in highly congested partitions receive lower service than the partitions with low number of users. In the proposed approach, however, the resources (hover time and bandwidth) are fairly shared between the users thus leading to a higher fairness index. From Fig. 4.4 we can also observe that, for higher values of σ_o (more uniform distribution), the fairness index for the proposed approach becomes closer to the weighted Voronoi case.

Fig. 4.5 shows the average number of users in each cell partition. Clearly, in the Voronoi case, the average number of users per cell significantly varies for different cell partitions. For instance,

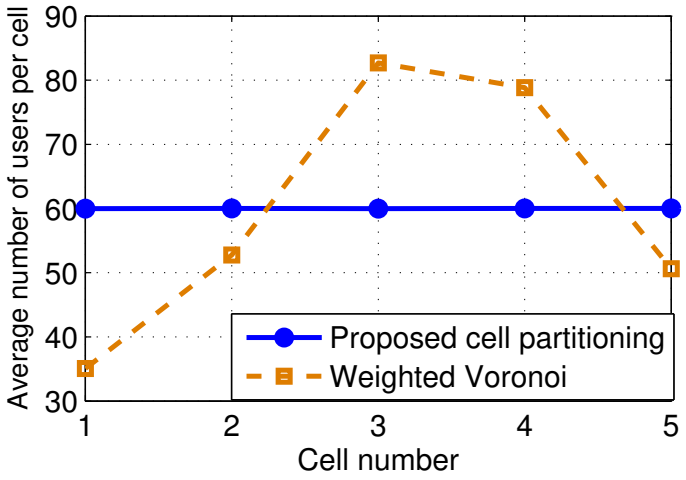


Figure 4.5: Average number of users per cell partition.

the average number of users in cell 3 is 2.5 times higher than cell 1. Consequently, compared to cell 3, users in cell 1 will receive lower data service from their associated UAV. However, in the proposed approach, the cell partitions associated with the UAVs are formed such that the number of users per cell be proportional to the bandwidth and hover time of the UAVs. In this case, given equal bandwidth and hover times of UAVs, the cell partitions contains an equal number of users. Therefore, our approach avoids generating unbalanced cell partitions and, hence, it leads to a higher level of fairness compared to the classical Voronoi approach.

Fig. 4.6 shows the average sum-logarithm (sum-log) of data service as the interference factor varies (for 5 UAVs, and $\sigma_o = 200$ m). The sum-log utility is a commonly used performance metric that accounts for both total system performance and individual fairness [129]. Clearly, our proposed approach outperforms the weighted Voronoi technique. Moreover, as the interference between UAVs decreases, the total data service that they can provide to the ground users increases. For instance, by decreasing β from 1 (full interference case) to 0 (interference-free), the sum-log data service increases by 14%.

In Fig. 4.7, we show how the total data service changes as the maximum hover time of the UAVs increases. As expected, by increasing the hover time of each UAV, the users can be served for a longer time and, hence, the total data service increases. Fig. 4.7 also compares the performance of deploying 5 UAVs versus 10 UAVs. Interestingly, we can see that the 5-UAVs case with 40 minuets hover time (for each UAV) outperforms the 10-UAVs case with 30 minuets hover time. As a result, in this case, deploying UAVs that has a 33% higher hover time is preferred than doubling the number of UAVs. In fact, increasing the number of UAVs results in a higher interference which reduces the maximum data service gain that can be typically achieved by using more UAVs. Therefore, depending on system parameters, using more capable UAVs (i.e. with longer flight time) to service ground users can be more beneficial than deploying more UAVs with shorter flight times.

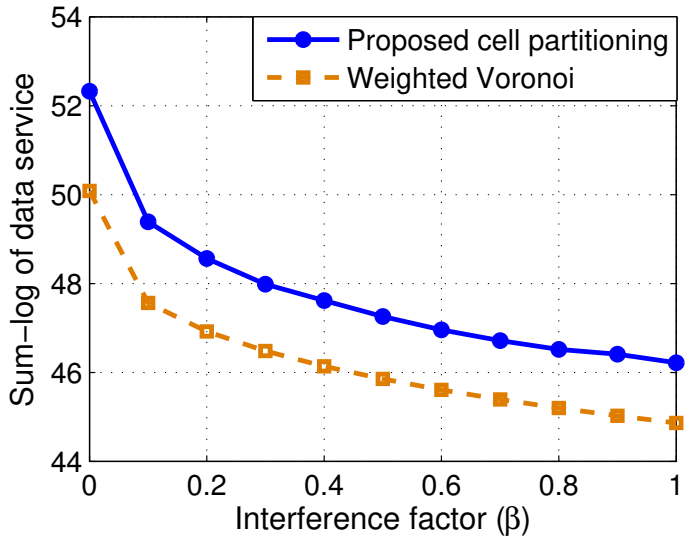


Figure 4.6: Average sum-log of data service as a function of the interference factor.

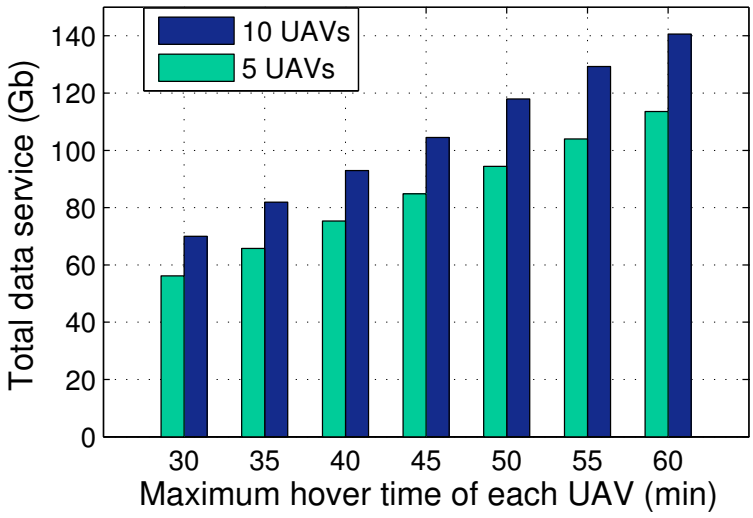


Figure 4.7: Total data service versus the maximum hover time of each UAV.

4.5.2 Results for Scenario 2

Here, we present the results for Scenario 2 in which the users are completely serviced using a minimum hover time. In this case, we consider a 10 Mb data service requirement for each user.

In Fig. 4.8, we show the total hover time versus the transmission bandwidth. Two bandwidth allocation schemes are considered, the optimal bandwidth allocation resulting from Proposition 2, and an equal bandwidth allocation. Clearly, by increasing the bandwidth, the total hover time required for serving the users decreases. In fact, a higher bandwidth can provide a higher the

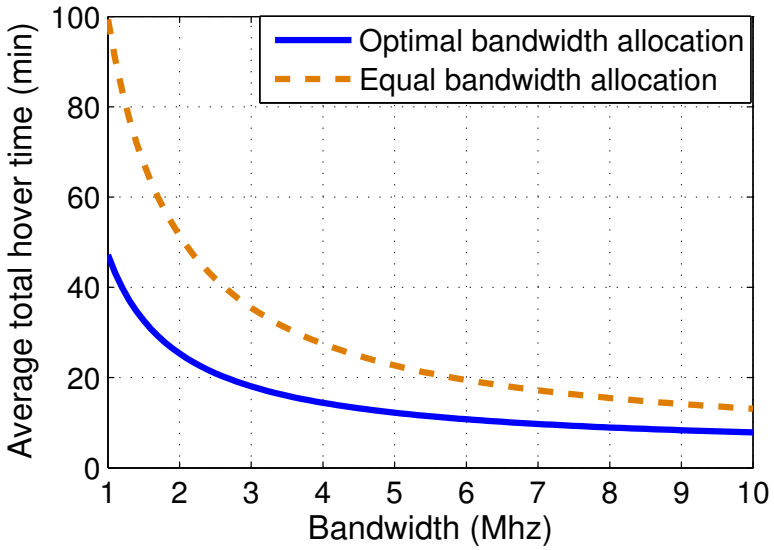


Figure 4.8: Average hover time versus bandwidth.

transmission rate and, hence, users can be serviced within a shorter time duration. From Fig. 4.8, we can see that, the optimal bandwidth allocation scheme can yield a 51% hover time reduction compared to the equal bandwidth allocation. This is due the fact that, according to Proposition 2, by optimally assigning the bandwidth to each user based on its demand and location, the total hover time of UAVs can be minimized.

Fig. 4.9 shows the average total hover time of the UAVs as the number of UAVs varies. This result corresponds to the interference-free scenario in which the UAVs operate on different frequency bands. Hence, the total bandwidth usage linearly increases by increasing the number of UAVs. From Fig. 4.9, we can see that the total hover time decreases as the number of UAVs increases. A higher number of UAVs corresponds to a higher number of cell partitions. Hence, the size of each cell partition decreases and the users will have a shorter distance to the UAVs. In addition, lower control time is required during serving a smaller and less congested cell. In fact, increasing the number of UAVs leads to a higher transmission rate, and lower control time thus leading to a lower hover time. For instance, as shown in Fig. 4.9, when the number of UAVs increases from 2 ot 6, the total hover time decreases by 53%. Nevertheless, deploying more UAVs in interference-free scenario results in a higher bandwidth usage. Therefore, there is a fundamental tradeoff between the hover time of UAVs and the bandwidth efficiency.

Fig. 4.10 shows the impact of control time on the total hover time for the proposed cell partitioning, as a result of Theorem 3 and the weighted Voronoi diagram. In both cases, we use the optimal bandwidth allocation scheme. Clearly, as the control time factor, α, increases, the total hover time also increases. From Fig. 4.10, we can see that, using our proposed optimal cell partitioning approach, the average total hover time can be reduced by around 20% compared to weighted Voronoi case. This is due to the fact that, unlike the weighted Voronoi, our approach also minimizes the control time while generating the cell partitions. We note that, the hover time difference between

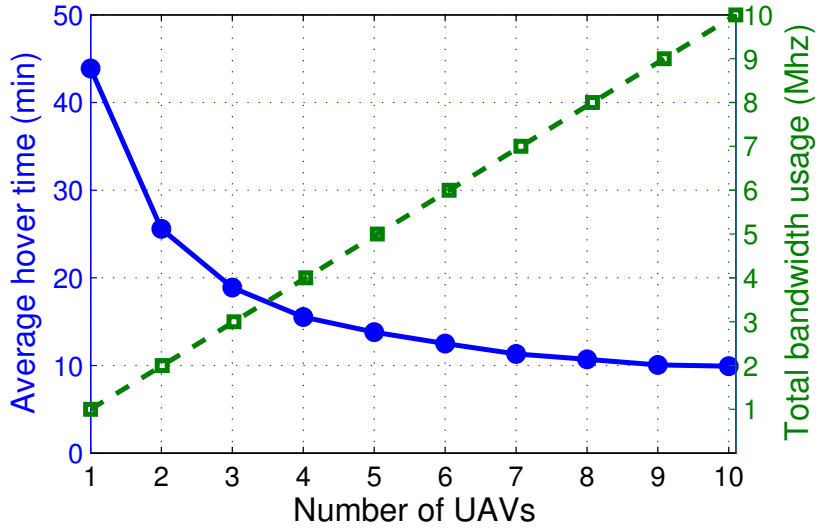


Figure 4.9: Average hover time versus number of UAVs and bandwidth usage.

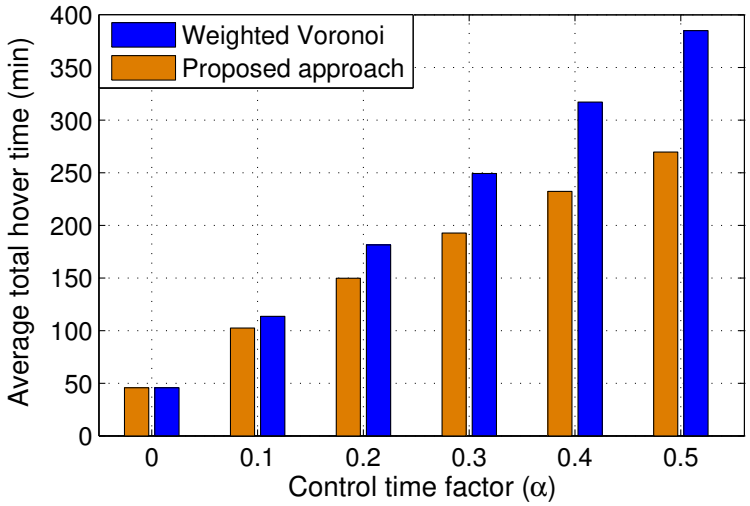


Figure 4.10: Average hover time versus control time factor (α) for $\sigma_o = 200$ m.

these two cases increases as α increases. In particular, as shown in Fig. 4.10, our approach yields around 32% hover time reduction when $\alpha = 0.5$.

In Fig. 4.11, we show the impact of interference on the hover time of UAVs. Clearly, the total hover time increases as the interference between the UAVs increases. This is due to the fact that a lower SINR leads to a lower transmission rate and, hence, a given UAV needs to hover for a longer time in order to completely service its users. For instance, the average hover time in the full interference case ($\beta = 1$) is 4.5 times larger than the interference-free case in which $\beta = 0$. Therefore, one can significantly reduce the hove time of UAVs by adopting interference mitigation techniques such as using orthogonal frequencies and scheduling of UAVs.

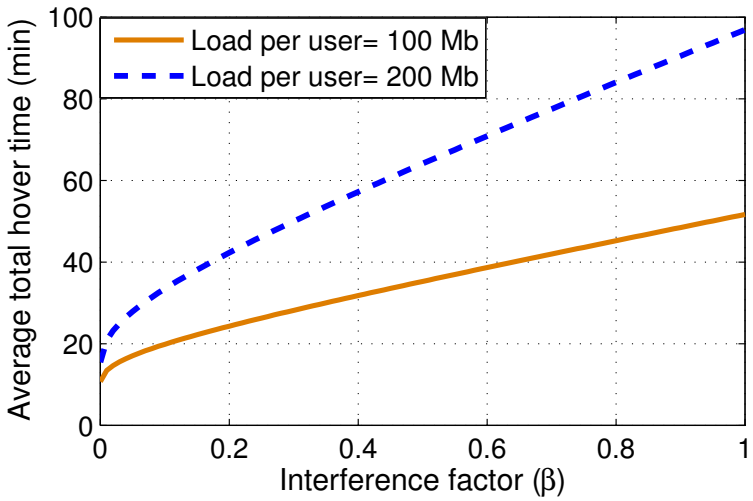


Figure 4.11: Average hover time versus interference factor.

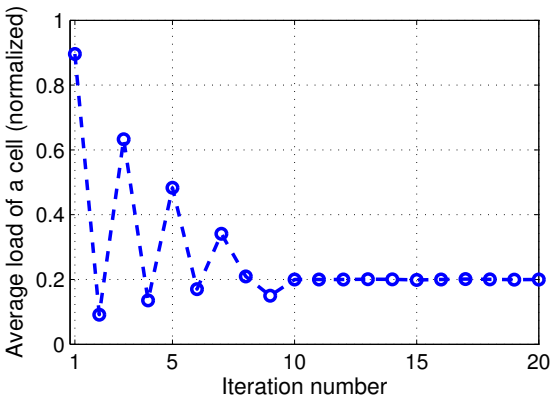


Figure 4.12: Convergence of Algorithm 1.

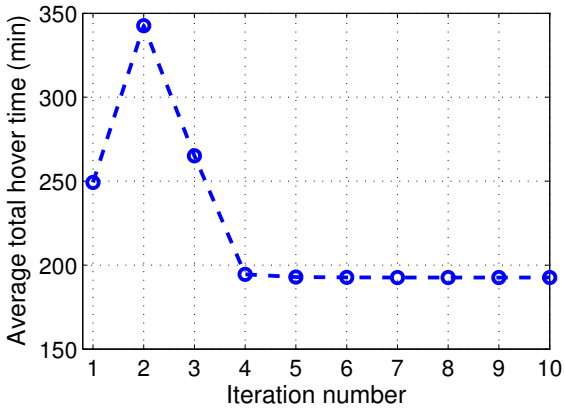


Figure 4.13: Convergence of Algorithm 2.

Finally, in Figures 4.12 and 4.13, we show the convergence of our proposed approach in Algorithms 1 and 2. Clearly, these algorithms converge after a reasonable number of iterations.

4.6 Summary

In this chapter, we have proposed a novel framework for optimizing UAV-enabled wireless networks while taking into account the flight time constraints of UAVs. In particular, we have investigated two UAV-based communication scenarios. First, given the maximum possible hover times of UAVs, we have maximized the average data service to the ground users under a fair resource allocation policy. To this end, using tools from optimal transport theory, we have determined the optimal cell partitions associated with the UAVs. In the second scenario, given the load require-

ments of users, we have minimized the average hover time of UAVs needed to completely serve the users. In this case, we have derived the optimal cell partitions as well as the optimal bandwidth allocation to the users that lead to the minimum hover time. The results have shown that, using our proposed cell partitioning approach, the users receive higher fair data service compared to the classical Voronoi case. Also, our results for Scenario 2 have revealed that the average hover time of UAVs can be significantly reduced by using the proposed approach.

4.7 Appendix A

A.1 Proof of Theorem 3

As shown in Proposition 5, there exist optimal cell partitions $\mathcal{A}_i$, $i \in \mathcal{M}$ which are solutions to the optimization problem in (4.49). Now, we consider two optimal partitions $\mathcal{A}_l$ and $\mathcal{A}_m$, and a point $\boldsymbol{v}_o = (x_o, y_o) \in \mathcal{A}_m$. Also, let $B_\epsilon(\boldsymbol{v}_o)$ be the intersection of $\mathcal{A}_m$ with a disk that has a center v_o and radius $\epsilon > 0$. To characterize the optimal solution of (4.49), we first generate new cell partitions $\widehat{\mathcal{A}}_i$ (a variation of optimal partitions) as follows:

$$\begin{cases} \widehat{\mathcal{A}}_m = \mathcal{A}_m \backslash B_\varepsilon(\boldsymbol{v}_o), \\ \widehat{\mathcal{A}}_l = \mathcal{A}_l \cup B_\varepsilon(\boldsymbol{v}_o), \\ \widehat{\mathcal{A}}_i = \mathcal{A}_i, \quad i \neq l, m. \end{cases} \tag{4.59}$$

Also, let $a_\varepsilon = \int_{B_\varepsilon(\boldsymbol{v}_o)} f(x,y)\mathrm{d}x\mathrm{d}y$, and $\widehat{a}_i = \int_{\widehat{\mathcal{A}}_i} f(x,y)\mathrm{d}x\mathrm{d}y$. Considering the optimality of $\mathcal{A}_i$, $i \in \mathcal{M}$, we have:

$$\begin{aligned}
&\sum_{i\in\mathcal{M}} \int_{\mathcal{A}_i} \left[\frac{Nu}{C_i^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_i(a_i) \\
&\overset{(a)}{\leq} \sum_{i\in\mathcal{M}} \int_{\widehat{\mathcal{A}}_i} \left[\frac{Nu}{C_i^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_i(\widehat{a}_i), \\
&\int_{\mathcal{A}_l} \left[\frac{Nu}{C_l^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_l(a_l) + \int_{\mathcal{A}_m} \left[\frac{Nu}{C_m^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_m(a_m) \\
&\leq \int_{\mathcal{A}_l \cup B_\varepsilon(v_o)} \left[\frac{Nu}{C_l^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_l(a_l + a_\epsilon) \\
&+ \int_{\mathcal{A}_m \backslash B_\varepsilon(v_o)} \left[\frac{Nu}{C_m^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_m(a_m - a_\epsilon) \\
&\int_{B_\varepsilon(v_o)} \left[\frac{Nu}{C_m^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_m(a_m) - g_m(a_m - a_\epsilon) \\
&\leq \int_{B_\varepsilon(v_o)} \left[\frac{Nu}{C_l^B(x,y)}\right] f(x,y)\mathrm{d}x\mathrm{d}y + g_l(a_l + a_\epsilon) - g_l(a_l),
\end{aligned} \tag{4.60}$$

where (a) comes from the fact that $\mathcal{A}_i$ is optimal and, hence, any variation of that ($\widehat{\mathcal{A}}_i$) cannot lead to a better solution.Now, we multiply both sides of the inequality in (4.60) by $\frac{1}{a_\varepsilon}$. Then, we take the limit when $\varepsilon \to 0$, and we use the following equality:

$$\lim_{\varepsilon\to 0}\frac{1}{a_\varepsilon}\int_{B_\varepsilon(v_o)}\left[\frac{Nu}{C_l^B(x,y)}\right]f(x,y)\mathrm{d}x\mathrm{d}y = \lim_{\varepsilon\to 0}\frac{\int_{B_\varepsilon(v_o)}\left[\frac{Nu}{C_l^B(x,y)}\right]f(x,y)\mathrm{d}x\mathrm{d}y}{\int_{B_\varepsilon(v_o)}f(x,y)\mathrm{d}x\mathrm{d}y}$$
$$= \frac{Nu}{C_l^B(x_o,y_o)}. \tag{4.61}$$

Subsequently, following from (4.60), we have:

$$\frac{Nu}{C_m^B(x_o,y_o)} + g_l'(a_m) \le \frac{Nu}{C_l^B(x_o,y_o)} + g_l'(a_l). \tag{4.62}$$

Note that, (4.62) provides the condition under which a point (x_o, y_o) is assigned to partition m rather than l. Therefore, the optimal cell partitions can be characterized as:

$$\mathcal{A}_i^* = \left\{(x,y)|\ \frac{Nu}{C_i^B(x,y)} + g_i'(a_i) \le \frac{Nu}{C_j^B(x,y)} + g_j'(a_j), \forall j \ne i \in \mathcal{M}\right\}. \tag{4.63}$$

Finally, using (4.41), the optimal average hover time of UAV i is:

$$\tau_i^* = \int_{\mathcal{A}_i^*}\frac{Nu}{C_i^B(x,y)}f(x,y)\mathrm{d}x\mathrm{d}y + g_i\left(\int_{\mathcal{A}_i^*}f(x,y)\mathrm{d}x\mathrm{d}y\right). \tag{4.64}$$

Chapter 5

Fundamental Performance Analysis of Unmanned Aerial Vehicle with Terrestrial Device-to-Device Network

5.1 Background, Related Works, and Contributions

As we discussed in previous chapters, owing to their agility and mobility of UAVs, they can be quickly and efficiently deployed to support cellular networks and enhance their QoS. On the one hand, UAV-based aerial base stations can be deployed to enhance the wireless capacity and coverage at temporary events or hotspots such as sport stadiums and outdoor events. On the other hand, they can be used in public safety scenarios to support disaster relief activities and to enable communications when conventional terrestrial networks are damaged [6].

The most significant existing body of work on UAV communications focuses on the air-to-ground channel modeling [15, 49, 50, 138]. For instance, in [15] and [50], the probability of line of sight (LoS) for air-to-ground communication as a function of the elevation angle and average height of buildings in a dense urban area was derived. The air-to-ground path loss model has been further studied in [138] and [49]. As discussed in [49], due to path loss and shadowing, the characteristics of the air-to-ground channel are shown to depend on the height of the aerial base stations.

To address the UAV deployment challenge, the authors in [8] derived the optimal altitude enabling a single, static UAV to achieve a maximum coverage radius. However, in this work, the authors simply defined a deterministic coverage by comparing the path loss with a specified threshold and did not consider the coverage probability. The work in [9] extends the results of [8] to the case of two UAVs while considering interference between the UAVs. In [29], the authors studied the optimal placement of UAVs for public safety communications in order to enhance the coverage performance. However, the results presented in [29] are based on simulations and there is no significant analytical analysis. The use of UAVs for supplementing existing cellular infrastructure was

discussed in [30] which provides a general view of practical considerations for integrating UAVs with cellular networks. The work in [13] considered the use of UAVs to compensate for the cell overload and outage in cellular networks. However, [30] did not provide any analysis on the coverage performance of UAVs and their optimal deployment methods. In [25], the authors investigated how to optimally move UAVs for improving connectivity of ad hoc networks. However, [25] only focused on an ad-hoc network and assumed that the UAV have complete information about the location of nodes. In [26], considering static ground users, the optimal trajectory and heading of UAVs equipped with multiple antennas for ground to air uplink scenario was derived.

For scenarios in which there is limited or no infrastructure support, beyond the use of UAVs, there has been considerable recent studies on D2D communications between wireless users over the licensed spectrum [139]. Such D2D communications has been shown improve coverage and capacity of existing wireless networks such as cellular systems. In particular, in hotspot areas or public safety scenarios, D2D will allow users to communicate directly with one another without significant infrastructure. D2D communications are typically deployed using underlaid transmission links which reuse existing licensed spectrum resources [140]. Therefore, deploying a UAV over a spectrum band that must be shared with an underlaid D2D network will introduce important interference management challenges. In the literature, there are some studies on the coexistence of the underlaid D2D and cellular communications with a single base station [141]. Furthermore, the authors in [142] exploited the interplay between the massive MIMO and underlaid D2D communications for a single cell case. The authors in [143] extended the previous work on the D2D/massive MIMO coexistence to the multi-cell scenario. However, none of theses prior work studied the coexistence of UAVs and underlaid D2D communications. In particular, a comprehensive analytical analysis to evaluate this coexistence in terms of different performance metrics, such as coverage and rate, is lacking in the current state-of-the-art [8, 25, 141–143].

Compared to the previous studies on the coexistence of D2D and cellular networks such as [142] and [143], the presence of an aerial UAV base station along with D2D links introduces new challenges. First, the channel modeling between the UAV and ground users will no longer be a classical fading channel, instead, it will be based on probabilistic LoS and NLoS links [15, 50], while the channel between a base station and the users will still follow a Rayleigh fading model. Second, unlike conventional, fixed base stations, the height of a UAVs is adjustable and this impacts the channel characteristics and the coverage performance. Third, the potential mobility of a UAV introduces new dimensions to the problem and the impact of such mobility on D2D and network performance must be analyzed. The prior studies on UAVs such as [8,9,13,15,25,29,30,49,50,138] have not addressed the third challenge. More specifically, the interplay between UAVs and D2D communications and the existing challenges and tradeoffs have not been investigated in these literature. This chapter will provide *the fundamental analysis* on the performance of UAV communication in the presence of underlaid D2D links.

In this chapter, we analyze the coverage and rate performance of UAV-based wireless communication in the presence of underlaid D2D communication links. In particular, we consider a network in which a single UAV must provide downlink transmission support to a number of users within a given area. In this area, a subset of the devices is also engaged in D2D transmissions that operate in

an underlay fashion over the UAV's transmission. We consider two types of users, namely downlink users (DUs) which receive data from the UAV, and D2D users which communicate directly with one another. Here, the UAV must communicate with the DUs while taking into account the potential interference stemming from the underlaid D2D transmissions. For this network, we analyze two key cases: static UAV and mobile UAV. Using tools from stochastic geometry, for both scenarios, we derive the average downlink coverage probabilities for DUs and D2D users and we analyze the impact of the UAV altitude and density of the D2D users on the overall performance. For the static case, we find the optimal values for the UAV altitude which leads to a maximum coverage probability for DUs. In addition, considering both DUs and D2D users, an optimal altitude which maximizes the system sum-rate is computed. Our results demonstrate that the optimal UAV altitude decreases as the density of D2D users increases. The results show that a maximum system sum-rate can be achieved if the UAV altitude is appropriately adjusted based on the D2D users' density. Furthermore, for a given UAV altitude, we show that an optimal value for the number of D2D users that maximizes the system sum-rate exists.

For the mobile UAV case, we assume that the UAV can travel over the area while stopping at some given locations in order to serve the downlink users. Using the disk covering problem, we find a minimum number of stop points that the UAV needs to to completely cover the area. This can be interpreted as the fastest way to cover the whole area with a minimum required transmit power. In addition, we analyze the tradeoff between the number of stop points, which is considered as delay here, and the coverage probability for the downlink users. Moreover, considering retransmissions at different time instances, we derive the overall outage probability for the D2D communications. We show that, in order to enhance the coverage for DUs, the UAV should stop in more locations over the target area which can, in turn, lead an increased delay for DUs and higher outage probability for D2D users. For example, our results show that for a given density of D2D users, to increase the DU coverage probability from 0.4 to 0.7, the number of stop points should be increased from 5 to 23. Furthermore, the number of stop points is shown to significantly depend on the number of D2D users. For instance, if the average number of D2D users in the area increases from 50 to 100, in order to maintain the DUs' coverage requirement, the number of stop points should be increased from 20 to 55.

The rest of this chapter is organized as follows. Section 5.2 presents the system model and describes the air-to-ground channel model. In Section 5.3, coverage probabilities for DUs and D2D users are derived for a single static UAV. Section 5.4 presents the performance evaluation for the mobile UAV which is used to provide full coverage for the target area. Section 5.5 presents the simulation results while Section 5.6 draws some conclusions.

5.2 System Model

Consider a circular area with a radius R_c in which a number of wireless users are deployed. In this area, a UAV (at low altitude platform) is deployed to act as a flying base station and to serve a subset of those users. In this network, the users are divided into two groups: downlink users located

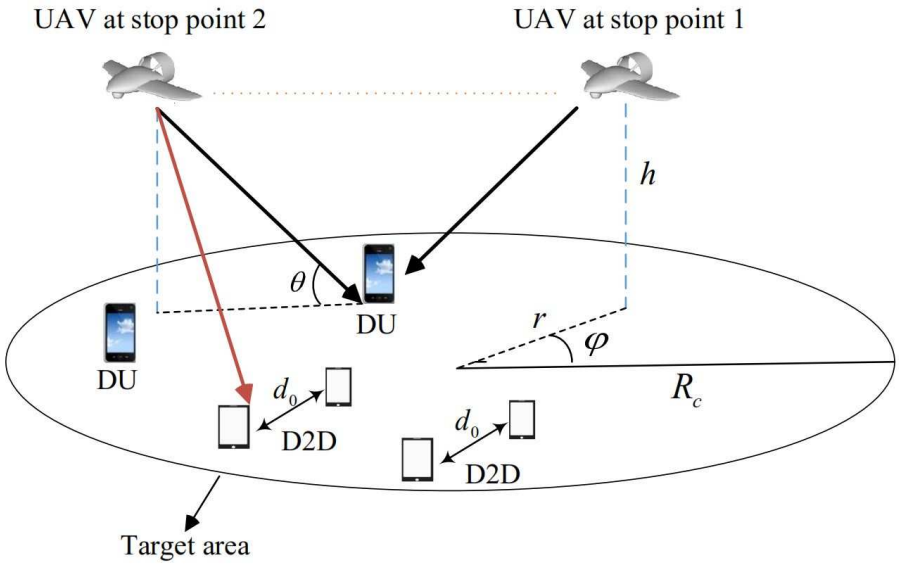


Figure 5.1: Network model including a UAV, downlink users and D2D.

uniformly in the cell with density λ_{du} (number of users per m^2), and D2D users whose distribution follows homogeneous Poisson point processes (PPP) Φ_{B} [144] with density of λ_d (number of D2D pairs per m^2). Note that, the average number of users in a given area is equal to the density of the users multiplied by the size of the area. Here, we focus on the downlink scenario for the UAV and we assume that the D2D users communicate in an underlay fashion. Furthermore, we assume that a D2D receiver connects to its corresponding D2D transmitter pair located at a fixed distance away from it in an isotropic direction [141]. Therefore, the received signals at the D2D receiver include the desired signal from the D2D transmitter pair and interference from the UAV and other D2D transmitters. A downlink user, on the other hand, receives the desired signal from the UAV but it also experiences interference from all the D2D transmitters. It should be noted that, in our model, the UAV provides service for downlink users (DUs) located inside a given, finite area with radius R_c. Nonetheless, we assume that the D2D users are spatiality distributed according to a PPP over an infinite area. In other words, each user receives interference from an infinite number of D2D transmitters. This is a typical assumption in PPP analysis which ensures that, the average amount of received interference from D2D transmitters does not depend on the location of the users [142, 145], and [146].

The signal to interference plus noise ratio (SINR) expression for a D2D receiver is:

$$\gamma_d = \frac{P_{r,d}}{I_d^c + I_u + N}, \tag{5.1}$$

where $P_{r,d}$ is the received signal power from the D2D transmitter, I_d^c is the total interference from other D2D users, I_u is the interference from the UAV, and N is the noise power. Moreover, we have:

$$P_{r,d} = P_d d_0^{-\alpha_d} g_0, \tag{5.2}$$

$$I_d^c = \sum_{i \neq 0} P_d {d_i}^{-\alpha_d} g_i, \tag{5.3}$$

$$I_d = \sum_{i} P_d {d_i}^{-\alpha_d} g_i, \tag{5.4}$$

where the index $i = 0$ is used for the selected D2D transmitter/receiver pair, g_0 and g_i are, respectively, the channel gains between a D2D receiver and its corresponding D2D transmitter, and the i^{th} interfering D2D transmitters. For the D2D transmission, we assume a Rayleigh fading channel model [141, 143] and [147]. P_d is the D2D transmit power which is assumed to be fixed and equal for all the users, d_i is the distance between a D2D receiver and the i^{th} D2D transmitter, d_0 is the fixed distance between the D2D receiver and transmitter of the selected D2D pair, and α_d is the path loss exponent between D2D users. Note that the received signal powers as well as the noise power are normalized by a path loss coefficient.

The SINR expression for a DU user that can connect to the UAV is:

$$\gamma_u = \frac{P_{r,u}}{I_d + N}, \tag{5.5}$$

where $P_{r,u}$ is the received signal power from the UAV and I_d is the total interference power from D2D transmitters. Finally, the SINR-based coverage probability for the downlink users and the D2D users is given by:

$$P_{\text{cov},du}(\beta) = \mathbb{P}\left[\gamma_u \geq \beta\right], \tag{5.6}$$

$$P_{\text{cov},d}(\beta) = \mathbb{P}\left[\gamma_d \geq \beta\right], \tag{5.7}$$

where γ_u and γ_d are, respectively, the SINR values at the location of the downlink users and the D2D users, and β is the SINR threshold.

5.2.1 Air-to-ground channel model

As discussed in [15] and [8], the ground receiver can receive three groups of signals including LoS, strong reflected non-line-of-sight (NLoS) signals, and multiple reflected components which cause multipath fading. These groups can be considered separately with different probabilities of occurrence as shown in [49] and [15]. Typically, it is assumed that the received signal is categorized in only one of those groups [8]. Each group has a specific probability of occurrence which is a function of environment, density and height of buildings, and elevation angle. Note that the probability of having the multipath fading is significantly lower than the LoS and NLoS groups [8]. Therefore, the impact of small scale fading can be neglected in this case [15]. One common approach for modeling air-to-ground propagation channel is to consider LoS and NLoS components along with their occurrence probabilities separately as shown in [15] and [49]. Note that for NLoS connections due to the shadowing effect and the reflection of signals from obstacles, path loss is higher than in LoS. Hence, in addition to the free space propagation loss, different excessive

path loss values are assigned to LoS and NLoS links. Depending on the LoS or NLoS connection between the user and UAV, the received signal power at the user location is given by [8]:

$$P_{r,u} = \begin{cases} P_u |X_u|^{-\alpha_u} & \text{LoS link}, \\ \eta P_u |X_u|^{-\alpha_u} & \text{NLoS link}, \end{cases} \tag{5.8}$$

where P_u is the UAV transmit power, $|X_u|$ is the distance between a generic user and the UAV, α_u is the path loss exponent over the user-UAV link, and η is an additional attenuation factor due to the NLoS connection. Here, the probability of LoS connection depends on the environment, density and height of buildings, the location of the user and the UAV, and the elevation angle between the user and the UAV. The LoS probability can be expressed as follows [8]:

$$P_{\text{LoS}} = \frac{1}{1 + C\exp(-B\left[\theta - C\right])}, \tag{5.9}$$

where C and B are constant values which depend on the environment (rural, urban, dense urban, or others) and θ is the elevation angle. Clearly, $\theta = \frac{180}{\pi} \times \sin^{-1}\left(\frac{h}{|X_u|}\right)$, $|X_u| = \sqrt{h^2 + r^2}$ and also, probability of NLoS is $P_{\text{NLoS}} = 1 - P_{\text{LoS}}$.
As observed from (9), the LoS probability increases as the elevation angle between the user and UAV increases.

Given this model, we will consider two scenarios: *a static UAV* and *a mobile UAV*. For each scenario, we will derive the coverage probabilities and average rate for DUs and D2D users. Once those metrics are derived, considering the D2D users density, we obtain optimal values for the UAV altitude that maximize the coverage probability and average rate.

5.3 Network with a Static UAV

In this section, we evaluate the coverage performance of the users in the scenario in which one UAV located at the altitude of h in the center of the area serves the downlink users in the presence of underlaid D2D communications. It can be shown that, for a uniform distribution of users over the given area, palcing the UAV in the center of the area can maximize the coverage probability of the downlink users.

5.3.1 Coverage probability for D2D users

Consider a D2D receiver located at (r, φ), where r and φ are the radius and angle in a polar coordinate system assuming that the UAV is located at the center of the area of interest. The distance between the D2D transmitter and its corresponding receiver is fixed and it is denoted by d_0. In this case, for underlaid D2D communication, the coverage probability for the D2D users can be derived as follows:

Theorem 7. The coverage probability for a D2D receiver, at the location (r, ϕ), connecting to its D2D transmitter located at a distance d_0 away from it, is given by:

$$P_{\text{cov},d}(r,\varphi,\beta) = \exp\left(\frac{-2\pi^2\lambda_d\beta^{2/\alpha_d}d_0^2}{\alpha_d \sin(2\pi/\alpha_d)} - \frac{\beta d_0^{\alpha_d} N}{P_d}\right) \times \left[P_{\text{LoS}} \exp\left(\frac{-\beta d_0^{\alpha_d} P_u |X_u|^{-\alpha_u}}{P_d}\right) + P_{\text{NLoS}} \exp\left(\frac{-\beta d_0^{\alpha_d} \eta P_u |X_u|^{-\alpha_u}}{P_d}\right)\right], \tag{5.10}$$

where $|X_u| = \sqrt{h^2 + r^2}$.

Proof. See Appendix B.1. □

From this theorem, we can make several key observations. First, considering the fact that the UAV creates interference on the D2D users, increasing the UAV altitude to increase its distance from the D2D users does not necessarily reduces the interference on the D2D users. As will be shown later by numerical simulations, by increasing the UAV altitude the D2D coverage probability decreases first, and then increases. This is due to the fact that, considering (5.9) and (5.10), although increasing the UAV altitude increases the path loss term, it also leads to a higher LoS probability. In general, the D2D users prefer to have the NLoS view towards the UAV and have a maximum distance from it, however, these two objectives conflicts with each other. Second, increasing the D2D transmit power (P_d), always enhances the D2D coverage probability, even in an interference limited scenario where noise is ignored. Typically, in the interference limited scenarios, increasing the transmit power of the D2D users does not improve the coverage performance due to the increased interference from other D2D transmitters. According to Theorem 1, although in the interference limited scenario ($N = 0$) the first multiplying term in (5.10) is independent of P_d due to the interference from D2D transmitters, the second term is an increasing function of P_d. Finally, the D2D coverage probability in (5.10) decreases when the UAV transmit power increases. To cope with this situation, the D2D users can increase their transmit power or reduce the fixed distance parameter (d_0). In addition, decreasing the D2D user density improves the coverage probability due to decreasing the interference. Note that the result presented in Theorem 18 corresponds to the coverage probability for a D2D user located at (r, φ). To compute the average coverage probability in the cell, we consider a uniform distribution of users over the area with $f(r,\varphi) = \frac{r}{\pi R_c^2},\ 0 \le r \le R_c\ ,\ 0 \le \varphi \le 2\pi$[1], where R_c is the radius of the desired circular area. Then, we compute the average over the desired area. The average coverage probability for D2D users will be:

[1]Note that the number of users has a Poisson distribution but their location follows the uniform distribution over the area.

$$\begin{aligned}\bar{P}_{\text{cov},d}(\beta) &= \mathbb{E}_{r,\varphi}\left[P_{\text{cov},d}(r,\varphi,\beta)\right] \\ &= \exp\left(\frac{-2\pi^2\lambda_d\beta^{2/\alpha_d}d_0^2}{\alpha_d\sin(2\pi/\alpha_d)} - \frac{\beta d_0^{\alpha_d}N}{P_d}\right) \\ &\times \int_0^{R_c} \mathbb{E}_{I_u}\left[\exp(\frac{-\beta d_0^{\alpha_d}I_u}{P_d})\right] f(r,\varphi)\mathrm{d}r\mathrm{d}\varphi \\ &= \exp\left(\frac{-2\pi^2\lambda_d\beta^{2/\alpha_d}d_0^2}{\alpha_d\sin(2\pi/\alpha_d)} - \frac{\beta d_0^{\alpha_d}N}{P_d}\right) \\ &\times \int_0^{R_c} \mathbb{E}_{I_u}\left[\exp(\frac{-\beta d_0^{\alpha_d}I_u}{P_d})\right] \frac{2r}{R_c^2}\mathrm{d}r. \end{aligned} \tag{5.11}$$

From (5.11), we can see that the average coverage probability for D2D users increases as the size of the area, R_c, increases. In fact, when the UAV serves a larger area, the average distance of D2D users from the UAV increases and on the average they receive lower interference from it. Next, we provide a special case for (5.11) in which the UAV has a very high altitude or very small transmit power.

Remark 1. For $P_u = 0$ or $h \to \infty$, the average coverage probability for the D2D users is simplified to [146]:

$$\bar{P}_{\text{cov,d}}(\beta) = \exp\left(\frac{-2\pi^2\lambda_d\beta^{2/\alpha_d}d_0^2}{\alpha_d\sin(2\pi/\alpha_d)} - \frac{\beta d_0^{\alpha_d}N}{P_d}\right), \tag{5.12}$$

Note that, (5.12) corresponds to the coverage probability in overlay D2D communication in which there is no interference between the UAV and the D2D transmitters. It should be noted that, this result is also related to the success probability in a bipolar ad hoc network [146].

5.3.2 Coverage Probability for Downlink Users

Here, we first derive the bound and lower bound for the downlink users' coverage probability.

Theorem 8. The lower bound and upper bound of the average coverage probability for DUs in the area of interest is given by:

$$\begin{aligned}\bar{P}^L_{\text{cov,du}}(\beta,h) &= \int_0^{R_c} P_{\text{LoS}}(r,h)L_I\left(\frac{P_u|X_u|^{-\alpha_u}}{\beta} - N\right)\frac{2r}{R_c^2}\mathrm{d}r \\ &+ \int_0^{R_c} P_{\text{NLoS}}(r,h)L_I\left(\frac{\eta P_u|X_u|^{-\alpha_u}}{\beta} - N\right)\frac{2r}{R_c^2}\mathrm{d}r, \end{aligned} \tag{5.13}$$

$$\bar{P}^{U}_{\text{cov,du}}(\beta,h)=\int_{0}^{R_c}P_{\text{LoS}}(r,h)U_I\left(\frac{P_u|X_u|^{-\alpha_u}}{\beta}-N\right)\frac{2r}{R_c^2}\mathrm{d}r$$
$$+\int_{0}^{R_c}P_{\text{NLoS}}(r,h)U_I\left(\frac{\eta P_u|X_u|^{-\alpha_u}}{\beta}-N\right)\frac{2r}{R_c^2}\mathrm{d}r, \tag{5.14}$$

where $\beta N < P_u||X_u||^{-\alpha_u}$, and for any $T>0$,

$$L_I(T)=\left[1-\frac{2\pi\lambda_d\Gamma(1+2/\alpha_d)}{\alpha_d-2}\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\right]$$
$$\times\exp\left(-\pi\lambda_d\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\Gamma(1+2/\alpha_d)\right), \tag{5.15}$$

$$U_I(T)=\exp\left(-\pi\lambda_d\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\Gamma(1+2/\alpha_d)\right). \tag{5.16}$$

Also, $\Gamma(t)=\int_{0}^{\infty}x^{t-1}e^{-x}\mathrm{d}x$ is the gamma function [148].

Proof. See Appendix B.2. □

From Theorem 8, we can first see that, for $T >> P_d$, given that $e^{-x}\approx 1-x$ when $x\to 0$, we have $U_I(T)=L_I(T)\approx 1-\pi\lambda_d\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\Gamma(1+2/\alpha_d)$. This means that the lower bound and upper bound become tighter for lower transmit power of D2D users. Moreover, from (15) and (16), when $\lambda_d\to\infty$, the number of D2D users tends to infinity and $U_I=L_I=0$. Consequently, the downlink users experience an infinite interference from the D2D users which results in $\bar{P}_{\text{cov,du}}=0$.

Furthermore, considering (4.9), (4.13), and (4.14), we can see that increasing the UAV altitude (h), can enhance the LoS probability and the coverage probability. On the other hand, due to increasing $|X_u|$, L_I and U_I decrease, and hence the coverage probability for downlink users decreases. Therefore, in order to achieve the maximum coverage, the altitude of the UAV should be carefully adjusted.

As per Theorem 8, increasing R_c decreases the average coverage probability for the downlink users. However, higher R_c results in a higher D2D average coverage probability. Moreover, the average coverage probability for downlink users decreases as the density of the D2D users increases. In this case, to improve the DUs coverage performance, one must increase P_u or reduce R_c. Next, we derive the DU coverage probability in the absence of the D2D users.

Proposition 6. For low density and transmit power of D2D users, the interference from D2D users is negligible compared to the UAV, then, the exact average coverage probability for the downlink users can be expressed as:

$$\bar{P}_{\mathrm{cov,du}}(\beta) = \int_0^{\min\{\left[(\frac{P_u}{\beta N})^{2/\alpha_u} - h^2\right]^{0.5}, R_c\}} P_{\mathrm{LoS}}(r)\frac{2r}{R_c^2}\mathrm{d}r + \int_0^{\min\{\left[(\frac{\eta P_u}{\beta N})^{2/\alpha_u} - h^2\right]^{0.5}, R_c\}} P_{\mathrm{NLoS}}(r)\frac{2r}{R_c^2}\mathrm{d}r. \tag{5.17}$$

Proof. For a DU located at (r, φ), the coverage probability in absence of D2D users becomes

$$\begin{aligned} P_{\mathrm{cov,du}}(r,\varphi,\beta) &= \mathbb{P}\left[\gamma_u \geq \beta\right] = P_{\mathrm{LoS}}(r)\mathbb{P}\left[\gamma_u \geq \beta|\mathrm{LoS}\right] \\ &\quad + P_{\mathrm{NLoS}}(r)\mathbb{P}\left[\gamma_u \geq \beta|\mathrm{NLoS}\right] \\ &= P_{\mathrm{LoS}}(r)\mathbb{1}\left[|X_u| \leq \left(\frac{P_u}{\beta N}\right)^{1/\alpha_u}\right] \\ &\quad + P_{\mathrm{NLoS}}(r)\mathbb{1}\left[|X_u| \leq \left(\frac{\eta P_u}{\beta N}\right)^{1/\alpha_u}\right]. \end{aligned} \tag{5.18}$$

Now, the average coverage probability is computed by taking the average of $P_{\mathrm{cov,du}}(r,\varphi,\beta)$ over the cell with the radius R_c.

$$\begin{aligned} P_{\mathrm{cov},du}(r,\varphi,\beta) &= \mathbb{E}_{r,\varphi}\left[P_{\mathrm{cov},du}(r,\varphi,\beta)\right] \\ &= \int_0^{\min\{\left[(\frac{P_u}{\beta N})^{2/\alpha_u} - h^2\right]^{0.5}, R_c\}} P_{\mathrm{LoS}}(r)\frac{2r}{R_c^2}\mathrm{d}r \\ &\quad + \int_0^{\min\{\left[(\frac{\eta P_u}{\beta N})^{2/\alpha_u} - h^2\right]^{0.5}, R_c\}} P_{\mathrm{NLoS}}(r)\frac{2r}{R_c^2}\mathrm{d}r. \end{aligned} \tag{5.19}$$

□

Proposition 6 gives the exact expression for the downlink users' coverage probability when the interference from D2D users, due to their low density and low transmit power, is negligible compared to the UAV. Therefore, the result in Proposition 6 shows the maximum achievable coverage performance for downlink users when the received signal from the UAV is dominant compared to the interference from the D2D transmitters.

5.3.3 System sum-rate

Now, we investigate the average achievable rates for the DUs and D2D users which can be expressed as in [142]:

$$\bar{C}_{du} = W\log_2(1+\beta)\bar{P}_{\mathrm{cov},du}(\beta), \tag{5.20}$$

$$\bar{C}_{d} = W\log_2(1+\beta)\bar{P}_{\mathrm{cov},d}(\beta), \tag{5.21}$$

where W is the transmission bandwidth. Note that, (5.20) and (5.21) represent lower bounds on the maximum achievable rate due to $\beta \leq \gamma_d$ and $\beta \leq \gamma_u$. However, for tractability, we use (5.20) and (5.21) to compute the average rates as done in [142]. Considering the whole DUs and D2D users in the cell, the system sum-rate, $\bar{C}_{\mathrm{sum}}$, can be derived as a function of the coverage probabilities and the number of users as follows:

$$\bar{C}_{\mathrm{sum}} = {R_c}^2\pi\lambda_{du}\bar{C}_{du} + {R_c}^2\pi\lambda_{d}\bar{C}_{d}. \tag{5.22}$$

Assuming $\mu = \frac{\lambda_{du}}{\lambda_d}$, we have

$$\bar{C}_{\mathrm{sum}} = \lambda_d {R_c}^2\pi \left[\mu\bar{P}_{\mathrm{cov},du}(\beta) + \bar{P}_{\mathrm{cov},d}(\beta)\right] W\log_2(1+\beta), \tag{5.23}$$

where ${R_c}^2\pi\lambda_d$ and ${R_c}^2\pi\lambda_{du}$ are the number of DUs and D2D users in the target area respectively.

From (23), observe that, on the one hand, $\bar{C}_{\mathrm{sum}}$ is directly proportional to λ_d, but on the other hand, it depends on the coverage probabilities of DUs and D2D users which both are decreasing functions of D2D user density. Therefore, in general, increasing λ_d does not necessarily enhance the rate. Note that, considering (4.11), (4.13), (4.14), and (4.23), for both $\lambda_d \to 0$ and $\lambda_d \to \infty$ cases the system sum-rate tend to zero. Hence, there is an optimum value for λ_d that maximizes $\bar{C}_{\mathrm{sum}}$.

According to (4.23), $\bar{C}_{\mathrm{sum}}$ is a function of the coverage probability and a logarithmic function of the threshold (β). The former is a decreasing function of β whereas the latter is an increasing function of β. In other words, although increasing the threshold is desirable for the rate due to increasing the logarithmic function, it also reduces the coverage probability. Therefore, in order to achieve a maximum rate, a proper value for the threshold can be adopted. It should be noted that, the SINR threshold, β, is typically fixed and cannot be set lower than the receiver sensitivity. However, the analysis of different values of β brings value in order to understand how one could change the SINR threshold value (in the future) through proper resource allocation or just system design (change in the number of users, etc).

5.4 Network with a Mobile UAV

Now, we assume that the UAV can move around the area of radius R_c in order to provide coverage for all the downlink users in the target area. In particular, we consider a UAV that moves over the

target area and only transmits at a given geographical location (area) which we hereinafter refer to as *stop points*. Each stop point represents a location over which the UAV stops and serves the present downlink users. Here, our first goal is to minimize the number of stop points (denoted by M) and determine their optimal location. The objective of the UAV is to cover the entire area and ensure that the coverage requirements for all DUs are satisfied with a minimum UAV transmit power and minimum number of stop points. In other words, we find the minimum number and location of stop points for the UAV to completely cover the area. We model this problem by exploiting the so-called *disk covering problem* [149]. In the disk covering problem, given a unit disk, the objective is to find the smallest radius required for M equal smaller disks to completely cover the unit disk. In the dual form of the problem, for a given radius of small disks, the minimum number of disks required to cover a bigger disk is found.

In Figure 5.2, we provide an illustrative example to show the mapping between the mobile UAV communication problem and the disk covering problem. In this figure, the center of small disks can be considered as the location of stop points and the radius of the disk is the coverage radius of the UAV. Using the disk covering problem analysis, in Table 5.1, we present, for different number of stop points, the minimum required coverage radius of a UAV for completely covering the target area [149, 150]. Thereby, using the dual disk covering problem, for a given maximum coverage radius of a UAV, we can find the minimum number of stop points for covering the entire area. The detailed steps for finding the minimum number of stop points are provided next.

First, we compute the coverage radius of the UAV based on the minimum requirement for the DU coverage probability. The coverage radius is defined as the maximum radius within which the coverage probability for all DUs (located inside the coverage range) is greater than a specified threshold, ϵ. In this case, the UAV satisfies the coverage requirement of each DU which is inside its coverage range. The maximum coverage radius for the UAV at an altitude h transmitting with a power P_u will be given by:

$$R_m = \max\{R | P_{\text{cov},du}(\beta, R) \geq \varepsilon, P_u, h\} = P^{-1}_{\text{cov},du}(\beta, \varepsilon), \tag{5.24}$$

where ε is the threshold for the average coverage probability in the cell (area covered by the UAV). Note that, a user is considered to be in coverage if it is in the coverage range of the UAV. The minimum required number of stop points for the full coverage is:

$$\begin{cases} L = \min\{M\}, \\ P_{\text{cov},du}(r, \varphi, \beta) \geq \varepsilon, \end{cases} \tag{5.25}$$

where M represents the number of stop points, the second condition guarantees that the area is completely covered by the UAV, and L is the minimum value for the number of stop points if the following condition holds:

$$R_{\min,L} \leq R_m \leq R_{\min,L-1} \rightarrow \min\{M\} = L. \tag{5.26}$$

By using Table 5.1, we see that, $R_{\min,L-1}$ and $R_{\min,L}$ are, respectively, the minimum radius required to cover the entire target area with $L-1$ and L disks. After finding the minimum M, we can reduce

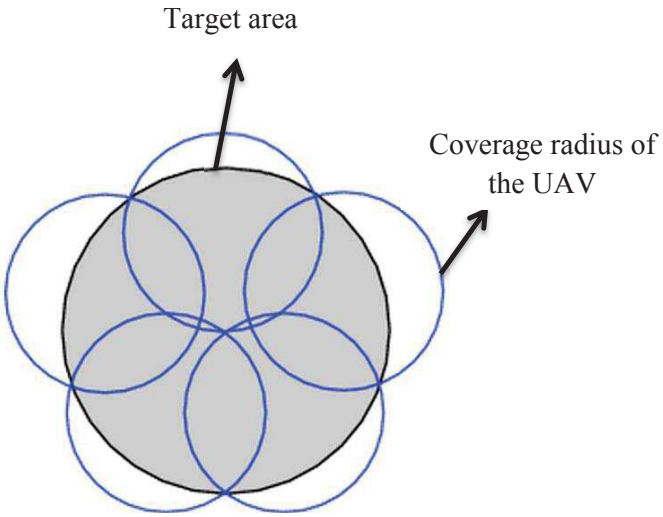


Figure 5.2: Five disks covering problem.

the UAV transmission power such that the coverage radius decreases to the minimum required radius ($R_{\min,L}$). In this way, the UAV transmit power is minimized. Thus, we have:

$$P_{u,\min} = \underset{P_u}{\operatorname{argmin}}\{P^{-1}_{\text{cov,du}}(\beta, \varepsilon) = R_{\min,L}|h\}, \tag{5.27}$$

where $P_{u,\min}$ is the minimum UAV transmit power. Thereby, the minimum number of stop points leads to a full coverage at a minimum time with a minimum required transmit power.

In summary, the proposed UAV deployment method that leads to the complete coverage with a minimum time and transmission power proceeds as follows. First, depending on the parameters of the problem such as density of users and threshold, we compute the maximum coverage radius of a UAV at the optimal altitude that can serve the DUs. Second, considering the size of target area, using the disk covering problem, we find the minimum required number of transmission points along with the coverage radius at each point. Third, we reduce the transmission power of UAV such that its maximum coverage radius becomes equal to the required coverage radius found in the previous step. Using the proposed method, the target area can be completely covered by the UAV with a minimum required transmit power and minimum number of stop points. Next, we investigate the impact of the number of stop points on the full coverage time of the downlink users, and the overall outage probability of the D2D users.

We consider the network during M time instances in which the UAV and D2D users will execute M retransmissions. Note that, our system model considers the downlink, therefore, the retransmissions are essentially from the UAV to the DUs, and from D2D transmitters to corresponding receivers. The moving UAV satisfies the coverage requirements of the downlink users in M retransmissions from different locations. Clearly, as the number of stop points (M) increases, the time required for UAV to completely cover the desired area, increases. Here, the time that the UAV needs to provide the full coverage for the area by visiting all the stop points, is called delay. Hence,

Table 5.1: Number and radii of disks in the covering problem.

