

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Traffic-aware Management of Communications Resources in Flying Networks

André Filipe Pinto Coelho

PH.D. THESIS

Doctoral Programme in Telecommunications (MAP-Tele)

Supervisor: Prof. Manuel Alberto Pereira Ricardo (Ph.D.)

Co-supervisor: Prof. Rui Lopes Campos (Ph.D.)



June 20, 2023

Abstract

The growing usage of a myriad of communications services and applications, including augmented and virtual reality, ultra-high definition video streaming, and holographic-type and mission-critical communications, is pushing the capacity of existing telecommunications infrastructures to the limit. Besides the large amount of traffic that the Radio Access Networks (RANs) and back-haul networks have to accommodate, support for ultra reliable and low-latency communications is a growing requirement. The goal is to reliably transmit data with low-latency so that billions of connected machines and humans are able to interact with real and virtual objects in real-time.

These requirements have been defined for 5G networks, which aim at providing differentiated services to multiple verticals, including massive Machine-Type Communications (mMTC), massive Internet-of-Things (mIoT), Ultra-Reliable Low-Latency Communications (URLLC), and enhanced Mobile Broadband (eMBB). Current Mobile Network Operators' infrastructures are limited in providing ubiquitous wireless connectivity. Moreover, their flexibility, availability, reliability, and capacity are not sufficient to satisfy the Quality of Service (QoS) levels expected by the network users in many scenarios, including temporary crowded events and disaster management scenarios.

Flying networks, consisting of Unmanned Aerial Vehicles (UAVs), have been proposed as an effective and flexible solution to provide communications on-demand. UAVs' ability to operate virtually anywhere, anytime as well as to hover, move in 3-Dimensional (3D) space, and carry payloads make them adequate platforms to act as communications nodes, including Wi-Fi Access Points (APs) and cellular Base Stations (BSs). However, in order to satisfy the heterogeneous QoS requirements of multiple communications services and applications, the flying networks should provide always-on wireless connectivity, high capacity, and controlled delays.

This thesis proposes a set of novel Algorithms and Mechanisms acting as Network Functions tailored for service-based Flying Multi-hop Networks (FMNs), named AM2Fly. AM2Fly enables the traffic-aware management of communications resources in FMNs, consisting of UAVs that perform the role of Flying Access Points (FAPs) and Flying Gateways (FGWs) to the Internet, for: 1) extending wireless coverage and capacity from a remote AP or BS; 2) offering heterogeneous QoS levels to ground devices, using a minimum number of UAVs; and 3) ensuring uninterrupted wireless communications paths in dynamic FMNs. AM2Fly benefits from the holistic view over the FMN, available in a Central Station that may take advantage of Edge or Cloud computing for dynamically reconfiguring the FMN.

The works found in the literature focus mainly on UAVs that act as flying communications relays and FAPs providing wireless connectivity to ground devices. However, the QoS levels offered are typically not considered or differentiated. This thesis proposes the integration of UAVs into

telecommunications infrastructures for creating FMNs that can be deployed anywhere, anytime, in order to meet the heterogeneous QoS requirements of any number and type of network slices.

Keywords: Aerial Networks, Flying Networks, Network Slicing, Predictive Routing, Quality of Service (QoS), Queue Size Management, UAV Placement, Unmanned Aerial Vehicles (UAVs).

Resumo

O uso crescente de um conjunto de serviços e aplicações de comunicações, incluindo realidade aumentada e virtual, transmissão de vídeo de ultra elevada definição, e comunicações holográficas e críticas, está a levar a capacidade das infraestruturas de telecomunicações existentes ao limite. Além da grande quantidade de tráfego que as redes de acesso rádio (RANs) e transporte têm de acomodar, o suporte de comunicações ultra fiáveis e de baixa latência é um requisito crescente. O objetivo é a transmissão de dados de forma fiável e com baixa latência para que milhões de máquinas ligadas entre si e humanos possam interagir com objetos reais e virtuais em tempo real.

Estes são requisitos definidos, por exemplo, para as redes 5G, que visam fornecer serviços diferenciados para múltiplos setores, incluindo comunicações massivas para máquinas (mMTC), Internet massiva das coisas (mIoT), comunicações de baixa latência ultra fiáveis (URLLC) e banda larga móvel aprimorada (eMBB). As infraestruturas atuais dos operadores de redes móveis são limitadas no fornecimento de conectividade sem fios ubíqua. Além disso, a sua flexibilidade, disponibilidade, fiabilidade e capacidade não são suficientes para satisfazer os níveis de Qualidade de Serviço (QoS) esperados pelos utilizadores da rede em muitos cenários, incluindo eventos temporários de grande dimensão e cenários de gestão de catástrofes.

As redes voadoras, compostas por Veículos Aéreos Não Tripulados (UAVs), têm sido propostas como uma solução eficaz e flexível para fornecer comunicações a pedido. A capacidade dos UAVs para operar virtualmente em qualquer lugar e a qualquer momento, bem como para pairar, mover-se em três dimensões (3D) e transportar carga, torna-os plataformas adequadas para atuar como nós de comunicações, incluindo Pontos de Acesso (APs) *Wi-Fi* e Estações-Base (BSs) celulares. No entanto, para satisfazer os requisitos de QoS heterogêneos de múltiplos serviços e aplicações de comunicações, as redes voadoras devem fornecer conectividade sem fios permanente, capacidade elevada e atrasos controlados.

Esta tese propõe um conjunto de novos algoritmos e mecanismos que atuam como funções para redes voadoras de múltiplo salto (FMNs) baseadas em serviços, designado de AM2Fly. O AM2Fly permite a gestão sensível ao tráfego dos recursos de comunicações em FMNs, compostas por UAVs que desempenham o papel de APs voadores (FAPs) e *gateways* voadoras (FGWs) para a Internet, para: 1) alargar a cobertura sem fios e a capacidade de um AP ou de uma BS remota; 2) oferecer níveis de QoS heterogêneos a dispositivos terrestres, usando um número mínimo de UAVs; e 3) garantir caminhos de comunicações sem fios ininterruptos em FMNs dinâmicas. O AM2Fly beneficia de uma visão holística da rede, disponível numa Estação Central que pode tirar partido de computação na periferia da rede ou na nuvem para reconfigurar dinamicamente a FMN.

Os trabalhos encontrados na literatura focam-se principalmente em UAVs que atuam como repetidores voadores de comunicações (FCRs) e FAPs que fornecem conectividade sem fios a dispositivos terrestres. No entanto, os níveis de QoS oferecidos não são tipicamente considerados ou diferenciados. Esta tese propõe a integração de UAVs em infraestruturas de telecomunicações para criar FMNs que podem ser implantadas em qualquer lugar e momento, por forma a satisfazer os requisitos heterogêneos de QoS de qualquer número e tipo de redes virtuais.

Palavras-chave: Encaminhamento Preditivo, Gestão do Tamanho de Filas de Espera, Posicionamento de UAVs, Qualidade de Serviço (QoS), Redes Aéreas, Redes Voadoras, Veículos Aéreos Não Tripulados (UAVs), Virtualização de Rede.

Funding

This thesis was financed by *Fundação para a Ciência e Tecnologia* (FCT) and co-financed by the European Social Fund through the Norte Portugal Regional Operational Programme (NORTE 2020), under the Ph.D. grant *SFRH/BD/137255/2018*.

The work presented herein was co-supported by the following research projects:

- *WISE – Traffic-aware Flying Backhaul Mesh Networks (POCI-01-0145-FEDER-016744)*, supported by the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization – COMPETE 2020 Programme, and by National Funds through the portuguese funding agency – FCT.
- *5Go – Products and Services for 5G Networks (POCI-01-0247-FEDER-024539)* and *CHIC – Cooperative Holistic view on Internet and Content (POCI-01-0247-FEDER-024498)*, supported by the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization – COMPETE 2020 and the Lisboa2020 under the PORTUGAL 2020 Partnership Agreement, and through the Portuguese National Innovation Agency (ANI).
- *DECARBONIZE – DEvelopment of strategies and policies based on energy and non-energy applications towards CARBON neutrality via digitalization for citIZens and society (NORTE-01-0145-FEDER-000065)*, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).



REPÚBLICA
PORTUGUESA

CIÊNCIA, TECNOLOGIA
E ENSINO SUPERIOR



Acknowledgments

I am deeply grateful to my supervisors, Professor Manuel Ricardo and Professor Rui Campos, for their guidance, support, and inspiration throughout my research work. Their valuable knowledge, experience, and advice were crucial to my success. I appreciate their encouragement to consistently write and pursue publications since the beginning of my Ph.D. work, which facilitated the process of completing my thesis. They were extraordinary mentors, and I sincerely thank both of them for the time and effort they invested in reviewing my papers and thesis.

I would like to thank INESC TEC and all its esteemed collaborators with whom I have had the privilege of working. I dedicate a special thanks to the Centre for Telecommunications and Multimedia (CTM) of INESC TEC for the institutional and personal support. INESC TEC and CTM provided a stimulating research environment that allowed me to grow as a researcher and made available all the necessary resources and conditions to carry out my Ph.D. work.

I express my gratitude to the Wireless Networks research group at CTM, especially for the insightful discussions and valuable comments provided throughout my Ph.D. work. I also thank all the undergraduate students I worked with, whose contributions were very important to validate the concepts developed during my Ph.D. work.

I would like to thank my friends for their support, confidence, and constant source of strength. I appreciate their presence, even in moments of my physical absence.

I am deeply grateful to my family, especially my parents and sister, for their invaluable support and guidance. Their unconditional love and encouragement were essential in keeping me motivated and achieving my goals. To my father, to whom I would like to dedicate this thesis.

Finally, I want to express my gratitude for the financial support provided by *Fundação para a Ciência e a Tecnologia* (FCT), under the Ph.D. grant SFRH/BD/137255/2018, which made my Ph.D. work possible.

This Ph.D. is the result of hard work and true dedication, but also of lucky moments that favored my path even when scientifically there seemed to be no way forward. I am grateful for all of that.

Thank you all for your invaluable contributions.

André Filipe Pinto Coelho.

*“If I have seen further than others,
it is by standing upon the shoulders of giants.”*

Isaac Newton

Contents

1	Introduction	1
1.1	Scope	1
1.2	Motivation	3
1.3	Problem Statement	4
1.4	Thesis Hypothesis	8
1.5	Objectives	8
1.6	AM2Fly Overview	8
1.7	Research Methodology	10
1.8	Original Contributions	12
1.9	Related Research Projects	18
1.10	Document Structure	22
2	An Algorithm for Gateway Placement in Flying Multi-hop Networks	25
2.1	Related Work	26
2.2	System Model and Problem Formulation	35
2.3	GWP Algorithm	37
2.4	Performance Evaluation	40
2.5	Discussion	45
2.6	Summary	47
3	A Mechanism for Proactive Queue Management in Flying Multi-hop Networks	49
3.1	Related Work	50
3.2	System Model and Problem Formulation	55
3.3	PQM Mechanism	56
3.4	Performance Evaluation	59
3.5	Discussion	64
3.6	Summary	66
4	An Algorithm for Placement and Queue Management in Flying Multi-hop Networks	67
4.1	Related Work	68
4.2	System Model and Problem Formulation	77
4.3	GPQM Algorithm	82
4.4	Performance Evaluation	89
4.5	Discussion	99

4.6	Summary	101
5	A Mechanism for Predictive Routing in Flying Multi-hop Networks	103
5.1	Related Work	104
5.2	System Model and Problem Formulation	126
5.3	RedeFINE Mechanism	127
5.4	Performance Evaluation of RedeFINE	131
5.5	I2R Routing Metric	137
5.6	Performance Evaluation of RedeFINE using I2R	140
5.7	Discussion	144
5.8	Summary	147
6	An Algorithm for Placement and Allocation of Communications Resources in Slicing-aware Flying Multi-hop Networks	149
6.1	Related Work	150
6.2	System Model and Problem Formulation	164
6.3	SurFABLE Algorithm	171
6.4	Performance Evaluation	174
6.5	Discussion	191
6.6	Summary	196
7	Conclusions	199
7.1	Work Developed Overview	199
7.2	Original Contributions	201
7.3	Future Work	205
A	Assessment of the Quality of the Solutions obtained for the Flying Gateway Placement	209
	References	211

List of Figures

1.1	Requirements of the recently released 5G networks (in light blue) and the emerging 6G networks (in dark blue) [1].	2
1.2	A multi-tier Flying Multi-hop Network (FMN) consisting of Unmanned Aerial Vehicles (UAVs) enabling communications in a disaster management scenario.	5
1.3	A multi-tier FMN consisting of UAVs providing Internet access in a music festival.	5
1.4	A multi-tier FMN consisting of UAVs providing wireless connectivity to ground devices.	9
1.5	Experimental validation of the WISE's communications solution.	19
1.6	Experimental validation of the ResponDrone's communications solution.	20
1.7	Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)'s Flying Access Point (FAP) and reference networking scenario used in the performance evaluation carried out in the 5Go project.	21
1.8	Experimental validation and demonstration of the CHIC's communications solution.	22
2.1	Wireless mesh network consisting of clients, routers/communications relays, and gateways connected to the Internet [2].	27
2.2	Representative wireless sensor network topology, where the communications paths consist of wireless links (dashed black lines) in a multi-hop architecture. The positions of the communications nodes (circles) and the offered traffic flows (solid green lines) are typically considered in the state of the art sink placement approaches, in order to meet the targeted network performance (modified from [3]).	29
2.3	Flying network consisting of a gateway UAV and Flying Mesh Access Points (FMAPs); the FMAPs provide wireless communications cells to the ground devices. The placement of the FMAPs is defined by the state of the art NetPlan algorithm [4].	29
2.4	Radio Access Network (RAN) designed by the NetPlan algorithm, depicting the FMAPs (red circles) and the wireless communications cells (green circles) offered to the ground devices (blue circles). The traffic demand of the ground devices is proportional to the size of their circles [4].	30
2.5	Placement of a flying communications relay UAV (red triangle) forwarding data traffic between two ground devices (blue squares) in an IEEE multi-rate network. The semicircles represent the maximum wireless ranges between each ground device and the flying communications relay UAV to enable the use of different IEEE 802.11 physical data rates [5].	31
2.6	Wireless networking scenario addressed by the algorithm proposed in [6], where a UAV is placed to serve a set of ground services with targeted data rate requirements.	32

2.7	Performance of Multi Unmanned Aerial Vehicle-based Particle Swarming Optimization Approach (MPA) proposed in [7].	33
2.8	Multi-tier Free-Space Optics (FSO) based wireless fronthaul/backhaul network, where ultra-dense small cells are overlaid with macro cells [8].	34
2.9	Broadband wireless backhaul network (wavy red lines) for remote oceans areas, enabled by tethered balloons acting as flying communications relays [9].	34
2.10	FMN providing on-demand broadband wireless connectivity to first responders in a disaster management scenario – the Flying Gateway (FGW) is highlighted by the blue circle.	36
2.11	FGW Placement Subspace (S_G) in 2-Dimensional (2D) space, which results from the intersection of the spheres, centered at each UAV, with radius equal to the maximum distance compliant with the target Signal-to-Noise Ratio (SNR) values.	38
2.12	Simulation scenarios used in the evaluation of the GateWay Placement (GWP) algorithm.	43
2.13	Scenario A – Aggregate throughput and delay results measured in the FGW. Position 8 was defined by the GWP algorithm.	44
2.14	Scenario B – Aggregate throughput and delay results measured in the FGW for λ_1 and λ_2 equal to 10% and 90% of the reference fair share L , respectively.	44
2.15	Scenario B – Aggregate throughput and delay results measured in the FGW for λ_1 and λ_2 equal to 25% and 75% of the reference fair share L , respectively.	44
3.1	Average end-to-end delay per flow and total goodput achieved when using the state of the art WQM solution in the wireless network topology depicted in Fig. 3.1a. Error bars represent maximum and minimum values considering at least three runs. WQM allows reducing the end-to-end delay by $8\times$ for one hop, $6\times$ for two hops, and $4\times$ for three hops, without compromising the goodput. The results confirm that large queues (<i>Default</i> approach considered) are typically not needed in wireless networks [10].	52
3.2	Simulation scenarios used in the evaluation of the Proactive Queue Management (PQM) mechanism.	58
3.3	Throughput and delay results for the User Datagram Protocol (UDP) flow generated by FAP ₁ , measured in the FGW, considering the simple scenario. PQM (green circles) is the proposed mechanism, whereas the baseline corresponds to Drop-Tail with the default static queue size of 1000 packets defined in Network Simulator 3 (ns-3).	62
3.4	Analysis on the optimal queue size (Q) for each FAP composing the simple scenario. $Q = 5$ enables the minimum delay and the maximum throughput; the results are expressed by means of the 25 th , 50 th , and 75 th Cumulative Distribution Function (CDF) percentiles, considering UDP traffic.	62
3.5	Aggregate throughput and delay results for UDP traffic, measured in the FGW, considering the complex scenario.	63
3.6	Aggregate throughput and delay results for Transmission Control Protocol (TCP) traffic, measured in the FGW, considering the complex scenario.	64
3.7	Theoretical Packet Loss Ratio (PLR) (PLR_i) versus traffic load ρ_i	65

4.1	UAV acting as a Base Station (BS) connected to a terrestrial BS by means of a Free-Space Optics(FSO) based wireless link [11].	68
4.2	Illustration of the dynamic placement approach proposed in [12], where a UAV acting as a flying BS serves a set of ground devices, each with a network queue. . .	69
4.3	Hierarchical flying network architecture consisting of UAVs in the bottom layer and manned aircrafts in the upper layer. On the left-hand side, a traffic source seeks to send text data, while on the right-hand side a second source seeks to send video data. The placement of the UAVs acting as relay nodes in the bottom layer, in order to forward the traffic to the upper layer, is determined taking into account the Quality of Service (QoS) requirements and the wireless links conditions [13].	70
4.4	Representative networking scenario consisting of a single UAV performing the role of flying communications relay between two ground devices acting as, respectively, source and destination of the data traffic. The authors have concluded that a flying communications relay following a given trajectory outperforms the network performance achieved by a static UAV [14].	71
4.5	Networking scenario addressed by the algorithm proposed in [15]. In order to reestablish a RAN due to a damaged or maintenance ground BS, a UAV-enabled BS is deployed on-demand at optimal altitude, while serving a set of ground Mobile Users (MUs).	72
4.6	Reference networking scenario considered in [16], which consists of a UAV acting as a BS transmitting data towards a ground User Equipment. The authors have demonstrated the advantages of using a flying BS for reducing the delay, especially due to the enhanced Line-of-Sight achieved; this is exacerbated when the UAV's altitude is optimized considering the User Equipment's location.	72
4.7	Hierarchical flying network architecture consisting of UAVs. Due to the limited wireless communications range, multiple flying communications relays and a single FGW are used to forward traffic between a Macrocell BS and the ground devices acting as User Equipment [17].	73
4.8	Software-Defined Networking (SDN) scenario, consisting of a UAV acting as a flying BS coexisting with a ground Wi-Fi Access Point (AP), considered in [18]. The authors have minimized the average M/M/1 queuing delay of the cellular clients, while guaranteeing the delay performance for the Wi-Fi clients, by optimizing the spectrum allocation, the UAV position, the association of cellular clients to the ground AP, and the cellular clients' traffic offloading.	74
4.9	Overview of the Multi-Unmanned Aerial Vehicle Deep Reinforcement Learning Scheduling Algorithm (MADRL-SA), considering a networking scenario consisting of three UAVs learning by observing the environment [19].	75
4.10	Flying network in which the UAVs act as Integrated Access and Backhaul (IAB)-nodes capable of forming access and backhaul networks; they share the spectral resources provided by a fixed BS acting as an IAB-donor [20].	77
4.11	FMN organized into a two-tier architecture. The FAPs provide wireless connectivity to the ground devices and the FGW forwards the data traffic to/from the Internet (dashed blue line). The FMN is controlled by a Central Station deployed in the Edge or in the Cloud. The UAVs use an out-of-band long-range wireless channel (dotted orange line) to exchange information with the Central Station.	78

4.12 Theoretical SNR and capacity values of the wireless links established between each FAP and the FGW, obtained by solving the optimization problem defined in Eq. (4.1). The colors of the markers that represent the FAPs are associated with the order of magnitude of the SNR values of the wireless links established with the FGW, considering the depicted color bar.	83
4.13 Reference scenario depicting the FGW Placement Subspace (S_G) in a 2D space, which results from the intersection of the circumferences, centered at each FAP, with radius equal to the maximum distance compliant with the minimum SNR required.	86
4.14 Information exchanged between the FMN, the FAP placement algorithm, and the Gateway Placement and Queue Management (GPQM) algorithm. The circled numbers from 1 to 5 define the sequential order considered. The FMN, including the UAVs positions and queue size, is reconfigured each $\Delta t \gg 1$ s.	87
4.15 Experimental SNR versus the Line-of-Sight distance of the wireless link, considering different UAV altitudes. The Free-Space Path Loss (FSPL) (Friis) and two-ray ground reflection models are represented by the dashed and dotted lines, respectively [21].	90
4.16 Wireless channel capacity modeled by a linear regression between the minimum SNR values and the data rate associated to the IEEE 802.11ac Modulation and Coding Scheme (MCS) indexes, considering 3 FAPs using the same wireless channel.	92
4.17 Performance results for 3 static FAPs (reference scenario of Fig. 4.13), considering UDP Poisson traffic.	95
4.18 Performance results for 3 static FAPs (reference scenario of Fig. 4.13), considering UDP <i>OnOff</i> traffic.	95
4.19 Performance results for 3 static FAPs (reference scenario of Fig. 4.13), considering TCP <i>BulkSend</i> traffic.	95
4.20 Performance results for 3 moving FAPs, under 5 random dynamic networking scenarios, considering UDP <i>OnOff</i> traffic.	96
4.21 Performance results for 3 moving FAPs, under 5 random networking scenarios, considering TCP <i>BulkSend</i> traffic.	97
4.22 Performance results for 6 moving FAPs, under 5 random networking scenarios, considering UDP Poisson traffic.	97
4.23 Performance results for 12 moving FAPs, under 5 random networking scenarios, considering UDP <i>OnOff</i> traffic.	97
4.24 Performance results for 12 moving FAPs, under 5 random networking scenarios, considering TCP <i>BulkSend</i> traffic.	98
4.25 Benchmarking results considering the FGW position, $P_0 = (x_0, y_0, z_0)$, the transmission power of the UAVs, P_T , and the queue size for each FAP $_i$, Q_i , determined by means of 1) the GPQM algorithm (green circles), 2) the gradient-based solver (solid black lines), and 3) the gradient-free solver (dashed yellow lines), considering the optimization problem formulated in Eq. (4.1), for 3, 6, and 12 FAPs generating UDP Poisson traffic, under 5 random networking scenarios for each number of FAPs. . .	98

5.1	Illustrative networking scenario of single-hop routing, where a UAV collects data from ground nodes and carries it to the destination by flying [22].	104
5.2	Illustrative networking scenario of multi-hop routing, where two UAVs forward data between multiple ground devices and a ground gateway [22].	106
5.3	Modified structure of the Optimized Link State Routing (OLSR)'s <i>HELLO</i> message. Information fields added or modified in the Predictive Optimized Link State Routing (P-OLSR) protocol implementation are highlighted in gray, including altitude, latitude, longitude, and the average relative speed associated with the communications nodes [23].	107
5.4	Illustration of SDN applied to a generic mobile networking scenario [24]. Considering an SDN-based architecture, the Temporospacial-SDN approach proposed in [25] takes advantage of the holistic view available in the SDN controllers to predict the time-varying network's conditions.	108
5.5	Routing algorithm based on Directional Particle Swarming Optimization (DPSO) [26].	110
5.6	Wireless communications paths defined by the Mobility Management Unmanned Aerial Vehicle-based Grouping (MMUG) routing protocol. Different UAV clusters are circled by the blue, yellow, orange, and green solid lines. The blue UAV cluster includes the source node, which is referred to as the Group Head Source Identifier (GH-S-ID), whereas the green UAV cluster includes the Group Head Destination Identifier (GH-D-ID) [27].	110
5.7	Routing discovery phase of the routing scheme based on fuzzy logic proposed in [28], where: 1) route delay ($Delay_{Route}$), initially set to zero, 2) route fitness (F_{Route}), equal to the lowest score among all UAVs in the route, and 3) $HopCount$, starting at zero, are considered as inputs to determine the $BestRoute$; they are updated at each hop. $BestRoute$ is the communications path with the highest fitness, the least number of hops, and the minimum delay.	112
5.8	Architecture of the two-phase topology control and the <i>Q</i> -learning based Topology-aware <i>Q</i> -Routing (TAQR) module of the Joint Topology Control and Routing (JTCR) protocol proposed in [29].	113
5.9	SDN-based flying network considered by the Optimized Predictive and Adaptive Routing (OPAR) and Load-Balanced Optimized Predictive and Adaptive Routing (LB-OPAR) algorithms. Messages are periodically exchanged between the controller and the UAVs. Based on three consecutive positions received from each UAV, the controller predicts the link lifetime, calculates the traffic load for each node, and solves an optimization problem to define the communications paths [30].	114
5.10	3-Dimensional (3D) <i>Q</i> -learning routing protocol, which combines Hop-By-Hop and Store-Carry-Forward approaches. The routing decisions are taken by a <i>Q</i> -learning based routing decision scheme, consisting of link quality prediction and routing decision modules [31].	116
5.11	Forwarding table calculation made by the PreQ-Learning routing algorithm [32]. . .	117
5.12	Deep Reinforcement Learning-based Geographical Routing protocol's architecture, depicting the sensing, learning, and routing decision phases [31].	118

5.13 Illustration of a disaster management scenario considered by the Packet Arrival Prediction routing protocol, where the red area has been subjected to an earthquake. UAVs are used to extend wireless connectivity from a remote BS. The transmission queue backlog is shown by the progressive bars above each UAV. Red denotes a long queue backlog, while yellow and green denote medium and small queue backlogs, respectively. Forwarding data traffic through a UAV with a red bar may lead to high PLR [33].	119
5.14 Deep Neural Network Routing (DNNR) solution proposed in [34].	121
5.15 Q-learning-based Topology-aware Routing (QTAR) proposed in [35]. QTAR includes modules for obtaining and estimating the communications nodes' location and for neighbor discovery, which allow building neighbor tables. Q-learning is used in the route discovery process, considering the two-hop neighbor information gathered.	122
5.16 Q-learning routing strategy employed by the Topology-Aware Resilient Routing based on Adaptive Q-learning (TARRAQ) protocol proposed in [36].	123
5.17 Representative networking scenario consisting of three nodes connected by means of asymmetric wireless links. Strong black line arrows are associated with a good connection (i.e., low PLR), while thin black line arrows are associated with a bad connection (i.e., high PLR). Better Approach To Mobile Ad hoc Networking (B.A.T.M.A.N.) IV rebroadcasts the Originator Messages (OGMs) with the highest Transmission Quality (TQ) value with respect to the originator. Herein, node B sends the TQ value received via node C in the OGMs associated with node A [37].	125
5.18 Distance between a moving communications node and two static communications nodes over time, in linear and logarithmic scale. Linear scale is the most suitable to evaluate distance changes.	128
5.19 Scenario used as reference case. For the initial position of FAP ₃ , the wireless links between FAP ₃ and both FAP ₂ and FGW are not available, due to the SNR threshold constraint.	130
5.20 Theoretical capacity values for the paths between FAP ₃ and the FGW. In order to reach the FGW, FAP ₃ uses as relay nodes FAP ₁ until $t = 60$ s and FAP ₂ from $t = 60$ s to $t = 130$ s. The path with the highest capacity at each instant is highlighted by the circle symbol.	130
5.21 Simulation scenarios considered for the Routing for Flying Multi-hop Networks (Re-deFINE) mechanism evaluation.	133
5.22 Scenario A: throughput results measured in the FGW, considering UDP Poisson traffic.	134
5.23 Scenario A: Packet Delivery Ratio (PDR) results measured in the FGW, considering UDP Poisson traffic.	135
5.24 Scenario A: delay CDF measured in the FGW, considering UDP Poisson traffic with $\lambda=73$ Mbit/s.	135
5.25 Scenario B: throughput results measured in the FGW, considering UDP Poisson traffic.	136
5.26 Scenario B: PDR results measured in the FGW, considering UDP Poisson traffic.	137
5.27 Scenario B: delay CDF measured in the FGW, considering UDP Poisson traffic with $\lambda=73$ Mbit/s.	137

5.28	Network graph illustrating the Transmitter-Receiver Conflict Avoidance interference model. If FAP ₈ is used as relay node, the network performance will be reduced up to 50%, since FAP ₈ is in the interference range of FAP ₂ and FAP ₆	139
5.29	Results for throughput and delay in the FGW. The results were obtained considering 5 FAPs generating traffic, and $\alpha = 1$	143
5.30	Results for throughput and delay in the FGW. The results were obtained considering 10 FAPs generating traffic, and $\alpha = 1$	144
5.31	The 25 th , 50 th , and 75 th percentiles of both the throughput Complementary Cumulative Distribution Function (CCDF) and delay CDF, considering different α values. The results consider 5 FAPs generating traffic.	144
6.1	Flying network, consisting of UAVs, used by a Mobile Network Operator to provide network slices to a Service Provider that offers services to ground devices in target geographical areas [38].	150
6.2	Illustration of IEEE 802.11e Enhanced Distributed Channel Access, where four Access Categories with different priorities (descending from the left-hand side to the right-hand side) are associated with voice (VO), video (VI), best-effort (BE), and background (BK) traffic (modified from [39]).	152
6.3	Network slicing architecture where three network slices share the physical network resources available, including a BS with cache storage connected to the Core network through a capacity-constrained backhaul link. The constants α , β , and γ represent the fractions of, respectively, total wireless channel bandwidth (in Hz), cache storage (contents), and backhaul link capacity (in bit/s) allocated to each network slice [40].	153
6.4	Illustrative resource allocation approach with different isolation levels proposed in [41] for three network slices.	154
6.5	Cooperative network consisting of a High-Altitude Platform Station (HAPS) and Low Altitude Platforms (LAPs) extending wireless coverage and improving the network performance offered to ground devices.	156
6.6	Network slicing-aware architecture proposed in [42]. It consists of UAVs, acting as Aerial Relay Nodes, and Road-Side Units (RSUs), managed by an SDN controller, which allow providing Resource Blocks (RBs) to ground devices.	157
6.7	UAV offering Machine-Type Communications (MTC) and broadband services to ground devices [43].	158
6.8	Illustration of the iterative optimization approach with a successive convex approximation technique proposed in [44] for slicing-aware flying networks.	159
6.9	Two-tier heterogeneous RAN where a single Macro BS (MBS), deployed at the center of the network, is aided by multiple Drone Small Cells (DSCs). Due to the higher Macro BS's transmission power, it provides wireless coverage over a large area; the Drone Small Cells, which allow for a short wireless range due to their lower transmission power, are placed within the 3D coverage area enabled by the Macro BS [45].	160
6.10	UAV, acting as a flying communications relay, enabling Ultra-Reliable Low-Latency Communications (URLLC) for a ground device with Line-of-Sight towards a BS blocked by an obstacle.	161

6.11	Wireless network consisting of a single UAV providing a URLLC-based communications service to a set of ground devices [46].	162
6.12	Network slicing-aware system model considered in [47], where a set of Service Providers rent wireless communications resources from a set of Mobile Network Operators. The wireless communications resources are made available by each Mobile Network Operator using a single UAV.	163
6.13	Network slicing approach proposed in [48], where two multiplexing methods – Dynamic Multiplexing and Orthogonal Slicing – are employed to enable URLLC and enhanced Mobile Broadband (eMBB) communications services.	164
6.14	Slicing-aware FMN consisting of UAVs. The FAPs, in the first tier, provide a wireless access network to the ground devices, while the FGWs, in the second tier, establish a wireless backhaul network with the FAPs and forward the traffic to/from the Internet. The UAV placement and communications resources allocation are jointly computed in a Central Station, deployed in the Cloud or at the Edge of the FMN.	165
6.15	Illustrative networking scenario (3D representation on the left-hand side) with multiple ground devices (blue squares) served by an FMN consisting of a FAP (green star) and an FGW (orange circle). The ground devices are associated with two network slices $s \in S$ made available in different subareas $a \in A^s$ at the base of cuboid C (2D representation on the right-hand side).	166
6.16	Wireless channel capacity modeled by linear regressions between the SNR and the data rate associated to the IEEE 802.11ac MCS indexes, considering channel bandwidth of 20 MHz.	168
6.17	Resource allocation performed by Slicing-aware Flying Access and Backhaul (SurFABLE), considering the minimum channel bandwidth equal to 20 MHz [38].	175
6.18	Simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP Poisson traffic.	182
6.19	Simulation average performance results for five networking scenarios consisting of 20 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP Poisson traffic.	183
6.20	Simulation average performance results for five networking scenarios consisting of 40 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP Poisson traffic.	184
6.21	Simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering TCP traffic.	185
6.22	Simulation average performance results for five networking scenarios consisting of 20 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering TCP traffic.	186
6.23	Simulation average performance results for five networking scenarios consisting of 40 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering TCP traffic.	187

6.24 Average bandwidth and number of UAVs used in networking scenarios consisting of 5, 20, and 40 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, including 95% confidence intervals.	188
6.25 Ground devices deployed in the experimental testbed.	189
6.26 Experimental wireless channel capacity modeled by linear regressions between SNR and data rates obtained in practice. The data labels at each marker refer to the Euclidean distances, in meters, between sender and receiver considered in the experimental measurements.	190
6.27 Experimental and simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (4 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP Constant Bit Rate (CBR) traffic and channel bandwidth of 60 MHz.	192
6.28 Experimental and simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (4 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP CBR traffic and channel bandwidth of 80 MHz.	193
A.1 Admissible FGW positions for maximum and minimum overall capacity $C(t_k)$ values (red square and green circle, respectively).	210

List of Tables

2.1	Data rate and minimum Signal-to-Noise Ratio (SNR) associated with Institute of Electrical and Electronics Engineers (IEEE) 802.11ac's Modulation and Coding Scheme (MCS) indexes 3 and 8 (channel bandwidth: 160 MHz and Guard Interval: 800 ns).	38
2.2	Summary of the main Network Simulator 3 (ns-3) simulation parameters used in the evaluation of the GateWay Placement (GWP) algorithm.	41
3.1	Relation between the minimum Received Signal Strength Indicator (RSSI), data rate of the targeted MCS indexes, and the link capacity values for 4 Flying Access Points (FAPs) sharing the transmission time, while using the IEEE 802.11ac standard; they were obtained through ns-3 simulations.	59
3.2	Summary of the main ns-3 simulation parameters used in the evaluation of the Proactive Queue Management (PQM) mechanism.	60
4.1	Minimum SNR required to select the IEEE 802.11ac MCS indexes 2, 5, and 7, and the corresponding fair share for 3 FAPs. A factor $\eta = 0.80$ is considered to take into account the efficiency of the Medium Access Control (MAC) protocol.	82
4.2	Summary of the main ns-3 simulation parameters used in the evaluation of the Gateway Placement and Queue Management (GPQM) algorithm.	94
5.1	Summary of the main ns-3 simulation parameters used in the evaluation of the Routing for Flying Multi-hop Networks (RedeFINE) mechanism.	132
6.1	Positions and network slice types associated with the ground devices considered in the performance evaluation carried out using ns-3 simulations and an experimental testbed.	177
6.2	Networking parameters used in the evaluation of the Slicing-aware Flying Access and Backhaul (SurFABLE) algorithm. IEEE 802.11 was used due to its ubiquity, cost-effectiveness, and flexibility for evaluating network performance in both simulation and experimentation, using the same exact conditions. Yet, the problem formulation and SurFABLE are agnostic to the technology used.	178
6.3	Networking parameters considered in the evaluation of the SurFABLE algorithm using ns-3 and experimental testbed.	179

List of Acronyms

pi Prediction Interval

si Sampling Interval

2D 2-Dimensional

3D 3-Dimensional

3GPP 3rd Generation Partnership Project

AODV Ad hoc On-Demand Distance Vector

AP Access Point

AQM Active Queue Management

B.A.T.M.A.N. Better Approach To Mobile Ad hoc Networking

BER Bit Error Rate

BS Base Station

CAPEX Capital Expenditure

CBR Constant Bit Rate

CCDF Complementary Cumulative Distribution Function

CDF Cumulative Distribution Function

CHU Cluster Head Unmanned Aerial Vehicle

CoDel Controlled Delay

CPE Customer Premises Equipment

CSMA/CA Carrier Sense Multiple Access with Collision Avoidance

eMBB enhanced Mobile Broadband

ETT Expected Transmission Time

ETX	Expected Transmission Count
FAP	Flying Access Point
FGW	Flying Gateway
FMAP	Flying Mesh Access Point
FMN	Flying Multi-hop Network
FSPL	Free-Space Path Loss
GPQM	Gateway Placement and Queue Management
GPS	Global Positioning System
GWP	GateWay Placement
HAPS	High-Altitude Platform Station
I2R	Inter-flow Interference-aware Routing
IAB	Integrated Access and Backhaul
IAWARE	Interference Aware
ICMP	Internet Control Message Protocol
IEEE	Institute of Electrical and Electronics Engineers
INESC TEC	Institute for Systems and Computer Engineering, Technology and Science
IoT	Internet of Things
IP	Internet Protocol
IPOPT	Interior Point OPTimizer
LAP	Low Altitude Platform
LB-OPAR	Load-Balanced Optimized Predictive and Adaptive Routing
LEPR	Link Stability Estimation-based Preemptive Routing
MAC	Medium Access Control
MADRL-SA	Multi-Unmanned Aerial Vehicle Deep Reinforcement Learning Scheduling Algorithm
MCS	Modulation and Coding Scheme
MIC	Metric of Interference and Channel-switching
mIoT	massive Internet-of-Things

ML	Machine Learning
mMTC	massive Machine-Type Communications
MMUG	Mobility Management Unmanned Aerial Vehicle-based Grouping
mmWave	Millimeter Wave
MPA	Multi Unmanned Aerial Vehicle-based Particle Swarming Optimization Approach
MTC	Machine-Type Communications
MU	Mobile User
NIC	Network Interface Card
ns-3	Network Simulator 3
O-RAN	Open Radio Access Network
OFDMA	Orthogonal Frequency-Division Multiple Access
OGM	Originator Message
OLSR	Optimized Link State Routing
OMNET++	Objective Modular Network Testbed in C++
OPAR	Optimized Predictive and Adaptive Routing
OPEX	Operational Expenditure
P-OLSR	Predictive Optimized Link State Routing
PARRoT	Predictive Ad-hoc Routing by Reinforcement Learning and Trajectory knowledge
PBARP3D	3D Position Based Adaptive Real-time routing Protocol
PDR	Packet Delivery Ratio
PI	Proportional Integral
PIE	Proportional Integral controller Enhanced
PLR	Packet Loss Ratio
PQM	Proactive Queue Management
PSO	Particle Swarm Optimization
QoS	Quality of Service
QTAR	Q-learning-based Topology-aware Routing

RAN	Radio Access Network
RB	Resource Block
RED	Random Early Detection
RedeFINE	Routing for Flying Multi-hop Networks
REST	Representational State Transfer
RSSI	Received Signal Strength Indicator
RSU	Road-Side Unit
RTT	Round-Trip Time
SDN	Software-Defined Networking
SEEA	Single Unmanned Aerial Vehicle Exploration-Exploitation deployment Approach
SLA	Service Level Agreement
SNR	Signal-to-Noise Ratio
SPA	Single Unmanned Aerial Vehicle Particle Swarming Optimization deployment Approach
STA	Station
SurFABLE	Slicing-aware Flying Access and Backhaul
TAQR	Topology-aware Q-Routing
TARRAQ	Topology-Aware Resilient Routing based on Adaptive Q-learning
TCP	Transmission Control Protocol
TQ	Transmission Quality
TX	Transmission
UAV	Unmanned Aerial Vehicle
UDP	User Datagram Protocol
URLLC	Ultra-Reliable Low-Latency Communications
VFMC	Virtual Force-based Mobility Control
WCETT	Weighted Cumulative Expected Transmission Time

List of Symbols

A	Area forming the base of cuboid C
A^s	Set of subareas $a \in A$ served by the network slice s
B	Constant channel bandwidth, in Hz
c	Speed of light in vacuum, in m/s
C	Cuboid within which the Unmanned Aerial Vehicles (UAVs) can be positioned
$c_{i,j}$	Bidirectional network capacity, in bit/s, provided by a given channel bandwidth block made available by UAV _{i} to node j
$C_{i,j}$	Capacity, in bit/s, of the wireless link available between UAV _{i} and node j
C^{MAX}	Maximum capacity, in bit/s, of the wireless communications channel considered
δ_t	Period considered by RedeFINE for updating forwarding tables, in seconds
Δt	Network reconfiguration period, in seconds
D_i	Average delay, in seconds, of the packets received from UAV _{i}
d_i^{MAX}	Maximum distance, in meters, between UAV _{i} and node j to ensure the selection of MCS _{i,j}
E	Set of directional wireless links available in the Flying Multi-hop Network (FMN)
E_T	Active wireless links in the FMN
f_i	Carrier frequency, in Hz, associated with UAV _{i}
F^f	Cost of FAP _{f}
F^g	Cost of FGW _{g}
$F_{i,j}$	Traffic flow transmitted by UAV _{i} towards node j
G	Directed graph representing the FMN
H	Maximum average delay allowed, in seconds
H^s	Maximum average packet delay, in seconds, of network slice s
f, g, i	Indexes associated to respectively Flying Access Point (FAP), Flying Gateway (FGW), and UAV

a, j, s	Symbols representing respectively a subarea, a node (from the set of FAPs, FGWs, and ground devices), and a network slice
L	Reference fair share, in bit/s, of the wireless channel's capacity
M	Distance, in meters, to be ensured between the position of the FGW and the position of any other UAV _{<i>i</i>} in the FMN
$\text{MCS}_{i,j}$	Modulation and Coding Scheme (MCS) index used in the wireless link established between UAV _{<i>i</i>} and node <i>j</i> . Index <i>j</i> is not specified when UAV _{<i>i</i>} is directly connected to UAV ₀ (i.e., one-hop wireless link)
N	Maximum number of UAVs in the FMN
P_a	Position of center of subarea $a \in A$
P_i	Position of UAV _{<i>i</i>}
P_N	Constant noise power, in dBm
$P_{R_{i,j}}$	Power received, in dBm, from UAV _{<i>i</i>} at node <i>j</i>
P_T	Transmission power, in dBm, of the UAVs
Q_i	Queue size, in packets, of UAV _{<i>i</i>}
r_a	Ground device located in center of subarea $a \in A$
R_i	Bitrate, in bit/s, of the traffic flow received from UAV _{<i>i</i>}
S	Packet size, in bytes
S_G	FGW placement subspace
SNR_i	Minimum Signal-to-Noise Ratio (SNR), in dB, required for selecting the targeted MCS index
$\text{SNR}_{i,j}$	SNR received, in dB, at node <i>j</i> from UAV _{<i>i</i>}
$T(I, E_T)$	Active topology of the FMN
T_i	Bitrate, in bit/s, of the traffic flow transmitted by UAV _{<i>i</i>}
t_k	Time instant <i>k</i> , in seconds, corresponding to a snapshot of the FMN
T^s	Minimum average data rate, in bit/s, of network slice <i>s</i>
UAV_i	Unmanned Aerial Vehicle <i>i</i>
V	Set of UAVs composing the FMN
W_i	Total number of wireless channels available at UAV _{<i>i</i>}
X, Y, Z	Length, width, and height of <i>C</i> , respectively, in meters
x_i, y_i, z_i	3-Dimensional (3D) Cartesian coordinates of UAV _{<i>i</i>}

Chapter 1

Introduction

1.1 Scope

The evolution of wireless networks over the years has been associated with an exponential growth in the usage of mobile devices and the emergence of robotics, automation, and Artificial Intelligence in different areas of society. While the first generations of wireless networks were designed to enable communications services for people, emerging wireless networks aim at paving the way for billions of machines connected to each other and interacting with humans [1]. Coupled with the growing usage of a myriad of communications services and applications in an ever-connected society, including augmented and virtual reality, ultra-high definition video streaming, and holographic-type and mission-critical communications, this is pushing the capacity considered in the design of existing Radio Access Networks (RANs) and backhaul networks to the limit. The challenges are exacerbated in disaster management scenarios, such as forest fires and floods, as well as in temporary crowded events, including music festivals and parades, where the fixed telecommunications infrastructures may collapse or may not meet the heterogeneous Quality of Service (QoS) levels required by dynamic and highly concentrated mobile devices.

Besides the large amount of traffic that the RANs and backhaul networks have to accommodate, the ability to ensure ultra-low latency communications is a challenge in the design of emerging wireless networks. The goal is to reliably exchange data with controlled delays so that communications between billions of interconnected machines are possible and humans are able to interact with remote devices in real-time, which includes not only the ability to see and hear the surrounding environment, but also touch and feel real and virtual objects. The interaction between humans and remote devices is only intuitive and natural if the feedback provided by the latter is compliant with the human reaction times. Typical applications include immersive multimedia services, interactive tactile Internet applications, and reliable and latency-sensitive remote control of machines [1, 49]. These requirements are envisioned in the recently released 5G networks

and the emerging 6G networks, which aim at enabling massive Machine-Type Communications (mMTC) [50, 51], massive Internet-of-Things services [52–54], Ultra-Reliable Low-Latency Communications (URLLC) [48, 55], and enhanced Mobile Broadband (eMBB) [56, 57]. Ensuring these services demands ubiquitous wireless connectivity and enhanced network performance, based on the target requirements depicted in Fig. 1.1 for 5G and 6G. This will require flexible wireless networks capable of being dynamically reconfigured and deployed on-demand.

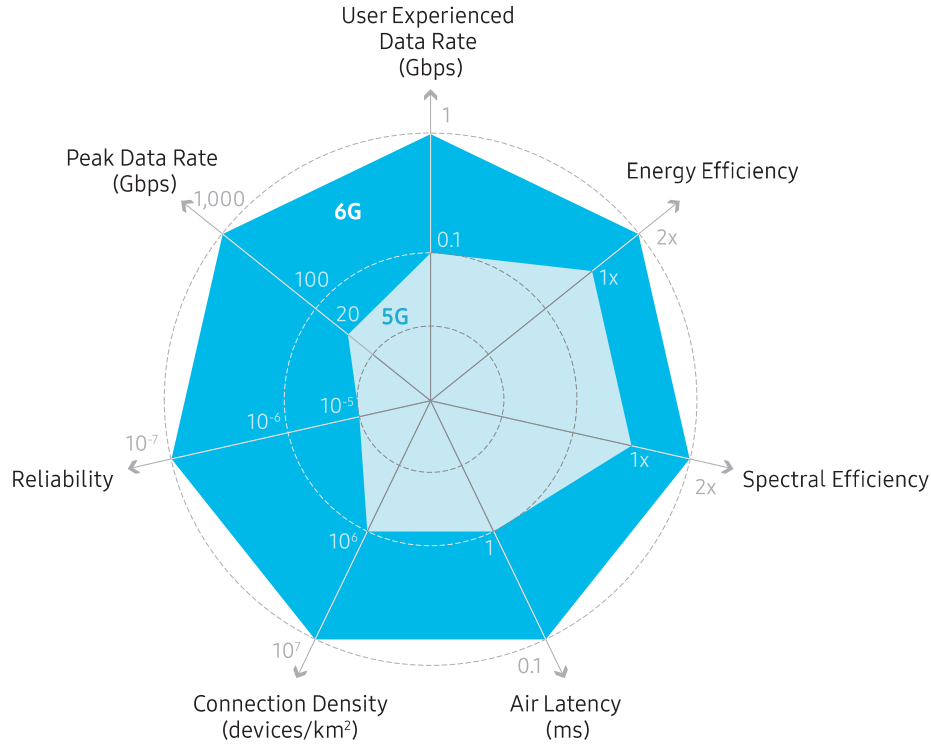


Figure 1.1: Requirements of the recently released 5G networks (in light blue) and the emerging 6G networks (in dark blue) [1].

At the same time, the usage of Unmanned Aerial Vehicles (UAVs) has become popular in several military and civilian applications [58]. Some examples are surveillance and reconnaissance, fire-fighting and rescue, remote sensing and exploration, logistics, and payload transportation [59]. A UAV is an aircraft that does not require a human on-board to carry out its control. It can be configured to follow a set of predefined waypoints or be remotely controlled in real-time. Typically, UAVs can be categorized as either Fixed-Wing Vehicles or Rotor-Wing Vehicles. Fixed-Wing Vehicles are not able to maintain a fixed position and need a platform for takeoff and landing; Rotor-Wing Vehicles are able to hover, while maintaining its position in a fixed location, and their takeoff and landing are performed vertically, which avoids using a dedicated platform. The advent of small and low-cost UAVs, associated with the electronics miniaturization, has paved the way

for the usage of swarms of UAVs. In a swarm, UAVs can cooperatively perform missions more efficiently when compared with UAVs working independently.

The capability of UAVs to be used on-demand as well as their ability to hover, move in 3-Dimensional (3D) space, and carry payload on-board make them suitable platforms to act as communications nodes, including Wi-Fi Access Points (APs) and cellular Base Stations (BSs) [60]. This enables the quick deployment of Flying Multi-hop Networks (FMNs) capable of providing wireless coverage and reinforcing the network capacity anywhere, anytime [22, 61–67]. Despite the potentially improved performance achieved with FMNs, they increase the complexity for the proper management of the UAVs, namely for defining their positions and the radio resources made available.

1.2 Motivation

The growing traffic demand has led to a significant revolution in the worldwide telecommunications infrastructures. Over the years, communications networks have evolved from circuit-switched to packet-switched networks; in particular, cellular networks have progressively evolved from 2G to 5G, targeting broadband, delay-sensitive, and ultra-reliable communications. At the same time, Mobile Network Operators have been pursuing scalable and flexible approaches for deploying wireless networks consisting of the minimum number of cellular BSs and APs, in order to meet different use cases' requirements with reduced Capital Expenditure (CAPEX) and Operational Expenditure (OPEX) [68]. Overall, besides the capacity for handling large amounts of traffic, RANs and backhaul networks are expected to meet the differentiated traffic requirements imposed by emerging communications services and applications, including high throughput and low-latency.

Current Mobile Network Operators' infrastructures are limited in providing ubiquitous and cost-effective wireless connectivity. They are designed to provide wireless connectivity for a given density of users per area unit and are deployed only where users are expected to be present. Remote areas characterized by a low user density typically have many zones without adequate wireless coverage. Moreover, the flexibility, availability, reliability, and capacity of current Mobile Network Operators' infrastructures may not be sufficient to meet the QoS levels expected by the network users in a myriad of application scenarios [69]. In disaster management scenarios, first responders require reliable mission-critical communications, including group communications and rich-media communications (e.g., image and video); they are increasingly relevant for enhanced situational awareness and cooperation between first responders in the theater of operations and the command centers. On the other hand, a best-effort wireless connectivity service may be sufficient to serve the victims. In temporary crowded events, thousands of people with personal devices seek ubiquitous Internet access so that they can use different communications services

and applications, such as social networks and streaming services, and exchange rich-media content. Wireless connectivity that allows meeting differentiated QoS requirements must be ensured for the staff, the media, and the spectators. Nowadays, the network requirements associated with these networking scenarios are addressed by Mobile Network Operators by means of network over-provisioning using fixed and mobile BSs deployed on ground vehicles, using the Cell-on-Wheels concept [70]. Fixed telecommunications infrastructures may collapse, fail to provide sufficient wireless coverage, or lack the required communications resources. Since the traffic generated by mobile devices is highly variable and non-uniform, static wireless networks may be either not enough to meet the targeted requirements or imply the waste of resources during periods of under-utilization. BSs deployed on ground vehicles provide more flexibility than fixed network infrastructures, but their movement is limited to roads and parking spaces. Moreover, they are expensive and limited in number, and imply determining an adequate position before moving them to the site so that their usage is effective in practice. Finally, the unpredictability associated to the number of devices and their mobility patterns in some scenarios may degrade the QoS offered when using fixed and mobile BSs mounted on ground vehicles.

FMNs, referred to in Section 1.1, represent a promising solution to enable wireless communications in disaster management scenarios and temporary crowded events. In disaster management scenarios, an FMN providing dynamic wireless connectivity according to the traffic demand, as depicted in Fig. 1.2, makes it possible using devices with communications capabilities by first responders and victims, while improving disaster management and human protection. In temporary crowded events, the dynamic crowd movements and variable traffic demand make FMNs a suitable solution to provide on-demand traffic-aware Internet access, as depicted in Fig. 1.3. Besides their flexibility and capability for being deployed on-demand, FMNs have significant advantages with respect to terrestrial networks. In particular, in FMNs the wireless links are characterized by a strong Line-of-Sight component, promoted by the altitude factor, which allows avoiding obstacles to the radio signal propagation. Considering these characteristics, UAVs are a promising and cost-effective solution to extend and reinforce existing network deployments on-demand. The lack of solutions addressing the management of communications resources in FMNs consisting of UAVs, the recent popularity of UAVs as well as the emerging Non-Terrestrial networking paradigm under discussion in 3rd Generation Partnership Project (3GPP) and other fora, including 6G, motivate the development of novel networking solutions for FMNs.

1.3 Problem Statement

In order to support heterogeneous communications services and applications, a possible approach is to deploy multiple independent wireless networks, each with custom configurations and network topologies so that the targeted QoS requirements associated with different use cases are

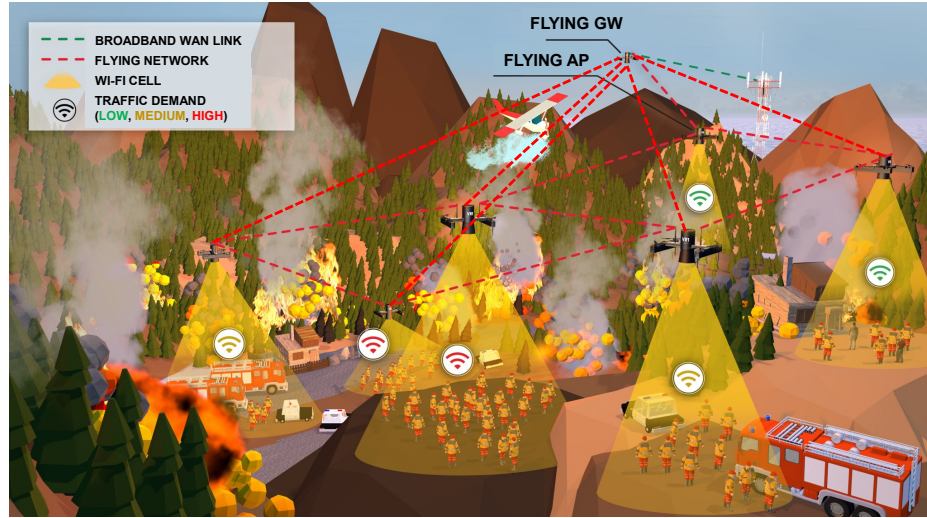


Figure 1.2: A multi-tier FMN consisting of UAVs enabling communications in a disaster management scenario.

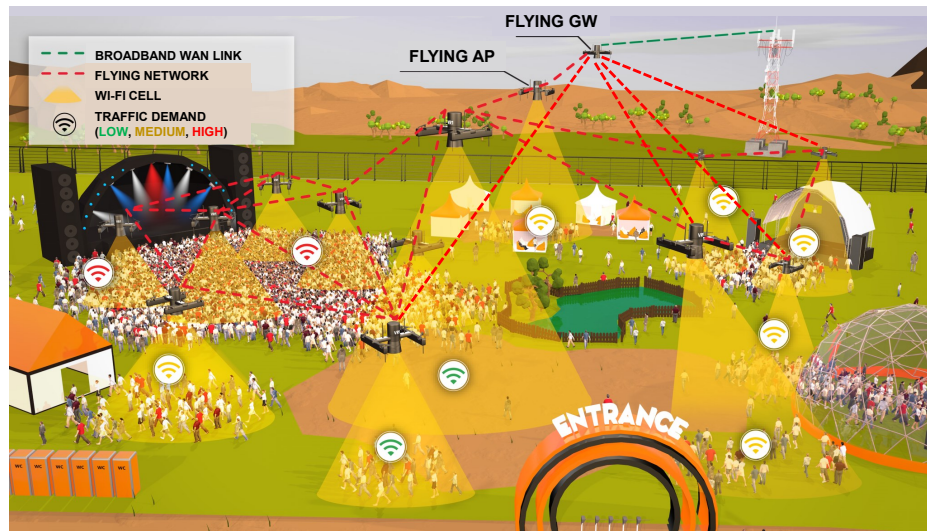


Figure 1.3: A multi-tier FMN consisting of UAVs providing Internet access in a music festival.

met. However, this is not a scalable approach and involves high CAPEX and OPEX. For instance, disaster management scenarios are unpredictable and may occur in remote areas, while temporary crowded events are typically short-lived and only take place at a very specific time of the year. For these reasons, static wireless network infrastructures are not the most suitable solution. On the other hand, the usage of fixed wireless communications infrastructures, such as Wi-Fi APs and cellular BSs, requires planning their location in advance considering a target network capacity for serving the maximum traffic expected. This may lead to different network planning challenges,

including: 1) under-used wireless networks; 2) capacity problems, if traffic exceeds the amount considered in the design phase; and 3) wireless coverage and network capacity problems, if the network infrastructures are damaged due to hostile conditions experienced in disaster management scenarios and dynamic behavior in temporary crowded events. Despite Mobile Network Operators may use mobile BSs mounted on ground vehicles, using the Cell-on-Wheels concept to deploy wireless networks on-demand and reinforce network capacity, they suffer from the same network planning challenges associated to fixed wireless communications infrastructures. Also, their limited mobility is not compliant with flexible and dynamic wireless networks over time.

FMNs consisting of UAVs have been proposed in the literature as a flexible approach to provide wireless connectivity on-demand. Fig. 1.2 and Fig. 1.3 illustrate a typical FMN architecture consisting in UAVs acting as Flying Access Points (FAPs) and Flying Gateways (FGWs) that cooperate with each other, in order to enable wireless coverage extension from a remote Mobile Network Operator's infrastructure. Cooperation requires reliable and stable wireless links between the UAVs themselves, and between the UAVs and the ground devices served. This is a challenge due to the highly dynamic nature of FMNs, especially when the UAVs move for adjusting their placement in 3D, in order to better serve the ground devices. Due to the UAV movement, degradation of the wireless links' Signal-to-Noise Ratio (SNR) occurs, frequent path disruptions arise, and the offered QoS is degraded. To address this problem, probe packets may be exchanged with high frequency for estimating the wireless links' quality over time. However, such approach introduces high network overhead, reduces the bandwidth available for exchanging data traffic, and is not scalable. In order to avoid wireless link disruptions, highly dense FMNs consisting of UAVs close to each other may also be considered for ensuring short wireless links with high SNR. However, this deployment approach leads to increased inter-flow interference¹ and requires a higher number of UAVs, thus leading to higher CAPEX and OPEX and increasing the complexity of the UAV placement problem.

The research on UAV placement has been focused on enhancing the wireless coverage, the number of ground devices served as well as improving the aggregate QoS offered by FAPs forming RANs. On the one hand, not considering nor differentiating the QoS levels offered to ground devices may result in the network capacity being insufficient to carry the offered traffic or not being fully exploited by ground devices with less demanding QoS requirements. On the other hand, even though the ground devices are directly affected by the QoS offered by the RAN, the back-haul network needs to be carefully designed so that heterogeneous performance guarantees are ensured. In addition to the ability for accommodating the offered traffic, FMNs providing wireless connectivity are expected to meet the strict low-latency requirements imposed by emerging communications services and applications. In order to enable low-latency communications, placing

¹Inter-flow interference occurs due to neighboring communications nodes competing for the same shared wireless channel.

UAVs close to the devices to be served, while enabling short-distance Line-of-Sight wireless links, represents an advantage over ground wireless networks with limited mobility. Yet, unprecedented challenges from the network configuration point of view also arise.

A critical aspect that must be addressed is the size of the queues configured in the communications nodes, which are used for accommodating transient traffic bursts offered to the network. A baseline approach lies on considering oversized queues, in order to minimize the number of dropped packets. The main disadvantage is the increased packet delay, leading to the well-known bufferbloat problem [71]. Conversely, small queues may be unable to accommodate the traffic offered by the devices, especially bursty traffic. When compared with wireless networks in general, FMNs bring up significant additional challenges, including: 1) potential increase in inter-flow interference, due to enhanced carrier-sense range induced by the altitude factor and lack of obstacles above the ground; and 2) frequent changes in the SNR of the wireless links due to the highly dynamic nature of FMNs, which induce time-varying network capacity. Queues with static size and the state of the art Active Queue Management (AQM) solutions are best suited for wired networks and static wireless networks, for which it is reasonable to assume that the links have approximately constant network capacity; this assumption is not applicable to dynamic FMNs.

Finally, the usage of multiple services with different performance requirements impose significant challenges to Mobile Network Operators, in order to design emerging wireless networks. Despite network slicing has emerged in 5G networks to enable the usage of multiple services with different performance requirements, by means of logical networks that provide specific network capabilities, fixed BSs are typically considered. This may not be feasible in scenarios characterized by limited network resources or dynamic behavior. On the other hand, despite the ongoing research on FMNs consisting of UAVs, existing solutions are focused on extending the wireless coverage and network capacity considering a best-effort approach. In order to meet heterogeneous QoS requirements using a minimum amount of communications resources, a QoS-aware FMN leveraged on UAVs that provide on-demand communications resources to meet the requirements of different network slices is a promising approach. Yet, accomplishing a slicing-aware FMN implies addressing the following research challenges:

- to meet the different QoS levels of multiple network slices on top of a shared physical wireless network infrastructure;
- to determine the minimum number of UAVs required and their placement, in order to simultaneously provide coverage with minimum target SNR values to multiple network slices made available in different geographical areas;
- to achieve a global admissible solution for the placement and allocation of communications resources, in a time interval compliant with the network reconfiguration period, while ensuring the minimum number of UAVs and wireless channels required to meet the QoS levels of

multiple network slices.

The approach followed in this thesis consists in addressing the UAV placement and communications resources' allocation with the ultimate goal of enabling a shared FMN to offer heterogeneous QoS levels.

1.4 Thesis Hypothesis

It is possible using a **dynamic flying multi-hop network**, partitioned into any number and type of **network slices** associated with **variable and heterogeneous QoS** levels, to serve multiple ground devices **without loss of connectivity**.

1.5 Objectives

The main goal of this thesis is the design of a solution addressing the traffic-aware management of communications resources in FMNs consisting of UAVs. The following three major specific objectives are defined:

- to extend wireless coverage and capacity from a remote AP or BS;
- to offer heterogeneous QoS levels to ground devices using the minimum number of UAVs and radio resources;
- to ensure uninterruptible wireless communications paths in dynamic FMNs.

1.6 AM2Fly Overview

This thesis proposes a set of novel Algorithms and Mechanisms acting as Network Functions tailored for service-based FMNs, named AM2Fly. AM2Fly enables the traffic-aware management of communications resources in FMNs consisting of UAVs acting as FAPs and FGWs. The Service Based Architecture concept has been introduced by 3GPP in the design of 5G systems [72]. In a Service Based Architecture, the network is managed by means of Network Functions. Network Functions are self-contained, modular software applications interconnected in a logically shared infrastructure, while providing services and exchanging data with each other. In this thesis we assume an FMN organized in two tiers, as depicted in Fig. 1.4. The first tier comprises the RAN, which consists of FAPs that provide wireless connectivity to ground devices. The second tier consists of FGWs, which are placed above the FAPs. The FGWs are in charge of forwarding traffic between the RAN and the Internet. The FMN is assumed to be controlled by a Central Station, which can be deployed anywhere in the Edge or in the Cloud. An out-of-band long-range

wireless link (e.g., based on Institute of Electrical and Electronics Engineers (IEEE) 802.11ah) is assumed for exchanging information between the Central Station and the UAVs. It enables an always-on control link and avoids introducing overhead in the data plane. The design of the out-of-band communications solution is beyond the scope of this thesis. Given the complexity of this FMN, new solutions are needed to dynamically manage its communications resources.

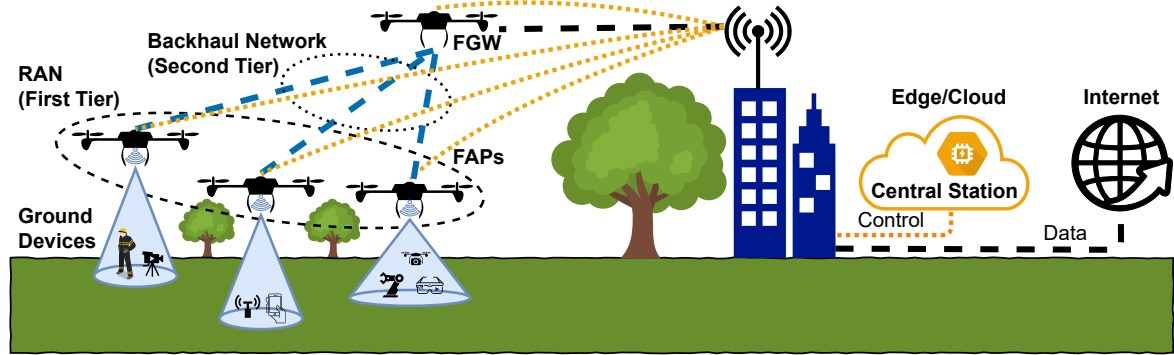


Figure 1.4: A multi-tier FMN consisting of UAVs providing wireless connectivity to ground devices.

Considering the FMN depicted in Fig. 1.4, AM2Fly allows: 1) to determine suitable 3D positions for UAVs acting as FAPs and FGWs; 2) to define adequate communications resources' allocation, in order to meet heterogeneous QoS levels using a shared FMN; and 3) to ensure seamless FMN reconfiguration so that wireless communications paths are always available to forward data traffic between the ground devices and the FGWs. The 3D positioning consists in determining the up-to-date positions for the UAVs according to the positions and the traffic demand of the devices to be served, being them the FAPs in the tier immediately below the FGWs or the ground devices. Based on the up-to-date positions taken for the UAVs, the communications resources' allocation includes defining the network queue sizes, transmission power levels, wireless channel capacity, and number of wireless channels to be used. Considering the actual and future positions of the FAPs and FGWs, determined by the algorithms running in the Central Station, the UAVs' forwarding tables and the time instants they shall be updated are defined, in order to enable a seamless FMN reconfiguration, while maximizing the QoS offered and minimizing the disruptions caused by the movement of the UAVs for repositioning themselves.

The FMN benefits from the synergy between the algorithms and mechanisms running in the Central Station. Each time instant corresponds to a snapshot of the physical FMN, where the FAPs and FGWs are hovering in fixed positions (i.e., zero flying speed), in order to serve the ground devices. The period defined for computing up-to-date decisions takes into account a trade-off between: 1) the stability of the FMN; 2) the time required for computing up-to-date decisions; 3) the time taken by the UAVs to reposition themselves according to the up-to-date decisions; and 4) the network performance gains obtained in practice when up-to-date decisions are considered.

We assume the FMN's update period is a configuration parameter. If the update period is set to a high value, the FMN is more stable but will take a longer time to react to changes in the QoS levels demanded by the ground devices; if the FMN's update period is set to a low value, the FMN reacts quicker to changes in the QoS levels associated to the ground devices, but the FMN is less stable and the FMN's short update period may not be enough to compute up-to-date solutions, which may compromise the performance obtained in practice.

AM2Fly paves the way for more resilient and sustainable wireless communications infrastructures, while avoiding network over-provisioning of static communications resources. This represents a step forward with respect to the current approach typically employed by Mobile Network Operators to reinforce wireless coverage and network capacity in disaster management scenarios and temporary crowded events. Moreover, AM2Fly allows that the same physical shared flying network infrastructure can be used to simultaneously serve different communications services and applications, which may be characterized by heterogeneous targeted QoS requirements.

1.7 Research Methodology

This section describes the research methodology used to address the problem of this thesis and to evaluate the performance of the designed solutions.

1.7.1 AM2Fly Design

Due to the complexity of the problem addressed by this thesis, AM2Fly was developed in two phases. Considering the state of the art solutions available to compute the placement of UAVs acting as APs and BSs, in the first phase we focused on the backhaul network, consisting of the wireless links available between the FAPs serving the ground devices and the FGW used to forward data traffic to/from the Internet.

First, the problem addressed in this thesis consisted in maximizing the aggregate throughput measured in the FGW. For that purpose, an FGW placement algorithm was designed to enable wireless links between the FAPs and the FGW with high enough capacity for accommodating the traffic forwarded to/from the Internet. Delay constraints were not taken into account for placing the FGW. Yet, after concluding that the usage of static size network queues in wireless networks leads to underutilized communications nodes and unnecessary high delays, we designed a proactive queue management mechanism for FMNs, able to define in advance the queue sizes for the FAPs according to the capacity of the wireless links established with the FGW and the traffic demand of the ground devices served.

Afterwards, building upon the algorithm proposed for the FGW placement and the novel mechanism proposed for the proactive queue management in FMNs, which were designed as independent algorithms not benefiting from the gains of each other as a joint solution, we went a step

further and designed an algorithm for jointly defining the placement of the FGW and the suitable queue sizes for the FAPs, in order to reliably carry high bitrates under controlled delays. The FGW placement and queue size management in FMNs allows for wireless links with high enough capacity and network queues with a large enough size so that traffic bursts are accommodated and controlled delays are achieved.

It is worth mentioning that the up-to-date decisions taken by AM2Fly are computed at periodic time instants, each corresponding to a snapshot of the FMN. Considering that moving the UAVs for reconfiguring the FMN can disrupt the wireless links, we designed a predictive routing mechanism for FMNs to enable uninterruptible communications paths between the FAPs and the FGWs. It allows selecting communications paths consisting of wireless links with SNR higher than a targeted level, while allowing for improved QoS offered to the ground devices and minimizing the disruption caused by the movement of the UAVs.

Based on the Service Based Architecture concept, AM2Fly consists of algorithms and mechanisms as Network Functions able to independently manage the communications resources in FMNs. For that purpose, the inputs and outputs of each algorithm and mechanism are combined in order to achieve a solution that takes advantage of the gains achieved by each one. They rely on a best-effort approach, in which a shared wireless channel is employed to maximize the aggregate performance in an FMN consisting of a single FGW. However, when a single FGW using a shared wireless channel does not allow for high enough capacity to accommodate the traffic demand and meet the delay constraints, the QoS offered may be degraded. The joint management of the communications resources, including their number, placement, and allocation, was addressed in the second phase of this thesis.

In order to ensure heterogeneous QoS guarantees on top of a shared physical FMN, in the second phase of this thesis we designed a network slicing-aware algorithm to support different communications services using a minimum amount of communications resources. It enables the joint placement and allocation of communications resources by determining the number and 3D positions of multiple FAPs and FGWs as well as the number of wireless channels to be made available, in order to satisfy the QoS requirements of any number and type of network slices. The proposed algorithm allows that AM2Fly enables meeting heterogeneous QoS levels associated with multiple network slices using the minimum number of UAVs and radio resources.

1.7.2 Performance Evaluation

Since a theoretical performance evaluation requires many simplifications, due to the high complexity of the problem addressed in this thesis, it potentially leads to unrealistic results. For this reason, the performance evaluation of AM2Fly was carried out by means of Network Simulator 3 (ns-3) [73] simulations and an experimental testbed.

On the one hand, the ns-3 simulations allow to evaluate the network performance by considering complex networking scenarios characterized by high variability in the communications nodes' positions and network configurations. The ns-3 simulator was chosen due to its accurate models for wireless networks and its wide acceptance by the scientific community. It allows taking into account the stochastic characteristics of wireless networks and the behavior of representative traffic generation models by means of a packet-driven simulation approach, thus leading to more realistic results. On the other hand, the experimental testbed allowed validating in the real-world the results obtained by means of ns-3 simulations.

Representative networking scenarios observed in the real-world were employed in the performance evaluation carried out. These scenarios were defined in terms of the number of UAVs used, their spatial distribution in 3D, and the requirements associated to the offered traffic. For that purpose, random dynamic networking scenarios consisting of different numbers of UAVs were defined, in order to consider small, medium, and large FMNs.

1.8 Original Contributions

The original contributions claimed by this thesis are:

1. **An Algorithm for Gateway Placement in FMNs (GWP).** The GateWay Placement (GWP) algorithm is proposed to define a suitable placement for the FGW in FMNs with controlled topology. GWP takes advantage of the information on the traffic demand of the ground devices and the positions of the FAPs, defined by a state of the art FAP placement algorithm, in order to define the position of the FGW. GWP ensures that the wireless link established between each FAP and the FGW has high enough capacity to accommodate the traffic exchanged with the ground devices over time.

The literature has been focused on the placement of UAVs as APs and BSs serving the ground devices, overlooking the backhaul network design. Even though the ground devices are directly affected by the QoS offered by the RAN, the backhaul network needs to be carefully designed in order to meet the variable traffic demand of the FAPs. GWP introduces a differentiating factor that lies on defining in advance the position of the FGW taking advantage of the knowledge of the future positions of the FAPs being served, whose mobility is controlled in a Central Station. The performance achieved using the GWP algorithm was evaluated through ns-3 simulations. The obtained results show gains up to 40% in the aggregate throughput, while the delay is reduced up to 26%. This contribution has been published in [21, 74, 75].

2. **A Mechanism for Proactive Queue Management in FMNs (PQM).** The Proactive Queue Management (PQM) mechanism is proposed to define in advance the queue size of the

UAVs over time, in order to maximize the throughput with stochastic delay guarantees. PQM benefits from the centralized and holistic view over the FMN available in a Central Station, including the positions of the FAPs and the QoS levels of the ground devices; the QoS levels are assumed to be defined in terms of the average traffic demand and the maximum average packet delay. Considering the future positions of the FAPs and the FGW, PQM ensures that the queue of each FAP has, at each time instant, the minimum size that allows accommodating the traffic sent towards the FGW.

It is worth noting that the GWP algorithm aims at ensuring target network capacity values for the wireless links established between the FAPs and the FGWs. Still, in a real-world communications node, the instantaneous traffic demand is difficult to estimate, including the maximum values and duration of traffic bursts. If the instantaneous traffic demand at a given FAP is greater than the network capacity available in the wireless link established with the FGW, the FAP holds packets in its queue. The PQM mechanism allows ensuring controlled delays by defining the minimum queue size for each FAP. PQM was designed upon the conclusion that the usage of static, oversized queues in wireless networks leads to underutilized communications nodes and unnecessary high delays without improving the aggregate throughput. By ensuring a minimum queue size, PQM allows the packets to be held in the queues for shorter time. Also, the queuing delay is decreased, the throughput is not compromised, and the shared medium usage is enhanced. This allows for improved overall network performance. When a target reliability level should be met in the QoS guarantees to be offered by the FMN, a trade-off for defining the queue size should be considered. A larger queue size allows for improved reliability at the expense of greater delay; a smaller queue size allows for shorter delay but lower reliability. This thesis is focused on ensuring throughput and delay guarantees only.

The design of a single queue management mechanism able to meet the variable characteristics of wireless networks has been pointed out in the literature as a research challenge, especially due to the need of fine-tuning several parameters for properly configuring the communications nodes. The queue size management challenge is exacerbated because state of the art AQM solutions have been designed for wired networks, which are static in nature. Wireless networks in general and FMNs in particular, due to their highly dynamic nature, bring up additional challenges that must be taken into account, including the variable channel capacity over time. PQM advances the state of the art by dynamically managing the queue size in advance according to the capacity of the wireless links and the traffic demand of the ground devices, leveraged on the knowledge about the future positions of the UAVs that compose the FMN. The performance achieved using the PQM mechanism was evaluated by means of ns-3 simulations. The obtained results show gains up to 28% in the

aggregate throughput, while the delay is reduced up to 96%. This contribution has been published in [76].

3. **An Algorithm for Placement and Queue Management in FMNs (GPQM).** The Gateway Placement and Queue Management (GPQM) algorithm results from the combination of the GWP algorithm and the PQM mechanism. On the one hand, the PQM algorithm aims at maximizing the throughput of the traffic carried between the FAPs and the FGW, while meeting the average delay constraint. Since PQM configures only the queue size of the communications nodes, it is unable to find a suitable solution that meets the average delay constraint for all flows offered by the FAPs if the capacity of the wireless links is not high enough. On the other hand, the GWP algorithm aims at ensuring that the wireless link established between each FAP and the FGW has high enough capacity for carrying the offered traffic. For that purpose, the GWP algorithm defines the FGW placement and the transmission power of the UAVs. However, GWP does not address the queue size management of the communications nodes, which may imply long queuing delays when the queue size is not adequate for the capacity of the wireless links available. As such, GWP is not able to provide average delay guarantees.

GPQM takes advantage of the information on the future positions of the FAPs, defined according to the traffic demand of the ground devices by a state of the art FAP placement algorithm that runs in a Central Station, in order to 1) determine the FGW position so that a wireless link able to accommodate the traffic forwarded by each FAP is established between each FAP and the FGW, and 2) compute the queue size for each FAP so that stochastic delay guarantees are provided. GPQM allows jointly defining an adequate placement for the FGW and a suitable queue size for the FAPs, which are aspects addressed independently by the state of the art solutions, without taking advantage of the gains allowed by each other. GPQM allows minimizing the packet service time and the traffic load in each FAP. The performance achieved using the GPQM algorithm was evaluated using ns-3 simulations, considering a realistic wireless channel model. The obtained results show gains up to 85% in the aggregate throughput, while the delay is reduced up to 74%. This contribution has been published in [77].

4. **A Mechanism for Predictive Routing in FMNs (RedeFINE).** The Routing for Flying Multi-hop Networks (RedeFINE) mechanism is proposed for selecting in advance the best wireless path for each FAP and the time instants each forwarding table is updated, acting in advance to the UAV placement. It allows for uninterruptible communications paths and includes a novel Inter-flow Interference-aware Routing (I2R) metric, tailored to centralized routing in FMNs. RedeFINE takes advantage of the centralized knowledge of the future positions and trajectories of the UAVs to predict the capacity of the wireless links. This holistic

knowledge is available in a Central Station. The Central Station is running an independent state of the art UAV placement algorithm. The state of the art UAV placement algorithm is used to define the optimal positions of the FAPs, in order to meet the traffic demand of the ground devices. RedeFINE can use the proposed I2R metric or any other routing metric suitable for centralized routing.

The RedeFINE mechanism represents a step forward with respect to existing routing solutions, in which the communications nodes only recover from link failures after they occur, which implies packet loss, throughput degradation, and non-zero reconfiguration times, due to the time wasted for updating the forwarding tables. Due to the highly dynamic behavior of FMNs, which induce frequent changes in the quality of the radio links and in the network topology, existing routing solutions need to constantly flood the network with route discovery packets, thus reducing the bandwidth available for data traffic. In order to address this challenge, RedeFINE and I2R do not use probe packets for neighbor discovery, interference estimation, and link quality prediction, benefiting from the information about the future positions of the FAPs defined in the Central Station. This allows reducing the bandwidth required for control traffic. The performance achieved using RedeFINE was evaluated using ns-3 simulations. The obtained results show gains up to 54% in the aggregated throughput without gains in the delay. This contribution has been published in [21, 78, 79].

