DEDICAT 6G

DEDICAT 6G: Dynamic coverage Extension and Distributed Intelligence for human Centric Applications with assured security, privacy and Trust: from 5G to 6G

Deliverable D4.2
Second release of mechanisms for dynamic coverage and connectivity extension

Project Details

Call	H2020-ICT-52-2020
Type of Action	RIA
Project start date	01/01/2021
Duration	36 months
GA No	101016499

Deliverable Details

Deliverable WP:	WP4
Deliverable Task:	Task T4.1, T4.2
Deliverable Identifier:	DEDICAT 6G_D4.2
Deliverable Title:	Second release of mechanisms for dynamic coverage and connectivity extension
Editor(s):	WINGS
Author(s):	Mickael Maman, Esteban Catte, Mohamed Sana (CEA), Anttonen Antti (VTT), Haeyoung Lee, Francois Carrez (UoS), Vasiliki Lamprousi, Charalampos Korovesis, Antonis Karaolanis, Vera Stavroulaki, Panagiotis Demestichas, Kostas Tsagkaris (WINGS), Javier Moreno, Javier Renart (ATOS), Yolanda Fernandez (TTI), Merkebu Girmay, Vasilis Maglogiannis, Dries Naudts(IMEC), Jihoon Yang, Klaus Moessner (TUC), Drazen Ribar (AIRBUS), Srdjan Penjivrag (VLF)
Reviewer(s):	IMEC and TTI
Contractual Date of Delivery:	30/09/2022
Submission Date:	30/09/2022
Dissemination Level:	PU
Status:	Final
Version:	V1.0
File Name:	DEDICAT6G_D4.2_Second_release_of_mechanisms_for_dynamic_coverage_and_connectivity_extension_v1.0

Disclaimer
The information and views set out in this deliverable are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Deliverable History

Version	Date	Modification
V1.0	*30/09/2022*	*Final version, submitted to EC through SyGMa*

Table of Content

List of Acronyms and Abbreviations

Acronym/Abbreviation	Definition
µS	Micro-service
3D	3-Dimensional
AI	Artificial Intelligence
AP	Access Point
API	Application Programming Interface
B5G	Beyond 5G
BS	Base Station
CEaaS	Coverage Extension as a Service
CEDM	Coverage Extension Decision Making
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
C-V2X	Cellular Vehicle to everything
DB	Data Base
DL	Downlink
DM	Decision Making
ENB	eNodeB
ETSI	European Telecommunications Standards Institute
FC	Functional Component
FG	Functional Group
IDDM	Intelligence Distribution Decision Making
IoT	Internet of Things
IQ	in-phase and quadrature-phase
ITS	Intelligent Transport Systems
ITU	International Telecommunication Union
IVI	Infrastructure to Vehicle Information
LoS	Line of Sight
LTE	Long Term Evolution
MAC	Media Access Control
MANO	MANagement and Orchestration
MAP	Mobile Access Point
MARL	Multi-Agent Reinforcement Learning
MBSFN	Multi-Broadcast Single-Frequency Network (MBSFN)
MCS	Mission Critical Service
MEC	Mobile Edge Computing
MIP	Mixed Integer Programming

ML	Machine Learning
mmW	Millimeter Wave
MQTT	MQ Telemetry Transport
NFV	Network Function Virtualization
NFVI	Network Functions Virtualization Infrastructure
NFV-O	Network Function Virtualization Orchestrator
NLoS	non-Line of Sight
NODM	Network operation Decision Making
NR	New Radio
OF	Objective Function
OSM	Open Source MANO
PoC	Proof of Concept
PPDR	Public protection & Disaster Relief
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
REST	Representational State Transfer
RSSI	Received Signal Strength Indicator
RSU	Roadside Unit
SDN	Software-Defined Networking
SDR	Software Defined Radio
SINR	Signal to Interference plus Noise Ratio
SLA	Service Level Agreement
srsRAN	Software Radio System RAN
SW	Software
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TR	Technology Recognition
UAV	Unmanned Aerial Vehicle
UDP	User Datagram Protocol
UE	User Equipment
UL	Uplink
UMI	Urban Micro base station
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to- Vehicle
V2X	Vehicle-to-everything
VNF	Virtualized Network Function
VRU	Vulnerable Road User

List of Figures

List of Tables

Executive Summary

The emerging 6G architecture will be dominated by operator services enabling dynamic topology changes of the network infrastructure to better adapt to time-varying user traffic needs. With dynamic and efficient infrastructure expansion, future 6G networks are envisaged to ensure that equal opportunity is offered to users regardless of location and mobility. This flexible infrastructure could be implemented by using Mobile Access Points (MAPs), which can be dynamically deployable based on the required service. A large volume of research has been conducted in recent years on MAPs such as cars, robots, and drones or Unmanned Aerial Vehicles (UAVs). DEDICAT 6G project develops dynamic coverage and connectivity extension mechanisms exploiting multiple types of MAPs for covering areas that cannot be easily reached (e.g. hard geo-morphology such as cave forest), where infrastructure, or additional capacity is required only for a finite, short amount of time (e.g. moving hotspots like festivals), or where regular network infrastructure has been damaged (e.g. after an emergency like earthquake, fire, terrorist attack). These mechanisms can be triggered by verticals for Coverage Extension as a Service (CEaaS)

