A Reliable and Redundant Communication System for Safe Urban Air Mobility Operations

Vom Promotionsausschuss der
Technischen Universität Hamburg

zur Erlangung des akademischen Grades

Doktor-Ingenieurin (Dr.-Ing.)

genehmigte Dissertation (Monografie)

von

Shashini Thamarasie Wanniarachchi

aus

Colombo, Sri Lanka

2025

Date of Oral Examination	July 16th, 2025
Chair of Examination Board	Prof. Dr. Sibylle Schupp Institute of Software Systems Hamburg University of Technology
First Examiner	Prof. Dr. rer. nat. Volker Turau Institute of Telematics Hamburg University of Technology
Second Examiner	Priv.-Doz. Dr.-Ing. habil. Rainer Grünheid Institute of Communications Hamburg University of Technology

Acknowledgment

I am deeply grateful to TUHH for the opportunity to pursue my PhD. My sincere gratitude goes to Prof. Volker Turau, for his guidance and kindness. Most importantly his dedication and unwavering support, even after his retirement, through regular visits, emails, and video calls meant I never felt without a mentor. His great sense of humor also created a nice working environment. His invaluable advices truly shaped both my academic and personal growth. My appreciation also goes to Dr. Rainer Grünheid and Prof. Sibylle Schupp for being part of my PhD examination committee and their valuable feedback.

I also wish to thank the entire Institute of Telematics starting from Dr. Markus Venzke for his insightful suggestions and friendly conversations, and the administrative help during the last few months. I am thankful for Christoph Weyer, for always being quick and attending to help with technical issues, and Kristina Ahlborn, who assisted me with administrative matters ever since my early days at the institute. I'm also grateful to my colleagues Florian, Peter, Christian, Amine, Ivonne, and Nisal for the support, valuable discussions shared and for creating a friendly working environment.

My appreciation extends to the ULTRAS project partners: Jan, Philipp, and Andreas for the great collaborations, and to colleagues at the Institute of Communication Networks for the enriching experiences we shared. A special thanks to Dr. Koojana Kuladinithi for being a friend, colleague, and a consistent source of support.

To my friends in Hamburg, thank you for the joyful weekends that helped me unwind. Heartfelt gratitude to Dinithi and Ramesh for being my family in Germany, and for their love, as well as their constant emotional and motivational support. I extend special appreciation to Dr. Asanga Udugama for all the help and discussions, specially around OMNeT.

I would not be here without the initial encouragement to pursue the academic path. For this I'm grateful to Prof. Thushara Weerawardena and Dr. Andreas Könsgen. To Prof. Anna Foster, your example has inspired me since my master's, and I'm grateful for your guidance and encouragement.

Lastly, my deepest thanks go to my beloved family, the most important people in my life and my pillar of strength. My parents Shirani and Gnanaratne Wanniarachchi, my brother Dulitha and his family. Your unwavering support, unconditional love since day one, belief in me, and constant motivation made everything possible. I'm grateful for everything!

Abstract

Urban Air Mobility (UAM) is a new concept that aims at supplementing current public transportation systems in dedicated use cases. UAM implementation requires complex integration of different engineering fields, including command, control, and communication for safe operation. This dissertation explores the aspect of communication and proposes solutions to enhance the performance to support reliable and redundant communication for safe UAM operations. Network planning, channel modeling, ground station infrastructure for redundancy provision, and interference management are investigated. Simulation results show that employing 5^{th} generation communication standards with efficient network and resource management, particularly with device-to-device technology, provides the needed degree of performance.

Table of Contents

TABLE OF CONTENTS

Chapter 1

Introduction

Rapid technological advancements are currently seen in the fields of logistics, urban development, air transportation, etc. This has enabled the development of innovative transport solutions such as urban air mobility (UAM). This concept emerging as a supplementary public transportation mode aims to provide a possible solution to the prevailing high demand for urban public transportation. UAM entails the integration of fully autonomous aerial vehicles as a means of passenger and cargo transit, employing low-altitude urban airspace. Although UAM is a feasible solution, realizing such a complicated system requires the complex integration of various technical fields. These include demand modeling, ground infrastructure planning, control, navigation, aviation, and communication. Careful analysis of each of the fields is fundamental in such an implementation. Safe operations are non-negotiable, given that public transportation is also an intended use case of the UAM concept.

While all these aspects play a vital role in UAM implementations, this dissertation focuses on the communication sector. Being fully autonomous, these aerial vehicles operate without an onboard pilot or remote controller making UAM a highly time-sensitive application. In this context, the entire air taxi navigation relies on control decisions informed by real-time data available to the vehicle. Due to this dependency, the aerial vehicle must be equipped with methodologies that ensure continuous and timely availability of data pertaining to the surrounding environment. This necessitates the availability of a reliable and efficient communication network in the UAM system. Such a communication framework can ensure the safe operation of aerial vehicles in complex, dynamic, and densely populated urban areas.

However, the communication requirements in a UAM system go beyond traditional aviation needs due to differences in the operational environments. To this end, specifically tailored communication requirements have arisen, and special attention must be paid to aspects such as applicable technologies, network planning, channel conditions, and the dynamic nature of the communication conditions. Moreover, in such a system, two kinds of communication are involved. First is the communication between the aerial vehicles, referred to as Air-to-Air (A2A) communication. This is the primary carrier of information exchange that is important to the system architecture. Secondly, an air-to-ground (A2G) communication system is used to communicate between an aerial vehicle and an autonomous ground station. This is required to ensure continuous system availability, as faults in any system are inevitable. The ground station is an autonomous monitoring and data-gathering entity with no control capabilities. This communication framework is portrayed in Fig. 1.1.

Accordingly, UAM communication systems necessitate effective management of both these communication sectors and the mitigation of interference and latency issues while ensuring reliability and availability. The adoption of 5th generation (5G) and beyond communication technologies offers the possibility to overcome these challenges. 5G enables seamless connectivity and provides ultra-reliable, low latent networks with improved throughput and high capacity. However, given the critical nature of the operation, further system adaptations and improvements are necessary to make the communication system achieve the needed degree of performance.

It is to be noted that the whole domain of security-related, highly sensitive issues associated with communication establishment in critical systems like UAM is acknowledged. However, this dissertation does not explore the scope of secure communications through any of the topics discussed.

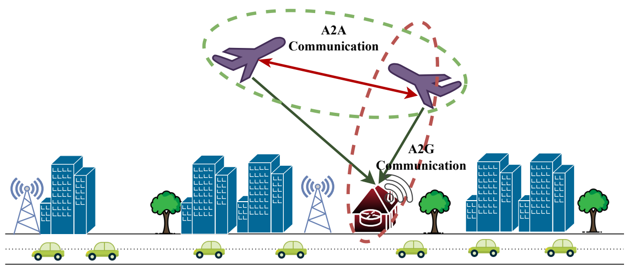


Figure 1.1: UAM communication system

1.1 Contributions

Given the above-discussed requirements, this dissertation is dedicated to investigating the aspects leading to establishing a reliable and redundant communication system for safe UAM operations. The existing research gaps in both the A2A and A2G communication sectors are explored. Possible solutions are proposed for these challenges and verified using simulation tools. 5G communication standards are considered as the technology provider. The considered aspects span multiple protocol layers, including the physical, medium access, and network layers. In terms of implementations, first a simulation environment is introduced as a contribution from this dissertation. An OMNeT++ based multi-agent simulation network for control and communication simulations is presented. For 5G protocol stack realization, the Simu5G library based on OMNeT++ is employed.

The channel model and the network plan play prominent roles when developing a communication system. In A2G communication, the direct applicability of existing channel models from the 3GPP 5G standardization is not viable due to the changes in the application environment conditions. Accordingly, the dissertation discusses and implements a channel model to suit low-altitude airspace communication. This is then extended into efficient and economical network planning. Here, a step-by-step approach is presented, which can be applied to different network planning requirements. Starting from mathematical models and using various heuristics, a network planning approach is proposed and verified with simulation results. The Hamburg metropolitan region is an application example, and a network plan to suit UAM communications in Hamburg is designed. Simulation results verify the enhanced system performance in terms of reliability and latency. This validates applying the approach to base station planning with careful consideration of the environment and network traffic conditions. Further analysis is conducted on applicable data types transmitted in a UAM scenario, such as images and videos. Based on simulation results, recommendations are made for the proper communication parameters that need to be employed.

It is to be noted that, regardless of using the most advanced technologies and system implementations, the possibility of running into faults in any system is inevitable. For this reason, the dissertation also contributes to redundant communication provision. While A2G communication is a redundant communication mode for A2A communications, conceptualizing the ground station architecture is essential to avoid issues such as single-point failure, enormous overhead, and network overload. As a solution, implementing a distributed ground station network architecture followed by a role delegation mechanism is investigated. This way, the mean time taken to repair the system is reduced. This ensures the highest possible system availability, addressing the issue of fault tolerance.

With regard to A2A communication, one of the main domains of research interests is interference management. The dissertation addresses this challenge and proposes solutions in two directions. First, the aspect of resource allocation is analyzed. Without restricting to static resource allocation approaches, the dissertation investigates means of dynamically adjusting the resource allocation strategy. This way, the communication system is updated based on prevailing network conditions. Accordingly, the reinforcement learning technique of Expected SARSA is studied, and a novel dynamic resource allocation technique is presented. Based on simulation results, more than double the system throughput improvement and more than 99.8% reliability provision is proven.

Secondly, the aspect of mode selection: deciding whether the communication happens via the direct device-to-device or infrastructure-based mode, is considered. An algorithm that evaluates the device proximity, SNR, and achievable bit rates is proposed. This is followed by a power allocation strategy based on the Expected SARSA reinforcement learning technique. The power control mechanism handles the trade-off between reducing interference while reaching high performance, grounded in the transmission power control. In combination, the mode selection and power control strategy demonstrate improvements in spectral efficiency and network throughput while maintaining high reliability and low delay. In addition, to further aid the interference management in the system, a tool that can be used for real-time tracking of network conditions, the *radio environment map*, is implemented for the UAM communication scenario.

1.2 Dissertation Structure

The remaining chapters of this dissertation are organized as follows. Chap. 2 discusses the UAM system in detail. This includes introducing the concept and extending it to present the components and requirements of the UAM system. Chap. 3 extensively discusses the existing communication technologies and the applicability of 5G to the application in the discussion. Chap. 4 is dedicated to introducing the research approach, including the simulation framework to suit the UAM scenario. In addition, this chapter provides an overview of the tools used and performance evaluation criteria applicable to the dissertation.

Chap. 5 delves into the challenge of realizing an A2G communication link specifically tailored to UAM communications by proposing a possible channel model. This is followed by a network planning approach, which leads to the proposal of an efficient and economical network plan. Next, in Chap. 6, a ground station network architecture and a concept of role delegation are discussed to ensure the highest possible system availability.

Chap. 7 and Chap. 8 both are focused on interference management in A2A communication. While Chap. 7 proposes a dynamic channel resource allocation strategy based on the Expected SARSA reinforcement algorithm, Chap. 8 analyses the challenge of mode selection along with a power allocation mechanism. Lastly, the dissertation is concluded with Chap. 9.

Chapter 2

Urban Air Mobility

This chapter sets the foundation for this dissertation with an overview of the background related to *urban air mobility* (UAM). From introducing the UAM concept and discussing the developments in the field, the chapter evolves into highlighting aspects such as the components of a UAM system and the system requirements.

2.1 Overview

According to the world population prospects 2024 by the United Nations Department of Economic and Social Affairs [Nat24], the global population of 8.2 billion at the moment in 2024 is expected to continue expanding for the upcoming 50 to 60 years with projections suggesting a peak of around 10.3 billion by mid-2080s. Moreover, The World Bank [The23] reports that more than half of the world's population is currently residing in urban areas, with urban populations projected to increase by 1.5 times by 2045. This refers to approximately 6 billion people residing in urban areas, accounting for close to 70%. Population growth leads to increased demands on resources, infrastructure, and services, generating significant challenges for sustainable development and resource provision on a global scale. To support this, the urban authorities must take action to improve city conditions, including efficient and affordable public transport modes, expanding green spaces, and enhancing air quality. Of these, special attention must be paid to public transportation as the rapid growth in population is leading to a high demand for public transit modes. Due to this increased demand, the traffic congestion becomes worse, causing frequent delays on the road for commuters. This has sparked a research interest in finding ways to reduce road traffic congestion and seeking efficient and innovative solutions to address this growing bottleneck in ground public transportation.

To cater to this requirement, extending the current ground transport network can be considered as a solution. However, such a ground network expansion is challenging due to several infrastructure and economic constraints. First, urban areas already consist of limited space, making it difficult to widen these roads or add new ones. Additionally, such expansions can be expensive, necessitating significant investment. These projects can also disrupt daily life, requiring lengthy construction periods. On the other hand, there are also environmental concerns related to such work, as road capacity enhancements reduce green spaces and contribute to air pollution. To this end, achieving the goal of ground traffic mitigation needs new policies, investments in improved infrastructure and public services, and encouragement for innovative alternative mobility solutions.

In terms of alternative transportation modes, several options can be considered. One option would be bike and scooter sharing, which is a micro-mobility option for short distances. Public transit expansion is another method achieved through metro, light rail, trams, etc. Water transit can also be considered for cities near water where ferries or water taxis can be employed as a mode of transportation. Carpooling and ride-sharing are other alternatives that would help reduce the number of vehicles on the road. An innovative solution among these is Urban Air Mobility (UAM), which uses autonomous aerial vehicles referred to as *air taxis* or drones to move passengers through the airspace. Each of these options sounds feasible, but each comes with certain disadvantages as well. Tbl. 2.1 is comparing the mentioned alternative transportation modes.

In comparison, it is clear that carpooling does not entirely help reduce ground traffic, while water transit and the sharing of bikes and scooters have geographical constraints. This narrows it down to the two options of public transit expansion and UAM. Between these two, UAM can be considered to deliver unique advantages over the public transit expansion, which is still a ground-based alternative. This is due to the ability of UAM to bypass ground traffic congestion entirely by offering an efficient and speedy mode of transportation, reduced infrastructure strain, and scalability for urban growth as it only requires less space on the ground for landing ports. As opposed to that, public transit expansion demands large ground areas for parking and rail road implementations.

Therefore, even though UAM has its challenges, such as high safety, technological, and regulatory approvals emphasizing airspace management, its potential to provide fast transportation across crowded urban areas, bypassing traffic congestion, makes it an appealing solution for future urban mobility. As a result, UAM is an emerging mode of public transportation concept aiming to reduce ground traffic accumulation and improve urban mobility within short

Transport Mode	Advantages	Disadvantages
Bike and Scooter sharing	Eco-friendly, space-efficient, flexible	Weather-dependent, limited to short distances, storage and safety issues
Public transit expansion	Moves larger number of people, reduces carbon emissions, cheap mode of transportation	System initiation and maintenance is expensive, limited flexibility, possibility of becoming overcrowded
Water transit	Leverages natural waterways, low emissions, scenic routes	Limited to coastal and cities with lakes or rivers, slow, requires docks and terminals
Carpooling and Ride-sharing	Reduce the number of cars on the road, affordable, flexible routes	contribute to traffic, dependent on demand, more emissions
UAM	Fast travel in urban areas, avoid road congestion, reduce land demand	High infrastructure cost, high safety requirements, limited capacity, requires regulation, noise, airspace management issues

Table 2.1: Alternative transport modes comparison

distances. The concept of UAM, therefore, entails the system of low-altitude air transportation focused on carrying passengers and goods and handling emergency services inside urban areas.

It is to be noted that UAM introduces new dynamics compared to traditional aviation due to unique operational, technological, and infrastructural demands. Operationally, UAM vehicles are designed for low-altitude, urban flights, navigating at heights where conventional aircraft typically don't operate. From a technical point of view, UAM is based on *Vertical Take-Off and Landing* (VTOL) technology, allowing them to vertically ascend and descend without requiring long runways as in conventional airplanes. This capability also shapes UAM infrastructure, where instead of airports, UAM requires only compact landing areas, like designated landing pads, supporting flexible integration within cities [NGS24].

When it comes to the topic of autonomous transportation systems, the first application that crosses our minds is autonomous driving. It is worth admitting that autonomous driving is a concept that has been under research and testing for many years now, and yet, the implementations are still in the early stages. Although Google laid the foundation of autonomous driving back in 2009 [Mar10], even after 15 years, there are still no vehicles in accordance with the full automation level 5 definition of [Int21]. At the moment, the tests are underway to fulfill the traffic management and collision avoidance mechanisms in automation level 4 [Int21], which only concerns autonomous driving in limited geographical areas [VD24].

In comparison with that, autonomous air taxis require an even higher level of traffic management and collision avoidance mechanisms, making the implementation significantly challenging. Autonomous air taxis operate in three-dimensional space with far fewer defined boundaries, requiring exceptionally sophisticated traffic management and collision avoidance systems. The aerial environment introduces challenges such as managing vertical and lateral spacing, accounting for variable weather conditions, and ensuring safe separation among multiple vehicles in dense urban areas. As a result, implementing reliable control and navigation systems for air taxis is notably more complex, as it demands precision and resilience far beyond that of autonomous driving systems on structured road networks. This is mainly due to the fact that the consequences of a passenger air taxi accident would be far more severe and critical than that of a ground vehicle accident.

The foundation for the UAM industry regulatory framework was set by the Federal Aviation Administration (FAA) in June 2016 by introducing operational rules for the civilian use of UAM vehicles. This initiative has fueled further exploration into the broader applications of aerial vehicles [WZZ20]. According to the EUROCONTROL Forecast 2024-2030, flight numbers are set to grow significantly, reaching 10.9 million by 2025, a 4.9% increase over the previous year. Growth is projected to average 2% annually beyond 2025, with over 12 million flights anticipated by 2030, signaling robust expansion and technological advancements in UAM [EUR24].

2.2 UAM System

With the advancements of automation technology in the 21st century, the current focus of UAM is mostly on highly automated aircraft. However, the UAM concept dates back to the early 1900s, when the idea of *flying cars* (Fig. 2.1) was developed. This evolved into the usage of helicopters in the late 1900s in the United States, followed by the usage of VTOL aircraft upgrading to *electric VTOL* (eVTOL) aircraft from the mid-2000s [CSF21]. Considering the air taxi structure, take-off and landing requirements, the interest in UAM is on aircraft that can vertically take off, hover, and land. To this end, in addition to helicopters, other VTOL aircraft, convertiplanes, gyrocopters, quadrocopters, and fixed-wing unmanned aerial vehicles (UAV), etc., are also of interest.

An air taxi can operate in various modes, including with a pilot on board, controlled remotely by a pilot on the ground, or fully autonomous without any human control [EHJ+20]. Each mode has distinct operational requirements

Figure 2.1: Convair: Model 118 flying car from San Diego Airspace Museum Archive

and safety protocols. Piloted air taxis rely on onboard decision-making, remotely piloted versions depend on real-time communication with ground control, and fully autonomous air taxis require advanced sensing, navigation, and collision avoidance systems to operate independently in complex urban airspaces. While UAM kept on progressing in the piloted aircraft sector, over the last decade, new research, implementations, and tests on unmanned aerial vehicles have continued to expand worldwide. The most common pilot-less aircraft design is the UAV, commonly known as the *drone*. This was first introduced as remotely operated and only for military purposes. Later developments leading to low-cost implementations opened UAV access to the general public [ZWZ19, GJV16]. This initiated applications in the entertainment industry, photography, agriculture, delivery, and surveillance, etc.

At the moment, the emerging trend is full automation, where the topic deals with unmanned aircraft that are not remote controller operated. The system entirely depends upon internal module interactions and real-time data exchange between the neighboring aircraft. Currently, the common applications of fully automated drones are seen in operations such as package delivery, remote sensing, agriculture, law and order, and emergency medical supply transportation in commercial and civilian sectors [ZZL16, WXZ+21]. One recent example of drone utilization was during the COVID-19 pandemic. These drones are very small in size (weighing less than 4.4 pounds) and meant to fly without people onboard in Line-of-sight (LOS) [GJV16]. According to the air and space law defined by the International Civil Aviation Organization, the traversable airspace is divided into seven classes [Web07] as depicted in Fig. 2.2. Here, the airspace, which is the segment of the atmosphere that is under the control or regulation of the country above where it is located, is named after the first seven letters of the alphabet according to the height levels. Out of that, the UAM aircraft occupy the low altitude airspace, also referred to as *class-G* airspace, and it is an uncontrolled airspace [Lab21].

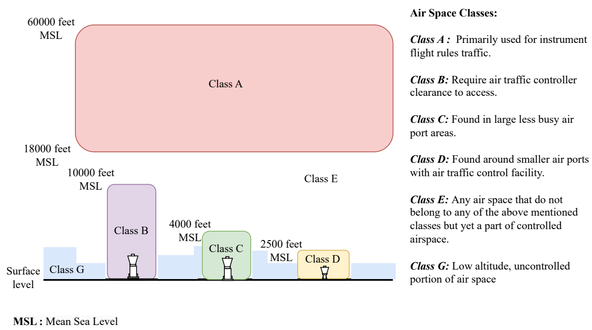


Figure 2.2: Standard Airspace classification according to [Web07]

The autonomous UAM solution is appropriate. Nevertheless, such implementations can contain both pre-evaluated and unforeseen risk situations associated with the airspace. When it comes to full automation in conjunction with UAM as a means of public transportation, the involved concerns are even higher compared to delivery and service drones. Because now, safe navigation is evidently non-negotiable. In addition, the involved traffic conditions are characterized by significant complexity and rapid changes. As a result, this concept requires the assessment of several key components spanning various technical fields. Air transportation regulations, airspace occupancy, trajectory planning, demand handling, control, navigation, sensing, and communication are a few examples of such technical disciplines that play an important role in handling the complications associated with a fully autonomous version.

2.3 System Architecture

The system architecture of UAM involves a complex integration of several key components that work together to enable efficient, safe, and scalable operations in urban environments. The central component of this architecture is the aircraft, which are typically autonomous or remotely piloted aerial vehicles designed to focus on urban air transit. Vertiports serve as air taxi take-off and landing pads that need to be strategically placed across the urban area. Additionally, ground stations are essential for air traffic coordination, as well as maintaining and managing regular communication with the air taxi to ensure safe UAM operations through real-time monitoring and control. To this end, these topics are further elaborated in this section.

2.3.1 Aircraft

From the different vehicle types used in UAM elaborated earlier in the chapter, UAVs are the main concern of this dissertation, and the focus is on full automation with no human pilot or remote controller involvement. These vehicles usually have four to eight propellers and rechargeable batteries. New technological concepts have been combined to make these vehicles much quieter than the helicopters [MVD+22], and command and control focusing passenger comfort have been a key priority. Passenger transportation uses larger versions of these UAVs and can weigh up to a few thousand kilograms depending on the structure, battery, propulsion systems, and the number of passengers [CSF21]. A typical air taxi can be a two to five-seater model [MVD+22]. At present, several companies have already built and tested some of the first UAM vehicles. EHang 116, Uber Elevate, Airbus A^3 Vahana, Boeing PAV, and Volocopter Volocity are some examples [MVD+22]. Fig.2.3 depicts a prototype of the air taxi model from Airbus referred to as *CityAirbus NextGen*, which has its first flight scheduled for later this year (2024) [Air24].

Figure 2.3: The prototype of an air taxi model - "CityAirbus NextGen" from Airbus. ©Airbus Helicopters 2024 Christian Keller

Considering the aircraft architecture, there are a few modules that are vital in the design to ensure smooth, autonomous operation. One such module is control and dynamics. This assures the proper control decision-making following the separation algorithms and navigating the air taxi in the planned trajectory, avoiding collisions. However, regardless of adhering to pre-planned trajectories, facing unexpected emergencies due to other air taxis or objects in the airspace is inevitable. Accordingly, the air taxi being occupied with a sensing and communication system that is highly available, coherent, and robust is paramount for unmanned air transportation [EHJ+20]. The sensing system needs to be occupied with several sensors such as a global positioning system (GPS), lidar sensor, camera, etc. This is to guarantee the possibility of multiple information reception from different sources. The fusion of these data creates information that could be applied in the command and control systems. Furthermore, by employing the communication system, this information needs to be exchanged between the neighboring air taxis so that all aircraft are revised in the surrounding environment.

2.3.2 Vertiports

For the air taxi system to operate seamlessly, one main aspect is the availability of infra-structure. The basic contributing infrastructural requirement is having a sufficient amount of take-off and landing locations. This module is referred to as a *vertiport*, and it is a landing pad that can accommodate several air taxis. Within each vertiport, individual spaces are designated for each air taxi, called *vertipads*, which are specifically sized for safe vertical take-offs and landings. The capacity and strategic placement of vertiports and vertipads are prominent for the UAM systems to ensure that air taxis can operate without congestion and delays [CSF21].

2.3.3 Ground Station

One other fundamental infrastructure component for UAM is the ground control center. As the focus is on full automation, this doesn't refer to a remote operation of the air taxi or the air traffic control tower as in the conventional air transport systems. The concept is similar, but rather an automated ground control center intended for monitoring purposes. A vital aspect of UAM for public transportation is guaranteeing safe maneuvering. Typically, air taxis are expected to interact with each other to exchange information. However, faults in any software system are unavoidable, and in the considered application, even a simple failure has the potential to cause critical aftereffects. Therefore, the purpose of the ground station is to aid with this as a redundant measure by providing a suitable infrastructure. The ground station can be equipped with a sensing system to monitor the airspace, maintaining a global overview by keeping track of air taxis and other objects, such as drones, balloons, birds, etc., movements that could be occupying the low-altitude airspace. In the context of air transportation, numerous emergency scenarios, such as adverse weather conditions and traffic conflicts, may suddenly arise, requiring prompt action to be taken. By regularly maintaining a global perspective on the surrounding environment, the ground station is capable of updating the air taxi with real-time information in case of such unsafe flying conditions. This facilitates safe maneuvering, preventing possible critical accidents.

2.4 System Requirements

In a UAM system, having only the appropriate aircraft architecture and infrastructure components is not sufficient to ensure safe operations. A range of specific requirements must be met to guarantee safety, which can be divided into two categories: pre-flight and in-flight requirements. Pre-flight requirements focus on operational readiness, including demand modeling, vertiport placement, flight route planning and deconfliction, etc. In-flight requirements ensure the aircraft's safe navigation, collision avoidance, and air taxi interactions. These considerations are paramount to the seamless and safe functioning of the UAM system and are described in this section. Fig.2.4 presents the modules inside the air taxi and the essential links between these modules for the efficient, robust, and safe operation of the air taxi network.

2.4.1 Pre-Flight Stage Requirements

In the pre-flight stage, one main aspect is demand modeling. A proper assessment of the demand needs to be done along with the other available public transportation modes and their usage. This helps identify high-demand routes and improve the integration points of UAM with other transit modes. Grounded on the analysis alongside the other public transportation modes, it's possible to deploy the UAM services strategically in such a way as to complement and enhance the overall transportation network, reducing traffic congestion and improving urban mobility. Moreover, in order to implement UAM in any urban space, planning the integration of the vertiport network into the existing urban environment is paramount. Being initiated as a means of passenger transportation, it is necessary to ensure the availability of vertiports with easy access to important venues in city districts. In addition to that, air taxis cannot fly for hours at a stretch and need to be recharged after navigating

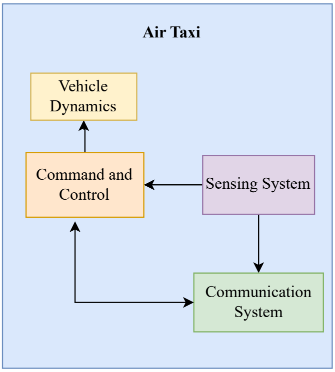


Figure 2.4: Air taxi module architecture showcasing the module interactions

for several kilometers. This service of recharging the batteries of air taxis can also be facilitated at vertiports. The requirement is to have several vertipads per vertiport, as, during non-peak times and nonoperational times, the air taxis need to be parked at these locations. One major concern with this would be the selection of locations in the city that are not close to airports as this can disrupt the airport operations, military bases, power plants, and generally crowded areas such as sports stadiums, etc. [CCB+21]. In air transport terminology, these areas are referred to as *no fly zones*. Accordingly, the top of tall buildings in the center of the city or open spaces in the city limits would be suitable locations for building vertiports similar to the concept of building helipads on top of hospitals and hotels, etc.

Another essential factor in the UAM pre-planning stage is trajectory planning and flight scheduling. This is fundamental in ensuring safe and efficient travel through densely populated airspaces without route conflictions. Trajectory planning involves determining the optimal routes between an origin and destination location pair and accounting for potential conflicts between air taxis and other air traffic. This process includes deconfliction strategies to prevent route overlaps and manage traffic. Scheduling flights relies on demand data and insights from the trajectory planning phase to determine departure times and route allocations, ensuring that air taxis operate smoothly, efficiently, and with minimal delays in busy urban environments.

2.4.2 In - Flight Stage Requirements

Even if the flights are planned after evaluating all the possibilities, as highlighted in the section 2.3, extensive control decision-making, continuous adaption on vehicle dynamics, and real-time interactions within the air taxi network are significant for safe operation. A functioning control and navigation system allowing both path and separation control is the key to facilitating this. The interactions between the control system and the vehicle dynamics are fundamental for updating the flight paths based on real-time conditions such as changes in traffic patterns and weather and unforeseen obstacles. That is, the control system must be able to supply accurate and timely control decisions for the air taxi path movement. An advanced control system thus allows UAM vehicles to operate autonomously, minimizing risks and enhancing reliability across densely populated, complex urban airspace.

Being fully autonomous, these air taxis depend entirely on the input data and the information exchange with other air taxis. For this reason, planning and managing the sensing and communication system and maintaining proper internal interactions with the control modules are crucial. Hence, in the sensing domain, the major emphasis is on sensor fusion and overall situational awareness handling, neighbor detection, and information exchange with command and control. For that matter, a UAM system is equipped with numerous sensing devices to gather information pertaining to the surrounding environment, including neighboring air taxis and other objects in the airspace. The ultimate goal is to provide real-time monitoring of the UAM vehicles to ensure safe air navigation. For safe navigation, the types of sensor data that are important for the UAM system that need to be accumulated through the sensor system include:

- Position Data - Latitude, longitude, and altitude
- Velocity and acceleration data
- Environmental data - Wind speed, temperature, and weather parameters

Different sensing systems have been analyzed over the years for this matter, and certain work in the literature concludes the combination of different strategies to achieve better surveillance. [NGS24] describes the requirements for such a system as :

- High accuracy

Data Type	Data Rate	Reliability	Latency
Payload Data (nC2)	Up to 50 Mbps	99.9%	ultra-low (<= 10 ms)
Non payload data (C2) - Downlink (ground station to air taxi	60 - 100 kbps	10^{-3} packet error rate	50 ms
Non payload data (C2) - Uplink (air taxi to ground station)	60 - 100 kbps	10^{-3} packet error rate	Not applicable

Table 2.2: UAV communication requirement specifications according to 3GPP standard TR 36.777 [3GP17b] presented in [WZZ20] and [NGS24]

- Enhanced coverage to include a vast airspace
- Seamless availability of data
- Low latency facilitating conditional awareness in real-time.
- Scalability to accommodate an increased number of air taxis

Furthermore, the authors provide a key performance indicator (KPI) comparison of such navigation systems. Considering aspects such as accuracy, integrity, continuity, and availability, the following navigation systems with indicated sensors or combinations have been recognized as best suited for UAM operations.

- Radar
- Inertial Navigation System (INS) - Consists of sensors including accelerometers and gyroscopes
- Global Navigation Satellite System (GNSS) [e.g.: Global Positioning System (GPS) or Galileo system] in combination with INS
- Terrestrial-based navigation - Using radio signals transmitted by terrestrial stations

For communication, on the other hand, significant aspects are, implementing communication between air taxis, communication network planning including base station placement, assessing for channel models and possible real-time interference such as shadowing and fading, overall communication system optimizations to minimize communication delays, improved system reliability and accounting for redundancy with air taxi to ground control station communication implementation. To this end, two types of communication are involved with UAM:

- Air Taxi - to - Air Taxi (A2A) Communication: The direct communication between the neighboring air taxis
- Air Taxi - to - Ground (A2G) Communication: The communication between the Air taxi and the ground control station that occurs through the infrastructure network.

These communication types will be elaborated further in the upcoming chapters. However, in terms of wireless communication applications for UAM, there are two types of data: payload data and non-payload data [ZWZ19]. In the scope of UAM, the payload data is called nC2 data and refers to the intended message transfers focusing on the application, passenger data, post-flight data, etc. This deals with data deployments in which achieving a high throughput is vital because the customers demand data rates similar to what they experience in general on the ground [EHJ+20]. The other requirements to satisfy for nC2 data links include secure communication links, location verification, operation in licensed spectrum, coverage overcoming urban obstacles to ensure reliable operations with *Received Signal Strength Indicator* (RSSI) of at least -75 dBm and scalability to handle the increase in UAM traffic, etc. [ARP20, NGS24]. On the other hand, non-payload data is referred to as C2 data in the industry and consists of data transmissions to aid with flight controls and safe navigation. In a remote piloted or unmanned aircraft situation, C2 communication needs to be reliable enough to ensure effective UAM operations [EHJ+20]. Accordingly, C2 communication requires ultra-high reliability and availability to ensure continuous control and tracking of UAVs with ultra-low latency for real-time monitoring and control [ARP20]. In comparison, depending on the purposes served by each data type, the disruption of C2 communication links is more severe than that of the nC2 communication links in the UAM domain [EHJ+20]. Tbl. 2.2 summarizes the KPIs based on 3GPP specification [3GP17b] as proposed in [WZZ20] and [NGS24].

Chapter 3

Communication System

The significance of the communication system within the context of UAM was discussed extensively in the section 2.4. As highlighted, the communication system for UAM demands numerous requirements, specifically due to real-time information transfer needs. Managing these issues can be difficult, necessitating the integration of an appropriate wireless technology into the system to be able to fully harness the UAM applications [ARP20]. For safer operations in UAM, a low latent, robust, and reliable communication would be of best interest. Hence, these aspects are to be first accounted for when planning the communication network. A specifically tailored custom communication link with such capabilities would be thus best suited for UAM. Nevertheless, due to such implementations being costly, the ideal solution would be to exploit and shape up the existing communication technologies to satisfy the requirements for safe navigation to the maximum achievable extent [EHJ+20, BDO+21].

3.1 Communication Technologies

In the context of wireless communications, a variety of communication technologies can be employed in the UAV communication paradigm depending on the applications and operation conditions. While these technologies are beneficial, certain challenges associated with some technologies hinder their direct applicability to UAV communications. In this section, diverse communication technologies, their benefits, and drawbacks in the domain of UAV communications are discussed.

For remotely piloted UAVs, the use of Institute of Electrical and Electronics Engineers (IEEE) standards is an appropriate and popular communication technology approach. Of those, WiFi technology can be commonly seen as a cost-effective, readily available solution [SWK+19]. However, only short-range communications are then possible, and it hinders the application beyond the visual line-of-sight (BVLOS) domain. ZigBee has the advantage of low power consumption, which is crucial for battery-operated systems in urban air mobility vehicles. Additionally, Zigbee's mesh networking capabilities allow devices to relay data through multiple nodes, creating a resilient communication network. However, Zigbee technology has the same challenge of limited range associated with it [BDO+21]. WiMAX can also be utilized to provide large coverage rates. Several works have been conducted concentrating on this technology and found it is suitable for agriculture and public safety applications [SI14, Rah14]. A second category is a direct link, where peer-to-peer communication is established between the UAV and the ground. This is a simple and low-cost solution. Nevertheless, this prone-to-interference system is non-scalable and has a limited range while it provides low data rates [WZZ20].

Ad-hoc networks are a subsequent possibility. Being a self-organizing and infrastructure-free network, it is adaptable, robust, and can handle high mobility scenarios favoring UAV communications. However, it is a costly network architecture with complex routing protocols and low spectrum efficiency [WZZ20]. An alternative category of communication technology suited to be applied for UAV communications is satellite communication. One non-terrestrial communication option under this category that provides a wide range of coverage is Geostationary Earth Orbit (GEO) satellites. The downside is the high latency and large signal attenuation. High Altitude Platform Communications, commonly referred to as HAPCom, is another possibility but can cause swift handovers leading to connectivity outages [BDO+21]. Moreover, the involvement of satellite communications could be costly and involve heavy and energy-consuming equipment [WZZ20].

In terms of other technologies, Ultra-wideband (UWB) technology is preferable due to its low power consumption and anti-multipath signaling, although a specific antenna design to suit UAV is needed due to electromagnetic pulse [HK17, WMM19]. LoRa or *Long Range* communication is another candidate which is rather suitable for UAV-to-UAV communications. The associated challenge is the operation in the same frequency range as air traffic management radars, which is 800 MHz, necessitating changes in allowed radar interference levels [TKB+17]. In the C-band, AeroMACS is a technology that utilizes an orthogonal frequency division multiplexing (OFDM) structure, which can work alongside satellite systems without causing interference and hence can be applied for UAV communications [ZTJ18]. However, this is primarily designed for airport ground operations and conventional aviation rather than low-altitude, dense urban air mobility scenarios. The *Automatic Dependent Surveillance Broadcast System*, commonly referred to as ADS-B, is a radio technology widely used in commercial aircraft surveillance purposes. This is applicable to UAVs. When the number of UAVs simultaneously occupying the airspace increases, the ADS-B spectrum gets saturated, and hence, for dense deployments, this system is not compatible [BDO+21].

Cellular Networks is the remaining type of communication technology suited to be applied for UAV communications. This scalable and cost-effective communication platform is proven to be of high performance [WZZ20]. Primarily, cellular networks effectively address comparable use cases found in terrestrial communications, making it appealing to be applied for the considered use case in airspace [ARP20]. Cellular networks are known to support reliable, high-speed connections essential for real-time communication needs in UAM, such as collision avoidance, vehicle tracking, etc. Moreover, they are scalable, allowing UAM communications to leverage existing infrastructure at low costs. To this end, considering the limitations and application scope of some of the above-discussed communication technologies, comparatively, the wireless cellular communication technologies can be regarded as the most rational solution for the UAM scenario [BDO+21]. However, Wi-Fi, direct link, satellite communication, etc., remain viable technologies for addressing scenarios such as remote areas lacking cellular coverage [WXZ+21] and the potential challenge associated with cellular communications, which is the interference with ground communications [WZZ20]. A full comparison of all the communication technologies is presented in Tbl. 3.1.

Over the years, the cellular communications sector evolved from generation to generation while improving its capabilities and providing efficient and reliable communication solutions. At the moment, the 6th generation (6G) communication standards are under development while 5G communication standards are being applied and evaluated in a variety of applications. The *International Telecommunication Union* (ITU), defining the standards, states that 5G communication concepts are capable of providing ultra-reliable and low latency communication with enhanced capacity [ITU17]. In comparison with the currently available cellular communication options, this capability of 5G communication makes it the most appropriate candidate for UAM communications [WXZ+21]. Even otherwise, it is accepted in general that 5G communication concepts are proven to be the most promising and cost-effective communication provider at the moment, specifically for time-sensitive applications. Therefore, in the next section, an overview of 5G communication and its capabilities alongside the evolution of cellular generation is discussed.

Technology	Advantages	Disadvantages
WiFi	Cost-effective , readily available	Limited to short-range communication
Zigbee	Low power consumption, increased reliability through mesh networking	Limited to short-range communication
WiMax	Large coverage rate	Limited applications
Direct Link	Low-cost solution	non-scalable, prone to interference, limited range, low data rates
Ad-hoc Network	Adaptable, robust, support high mobility scenarios	Costly network architecture, complex routing protocols, low spectrum efficiency
GEO Satellites	Wide range of coverage	High latency, large signal attenuation, costly, heavy energy consumption
HAPCom	Swift handovers	Connectivity outages, expensive, heavy energy consumption
UWB Technology	Low power consumption and anti-multipath signaling	Specific antenna designs are required
LoRa	Long range communications enabling UAV-UAV communications	Operation in the same frequency range as air traffic management radars
AeroMACS	Possibility to work alongside satellite communication without causing interference	Designed for high altitude operations
ADSB	An established, standardized system for real-time position reporting	Non-scalable
Cellular Communications	Scalable, cost-effective, enhanced performance	Interference with ground communications

Table 3.1: Comparison of various communication technologies in application to UAM communications

3.2 Overview on 5G Communication

Before investigating the expectations and capabilities of 5G communication, it is also important to briefly discuss the history and evolution of cellular communication. Fig.3.1 presents the evolution of cellular generations over the last 40 years and what is expected for the future. Mobile or portable devices being connected to wireless networks and establishment of communication began in the 1980s, which is referred to as the 1^{st} generation (1G) mobile networks. These were based on analog technologies, and in North America, this technology was called *Advanced Mobile Phone Systems* (AMPS). This system used frequency modulation in 30 kHz channels, and multiple access was gained through frequency division multiple access [Wei16]. Other examples of technologies developed around the world in the 1G are *A-Netz* to *C-Netz* in Germany, *Radiocom 2000* in France, *Nordic Mobile Telephone* (NMT) by Nordic countries and *Total Access Communication System* (TACS) by the United Kingdom. In the 1990s, the 2nd generation (2G) of mobile communications was introduced. In 2G, the first digital systems were invented, including voice, short message service (SMS), and data services. The primary standard in the 2G system is the global system for mobile communication (GSM) developed by the *European Telecommunications Standards Institute* (ETSI). This defines a digital, circuit-switched network specifically engineered to optimize full-duplex voice transmissions. Over time, the system broadened to include packet data transmissions via general packet radio service (GPRS) and enhanced data rates for GSM evolution, commonly known as EDGE.

Mid to end of the 20th century, the developments leading to the standardization of the next generation of communication, which is 3rd generation (3G) communication, originated in several regions of the world. This led to the formation of the *3rd Generation Partnership Project* (3GPP), which is now the responsible organization for developing technical specifications pertaining to any mobile communication standard. A ready-for-development specification published by 3GPP is termed a *Release*. 3GPP developed the 3G standardization based on an evolved GSM core network (CN) in the 2000s, and this was Release 99. From the year 2001, they came up with enhancements to the 3G system under different releases. The expectation of 3G was

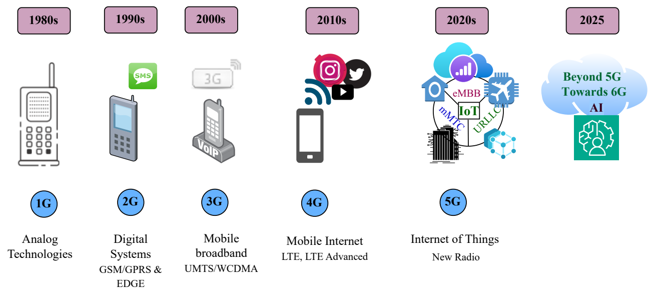


Figure 3.1: Evolution of cellular generations

to include multimedia support and high-speed packet data. The technology of 3G is commonly known as *Universal Mobile Telecommunication Service* (UMTS), which introduced new cellular capabilities such as video calling. 3G communication is based on code division multiple access (CDMA) but with a wider bandwidth and, therefore, also referred to as wideband CDMA (WCDMA).

The system kept on evolving to long-term evolution (LTE) in the 2010s, shifting to all-IP packet-switched networks utilizing OFDM. This was a major development in telecommunication history, and it introduced the concepts of scalable channel bandwidths, peak data rates up to 100 Mbps, smooth handovers, and high-quality multimedia support. LTE-Advanced is, however, the 1st technology to fall into the 4th generation (4G) of communication standards. This came into effect with the Release 10 of the 3GPP standards. New additions such as multiple input multiple outputs (MIMO), device-to-device (D2D) communication, and machine-type communications are the highlights of 4G. From Release 13, standardization of LTE-Advanced Pro began as a step towards 5G. This included vehicle-to-everything communication and full duplex MIMO (FD-MIMO) capabilities. Over the last few years, from 2020 onwards, a lot of developments have been made, and the 5G standard enabling the Internet of Things (IoT) as its main technology component was introduced with 3GPP Release 16. The first feature drop was called *New Radio* (NR). At the moment, work beyond 5G towards 6G standardization is in progress, and it is expected to have full implementation of 6G by 2030 with improved performance and new capabilities involving artificial intelligence (AI).

The standardization work for 5G already began towards the end of the last decade. Release 15 of the 3GPP technical specifications [3GP19a] defines the expectations of 5G communications. Accordingly, 5G is proposed to be addressing a wide set of new capabilities. These are enhanced mobile broadband (eMBB), ultra-reliable and low latency communication (URLLC), and massive machine type communication (mMTC). Moreover, different deployment and coverage scenarios, such as indoor, outdoor, rural, and urban, have also been taken into account. For eMBB, the intention is achieving extreme data rates of multi-Gbps and enhanced capacity leading to 10 Tbps per km^2. When it comes to mMTC, ultra-low battery life of up to around 10+ years, low complexity, and ultra-high density (around 1 million nodes per km^2) are envisioned. With URLLC, mission-critical control is addressed, and strong security, ultra-high reliability, and ultra-low latency of as low as 1 ms are the expectations. It is to be noted that these values are theoretical extremes, and practically achieving these enhanced values is not realistic due to real-time challenges such as channel disruptions and interference. However, several mechanisms to support URLLC in 5G are discussed in

[3GP22]. These mechanisms include:

- Redundant transmission, meaning duplicate user packet transmission simultaneously via disjoint user plane paths to the receiver
- Quality of service (QoS) monitoring in different granularities
- Dynamic division of packet delay budget and enhancement of session continuity where the user plane path of low latency services is optimized

With such advancements in 5G technology, a few examples of 5G use cases are :

- Telepresence
- Ultra high-definition (HD) video streaming
- Virtual reality
- Smart city and smart homes
- Object tracking
- Remote sensor systems
- Wearable devices for fitness tracking

Device-to-Device Communication

One main technological aspect supported by 5G, which is also highly relevant to the work discussed in this dissertation, is the Device-to-Device (D2D) communication capability. D2D communication refers to the possibility of direct communication establishment between the devices without the intervention of the infrastructure network. Traditionally, all cellular communications occur through the base station, and this is referred to as infrastructure-based communication. As the communication needs increase, the base station or the radio network can be overwhelmed due to the heavy load, leading to high delays and less efficient communications. As a solution, D2D communication is introduced, where the devices can directly interact between themselves depending on their proximity to each other. This helps improve the spectral efficiency and data rates and also reduces the communication delay.

D2D communication supports a wide range of applications that have highly reliable and low-latency communication needs. Examples of such applications include public safety, IoT, smart cities, etc. In terms of public safety, during emergency situations where the base station is out of service, communication can still be managed through direct D2D connections. For smart cities, D2D communication has paved the way for vehicle-to-everything communication, making intelligent traffic management possible. Further applications include smart grids and autonomous drone networks similar to the application under discussion in this dissertation. The A2A communication aspect is enabled via D2D communication technology. Edge computing and proximity-based social network services stand among the vast range of other applications. D2D communication is expected to play a critical role in future wireless communications, leading to seamless and robust connectivity. The topic of D2D communication is discussed broadly in the upcoming Sect. 7.1.1 in relation to the A2A communication aspect.

3.3 5G Network Infrastructure

As discussed in the previous section, 5G technology has been introduced, with numerous new capabilities and performance upgrades to support a variety of services in different network conditions. In order to achieve that proficiency, changes were made to the network infrastructure and the protocol stack implementations. In this section, 5G network architecture, its components, protocol stack, and radio physical layer aspects are discussed.

3.3.1 Network Architecture

[3GP19a] discuss two network architectures for 5G communication. The first deployment scenario is *Non-Standalone New Radio* (NSA NR) and refers to the architecture in which the 5G radio access network (RAN) is connected with the CN of the 4G network. In this architecture, the base station of 4G network, referred to as enodeB (eNB) is connected via the X2 logical point-to-point interface to the base station of the 5G network, referred to as gnodeB (gNB). The concept of connectivity is called dual connectivity, and this allows the user equipment (UE) to be connected to two nodes simultaneously, and the 5G connection is initiated after the 4G connection setup. This way, the eNB is the master node, while the gNB is a secondary node. This was regarded as a temporary step towards the full deployment of 5G communication. Fig.3.2 demonstrates the NSA NR network architecture. The mobility management entity/serving gateway (MME/S-GW) performs the CN functionalities and belongs to CN, which is referred to as the evolved packet core (EPC). The logical many-to-many interface between the MME/S-GW and eNB is the S1 interface, while the one between MME/S-GW and gNB is the S1-U interface. X2-U is the interface between two gNBs.

The second architecture, according to [3GP19a], is the *Standalone New Radio* (SA NR) and is depicted in Fig.3.3. The 5G system consists of the UE, RAN, and the CN. This is considered the full 5G deployment, and it is not dependent on any component from the 4G network. The RAN of SA NR is called *Next Generation Radio Access Network* (NG-RAN) while the CN is referred to as *5G core* (5GC). The connection between these two entities is maintained by the NG interface, while between the gNBs, it is the Xn interface. A subset of functionalities of NG protocol are UE context management, Packet Data Unit (PDU) session management, paging function, and mobility management. In the case of the Xn protocol, configuration data update, handover preparation and cancellation, data forwarding control, and energy saving are a few of the many functionalities.

According to [3GP19a], the specialty in 5GC is the elements being defined as network functions (NF) and seen as a set of interconnected NFs. This offers modularity and reusability. Examples of NFs are the application function (AF) handling applications, the access and mobility management function (AMF) accessing the UE and access network, and the session management function (SMF) that accesses the user plane function (UPF). UPF is responsible for

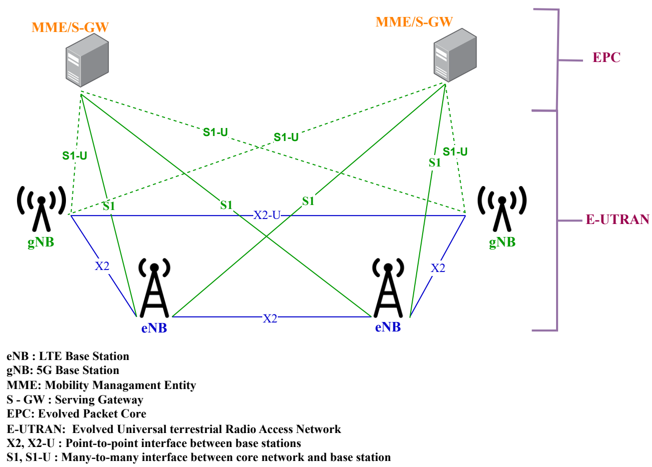


Figure 3.2: Non Stand Alone New Radio architecture based on [3GP19a]

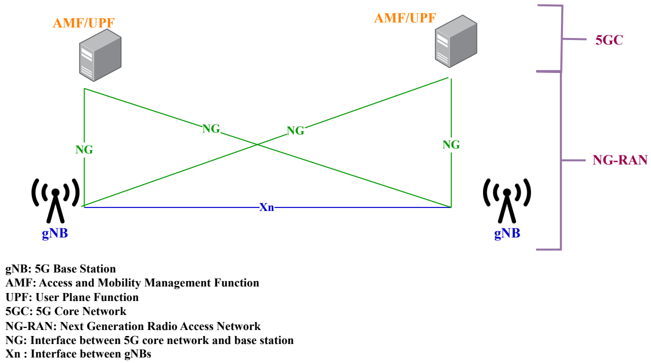


Figure 3.3: Stand Alone New Radio architecture based on [3GP19a]

handling user data, and thus, it is the entity in between the access network and the data network. The main tasks performed by the UPF include packet routing and forwarding, traffic usage reporting, interconnecting external PDU sessions to a data network, packet inspection and policy enforcement, downlink packet buffering, and QoS handling.

The RAN of the 5G system consists of only one entity, which is the gNB. Examples of functions performed by the gNB are: radio bearer control, radio admission control, scheduling of UEs, connection setup and release, session management, QoS flow management, and measurement reporting for mobility and scheduling. When it comes to the protocol stack of the RAN in 5G, it is divided into the control plane and the user plane. Each plane carries a different type of traffic and conceptually an overlay network. The user plane carries the network user traffic while the control plane carries the signaling traffic [3GP19a].