5. **An Algorithm for Placement and Allocation of Communications Resources in Slicing-aware FMNs (SurFABLE).** The Slicing-aware Flying Access and Backhaul (SurFABLE) algorithm is proposed to enable slicing-aware FMNs. For that purpose, it makes the joint computation of the amount of communications resources needed, namely the number of FAPs and FGWs, and their placement in 3D space. SurFABLE allows ensuring that 1) the heterogeneous QoS levels of different network slices are met and 2) the capacity of each UAV in terms of the total number of wireless channels available is not exceeded. The resulting FMN allows extending wireless coverage using a multi-hop topology and offers multiple network slices that ensure the targeted QoS levels to ground devices using the minimum number of UAVs and radio resources. SurFABLE uses 1) PQM to compute a suitable queue size for each FAP and 2) RedeFINE to define the forwarding tables of the FAPs.

State of the art UAV placement algorithms have been especially designed for determining the UAV positions that maximize the wireless coverage or the QoS offered to the ground devices using a best-effort approach. When it comes to slicing in FMNs, the literature has been focused on UAVs acting as clients of telecom infrastructures. Existing works that take advantage of UAVs for providing on-demand network slices aim at improving energy-efficiency, fair coverage, and resource allocation efficiency. However, they do not consider the joint placement and allocation of communications resources. SurFABLE aims at paving

the way to integrate UAVs into Mobile Network Operators' infrastructures as on-demand communications resources that can be deployed anywhere, anytime, while meeting different QoS levels of any number and type of network slices that coexist on top of a shared physical flying network infrastructure. The performance achieved using SurFABLE was evaluated by means of ns-3 simulations and an experimental testbed. The obtained results show that SurFABLE enables meeting the heterogeneous QoS levels of multiple network slices using the minimum number of UAVs and radio resources. This contribution has been published in [38, 80].

As result of this thesis, eleven conference papers and five journal papers were produced:

1. **André Coelho**, Rui Campos, and Manuel Ricardo. Traffic-aware Gateway Placement and Queue Management in Flying Networks. *Ad Hoc Networks*, 138:103000, 2023.
2. **André Coelho**, João Rodrigues, Helder Fontes, Rui Campos, and Manuel Ricardo. An Algorithm for Placing and Allocating Communications Resources Based on Slicing-Aware Flying Access and Backhaul Networks. *IEEE Access*, 10:128923–128942, 2022.
3. Hugo Rodrigues, **André Coelho**, Manuel Ricardo, and Rui Campos. Joint Energy and Performance Aware Relay Positioning in Flying Networks. *IEEE Access*, 10:1–16, 2022.
4. Hugo Rodrigues, **André Coelho**, Manuel Ricardo, and Rui Campos. Energy-aware Relay Positioning in Flying Networks. *International Journal of Communication Systems*, pages 1–17, 2022.
5. **André Coelho**, Helder Fontes, Rui Campos, and Manuel Ricardo. Placement and Allocation of Communications Resources in Slicing-aware Flying Networks. In *Wireless On-Demand Network Systems and Services Conference (WONS)*, pages 1–8. IFIP, 2022.
6. David Maia, **André Coelho**, and Manuel Ricardo. Obstacle-aware On-demand 5G Network using a Mobile Robotic Platform. In *2022 18th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 470–473. IEEE, 2022.
7. Max Friedrich, Teemu Joonas Lieb, Annette Temme, Eduardo N. Almeida, **André Coelho**, and Helder Fontes. ResponDrone – A Situation Awareness Platform for First Responders. In *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)*, pages 1–7. IEEE, 2022.
8. Eduardo Nuno Almeida, **André Coelho**, José Ruela, Rui Campos, and Manuel Ricardo. Joint traffic-aware UAV placement and predictive routing for aerial networks. *Ad Hoc Networks*, 118:102525, 2021.

9. **André Coelho**, Helder Fontes, Rui Campos, and Manuel Ricardo. Traffic-aware Gateway Placement for High-capacity Flying Networks. In *IEEE Vehicular Technology Conference*. IEEE, 2021.
10. Gonçalo Santos, João Martins, **André Coelho**, Helder Fontes, Manuel Ricardo, and Rui Campos. A Fast Gateway Placement Algorithm for Flying Networks. In *IEEE Vehicular Technology Conference*. IEEE, 2021.
11. **André Coelho**, Filipe Ribeiro, Helder Fontes, Manuel Ricardo, and Rui Campos. On the Experimental Evaluation of a Flying Network for Coverage and Capacity Extension in 5G+ Scenarios. In *INForum 2021 (Peer-reviewed Portuguese Symposium on Informatics Engineering)*, 2021.
12. **André Coelho**, Rui Campos, and Manuel Ricardo. Proactive Queue Management for Flying Networks. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*. IEEE, 2020.
13. **André Coelho**, Eduardo Nuno Almeida, José Ruela, Rui Campos, and Manuel Ricardo. A Routing Metric for Inter-flow Interference-aware Flying Multi-hop Networks. In *2019 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6. IEEE, 2019.
14. Renato Cruz, **André Coelho**, Rui Campos, and Manuel Ricardo. A Theoretical Model for Planning NB-IoT Networks. In *International Conference on Wireless and Mobile Computing, Networking and Communications*, pages 31–34. IEEE, 2019.
15. **André Coelho**, Mario Lopes, Bruno Ferreira, Rui Campos, and Manuel Ricardo. Experimental Evaluation of Shore to Unmanned Surface Vehicle Wi-Fi Communications. In *IFIP Wireless Days*, pages 86–91. IFIP, 2018.
16. **André Coelho**, Eduardo N. Almeida, Pedro Silva, José Ruela, Rui Campos, and Manuel Ricardo. RedeFINE: Centralized Routing for High-capacity Multi-hop Flying Networks. In *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 75–82, 10 2018.

In addition, multiple M.Sc. dissertations and summer internships were carried out in the context of this thesis and under the supervision of the author, including:

1. **Energy and Performance Aware Multi-Drone Placement for Sustainable Flying Networks**. Pedro Ribeiro. M.Sc. dissertation. Faculdade de Engenharia, Universidade do Porto, 2023 (ongoing).

2. **5G Mobile Open Radio Access Network.** Gonalo Queirs. M.Sc. dissertation. Faculdade de Engenharia, Universidade do Porto, 2023 (ongoing).
3. **Slicing-aware Flying Communications Network.** Joo Rodrigues. M.Sc. dissertation. Faculdade de Engenharia, Universidade do Porto, 2022.
4. **Control and Positioning of a 5G Radio Access Node Deployed in a Mobile Robotic Platform.** David Maia. M.Sc. dissertation. Faculdade de Engenharia, Universidade do Porto, 2022.
5. **Resource Planning Optimization in Flying Communications Networks.** David Pinto, Ivan Cojocar, Manuel Silva, and Rafael Morais. Summer internship. Institute for Systems and Computer Engineering, Technology and Science (INESC TEC) and Faculdade de Engenharia, Universidade do Porto, 2021.
6. **Gateway Positioning in Flying Networks.** Hugo Rodrigues. M.Sc. dissertation. Faculdade de Engenharia, Universidade do Porto, 2020.
7. **Traffic-aware Flying Communications Node Controller.** Eduardo Correia, Gonalo Santos, Joo Carlos Martins, and Joo Miguel Martins. Summer internship. INESC TEC and Faculdade de Engenharia, Universidade do Porto, 2020.
8. **Routing for Flying Networks using Software-Defined Networking.** Andr Oliveira. M.Sc. dissertation. Faculdade de Engenharia, Universidade do Porto, 2019.

The multiple M.Sc. dissertations and summer internships developed within the scope of this thesis promoted the advanced training of undergraduate students.

1.9 Related Research Projects

Multiple research projects developed at INESC TEC [88] have been directly related to this thesis. These research projects, which are briefly described in the following subsections, allowed us to: 1) identify real-world problems and motivate the contributions claimed by this thesis; 2) validate and evaluate experimentally the solutions developed; and 3) disseminate the results of this thesis with the scientific community.

1.9.1 WISE

The WISE [89] project, funded by *Fundao para a Cincia e a Tecnologia*, developed a new communications solution based on the concept of Flying Backhaul Mesh Network. At the core of the Flying Backhaul Mesh Network concept is the usage of rotor-wing UAVs that carry

communications nodes on-board able to form a mobile and physically reconfigurable wireless backhaul mesh network consisting of FAPs. The proposed network is self-organized and the FAPs are able to position themselves according to the data traffic generated by ground devices and the need to relay the traffic towards the Internet. In order to validate the WISE concept in the real-world, a testbed was designed. It includes a UAV that acts as a FAP to the ground devices, as depicted in Fig. 1.5.

The predictive routing mechanism for FMNs claimed as a contribution of this thesis was evaluated in the WISE project.



Figure 1.5: Experimental validation of the WISE's communications solution.

1.9.2 ResponDrone

ResponDrone was an international project, co-funded by the European Union and the Korean government, that developed an integrated solution for first responders so that they are able to easily operate several drones with multiple synchronized missions for enhanced situational awareness and human protection.

The ResponDrone solution offers a Communications Relaying Service that uses UAVs to transport the INESC TEC's Payload Communications Relay, as depicted in Fig. 1.6. The Payload Communications Relays form a private flying network that provides on-demand wireless communications to first responders and victims management scenarios. This network allows first responders to access the Internet, communicate among themselves, use the ResponDrone mobile app running on their smartphones, and relay data to/from the command centers and the Cloud.

The ResponDrone project made it possible to develop a high Technology Readiness Level FAP carried by an industrial-grade UAV and integrate it with other solutions developed by different project partners as a complete system. This allowed evaluating the feasibility of applying the

contributions of this thesis in the real-world. In addition, ResponDrone enabled collecting an experimental communications dataset using UAVs, which allowed validating in practice the wireless channel model considered in the design of the contributions of this thesis.

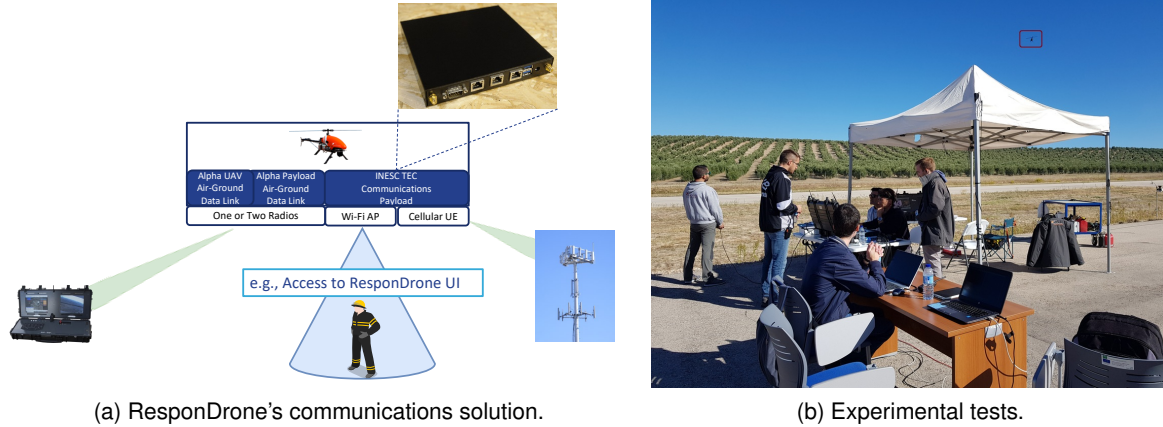


Figure 1.6: Experimental validation of the ResponDrone's communications solution.

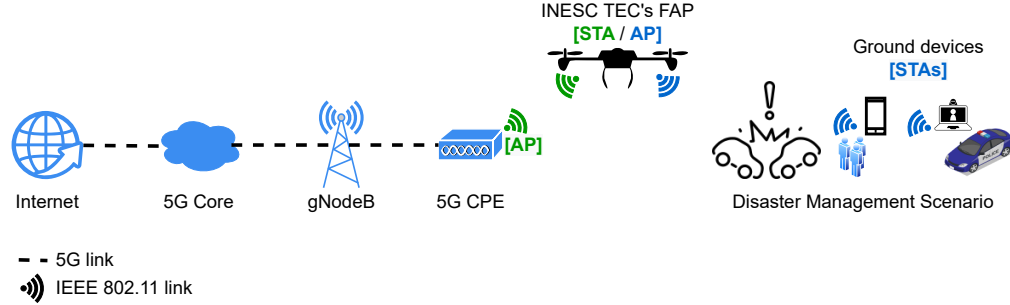
1.9.3 5Go

5Go was a project funded by Portugal 2020 that designed, integrated, and validated a new set of communications services and products for 5G networks, including RAN, Core Network, and verticals.

The component developed by INESC TEC in 5Go was the FAP, which is a direct result of this thesis. The FAP takes advantage of a rotor-wing UAV that carries a communications node on board, in order to extend the coverage and network capacity in 5G and beyond networking scenarios, including disaster management scenarios, as depicted in Fig. 1.7. For that purpose, the FAP is able to autonomously position in a 3D space according to the decisions taken by a FAP placement algorithm so that the traffic demand of the ground devices is met. The FAP includes a computing unit equipped with a Network Interface Card (NIC) acting as an AP or BS, which enables a RAN that provides Internet connectivity to the ground devices, and an additional NIC connected to a backhaul network, which takes advantage of a 5G Customer Premises Equipment (CPE) to forward the traffic to/from the Internet by means of 5G. Multiple FAPs can coexist to form an FMN, and they can be integrated with existing terrestrial networks, including APs and BSs.

1.9.4 CHIC

CHIC, Cooperative Holistic View on Internet and Content, was a Portugal 2020 project that developed, tested, and demonstrated new processes, products, and services for the audiovisual



(a) Reference disaster management scenario considered in the performance evaluation of the INESC TEC's FAP.



(b) UAV acting as a FAP.

Figure 1.7: INESC TEC's FAP and reference networking scenario used in the performance evaluation carried out in the 5Go project.

and multimedia sectors.

INESC TEC was in charge of the design, development, and validation of the CHIC's Wi-Fi network infrastructure, which took advantage of UAVs acting as FAPs and their cooperation with existing terrestrial networks. To implement and demonstrate the CHIC's communications solution, a prototype was built. It consisted of a rotor-wing UAV carrying a communications node acting as FAP, which was able to extend the Internet connectivity from a remote location (e.g., cellular BS) and provide a RAN to the devices in crowded events (e.g., music festival), as shown in Fig. 1.8.

The mechanism for proactive queue management in FMNs, which is one of the contributions claimed by this thesis, was validated in the CHIC project.

1.9.5 Summary

The research projects presented herein allowed to promote the integration of the contributions achieved in this thesis with state of the art solutions, while creating value and impact through the integration, validation, and evaluation of the developed solutions in the real-world. Moreover, the validation of the contributions of this thesis within the context of research projects paved the way



Figure 1.8: Experimental validation and demonstration of the CHIC's communications solution.

for the production of publications in national and international scientific conferences and journals, promoting not only the advance of the state of the art but also the dissemination of the research projects within the international scientific community.

1.10 Document Structure

This thesis consists of seven chapters considering the structure presented herein.

Chapter 2 presents the algorithm for placement of gateways in FMNs (GWP). The rationale of GWP and a theoretical analysis for a representative scenario are presented. ns-3 simulation results demonstrate the gains achieved when placing the FGW using GWP against different placement approaches, including 1) placing the FGW for maximizing the SNR achieved per FAP, 2) placing the FGW in the geometric center of the venue, and 3) placing the FGW in random positions over time.

Chapter 3 introduces the mechanism for proactive queue management in FMNs (PQM). The rationale of PQM and a theoretical analysis for a representative scenario are described. ns-3 simulation results demonstrate the gains achieved by PQM when compared with state of the art counterpart solutions, including Random Early Detection (RED), BLUE, Proportional Integral (PI), Proportional Integral controller Enhanced (PIE), and Controlled Delay (CoDel).

Chapter 4 explains the algorithm for placement and queue management in FMNs (GPQM). The rationale of GPQM and a theoretical analysis for a representative scenario are presented. ns-3 simulation results using a realistic wireless channel model taken in a real-world testbed show the gains of GPQM against multiple state of the art counterpart approaches, including 1) placing the FGW for maximizing the SNR achieved per FAP and 2) placing the FGW in the geometric

center of the venue, while using either RED or CoDel as the AQM solutions. Moreover, the quality of the solutions taken by GPQM are assessed against the solutions obtained by means of state of the art solvers for the optimization problem formulated.

Chapter 5 describes the mechanism for predictive routing in FMNs (RedeFINE). The rationale of RedeFINE and a theoretical analysis for a representative scenario are presented. ns-3 simulation results show the gains of RedeFINE against different state of the art counterparts, including the static forwarding approach, and the Ad hoc On-Demand Distance Vector (AODV) and Optimized Link State Routing (OLSR) routing protocols, as well as the Expected Transmission Count (ETX), Airtime, and Euclidean distance routing metrics.

Chapter 6 refers to the algorithm for placement and allocation of communications resources in slicing-aware FMNs (SurFABLE). The rationale of SurFABLE is explained. In addition, a theoretical analysis of the proposed channel assignment approach is made for a representative scenario. ns-3 simulations and experimental results show that SurFABLE allows meeting the targeted QoS levels associated with multiple network slices using the minimum number of UAVs and radio resources.

Finally, Chapter 7 draws the main conclusions, reviews the work done and the main original contributions of this thesis, and points out directions for future work related with this thesis.

Chapter 2

An Algorithm for Gateway Placement in Flying Multi-hop Networks

Following the advent of Flying Multi-hop Networks (FMNs) consisting of Unmanned Aerial Vehicles (UAVs), several algorithms have been proposed in the literature to determine the UAV placement and improve the network performance in terms of coverage, Signal-to-Noise Ratio (SNR) provided, number of users served, and aggregate Quality of Service (QoS) offered. Still, state of the art works have been focused on UAVs acting as Access Points (APs) and Base Stations (BSs) in Radio Access Networks (RANs). Even though the ground devices are directly affected by the QoS provided by the RAN, the backhaul network needs to be carefully designed in order to accommodate the traffic forwarded by the Flying Access Points (FAPs) to/from the Internet, while ensuring the targeted QoS levels. This aspect has been overlooked in the state of the art.

Considering the focus of the literature on UAV placement algorithms for RANs, in this thesis we went a step further by proposing an algorithm for the placement of gateways in FMNs, named GateWay Placement (GWP). GWP is a centralized, traffic-aware algorithm targeting FMNs with controlled topology, which enables backhaul communications paths with high enough capacity to accommodate the traffic demand of the ground devices. GWP assumes the traffic demand and the future positions of the FAPs are provided by a state of the art FAP placement algorithm, such as the one proposed in the WISE project [4], running in parallel to define the FAP positions so that the traffic demand of the ground devices is met.

We start this chapter by presenting the related work on gateway placement solutions for wireless mesh networks, wireless sensor networks, and flying networks. Then, we define the system model and formulate the problem. After that, we describe the GWP algorithm, including its rationale and a theoretical analysis for a representative scenario. Afterwards, we present the

performance evaluation of the GWP algorithm, including the simulation setup, the simulation scenarios, the performance metrics considered, and the simulation results obtained. At the end of the chapter, we summarize the contents presented.

2.1 Related Work

Considering the Flying Gateway (FGW) placement problem in FMNs, in this section we review the related work on placement in wireless networks, wireless sensor networks, and flying networks.

2.1.1 Wireless Networks

In the literature, gateway placement in wireless networks has been a common problem. The gateway placement problem typically consists of identifying a subset of communications nodes that are candidates for acting as gateways among a set of communications nodes connected in a wireless mesh network, as depicted in Fig. 2.1.

In [90], an algorithm has been proposed for selecting the set of gateways in a wireless mesh network while taking into account delay, traffic load per node, and gateway capacity constraints. The algorithm employs an incremental clustering approach that allows identifying gateways and their clients in an iterative manner. A similar work has been presented in [2]. The problem has been formulated as a Capacitated Facility Location Problem, with two optimization objectives subject to QoS requirements: minimizing the number of gateways and minimizing the number of hops to reach the gateways. An algorithm, named Zero-Degree, has been proposed to solve this problem, while satisfying all constraints and considering the distribution of gateways (i.e., aggregate or sparse) as well as their location with respect to the available wired network and Internet connection points. In [91], the gateway placement and spatial reuse in wireless mesh networks has been formulated as an optimization problem. The authors had jointly considered interference, routing, and scheduling in the gateway placement problem, which are aspects typically addressed independently.

Existing works on placement and selection of gateways in wireless mesh networks have been reviewed in [92], including solutions that use multiple Network Interface Cards (NICs) and multiple wireless channels to increase the network throughput and capacity. Multiple works addressing Single Radio Single Channel, Single Radio Multi-Channel, and Multi-Radio Multi-Channel have also been presented. Moreover, a discussion on the benefits and drawbacks of heuristic, meta-heuristic, and operational research-based approaches has been carried out. While heuristic and metaheuristic-based approaches allow for quasi-optimal solutions, operational research-based methods lead to optimal solutions. Centralized approaches typically allow for better performance

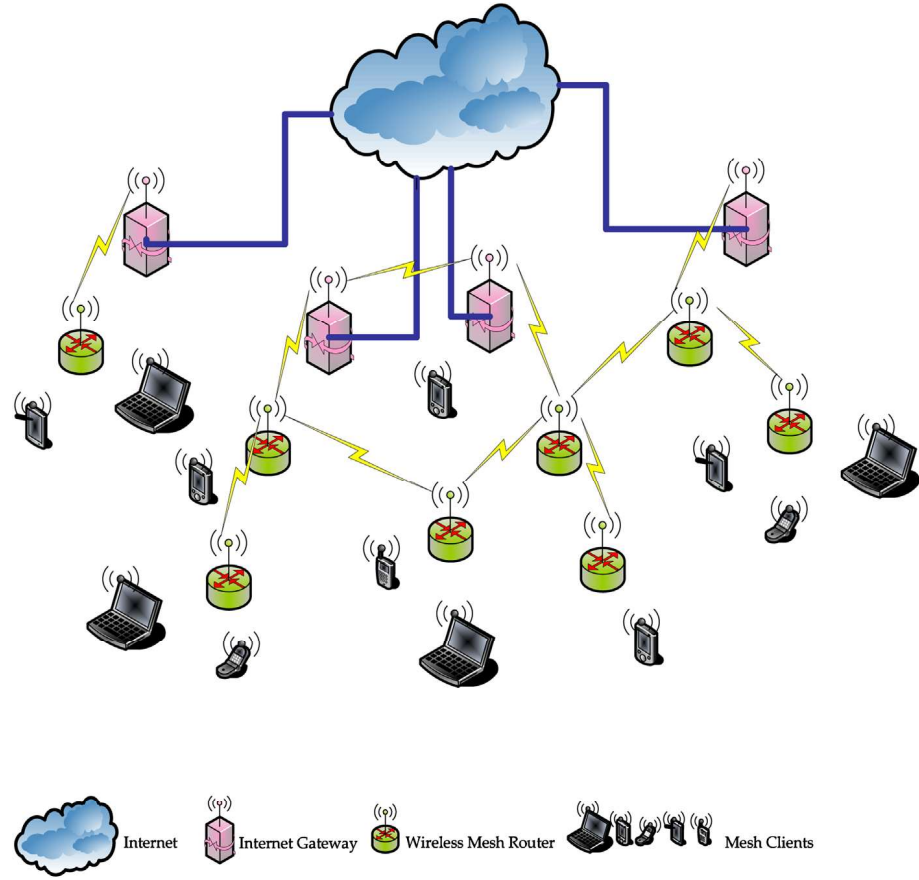


Figure 2.1: Wireless mesh network consisting of clients, routers/communications relays, and gateways connected to the Internet [2].

in small-sized networks, whereas distributed methods lead to optimal solutions in reasonable times for large-sized networks.

Despite considering the traffic demand and the capacity of the wireless links, state of the art works assume a fixed position for the gateways. In [93], the authors have concluded that the gateway placement and the transmission power have a significant impact on the network throughput. They have evaluated three heuristics: 1) *H1*, which defines the minimum transmission power for the devices and uses the minimum hop count metric for placing the gateways; 2) *H2*, which places the gateways and defines a minimum transmission power to meet a targeted Signal-to-Interference plus Noise Ratio; and 3) *H3*, which uses the shortest path algorithm to compute the minimum weight path from any device to the gateway with the ultimate goal of minimizing the sum of all the shortest paths' weights. Due to the short radio range of the wireless links established when using *H1*, multiple gateways are typically selected. *H2* is also suboptimal, since it selects the same gateways' positions for multiple transmission power levels. *H3*, which allows

selecting the smallest number of gateways, is the best of the three heuristics evaluated. However, the three heuristics do not consider the devices' traffic demand.

Even though different studies have been carried out over the years to address the gateway placement problem in wireless networks, most of them aim at minimizing the number of gateways while defining their placement considering a set of predefined locations; they are not optimal solutions for improving target network performance metrics, including throughput and delay.

2.1.2 Wireless Sensor Networks

An analogy between the gateway placement and the sink placement in the context of wireless sensor networks, as depicted in Fig. 2.2, can be made. Both are in charge of receiving all the traffic generated within a network. In [3], several sink placement approaches have been introduced, and their benefits and drawbacks have been discussed. Random Sink Placement has been used as a worst-case approach. Geographic Sink Placement has been considered as a baseline approach. Intelligent Sink Placement has been proposed as the optimal sink placement approach. Because Intelligent Sink Placement is computationally expensive, a Genetic Algorithm-based Sink Placement approach has been discussed and shown to be a suitable approximation for optimal sink placement. The Geographic Sink Placement approach is adequate for wireless sensor networks with unknown sensor nodes' positions, whereas the Genetic Algorithm-based Sink Placement approach is best suited for wireless sensor networks with known sensor nodes' positions. However, in wireless sensor networks the traffic demand is significantly lower and it is not generated continuously over time. Moreover, the sink placement in wireless sensor networks aims at minimizing the energy consumption or the delay, and increase the network lifetime rather than enabling high-capacity wireless communications paths; this is a key-requirement in FMNs serving emerging communications services and applications.

2.1.3 Flying Networks

The literature on UAV placement in FMNs has been focused on defining the positions of 1) FAPs, considering wireless coverage, number of ground devices served, and QoS offered to ground devices [4, 21, 52, 64–67, 94–98], and 2) UAVs forwarding traffic between ground devices [5, 7, 99]. When it comes to the FAP positioning in FMNs (cf. Fig. 2.3), a reference algorithm, named NetPlan, has been proposed in [4].

NetPlan is based on the Potential Fields technique [100], which lies on Potential Fields generators that apply forces on the FAPs; forces are attractive to areas with high traffic demand and repulsive to areas with low traffic demand. In order to improve the aggregate throughput, the FAPs closer to ground devices with higher traffic demand establish smaller Wi-Fi cells, whereas the remaining FAPs establish larger Wi-Fi cells so that all devices are served, as shown in Fig. 2.4

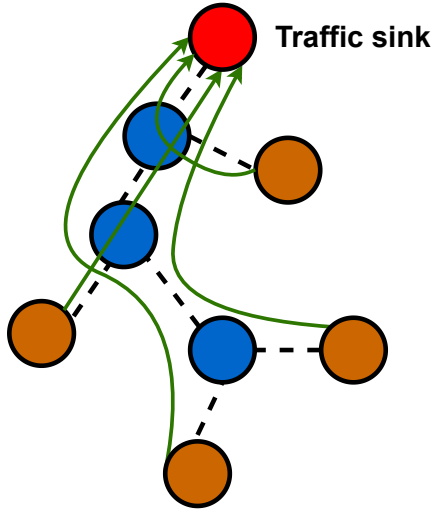


Figure 2.2: Representative wireless sensor network topology, where the communications paths consist of wireless links (dashed black lines) in a multi-hop architecture. The positions of the communications nodes (circles) and the offered traffic flows (solid green lines) are typically considered in the state of the art sink placement approaches, in order to meet the targeted network performance (modified from [3]).

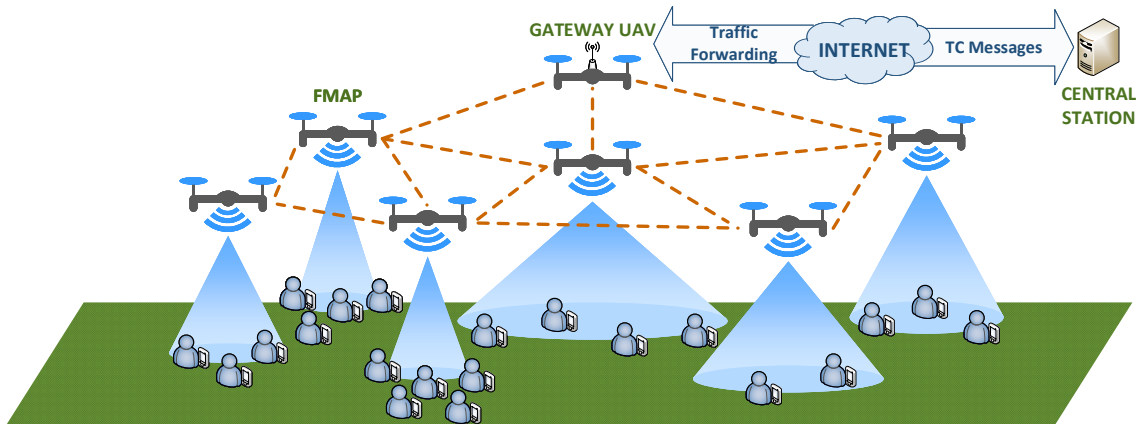
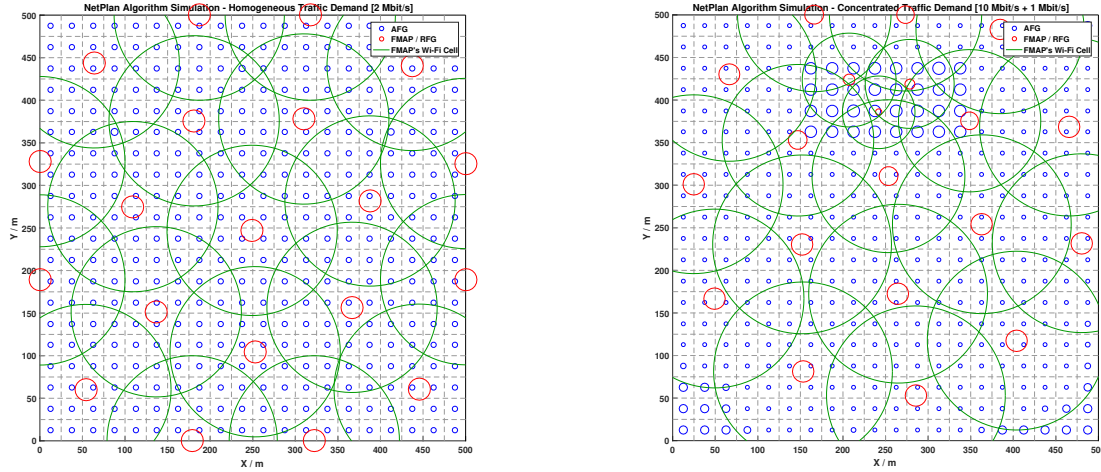


Figure 2.3: Flying network consisting of a gateway UAV and Flying Mesh Access Points (FMAPs); the FMAPs provide wireless communications cells to the ground devices. The placement of the FMAPs is defined by the state of the art NetPlan algorithm [4].

for homogeneous and concentrated traffic demand. However, the authors of NetPlan have been focused on placing the FAPs in order to meet the traffic demand of ground devices. The design of the wireless backhaul network has not been addressed.

Concerning the second type of approach, in [5], the authors have studied the optimal position



(a) Homogeneous traffic demand, in which the FMAPs provide equal sized wireless communications cells.

(b) Heterogeneous traffic demand, in which the FMAPs provide smaller sized and denser wireless communications cells in areas of higher traffic demand.

Figure 2.4: RAN designed by the NetPlan algorithm, depicting the FMAPs (red circles) and the wireless communications cells (green circles) offered to the ground devices (blue circles). The traffic demand of the ground devices is proportional to the size of their circles [4].

for a UAV acting as a communications relay between two ground devices in a multi-rate Wi-Fi network, as depicted in Fig. 2.5. Theoretical analysis and Network Simulator 3 (ns-3) [73] simulations allowed to conclude that placing the UAV in the geometric center of the ground devices does not allow maximizing the throughput achieved. Considering the reference networking scenario depicted in Fig. 2.5, the authors have evaluated the network performance for three possible alternative UAV placement approaches:

1. the data rate of the wireless link established between the UAV and the ground device at a greater Euclidean distance is reduced, but the data rate of the wireless link established between the UAV and the ground device at a shorter Euclidean distance is not improved. In such case, the placement of the UAV in the geometric center of the two ground devices maximizes the network performance.
2. the data rate of the wireless link established between the UAV and the ground device at a shorter Euclidean distance is improved, while the data rate of the wireless link established between the UAV and the ground device at a greater Euclidean distance is not reduced.
3. the data rate of the wireless link established between the UAV and the ground device at a shorter Euclidean distance is improved, but the data rate of the wireless link established between the UAV and the ground device at a greater Euclidean distance is reduced.

However, the obtained conclusions are limited to a networking scenario consisting of two ground devices and a UAV only.

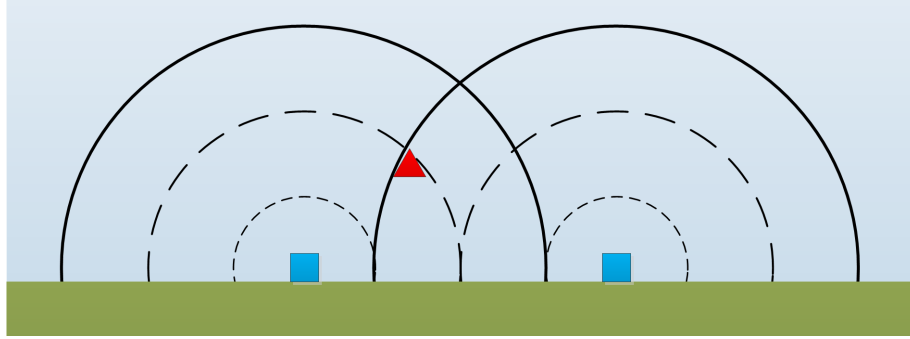


Figure 2.5: Placement of a flying communications relay UAV (red triangle) forwarding data traffic between two ground devices (blue squares) in an IEEE multi-rate network. The semicircles represent the maximum wireless ranges between each ground device and the flying communications relay UAV to enable the use of different IEEE 802.11 physical data rates [5].

In [6], the authors have proposed an on-demand density-aware UAV 3-Dimensional (3D) deployment algorithm to increase the number of ground devices served by a single UAV, as depicted in Fig. 2.6. The problem has been formulated as an NP-complete binary Knapsack problem¹, while satisfying the target data rate associated with each ground device and taking into account capacity and positioning constraints for a UAV. The algorithm aims at determining the horizontal position of the UAV, in order to guarantee the targeted data rate for each ground device. Through numerical simulation results, the authors have concluded that their algorithm allows ensuring the targeted data rate for four different types of traffic associated with High-Definition video streaming, online gaming, web browsing, and Voice over Internet Protocol services. However, the algorithm assumes that the UAV is deployed at a constant altitude and is only focused on the access network provided to the ground devices; the wireless backhaul network design has not been addressed.

In [7], the optimal UAV relay deployment problem has been addressed. Considering an unknown channel model and ground devices' positions, three model-free approaches have been proposed to solve the problem: 1) Multi Unmanned Aerial Vehicle-based Particle Swarming Optimization Approach (MPA); 2) Single Unmanned Aerial Vehicle Particle Swarming Optimization deployment Approach (SPA); and 3) Single Unmanned Aerial Vehicle Exploration-Exploitation deployment Approach (SEEA). According to the simulation results, MPA allows the fastest convergence rate and a high probability of achieving the optimal solution for multiple UAVs, as depicted in Fig. 2.7. The convergence rate is defined as the number of *Monte Carlo* runs required to

¹Knapsack problem relies on combinatorial optimization, in which for a given a set of items, each with a weight and a value, the objective is determining the number of items to select so that the total weight is less than or equal to a target limit and the total value is as large as possible.

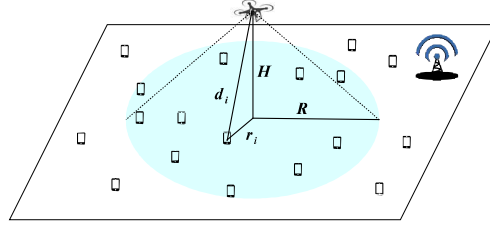
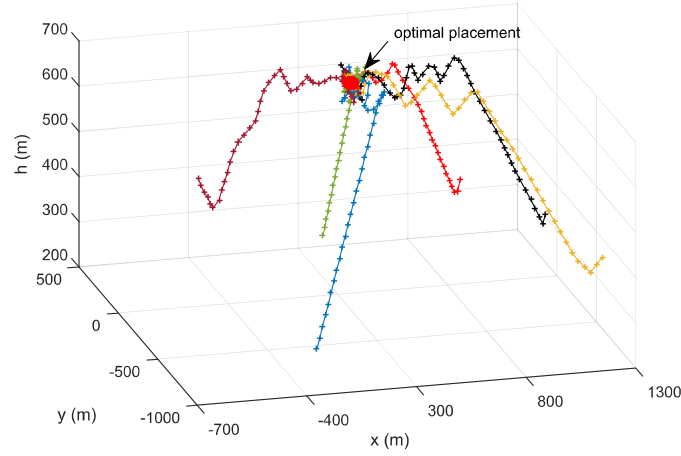


Figure 2.6: Wireless networking scenario addressed by the algorithm proposed in [6], where a UAV is placed to serve a set of ground services with targeted data rate requirements.

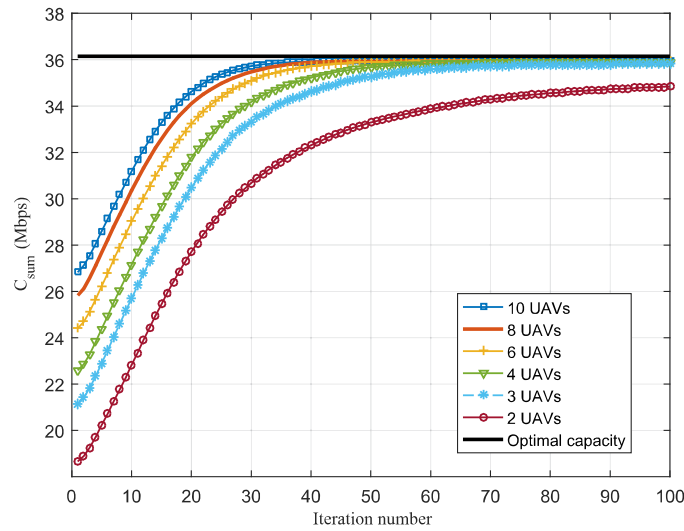
reach the optimal UAV solution. The smaller the number of UAVs, the lower the convergence rate achieved by MAP. This is because MPA lies in multiple UAVs to converge to the optimal solution, which cooperatively exchange information and learn by performing real-time measurements of the network capacity. For a single UAV, SPA is able to carry out a solution search similar to MPA. However, since a single UAV is employed to assess all possible positions at each iteration, SPA spends more time to determine a suitable solution. This limitation is addressed by SEEA considering a trade-off between searching accuracy and searching time. However, when compared to MPA, the convergence rate of SEEA is lower. Since the proposed model-free approaches rely on real-time measurements, their main drawback is the convergence time required to achieve an optimal solution when compared with a model-based approach. The authors have envisioned the combination of model-free and model-based approaches as the most suitable solution in practice. Thus, when it comes to the UAV placement problem, the literature has been focused on UAVs as APs and BSs serving ground devices, overlooking the backhaul network design.

Even though the network users are directly affected by the QoS provided by the access network, the backhaul network needs to be carefully designed in order to meet the variable traffic demand of the FAPs. Due to the emergence of flying networks comprised of flying platforms such as UAVs, balloons, High-Altitude Platform Stations (HAPSs), and Low Altitude Platforms (LAPs), as well as the recent advances in wireless communications technologies, the use of the vertical backhaul flying network depicted in Fig. 2.8, for transporting data traffic between access and core networks, has been proposed in [8]. The proposed flying network consists of Small Cell BSs combined with macro cells in a given geographical area. The Small Cell BSs with Line-of-Sight to the flying network exchange data traffic through Free-Space Optics-based links. On the other hand, Small Cell BSs with Line-of-Sight blocked to the flying network are served by nearby Small Cell BSs able to forward data traffic to the flying network.

In [9], the usage of tethered balloons has been proposed for enabling broadband wireless communications at remote ocean areas, as depicted in Fig. 2.9. The authors have created an FMN with proven ranges exceeding 50 km, while proposing a positioning control approach that



(a) Trajectories followed by six UAVs (each associated with a different color) starting from random positions until the optimal position is reached.



(b) Impact of the number of UAVs on the convergence to the optimal solution, considering average results for 1000 *Monte Carlo* runs. The smaller the number of UAVs, the lower the convergence rate achieved.

Figure 2.7: Performance of MPA proposed in [7].

allows determining the optimal position for each communications relay node in the network. The proposed approach allows maximizing the capacity of the flying network, given the positions of the end nodes, the number of communications relay nodes, and the signal reflections on the sea surface. The communications relay node positioning problem has been formulated as an optimization problem and solved using the Particle Swarm Optimization (PSO) technique. Results have shown that the proposed approach can assign the optimal height values for each flying com-

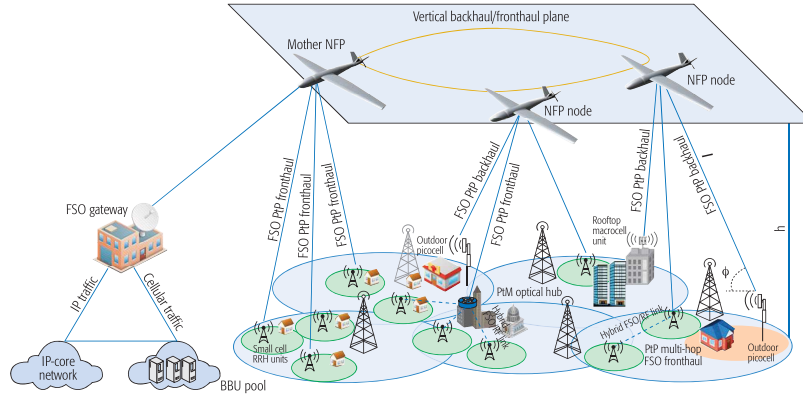


Figure 2.8: Multi-tier Free-Space Optics (FSO) based wireless fronthaul/backhaul network, where ultra-dense small cells are overlaid with macro cells [8].

munications relay, ensuring the lowest propagation losses and maximizing network throughput when compared with random and fixed height approaches. The main limitation of the proposed approach lies on the limited mobility of the tethered balloons, for which only altitude is controlled.

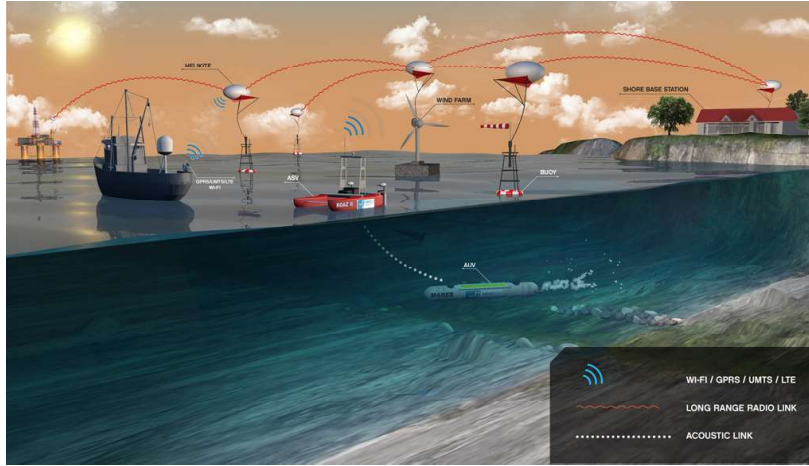


Figure 2.9: Broadband wireless backhaul network (wavy red lines) for remote oceans areas, enabled by tethered balloons acting as flying communications relays [9].

Motivated by the increasing usage of UAVs, especially due to their mobility in 3D space, FMNs have been proposed as a suitable solution to provide communications anytime, anywhere. Yet, in order to accommodate the traffic demand of RANs, backhaul networks deployed on-demand need to be highly dynamic and modular, while providing high enough capacity. For addressing this problem, this thesis proposes the GWP algorithm. GWP allows determining traffic-aware 3D positions for FGWs, taking advantage of the controlled mobility over UAVs. Hereafter, this will be the subject of this chapter, starting with the system model and the problem formulation.

2.2 System Model and Problem Formulation

The FMN consists of N UAVs that are controlled by a Central Station. Two types of UAVs are assumed to compose the FMN, as depicted in Fig. 2.10: 1) the FAPs, which provide Internet access to ground devices; 2) the FGW, which connect the FMN to the Internet. The FMN is organized into a multi-tier architecture, being especially targeted at taking advantage of short range high-directional air-air radio links, which provide high bandwidth channels. The Central Station can be deployed anywhere on the Edge or in the Cloud. The Central Station is in charge of periodically: 1) defining the updated positions of the FAPs by running a state of the art FAP placement algorithm [4] so that the FAPs meet the traffic demand of the ground devices, which is measured by the FAPs themselves and transmitted to the Central Station; and 2) determining the updated FGW position, in order to enable links that accommodate the traffic forwarded to/from the FAPs. The Central Station benefits from a holistic view over the network, including the updated positions and the traffic demand of the FAPs. This information is used to compute in advance the position of the FGW. Finally, the Central Station sends the updated positions to both the FAPs and the FGW, which configure and position themselves accordingly.

At the discrete time instant $t_k = k \cdot \Delta t, k \in \mathbb{N}_0$, where $\Delta t \gg 1s$ is the update period defined by the FAP placement algorithm, the FMN is represented by a directed graph $G(t_k) = (V, E(t_k))$. $V = \{0, \dots, N-1\}$ is the set of UAVs, which are placed at $P_i = (x_i, y_i, z_i)$ inside a cuboid x^{MAX} long, y^{MAX} wide, and z^{MAX} high. $E(t_k) \subseteq V \times V$ is the set of directional links between UAV $_i$ and UAV $_j$ at t_k , where $i, j \in V$, and $(i, j) \in E(t_k)$. The wireless channel between each pair of UAVs is modeled by the Free-Space Path Loss (FSPL) model, since a strong Line-of-Sight component dominates the wireless links between UAVs flying above the ground. Let us assume that UAV $_i$, $i \in \{1, \dots, N-1\}$, performs the role of FAP and transmits a traffic flow of bitrate $T_i(t_k)$ bit/s during time slot t_k towards UAV $_0$, acting as an FGW. The tree $T(V, E_T)$ is a sub-graph of G , where $E_T \subset E$ is the set of direct links between UAV $_i$ and UAV $_0$. This tree defines the FMN active topology. The flow $F_{i,0}$ with $T_i(t_k)$ bit/s demands a minimum capacity $C_{i,0}(t_k)$, in bit/s, for the wireless link available between UAV $_i$ and UAV $_0$ at time t_k . $F_{i,0}$ is received at UAV $_0$ from UAV $_i$ with bitrate $R_i(t_k)$ bit/s. The maximum channel capacity is equal to C^{MAX} bit/s. The wireless medium is shared and we assume that every UAV $_i$ can listen to any other UAV $_j$, including UAV $_0$. For this reason, the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism is employed for Medium Access Control (MAC), which is in charge of avoiding collisions of frames by enabling transmissions only when the channel is sensed to be idle.

Considering $C_{i,0}(t_k)$ as the capacity of the bidirectional wireless link established between UAV $_i$ and UAV $_0$ at time t_k , and $N-1$ UAVs generating a traffic flow $F_{i,0}$ with bitrate $T_i(t_k)$ bit/s towards UAV $_0$, we aim at determining at any time instant t_k the position of UAV $_0$, $P_0 = (x_0, y_0, z_0)$, and the transmission power P_T of the UAVs, considering P_T^{MAX} as the maximum transmission power

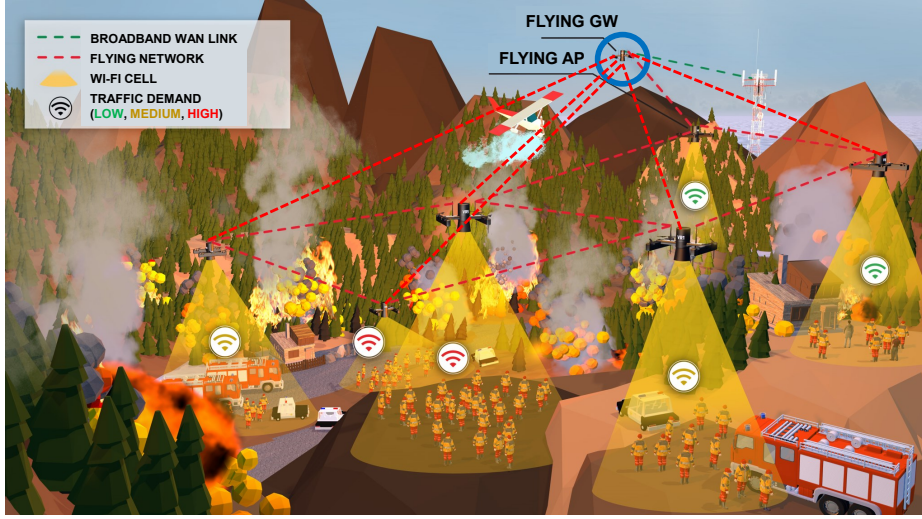


Figure 2.10: FMN providing on-demand broadband wireless connectivity to first responders in a disaster management scenario – the FGW is highlighted by the blue circle.

allowed for the wireless technology used, so that $C_{i,0}(t_k)$ is high enough to accommodate $T_i(t_k)$ bit/s and the overall network capacity, $C(t_k) = \sum_{i=1}^{N-1} C_{i,0}(t_k)$, is minimized. By minimizing $C(t_k)$, a lower transmission power is required, which in turn allows decreasing the interference between the communications nodes and the energy consumption, thus improving the overall network performance. Our objective function is defined in Eq. (2.1a), together with the constraints of the optimization problem.

$$\underset{P_T, (x_0, y_0, z_0)}{\text{minimize}} \quad C(t_k) = \sum_{i=1}^{N-1} C_{i,0}(t_k) \quad (2.1a)$$

$$\text{subject to: } 0 \leq P_T \leq P_T^{\text{MAX}} \quad (2.1b)$$

$$C(t_k) \leq C^{\text{MAX}} \quad (2.1c)$$

$$0 < T_i(t_k) \leq C_{i,0}(t_k), i \in \{1, \dots, N-1\} \quad (2.1d)$$

$$0 \leq x_i \leq x^{\text{MAX}}, i \in \{0, \dots, N-1\} \quad (2.1e)$$

$$0 \leq y_i \leq y^{\text{MAX}}, i \in \{0, \dots, N-1\} \quad (2.1f)$$

$$0 \leq z_i \leq z^{\text{MAX}}, i \in \{0, \dots, N-1\} \quad (2.1g)$$

$$(0, i), (i, 0) \in E(t_k), i \in \{1, \dots, N-1\} \quad (2.1h)$$

$$(x_0, y_0, z_0) \neq (x_i, y_i, z_i), i \in \{1, \dots, N-1\} \quad (2.1i)$$

$$\begin{aligned}
C &= B \times \log_2(1 + \text{SNR}_1) + B \times \log_2(1 + \text{SNR}_2) + B \times \log_2(1 + \text{SNR}_{N-1}) = \\
&B \times \log_2 \left[(1 + \text{SNR}_1) \times (1 + \text{SNR}_2) \times \dots \times (1 + \text{SNR}_{N-1}) \right] = \\
&B \times \log_2(u) \quad (2.2)
\end{aligned}$$

where:

$$u = \prod_{i=1}^{N-1} (1 + \text{SNR}_i)$$

The objective function defined in Eq. (2.1a) aims at minimizing the overall capacity $C(t_k)$ of the wireless links established between all the FAPs and UAV₀ (FGW) at each instant t_k . The wireless links' capacity is expressed by means of the Shannon-Hartley theorem. The Shannon-Hartley theorem allows computing the maximum amount of information that can be carried per unit of time by a communications channel subject to a certain SNR_{*i*}, using a given channel bandwidth B , as follows: $C = B \times \log_2(1 + \text{SNR}_i)$. This equation is a non-convex function [101, 102]. Since the objective function consists of the sum of multiple non-convex functions, each representing the capacity for the wireless link established between a FAP and the FGW computed using the Shannon-Hartley theorem, then it is a non-convex function too. The proof is presented in Eq. (2.2). A non-convex problem is non-deterministic polynomial-time hard (NP-hard) [103]. For this reason, we propose a heuristic algorithm to solve the FGW placement problem.

2.3 GWP Algorithm

The GWP algorithm is presented in this section, including its rationale and a theoretical analysis for a representative scenario.

2.3.1 Rationale

The GWP algorithm takes advantage of the centralized view of the FMN available at the Central Station, which is obtained as referred to in Section 2.2. For the sake of simplicity, we omit t_k in what follows. Considering the future positions of UAV_{*i*} and the bitrate of the traffic flow $F_{i,0}$, T_i , we aim at guaranteeing that a wireless link towards UAV₀ (FGW) has an SNR, SNR_{*i*}, that enables the usage of a Modulation and Coding Scheme (MCS) index, MCS_{*i*}, capable of transmitting T_i bit/s. Conceptually, if MCS_{*i*} is ensured, then the network will be able to accommodate T_i bit/s, which maximizes the amount of information received in the FGW. The selection of MCS_{*i*} imposes a minimum SNR_{*i*}, considering a constant noise power P_N . Then, if the transmission power P_T is known, we can compute the maximum distance d_i^{MAX} between UAV_{*i*} and UAV₀, using the FSPL

model defined in Eq. (2.3) in dB, where f_i is the carrier frequency and c represents the speed of light in vacuum.

$$\text{SNR}_i = P_T - 20\log_{10}(d_i^{\text{MAX}}) - 20\log_{10}(f_i) - 20\log_{10}\left(\frac{4 \times \pi}{c}\right) - P_N \quad (2.3)$$

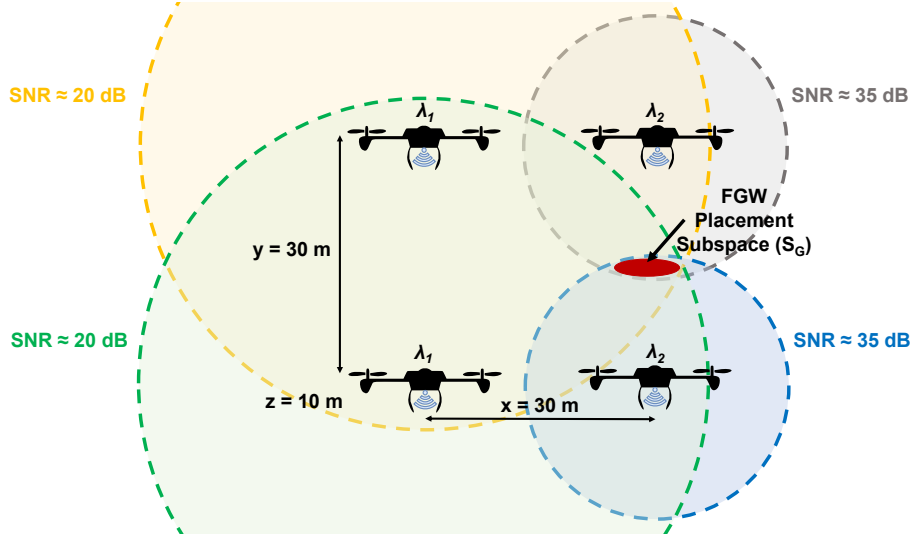


Figure 2.11: FGW Placement Subspace (S_G) in 2-Dimensional (2D) space, which results from the intersection of the spheres, centered at each UAV, with radius equal to the maximum distance compliant with the target SNR values.

In 3D, d_i^{MAX} corresponds to the radius of a sphere centered at UAV _{i} , inside which UAV₀ should be placed. Considering $N - 1$ UAVs, the placement subspace for positioning UAV₀ is defined by the intersection of the corresponding spheres $i \in \{1, \dots, N - 1\}$; we refer to this subspace as the FGW Placement Subspace, S_G , as depicted in Fig. 2.11. In order to simplify the process of calculating S_G , we propose Algorithm 1, which iteratively allows obtaining the point $P_0 = (x_0, y_0, z_0)$ for positioning UAV₀ and the transmission power P_T that we assume to be the same for all UAVs.

Table 2.1: Data rate and minimum SNR associated with Institute of Electrical and Electronics Engineers (IEEE) 802.11ac's MCS indexes 3 and 8 (channel bandwidth: 160 MHz and Guard Interval: 800 ns).

MCS index	MCS data rate (Mbit/s)	SNR (dB)
3	234	20
8	702	35

Algorithm 1: GWP algorithm

Input: Traffic demand T_i , in bit/s, of each FAP $_i$ and **position** P_i , in 3D Cartesian coordinates, of each FAP $_i$

Output: **Transmission (TX) power** P_T , in dBm, for all UAVs and **FGW position** P_0 in 3D Cartesian coordinates

```

/* 0dBm TX power. */
1  $P_T \leftarrow 0$ 
/* Iterate up to maximum allowed TX power -> O(n). */
2 while  $P_T \leq P_T^{MAX}$  do
    /* Compute the FGW position, with  $i \in \{1, \dots, N-1\}$ . SNR $_i$  is determined taking into
       account  $T_i$  by means of a lookup table similar to Table 2.1. */
3   computeFgwPos( $[x_i]$ ,  $[y_i]$ ,  $[z_i]$ ,  $[SNR_i]$ ,  $P_T$ )
    /* i.e.,  $(x_0, y_0, z_0) \in S_G$ . */
4   if  $(x_0, y_0, z_0) \neq \emptyset$  then
       /* TX power for all UAVs and FGW position. */
5       return  $P_T, (x_0, y_0, z_0)$ 
6   end
7   else
       /* Increase TX power by 1dBm. */
8        $P_T \leftarrow P_T + 1$ 
9   end
10 end
/* Compute and return  $x_0, y_0$ , and  $z_0 \rightarrow O(1)$ . In the best-case scenario, it has a time
   complexity of  $O(1)$ ; however, this depends on the algorithm employed to solve the
   system of equations. */
11 Function computeFgwPos ( $[x_i]$ ,  $[y_i]$ ,  $[z_i]$ ,  $[SNR_i]$ ,  $P_T$ ) :
    /* In the form of the system of Eq. (2.4). */
12    $(x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2 \leq \left(10^{\frac{K+P_T-SNR_1}{20}}\right)^2$ ;
13   ...;
14    $(x_0 - x_{N-1})^2 + (y_0 - y_{N-1})^2 + (z_0 - z_{N-1})^2 \leq \left(10^{\frac{K+P_T-SNR_{N-1}}{20}}\right)^2$ ;
15    $K = -20\log_{10}(f_i) - 20\log_{10}\left(\frac{4\pi}{c}\right) - P_N$ ;
16   return  $(x_0, y_0, z_0)$ ;

```

2.3.2 Theoretical Analysis for a Representative Scenario

Without loss of generality, we now exemplify the execution of Algorithm 1 for the representative scenario shown in Fig. 2.11; the algorithm is generic and may be applied to any traffic demand and number of FAPs. The scenario of Fig. 2.11 consists of four FAPs that are placed at the vertices of a square of side 30 m, hovering at an altitude of 10 m. We assume the use of the IEEE 802.11ac standard with one spatial stream, Guard Interval of 800 ns, and channel bandwidth of 160 MHz (channel 50 at 5250 MHz). Let us consider that the demanded capacity for the right-side FAPs is 702 Mbit/s, which is associated to the IEEE 802.11ac MCS index 8, and the demanded

capacity for the left-side FAPs is 234 Mbit/s, which is associated to the IEEE 802.11ac MCS index 3 [104]. These demanded capacity values were selected in order to exemplify the execution of Algorithm 1 for a representative networking scenario where the right-hand side FAPs have a traffic demand up to three times higher than the left-hand side FAPs, due to a greater number of ground devices served by the first FAPs. Taking into account the minimum SNR and the theoretical data rate of the IEEE 802.11ac MCS indexes, from Table 2.1 we conclude that the target SNR values in dB are respectively 20 dB for the left-side FAPs and 35 dB for the right-side FAPs, which were computed considering the minimum Received Signal Strength Indicator (RSSI) values proposed in [104] and a -85 dBm noise floor power.

$$\begin{cases} (x_0 - 30)^2 + y_0^2 + (z_0 - 10)^2 \leq \left(10^{\frac{K - P_T - 35}{20}}\right)^2 \\ (x_0 - 30)^2 + (y_0 - 30)^2 + (z_0 - 10)^2 \leq \left(10^{\frac{K - P_T - 35}{20}}\right)^2 \\ x_0^2 + (y_0 - 30)^2 + (z_0 - 10)^2 \leq \left(10^{\frac{K - P_T - 20}{20}}\right)^2 \\ x_0^2 + y_0^2 + (z_0 - 10)^2 \leq \left(10^{\frac{K - P_T - 20}{20}}\right)^2 \\ K = -20 \log_{10}(5250 \times 10^6) - 20 \log_{10}\left(\frac{4 \times \pi}{3 \times 10^8}\right) - (-85) \end{cases} \quad (2.4)$$

Solving the system of Eq. (2.4), which is derived from Eq. (2.3) considering d_i^{MAX} as the Euclidean distance between UAV_0 and UAV_i , we conclude that a suitable placement for the FGW is $(x_0, y_0, z_0) \approx (23.3, 15.4, 3.3)$ for a transmission power $P_T = 22$ dBm. P_T , which is the fine-tuning parameter in Eq. (2.4), is initially set to 0 dBm; then, it is iteratively increased by 1 dBm until a valid solution for the FGW position is found. If no solution is found, Algorithm 1 is terminated. In order to achieve a solution for the FGW placement, multiple FGW should be considered or the RAN placement should be adjusted, aiming at enabling shorter wireless links established with FAPs configured with transmission power P_T lower than or equal to P_T^{MAX} . In the system of Eq. (2.4), the carrier frequency f_i being used is considered: 5250 MHz herein; however, our conclusions are independent of f_i .

2.4 Performance Evaluation

The FMN performance achieved using the GWP algorithm is presented in this section, including the simulation setup, the simulation scenarios, and the performance metrics considered.

Table 2.2: Summary of the main ns-3 simulation parameters used in the evaluation of the GWP algorithm.

Parameter	Value
Simulation time	30 s (bootstrap) + 130 s
Wi-Fi standard	IEEE 802.11ac
Wi-Fi mode	Ad Hoc
Wi-Fi channel	50 (5250 MHz)
Channel bandwidth	160 MHz
Guard Interval	800 ns
Number of antennas	1
Propagation delay	Constant speed
Propagation model	FSPL (Friis)
Remote station manager	<i>IdealWifiManager</i>
Mobility model	Waypoint
Traffic type	UDP Poisson
Packet size	1400 Bytes
Traffic control	CoDel

2.4.1 Simulation Setup

In order to evaluate the FMN performance achieved with the GWP algorithm, the ns-3 simulator was used. A NIC was configured on each UAV in Ad Hoc mode, using the IEEE 802.11ac standard in channel 50, with channel bandwidth of 160 MHz, and Guard Interval of 800 ns. One spatial stream was used for all inter-UAV links. The traffic generated was User Datagram Protocol (UDP) Poisson for a constant packet size of 1400 bytes. The data rate was automatically defined by the *IdealWifiManager* mechanism. The traffic generation was only triggered after 30 s of simulation, in order to ensure a stable state, with a total simulation time of 130 s. The Controlled Delay (CoDel) algorithm [105], which is a Linux-based queuing discipline that considers the time that packets are held in the transmission queue to discard packets, was used; it allows mitigating the bufferbloat problem. The default parameters of CoDel in ns-3 were employed [106]. The main simulation parameters used in the evaluation of GWP are detailed in Table 2.2.

2.4.2 Simulation Scenarios

In addition to the FGW position that was computed using the GWP algorithm, other positions for the FGW in the venue depicted in Fig. 2.12a were evaluated, in order to show the performance gains obtained when using the GWP algorithm; the seven additional positions considered are depicted in Fig. 2.12a and hereafter referred to as Scenario A. Position 1 to position 7 were defined

to allow an inter-position distance of 7.5 m; they aimed at exploring the vertical and horizontal corridors of the venue. We define as baseline the FGW placed in the FAPs center (i.e., three-coordinates average considering all FAPs), which is the position that maximizes the SNR of the wireless links between each FAP and the FGW; the placement of the FGW in the central position is commonly considered in the literature [4]. In turn, position 8 represents the FGW placement derived from the system of Eq. (2.4). In order to evaluate the performance achieved when using the GWP algorithm in a typical crowded event, a more complex scenario, depicted in Fig. 2.12b and hereafter named Scenario B, was also considered. It represents an FMN consisting of 10 FAPs and 1 FGW, inside a cuboid of dimensions $80\text{ m} \times 80\text{ m} \times 20\text{ m}$. The FAPs were randomly positioned in order to form two zones with different traffic demand: λ_1 bit/s and λ_2 bit/s, as illustrated in Fig. 2.12b. The traffic demand of the FAPs in the ns-3 simulations was defined assuming a reference fair share $L = (780\text{ Mbit/s})/(N - 1)$, where 780 Mbit/s is the data rate associated to the maximum MCS index of the IEEE 802.11ac standard and $N - 1$ is the number of FAPs generating traffic, considering one spatial stream, channel bandwidth of 160 MHz, and Guard Interval of 800 ns. For Scenario A (cf. Fig. 2.12a), we considered traffic demand $\lambda_2 = 3 \times \lambda_1 = 0.75 \times L$. For Scenario B (cf. Fig. 2.12b), two different traffic demand combinations were considered: i) $\lambda_2 = 9 \times \lambda_1 = 0.9 \times L$, and ii) $\lambda_2 = 3 \times \lambda_1 = 0.75 \times L$.

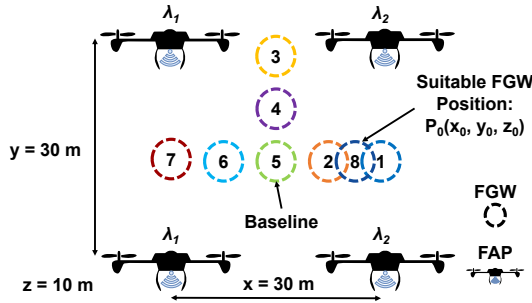
Since the GWP algorithm relies on knowing in advance the positions of the FAPs, provided by the FAP placement algorithm in a real-world deployment, instead of generating the random waypoints during the ns-3 simulation we used *BonnMotion* [107], a mobility scenario generation tool. The resulting waypoints were considered to compute in advance the FGW position using the GWP algorithm. Finally, the computed FGW positions over time and the generated scenarios were imported to ns-3, with a sampling period of 1 s. The *WaypointMobilityModel* module was used in ns-3 to place the FAPs and the FGW in the positions generated by *BonnMotion* and defined by the GWP algorithm, respectively.

The GWP algorithm was evaluated considering three counterpart approaches: 1) the FGW placed at the position that maximizes the SNR of the wireless links between each FAP and the FGW; 2) the FGW placed in the center of the venue; and 3) the FGW randomly positioned over time, following the *RandomWaypointMobility* model with velocity uniformly distributed between 0.5 m/s and 3 m/s, inside the venue of Fig. 2.12b.

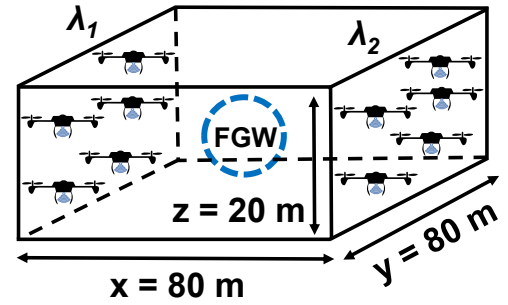
2.4.3 Performance metrics

The performance achieved with the GWP algorithm was evaluated considering two metrics:

- **Aggregate throughput:** the average number of bits received per second by the FGW.



(a) Scenario A, in which different positions for the FGW were evaluated. Position 8 corresponds to the FGW position computed by means of the GWP algorithm, while position 5 corresponds to the baseline – FGW placed in the FAPs center.



(b) Scenario B, in which 10 FAPs were randomly positioned in order to form two zones with different traffic demand: λ_1 and λ_2 . The baseline, represented by a dashed circumference, corresponds to the FGW placed in the FAPs center.

Figure 2.12: Simulation scenarios used in the evaluation of the GWP algorithm.

- **Delay:** the average time taken by the packets to reach the sink application of the FGW since the instant they were generated by the source application of each FAP, including queuing, transmission, and propagation delays.

2.4.4 Simulation results

The simulation results are presented in this section. The results were obtained after 20 simulation runs for each traffic demand combination considered (cf. Section 2.4.2), under the same exact networking conditions. The results are expressed using average values and they are represented using the Cumulative Distribution Function (CDF) for the delay and by the Complementary Cumulative Distribution Function (CCDF) for the aggregate throughput. The CCDF $F'(x)$ represents the percentage of time for which the average aggregate throughput was higher than x , while the CDF $F(x)$ represents the percentage of time for which the average delay was lower than or equal to x .

Regarding Scenario A, when the FGW is placed according to the GWP algorithm (Position 8 in Fig. 2.12a), the aggregate throughput is improved 24% for the 90th percentile and 21% for the 50th percentile (median), with respect to the baseline (i.e., the FGW placed in the FAPs center). In parallel, the delay is decreased 26% for both the 90th and 50th percentiles (cf. Fig. 2.13). The similar results obtained for Position 2 and Position 8, which are depicted in Fig. 2.13, are justified by the closer distance between these positions; note that Position 2 was obtained by chance, while Position 8 resulted from the GWP algorithm. In order to meet the higher traffic demand of the right-side FAPs, the GWP algorithm places the FGW closer to them, in order to improve the SNR of the communications links and enable the selection of higher MCS indexes. This

improves the overall FMN performance and the shared medium usage – the packets are held in the transmission queues for shorter time, the transmission delay decreases, and the throughput increases.

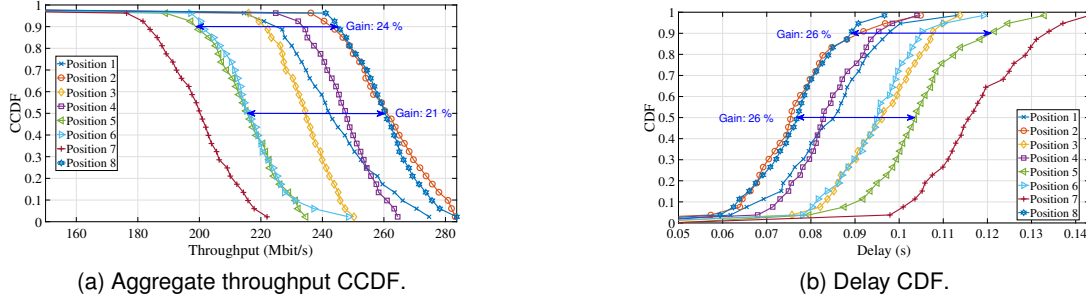


Figure 2.13: Scenario A – Aggregate throughput and delay results measured in the FGW. Position 8 was defined by the GWP algorithm.

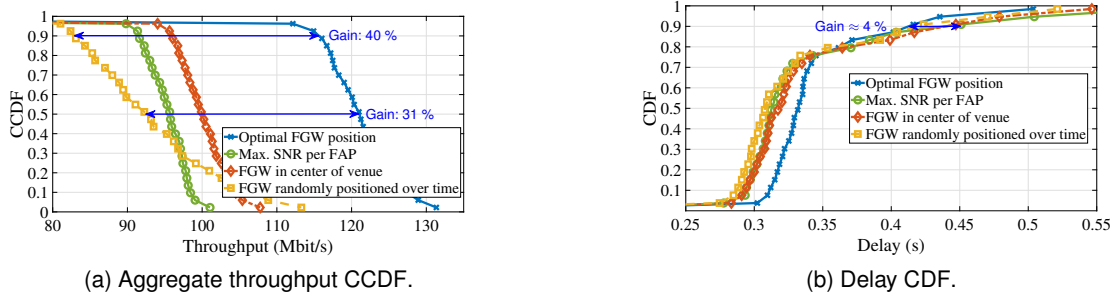


Figure 2.14: Scenario B – Aggregate throughput and delay results measured in the FGW for λ_1 and λ_2 equal to 10% and 90% of the reference fair share L , respectively.

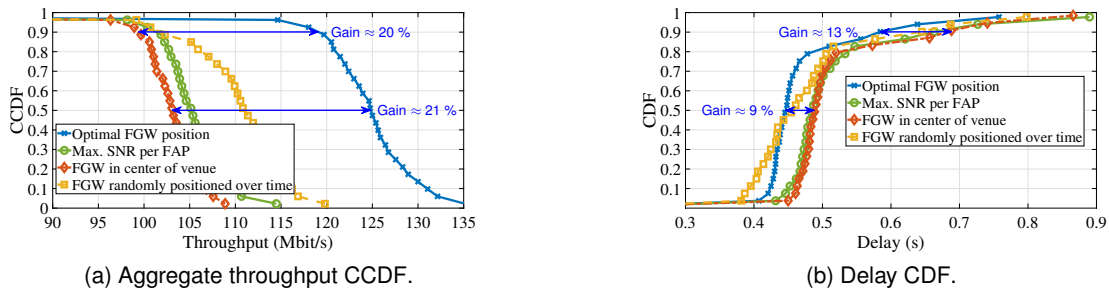


Figure 2.15: Scenario B – Aggregate throughput and delay results measured in the FGW for λ_1 and λ_2 equal to 25% and 75% of the reference fair share L , respectively.

With respect to Scenario B, when λ_1 and λ_2 are respectively equal to 10% and 90% of the reference fair share L , the GWP algorithm allows improving the aggregate throughput up to 40%, considering the 90th and 50th percentiles, while the delay is reduced up to 4% (cf. Fig. 2.14). When λ_1 and λ_2 are respectively equal to 25% and 75% of L , the GWP algorithm also improves the aggregate throughput in 20% with respect to the 90th percentile and 21% for the 50th percentile; the delay is reduced 13% for the 90th percentile and 9% for the 50th percentile (cf. Fig. 2.15). These results validate the effectiveness of the GWP algorithm and confirms that the FMN performance can be improved by dynamically adjusting the position of the FGW over time, considering information about the traffic demand and the positions of the FAPs.

It is worth mentioning that the networking scenarios employed in the performance evaluation of GWP were defined in order to ensure the FAPs listen to each other, thus avoiding the hidden node problem. In fact, considering the minimum transmission power values of 1) 22 dBm for Scenario A (cf. Fig. 14), and (2) 24 dBm for Scenario B (cf. Fig. 16) (defined by the GWP algorithm), consisting of a pair of FAPs placed on diagonally opposite sides, it can be concluded that the SNR values at the FAPs are, respectively, 28 dB and 21 dB (considering the FSPL model and a -85 dBm noise floor power). These SNR values are high enough to ensure the FAPs listen to each other. In order to avoid the hidden node problem in real-world deployments where networking scenarios are not fully controlled, it is worthwhile to improve the GWP algorithm so that the same transmission power defined for the FAPs is high enough not only to ensure the minimum SNR required to meet the FAPs' traffic demand, but also to ensure that the FAPs listen to each other. Since our focus is the traffic-aware UAV placement problem, the MAC challenges are left for future work.

2.5 Discussion

The GWP algorithm allows determining suitable positions for FGWs so that the Euclidean distances to the FAPs allow ensuring wireless links with the targeted SNR values. By ensuring the targeted SNR values, GWP aims at inducing the MAC automatic rate mechanism to select for each wireless links an MCS index with a physical data rate higher than or equal to the traffic demand. Since GWP does not control the operation of the MAC automatic rate mechanism used by the communications nodes, the real capacities of the wireless links can present slight changes when compared with the theoretical values, especially due to deviations in the SNR in some real-world networking scenarios. In order to address the problem, the use of an SNR margin with respect to the theoretical values obtained by means of the FSPL model or any other wireless propagation model may be considered in practice.

When running the GWP algorithm, it may happen that:

- (a) it is possible determining an intersection volume between the spheres representing the transmission ranges of all FAPs (i.e., FGW placement subspace) for transmission power lower than or equal to the maximum transmission power allowed by the NICs carried on-board the UAVs;
- (b) it is not possible finding an intersection volume between the spheres associated to all FAPs. This may occur especially due to practical transmission power levels constraints associated to the NICs carried on-board the UAVs, considering that GWP uses transmission power as the fine-tuning parameter to achieve a suitable FGW position.

When it is not possible defining an intersection volume considering all FAPs, the use of multiple FGWs may be considered. For that purpose, different clusters of FAPs may be formed and a dedicated FGW, with position defined by means of the GWP algorithm, may be assigned to each cluster. In such case, some FAPs associated to a cluster may also be suitable to be served by the FGW of a nearby cluster. Moreover, if the wireless channels used by the FGWs associated to some clusters have sufficient remaining network capacity to accommodate more traffic, load balancing among different FGWs may be considered.

The use of multiple FGWs may impose additional research challenges regarding defining the clusters of FAPs with the ultimate goal of maximizing the aggregate network performance considering all FGWs. This may include multiple trade-offs in the network design, including using a dedicated wireless channel per FGW or different channels for all FGWs. Sharing a wireless channel may involve taking into account the hidden node problem. It can also be discussed whether all FGWs use a single remote BS/AP or different BSs/APs to forward data traffic to/from the Internet, whose position can be also taken into account in the FGW placement problem. Finally, a single FGW, placed by the GWP algorithm closer to the network Edge, while using a high-capacity wireless channel that allows aggregating all data forwarded to/from UAVs acting as flying communications relays for each cluster of FAPs, may be also considered; this is especially suitable for coverage and capacity extension in remote networking scenarios. Evolving GWP for multiple FGWs and addressing the associated research challenges are left for future work. It is worth mentioning that the current version of the GWP algorithm can also be used to define the positioning of an FGW in a networking scenario consisting of fixed terrestrial Wi-Fi APs or cellular BSs.

In the GWP algorithm we assume that the overhead of UDP, Internet Protocol (IP), and MAC packet headers is negligible; this is compliant with emerging wireless communications technologies, such as IEEE 802.11ax and 5G New Radio, where Orthogonal Frequency-Division Multiple Access (OFDMA) and frame aggregation mechanisms improve the MAC efficiency [108]. Moreover, GWP provides the same output whether downlink or uplink traffic is considered, since all the UAVs are configured with the same transmission power and the wireless channel is assumed to

be symmetric. This paves the way for using the GWP algorithm in emerging networking scenarios where symmetric traffic applications are growing [109], such as social networks, video streaming, and online gaming. In networking scenarios with asymmetric wireless links, the lowest capacity among the two wireless link directions should be considered.