This deliverable reports the activity conducted in the work package 4 on the design and development of mechanisms for the dynamic coverage and connectivity extension through the exploitation of innovative devices (e.g., drones, robots, connected cars, other mobile assets like forklifts in a warehouse, etc.). D4.2 provides a second release of the mechanisms and components for dynamic coverage and connectivity extension, including knowledge building. This deliverable is an extension of D4.1. These innovative components are fueling the lab integration reported in WP6.

First, D4.2 deliverable describes the DEDICAT 6G functional architecture for CEaaS, highlighting the interactions between the DEDICAT 6G platform and a legacy 5G network infrastructure and provides an overview of the functionalities that the DEDICAT 6G platform will offer for CEaaS. One important feature is the context awareness, the ability to monitor what is happening in the network and the system as a whole, to constantly deduce the current status and whether decision making is required. Another important aspect is related to decision making components, to enable the system to make decisions on the MAPs deployment strategies. This includes making decisions about intelligence distribution (i.e. service placement), network operation (i.e. MAP-UE association and Radio Access Technology (RAT) selection) and coverage extension (e.g. MAP and swarm operation).

Second, D4.2 describes in detail several context and situation awareness functionalities (i.e. device discovery, technology recognition and VRUs tracking). These modules provide essential information for making decisions exploiting MAPs.

- Device discovery module is based on a new type of enhanced IoT devices estimating the number of users in a coverage area by short-range signal scanning techniques. Thanks to the information monitored and collected by the IoT sensing nodes, the backend of the micro-service generates detection chart reporting the history of device detections that have been performed by each of the nodes and real-time heat maps of the occupation by connected and non-connected users present in the area of interest.
- Technology recognition module can recognize a wide range of wireless technologies operating in the 5.9GHz ITS band (LTE, 5G NR, WiFi, C-V2X PC5 and ITS-G5) and can use the statistics of each identified technology to estimate and predict the traffic characteristics of each technology and to implement flexible and dynamic spectrum management and utilization.
- Vulnerable Road Users (VRUs) tracking module can track the movement of VRUs on

the road or obscured subjects in real-time through camera sensors installed in the roadside unit located at an intersection and warns the driver of a hazardous situation via wireless vehicle-based infrastructure communication.

Third, D4.2 describes and evaluates several network operation decision making functionalities for MAP/UE association. The first option in a centralized UE association and MAP placement making a trade-off between the network cost minimization (number of MAPs) and user utility (spectral efficiency) maximization in a single framework where user utility is both an optimization constraint and optimization sub-objective. The second option is a distributed UE association and MAP placement deciding the number and the optimal location of MAPs, which maximizes the throughput and ratio of well-deserved users while minimizing the number of drones deployed and the execution time in sub-6GHz and mm-wave band. The last option provides heterogeneous MAP-assisted networks with machine learning to maximize the Quality of Service (QoS) satisfaction level and the energy efficiency by jointly optimizing user association and power allocation under wireless backhaul link capacity constraint in highly crowded areas with heavy traffic loads. Another network operation decision making functionality described in D4.2 is the management and configuration of vehicular based MAPs (i.e. RAT selection and configuration based on MAP capabilities, applications/services traffic demands, and/or identified characteristics of the wireless environment).

Finally, D4.2 also describes coverage extension decision making (i.e. MAP operation managing mobility and swarm operation) and assistance to Intelligence distribution in the coverage extension. Intelligence is every software-based part that can consume computational resources available on the nodes of a B5G network scheme. Three types of intelligence can be defined: (i) DEDICAT 6G Functional Component (FC) instances assisting the placement optimization and deploying concrete FCs instances along the network depending on the needs, (ii) NFV Orchestration of VNFs in the instantiation of network slices to enable end-to-end 5G/B5G connectivity and (iii) verticals apps impacting the coverage extension to guarantee the service. Concerning the MAP mobility management, D4.2 proposes an algorithm for robot based MAPs, finding the path or trajectory that each MAP should follow to reach the target position and the selection of nearby docking/charging stations.

1 Introduction

One of the DEDICAT 6G project challenges is the need for expanded coverage and dynamic connectivity. It is important to ensure that equal opportunity is offered to citizens and businesses regardless of location. The expansion of the activity area of people and things, anywhere and anytime, requires the efficient expansion of the infrastructure. DEDICAT 6G proposes to exploit access nodes delivering B5G/6G features, dynamically deployable, (i.e., mobile base stations), which can deliver the needed service at the needed area (e.g., in a wide-scale and reliable wireless manner, during disasters and temporary events).