In the control plane (see Fig.3.4), the NAS stands for Non-Access Stratum control protocol. This consists of the protocols that are not relevant to the access network, such as authentication, mobility management, and control of security. This terminates at the AMF in the 5G CN. In the access network, RRC denotes radio resource control. Some of its features comprise broadcasting system information to the Access Stratum (AS), which manages the interaction between the UE and the radio interface, as well as to the NAS, which oversees higher-level functions such as mobility and session man-

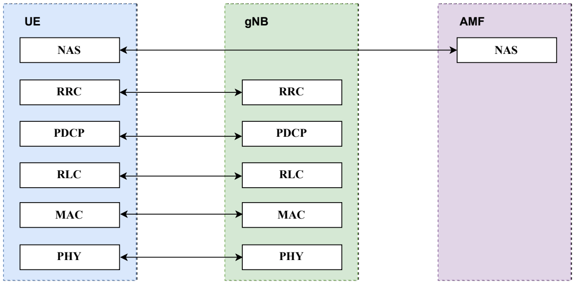


Figure 3.4: Control plane of 5G access network based on [3GP19a]

agement. It also includes paging, managing the RRC connection between the UE and the NG-RAN, and facilitating mobility functions like handover and context transfer. The PDCP, or the packet data convergence protocol, performs sequence numbering, ciphering and deciphering, control data transfer, and duplication detection, etc. The protocol layer of radio link control (RLC) is responsible for sending PDUs from upper layers, segmentation, duplicate detection, etc. One of the most important protocols, which is the medium access control (MAC) functions, consists of mapping of transport and logical channels, error correction, priority handling, etc. The role of the lowest protocol layer, which is the physical (PHY) layer, is modulation and demodulation of the signal on the radio interface [3GP19a].

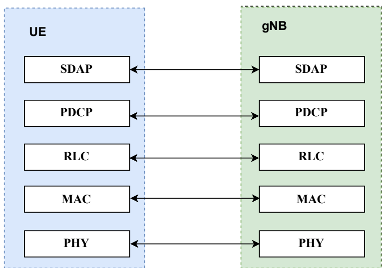


Figure 3.5: User plane of 5G access network based on [3GP19a]

In the user plane demonstrated in Fig.3.5 according to [3GP19a], there are five sublayers. As opposed to the control plane, the RRC layer is replaced by the service data adaptation protocol (SDAP) in the user plane. This layer is responsible for handling QoS flow. The functionalities of the PDCP layer in the user plane differ from that of the control plane and include sequence numbering, header compression and decompression and, user data transfers, etc. The roles of RLC, MAC, and PHY layers do not vary regardless of the control plane or user plane.

3.3.2 Physical Layer Aspects

Major changes are done in the physical layer aspects of 5G communication to support its wide range of possibilities and are discussed in [3GP19a]. OFDM with cyclic prefix (CP) similar to LTE is employed here but in both uplink and downlink directions. In addition, 5G communication can be operated in two frequency ranges. The frequency range 1 (FR1) is the sub6, meaning deployed below 6 GHz. The second frequency range (FR2) is called 5G millimeter wave (mmW), which operates in the mmW spectrum, that is, between 24.25 GHz to 52.6 GHz, providing high throughputs. For each of these frequency ranges, a specific set of bandwidths is also supported, where in FR1, it is from 5 MHz to 100 MHz, while in FR2, it is from 50 MHz to 400 MHz. In accordance with the principles of radio signal transmission, increasing frequency results in a reduced range due to free-space and tropospheric attenuation. As a result, in the FR2, it is confined to short-range and LOS communications [EHJ+20].

An emerging technology to realize this adaptability of frequencies is the flexible frame structure introduced in 5G with variable sub-carrier spacings (SCS). To this end, 15, 30, 60, 120, and 240 kHz can be used as the SCS, which can be applied using the numerology index (0, 1, 2, 3, and 4, respectively) and thus referred to as multiple numerology support. Similar to the concept of LTE, 12 subcarriers form one resource block (RB), and the channel bandwidth constitutes a number of RBs. In the time domain, radio frames are divided into 10 ms each with 10 subframes of 1 ms in each radio frame. A unit of subcarrier or one OFDM symbol is defined as a resource element (RE), and a slot constitutes 14 symbols in the case of normal CP. This slot structure, according to [3GP19a], is elaborated in Fig.3.6.

3.4 Related Work

In recent years, a growing enthusiasm towards UAV-related research has been widely observed. The attention of these research were mostly not in the scope

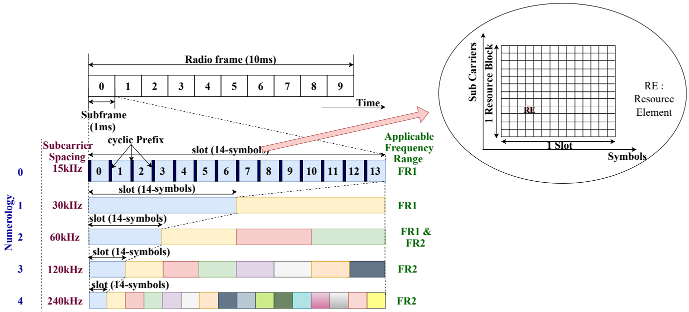


Figure 3.6: Slot structure in 5G based on [3GP19a]

of UAM as passenger transportation but drones in other applications such as package delivery and public safety. As UAV is a topic that incorporates numerous technical disciplines, the research done in the field tends to highlight multiple perspectives individually as well as an interdisciplinary approach. Among these, the field of communication has been recognized as an important aspect and studied and discussed to a greater extent. While some specifically address 5G communication due to its suitability for UAV applications given the performance expectations, in some work, multiple communication technologies are evaluated and dealt with [BDO+21, ZWZ19]. While the majority of work found in the literature is focused on UAVs, the authors in [BDO+21] present a discussion on wireless network applications for aerial vehicles, focusing on UAM as a means of passenger transportation. Without being restricted to a single communication technology, different wireless communication systems are compared. Despite that, the target here has been remote pilot-operated aerial UEs.

The field of communication in relation to UAVs can be discussed under two main divisions:

- Cellular-connected UAVs
- Cellular-assisted UAVs

The first category being *cellular-connected UAVs*, basically analyzes and identifies the possibilities and challenges of the impact of communication networks for UAV systems considering UAV as a UE [BDO+21, CCB+21, ZWZ19, GJV16]. On the other hand, there's the category of *UAV-assisted wireless communications*, where the UAV is employed as an aerial communication platform to enhance coverage and overall performance of the wireless communication network [FQD+19, WXZ+21, ZWZ19].

3.4.1 Cellular-Connected UAVs

Cellular-connected UAVs integrate UAVs as UEs into the cellular communication network. The major advantage here is the ability to manage interactions with fellow UAVs and ground controllers regardless of the range of operations. This further enables the ability to monitor and control the UAV globally from a remote pilot based on the ground. Moreover, the enhancements in cellular technologies widen the application horizon of UAVs with improved performance. In addition, cellular-based localization is also an advantage that the UAV paradigm can gain by being a part of the cellular-connected UAVs. One other main aspect is being a cost-effective solution, as the system reuses the already available communication infrastructure [ZLZ19].

The recognition of UAV as a UE was done by 3GPP in 2017 as a part of their technical specifications for Release 15. The vehicle was denoted as *aerial UE*. Through this study item, 3GPP has analyzed the possibility of employing LTE networks to serve aerial vehicles, considering them as UEs in airspace [3GP17b]. The use cases, design considerations, deployment scenarios, channel modeling, and identification of performance requirements have been the main expectations of this investigation. The applicable height of UAVs has been recognized to be from the ground to 300 m, while the traffic requirements and channel models have been based on [3GP17a], which was a part of Release 14. Since then, the studies on cellular-connected UAVs have evolved, and the initial research done in the field is found in [ARP17, ARP18, ARP19a]. A spectrum shared with the ground users has been considered, and the means to improve the efficiency of the network in the presence of UAVs is discussed. The influence of ground parameters such as base station antenna tilting and height adjustments are also studied. Cellular networks behave differently in the airspace compared to ground networks. Consequently, it is also possible for the aerial UE to connect to a distant base station due to LOS conditions [ARP17]. The downside is that, as the height increases, the aerial UEs can detect more cells, leading to high interference [Qua17].

Various technologies in accordance with LTE have also been investigated with aerial UE use cases. One example is the application of massive MIMO for UAV study in [GGRG+18]. The authors conclude the ability to establish reliable communication operations for UAVs and that the performance of ground users can be affected by this. Another example is applying D2D communication. UAV-to-UAV communication link performance is analyzed in [ARP19b]. The focus is on an uplink power spectrum, and the authors have applied distance-proportional power control. Their conclusions suggest that this power control application is able to limit the interference of UAV links on the ground network. Another exploration on uplink cellular-connected UAVs is found in [MWZ18]. The uplink transmission from the UAV to the base station is examined with a focus on minimizing the UAV's interference through the joint optimization of cell association, resource block allocation, and transmit power.

It is to be noted that in addition to academia, numerous studies in UAV communications have also been conducted in the industry. For example, Qualcomm in [Qua17] studies the highly dense UAV networks in terms of the impact of power control on uplink interference and throughput while also evaluating the signal-to-interference-plus-noise ratio (SNR) distribution. Nokia, Intel, AT & T are among the companies that initiated cellular-connected UAV testings [WXZ+21]. Nevertheless, all of these investigations on cellular-connected UAVs were carried out with a focus on the LTE. Towards the end of 2019, with the standardization of 5G, research focusing on aerial UEs in application with 5G communication studies began. One such example is [ARP20]. The authors investigate the applicability of existing 5G ground networks to serve the aerial UEs. Based on a stochastic geometry model developed, an analytical analysis is carried out on the UAV altitudes, base station density, different environment types, antenna parameters, etc. The findings further prove the suitability of existing 5G networks for UAV deployments. Moreover, mechanisms on how the antennas can be adjusted to block LOS and increase the range of operating altitude of the UAVs are suggested.

Another approach is to evaluate each of the three use cases of 5G communication alongside aerial UEs. There are associated challenges in each use case, which create the crucial need to shape the communication network to cater to such various requirements in 3D space. This includes efficient handover management to cater to URLLC scenarios, resource allocation in mMTC scenarios, and antenna re-configurations and channel adjustments to satisfy eMBB. On the other hand, there are benefits for each use case due to UE being in a 3D space as well. For example, due to the high altitude of operations, the UE is able to adjust the positions themselves to avoid signal blockage and work in the mmWave range, facilitating eMBB [WXZ+21, BDO+21].

Alongside Releases 17,18, and 19, 3GPP introduced several study items focusing on cellular-connected UAVs integrating 5G technology. In particular, attention was paid to interference management, mobility, and radio measurements. This introduced techniques such as height-specific configurations to be applied for aerial UE, flight path reporting for radio resource planning, beam forming in FR1 for uplink transmissions, and aerial UE iden-

Specification	Release No.	Title	Reference
TR 22.829	17	Enhancement for Unmanned Aerial Vehicles	[3GP19b]
TR 23.754	17	Study on supporting Unmanned Aerial Systems (UAS) connectivity, Identification and tracking	[3GP21b]
TR 23.755	17	Study on application layer support for Unmanned Aerial Systems (UAS)	[3GP21a]
TR 38.876	18	Air-to-ground network for NR	[3GP24b]
TS 22.125	19	Uncrewed Aerial System (UAS) support in 3GPP	[3GP24c]

Table 3.2: 3GPP study items focusing 5G aerial UEs

tification via the functionality of broadcasting UAV ID, etc. The specific study items, along with their focus, are listed in Tbl.3.2. Here, *TR* refers to Technical Report while *TS* refers to Technical Specification. Ever since the groundwork for 6G communications has been laid, in terms of 5G and beyond technologies, AI has gained a lot of attention. In the meantime, the applications of aerial UEs have to deal with a lot of data management and the need to carry out real-time decision-making and predictions, which is why having an AI-based communication system would be beneficial. As a result, for aerial UE communications, integrating AI has been recognized as a key driving force [GdKS+19, LCS+19].

3.4.2 Cellular-Assisted UAVs

Cellular-assisted UAVs transform communication technologies into a new dimension by enabling more efficient data transmission and connectivity. The concept involves integrating UAVs into the communication networks as an aerial communication platform where a lightweight base station or a relay is mounted onto the UAV. The UAVs incur low manufacturing costs, enabling the possibility to use them this way to effectively support the terrestrial cellular networks [ZWZ19]. This way, the limitations of the traditional LOS radio systems can be surpassed significantly, leading to enhancements in real-time communication. Furthermore, unlike fixed terrestrial base stations, these can be deployed promptly on demand and hence suited for connectivity provisions in temporary events like concerts, fares, etc., or emergency events such as rescue missions. Moreover, owing to their controllability and mobility, cellular-assisted UAVs have the potential to enhance communication system performance in terms of coverage, spectral efficiency, QoS and load balancing, etc., through dynamic repositioning of the base station [ZZL16, BCD20, SSC+17].

As a relay, a UAV can be employed to extend coverage and assist with communications. In particular, this addresses the cases of blockage of communications between a terrestrial user and base station or when the user and the ground base station are located far away from each other [WXZ+21, DdFE16]. Moreover, UAVs can be deployed as access points for data collection from ground networks as a part of sensor or IoT networks [ZWZ19]. In cellular-assisted UAV communications, the system model could be a UAV-enabled relay, UAV-enabled downlink, UAV-enabled uplink, UAV-enabled multicast, or Multi-UAV interference channel, where there's co-channel interference from others in the presence of simultaneously communicating multiple UAVs [ZZZ18b]. With regard to the usage of UAVs as an aerial base station, there are two deployment scenarios in which the UAV is used as a stationary aerial platform and also as a flying aerial platform, depending on the application. In the case of flying base stations, attention must be paid to deciding what flying speed to be applied based on the scenario that the UAV base station is providing coverage for [ZWZ19].

Several advantages of utilizing cellular-assisted UAV communications were discussed in detail in this section. Nevertheless, this technology is also involved with numerous challenges that have been investigated continuously in the research community [FQD+19, ZWZ19, ALZ+22]. For the completeness of the discussion in this section, some of these associated challenges and the proposed solutions from the literature are presented. One of the main critical issues is the positioning of UAVs, which is owing to the ability to place UAVs at variable heights. The effect of this is that, depending on the UAV base station height, the channel conditions could vary. Several optimization techniques focusing on maximizing throughput, minimizing total transmit power, minimizing the number of aerial base stations, etc., can be found in literature [MSBD16a, LZZL17] as solutions for this.

Mobility or trajectory planning is another concern associated with cellular-assisted UAV communications. Based on mobility, performance enhancement can be expected but is limited to hardware availability. As a solution, cooperative UAV networks with multiple UAVs, usage of distributed algorithms to determine mobility direction and game theory-based techniques have been proposed [ZGD+18, FDH17, FDH18, LZZ16]. Another important aspect is energy efficiency, and the proposed solutions include reducing transmission power and developing an optimal transmission scheduling [MSBD16b, ZZZ18a]. Further challenges, such as recharging aerial base stations, cost models, and front hauling and access communication links, are discussed in detail in [FQD+19]. The channel model, handover, and interference depending on the UAV density are also some topics that have gained research interest over the years [BCD20, KG20].

The inclusion of cellular-assisted UAVs into the 3GPP standardization occurred with Release 17 [3GP17b, 3GP19b]. In their specifications, with considerations for existing network conditions, the coverage up to a maximum of 300 m altitude is described [3GP17b]. Meanwhile, at the industry level, several companies like Nokia and Huawei initiated testing of aerial base stations along with LTE deployments [FQD+19]. Over the last few years, alongside 5G-related technical aspects such as massive MIMO, mmWave, and cognitive radio, new horizons in cellular-assisted UAV communications were also explored. Among many other research works done in the field, in [WXZ+21], the topic is discussed under the three use cases of 5G communications, each highlighting the issues and the research directions. From the initial investigations with LTE up to deployments in 5G networks, the UAV-assisted networks thus have evolved. At the moment, along with the standardization work in 6G and its involvement with AI, a rising interest in research is seen on cellular-assisted UAV communications towards upgrading the performance concentrating on channel modeling, UAV positioning, and trajectory optimizations [ALZ+22].

3.5 Discussion

After careful consideration of the requirements of fully autonomous UAM in application to public transportation, safe navigation was evidently the priority. To this end, the communication network was recognized as a key enabler due to system dependability on air taxi interactions. Autonomous UAM is a futuristic technology, and research is currently underway to aim for error-free, safe, and robust communication in the UAM paradigm. Although very limited work was seen aiming at UAM as a means of a passenger carrier and in the fully autonomous domain, with regard to UAV communications, both cellular-connected and cellular-assisted streams have been under research for years, focusing on various aspects. Both sectors have evolved alongside communication technologies and show potential for the next generation of communication networks involving AI in performance enhancements.

The usage of UAM for passenger transportation falls into the cellular-connected UAV category. A considerable amount of assessment on associated challenges under the topic is found in the literature. Although the focus hasn't been exactly on UAM as a passenger carrier, these previous studies provide insights into the areas of interest in the communication domain of UAM. In principle, a robust, reliable, and low-latent communication system is the criterion. Various factors that could contribute to achieving these objectives have been investigated in prior research. Despite the significant progress already made, key research areas in this scope applicable to both air-to-ground (A2G) and air-to-air (A2A) communication remain:

- Channel Modeling
- Base Station Network Planning
- Resource Allocation
- Communication Redundancy Assurance
- Interference Management

Consequently, this dissertation is dedicated to exploring these topics in relation to the application of UAM as a passenger transportation system. Both aspects related to A2A and A2G communication are investigated, and techniques and suggestions on how to enhance UAM communications are discussed in detail.

Chapter 4

Simulation of Urban Air Mobility

This chapter is intended to cover the techniques and tools employed throughout the work presented in this dissertation. First, the research approach is detailed, along with an evaluation of similar approaches found in the literature. This is followed by a discussion on the associated tools and frameworks, and finally, the performance evaluation criteria considered in relation to the upcoming chapters are elaborated.

4.1 Research Approach

As detailed throughout the Chap.2, this dissertation is centered around the topic of UAM. The specific case addressed is employing autonomous aerial vehicles in UAM as a means of public transportation. In Sect. 2.1, the developments in the autonomous driving sector were discussed, and it was clear that although autonomous driving is a concept that had been under research and testing for many years, the implementations under the full automation level 5 are not available yet. Compared to that, autonomous air taxi is a far more critical application with even higher levels of requirements on collision avoidance and traffic managing mechanisms to ensure safe navigation. This generates the need for extensive testing before the actual implementation of air taxis as a passenger carrier. Nevertheless, carrying out real-world testing scenarios of such a system would also be extremely risky for the general public. On the other hand, real-world testing incurs high costs associated with safety assurance, compliance with regulatory approvals, and complex infrastructure requirements for air space integration, as well as vehicle maintenance.

To this end, before diving into the actual implementations of air taxi systems, it is indispensable to test the air taxi paradigm with extensive simulation analyses. This way, it is possible to model real-world scenarios to test new protocols and implementations in a controlled environment for a low cost, with high scalability and reduced complexity, leading to no critical effects [SGB12, MA22, ASZS24]. Accordingly, the work presented in this dissertation is based on simulations performed with a focus on realistic conditions to suit UAM. In the scope of UAM simulations, as many different disciplines are involved, different tool chains and frameworks focusing on multiple aspects related to UAM have been implemented and tested. For example, evaluating UAM traffic systems with regard to trajectories [NBdA+22], travel times [RFBA21], and congestion [PLM+21] etc., can be found in the literature. On the other hand, UTSim [ASAM+19] is a Unity platform-based UAV simulator for integrating communication, control, sensing, and navigation. Studying different types of UAVs and environmental conditions, different communication protocols, and control algorithms are enabled through UTSim. [CCB+21] is another example of a simulator provision for sensing and communication simulations in this scope. The authors developed a simulator based on microservices, considering pilots, UAVs, ground control stations, surveillance, communication, and external environment effects such as weather conditions as main contributors to the system. Both systems are promising, but the downside is that the former is only communication- and control-oriented while the latter is only communication-oriented. Thus, neither has given any consideration to demand modeling and trajectory planning data integration.

Consequently, it is evident that, in the context of UAM in application to passenger transportation, a simulator integrating the demand and trajectory information alongside control and communication is lacking in the literature. This created the need for a simulation framework integrating all aspects of UAM. The challenge here is to find a single simulator that would support both communication and control functionalities with easy integration into a toolchain that takes demand modeling and trajectory planning into account. Modeling control algorithms is often a complex and challenging task. In general, for control-related developments, the common interactive environment used to model and simulate control algorithms is MATLAB. It provides control systems toolboxes with pre-built functions and blocks for designing, analyzing, and simulating control systems. An added advantage here is that MATLAB enables the easy integration of such developments into frameworks based on other programming languages, such as C or C++, with the in-built converter MATLAB Coder [BHW+23]. This way, the control developments don't have to be limited to MATLAB, and they can also be merged into other platforms easily.

On the other hand, communication network modeling is also a complex task as protocol stack development must be handled in relation to the specific communication technology. When communication network modeling is individually taken into account, numerous simulators are available for general simulations and customized for specific applications. Some of the most popular simulators are:

- Network Simulator version 2 (NS-2) [IH08]
- Network Simulator version 3 (NS-3) [RH10]
- Optimized Network Engineering Tools (OPNET) [Xin99]
- Network Based Environment for Modeling Simulation (NetSim) [Tet04]
- Objective Modular Network Testbed in C++ (OMNeT++) [Var01]
- Research and Analysis of Load flow (REAL) [Kes88]
- Java-based Simulation environment (J-Sim) [SCH+05]

These are all discrete event simulators where the operation of a system is represented as a sequence of discrete events that occur at specific points in time. Several surveys conducted to compare these network simulators can be found in existing research [CO15, SGB12, ASZS24]. After a thorough investigation, these previous works highlight the advantages and disadvantages of each of the network simulators. This comparison is summarized in Tbl. 4.1. Accordingly, in terms of application scope, being open source, or with regard to the involvement from the research community, some simulators are better compared to others, while each carries some disadvantages. Thus, none of the simulators can be generalized as the best simulator available. Consequently, based on the application requirements and the availability of libraries to support the development and integration of the desired framework, a simulator must be picked.

4.2 Simulation Environment

Considering similar applications to UAM, the framework *Veins* [SGD11] is an example of a model library based on OMNeT++, providing the ability to simulate communication in ground vehicle scenarios. However, this is limited to 2D environments and, therefore, cannot be directly integrated into

---

Extension of previously published "J. Berling, P. Hastedt, S.T. Wanniarachchi, A. Vieregg, C. Gertz, V. Turau, H. Werner, and V. Gollnick. A modular urban air mobility simulation toolchain with dynamic agent interaction. German Aerospace Congress 2022, Dresden, Feb 2023. doi: 10.25967/570247." [BHW+23]

Simulator	Language	Advantages	Disadvantages
NS-2	C++	Support for IP protocols, open source, well-documented	Limited scalability, outdated architecture
NS-3	C++	Python scripting support, better scalability ,open source	Some models are still under development
OPNET	C++	User-friendly graphical user interface (GUI), support variety of protocols	Commercial network simulator
NetSim	Java	User friendly	Limited scalability, commercial network simulator
OMNeT++	C++	Modular, highly flexible, GUI support, open source, a large active user community	Detailed knowledge of C++ is required
REAL	C	Light-weight, open source	Limited protocol support, less scalable
J-Sim	Java	Support sensor network simulations, open source	Limited documentation, small user community

Table 4.1: Comparison of network simulators

a UAM scenario. To this end, the need for a simulation framework for UAM simulations, that is, a framework capable of simulating large 3D scenarios, was identified. To cater to this, the development of such a framework under the UAM-based project *ULTRAS* (Urbane Lufttransportsimulation) was done in cooperation with the Institute of Control Systems (ICS) of the Hamburg University of Technology [BHW+23]. This framework was implemented with the possibility of integrating as a component in the *Remote Component Environment* (RCE) toolchain in ULTRAS [BHW+23], which is an extension of the RCE tool developed for the integrated modeling of urban air transportation systems in [DSB+20]. The ULTRAS RCE toolchain is a collaborative development by the four institutes, Institute of Air Transport Systems, ICS, Institute for Transport Planning and Logistics, and Institute of Telematics, in which each member institute was responsible for different tools in the toolchain. The basic RCE toolchain from ULTRAS [BHW+23] is depicted in Fig. 4.1. All the tools from the toolchain will be briefly discussed here. Nevertheless, it is to be noted that the key component relevant to this dissertation is the simulation tool, which is represented as tool number 10, and except for that, the other tools were developed by the other member institutes of the ULTRAS group.

Within the toolchain, first, the passenger demand between Hamburg city districts is estimated based on the city transport model. This calculation integrates existing data, approximating UAM demand as a proportion of the

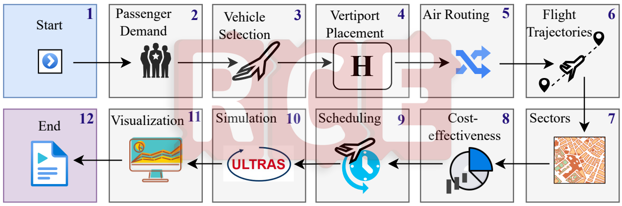


Figure 4.1: RCE tool chain from ULTRAS according to [BHW+23]

already modeled demand for cars and public transportation. This enables realistic demand forecasting aligning with existing mobility patterns and local transport behaviors. In the next stage, vehicle selection from a vehicle database sufficient to cover the considered distances is done. For vertiport placement in Hamburg, a straightforward approach is followed, where each of the city's 103 districts is assigned one vertiport. Using a geographic information software (GIS) program, these vertiports are positioned near the district centroids while avoiding restricted areas such as airports and no-fly zones. This ensures optimal placement that balances accessibility with compliance to safety and regulatory constraints [BHW+23].

In the subsequent phase, 2D routes for each origin-destination pair are generated based on a least-cost path algorithm. These predefined routes guide movements and ensure operational efficiency. In terms of nodes, each node is considered to be connected to its eight nearest neighbors. Particular considerations here are bypassing no-fly zones and maintaining conflict-free parallel paths. As a result, a total of 1,064 routes are established, providing a structured and safe navigation framework for vehicles. Then, on the designed routes, 4D trajectories are created by the trajectory planning tool. In the sector tool, the area is divided into sectors, and capacities are evaluated based on the planned trajectories. In the 8th module, which is the cost-effectiveness, the costs of UAM trajectories and vehicle operations are calculated based on their intended usage, allowing for effective economic trade-offs in scheduling decisions. This approach enables resource allocation and operational planning enhancements while balancing economic considerations. The cost of the trajectories and the cost of operating a UAM vehicle are calculated on the basis of planned use. Then, the flight scheduling is done in the scheduling module [BHW+23].

The next tool is the simulation, which is the tool handling the ULTRAS simulation framework, and it is the most relevant module for this dissertation out of all the available tools. The main objective of the ULTRAS simulation framework is to incorporate control and communication aspects that suit the application. With the integration into the toolchain, the framework also takes the trajectory and demand data into consideration. This way, the simulation environment models the entire scope of requirements in UAM. Nonetheless, the simulation framework is not dependent on the RCE toolchain, and with simple input data provision, individual simulation runs can be performed. Considering the comparisons from the literature mentioned above with regard to communication network modeling and the capability of converting control algorithms facilitating easy integration, the framework was decided to be implemented based on OMNeT++ discrete event simulator. The specific reasoning for the decision is elaborated in the next subsection. For the majority of the work presented in this dissertation, the ULTRAS simulation framework was used as the simulation environment. The final tool, which is the visualization, is intended to provide the evaluation results corresponding to travel time differences, energy consumption, etc. [BHW+23].

4.2.1 OMNeT++

OMNeT++ stands for *Objective Modular Network Testbed in C++*, which is an open-source simulator based on C++ programming language. Rather than classifying as a network simulator, this is considered a generic simulation framework for modeling and testing complex distributed systems [AVK20]. OMNeT++ enables large-scale simulations that can possibly be built with reusable pre-built modules. This flexible, modular simulation framework offers an *Integrated Development Environment* (IDE) that allows users to efficiently create and customize simulation models. The simulations can be run both in a graphical user interface (GUI) or in a command line interface. With the GUI, users can animate and visualize the simulation in 2D or 3D and explore the underlying condition of the model. This allows users to efficiently debug and understand the simulation. The simulation also provides an event log, ultimately making the events easily traceable and saving time for the user. The simulations can be run for a specified time interval or event-wise under different simulation speeds [Var01]. One very important feature related to OMNeT++ is the active research community. This community engagement has enabled the maintenance of OMNeT++ to stay up-to-date alongside emerging technologies, delivering a flexible simulation platform to test complex network scenarios effectively.

OMNeT++ is a structured framework with a cascade of modules starting from the network module at the topmost level. A network module consists of several sub-modules that are defined as compound or simple modules, which are both regarded as instances of the OMNeT modules. The module structure in OMNeT++ is depicted in Fig. 4.2. The simple and compound module definitions and network definitions are done using the OMNeT++ network description language, commonly referred to as NED. A compound module consists of a group of simple modules. A simple module is the lowest in the hierarchy and is written in C++. The modules have parameters through which these are customizable. The simple modules communicate with each other by passing messages via the gates or by directly sending the messages to the destination module. These messages can contain arbitrary data corresponding to packets in a network. Messages always originate and end in simple modules connected via numerous simple and compound module connections [Var01].

OMNeT++ has been the base for numerous model frameworks and simulation models pertaining to queuing, internet protocols, wireless networks, peer-to-peer networks, etc. Out of those, one of the main model frameworks in existence is INET. For communication network designs, the INET module provides network components such as routers, network devices, protocol layers, background maps, etc. [AVK20]. Accordingly, one important feature in OMNeT++ is the availability of such model libraries that can be easily integrated into new framework implementations for different applications without having to design the network infrastructures from the baseline. To this end, the framework Simu5G [NSTV20] is an example where a module library for 5G communication component modeling has been provided.

4.2.2 Simu5G

Simu5G is a model library built on OMNeT++ designed explicitly for simulating 5G networks. It consists of all the protocol layers pertaining to

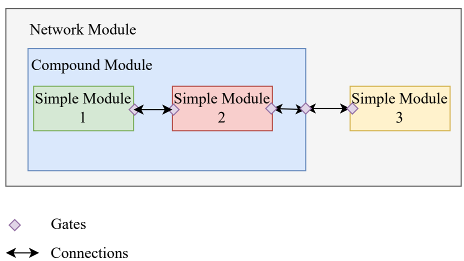


Figure 4.2: Module Structure in OMNeT++

the user plane and the core network of the 5G architecture discussed in Sect. 3.3.1. Control plane modeling is not included in the current implementations. Simu5G is a tool designed to assess the performance of 5G communications in various 5G deployment scenarios and configurations within an end-to-end framework incorporating advanced TCP/IP networking within systems utilizing 5G NR layer-2 interfaces [NSTV20]. In protocol layer design, Simu5G has made use of the network components from INET, as elaborated in the previous section.

Modeling of 5G communication in both frequency division duplexing (FDD) and time division duplexing (TDD) is possible with Simu5G. Handover, D2D communication extended to cellular vehicle-to-everything (C-V2X) standard, simulating inter-cell interference and dual connectivity are some of the simulating possibilities that are enabled. Support for different UE mobility scenarios is another key feature of Simu5G. Following the 3GPP standardization Release 16 [3GP22] dedicated to 5G communication, Simu5G has included realistic channel condition modeling, including shadowing and fading conditions. The 5G deployment environments, such as Urban Macro, Urban Micro, Rural Area, etc., can all be simulated in Simu5G to suit the respective applications. Resource scheduling, carrier aggregation, and multiple numerologies are a few other 5G-related technologies that are features of Simu5G [NSTV20].

The key modules of Simu5G, according to [NSTV20], are as depicted in Fig.4.3. The two main compound modules are the UE and gNB. In a simulation, an unlimited number of gNB and UE instances can be utilized. The communications among the UE and the gNB take place via the layer two components of the Open Systems Interconnection (OSI) protocol reference model. In 5G, both the layer 1 and 2 functionalities are narrowed down

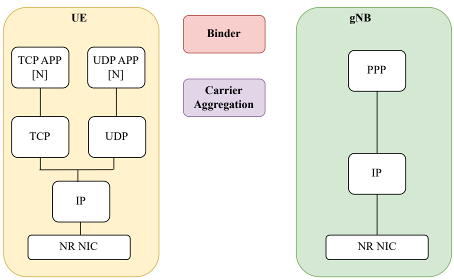


Figure 4.3: Key modules of Simu5G according to [NSTV20]

into a protocol stack consisting of 4 layers, PDCP, RLC, MAC, and PHY, as described earlier in 3.3.1. This 5G protocol stack is implemented inside the compound module NR NIC, which stands for *New Radio Network Interface Card* in simu5G for both gNB and UE. The UE consists of all the layers, from the application layer to the physical layer, along with IP and TCP/UDP layers. The NR features are embedded in the *NrNicUe*, which is the specific NIC for the UE. In the case of the gNB, protocol layers up to the IP layer are included. From the IP layer, there are two interfaces, notably, the point-to-point (PPP) interface for the CN wired connection and the NR interface to the protocol stack inside NR NIC called the *NrNicGnb*. To allow dual connectivity, in simu5G, both LTE and NR protocol stacks are included inside the NrNicUe module.

In addition to that, there is a *Binder* module that stores global data such as frequencies used by gNBs, UE multicast groups, etc. This enables easy access to this information for system functions and abstracting the control plane functionalities. The module *carrier aggregation* handles all the information related to carrier components (CC) in the network, which are the parts of frequency characterized by a carrier frequency, numerology, available number of RBs, and whether the communication is FDD or TDD along with the slot format for TDD. In NR, simultaneous communication through multiple CCs is possible, and this functionality is enabled through the carrier aggregation module. The idea is that communication is only permitted through the CCs supported by both the UE and gNB. Both these modules, Binder and Carrier Aggregation, are only available as a single copy in Simu5G [NSTV20].

Considering the implementation details, the PDCP layer of the NR stack is developed in the *NrPdcp* module in Simu5G. Header compression and the assigning of the connection identifier are the main functionalities performed by the PDCP protocol. To enable this in Simu5G, once an IP packet has arrived, the connection identifier referred to as *Logical Connection Identifier* (LCID) consisting of the source and destination addresses and ports is attached to the packet. Then, a PDCP PDU is generated and sent to the next layer, which is RLC. In the case of reception from RLC, decapsulation, and transfer of the packet to the upper layer are handled by the NrPdcp. In terms of RLC layer development, *NrRlc* module is capable of modeling all three modes of RLC: transparent mode, unacknowledged mode, and acknowledged mode. In the transmission process, RLC PDUs are temporarily stored and fetched by the MAC layer during each transmission time interval (TTI) as part of its scheduling process. On the reception side, incoming PDUs are held in buffers until all segments of a PDCP PDU are reassembled, after which they are forwarded to the PDCP layer. This ensures orderly data flow and efficient handling of protocol operations [NSTV20].

The MAC layer in Simu5g is regulated to repeat every TTI. In MAC scheduling, a set of backlogged UEs are taken into account, and one CC at a time is allocated. Different types of schedulers can be enabled, such as Maximum Carrier to Interference (Max C/I) and preferential fair (PF) to allocate resource blocks (RB) to the UEs. After each CC is scheduled, a global scheduling list is obtained. For each element in the list, a MAC transport Block (TB) is built by the MAC layer in the downlink case or issues a scheduling grant in the UL case. The hybrid automatic repeat request (H-ARQ) procedure is also supported in Simu5G. The lowest layer, the PHY layer, is modeled in such a way as to simulate the impact of signal transmission and interference on MAC TB. The model works by embedding each MAC TB within an *Air Frame*, which is a message in OMNeT++ that is transferred to the destination module. At the destination, a channel model is applied to calculate the received power and the signal-to-interference plus noise ratio (SINR) on each TB via a specific function defined in these modules. Then, the Block Error Rate (BLER) curves from 3GPP documents or a link simulator are applied to estimate the probability with which the data is received on each RB belonging to the current transmission [NSTV20].

In terms of the core network, the UPF implementation is provided in simu5G. This facilitates IP packet routing between the data network and the gNB. As depicted in Fig. 4.4, Simu5G models this in two modes:

- Standalone mode - 5G only deployment
- Dual Connectivity mode - 5G and 4G coexistence, where eNB acts as the master node while gNB is the secondary node. LTE traffic is routed through eNB, while 5G traffic is routed through eNB first and reaches gNB via the X2 interface

4.2.3 ULTRAS Simulation Framework

ULTRAS simulation framework is built on OMNeT++, considering its modular, extensible, and component-based nature. One advantage of using OMNeT++ is the ability to use the Simu5G model library for 5G communication protocol development. On the other hand, since the simulator is built on the C++ programming language, integrating the air taxi control and dynamic modeling from MATLAB was straightforward through code conversions. For the implementation of the framework, the Hamburg metropolitan area was considered as a use case. The ULTRAS framework takes trajectory and settings in the format of extensible markup language (XML) files as the input. The ULTRAS framework outputs the simulated trajectories and other control

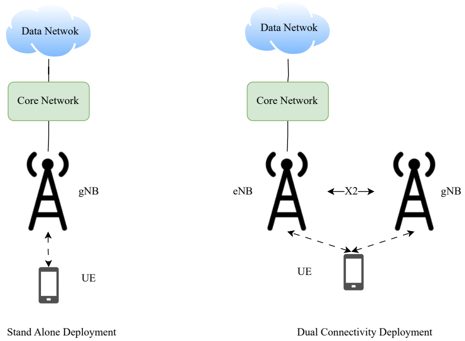


Figure 4.4: The stand-alone and dual connectivity deployment scenarios according to [NSTV20]

and communication-related data in the comma-separated value (CSV) format. To this end, these results can be easily processed with standard Python libraries to generate the output in XML format [BHW+23].

Air Taxi Network

In the simulation environment of the air taxi framework, the major components are the loader module, flight schedule manager, ground station, and

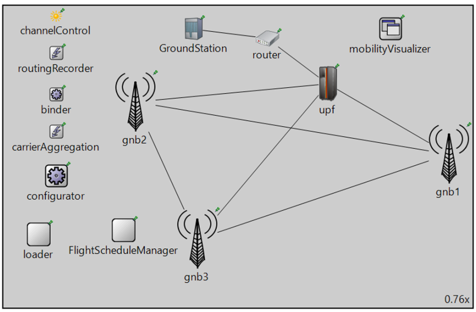


Figure 4.5: Module architecture of the ULTRAS simulation framework

the modules supporting visualization. The loader module is responsible for reading the XML files and converting them to C++ understandable format. The flight schedule manager spawns the flights into the simulation environment following the planned flight schedule. In general, the framework was developed, including ideal conditions with the possibility of extension into complex communication protocols and control algorithms. Hence, in the basic framework, the communication was implemented as a direct link. Later, for the 5G modeling, the 5G network infrastructure from Simu5G was integrated into the ULTRAS framework. The module architecture comprising of 5G components from OMNeT simulation environment is depicted in Fig.4.5.

The ground station models the ground control station for redundancy purposes as elaborated under the system architecture in Sect. 2.3. It is a compound module and consists of a sensor system, a communication module, and modules supporting visualization. The idea is for air taxis to regularly communicate with the ground station, and based on the information received, the ground station will provide global awareness for the air taxi network. Accordingly, air taxi interactions and monitoring are supported by the ground station. To facilitate this, communication to and from the air taxi is enabled. The communication module in the ground station can be extended to include any communication protocol. As per the sensor system, a ground radar can be included, and the data from the radar and the information received from the air taxis can be fused together to create a global situation awareness map that can be communicated to the air taxi network. This further includes the possibility of providing weather and other emergency information that would be relevant for air taxis. Fig.4.6 demonstrates the module architecture inside the ground station from the ULTRAS simulation environment.

Air Taxi

The air taxi is a compound module in the network and is designed as a mobile node in OMNeT++. The module architecture inside the air taxi is demonstrated in Fig.4.7. For the mobility of the air taxi, a new mobility model extending the mobility base of INET was developed so that the air taxi follows the planned trajectory. The air taxi module consists of a dynamics module, a guidance and navigation (GNC) module, a sensor system, and a communication module. Dynamics and GNC are navigation and control-specific modules implemented by the ICS. The Dynamics module simulates vehicle behavior to suit four different air taxi models, namely, without dynamics, simple kinematics, simple dynamics, and non-linear dynamics models. For vehicle control, a cascaded controller structure with an inner path controller and an outer separation controller is employed [BHW+23].

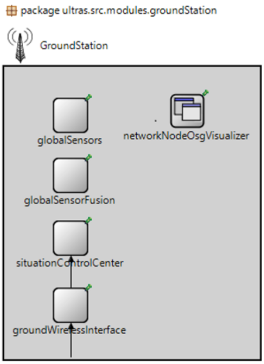


Figure 4.6: Module architecture inside the ground station from the ULTRAS simulation framework

The sensor system consists of three main modules, namely: sensors, sensor fusion and local situation picture. The sensor module models the behavior of sensing the surroundings for collision detection purposes. This is modeled including the GPS module, which can be extended into other sensors such

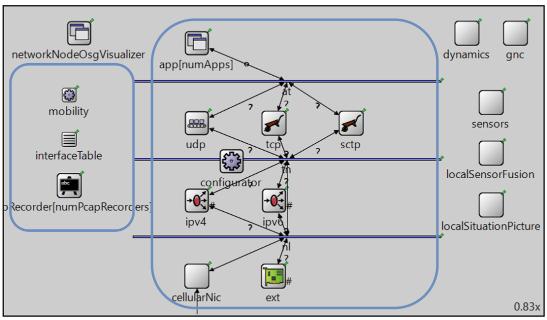


Figure 4.7: Module architecture inside the air taxi from the ULTRAS simulation framework

as lidar, camera, etc. Sensor fusion is used when several sensor types are available to fuse the data and create a single output. This output is made available to the air taxi control and navigation systems via the local situation picture module. This module basically means the background map of the environment, which is local to the air taxi. Each air taxi creates its local situation picture. This consists of a combination of data from the sensor system and the extracted details about the air taxi obtained from the ground station, providing a global overview of the air taxi network. The local situation picture of the air taxis is shared among each other and is also used as a module that provides information for control decision-making.

The communication module was developed with the possibility of extending it to support any communication technology as required. For the experiments with regard to this dissertation, the case of integrating the 5G protocol stack using the Simu5G model library was considered. In Fig.4.7, the 5G protocol stack and other 5G-related components from Simu5G are highlighted. Modeling of both A2A communication for air taxi interactions and A2G communication for air taxi-to-ground redundant communication is possible with the ULTRAS simulation framework. This ensures air taxi information management for collision avoidance and maneuvering safety. In the A2A communication, the D2D communication capability of 5G technology has been utilized. Accordingly, the air taxis in close proximity are identified, and they can communicate with each other without the intervention of the base station. The information that is communicated in both communication types is the position and velocity data and the local situation picture consisting of the own neighbor data. Both information flows happen in a periodic manner. Internal connections exist between communication and the GNC module via the local situation picture module such that the neighbor data can be updated at the GNC via OMNeT messages.

4.3 Performance Evaluation Metrics

In this section, the performance evaluation criteria that are being used throughout this dissertation are introduced.

Delivery Ratio

It is the ratio between the total number of data packets successfully delivered at the destination to the total number of data packets sent by the sender and given by Eqn. 4.1. This is a metric that indicates the reliability of the system by providing a measure of how much of the data transmitted by the sender is actually delivered at the destination. In UAM communications, this is a vital measure, as it is necessary to ensure seamless and reliable data exchange between air taxis and ground infrastructure. Being fully autonomous, the control and navigation decisions are aided by maintaining real-time coordination between the entities, and therefore, high delivery rates are essential for a UAM system.

$$DeliveryRatio = \frac{NumberofReceivedMessages}{NumberofTransmittedMessages} \tag{4.1}$$

Delivery Delay

This metric relates to the time duration required by a data packet for a traverse from the sender to the receiver in a communication network. This is often measured in milliseconds (ms) and is commonly referred to as latency. In the context of UAM communications, low latency is highly important to guarantee real-time data exchange between the air taxis and the ground control stations. Autonomous air taxis make navigation decisions based on real-time information about the surrounding environment. Therefore, timely information availability is vital for UAM communication to avoid critical aftereffects and to ensure safe and smooth operations in urban airspace.

Throughput

A measure of how much data is successfully transmitted over a communication channel or network in a given period of time. It is generally expressed in bits per second (bps). The throughput calculation is given by the Shannon capacity theorem expressed in Eqn. 4.2, where B is the bandwidth, and SNR is the signal-to-noise ratio. For UAM communications, maintaining high throughput is critical to support reliable data exchange between air taxis and ground stations. This metric ensures communication links can handle the data demands required for efficient and safe air traffic management.

$$Throughput = B\log_2(1 + SNR) \tag{4.2}$$

Signal to Noise Ratio

The signal-to-noise ratio, commonly referred to as SNR, is defined as the ratio of the power of a signal (the proper or desired information) to the power of the noise (unwanted interference or background noise). SNR is expressed in decibels (dB) and given by Eqn. 4.3. N_o is the additive white Gaussian noise, and $\sum I$ is the summation of total interference at the receiving end. In communication systems, maintaining a high SNR reflects less interference, enabling reliable and efficient communication, leading to low latency. This is highly important for ensuring safe UAM operations.

$$SNR = \frac{TransmittedPower}{(N_o + \sum I)} \tag{4.3}$$

4.4 Discussion

While extensive studies on UAM systems are evidently significant before starting the implementations due to safety concerns, the real-world testing of such systems would also be risky, costly, doesn't accommodate repeatability, and is less scalable. To this end, simulations are a viable solution, providing the user the capability to model the expected real-world scenario in a rather controlled environment, allowing repeatability at a low cost, and allowing the testing of the worst-case scenarios without leading to critical effects. However, in modeling such a simulator, it is important to take all aspects concerning the system of interest into account. With the focus being UAM as a means of passenger carrier, demand modeling is of high prominence, and the simulator must be able to acknowledge this data. Moreover, the inclusion of planned trajectories, navigation and control of the air taxi, communication within as well as interactions between air taxis, and redundancy handling must be the major components of the simulation environment to ensure a safe and reliable UAM system.

Various simulators to cater to UAM focusing on UAVs have already been studied and implemented. Nevertheless, these simulators are built with the aim of one or several factors of interest but not all contributing aspects. To this end, introduced in this chapter is a simulation framework developed on OMNeT++ discrete event simulator as a component of an overall toolchain for modeling UAM systems. This simulation framework, referred to as the ULTRAS simulation framework, is not dependable on the toolchain and can work independently with trajectory input data. This way, it models the control and communication integration while taking trajectory planning and demand modeling data as inputs. As a use case, the Hamburg metropolitan area was taken into account. After a detailed overview of the available simulators, the ULTRAS simulation framework was discussed in detail in this chapter, and explanations of each of the modules involved were given. In addition to that, the performance evaluation criteria that will be used for evaluating the topics of interest in this dissertation were also presented in this chapter.

Chapter 5

Urban Air Mobility Communication Network Planning

The significance and impact of a seamless communication system within a UAM environment were emphasized in Chap. 2. The operating environments in relation to UAM introduce distinct technical challenges for communication, leading to a variety of open research questions as highlighted in Sect. 3.5. All these individual queries ultimately contribute to realizing the main challenge of establishing a robust, low-latent, and reliable communication system. Considering a time-sensitive application similar to UAM, both Air-to-Ground (A2G) and Air-to-Air (A2A) communication capabilities need to be established and optimized to cater to the requirement of uninterrupted availability of communication services. After a thorough analysis of the available communication technologies as presented in Chap. 3, it was evident that the most suitable candidate for ensuring the communication needs of such an environment is 5G communication technology. It is apparent that even to make the best of applying the 5G communication paradigm, one of the key concerns is the ground network planning [SSH20, ZRPH+21]. The number of base stations to be placed, the positioning, and the placement topology are some of the prominent factors that have an influence on communication performance, specifically focusing on the network layer of the protocol stack. Ideally, a dense deployment of base stations to satisfy the required demand would be the solution. Nevertheless, practically, this would be costly in terms of implementation and maintenance and energy inefficient, making it a less viable option for achieving efficient and sustainable communication systems.

Significant extension of previously published "S. T. Wanniarachchi and V. Turau, "A Study on the Influence of 5G Network planning on communication in Urban Air Mobility," 2023 IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), Boston, MA, USA, 2023, pp. 394-399, doi: 10.1109/WoWMoM57956.2023.00070." [WT23b]

Thus, exploring means to achieve a ground base station network with a low number of base stations while still meeting the required level of performance is significant.

Another important feature in terms of communication network establishment for a UAM scenario is setting up a realistic A2G communication link tailored to the specific application requirements. Being a propagation characteristic, the focus here is on the physical layer of the communication protocol stack. Ensuring the availability of an appropriate A2G propagation model is essential for handling reliable resource allocation, modulation, waveforms, etc. [KGM+20]. In general, the communication technologies are introduced aiming at ground users and hence, the channel model designs are confined to considering the configurations applicable for terrestrial user equipment (UE) scenarios. On the other hand, the channel models designed and utilized for conventional aircraft communications primarily focus on high-altitude operations. Due to this difference in altitude, the factors influencing signal propagation from the transmitter to the receiver, such as reflections, diffraction, and scattering, differ significantly when compared to low-altitude operations [KGM+20]. Thus, neither of these channel conditions can be directly integrated into the UAM circumstances that occupy the low-altitude airspace. This requires studying the contributing factors and defining a channel model suitable for the UAM communication system.

Contributions

Considering above, this chapter primarily focuses on the topic of communication network planning. First, the topic of channel modeling is elaborated, presenting the work already done in this domain and 3GPP standardization on 5G, and proceeds to introduce an A2G channel model tailored to suit the specific application. Next, the ground network infrastructure establishment is discussed in detail. Starting with an analysis of the previous studies from the literature, the section extends into presenting an approach for ground communication network design. This is followed by an evaluation focusing on the Hamburg metropolitan region as an application, proposing a base station network with the number of base stations and the placement. By performing simulations for several use cases under the considered application, the proposed network architecture is verified. In terms of simulations, the ULTRAS simulation framework introduced in Sect. 4.2 is used as the simulation environment. To summarize, the contributions of the chapter are as follows.

- Challenge Addressed: Establishing a robust, low-latent, and reliable communication system for UAM
- Approach: Efficient and effective communication network planning
- Solution:
  1. A2G channel Modeling
  2. Developing a base station (gNodeB (gNB) in 5G) network design

5.1 Channel Modeling for UAM communications

The communication emphasized here is A2G, specifically between the aerial UE and the ground station, making the A2G channel the primary focus of discussion. The channel conditions vary highly in terms of the applicable environment. For that reason, a channel model designed for terrestrial communication or conventional air traffic scenarios is not suitable for application in UAM communications. In UAM, air taxis occupy low-altitude airspace in urban environments. Therefore, the aircraft is in close proximity to obstacles, causing signal blockages leading to attenuation [EHJ+20]. A specifically tailored channel model must thus be designed to suit UAM and reflect particular conditions, such as path loss constraints. In this section, these aspects are examined in detail.

5.1.1 Related Work

The propagation channels for UAM or UAV to ground communication is not a topic that has been comprehensively studied as compared to ground communication channels. However, with the popularity of UAM over the last few years, the study of A2G channel modeling has also attracted significant research interest. In general, A2G channel modeling has been studied since manually operated radiotelegraphs in the early 1900s [KGM+20]. Starting from the usage of lower frequency bands, this has evolved into military and civilian aeronautical A2G communication that happens in the L-band (belonging to Ultra High Frequency (UHF) range) over the last few decades. In addition to that, A2G channels for satellite and high-altitude platform communications have also been extensively studied in the literature [KGM+20].

Several surveys conducted in the field of A2G channel modeling can be found in previous research. [KGM+20] is a recent study that examines the latest channel measurement campaigns and modeling initiatives designed to characterize A2G channels for UAVs. The research also addresses the challenges in this area and suggests potential improvements concerning UAV propagation channels. [KCZ+18] is another example survey done in a similar scope. The authors discuss both empirical and analytical channel models and conclude that empirical studies with measurements are necessary for channel modeling as analytical modeling may not always describe the real channel conditions. Two other rather old surveys done on A2G communication channel studies are [Mat12, MS15]. Alongside these surveys, there is a range of other studies that focus on various aspects of A2G channel modeling.

In the context of UAV and UAM A2G channels, existing studies are generally done under two main categories. The *payload communications* category focuses on the transmission of data and information from the UAV to ground stations, which includes video feeds, sensor data, and other payload-related communications. The second category, *Control and Non-Payload Communications* (CNPC), encompasses the communication necessary for the operation and control of UAVs, including command and control (C2) signals and telemetry data. In [Mat12], CNPC A2G communications are studied, and L-band and C-band are suggested as the potential bands to be considered. A later study in 2017, also focused on CNPC, is [MS17], and here, the channel characterization is studied with the aim of generating an empirical A2G channel model.