It is worth mentioning that in [85] a successor of the GWP algorithm has been proposed to solve the FGW placement problem using an optimization-based approach. Yet, the authors have modeled the wireless links established between the FAPs and the FGW using the approach originally proposed by the GWP algorithm. This has allowed overcoming the NP-hard nature of the problem referred to in Section 2.2. For that purpose, they have formulated an optimization problem able to determine the minimum transmission power required for ensuring a minimum SNR and inducing a targeted MCS index capable of accommodating the offered traffic.

2.6 Summary

When it comes to the UAV placement problem, existing solutions are focused on enhancing the wireless coverage, the number of ground devices served, and the QoS offered by the RAN. The placement of UAVs acting as flying communications relays and FGWs to/from the Internet has been overlooked in the literature, where an unconstrained backhaul network is commonly considered. In this chapter we presented GWP, a centralized traffic-aware FGW placement algorithm for FMNs. GWP takes advantage of the knowledge of the traffic demand and the future positions of the FAPs, controlled by a state of the art FAP placement algorithm running in a Central Station, to enable backhaul communications paths with high enough capacity to accommodate the traffic exchanged between the RAN and the Internet. For that purpose, the GWP algorithm ensures that the wireless link established between each FAP and the FGW has a minimum SNR that enables using of an MCS index with a physical data rate capable of carrying a given traffic demand; this allows maximizing the amount of bits received in the FGW. The performance evaluation of FGW was carried out through ns-3 simulations. The obtained results show gains up to 40% in the aggregate throughput, while the delay is reduced up to 26%.

Even though GWP is able to reduce the overall network delay, it is not designed to meet strict delay constraints for the transported traffic. For ensuring delay-sensitive communications, the flexibility provided by UAVs to be placed so that short-distance Line-of-Sight wireless links are ensured is a step forward [110]. Yet, new challenges from the network configuration point of view should be taken into account. In particular, the size of the network queues configured in the communications nodes, which are used for accommodating transient traffic bursts offered to the network, needs to be managed. A mechanism addressing the dynamic queue size management to enable controlled delays in FMNs is worth considering and is the subject of the next chapter.

Chapter 3

A Mechanism for Proactive Queue Management in Flying Multi-hop Networks

In Chapter 2, we presented the algorithm proposed by this thesis for the placement of gateways in Flying Multi-hop Networks (FMNs), named GateWay Placement (GWP). GWP allows maximizing the aggregate throughput measured in the Flying Gateway (FGW). Yet, it does not take into account delay constraints for placing the FGW. The interest in low-latency communications is emerging. As such, besides the ability to accommodate the traffic demand, FMNs are expected to meet strict low-latency requirements. The goal is to reliably transmit high bitrates under controlled delays, thus enabling human and machine-oriented communications. A critical aspect that must be addressed is the delay introduced by network queues. The literature has been focused on addressing the queue size in wired networks. Wireless networks bring up additional challenges due to their dynamic nature. This is exacerbated in FMNs.

In this chapter, we present the Proactive Queue Management (PQM) mechanism for FMNs proposed by this thesis. PQM takes advantage of the knowledge of the future positions of the Flying Access Points (FAPs) and FGWs composing the FMN, in order to define the queue size for each FAP that allows maximizing the aggregate throughput with stochastic delay guarantees. We assume the FAPs' positions are defined by a state of the art FAP placement algorithm, while the FGWs' positions are defined by the GWP algorithm presented in Chapter 2. Since PQM configures only the queue size of the communications nodes, it represents a step forward with respect to state of the art solutions on network queue size management, which typically require the fine-tuning of several parameters and are designed for static networks.

We start the chapter by presenting the related work on solutions for queue size management and active queue management in delay-sensitive wireless networks. Next, we define the system model, formulate the problem, describe the PQM mechanism, and refer to the performance evaluation when PQM is used in an FMN. We end the chapter by summarizing the contents presented.

3.1 Related Work

The increasing usage of a variety of communications services, such as interactive multimedia services, interactive tactile Internet applications, and reliable and latency-sensitive remote control of machines, imposes orthogonal requirements in the design of Radio Access Networks (RANs) and backhaul networks [1, 49]. Besides the large amount of traffic that the RANs and the backhaul networks should accommodate, the interest in low-latency communications is emerging. This brings significant challenges from the network configuration point of view, in order to ensure low delays and high throughput simultaneously [111]. These requirements are envisioned in the recently released 5G and the emerging 6G networks, aimed at providing differentiated services for different verticals, including massive Internet-of-Things, massive Machine-Type Communications (mMTC), Ultra-Reliable Low-Latency Communications (URLLC), and enhanced Mobile Broadband (eMBB).

A critical aspect that should be addressed to meet the requirements of emerging communications services is the delay introduced by the queues installed on communications nodes. Queues are used to accommodate temporary traffic bursts offered to the network; they are also a problem in emerging networks, where different services should coexist in Quality of Service (QoS) classes sharing the same network queues. Large queues are typically used to improve throughput by reducing the number of dropped packets; however, they result in long queuing delays, a problem commonly known as bufferbloat [71]. Conversely, small queues can lead to communications nodes not being able to cope with the offered traffic. The ultimate goal lies in designing queues large enough to accommodate traffic bursts, while minimizing the queuing delay. The literature has been focused on addressing the queue size management in wired networks. However, wireless networks raise additional challenges, including 1) shared channels, in which the available capacity is limited by the number of neighboring nodes, and 2) frequent changes in the wireless links' Signal-to-Noise Ratio (SNR), which lead to network topology reconfiguration and variable network capacity over time. These constraints dictate that queues of static size are not suitable for wireless networks, which are assumed to be dynamic by nature. The queue size management challenge is exacerbated in FMNs, which are highly dynamic especially when it comes to the communications nodes' positions.

3.1.1 Queue Size Management in Wireless Networks

With the growing usage of applications with strict requirements regarding throughput and delay, the research on queue size management in wireless networks is emerging [111–113]. In [114], the performance of three different queue size management algorithms has been evaluated: emulated Bandwidth-Delay Product, Adaptive Limit Tuning, and A*. The authors have concluded

that IEEE 802.11 networks introduce a variety of challenges when it comes to queue size management. First, IEEE 802.11 transmissions result in significant differences between the packet service times at each Station (STA). Since IEEE 802.11 ensures that all STAs have an equal number of transmission opportunities, the average packet service time at a given STA is one order of magnitude longer when ten other STAs are active, considering as baseline a single active STA. As such, the queue size at each STA should be dynamically defined based on the number of active STAs in the network. In addition to changes in the average service time, the distribution of the packet service times depends on the traffic demand. This directly influences the occurrence of traffic bursts and the required queue size. Second, IEEE 802.11 STAs dynamically control packet losses in non-saturated channels by selecting suitable Modulation and Coding Scheme (MCS) indexes. The Medium Access Control (MAC) automatic rate adaptation, in which the maximum physical data rate may vary significantly according to the MCSs indexes used, motivate a dynamic queue size. Thirdly, the ongoing IEEE 802.11 improvements aim at continuously enhancing the throughput achieved in wireless networks by aggregating multiple packets into larger frames. This creates a trade-off between throughput and queue size, as the latter directly influences the packets made available to be aggregated into frames before their transmission. For these reasons, using a predefined fixed queue size on all communications nodes is not the most suitable approach in wireless networks.

In [10], an aggregation-aware queue management scheme for wireless networks, called WQM, has been proposed. WQM differentiates good and bad queues. Good queues are required to accommodate traffic bursts, whereas bad queues introduce network delay without gains in throughput. WQM uses data traffic being exchanged as a probe for network measurements, which avoids using dedicated probe packets and introducing overhead. In order to estimate the channel capacity, WQM periodically calculates the time required to empty the queue based on the current transmission data rate. It adjusts the queue size to ensure the targeted QoS, while reducing queuing delays and fulfilling the available network capacity. When compared with the default queue size used in *Linux* based communications nodes, for example, WQM reduces the end-to-end delay up to eight times with a slight decrease in network goodput, as depicted in Fig. 3.1. It also has lower delay when compared with Controlled Delay (CoDel) [105].

A proactive queue management algorithm has been proposed in [115]. It allows managing the queue size in order to achieve low queuing delay, high nodes' usage, and low packet loss. The average queue length is considered as a congestion indicator. Periodically, according to a predefined Sampling Interval (si) in seconds, the instantaneous queue size is measured. si is defined according to a Prediction Interval (pi). Considering pi/si measurements of the instantaneous queue length, the average queue length over the previous prediction interval is computed at the start of each pi . The average value obtained is used to predict the average queue length for the next prediction interval, using the Recursive Least Square algorithm. Recursive Least Square

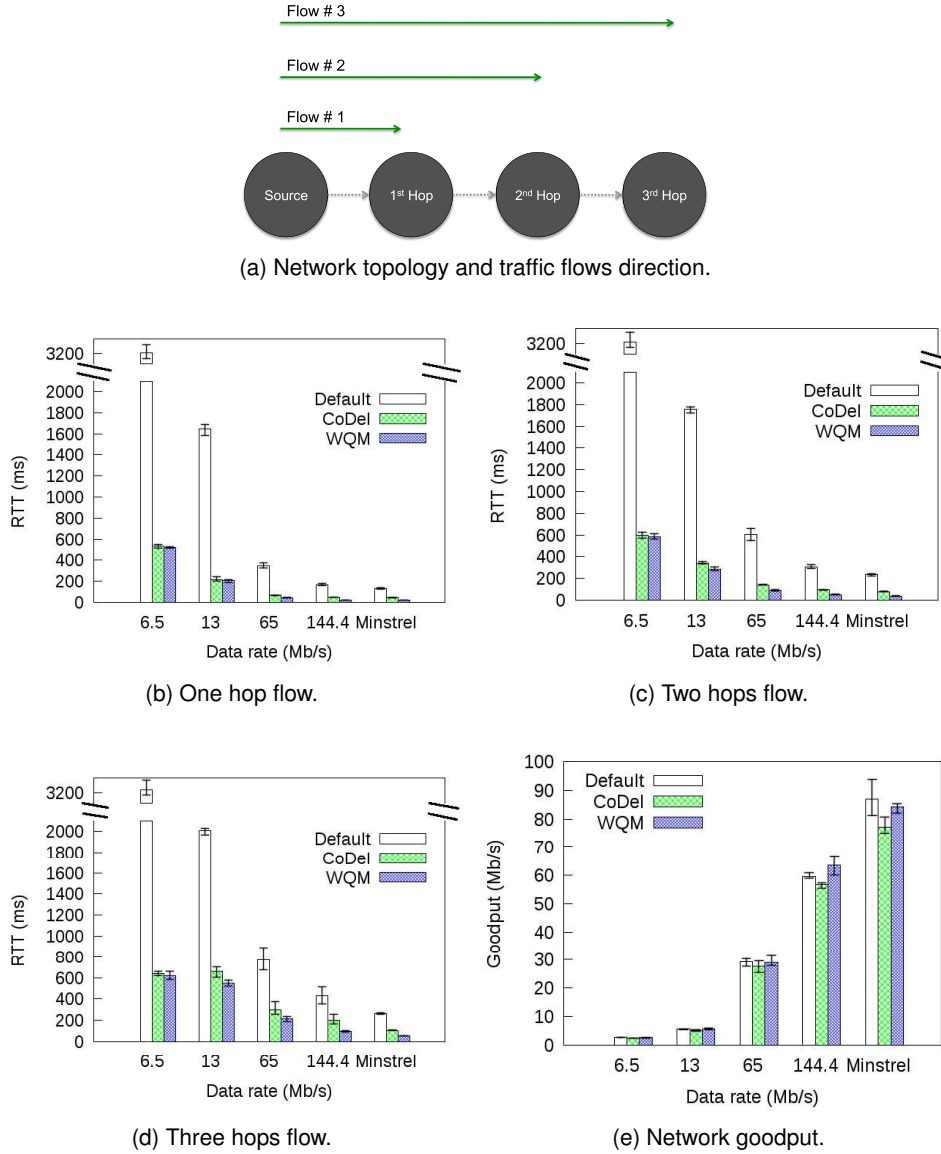


Figure 3.1: Average end-to-end delay per flow and total goodput achieved when using the state of the art WQM solution in the wireless network topology depicted in Fig. 3.1a. Error bars represent maximum and minimum values considering at least three runs. WQM allows reducing the end-to-end delay by $8\times$ for one hop, $6\times$ for two hops, and $4\times$ for three hops, without compromising the goodput. The results confirm that large queues (*Default* approach considered) are typically not needed in wireless networks [10].

allows considering an adaptive learning function based on the average queue lengths computed. The packet dropping probability is determined based on the deviation between the predicted average queue length and the stable queue length. The instantaneous queue length q_i is compared to the stable queue length q_0 at each packet arrival. If q_i is lower than q_0 , the incoming packet

is enqueued; otherwise, it is dropped according to the packet dropping probability determined. This solution does not allow ensuring differentiated services, including different delay guarantees. Moreover, the packet drop probability is computed only once at the beginning of each pi and stays constant throughout that interval. The pi value should consider a trade-off between computational cost and solution responsiveness. On the one hand, a higher pi value results in a slower response to changes in the offered traffic. This can cause unstable behavior and compromise performance, especially in wireless links with varying network capacity over time, including FMNs. On the other hand, a low pi value results in a quick response to changes in the offered traffic at the expense of an increase in the computational cost for rapidly updating the packet dropping probability.

3.1.2 Active Queue Management

In order to avoid using fixed size queues on the communications nodes, different Active Queue Management (AQM) algorithms have also been proposed in the literature. They aim at dropping packets based on different strategies, including queue size, packet size, queuing delay, link utilization, and traffic classification. Some examples are Drop-Tail, Random Early Detection (RED), BLUE, Proportional Integral (PI), CoDel, and Proportional Integral controller Enhanced (PIE). Drop-Tail, which is the default queue mechanism in *Linux* based communications nodes, employs a static queue size; the packets are only enqueued until this value is reached. The main disadvantage of Drop-Tail lies on the fact that packets can be held in the queue for long time if the size of the queue is not correctly defined, especially on links in which the network capacity is significantly lower than the offered traffic.

RED computes average queue size using a low-pass filter with an exponential weighted moving average [116]. The average queue size is compared with minimum and maximum thresholds; every arriving packet is marked when the average queue size exceeds the maximum threshold. The average queue size does not significantly exceed the maximum threshold if marked packets are dropped or the source nodes cooperate between themselves. When the average queue size falls between the minimum and maximum thresholds, each arriving packet is marked with probability p_a , where p_a depends on the average queue size. One of the limitations of RED lies in estimating congestion based on queue occupation. Although the presence of a full queue is associated with congestion, it does not allow to infer the number of communications nodes and flows competing for a wireless link. For instance, a single source transmitting temporarily at a data rate greater than the capacity of the link may have the same effect on congestion as multiple communications nodes generating traffic dictate. Moreover, RED requires configuring a wide set of parameters for adequate performance in different congestion scenarios. Fine-tuning these parameters is a challenge, especially in highly dynamic networks.

In order to avoid the need of fine-tuning several parameters, BLUE has been proposed in [117]. The rationale of BLUE lies in performing queue management based on packet loss and

network usage rather than considering queue occupancy for congestion detection. When packets are queued, BLUE sets a probability to drop them. If the queue is consistently dropping packets because it is full, BLUE increases the drop probability; therefore, it increases the rate at which a congestion notification is sent back to the source communications nodes. If the queue is empty or if the link is idle, the drop probability defined by BLUE is reduced. This allows BLUE to learn the rate at which it should send a congestion notification. Yet, BLUE is designed to tackle non-responsive traffic, where the data rate is limited to a fixed amount of the bottleneck bandwidth of the link or a fair share of the capacity of the link.

An additional limitation of RED, intrinsic to its controller, is the slow response time. In [118], two AQM alternatives have been proposed: proportional and PI controllers [118]. The former is easier to implement than RED and the latter improves network performance with a complexity comparable to RED. Both controllers enable lower response times than the RED controller, while the PI controller attains superior performance in controlling the steady-state value of the queue size. The PI controller aims at ensuring a predefined reference value for the queue size. This reference value is a configuration parameter that allows ensuring controlled Round-Trip Times (RTTs) and packet loss values. The steady-state regulation is achieved by the integral (I) action of the PI controller. In turn, the controller's proportional (P) component allows ensuring stability margins and reaction without a precise prior knowledge of Transmission Control Protocol (TCP) load and propagation delay. However, PI reacts slowly to changes, which makes it not suitable for high dynamic networks and variable traffic demand.

Contrary to the previous AQM algorithms, CoDel relies on the sojourn time – the time that the packets are held in the queue [105]. As a more reliable and precise indicator of congestion, CoDel's authors have used the local minimum queuing delay. Moreover, rather than maintaining a window of values to compute the minimum delay value, they concluded that it is sufficient to use a single state variable whose value represents how long the minimum has been above or below the targeted queuing delay. When a packet arrives and the queue is full, the packet is dropped. CoDel detects persistent delays by tracking the local minimum queuing delay that packets experience. To avoid congestion, the minimum value must have been experienced within the most recent interval. When the queuing delay exceeds the target value for at least an interval, a packet is dropped, and the next drop time is determined by a control law. The controller stops dropping packets when the queuing delay is below the targeted value. As such, CoDel requires time-stamping the packets at their arrival in the queue, which implies significant computational processing for high traffic load. In addition, CoDel requires defining two configuration parameters: *target* and *interval*. *Target* stands for the acceptable queuing delay, while *interval* should be a value in the order of the worst-case RTT associated to the bottleneck link. The authors have determined values for *target* and *interval* that allow for high utilization and controlled delays, considering different link's bandwidths, RTTs, and traffic loads. Defining suitable values in highly dynamic networks, including FMNs, is

a challenge.

Similarly to CoDel, PIE aims at controlling the average queuing delay [119]. It combines the benefits of both RED and CoDel: PIE is easier to implement than RED and allows controlling the delay as CoDel. Similarly to RED, PIE randomly drops packets at the onset of congestion. However, the congestion detection is based on the queuing delay, as CoDel, rather than queue occupancy as most of the state of the art AQM mechanisms do, including RED. Moreover, PIE also uses information regarding queuing delay variation for determining congestion levels.

Overall, as stated in [120], designing a single optimal queue management solution that suits the variable characteristics of wireless networks is not straightforward, especially due to the need of fine-tuning many parameters. Therefore, a proactive queue management solution fed by a holistic view over the network and able to define in advance the queue size according to the network topology and traffic demand of the communications nodes, while addressing the variable wireless channel characteristics over time, is worth considering. This is the subject of this chapter.

3.2 System Model and Problem Formulation

The FMN is assumed to have an architecture similar to the one considered in Chapter 2. We are focused on the backhaul network, consisting of the wireless links used to forward traffic between the FAPs and the FGW connecting the FMN to the Internet. While in Chapter 2 we were focused on defining the placement for the FGW, herein we address the queue size management so that controlled delays are ensured.

The FMN consists of N Unmanned Aerial Vehicles (UAVs) that are controlled by a Central Station. Two types of UAVs are assumed to compose the FMN: 1) the FAPs providing wireless connectivity to the ground devices; 2) the FGW forwarding data traffic between the FMN and the Internet. The Central Station can be deployed anywhere on the Internet. It is responsible for periodically: 1) defining the updated positions of the FAPs by running a state of the art FAP placement algorithm, such as the algorithm proposed in [4], in order to meet the traffic demand of the ground devices; 2) calculating the FGW position by means of an FGW placement algorithm, such as the one proposed in Chapter 2; and 3) determining the queue size for each FAP and the time instants it shall be updated by employing PQM. The Central Station has a holistic view over the FMN, including the updated positions and the traffic demand of the FAPs, which are assumed to be provided by the FAP placement algorithm; this information is used to compute in advance the FGW position and the queue size for each FAP over time. The Central Station is in charge of sending the updated positions to both the FAPs and FGW as well as the queue size for each FAP, which position and reconfigure themselves accordingly.

At the discrete time instant $t_k = k \cdot \Delta t, k \in \mathbb{N}_0$, where $\Delta t \gg 1s$ is defined according to the FMN update period imposed by the FAP placement algorithm, the FMN is represented by a directed

graph $G(t_k) = (V, E(t_k))$, where $V = \{0, \dots, N-1\}$ is the set of UAVs, with UAV_{*i*} positioned at $P_i = (x_i, y_i, z_i)$, $E(t_k) \subseteq V \times V$ is the set of directional wireless links between UAV_{*i*} and UAV_{*j*} at t_k , $i, j \in V$, and $(i, j) \in E(t_k)$. The wireless channel between two UAVs is modeled by the Free-Space Path Loss (FSPL) model, which is justified by the strong Line-of-Sight component that dominates the wireless links between UAVs flying above the ground. We define $C_{i,j}(t_k)$ as the capacity, in bit/s, of the wireless channel available between UAV_{*i*} and UAV_{*j*} at time t_k . UAV_{*i*}, $i \in \{1, \dots, N-1\}$, performs the role of FAP and transmits a traffic flow of bitrate $T_i(t_k)$ bit/s during time slot t_k towards UAV₀, which acts as FGW. The FMN active topology is defined by a tree $T(V, E_T)$ that is a sub-graph of G , where $E_T \subset E$ is the set of direct links between UAV_{*i*} and UAV₀. The flow $F_{i,0}$ is received at UAV₀ from UAV_{*i*} with bitrate $R_i(t_k)$ bit/s and delay $D_i(t_k)$ seconds. The wireless medium is shared and we assume that every UAV_{*i*} is in the same collision domain, including UAV₀. The MAC mechanism is employed for MAC.

Considering the throughput $R_i(t_k)$ as the bitrate of the flow $F_{i,0}$ at time t_k , $D_i(t_k)$ as the average delay of the packets composing the flow $F_{i,0}$ at time t_k , and $N-1$ FAPs generating traffic towards UAV₀ (FGW), we aim at determining at any time instant t_k the queue size Q_i of each FAP_{*i*} such that the aggregate throughput $R(t_k) = \sum_{i=1}^{N-1} R_i(t_k)$ is maximized and the average delay constraint $D_i(t_k) \leq H$ is met. The objective function is defined in Eq. (3.1a), together with the constraints of the optimization problem.

$$\begin{aligned} \underset{Q_i}{\text{maximize}} \quad & R(t_k) = \sum_{i=1}^{N-1} R_i(t_k) \end{aligned} \quad (3.1a)$$

$$\text{subject to:} \quad Q_i(t_k) \geq 0, i \in \{1, \dots, N-1\} \quad (3.1b)$$

$$T_i(t_k) > 0, i \in \{1, \dots, N-1\} \quad (3.1c)$$

$$R_i(t_k) \leq T_i(t_k), i \in \{1, \dots, N-1\} \quad (3.1d)$$

$$D_i(t_k) \leq H, i \in \{1, \dots, N-1\} \quad (3.1e)$$

3.3 PQM Mechanism

The PQM mechanism is presented in this section, including its rationale and a theoretical analysis for a representative scenario.

3.3.1 Rationale

PQM benefits from the centralized view of the FMN available in the Central Station. In order to present its rationale, for the sake of simplicity we will omit t_k thereafter. Considering the future positions of FAP_{*i*} and the bitrate of the traffic flow $F_{i,0}$, we aim at ensuring that the queue of each

FAP has, at each time instant, the minimum size Q_i that allows accommodating the traffic sent towards UAV₀. Conceptually, if Q_i is minimal, but enough to accommodate the offered traffic, then the packets are held in the queue for shorter time, the queuing delay decreases, the throughput is not compromised, and the shared medium usage is enhanced, thus improving the overall FMN performance. We assume knowing beforehand the traffic demand of each FAP; this information is provided by a state of the art FAP placement algorithm in charge of positioning the FAPs according to the traffic demand of the ground devices [4]. For the sake of simplicity, but without loss of generality, herein we exemplify the application of PQM using the M/D/1 queuing model, according to the Kendall's notation [121]. The data packets generated by the ground devices being served by the FAPs have mutually independent inter-arrival times; nevertheless, PQM is valid for any traffic model. For that purpose, the traffic consists of packets with constant size S and their arrival is modeled by a Poisson process. The packet inter-arrival times for each FAP are exponentially distributed according to λ_i packet/s, for a constant service time defined by $\mu_i = \frac{C_{i,0}}{S}$, while the traffic load of each FAP is expressed by $\rho_i = \frac{\lambda_i}{\mu_i}$. The number of packets in each FAP's queue is computed using Eq. (3.2) and the average packet delay can be computed using Eq. (3.3) [121], which includes the waiting time in the queue plus the service time.

$$Q_i = \frac{1}{2} \times \frac{\rho_i^2}{1 - \rho_i} \quad (3.2)$$

$$D_i = \frac{2 - \rho_i}{2 \times \mu_i \times (1 - \rho_i)} \quad (3.3)$$

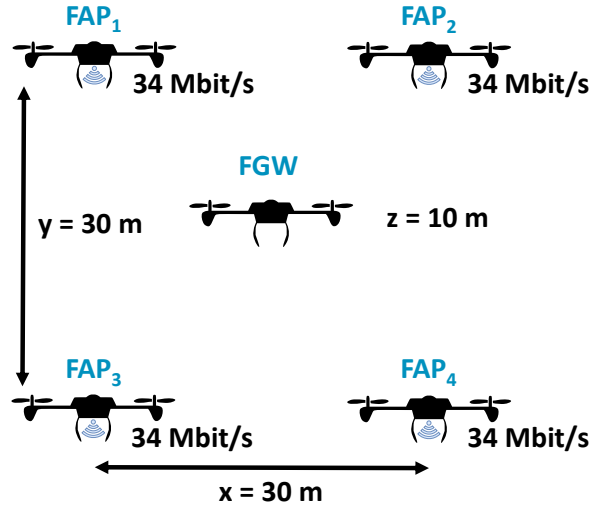
3.3.2 Theoretical Analysis for a Representative Scenario

In order to exemplify the application of PQM, let us perform a theoretical analysis for the representative scenario depicted in Fig. 3.2a. It consists of 4 FAPs placed in a square of side 30 m, hovering at an altitude of 10 m. For the sake of simplicity, the FAPs are assumed to have the same traffic demand; for this reason, the FGW is placed in the FAPs center (i.e., $(x, y, z) = (15 \text{ m}, 15 \text{ m}, 10 \text{ m})$), in order to enable links with the same capacity, which is according to the rationale behind the algorithm we have proposed in [74]. We assume the IEEE 802.11ac technology, one spatial stream, Guard Interval of 800 ns, and channel bandwidth of 160 MHz.

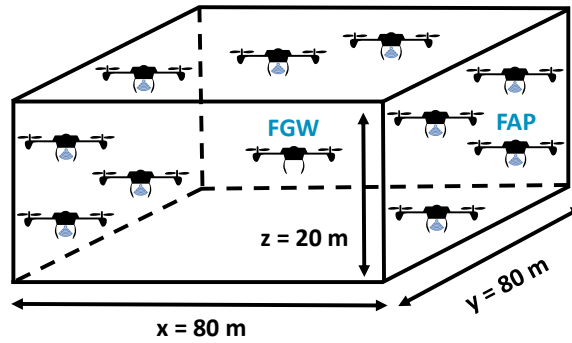
For the scenario of Fig. 3.2a, the Euclidean distance between each FAP and the FGW is $\approx 21 \text{ m}$. According to Eq. (3.4), assuming the FSPL model, this corresponds to $\approx -73 \text{ dBm}$ for the Received Signal Strength Indicator (RSSI), considering a transmission power $P_T = 0 \text{ dBm}$. Taking into account the relation between RSSI and the IEEE 802.11ac MCS indexes, we conclude that an RSSI of -73 dBm enables 175.5 Mbit/s as the maximum theoretical data rate for the flow between each FAP and the FGW, as presented in Table 3.1. Since the transmission time is shared by 4 FAPs generating traffic, each one hop away from the FGW, this results in a fair share $L =$

$0.85 \times \frac{175.5}{4} \approx 37$ Mbit/s for the capacity of the wireless channel between each $FAP_i, i \in \{1, 2, 3, 4\}$, and the FGW; we assume $\eta = 0.85$ as the efficiency of the MAC scheme, including the overhead of the transport protocol, Internet Protocol (IP), and MAC packet headers. The relation between RSSI and the IEEE 802.11ac MCS indexes, and the factor η were defined by running independent Network Simulator 3 (ns-3) simulations [74].

$$\frac{P_{R_{i,0}}}{P_T} = \left(\frac{c}{4\pi \times d_i^{MAX} \times f_i} \right)^2 \quad (3.4)$$



(a) Simple scenario consisting of 4 static FAPs generating traffic towards the FGW. The traffic demand is close to 90% of the fair share of the wireless channel capacity.



(b) Complex scenario consisting of 10 FAPs in movement generating traffic towards the FGW. The traffic demand of each FAP is equal to 18 Mbit/s, enabling $\rho < 1$ at any time instant.

Figure 3.2: Simulation scenarios used in the evaluation of the PQM mechanism.

Table 3.1: Relation between the minimum RSSI, data rate of the targeted MCS indexes, and the link capacity values for 4 FAPs sharing the transmission time, while using the IEEE 802.11ac standard; they were obtained through ns-3 simulations.

Minimum RSSI (dBm)	MCS data rate (Mbit/s)	Link capacity (Mbit/s)
-77	117	$117/4 \approx 29$
(-75)	175.5	$175.5/4 \approx 44$
-71	234	$234/4 \approx 58.5$

In this theoretical analysis, we assume a constant packet size of 1400 bytes, which is close to the Maximum Transmission Unit of 1500 bytes typical in data packets composing the Internet traffic [122], and an average bitrate equal to $90\% \times L \approx 34$ Mbit/s, at the onset of channel saturation. This results in $\lambda_i = \frac{34 \text{ Mbit/s}}{1400 \times 8 \text{ bits/packet}} \approx 3036 \text{ packets/s}$ and $\mu_i = \frac{37 \text{ Mbit/s}}{1400 \times 8 \text{ bits/packet}} \approx 3304 \text{ packets/s}$, leading to $\rho_i \approx 0.9$ and $Q_i = 5$ packets, $i \in \{1, 2, 3, 4\}$, according to Eq. (3.2). We argue that this queue size, which in fact represents the average number of packets in the queue, allows packets to be held in the queue for shorter time, thus decreasing the queuing delay and maximizing the throughput, when compared to the baseline that does not employ any queue management mechanism.

3.4 Performance Evaluation

The FMN performance achieved when employing PQM is presented in this section, including the simulation setup, the simulation scenarios, and the performance metrics evaluated.

3.4.1 Simulation Setup

ns-3 was used to evaluate the FMN performance achieved with PQM; for that purpose, a *DropTailQueue* with size scheduled in advance was implemented. On each UAV, a Network Interface Card (NIC) was configured in Ad Hoc mode, using the IEEE 802.11ac technology in channel 50, which enables channel bandwidth of 160 MHz and Guard Interval of 800 ns. A single spatial stream was used for the links between the FAPs and the FGW. User Datagram Protocol (UDP) Poisson and TCP *BulkSend* traffic, with packet size $S = 1400$ bytes, was generated during 130 s of simulation time, for each scenario. The data rate was defined by the *IdealWifiManager* mechanism. The main ns-3 simulation considered in the evaluation of PQM are presented in Table 3.2.

Table 3.2: Summary of the main ns-3 simulation parameters used in the evaluation of the PQM mechanism.

Parameter	Value
Simulation time	30 s (bootstrap) + 130 s
Wi-Fi standard	IEEE 802.11ac
Wi-Fi mode	Ad Hoc
Wi-Fi channel	50 (5250 MHz)
Channel bandwidth	160 MHz
Transmission (TX) power	Up to 10 dBm
Guard Interval	800 ns
Number of antennas	1
Propagation delay	Constant speed
Propagation model	FSPL (Friis)
Remote station manager	<i>IdealWifiManager</i>
Mobility model	Waypoint
Traffic type	UDP Poisson
Packet size	1400 Bytes

3.4.2 Simulation Scenarios

The performance of PQM was evaluated in two scenarios. The first scenario, which was described in Section 3.3.2, is depicted in Fig. 3.2a. It aims at comparing the FMN performance achieved when employing PQM against the baseline, which does not employ any AQM algorithm – Drop-Tail with the default static queue size of 1000 packets [123]. In addition, a sensitivity analysis on Q_i was performed, in order to verify whether the queue size obtained using Eq. (3.2) is the optimal value.

In order to evaluate the FMN performance achieved in typical scenarios observed in the real-world, 8 random complex networking scenarios were also considered. In these scenarios, 10 FAPs were moving according to the Random Waypoint Mobility model with a speed uniformly distributed between 0.5 m/s and 3 m/s inside the venue of dimensions 80 m $\times$ 80 m $\times$ 20 m that is depicted in Fig. 3.2b. At each time instant, the FGW was placed in the center of the FAPs (i.e., three-coordinates average considering all FAPs). Since PQM relies on knowing in advance the positions of the FAPs, provided by a FAP placement algorithm [4] in a real-world deployment, we used *BonnMotion* [107], a mobility scenario generation tool. The resulting waypoints were considered to compute in advance the FGW position [74] and the queue size for each FAP. Finally, the FGW position, the queue size, and the generated scenarios for the whole simulation time were imported to ns-3, with a sampling period of 1 s. The *WaypointMobilityModel* module was used in

ns-3 to place the FAPs in the positions generated by *BonnMotion* and the FGW in the position that was computed. The transmission power of the NICs was set to 10 dBm, which is high enough to enable single-hop communications between the FAPs and the FGW, so that we can focus on the queue management of the source nodes only.

3.4.3 Performance Metrics

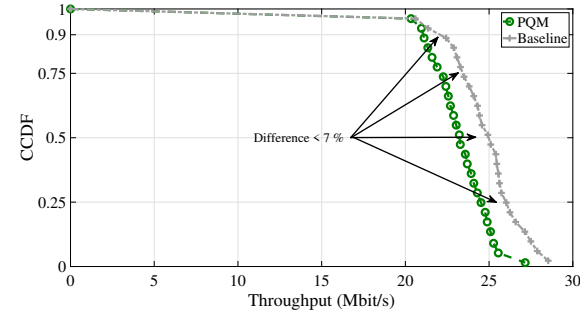
PQM was evaluated against five state of the art AQM algorithms: 1) RED, which drops packets based on the queue size; 2) BLUE and 3) PI, which rely on dropping probabilities; 4) PIE and 5) CoDel, which take dropping decisions based on the time that packets are held in the queue. In order to enable a fair comparison, CoDel was configured to use the same queue size as its state of the art counterparts: 25 packets. The performance evaluation considers two performance metrics:

- **Aggregate throughput:** the average number of bits received per second by UAV₀ (FGW).
- **Delay:** the average time taken by the packets to reach the sink application of UAV₀ since the instant they were generated by the source application of each FAP, including queuing, transmission, and propagation delays.

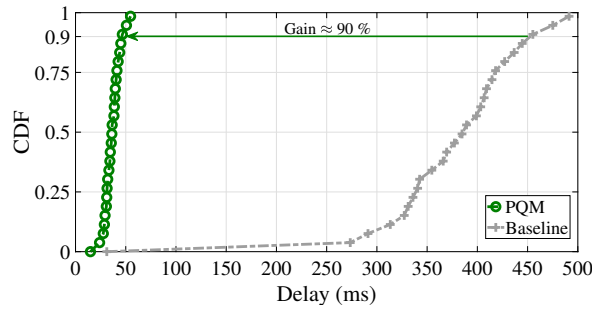
3.4.4 Simulation Results

The simulation results were obtained after 20 runs for each scenario defined in Section 3.4.2, under the same exact networking conditions. The results are expressed as average values, and they are represented using the Cumulative Distribution Function (CDF) for the delay and by the Complementary Cumulative Distribution Function (CCDF) for the aggregate throughput. The CCDF $F'(x)$ represents the percentage of time for which the aggregate throughput was higher than x , while the CDF $F(x)$ represents the percentage of time for which the delay was lower than or equal to x .

Regarding the simple scenario, we focus our analysis on the achieved throughput and delay with respect to a single FAPs – FAP₁ – measured in the FGW. When the queue size is set to 5 packets by PQM, which was obtained following the theoretical analysis presented in Section 3.3.2 for traffic equal to 90% of the fair share of the wireless channel capacity, the delay is reduced up to 91% for approximately the same throughput (Fig. 3.3), considering the 25th, 50th, 75th, and 90th percentiles and the Drop-Tail queue management mechanism using its default configuration as the baseline [123]. PQM sets the queue size to the average number of packets expected to be in the queue at each time instant $t_k = k \cdot \Delta t, k \in \mathbb{N}_0$, where $\Delta t \gg 1s$ is defined considering the FMN update period imposed by the FAP placement algorithm, and assuming an M/D/1 queuing model. This ensures the shortest possible queue size, thus avoiding that packets are held in the



(a) Throughput CCDF.



(b) Delay CDF.

Figure 3.3: Throughput and delay results for the UDP flow generated by FAP_1 , measured in the FGW, considering the simple scenario. PQM (green circles) is the proposed mechanism, whereas the baseline corresponds to Drop-Tail with the default static queue size of 1000 packets defined in ns-3.

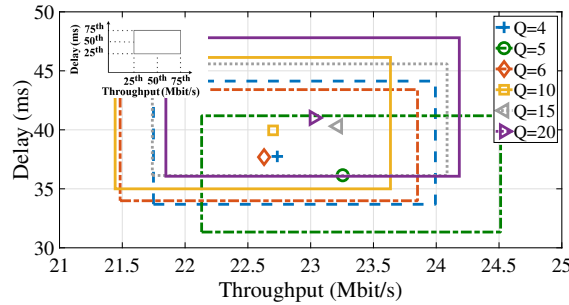
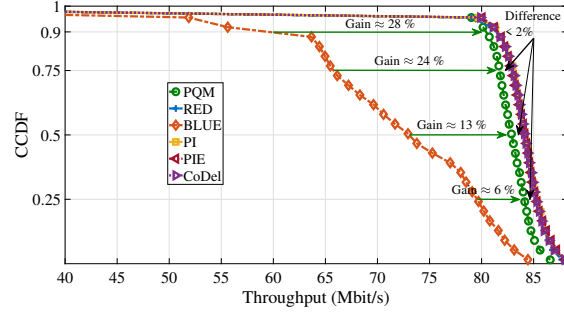


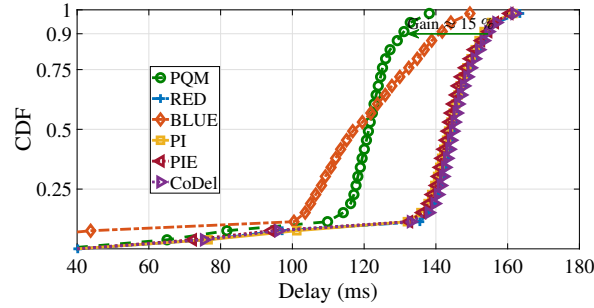
Figure 3.4: Analysis on the optimal queue size (Q) for each FAP composing the simple scenario. $Q = 5$ enables the minimum delay and the maximum throughput; the results are expressed by means of the 25th, 50th, and 75th CDF percentiles, considering UDP traffic.

transmission queue for long time, which significantly increases the delay of all packets without significant gains in the amount of bits received in the FGW. This can be concluded from the difference between the throughput for PQM and Drop-Tail in Fig. 3.3a. From the sensitivity analysis on the optimal queue size, whose simulation results are presented in Fig. 3.4 for values between

4 and 20 packets, it can be concluded that a queue size of 5 packets is in fact optimal, allowing to achieve the maximum throughput with minimum delay.



(a) Throughput CCDF.



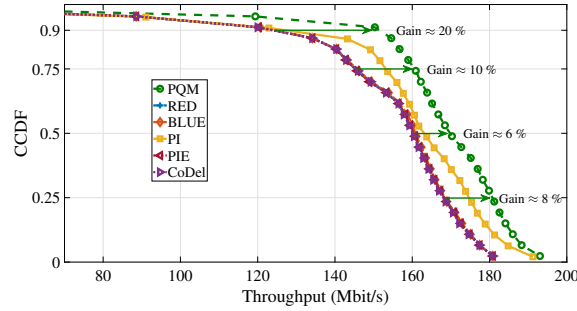
(b) Delay CDF.

Figure 3.5: Aggregate throughput and delay results for UDP traffic, measured in the FGW, considering the complex scenario.

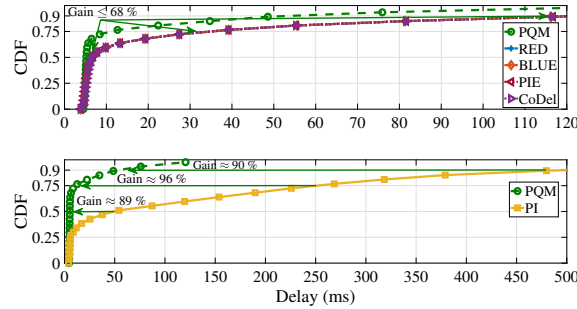
The results for the complex scenario, where 8 random dynamic topologies during 130 s simulation time were considered, are depicted in Fig. 3.5 and Fig. 3.6, for UDP and TCP traffic, respectively. The dynamic topologies motivate queues with different sizes over time, due to variable links capacity that induce changes in μ . We focus our analysis on the results for the 25th, 50th, 75th, and 90th percentiles of both aggregate throughput and delay. PQM improves the aggregate throughput up to 28% for UDP and 20% for TCP, while the delay is decreased up to 17% and 96% for UDP and TCP, respectively. Overall, PQM enables delay values lower than its counterparts (cf. 75th and 90th percentiles in Fig. 3.5b), for approximately the same throughput (the difference is lower than 2% for UDP traffic) in the worst-case scenario, as depicted in Fig. 3.5a. Hence, PQM is the solution that enables the highest throughput for the minimum delay. These results validate PQM and corroborate our research hypothesis: the FMN performance is improved by defining in advance the FAPs' queue size over time, considering both 1) the positions of the FAPs, and 2) the offered traffic, which are provided by the FAP placement algorithm.

3.5 Discussion

Even though the wireless link established between the FAPs and FGWs may have sufficient network capacity to accommodate the average traffic demand of the ground devices served, in real-world communications nodes estimating the instantaneous traffic demand is challenging, including the maximum values and duration of traffic bursts. If the instantaneous traffic demand at a given FAP exceeds the available network capacity in the wireless link established with the FGW, the FAP holds packets in its queue. By defining the minimum queue size for each FAP, the PQM mechanism allows for controlled delays.



(a) Throughput CCDF.



(b) Delay CDF.

Figure 3.6: Aggregate throughput and delay results for TCP traffic, measured in the FGW, considering the complex scenario.

$$\text{PLR}_i = p_{b_i} = P[Q_i] = \frac{1 - \rho_i}{1 - \rho_i^{Q_i+1}} \times \rho_i^{Q_i} \quad (3.5)$$

The sensitivity analysis made in the performance evaluation of PQM (cf. Fig. 3.4) proved that the queue size Q_i obtained by means of Eq. (3.2) allows for maximum aggregate throughput with minimum average delay, when compared with other possible queue size values. Q_i , which represents the average number of packets in the queue of FAP_{*i*}, considers the M/D/1 queuing

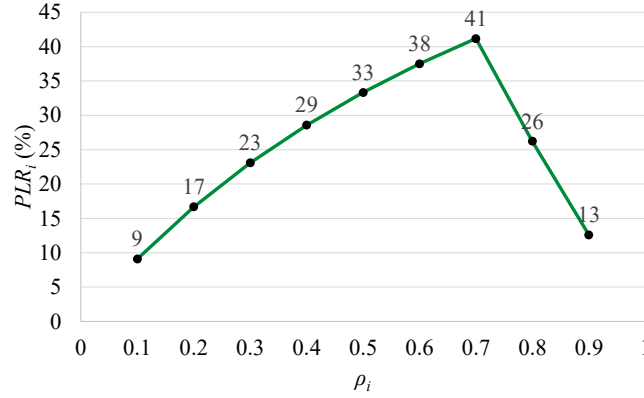


Figure 3.7: Theoretical Packet Loss Ratio (PLR) (PLR_i) versus traffic load ρ_i .

model and depends directly on ρ_i . The PLR (PLR_i) of FAP_i for different values of ρ_i , when employing Q_i as the queue size, can be computed using Eq. (3.5). Although the queue configured in FAP_i is represented by the M/D/1/ Q_i queuing model, for calculating PLR_i we take into account the M/M/1/ Q_i queuing model, since the blocking probability p_{b_i} does not depend on the packet service time distribution [121]. We conclude that PLR_i is up to 41% for $\rho_i = 0.7$, as depicted in Fig. 3.7. As the traffic load ρ_i associated with FAP_i increases, the average number of packets in its queue increases exponentially. For ρ_i between 0.1 and 0.7 the average number of packets Q_i is 1, whereas for ρ_i equal to 0.8 and 0.9 Q_i is respectively equal to 2 and 4. This increase in the number of packets that the queue of FAP_i can accommodate justifies the decrease of the PLR_i depicted in Fig. 3.7.

Despite Q_i enables maximum aggregate throughput with minimum average delay, the queue size of each FAP_i should be fine-tuned when low PLR_i is considered for Service Level Agreements (SLAs), such as to enable ultra-reliable communications. Despite ensuring ultra-reliable communications is beyond the scope of this thesis, in order to reduce PLR_i , Q_i should be increased, at the expense of a longer average delay D_i . On the other hand, since $\rho_i = \lambda_i / \mu_i = \lambda_i / (C_{i,0} / (S \times 8))$, if the capacity $C_{i,0}$ is increased, then ρ_i will be decreased, thus contributing to reduce PLR_i , as depicted in Fig. 3.7. This can be achieved by defining a suitable FGW placement and the transmission power P_T , in order to ensure low delay and low PLR values simultaneously. Therefore, the variables P_T and Q_i are worthy of being defined to improve the overall FMN performance. A solution to overcome this challenge is presented in Chapter 4. When multiple traffic flows have different QoS requirements, multiple queues with different sizes can be used on FAP_i , one for each flow, following the concept proposed in [124]. In this sense, queues with larger size than Q_i can be configured for flows that require low PLR_i , while queues with smaller size than Q_i can be configured to achieve maximum aggregate throughput with minimum average delay for other flows.

3.6 Summary

The ability to ensure low-latency communications is a challenge in the design of emerging wireless networks. The ultimate goal consists in reliably exchanging data with controlled delays so that communications between billions of interconnected machines are possible and humans can interact with remote devices in real-time. Wireless networks in general and flying networks in particular, due to their dynamic nature, bring up significant challenges, including frequent changes in the SNR of the wireless links, which induce time-varying network capacity. For these reasons, queues with static size and the state of the art AQM algorithms, which are designed for wired networks, are not adequate for FMNs. This chapter presented PQM, a proactive queue management mechanism for FMNs with controlled topology. Building upon the conclusion that the usage of static size network queues in wireless networks leads to unnecessary high delays, PQM takes advantage of the information provided by a state of the art FAP placement algorithm, namely the future positions of FAPs serving ground devices and their traffic demand, to define in advance the queue size of the FAPs over time. The FMN performance when PQM is employed was evaluated by means of ns-3 simulations. The achieved results show gains up to 28% in aggregate throughput, while the average delay is reduced up to 96%.

PQM configures only the queue size of the communications nodes, representing a step forward with respect to the existing solutions, which typically require fine-tuning several parameters. Still, it is unable to find a suitable solution that meets the average delay constraint for all flows offered by the FAPs if the capacity of the wireless links is not high enough. A solution able to ensure wireless links that meet the traffic demand of the FAPs, while benefiting from the network capacity information for computing the queue size so that strict delay requirements are met, is worth considering. This is the subject of the next chapter.

Chapter 4

An Algorithm for Placement and Queue Management in Flying Multi-hop Networks

In Chapter 2, we proposed the GateWay Placement (GWP) algorithm to address the Flying Gateway (FGW) placement problem. In Chapter 3, we proposed the Proactive Queue Management (PQM) mechanism to meet strict low-latency requirements in Flying Multi-hop Networks (FMNs). GWP is focused on the FGW placement problem only and PQM is limited to the proactive configuration of the queue size without considering the network capacity information for computing the queue size. Besides the ability to accommodate the offered traffic, FMNs are expected to meet strict low-latency requirements. In this chapter, we jointly address the FGW placement and queue size management problem. For that purpose, we propose the traffic-aware Gateway Placement and Queue Management (GPQM) algorithm.

Taking advantage of the information on the future positions of the Flying Access Points (FAPs), defined by a state of the art placement algorithm to meet the traffic demand of the ground devices served, the proposed GPQM algorithm allows 1) determining the FGW position and 2) calculating the queue size for each FAP. The FGW placement and queue size management at periodic time instants enables minimizing the packet service time and the traffic load in each FAP, while accommodating the offered traffic with stochastic packet delay guarantees.

We start this chapter by presenting the related work on joint Unmanned Aerial Vehicle (UAV) placement and resource management approaches proposed in the literature. Then, we propose a system model and formulate the problem. Afterwards, we describe the GPQM algorithm and refer to its performance evaluation, including the methodology, the simulation setup, and the obtained results. Before ending the chapter, we discuss the benefits and drawbacks of the GPQM algorithm. The chapter ends with a summary of the contents presented.

4.1 Related Work

Low-latency communications have been considered one of the pillars in the recently released 5G networks and the emerging 6G networks [125]. In order to enable low-latency communications, placing UAVs close to the devices to be served, while enabling short-distance Line-of-Sight wireless links, represents an advantage over ground wireless networks with limited mobility. Yet, the joint UAV placement and queue size management is a challenge. GPQM is the first algorithm addressing the joint FGW placement and queue size management in FMNs, as demonstrated by the literature review on joint UAV placement and resource management presented below. In what follows, we start by presenting state of the art solutions focused on delay-aware UAV positioning; then, we refer to solutions addressing the joint UAV placement and resource management enabling delay-aware communications.

In [11], a latency-aware UAV placement algorithm has been proposed. Considering a network consisting of a terrestrial Base Station (BS) and a UAV also acting as a flying BS, connected to each other by means of Free-Space Optics-based backhaul link, as depicted in Fig. 4.1, the algorithm aims at determining the UAV 3-Dimensional (3D) position, user association, and bandwidth allocation between the UAV and terrestrial BS. The objective is to minimize the ground devices' average latency and meet their Quality of Service (QoS) requirements; however, the queue management problem has not been addressed.

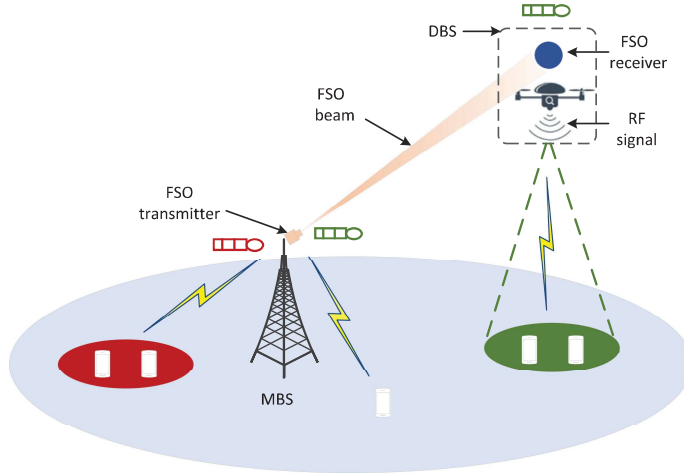


Figure 4.1: UAV acting as a BS connected to a terrestrial BS by means of a Free-Space Optics(FSO) based wireless link [11].

A dynamic flying BS placement approach has been proposed in [12] to minimize the queuing delay in a network consisting of a UAV serving ground devices, as depicted in Fig. 4.2. In

response to dynamic wireless traffic, the flying BS placement is dynamically defined for adapting the transmission data rate. The authors have evaluated three different scenarios regarding wireless traffic. The first considers that traffic can be predicted. The second considers that the flying BS is aware of the probability density function of the packet arrival process. Finally, the third considers that the number of arriving packets or the corresponding statistical process are unknown. The authors have designed a different dynamic flying BS placement approach for each scenario to minimize the average queuing delay, subject to realistic limits imposed by the battery life and velocity of the flying BS. They have concluded that, due to the dynamism imposed by variable wireless traffic, conventional trajectory design based on the sequential convex optimization technique is not able to solve the problem, while queue-aware UAV placement based on the Lyapunov optimization theory is incapable of finding the minimum-delay solution. As such, the queuing delay minimization problem was transformed into a Markov Decision Process-based model. However, the traffic demand and delay constraints for the ground nodes have not been considered. Overall, as stated in [120], designing a single optimal queue management algorithm that suits the variable characteristics of wireless networks for different congestion scenarios is not straightforward, especially due to the need of fine-tuning a wide set of parameters.

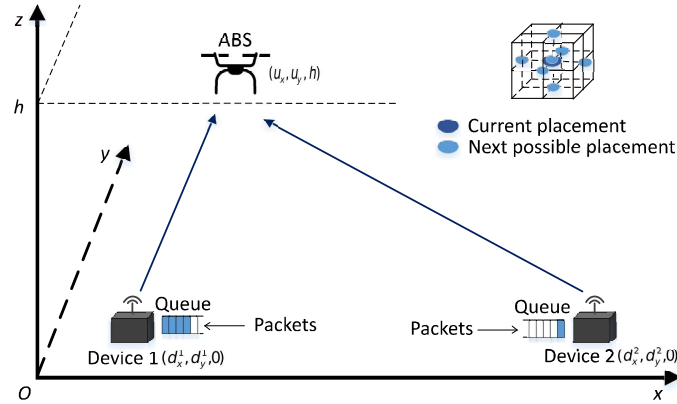


Figure 4.2: Illustration of the dynamic placement approach proposed in [12], where a UAV acting as a flying BS serves a set of ground devices, each with a network queue.

In [13], a hierarchical flying network has been employed. It is named Manned-and-Unmanned network because it employs both manned aircrafts in the upper layer and UAVs in the bottom layer, as depicted in Fig. 4.3. The manned aircrafts have advanced processing capabilities and are capable of communicating over long distances. They act as control nodes for running computationally intensive algorithms in real-time, including Deep Learning, while the UAVs can form different swarm topologies. Each control node communicates with a reduced number of pre-selected UAVs acting as gateway nodes. If a multi-swarm architecture is employed, the gateway nodes can serve multiple swarms. They typically have low mobility and better communications ca-

pabilities. Gateway UAVs can operate in multiple frequency bands; for example, a band to enable communications between the UAVs themselves and other for communications with the manned aircrafts. The authors have been focused on determining the optimal position for relay nodes, in order to meet the QoS requirements of UAV swarm applications. For that purpose, they have designed a queuing model based on packet priority levels, in order to meet the delay constraints of each packet over multi-hop paths. At each hop, information such as packet priority, multi-hop information, Time-To-Live, and Signal-to-Interference plus Noise Ratio are captured for any arriving or locally generated packet. If the node is in the tail of a path, it forwards the packet to the next hop. If the node is a relay, it determines the service time, Time-To-Live, and packet arrival rate. The queuing model has been employed to compute multi-hop delay and Packet Loss Ratio (PLR). The UAV relay positioning scheme allows optimizing the aggregate throughput.

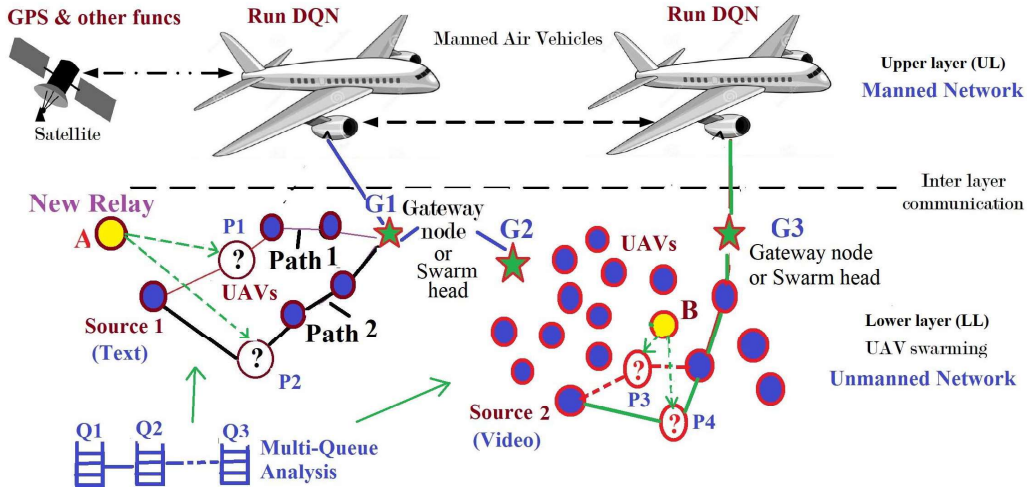


Figure 4.3: Hierarchical flying network architecture consisting of UAVs in the bottom layer and manned aircrafts in the upper layer. On the left-hand side, a traffic source seeks to send text data, while on the right-hand side a second source seeks to send video data. The placement of the UAVs acting as relay nodes in the bottom layer, in order to forward the traffic to the upper layer, is determined taking into account the QoS requirements and the wireless links conditions [13].

In [14], queuing in a UAV acting as a flying communications relay, as depicted in Fig. 4.4 has been modeled as a stochastic process using queuing theory, assuming that the transmission process follows the *Bernoulli* distribution with geometrically distributed successful transmission time. The authors have used the multiple or compound *Bernoulli* process for both the traffic arrival and packet transmission processes, which allow representing typical bursty traffic and a Line-of-Sight based channel model, respectively. The analytic model proposed is similar to a system with two consecutive back-to-back queuing models associated with the transmission processes of the source-relay and relay-destination links. It allows calculating the average end-to-end packet

delay and packet drop probability due to buffer overflow under bursty traffic. The accuracy of the proposed model has been validated through numerical simulations, confirming that a buffer-aided mobile relaying system outperforms static relaying in terms of throughput and end-to-end average packet delay. However, the conclusions are limited to the networking scenario depicted in Fig. 4.4, which consists of a single UAV acting as a flying communications relay and two ground devices.

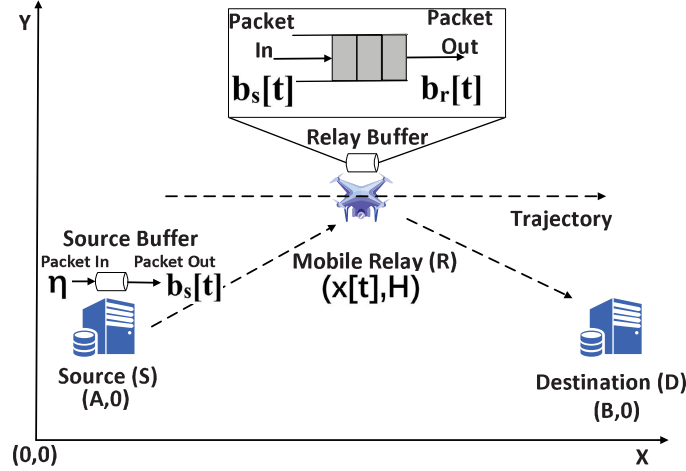


Figure 4.4: Representative networking scenario consisting of a single UAV performing the role of flying communications relay between two ground devices acting as, respectively, source and destination of the data traffic. The authors have concluded that a flying communications relay following a given trajectory outperforms the network performance achieved by a static UAV [14].

An algorithm for UAV-enabled BS placement has been proposed in [15] to provide QoS in terms of delay and coverage probability. The path-loss model for backhaul and fronthaul wireless links has been considered. In order to minimize the delay, while ensuring coverage and meeting the backhaul and fronthaul traffic load, an optimization problem has been formulated to determine the optimal UAV's altitude, as shown in Fig. 4.5. Despite the formulated optimization problem is non-convex, the objective function has been demonstrated to be unimodal (i.e., there is only a maximum value) with respect to UAV's height. To determine the optimal UAV's height with logarithmic time complexity, the Backhaul and Delay aware Positioning of Unmanned Aerial Vehicle-enabled Base Station algorithm has been proposed.

Finally, in [16], a flying network, consisting of a UAV acting as a BS sending data towards a ground User Equipment, has been considered. The reference networking scenario considered by the authors is depicted in Fig. 4.6. The delay components of cellular networks and how they contribute to end-to-end delay have been explained, including the queuing delay. Theoretical analysis has demonstrated the advantages of using UAVs for reducing the delay, especially due to the increased Line-of-Sight achieved. Moreover, the authors have concluded that optimizing

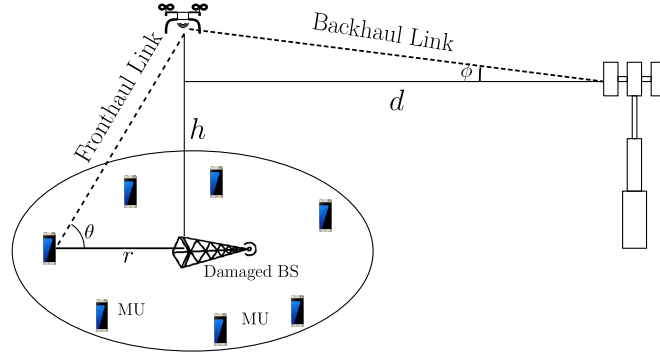


Figure 4.5: Networking scenario addressed by the algorithm proposed in [15]. In order to reestablish a Radio Access Network (RAN) due to a damaged or maintenance ground BS, a UAV-enabled BS is deployed on-demand at optimal altitude, while serving a set of ground Mobile Users (MUs).

the UAV's altitude considering the location of the User Equipment allows reducing the delay.

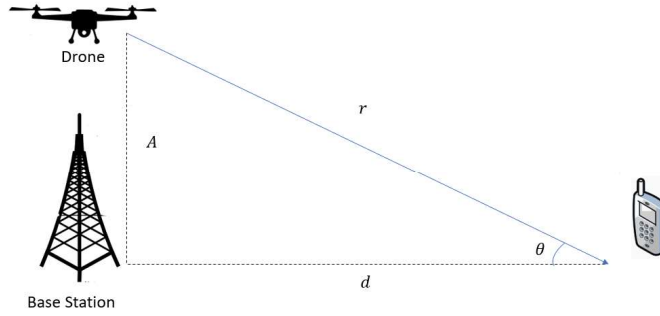


Figure 4.6: Reference networking scenario considered in [16], which consists of a UAV acting as a BS transmitting data towards a ground User Equipment. The authors have demonstrated the advantages of using a flying BS for reducing the delay, especially due to the enhanced Line-of-Sight achieved; this is exacerbated when the UAV's altitude is optimized considering the User Equipment's location.

In [17], a resource allocation mechanism has been proposed to minimize the packet transmission delay in a 3D cell network consisting of UAVs organized into a multi-layer architecture. First, the authors have designed a hierarchical UAV network using a *Voronoi* diagram. A *Voronoi* diagram results from partitioning a plane consisting of multiple points into convex polygons, each containing one generating point, where every point in a given polygon is closer to its generating point than to any other. Considering the networking scenario shown in Fig. 4.7, the authors have assumed that in a *Poisson-Voronoi* based system, the expected number of UAVs covered by a macro BS can be calculated considering its coverage radius. In addition, the authors have assumed that the expected number of User Equipments covered by each UAV, as well as the length

of the corresponding wireless links and the number of UAV layers, can be determined by considering the minimum transmission range required for co-channel interference avoidance, which is assumed to represent the distance between adjacent UAV layers. A non-Line-of-Sight wireless channel associated with a Rayleigh fading model for signal and interference was assumed, which leads to a random channel power gain modeled as a unity mean exponential distribution. The average packet transmission delay has been modeled by means of an M/G/1¹ queue per UAV. Finally, the packet transmission delay minimization problem has been formulated in a two-layer UAV network and solved by gradient descent using the bisection method.

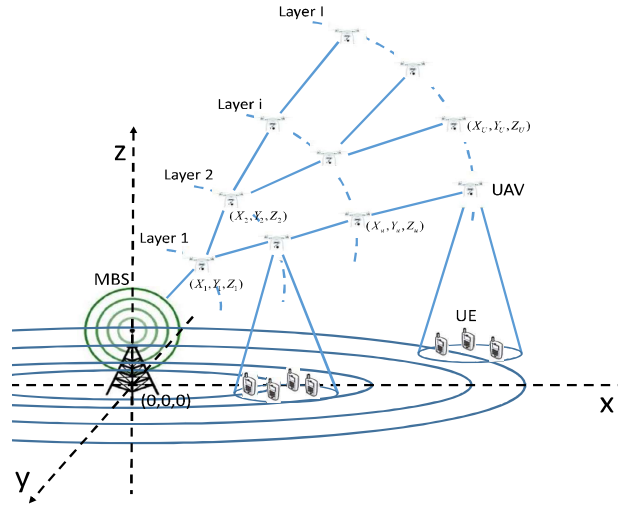


Figure 4.7: Hierarchical flying network architecture consisting of UAVs. Due to the limited wireless communications range, multiple flying communications relays and a single FGW are used to forward traffic between a Macrocell BS and the ground devices acting as User Equipment [17].

In [18], the queuing delay of an on-demand Software-Defined Networking (SDN) based UAV-mounted BS coexisting with a Wi-Fi Access Point (AP) has been investigated, as depicted in Fig. 4.8. For modeling the delay of cellular and Wi-Fi users, an M/M/1 queuing model per user has been considered. An SDN controller is in charge of globally minimizing the average queuing delay of cellular users while maintaining a maximum controlled queuing delay for the Wi-Fi users. For that purpose, it optimizes spectrum allocation, UAV's position, association of cellular users to the AP, and traffic offloading rate from the cellular users to both the UAV-mounted BS and the AP. By positioning the UAV, the authors have concluded that there is a significant difference between the objectives of sum-rate maximization and average delay minimization. The optimization problem

¹ In Kendall's notation, the M/G/1 queuing model assumes that the packet arrivals are modeled by a Poisson process and the service times have a General distribution, while considering a single server. M/G/1 is an extension of the M/M/1 queuing model, where the service times are exponentially distributed [121].

has been formulated as a non-convex mixed-integer problem; due to its complexity, a heuristic-based approach lying on block coordinate descent and Sequential Convex Approximation has been proposed.

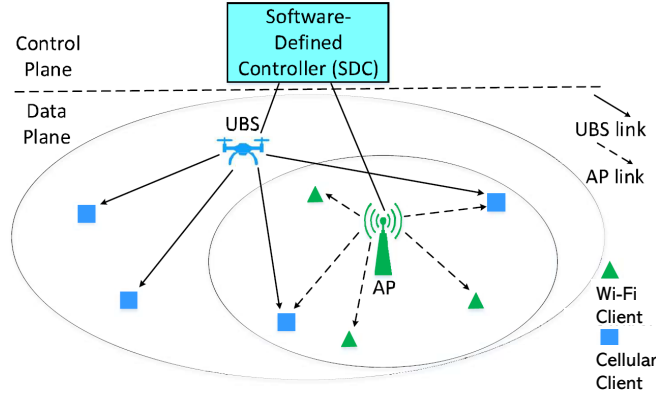


Figure 4.8: SDN scenario, consisting of a UAV acting as a flying BS coexisting with a ground Wi-Fi AP, considered in [18]. The authors have minimized the average M/M/1 queuing delay of the cellular clients, while guaranteeing the delay performance for the Wi-Fi clients, by optimizing the spectrum allocation, the UAV position, the association of cellular clients to the ground AP, and the cellular clients' traffic offloading.

A joint UAV placement and queue scheduling algorithm for admission and congestion control has been proposed in [126]. The problem involves deploying multiple UAVs to serve a set of ground devices. The problem has been subdivided into two distinct components. The first component is the UAV placement; it is accomplished by computing a scheduling cost for each area to be served, which allows identifying areas with higher traffic demand. The second component involves addressing admission and congestion control and QoS provisioning. The solution allows for shorter delays and a higher probability of guaranteed QoS. Yet, it is focused on multiple UAVs that act as flying communications relays for ground nodes, and the queue scheduling lies on assigning different priorities to the network packets; the queue size management has not been addressed. In [127], a dynamic trajectory control algorithm for fixed-wing UAVs providing wireless connectivity to ground nodes has been proposed. It defines the trajectories of the UAVs so that the queue occupancy of each UAV is lower than a given threshold, in order to increase the Packet Delivery Ratio (PDR) for the wireless links established between the UAVs. However, the trajectories of the UAVs are defined to ensure radio coverage for the ground nodes and the QoS requirements have not been taken into account. In addition, since the trajectories of the UAVs are readjusted only when link congestion occurs, it constitutes a reactive approach.

The joint communications schedule and velocity control of multiple UAVs has been formulated as a Multi-agent Markov Decision Process in [19]. The goal is to reduce packet loss caused by queue overflowing and communications failures. The Multi-agent Markov Decision Process

network state includes ground sensor battery levels and data queue lengths, channel conditions, visit time, and waypoints along the UAVs' trajectory. The UAVs are in charge of scheduling ground sensors' data transmissions, determining modulation schemes, and changing the patrol velocity. To train a UAV based on the actions of the remaining UAVs, a local action recording mechanism, in which each ground sensor stores all UAVs' historical visits, is devised. The records, which comprise the previous scheduling decisions of the other UAVs, are received by the UAV that schedules the ground sensor to transmit data. For that purpose, a Multi-Unmanned Aerial Vehicle Deep Reinforcement Learning Scheduling Algorithm (MADRL-SA) has been proposed so that UAVs are able to minimize the system's packet loss. MADRL-SA enables that each UAV learns to coordinate its actions. In Fig. 4.9, MADRL-SA considers three UAVs, each of which running a Deep Q-Networks algorithm, while learning through interaction with the environment. For example, UAV₃ schedules a ground sensor, receives its visiting time, and estimates the difference between its visiting times. Based on the time difference calculated, the UAVs that visited the same ground sensor to improve their scheduling process. If the calculated time difference is high, the UAV is induced to schedule the ground sensor for the next time interval, in order to prevent the ground sensor's queue starts dropping packets. The ultimate goal is to promote that different UAVs are in charge of scheduling different ground sensors. On the other hand, if a ground sensor was recently visited by a UAV, MADRL-SA aims at inducing that no other UAV will visit that sensor in the next time interval.

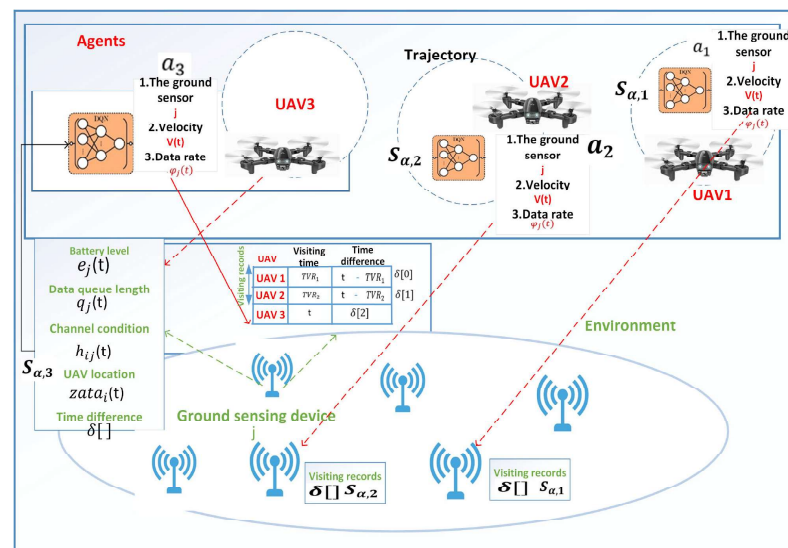


Figure 4.9: Overview of the MADRL-SA, considering a networking scenario consisting of three UAVs learning by observing the environment [19].

In [20], the authors have proposed an algorithm for interference control in Integrated Access

and Backhaul (IAB) based flying networks; IAB has been initially defined in 3rd Generation Partnership Project (3GPP) Release 16 for establishing backhaul networks using spectral resources made available by fixed BSs [128], as depicted in Fig. 4.10. The algorithm jointly optimizes the association between ground devices and BSs, power allocation, and 3D deployment of UAVs. Two spatial based deployment approaches for UAVs have been proposed: distributed UAVs and Drone Antenna Array, considering the geographical distribution of the ground devices. For the first approach, the UAV deployment in 3D was considered. For the second approach, the Drone Antenna Array design parameters have been defined in terms of antenna's array orientation, antenna's element spacing, and 3D position of the antenna array's center. The problem has been formulated to maximize the network data rate. Due to the interdependence of the optimization variables considered, the main problem has been divided into two sub-problems: 1) defining the association between ground devices and BSs as well as the power allocation for access and backhaul networks, considering a fixed spatial configuration for the UAVs; 2) determining the spatial configuration for the UAVs and power allocation considering fixed association between ground devices and BSs. The two sub-problems have been solved by means of heuristic-based algorithms. Through *Monte Carlo* simulation results, the authors have concluded that the algorithm allows improving up to 3 times the Signal-to-Interference plus Noise Ratio in the ground devices and 7 times the network data rate, when compared with a baseline deployment approach not using UAVs. Moreover, the authors have confirmed that sharing the same wireless channel for access and backhaul networks (in-band operation) allows improving both wireless coverage and network capacity. Finally, for the Drone Antenna Array-based approach, they have demonstrated that the network performance increases directly proportional to the number of antenna's elements, while the computational complexity is independent of the number of UAVs used. However, the algorithm does not allow ensuring QoS guarantees.

Based on the literature review presented herein, we conclude the state of the art works have been focused on maximizing coverage and aggregate network performance, including throughput maximization and delay minimization. Besides UAV placement, they have been optimizing different components, including resource allocation, path and trajectory planning, transmission power, channel assignment, and traffic offloading, using heuristic, metaheuristic, game theory, exhaustive search, Machine Learning (ML), and stochastic based approaches. However, they do not address the UAV 3D placement and queue size management simultaneously. Taking advantage of the deterministic knowledge about the future positions of the FAPs, GPQM is able to determine the FGW position and the queue size for each FAP. The ultimate goal is to ensure wireless links with high enough capacity to accommodate the traffic forwarded to/from the FAPs, while providing stochastic packet delay guarantees. The joint FGW placement and queue size management performed by the GPQM algorithm proposed by this thesis allows minimizing the packet service time and the traffic load in each FAP, thus improving the FMN performance. GPQM represents a

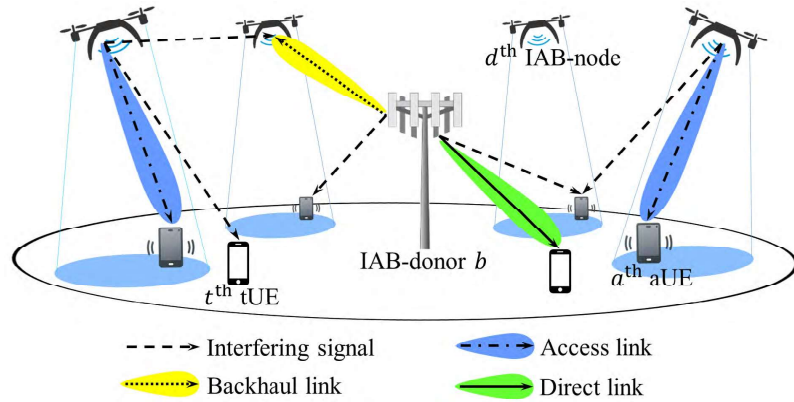


Figure 4.10: Flying network in which the UAVs act as IAB-nodes capable of forming access and backhaul networks; they share the spectral resources provided by a fixed BS acting as an IAB-donor [20].

step forward to the state of the art algorithms that rely on mobility prediction-based approaches or real-time measurements, which typically imply long convergence times to achieve a suitable solution.

4.2 System Model and Problem Formulation

The FMN is assumed to be organized into a two-tier architecture, as depicted in Fig. 4.11. This architecture is especially targeted at taking advantage of short range high-directional wireless links, providing high bandwidth wireless channels and low inter-flow interference. Two types of UAVs are assumed to compose the FMN: 1) the FAPs, which provide wireless connectivity to the ground devices, and 2) the FGW, which forwards the traffic to/from the Internet. The FAPs and the FGW establish an FMN able to 1) extend the radio coverage from a remote AP or BS and 2) provide on-demand wireless connectivity over a given area. The first tier consists of the access network, which is made up of FAPs enabling high-capacity small cells. The FAPs are dynamically placed and configured according to the traffic demand of the ground devices by means of a state of the art FAP placement algorithm. For that purpose, the FAPs periodically take a snapshot of the first tier, including the amount of traffic that the ground devices are offering to the network and send this information to a Central Station, where the FAP placement algorithm is running. The second tier consists of the FGW that is placed above the FAPs. The FGW forwards the traffic to/from the Internet by means of a dedicated broadband wireless link with QoS ensured by the core network, typically managed by a Mobile Network Operator. The traffic demand of each FAP must be higher than or equal to the sum of the traffic offered by the ground devices connected to it.

The Central Station can be deployed anywhere in the Edge or in the Cloud. The Central Station is responsible for controlling the FMN by periodically: 1) defining the positions of the FAPs according to the traffic offered by the ground devices; 2) determining the updated positions of the FGW, in order to enable wireless links able to meet the traffic demand of the FAPs; and 3) calculating the queue size for each FAP, in order to provide stochastic delay guarantees. The Central Station is in charge of sending the updated positions to the FAPs and the FGW, and the queue size to each FAP, at periodic time instants. Based on the information received, the FAPs and the FGW position and configure themselves accordingly. An out-of-band long-range wireless channel (e.g., based on IEEE 802.11ah) is used for exchanging information between the Central Station and the UAVs. It enables an always-on control link and avoids introducing overhead in the backhaul and access networks formed by the UAVs. The design of the control link is beyond the scope of this thesis.

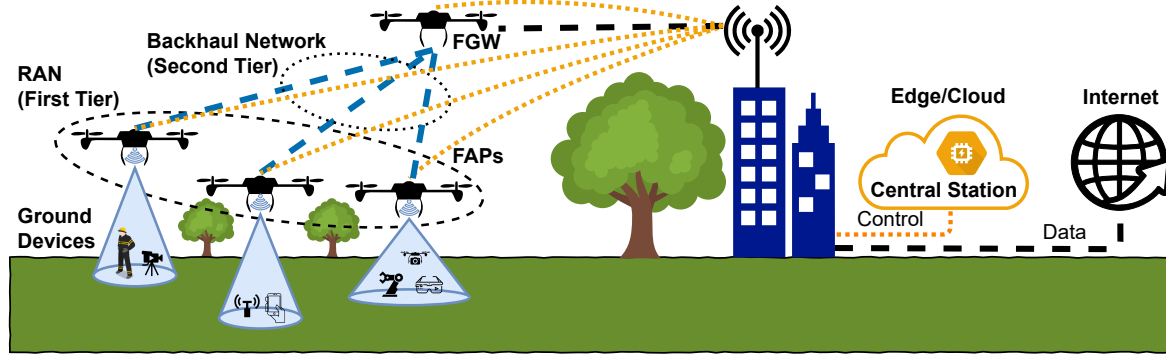


Figure 4.11: FMN organized into a two-tier architecture. The FAPs provide wireless connectivity to the ground devices and the FGW forwards the data traffic to/from the Internet (dashed blue line). The FMN is controlled by a Central Station deployed in the Edge or in the Cloud. The UAVs use an out-of-band long-range wireless channel (dotted orange line) to exchange information with the Central Station.