Number of stop points	**Minimum required coverage radius ($R_{\min}$)**
$M = 1, 2$	R_c
$M = 3$	$\frac{\sqrt{3}}{2}R_c$
$M = 4$	$\frac{\sqrt{2}}{2}R_c$
$M = 5$	$0.61R_c$
$M = 6$	$0.556R_c$
$M = 7$	$0.5R_c$
$M = 8$	$0.437R_c$
$M = 9$	$0.422R_c$
$M = 10$	$0.398R_c$
$M = 11$	$0.38R_c$
$M = 12$	$0.361R_c$

the delay depends on the travel time of the UAV between the stop points, and the time that UAV spends at each stop point for transmissions. Thus, the delay can be written as:

$$\tau = T_{tr} + MT_s, \tag{5.28}$$

where T_{tr} is the total UAV travel time, M is the number of stop points, and T_s is the time that the UAV stays at each stop point. Clearly, the travel time depends on the travel distance and location of the stop points, and the speed of the UAV. The total travel time will clearly increase as the number of stop points increases. However, in general, the exact relationship between T_{tr} and M strongly depends on the locations of the stop points which do not necessary follow a fixed path/distribution for different values of M. As an example, it can be shown that the exact travel time for $M = 3$ and $M = 4$ is $\frac{\sqrt{3}R_c}{v}$ and $\frac{3R_c}{v}$ respectively, where v is the speed of the UAV, and R_c is the radius of the desired area. The residence time, T_s, depends on the multiple access method. If the UAV adopts a time division multiple access (TDMA) technique, the residence time will be a function of the number of stop points. Note that, a higher number of stop points corresponds to a smaller coverage region of the UAV. Hence, at each stop point, the UAV needs to provide service for a fewer number of users. Therefore, by increasing the number of stop points, the residence time can be decreased in the TDMA case. Considering a uniform distribution of the users, the residence time is approximately computed as:

$$T_s \approx T_{s,1}\frac{R^2_{\min}(M)}{R^2_c}U, \tag{5.29}$$

where $T_{s,1}$ is the service time of UAV for each downlink user, U is the number of downlink users, $R_{\min}$ is the coverage radius of the UAV which depends on M, the number of the stop points, and R_c is the radius of the desired area. However, if the UAV uses a frequency division multiple access (FDMA) technique, the users can be served simultaneously. In other words, the UAV does not

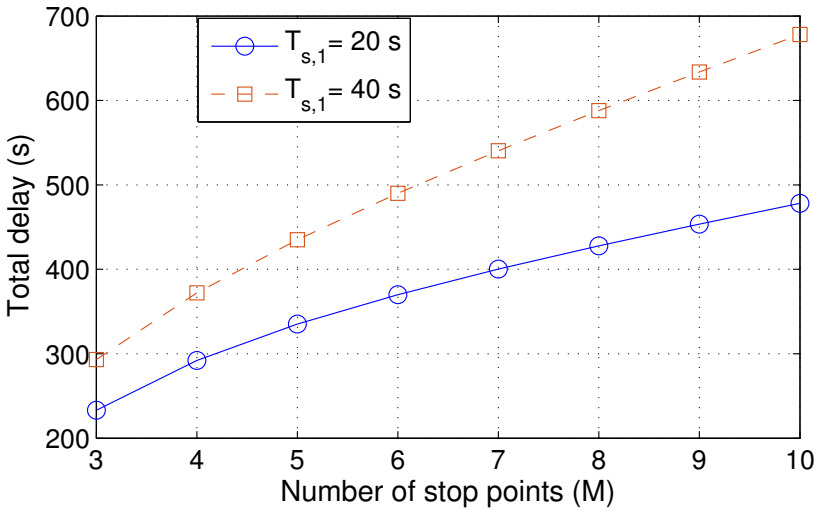


Figure 5.3: Total delay increases as the number of stop points.

need to use different time slots to serve the users. Therefore, if users are of homogeneous traffic type, the residence time of the UAV at each stop point does not depend on the number of the users, and hence it can be fixed. In this case, the residence time at each stop point will be constant and it does not depend on the coverage radius of the UAV and the number of stop points. As a result, $T_s = T_{s,1}$. In our model, we have considered FDMA for multiple access. Hence, the residence time is the same for all values of M. In Figure 5.3, we have shown the total delay versus the number of stop points for two values of residence time, and $v = 10$ m/s. As expected, the total delay increases as the number of stop points increases. Moreover, when the residence time of the UAV at each stop point increases, the additional delay due to a higher number of stop points increases. As we can see from Figure 5.3, for $T_{s,1} = 20$ s, the delay increases from 230 s to 480 s if the number of stop points increases from 3 to 10. However, for $T_{s,1} = 40$ s the delay increases from 295 s to 690 s. Clearly, the delay and the number of stop points are directly related. It should be noted that, for our simulations, we consider the number of stop points as delay.

Next, we derive the overall outage probability for a typical D2D user in M time instances for the mobile UAV case. The outage probability is the probability of having at least one failure during M retransmissions. Assume that the relative location of the i^{th} stop point with respect to the D2D user is (r_i, h_i) where r_i is the distance between the projection of the UAV on the ground and D2D user and h_i is the UAV altitude. Clearly, the distance between the user and UAV is $|X_{u,i}| = \sqrt{h_i^2 + r_i{}^2}$. For different time slots, the Rayleigh fading changes and can be considered independent [145]. However, since the location of the D2D users do not significantly change during the multiple time slots, the interference from the D2D users are correlated. Then, the overall outage probability for D2D users can be found in the next theorem.

Theorem 9. The overall outage probability for D2D users in M retransmissions considering moving UAV is given by:

$$P_{out,d} = 1 - \exp\left(-\lambda_d \int_{R^2} \left[1 - \left(\frac{1}{1 + \frac{\beta|x|^{-\alpha_d}}{d_0^{-\alpha_d}}}\right)^M\right] \mathrm{d}x\right) \times \prod_{i=1}^{M} \mathbb{E}_{I_{u,i}}\left[\exp\left(\frac{-d_0^{\alpha_d}\beta I_{u,i}}{P_d}\right)\right] \exp\left(\frac{-d_0^{\alpha_d}\beta MN}{P_d}\right), \tag{5.30}$$

where M is the number of retransmissions, $I_{u,i}$ is the interference from the UAV at i^{th} retransmission, and $E_{I_{u,i}}(.)$ is:

$$\begin{aligned} &\mathbb{E}_{I_{u,i}}\left[\exp\left(\frac{-d_0^{\alpha_d}\beta I_{u,i}}{P_d}\right)\right] \\ &= P_{\mathrm{LoS},i} \exp\left(\frac{-\beta d_0^{\alpha_d} P_u |X_{u,i}|^{-\alpha_d}}{P_d}\right) \\ &+ P_{\mathrm{NLoS},i} \exp\left(\frac{-\beta d_0^{\alpha_d} \eta P_u |X_{u,i}|^{-\alpha_d}}{P_d}\right). \end{aligned} \tag{5.31}$$

Proof. See Appendix B.3.

□

From Theorem 9, we can observe that, increasing M leads to a higher outage probability. In fact, as the number of stop points increases, the UAV creates a stronger interference on the D2D users. Consequently, $P_{out,d}$ tends to 1 for $M \to \infty$. However, the higher number of stop points for UAV enhances the coverage performance of the downlink users. Hence, a tradeoff between coverage performance of downlink users and the outage of D2D communications should be taking into account. Moreover, Theorem 3 shows that, in order to guarantee that the outage probability does not exceed a specified threshold for different values of M, we should adaptively reduce the distance between the D2D transmitter and receiver (d_0), or have orthogonal spectrum.

5.5 Simulation Results and Analysis

5.5.1 The static UAV scenario

First, we compare our analytical results of the coverage probabilities with the simulation results. Table 5.2 lists parameters used in the simulation and statistical analysis. These parameters are set based on typical values such as in [8] and [142]. Here, we will analyze the impact of the various parameters such as the UAV altitude, D2D density, and SINR threshold on the performance evaluation metrics.

Table 5.2: Simulation parameters in the UAV-D2D network.

Description	**Parameter**	**Value**
UAV transmit power	P_u	5 W
D2D transmit power	P_d	100 mW
Path loss coefficient	K	-30 dB
Path loss exponent for UAV-user link	α_d	2
Path loss exponent for D2D link	α_u	3
Noise power	N	-120 dBm
Bandwidth	W	1 MHz
D2D pair fixed distance	d_0	20 m
Excessive attenuation factor for NLoS	η	20 dB
Parameters for dense urban environment	B, C	0.136, 11.95

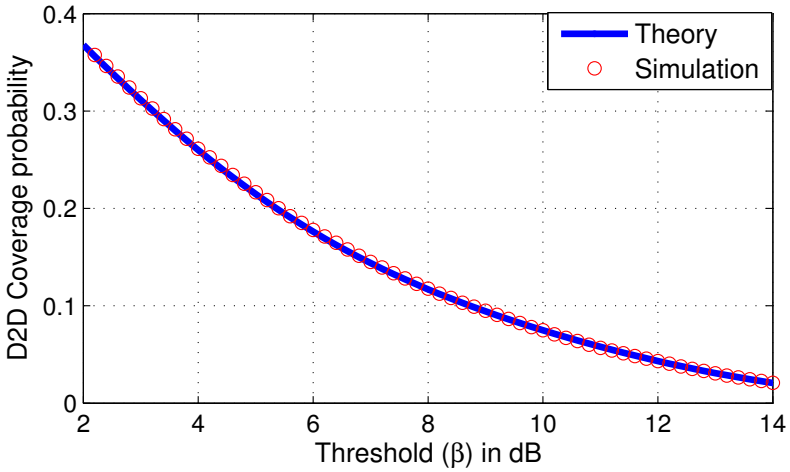


Figure 5.4: D2D coverage probability vs. SINR threshold

In Figures 5.4 and 5.5, we show, respectively, the D2D coverage probability and the lower and upper bounds for the DU coverage probability for different SINR detection threshold values. From these figures, we can clearly see that, the analytical and simulation results for D2D match perfectly and the analytical bounds for DU coverage probability and the exact simulation results are close. Figures 5.4 and 5.5 show that, by increasing the threshold, the coverage probability for D2D users and DUs will decrease.

Figure 5.6 illustrates the system sum-rate (Gbps) versus the threshold for 1 MHz transmission bandwidth, $\lambda_{du} = 10^{-4}$, $h = 500$ m, and two different values of λ_d. By inspecting (23) in Section III, we can see that the rate depends on the coverage probability, which is a decreasing function of the threshold, β, and an increasing logarithmic function of it. Clearly, for high values of β, the received SINR cannot exceed the threshold and, thus, the coverage probabilities tend to zero. On the other hand, according to (20) and (21), as β increases, $\log_2(1+\beta)$ increases

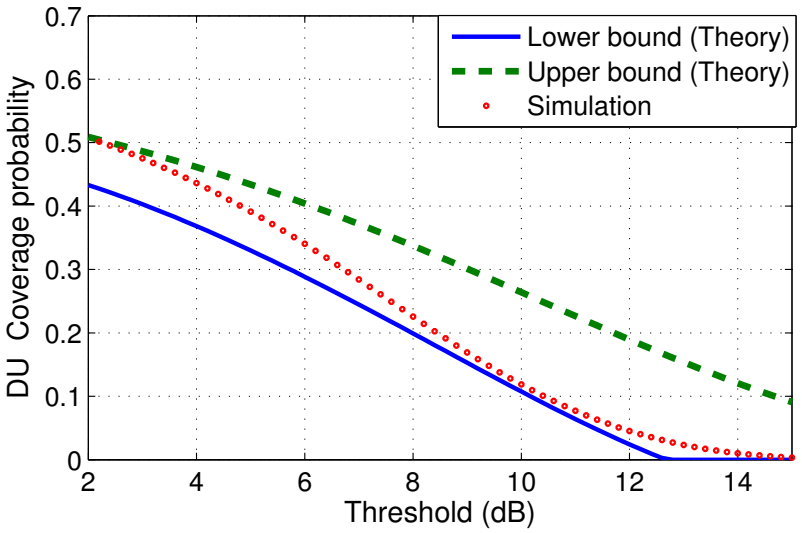


Figure 5.5: DU coverage probability vs. SINR threshold.

accordingly. However, since the coverage probability exponentially decreases but $\log_2(1+\beta)$ increases logarithmically, the average rate tends to zero for the high values of β. Furthermore, for $\beta \to 0$, since $\log_2(1+\beta)$ tends to zero and the coverage probabilities approach one, the rate becomes zero.

Figure 5.7 shows the impact of D2D density on the sum-rate. In this figure, we can see that a low D2D density yields low interference. However, naturally, decreasing the number of D2D users in an area will also decrease the sum-rate. For high D2D density, high interference reduces the coverage probability and consequently the data rate for each user. However, since the sum-rate is directly proportional to the number of D2D users, increasing the D2D density can also improve the sum-rate. According to the Figure 6, as the density of downlink users increases, the optimal λ_d that maximizes the sum-rate decreases. This is due to the fact that, as λ_{du} increases, the contribution of DUs in the system sum-rate increases and hence increasing the rate of each DU enhances the system sum-rate. To increase the rate of a DU, the number of D2D users as the interference source for DUs should be reduced. As a result, the optimal λ_d decreases as as λ_{du} increases. For instance as shown in the figure, by increasing λ_{du} from 10^{-4} to 4×10^{-4}, the optimal λ_d decreases from 0.9×10^{-4} to 0.3×10^{-4}.

It is important to note that the value of the fixed distance, d_0, between the D2D pair significantly impacts the rate performance. Figure 5.8 shows the $\bar{C}_{\text{sum}}$ as a function of the density of D2D users and d_0. From this figure, we can see that, the rate increases as the fixed distance between a D2D receiver and its corresponding transmitter decreases. Moreover, the optimal D2D density which leads to a maximum $\bar{C}_{\text{sum}}$, increases by decreasing d_0. In fact, for lower values of d_0 we can have more D2D users in the network. For instance, by reducing d_0 from 8 m to 5 m, the optimum average number of D2D users increases by a factor of 3.

Figure 5.9 shows the coverage probability for DUs and D2D users as a function of the UAV altitude.

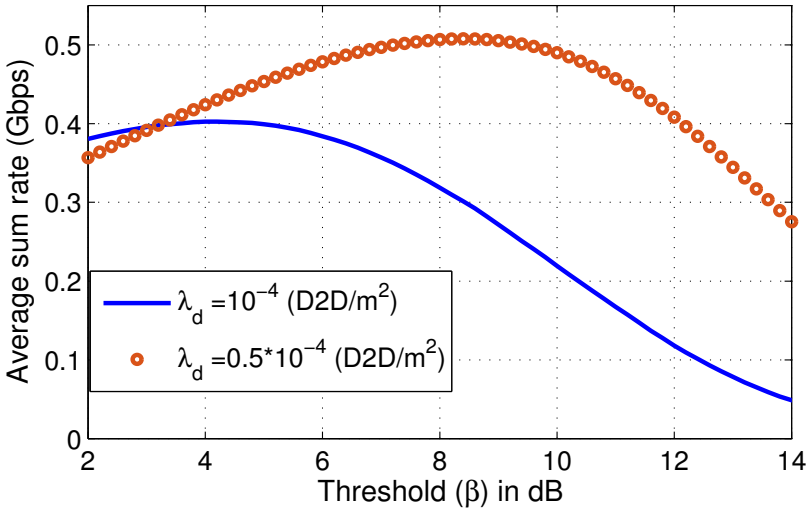


Figure 5.6: System sum-rate vs. SINR threshold.

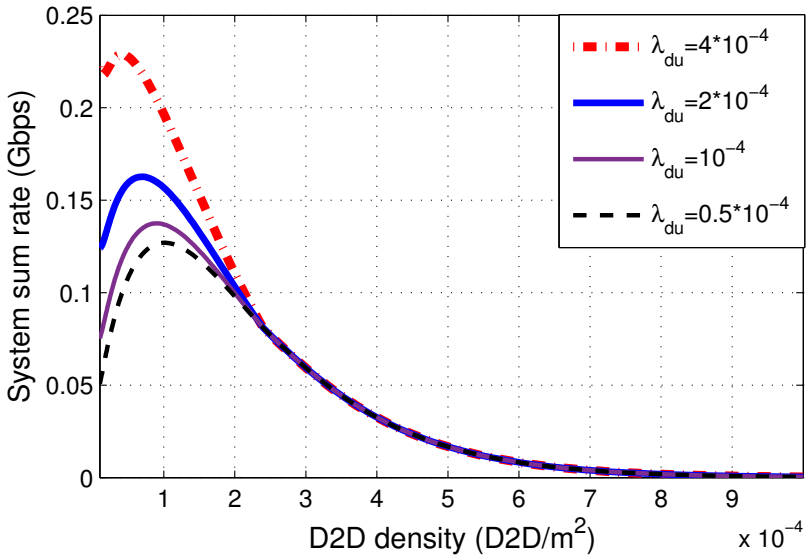


Figure 5.7: System sum-rate vs. D2D density (number of D2D pairs per m^2).

From the DUs' perspective, the UAV should be at an optimal altitude such that it can provide a maximum coverage. In fact, the UAV should not position itself at very low altitudes, due to high shadowing and a low probability of LoS connections towards the DUs. On the other hand, at very high altitudes, LoS links exist with a high probability but the large distance between UAV and DUs results in a high the path loss. As shown in Figure 5.9, for $h = 500\,\mathrm{m}$ the DU coverage probability is maximized. Note that from a D2D user perspective, the UAV creates interference on the D2D receiver. Therefore, D2D users prefer the UAV to be at an altitude for which it provides a minimum coverage radius. As seen in Figure 8, for $h \to \infty$, the D2D users achieve the maximum

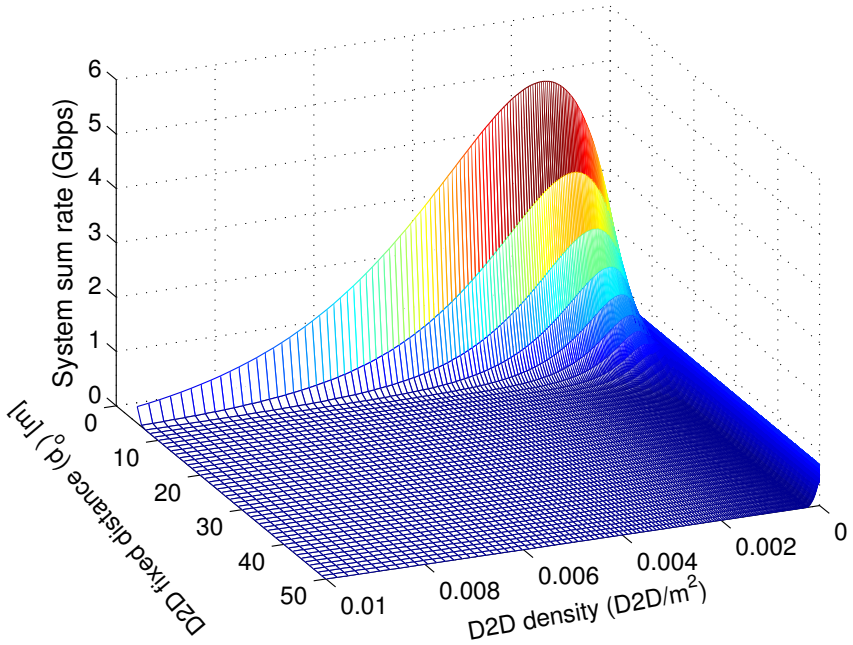


Figure 5.8: System sum-rate vs. D2D density and d_0.

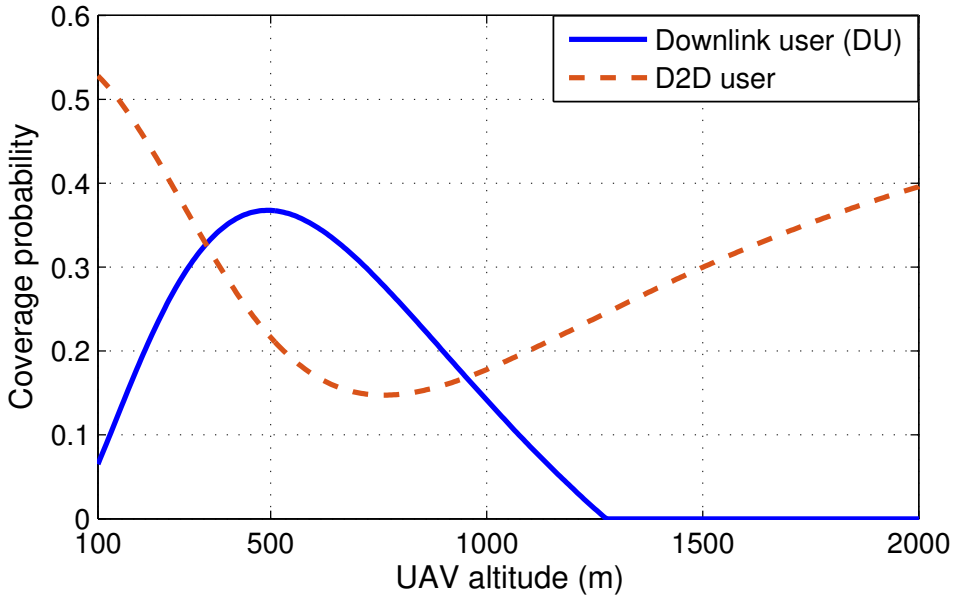


Figure 5.9: Coverage probability vs. UAV altitude.

performance. However, $h = 800\,\text{m}$ results in a minimum D2D coverage probability due the high interference from the UAV.

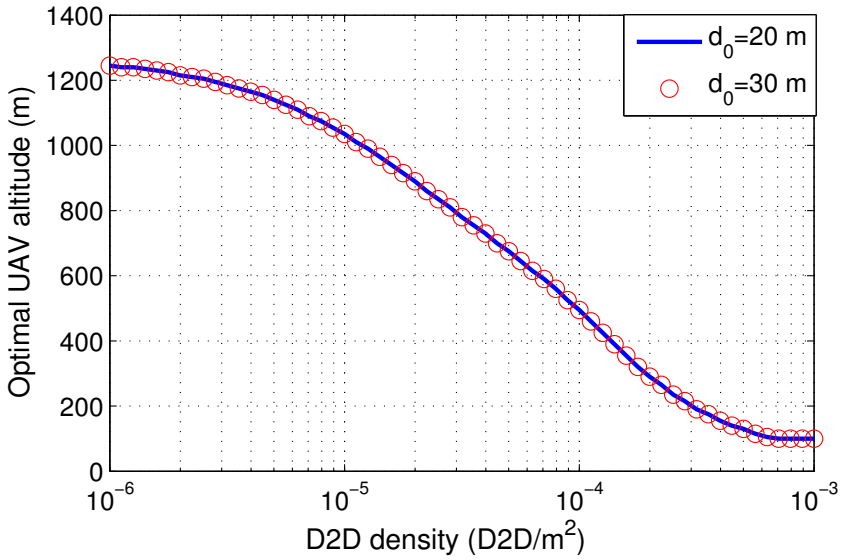


Figure 5.10: Optimal UAV altitude vs. D2D density.

Figure 5.10 shows the optimal UAV altitude that maximizes DU coverage probability versus the D2D users' density. As we can see from Figure 5.10, the optimal UAV altitude for downlink users decreases as the number of D2D users increases. This is due to the fact that a higher density of D2D users creates higher interference on the downlink users, and consequently the UAV reduces its altitude to improve SINR value for the downlink users. In other words, the UAV positions itself closer to the downlink users to cope with the high interference caused by the increased number of D2D users. From Figure 5.10, we can see that, the optimal UAV altitude is independent of the fixed distance, d_0, between the D2D transmitter and receiver pair. In fact, the distance between D2D users does not affect the amount interference generated on the downlink users. Therefore, the optimal altitude of the UAV does not change if d_0 changes.

Figure 5.11 shows $\bar{C}_{\text{sum}}$ versus the UAV altitude for different values of the fixed distance, d_0, the fixed distance between a D2D transmitter/receiver pair. The optimum values for the height which lead to a maximum $\bar{C}_{\text{sum}}$ are around $300\,\text{m}$, $350\,\text{m}$, and $400\,\text{m}$ for $d_0 = 20\,\text{m}, 25\,\text{m}$ and $30\,\text{m}$. Note that the optimal h that maximizes the sum-rate depends on the density of DU and D2D users. From Figure 5.11, considering $d_0 = 20\,\text{m}$ as an example, we can see that for $h > 1300\,\text{m}$, the system sum-rate starts increasing. This stems from the fact that the DU coverage probability tends to zero and, thus, only D2D users impact $\bar{C}_{\text{sum}}$. Hence, as the UAV moves up in altitude, the interference on D2D users decreases and $\bar{C}_{\text{d}}$ increases. Moreover, for $300\,\text{m} < h < 1300\,\text{m}$, Figure 5.11 shows that the coverage probability and, consequently, the average rate for the downlink users decrease as the altitude increases. However, increasing the UAV altitude reduces the interference on the D2D users and improves the average rate for D2D users. In addition, in this range of h, since DUs have more contributions on $\bar{C}_{\text{sum}}$ than the D2D users, $\bar{C}_{\text{sum}}$ is a decreasing function of altitude.

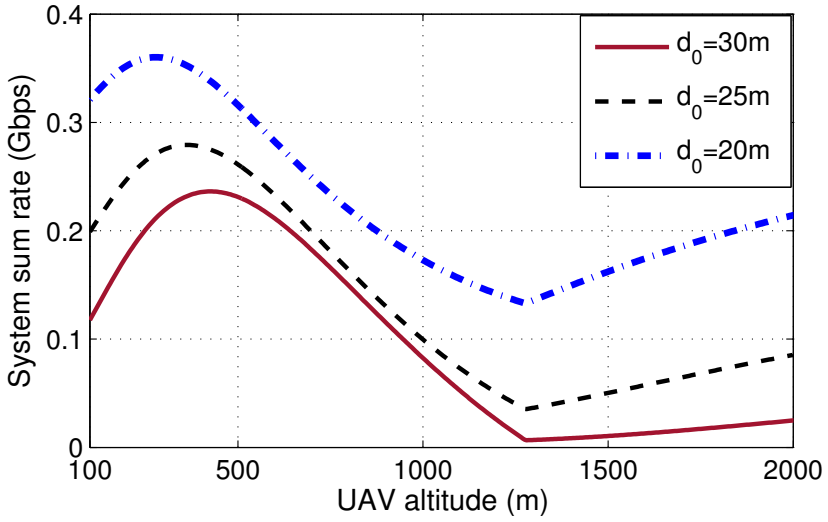


Figure 5.11: System sum-rate vs. UAV altitude.

5.5.2 The mobile UAV scenario

Next, we study the mobile UAV scenario. In this case, the UAV can satisfy the coverage requirement for all the DUs. In fact, the UAV moves over the target area and attempts to serve the DUs at the stop points to guarantee that all the DUs will be in its coverage radius.

Figure 5.12 shows the coverage radius of the mobile UAV when it is located at the optimal altitude as the D2D density varies. As expected, the coverage radius decreases as the D2D density increases. For instance, for $\varepsilon = 0.6$, when λ_d increases from 10^{-5} to 10^{-4}, the coverage radius decreases from 1600 m to 300 m. Moreover, by reducing the minimum coverage requirement of DUs, the UAV can cover a larger area. For instance, reducing ϵ from 0.6 to 0.4 increases the UAV coverage radius from 290 m to 380 m for $\lambda_d = 10^{-4}$. Note that, since the main goal of the UAV is to provide coverage for the entire target area, to compensate for the low coverage radius, we should increase the number of stop points for serving the DUs and consequently the full coverage time increases.

In Figure 5.13, we show the minimum number of stop points as a function of the D2D user density. In this figure, we can see that, as expected, the number of stop points must increase when the density of D2D users increases. In fact, to overcome the higher interference caused by increasing the number of D2D users, the UAV will need more stop points to satisfy the DUs' coverage constraints. For instance, when λ_d increases from 0.2×10^{-4} to 0.8×10^{-4}, the number of stop points must be increased from 3 to 8. Note that, when computing the minimum number of stop points for each λ_d, we considered optimal values for the UAV altitude such that it can provide a maximum coverage for the DUs. Therefore, the UAV's altitude changes according to the D2D density. Moreover, as seen from Figure 5.13, the minimum number of stop points remains constant for a range of λ_d. This is due to the fact that the number of stop points is an integer and hence, for different values of λ_d, the integer value will be the same. However, although the minimum number of stop points

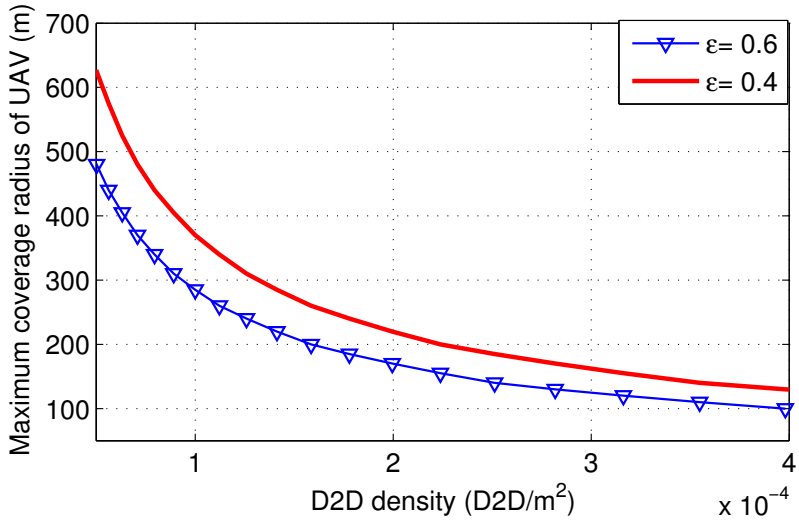


Figure 5.12: Maximum UAV coverage radius vs. D2D density (number of D2D pairs per m^2).

for two different D2D densities are the same, the UAV can transmit with lower power in the case of lower D2D density.

In Figure 5.14, we show the minimum number of stop points as a function of the UAV altitude for $\lambda_d = 10^{-4}$. Figure 5.14 shows that, for some values of h which correspond to the optimal UAV altitude, the minimum number of stop points is minimized. For example, the range of optimal h for $\epsilon = 0.4$ and $\epsilon = 0.6$ is, respectively, $400\,\mathrm{m} < h < 500\,\mathrm{m}$ and $300\,\mathrm{m} < h < 350\,\mathrm{m}$. As expected, the minimum number of stop points is lower for the lower value of ϵ.

Figure 5.15 shows the tradeoff between the downlink coverage probability and the delay which is considered to be proportional to the number of stop points. In Figure 5.15, we can see that, in order to guarantee a higher coverage probability for DUs, the UAV should stop at more locations. As observed in this Figure, for $\lambda_d = 10^{-4}$, to increase the DU coverage probability from 0.4 to 0.7, the number of stop points should increase from 5 to 23. For a higher number of stop points, the UAV is closer to the DUs and, thus, it has a higher chance of LoS. However, on the average, a DU should wait for a longer time to be covered by the UAV that reaches its vicinity. In addition, as the density of D2D users increases, the number of stop points (delay) increases especially when a high coverage probability for DUs must be satisfied. For instance, if λ_d increases from 0.5×10^{-4} to 10^{-4}, or equivalently from 50 to 100 for the given area, the number of stop points should increase from 4 to 9 to satisfy a 0.5 DU coverage probability, and from 20 to 55 for a 0.8 coverage requirement.

Figure 5.16 shows the overall outage probability for D2D users versus the number of retransmissions. As the number of retransmissions (time slots) increases, the overall outage probability also increases. In other words, for higher number of time slots, the possibility that a failure happens during retransmissions, increases. Furthermore, since the UAV is an interference source for the D2D users, the higher number of stop points leads to a higher outage probability. From Figure

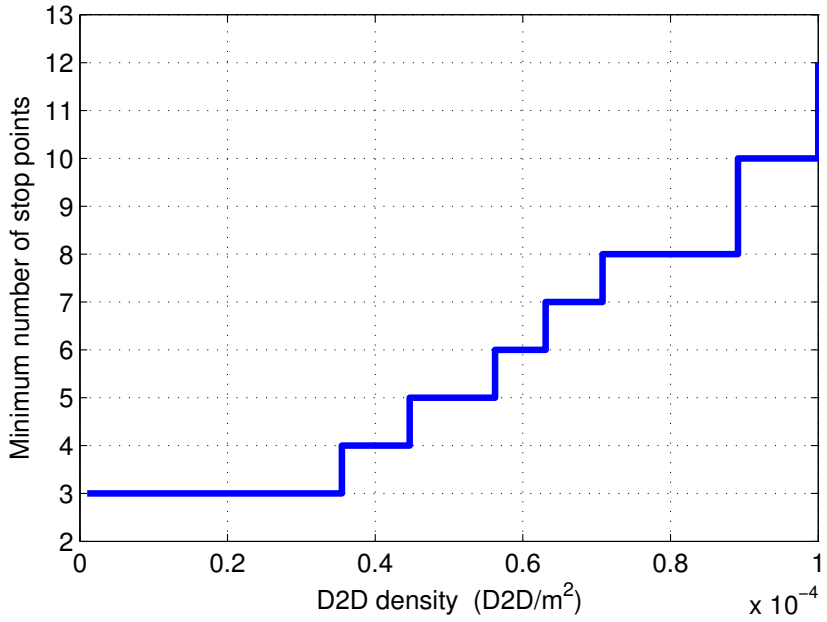


Figure 5.13: Number of stop points vs. D2D density.

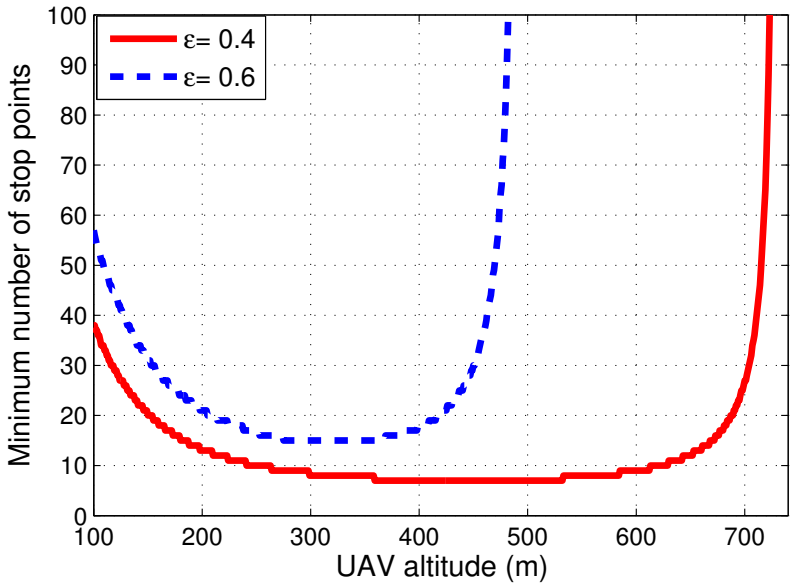


Figure 5.14: Minimum number of stop points vs. UAV altitude.

5.16, we can see that, the increase in the outage probability of D2D users due to the UAV is 0.20 for $M = 3$, and is 0.38 for $M = 7$. Therefore, when the number of stop points increases due to the higher density of D2D users or a higher coverage requirement of the downlink users, the D2D communications are more prone to a failure.

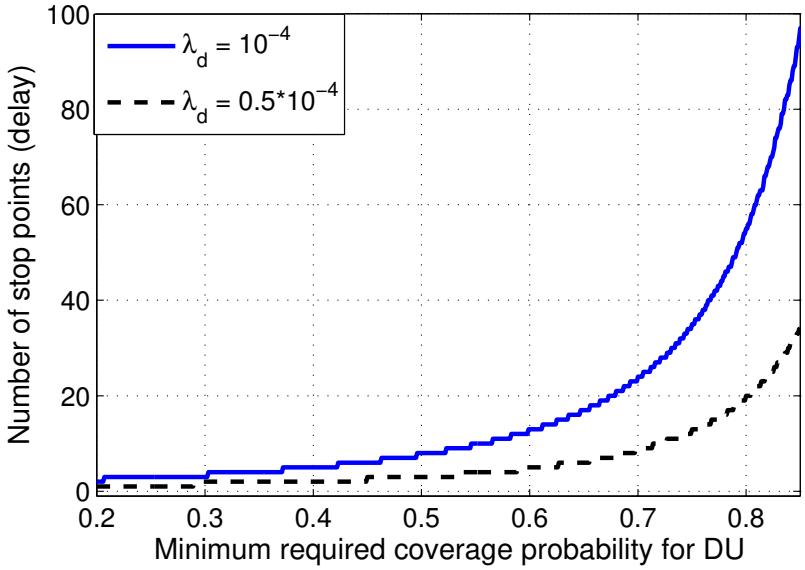


Figure 5.15: Minimum number of stop points vs. coverage probability (coverage-delay tradeoff)

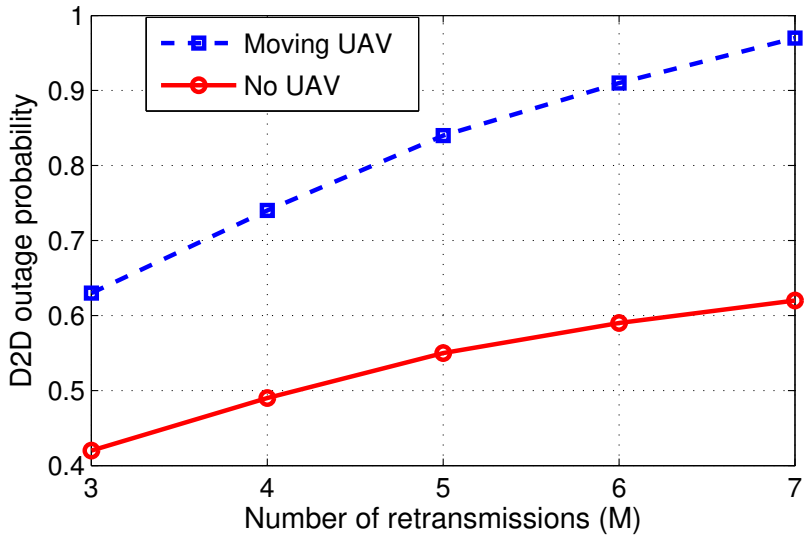


Figure 5.16: Overall D2D outage probability vs. number of retransmissions.

5.6 Summary

In this chapter, we have studied the performance of a UAV that acts as a flying base station in an area in which users are engaged in the D2D communication. We have considered two types of users: in the network: the downlink users served by the UAV and D2D users that communicate directly with one another. For both types, we have derived tractable expressions for the coverage

probabilities as the main performance evaluation metrics. The results have shown that a maximum system sum-rate can be achieved if the UAV altitude is appropriately adjusted based on the D2D users density. In the mobile UAV scenario, using the disk covering problem, the entire target area (cell) can be completely covered by the UAV in a shortest time with a minimum required transmit power. Moreover, in this case, we have derived the overall outage probability for D2D users, and have shown that the outage probability increases as the number of stop point increases. Finally, we have analyzed the tradeoff between the coverage and the time required for covering the entire target area (delay) by the mobile UAV. The results have shown that, the number of stop points must be significantly increased as the minimum coverage requirement for DUs increases.

5.7 Appendix B

B.1 Proof of Theorem 7

$$\begin{aligned} P_{\mathrm{cov},d}(r,\varphi,\beta) &= \mathbb{P}\left[\gamma_d \geq \beta\right] = \mathbb{P}\left[\frac{P_d d_0^{-\alpha_d} g}{I_d^c + I_u + N} \geq \beta\right] \\ &= \mathbb{P}\left[g \geq \frac{\beta d_0^{\alpha_d}(I_d^c + I_u + N)}{P_d}\right] \\ &\overset{(a)}{=} \mathbb{E}_{I_u, I_d^c}\left[\exp(\frac{-\beta d_0^{\alpha_d}(I_d^c + I_u + N)}{P_d})\right] \\ &\overset{(b)}{=} \mathbb{E}_{I_u}\left[\exp(\frac{-\beta d_0^{\alpha_d} I_u}{P_d})\right] \mathbb{E}_{I_d^c}\left[\exp(\frac{-\beta d_0^{\alpha_d} I_d^c}{P_d})\right] \\ &\times \exp\left(\frac{-\beta d_0^{\alpha_d} N}{P_d}\right), \end{aligned} \tag{5.32}$$

where g is an exponential random variable with a mean value of one (i.e. $g \sim \exp(1)$), (a) follows from the exponential distribution of g based on the Rayleigh fading assumption, and taking the expectation over I_u and I_d^c (as random variables). Step (b) comes from the fact that I_u and I_d^c are independent because the interference stems from different sources which are spatially uncorrelated.

Here, $\mathbb{E}_{I_u}$ and $\mathbb{E}_{I_d^c}$ are given by:

$$\begin{aligned} \mathbb{E}_{I_u}\left[\exp(\frac{-\beta d_0^{\alpha_d} I_u}{P_d})\right] &= P_{\mathrm{LoS}} \exp\left(\frac{-\beta d_0^{\alpha_d} P_u |X_u|^{-\alpha_u}}{P_d}\right) \\ &+ P_{\mathrm{NLoS}} \exp\left(\frac{-\beta d_0^{\alpha_d} \eta P_u |X_u|^{-\alpha_u}}{P_d}\right), \end{aligned} \tag{5.33}$$

$$\mathbb{E}_{I_d^c}\left[\exp(\frac{-\beta d_0^{\alpha_d} I_d^c}{P_d})\right] = \mathbb{E}_{d_i, g_i}\left[\prod_i \exp(\frac{-\beta d_0^{\alpha_d}}{P_d} P_d {d_i}^{-\alpha_d} g_i)\right]$$

$$\stackrel{(a)}{=} \exp\left(\frac{-2\pi^2\lambda_d\beta^{2/\alpha_d}d_0^2}{\alpha_d \sin(2\pi/\alpha_d)}\right), \tag{5.34}$$

where the details of step (a) follow directly from the results in [145].

Finally, using (5.32-5.34) Theorem 18 is proved.

B.2 Proof of Theorem 8

The coverage probability for a cellular user located at (r, φ) is written as:

$$\begin{aligned} P_{\text{cov},du}(r,\varphi,\beta) &= \mathbb{P}\left[\gamma_u \geq \beta\right] = P_{\text{LoS}}(r)\mathbb{P}\left[\frac{P_u r^{-\alpha_u}}{I_d + N} \geq \beta\right] \\ &+ P_{\text{NLoS}}(r)\mathbb{P}\left[\frac{\eta P_u r^{-\alpha_u}}{I_d + N} \geq \beta\right] \\ &= P_{\text{LoS}}(r)\mathbb{P}\left[I_d \leq \frac{P_u r^{-\alpha_u} - \beta N}{\beta}\right] \\ &+ P_{\text{NLoS}}(r)\mathbb{P}\left[I_d \leq \frac{\eta P_u r^{-\alpha_u} - \beta N}{\beta}\right]. \end{aligned} \tag{5.35}$$

Note that, there is no closed-form expression for the cumulative distribution function (CDF) of the interference from D2D users [151] and [152]. Here, we provide lower and upper bounds for the CDF of interference. First, we divide the interfering D2D transmitters into two subsets [145]:

$$\begin{cases} \Phi_1 = \{\Phi_{\text{B}} | P_d {d_i}^{-\alpha_d} g_i \geq T\}, \\ \Phi_2 = \{\Phi_{\text{B}} | P_d {d_i}^{-\alpha_d} g_i \leq T\}, \end{cases} \tag{5.36}$$

where T is a threshold which is used to derive the CDF of the interference from D2D users.
Now, considering the interference power from D2D users located in Φ_1 and Φ_2 as I_{d,Φ_1} and I_{d,Φ_2}, we have:

$$\begin{aligned} \mathbb{P}\left[I_d \leq T\right] &= \mathbb{P}\left[I_{d,\Phi_1} + I_{d,\Phi_2} \leq T\right] \leq \mathbb{P}\left[I_{d,\Phi_1} \leq T\right] \\ &= \mathbb{P}\left[\Phi_1 = 0\right] = \mathbb{E}\left[\prod_{\Phi_B} \mathbb{P}(P_d {d_i}^{-\alpha_d} g_i < T)\right] \\ &= \mathbb{E}\left[\prod_{\Phi_B} \mathbb{P}(g_i < \frac{T {d_i}^{\alpha_d}}{P_d})\right] \stackrel{(a)}{=} \mathbb{P}\left[\prod_{\Phi_B} 1 - \exp(-\frac{T {d_i}^{\alpha_d}}{P_d})\right] \\ &\stackrel{(b)}{=} \exp\left(-\lambda_d \int_0^\infty \exp(-\frac{T r^{\alpha_d}}{P_d}) r \mathrm{d}r\right) \end{aligned}$$

$$= \exp\left(-\pi\lambda_d\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\Gamma(1+2/\alpha_d)\right), \tag{5.37}$$

where (a) and (b) come from the Rayleigh fading assumption and PGFL of the PPP.

The upper bound is derived as follows:

$$\begin{aligned}
\mathbb{P}\left[I_d \le T\right] &= 1 - \mathbb{P}\left[I_d \ge T\right] \\
&= 1 - \Big(\mathbb{P}\left[I_d \ge T | I_{d,\Phi_1} \ge T\right]\mathbb{P}\left[I_{d,\Phi_1} \ge T\right] \\
&\qquad + \mathbb{P}\left[I_d \ge T | I_{d,\Phi_1} \le T\right]\mathbb{P}\left[I_{d,\Phi_1} \le T\right]\Big) \\
&= 1 - \Big(\mathbb{P}\left[I_{d,\Phi_1} \ge T\right] + \mathbb{P}\left[I_d \ge T | I_{d,\Phi_1} \le T\right] \\
&\qquad \times \mathbb{P}\left[I_{d,\Phi_1} \le T\right]\Big) \\
&= 1 - \Big(1 - \mathbb{P}\left[\Phi_1 = 0\right] + \mathbb{P}\left[I_d \ge T | I_{d,\Phi_1} \le T\right] \\
&\qquad \times \mathbb{P}\left[\Phi_1 = 0\right]\Big) \\
&= \mathbb{P}\left[\Phi_1 = 0\right]\Big(1 - \mathbb{P}\left[I_d \ge T | \Phi_1 = 0\right]\Big).
\end{aligned} \tag{5.38}$$

Also,

$$\begin{aligned}
\mathbb{P}\left[I_d \ge T | \Phi_1 = 0\right] &\overset{(a)}{\le} \frac{\mathbb{E}\left[I_d \ge T | \Phi_1 = 0\right]}{T} \\
&= \frac{1}{T}\mathbb{E}\left[\sum_{\Phi} P_d {d_i}^{-\alpha_d} g_i \mathbb{1}(P_d {d_i}^{-\alpha_d} g_i \le T)\right] \\
&= \frac{1}{T}\mathbb{E}_{d_i}\left[\sum_{\Phi} P_d {d_i}^{-\alpha_d}\mathbb{E}_{g_i}\left[g_i \mathbb{1}(g_i \le \frac{T {d_i}^{\alpha_d}}{P_d})\right]\right] \\
&= \frac{1}{T}\mathbb{E}_{d_i}\left[\sum_{\Phi} P_d {d_i}^{-\alpha_d}\left[\int_0^{\frac{T {d_i}^{\alpha_d}}{P_d}} g e^{-g}\mathrm{d}g\right]\right] \\
&= \frac{2\pi P_d \lambda_d}{T}\int_0^{\infty} r^{-\alpha_d}\left(\int_0^{\frac{T r^{\alpha_d}}{P_d}} g e^{-g}\mathrm{d}g\right) r\mathrm{d}r \\
&= \frac{2\pi\lambda_d\Gamma(1+2/\alpha_d)}{\alpha_d - 2}\left(\frac{T}{P_d}\right)^{-2/\alpha_d},
\end{aligned} \tag{5.39}$$

where (a) is based on the Markov's inequality which is stated as follows: for any non-negative integrable random variable X and positive L, $P(X \geq L) \leq \frac{\mathbb{E}[X]}{L}$. Also, $\mathbb{1}(.)$ is the indicator function which can only be equal to 1 or 0. Hence, the lower (L_I) and upper (U_I) bounds for the CDF of interference become:

$$L_I(T) = \left[1 - \frac{2\pi\lambda_d \Gamma(1+2/\alpha_d)}{\alpha_d - 2}\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\right] \times \exp\left(-\pi\lambda_d\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\Gamma(1+2/\alpha_d)\right), \tag{5.40}$$

$$U_I(T) = \exp\left(-\pi\lambda_d\left(\frac{T}{P_d}\right)^{-2/\alpha_d}\Gamma(1+2/\alpha_d)\right). \tag{5.41}$$

Thus, we have $L_I(T) \leq \mathbb{P}\{I_d \leq T\} \leq U_I(T)$.

Finally, considering (34), (39), and (40), the lower bound and upper bound of the average coverage probability for DUs in the cell is expressed as:

$$\bar{P}^L_{\text{cov,du}}(\beta) = \int_0^{R_c} P_{\text{LoS}}(r) L_I\left(\frac{P_u|X_u|^{-\alpha_u}}{\beta} - N\right)\frac{2r}{R_c^2}\mathrm{d}r + \int_0^{R_c} P_{\text{NLoS}}(r) L_I\left(\frac{\eta P_u|X_u|^{-\alpha_u}}{\beta} - N\right)\frac{2r}{R_c^2}\mathrm{d}r, \tag{5.42}$$

$$\bar{P}^U_{\text{cov,du}}(\beta) = \int_0^{R_c} P_{\text{LoS}}(r) U_I\left(\frac{P_u|X_u|^{-\alpha_u}}{\beta} - N\right)\frac{2r}{R_c^2}\mathrm{d}r + \int_0^{R_c} P_{\text{NLoS}}(r) U_I\left(\frac{\eta P_u|X_u|^{-\alpha_u}}{\beta} - N\right)\frac{2r}{R_c^2}\mathrm{d}r, \tag{5.43}$$

and Theorem 8 is proved.

B.3 Proof of Theorem 9

Consider $\gamma_{d,i}$ and g_i, respectively, the SINR and the channel gain (with exponential distribution) at i^{th} retransmission, for $1 \leq i \leq M$. The outage probability is the probability of having at least one failure during M retransmissions. Then, we have:

$$P_{out,d} = 1 - \mathbb{P}\left[\gamma_{d,1} \geq \beta, ..., \gamma_{d,M} \geq \beta\right]$$

$$
\begin{aligned}
&= 1 - \mathbb{P}\left[\frac{P_d d_0^{-\alpha_d} g_1}{I_{d,1}^c + I_{u,1} + N} \geq \beta, ..., \frac{P_d d_0^{-\alpha_d} g_M}{I_{d,M}^c + I_{u,M} + N} \geq \beta\right] \\
&= 1 - \mathbb{P}\Bigg[g_1 \geq \frac{d_0^{\alpha_d}\beta(I_{d,1}^c + I_{u,1} + N)}{P_d}, ... \\
&..., g_M \geq \frac{d_0^{\alpha_d}\beta(I_{d,M}^c + I_{u,M} + N)}{P_d}\Bigg] \\
&\overset{(a)}{=} 1 - \mathbb{E}\left[\prod_{i=1}^{M} \exp\left(\frac{-d_0^{\alpha_d}\beta(I_{d,i}^c + I_{u,i} + N)}{P_d}\right)\right] \\
&\overset{(b)}{=} 1 - \mathbb{E}\left[\prod_{i=1}^{M} \exp\left(\frac{-d_0^{\alpha_d}\beta I_{d,i}^c}{P_d}\right)\right] \mathbb{E}\left[\prod_{i=1}^{M} \exp\left(\frac{-d_0^{\alpha_d}\beta I_{u,i}}{P_d}\right)\right] \\
&\qquad \times \exp\left(\frac{-d_0^{\alpha_d}\beta M N}{P_d}\right),
\end{aligned} \tag{5.44}
$$

where (a) follows the assumption that the fading is independent in different retransmissions, and step (b) comes from the fact that interference due to D2D users, interference from UAV, and noise are all independent. Also,

$$
\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{M} \exp\left(\frac{-d_0^{\alpha_d}\beta I_{d,i}^c}{P_d}\right)\right] &= \mathbb{E}\left[\exp\left(\frac{-d_0^{\alpha_d}\beta \sum_{i=1}^{M} I_{d,i}^c}{P_d}\right)\right] \\
&\overset{(c)}{=} \exp\left(-\lambda_d \int_{R^2} \left[1 - \left(\frac{1}{1 + \frac{\beta|x|^{-\alpha_d}}{d_0^{-\alpha_d}}}\right)^{M}\right] \mathrm{d}x\right),
\end{aligned} \tag{5.45}
$$

where details of (c) can be found in [145] where the correlation between D2D interference in different retransmissions is taken into account. Finally,

$$
\begin{aligned}
&\prod_{i=1}^{M} \mathbb{E}_{I_{u,i}}\left[\exp\left(\frac{-d_0^{\alpha_d}\beta I_{u,i}}{P_d}\right)\right] \\
&\overset{(d)}{=} \prod_{i=1}^{M}\Bigg[P_{\mathrm{LoS},i} \exp\left(\frac{-\beta d_0^{\alpha_d} P_u |X_{u,i}|^{-\alpha_d}}{P_d}\right) \\
&\qquad + P_{\mathrm{NLoS},i} \exp\left(\frac{-\beta d_0^{\alpha_d}\eta P_u |X_{u,i}|^{-\alpha_d}}{P_d}\right)\Bigg],
\end{aligned} \tag{5.46}
$$

where step (d) is based on the fact that the interference from the UAV can be treated as independent in different retransmissions.

Finally, using (5.44-5.46), Theorem 9 is proved.

Chapter 6

Cache-Enabled Unmanned Aerial Vehicles in Wireless Networks

6.1 Background, Related Works, and Contributions

In the previous chapters we analyzed various challenges related to UAV-based wireless communications. In this chapter, we introduce a potential application of UAVs in wireless networks that can yield a significant in their quality of service. In particular, we investigate the use of cache-enabled UAVs that can assist the cellular networks in various scenarios by traffic offloading and capacity enhancement.

The next-generation of cellular systems is expected to be largely user centric and, as such, it must be cognizant of human-related information such as users' behavior, mobility patterns, and quality-of-experience (QoE) expectations [153]. One promising approach to introduce such wireless network designs with human-in-the-loop is through the use of cloud radio access networks (CRANs) [154]. In CRANs, a central cloud processor can parse through the massive users' data to learn the users' information such as content request distribution and mobility patterns, then, determine how to manage resources in the network. However, an effective exploration of human-in-the-loop features in a CRAN faces many challenges that range from effective predictions to user behavior tracking, effective caching, and optimized resource management.

The existing literature has studied a number of problems related to caching in CRANs and heterogeneous networks with human-in-the-loop such as [155–161]. In [155], the authors proposed a content pushing strategy for traffic offloading that accounts for both content preference and sharing willingness of human users. The authors in [156] proposed an echo state network to predict the users' content request distribution and mobility patterns in CRANs. The work in [157] proposed a caching-based milimiter wave (mmWave) framework, in which base stations pre-store video contents and service users with a high data rate. In [158], a caching framework aimed at fully exploiting the potential of such CRAN systems through cooperative hierarchical caching is

proposed. The authors in [159] investigated the fully cooperative caching case during which a centralizer helps all base stations to make caching decisions. The work in [160] studied the use of cache-enabled small base stations to alleviate the load of backhaul. In [161], a content caching strategy is proposed to jointly minimize the cell average outage probability and fronthaul usage in CRANs. However, most of these existing caching works [155–161] were typically restricted to static networks without mobility and ultra dense users. Note that, in these contributions [155–161], the cached content is stored at the terrestrial static base stations. However, in an area with ultra dense users and high rise buildings (i.e. stadium or hotspots), the static ground base stations with caching may not be able to meet high capacity demands of the users. Moreover, caching at ground base stations may not be effective in serving the mobile users once they move outside the coverage range of the ground base stations. For instance, whenever a mobile user moves to a new cell, its requested content may not be available at the new base station and, consequently, the users cannot be serviced properly. In such case, to effectively serve mobile user, the requested content needs to be cached at multiple base stations which may not be efficient. Therefore, there is a need to introduce a more flexible base station that can boost the capacity and track the users' mobility patterns so as to improve the caching efficiency. To this end, unmanned aerial vehicles can be used as flying base stations to dynamically cache the popular contents, track the mobility pattern of the corresponding users then, effectively serve them. In this case, due to the high altitude and flexible deployment of the UAVs, they can establish reliable communication links to the users by mitigating the blockage effect.

The use of UAVs for enhancing wireless communications in cellular and ad hoc networks was studied in [4, 7, 9, 10, 12, 18, 162, 163]. However, this existing literature [4, 7, 9, 10, 12, 18, 162, 163] was focused on performance analysis and did not consider prediction of user-centric behavioral patterns such as mobility nor it study the use of UAVs for caching purposes. The prediction of the users' mobility patterns can enable the UAVs to effectively move and provide service for the ground users. Moreover, when UAVs are considered within a CRAN system, the network must take into account the fact that the fronthaul links that connect the UAVs to the cloud will be capacity-limited. This is due to the fact that the bandwidth of the UAVs fronthaul links is limited. To overcome this limited-fronthaul capacity challenge, one can use content caching techniques to proactively download and cache content at the UAVs during off peak hours or when the UAVs are back at their docking stations. The use of caching enables the UAVs directly transmit the content to its requested user, thus reducing the fronthaul traffic load.

Some recent works such as in [156] and [164–169] have studied a number of ideas related to the predictions of human behavior in wireless networking scenarios. In [164] and [165], the authors proposed a prediction algorithm for the users' mobility patterns based on a deep learning algorithm and a semi-Markov process. In [166], a type of user-initiated network is proposed for cellular users to trade data plans by leveraging personal hotspots with users' smartphones. The work in [167] investigated the predictable degree of users' mobility patterns. Nevertheless, the mobility prediction works in [164–167] focused only on the prediction phase and did not study how the users' mobility patterns can be used to optimize the wireless performance using user-centric caching and resource allocation techniques. The authors in [168] developed a data extraction method using the Hadoop

platform to predict content popularity. The work in [169] proposed a fast threshold spread model to predict the future access patterns of multimedia content based on social information. However, the works in [168] and [169] do not consider the complexity of using the users' traffic data while predicting the users' content popularity. In particular, in a dense network with high traffic demands, the proposed approaches in [168] and [169] are not practical as they require each base station to store all users' traffic data to predict the users' content popularity. In contrast, our proposed conceptor based ESN algorithm can record the users' historical traffic data and, consequently, use them to predict the users' behavior. In this case, the complexity of the learning algorithm will be significantly lower compared to the algorithms without ability of recording historical data. The most related work here is our work [156] in which we exploit the echo state network to predict the content request distribution. However, in [156], the ESN-based algorithm can only predict one non-linear system. Here, we proposed an *conceptor*-based ESN algorithm which can separate the users' behavior into different patterns and learn these patterns independently, thus leading to a significant improvement in the accuracy of predictions. For example, in order to predict the user's mobility patterns, the ESN algorithm in [156] will collect all users' mobility data from Monday to Sunday to train the ESN as a prediction system. However, the conceptor-based ESN algorithm can separate the data from Monday to Sunday into seven patterns and use the data of seven patterns to train the conceptor-based ESN as seven independent prediction systems. In this case, the conceptor-based ESN can use seven independent prediction system to predict the user's mobility during each day. Since in reality the user's mobility pattern in each day will be different, the conceptor-based ESN algorithm can improve the accuracy of the prediction.