One of the prominent topics to be discussed under channel modeling is channel characteristics. Typically, path loss, shadowing, fading, delay spread, and Doppler frequency spread are studied as channel characteristics. Path loss is a widely addressed topic in the literature. Path loss is the phenomenon of the reduction in signal strength as it travels through the channel from the transmitter to the receiver. This loss can occur due to various factors, such as distance, obstacles like buildings or trees, and atmospheric conditions. In simple terms, the farther a signal has to travel or the more obstacles it encounters, the weaker it becomes when it reaches its destination. Several methods for path loss estimation have been extensively discussed in the literature, ranging from ray tracing to proposing machine learning-based techniques. For A2G channels, path loss modeling is initiated from the free space path loss model (FSPL) in decibels based on Friss Transmission Formula in free space and given by Eqn. 5.1

$$\text{FSPL} = 20\log_{10}(d) + 20\log_{10}(f) + 20\log_{10}\left(\frac{4\pi}{c}\right) \tag{5.1}$$

Where:

- FSPL is the free space path loss (in dB)
- d is the distance between the transmitter and receiver (in m).

- f is the frequency of the signal (in Hz).
- c is the speed of light in vacuum ($c \approx 3 \times 10^8$ m/s).

For the purpose of UAM communication channel modeling, the FSPL model is extended based on the log-distance path loss model given by Eqn. 5.2. This is due to the significance of the inclusion of variations in signal paths that are influenced greatly by altitude and environmental conditions. The equation incorporates the path loss exponent (PLE) denoted by n to represent the increase in path loss as the altitude grows [AT19]. Typically, for free space, n is 2, and for urban areas, it is considered to be varying between 3-4. However, because this value varies depending on the environment, this estimation of PLE has been recognized as one major component contributing to channel modeling in UAM communications and has been studied under several previous works in the field. In [WZCL21], the Bayesian filtering-based method is proposed as a dynamic PLE estimation approach. In [AT19], for different UAV heights, PLE calculations for both line of sight (LOS) and non line of sight (NLOS) conditions have been done based on ray tracing simulations. In general, path loss is defined separately for LOS and NLOS conditions, as each scenario exhibits distinct propagation characteristics and impacts signal strength differently. Not only the altitude and environment but also the orientation of the UAV antennas can affect the PLE values [KGM+20].

$$\mathrm{PL}(d) = \mathrm{PL}(d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) \tag{5.2}$$

Where:

- $\mathrm{PL}(d)$ is the path loss at distance d.
- $\mathrm{PL}(d_0)$ is the path loss at a reference distance d_0 (typically 1 meter).
- n is the path loss exponent

Path loss can also be studied in association with shadowing [KGM+20]. Shadowing is the slow variation in received signal strength over a large distance caused by obstacles in the environment, which obstruct and reflect radio waves as they propagate from the transmitter to the receiver. This is also called large-scale fading, and variations in path loss based on different environments have been extensively studied in terrestrial networks for urban macro and micro-cellular scenarios [ZLT+15, SRR+16]. Similarly, numerous studies on the A2G communication channel are available in the literature. In [QMTN06], a statistical model focusing on sub-6 bands has been introduced for dense urban environments. Their findings suggest that representing the channel in terms of height and elevation angle is rather suitable for A2G channel modeling. The LOS probability (the likelihood that a direct, unobstructed path exists between a transmitter and a receiver) for A2G channels has been recognized as high, and shadowing has been identified to be less compared to terrestrial channels as mostly the shadowing in A2G channels is occurring from buildings in the environment. An analytical model for A2G communication channels under similar environmental and frequency conditions has been proposed in recent work by [SNH20]. The study is conducted based on model terms derived through a ray tracing simulation analysis. Further, this work has highlighted the significance of frequency and UAV height on the path loss and LOS probability.

Among other work done in the field, [CGBR+21] discusses both channel characteristics for A2A and A2G communications. The authors address the channel modeling under the three categories of UAV-to-UAV, UAV-to-base station, and UAV-to-UE. The key challenges in terms of channel modeling have been recognized as the variable heights, mobility of the UAV, rain attenuation, frequencies, and environments. The study is divided into two parts: large-scale statistics covering path loss and shadow fading and small-scale statistics, such as delay spread. With an analysis based on case studies done for different frequency ranges and environmental conditions, the machine learning technique of artificial neural networks is proposed as a suitable methodology for path loss predictions. [CRPY+19] is a rather new study focused on LTE communications. The authors proposed an empirical A2G channel model based on LTE channel measurements and extracting channel impulse response. Measurements are done for different UAV heights in a sub-urban scenario. The channel characteristics of fading, Doppler effect, delay spread, and shadowing are examined in addition to the path loss. The channel is recognized to be more LOS at higher heights with less multi-path propagation characteristics. [ZYB+21] proposes a general map-based deterministic A2G path loss prediction method using a ray tracing technique focusing on mmWave communications. Under the three propagation conditions of LOS, diffraction, and reflection, the analysis has been done for various environments, frequencies, and UAV height conditions. The method is proven to be applicable for both low and high-altitude conditions.

Alongside the work in academia, 3GPP has also paid attention to developing A2G channel models. This is done in conjunction with the 3GPP Release 15, which extended the work from the Release 14 study item, TR 38901 version 14 [3GP17a]. In TR 38901 version 14, the channel model focusing on 0.5 GHz to 100 GHz frequencies is explored. Rural Macro, Urban Macro, Urban Micro-street, and in-office environmental conditions are taken into account for a maximum UE height of 22.5 m. The path loss models are initiated from

FSPL models. Both LOS and NLOS conditions with log-normal shadow fading distribution are examined. This study item outputs the LOS probability and path loss model equations. With the study item TR 36.777 [3GP17b] from Release 15, 3GPP examines the channel models for aerial communications. This is an extension from Release 14 to include higher altitude UEs derived from the channel model definitions of TR 38901 version 14. The focus is LTE, and the deployment scenarios of urban macro, urban micro, and rural macro have been studied. The definitions included in the study have focused on LOS probability, path loss models, and small-scale fading models [KGM+20]. The distance and height definitions pertaining to channel model developments in 3GPP are elaborated in Fig. 5.1. The 2D and 3D distances between the UE and the base station (BS) are d_{2D} and d_{3D}, respectively. In addition to that, the h_{UE} stands for UE height from the ground level, whereas h_{BS} refers to the base station height. It is to be noted that in TR 36.777, the definitions are done in such a way that the UE heights can be adjusted to be both higher or lower than the base station height. Accordingly, a UAV with a maximum height of 300 m and a maximum horizontal speed of 160 km/h is the target scenario. The standard highlights the increment in LOS probability as the altitude rises.

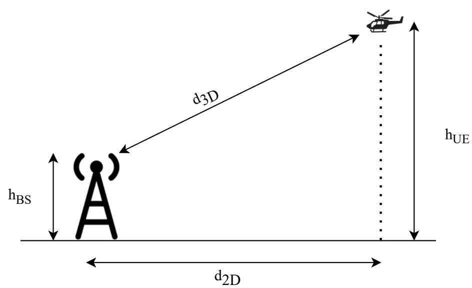


Figure 5.1: Distance definitions for channel model development according to [3GP17b]

5.1.2 A2G channel model

The A2G channel modeling approach applied for communication between the air taxi and the ground station is discussed in this subsection. Hamburg metropolitan area is the use case. As elaborated in Sect. 3.1, 5G communication is employed. Accordingly, the implementation of the A2G channel model for this specific case should follow the principles of 5G communication channel modeling, considering the conditions of an urban environment. In 3GPP standardization for 5G communications, the channel models for terrestrial networks are reused and modified in later releases. The latest version available is version 18 of TR 38.901 under the 3GPP Release 18 [3GP24a]. Thus, the 5G specific channel models are the same as compared to TR 38901 version 14 [3GP17a] and developed for a UE height range up to 22.5 m.

As elaborated in Sect. 4.2.3, simu5G is used in 5G communication protocol design for the purpose of simulations in this dissertation. In the existing simu5G library, the channel modeling is also done based on TR 38.901 and is limited to ground network scenarios. The UE height consideration for the urban macro scenario is only up to 22.5 m. In the desired application, the UE height is variable and can go up to a height of around 600 m. Consequently, the channel model needs to have path loss models for such heights, facilitating the dynamic change in the applied path loss model according to the specific height scenario. Thus, the direct application of the TR 38.901 standard documentation is not possible for the considered UAM application.

To this end, the TR 36.777, centered around LTE-based aerial communications, is considered for designing channel models for UAM. Since the equations for TR 36.777 are developed based on version 14 of TR 38.901, applying them directly in this context is deemed a reasonable approach. The channel modeling in TR 36.777 is done pertaining to three height categories:

- 1.5 m $< h_{UE} <$ 22.5 m
- 22.5 m $< h_{UE} <$ 100 m
- 100 m $< h_{UE} <$ 300 m

Accordingly, first, the modeling of the communication channel is done under these three height categories in the case of urban macro environments. Tbl. 5.1 demonstrates LOS probabilities (denoted by P_{LOS}) according to TR 36.777 [3GP17b] under each height level of consideration. For heights above 100 m, the LOS probability is assumed to be 100% as the LOS condition improves with the height increases. Tbl. 5.2 presents the path loss definitions for the LOS case, and Tbl. 5.3 depicts the path loss definitions for NLOS conditions according to the standardization document [3GP17b] used in the channel modeling. Refer to Fig. 5.1 for the illustrations of h_{UE}, h_{BS} , d_{3D} and d_{2D} values used in the channel model equations. The unit of all these distance and height values is meters.

LOS Probability	UE Height Range
$P_{\text{LOS}} = \begin{cases} \frac{18}{d_{2D}} + \exp\left(-\frac{d_{2D}}{63}\right)\left(1-\frac{18}{d_{2D}}\right), & d_{2D} \leq 18\,\text{m}, \\ \frac{1}{1+C'(h_{UE})\frac{5}{4}\left(\frac{d_{2D}}{100}\right)^{3}\exp\left(-\frac{d_{2D}}{150}\right)}, & 18\,\text{m} < d_{2D}. \end{cases}$ Where: $C'(h_{UE}) = \begin{cases} 0, & h_{UE} \leq 13\,\text{m}, \\ \left(\frac{h_{UE}-13}{10}\right)^{1.5}, & 13\,\text{m} < h_{UE} \leq 23\,\text{m}. \end{cases}$	$1.5\,\text{m} < h_{UE} < 22.5\,\text{m}$
$P_{\text{LOS}} = \begin{cases} 1, & \text{if } d_{2D} \leq d_1, \\ \frac{d_1}{d_{2D}} + \exp\left(-\frac{d_{2D}}{p_1}\right)\left(1-\frac{d_1}{d_{2D}}\right), & \text{if } d_{2D} > d_1. \end{cases}$ $p_1 = 4300\log_{10}(h_{UE}) - 3800$ $d_1 = max(460\log_{10}(h_{UE}) - 700, 18)$, ($d_1$ in metes)	$22.5\,\text{m} < h_{UE} < 100\,\text{m}$
1	$100\,\text{m} < h_{UE} < 300\,\text{m}$

Table 5.1: The LOS probability definitions for urban macro scenario according to [3GP17b]

However, in the specified application scenario, the air taxi can ascend beyond 300 m, reaching around 600 m, necessitating consideration of this height level in channel modeling. As a result, when implementing the channel model by extending those available in Simu5G, the path loss characteristics for A2G wireless channels exceeding heights of 300 m are derived from the work of [AT19]. The decision to utilize this work stemmed from its provision of calculations for the PLE values and shadow fading values under different height conditions, which are derived from empirical measurements. This approach ensures that the channel model applied in this work is based on real-world data, enhancing its accuracy and applicability to different operational heights to suit the UAM application in discussion.

Accordingly, for the case of air taxi heights above 300 m, the LOS probability is assumed to be 1, the same as it is for the case until 300 m from the standard [3GP17b]. The PLE for the LOS case is given by Eqn. 5.3, and the path loss model for the LOS scenario is represented by Eqn. 5.4. d_0, the reference distance is considered to be the relative height. It is calculated as the height difference between the base station (BS) and UE as given by Eqn. 5.5. $\text{P}_{\text{FSPL}}(d_0)$ given by Eqn. 5.6 is the free space path loss at a reference distance

Pathloss in dB, f_c in GHz and d in m	UE Height Range
$$PL_{\text{LOS}} = \begin{cases} PL_1, & 10\,\text{m} < d_{2D} < d'_{BP} \\ PL_2, & d'_{BP} < d_{2D} < 5\,\text{km} \end{cases}$$ $PL_1 = 28.0 + 22\log_{10}(d_{3D}) + 20\log_{10}(f_c)$ $PL_2 = 28.0 + 40\log_{10}(d_{3D}) + 20\log_{10}(f_c) - 9\log_{10}((d'_{BP})^2 + (h_{BS} - h_{UE})^2)$ ■ $d'_{BP} = 4 \times h'_{BS} \times h'_{UE} \times (\frac{f_c}{c})$ ■ f_c = center frequency in Hz ■ c = 3.0×10^8 m/s ■ $h'_{BS} = h_{BS} - h_E$ ■ $h'_{UE} = h_{BS} - h_E$ ■ h'_{BS} -> effective antenna height at base station ■ h'_{UE} -> effective antenna height at UE ■ h_E -> effective environment height ■ For Urban macro scenario, $$h_E = \begin{cases} 1\,\text{m} & \text{with probability } P \\ \text{Uniform}(12, 15, \ldots, h_{UE} - 1.5) & \text{otherwise.} \end{cases}$$ where, $P = \frac{1}{1+C(d_{2D}, h_{UE})}$ with, $$C(d_{2D}, h_{UE}) = \begin{cases} 0, & h_{UE} < 13\,\text{m} \\ \left(\frac{h_{UE}-13}{10}\right)^{1.5} g(d_{2D}) & 13\,\text{m} \leq h_{UE} \leq 23\,\text{m} \end{cases}$$ where, $$g(d_{2D}) = \begin{cases} 0 & d_{2D} \leq 18\,\text{m} \\ \frac{5}{4}\left(\frac{d_{2D}}{100}\right)^3 \exp\left(\frac{-d_{2D}}{150}\right) & d_{2D} > 18\,\text{m} \end{cases}$$	$1.5\,\text{m} < h_{UE} < 22.5\,\text{m}$
$PL_{LOS} = 28.0 + 22\log_{10}(d_{3D}) + 20\log_{10}(f_c)$	$22.5\,\text{m} < h_{UE} < 300\,\text{m}$ $d_{2D} \leq 4\,\text{km}$

Table 5.2: Path loss model for LOS case in urban macro scenario according to [3GP17b]

Pathloss in dB, f_c in GHz and d in m	UE Height Range
$PL_{NLOS} = \max(PL_{LOS}, PL'_{NLOS})$ for $10\,\text{m} \leq d_{2D} \leq 5\,\text{km}$ $PL'_{NLOS} = 13.54 + 39.08\log_{10}(d_{3D}) + 20log(f_c) - 0.6(h_{UE} - 1.5)$	$1.5\,\text{m} < h_{UE} <$ $22.5\,\text{m}$
$PL_{NLOS} = -17.5 + (46 - 7\log_{10}(h_{UE}))\log_{10}(d_{3D}) + 20\log_{10}(\frac{40\pi f_c}{3})$	$10\,\text{m} < h_{UE} <$ $100\,\text{m}$ $d_{2D} \leq 4\,\text{km}$

Table 5.3: Path loss model for NLOS case in the urban macro scenario according to [3GP17b]

d_0, calculated using FSPL model (Eqn. 5.1).

$$PLE_{\text{LOS}} = 0.0059 \cdot h_{\text{UE}} + 1.9363 \tag{5.3}$$

$$PL_{\text{LOS}} = P_{\text{FSPL}}(do) + 10 \cdot PLE_{\text{LOS}} \cdot \log_{10}\left(\frac{d_{3D}}{do}\right) \tag{5.4}$$

Where:

$$do = \text{Relative Height} = h_{\text{UE}} - h_{\text{BS}} \tag{5.5}$$

$$P_{\text{FSPL}}(do) = 20 \cdot \log_{10}(do) + 20 \cdot \log_{10}(f_c) - 147.56 \tag{5.6}$$

For the NLOS case, the PLE is given by Eqn. 5.7 while the path loss model is elaborated in Eqn. 5.8. The same equations as Eqn. 5.5 and Eqn. 5.6 are applied for the relative height and the path loss at a reference distance, $P_{\text{FSPL}}(d_0)$.

$$PLE_{\text{NLOS}} = 0.0167 \cdot h_{\text{UE}} + 5.2995 \tag{5.7}$$

$$PL_{\text{NLOS}} = P_{\text{FSPL}}(do) + 10 \cdot PLE_{\text{NLOS}} \cdot \log_{10}\left(\frac{d_{3D}}{do}\right) \tag{5.8}$$

Moreover, according to the shadow fading values presented in [AT19], the shadow fading value calculation for the UE height above 300 m is also modeled. These are demonstrated by Eqn. 5.9 and Eqn. 5.10, respectively, for the LOS and NLOS cases.

$$\text{Shadow Fading}_{\text{LOS}} = -0.0086 \cdot h_{\text{UE}} + 1.473946 \quad (5.9)$$

$$\text{Shadow Fading}_{\text{NLOS}} = -0.00911 \cdot h_{\text{UE}} + 7.955077 \quad (5.10)$$

Therefore, in the implementation, the TR 36.777 [3GP17b] and the work of [AT19] are used in collaboration to achieve adequate channel model conditions compatible with the UAM application. These models are developed as a new C++ class in Simu5G by extending the existing channel model class based on TR 38.901. In the adaptation of the A2G channel, the height of the UE is recorded, and then the appropriate channel model from the described set of equations is applied accordingly.

In addition to this path loss model, a fading model is also applied to the channel model calculations to model channel characteristics. The case of Rayleigh fading, which is a model also provided with Simu5G, is used. This phenomenon describes the rapid amplitude fluctuations of a signal as it propagates through a shorter distance caused by scattering or multiple paths due to reflections or diffraction. This kind of fading is also referred to as small-scale fading. Moreover, the phenomenon of shadowing is also taken into account. This facilitates the modeling of real-world situations of mobile communication signals being obstructed by various obstacles in the environment, leading to multi-path propagation.

To this end, for the research presented in this dissertation, these channel models are thus utilized as the A2G channel model for simulating air taxis in the urban macro scenario in application to the Hamburg metropolitan region.

5.2 Base Station Network Planning for UAM Communications

Base station network planning refers to the efficient placement of base stations in a telecommunication network such that the coverage, capacity, quality of service (QoS), and cost-effectiveness are optimized. This is one of the crucial aspects of designing and deploying wireless communication networks. A base station acts as the hub that connects the UEs to the network infrastructure, enabling communication between the involved entities. Wide, reliable coverage is a priority in such a network and can be achieved through effective base station network planning. Specifically, in 5G communication, base station network planning is even more critical due to the shorter ranges of the base stations requiring a higher density of base station placement. Moreover, the application in discussion is UAM. To this end, it is not feasible to directly apply the base station network plan of ground network architecture from a terrestrial network scenario. However, the existing terrestrial networks offer the best option for quickly incorporating connectivity into UAM scenarios [PMRG24]. Nonetheless, achieving this will require new approaches in the design and comprehensive planning of the network. To support this, experiments must focus on UEs at high altitudes, and the base station network must be specifically tailored and validated to meet the needs of the application.

With a well-structured connectivity and coverage evaluation, the required number of base stations and proper placement, leading to an appropriate base station layout, can be realized. This way, the expected communication performance can be accomplished as a cost-effective and sustainable solution. This challenge is discussed in this section. A proposal for an approach to designing a base station network plan with a low number of base stations still reaching acceptable communication performance is made. Hamburg metropolitan region is considered as an application.

5.2.1 Related Work

The research community around the 5G communication network domain is currently paying a lot of attention to communication QoS upgrades in applications to futuristic technologies such as UAM. As discussed earlier in previous chapters, the general focus here has been the cargo or emergency service UAM and not much as a mode of public transportation. A broader perspective focusing on applications in all automotive, rail, and air transport is presented in [FVH+21]. The core concepts introduced alongside 5G *New Radio* (NR) standards applicable for transport-related use cases are studied, and proposals on how 5G technological components can further be utilized for enhanced QoS provisions are made. An advanced positioning technique incorporated with received reference signal measures from multiple base stations is recommended as a successful solution for UAM. In [YM20], with a focus on the network slicing principle introduced in 5G, a new concept *Airslice* to support various QoS requirements in UAVs is proposed. [HO22] is another recent research done in this domain that studies the major issues in URLLC and reviews the system performance requirements, specifically focusing on UAV communications. Moreover, [BDO+21] is another example that surveys use cases and connectivity demands for aerial vehicles. With a focus on the passenger transportation aspect of UAM, the authors formulate means to cater to the connectivity demands by combining different wireless technologies.

As such, numerous means of augmenting QoS requirements to serve aerial transportation or UAVs have been under discussion and even addressed in

[3GP17b] in relation to 3GPP release 14 on LTE. One observation here is the fact that without improvements in the base station operation, the aerial communication establishment will always end up with some severe limitations. One issue is inter-cell interference, while the other is that both aerial UE and terrestrial UE are prone to encounter challenges during cell handover or when selecting a new cell. [PLRK22] have recognized the cause of these issues to be limited LOS and the sidelobe beam-based LTE unmanned aerial vehicle operations. To this end, first, the authors propose deploying an additional antenna array at the base station with an upward mechanical tilt. This technique of creating aerial cells over the ground network is done as a means of improving the reception power of aerial UE signals. This way, the inter-cell interference generated by ground UEs is managed. The downside here is the system is expensive and complex. As a result, the authors recommend a concept using the same antenna array for both aerial and ground UEs but with a vertical beam steering on the down-tilted antenna array. This proposal has the limitation of increased path loss but has proven to be performing better than the approach from [3GP17b].

The above highlights the significance of handling the challenge of connectivity, specifically in relation to the ground base station network. Despite that, the amount of work done in the scope of investigating the influence of ground network planning aimed at UAV communications is limited, in particular, for the case of 5G communications. Even the work explored previously is centered around LTE. To this end, the previous research on base station planning focusing on ground users is first examined. One area of research here is the coverage optimization of base stations by employing various deployment techniques. One of the early research done focusing on this topic is [SZL17]. The proposed network architecture combines macro and micro base stations to enhance coverage. Experiments are conducted under both lossless and lossy conditions, analyzing various patterns and numbers of base station placements. The goal has been to evaluate and enhance the system's performance under different scenarios.

The contributions in [SSH20] is another work falling into the same category with a focus on 5G communication. The authors have proposed two mathematical models for automatic ground network planning. The objective has been to optimize the placement of gNBs, aiming to determine the minimum number required while ensuring that the desired coverage is fully achieved. The models have been tested for deployments in new ground locations without existing gNBs as well as for locations with pre-existing gNBs. These models have been validated for the deployment scenarios of urban, suburban, and rural areas, and the recommendation is to include the models in network planning tools for enhanced ground infrastructure design. Another example using a similar approach is [RNA21]. A capacity and coverage-based planning approach based on [3GP17a] urban macro scenario is discussed. The focus has been on ground users. Two separate analyses have been conducted to study the number of base station requirements, considering a capacity-based analysis and a coverage-based analysis. Nevertheless, the aspects of low latency or cost-effectiveness have not been taken into account under their design considerations. Among other work, [Tun21] addresses radio network planning and optimization, focusing on power allocation in 5G terrestrial user scenarios.

In addition to these works from academia, radio network planning tools are being widely used in the industry. A few examples of such tools are Atoll Radio Planning Tool [Rio19, MB14], RadioPlanner [Wir17], and ASSET Radio [TEO22], employed in the industry. Of these, Atoll Radio Planning is predominantly employed in the industry [SSH20]. Conventional network planning principles were applied as the basis for the first radio network

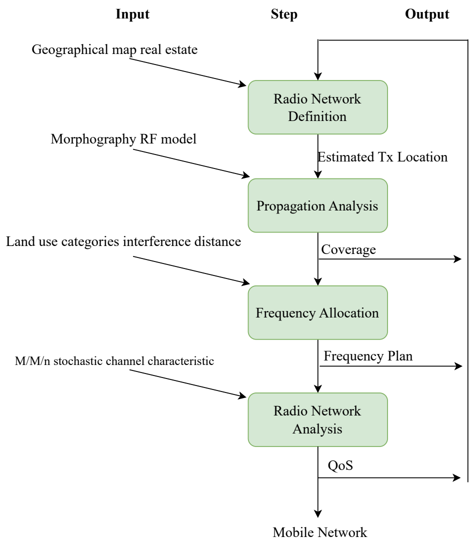


Figure 5.2: The traditional cellular network planning strategy according to [Tut98]

planning tools in the late 1900s. However, due to limited consideration given to network capacity issues and the inability to gain overall optimization due to non-integrated design steps, in [Tut98], a new integrated network planning paradigm is introduced, and this is being studied in new network planning tool designs such as Atoll Radio Planning Tool [SSH20]. The steps followed in conventional network planning can be explained as in Fig. 5.2 based on [Tut98]. This is an analytical radio frequency design-oriented approach dating back to 1986 and consists of four stages. In the first stage of radio network definition, cell sites are chosen manually, typically following the hexagonal grid design. Based on the transmitter parameters, then the propagation analysis is done. In the third stage, the network capacity issues are handled and is called frequency allocation. The final stage is radio network analysis, where a QoS evaluation is performed.

However, due to the limited focus on network capacity issues and the lack of overall optimization resulting from non-integrated design steps, a new integrated network planning paradigm is proposed in [Tut98]. In the integrated planning method, the four modules, radio transmission, mobile subscriber, system architecture, and resource management, are input parameters, and in the automatic network design module, the network configurations are generated. This way, the interconnections and dependencies between the input modules are captured in the network design, and thus, the network capacity issues are addressed. This approach is depicted in Fig. 5.3.

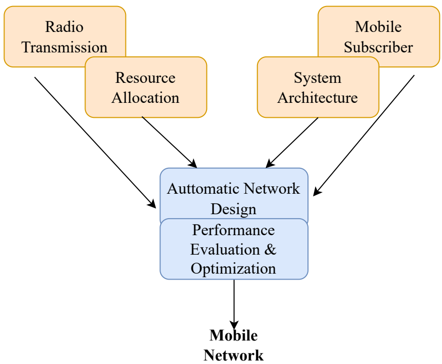


Figure 5.3: The integrated network planning approach introduced by [Tut98]

Based on the discussed strategies, it is apparent that a wide range of work in both academia and industry has been done focusing on network planning in the domain of 5G as well as on other communication technologies. However, these are all focused on terrestrial users and thus cannot be directly integrated into an aerial UE communication scenario due to changes in the channel conditions.

An alternative suggestion from the previous research in this scope is to use UAV-mounted base stations for connectivity enhancements as elaborated before under Cellular-Assisted UAVs in 3.4.2. This proposal is highlighted in the work of [DA21], where a machine learning-based approach is used for aerial base station placement optimizations. The results indicate that the technique is promising for UAM-based applications. Despite that, the motivation for the work in this chapter is a cost-effective base station architecture, and therefore, an extension of the existing ground network architecture is intended. To this end, the aerial base station architecture is not taken into consideration in this work. It is further to be noted that a relatively new research investigating the connectivity enhancements for UAM concentrating on terrestrial base station networks while using a minimum number of gNBs is addressed in [PMRG24]. With power adjustments leading to transmit power limitations and specifically designed antennas, a base station enhancement is proposed. These base stations used in their implementations are referred to as *enhanced-base stations*. The authors propose two optimization problems with the objective of reaching a target average throughput and minimum throughput per UE. This approach is proven to enhance connectivity while limiting the needed number of enhanced base stations in the UAM communication network. Despite that, the work presented in this chapter is based on [WT23b], which was completed before the work in [PMRG24], and hence the solutions suggested by the authors of [PMRG24] are not taken into account when compiling this chapter.

5.2.2 Network Planning Approach

The above-highlighted considerations in facilitating connectivity in the UAM scenario are discussed in this section with an approach to model a base station network. In network planning, the first step is recognizing the influential factors on UAM communications in relation to the gNB network. This way, the impact of the gNB layout on the performance of UAM communications can be evaluated. With this regard, the two directly influencing factors are the number of base stations and the placement strategy. Given the specific performance requirements, a dense deployment of gNBs is one possible option to achieve enhanced coverage and capacity. Nevertheless, this solution is costly in terms of implementation and maintenance. This limits its applicability because streamlining a cost-effective process is a key requirement. To this end, the prevailing challenge can be framed as an optimization problem with the number of gNBs and the gNB placement as the decision variables. In terms of objective functions applicable to the specific case, there exist two conflicting objectives that need to be optimized.

- Cost - Effectiveness - Minimize the total number of gNBs deployed
- Network performance - Maximize the performance in terms of latency and reliability

The challenge involves multiple decision-making criteria, leading to a trade-off between two conflicting objective functions. As these objectives need to be optimized simultaneously, the problem can be defined as a multi-objective optimization problem. The constraints here include coverage, meaning each UE must belong to the coverage area of at least one base station and gNB capacity for user connections. The said problem can thus be expressed as a linear optimization problem with linear objective functions and constraints and be solved using linear programming techniques. Alternatively, it can also be formulated as a quadratic optimization problem with quadratic non-linear relationships and addressed using quadratic programming methods.

However, such an approach would be computationally expensive to solve. It consists of a multi-objective optimization with non-linear performance metrics and interdependencies, such as the influence of gNB placement on both delay and reliability, which complicates the optimization process. Solving problems of this nature often requires iterative algorithms to explore a vast solution space. This leads to significant computational demands. In real-time, therefore, this can be infeasible without significant computational resources or heuristics due to associated computational overhead. Heuristics can simplify the problem by providing approximate solutions that significantly reduce computational time while still reaching acceptable levels of performance.

Therefore, through the work described in this section, rather than attempting to directly solve the optimization problems, heuristics are proposed to improve the metrics. Accordingly, first, all the influential factors on the network performance are recognized in order to determine the applicable heuristics. Inherently, the main influencing factor is the number of gNBs. It is evident that increasing the number of gNBs enhances coverage, reduces latency, and improves overall capacity by distributing the load across multiple nodes. However, even if that condition is satisfied, without considering a proper gNB placement, the network would not yield the expected high communication performance. Planning the network following strategic placement techniques leads to enhanced signal strength and reduced latency and ensures efficient resource utilization. Factors such as terrain, no-fly zones, and existing buildings should be considered when locating the base stations. From a technical point of view, proper gNB placement facilitates low interference and better handover processes between cells, ensuring seamless connectivity for users as they move through the urban network. On the other hand, this is important as a means of fulfilling a social responsibility related to health, safety, and environmental concerns. For this reason, gNB placement topologies must be studied and evaluated to realize the appropriate placement.

At the same time, the number of airframes handled by a gNB is another essential aspect. By measuring the number of airframes received at each gNB, it is possible to identify the gNBs that are least contributing to the ongoing communication in the network. This helps in reducing the overall network size. Alongside that, a different dimension to look into is what happens if the number of airframes exceeds the gNB's capacity. That is, if the gNB has to serve an excessive amount of UEs. This situation can result in a congested network. This degrades the service quality, impacting the overall user experience. Therefore, this is a contributing factor in terms of communication performance. Considering the presented influential factors, the complexity of the system, and the resource availability, heuristics are proposed. These heuristics are aimed at simplifying the search domain by applying practical, rule-based iterative improvements to the system. The proposed heuristics include:

- Greedy Reduction
- Hybrid Placement Strategy
- Load-based optimization
- Load Redistribution

Greedy Reduction

This heuristic involves initiating the network configuration with random placement of a large number of gNBs and experimenting with reducing the number of gNBs while monitoring the performance metrics. This is a very simple approach and sets the ground for further refinement in the base station placement and performance improvements. The challenge here is having a random placement, which can end up in an inefficient initial base station network.

Hybrid Placement Strategy

This heuristic refers to experimenting with different base station placement techniques leading to a suitable base station layout. For example, manual and grid-based placements are two options. In the grid-based placement, uniform distribution of the gNBs is facilitated by applying a structured approach to divide the network area onto a grid and placing gNBs at the grid intersection points. This method ensures uniform coverage. However, real-world constraints such as environmental conditions are completely ignored. On the other hand, in manual placement, network area and traffic demand knowledge are incorporated into refining the placement. Therefore, this technique tailors a layout specific to the needs and demands of the network area, while the downside is that the method is time-consuming.

Load-based Optimization

Under this heuristic, the number of messages that can be handled at each base station is evaluated. The technique intends to analyze the base station locations with high traffic demand and figure out the idle base station locations that are completely useless due to no traffic involvement. Through several iterations, this method helps identify a scaled-down network plan by investigating the real-time performance.

Load Redistribution

With this heuristic, the heavily used base stations are replaced with multiple gNBs to facilitate load balancing. Overburdened gNBs leading to communication delays are recognized, and the load is balanced by replacing the base station with two or more closely placed gNBs, depending on the demand. This is a simple technique to address individual performance issues but also contributes to increased costs due to introducing additional gNBs.

Proposed Network Planning Strategy

While each heuristic individually contributes to a better network configuration in its ways, each also comes with certain limitations. Accordingly, it is beneficial to develop a strategy where these individual heuristics are integrated and applied in a sequential procedure. This way, all the influential factors are taken into account, leading to a refined base station placement with reduced costs while achieving required performance targets. To this end, a step-by-step straightforward procedure for designing the base station plan is proposed in this section, involving the discussed heuristics. The described procedure addresses each heuristic in the order of importance of each, considering the influence of each factor on the performance requirements and cost-effectiveness.

Accordingly, the first heuristic in the approach is greedy reduction, which is involved with the number of gNBs. A significant concern here is realizing a starting value for the number of gNBs. This cannot be done randomly due to the lack of knowledge about the demand that needs to be handled. As a result, the chosen values may not reflect the associated conditions and influential factors. To this end, a viable solution is applying link budget calculations adopted from 3GPP TR 38.901 version 18 [3GP24a]. Thus, these calculations are the initiation point in the network planning approach. Next, starting from the resulting number of gNBs from link budget calculation, greedy reduction is applied. Then, hybrid placement strategies are tested for the resulting number of gNBs, and a placement technique is finalized. Then, load-based optimization is performed on this network, which produces a scaled-down network architecture with a reduced number of gNBs after removing the idle gNBs. To that result, as the final step, load distribution is applied, and this yields the final gNB layout. This procedure is demonstrated in Fig. 5.4. This proposed strategy is discussed in detail with reference to an application in the upcoming section.

5.3 Application

This section presents an application of the step-by-step procedure for the gNB network plan proposed in the previous subsection. As the application setting, UAM implementation for the Hamburg metropolitan region is chosen. Emphasizing specifically a UAM scenario utilizing 5G communication, simulations are conducted for this purpose by applying the ULTRAS simulation framework (Sect. 4.2.3). For the air taxis, the trajectories from the toolchain catering to the calculated demand to suit the Hamburg metropolitan region are used. The flight dynamics and control structures are also incorporated from the ULTRAS simulation framework. Alongside that, the support of Simu5G library [NSTV20] is employed for 5G protocol stack development of the gNB, UE, and the 5G related network entities such as user plane function during the network planning. The communication from the aerial UE to the ground station is focused. The channel model elaborated in Sect. 5.1.2 is applied as the A2G communication model for interactions between the UE and the ground station.

In terms of simulations, the considered simulation scenarios and the simulation time durations determine the possibility of obtaining accurate and

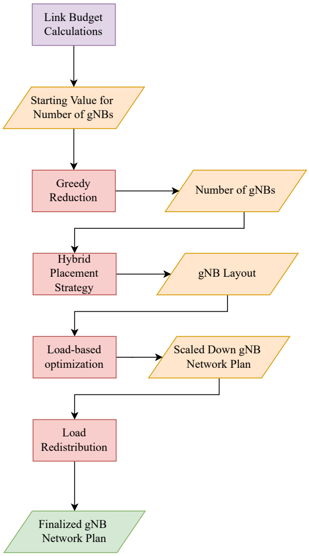


Figure 5.4: The proposed network planning approach

meaningful results. The flight schedule from the ULTRAS simulation environment captures the schedule for an entire day. With regard to the respective time of the day, the demand varies. Accordingly, the flight schedule shows a variation of the planned number of simultaneous air taxis throughout the day. The ultimate goal here is to estimate the demand to be served for better network planning and the number of users directly contributing to the capacity constraints. Therefore, for the evaluations in this chapter, it is essential to cover the daily variation of air taxis in the network. However, it is impractical to consider the entire 24-hour duration for communication simulations, as it takes a longer duration in real-time to simulate this. Therefore, only certain portions of the day representing different flight distribution situations are taken into account for testing. For that reason, first, a simulation for a 24-hour time period under ideal conditions (without applying any channel model or 5G communication specifics) is run, and the flight distribution is observed

Scenario	Time period of the day
Scenario 1	21600 s - 23600 s (06:00 am to 06:33 am)
Scenario 2	39946 s - 41946 s (11:05 am to 11:39 am)
Scenario 3	63334 s - 65334 s (17:35 pm to 18:08 pm)

Table 5.4: Simulation scenario definitions

throughout the day. Based on this distribution, different patterns of variations in the active number of air taxis in the airspace are identified. For instance, a rapid increase in the number of air taxis or a sudden drop in the number is considered significant.

Three different scenarios pertaining to three different time windows are chosen. Each time window has a duration of 2000 s. Fig. 5.5 depicts the air taxi distribution throughout the simulated period of 24 hours, and the three selected scenarios are highlighted in the figure. The first scenario illustrates a rapid growth in the number of air taxis. The second scenario is selected to showcase a symmetric variation in the air taxi numbers. The third scenario demonstrates a rise to the peak value of the daily count of air taxis (110), followed by a sharp decrease in the number. The respective time windows of each scenario are presented in Tbl. 5.4. The study is limited to these three time windows. However, by examining three distinct variations of air taxi schedules, including the case with the maximum number of air taxis available in the environment, a generalized perspective of the network is captured in the simulations. This approach thus ensures that the tests account for a range of conditions, providing a comprehensive view of network performance.

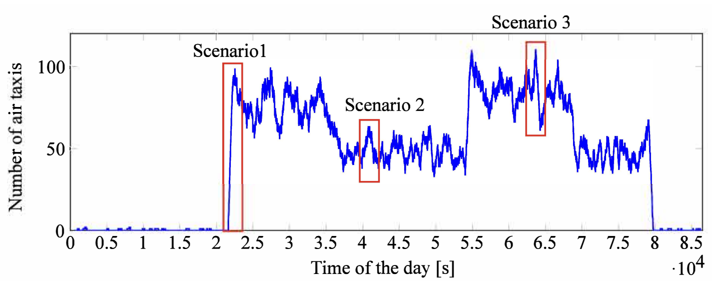


Figure 5.5: The air taxi distribution over a period of 24 hours highlighting the three chosen simulation scenarios

5.3.1 Simulation: Network Planning Approach

This section demonstrates the simulation of applying the proposed gNB planning approach to the Hamburg metropolitan region. As highlighted in Fig 5.4, starting from the link budget calculations, the step-by-step procedure following each proposed heuristic is applied for dimensions and conditions satisfying the Hamburg metropolitan region. First, the calculations and considerations under each step are described. Later, the results leading to the cost-effective gNB plan to suit the Hamburg Metropolitan Region are discussed.

Link Budget Calculations

As explained before, the baseline or the initial value for the number of base stations is set by the link budget calculations following the 3GPP standardization document TR 38.901 [3GP24a]. Although this TR is focused on ground user scenarios, the same calculations are applied to realize a suitable starting point. Regardless of the 2D or 3D space of deployment, a general feasible value for the number of base stations is intended. The maximum allowable path loss (MAPL) is calculated first. It represents the maximum amount of signal loss that can occur between a transmitter and receiver while still allowing successful communication. In communication network design, MAPL is essential to determine the coverage area, cell size, and gNB placement. For the specific case in discussion, MAPL is calculated in decibels for NLOS conditions considering the equivalent isotropically radiated power (EIRP), receiver antenna gain (G_{UE}), sensitivity of the receiver ($P_{UE_{sensitivity}}$), interference margin (IM) and shadowing margin (SM) into account. The MAPL is given by Eqn. 5.11.

Here, EIRP (given by Eqn. 5.12) is a measure of the power radiated by an antenna in a specific direction as if it is emitted by an isotropic antenna (an ideal antenna that radiates equally in all directions). This represents the maximum power that an antenna can emit in any given direction, taking into account the transmitter power ($P_{TX_{BS}}$) and the gain of the antenna (G_{BS}). $P_{UE_{sensitivity}}$ is represented by Eqn. 5.13, and it is the minimum signal strength that a receiver can detect and still successfully decode the transmitted data with an acceptable level of error. It indicates how weak a signal the receiver can handle while maintaining reliable communication. Here, G_{BS} stands for gNB antenna gain, NF_{UE} is the user equipment noise figure, while SNR is the signal-to-noise ratio defining the achievable signal strength in relation to the background noise level. The formula for SNR is illustrated by Eqn. 5.14, where α_S is the Shannon capacity scaling factor, a parameter that influences the maximum data rate of a communication channel, and w is the bandwidth. Together, the inverse proportionality in this relationship conveys the idea that expanding the bandwidth of a communication system generally increases the noise power. Of the other parameters, IM is the additional buffer built into a communication system to account for potential interference from other signals. SM defines the additional tolerance in signal strength to compensate for fluctuations caused by obstacles in the environment, such as tall buildings, trees, or other structures in a city.

$$MAPL = EIRP + G_{\text{UE}} - P_{\text{UE}_{\text{sensitivity}}} - IM - SM \tag{5.11}$$

$$EIRP = P_{\text{TX}_{\text{BS}}} + G_{\text{BS}} - \text{cable loss} \tag{5.12}$$

$$P_{\text{UE}_{\text{sensitivity}}} = -174 + 19\log(G_{\text{BS}}) + NF_{\text{UE}} - SNR \tag{5.13}$$

$$SNR = 2^{\frac{1}{\alpha_S \cdot w}} \tag{5.14}$$

Based on the resulting MAPL value, the 3D distance between the UE and the gNB (d_{3D}) is calculated by solving Eqn. 5.15. PL_{NLOS} stands for path loss for the NLOS case and is equal to the MAPL calculated earlier in 5.11. This is the same mathematical model for the NLOS path loss calculation for the case of 1.5 m < h_{UE} < 22.5 m in Sect. 5.1.2. This model helps in predicting how much signal strength will be lost as a signal travels from a transmitter to a receiver under conditions where there are obstacles in the LOS between them. For generalization, an architecture with a square grid placement of gNBs is assumed for calculations, although the real implementation of UAM conditions is different from this, with regard to the h_{UE} value range. h_{UE} is the height of the UE while f_c is the carrier frequency.

$$\begin{aligned} PL_{NLOS}(d_{3D}) = \; & 13.54 + 39.08\log_{10}(d_{3D}) + \\ & 20\log_{10}(f_c) - 0.6(h_{UE} - 1.5) \end{aligned} \tag{5.15}$$

The solution of Eqn. 5.15 is then applied on Eqn. 5.16 to calculate the cell radius value where h_{BS} is the gNB height. This formula is derived from the Pythagorean theorem in three-dimensional space.

$$CellRadius = sqrt[d_{3D}^2 - (h_{BS} - h_{UE})^2] \tag{5.16}$$

Finally, the number of base stations is calculated via Eqn. 5.17. The above-calculated cell radius value is used here in such a way that the network area is

Parameter	Value
UE height (h_{UE})	1.5 m
gNB height (h_{BS})	25 m
Carrier frequency (f_c)	2 GHz
Bandwidth (w)	20 MHz
UE antenna gain (G_{UE})	0 dBi
gNB antenna gain	18 dBi
cable loss	2 dB
UE noise figure	7 dB
gNB Tx Power	46 dBm
Shannon capacity scaling factor (α_S)	0.65
Interference Margin (IM)	5 dB
Shadowing Margin (SM)	6.22 dB
Network Area	42 km x 40 km

Table 5.5: Parameter values for link budget calculations

divided into an equidistant square grid.

$$NumberOfBaseStations = \frac{NetworkArea}{(2 \times CellRadius)^2} \tag{5.17}$$

The described link budget calculations resulted in the number of required gNBs to be approximately 47. The considered values for each parameter in all calculations from 5.11 until 5.17 are presented in Tbl. 5.5. These values are assumed for the calculation following the default values provided in the 3GPP standardization document TR 38.901 [3GP24a]. For the network area parameter, the value is chosen to suit the area dimensions of the city of Hamburg.

Greedy Reduction

An approximation for the required number of gNBs was identified through the link budget calculations presented above. On this basis, different numbers of gNBs are decided to be considered in simulations. This is a greedy reduction approach. Thus, three instances are defined around the calculated value of 47. The chosen values for the number of gNBs are 30, 40, and 50. It is to be noted that a random gNB placement is considered here.

Hybrid Placement Strategy

Under the heuristic of the hybrid placement strategy, two placements are studied. The first placement strategy is *manual placement*. Here, first, a full simulation run of the ULTRAS framework for a period of 24 hours is observed from time to time randomly during the simulation period. Based on the observations, the popular locations in the city are identified depending on the frequency of the air taxis scheduled during the day. Then, the map of Hamburg is further followed to avoid placing the gNBs in restricted areas and proximity to sensitive locations such as hospitals and schools, airports, environmentally protected sites, etc. Depending on the locations identified with these two studies, a manual gNB placement is carried out. An example of such a placement is showcased in Fig. 5.6.

The second kind of placement is a grid placement. In this case, the physical

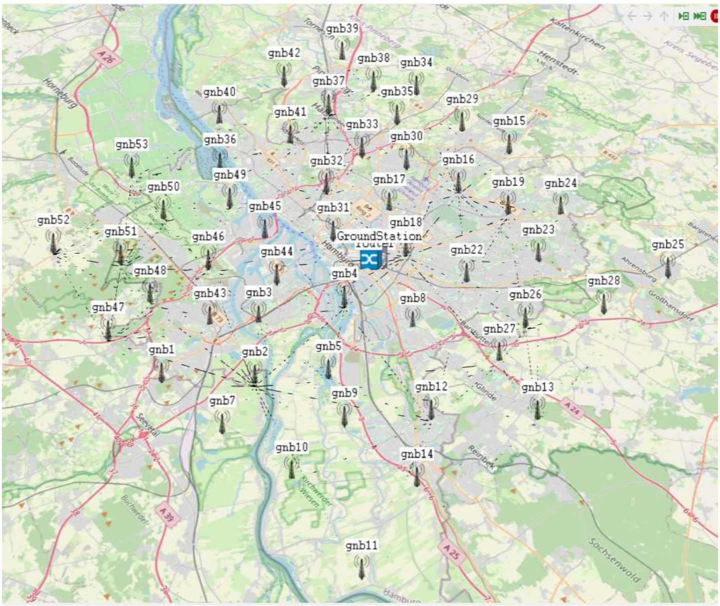


Figure 5.6: Example of gNB manual placement from the ULTRAS simulation environment focusing on Hamburg metropolitan region

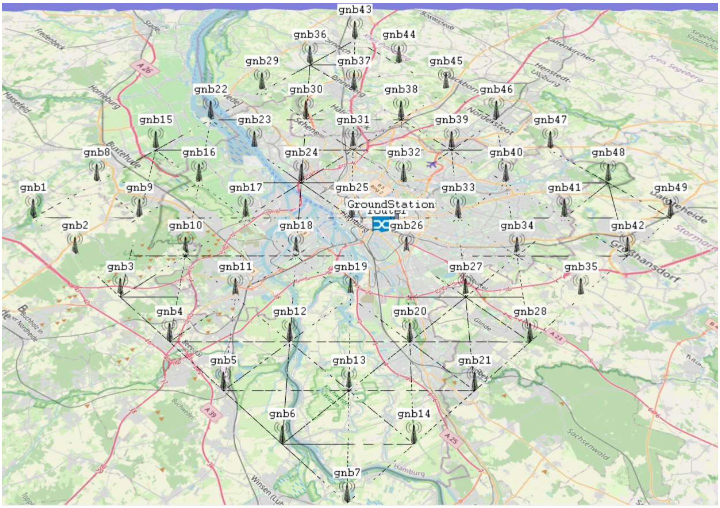


Figure 5.7: Example of gNB grid placement from the ULTRAS simulation environment focusing on Hamburg metropolitan region

constraints in the city environment are not taken into account. Instead, the gNBs are placed in an equidistant square grid architecture as depicted in Fig. 5.7.

Load balancing

Delving further into reducing the number of gNBs required, the heuristic of load balancing is applied. Here, the third influencing factor, the involvement of the placed gNBs in message transmissions between the air taxi and the ground station, is investigated. To analyze this, the number of airframes handled by each gNB is recorded. An airframe in the context of Simu5G typically refers to the time structure used to organize data transmission over the radio interface. These frames are part of the physical layer transmission scheduling.

Load Redistribution

The final heuristic is load redistribution, which explores the effect of load sharing. Depending on the placement, certain gNBs can be in the range

Heuristic	Scenario	Description
Greedy Reduction	2	Simulating 30,40 and 50 number of gNBs
Hybrid Placement Strategy	All three	Studying two gNB placement techniques (manual and grid)
Load-based Optimization	All three	Examining the number of airframes received at each gNB
Load Redistribution	3	Investigating the effect of sharing the load by replacing a high demand gNBs with closely located two gNBs

Table 5.6: Performed simulations

of a large number of air taxis, hence having to serve multiple aerial UEs simultaneously. Due to this reason, the amount of traffic to be handled through these gNBs increases, causing congestion and leading to increased transmission delays. Although a key concern is to reduce the number of gNBs, maintaining reliable and low latent communication is the underlying requirement in UAM communication.

Simulations

A set of simulation runs focusing on the application setting is conducted. The scenarios defined in Tbl. 5.4 are simulated, depending on the scope of expected results for the analysis. Tbl. 5.6 gives an overview of the performed simulations. The used parameter values in the simulations are described in Tbl. 5.7.

The application being UAM, a low signal attenuation is intended. The channel bandwidth and carrier frequency values are chosen to serve this purpose. This helps in managing signal strength over extended ranges, ensuring reliable connectivity. It also helps reduce the impact of interference and environmental obstacles, which is essential for safe UAM operation in urban settings. As per transmitted messages, exchanging position data for localization is considered. Thus, the message contains the position coordinates of each air taxi. This accounts for a very small message size. Hence, it is assumed to not exceed 40 bytes and 40 is chosen as the message size. The air taxis communicate with the ground stations periodically every 2 s.

5.3.2 Analysis of Results

The simulation results from the application of the discussed network planning approach based on the proposed heuristics to the Hamburg Metropolitan region are elaborated in this section. Following the results, a recommendation for an efficient and cost-effective gNB network architecture suitable for

Parameter	Value
Carrier frequency	2 GHz
Channel Bandwidth	20 MHz
Message Size	40 bytes
Message Interval	Every 2 s

Table 5.7: Parameter values used for the simulations

UAM operations in Hamburg is presented. The analysis is done based on the delivery ratio and delivery delay. These performance indicators are chosen to get an understanding of the reliability of the communication system and the latency involved, which are the two performance sectors that are expected to be enhanced with the applied gNB network. Each simulation is repeated 10 times. The results presented are the 90th percentile values reached after analyzing all 10 simulation result sets for each. The goal is to gain a clearer insight into the system's actual performance while minimizing the influence of outliers on the results. It's important to highlight that careful consideration is given to the statistic variance, with values falling within a range of 0 to 5.

Greedy Reduction

Only scenario 2 from Tbl. 5.4 is taken into account here. This scenario depicts a symmetric variation in the number of simultaneously active air taxis. It is decided that it is sufficient to test only this one scenario as it provides a controlled and consistent baseline. Simulations are done by varying the number of gNBs in the network to be 30,40 and 50. The results obtained for the 90th percentile values pertaining to the delivery ratio (Fig. 5.8(a)) and delivery delay (Fig. 5.8(b)) are depicted in Fig. 5.8. A gradual improvement in performance is noted as the number of gNBs increases from 30 to 50. This outcome aligns with the principle that a more significant number of gNBs in a network typically leads to enhanced system performance. Accordingly, the outcome of the first step of the procedure suggests using 50 gNBs.