WP4 addresses the design and development of mechanisms for the dynamic coverage and connectivity extension through the exploitation of innovative devices (e.g., drones, robots, connected cars and automated guided vehicles.). The overall aim is to enable the dynamic, opportunistic set up of dynamic coverage and connectivity extensions for covering areas that cannot be easily reached, where infrastructure is required only for a finite, short amount of time, or where regular network infrastructure has been damaged e.g., due to terrorist actions or natural disasters. These will be combined with the mechanisms for dynamic, distributed intelligence, caching and storage (developed in WP3) and secure and trusted exchange of information (developed in WP5). We will consider a wide range of factors such as the location and trajectory of devices that can play the role of temporary MAPs, communication capabilities, locations of docking and charging stations of the candidate mobile access points. In this direction, the main objectives of this WP are:

- To formulate and develop algorithms for deciding on the creation and optimal configuration of the dynamic coverage and connectivity extensions in terms of the location of MAPs, the communication configuration of the MAPs and the allocation of nodes to MAPs;
- To formulate and develop mechanisms for knowledge building to support the decision-making process;
- To develop a laboratory proof of concept platform (including prototypes as well as simulations) to validate the technology enablers;
- To implement MAPs using connected cars, drones and robots as MAPs.

1.1 Scope

Task T4.1 investigates the different strategies which can be applied at Radio Access Network (RAN) level to extend the coverage of the network by means of MAPs. This includes how heterogeneous resources can be allocated to MAPs in order to provide the required coverage and connectivity extension and to ensure the QoS expected by the mobile nodes. It also addresses preliminary issues on the way the backhauling link is set up between MAPs and the fixed architecture. A central aspect in this task relates to the spectrum sharing options which can be used for both the air interface and the backhauling, as well as to the MAPs multi-*Radio Access Technologies (RAT)* capabilities as this would yield to different levels of achievable performance (capacity, latency, etc.) in the area of the coverage and connectivity extension.

Task 4.2 focuses on the specification and development of algorithms for deciding on the creation and optimal configuration of dynamic coverage/service extensions through the exploitation of mobile nodes/mobile access points. This also includes the specification of mechanisms for knowledge building for supporting the decision making on the creation and optimal configuration of dynamic coverage/service extensions. Task 4.2 proposes solutions deciding the deployment of MAPs and finding the optimal (i) positions to which each MAP entity should move, and the path or trajectory that each MAP entity should follow so as to reach the target position; (ii) Configuration of the radio network of the MAP entities; (iii) Allocation of nodes to MAPs; (iv) Selection of nearby docking/charging stations for drone MAP so as to ensure connectivity of the appropriate QoS to mobile nodes. This decision and optimisation process takes into account knowledge and information on the context, the capabilities of UE and MAP entities and the potential policies.

D4.2 provides a second release of the mechanisms and components for dynamic coverage and connectivity extension, including knowledge building. This deliverable is an extension of D4.1. These innovative components are fueling the lab integration reported in WP6.

1.2 Structure

Section 2 describes the DEDICAT 6G framework for Coverage Extension as a Service (CEaaS). It includes its integration in the DEDICAT 6G functional architecture from WP2, its interaction with a legacy 5G network infrastructure and its deployment procedure. Section 3 describes context awareness mechanisms for CEaaS (i.e. technology recognition, traffic characterization, device discovery and vulnerable road user tracking). Section 4 focuses on network operation decision making (i.e. several strategies for MAP/UE Association and MAP configuration), coverage extension decision making (i.e. MAP operation managing mobility and swarm operation) and assistance to Intelligence distribution in the coverage extension. Finally, Section 5 concludes the deliverable.

2 DEDICAT 6G Architecture for CEaaS

DEDICAT 6G project develops dynamic coverage and connectivity extension mechanisms exploiting multiple types of MAPs for covering areas that cannot be easily reached (e.g. hard geo-morphology such as cave forest), where infrastructure, or additional capacity is required only for a finite, short amount of time (e.g. moving hotspots like festivals), or where regular network infrastructure has been damaged (e.g. after an emergency like earthquake, fire, terrorist attack). These mechanisms can be triggered by verticals for Coverage Extension as a Service (CEaaS).

2.1 DEDICAT 6G Functional Architecture Overview

Figure 1 depicts an overview of the DEDICAT 6G functional architecture for CEaaS, highlighting the interactions between the DEDICAT 6G platform and a legacy 5G network infrastructure. We consider a hierarchical network organized in four mains domains: Access, far edge, edge and core cloud.