In this chapter, we develop a novel framework that leverages user-centric information, such as content request distribution and mobility patterns, to effectively deploy cache-enabled UAVs while maximizing the users' QoE using a minimum total transmit power of the UAVs. The adopted QoE metric captures human-in-the-loop features such as transmission delay and the users' perceptions on the rate requirement, depending on their device type. In the proposed framework, the cloud can accurately predict the content request distribution and mobility patterns of each user. These predictions of user's behavior can then be used to find the optimal locations and content caching strategies for the UAVs. Unlike previous studies such as [164–169] that predict the users' behavior using only one non-linear system, we propose a conceptor-based echo state network (ESN) approach to perform users' behavior prediction. Such an ESN model with conceptors enables the cloud to separate the users' behaviors into different patterns and learn these patterns independently, thus leading to a significant improvement in the accuracy of predictions. Moreover, unlike previous studies such as [4, 10, 12, 18, 162, 163, 170] that consider the deployment of the UAVs assuming static users, we study the deployment of cache-enabled UAVs in CRANs with mobile users. In the proposed CRANs model, we derive the user-UAV association, the optimal locations of the UAVs as well as the content to cache at the UAVs. In this work, *we analyze the use of caching at the level of UAVs, given ESN-based predictions on the users' behavior*. To evaluate the performance of the proposed approach, we use real data from Youku for content requests as well as realistic measured mobility data from the Beijing University of Posts and Telecommunications for mobility simulations. Simulation results show that the proposed algorithm can yield 33.3% gain in terms of the average transmit power of the UAVs compared to a baseline algorithm without cache. More-

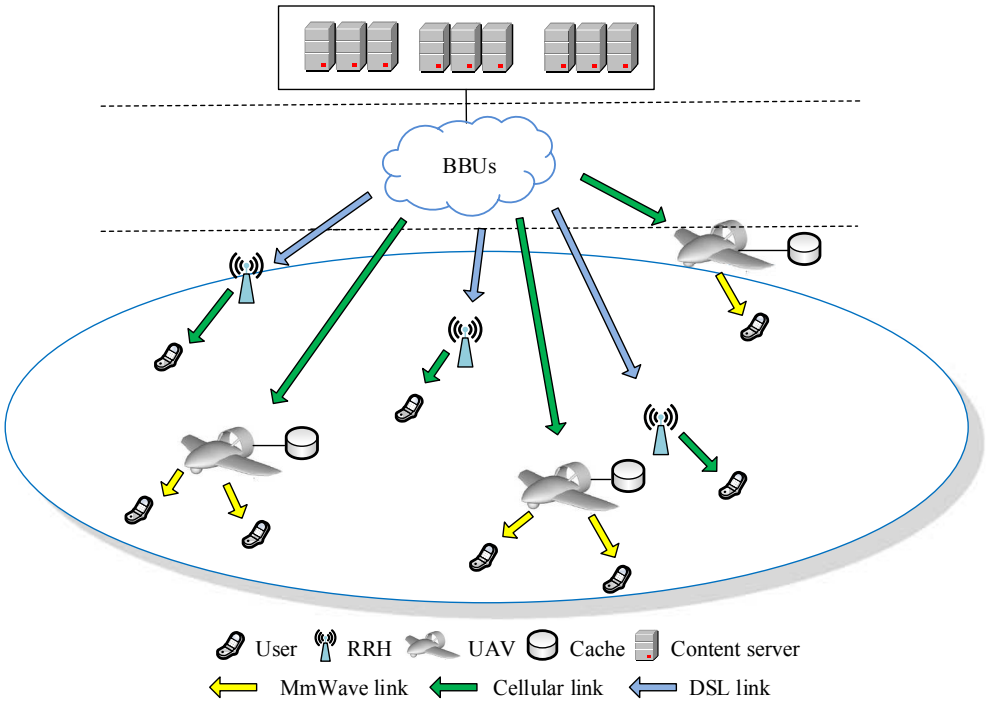


Figure 6.1: A CRAN with cache-enabled UAVs.

over, the proposed algorithm can also yield 59.6% gain in terms of the percentage of the users with satisfied QoE compared to a benchmark scenario without UAVs.

The rest of this chapter is organized as follows. The system model and problem formulation are presented in Section 6.2. The conceptor ESN for content request distribution and mobility patterns predictions is proposed in Section 6.3. The proposed approach for user-UAV association, content caching, and optimal location of each UAV is presented in Section 6.4. In Section 6.5, we provide numerical and simulation results. Finally, conclusions are drawn in Section 6.6.

6.2 System Model and Problem Formulation

Consider the downlink of a CRAN system servicing a set $\mathcal{U}$ of U mobile users via a set $\mathcal{R}$ of R remote radio heads (RRHs) acting as distributed antennas. The RRHs are grouped into E clusters using K-*mean clustering* approach [171] so that zero-forcing beamforming (ZFBF) [172] can be used to service the users. In this system, a set $\mathcal{K}$ of K UAVs equipped with cache storage units can be deployed to act as flying cache-enabled RRHs to serve the ground users along with the terrestrial RRHs. For the UAV-to-users communication links, since the high altitude of the UAVs can significantly reduce the blocking effect due to obstacles, we consider air-to-ground UAV transmissions using the millimeter wave (mmWave) frequency spectrum. Meanwhile, the terrestrial RRHs transmit over the cellular band and are connected to the cloud's pool of the baseband units (BBUs)

Table 6.1: List of notations

Notation	Description	Notation	Description
U	Number of users	C	Number of contents stored at UAV cache
K	Number of UAVs	F	Number of intervals in each time slot
R	Number of RRHs	H	Number of time slots to collect user mobility
P_R	Transmit power of RRHs	$P_{t,ki}$	Transmitted power of UAV or RRH
N	Number of contents	τ, Δ_τ	Time slot index, Time slot duration
$l_{t,ki}$	Path loss of UAVs-users	$d_{t,ki}$	Distance between RRHs or UAVs and users
$x_{\tau,k}, y_{\tau,k}, h_{\tau,k}$	Coordinates of UAVs	$\delta_{S_i,n}$	Rate requirement of device type
L_{FS}	Free space path loss	d_0	Free-space reference distance
f_c	Carrier frequency	$l^F_{t,ki}$	Path loss of fronthaul links
μ_{LoS}, μ_{NLoS}	Path loss exponents	$\chi_{\sigma_{\text{LoS}}}, \chi_{\sigma_{\text{NLoS}}}$	Shadowing random variable
$\gamma^{\text{V}}_{t,ki}, \gamma^{\text{H}}_{t,ki}$	SINR of user i	$L^{\text{LoS}}_{t,k}$, $L^{\text{NLoS}}_{t,k}$	LoS/NLoS path loss from the BBUs to UAV k
t, Δ_t	Small interval, interval duration	$l^{\text{LoS}}_{t,k}$, $l^{\text{NLoS}}_{t,k}$	LoS/NLoS path loss from UAV k to users
c	Speed of light	$h_{t,ki}$	Channel gains between RRHs k and user i
$\bar{D}_{\tau,i,n}$	Delay	$C^F_{\tau,ki}$	Fronthaul rate of UAV or RRH k
$C^{\text{V}}_{\tau,ki}$	Rate of UAV-user link	$C^{\text{H}}_{\tau,qi}$	Rate of RRH-user link
$Q_{\tau,i,n}$	QoE of each user i	T	Number of time slots for caching update
$x_{t,i}, y_{t,i}$	Coordinates of users	P_B	Transmit power of the BBUs

via capacity-constrained, digital subscriber line (DSL) fronthaul links. Further, the cloud connects to the content servers via fiber backhaul links. The transmissions from the cloud to the UAVs occurs over wireless fronthaul links using the licensed cellular band. Consequently, the UAVs' wireless fronthaul links may interfere with the transmission links from the RRHs to the users when the user's requested content needs to be transmitted from the content server.

In our model, the content server stores a set $\mathcal{N}$ of all N contents required by all users. The contents are of equal size L. Caching at the UAVs, referred to as "UAV cache" hereinafter, will be used to store the popular content that the users request. By caching predicted content, the transmission delay from the content server to the UAVs can be significantly reduced as each UAV can directly transmit its stored content to the users. Different from caching at the RRHs or BBUs, caching at UAVs allows servicing mobile users when their QoE requirement cannot be satisfied by the RRHs. We denote the set of C cached contents in the storage units of UAV k by $\mathcal{C}_k$, where $C \leq N$ and $k \in \mathcal{K}$. For simplicity, we assume that each user can request at most one content during each specified time slot τ. We also let Δ_τ be the duration of time slot τ that also represents the maximum transmission duration of each content. The maximum transmission duration Δ_τ is determined by the proposed algorithm in Section 6.4. We assume that the content stored at the UAV cache will be refreshed every period that consists of T time slots and this UAVs caching is performed at off peak hours when the UAVs return to their cloud-based docking stations for purposes such as battery charge. Table I provides a summary of the notations used in this chapter.

6.2.1 Mobility Model

In our system, we assume that the users can move continuously. In this case, we consider a realistic model for periodic, daily, and *pedestrian mobility patterns* according to which each user will regularly visit a certain location of interest. For example, certain users will often go to the same

office for work at the same time during weekdays. The locations of each user are collected by the BBUs once every H time slots. Here, the duration of H time slots is considered a period of one hour for each user. In addition, we assume that each user moves between two collected locations at a constant speed. The mobility pattern for each user will then be used to determine the content that must be cached as well as the optimal location of each UAV which will naturally impact the QoE of each user.

In this model, the associations of the mobile users with the UAVs or the RRHs can change depending on the QoE requirement. Since the users are moving continuously, the locations of the UAVs must change accordingly so as to serve the users effectively. However, for tractability, we assume that the UAVs will remain static during each content transmission. In essence, the UAVs will update their locations according to the mobility of the users after each content transmission is complete at a current location.

6.2.2 Transmission Model

Next, we introduce the models for transmission links between BBUs and UAVs, UAVs and users, and RRHs and users. For ease of exposition, a time slot τ is discretized into F equally spaced time intervals Δ_t, i.e., $\Delta_\tau = F\Delta_t$. The time interval Δ_t is chosen to be sufficiently small so that each user's location can be considered constant during t as in [162].

UAVs-Users Links

The mmWave propagation channel of the UAVs-user link is modeled using the standard log-normal shadowing model of [173]. The standard log-normal shadowing model can be used to model the line-of-sight (LoS) and non-line-of-sight (NLoS) links by choosing specific channel parameters. Therefore, the LoS and NLoS path loss of UAV k located at $(x_{\tau,k}, y_{\tau,k}, h_{\tau,k})$ transmitting a content to user i at interval t of time slot τ is [174] (in dB):

$$l_{t,ki}^{\text{LoS}}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right) = L_{FS}\left(d_0\right) + 10\mu_{\text{LoS}} \log\left(d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)\right) + \chi_{\sigma_{\text{LoS}}}, \tag{6.1}$$

$$l_{t,ki}^{\text{NLoS}}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right) = L_{FS}\left(d_0\right) + 10\mu_{\text{NLoS}} \log\left(d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)\right) + \chi_{\sigma_{\text{NLoS}}}, \tag{6.2}$$

where $\boldsymbol{w}_{\tau,t,k} = [x_{\tau,k}, y_{\tau,k}, h_{\tau,k}]$ is the coordinate of UAV k during time slot τ with $h_{\tau,k}$ being the altitude of UAV k at time slot τ. Also, $\boldsymbol{w}_{\tau,t,k} = [x_{t,i}, y_{t,i}]$ is the time-varying coordinate of user i at interval t. $L_{FS}\left(d_0\right)$ is the free space path loss given by $20\log\left(d_0 f_c 4\pi/c\right)$ with d_0 being the free-space reference distance, f_c being the carrier frequency and c being the speed of light. $d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right) = \sqrt{\left(x_{t,i} - x_{\tau,k}\right)^2 + \left(y_{t,i} - y_{\tau,k}\right)^2 + h_{\tau,k}^2}$ is the distance between user i and UAV k. μ_{LoS} and μ_{NLoS} are the path loss exponents for LoS and NLoS links. $\chi_{\sigma_{\text{LoS}}}$ and $\chi_{\sigma_{\text{NLoS}}}$

are the shadowing random variables which are, respectively, represented as the Gaussian random variables in dB with zero mean and σ_{LoS}, σ_{NLoS} dB standard deviations.

In our model, the probability of LoS connection depends on the environment, density and height of buildings, the locations of the user and the UAV, and the elevation angle between the user and UAV. The LoS probability is given by [170] and [163]:

$$\Pr\left(l_{t,ki}^{\text{LoS}}\right) = \left(1 + X \exp\left(-Y\left[\phi_t - X\right]\right)\right)^{-1}, \tag{6.3}$$

where X and Y are constants which depend on the environment (rural, urban, dense urban, or others) and $\phi_t = \sin^{-1}\left(h_{\tau,k}/d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)\right)$ is the elevation angle. Clearly, the average path loss from the UAV k to user i at interval t is [163]:

$$\bar{l}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right) = \Pr\left(l_{t,ki}^{\text{LoS}}\right) \times l_{t,ki}^{\text{LoS}} + \Pr\left(l_{t,ki}^{\text{NLoS}}\right) \times l_{t,ki}^{\text{NLoS}}, \tag{6.4}$$

where $\Pr\left(l_{t,ki}^{\text{NLoS}}\right) = 1 - \Pr\left(l_{t,ki}^{\text{LoS}}\right)$. Based on the path loss, the average signal-to-noise ratio (SNR) of user i located at $\boldsymbol{w}_{\tau,t,i}$ from the associated UAV k at interval t is given by:

$$\gamma_{t,ki}^{\text{V}} = \frac{P_{t,ki}}{10^{\bar{l}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)/10}\sigma^2}, \tag{6.5}$$

where $P_{t,ki}$ is the transmit power of UAV k to user i at time t, and σ^2 is the variance of the Gaussian noise. We assume that the total bandwidth available for each UAV is B_V which is equally divided among the associated users. The channel capacity between UAV k and user i for each content transmission will be $C_{\tau,ki}^{\text{V}} = \frac{1}{F_{\tau,i}} \sum_{t=1}^{F_{\tau,i}} \frac{B_V}{U_k} \log_2\left(1 + \gamma_{t,ki}^{\text{V}}\right)$, where U_k is the number of the users associated with UAV k and $F_{\tau,i}$ is the number of the intervals that user i uses to receive a content during time slot τ.

BBUs-UAVs Ground-to-Air Links

For the BBUs-UAVs (ground-to-air) link, we consider probabilistic LoS and NLoS links over the licensed band. Since the distance of the UAVs fronthaul link may be larger compared to the distance of the UAV-user link, the cellular band can provide a more reliable transmission and a smaller path loss compared to the mmWave channel. In such a model, NLoS links experience higher attenuations than LoS links due to the shadowing and diffraction loss. The LoS and NLoS path loss from the BBUs to UAV k at time t of time slot τ can be given by [170]:

$$L_{t,k}^{\text{LoS}} = d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,B}\right)^{-\beta}, \tag{6.6}$$

$$L_{t,k}^{\text{NLoS}} = \eta d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,B}\right)^{-\beta}, \tag{6.7}$$

where $\boldsymbol{w}_{\tau,t,B} = [x_B, y_B]$ is the location of the BBUs, and β is the path loss exponent. The LoS connection probability and the average SNR of the link between the BBUs and UAV k can be calculated using (6.3)-(6.5).

RRHs-Users Links

In our model, RRHs are grouped into E clusters. Then, the RRHs in each cluster use ZFBF to improve the users' rates. The received signals of the users associated with RRHs cluster q at interval t is:

$$\boldsymbol{b}_{t,q} = \sqrt{P_R}\boldsymbol{H}_{t,q}\boldsymbol{F}_{t,q}\boldsymbol{a}_{t,q} + \boldsymbol{n}, \tag{6.8}$$

where $\boldsymbol{H}_{t,q} \in \mathbb{R}^{U_q \times R_q}$ is the path loss matrix with U_q being the number of users associated with RRH cluster q, and R_q is the number of RRHs' antennas. P_R is the transmit power of each RRH which is assumed to be equal for all RRHs. $\boldsymbol{a}_{t,q} \in \mathbb{R}^{U_q \times 1}$ is the transmitted content at interval t and $\boldsymbol{n}_{t,q} \in \mathbb{R}^{U_q \times 1}$ is the noise power. Also, $\boldsymbol{F}_{t,q} = \boldsymbol{H}_{t,q}^{\mathrm{H}} \left(\boldsymbol{H}_{t,q}\boldsymbol{H}_{t,q}^{\mathrm{H}}\right)^{-1} \in \mathbb{R}^{R_q \times U_q}$ is the beamforming matrix [175]. We also assume that the bandwidth of each user associated wth the RRHs is B. Then, the received signal-to-interference-plus-noise-ratio (SINR) of user i in cluster $\mathcal{M}_q$ at interval t will be:

$$\gamma_{t,qi}^{\mathrm{H}} = \frac{P_R\left\|\boldsymbol{h}_{t,qi}\boldsymbol{f}_{t,qi}\right\|^2}{\underbrace{\sum_{j=1,j\neq q}^{E}\sum_{u\in\mathcal{U}_j} P_R\left\|\boldsymbol{h}_{t,ji}\boldsymbol{f}_{t,ju}\right\|^2}_{\text{other cluster RRHs interference}} + \underbrace{P_B g_{t,Bi} d_{t,Bi}^{-\beta}}_{\text{wireless fronthaul interference}} + \sigma^2},$$

where $\mathcal{M}_j$ is the set of the RRHs in group j, $\mathcal{U}_j$ is the set of the users associated with the RRHs in group j, $\boldsymbol{h}_{t,qi} \in \mathbb{R}^{1\times R_q}$ is the channel gain between the RRHs in cluster $\mathcal{M}_q$ and user i with $h_{t,ki} = g_{t,ki} d_{t,ki}\left(x_i, y_i\right)^{-\beta}$, $g_{t,ki}$ is the Rayleigh fading parameter at interval t, and $d_{t,ki}\left(x_i, y_i\right) = \sqrt{\left(x_{t,k} - x_{t,i}\right)^2 + \left(y_{t,k} - y_{t,i}\right)^2}$ is the distance between RRH k and user i at interval t. $\boldsymbol{f}_{t,qi} \in \mathbb{R}^{R_q \times 1}$ is the beamforming vector. Given (9), the channel capacity between RRH cluster $\mathcal{M}_q$ and user i for each content transmission is:

$$C_{\tau,qi}^{\mathrm{H}} = \frac{1}{F_{\tau,i}}\sum_{t=1}^{F_{\tau,i}} B\log_2\left(1 + \gamma_{t,qi}^{\mathrm{H}}\right). \tag{6.9}$$

6.2.3 Quality-of-Experience Model

Given the proposed models in the previous subsections, here, we present the QoE model for each user. The *quality-of-experience* of each user is formally defined as a concrete human-in-the-loop metric that captures each user's data rate, delay, and device type.

Delay

In the considered CRAN system, contents can be transmitted to the users via three types of links: (a) content server-BBUs-RRHs-user, (b) content server-BBUs-UAV-user, and (c) UAV cache-user. The backhaul link connecting the cloud to the core network is assumed to be fiber and, therefore, its delay is neglected. We assume that the capacity of the wired fronthaul links between the BBUs

and the RRHs is limited to a maximum rate of v_F for all users. Consequently, the fronthaul rate for each user receiving a content from the RRHs will be $v_{FU} = v_F/N_{FR}$ with N_{FR} being the number of the users receive contents from the RRHs. Thus, the delay of a user i receiving content n over the three types of links at each time slot τ can be written as:

$$D_{\tau,i,n} = \begin{cases} \frac{L}{v_{FU}} + \frac{L}{C^{\mathrm{H}}_{\tau,qi}}, & \text{link } (a), \\ \frac{L}{C^{F}_{\tau,k}} + \frac{L}{C^{\mathrm{V}}_{\tau,ki}}, & \text{link } (b), \\ \frac{L}{C^{\mathrm{V}}_{\tau,ki}}, & \text{link } (c), \end{cases} \tag{6.10}$$

where $C^{F}_{\tau,k}$ is the rate of content transmission from the BBUs to UAV k which is calculated analogously to (6.2.2) and (6.9). Next, we derive the lower bound on the delay that each user can tolerate for each content transmission.

Proposition 7. The lower bound of the delay for each user i receiving content n are given by:

$$\min\left\{\frac{L}{v_F}, \frac{L}{C^{\max}_K}\right\} \leq D_{\tau,i,n}, \tag{6.11}$$

where $C^{\max}_K = B_V \log_2\left(1 + \frac{P_{\max}}{10^{(L_{FS}(d_0)+10\mu_{\mathrm{LoS}}\log(h_{\min})-4\sigma_{\mathrm{LoS}})/10}\sigma^2}\right)$ with $P_{\max}$ being the maximum transmit power of each UAV, and $h_{\min}$ being the minimum altitude of the UAV.

Proof. See Appendix C.1. □

From Proposition 7, we can see that the minimum delay of each user depends on the rate of the fronthaul links and the maximum transmit power of the UAVs. Therefore, we can improve the QoE of each user by adjusting the UAV's transmit power. In particular, as the number of users increases and the rate of fronthaul links decreases, the QoE requirement of users can be satisfied by adjusting the UAVs' transmit power. Note that, the upper bound of the delay Δ_τ is set by the system requirement. Using the results of Proposition 7, we can categorize the sensitivity to the delay into five groups using the popular mean opinion score (MOS) model [176] which is often used to measure the QoE of a wireless user. The mapping between delay and MOS model [176] is given by:

$$\bar{D}_{\tau,i,n} = \frac{\Delta_\tau - D_{\tau,i,n}}{\Delta_\tau - \min\left\{\frac{L}{v_F}, \frac{L}{C^{\max}_K}\right\}}, \tag{6.12}$$

which is shown in Table 6.2.

Device Type

The screen size of each device type of the user will also affect the QoE perception of the user, especially for video-oriented applications. Indeed, users who own devices that have larger screens (such as tablets) will be more sensitive to QoE compared to those who own smaller devices (such

Table 6.2: Mean opinion score model

QoE	Poor	Fair	Good	Very Good	Excellent
Interval scale	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1

as small smartphones). We capture the impact of the screen size of each user i using a parameter S_i that reflects the diameter length of the user's device. Typically, devices with a larger screen size, can display content at a higher resolution thus requiring a higher data rate. We assume that the rate requirement of user i with device S_i receiving a content n at interval t is $\delta_{S_i,n} = S_i\hat{C}_n$, where $\hat{C}_n$ is the rate requirement of each user receiving content n. The mapping from the rate requirement of user device to the MOS model is:

$$V_{t,i} = \begin{cases} 1, & j \geq \delta_{S_i,n}, \\ 0, & j < \delta_{S_i,n}, \end{cases} \tag{6.13}$$

where $j \in \left\{C^{\mathrm{V}}_{t,ki}, C^{\mathrm{H}}_{t,qi}\right\}$. Here, $V_{t,i} = 1$ indicates that the user's data rate satisfies the requirement of its device type and 0 represents the user's data rate cannot satisfy the requirement. In this case, the rate requirement of of a user's device is mapped to the MOS. The QoE of each user i receiving content n at time slot τ can be given by [176]:

$$Q_{\tau,i,n} = \zeta_1 \bar{D}_{\tau,i,n} + \frac{\zeta_2}{F_{\tau,i}} \sum_{t=1}^{F_{\tau,i}} V_{t,i}, \tag{6.14}$$

where q_1 and q_2 are weighting parameters with $\zeta_1 + \zeta_2 = 1$.

6.2.4 Problem Formulation

Here, we first find the minimum rate required to meet the QoE requirement of each user associated with the UAVs. Next, we determine the minimum transmit power of each UAV required to meet the QoE threshold of the associated users. Finally, we formulate the minimization problem. From Table 6.2, we can see that, for $0.8 \leq \bar{D}_{\tau,i,n} \leq 1$, the MOS of delay will be "Excellent", which means that the delay is minimized. In this case, $\bar{D}_{\min} = 0.8$ is the minimum value that maximizes the delay component of user i's QoE, during the transmission of a given content n. We define the rate that achieves the optimal delay as the delay requirement and also, define the rate that meets the rate requirement of device as the device rate requirement. Consider the transmission between a UAV k located at $\boldsymbol{w}_{\tau,t,k}$ and a user i located at coordinates $\boldsymbol{w}_{\tau,t,i}$. From (6.10), the delay requirement for UAV k transmitting content n to user i at time slot τ is:

$$C^R_{\tau,ki,n} = \begin{cases} \dfrac{L}{\left(\Delta_\tau - \bar{D}_{\min}\left(\Delta_\tau - \min\left\{\frac{L}{v_F}, \frac{L}{C^{\max}_k}\right\}\right) - \frac{L}{C^F_{\tau,k}}\right)}, & n \notin \mathcal{C}_k, \\ \dfrac{L}{\left(\Delta_\tau - \bar{D}_{\min}\left(\Delta_\tau - \min\left\{\frac{L}{v_F}, \frac{L}{C^{\max}_k}\right\}\right)\right)}, & n \in \mathcal{C}_k. \end{cases} \tag{6.15}$$

From (6.15), we can see that, by storing content n at cache of UAV k, the delay requirement for minimizing delay decreases.

Let $\delta_{S_i,n}$ be the device rate requirement of user i associated with a UAV. Clearly, the QoE is maximized when $C^{\mathrm{V}}_{t,ki} \geq \max\left\{C^{R}_{\tau,ki,n}, \delta_{S_i,n}\right\}$. Hence, the minimum rate required to maximize the user's QoE is $\delta^{R}_{i,n} = \max\left\{C^{R}_{\tau,ki,n}, \delta_{S_i,n}\right\}$. Based on (6.5), the minimum transmit power needed to guarantee the QoE requirement of user i receiving content n at interval t is:

$$P^{\min}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \delta^{R}_{i,n}, n\right) = \left(2^{\delta^{R}_{i,n}U_k/B_V} - 1\right)\sigma^2 10^{\bar{l}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)/10}. \tag{6.16}$$

From (6.16), we can see that the minimum transmit power of UAV k transmitting content n to user i depends on the UAV's location, the rate needed to satisfy the QoE requirement of user i, and the transmitted content n.

Given this system model, our goal is to find an effective deployment of cache-enabled UAVs to enhance the QoE of each user while minimizing the transmit power of the UAVs. This problem involves predicting the content request distribution and periodic locations for each user, finding the optimal contents to cache at the UAVs, determining the users' associations and adjusting the locations[1], and transmit power of the UAVs. This problem can be formulated as follows:

$$\min_{\mathcal{C}_k, \mathcal{U}_{\tau,k}, \boldsymbol{w}_{\tau,t,k}} \sum_{\tau=1}^{T} \sum_{k\in\mathcal{K}} \sum_{i\in\mathcal{U}_{\tau,k}} \sum_{t=1}^{F_{\tau,i}} P^{\min}_{\tau,t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \delta^{R}_{i,n}, n_{\tau,i}\right), \tag{6.17}$$

$$\text{s. t. } h_{\min} \leq h_{\tau,k}, k \in \mathcal{K}, \tag{6.17a}$$

$$m \neq j, m, j \in \mathcal{C}_k, \mathcal{C}_k \subseteq \mathcal{N}, k \in \mathcal{K}, \tag{6.17b}$$

$$0 < P^{\min}_{\tau,t,ki} \leq P_{\max}, i \in \mathcal{U}, k \in \mathcal{K}, \tag{6.17c}$$

where $P^{\min}_{\tau,t,ki}$ is the minimum transmit power of UAV k to user i at interval t during time slot τ. $n_{\tau,i}$ is the content that user i requests at time slot τ, $\mathcal{U}_{\tau,k}$ is the set of the users that are associated with UAV k at time slot τ. $h_{\min}$ is the minimum altitude that each UAV can reach at time slot τ. Here, (6.17b) captures the fact that each cache storage unit at the UAV stores a single, unique content, and (6.17c) indicates that the transmit power of the UAVs should be minimized. Since the problem as per (6.17) is to satisfy the rate needed for meeting each user's QoE requirement during the next time period T, the predictions of users behavior will directly impact the solution. From (6.17), we can see that the prediction of the users' mobility patterns enable the BBUs to find the optimal locations of the UAVs. Moreover, by predicting the users' content request distribution the BBUs can determine the most popular content to cache at the UAVs. In addition, since the content transmission link affects the transmit power that the UAV must use to satisfy the user's QoE requirement, the problem of minimizing the transmit powers of the UAVs in (6.17) inherently incorporates the caching constraints.

[1]Typically, the speed of a UAV can reach up to 30 m/s while the average speed of each pedestrian ground user is less than 2 m/s. Therefore, in our model, we ignore the time duration that each UAV uses to change its location.

6.3 Conceptor Echo State Networks for Content and Mobility Predictions

In this section, we propose a prediction algorithm using the framework of ESN with conceptors, to find the users' content request distributions and their mobility patterns. The predictions of the users' content request distribution and their mobility patterns will then be used in Section 6.4 to find the user-UAV association, optimal locations of the UAVs and content caching at the UAVs. Echo state networks are a special type of recurrent neural networks designed for performing non-linear systems forecasting [177]. The ESN architecture is based on a randomly connected recurrent neural network, called reservoir, which is driven by a temporal input. The state of the reservoir is a rich representation of the history of the inputs so that a simple linear combination of the reservoir units is a good predictor of the future inputs. In our model, the reservoir will be combined with the input to store the users' context information and will also be combined with the trained output matrix to output the predictions of the users' content request distribution and mobility patterns. Here, a user's *context* is defined as the current state and attribute of a user including time, week, gender, occupation, age, and device type (e.g., tablet or smartphone). Therefore, an ESN-based approach can use the users' context to predict the corresponding behavior such as content request and mobility.

Compared to traditional neural network and deep learning approaches such as in [164], an ESN-based approach can quickly learn the mobility pattern and content request distribution without requiring significant training data due to the use of the echo state property. However, traditional ESN-based prediction algorithms such as in [156] can be trained to predict only one mobility pattern for each user. In particular, to predict the weekly mobility pattern of each user using the traditional ESN approach, the users' context information for an entire week need to be used as input of the ESNs that act as one non-linear system. In this conventional ESN approach, it is not possible to separate the users' contexts in a week into several days and train the ESNs to predict the user's mobility in each day with one specific non-linear system. To enable the ESN algorithm to predict the user's mobility pattern and content request distribution with various non-linear systems, the notion of a *conceptor* as defined in [178], is an effective solution that allows characterizing the ESN's reservoir. Conceptors enable an ESN to perform a large number of mobility and content request patterns predictions. Moreover, new patterns can be added to the reservoir of the ESN without interfering with previously acquired ones. For each ESN algorithm, an ESN can record a limited number of history input data due to the echo state property of each ESN. Consequently, the learned pattern will be removed as the recorded input data is updated. Here, we call the ability of recording a limited number of history input data as the *memory* of the ESN's reservoir. The idea of a conceptor can be used to allocate any free memory of an ESN's reservoir to the new learned patterns of the mobility and content request distribution.

Next, we first introduce the components of a conceptor ESN-based prediction algorithm. Then, we formulate the conceptor ESN algorithm to predict the content request distribution and mobility patterns of the users.

6.3.1 Conceptor ESN Components

The conceptor ESN-based prediction approach consists of five components: a) agents, b) input, c) output, d) ESN model, and e) conceptor. Since the content request and mobility pattern are user-specific, we design the specific components for the algorithms of the content request distribution and mobility pattern predictions, separately.

Content request distribution prediction

The content request distribution prediction algorithm has the following components:

• *Agent*: The agent in our ESNs is the cloud. Since each ESN scheme typically performs a content request distribution prediction for just one user, the cloud's BBUs must implement U conceptor ESN algorithms.

• *Input:* The conceptor ESN takes input as a vector $\boldsymbol{x}_{t,j} = [x_{tj1}, \cdots, x_{tjN_x}]^{\mathrm{T}}$ that represents the context of user j at time t which includes gender, occupation, age, and device type (e.g., tablet or smartphone). Here, N_x is the number of properties that constitute the context information of user j. The vector $\boldsymbol{x}_{t,j}$ is then used to determine the content request distribution $\boldsymbol{y}_{t,j}$ for user j. Note that, the input of the ESNs is the information related to the users' content requests. Our goal is to predict the content request distribution using the context of each user.

• *Output:* The output of the content request distribution prediction ESN at time t is a vector of probabilities $\boldsymbol{y}_{t,j} = [p_{tj1}, p_{tj2}, \ldots, p_{tjN}]$ that represents the probability distribution of content request of user j with p_{tjn} being the probability that user j requests content n at time t.

• *ESN Model:* An ESN model for each user j can find the relationship between the input $\boldsymbol{x}_{t,j}$ and output $\boldsymbol{y}_{t,j}$, thus building the function between the user's context and the content request distribution. Mathematically, the ESN model consists of the output weight matrix $\boldsymbol{W}_j^{\alpha,\text{out}} \in \mathbb{R}^{N \times N_w}$ and the dynamic reservoir containing the input weight matrix $\boldsymbol{W}_j^{\alpha,\text{in}} \in \mathbb{R}^{N_w \times N_x}$, and the recurrent matrix $\boldsymbol{W}_j^{\alpha} \in \mathbb{R}^{N_w \times N_w}$ with N_w being the number of the dynamic reservoir units. For each user j, the dynamic reservoir will be combined with the input $\boldsymbol{x}_{t,j}$ to store the history context of user j. The output weight matrix $\boldsymbol{W}_j^{\alpha,\text{out}}$ with the reservior is trained to approximate the prediction function. The ESN model of user j is initially randomly generated following a uniform distribution. To ensure that the reservoir has the echo state property, $\boldsymbol{W}_j^{\alpha}$ is defined as a sparse matrix with a spectral radius less than one [179].

• *Conceptors:* For content request distribution prediction, we collect the users' context information and the corresponding content requests during the same time slots for different weeks to train one content request distribution. We refer to each content request distribution as one prediction pattern. Given a sequence of the reservoir states $\boldsymbol{v}_j^i = \left[\boldsymbol{v}_{1,j}^i, \ldots, \boldsymbol{v}_{t,j}^i\right]$ with $\boldsymbol{v}_{t,j}^i = \left[v_{t,j1}^i, \ldots, v_{t,jN_w}^i\right]^{\mathrm{T}}$ being the reservoir state of prediction pattern i at time t and the state correlation matrix $\boldsymbol{R}_j^i =$

$\mathbb{E}\left[\boldsymbol{v}_{t,j}^{i}\left(\boldsymbol{v}_{t,j}^{i}\right)^{\mathrm{T}}\right]$, the conceptor of prediction pattern i will be [178]:

$$\boldsymbol{M}_{j}^{i}=\boldsymbol{R}_{j}^{i}\left(\boldsymbol{R}_{j}^{i}+\chi^{-2}\boldsymbol{I}\right)^{-1}, \tag{6.18}$$

where χ is *aperture* defined in [178]. The aperture χ needs to be appropriately set for accurately learning several mobility patterns. When the aperture is small, the reservoir of the ESN slightly changes for learning each new pattern. However, for a large aperture, the reservoir of the ESN changes significantly.

Mobility pattern prediction

The components of mobility pattern prediction algorithm are:

• *Agents*: The agents in our conceptor ESNs are the BBUs. Since each ESN scheme typically performs mobility prediction for only one user, the BBUs must also implement U conceptor ESN algorithms.

• *Input:* $\boldsymbol{m}_{t,j}=\left[m_{tj1},\cdots,m_{tjN_x+1}\right]^{\mathrm{T}}$ represents the current location of user j and the context of this user at time t. Using input $\boldsymbol{m}_{t,j}$, the future locations of user j can be predicted.

• *Output:* $\boldsymbol{s}_{t,j}=\left[s_{tj1},\cdots,s_{tjN_s}\right]^{\mathrm{T}}$ represents the predicted locations of user j in the next time slots, where N_s is the number of locations in the next N_s time duration H.

• *ESN Model:* The ESN model of mobility prediction consists of the output weight matrix $\boldsymbol{W}_{j}^{\text{out}}\in\mathbb{R}^{N_s\times N_w}$, the dynamic reservoir containing the input weight matrix $\boldsymbol{W}_{j}^{\text{in}}\in\mathbb{R}^{N_w\times N_x+1}$, and the recurrent matrix $\boldsymbol{W}_{j}\in\mathbb{R}^{N_w\times N_w}$. The generation of the mobility prediction ESN model is similar to the one in the content request distribution prediction case.

• *Conceptors:* For mobility pattern prediction, we consider each user's mobility in each day during one week as one prediction pattern. The expression of the conceptors is the same as the one for the content request distribution given in (6.18).

6.3.2 Conceptor ESN Algorithm for Content and Mobility Predictions

Here, we present the proposed conceptor ESN algorithm to predict the content request distribution and mobility. The proposed algorithm consists of two stages: training and prediction stages.

Training Stage

The dynamic reservoir state $\boldsymbol{v}_{t,j}^{i}$ of prediction pattern i for user j at time t which is used to store the states of user j is given by [179]:

$$\boldsymbol{v}_{t,j}^{i}=f\left(\boldsymbol{W}_{j}^{\alpha}\boldsymbol{v}_{t-1,j}^{i}+\boldsymbol{W}_{j}^{\alpha,\text{in}}\boldsymbol{x}_{t,j}\right), \tag{6.19}$$

where $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. Note that, we consider the input and corresponding prediction output as a training data. In this case, we use N_{tr} training data that consists of N_{tr} users' contexts and the corresponding content request to calculate the conceptors and train the output weight matrix $\boldsymbol{W}_j^{\alpha,\text{out}}$. Based on N_{tr} training data and (6.19), the reservoir states before update for each prediction pattern j is $\boldsymbol{v}_{\text{old},j}^i = \left[0, \boldsymbol{v}_{1,j}^i, \ldots, \boldsymbol{v}_{N_{tr}-1,j}^i\right]$ and the updated reservoir states are $\boldsymbol{v}_j^i = \left[\boldsymbol{v}_{1,j}^i, \ldots, \boldsymbol{v}_{N_{tr},j}^i\right]$. The matrix $\boldsymbol{v}_{\text{old},j}^i$ will be used to train an *input simulation matrix* $\boldsymbol{D}_j \in \mathbb{R}^{N_w \times N_w}$ and $\boldsymbol{v}_j^i$ will be combined with the updated reservoir states of other prediction patterns to train the output weight matrix.

Then, $\boldsymbol{D}_j$ will be combined with output weight matrix $\boldsymbol{W}_j^{\alpha,\text{out}}$ to predict the content request distribution pattern for each user. For each added learning pattern i of each user j, the update of $\boldsymbol{D}_j$ will be [178]:

$$\boldsymbol{D}_j = \boldsymbol{D}_{\text{old},j} + \boldsymbol{D}_{\text{inc},j}^i, \tag{6.20}$$

where $\boldsymbol{D}_{\text{inc},j}^i = \left(\left(\boldsymbol{S}\boldsymbol{S}^{\text{T}}/N_{tr} + \chi^{-2}\boldsymbol{I}\right)^{\dagger}\boldsymbol{S}\boldsymbol{T}^{\text{T}}/N_{tr}\right)^{\text{T}}$ with $\boldsymbol{S} = \boldsymbol{F}_j^{i-1}\boldsymbol{v}_{\text{old},j}^i$ and $\boldsymbol{T} = \boldsymbol{W}_j^{\alpha,\text{in}}\boldsymbol{x}_j^i - \boldsymbol{D}_{\text{old},j}\boldsymbol{v}_{\text{old},j}^i$. Here, $\boldsymbol{F}_j^{i-1} = \neg \vee \left\{\boldsymbol{M}_j^1, \ldots, \boldsymbol{M}_j^{i-1}\right\}$ is the free memory of the reservoir with $\neg$ and $\vee$ being the boolean operators [178], and $\boldsymbol{x}_j^i = \left[\boldsymbol{x}_{1,j}^i, \ldots, \boldsymbol{x}_{N_{tr},j}^i\right]$ is the input sequences of prediction pattern i. During the learning of each pattern i of user j, the conceptor $\boldsymbol{M}_j^i$ can be computed using (6.18).

In our proposed ESN algorithm, the output weight matrix $\boldsymbol{W}_j^{\alpha,\text{out}}$ is trained in an offline manner using ridge regression [179] to approximate the prediction function which is given by:

$$\boldsymbol{W}_j^{\alpha,\text{out}} = \boldsymbol{y}_j\boldsymbol{v}_j^{\text{T}}\left(\boldsymbol{v}_j^{\text{T}}\boldsymbol{v}_j + \lambda^2\mathbf{I}\right)^{-1}, \tag{6.21}$$

where $\boldsymbol{v}_j = \left[\boldsymbol{v}_j^1, \boldsymbol{v}_j^2, \ldots, \boldsymbol{v}_j^{N_M}\right]^{\text{T}}$ with $\boldsymbol{v}_j^i = \left[\boldsymbol{v}_{1,j}^i, \ldots, \boldsymbol{v}_{N_{tr},j}^i\right]$ being the reservoir state sequence of prediction pattern i for user j, λ is the learning rate, and N_M being the number of the prediction patterns of each user's content request distribution. In (6.21), $\boldsymbol{v}_j^i$ can also be used to calculate the conceptor $\boldsymbol{M}_j^i$ for prediction pattern i of user j.

Prediction Stage

Based on the learning stage, we can use the input simulation matrix $\boldsymbol{D}_j$, conceptors $\boldsymbol{M}_j = \left[\boldsymbol{M}_j^1, \ldots, \boldsymbol{M}_j^{N_M}\right]$, and output weight matrix $\boldsymbol{W}_j^{\alpha,\text{out}}$ to obtain the corresponding predictions. In the prediction stage, the reservoir state of pattern i of user j is [178]:

$$\boldsymbol{v}_{t,j}^i = \boldsymbol{C}_j^i f\left(\boldsymbol{W}_j^{\alpha}\boldsymbol{v}_{t-1,j}^i + \boldsymbol{D}_j\boldsymbol{v}_{t-1,j}^i\right). \tag{6.22}$$

From (6.22), we can see that the conceptor of pattern j, $\boldsymbol{C}_j$, controls the update of the reservoir states. By changing the conceptor $\boldsymbol{C}_j$, the ESN can predict different patterns in one ESN architecture. The prediction of content request distribution i for user j can be given by:

$$\boldsymbol{y}_{t,j} = \boldsymbol{W}_j^{\alpha,\text{out}}\boldsymbol{v}_{t,j}^i. \tag{6.23}$$

Table 6.3: Proposed conceptor ESN prediction algorithm

Inputs: N_{tr} training data,
Initialize: $\boldsymbol{W}_j^{\alpha,\text{in}}$, $\boldsymbol{W}_j^{\alpha}$, $\boldsymbol{W}_j^{\alpha,\text{out}}$, $\boldsymbol{y}_j = 0$, $\boldsymbol{D}_j = 0$.

Training Stage:

for each prediction pattern i **do**.

if reservoir memory space $\boldsymbol{F}_j^{i-1} > 0$ **do**.

(a) BBUs collect the states $\boldsymbol{v}_{\text{old},j}^i$ and $\boldsymbol{v}_j^i$ to update $\boldsymbol{D}_j$, using (6.20).

(b) BBUs use the states $\boldsymbol{v}_j^i$ to calculate the conceptor $\boldsymbol{C}_j^i$ using (6.18).

else

(c) increase reservoir matrix $\boldsymbol{W}_j^{\alpha}$, re-train all prediction patterns.

end if

end for

(d) BBUs collect states for all patterns $\boldsymbol{v}_j$ to train $\boldsymbol{W}_j^{\alpha,\text{out}}$, by (6.21).

Prediction Stage:

(a) BBUs chooses the conceptor to obtain the corresponding reservoir state, using (6.22).

(b) Get the prediction of content request distribution based on (6.23) .

Output: Prediction $\boldsymbol{y}_{t,j}$

From (6.22) and (6.23), we can see that the conceptor ESN algorithm exploits an input simulation matrix $\boldsymbol{D}_j$ to control the memory of ESN reservoir. The conceptor ESN algorithm for predicting the content request distribution of each user j is shown in Table 6.3.

As shown in Table 6.3, the proposed conceptor ESN algorithm can learn each prediction pattern by a unique non-linear system. This property of the proposed algorithm enables the ESNs to perform the users' behavior predictions using different non-linear systems during different time periods. Furthermore, using the proposed algorithm, one can have the information of the reservoir memory and extract a specific prediction pattern from the learned patterns.

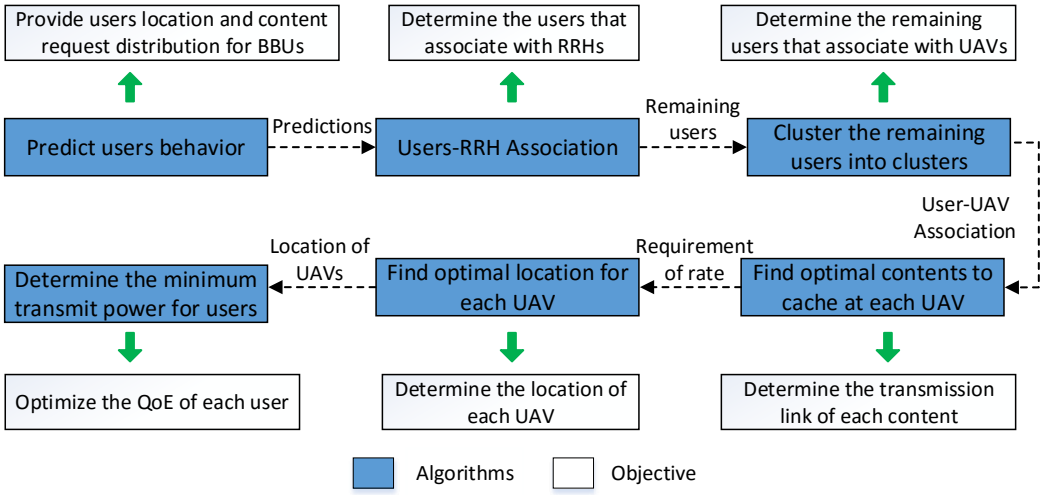


Figure 6.2: The procedure used for solving the optimization problem in (6.17).

6.4 Optimal Location and Content Caching for UAVs

In this section, we use the content request distribution and mobility patterns predictions resulting from the proposed conceptor ESN algorithm in Section 6.3 to solve the problem in (6.17). In our model, a subset of the users selected by the BBUs are connected to the RRHs. The remaining users are clustered into K clusters and each UAV provides service for one cluster. Based on the associations and predictions, we determine which contents to cache at each UAV and find the optimal location of each UAV. Finally, we analyze the implementation and complexity of the proposed algorithm. Fig. 6.2 summarizes the proposed framework that is used to solve the problem in (6.17).

6.4.1 Users-RRH Association

We find the user-RRH association based on the predicted users' locations at the next time interval. Clearly, the prediction accuracy of the users' locations will directly affects the users association. A user is associated with RRHs if the following condition is satisfied:

Theorem 10. Given minimum $\bar{D}_{\min}$ and device screen size S_i of each user i, user i will be associated with a cluster k of RRHs if the following rate requirement is satisfied:

$$C^{H}_{t,qi} \geq \max\left\{\frac{L}{\left(\Delta_\tau - \bar{D}_{\min}\left(\Delta_\tau - \min\left\{\frac{L}{v_F}, \frac{L}{C^{\max}_k}\right\}\right) - \frac{L}{v_{FU}}\right)}, \delta_{S_i,n}\right\}. \tag{6.24}$$

Proof. See Appendix C.2. □

From Theorem 10, we can see that the user-RRH association depends on the fronthaul rate of each user, the delay requirement, and the device rate requirement. From (6.24), we can see that the fronthaul rate of each user decreases as the number of the users associated with the RRHs increases. Clearly, the decrease of the fronthaul rate for each user will improve the delay requirement. Note that, the energy consumption of the RRHs is not considered in our optimization problem as the RRHs can have continuous power supply on the ground while the UAVs are powered by on-board batteries with limited energy. Therefore, it is natural to allow the users to first associate with the RRHs when the RRHs can satisfy the users' QoE requirements.

6.4.2 Optimal Content Caching for UAVs

In our model, the remaining users who are not associated with RRHs, will be served by the UAVs. In this case, the users-UAVs associations need to be determined. To this end, we use K-mean clustering approach [171] in which the users are clustered into K groups. By implementing the K-mean clustering approach, the users that are close to each other will be grouped into one cluster. Thereby, each UAV services one cluster and the user-UAV association will be determined. Then, based on the UAV association, we find the optimal contents to cache at each UAV. The content caching will reduce the transmission delay and, hence, decrease the delay requirement. From (6.15), we can see that, optimal contents to store at the UAV cache lead to maximum reduction of the UAV's transmit power. The reduction of UAV transmit power is caused by the decrease of the delay requirement. Let $\boldsymbol{p}_{j,i} = [p_{j,i1}, p_{j,i2}, \ldots, p_{j,iN}]$ be the content request distribution of user i during period j that consists of H time slots. The optimal contents that will be stored at each UAV cache can be determined based on the following theorem.

Theorem 11. The optimal set of contents $\mathcal{C}_k$ to cache at each UAV k during period T is:

$$\mathcal{C}_k = \arg\max_{\mathcal{C}_k} \sum_{j=1}^{T/H} \sum_{\tau=1}^{H} \sum_{i \in \mathcal{U}_{\tau,k}} \sum_{n \in \mathcal{C}_k} (p_{j,in} \Delta P_{j,\tau,ki,n}), \tag{6.25}$$

where $\Delta P_{j,\tau,ki,n} =$

$$\begin{cases} P^{\min}_{\tau,ki}\left(C^R_{\tau,ki}\right)_{n \notin \mathcal{C}_k} - P^{\min}_{\tau,ki}\left(C^R_{\tau,ki}\right)_{n \in \mathcal{C}_k}, C^R_{\tau,ki,n \notin \mathcal{C}_k} \geq \delta_{S_i,n}, \\ P^{\min}_{\tau,ki}(\delta_{S_i,n})_{n \notin \mathcal{C}_k} - P^{\min}_{\tau,ki}\left(C^R_{\tau,ki}\right)_{n \in \mathcal{C}_k}, \delta_{S_i,n} > C^R_{\tau,ki,n \notin \mathcal{C}_k}, \end{cases}$$

with $P^{\min}_{\tau,ki}\big(\boldsymbol{w}_{\tau,t,k}, C^R_{\tau,ki}, n\big)$ being simplified to $P^{\min}_{\tau,ki}\big(C^R_{\tau,ki}\big)$.

Proof. See Appendix C.3. □

From Theorem 11, we can see that when the fronthaul rates of all users are the same, the transmit power reduction $\Delta P_{j,\tau,ki,n}$ will be a constant. Subsequently, the optimal content caching becomes

$\mathcal{C}_k = \arg\max_{\mathcal{C}_k} \sum_{j=1}^{T/H} \sum_{\tau=1}^{H} \sum_{i \in \mathcal{U}_{\tau,k}} \sum_{n \in \mathcal{C}_k} p_{j,in}$ which corresponds to the result given in [156]. From Theorem 11, we can see that the content caching depends on the pre-knowledge of users association as well as the content request distribution of each user. Therefore, by predicting the mobility pattern and content request distribution for each user, we can determine the optimal content to cache.

6.4.3 Optimal Locations of UAVs

Here, we determine the optimal UAVs' locations where the UAVs can serve their associated users using minimum transmit power. Once each UAV selects the suitable contents to cache, the transmission link (BBUs-UAV-user or UAV-user) for each content and the delay requirement $C^R_{\tau,ki,n}$ in (6.15) are determined. In this case, the rate $\delta^R_{i,n}$ which is used to meet the QoE requirment of each user is also determined. Next, we derive a closed-form expression for the optimal location of UAV k during time slot τ in two special cases.

Theorem 12. To minimize the transmit power of UAV k, the optimal locations of UAV k during time slot τ for cases: a) UAV k positioned at low altitudes compared to the size of its corresponding coverage, $h^2_{\tau,k} \ll (x_{t,i} - x_{\tau,k})^2 + (y_{t,i} - y_{\tau,k})^2$ and $\mu_{\text{NLoS}} = 2$, b) UAV k is placed at high altitudes compared to the size of its corresponding coverage, $h^2_{\tau,k} \gg (x_{t,i} - x_{\tau,k})^2 + (y_{t,i} - y_{\tau,k})^2$, are given by:

$$x_{\tau,k} = \frac{\sum_{i \in \mathcal{U}_{\tau,k}} \sum_{t=1}^{F_{\tau,i}} x_{t,i} \psi_{t,ki}}{\sum_{i \in \mathcal{U}_{\tau,k}} \sum_{t=1}^{F_{\tau,i}} \psi_{t,ki}}, y_{\tau,k} = \frac{\sum_{i \in \mathcal{U}_{\tau,k}} \sum_{t=1}^{F_{\tau,i}} y_{t,i} \psi_{t,ki}}{\sum_{i \in \mathcal{U}_{\tau,k}} \sum_{t=1}^{F_{\tau,i}} \psi_{t,ki}}, \tag{6.26}$$

where $\psi_{t,ki} = \left(2^{\delta^R_{i,n}/B} - 1\right) \sigma^2 10^{(L_{FS}(d_0) + \chi_\sigma)/10}$ with $\sigma = \begin{cases} \sigma_{\text{NLoS}}, & \text{for case } a), \\ \sigma_{\text{LoS}}, & \text{for case } b). \end{cases}$

Proof. See Appendix C.4. □

Using Theorem 12, we can find the optimal locations of the UAVs given the users association and altitude $h_{\tau,k}$ for the two special cases. For more generic cases, it is highly challenging to find the optimal UAVs' locations using derivation, since the UAV's altitude depends on the x and y coordinates of the UAV. Therefore, we use a learning algorithm given in [180] and [181] to find a sub-optimal solution. The learning algorithm can learn the network state and exploit different actions to adapt the UAV's location according to the network. After the learning step, each UAV will find a sub-optimal location to service the users in a power efficient way.

6.4.4 Implementation and Complexity

The complexity of the proposed algorithm pertains to two components: the conceptor ESN algorithm and the optimization algorithm. In the conceptor ESN algorithm, the ESN needs to be trained and hence, the complexity of the training depends on the users' data. The complexity of implementing the conceptor-based ESN algorithm depends on the number of users and the number of needed predictions. Since predictions occur once every H, during a period T, the ESN algorithm needs to be executed $\frac{T}{H}$ times. Since the ESN needs to predict the users' content request distribution and mobility patterns, the total complexity of the ESN algorithm is $O(U \times 2T/H)$.

The optimization algorithm can be divided into three algorithms: 1) user-UAV association, 2) caching optimization, and 3) optimal location of the UAVs. The complexity of the user-UAV association is $O\left(U^{2K+1} \log U\right)$ [182]. The complexity will be significantly decreased since the number of users that associated with UAVs decreases as the users will first associate with RRHs. For each UAV k, the complexity of the caching optimizatation algorithm depends on the number of the associated users U_k, the number of the contents N and the number of predictions during a period T, $\frac{T}{H}$. Therefore, the complexity of the caching algorithm can be given as $O\left(U_k \times N \times \frac{T}{H}\right)$. However, in practical scenarios, only a few contents have high request probabilities. In this practical case, the complexity of the caching algorithm can be significantly reduced. Finally, for each UAV k, the complexity of the optimal UAV algorithm depends on the number of the associated users U_k and the locations of each associated user during each time slot. Therefore, the complexity of the optimal UAV location algorithm is $O\left(U_k \times F\right)$.

6.5 Simulation Results

For our simulations, the content request data that the ESN uses to train and predict content request distribution is obtained from *Youku* of *China network video index*[2]. Here, one circular CRAN area with a radius $r = 500$ m is considered with $U = 70$ uniformly distributed users and $R = 20$ uniformly distributed RRHs. The detailed parameters are listed in Table V. Actual pedestrian mobility data is measured from 100 students at the *Beijing University of Posts and Telecommunications*. We recorded the daily mobility pattern of each student and collected their locations every hour during 9:00 am - 12:00 pm in a period of over two months. For comparison purposes, we investigate: a) optimal algorithm that has a prior knowledge of the accurate user's mobility patterns and content request distribution, b) ESN algorithm in [156] to predict the content request distribution and mobility pattern, and c) random caching with ESN algorithm in [156] to predict content request distribution. All statistical results are averaged over 5000 independent runs. The accuracy of ESN prediction is measured by normalized root mean square error [178].

Fig. 6.3 shows how the memory of the conceptor ESN reservoir changes as the number of the mobility patterns that were learned by the conceptor ESN varies. In this figure, the green curve

[2]The data is available at `http://index.youku.com/`.

Table 6.4: System parameters in the cache-enabled UAV network

Parameter	Value	Parameter	Value	Parameter	Value
F	1000	Y	0.13	P_B	30 dBm
X	11.9	N	25	P_R	20 dBm
$\chi_{\sigma_{\text{LoS}}}$	5.3	H	10	$P_{\max}$	20 W
N_{tr}	1000	d_0	5 m	σ^2	-95 dBm
N_s	12	λ	0.01	$h_{\min}$	100 m
N_x	4	β	2	B	1 MHz
μ_{LoS}	2	μ_{NLoS}	2.4	$\delta_{S_i,n}$	5 Mbit/s
χ	15	ζ_1	0.5	f_c	38 GHz
$\chi_{\sigma_{\text{NLoS}}}$	5.27	η	100	B_v	1 GHz
K	5	C	1	L	1 Mbit
T	120	ζ_2	0.5	N_w	1000

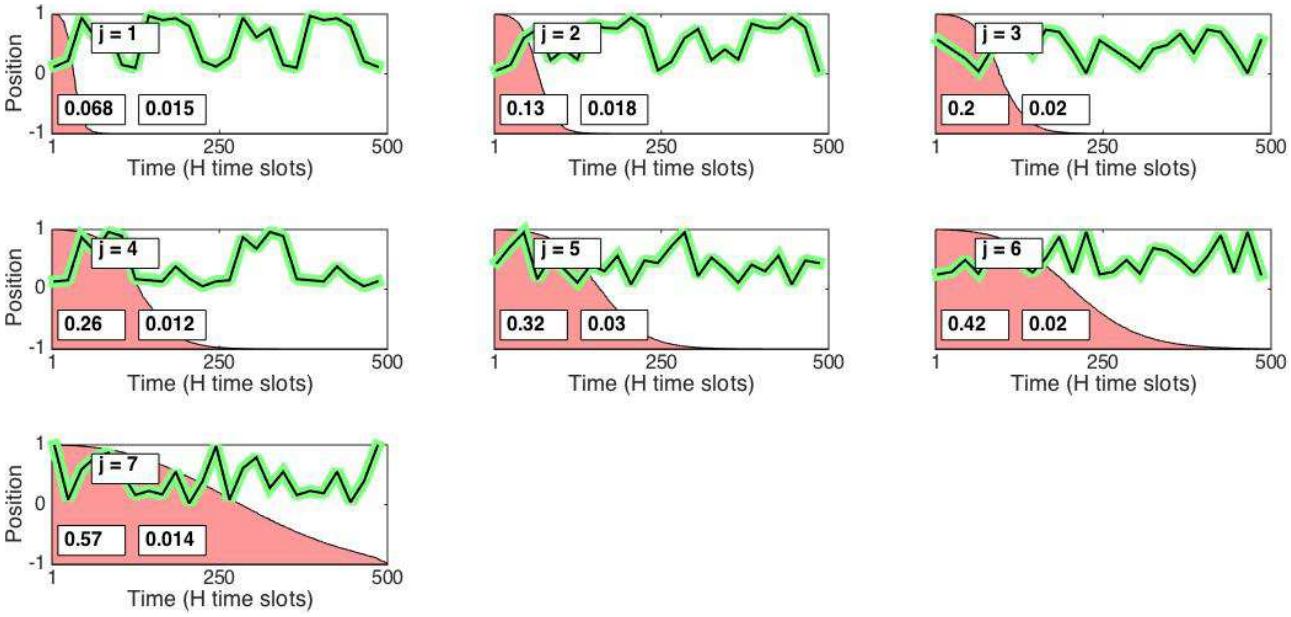


Figure 6.3: Mobility patterns predictions of Conceptor ESN algorithm.

represents the conceptor ESN prediction, the black curve is the real positions, top rectangle j is the index of the mobility pattern learned by ESN, the legend on the bottom left shows the total reservoir memory used by ESN and the legend on the bottom right shows the normalized root mean square error of each mobility pattern prediction. Here, one mobility pattern represents the users' trajectory in one day and the colored region represents the memory used by the conceptor ESN. In Fig. 6.3, we can see that the memory usage increases as the number of the learned mobility patterns increases. This is due to the fact the conceptor ESN uses a limited memory to learn mobility patterns. From Fig. 6.3, we can also see that the conceptor ESN uses less memory for learning mobility pattern 2 compared to pattern 6. In fact, mobility pattern 2 is similar to

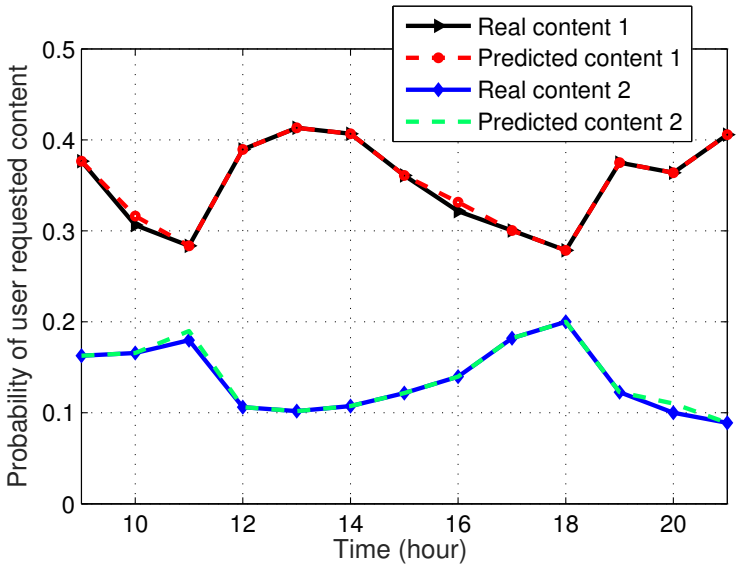


Figure 6.4: Content request probability predictions.

mobility pattern 1, and, hence, the conceptor ESN requires only a small amount of memory to learn mobility pattern 2. However, the conceptor ESN needs to use more memory to learn mobility pattern 6. Clearly, when a new mobility pattern needs to be learned, the proposed approach needs to learn the difference between the learned mobility patterns and the new one.