Let us assume that $\Delta t \gg 1s$ corresponds to the update period defined by the FAP placement algorithm. At the discrete time instant $t_k = k \cdot \Delta t, k \in \mathbb{N}_0$, the FMN is represented by a directed graph $G(t_k) = (V, E(t_k))$, where $V = \{0, \dots, N-1\}$ is the set of UAVs placed at $P_i = (x_i, y_i, z_i)$ inside a cuboid x^{MAX} long, y^{MAX} wide, and z^{MAX} high. $E(t_k) \subseteq V \times V$ is the set of directional wireless links available between UAV_i and UAV_j at time t_k , with $(i, j) \in E(t_k)$. The wireless channel between each pair of UAVs is modeled by the Free-Space Path Loss (FSPL) model [129], which is justified by the strong Line-of-Sight component that dominates the wireless links established between UAVs [21]. $C_{i,j}(t_k)$ defines the capacity, in bit/s, of the wireless link available between UAV_i and UAV_j at time t_k , considering a constant channel bandwidth B in Hz. We assume that the maximum link capacity is equal to the data rate associated to the Modulation and Coding Scheme (MCS) index MCS_i selected by each UAV_i over the number of UAVs that use the same wireless channel.

MCS_{*i*} requires a minimum SNR_{*i*}, considering a constant noise power P_N .

UAV_{*i*}, $i \in \{1, \dots, N-1\}$, which is placed one hop away from UAV₀, acts as a FAP and transmits, during time interval t_k , a traffic flow with bitrate $T_i(t_k)$ bit/s and constant packet size S , in bytes, towards UAV₀, which performs the role of FGW. The active topology of the FMN is defined by a tree $T(V, E_T)$ that is a sub-graph of G , where $E_T \subseteq E$ is the set of direct wireless links between UAV_{*i*} and UAV₀, which is the root of the tree. The flow $F_{i,0}$, with bitrate $T_i(t_k)$ bit/s, demands a minimum capacity $C_{i,0}(t_k)$, in bit/s, for the wireless link available between UAV_{*i*} and UAV₀, at time t_k . $F_{i,0}$ is received at UAV₀ from UAV_{*i*} with bitrate $R_i(t_k)$ bit/s and average delay $D_i(t_k)$ seconds, at time t_k . Herein, we consider average delay values for illustrative purposes, but a Service Level Agreement (SLA) established with a Mobile Network Operator can also refer maximum values (e.g., the longest packet delay among all packet delay values) or median values (e.g., 50th percentile of the frequency distribution of packet delays) [130]. The wireless channel is assumed to be symmetric. In networking scenarios with asymmetric wireless links, the lowest capacity among the two directions should be considered. The maximum capacity of the wireless channel is equal to C^{MAX} bit/s. The wireless channel is shared and we assume that every UAV_{*i*} can listen to any other UAV_{*j*}, including UAV₀. This is a realistic assumption, since an FMN is expected to provide wireless connectivity over a confined area of a few hundred square meters (e.g., disaster area or music festival venue) using IEEE 802.11 or 4G/5G communications technologies, which provide a high enough communications range, while benefiting from an improved Line-of-Sight component when compared with terrestrial networks. The Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism is employed for Medium Access Control (MAC), which is in charge of avoiding collisions of frames by enabling transmissions only when the wireless channel is sensed to be idle. Considering $N-1$ FAPs, each generating a traffic flow $F_{i,0}$ with bitrate $T_i(t_k)$ bit/s towards UAV₀, we aim at determining, at each periodic time instant t_k , the position of UAV₀, $P_0 = (x_0, y_0, z_0)$, the transmission power of each FAP_{*i*}, P_T , considering P_T^{MAX} as the maximum transmission power allowed for the wireless technology used, and the queue size for each FAP_{*i*}, Q_i , so that: 1) $C_{i,0}(t_k)$ is high enough for accommodating $T_i(t_k)$ bit/s and $C(t_k) = \sum_{i=1}^{N-1} C_{i,0}(t_k)$ is minimized; 2) the average delay $D_i(t_k)$ is below a given threshold H . Our objective function, which minimizes the overall capacity required to carry the traffic offered by the FAPs, is defined in Eq. (4.1a), together with the constraints of the optimization problem. Using the minimum capacity $C_{i,0}(t_k)$ aims at reducing the number of communications resources required to carry $T_i(t_k)$ bit/s. In our optimization problem, we employ the transmission power P_T as the fine-tuning communications resource. However, this rationale is valid for any other communications resource and corresponding optimization problem formulation, such as when considering the channel bandwidth B as the fine-tuning parameter for computing the capacity $C_{i,0}(t_k)$ by means of the Shannon–Hartley theorem [131]. A discussion on the quality of the admissible solutions obtained, considering a reference networking scenario, is presented in Appendix A.

$$\underset{P_T, (x_0, y_0, z_0), Q_i}{\text{minimize}} \quad C(t_k) = \sum_{i=1}^{N-1} C_{i,0}(t_k) \quad (4.1a)$$

subject to:

$$0 \leq P_T \leq P_T^{MAX} \quad (4.1b)$$

$$C(t_k) \leq C^{MAX} \quad (4.1c)$$

$$0 < T_i(t_k) \leq C_{i,0}(t_k), \quad i \in \{1, \dots, N-1\} \quad (4.1d)$$

$$Q_i(t_k) \geq 0, \quad i \in \{1, \dots, N-1\} \quad (4.1e)$$

$$D_i(t_k) < H, \quad i \in \{1, \dots, N-1\} \quad (4.1f)$$

$$0 \leq x_i \leq x^{MAX}, \quad i \in \{0, \dots, N-1\} \quad (4.1g)$$

$$0 \leq y_i \leq y^{MAX}, \quad i \in \{0, \dots, N-1\} \quad (4.1h)$$

$$0 \leq z_i \leq z^{MAX}, \quad i \in \{0, \dots, N-1\} \quad (4.1i)$$

$$(0, i), (i, 0) \in E(t_k), \quad i \in \{1, \dots, N-1\} \quad (4.1j)$$

$$(x_0 - x_i)^2 + (y_0 - y_i)^2 + (z_0 - z_i)^2 > M^2, \quad i \in \{1, \dots, N-1\} \quad (4.1k)$$

In the optimization problem formulated in Eq. (4.1), the following constraints are considered for any time instant t_k :

- Eq. (4.1b) guarantees that the computed transmission power P_T is between 0 and the maximum value P_T^{MAX} allowed for the wireless technology used.
- Eq. (4.1c) ensures that the aggregate capacity $C(t_k)$ of the wireless links established with UAV₀ is lower than or equal to the maximum capacity C^{MAX} of the shared wireless channel used.
- Eq. (4.1d) assures that the capacity of the wireless link established between UAV_i and UAV₀ is higher than or equal to the bitrate of the traffic flow transmitted by UAV_i.
- Eq. (4.1e) guarantees that the computed queue size for UAV_i is not negative.
- Eq. (4.1f) ensures that the average delay of the packets generated by UAV_i is below a given threshold H , which is defined according to the QoS levels to be offered by the FMN.
- Eq. (4.1g), Eq. (4.1h), Eq. (4.1i) ensure that the position $P_i = (x_i, y_i, z_i)$ of UAV_i is within the dimensions x^{MAX} , y^{MAX} , and z^{MAX} that characterize the 3D space where the FMN should be deployed.

- Eq. (4.1j) assures that a wireless link capable of being established between UAV₀ and UAV_i is always available.
- Eq. (4.1k) guarantees that the position computed for UAV₀ is M meters away from the position of any UAV_i composing the FMN, in order to avoid collisions between UAVs.

When it comes to the 3D space where the UAVs should be deployed, in the current version of our problem formulation we generically consider that the position $P_i = (x_i, y_i, z_i)$ of each UAV_i is within zero and maximum values x^{MAX} , y^{MAX} , and z^{MAX} , which characterize the dimensions of the 3D space available. For safety reasons in a real-world deployment, the minimum value considered for the altitude z_i of UAV_i should assume a minimum safety value greater than 0 m, so that it is positioned above the height of any ground device. This should be defined by the constraint in Eq. (4.1g). Likewise, in order to avoid collisions with any obstacles present at a given location, the 3D space available for placing the UAVs can be defined by means of the constraints defined by Eq. (4.1g), Eq. (4.1h), and Eq. (4.1i). Although ensuring safety when placing UAVs is a relevant research topic, it is beyond the scope of this thesis. Constraints regarding the energy consumption of the UAVs are also not considered in this thesis. This research challenge was approached in [81, 82].

Solving optimization problems typically requires overcoming many challenges, including non-differentiable functions and constraints, large dimensionality, and multiple local minima. Existing solvers can be divided into two main groups: gradient-based and gradient-free solvers. Gradient-based solvers are best suited at finding local minima for high-dimension, non-linear constrained, and convex optimization problems. However, they are not adequate for non-convex optimization problems. In addition, their computational complexity typically grows exponentially with the problem size [132]. Gradient-free solvers are better suited for non-convex problems. Nevertheless, contrary to gradient-based solvers, gradient-free solvers do not guarantee optimal solutions, but only acceptable solutions.

Moreover, the variables that influence the values of P_T , $P_0 = (x_0, y_0, z_0)$, and Q_i , in the problem formulated in Eq. (4.1) include: 1) the capacity of the wireless links between each FAP_i and UAV₀ (FGW), at time t_k ; 2) the number of UAVs that use the same wireless channel; 3) the behavior of the MAC protocol; and 4) the interference between UAVs. These factors are difficult to model mathematically, which makes it difficult to obtain optimal or quasi-optimal solutions in practice. In order to demonstrate that, for the sake of theoretical analysis, we solved the optimization problem using the *GEKKO* optimization suite [133] for a networking scenario consisting of three UAVs. We took into account the Shannon-Hartley theorem [131] for determining the capacity $C_{i,0}(t_k)$ of the wireless link established between each FAP_i and UAV₀, while the queue of each FAP was modeled using the M/D/1 queuing model [121], considering a constant packet size S equal to 1400 bytes. SNR_i was computed by means of the FSPL model, for channel bandwidth B equal to

Table 4.1: Minimum SNR required to select the IEEE 802.11ac MCS indexes 2, 5, and 7, and the corresponding fair share for 3 FAPs. A factor $\eta = 0.80$ is considered to take into account the efficiency of the MAC protocol.

SNR (dB)	MCS index	Data Rate (Mbit/s)	Fair share (Mbit/s)
15	2	175.5	$0.80 \times 175.5/3 \text{ FAPs} = 50$
27	5	468	$0.80 \times 468/3 \text{ FAPs} = 133$
35	7	585	$0.80 \times 585/3 \text{ FAPs} = 166$

160 MHz, 5250 MHz as the carrier frequency f_i , and constant noise power P_N equal to -85 dBm. This allowed us to obtain the point $(x_0, y_0, z_0) = (46.4, 12.3, 10)$ m for the FGW placement, UAVs' transmission power $P_T = 0$ dBm, and queue size Q_1 , Q_2 , and Q_3 equal to respectively 1, 111, and 133 packets. The SNR and capacity of the wireless links obtained are presented in Fig. 4.12.

Due to the non-convex nature of the objective function, as referred to in Chapter 2, it is known that the solution obtained can be a local optimum only. Moreover, since the Shannon-Hartley theorem only provides a theoretical maximum value for the capacity of the wireless links, the capacity obtained in practice in IEEE 802.11 links is substantially lower, especially due to the influence of neighboring nodes competing to access the medium and the overhead associated with the MAC protocol. For these reasons, the obtained theoretical results lead to target SNR values lower than the ones required in practice to achieve the same channel capacity, when considering equal channel bandwidth and noise power. This can be concluded by comparing them with the minimum SNR values required for using the IEEE 802.11ac MCS indexes with a data rate close to the FAPs' traffic demand, which are presented in Table 4.1. Furthermore, the time required to obtain a local optimal or quasi-optimal solution using a state of the art solver raises exponentially as both the number of UAVs and dimensions of the 3D space where the UAVs can be deployed increase. This may avoid obtaining timely solutions in dynamic FMNs that need to be reconfigured in short periods of time. For these reasons, we propose a heuristic algorithm that takes into account domain knowledge to find the adequate traffic-aware solution for the problem formulated in Eq. (4.1).

4.3 GPQM Algorithm

The rationale of the GPQM algorithm and a theoretical analysis for a representative scenario are presented in this section.

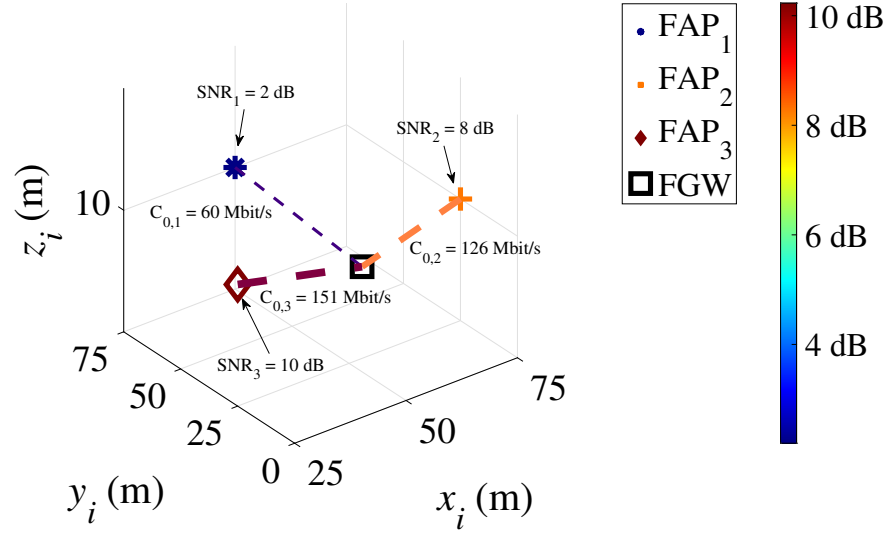


Figure 4.12: Theoretical SNR and capacity values of the wireless links established between each FAP and the FGW, obtained by solving the optimization problem defined in Eq. (4.1). The colors of the markers that represent the FAPs are associated with the order of magnitude of the SNR values of the wireless links established with the FGW, considering the depicted color bar.

4.3.1 Rationale

The GPQM algorithm results from the combination of the GWP algorithm and the PQM mechanism presented in Chapter 2 and Chapter 3, respectively. On the one hand, the PQM mechanism aims at maximizing the throughput of the traffic carried between the FAPs and the FGW, while meeting the average delay constraint. Since PQM configures only the queue size of the communications nodes, it is unable to find a suitable solution that meets the average delay constraint for all flows offered by the FAPs, if the capacity of the wireless links is not high enough. On the other hand, the GWP algorithm aims at ensuring that the wireless link established between each FAP and the FGW has high enough capacity for carrying the offered traffic. For that purpose, the GWP algorithm defines the FGW placement and the transmission power of the UAVs. However, it does not address the queue size management of the communications nodes, which may imply long queuing delays when the queue size is not adequate for the capacity of the wireless links available. As such, GWP is not able to provide average delay guarantees. Building upon the conclusion that the usage of static, oversized queues in IEEE 802.11 networks leads to underutilized communications nodes and unnecessary high delays without improving the aggregate throughput [114], GPQM defines: 1) the FGW position, in order to enable a wireless link between each FAP_{*i*} and the FGW that meets the traffic demand of FAP_{*i*}; and 2) calculating the queue size for each FAP_{*i*}, in order to provide stochastic delay guarantees. The GPQM algorithm allows packets

to be held in the queue of each FAP_i for a short time, while maximizing the aggregate throughput measured in the FGW and improving the shared channel usage. Overall, this allows improving the performance of the FMN. In what follows, we explain the GPQM algorithm, which is presented in Algorithm 2.

GPQM takes advantage of the centralized view of the FMN available at the Central Station, where the FMN is controlled. For that purpose, the Central Station takes into account the information collected and transmitted periodically by the FAPs themselves, including the traffic demand and the positions of the ground devices. Based on this information, the FAP placement algorithm, which is beyond the scope of this thesis, determines the FAPs' positions and the cell ranges of the access network, in order to meet the traffic demand of the ground devices. Before sending the updated information to the FAPs, the GPQM algorithm computes the FGW position and the queue size for each FAP. For the sake of simplicity, we omit t_k in what follows. Considering the future positions of FAP_i and the bitrate of the traffic flow $F_{i,0}$, T_i bit/s, GPQM aims at guaranteeing that a wireless link available between FAP_i and UAV_0 (FGW) has a minimum SNR_i . The minimum SNR_i is defined to enable the usage of a targeted MCS_i index, characterized by a physical data rate higher than or equal to T_i bit/s, while allowing an average packet delay D_i lower than H seconds. Conceptually, if MCS_i is ensured, then the network will be able to accommodate T_i bit/s, allowing to maximize the amount of bits received in the FGW and to minimize the average delay D_i .

The selection of MCS_i imposes a minimum SNR_i , considering a constant noise power P_N . If the transmission power P_T is known, we can compute the maximum admissible distance d_i^{MAX} between FAP_i and UAV_0 , using the FSPL model defined in Eq. (4.2) in dB, where f_i is the carrier frequency, π is the mathematical constant defined in Euclidean geometry as the ratio of a circle's circumference to its diameter, and c represents the speed of light in vacuum.

$$SNR_i = P_T - 20\log_{10}(d_i^{MAX}) - 20\log_{10}(f_i) - 20\log_{10}\left(\frac{4 \times \pi}{c}\right) - P_N \quad (4.2)$$

In 3D space, d_i^{MAX} corresponds to the radius of a sphere, centered at FAP_i , inside which UAV_0 should be placed. Considering $N - 1$ UAVs, the placement subspace for positioning UAV_0 is defined by the intersection of the corresponding spheres $i \in \{1, \dots, N - 1\}$; we refer to this subspace as the FGW Placement Subspace, S_G , as depicted in Fig. 4.13. S_G is computed using GPQM, which iteratively allows obtaining the point $P_0 = (x_0, y_0, z_0)$ for positioning UAV_0 and the minimum transmission power P_T that is assumed to be the same for all UAVs. If no solution is found, Algorithm 2 is terminated, since it constitutes an access network planning problem; in order to achieve a solution for the FGW placement, additional FAPs or communications relays should be included in the FMN, aiming at enabling shorter wireless links established between UAVs configured with transmission power P_T lower than or equal to P_T^{MAX} .

In addition, the GPQM algorithm ensures that FAP_i is configured, at each periodic time instant

Algorithm 2: GPQM algorithm

Input: Traffic demand T_i , in bit/s, of each FAP $_i$, **packet size** S in bytes, and **position** P_i , in 3D Cartesian coordinates, of each FAP $_i$

Output: **Transmission (TX) power** P_T , in dBm, for all UAVs, **FGW position** P_0 in 3D Cartesian coordinates, and **queue size** Q_i , in number of packets, for each FAP $_i$

```

/* 0dBm TX power. */
1  $P_T \leftarrow 0$ 
/* FAP $_i$  packet inter-arrival time. */
2  $\lambda_i \leftarrow \frac{T_i}{S \times 8 \text{ bits/packet}}$ 
/* Iterate up to maximum allowed TX power -> O(n). */
3 while  $P_T \leq P_T^{MAX}$  do
    /* Compute the FGW position, considering  $i \in \{1, \dots, N-1\}$ . SNR $_i$  is determined taking into
       account  $T_i$  by means of a lookup table similar to Table 4.1. */
    4 computeFgwPos( $[x_i], [y_i], [z_i], [\text{SNR}_i], P_T$ )
    /* i.e.,  $(x_0, y_0, z_0) \in S_G$ . */
    5 if  $(x_0, y_0, z_0) \neq \emptyset$  then
        /* FAP $_i$  packet service time. */
        6  $\mu_i \leftarrow \frac{C_{i,0}}{S \times 8 \text{ bits/packet}}$ 
        /* FAP $_i$  traffic load. */
        7  $\rho_i = \frac{\lambda_i}{\mu_i}$ 
        /* FAP $_i$  queue size. */
        8  $Q_i = \frac{1}{2} \times \frac{\rho_i^2}{1 - \rho_i}$ 
        /* FAP $_i$  delay. */
        9  $D_i = \frac{2 - \rho_i}{2 \times \mu_i \times (1 - \rho_i)}$ 
        /* FAP $_i$  delay lower than maximum delay  $H$ . */
        10 if  $D_i < H$  then
            /* TX power for all UAVs, FGW position, and queue size for each FAP $_i$ . */
            11 return  $P_T, (x_0, y_0, z_0), Q_i$ 
        12 end
    13 end
    14 else
        /* Increase TX power by 1dBm. */
        15  $P_T \leftarrow P_T + 1$ 
    16 end
    /* If  $D_i < H$  condition (line 10) is not met, the GPQM algorithm ends when the maximum
       transmission power level  $P_T^{MAX}$  allowed for the wireless technology used is reached;
       GPQM returns TX power for all UAVs, FGW position, and queue size for each FAP $_i$ .
       */
17 end
/* Compute and return  $x_0, y_0$ , and  $z_0$  -> O(1). In the best-case scenario, it has a time
   complexity of O(1); however, this depends on the algorithm employed to solve the
   system of equations. */

```

18 Function **computeFgwPos** ($[x_i], [y_i], [z_i], [\text{SNR}_i], P_T$) :

```

/* In the form of the system of Eq. (4.5). */
19  $(x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2 \leq \left(10^{\frac{K + P_T - \text{SNR}_1}{20}}\right)^2$ ;
20 ...;
21  $(x_0 - x_{N-1})^2 + (y_0 - y_{N-1})^2 + (z_0 - z_{N-1})^2 \leq \left(10^{\frac{K + P_T - \text{SNR}_{N-1}}{20}}\right)^2$ ;
22  $K = -20 \log_{10}(f_i) - 20 \log_{10}\left(\frac{4 \times \pi}{c}\right) - (P_N)$ ;
23 return  $(x_0, y_0, z_0)$ ;

```

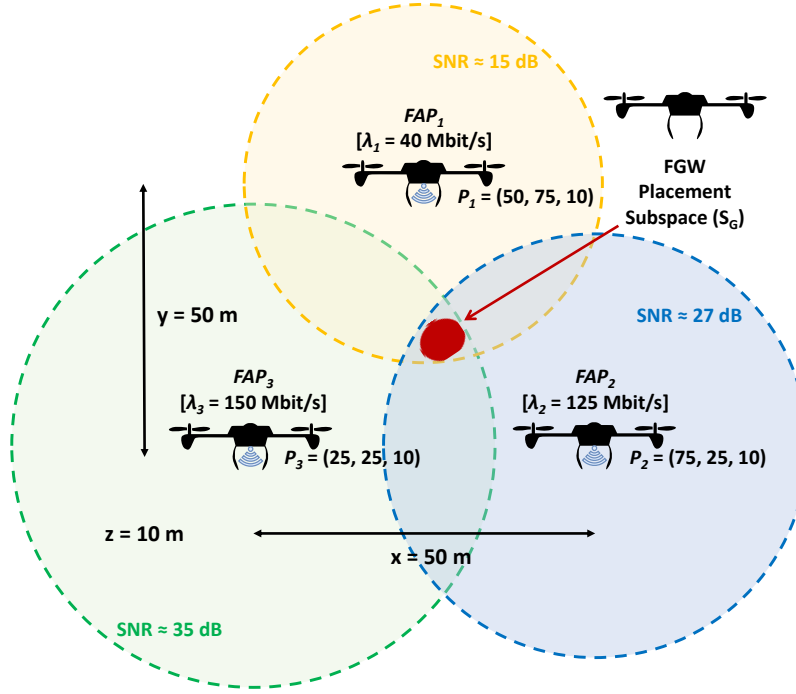


Figure 4.13: Reference scenario depicting the FGW Placement Subspace (S_G) in a 2-Dimensional (2D) space, which results from the intersection of the circumferences, centered at each FAP, with radius equal to the maximum distance compliant with the minimum SNR required.

t_k , with a queue size Q_i able to accommodate the traffic sent towards UAV₀. We assume the M/D/1 queuing model, according to the Kendall's notation [121], considering the data packets generated by the users being served by the FAPs have constant size S and mutually independent inter-arrival times. The packet arrival at each FAP _{i} is assumed to follow a Poisson process. The inter-arrival times follow an exponential distribution whose average is λ_i packet/s. The packet service time is defined by $\mu_i = C_{i,0}/(S \times 8)$ packet/s, while the traffic load of each FAP is expressed by $\rho_i = \lambda_i/\mu_i$. The average number of packets Q_i in each FAP's queue is computed using Eq. (4.3) and the average packet delay is computed using Eq. (4.4) [121], which includes the waiting time in the queue plus the packet service time. By enabling FAP _{i} to have a network queue of size Q_i , we aim at providing an average delay D_i below a given threshold H during t_k .

$$Q_i = \frac{1}{2} \times \frac{\rho_i^2}{1 - \rho_i} \quad (4.3)$$

$$D_i = \frac{2 - \rho_i}{2 \times \mu_i \times (1 - \rho_i)} \quad (4.4)$$

The pseudocode presented in Algorithm 2 is considered as reference for analyzing the time

complexity of the GPQM algorithm. Considering the partial time complexities presented above line 3 and line 14, namely $O(n)$ and $O(1)$, we conclude the GPQM algorithm has time complexity $O(n)$. n represents the transmission power levels required for achieving a valid solution for the problem. P_T , initially set to 0 dBm, is successively increased by 1 dBm up to n times, until a valid solution is found. The GPQM algorithm always converges to a valid solution, since it is a heuristic algorithm leveraged on the theoretical domain knowledge that improving the transmission power P_T of the FAPs leads to higher SNR and improved network capacity for the wireless links. As such, the spheres that represent the transmission ranges of the FAPs always intersect between themselves at a certain transmission power level P_T , unless the maximum transmission power level P_T^{MAX} allowed for the wireless technology used is reached, which is not a limitation of the algorithm itself. The information exchanged between the FMN, the FAP placement algorithm, and the GPQM algorithm is depicted in Fig. 4.14, including the sequential order considered in each FMN reconfiguration.

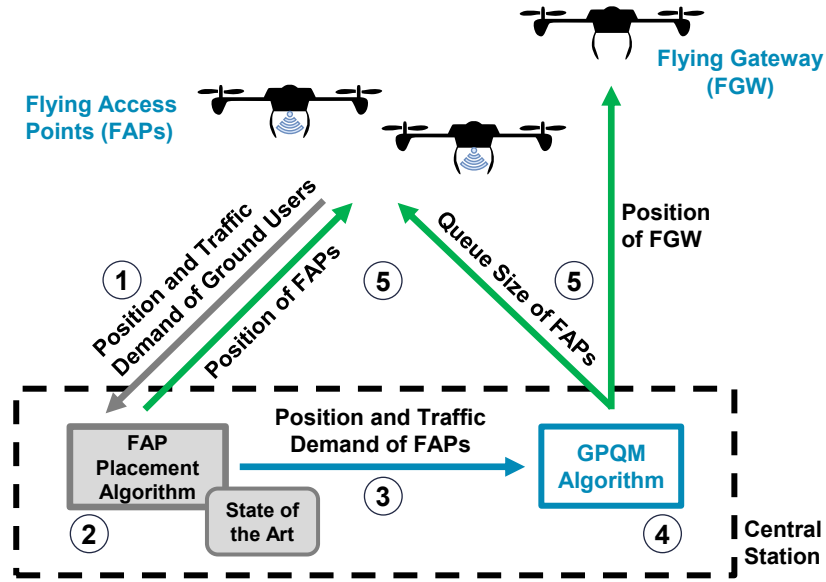


Figure 4.14: Information exchanged between the FMN, the FAP placement algorithm, and the GPQM algorithm. The circled numbers from 1 to 5 define the sequential order considered. The FMN, including the UAVs positions and queue size, is reconfigured each $\Delta t \gg 1$ s.

4.3.2 Theoretical Analysis for a Representative Scenario

Herein, we exemplify the execution of the GPQM algorithm for the simple scenario shown in Fig. 4.13. It consists of three FAPs that are placed in square of side 50 m, hovering at an altitude of 10 m. In this theoretical analysis, we consider the use of the IEEE 802.11ac technology with

one spatial stream, channel bandwidth of 160 MHz, and Guard Interval of 800 ns. Let us assume that the demanded capacity values for FAP₁, FAP₂, and FAP₃ are, respectively, 175.5 Mbit/s, 468 Mbit/s, and 585 Mbit/s, which correspond respectively to the MCS indexes 2, 5, and 7. These demanded capacities were selected in order to exemplify the traffic-awareness of the GPQM algorithm when the FAPs have different traffic demand values, which may occur in the real-world due to a different number of ground devices served. Taking into account that a minimum SNR value is required to use a target MCS index, which is characterized by a theoretical data rate, in Table 4.1 we conclude that the target SNR values are 15 dB for FAP₁, 27 dB for FAP₂, and 35 dB for FAP₃.

Solving the system of Eq. (4.5), which is derived from Eq. (4.2), and considering d_i^{MAX} as the Euclidean distance between UAV₀ (FGW) and each FAP_{*i*}, we conclude that a suitable position for the FGW is $(x_0, y_0, z_0) = (39.0, 21.9, 9.2)$ m, for a transmission power $P_T = 20$ dBm. In the system of Eq. (4.5), P_T is the fine-tuning parameter, so that we can find at least a point $(x_0, y_0, z_0) \in S_G$. P_T is initially set to 0 dBm. Then, it is iteratively increased by 1 dBm until a valid solution is found or the maximum transmission power allowed P_T^{MAX} is reached. The factor K is computed considering the carrier frequency f_i used and the constant noise power P_N : 5250 MHz and -85 dBm, respectively.

$$\begin{cases} (x_0 - 50)^2 + (y_0 - 75)^2 + (z_0 - 10)^2 \leq \left(10^{\frac{K+P_T-15}{20}} \right)^2 \\ (x_0 - 75)^2 + (y_0 - 25)^2 + (z_0 - 10)^2 \leq \left(10^{\frac{K+P_T-27}{20}} \right)^2 \\ (x_0 - 25)^2 + (y_0 - 25)^2 + (z_0 - 10)^2 \leq \left(10^{\frac{K+P_T-35}{20}} \right)^2 \\ K = -20\log_{10}(5250 \times 10^6) - 20\log_{10}\left(\frac{4 \times \pi}{3 \times 10^8}\right) - (-85) \end{cases} \quad (4.5)$$

Since the wireless channel is shared by $N - 1$ FAPs generating traffic towards UAV₀ (FGW), in order to define the bitrate of the traffic flow transmitted by the FAPs, we assumed a reference fair share $L = 0.80 \times (585 \text{ Mbit/s}) / (N - 1)$ of the maximum channel capacity allowed, where $\eta = 0.80$ is the efficiency of the MAC protocol, including the overhead of the transport protocol, Internet Protocol (IP), and MAC packet headers. 585 Mbit/s is the data rate associated to the MCS index 7, which is the highest MCS index to be ensured by the FMN so that the traffic demand of FAP₃ is met. The minimum SNR required to use each IEEE 802.11ac MCS index (cf. Table 4.1) and the factor η were obtained by means of Network Simulator 3 (ns-3) simulations, considering a FAP generating User Datagram Protocol (UDP) Poisson traffic towards the FGW, for a constant transmission power equal to 20 dBm. For that purpose, the FAPs was moving away from the FGW, in order to increase the Euclidean distance, leading to the degradation of the wireless link's

SNR, and enabling the selection of multiple MCS indexes. This allowed establishing the relation between the wireless link's SNR values and the selected MCS indexes.

For the networking scenario depicted in Fig. 4.13, let us assume that the bitrate of the traffic offered by the FAPs is equal to 40 Mbit/s, 125 Mbit/s, and 150 Mbit/s, which correspond, respectively, to approximately 25%, 75%, and 90% of the reference fair share L (166 Mbit/s) of the maximum channel capacity allowed (cf. Table 4.1). In order to compute the queue size, a constant packet size S equal to 1400 bytes is considered in this example, which is close to the Maximum Transmission Unit of 1500 bytes typical in data packets composing the Internet traffic [122]. Considering the specific case of FAP₃ for illustrative purposes, this results in $\lambda_3 = (150 \text{ Mbit/s}) / (1400 \times 8 \text{ bits/packet}) = 13393 \text{ packet/s}$ and $\mu_3 = (166 \text{ Mbit/s}) / (1400 \times 8 \text{ bits/packet}) = 14821 \text{ packet/s}$, leading to $\rho_3 \approx 0.9$ and $Q_3 = 5 \text{ packets}$, according to Eq. (4.3). We argue that this queue size allows that packets are held in the transmission queue for shorter time, thus decreasing the queuing delay and maximizing the throughput, when compared to the baseline that does not employ any queue management mechanism. The same rationale is employed for the remaining FAPs that compose the FMN.

4.4 Performance Evaluation

The performance evaluation when the GPQM algorithm is used is presented in this section. Firstly, we describe the evaluation methodology. Secondly, we characterize the simulation setup. Thirdly, we present and discuss the network performance results obtained in simulation. Finally, we compare the GPQM algorithm with state of the art solvers.

4.4.1 Evaluation Methodology

Since 1) a purely theoretical performance evaluation would require many simplifications, due to the high complexity of the problem, leading to unrealistic results, and 2) an experimental evaluation would be limited, with reduced scalability, and expensive, due to the inherent complexity of setting up a testbed consisting of multiple UAVs, we carried out a performance evaluation based on ns-3 simulations. The ns-3 simulator [73] was chosen due to its accurate models for wireless networks and its wide acceptance within the scientific community.

Although the ns-3 simulator includes a set of theoretical propagation models tailored for wireless networks, in order to improve the accuracy and realism of the performance evaluation presented herein, the wireless channel between each FAP_{*i*} and the FGW was characterized based on the experimental model proposed in [21], where it was analyzed in terms of the path-loss and fast-fading components. Regarding the path loss component, it was concluded that the FSPL model is the best suited to characterize the wireless links in FMNs. The FSPL model (cf. dashed

green line in Fig. 4.15) expresses that the power loss is proportional to the square of the distance between the transmitter and the receiver communications nodes, while proportional to the square of the carrier frequency used. This model fits the experimental SNR values measured in an FMN, which is justified by the strong Line-of-Sight component induced by the UAV altitude, as suggested in the literature [134]. For comparison purposes, the two-ray ground reflection model (cf. dotted blue line in Fig. 4.15) is also depicted. It takes into account a component reflected on the ground and the influence of the antennas' height, in addition to the power loss induced by the Line-of-Sight component. For distances between the transmitter and the receiver less than the height of the transmitter's antenna, two waves are added in a constructive way, increasing the power of the received signal. As the distance increases, these waves add up constructively and destructively, giving rise to up-fade and down-fade lobes in the power of the received signal, as depicted in Fig. 4.15.

On the other hand, the *Rician* fast-fading component, commonly called K-factor, which was characterized in [21], is also considered. The K-factor constitutes the ratio of the received power in the dominant component over the scattered power. We take into consideration worst-case scenarios where the wireless links established between the UAVs are modeled by the FSPL model and the *Rician* fast-fading with a low K-factor value. This aims at considering the wireless conditions sensed by UAVs when they are flying at low altitudes or placed on a ground platform.

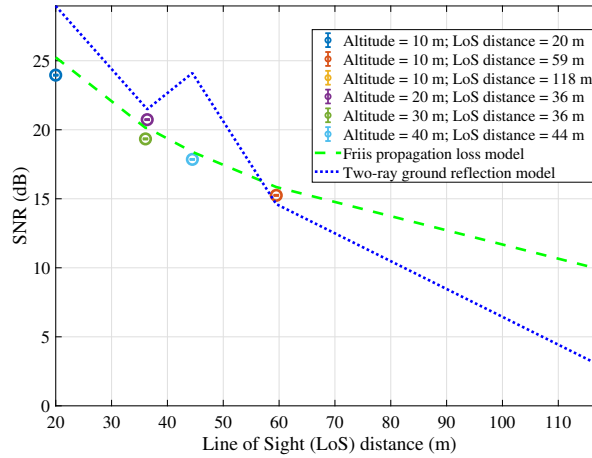


Figure 4.15: Experimental SNR versus the Line-of-Sight distance of the wireless link, considering different UAV altitudes. The FSPL (Friis) and two-ray ground reflection models are represented by the dashed and dotted lines, respectively [21].

As a step forward to the original performance validation of the GWP algorithm and the PQM mechanism, which considered a theoretical radio propagation model only, GPQM was evaluated under a realistic wireless channel model, based on experimental measurements collected in a

testbed. Moreover, different traffic generation models were considered, including UDP Poisson, UDP *OnOff*, and Transmission Control Protocol (TCP) *BulkSend*, allowing to emulate the conditions imposed by real-world applications in multiple networking scenarios, which also goes beyond the evaluation of the original versions of PQM and GWP performed using only UDP Poisson traffic. The reference scenario, which is depicted in Fig. 4.13, aimed at evaluating the network performance when employing the FGW placement and the queue size derived in the theoretical analysis presented in Section 4.3.2. Also, GPQM was evaluated under three sets of 5 random dynamic networking scenarios consisting of different numbers of FAPs, in order to consider small, medium, and large FMNs, respectively, according to the number of UAVs expected to be used in practice [135].

In order to evaluate the network performance when GPQM is used, two placement approaches and two state of the art traffic control algorithms were considered. The counterpart placement approaches include: 1) the FGW placed in the FAPs center (i.e., three-coordinates average considering all FAPs), which is the position that maximizes the SNR of the wireless links established between the FAPs and the FGW; and 2) the FGW placed in the center of the venue. The counterpart traffic control algorithms include: 1) Random Early Detection (RED), which drops packets based on the queue size; and 2) Controlled Delay (CoDel), which takes dropping decisions based on the time the packets are held in the queue. RED and CoDel were employing their default parameters in ns-3 [136], but the packet size value S was provided, in order to enable a fair comparison with the GPQM algorithm. The performance evaluation carried out considers two performance metrics:

- **Aggregate throughput:** the average number of bits received per second by UAV₀ (FGW).
- **Delay:** the average time taken by the packets to reach the sink application of UAV₀ since the instant of time they were generated by the source application of each FAP_{*i*}, which includes queuing, transmission, and propagation delays.

Moreover, in order to assess the quality of the solutions obtained by means of the GPQM algorithm, we solved the optimization problem formulated in Eq. (4.1) and compared the network performance achieved when employing the solutions obtained with: 1) Interior Point OPTimizer (IPOPT) [137], which is a gradient-based solver; and 2) Particle Swarm Optimization (PSO) [138], which is a gradient-free solver. IPOPT is targeted for large-scale nonlinear constrained optimization problems defined in n dimensions, which aim at minimizing an objective function $f(x)$ subject to a set of constraints. The *GEKKO* optimization suite [133], which employs IPOPT as the default solver, was used to solve the optimization problem considering this approach. In turn, PSO is inspired in the social behavior of a flock of birds. It consists of a set of s candidate solutions, called particles, which are moved in an n -dimensional hyperspace, where n is the number of parameters to be optimized. The particles are moved taking into account their position and velocity, which

are updated in each iteration. The movement of each particle is influenced by its best-known local position and thrives towards the best global positions found by all particles.

For solving the optimization problem formulated in Eq. (4.1), we assumed that the capacity $C_{i,0}(t_k)$ of each wireless link results from the fair share associated to the MCS index selected by each FAP_i (cf. Table 4.1), which requires a minimum SNR value, following the rationale of GPQM. We modeled the relation between SNR and fair share as a continuous function using a linear regression, which is a function that closely fits the data; it is depicted in Fig. 4.16 for 3 FAPs that use the same wireless channel. In order to compute the theoretical SNR values, we employed the FSPL model (cf. Eq. (4.2)), considering constant noise power, whereas the queue size and the average packet delay were determined by means of the M/D/1 queuing model (cf. Eq. (4.3) and Eq. (4.4)), which assumes the data packets have constant size and mutually independent inter-arrival times. This enables a fair comparison with the GPQM algorithm, random networking scenarios consisting of 3, 6, and 12 static FAPs were considered; in this theoretical analysis, static FAPs were considered due to the challenges imposed by dynamic networking scenarios to the gradient-based solvers, including non-differentiable functions and constraints, multiple local minima, and the need to specify the initial values and set the variable bounds.

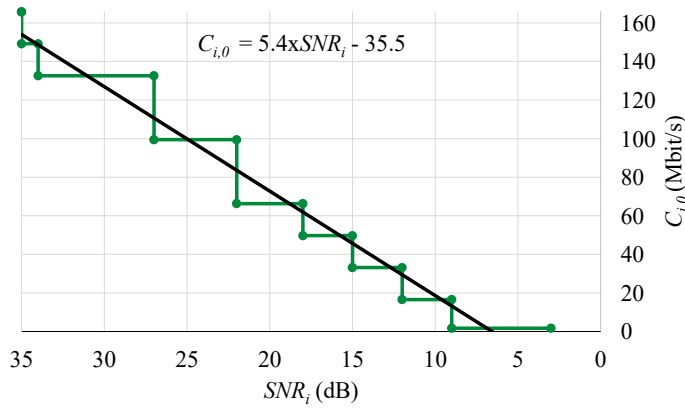


Figure 4.16: Wireless channel capacity modeled by a linear regression between the minimum SNR values and the data rate associated to the IEEE 802.11ac MCS indexes, considering 3 FAPs using the same wireless channel.

4.4.2 Simulation Setup

In order to evaluate the performance of GPQM in ns-3, the FAPs and the FGW were configured to use a Network Interface Card (NIC) in Ad Hoc mode, considering the IEEE 802.11ac technology in channel 50, which enables channel bandwidth 160 MHz and Guard Interval of 800 ns. The wireless link between each FAP_i and the FGW was employing a single spatial stream; due to

the strong Line-of-Sight, the spatial diversity provided by the Multiple-Input and Multiple-Output technique is not a relevant advantage for the wireless links established between UAVs. Three different traffic types were considered: 1) UDP Poisson, which is a traffic model widely used in queuing theory, since the aggregation of multiple independent flows constitutes a Poisson arrival process; 2) UDP *OnOff*, in order to represent traffic sources that include active and idle periods, such as *Web* servers; and 3) TCP *BulkSend*, which is made up of multiple bursts of consecutive packets, typically associated to a File Transfer Protocol session [139]. The data rate was defined by the *IdealWifiManager* auto rate mechanism [140]. The main simulation parameters used are summarized in Table 4.2.

The FAPs were moving according to the Random Waypoint Mobility model, with speed uniformly distributed between 0.5 m/s and 3 m/s, in order to consider time-varying capacity for the wireless links. It is worth mentioning that dynamic networking scenarios consisting in FAPs were only considered for evaluating the network performance achieved GPQM against the state of the art counterparts; in order to assess the quality of the solutions obtained by GPQM, by comparing ns-3 simulations results with the values determined using state of the art solvers for representative networking scenarios, we considered static FAPs, due to the reasons referred at the end of Section 4.4.1. The UAVs were placed inside a venue with dimensions $x^{MAX} = 100$ m long, $y^{MAX} = 100$ m wide, and $z^{MAX} = 20$ m high. The traffic demand of the FAPs was variable, according to the networking scenario employed, but always lower than the maximum capacity of the shared wireless channel. Since the GPQM algorithm relies on knowing in advance the positions of the FAPs, which in a real-world deployment are provided by the FAP placement algorithm running at the Central Station, we used *BonnMotion* [107], which is a generator of mobile scenarios. The resulting waypoints, consisting of the future positions of the FAPs at each periodic time instant t_k , were used to compute in advance the FGW position and the queue size for each FAP over time. Finally, the FGW position, the position and queue size for each FAP, and the transmission power for the UAVs for the whole simulation time, with a sampling period of 1 s, were imported to ns-3. The *WaypointMobilityModel* [141] was used to place the FAPs and the FGW over time in the simulation environment, according to the positions generated by *BonnMotion* and defined by GPQM, respectively, which were exported to ns-3. To configure the queue size at each FAP over time, a *DropTailQueue* with size scheduled in advance was implemented. The simulation results consider 20 runs for each scenario described in Section 4.4.1, under the same exact networking conditions.

4.4.3 Simulation Results

The obtained simulation results are the average values for the two performance metrics defined in Section 4.4.1, computed for each second of the simulation runs. The set of average values obtained for each performance metric is represented by means of the Cumulative Distribution

Table 4.2: Summary of the main ns-3 simulation parameters used in the evaluation of the GPQM algorithm.

Parameter	Value
Simulation time	30 s (bootstrap) + 70 s
Wi-Fi standard	IEEE 802.11ac
Wi-Fi mode	Ad Hoc
Wi-Fi channel	50 (5250 MHz)
Channel bandwidth	160 MHz
Guard Interval	800 ns
Number of antennas	1
Propagation delay	Constant speed
Propagation model	FSPL (Friis)
<i>Rician</i> K-factor	13 dB
Remote station manager	<i>IdealWifiManager</i>
Mobility model	Waypoint
Packet size	1400 bytes
Traffic types	UDP Poisson and <i>OnOff</i> , and TCP <i>BulkSend</i>

Function (CDF) for the packet delay and by the Complementary Cumulative Distribution Function (CCDF) for the aggregate throughput. The CDF $F(x)$ represents the percentage of samples for which the delay was lower than or equal to x , while the CCDF $F'(x)$ represents the percentage of samples for which the aggregate throughput was higher than x . The analysis presented herein takes into account the results for the 90th percentile.

4.4.3.1 Reference Scenario

Regarding the reference scenario (cf. Fig. 4.13), we conclude that GPQM allows improving the aggregate throughput up to 85%, considering UDP Poisson traffic (cf. Fig. 4.17), while the average delay is decreased up to 60%. Similarly, for UDP *OnOff* traffic (cf. Fig. 4.18), the average delay is decreased up to 56%, while the aggregate throughput is improved up to 80%. When considering TCP *BulkSend* traffic (cf. Fig. 4.19), the average delay is decreased up to 38%, while the aggregate throughput is improved in 11%. GPQM is the solution that provides the best network performance. This is achieved by ensuring the smallest possible queue size for each FAP_{*i*}, thus avoiding packets from being held in the transmission queue for a long time, which significantly increases the delay of all packets without relevant gains in the aggregate throughput.

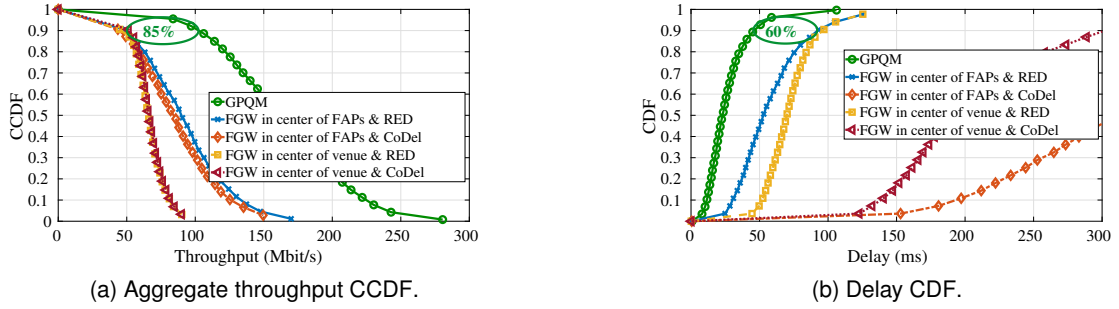


Figure 4.17: Performance results for 3 static FAPs (reference scenario of Fig. 4.13), considering UDP Poisson traffic.

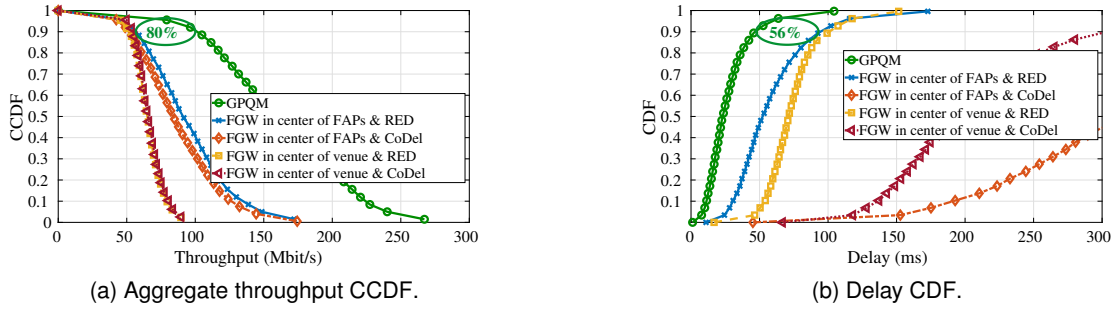


Figure 4.18: Performance results for 3 static FAPs (reference scenario of Fig. 4.13), considering UDP *OnOff* traffic.

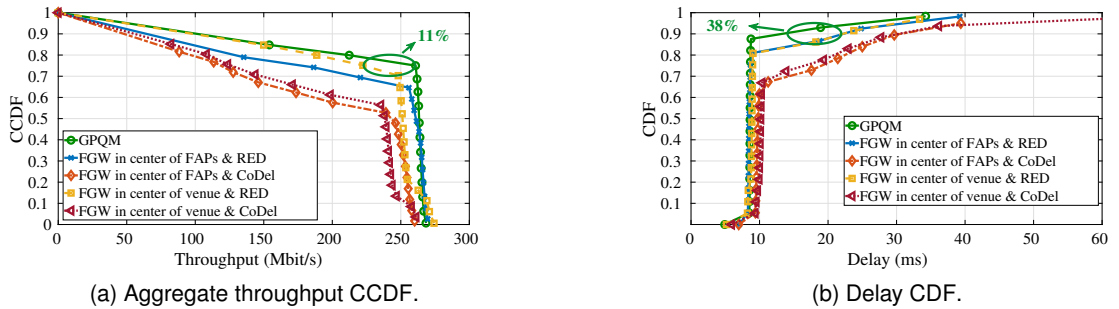


Figure 4.19: Performance results for 3 static FAPs (reference scenario of Fig. 4.13), considering TCP *BulkSend* traffic.

4.4.3.2 Dynamic Scenarios

The results for the dynamic scenarios, which consisted of 3, 6, and 12 FAPs in movement during 70 s simulation time, are depicted in Fig. 4.20, Fig. 4.21, Fig. 4.22, Fig. 4.23, and Fig. 4.24,

considering UDP Poisson, UDP *OnOff*, and TCP *BulkSend* traffic types. GPQM improves the aggregate throughput up to 82% while reducing the average delay up to 58% (cf. Fig. 4.20). In the worst-case scenario, GPQM enables approximately the same aggregate throughput for a decrease in the average delay up to 15% (cf. Fig. 4.21b), which was observed for TCP *BulkSend* traffic, considering 3 FAPs. The gains in the average delay for TCP are higher when the number of FAPs increases, as expressed by the gain of 22% in Fig. 4.24b. The lower gains for TCP traffic with respect to UDP traffic are justified by the fact that TCP already includes a congestion control mechanism. The gains are even more relevant if we consider CoDel as the counterpart queue management algorithm. Overall, CoDel allows achieving high throughput, at the expense of increased average delay. Taking into account the results for UDP *OnOff* traffic and 12 FAPs, the aggregate throughput achieved by GPQM is approximately 5% lower with respect to CoDel; however, GPQM allows decreasing the average delay up to 74% (cf. Fig. 4.23).

The performance evaluation carried out allows concluding that GPQM, in the worst-case scenario, enables approximately the same aggregate throughput for a lower average delay. In addition, we can infer that the influence of GPQM on performance improvement is lower when the number of FAPs increases. Yet, the number of FAPs that compose real-world FMNs is typically limited, motivating the use of GPQM in practice.

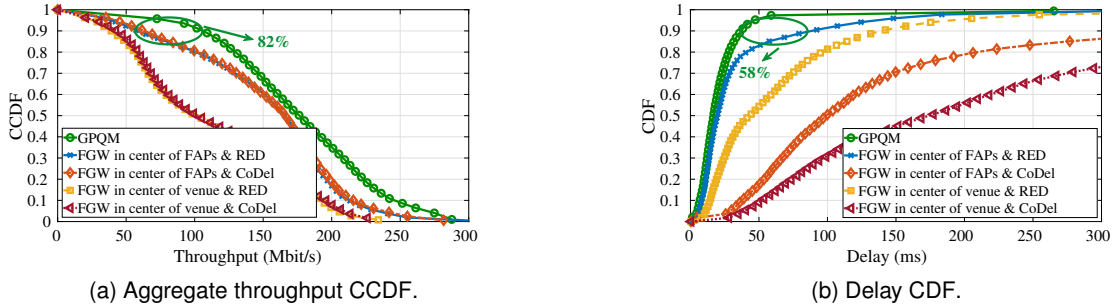


Figure 4.20: Performance results for 3 moving FAPs, under 5 random dynamic networking scenarios, considering UDP *OnOff* traffic.

4.4.4 GPQM Benchmarking

Taking into account the FGW position, $P_0 = (x_0, y_0, z_0)$, the transmission power of the UAVs, P_T , and the queue size for each FAP_{*i*}, Q_i , determined by means of 1) the GPQM algorithm, 2) the gradient-based solver IPOPT, and 3) the gradient-free solver PSO, ns-3 simulations were carried out to evaluate the FMN performance and assess the quality of the solutions obtained. ns-3 allows simulating the stochastic characteristics of wireless networks and the behavior of representative traffic generation models, which are too complex to be modeled mathematically. For

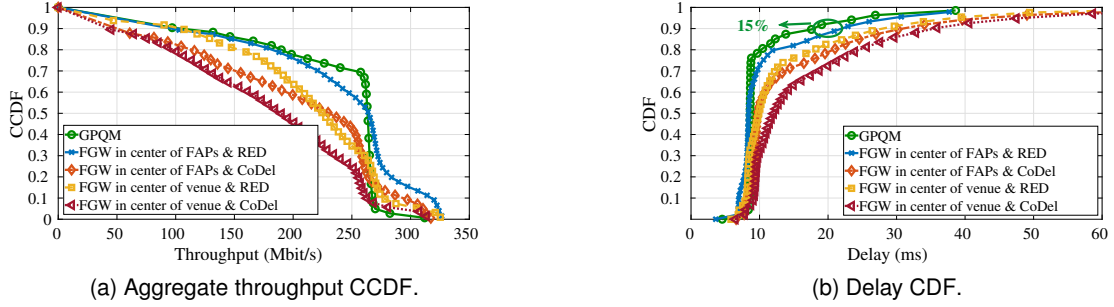


Figure 4.21: Performance results for 3 moving FAPs, under 5 random networking scenarios, considering TCP *BulkSend* traffic.

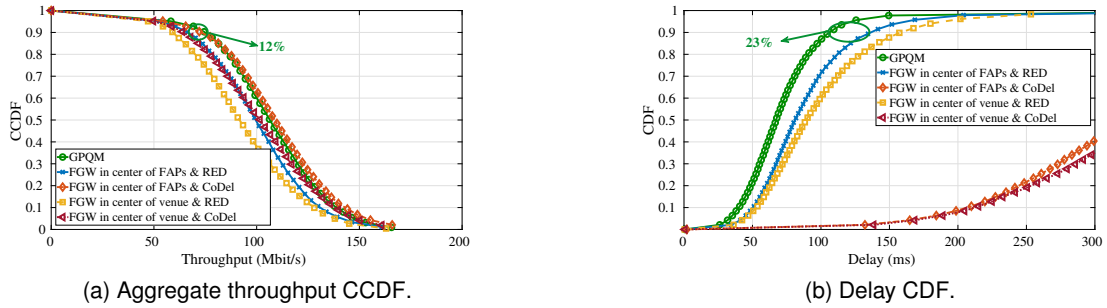


Figure 4.22: Performance results for 6 moving FAPs, under 5 random networking scenarios, considering UDP Poisson traffic.

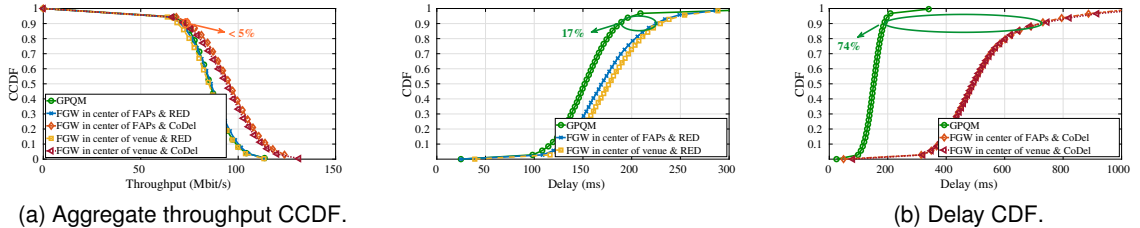


Figure 4.23: Performance results for 12 moving FAPs, under 5 random networking scenarios, considering UDP *OnOff* traffic.

that purpose, it takes advantage of simulation models that represent abstractions of communications nodes, allowing to accurately evaluate the performance of wireless networks by means of a packet-driven approach. We observed in practice that, as the number of UAVs increases, the time required to solve the optimization problem defined in Eq. (4.1) grows exponentially, which makes it difficult to obtain solutions for dynamic FMNs consisting of UAVs in movement. For this reason, we considered static networking scenarios, consisting of 3, 6, and 12 FAPs in fixed positions, for

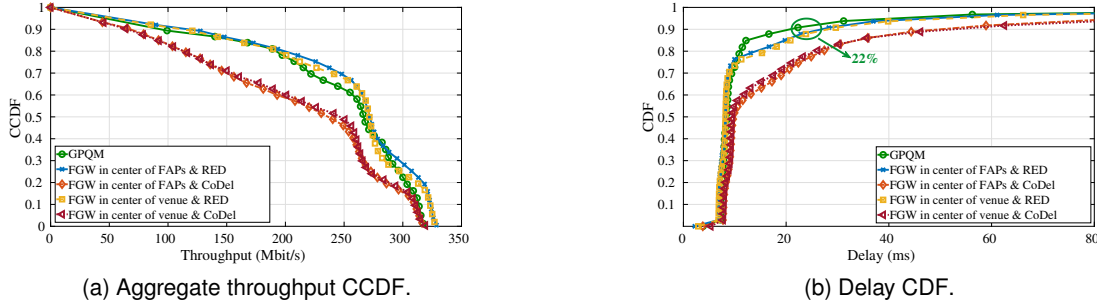


Figure 4.24: Performance results for 12 moving FAPs, under 5 random networking scenarios, considering TCP *BulkSend* traffic.

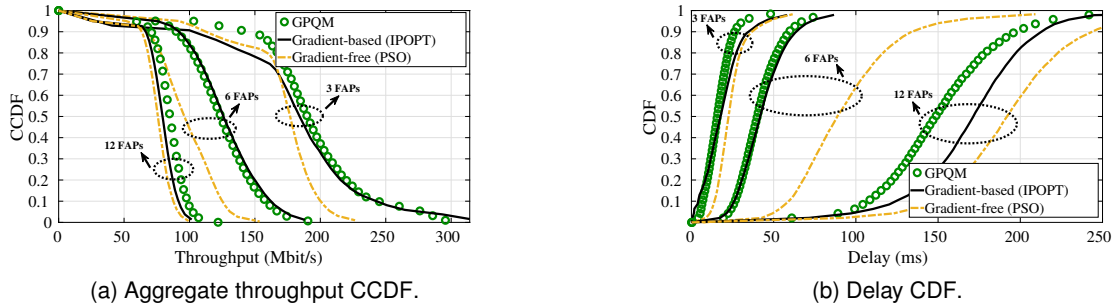


Figure 4.25: Benchmarking results considering the FGW position, $P_0 = (x_0, y_0, z_0)$, the transmission power of the UAVs, P_T , and the queue size for each FAP_{*i*}, Q_i , determined by means of 1) the GPQM algorithm (green circles), 2) the gradient-based solver (solid black lines), and 3) the gradient-free solver (dashed yellow lines), considering the optimization problem formulated in Eq. (4.1), for 3, 6, and 12 FAPs generating UDP Poisson traffic, under 5 random networking scenarios for each number of FAPs.

benchmarking purposes. It is important to note that this limitation, which led to the use of static networking scenarios in this evaluation, was imposed by IPOPT and PSO; the GPQM algorithm was evaluated considering dynamic networking scenarios in Section 4.4.3.2. The benchmarking results are depicted in Fig. 4.25. The plots clearly show that the GPQM algorithm (green circles) leads to the best network performance. While the solutions obtained using a gradient-free solver (PSO) provide the worst network performance (dashed yellow lines), since they are only acceptable solutions, the solutions obtained using a gradient-based solver (IPOPT) (solid black lines) are closer to the results obtained using GPQM. Yet, solving an optimization problem using a gradient-based solver implies specifying initial values, setting variable bounds, and scaling variables and equations. For different initial values, different solutions are achieved. However, the convergence to the optimal solutions relies on good directional information (i.e., good derivatives),

for which providing an initial point close to a local optimum and compliant with all constraints of the optimization problem may improve the solver performance. Still, setting the initial values is not straightforward, since there is no information about the regions where the solutions of the optimization problem are located. This makes it difficult to use gradient-based solvers in networks with dynamic topology, for which the initial values must be defined for each topology reconfiguration. GPQM employs a heuristic approach based on the theoretical domain knowledge that improving the transmission power of the communications nodes leads to higher SNR for the wireless links and improved network capacity. This is the main differentiating factor between GPQM and state of the art solvers, since the latter change the values of the variables to be optimized without domain knowledge about their influence on network performance.

4.5 Discussion

The performance evaluation carried out allows concluding significant gains in the QoS offered by the FMN when the GPQM algorithm is used. GPQM is especially suitable for highly dynamic FMNs, which motivate adjustments in the FGW placement over time, in order to meet the traffic demand of the FAPs, as well queues with dynamic size, for addressing the changes in the packet service time that are induced by the time-varying capacity of the wireless links. The use of GPQM is especially useful in networks at the onset of congestion, typical in crowded and disaster management scenarios. The GPQM algorithm was designed for networking scenarios consisting of a single FGW; however, it can also be employed in networking scenarios with multiple FGWs, which should be configured to use orthogonal wireless channels, in order to minimize interference. This is especially relevant when 1) a single wireless channel does not provide high enough capacity to meet the traffic demand of all FAPs in the network and 2) the FAPs are placed over a wide area, such that a FAP may get out of reach of the FGW it is currently assigned to. For that purpose, clusters of closer FAPs can be formed, and for each cluster of FAPs a dedicated FGW may be assigned. In this sense, the GPQM algorithm can be employed to define the position of the FGW of each cluster of FAPs as well as the FAPs' queue size. The overall network performance will be improved if the multiple FGWs are able to use different connections to the Internet (e.g., cellular BSs and tethered drones).

GPQM provides the same output whether downlink or uplink traffic is considered, since all the UAVs are configured with the same transmission power and the wireless channel is assumed to be symmetric. In this sense, for the downlink traffic, the FGW may include a queue for each FAP and take advantage of a scheduling mechanism to select which queue to remove a packet from over time, in order to ensure fairness. This paves the way to use the GPQM algorithm in emerging networking scenarios where symmetric traffic applications are growing [109], such as social networks, video streaming, and online gaming. In networking scenarios with asymmetric

wireless links, the lowest capacity among the two directions should be considered for placing the FGW.

The position of the FGW and the queue size of the FAPs are computed by the GPQM algorithm for discrete time instants. Each time instant corresponds to a snapshot of the FMN where the FGW and the FAPs are hovering in fixed positions (i.e., zero flying speed), in order to serve the ground devices. The periodic time instants considered by the GPQM algorithm are defined based on the update period of the state of the art FAP placement algorithm used. Considering discrete time instants for computing up-to-date decisions for both the UAV placement and queue management allows considering a trade-off between: 1) the stability of the FMN; 2) the time required by the algorithms for computing the decisions; 3) the time taken by the UAVs to reconfigure themselves according to the up-to-date decisions; and 4) the network performance gains obtained in practice. If the update period is set to a high value, the FMN will be more stable but will take a longer time to react to changes in the QoS levels demanded by the ground devices. Conversely, if the update period is set to a low value, the FMN will react quicker to changes in the QoS levels demanded by the ground devices. However, the FMN will be less stable, possibly making it impossible for UAVs to position themselves as fast as the decisions computed by the algorithms, which may result in negligible or no gains achieved in practice.

It must be noted that, since the transmission power P_T is the fine-tuning parameter to achieve a suitable FGW position, in networking scenarios consisting of 1) multiple FAPs placed over a wide area and 2) FAPs with traffic demand close to or higher than the maximum channel capacity, P_T will tend to infinite, which is not feasible in a real-world deployment. In order to address this challenge, the evolution of the GPQM algorithm for networking scenarios consisting of flying communications relays able to forward data traffic between them, thus enabling short-range wireless links, is worth considering. For that purpose, the flying communications relays should be configured to form multi-hop wireless paths using different communications channels, in order to mitigate the inter-flow interference and enhance the overall capacity of the FMN, while employing reduced transmission power values.

The networking scenarios employed in the evaluation of GPQM were defined so that the FAPs listen to each other, in order to avoid the hidden node problem. In order to avoid the hidden node problem in real-world FMNs, where the scenarios are not fully controlled, the GPQM algorithm is worth improving so that the transmission power defined for the FAPs is high enough not only to ensure the minimum SNR required to meet the traffic demand of the FAPs, but also to ensure the FAPs listen to each other. Since the traffic-aware gateway placement and queue management problem is our focus, the MAC challenges are left for future work.

The performance evaluation of the GPQM algorithm, which was carried out under realistic channel conditions for multiple networking scenarios, allows validating GPQM and confirm that the performance of the FMN is improved by defining in advance 1) the FGW position and 2) the

queue size for each FAP, considering both the positions of the FAPs and their traffic demand over time.

4.6 Summary

In addition to the ability to accommodate the offered traffic, FMNs providing wireless connectivity are expected to meet strict low-latency requirements. To ensure low-latency communications, placing UAVs close to the devices to be served, while enabling short-distance Line-of-Sight wireless links, represents an advantage over ground wireless networks with limited mobility. Yet, new challenges from the network configuration point of view arise. In particular, the size of the queues configured in the communications nodes, which are used for accommodating transient traffic bursts offered to the network needs to be defined. A baseline approach lies on considering oversized queues, in order to minimize the number of dropped packets. The main disadvantage is the increased packet delay, leading to the well-known bufferbloat problem. Conversely, small queues may be unable to accommodate the traffic offered by the devices, especially bursty traffic. When compared with wireless networks in general, FMNs bring up additional challenges, including: 1) potential increase in inter-flow interference, due to enhanced wireless range induced by the altitude factor and absence of obstacles; and 2) frequent changes in the SNR of the wireless links due to the highly dynamic nature of FMNs, which induce time-varying network capacity. For these reasons, queues with static size and the state of the art Active Queue Management (AQM) algorithms, which are designed for wired networks, are typically not adequate for FMNs.

In order to address the queue size management in FMNs, we proposed in Chapter 3 the PQM mechanism, which allows providing stochastic delay guarantees. However, PQM is limited to the proactive management of the queue size. If the capacity of the wireless links is insufficient to accommodate the offered traffic, it may be impossible to determine a suitable solution that satisfies the average delay constraint for all the traffic flows. The GWP algorithm, described in Chapter 2, solves the problem by ensuring that the wireless link available between each FAP and the FGW has a sufficient network capacity for carrying the offered traffic. It defines the transmission power for the UAVs and the FGW placement. Considering that the use of large queues in IEEE 802.11 networks lead to underused communications nodes and unnecessary high delays without enhancing the aggregate throughput, in this chapter we went a step forward and presented a traffic-aware gateway placement and queue management algorithm for FMNs, called GPQM. Taking advantage of the information on the future positions of the FAPs, defined by a state of the art FAP placement algorithm so that the traffic demand of the ground devices are met, the GPQM algorithm is able to 1) determine the FGW position, in order to accommodate the offered traffic, and 2) compute the queue size for each FAP, in order to enable stochastic delay guarantees. The performance evaluation of GPQM, carried out by means of ns-3 simulations considering a realistic

wireless channel model built upon experimental results collected in a testbed, show gains up to 85% in the aggregate throughput, while reducing the average packet delay up to 74%.

The GPQM algorithm allows jointly defining the adequate placement for the FGWs and a suitable queue size for the FAPs, which are aspects addressed independently by the state of the art counterparts, without taking advantage of the gains allowed by a joint solution. GPQM allows minimizing the packet service time and the traffic load in each FAP, thus improving the overall network performance. Since GPQM leverages on the holistic knowledge about the FMN, made available in the Central Station, this constitutes a differentiating factor with respect to state of the art algorithms that rely on mobility prediction-based approaches or real-time measurements, which typically imply long convergence times to achieve a suitable solution. It is worth mentioning that the up-to-date decisions are taken by the GPQM algorithm for discrete time instants, each corresponding to a snapshot of the FMN. Considering that when the UAVs move for repositioning themselves the wireless links can be disrupted, a routing mechanism that enables the seamless FMN reconfiguration is worth considering and is the subject of the next chapter.

Chapter 5

A Mechanism for Predictive Routing in Flying Multi-hop Networks

The Gateway Placement and Queue Management (GPQM) algorithm presented in Chapter 4 considers the future positions of the Flying Access Points (FAPs), defined by means of a state of the art FAP placement algorithm, as input for determining the Flying Gateway (FGW)'s position and the FAPs' queue sizes. A challenge that should be addressed is the seamless Flying Multi-hop Network (FMN) reconfiguration when the Unmanned Aerial Vehicles (UAVs) move for repositioning themselves. Each FMN reconfiguration occurs when the state of the FAP placement algorithm and the GPQM algorithm compute up-to-date decisions; these decisions are computed at periodic time instants and may imply FMN topology changes. In order to avoid disruptions in the communications paths, this thesis proposes the predictive Routing for Flying Multi-hop Networks (RedeFINE) mechanism. Knowing the initial and final positions of the FAPs and FGW, RedeFINE calculates their future trajectories and determines the forwarding tables of the FAPs and FGW as well as the time instants they shall be updated. RedeFINE enables a seamless FMN reconfiguration by selecting high-capacity communications paths between UAVs, while simultaneously avoiding path disruptions. This allows for improved Quality of Service (QoS) offered to the ground devices whilst minimizing the disruption caused by the movement of the UAVs. Complementarily, this thesis proposes a novel Inter-flow Interference-aware Routing (I2R) metric, tailored for centralized routing in FMNs.

This chapter starts with the related work on routing paradigms, protocols, and metrics for wireless networks, with a special focus on predictive routing solutions for FMNs. After that, we describe the system model and formulate the problem. Then, we explain the RedeFINE concept and its performance evaluation, including the simulation setup, the networking scenarios employed, the performance metrics considered, and the simulation results obtained. Following that, we introduce the I2R metric, including its motivation and rationale, and the performance evalua-

tion when RedeFINE uses the I2R metric. Finally, we summarize the contents presented in this chapter.

5.1 Related Work

FMNs impose significant routing challenges, especially when the UAVs move for repositioning themselves and better serve the ground devices. On the one hand, due to the variations in the Signal-to-Noise Ratio (SNR) of the wireless links when the UAVs change their positions, disruptions in the communications paths may occur. On the other hand, there is inter-flow interference between neighboring UAVs that compete for the same shared wireless channel. In this section, we start by presenting the main routing paradigms used in flying networks. Then, we review the state of the art on predictive routing solutions for flying networks. Finally, existing routing metrics for wireless networks are presented.

5.1.1 Routing Paradigms for Flying Multi-hop Networks

Existing routing solutions for FMNs are based on the protocols initially proposed for Mobile Ad hoc Networks and Vehicular Ad hoc Networks. They may be divided into two main paradigms: single-hop routing and multi-hop routing. Multiple solutions tailored for FMNs have been proposed in the literature [142–151]. In this section, we give special emphasis to multi-hop routing.

In single-hop routing, a static pre-defined routing table is configured in the communications nodes and used to perform the routing decisions. Within the context of flying networks, single-hop routing typically considers UAVs that load data from ground devices and carry it to the destination by flying, as depicted in Fig. 5.1. This routing paradigm is best suited to transport delay-tolerant data. Since fault detection of invalid routes is typically not available, single-hop routing protocols are mainly suitable for networks with static topology, including fixed ground devices.

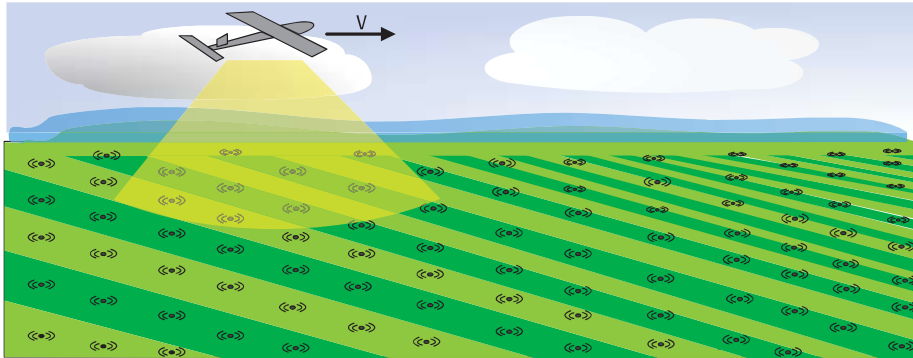


Figure 5.1: Illustrative networking scenario of single-hop routing, where a UAV collects data from ground nodes and carries it to the destination by flying [22].

In multi-hop routing, the packets are forwarded hop by hop, through communications paths consisting of multiple communications nodes, as depicted in Fig. 5.2. Multi-hop routing solutions include topology-based and position-based routing protocols. Topology-based routing protocols use metrics such as hop-count and the quality of the wireless links to perform the routing decisions; they are classified as proactive, reactive, and hybrid [145]. Proactive routing protocols (e.g., Optimized Link State Routing (OLSR) [152]) aim at ensuring that the routing tables of all communications nodes are always up-to-date; with this the communications paths can be defined in advance and the time to forward packets is reduced. This is typically achieved by means of an update mechanism that uses the broadcast of *HELLO* packets for neighbor discovery and the flooding of topology control packets for route announcement. However, proactive protocols consume significant radio resources and react slowly to network topology changes; hence, they are not suitable for highly dynamic networks, such as FMNs [142, 143]. Conversely, reactive routing protocols (e.g., Ad hoc On-Demand Distance Vector (AODV) [153]) follow an on-demand route establishment approach, with suitable communications paths only defined when packets need to be sent. Reactive routing protocols aim at addressing the overhead problem introduced by the proactive routing protocols; the price paid is the latency inherent to the routing discovery process. Therefore, reactive routing protocols are best suited for highly dynamic networks with low traffic demand [143, 145]. Finally, hybrid routing protocols (e.g., Zone Routing Protocol [154]) consider a combination of the proactive and reactive routing approaches, which aim at reducing the overhead of the proactive routing protocols and the delay introduced by the reactive routing protocols. Hybrid protocols are especially suitable for large networks, which may be divided into a set of zones, where intra-zone routing may be done using a proactive routing protocol, while inter-zone routing may be done using a reactive routing protocol. The position-based routing protocols use geographical information, typically determined in real-time, to perform the routing decisions. Obtaining the position of the UAVs in real time with high accuracy (e.g., using Global Positioning System (GPS)-based mechanisms) is challenging, especially due to the low payload capacity and energy available on-board the UAVs, which imply the usage of computing units with limited processing capabilities. A routing approach that takes advantage of the information about the trajectory or positions defined in advance for the UAVs is worth considering.

5.1.2 Predictive Routing Solutions for Flying Multi-hop Networks

In this section, we review the related work on predictive routing solutions for FMNs. Concerning the predictive routing approach, many solutions proposed in the literature rely on knowing the positions of the UAVs over time, which are inferred based on their speed, moving direction, and suitable mobility models [155–159]. Predictive routing solutions aim at addressing the fast network topology changes of highly dynamic networks, which is a drawback of the proactive routing

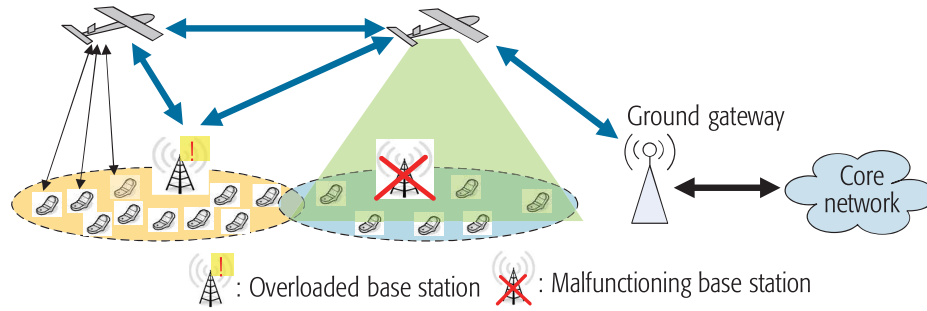


Figure 5.2: Illustrative networking scenario of multi-hop routing, where two UAVs forward data between multiple ground devices and a ground gateway [22].

protocols, while improving the routing discovery process of the reactive routing protocols. However, the implementation of solutions that use position-based information in real-time implies strict hardware requirements, in order to determine the location of the UAVs with high accuracy. In what follows, we describe each of the relevant solutions found in the state of the art, grouped by heuristic and Machine Learning (ML)-based prediction approaches. At the end of each section, we summarize the main drawbacks associated with each group of solutions.

5.1.2.1 Heuristic-based Prediction

In [160], the Link Stability Estimation-based Preemptive Routing (LEPR) protocol has been proposed. It is based on the AODV protocol. The concept of the LEPR protocol lies in computing different stable link-disjoint paths using a novel link stability metric, in order to determine alternative reliable paths able to replace failed links. It uses GPS information and employs mobility prediction models to estimate the stability of the wireless links and minimize the delay of the routing discovery process, typical in reactive routing protocols. LEPR is targeted for flying networks where the movement of the UAVs induce frequent link failures and may compromise the wireless connectivity. The link stability metric considers the past, present, and future status of the wireless links, inferred based on the link quality, the distance between the communications nodes, and a mobility prediction factor, respectively. LEPR consists of two processes: reactive route discovery and semi-proactive route maintenance. During the route discovery process, the source node selects multiple link-disjoint paths, in which no common links are shared. Based on the link stability metric, the source node selects the most reliable path for data forwarding. In order to provide awareness to the semi-proactive route maintenance, constant monitoring of the stability of the path in use is carried out, in order to reduce the possibility of a connectivity disruption due to a link failure. Every communications node that is part of the communications paths in use is able to send Route Switch over messages, in order to inform the source node to prepare

for a potential disconnection. In such case, LEPR determines if an alternative path with a greater quality exists; a rerouting operation is initiated otherwise. Network Simulator 3 (ns-3) simulation results have shown the improved performance of LEPR when compared with AODV and Dynamic Source Routing, especially in terms of Packet Delivery Ratio (PDR), delay, and control overhead for high and low mobility scenarios.

A similar approach, with a strong implementation-oriented and experimental component, has been presented in [23]. Therein, the Predictive Optimized Link State Routing (P-OLSR) protocol has been proposed to improve the performance of the OLSR protocol in flying networks. In order to predict topology changes and react before link disruptions occur, P-OLSR uses GPS information available in the UAVs and improves the Expected Transmission Count (ETX) routing metric to consider the direction and relative speed of the UAVs. The modified structure of the OLSR's *HELLO* message employed by the P-OLSR protocol is depicted in Fig. 5.3. Comparative field experiments between P-OLSR and OLSR have been conducted, considering two fixed-wing UAVs and a ground station. Through experimental results the authors have concluded that P-OLSR is able to quickly react to network topology changes and offer reliable multi-hop communications paths in scenarios where OLSR typically fails. To evaluate the performance of P-OLSR in large networks, Medium Access Control (MAC) layer-based emulations have been considered. Yet, the overhead inherent to the proactive routing protocols has not been addressed by P-OLSR.