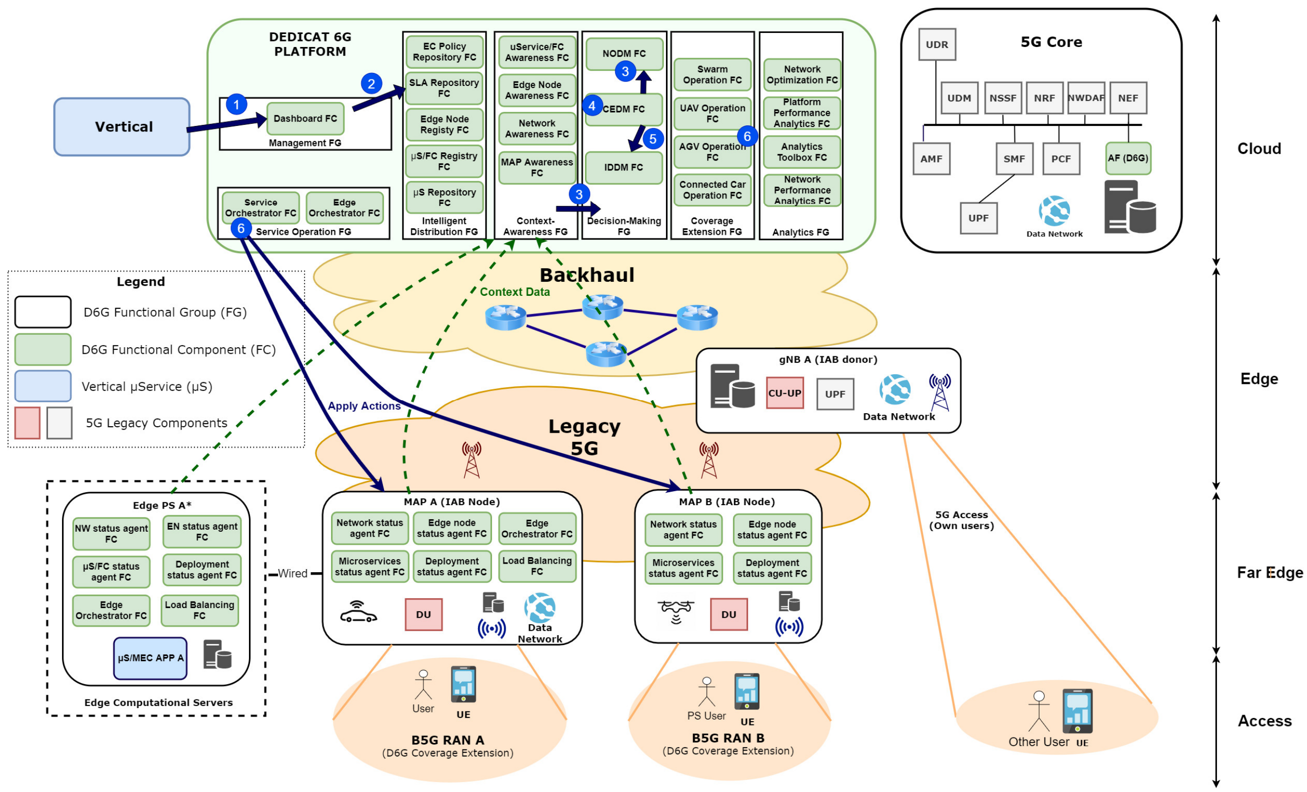


Figure 1: DEDICAT 6G functional Model

One important feature is the ability to monitor what is happening in the network and the system as a whole, to constantly deduce the current status and whether decision making is required. The system must obtain and monitor information on network goals and objectives to be achieved, as well as potential policies. Information on capabilities of network elements, MAPs and edge devices in terms of communication networking (e.g. RAT and spectrum, capacity, and coverage), physical movement, the type of the MAP, computation capabilities, storage capabilities and available power must also be monitored. The system has to maintain information and knowledge on the context in terms of computation tasks, power consumption requirements, a set of mobile nodes that needs coverage, mobility and traffic profiles of the different nodes, radio quality experienced by UEs, options for connecting to wide area networks, the locations of docking and charging stations for MAPs and the current locations of the terminals and MAPs elements.

Another important aspect is to enable the system to make decisions on the MAPs deployment strategies. The key aims of Decision Making (DM) components are to produce the optimal placement of intelligence in terms of data and computation as micro-services, the optimal configuration of the radio network of the MAPs and the MAPs paths (trajectories) that need to be followed, in order to offer adequate QoS levels, in terms of service availability, performance and reliability.

The objective of Coverage Extension DM (CEDM) for static deployment is to find static places for one or more MAPs that fulfil the desired optimization objective. For dynamic deployment, the path between the initial and final destinations is also of interest and some related parameters, such as MAP speed, may be worth for maximizing the operation lifetime. Optimizing the trajectories of the UAVs can be very important in dynamically changing environments. When a MAP is deployed to provide temporary wireless coverage, it is operated as a relay node to relay traffic between a nearby macro Base Station (BS) and end users and the limit of backhaul link capacity can be very challenging. In addition, MAP uses on-board batteries to power its operation (i.e., moving and providing wireless links). Thus, the limitations of the backhaul capacity and battery power will impact decisions and network performance.