In Fig. 6.4, we show the variations of two content request probabilities of a selected user during one day. The user is randomly chosen from the set of users in the network. From Fig. 6.4, we can see that, the probability with which this user requests content 1 decreases during working hours (9:00-11:00 and 14:00-18:00) and increases at all other times. Similarly, the request probability of content 2 increases during working hours and decreases during the rest of the time. This is due to the fact that content 1 is an entertainment content while content 2 is a work-related content. Fig. 6.4 also shows that the sum of the probability with which this user requests content 1 and content 2 exceeds 0.5 during each hour. This is because the user always requests a small amount of contents during one day.

Fig. 6.5 shows how the total transmit power of the UAVs in a time period changes as the number of the users varies. In Fig. 6.5, we can see that the total UAV transmit power of all algorithms increases as the number of the users increases. This is due to the fact that the number of the users associated with the RRHs and the capacity of the wireless fronthaul link of UAVs are limited. Therefore, the UAVs need to increase their transmit power to satisfy the QoE requirement of each user. From Fig. 6.5a, we can also see that the proposed approach can reduce the total transmit power of the UAVs of about 16.7% compared to the ESN algorithm used to predict the content request and mobility for a network with 70 users. This is because the conceptor ESN that separates the users' behavior into multiple patterns and uses the conceptor to learn these patterns, can predict the users' behavior more accurately compared to the ESN algorithm. Fig. 6.5b shows that the proposed algorithm can yield, respectively, 33.3% and 20% gains with respect to reducing the total

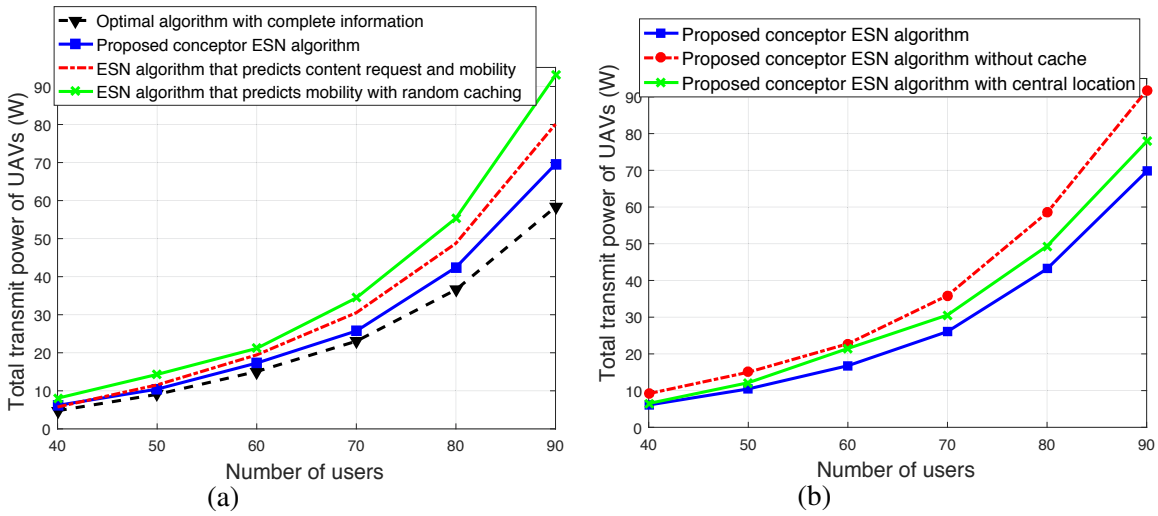


Figure 6.5: Total transmit power as the number of users varies ($K = 5$ and $C = 1$.)

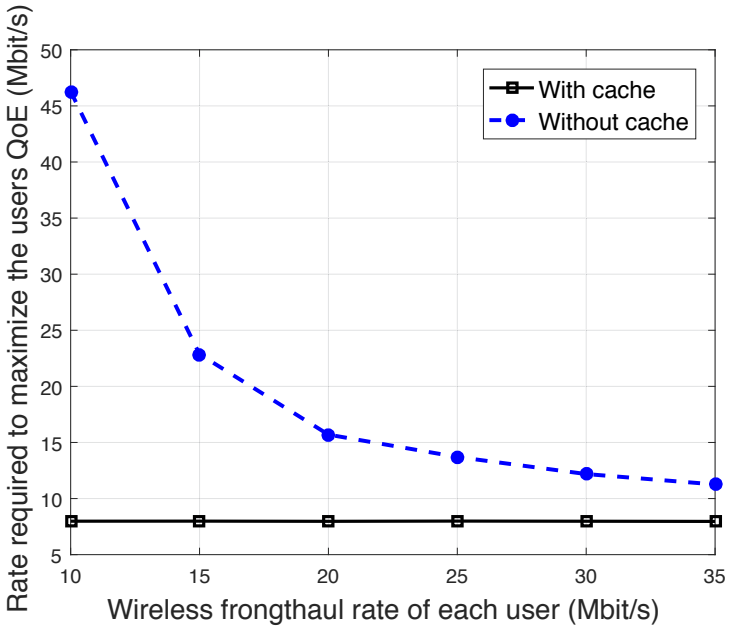


Figure 6.6: Rate required to maximize the users QoE as the fronnthaul rate of each user changes.

transmit power compared to the proposed algorithm without cache and the proposed algorithm without optimizing the UAVs' locations for a network with 80 users.

Fig. 6.6 shows the rate needed for satisfying the QoE requirement of each user versus the wireless fronthaul rate of each user. In this figure, the black and blue lines represent, respectively, the rate requirement of a user that receives a content from the UAV cache and the BBUs. In Fig. 6.6, we can see that the rate required to maximize the users QoE of BBUs-UAV-user link decreases as the

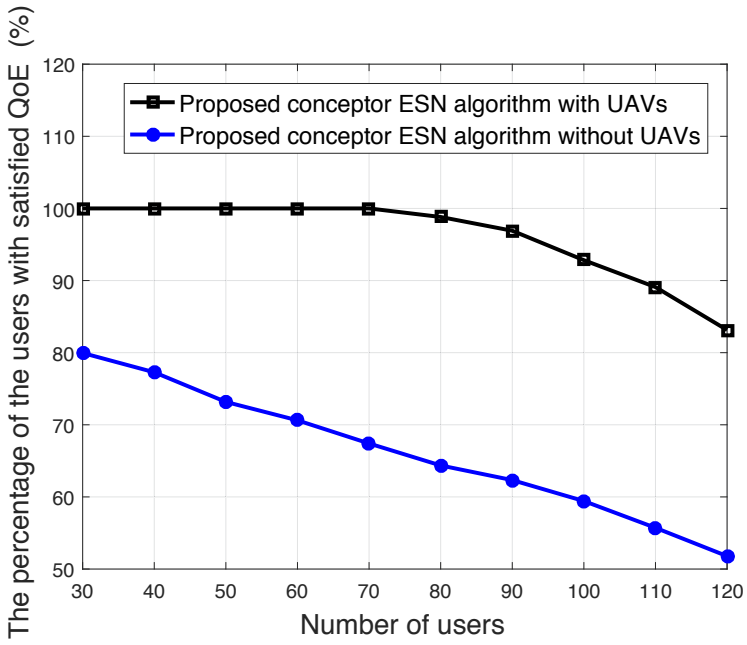


Figure 6.7: The percentage of the users QoE that is maximized as the number of users varies

wireless fronthaul rate increases. However, the rate needed to maximize the user's QoE of UAV-user link not change when the fronthaul rate varies. Clearly, the use of caching at the UAVs can significantly reduce the rate required to reach the QoE threshold of each user when the wireless fronthaul rate for each user is low.

In Fig. 6.7, we show how the percentage of users with satisfied QoE requirement changes as the number of the users varies. From Fig. 6.7, we can see that the percentage of the satisfied users decreases as the number of the users increases. However, using the proposed approach, the QoE remains maximum for all number of users when the number of the users increases from 30 to 70. In particular, the proposed algorithm can yield a gain of 59.6% gain in terms of the percentage of the users with satisfied QoE compared to the proposed algorithm without UAVs for the network with 120 users. This is due to the fact that the UAVs can maximize the users' QoE when the RRHs are not able to satisfy the QoE requirements.

In Fig. 6.8, we show how the average minimum transmit power of UAVs changes as the number of the UAVs varies. From Fig. 6.8, we can see that the average minimum transmit power of each UAV decreases as the number of the UAVs increases. In particular, using the proposed algorithm, the average transmit power of the UAVs decreases by 86% when the number of UAVs increases from 3 to 7. This is due to the fact that for a higher number of UAVs the number of users associated with each UAV decreases, and, hence, the average transmit power per UAV also decreases. As shown in Fig. 6.8, the proposed approach becomes closer to the optimal one as the number of UAVs increases. The reason is that the location prediction error is higher for a lower number of UAVs (or equivalently the clusters).

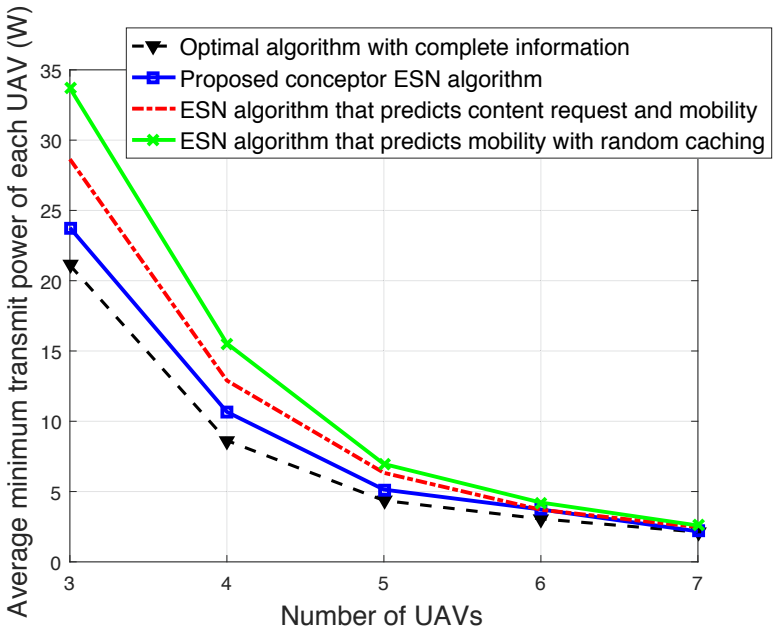


Figure 6.8: Average minimum transmit power as the number of UAVs changes ($U = 70$ and $C = 1$).

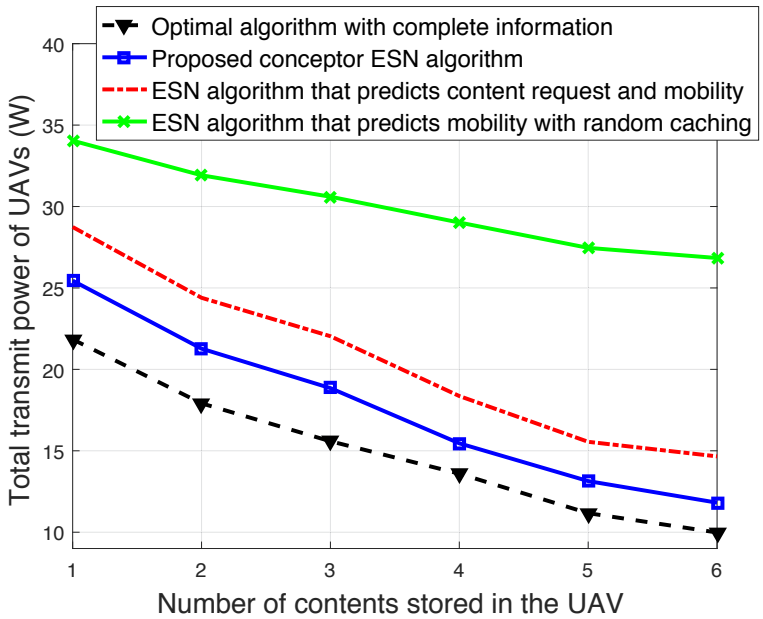


Figure 6.9: Total transmit power as the number of the contents stored in a UAV cache varies ($U = 70$ and $K = 5$).

Fig. 6.9 shows the total transmit power of the UAVs as a function of the number of the contents

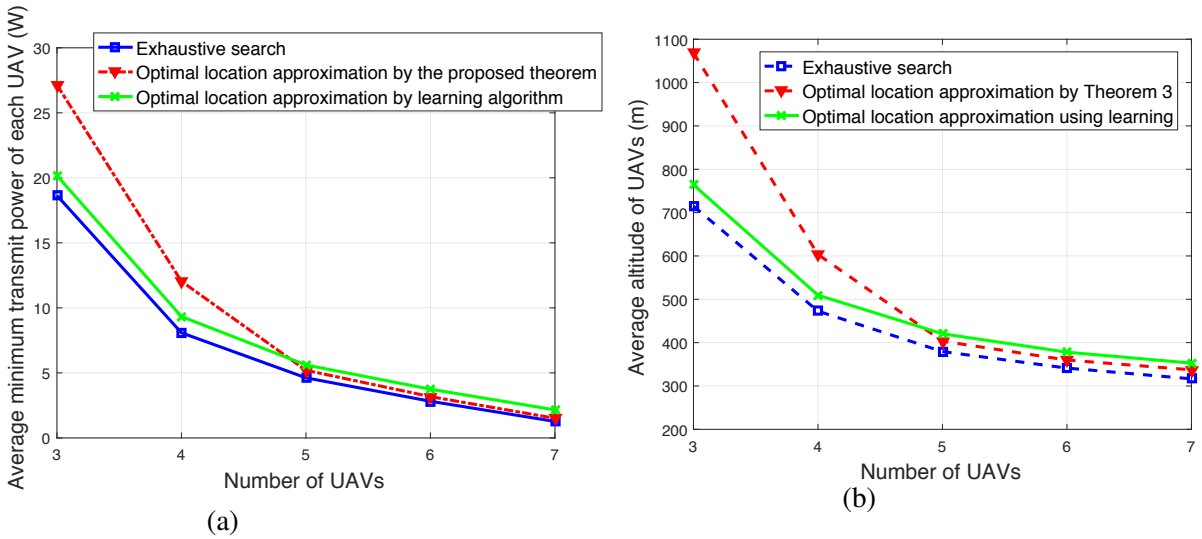


Figure 6.10: Average minimum transmit power and average altitude vs. the number of UAVs.)

stored at the UAV cache. As shown in Fig. 6.9, the total transmit powers of all considered algorithms increase as the number of storage units increases. The reason is that the probability that the requested contents of the users are stored at the UAV cache increases, and, consequently, the UAV will directly transmit the requested contents to the users. Fig. 6.9 also shows that the ESN approach that predicts the content request and mobility can yield up to 49% power reduction compared to the ESN approach that predicts the mobility with the random caching scheme.

In Fig. 6.10, we show how the average transmit power and average altitude of the UAVs change as the number of UAVs varies. In this case, we compare the result of our proposed approach with the optimal result obtained by an exhaustive search method. For the learning algorithm, the interval of the neighboring action of each coordinate is 3 m. Figs. 6.10a shows that the optimal location of the UAV approached by Theorem 12 has only 6.4% deviation compared to the exhaustive search. Furthermore, as shown in Fig. 6.10b, by increasing the number of UAVs from 3 to 7, the average altitude of the UAVs decreases from 1080 m to 332 m in the proposed algorithm case. This is due to the fact that for a higher number of the UAVs, each UAV needs to provide coverage for a smaller area and, hence, it can be deployed at a lower altitude. From Figs. 6.10a and 6.10b, we can also see that, as the number of the UAVs increases, the result of Theorem 12 approaches the optimal solution that is obtained by the exhaustive search. This is due to the fact that for a higher number of UAVs, the coverage area of each UAV decreases, and, hence, the approximation condition in Theorem 12 will hold with a tighter bound.

6.6 Summary

In this chapter, we have proposed a novel framework that uses flying UAVs to provide service for the mobile users in a CRAN system. First, we have presented an optimization problem that seeks to guarantee the QoE rquirement of each user using the minimmum transmit power of the UAVs.

Next, to solve this problem, we have developed a novel algorithm based on the echo state networks and concepters. The proposed algorithm allows predicting the content request distribution of each user with limited information on the network state and user context. The proposed algorithm also enables the ESNs separate the users behavior into several patterns and learn these patterns with various non-linear systems. Simulation results have shown that the proposed approach yields significant performance gains in terms of minimum transmit power compared to conventional ESN approaches.

6.7 Appendix C

C.1 Proof of Proposition 7

From (6.10), we can see that the delay of link (b) is larger than that of link (c) and the minimum delay of the link (a) is L/v_F. Hence, we only need to consider the delay values between L/v_{FU} and $L/C^{\mathrm{V}}_{\tau,ki}$, $k \in \mathcal{K}$. To maximize $C^{\mathrm{V}}_{\tau,ki}$, we consider $d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right) = h$, and $P_{t,ki} = P_{\max}$. Then, the rate of the UAV-user link $C^{\mathrm{V}}_{\tau,ki}$ is given by:

$$\begin{aligned} C^{\mathrm{V}}_{\tau,ki} &= B_V \log_2\left(1 + \frac{P_{t,ki}}{10^{\bar{l}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)/10}\sigma^2}\right) \\ &\leq B_V \log_2\left(1 + \frac{P_{\max}}{10^{l^{\mathrm{LoS}}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)/10}\sigma^2}\right) \\ &\overset{(a)}{\leq} B_V \log_2\left(1 + \frac{P_{\max}}{10^{\left(L_{FS}(d_0) + 10\mu_{\mathrm{LoS}}\log(h) - 4\sigma_{\mathrm{LoS}}\right)/10}\sigma^2}\right), \end{aligned} \tag{6.27}$$

where (a) follows from the fact that with a probability close to one (greater than 99.99%), the Gaussian random variable $\chi_{\sigma_{\mathrm{LoS}}}$ will have a value larger than $-4\sigma_{\mathrm{LoS}}$. From (6.27), we can see that, as h increases, the capacity $C^{\mathrm{V}}_{\tau,ki}$ decreases. Therefore, we set $h = h_{\min}$. This completes the proof.

6.7.1 Proof of Theorem 10

Based on (6.12) and $\bar{D}_{\min}$, the delay is $D_{\tau,i,n} = \Delta_\tau - \bar{D}_{\min}\left(\Delta_\tau - \min\left\{\frac{L}{v_F}, \frac{L}{C^{\max}_k}\right\}\right)$, and, hence, the delay requirement for RRH cluster q transmitting content n to user i during time slot τ will be:

$$C^R_{\tau,qi} = \frac{L}{\Delta_\tau - \bar{D}_{\min}\left(\Delta_\tau - \min\left\{\frac{L}{v_F}, \frac{L}{C^{\max}_k}\right\}\right) - \frac{L}{v_{FU}}}. \tag{6.28}$$

Therefore, the delay requirement during each interval is equal to $C^R_{\tau,qi}$. Since the device rate requirement is $\delta_{S_i,n}$, the rate of RRH cluster q transmitting content n to user i, $C^{\mathrm{H}}_{\tau,qi}$ must satisfy $C^{\mathrm{H}}_{t,qi} \geq \max\left\{C^R_{\tau,qi}, \delta_{S_i,n}\right\}$. This completes the proof.

C.2 Proof of Theorem 11

Since the delay requirement $C^R_{\tau,ki,n}$ depends on the contents at the UAV cache, it can be written as $\delta^R_{i,n} = \max\left\{C^R_{\tau,ki,n(n\in\mathcal{C}_k)}, C^R_{\tau,ki,n(n\notin\mathcal{C}_k)}, \delta_{S_i,n}\right\}$. Let $P^{\min}_{\tau,ki} = \sum_{t=1}^{F_{\tau,i}} P^{\min}_{j,\tau,t,ki}$. Then the reduction of UAV transmit power by content caching during time slot τ of period j will be:

$$\Delta P_{j,\tau,ki,n}= \begin{cases} P^{\min}_{\tau,ki}\left(C^R_{\tau,ki}\right)_{n\notin\mathcal{C}_k} - P^{\min}_{\tau,ki}\left(C^R_{\tau,ki}\right)_{n\in\mathcal{C}_k}, C^R_{\tau,ki,n\notin\mathcal{C}_k} \geq \delta_{S_i,n}, \\ P^{\min}_{\tau,ki}(\delta_{S_i,n})_{n\notin\mathcal{C}_k} - P^{\min}_{\tau,ki}\left(C^R_{\tau,ki}\right)_{n\in\mathcal{C}_k}, \quad \delta_{S_i,n} > C^R_{\tau,ki,n\notin\mathcal{C}_k}. \end{cases}$$

Considering the fact that the content request distribution changes once every H time slots, the power minimization problem for UAV k during a period that consists of H time slots is:

$$\begin{aligned} &\min_{\mathcal{C}_k} \sum_{\tau=1}^{T} \sum_{i\in\mathcal{U}_{\tau,k}} P^{\min}_{\tau,ki} \min_{\mathcal{C}_k} = \sum_{j=1}^{T/H} \sum_{\tau_j=1}^{H} \sum_{i\in\mathcal{U}_{\tau,k}} P^{\min}_{\tau_j,ki} \\ &= \min_{\mathcal{C}_k} \sum_{j=1}^{T/H} \sum_{\tau=1}^{H} \sum_{i\in\mathcal{U}_{\tau,k}} P^{\min}_{j,\tau,ki} \overset{(a)}{\Leftrightarrow} \max_{\mathcal{C}_k} \sum_{j=1}^{T/H} \sum_{\tau=1}^{H} \sum_{i\in\mathcal{U}_{\tau,k}} \Delta P_{j,\tau,ki,n}, \\ &\overset{(b)}{=} \max_{\mathcal{C}_k} \sum_{j=1}^{T/H} \sum_{\tau=1}^{H} \sum_{i\in\mathcal{U}_{\tau,k}} \left(\sum_{n\in\mathcal{C}_k} (p_{j,in}\Delta P_{j,\tau,ki,n}) + \sum_{n\notin\mathcal{C}_k} (p_{j,in}\Delta P_{j,\tau,ki,n}) \right), \\ &= \max_{\mathcal{C}_k} \sum_{j=1}^{T/H} \sum_{\tau=1}^{H} \sum_{i\in\mathcal{U}_{\tau,k}} \sum_{n\in\mathcal{C}_k} (p_{j,in}\Delta P_{j,\tau,ki,n}), \end{aligned}$$

where (a) follows the fact that minimizing the transmit power of the UAVs is equivalent to maximizing the reduction of the UAVs' transmit power caused by caching, and (b) is obtained by computing the average power reduction using content request probability distribution of each user. This completes the proof.

C.3 Proof of Theorem 12

At very low altitudes, $h^2_{\tau,k} \ll (x_{t,i} - x_{\tau,k})^2 + (y_{t,i} - y_{\tau,k})^2$, $\frac{h_{\tau,k}}{d_{t,ki}(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i})} \approx 0$ leading to $\phi_{\bar{t}} = 0^\circ$, and, hence, $\Pr\left(l^{\text{NLoS}}_{t,ki}\right) = 1$. Thus, we have $\bar{l}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right) = l^{\text{NLoS}}_{t,ki}$ and (6.16) can be rewritten as $P^{\min}_{\tau,t,ki} = \left(2^{\delta^R_{i,n}/B} - 1\right)\sigma^2 10^{\left(L_{FS}(d_0) + \chi_{\sigma_{\text{NLoS}}}\right)/10} d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k}, \boldsymbol{w}_{\tau,t,i}\right)^{\mu_{\text{NLoS}}}$.

Now, we find the optimal location $(x_{\tau,k}, y_{\tau,k})$ of UAV k during time slot τ in order to minimize

$\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}P^{\min}_{\tau,t,ki}$. In this case, the derivation of $\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}P^{\min}_{\tau,t,ki}$ with respect to $x_{\tau,k}$ is given by:

$$\frac{\partial\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}P^{\min}_{\tau,t,ki}}{\partial x_{\tau,k}}=\frac{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\partial P^{\min}_{\tau,t,ki}}{\partial x_{\tau,k}}=$$
$$\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\mu_{\text{NLoS}}(x_{\tau,k}-x_{t,i})\psi_{t,ki}\Big((x_{\tau,k}-x_{t,i})^2+(y_{\tau,k}-y_{t,i})^2+h^2_{\tau,k}\Big)^{\frac{\mu_{\text{NLoS}}}{2}-1}. \tag{6.29}$$

As $\mu_{\text{NLoS}}=2$, (6.29) is simplified to $\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}2(x_{\tau,k}-x_{t,i})\,\psi_{t,ki}=0$. As a result, $x_{\tau,k}=\frac{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}x_{t,i}\psi_{t,ki}}{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\psi_{t,ki}}$.

Likewise, we can show that $y_{\tau,k}=\frac{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}y_{t,i}\psi_{t,ki}}{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\psi_{t,ki}}$.

For case b), since $h^2_{\tau,k}\gg(x_{t,i}-x_{\tau,k})^2+(y_{t,i}-y_{\tau,k})^2$, $d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k},\boldsymbol{w}_{\tau,t,i}\right)\approx h_{\tau,k}$ and, hence, $\frac{h_{\tau,k}}{d_{t,ki}\left(\boldsymbol{w}_{\tau,t,k},\boldsymbol{w}_{\tau,t,i}\right)}\approx 1\rightarrow\phi_t=90^\circ$. Consequently, $\Pr\left(l^{\text{LoS}}_{t,ki}\right)=1$. Then, we have $\bar{l}_{t,ki}\left(\boldsymbol{w}_{\tau,t,k},\boldsymbol{w}_{\tau,t,i}\right)=l^{\text{LoS}}_{t,ki}$. The derivation of $\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}P^{\min}_{\tau,t,ki}$ will be:

$$\frac{\partial\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}P^{\min}_{\tau,t,ki}}{\partial x_{\tau,k}}=\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\mu_{\text{LoS}}\left(x_{\tau,k}-x_{t,i}\right)\psi_{t,ki}$$
$$\times\Big((x_{\tau,k}-x_{t,i})^2+(y_{\tau,k}-y_{t,i})^2+h^2_{\tau,k}\Big)^{\frac{\mu_{\text{LoS}}}{2}-1}$$
$$\approx\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\mu_{\text{LoS}}\left(x_{\tau,k}-x_{t,i}\right)\psi_{t,ki}h^{\mu_{\text{LoS}}-2}_{\tau,k}=0.$$

As a result, $x_{\tau,k}=\frac{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}x_{t,i}\psi_{t,ki}}{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\psi_{t,ki}}$ and $y_{\tau,k}=\frac{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}y_{t,i}\psi_{t,ki}}{\sum_{i\in\mathcal{U}_{\tau,k}}\sum_{t=1}^{F_{\tau,i}}\psi_{t,ki}}$. This completes the proof.

Chapter 7

Communications and Control for Wireless Drone-Based Antenna Array

7.1 Background, Related Works, and Contributions

The use of drones is growing rapidly across many domains including delivery, communications, surveillance, and search and rescue in emergency operations [7, 61, 80, 183–185]. In wireless networks, drones can be used as flying base stations to provide reliable and cost-effective wireless connectivity [7, 12, 15, 80, 124, 183, 184, 186, 187]. Due to their flexibility, agility, and mobility, drones can support reliable, cost-effective, and high data rate wireless communications for ground users. In particular, during major public events such as Olympic games that generate a substantial demand for communication, there is a need to supplement the limited capacity and coverage capabilities of existing cellular networking infrastructure. In such scenarios, drone-based wireless communication is an ideal solution. For instance, AT&T and Verizon are planning to use flying drones to boost the Internet coverage for the college football national championship and the Super Bowl [34]. Drones can also play a key role in enabling wireless connectivity in other key scenarios such as public safety, and Internet of Things (IoT) scenarios [7]. To effectively leverage drones for wireless networking applications, one must address a number of challenges that include optimal placement of drones, path planning, resource management, control, and flight time optimization [7, 12, 80].

There has been a recent surge of literature discussing the use of drones for wireless communication purposes [7, 12, 15, 27, 80, 124, 183, 184, 186, 188–190]. For instance, in [183], the authors studied the optimal 3D placement of UAVs for maximizing the number of covered users with different quality-of-service (QoS) requirements. The works in [80] and [124] studied path planning and optimal deployment problems for UAV-based communications and computing. The work in [186] proposed a framework for the optimal placement and distribution of UAVs to minimize the overall delay in a UAV-assisted wireless network. A comparison between the performance of aerial base stations and terrestrial base stations in terms of average sum rate and transmit power is presented

in [27]. In [188], a polynomial-time algorithm for the optimal placement of drones that provide coverage for ground terminals is proposed.

One of the fundamental challenges in drone-based communications systems is the limited flight endurance of drones. Naturally, flying drones have a limited amount of on-board energy which must be used for transmission, mobility, control, data processing, and payloads purposes. Consequently, the flight duration of drones is typically short and can be insufficient for providing a long-term, continuous wireless coverage. Furthermore, due to the limited transmit power of drones, providing long-range, high rate, and low latency communications can be challenging in drone-enabled wireless systems. In this regard, a key performance metric in drone-enabled wireless networks is *airborne service time*, which is defined as the time needed for servicing ground users. The service time directly impacts the flight time of drones as well as the quality-of-service (i.e., delay) for ground users. From the drones' perspective, a lower service time corresponds to a shorter flight time as well as less energy consumption. From the users' point of view, a lower service time is also needed as it directly yields lower latency. To address the flight time challenges of drones, the work in [189] minimized the hover time of drone base stations by deriving the optimal cell association schemes. However, the model in [189] is limited to static single-antenna drones. In [190], the trajectory and mission completion time of a single UAV that serves ground users are optimized. However, the work in [190] does not analyze a scenario with multiple UAVs.

One promising approach to provide high data rate and low service time is to utilize multiple drones within an antenna array system composed of multiple single-antenna drones [97]. Compared to conventional antenna array systems, a drone-based antenna array has the following advantages. First, the number of antenna elements (i.e., drones) is not limited by space constraints. Second, the gain of the drone-based antenna array can be increased by adjusting the array element spacing. Third, the mobility and flexibility of drones enable an effective mechanical beam-steering in any three-dimensional (3D) direction. Clearly, a high gain drone-based antenna array can provide high data rate wireless services to ground users thus reducing the service time.

In [97], the authors studied the design of a UAV-based antenna array for directivity maximization. However, the approach presented in [97] is based on a heuristic and a computationally demanding evolutionary algorithm. Moreover, the service time analysis is ignored in [97]. In [191], the authors derived the asymptotic capacity of an airborne multiple-input-multiple-output (MIMO) wireless communication system. However, the work in [191] considers fixed positions for the antenna elements of the transmitter and the receiver. Furthermore, this work does not analyze the control aspect of drones which is essential in designing drone-based MIMO systems.

In fact, none of the previous works on drone communications, such as in [7, 12, 15, 27, 61, 80, 97, 124, 183, 184, 186, 188–191], has studied the use of a drone-based antenna array system for service time minimization in wireless networks.

We note that, there exist some studies on time-optimal motion planning [192–195]. However, most of the previous works do not address the time-optimal control problem of quadrotor drones. While the authors in [195] consider a quadrotor drone in their model, they ignore the effect of external forces on the control time. Furthermore, the approach in [195] is based on a genetic algorithm

which is computationally demanding [196]. Unlike our work, the work in [195] ignores the communication aspects of drones, and does not capture the impact of control time on the performance of drone-enabled wireless networks. In fact, compared to [195], our proposed framework comprises both communication and control aspects of drones and it is analytically tractable. Next, we present the dynamics of quadrotor drones which will be used to analyze our time-optimal control problem.

The main contribution of this chapter is a novel framework for deploying and operating a drone-based antenna array system that delivers wireless service to a number of ground users within a minimum time. In particular, we minimize the service time that includes both the transmission time and the control time needed to control the movement and orientation of the drones. To this end, we minimize the transmission time, by optimizing the drones' locations, as well as the control time that the drones need to move between these optimal locations. To minimize the transmission time, first, we determine the optimal drone spacing for which the array directivity is maximized. In this case, using perturbation theory [98], we solve the drone spacing optimization problem by successively solving a number of perturbed convex optimization problems. Next, given the derived drone spacing, we optimally adjust the locations of the drones according to the position of each ground user. In order to serve different users, the drones must dynamically move between the derived optimal locations, during the control time period. To minimize the control time of quadrotor drones, we determine the optimal speeds of rotors such that the drones can update their positions and orientations within a minimum time. In this case, using *bang-bang* control theory [99], we derive a closed-form expression for the minimum control time as a function of external forces (e.g., wind and gravity), the drone's weight, and the destinations of drones. Our results show that the proposed drone antenna array approach can significantly reduce the service time and improve the spectral and energy efficiency of the network. In particular, our approach yields 48% improvement in spectral efficiency compared to a case in which the same number of drones are deployed separately. The results also reveal a tradeoff between the control time and transmission time while varying the number of drones.

7.2 System Model and General Problem Formulation

Consider a set $\mathcal{L}$ of L single-antenna wireless users located within a given geographical area. In this area, a set $\mathcal{M}$ of M quadrotor drones are used as flying access points to provide downlink wireless service for ground users. The M drones will form an antenna array in which each element is a single-antenna drone, as shown in Fig. 7.1. For tractability, we consider a linear antenna array whose elements are symmetrically excited and located about the origin of the array as done in [197]. The results that we will derive for the linear array case can provide a key guideline for designing more complex 2D and 3D array configurations. The 3D location of drone $m \in \mathcal{M}$ and of user $i \in \mathcal{L}$ is given by $(x_i^{\mathrm{u}}, y_i^{\mathrm{u}}, z_i^{\mathrm{u}})$, and the location of drone m while serving user i is $(x_{m,i}, y_{m,i}, z_{m,i})$. To avoid collisions[1], we assume that adjacent drones in the array are separated

[1] Virtual force field approaches such as in [198] can also be used to avoid collisions.

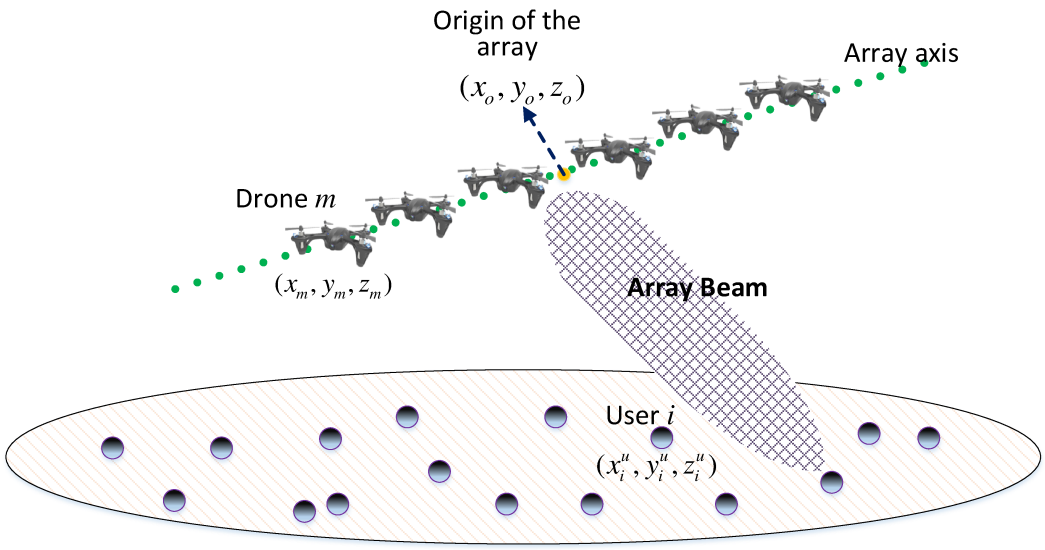


Figure 7.1: Drone-based antenna array.

by at least $D_{\min}$. Let a_m and β_m be the amplitude and phase of the signal (i.e. excitation) at element m in the array. Let $d_m = \sqrt{(x_{m,i} - x_o)^2 + (y_{m,i} - y_o)^2 + (z_{m,i} - z_o)^2}$ be the distance of drone m from the origin of the array whose 3D coordinate is (x_o, y_o, z_o). The magnitude of the far-field radiation pattern of each element is $w(\theta, \phi)$, where θ and ϕ are the polar and azimuthal angles in the spherical coordinate.

To serve ground users distributed over a geographical area, the drones will dynamically change their positions based on each user's location. Such repositioning is needed for adjusting the distance and beam direction of the antenna array to each ground user. Note that, unlike a classical linear phased array that uses electronic beam steering, the proposed drone-based antenna array relies on the repositioning of drones. This is due to the fact that, in the drone antenna array, precisely adjusting the elements' phase is more challenging than the phased array whose elements are directly connected. In addition, a linear phased array cannot perform 3D beam steering. Hence, in our model, the drones dynamically adjust their positions in order to steer the beam towards ground users. Clearly, the *service time*, which is the time needed to serve the ground users, depends on the transmission time and the control time during which the drones must move and stabilize their locations. The transmission time is inversely proportional to the downlink data rate which depends on the signal-to-noise-ratio (SNR) which is, in turn, function of the array's beamforming gain.

The service time is an important metric for both users and drones. A lower service time yields a lower delay and, hence, higher quality-of-service for the users. Also, the service time is directly related to spectral efficiency as it depends on data rate and transmission bandwidth. For drones, a lower service time corresponds to a shorter flight time and less energy consumption. In fact, minimizing the service time improves both energy and spectral efficiency. Therefore, our goal is to minimize the total service time of the ground users by optimally adjusting the drones' locations,

within a minimum control time, that can provide a maximum data rate.
For drone-to-ground communications, we consider a line-of-sight (LoS) propagation model as done in [80] and [190]. Such a channel model is reasonable here as the effect of multipath is significantly mitigated due to the high altitude of drones and using beamforming [190]. The transmission rate from the drone antenna array to ground user i in a far-field region is given by [190]:

$$R_i(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i) = B\log_2\left(1 + \frac{r_i^{-\alpha} P_t K_o G_i(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i)}{\sigma^2}\right), \tag{7.1}$$

where $\boldsymbol{x}_i = [x_{m,i}]_{M\times 1}$, $\boldsymbol{y}_i = [y_{m,i}]_{M\times 1}$, $\boldsymbol{z}_i = [z_{m,i}]_{M\times 1}$, $m \in \mathcal{M}$ representing the 3D coordinates of the drones while serving user i. B is the transmission bandwidth, r_i is the distance between the origin of the array and user i, P_t is the total transmit power of the array, σ^2 is the noise power, and K_o is the constant path loss coefficient. $G_i(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i)$ is the gain of the antenna array towards the location of user i, given by [199]:

$$G_i(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i) = \frac{4\pi |F(\theta_i, \phi_i)|^2 w(\theta_i, \phi_i)^2}{\int_0^{2\pi}\int_0^{\pi} |F(\theta, \phi)|^2 w(\theta, \phi)^2 \sin\theta \mathrm{d}\theta \mathrm{d}\phi}\eta, \tag{7.2}$$

where $0 \leq \eta \leq 1$ is the antenna array efficiency which is multiplied by directivity to compute the antenna gain. In fact, the antenna gain is equal to the antenna directivity multiplied by η. In (7.2), $F(\theta, \phi)$ is the array factor which can be written as [199]:

$$F(\theta, \phi) = \sum_{m=1}^{M} a_m e^{j[k(x_{m,i}\sin\theta\cos\phi + y_{m,i}\sin\theta\sin\phi + z_{m,i}\cos\theta) + \beta_m]}, \tag{7.3}$$

where $k = 2\pi/\lambda$ is the phase constant, and λ is the wavelength. Note that, the overall radiation pattern of the antenna array is equal to $F(\theta, \phi)w(\theta_i, \phi_i)$ which follows from the pattern multiplication rule [199].

Now, the total time that the drones spend to service the ground users will be:

$$T_{\text{tot}} = \sum_{i=1}^{L} \frac{q_i}{R_i(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i)} + T_i^{\text{crl}}(\boldsymbol{V}, \boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i), \tag{7.4}$$

where T_{tot} shows the total service time, q_i is the load of user i which represents the number of bits that must be transmitted to user i. T_i^{crl} is the control time during which the drones adjust their locations according to the location of ground user i. In particular, T_i^{crl} captures the time needed for updating the drones' locations from state $i-1$ (i.e., locations of drones while serving user $i-1$, $i > 1$) to state i. The control time is obtained based on the dynamics of the drones and is a function of control inputs, external forces, and the movement of drones. In fact, each drone needs a vector of control inputs in order to move from its initial location to a new location while serving different users. For quadrotor drones, the rotors' speeds are commonly considered as control inputs. Therefore, in (7.4), we have $\boldsymbol{V} = [v_{mn}(t)]_{M\times 4}$ with $v_{mn}(t)$ being the speed of

rotor n of drone m at time t. The maximum speed of each rotor is $v_{\max}$. In this case, one can minimize the control time of the drones by properly adjusting the rotors' speeds. In Section 7.4, we will provide a detailed analysis of the control time given the drones' dynamics.

Given this model, our goal is to minimize the total service time of drones by finding the optimal locations of the drones with respect to the center of the array, as well as the optimal control inputs. Our optimization problem, in its general form, is given by:

$$\underset{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{V}}{\text{minimize}} \sum_{i=1}^{L} \frac{q_i}{R_i(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i)} + T_i^{\text{crl}}(\boldsymbol{V}, \boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i), \tag{7.5}$$

$$\text{st. } d_{m+1} - d_m \geq D_{\min}, \ \forall m \in \mathcal{M}\backslash\{M\}, \tag{7.6}$$

$$0 \leq v_{mn}(t) \leq v_{\max}, \ \forall m \in \mathcal{M}, n \in \{1, ..., 4\}, \tag{7.7}$$

where $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ are matrices whose rows i are, respectively, vectors $\boldsymbol{x}_i$, $\boldsymbol{y}_i$, and $\boldsymbol{z}_i$, $\forall i \in \mathcal{L}$. The constraint in (7.6) indicates that the minimum separation distance between two adjacent drones must be greater than $D_{\min}$ to avoid collision. (7.7) represents the constraints on the speed of each rotor. Note that, the first term in (7.5) represents the transmission time which depends on the drones' locations. The second term, T_i^{crl}, is the control time which is a function of the rotors' speeds as well as the drones' locations. Solving (7.5) is challenging as it is highly nonlinear due to (7.2). Moreover, as we can see from (7.3), the array factor is a complex function of the array element's positions. In addition, due to the nonlinear nature of quadrotor's dynamic system, finding the optimal control inputs is a challenging task, as will be discussed in Section 7.4.

We note that, considering a narrow-beam antenna array communication, (7.5) can be solved by separately optimizing drones' locations and rotors' speeds. In the narrow-beam case, the drone array must perfectly steer its beam towards each ground user. Hence, we can first determine the optimal drones' positions and, then, optimize the rotors' speeds to move to these optimal positions within a minimum time. Our approach for solving (7.5) includes two key steps. First, given the location of any ground user, we optimize the locations of the drones in the linear array to minimize the transmission time. Thus, given L ground users, we will have L sets of drones' locations. In the second step, using the result of the first step, we determine the drones' optimal control strategy to update their locations within a minimum time. Hence, the solution of the transmission time optimization problem (in the first step) is used as inputs to the time-optimal control problem (in the second step). While, in general, this approach leads to a suboptimal solution, it is analytically tractable and practically easy to implement. Next, we will optimize the location of drones to achieve a minimum transmission time for any arbitrary ground user.

7.3 Optimal Positions of Drones in Array for Transmission Time Minimization

In this section, we determine the optimal positions of the drones in the array based on the location of each user such that the transmission time to the user is minimized. Clearly, given (7.1), (7.2), and (7.4), to minimize the transmission time, we need to maximize the array gain (i.e., directivity) towards each ground user.

Without loss of generality, we consider an even number of drones. For an odd number of drones, the same analysis will still hold. Now, the array factor for M drones located on the x-axis of the Cartesian coordinate can be given by:

$$\begin{aligned} F(\theta,\phi) &= \sum_{m=1}^{M} a_m e^{j[kx_m \sin\theta\cos\phi+\beta_m]} \\ &\overset{(a)}{=} \sum_{n=1}^{M/2} a_n \left(e^{j[kd_n \sin\theta\cos\phi+\beta_n]} + e^{-j[kd_n \sin\theta\cos\phi+\beta_n]}\right) \\ &\overset{(b)}{=} 2\sum_{n=1}^{N} a_n \cos\left(kd_n \sin\theta\cos\phi+\beta_n\right), \end{aligned} \tag{7.8}$$

where $N = M/2$, and d_n is the distance of element $n \in \mathcal{N} = \{1, 2, ..., N\}$ from the center of the array (origin). Also, (a) follows from the fact that the array is symmetric with respect to the origin, and (b) is based on the Euler's rule.

Now, we can maximize the directivity of the array by optimizing $d_n, \forall n \in \mathcal{N}$:

$$\underset{d_n,\forall n\in\mathcal{N}}{\text{maximize}} \frac{4\pi|F(\theta_{\max},\phi_{\max})|^2 w(\theta_{\max},\phi_{\max})^2}{\int_0^{2\pi}\int_0^{\pi}|F(\theta,\phi)|^2 w(\theta,\phi)^2 \sin\theta \mathrm{d}\theta\mathrm{d}\phi}, \tag{7.9}$$

where $(\theta_{\max}, \phi_{\max})$ are the polar and azimuthal angles at which the total antenna pattern $F(\theta,\phi)w(\theta,\phi)$ has a maximum value. Clearly, solving (7.9) is challenging due to the non-linearity and complex expression of the objective function of this optimization problem. Moreover, this problem is non-convex and, hence, cannot be exactly solved using classical convex optimization methods. Next, we solve (7.9) by exploiting the perturbation technique [197]. In general, perturbation theory aims at finding the solution of a complex problem, by starting from the exact solution of a simplified version of the original problem [98]. This technique is thus useful when dealing with nonlinear and analytically intractable optimization problems such as (7.9).

7.3.1 Perturbation Technique for Drone Spacing Optimization

To optimize the distance between drones, we first consider an initial value for the distance of each drone from the origin. Then, we find the optimal perturbation value that must be added to this

initial value[2]. Let d_n^0 be initial distance of drone n from the origin, then, the perturbed distance will be:

$$d_n = d_n^0 + e_n, \tag{7.10}$$

where $e_n << \lambda$, with λ being the wavelength, is the perturbation value. Given (7.10), the array factor can be approximated by:

$$\begin{aligned} F(\theta,\phi) &= 2\sum_{n=1}^{N} a_n \cos\left(k(d_n^0+e_n)\sin\theta\cos\phi+\beta_n\right) \\ &= 2\sum_{n=1}^{N} a_n \cos\left[\left(kd_n^0\sin\theta\cos\phi+\beta_n\right)+ke_n\sin\theta\cos\phi\right] \\ &\overset{(a)}{\approx} \sum_{n=1}^{N} 2a_n\cos\left(kd_n^0\sin\theta\cos\phi+\beta_n\right) \\ &\quad -\sum_{n=1}^{N} 2a_nke_n\sin\theta\cos\phi\sin\left(kd_n^0\sin\theta\cos\phi+\beta_n\right), \end{aligned} \tag{7.11}$$

where in (a) we used the trigonometric properties, and the fact that $\sin(x)\approx x$ for small values of x. Clearly, given $e_n << \lambda$, the numerator of (7.9) can be computed based on the values of d_n^0, $\forall n\in\mathcal{N}$. Hence, given d_n^0, our optimization problem in (7.9) can be written as:

$$\min_{\boldsymbol{e}} \int_0^{2\pi}\int_0^{\pi} F(\theta,\phi)^2 w(\theta,\phi)^2\sin\theta \mathrm{d}\theta\mathrm{d}\phi, \tag{7.12}$$

$$\text{s.t. } d_{n+1}^0+e_{n+1}-d_n^0-e_n \geq D_{\min},\ \forall n\in\mathcal{N}\setminus\{N\}, \tag{7.13}$$

where $\boldsymbol{e}$ is the perturbation vector having elements e_n, $n\in\mathcal{N}$.

For brevity, we define the following functions:

$$F^0(\theta,\phi) = \sum_{n=1}^{N} a_n\cos\left(kd_n^0\sin\theta\cos\phi+\beta_n\right), \tag{7.14}$$

$$I_{\text{int}}(x) = \int_0^{2\pi}\int_0^{\pi} x\sin\theta\mathrm{d}\theta\mathrm{d}\phi. \tag{7.15}$$

Theorem 13. The optimization problem in (7.12) is convex, and the optimal perturbation vector is the solution of the following system of equations:

$$\begin{cases} \boldsymbol{e} = \boldsymbol{G}^{-1}[\boldsymbol{q}+\boldsymbol{\mu}_{\boldsymbol{\mathcal{L}}}], \\ \mu_n\left(e_n-e_{n+1}+D_{\min}+d_n^0-d_{n+1}^0\right)=0,\ \forall n\in\mathcal{N}\setminus\{N\}, \\ \mu_n\geq 0,\ \forall n\in\mathcal{N}\setminus\{N\}. \end{cases} \tag{7.16}$$

[2]We assume that the distance between adjacent drones is greater than the minimum distance required for collision avoidance.

where $\boldsymbol{G} = [g_{m,n}]_{N\times N}$ is an $N \times N$ matrix with:

$$g_{m,n} = I_{\text{int}}\Bigg(a_m a_n (k \sin\theta \cos\phi w(\theta,\phi))^2 \\ \times \sin\left(kd_n^0 \sin\theta\cos\phi + \beta_n\right) \sin\left(kd_m^0 \sin\theta\cos\phi + \beta_m\right)\Bigg), \tag{7.17}$$

and $\boldsymbol{q} = [q_n]_{N\times 1}$ whose elements are given by:

$$q_n = I_{\text{int}}\Bigg(a_n k \sin\theta\cos\phi w(\theta,\phi) F^0(\theta,\phi) \times \sin\left(kd_n^0 \sin\theta\cos\phi + \beta_n\right)\Bigg). \tag{7.18}$$

In (7.16), $\boldsymbol{\mu_{\mathcal{L}}}$ is a vector of Lagrangian multipliers, whose element n is $\boldsymbol{\mu_{\mathcal{L}}}(n) = \mu_{n+1} - \mu_n$, with μ_n being a Lagrangian multiplier associated with constraint n.

Proof. See Appendix D.1. □

Using Theorem 13, we can update the distance of each drone from the origin as follows:

$$\boldsymbol{d}^1 = \boldsymbol{d}^0 + \boldsymbol{e}^*, \tag{7.19}$$

where $\boldsymbol{d}^1 = [d_n^1]_{N\times 1}$, and $\boldsymbol{d}^0 = [d_n^0]_{N\times 1}$, $n \in \mathcal{N}$.

Clearly, $\boldsymbol{d}^1$ leads to a better solution than $\boldsymbol{d}^0 = [d_n]_{N\times 1}$. In fact, we can proceed and further improve the solution to (7.12) by updating $\boldsymbol{d}^1$. In particular, at step update $r \in \mathbb{N}$, we find $\boldsymbol{d}^{(r)}$ as follows:

$$\boldsymbol{d}^{(r)} = \boldsymbol{d}^{(r-1)} + \boldsymbol{e}^{*(r)}, \tag{7.20}$$

where $\boldsymbol{e}^{*(r)}$ is the optimal perturbation vector at step r which is obtained based on $\boldsymbol{d}^{(r-1)}$.

Note that, at each step, the objective function in (7.12) decreases. Since the objective function is monotonically decreasing and bounded from below, the solution converges after several updates. We note that due to the approximation used in (7.11), the solution may not be a global optimal. Nevertheless, as we can see from Theorem 4, it is analytically tractable and, hence, it has a low computational complexity. Here, we use $\boldsymbol{d}^*$ to represent the vector of nearly-optimal distances of drones from the original of the array. Next, we use $\boldsymbol{d}^*$ to determine the optimal 3D locations of the drones that result in a maximum array directivity towards a given ground user.

7.3.2 Optimal Locations of Drones

Here, following from Subsection 7.3.1, we derive the optimal 3D positions of drones that yields a maximum directivity of the drone-based antenna array. Let $(x_i^{\text{u}}, y_i^{\text{u}}, z_i^{\text{u}})$ and (x_o, y_o, z_o) be, respectively, the 3D locations of user $i \in \mathcal{L}$ and the origin of the antenna array.

Without loss of generality, we translate the origin of our coordinate system to the origin of the antenna array. In other words, we assume that the arrays' center is the origin of our translated

coordinate system. In this case, the 3D location of user i will be $(x_i^{\mathrm{u}} - x_o, y_i^{\mathrm{u}} - y_o, z_i^{\mathrm{u}} - z_o)$. Subsequently, the polar and azimuthal angles of user i in the spherical coordinate (with an origin of antenna array) are given by:

$$\theta_i = \cos^{-1}\left[\frac{z_i^u - z_o}{\sqrt{\left(x_i^u - x_o\right)^2 + \left(y_i^u - y_o\right)^2 + \left(z_i^u - z_o\right)^2}}\right], \tag{7.21}$$

$$\phi_i = \sin^{-1}\left[\frac{y_i^u - y_o}{\sqrt{\left(x_i^u - x_o\right)^2 + \left(y_i^u - y_o\right)^2}}\right]. \tag{7.22}$$

Now, the optimal locations of the drones in the antenna array is given as follows.

Theorem 14. The optimal locations of the drones for maximizing the directivity of the drone-based antenna array towards a given ground user will be:

$$\left(x_m^*, y_m^*, z_m^*\right)^T = \begin{cases} \boldsymbol{R}_{\text{rot}}\left(d_m^* \sin\alpha_o \cos\gamma_o, d_m^* \sin\alpha_o \sin\beta_o, d_m^* \cos\alpha_o\right)^T, & \text{if } m \leq M/2, \\ -\boldsymbol{R}_{\text{rot}}\left(d_m^* \sin\alpha_o \cos\gamma_o, d_m^* \sin\alpha_o \sin\gamma_o, d_m^* \cos\alpha_o\right)^T, & \text{if } m > M/2, \end{cases} \tag{7.23}$$

where α_o and γ_o are the initial polar and azimuthal angles of drone $m \leq M/2$ with respect to the array's center. $\boldsymbol{R}_{\text{rot}}$ is the rotation matrix for updating drones' positions, given by:

$$\boldsymbol{R}_{\text{rot}} = \begin{pmatrix} a_x^2(1-\delta)+\delta & a_x a_y(1-\delta)-\lambda a_z & a_x a_z(1-\delta)+\lambda a_y \\ a_x a_y(1-\delta)+\lambda a_z & a_y^2(1-\delta)+\delta & a_y a_z(1-\delta)-\lambda a_x \\ a_x a_z(1-\delta)-\lambda a_y & a_y a_z(1-\delta)+\lambda a_x & a_z^2(1-\delta)+\delta \end{pmatrix}, \tag{7.24}$$

where $\delta = \|\boldsymbol{q_i} \cdot \boldsymbol{q_{\max}}\|$, $\lambda = \sqrt{1-\delta^2}$, $\boldsymbol{q_i} = \begin{pmatrix} \sin\theta_i \cos\phi_i \\ \sin\theta_i \sin\phi_i \\ \cos\theta_i \end{pmatrix}$, $\boldsymbol{q}_{\max} = \begin{pmatrix} \sin\theta_{\max} \cos\phi_{\max} \\ \sin\theta_{\max} \sin\phi_{\max} \\ \cos\theta_{\max} \end{pmatrix}$. Moreover, a_x, a_y, and a_z are the elements of vector $\boldsymbol{a} = \left(a_x, a_y, a_z\right)^T = \boldsymbol{q_i} \times \boldsymbol{q_{\max}}$.

Proof. See Appendix D.2. □

Using Theorem 14, we can find the optimal locations of the drones such that the directivity of the drone-based antenna array is maximized towards any given ground user. Moreover, this theorem can be used to dynamically update the drones' positions for beam steering while serving different ground users.

Thus far, we have determined the optimal locations of the drones in the antenna array to maximize the directivity of the array towards any given ground user. Therefore, the data rate is maximized and, hence, the transmission time for serving the user is minimized. In Algorithm 1, we have summarized the key steps needed for optimizing the locations of drones with respect to the center of the array.

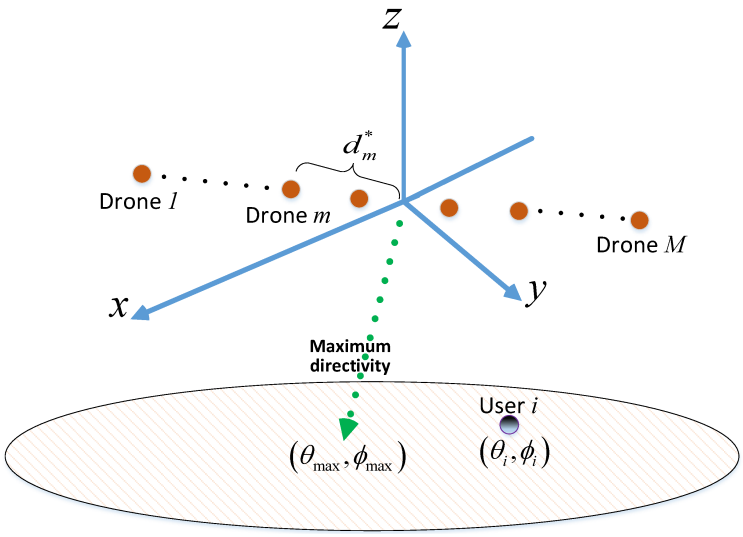


Figure 7.2: Illustrative figure for Theorem 2.