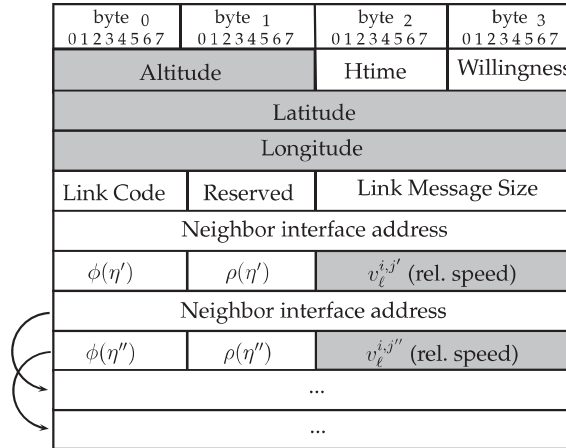


Figure 5.3: Modified structure of the OLSR's *HELLO* message. Information fields added or modified in the P-OLSR protocol implementation are highlighted in gray, including altitude, latitude, longitude, and the average relative speed associated with the communications nodes [23].

In [25], the authors have proposed extending the Software-Defined Networking (SDN) paradigm to include a joint time-spatial component, as part of a Google's SDN operating system. The authors have envisioned that a UAV-based mesh and Internet backhaul network may be associated with disruptive, albeit predictable, events, such as changes in the positions of the UAVs over time.

They have proposed the use of centralized intelligence to orchestrate the network state in advance by means of a Temporospatial-SDN approach. Temporospatial-SDN takes into account the ability provided by SDN controllers to predict the time-varying characteristics of wireless links and communications nodes, allowing network reconfiguration to be scheduled rather than made in a reactive manner, in order to avoid link disruptions. The SDN concept applied to a generic wireless networking scenario is depicted in Fig. 5.4. Using third-party software, the SDN controllers under the Temporospatial-SDN approach are able to predict the future low level network status, based on the physical positions of the UAVs and their antennas, while considering a holistic view over the network. The performance evaluation carried out through ns-3 simulations has shown that the Temporospatial-SDN approach can completely avoid connectivity disruptions by predicting changes in wireless links and proactively rerouting the network flows. However, the authors of Temporospatial-SDN have not considered any QoS requirement to reconfigure the network and perform traffic forwarding.

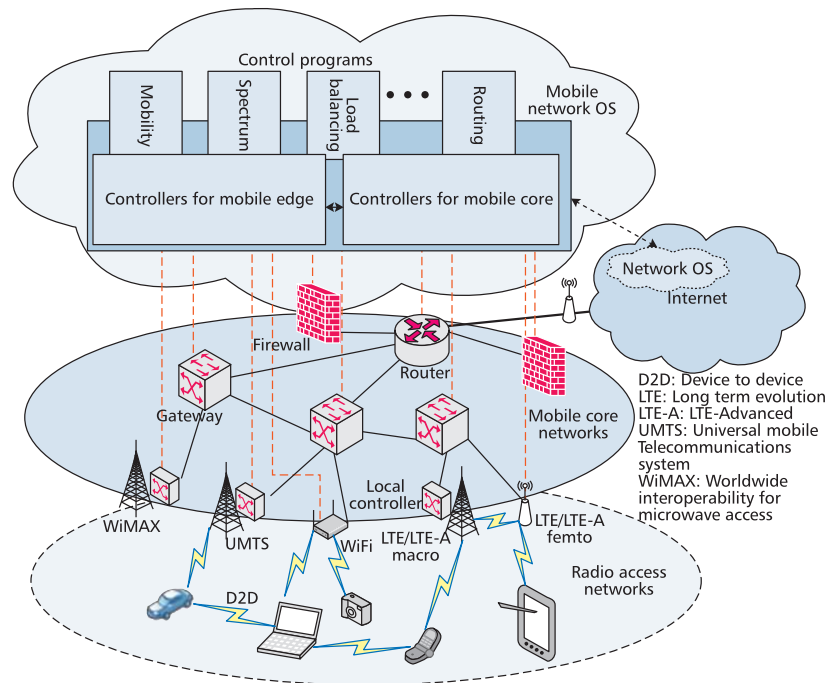


Figure 5.4: Illustration of SDN applied to a generic mobile networking scenario [24]. Considering an SDN-based architecture, the Temporospatial-SDN approach proposed in [25] takes advantage of the holistic view available in the SDN controllers to predict the time-varying network's conditions.

In [26], a centralized routing algorithm for flying networks, named Directional Particle Swarming Optimization, has been proposed. Directional Particle Swarming Optimization is aided by a Swarm Intelligence-enabled SDN controller, which is in charge of mobility estimation and prediction. Extended Kalman Filter and Particle Swarm Optimization (PSO) based approaches are

used to overcome the high mobility and improve the QoS in flying networks. The Extended Kalman Filter is employed to estimate and predict the UAVs' position, whose results are used by the centralized SDN controller or Cluster Head Unmanned Aerial Vehicle (CHU) to assess the wireless links' quality by building a Routing Estimated Cost function. PSO is employed to overcome the non-deterministic polynomial-time hard (NP-hard) nature of the routing problem. The transmission Routing Estimated Cost associated with each UAV's request is able to address frequent network topology changes. The SDN controller or CHUs compute optimal routing paths for CHUs using global network information. While the source and destination UAVs are both within the coverage area of the same CHU, additional independent routing decisions are performed by the CHUs, in order to reduce the overall Routing Estimated Cost. Considering Fig. 5.5, when a routing request is received by a Swarm Intelligence-enabled CHU, it generates the initial particles. If the generated particles do not meet the termination conditions defined, the CHU continues determining each particle's fitness function value, neighbor optimal particles, and history optimal particles, while updating each particle. Besides the CHUs, the Swarm Intelligence-enabled Centralized Controller is able to define the shortest-path for enabling communications between CHUs by means of the Directional Particle Swarming Optimization algorithm. In order to generate particles, determine fitness values, and update each particle, the EKF-based mobility estimation and prediction module is employed. The generating and updating processes benefit from the path constraint imposed by a forwarding zone. When the fitness value satisfies the termination conditions defined, such as a maximum Routing Estimated Cost for each hop, the routing decisions are taken. Simulation results show that Directional Particle Swarming Optimization outperforms the state of the art counterparts, including Geographic Position Mobility Oriented Routing [161] and Adaptive Forwarding Protocol [162], in terms of throughput, delay, and PDR. Geographic Position Mobility Oriented Routing uses a Gauss-Markov mobility model to predict the movement of the UAVs, while Adaptive Forwarding Protocol lies in an adaptive forwarding scheme based on forwarding probability and forwarding zone criterion; due to the enhanced performance of Directional Particle Swarming Optimization, Geographic Position Mobility Oriented Routing and Adaptive Forwarding Protocol are not detailed herein.

The Mobility Management Unmanned Aerial Vehicle-based Grouping (MMUG) routing protocol has been proposed in [27] for flying networks serving biomedical applications, including human body sensors used to monitor vital signs. The flying network is assumed to support inter-UAV communications as well as communications between ground stations and UAVs, as depicted in Fig. 5.6. In the proposed MMUG routing protocol, groups of UAVs are formed with a group head being assigned; the node closest to the ground station is called the group head. All UAVs must be connected to a group header. If any UAV is outside the wireless range of any group head, it forms a new group and connects to the nearest UAV. The MMUG routing protocol is used to forward data traffic between the flying network and remote Base Stations (BSs). The routing met-

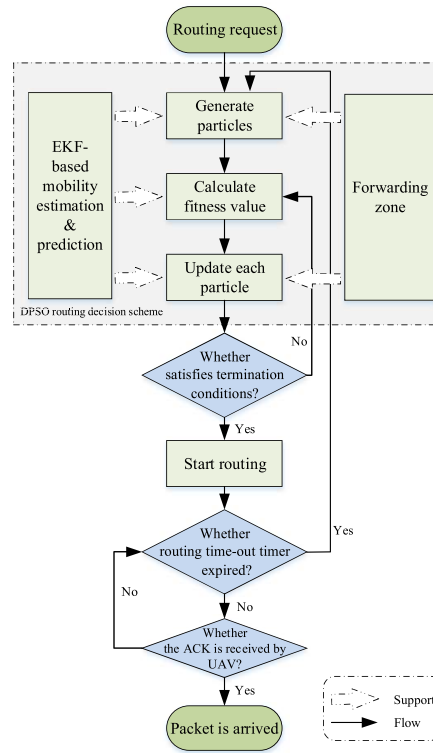


Figure 5.5: Routing algorithm based on Directional Particle Swarming Optimization (DPSO) [26].

ric proposed for choosing optimal communications paths is able to consider minimum bandwidth, minimum delay, and minimum path length. Simulation results obtained in Network Simulator 2.35 show that the MMUG routing protocol allows for lower resource usage, lower cost, and greater network efficiency, while reducing control and routing overhead when compared with the state of the art LEPR protocol [163] and Energy-Efficient AODV routing protocol.

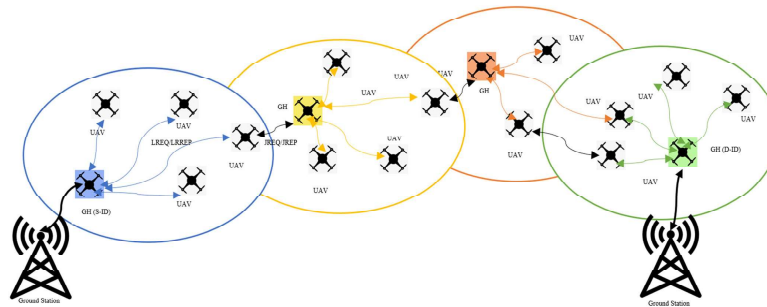


Figure 5.6: Wireless communications paths defined by the MMUG routing protocol. Different UAV clusters are circled by the blue, yellow, orange, and green solid lines. The blue UAV cluster includes the source node, which is referred to as the Group Head Source IDentifier (GH-S-ID), whereas the green UAV cluster includes the Group Head Destination IDentifier (GH-D-ID) [27].

In [28], a routing scheme based on fuzzy logic has been proposed for flying networks. The proposed routing scheme considers multiple parameters, such as link quality, node stability, and node energy, in order to create stable communications paths and improve the PDR. In addition, it aims at increasing the network's lifetime by balancing the energy consumption of the nodes. Two phases are considered: 1) the route discovery phase and 2) the route maintenance phase. The route discovery phase is triggered when the source node has data packets to send to a destination node and its forwarding table does not have a suitable entry. For that purpose, the source node broadcasts a Route Request message to the neighboring nodes. The neighboring nodes calculate their scores based on different parameters, such as movement direction, residual energy, link quality, and node stability. Then, they can rebroadcast the Route Request message if they obtain the required score to be suitable to forward traffic; this score-based approach aims at reducing the overhead associated with the broadcast of Route Request messages. Moreover, the authors have designed a fuzzy system in the route discovery phase, in order to choose communications paths with greater robustness, less delay, and fewer hops. The process is depicted in Fig. 5.7. The route maintenance phase consists of two steps: preventing communications paths disruption and recovering disrupted communications paths. Network Simulator 2.35 was used to evaluate the performance of the proposed routing scheme. The obtained results have demonstrated that the proposed routing scheme allows for improved performance in terms of delay, PDR, stability of the communications paths, and energy consumption, at the expense of a higher routing overhead, when compared with the Energy-efficient Connectivity-aware Data Delivery [164], LEPR, and AODV routing protocols. Energy-efficient Connectivity-aware Data Delivery and LEPR are inspired by the AODV routing protocol. When it comes to flying networks, LEPR neglects delay, UAV power-constraints, node stability, and the number of hops in the route discovery process. On the other hand, Energy-efficient Connectivity-aware Data Delivery does not consider the movement direction of the UAVs in the route selection process, and the link quality and node stability aspects in the route discovery process.

In order to ensure a stable flying network, it is necessary to maintain strong relationships between neighboring UAVs by managing their relative distance, resulting direction, and velocity. To achieve that, a Joint Topology Control and Routing protocol consisting of three modules has been proposed in [29], as depicted in Fig. 5.8. The first module is named Virtual Force-based Mobility Control (VFMC). VFMC employs two distinct virtual forces: the Mobile Ground Target-Discovery Force, which aims at maximizing the coverage provided to Mobile Ground Targets, and the adaptively weighted Topology Formation Force in order to ensure a targeted Signal-to-Interference plus Noise Ratio level for the inter-UAV wireless links. By employing virtual forces at each time slot, each UAV is able to estimate a Net Virtual Force able to determine the optimal mobility information. The VFMC module optimizes the interval between *HELLO* packets to ensure quick reaction to network topology changes and minimize the control overhead. The

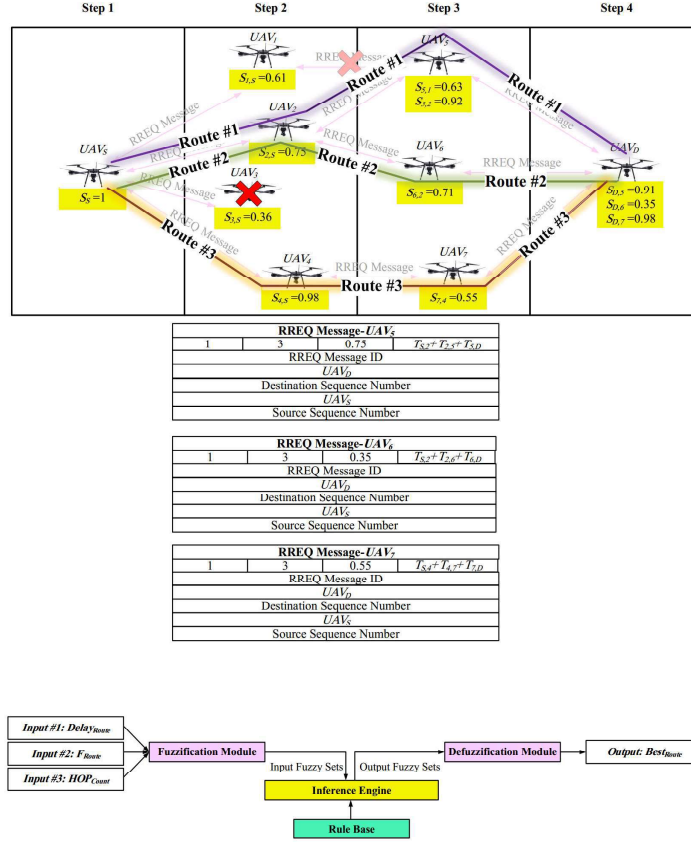


Figure 5.7: Routing discovery phase of the routing scheme based on fuzzy logic proposed in [28], where: 1) route delay ($Delay_{Route}$), initially set to zero, 2) route fitness (F_{Route}), equal to the lowest score among all UAVs in the route, and 3) $HopCount$, starting at zero, are considered as inputs to determine the $Best_{Route}$; they are updated at each hop. $Best_{Route}$ is the communications path with the highest fitness, the least number of hops, and the minimum delay.

second module is named Energy-efficient Mobility-aware Fuzzy Clustering, which employs the mobility information supplied by the VFMC module to partition the topology into different clusters for data traffic aggregation. This concept minimizes the number of agents by aggregating data packets at each CHU, enabling fewer transmissions and less MAC contention, while providing higher PDR. The Topology-aware Q-Routing is the third module of the Joint Topology Control and Routing protocol. Topology-aware Q-Routing lies in multi-objective optimization for inter-cluster position-based multi-hop routing. It enables that source CHUs transmit Aggregated Data Packets to BS by choosing the communications path that minimizes congestion, energy holes, loops, and link failures. Q-learning has been employed to allow determining an optimal communications path. Also, a state exploration and exploitation technique, based on the interaction between the average Neighbor Intimacy, Packet Travel Time, Packet Travel Speed, and link lifetime, has been

designed. The proposed technique balances exploration with exploitation, in order to prevent local optima. It also allows for improved convergence and decreased amount of retransmissions. Finally, one-hop delay, path stability defined by Neighbor Intimacy, and the Residual Energy of UAVs have been considered to define a multi-objective reward function. Through *MATLAB*'s simulations the authors have concluded that the developed reward function permits a higher average reward than *Q*-learning-based Topology-aware Routing (QTAR) [35], Mobility Control-based Clustering [165], and the Mission-Critical Flying Ad Hoc Network Operation algorithm [166]. Moreover, the Topology-aware *Q*-Routing module is able to avoid routing holes, routing loops, failure-state, and wireless link failures. The state of the art counterparts are not described herein since they follow similar approaches as Joint Topology Control and Routing (JTCR); they have been selected by its authors to ensure a fair performance assessment.

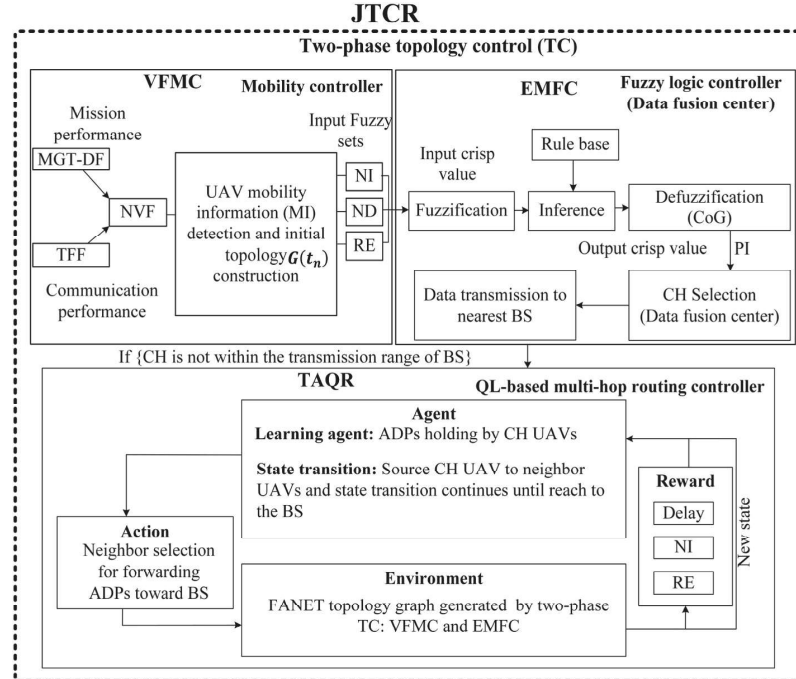


Figure 5.8: Architecture of the two-phase topology control and the *Q*-learning based Topology-aware *Q*-Routing (TAQR) module of the Joint Topology Control and Routing (JTCR) protocol proposed in [29].

In [167], the OPAR solution for cooperative flying networks has been proposed. Besides the path length, OPAR takes into account the path lifetime, in order to reduce the overhead introduced for updating the forwarding tables. The authors have concluded that considering the path lifetime allows for improved throughput, flow completion time, and flow success rate. The flow completion time represents the interval between the transmission of the first packet and the delivery of the

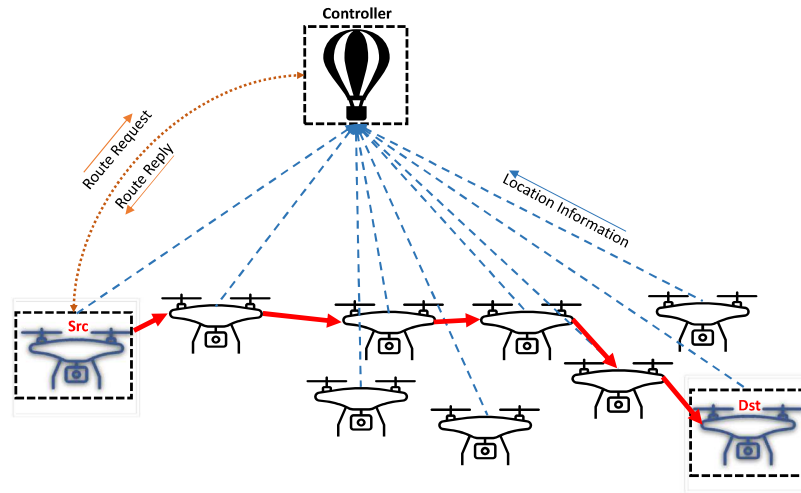


Figure 5.9: SDN-based flying network considered by the Optimized Predictive and Adaptive Routing (OPAR) and Load-Balanced Optimized Predictive and Adaptive Routing (LB-OPAR) algorithms. Messages are periodically exchanged between the controller and the UAVs. Based on three consecutive positions received from each UAV, the controller predicts the link lifetime, calculates the traffic load for each node, and solves an optimization problem to define the communications paths [30].

last packet for each flow, while the flow success rate is the portion of flows successfully delivered to the destination. Moreover, they have concluded that considering the path lifetime and the path length can lead to traffic load-balancing issues. As such, for a shared wireless channel, this problem may result in decreased throughput and high flow completion time, particularly in saturated networks. In order to address the problem, the same authors have proposed in [30] the successor of OPAR – LB-OPAR. Besides the path lifetime and the path length, the LB-OPAR algorithm considers the network load to select the communications paths. Linear Programming optimization has been considered to formulate the routing problem in an SDN based flying network. The formulated optimization problem jointly considers path length, path lifetime, and network load. The path lifetime is defined as the sum of the lifetimes associated with all links that integrate a communications path; it is affected by the mobility of the communications nodes. A link lifetime prediction algorithm has been proposed to provide awareness to LB-OPAR. In turn, the path load is defined as the maximum load for all nodes in the communications path. The load of each node is defined as the number of neighboring nodes that generate or forward data packets. A network controller is assumed to know the location of the communications nodes, as depicted in Fig. 5.9. The optimization problem consists of binary variables. Binary Linear Programming problems are non-deterministic polynomial-time complete (NP-complete). However, the authors of LB-OPAR have demonstrated that the proposed algorithm is able to determine a solution for the proposed problem. The network performance when OPAR and LB-OPAR are used has been

carried out through ns-3 simulations. LB-OPAR allows minimizing the path length and the path load, while maximizing the path lifetime. However, this may not allow maximizing the QoS levels offered in practice, since the wireless links may not provide a high enough network capacity to accommodate the offered traffic.

Despite the multiple routing solutions tailored for FMNs proposed in the literature, they typically employ position-aware and topology-aware distributed approaches, similar to the ones designed for Mobile Ad hoc Networks and Vehicular Ad hoc Networks [168–170]. They typically impose a large overhead for constantly updating the global network information based on changes of the flying network determined in real-time or require a time-consuming convergence to define the forwarding tables, especially in large networks. Routing protocols suitable for flying networks should have low complexity and low overhead. Yet, the challenges inherent to FMNs, including quick degradation of the SNR of the wireless links and frequent network topology changes, caused by moving UAVs, need to be addressed. Routing solutions for FMNs should consider a trade-off between the network overhead and the delay introduced for updating the forwarding tables, as well as the capability to react to the FMNs' highly dynamic nature.

5.1.2.2 Machine Learning-based Prediction

A 3-Dimensional (3D) *Q*-learning routing protocol for flying networks has been proposed in [31], in order to allow for PDR guarantees and improved QoS. In the 3D *Q*-learning routing protocol, Hop-By-Hop and Store-Carry-Forward routing approaches are combined using a *Q*-learning based routing decision technique, as depicted in Fig. 5.10. The quality of the wireless links is characterized using a prediction module, while a *Q*-learning based module is employed to compute the routing decisions in an adaptive manner. The goal is to maximize the total expected rewards for each UAV. Multi-hop forwarding is a component of Hop-By-Hop routing, for which each UAV is responsible for its own Hop-By-Hop decision. Each UAV triggers a routing request when a packet needs to be sent; the routing decision module then replies with a routing decision. The location information, packet arrival information, and packet transmission information are considered as inputs to the 3D *Q*-learning routing protocol. Considering this information, each UAV computes the Degree Towards the Ground Station and Estimated Next Packet Arrival Time values, where a Packet Arrival Prediction algorithm is employed to estimate the packet arrival interval for the neighboring UAVs, namely whether it is small or large. The UAVs decide the routing approach to be used whenever they need to send a packet: if it is Store-Carry-Forward, the UAVs store and carry the packets by moving towards the ground station; otherwise, the UAVs forward the packets to neighboring UAVs. Each UAV calculates a reward when a data packet is delivered to the ground station or reaches a predefined expiration time. The 3D *Q*-learning routing protocol is assumed to run at every UAV, enabling a distributed yet cooperative routing paradigm.

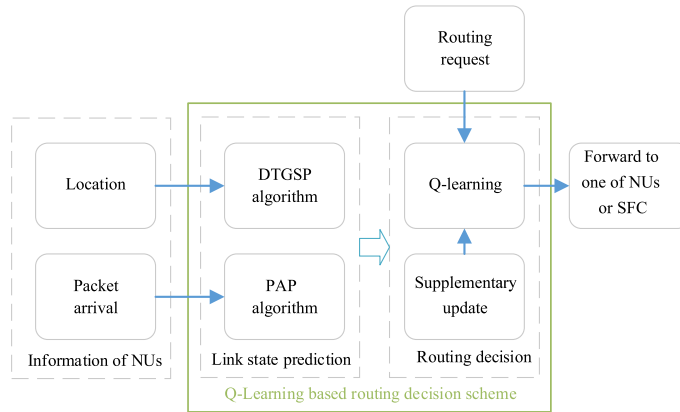
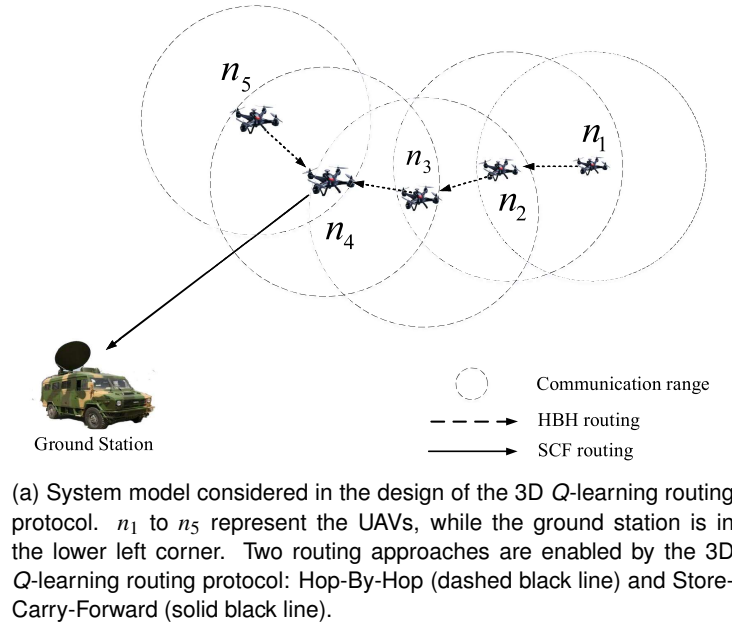


Figure 5.10: 3D Q -learning routing protocol, which combines Hop-By-Hop and Store-Carry-Forward approaches. The routing decisions are taken by a Q -learning based routing decision scheme, consisting of link quality prediction and routing decision modules [31].

In [32], an enhanced Q -learning routing algorithm based on trajectory prediction, named Pre Q -Learning, has been proposed. Considering that routing in multi-hop wireless networks can be reduced to a Markov Decision Process, as it depends on the status of each node and the choice of the next hop, Reinforcement Learning based on Markov Decision Process, namely Q -Learning, has been employed to perform routing of data packets in flying networks. Q -learning lies in choosing the action with the highest Q -value, considering the environment's reward and the current state. In the proposed Pre Q -Learning routing algorithm, routing decisions are based

on a Q -table built for each pair of states and actions. PreQ-Learning improves the convergence of the Q -learning algorithm by introducing a mechanism for computing the Q -value in advance, based on trajectory prediction. In order to identify whether two nodes may become neighbors, the employed trajectory prediction approach considers a Kalman Filter. When a node in a dynamic network topology has a new neighbor node, it updates the destination's Q -value using the Q -learning method. It has been demonstrated that the Kalman Filter-based algorithm is able to accurately predict trajectories. The PreQ-Learning routing algorithm is based on the proactive Q -learning Optimized Link State Routing protocol, which aims at addressing the high mobility and communications requirements in flying networks. The Q -learning Optimized Link State Routing protocol maintains a Q -table with the Q -value associated with each node; this allows building a forwarding table starting from the next hop with the highest Q -value. When using the PreQ-Learning routing algorithm, each node keeps both a preset Q -table and a routing Q -table, as depicted in Fig. 5.11. Three processes are involved in the computation of the routing Q -table. First, each node determines if the next hop predicted in advance is within its communications range. If so, the computed preset Q -value is used as the default value for the associated routing Q -value. After setting the default routing Q -value, the node computes the routing Q -value from its neighbors to the destination. The previously used default Q -value is removed from the default preset Q -table. In contrast to the Q -learning routing method, the iteration of the Q -value for newly connected nodes is carried out in advance, which allows using an initial routing Q -value and converge faster to the optimal routing choice. As such, a suitable communications path can be chosen. Finally, the default preset Q -table is replaced according to the predicted results. The performance evaluation of the PreQ-Learning routing algorithm has been carried out by means of ns-3 simulations against QL-OLSR and OLSR, showing that PreQ-Learning is able to achieve improved network performance in terms PDR and average delay.

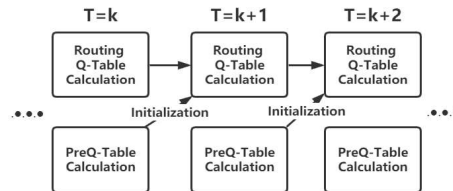


Figure 5.11: Forwarding table calculation made by the PreQ-Learning routing algorithm [32].

In [171], a model-free Reinforcement Learning-based geographical routing protocol has been presented. A Markov Decision Process has been considered to simplify the modeling of the routing decision process as a Reinforcement Learning problem. The proposed protocol consists of four phases: sensing, feature extraction, learning, and routing decision. Fig. 5.12 depicts the routing protocol's architecture. The purpose of the sensing phase is to collect features. Deep

Reinforcement Learning is used for training the nodes. During the sensing phase, each UAV broadcasts its own location through periodic *HELLO* packets, while maintaining a neighbor table. The neighbor table contains the stability of the wireless links available between the node and its neighbors for a given time interval, the expected residual energy of the neighbors, and the geographical positions of the neighbors. Additional information, namely the source node, the last-hop node, the destination node and its position, and the set of nodes that have been traversed, is included in the data packet during the packet-forwarding process; this information aims at creating feature values and prevent the creation of loops. A Deep Reinforcement Learning-based Geographical Routing protocol has been designed based the proposed Reinforcement Learning model. It is run whenever a node receives a packet, in which the routing decision lies in choosing the neighbor node to forward the packet to, while the state is moved from the current node to the next node; the process is repeated until the packet delivery fails or succeeds. Numerical simulation results have demonstrated that the Deep Reinforcement Learning-based Geographical Routing protocol ensures improved robustness and reliability, making it suitable for dynamic, energy-limited flying networks.

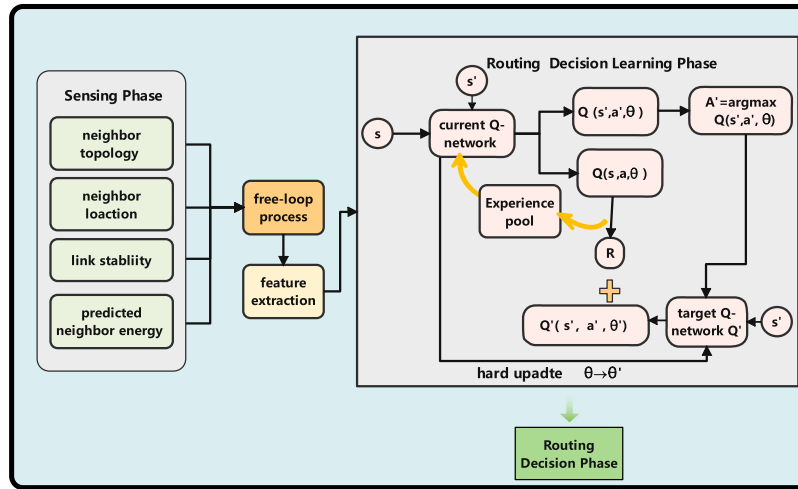


Figure 5.12: Deep Reinforcement Learning-based Geographical Routing protocol's architecture, depicting the sensing, learning, and routing decision phases [31].

Due to the dynamic topology and traffic demand, transmission buffer overflow and packet loss are a challenge in flying networks. Traffic prediction may be considered to solve this issue. In [33], a Packet Arrival Prediction routing protocol for flying networks has been proposed. It employs an ML method and a joint decision approach to dynamically perform routing decisions and improve the reliability of wireless communications paths. The authors have designed a Packet Arrival Prediction module based on Long Short-Term Memory to predict the data traffic in flying networks. They have used a Routing Decision Factor to jointly assess each wireless link, while taking into

account the mobility inherent to flying networks. In order to determine the communications paths, the routing decision problem has been formulated as an optimization problem. The authors have assumed that the packet arrival rate can be predicted and is location-aware, taking into account that: 1) changes in the packet arrival rate are typically raised by events, such as natural disasters and crowd's movements; 2) events are typically regional, associated with the areas where the event occurs; and 3) changes in the packet arrival rate are typically gradual in time and not instantaneous. The red area in Fig. 5.13 depicts a disaster area where infrastructures have been destroyed. The UAVs are used to provide wireless connectivity on-demand. The packet arrival rate at the UAVs is assumed to increase gradually as they approach the disaster area, where a higher traffic demand is expected. A Recurrent Neural Network has been employed to consider different inputs for long periods of time by applying the Long Short-Term Memory model. In order to reach a joint decision and high-performance solution, the routing decision takes into account the predicted packet arrival rate, the channel quality, and both the existing and future queue backlogs. Simulation results have demonstrated that the proposed Packet Arrival Prediction routing protocol outperforms the state of the art counterparts regarding PDR and delay.

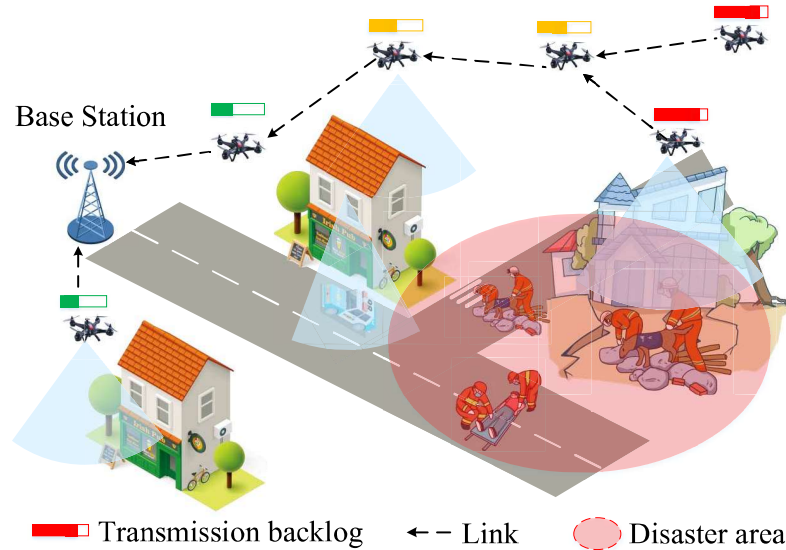


Figure 5.13: Illustration of a disaster management scenario considered by the Packet Arrival Prediction routing protocol, where the red area has been subjected to an earthquake. UAVs are used to extend wireless connectivity from a remote BS. The transmission queue backlog is shown by the progressive bars above each UAV. Red denotes a long queue backlog, while yellow and green denote medium and small queue backlogs, respectively. Forwarding data traffic through a UAV with a red bar may lead to high Packet Loss Ratio (PLR) [33].

Predictive Ad-hoc Routing by Reinforcement Learning and Trajectory knowledge (PARRoT) has been proposed in [172]. PARRoT is an ML-enabled routing protocol that employs a cross-

layer strategy, while leveraging on knowledge about the relative mobility of different agents across a specified prediction window. For that purpose, each agent calculates its future position depending on its current position. The result is shared with the remaining nodes. User Datagram Protocol (UDP) routing messages are used by PARRoT in order to adapt to dynamic network topologies. Periodically, each PARRoT node generates messages based on a predetermined interval, which are forwarded throughout the network. Similarly to decentralized routing protocols like Better Approach To Mobile Ad hoc Networking (B.A.T.M.A.N.) and unlike path planning-based protocols such as OLSR, each node has only a partial view of the entire network. In order to forward messages to a given destination, each node considers only its next-hop neighbors. The implicit maintenance of several candidate paths for each destination allows PARRoT to rapidly recover after link failures occur. Objective Modular Network Testbed in C++ (OMNET++) simulation results have demonstrated a greater packet delivery ratio and a shorter delay achieved by PARRoT in comparison to AODV, OLSR, Greedy Perimeter Stateless Routing [173], and B.A.T.Mobile [174] state of the art routing protocols.

A Deep Neural Network Routing-based approach has been proposed in [34]. It involves routing discovery and data transmission phases, as depicted in Fig. 5.14. In routing discovery, the BS and the UAVs broadcast their positions and speeds, estimate their positions for the next-time interval, and find whether the UAVs are inside or outside a cylinder, which is used to represent the 3D space. UAVs jointly forward data packets during data transmission using either Deep Neural Network forwarding or greedy forwarding. If the UAVs are outside the cylinder, they use DNN to select the next-hop node; otherwise, they employ greedy forwarding. In real-world flying networks, data packets typically are characterized by a Time-To-Live parameter. After the defined Time-To-Live expires, data packets are no longer valid. Therefore, a Time-To-Live cache approach and retransmissions are also employed. Each data packet in Deep Neural Network Routing includes a timestamp. If data packets exceed the defined Time-To-Live, they are removed from the cache. As such, if any UAV does not get a reply for a packet it has transmitted, the UAV will retransmit the data packet once to prevent exceeding the Time-To-Live. ns-3 simulations have shown that Deep Neural Network Routing allows for improved performance in terms of PDR and energy consumption, when compared with the state of the art counterparts P-OLSR [23], extended random AB routing algorithm in 3D space [175], Fuzzy logic-based Reliable Routing Protocol [176], and 3D Position Based Adaptive Real-time routing Protocol (PBARP3D) [177]. P-OLSR is an extended version of OLSR, AB routing algorithm in 3D space, and PBARP3D are representatives of 3D routing protocols, while the Fuzzy logic-based Reliable Routing Protocol is an ML-based routing protocol; they have been selected by the Deep Neural Network Routing solution's authors to enable a fair performance assessment.

In [35], the QTAR protocol for flying networks has been proposed. The concept is depicted in Fig. 5.15. QTAR is aided by UAV position information, which can be obtained by means of

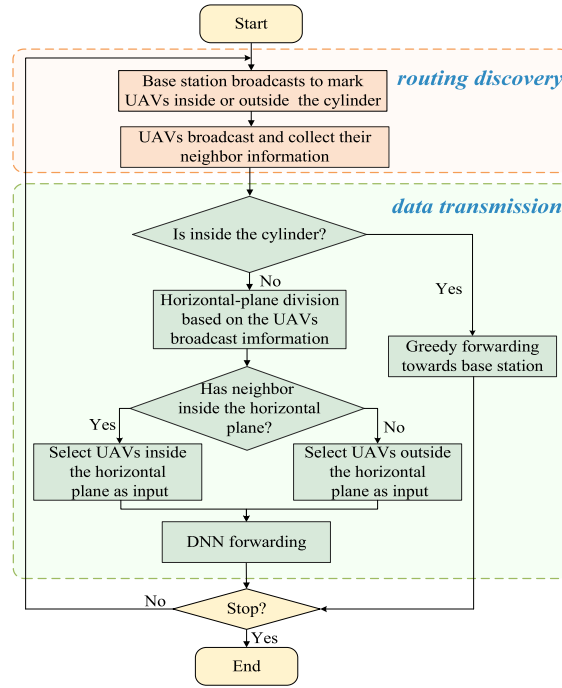


Figure 5.14: Deep Neural Network Routing (DNNR) solution proposed in [34].

GPS. Two-hop neighbor information is used to characterize the environment and nodes' states for topology control purposes, including node position, delay, velocity, and residual energy level. The two-hop based geographic routing approach allows for improved routing efficiency. QTAR dynamically adjusts the Link Duration, *HELLO* packets' period, and network topology, considerably reducing the communications delay. One of the main advantages of the two-hop neighbor information is to reduce the number of hops between the source and the destination, while aiding in re-estimating neighbor relations for routing decisions and selecting the most reliable next-hop node. The proposed adaptive *Q*-learning method enables learning from dynamic network environment, while dynamically adjusting the learning rate and the reward. Finally, QTAR promotes load balance by taking into account the node energy level of the communications nodes in *Q*-learning. For that purpose, it selects alternative paths to balance the network's energy consumption. *MATLAB* simulations have shown that QTAR outperforms the state of the art counterparts following similar position-based routing approaches, namely QGeo [178] and Greedy Perimeter Stateless Routing [173], in terms of PDR, delay, energy consumption, and network lifetime, despite having a greater overhead.

The Topology-Aware Resilient Routing based on Adaptive *Q*-learning (TARRAQ) protocol has been proposed in [36]. First, a queuing service framework for assessing link establishment and disconnection between the UAVs has been designed. In addition, closed-form expressions for

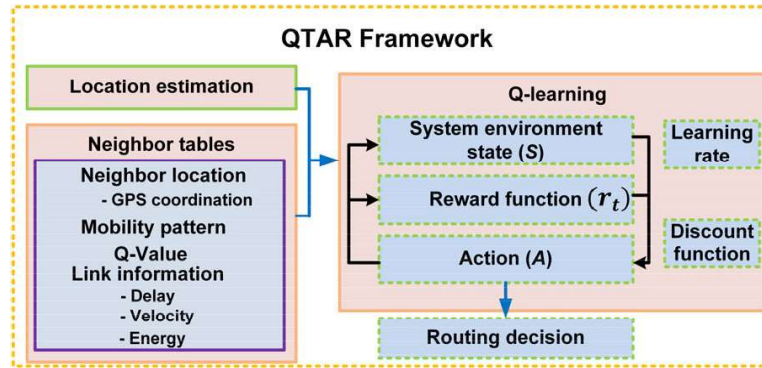


Figure 5.15: QTAR proposed in [35]. QTAR includes modules for obtaining and estimating the communications nodes' location and for neighbor discovery, which allow building neighbor tables. Q-learning is used in the route discovery process, considering the two-hop neighbor information gathered.

the Link Duration, Neighbors' Change Rate, and Neighbors' change Inter-arrival Time have been obtained. TARRAQ takes into account three phases: 1) neighbor discovery, 2) neighbor maintenance, and 3) relay selection. In the first two phases, the residual Link Duration, estimated by a Kalman Filter, serves as an Expiration Timer for available neighbors, while the Neighbors' Change Rate and Neighbors' change Inter-arrival Time are considered to compute the sensing interval. All UAVs are capable of dynamically adjusting their sensing scheme based on dynamic performance requirements; as a result, reliable information regarding neighbors can be collected. In the third phase, the neighbor information gathered is treated in the form of finite states. An adaptive Q-learning strategy enables that UAVs perform distributed, autonomous, and adaptive routing decisions. The reward function has been designed taking into account the wireless links' quality, residual energy, and neighbor distance, in order to determine stable communications paths with reduced PLR, fewer hops, and lower energy consumption. In addition, the action selection, learning rate, and discount factor are fine-tuned based on the residual Link Duration, in order to enable adaptive learning in dynamic environments. The proposed Q-learning strategy is shown in Fig. 5.16. The performance evaluation of TARRAQ has considered three state of the art counterparts following similar approaches: Greedy Perimeter Stateless Routing based on Energy-Efficient *HELLO* [179], Mobility Prediction-based Virtual Routing [180], and QTAR [35]. Through *MATLAB*'s simulations the authors have concluded the superior performance of TARRAQ regarding PDR, routing overhead, and energy consumption.

The use of ML-based routing protocols is typically associated with increased computation complexity and demanding computational power. This is a challenge in power-constrained UAVs. Moreover, ensuring the cooperation between multiple UAVs, acting as independent ML agents,

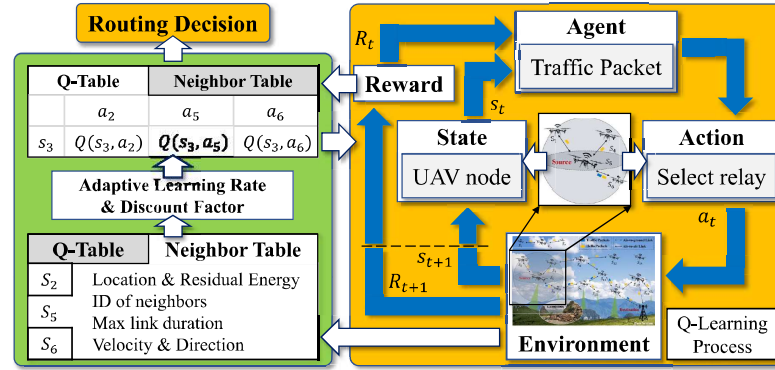


Figure 5.16: Q-learning routing strategy employed by the TARRAQ protocol proposed in [36].

for jointly optimizing the network performance is a challenge, due to the typically distributed nature of the ML-based routing protocols. Most of the predictive routing solutions proposed in the literature aim at maximizing the communications paths' lifetime, while taking into account the energy available at the UAVs and improving the reliability of the wireless links by employing mobility prediction-based approaches. The deterministic knowledge about the future positions of the UAVs, calculated in advance by centralized UAV placement algorithms to meet the traffic demand of the ground devices served, has not been taken into account in the literature concerning network routing. Moreover, the improved wireless propagation conditions in FMNs are worthy of being considered to achieve predictive routing decisions. Specifically, since the wireless links between the UAVs are characterized by a strong Line-of-Sight component, induced by the potential absence of obstacles above the ground, simplified wireless propagation models, such as the Free-Space Path Loss (FSPL) model, may be used to estimate 1) the SNR of the wireless links [134] as well as 2) the interference between concurrent flows forwarded through multi-hop communications paths enabled by UAVs positioned close to each other. From the network planning point of view, a holistic and centralized view over the network allows predicting potential changes in the wireless links' quality and estimating inter-flow interference caused by neighbor UAVs. When it comes to network routing, defining in advance the instants for reconfiguring the UAVs' forwarding tables paves the way to avoid disruptions in the communications paths and improve the QoS offered. In the following section, we review the main state of the art routing metrics proposed for wireless routing protocols.

5.1.3 Routing Metrics for Flying Multi-hop Networks

The ability to offer the QoS expected by the users is a targeted requirement in FMNs. In FMNs, the usage of UAVs carrying Wi-Fi Access Points (APs) and cellular BSs may not only provide adequate wireless coverage but also latency-constrained communications paths with high

enough capacity to accommodate the traffic demand of the ground devices. However, many UAVs need to be deployed close to each other. This may end up in a high level of interference between neighboring UAVs, which must be addressed for improved network performance. As referred to in Section 5.1.1, the existing routing solutions for FMNs are based on the protocols employed in Mobile Ad hoc Networks and Vehicular Ad hoc Networks, and most of them have proposed their own routing metrics. The routing metrics take into account different factors affecting network performance, including the quality and stability of the wireless links, and the energy available at the communications nodes. On the one hand, the heuristic-based solutions presented in Section 5.1.2.1 use routing metrics that may imply large overhead, time-consuming topology estimation, and complex path computation, due to their distributed nature. On the other hand, the ML-based routing solutions referred to in Section 5.1.2.2 may introduce a high overhead during exploration, used for determining emerging optimal communications paths in real-time, especially when the network topology is prone to significant changes [151]. An inherent challenge lies in achieving a trade-off between the exploration and exploitation phases, while addressing the dynamism of the network topology. In the following, we present specific routing metrics proposed in the literature for wireless networks.

ETX [181] aims at estimating the number of transmissions required to successfully deliver a packet through a wireless link by capturing the PLR and link length. Similarly, Expected Transmission Time (ETT) [182] aims at estimating the MAC layer duration to successfully transmit a packet through a wireless link. ETT is an improved version of ETX, which takes into account the quality and capacity of the wireless links. However, neither ETX nor ETT consider the effect of interference in the wireless links. Weighted Cumulative Expected Transmission Time (WCETT) [182] has been the first multi-channel routing metric that took into account the effect of interference. WCETT improves ETT by considering intra-flow interference, in order to select communications paths consisting of wireless links using as many channels as possible. However, WCETT is not isotonic, which means that the order of the weights of two paths is not preserved if they are appended or prefixed by a third path; hence, WCETT does not allow for paths without loops. Additionally, WCETT does not capture inter-flow interference, which makes it not appropriate for highly dense networks. Metric of Interference and Channel-switching (MIC) [183] represents a step forward to WCETT, since it is isotonic and takes into account intra-flow and inter-flow interference; nevertheless, MIC is designed for static networks, since it does not consider time-varying interference, such as due to changes in the received power and the amount of traffic offered by the communications nodes. The Interference Aware (iAWARE) [184] routing metric uses SNR and Signal-to-Interference plus Noise Ratio to estimate inter-flow and intra-flow interference. However, iAWARE is not isotonic.

These metrics are adopted by most of the state of the art routing protocols used in generic wireless networks. An exception is the B.A.T.M.A.N. routing protocol, which has introduced a

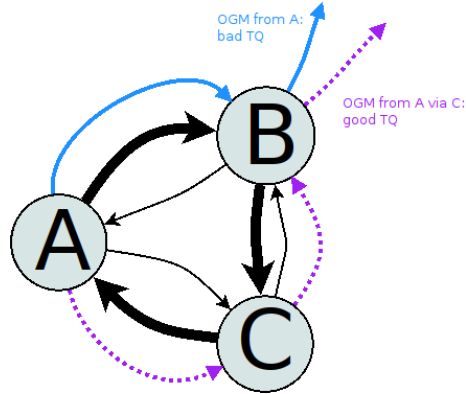


Figure 5.17: Representative networking scenario consisting of three nodes connected by means of asymmetric wireless links. Strong black line arrows are associated with a good connection (i.e., low PLR), while thin black line arrows are associated with a bad connection (i.e., high PLR). B.A.T.M.A.N. IV rebroadcasts the Originator Messages (OGMs) with the highest Transmission Quality (TQ) value with respect to the originator. Herein, node *B* sends the TQ value received via node *C* in the OGMs associated with node *A* [37].

new routing metric [185]. In B.A.T.M.A.N., the communications nodes broadcast OGMs to inform the neighborhood about their existence. OGMs allow to infer the quality of the wireless links and the congestion of the network. In particular, B.A.T.M.A.N. IV [37] considers two components to assess the wireless link's quality: 1) transmission link quality and 2) receiving link quality. The receiving link quality refers to the probability of successfully delivering a packet to a node. The transmission link quality describes the probability of a transmission to a neighboring node is successful. Only the transmission link quality influences the routing decisions. OGMs are used to inform neighboring nodes about the transmission link quality of each node. B.A.T.M.A.N. IV has introduced a TQ field in the OGMs. Before an OGM is broadcasted after being generated, its TQ field is set to a maximum value. After receiving an OGM, each node changes the TQ value and rebroadcasts it. For locally changing the TQ value of each OGM, every communications node considers that $TQ = TQ_{incoming} \times TQ_{local}$. As such, each node in the network is aware of the transmission quality towards the source node. For the representative networking scenario depicted in Fig. 5.17, consisting of three communications nodes connected through asymmetric wireless links, the data packets from node *A* have a high probability of being successfully delivered to node *B*, but the TQ value broadcasted by node *B* is lower due to a high PLR towards node *A*. On the other hand, the data packets generated by node *A* that are forwarded through node *C* have a low probability of arriving at node *B* due to the high PLR, despite they have a higher TQ value. As such, B.A.T.M.A.N. IV rebroadcasts the received OGMs with the highest TQ value with respect to the originator. Herein, node *B* reflects in the message from node *A* the TQ value received via node *C*, while broadcasting OGMs with the highest TQ only.

All the state of the art routing metrics presented herein use probe packets for neighbor discovery in real-time. Additionally, inter-flow interference in FMNs is not a local concept, since it may be caused by different communications nodes along a communications path. A routing metric fed by a holistic and centralized view over the network is worth considering, in order to make it possible estimating in advance the quality of the wireless links based on the UAVs' positions defined by centralized placement algorithms. A routing protocol employing a routing metric fed by a holistic view over the network requires only control packets for periodically sending up-to-date forwarding tables to the UAVs. Considering that each FMN's reconfiguration is transitory and relatively short, when compared to the longer interval in which the UAVs are in fixed positions, such routing approach allows reducing the overhead introduced in the FMN and enables reconfiguring the FMN without loss of connectivity.

5.2 System Model and Problem Formulation

Herein, we assume the architecture of the FMN considered in Chapter 2, Chapter 3, and Chapter 4. The FMN is controlled by a Central Station, which is in charge of defining the UAV placement and computing the forwarding tables of the UAVs. The UAVs act as SDN switches. The Central Station benefits from a holistic knowledge about the network, including the UAVs geographical coordinates, assumed to be defined by UAV placement algorithms.

At the discrete time instant $t_k = k \cdot \delta t$, where $k \in \mathbb{N}_0$ and $\delta t \in \mathbb{R}$, the FMN is represented by means of a graph $G(t_k) = (V, E(t_k), w(t_k))$, where V represents the set of UAVs forming the FMN, $E(t_k) \subseteq V \times V$ represents the communications links, and $w(t_k)$ represents the cost assigned to the communications links that compose G . $(i, j)_{t_k} \in E(t_k)$ represents the directional communications link available from UAV_{*i*} to UAV_{*j*} at t_k , where $i, j \in V$, and $w_{i,j}(t_k)$ represents the cost of link $(i, j)_{t_k}$. $P_i(t_k) = (x_i, y_i, z_i)_{t_k}$ represents the position of UAV_{*i*} at time t_k and depends on the UAV initial position at instant t_0 and its associated velocity and acceleration vectors defined by the Central Station in charge of controlling the UAV positions. The availability of the wireless links connecting the UAVs changes over time. The directional wireless communications link $(i, j)_{t_k}$ exists if and only if the power $P_{R_i}(t_k)$ received by UAV_{*i*} at time t_k over the noise power N_i satisfies Eq. (5.1), which defines that the SNR should be higher than the threshold S . The received power at UAV_{*i*}, $P_{R_i}(t_k)$, results from the FSPL model defined in Eq. (5.2), where $P_{T_j}(t_k)$ refers to the transmission power of UAV_{*j*} at time t_k , $\lambda_{i,j}$ denotes the link wavelength, and $d_{i,j}(t_k)$ expresses the Euclidean distance between UAV_{*i*} and UAV_{*j*} at time t_k .

$$\frac{P_{R_i}(t_k)}{N_i} > S \quad (5.1)$$

$$\frac{P_{R_i}(t_k)}{P_{T_j}(t_k)} = \left(\frac{\lambda_{i,j}}{4\pi \times d_{i,j}(t_k)} \right)^2 \quad (5.2)$$

We define a path as a set of adjacent links connecting UAV_{*i*} to UAV₀ (FGW). Multiple paths may be available for UAV_{*i*} at time t_k , but only one of them is used, in order to minimize the inter-flow interference, which constitutes a major cause of network performance degradation in shared wireless channels. We also define $C_{i,j}(t_k)$ as the capacity, in bit/s, of the communications link available between UAV_{*i*} and UAV_{*j*} at time t_k , considering a constant value for the link bandwidth $B_{i,j}$ in Hz; Shannon-Hartley theorem is used for this purpose, as given by Eq. (5.3).

$$C_{i,j}(t_k) = B_{i,j} \times \log_2 \left(1 + \frac{P_{R_i}(t_k)}{N_i} \right) \quad (5.3)$$

Considering the throughput $R_i(t_k)$, in bit/s, as the bit rate of the flow from UAV_{*i*}, $i \in \{1, \dots, N-1\}$ received at the FGW at time t_k , and $N-1$ UAVs generating traffic towards the FGW, we aim at defining the forwarding table of every UAV_{*i*} at each time instant t_k to maximize the amount of bits received by the FGW during the subsequent time interval δt (e.g., $\delta t = 1$ s). Our objective function is defined in Eq. (5.4).

$$\text{maximize } \sum_{i=1}^{N-1} R_i(t_k) \times \delta t \quad (5.4)$$

The factors influencing $R_i(t_k)$ include the capacity of the path used by UAV_{*i*}, which should be limited by the link in the path having the smallest capacity, the number of flows traversing the links, the medium access protocol behavior, and the interference between the communications nodes. The capacity of the wireless links is calculated using the Shannon–Hartley theorem [131]. The Shannon-Hartley theorem is a non-convex function [101, 102]. Considering that the objective function defined in Eq. (5.4) is associated with multiple non-convex functions, the optimization problem is non-convex too. As referred to in Chapter 4, a non-convex optimization problem is NP-hard [103]. In order to solve the problem, for each time instant t_k we will find a communications path for each UAV_{*i*} so that we meet the objective function defined in Eq. (5.4).

5.3 RedeFINE Mechanism

In this section, the RedeFINE mechanism is described, along with its rationale and a theoretical analysis for a representative scenario.

5.3.1 Rationale

RedeFINE was designed to take advantage of the centralized view of the FMN available at the Central Station. By considering 1) the future positions of the FMN UAVs, which are computed at the Central Station to meet the traffic demand of the ground devices, and 2) the velocity of the UAVs following a straight line between the origin and destination points, RedeFINE selects periodically the optimal path for each UAV. We define the optimal path as the one with the lowest cost. In RedeFINE we define the cost of a link $w_{i,j}(t_k) = d_{i,j}(t_k)$, considering the Euclidean distance as the link cost; the path cost is the sum of the costs of its composing links. As the distance decreases, the capacity of the link (cf. Eq. (5.3)) increases. Using distance as path cost allows fulfilling the objective function defined in Eq. (5.4). RedeFINE assumes a strong Line-of-Sight component between the communications nodes, which characterizes the communications links between UAVs flying above the ground. In order to find the shortest path between UAVs, RedeFINE employs the Dijkstra's algorithm [186].

Although RedeFINE was designed for defining the backhaul wireless paths between the FAPs and the FGW, taking into account networking scenarios composed mainly of Internet traffic, it can also be used for defining backhaul wireless paths that allow communications between the ground devices served by the FMN. For that purpose, the FAPs' forwarding tables should be computed so that the FAPs forward the packets received from the ground devices 1) directly to the FAP serving the destination ground device or 2) to the FGW as the default forwarding decision. RedeFINE benefits from the centralized and holistic view provided by the Central Station, which knows the FAPs that are serving the ground devices over time.

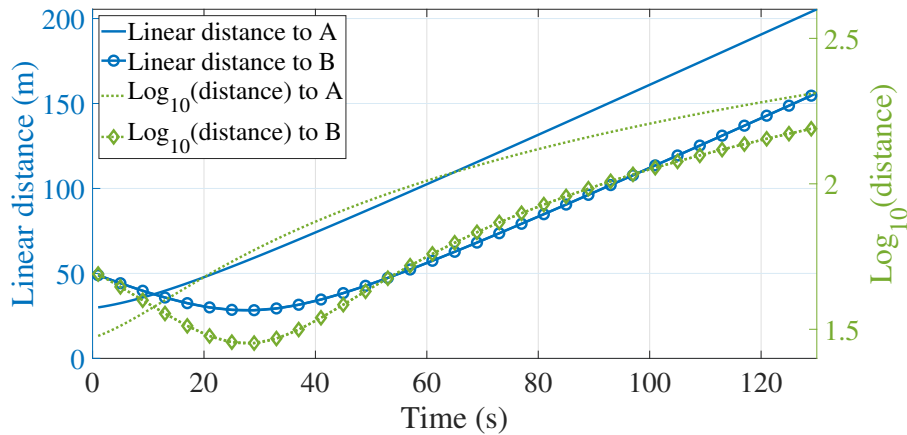


Figure 5.18: Distance between a moving communications node and two static communications nodes over time, in linear and logarithmic scale. Linear scale is the most suitable to evaluate distance changes.

RedeFINE uses the linear Euclidean distance between neighbors as routing metric. An alternative metric would be the logarithmic distance; however, the linear scale has been confirmed to be most suitable to evaluate the effect of distance on the link quality, as depicted in Fig. 5.18. Fig. 5.18 depicts the distance variation over time, in both linear (continuous lines) and logarithmic (dashed lines) scale, between a moving node and two static communications nodes – node A (without marker) and node B (with marker). In the linear scale each unit of change between both curves is represented by the same distance (cf. Y-axis on the left hand-side of Fig. 5.18). Conversely, in the logarithmic scale the distance (cf. Y-axis on the right hand-side of Fig. 5.18) is only the same for equal relative changes. Thus, while in the linear scale the difference between 100 m and 50 m is the same as the difference between 150 m and 100 m, in the logarithmic scale the difference between these values is not equal. On the other hand, in the logarithmic scale the difference between, for instance, 200 m and 100 m is the same as between 100 m and 50 m. Therefore, the linear scale is preferable to be used as routing metric, in order to characterize the link quality based on the distance.

5.3.2 Theoretical Analysis for a Representative Scenario

In order to demonstrate the RedeFINE concept and perform a preliminary evaluation, an analysis for a representative scenario is presented herein. For the sake of simplicity, we consider the scenario depicted in Fig. 5.19. It is formed by: 1) an FGW, at the highest altitude; 2) two relay FAPs, FAP_1 and FAP_2 , at the intermediate tier; and 3) a moving FAP, FAP_3 , which is placed at the lowest altitude and is following a straight line at 0.5 m/s in the direction from FAP_1 to FAP_2 , between $t = 0$ s and $t = 130$ s. The initial position of UAV_i is represented by $P_i(0) = (x_i, y_i, z_i)_0$ and the constant velocity is denoted by $V_i = (v_{x_i}, v_{y_i}, v_{z_i})$. The initial cost of the wireless link between UAV_i and UAV_j is represented by $w_{i,j}(0)$, which is the linear Euclidean distance between both UAVs considering their initial positions. The wireless link availability is restricted by a 5 dB SNR threshold, that is $S \approx 3.16$ (cf. Eq. (5.1)). The SNR is derived from the FSPL propagation model (cf. Eq. (5.2)). We assume that the maximum capacity for the wireless links is given by the Shannon-Hartley theorem (cf. Eq. (5.3)). The maximum capacity of a wireless link is computed considering an average noise power equal to -85 dBm. In turn, the capacity of a path is restricted by the link with lower capacity among the set of links forming the path. Additionally, we consider the transmission power equal to 0 dBm, the carrier frequency equal to 5250 MHz, and the channel bandwidth equal to 160 MHz, which are values compatible with the IEEE 802.11ac standard. Based on these assumptions, the theoretical maximum capacity values for the paths between FAP_0 and the FGW are plotted in Fig. 5.20. With RedeFINE, in order to reach the FGW, FAP_3 uses as relay nodes FAP_1 until $t = 60$ s and FAP_2 from $t = 60$ s to $t = 130$ s. These are the highest-capacity paths, which are highlighted by the circle symbol in Fig. 5.20.

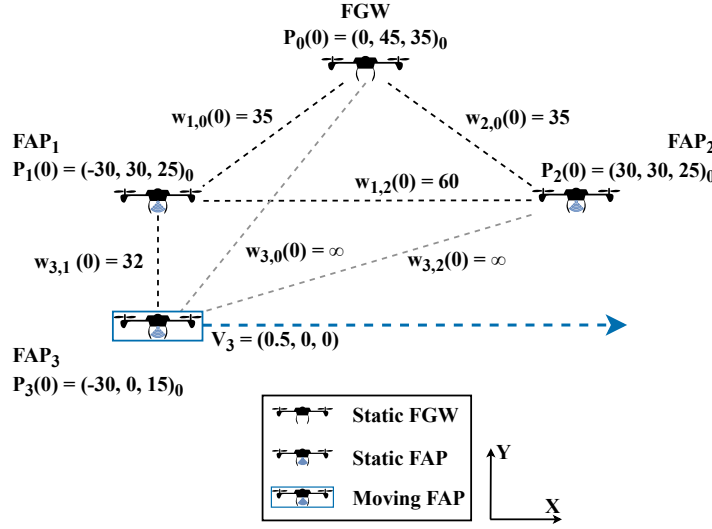


Figure 5.19: Scenario used as reference case. For the initial position of FAP₃, the wireless links between FAP₃ and both FAP₂ and FGW are not available, due to the SNR threshold constraint.

Considering as baseline the static routing configuration that uses FAP₁ as relay node for the communications between FAP₃ and the FGW during 130 s, RedeFINE allows a gain of $\approx 20\%$ regarding the total amount of bits received in the FGW. The areas under the curves in Fig. 5.20 give the total amount of bits carried by the respective path. We compute the amount of information received on each case (static and RedeFINE) as $\sum_{k=0}^{130} R_0(t_k) \times \delta t$, considering $\delta t = 1$ s. In this analysis, FAP₁ and FAP₂ do not generate traffic; they only forward data traffic received from FAP₃.

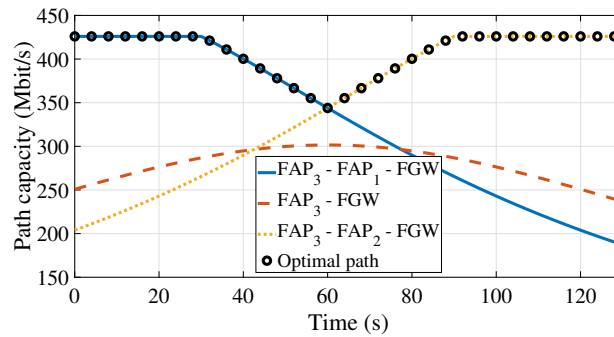


Figure 5.20: Theoretical capacity values for the paths between FAP₃ and the FGW. In order to reach the FGW, FAP₃ uses as relay nodes FAP₁ until $t = 60$ s and FAP₂ from $t = 60$ s to $t = 130$ s. The path with the highest capacity at each instant is highlighted by the circle symbol.

5.4 Performance Evaluation of RedeFINE

The methodology followed in the performance evaluation of RedeFINE is presented herein, including the simulation setup, the simulation scenarios, and the performance metrics used.

5.4.1 Simulation Setup

In order to evaluate the performance of RedeFINE in more complex networking scenarios, the ns-3 simulator was used. The FMN consisted of 1 FGW in the first tier, 16 FAPs in the second tier, and 4 FAPs in the third tier generating traffic in descending order of altitude respectively. The FAPs in the third tier were moving according to the *ConstantVelocityMobilityModel* at 0.5 m/s during 130 s. The FGW and the FAPs acting as relay nodes are static. A Network Interface Card (NIC) on each UAV was configured in Ad Hoc mode, using the IEEE 802.11ac standard in channel 50, with channel bandwidth of 160 MHz, and Guard Interval of 800 ns. The wireless links between UAVs were modeled by the FSPL model; only links with SNR higher than or equal to 5 dB were considered. The transmission power of the NICs was set to 0 dBm and the UAVs positions were defined in order to guarantee, at least, ≈ 15 dB of SNR for the links between neighbors in their initial positions.

One spatial stream was used for all inter-UAV IEEE 802.11ac links. With 1 spatial stream, the data rate corresponding to the maximum Modulation and Coding Scheme (MCS) index is 780 Mbit/s. Since the FMN includes FAPs acting as relay nodes, the FGW and the 4 FAPs in the third tier are 2 hops away; this results in $780/4/2 = 97.5$ Mbit/s for the maximum achievable data rate per flow. Based on that, the maximum offered load for each scenario was set to: $\lambda_1 = 0.30 \times 97.5 \approx 24$ Mbit/s, and $\lambda_2 = 0.75 \times 97.5 \approx 73$ Mbit/s. The traffic generated was UDP, for a constant packet size of 1400 bytes, with the arrival process modeled as Poisson. The IEEE 802.11ac data rate was automatically defined by the *IdealWifiManager* mechanism. The traffic generation and the FAPs' movement were only triggered after 30 s of simulation, in order to ensure a stable state. A summary of the ns-3 simulation parameters considered is presented in Table 5.1.

5.4.2 Simulation Scenarios

Two simulation scenarios were used to evaluate the performance of RedeFINE and the state of the art routing solutions in typical crowded events. Scenario A (cf. Fig. 5.21a) considered the movement away from a central point, in order to reproduce the dispersion of people from a central point to different points of interest in the event. Scenario B (cf. Fig. 5.21b) considered a linear movement along the event venue, in order to exemplify the movement of people between two points of interest in the event area.

Table 5.1: Summary of the main ns-3 simulation parameters used in the evaluation of the RedeFINE mechanism.

Parameter	Value
Simulation time	30 s (bootstrap) + 130 s
Wi-Fi standard	IEEE 802.11ac
Wi-Fi mode	Ad Hoc
Wi-Fi channel	50 (5250 MHz)
Channel bandwidth	160 MHz
Transmission (TX) power	0 dBm
Guard Interval	800 ns
Number of antennas	1
Propagation delay model	Constant speed
Propagation loss model	FSPL (Friis)
Remote station manager	<i>IdealWifiManager</i>
Mobility model	Constant velocity
Third tier's FAPs speed	0.5 m/s
Traffic type	UDP Poisson
Packet size	1400 bytes

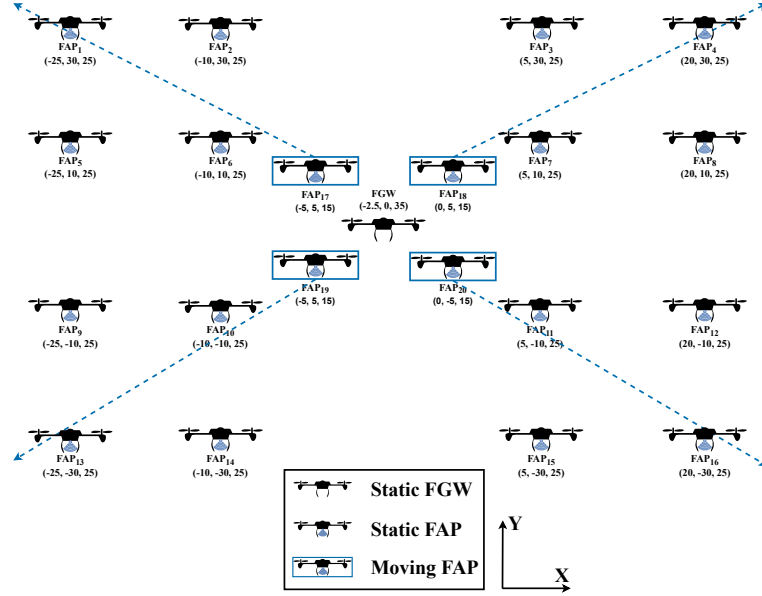
5.4.3 Performance Metrics

RedeFINE was evaluated in simulation environment against two state of the art distributed routing protocols representative of the reactive and proactive routing paradigms – AODV and OLSR, respectively – and static routing, for which the forwarding tables are maintained throughout the whole simulation time.

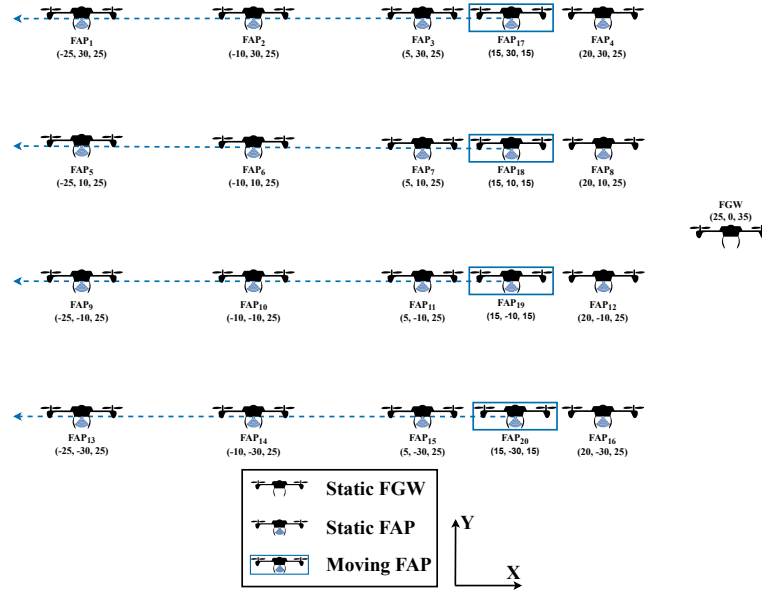
Since the default implementations of AODV and OLSR in the latest release of ns-3 do not employ a link quality-based routing metric, we used alternative implementations for both protocols, publicly available in [187] and [188], respectively. They use the ETX routing metric[181]. This enables a fair comparison with RedeFINE. ETX represents the predicted number of transmissions required to send a packet over a link, including retransmissions.

The performance evaluation of RedeFINE considers three performance metrics:

- **Aggregate throughput** – The average number of bits received per second by the FGW.
- **PDR** – The number of packets received by the FGW divided by the number of packets generated by the FAPs, measured at each second.



(a) Scenario A: movement away from a central point.

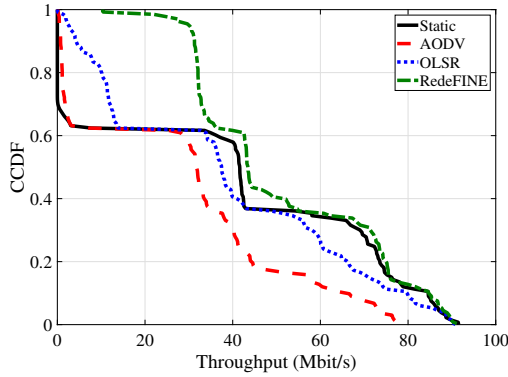


(b) Scenario B: linear movement along the event venue.

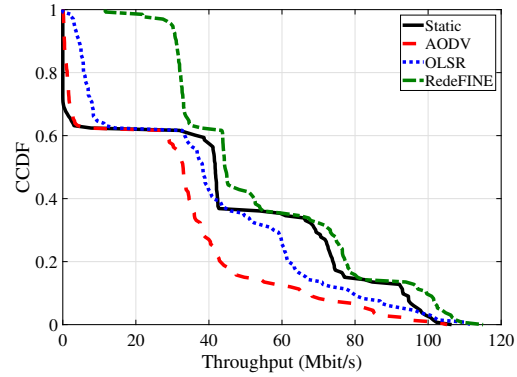
Figure 5.21: Simulation scenarios considered for the RedeFINE mechanism evaluation.

- **Delay** – Time taken by the packets to reach the application layer of the FGW since the instant they were generated, measured at each second; it includes queuing, transmission, and propagation delays.

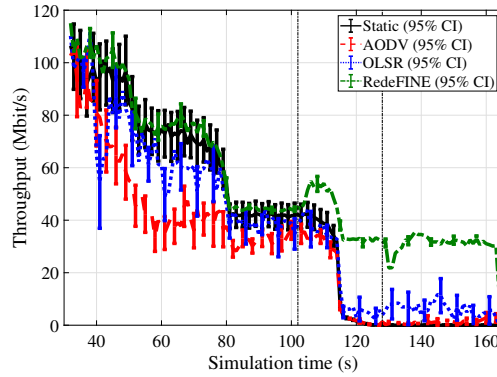
5.4.4 Simulation Results



(a) Throughput Complementary Cumulative Distribution Function (CCDF) for $\lambda=24$ Mbit/s.



(b) Throughput CCDF for $\lambda=73$ Mbit/s.



(c) Throughput versus simulation time for $\lambda=73$ Mbit/s. The vertical lines represent the path change events defined by RedeFINE.

Figure 5.22: Scenario A: throughput results measured in the FGW, considering UDP Poisson traffic.

The simulation results obtained for both scenarios are presented in this section. The results were obtained considering 20 simulation runs for each scenario. The results are expressed by means of the CDF for the average delay and by the CCDF for the aggregate throughput and PDR. The CDF $F(x)$ represents the percentage of simulation time for which the average delay was lower or equal to x , whereas the CCDF $F'(x)$ represents the percentage of simulation time for which the average aggregate throughput or PDR was higher than x . The average aggregate throughput measured in the FGW versus the simulation time is also presented.

The CCDF representations for the aggregate throughput in the FGW (cf. Fig. 5.22 and Fig. 5.25), considering different packet generation rates, show that RedeFINE allows higher

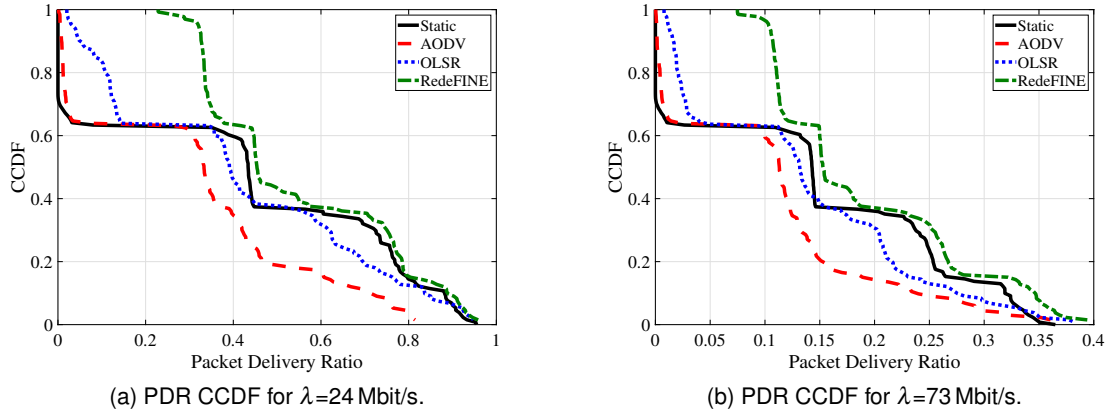


Figure 5.23: Scenario A: PDR results measured in the FGW, considering UDP Poisson traffic.

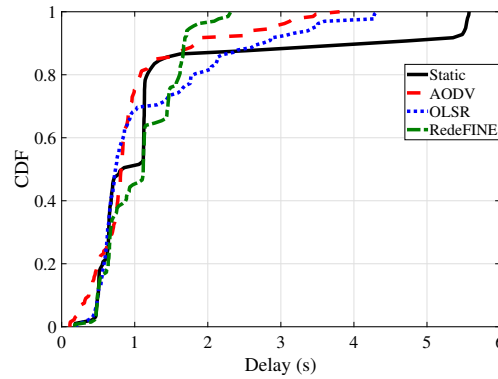


Figure 5.24: Scenario A: delay Cumulative Distribution Function (CDF) measured in the FGW, considering UDP Poisson traffic with $\lambda=73$ Mbit/s.

throughput than its state of the art counterparts. RedeFINE overcomes the state of the art solutions mainly in the last stage of the FAPs movement; the static route to the FGW becomes unusable, while AODV and OLSR switch in short time periods between different relay nodes, which decreases the throughput substantially, as depicted in Fig. 5.22c and Fig. 5.25c. RedeFINE enables throughput higher than 10 Mbit/s for the whole simulation time, which demonstrates its superior performance. This is due to the selection of stable paths with minimal length, formed by links with high SNR, high capacity, and low delay. The area under the CCDF curves for the aggregate throughput gives the total number of bits received in the FGW. For Scenario A, we conclude that RedeFINE increases the number of bits received up to 27%, 47%, and 31% when compared with static routing, AODV, and OLSR, respectively; for Scenario B, RedeFINE increases the number of bits received up to 37%, 54%, and 46%, respectively. The CCDF plots for PDR follow the same pattern as the CCDF for the aggregate throughput. Overall, the PDR in Scenario A (cf. Fig. 5.23)

is higher than in Scenario B (cf. Fig. 5.26). This is justified by the links with higher capacity in Scenario A, due to the lower distances between the FAPs and the FGW. Nevertheless, RedeFINE allows higher PDR than static routing, AODV, and OLSR for both scenarios. Regarding the delay, RedeFINE does not achieve the minimum values that its state of the art counterparts do, as depicted in Fig. 5.24 and Fig. 5.27. However, this is justified by the higher PDR achieved by RedeFINE. While for static routing, AODV, and OLSR more packets are dropped, with RedeFINE the network is subject to a higher level of congestion; hence, packets are held in the transmission queue longer.