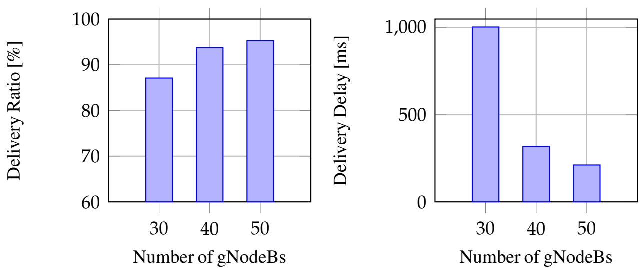
Hybrid Placement Strategy

For the second procedure step, the factor of consideration is the heuristic gNB placement. Following the previous results, a gNB placement architecture for 50 gNBs placed manually and a square grid with 49 gNBs are simulated for all scenarios from Tbl. 5.4. This is because the placement of gNBs can have varying effects depending on the distribution of air taxis. Therefore, it is essential to test all three scenarios to capture these potential variations in performance. The 90th percentile values for the delivery ratio and delivery delay value comparison for the manual and grid placement techniques are depicted in Fig. 5.9 and Fig. 5.10 respectively. It is clearly observed that results improve when manual placement, considering external environmental conditions, is applied. However, in scenario 3, delivery delay variation in both the manual and grid-based approaches shows similar results. In contrast, for the delivery ratio in the same scenario, manual placement shows better performance, although the difference is minimal. These findings demonstrate that gNB placement has a direct impact on system performance, highlighting the importance of carefully considering the external environment during the placement process.

Load Balancing

In this step of the procedure, the exploration of the involvement of each gNB in the message transmissions is done as a means of reducing the involved costs. From the previous two simulations, it was identified that the number of gNBs to consider must be 50, along with the manual placement. As a result, here, these conditions are applied to all three scenarios from Tbl. 5.4, leading to an overall evaluation of the system. Results for each cell ID pertaining to the 50 gNBs in the environment are depicted in Fig. 5.11. Two major observations are made:

- 11 gNBs did not receive any airframe under any scenario of consideration



(a) Delivery Ratio (b) Delivery Delay

Figure 5.8: Comparison of performance for different numbers of gNodeBs

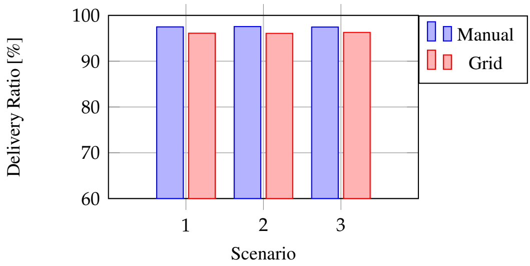


Figure 5.9: 90th percentile value for the delivery ratio of the manual (50 gNodeBs) and grid (49 gNodeBs) based architecture for the three scenarios

- Some gNBs are only involved in certain scenarios (e.g.: cell ID 2 receives airframes in scenarios 1 and 3 but not in scenario 2)

This simulation made it possible to generalize the gNB network plan by removing the 11 gNBs that are not at all used in message transmissions. This ended up with a gNB network plan consisting of 39 gNBs. Due to previous recognition of proper placement based on environmental conditions, the gNB positioning is reevaluated. The vertiport placement information from the work in [BHW+23] is also considered. Accordingly, there is a vertiport placed in the middle of every city district in Hamburg, accounting for 103 vertiports. Concurrently, the following two factors are also explored when determining the gNB placement for the network architecture design.

- Number of air taxis handled by each vertiport during the day

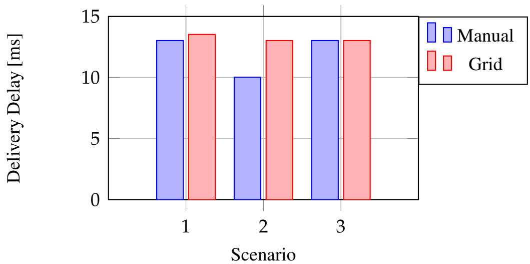


Figure 5.10: 90th percentile value for the delivery delay of the manual (50 gNodeBs) and grid (49 gNodeBs) based architecture for the three scenarios

- The distribution of air taxis in the airspace depending on originating and destination vertiports

These distributions emphasized the locations with a possibility of high density of air taxi expectations during the day. This is done with the intention of getting a more generic picture suitable for the overall UAM network. After that, another set of simulations to examine the airframe reception under all three scenarios for the new scaled-down gNB network plan with 39 gNBs is done. It is a measure to recognize the potential for further re-locations to replace the least utilized gNodeBs within the new plan. Fig. 5.12 illustrates the obtained results from the simulation. This revealed that cell IDs 9 and 37 are still not involved at all with the communications. For this reason, these two gNBs are removed from the network. Consequently, these simulations concluded a scaled-down gNB network plan with 37 gNBs.

Load Redistribution

Other findings emerged from the previous simulation results shown in Fig.5.12 revealed that in cell ID 24, a peak in the number of received airframes is consistently observed across all three scenarios, indicating that it is the gNB with the highest level of interactions overall. Building on this insight, this cell is chosen to test the load redistribution as the final step of the procedure. Consequently, cell ID 24 is substituted with two nearby gNBs positioned around its original coordinates. The simulation is conducted only for the scenario 3. This decision is grounded on the observation that, in the previous results, all three scenarios exhibited a similar range in the number of received airframes, with scenario 3 yielding the highest value. Thus, limiting the test to

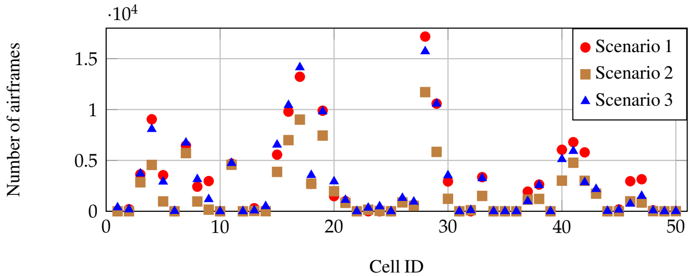


Figure 5.11: Number of airframes observed in 50 gNodeB manual placement case for the three scenarios

scenario 3 is deemed a valid approach. Results are evaluated as a comparison between two configurations:

- Configuration 1: The gNB network before load sharing - 37 gNBs
- Configuration 2: The gNB network after load sharing - 38 gNBs

The evaluation is done under delivery ratio and delivery delay aspects. The delivery ratio remained unaffected by load sharing, with values for both configurations aligning close to 97%. Generally, in wireless transmissions, the packet drops are caused by the associated channel conditions. Therefore, the principles, such as load sharing, are not sufficient to improve performance. However, in terms of delivery delay, depicted in Fig. 5.13, a notable reduction is recorded, decreasing from approximately 13 ms to 10 ms. For time-sensitive applications like UAM, low latency is crucial, and the results prove that load sharing can play a vital role in achieving this objective. It is thus evident that the application of such load sharing for the rest of the heavily used gNBs would be a better solution to reach the gNB network with ideal low latency conditions. Despite that, one constraint under consideration is the associated cost. Increasing the number of gNBs incurs added costs in terms of implementation as well as maintenance. Each additional gNB requires investment in infrastructure, setup, and operations. Therefore, it is essential to realize the right balance between the number of gNBs and system performance.

To this end, by applying the proposed heuristic-based procedure for gNB network planning for the Hamburg metropolitan region, a gNB network architecture consisting of 38 gNBs is finalized based on the trade-off between cost and performance. This configuration is deemed a suitable network plan

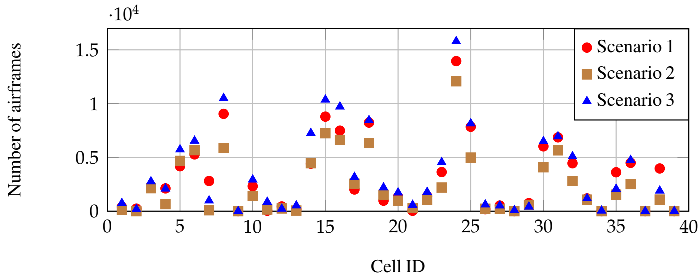


Figure 5.12: Number of airframes observed for the network plan with 39 gNodeBs under manual placement for the three scenarios

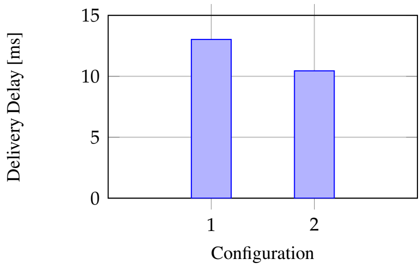


Figure 5.13: Comparison of the delivery delay for the two gNodeB configurations; 1: before load sharing. 2: after load sharing

to serve Hamburg during an application of UAM. It is to be noted that, in a gNB network plan, for dealing with emergency reasons, it would be better to have gNBs, specially within the city boundaries. Therefore, such gNBs cannot be ignored, although they are out of the scope of this work.

This gNB network plan consisting of 38 gNBs is given in Fig. 5.14. The 90^{th} percentile value for delivery ratio for this setup is recorded to be 97%. This is an acceptable value for wireless communication scenarios with a

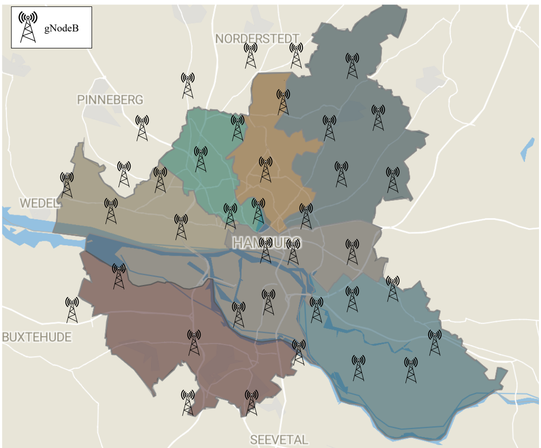


Figure 5.14: Finalized gNodeB placement for the city of Hamburg

high number of users in the presence of interference conditions such as shadowing and fading and using UDP transmissions. This value can thus be further improved by employing TCP retransmission mechanisms, which leads to improved reliability, although this wasn't tested during this work. As TCP was originally designed for wired communication and in dealing with wireless communication, it is common to rely on UDP protocols [FZZ23], the current evaluation only considered UDP. On the other hand, the 90$^{\text{th}}$ percentile value for the delivery delay is recognized as approximately 10 ms. In [HO22], the authors highlight the requirements for ultra-reliable low-latency communications, and for the application of UAM, 10 ms is stated as the required latency level. This complements the applicability of the proposed gNB network to gain the required system performance suitable for UAM. Additional analysis of the results is done through a delay spread observation over the span of time duration pertaining to scenario 3. The following observations are made:

- For 95% of the time, the delivery delay is less than 50 ms
- The average delivery delay accounted for 60 ms
- Delay value range: 4 ms to 7 s
- Variance: 7.4% for the span of observed 2000 s

The obtained results can further be justified based on the system requirement definitions of [BHW+23]. According to data provided in [BHW+23], the minimum separation distances have been identified to be 600 m laterally and 45 m vertically. For instance, assuming the speed of the air taxi to be 100 km/h, the associated latency of 60 ms on average does not pose a significant risk in maintaining safe operations under these circumstances.

5.4 Use Cases

For the purpose of evaluating the network plan resulting from the application of the proposed network planning procedure in Hamburg, several use cases presented in Tbl. 5.8 are discussed in this section. For different types of data associated with sensing in the UAM system, use cases are defined and evaluated. Data types pertain to covering GPS and cameras as sensors. Packet sizes and other general parameters applied in the simulations are defined in Tbl. 5.9. The three data types are: Coordinate data, Image data, and Video data. The results for delivery ratio (Fig. 5.15(a)) and delivery delay (Fig. 5.15(b)) comparison for different use cases are provided in Fig.5.15.

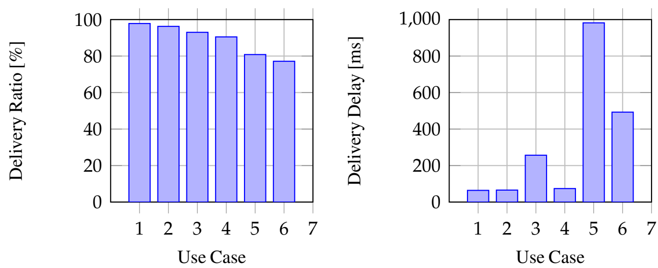
Use Case	Data Type	Bandwidth (MHz)	Period (s)
1	Coordinates	5	2
2	Coordinates	20	2
3	Image	20	uniform(2,5)
4	Image	50	uniform(2,5)
5	Video	100	uniform(2,5)
6	Video	100	uniform(15,18)

Table 5.8: Use Cases

Parameter	Value
Tested scenario	Scenario 2
Carrier frequency	2 GHz
Coordinates data size	40 bytes
Image data size	A file size uniformly distributed between 130 - 150 KB
Video data size	A file size uniformly distributed between 1 - 2 MB
Performance evaluation criteria	Delivery Ratio and Delivery Delay

Table 5.9: Used general parameters for simulating use case

The communication protocol layers have maximum transfer unit values. Due to these limitations, if the file size is higher than the maximum transfer unit value, the file is split into multiple packets and transmitted. This fragmentation leads to a significant number of packets being transmitted within a



(a) Delivery Ratio (b) Delivery Delay

Figure 5.15: Comparison of different use cases

Use Case	Number of resource blocks	Subcarrier spacing (kHz)
3	50	30
4	250	15
Additional Simulation	125	30

Table 5.10: Subcarrier spacing and resource block usage for image data transmission

short period, increasing network congestion. Image and video files are large in size compared to the coordinate data, creating the need for fragmentation. Consequently, network congestion rises and the packet loss and delay are high when transmitting image or video files compared to coordinates. Despite this, according to results, the packet loss for images remains relatively minor, while for videos, a noticeable increase in packet loss is observed. Nonetheless, missing a few frames in a video does not significantly impact its playback, as it remains functional even at a lower frame rate. A slight reduction in smoothness can be expected, but the overall content and quality of the video typically remain intact.

Parameter-wise, the first analysis is done on channel bandwidth values when transmitting different types of data in use cases 1 to 4. Given the high costs associated with hardware for higher bandwidths, this analysis aims to explore the feasibility of using a lower bandwidth. For coordinate data (in use cases 1 and 2), adjusting the channel bandwidth from 5 MHz to 20 MHz did not lead to significant changes in delivery ratio or delivery delay. However, in use cases 3 and 4, which involved transmitting images, a notable improvement in delivery delay is seen with the increase in bandwidth. The bandwidth is influenced by both the number of resource blocks and the subcarrier spacing. Bandwidth in MHz is defined as Eqn. 5.18 where μ stands for numerology index. It is defined according to the subcarrier spacing as elaborated before in Fig. 3.6.

$$B = 2^{\mu} \times \frac{NumberofResourceBlocks}{5} \tag{5.18}$$

Tbl. 5.10 exhibits the applied number of resource blocks and subcarrier spacing for use cases 3 and 4. In addition to entirely increasing the bandwidth, it is noteworthy to gain insights into the effect of applying a wider subcarrier spacing on the communication system performance. For this reason, an additional simulation is done using image data parameters by applying a 30 kHz sub-carrier spacing and 125 resource blocks as defined in Tbl. 5.10 under the *Additional Simulation* use case. Here, the same 50 MHz bandwidth is maintained along with other general parameters but with a higher sub-carrier

Data Type	Bandwidth (MHz)	Period (s)
Coordinates only	5	2
Coordinate and Image	50 (Subcarrier spacing 30 kHz)	uniform (2,5)
All data types	100	uniform (15,18)

Table 5.11: Recommendations for various data transmissions in UAM communications

spacing and a reduced number of resource blocks appropriate to the bandwidth calculation. The obtained results are as follows.

- Delivery Ratio: 72.5%
- Delivery Delay: 55 ms

Compared to the results from Fig. 5.15(a) of use case 4, the delivery ratio shows a reduction. Nevertheless, the delivery delay values show an improvement, reaching 55 ms as opposed to the 73.4 ms obtained earlier (Fig. 5.15(b). This result clearly demonstrates the gain in latency by applying a wider subcarrier spacing. Based on these findings, even a low bandwidth of 5 MHz provides adequate performance for coordinates, while images require a higher bandwidth of at least 50 MHz (with a higher subcarrier spacing) to meet the application's performance needs.

In use cases 5 and 6, the parameter studied is the period. Here, the impact of the time interval between consecutive video file transmissions is examined. The channel bandwidth is set to 100 MHz, given the file sizes are in the MB range. For use case 5, representing a shorter interval, the period ranged between 2 to 5 seconds, while for use case 6, it ranged between 15 to 18 seconds, demonstrating a longer interval. As illustrated in Fig. 5.15, increasing the period positively influences the communication system, resulting in improved delivery delay.

Recommendations

It is clear that both bandwidth and the period significantly influence the delivery delay. Conversely, the delivery ratio does not appear to benefit from changes in these parameters. However, in every use case, a delivery ratio in the range of 80% - 90% is seen. The delivery ratio primarily depends on network congestion, which is influenced by the number of packets and losses in the wireless channel. To further improve this, the TCP protocol can be applied with retransmission techniques. Overall, to ensure the system can handle all three data types efficiently, a high subcarrier spacing (e.g., 30 kHz) with a high bandwidth (e.g., 100 MHz) and a longer period is recommended. If the data consists only of coordinates and images, a shorter period with a moderate bandwidth (e.g., 50 MHz) suffices. For systems exclusively transmitting coordinate data, hardware costs can be reduced further by utilizing a very low bandwidth, such as 5 MHz. This is summarized in Tbl. 5.11.

5.5 Conclusion

Time-sensitive applications, such as UAM, demand reliable and low-latent communication systems. However, directly applying the ground base station networks to facilitate UAM communications is not suitable due to the differences in the requirements. UAM deals with aerial vehicles occupying the low-altitude air space that traverse at varying altitudes and speeds. This nature of movement significantly affects the LOS conditions and propagation characteristics compared to ground-based users. The existing terrestrial networks are aimed at ground users, whereas to cater to UAM communications, much larger coverage areas and more resilient connections need to be taken into account to address the rapid movement and environmental changes.

Within this context, the design of the terrestrial network architecture plays a critical role. It serves as the foundation for effective and reliable data exchange. As a matter of fact, the influential factors related to the ground network plan must be recognized. Following that, a specifically tailored gNB network plan can be realized to facilitate UAM communications with expected performance. One additional factor to be considered is the cost-effectiveness of the system because dense deployments of gNBs are expensive. Accordingly, this overall challenge can be presented as a multi-objective optimization problem and solved using linear or quadratic optimization techniques. However, such methodologies are computationally expensive to implement. One possible solution in this situation is to propose heuristics and improve the involved metrics. To this end, this study explored the impact of various factors on the overall network performance, including the number and placement of gNBs, as well as other system parameters. Based on these insights, heuristics were proposed, leading to a step-by-step procedure for realizing an efficient and cost-effective gNB network plan.

Another important feature for UAM communications is the A2G channel model considered in the aerial UE to ground station communications. The channel models already available for ground users or conventional aerial communications cannot be integrated because they do not account for the unique propagation conditions, mobility patterns, and interference challenges specific to UAM environments. These models are typically optimized for either low-altitude stationary users or high-altitude communications and fail to address the varying altitudes, high speeds, and urban obstacles encountered in UAM scenarios. Accordingly, a specifically tailored A2G channel model is identified as a requirement, and a design based on 3GPP TR 36.777, TR 38.901, and [AT19] is proposed through this work.

As an application of the proposed network plan and the channel model, simulations were done focusing on the Hamburg Metropolitan region. Following the proposed network planning procedure, a gNB network plan with 38 gNBs is proposed to suit a UAM implementation in Hamburg. The simulation findings indicate that gNB positioning must follow geographical conditions as well as traffic demand in order to reach the trade-off between the required number of base stations and the system performance. Nevertheless, in terms of communication latency, investigations into extreme scenarios with rapid air taxi movements are recommended, as the achieved average delay of 60 ms is only marginally adequate for the application.

For the purpose of evaluating the proposed network plan for Hamburg, specific use cases were simulated. The use cases addressed different data types associated with UAM communications, such as coordinates, images, and video. The main purpose here was to examine the influence of channel bandwidth and period when transmitting different data types. Based on these results, recommendations were made to suit various data transmissions. These results revealed that additional means to mitigate data losses in video transmissions are needed. Overall, the proposed gNB network configuration is an applicable solution for implementing a UAM communication system in Hamburg, validated through results demonstrating performance enhancements. Further optimizations through load sharing and increased gNB deployment are possible, provided they remain financially viable. This approach underscores that by paying careful attention to key factors influencing UAM communications, radio network planning can be realized, leading to enhanced UAM communication performance.

Chapter 6

Fault-Tolerant Communication for Urban Air Mobility

The complexity of the autonomous UAM system is highlighted in Chap. 2, particularly in ensuring that specific requirements are met to maintain system safety. The system in this context refers to an urban environment showcasing simultaneous occupancy of the airspace by multiple air taxis. In terms of safety assurance, air taxi navigation in planned flight routes is expected to maintain the required separation margins.

6.1 Overview

In this chapter, with a specific focus on UAM communication networks, the scope of safety assurance in autonomous air taxis is addressed. It is to be noted that only the safety sector is explored here and not the communication security domain. While security is highly relevant in such a critical application, as there are encryption and well-established technologies to deal with cyber-attacks and other possible security-related issues, this is out of the scope of the discussion in this chapter. With regard to safety assurance, the identified problem related to communication is making the communication system redundant to enable the maximum possible communication link availability. To this end, the chapter first sets the background and presents an overview of the related work done in the field. This is then followed by delving into the core concepts related to the topic in discussion. The chapter then proceeds to introduce the specific contributions of this work, including

Significant extension of previously published "S. T. Wanniarachchi and V. Turau, "A Fault-Tolerant Distributed Air-to-Ground Communication Architecture for Urban Air Mobility," 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT), Pafos, Cyprus, 2023, pp. 639-646, doi: 10.1109/DCOSS-IoT58021.2023.00102." [WT23a] a fault-tolerant communication system with a role delegation mechanism to ensure communication availability, leading to safe UAM operations.

6.1.1 Motivation

When discussing autonomous systems, the first application that often comes to mind is self-driving vehicles. The developments in this field are already covered in Chap. 2, and it was acknowledged that this is still an emerging trend. Compared to that, being fully autonomous, traversing the airspace, and intending to carry passengers, the safety requirements in UAM are many folds higher. This makes it crucial to implement strict safety protocols and real-time monitoring systems to ensure safe navigation and mitigate any potential risks associated with autonomous flight operations.

For instance, a higher level of air traffic management is expected in airspace occupancy due to no-flight zones and limited flight routes. Air traffic management should also ensure that user demand requirements are met while planning flight trajectories. A key focus is on managing deconfliction along scheduled flight routes to ensure safe and efficient airspace usage. On the other hand, as a fully autonomous system, effective information management is essential to ensure safe navigation maintaining expected separation distance between the air taxis. This further extends to facilitating coordination with ground control centers for additional information provisions such as emergency situations, weather conditions, unsafe routes, etc. Additionally, automated control systems with built-in redundancy and fault-tolerant mechanisms are required to ensure continued safe operation in the event of a malfunction. As such, it is apparent that the realization of safe navigation in autonomous UAM entails a diverse set of contributing factors, namely:

- Airspace management
- Information management
- Coordination with the ground control station
- Vehicle dynamics optimization
- Automated control systems
- Evaluation of socio-economic feasibility

Each of these factors plays a critical role in ensuring the comprehensive development and operational success of UAM systems. As a direct approach, each factor can be addressed and enhanced individually. However, there exist inter-dependencies between these factors, which cannot be ignored or treated in isolation. To this end, effectively managing these factors in an integrated approach is essential in creating a resilient and well-coordinated UAM system that meets both operational demands and safety standards.

Air Traffic Management (ATM)

In the exploration of mechanisms to manage the factors in safe UAM operations, reviewing similar concepts used in conventional air traffic is a first step. In terms of conventional air transport, the standards from the International Civil Aviation Organization (ICAO) [Web07] are applied for the safe management of aircraft in the airspace. As explained in Chap. 2 and demonstrated in Fig. 2.2, ICAO addresses the navigable airspace under seven classes. Of these, the first six: A-F classes are regulated by the ICAO's Air Traffic management (ATM) systems [Int24]. As a concept, ATM evolved over time with the development of air traffic control systems. The ATMs are responsible for managing the air traffic within the airspaces, ensuring that aircraft can travel safely and efficiently across them. The key components of an ATM system include:

- Air traffic control
- Airspace management
- Flight planning and route management
- Air traffic flow management
- Surveillance systems
- Automation and decision support systems

Accordingly, it is evident that the main elements in ATM systems clearly align with the factors previously identified in UAM systems as essential for safe navigation. Nevertheless, ATM has traditionally been designed for manned aircraft and is primarily intended to manage longer flight routes, including intercontinental air traffic. That is, the ATM systems and protocols focus on high-altitude, long-duration flights across long distances. Therefore, they cannot directly be applied to UAM for several reasons related to the unique characteristics of UAM operations, including:

- Different airspace usage - UAM occupies the 7th airspace class, which is class G, occupying the low-altitude airspace

- UAM relates to semi or fully autonomous aircraft
- UAM concerns a high volume of short-range air taxis that traverse within urban environments
- Differences in used infrastructure - UAM uses vertical take-off and landing (VTOL) aircraft and lands on vertiports as opposed to conventional air crafts designed for bigger airports and runaways

Unmanned aircraft system traffic management (UTM)

Due to such differences limiting the applicability of ATM into UAM, the Federal Aviation Administration (FAA), in collaboration with the National Aeronautics and Space Administration (NASA) and other private contributors, have developed the concepts of *Unmanned aircraft system Transport Management* (UTM) to handle the challenging task of managing the UAM system [Fed24]. UTM manages the challenges of operating autonomous and semi-autonomous aircraft in complex, low-altitude, and high-density airspace, particularly in urban environments. That is, the UTM concept comprises all the systems that enable an air taxi to depart from its origin, navigate through the airspace, and arrive safely at its destination. It is to be noted that UTM heavily relies on automation and communication networks to enable the safe operation of UAM. Other key characteristics of UTM include:

- Low altitude operations
- Real-time coordination
- Integration with ground systems such as vertiports
- Scalability to address a dense distribution of air taxis

With the rapidly evolving technologies related to UAM, UTM is a topic that has been widely addressed and discussed in the literature. [Lab21] is one such example, and the author explores UTM and regards it as a network of interconnected components that operates by enabling coordinated interaction among individuals, data, tools, and resources. The integration of various components such as air, terrestrial, and communication systems and protocols is identified as the key requirement in UTM. It is further highlighted that the primary objective of a UTM system is to enable the exchange of data and information while safely managing and overseeing UAM with different features and ensuring their seamless integration with other airspace users. Accordingly, the goal of the broad range of aspects considered under the topic of UTM can be simply narrowed down to managing the UAM airspace both strategically and tactically. To this end, the two branches of airspace management can be defined as:

- Strategic Airspace Management (SAM) - Efficient allocation of airspace to users through strategic planning and segmentation

- Tactical Airspace Management (TAM) - Facilitating air taxi separation and collision avoidance mechanisms supported by robust information management.

To achieve SAM in the UAM system, several key aspects must be considered. A few examples are:

- Trajectory planning and flight scheduling

- Traffic density analysis and forecasting

- Conflict detection and resolution systems

- Defining safety buffer zones and separation standards

On the other hand, in terms of satisfying TAM in an UAM environment, the following approaches can be applied :

- Resilient information management

- Sense-and-avoid systems for robust collision avoidance

- Autonomous navigation and guidance systems

- Environmental sensing and weather adaptation

- Emergency response and contingency planning

UTM Functionalities

Based on the SAM and TAM aspects stated above, it is apparent that handling TAM and SAM together can fulfill UTM for safe UAM operations. It is further clear that UTM is an umbrella term that comprises several functions that contribute to a successful UTM. To streamline understanding, these functions can broadly be categorized and addressed under three major safety-related functional areas. Drawn from industry proposals and literature, [Lab21] outlines these three distinct subdivisions of UTM functionalities:

- SCF - Safety Critical Functions: Functions that lead to severe damages if there are disruptions or errors.

- SRF - Safety Related Functions: Functions that still play a major role in ensuring safety but do not cause any major destruction in case of failure.

- OSF - Operational Support Functions: Information systems and tools for safe implementation of UAM missions.

The difference between SCF and SRF can be explained with the following examples. An instance of a failure in a SCF is when the localization data of neighboring air taxis is unavailable at an air taxi. This disrupts the control and navigation decision-making and can cause critical effects, even leading up to the worst-case scenario of air taxi collisions. In contrast, a failure in a SRF is an inaccuracy in a planned trajectory. Due to these errors in the trajectory, the air taxi could end up in a no-fly zone. When the resulting effects from the two examples are compared, it is seen that, while the latter does not have immediate consequences, the former poses a significant risk, as the unavailability of neighbor data can result in catastrophic outcomes. Therefore, it is clear that preventing SCF failures is paramount, particularly in passenger transportation. On the other hand, the third functionality, OSF, refers to the additional data provided to the system, including weather updates, maps, coverage details, and other relevant data supporting the navigation. Based on this, it is apparent that SCFs are important for safe UAM operations, making it the most critical UTM functionality out of the three. For the considered case of UAM communications, two major SCFs can be identified:

- The exchange of sensor data: This involves collecting localization-related data from the onboard sensors and communicating this between the air taxis. This is essential for collision avoidance, and trajectory re-planning in UAM.

- Flight Information Management (FIM): This provides important information such as planned trajectories, air routes, vertiport locations, and vehicle types to the air taxi before the mission begins.

While operations can commence based on information from FIM, responding to unexpected situations requires rapid decision-making. In a fully autonomous context, this can be achieved through the information exchanged between the air taxis. To aid with this, real-time sensor data collection and processing inside the air taxis must be done. For instance, in an emergency, the air taxi may need to deviate from its planned trajectory as part of a disaster recovery strategy. These decisions rely on real-time localization data. This includes the air taxi's position (latitude, longitude, altitude), velocity, and acceleration, collectively referred to as the *locally sensed environment* (LSE). Consequently, the SCF of sensor data management plays a key role in decision-making related to control and segregation, ensuring the safe maneuvering of an autonomous air taxi. This concludes that in comparison, while both SCFs are extremely important for UAM operations, the sensor data exchange is a significant challenge for safety assurance in UAM.

6.1.2 Contributions

This work is dedicated to addressing the challenge of sensor data management SCF for UTM. As UTM is an umbrella term for functions of both SAM and TAM sectors, attention must be paid to both. However, what is more relevant to the challenge highlighted here is TAM. This is because the fundamental module influencing the sensor data exchange in an UAM system is communication. The prompt availability of sensor data relies heavily on a robust and reliable communication system. Aligning with this concern, TAM's focus on resilient information management can be correlated with the SCF of sensor data management. This correlation helps narrow down a solution to address the identified SCF. For this purpose, various methods for resilient information management are first explored, and a solution proposal is tailored. Some examples of resilient information management strategies include:

- Redundant communication channel establishment
- Decentralized data storage and cloud integration
- Fault-Tolerant data management systems
- Data quality monitoring and validation

While acknowledging the relevance of each method in achieving resilient information management, due to the complexity of analyzing all aspects together, as a first step, this chapter only proposes a solution based on one of the above-mentioned aspects. In general, faults are unavoidable in any intricate system. Therefore, it is essential to implement strategies that can mitigate these faults while still maintaining an acceptable level of service. Therefore, in an UAM context, ensuring fault tolerance in the components responsible for facilitating sensor data exchange becomes imperative. In other words, this underscores the importance of developing a fault-tolerant communication system for UAM. This not only enhances communication availability but also safeguards UAM operations. By prioritizing fault tolerance, the UAM system can maintain continuous operations and adaptability in case of potential disruptions, ultimately ensuring safer and more efficient air travel for passengers.

For this reason, out of the different strategies mentioned for achieving resilient information management for UAM, providing redundancy to the communication system is selected for this dissertation. While acknowledging redundancy is only one way to provide safety, a redundant communication provision is presented in this chapter as a solution for the challenge of sensor data management SCF. Moreover, redundant communication has already been identified as an open research direction in UAM communications research under the Sect. 3.5, making it further relevant to be discussed in this context. To summarize:

- **Challenge Addressed:** The safety-critical function of sensor data management in UAM
- **Approach:** Focusing on *Resilient Information Management* under tactical airspace management sector achieved through *Redundant Communication Provision*
- **Solution:**
  1. The introduction of a concept for a distributed ground station system architecture to enable Air-to-Ground (A2G) communication to aid with redundancy provision for Air-to-Air (A2A) communication system
  2. A role delegation mechanism for the proposed distributed ground station architecture to ensure fault-tolerance

6.2 Related Work

Fault tolerance is a critical and widely discussed topic in relation to UAM. But the core technical component addressed in terms of fault tolerance or redundancy provision is the navigation and control aspects in UAM [XCL24, CCL15, XZ24]. It is apparent how crucial collision avoidance is for UAM, and the main contributor in this area is control and navigation. Although information management plays a vital role in supporting control decision-making, the literature lacks work done focusing on redundancy provision for UAM communications. This, on the one hand, is due to the limited amount of work concerning fully autonomous air taxis for public transportation. The research community is still focused majorly on UAVs, and thus, it revolves around piloted or remotely operated aerial vehicles, in which the communication functionalities are not as dominant as in a fully autonomous scenario.

In terms of communication in general, redundancy provision is addressed with reference to routing protocols to enhance data delivery. One of the major techniques employed is the transmitting node generating a set number of duplicates of a data packet and sending them along multiple paths to reach the destination. This approach is known as multicopy multipath (MCMP) routing [HWYL17]. This method contributes to improving data delivery. At the same time, the authors highlight the associated issues with the MCMP, such as high overhead and buffer overflows leading to performance degradations if the redundancy provision isn't properly controlled. Another redundancy provision mechanism discussed is the dynamically generated redundancy based on the current needs assessed by individual nodes. This adaptive method is preferred over a static one; however, optimizing it is challenging due to the inherent difficulty of calculating data delivery probability when redundant transmissions are correlated. To this end, a single-copy multipath (SCMP) routing method, which is referred to as a virtual redundancy scheme, has been introduced. The system is flexible to accommodate multiple paths, but a node will receive and forward a data packet only once, reducing the overhead and possible buffer overflows. [JJ24] is an example where these approaches for redundancy provision in routing are tested for a vehicular communication scenario where the vehicle-to-everything (V2X) communication performance is enhanced by employing UAV-assisted device-to-device (D2D) technology.

Another research direction in this scope is redundancy provision for radio access networks (RAN). Work has been done since UMTS technology [VVSC02], on improving redundancy in the RAN or focusing on the base station. The developments in the RAN architecture were since carried on with redundancy enhancements, and in Release 16 of the specification [3GP22] also, this issue is addressed. Multiple proposals have been made for duplicating the user plane between the RAN and the 5G core network. When RANs are virtualized in small functional units, it becomes easy to evaluate and respond to each individual failure. Going in the same direction, [WSK+23] introduces redundant network functions. This is a strategy to improve the redundancy of RAN by utilizing partially ready redundant network functions that can swiftly take over when a primary function fails. This is supported by the Open Radio Access Network architecture [All24], where the RAN is subdivided into a central unit, distributed unit, and radio unit, which are virtual network functions.

As such, redundancy provisions and managing energy efficiency in such systems when redundant components are active have been discussed in the context of communication research. In terms of UAM or UAV communication research, ultra-reliability assurance has been the closest topic to the current topic of interest that has gained the attention of the research community. Although system resilience in terms of introducing spare components is out of the scope of such work, methods for assuring system reliability without being disrupted by faults are suggested. One such example is the work of [YJS+18]. The authors examine the challenges in achieving highly reliable communication for UAV swarms, comparing 3GPP and alternative communication technologies across both urban and rural settings. For urban environments, a remotely controlled system using 3GPP-based technologies is considered, whereas rural areas rely on a ground control station that employs long-range (LoRa) with minimal bandwidth and Wi-Fi with limited range. The paper concludes that the usage of LoRa communication improves reliability in dense networks over long coverage areas, while LTE provides high reliability with enhanced connectivity based on experiments performed on their UAV-swarming platform. The results are promising for a better understanding of UAV communication within urban areas. Nevertheless, it does not address 5G communication protocols or redundancy measures.

Among other research, [MBT17] explores a crowd surveillance application using a UAV-based IoT system. This application is similar to the UAM use case under discussion, as both focus on real-time data offloading from UAVs to a central server. Furthermore, it includes UAVs equipped with various sensors, such as cameras and GPS. However, their work does not address critical factors like sensor data exchange, safety-related elements, or redundancy assurance. In a UAM or UAV system with data offloading to a central server, which acts as a control station, in the classical UAV approach, the control center is handled by a human controller, and thus, redundancy provision is not a critical issue. Nevertheless, in an autonomous UAM scenario, where the air taxi is carrying passengers, enabling redundancy components to ensure consistent system availability is prominent. To achieve this, the communication system must be made fault-tolerant due to the communication between these components being the backbone that guarantees safe maneuvering in an autonomous UAM scenario. That is, the message transmissions must be made redundant.

6.3 Core Concepts

This work focuses on solving the challenge of sensor data management for safe UAM. The solution proposal entails looking into resilient information management as a TAM measure with a specific focus on redundant communication provision. It was clear from the literature survey that the redundant communication provision in relation to autonomous UAM communications is an overlooked area. Consequently, this topic is addressed in this chapter, and this section outlines the core concepts underlying the topic. That is, the challenge, solution approach, and background related to the solution are elaborated in this section.

6.3.1 Sensor Data Management

The challenge examined is the SCF of sensor data exchange. Therefore, in this subsection, the methodology employed in handling sensor data in the UAM system is presented. A method inspired by the multi-access edge computing (MEC) [ETS14] framework developed by the European Telecommunication Standards Institute (ETSI) is applied for the sensor data handling. Following this approach, safety-critical data is continuously gathered and shared among air taxis within the network. Each air taxi is equipped with a set of sensors consisting of GPS, lidar, and cameras, etc. It is to be noted that sensor data collection is also a process operating in a fault-tolerant manner (e.g., through active replication) to ensure consistent data collection. The sensor data is then processed into a background map, creating a representation LSE of the air taxi. MEC's concept is mirrored in this approach, as data collection and processing are performed directly on the air taxi at the network edge rather than on a centralized server. The resulting map, termed the *Locally Sensed Environment Picture* (LSEP), is similar to the *Local Dynamic Map* (LDM) of Intelligent Transport System (ITS) framework defined in [ETS10]. The situational awareness among neighboring air taxis is then enabled through LSEP exchange, consistent with the concept of *Collective Perception Messages* of ETSI standards [ETS19].

6.3.2 Resilient Information Management

In this subsection, an approach leading from resilient information management up to highlighting the relevance of redundancy provision is explored. As evident from the analysis up to this point, the communication between air taxis is a key influential factor for the sensor data management of the UAM network. This initiated the need to address the challenge of sensor data management, specifically aimed at enhancing information management. For instance, LSEP exchange is the source providing awareness on the surroundings of the air taxis. In case of a missing LSEP, the control systems inside the autonomous air taxi lack information for decision-making. This could lead to deviations from the planned trajectory, which can be a severe hazard. This necessitates the need to enable regular and timely LSEP exchange between air taxis, making strong coordination between air taxis essential.

Advanced 5G technologies supporting high bandwidth, low latency, and improved network capacity [BDO+21] can be considered as an option for a suitable communication technology to explore. Particularly, this is enabled through network-assisted D2D communication or cellular vehicle-to-everything [3GP22], which provides reliable direct links between the air taxis. In this mode of communication, the air taxis directly interact with each other without the intervention of the base station, which helps reduce latency, improve network coverage, enhance spectral efficiency, etc. This is referred to as Air-to-Air communication in this chapter and is denoted by A2A. Ideally, A2A links would exclusively handle communication in an autonomous UAM scenario. However, due to potential factors like LOS loss, high mobility of air taxis, synchronization issues, interference from nearby air taxis, resource allocation issues, or adverse weather, A2A communication link failures can occur. As described before, an A2A connection loss prevents the LSEP exchange between air taxis, resulting in severe after-effects. Thus, it is clear that managing the availability of A2A links is important for safe autonomous UAM operations.

6.3.3 Redundancy

The above discussion highlighted the need for redundancy provision in A2A communication. Redundancy refers to enabling measures that enhance the dependability of modules in a system. When a system is equipped with redundant components, even during failures or malfunctions, the system can still continue operations maintaining reliability. This strategy ensures the maximum possible availability of the system functions, reinforcing the overall resilience and stability. The process of facilitating and managing redundancy within a system is referred to as *Fault-Tolerance*. Simply put, fault tolerance implies making a system dependable. This dependability is assessed by two metrics: *reliability* referring to the probability R(t) that the system remains operational throughout the time interval [0,t] and *availability*, the time segment over which the system is operational within a specified time frame. Addressing redundancy can be done in several forms [KK07]. Namely:

- Hardware Redundancy - Using additional hardware for the purpose of detecting and compensating failures
- Software Redundancy - Producing a secondary version of the software to be used in case of a failure in the primary version

- Information Redundancy - Adding check bits to the original data bits as error detection and correction codes
- Time Redundancy - Re-running the same application on the same hardware

When it comes to communication, as explained in the Sect. 6.2, various aspects of redundancy provisions can be facilitated. For example, redundancy in routing protocols, inclusion of redundant bits, implementation of dual communication channels, backup communication hardware component implementation, and redundancy in RAN protocols, etc., can be taken into account. It is to be noted that routing protocols and redundant bit inclusion solutions have already gained attention in the research community, although they have not been applied to UAM scenarios. For this reason, rather than revisiting already established wired and wireless communication protocols to realize reliable communication establishment, looking into backup communication channels and modes is intended here. Drawing from that, out of the identified types of redundancy, the type most relevant to the UAM communication system is hardware redundancy. This involves incorporating duplicate hardware components, such as backup communication devices. This way, if one component fails, another can immediately take over, minimizing the risk of communication breakdown. Hardware redundancy can further be addressed under two categories[KK07] :

- Static Hardware Redundancy - Having multiple components executing the same function
- Dynamic Hardware Redundancy - Having backup components that activate in the event of a failure

Based on this categorization, the solution discussed in this chapter incorporates a static hardware redundancy provision method by introducing a redundant communication medium for inter-air taxi communication. That is a redundant mode of communication system, such that the communication links remain operational with a high probability. Simply put, it is intended to provide a copy of the same information communicated through the A2A communication system through a backup communication system. To this end, the introduced solution proposes integrating an Air-to-Ground (A2G) communication setup into the UAM communication paradigm. Here, a periodic data transfer occurs between air taxis and a ground station, ensuring ongoing information management within the network as elaborated in Fig. 6.1.

In this approach, the LSEP from each air taxi is periodically sent to a ground station, creating a centralized, complete situational overview termed

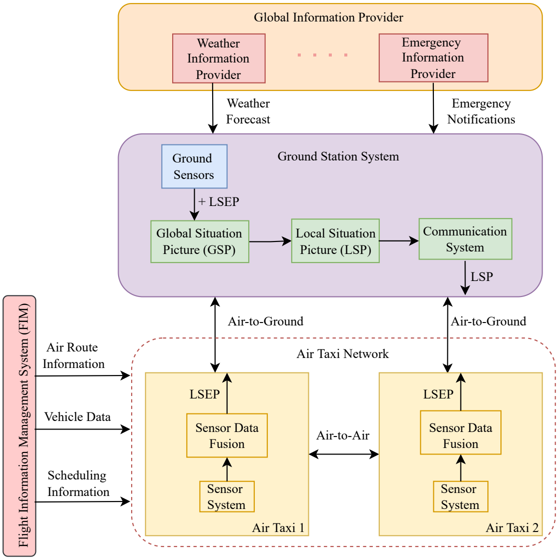


Figure 6.1: General System Architecture

the *Global Situation Picture* (GSP), similar to data aggregation by roadside units in ITS [ETS19]. The GSP results from fusing ground sensor data with LSEPs from all air taxis. Primarily, air taxis only need the data relevant to their immediate surroundings. This makes the distribution of the full GSP to each air taxi inefficient and bandwidth-intensive, specially considering the high frequency at which GSPs are generated. To this end, a map local to each air taxi referred to as the *Local Situation Picture* (LSP) is projected from the GSP. This LSP is then transmitted back to the air taxis periodically, offering each vehicle a reliable view of the surroundings local to it. Furthermore, having a ground station setup also supports remote monitoring of the air taxi network, enhancing system safety, as noted in prior research [NCSG18].

An important distinction highlighted in fault-tolerant systems research is the difference between the terms *fault* and *error*. Failures or faults are the defects that are identified in hardware components. In contrast, errors are the resulting effect of these faults. For instance, the false positive or duplicate signal reception is an error resulting from a system facing a failure. In the presented work, addressing the detection of errors and other types of redundancies, including software redundancy, is out of scope. Additionally, the system includes a FIM component to satisfy the second SCF identified earlier in the chapter. The FIM supplies static data on air routes, schedules, and vehicle information, although this is outside the scope of the present work.

6.3.4 Ground Station Network

When discussing the ground station network, it is essential to distinguish that the concept of a ground station is not the same as a typical base station in a communication network. While a base station functions primarily as a communication relay, the ground station is an integrated hardware-software component designed to enhance communication system reliability by ensuring redundancy. The ground station is, hence, not a communication component, but it contains communication modules to initiate communication with air taxis. The purpose and functions of a ground station were previously described in subsection 2.3.3. This design flexibility allows the ground station setup to operate alongside any communication infrastructure used for air taxis. Thus, irrespective of the ground station's architecture, the number of base stations varies depending on the communication standard in use. For clarity, base stations are not depicted in this section's figures.

The core objective here is implementing a ground station to enhance fault tolerance within the UAM communication framework. Accordingly, the most straightforward design employs a centralized ground station setup, as shown in Fig. 6.2. Only one ground station is present, and this single ground

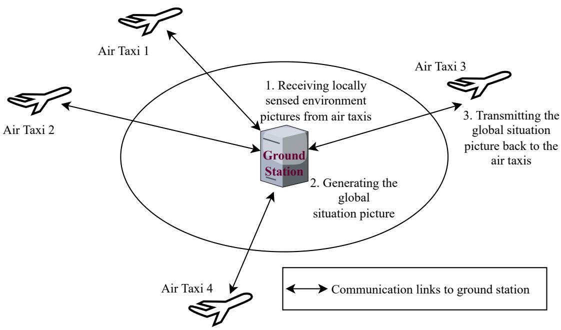


Figure 6.2: First option - Ground station architecture with a single ground station

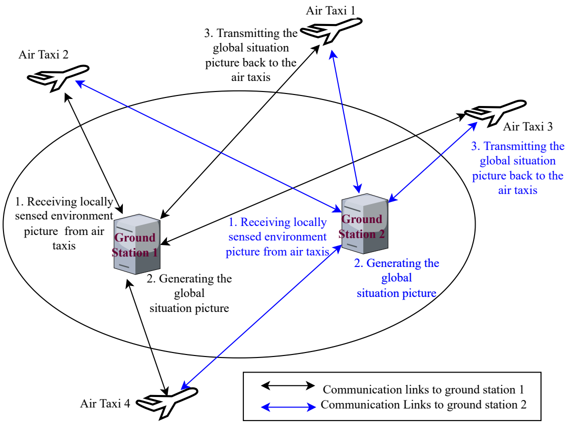


Figure 6.3: Second option - Ground station architecture with two ground stations

station aggregates the LSEP data from all air taxis (using the infrastructure communication through base stations), processes it, and generates a GSP. Then, the LSPs are projected from this GSP and sent back to each air taxi. Although it is very easy to implement this, being a simple concept, the design has several limitations. First, handling individual LSP processing within a single ground station demands substantial computational power, as it must aggregate sensor data and generate separate LSPs for each air taxi in a large UAM network. Managing this frequent process through a single ground station risks network and system overload. On account of that, applying the MEC concept appears to be more beneficial for distributed load handling in the UAM communication system. Moreover, while the ground station introduces redundancy to the communication framework, having only a single ground station leads to a single point of failure (SPF) risk in the ground network, impacting system *availability*. Availability is determined by metrics of *mean time to repair* (MTTR) and *mean time to failure* (MTTF). In general, the MTTR is a high value [KK07]. This means having a centralized architecture disrupts the system availability, compromising the information management in the UAM system. Therefore, the centralized ground station architecture is not a viable solution.

To address these limitations, a distributed ground station architecture is devised, as shown in Fig. 6.3. The figure depicts an example scenario with two ground stations. In this setup, each ground station independently collects LSEP data from the air taxis, processes it to generate the GSP, projects, and transmits the LSPs back to all air taxis. In the event of one ground station failure, air taxis can still rely on the other operational ground stations. This way, static redundancy is provided to the UAM communication system through additional hardware to replace the failed unit. This approach effectively addresses the SPF issue identified in the previous model, enhancing system availability unless all ground stations were to fail simultaneously. However, the likelihood of simultaneous failures diminishes exponentially as the number of ground stations increases, making this a minor concern. Nevertheless, challenges related to bandwidth inefficiency persist due to the repeated transmission of messages, which also causes substantial message overhead. These factors indicate that the second design is also inadequate for establishing a network of ground stations in a UAM system, particularly when considering capacity limitations.

To accommodate the previously mentioned drawbacks, a distributed ground station architecture is proposed. This architecture features multiple ground stations that do not overload the system while still providing redundancy. In this model, the responsibilities for generating the GSP and LSP are assigned to different ground stations. In the event of a ground station failure, its role can be seamlessly transferred to another ground station. This approach reduces the MTTR and enhances overall availability. A detailed discussion of this architecture is presented in the next section, Sect. 6.4. The advantages of implementing a distributed ground station system in a UAM framework include:

- Ensuring availability of ground network
- Ability to distribute service provisions to air taxis
- Reduced computational complexity in the network
- Ability to include area-specific alerts for air taxis

6.4 Ground Station Architecture

This section elaborates on the primary solution proposal in this chapter. That is, proposing a distributed ground station architecture to establish A2G communication as a redundant measure for the A2A UAM communication system. As discussed in the previous sections, the specific requirements to satisfy in the system design are recognized as:

- Redundancy provision to existing A2A communication network
- System not leading to a SPF issue
- A distributed network architecture
- Not overloading the system with an extensive amount of message transmissions

To this end, the discussed approach involves establishing two types of ground stations in the system:

- Main Ground Station (MGS)
- Sub Ground Station (SGS)

The MGS is primarily responsible for creating the GSP. The generation method is discussed later in this chapter. Each SGS has a defined scope of responsibility, which includes a designated set of air taxis under its focus. The methodology for determining this scope is also discussed later in this chapter. Within its area of responsibility, an SGS performs two key tasks:

- Receives the locally sensed environment pictures from the assigned air taxis, similar to the MEC concept
- Generates local situation picture based on the global situation picture received from the MGS and transmits them to the air taxis in its scope of responsibility

The network architecture is expected to contain one MGS and several SGSs distributed throughout the network. The ground area is divided into regions based on city-specific criteria (e.g., city districts), with an SGS set up in each region. This strategy minimizes additional infrastructure costs while enhancing system availability and ensuring economic viability. All ground stations are interconnected via a wired network that utilizes IP protocols, providing high bandwidth. Fig. 6.4 illustrates the proposed architecture.

Fig. 6.5 illustrates the A2G communication model with an air taxi and an SGS as an example. The involved steps are as follows:

1. An air taxi creates the locally sensed environment picture by sensor fusion of locally sensed environment data.
2. The air taxi transmits the locally sensed environment picture to the sub ground station it belongs to.

3. The sub ground station sends the received locally sensed environment pictures to the main ground station.

4. The main ground station generates the global situation picture by the fusion of ground sensor data and locally sensed environment pictures.

5. The main ground station sends the global situation picture to the sub ground stations.

6. Each sub ground station projects the global situation picture into local situation pictures for each air taxi in its scope of responsibility.

7. Sub ground station transmits the local situation pictures to the respective air taxis.

Within each SGS, an LSP is generated for every air taxi within its designated scope of responsibility. The system also allows for the inclusion of area-specific alerts, such as emergency information or, weather forecast information, etc., in the LSP. This LSP is tailored to the local surroundings of the air taxi and generally contains information about all nearby air taxis within a specified radius, similar to the concept of a LDM. This way, upon the reception of the LSP from the ground station, the air taxi can update its control system and perform necessary navigation changes as required. Ideally, this information must also be received individually from the nearby air taxis through the A2A communication system. With the provision of LSP from the

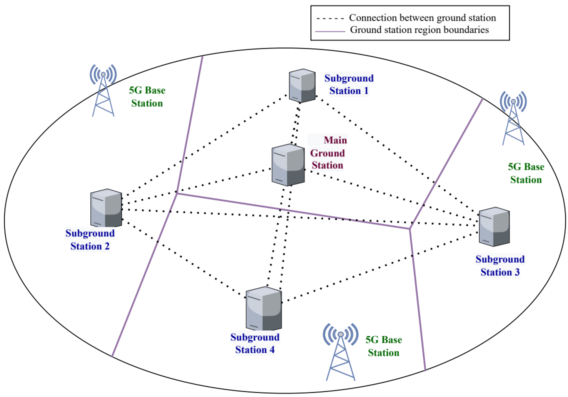


Figure 6.4: Proposed ground station architecture

ground station, further verification of this information received from A2A links is possible. In addition to that, the inclusion of any missing information due to A2A communication link failures is made possible. This leaves out the risk of collisions as a result of insufficient data availability due to issues with the A2A links.