Algorithm 4 Optimizing drones' locations for maximum array gain towards user i.

1: **Inputs:** Locations of user i, $(x_i^{\mathrm{u}}, y_i^{\mathrm{u}}, z_i^{\mathrm{u}})$, and origin of array, (x_o, y_o, z_o).
2: **Outputs:** Optimal drones' positions, $(x_{m,i}^*, y_{m,i}^*, z_{m,i}^*)$, $\forall m \in \mathcal{M}$.
3: Set initial values for distance between drones, $\boldsymbol{d}$.
4: Find $\boldsymbol{e}^*$ by using (7.16)-(7.18).
5: Update $\boldsymbol{d}$ based on (7.19).
6: Repeat steps (4) and (5) to find the optimal spacing vector $\boldsymbol{d}^*$.
7: Use (7.21)-(7.58) to determine (x_m^*, y_m^*, z_m^*), $\forall m \in \mathcal{M}$.

Hence, using Algorithm 1, we can determine the optimal locations of the array's drones with respect to each ground user. To serve multiple users spread over a given geographical area, the drones must dynamically move between these determined optimal locations. This, in turn, yields a control time for drone movement that must be optimized. From (7.5), we can see that the service time decreases by reducing the control time. Therefore, next, using the determined drones' locations in Section III, we minimize the control time of the drones.

7.4 Time-Optimal Control of Drones

Here, our goal is to minimize the control time that the drones spend to move between the optimal locations which are determined in Section 7.3. In particular, we derive the optimal rotors' speeds for which the quadrotor drones can move and stabilize their positions within a minimum time. Moreover, we account for *wind effects* while analyzing the drones' stability in the proposed drone-based antenna array system.

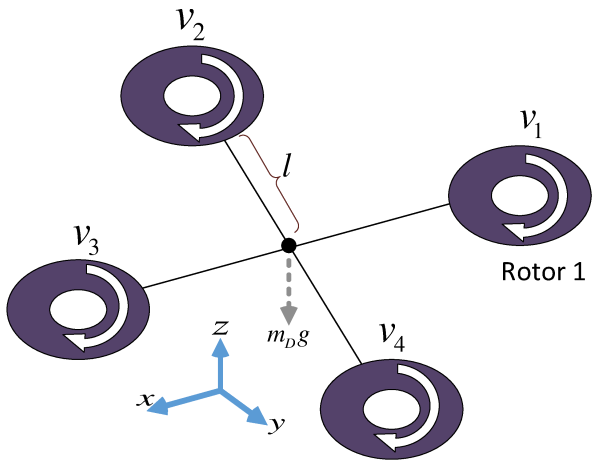


Figure 7.3: A quadrotor drone.

7.4.1 Dynamic Model of a Quadrotor Drone

Fig. 7.3 shows an illustrative example of a quadrotor drone. This drone has four rotors that can control the hovering and mobility of the drone. In particular, by adjusting the speed of these rotors, the drone can hover and move horizontally or vertically. Let (x, y, z) be the 3D position of the drone. Also, we use $(\psi_{\rm r}, \psi_{\rm p}, \psi_{\rm y})$ to represent the roll, pitch, and yaw angles that capture the orientation (i.e., attitude) of the drone. The speed of rotor $i \in \{1, 2, 3, 4\}$ is given by v_i. For a quadrotor drone, the total thrust and torques that lead to the roll, pitch, and yaw movements are related to the rotors' speeds by [200]:

$$\begin{pmatrix} T_{\rm tot} \\ \kappa_1 \\ \kappa_2 \\ \kappa_3 \end{pmatrix} = \begin{pmatrix} \rho_1 & \rho_1 & \rho_1 & \rho_1 \\ 0 & -l\rho_1 & 0 & l\rho_1 \\ -l\rho_1 & 0 & l\rho_1 & 0 \\ -\rho_2 & \rho_2 & -\rho_2 & \rho_2 \end{pmatrix} \begin{pmatrix} v_1^2 \\ v_2^2 \\ v_3^2 \\ v_4^2 \end{pmatrix}, \tag{7.25}$$

where $T_{\rm tot}$ is the total thrust generated by the rotors, κ_1, κ_2, and κ_3 are the torques for roll, pitch and yaw movements. ρ_1 and ρ_2 are lift and torque coefficients, and l is the distance from each rotor to the center of the drone.

Now, we write the dynamic equations of a quadrotor drone in presence of an external wind force as follows[3]:

$$\ddot{x} = (\cos\psi_r \sin\psi_p \cos\psi_y + \sin\psi_r \sin\psi_y) \frac{T_{\rm tot}}{m_D} + \frac{F_x^{\rm W}}{m_D}, \tag{7.26}$$

$$\ddot{y} = (\cos\psi_r \sin\psi_p \sin\psi_y + \sin\psi_r \cos\psi_y) \frac{T_{\rm tot}}{m_D} + \frac{F_y^{\rm W}}{m_D}, \tag{7.27}$$

$$\ddot{z} = (\cos\psi_r \cos\psi_p) \frac{T_{\rm tot}}{m_D} - g + \frac{F_z^{\rm W}}{m_D}, \tag{7.28}$$

[3]Note that, here, drag coefficients are assumed to be negligible.

$$\ddot{\psi}_{\mathrm{r}} = \frac{\kappa_2}{I_x}, \tag{7.29}$$

$$\ddot{\psi}_{\mathrm{p}} = \frac{\kappa_1}{I_y}, \tag{7.30}$$

$$\ddot{\psi}_{\mathrm{y}} = \frac{\kappa_3}{I_z}, \tag{7.31}$$

where m_D is the mass of the drone, and g is the gravity acceleration. F_x^{W}, F_y^{W}, and F_z^{W} are the wind forces in positive x, y, and z directions. Also, I_x, I_y, I_z are constant values which represent the moments of inertia along x, y, and z directions.

Given the dynamic model of the drone, we aim to find the optimal speeds of the rotors such that the drone moves from an initial location (x_I, y_I, z_I) to a new location (x_D, y_D, z_D) within a minimum time duration. Under such optimal control inputs (i.e., rotors' speed), the time needed for each UAV to update its location based on the users' locations will be minimized. Note that the drone must be stationary at its new location and it does not move in x, y, or z direction. Let $(x(t), y(t), z(t))$ and $(\psi_{\mathrm{r}}(t), \psi_{\mathrm{p}}(t), \psi_{\mathrm{y}}(t))$ be the 3D location and orientation of the drone at time $t \in [0, T_{I,D}]$, with $T_{I,D}$ being the total control time for moving from location I to location D. Now, we can formulate our time-optimal control problem for a drone, moving from location I to location D, as follows:

$$\underset{[v_1(t), v_2(t), v_3(t), v_4(t)]}{\text{minimize}} \; T_{I,D}, \tag{7.32}$$

$$\text{st.} \;\; |v_i(t)| \leq v_{\max}, \;\; \forall i \in \{1, ..., 4\}, \tag{7.33}$$

$$(x(0), y(0), z(0)) = (x_I, y_I, z_I)\,, \tag{7.34}$$

$$(x(T_{I,D}), y(T_{I,D}), z(T_{I,D})) = (x_D, y_D, z_D)\,, \tag{7.35}$$

$$(\dot{x}(T_{I,D}), \dot{y}(T_{I,D}), \dot{z}(T_{I,D})) = (0, 0, 0)\,, \tag{7.36}$$

where $[v_1(t), v_2(t), v_3(t), v_4(t)]$ represents the rotors' speeds at time t. In (7.33), $v_{\max}$ is the maximum possible speed of each rotor. Constraints (7.34) and (7.35) show the initial and final location of the drone, (7.36) indicates that the drone will be stationary at its final location. Here, without loss of generality, we assume $(\psi_r(0), \psi_p(0), \psi_y(0)) = (0, 0, 0)$. Note that, in (7.5), the control time for serving user i, T_i^{crl}, is equal to the maximum control time among the drones that update their positions according to the user. In this case, the initial and final locations of each drone are determined based on Algorithm 1.

Our problem in (7.32) is difficult to solve due to its non-linear nature, and coupled relation of the dynamic system parameters as well as the infinite number of optimization variables given the continuous time interval $[0, T_{I,D}]$. Consequently, in general, the exact analytical solution to such nonlinear time-optimal control problem may not be explicitly derived as pointed out in [194] and [195]. In such case, one needs to use tractable and computationally feasible methods to solve time-optimal control problems. To provide a tractable solution to our time-optimal control problem in (7.32), we decompose the movements and orientation changes of drones. In particular, we minimize the time durations needed for orientation adjustment and displacement of the drone,

separately. While this approach yields a suboptimal solution, it can be used to derive a closed-form expression for the control inputs (i.e., rotors' speeds) in (7.32) and, thus, it is remarkably easy to implement. In addition, the computational time, which is a key constraint in wireless drone systems, can be significantly reduced.

Now, we aim to derive the optimal speeds of rotors for which the drone can update its locations within a minimum time duration. To this end, we first present the following lemma from control theory [99] which will be then used to derive the optimal rotors' speeds.

Lemma 1. (From [99]): Consider the state space equations for an object within time duration $[0, T]$:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}u(t),\ u_{\min} \leq u(t) \leq u_{\max}, \tag{7.37}$$

$$\boldsymbol{x}(0) = \boldsymbol{x}_1, \tag{7.38}$$

$$\boldsymbol{x}(T) = \boldsymbol{x}_2, \tag{7.39}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^{N_s}$ is the state vector of the object at time $t \in [0, T]$, N_s is the number of state's elements. $u(t)$ is a bounded control input with $u_{\max}$ and $u_{\min}$ being its maximum and minimum values. $\boldsymbol{A} \in \mathbb{R}^{N_s \times N_s}$ and $\boldsymbol{b} \in \mathbb{R}^{N_s}$ are given constant matrices. $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are the initial and final state of the object. Then, the optimal control input that leads to a minimum state update time T^* is given by [99]:

$$u^*(t) = \begin{cases} u_{\max}, & t \leq \tau, \\ u_{\min}, & t > \tau, \end{cases} \tag{7.40}$$

where τ is called the switching time at which the control input changes. In this case, the control time decreases by increasing $u_{\max}$ and/or decreasing $u_{\min}$.

Lemma 1 provides the solution to the time-optimal control problem for a dynamic system which is characterized by (7.37)-(7.39). In particular, the optimal control solution given in (7.40) is refereed to as *bang-bang* solution [99]. In this case, the optimal control input is always at its extreme value (i.e. maximum or minimum). Next, we provide a new lemma (Lemma 2) which will be used along with Lemma 1 to solve (7.32).

Lemma 2. Consider a drone that needs to move towards a given location D (as shown in Fig.7.4), with a coordinate $\boldsymbol{P}_D = (x_D, y_D, z_D)$, in presence of an external force $\boldsymbol{F}_{\text{ex}} = (F_{\text{ex},x}, F_{\text{ex},y}, F_{\text{ex},z})$. The drone's orientation that leads to a movement with the maximum acceleration towards $\boldsymbol{P}_D$ will be:

$$\psi_{\text{p}}^D = \cos^{-1}\left[\frac{A\cos\theta_D - |\boldsymbol{F}_{\text{ex}}|\cos\theta_{\text{ex}}}{F}\right], \tag{7.41}$$

$$\psi_{\text{r}}^D = \tan^{-1}\left(\tan\beta \times \sin\psi_p^D\right), \tag{7.42}$$

$$\psi_{\text{y}}^D = 0, \tag{7.43}$$

where
$A = \left[F^2 + |\boldsymbol{F}_{\text{ex}}|^2 + 2F|\boldsymbol{F}_{\text{ex}}|\cos\left(\gamma + \sin^{-1}\left(\frac{|\boldsymbol{F}_{\text{ex}}|}{F}\sin\gamma\right)\right)\right]^{1/2}$, $\beta = \phi_D - \sin^{-1}\left[\frac{|\boldsymbol{F}_{\text{ex}}|\sin\theta_{\text{ex}}\sin(\phi_D - \phi_{\text{ex}})}{F\sin\psi_p^D}\right]$,

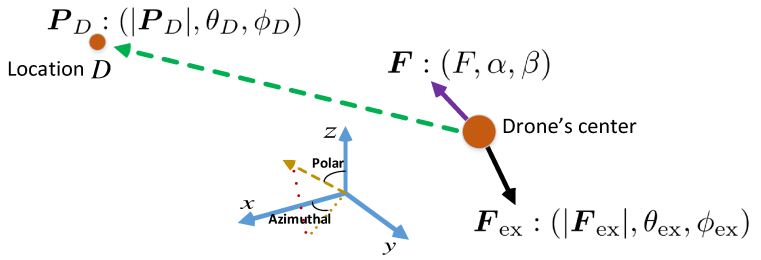


Figure 7.4: Drone's movement in presence of an external force.

$\gamma = \cos^{-1}\left(\frac{\boldsymbol{F}_{\text{ex}}.\boldsymbol{P}_D}{|\boldsymbol{F}_{\text{ex}}||\boldsymbol{P}_D|}\right)$, and F is the magnitude of the maximum force of the drone. $|\boldsymbol{F}_{\text{ex}}|$ represents the magnitude of vector $\boldsymbol{F}_{\text{ex}}$, $\theta_{\text{ex}} = \cos^{-1}\left(\frac{F_{\text{ex},z}}{|\boldsymbol{F}_{\text{ex}}|}\right)$, $\phi_{\text{ex}} = \tan^{-1}\left(\frac{F_{\text{ex},y}}{F_{\text{ex},x}}\right)$, $\phi_D = \tan^{-1}\left(\frac{y_D}{x_D}\right)$, and $\theta_D = \cos^{-1}\left(\frac{z_D}{|\boldsymbol{P}_D|}\right)$.

Proof. See Appendix D.3. □

Lemma 2, can be used to determine the optimal orientation of the drone that enables it to move towards any given location in presence of external forces. Next, using Lemmas 1 and 2, we derive the speed of each drone's rotor for which the control time is minimized. In this case, we find the rotors' speeds at several pre-defined *stages* in which the drone updates its position or ordination.

Theorem 15. The optimal speeds of rotors with which a drone can move from location $(0, 0, 0)$, and $(0, 0, 0)$ orientation, to location (x_D, y_D, z_D) within a minimum control time are given by:

$$\text{Stage 1:} \begin{cases} v_2 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}} v_{\max}, v_4 = v_{\max}, & \text{if } 0 < t \leq \tau_1, \\ v_4 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}} v_{\max}, v_2 = v_{\max}, & \text{if } \tau_1 < t \leq \tau_2, \\ v_1 = 0, v_2 = v_4 = \frac{1}{\sqrt{2}} v_{\max}, v_3 = v_{\max}, & \text{if } \tau_2 < t \leq \tau_3, \\ v_3 = 0, v_2 = v_4 = \frac{1}{\sqrt{2}} v_{\max}, v_1 = v_{\max}, & \text{if } \tau_3 < t \leq \tau_4. \end{cases} \tag{7.44}$$

$$\text{Stage 2:} \quad v_1 = v_2 = v_3 = v_4 = v_{\max}, \quad \text{if } \tau_4 < t \leq \tau_5. \tag{7.45}$$

$$\text{Stage 3:} \begin{cases} v_2 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}} v_{\max}, v_4 = v_{\max}, & \text{if } \tau_5 < t \leq \tau_6, \\ v_4 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}} v_{\max}, v_2 = v_{\max}, & \text{if } \tau_6 < t \leq \tau_7, \\ v_1 = 0, v_2 = v_4 = v_{\max}, v_3 = v_{\max}, & \text{if } \tau_7 < t \leq \tau_8, \\ v_3 = 0, v_2 = v_4 = \frac{1}{\sqrt{2}} v_{\max}, v_1 = v_{\max}, & \text{if } \tau_8 < t \leq \tau_9. \end{cases} \tag{7.46}$$

$$\text{Stage 4:} \quad v_1 = v_2 = v_3 = v_4 = v_{\max}, \text{ if } \tau_9 < t \leq \tau_{10}. \tag{7.47}$$

$$\text{Stage 5:} \begin{cases} v_2 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}} v_{\max}, v_4 = v_{\max}, & \text{if } \tau_{10} < t \leq \tau_{11}, \\ v_4 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}} v_{\max}, v_2 = v_{\max}, & \text{if } \tau_{11} < t \leq \tau_{12}, \\ v_1 = 0, v_2 = v_4 = \frac{1}{\sqrt{2}} v_{\max}, v_3 = v_{\max}, & \text{if } \tau_{12} < t \leq \tau_{13}, \\ v_3 = 0, v_2 = v_4 = \frac{1}{\sqrt{2}} v_{\max}, v_1 = v_{\max}, & \text{if } \tau_{13} < t \leq \tau_{14}. \end{cases} \tag{7.48}$$

Algorithm 5 Steps for minimizing the service time by solving (7.5).

1: **Inputs:** Locations of users, $(x_i^{\mathrm{u}}, y_i^{\mathrm{u}}, z_i^{\mathrm{u}})$, $\forall i \in \mathcal{L}$, and origin of array, (x_o, y_o, z_o).
2: **Outputs:** Optimal drones' positions, $(x_{m,i}^*, y_{m,i}^*, z_{m,i}^*)$, rotors' speeds, $v_{mn}(t), \forall m \in \mathcal{M}, \forall i \in \mathcal{L}, n \in \{1, ..., 4\}$, and total service time.
3: Using Algorithm 1, find the optimal locations of drones with respect to each user, $(x_{m,i}^*, y_{m,i}^*, z_{m,i}^*)$.
4: Using Theorem 3 and Lemma 2, for each drone, determine the rotors' speeds for moving from $(x_{m,i-1}^*, y_{m,i-1}^*, z_{m,i-1}^*)$ to $(x_{m,i}^*, y_{m,i}^*, z_{m,i}^*)$.
5: Compute the total service time based on (7.5), (7.32), and (7.50).

$$\text{Stage 6:} \;\; v_1 = v_2 = v_3 = v_4 = v_{\mathrm{F}}, \;\text{if}\; t > \tau_{14}. \tag{7.49}$$

Also, the total control time of the drone can be given by:

$$\begin{aligned} T_{I,D} =& \sqrt{2d_D\Big(\frac{m_D}{A_{s2}} - \frac{m_D}{A_{s4}}\Big)} + \frac{2}{v_{\max}}\Big[\sqrt{\frac{\Delta\psi_{\mathrm{p},1} I_y}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{r},1} I_x}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{p},3} I_y}{l\rho_1}} \\ &+ \sqrt{\frac{\Delta\psi_{\mathrm{r},3} I_x}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{p},5} I_y}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{r},5} I_x}{l\rho_1}}\Big], \end{aligned} \tag{7.50}$$

where $v_{\max}$, v_{in}, and v_{F} are, respectively, the maximum, the initial, and the final speeds of rotors. m_D is the drone's mass, $\Delta\psi_{\mathrm{r},i}$ and $\Delta\psi_{\mathrm{p},i}$ are the roll and pitch changes in Stage i. d_D is the distance between the initial and final locations of the drone. $\tau_1, ..., \tau_{14}$ are the switching times at which the rotors' speeds changes. The values of switching times and v_F are provided in the proof of this theorem.

Proof. See Appendix D.4. □

In Theorem 15, Stages 1, 3, and 5 correspond to the orientation changes, Stages 2 and 4 are related to the drone's displacement, and Stage 6 represents the drone's stability condition. Note that v_{F} is adjusted such that the drone's stability is ensured at its final location. In (7.50), A_{s2} and A_{s4} are, respectively, the total forces towards the drone's destination at Stages 2 and 4.

Using Theorem 15, we can find the speeds of the rotors (at different time instances) that enable each to move towards its destination within a minimum time. In this case, the control time depends on the destination of the drone, external forces (e.g. wind and gravity), the rotors' speed, and the drone's weight.

In summary, our approach for minimizing the service time, which is composed of the transmission time and the control time, is as follows. In the first step, using the approach in Section III, we minimize the transmission time for each ground user by optimizing the positions of drones with respect to the ground users. Then, based on these determined optimal drones' locations, we minimize the control time needed for adjusting the movement and orientations of drones. In Algorithm 5, we summarize our approach for minimizing the service time by solving (7.5).

7.5 Simulation Results and Analysis

For our simulations, we consider a number of ground users uniformly distributed within a square area of size $1\,\text{km} \times 1\,\text{km}$. Unless stated otherwise, the number of users is 100, and the number of drones[4] that form a linear array is assumed to be 10. The main simulation parameters are given in Table 7.1. We compare the performance of our drone-based antenna array system with a case in which the same number of drones are deployed separately to serve ground users (hereinafter, this case is referred to as *multi-drones*). In the multi-drone scenario, multiple drones are separately deployed based on a grid-based deployment. In this case, each drone serves its associated ground users which are determined based on the minimum distance association rule. For the multi-drones case, we do not consider any control time. All statistical results are averaged over a large number of independent runs.

Fig. 7.5 shows the total service time for the drone antenna array and the multi-drone case. Fig. 7.5 shows that, for a given bandwidth, our proposed drone antenna array outperforms the multi-drone case in terms of service time. This is because, in the proposed approach, the drones' locations are adjusted such that the array antenna gain towards each user is maximized, hence reducing the transmission time. Fig. 7.5 also shows the tradeoff between bandwidth and service time. Clearly, the service time decreases by using more bandwidth which effectively provides a higher data rate. Fig. 7.5 shows that the drone antenna array improves spectral efficiency compared to the multi-drones case. For instance, to achieve 10 minutes of service time, the drone antenna array will require 48% less bandwidth than in the multi-drones scenario.

In Fig. 7.6, we show the impact of the number of users on the service time. Clearly, the service time increases as the number of users increases. For a higher number of users, the drones must deliver a higher data service which results in a higher transmission time. Moreover, in the drone antenna array case, the control time also increases while increasing the number of users. Fig. 7.6 shows that our proposed drone antenna array system outperforms the multi-drones case for various number of users. For instance, using our approach, the average service time can be reduced by 20 min (or 30%) while serving 400 users. Hence, compared to the multi-drones case, the drones need to fly for 20 min less in order to service the ground users. Meanwhile, the users can receive faster wireless services while exploiting the proposed drone antenna array system.

Fig. 7.7 shows how the control, transmission, and service times resulting from the proposed approach for different numbers of drones in the array. As the number of drones increases, the control time increases. In contrast, the transmission time (for 10 MHz bandwidth) decreases due to the increase of the array gain. Fig. 7.7 shows that, by increasing the number of drones from 10 to 30, the average control time increases by 20% while the average transmission time decreases by 36%. Therefore, there is a tradeoff between the transmission time and the control time as a function of the number of drones in the array. In this figure, the decrease in the transmission time is more than the additional control time and, thus, the total service time decreases by increasing the number of drones.

[4] In our simulations, each drone in the array has an omni-directional antenna, as in [97, 191].

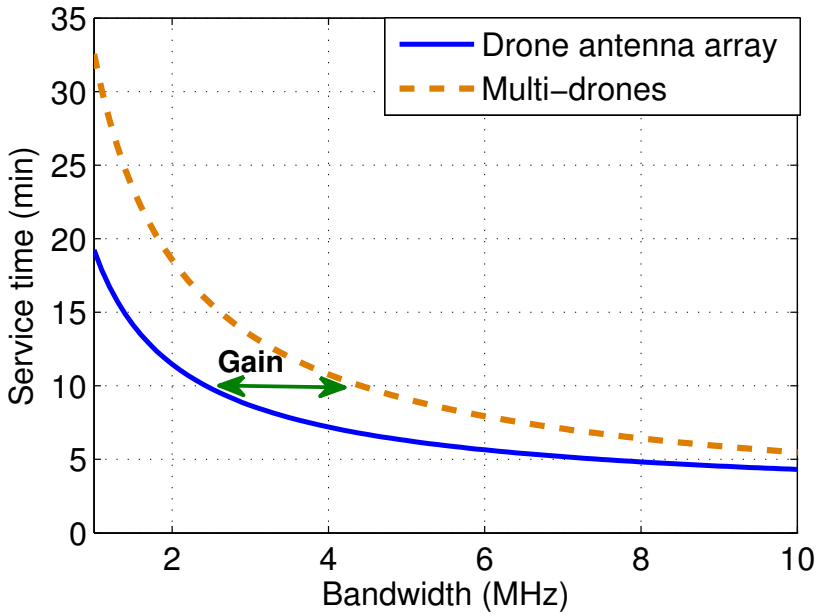


Figure 7.5: Service time vs. bandwidth for the drone antenna-array and multi-drones cases.

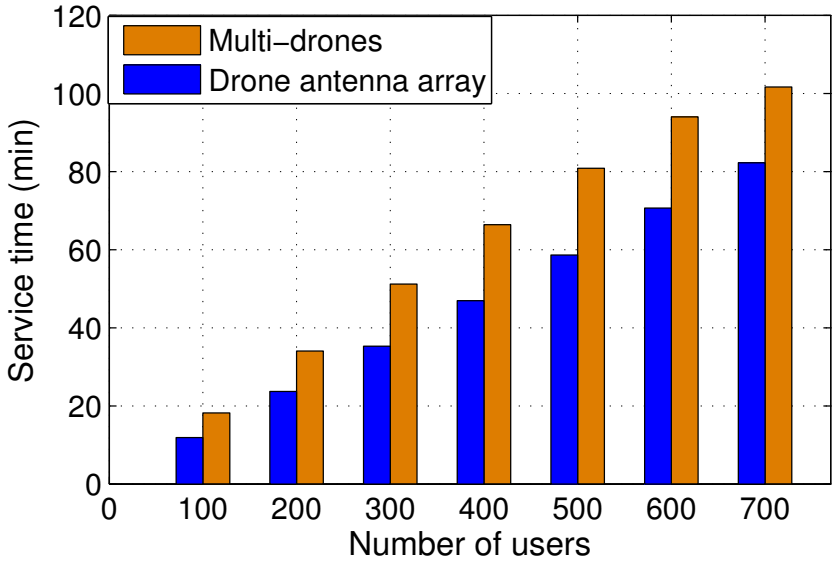


Figure 7.6: Service time vs. number of users for the drone antenna array and multi-drones (2MHz bandwidth).

In Fig. 7.8, we show how the number of users impacts the control time. As we can see from this figure, the control time increases while serving more users. This is due to the fact that, for a higher number of users, the drone-array must move more in order to steer its beam toward the users. The

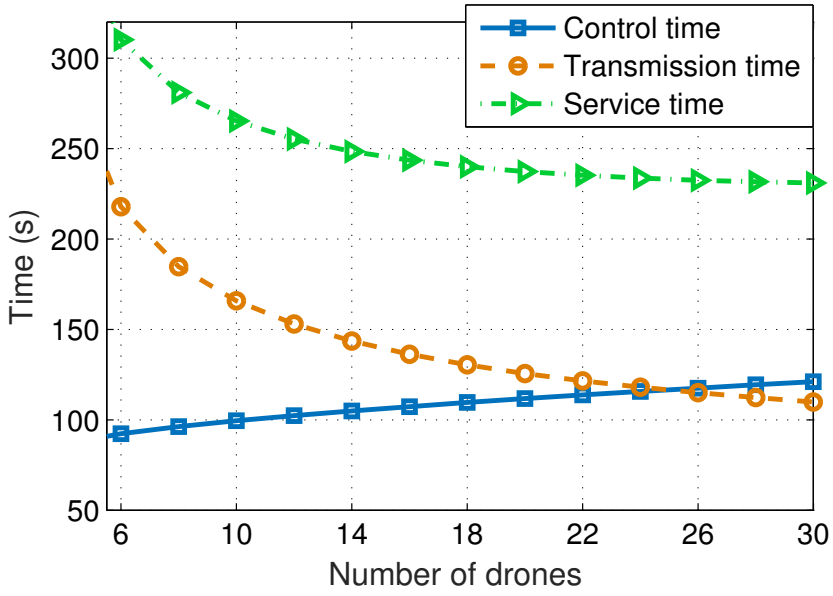


Figure 7.7: Control, transmission, and service times vs. number of drones.

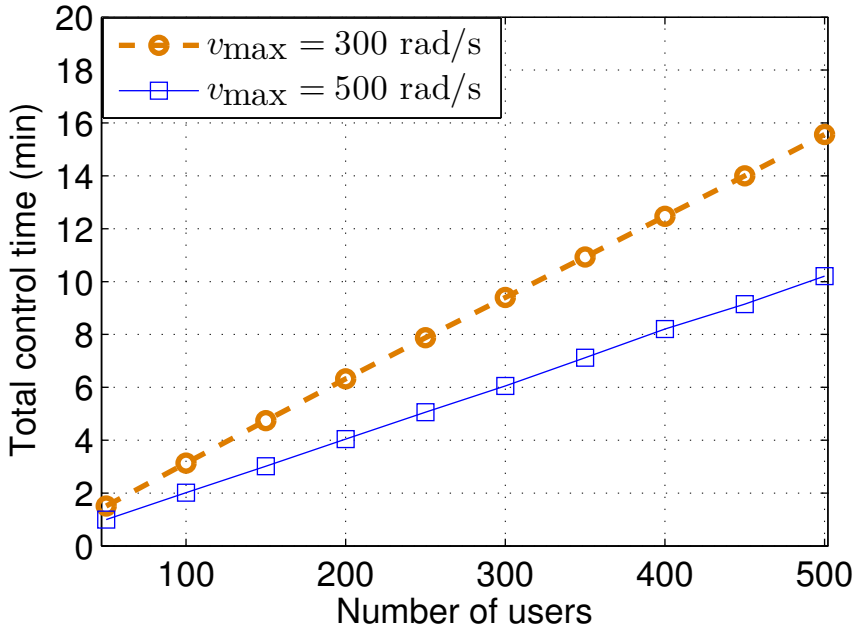


Figure 7.8: Total control time vs. number of users.

control time can be reduced by increasing the maximum speed of the rotors, which is in agreement with Theorem 3. For instance, increasing the maximum rotors' speed from 300 rad/s to 500 rad/s yields around 35% control time reduction when serving 200 users.

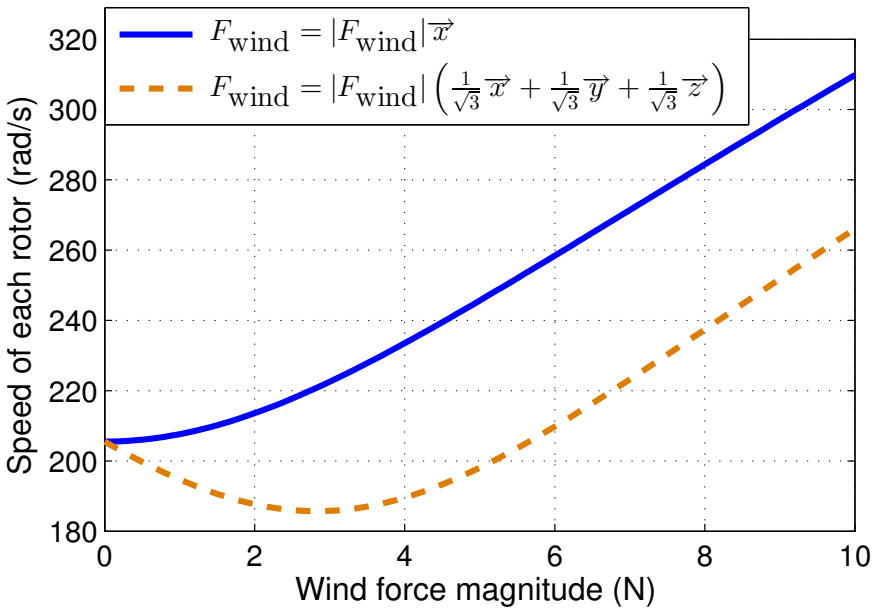


Figure 7.9: Speed of each rotor vs. wind force under the drone's stability condition.

Fig. 7.9 represents the speeds of the rotors needed to ensure the drone's stability in presence of wind, obtained using (7.77). Clearly, the drone is stable when its total force which is composed of the wind force, gravity, and the drone force is zero. For $F_{\text{wind}} = |F_{\text{wind}}|\overrightarrow{x}$, the rotor's speed must increase as the wind force increases. In the $F_{\text{wind}} = |F_{\text{wind}}|\left(\frac{1}{\sqrt{3}}\overrightarrow{x} + \frac{1}{\sqrt{3}}\overrightarrow{y} + \frac{1}{\sqrt{3}}\overrightarrow{z}\right)$ case, however, the rotor's speed first decreases, and then increases. This is because, when $|F_{\text{wind}}| \leq 3\,\text{N}$, the wind force helps hovering the drone by compensating for the gravity. Hence, the drone's force can be decreased by decreasing the speed of its rotors. For $|F_{\text{wind}}| > 3\,\text{N}$, the rotor's speed start increasing such that the total force on the drone becomes zero. This result also implies that, in some cases (depending on the magnitude and direction of wind), wind can facilitate hovering of the drone by overcoming the gravity force. However, in case of strong winds, the drone's stability may not be guaranteed by adjusting the speed of the rotors. This is because the drone force, which is limited by the maximum rotors' speeds, cannot overcome the external forces induced by wind and gravity forces.

7.6 Summary

In this chapter, we have proposed a novel framework for employing a drone-enabled antenna array system that can provide wireless services to ground users within a minimum time. To this end, we have minimized the transmission time as well as the control time needed for changing the locations and orientations of the drones. First, we have optimized the positions of drones (as the array elements) within the antenna array such that the transmission time for each user is minimized. Next, given the determined locations of drones, we have minimized the control time of the quadrotor drones by optimally adjusting the rotors' speeds. Our results have shown that the

Table 7.1: Main simulation parameters.

Parameter	Description	Value
f_c	Carrier frequency	700 MHz
P_i	Drone transmit power	0.1 W
N_o	Total noise power spectral density	-160 dBm/Hz
N	Number of ground users	100
(x_o, y_o, z_o)	Array's center coordinate	(0,0,100) in meters
q_i	Load per user	100 Mb
α	Pathloss exponent	3
I_x, I_y	Moments of inertia	4.9×10^{-3}kg.m^2 [201]
m_D	Mass of each drone	0.5 kg
l	Distance of a rotor to drone's center	20 cm
ρ_1	lift coefficient	2.9×10^{-5} [201]
$\beta_m - \beta_{m-1}$	Phase excitation difference for two adjacent antennas	$\frac{\pi}{5(M-1)}$

proposed drone antenna array with the optimal configuration yields a significant improvement in terms of the service time, spectral and energy efficiency. In addition, our results have revealed key design guidelines and fundamental tradeoffs for leveraging in an antenna array system. To our best knowledge, this is the first comprehensive study on the joint communications and control of drone-based antenna array systems.

Appendix D

D.1 Proof of Theorem 13

First, we find $F^2(\theta, \phi)$ by using (7.11):

$$F^2(\theta,\phi) = \left[2F^0(\theta,\phi)\right]^2 + \left[2\sum_{n=1}^{N} a_n k e_n \sin\theta\cos\phi \sin\left(kd_n^0 \sin\theta\cos\phi + \beta_n\right)\right]^2 - 8F^0(\theta,\phi)\sum_{n=1}^{N} a_n k e_n \sin\theta\cos\phi\sin\left(kd_n^0\sin\theta\cos\phi + \beta_n\right).$$

Subsequently, our objective function in (7.12) can be written as:

$$I_{\text{int}}\left(F^2(\theta,\phi)w^2(\theta,\phi)\right) = 4\left[\boldsymbol{e}^T\boldsymbol{G}\boldsymbol{e} - 2\boldsymbol{e}^T\boldsymbol{q} + I_{\text{int}}\left(F_0^2(\theta,\phi)w^2(\theta,\phi)\right)\right], \tag{7.51}$$

where $\boldsymbol{G}$ and $\boldsymbol{q}$ are given in (7.17) and (7.18). Clearly, (7.51) is a quadratic function of $\boldsymbol{e}$. Therefore, (7.51) is convex if and only if $\boldsymbol{G}$ is a positive semi-definite matrix. Given (7.17), we have:

$$\boldsymbol{y}^T\boldsymbol{G}y = \sum_{n=1}^{N} y_n \sum_{m=1}^{N} y_m g_{m,n}. \tag{7.52}$$

Now, in (7.17), let us define

$$z_n = a_n k \sin\theta \cos\phi w(\theta,\phi) \sin\left(k d_n^0 \sin\theta\cos\phi + \beta_n\right), \tag{7.53}$$

then, using (7.52), we have:

$$\boldsymbol{y}^T \boldsymbol{G} \boldsymbol{y} = I_{\text{int}} \left(\left[\sum_{n=1}^{N} z_n y_n \right]^2 \right). \tag{7.54}$$

In (7.15), we can see that $I_{\text{int}}(x) \geq 0$ for $x \geq 0$. Hence, from (7.54), we can conclude that $\boldsymbol{y}^T \boldsymbol{G} \boldsymbol{y} \geq 0$. Therefore, $\boldsymbol{G}$ is positive semi-definite and the objective function in (7.12) is convex. Moreover, the constraints in (7.13) are affine functions which are convex. Hence, this optimization problem is convex. Now, we find the optimal perturbation vector $\boldsymbol{e}$ by using Karush-Kuhn-Tucker (KKT) conditions. The Lagrangian function will be:

$$\mathcal{L} = \boldsymbol{e}^T \boldsymbol{G} \boldsymbol{e} - 2\boldsymbol{e}^T \boldsymbol{q} + I_{\text{int}} \left(F_0^2(\theta,\phi) w^2(\theta,\phi) \right) + \sum_{n=1}^{N-1} \mu_n \left(e_n - e_{n+1} + D_{\min} + d_n^0 - d_{n+1}^0 \right), \tag{7.55}$$

where $\mu_n \geq 0$, $n = 1, ..., N-1$ are the Lagrange multipliers.

The necessary and sufficient (due to the convexity of the problem) KKT conditions for finding the optimal perturbation vector $\boldsymbol{e}$ are given by:

$$\nabla_{\boldsymbol{e}} \left[\mathcal{L} \right] = 0, \tag{7.56}$$

which leads to $\boldsymbol{e} = \boldsymbol{G}^{-1}[\boldsymbol{q} + \boldsymbol{\mu}_{\boldsymbol{\mathcal{L}}}]$, with $\boldsymbol{\mu}_{\boldsymbol{\mathcal{L}}}$ being a $(N-1) \times 1$ vector whose element n is $\boldsymbol{\mu}_{\boldsymbol{\mathcal{L}}}(n) = \mu_{n+1} - \mu_n$. Based on the complementary slackness conditions, we have:

$$\begin{cases} \mu_n \left(e_n - e_{n+1} + D_{\min} + d_n^0 - d_{n+1}^0 \right) = 0, \ \forall n \in \mathcal{N} \setminus \{N\}, \\ \mu_n \geq 0, \ \forall n \in \mathcal{N} \setminus \{N\}. \end{cases} \tag{7.57}$$

Finally, the optimal perturbation vector, $\boldsymbol{e}^*$, can be determined by solving (7.56) and (7.57). This proves the theorem.

D.2 Proof of Theorem 14

In Subsection III-A, we have derived the optimal distance of drones from the origin that leads to a maximum array directivity. First, we consider an initial (or arbitrary) orientation, as shown in Figure 7.2. Let d_m^* be the optimal distance of drone $m \leq M/2$ from the array's center, α_o and γ_o be the initial polar and azimuthal angles of the drone. Based on the considered drones' locations, let $(\theta_{\max}, \phi_{\max}) = \text{argmax}\left[F(\theta,\phi)w(\theta,\phi)\right]$ be a direction at which the directivity of the array is maximized. Our goal is to achieve the maximum directivity at a given direction (θ_i, ϕ_i) corresponding to user i. Therefore, we need to change the locations of the drones such that $\theta_i = \theta_{\max}$, and $\phi_i = \phi_{\max}$. To this end, we align the unit vector $(1, \theta_{\max}, \phi_{\max})$ with $(1, \theta_i, \phi_i)$ in the spherical coordinate and, then, we update the drones' positions accordingly. In the Cartesian coordinate system, we need to rotate vector $\boldsymbol{q}_{\max} = \left(\sin\theta_{\max}\cos\phi_{\max}, \sin\theta_{\max}\sin\phi_{\max}, \cos\theta_{\max}\right)^T$ such that it becomes aligned with $\boldsymbol{q_i} = \left(\sin\theta_i\cos\phi_i, \sin\theta_i\sin\phi_i, \cos\theta_i\right)^T$.

The rotation matrix for rotating a vector $\boldsymbol{u}$ about another vector $\boldsymbol{a} = \left(a_x, a_y, a_z\right)^T$, with a ω rotation angle, is [202]:

$$\boldsymbol{R}_{\text{rot}} = \begin{pmatrix} a_x^2(1-\cos\omega)+\cos\omega & a_x a_y(1-\cos\omega)-a_z\sin\omega & a_x a_z(1-\cos\omega)+a_y\sin\omega \\ a_x a_y(1-\cos\omega)+a_z\sin\omega & a_y^2(1-\cos\omega)+\cos\omega & a_y a_z(1-\cos\omega)-a_x\sin\omega \\ a_x a_z(1-\cos\omega)-a_y\sin\omega & a_y a_z(1-\cos\omega)+a_x\sin\omega & a_z^2(1-\cos\omega)+\cos\omega \end{pmatrix}, \tag{7.58}$$

In our problem, the rotation between $\boldsymbol{q}_{\max}$ and $\boldsymbol{q_i}$ can be done about the normal vector of these vectors, with the rotation angle being the angle between $\boldsymbol{q}_{\max}$ and $\boldsymbol{q_i}$. Hence, based on the dot-product and cross-product of vectors, we use $\boldsymbol{a} = \boldsymbol{q_i} \times \boldsymbol{q_{\max}}$, and $\omega = \cos^{-1}(\boldsymbol{q_i} \cdot \boldsymbol{q_{\max}})$ to find the rotation matrix in (7.58). Now, we update the locations of drones using the rotation matrix. Clearly, for $m \leq M/2$, the initial location of drone m in the Cartesian coordinate is $\left(d_m^* \sin\alpha_o \cos\gamma_o, d_m^* \sin\alpha_o \sin\beta_o, d_m^* \cos\alpha_o\right)^T$. As a result, the optimal locations of drones for serving user i is given by:

$$\left(x_m^*, y_m^*, z_m^*\right)^T = \boldsymbol{R}_{\text{rot}} \left(d_m^* \sin\alpha_o \cos\gamma_o, d_m^* \sin\alpha_o \sin\beta_o, d_m^* \cos\alpha_o\right)^T, \text{if } m \leq M/2. \tag{7.59}$$

Finally, due to the symmetric configuration of the antenna array about the origin, the optimal locations of drones m when $m > M/2$ are as follows:

$$\left(x_m^*, y_m^*, z_m^*\right)^T = -\boldsymbol{R}_{\text{rot}} \left(d_m^* \sin\alpha_o \cos\gamma_o, d_m^* \sin\alpha_o \sin\beta_o, d_m^* \cos\alpha_o\right)^T, \text{if } m \leq M/2. \tag{7.60}$$

This completes the proof.

D.3 Proof of Lemma 2

To maximize the drone's acceleration towards the given location D, we need to maximize the total force in the direction of $\boldsymbol{P}_D$. Considering the center of the drone as the origin of the Cartesian and spherical coordinate systems, we can present the vectors of forces and the movement as in Fig. 7.4. In this figure, based on the Cartesian-to-spherical coordinates transformation, the polar and azimuthal angles in the spherical coordinate are given by $\theta_{\text{ex}} = \cos^{-1}\left(\frac{F_{\text{ex},z}}{|\boldsymbol{F}_{\text{ex}}|}\right)$, $\phi_{\text{ex}} = \tan^{-1}\left(\frac{F_{\text{ex},y}}{F_{\text{ex},x}}\right)$, $\phi_D = \tan^{-1}\left(\frac{y_D}{x_D}\right)$, and $\theta_D = \cos^{-1}\left(\frac{z_D}{|\boldsymbol{P}_D|}\right)$. Let α and β be, respectively, the polar and azimuthal angles of the drone's force. Here, we seek to determine α and β such that the drone can move towards location D with a maximum acceleration (i.e., maximum total force). In this case, the total force $\boldsymbol{F}_{\text{ex}} + \boldsymbol{F}$ must be in the same direction as $\boldsymbol{P}_D$. Let γ be the angle between $\boldsymbol{F}$ and $\boldsymbol{P}_D$, and q be the angle between $\boldsymbol{F}_{\text{ex}}$ and $\boldsymbol{P}_D$. To ensure that $\boldsymbol{F}_{\text{ex}} + \boldsymbol{F}$ is in the direction of $\boldsymbol{P}_D$, we should have:

$$|\boldsymbol{F}_{\text{ex}}| \sin\gamma = |\boldsymbol{F}| \sin q = F \sin q. \tag{7.61}$$

Also, using the inner product formula, γ is given by:

$$\gamma = \cos^{-1}\left(\frac{\boldsymbol{F}_{\text{ex}} \cdot \boldsymbol{P}_D}{|\boldsymbol{F}_{\text{ex}}||\boldsymbol{P}_D|}\right). \tag{7.62}$$

As a result, q will be:

$$q = \sin^{-1}\left(\frac{|\boldsymbol{F}_{ex}|}{|\boldsymbol{F}|}\sin\left[\cos^{-1}\left(\frac{\boldsymbol{F}_{ex}.\boldsymbol{P}_D}{|F_{ex}||P_D|}\right)\right]\right). \tag{7.63}$$

Now, based on the law of cosines, the total force magnitude is equal to:

$$A \triangleq |\boldsymbol{F}_{\text{ex}} + \boldsymbol{F}| = \left[F^2 + |\boldsymbol{F}_{\text{ex}}|^2 + 2F|\boldsymbol{F}_{\text{ex}}|\cos\left(\gamma + \sin^{-1}\left(\frac{|\boldsymbol{F}_{\text{ex}}|}{F}\sin\gamma\right)\right)\right]^{1/2}. \tag{7.64}$$

By projection $(\boldsymbol{F}_{\text{ex}} + \boldsymbol{F})$, $\boldsymbol{F}_{\text{ex}}$, and $\boldsymbol{F}$ on z-axis and $x-y$ plane, we have:

$$A\cos\theta_D = |\boldsymbol{F}_{\text{ex}}|\cos\theta_{\text{ex}} + F\cos\alpha, \tag{7.65}$$

$$|\boldsymbol{F}_{\text{ex}}|\sin\theta_{\text{ex}}\sin\left(\phi_D - \phi_{\text{ex}}\right) = F\sin\alpha\sin\left(\phi_D - \beta\right). \tag{7.66}$$

Subsequently, we obtain α and β as follows:

$$\alpha = \cos^{-1}\left[\frac{A\cos\theta_D - |\boldsymbol{F}_{\text{ex}}|\cos\theta_{\text{ex}}}{F}\right], \tag{7.67}$$

$$\beta = \phi_D - \sin^{-1}\left[\frac{|\boldsymbol{F}_{\text{ex}}|\sin\theta_{\text{ex}}\sin\left(\phi_D - \phi_{\text{ex}}\right)}{F\sin\psi_p^D}\right]. \tag{7.68}$$

Finally, considering the fact that the drone's force is perpendicular to its rotors' plane, as well as using the transformation between body-frame and earth-frame, the drone's orientation can be given by[5]:

$$\psi_{\text{p}}^D = \alpha, \psi_{\text{r}}^D = \tan^{-1}\left(\tan\beta \times \sin\psi_p^D\right), \psi_{\text{y}}^D = 0, \tag{7.69}$$

which proves Lemma 2.

7.6.1 D.4 Proof of Theorem 15

Let $s(t)$ be the distance that the drone moves towards destination D at time t. We define state $\boldsymbol{g}(t) = [s(t), \dot{s}(t)]^T$, and provide the following equation:

$$\dot{\boldsymbol{g}}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\boldsymbol{g}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}a_D(t), \tag{7.70}$$

where $a_{\text{min}} \leq a_D(t) \leq a_{\text{max}}$ is the drone's acceleration towards D, with a_{min} and a_{max} being the minimum and maximum values of $a_D(t)$. Clearly, the drone can reach the destination and stop at D within duration T, if $\boldsymbol{g}(T) = [0,0]^T$. Based on Lemma 1, T is minimized when $a_D(t) = \begin{cases} a_{\text{max}}, & 0 < t \leq \tau, \\ a_{\text{min}}, & \tau < t \leq T. \end{cases}$. Now, we find τ by using kinematic equations that describe an object's motion. Let d_D be the distance between the initial and the final locations of the drone. Clearly, the drone's displacement until $t = \tau$ is equal to $\frac{1}{2}a_{\text{max}}\tau^2$. During $\tau < t \leq T$, the displacement will be $\frac{1}{2}a_{\text{min}}(T-\tau)^2 + a_{\text{max}}\tau(T-\tau)$. Hence, the total drone's disparagement is:

[5]Note that, we consider $(0,0,0)$ as the initial drone's orientation. Here, to change the orientation, we first update the pitch angle and, then, we update the roll angle.

$$d_D = \frac{1}{2}a_{\max}\tau^2 + \frac{1}{2}a_{\min}(T-\tau)^2 + a_{\max}\tau(T-\tau). \tag{7.71}$$

Also, considering the fact that drone stops (i.e. zero speed) at $t = T$, we have:

$$a_{\max}\tau + a_{\min}(T-\tau) = 0, \tag{7.72}$$

According to (7.71) and (7.72), the total control time, T, and the switching time can be found by:

$$T = \sqrt{2d_D(\frac{1}{a_{\max}} - \frac{1}{a_{\min}})}, \tag{7.73}$$

$$\tau = \frac{a_{\min}}{a_{\min} - a_{\max}}T. \tag{7.74}$$

As we can see from (7.73), T can be minimized by maximizing $a_{\max}$ and minimizing $a_{\min}$. To this end, we will adjust the drone's orientation as well as the rotors' speeds. Each drone's orientation can be determined by using Lemma 2. Also, given (7.25)-(7.28), we can show that the optimal speeds of the rotors are $v_1 = v_2 = v_3 = v_4 = v_{\max}$.

To adjust the drone's orientation within a minimum time, we minimize the time needed for the pitch and roll updates. Using a similar approach as in (7.70), and considering (7.25), (7.29), (7.30), and zero yaw angle (i.e. $v_2^2 + v_4^2 = v_1^2 + v_3^2$), the optimal rotors' speeds can be given by:

$$\text{positive change of pitch angle: } \begin{cases} v_2 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}}v_{\max}, v_4 = v_{\max}, & \text{if } 0 < t \leq \tau_1, \\ v_4 = 0, v_1 = v_3 = \frac{1}{\sqrt{2}}v_{\max}, v_2 = v_{\max}, & \text{if } \tau_1 < t \leq \tau_2, \end{cases} \tag{7.75}$$

$$\text{positive change of roll angle: } \begin{cases} v_1 = 0, v_2 = v_4 = \frac{1}{\sqrt{2}}v_{\max}, v_3 = v_{\max}, & \text{if } \tau_2 < t \leq \tau_3, \\ v_3 = 0, v_2 = v_4 = \frac{1}{\sqrt{2}}v_{\max}, v_1 = v_{\max}, & \text{if } \tau_3 < t \leq \tau_4, \end{cases} \tag{7.76}$$

Therefore, in the first Stage, the drone changes its orientation such that it can move towards D in presence of external forces (e.g., gravity and wind). In the second Stage, the drone moves with a maximum acceleration. In Stage 3, the drone's orientation changes to minimize the acceleration towards D. In Stage 4, the drone moves with a minimum acceleration. In Stages 5 and 6, the drone's orientation and the rotors' speeds are adjusted to ensure the stability of drone at D. Clearly, the drone will be stable when its total force, A given in (7.64), is zero. Hence, we must have $F = |\boldsymbol{F}_{\text{ext}}|$. Now, using (7.25) with $T_{\text{tot}} = |\boldsymbol{F}_{\text{ext}}|$, the rotors' speed in the stable stage (i.e., Stage 6) will be:

$$v_{\text{F}} = \sqrt{\frac{|\boldsymbol{F}_{\text{ext}}|}{4\rho_1}}. \tag{7.77}$$

The rotors' speed in Stages 1-6 are given in (9.11)-(7.49).

In order to find the switching times, we use the dynamic equations of the drone given in (7.25-7.29). For instance, in Stage 1, the time needed for a $\Delta\psi_{\text{p},1}$ pitch angle change can be obtained

using (7.25) and (7.29). In this case, given the rotors' speed in (9.11), and the dynamic equations of the drone, we can find τ_1 and τ_2 as:

$$\tau_1 = \frac{1}{v_{\max}} \sqrt{\frac{\Delta\psi_{\mathrm{p},1} I_y}{l\rho_1}}, \tag{7.78}$$

$$\tau_2 = 2\tau_1. \tag{7.79}$$

where $\Delta\psi_{\mathrm{p},1}$ is the change of pitch angle at Stage 1. Likewise, τ_3 and τ_4 can also be determined.

In Stage 2, the time needed for moving within a $d_{\mathrm{s}2}$ distance is given by:

$$t_{s2} = \sqrt{\frac{2 d_{s2} A_{s2}}{m_D}}, \tag{7.80}$$

where A_{s2} is the total force towards the drone's destination at Stage 2 which can be determined using (7.64). Subsequently, we can find the switching time by $\tau_5 = \tau_4 + t_{s2}$.

The switching times in Stages 3-5 can be determined by adopting the similar approach used in Stages 1 and 2. Note that, τ_{14} represents the total control time the drone, which can be determined based on (7.73) and (7.79) as follows:

$$T_{I,D} = \tau_{14} = \sqrt{2 d_D \Big(\frac{m_D}{A_{s2}} - \frac{m_D}{A_{s4}} \Big)} + T^O, \tag{7.81}$$

where A_{s4} is the total force on the drone as Stage 4. T^O is the total control time needed for the orientation changes in Stages 1,3, and 5, given by:

$$T^O = \frac{2}{v_{\max}} \Big[\sqrt{\frac{\Delta\psi_{\mathrm{p},1} I_y}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{r},1} I_x}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{p},3} I_y}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{r},3} I_x}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{p},5} I_y}{l\rho_1}} + \sqrt{\frac{\Delta\psi_{\mathrm{r},5} I_x}{l\rho_1}} \Big].$$

where $\Delta\psi_{\mathrm{p},i}$, $\Delta\psi_{\mathrm{r},i}$ are the pitch and roll changes in Stage i. This completes the proof.

Chapter 8

Optimal Transport Theory for Cell Association in UAV-Enabled Cellular Networks

8.1 Background, Related Works, and Contributions

The deployment of UAVs is an effective technique for improving the quality-of-service (QoS) of wireless cellular networks due to their inherent ability to create line-of-sight (LoS) communication links [5, 7, 11, 15, 61, 203]. One of important challenges in UAV-based communications is cell (or user) association. In [16], the authors analyzed the user-UAV assignment for capacity enhancement of heterogeneous networks. However, this work is limited to the case in which users are uniformly distributed within a geographical area. In [127], the authors proposed a power-efficient cell association scheme while satisfying the rate requirement of users in cellular networks. However, in [127], the authors do not consider the presence of UAVs and their objective function does not account for network delay. In [125], the optimal deployment and cell association of UAVs are determined with the goal of minimizing the UAVs' transmit power while satisfying the users' rate requirements. However, the work in [125] mainly focused on the optimal deployment of the UAVs and does not analyze the existence and characterization of the cell association problem. Therefore, our work is different from [125] in terms of the system model, the objective function, the problem formulation as well as analytical results. In fact, none of the previous studies in [5,7,11,15,16,61,125,127,203], addressed the delay-optimal cell association problem considering both UAVs and terrestrial base stations, for any arbitrary distribution of users.

The main contribution of this chapter is to introduce a novel framework for delay-optimal cell association in a cellular network in which both UAVs and terrestrial BSs co-exist. In particular, given the locations of the UAVs and terrestrial BSs as well as any general spatial distribution of users, we find the optimal cell association by exploiting the framework of *optimal transport theory*

[100]. Within the framework of optimal transport theory, one can address cell association problems for any general spatial distribution of users. In fact, the main advantage of optimal transport theory is to provide tractable solutions for a variety of cell association problems in wireless networks. In our problem, we first prove the existence of the optimal solution to the cell association problem, and, then, we characterize the solution space. The results show that, our approach results in a significantly lower delay compared to a conventional signal strength-based association.

8.2 System Model and Problem formulation

Consider a geographical area $\mathcal{D} \subset \mathbb{R}^2$ in which K terrestrial BSs in set $\mathcal{K}$ are deployed to provide service for ground users that are spatially distributed according to a distribution $f(x, y)$ over the two-dimensional plane. In addition to the terrestrial BSs, M UAVs in set $\mathcal{M}$ are deployed as aerial base stations to enhance the capacity of the network. We consider a downlink scenario in which the BSs and the UAVs use a frequency division multiple access (FDMA) technique to service the ground users. The locations of BS $i \in \mathcal{K}$ and UAV $j \in \mathcal{M}$ are, respectively, given by (x_i, y_i, h_i) and $(x_j^{\text{uav}}, y_j^{\text{uav}}, h_j^{\text{uav}})$, with h_i and h_j^{uav} being the heights of BS i and UAV j. The maximum transmit powers of BS i and UAV j are P_i and P_j^{uav}. Let W_i and W_j be the total bandwidth available for each BS i and UAV j. Our performance metric is the *transmission delay*, which is referred to as the time needed for transmitting a given number of bits. In this case, the delay is inversely proportional to the transmission rate. We use A_i and B_j to denote, respectively, the area (cell) partitions in which the ground users are assigned to BS i and UAV j. Hence, the geographical area is divided into $M + K$ disjoint partitions each of which is served by one of the BSs or the UAVs.

Given this model, our goal is to minimize the average network delay by optimal partitioning of the area. Based on the spatial distribution of the users, we determine the optimal cell associations to minimize the average network delay. Note that, the network delay significantly depends on the cell partitions due to the following reasons. First, the cell partitions determine the service area of each UAV and BS thus impacting the channel gain that each user experiences. Second, the number of users in each partition depends on the cell partitioning. In this case, since the total bandwidth is limited, the amount of bandwidth per user decreases as the number of users in a cell partition increases. Thus, users in the crowded cell partitions achieve a lower throughput which results in a higher delay. Next, we present the channel models.