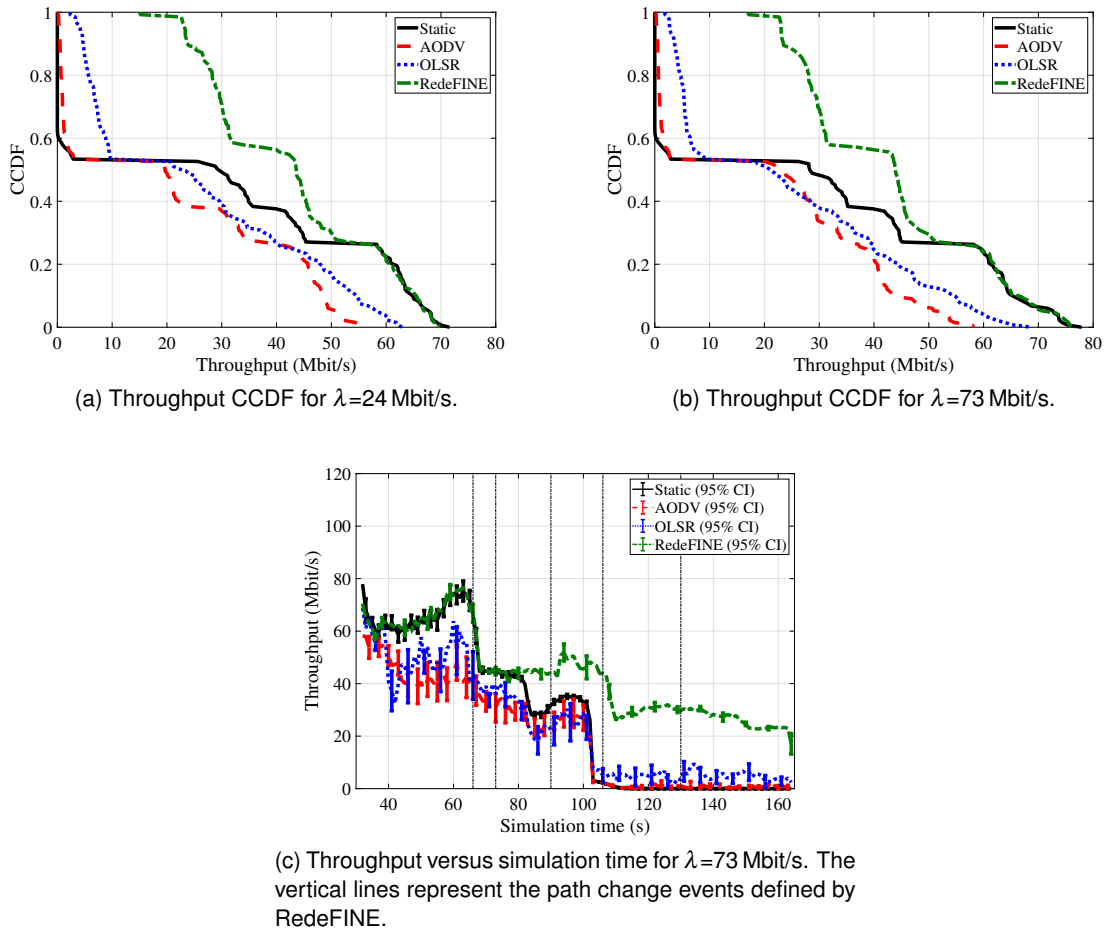


Figure 5.25: Scenario B: throughput results measured in the FGW, considering UDP Poisson traffic.

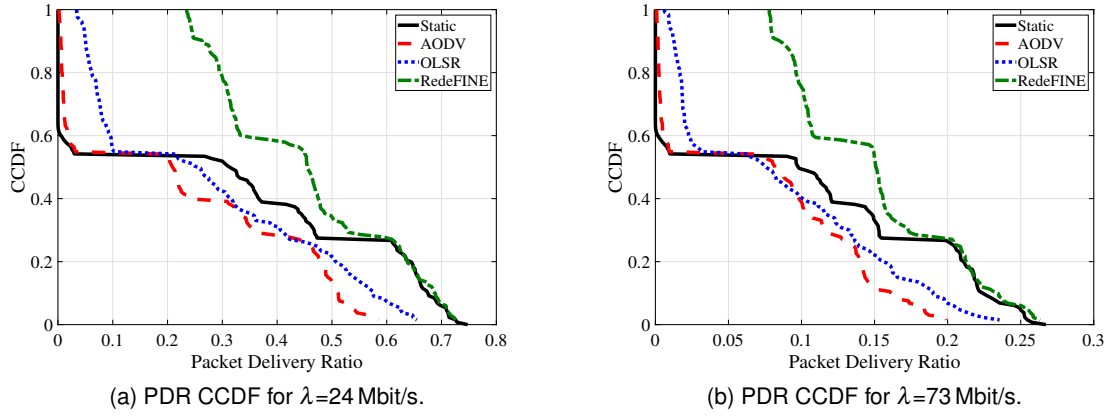


Figure 5.26: Scenario B: PDR results measured in the FGW, considering UDP Poisson traffic.

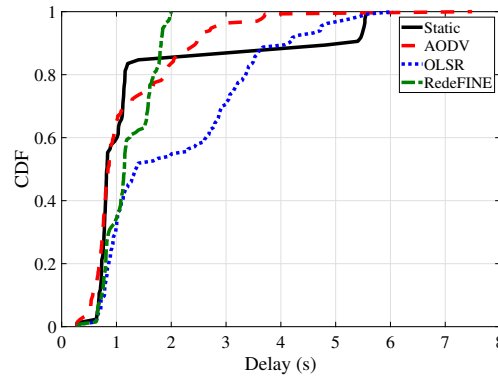


Figure 5.27: Scenario B: delay CDF measured in the FGW, considering UDP Poisson traffic with $\lambda=73$ Mbit/s.

5.5 I2R Routing Metric

Setting up a multi-hop wireless network with broadband links is a challenging task. Due to transmissions through shared wireless channels, which characterize IEEE 802.11-based multi-hop networks, the interference caused by the traffic being transmitted by neighboring communications nodes must be taken into account. This type of interference is commonly named inter-flow interference. The problem is exacerbated in crowded events, such as fun parks, parades, and festivals, where many ground devices need to be served, requiring an FMN with many UAVs close to each other to address the traffic demand. However, typically, estimating inter-flow interference is not straightforward, since it depends on the radio propagation environment and the communications nodes in carrier-sense range.

In Section 5.3, we presented RedeFINE, a centralized routing solution for FMNs that uses the

Euclidean distance between the UAVs as the routing metric to define in advance the forwarding tables and the instants they shall be updated in the UAVs. RedeFINE takes advantage of the holistic knowledge provided by a Central Station, which is responsible for defining the locations of the UAVs to meet the traffic demand from the ground devices. However, the Euclidean distance routing metric is not suitable for highly dense networks, since the interference between neighboring UAVs is not captured; in order to address this problem, we propose the I2R metric, designed for centralized routing in FMNs. I2R does not require probe packets and enables the configuration of paths with minimal Euclidean distance formed by UAVs with the lowest number of neighbors in carrier-sense range, thus minimizing inter-flow interference in the FMN. I2R advances the state of the art by avoiding the usage of probe packets for neighbor discovery and interference estimation. Due to the highly dynamic behavior of FMNs, which induce frequent changes in the quality of the wireless links and in the network topology, existing routing solutions need to constantly flood the network with route discovery packets, thus reducing the bandwidth available for data traffic. In I2R, since a strong Line-of-Sight component characterizes the communications links between UAVs flying above the ground, we assume the wireless links can be modeled by the FSPL model, which is used to estimate the SNR and identify the neighbors in carrier-sense range without any additional signaling. In the following, we present the rationale of I2R and a theoretical analysis for a reference scenario.

5.5.1 Rationale

The I2R metric consists of two factors: the distance-aware factor and the inter-flow interference-aware factor. The distance-aware factor is based on the Euclidean distance of the links between each pair of UAVs, at time instant t_k , which is denoted by $d_{i,j}(t_k)$. As such, this factor includes the sum of the Euclidean distances of the set of links forming a path p , considering in advance the future trajectories that UAVs will follow, which are computed and pre-defined by the Central Station to fulfill the traffic demand of the ground devices. Using the Euclidean distance as part of the routing metric is compliant with the objective of selecting high-capacity paths, since the link capacity increases as the Euclidean distance decreases, according to the Shannon-Hartley theorem. $d_{i,j}(t_k)$ is normalized to the maximum Euclidean distance among all the usable links of the FMN, at t_k . In turn, the inter-flow interference aware factor is a value $\gamma(t_k)$ that is added to the Euclidean distance of the link between UAV_{*i*} and UAV_{*j*}, at t_k . $\gamma_j(t_k)$ represents the number of neighboring communications nodes of UAV_{*j*}, excluding UAV_{*i*}, at t_k . We assume as neighboring communications nodes the UAVs in carrier-sense range that are generating or forwarding traffic. $\gamma_j(t_k)$ considers that the level of interference is equal either the neighboring communications nodes are close or far away, as the Transmitter-Receiver Conflict Avoidance model [189] denotes. $\gamma_j(t_k)$ is normalized to the maximum number of neighbors that any UAV composing the FMN has at t_k .

The path cost using I2R is defined in Eq. (5.5), where $0 \leq \alpha \leq 1$ is a tunable parameter that weights the influence of the distance-aware and interference-aware factors. To compute the path between any UAV and the FGW, the Dijkstra's algorithm [186] is used. Considering the reference case depicted in Fig. 5.28, I2R uses the factor γ to increase the cost of the links $FAP_2 \rightarrow FAP_8$ and $FAP_6 \rightarrow FAP_8$, since FAP_8 is in the interference range of FAP_2 and FAP_6 .

$$I2R = (1 - \alpha) \times \sum_{\forall (i,j) \in P} d_{i,j}(t_k) + \alpha \times \sum_{\forall j \in P} \gamma_j(t_k) \quad (5.5)$$

5.5.2 Theoretical Analysis for a Reference Scenario

Most state of the art routing solutions do not consider the effect on the network performance of concurrent flows transmitted over different wireless links that use the same communications channel. However, it leads to a significant reduction of the capacity of the wireless links, since the flows compete for the same communications channel resources in time and frequency [189]. Taking into account the IEEE 802.11 MAC protocol, for a packet transmission to be successful neither the transmitter nor the receiver should be interfered by other communications nodes. Hence, the transmissions on links $(i, j)_{t_k}$ and $(k, l)_{t_k}$ are both successful at t_k if and only if both i and j are outside the interference range of k and l at t_k . This is expressed by the Transmitter-Receiver Conflict Avoidance interference model.

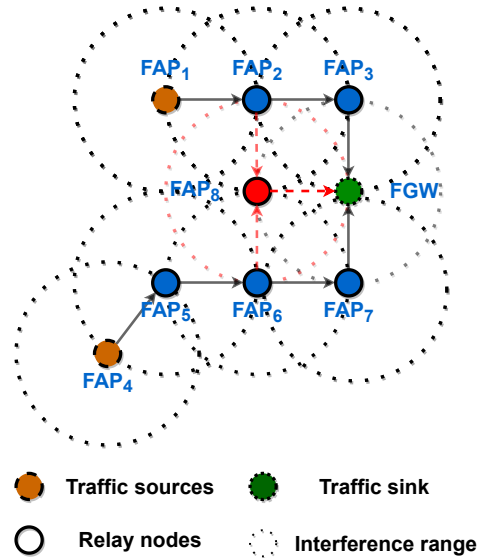


Figure 5.28: Network graph illustrating the Transmitter-Receiver Conflict Avoidance interference model. If FAP_8 is used as relay node, the network performance will be reduced up to 50%, since FAP_8 is in the interference range of FAP_2 and FAP_6 .

In order to demonstrate how the selection of relay nodes affects the network performance, let us analyze a reference case. For the sake of simplicity, we consider the scenario depicted in Fig. 5.28. It is formed by: 1) two FAPs generating traffic – FAP_1 and FAP_4 ; 2) one FGW; and 3) six FAPs able to forward data traffic. The interference range of each node is represented by a dashed circumference around that node. Firstly, we consider two paths for the flows between the FAPs generating traffic and the FGW: $p_1 : FAP_1 \rightarrow FAP_2 \rightarrow FAP_3 \rightarrow FGW$ and $p_2 : FAP_4 \rightarrow FAP_5 \rightarrow FAP_6 \rightarrow FGW$. For these paths, there is no inter-flow interference, hence the throughput achieved by each flow is only limited by the link with the lowest capacity among the ones forming the path. Conversely, if FAP_8 is chosen to be part of a path $p'_1 : FAP_1 \rightarrow FAP_2 \rightarrow FAP_8 \rightarrow FGW$, since FAP_8 is in the interference range of FAP_6 , the links $FAP_2 \rightarrow FAP_8$ and $FAP_5 \rightarrow FAP_6$ become mutually interfered. Therefore, the network performance is reduced up to 50%, when compared with the previous routing configuration. This reference case motivates the use of I2R to minimize the inter-flow interference and improve the performance of RedeFINE.

5.6 Performance Evaluation of RedeFINE using I2R

The performance evaluation of the I2R metric is presented in this section, including the simulation setup, the simulation scenarios, and the performance metrics considered.

5.6.1 Simulation Setup

The ns-3 simulator was used to evaluate the proposed routing metric in complex networking scenarios formed by an FMN consisting of 1 FGW and 20 FAPs. In each UAV, a NIC was configured in Ad Hoc mode, using the IEEE 802.11ac standard in channel 50, which allows channel bandwidth of 160 MHz. The data rate was defined by the *IdealWifiManager* mechanism. The wireless links between the UAVs were modeled by the FSPL model; only links with SNR above 5 dB were considered as usable. The transmission power of the NICs was set to 0 dBm.

One IEEE 802.11ac spatial stream was used for the wireless links. With one spatial stream, the data corresponding to the maximum MCS index is 780 Mbit/s, considering Guard Interval of 800 ns. Taking into account the dimensions of the simulation scenarios, we assume an average number of 2 hops between the FAPs generating traffic and the FGW; this results in $780/N_{tx}/2$ Mbit/s for the maximum achievable data rate per flow, where N_{tx} denotes the number of FAPs generating traffic. Based on that, the maximum offered load for each scenario was set to 75% of the maximum achievable data rate per flow, for a total number of FAPs generating traffic between 5 Mbit/s and 10 Mbit/s. The traffic generated was UDP with the arrival process modeled as Poisson, for a constant packet size of 1400 bytes; the traffic generation was only triggered after 30 s of simulation, in order to ensure a stable state. In addition, different values for the tunable parameter α , between 0.2 and 1, were considered. The Controlled Delay (CoDel) algorithm [105], which is

a Linux-like queuing discipline that considers the time that packets are held in the transmission queue to decide when to discard packets, was used; it allows mitigating the bufferbloat problem. The default parameters of CoDel in ns-3 were employed. A summary of the ns-3 simulation parameters used in the performance evaluation of RedeFINE using the I2R metric is presented in Table 5.1.

5.6.2 Simulation Scenarios

Five scenarios, in which the UAVs were moving according to the Random Waypoint Mobility model, were generated to evaluate the performance of RedeFINE using the I2R metric in typical crowded events. Under the Random Waypoint Mobility model, each UAV chooses a random destination and a speed uniformly distributed between a minimum and a maximum value. Then, the UAV moves to the chosen destination at the selected speed; upon arrival, the UAV stops for a specified period of time and repeats the process for a new destination and speed [190]. Since I2R relies on knowing in advance the movements of the UAVs, instead of generating the random movements during the ns-3 simulations, we used *BonnMotion* [107], a mobility scenario generation tool. *BonnMotion* was set to create Random Waypoint 3D movements for 21 UAVs (20 FAPs and 1 FGW) within a box of dimensions $80\text{ m} \times 80\text{ m} \times 25\text{ m}$ during 160 s, considering a velocity between 0.5 m/s and 3 m/s for the UAVs. These scenarios were used to compute in advance the forwarding tables and the instants they shall be updated in the UAVs. Both the forwarding tables and the generated scenarios were finally imported to ns-3 with a sampling period of 1 s. To employ mobility to the UAVs, based on the generated scenarios, the *WaypointMobilityModel* module was used in ns-3.

5.6.3 Performance Metrics

RedeFINE using the Euclidean distance, Airtime, and I2R metrics was evaluated against two state of the art distributed routing protocols representative of the reactive and proactive routing paradigms – AODV [187] and OLSR [188], respectively, using the ETX routing metric. The Euclidean distance and I2R metrics rely on the centralized and holistic view of the network provided by the central station, while Airtime – the default routing metric specified in the IEEE 802.11s standard [191] – expresses the amount of channel resources consumed for transmitting a frame over a link. This allows comparing the predictive decisions achieved by means of the Euclidean distance and I2R metrics with a routing metric that uses real measurements to estimate data rate, overhead, and frame error ratio of the communications links, as Airtime does. Since the theoretical calculation of the Airtime resulting costs is not straightforward, we exported them from ns-3, by running previous simulations for each one of the generated scenarios. Afterwards, we employed the Dijkstra's algorithm to find the shortest paths between each FAP and the FGW, considering

a sampling period of 1 s and the corresponding routing metric. The ETX routing metric was used by AODV and OLSR in our performance evaluation. ETX represents the expected number of transmissions required to send a packet over a link, including retransmissions. This allows a fair performance comparison between AODV and OLSR with RedeFINE, considering a quality-based routing metric for each solution.

Our performance evaluation considers aggregate throughput and delay as the performance metrics, which are described in Section 5.4.3.

5.6.4 Simulation Results

The simulation results are presented in this section. The results were obtained after 20 simulation runs for each experimental combination, including different α values and different number of FAPs generating traffic. The results are expressed using average values, considering five random scenarios, as stated in Section 5.6.2. They are represented by means of the CDF for the delay and by the CCDF for the aggregate throughput, including the values for the 25th, 50th, and 75th percentiles. The CDF $F(x)$ represents the percentage of simulation time for which the average delay was lower or equal to x , while the CCDF $F'(x)$ represents the percentage of simulation for which the average aggregate throughput was higher than x . Finally, the influence of the tunable parameter α on the FMN performance is also evaluated.

When 5 FAPs are used as traffic sources (cf. Fig. 5.29), the usage of the I2R metric improves the delay achieved by RedeFINE using the Euclidean distance in approximately 22%, OLSR and AODV using ETX in 21% and 15%, respectively, and RedeFINE using Airtime in 10%. These values are obtained considering the average delay of the packets received in the FGW for the different solutions. The outperforming results of RedeFINE using the I2R metric are justified by the selection of paths formed by UAVs with reduced number of neighbors that are generating or forwarding traffic. Regarding the total amount of bits received in the FGW, RedeFINE using the I2R metric provides a gain up to 7% when compared with RedeFINE using the Euclidean distance. In turn, when compared with AODV and OLSR using the ETX routing metric, and RedeFINE using Airtime, the gains are even more relevant: approximately 45%, 17%, and 28%, respectively. When 10 FAPs are generating traffic (cf. Fig. 5.30), RedeFINE using the I2R metric improves delay in approximately 10% with respect to RedeFINE using the Euclidean distance, while the gain over AODV and OLSR using ETX is approximately 18% and 13%, respectively. The gain in delay of I2R over the Airtime routing metric applied to RedeFINE is negligible. Regarding the total amount of bits received in the FGW, the gain of RedeFINE using I2R over OLSR using ETX is still approximately 18%, and over RedeFINE using the Euclidean distance is negligible ($\approx 4\%$). Conversely, the gain over AODV using ETX is increased to approximately 68%, while with respect to RedeFINE using Airtime it is approximately 31%.

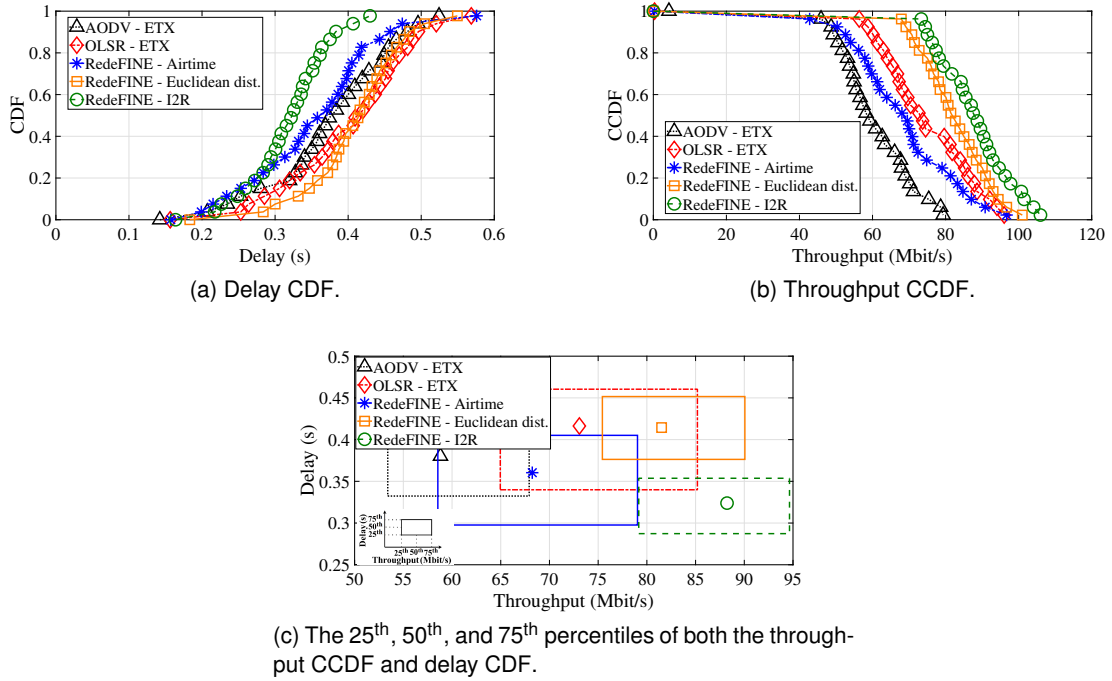


Figure 5.29: Results for throughput and delay in the FGW. The results were obtained considering 5 FAPs generating traffic, and $\alpha = 1$.

The relation between aggregate throughput and delay for the different combinations of protocols and routing metrics is depicted in Fig. 5.29c and Fig. 5.30c, where the 25th, 50th, and 75th percentiles of both the throughput CCDF and delay CDF are represented. Overall, the gains in delay and throughput of RedeFINE using I2R are reduced when the number of transmission FAPs increases. The performance evaluation carried out allowed to conclude that I2R selects preferably as relay nodes the FAPs that are also sources of traffic; for instance, in a scenario where 5 FAPs are generating traffic, if any of these FAPs need a relay to reach the FGW, then I2R will give preference to any of the remaining 4 FAPs that are generating traffic, thus avoiding that a sixth FAP introduces interference in the FMN. This effect is faded when the number of FAPs in the FMN increases.

Regarding the tunable parameter α , it must be set to a value close to 1 for higher throughput and lower delay values. As α decreases, the performance worsens, as exacerbated by $\alpha = 0.2$, in Fig. 5.31. This demonstrates how the selection of paths formed by the minimum number of neighboring UAVs in carrier-sense range contributes to improve the performance of an FMN rather than the selection based only on the Euclidean distance.

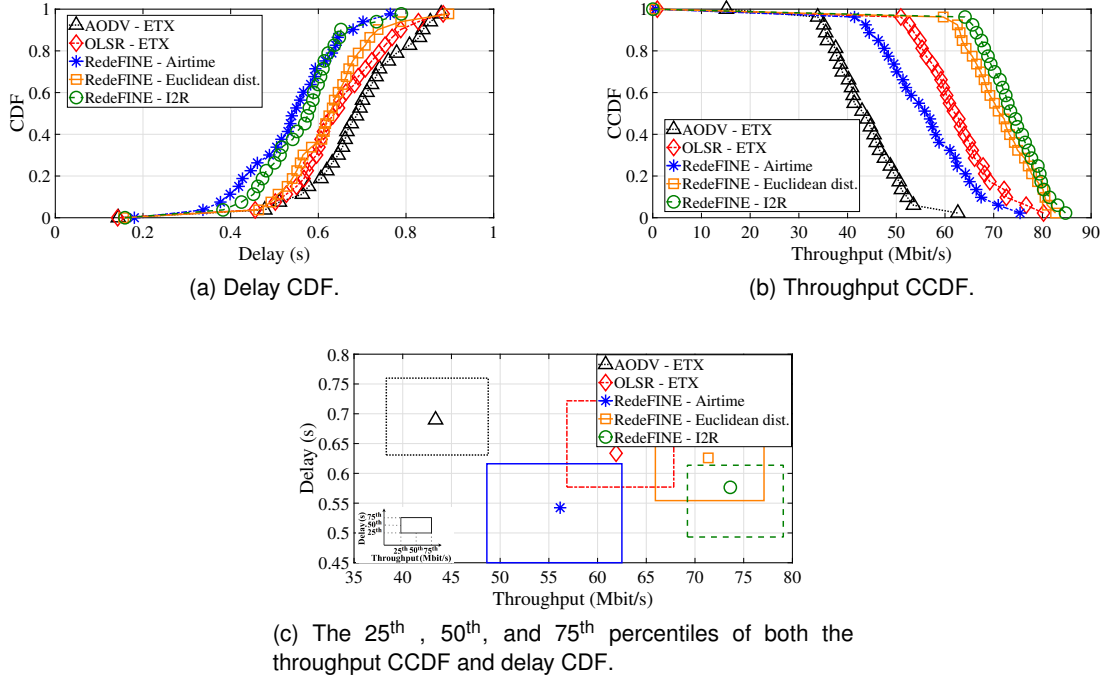


Figure 5.30: Results for throughput and delay in the FGW. The results were obtained considering 10 FAPs generating traffic, and $\alpha = 1$.

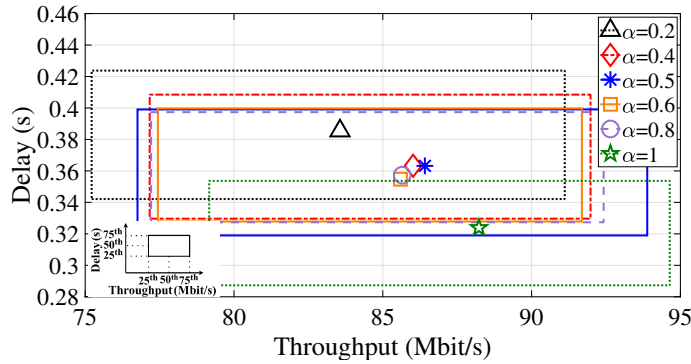


Figure 5.31: The 25th, 50th, and 75th percentiles of both the throughput CCDF and delay CDF, considering different α values. The results consider 5 FAPs generating traffic.

5.7 Discussion

The performance evaluation of RedeFINE shows that it allows improving the performance of an FMN, especially regarding aggregate throughput and PDR. RedeFINE outperforms AODV and OLSR, even when they use the link quality-based metric ETX. These results are due to the selection of stable paths with minimal length, formed by links with high SNR, and consequently

with high capacity and low delay. RedeFINE potentially allows for even better performance in scenarios where the UAVs moves at higher speeds. In those scenarios, we concluded that the distributed routing protocols are not be able to update the routing tables frequently enough to properly react to the topology changes.

The main idea behind RedeFINE is the ability to know in advance the future state of the FMN, based on the UAV placement computed at the Central Station. In order to achieve similar results using a distributed routing protocol, an option would be to reduce the interval between *HELLO* packets, so that the routing tables could be updated frequently enough to support high mobility scenarios. However, the price paid would be the reduction of the bandwidth available for data traffic. RedeFINE does not use probe packets for neighbor discovery and link sensing, and the routing tables are computed centrally, allowing to reduce the computational power on board of the UAVs; control traffic consists only of the routing tables sent from the Central Station to the UAVs. Nevertheless, the overhead of the control protocol must be studied in the future, in order to define the optimal period for sending the routing tables from the Central Station to the UAVs; it must guarantee the synchronism of the whole system, but simultaneously taking into account the bandwidth wasted. Additionally, RedeFINE must be concerned with the stability of the system, avoiding changing the routing tables in the same UAV in short time intervals, mainly when no significant gains are anticipated.

Moreover, the I2R metric outperforms the link quality-based routing metrics ETX and Airtime, as well as the Euclidean distance routing metric. This can be justified by the fact that the ETX and Airtime routing metrics provide estimations based on small probe packets. As consequence, the link loss ratio is underestimated, especially if high data rates are used, as IEEE 802.11ac does. When the Euclidean distance is used as routing metric, it may result in multiple communications nodes close to each other transmitting simultaneously, causing inter-flow interference within the FMN. In the worst-case scenario, I2R enables better delay than the Euclidean distance routing metric for the same total amount of bits received in the FGW. RedeFINE using the I2R metric represents an inter-flow interference-aware approach that estimates in advance the number of neighboring UAVs without using probe packets for neighbor discovery, link sensing, and interference estimation, based on the decisions performed by the Central Station that defines the FMN placement. Since forwarding tables are computed centrally, the computational power on board of the UAVs can be reduced. In its current version, I2R does not consider the amount of traffic generated by the UAVs and intra-flow interference. In addition, the improvement of I2R for an FMN consisting of multiple FGWs is worth be considered.

It is worth mentioning that the algorithms and mechanisms proposed in this thesis were evaluated as independent solutions, in order to assess the performance gains achieved by each one. This is in line with the proposed AM2Fly solution, which lies in independent Network Functions to manage a service-based FMN. However, in order to discuss the combination of the RedeFINE

mechanism, the GPQM algorithm, and a state of art centralized FAP placement algorithm as a potential holistic solution, we can consider the networking scenario depicted in Fig. 5.19. In this network scenario, FAP_1 , FAP_2 , and the FGW are in fixed positions, while FAP_3 is moving. During its movement, FAP_3 first uses FAP_1 and then FAP_2 to reach the FGW through two independent multi-hop communications paths. RedeFINE runs after the GPQM algorithm, which is responsible for defining: 1) the position of the FGW considering the positions of FAP_1 and FAP_2 , defined by a state of the art FAP placement algorithm, as well as 2) the queue size for each FAP, in order to allow stochastic delay guarantees. The 3D space in which the FGW can be placed results from the intersection of the spheres associated with FAP_1 and FAP_2 , each with a radius equal to the minimum Euclidean distance that allows guaranteeing the targeted SNR level to induce an MCS index with a data rate capable of accommodating the offered traffic; the greater the radius of the spheres the greater the allowed intersection volume for placing the FGW.

Let us focus on FAP_3 . It may happen that: a) FAP_3 is capable of being connected directly to the FGW at certain positions along its trajectory, if the sphere associated with FAP_3 virtually intercepts the spheres of FAP_1 and FAP_2 , while the FGW is placed at the intersection of the spheres associated with the three FAPs; b) there is no possibility of a direct wireless connection between FAP_3 and the FGW, during the movement of the former. In this case, FAP_1 and FAP_2 are used as flying communications relays for FAP_3 . As the traffic offered by FAP_3 is carried by FAP_1 or FAP_2 , the optimal position of the FGW at the beginning and end of the trajectory can be different; the FGW must be closer to FAP_1 at the beginning of the trajectory and closer to FAP_2 at the end of the trajectory, due to the increase in the amount of traffic they forward to/from FAP_3 in each case, according to the forwarding tables defined by RedeFINE. As the movement of the UAVs is not instantaneous – in practice it is limited by physical constraints, the FGW position only allows improving the network performance in the initial and final instants; this is in line with the model used, which assumes discrete time instants associated with FMN's snapshots. RedeFINE enables seamless FMN reconfiguration by selecting high-capacity communications paths between the UAVs while simultaneously avoiding disruptions in the communications paths. However, for the reasons mentioned herein, it may not allow ensuring QoS guarantees during the FMN's reconfiguration period. To meet the targeted QoS requirements when the UAVs are moving to reconfigure the FMN, GPQM and RedeFINE should be configured with a period compliant with the speed of the UAVs, as it defines how they change their positions over time. On the other hand, the period defined for computing up-to-date decisions should consider the time needed to compute up-to-date decisions; if the update period is set to a low value, it may not be enough to compute up-to-date decisions. Moreover, despite a low period theoretically leads to an FMN more prone to react to the dynamic wireless channel's conditions, in practice it may not lead to performance gains, since the FMN's reconfiguration is transitory and relatively short, when compared to the longer interval in which the UAVs are fixed in QoS-aware positions.

5.8 Summary

FMNs impose significant routing challenges, especially when the UAVs move for repositioning themselves and reconfiguring the FMN according to the decisions of centralized UAV placement algorithms. On the one hand, due to the dynamic SNR of the wireless links, disruptions of the inter-UAV communications paths may occur. Moreover, the network capacity provided by the communications paths may be insufficient to accommodate the offered traffic. On the other hand, if many UAVs are placed close to each other to maximize the network capacity offered, there is inter-flow interference between neighboring UAVs, which degrades the network performance.

In order to enable inter-flow interference-aware, high-capacity wireless links between neighboring UAVs without loss of connectivity, this chapter presented RedeFINE, a centralized routing solution for FMNs. RedeFINE takes advantage of the holistic knowledge provided by a Central Station to predict the future state of the network, in order to perform the routing decisions in advance. Based on an analysis of a reference case and the simulation results for two representative scenarios, we demonstrated the superior performance of RedeFINE regarding throughput and PDR when compared to static routing, AODV, and OLSR. The results showed gains up to 54%. Regarding delay, RedeFINE did not achieve the values obtained for its state of the art counterparts; however, the longer delay is justified by the higher PDR achieved with RedeFINE, which increases network congestion and the time packets are held in the transmission queue. Also, this chapter presented an interference-aware routing metric especially designed for centralized routing in FMNs. I2R does not require any probe packets and enables the configuration of paths with minimal Euclidean distance formed by UAVs with the lowest number of neighbors in carrier-sense range, thus minimizing inter-flow interference in the FMN. Based on the simulation results for five random scenarios and different number of FAPs generating traffic, we demonstrated the superior performance of I2R with respect to the routing metrics ETX, Airtime, and Euclidean distance. Simulation results show that I2R improves the performance of RedeFINE up to 22% in delay and up to 7% in throughput.

Until this chapter, we have assumed that a shared wireless channel is employed to maximize the aggregate network performance, considering an FMN consisting of a single FGW. In networking scenarios where a single FGW using a shared wireless channel does not allow for high enough capacity to accommodate the traffic offered and meet the delay constraints, this approach may compromise the QoS levels offered by the FMN. In the next chapter, we propose an algorithm for the placement and allocation of the communication resources, namely the number of UAVs, required to meet heterogeneous QoS levels using a shared physical FMN.

Chapter 6

An Algorithm for Placement and Allocation of Communications Resources in Slicing-aware Flying Multi-hop Networks

In the design of the algorithms and mechanisms proposed in the previous chapters, we assumed a Flying Multi-hop Network (FMN) where a single shared wireless channel is employed to improve the aggregate network performance, considering a single Flying Gateway (FGW). In networking scenarios where a single FGW using a shared wireless channel does not allow for high enough capacity to accommodate the traffic demand and meet the delay constraints, this approach may compromise the Quality of Service (QoS) levels offered by the FMN. In addition, the algorithms and mechanisms proposed in the previous chapters do not address the joint Flying Access Point (FAP) and FGW placement problem.

Network slicing emerged in 5G networks to enable the coexistence of multiple virtual networks on top of a shared physical communications infrastructure, while supporting different services with heterogeneous QoS levels [192]. In order to enable slicing-aware FMNs, in this chapter we propose the Slicing-aware Flying Access and Backhaul (SurFABLE) algorithm. SurFABLE carries out the joint placement and allocation of communications resources so that the cost of deploying a slicing-aware FMN is minimized. Based on the information provided by a Service Level Agreement (SLA), namely target QoS levels and coverage area, SurFABLE allows determining the minimum number of FAPs and FGWs required, their 3-Dimensional (3D) positions, and the number of wireless channels to be made available for any number and type of network slices.

We start the chapter by reviewing the state of the art on network slicing, with a special focus on solutions for slicing-aware FMNs. Then, we describe the system model and formulate the problem. Afterwards, we explain the SurFABLE algorithm, refer to its performance evaluation, which was carried out by means of Network Simulator 3 (ns-3) simulations and an experimental

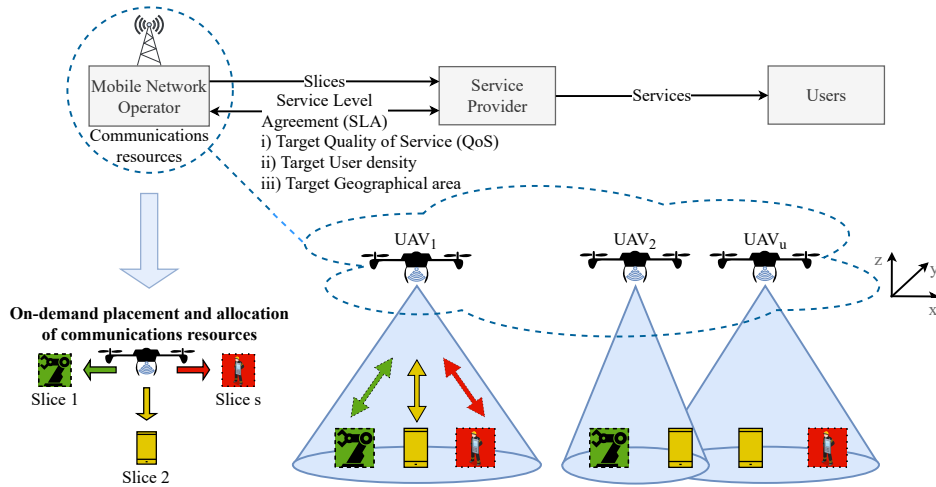


Figure 6.1: Flying network, consisting of Unmanned Aerial Vehicles (UAVs), used by a Mobile Network Operator to provide network slices to a Service Provider that offers services to ground devices in target geographical areas [38].

testbed, and discuss the design and performance results achieved by SurFABLE. The chapter ends with a summary of the contents presented.

6.1 Related Work

In this section, we begin by providing an overview of network slicing, and then place special emphasis on the state of the art solutions for slicing-aware FMNs. The network slicing concept allows that Mobile Network Operators lease portions of the communications resources that compose their network infrastructure, including data centers, cell sites, and fronthaul and backhaul networks, to Service Providers or Virtual Mobile Network Operators which offer services with targeted QoS to their customers, acting as network users; this is depicted in Fig. 6.1. Network slices are characterized at a high-level in the form of an SLA, which is established between Service Providers and Mobile Network Operators. An SLA specifies the service requirements for a network slice including QoS levels, number of devices served, and geographical areas covered [193]. Network slicing allows reducing costs by avoiding the need for permanently assigned communications resources, from the Radio Access Network (RAN) to the core network, while enabling improved flexibility, scalability, and efficiency in the management of the communications infrastructures over time [194].

The literature on network slicing has been mainly focused on ensuring fairness and efficient resource management, while meeting the targeted service requirements. Yet, a fixed communications infrastructure is typically assumed [193], which may not be feasible in scenarios characterized by limited network resources or dynamic behavior. The use of flexible wireless networks, ca-

pable of providing coverage and communications resources on-demand, is a promising approach for Mobile Network Operators, which may be able to meet the Service Providers requirements in a myriad of application scenarios. In disaster management scenarios, such as forest fires and floods, the first responders require reliable, mission-critical communications, while a best-effort wireless service may be provided to the victims. In crowded scenarios, such as outdoor festivities, wireless connectivity with different QoS requirements must be ensured for the staff, the media, and the spectators. Nowadays, these scenarios are addressed by Mobile Network Operators using network over-provisioning with fixed and mobile Base Stations (BSs) deployed on ground vehicles (Cell-on-Wheels concept [70]). However, these BSs may collapse, fail to provide sufficient wireless coverage, or lack the required communications resources. Moreover, the unpredictability associated to the number of devices and their mobility may degrade the QoS offered.

The 3rd Generation Partnership Project (3GPP) defines a network slice as a logical network that provides specific network capabilities and target performance requirements, such as throughput and delay [192]. The 3GPP standardization process on network slicing has been approached in [195], while key principles, enabling technologies, and open research challenges related to network slicing have been envisioned in [196]. In these works, network slicing has been considered a major enabler for 5G networks, but the provisioning of network slices on-demand and the usage of UAVs to ensure a reconfigurable wireless network have not been envisioned.

Multiple approaches have been proposed in the literature to allow for QoS guarantees in wireless networks. A reference example has been introduced by IEEE 802.11e, which employs a contention-based channel access method, called Enhanced Distributed Channel Access, to differentiate traffic with different priorities [197]. In IEEE 802.11e Enhanced Distributed Channel Access, traffic differentiation is achieved through the assignment of different contention parameters to each Access Category. An IEEE 802.11e Enhanced Distributed Channel Access-based Station (STA) is able to accommodate a maximum of eight device priorities that are mapped to four Access Categories. Each Access Category competes for channel access with a unique Arbitration Inter-Frame Space and Contention Window configuration. Unlike Distributed Coordination Function, which uses the Distributed Coordination Function Interframe Space for a STA to access the channel, Enhanced Distributed Coordination Function employs a different Arbitration Inter-Frame Space for each Access Category. As depicted in Fig. 6.2, each Access Category is implemented by means of distinct queues per STA supporting IEEE 802.11e Enhanced Distributed Channel Access. Each queue acts as a virtual STA and competes independently for channel access. When a collision between separate queues of the same STA occurs, the channel access is assigned to the queue with the highest priority while the lower priority queues contend. On the one hand, IEEE 802.11e Enhanced Distributed Channel Access provides the same QoS to flows associated to the same traffic type, since only four virtual queues are available per STA. On the other hand, it may be unable to meet the targeted QoS guarantees, for instance if the capacity of the wireless links is

not enough or the network configuration employed is not adequately defined. For these reasons, IEEE 802.11e Enhanced Distributed Channel Access does not provide high enough scalability and the flexibility to meet the QoS levels of different network slices, especially when the number and type of network slices increase.

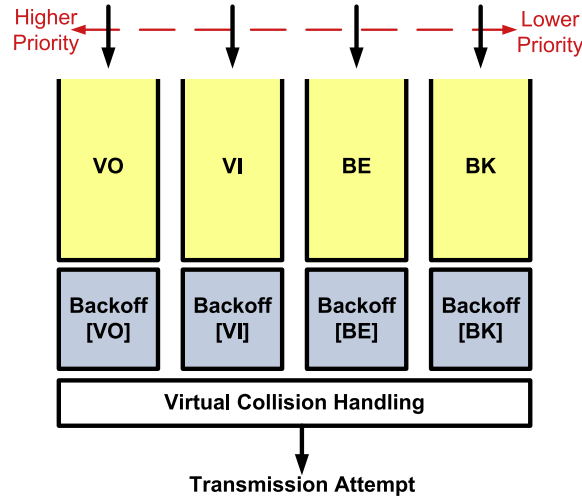


Figure 6.2: Illustration of IEEE 802.11e Enhanced Distributed Channel Access, where four Access Categories with different priorities (descending from the left-hand side to the right-hand side) are associated with voice (VO), video (VI), best-effort (BE), and background (BK) traffic (modified from [39]).

In [40], network slicing has been formulated as a bi-convex¹ optimization problem that considers backhaul network capacity, available storage, and achievable bandwidth, as depicted in Fig. 6.3. It aims at minimizing: 1) the aggregate network slice load per BS and 2) the delay for all network slices. The authors have designed two alternative minimization algorithms to solve the problem. The first is a centralized scheme suitable for 5G network virtualization and orchestration technologies, whereas the second lies in a distributed approach in which all network slices coordinate between themselves to reach a solution for the problem. Despite the first algorithm does not guarantee global convergence, numerical simulations have allowed the authors to conclude it allows achieving a global optimal solution, when compared with exhaustive search in a simple networking scenario consisting of two tenants. Moreover, the results demonstrate that, as the number of tenants increases, the algorithm allows achieving the same performance as the IpOPT solver [137]. However, a single BS in a fixed position has been considered. The placement of the BS in suitable positions, according to the network slice requirements, has not been addressed.

¹Bi-convex is a generalization of convex optimization where the objective function and the set of constraints can be biconvex. In a convex function, the line segment between any two points on its graphical representation lies above those two points.

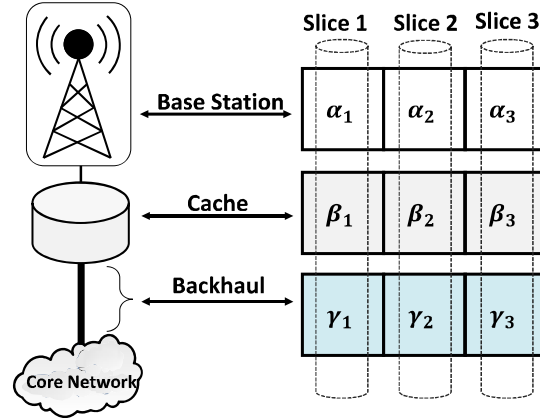


Figure 6.3: Network slicing architecture where three network slices share the physical network resources available, including a BS with cache storage connected to the Core network through a capacity-constrained backhaul link. The constants α , β , and γ represent the fractions of, respectively, total wireless channel bandwidth (in Hz), cache storage (contents), and backhaul link capacity (in bit/s) allocated to each network slice [40].

In [41], the authors have proposed different spectrum sharing strategies for a 5G slicing-aware network. Additionally, a business-driven model for determining network resource requirements per tenant has been proposed. Following the model, multiple spectrum allocation approaches for RAN slicing have been created, considering varying levels of resource isolation, as depicted in Fig. 6.4. The approach focus on capacity-oriented guarantees based on resource allocation rather than throughput-oriented guarantees. Each approach enables that the infrastructure provider has different degrees of flexibility for allocating the network resources. Considering snapshots of a 5G slicing-aware network consisting of small cells, the spectrum allocation approaches have been assessed. The results show that no-isolation solutions provide the highest network performance due to more efficient resource consumption. Approaches based on per-network capacity guarantees, rather than per-cell guarantees, allow for greater adaptability to geographical traffic changes, while leading to higher performance for low network resource usage and high traffic correlation between different network slices. Highly correlated traffic may be associated with network slices that provide different communications services from the same tenant or the same communications service from different tenants. Poorly correlated traffic is typically associated with network slices owned by different tenants that provide heterogeneous communication services. On the other hand, intra-cell or inter-cell isolation-based approaches allow providing equal guarantees against inter-slice interference. However, for high resource occupancy and low traffic correlation, intra-cell isolation results in improved performance due to a greater flexibility to meet the demand of the network slices. Multiple BSs have been considered, but their positioning in suitable positions has not been studied too.

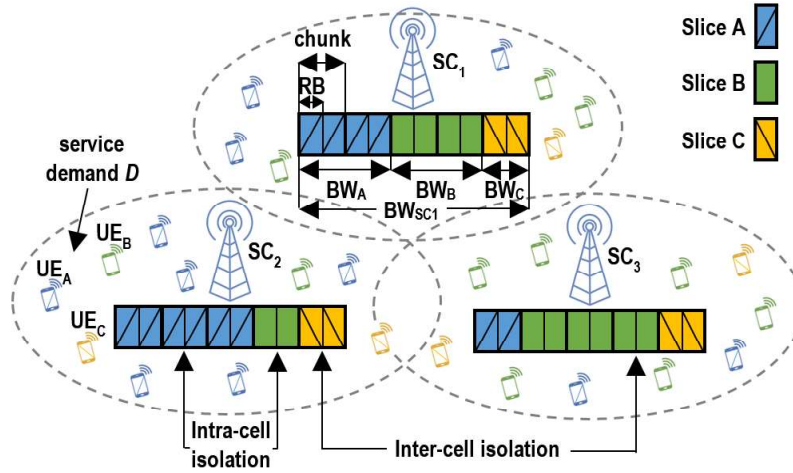


Figure 6.4: Illustrative resource allocation approach with different isolation levels proposed in [41] for three network slices.

Providing network slices, taking into account wireless coverage and the traffic demand requirements of the RAN, is a problem that has been addressed in [193, 194, 198]. Still, a fixed wireless communications infrastructure has been considered and the joint placement and allocation of communications resources on-demand has not been addressed. Due to the large number of variables involved, network slicing problems are complex and typically solved using either optimization [199], game theory [200], and evolutionary and heuristic algorithms [201, 202], which have been proposed for managing the network slices' lifecycle, including preparation, planning, and operation, while ensuring efficient resource management. Moreover, a special emphasis has been given to Artificial Intelligence and Machine Learning (ML) based techniques [57, 203–206], specifically for predicting resource requirements according to the network traffic load, minimizing the number of resources to be allocated, and reducing costs. Still, the authors have raised scalability and practicability challenges for applying Artificial Intelligence and ML based techniques to perform network slicing, especially regarding the high volume of data, massive number of devices, variable network size, and high time complexity when large data sets are employed; these challenges are exacerbated when a centralized based approach is employed. This motivates the design of heuristic-based solutions with less complexity.

When there is no wireless network infrastructure available or there is a need to enhance the coverage and capacity of existing networks, flying networks, made up of UAVs acting as mobile BSs and Access Points (APs), emerged as an adequate solution. For that purpose, different UAV placement algorithms have been proposed, in order to improve the QoS offered or maximize the coverage and the number of ground devices served, as referred to in Chapter 2. However, the state of the art solutions typically lie on a best-effort approach or are optimized for a single com-

munications service. Due to the particular characteristics of flying networks when compared with terrestrial networks – namely the improved wireless channel reliability promoted by the potential absence of obstacles above the ground –, flying networks have been envisioned in the literature as an enabler for Ultra-Reliable Low-Latency Communications (URLLC) communications in 6G networks [207]. Yet, the joint UAV placement and allocation of communications resources, while enabling the usage of different communications services with heterogeneous QoS requirements, has not been addressed for multiple UAVs so far. In the survey of the state of the art presented herein, we give special emphasis to slicing-aware solutions for flying networks.

In [208], a flying network consisting of a High-Altitude Platform Station (HAPS) and Low Altitude Platforms (LAPs) has been proposed to extend wireless coverage and improve the network performance offered to ground devices, as depicted in Fig. 6.5. The total data rate has been maximized by optimizing the downlink and uplink communications. For the downlink, the authors have optimized the altitude of the LAPs and the spectrum allocated for the wireless links established between the HAPS, the LAPs, and the ground devices served. For the uplink, the authors have taken into account a decoding threshold for the Successive Interference Cancellation technology employed and have optimized the power allocation for each ground device. The spectrum allocated for the ground devices, the LAPs, and the HAPS has also been considered. For downlink optimization, the sub-carrier allocation problem for the wireless links established between the HAPS and the LAPs has been solved by means of the Hungarian algorithm. In turn, the sub-carrier allocation for the wireless links established between each LAP and the ground devices has been solved using relax-and-round as well as Sequential Convex Approximation based techniques. The Fibonacci Search technique has been employed to optimize the altitude of the LAPs. Finally, an iterative optimization approach, in which the altitude and the spectrum allocation are optimized sequentially, has been designed. For uplink optimization, the spectrum allocation problems for the wireless links established between the ground devices and each LAP serving them has been addressed by means of the Greedy method. Similarly, the Hungarian method has been used to solve the spectrum allocation for the wireless links established between the LAPs and the HAPS. Moreover, the authors have investigated the optimal power allocation for the ground devices based on the optimized spectrum allocation and considering the Successive Interference Cancellation decoding threshold. Even when the altitude of the LAPs and the resource allocation are not optimized, through simulation results the authors have concluded that the proposed cooperative network outperforms a single HAPS-based overlay network. In addition, by improving the altitude of the LAPs and spectrum allocation in the downlink, the total data rate has been improved by 43% and 32%, respectively. The authors have also studied the influence of Non-Orthogonal Multiple Access on network performance in the uplink by optimizing spectrum and power allocation, for which the total data rate has been improved by 13% compared to Orthogonal Frequency-Division Multiple Access (OFDMA). However, the solution does not al-

low meeting heterogeneous QoS requirements. Moreover, since only the altitude component has been considered, the placement computed may not allow maximizing the QoS offered in practice and imply the usage of more radio resources when compared with a solution that defines a suitable placement in 3D space.

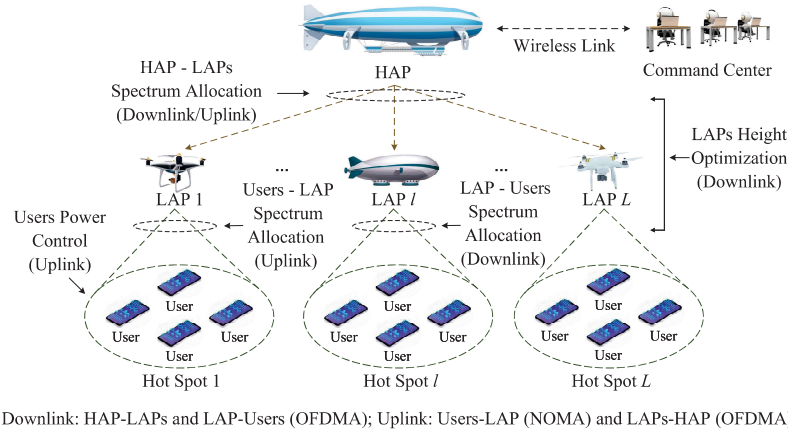


Figure 6.5: Cooperative network consisting of a HAPS and LAPs extending wireless coverage and improving the network performance offered to ground devices.

In [42], a network scheme for performing network slicing in a 5G vehicular network has been proposed to enhance the allocation of communication resources. A network architecture consisting of UAVs acting as Aerial Relay Nodes and Road-Side Units (RSUs) providing Resource Blocks (RBs) to ground devices has been proposed; it is depicted in Fig. 6.6. The position of each Aerial Relay Node is optimized by employing a fuzzy multi-attribute decision-making method. Each Aerial Relay Node has a set of RBs, whilst each RSU maintains Local Virtual Resource Pools of Local RBs and Shared RBs. Moreover, a Software-Defined Networking (SDN) controller has a global Virtual Resource Pool with information about the Shared RBs of the RSUs. The QoS and Signal-to-Interference plus Noise Ratio of the ground devices are monitored to determine the satisfaction level of the services provided. If it exceeds a predefined threshold value, the remaining Local Resource Blocks of the Aerial Relay Nodes or the RSUs, respectively, are able to meet the service needs; otherwise, the Aerial Relay Nodes or RSUs provide more RBs from the Local Virtual Resource Pool or Virtual Resource Pool, respectively. The performance evaluation carried out through ns-3 simulations has demonstrated that greater performance is achieved by optimizing the placement of the Aerial Relay Nodes. The solution has been designed to provide different network slices, but the authors have been focused on maximizing the aggregate network performance for each network slice without considering differentiated QoS guarantees.

A coverage-aware geometric placement approach for a UAV has been proposed in [43]. It consists of a Dynamic Standalone Drone Small Cell, employing a Dynamic Horizontal Opportunis-

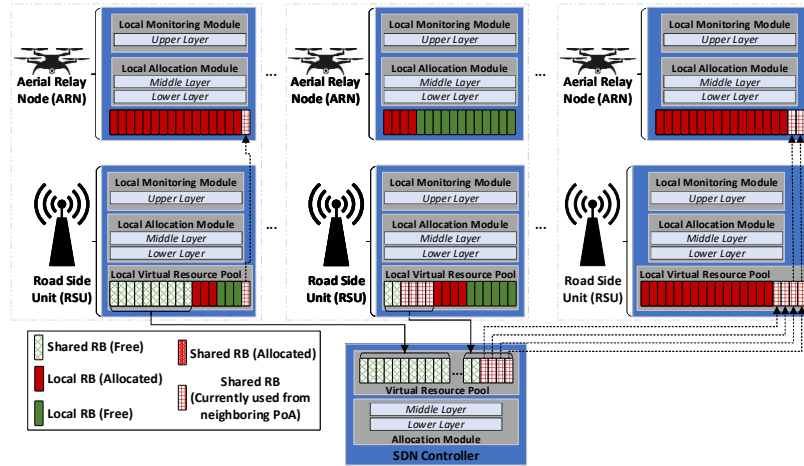


Figure 6.6: Network slicing-aware architecture proposed in [42]. It consists of UAVs, acting as Aerial Relay Nodes, and RSUs, managed by an SDN controller, which allow providing RBs to ground devices.

tic Positioning technique, for two network slices serving massive Machine-Type Communications (mMTC) and enhanced Mobile Broadband (eMBB) based services, as depicted in Fig. 6.7. eMBB has been assumed to have high priority and reliability requirements, while coexisting with random access mMTC-based ground devices. The objective has consisted in maximizing the average data rate for eMBB-based ground devices by jointly optimizing the network slicing allocation by priority class as well as the UAV altitude, cell size, and antenna's beamwidth. The wireless resources and UAV deployment have been addressed solving a geometry-based problem considering random locations for the eMBB-based ground devices. The UAV's position has been defined in 3D. Yet, since the usage of directional antennas limits changes in altitude, a constant altitude has been considered. The priority class associated with the ground devices and their location have been assumed to be known by the UAV, in order to position itself in the center of the Smallest Bounding Circle containing all the active eMBB-based ground device. Dynamic Horizontal Opportunistic Positioning allows achieving the shortest horizontal distance between the UAV and any ground device. The authors have shown that using a UAV allows improving the propagation environment by eliminating large scale shadowing induced by obstacles blocking Line-of-Sight with the ground devices. Through *Monte Carlo* simulations, they have demonstrated that their network slicing-aware approach outperforms a static deployment consisting of fixed LAPs or HAPs, especially in urban scenarios associated with high reliability requirements. However, Dynamic Horizontal Opportunistic Positioning is tailored for use cases with two network slices only, where a single UAV is deployed to maximize the average data rate of ground devices belonging to an eMBB network slice coexisting with ground devices associated with an mMTC network slice.

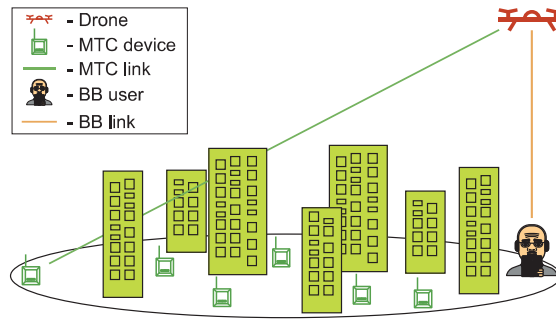


Figure 6.7: UAV offering Machine-Type Communications (MTC) and broadband services to ground devices [43].

In [44] the authors present a follow-up of the work developed in [209], consisting in a proactive network slicing-aware approach for URLLC and eMBB multiplexing in flying networks. Taking into account mobile ground devices and limited UAV wireless coverage, a dynamic flying network has been employed to improve the QoS offered. The authors have formulated the network slicing problem in flying networks as a sequential decision problem with the objective of enabling energy-efficient eMBB based services with fairness, while meeting URLLC based requirements for the UAV control and non-payload information exchange. Assuming there is a lack of accurate information about the location of the ground devices and channel gain models, a distributed learning method has been used to predict the ground devices' locations. Moreover, they have proposed a proactive approach to mitigate deviations between the network slices' supply and network slices' demand, considering the future location of the ground devices. Based on a Deep Neural Network, they have built accurate and analytically tractable channel gain models, which are highly dependent on mobile ground devices and UAVs, while having a powerful non-linear function approximation ability. Taking into account the Lyapunov approach, the problem has been decomposed into multiple mixed-integer non-convex optimization sub-problems. In order to address the mixed-integer and non-convexity nature, an iterative optimization approach and a sequential convex approximation technique have been used, respectively; this is depicted in Fig. 6.8. The authors have validated the learning techniques' accuracy as well as the effectiveness of the Lyapunov-based optimization approach through simulations. Still, the work has been focused on ensuring aggregate QoS guarantees. It is unable to meet heterogeneous QoS levels for different flows associated with the same traffic type. Moreover, the joint placement and allocation of communications resources considering multiple UAVs has not been addressed.

In [45], the authors have addressed uplink radio resource slicing using multiple Drone Small Cells. They have proposed a hybrid multi-Drone Small Cell placement and device association approach focused on reducing the amount of resources used while considering heterogeneous QoS requirements associated with mobile ground devices and machine-type communications services.

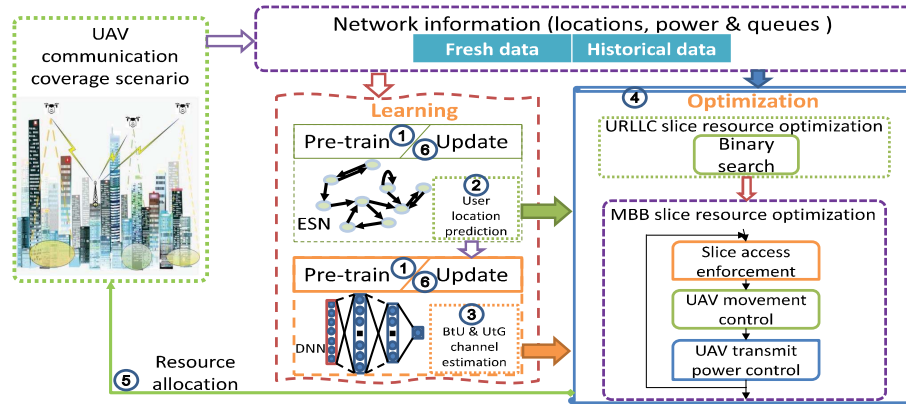


Figure 6.8: Illustration of the iterative optimization approach with a successive convex approximation technique proposed in [44] for slicing-aware flying networks.

To define the multi-UAV 3D placement, an interference-aware graph model has been considered. The authors have formulated the problem as a joint optimization of multi-UAV placement and ground devices' association for uplink resource slicing with heterogeneous QoS requirements. Given the NP-hard nature of the problem, they have evaluated how multi-Drone Small Cells affect resource usage. A two-tier uplink heterogeneous RAN has been considered, in which a single Macro BS has been deployed in the center of the network and overlaid by multiple Drone Small Cells, as shown in Fig. 6.9. The high transmission power of the Macro BS aims for wide-area wireless coverage. Drone Small Cells with low transmission power and a limited wireless coverage have been positioned in the 3D coverage area enabled by the Macro BS. When covered by a Drone Small Cell, the ground devices can connect to either the Drone Small Cell or the Macro BS. The offered traffic is suitable to be sent through three types of wireless links: 1) ground devices to Macro BS (G2M); 2) ground devices to Drone Small Cells (G2D); and 3) Drone Small Cells to Macro BS (D2M). The authors have designed an approach with fine-tunable complexity based on adaptive control policy for the Drone Small Cell's altitude, in order to balance uplink coverage and resource consumption, while dynamically adjusting the Drone Small Cell deployment for improved performance. A heuristic approach, considering constant flying altitude while maximizing the wireless coverage, has been proposed for determining approximate solutions to the problem with reduced complexity. The joint problem has been transformed into a weight-based *clique* problem² in each approximation, while a performance-aware *clique* algorithm has been considered to solve the problem. Numerical simulation results for low and high device density scenarios have been carried out to evaluate the solution's performance, thus demonstrating that it outperforms different state of the art approaches. Still, the work has been focused on the access network only

²A *clique* problem lies in finding *cliques* in a graph. *Cliques* are subsets of vertices, adjacent to each other, which define complete subgraphs.

and a fixed altitude has been considered to reduce the problem complexity, which may limit the quality of the solutions obtained.

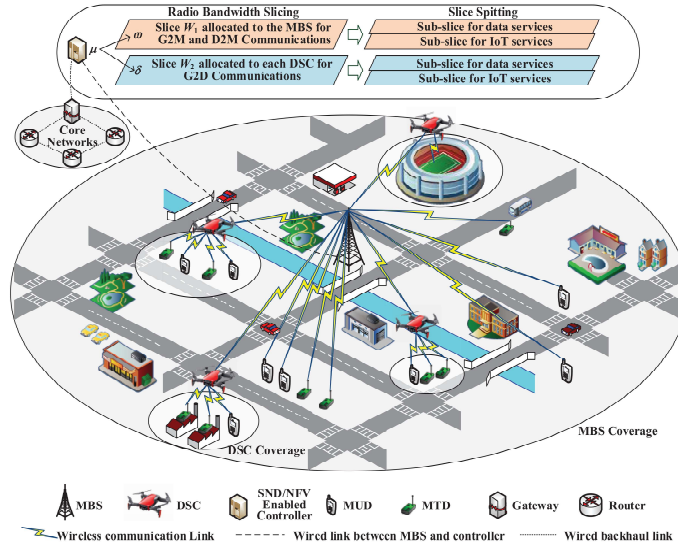


Figure 6.9: Two-tier heterogeneous RAN where a single Macro BS (MBS), deployed at the center of the network, is aided by multiple Drone Small Cells (DSCs). Due to the higher Macro BS's transmission power, it provides wireless coverage over a large area; the Drone Small Cells, which allow for a short wireless range due to their lower transmission power, are placed within the 3D coverage area enabled by the Macro BS [45].

In [55], a UAV has been used as a flying communications relay to enable a URLLC-based wireless link between a BS and a ground device with Line-of-Sight compromised due to an obstacle between them, as depicted in Fig. 6.10. The wireless link established between the BS and the ground device is referred to as the downlink link, which is used to carry control packets for the ground device performing a delay-sensitive task. The downlink's latency and decoding error probability have been controlled in order to move the ground device in a timely and reliable manner. On the other hand, the uplink is used to carry the large amount of data acquired by the ground device. The authors have been focused on maximizing the average uplink's data rate while meeting the URLLC requirements in the downlink. The problem has been divided into three sub-problems: time allocation, bandwidth allocation, and UAV position optimization. Due to the complexity of these sub-problems, only approximate solutions have been determined, where an iterative approach for solving the formulated problem has been proposed. Numerical simulation results have allowed the authors to conclude that the proposed solution, based on partial optimization, allows achieving performance gains. However, besides ensuring a single URLLC network slice using only one UAV, the iterative nature of the proposed algorithm leads to approximate solutions only.

An optimization-based approach, using a single UAV to maximize the average data rate and

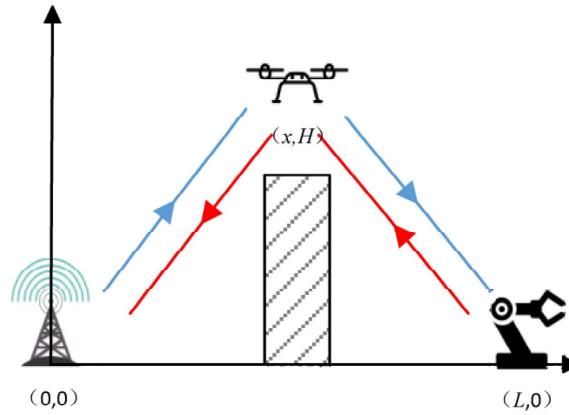


Figure 6.10: UAV, acting as a flying communications relay, enabling URLLC for a ground device with Line-of-Sight towards a BS blocked by an obstacle.

minimize the transmission power in out of coverage areas, has been proposed in [46], as depicted in Fig. 6.11. The proposed approach aims at meeting the requirements of URLLC-based services. The authors have used the Gaussian Process Regression technique for capturing the network dynamic behavior and estimating URLLC-based traffic characteristics in an online manner. Gaussian Process Regression is an active learning method suitable for capturing uncertainty in time varying processes. It allows the UAVs predicting latency-sensitive data packets associated with Internet of Things (IoT) devices, while performing efficient resource allocation, in terms of RBs and transmission power, and defining an optimal UAV deployment approach. Unlike other state of the art prediction algorithms, such as Long Short-Term Memory, Gaussian Process Regression is suitable for small datasets with random components. Considering the strict constraints associated with the URLLC requirements, the authors have formulated a joint optimization problem for maximizing the average aggregate data rate while minimizing the transmission power of the UAV. The problem has been considered NP-hard. As such, the authors have divided it into three sub-problems – transmission power, RB allocation, and UAV placement –, addressed with a low-complexity near-optimal sequential minimization approach. Based on *Monte-Carlo* simulation results, the authors have concluded that the solution allows for performance improvement up to 24% when compared with 1) a maximum power-based approach, which is the worst-case scenario regarding resource consumption, and 2) a random-based approach, which considers random placement for the UAVs.

In [47], a flying network consisting of UAVs, in which Service Providers rent wireless resources from Mobile Network Operators using virtualization technologies, has been proposed. The system model is depicted in Fig. 6.12. The authors have assumed that each Mobile Network Operator deploys a single UAV equipped with a communications node to enable wireless connectivity in areas where there is no terrestrial network infrastructure. The UAVs are assumed to operate

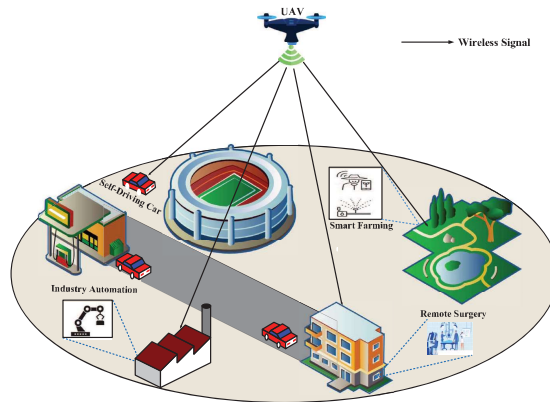


Figure 6.11: Wireless network consisting of a single UAV providing a URLLC-based communications service to a set of ground devices [46].

at different frequency bands, in order to avoid interference between the communications cells formed. Considering OFDMA, the total bandwidth available at each UAV has been subdivided into a set of sub-channels. The objective lies in maximizing the utility associated with Service Providers when it comes to the cost of the resources employed (sub-channels and transmission power), while meeting the QoS requirements of the mobile ground devices served. The authors have formulated a mixed-integer, non-linear and non-convex problem; due to its NP-hard nature, it has been divided into two sub-problems: 1) device association and 2) resource sharing. The authors have validated their solution through numerical simulations, concluding it outperforms the state of the art counterparts. However, the proposed solution assumes a single UAV for each Mobile Network Operator, which may not be sufficient to meet the QoS levels associated with the network slices to be made available. Finally, the authors have considered device association and resource sharing as two independent problems, which may compromise the quality of the solutions obtained when compared with a joint optimization approach.

In [48], the authors have proposed a dynamic Bandwidth Part allocation approach with two multiplexing methods: Dynamic Multiplexing and Orthogonal Slicing. The proposed approach aims at minimizing the effect of the URLLC traffic on the eMBB traffic in a flying network with different QoS requirements, especially for receiving ultra-reliable command and control flight messages with low latency as well as sending video streaming broadband traffic. Fig. 6.13 depicts the multiplexing process of URLLC traffic over eMBB traffic, considering the Dynamic Multiplexing and Orthogonal Slicing methods. eMBB traffic is considered to be scheduled in the time and frequency domains considering multiple RBs. For URLLC Dynamic Multiplexing-based, the packet transmissions are assumed to be scheduled as soon as packets arrive. The authors have considered that the Bandwidth Part size is equal to the bandwidth of a single carrier, which allows

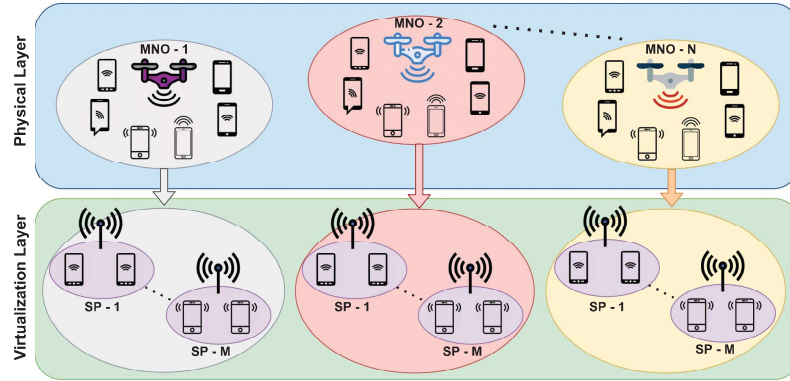


Figure 6.12: Network slicing-aware system model considered in [47], where a set of Service Providers rent wireless communications resources from a set of Mobile Network Operators. The wireless communications resources are made available by each Mobile Network Operator using a single UAV.

minimizing the impact of puncturing on the eMBB traffic³. As depicted at the top of Fig. 6.13 for a given Bandwidth Part, the URLLC traffic is considered to be scheduled uniformly within the allotted Bandwidth Part. The resources for URLLC and eMBB are orthogonally sliced in Orthogonal Slicing-based URLLC multiplexing, as depicted at the bottom of Fig. 6.13. As such, a predefined Bandwidth Part size is allocated to support URLLC traffic within the eMBB slot, while the remaining Bandwidth Part size uses the other carrier bandwidth resources available. The greater the URLLC traffic load the lower the eMBB throughput for both Dynamic Multiplexing and Orthogonal Slicing based methods. The tradeoff between eMBB throughput and URLLC multiplexing has been studied by the authors for ensuring effective Bandwidth Part allocation, taking into account the influence of physical layer parameters, such as Modulation and Coding Scheme (MCS) indexes and code block group size. Through *Monte Carlo* simulations, the authors have demonstrated that the Orthogonal Slicing-based method allows enhancing the eMBB throughput subject to a latency constraint for deterministic and predictable URLLC traffic. In contrast, Dynamic Multiplexing has a limited error-correcting capability against the puncturing effect of URLLC. Moreover, they have concluded that the MCS index used influences the prominence of the eMBB traffic over the URLLC traffic. The authors have focused on studying the influence of physical layer parameters and scheduling approaches to ensure the targeted QoS requirements in a flying network; however, they have not addressed determining the amount of communication resources required and the placement of the UAVs.

When it comes to network slicing in flying networks, the literature has been mostly focused

³A suitable state of the art puncturing strategy lies in exploiting the similarity between the symbols of the URLLC and eMBB traffic so that the eMBB symbols are similar to the URLLC symbols transmitted [210].

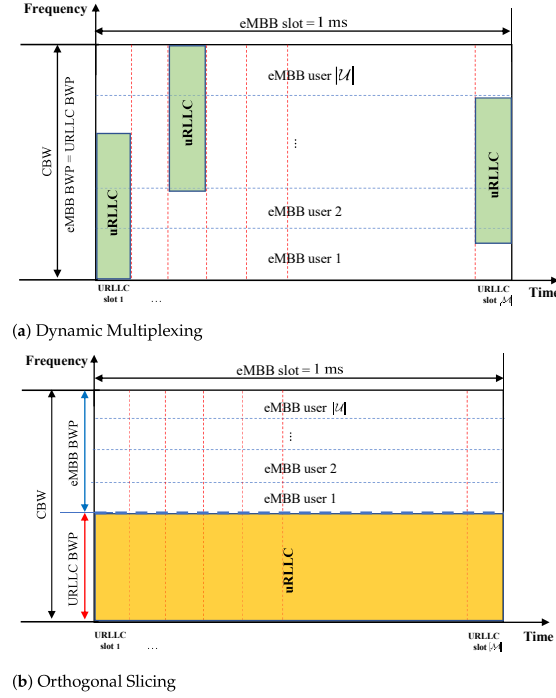


Figure 6.13: Network slicing approach proposed in [48], where two multiplexing methods – Dynamic Multiplexing and Orthogonal Slicing – are employed to enable URLLC and eMBB communications services.

on UAVs that act as clients of Mobile Network Operators' infrastructures providing wireless connectivity to ground devices [211]. However, the QoS levels offered have not been considered or differentiated. Conversely, this thesis proposes the integration of UAVs into Mobile Network Operators' infrastructures as communications resources that can be deployed anywhere, anytime, in order to meet the heterogeneous QoS requirements of any number and type of network slices, using a shared physical FMN.

6.2 System Model and Problem Formulation

The FMN considered herein allows extending wireless coverage using a multi-hop topology and offering multiple network slices with target QoS levels to ground devices located in given geographical areas. The FMN is assumed to be organized into a two-tier architecture, as depicted in Fig. 6.14. The FMN's architecture is similar to the one considered in the previous chapters of this thesis, but considers a network-slicing aware approach and addresses the joint placement of multiple FAPs and FGWs. Two types of UAVs compose the FMN: 1) the FAPs, which provide coverage-aware network slices to the ground devices; and 2) the FGWs, which forward the traffic to/from the Internet using wireless broadband links, while preserving the QoS guarantees of

the network slices offered by the FAPs to the ground devices. The first tier consists of the access network, which is composed of the FAPs; the second tier consists of the backhaul network, composed of FGWs.

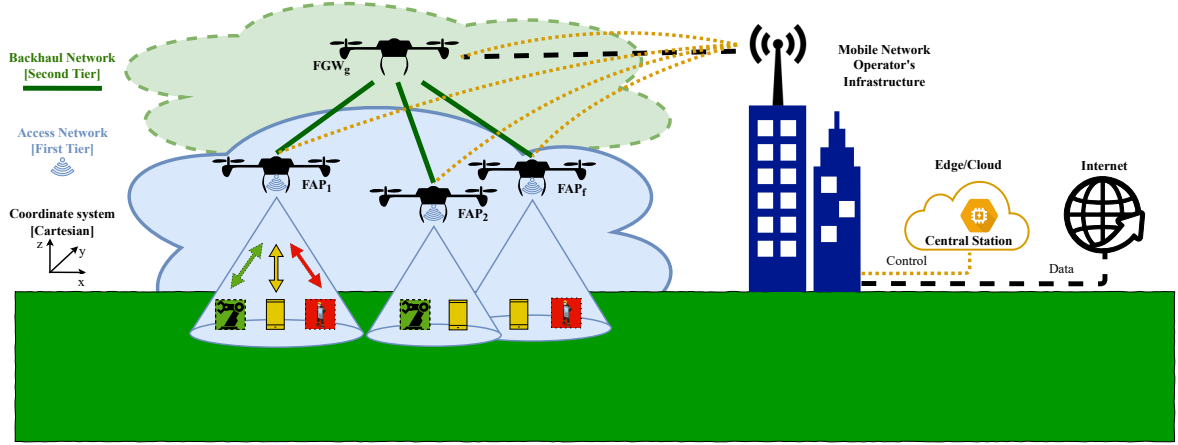


Figure 6.14: Slicing-aware FMN consisting of UAVs. The FAPs, in the first tier, provide a wireless access network to the ground devices, while the FGWs, in the second tier, establish a wireless backhaul network with the FAPs and forward the traffic to/from the Internet. The UAV placement and communications resources allocation are jointly computed in a Central Station, deployed in the Cloud or at the Edge of the FMN.