With the deployment of MAPs comes the problem of dynamic association of multiple users to multiple access points (fixed or moving). Network Operation DM (NODM) optimization includes the most appropriate allocation of users to MAPs as well as selection of nearby docking/charging stations for drone and robot MAPs to ensure connectivity with the appropriate QoS to users. The selection process results of a tri-partite negotiation between the UE (the initiator), BSs and MAPs. Prior to the negotiation, the DM at the UE side starts to collect information and decides the best association. When UEs are already associated to BS/MAPs, the DM decides whether it keeps the association to the existing MAP/BS or to handover to another one. For UEs capable of multi-connectivity, multiple eligible BSs and/or MAPs can be decided. In multiple Radio Access Technologies (RAT) networks, efficient RAT selection is also crucial for resource utilization, QoS and user satisfaction. The RAT resource configuration realizes harmless coexistence between the technologies.

Intelligence Distribution DM (IDDM) brings higher levels of flexibility, programmability and scalability in network management and orchestration, with Software-Defined Networking (SDN) and Network Function Virtualization (NFV). SDN encompasses a set of network management-oriented techniques that focuses primarily on decoupling the actions of the control plane logic from the data plane. This principle allows a high level of flexibility in the system while the logical control of the network can easily be centralized in an external entity, the SDN controller. NFV virtualizes traditional network functions, typically deployed on physical equipment, into software that can be run on servers or devices with computing capabilities, such as Virtual Network Functions (VNFs) or Container Network Functions (CNFs). From an architectural point of view, IDDM is essential in addressing the scalability and dynamism issues of managing computing resources in a B5G/6G scenario, which is particularly critical at the edge and domains. Additionally, IDDM is not only able to consider the distribution of NFV-related intelligence, but also the IT resource consumption of some required DEDICAT~6G Functional Components (FC) instances, where specific functionalities are integrated, and vertical applications to enable the correct deployment of CEaaS.

2.2 CEaaS System Use-case

The DEDICAT 6G system provides CEaaS to vertical business applications, meaning that a third-party application would request explicitly an extension of the radio coverage with specific QoS requirements in order to be able to run the application while fulfilling its QoS.

The various steps between Functional Components (FC) are illustrated in Figure 1 and described as follows:

1. A vertical application (e.g. a public concert organizer) requests a temporary coverage extension for the duration of its event in order to allow video sharing, and online video coverage. The request is coupled with a set of technical constraints, e.g. number of targeted attendees, event location, type of traffic and target traffic per UE.
2. This step focuses on the Intelligence Distribution Functional Group (FG) like registries and repositories for the Mobile Edge Computing (MEC) and associated look-up /discovery functions plus Service Level Agreements (SLA) and migration policies storage. An SLA negotiation takes place between both parties, leading to a set of technical objectives to be met by the DEDICAT 6G platform. This set is the main input to the CEDM.
3. The CEDM instruments the NODM with QoS targets, which in turn will configure the network in order to fulfil the QoS objective (e.g. involving network slicing).
4. The CEDM decides about the nature of MAP involved and their deployment (number of MAPs and their location). Edge Node Awareness (and related agents) and Edge Node Registry FC are involved in building up a context used for the IDDM decision process.
5. The CEDM instruments the IDDM where information distribution is also required, either to support the deployment and execution of specific vertical FCs or to support purely telecom-related aspects. The Edge Node Awareness FC and associated agents are involved in creating a context used for the IDDM decision process.
6. Finally, the operation of the MAPs, services, and MEC is performed. Coverage Extension FG supports the dynamic coverage extension like MAP dynamic ad-hoc routing, autonomy management, placement management. Service operation translates decisions into readable instructions or commands for external agents, e.g. NFV Orchestrator. The load balancing functionality is applied to the service and networking components required for the distributed nature of the micro services to be on boarded.

2.3 DEDICAT 6G Framework for CEaaS

Figure 2 illustrates the implementation of DEDICAT 6G framework for CEaaS. In this section, we mainly focus on context awareness (i.e. knowledge about UEs and RATs), CEDM FCs (i.e. MAP and swarm operation) and NODM FCs (i.e. MAP-UE association and RAT selection).