8.2.1 UAV-User and BS-User path loss models

In UAV-to-ground communications, the probability of having LoS links to users depends on the locations, heights, and the number of obstacles, as well as the elevation angle between a given UAV and it's served ground user. In our model, we consider a commonly used probabilistic path loss model provided by International Telecommunication Union (ITU-R), and the work in [15]. The path loss between UAV j and a user located at (x, y) is [15]:

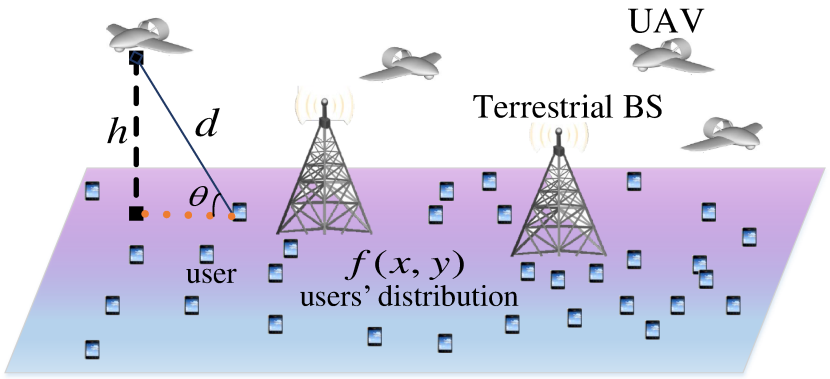


Figure 8.1: Network model.

$$\Lambda_j = \begin{cases} K_o^2 (d_j/d_o)^2 \mu_{\text{LoS}}, & \text{LoS link,} \\ K_o^2 (d_j/d_o)^2 \mu_{\text{NLoS}}, & \text{NLoS link,} \end{cases} \tag{8.1}$$

where $K_o = \left(\frac{4\pi f_c d_o}{c}\right)^2$, f_c is the carrier frequency, c is the speed of light, and d_o is the free-space reference distance. Also, μ_{LoS} and μ_{NLoS} are different attenuation factors considered for LoS and NLoS links. $d_j = \sqrt{(x - x_j^{\text{uav}})^2 + (y - y_j^{\text{uav}})^2 + h_j^{\text{uav}2}}$ is the distance between UAV j and an arbitrary ground user located at (x, y). For the UAV-user link, the LoS probability is [15]:

$$\mathbb{P}_{\text{LoS},j} = \alpha \left(\frac{180}{\pi}\theta_j - 15\right)^{\gamma}, \quad \theta_j > \frac{\pi}{12}, \tag{8.2}$$

where $\theta_j = \sin^{-1}(\frac{h_j}{d_j})$ is the elevation angle (in radians) between the UAV and the ground user. Also, α and γ are constant values reflecting the environment impact. Note that, the NLoS probability is $\mathbb{P}_{\text{NLoS},j} = 1 - \mathbb{P}_{\text{LoS},j}$.

Considering $d_o = 1$ m, the average path loss is $K_o {d_j}^2 [\mathbb{P}_{\text{LoS},j}\mu_{\text{LoS}} + \mathbb{P}_{\text{NLoS},j}\mu_{\text{NLoS}}]$. Therefore, the received signal power from UAV j considering an equal power allocation among its associated users will be:

$$\bar{P}_{r,j}^{\text{uav}} = P_j^{\text{uav}} / \left(N_j^{\text{uav}} K_o {d_j}^2 [\mathbb{P}_{\text{LoS},j}\mu_{\text{LoS}} + \mathbb{P}_{\text{NLoS},j}\mu_{\text{NLoS}}]\right), \tag{8.3}$$

where P_j^{uav} is the UAV's total transmit power, and $N_j^{\text{uav}} = N \iint_{B_j} f(x, y) \mathrm{d}x\mathrm{d}y$ is the average number of users associated with UAV j, with N being the total number of users. For the BS-user link, we use the traditional path loss model. In this case, the received signal power from BS i at user's location (x, y) will be:

$$P_{r,i} = P_i K_o^{-1} d_i^{-n} / N_i, \tag{8.4}$$

where d_i is the distance between BS i and a given user, $N_i = N \iint_{A_i} f(x, y) \mathrm{d}x\mathrm{d}y$ is the average number of users associated with BS i, and n is the path loss exponent.

8.2.2 Problem formulation

Given the average received signal power in the UAV-user communication, the average throughput of a user located at (x, y) connecting to a UAV j can be approximated by:

$$C_j^{\text{uav}} = \frac{W_j}{N_j^{\text{uav}}} \log_2\big(1 + \frac{\bar{P}_{r,j}^{\text{uav}}}{\sigma^2}\big), \tag{8.5}$$

where σ^2 is the noise power for each user which is linearly proportional to the bandwidth allocated to the user.

The throughput of the user if it connects to a BS i is:

$$C_i = \frac{W_i}{N_i} \log_2\big(1 + \frac{P_{r,i}}{\sigma^2}\big). \tag{8.6}$$

Now, let $\mathcal{L} = \mathcal{K} \cup \mathcal{M}$ be the set of all BSs and UAVs. Also, here, the location of each BS or UAV is denoted by $\boldsymbol{s}_k$, $k \in \mathcal{L}$. We also consider $D_k = \begin{cases} A_k, & \text{if } k \in \mathcal{K}, \\ B_k, & \text{if } k \in \mathcal{M}, \end{cases}$ denoting all the cell partitions, and $Q\left(\boldsymbol{v}, \boldsymbol{s}_k, D_k\right) = \begin{cases} b/C_k, & \text{if } k \in \mathcal{K}, \\ b/C_k^{\text{uav}}, & \text{if } k \in \mathcal{M}, \end{cases}$ where $\boldsymbol{v} = (x, y)$ is the 2D locations of the ground users, and b is the number of bits that must be transmitted to location $\boldsymbol{v}$. Then, our optimization problem that seeks to minimize the average network delay over the entire area will be:

$$\min_{D_k} \sum_{k \in \mathcal{L}} \int_{D_k} Q\left(\boldsymbol{v}, \boldsymbol{s}_k, D_k\right) f(x, y) \mathrm{d}x \mathrm{d}y, \tag{8.7}$$

$$\text{s.t.} \bigcup_{k \in \mathcal{L}} D_k = \mathcal{D}, \;\; D_l \cap D_m = \emptyset, \;\; \forall l \neq m \in \mathcal{L}. \tag{8.8}$$

where both constraints in (8.8) guarantee that the cell partitions are disjoint and their union covers the entire area, $\mathcal{D}$.

8.3 Optimal Transport Theory for Cell Association

Given the locations of the BSs and the UAVs as well as the distribution of the ground users, we find the optimal cell association for which the average delay of the network is minimized. Let $g_k(z) = \frac{Nz}{W_k}$, with W_k being the bandwidth for each BS or UAV k and z is a generic argument. Also, we consider:

$$F(\boldsymbol{v}, \boldsymbol{s}_k) = \begin{cases} b/\log_2\left(1 + P_{r,k}(\boldsymbol{v}, \boldsymbol{s}_k)/\sigma^2\right), & \text{if } k \in \mathcal{K}, \\ b/\log_2\left(1 + \bar{P}_{r,j}^{\text{uav}}(\boldsymbol{v}, \boldsymbol{s}_k)/\sigma^2\right), & \text{if } k \in \mathcal{M}. \end{cases} \tag{8.9}$$

Now, the optimization problem in (8.7) can be rewritten as:

$$\min_{D_k} \sum_{k\in\mathcal{L}} \int_{D_k} \left[g_k \left(\int_{D_k} f(x,y)\mathrm{d}x\mathrm{d}y \right) F(\boldsymbol{v}, \boldsymbol{s}_k) \right] f(x,y)\mathrm{d}x\mathrm{d}y, \tag{8.10}$$

$$\text{s.t.} \bigcup_{k\in\mathcal{L}} D_k = \mathcal{D}, \;\; D_l \cap D_m = \emptyset, \;\; \forall l \neq m \in \mathcal{L}, \tag{8.11}$$

where D_k is the cell partition of each BS or UAV k.

Solving the optimization problem in (8.10) is challenging and intractable due to various reasons. First, the optimization variables $D_k, \forall k \in \mathcal{L}$, are sets of continuous partitions which are mutually dependent. Second, $f(x,y)$ can be any generic function of x and y that leads to the complexity of the given two-fold integrations. To overcome these challenges, next, we model this problem by exploiting *optimal transport theory* [100] in order to characterize the solution.

Optimal transport theory [100] allows analyzing complex problems in which, for two probability measures f_1 and f_2 on $\Omega \subset \mathbb{R}^n$, one must find the optimal transport map T from f_1 to f_2 that minimizes the following function:

$$\min_{T} \int_{\Omega} c\left(x, T(x)\right) f_1(x)\mathrm{d}x; \;\; T : \Omega \to \Omega, \tag{8.12}$$

where $c(x, T(x))$ denotes the cost of transporting a unit mass from a location x to a location $T(x)$.

Our cell association problem can be modeled as a semi-discrete optimal transport problem. In this case, the users follow a continuous distribution, and the base stations can be considered as discrete points. Then, we need to map the users to the BSs and UAVs such that the total cost function is minimized. In this case, the optimal cell partitions are directly determined by the optimal transport map [134]. Next, we prove the existence of the optimal solution to the problem in (8.10).

Theorem 16. The optimization problem in (8.10) admits an optimal solution given $N \neq 0$, and $\sigma \neq 0$.

Proof. Let $a_k = \int_{D_k} f(x,y)\mathrm{d}x\mathrm{d}y$, and for $\forall k \in \mathcal{L}$,
$E=\left\{ \boldsymbol{a} = (a_1, a_2, ..., a_{K+M}) \in \mathbb{R}^{K+M}; a_k \geq 0, \sum_{k=1}^{K+M} a_k = 1 \right\}$. Now, considering $f(x,y) = f(\boldsymbol{v})$ and $c\left(\boldsymbol{v}, \boldsymbol{s}_k\right) = g_k(a_k) F\left(\boldsymbol{v}, \boldsymbol{s}_k\right)$, for any given vector $\boldsymbol{a}$, problem (8.10) can be considered as a classical semi-discrete optimal transport problem. First, we prove that $c\left(\boldsymbol{v}, \boldsymbol{s}\right)$ is a semi-continuous function. Considering the fact that $\boldsymbol{s}_k$ is discrete, we have: $\lim_{(\boldsymbol{v},\boldsymbol{s})\to(\boldsymbol{v}^*,\boldsymbol{s}_k)} F\left(\boldsymbol{v}, \boldsymbol{s}\right) = \lim_{\boldsymbol{v}\to\boldsymbol{v}^*} F\left(\boldsymbol{v}, \boldsymbol{s}_k\right)$. Note that, given any $\boldsymbol{s}_k$, k belongs to only of $\mathcal{K}$ and $\mathcal{M}$ sets. Given $\boldsymbol{s}_k$, $F(\boldsymbol{v}, \boldsymbol{s}_k)$ is a continuous function of $\boldsymbol{v}$. Then, considering the fact that given a_k, $g_k(a_k)$ is constant, we have $\lim_{(\boldsymbol{v},\boldsymbol{s})\to(\boldsymbol{v}^*,\boldsymbol{s}_k)} g_k(a_k) F\left(\boldsymbol{v}, \boldsymbol{s}\right) = g_k(a_k) F\left(\boldsymbol{v}^*, \boldsymbol{s}_k\right)$. Therefore, $c(\boldsymbol{v}, \boldsymbol{s})$ is a continuous function and, hence, is also a lower semi-continuous function. Now, we use the following lemma from optimal transport theory:

Lemma 3. Consider two probability measures f and λ on $\mathcal{D} \subset \mathbb{R}^n$. Let f be continuous and $\lambda = \sum_{k\in\mathbb{N}} a_k \delta_{\boldsymbol{s}_k}$ be a discrete probability measure. Then, for any lower semi-continuous cost function,

there exists an optimal transport map from f to λ for which $\int_{\mathcal{D}} c\left(x, T(x)\right) f(x) \mathrm{d}x$ is minimized [134].

Considering Lemma 1, for any $\boldsymbol{a} \in E$, the problem in (8.10) admits an optimal solution. Since E is a unit simplex in $\mathbb{R}^{M+K}$ which is a non-empty and compact set, the problem admits an optimal solution over the entire E. □

Next, we characterize the solution space of (8.10).

Theorem 17. To acheive the delay-optimal cell partitions in (8.10), each user located at (x, y) will satisfy the following:

$$k = \arg\min_{l \in \mathcal{L}} \big\{ \frac{N}{W_l} \int_{D_l} F(\boldsymbol{v}, \boldsymbol{s}_l) f(x,y) \mathrm{d}x\mathrm{d}y + \frac{N a_l}{W_l} F(\boldsymbol{v}_o, \boldsymbol{s}_l) \big\}, \tag{8.13}$$

Given (8.13), the optimal cell partition D_k includes all the points which are assigned to BS (or UAV) k.

Proof. As proved in Theorem 1, there exist optimal cell partitions D_k, $k \in \mathcal{L}$ which are the solutions to (8.10). Now, consider two partitions D_l and D_m, and a point $\boldsymbol{v}_o = (x_o, y_o) \in D_l$. Also, let $B_\epsilon(\boldsymbol{v}_o)$ be a ball with a center $\boldsymbol{v}_o$ and radius $\epsilon > 0$. Now, we generate the following new cell partitions $\widehat{D}_k$ (which are variants of the optimal partitions):

$$\begin{cases} \widehat{D}_l = D_l \backslash B_\varepsilon(\boldsymbol{v}_o), \\ \widehat{D}_m = D_m \cup B_\varepsilon(\boldsymbol{v}_o), \\ \widehat{D}_k = D_k, \quad k \neq l, m. \end{cases} \tag{8.14}$$

Let $a_\varepsilon = \int_{B_\varepsilon(\boldsymbol{v}_o)} f(x,y) \mathrm{d}x\mathrm{d}y$, and $\widehat{a}_k = \int_{\widehat{D}_k} f(x,y) \mathrm{d}x\mathrm{d}y$. Considering the optimality of D_k, $k \in \mathcal{L}$, we have:

$$\begin{aligned} &\sum_{k \in \mathcal{K}} \int_{D_k} g_k\left(a_k\right) F(\boldsymbol{v}, \boldsymbol{s}_k) f(x,y) \mathrm{d}x\mathrm{d}y \\ &\overset{(a)}{\leq} \sum_{k \in \mathcal{K}} \int_{\widehat{D}_k} g_k\left(\widehat{a}_k\right) F(\boldsymbol{v}, \boldsymbol{s}_k) f(x,y) \mathrm{d}x\mathrm{d}y. \end{aligned} \tag{8.15}$$

Now, canceling out the common terms in (9.52) leads to:

$$\begin{aligned} &\int_{D_l} g_l\left(a_l\right) F(\boldsymbol{v}, \boldsymbol{s}_l) f(x,y) \mathrm{d}x\mathrm{d}y + \int_{D_m} g_m\left(a_m\right) F(\boldsymbol{v}, \boldsymbol{s}_m) f(x,y) \mathrm{d}x\mathrm{d}y \\ &\leq \int_{D_m \cup B_\varepsilon(\boldsymbol{v}_o)} g_m\left(a_m + a_\varepsilon\right) F(\boldsymbol{v}, \boldsymbol{s}_m) f(x,y) \mathrm{d}x\mathrm{d}y \\ &+ \int_{D_l \backslash B_\varepsilon(\boldsymbol{v}_o)} g_l\left(a_l - a_\varepsilon\right) F(\boldsymbol{v}, \boldsymbol{s}_l) f(x,y) \mathrm{d}x\mathrm{d}y, \\ &\int_{D_l} \left(g_l\left(a_l\right) - g_l\left(a_l - a_\varepsilon\right)\right) F(\boldsymbol{v}, \boldsymbol{s}_l) f(x,y) \mathrm{d}x\mathrm{d}y \end{aligned}$$

$$+\int_{B_\varepsilon(\boldsymbol{v}_o)} g_l\left(a_l - a_\varepsilon\right) F(\boldsymbol{v}, \boldsymbol{s}_l) f(x,y)\mathrm{d}x\mathrm{d}y$$
$$\leq \int_{D_m} \left(g_m\left(a_m + a_\varepsilon\right) - g_m\left(a_m\right)\right) F(\boldsymbol{v}, \boldsymbol{s}_m) f(x,y)\mathrm{d}x\mathrm{d}y$$
$$+\int_{B_\varepsilon(\boldsymbol{v}_o)} g_m\left(a_m + a_\varepsilon\right) F(\boldsymbol{v}, \boldsymbol{s}_m) f(x,y)\mathrm{d}x\mathrm{d}y, \tag{8.16}$$

where (a) comes from the fact that D_k, $\forall k \in \mathcal{L}$ are optimal and, hence, any variation of such optimal partitions, shown by $\widehat{D}_k$, cannot lead to a better solution. Now, we multiply both sides of the inequality in (9.54) by $\frac{1}{a_\epsilon}$, take the limit when $\epsilon \to 0$, and use the following equalities:

$$\lim_{\varepsilon \to 0} a_\varepsilon = 0, \tag{8.17}$$
$$\lim_{a_\varepsilon \to 0} \frac{g_l(a_l) - g_l(a_l - a_\varepsilon)}{a_\varepsilon} = g_l'(a_l), \tag{8.18}$$
$$\lim_{a_\varepsilon \to 0} \frac{g_m(a_m + a_\varepsilon) - g_m(a_m)}{a_\varepsilon} = g_m'(a_m), \tag{8.19}$$

then we have:

$$g_l'\left(a_l\right) \int_{D_l} F(\boldsymbol{v}, \boldsymbol{s}_l) f(x,y)\mathrm{d}x\mathrm{d}y + g_l\left(a_l\right) F(\boldsymbol{v}_o, \boldsymbol{s}_l)$$
$$\leq g_m'\left(a_m\right) \int_{D_m} F(\boldsymbol{v}, \boldsymbol{s}_m) f(x,y)\mathrm{d}x\mathrm{d}y + g_m\left(a_m\right) F(\boldsymbol{v}_o, \boldsymbol{s}_m). \tag{8.20}$$

Now, given $g_k(z) = \frac{Nz}{W_k}$, we can compute $g_l'(a_l) = \left.\frac{dg_l(z)}{dz}\right|_{z=a_l} = \frac{N}{W_k}$, then, using $a_k = \int_{D_k} f(x,y)\mathrm{d}x\mathrm{d}y$ leads to:

$$\frac{N}{W_l} \int_{D_l} F(\boldsymbol{v}, \boldsymbol{s}_l) f(x,y)\mathrm{d}x\mathrm{d}y + \frac{Na_l}{W_l} F(\boldsymbol{v}_o, \boldsymbol{s}_l)$$
$$\leq \frac{N}{W_m} \int_{D_m} F(\boldsymbol{v}, \boldsymbol{s}_m) f(x,y)\mathrm{d}x\mathrm{d}y + \frac{Na_m}{W_m} F(\boldsymbol{v}_o, \boldsymbol{s}_m), \tag{8.21}$$

□

Theorem 8.13 provides a precise cell association rule for ground users that are distributed following any general distribution $f(x,y)$. In fact, this theorem represents the condition under which the user is assigned to a BS or UAV l. Under the special case of a uniform distribution of the users, the result in Theorem 8.13 leads to the classical SNR-based association in which users are assigned to base stations that provide strongest signal. From Theorem 2, we can see that there is a mutual dependence between a_l and D_l (i.e. cell association), $\forall l \in \mathcal{L}$. To solve the equation given in Theorem 2, we adopt an iterative approach which is shown to converge to the global optimal solution [134]. In this case, we start with initial cell partitions (e.g. Voronoi diagram), and iteratively update the cell partitions based on Theorem 8.13.

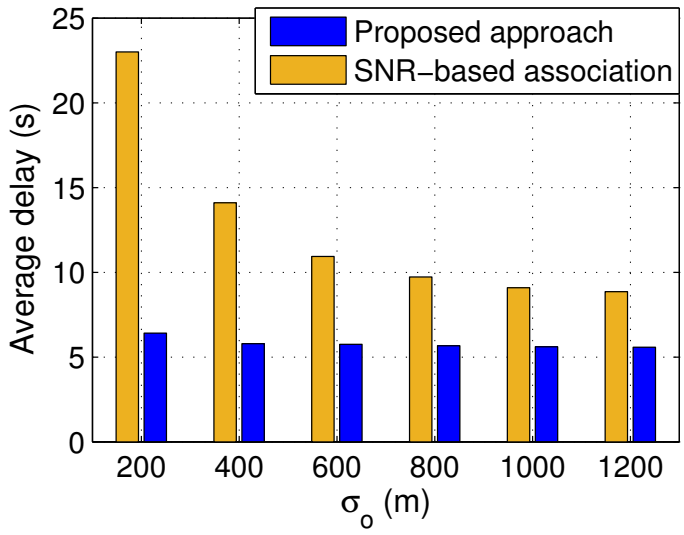
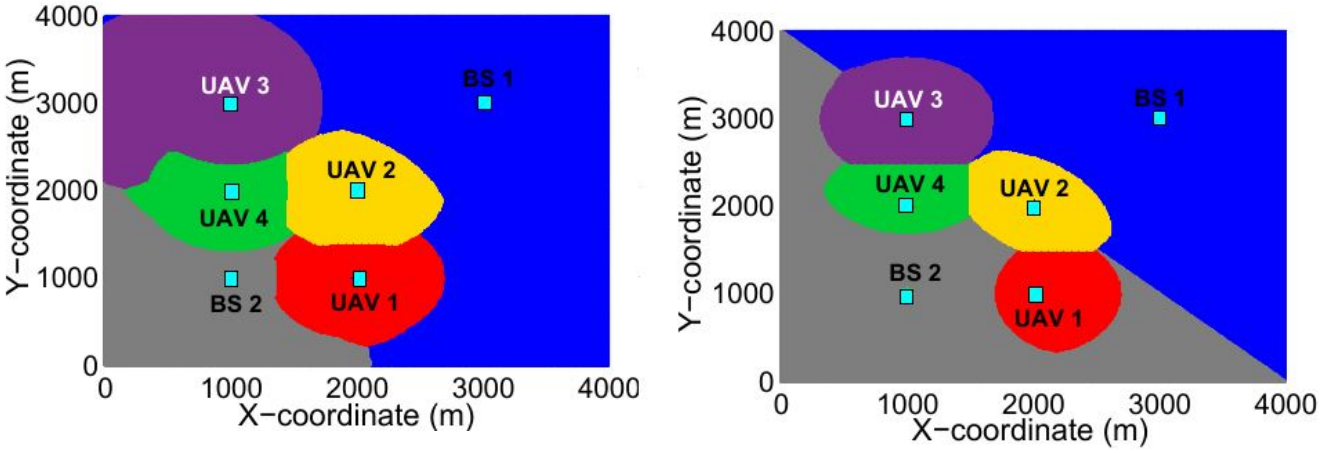


Figure 8.2: Average network delay per 1Mb data transmission.

8.4 Simulation Results and Analysis

For our simulations, we consider an area of size $4\,\text{km} \times 4\,\text{km}$ in which 4 UAVs and 2 macrocell base stations are deployed based on a traditional grid-based deployment. The ground users are distributed according to a truncated Gaussian distribution with a standard deviation σ_o. This type of distribution which is suitable to model a hotspot area. The simulation parameters are given as follows. f_c =2 GHz, transmit power of each BS is 40 W, and transmit power of each UAV is 1 W. Also, $N = 300$, $W_j = W_i = 1\,\text{MHz}$, and the noise power spectral density is -170 dBm/Hz. We consider a dense urban environment with $n = 3$, $\mu_{\text{LoS}} = 3\,\text{dB}$, $\mu_{\text{NLoS}} = 23\,\text{dB}$, $\alpha = 0.36$, and $\gamma = 0.21$ [15]. The heights of each UAV and BS are, respectively, 200 m and 20 m [7, 15, 16]. All statistical results are averaged over a large number of independent runs.

In Fig. 8.2, we compare the delay of the proposed cell association with the traditional SNR-based association. We consider a truncated Gaussian distribution with a center (1300 m,1300 m), and σ_o varying from 200 m to 1200 m. Lower values of σ_o correspond to scenarios in which users are more concentrated around the hotspot center. Fig. 8.2 shows that the proposed cell association significantly outperforms the SNR-based association in terms of the average delay. For low σ_o values, the average delay decreases by 72% compared to the SNR-based association. This is due to the fact that, in the proposed approach, the impact of network congestion is taken into account. Hence, the proposed approach avoids creating highly loaded cells. In contrast, an SNR-based association can yield highly loaded cells. As a result, in the congested cells, each user will receive a low amount of bandwidth that leads a low transmission rate or equivalently high delay. In fact, compared to the SNR-based association case, our approach is more robust against network congestion and its performance is significantly less affected by changing σ_o.



(a) Proposed delay-optimal cell partitions.

(b) SNR-based association.

Figure 8.3: Cell partitions associated to UAVs and BSs given the non-uniform spatial distribution of users.

As an illustrative example, Fig. 8.3 shows the locations of the BSs and UAVs as well as the cell partitions obtained using SNR-based association and the proposed delay-optimal association. In this case, users are distributed based on a 2D truncated Gaussian distribution with mean values of (1300 m,1300 m), and $\sigma_o = 1000$ m. As shown in Fig. 8.3, the size and shape of cells are different in these two association approaches. For instance, the red cell partition in the proposed approach is smaller than the SNR-based case. In fact, the red partition in the SNR-based approach is highly congested and, consequently, its size is reduced in the proposed case so as to decrease the congestion as well as the delay.

8.5 Summary

In this chapter, we have proposed a novel framework for delay-optimal cell association in UAV-enabled cellular networks. In particular, to minimize the average network delay based on the users' distribution, we have exploited optimal transport theory to derive the optimal cell associations for UAVs and terrestrial BSs. Our results have shown that, the proposed cell association approach results in a significantly lower network delay compared to an SNR-based association.

Chapter 9

Beyond 5G with UAVs: Foundations of a 3D Wireless Cellular Network

9.1 Background, Related Works, and Contributions

Recent reports from the federal aviation administration (FAA) show that the number of unmanned aerial vehicles (UAVs), also known as drones, will exceed 7 million in 2020. From a wireless communications perspective, drones will have two key roles: aerial base stations (BSs), and user equipments (UEs) [12, 184, 204]. Due to their aerial nature and flexible operation, drone-BSs can support broadband communications to under-developed areas and provide hotspot wireless coverage during temporary events [11, 12, 184, 205]. Meanwhile, drones can also act as user equipments (i.e., drone-UEs) that must connect to a wireless network so as to operate. In particular, wireless-connected drone-UEs can be used for wide range of applications such as package delivery, remote sensing, and virtual reality.

Wireless networking with drones faces a number of challenges. For instance, for drone-BSs, the key design problems include 3D deployment of drones, performance analysis, resource allocation, and cell association. In the drone-UEs scenario, there is a need for reliable and low latency communications so as to efficiently control drones' operations. However, existing terrestrial cellular networks have been primarily designed for supporting ground users and are not able to effectively serve aerial users. In fact, terrestrial BSs may not be able to meet the low-latency and reliable communication requirements of drone-UEs due to blockage effects and the specific design of BSs' antennas which are not suitable for supporting users at high-elevation angles. Furthermore, in areas with geographical constraints, terrestrial base stations may not be available to provide wireless service to drone-UEs. In such case, the deployment of aerial drone-BSs is a promising opportunity for providing reliable wireless connectivity for drone-UEs. Clearly, to support the key roles of drones in wireless networking applications, there is a need for developing the novel concept of a *3D cellular network* that can incorporate both drone-BSs and drone-UEs.

Recent studies on drone communications have investigated various design challenges that include performance characterization, 3D deployment, user-to-drone association, and cellular-connected UAVs. For instance, in [18], the authors proposed an algorithm for jointly optimizing the locations and number of drones to maximize wireless coverage. The work in [206] studied the optimal 3D deployment of UAVs for maximizing the number of covered ground users under quality-of-service constraints. In [207], the authors proposed a framework for strategic placement of drone-BSs for a large-scale ground network. However, the prior studies on deployment of UAV base stations ignore the existence of flying drone-UEs. In addition, the work in [205] proposed an algorithm for maximizing the sum-rate of ground users by joint optimization of user-to-drone-BSs association and wireless backhaul bandwidth allocations. The work in [189] proposed a novel cell association approach that maximizes the total data delivered to ground users by drone-BSs that have limited flight endurance. However, the previous works on user association problems in drone networks are limited to ground users and do not consider 3D aerial users.

While there exists a number of studies on cellar-connected drone-UEs [208–210], the potential use of drone-BSs for serving drone-UEs has not been considered. For example, in [208], the authors studied the coexistence of drone-UEs and ground users in cellular networks and characterized the downlink coverage performance. The work in [209] proposed an interference-aware path planning approach for drone-UEs with the goal of minimizing their latency and interference. In [210], the authors proposed a trajectory design scheme for minimizing the mission time of a single drone-UE. However, none of these previous works [11, 12, 18, 184, 189, 205–210], studied a 3D wireless network of co-existing aerial base stations and users (i.e., drone-BSs and drone-UEs) while addressing the deployment and latency-aware cell association problems.

The main contribution of this chapter is to introduce the novel concept of a fully-fledged drone-based 3D cellular network that incorporates drone-UEs, low-altitude platform (LAP) drone-BSs, and high-altitude platform (HAP) drones. In this new 3D cellular network architecture, we propose a framework for addressing the two fundamental problems of network planning and 3D cell association. In particular, our proposed framework includes a tractable approach for three-dimensional placement and frequency planning for drone-BSs, as well as a latency-minimal 3D cell association scheme for servicing drone-UEs. For deployment, we introduce a new approach based on truncated octahedron cells that determines the minimum number of drone-BSs that can cover a 3D space, along with their locations. Furthermore, for frequency planning in the proposed 3D wireless network, we derive an analytical expression for the feasible integer frequency reuse factors. To perform latency-minimal 3D cell association, first, we estimate the spatial distribution of drone-UEs by using a kernel density estimation method. Then, given the locations of drone-BSs and the distribution of drone-UEs, we find the optimal 3D cell association for which the total latency of serving drone-UEs is minimized. In this case, we analytically characterize the optimal 3D cell partitions by exploiting tools from optimal transport theory. Our results show that the proposed approach significantly reduces the latency of serving drone-UEs, compared to classical cell association approach that uses signal-to-interference-plus-noise ratio (SINR) criterion. In particular, our approach yields around 46% reduction in the average total latency compared to the SINR-bases association. The results also reveal that our latency-optimal cell association improves

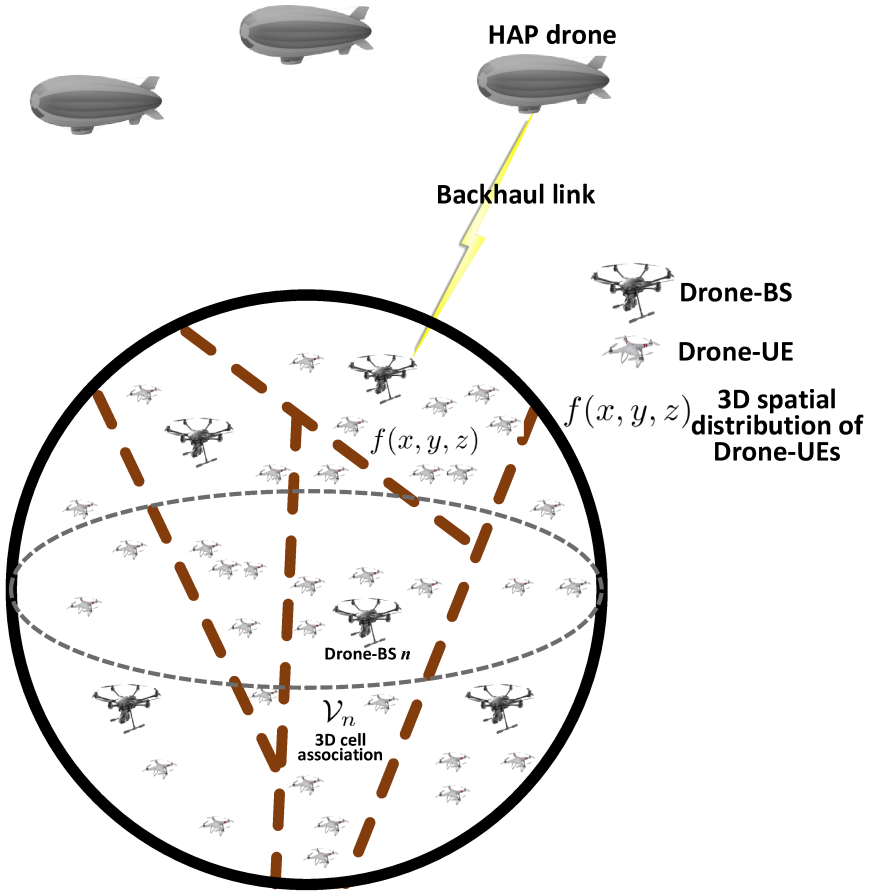


Figure 9.1: The proposed 3D wireless network with drone-BSs, drone-UEs, and HAP drones.

spectral efficiency of the considered 3D wireless network with drones.

The rest of this chapter is organized as follows. In Section 9.2, we present the system model. In Section 9.3, the three-dimensional placement of drone-BSs is investigated. In Section 9.4, we describe our approach for estimating the spatial distribution of drone-UEs. Section 9.5 presents the proposed latency-optimal cell association scheme. Simulation results are provided in Section 9.6 and conclusions are drawn in Section 9.7.

9.2 System Model

Consider a 3D cellular network composed of L drone users, N LAP drone base stations, and a number of HAP drones, as shown in Figure 9.1. We represent the sets of drone-UEs, and drone-BSs, respectively, by $\mathcal{L}$, and $\mathcal{N}$. In this aerial network, drone-BSs serve drone-UEs in the downlink, and *HAP drones provide a wireless backhaul connectivity* [211] for drone-BSs. The key advantage of HAP drones is their ability to adjust their positions according to the locations of drone-BSs.

In addition, due to their high altitudes, HAPs can establish LoS backhaul links to the drone-BSs. Here, the deployment of drone-BSs is performed based on a 3D cellular structure which will be presented in Section III. For backhaul connectivity, we assume that each drone-BS connects to its closest HAP that can provide a maximum rate. We denote the backhaul transmission rate for drone-BS n by C_n, which is assumed to be given in our model[1]. Drone-BS n uses transmit power P_n bandwidth B_n in order to serve its associated flying drone-UEs. Let $f(x, y, z)$ be the spatial probability density function of drone-UEs which represents the probability that each drone-UE is present around a 3D location (x, y, z). In our model, drone-BSs use machine learning tools to estimate the spatial probability distribution of drone-UEs, for a certain period of time, based on any available prior information about the drone-UEs' locations. By performing such estimation, the network will no longer need to continuously track the locations of flying drone-UEs thus alleviating the associated overhead. To find the 3D cell association when serving drone-UEs, we partition the space into N 3D cells each of which representing a volume that must be serviced by one drone-BS. Let $\mathcal{V}_n$ be a 3D space (i.e., 3D cell) associated to drone-BS n that serves drone-UEs located within this cell. The average number of drone-UEs inside $\mathcal{V}_n$ is given by:

$$K_n = L \int_{\mathcal{V}_n} f(x, y, z) \mathrm{d}x \mathrm{d}y \mathrm{d}z. \tag{9.1}$$

We assume that each drone-BS adopts a frequency division multiple access (FDMA) technique (as done in [189] and [212]) when servicing its associated drone-UEs. Hence, the average downlink transmission rate from a drone-BS n to a drone-UE located at (x, y, z) is:

$$R_n(x, y, z) = \frac{B_n}{K_n} \log_2 \big(1 + \gamma_n(x, y, z)\big), \tag{9.2}$$

where $\frac{B_n}{K_n}$ is the amount of bandwidth for servicing each drone-UE located in $\mathcal{V}_n$, which is determined by sharing the total bandwidth among the drone-UEs. $\gamma_n(x, y, z)$ is the SINR for a drone-UE located at (x, y, z) served by drone-BS n.

We consider the *average latency* in servicing drone-UEs as our main performance metric. In particular, we consider transmission latency in drone-BSs to drone-UEs communications, backhaul latency in drone-BSs to HAP drones links, and computation latency for drone-BSs that serve drone-UEs. The transmission latency for a drone-UE located at (x, y, z) which is served by drone-BS n is:

$$\tau_n^{\mathrm{Tr}}(x, y, z, K_n) = \frac{\beta}{R_n(x, y, z)}, \tag{9.3}$$

where β is the number of bits per packet that must be transmitted to each drone-UE.

The backhaul latency depends on the load of drone-BSs and the backhaul transmission rates. In this case, the average backhaul latency in drone-BS n to its corresponding HAP-drone communications

[1]The backhaul rate can be easily calculated based on the locations of HAPs and drone-BSs, transmit power of HAPs, and bandwidth of backhaul links.

is given by:

$$\tau_n^{\mathrm{B}}(K_n) = \frac{\beta L \int_{\mathcal{V}_n} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z}{C_n} = \frac{\beta K_n}{C_n}, \tag{9.4}$$

where C_n is the maximum backhaul transmission rate for drone-BS n, and $\beta L \int_{\mathcal{V}_n} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$ represents the average load on drone-BS n.

The computation time depends on the data size (i.e., load) that must be processed in each drone-BS, and the processing speed. To model the computational latency at drone-BS n, we use function $g_n(\beta K_n)$ with βK_n being the total data size that must be processed at the drone-BS. Therefore, the total latency experienced by any arbitrary drone-UE located at (x, y, z) while being served by drone-BS n can be given by:

$$\tau_n^{\mathrm{tot}}(x,y,z,K_n) = \tau_n^{\mathrm{Tr}}(x,y,z,K_n) + \tau_n^{\mathrm{B}}(K_n) + g_n(\beta K_n), \tag{9.5}$$

Given this model, our goal is to minimize the average latency of drone-UEs by finding the optimal 3D cell association in drone-BSs to drone-UEs communications. In particular, given the locations of drone-BSs which are deployed based on a 3D cellular structure (in Section III), and the estimated spatial distribution of drone-UEs (in Section IV), we determine the optimal 3D cell partitions $\mathcal{V}_n$, $\forall n \in \mathcal{N}$ that lead to a minimum average latency for drone-UEs. In this regard, our 3D cell association optimization problem can be posed as follows:

$$\min_{\mathcal{V}_1,\ldots,\mathcal{V}_N} \sum_{n=1}^{N} \left[\int_{\mathcal{V}_n} \tau_n^{\mathrm{Tr}}\big(x,y,z,K_n\big) f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + \tau_n^{\mathrm{B}}(K_n) + g_n(\beta K_n) \right], \tag{9.6}$$

$$\text{s.t. } \mathcal{V}_l \cap \mathcal{V}_m = \emptyset, \ \ \forall l \neq m \in \mathcal{N}, \tag{9.7}$$

$$\bigcup_{n\in\mathcal{N}} \mathcal{V}_n = \mathcal{V}, \tag{9.8}$$

where $K_n = L \int_{\mathcal{V}_n} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$ is the average number of drone-UEs in $\mathcal{V}_n$ which depends on the 3D cell association, and $\mathcal{V}$ is the entire considered space in which drone-UEs can fly. The constraints in (9.7) and (9.8) ensure that the 3D association spaces are disjoint and their union covers the considered space $\mathcal{V}$.

In Figure 9.2, we summarize the key steps for developing our proposed drone-based 3D cellular network architecture. First, we plan the network deployment of drone-BSs based on a truncated octahedron scheme that can ensure full coverage with a minimum number of drone-BSs. Second, using some available information about the locations' history of drone-UEs, we estimate the 3D spatial distribution of the drone-UEs for a given period of time. Finally, given the locations of drone-BSs and the spatial distribution of drone-UEs, we derive an optimal 3D cell association rule for which the latency of servicing drone-UEs is minimized.

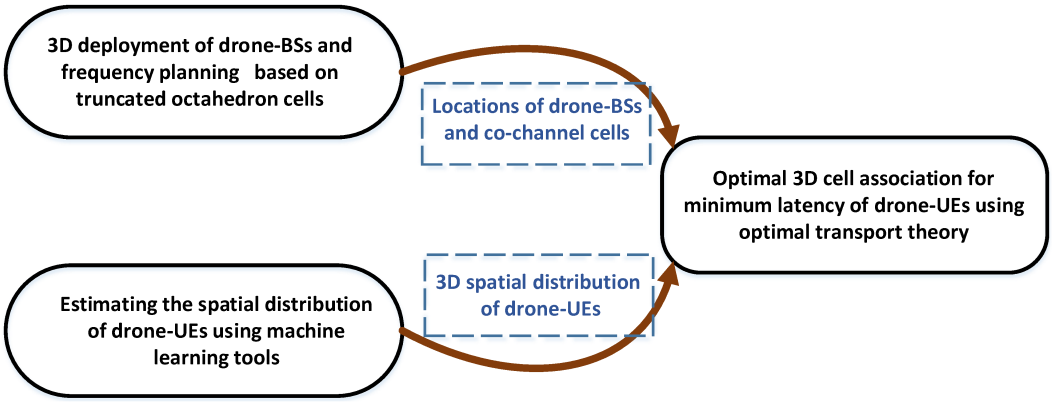


Figure 9.2: Our proposed framework for designing the 3D cellular network.

9.3 Three-dimensional Network Planning of Drone-BSs: A Truncated Octahedron Structure

To perform 3D network planning, we propose a framework for the 3D deployment of drone-BSs and associated frequency planning. In particular, we use the notion of truncated octahedron structure to determine the drone-BSs' locations as well as the feasible integer frequency factors that allow finding co-channel interfering drone-BSs.

In traditional ground cellular networks, hexagonal cell shapes are used while deploying base stations. This is due to the fact that, a 2D space can be fully covered (i.e., without any gaps) by non-overlapping hexagons. While triangle and square cells are also able to tessellate a 2D area, the hexagonal cell is preferred in cellular wireless network planning due to the following reasons. First, the hexagonal shape has a larger area than the square and the triangle, hence less cells will be needed to cover a geographical area. Second, the hexagonal cell reasonably approximates the circular radiation pattern of an omni-directional antenna base station.

Inspired by 2D cellular networks, we propose a framework for the deployment of a 3D cellular network. In three dimensions, the regular polyhedron geometric shapes that can tessellate the space (i.e., fill the 3D space entirely) include cube, hexagonal prism, rhombic dodecahedron, and truncated octahedron [213]. Among these 3D shapes, the truncated octahedron is the closest approximation of a sphere. Moreover, the number of polyhedron required for completely covering a 3D space is minimized by adopting the truncated octahedron [213]. Therefore, in our model, we use the truncated octahedron structure for deploying the drone-BSs.

The truncated octahedron is a polyhedron in three dimensions with regular polygons faces. As we can see from Fig. 9.3, the truncated octahedron has 14 faces with 8 regular hexagonal and 6 square, 24 vertices, and 36 edges [214]. The key feature of the truncated octahedron is that it can tessellate the three-dimensional Euclidean space. In other words, the 3D space can be completely filled with multiple copies of the truncated octahedron without any overlap. We exploit this feature of the

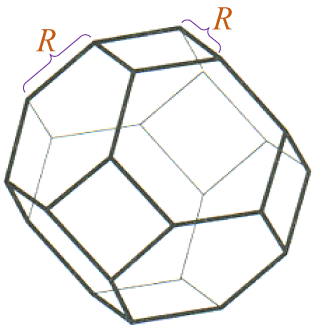


Figure 9.3: Truncated octahedron in 3D.

Figure 9.4: Deployment of drone-BSs based on truncated octahedron cells.

truncated octahedron in our 3D cellular network deployment with drone-BSs.

The deployment of drone-BSs needs to be done such that the entire desired space is covered. To this end, we first completely fill the given space with an arrangement of multiple truncated octahedron cells. Then, we place each drone-BS at the center of each truncated octahedron, as shown in Fig. 9.4 as an illustrative example. Our proposed deployment approach can ensure full coverage for a given 3D space and is also easy to implement and tractable. Moreover, our approach facilitates frequency planning in 3D cellular networks by deriving analytical expressions for the feasible integer reuse factors. Next, we determine the locations of drone-BSs based on the proposed truncated octahedron cell structure.

Theorem 18. The three-dimensional locations of drone-BSs in the proposed 3D cellular network are given by:

$$\boldsymbol{P}_{\{a,b,c\}} = \big[x_o, y_o, z_o\big] + \sqrt{2}R\Big[a+b-c, -a+b+c, a-b+c\Big], \tag{9.9}$$

where a, b, c are integers chosen from set $\{..., -2, -1, 0, 1, 2, ...\}$, and R is the edge length of the considered truncated octahedrons. $[x_o, y_o, z_o]$ is the Cartesian coordinates of a given reference location (e.g., center of a specified space).

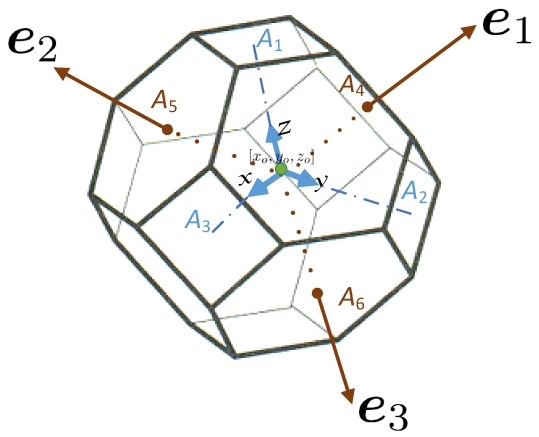


Figure 9.5: Coordinate systems in drone-BSs deployment.

Proof. For the deployment of drone-BSs, we first create a 3D lattice of truncated octahedrons and then, place each drone-BS at the center of each truncated octahedron. Hence, to determine the locations of drone-BSs, we need to find the center of truncated octahedrons.

Let $[x_o, y_o, z_o]$ be the center of the first truncated octahedron in Cartesian coordinates with the x, y, and z directions being perpendicular to square faces A_3, A_2, and A_1 as shown in Figure 9.5. We find a new coordinate system whose integer coordinates are the center of the truncated octahedrons. By moving, in integer value steps, along the axes of this coordinate system, we can reach the center of the truncated octahedrons. We consider a coordinate system whose axes ($\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$) are vertically outward the hexagonal faces, A_4, A_5, and A_6. Now, we find the Euclidean length of each unit axis of this coordinate system. The distance between the center of the truncated octahedron (e.g., $[x_o, y_o, z_o]$) and each hexagonal face is $R\sqrt{6}/2$ [214]. Therefore, the distance from $[x_o, y_o, z_o]$ to the center of an adjacent truncated octahedron connecting to face A_4 is $R\sqrt{6}$. As a result, each unit on axis $\boldsymbol{e}_1$ (and also $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$) must be $2R\sqrt{6}$. It can be easily verified that the centers of the truncated octahedrons in the 3D lattice are the integer coordinates of the ($\boldsymbol{e}_1$,$\boldsymbol{e}_2$,$\boldsymbol{e}_3$) coordinate system. Hence, the 3D location of each drone-BS can be represented by a triple (a, b, c) with a, b, and c being integers. The position of a drone-BS obtained by $\{a, b, c\}$ is given by:

$$\boldsymbol{P}_{\{a,b,c\}} = a\boldsymbol{e}_1 + b\boldsymbol{e}_2 + c\boldsymbol{e}_3. \tag{9.10}$$

Now, we need to represent $\boldsymbol{P}_{\{a,b,c\}}$ using Cartesian coordinates. To this end, we find the projection of $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$ on the x, y, and z axes. With some geometric calculations and using the fact that the dihedral angle (i.e., angle between two intersecting planes) between the adjacent square face and hexagonal face is $\cos^{-1}(\frac{-1}{\sqrt{3}})$ [214], we obtain:

$$\begin{cases} \boldsymbol{e}_1 = R\sqrt{6}\big(\frac{-1}{\sqrt{3}}\boldsymbol{x} + \frac{1}{\sqrt{3}}\boldsymbol{y} + \frac{1}{\sqrt{3}}\boldsymbol{z}\big), \\ \boldsymbol{e}_2 = R\sqrt{6}\big(\frac{1}{\sqrt{3}}\boldsymbol{x} + \frac{-1}{\sqrt{3}}\boldsymbol{y} + \frac{1}{\sqrt{3}}\boldsymbol{z}\big), \\ \boldsymbol{e}_3 = R\sqrt{6}\big(\frac{1}{\sqrt{3}}\boldsymbol{x} + \frac{1}{\sqrt{3}}\boldsymbol{y} + \frac{-1}{\sqrt{3}}\boldsymbol{z}\big). \end{cases} \tag{9.11}$$

Finally, using (9.10) and (9.11), the 3D locations of drone-BSs in Cartesian coordinates, with

respect to the reference position $[x_o, y_o, z_o]$ are given by:

$$\boldsymbol{P}_{\{a,b,c\}} = \big[x_o, y_o, z_o\big] + \sqrt{2}R\Big[a+b-c, -a+b+c, a-b+c\Big], \tag{9.12}$$

which proves the theorem. □

Using Theorem 18, we can find the 3D coordinates of drone-BSs which are deployed at the centers of truncated octahedrons. Moreover, as shown next, Theorem 18 allows us to determine the frequency reuse factor as well as interfering drone-BSs in the proposed 3D cellular network.

Theorem 19. In the considered 3D cellular network, any feasible integer frequency reuse factors can be determined by solving the following equations:

$$\begin{cases} q = \sqrt{\dfrac{\big[3(n_1^2+n_2^2+n_3^2) - 2(n_1n_2+n_1n_3+n_2n_3)\big]^3}{27}}, \\ q = \sqrt{\dfrac{\big[3(m_1^2+m_2^2+m_3^2) - 2(m_1m_2+m_1m_3+m_2m_3)\big]^3}{64}}, \end{cases} \tag{9.13}$$

where q is the frequency reuse factor which is a positive integer. n_1, n_2, n_3, m_1, m_2, and m_3 are integers that satisfy (9.13) by generating feasible frequency reuse factors.

Proof. We consider a truncated octahedron cell with 14 faces, as a reference cell. In this case, the number of first tier co-channel interfering cells is 14. Since the distance between centers of the reference cell and its adjacent cell is varies depending on the connecting face (i.e., hexagonal or square face), we consider two different co-channel distances (i.e., reuse distances). Let D_u and D_l be two different reuse distances to different interfering cells.

Assume that the center of a co-channel cell at a distance D_l is located at a positive integer coordinate (n_1, n_2, n_3) in our defined coordinate system $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. Now, using (9.9) in Theorem 18 leads to:

$$\begin{aligned} D_l &= \sqrt{2}R\sqrt{(n_1+n_2-n_3)^2 + (-n_1+n_2+n_3)^2 + (n_1-n_2+n_3)^2} \\ &\overset{(a)}{=} R\sqrt{6(n_1^2+n_2^2+n_3^2) - 4(n_1n_2+n_1n_3+n_2n_3)}, \end{aligned} \tag{9.14}$$

where in (a) we used algebraic identities.

Similar to 2D cellular networks, the frequency reuse factor is equal to the number of non-interfering cells within a cluster of cells. Hence, cells within each cluster will use different frequency bands. To find the frequency reuse factor in the 3D network, we compute the number of non-interfering cells that create one 3D cluster. Clearly, for the reference cell, a co-channel interfering cell is located in the adjacent cluster. Here, a given space is covered by multiple clusters of truncated octahedron cells. In addition, any space can be fully covered by a number of arbitrary-sized truncated

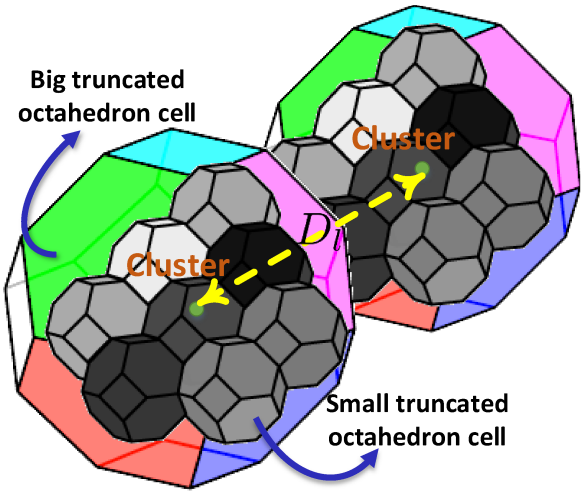


Figure 9.6: Clusters of truncated octahedron cells.

octahedrons. Therefore, we can replace each cluster of cells with a big truncated octahedron cell (as illustrated in Fig. 9.6) of the same volume. In this case, the centers of two co-channel cells are also the centers of two adjacent big truncated octahedron cells, as shown in Fig. 9.6. These two big cells can be connected to each other either from their hexagonal face (reuse distance D_l) or square face (reuse distance D_u). For the hexagonal case, the edge length of the big cells, R_B, is related to the reuse distance distance by:

$$R_B = \frac{D_l}{\sqrt{6}}. \tag{9.15}$$

The number of cells per cluster is equivalent to the volume ratio of the big cell (i.e., cluster) to one truncated octahedron cell:

$$q = \frac{V_B}{V_S} \overset{(a)}{=} \frac{8\sqrt{2}R_B^3}{8\sqrt{2}R^3} = \left(\frac{D_l}{\sqrt{6}R}\right)^3 \overset{(b)}{=} \sqrt{\frac{\left[3(n_1^2+n_2^2+n_3^2)-2(n_1n_2+n_1n_3+n_2n_3)\right]^3}{27}}, \tag{9.16}$$

where V_B and V_S are, respectively, the volumes of one cluster (e.g., big truncated octahedron) and a truncated octahedron cell. (a) follows from the volume of the truncated octahedron as a function of its edge length [214], and (b) follows from (9.14).

For two big cells connecting from their square faces, we have:

$$D_u = R\sqrt{6(m_1^2+m_2^2+m_3^2)-4(m_1m_2+n_1m_3+m_2m_3)}, \tag{9.17}$$

$$R_B = \frac{D_u}{2\sqrt{2}}. \tag{9.18}$$

Then, the integer frequency reuse will be:

$$q = \frac{V_B}{V_S} = \left(\frac{D_u}{2\sqrt{2}R}\right)^3 = \sqrt{\frac{\left[3(m_1^2+m_2^2+m_3^2)-2(m_1m_2+m_1m_3+m_2m_3)\right]^3}{64}}. \tag{9.19}$$

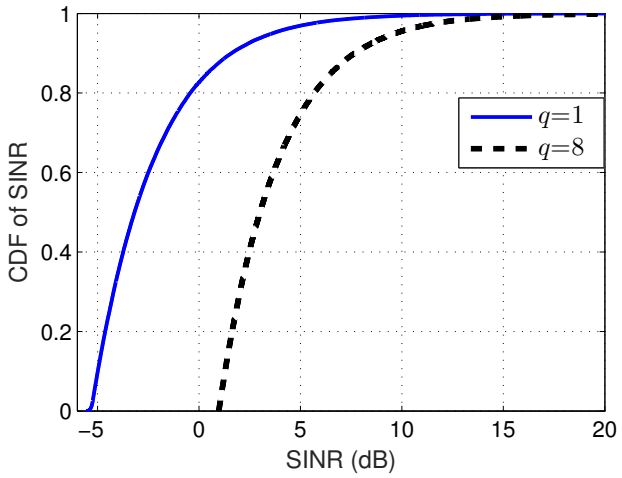


Figure 9.7: CDF of drone-UEs' SINR in a 3D cell for two different frequency reuse factors.

Since the number of cells per cluster represents the frequency reuse factor is a positive integer, (n_1, n_2, n_3) and (m_1, m_2, m_3) must generate an integer in (9.16) and (9.19). Hence, Theorem 19 is proved. □

Theorem 19 can be used to determine the feasible integer frequency reuse factors in the considered 3D network. In addition, while performing frequency planning, the 3D locations of co-channel cells (i.e., drone-BSs) can be identified. As an example, the frequency reuse of one is obtained by considering $(n_1, n_2, n_3) = (1, 0, 0)$, and $(m_1, m_2, m_3) = (1, 1, 0)$. In fact, $q = 1$ corresponds to a worst-case scenario in which all the drone-BSs will interfere with each other. In this case, the locations of co-channel interfering drone-BSs corresponding to a reference cell with an edge length R and center $(0, 0, 0)$ are the columns of the following matrix:

$$\boldsymbol{H} = \sqrt{2}R \begin{pmatrix} 1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & 2 & 0 & 0 & -2 & 0 & 0 \\ -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 0 & 2 & 0 & 0 & -2 & 0 \\ 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 0 & 0 & 2 & 0 & 0 & -2 \end{pmatrix}, \tag{9.20}$$

where each column of matrix $\boldsymbol{H}$ represents a 3D location of one co-channel drone-BS.

In summary, our approach for 3D deployment and frequency planning of drone-BSs can proceed as follows. We deploy the first drone-BS as a reference cell in a specified space of interest. Then, using our truncated octahedron model with parameter R, we use Theorem 18 to find the locations of other drone-BSs with respect to the reference cell. In this case, each drone-BS is located at the center of one truncated octahedron cell. This results in a truncated octahedron tessellation that covers a given space without any gap or overlap. For frequency planning, we use Theorem 19 to find the feasible frequency reuse factors. Then, for any given frequency reuse factor, we determine the sets of co-channel cells in the network. This, in turn, enables us to compute the SINR and transmission latency (which is used in our optimization problem in (9.6)) at any location in the 3D space.

To show the impact of the frequency reuse factor on the SINR of drone-UEs, in Fig. 9.7, we plot the cumulative distribution function (CDF) of drone-UEs' SINR in a 3D cell with a $R = 400\,\text{m}$. As we expect, drone-UEs experience higher SINR for a higher frequency reuse factor (i.e., q). However, a case with a frequency reuse factor 8 requires eight time more bandwidth compared to the case of frequency reuse 1.

9.4 Estimation of the Spatial Distribution of Drone-UEs

Since drone-UEs cannot continuously report their locations due to excessive overhead costs, we need to design a machine learning based mechanism for estimating the locations of drone-UEs using sparse information. Therefore, we assume that each drone-UE is able to report its location at each T seconds. Then, using that, we estimate the spatial distribution of drone-UEs which remains valid for the next T seconds. We should note that, during T seconds, the location of each drone-UE is changing due to its mobility. However, the distribution of drone-UEs is fixed so that we can use our estimation for the period of T seconds. To this end, we develop a nonparametric model for $f(x, y, z)$ using a kernel density estimation (KDE) [215]. In case of parametric density estimation methods, if one uses a poor assumption for the density model, it results in a poor estimation performance. However, nonparametric methods are not sensitive to such poor assumptions.