If an air taxi sends an LSEP to a certain SGS, this air taxi is considered as belonging to the scope of responsibility of the specific SGS. Accordingly, this is the principle followed by the SGS to recognize the air taxis belonging to its scope of responsibility. All the air taxis that initiate an LSEP transmission to the SGS are registered by that SGS into the group of air taxis of the SGS's scope of responsibility. To find the nearest SGS, an air taxi must perform a distance calculation based on its position. The operational area for each SGS is anticipated to range from 50 to 100 square kilometers (e.g., a city district). Therefore, once an air taxi enters the SGS's scope of responsibility, it is expected to remain there for several seconds. This leads to the conclusion that the positioning calculations will be conducted less frequently than the message transmission frequency. This prevents the computational complexity inside the air taxi communication system. To assist air taxis in identifying SGSs, the location coordinates of the SGSs are provided as inputs. Additionally, since the air taxis are equipped with powerful processors, potential computational delays that could hinder the real-time provision of information at the SGSs can be mitigated.

6.5 Role Delegation Mechanism

The previous sections highlighted the requirement of an A2G communication system as a redundant communication mode for A2A communication

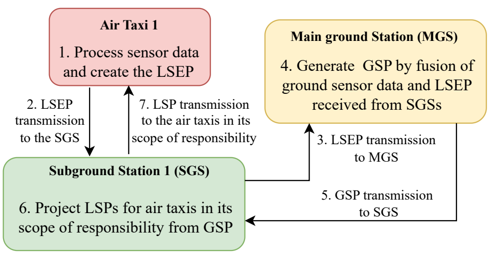


Figure 6.5: Proposed methodology for air-to-ground communication

in UAM. Accordingly, a ground station architecture that enables this A2G communication establishment was presented. It is evident that the main objective is thus satisfied, and a high probability of communication availability is facilitated. Nevertheless, it is to be noted that this doesn't solve the entire challenge of maintaining system availability. This is due to the fact that the ground stations in discussion could also be prone to failures, and this could interrupt the A2G communication setup. This way, in case an A2A communication link also fails, the air taxis will not have a source of information provision, which may lead to catastrophic aftereffects. Therefore, making the ground station system fault-tolerant is regarded as a crucial requirement leading to managing the SCF of sensor data exchange in a UAM system. Consequently, this section outlines a fault-tolerant framework addressing the potential failures of ground stations.

A fault-tolerant system must detect failures and quickly initiate recovery processes to maintain high system availability. The suggested protocol focuses on achieving a low MTTR and facilitating rapid system recovery in the event of a ground station failure. Three types of failures are considered:

- Failure of the Main Ground Station
- Failure of Sub Ground Stations
- Simultaneous failure of both Main Ground Station and Sub Ground Stations

To establish a clear understanding, several key terms are introduced, such as:

- Heart Beat Protocol
- Sub Ground Station Priority List (SGSP list)

Heartbeat Protocol (HBP)

A heartbeat protocol is a mechanism used in distributed systems to monitor the health and availability of nodes or components within a network. It involves sending regular, short *heartbeat* messages between nodes to confirm that they are active and functioning correctly. Here, a heartbeat signal is used between the MGS and SGSs as well as among the SGSs themselves. The heartbeat signal serves as a periodic signal that indicates the operational status of the MGS or SGSs. In this approach, the heartbeat signal functions as a failure detection mechanism.

Sub Ground Station Priority List

A coherent approach to role delegation among the components is crucial for a successful fault-tolerant system. For instance, in this network, all the ground stations need to agree upon which ground station will take up the role of the next MGS in case of a failure. To this end, the underlying problem of choosing a new MGS can be framed as a *consensus problem* in distributed systems, which is defined as applying a commonly agreed-upon rule by all the participating components in attempting to reach an agreement in making a certain collective decision [Lyn96]. With this regard, different algorithms can be applied when making the decision. Some examples are real-time voting by the SGSs, Priority advertisement by each SGS based on a certain predefined metric and choosing the next MGS based on that, Using a predefined sequence of role delegation, proximity-based selection, applying advanced distributed consensus algorithms such as Byzantine Agreement [Lyn96], etc.

Out of these many possibilities, in the presented work, a deterministic approach is applied where a predefined order is used to take over the responsibilities of the MGS. The objective is to minimize the time taken for the role delegation procedure as much as possible. In this regard, applying real-time voting and measurement-based techniques won't be beneficial as they are time-consuming. Therefore, the pre-assigned sequence is applied such that, in case of a failure, the next SGS is automatically selected. To this end, a list known as the sub ground station priority list (SGSP list) is created to determine the order of role delegation from the MGS to the SGSs in the event of MGS failure. Given an infrastructure-based communication system is utilized, the location of the new MGS remains unaffected by factors such as transmission range or distance to the previous MGS. This allows the role delegation priority list to be based on any site-specific criteria. The SGSP list is provided to each SGS at the system launch, indicating each SGS's priority status for taking over the MGS responsibilities. It is to be noted that, the role delegation involved with this proposal is not simply a handover. It is a process that needs to be handled by considering various factors by the SGSs, as the MGS is not functioning at this point. Accordingly, the provision of the SGSP list is essential to prevent multiple SGSs from simultaneously attempting to assume the role of the failed MGS, which could compromise network consistency.

Fault Tolerant Protocol

With the terms being clarified, the fault-tolerant protocol is defined in this subsection. Each failure type is explained with the involved MTTR and the algorithm.

Main Ground Station Failure

When the main ground station fails to transmit the heartbeat signal to the SGSs, this absence is interpreted as a failure. Each SGS waits for a specified time period, known as the *mean time until failure recognition* (MTUFR), during which the heartbeat signal is expected from the MGS. Once a failure is confirmed (Absence of a heartbeat signal for a period of MTUFR), the SGS checks the SGSP list, and the SGS with the highest priority takes over the MGS role, broadcasting a message to indicate its new status. All remaining active SGSs must acknowledge this transition. The MGS waits in expectation of an acknowledgment for a time period, TACK. In case of a missing acknowledgment during this period, one retransmission is done, and if this is also not acknowledged, then the system reports the corresponding SGS as a failed SGS. This new MGS announcement broadcast is also received by the air taxis, leading them to be notified about the status of the ground station network. Additionally, the SGSP list needs to be updated to reflect the current configuration in preparation for future failure scenarios. The new MGS is responsible for updating the SGSP list and distributing it across the ground station network. Each SGS must confirm the reception of the new SGSP list to prevent confusion between the old and updated versions. Here again, the same retransmission procedure as before is applied in case of a missing acknowledgment during the period TACK. This process is elaborated in Alg. 1, and this algorithm is applied to all active SGSs.

The time taken from the initial check of the SGSP list until the new MGS receives acknowledgments from the SGSs is termed the *mean time to recover* (MTTRE). Based on this discussion, the MTTR is defined in Eq. 6.1, indicating that A2G communication is unavailable during this period. The MTTR associated with this role delegation process is negligible compared to the MTTR required for reinstating a failed ground station.

$$MTTR_{MGSFailure} = MTUFR + MTTRE \tag{6.1}$$

Once a new MGS takes on its responsibilities, A2G communication will resume. Based on how the algorithm is designed, at any moment, an SGS has the potential to upgrade its status to MGS. When an SGS assumes the MGS role, it is exempted from its prior SGS duties. As a result, air taxis will no longer consider this SGS's position when selecting which SGS to communicate with. Instead, air taxis will identify the nearest available SGS from the remaining options, and the transmission of LSEP and LSP creation

Algorithm 1 MGS Failure: Role Delegation Procedure

```
Start
STEP 1: MGS failure recognition
wait for the MGS HBP signal for a time period of MTUFR
if HBP signal received then
    continue operations
else
    STEP 2: Role Delegation - Actions during the time period MTTRE
    check SGSP list
    if priority == 1 then
        report new role as MGS via broadcast
        wait for acknowledgment from each SGS for a time period TACK
        if acknowledgment is received then
            continue operations
        else
            STEP 3: Retransmission in case of missing acknowledgment
            resend the message to SGSs that didn't acknowledge
            wait again for the acknowledgment for a period TACK
            if acknowledgment received then
                continue operations
            else
                Report the particular SGS as failed
            end if
        end if
        send updated SGSP list to SGSs
        wait for the SGSP list Acknowledgments from SGSs for a time
        TACK
        if acknowledgment received then
            continue operations
        else
            apply STEP 3
        end if
    else
        Process new MGS announcement
        acknowledge the reception of the message from the new MGS
        acknowledge the reception of the updated SGSP list
    end if
    continue operations
end if
End
```

will proceed with the newly chosen SGS. The overall procedure undertaken by an SGS is illustrated in Fig. 6.7.

Furthermore, it is acknowledged that the new MGS may encounter a failure during the role delegation process. Such occurrences can be tied to the update message of the SGSP list. After the MTUFR has elapsed, the remaining SGSs should receive the SGSP list update from the new MGS. If this update is not received, the next highest priority SGS from the previous list will issue a status request to the new MGS. If a response is not obtained within a designated time frame, the new MGS will be considered failed as well, and the SGS that sent the status request (the next in priority) will take over the MGS role and execute the previously outlined procedures.

Sub Ground Station Failure

The second scenario of consideration involves the failure of SGSs. The MGS receives regular HBP status updates from each SGS. Consequently, after a designated period known as the *mean time to identify failure* (MTTIF), the MGS can detect any failed SGSs if a HBP signal from an SGS is not received. Upon identifying this failure, the MGS broadcasts a notification indicating which SGS has become non-operational. This communication is important for air taxis to locate the nearest active SGS. Additionally, the MGS is responsible for updating the SGSP list as previously outlined. This procedure is presented in Alg. 2 and is applied inside the MGS.

This role delegation can proceed in such a way that the probability of A2G communication availability is very high. However, a challenge arises when multiple sub ground stations fail, increasing the workload on the remaining SGSs and gradually increasing the computational complexity. To manage this, a threshold must be established: the minimum number of active SGSs required to maintain successful A2G communication. This threshold needs to be investigated and validated through experiments. Once this threshold is reached, an active SGS will notify the network about service unavailability until the entire system is restored. In Alg. 2, the threshold check comparison with the number of SGSs is done to achieve this. When multiple ground stations have failed, the system cannot continue operations. This is called system shutdown and is further elaborated later in this chapter.

Regarding the MTTR in this SGS failure situation, the duration required for the MGS to complete both tasks of reporting the failed SGS and updating the SGSP list is referred to as the *mean time to update* (MTTU). Therefore, the MTTR can be represented as Eq. 6.2

$$MTTR_{SGSFailure} = MTTIF + MTTU \tag{6.2}$$

Algorithm 2 SGS Failure

```
Start
STEP 1: SGS Failure Recognition
wait for the SGS HBP signal for a period of MTTIF
if HBP signal received then
    continue operations
else
    STEP 2: Heartbeat not received - Actions for time period MTTU
    if Number of SGSs >= Threshold then
        report SGS failure via broadcast
        update SGSP list
        send updated SGSP list to SGSs
        wait for SGSP acknowledgments for a period TACK
        if acknowledgment received then
            continue operations
        else
            resend the SGSP list
            wait for SGSP list acknowledgments for a period TACK
            if acknowledgment received then
                continue operations
            else
                reports the SGS as failed
                continue operations
            end if
        end if
    else
        announces system shutdown
    end if
end if
End
```

Simultaneous Failure of Main Ground Station and Sub Ground Station

The third type is the simultaneous failure of the MGS along with one or more SGSs. Previously, when only SGSs experienced failures, the MGS handled role delegation. As opposed to that, in this case, the remaining SGSs must identify the failures independently as now even the MGS has failed. Initially, the failure of the MGS is tackled. Based on the reception of the HBP signal from the MGS, the active SGSs can detect MGS failure. In such an event, the SGS with the highest priority in the SGSP list assumes the MGS role and notifies the air taxis and other SGSs according to Alg. 1. At the same time, as each of the SGS is exchanging HBP signals, the simultaneous failure of an SGS is also identified by the same SGS that assumed the duties as the new MGS. In this event, it will also initiate applying Alg. 2, as long as the remaining number of active SGSs lies below the threshold.

To this end, both algorithms are applied simultaneously, and the MTTR is slightly higher than the MTTR of the individual cases. Even still, the MTTR is expected to be only a few milliseconds. Given that the ground station network operates on wired connections, message transmissions for role delegation can be completed within microseconds. Thus, including decision-making and reception of acknowledgments, the MTTR can still be maintained to be a few milliseconds. This, however, needs to be validated experimentally. The MTUFR remains to be the same value here, as periodic HBPs are used, and therefore, during the same time period, both failures can be recognized. The rest of the procedures include informing about the new MGS takeover, reporting the failed SGS, and updating the SGSP list. The time taken to this is denoted here as *mean time for simultaneous update* (MTFSU). This modifies the MTTR equation in the third failure type to be Eq. 6.3. If the failed SGS is the one with the top priority, the next in line takes over, and this process continues.

$$MTTR_{BothMGSandSGSFailure} = MTUFR + MTFSU \tag{6.3}$$

System Shutdown

The possibility of A2G service unavailability was explained earlier under the SGS failure situation. This is an inevitable scenario, as any system is prone to faults. The algorithm makes sure that the probability of running into faults is reduced as much as possible. But 100% assurance cannot be granted. For this reason, an A2G communication system shutdown is a possibility. However, it is assumed that during this time, the A2A communication will continue to function, meaning the brief period during which A2G communication is unavailable will not critically impact the system, although actual values need experimental validation. It should also be noted that the possibility is there for both communication systems to potentially be offline simultaneously. This is the most critical situation in a UAM scenario due to the unavoidable event of failures. Nevertheless, this situation is not covered here.

Reinstating operations

Once the ground stations are operational again, a new update message is broadcasted. Upon receiving this message, the system resets to its default state. If the MGS or any failed SGS is restored before the threshold (the minimum number of SGS required for system continuation) is reached, the restored entity will send a message to inform other ground stations of its status. In instances where an SGS has been reinstated, the MGS broadcasts a message, resets the SGSP list, and operations continue. This also lets the air taxis reset the SGS location list to the default set. If the MGS is restored, after notifying the other ground stations of its recovery, the entire system, including the SGSP list, resets to default and resumes functioning.

6.6 Conclusion and Outlook

In an autonomous UAM scenario, one of the major challenges to handle is the SCF of sensor data exchange. As all decisions must be made based on observing the surrounding environment by the air taxi itself, it is necessary to ensure a seamless sensor data exchange in the UAM system. To realize this, regular message transmissions between the nearby air taxis can be enabled. As the technology provider, 5G D2D communication is a suitable candidate to establish the A2A communication. Nevertheless, faults in any system are inescapable, and thus, the same applies to the communication system. The A2A communication in UAM can be faced by the unavailability of LOS, resource allocation issues, etc., that could hinder continuous information management. Communication link failure is crucial for autonomous UAM scenarios as it is directly related to deteriorating the safety assurance in UAM operations.

While there has been significant research exploring communication aspects in UAM, there is limited work focusing on ensuring fault tolerance. Given the severity of the issue in UAM implementations, this chapter addressed the vulnerabilities in A2A communication and proposed a solution focusing on the TAM sector, specifically aiming at resilient information management achieved through establishing a redundant communication mode. This proposal encompassed the incorporation of an A2G communication system with a distributed ground network architecture. Here, considerations were made not to overload the system with regard to functionalities assigned to the ground station. Furthermore, attention was paid to avoiding the single-point failure of the system. A ground station network architecture applicable to the Hamburg metropolitan region is suggested as an outcome of the work and presented in Fig. 6.6. The city districts in Hamburg are taken as a criterion, and seven

SGSs are proposed so that each city district consists of an SGS. One MGS is planned for the city center, as demonstrated.

It was further identified that simply having an A2G communication network as a redundant communication medium for A2A communication does not solve the challenge of maintaining system availability. The potential for sudden failures in ground stations must also be considered to guarantee system availability. Accordingly, a role delegation concept between these stations is suggested to ensure service continuity. This approach aims to achieve an MTTR that is significantly lower than the MTTR associated with restoring failed communication links. To avoid system overload, it is necessary to determine capacity thresholds and identify the minimum number of ground stations required for safe system operation. However, it is to be noted that, even with such an algorithm, it is only possible to increase the probability of maintaining communication system availability even if certain system components have run into failures. It is not possible to guarantee 100% system availability. To this end, a complete system shutdown for a short period of time is also expected, as there is also a chance of simultaneous faults in many ground stations.

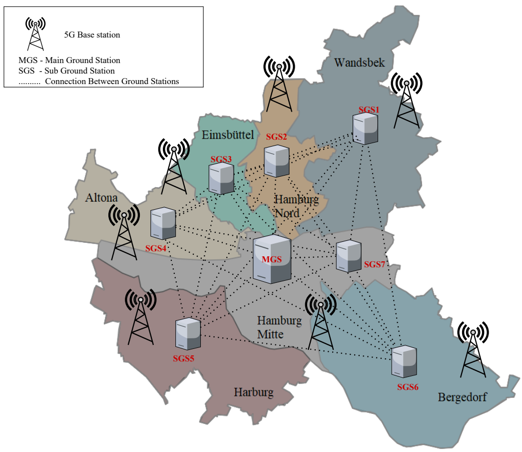


Figure 6.6: Ground station architecture applied to Hamburg metropolitan region

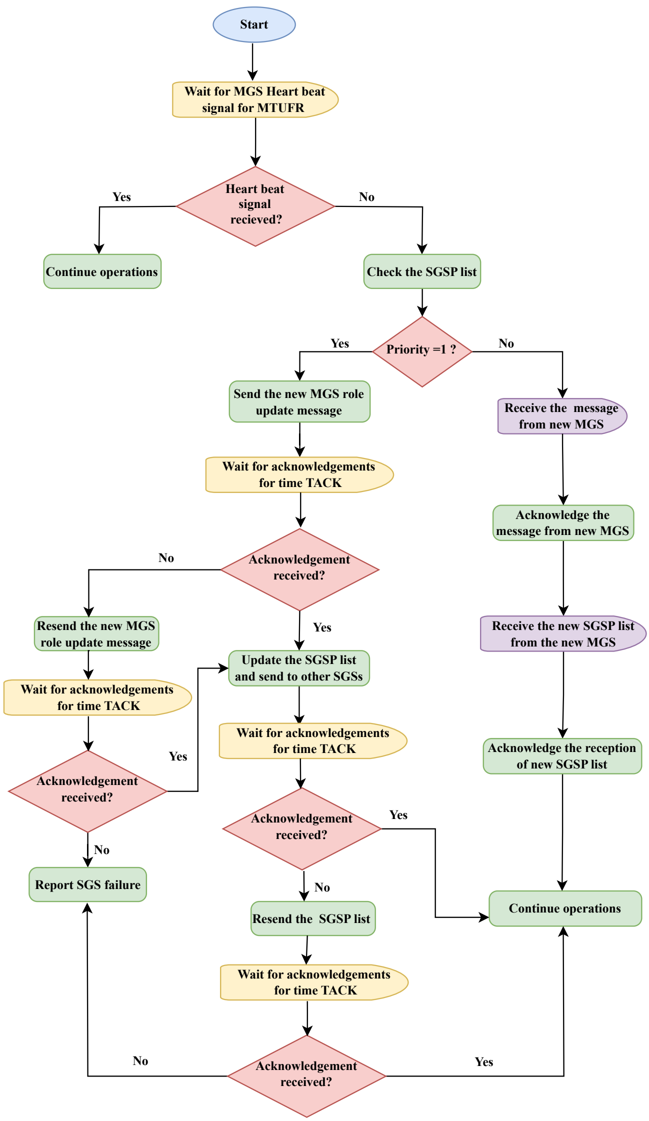


Figure 6.7: Role delegation procedure followed by a sub ground station

Chapter 7

Dynamic Resource Allocation for Air-to-Air Communication

Previously discussed topics have evolved, highlighting the essential role of uninterrupted, reliable, and low-latency communication in the air taxi system for achieving safe, efficient, and coordinated movement within urban environments. This communication framework underpins critical functions such as real-time situational awareness, collision avoidance, and urban air traffic management, all of which are integral to supporting reliability of UAM. When describing the in-flight system requirements of UAM in Sect. 2.4.2, two kinds of communication associated with the UAM system were expressed: Air-to-Air (A2A) and Air-to-Ground (A2G) communication. As UAM operations expand, the demand for robust A2A communication grows, necessitating a system that is resilient, fault-tolerant, and capable of supporting high-frequency data exchange in dense and complex urban air spaces. These considerations establish A2A communication as a fundamental element, ensuring UAM networks can meet the rigorous demands of urban transport infrastructure.

Towards the realization of a reliable A2A communication network for UAM, this chapter explores another direction highlighted in Sect. 3.5 as an open research question. That is the problem of resource allocation. Here, specifically, D2D communication is addressed as the topic evolves around A2A communication. Resource allocation is a vital aspect of D2D communication as it directly impacts the interference of the communication system. To this end, the chapter deals with the physical and medium access control

Extension of previously published "Wanniarachchi, S.T., Turau, V. (2024). Dynamic Resource Allocation for 5G Device-to-Device Communication Based on Expected SARSA. In: Castañeda, A., Enea, C., Gupta, N. (eds) Networked Systems. NETYS 2024. Lecture Notes in Computer Science, vol 14783. Springer, Cham. https://doi.org/10.1007/978-3-031-67321-4_1."[WT24]

layers of the communication protocol stack. It is to be noted that interference management is also an open research direction according to the aforementioned Sect. 3.5, making it viable to be studied further. Effective interference management is essential in networks where D2D communication operates alongside traditional cellular systems. Inadequate control of interference can result in increased latency, lower aggregate data rates, diminished throughput, and reduced spectral efficiency. To this end, in this chapter, resource allocation is explored as a means of handling the interference.

This chapter discusses the background related to resource allocation and interference management in 5G D2D networks. Then, it proceeds to explore existing literature on resource allocation techniques. This is followed by a detailed solution, including a dynamic resource allocation strategy for 5G D2D users, employing the Expected SARSA reinforcement learning algorithm. Simulation results are used to demonstrate that the proposed algorithm significantly enhances system performance in terms of throughput and reliability, illustrating the benefits of applying dynamic resource allocation for managing interference in D2D communications.

7.1 Overview

This section establishes the foundational context related to the chapter. First, the D2D communication technology is introduced. This is then followed by exploring challenges in D2D communication and a discussion of resource allocation strategies.

7.1.1 Device-to-Device Communication

Rapid technological progress and constant innovation are affecting people's needs. As this evolution continues, smart devices and connectivity have become daily requirements. Furthermore, urban areas are experiencing significant growth in population density, calling for a high demand for continuous mobile connectivity. Each year, this surge in communication requirements and mobile device usage intensifies substantially. The current projections on smartphone and internet usage data state that the global smartphone usage by the year 2027 is expected to be around 7.7 billion worldwide [Inc21].

Today, communication applications extend well beyond basic voice or data exchanges between individuals. In various sectors like education, transportation, healthcare, security, agriculture, and manufacturing etc., reliable data transmission among devices is crucial for efficient service delivery. Traditionally, cellular networks manage each communication link via a central base station referred to as *infrastructure-based communication*. Given the growing demand, if all traffic were to route through base stations, it would impose an overwhelming load, leading to increased transmission delays, reduced throughput, lower aggregate data rates, and a decline in spectral efficiency [GTK21]. To address these issues, the 5G communication framework includes the solution of D2D communication. Through D2D communication, user equipment (UE) can communicate directly without base station mediation. This direct mode of communication between two UEs is depicted in Fig. 7.1. This approach reduces latency significantly, ensuring quicker data transmissions with low power consumption, which is vital for time-sensitive applications. Additionally, it alleviates the burden on base stations, improving network performance and accommodating higher user volumes effectively, specially in densely populated areas.

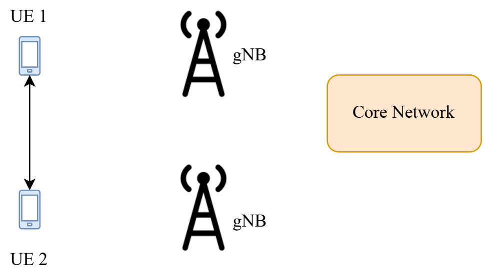


Figure 7.1: The direct mode communication illustration according to [3GP13]

Technical reports TR 23803 [3GP13], and TR 23703 [3GP14] from the 3GPP Release 12 outline two main modes of this direct communication:

- Network independent direct communication
- Network authorized direct communication

Of these, the first mode of D2D communication operates without needing base station network authorization. Direct communication between nearby devices is enabled solely based on information locally available to the UE. However, this mode is restricted to proximity services (ProSe)-enabled public safety UE. In contrast, the second mode requires network assistance from the base station for resource allocation and detecting nearby devices, making it accessible to any ProSe-enabled user within the same radio access network. Considering the applicability to UAM, this chapter focuses on network-authorized D2D communication. Here, the radio resources can be used by both the D2D user as well as the cellular user simultaneously, and only a single link (uplink or downlink) is associated with the transmission as opposed to traditional communication. These two gains achieved through enabling the D2D communication are referred to as *resuse gain* and *hop gain* respectively [FDM+12].

It is clear that the 5G D2D communication is designed to mitigate key issues like high delay, low throughput, and limited sum rates often found in legacy networks, with the goal of establishing a more efficient and robust communication system. Nonetheless, several core challenges associated with D2D communication, as identified in [AP18], continue to affect user experience. These challenges include:

- Peer discovery
- Resource allocation
- Resource management
- Interference management
- power control

Accordingly, although the challenges associated with D2D communications appear to be five different types of challenges that can be addressed individually, these can broadly be categorized into two, given the dependability of some challenges on each other. This is due to the fact that interference management can be achieved through several ways, such as:

- Resource Allocation - Assigning dedicated or shared spectrum resources to D2D links to minimize conflicts with cellular users
- Power Control - Adjusting transmission power to limit interference while maintaining the required quality of service in communication
- Orthogonal Frequency Reuse - Allocating different frequency bands to nearby D2D pairs and cellular users to avoid overlap.
- Beamforming - Directing signal energy toward specific devices to reduce interference with unintended receivers

- Clustering and grouping - Grouping nearby D2D users to manage interference within and outside the cluster
- Resource management - Application of scheduling techniques to manage resource usage by the users
- Mode Selection - Determining the communication mode (D2D or infrastructure mode) depending on interference level, proximity, network conditions, etc.

Based on this, it is clear that the factors of resource allocation, resource management, and power control are key elements that lead to efficient interference management. This dependability makes it viable to address these three topics in the context of interference management. Peer discovery, on the other hand, can be treated as an independent challenge. Consequently, this discussion is limited to the two major categories of challenges associated with D2D communication, which are:

- Peer Discovery
- Interference Management

Peer Discovery

Peer discovery involves identifying nearby D2D-enabled UEs that can establish a communication link. This process enhances resource efficiency, reduces communication delays, and enables diverse applications such as location-based services, disaster recovery networks, and direct sharing of local content among devices. In network-authorized mode, the cellular network or the gNodeB (gNB) assists the UE in identifying peers by providing proximity and channel allocation information. In the network-independent mode, the UEs detect peers using beacon signal transmissions [3GP13, 3GP14]. However, this fundamental procedure in D2D communication also poses several challenges. Mainly, the efficient handling of energy during the process is vital. Furthermore, accurate and timely detection of peers must be done through varying mobility and interference conditions while improving security in the system to prevent malicious attacks. Although peer discovery is a vital component in D2D communication and remains a crucial area of study, this chapter focuses only on interference management.

Interference Management

Interference management aims to minimize interference between cellular and D2D users. This is crucial in network-authorized D2D communication as

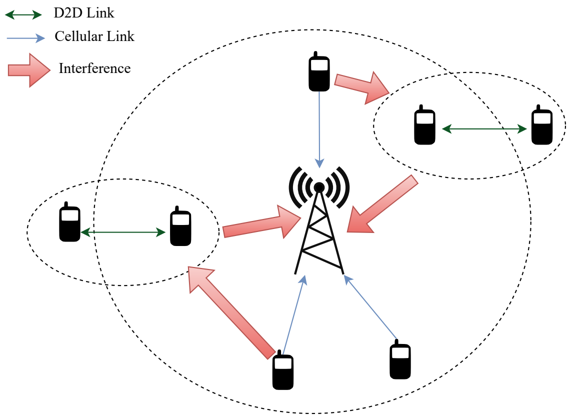


Figure 7.2: Interference scenario with D2D and cellular users

both the cellular and D2D UEs are simultaneously occupying the resources. A general interference scenario is depicted in Fig. 7.2. Depending on the type of resources used by the D2D transmissions, the possibility is there for the generation of both intercell and intracell interferences. Depending on the resource block allocation and duplex schemes used for the D2D communication links, intercell interference could occur. When the D2D link is utilizing cellular uplink resources, a cellular UE of a neighboring cell in the transmitting mode can cause strong interference on the receiving UE of the D2D pair. On the other hand, when the D2D link is using cellular downlink resources, then the transmitting D2D UE can cause strong interference on the receiving cellular UE in the neighboring cell that is utilizing the same downlink resources [FDM+12]. The two intercell interference scenarios according to [FDM+12] are presented in Fig. 7.3. On the other hand, intracell interference occurs when multiple D2D pairs and cellular users within the same cell utilize overlapping resources.

In order to manage the interference, various strategies were pointed out previously. However, while handling all the factors is important, this discussion is confined only to resource handling. Here, *resource* refers to the channel, spectrum, and power allocated to users. Allocating orthogonal or minimally overlapping resources to D2D UE pairs and cellular UEs is taken into account in terms of resource allocation. Resource management in wireless communication systems involves resource block scheduling to improve network

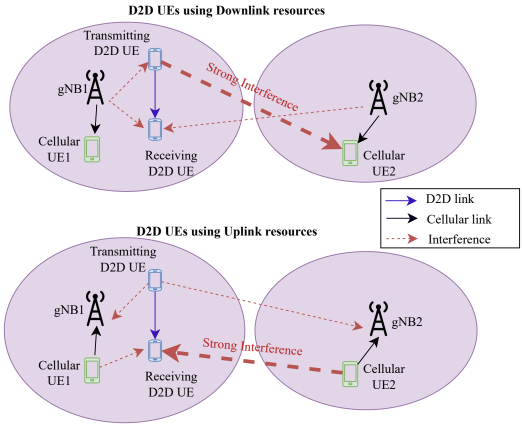


Figure 7.3: Two possibilities of intercell interference according to [FDM+12]

performance and ensure fair resource distribution. Traditional scheduling algorithms such as round-robin, proportional fair, and best channel quality indicator (BCQI) are widely used for allocating resource blocks based on fairness, efficiency, or channel conditions. Advanced strategies, including dynamic switching between scheduling modes and adaptable resource block configurations, are also applicable. In terms of power control, energy harvesting techniques are explored as sustainable solutions for transmission power control. These methods enable devices to operate autonomously over extended periods, reducing dependence on conventional energy sources. Although this is not relevant to the considered application of UAM, it is an interesting power management solution for other applications. Moreover, incorporating energy harvesting mechanisms into D2D systems can significantly support beacon transmission in peer discovery or direct D2D data exchanges to utilize low power ensuring both operational efficiency and environmental sustainability. Despite its importance, this chapter explores only the aspect of resource allocation, excluding detailed discussions on resource management and power control.

7.1.2 Resource Allocation

Resource allocation refers to the process of assigning and distributing available resources, such as spectrum, power, or computational capacity, to different users, devices, or processes to enhance system performance and cater to specific requirements. It focuses on addressing competing demands while adhering to constraints like capacity, fairness, and quality of service. The topic of resource allocation discussed here involves managing how spectrum and power are distributed to enhance network efficiency, minimize interference, and ensure reliability across the D2D communication system. Power allocation is vital for energy conservation and signal quality. On the other hand, spectrum utilization has a direct impact on network capacity, throughput, and overall efficiency. According to the two D2D modes outlined by 3GPP, spectrum utilization in D2D communication can be elaborated under modes [GZ18]:

- Inband communication
- Outband communication

Inband communication operates within the licensed spectrum, offering higher communication quality, improved system throughput, and better spectrum efficiency. This inband mode is further classified into:

- Overlay mode - A separate spectrum is allocated for D2D users
- Underlay mode - D2D users share a portion of the cellular spectrum

In the *overlay* mode, spectrum allocation is separated for D2D and cellular users, assigning each a dedicated, non-overlapping portion of the spectrum [GPG+22]. Due to low interference, overlay spectrum sharing enhances the data rates of both D2D and legacy networks, although spectrum utilization may be less efficient. In contrast, the *underlay* mode allows D2D users to share a portion of the cellular spectrum, leading to increased spectral efficiency, higher data rates, and improved throughput. However, this sharing also results in higher interference levels [GPG+22]. A comparative study, [HC20], analyzes the overlay, underlay, and traditional cellular modes by examining signal-to-noise ratio (SNR), throughput, and transmission delay. Findings indicate that overlay mode performs best in terms of SNR, delay, and throughput, followed by underlay, with the traditional mode yielding the lowest performance, further supporting the observations of [GPG+22].

The second category of spectrum utilization under D2D is *outband* communication which utilizes the unlicensed spectrum. This means that no cellular resources are required for operation [GZ18]. This prevents the system from being prone to interference from cellular users, but still, the involved interference is high because the unlicensed spectrum is quite congested by the other communication technologies.

7.1.3 Contributions

With the related aspects being introduced, this subsection narrows down the topic and elaborates on the contributions discussed in the upcoming sections of the chapter. As discussed before, this chapter concentrates on resource allocation in relevance to interference management within network-authorized D2D communication. D2D communication is considered due to its relevance to A2A communication in UAM. The particular emphasis is on inband underlay D2D mode, as maximizing spectrum utilization is essential. Moreover, the overlay mode is unsuitable for upcoming wireless networks that expect a high user demand because it restricts the number of users that can be supported in the network due to limited orthogonal resources [ZZW+19].

The solution approach here focuses on reinforcement learning techniques, although an analysis of all possible resource allocation techniques will be presented in the next section of this chapter. Reinforcement learning and the associated algorithms will be elaborated on in detail in the upcoming sections. To this end, the main contribution of this chapter is to propose a dynamic resource allocation technique that employs Expected SARSA reinforcement learning to mitigate interference in D2D communication. To summarize:

- **Challenge Addressed :** Improving resource allocation strategies leading to interference management in network authorized D2D communication in application to A2A communication in UAM
- **Approach :** Focusing on reinforcement learning techniques for resource allocation
- **Solution :** A dynamic resource allocation technique based on the Expected SARSA reinforcement learning technique

7.2 Related Work

Interference management and enhancing throughput in D2D communication have gathered significant attention within the research community over recent years. In particular, various resource allocation methods have been explored to address these challenges. Effectively minimizing interference between cellular and D2D users largely depends on the approach used to assign resources such as channels, spectrum, and power to users. [GPG+22] offers a comprehensive taxonomy of solutions that can serve as resource allocation schemes tailored explicitly for D2D communications. In this section, these different techniques are elaborated according to the taxonomy provided in [GPG+22].

Game-Theory Based Resource Allocation

The first category consists of game-theory based resource allocation schemes. Game theory, a branch of applied mathematics, is frequently utilized as a tool to analyze and optimize system processes effectively [BJF22]. Numerous applications of game theory can be found in the literature focused on identifying optimal solutions for resource allocation in D2D communication. For example, [RG20] proposes a *Stackelberg* game theory-based algorithm tailored for heterogeneous networks, which has been validated through performance and throughput analysis. Another study, [WZCZ21], also applies Stackelberg game theory, integrating an interference alignment algorithm to support D2D communication in underlay mode, achieving high sum rates and reduced time complexity. Authors of [MZOA21] apply game theory with mixed strategies in D2D-enabled vehicular networks, optimizing system computational capacity. Additionally, *Coalition* game theory is applied as a resource allocation strategy in D2D communication with both co-channel and cross-channel interference in [GT22], resulting in maximized system sum rates and secure resource allocation. A comparison of these game-theory based resource allocation techniques is summarized in Tbl. 7.1.

Graph-Theory Based Resource Allocation

The second category comprises graph-theory based resource allocation methods, with numerous studies employing the *coloring principle* and *centrality concept* of graph theory. For instance, both [XYYY19] and [ZCHL22] utilize the coloring principle. In [XYYY19], power control and node priority calculations are applied to optimize the communication index, resulting in enhanced fairness alongside reduced power consumption and delay. [ZCHL22] uses an advanced graph coloring method to form interference graphs in secondary clustering for D2D users, further incorporating the *Hungarian algorithm* for optimal channel allocation, which improves system throughput and access rate. Similarly, authors of [JKP+21] apply the centrality concept for D2D pair selection and a modified resource allocation approach, achieving a high

Author and Year	Game-Theory	System Model	Outcome
Sun *et al.* [SWL19], 2019	Coalition formation game and whale optimization based power allocation	D2D communication in underlay mode	Near-optimal throughput and low complexity
Dun *et al.* [DYJ+19], 2019	A distributed algorithm modeled as a non cooperative game	D2D communication in underlay mode	Improved throughput and guaranteed quality of service
Ghosh *et al.* [GD20], 2020	Cooperative game theory	Cluster based D2D system	Low power consumption, enhanced SNR and spectral efficiency
Rathi *et al.* [RG20], 2020	Stackelberg game theory	Heterogeneous networks	Analysis on system performance and throughput
Wang *et al.* [WZCZ21], 2021	Stackelberg game and interference alignment algorithm	D2D communication in underlay mode	High sum rate and low time complexity
Mensah *et al.* [MZOA21], 2021	Game with mixed strategies	D2D enabled vehicular network	Optimized system computation capacity
Gupta *et al.* [GT22], 2022	Coalition game	D2D with co-channel and cross-channel interference	Maximized system sum rate and secrecy-ensured resource allocation

Table 7.1: Game-Theory based resource allocation schemes for D2D communication

maximum sum rate and reduced frequency range. A comparison of these methods is presented in Tbl. 7.2.

Joint Channel-based Resource Allocation

The third category of resource allocation schemes identified by [GPG+22] involves joint channel-based mechanisms. An example of this approach is presented in [ASKC19], where a Non-Orthogonal Multiple Access (NOMA) framework is applied to D2D communication. The study employs the *Kuhn-Munkres* method for sub-channel allocation and uses *Karush-Kuhn-Tucker* conditions for power allocation. System analysis under varied network conditions demonstrated improvements in both energy efficiency and throughput. Another approach, discussed in [DSS+21], focuses on a multi-user D2D

Author and Year	Graph-Theory	System Model	Outcome
Chang *et al.* [CLS+19], 2019	Branch and bound based algorithm	Using a high transmission power on the D2D users with minimum data requirements	High weighted sum data rate
Yuhong *et al.* [XYYY19], 2019	Graph theory coloring principle	Power control and node priority calculation to optimize communication index	Improved fairness, low power consumption and low delay
Jeon *et al.* [JKP+21], 2021	Centrality concept of graph theory	D2D pair selection based on graph theory and a modified resource allocation scheme	High maximum sum rate and reduction in the frequency range
Zhou *et al.* [ZCHL22], 2022	Improved graph coloring	Interference graph formation by secondary clustering of D2D user nodes. Hungarian algorithm for optimal channel allocation	Improved system throughput and access rate

Table 7.2: Graph-Theory based resource allocation schemes for D2D communication

underlay cellular network, introducing a matrix channel allocation algorithm that ensures each D2D user is assigned at least one cellular sub-channel. Simulation results indicate that this approach yields high spectrum utilization and efficiency. Additionally, [LWLL20] utilizes the *Dinkelbach* algorithm through an iterative method, achieving maximized energy spectral efficiency by optimizing joint channel selection.

NOMA Based Resource Allocation

Another subset of resource allocation methods for D2D communication includes NOMA-based approaches [GPG+22]. NOMA, a power-based multiple access technique, enables multiple users to access the same time, frequency, or code resources by assigning different power levels. This approach is designed to enhance the network's spectral efficiency in 5G environments [BTTG19]. An example of NOMA-based resource allocation is [CHCX21], where performance is measured by quality of experience. Here, sub-channel assignment and power allocation are optimized through an iterative process involving an alternating optimization algorithm combined with a constrained convex procedure technique. Two additional studies in this category include [BKT22] and [BTTG19]. In the first, an interference management paradigm is presented for uplink NOMA, while in the second, a joint channel allocation and power control mechanism for femtocell users is developed using a cognitive radio-based non-orthogonal multiple access model.

Machine Learning Based Resource Allocation

The next category of resource allocation algorithms for D2D communication encompasses machine learning (ML) based methods. Among the various studies in this area, work in [ZLYY20] proposes a neural network-driven scheme for underlay D2D users, enabling them to reuse cellular resources through a non-linear optimization algorithm. This approach demonstrates high performance with reduced computational complexity. Another study, [LL21], combines a heuristic equally reduced power scheme with a deep neural network-based approach for transmitting power allocation across channels, achieving near-optimal sum rates. Additionally, [BV23] offers a framework for analyzing effective ML-based resource allocation strategies specifically for mmWave D2D networks.

Deep Reinforcement Learning Based Resource Allocation

As detailed above, various approaches are available to address resource allocation in D2D communication. Yet, one of the most widely adopted techniques today is deep reinforcement learning (DRL). These methods utilize DRL techniques, such as Q-learning, to train D2D users, enabling them to independently select resources like power and channels for D2D networks. For instance, in [YJL21], a deep Q-network is proposed where the D2D transmitter selects channel resources and transmits power based on the positions of mobile devices and actions of other D2D users, maximizing the overall throughput of D2D communication. In a different study [KMMH21], a double-dueling deep Q-network is employed, where a centralized controller interacts with the environment to determine an optimal resource allocation strategy, delivering near-optimal outcomes.

The double deep Q-network (DDQN) has been widely utilized in prior research to address D2D communication resource allocation. In [XPG+22], DDQN with priority sampling (Pr-DDQN) enables D2D users to identify essential features, facilitating the selection of suitable channel resources and transmit power for high data rates. In [YLL+22], DDQN is applied to enable D2D pairs to learn optimal strategies based on local data about achievable rates and power usage, yielding better convergence results. Another example is [QG23], where DDQN is integrated with an energy-harvesting model to optimize resource allocation in D2D underlay mode, achieving enhanced energy efficiency. This vast range of research studies on DRL techniques for resource allocation is summarized in Tbl. 7.3.

Other Resource Allocation Techniques

In addition to the aforementioned methodologies, a comparative study conducted in [JJC+21] evaluates a Q-learning-based power allocation scheme against a neural network model that is trained using data generated from the Q-learning approach. The findings indicate that both strategies can achieve high system throughput, while the neural network-based model offers reduced time costs compared to the Q-learning method. The research presented in [SLO+22] introduces an innovative approach that integrates Stackelberg game theory with Q-learning. In this model, the learning algorithm is directed by a pivotal Stackelberg Q-value derived from the equilibrium established within the Stackelberg game. This enables D2D users to allocate resources in a decentralized fashion and facilitates quicker system convergence. This method has been shown to enhance average utility, decrease training time, and increase channel capacity. Furthermore, [RRNY21] proposes a strategy to switch between communication modes based on the communication distances. The algorithm determines whether to operate in overlay mode (for short distances) or underlay mode (for longer distances) based on a specified distance threshold. Additionally, system throughput is evaluated by adjusting the transmit power. The results demonstrate improved SNR values and throughput while minimizing interference.

7.3 Associated Concepts

The objective of the work is to dynamically allocate channel resources for the D2D UEs, aiming to minimize interference in the underlay D2D scenario. In this section, associated concepts are discussed. The following resources are considered channel resources to be dynamically allocated.

- Number of resource blocks (NRB)
- Numerology Index (NI)
- Transmission Power (Tx Power)

These parameters are chosen as they have a direct impact on the system's performance. The number of resource blocks and the numerology index govern the channel bandwidth, as highlighted before in Eqn. 5.18. The transmission power directly affects the SNR value according to the SNR formula

Author and Year	DRL Technique	System Model	Outcome
Bi *et al.* [BZ20], 2020	Deep Q-learning	Power control considering the channel state as a finite state Markov channel	Maximizing system capacity and quality of user experience
Yu *et al.* [YJL21], 2021	Deep Q-network	D2D transmitter determines the channel resource and transmit power based on locations of mobile devices and actions taken by other D2D users	Maximizing the overall effective throughput of D2D communication
Huang *et al.* [HYH+21], 2021	Double deep Q-network (DDQN)	D2D pairs find the unused time slots within a frame and reuse the time slots of the cellular user	Achieving maximum sum throughput and fairness between D2D pairs
Kai *et al.* [KMMH21], 2021	Double-dueling-deep Q-network (D3QN)	A centralized controller interacts with the environment to find the best resource allocation approach	Near-optimal performance
Xiang *et al.* [XPG+22], 2022	DDQN with priority sampling (Pr-DDQN)	D2D users learn the significant features through Pr-DDQN and selects the channel resources and transmit power to use	High achievable rate
Yuan *et al.* [YLL+22], 2022	Double deep Q-network (DDQN)	D2D pairs learn the optimal strategy based on local information regarding the difference of achievable rate of D2D users and used power	Better convergence performance
Qi *et al.* [QG23], 2023	DDQN algorithm	DDQN in combination with energy harvesting model for optimization of resource allocation in D2D underlay mode	High energy-efficient performance

Table 7.3: DRL based resource allocation schemes for D2D communication

described in Eqn. 4.3. The channel Bandwidth (Channel BW) and the SNR contribute to the system throughput calculation according to the *Shannon Capacity Theorem* as highlighted by the Eqn. 4.2. System throughput can be used as a measure of interference level in a communication system because interference directly impacts the successful transmission and reception of data. High interference levels typically degrade signal quality, leading to increased packet loss, retransmissions, or lower SNR, which reduce the effective throughput. Conversely, lower interference allows for cleaner transmissions, higher SNR, and more efficient use of bandwidth, resulting in improved throughput. Therefore, monitoring throughput offers an indirect but practical way to assess the prevailing interference in the system. To this end, system throughput is chosen as the main performance indicator.

Given the above conditions, this problem can be characterized as a Markov Decision Process (MDP) due to its iterative decision-making nature. Changes in channel resources directly influence throughput, and this feedback-driven process thus inherently involves states, actions, and rewards, which are the characteristics of an MDP. The Markov property holds since each decision depends only on the current state and action, not prior states. Moreover, the system exhibits stochastic behavior because interference levels and throughput are influenced by unpredictable factors like network traffic and user mobility. Therefore, applying reinforcement learning, which optimizes the system through a trial and error based approach, is suitable for the problem under discussion.

7.3.1 Reinforcement Learning

Reinforcement learning (RL) algorithms learn to make optimal decisions by observing the environment and selecting the most suitable actions for each situation to achieve the best possible results. RL is dependent on five elements, and understanding the function and significance of each is fundamental to comprehending RL. The five elements are:

1. State (S_t) - The state at any given moment
2. Action (A_t) - Action performed
3. Reward (R_t) - Reward for performing the action
4. Agent - The entity that learns by taking actions and receiving feedback
5. Environment - The external system within which the agent operates and interacts

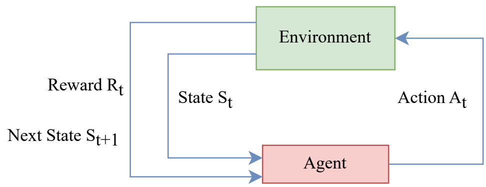


Figure 7.4: Agent and environment interaction in reinforcement learning

In RL, the policy, denoted as π, is a fundamental concept. Here, an agent considers a state as input and seeks the optimal policy based on the reward obtained by taking specific actions. The policy can be either deterministic or stochastic. A task can be represented as an MDP using the five key RL elements, where the agent's policy acts as a map from each possible state to a probability distribution across potential actions. The agent iteratively refines the policy by interacting with the environment and collecting rewards. This process of interactions between the agent and the environment is depicted in Fig. 7.4. The goal is to find the optimal policy which maximizes the expected cumulative reward over time. The optimal policy can be calculated using offline methods like *dynamic programming* or through online, experience-based approaches, such as *Monte Carlo methods* and *temporal difference methods* (TDM). Online methods generally converge more quickly since they do not require exploration of the entire state space, unlike offline approaches [vSvHWW09].

In TDM, an agent interacts with its environment, generating a trajectory of visited states, actions taken, and rewards received. This sequence is then used to update the agent's policy. TDM can operate as either *off-policy* or *on-policy*. In off-policy learning, such as with the Q-learning algorithm, two separate policies are employed: a *behavior policy* for exploring the environment and an *estimation policy* focused on optimization [vSvHWW09]. Conversely, in on-policy strategies, the behavior and estimation policies are the same, meaning the algorithm updates the policy as it follows it. SARSA serves as an example of this on-policy approach.

The name SARSA is derived from the five elements that drive its update rule: current **S**tate (S_t), current **A**ction (A_t), received **R**eward (R), next **S**tate (S_{t+1}), and next **A**ction (A_{t+1}). The Q-value, Q(S_t, A_t), represents the expected return from taking a specific action in a given state under the policy. Since the policy continually adapts, SARSA is advantageous in dynamic environments where on-policy exploration aligns closely with environmental

RL Technique	Advantages	Disadvantages
Dynamic Programming	Provides optimal solutions under known environment models, fast convergence, useful for solving Markov Decision Processes	Requires complete knowledge of the environment, computationally expensive for large state spaces, poor scalability in complex problems
The Monte Carlo method	Does not require prior knowledge of the environment and easy to implement,	High variance in estimates and inefficient for continuous tasks
Q-Learning	Model-free, no need to know the environment, off-policy learning, suitable for stochastic and dynamic environments	Prone to instability due to greedy action selection, slow convergence in large state spaces
SARSA	On-policy learning, stable convergence, adapts to the current policy, reducing exploratory risk	Requires balancing exploration and exploitation, slower than off-policy methods like Q-Learning, the possibility of converging to suboptimal policies if exploration is insufficient
Expected SARSA	Reduces variance by considering the expected value of future states, smoother convergence in stochastic environments	Requires calculation of expected value over all actions, computationally complex than SARSA

Table 7.4: Comparison of various reinforcement learning techniques

shifts. However, SARSA's drawback is that, the dependence on the next action, A_{t+1}, introduces an additional variance, which can slow convergence [vSvHWW09]. Expected SARSA extends SARSA with a solution to the additional variance issue imposed by SARSA.

A comparison of different reinforcement learning techniques is provided in Tbl. 7.4. Based on the advantages and disadvantages discussed, it is clear that each method consists of both strengths and weaknesses, making them applicable to different types of problems and environments. The choice depends on factors like environmental knowledge, computational resources, and task requirements. In terms of the problem addressed in this chapter, it is to be noted that this problem involves a continuous learning process. That means, the system learns and adapts iteratively over time, refining its understanding as new data or experiences become available. Given this stochastic nature, the reinforcement learning technique of Expected SARSA is particularly suited for improving this system based on the reinforcement learning technique comparison elaborated in Tbl. 7.4. Expected SARSA facilitates continuous learning, allowing the system to adapt iteratively and refine its performance over time.

7.3.2 Expected SARSA

Expected SARSA extends SARSA by using a weighted average of Q-values for all possible actions in the next state to estimate the Q-value. This approach reduces variance, allowing Expected SARSA to converge faster with less training data required [vSvHWW09]. Consequently, Expected SARSA offers an efficiency advantage over SARSA. Expected SARSA operates as follows: first, the agent observes the current state, S_t, and selects an action, A_t, based on the policy π. The policy is termed *ϵ-greedy*, where $0 \leq \epsilon \leq 1$. With probability $1-\epsilon$, the agent chooses an action that maximizes the estimated Q-value, Q(S_t, A_t), to exploit for the highest anticipated reward. Otherwise, it selects a random action, engaging in *exploration*. The agent then observes the reward, R, for the chosen action and transitions to the next state, S_{t+1}. The Q-value update in Expected SARSA follows the rule:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R + \gamma(\sum_{A_t \in \delta_t} \pi(A_t|S_{t+1})Q(S_{t+1}, A_t) - Q(S_t, A_t))],$$

The Q-value of the current state-action pair is estimated with α as the learning rate. In this context, δ_t represents the set of all possible actions in state S_t, and γ is the discount factor, which adjusts returns to prioritize actions that yield the most efficient rewards. Consequently, Q(S_t, A_t) denotes the expected cumulative reward for taking a particular action in a given state aligned with the policy. The algorithm then updates the policy π for the current state based on these estimated Q-values. Expected SARSA offers flexibility, as it can also be applied as an off-policy approach. The Expected SARSA algorithm is given by Alg. 3.