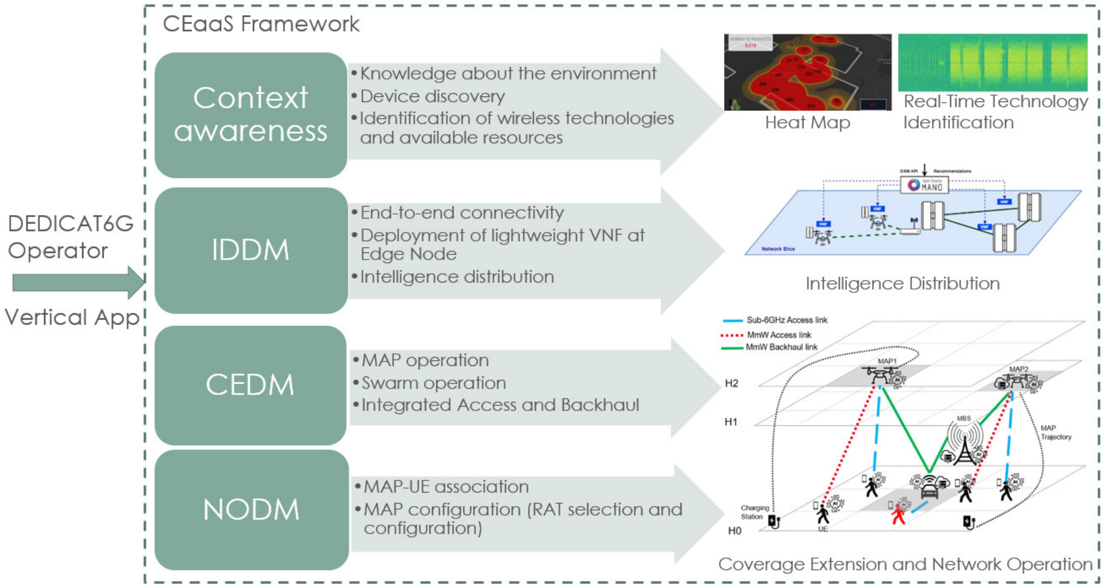


Figure 2: DEDICAT 6G framework for CEaaS

2.3.1 Context Awareness

DEDICAT 6G has developed a context awareness module for device discovery. A new type of enhanced Internet of Things (IoT) devices is proposed to estimate the number of users in a coverage area by short-range signal scanning techniques (e.g., Bluetooth [1] or Wi-Fi signals [2]). These devices monitor what is happening and draw a heat map related to the number of users present at a given time at a designated point. This strategy works like a people counter [3], whose application is to detect a crowd at a given location. It can be used to detect a possible overloading of communication networks due to the presence of a high number of users. The people counter is performed by IoT devices through the knowledge of media access control (MAC) addresses detected and the received signal strength in the coverage area [3]. To collect the data generated by each IoT device, the context awareness module uses a standard messaging MQTT (Message Queuing Telemetry Transport) protocol. It simplifies the collection of data from the sensors, the publication of the different values obtained and the remote configuration of the nodes. Depending on the data analysis, this context awareness module can explicitly request CEaaS, support CEDM/IDDM/NODM decisions and make MAP deployments more efficient.

Bernardos et al. [4] predicted that spectrum management and utilization will become more flexible and dynamic compared to today's static and conservative approaches. In the 5.9 GHz ITS band, multiple RATs, such as 5G NR, LTE, and Wi-Fi, can be used for vehicle-based MAP deployment along with the incumbent Vehicle-to-everything (V2X) technologies. DEDICAT 6G proposes a knowledge-building process which is used to sense the traffic from co-located RATs. The knowledge-building process deploys a Technology Recognition (TR) solution that can recognize a wide range of wireless technologies. The proposed TR model is formulated by extending the TR model of [5] considering co-located LTE and Wi-Fi transmitters operating at 2.4 GHz. In this paper, it proposes and evaluates a TR model that can classify LTE, 5G NR, and ITS-G5 transmissions in the ITS band. The statistics of each identified technology can be used to estimate and predict the traffic characteristics of each technology [6], which can be used as an input to develop a spectrum sharing scheme.

2.3.2 Network Operation Decision Making

Network Operation Decision Making considers both MAP-UE Association and MAP configuration. A large volume of research has been conducted on UE-BS association for fixed BSs. For example, Sana et al. [7] studied user association based on multi-agent reinforcement learning in which users act as independent agents based on their local observations and learn to autonomously coordinate their actions in order to optimize the network sum-rate. DEDICAT 6G proposed an extension of this work in [8] for deciding the number and the optimal location of MAPs which maximizes the throughput and ratio of well-deserved users while minimizing the number of drones deployed and the execution time. We optimize the user association to map UE to both sub-6GHz and mm-wave bands and to guarantee the UEs' QoS requirement. Our proposed solution jointly considers inter-cell and intra-cell interference, user mobility, and user traffic request during the optimization. As a result, the proposed solution adapts well by design to dynamic networks with mobility, dynamic traffic, and varying number of UEs. Other MAP-UE association strategies are investigated in DEDICAT 6G framework. A user association strategy is proposed in heterogeneous MAP-assisted networks with machine learning. The objective is to maximize the QoS satisfaction level and the energy efficiency by jointly optimizing user association and power allocation under wireless backhaul link capacity constraint in highly crowded areas with heavy traffic loads. An important characteristic of MAP placement is awareness of different trade-offs, e.g. between MAP swarm size and target user utility maximization with given optimization constraints. In regard to this, also the centralization degree of the MAP placement algorithm is important which can vary from fully centralized to fully distributed frameworks. In general, some parts of the selected multi-objective functions (e.g. network cost minimization) may be more reasonable to perform in a centralized fashion while some other parts in a distributed fashion (e.g. user utility maximization).