The distribution of drone-UEs changes with time. Nevertheless, since we assume that this distribution is fixed within an interval of T seconds, we sample the location of each drone-UE every T seconds, and use it to estimate $f(x, y, z)$. This reduces overhead compared to the case in which the system knows the location of drone-UEs at every time instant. We consider some small regions $\mathcal{R}$ where each drone lies in with probability p. Hence, the number of drone-UEs in this region K follows a binomial distribution, i.e.,

$$\Pr(K) = \frac{L!}{(L-K)!K!} p^K (1-p)^{L-K}, \tag{9.21}$$

For a binomial distribution, we know that the mean is $\mathbb{E}(\frac{K}{L}) = p$. Thus, we can write:

$$\lim_{L\to\infty} \frac{K}{L} = p. \tag{9.22}$$

Therefore, for a large L, we can assume $K = Lp$. Since $\mathcal{R}$ is a small region, we can assume that $f(x, y, z), \forall (x, y, z) \in \mathcal{R}$ is constant, and hence:

$$p = \int_{\mathcal{R}} f(x, y, z) \mathrm{d}x\mathrm{d}y\mathrm{d}z = f(x, y, z)\, \mathcal{V}_{\mathcal{R}}, \tag{9.23}$$

where $\mathcal{V}_{\mathcal{R}}$ is the volume of region $\mathcal{R}$. Combining (9.22) and (9.23), we can write:

$$f(x, y, z) = \frac{K}{L\mathcal{V}_{\mathcal{R}}}. \tag{9.24}$$

If we define a small region $\mathcal{R}$ as a cube:

$$\mathcal{C}(\frac{x}{h_x}, \frac{y}{h_y}, \frac{z}{h_z}) = \begin{cases} 1, & \max\{|\frac{x}{h_x}|, |\frac{y}{h_y}|, |\frac{z}{h_z}|\} \leq 1/2, \\ 0, & \text{otherwise}, \end{cases} \tag{9.25}$$

then, we can write the total number of users inside this cube as:

$$K = \sum_{i=1}^{L} \mathcal{C}\left(\frac{x - x_i}{h}, \frac{y - y_i}{h}, \frac{z - z_i}{h}\right) = L h_x h_y h_z f(x, y, z). \tag{9.26}$$

Since the volume of the cube in (9.25) is $h_x \cdot h_y \cdot h_z$, we can write the density function as:

$$f(x, y, z) = \frac{1}{L} \sum_{i=1}^{L} \frac{1}{h_x h_y h_z} \mathcal{C}\left(\frac{x - x_i}{h_x}, \frac{y - y_i}{h_y}, \frac{z - z_i}{h_z}\right), \tag{9.27}$$

which can be interpreted as L cubes with the volume $h_x \cdot h_y \cdot h_z$ centered at each data point. Also, h_x, h_y, and h_z are the widths of the kernel in dimensions x, y, and z, respectively. To remove the discontinuity of cubes in the space, we use Gaussian kernels [216]. If we approximate each cube in (9.27) with a Gaussian kernel, we have:

$$\hat{f}(x, y, z) = \frac{1}{L} \sum_{i=1}^{L} \frac{1}{\sqrt{(2\pi)^3 h_x h_y h_z}} e^{-\left(\frac{(x-x_i)^2}{h_x} + \frac{(y-y_i)^2}{h_y} + \frac{(z-z_i)^2}{h_z}\right)}. \tag{9.28}$$

$\hat{f}(x, y, z)$ is not equal to $f(x, y, z)$, for two reasons. First, L is a finite number, and second, the Gaussian kernel is an approximation of the cube in (9.25). However, we will see that this estimation has small errors even when the value of L is not large. We assume that x, y, and z are uncorrelated, and hence, all the off-diagonal elements of the covariance matrix are zero. Here, the parameters h_x, h_y, and h_z have a major effect on the accuracy of the estimation and need to be estimated. The criteria for accuracy of kernel density estimation is the mean integrated squared error (MISE) and for our problem, it is given by:

$$e = \mathbb{E}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\hat{f}(x, y, z; h_x, h_y, h_z) - f(x, y, z)\right)^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z\right]. \tag{9.29}$$

Since the MISE is not a mathematically tractable expression except in special cases, we have to use approximation methods for approximating it. To this end, we first write MISE as:

$$\mathbb{E}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}^2(x, y, z; h_x, h_y, h_z) + f^2(x, y, z) - 2\hat{f}(x, y, z; h_x, h_y, h_z) f(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z\right], \tag{9.30}$$

where h_x, h_y, and h_z are solutions to the following minimization problem:

$$[h_x, h_y, h_z] = \arg\min \mathbb{E}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}^2(x, y, z) - 2\hat{f}(x, y, z; h_x, h_y, h_z) f(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z\right], \tag{9.31}$$

where $f^2(x,y,z)$ has been omitted since it is a constant in the minimization problem. We can approximate (9.31) using leave-one-out cross-validation (LOOCV) methods. To this end, we first build a model for $\hat{f}(x,y,z;h)$ using the locations of all drone-UEs except one [217]. Then, we find the log-likelihood for the remaining drone-UEs' locations using the current model. We repeat this operation and take an average with L log-likelihood values, i.e.,

$$\mathcal{L}(h_x,h_y,h_z) = \frac{1}{L}\sum_{j=1}^{L} \hat{f}_{-j}(X_j,Y_j,Z_j;h_x,h_y,h_z) \tag{9.32}$$

$$= \frac{1}{L}\sum_{\substack{i=1\\ i\neq j}}^{L} \frac{1}{\sqrt{(2\pi)^3 h_x h_y h_z}} e^{-\left(\frac{(X_j-x_i)^2}{h_x}+\frac{(Y_j-y_i)^2}{h_y}+\frac{(Z_j-z_i)^2}{h_z}\right)}. \tag{9.33}$$

It can be shown [218, 219] that:

$$\mathbb{E}\Big[\hat{f}(x,y,z;h_z,h_y,h_z)\Big] = \mathcal{L}(h_x,h_y,h_z), \tag{9.34}$$

and since

$$\mathbb{E}\Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}(x,y,z;h_x,h_y,h_z)f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z\Big] = \mathbb{E}\Big[\hat{f}(x,y,z;h_z,h_y,h_z)\Big], \tag{9.35}$$

we can find h_x, h_y, and h_z by a cross-validation method as:

$$h_x,h_y,h_z = \arg\min \mathbb{E}\Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}^2(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z - 2\mathcal{L}(h_x,h_y,h_z)\Big]. \tag{9.36}$$

Hence, we can say that $-\mathcal{L}(h_x,h_y,h_z)$ is a biased estimator of MISE. Therefore, it can predict the location of the minimum MISE, and using that, we can find the optimal h_x, h_y, and h_z. Algorithm 6 summarizes the estimation of $f(x,y,z)$ using location of drone-UEs in each T seconds.

Fig. 9.8 shows the MISE in case of symmetric kernels ($h_x = h_y = h_z = h$) for 30 drone-UEs for different values of h. As we can see, our algorithm can potentially minimize the MISE to a value of 7.9554×10^{-04}. Fig. 9.9 shows the negative log-likelihood function. As we can see from Figs. 9.8 and 9.9, $-\mathcal{L}(h_x,h_y,h_z)$ is a biased estimator of MISE, and hence, we can use $-\mathcal{L}(h_x,h_y,h_z)$ to find the optimal h_x,h_y,h_z. Fig. 9.9 shows that, by means of LOOCV method, the MISE for our PDF estimation is 5.7221×10^{-04} which is close to the MISE lower bound that is 7.9554×10^{-04}.

In summary, our approach for estimation of drone-UE spatial distribution is as follows. We collect the location of drone-UEs at each T seconds. Then, we estimate the distribution of drone-UEs to use it for 3D cell association during the next T seconds. We adopt an accuracy metric for our density estimation and use it to find width of the kernels. We showed that our approach is able to estimate the spatial distribution of drone-UEs with a near optimal accuracy.

Algorithm 6 drone-UEs' distribution estimation algorithm

Input: drone-UEs location $(X_1, Y_1, Z_1) \cdots, (X_L, Y_L, Z_L)$
Output: $\hat{f}(x, y, z)$
  Initialize: $\mathcal{H} \leftarrow$ set of candidate for $\{h_x, h_y, h_z\}$, $\mathcal{L}(h_{\text{best}}) \leftarrow \infty$
  for $h_x, h_y, h_z \in \mathcal{H}$ **do**
    for $j = 1, \cdots, L$ **do**
      Build a model using (9.28) with $X_i, i \in \{1, \cdots, L\}, i \neq j$
      sum$\leftarrow$ sum$+\frac{1}{2}\log h_x + \frac{1}{2}\log h_y + \frac{1}{2}\log h_z + \left(\frac{(X_j - x_i)^2}{h_x} + \frac{(Y_j - y_i)^2}{h_y} + \frac{(Z_j - z_i)^2}{h_z}\right) + \frac{3}{2}\log(2\pi)$
    end for
    $\mathcal{L}(h_x, h_y, h_z) \leftarrow \frac{1}{L}$sum
    if $\mathcal{L}(h_x, h_y, h_z) \leq \mathcal{L}(h_{\text{best}})$ **then**
      $h_{\text{best}} \leftarrow h_x, h_y, h_z$
    end if
  end for
  $h_x, h_y, h_z \leftarrow h_{\text{best}}$
  return $\hat{f}(x, y, z; h_x, h_y, h_z)$ in (9.28) as drone-UEs PDF

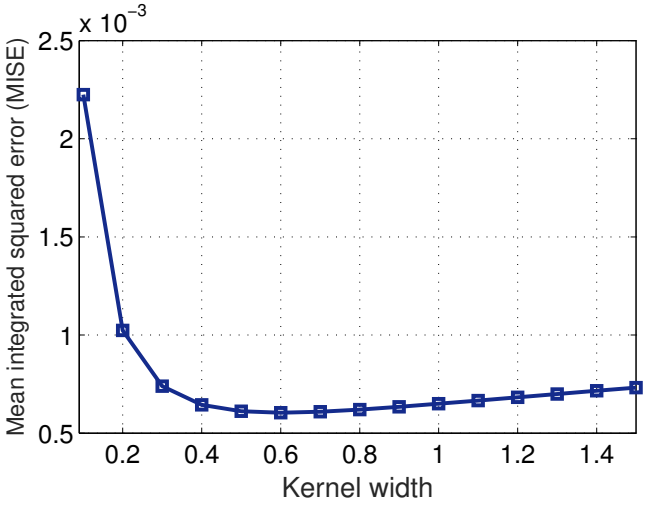


Figure 9.8: MISE for symmetric kernel widths ($h_x = h_y = h_z = h$).

9.5 Optimal 3D Cell Association for Minimum Latency

In Sections III and IV, we determined the locations of drone-BSs and the spatial distribution of drone-UEs. Here, we use this information (i.e., drone-BSs' locations and drone-UEs' distribution) to explicitly formulate our latency-optimal 3D cell association problem.

$$\min_{\mathcal{V}_1,\ldots,\mathcal{V}_N} \sum_{n=1}^{N} \left[\int_{\mathcal{V}_n} \frac{\beta K_n}{B_n \log_2\left(1 + \gamma_n(x, y, z)\right)} \hat{f}(x, y, z) \mathrm{d}x\mathrm{d}y\mathrm{d}z + \frac{\beta K_n}{C_n} + g_n(\beta K_n) \right], \tag{9.37}$$

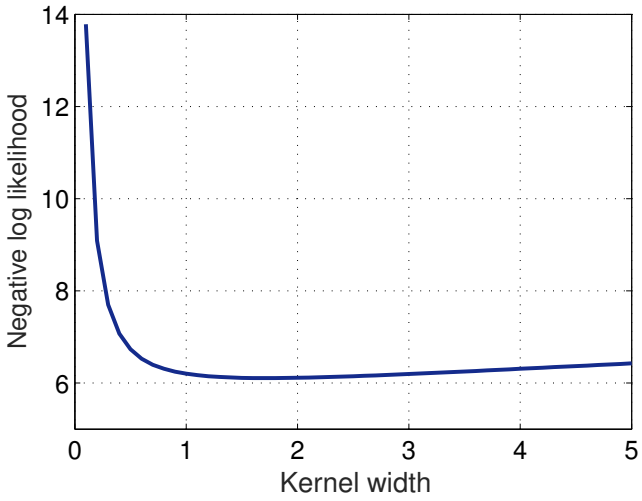


Figure 9.9: LOOCV method for finding an optimal kernel width (h).

$$\text{s.t.} \quad K_n = L \int_{\mathcal{V}_n} \hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z, \tag{9.38}$$

$$\mathcal{V}_l \cap \mathcal{V}_m = \emptyset, \ \ \forall l \neq m \in \mathcal{N}, \tag{9.39}$$

$$\bigcup_{n\in\mathcal{N}} \mathcal{V}_n = \mathcal{V}, \tag{9.40}$$

where $\gamma_n(x,y,z)$ is the downlink SINR for a drone-UE located at (x,y,z) which is served by drone-BS n. Considering a practical bounded path loss model [220] for air-to-air communications, the SINR can be given by:

$$\gamma_n(x,y,z) = \frac{\eta P_n [1 + d_n(x,y,z)]^{-\alpha}}{\sum\limits_{u\in\mathcal{I}_{\text{int}}} \eta P_u [1 + d_u(x,y,z)]^{-\alpha} + N_o B_n}, \tag{9.41}$$

$$d_n(x,y,z) = \sqrt{(x-x_n)^2 + (y-y_n)^2 + (z-z_n)^2}, \tag{9.42}$$

$$d_u(x,y,z) = \sqrt{(x-x_u)^2 + (y-y_u)^2 + (z-z_u)^2}, \ \ u \in \mathcal{I}_{\text{int}}, \tag{9.43}$$

where α is the path loss exponent, N_o is the noise power spectral density, η is the path loss constant, and (x_n, y_n, z_n) is the 3D location of drone-BS n. $d_n(x,y,z)$ and $d_u(x,y,z)$ are, respectively, the distance of drone-BSs n and u with a drone-UE located at (x,y,z). Also, $\mathcal{I}_{\text{int}}$ is the set of co-channel interfering drone-BSs that operate over the same frequency band as drone-BS n.

Solving (9.37) is challenging since the optimization variables $\mathcal{V}_n$, $\forall n \in \mathcal{N}$, are continuous 3D association spaces which are mutually dependent. Furthermore, the fact that the size and shape of these 3D association spaces are unknown, exacerbates the complexity. In addition, the objective function in (9.37) does not have a closed-form expression thus making the problem intractable. Consequently, employing traditional optimization techniques (e.g., convex optimization) are not

sufficient to solve (9.37). Here, we tackle our 3D space association by exploiting optimal transport theory. In particular, first, we prove the existence of an optimal solution to (9.37) and, then, we completely characterize the solution space. We note that, compared to our previous work in [189], this work is different in terms of the system model, the 3D cell association optimization problem, as well as the solution.

Optimal transport theory is a mathematical tool that is used to find an optimal mapping between two arbitrary probability measures [96]. More specifically, in a semi-discrete optimal transport problem, a continuous probability density function must be mapped to a discrete probability measure. In such a semi-discrete case, the optimal transport map will optimally partition the continuous distribution and assign each partition to one point in the discrete probability measure (which, in our case, is the discrete set of drone-BSs).

Our cell association problem can be modeled as a semi-discrete optimal transport problem in which the source measure (drone-UEs' distribution) is continuous while the destination (distribution of drone-BSs) is discrete. Then, the optimal 3D cell partitions are obtained by optimally mapping the drone-UEs to drone-BSs.

Lemma 4. The optimization problem in (9.37) admits an optimal solution for any semi-continuous function $g_n(.)$, $n \in \mathcal{N}$.

Proof. Consider $K_n = L \int_{\mathcal{V}_n} \hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$ and the following simplex:

$$S = \left\{ \boldsymbol{K} = (K_1, K_2, ..., K_N) \in \mathbb{R}^N; \sum_{n=1}^{N} K_n = L, K_n \geq 0, \forall n \in \mathcal{N} \right\}. \tag{9.44}$$

Given any vector $\boldsymbol{K}$, the optimization problem in (9.37) can be represented by:

$$\min_{\mathcal{V}_1,...,\mathcal{V}_N} \sum_{n=1}^{N} \int_{\mathcal{V}_n} c(\boldsymbol{v}, \boldsymbol{s}_n) \hat{f}(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}, \tag{9.45}$$

$$\text{s.t.} \quad \int_{\mathcal{V}_n} \hat{f}(\boldsymbol{v}) \mathrm{d}\boldsymbol{v} = \frac{K_n}{L}, \tag{9.46}$$

$$\mathcal{V}_l \cap \mathcal{V}_m = \emptyset, \ \ \forall l \neq m \in \mathcal{N}, \ \ \bigcup_{n \in \mathcal{N}} \mathcal{V}_n = \mathcal{V}, \tag{9.47}$$

where $\boldsymbol{s_n}$ is the location of drone-BS n, $\boldsymbol{v} = (x,y,z)$, and $c(\boldsymbol{v}, \boldsymbol{s}_n) = \frac{\beta K_n}{B_n \log_2\left(1+\gamma_n(x,y,z)\right)} + \frac{L}{K_n}\left(\frac{\beta K_n}{C_n} + g_n(\beta K_n)\right)$.

This optimization problem is equivalent to the following semi-discrete optimal transport problem:

$$\min_{T} \int_{\mathcal{V}} c\left(\boldsymbol{v}, \boldsymbol{s}\right) \hat{f}(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}, \ \ \boldsymbol{s} = T(\boldsymbol{v}), \tag{9.48}$$

where $\boldsymbol{s}$ is the location of a drone-BS, and $T(.)$ is the transport map which is related to 3D cell partition $\mathcal{V}_n$ by:

$$\left\{T(\boldsymbol{v}) = \sum_{n=1}^{N} \boldsymbol{s}_n \mathbb{1}_{\mathcal{V}_n}(\boldsymbol{v}); \int_{\mathcal{V}_n} \hat{f}(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = \frac{K_n}{L}\right\}. \tag{9.49}$$

Considering the fact that for any semi-discrete optimal transport problem with a lower semi-continuous cost function an optimal transport map exists [96, 134], (9.45) admits an optimal solution for any $\boldsymbol{K} \in S$. Also, since S is a simplex (which is a non-empty and compact set), problem (9.37) admits an optimal solution over the entire S. This proves Lemma 1. □

Next, given the existence of the optimal solution, we characterize the solution.

Theorem 20. The optimal 3D cell association for drone-BS l, that leads to a minimum average latency in (9.37), is given by:

$$\mathcal{V}_l^* = \Big\{(x,y,z)\Big|\alpha_l + \frac{K_l}{L}h_l(x,y,z) + \frac{\beta}{C_l} + g_l'(\beta K_l) \\ \leq \alpha_m + \frac{K_m}{L}h_m(x,y,z) + \frac{\beta}{C_m} + g_m'(\beta K_m),\ \forall l \neq m\Big\}, \tag{9.50}$$

where $h_l(x,y,z) \triangleq \frac{\beta}{B_l \log_2\left(1+\gamma_l(x,y,z)\right)}$, and $\alpha_l \triangleq \int_{\mathcal{V}_l} h_l(x,y,x)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$.

Proof. See Appendix E.1. □

Using Theorem 20, we can determine the optimal 3D cell partitions associated with each drone-BS that ensure the minimum average latency for drone-UEs. From (9.50), we can see how the optimal 3D association depends on various network's parameters such as the distribution of drone-UEs, locations of drone-BSs, backhaul data rate, load of the network, and the computational speed. Based on these parameters, Theorem 20 is utilized to optimally partition a specified space and determine a minimum latency 3D cell association scheme. In this case, to minimize the average latency, a drone-BS with a faster backhaul link and computational capabilities, or higher bandwidth and transmit power will serve more drone-UEs.

To solve (9.50), we propose the iterative algorithm shown in Algorithm 7. This algorithm, based on [134], converges to the optimal solution within a reasonable number of iterations [134]. Algorithm 7 for solving (9.50) that finds the optimal 3D cell partitions proceeds as follows. The inputs are the 3D spatial distribution of drone UEs, number of drone-UEs, load, locations of the drone-BSs, computation time function, and the number of iterations, Q. In Algorithm 7, t represents the iteration number. First, we generate initial 3D cell partitions $\mathcal{V}_l^{(t)}$, and set $\psi_l^{(t)}(x,y,z) = 0, \forall l \in \mathcal{N}$, with $\psi_l^{(t)}(x,y,z)$ being a pre-defined parameter which is used to update the cell partitions. Next, we update $\psi_l^{(t+1)}(x,y,z)$, and compute K_l in step 6. In step 8, we update the partitions based on (9.50). Finally, we obtain the optimal 3D cell partitions and associations, at the end of the iteration.

Algorithm 7 Iterative algorithm for finding the optimal 3D cell association.

Inputs: $\hat{f}(x,y,z)$, β, Q, L, Locations of drone-BSs, C_l, $g_l(.)$, $\forall l \in \mathcal{N}$.
Outputs: $\mathcal{V}_l^*, \forall l \in \mathcal{N}$.
Set $t=1$, generate an initial cell partitions $\mathcal{V}_l^{(t)}$, and set $\psi_l^{(t)}(x,y,z)=0, \forall l \in \mathcal{N}$.
while $t < Q$ **do**
    Compute $\psi_l^{(t+1)}(x,y,z) = \begin{cases} [1-1/t]\,\psi_l^{(t)}(x,y,z), & \text{if } (x,y,z) \in \mathcal{V}_l{}^{(t)}, \\ 1-[1-1/t]\left(1-\psi_l^{(t)}(x,y,z)\right), & \text{otherwise.} \end{cases}$
    Compute $K_l = \int_{\mathcal{V}} \left(1-\psi_l^{(t+1)}(x,y,z)\right) \hat{f}(x,y,z) \mathrm{d}x\mathrm{d}y\mathrm{d}z, \forall l \in \mathcal{N}$.
    $t \rightarrow t+1$.
    Update cell partitions using (9.50).
end while
$\mathcal{V}_l^* = \mathcal{V}_l^{(t)}$,

In summary, our approach for deployment and latency-optimal cell association in the proposed 3D cellular network is as follows. First, using the proposed truncated octahedron approach, and Theorems 1 and 2 in Section III, we determine the locations of drone-BSs as well as the co-channel cells. Then, in Section IV, we estimate the spatial distribution of drone-UEs using kernel method presented in Algorithm 1. Finally, based on the determined locations of drone-BSs and the drone-UEs' distribution, we use Algorithm 2 to derive the optimal 3D cell association for which the average total latency of serving drone-UEs is minimized.

9.6 Simulation Results and Analysis

For our simulations, we consider a cubic space of size $3\,\text{km} \times 3\,\text{km} \times 3\,\text{km}$ in which 18 drone-BSs are deployed based on the proposed truncated octahedron approach to serve drone-UEs. We determine the locations of drone-BSs by using (9.12) with parameters $a \in \{-1,0,1\}, b \in \{-1,0,1\}$, $c \in \{0,1\}$, and $R = 400\,\text{m}$. We randomly generate a sample (i.e., a realization of a continuous distribution) of drone-UEs' locations based on a three-dimensional truncated Gaussian distribution with a specified mean and variance values. These locations' samples are then used to estimate the spatial distribution of drone-UEs using Algorithm 6. For the computation time, we consider a quadratic function of data size (i.e., load on each drone-BS), but our approach can accommodate any other arbitrary function. Here, the computation time for drone-BS n is $g_n(\beta K_n) = \frac{(\beta K_n)^2}{\omega_n}$, with ω_n being the processing speed of drone-BS n. Unless states otherwise, we use the simulation parameters listed in Table I. We compare our proposed 3D cell association with the classical SINR-based association (i.e., weighted Voronoi diagram) baseline. All statistical results are averaged over a large number of independent runs.

Fig. 9.10 shows the average total latency as a function of the number of drone-UEs for the proposed 3D cell association and the SINR-based association schemes. As we can see from this figure, the total latency increases by increasing the number of drone-UEs. A higher number of drone-UEs

Table 9.1: Simulation parameters.

Parameter	Description	Value
f_c	Carrier frequency	2 GHz
P_n	Drone-BS transmit power	0.5 W
N_o	Noise power spectral density	-170 dBm/Hz
L	Number of drone-UEs	200
B_n	Bandwidth for each drone-BS	10 MHz
α	Path loss exponent	2
η	Path loss constant	1.42×10^{-4}
β	Packet size for drone-UE	10 kb
q	Frequency reuse factor	1
C_n	Backhaul rate for drone-BS n	$(100 + n)$ Mb/s
ω_n	Computation constant (i.e., speed) for each drone-BS	10^2 Tb/s
μ_x, μ_y, μ_z	Mean of the truncated Gaussian distribution in x, y, and z directions	1000 m, 1000 m, 1000 m
$\sigma_x, \sigma_y, \sigma_z$	Standard deviation of the distribution in x, y, and z directions	600 m, 600 m, 600 m

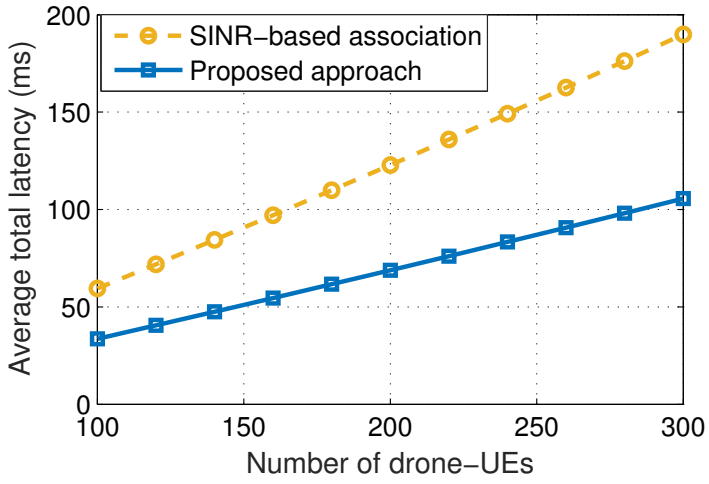


Figure 9.10: Average total latency vs. number of drone-UEs.

leads to a higher network congestion which, in turn, increases transmission time, backhaul latency, and computation time. Fig. 9.10 shows that, when the number of drone-UEs increases from 200 to 300, the total latency increases by 56% and 42% for the SINR-based association and the proposed approach. Moreover, we can see that our proposed approach significantly reduces the latency compared to the SINR association case. This is due to the fact that, in our approach, besides SINR, the impact of congestion on the transmission, backhaul, and computational latencies is also taken into account. The proposed approach avoids creating highly congested 3D cell partitions that can cause excessive latency. From Fig. 9.10, we can see that our approach yields around 46% reduction in the average total latency compared to the SINR-based association.

Fig. 9.11 shows how the latency can be reduced by increasing the transmission bandwidth. By using more bandwidth, the transmission rate increases and, hence, the transmission latency decreases. Fig. 9.11 also reveals that our approach significantly enhances spectral efficiency compared to the SINR-based association. In essence, compared to the SINR case, the proposed approach requires

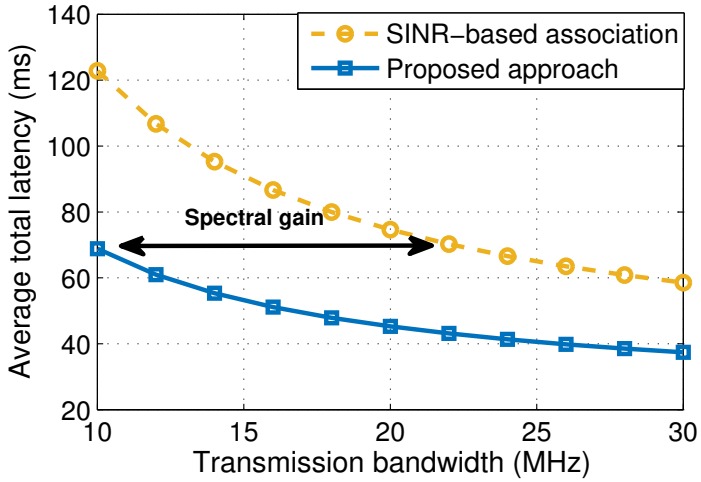


Figure 9.11: Average total latency vs. transmission bandwidth.

less transmission bandwidth in order to meet a certain latency requirement. For instance, as we can see from Fig. 9.11, to ensure a 70 ms maximum total latency, our approach requires 57% less bandwidth than the SINR-based association scheme. Another observation from Fig. 9.11 is that the rate of latency reduction decreases as the bandwidth increases. This is because in large bandwidth scenarios, the transmission latency can be smaller than the computation and backhaul latencies. Thus, the impact of decreasing the transmission latency on the total latency is relatively minor.

Fig. 9.12 shows the impact of drone-UEs' load on the transmission, computation, and backhaul latencies. As expected, these three types of latency increase when the load of drone-UEs increases. Nevertheless, the rate of increase is different for different types of latency. For instance, in Fig. 9.12, the increase rate of the transmission latency is higher than that of computational latency and backhaul latency. The impact of load on each type of latency depends on two factors: 1) the function that directly relates the load to the latency, and 2) the 3D cell partitions which are related to load by (9.50). In fact, while varying load, the cell partitions and different component of latency dynamically change such that the total latency is minimized.

In Fig. 9.13, we show how any error in estimating the spatial distribution of drone-UEs can affect the system performance in terms of latency. As an example, in this figure, we show the additional latency that can be caused by an error in estimating the mean value of the drone-UEs' distribution. Clearly, the total latency significantly depends on the 3D cell partitions which themselves are a function drone-UEs' distribution. Therefore, an estimation error of drone-UEs' distribution leads to a deviation from the optimality of cell partitions. Consequently, such estimation error will increase the latency. For instance, from Fig. 9.13, we can observe around 27% increase in the average total latency when increasing the error up to 400 m. Hence, while estimating the distribution of drone-UEs, it is crucial to adopt high-accuracy models, as we did in Section IV.

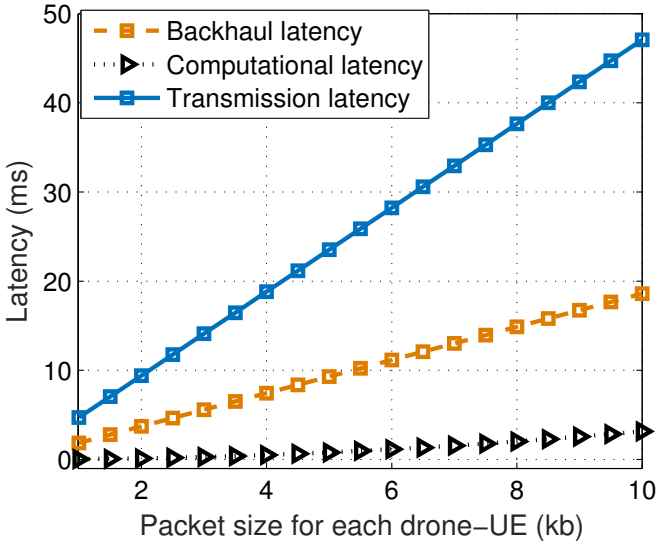


Figure 9.12: Transmission, backhaul, and computation latency vs. load of each drone-UE in the proposed approach.

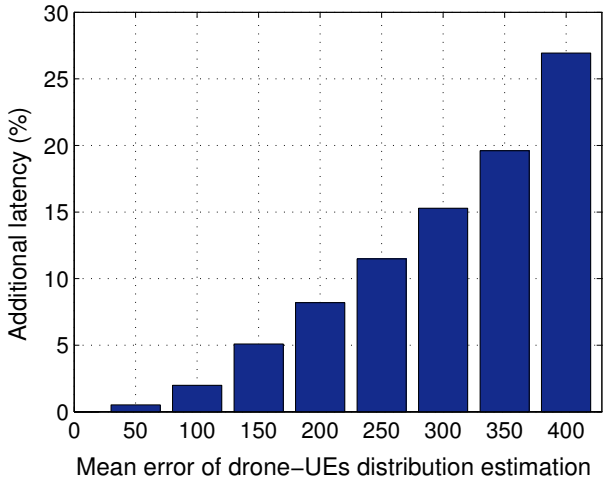


Figure 9.13: Additional latency due to estimation error in mean of drone-UEs' distribution.

9.7 Summary

In this chapter, we have introduced a novel framework for cell association and deployment in 3D cellular networks with drone-BSs and drone-UEs. We have proposed a tractable method for the 3D deployment of drone-BSs and solved the problem of cell association with the goal of minimizing the latency of drone users. For deployment, we have determined the drone-BSs' locations based on a truncated octahedron structure and derived the feasible frequency reuse factor in the considered 3D network. For latency-minimal cell association, first, we have estimated the spatial distribution

of the drone-UEs using the kernel density estimation method. Then, using the estimated distribution of drone-UEs and the location of drone-BSs, we have derived the optimal cell association of drone-UEs using optimal transport theory such that the latency for drone-UEs is minimized. Our results have shown that the proposed approach significantly reduces the latency of drone-UEs compared to the classical SINR-based association. Furthermore, the proposed latency-optimal cell association improves the spectral efficiency of the 3D drone-enabled wireless networks.

Appendix E

E.1 Proof of Theorem 20

In Lemma 1, we proved the existence of the optimal 3D cell partitions $\mathcal{V}_n$, $n \in \mathcal{N}$. Now, consider two 3D partitions $\mathcal{V}_l$ and $\mathcal{V}_m$, and a point $\boldsymbol{v}_o = (x_o, y_o, z_o) \in \mathcal{V}_l$. Also, let $B_\epsilon(\boldsymbol{v}_o)$ be a ball with a center $\boldsymbol{v}_o$ and radius $\epsilon > 0$. Now, we generate the following new 3D partitions $\widehat{\mathcal{V}}_n$ (which are variants of the optimal partitions):

$$\begin{cases} \widehat{\mathcal{V}}_l = \mathcal{V}_l \backslash B_\varepsilon(\boldsymbol{v}_o), \\ \widehat{\mathcal{V}}_m = \mathcal{V}_m \cup B_\varepsilon(\boldsymbol{v}_o), \\ \widehat{\mathcal{V}}_n = \mathcal{V}_n, \quad n \neq l, m. \end{cases} \tag{9.51}$$

Let us define $p_1(K_n) \triangleq K_n$, $p_2(K_n) \triangleq \frac{\beta K_n}{C_n}$, $K_\varepsilon = L \int_{B_\varepsilon(\boldsymbol{v}_o)} \hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$, and $\widehat{K}_n = L \int_{\widehat{\mathcal{V}}_n} \hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$. Considering the optimality of $\mathcal{V}_n$, $n \in \mathcal{N}$, we have:

$$\begin{aligned} &\sum_{n\in\mathcal{N}} \int_{\mathcal{V}_n} p_1\left(K_n\right) h_n(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(K_n) + g_n(\beta K_n) \\ &\overset{(a)}{\leq} \sum_{n\in\mathcal{N}} \int_{\widehat{\mathcal{V}}_n} p_1\left(\widehat{K}_n\right) h_n(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(\widehat{K}_n) + g_n(\beta \widehat{K}_n). \end{aligned} \tag{9.52}$$

Canceling out the common terms in (9.52) leads to:

$$\begin{aligned} &\int_{\mathcal{V}_l} p_1\left(K_l\right) h_l(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(K_l) + g_l(\beta K_l) \\ &+ \int_{\mathcal{V}_m} p_1\left(K_m\right) h_m(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(K_m) + g_m(\beta K_m) \\ &\leq \int_{\mathcal{V}_m \cup B_\varepsilon(\boldsymbol{v}_o)} p_1\left(K_m + K_\epsilon\right) h_m(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(K_m) + g_m(\beta(K_m + K_\epsilon)) \\ &+ \int_{\mathcal{V}_l \backslash B_\varepsilon(\boldsymbol{v}_o)} p_1\left(K_l - K_\epsilon\right) h_l(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(K_l - K_\epsilon) + g_l(\beta(K_l - K_\epsilon)), \end{aligned} \tag{9.53}$$

$$\int_{\mathcal{V}_l} \left(p_1\left(K_l\right) - p_1\left(K_l - K_\epsilon\right)\right) h_l(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(K_l) - p_2(K_l - K_\epsilon)$$

$$
\begin{aligned}
&+ g_l(\beta K_l) - g_l(\beta(K_l - K_\epsilon)) + \int_{B_\varepsilon(\boldsymbol{v}_o)} p_1\left(K_l - K_\epsilon\right) h_l(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z \\
&\leq \int_{\mathcal{V}_m} \left(p_1\left(K_m + K_\epsilon\right) - p_1\left(K_m\right)\right) h_l(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + p_2(K_m + K_\epsilon) - p_2(K_m) \\
&+ g_m(\beta(K_m + K_\epsilon)) - g_m(\beta K_m) + \int_{B_\varepsilon(\boldsymbol{v}_o)} p_1\left(K_m + K_\varepsilon\right) h_m(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z, \quad (9.54)
\end{aligned}
$$

where (a) comes from the fact that $\mathcal{V}_n$, $\forall n \in \mathcal{N}$ are optimal 3D partitions and, thus, any variation of such optimal partitions, shown by $\widehat{\mathcal{V}}_n$, does not lead to a better solution.

Note that, $K_\epsilon = L\int_{B_\epsilon(\boldsymbol{v}_o)} \hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$. Now, we multiply both sides of the inequality in (9.54) by $\frac{1}{K_\epsilon}$ and take the limit when $\epsilon \to 0$. Then, we use the following equalities:

$$\lim_{\varepsilon \to 0} K_\epsilon = 0, \quad (9.55)$$

$$\lim_{K_\epsilon \to 0} \frac{p_1(K_l) - p_1(K_l - K_\epsilon)}{K_\epsilon} = p_1'(K_l), \quad (9.56)$$

$$\lim_{K_\epsilon \to 0} \frac{p_1(K_m + K_\epsilon) - p_1(K_m)}{K_\epsilon} = p_1'(K_m), \quad (9.57)$$

$$
\begin{aligned}
\lim_{K_\epsilon \to 0} \frac{\int_{B_\epsilon(\boldsymbol{v}_o)} p_1\left(K_l - K_\epsilon\right) h_l(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z}{K_\epsilon} &= \frac{p_1\left(K_l\right) h_l(\boldsymbol{v}_o)}{L}, \\
\lim_{K_\epsilon \to 0} \frac{\int_{B_\varepsilon(\boldsymbol{v}_o)} p_1\left(K_m + K_\epsilon\right) h_m(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z}{K_\epsilon} &= \frac{p_1\left(K_m\right) h_m(\boldsymbol{v}_o)}{L}.
\end{aligned}
$$

Finally, using (9.55)-(9.58), we obtain:

$$
\begin{aligned}
&p_1'\left(K_l\right) \int_{\mathcal{V}_l} h_l(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + \frac{1}{L} p_1\left(K_l\right) h_l(\boldsymbol{v}_o) + p_2'(K_l) + g_l'(\beta K_l) \\
&\leq p_1'\left(K_m\right) \int_{\mathcal{V}_m} h_m(x,y,z)\hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + \frac{1}{L} p_1\left(K_m\right) h_m(\boldsymbol{v}_o) + p_2'(K_m) + g_m'(\beta K_m).
\end{aligned} \quad (9.58)
$$

Note that, in $p_1'(K_l)$, the derivative is taken with respect to a single variable which is written as $p_1'(K_l) = \left.\frac{dp_1(t)}{dt}\right|_{t=K_l}$.

We can further proceed to derive a tractable expression for (9.58):

Given $p_1(K_l) = K_l$, we can compute $p_1'(K_l) = 1$, then, using $K_l = \int_{\mathcal{V}_l} \hat{f}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$ leads to:

$$\alpha_l + \frac{1}{L} K_l h_l(\boldsymbol{v}_o) + \frac{\beta}{C_l} + g_l'(\beta K_l) \leq \alpha_m + \frac{1}{L} K_m h_m(\boldsymbol{v}_o) + \frac{\beta}{C_m} + g_m'(\beta K_m), \quad (9.59)$$

As a result, each optimal 3D cell association can be represented by:

$$\mathcal{V}_l^* = \Big\{(x,y,z)\Big|\alpha_l + \frac{K_l}{L}h_l(x,y,z) + \frac{\beta}{C_l} + g_l'(\beta K_l) \\ \leq \alpha_m + \frac{K_m}{L}h_m(x,y,z) + \frac{\beta}{C_m} + g_m'(\beta K_m),\ \forall l \neq m\Big\}, \tag{9.60}$$

which completes the proof of Theorem 20.

Chapter 10

Sum-Rate Analysis for High Altitude Platform (HAP) Drones with Tethered Balloon Relay

10.1 Background, Related Works, and Contributions

Satellite and terrestrial cellular communications are the two most widely used communication systems for providing global connectivity to mobile ground users. While satellites can deliver wireless service to users in remote areas, their spectral efficiency is limited by their large footprints [221]. Meanwhile, terrestrial communication systems cannot guarantee a reliable service for users in remote, rural areas, due to the lack of infrastructure nodes, such as base stations (BSs). High altitude platform (HAP) drones can substantially extend the coverage of terrestrial networks by establishing line-of-sight (LoS) links and adjusting their altitude [10].

To exploit the spatial dimension and enhance spectral efficiency, HAPs will typically rely on highly directive antennas to communicate with ground stations [222]. In a single HAP system with multiple antennas at the transmitter, a spatial multiplexing gain cannot be typically achieved due to a high correlation between parallel paths [223]. However, the deployment of multiple spatially separated HAPs can be a promising solution to exploit spatial multiplexing and boost spectral efficiency. In particular, by using a large number of antennas at the HAPs, one can provide a precise beamforming which is a key requirement for spatial multiplexing. To this end, channel state information (CSI) at the transmitter (CSIT) is required [223], [224]. However, in HAP drone systems, acquiring precise CSIT is challenging due to the high altitudes and the movement of the drones. Consequently, in HAPs-to-ground stations (GSs) communications[1] exploiting spatial multiplexing,

[1] An HAP drones-GSs wireless system in which each HAP drone carries a dedicated symbol for each GS can be modeled as an X network. The X network houses all possible channel models such as the interference channel, the Y channel, and the Z channel. Beyond offering a generalized structure, an X network offers a maximum capacity as

which can yield a maximum possible sum-rate, is also challenging.
One practical approach to achieve a maximum possible sum-rate for HAP drones-to-GSs communications is via the use of interference alignment (IA) schemes [224–226]. In particular, at a high SNR regime, which is a typical case in HAP communications, IA can achieve a maximum sum-rate by restricting interference beams to a smaller subspace that does not overlap with the desired signal space [224]. Unlike terrestrial wireless systems in which the BSs' positions are fixed, acquiring CSIT to implement IA in HAPs-GSs communication systems is challenging due to imperfect HAP drone stabilization. As a result, the lack of exact CSI at HAP drones can yield a significant degradation of the sum-rate performance. Nevertheless, with the use of relays, it is possible to achieve maximum sum-rate when CSIT is not available [227].
Unlike the time domain realization of IA, the performance of IA in the spatial domain is limited due to the restrictions in designing precoding matrices [224]. Therefore, the maximum possible sum-rate can be achieved by implementing relay-assisted IA in the time domain. In this case, one can use popular relaying mechanisms such as amplify and forward (AF) or decode and forward (DF) [227]. An AF relaying scheme with multiple relays achieves an upper bound for the DoF of a generalized X network [227]. Hence, by adopting an AF relaying scheme such as the one in [227], multiple relays can be used to achieve maximum possible sum-rate of the system where HAP drones have knowledge of CSI. Similar results can be achieved with a DF-based relaying scheme with single relay in a two-user X channel [228]. However, the previous works in [227] and [228] did not investigate the use of a DF relaying mechanism for IA in an HAP drones system. The main contribution of this chapter is a novel framework for maximizing the sum-rate of a relay-aided HAP drones wireless system when the CSI is not available. In particular, to achieve the maximum sum-rate, we propose a DF scheme involving M HAP drones, N ground receivers, and one relay with $(M-1)\times(N-1)$ antennas. In this scenario, we show that it is possible to achieve the maximum possible sum-rate by exploiting the IA scheme. Moreover, we derive a closed-form analytical expression for the capacity of an $M \times N$ X channel with tethered balloon relay. Simulation results verify our analytical results and show that a significant sum-rate gain can be achieved by using the proposed scheme.

10.2 System Model

Consider a geographical area with N GSs (or receivers) and a tethered balloon attached to a control station, as shown in Fig. 10.1. This control station provides the power required to operate the tethered balloon. Meanwhile, the GSs receive data from M HAPs, which are located at altitudes within the range of 17-22 km [221]. Each HAP and receiver houses A antennas while the tethered balloon has $(M-1)\times(N-1)$ antennas. Each GS receives data from each HAP, forming an $M\times N$ X network. Unlike [222], which uses a frequency duplexing technique to avoid the interference in HAP communications, our proposed model uses a relay that operates in half duplex mode in the same frequency band.

compared to other channel models such as the interference channel [224].

10.2.1 Channel Model

For terrestrial communications, the channel is typically modeled as Rayleigh in urban areas and Rician in suburban scenarios. However, in air-to-ground communications, the channel has different characteristics [229]. In urban environments, the air-to-ground channel experiences Rician fading due to the presence of LoS links. In suburban areas, a Rayleigh fading is experienced due to the presence of reflected signals which are stronger than LoS signals [229].

Here, we adopt a Rician channel model in which both LoS and non-LoS (NLoS) paths are considered. Therefore, the channel gain matrix can be represented as [230]:

$$\boldsymbol{H} = \sqrt{\frac{\kappa}{1+\kappa}}\boldsymbol{H}_{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}}\boldsymbol{H}_{\text{NLoS}}, \tag{10.1}$$

where $\boldsymbol{H}_{\text{LoS}}$ and $\boldsymbol{H}_{\text{NLoS}}$ represent, respectively, the channel matrices for LoS and non-LoS communication. The Rician factor κ is given by, $\kappa = \frac{\sigma^2_{\text{LoS}}}{\sigma^2_{\text{NLoS}}}$ [231], where σ^2_{LoS} and σ^2_{NLoS} are the power of LoS path and NLoS path, respectively. For our model, we consider A antennas at each HAP and GS. The role of HAP and GS as transmitter or receiver can be reversed, using the reciprocity property. The static MIMO channel, excluding the path loss, is given by [231]:

$$\bar{\boldsymbol{H}}_{\text{LoS}} = \begin{bmatrix} 1 \\ e^{j2\pi\frac{d_R}{\lambda}\sin(\theta_A)} \\ \vdots \\ e^{j2\pi\frac{d_R}{\lambda}(M-1)\sin(\theta_A)} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ e^{j2\pi\frac{d_T}{\lambda}\sin(\theta_D)} \\ \vdots \\ e^{j2\pi\frac{d_T}{\lambda}(N-1)\sin(\theta_D)} \end{bmatrix}^T, \tag{10.2}$$

where d_R and d_T are the antenna spacing at the receiver and transmitter ($L \gg d_R, d_T$), and λ is the wavelength. Also, θ_A and θ_D represent, respectively, the angle-of-arrival at the receiver and angle-of-departure at the transmitters. We also note that the NLoS MIMO channel follows a Rayleigh distribution.

10.3 DF relay aided Interference Alignment for systems without CSIT

Knowledge of CSIT is a key requirement for designing a precoder at the HAPs. However, at high altitudes acquiring a precise CSIT is challenging due to the difficulty in stabilizing the aerial platform that is affected by wind and other natural factors. In such a case, to achieve IA, a feasible solution is to accommodate a tethered balloon relay to perform DF operation. A DF scheme will require at least M antennas to decode M transmitted symbols. However, the DF relay must accommodate $(M-1)\times(N-1)$ antennas to perform IA. By exploiting temporal domain characteristics, we can transmit MN number of symbols in $M+N-1$ time slots for transmitter and receiver.

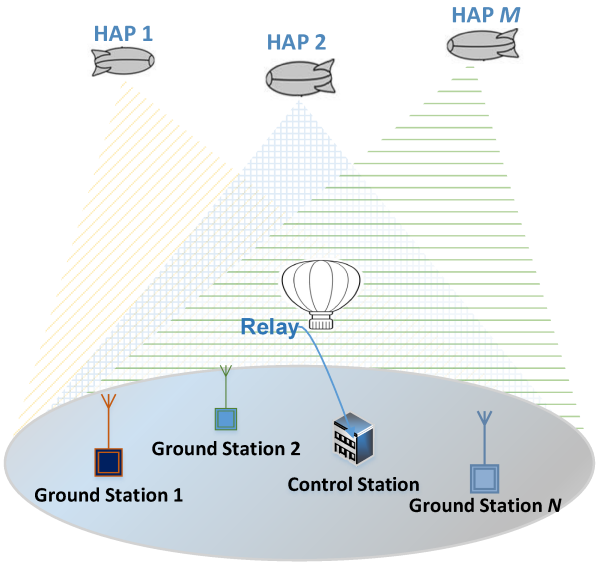


Figure 10.1: System model.

10.3.1 Feasibility of DF in tethered balloon relay

For IA, the tethered balloon must have $(M-1)\times(N-1)$ antennas. To tightly pack these antennas in a tethered balloon relay, the HAPs must be placed sufficiently apart so that the links from HAPs to GSs become uncorrelated. We deduce the minimal separation required between HAPs, when they are located at 18 km above earth, operating at 48 GHz [221]. From the concept of LoS MIMO [223], we have:

$$d_{\text{HAP}} d_{\text{GS}} = \frac{L\lambda}{\beta}, \tag{10.3}$$

where L is the distance between an HAP and the GS, β represents the degrees of freedom. Also, d_{HAP} and d_{GS}, respectively, represent the inter-HAP and inter-GS distances. To determine the HAP separation distance, we choose the following parameters: $d_{\text{GS}} = 0.1\,\text{m}$, $L = 18\,\text{km}$, $\lambda = 0.00625\,\text{m}$, and $\beta = 1$. Now, by using (10.3), we find that the HAPs must be spaced 1125 m apart to ensure that the channels will be uncorrelated. In this case, each HAP drone-GS wireless communication channel will be full rank which allows sending data over multiple paths.

The single tethered balloon relay performs the DF operation and operates in half-duplex mode as shown in [227]. The communication between HAPs and GSs is carried out in two phases, direct transmission and relay aided transmission. In the first phase, GSs receive signal directly from the HAP drones while the relay remain silent. In this case, HAP drones transmit data to one specific GS during one time slot. In the second phase, tethered balloon relay is active and transmits data to the GS after precoding. Therefore, the GS receives signals from the relay and the first transmitter.

10.3.2 Capacity of Rician X network

The asymptotic sum-rate of a network as a function of the signal-to-noise-ratio (SNR) can be expressed as [224, 226]:

$$C = \beta \cdot \log(\gamma) + \mathcal{O}(\log(\gamma)), \tag{10.4}$$

where γ is the SNR value at a given receiver.

The DoF for an X network with M transmitters and N receivers each with A antennas, is equal to $\beta = \frac{MNA}{M+N-1}$. However, we consider a relay-aided system that uses a DF relaying mechanism. We denote the channel matrices between the relay and HAP i by $\boldsymbol{H}_i$, and between GS j and the relay by $\boldsymbol{G}_j$. Since the Rician factor of the HAP-to-relay link is greater than that of the relay-to-GS link, we model $\boldsymbol{H}_i$ and $\boldsymbol{G}_j$ with different κ values. In this case, the small-scale fading matrices $\bar{\boldsymbol{H}}_i$ and $\bar{\boldsymbol{G}}_j$ are obtained by replacing κ in (1) by κ_i^u and κ_j^l, respectively. After adding the path loss components to $\bar{\boldsymbol{H}}_i$ and $\bar{\boldsymbol{G}}_j$, we get $\boldsymbol{H}_i = \frac{\alpha_i}{(d_{Ri})^2}\bar{\boldsymbol{H}}_i$ and $\boldsymbol{G}_j = \frac{\psi_j}{(d_{jR})^2}\bar{\boldsymbol{G}}_j$, where α_i and ψ_j represent, respectively, the channel gains in $\boldsymbol{H}_i$ and $\boldsymbol{G}_j$ at a 1 m reference distance. Also, d_{Ri} and d_{jR} are the link distances in HAP i-tethered balloon and tethered balloon-GS j communications. In order to find capacity of the considered relay aided HAPs-GSs system, we first calculate the SNR of each stream. Here, we have two phases: HAPs-tethered balloon communications, and tethered balloon-GSs communications. We begin with the first phase, and the same steps can be used to find the sum-rate of the second phase. The zero-forcing detection (ZF) SNR of the k^{th} parallel channel, γ_k, is given by [230]:

$$\gamma_k = \frac{\Gamma}{[\boldsymbol{W}^{-1}]_{k,k}}, \tag{10.5}$$

where Γ is the transmit power per symbol, and $\boldsymbol{W} = \boldsymbol{H}^H\boldsymbol{H}$. Now, given the knowledge of ZF SNR, we derive the explicit expression for ZF-capacity as follows.

Theorem 21. The capacity of a relay aided multiple HAPs-GSs communication is given by:

$$C = \frac{MN}{M+N-1}\min\Big(\sum_{i=1}^{M}\log_2\left(1+\Gamma_s\boldsymbol{h}_{i1}^H\boldsymbol{Q}_i\boldsymbol{h}_{i1}\right), \sum_{j=1}^{N}\log_2\left(1+\Gamma_s\boldsymbol{g}_{j1}^H\boldsymbol{W}_j\boldsymbol{g}_{j1}\right)\Big), \tag{10.6}$$

where $\boldsymbol{Q}_i$ and $\boldsymbol{W}_j$ are given by:

$$\boldsymbol{Q}_i = [\boldsymbol{I} - \tilde{\boldsymbol{H}}_i(\tilde{\boldsymbol{H}}_i^H\tilde{\boldsymbol{H}}_i)^{-1}\tilde{\boldsymbol{H}}_i^H], \tag{10.7}$$

$$\boldsymbol{W}_j = [\boldsymbol{I} - \tilde{\boldsymbol{G}}_j(\tilde{\boldsymbol{G}}_j^H\tilde{\boldsymbol{G}}_j)^{-1}\tilde{\boldsymbol{G}}_j^H], \tag{10.8}$$

with $\tilde{\boldsymbol{H}}_i$, $\boldsymbol{h}_{i1}$ and $\tilde{\boldsymbol{G}}_j$, $\boldsymbol{g}_{j1}$ obtained from channel matrices $\boldsymbol{H}_i$ and $\boldsymbol{G}_j$ based on $\boldsymbol{H}_i = [\boldsymbol{h}_{i1}\tilde{\boldsymbol{H}}_i]$ and $\boldsymbol{G}_j = [\boldsymbol{g}_{j1}\tilde{\boldsymbol{G}}_j]$.

Proof. In the proposed model, we have two phases: HAPs-tethered balloon and tethered balloon-GSs communications. We first find the SNR of HAP i-to-relay link, γ_i, and relay-to-GS j, γ_j. Then, we proceed to calculate the sum-rate. In general, the SNR of the first stream at the receiver is given by:

$$\gamma_1 = \frac{\frac{E_s}{\sigma^2 N_T}}{\left[(\boldsymbol{H}^H \boldsymbol{H})^{-1}\right]_{1,1}}, \tag{10.9}$$

where $\frac{E_s}{N_T}$ is the transmitted energy per symbol, and σ^2 is the noise power. From [232], we know that $[\boldsymbol{W}^{-1}]_{1,1}$ can be found from the elements of $\boldsymbol{H}$. That is,

$$\left[(\boldsymbol{H}^H \boldsymbol{H})^{-1}\right]_{1,1} = \boldsymbol{h}_1^H [\boldsymbol{I} - \tilde{\boldsymbol{H}} (\tilde{\boldsymbol{H}}^H \tilde{\boldsymbol{H}})^{-1} \tilde{\boldsymbol{H}}^H] \boldsymbol{h}_1, \tag{10.10}$$

where, $\boldsymbol{H} = [\boldsymbol{h}_1 \tilde{\boldsymbol{H}}]$. Now, substituting (10.10) in (10.9) leads to:

$$\gamma_1 = \frac{E_s}{\sigma^2 N_T} \boldsymbol{h}_1^H [\boldsymbol{I} - \tilde{\boldsymbol{H}} (\tilde{\boldsymbol{H}}^H \tilde{\boldsymbol{H}})^{-1} \tilde{\boldsymbol{H}}^H] \boldsymbol{h}_1. \tag{10.11}$$

By substituting $\boldsymbol{H}_i$ for $\boldsymbol{H}$ in (10.11), we get ϕ_i:

$$\phi_i = \frac{E_{\text{HAP}}}{\sigma^2 N_T} \boldsymbol{h}_{i1}^H [\boldsymbol{I} - \tilde{\boldsymbol{H}}_i (\tilde{\boldsymbol{H}}_i^H \tilde{\boldsymbol{H}}_i)^{-1} \tilde{\boldsymbol{H}}_i^H] \boldsymbol{h}_{i1}, \tag{10.12}$$

where $\mathbb{E}_{\text{HAP}}$ is the transmitted power of each HAP. Similarly, the SNR in GS j is:

$$\varphi_j = \frac{E_{\text{BL}}}{\sigma^2 N_T} \boldsymbol{g}_{j1}^H [\boldsymbol{I} - \tilde{\boldsymbol{G}}_j (\tilde{\boldsymbol{G}}_j^H \tilde{\boldsymbol{G}}_j)^{-1} \tilde{\boldsymbol{G}}_j^H] \boldsymbol{g}_{j1}, \tag{10.13}$$

where E_{BL} is the transmitted power of tethered balloon, associated with $\boldsymbol{G}_j$. The overall system capacity can be defined as in [233]:

$$C_{DF}(\gamma) = \beta \cdot \min(C_1(\gamma), C_2(\gamma)), \tag{10.14}$$

where $C_1 = \sum_{i=1}^{M} \log(1 + \phi_i)$ is the sum-rate between the HAPs-tethered balloon link and $C_2 = \sum_{j=1}^{N} \log(1 + \varphi_j)$ is the sum-rate between the tethered balloon-GSs link. Finally, by substituting (10.12) and (10.13) in (10.14) this theorem is proved. □

Using Theorem 21, it is possible to analyze the impact of κ_i^u, κ_j^l, d_{Ri}, and M on the capacity of the Rician X-channel in the HAP drones wireless system.

Remark 2. *At high* SNR*, the system fails to achieve a higher sum-rate. This is because, at larger κ_i^u and κ_j^l, the columns of $\boldsymbol{H}$ will be correlated. Hence, IA will fail to achieve a maximum capacity for higher κ_i^u and κ_j^l values.*

Remark 3. *When $M = N$ and κ_i^u=κ_j^l, the capacity of the system is maximum when $\frac{E_{HAP} \times (d_{Ri})^2}{\alpha_i} = \frac{E_{BL} \times (d_{jR})^2}{\psi_j}$. Also, when the transmitted power $E_{HAP} = E_{BL}$, the maximum capacity is obtained when the tethered balloon is at the center of the HAPs-GSs link.*

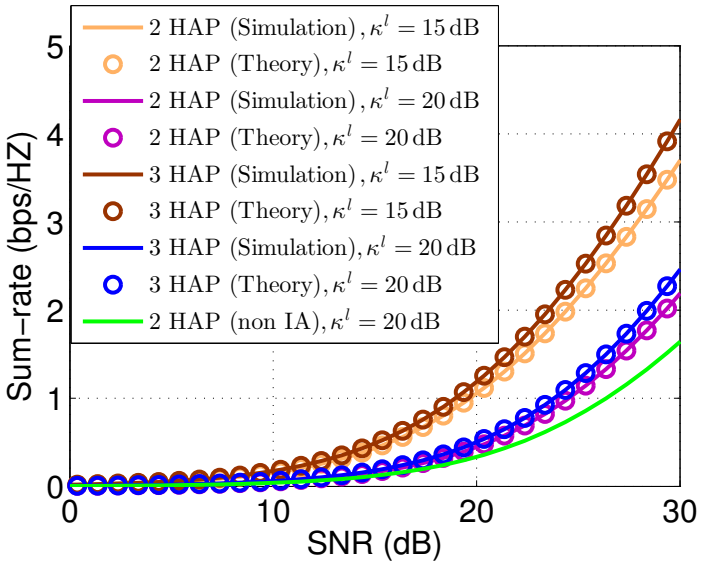


Figure 10.2: Sum-rate vs. SNR.