7.4 System Implementation

The basic network setup is structured similarly to Fig. 7.2 with a single cell. It includes one gNB, a D2D pair, and a variable number of interfering devices as necessary for the experiments, and these details will be provided later in the chapter. Communication setup relies on the urban macro propagation model and path loss model outlined in the 3GPP 5G standard documentation [3GP17a]. The implementation is conducted in OMNeT++ using the Simu5G framework. To simulate realistic conditions, Rayleigh fading and shadowing effects are incorporated, while other SNR influencing parameters such as thermal noise, attenuation, and fading remain unmodified compared to modeling in Simu5G.

Algorithm 3 Expected SARSA

```
Input: α learning rate, ϵ random action probability, γ discount factor
Initialize Q(s,a) arbitrarily for all states s and actions a
Initialize starting state s
repeat
    Choose action a from state s using policy π
    Take action a, observe reward R and next state s'
    Calculate the expected value for the next state:
        V_{s'} = Σ_a π(a|s') · Q(s',a)
    Update Q-value:
        Q(s,a) ← Q(s,a) + α [R + γV_{s'} − Q(s,a)]
    Set s ← s'
until s is a terminal state
end
```

D2D communication simulation with Simu5G

In Simu5G, D2D communication is made possible by employing uplink resources. This approach is adopted because uplink traffic typically experiences lower utilization compared to downlink traffic and is less affected by interference. The possibility is there to model both point-to-point involving one transmitting D2D UE and a receiving D2D UE as well as transmitting to a multi-cast group of receiving D2D UEs referred to as point-to-multi point D2D communication in Simu5g. Here, acknowledgments are enabled but limited to the point-to-point case. In enabling D2D communication capabilities, a few modifications specific to D2D communications have been done, but the same protocol stack previously discussed under subsection 4.2.2 is applied [NSTV20]. The D2D communication discussed in the chapter concerns only point-to-point D2D communication. These modifications according to [NSTV20] are:

- Adding a flow direction ID into the UE's PDCP layer packet to indicate the communication mode. The model further provides a mode-switching possibility such that the flow direction can be dynamically updated depending on communication range or interference levels, etc.

- Introducing the Hybrid Automatic Repeat Request (H-ARQ) technique for point-to-point communication. The acknowledgment in H-ARQ is sent back directly to the transmitting D2D UE. Simultaneously, a copy of this acknowledgment is also sent to the gNB as the gNB needs to keep track of if the D2D communication is successful.

Action (A_t)	NRB	NI	Channel BW (MHz)	Tx Power (dBm)
1	125	0	25	23
2	100	0	20	23
3	50	1	20	26
4	25	0	5	26
5	75	1	30	26

Table 7.5: Action space for Expected SARSA algorithm

- Dynamic scheduling of D2D transmissions is done. This lets multiple D2D transmitters utilize the same resource blocks concurrently within a single transmission time interval, minimizing total spectrum usage.

Simu5G is widely utilized in D2D communication studies, as demonstrated by [AJ22] and [MLC+22]. [AJ22] evaluates infrastructure versus D2D modes, concluding that D2D enhances spectrum utilization and total throughput. Meanwhile, [MLC+22] leverages Simu5G to examine D2D communication quality within vehicular networks, focusing on the BCQI metric to assess performance. These studies underscore Simu5G's role in analyzing D2D communication's impact on network performance in diverse scenarios.

7.4.1 Modeling of Expected SARSA Algorithm

With the foundation of Expected SARSA and network architecture established, the specific components, states, actions, rewards, and other parameters of the Expected SARSA system are formally defined.

Action space

As explained before, the three parameters: NRB, NI, and the Tx Power, are dynamically adjusted. Accordingly, five distinct actions involving combinations of these parameters are identified. This action space is defined in Tbl. 7.5. The values within this action space are selected based on preliminary experiments conducted without the application of Expected SARSA, focusing on minimizing the resources utilized.

State space

As the state space, the main performance indicator *throughput* is chosen. Throughput is a suitable state space for this problem because it encapsulates the system's overall performance and directly reflects the impact of dynamic resource adjustments (e.g., bandwidth and transmission power). Throughput inherently depends on interference levels, resource allocation, and system conditions, providing a comprehensive representation of the network's current state. As the key outcome metric, monitoring throughput allows the reinforcement learning algorithm to adapt decisions iteratively, ensuring alignment with the objective of enhancing resource usage while maintaining or improving performance.

Initially, the achievable throughput is evaluated under the default conditions provided by Simu5G. Based on this analysis, two threshold values are established, and three states are defined in relation to these two threshold levels:

- Upper Threshold = 250 Mbps
- Lower Threshold = 100 Mbps

The achieved throughput is considered as the state level, and it is compared with the threshold values, and the state space (S_t) is defined as

$$State(S_t) = \begin{cases} 2 & Throughput \geq UpperThreshold \\ 1 & LowerThreshold \leq Throughput < UpperThreshold \\ 0 & Throughput < LowerThreshold \end{cases}$$

Reward Function

Throughput is again utilized in the reward function definition. By enhancing throughput, the system improves data transmission rates and overall communication efficiency, leading to better performance and user satisfaction. In this problem, maximizing throughput stimulates the agent to make decisions that enhance network performance. This reward mechanism encourages actions that improve system efficiency while considering factors that affect throughput. To this end, the same threshold levels of throughput defined before under the state space are utilized for the reward function. Accordingly, the following rewards are assigned.

$$Reward(R) = \begin{cases} +80 & Throughput \geq UpperThreshold \\ 0 & LowerThreshold \leq Throughput < UpperThreshold \\ -30 & Throughput < LowerThreshold \end{cases}$$

After a communication cycle, if the achieved throughput is higher than the upper threshold, a reward is given as +80. If the throughput falls between the lower and upper threshold bounds, then no reward is granted (R = 0). As opposed to passing the upper threshold, if the throughput lies below the lower threshold, then it is no longer a reward but a penalty, which is -30. There exist no technical criteria for allocating reward values. Therefore, the used reward values are determined through experimental evaluations aimed at fine-tuning the algorithm to reach a balance between the actions taken and the system performance. The reward for throughput values above the upper threshold is set to be a comparatively very high value to encourage the selection of actions that lead to high system performance. Similarly, for the medium level, a neutral reward (R=0) signifies adequate but not exceptional performance, guiding the system to improve without huge penalties. Finally, the penalty discourages configurations, leading to poor performance and forcing the system to avoid taking actions that cause inefficiency.

Parameter based Exploration

In Expected SARSA, the parameter ϵ controls the balance between exploration and exploitation. A higher ϵ increases exploration, encouraging the algorithm to try out more actions. Conversely, a lower ϵ favors exploitation, where the algorithm tends to select actions based on past experiences, leading to maximized rewards. This balance allows the algorithm to both learn from new experiences and refine its knowledge from prior ones. In general, the ϵ value is a static parameter and not changed during the simulation. However, given the stochastic nature of the considered application, the environment keeps changing rapidly. In such a situation, having a static ϵ value could result in potential bias in the algorithm. This is due to the fact that if the system performance is deteriorating and a low ϵ is used, it will overly rely on exploitation, hindering the optimization. Conversely, if the system performance is high and a high ϵ is chosen, it introduces unnecessary randomness, which can degrade performance.

This shows that a fixed epsilon might not be suitable for dynamic environments where the balance between exploration and exploitation needs to be adjusted based on the prevailing system state. Therefore, in addition to defining the state, action, and reward spaces and initiating the learning process, a parameter-based exploration with regard to the ϵ is also done. Here, the parameter ϵ is varied based on the achieved throughput to mitigate any possible bias in the algorithm. As the state changes, ϵ is adjusted towards exploitation. That is, a lower ϵ value is employed if throughput is favorable, while a higher value is applied when throughput declines.

Other Parameters

As per other parameters relevant to the Expected SARSA algorithm, the following parameter values are chosen based on the initial experiment analysis

- learning rate $\alpha = 0.5$
- discount factor $\gamma = 0.99$

Learning Model

The modeling of the Expected SARSA algorithm in OMNeT++ is done by following the approach outlined in [MT23]. At each communication cycle, the agent observes the current state and selects channel resources. The Expected SARSA algorithm initiates once the initial SNR measurements are recorded. After computing the throughput with the current parameters, the agent receives a corresponding reward. The next state is then determined based on the current throughput. Subsequently, the next action is chosen in accordance with the policy, and the Q-value table is updated. During initialization, the values of the Q-value table are set to a high value, promoting exploration during the initial iterations of the learning process. The module *component carrier* in the Simu5G architecture is the module responsible for initializing the parameter values of channel resources. It then reads the new action value, which establishes the NRB, NI, and Tx power to be utilized in the subsequent communication cycle. The component carrier signals the module binder (previously discussed in subsection 4.2.2) to register the new carrier, ensuring that the channel model and other system components are updated with the new parameters. Consequently, in each communication cycle, the resource parameters are updated based on the experience gained. This procedure is depicted as a flow chart in Fig. 7.5. As the objective of the policy is to augment the system throughput, the agent is inclined to select actions that promote performance enhancements. The algorithm is depicted in Alg. 4.

The carrier frequency is another parameter investigated within the Simu5G setup, where configuration occurs through the *component carrier* module in Simu5G. The Expected SARSA algorithm aligns with Simu5G's update interval, meaning frequency adjustments influence the communication framework, as both the D2D transmitter and receiver must operate on a shared, gNB-supported carrier frequency. Adjusting this frequency necessitates a deeper exploration of Simu5G's physical and MAC layer implementations. Consequently, carrier frequency adjustments or mode switching between underlay and overlay modes are not applied within this scope.

Algorithm 4 Expected SARSA algorithm for dynamic resource allocation

Input: α learning rate, ϵ Exploration rate, γ discount factor
Initialize
Q-table: $Q[\text{state}, \text{action}]$ with initial value
Set initial state $state$
loop
  Step 1: Action Selection

$$Probability = \begin{cases} \epsilon & \text{Random Action Selection} \\ 1-\epsilon & \text{Choose Greedy Action:} a = \arg\max_{a'} Q[\text{state}, a'] \end{cases}$$

  Step 2: Execute Action
  Execute action in the environment
  Observe $reward$ and $nextState$
  Step 3: Calculate Expected Q-value
  Find the Maximum Q-value from the Q-table: **Returns** action with maximum Q-value
  Compute action probabilities for $nextState$:

$$\text{Probability[action]} = \begin{cases} \frac{\epsilon}{\text{Number of Actions}}, & \text{for non-greedy actions} \\ 1-\epsilon+\frac{\epsilon}{\text{Number of Actions}}, & \text{for the greedy action} \end{cases}$$

  Compute:

$$\text{nextStateExpectation} = \sum_{\text{actions}} \text{Probability[action]} \times Q[\text{nextState}, \text{action}]$$

  Step 4: Update Q-value
  Update:

$$Q[\text{state}, a] \leftarrow Q[\text{state}, a] + \alpha \times (\text{reward} + \gamma \times \text{nextStateExpectation} - Q[\text{state}, a])$$

  Step 5: Update State:
  $state \leftarrow nextState$
end loop

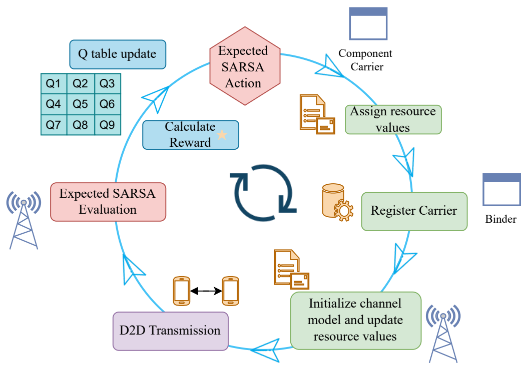


Figure 7.5: Resource allocation procedure with Expected SARSA application

Parameter	Value
Carrier Frequency (f_c)	2 GHz
Distance between D2D UEs and gNB	50 m
UE Height	55 m
gNB Tx Power	46 dBm
Packet Size	1000 Bytes
Transmission Protocol	UDP
Simulation Duration	30 s
Number of Repetitions	30

Table 7.6: Parameter values for experiments

7.4.2 Experiments

For the general network implementation, the example network configuration and default parameter settings for D2D communication provided in Simu5G example scenarios are applied. The UE was set at a height of 55 m to replicate the UAM scenario. The used parameters and simulation conditions common to all experiments are given in Tbl. 7.6

Different types of experiments are carried out to evaluate the algorithm under varying conditions, such as the mobility of the D2D users and the

Experiment	Description
Throughput Analysis	A throughput analysis between the proposed algorithm and default case from Simu5G
SNR Analysis	Complimentary Cumulative Distribution function of the SNR is compared in the presence of interfering devices between the proposed algorithm and default case from Simu5G
Reliability and Delay Analysis	Comparing the Delivery Ratio and the Delivery Delay between the proposed algorithm and default case from Simu5G
Analysis of Impact of Mobility	Comparing algorithm's performance when the D2D UEs are mobile (A2A use case) between the proposed algorithm and default case from Simu5G

Table 7.7: Experiments conducted to evaluate the performance of Expected SARSA

distance between the UEs, etc. The conducted experiments are summarized in Tbl. 7.7. A detailed description of the results will be provided in the upcoming section.

7.5 Analysis of Results

This section presents the results from the OMNeT++ simulations conducted to evaluate the proposed concept. The results shown are derived from averaging statistics across 30 independent repetitions for each experiment, with a consideration of 95% confidence interval. For the performance evaluation of the algorithm, several metrics, such as the system throughput calculated according to Eqn. 4.2, SNR based on Eqn. 4.3, message transmission delay and delivery ratio defined by Eqn. 4.1, are analyzed.

7.5.1 Experiment: Throughput Analysis

Initially, the performance of the proposed algorithm is compared against the default D2D scenario provided in Simu5G. Parameters considered for the default scenario are given in Tbl. 7.8. The D2D UEs are considered to be stationary. To provide a more detailed evaluation, the throughput performance is analyzed in relation to two additional parameters. The first parameter is the distance between the D2D pair. The purpose is to explore the influence of increasing the D2D pair distance, as the proximity between the two devices is a main contributing factor to the performance of D2D communication. Three distance instances are chosen: 20 m, 40 m and 60 m.

The second parameter is the number of interfering devices in the environment. As the ultimate goal of handling resource allocation problems is

Parameter	Value
Carrier Frequency (f_c)	2 GHz
Distance between D2D UEs and gNB	50 m
gNB Tx Power	46 dBm
D2D Transmission Power	26 dBm
Number of Resource Blocks	50
Numerology Index	0
D2D User Mobility	Disabled (stationary UEs)
UE Height	55 m

Table 7.8: Parameter values for default scenario

to minimize the associated interference, the realistic approach to test the algorithm is in the presence of interfering devices. This is assessed under three conditions:

- No interference
- Presence of 10 interfering devices

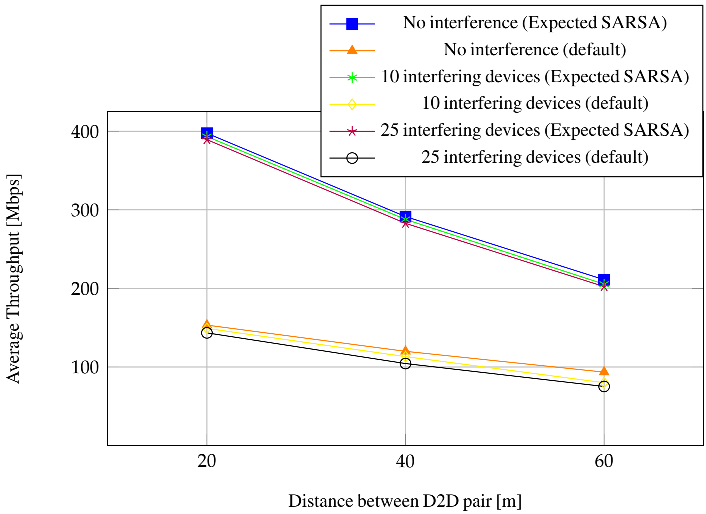


Figure 7.6: Average D2D throughput comparison for various distances between the D2D pair with 10 and 25 interfering devices

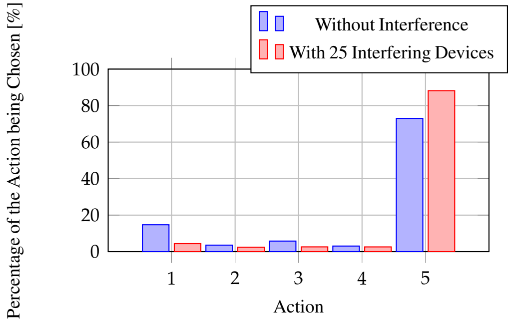


Figure 7.7: Comparison of the percentage of each action being chosen during the dynamic resource allocation procedure when the distance between the D2D pair is 60 m for the two cases: no interference and in the presence of 25 interfering devices

- Presence of 25 interfering devices

For this purpose, the background cell traffic generation model from Simu5G is utilized, with all interfering devices assumed to be stationary. The outcome of the experiments is presented in Fig. 7.6. Accordingly, the average throughput comparison reveals a significant improvement (a more than double increase) in throughput across all cases using the proposed algorithm. Notably, the difference in throughput values between no interference and interference scenarios is reduced in the proposed system compared to the default case, highlighting that the Expected SARSA algorithm effectively maintains performance even with substantial interference from 25 interfering devices. Throughput performance also degrades as the distance between the D2D pair increases to 60 m, aligning with the general observation that the shorter the distance between the D2D devices, the better the performance. However, even at a greater distance of 60 m, the Expected SARSA based approach achieves a throughput exceeding 200 Mbps, compared to under 100 Mbps in the default case. To validate this, results from two instances are further discussed in terms of chosen action throughout the simulation. For this, one result instance pertaining to no interference and in the presence of 25 interfering devices from a distance of 60m are compared.

These results are depicted in Fig. 7.7. According to the action space definitions in Tbl. 7.5, action number 5 corresponds to the highest allocated channel bandwidth (30 MHz). The results in Fig. 7.7 reveals that under interference, the algorithm selects action number 5 more frequently than in

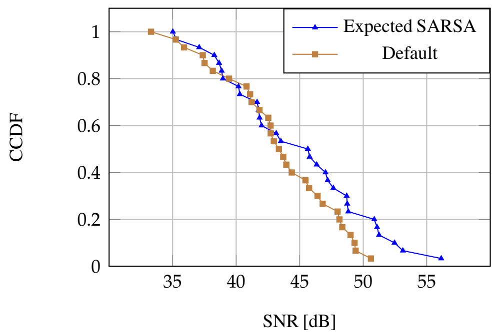


Figure 7.8: Complimentary cumulative distributed function (CCDF) plot of measured SNR with 25 interfering devices for a 20 m D2D pair distance

interference-free conditions (88 % of the time in the former as compared to 72 % in the latter). This behavior demonstrates the algorithm's capacity to adapt under varying conditions. Additionally, the observed action selection percentages suggest that consistently allocating a high static bandwidth may appear as a feasible approach for maintaining system performance. However, such a strategy could lead to unnecessary resource wastage under favorable system conditions, where reduced resource usage would suffice to achieve acceptable performance levels.

7.5.2 Experiment: SNR analysis

Additionally, an analysis of SNR results is performed specifically for the scenario with 25 interfering devices and a 20 m distance between the D2D pair. The results are depicted in Fig. 7.8. This result further supports the effectiveness of the proposed algorithm. The complementary cumulative distribution function (CCDF) of SNR values shows that the CCDF obtained with the proposed algorithm is slightly shifted toward higher SNR values, indicating better performance in the presence of 25 interfering devices compared to the default scenario. This result demonstrates that the Expected SARSA algorithm used in the proposed approach substantially enhances interference management in D2D communication.

7.5.3 Experiment: Reliability and Delay Analysis

Beyond throughput and SNR, the delivery ratio and delivery delay are also evaluated to assess reliability and delay performance. This analysis is conducted with a D2D pair distance of 60 m to test the system under interfering conditions. For interference, only the scenario with 25 interfering devices is considered. The average delivery ratio and average delivery delay are calculated with a 95% confidence interval.

Fig. 7.9 and Fig. 7.10 present the average delivery ratio and delivery delay results with 95% confidence intervals, respectively. Both metrics show similar trends for the proposed algorithm and the default scenario in each case. The delivery ratio shows a slight improvement with the proposed algorithm under no interference conditions while remaining consistent with the default scenario when interference is present. Generally, achieving a delivery ratio of 99% or higher is expected in D2D communication scenarios, and the values obtained here are for a relatively large D2D pair distance of 60 m. Thus, it is evident that the proposed algorithm does not hinder but instead supports maintaining or improving system reliability, even at an increased D2D pair distance.

In contrast, delivery delay values for the proposed algorithm show a minor increase compared to the default scenario, though the difference is very low. This slight increase in delivery delay is balanced by the significant throughput improvement achieved without compromising reliability. Although a delivery delay of around 20 ms is not ideal for a D2D communication scenario, similar values are observed even in the default case. This indicates the need for further enhancements in terms of delivery delay at higher D2D distances.

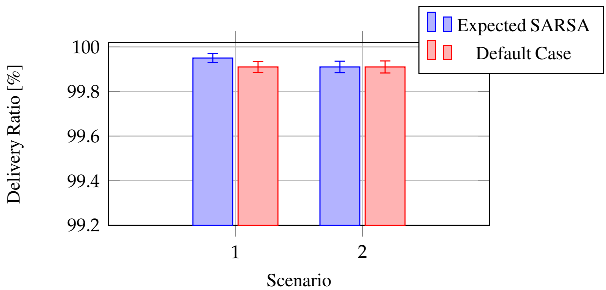


Figure 7.9: Comparison of performance of delivery ratio when the distance between the D2D pair is 60 m for the two scenarios: 1 - no interference and 2 - in the presence of 25 interfering devices

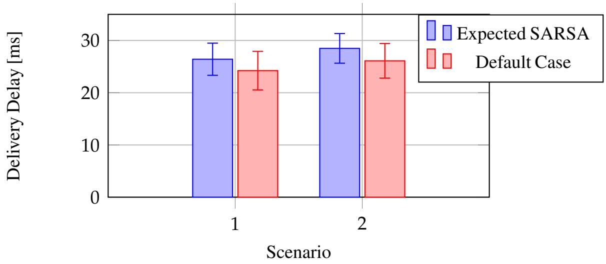


Figure 7.10: Comparison of performance of delivery delay when the distance between the D2D pair is 60 m for the two scenarios: 1 - no interference and 2 - in the presence of 25 interfering devices

Nonetheless, given that this aspect is not a primary focus of the current study, further analysis has not been conducted.

7.5.4 Experiment: Analysis of the Impact of Mobility

Additionally, an experiment is conducted to observe the algorithm's performance under mobility. The experiments leading to this point considered a stationary D2D UE scenario to evaluate the proposed algorithm. As seen from the results, the algorithm significantly contributes to enhancing the D2D system performance. In this set of experiments, the performance in the presence of mobility, which is crucial due to the application being UAM, is evaluated. For this reason, this analysis is aimed at the specific application scenario of A2A communication. Throughput is assessed under three cases:

- Case 1: D2D users are mobile with no interfering devices
- Case 2: D2D users are mobile with 25 stationary interfering devices
- Case 3: Both D2D users and interfering devices are mobile

The *MassMobility* model from the INET framework, which induces random motion in nodes with a mass, is applied to both D2D users and interfering devices. This mobility model is chosen to support the UEs being air taxis in the considered A2A communication scenario. A speed of 20 m/s is set for all UEs at a height of 100 m and movement is confined to a 100 m x 100 m area to ensure continuous D2D communication possibility. The trajectories from ULTRAS simulation tool are not used here to guarantee D2D communication availability during the entire simulation. Table 7.9 provides the throughput

Scenario	Case 1	Case 2	Case 3
Expected SARSA	294.25 Mbps	292.78 Mbps	293.78 Mbps
Default	113.60 Mbps	114.23 Mbps	115.33 Mbps

Table 7.9: Throughput values in mobility scenario

results for the three cases. In the mobility model, the random movement of devices causes the distance between D2D users to vary and can lead to fluctuating interference levels. Nonetheless, even in the case where both D2D users and interfering devices are mobile, the system achieves a comparatively high throughput. On the one hand, this shows that, although it seemed to be a major contributing factor, the influence of mobility is not significant in terms of performance. Nevertheless, this outcome further validates the algorithm's adaptability and effectiveness in managing interference and maintaining high throughput in dynamic D2D communication scenarios.

Other Findings

An important consideration for the proposed concept is its computational complexity. To this end, several observations on the computational complexity are also carried out. Accordingly, compared to the default system, the proposed algorithm requires approximately three times more simulation time and incurs around 15% additional memory usage. Specifically, the algorithm takes about 1 minute in real time to simulate a 30-second scenario, with memory consumption reaching around 38% of total RAM (16 GB). However, these values remain within a tolerable range, suggesting that the added computational demand is manageable. It is to be noted that in the real-time application of the algorithm, the computational burden would not be too high because, in general, UAM systems are equipped with high-performance computers. On the other hand, the adaptation rate of the algorithm can be changed by introducing a sleep state, such that the frequency of resource allocation occurrence is reduced. Consequently, despite its relatively higher complexity, the algorithm's computational requirements are deemed acceptable for the intended application.

7.6 Conclusion

In terms of UAM, A2A communication involves the technology of D2D supported by 5G communication standards. A2A communication is crucial in UAM for information management, and thus, the enhancement of A2A communication performance is important. 5G D2D communication offers solutions to meet the rising communication demands of today. This helps in reducing associated communication delays, reduces overloading the base station, improves throughput and reliability, etc. However, D2D communication brings several challenges that hinder the system's performance and affect the seamless and reliable communication provision. Two kinds of D2D communication possibilities were presented, and the analysis was conducted only for the network-authorized D2D scenario. In this, the D2D UEs share the uplink or downlink resources of the cellular users, and hence, the system is prone to intercell interference depending on which resources are shared. Therefore, one major challenge in D2D communication is managing interference, which is essential to prevent increased system delay, low sum rate, diminished throughput, and reduced spectral efficiency.

To this end, the problem of interference management was addressed in this chapter. As a solution, different strategies were discussed, and the chapter was confined to focusing specifically on resource allocation strategy as a means of managing interference in D2D communication. Several existing resource allocation techniques employed in literature were analyzed, and a dynamic approach to allocating resources, such as a number of resource blocks, numerology index, and the transmission power using the Expected SARSA RL algorithm, was proposed. The system was evaluated based on key performance metrics, including achieved throughput, SNR, delivery ratio, and delivery delay. Results demonstrated improved system throughput and high reliability. A slight degradation in delivery delay was observed, indicating the need for further analysis of the delivery delay in 5G D2D communication domain.

The influence of mobility and the interfering devices was also considered due to the application being A2A communication for UAM. The algorithm shows adaptability in varying network conditions by performing well even during high interference scenarios, while the influence of mobility was not significant. The computational complexity of the algorithm was also observed to be in a tolerable range, proving its applicability to real UAM scenarios. Overall, the results from extensive simulations confirmed the suitability of applying the Expected SARSA-based dynamic resource allocation algorithm for enhancing performance while achieving low interference in 5G D2D communication scenarios.

Chapter 8

Adaptive Communication Strategies for Air-to-Air Networks

This chapter extends the discussion on Air-to-Air (A2A) communication from the previous Chap. 7. There, the importance of effective interference management (a key research direction identified in the sect. 3.5) was elaborated in detail, with resource allocation as a possible solution. In the current chapter, the aspect of interference management exploration is continued, focusing on factors leading to adaptive communication strategies. Here, adaptive communication refers to dynamically adjusting communication configurations to enhance network performance in response to changing environmental and network conditions. Dealing with interference management, this chapter also addresses the communication protocol stack's physical and medium access control layers. Within the context of A2A communication in UAM, this can involve:

- Assessing Radio Environment - Tracking the signal quality and interference in real-time.
- Choosing the mode of operation - Selecting between the device-to-device (D2D) mode and infrastructure mode (IM) as the mode of communication based on criteria like distance, signal-to-noise ratio (SNR), network load, etc.
- Optimizing resource use - Efficient allocation of spectrum, power, and channels
- Adapting to mobility - Changing connectivity scenarios depending on mobility dynamics, varying altitudes, etc.

The chapter only investigates two of these methodologies leading to adaptive communication planning. These are chosen to suit the continuation of the ongoing discussion on A2A communication in this dissertation. To this end, the chapter is majorly sectioned into two main topics. First, the facilitation of interference detection through tracking signal quality and interference via the *Radio Environment Map* (REM) generation is discussed. Secondly, the aspect of mode selection, highlighted in subsection 7.1.1 as a contributing factor for interference management in D2D communication, is investigated. An algorithm for mode selection is introduced by considering the factors distance, SNR, and achievable bits on a single resource block. Furthermore, a power control mechanism based on the Expected SARSA reinforcement learning algorithm is also presented in conjunction with mode selection. To summarize, the contributions of the chapter are as follows.

- **Challenge Addressed:** Interference management in D2D communication in application to A2A communication in UAM
- **Approach:** Adaptive communication network planning for A2A communication
- **Solution:**
  1. REM generation
  2. Mode selection and power control

Thus, the chapter covers both these topics, each with a discussion of the background and overview of existing literature. Then, the simulation implementation and experiment details are discussed. This is followed by a result analysis of the cooperative effect of mode selection and power control as a solution to the problem of interference management in A2A communications.

8.1 Radio Environment Maps

Radio Environment Maps (REMs) are essential tools in modern communication systems, including 5G and beyond. They comprehensively represent the radio environment, providing valuable information for enabling more intelligent network management and network resource usage enhancements [CBSZM18]. REM is generally considered a framework for representing spectral allocation, commonly applied in monitoring, accessing, and sharing spectrum resources [ZZX+22]. The primary uses of REM in communication can be elaborated as follows:

- Spectrum Management - Usage of REM data for dynamic spectrum resource allocation and enabling efficient spectrum sharing by identifying unused frequency bands
- Interference Management - Provision of interference maps to detect and mitigate interference sources
- Network Planning and Deployment - Assisting in base station and other network infrastructure placement to enhance coverage and capacity
- Quality of Service (QoS) Enhancement - Continuous monitoring of the radio environment to ensure that QoS meets user expectations.
- Mobility Management - Managing handover decision-making of mobile networks by providing up-to-date information on signal strength, interference, and network topology.

As such, REMs are essential for enhancing the performance and efficiency of modern communication systems. In fully autonomous UAM, the operational environment is highly dynamic due to obstacles or interference from other devices. Furthermore, the communication resources are limited, necessitating efficient resource management. For this purpose, the accessibility to real-time information on channel conditions, including used frequency bands and other parameters, is essential. Dynamically changing the communication parameters to reflect real-time channel conditions ensures reliable communication performance and contributes to the overall safety and efficiency of the UAM system. In this regard, generating REMs is advantageous for the overall UAM communication system.

8.1.1 Related Work

REMs act as an enabler for cognitive radio networks, and a lot of work done in the field can be found in the literature. Cognitive radio networks are advanced intelligent communication systems designed to optimize a communication system's spectrum utilization. These networks are supposed to learn and adapt the operating parameters to enhance overall network performance by employing cognitive capabilities [BCG+12]. Cognitive radio devices are expected to expand communication spectrum access by identifying unused portions of licensed bands. They monitor primary licensed user activity on these bands and ensure the seamless operation of secondary unlicensed users on the same licensed band by leaving the band when needed to avoid disrupting primary users. This principle is called dynamic spectrum access (DSA) [AY18]. Additionally, they adapt transmission power and modulation schemes to prevent interference with primary users on adjacent channels [NA14]. As a result, to improve the efficiency of the DSA process, the usage of REM to keep the information up-to-date regularly was suggested by authors of [ZLR06]. REM is, therefore, considered a data source for dynamically collecting information about the radio environment and facilitating the cognitive radio network to initiate the adaptations to suit better spectrum utilization [BCG+12].

Focusing on REM generation, different techniques have been proposed based on different analytical models or information availability. Accordingly, there are two types of REM generation, *direct methods* and *indirect methods* [DPR+11]. In the direct method, the data is collected from different points in the area, and estimating the values pertaining to any point through interpolation is done [AY18]. Spatial interpolation is the commonly utilized technique in direct methods to estimate the values for the interpolated points [DWL22]. Classical spatial interpolation techniques such as Kriging [DKG11], thin plate splines [MBG09], and nearest neighbor [BPRAS11] etc., have been applied in previous work. In the case of indirect methods, they first determine the locations and key parameters of transmitters. These parameters are then used with a propagation model to predict signal levels across the target area. These methods are highly dependent on localization techniques [AY18]. The authors of [AY18] address indirect REM generation methods where they compare two transmitter parameter estimation methods based on received signal strength difference and received signal strength (RSS). Regarding direct REM generation, [DHJA17] introduces an architecture for the REM generation platform, including a distributed set of sensor nodes and the central processing unit. This work is conducted focusing on spectrum sharing in 5G networks.

In [BCG+12], a layered REM architecture is proposed to cater to a heterogeneous LTE network consisting of macrocells and femtocells. The architecture is referred to as *FARAMIR's layered REM architecture* and includes a measurement device, data storage and collection unit, and an REM management unit that generates the REM. One of the main challenges addressed here in relation to cellular communications is reducing the associated signaling overhead, as REM management contributes to increased signaling overhead in the system. As a solution, the authors introduce an entity: time of validity for REM data. To this end, the statistical characteristics of the raw data would be held in the system longer and shared while the raw data would be only available locally. As a layered architecture, the REM components are placed in different network layers or terminals of the communication system. A layered architecture is essential to address the limitations of certain network nodes, such as restricted computation power or memory, and to avoid unnecessary data dissemination. REM information is managed at the smallest or least centralized nodes within the hierarchy to optimize efficiency.

Due to the importance of REMs in cognitive radio systems, means of maintaining up-to-date REMs are often discussed in the literature. One such example is the work of [NA14] aiming at frequency reuse in geo-space. The main challenges in this have been identified as the QoS provision and coexistence alongside other licensed users and cognitive radio devices. In realizing this, the authors propose to work in conjunction with information from an REM, which works as a supportive network infrastructure. That is, the cognitive radio network would be making transmission decisions based on the REM, and thus, the accuracy of the information provided by the REM must be very high. To achieve this, it is proposed that the REM be updated in real-time by employing high-fidelity dynamic radio propagation predictions and incorporating readings from cognitive radios and sensors via pilot channels. Here, the REM is expected to create a real-time heat map or signal strength map, including the location, antenna positions, equivalent isotropically radiated power, and antenna patterns. By accurate predictions on primary user transmissions made by the REM, the coexistence of the cognitive radio on the licensed spectrum can be achieved while eliminating the hidden node problem.

8.1.2 REM Generation

As seen from the previous subsection, REM generation is widely used, aiming at better spectrum resource utilization in cognitive radio systems. However, in this chapter, the REM generation is not addressed as a means of spectrum management but as a common entity that can aid in multiple communication requirements. In terms of an A2A communication network in a UAM system that is fully automated, REM data can be helpful in many aspects ranging from interference management, network planning, and QoS enhancement. It is to be noted that while higher spectrum utilization is very important given the limited resource availability, the chapter rather looks into seamless communication establishment. The decision-making process in an autonomous UAM system depends entirely on communication between network entities. In light of that, the main focus of this chapter is addressing interference management in A2A communication.

For proper interference handling, up-to-date knowledge of the communication network behavior is useful. When communication resource allocation or parameter selection is done, this can reflect the current communication network status, leading to improved system performance and reduced communication interruptions. For this reason, continuous availability of real-time network performance data is required, and REM facilitates this. Accordingly, in this work, REM acts as an information source or database that keeps track of the current network conditions. REM can be accessed and used by various modules of the communication system. This way, the REM data can assist in enhancing communication system performance, specifically leading to interference management.

In the REM architecture discussed in this chapter, the data collection is centrally handled by a global entity, and the REM generation is carried out separately at the ground control station. To this end, the REM generation procedure followed here is also a layered architecture similar to that in [BCG+12]. The REM generation and distribution procedure associated with this chapter is elaborated in Fig. 8.1.

One of the many possible applications of REM data usage is the main focus of this work, which is the mode selection procedure. That is the decision-making on whether the communication is to be conducted via direct communication between the air taxis or through the infrastructure network. This decision is made inside the gNB for the D2D user pairs that request resource allocation in the network-assisted D2D communication scenario. When making the decision on mode selection, having prior knowledge of the current status of the network is advantageous. Depending on the communication mode, the transmission parameters can be adjusted for better system performance. On the other hand, if the performance keeps degrading, the mode can be switched, expecting better performance. This requires accessibility to real-time network conditions, and this can be facilitated through the REM. The aspect of mode selection, in combination with the power control, is the main objective of this work, as highlighted at the beginning of this chapter. This will be discussed later. REM is, therefore, only a secondary outcome of the chapter, which is introduced into the UAM communication system as a tool for real-time network data provision. The implementation is done with the ability to include more data as necessary and use it in any suitable application in the scope of UAM communication or communication networks in general without restricting it only to mode selection.

The existing D2D communication example from Simu5G is considered the network scenario for implementing the REM generation procedure. In the module architecture, the REM data collection is done by a module located inside the 5G stack of the 5G gNB. Although further usage of the REM generated in these simulations is not discussed in this section, as the chapter deals with the mode selection procedure, the data collection was done in conjunction with this. As highlighted before, mode selection is handled by the gNB, and hence, mode selection is a module inside the 5G protocol stack of the gNB. The required data collection for REM generation is thus handled through the same mode selection entity. This extends the possibility

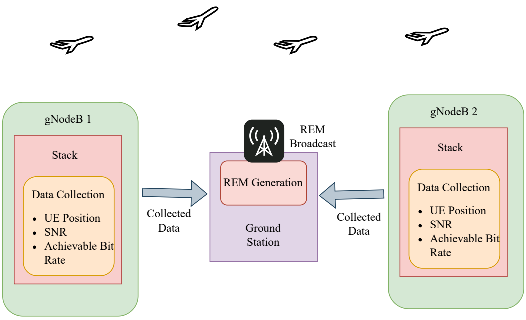


Figure 8.1: REM generation and distribution procedure

of further usage of REM and the reuse of data, as the mode selection algorithm is also required to gather local D2D user performance data for mode selection decision-making. With this regard, the following data are considered as REM measurements:

- Position coordinates of the users
- Signal-to-noise ratio
- Achievable bit rate

These data types are selected based on the requirements of the mode selection algorithm. However, this system has the flexibility for future extensions to accommodate additional data as necessary to suit any considered application. When a user Equipment (UE) enters a network area, it scans for the available cells for connection establishment. In response to this, gNB assigns resources and configures the connection. Therefore, the gNB has information about the UEs that it serves at any given point. This enables the possibility of relevant information gathering. Through channel state information exchanges or other reference signal measurement exchanges, the gNB can infer the location information of the UE. The SNR and the achievable bit rates are measured when assessing communications. This data is, therefore, collected and stored in the gNB periodically for all the UEs belonging to it. This data collected within each gNB is subsequently transmitted to a central REM generation entity located at the ground control station. The REM generator inside the ground station aggregates data from all gNBs to create a comprehensive REM for the entire network. This way, all the contributing UEs and network conditions are included in the REM. Unlike traditional REM formation approaches, the map discussed here is not visualized as a heat map but stored as a data structure containing position-specific network information.

This REM generation is a periodic procedure. Once the REM is generated, it is broadcast across the network to be used by UEs, such as air taxis, to aid with communication-related decisions. gNBs also leverage this REM for decision-making processes, including mode selection. Unlike previously discussed systems, this method does not rely on spectrum-based RSS measurements. Instead, it focuses on collecting communication-related data associated with different users at various locations. Consequently, interpolation techniques are unnecessary. The implementation of REM is completed up to broadcasting the generated REM to the network and possible to be included in making decisions leading to adapting network conditions.

8.2 Mode Selection and Power Control

The challenge addressed in the chapter is interference management in D2D networks in application to A2A communication. As a solution, mode selection is suggested as the main contribution of this chapter. Mode selection in this context refers to the communication mode, that is, whether the communication occurs through the D2D mode or the IM. In general, this decision is taken by the base station responsible for allocating resources to the UE based on signal quality checks. Different algorithms to carry out mode selection can be applied, and through this work, an algorithm that considers the following parameters is discussed.

- Proximity of D2D pair
- Signal-to-noise ratio
- Achievable bit rate

In addition to introducing an algorithm for mode selection, the section further extends into handling the transmit power of the UE. Transmission power is a primary contributing factor for interference in the system, and thus, proper power management is required to reduce the interference. However, the system performance in terms of achievable throughput also depends on the transmission power. Therefore, having a very low transmission power is also not ideal. Hence, this trade-off needs to be analyzed when determining the applicable transmission power. To this end, a power control mechanism based on the expected SARSA reinforcement learning algorithm is proposed here to facilitate this.

8.2.1 Mode Selection

Mode selection in D2D communication is important in determining the interaction method between devices to enhance performance, optimize resource utilization, and reduce interference. In D2D capable systems, devices can communicate directly with one another or through a central entity, such as a base station or access point. Device proximity, channel conditions, interference levels, and the availability of network resources influence the mode selection process. D2D systems can achieve higher data rates, lower latency, better spectrum efficiency, and improved interference management by making the appropriate mode selection decisions. This capability is particularly important in modern wireless networks, especially in dense urban areas and advanced technologies like 5G.

As explained previously in Sect. 7.1.1, depending on the resource usage direction (Downlink or Uplink), various intercell interference conditions are associated with a network-assisted D2D communication system. To this end, mode selection is one of the main factors impacting how the interference is distributed and controlled within the network. On the one hand, by enabling D2D mode between devices in close proximity, the transmissions do not require high power to reach the destination. This is as opposed to the traditional cellular links that rely on higher power to reach the base station. This way, the transmission power needs are reduced, leading to the degradation of the interference caused between the neighboring devices. On the other hand, if the D2D link is experiencing high interference, mode selection enables switching the communication to the IM, leveraging the communication through the base station, mitigating interference, and enhancing network performance.

Therefore, mode selection is a design aspect of cellular network-assisted D2D communications and is made adaptively in the planning and resource allocation stage considering proximity, network load, channel, and interference conditions [FDM+12]. Handling the mode selection is, therefore, a task of the network, meaning the base station is responsible for deciding on communication mode as it has access to the UE information, interference levels, channel and buffer conditions, traffic status, etc. When dealing with mode selection, two main important factors must be considered [FDM+12].

1. Determining the appropriate timescale for mode selection and channel quality assessment: Rapidly changing radio conditions require a highly adaptive approach, while minimizing measurement and signaling overhead is essential to maintain system efficiency.

2. Identifying suitable measurement and reporting strategies, along with hybrid algorithms combining periodic and event-triggered approaches:

To enable effective decision-making between direct D2D links and conventional cellular links.

Accordingly, [FDM+12] provides three alternatives of mode selection techniques suitable for LTE in which three different approaches are applied by the base station to allocate mode and resources to the UE in three different time scales. The same techniques can be applied in 5G. Therefore, these three alternatives are explored first. In the alternative one, referred to as *ALT1*, the time scale is associated with LTE's fixed scheduling and transmission time interval (TTI), which is 1 ms. That is, the base stations collect channel measurements during the 1 ms period and apply the mode selection and resource allocation periodically at the TTI. However, this mode is associated with a scaling issue as the signaling overhead grows proportionally in terms of channel measurements when the number of D2D pairs increases.

As a solution, in the second alternative, referred to as *ALT2*, a time scale of hundreds of milliseconds is considered. Two variations of ALT2 are introduced as ALT2/A and ALT2/B. In A, a pool of resources is allocated to the UEs, while in B, dedicated resource allocation for uplink and downlink data streams is done. In the third alternative, *ALT3*, an even longer time scale is allowed for decision-making, further reducing the signaling overhead. Here, a dedicated resource reservation for a set or all associated D2D pairs is made. As a result, medium access mechanisms are required to access the resource pool and maintain uninterrupted services. The three alternatives according to [FDM+12] are depicted in Fig. 8.2. It is also to be noted that hybrid versions of these alternatives are also possible by mapping time scales and allocation techniques of different alternatives and versions.

While being beneficial in interference management, mode selection also contributes significantly to enhancing the overall performance of a D2D communication system in several other ways. This is because mode selection leads the network to adapt the communication method based on current network conditions, ultimately enhancing the overall performance. First, spectral efficiency is improved through mode selection by allowing for better spectrum utilization. By selecting the communication mode, whether direct D2D or through the IM, users can avoid interference with neighboring devices and make more efficient use of the available bandwidth. Secondly, throughput or the achievable bit rate is enhanced through mode selection by reducing interference. When devices use the D2D mode instead of using the base station, the link quality is often better due to the reduced distance and lower transmission power requirements. Interference reduction helps in overall throughput enhancement. Moreover, mode selection also enables dynamic resource allocation, facilitating high-priority applications and further

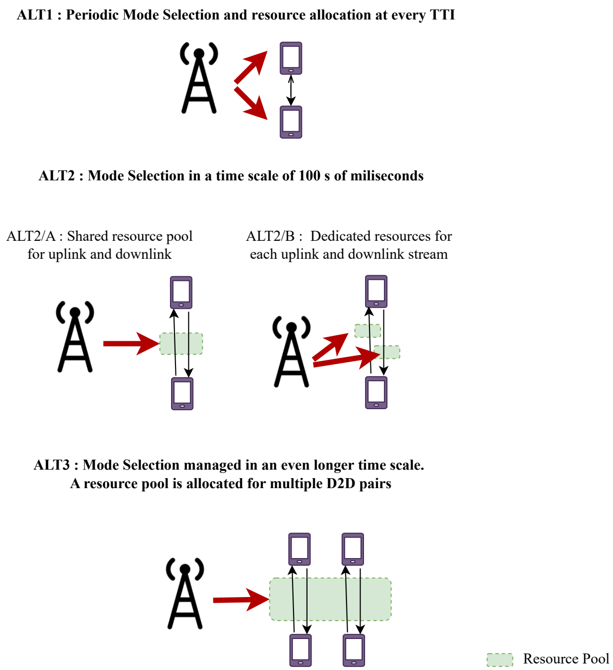


Figure 8.2: Mode Selection and resource allocation alternatives according to [FDM+12]

improving throughput.

Mode Selection also leads to improved energy efficiency in the network. By switching to D2D mode when the devices are in close proximity, energy consumption can be reduced due to the possibility of applying a low transmission power, as the devices don't need to engage with the base station for relaying data. Additionally, energy is saved by reducing the overhead associated with continuous signaling and resource management, as mode selection allows for fewer retransmissions and more efficient resource usage. Although not relevant to applications such as UAM, which is a high energy consumption application in general, energy efficiency is a significant benefit of mode selection. Moreover, reliability and latency are improved through intelligent mode selection. In scenarios with stringent reliability and latency requirements (such as in ultra-reliable low-latency communications), mode selection can prioritize D2D communication, which can be faster and more reliable than routing traffic through a base station. By selecting the best communication mode based on real-time channel conditions, interference levels, and network congestion, mode selection helps reduce delays. It increases the likelihood of successful transmissions, ensuring a more stable system.

8.2.2 Related Work

Mode selection can be achieved by considering various network factors, and many different algorithms applied for mode selection can be found in the literature. In [XS10], a mode selection technique based on path loss is proposed. The algorithm works by choosing the D2D communication mode if the path loss between the D2D user pair is less than that of the path loss between the users and the base station. Although this approach is simple and easy to follow, the authors highlight that without the involvement of interference coordination, the system performance is not enhanced enough. A similar approach is seen in [OF19], where the path loss is taken into account, and the D2D communication mode is chosen if the two users are in line-of-sight (LOS) within a defined area surrounding the transmitter. Otherwise, the communication occurs through the LOS base station within the considered region. The issue in such systems is that the focus is only on path loss. Thus, it doesn't reflect the channel quality or interference factors when making mode selection decisions.

Authors in [TFK+18] highlight that the mode selection is related to both the distance between the D2D user pair as well as the distance between the user and the base station. A work in line with that is the [EHA14], where authors propose a mode selection algorithm based on biased link quality and distances. The network parameters involved are the bias factor for traffic offload control, D2D link distance, and the distance between the D2D UE and the serving base station. This approach offers advantages by accounting for link distances, inherently contributing to interference mitigation to the cellular uplink as opposed to works like [ZZW+19] where the mode selection decision is based only on the D2D pair proximity, adding a bias in D2D user distance calculation.

In [JYD+09], the mode selection decision is made based on channel quality. The channel state information is considered so that the mode selection and resource allocation decision can be performed through the base station. The authors use the metric sum rate of the D2D connection and the cellular connection to determine the mode of operation. On a similar approach, in [SGAKSS18], RSS is considered as the metric to decide on the mode selection choice. The resource allocation in this work is done by focusing on improving the sum rate while maintaining a low SNR. The algorithm works in such a way that it first checks the distance to see if the D2D pair belongs within the D2D communication range. Then, the RSS is predicted to check if it is lower than the threshold. If yes, the cellular mode is chosen, and if not, the system proceeds with resources and power allocation to D2D transmission. Another example is [DYRJ10], where the mode selection algorithm takes into account the D2D as well as cellular link quality along with the interference.

While these works have only considered one or two parameters for mode selection decision-making, the works in [KSGR19] and [HC20] present mode selection algorithms that consider multiple factors. In most previously elaborated works, the channel quality or the interference factor has been completely ignored. Authors in both [KSGR19, HC20] point out that path loss, RSS, the physical distance between the users, channel quality, and interference all contribute to the mode selection procedure, and thus, the algorithm must constitute these factors. In [KSGR19], a 3-tier heterogeneous cellular network is addressed, which contains macro base stations, small cells, and D2D communication scenarios. The system takes three modes of operation into account:

- Cellular mode
- Sharing mode - cellular and D2D users share the same spectrum resources
- D2D mode

The proposed mode selection algorithm in [KSGR19] prioritizes determining the communication mode between cellular and D2D in the initial phase. The next decision is based on the proximity of the D2D pair. If the pair is not sufficiently close, the cellular mode is selected. For pairs deemed close enough, the algorithm evaluates the direction of resource reuse, distinguishing between uplink and downlink scenarios. In the uplink case, the proximity of the D2D pair to the base station is considered. If the pair is located near the base station, the D2D mode is assigned; otherwise, a sharing mode is selected. Conversely, in the downlink, the proximity of the D2D pair to the cellular user is assessed. If the pair is close to the cellular user, the D2D mode is chosen, while the sharing mode is assigned if the proximity condition is not satisfied.

On the other hand, in [HC20], the channel quality, the distance between the D2D pair, and the distance between the base station and the D2D user pair are considered. In this proposal, the D2D users are first grouped following a location distribution criterion. The base station collects the channel state information, and spectrum resources are allocated according to user requirements. The transmission mode is assigned based on a priority order of communication needs, and the three modes, cellular, D2D, and D2D multiplexing, are considered.

In addition to such work, another direction that can be seen jointly studied, along with mode selection in literature, is the transmission power control. For instance, when devices operate in direct communication mode and are in close proximity, lower transmission power is required compared to cellular mode. This results in extended device battery life and decreased interference across the network. The reduced interference further enhances overall system capacity and improves spectrum utilization [HCLK10]. Accordingly, the research in [CGL13] explores distributed power control and link selection in D2D-enabled cellular networks under power and SINR constraints. An iterative algorithm is introduced to reduce power consumption by temporarily deactivating links when power requirements exceed set limits.

In [HCLK10], the authors address the problem of minimizing total uplink transmitted power while ensuring that each mobile device meets its individual carrier-to-interference ratio (CIR) requirements. This minimization is conducted over the set of power vectors and base station assignments. The authors extend the system equations to incorporate the possibility of D2D communication in time division duplexing, utilizing the same radio resources as cellular uplink operations in frequency division duplexing. This approach aims to jointly optimize each user device's communication mode and transmission power. In [WZZY13], the transmission mode is chosen based on the channel state in the initial stage, and then the dynamic transmission mode adjustment is made based on the service history. The problem addressed is a joint mode selection and resource allocation problem, aiming at system throughput enhancements considering transmit power. A similar work found in the literature is [GST+14]. The authors formulate the joint resource block scheduling, mode selection, and power control for D2D users with the objective of minimizing the total transmit power of all D2D users while ensuring that the target aggregate throughput for these users in a single cell scenario is met. A similar work focusing on a multi-cell scenario is [BFA11].