In 6G networks, different technologies can operate in wide range of licensed and unlicensed spectrum, as far as efficient coexistence scheme is deployed [9]. In this direction, DEDICAT 6G develops a knowledge-building process for deriving information about the wireless environment that can be used for RAT selection and configuration, as well as enabling coexistence and incumbent protection. For a newly activated MAP, specific RATs can be selected based on MAP capabilities, applications/services traffic demands, and/or identified characteristics of the wireless environment. As multiple MAPs with multi-RAT capabilities can be dynamically set up close to each other, it is important to set up coexistence mechanisms to ensure fairness, optimal sharing of the spectral resources between different co-located RATs, and protection of potential incumbents like C-V2X PC5 and ITS-G5. Hence, the selected RAT(s) has to be properly configured by a RAT configuration scheme which is used to optimally utilize the available spectrum, enabling harmless coexistence and incumbent protection. DEDICAT 6G aims to formulate efficient RAT selection and configuration schemes based on the results obtained in the knowledge building.

2.3.3 Coverage Extension Decision Making

Coverage Extension Decision Making considers MAP operation and swarm operation. MAP operation includes managing mobility, adjusting orientation and tilt, and receiving status reports on ongoing operations. Swarm operation FC performs self-organizing functions as an ad hoc network (including access and backhaul links), MAP placement, MAP path optimization to achieve placement within the constraints of the CEDM coverage area, and self-management of resources, including autonomy (e.g. how often and where a MAP should return to its docking station).

DEDICAT 6G studies a complex optimization problem to decide on the deployment of

MAPs to provide appropriate QoS connectivity to mobile nodes and find (i) the optimal positions to which each MAP should move, (ii) the path or trajectory that each MAP should follow to reach the target position, (iii) the selection of nearby docking/charging stations for the MAP drones and (iv) the configuration of the MAPs' radio network.

DEDICAT 6G first focuses on how to deploy MAPs, i.e. the optimal number and 3D positions of MAPs in continuous space. MAPs can then dynamically move close to the area where users are concentrated or where more service is requested. Thus, MAPs trajectories can take into account typical trajectory constraints, application-based deployment optimization, and ground user mobility in dynamic and changing environments.

DEDICAT 6G also involves a joint optimization of NODM and CEDM (i.e. multiple UAVs 3D placement and user association) taking into account (i) MAP mobility, (ii) MAP radio network configuration, (iii) node allocation to MAPs and (iv) the resource allocation for wireless access and backhaul.

2.3.4 Intelligence Distribution Decision Making

The DEDICAT 6G project assumes as intelligence every software-based part that can consume computational resources available on the nodes of a B5G network scheme. Three different types of intelligence are considered in this context, DEDICAT 6G FC instances, VNFs based on ETSI's NFV principles, and Verticals apps to exploit the MEC capabilities of the 5G ecosystem. In CEaaS, it is especially relevant to put the spotlight on the first two types of intelligence. First, it is essential to deploy concrete FCs along the network fabric to perform their corresponding functionalities in order to enable CEaaS. For this purpose, the DEDICAT 6G architecture relies on the Intelligence Distribution Decision Making FC (IDDM FC) to calculate the best locations to accommodate the intelligence taking into account the current state of the computational and network contexts. DEDICAT 6G considers as one of its main pillars the research and development of novel intelligence distribution techniques to be applied in a dynamic and changing network ecosystem. And secondly, combination of control-data decoupling in network management and virtualization-based orchestration techniques is essential for successful end-to-end connectivity and coverage extension in B5G networks. DEDICAT 6G investigates techniques to automate the correct configuration of a connection from the core platform in coverage extension operations. We focus on instructing NFV Orchestrator (NFV-O) to set up the service/network portion to enable ad-hoc connectivity when the connectivity area needs to be extended or moved. We also involve deploying lightweight VNFs at edge nodes or far edge nodes, such as drones or robots. DEDICAT 6G supports the configuration of 5G/RAN components and enables run-time operations or post-installation reconfigurations.

Last but not least, IDDM FC will not directly implement the changes in the system but will provide optimal recommendations. The Service Orchestrator FC is in charge of first translating those instructions into the correct form, and ultimately interacting with the external entities to implement the corresponding actions. Hence, DEDICAT 6G is investigating and developing concrete modules to support such interactions with some of the trending 5G-based control plane tools, such as ETSI OpenSource MANO or Kubernetes.

3 Context awareness mechanisms for dynamic coverage and connectivity extension

In this section, we describe several context and situation awareness functionalities (i.e. device discovery, technology recognition and VRUs tracking). These modules provide essential information for making decisions exploiting MAPs.