At the discrete time instant $t_k = k \cdot \Delta t, k \in \mathbb{N}_0$, where $\Delta t \gg 1$ s, the FMN can be reconfigured. Let $j \in J$ represent a communications node, being it a UAV $i \in V$ or a ground device $r \in R$ positioned within a cuboid C with dimensions X long, Y wide, and Z high, as depicted in Fig. 6.15, where UAVs $V \subset J$ and $R \subset J$.

Cuboid C is subdivided into a set of N equal and smaller fixed-size sub-cuboids, where $n \in N$ represents a sub-cuboid in which center a UAV may be located (cf. left-hand side of Fig. 6.15). Let $P_i = (x_i, y_i, z_i)$ represent the position of UAV $i \in V$. When used, UAV i can act as a FAP f , which is responsible for providing wireless connectivity to ground devices located at the base of cuboid C , or as a FGW g , which is used to forward the traffic between the FAPs and the Internet. Let V^f represent the set of FAPs and V^g represent the set of FGWs. Consequently, FAP f belongs to the subset $V^f \subset V$, whereas FGW g belongs to the subset $V^g \subset V$. UAV i has available a total number of W_i wireless channels with constant bandwidth B , in Hz. We assume that B is the minimum channel bandwidth block per wireless channel that can be used to carry traffic and depends on the technology employed. The base of cuboid C is subdivided into a given number of subareas of fixed size (cf. right-hand side of Fig. 6.15). Let $a \in A$ represent a subarea, where A is the set of subareas that compose the base of the cuboid C . Subarea $a \in A$ is centered at $P_a = (x_a, y_a, 0)$, where there may be one ground device at maximum.

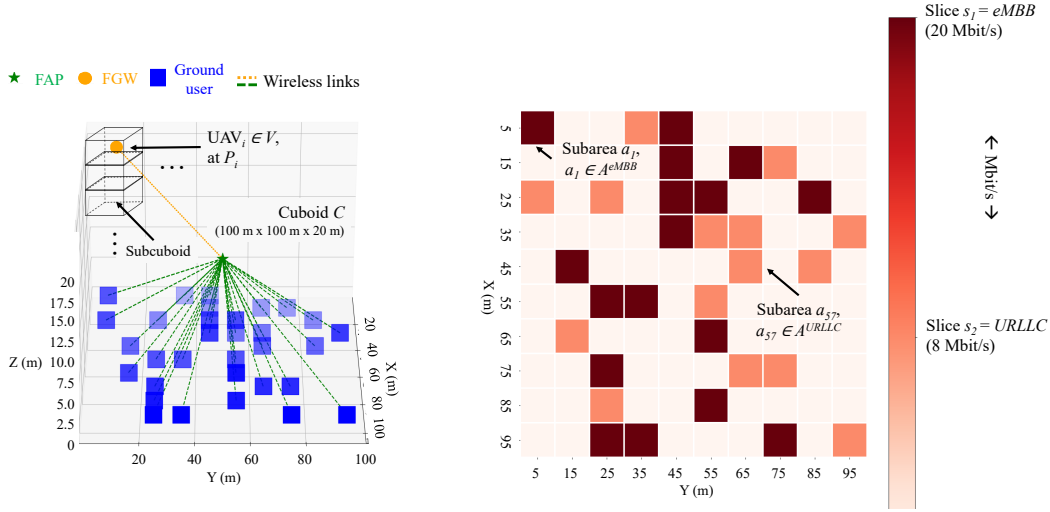


Figure 6.15: Illustrative networking scenario (3D representation on the left-hand side) with multiple ground devices (blue squares) served by an FMN consisting of a FAP (green star) and an FGW (orange circle). The ground devices are associated with two network slices $s \in S$ made available in different subareas $a \in A^s$ at the base of cuboid C (2D representation on the right-hand side).

The wireless channel between the UAVs themselves and the UAVs and the ground devices is modeled by the Free-Space Path Loss (FSPL) model. We assume the wireless channel is symmetric. The power $P_{R_{i,j}}$ received at communications node j from UAV $_i$, in dBm, is given by Eq. (6.1), where P_{T_i} is the transmission power of UAV $_i$, in dBm, and G_{T_i} and G_{R_j} are respectively the antenna gains of UAV $_i$ and node j , in dBi. The path loss component $P_{L_{i,j}}$, in dB, is computed by means of Eq. (6.2), where $d_{i,j}$ is the Euclidean distance between UAV $_i$ and node j , in meters, f_i is the carrier frequency used by UAV $_i$, in Hz, and c is the speed of light in vacuum, in m/s.

$$P_{R_{i,j}} = P_{T_i} + G_{T_i} + G_{R_j} - P_{L_{i,j}} \quad (6.1)$$

$$P_{L_{i,j}} = 20 \cdot \log_{10}(d_{i,j}) + 20 \cdot \log_{10}(f_i) + 20 \cdot \log_{10}\left(\frac{4 \cdot \pi}{c}\right) \quad (6.2)$$

The Signal-to-Noise Ratio (SNR) measured at node j from UAV $_i$, in dB, is given by $\text{SNR}_{i,j} = P_{R_{i,j}} - P_{N_{i,j}}$, where $P_{N_{i,j}}$ is the noise power, in dBm, which we assume to be constant for the channel bandwidth B . The capacity provided by each wireless channel is equal to the data rate enabled by the MCS index used for the wireless link established between UAV $_i$ and node j ($\text{MCS}_{i,j}$). The use of $\text{MCS}_{i,j}$ imposes a minimum $\text{SNR}_{i,j}$, considering the constant noise power $P_{N_{i,j}}$.

The number of independent wireless channels provided by UAV $_i$ to node j at time instant t_k is denoted by $w_{i,j}(t_k)$. $w_{i,j}(t_k)$ must be lower than or equal to the total number W_i of wireless

channels available at UAV_{*i*}, as defined in Eq. (6.3).

$$\sum_{j \in J} w_{i,j}(t_k) \leq W_i, \forall i \in V \quad (6.3)$$

The number of UAVs providing resources to each node $j \in J$, whether it is another UAV or a ground device, at time instant t_k is denoted by $K_{i,j}(t_k)$. We assume that node $j \in J$ is served by one and only one UAV_{*i*}, as stated in Eq. (6.4).

$$K_{i,j}(t_k) = 1, \forall i \in V, \forall j \in J \quad (6.4)$$

The indicator function $1_i(t_k)$, defined in Eq. (6.5), denotes whether UAV_{*i*} serves any node $j \in J$, be it another UAV or a ground device, at time instant t_k .

$$1_i(t_k) = \begin{cases} 1, & \text{if } \sum_{j \in J} w_{i,j}(t_k) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (6.5)$$

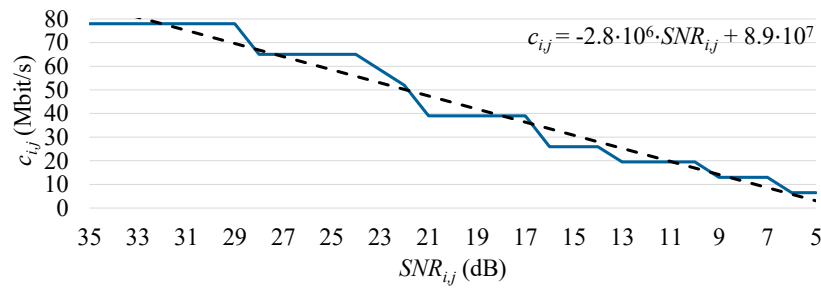
Let us consider that a Service Provider rents a set of network slices from a Mobile Network Operator, in order to offer online services to the ground devices located in area A , as depicted in Fig. 6.15. Let $s \in S$ represent a network slice, where S is the set of network slices. We assume that each subarea $a \in A$, occupied by a ground device in its center, is associated with a single network slice s , but a network slice s can cover multiple subareas $a \in A$. As such, network slice s enables the use of a service made available in area $A^s \subset A$. The area A^s associated with network slice s is the union of a set of fixed-size subareas $a \in A$. Let r_a represent the ground device located in the center of area $a \in A^s$. In turn, $j \in J^s$ represents a communications node (ground device or UAV) that either uses or provides communications resources to network slice $s \in S$.

The average traffic demand enabled in subarea $a \in A^s$ by network slice s is T^s . As such, for each node $j \in J^s$, the average data rate when served by a UAV_{*i*} must be higher than or equal to the average data rate T^s demanded by network slice s , as given by Eq. (6.6). $c_{i,j}(t_k)$ represents the bidirectional network capacity provided by a channel bandwidth block, in terms of the amount of bit/s carried between UAV_{*i*} and node j .

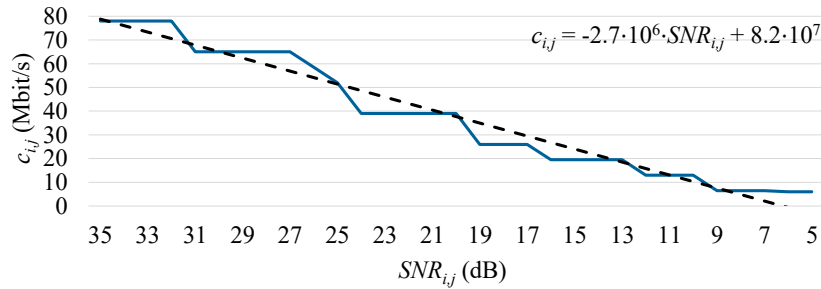
$$\sum_{i \in V} c_{i,j}(t_k) \cdot w_{i,j}(t_k) \geq T^s, \forall j \in J^s, \forall s \in S \quad (6.6)$$

The relation between the minimum SNR_{*i,j*} from UAV_{*i*} measured in node $j \in J^s$ required for using MCS_{*i,j*} is considered, taking into account target Bit Error Rate (BER) values according to the requirements of network slice s . For improved reliability, higher SNR values must be ensured so that a lower BER is achieved [212]. For illustrative purposes, BER equal to 10^{-10} is considered for a URLLC network slice type, while BER equal to 10^{-5} is employed for an eMBB network slice type. The relation between SNR_{*i,j*} and MCS_{*i,j*} for different target BER values is presented in

Fig. 6.16, considering the IEEE 802.11ac standard, Guard Interval of 800 ns, channel bandwidth of 20 MHz, and transmission power of 20 dBm. The minimum $SNR_{i,j}$ required to use each $MCS_{i,j}$ index was obtained by means of ns-3 simulations, considering node j generating User Datagram Protocol (UDP) Poisson traffic towards UAV $_i$. For each target BER value, an ns-3 simulation run, in which the UAV $_i$ was moving away from the node j , was carried out. The successively increased Euclidean distance between UAV $_i$ and node j led to the degradation of $SNR_{i,j}$ and induced the selection of suitable $MCS_{i,j}$ indexes. This allowed establishing the relationship between different $SNR_{i,j}$ values and the $MCS_{i,j}$ indexes selected for each target BER value. The *IdealWifiManager* mechanism was used in ns-3 as the Medium Access Control (MAC) automatic rate selection mechanism, since it allows configuring target BER values. The wireless link established between node j and UAV $_i$ was modeled by the FSPL model.



(a) Theoretical capacity versus SNR for an eMBB network slice, considering the target BER equal to 10^{-5} .



(b) Theoretical capacity versus SNR for a URLLC network slice, considering the target BER equal to 10^{-10} .

Figure 6.16: Wireless channel capacity modeled by linear regressions between the SNR and the data rate associated to the IEEE 802.11ac MCS indexes, considering channel bandwidth of 20 MHz.

Since the relation between $SNR_{i,j}$ and $MCS_{i,j}$ is stepwise (cf. solid blue lines in Fig. 6.16), making the problem intractable and complex to solve mathematically, we modeled it as a continuous function using a linear regression (cf. dashed black lines in Fig. 6.16), which is the function that closely fits the data. $c_{i,j}(t_k)$ changes according to the location of node $j \in J^s$ and the position of UAV $_i$, since both influence $SNR_{i,j}$.

For the node $j \in J^s$, the traffic being forwarded by UAV_{*i*} is modeled by an M/D/1 queue $Q_{i,j}$ (Poisson arrival, deterministic service time, 1 server) [121]. Traffic arrives at queue $Q_{i,j}$ with arrival rate $\lambda_{i,j}$ packet/s and is served with a service rate $\mu_{i,j}$ packet/s. The average delay $D_{i,j}(t_k)$ of a packet generated by node $j \in J^s$ at time instant t_k is computed using Eq. (6.7).

$$D_{i,j}(t_k) = \frac{1}{\mu_{i,j}} \cdot \frac{\rho_{i,j}}{2 \cdot \mu_{i,j} \cdot (1 - \rho_{i,j})} \cdot 1_i(t_k), \forall i \in V, \forall j \in J^s, \forall s \in S \quad (6.7)$$

where:

$$\rho_{i,j} = \frac{\lambda_{i,j}}{\mu_{i,j}} < 1, \lambda_{i,j} \neq 0, \forall i \in V, \forall j \in J^s, \forall s \in S$$

$$\mu_{i,j} = \sum_{i \in V} c_{i,j} \cdot w_{i,j}, \forall j \in J^s, \forall s \in S$$

For each node $j \in J^s$, the average packet delay when served by a UAV_{*i*} must be lower than or equal to the maximum average packet delay H^s associated with network slice s , as given by Eq. (6.8).

$$D_{i,j} \cdot 1_i(t_k) \leq H^s, \forall i \in V, \forall j \in J^s, \forall s \in S \quad (6.8)$$

Herein, we consider average QoS values for illustrative purposes, but an SLA established with a Mobile Network Operator can also refer maximum values (e.g., the longest packet delay among all packet delay values) or median values (e.g., 50th percentile of the frequency distribution of packet delays) [130].

The problem addressed herein consists in minimizing the cost of deploying a slicing-aware FMN able to meet the coverage and QoS levels of any network slice $s \in S$, including a minimum average data rate T^s and a maximum average packet delay H^s . Solving this problem includes determining the minimum number of UAVs to use, including FAPs and FGWs, their 3D positions, and the number of wireless channels they provide. We assume that there is a set of N predefined 3D positions where potential UAVs are placed. Placing FAP_{*f*} $\in V^f$ at P_i , at time instant t_k , has a fixed cost F^f , where F^f is a constant associated with the usage of FAP_{*f*}. Similarly, placing an FGW $g \in V^g$ at P_i , at time instant t_k , has a fixed cost F^g , where F^g is a constant associated with the usage of FGW_{*g*}. The costs F^f and F^g may be defined according to multiple criteria, such as the cost of the hardware carried on board the UAV, the type of UAV (FAP or FGW), and the operating cost of each UAV. The solution aims at ensuring that each FAP provides sufficient communications resources to meet the requirements of the ground devices served, each associated with a given network slice. In turn, the communications resources provided by each FAP must also be made available in the wireless link established between the FAP and the corresponding FGW, in order to meet the QoS requirements associated to each network slice in the access and backhaul networks.

Accomplishing a slicing-aware FMN involves multiple research challenges, including: 1) meeting the different QoS levels of multiple network slices on top of a shared physical wireless network infrastructure; 2) determining the minimum number of UAVs required and their placement, in order to simultaneously provide coverage with the minimum targeted SNR values to multiple network slices made available in different geographical areas; and 3) achieving a global admissible solution for the placement and allocation of communications resources, formulated as an optimization problem, in a time interval compliant with the network reconfiguration period, while ensuring the minimum number of UAVs and wireless channels required to meet the QoS levels of multiple network slices. The optimization problem, including the objective function and the constraints, is defined in Eq. (6.9). In the optimization problem formulated in Eq. (6.9), the objective function defined in Eq. (6.9a) minimizes the overall cost of deploying a slicing-aware FMN, which consists of the sum of the costs associated with the FGWs and FAPs used at any time instant t_k . Solving the problem includes identifying the potential UAVs that are actually used and their role (whether FAPs or FGW) in the slicing-aware FMN, considering the following constraints at any time instant t_k :

- Eq. (6.9b) ensures that the number of wireless channels provided by FGW_g to FAP_f is lower than or equal to the total number of wireless channels available at FGW_g .
- Eq. (6.9c) ensures that the capacity, in bit/s, of the wireless links established between FGW_g and FAP_f is higher than or equal to the capacity, in bit/s, of the wireless link established between FAP_f and the ground device in the center of subarea $a \in A^s$, in order to meet the QoS requirements of the network slice $s \in S$.
- Eq. (6.9d) ensures that FAP_f is served by one and only one FGW_g .
- Eq. (6.9e) ensures that the number of wireless channels provided by FAP_f to the ground device in the center of subarea $a \in A^s$ is lower than or equal to the total number of wireless channels available at FAP_f .
- Eq. (6.9f) ensures that the capacity, in bit/s, of the wireless links established between FAP_f and the ground device in the center of subarea $a \in A^s$ is higher than or equal to the minimum average data rate enabled by network slice $s \in S$.
- Eq. (6.9g) ensures that the ground device in the center of subarea $a \in A^s$ is served by one and only one FAP_f .
- Eq. (6.9h) ensures that the average packet delay, in seconds, carried between the ground device in the center of subarea $a \in A^s$ and FGW_g , through FAP_f , is lower than or equal to the maximum average packet delay enabled by network slice $s \in S$.

- Eq. (6.9i) ensures that a potential UAV_i can perform one and only one role, FAP or FGW.

$$\text{minimize}_{w_{g,f}(t_k), 1_g(t_k), w_{f,r_a}(t_k), 1_f(t_k)} \sum_{g \in V^g} F^g \cdot 1_g(t_k) + \sum_{f \in V^f} F^f \cdot 1_f(t_k), \forall a \in A \quad (6.9a)$$

subject to:

$$\sum_{f \in V^f} w_{g,f}(t_k) \leq W_g, \forall g \in V^g \quad (6.9b)$$

$$\sum_{g \in V^g} c_{g,f}(t_k) \cdot w_{g,f}(t_k) \geq c_{f,r_a}(t_k) \cdot w_{f,r_a}(t_k), \forall f \in V^f, \forall a \in A^s, \forall s \in S \quad (6.9c)$$

$$K_{g,f}(t_k) = 1, \forall g \in V^g, \forall f \in V^f \quad (6.9d)$$

$$\sum_{a \in A^s} w_{f,r_a}(t_k) \leq W_f, \forall f \in V^f \quad (6.9e)$$

$$\sum_{f \in V^f} c_{f,r_a}(t_k) \cdot w_{f,r_a}(t_k) \geq T^s, \forall a \in A^s, \forall s \in S \quad (6.9f)$$

$$K_{f,r_a}(t_k) = 1, \forall f \in V^f, \forall a \in A^s, \forall s \in S \quad (6.9g)$$

$$D_{f,r_a} \cdot 1_f(t_k) + D_{g,f} \cdot 1_g(t_k) \leq H^s, \forall f \in V^f, \forall g \in V^g, \forall a \in A^s, \forall s \in S \quad (6.9h)$$

$$1_f(t_k) + 1_g(t_k) \leq 1, \forall f \in V^f, \forall g \in V^g \quad (6.9i)$$

The objective function in Eq. (6.9a) allows minimizing the number of UAVs used, since we are focused on minimizing the cost of deploying a slicing-aware FMN. Implicitly, we can take advantage of all wireless channels available at each UAV. In order to improve efficiency, the objective function can be evolved as future research to optimize the number of wireless channels used by the UAVs.

6.3 SurFABLE Algorithm

The SurFABLE algorithm is described in this section, including its rationale and a theoretical analysis on the employed channel assignment approach for a representative scenario.

6.3.1 Rationale

In order to enable the coverage extension from a remote AP or BS, while meeting the requirements of different network slices, the placement of UAVs as well as the allocation of communications resources in FMNs have to be defined. This thesis proposes SurFABLE, an algorithm that makes the joint placement and allocation of communications resources in slicing-aware FMNs. SurFABLE allows the joint computation of the amount of communications resources

needed, namely the number of UAVs acting as FAPs and FGWs, and their placement in 3D space. SurFABLE is intended to run in a Central Station deployed in the Cloud or at the Edge of the FMN. The Central Station is in charge of sending the up-to-date UAV positions and network configuration computed by SurFABLE. The FAPs and the FGWs position and reconfigure themselves based on the information received. For exchanging data between the Central Station and the UAVs, an out-of-band long-range wireless channel (e.g., based on IEEE 802.11ah) is used. It allows for an always-on control link, even when the FMN is being reconfigured, and avoids introducing overhead in the access and backhaul network. The design of the out-of-band wireless communications solution is beyond the scope of this thesis.

Inspired by the Capacitated Facility Location Problem [213], a classical optimization problem that aims at selecting the best among potential locations for a factory or warehouse, SurFABLE initially represents the venue where UAVs can be deployed as a cuboid C (step 1 of Algorithm 3). Cuboid C is divided into a set of N smaller sub-cuboids, each associated with a potential UAV deployed at P_i (step 2), and M ground subareas (step 3), each occupied by a ground device in the central position. Any potential UAV is suitable to act as a real FAP or FGW, if it is part of the final solution determined by SurFABLE. Each ground device has associated a known traffic demand, in bit/s, and a maximum average packet delay, in seconds; both are values associated with the SLA of a given network slice $s \in S$.

SurFABLE calculates the SNR of the wireless links that can be established between each potential UAV and the ground device in the center of each subarea $a \in A^s$, as well as between the potential UAVs themselves (step 4). Then, it calculates the network capacity that can be achieved when using the minimum channel bandwidth block that each potential UAV can provide (step 5). The minimum channel bandwidth block is a configuration parameter that can be specified either in terms of the number of OFDMA Resource Units for IEEE 802.11ax and 5G New Radio or in terms of the number of 20 MHz wireless channels for IEEE 802.11.

SurFABLE uses a state of the art solver to identify the potential UAVs that shall be actually used their role in the FMN (FAP or FGW), and the number of wireless channels to be made available by each UAV. Each potential UAV has a given cost and a limited number of wireless channels available. The objective is to minimize the cost of deploying a slicing-aware FMN. The solver must simultaneously ensure that 1) each network slice's SLA is met and 2) the capacity of each UAV in terms of the total number of wireless channels available is not exceeded (step 6).

The *Gurobi*TM [214] optimizer is used in the current version of SurFABLE. A cost of 1000 and a set of 8×20 MHz wireless channels per potential UAV, which enable a total channel bandwidth of 160 MHz, were considered. As shown in Fig. 6.15, the resulting solution determined by SurFABLE consists of the UAVs acting as FAPs or FGWs, which are associated with a known 3D location, and the number of wireless channels that they must provide to each subarea. When a potential UAV does not provide resources to any of the subareas, its cost is zero, and it is not used in the

Algorithm 3: SurFABLE algorithm**Input:**

- i): Venue dimensions (m^3)
- ii): Number of wireless channels and channel bandwidth (Hz) available per potential UAV
- iii): Device density (devices/ m^2) per network slice
- iv): Traffic demand (bit/s) per network slice
- v): Maximum packet delay (s) per network slice

Output:

- i): FAPs and FGW positions (3D Cartesian coordinates)
- ii): Number of wireless channels provided by each FAP and FGW
- iii): Association between FGWs, FAPs and ground devices served

Steps:

- 1: Represent the venue as a cuboid C
- 2: Discretize cuboid C into N sub-cuboids centered at P_i
- 3: Discretize base of C into M subareas $a \in A$
- 4: Compute $SNR_{i,j}$ for the wireless link available between the potential UAVs themselves, and with the ground devices
- 5: Compute the network capacity $c_{i,j}$ that each potential UAV can provide using the available wireless links
- 6: Solve optimization problem formulated in Eq. (6.9) using a state of the art solver
- 7: Assign the wireless channels that reduce the bandwidth used
- 8: Define the queue size and forwarding table for each FAP
- 9: Reconfigure the FMN

FMN.

As a precise network resource allocation is not possible in some technologies, such as IEEE 802.11, where the channel bandwidth must be an integer multiple of 20 MHz, SurFABLE allocates resources employing a heuristic-based channel assignment approach (step 7). This channel assignment approach also allows for a reduction in the overall bandwidth used while providing QoS guarantees. To that end, each wireless channel is assigned to as many subareas as possible, in order to reduce the overall bandwidth used by sharing the available spectral resources. The channel bandwidth allocated to each subarea is computed by multiplying the number of wireless channels derived from the optimization problem defined in Eq. (6.9) with the minimum channel bandwidth block enabled by the technology used, as it will be explained in Section 6.3.2.

The spatial component of the problem solved by SurFABLE is related with the number of communications resources to allocate to each subarea. For that purpose, we consider that each ground device, placed in the center of a given subarea $a \in A^s$, represents the aggregate traffic generated within that subarea. We considered this approach since we assume the SLA established between a Service Provider and a Mobile Network Operator includes information on the number of devices per area unit, in device/ m^2 . Moreover, a precise location of the ground de-

vices is typically not known and can include some deviations between the expected and the real positions over time. The discretization of the space into subareas allows overcoming these challenges, and it can be fine-tuned according to multiple criteria; this aspect is left for future work.

Regarding the spectrum component, SurFABLE was designed with recently released and emerging wireless technologies in mind, such as 5G New Radio and IEEE 802.11ax, in which the number of spectral resources is expressed as the number of Resource Units, called generically as wireless channels herein. This allows accurately allocating the number of wireless channels computed by SurFABLE using minimum channel bandwidth, since each Resource Unit is typically characterized by a few kHz of channel bandwidth. The smaller the bandwidth of the available channels the more efficient SurFABLE is, since it allows reducing the underused spectrum per wireless channel. When legacy technologies such as IEEE 802.11 are used, which do not allow to define a precise channel bandwidth per subarea, the channel assignment approach employed by SurFABLE allows reducing the overall bandwidth required. This makes SurFABLE agnostic to the technology used by the FMN.

6.3.2 Theoretical Analysis for a Representative Scenario

In order to exemplify the channel assignment approach made by SurFABLE, let us consider the illustrative example for a network slice $s \in S$ shown in Fig. 6.17. Without the channel assignment approach employed by the SurFABLE algorithm, the total bandwidth required is 140 MHz, as shown in Fig. 6.17a. This baseline channel assignment approach takes into account different 20 MHz wireless channels assigned to each subarea $a \in A^s$. The higher the number of subareas $a \in A^s$ the higher the total channel bandwidth required – each row in Fig. 6.17a corresponds to a subarea $a \in A^s$. However, because the bandwidth used for each subarea $a \in A^s$ is far from the full channel bandwidth available, this baseline channel assignment approach has a reduced spectral efficiency. With SurFABLE's channel assignment, the total bandwidth used is reduced to 100 MHz. This is accomplished by using the maximum available bandwidth of each wireless channel while assigning the same wireless channel to the maximum number of subareas $a \in A^s$, as shown in Fig. 6.17b. The queue size management is performed by the Proactive Queue Management (PQM) mechanism, described in Chapter 3. The forwarding tables of the FAPs, used to exchange data traffic between the ground devices and the assigned FGW in a two-hop architecture, are defined by the Routing for Flying Multi-hop Networks (RedeFINE) mechanism, presented in Chapter 5. Finally, the FMN is reconfigured accordingly.

6.4 Performance Evaluation

The performance evaluation of the SurFABLE algorithm is described in this section. Firstly, we explain the evaluation methodology. Secondly, we characterize the networking scenarios and

Preliminary number of channels before SurFABLE's channel assignment	Preliminary number of channels (rounding to upper integer)	Preliminary minimum bandwidth required (MHz)
1.117	2	2 x 20 = 40
0.937	1	1 x 20 = 20
1.490	2	2 x 20 = 40
0.769	1	1 x 20 = 20
0.495	1	1 x 20 = 20
Total network resources	7	140

(a) Resources required before channel assignment. Each row is associated to a subarea $a \in A^s$.

Final number of channels after SurFABLE's channel assignment	Final number of channels (rounding to upper integer)	Final minimum bandwidth required (MHz)
1.117 + 0.769 = 1.886	2	2 x 20 = 40
0.937	1	1 x 20 = 20
1.490 + 0.495 = 1.985	2	2 x 20 = 40
Total network resources	5	100

(b) Resources required after channel assignment. The quantities of wireless channels stated in Fig. 6.17a are added together according to the color patterns.

Figure 6.17: Resource allocation performed by SurFABLE, considering the minimum channel bandwidth equal to 20 MHz [38].

the baseline solutions considered. Thirdly, we present and discuss the ns-3 simulation results. Fourthly, we detail the experimental setup. Fifthly, we present and discuss the experimental results.

6.4.1 Methodology

A theoretical performance evaluation requires many simplifications, due to the high complexity of the problem addressed. It potentially leads to unrealistic results. For this reason, the performance evaluation of the SurFABLE algorithm was carried out by means of ns-3 simulations and an experimental testbed. On the one hand, the ns-3 simulations allowed to evaluate the performance of SurFABLE in complex networking scenarios, characterized by high variability regarding the communications nodes' positions and network configurations. The ns-3 simulator [73] was chosen due to a) its accurate models for wireless networks that lead to realistic results and b) its wide acceptance by the scientific community. On the other hand, the experimental testbed allows validating in the real-world the results obtained by means of ns-3 simulations for representative networking scenarios.

The performance evaluation considers three performance metrics:

- **Throughput:** the average number of bits received per second by the sink application at the FGWs, in bit/s.
- **Packet Delivery Ratio (PDR):** the average ratio between the number of packets received by the sink application at the FGWs and the number of packets generated by the source application at the ground devices.
- **Delay:** the average time interval since the packets are generated by the source application at the ground devices until they reach the sink application at the FGWs. It includes queuing, transmission, and propagation delays.

The average values of throughput, PDR, and packet delay for each network slice were measured every second for each performed experiment. The results are represented by means of the Cumulative Distribution Function (CDF) for the packet delay and by the Complementary Cumulative Distribution Function (CCDF) for the throughput and PDR. The CDF $F(x)$ represents the percentage of samples for which the delay is lower than or equal to x , while the CCDF $F'(x)$ represents the percentage of samples for which the throughput and PDR are higher than x .

6.4.2 Networking Scenarios

Three sets of networking scenarios were considered: 1) five scenarios consisting of 5 ground devices; 2) five scenarios consisting of 20 ground devices; and 3) five scenarios consisting of 40 ground devices. In all networking scenarios, each ground device was randomly associated to a network slice type (whether eMBB or URLLC) and randomly positioned on the ground of a venue with dimensions up to $X = 100$ m, $Y = 100$ m, $Z = 20$ m. The number of subareas was determined by taking into account the occupation of the base of the cuboid C (area A) equal to 5%, 20%, and 40% of the total area available. Each ground device was identified by a traffic demand T^s equal to 4 Mbit/s, 8 Mbit/s or 20 Mbit/s, which correspond to respectively 5%, 10%, and 25% of the data rate associated with the highest MCS index for channel bandwidth of 20 MHz, Guard Interval of 800 ns, and single spatial stream wireless links (78 Mbit/s). In the ns-3 simulations, we assumed $BER^{eMBB} = 10^{-5}$ and $BER^{URLLC} = 10^{-10}$, which define the minimum SNR values required for transmitting a frame using any MCS index with the network configuration employed.

We consider that each networking scenario consists of static ground devices, each representing the aggregate traffic generated within the subarea it is placed. Each networking scenario corresponds to a snapshot of the FMN at $t_k = k \cdot \Delta t, k \in \mathbb{N}_0$, where $\Delta t \gg 1$ s is the network reconfiguration period. In a real-world deployment, each network reconfiguration entails solving the problem again to find an up-to-date solution. As already referred to in previous chapters, Δt is a networking scenario parameter that can be adjusted to achieve a trade-off between the

Table 6.1: Positions and network slice types associated with the ground devices considered in the performance evaluation carried out using ns-3 simulations and an experimental testbed.

Scenario	Ground devices characterization: {Cartesian coordinates, network slice type}
Scenario 1	{{(26, 32), eMBB}, {(37, 37), eMBB}, {(2, 22), eMBB}, {(7, 18), URLLC}, {(14, 35), URLLC}}
Scenario 2	{{(6, 30), URLLC}, {(34, 20), URLLC}, {(26, 24), URLLC}, {(35, 17), eMBB}, {(34, 32), eMBB}}
Scenario 3	{{(38, 25), eMBB}, {(30, 37), URLLC}, {(26, 35), URLLC}, {(15, 37), eMBB}, {(24, 1), URLLC}}
Scenario 4	{{(6, 28), URLLC}, {(29, 33), URLLC}, {(19, 9), URLLC}, {(34, 34), eMBB}, {(33, 17), eMBB}}
Scenario 5	{{(2, 14), eMBB}, {(14, 27), eMBB}, {(28, 0), URLLC}, {(17, 25), URLLC}, {(24, 32), eMBB}}

stability of the FMN, the performance requirements defined by the SLAs of the network slices offered, and the time required to determine a suitable solution for the placement and allocation of communications resources.

The details regarding the network configurations employed in the evaluation using ns-3 simulations and the experimental testbed are presented in Table 6.2.

Two network slice types were considered:

- **eMBB network slice**, designed to enable the use of rich-media applications (e.g., video streaming) with an average throughput of 20 Mbit/s per ground device (subarea) and an average delay up to 100 ms.
- **URLLC network slice**, designed to enable the use of mission-critical applications (e.g., communications for first-responders) with an average throughput of 4 Mbit/s (evaluation using ns-3 and the experimental testbed), and 8 Mbit/s (evaluation in ns-3 only) per ground device (subarea) and average delay up to 10 ms.

The lower target throughput values (4 Mbit/s) considered in the joint evaluation in ns-3 and in the experimental testbed were due to the reduced channel bandwidth enabled by the hardware used in the testbed. This motivated the definition of less demanding throughput requirements to reach at least one possible solution for the networking scenarios employed. It is worth noting that this does not compromise the validity of the performance evaluation carried out, since the same exact conditions were considered in the networking scenarios subject to joint evaluation in ns-3 and using the testbed.

The two types of network slices were defined to take into account networking performance requirements imposed by representative applications. Yet, the optimization problem formulated

Table 6.2: Networking parameters used in the evaluation of the SurFABLE algorithm. IEEE 802.11 was used due to its ubiquity, cost-effectiveness, and flexibility for evaluating network performance in both simulation and experimentation, using the same exact conditions. Yet, the problem formulation and SurFABLE are agnostic to the technology used.

Parameter	Value
Simulation time per run	70 s
Number of runs	20
Wi-Fi standard	IEEE 802.11ac
Wi-Fi mode	Infrastructure
Maximum channel bandwidth	160 MHz
Minimum channel bandwidth	20 MHz
Transmission (TX) power	20 dBm
Guard Interval	800 ns
Number of antennas	1
Propagation model	FSPL (Friis)
Application traffic	UDP Poisson and Transmission Control Protocol (TCP)
Offered bitrate per slice	8 Mbit/s (URLLC), 20 Mbit/s (eMBB)
Packet size	1400 Bytes
MAC auto rate	<i>IdealWifiManager</i>

in Eq. (6.9) and the SurFABLE algorithm are valid for any number and type of network slices, as well as different network performance requirements. This is a step forward with respect to the literature, such as the approach used by IEEE 802.11e, which provides four different QoS classes only, according to the traffic type, but does not allow ensuring different QoS guarantees for traffic flows of the same type.

6.4.3 Baseline Solutions

We considered three baseline solutions in the performance evaluation. All the baseline solutions ensure that each used UAV provides network resources only to subareas that belong to the same network slice $s \in S$ so that resource isolation is guaranteed. This ensures that each network slice represents an independent network in practice.

The following baseline solutions were considered:

- **k-means clustering.** This baseline solution, which uses the state of the art k -means clustering algorithm [215], is proposed by this thesis for solving the allocation and placement of communications resources as two independent sub-problems. It allows providing each

Table 6.3: Networking parameters considered in the evaluation of the SurFABLE algorithm using ns-3 and experimental testbed.

Parameter	Value
Experiment time per run	10 s
Number of runs	6
Wi-Fi standard	IEEE 802.11n
Wi-Fi mode	Infrastructure
Maximum channel bandwidth	40 MHz
Minimum channel bandwidth	20 MHz
TX power	20 dBm
Guard Interval	800 ns
Number of antennas	1
Application traffic	UDP Constant Bit Rate (CBR)
Offered bitrate per slice	4 Mbit/s (URLLC), 20 Mbit/s (eMBB)
Packet size	1400 Bytes
MAC auto rate	<i>MinstrelHtWifiManager</i>

subarea with at least the same amount of channel bandwidth calculated by SurFABLE. However, it defines a non QoS-aware placement for the FAPs. This potentially leads to wireless links not having a high enough SNR to allow for a network capacity capable of accommodating the offered traffic. Moreover, it requires as many UAVs as the number of clusters of subareas formed.

First, the k -means clustering algorithm defines $|K|$ random positions as clusters' centroids. Then, it assigns each subarea $a \in A^s$ to the nearest cluster by calculating distances to clusters' centroids. After that, it determines the up-to-date cluster's centroid by computing the average position among the assigned subareas, which is the position where a FAP must be placed, at 10 m altitude – since the altitude of the cuboid C considered for deploying the UAVs is 20 m, in the k -means clustering algorithm we considered the FAPs, in the first tier, are placed at 10 m altitude, while the FGWs, in the second tier, can be placed up to 20 m altitude.

Each FAP must provide the cluster's subareas with the minimum channel bandwidth computed by SurFABLE for the same exact subareas, using up to 160 MHz of channel bandwidth. When any FAP does not have enough channel bandwidth available, $|K|$, initially set to 1, is successively increased by 1 and the k -means clustering algorithm is run again until all FAPs are able to provide the required channel bandwidth. The algorithm is run independently for each network slice $s \in S$, in order to ensure resource isolation. As such, the number of

clusters ($|K|$) can be different for each network slice $s \in S$.

After the placement of the FAPs is determined by means of the k -means algorithm, the GateWay Placement (GWP) algorithm [79] is used to determine the placement of an FGW serving all the FAPs associated with the same network slice $s \in S$. GWP enables wireless links between the FAPs and the FGW with high enough capacity to accommodate the traffic forwarded by the FAPs. Each FGW is configured with channel bandwidth of 160 MHz, which is the maximum channel bandwidth enabled by legacy IEEE standards. In this solution each UAV serves a single network slice. GWP implies the use of a single FGW for each network slice; independently allocating and placing communications resources to the access and backhaul networks increases the number of UAVs used.

- **Geometric center.** This baseline solution provides an independent network for the ground devices that belong to each network slice. It considers a best-effort coverage-aware UAV deployment approach, without calculating the amount of communications resources required to meet the targeted QoS requirements.

This baseline solution uses one FAP and one FGW for each network slice $s \in S$, placed in the geometric center of all subareas $a \in A^s$. For each network slice, the FAP is hovering at altitude of 10 m, while the FGW is hovering exactly above the FAP at altitude of 20 m. The FAPs make available a wireless channel with the minimum channel bandwidth computed by SurFABLE, while the FGW provides the maximum possible channel bandwidth in IEEE 802.11, equal to 160 MHz. From the communications perspective, placing a UAV at the geometric center of all ground devices that belong to the same network slice $s \in S$ allows maximizing the SNR offered to them.

- **Per network slice QoS-aware.** This baseline solution allows determining the joint allocation and placement of communications resources for each slide independently.

The per network slice QoS-aware baseline solution considers that the SurFABLE algorithm runs independently for each network slice $s \in S$. As such, it provides the exact amount of communications resources required for each network slice $s \in S$, while not sharing communications resources among different network slices. In the same way as the geometric center baseline solution, it provides an independent network for each network slice $s \in S$; however, this solution is aware of the QoS requirements associated with each network slice $s \in S$.

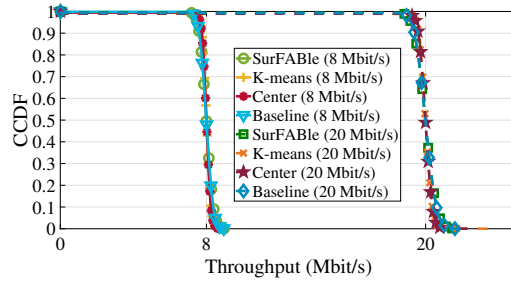
It must be noted that the slicing-aware approach employed by SurFABLE is achieved using different wireless channels with channel bandwidth equal to 20 MHz (minimum channel bandwidth block allowed in IEEE 802.11). As such, considering an illustrative example where a channel bandwidth of 60 MHz is required to serve the ground devices, SurFABLE uses three wireless

channels of 20 MHz, totaling 60 MHz. On the other hand, since the k -means clustering and geometric center baseline solutions use a single wireless channel per UAV, a channel bandwidth of 80 MHz is required, because it is not possible to configure a channel bandwidth of 60 MHz in IEEE 802.11. For this reason, the k -means and geometric center baseline solutions will tend to use more communications resources than the SurFABLE algorithm.

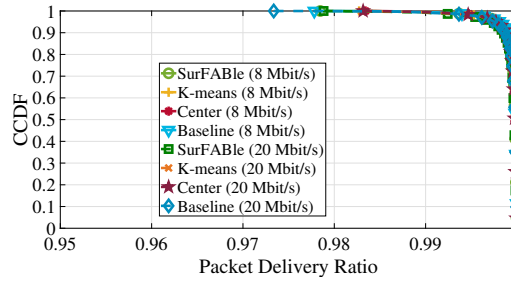
6.4.4 Simulation Results

The simulation results considering UDP traffic are presented in Fig. 6.18, Fig. 6.19, and Fig. 6.20, whereas the simulation results for TCP traffic are presented in Fig. 6.21, Fig. 6.22, and Fig. 6.23. They demonstrate that SurFABLE (represented by the green circle and square markers) enables meeting the target QoS levels associated with the eMBB (20 Mbit/s) and URLLC (8 Mbit/s) network slices. For 5 subareas and UDP traffic, all solutions allow meeting the target QoS levels (cf. Fig. 6.18), especially due to the low demand in terms of communications resources that characterize these scenarios. For this reason, the k -means clustering (represented by the yellow plus and orange crossed markers) and geometric center (represented by the reddish asterisk and star markers) baseline solutions over-provision communications resources, enabling delay lower than 0.6 ms. The network performance achieved when using k -means clustering and geometric center degrades significantly for TCP traffic when considering 5 subareas (cf. Fig. 6.21), and for UDP and TCP as the number of subareas increases (cf. Fig. 6.19, Fig. 6.20, Fig. 6.22, and Fig. 6.23). On the one hand, the increase in the number of subareas leads to a greater demand for communications resources. On the other hand, due to its congestion control mechanism, TCP dynamically adjusts the traffic offered to the Internet Protocol (IP) layer so that the channel is fully utilized but not saturated. This justifies the lower throughput, PDR, and delay measured when using the baseline solutions (cf. Fig. 6.21 to Fig. 6.23). In turn, the per network slice QoS-aware baseline solution (blue downward-pointing triangle and diamond markers) brings the network performance closer to SurFABLE in all scenarios, at the expense of using more UAVs.

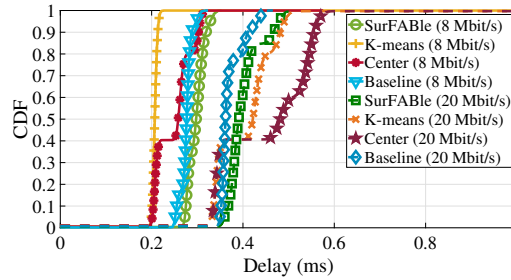
When comparing the solutions in terms of average bandwidth and number of UAVs used (cf. Fig. 6.24), the results show that SurFABLE (represented by the green horizontal line) uses only two UAVs (one FAP plus one FGW) and requires the least amount of channel bandwidth. This is achieved by ensuring a suitable placement for the FAPs and FGWs, and the allocation of multiple wireless channels with the minimum channel bandwidth block available (20 MHz in IEEE 802.11), which allows reducing the wasted channel bandwidth per wireless channel. On the other hand, although the geometric center baseline solution (represented by the red diagonal line) uses twice the number of UAVs (two FAPs plus two FGWs) with respect to SurFABLE – Fig. 6.24, the network performance is significantly degraded as the number of subareas grows, as observed in the performance results from Fig. 6.18 to Fig. 6.23. The k -means clustering algorithm (represented by the orange vertical line in Fig. 6.24) defines a non-QoS aware placement for the FAPs.



(a) Throughput CCDF.



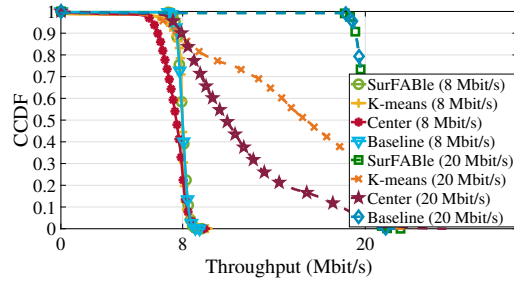
(b) PDR CCDF.



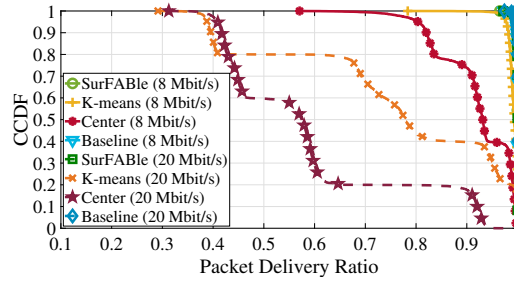
(c) Delay CDF.

Figure 6.18: Simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP Poisson traffic.

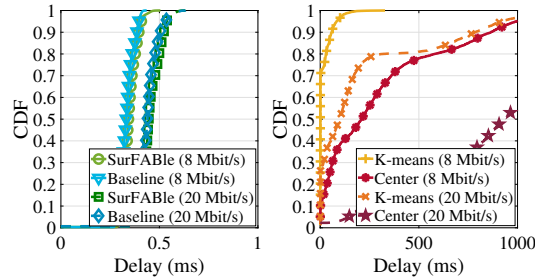
Even though it uses the highest number of UAVs and channel bandwidth, this does not result in better network performance. From the communications perspective, the k -means clustering algorithm provides the same SNR to all subareas within the same cluster, but it does not guarantee target SNR values, due to its non QoS-aware placement baseline solution. As such, the wireless links may not have the minimum SNR necessary to induce the MCS indexes capable of meeting the targeted performance requirements in some networking scenarios. Moreover, the channel bandwidth calculated using SurFABLE, which considers the joint placement and allocation of communications resources, is insufficient to achieve a network capacity high enough to accommodate the traffic offered when the independent resource allocation and placement made



(a) Throughput CCDF.



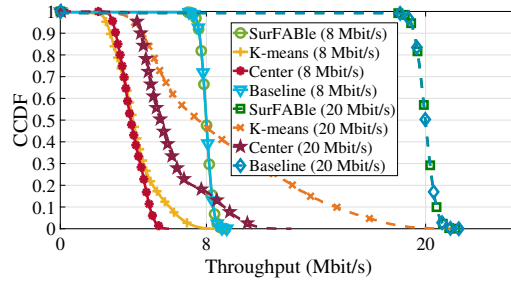
(b) PDR CCDF.



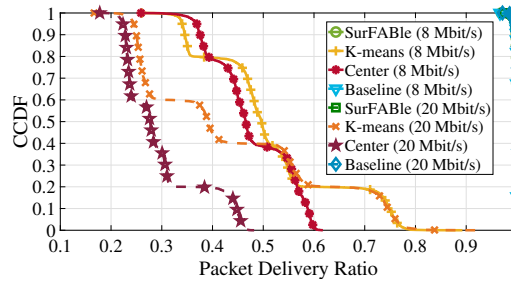
(c) Delay CDF.

Figure 6.19: Simulation average performance results for five networking scenarios consisting of 20 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP Poisson traffic.

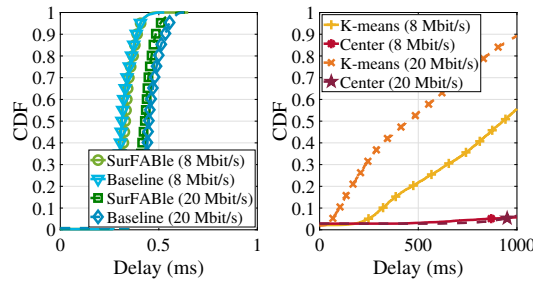
by the k -means clustering algorithm is employed. The high resource usage inherent to k -means clustering is exacerbated because the wireless channel made available by each FAP is not shared among subareas that belong to different clusters. This leads to wasted bandwidth in underused wireless channels. The performance degradation observed when employing the k -means and geometric center baseline solution is overcome by the per network slice QoS-aware baseline solution – Fig. 6.19 to Fig. 6.23. It uses the SurFABLE algorithm to compute the amount of channel bandwidth and placement of the UAVs for each network slice independently. However, the independent placement and allocation of communications resources employed by the per network slice QoS-aware baseline solution's results in four UAVs required (cf. blue wave line in Fig. 6.24b)



(a) Throughput CCDF.



(b) PDR CCDF.

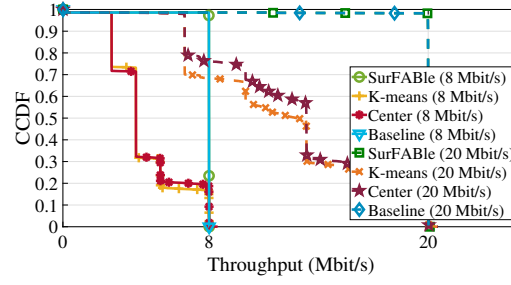


(c) Delay CDF.

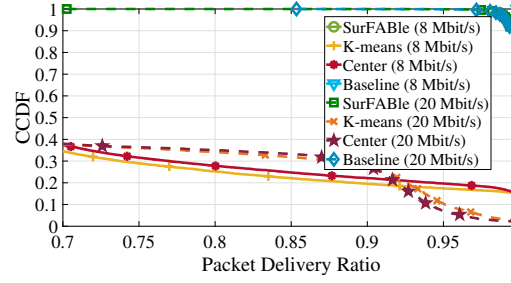
Figure 6.20: Simulation average performance results for five networking scenarios consisting of 40 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP Poisson traffic.

– twice as many UAVs when compared to SurFABLE, which employs a slicing-aware approach.

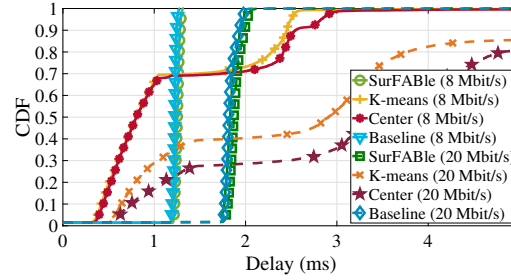
Despite the minimum amount of communications resources used, SurFABLE allows achieving higher network performance. This is attained by ensuring a suitable placement of the FAPs and FGWs, and the allocation of a given number of channels with the minimum channel bandwidth block possible for the wireless communications technology employed (multiples of 20 MHz in the case of IEEE 802.11). This represents a significant advantage when compared with approaches that aim at improving the aggregate network performance offered, as the geometric center baseline solution. The geometric center baseline solution assigns to all subareas of each network slice a single wireless channel with up to 160 MHz bandwidth, which is not enough in some networking



(a) Throughput CCDF.



(b) PDR CCDF.



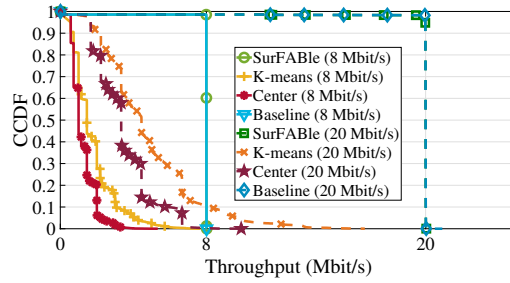
(c) Delay CDF.

Figure 6.21: Simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering TCP traffic.

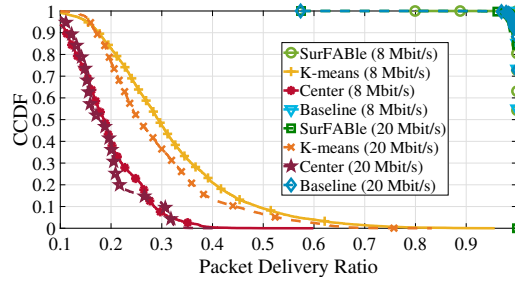
scenarios and leads to underused channel bandwidth in others.

6.4.5 Experimental Setup

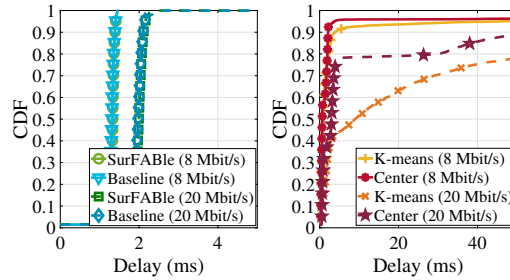
The performance achieved when using the SurFABLE algorithm was also evaluated by means of an experimental testbed. The ground devices consisted of *Raspberry Pi* devices running the *Raspberry Pi* Operating System [216], each equipped with a *Panda Wireless N600* [217] IEEE 802.11 Network Interface Card (NIC) connected through Universal Serial Bus; the NICs were configured in Infrastructure mode as IEEE 802.11 STA. Each FAP consisted of a *TP-Link TL-WR902AC v3* [218] IEEE 802.11 AP with two built-in independent NICs, running the *OpenWrt*



(a) Throughput CCDF.



(b) PDR CCDF.

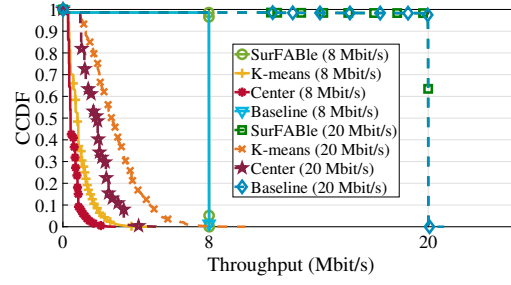


(c) Delay CDF.

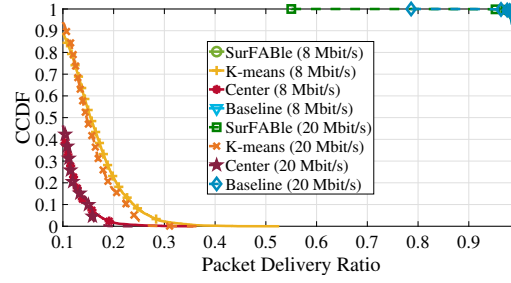
Figure 6.22: Simulation average performance results for five networking scenarios consisting of 20 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering TCP traffic.

Operating System [219]. All the communications nodes were able to operate at 2.4 GHz and 5 GHz. The *Raspberry Pi* devices were selected due to their cost-effectiveness and portability. In addition, the *TP-Link TL-WR902AC v3* [218] AP model was selected due to its low weight, small dimensions, and low power consumption, which make it suitable to be used as a FAP. The experimental testbed is depicted in Fig. 6.25. The network configuration parameters of the experimental testbed are presented in Table 6.3.

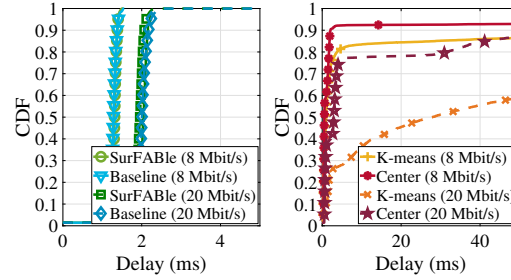
Without loss of generality, we carried out the performance evaluation of SurFABLE in the real-world using an access network consisting of ground devices only, in order to rule out the challenges associated with scalability and logistics that a greater number of communications nodes



(a) Throughput CCDF.



(b) PDR CCDF.



(c) Delay CDF.

Figure 6.23: Simulation average performance results for five networking scenarios consisting of 40 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, considering TCP traffic.

in a more complex network would require. Line-of-Sight between the communications nodes was ensured, in order to reproduce the FSPL model considered in the theoretical system modeling.

6.4.6 Experimental Results

In order to evaluate the performance achieved when using the SurFABLE algorithm, the wireless channel was first modeled in terms of the data rate achieved for different SNR values, as described in Section 6.4.6.1. Then, the network performance was assessed in terms of throughput, PDR, and delay, considering five networking scenarios consisting of five static ground devices, as presented in Section 6.4.6.2. The *iPerf3* tool [220] was used to 1) generate UDP CBR

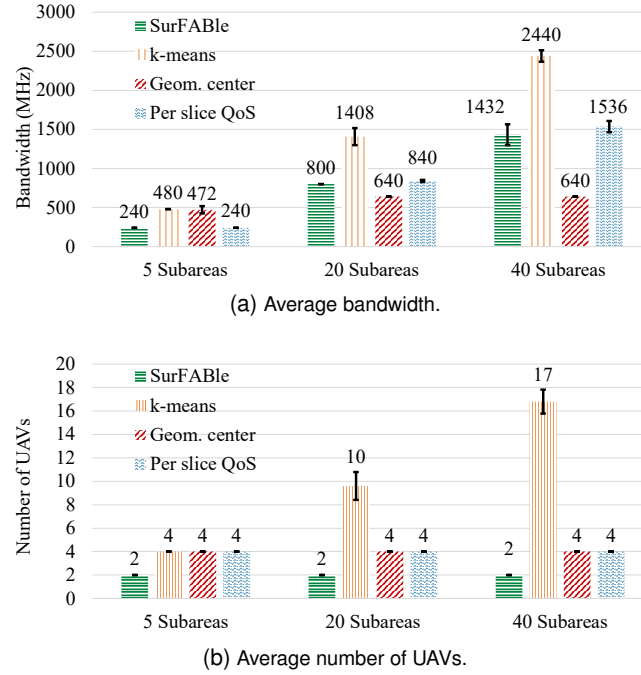


Figure 6.24: Average bandwidth and number of UAVs used in networking scenarios consisting of 5, 20, and 40 subareas randomly associated to URLLC (8 Mbit/s) and eMBB (20 Mbit/s) network slices, including 95% confidence intervals.

traffic from the ground devices towards the FAPs and 2) measure the throughput and PDR. The *ping* tool [221] was used to measure the Round-Trip Time (RTT); herein, we assume that the delay is equal to half the measured RTT. The Internet Control Message Protocol (ICMP) packets issued by *ping* were generated with a period of 100 ms, with no other traffic being exchanged simultaneously in the network.

6.4.6.1 Channel Modeling

In order to determine the data rate achieved in practice for different SNR values, experimental measurements were performed for different distances between a ground device and a FAP. The measurements were carried out using two ground devices in Line-of-Sight at different distances from each other, as presented in the data labels of Fig. 6.26.

Considering the set of SNR values and experimental data rates measured, linear regressions, which closely fit the experimental data obtained, were plotted. The linear regressions allow to define a continuous function for modeling the stepwise relation between SNR and the MCS indexes. This allowed to compute the data rate expected in practice for different SNR values by means of the empirical function obtained. Multiple measurements were carried out in a low noise environment and in a noisier environment, which led to respectively: 1) a more optimistic linear

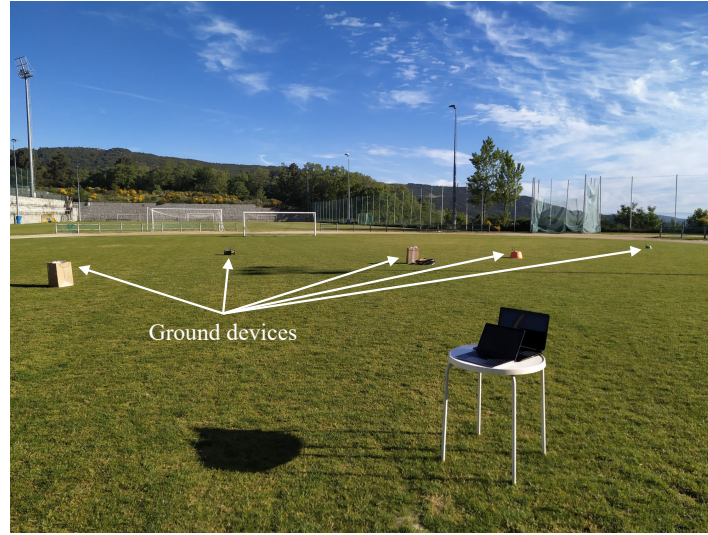
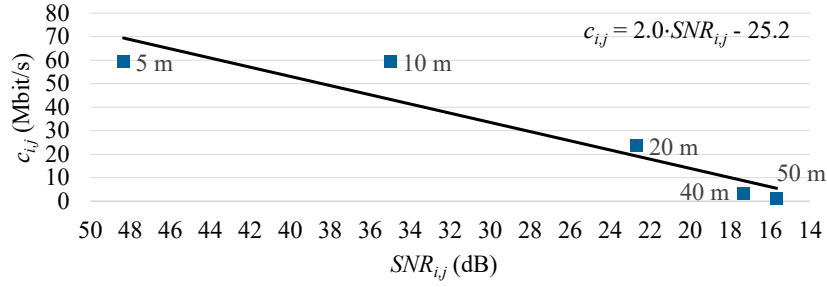


Figure 6.25: Ground devices deployed in the experimental testbed.

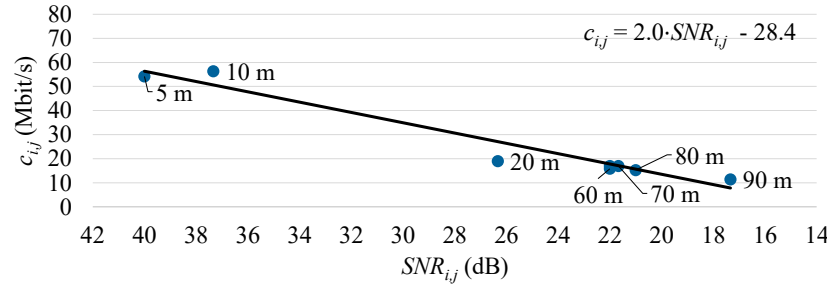
regression (cf. Fig. 6.26a), in which a lower SNR value is required to achieve a target data rate; and 2) a more pessimistic linear regression (cf. Fig. 6.26b), which implies greater SNR values (1.6 dB margin) to achieve the same target data rate. The different conditions of the considered environments were due to the interference from neighboring networks. The obtained linear regressions, which represent a real-world model for the wireless channel, were considered by SurFABLE (cf. step 5 in Algorithm 3) to compute the joint placement and allocation of the communications resources in the networking scenarios considered in the performance assessment carried out using the experimental testbed. The more optimistic linear regression was considered for the eMBB network slice, due to its low demanding reliability requirements. On the other hand, the more pessimistic linear regression was considered for the URLLC network slice, in order to meet its greater reliability requirements.

6.4.6.2 Performance Assessment

The performance assessment considered five networking scenarios, each of them consisting of five static ground devices randomly positioned in a $X = 40$ m, $Y = 40$ m area. The ground devices were randomly assigned to the URLLC and eMBB network slices. Each network slice consisted of a minimum of two ground devices for each networking scenario. For the sake of reproducibility, the positions and network slice types associated with the ground devices considered in the performance evaluation are presented in Table 6.1. The baseline solution considers a FAP placed in the geometric center of the positions of all the ground devices that belong to the same network slice. In addition, it provides each network slice with the same channel bandwidth computed by SurFABLE.



(a) Experimental capacity versus SNR for the eMBB network slice.



(b) Experimental capacity versus SNR for the URLLC network slice. It was obtained in a noisier environment with more interference from neighboring networks, which allowed to determine a more pessimistic linear regression – a higher SNR is required to achieve the same target data rate when compared with Fig. 6.26a, in order to meet the more demanding reliability requirements of the URLLC network slice.

Figure 6.26: Experimental wireless channel capacity modeled by linear regressions between SNR and data rates obtained in practice. The data labels at each marker refer to the Euclidean distances, in meters, between sender and receiver considered in the experimental measurements.

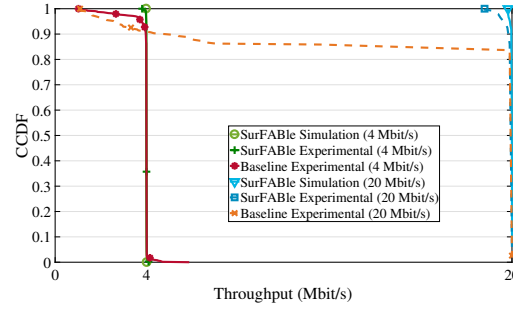
Considering the access network only, the solution obtained with SurFABLE consists of a single FAP_f for each network scenario, while the baseline uses two FAPs (one FAP_f for each network slice). The experimental results obtained in the networking scenarios that use a total channel bandwidth equal to 60 MHz and 80 MHz are presented in Fig. 6.27 and Fig. 6.28, respectively. The performance results obtained in simulation, considering the same exact conditions, are also presented for the SurFABLE algorithm. While SurFABLE and the baseline allow to meet the traffic demand of the URLLC network slice, with a slight degradation observed in the baseline (cf. 95th percentile in Fig. 6.27a, Fig. 6.27b, Fig. 6.28a and Fig. 6.28b), the throughput and PDR are significantly degraded for the eMBB network slice, due to the greater traffic demand associated with this network slice. The baseline offers the same channel bandwidth as SurFABLE, and uses twice as many UAVs, but the communications resources are not placed taking into account the QoS requirements of the network slices. Placing a FAP at the geometric center of the ground devices that belong to the same network slice allows providing the same SNR to each of them, but it is insufficient to ensure that the wireless links have high enough capacity to carry the offered traffic in some

networking scenarios. On the other hand, the delay achieved with the baseline is lower than the delay measured when using SurFABLE (cf. Fig. 6.27c and Fig. 6.28c). This is justified by the air-time fairness ensured by the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism of IEEE 802.11. In fact, when employing SurFABLE, the ground devices that belong to different network slices can be served by the same FAP. However, since the CSMA/CA mechanism assigns the same transmission opportunities to all the ground devices that use the same wireless channel, in order to ensure fairness in the medium access, the ground devices with a lower SNR value take longer to transmit the same amount of information. Moreover, it is worth mentioning that SurFABLE ensures a minimum SNR and channel bandwidth to accommodate the traffic offered by each ground device. On the other hand, placing a FAP at the geometric center of the ground devices associated to each network slice may lead to a higher SNR offered to some ground devices and induce higher MCS indexes when compared with SurFABLE, which ends up in a shorter transmission time enabled by the geometric center baseline solution. However, it must be noted that SurFABLE does not violate the QoS requirements associated with each network slice.

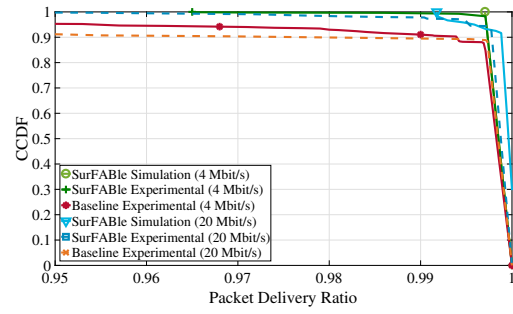
Considering the curves presented in Fig. 6.27 and Fig. 6.28, the deviation between the ns-3 simulation (represented by the green circle and blue downward-pointing triangle markers) and experimental (represented by the green plus and blue square markers) results is negligible, especially for throughput (cf. Fig. 6.27a and Fig. 6.28a) and PDR (cf. Fig. 6.27b and Fig. 6.28b). When it comes to the delay metric (cf. Fig. 6.27c and Fig. 6.28c), the simulation results for the SurFABLE algorithm are more optimistic than the experimental results. This may be justified by slightly differences between the simulation and experimental setup, including hardware particularities and associated computational delays. Moreover, as the *ping* tool was used, we assume that the delay is half of the RTT measured in practice. As in practice the wireless links are asymmetric, this may explain the deviation observed. However, it is worth noting that the delay results are of the same order of magnitude and show that SurFABLE does not violate the maximum average packet delay requirements associated with each network slice. Overall, the reduced deviation between the simulation and the experimental results demonstrates the accuracy of ns-3 for evaluating the performance of wireless networks and validates the performance evaluation carried out for multiple networking scenarios, whose results are presented from Fig. 6.18 to Fig. 6.23.

6.5 Discussion

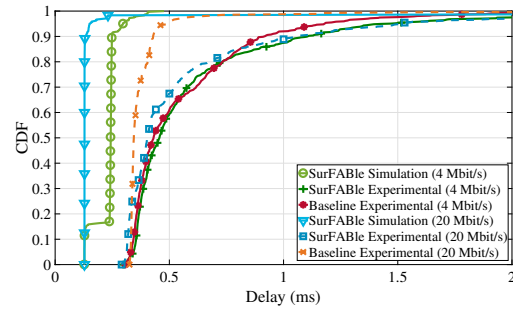
In this chapter, we presented SurFABLE, an algorithm that allows the computation of the joint placement and allocation of communications resources in slicing-aware FMNs. In the SurFABLE design, some theoretical assumptions were made to formulate and solve the problem.



(a) Throughput CCDF.



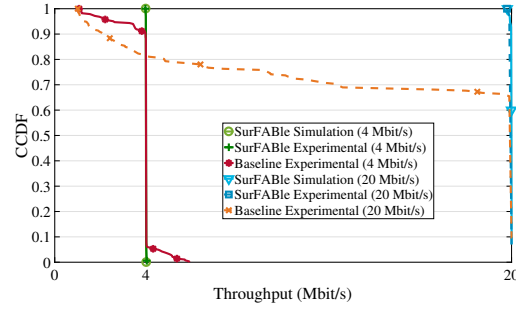
(b) PDR CCDF.



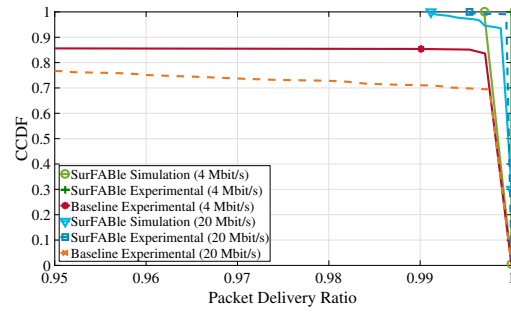
(c) Delay CDF.

Figure 6.27: Experimental and simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (4 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP CBR traffic and channel bandwidth of 60 MHz.

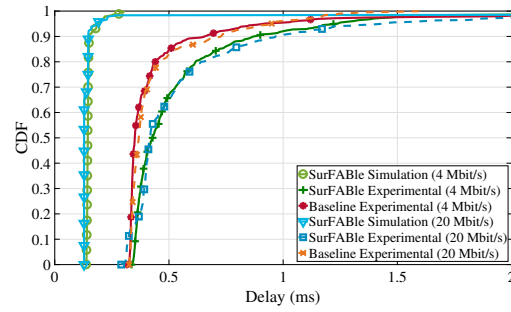
First, the FSPL model was used. This propagation model is commonly considered in the literature to characterize the wireless channel available between UAVs and ground devices, due to the strong Line-of-Sight component induced by the UAV altitude [21]. The FSPL model takes into account the 3D Euclidean distance between the UAVs and ground devices for determining the received power. Moreover, since we assume the use of omnidirectional antennas, which conceptually radiate equal power in all directions, the influence of heterogeneous radiation patterns and the heading of the UAVs are not considered. Yet, SurFABLE is agnostic to the radio propaga-



(a) Throughput CCDF.



(b) PDR CCDF.



(c) Delay CDF.

Figure 6.28: Experimental and simulation average performance results for five networking scenarios consisting of 5 subareas randomly associated to URLLC (4 Mbit/s) and eMBB (20 Mbit/s) network slices, considering UDP CBR traffic and channel bandwidth of 80 MHz.

tion model used. Different propagation models can be considered, according to the environment where the UAVs are deployed, the wireless communications technology used, and the network configuration employed.

In order to meet the reliability requirements of different network slices, SurFABLE allows considering different target BER values. This is depicted in Fig. 6.16, in which the wireless channel capacity is modeled by different linear regressions. Although we consider target BER values in the system modeling, the performance evaluation presented herein does not assess the reliabil-

ity requirements achieved with SurFABLE with an accuracy of 10^{-9} . The evaluation of the BER guarantees enabled by SurFABLE is left for future work, since herein we are focused on ensuring throughput and delay guarantees.

On the one hand, the throughput requirements are ensured by allocating to all subareas an amount of communications resources that provide high enough capacity to accommodate the offered traffic. On the other hand, the delay requirements are met by defining the minimum possible queue size for the network capacity of the wireless channel available, considering the M/D/1 queuing model for simplicity. The M/D/1 queuing model takes into account that the data packets have constant size and independent inter-arrival times. The obtained results show that the M/D/1 queuing model allows meeting the performance requirements for both UDP and TCP traffic. Yet, SurFABLE is agnostic to the queuing model used. Different queuing models can be considered, according to the traffic generated by the applications served by the network slices.

Requirements regarding traffic demand, in bit/s, and delay, in seconds, are considered by SurFABLE for each network slice. These are important metrics from the network performance point of view, and they are fulfilled by ensuring minimum SNR values for the wireless links established, while defining suitable positions for the UAVs. It should be noted that the SNR experienced by the UAVs in real-world flying networks may present some difference with respect to the theoretical values considered by SurFABLE, namely due to deviations in the location of ground devices and interference experienced by the UAVs. This may affect network performance. To overcome the problem, the use of an SNR margin with respect to the theoretical values computed by means of the FSPL model may be considered in practice. This aspect is left for future research. Moreover, SurFABLE assumes that the wireless channel is symmetric. In networking scenarios with asymmetric wireless links, the lowest network capacity among the two directions should be considered.

For convenience, the performance evaluation of SurFABLE presented herein considers legacy IEEE 802.11 standards (IEEE 802.11n and IEEE 802.11ac). From the spectral point of view, this represents a worst-case evaluation approach, due to the relatively high minimum channel bandwidth block of 20 MHz possible. It is worth mentioning the lower the minimum channel bandwidth block enabled by the technology used the higher the efficiency of SurFABLE. This makes SurFABLE especially suitable for OFDMA-based technologies, such as 5G New Radio and IEEE 802.11ax, which allow to assign a given number of channel bandwidth blocks of a few kHz to individual devices in a more efficient manner. For these reasons, the results presented may be considered as a baseline for comparison with future research works that employ SurFABLE in flying networks using more efficient wireless communications technologies.

Finally, in order to limit the number of admissible solutions when solving the optimization problem formulated in Eq. (4.1), the 3D space where the UAVs can be placed is divided by SurFABLE into N equal and small fixed-size sub-cuboids so that a solution is obtained in a timely manner, according to the network reconfiguration period. In its current version, the size assumed by Sur-

FABLE for each sub-cuboid takes into account that changing the position of a UAV a few meters has negligible impact on network performance in practice. The size of each sub-cuboid should decrease as the network performance requirements increase, in order to improve the accuracy of the solution obtained, at the expense of a longer time to solve the optimization problem. The sub-cuboid side size is a configuration parameter, and can be changed according to the networking scenario and the requirements of the network slices considered.

When considering 10 m as the sub-cuboid side size, the average time spent to solve the problem was 3.65 ± 0.76 s (95% confidence interval) considering five networking scenarios consisting of 5 subareas (the least complex networking scenarios considered). For five networking scenarios consisting of 40 subareas (the most complex networking scenarios considered), the average time spent by SurFABLE increased to 54.12 ± 11.43 s (95% confidence interval). These results were achieved using an 11th Generation Intel® Core™ i5-1135G7 processor running at 2.40 GHz with 16 GB of RAM, which could be deployed at the Edge of the FMN. This makes SurFABLE suitable for use in the real-world, where the SLAs established between Service Providers and Mobile Network Operators are not expected to change minute by minute, as we assumed for defining the flying network reconfiguration period $\Delta t \gg 1$ s. The fine-tuning of the sub-cuboid side size and Δt is beyond the scope of this thesis.

Despite the theoretical assumptions made, ns-3 enabled a realistic performance evaluation, carried out under different networking scenarios. The ns-3 simulator allows taking into account the stochastic characteristics of wireless networks and the behavior of representative traffic generation models by means of a packet-driven simulation approach. In the ns-3 simulations performed, we considered Poisson and CBR as the traffic generation models, taking into account the assumptions made in the system modeling; however, any other representative traffic generation models, which should be selected according to the characteristics of the traffic generated by the communications services and online applications to be served by the network slices, can be integrated into the SurFABLE algorithm.

The SurFABLE algorithm presented herein takes advantage of the PQM mechanism, described in Chapter 3, to address the queue size management and ensure delay guarantees for each network slice. Moreover, the forwarding tables of the FAPs, used to exchange data traffic between the ground devices and the assigned FGWs in the two-hop architecture considered by SurFABLE, are defined by the RedeFINE mechanism, presented in Chapter 5. RedeFINE avoids disruptions in the communications paths when the UAVs move for reconfiguring the network according to the decisions taken by SurFABLE. SurFABLE is designed upon the GWP and Gateway Placement and Queue Management (GPQM) algorithms, explained respectively in Chapter 2 and Chapter 4, since it considers targeted QoS requirements, defined by the SLAs associated with each network slices offered, to place the UAVs. The UAVs are placed in order to guarantee wireless links with the minimum an SNR values that allow inducing the usage of an MCS indexes capable of accom-

modating the offered traffic. Still, it is worth mentioning that the GWP and GPQM algorithms are also suitable to be used in a slicing-aware RAN designed by a state of the art algorithm. For that purpose, the single FGW employed should be capable of carrying multiple NICs, in order to enable a high amount of channel bandwidth able to meet the network slices' QoS guarantees through the use of different wireless channels. SurFABLE represents a step forward to such approach, since it allows determining the joint placement of multiple FGWs and FAPs, while reducing the amount of communications resources used by the slicing-aware FMN.

Regardless of the baseline solutions considered, the results obtained in the performance evaluation carried out show that SurFABLE allows meeting the QoS levels associated with the network slices, using the minimum number of UAVs. Furthermore, the performance evaluation carried out using the experimental testbed allows confirming the results obtained by means of ns-3 simulations and validated the SurFABLE algorithm in the real-world.

Despite minimizing the energy consumption of the UAVs to improve the network lifetime is a relevant research challenge, it is beyond the scope of this thesis. As future work, we aim at evolving SurFABLE to improve the energy efficiency of slicing-aware FMNs, building upon our previous research works in [81, 82].

6.6 Summary

The challenges imposed to Mobile Network Operators have been exacerbated by the use of multiple services with different performance requirements. In disaster management scenarios, such as forest fires and floods, first responders demand mission-critical, reliable communications, while victims may be served using a best-effort wireless coverage service. In crowded scenarios, such as outdoor festivities, the staff, the media, and the spectators have different QoS requirements. Mobile Network Operators typically overcome the challenges imposed in these scenarios through network over-provisioning with fixed and mobile BSs leveraging the Cell-on-Wheels concept [70]. Nevertheless, they may fail providing adequate wireless coverage or lack the required communications resources. In addition, the unpredictability associated with the number of mobile devices and their mobility may affect the QoS provided. Despite network slicing has emerged in 5G networks to enable the use of different communications services on top of a shared network infrastructure, fixed BSs are typically considered for satisfying the QoS requirements of different network slices. This may not be feasible in scenarios characterized by limited network resources or dynamic behavior. FMNs, made up of UAVs acting as mobile BSs and APs, have emerged to provide on-demand wireless connectivity anywhere, anytime. However, they are typically designed for improving the overall network performance considering a best-effort approach or are optimized for a given communications service. The QoS levels offered are not considered or differentiated. Moreover, the joint UAV placement and allocation of communications resources for

enabling the usage of different communications services with heterogeneous QoS requirements for multiple UAVs has not been addressed so far.