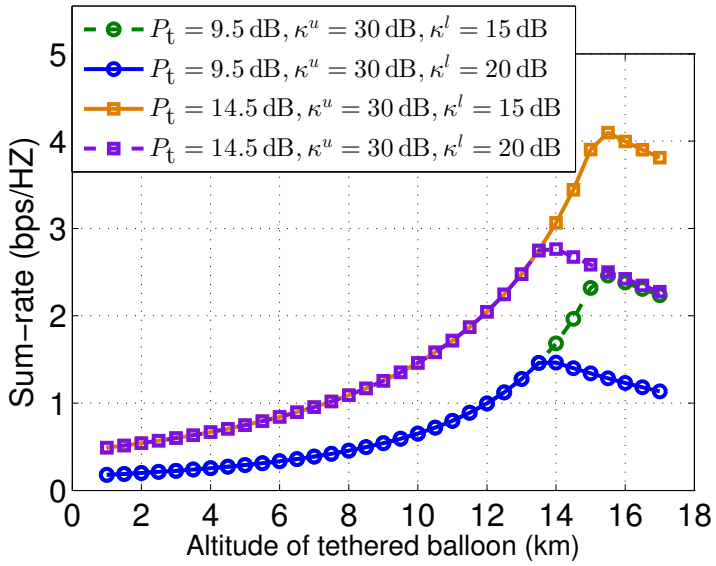


Figure 10.3: Sum-rate vs. tethered balloon's altitude.

10.4 Numerical Results

Here, we evaluate the sum-rate performance of the relay aided multiple HAPs-to-GSs communications. HAPs are located at an altitude of 18 km above earth and each GS receives signals from all HAPs. We consider an X channel, consisting of $M = 2$ and 3 HAPs, a tethered balloon, and 3 ground stations. Fig. 10.2 shows the sum-rate of the Rician channel as a function SNR for dif-

ferent Rician factors and number of HAPs. As we can see from this figure, the sum-rate obtained using simulations (the curves shown without a circle marker) is in agreement with the analytical derivation. Without the use of a tethered balloon, the sum-rate of the considered system decreases as the maximum DoF cannot be achieved. While exploiting IA, however, the system is equivalent to a system in which CSI is known at the transmitters and, hence, a higher sum-rate is achieved. In this analysis, we consider $\kappa^u = 30$ dB and $d_{SR} = 1$ km and we vary κ^l and M. As expected, the sum-rate increases by using the IA solution and decreasing the Rician factor, κ^l. For instance, at SNR $= 25$ dB, the channel with $\kappa^l = 20$ dB offers a sum-rate gain of 34.8%. Interestingly, under the same SNR, the sum-rate increases by up to 1.7 times when $\kappa^l = 15$ dB. From Fig. 10.2, we can also see that, as κ increases, the sum-rate degradation occurs at a higher SNR. This is due to the fact that increasing κ increases the correlation between channels, which results in a degradation of the asymptotic performance.

From Fig. 10.3, we can see that sum-rate increases as the tethered balloon moves away from the HAPs. Similarly, the sum-rate decreases with κ^l, provided that the HAPs are located sufficiently apart. This is due to the fact that exploiting IA while deploying a tethered balloon relay allows achieving a maximum DoF in the system. Fig. 10.3 shows that the optimal altitude of tethered balloon at which the HAPs-tethered balloon-GSs link has the maximum sum-rate (d_{RD}^{opt}) does not change by varying the transmit power. As we can see from this figure, given $d_{SD} = 18$ km, a channel with $P_t = 14.5$ dB, $\kappa^u = 30$ dB and $\kappa^l = 15$ dB offers maximum sum-rate, when $d_{RD}^{\text{opt}} = 15.5$ km. Also, when $\kappa^u = 30$ dB and $\kappa^l = 20$ dB, the optimal relay's altitude is 14 km. In fact, as the difference between κ^u and κ^l become smaller, d_{RD}^{opt} converges towards $\frac{d_{SD}}{2}$.

10.5 Summary

In this chapter, we have proposed an effective interference alignment scheme for maximizing the sum-rate of HAPs-ground stations communications assisted by a tethered balloon relay. In particular, we have considered the half-duplex relaying scheme using a tethered balloon relay to achieve maximum DoF in HAPs-ground stations communications, when the HAPs lack the knowledge of CSI. Our results have shown that, using a tethered balloon relay for exploiting IA provides a significant sum-rate gain in the HAP-based wireless system that uses multiple interfering drones.

Chapter 11

Conclusions and Open Problems

The main body of our work in this dissertation can be summarized as follows. In the first chapter, we have presented the main benefits, applications, and challenges of UAV-enabled wireless networks. Then, we have investigated a number of these technical challenges as well as potential applications of UAVs in wireless networks. In particular, in Chapter 2, we have studied the optimal 3D placement of multiple UAVs used as aerial base stations to provide a maximum coverage. To this end, first, we have derived the downlink coverage probability for of each UAV as a function of the altitude and the antenna gain. Then, using circle packing theory, we have determined the 3D locations of the UAVs such that the total coverage area of the UAVs is maximized. In Chapter 3, we have investigated the efficient deployment and mobility of multiple UAVs used to collect data from ground IoT devices. In particular, to enable reliable uplink communications for IoT devices with a minimum total transmit power, we have proposed a novel framework to jointly optimize the 3D placement and mobility of the UAVs, device-UAV association, and uplink power control. Moreover, based on the activation process of the IoT devices in the time varying IoT network, we have derived the time instances at which the UAVs must update their locations. In Chapter 4, we have analyzed UAV communications under flight time considerations. In particular, we have proposed a novel framework for optimizing the performance of a UAV-based wireless system in terms of the amount of data transmitted to users as well as UAVs' hover duration. To this end, using the mathematical framework of optimal transport theory, we have derived the optimal cell associations that lead to a maximum data service with a minimum required hover time based on any arbitrary spatial distribution of ground users. In Chapter 5, we have performed the fundamental performance of a UAV-based communication system in coexistence with a terrestrial network. In particular, we have considered the coexistence between a UAV, that is transmitting data in the downlink, and an underlaid D2D communication network is considered. In this case, we have derived a tractable analytical framework for the coverage and rate analysis. In Chapter 6 of this dissertation, we have developed a novel framework for effectively deploying cache-enabled UAVs in wireless networks. In particular, by exploiting user-centric information, the cache-enabled UAVs are intelligently used to maximize the users' quality of experience. In Chapter 7, we have proposed a new framework for designing a drone-based antenna array system that can provide wireless services to ground users

within a minimum time. In Chapter 8, we have proposed a delay-optimal cell association scheme in a UAV-assisted cellular networks. In Chapter 9, we have introduced the novel concept of a fully-fledged drone-based 3D cellular network while proposing a new framework for addressing two fundamental problems of deployment and 3D cell association in such a 3D cellular network. In the last chapter of this dissertation, we have investigated an effective interference alignment technique for maximizing the sum-rate of HAPs-to-user communications assisted by a tethered balloon relay.

While we have addressed several key challenges in UAV-based wireless communication systems, some of our results need to be further extended and a number of new directions can be explored. Therefore, in this chapter, we also present an overview of the future directions of this dissertation.

11.1 Summary

11.1.1 UAVs in Wireless Networks: Applications, and Challenges

In Chapter 1, we have presented a comprehensive overview on the potential benefits and applications of UAVs in wireless communications. The applications were drawn from a variety of scenarios, that include imminent use cases, such as for public safety scenarios or hotspot coverage, as well as more futuristic applications such as the use of UAVs as caching apparatus or IoT enablers. Moreover, the important challenges and the fundamental tradeoffs in UAV-enabled wireless networks have been thoroughly investigated. In particular, we have explored the key UAV challenges such as three-dimensional deployment, performance analysis, air-to-ground channel modeling, and energy efficiency.

11.1.2 Efficient Deployment of Multiple Unmanned Aerial Vehicles for Optimal Wireless Coverage

In Chapter 2, we have studied the optimal deployment of multiple UAVs equipped with directional antennas used as aerial base stations. First, the downlink coverage probability was derived based on the probabilistic LoS/NLoS links and considering the shadow fading. Next, given a desired geographical area which needs to be covered by multiple UAVs, an efficient deployment approach was proposed based on the circle packing theory that leads to a maximum coverage while each UAV uses a minimum transmit power. The results have shown that, the optimal altitude and location of the UAVs can be determined based on the number of available UAVs and the gain/beamwidth of the directional antennas.

11.1.3 Mobile Unmanned Aerial Vehicles for Energy-Efficient Internet of Things Communications

In Chapter 3, we have proposed a novel framework for efficiently deploying and moving UAVs to collect data in the uplink from ground IoT devices. In particular, we have determined the jointly optimal UAVs' locations, device association, and uplink power control of the IoT devices such that the total transmit power of the devices under their SINR constraints is minimized. In addition, we have investigated the effective movement of the UAVs to collect the IoT data in a time-varying IoT network. For this case, based on the devices' activation process, we have derived the update time instances at which the UAVs must update their locations. Furthermore, we have obtained the optimal trajectories that are used by the UAVs to dynamically serve the IoT devices with a minimum energy consumption. The results have shown that by intelligently moving and deploying the UAVs, the total transmit power of the devices significantly decreases compared to the case with pre-deployed stationary aerial base stations. Moreover, there is a fundamental tradeoff between the number of updates, the UAVs' mobility, and the devices' transmit power.

11.1.4 Hover Time Optimization in UAV-Enabled-Wireless Networks

In Chapter 4, we have proposed a novel framework for optimizing UAV-enabled wireless networks while taking into account the flight time constraints of UAVs. In particular, we have investigated two UAV-based communication scenarios. First, given the maximum possible hover times of UAVs, we have maximized the average data service to the ground users under a fair resource allocation policy. To this end, using tools from optimal transport theory, we have determined the optimal cell partitions associated with the UAVs. In the second scenario, given the load requirements of users, we have minimized the average hover time of UAVs needed to completely serve the users. In this case, we have derived the optimal cell partitions as well as the optimal bandwidth allocation to the users that lead to the minimum hover time. The results have shown that, using our proposed cell partitioning approach, the users receive higher fair data service compared to the classical Voronoi case. Also, our results for Scenario 2 have revealed that the average hover time of UAVs can be significantly reduced by using the proposed approach.

11.1.5 Fundamental Performance Analysis of Unmanned Aerial Vehicle with Terrestrial Device-to-Device Network

In Chapter 5, we have studied the performance of a UAV that acts as a flying base station in an area in which users are engaged in the D2D communication. We have considered two types of users: in the network: the downlink users served by the UAV and D2D users that communicate directly with one another. For both types, we have derived tractable expressions for the coverage probabilities as the main performance evaluation metrics. The results have shown that a maximum system sum-rate can be achieved if the UAV altitude is appropriately adjusted based on the D2D

users density. In the mobile UAV scenario, using the disk covering problem, the entire target area (cell) can be completely covered by the UAV in a shortest time with a minimum required transmit power. Moreover, in this case, we have derived the overall outage probability for D2D users, and have shown that the outage probability increases as the number of stop point increases. Finally, we have analyzed the tradeoff between the coverage and the time required for covering the entire target area (delay) by the mobile UAV. The results have shown that, the number of stop points must be significantly increased as the minimum coverage requirement for DUs increases.

11.1.6 Cache-Enabled Unmanned Aerial Vehicles in Wireless Networks

In Chapter 6, we have proposed a novel framework that uses flying UAVs to provide service for the mobile users in a CRAN system. First, we have presented an optimization problem that seeks to guarantee the QoE requirement of each user using the minimmum transmit power of the UAVs. Next, to solve this problem, we have developed a novel algorithm based on the echo state networks and conceptors. The proposed algorithm allows predicting the content request distribution of each user with limited information on the network state and user context. The proposed algorithm also enables the ESNs separate the users behavior into several patterns and learn these patterns with various non-linear systems. Simulation results have shown that the proposed approach yields significant performance gains in terms of minimum transmit power compared to conventional ESN approaches.

11.1.7 Communications and Control for Wireless Drone-Based Antenna Array

In Chapter 7, we have proposed a novel framework for employing a drone-enabled antenna array system that can provide wireless services to ground users within a minimum time. To this end, we have minimized the transmission time as well as the control time needed for changing the locations and orientations of the drones. First, we have optimized the positions of drones (as the array elements) within the antenna array such that the transmission time for each user is minimized. Next, given the determined locations of drones, we have minimized the control time of the quadrotor drones by optimally adjusting the rotors' speeds. Our results have shown that the proposed drone antenna array with the optimal configuration yields a significant improvement in terms of the service time, spectral and energy efficiency. In addition, our results have revealed key design guidelines and fundamental tradeoffs for leveraging in an antenna array system. To our best knowledge, this is the first comprehensive study on the joint communications and control of drone-based antenna array systems.

11.1.8 Optimal Transport Theory for Cell Association in UAV-Enabled Cellular Networks

In Chapter 8, we have proposed a novel framework for delay-optimal cell association in UAV-enabled cellular networks. In particular, to minimize the average network delay based on the users' distribution, we have exploited optimal transport theory to derive the optimal cell associations for UAVs and terrestrial BSs. Our results have shown that, the proposed cell association approach results in a significantly lower network delay compared to an SNR-based association.

11.1.9 3D Cellular Networks of Drones: Deployment and Latency Analysis

In Chapter 9, we have proposed a novel framework for cell association and deployment in 3D cellular networks with drone-BSs and drone-UEs. We have described a tractable method for 3D deployment of drone-BSs and solved the problem of user-cell association while minimizing the delay of drone users. In addition, based on the locations of drone-BSs and drone-UEs' distribution, we have derived the optimal cell association of drone-UEs using optimal transport theory such that the latency for drone-UEs are minimized. Our results have shown that the proposed approach significantly reduces the latency of drone-UEs compared to the classical SINR-based association. We have also shown that the proposed latency-optimal cell association improves the spectral efficiency of the 3D drone-enabled wireless networks.

11.1.10 Sum-Rate Analysis for High Altitude Platform (HAP) Drones with Tethered Balloon Relay

In Chapter 10, we have proposed an effective interference alignment scheme for maximizing the sum-rate of HAPs-ground stations communications assisted by a tethered balloon relay. In particular, we have considered the half-duplex relaying scheme using a tethered balloon relay to achieve maximum DoF in HAPs-ground stations communications, when the HAPs lack the knowledge of CSI. Our results have shown that, using a tethered balloon relay for exploiting IA provides a significant sum-rate gain in the HAP-based wireless system that uses multiple interfering drones.

11.2 Open Problems

Despite a considerable number of studies on UAV communications, there are still many key open problems that must be investigated.

11.2.1 UAV Channel Modeling

For air-to-ground channel modeling, there are several key open problems. First and foremost, there is a need for more realistic channel models that stem from real-world measurements [92]. While efforts in this regard already started, most of them remain limited to a single UAV or to very specific environments. A broader campaign of channel measurements that can cut across urban and rural areas, as well as various operational environments (e.g., weather conditions) is needed. Such experimental work can complement the existing, mostly ray tracing simulation based results. Moreover, the simulation results can also be expanded to model small-scale fading A2G communications. In addition, as UAVs become more commonly used as flying base stations, drone-UEs, or even for backhaul support, one must have more insights on air-to-air channel modeling. In particular, there is a need for an accurate UAV-to-UAV channel model that can capture time-variation of channel and Doppler effect due to mobility of UAVs. Furthermore, multipath fading in air-to-air communications needs to be characterized while considering UAVs' altitude as well as antennas' movement.

11.2.2 UAV Deployment

In terms of open problems for UAV deployment, there is a need for new solutions to optimal 3D placement of UAVs while accounting for their unique features. For instance, one of the key open problems is the optimal 3D placement of UAVs in presence of terrestrial networks. For instance, there is a need to study how UAVs must be deployed in coexistence with cellular networks while considering mutual interference between such aerial and terrestrial systems. Other key open problems in deployment include: 1) joint optimization of deployment and bandwidth allocation for low latency communications, 2) joint optimal 3D placement and cell association for flight time minimization, and 3) obstacle aware deployment of UAVs for maximizing wireless coverage.

11.2.3 Performance Analysis

For performance analysis, there are numerous problems that can still be studied. For instance, one must completely characterize the performance of UAV-enabled wireless networks, that consist of both aerial and terrestrial users and base stations, in terms of coverage and capacity. In particular, there is a need for tractable expressions for coverage probability and spectral efficiency in heterogeneous aerial-terrestrial networks. Moreover, fundamental performance analysis needs to be done to capture inherent tradeoffs between spectral efficiency and and energy efficiency in UAV networks. Another open problem is to evaluate the performance of UAV-enabled wireless networks while incorporating the mobility of UAVs. The fundamental analysis of such mobile wireless networks involves capturing the spatial and temporal variations of various performance metrics in the network. For instance, there is a need to study how the trajectory of UAVs impacts their performance in terms of throughput, latency, and energy efficiency. Finally, the effect of

dynamic scheduling on the performance of UAV communication systems needs to be analyzed.

11.2.4 Planning Cellular Networks with UAVs

An efficient network planning with UAVs requires addressing a number of key problems. For example, what is the minimum number of UAVs needed to provide a full coverage for given a geographical area that is partially covered by ground base stations. Solving such problems is particularity challenging when the geographical area of interest does not have a regular geometric shape (e.g., disk or square). Another design problem is the backhaul-aware deployment of UAVs while using them as aerial base stations. In this case, while deploying UAV-BSs, one must consider both the backhaul connectivity of UAVs and their users' quality-of-service. Other important open problems include: 1) performing efficient frequency planning when both ground and aerial BSs and users exist, 2) developing new approaches to dynamically provision UAVs on the fly whenever they join network, and 3) designing robust and adaptive network planning techniques that can account for highly mobile drone-UEs. Last but not least, it is imperative to analyze the signaling overhead associated with the deployment of both UAV-BSs and UAV-UEs, while characterizing how that overhead can affect performance.

11.2.5 Resource Management in UAV Networks

Resource management is another key research problem in UAV-based communication systems. In particular, there is a need for a framework that can dynamically manage various resources including bandwidth, energy, transmit power, UAV's flight time, and number of UAVs, among others. For instance, how to adaptively adjust the transmit power and trajectory of a flying UAV that serves ground users. In this case, a key problem is to provide optimal bandwidth allocation mechanisms that can capture the impact of UAVs' locations, mobility, LoS interference, and traffic distribution of ground users. Also, there is a need for designing efficient scheduling techniques to mitigate interference between aerial and terrestrial base stations in a UAV-assisted cellular network. In addition, one must analyze dynamic spectrum sharing in a heterogeneous network of both flying and ground base stations. Finally, adopting suitable frequency bands (e.g., WiFi, LTE bands) for UAV operations is of important design problems.

11.2.6 UAV Trajectory Optimization

While the potential mobility of UAVs provides promising opportunities, it introduces new challenges and technical problems. In a UAV-assisted wireless network, the trajectory of UAVs needs be optimized with respect to key performance metrics such as in throughput, energy and spectral efficiency, and delay. Furthermore, trajectory optimization problems must account for the dynamic aspects and type of UAVs. The open problems in UAV trajectory optimization include: 1) UAV tra-

jectory optimization under the goal of jointly optimizing communication throughput and the UAV's energy consumption, 2) joint optimization of multiuser communication scheduling and UAV trajectory for minimizing delay, 3) trajectory optimization for maximizing reliability in UAV-enabled wireless networks, and 4) joint control, communication, and trajectory optimization of UAVs for flight time minimization.

11.2.7 Drone-UEs Scenarios

Naturally, flying drones that act as users within cellular networks can introduce new design challenges. In particular, while using drone-UEs in wireless networks, one must account for mobility, LoS interference, handover, energy constraints, and low-latency control of drones. In this regard, key open problems in drone-UEs communications include: 1) developing robust interference mitigation techniques for massive drone-UEs deployment scenarios, 2) designing dynamic handover mechanisms to manage frequent handovers due to mobility, 3) providing accurate ground-to-air channel models for BSs-to-drone communications, 4) proposing new scheduling schemes while considering battery limitations of drones, 5) designing effective solutions that allow meeting URLLC requirements for drone-UEs, and 6) Analyzing application-specific quality-of-service measures.

Bibliography

[1] K. P. Valavanis and G. J. Vachtsevanos, *Handbook of unmanned aerial vehicles*. Springer Publishing Company, Incorporated, 2014.

[2] A. Hanscom and M. Bedford, "Unmanned aircraft system (uas) service demand 2015-2035."

[3] A. Puri, "A survey of unmanned aerial vehicles (UAV) for traffic surveillance," *Department of computer science and engineering, University of South Florida*, 2005.

[4] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile Internet of Things: Can UAVs provide an energy-efficient mobile architecture?" in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, USA, Dec. 2016.

[5] R. Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3-D placement of an aerial base station in next generation cellular networks," in *Proc. of IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, May. 2016.

[6] I. Bucaille, S. Hethuin, A. Munari, R. Hermenier, T. Rasheed, and S. Allsopp, "Rapidly deployable network for tactical applications: Aerial base station with opportunistic links for unattended and temporary events absolute example," in *Proc. of IEEE Military Communications Conference (MILCOM)*, San Diego, CA, USA, Nov. 2013.

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

[8] A. Hourani, K. Sithamparanathan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Communication Letters*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[9] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Drone small cells in the clouds: Design, deployment and performance analysis," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, USA, Dec. 2015.

[10] ——, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Communications Letters*, vol. 20, no. 8, pp. 1647–1650, Aug. 2016.

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

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

[13] S. Rohde and C. Wietfeld, "Interference aware positioning of aerial relays for cell overload and outage compensation," in *Proc. of IEEE Vehicular Technology Conference (VTC)*, Quebec, QC, Canada, Sept. 2012.

[14] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Communications and control for wireless drone-based antenna array," *available online: arxiv.org/abs/1712.10291*, 2017.

[15] A. Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, Austin, TX, USA, Dec. 2014.

[16] V. Sharma, M. Bennis, and R. Kumar, "UAV-assisted heterogeneous networks for capacity enhancement," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1207–1210, June 2016.

[17] V. V. Chetlur and H. S. Dhillon, "Downlink coverage analysis for a finite 3D wireless network of unmanned aerial vehicles," *available online: arxiv.org/abs/1701.01212*, 2017.

[18] E. Kalantari, H. Yanikomeroglu, and A. Yongacoglu, "On the number and 3D placement of drone base stations in wireless cellular networks," in *Proc. of IEEE Vehicular Technology Conference*, 2016.

[19] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (UAVs) for energy-efficient Internet of Things communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, Nov. 2017.

[20] 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 Journal on Selected Areas in Communications, to appear*, 2017.

[21] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3gpp machine-to-machine communications," *IEEECommunications Magazine*, vol. 49, no. 4, pp. 66–74, April. 2011.

[22] H. S. Dhillon, H. Huang, and H. Viswanathan, "Wide-area wireless communication challenges for the internet of things," *available online: arxiv.org/abs/1504.03242.*, 2015.

[23] Facebook, "Connecting the world from the sky," Facebook Technical Report, 2014.

[24] "Paving the path to 5G: Optimizing commercial LTE networks for drone communication," *available online: https://www.qualcomm.com/news/onq/2016/09/06/paving-path-5g-optimizing-commercial-lte-networks-drone-communication.*

[25] Z. Han, A. L. Swindlehurst, and K. Liu, "Optimization of MANET connectivity via smart deployment/movement of unmanned air vehicles," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3533–3546, Dec. 2009.

[26] F. Jiang and A. L. Swindlehurst, "Optimization of UAV heading for the ground-to-air uplink," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 5, pp. 993–1005, June 2012.

[27] A. M. Hayajneh, S. A. R. Zaidi, D. C. McLernon, and M. Ghogho, "Drone empowered small cellular disaster recovery networks for resilient smart cities," in *Proc. of IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*, June 2016.

[28] M. M. Azari, F. Rosas, K. C. Chen, and S. Pollin, "Joint sum-rate and power gain analysis of an aerial base station," in *Proc. of IEEE GLOBECOM Workshops*, Dec. 2016.

[29] J. Kosmerl and A. Vilhar, "Base stations placement optimization in wireless networks for emergency communications," in *Proc. of IEEE International Conference on Communications (ICC)*, Sydney, Australia, June. 2014.

[30] K. Daniel and C. Wietfeld, "Using public network infrastructures for UAV remote sensing in civilian security operations," DTIC Document, Tech. Rep., Mar. 2011.

[31] M. N. Soorki, M. Mozaffari, W. Saad, M. H. Manshaei, and H. Saidi, "Resource allocation for machine-to-machine communications with unmanned aerial vehicles," in *IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016.

[32] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Optimal transport theory for cell association in UAV-enabled cellular networks," *IEEE Communications Letters*, vol. 21, no. 9, pp. 2053–2056, Sep. 2017.

[33] P. G. Sudheesh, M. Mozaffari, M. Magarini, W. Saad, and P. Muthuchidambaranathan, "Sum-rate analysis for high altitude platform (HAP) drones with tethered balloon relay," *IEEE Communications Letters, Early access*, 2017.

[34] "AT&T detail network testing of drones in football stadiums," *available online: https://www.androidheadlines.com/2016/09/att-detail-network-testing-of-drones-in-football-stadiums.html.*

[35] A. Merwaday and I. Guvenc, "UAV assisted heterogeneous networks for public safety communications," in *Proc. of IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, March 2015.

[36] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (IoT): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[37] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, "Towards massive machine type cellular communications," *IEEE Wireless Communications Magazine, to appear*, 2016.

[38] Y. Pang, Y. Zhang, Y. Gu, M. Pan, Z. Han, and P. Li, "Efficient data collection for wireless rechargeable sensor clusters in harsh terrains using UAVs," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Austin, TX, USA, Dec. 2014, pp. 234–239.

[39] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, Feb. 2014.

[40] X. Ge, S. Tu, G. Mao, C. X. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 72–79, Feb. 2016.

[41] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "Mmwave massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 13–21, Oct. 2015.

[42] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: challenges and solution approaches," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 22–31, Oct. 2015.

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

[44] M. Alzenad, M. Z. Shakir, H. Yanikomeroglu, and M.-S. Alouini, "Fso-based vertical backhaul/fronthaul framework for 5g+ wireless networks," *available online: arxiv.org/abs/1607.01472*, 2016.

[45] A. Y. Javaid, W. Sun, V. K. Devabhaktuni, and M. Alam, "Cyber security threat analysis and modeling of an unmanned aerial vehicle system," in *Proc. IEEE Conference on Technologies for Homeland Security (HST)*, 2012, pp. 585–590.

[46] K. Hartmann and C. Steup, "The vulnerability of uavs to cyber attacks-an approach to the risk assessment," in *Proc. IEEE International Conference on Cyber Conflict (CyCon)*, 2013, pp. 1–23.

[47] A. Zajić, *Mobile-to-mobile wireless channels*. Artech House, 2012.

[48] Y. Zheng, Y. Wang, and F. Meng, "Modeling and simulation of pathloss and fading for air-ground link of HAPs within a network simulator," in *Proc. of IEEE International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, Beijing, China, Oct. 2013.

[49] J. Holis and P. Pechac, "Elevation dependent shadowing model for mobile communications via high altitude platforms in built-up areas," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 4, pp. 1078–1084, April 2008.

[50] Q. Feng, E. K. Tameh, A. R. Nix, and J. McGeehan, "Modelling the likelihood of line-of-sight for air-to-ground radio propagation in urban environments," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, San Diego, CA, USA, Nov. 2006.

[51] K. Daniel, M. Putzke, B. Dusza, and C. Wietfeld, "Three dimensional channel characterization for low altitude aerial vehicles," in *Proc. of IEEE International Symposium on Wireless Communication Systems*, Sep. 2010, pp. 756–760.

[52] E. Yanmaz, R. Kuschnig, and C. Bettstetter, "Channel measurements over 802.11a-based UAV-to-ground links," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2011.

[53] K. Sasloglou, I. A. Glover, V. Gazis, P. Kikiras, K. Mathioudakis, and I. Andonovic, "Empirical channel models for optimized communications in a network of unmanned ground vehicles," in *Proc. IEEE International Symposium on Signal Processing and Information Technology*, Dec. 2013.

[54] H. Shakhatreh, A. Khreishah, J. Chakareski, H. B. Salameh, and I. Khalil, "On the continuous coverage problem for a swarm of UAVs," in *Proc. of IEEE 37th Sarnoff Symposium*, Sep. 2016, pp. 130–135.

[55] A. M. Hayajneh, S. A. R. Zaidi, D. C. McLernon, and M. Ghogho, "Optimal dimensioning and performance analysis of drone-based wireless communications," in *Proc. of IEEE GLOBECOM Workshops*, Dec. 2016.

[56] K. Gomez, A. Hourani, L. Goratti, R. Riggio, S. Kandeepan, and I. Bucaille, "Capacity evaluation of aerial LTE base-stations for public safety communications," in *Peoc. IEEE European Conference on Networks and Communications (EuCNC)*, June 2015.

[57] S. Jia and Z. Lin, "Modeling unmanned aerial vehicles base station in ground-to-air cooperative networks," *IET Communications*, 2017.

[58] ITU-R, "Rec. p.1410-2 propagation data and prediction methods for the design of terrestrial broadband millimetric radio access systems," *Series, Radiowave propagation*, 2003.

[59] P. Zhan, K. Yu *et al.*, "Wireless relay communications using ann unmanned aerial vehicle," in *Proc. of IEEE 7th Workshop on Signal Processing Advances in Wireless Communications (SPAWC'06)*. IEEE, 2006, pp. 1–5.

[60] E. P. De Freitas, T. Heimfarth, I. F. Netto, C. E. Lino, C. E. Pereira, A. M. Ferreira, F. R. Wagner, and T. Larsson, "UAV relay network to support wsn connectivity," in *Proc. of International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, 2010, pp. 309–314.

[61] D. Orfanus, E. P. de Freitas, and F. Eliassen, "Self-organization as a supporting paradigm for military UAV relay networks," *IEEE Communications Letters*, vol. 20, no. 4, pp. 804–807, 2016.

[62] F. Jiang and A. L. Swindlehurst, "Dynamic UAV relay positioning for the ground-to-air uplink," in *2010 IEEE Globecom Workshops*, Dec. 2010.

[63] Z. Han, A. L. Swindlehurst, and K. Liu, "Optimization of MANET connectivity via smart deployment/movement of unmanned air vehicles," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3533–3546, Dec. 2009.

[64] A. I. Alshbatat and L. Dong, "Performance analysis of mobile ad hoc unmanned aerial vehicle communication networks with directional antennas," *International Journal of Aerospace Engineering*, vol. 2010, 2011.

[65] W. Guo, C. Devine, and S. Wang, "Performance analysis of micro unmanned airborne communication relays for cellular networks," in *Proc. of 9th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP)*. IEEE, 2014, pp. 658–663.

[66] P. Zhan, K. Yu, and A. L. Swindlehurst, "Wireless relay communications with unmanned aerial vehicles: Performance and optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 3, pp. 2068–2085, July 2011.

[67] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Modeling and performance analysis of clustered device-to-device networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4957–4972, July 2016.

[68] M. Afshang and H. S. Dhillon, "Poisson cluster process based analysis of hetnets with correlated user and base station locations," *IEEE Transactions on Wireless Communications, Early access*, 2018.

[69] C. Zhang and W. Zhang, "Spectrum sharing for drone networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 136–144, Jan. 2017.

[70] S. Mumtaz, S. Huq, K. Mohammed, A. Radwan, J. Rodriguez, and R. L. Aguiar, "Energy efficient interference-aware resource allocation in LTE-D2D communication," in *Proc. of IEEE International Conference on Communications (ICC)*, Sydney, Australia, June. 2014.

[71] B. Uragun, "Energy efficiency for unmanned aerial vehicles," in *Proc. of IEEE 10th International Conference on Machine Learning and Applications and Workshops (ICMLA)*, vol. 2, Honolulu, HI, USA, Dec. 2011, pp. 316–320.

[72] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, to appear 2017.

[73] T. X. Tran, A. Hajisami, and D. Pompili, "Cooperative hierarchical caching in 5G cloud radio access networks (C-RANs)," *available online: arxiv.org/abs/1602.02178*, Jan. 2016.

[74] D. Zorbas, T. Razafindralambo, F. Guerriero *et al.*, "Energy efficient mobile target tracking using flying drones," *Procedia Computer Science*, vol. 19, pp. 80–87, June. 2013.

[75] S. R. Anton and D. J. Inman, "Performance modeling of unmanned aerial vehicles with on-board energy harvesting," in *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*. International Society for Optics and Photonics, 2011, pp. 79 771H–79 771H.

[76] P. J. Vincent, M. Tummala, and J. McEachen, "An energy-efficient approach for information transfer from distributed wireless sensor systems," in *Proc. of IEEE International Conference on System of Systems Engineering (IEEE/SMC)*. IEEE, 2006, pp. 6–pp.

[77] M. S. Sharawi, D. N. Aloi, O. Rawashdeh *et al.*, "Design and implementation of embedded printed antenna arrays in small UAV wing structures," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 8, pp. 2531–2538, 2010.

[78] E. T. Ceran, T. Erkilic, E. Uysal-Biyikoglu, T. Girici, and K. Leblebicioglu, "Optimal energy allocation policies for a high altitude flying wireless access point," *available online: arxiv.org/abs/:1509.08304*, 2015.

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

[80] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for UAV-enabled multiple access," *available online: https://arxiv.org/abs/1704.01765*, 2017.

[81] C. D. Franco and G. Buttazzo, "Energy-aware coverage path planning of UAVs," in *Proc. of IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. IEEE, 2015, pp. 111–117.

[82] E. I. Grøtli and T. A. Johansen, "Path planning for UAVs under communication constraints using splat! and milp," *Journal of Intelligent & Robotic Systems*, vol. 65, no. 1-4, pp. 265–282, 2012.

[83] J. Tisdale, Z. Kim, and J. K. Hedrick, "Autonomous UAV path planning and estimation," *IEEE Robotics & Automation Magazine*, vol. 16, no. 2, pp. 35–42, 2009.

[84] M. Mozaffari, A. Broumandan, K. O'Keefe, and G. Lachapelle, "Weak GPS signal acquisition using antenna diversity," *Navigation*, vol. 62, no. 3, pp. 205–218, 2015.

[85] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for Multi-UAV enabled wireless networks," *available online: https://arxiv.org/pdf/1705.02723.pdf, year=2017*.

[86] K. Doğançay, "UAV path planning for passive emitter localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1150–1166, 2012.

[87] A. Rucco, A. P. Aguiar, and J. Hauser, "Trajectory optimization for constrained UAVs: A virtual target vehicle approach," in *Proc. IEEE International Conference on Unmanned Aircraft Systems (ICUAS)*, June 2015.

[88] J. S. Bellingham, M. Tillerson, M. Alighanbari, and J. P. How, "Cooperative path planning for multiple UAVs in dynamic and uncertain environments," in *Proc. IEEE Conference on Decision and Control*, Dec. 2002.

[89] J. How, Y. Kuwata, and E. King, "Flight demonstrations of cooperative control for UAV teams," in *AIAA 3rd" Unmanned Unlimited" Technical Conference, Workshop and Exhibit*, 2004, p. 6490.

[90] J. Tisdale, Z. Kim, and J. K. Hedrick, "Autonomous UAV path planning and estimation," *IEEE Robotics Automation Magazine*, vol. 16, no. 2, pp. 35–42, June 2009.

[91] P. Chandler, S. Rasmussen, and M. Pachter, "UAV cooperative path planning," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2000, p. 4370.

[92] W. Khawaja, I. Guvenc, D. Matolak, U.-C. Fiebig, and N. Schneckenberger, "A survey of air-to-ground propagation channel modeling for unmanned aerial vehicles," *available online: arxiv.org/abs/1801.01656*, 2018.

[93] Z. Gáspár and T. Tarnai, "Upper bound of density for packing of equal circles in special domains in the plane," *Periodica Polytechnica Civil Engineering*, vol. 44, no. 1, pp. 13–32, 2000.

[94] S. Niu, J. Zhang, F. Zhang, and H. Li, "A method of UAVs route optimization based on the structure of the highway network," *International Journal of Distributed Sensor Networks*, 2015.

[95] K. Dorling, J. Heinrichs, G. G. Messier, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 70–85, Jan 2017.

[96] C. Villani, *Topics in optimal transportation*. American Mathematical Soc., 2003, no. 58.

[97] J. Garza, M. A. Panduro, A. Reyna, G. Romero, and C. d. Rio, "Design of UAVs-based 3D antenna arrays for a maximum performance in terms of directivity and SLL," *International Journal of Antennas and Propagation*, vol. 2016, Aug. 2016.

[98] J. F. Bonnans and A. Shapiro, "Optimization problems with perturbations: A guided tour," *SIAM review*, vol. 40, no. 2, pp. 228–264, 1998.

[99] L. C. Evans, "An introduction to mathematical optimal control theory," *Lecture Notes, University of California, Department of Mathematics, Berkeley*, 2005.

[100] C. Villani, *Topics in optimal transportation*. American Mathematical Soc., 2003, no. 58.

[101] E. P. De Freitas, T. Heimfarth, A. Vinel, F. R. Wagner, C. E. Pereira, and T. Larsson, "Cooperation among wirelessly connected static and mobile sensor nodes for surveillance applications," *Sensors*, vol. 13, no. 10, pp. 12 903–12 928, Sep. 2013.

[102] I. Bekmezci, O. K. Sahingoz, and Ş. Temel, "Flying ad-hoc networks (FANETs): A survey," *Ad Hoc Networks*, vol. 11, no. 3, pp. 1254–1270, 2013.

[103] K. Venugopal, M. C. Valenti, and R. W. Heath Jr, "Device-to-device millimeter wave communications: Interference, coverage, rate, and finite topologies," *available online: arxiv.org/abs/:1506.07158*, 2015.

[104] C. A. Balanis, *Antenna theory: analysis and design*. John Wiley & Sons, 2016.

[105] V. V. Chetlur and H. S. Dhillon, "Downlink coverage analysis for a finite 3D wireless network of unmanned aerial vehicles," *available online: arxiv.org/abs/1701.01212*, 2017.

[106] N. Abuzainab, W. Saad, and H. V. Poor, "Cognitive hierarchy theory for heterogeneous uplink multiple access in the Internet of Things," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Barcelona, June 2016.

[107] C.-Y. Tu, C.-Y. Ho, and C.-Y. Huang, "Energy-efficient algorithms and evaluations for massive access management in cellular based machine to machine communications," in *Proc. of Vehicular Technology Conference (VTC)*, San Francisco, CA, USA, Sep. 2011.

[108] K.-C. Chen and S.-Y. Lien, "Machine-to-machine communications: Technologies and challenges," *Ad Hoc Networks*, vol. 18, pp. 3–23, July. 2014.

[109] M. Tavana, V. Shah-Mansouri, and V. W. S. Wong, "Congestion control for bursty M2M traffic in LTE networks," in *Proc. of IEEE International Conference on Communications (ICC)*, London, UK, June 2015.

[110] X. Jian, X. Zeng, Y. Jia, L. Zhang, and Y. He, "Beta/M/1 model for machine type communication," *IEEE Communications Letters*, vol. 17, no. 3, pp. 584–587, Mar. 2013.

[111] 3GPP, "Study on RAN improvements for machine type communication," *TR 37.868*, Sept. 2011.

[112] A. K. Gupta and S. Nadarajah, *Handbook of beta distribution and its applications*. CRC press, 2004.

[113] S. Z. Selim and M. A. Ismail, "K-means-type algorithms: a generalized convergence theorem and characterization of local optimality," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 1, pp. 81–87, 1984.

[114] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[115] R. Sun, M. Hong, and Z. Q. Luo, "Joint downlink base station association and power control for max-min fairness: Computation and complexity," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1040–1054, June. 2015.

[116] R. E. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems, Revised Reprint*. SIAM, 2009.

[117] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[118] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta numerica*, vol. 4, pp. 1–51, 1995.

[119] K. Scheinberg, "An efficient implementation of an active set method for SVMs," *Journal of Machine Learning Research*, vol. 7, no. Oct, pp. 2237–2257, 2006.

[120] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[121] C. Di Franco and G. Buttazzo, "Energy-aware coverage path planning of UAVs," in *Proc. of IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, Vila Real, Portugal, Apr. 2015.

[122] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *available online: arxiv.org/abs/1608.01828*, 2016.

[123] A. Filippone, *Flight performance of fixed and rotary wing aircraft*. Elsevier, 2006.

[124] S. Jeong, O. Simeone, and J. Kang, "Mobile edge computing via a UAV-mounted cloudlet: Optimal bit allocation and path planning," *available online: https://arxiv.org/abs/1609.05362.*, 2016.

[125] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Optimal transport theory for power-efficient deployment of unmanned aerial vehicles," in *Proc. of IEEE International Conference on Communications (ICC)*, May 2016.

[126] F. Aurenhammer, "Voronoi diagrams a survey of a fundamental geometric data structure," *ACM Computing Surveys (CSUR)*, vol. 23, no. 3, pp. 345–405, 1991.

[127] A. Silva, H. Tembine, E. Altman, and M. Debbah, "Optimum and equilibrium in assignment problems with congestion: Mobile terminals association to base stations," *IEEE Transactions on Automatic Control*, vol. 58, no. 8, pp. 2018–2031, Aug. 2013.

[128] H. Ghazzai, "Environment aware cellular networks," *available online: http://repository.kaust.edu.sa/kaust/handle/10754/344436*, Feb. 2015.

[129] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2706–2716, June 2013.

[130] K. Son, S. Chong, and G. D. Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3566–3576, July 2009.

[131] W. Li, S. Wang, Y. Cui, X. Cheng, R. Xin, M. A. Al-Rodhaan, and A. Al-Dhelaan, "Ap association for proportional fairness in multirate wlans," *IEEE/ACM Transactions on Networking*, vol. 22, no. 1, pp. 191–202, Feb. 2014.

[132] F. Santambrogio, "Optimal transport for applied mathematicians," *Birkäuser, NY*, 2015.

[133] L. Ambrosio and N. Gigli, "A user's guide to optimal transport," in *Modelling and optimisation of flows on networks*. Springer, 2013, pp. 1–155.

[134] G. Crippa, C. Jimenez, and A. Pratelli, "Optimum and equilibrium in a transport problem with queue penalization effect," *Advances in Calculus of Variations*, vol. 2, no. 3, pp. 207–246, 2009.

[135] Z.-J. Shi, "Convergence of line search methods for unconstrained optimization," *Applied Mathematics and Computation*, vol. 157, no. 2, pp. 393–405, 2004.

[136] M. C. Achtelik, J. Stumpf, D. Gurdan, and K. M. Doth, "Design of a flexible high performance quadcopter platform breaking the MAV endurance record with laser power beaming," in *Proc. of IEEE International Conference on Intelligent Robots and Systems*, Sep. 2011.

[137] R. Jain, D.-M. Chiu, and W. R. Hawe, *A quantitative measure of fairness and discrimination for resource allocation in shared computer system*. tech. rep., Digital Equipment Corporation, DEC-TR-301, 1984, vol. 38.

[138] Q. Feng, J. McGeehan, E. K. Tameh, and A. R. Nix, "Path loss models for air-to-ground radio channels in urban environments," in *Proc. of IEEE Vehicular Technology Conference (VTC)*, Melbourne, Vic, Australia, May 2006.

[139] E. Yaacoub and O. Kubbar, "Energy-efficient device-to-device communications in LTE public safety networks," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM), Workshop on Green Interntet of Things*, Anaheim, CA, USA, Dec. 2012.

[140] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communication Magazine*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[141] N. Lee, X. Lin, J. G. Andrews, and R. Heath, "Power control for D2D underlaid cellular networks: Modeling, algorithms, and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 1–13, Feb. 2015.

[142] S. Shalmashi, E. Björnson, M. Kountouris, K. W. Sung, and M. Debbah, "Energy efficiency and sum rate tradeoffs for massive MIMO systems with underlaid device-to-device communications," *available online: arxiv.org/abs/1506.00598.*, 2015.

[143] X. Lin, R. Heath, and J. Andrews, "The interplay between massive MIMO and underlaid D2D networking," *IEEE Transactions on Wireless Communications,*, June. 2015.

[144] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[145] M. Haenggi and R. K. Ganti, *Interference in large wireless networks*. Foundations and Trends in Networking, 2009.

[146] F. Baccelli and B. Blaszczyszyn, "Stochastic geometry and wireless networks, volume ii-applications," 2009.

[147] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Modeling and performance analysis of clustered device-to-device networks," *available online: arxiv.org/abs/:1508.02668*, 2015.

[148] E. Artin, *The gamma function*. Courier Dover Publications, 2015.

[149] R. Kershner, "The number of circles covering a set," *American Journal of mathematics*, pp. 665–671, 1939.

[150] G. F. Tóth, "Thinnest covering of a circle by eight, nine, or ten congruent circles," *Combinatorial and computational geometry*, vol. 52, no. 361, p. 59, 2005.

[151] R. K. Ganti, "A stochastic geometry approach to the interference and outage characterization of large wireless networks," Ph.D. dissertation, University of Notre Dame, 2009.

[152] S. P. Weber, X. Yang, J. G. Andrews, and G. De Veciana, "Transmission capacity of wireless ad hoc networks with outage constraints," *IEEE Transactions on Information Theory,*, vol. 51, no. 12, pp. 4091–4102, Nov. 2005.

[153] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, "User-centric ultra-dense networks for 5G: Challenges, methodologies, and directions," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 78–85, April 2016.

[154] M. Peng, Y. Sun, X. Li, Z. Mao, and C. Wang, "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 2282–2308, Thirdquarter 2016.

[155] Y. Pan, C. Pan, H. Zhu, Q. Z. Ahmed, M. Chen, and J. Wang, "On consideration of content preference and sharing willingness in D2D assisted offloading," *IEEE J. Select. Areas Commun. (JSAC), Special Issue on Human-In-The-Loop Mobile Networks*, to appear, 2017.

[156] M. Chen, W. Saad, C. Yin, and M. Debbah, "Echo state networks for proactive caching in cloud-based radio access networks with mobile users," *available online: arxiv.org/abs/1607.00773*, July. 2016.

[157] J. Qiao, Y. He, and S. Shen, "Proactive caching for mobile video streaming in millimeter wave 5G networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 7187–7198, Oct. 2016.

[158] T. X. Tran and D. Pompili, "Octopus: A cooperative hierarchical caching strategy for radio access networks," *available online: arxiv.org/abs/1608.00067*, July 2016.

[159] Y. Guo, L. Duan, and R. Zhang, "Cooperative local caching under heterogeneous file preferences," *IEEE Transactions on Communications*, to appear, 2016.

[160] E. Bastug, M. Bennis, M. Kountouris, and M. Debbah, "Cache-enabled small cell networks: Modeling and tradeoffs," *EURASIP J. Wireless Commun. Netw.,Special Issue Tech. Adv. Design Deployment Future Heterogeneous Netw.*, vol. 2015, no. 1, Feb 2015.

[161] Z. Ye, C. Pan, H. Zhu, and J. Wang, "Tradeoff caching strategy of outage probability and fronthaul usage in Cloud-RAN," *available online: arxiv.org/abs/1611.02660*, Nov. 2016.

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

[163] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Austin, TX, USA, Dec. 2014.

[164] N. T. Nguyen, Y. Wang, H. Li, X. Liu, and Z. Han, "Extracting typical users' moving patterns using deep learning," in *Proc. of IEEE Global Communication Conference (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012.

[165] J. K. Lee and J. C. Hou, "Modeling steady-state and transient behaviors of user mobility: formulation, analysis, and application," in *proc. of ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2006, pp. 85–96.

[166] X. Wang, L. Duan, and R. Zhang, "User-initiated data plan trading via a personal hotspot market," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7885–7898, Nov. 2016.

[167] C. Song, Z. Qu, N. Blumm, and A. L. Barabasi, "Limits of predictability in human mobility." *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[168] E. Bastug, M. Bennis, E. Zeydan, M. A. Kader, I. A. Karatepe, A. S. Er, and M. Debbah, "Big data meets telcos: A proactive caching perspective," *Journal of Communications and Networks*, vol. 17, no. 6, pp. 549–557, Dec. 2015.

[169] D. A. Soysa, D. G. Chen, O. C. Au, and A. Bermak, "Predicting YouTube content popularity via Facebook data: a network spread model for optimizing multimedia delivery," in *Proc. of IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, Singapore, April. 2013.

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

[171] F. Hoppner and F. Klawonn, *Clustering with Size Constraints*. Springer Berlin Heidelberg, 2008.

[172] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, March. 2006.

[173] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 2002.

[174] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1850–1859, April. 2013.

[175] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, and S. Shamai, "Cooperative multicell zero-forcing beamforming in cellular downlink channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3206–3219, June. 2009.

[176] K. Mitra, A. Zaslavsky, and C. Ahlund, "Context-aware QoE modelling, measurement and prediction in mobile computing systems," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 920–936, Dec. 2015.

[177] J. Herbert and H. Harald, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, 2004.

[178] H. Jaeger, "Controlling recurrent neural networks by conceptors," *Eprint Arxiv*, 2014.

[179] M. Lukoševicius, *A Practical Guide to Applying Echo State Networks*. Springer Berlin Heidelberg, 2012.

[180] M. Bennis, S. Perlaza, P. Blasco, Z. Han, and H. Poor, "Self-organization in small cell networks: A reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3202–3212, June. 2013.

[181] M. Chen, W. Saad, and C. Yin, "Echo state networks for self-organizing resource allocation in LTE-U with uplink-downlink decoupling," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, Jan. 2017.

[182] M. Inaba, N. Katoh, and H. Imai, "Applications of weighted voronoi diagrams and randomization to variance-based k-clustering," in *Pro. of Annual Symposium on Computational Geometry*, 1994.

[183] M. Alzenad, A. El-Keyi, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station for maximum coverage of users with different QoS requirements," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 38–41, Feb. 2018.

[184] Q. Wu, J. Xu, and R. Zhang, "UAV-enabled aerial base station (BS) III/III: Capacity characterization of UAV-enabled two-user broadcast channel," *available online: arxiv.org/abs/1801.00443*, 2018.

[185] S. Abeywickrama, L. Jayasinghe, H. Fu, and C. Yuen, "RF-based direction finding of small unmanned aerial vehicles using deep neural networks," *available online: arxiv.org/abs/1712.01154*, 2018.

[186] V. Sharma, R. Sabatini, and S. Ramasamy, "UAVs assisted delay optimization in heterogeneous wireless networks," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2526–2529, Dec. 2016.

[187] Y. A. Nijsure, G. Kaddoum, G. Gagnon, F. Gagnon, C. Yuen, and R. Mahapatra, "Adaptive air-to-ground secure communication system based on ADS-B and wide-area multilateration," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3150–3165, May 2016.

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

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

[190] Y. Zeng, X. Xu, and R. Zhang, "Trajectory optimization for completion time minimization in UAV-enabled multicasting," *IEEE Transactions on Wireless Communications, Early access*, 2018.

[191] W. Su, J. D. Matyjas, M. J. Gans, and S. Batalama, "Maximum achievable capacity in airborne MIMO communications with arbitrary alignments of linear transceiver antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5584–5593, Nov. 2013.

[192] J. E. Bobrow, S. Dubowsky, and J. Gibson, "Time-optimal control of robotic manipulators along specified paths," *The international journal of robotics research*, vol. 4, no. 3, pp. 3–17, Sep. 1985.

[193] W. S. Newman, "Robust near time-optimal control," *IEEE Transactions on Automatic Control*, vol. 35, no. 7, pp. 841–844, July 1990.

[194] T.-S. Chung and C.-J. Wu, "A computationally efficient numerical algorithm for the minimum-time control problem of continuous systems," *Automatica*, vol. 28, no. 4, pp. 841–847, July 1992.

[195] L.-C. Lai, C.-C. Yang, and C.-J. Wu, "Time-optimal control of a hovering quad-rotor helicopter," *Journal of Intelligent & Robotic Systems*, vol. 45, no. 2, pp. 115–135, Feb. 2006.

[196] M. Afshang, M. S. Helfroush, and A. Zahernia, "Gabor filter parameters optimization for texture classification based on genetic algorithm," in *Proc. of International Conference on Machine Vision*, Dubai, United Arab Emirates, Dec. 2009.

[197] D. K. Cheng, "Optimization techniques for antenna arrays," *Proceedings of the IEEE*, vol. 59, no. 12, pp. 1664–1674, Dec. 1971.

[198] U. Challita and W. Saad, "Network formation in the Sky: Unmanned aerial vehicles for multi-hop wireless backhauling," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, Singapore, Dec. 2017.

[199] W. L. Stutzman and G. A. Thiele, *Antenna theory and design*. John Wiley & Sons, 2012.

[200] S. Vaidyanathan and C.-H. Lien, *Applications of Sliding Mode Control in Science and Engineering*. Springer, 2017, vol. 709.

[201] Y. Mutoh and S. Kuribara, "Control of quadrotor unmanned aerial vehicles using exact linearization technique with the static state feedback," *Journal of Automation and Control Engineering*, vol. 4, no. 5, pp. 340–346, Oct. 2016.

[202] T. Bajd, M. Mihelj, and M. Munih, "Rotation and orientation," in *Introduction to Robotics*. Springer, 2013, pp. 9–36.

[203] J. Lyu, Y. Zeng, and R. Zhang, "Cyclical multiple access in UAV-aided communications: A throughput-delay tradeoff," *available online: arxiv.org/abs/1608.03180*, 2016.

[204] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *available online: arxiv.org/abs/1803.00680*, 2018.

[205] E. Kalantari, I. Bor-Yaliniz, A. Yongacoglu, and H. Yanikomeroglu, "User association and bandwidth allocation for terrestrial and aerial base stations with backhaul considerations," in *Proc. IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Montreal, QC, Canada, Oct. 2017.

[206] M. Alzenad, A. El-Keyi, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station for maximum coverage of users with different QoS requirements," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 38–41, Feb. 2018.

[207] F. Lagum, I. Bor-Yaliniz, and H. Yanikomeroglu, "Strategic densification with UAV-BSs in cellular networks," *IEEE Wireless Communications Letters, Early access*, 2017.

[208] M. M. Azari, F. Rosas, A. Chiumento, and S. Pollin, "Coexistence of terrestrial and aerial users in cellular networks," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM) Workshops*, Singapore, Dec. 2017.

[209] U. Challita, W. Saad, and C. Bettstetter, "Cellular-connected UAVs over 5G: Deep reinforcement learning for interference management," *available online: arxiv.org/abs/1801.05500*, 2018.

[210] S. Zhang, Y. Zeng, and R. Zhang, "Cellular-enabled UAV communication: Trajectory optimization under connectivity constraint," in *Proc. of IEEE International Conference on Communications (ICC), to appear*, Kansas City, USA, May. 2018.

[211] J. Horwath, N. Perlot, M. Knapek, and F. Moll, "Experimental verification of optical backhaul links for high-altitude platform networks: Atmospheric turbulence and downlink availability," *International Journal of Satellite Communications and Networking*, vol. 25, no. 5, pp. 501–528, 2007.

[212] O. Lysenko, S. Valuiskyi, P. Kirchu, and A. Romaniuk, "Optimal control of telecommunication aeroplatform in the area of emergency," *Information and Telecommunication Sciences*, no. 1, 2013.

[213] S. Alam and Z. J. Haas, "Coverage and connectivity in three-dimensional networks," in *Proc. of annual international conference on Mobile computing and networking*, Los Angeles, CA, USA, Sep. 2006.

[214] H. S. M. Coxeter, *Regular polytopes*. Courier Corporation, 1973.

[215] C. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer Information Science and Statistics, 2007.

[216] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151–1163, Jul 2002.

[217] A. Z. Zambom and R. Dias, "A review of kernel density estimation with applications to econometrics," *arXiv preprint arXiv:1212.2812*, 2012.

[218] B. A. Turlach, "Bandwidth selection in kernel density estimation: A review," in *CORE and Institut de Statistique*.

[219] T. Duong and M. L. Hazelton, "Cross-validation bandwidth matrices for multivariate kernel density estimation," *Scandinavian Journal of Statistics*, vol. 32, no. 3, pp. 485–506, 2005.

[220] J. Liu, M. Sheng, L. Liu, and J. Li, "Effect of densification on cellular network performance with bounded pathloss model," *IEEE Communications Letters*, vol. 21, no. 2, pp. 346–349, 2017.

[221] A. Mohammed, A. Mehmood, F. N. Pavlidou, and M. Mohorcic, "The role of high-altitude platforms (HAPs) in the global wireless connectivity," *Proceedings of the IEEE*, vol. 99, no. 11, pp. 1939–1953, Nov. 2011.

[222] Y. Liu, D. Grace, and P. D. Mitchell, "Exploiting platform diversity for GoS improvement for users with different high altitude platform availability," *IEEE Transactions on Wireless Communications*, vol. 8, no. 1, pp. 196–203, Jan. 2009.

[223] A. Irish, F. Quitin, U. Madhow, and M. Rodwell, "Sidestepping the Rayleigh limit for LoS spatial multiplexing: A distributed architecture for long-range wireless fiber," in *Proc. of ITA*, San Diego, CA, USA, Feb. 2013, pp. 1–6.

[224] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom of wireless x networks," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3893–3908, Sept. 2009.

[225] F. Pantisano, M. Bennis, W. Saad, M. Debbah, and M. Latva-aho, "Interference alignment for cooperative femtocell networks: A game-theoretic approach," *IEEE Transactions on Mobile Computing*, vol. 12, no. 11, pp. 2233–2246, Nov. 2013.

[226] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[227] Y. Tian and A. Yener, "Guiding blind transmitters: Degrees of freedom optimal interference alignment using relays," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 4819–4832, Aug. 2013.

[228] H. Zebardast, A. Tajer, B. Maham, and M. Rezaee, "Relay X channels without channel state information at the transmit sides: Degrees of freedom," in *Proc. of WCNC*, New Orleans, LA, March 2015, pp. 493–497.

[229] A. Zajić, *Mobile-to-mobile wireless channels*. Artech House, 2012.

[230] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge university press, 2003.

[231] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM wireless communications with MATLAB*. Wiley & Sons, 2010.

[232] M. Kiessling and J. Speidel, "Analytical performance of mimo zero-forcing receivers in correlated rayleigh fading environments," in *Proc. of SPAWC*. IEEE, 2003, pp. 383–387.

[233] G. Levin and S. Loyka, "Amplify-and-forward versus decode-and-forward relaying: Which is better?" 2012.