The work in [NR16] has also contributed to the literature by proposing a distributed joint power control and mode selection algorithm. The focus is on a single-cell scenario, and minimization of aggregated transmit power of all users while achieving the minimum throughput requirements of both cellular and D2D users. The UEs have the capability to choose different transmission modes on their allocated resource blocks, and the power allocation, along with the mode selection, is done based on the total throughput for each user. In this algorithm, each D2D user begins by selecting its communication mode for each assigned resource block based on the current effective interference experienced on those blocks. Following this, the user adjusts its transmit power on the assigned resource blocks according to the selected mode to meet its target aggregate throughput. This process iterates until the transmit power levels and mode chosen for the D2D user reach convergence.

Recently, with the research interest in machine learning and reinforcement learning-based approaches, mode selection has also been studied by applying such algorithms [ICVP21]. One example is [AHID24], where a greedy algorithm is developed to select the transmission mode. This further allows switching between the D2D and relay-aided D2D modes (for D2D connections that exceed the range of direct communication), which are the two transmission modes considered in the work. This is then extended into a channel and power allocation strategy based on Multi-Agent Q-learning. As such, different mode selection techniques have been studied in the context of D2D communication.

8.2.3 Mode Selection Algorithm

The dependability of mode selection on various network factors has been highlighted by prior research work, as elaborated in the previous section. Based on that, in this section, a mode selection algorithm that takes three such parameters into account is introduced. These parameters are:

- Distance between the D2D pair
- Signal-to-noise ratio
- Achievable bit rate

As the main intention of the work is to manage the interference, SNR is considered as one parameter. While reducing interference, achieving enhanced system performance is expected. For that reason, the achievable bit rate is included as a parameter of consideration. In addition to that, the system is further expected to control the transmission power, which will be explained in the upcoming section. As the proximity of the D2D user pair directly impacts the transmission power requirement, it is important to include the distance as a parameter in the mode selection decision.

The implementation is carried out through the Simu5G framework, and a network scenario example available in simu5G is used with modifications to make it flying UEs to replicate air taxis. The reason for this network scenario usage rather than using the ULTRAS simulation environment (see Sect. 4.2.3) is that in ULTRAS, the trajectories are planned to suit the transportation demand in the Hamburg metropolitan region, and D2D communication-specific device pair distances and conditions cannot be guaranteed all the time. Therefore, for the purpose of verifying the algorithm, a rather controlled D2D network scenario is used. To this end, an extension is done to an already available D2D communication scenario in Simu5G.

Simu5G has a built-in mode selection principle that allows the user to enable mode switching. The D2D communication protocol in Simu5G is designed in such a way that D2D users always share the uplink cellular resources. Therefore, the applied mode selection technique, referred to as *mode selection best channel quality index*, takes the achievable bits on a single resource block of uplink direction into account. The mode selection works inside the gNB as a simple module in the protocol stack. As the gNB protocol stack has access to all the information of the UEs in its coverage range, the measurement data acquisition for mode selection is made easy. The mode selection decision is made every 0.1 s to reduce the signaling overhead associated with rapid mode switching. In this work, this already available mode selection technique is extended and modified. Two versions of the algorithm are presented, one with the focus of only using uplink resources for D2D communication as implemented in Simu5G to suit our verification process (See Fig. 8.3) and one algorithm that can be applied to a general case where the UE is allowed to share either of the uplink or downlink cellular resources (See Fig. 8.4).

In the algorithm employed for the experiments in this chapter, as illustrated in Fig. 8.3, the process begins with analyzing the distance between the D2D pair. Here, d_{D2D} represents the D2D pair distance, while $thresh_{D2D}$ denotes the preset threshold value. Preliminary experiments are conducted to determine all threshold values associated with the algorithm. If the distance between the D2D pair exceeds the threshold, the system directly assigns the cellular mode without further checks. Conversely, if the distance satisfies the threshold condition, the algorithm proceeds to compare the SNR of the uplink, SNR_{UL} with the SNR of D2D, SNR_{D2D}. If the uplink SNR is higher, the cellular mode is allocated. Otherwise, the system further checks if the D2D SNR exceeds the allowed threshold, SNR_{thresh}. If the D2D SNR is below the threshold, the cellular mode is chosen. If the SNR of D2D is greater than the threshold, the algorithm compares the achievable bit rates on a single resource block for both the uplink, $bits_{UL}$ and D2D, $bits_{D2D}$. If the uplink achieves a higher bit rate, cellular mode is selected; otherwise, D2D communication is permitted.

On the other hand, in the general algorithm, once the distance comparison is done, the resource reuse direction is checked. In the case of uplink (UL), a

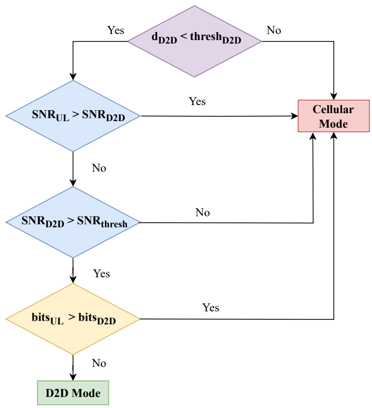


Figure 8.3: Mode Selection algorithm applied in simulation experiments

further distance check is done comparing the distance between the D2D pair and base station, d_{D2BS} and the set threshold of distance, $thresh_{D2BS}$. If the condition is satisfied by distance below the threshold, the same procedure as in the previous algorithm is followed. In case the resource reuse direction is downlink (DL), then the distance from the D2D UE to the cellular user is taken into account in comparison to a threshold. These measurements are denoted as d_{D2CUE} and $thresh_{D2CUE}$, respectively. This is then followed by the same procedure of SNR and achievable bit rate check, with the only difference being in the first SNR check. Here, the D2D SNR is compared against the downlink SNR, which is represented as SNR_{DL} in the algorithm. It is to be noted that the gNB can acquire these distance values by the channel state information or reference signal measurements as well as by utilizing a global positioning system (GPS).

Reasons for the Parameter Check Order

The primary criterion governing the feasibility of D2D communication is the D2D pair distance. If the two devices are not close enough, the path loss incurred is high, making IM the better choice of communication between these devices. Therefore, in the algorithm, the distance between the devices is evaluated first. If the distance threshold is exceeded, then the IM mode is

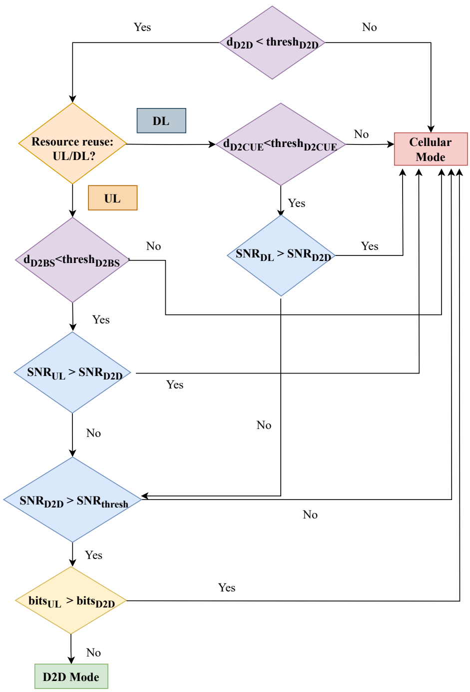


Figure 8.4: Mode Selection algorithm possible to to be applied for a general case where D2D users are allowed to share either uplink or downlink resources

directly assigned at the first step itself. If the distance criterion is met, then the SNR is assessed. This is a metric that indicates the link quality, reflecting on the reliability of the connection. Accordingly, this check verifies if the D2D mode can provide a more reliable connection compared to IM in the specific situation. Finally, if both the distance and SNR conditions are satisfied, the achievable bit rate is evaluated. This is the final factor of consideration, and this determines which of the two modes will enhance the data throughput, reaching the expected degree of performance.

8.2.4 Power Control Mechanism

Once the mode selection is applied and the mode is allocated as D2D communication, a power control mechanism is applied in order to limit the transmit power to degrade the interference. Studying power control in conjunction with mode selection was widely seen in previous works, and several techniques have already been tested and verified. To this end, the chapter contributes by introducing a power control mechanism based on the Expected SARSA reinforcement learning algorithm (see Sect. 7.3.2), which was introduced and applied previously in Chap. 7. The specific components and parameters of the Expected SARSA algorithm are as follows.

Action Space

The action space here is the transmit power that is being applied to the D2D user. The 3GPP Technical specification, TS 38101-2 [3GP24d] attached to release 18 of the 3GPP standards defines the transmit power classes applicable to UEs. Accordingly, vehicular UEs belong to power class 2, and the maximum power allowed is 26 dBm. The default transmit power applied in Simu5G is also 26 dBm. Based on this, seven action values are defined for the action space in the power control reinforcement learning algorithm. The action space:

$$\text{Action Space} = \{20, 21, 22, 23, 24, 25, 26\}$$

State Space

Simply considering a single parameter is insufficient to represent the state space effectively. The state parameter should reflect the multidimensional nature of the environment and the complex interactions between variables that can influence decision-making. For this reason, as the state space for this algorithm, a combination of two parameter values, *Transmit Power* and *Throughput*, is chosen.

Transmit power plays a key role in determining interference levels as it directly affects the SNR and overall system performance. Including it in the state ensures the algorithm has a starting point for adjusting power levels and can respond to changes in the environment or communication needs. Throughput reflects how effectively the system delivers data. By considering throughput in state space, the algorithm can optimize power settings to balance interference control with strong system performance, ensuring efficient data transmission. The throughput calculations are done following the Shannon capacity theorem given by Eqn. 4.2 The limits for upper and lower thresholds for throughput and power calculations are done based on initial experiments. Considering how the high and low power and throughput scenarios could affect the system, three states are defined as inefficient, efficient, and priority on performance scenarios:

- State 0 - Inefficient scenario:
  1. Power: High (Power > 24 dBm)
  2. Throughput: Low (Throughput < 130 Mbps)
- State 1 - Efficient scenario:
  1. Power: Low (Power < 22 dBm)
  2. Throughput: Medium (130 Mbps < Throughput < 140 Mbps)
- State 2 - Priority on performance scenario:
  1. Power: Medium (22 dBm < Power < 24 dBm)
  2. Throughput: High (Throughput > 140 Mbps)

Reward Function

As the reward function defines the objective of the learning process, a reward function consisting of the same two parameters of throughput and the transmit power is defined as in Eqn.8.1.

$$\text{Reward Function (RF)} = \frac{\text{Throughput}}{\text{Transmit Power}} \tag{8.1}$$

By choosing this ratio as the reward function, the reinforcement learning agent learns to balance the system performance with power usage. Throughput ensures that the reward reflects the system's performance in delivering data effectively. Incorporating power in the denominator discourages excessive power usage, promoting efficient operation without compromising communication performance. Moreover, lower power levels mitigate interference with other users in the network, improving spectral efficiency. To this end, using this reward function encourages the agent to choose actions favoring low interference while maintaining adequate throughput performance.

Accordingly, considering the possible value expectation for the reward function, two upper and lower threshold values are defined, and three reward values are allocated, as shown below.

$$Reward(R) = \begin{cases} +10 & RF \geq UpperThreshold \\ 0 & LowerThreshold \leq RF < UpperThreshold \\ -20 & RF < LowerThreshold \end{cases}$$

If the reward function output produces a value higher than the upper threshold, a reward of +10 is given to reinforce this desirable outcome. If the value falls between the upper and lower thresholds, no reward or penalty is applied ($R = 0$), as this range is considered acceptable but not exceptional. However, if the reward function output drops below the lower threshold, it incurs a penalty of -20 to discourage such poor performance. These specific reward values are determined through experimental testing to fine-tune the algorithm to achieve a balance between its actions and overall system performance.

Parameter Based Exploration

Once again, the parameter ϵ that governs the balance between exploration and exploitation is made dynamic to avoid the potential bias in the system, the same as in the Expected SARSA application of Chap. 7. Similarly, the environment keeps changing in this scenario, so the ϵ must also be changed to favor exploration (higher ϵ) or exploitation (lower ϵ) accordingly. To this end, the parameter ϵ depends on the throughput level, and the ϵ value is varied between 0 and 1 dynamically during the Expected SARSA learning algorithm.

Other Parameters

The following parameters are set.

- learning rate $\alpha = 0.5$
- discount factor $\gamma = 0.99$

Learning Model

The Expected SARSA algorithm for power allocation also runs in the same module that carries out the mode selection inside the 5G protocol stack. Accordingly, every 1 s, if the communication mode is chosen to be D2D, the system applies the Expected SARSA reinforcement learning algorithm to determine the applicable transmit power. The system initializes with state 0, and the default power value of 26 dBm from Simu5G is applied in the initial state. In Simu5G, the D2D transmit power value is set in the physical layer function of the D2D UE. Therefore, after running the algorithm (Alg.3) and determining the power value, the physical layer of the UE is updated with this value, where it then sets the applicable new transmit power for the transmission continuation.

8.2.5 Analysis of Results

This section presents the experiment scenarios employed to verify the mode selection and power control mechanisms discussed above, followed by an analysis of the results.

Experiment: Initial experiments to set threshold values to the mode selection algorithm

As the first set of experiments, a simple D2D user scenario with a D2D transmitter and a receiver with a base station in the network is considered, as shown in Fig. 8.5. The D2D pair is set to be stationary. Initially, the D2D pair is set at a distance of 50 m from the gNB, and the distance between the user pair is set to be 20 m. Next, the distance between the D2D user pair is varied, and the achievable SNR in each case corresponding to the distance between the gNB and the user pair variation is recorded until a maximum D2D pair distance of 600 m. In the case of uplink, the SNR for each device is recorded, while in D2D, the values from the transmitting device are considered. In terms of D2D pair distance, as the network environment is 3D, both lateral and vertical distance differences are considered. Accordingly, the SNR threshold is set to be 10 dBm, and the D2D distance threshold is set to be 300 m.

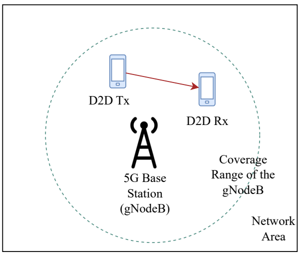


Figure 8.5: Network setup for initial experiment

During this set of experiments, the following observations regarding the network environment are made:

- For the lateral distance variation in D2D pair:
    1. Upto a distance of 400 m D2D pair distance, D2D SNR > Uplink SNR

2. For the entire 600 m distance gap, the recorded SNR values are positive.

- For the vertical distance variation in D2D pair:
  1. Up to a distance of 300 m D2D pair distance for heights of the D2D users more than 100 m from ground level, D2D SNR > Uplink SNR
  2. Up to a distance of 400 m D2D pair distance for heights of the D2D users less than 100 m from ground level, D2D SNR > Uplink SNR

- Between lateral and vertical distance changes, SNR values vary in the same range for similar distance values. Therefore, altitude and D2D pair distance need not be separately considered as parameters in mode selection criteria.

- Regardless of whether the two devices are at equal distances from the gNB or not, for the same distance between two D2D users, the same SNR value is recorded because, for the D2D communication, the distance between the base station and the user pair is irrelevant.

- Increasing the height of the UE pair from the ground level doesn't affect the D2D communication performance significantly. However, in the IM communication scenario, this height influences the performance because the SNR increases as the LOS conditions come into play. This is due to the fact that the D2D communication depends solely on the communication distance between the pair.

***Experiment*: Spectral Efficiency Analysis**

The decision to operate in either D2D mode or IM in D2D communication significantly influences how effectively the available spectrum is utilized. This characteristic, known as spectral efficiency, represents the amount of data that can be transmitted over a given bandwidth within a specific time interval. Therefore, spectral efficiency is a critical metric for evaluating the performance of communication systems and is mathematically defined as shown in Eqn. 8.2 and measured in bits/s/Hz.

$$SpectralEfficiency = \frac{Throughput}{Bandwidth} \tag{8.2}$$

By using an advanced mode selection technique, the system interference can be degraded, signal quality can be enhanced, and better resource allocation

Parameter	Value
UE height	200 m
gNB height	25 m
Carrier frequency	2 GHz
Bandwidth (w)	10 MHz
gNB Tx power	40 dBm
Number of D2D users	2
Mobility	Linear Mobility
UE speed	30 mps
Message size	1000 Bytes
Number of simulation repetitions	30

Table 8.1: Network Parameter values for spectral efficiency evaluation

can be facilitated, all of which contribute to improved spectral efficiency. To this end, in order to verify the mode selection algorithm presented in this chapter, an experimental spectral efficiency analysis is done. The used experimental setup is the same as in Fig. 8.5 with a single D2D pair. Mobility is enabled for the pair using linear mobility. The UE height is set to be 200 m to replicate the UAM scenario. The parameters according to Tbl. 8.1 are applied in the simulation.

Here, the proposed algorithm's performance is compared against the default mode selection algorithm available in Simu5G. Moreover, the effect of the proposed power allocation technique is also taken into account. Thus, the results represent three scenarios: mode selection algorithm of Simu5G (*MS default*), proposed mode selection algorithm without power control (*MS proposed*), and proposed mode selection algorithm with power control (*MS proposed with PC*). Thirty simulation repetitions each are done, and the spectral efficiency results for the 95% confidence interval are presented in Fig. 8.6.

The D2D communication here reuses uplink resources as expressed previously in the chapter, and in this case, the spectral efficiency is degraded due to the interference, but it will still be better than purely cellular systems. In Fig. 8.6, it is seen that by using the proposed mode selection scheme, the spectral efficiency is only improved minutely. However, with transmission power control, a significant increase in the spectral efficiency is observed. This is due to the fact that, by controlling the transmission power, the interference is managed, leading to better spectrum utilization. Therefore, these results verify the mode selection and power control algorithm's suitability to be applied for the discussed D2D communication scenario to mitigate interference and achieve high performance.

Experiment: Interference Analysis

In the next set of experiments, cellular users are introduced into the network to analyze the interference conditions. The network setup is visualized in Fig. 8.7. To focus on the impact of D2D users on system interference, the D2D users are configured to remain stationary, ensuring that D2D communication is consistently maintained. Similarly, the cellular users are also kept stationary. Notably, Cellular User 2 (Cellular UE2) is deliberately positioned far from the gNB, although still within the gNB's coverage area, while Cellular User 1 (Cellular UE1) is placed closer. This setup ensures that communication involving these cellular users does not occur via the D2D communication mode. The other network parameters are kept the same as elaborated in Tbl. 8.1.

It was proven from the previous experiment that the mode selection, along with the power allocation mechanism, provides better system performance. Therefore, in this section of experiments, the results are compared only between the default mode selection technique in Simu5G and the new mode selection technique with power allocation. For this reason, in the remainder of the chapter, *proposed MS Algorithm* refers to the proposed algorithm with power allocation technique even if not explicitly mentioned. All three uplink, downlink, and D2D SNR measurement recordings are taken from the gNB to evaluate the interference in the system. The average SNR results for the 95% confidence interval recorded for 30 repetitions of simulations are presented in Fig. 8.8. Given the D2D pair is utilizing uplink resources, according to Fig. 7.3, the most crucial interference scenario here is the uplink interference

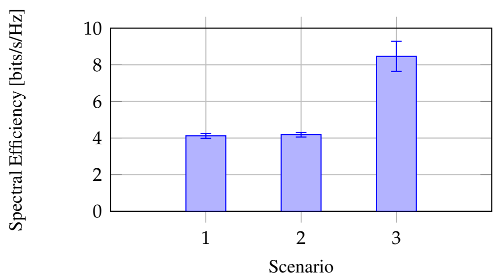


Figure 8.6: Comparison of performance of spectral efficiency for the three scenarios: 1 - MS default, 2 - MS proposed, and 3 - MS proposed with PC

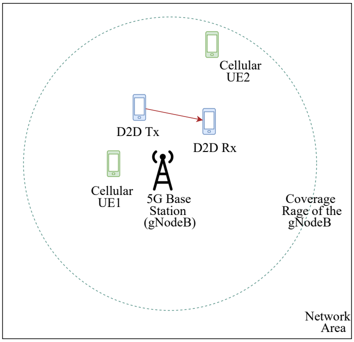


Figure 8.7: Network setup with cellular users

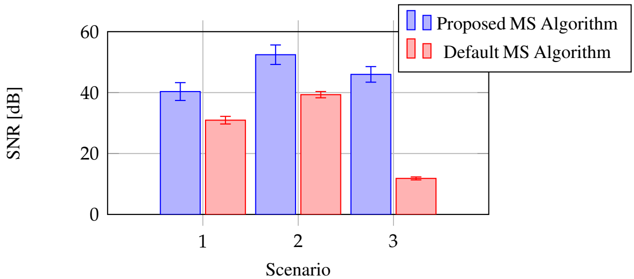


Figure 8.8: Comparison of signal-to-noise ratio in the system in 1 - Uplink, 2 - Downlink, and 3- D2D, when D2D and cellular users coexist in the network

on the gNB caused by the D2D users. The results show that all three uplink SNR, downlink SNR, and D2D SNR scenarios have improved drastically with the proposed MS algorithm, meaning the interference has been significantly reduced by controlling the transmission power.

Experiment: Reliability and Delay Assessment

To further prove the performance enhancements through the proposed MS algorithm, a reliability and delay assessment comparison between the default and the proposed algorithms is done. These results were also captured during the same experiment as before. Therefore, the results are presented for the 95% confidence intervals based on the 30 simulation repetitions. The delivery ratio and delivery delay results are represented in Fig. 8.9 and Fig. 8.10, respectively.

A significant improvement in the delivery ratio and a reduction in delivery delay are observed in the presented results. However, achieving a 97% delivery ratio cannot be deemed satisfactory, as the reliability requirements for systems like UAM communication exceed 99.9%. The results show that the delivery ratio's error bar includes a 2% confidence interval, indicating a more significant variability in data between 95% and 99%. The 95th percentile value of the delivery ratio is also calculated, yielding a delivery ratio of 99.5%. That is, reaching at least a 99.5% of reliability is only guaranteed in 5% of the cases. This indicates that further improvements in terms of delivery ratio are required. For example, retransmission techniques with TCP protocols could improve reliability.

Regarding delivery delay, earlier results from Chap. 7 also highlighted that, in general, D2D communication delay tends to be in the range of 20 ms. Compared with those results and the application of the default mode selection algorithm, the new mode selection algorithm, combined with the power control mechanism, demonstrates a substantial reduction in delivery delay. This is the result of establishing communication through enhanced transmission conditions. By applying the most suitable communication mode and power levels, the signal quality is improved, and fast data transmissions are enabled. Therefore, the new algorithm can be considered to improve both the reliability and delay status of the D2D communication system compared to the default, although there's a need for further improvements.

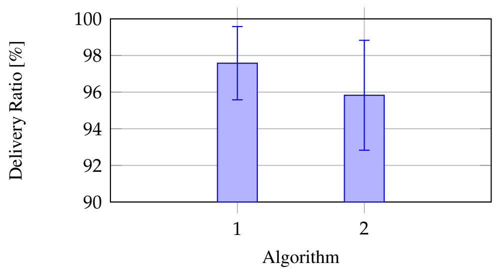


Figure 8.9: Comparison of performance of message delivery ratio with the application of: 1 - Proposed MS Algorithm and 2 - Default MS algorithm

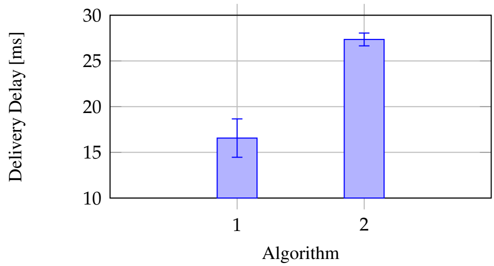


Figure 8.10: Comparison of performance of message delivery delay with the application of: 1 - Proposed MS Algorithm and 2 - Default MS algorithm

Experiment: Application

Based on the previous results, the viability of the new mode selection and power allocation algorithm has been proven compared to the default algorithm. In this section, to demonstrate the algorithm's behavior in a dense urban environment to replicate the UAM scenario, a multi-pair D2D user scenario in the presence of multiple cellular users is implemented. This is because, in general, for UAM user scenarios, a large number of air taxis occupy the air space simultaneously. Mass mobility is applied to all UEs instead of linear mobility to bring in further realistic conditions. Mass mobility is considered because it models the random motion behavior of an object with a mass, which suits the UE in the considered application, which is an air taxi. The ULTRAS trajectories are not considered here to enable the maximum possible D2D communication for testing purposes. The network setup extends Fig. 8.7 to include 10 D2D user pairs and 10 cellular users. Cellular users are placed far apart to avoid enabling communication through D2D mode. Both user types are placed at an altitude of 200 m to replicate the UAM scenario. The purpose of this setup is to evaluate the D2D communication system in the presence of interfering aerial users. Ground cellular users are not considered for the experiment. Considered parameters are listed in Tbl. 8.2.

Results for the spectral efficiency for each user pair are depicted in Fig. 8.11. It is clear from the figure that, on average, the spectral efficiency measured on each user pair varies around the same range. This indicates stable performance across the users, showing that similar resource and channel conditions are applied to each pair. This means that, due to the application of the new mode selection and power control method, the interference between the neighboring devices is reduced, making more efficient usage of the available bandwidth.

Parameter	Value
UE height	200 m
gNB height	25 m
Carrier frequency	2 GHz
Bandwidth (w)	10 MHz
gNB Tx power	40 dBm
Number of D2D user pairs	10
Number of cellular users	10
Mobility	Mass Mobility
Mobility change interval	0.1s
UE speed	30 mps
D2D transmitter initial movement heading	0 deg
D2D receiver initial movement heading	270 deg
Cellular transmitter initial movement heading	60 deg
Cellular receiver initial movement heading	180 deg
Message size	1000 Bytes
Network area	750m x 750m
Number of simulation repetitions	30

Table 8.2: Network Parameter values for mode selection application

However, the confidence interval recorded is vast due to the high variability in recorded data. The 95th percentile value is also plotted in the same figure. This is an indication that although most results show moderate performance, certain cases achieve significantly higher values. Moreover, the 95th percentile in each case lies closer to the mean, suggesting that most data points are within a reasonable range, ensuring fairness and the high confidence interval must have resulted from the few extreme outliers.

Fig 8.12 demonstrates the SNR results. Values for each pair vary in the same range as obtained before in the SNR comparison between the default and proposed methods (Fig 8.8). Here, the confidence interval is narrow, indicating low variability in SNR, suggesting that well-optimized power control is applied across all users. This supports the behavior of the applied power control strategy based on Expected SARSA and the mode selection technique, contributing to uniform link quality. In addition to that, the reliability and the latency for each D2D pair in the dense user scenario are also assessed. The delivery ratio, the measure of the reliability, is shown in Fig. 8.13. Here, the average delivery ratio lies around 97%, with a narrow confidence interval indicating low variability in data. The 95th percentile value is also plotted for each case in the same figure, resulting in values above 99.5%. The results

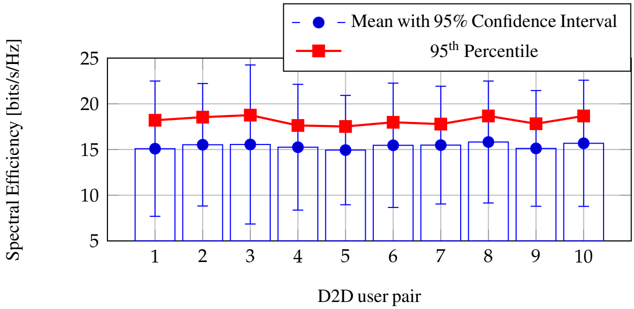


Figure 8.11: Spectral Efficiency results for the 10 D2D user pairs

are identical to the previous comparison between the algorithms given by Fig. 8.9. This implies that the algorithm performs well even in a dense network and does not degrade the system's reliability. Furthermore, with 99.5% of delivery ratio as the 95$^{\text{th}}$ percentile, it suggests that the system can perform very reliably under ideal conditions. However, as suggested earlier in this section, the retransmission mechanism can be combined in order to reach higher reliability, guaranteeing more than 99.9% delivery ratios.

On the other hand, in delivery delay (See Fig. 8.14), except for one user pair, all other users show reasonably low delay values (less than 20 ms) aligning with the results of Fig. 8.10. The slightly high delay on average in the first user pair could be due to communication distance, as all the users are mobile. The confidence interval is not as narrow as in the single-user pair case. However, it is also not a significantly high confidence interval and indicates

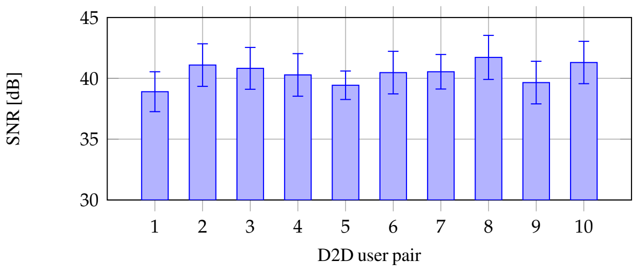


Figure 8.12: SNR results for the 10 D2D user pairs

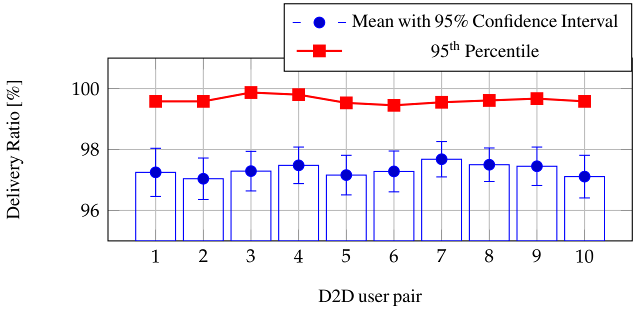


Figure 8.13: Delivery ratio results for the 10 D2D user pairs

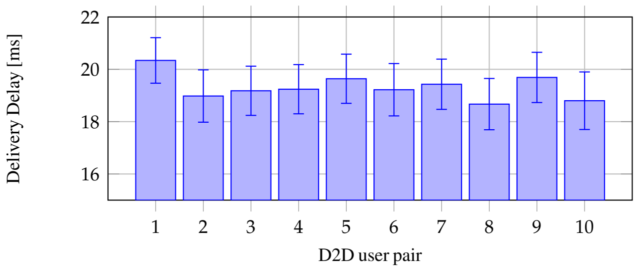


Figure 8.14: Delivery delay results for the 10 D2D user pairs

fluctuations in delay across different transmissions. The consistent delay values among the users suggest that network conditions are well managed even with multiple users, making the system suitable for a dense deployment while still maintaining low latency. In general, all the presented results here verify that, even in a dense deployment of users, the proposed algorithm's performance aligns with the case of a single-pair scenario, complementing the scalability, which is essential in a communication system like UAM. Furthermore, it suggests that, by controlling transmission power, the interference conditions are well-managed, and the proposed mode selection algorithm with the Expected SARSA-based power control technique is thus suitable to be applied in a time-sensitive application like UAM, even with a dense deployment of air taxis.

8.3 Conclusion

This chapter expanded on addressing the interference challenge associated with A2A communication. In line with that, various strategies for reducing system interference in a D2D communication scenario were explored, focusing on adaptive communication strategies. Two key aspects of adaptive network planning were discussed in detail.

First, a Radio Environment Map (REM) was introduced as a tool for providing network-related information to adapt network conditions dynamically. This approach is specially valuable in UAM's A2A communication systems, where autonomous decision-making is crucial. Implementing such an REM was examined as a foundational tool for improving communication efficiency.

Secondly, as the main contribution of the chapter, mode selection was investigated. The communication mode, whether D2D or IM communication is used, significantly impacts system performance. Intercell interference becomes a critical concern as D2D communication reuses uplink resources, as assumed in the simulation experiments. A mode selection algorithm based on the proximity of D2D users, SNR, and achievable bit rate was proposed to address this. A transmission power control technique using the Expected SARSA reinforcement learning algorithm was also introduced to mitigate interference further. Given the proximity of D2D users, the application of lower transmission power becomes feasible.

Simulation results demonstrated that the proposed mode selection algorithm, in combination with the power control technique, can substantially enhance the system's performance. This includes achieving higher spectral efficiency, improved reliability, reduced interference, and lower latency, validating the proposed solution's applicability to an A2A communication scenario in UAM. It was applied to a dense deployment scenario with 10 A2A pairs and 10 aerial cellular users to validate the algorithm further. The results demonstrated that, even with multiple users, the proposed algorithm performs well without significant performance degradation, proving scalability. As established through the simulation results, overall, despite the increased load, the proposed algorithm helps enhance the A2A communication system's performance, proving its suitability in application to a real-time UAM scenario.

Chapter 9

Conclusion

An alternative mode of public transit to cater to the existing high demand for public transportation is UAM. The concept entails employing autonomous aerial vehicles referred to as *air taxis* to move passengers and goods through low-altitude airspace. Nevertheless, such a system implementation requires complex integration of several key technical fields. These include air transportation, guidance, navigation and control, ground infrastructure and demand modeling, and communication that work together to enable efficient and safe UAM operations. To this end, the dissertation explored the key sector of communication and investigated different aspects of communication requirements and enhancements to support efficient and effective UAM systems. Communication plays a significant role since the system is fully autonomous. Air taxi navigation depends on the control decisions based on input data and the interactions between the air taxis. Hence, a reliable, low latent, and redundant communication system is a prominent requirement for UAM. Accordingly, the contributions of this dissertation can be expressed under two main categories: air-to-air (A2A) communication and air-to-ground (A2G) communication. The contributions span multiple protocol layers, including the physical, medium access, and network layers.

In communication establishment, the most critical decision is the choice of communication technology. The 5th generation (5G) cellular communication standard claims to provide ultra-reliable, low-latency communications with enhanced capacity and support for device-to-device (D2D) communications. Alongside the enhanced performance requirements, D2D communication is vital in this setup, as air taxi interactions are supposed to occur through direct D2D links. Therefore, a 5G-based UAM communication system was considered. A simulation-based implementation and verification process was applied in the dissertation. Here, contributions are made regarding a simulation environment for simulating UAM communication capabilities alongside control modeling. The framework, referred to as ULTRAS, is based on OMNeT++ discrete event simulator, and for the 5G protocol stack, the Simu5G library associated with OMNeT++ was applied. The ULTRAS framework and extensions for communication networks available in Simu5G were employed in the implementations. The system performance was analyzed based on reliability, latency, interference, and spectral efficiency metrics.

With the preliminaries being set, the dissertation first analyzed network planning. A cost-effective network implementation approach based on heuristics was proposed. As an application of this method, the Hamburg metropolitan region was considered, and a network architecture to suit UAM implementation in Hamburg was designed. In terms of A2G communication, a channel model to fit the unique operational requirements of UAM communications was proposed based on existing 3GPP standards and previous research. This communication setup was then tested for different scenarios of various data types applicable to UAM communications. Recommendations on network parameters to suit these data types were made. The simulation results indicated that careful positioning of base stations, following geographical conditions and traffic demand, can result in an economical base station network plan yielding enhanced system performance. Results showed the applicability of this network architecture in a real UAM scenario by providing a 97% delivery ratio and 10 ms delivery delay on average.

It is to be noted that in the dissertation, A2G communication is considered as a means of providing redundancy for A2A communication. This is because communication disruptions cannot be tolerated under such critical operational conditions. However, faults in any system are inevitable. In recognition of that, a fault-tolerant distributed ground station architecture concept for A2G communication consisting of a main ground station and several subground stations is proposed. This was then applied to the considered use case of the Hamburg metropolitan region. This dissertation also proposed a role delegation principle for the ground station architecture to further improve communication availability. This way, the communication infrastructure can ensure the maximum possible service provision while achieving a low mean time to repair the UAM communication system in case of failure.

Regarding A2A communication, the dissertation explored the critical issue of interference management in D2D connections. Two aspects leading to interference management were discussed. Firstly, a novel dynamic approach for the resource allocation problem is proposed based on the Expected SARSA reinforcement learning. Here, resources such as the number of resource blocks, numerology index, and transmit power were dynamically allocated. Simulations were conducted under various peer distances, mobility, and interference conditions. The results demonstrated that the dynamic choice of resources based on real-time network conditions leads to an improved system throughput and enhanced reliability ($> 99.8\%$ delivery ratio) while mitigating interference. It is also to be noted that employing a reinforcement learning algorithm in the simulations did not incur a significant computational burden. This shows that the computational complexity should be tolerable even in actual implementations, as aerial vehicles are equipped with high-performance computers. Nevertheless, the associated latency showed an increment compared to the default case. However, it was apparent that even in the default case, the delay was not small enough (> 20 ms) for a D2D system. This indicates the need for a protocol enhancement study in 5G concerning latency, which was out of the scope of the research focus in this dissertation.

Next, for the same interference problem, an adaptive communication strategy was investigated. Here, first, an information provision tool implementation was discussed. Secondly, a mode selection algorithm with a power control strategy was proposed. The information provision tool is the radio environment map (REM), where network-related real-time information mapping can be done. In terms of mode selection, the communication mode between D2D or infrastructure-based communication decisions is made. For this purpose, the new algorithm considers user proximity, signal-to-noise ratio, and achievable bit rate. The power allocation mechanism is based on the expected SARSA reinforcement learning algorithm. The aim is to reduce the transmission power involved as a means of mitigating interference conditions. The results suggest that enabling adaptability through communication mode selection and transmission power control helps achieve efficient and reliable network conditions. The algorithm yielded significantly high spectral efficiency (>8), a 99.5% delivery ratio, and a delivery delay reduction below 20 ms. The algorithm was applied to a dense UAM deployment scenario for further validation. It was seen that, even with multiple users, the new algorithm performed well without significant performance degradation, demonstrating scalability.

Overall, the contributions made through the dissertation indicate that proper network planning, channel modeling, enabling redundant communication modes, and keeping track of the communication network conditions lead to enhanced performance. Dynamically adapting communication mode, resource allocation, and transmit power to reflect changes in the network conditions enhances throughput, spectral efficiency, and reliability while reducing latency and interference. It is to be noted that although the Hamburg metropolitan region was considered the application environment for some of the topics discussed in this dissertation, the proposals made are not confined to Hamburg. These algorithms and methodologies are general solutions for specific challenges in the UAM communications domain and can be applied to any urban area with considerations for a UAM implementation.

As proven through results, the proposed algorithms and methodologies in the dissertation lead to establishing a reliable and redundant communication system for safe UAM operations. Nonetheless, while UAM is a feasible solution, when it comes to actual implementations, a number of improvements and multiple testing scenarios in each involved technical field are required. For example, along with all the communication system modifications, a reliability $> 99.9\%$ must always be ensured. Equal attention must also be paid to the communication security sector to prevent cyber threats that could be critical. Therefore, from a personal point of view, even though the simulation results suggest robust information management, the real UAM implementations are likely to take more than a decade.

Bibliography

[3GP13] 3GPP. Technical specification group services and system aspects; feasibility study for proximity services. Techn. rep. 22803, 3GPP, 2013.

[3GP14] 3GPP. Technical specification group services and system aspects; study on architecture enhancements to support proximity services. Techn. rep. 23703, 3GPP, 2014.

[3GP17a] 3GPP. Study on channel model for frequencies from 0.5 to 100 GHz. Technical report (TR) 38.901, 3GPP, 2017. URL: https://www.3gpp.org/ftp/Specs/archive/38_series/38.901.

[3GP17b] 3GPP. Technical specification group radio access network; study on enhanced lte support for aerial vehicles. Technical report (TR) 36.777, 3GPP, 2017. URL: https://www.3gpp.org/ftp/Specs/archive/36_series/36.777/.

[3GP19a] 3GPP. Release 15 description; summary of Rel-15 work items. Technical report (TR) 21.915, 3GPP, 2019. URL: https://www.3gpp.org/release-15.

[3GP19b] 3GPP. Technical specification group services and system aspects;enhancement for unmanned aerial vehicles. Technical report (TR) 22.829, 3GPP, 2019. URL: https://www.3gpp.org/ftp/Specs/archive/22_series/22.829/.

[3GP21a] 3GPP. Technical specification group services and system aspects; study on application layer support for unmanned aerial systems (UAS). Technical report (TR) 23.755, 3GPP, 2021. URL: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3588.

[3GP21b] 3GPP. Technical specification group services and system aspects;study on supporting unmanned aerial systems (UAS) connectivity, identification and tracking. Technical report (TR) 23.754, 3GPP, 2021. URL: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3575.

[3GP22] 3GPP. Release 16 description; summary of Rel-16 work items. Technical report (TR) 21.916, 3GPP, 2022. URL: https://www.3gpp.org/release-16.

[3GP24a] 3GPP. Study on channel model for frequencies from 0.5 tpo 100 ghz (release 18). Technical report (TR) 38.901, 3GPP, 2024. URL: https://www.3gpp.org/release-18.

[3GP24b] 3GPP. Technical specification group radio access network; NR; air-to-ground network for NR. Technical report (TR) 38.876, 3GPP, 2024. URL: https:

//portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4062.

[3GP24c] 3GPP. Technical specification group services and system aspects;uncrewed aerial system (UAS) support in 3GPP. Technical specification (TS) 22.125, 3GPP, 2024. URL: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3545.

[3GP24d] 3GPP. User equipment (UE) radio transmission and reception; part 2: Range 2 standalone. Technical specifiction(TS) 38.101-2, 3GPP, 2024. URL: https://www.3gpp.org/ftp/Specs/archive/38_series/38.101-2/.

[AHID24] A.M.H. Alibraheemi, M. N. Hindia, T.F.T. M. N. Izam, and K. Dimyati. Spectrum efficient mode selection and resource allocation optimization for D2D communication in HetNet: A multi-agent Q-learning approach. *IEEE Access*, 12:131217–131229, 2024. doi:10.1109/ACCESS.2024.3447471.

[Air24] Airbus. CityAirbus NextGen makes its debut, March 2024. Last visited: 09/10/2024. URL: https://www.airbus.com/en/newsroom/press-releases/2024-03-cityairbus-nextgen-makes-its-debut.

[AJ22] M. Q. Abdulqadir and O. A. Ali Janaby. Capacity and throughput enhancement for minimum distance using D2D scheme with TCP for 5G system. In *2022 8th Int. Conf. on Contemporary Info. Tech. and Mathematics*, pages 34–38, 2022. doi:10.1109/ICCITM56309.2022.10031989.

[All24] O-RAN Alliance. O-RAN work group 1 (use cases and overall architecture) o-RAN architecture description. Technical Specification (TS) O-RAN.WG1.OAD-R003-v12.00, O-RAN Alliance, 2024. URL: https://specifications.o-ran.org.

[ALZ+22] G. Amponis, T. Lagkas, M. Zevgara, G. Katsikas, T. Xirofotos, I. Moscholios, and P. Sarigiannidis. Drones in B5G/6G networks as flying base stations. *Drones*, 6(2), 2022. URL: https://www.mdpi.com/2504-446X/6/2/39.

[AP18] A. Austine and R. S. Pramila. A review on recent research directions in device to device communication. In *3rd Int. Conf. on Inventive Computation Tech.*, 2018. doi:10.1109/ICICT43934.2018.9034350.

[ARP17] M. M. Azari, F. Rosas, and S. Pollin. Coexistence of terrestrial and aerial users in cellular networks. In *2017 IEEE Globecom Workshops (GC Wkshps)*, page 1–6. IEEE, December 2017. URL: http://dx.doi.org/10.1109/GLOCOMW.2017.8269068, doi:10.1109/glocomw.2017.8269068.

[ARP18] M. M. Azari, F. Rosas, and S. Pollin. Reshaping cellular networks for the sky: Major factors and feasibility. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–7, 2018. doi:10.1109/ICC.2018.8422685.

[ARP19a] M. M. Azari, F. Rosas, and S. Pollin. Cellular connectivity for UAVs: Network modeling, performance analysis, and design guidelines. *IEEE Transactions on Wireless Communications*, 18(7):3366–3381, 2019. doi:10.1109/TWC.2019.2910112.

[ARP19b] M. M. Azari, F. Rosas, and S. Pollin. UAV-to-UAV communications in cellular networks, 2019. URL: https://arxiv.org/abs/1912.07534, arXiv:1912.07534.

[ARP20] M. M. Azari, F. Rosas, and S. Pollin. *Performance Analysis for Cellular-Connected UAVs*, chapter 4, pages 103–138. John Wiley & Sons, Ltd, 2020. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119575795.ch4, doi:https://doi.org/10.1002/9781119575795.ch4.

[ASAM+19] A. Almousa, B. Sababha, N. Al-Madi, A. Barghouthi, and R. Younisse. UTSim: A framework and simulator for UAV air traffic integration, control, and communication. *International Journal of Advanced Robotic Systems*, 16:172988141987093, 09 2019. doi:10.1177/1729881419870937.

[ASKC19] S. Alemaishat, O. A. Saraereh, I. Khan, and B. J. Choi. An efficient resource allocation algorithm for D2D communications based on NOMA. *IEEE Access*, 7, 2019. doi:10.1109/ACCESS.2019.2937401.

[ASZS24] G. H. Adday, S. K. Subramaniam, Z. A. Zukarnain, and N. Samian. Investigating and analyzing simulation tools of wireless sensor networks: A comprehensive survey. *IEEE Access*, 12:22938–22977, 2024. URL: https://api.semanticscholar.org/CorpusID:267540624.

[AT19] G.E. Athanasiadou and G. V. Tsoulos. Path loss characteristics for UAV-to-ground wireless channels. In *2019 13th European Conference on Antennas and Propagation (EuCAP)*, pages 1–4, 2019.

[AVK20] Antonio A. Virdis and M. Kirsche. *Recent Advances in Network Simulation*. Springer, 10 2020. URL: https://books.google.com/books/about/Recent_Advances_in_Network_Simulation.html?hl=&id=9ymXzQEACAAJ.

[AY18] S. Alfattani and A. Yongacoglu. Indirect methods for constructing radio environment map. In *2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)*, pages 1–5, 2018. doi:10.1109/CCECE.2018.8447654.

[BCD20] M. Banagar, V. V. Chetlur, and H. S. Dhillon. *Stochastic Geometry-Based Performance Analysis of Drone Cellular Networks*, chapter 9, pages 231–254. John Wiley & Sons, Ltd, 2020. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119575795.ch9, doi:https://doi.org/10.1002/9781119575795.ch9.

[BCG+12] J.V.D. Beek, T. Cai, S. Grimoud, B. Sayrac, P. Mahonen, J. Nasreddine, and J. Riihijarvi. How a layered rem architecture brings cognition to today's mobile networks. *IEEE Wireless Communications*, 19(4):17–24, 2012. doi:10.1109/MWC.2012.6272419.

[BDO+21] A. Baltaci, E. Dinc, M. Ozger, A. Alabbasi, C. Cavdar, and D. Schupke. A survey of wireless networks for future aerial communications (FACOM). *IEEE Comm. Surv. & Tut.*, 23(4):2833–2884, 2021. doi:10.1109/COMST.2021.3103044.

[BFA11] M. Belleschi, G. Fodor, and A. Abrardo. Performance analysis of a distributed resource allocation scheme for D2D communications. In *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, pages 358–362, 2011. doi:10.1109/GLOCOMW.2011.6162471.

[BHW+23] J. Berling, P. Hastedt, S.T. Wanniarachchi, A. Vieregg, C. Gertz, V. Turau, H. Werner, and V. Gollnick. A modular urban air mobility simulation toolchain with dynamic agent interaction. *German Aerospace Congress 2022, Dresden*, Feb 2023. doi:10.25967/570247.

[BJF22] M. Bamdad, S. Jamali, and R. Fotohi. Dynamic allocation of 5G network resources using game theory. 09 2022. doi:10.21203/rs.3.rs-2066340/v1.

[BKT22] I. Budhiraja, N. Kumar, and S. Tyagi. Ishu: Interference reduction scheme for D2D mobile groups using uplink NOMA. *IEEE Transactions on Mobile Computing*, 21(9), 2022. doi:10.1109/TMC.2021.3051670.

[BPRAS11] L. Bolea, J. Perez-Romero, R. Agusti, and O. Sallent. Context discovery mechanisms for cognitive radio. In *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pages 1–5, 2011. doi:10.1109/VETECS.2011.5956366.

[BTTG19] I. Budhiraja, S. Tyagi, S. Tanwar, and N. Kumarand M. Guizani. Cross layer NOMA interference mitigation for femtocell users in 5G environment. *IEEE Transactions on Vehicular Technology*, 68(5):4721–4733, 2019. doi:10.1109/TVT.2019.2900922.

[BV23] N. M. Bilal and T. Velmurugan. A review of efficient resource allocation for mm-wave D2D communication in cellular networks using ML algorithm. In *2023 3rd Int. Conf. on Pervasive Comp. and Social Net.*, pages 1307–1312, 2023. doi:10.1109/ICPCSN58827.2023.00220.

[BZ20] Z. Bi and W. Zhou. Deep reinforcement learning based power allocation for D2D network. In *2020 IEEE 91st Vehicular Technology Conference*, pages 1–5, 2020. doi:10.1109/VTC2020-Spring48590.2020.9129537.

[CBSZM18] V. Chowdappa, C. Botella, J. J. Samper-Zapater, and R. J. Martinez. Distributed radio map reconstruction for 5G automotive. *IEEE Intelligent Transportation Systems Magazine*, 10(2):36–49, 2018. doi:10.1109/MITS.2018.2806632.

[CCB+21] D. Carramiñana, I. Campaña, L. Bergesio, A. M. Bernardos, and J. A. Besada. Sensors and communication simulation for unmanned traffic management. *Sensors*, 21(3), 2021. URL: https://www.mdpi.com/1424-8220/21/3/927, doi:10.3390/s21030927.

[CCL15] J. Chen, D. Cao, and X. Lv. Design of flexray-based communication on triplex redundancy flight control computer. In *2015 Chinese Automation Congress (CAC)*, pages 1969–1974, 2015. doi:10.1109/CAC.2015.7382827.

[CGBR+21] Z. Cui, K. Guan, C. Briso-Rodríguez, B. Ai, Z. Zhong, and C. Oestges. Channel modeling for UAV communications: State of the art, case studies, and future directions, 2021. URL: https://arxiv.org/abs/2012.06707, arXiv:2012.06707.

[CGL13] Y. Cheng, Y. Gu, and X. Lin. Combined power control and link selection in device-to-device enabled cellular systems. *IET Communications*, 7:1221–1230, 2013. doi:10.1049/iet-com.2012.0769.

[CHCX21] L. Chen, B. Hu, S. Chen, and G. Xu. QoE-driven resource allocation for D2D underlaying NOMAcellular networks. In *IEEE wireless Com. and Net. Conf.*, 2021. doi:10.1109/WCNC49053.2021.9417558.

[CLS+19] H. Chang, L. Liu, H. Song, A. Pidwerbetsky, A. Berlinsky, J. Ashdown, K. Turck, and Y.Yi. Maximizing system throughput in D2D networks using alternative DC programming. In *2019 IEEE Global Com. Conf.*, pages 1–6, 2019. doi:10.1109/GLOBECOM38437.2019.9013396.

[CO15] G. Chengetanai and G.B. O'Reilly. Survey on simulation tools for wireless mobile ad hoc networks. In *2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pages 1–7, 2015. doi:10.1109/ICECCT.2015.7226167.

[CRPY+19] X. Cai, J. Rodríguez-Piñeiro, X. Yin, N. Wang, B. Ai, G. F. Pedersen, and A. P. Yuste. An empirical air-to-ground channel model based on passive measurements in lte. *IEEE Transactions on Vehicular Technology*, 68(2):1140–1154, 2019. doi:10.1109/TVT.2018.2886961.

[CSF21] P. A. Cohen, A. S. Shaheen, and M. E. Farrar. Urban air mobility: History, ecosystem, market potential, and challenges. *IEEE Transactions on Intelligent Transportation Systems*, 22(9):6074–6087, 2021. doi:10.1109/TITS.2021.3082767.

[DA21] D.Mandloi and R. Arya. Seamless connectivity with 5G enabled unmanned aerial vehicles base station using machine programming approach. *Expert Systems*, 39, 09 2021. doi:10.1111/exsy.12828.

[DdFE16] D. Dalimir, E. P. de Freitas, and F. Eliassen. Self-organization as a supporting paradigm for military UAV relay networks. *IEEE Communications Letters*, 20(4):804–807, 2016. doi:10.1109/LCOMM.2016.2524405.