We propose SurFABLE, an algorithm that allows the joint computation of the amount of communications resources needed, namely the number of UAVs acting as FAPs and FGWs and their positions in FMNs. Inspired by Capacitated Facility Location Problem, a classical optimization problem that aims at selecting the best among potential locations for a factory or warehouse, SurFABLE determines the number of UAVs acting as FAPs and FGWs and their positions, among a set of candidate locations, while defining the number of channel bandwidth blocks to allocate to the ground subareas where the network slices should be made available. The performance evaluation of SurFABLE was carried out by means of ns-3 simulations and an experimental testbed. ns-3 simulation results show that SurFABLE allows meeting heterogeneous QoS levels associated with different network slices using the minimum number of UAVs. Moreover, experimental results confirm the results obtained by means of ns-3 simulations for representative networking scenarios, while validating SurFABLE in the real-world.

In this chapter we demonstrated the feasibility of using a dynamic FMN, partitioned into any number and type of network slices associated with variable and heterogeneous QoS levels. Together with the solutions presented in the previous chapters, this validates our research hypothesis defined in Chapter 1.

Chapter 7

Conclusions

In this chapter, we present an overview of the work developed, review the main contributions of this thesis, and refer to research challenges that may be considered as part of future work.

7.1 Work Developed Overview

The widespread advent of small and low-cost Unmanned Aerial Vehicles (UAVs) has paved the way for the usage of swarms of interconnected UAVs. When compared with independent UAVs, UAVs in a swarm can cooperatively perform many tasks in a more efficient manner. This gave rise to the concept of Flying Multi-hop Network (FMN). The flexibility and modularity provided by an FMN, particularly the possibility of a rapid deployment and the ability of UAVs to adjust their positions in 3-Dimensional (3D) space, represent a step-forward with respect to terrestrial networks. In particular, wireless links in FMNs have a strong Line-of-Sight component, induced by the altitude factor, which allows avoiding obstacles to the radio signal propagation. Given these characteristics, FMNs are a promising and cost-effective solution for extending and reinforcing existing network deployments on-demand in a myriad of networking scenarios, including disaster management and temporary crowded events. The lack of solutions addressing the traffic-aware management of communications resources in FMNs consisting of UAVs motivates the development of novel networking solutions. This is further motivated by the recent popularity of UAVs and the emerging Non-Terrestrial networking paradigm under discussion in 3rd Generation Partnership Project (3GPP) and other fora, including 6G. These solutions aim at enabling communications for people and machines connected to each other.

The main goal of this thesis was the design of a solution addressing the traffic-aware management of communications resources in FMNs consisting of UAVs, in order to: 1) extend wireless coverage and capacity from a remote Access Point (AP) or Base Station (BS); 2) offer heterogeneous Quality of Service (QoS) levels to ground devices using the minimum number of UAVs and

radio resources; and 3) ensure uninterruptible wireless communications paths in dynamic FMNs. To attain the main goal the following research challenges were addressed:

- meet the different QoS levels of multiple network slices on top of a shared physical wireless network infrastructure;
- determine the minimum number of UAVs required and their placement, in order to simultaneously provide coverage with minimum target Signal-to-Noise Ratio (SNR) values to multiple network slices made available in different geographical areas;
- achieve a global admissible solution for the placement and allocation of communications resources, in a time interval compliant with the network reconfiguration period, while ensuring the minimum number of UAVs and wireless channels required to meet the QoS levels of multiple network slices.

In order to solve the problem addressed in this thesis, we proposed different algorithms and mechanisms acting as Network Functions tailored for service-based FMNs (AM2Fly). AM2Fly considers the inputs and outputs of each Network Function to achieve an admissible solution for the traffic-aware management of communications resources in FMNs. Due to the high complexity of the problem, the work was split in two parts. In the first part of the work, we have developed a set of algorithms and mechanisms considering a best-effort approach. First, the problem consisted in maximizing the aggregate throughput measured in the Flying Gateway (FGW). For that purpose, the GateWay Placement (GWP) algorithm was designed to enable wireless links between the Flying Access Points (FAPs) and the FGW with high enough capacity for accommodating the traffic forwarded to/from the Internet. Then, we designed the Proactive Queue Management (PQM) mechanism for FMNs, capable of defining in advance the FAPs' queue size according to the capacity of the wireless links established with the FGW and the traffic demand of the ground devices served. Since GWP and PQM were designed as independent algorithms, thus not benefiting from the gains of each other, we went a step forward and designed an algorithm for defining the placement of the FGW and suitable queue sizes for the FAPs, named Gateway Placement and Queue Management (GPQM). GPQM allows for wireless links with high enough capacity and network queues with a large enough size so that traffic bursts can be accommodated, while ensuring controlled delays for the transported traffic. The seamless FMN reconfiguration when the UAVs move for repositioning themselves is a challenge that may degrade the QoS offered, especially due to changes in the quality of the wireless links and topology changes. As such, we concluded the first phase of this thesis with the design of the Routing for Flying Multi-hop Networks (RedeFINE) mechanism. RedeFINE allows for improved QoS offered to the ground devices whilst avoiding disruptions in the communications paths when the UAVs move for reconfiguring the FMN.

In order to ensure heterogeneous QoS guarantees on top of a shared physical flying network infrastructure, in the second part of this thesis we designed a network slicing-aware approach

so that an FMN is able to support different communications services using the minimum amount of communications resources. For that purpose, we designed the Slicing-aware Flying Access and Backhaul (SurFABLe) algorithm, which makes the joint placement and allocation of communications resources by determining the number and 3D positions of multiple FAPs and FGWs as well as the number of wireless channels to be made available. SurFABLe allows meeting the requirements of any number and type of network slices using the minimum number of UAVs. The performance evaluation of AM2Fly was carried out by means of Network Simulator 3 (ns-3) simulations and an experimental testbed, considering representative networking scenarios observed in the real-world. The performance results obtained showed that the algorithms and mechanisms proposed by this thesis are able to meet heterogeneous QoS levels, while using a shared FMN, representing a step forward to the baselines and state of the art counterparts considered. Moreover, they validate the research hypothesis of this thesis: **it is possible using a dynamic flying multi-hop network, partitioned into any number and type of network slices associated with variable and heterogeneous QoS levels, to serve multiple ground devices without loss of connectivity.**

By placing and allocating communications resources in FMNs consisting of UAVs, this thesis paves the way for wireless communications infrastructure more resilient and efficient, while avoiding network over-provisioning of static communications resources. This represents a step forward to the approaches typically employed by Mobile Network Operators to reinforce the wireless coverage and network capacity in disaster management and temporary crowded events, which consider mobile cells mounted on ground vehicles, characterized by network planning challenges and limited mobility. Moreover, AM2Fly allows that the same shared physical FMN can be used to simultaneously serve different communications services and applications, which may be characterized by heterogeneous target QoS requirements.

7.2 Original Contributions

This thesis resulted in five original contributions. In this section, we recall each of them.

7.2.1 An Algorithm for Gateway Placement in FMNs (GWP)

The GWP algorithm is proposed to define a suitable placement for the FGW in FMNs with controlled topology. GWP takes advantage of the information on the traffic demand of the ground devices and the positions of the FAPs, defined by a state of the art FAP placement algorithm, in order to define the position of the FGW. GWP ensures that the wireless link established between each FAP and the FGW has high enough capacity to accommodate the traffic exchanged with the ground devices over time.

The state of the art solutions have been focused on the placement of UAVs as APs and BSs serving the ground devices, overlooking the backhaul network design. Even though the ground devices are directly affected by the QoS offered by the Radio Access Network (RAN), the backhaul network needs to be carefully designed in order to meet the variable traffic demand of the FAPs. GWP introduces a differentiating factor that lies on defining in advance the position of the FGW by taking advantage of the knowledge of the future positions of the FAPs being served, whose mobility is controlled in a Central Station.

7.2.2 A Mechanism for Proactive Queue Management in FMNs (PQM)

The PQM mechanism is proposed to define in advance the queue size of the UAVs over time, in order to maximize the throughput with stochastic delay guarantees. PQM benefits from the centralized and holistic view over the FMN available in a Central Station, including the positions of the FAPs and the QoS levels of the ground devices; the QoS levels are assumed to be defined in terms of the average traffic demand and the maximum average packet delay. Considering the future positions of the FAPs and the FGW, PQM ensures that the queue of each FAP has, at each time instant, the minimum size that allows accommodating the traffic sent towards the FGW.

It is worth noting that the GWP algorithm aims at ensuring target network capacity values for the wireless links established between the FAPs and the FGWs. Still, in a real-world communications node, the instantaneous traffic demand is difficult to estimate, including the maximum values and duration of traffic bursts. If the instantaneous traffic demand at a given FAP is greater than the network capacity available in the wireless link established with the FGW, the FAP holds packets in its queue. The PQM mechanism allows ensuring controlled delays by defining the minimum queue size for each FAP. PQM was designed upon the conclusion that the usage of static, oversized queues in wireless networks leads to underutilized communications nodes and unnecessary high delays without improving the aggregate throughput. By ensuring a minimum queue size, PQM allows the packets to be held in the queues for shorter time. Also, the queuing delay is decreased, the throughput is not compromised, and the shared medium usage is enhanced. This allows for improved overall network performance. When a target reliability level should be met in the QoS guarantees to be offered by the FMN, a trade-off for defining the queue size should be considered. A larger queue size allows for improved reliability at the expense of greater delay; a smaller queue size allows for shorter delay but lower reliability. PQM is focused on providing throughput and delay guarantees only.

The design of a single queue management mechanism able to meet the variable characteristics of wireless networks has been pointed out in the literature as a research challenge, especially due to the need of fine-tuning several parameters for properly configuring the communications nodes. This is exacerbated by the fact that state of the art Active Queue Management (AQM) solutions have been designed for wired networks, which are static in nature. Wireless networks in

general and FMNs in particular, due to their highly dynamic nature, bring up additional challenges that must be taken into account, including the variable channel capacity over time. PQM goes beyond the state of the art by dynamically managing the queue size in advance, according to the capacity of the wireless links and the traffic demand of the ground devices, leveraged on the knowledge about the future positions of the UAVs that compose the FMN.

7.2.3 An Algorithm for Placement and Queue Management in FMNs (GPQM)

The GPQM algorithm results from the combination of the GWP algorithm and the PQM mechanism. On the one hand, the PQM algorithm aims at maximizing the throughput of the traffic carried between the FAPs and the FGW, while meeting the average delay constraint. Since PQM configures only the queue size of the communications nodes, it is unable to find a suitable solution that meets the average delay constraint for all flows offered by the FAPs if the capacity of the wireless links is not high enough. On the other hand, the GWP algorithm aims at ensuring that the wireless link established between each FAP and the FGW has high enough capacity for carrying the offered traffic. For that purpose, the GWP algorithm defines the FGW placement and the transmission power of the UAVs. However, GWP does not address the queue size management of the communications nodes, which may imply long queuing delays when the queue size is not adequate for the capacity of the wireless links available. As such, GWP is not able to provide average delay guarantees.

The GPQM algorithm takes advantage of the information on the future positions of the FAPs, defined according to the traffic demand of the ground devices by a state of the art FAP placement algorithm, in order to: 1) determine the FGW position so that a wireless link able to accommodate the traffic forwarded by each FAP is established between each FAP and the FGW; and 2) compute the queue size for each FAP so that stochastic delay guarantees are provided. GPQM allows jointly defining an adequate placement for the FGW and a suitable queue size for the FAPs, which are aspects addressed independently by the state of the art solutions, without taking advantage of the gains allowed by each other.

7.2.4 A Mechanism for Predictive Routing in FMNs (RedeFINE)

The RedeFINE mechanism is proposed for selecting in advance the best wireless path for each FAP and the time instants each forwarding table is updated, acting in advance to the UAV placement. It allows for uninterruptible communications paths and includes a novel Inter-flow Interference-aware Routing (I2R) metric, tailored to centralized routing in FMNs. RedeFINE takes advantage of the centralized knowledge of the future positions and trajectories of the UAVs to predict the capacity of the wireless links. This holistic knowledge is available in a Central Station. The Central Station is running an independent state of the art UAV placement algorithm. The

state of the art UAV placement algorithm is used to define the optimal positions of the FAPs, in order to meet the traffic demand of the ground devices. RedeFINE can use the proposed I2R metric or any other routing metric suitable for centralized routing.

The proposed predictive routing mechanism represents a step forward with respect to existing routing solutions, in which the communications nodes only recover from link failures after they occur, which implies packet loss, throughput degradation, and non-zero reconfiguration times, due to the time wasted for updating the forwarding tables. Due to the highly dynamic behavior of FMNs, which induce frequent changes in the quality of the radio links and in the network topology, existing routing solutions need to constantly flood the network with route discovery packets, thus reducing the bandwidth available for data traffic. In order to address this challenge, RedeFINE and I2R do not use probe packets for neighbor discovery, interference estimation, and link quality prediction, benefiting from the information about the future positions of the FAPs available in the Central Station. This allows reducing the bandwidth required for control traffic.

7.2.5 An Algorithm for Placement and Allocation of Communications Resources in Slicing-aware FMNs (SurFABLE)

The SurFABLE algorithm is proposed to enable slicing-aware FMNs. For that purpose, it makes the joint computation of the amount of communications resources needed, namely the number of FAPs and FGWs, and their placement in 3D space. SurFABLE ensures that 1) the heterogeneous QoS levels of different network slices are met and 2) the capacity of each UAV in terms of the total number of wireless channels available is not exceeded. The resulting FMN allows extending wireless coverage using a multi-hop topology and offers multiple network slices that ensure the targeted QoS levels to ground devices using the minimum number of UAVs and radio resources. SurFABLE uses 1) PQM to compute a suitable queue size for each FAP and 2) RedeFINE to define the forwarding tables of the FAPs.

State of the art UAV placement algorithms have been especially designed for determining the UAV positions that maximize the wireless coverage or the QoS offered to the ground devices using a best-effort approach. When it comes to slicing in FMNs, the literature has been focused on UAVs acting as clients of Mobile Network Operators' infrastructures and the QoS levels offered are not considered or differentiated. The state of the art works that take advantage of UAVs for providing on-demand network slices aim at improving energy-efficiency, fair coverage, and resource allocation efficiency. However, they do not consider the joint placement and allocation of communications resources. SurFABLE aims at paving the way to integrate UAVs into Mobile Network Operators' infrastructures as on-demand communications resources that can be deployed anywhere, anytime, while meeting different QoS levels of any number and type of network slices that coexist on top of a shared physical flying network infrastructure.

7.3 Future Work

In this section, we point out research challenges that may be considered as part of future work.

Enable the placement of multiple FGWs serving clusters of FAPs. Multiple FGWs may be required when: 1) a single wireless channel does not provide high enough capacity to meet the traffic demand of all FAPs in the network and 2) the FAPs are placed over a wide area, such that a FAP may get out of reach of the FGW it is currently assigned to. Clusters of closer FAPs can be formed for this purpose, and a dedicated FGW can be assigned to each cluster of FAPs. The GWP and GPQM algorithms may be evolved to define clusters of FAPs and positioning of the FGW associated to each cluster may be considered. The placement of multiple FGWs so that the overall network performance is improved may be considered as well; this may imply using different connections to the Internet (e.g., cellular BSs, Wi-Fi APs, and tethered drones).

Design an energy-aware placement algorithm for multiple UAVs. In [81, 82, 222], it was concluded that moving UAVs at optimal speed values results in lower power consumption than placing the UAVs in hovering state. Yet, the solutions proposed so far considering this conclusion are tailored for a single UAV. Building upon the algorithms proposed in this thesis for the placement of UAVs in hovering state, future work may consider defining suitable trajectories and speed values for multiple UAVs, in order to minimize the energy spent by them for propulsion, without compromising the QoS offered. This work is already being developed in the context of an M.Sc. dissertation under the supervision of the author.

Ensure seamless QoS-aware handovers for the ground devices. The predictive routing mechanism for FMNs proposed by this thesis ensures uninterrupted wireless communications paths between the FAPs and the FGWs. Evolving the concept so that uninterrupted communications are ensured for the ground devices served by the FAPs may be considered as future work. The concept of Light Virtual Access Points [223] may be explored, which abstracts the association state away from the physical AP itself. Considering a Software-Defined Networking (SDN) controller to which all the FAPs are connected, the concept allows creating a virtual FAP for each ground device. As such, each ground device is only aware of the connection to a single virtual FAP, even if the physical FAP to which the ground devices is connected to changes due to an FMN reconfiguration. When it comes to seamless wireless connectivity, this challenge has been addressed in [224], but employing AM2Fly to ensure QoS guarantees is worth considering.

Specify a UAV position-aided directional wireless backhaul network. In this thesis we considered omnidirectional antennas for establishing wireless links. The use of high-frequency bands, such as Millimeter Wave (mmWave), THz, and visible-light based spectrum, will enable establishing ultra-high bandwidth wireless backhaul links, using directional antennas. Directional antennas are characterized by a limited beamwidth and the achievable throughput is variable across the entire beamwidth (i.e., the throughput is greater at 0° of the antenna's beamwidth and decreases sideways). As such, the yaw of each UAV has a direct impact on the throughput achieved in the wireless links established between peer UAVs. From the network performance point of view, the optimal solution is achieved by using for each wireless link a pair of dedicated UAVs with their antennas perfectly aligned with each other; however, this approach requires more UAVs. As such, a tradeoff between network performance and the number of UAV should be considered. Yet, there are multiple open research challenges:

- Determine the maximum number of high-frequency wireless links a UAV is able to establish with its peer UAVs, considering available bandwidth, number of transceivers and antennas on-board, interference, power consumption, and UAV's payload weight and dimensions constraints.
- Define the yaw of each UAV so that the minimum number of UAVs is used, while wireless links with high enough capacity are ensured between peer UAVs.
- Align the radio beams between peer UAVs in a timely manner so that the high-frequency wireless links are not disrupted when the UAVs change their positions.
- Manage the available communications resources and network configuration in both time and space dimensions so that end-to-end paths ensured by means of high-directional wireless links established between peer UAVs are always available.
- Characterize how the UAV altitude affects the high-frequency communications channel modeling, including due to potential exposure to a high level of noise power.
- Investigate the influence of the UAV's operation, including propulsion, engine vibration, and electromagnetic interference, in the performance of high-frequency wireless links.

Design an obstacle-aware UAV placement algorithm. AM2Fly does not consider the presence of obstacles to radio signal propagation. Ensuring Line-of-Sight is a critical requirement in next-generation high-capacity wireless networks, especially when using wireless communications technologies relying on mmWave, THz, and visible light spectrum. Any obstruction to Line-of-Sight can compromise the wireless coverage and the network performance offered in practice. An obstacle-aware UAV placement algorithm, consisting of both obstacle and network slicing aware components, is worth designing. In particular,

providing the UAVs with vision capability (e.g., video cameras) to see the environment and each other, while employing radio frequency probes to sense the wireless medium, will allow determining the number of UAVs required and their positions so that Line-of-Sight wireless links are guaranteed and the QoS levels associated with multiple network slices are met.

Integrate AM2Fly in a Mobile Network Operator's 5G network. In 5G networks, 3GPP considers a Service Based Architecture, in which the control plane lies in a set of interconnected Network Functions, each with authorization to access the services of others. Network Functions are self-contained, independent software-based modules exchanging data with each other. Each Network Function exposes its functionality and data by means of a Service Based Interface (SBI) based on Representational State Transfer (REST) and Hypertext Transfer Protocol. Integrating AM2Fly in a Mobile Network Operator's 5G network, in order to enable the traffic-aware management of a RAN enabled by an FMN is worth addressing as future work. This challenge has been conceptually studied in [83, 225] for a 5G network enabled by a mobile robotic platform.

Moreover, deploying the algorithms and mechanisms proposed in this thesis as *xApps* running in a RAN Intelligent Controller specified by the Open Radio Access Network (O-RAN) Alliance [226] may be considered. The O-RAN Alliance focuses on making emerging RANs more open, virtualized, and interoperable, taking advantage of disaggregated RAN components based on Network Functions. The RAN Intelligent Controller is a software-based component in charge of controlling and optimizing the RAN functions in an O-RAN based architecture by means of software applications called *xApps*, which may be used to provide awareness to AM2Fly. This work is already being developed in the context of an M.Sc. dissertation under the supervision of the author.

Compare the performance of AM2Fly against Machine Learning (ML) based solutions. ML techniques, including Reinforcement Learning and Deep Reinforcement Learning, are being used as a promising alternative to solve wireless networking problems, including UAV placement [227, 228]. ML-based techniques are considered a flexible solution to optimize the management of communications resources in wireless networks, considering different environments and QoS requirements, while learning from experience rather than modeling the wireless channel. ML-based solutions have the potential to automatically fine-tune the most relevant variables that influence the traffic-aware management of communications resources in FMNs. Thus, it is worth comparing the FMN performance achieved when using ML-based solutions with the results obtained using AM2Fly.

Appendix A

Assessment of the Quality of the Solutions obtained for the Flying Gateway Placement

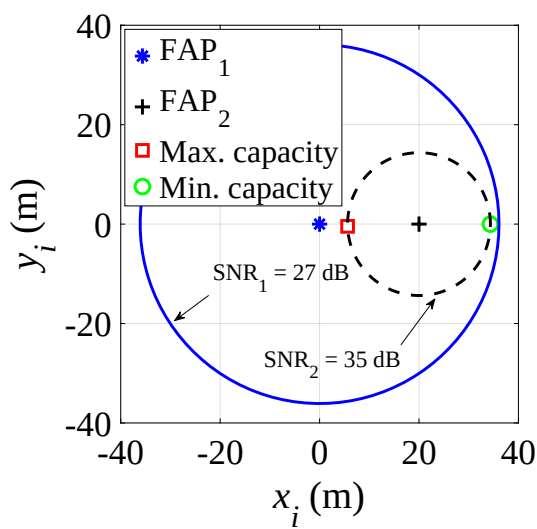
This appendix discusses the quality of the admissible solutions for the Flying Gateway (FGW) placement obtained when solving the optimization problem defined in Eq. (4.1). Let us consider the two representative networking scenarios depicted in Fig. A.1, which consist of two Flying Access Points (FAPs) placed at different distances from each other. They represent the two possibilities for which at least one solution for the FGW placement can be determined: fully overlapping circumferences and partially overlapping circumferences. When the circumferences associated with the FAPs do not intersect each other, there is no solution for the problem. The circumference associated with each FAP_i represents the maximum distance within which the FGW can be placed, so that the target capacity $C_{i,0}(t_k)$ demanded for the wireless link available between FAP_i and the FGW is guaranteed. A larger circumference is associated to a lower Signal-to-Noise Ratio (SNR) value, which enables a lower capacity $C_{i,0}(t_k)$, according to $B \cdot \log_2 \left(1 + 10^{\frac{SNR}{10}} \right)$, given by the Shannon-Hartley theorem [131]. For multiple FAPs, the FGW should be placed within the intersection of all the FAPs' circumferences. The SNR is determined by means of the Free-Space Path Loss (FSPL) model. In this theoretical analysis, we considered B equal to 160 MHz, P_T set to 20 dBm, 5250 MHz as the carrier frequency f_i , and constant noise power P_N equal to -85 dBm.

Admissible positions for the FGW placement are depicted by the green circles. In addition, admissible positions if the optimization problem aimed at maximizing the overall capacity $C(t_k)$ are represented by means of the red squares for comparison purposes, since maximization is a typical state of the art approach in optimization problems that aim at improving the overall network performance.

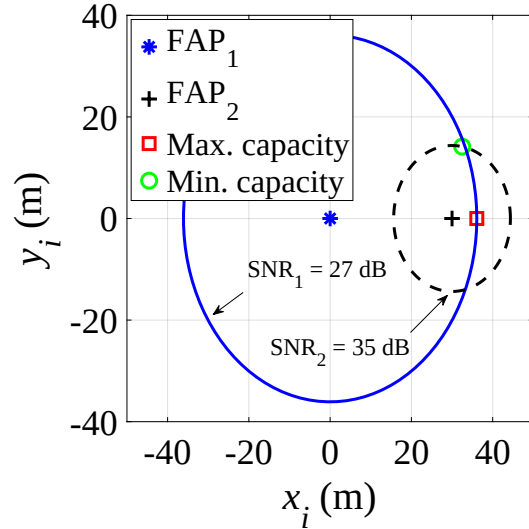
We conclude that, by minimizing the overall capacity $C(t_k)$ in our objective function, a suitable

solution is obtained, since the FGW is placed in positions (green circles) that ensure the minimum targeted capacity $C_{i,0}(t_k)$ for each FAP_{*i*}, which is limited by the circumferences (worst-case positions). On the other hand, if the positions associated with the maximum capacity (red squares) were chosen, by maximizing the solution in our objective function, the FGW would be placed closer to one of the FAPs, but from the perspective of the remaining FAPs the FGW would be placed over the corresponding circumference, which also represents a worst-case position as it occurs when minimizing the overall capacity $C(t_k)$.

For these reasons, we argue that minimizing the objective function defined in Eq. (4.1a) allows obtaining a suitable position for the FGW placement, using the minimum amount of communications resources, in contrast when the objective function is maximized. In order to take into account deviations between the theoretical and experimental values of $T_i(t_k)$ and $C_{i,0}(t_k)$ in real-world networking scenarios and accommodate bursty traffic, the value defined for $C_{i,0}(t_k)$ should include a margin with respect to the value of $T_i(t_k)$. This margin can be adjusted at each Flying Multi-hop Network (FMN) reconfiguration; its fine-tuning is left for future work.



(a) FAPs' transmission ranges (circumferences) **completely overlapping** each other.



(b) FAPs' transmission ranges (circumferences) **partially overlapping** each other.

Figure A.1: Admissible FGW positions for maximum and minimum overall capacity $C(t_k)$ values (red square and green circle, respectively).

References

- [1] Samsung Research. The Next Hyper-Connected Experience for All. https://cdn.codeground.org/nsr/downloads/researchareas/20201201_6G_Vision_web.pdf, 7 2020. Accessed on November 11, 2022.
- [2] Mojtaba Seyedzadegan, Mohamed Othman, Borhanuddin Mohd Ali, and Shamala Subramaniam. Zero-Degree algorithm for Internet GateWay placement in backbone wireless mesh networks. *Journal of Network and Computer Applications*, 36(6):1705–1723, 2013.
- [3] Wint Yi Poe and Jens B Schmitt. Minimizing the Maximum Delay in Wireless Sensor Networks by Intelligent Sink Placement. *Distributed Computer Systems Lab University of Kaiserslautern*, 67655:1–20, 2007.
- [4] Eduardo Nuno Almeida, Rui Campos, and Manuel Ricardo. Traffic-Aware Multi-Tier Flying Network: Network Planning for Throughput Improvement. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 04 2018.
- [5] Erlend Larsen, Lars Landmark, and Øivind Kure. Optimal UAV Relay Positions in Multi-Rate Networks. In *2017 Wireless Days*, pages 8–14. IEEE, 03 2017.
- [6] Chuan-Chi Lai, Chun-Ting Chen, and Li-Chun Wang. On-Demand Density-Aware UAV Base Station 3D Placement for Arbitrarily Distributed Users With Guaranteed Data Rates. *IEEE Wireless Communications Letters*, 8(3):913–916, 2019.
- [7] Xijian Zhong, Yan Guo, Ning Li, Yancheng Chen, and Shanling Li. Deployment Optimization of UAV Relay for Malfunctioning Base Station: Model-Free Approaches. *IEEE Transactions on Vehicular Technology*, 68(12):11971–11984, 12 2019.
- [8] Mohamed Alzenad, Muhammad Z Shakir, Halim Yanikomeroglu, and Mohamed-Slim Alouini. FSO-Based Vertical Backhaul/Fronthaul Framework for 5G+ Wireless Networks. *IEEE Communications Magazine*, 56(1):218–224, 2018.
- [9] Filipe B. Teixeira, Rui Campos, and Manuel Ricardo. Height Optimization in Aerial Networks for Enhanced Broadband Communications at Sea. *IEEE Access*, 8:28311–28323, 2020.
- [10] Ahmad Showail, Kamran Jamshaid, and Basem Shihada. WQM: An Aggregation-Aware Queue Management Scheme for IEEE 802.11n Based Networks. In *Proceedings of the 2014 ACM SIGCOMM Workshop on Capacity Sharing*, pages 15–20. ACM New York, NY, USA, 2014.

- [11] Shuai Zhang and Nirwan Ansari. Latency Aware 3D Placement and User Association in Drone-Assisted Heterogeneous Networks With FSO-Based Backhaul. *IEEE Transactions on Vehicular Technology*, 70(11):11991–12000, 2021.
- [12] Tong Bai, Cunhua Pan, Jingjing Wang, Yansha Deng, Maged ElKashlan, Arumugam Nallanathan, and Lajos Hanzo. Dynamic Aerial Base Station Placement for Minimum-Delay Communications. *IEEE Internet of Things Journal*, 8(3):1623–1635, 2021.
- [13] A.M. Koushik, Fei Hu, and Sunil Kumar. Deep Q-Learning-Based Node Positioning for Throughput-Optimal Communications in Dynamic UAV Swarm Network. *IEEE Transactions on Cognitive Communications and Networking*, 5(3):554–566, 2019.
- [14] Rukhsana Ruby, Hailiang Yang, Quoc-Viet Pham, and Kaishun Wu. Delay Performance of UAV-Based Buffer-Aided Relay Networks under Bursty Traffic: Mobile or Static? In *2021 IEEE 22nd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 51–60. IEEE, 2021.
- [15] Sidharth Kumar, Suraj Suman, and Swades De. Backhaul and Delay-aware Placement of UAV-enabled Base Station. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 634–639. IEEE, 2018.
- [16] Mouayad Horani and Mazen O. Hasna. Latency Analysis of UAV based Communication Networks. In *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 385–390, 2018.
- [17] Jin Li and Youngnam Han. Optimal Resource Allocation for Packet Delay Minimization in Multi-Layer UAV Networks. *IEEE Communications Letters*, 21(3):580–583, 2017.
- [18] Muntadher A. Ali, Yong Zeng, and Abbas Jamalipour. Software-Defined Coexisting UAV and WiFi: Delay-Oriented Traffic Offloading and UAV Placement. *IEEE Journal on Selected Areas in Communications*, 38(6):988–998, 2020.
- [19] Yousef Emami, Bo Wei, Kai Li, Wei Ni, and Eduardo Tovar. Joint Communication Scheduling and Velocity Control in Multi-UAV-Assisted Sensor Networks: A Deep Reinforcement Learning Approach. *IEEE Transactions on Vehicular Technology*, 70(10):10986–10998, 2021.
- [20] Abdurrahman Fouda, Ahmed S. Ibrahim, İsmail Güvenç, and Monisha Ghosh. Interference Management in UAV-Assisted Integrated Access and Backhaul Cellular Networks. *IEEE Access*, 7:104553–104566, 2019.
- [21] Eduardo Nuno Almeida, **André Coelho**, José Ruela, Rui Campos, and Manuel Ricardo. Joint traffic-aware UAV placement and predictive routing for aerial networks. *Ad Hoc Networks*, 118:102525, 2021.
- [22] Yong Zeng, Rui Zhang, and Teng Joon Lim. Wireless Communications with Unmanned Aerial Vehicles: Opportunities and Challenges. *IEEE Communications Magazine*, 54(5):36–42, 2016.

- [23] Stefano Rosati, Karol Kruzelecki, Gregoire Heitz, Dario Floreano, and Bixio Rimoldi. Dynamic Routing for Flying Ad Hoc Networks. *IEEE Transactions on Vehicular Technology*, 65(3):1690–1700, 2016.
- [24] Tao Chen, Marja Matinmikko, Xianfu Chen, Xuan Zhou, and Petri Ahokangas. Software Defined Mobile Networks: Concept, Survey, and Research Directions. *IEEE Communications Magazine*, 53(11):126–133, 2015.
- [25] Brian Barritt, Tatiana Kichkaylo, Ketan Mandke, Adam Zalcman, and Victor Lin. Operating a UAV Mesh & Internet Backhaul Network using Temporospacial SDN. In *2017 IEEE Aerospace Conference*, pages 1–7. IEEE, 2017.
- [26] Qihui Wu, Min Zhang, Chao Dong, Yong Feng, Yanli Yuan, Simeng Feng, and Tony Q. S. Quek. Routing protocol for heterogeneous FANETs with mobility prediction. *China Communications*, 19(1):186–201, 2022.
- [27] J. Vijitha Ananthi and P. Subha Hency Jose. Mobility Management UAV-based Grouping routing protocol in flying ad hoc networks for biomedical applications. *International Journal of Communication Systems*, 36(1):e5362, 2023.
- [28] Sang-Woong Lee, Saqib Ali, Mohammad Sadegh Yousefpoor, Efat Yousefpoor, Pooia Lalbakhsh, Danial Javaheri, Amir Masoud Rahmani, and Mehdi Hosseinzadeh. An Energy-Aware and Predictive Fuzzy Logic-Based Routing Scheme in Flying Ad Hoc Networks (FANETs). *IEEE Access*, 9:129977–130005, 2021.
- [29] Muhammad Morshed Alam and Sangman Moh. Joint topology control and routing in a UAV swarm for crowd surveillance. *Journal of Network and Computer Applications*, 204:103427, 2022.
- [30] Mohammed Gharib, Fatemeh Afghah, and Elizabeth Serena Bentley. LB-OPAR: Load balanced optimized predictive and adaptive routing for cooperative UAV networks. *Ad Hoc Networks*, 132:102878, 2022.
- [31] Min Zhang, Chao Dong, Simeng Feng, Xin Guan, Huichao Chen, and Qihui Wu. Adaptive 3D Routing Protocol for Flying Ad Hoc Networks Based on Prediction-Driven Q-Learning. *China Communications*, 19(5):302–317, 2022.
- [32] Haokeke Ye and Jing Liu. An Enhanced Q-Learning Routing Algorithm Based on Trajectory Prediction for UAV Networks. In *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5, 2021.
- [33] Min Zhang, Chao Dong, Peng Yang, Ting Tao, Qihui Wu, and Tony Q. S. Quek. Adaptive Routing Design for Flying Ad Hoc Networks. *IEEE Communications Letters*, 26(6):1438–1442, 2022.
- [34] HongGuang Zhang, Tao Wang, TingTing Liu, QingSong Zhang, and YuanAn Liu. Deep Neural Network Routing with Dynamic Space Division for 3D UAV FANETs. *Wirel. Pers. Commun.*, 125(3):2003–2028, 2022.

- [35] Muhammad Yeasir Arafat and Sangman Moh. A Q-Learning-Based Topology-Aware Routing Protocol for Flying Ad Hoc Networks. *IEEE Internet of Things Journal*, 9(3):1985–2000, 2022.
- [36] Yanpeng Cui, Qixun Zhang, Zhiyong Feng, Zhiqing Wei, Ce Shi, and Heng Yang. Topology-Aware Resilient Routing Protocol for FANETs: An Adaptive Q-Learning Approach. *IEEE Internet of Things Journal*, 9(19):18632–18649, 2022.
- [37] Open-Mesh. B.A.T.M.A.N. IV (TQ). https://www.open-mesh.org/projects/batman-adv/wiki/BATMAN_IV#BATMAN-IV-TQ, 2022. Accessed on December 30, 2022.
- [38] **André Coelho**, Helder Fontes, Rui Campos, and Manuel Ricardo. Placement and Allocation of Communications Resources in Slicing-aware Flying Networks. In *Wireless On-Demand Network Systems and Services Conference (WONS)*, pages 1–8. IFIP, 2022.
- [39] Katarzyna Kosek-Szott, Marek Natkaniec, and Andrzej R. Pach. A simple but accurate throughput model for IEEE 802.11 EDCA in saturation and non-saturation conditions. *Computer Networks*, 55(3):622–635, 2011.
- [40] Phuong Luu Vo, Minh N. H. Nguyen, Tuan Anh Le, and Nguyen H. Tran. Slicing the Edge: Resource Allocation for RAN Network Slicing. *IEEE Wireless Communications Letters*, 7(6):970–973, 2018.
- [41] Pablo Muñoz, ñscar Adamuz-Hinojosa, Jorge Navarro-Ortiz, Oriol Sallent, and Jordi Pérez-Romero. Radio Access Network Slicing Strategies at Spectrum Planning Level in 5G and Beyond. *IEEE Access*, 8:79604–79618, 2020.
- [42] Emmanouil Skondras, Emmanouel T Michailidis, Angelos Michalas, Dimitrios J Vergados, Nikolaos I Miridakis, and Dimitrios D Vergados. A Network Slicing Framework for UAV-Aided Vehicular Networks. *Drones*, 5(3):70, 2021.
- [43] Igor Donevski, Jimmy Jessen Nielsen, and Petar Popovski. Standalone Deployment of a Dynamic Drone Cell for Wireless Connectivity of Two Services. In *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–7. IEEE, 2021.
- [44] Peng Yang, Xing Xi, Kun Guo, Tony Q. S. Quek, Jingxuan Chen, and Xianbin Cao. Proactive UAV Network Slicing for URLLC and Mobile Broadband Service Multiplexing. *IEEE Journal on Selected Areas in Communications*, 39(10):3225–3244, 2021.
- [45] Hang Shen, Qiang Ye, Weihua Zhuang, Weisen Shi, Guangwei Bai, and Geng Yang. Drone-Small-Cell-Assisted Resource Slicing for 5G Uplink Radio Access Networks. *IEEE Transactions on Vehicular Technology*, 70(7):7071–7086, 2021.
- [46] Shashi Raj Pandey, Kitae Kim, Madyan Alsenwi, Yan Kyaw Tun, Zhu Han, and Choong Seon Hong. Latency-Sensitive Service Delivery With UAV-Assisted 5G Networks. *IEEE Wireless Communications Letters*, 10(7):1518–1522, 2021.

- [47] Yan Kyaw Tun, Yu Min Park, Tra Huong Thi Le, Zhu Han, and Choong Seon Hong. A Business Model for Resource Sharing in Cell-Free UAVs-Assisted Wireless Networks. *IEEE Transactions on Vehicular Technology*, pages 1–1, 2022.
- [48] Minsig Han, Jaewon Lee, Minjoong Rim, and Chung G. Kang. Dynamic Bandwidth Part Allocation in 5G Ultra Reliable Low Latency Communication for Unmanned Aerial Vehicles with High Data Rate Traffic. *Sensors*, 21(4), 2021.
- [49] Meryem Simsek, Adnan Aijaz, Mischa Dohler, Joachim Sachs, and Gerhard Fettweis. 5G-Enabled Tactile Internet. *IEEE Journal on Selected Areas in Communications*, 34(3):460–473, 03 2016.
- [50] Lingfeng Shen, Ning Wang, Zhengyu Zhu, Yajun Fan, Xiang Ji, and Xiaomin Mu. UAV-enabled Data Collection for mMTC Networks: AEM Modeling and Energy-Efficient Trajectory Design. In *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2020.
- [51] Saifur Rahman Sabuj, Derek Kwaku Pobi Asiedu, Kyoung-Jae Lee, and Han-Shin Jo. Delay Optimization in Mobile Edge Computing: Cognitive UAV-Assisted eMBB and mMTC Services. *IEEE Transactions on Cognitive Communications and Networking*, 8(2):1019–1033, 2022.
- [52] Xin Liu and Xueyan Zhang. Rate and Energy Efficiency Improvements for 5G-Based IoT With Simultaneous Transfer. *IEEE Internet of Things Journal*, 6(4):5971–5980, 2019.
- [53] Xin Liu and Xueyan Zhang. NOMA-Based Resource Allocation for Cluster-Based Cognitive Industrial Internet of Things. *IEEE Transactions on Industrial Informatics*, 16(8):5379–5388, 2020.
- [54] Xin Liu, Xiangping Bryce Zhai, Weidang Lu, and Celimuge Wu. QoS-Guarantee Resource Allocation for Multibeam Satellite Industrial Internet of Things With NOMA. *IEEE Transactions on Industrial Informatics*, 17(3):2052–2061, 2021.
- [55] Yeming Cai, Xu Jiang, Mingqian Liu, Nan Zhao, Yunfei Chen, and Xianbin Wang. Resource Allocation for URLLC-Oriented Two-Way UAV Relaying. *IEEE Transactions on Vehicular Technology*, 71(3):3344–3349, 2022.
- [56] Xing Xi, Xianbin Cao, Peng Yang, Jingxuan Chen, Tony Q. S. Quek, and Dapeng Wu. Network Resource Allocation for eMBB Payload and URLLC Control Information Communication Multiplexing in a Multi-UAV Relay Network. *IEEE Transactions on Communications*, 69(3):1802–1817, 2021.
- [57] Wen Wu, Conghao Zhou, Mushu Li, Huaqing Wu, Haibo Zhou, Ning Zhang, Xuemin Sherman Shen, and Weihua Zhuang. AI-Native Network Slicing for 6G Networks. *IEEE Wireless Communications*, 29(1):96–103, 2022.
- [58] N. H. Motlagh, M. Bagaa, and T. Taleb. UAV-Based IoT Platform: A Crowd Surveillance Use Case. *IEEE Communications Magazine*, 55(2):128–134, 02 2017.

- [59] D. Lee, A. Franchi, H. I. Son, C. Ha, H. H. Büthoff, and P. R. Giordano. Semiautonomous Haptic Teleoperation Control Architecture of Multiple Unmanned Aerial Vehicles. *IEEE/ASME Transactions on Mechatronics*, 18(4):1334–1345, 08 2013.
- [60] Minghui Dai, Ning Huang, Yuan Wu, Jie Gao, and Zhou Su. Unmanned Aerial Vehicles Assisted Wireless Networks: Advancements, Challenges, and Solutions. *IEEE Internet of Things Journal*, pages 1–1, 2022.
- [61] Ayon Chakraborty, Eugene Chai, Karthikeyan Sundaresan, Amir Khojastepour, and Sampath Rangarajan. SkyRAN: A Self-Organizing LTE RAN in the Sky. In *14th International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, pages 280–292. ACM New York, NY, USA, 2018.
- [62] Irem Bor-Yaliniz and Halim Yanikomeroglu. The New Frontier in RAN Heterogeneity: Multi-Tier Drone-Cells. *IEEE Communications Magazine*, 54(11):48–55, 2016.
- [63] Cihan Tugrul Cicek, Hakan Gultekin, Bulent Tavli, and Halim Yanikomeroglu. UAV Base Station Location Optimization for Next Generation Wireless Networks: Overview and Future Research Directions. In *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*, pages 1–6. IEEE, 2019.
- [64] Mohamed Alzenad, Amr El-Keyi, and Halim Yanikomeroglu. 3-D Placement of an Unmanned Aerial Vehicle Base Station for Maximum Coverage of Users With Different QoS Requirements. *IEEE Wireless Communications Letters*, 7(1):38–41, 2018.
- [65] Kirtan Gopal Panda, Shrayan Das, and Debarati Sen. Efficient UAV Placement Strategy for Guaranteed QoS Demand. In *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, pages 1–5. IEEE, 2020.
- [66] Mohammad Mozaffari, Walid Saad, Mehdi Bennis, and Merouane Debbah. Drone Small Cells in the Clouds: Design, Deployment and Performance Analysis. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2015.
- [67] Elham Kalantari, Muhammad Zeeshan Shakir, Halim Yanikomeroglu, and Abbas Yongacoglu. Backhaul-aware robust 3D drone placement in 5G+ wireless networks. In *2017 IEEE international conference on communications workshops (ICC workshops)*, pages 109–114. IEEE, 2017.
- [68] Z. Pi, J. Choi, and R. Heath. Millimeter-wave gigabit broadband evolution toward 5G: fixed access and backhaul. *IEEE Communications Magazine*, 54(4):138–144, April 2016.
- [69] Fengxiao Tang, Xuehan Chen, Ming Zhao, and Nei Kato. The Roadmap of Communication and Networking in 6G for the Metaverse. *IEEE Wireless Communications*, pages 1–15, 2022.
- [70] Jyh-Shyan Huang and Yao-Nan Lien. Challenges of Emergency Communication Network for Disaster Response. In *2012 IEEE International Conference on Communication Systems (ICCS)*, pages 528–532. IEEE, 2012.

- [71] Jim Gettys. Bufferbloat: Dark Buffers in the Internet. *IEEE Internet Computing*, 15(3):96–96, 2011.
- [72] 3GPP Portal. System Architecture for the 5G System (5GS). <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>, 2022. Accessed on January 16, 2023.
- [73] ns-3 website. ns-3 Network Simulator. <https://www.nsnam.org/>, 2023. Accessed on 06/12/2022.
- [74] **André Coelho**, Helder Fontes, Rui Campos, and Manuel Ricardo. Traffic-aware Gateway Placement for High-capacity Flying Networks. In *IEEE Vehicular Technology Conference*. IEEE, 2021.
- [75] **André Coelho**, Filipe Ribeiro, Helder Fontes, Manuel Ricardo, and Rui Campos. On the Experimental Evaluation of a Flying Network for Coverage and Capacity Extension in 5G+ Scenarios. In *INForum 2021 (Peer-reviewed Portuguese Symposium on Informatics Engineering)*, 2021.
- [76] **André Coelho**, Rui Campos, and Manuel Ricardo. Proactive Queue Management for Flying Networks. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*. IEEE, 2020.
- [77] **André Coelho**, Rui Campos, and Manuel Ricardo. Traffic-aware Gateway Placement and Queue Management in Flying Networks. *Ad Hoc Networks*, 138:103000, 2023.
- [78] **André Coelho**, Eduardo N. Almeida, Pedro Silva, José Ruela, Rui Campos, and Manuel Ricardo. RedeFINE: Centralized Routing for High-capacity Multi-hop Flying Networks. In *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 75–82, 10 2018.
- [79] **André Coelho**, Eduardo Nuno Almeida, José Ruela, Rui Campos, and Manuel Ricardo. A Routing Metric for Inter-flow Interference-aware Flying Multi-hop Networks. In *2019 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6. IEEE, 2019.
- [80] **André Coelho**, João Rodrigues, Helder Fontes, Rui Campos, and Manuel Ricardo. An Algorithm for Placing and Allocating Communications Resources Based on Slicing-Aware Flying Access and Backhaul Networks. *IEEE Access*, 10:128923–128942, 2022.
- [81] Hugo Rodrigues, **André Coelho**, Manuel Ricardo, and Rui Campos. Joint Energy and Performance Aware Relay Positioning in Flying Networks. *IEEE Access*, 10:1–16, 2022.
- [82] Hugo Rodrigues, **André Coelho**, Manuel Ricardo, and Rui Campos. Energy-aware Relay Positioning in Flying Networks. *International Journal of Communication Systems*, pages 1–17, 2022.
- [83] David Maia, **André Coelho**, and Manuel Ricardo. Obstacle-aware On-demand 5G Network using a Mobile Robotic Platform. In *2022 18th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 470–473. IEEE, 2022.

- [84] Max Friedrich, Teemu Joonas Lieb, Annette Temme, Eduardo N. Almeida, **André Coelho**, and Helder Fontes. ResponDrone – A Situation Awareness Platform for First Responders. In *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)*, pages 1–7. IEEE, 2022.
- [85] Gonçalo Santos, João Martins, **André Coelho**, Helder Fontes, Manuel Ricardo, and Rui Campos. A Fast Gateway Placement Algorithm for Flying Networks. In *IEEE Vehicular Technology Conference*. IEEE, 2021.
- [86] Renato Cruz, **André Coelho**, Rui Campos, and Manuel Ricardo. A Theoretical Model for Planning NB-IoT Networks. In *International Conference on Wireless and Mobile Computing, Networking and Communications*, pages 31–34. IEEE, 2019.
- [87] **André Coelho**, Mario Lopes, Bruno Ferreira, Rui Campos, and Manuel Ricardo. Experimental Evaluation of Shore to Unmanned Surface Vehicle Wi-Fi Communications. In *IFIP Wireless Days*, pages 86–91. IFIP, 2018.
- [88] INESC TEC. INESC TEC – Institute for Systems and Computer Engineering, Technology and Science. <https://www.inesctec.pt/en>, 2022. Accessed on January 16, 2023.
- [89] INESC TEC. WISE – Traffic-aware Flying Backhaul Mesh Networks. <http://wise.inesctec.pt/>, 2023. Accessed on January 16, 2023.
- [90] TANG Maolin. Gateways Placement in Backbone Wireless Mesh Networks. *International Journal of Communications, Network and System Sciences*, 2(01):44, 2009.
- [91] Valerio Targon, Brunilde Sansò, and Antonio Capone. The joint Gateway Placement and Spatial Reuse Problem in Wireless Mesh Networks. *Computer Networks*, 54(2):231–240, 2010.
- [92] Mohsen Jahanshahi and Arash Bozorgchenani. Gateway Placement and Selection Solutions in WMNs: A Survey. *arXiv preprint arXiv:1906.06774*, 2019.
- [93] Skanda N Muthaiah and C Rosenberg. Single Gateway Placement in Wireless Mesh Networks. *Proc. ISCN*, 8, 2008.
- [94] Yong Zeng, Rui Zhang, and Teng Joon Lim. Throughput Maximization for UAV-Enabled Mobile Relaying Systems. *IEEE Transactions on Communications*, 64(12):4983–4996, 2016.
- [95] E. N. Almeida, K. Fernandes, F. Andrade, P. Silva, R. Campos, and M. Ricardo. A Machine Learning Based Quality of Service Estimator for Aerial Wireless Networks. In *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–6. IEEE, 10 2019.
- [96] Qingqing Wu, Yong Zeng, and Rui Zhang. Joint Trajectory and Communication Design for Multi-UAV Enabled Wireless Networks. *IEEE Transactions on Wireless Communications*, 17(3):2109–2121, 2018.

- [97] Edgar Arribas, Vincenzo Mancuso, and Vicent Cholvi. Coverage Optimization with a Dynamic Network of Drone Relays. *IEEE Transactions on Mobile Computing*, 19(10):2278–2298, 2020.
- [98] Xijian Luo, Jun Xie, Liqin Xiong, Zhen Wang, and Chenjing Tian. 3-D Deployment of Multiple UAV-Mounted Mobile Base Stations for Full Coverage of IoT Ground Users with Different QoS Requirements. *IEEE Communications Letters*, 26(12):3009–3013, 2022.
- [99] Xijian Zhong, Yan Guo, Ning Li, and Yancheng Chen. Joint Optimization of Relay Deployment, Channel Allocation, and Relay Assignment for UAVs-Aided D2D Networks. *IEEE/ACM Transactions on Networking*, 28(2):804–817, 2020.
- [100] O. Khatib. Real-time Obstacle Avoidance for Manipulators and Mobile Robots. In *Proceedings. 1985 IEEE International Conference on Robotics and Automation*, volume 2, pages 500–505. IEEE, 1985.
- [101] G. Tychogiorgos, A. Gkelias, and K. K. Leung. Towards a Fair Non-convex Resource Allocation in Wireless Networks. In *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 36–40. IEEE, 2011.
- [102] Georgios Tychogiorgos. *Non-convex Resource Allocation in Communication Networks*. PhD thesis, Imperial College London, 2013.
- [103] P.M. Pardalos and G. Schnitger. Checking Local Optimality in Constrained Quadratic Programming is NP-hard. *Operations Research Letters*, 7(1):33–35, 1988.
- [104] Wireless LAN Professionals. MCS Table and How to Use It. <https://wlanprofessionals.com/mcs-table-and-how-to-use-it/>, 2023. Accessed on January 16, 2023.
- [105] Kathleen Nichols and Van Jacobson. Controlling Queue Delay. *Queue*, 10(5):20, 2012.
- [106] ns-3 website. CoDel Queue Disc. <https://www.nsnam.org/docs/models/html/codel.html>, 2023. Accessed on January 16, 2023.
- [107] Nils Aschenbruck, Raphael Ernst, Elmar Gerhards-Padilla, and Matthias Schwamborn. BonnMotion: A Mobility Scenario Generation and Analysis Tool. In *Proceedings of the 3rd international ICST conference on simulation tools and techniques*, page 51, 2010.
- [108] M. Shahwaiz Afaqui, Eduard Garcia-Villegas, and Elen Lopez-Aguilera. IEEE 802.11ax: Challenges and Requirements for Future High Efficiency WiFi. *IEEE Wireless Communications*, 24(3):130–137, 2016.
- [109] H. Elshaer, F. Boccardi, M. Dohler, and R. Irmer. Downlink and Uplink Decoupling: a Disruptive Architectural Design for 5G Networks. In *2014 IEEE Global Communications Conference*, pages 1798–1803. IEEE, 12 2014.
- [110] Changyang She, Chenxi Liu, Tony Q. S. Quek, Chenyang Yang, and Yonghui Li. UAV-Assisted Uplink Transmission for Ultra-Reliable and Low-Latency Communications. In *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE, 2018.

- [111] Dhulfiqar A. Alwahab and Sándor Laki. A Simulation-Based Survey of Active Queue Management Algorithms. In *Proceedings of the 6th International Conference on Communications and Broadband Networking*, pages 71–77, 2018.
- [112] Vivek Jain, Viyom Mittal, KS Shravya, and Mohit P Tahiliani. Implementation and validation of BLUE and PI queue disciplines in ns-3. *Simulation Modelling Practice and Theory*, 84:19–37, 2018.
- [113] Nader Bouacida and Basem Shihada. Practical and Dynamic Buffer Sizing Using *Learn-Queue*. *IEEE Transactions on Mobile Computing*, 18(8):1885–1897, 2018.
- [114] Tianji Li, Douglas Leith, and David Malone. Buffer Sizing for 802.11-Based Networks. *IEEE/ACM Transactions on Networking*, 19(1):156–169, 2010.
- [115] P.G. Kulkarni, S.I. McClean, G.P. Parr, and M.M. Black. Proactive Predictive Queue Management for improved QoS in IP Networks. In *International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL'06)*, pages 7–7, 2006.
- [116] Sally Floyd and Van Jacobson. Random Early Detection Gateways for Congestion Avoidance. *IEEE/ACM Transactions on networking*, 1(4):397–413, 1993.
- [117] Wu-chang Feng, Dilip Kandlur, Debanjan Saha, and Kang Shin. BLUE: A New Class of Active Queue Management Algorithms. Technical report, Technical Report CSE-TR-387-99, University of Michigan, 1999.
- [118] Chris V Holot, Vishal Misra, Don Towsley, and Wei-Bo Gong. On Designing Improved Controllers for AQM Routers Supporting TCP Flows. In *Proceedings IEEE INFOCOM 2001*, volume 3, pages 1726–1734. IEEE, 2001.
- [119] Rong Pan, Preethi Natarajan, Chiara Piglione, Mythili Suryanarayana Prabhu, Vijay Subramanian, Fred Baker, and Bill VerSteeg. PIE: A Lightweight Control Scheme to Address the Bufferbloat Problem. In *2013 IEEE HPSR*, pages 148–155. IEEE, 2013.
- [120] Ahmad Showail, Kamran Jamshaid, and Basem Shihada. Buffer Sizing in Wireless Networks: Challenges, Solutions, and Opportunities. *IEEE Communications Magazine*, 54(4):130–137, 2016.
- [121] Dimitri Bertsekas and Robert Gallager. *Data Networks (2nd Edition)*. Prentice-Hall, Inc., USA, 1992.
- [122] Tal Shapira and Yuval Shavitt. FlowPic: Encrypted Internet Traffic Classification is as Easy as Image Recognition. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 680–687. IEEE, 04 2019.
- [123] ns-3 website. Fifo Queue Disc. <https://www.nsnam.org/docs/models/html/fifo.html>, 2023. Accessed on January 16, 2023.

- [124] Toke Høiland-Jørgensen, Paul E. McKenney, Dave Täht, James Gettys, and Eric Dumazet. The Flow Queue CoDel Packet Scheduler and Active Queue Management Algorithm. *RFC*, 8290:1–25, 2018.
- [125] Changyang She, Chenxi Liu, Tony Q. S. Quek, Chenyang Yang, and Yonghui Li. Ultra-Reliable and Low-Latency Communications in Unmanned Aerial Vehicle Communication Systems. *IEEE Transactions on Communications*, 67(5):3768–3781, 2019.
- [126] Vishal Sharma and Rajesh Kumar. UAVs assisted queue scheduling in ground ad hoc networks. *International Journal of Ad Hoc and Ubiquitous Computing*, 30(1):1–10, 2019.
- [127] Zubair Md Fadlullah, Daisuke Takaishi, Hiroki Nishiyama, Nei Kato, and Ryu Miura. A Dynamic Trajectory Control Algorithm for Improving the Communication Throughput and Delay in UAV-Aided Networks. *IEEE Network*, 30(1):100–105, 2016.
- [128] Abdurrahman Fouda, Ahmed S. Ibrahim, Ismail Guvenc, and Monisha Ghosh. UAV-Based In-Band Integrated Access and Backhaul for 5G Communications. In *2018 IEEE 88th Vehicular Technology Conference (VTC2018-Fall)*, pages 1–5. IEEE, 2018.
- [129] Harald T Friis. A note on a Simple Transmission Formula. *Proceedings of the I.R.E. and Waves and Electrons*, 34(5):254–256, 1946. https://capmimo.ece.wisc.edu/capmimo_papers/friis_original_1946.pdf.
- [130] T Janevski, M Jankovic, and S Markus. Quality of Service Regulation Manual. *Telecommunication development Bureau*, page 173, 2017.
- [131] Ralph VL Hartley. Transmission of Information. *Bell System technical journal*, 7(3):535–563, 1928.
- [132] Raghavendra V Kulkarni and Ganesh Kumar Venayagamoorthy. Particle Swarm Optimization in Wireless-Sensor Networks: A Brief Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 41(2):262–267, 2010.
- [133] Logan DR Beal, Daniel C Hill, R Abraham Martin, and John D Hedengren. Gekko optimization suite. *Processes*, 6(8):106, 2018.
- [134] Aziz Altaf Khuwaja, Yunfei Chen, Nan Zhao, Mohamed-Slim Alouini, and Paul Dobbins. A survey of channel modeling for UAV communications. *IEEE Communications Surveys & Tutorials*, 20(4):2804–2821, 2018.
- [135] M. Mozaffari, W. Saad, M. Bennis, Y. Nam, and M. Debbah. A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems. *IEEE Communications Surveys & Tutorials*, 21(3):2334–2360, 2019.
- [136] ns-3 website. Traffic Control Layer. <https://www.nsnam.org/docs/models/html/traffic-control.html>, 2023. Accessed on January 16, 2023.
- [137] Andreas Wächter and Lorenz T Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106(1):25–57, 2006.

- [138] J. Kennedy and R. Eberhart. Particle Swarm Optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4, 1995.
- [139] Hossein Bidgoli. *The Handbook of Computer Networks, Volume 2, LANs, MANs, WANs, the Internet, and Global, Cellular, and Wireless Networks*. John Wiley & Sons, 2008.
- [140] ns-3 website. IdealWifiManager. <https://www.nsnam.org/docs/models/html/wifi-design.html#idealwifimanager>, 2023. Accessed on January 16, 2023.
- [141] ns-3 website. Mobility. <https://www.nsnam.org/docs/models/html/mobility.html>, 2023. Accessed on January 16, 2023.
- [142] Ilker Bekmezci, Ozgur Koray Sahingoz, and Şamil Temel. Flying Ad-Hoc Networks (FANETs): A survey. *Ad Hoc Networks*, 11(3):1254–1270, 2013.
- [143] Ozgur Koray Sahingoz. Networking Models in Flying Ad-Hoc Networks (FANETs): Concepts and Challenges. *Journal of Intelligent & Robotic Systems*, 74(1-2):513–527, 2014.
- [144] Md Hasan Tareque, Md Shohrab Hossain, and Mohammed Atiquzzaman. On the Routing in Flying Ad hoc Networks. In *Computer Science and Information Systems (FedCSIS), 2015 Federated Conference on*, pages 1–9. IEEE, 2015.
- [145] J. Jiang and G. Han. Routing Protocols for Unmanned Aerial Vehicles. *IEEE Communications Magazine*, 56(1):58–63, 01 2018.
- [146] Demeke Shumeye Lakew, Umar Sarad, Nhu-Ngoc Dao, Woongsoo Na, and Sungrae Cho. Routing in Flying Ad Hoc Networks: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, pages 1–1, 2020.
- [147] Antonio Guillen-Perez, Ana-Maria Montoya, Juan-Carlos Sanchez-Aarnoutse, and Maria-Dolores Cano. A Comparative Performance Evaluation of Routing Protocols for Flying Ad-Hoc Networks in Real Conditions. *Applied Sciences*, 11(10), 2021.
- [148] Ali H. Wheeb, Rosdiadee Nordin, Asma' Abu Samah, Mohammed H. Alsharif, and Muhammad Asghar Khan. Topology-Based Routing Protocols and Mobility Models for Flying Ad Hoc Networks: A Contemporary Review and Future Research Directions. *Drones*, 6(1), 2022.
- [149] Georgios A. Kakamoukas, Panagiotis G. Sarigiannidis, and Anastasios A. Economides. FANETs in Agriculture - A routing protocol survey. *Internet of Things*, 18:100183, 2022.
- [150] Omer T. Abdulhae, Jit Singh Mandeep, and Mt Islam. Cluster-Based Routing Protocols for Flying Ad Hoc Networks (FANETs). *IEEE Access*, 10:32981–33004, 2022.
- [151] Arnau Rovira-Sugranes, Abolfazl Razi, Fatemeh Afghah, and Jacob Chakareski. A review of AI-enabled routing protocols for UAV networks: Trends, challenges, and future outlook. *Ad Hoc Networks*, 130:102790, 2022.
- [152] Thomas H. Clausen and Philippe Jacquet. RFC 3626: Optimized Link State Routing Protocol (OLSR). RFC 3626, 2003.

- [153] Charles Perkins, Elizabeth Belding-Royer, and Samir Das. RFC 3561: Ad hoc On-demand Distance Vector (AODV) routing. RFC 3561, 2003.
- [154] Nicklas Beijar. Zone Routing Protocol (ZRP). *Networking Laboratory, Helsinki University of Technology, Finland*, 9(1):12, 2002.
- [155] Xianfeng Li and Junxian Huang. ABPP: An Adaptive Beacon Scheme for Geographic Routing in FANET. In *2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, pages 293–299. IEEE, 2017.
- [156] Ganbayar Gankhuyag, Anish Prasad Shrestha, and Sang-Jo Yoo. Robust and Reliable Predictive Routing Strategy for Flying Ad-Hoc Networks. *IEEE Access*, 5:643–654, 2017.
- [157] Min Song, Jing Liu, and Shixin Yang. A Mobility Prediction and Delay Prediction Routing Protocol for UAV Networks. In *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–6. IEEE, 2018.
- [158] Mehrdad Khaledi, Arnau Rovira-Sugranes, Fatemeh Afghah, and Abolfazl Razi. On Greedy Routing in Dynamic UAV Networks. In *2018 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*, pages 1–5. IEEE, 2018.
- [159] Benjamin Sliwa, Stefan Falten, and Christian Wietfeld. Performance Evaluation and Optimization of B.A.T.M.A.N. V Routing for Aerial and Ground-Based Mobile Ad-Hoc Networks. In *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pages 1–7. IEEE, 2019.
- [160] Xianfeng Li and Jiaojiao Yan. LEPR: Link Stability Estimation-based Preemptive Routing Protocol for Flying Ad Hoc Networks. In *2017 IEEE Symposium on Computers and Communications (ISCC)*, pages 1079–1084. IEEE, 2017.
- [161] Lin Lin, Qibo Sun, Shangguang Wang, and Fangchun Yang. A Geographic Mobility Prediction Routing Protocol for Ad Hoc UAV Network. In *2012 IEEE Globecom Workshops*, pages 1597–1602. IEEE, 2012.
- [162] Wang Qingwen, Liu Gang, Li Zhi, and Qi Qian. An Adaptive Forwarding Protocol for Three Dimensional Flying Ad Hoc Networks. In *2015 IEEE 5th International Conference on Electronics Information and Emergency Communication*, pages 142–145. IEEE, 2015.
- [163] Xianfeng Li and Jiaojiao Yan. LEPR: Link Stability Estimation-based Preemptive Routing protocol for Flying Ad Hoc Networks. In *2017 IEEE Symposium on Computers and Communications (ISCC)*, pages 1079–1084. IEEE, 2017.
- [164] Omar Sami Oubbati, Mohammad Mozaffari, Nouredine Chaib, Pascal Lorenz, Mohammed Atiquzzaman, and Abbas Jamalipour. ECaD: Energy-efficient routing in flying ad hoc networks. *International Journal of Communication Systems*, 32(18):e4156, 2019.
- [165] Xiao Cheng, Chao Dong, Haipeng Dai, and Guihai Chen. MOOC: A Mobility Control Based Clustering Scheme for Area Coverage in FANETs. In *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, pages 14–22. IEEE, 2018.

- [166] Do-Yup Kim and Jang-Won Lee. Joint Mission Assignment and Topology Management in the Mission-Critical FANET. *IEEE Internet of Things Journal*, 7(3):2368–2385, 2020.
- [167] Mohammed Gharib, Fatemeh Afghah, and Elizabeth Bentley. OPAR: Optimized Predictive and Adaptive Routing for Cooperative UAV Networks. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6. IEEE, 2021.
- [168] Qianqian Sang, Honghai Wu, Ling Xing, and Ping Xie. Review and Comparison of Emerging Routing Protocols in Flying Ad Hoc Networks. *Symmetry*, 12(6), 2020.
- [169] Rezoan Ahmed Nazib and Sangman Moh. Routing Protocols for Unmanned Aerial Vehicle-Aided Vehicular Ad Hoc Networks: A Survey. *IEEE Access*, 8:77535–77560, 2020.
- [170] Muhammad Asghar Khan, Alamgir Safi, Ijaz Mansoor Qureshi, and Inam Ullah Khan. Flying Ad-Hoc Networks (FANETs): A Review of Communication architectures, and Routing protocols. In *2017 First International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT)*, pages 1–9, 2017.
- [171] Yanan Zhang and Hongbing Qiu. DDQN with Prioritized Experience Replay-Based Optimized Geographical Routing Protocol of Considering Link Stability and Energy Prediction for UANET. *Sensors*, 22(13), 2022.
- [172] Benjamin Sliwa, Cedrik Schöler, Manuel Patchou, and Christian Wietfeld. PARRoT: Predictive Ad-hoc Routing Fueled by Reinforcement Learning and Trajectory Knowledge. In *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, pages 1–7. IEEE, 2021.
- [173] Brad Karp and H. T. Kung. GPSR: Greedy Perimeter Stateless Routing for Wireless Networks. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, MobiCom '00, page 243–254, New York, NY, USA, 2000. ACM New York, NY, USA.
- [174] Benjamin Sliwa, Daniel Behnke, Christoph Ide, and Christian Wietfeld. B.A.T.Mobile: Leveraging Mobility Control Knowledge for Efficient Routing in Mobile Robotic Networks. In *2016 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6, 2016.
- [175] A.E. Abdallah, T. Fevens, and J. Opatrny. High Delivery Rate Position-based Routing Algorithms for 3D Ad hoc Networks. *Computer Communications*, 31(4):807–817, 2008.
- [176] Solmaz Ghasemnezhad and Ali Ghaffari. Fuzzy Logic Based Reliable and Real-Time Routing Protocol for Mobile Ad hoc Networks. *Wireless Personal Communications*, 98(1):593–611, 2018.
- [177] Fariborz Entezami and Christos Politis. Three-Dimensional Position-Based Adaptive Real-Time Routing Protocol for wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2015(1):1–9, 2015.
- [178] Woo-Sung Jung, Jinhyuk Yim, and Young-Bae Ko. QGeo: Q-Learning-Based Geographic Ad Hoc Routing Protocol for Unmanned Robotic Networks. *IEEE Communications Letters*, 21(10):2258–2261, 2017.

- [179] Imtiaz Mahmud and You-Ze Cho. Adaptive Hello Interval in FANET Routing Protocols for Green UAVs. *IEEE Access*, 7:63004–63015, 2019.
- [180] Menglei Jiang, Qixun Zhang, Zhiyong Feng, Zhu Han, and Wei Li. Mobility Prediction Based Virtual Routing for Ad Hoc UAV Network. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2019.
- [181] Douglas S. J. De Couto, Daniel Aguayo, John Bicket, and Robert Morris. A High-Throughput Path Metric for Multi-hop Wireless Routing. *Wirel. Netw.*, 11(4):419–434, July 2005.
- [182] Richard Draves, Jitendra Padhye, and Brian Zill. Routing in Multi-Radio, Multi-Hop Wireless Mesh Networks. In *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking, New York, NY, USA, 2004*.
- [183] Yaling Yang, Jun Wang, and Robin Kravets. Interference-aware Load Balancing for Multihop Wireless Networks. Technical report, University of Illinois at Urbana-Champaign, 2005.
- [184] Anland Prabhn Subramanian, Milind M Buddhikot, et al. Interference Aware Routing in Multi-Radio Wireless Mesh Networks. In *IEEE workshop on wireless mesh networks (WiMesh)*, volume 9, pages 55–63. IEEE, 2006.
- [185] Axel Neumann, Corinna Aichele, Marek Lindner, and Simon Wunderlich. Better approach to mobile ad-hoc networking (BATMAN). *IETF draft*, pages 1–24, 2008.
- [186] S Skiena. Dijkstra’s algorithm. *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica, Reading, MA: Addison-Wesley*, pages 225–227, 1990.
- [187] ns3-aodv-etx repository. AODV ns-3 module with the ETX quality-aware metric. <https://github.com/neje/ns3-aodv-etx>, 4 2018. Accessed on January 16, 2023.
- [188] Thor-Framework repository. OLSR ns-3 module with the ETX quality-aware metric. <https://github.com/igorcompuff/ns-3.26>, 5 2017. Accessed on January 16, 2023.
- [189] Chiraz Houaidia, Hanen Idoudi, Adrien Van Den Bossche, Leila Azouz Saidane, and Thierry Val. Inter-flow and intra-flow interference mitigation routing in wireless mesh networks. *Computer Networks*, 120:141–156, 2017.
- [190] Tracy Camp, Jeff Boleng, and Vanessa Davies. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing*, 2(5):483–502, 2002.
- [191] G. R. Hiertz, D. Denteneer, S. Max, R. Taori, J. Cardona, L. Berleemann, and B. Walke. IEEE 802.11s: The WLAN Mesh Standard. *IEEE Wireless Communications*, 17(1):104–111, 02 2010.
- [192] 3GPP Portal. Management and Orchestration; Concepts, Use cases and Requirements. <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3273>, 2022. Accessed on January 16, 2023.

- [193] Quang-Trung Luu, Sylvaine Kerboeuf, and Michel Kieffer. Uncertainty-Aware Resource Provisioning for Network Slicing. *IEEE Transactions on Network and Service Management*, 18(1):79–93, 2021.
- [194] Quang-Trung Luu, Sylvaine Kerboeuf, and Michel Kieffer. Admission Control and Resource Reservation for Prioritized Slice Requests with Guaranteed SLA under Uncertainties. *IEEE Transactions on Network and Service Management*, pages 1–1, 2022.
- [195] Alexandros Kalokylos. A Survey and an Analysis of Network Slicing in 5G Networks. *IEEE Communications Standards Magazine*, 2(1):60–65, 2018.
- [196] Ibrahim Afolabi, Tarik Taleb, Konstantinos Samdanis, Adlen Ksentini, and Hannu Flinck. Network Slicing and Softwarization: A Survey on Principles, Enabling Technologies, and Solutions. *IEEE Communications Surveys & Tutorials*, 20(3):2429–2453, 2018.
- [197] Haitao Wu, Xin Wang, Qian Zhang, and Xuemin Shen. IEEE 802.11e Enhanced Distributed Channel Access (EDCA) Throughput Analysis. In *2006 IEEE International Conference on Communications*, volume 1, pages 223–228. IEEE, 2006.
- [198] Quang-Trung Luu, Sylvaine Kerboeuf, Alexandre Mouradian, and Michel Kieffer. Radio Resource Provisioning for Network Slicing with Coverage Constraints. In *2020 IEEE International Conference on Communications (ICC)*, pages 1–7. IEEE, 2020.
- [199] Pablo Caballero, Albert Banchs, Gustavo de Veciana, and Xavier Costa-Pérez. Multi-Tenant Radio Access Network Slicing: Statistical Multiplexing of Spatial Loads. *IEEE/ACM Transactions on Networking*, 25(5):3044–3058, 2017.
- [200] Salvatore D’Oro, Francesco Restuccia, Tommaso Melodia, and Sergio Palazzo. Low-Complexity Distributed Radio Access Network Slicing: Algorithms and Experimental Results. *IEEE/ACM Transactions on Networking*, 26(6):2815–2828, 2018.
- [201] Bin Han, Ji Lianghai, and Hans D. Schotten. Slice as an Evolutionary Service: Genetic Optimization for Inter-Slice Resource Management in 5G Networks. *IEEE Access*, 6:33137–33147, 2018.
- [202] Salvatore D’Oro, Francesco Restuccia, Alessandro Talamonti, and Tommaso Melodia. The Slice Is Served: Enforcing Radio Access Network Slicing in Virtualized 5G Systems. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, pages 442–450. IEEE, 2019.
- [203] Hongyu Xiang, Shi Yan, and Mugen Peng. A Realization of Fog-RAN Slicing via Deep Reinforcement Learning. *IEEE Transactions on Wireless Communications*, 19(4):2515–2527, 2020.
- [204] Giuseppe Faraci, Christian Grasso, and Giovanni Schembra. Design of a 5G Network Slice Extension With MEC UAVs Managed With Reinforcement Learning. *IEEE Journal on Selected Areas in Communications*, 38(10):2356–2371, 2020.

- [205] Johanna Andrea Hurtado Sánchez, Katherine Casilimas, and Oscar Mauricio Caicedo Rendon. Deep Reinforcement Learning for Resource Management on Network Slicing: A Survey. *Sensors*, 22(8), 2022.
- [206] Deborsi Basu, Soumyadeep Kal, Uttam Ghosh, and Raja Datta. SoftDrone: Softwarized 5G assisted drone networks for dynamic resource sharing using machine learning techniques. *Computers and Electrical Engineering*, 101:107962, 2022.
- [207] Antonino Masaracchia, Yijiu Li, Khoi Khac Nguyen, Cheng Yin, Saeed R. Khosravirad, Daniel Benevides Da Costa, and Trung Q. Duong. UAV-Enabled Ultra-Reliable Low-Latency Communications for 6G: A Comprehensive Survey. *IEEE Access*, 9:137338–137352, 2021.
- [208] Yixin He, Dawei Wang, Fanghui Huang, Ruonan Zhang, Xin Gu, and Jianping Pan. Downlink and Uplink Sum Rate Maximization for HAP-LAP Cooperated Networks. *IEEE Transactions on Vehicular Technology*, pages 1–1, 2022.
- [209] Peng Yang, Xing Xi, Tony Q S Quek, Jingxuan Chen, Xianbin Cao, and Dapeng Wu. Repeatedly Energy-Efficient and Fair Service Coverage: UAV Slicing. In *2020 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2020.
- [210] Mohammed Almekhlafi, Mohaned Chraiti, Mohamed Amine Arfaoui, Chadi Assi, Ali Ghrayeb, and Amira Alloum. A Downlink Puncturing Scheme for Simultaneous Transmission of URLLC and eMBB Traffic by Exploiting Data Similarity. *IEEE Transactions on Vehicular Technology*, 70(12):13087–13100, 2021.
- [211] Zhenhui Yuan and Gabriel-Miro Muntean. AirSlice: A Network Slicing Framework for UAV Communications. *IEEE Communications Magazine*, 58(11):62–68, 2020.
- [212] Praveenkumar Korrai, Eva Lagunas, Shree Krishna Sharma, Symeon Chatzinotas, Ashok Bandi, and Björn Ottersten. A RAN Resource Slicing Mechanism for Multiplexing of eMBB and URLLC Services in OFDMA Based 5G Wireless Networks. *IEEE Access*, 8:45674–45688, 2020.
- [213] Sanjay Melkote and Mark S. Daskin. Capacitated facility location/network design problems. *European Journal of Operational Research*, 129(3):481–495, 2001.
- [214] Gurobi Optimization. Gurobi Optimizer. <https://www.gurobi.com/solutions/gurobi-optimizer/>, 2022. Accessed on January 16, 2023.
- [215] Aristidis Likas, Nikos Vlassis, and Jakob J. Verbeek. The global k -means clustering algorithm. *Pattern Recognition*, 36(2):451–461, 2003.
- [216] Raspberry Pi Foundation. Raspberry Pi for home. <https://www.raspberrypi.org/>, 2022. Accessed on January 16, 2023.
- [217] Panda Wireless. Panda Wireless © N600 Dual Band (2.4 GHz & 5 GHz) Wireless-N USB Adapter With Dual 5 dBi Antennas. <http://www.pandawireless.com/pandaN600ant.html>, 2022. Accessed on January 16, 2023.

- [218] TP-Link USA Corporation. TL-WR902AC V3. <https://www.tp-link.com/us/support/download/tl-wr902ac/v3/>, 2022. Accessed on December 06, 2022.
- [219] OpenWrt website. OpenWrt Project. <https://openwrt.org/>, 2022. Accessed on December 06, 2022.
- [220] iPerf3 website. iPerf – The ultimate speed test tool for TCP, UDP and SCTP. <https://iperf.fr/>, 2022. Accessed on January 16, 2023.
- [221] die.net website. ping(8) - Linux man page. <https://linux.die.net/man/8/ping>, 2022. Accessed on January 16, 2023.
- [222] João Rodrigues. Slicing-aware Flying Communications Network. Master's thesis, Faculdade de Engenharia, Universidade do Porto, Portugal, 2022.
- [223] Julius Schulz-Zander, Lalith Suresh, Nadi Sarrar, Anja Feldmann, Thomas Hühn, and Ruben Merz. Programmatic Orchestration of WiFi Networks. In *2014 USENIX Annual Technical Conference*, pages 347–358. USENIX, 2014.
- [224] André Oliveira. Routing for Flying Networks using Software-Defined Networking. Master's thesis, Faculdade de Engenharia, Universidade do Porto, Portugal, 2019.
- [225] David Maia. Control and Positioning of a 5G Radio Access Node Deployed in a Mobile Robotic Platform. Master's thesis, Faculdade de Engenharia, Universidade do Porto, Portugal, 2022.
- [226] O-RAN Alliance. O-RAN Specifications. <https://www.o-ran.org/specifications>, 2023. Accessed on January 16, 2023.
- [227] Eduardo Nuno Almeida, Rui Campos, and Manuel Ricardo. Traffic-Aware UAV Placement using a Generalizable Deep Reinforcement Learning Methodology. In *2022 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6. IEEE, 2022.
- [228] Rúben Queirós, Eduardo Nuno Almeida, Helder Fontes, José Ruela, and Rui Campos. Wi-Fi Rate Adaptation using a Simple Deep Reinforcement Learning Approach. In *2022 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–3. IEEE, 2022.