[DHJA17] E. Dutkiewicz, Y. He, B. A. Jayawickrama, and H. V. Abeywickrama. Radio environment maps generation and spectrum sensing testbed for spectrum sharing in 5G networks. In *2017 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, pages 33–36, 2017. doi:10.1109/APWC.2017.8062233.

[DKG11] E. Dall'Anese, S. Kim, and G.B. Giannakis. Channel gain map tracking via distributed kriging. *IEEE Transactions on Vehicular Technology*, 60(3):1205–1211, 2011. doi:10.1109/TVT.2011.2113195.

[DPR+11] I. Dagres, A. Polydoros, J. Riihijärvi, J. Nasreddine, P. Mähönen, L. Gavrilovska, V. Atanasovski, J. van de Beek, B. Sayrac, S. Grimoud, M. López, J. Perez-Romero, R. Agusti, and F. Casadevall. Radio environmental maps: information models and reference model. document number d4.1. Technical report, GRCM - Grup de Recerca en Comunicacions Mòbils176394researchGroup[...] and CCABA - Centre de Comunicacions Avançades de Banda Ampla177711researchGroup, 04 2011. URL: http://www.ict-faramir.eu/fileadmin/user_upload/deliverables/FARAMIR-D4.1-Final.pdf.

[DSB+20] M. Niklaßand N. Dzikus, M. Swaid, J. Berling, B. Lührs, A. Lau, I. Terekhov, and V. Gollnick. A collaborative approach for an integrated modeling of urban air transportation systems. *Aerospace*, 7(5), 2020. URL: https://www.mdpi.com/2226-4310/7/5/50, doi:10.3390/aerospace7050050.

[DSS+21] Y. Du, Y. Shao, Y. Shi, W. Huang, and K. Sun. Joint power and channel allocation for D2D underlaid multi-user cellular networks. In *2021 IEEE Int. Conf. on Artificial Intelligence and Computer App.*, pages 219–223, 2021. doi:10.1109/ICAICA52286.2021.9498151.

[DWL22] Y. Du, H. Wang, and J. Liu. Radio environment map construction based on random forest regression. In *2022 IEEE 22nd International Conference on Communication Technology (ICCT)*, pages 551–556, 2022. doi:10.1109/ICCT56141.2022.10072831.

[DYJ+19] H. Dun, F. Ye, S. Jiao, Y. Li, and T. Jiang. The distributed resource allocation for D2D communication with game theory. In *2019 IEEE-APS Topical Conference on Antennas and Propagation in Wireless communnication*, pages 104–108, 2019. doi:10.1109/APWC.2019.8870437.

[DYRJ10] K. Doppler, C. Yu, C. B. Ribeiro, and P. Jänis. Mode selection for device-to-device communication underlaying an LTE-Advanced network. In *2010 IEEE Wireless Communication and Networking Conference*, pages 1–6, 2010. doi:10.1109/WCNC.2010.5506248.

[EHA14] H. ElSawy, E. Hossain, and M. Alouini. Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks. *IEEE Transactions on Communications*, 62(11):4147–4161, 2014. doi:10.1109/TCOMM.2014.2363849.

[EHJ+20] M. Cenk Ertürk, N. Hosseini, H. Jamal, A. Şahin, D. Matolak, and J. Haque. Requirements and technologies towards UAM: Communication, navigation, and surveillance. In *2020 Integrated Communications Navigation and Surveillance Conference (ICNS)*, pages 2C2–1–2C2–15, 2020. doi:10.1109/ICNS50378.2020.9223003.

[ETS10] ETSI. Intelligent transport systems (ITS); communications architecture. Technical report (TR) ETSI EN 302 665 v1.1.1, ETSI, 2010. Last visited: 12/11/2024. URL: https://www.etsi.org/302665.

[ETS14] ETSI. Multi-access edge computing (MEC), 2014. Last visited: 12/11/2024. URL: https://www.etsi.org/technologies/multi-access-edge-computing.

[ETS19] ETSI. Intelligent transport systems (ITS); vehicular communications; basic set of applications; analysis of the collective perception service (CPS); release 2. Technical report (TR) ETSI TR 103 562 V2.1.1, ETSI, 2019. Last visited: 12/11/2024. URL: https://www.etsi.org/3103562.

[EUR24] EUROCONTROL. Eurocontrol forecast 2024-2030, 2024. Last visited: 13/11/2024. URL: https://www.eurocontrol.int/publication/eurocontrol-forecast-2024-2030.

[FDH17] A. Fotouhi, M. Ding, and M. Hassan. Dynamic base station repositioning to improve spectral efficiency of drone small cells. In *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–9, 2017. doi:10.1109/WoWMoM.2017.7974285.

[FDH18] A. Fotouhi, M. Ding, and M. Hassan. Flying drone base stations for macro hotspots. *IEEE Access*, 6:19530–19539, 2018. doi:10.1109/ACCESS.2018.2817799.

[FDM+12] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi. Design aspects of network assisted device-to-device communications. *IEEE Communications Magazine*, 50(3):170–177, 2012. doi:10.1109/MCOM.2012.6163598.

[Fed24] Federal Aviation Administration. Unmanned aircraft system traffic management (UTM), 2024. Last Visited : 15/11/2024. URL: https://www.faa.gov/uas/advanced_operations/traffic_management#:~:text=UTM%20is%20how%20airspace%20will,communicate%20real%2Dtime%20airspace%20status.

[FQD+19] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan. Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges. *IEEE Communications Surveys & Tutorials*, 21(4):3417–3442, 2019. doi:10.1109/COMST.2019.2906228.

[FVH+21] G. Fodor, J. Vinogradova, P. Hammarberg, K.K. Nagalapur, Q. Zhiqiang, H. Do, R. Blasco, and M. U. Baig. 5G new radio for automotive, rail, and air transport. *IEEE Communications Magazine*, 2021. URL: https://arxiv.org/abs/2101.08874, doi:10.48550/ARXIV.2101.08874.

[FZZ23] J. Fei, X. Zhu, and R. Zhang. Evaluation of TCP variants on dynamic UAV networks. In *2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 11–16, 2023. doi:10.1109/CSCWD57460.2023.10152635.

[GD20] S. Ghosh and D. De. CG-D2D: Cooperative game theory based resource optimization for D2D communication in 5G wireless network. In *2020 Fifth Int. Conf. on Research in Computational Intelligence and comm. Networks*, pages 171–176, 2020. doi:10.1109/ICRCICN50933.2020.9296163.

[GdKS+19] D. Gunduz, P. de Kerret, N. D. Sidiropoulos, D. Gesbert, C. Murthy, and M. van der Schaar. Machine learning in the air, 2019. URL: https://arxiv.org/abs/1904.12385, arXiv:1904.12385.

[GGRG+18] G. Geraci, A. Garcia-Rodriguez, L. G. Giordano, D. López-Pérez, and E. Björnson. Understanding UAV cellular communications: From existing networks to massive MIMO. *IEEE Access*, 6:67853–67865, 2018. doi:10.1109/ACCESS.2018.2876700.

[GJV16] L. Gupta, R. Jain, and G. Vaszkun. Survey of important issues in UAV communication networks. *IEEE Communications Surveys & Tutorials*, 18(2):1123–1152, 2016. doi:10.1109/COMST.2015.2495297.

[GPG+22] S. Gupta, R. Patel, R. Gupta, S. Tanwar, and N. Patel. A survey on resource allocation schemes in device-to-device communication. In *2022 12th Int. Conf. on Cloud Comp., Data Science & Eng. (Confluence)*, pages 140–145, 2022. doi:10.1109/Confluence52989.2022.9734183.

[GST+14] C. Gao, X. Sheng, J. Tang, W. Zhang, S. Zou, and M. Guizani. Joint mode selection, channel allocation and power assignment for green device-to-device communications. In *2014 IEEE International Conference on Communications (ICC)*, pages 178–183, 2014. doi:10.1109/ICC.2014.6883315.

[GT22] R. Gupta and S. Tanwar. Interference mitigation and secrecy-ensured D2D resource allocation scheme using game theory. In *IEEE INFOCOM 2022 - IEEE Conf. on Computer comm. Workshops*, pages 1–6, 2022. doi:10.1109/INFOCOMWKSHPS54753.2022.9798264.

[GTK21] R. Gupta, S. Tanwar, and N. Kumar. Secrecy-ensured NOMA-based cooperative D2D-aided fog computing under imperfect CSI. *Jour. of Info. Sec. App.*, 59, 2021. doi:https://doi.org/10.1016/j.jisa.2021.102812.

[GZ18] J. Gui and K. Zhou. Cellular throughput optimization by game-based power adjustment and outband D2D communication. *EURASIP Journal on Wireless Comm. and Net.*, 2018, 2018. doi:10.1186/s13638-018-1275-2.

[HC20] G. Hou and L. Chen. D2D communication mode selection and resource allocation in 5G wireless networks. *Computer communications*, 155:244–251, 2020. doi:https://doi.org/10.1016/j.comcom.2020.03.025.

[HCLK10] S. Hakola, T. Chen, J. Lehtomäki, and T. Koskela. Device-to-device (D2D) communication in cellular network - performance analysis of optimum and practical communication mode selection. In *2010 IEEE Wireless Communication and Networking Conference*, pages 1–6, 2010. doi:10.1109/WCNC.2010.5506133.

[HK17] M. Hirose and T. Kobayashi. Measurements of ultra-wideband propagation within a small aircraft for replacing wire harnesses. In *2017 IEEE Radio and Antenna Days of the Indian Ocean (RADIO)*, pages 1–2, 2017. doi:10.23919/RADIO.2017.8242220.

[HO22] Y. Hwang and S. Oh. A study on ultra-reliable and low-latency communication technologies for 5G & 6G services. In *13th Int. Conf. on Inf. and Comm. Tech. Convergence (ICTC)*, pages 1207–1209, 2022. doi:10.1109/ICTC55196.2022.9952510.

[HWYL17] Y. Han, H. Wu, Z. Yang, and D. Li. A new data transmission strategy in mobile D2D networks—deterministic, greedy, or planned opportunistic routing? *IEEE Transactions on Vehicular Technology*, 66(1):594–609, 2017. doi:10.1109/TVT.2016.2540641.

[HYH+21] J. Huang, Y. Yang, G. He, Y. Xiao, and J. Liu. Deep reinforcement learning-based dynamic spectrum access for D2D communication underlay cellular networks. *IEEE Communication Letters*, 25(8):2614–2618, 2021. doi:10.1109/LCOMM.2021.3079920.

[ICVP21] I. Ioannou, C. Christophorou, V. Vassiliou, and A. Pitsillides. 5G D2D transmission mode selection performance & cluster limits evaluation of distributed ai and ml techniques. In *2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, pages 70–80, 2021. doi:10.1109/COMNETSAT53002.2021.9530792.

[IH08] T. Issariyakul and E. Hossain. *Introduction to Network Simulator NS2*. Springer Publishing Company, Incorporated, 1 edition, 2008.

[Inc21] The Radicati Group Inc. Mobile statistics report 2021-2025, 2021. Last visited: 11/11/2024. URL: https://www.radicati.com.

[Int21] SAE International. Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Standard J3016, SAE International, 2021. Last Visited: 11/11/24. URL: https://www.sae.org/standards/content/j3016_202104.

[Int24] International Civil Aviation Organization. Air traffic management (ATM), 2024. Last visited: 20/10/2024. URL: https://www.icao.int/safety/airnavigation/Pages/atm.aspx.

[ITU17] ITU. Minimum technical performance requirements for IMT-2020 radio interfaces. Technical report, Interntional Telecommunications Union, 2017.

[JJ24] A. P. Jain and S. Joshi. UAV- enabled D2D vehicular communication network: Link selection and outage analysis. In *2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)*, pages 1–5, 2024. doi:10.1109/VTC2024-Spring62846.2024.10683161.

[JJC+21] J.Zheng, S. Jiang, W. Chen, F. Zhou, and L. Yang. A neural network based power allocation algorithm for D2D communication in cellular networks. In *2021 IEEE Global Com. Conf.*, pages 1–6, 2021. doi:10.1109/GLOBECOM46510.2021.9685907.

[JKP+21] H. Jeon, B. Koo, S. Park, J. Park, and C. Chae. Graph-theory-based resource allocation and mode selection in D2D communication systems: The role of full-duplex. *IEEE wireless Communication Letters*, 10(2):236–240, 2021. doi:10.1109/LWC.2020.3025312.

[JYD+09] P. Jänis, C. Yu, K. Doppler, C. Ribeiro, C. Wijting, K. Hugl, O. Tirkkonen, and V. Koivunen. Device-to-device communication underlaying cellular communications systems. *IJCNS*, 2:169–178, 01 2009.

[KCZ+18] A. A. Khuwaja, Y. Chen, N. Zhao, M. Alouini, and P. Dobbins. A survey of channel modeling for UAV communications. *IEEE Communications Surveys & Tutorials*, 20(4):2804–2821, 2018. doi:10.1109/COMST.2018.2856587.

[Kes88] S. Keshav. Real: A network simulator. Technical Report UCB/CSD-88-472, EECS Department, University of California, Berkeley, Dec 1988. Last visited: 06/10/2024. URL: http://www2.eecs.berkeley.edu/Pubs/TechRpts/1988/5316.html.

[KG20] A. Kumbhar and I. Guvenc. *UAV Placement and Aerial–Ground Interference Coordination*, chapter 10, pages 255–282. John Wiley & Sons, Ltd, 2020. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119575795.ch10, doi:https://doi.org/10.1002/9781119575795.ch10.

[KGM$^+$20] W. Khawaja, I. Guvenc, D.W. Matolak, U. Fiebig, and N. Schneckenberger. *A Survey of Air-to-Ground Propagation Channel Modeling for Unmanned Aerial Vehicles*, chapter 2, pages 17–70. John Wiley & Sons, Ltd, 2020. doi:https://doi.org/10.1002/9781119575795.ch2.

[KK07] Israel Koren and C. Mani Krishna. *Fault-tolerant systems*. Denise Penrose, San Francisco (CA), first edition edition, 2007.

[KMMH21] C. Kai, X. Meng, L. Mei, and W. Huang. Deep reinforcement learning based user association and resource allocation for D2D-enabled wireless networks. In *2021 IEEE/CIC International Conference on communication in China*, pages 1172–1177, 2021. doi:10.1109/ICCC52777.2021.9580261.

[KSGR19] M. Kamruzzaman, N. I. Sarkar, J. Gutierrez, and S. K. Ray. A mode selection algorithm for mitigating interference in D2D enabled next-generation heterogeneous cellular networks. In *2019 International Conference on Information Networking (ICOIN)*, pages 131–135, 2019. doi:10.1109/ICOIN.2019.8718182.

[Lab21] N. S. Labib. *Distributed unmanned aerial vehicles traffic management system*. PhD thesis, Universite Du Luxembourg: The Faculty of Science, Technology and Medicine, 2021.

[LCS$^+$19] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y. A. Zhang. The roadmap to 6G: AI empowered wireless networks. *IEEE Communications Magazine*, 57(8):84–90, 2019. doi:10.1109/MCOM.2019.1900271.

[LL21] W. Lee and K. Lee. Resource allocation scheme for guarantee of QoS in D2D communications using deep neural network. *IEEE Communication Letters*, 25(3):887–891, 2021. doi:10.1109/LCOMM.2020.3042490.

[LWLL20] X. Lyu, Y. Wang, Z. Li, and M. Liu. Joint power and channel allocation based on mobility and interest aware D2D cache in HetNets. In *2020 IEEE wireless Com. and Net. Conf. Workshops*, pages 1–6, 2020. doi:10.1109/WCNCW48565.2020.9124859.

[Lyn96] N. A. Lynch. *Distributed Algorithms*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1996.

[LZZ16] J. Lyu, Y. Zeng, and R. Zhang. Cyclical multiple access in UAV-aided communications: A throughput-delay tradeoff. *IEEE Wireless Communications Letters*, 5(6):600–603, December 2016. URL: http://dx.doi.org/10.1109/LWC.2016.2604306, doi:10.1109/lwc.2016.2604306.

[LZZL17] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim. Placement optimization of UAV-mounted mobile base stations. *IEEE Communications Letters*, 21(3):604–607, 2017. doi:10.1109/LCOMM.2016.2633248.

[MA22] A. Musa and I. Awan. Functional and performance analysis of discrete event network simulation tools. *Simulation Modelling Practice and Theory*, 116:102470, 2022. URL: https://www.sciencedirect.com/science/article/pii/S1569190X2100157X, doi:https://doi.org/10.1016/j.simpat.2021.102470.

[Mar10] J. Markoff. Google cars drive themselves, in traffic, Oct 2010. Last visited: 11/11/2024. URL: https://www.nytimes.com/2010/10/10/science/10google.html.

[Mat12] D. W. Matolak. Air-ground channels & models: Comprehensive review and considerations for unmanned aircraft systems. In *2012 IEEE Aerospace Conference*, pages 1–17, 2012. doi:10.1109/AERO.2012.6187152.

[MB14] M. Mohammed and K. Bilal. LTE radio planning using Atoll radio planning and optimization software. *International Journal of Science and Research (IJSR)*, 3:1460–1464, 10 2014.

[MBG09] G. Mateos, J. Bazerque, and G. B. Giannakis. Spline-based spectrum cartography for cognitive radios. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pages 1025–1029, 2009. doi:10.1109/ACSSC.2009.5470044.

[MBT17] N. H. Motlagh, M. Bagaa, and T. Taleb. UAV-based IoT platform: A crowd surveillance use case. *IEEE Comm. Mag.*, 55(2):128–134, 2017. doi:10.1109/MCOM.2017.1600587CM.

[MLC+22] A. V. Militaru, R. G. Lazar, C. F. Caruntu, C. R. Comsa, and I. Bogdan. Analysis of message flow transmissions for an inter-vehicle communication scenario. In *2022 14th Int. Conf. on Electronics, Computers and Artificial Intelligence*, pages 1–6, 2022. doi:10.1109/ECAI54874.2022.9847482.

[MS15] D. W. Matolak and R. Sun. Unmanned aircraft systems: Air-ground channel characterization for future applications. *IEEE Vehicular Technology Magazine*, 10(2):79–85, 2015. doi:10.1109/MVT.2015.2411191.

[MS17] D. W. Matolak and R. Sun. Air–ground channel characterization for unmanned aircraft systems—part I: Methods, measurements, and models for over-water settings. *IEEE Transactions on Vehicular Technology*, 66(1):26–44, 2017. doi:10.1109/TVT.2016.2530306.

[MSBD16a] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah. Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. *IEEE Communications Letters*, 20(8):1647–1650, 2016. doi:10.1109/LCOMM.2016.2578312.

[MSBD16b] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah. Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs. *IEEE Transactions on Wireless Communications*, 15(6):3949–3963, 2016. doi:10.1109/TWC.2016.2531652.

[MT23] I. Mantilla and V. Turau. Comparison of WiFi interference mitigation strategies in DSME networks: Leveraging reinforcement learning with Expected SARSA. 2023. doi:10.1109/MeditCom58224.2023.10266605.

[MVD+22] A. Menichino, V. Di Vito, B. Dziugiel, A. Liberacki, H. Hesselink, and M. Giannuzzi. Urban air mobility perspectives over mid-term time horizon: Main enabling technologies readiness review. In *2022 Integrated Communication, Navigation and Surveillance Conference (ICNS)*, pages 1–13, 2022. doi:10.1109/ICNS54818.2022.9771511.

[MWZ18] W. Mei, Q. Wu, and R. Zhang. Cellular-connected UAV: Uplink association, power control and interference coordination. In *2018 IEEE Global Communications Conference (GLOBECOM)*, pages 206–212, 2018. doi:10.1109/GLOCOM.2018.8647820.

[MZOA21] R. N. K. Mensah, L. Zhiyuan, A. A. Okine, and J. M. Adeke. A game-theoretic approach to computation offloading in software-defined D2D-enabled vehicular networks. In *2021 2nd Info. Comm. Tech. Conf.*, pages 34–38, 2021. doi:10.1109/ICTC51749.2021.9441652.

[NA14] E. O. Nuallain and C. Anderson. Real-time radio environment mapping for cognitive radio. In *2014 National Wireless Research Collaboration Symposium*, pages 64–68, 2014. doi:10.1109/NWRCS.2014.16.

[Nat24] United Nations. World population prospects 2024: Summary of results. Techn. rep. UN DESA/POP/2024/TR/NO. 9., New York: United Nations., 2024. Last visited: 11/11/2024. URL: https://desapublications.un.org/publications/world-population-prospects-2024-summary-results.

[NBdA+22] E. C. Pinto Neto, D. M. Baum, J. R. de Almeida, J. B. Camargo, and P. S. Cugnasca. A trajectory evaluation platform for urban air mobility (UAM). *IEEE Transactions on Intelligent Transportation Systems*, 23(7):9136–9145, 2022. doi:10.1109/TITS.2021.3091411.

[NCSG18] V. C. Nneji, M. L. Cummings, A. J. Stimpson, and K. H. Goodrich. Functional requirements for remotely managing fleets of on-demand passenger aircraft. In *AIAA Aerospace Sciences Meeting*, page 2007, 2018.

[NGS24] F. Naeem, V. Gollnick, and C. Schmitt. *5G-Enabled Architectural Imperatives and Guidance for Urban Air Mobility: Enhancing Communication, Navigation, and Surveillance*. AIAA, 2024. URL: https://arc.aiaa.org/doi/abs/10.2514/6.2024-3783, doi:10.2514/6.2024-3783.

[NR16] E. Naghipour and M. Rasti. A distributed joint power control and mode selection scheme for D2D communication underlaying LTE-A networks. In *2016 IEEE Wireless Communications and Networking Conference*, pages 1–6, 2016. doi:10.1109/WCNC.2016.7565155.

[NSTV20] G. Nardini, D. Sabella, G. Steaand P. Thakkar, and A. Virdis. Simu5G–An OMNeT++ library for end-to-end performance evaluation of 5G networks. *IEEE Access*, 8:181176–181191, 2020. doi:10.1109/ACCESS.2020.3028550.

[OF19] O. E. Ochia and A. O. Fapojuwo. Energy and spectral efficiency analysis for a device-to-device-enabled millimeter-wave OFDMA cellular network. *IEEE Transactions on Communications*, 67(11):8097–8111, 2019. doi:10.1109/TCOMM.2019.2935728.

[PLM+21] A. Pukhova, C. Llorca, A. Moreno, C. Staves, Q. Zhang, and R. Moeckel. Flying taxis revived: Can urban air mobility reduce road congestion? *Journal of Urban Mobility*, 1:100002, 2021. URL: https://www.sciencedirect.com/science/article/pii/S2667091721000029, doi:https://doi.org/10.1016/j.urbmob.2021.100002.

[PLRK22] K. Park, J. Lee, H. Ryu, and Y. Kim. A novel cell deployment for UAM communications in 5G-advanced network. In *2022 IEEE Globecom Workshops (GC Wkshps)*, pages 1431–1436, 2022. doi:10.1109/GCWkshps56602.2022.10008545.

[PMRG24] A. Piccioni, A. Marotta, C. Rinaldi, and F. Graziosi. Enhancing mobile networks for urban air mobility connectivity. *IEEE Networking Letters*, 6(2):110–114, 2024. doi:10.1109/LNET.2024.3390610.

[QG23] Y. Qi and S. Geng. Deep-reinforcement-learning-based resource allocation for energy harvesting D2D communication. In *2023 4th International Conference on Electronic Communication and Artificial Intelligence*, pages 85–88, 2023. doi:10.1109/ICECAI58670.2023.10176743.

[QMTN06] F. Qixing, J. McGeehan, E. K. Tameh, and A. R. Nix. Path loss models for air-to-ground radio channels in urban environments. In *2006 IEEE 63rd Vehicular Technology Conference*, volume 6, pages 2901–2905, 2006. doi:10.1109/VETECS.2006.1683399.

[Qua17] Qualcomm Technologies, Inc. LTE unmanned aircraft systems trial report. Technical report, Qualcomm Technologies, Inc., 2017. URL: https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/lte_unmanned_aircraft_systems_trial_report_v1.0.1_0.pdf.

[Rah14] M. A. Rahman. Enabling drone communications with WiMAX technology. In *IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications*, pages 323–328, 2014. doi:10.1109/IISA.2014.6878796.

[RFBA21] R. Rothfeld, M. Fu, M. Balać, and C. Antoniou. Potential urban air mobility travel time savings: An exploratory analysis of munich, paris, and san francisco. *Sustainability*, 13(4), 2021. URL: https://www.mdpi.com/2071-1050/13/4/2217, doi:10.3390/su13042217.

[RG20] R. Rathi and N. Gupta. Device to device communication using stackelberg game theory approach. In *2020 Research, Innovation, Knowledge Management and Technology Application for Business Sustainability*, pages 100–103, 2020. doi:10.1109/INBUSH46973.2020.9392124.

[RH10] G. F. Riley and T. R. Henderson. *The ns-3 Network Simulator*, pages 15–34. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. doi:10.1007/978-3-642-12331-3_2.

[Rio19] R. Rios. 5G Network Planning and Optimisation Using Atoll . Master thesis, Universitat Politècnica de Catalunya, Barcelona, Spain,, 2019.

[RNA21] P. Rahmawati, M. I. Nashiruddin, and M. A.Nugraha. Capacity and coverage analysis of 5G NR mobile network deployment for Indonesia's urban market. In *IEEE Int. Conf. on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, pages 90–96, 2021. doi:10.1109/IAICT52856.2021.9532574.

[RRNY21] K. S. Randhava, M. Roslee, P. U. Nmenme, and Z. Yusoff. Radio resource allocation for interference management in device to device (D2D) 5G networks. In *2021 Int. Conf. on Electrical, Comm. , and Computer Eng.*, pages 1–6, 2021. doi:10.1109/ICECCE52056.2021.9514116.

[SCH+05] A. Sobeih, W. Chen, J. C. Hou, L. Kung, N. Li, H. Lim, H. Tyan, and H. Zhang. J-sim: a simulation environment for wireless sensor networks. In *38th Annual Simulation Symposium*, pages 175–187, 2005. doi:10.1109/ANSS.2005.27.

[SGAKSS18] S Sobhi-Givi, H.Kalbkhani A. Khazali, M.G. Shayesteh, and V. Solouk. Joint mode selection and resource allocation in D2D communication based underlaying cellular networks. *Telecommunication Systems*, 67, 2018. doi:10.1007/s11235-017-0320-5.

[SGB12] S.Siraj, A. K. Gupta, and R. Badgujar. Network simulation tools survey. In *International Journal of Advanced Research in Computer and Communication Engineering*, volume 1, 2012. URL: https://api.semanticscholar.org/CorpusID:212433099.

[SGD11] C. Sommer, R. German, and F. Dressler. Bidirectionally coupled network and road traffic simulation for improved IVC analysis. *IEEE Transactions on Mobile Computing*, 10(1):3–15, 2011. doi:10.1109/TMC.2010.133.

[SI14] S. Song and B. Issac. Analysis of Wifi and Wimax and wireless network coexistence. *International journal of Computer Networks& amp; Communications*, 6(6):63–77, November 2014. URL: http://dx.doi.org/10.5121/ijcnc.2014.6605, doi:10.5121/ijcnc.2014.6605.

[SLO+22] D. Shi, L. Li, T. Ohtsuki, M. Pan, Z. Han, and H. V. Poor. Make smart decisions faster: Deciding D2D resource allocation via stackelberg game guided multi-agent deep reinforcement learning. *IEEE Transactions on Mobile Computing*, 21:4426–4438, 2022. doi:10.1109/TMC.2021.3085206.

[SNH20] N. E. Safwat, F. Newagy, and I. M. Hafez. Air-to-ground channel model for UAVs in dense urban environments. *IET Communications*, 14(6):1016–1021, 2020. URL: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-com.2019.0964, doi:https://doi.org/10.1049/iet-com.2019.0964.

[SRR+16] S. Sun, T. S. Rappaport, S. Rangan, T. A. Thomas, A. Ghosh, I. Z. Kovacs, I. Rodriguez, O. Koymen, A. Partyka, and J. Jarvelainen. Propagation path loss models for 5G urban micro- and macro-cellular scenarios. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–6, 2016. doi:10.1109/VTCSpring.2016.7504435.

[SSC+17] V. Sharma, K. Srinivasan, H. C. Chao, K. Hua, and W. Cheng. Intelligent deployment of UAVs in 5G heterogeneous communication environment for improved coverage. *Journal of Network and Computer Applications*, 85:94–105, 2017. Intelligent Systems for Heterogeneous Networks. URL: https://www.sciencedirect.com/science/article/pii/S1084804516303034, doi:https://doi.org/10.1016/j.jnca.2016.12.012.

[SSH20] P. Seda, M. Seda, and J. Hosek. On mathematical modelling of automated coverage optimization in wireless 5G and beyond deployments. *Applied Sciences*, 10(24), 2020. URL: https://www.mdpi.com/2076-3417/10/24/8853, doi:10.3390/app10248853.

[SWK+19] J. Sae, R. Wiren, J. Kauppi, H. Maattanen, J. Torsner, and M. Valkama. Public lte network measurements with drones in rural environment. In *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pages 1–5, 2019. doi:10.1109/VTCSpring.2019.8746290.

[SWL19] Y. Sun, F. Wang, and Z. Liu. Coalition formation game for resource allocation in D2D uplink underlaying cellular networks. *IEEE Communication Letters*, 23(5):888–891, 2019. doi:10.1109/LCOMM.2019.2908152.

[SZL17] X. Sui, H. Zhang, and Y. Lv. Coverage performance analysis of grid distribution in heterogeneous network. In *IEEE 17th Int. Conf. on Communication Technology (ICCT)*, pages 1424–1428, 2017. doi:10.1109/ICCT.2017.8359867.

[TEO22] TEOCO. Asset radio - radio network planning tool, 2022. Last visited: 22/10/2024. URL: https://teocoaircom.com/asset-radio/.

[Tet04] Tetcos. Tetcos: NetSim - Network Simulation Software, India — tetcos.com, 2004. Last visited: 01/10/2024. URL: https://www.tetcos.com/index.html.

[TFK+18] F. Tang, Z. M. Fadlullah, N. Kato, F. Ono, and R. Miura. AC-POCA: Anticoordination game based partially overlapping channels assignment in combined UAV and D2D-based networks. *IEEE Transactions on Vehicular Technology*, 67(2):1672–1683, 2018. doi:10.1109/TVT.2017.2753280.

[The23] The World Bank. Urban development overview, 2023. Last visited: 11/11/2024. URL: https://www.worldbank.org/en/topic/urbandevelopment/overview.

[TKB+17] V. Tikhvinskiy, P. Korchagin, G. Bochechka, A. Gryazev, and A. Aytmagambetov. Spectrum sharing in 800 mhz band: Experimental, estimation of lora networks and air traffic control radars co-existence. In *2017 International Symposium on Electromagnetic Compatibility - EMC EUROPE*, pages 1–6, 2017. doi:10.1109/EMCEurope.2017.8094705.

[Tun21] H. Myo Tun. Radio network planning and optimization for 5G telecommunication system based on physical constraints. *Journal of Computer Science Research*, 3, 02 2021. doi:10.30564/jcsr.v3i1.2701.

[Tut98] K. Tutschku. Demand-based radio network planning of cellular mobile communication systems. In *Proceedings. IEEE INFOCOM '98, the Conference on Computer Communications. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Gateway to the 21st Century (Cat. No.98*, volume 3, pages 1054–1061 vol.3, 1998. doi:10.1109/INFCOM.1998.662915.

[Var01] A. Varga. The OMNET++ discrete event simulation system. *Proc.ESM*, 9, 2001.

[VD24] P. Valdes-Dapena. Waymo recalls driverless cars to make them less likely to drive into poles | cnn business, Jun 2024. URL: https://www.cnn.com/2024/06/13/business/waymo-recalls-driverless-cars-poles/index.html.

[vSvHWW09] H. van Seijen, H. van Hasselt, S. Whiteson, and M. Wiering. A theoretical and empirical analysis of expected sarsa. In *2009 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning*, pages 177–184, 2009. doi:10.1109/ADPRL.2009.4927542.

[VVSC02] K. Venken, I. G. Vinagre, R. Sigle, and J. D. Cervera. Enabling network redundancy in the radio access network. In *Third International Conference on 3G Mobile Communication Technologies*, pages 27–31, 2002. doi:10.1049/cp:20020355.

[Web07] L. Weber. International civil aviation organization : An introduction. *Air and Space Law 32.4*, 2007.

[Wei16] S. Weinstein. Chapter 1 - introduction to digital transmission. In Sarah Kate Wilson, Stephen Wilson, and Ezio Biglieri, editors, *Academic Press Library in Mobile and Wireless Communications*, pages 1–48. Academic Press, Oxford, 2016. URL: https://www.sciencedirect.com/science/article/pii/B9780123982810000016, doi:https://doi.org/10.1016/B978-0-12-398281-0.00001-6.

[Wir17] Wireless Planning. Radioplanner 3.0 RF planning tool, 2017. Last visited: 22/10/2024. URL: https://www.wireless-planning.com/radioplanner.

[WMM19] Y. Wang, L. Ma, and Z. Meng. Effects of UWB electromagnetic pulse on UAV data link system. In *2019 IEEE 6th International Symposium on Electromagnetic Compatibility (ISEMC)*, pages 1–4, 2019. doi:10.1109/ISEMC48616.2019.8986049.

[WSK+23] L. Wernet, F. Siegmund, R. Kundel, T. Meuser, and R. Steinmetz. Fault-tolerance in future ran: Enabling redundancy for efficient recovery of RAN network functions. In *2023 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, pages 105–110, 2023. doi:10.1109/NFV-SDN59219.2023.10329718.

[WT23a] S. T. Wanniarachchi and V. Turau. A fault-tolerant distributed air-to-ground communication architecture for urban air mobility. In *2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*, pages 639–646, 2023. doi:10.1109/DCOSS-IoT58021.2023.00102.

[WT23b] S. T. Wanniarachchi and V. Turau. A study on the influence of 5G network planning on communication in urban air mobility. In *2023 IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 394–399, 2023. doi:10.1109/WoWMoM57956.2023.00070.

[WT24] S. T. Wanniarachchi and V. Turau. Dynamic resource allocation for 5G device-to-device communication based on expected SARSA. In Armando Castañeda, Constantin Enea, and Nirupam Gupta, editors, *Networked Systems*, pages 231–246, Cham, 2024. Springer Nature Switzerland. doi:https://doi.org/10.1007/978-3-031-67321-4_16.

[WXZ+21] Q. Wu, J. Xu, Y. Zeng, D. W. K. Ng, N. Al-Dhahir, R. Schober, and A. L. Swindlehurst. A comprehensive overview on 5G-and-beyond networks with UAVs: From communications to sensing and intelligence. *IEEE Journal on Selected Areas in Communications*, 39(10):2912–2945, 2021. doi:10.1109/JSAC.2021.3088681.

[WZCL21] P. Wojcicki, T. Zientarski, M. Charytanowicz, and E. Lukasik. Estimation of the path-loss exponent by bayesian filtering method. *Sensors*, 21(6), 2021. URL: https://www.mdpi.com/1424-8220/21/6/1934, doi:10.3390/s21061934.

[WZCZ21] D. Wang, S. Zhang, Q. Cheng, and X. Zhang. Joint interference alignment and power allocation based on stackelberg game in device-to-device communications underlying cellular networks. *IEEE Access*, 9:81651–81659, 2021. doi:10.1109/ACCESS.2021.3086148.

[WZZ20] Q. Wu, Y. Zeng, and R. Zhang. *UAV Communications for 5G and Beyond*, chapter 1, pages 1–16. John Wiley & Sons, Ltd, 2020. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119575795.ch1, doi:https://doi.org/10.1002/9781119575795.ch1.

[WZZY13] S. Wen, X. Zhu, X. Zhang, and D. Yang. QoS-aware mode selection and resource allocation scheme for device-to-device (D2D) communication in cellular networks. In *2013 IEEE International Conference on Communications Workshops (ICC)*, pages 101–105, 2013. doi:10.1109/ICCW.2013.6649209.

[XCL24] G. Xin, W. Chen, and LiuCong. Typical redundancy architecture design of UAV flight control system. In *2024 IEEE 6th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, volume 6, pages 514–518, 2024. doi:10.1109/IMCEC59810.2024.10575698.

[Xin99] C. Xinjie. Network simulations with OPNET. In *WSC'99. 1999 Winter Simulation Conference Proceedings. 'Simulation - A Bridge to the Future' (Cat. No.99CH37038)*, volume 1, pages 307–314 vol.1, 1999. doi:10.1109/WSC.1999.823089.

[XPG+22] H. Xiang, J. Peng, Z. Gao, L. Li, and Y. Yang. Multi-agent power and resource allocation for D2D communications: A deep reinforcement learning approach. In *2022 IEEE 96th Vehicular Technology Conference*, pages 1–5, 2022. doi:10.1109/VTC2022-Fall57202.2022.10012889.

[XS10] H. Xing and S.Hakola. The investigation of power control schemes for a device-to-device communication integrated into OFDMA cellular system. In *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1775–1780, 2010. doi:10.1109/PIMRC.2010.5671643.

[XYYY19] Y. Xue, Z. Yang, W. Yang, and J. Yang. D2D resource allocation and power control algorithms based on graph coloring in 5G IoT. In *2019 Comp., Comm. and IoT App.*, pages 17–22, 2019. doi:10.1109/ComComAp46287.2019.9018806.

[XZ24] G. Xin and Zhan'lei. Research on redundancy management algorithm of UAV three-redundancy flight control computer. In *2024 IEEE 6th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, volume 6, pages 557–561, 2024. doi:10.1109/IMCEC59810.2024.10575527.

[YJL21] S. Yu, Y. J. Jeong, and J. W. Lee. Resource allocation scheme based on deep reinforcement learning for device-to-device communications. In *2021 Int. Conf. on Info. Net.*, pages 712–714, 2021. doi:10.1109/ICOIN50884.2021.9333953.

[YJS+18] Z. Yuan, J. Jin, L. Sun, K. Chin, and G. Muntean. Ultra-reliable IoT communications with UAVs: A swarm use case. *IEEE Comm. Mag.*, 56(12):90–96, 2018. doi:10.1109/MCOM.2018.1800161.

[YLL+22] Y. Yuan, Z. Li, Z. Liu, Y. Yang, and X. Guan. Double deep Q-Network based distributed resource matching algorithm for D2D communication. *IEEE Transactions on Vehicular Technology*, 71(1):984–993, 2022. doi:10.1109/TVT.2021.3130159.

[YM20] Z. Yuan and G. Muntean. Airslice: A network slicing framework for UAV communications. *IEEE Comm. Mag.*, 58(11):62–68, 2020. doi:10.1109/MCOM.001.1900325.

[ZCHL22] X. Zhou, G. Chen, Y. Hu, and X. Li. D2D interference management and resource allocation scheme based on improved graph coloring. In *2022 Int. Conf. on AI, Info. pro. and Cloud Comp.*, pages 184–187, 2022. doi:10.1109/AIIPCC57291.2022.00047.

[ZGD+18] J. Zhao, F. Gao, G. Ding, T. Zhang, W. Jia, and A. Nallanathan. Integrating communications and control for UAV systems: Opportunities and challenges. *IEEE Access*, 6:67519–67527, 2018. doi:10.1109/ACCESS.2018.2879637.

[ZLR06] Y. Zhao, B. Le, and J. H. Reed. Chapter 11 - network support: The radio environment map. In Bruce A. Fette, editor, *Cognitive Radio Technology*, pages 337–363. Newnes, Burlington, 2006. URL: https://www.sciencedirect.com/science/article/pii/B978075067952750012X, doi:https://doi.org/10.1016/B978-075067952-7/50012-X.

[ZLT+15] H. Zheng, W. Li, L. Tian, C. Xu, F. Huang, and J. Zhang. Path loss models for urban macro cell scenario at 3.35, 4.9 and 5.4 ghz. In *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 2229–2233, 2015. doi:10.1109/PIMRC.2015.7343668.

[ZLYY20] L. Zhu, C. Liu, J. Yuan, and G. Yu. Machine learning-based resource optimization for D2D communication underlaying networks. In *2020 IEEE 92nd Vehicular Techn. Conf.*, pages 1–6, 2020. doi:10.1109/VTC2020-Fall49728.2020.9348830.

[ZLZ19] Y. Zeng, J. Lyu, and R. Zhang. Cellular-connected UAV: Potential, challenges, and promising technologies. *IEEE Wireless Communications*, 26(1):120–127, 2019. doi:10.1109/MWC.2018.1800023.

[ZRPH+21] H. Zhu, J. Rodríguez-Piñeiro, Z. Huang, T. Domínguez-Bolaño, X. Cai, X. Yin, J. Lee, and D. Matolak. On the end-to-end latency of cellular-connected UAV communications. In *2021 15th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5, 2021. doi:10.23919/EuCAP51087.2021.9411072.

[ZTJ18] M. Zolanvari, M. A. Teixeira, and R. Jain. Analysis of AeroMACS data link for unmanned aircraft vehicles. In *2018 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 752–759, 2018. doi:10.1109/ICUAS.2018.8453394.

[ZWZ19] Y. Zeng, Q. Wu, and R. Zhang. Accessing from the sky: A tutorial on UAV communications for 5G and beyond. *Proceedings of the IEEE*, 107(12):2327–2375, 2019. doi:10.1109/JPROC.2019.2952892.

[ZYB+21] Q. Zhu, M. Yao, F. Bai, X. Chen, W. Zhong, B. Hua, and X. Ye. A general altitude-dependent path loss model for UAV-to-ground millimeter-wave communications. *Frontiers of Information Technology & Electronic Engineering*, 22:767–776, 06 2021. doi:10.1631/FITEE.2000497.

[ZZL16] Y. Zeng, R. Zhang, and T. J. Lim. Wireless communications with unmanned aerial vehicles: opportunities and challenges. *IEEE Communications Magazine*, 54(5):36–42, 2016. doi:10.1109/MCOM.2016.7470933.

[ZZW+19] D. Zhai, R. Zhang, Y. Wang, H. Sun, L. Cai, and Z. Ding. Joint user pairing, mode selection, and power control for D2D-capable cellular networks enhanced by nonorthogonal multiple access. *IEEE Internet of Things Journal*, 6(5):8919–8932, 2019. doi:10.1109/JIOT.2019.2924513.

[ZZX+22] P. Zhen, B. Zhang, Y. Xu, Z.Chen, H. Wang, and D. Guo. Radio environment map construction based on gaussian process with positional uncertainty. *IEEE Wireless Communications Letters*, 11(8):1639–1643, 2022. doi:10.1109/LWC.2022.3170147.

[ZZZ18a] C. Zhan, Y. Zeng, and R. Zhang. Energy-efficient data collection in UAV enabled wireless sensor network. *IEEE Wireless Communications Letters*, 7(3):328–331, 2018. doi:10.1109/LWC.2017.2776922.

[ZZZ18b] J. Zhang, Y. Zeng, and R. Zhang. UAV-enabled radio access network: Multi-mode communication and trajectory design. *IEEE Transactions on Signal Processing*, 66(20):5269–5284, 2018. doi:10.1109/TSP.2018.2866384.

List of Acronyms

1G	1st Generation Cellular
2D	Two Dimension
2G	2nd Generation Cellular
3D	Three Dimension
3G	3rd Generation Cellular
3GPP	3rd Generation Partnership Project
4D	Four Dimension
4G	4th Generation Cellular
5G	5th Generation Cellular
5GC	5th Generation Core Network
6G	6th Generation Cellular
A2A	Air-to-Air Communication
A2G	Air-to-Ground Communication
ADS-B	Automatic Dependent Surveillance Broadcast
AI	Artificial Intelligence
Alg	Algorithm
AMF	Application Management Function
AMPS	Advanced Mobile Phone System
AS	Access Stratum
ATM	Air Traffic Management
BCQI	Best Channel Quality Indicator
BLER	Block Error Rate
BVLOS	Beyond Visible Line of Sight
BW	Bandwidth
C2	Command and Control
CC	Carrier Component
CCDF	Complementary Cumulative Distribution Function
CDMA	Code Division Multiple Access
CIR	Carrier-to-Interference Ratio
CN	Core Network
CNPC	Control and Non-Payload Communications
CP	Cyclic Prefix
CSV	Comma-Seperated Value
C-V2X	Cellular Vehicle-to-Everything
D2D	Device-to-Device Communication
D3QN	Double Dueling Deep Q-Network
DDQN	Double Deep Q-Network
DL	Downlink
DRL	Deep Reinforcement Learning
DSA	Dynamic Spectrum Access
EDGE	Enhanced Data rates for GSM Evolution
EIRP	Equivalent Isotropically Radiated Power
eMBB	Enhanced Mobile Broadband
eNB	Evolved Node B - enodeB
EPC	Evolved Packet Core
Eqn	Equation
ETSI	European Telecommunication Standards Institute
eVTOL	electric Vertical Take Off and Landing
FAA	Federal Aviation Administration
FDD	Frequency Division Duplex
FD-MIMO	Full Duplex Multiple Input Multiple Output
Fig	Figure
FIM	Flight Information Management
FR1	Frequency Range 1
FR2	Frequency Range 2
FSPL	Free Space Path Loss
GEO	Geostationary Earth Orbit
GIS	Geographic Information Software
gNB	gNodeB

GNSS	Global Navigation Satellite System
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications
GSP	Global Situation Picture
GUI	Graphical User Interface
HAPCom	High Altitude Platform Communication
HBP	Heartbeat Protocol
H-ARQ	Hybrid Automatic Repeat Request
HD	High-Definition
ICAO	International Civil Aviation Organization
ICS	Institute of Control Systems
ID	Identity Number
IDE	Integrated Development Environment
IEEE	Institute of Electrical and Electronics Engineers
IM	Infrastructure-Based Mode
INS	Inertial Navigation System
IoT	Internet of Things
IP	Internet Protocol
ITS	Intelligent Transport System
ITU	International Telecommunication Union
J-Sim	Java-Based Simulation environment
KPI	Key Performance Indicator
LCID	Logical Connection Identifier
LoRa	Long Range Communication
LDM	Local Dynamic Map
LOS	Line of Sight
LSE	Locally Sensed Environment
LSEP	Locally Sensed environment Picture
LSP	Local Situation Picture
LTE	Long Term Evolution
MAC	Medium Access Control
MAPL	Maximum Allowable Path Loss
MAX C/I	Maximum Carrier to Interference
MCMP	Multicopy Multipath
MDP	Markov Decision Process
MEC	Multi-access Edge Computing
MGS	Main Ground Station
MIMO	Multiple Input Multiple Output
ML	Machine Learning
MME	Mobility Management Entity
mMTC	Massive Machine Type Communication
mMW	milimeter Wave
NAS	Non-Access Stratum
NASA	National Aeronautics and Space Administration
NED	Network Description
NetSim	Network based Environment for Modeling Simulations
NF	Network Function
NG RAN	Next Generation Radio Access Network
NI	Numerology Index
NIC	Network Interface Card
NLOS	Non Line of Sight
NMT	Nordic Mobile Telephone
NOMA	Non-Orthogonal Multiple Access
NR	New Radio
NRB	Number of Resource Blocks
NS-2	Network Simulator version 2
NS-3	Network Simulator version 3
NSA NR	Non-Standalone New Radio
OFDM	Orthogonal Frequency Division Multiplexing
OMNeT++	Objective Modular Network Testbed in C++
OPNET	Optimized Network Engineering Tools
OSI	Open Systems Interconnections reference model
OSF	Operational Support Functions
PC	Power Control
PDCP	Packet Data Convergence Protocol
PDU	Packet Data Unit

PF	Preferential Fair
PHY	Physical layer
PLE	Path Loss Exponent
PPP	Point-to-Point
Pr-DDQN	Double Deep Q-Network with Priority Sampling
ProSe	Proximity Services
QoS	Quality of Service
RAN	Radio Access Network
RB	Resource Block
RCE	Remote Component Environment
RE	Resource Element
REAL	Research and Analysis of Load flow
REM	Radio Environment Map
RL	Reinforcement Learning
RLC	Radio Link Control
RRC	Radio Resource Control
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
Rx	Reception
SAM	Strategic Airspace Management
SA NR	Standalone New Radio
SARSA	State-Action-Reward-Next State-Next Action
SCF	Safety Critical Functions
SCMP	Single-copy Multipath
SCS	Sub Carrier Spacing
SDAP	Service Data Adaptation Protocol
SGS	Sub Ground Station
SGSP list	Sub Ground Station Priority list
S-GW	Serving Gateway
SINR	Signal-to-Interference Plus Noise Ratio
SMF	Session Management Function
SMS	Short Message Service
SNR	Signal-to-Noise Ratio
SPF	Single Point Failure
SRF	Safety Related Functions
TACS	Total Access Communication System
TAM	Tactical Airspace Management
TB	Transport Block
Tbl	Table
TCP	Transmission Control protocol
TDD	Time Division Duplex
TDM	Temporal Difference Methods
TR	Technical Report
TS	Technical Specification
TTI	Transmission Time Interval
Tx	Transmission
UAM	Urban Air Mobility
UAS	Uncrewed Aerial System
UAV	Unmanned Aerial Vehicle
UDP	User Datagram Protocol
UE	User Equipment
UHF	Ultra High Frequency
UL	Uplink
ULTRAS	Urbane Lufttransportsimulation
UMTS	Universal Mobile Telecommunication Service
UPF	User Plane Function
URLLC	Ultra-Reliable Low Latency Communication
UTM	Unmanned aircraft system Traffic Management
UWB	Ultra-Wideband
V2X	Vehicle-to-Everything
VTOL	Vertical Take-Off and Landing
WCDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access
XML	Extensible Markup Language

List of Symbols

α	Learning Rate
α_S	Shannon capacity scaling factor
a	Action
A_t	Current Action
A_{t+1}	Next Action
B	Bandwidth
$bits_{D2D}$	Achievable bit rate on a single resource block in device-to-device communication
$bits_{UL}$	Achievable bit rate on a single resource block in uplink
c	Speed of light in vacuum
d	Distance between the transmitter and receiver
d_0	Reference distance
d_{2D}	2D distance between the UE and base station
d_{D2BS}	Distance between the device-to-device pair and base station
d_{D2CUE}	Distance between the device-to-device pair and cellular user
d_{3D}	3D distance between the UE and base station
d'_{BP}	Breakpoint distance
d_{D2D}	Device-to-device pair distance
δ_t	Set of all possible actions in state S_t
ϵ	Random action probability
f	Frequency of the signal
f_c	Carrier frequency
γ	Discount factor
G_{BS}	gNB antenna gain
G_{UE}	Receiver antenna gain
h_{BS}	Base station height
h'_{BS}	Effective antenna height at the base station
h_E	Effective antenna height
h_{UE}	Height of the UE from ground level
h'_{UE}	Effective antenna height at the UE

LIST OF SYMBOLS

$\sum I$	Total interference
IM	Interference margin
μ	Numerology index
$MTFSU$	Mean time for simultaneous update
$MTTF$	Mean time to failure
$MTTIF$	Mean time to identify failure
$MTTR$	Mean time to repair
$MTTRE$	Mean time to recover
$MTTU$	Mean time to update
$MTUFR$	Mean time until failure recognition
n	Path loss exponent
N_o	Additive white Gaussian noise
NF_{UE}	Noise figure of the user equipment
π	Policy in reinforcement learning
$P_{FSPL}(d_0)$	Free space path loss at a reference distance d_0
P_{LOS}	Line of sight probability
$PL(d)$	Path loss at distance d
$PL(d_0)$	Path loss at a reference distance d_0
PL_{NLOS}	Path loss for non-line-of-sight case
PLE_{LOS}	Path loss exponent for the line-of-sight case
PLE_{NLOS}	Path loss exponent for the non line-of-sight case
$P_{TX_{BS}}$	gNB Transmission Power
$P_{UE_{sensitivity}}$	Receiver sensitivity
SNR_{D2D}	Signal-to-noise ratio of device-to-device communication
SNR_{DL}	Downlink signal-to-noise ratio
SNR_{thresh}	Threshold of the signal-to-noise ratio
SNR_{UL}	Uplink signal-to-noise ratio
$Q(S_t, A_t)$	Q-Value
R	Reward
s	state
s'	Next state
SM	Shadow margin
S_t	Current State

S_{t+1}	Next State
$TACK$	Waiting time during acknowledgment expectation
$thresh_{D2D}$	Device-to-device distance threshold
$thresh_{D2BS}$	Distance threshold between device-to-device pair and base station
$thresh_{D2CUE}$	Distance threshold between device-to-device pair and cellular user
$V_{s'}$	Next state expectation
w	Bandwidth