3.1 Knowledge Building

An efficient spectrum sensing mechanism is required to enable dynamic spectrum management of 6G networks [9]. In this section, we present the performance of the Technology Recognition (TR) based spectrum sensing model used for the environment knowledge building process. Figure 3 shows the overall procedures of the proposed technology identification and characterization solution that is used for the knowledge building process. Figure 3 shows that transmissions from C-V2X PC5, ITS-G5, Wi-Fi, LTE, and 5G NR transmitters are received and pre-processed before the TR process.

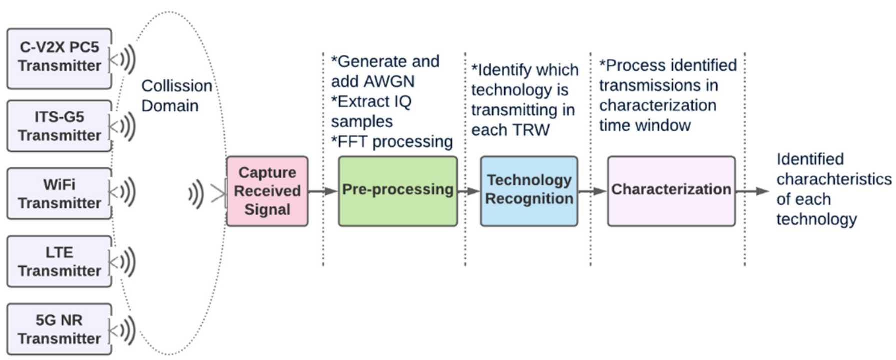


Figure 3: Architecture of proposed technology identification and characterization process.

3.1.1 Technology Recognition

The TR model is trained and validated based on dataset collected from LTE, WiFi, 5G NR, C-V2X PC5 and ITS-G5 technologies. In the data collection phase, samples of each signal type are collected at a sampling rate of 20 Msps and a center frequency of 5.9 GHz with 10 MHz bandwidth. For each technology, the dataset and a description of the dataset collection are available as an open source[1] .

Based on the data collection procedure, 6 different dataset clusters are collected at sampling rates of 1, 5, 10, 15, 20, and 25 Msps. The dataset size at each considered sampling rate is 7500 X M, where M can be 44, 220, 440, 660, 880, and 1100 for a sampling rate of 1, 5, 10, 15, 20, and 25 Msps, respectively. Figure 4 shows a spectrogram plot of sample signals taken from the dataset of each class at a 20 Msps sampling rate for a 1 s duration.

[1] https://gitlab.ilabt.imec.be/mgirmay/tech-rec-its-band

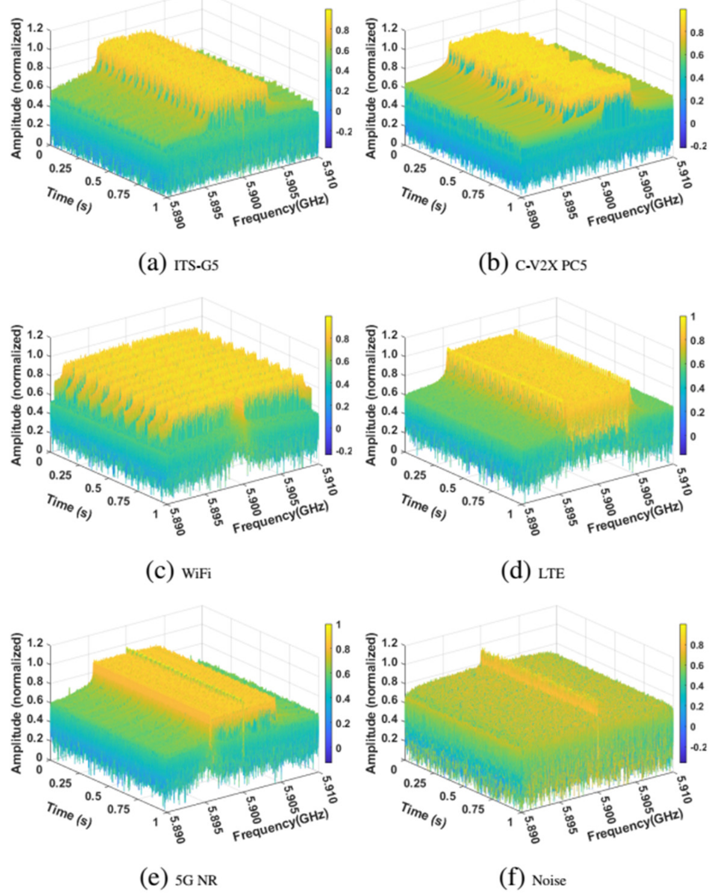


Figure 4: Spectrogram plot of sample signals taken from the dataset of each class at a 20 Msps sampling rate for a 1 s.

The collected in-phase and quadrature-phase (IQ) samples are pre-processed by introducing Additive White Gaussian Noise (AWGN) with -15, -10, …, 30 dB Signal to Noise Ratio (SNR) levels. This pre-processing is introduced to evaluate the performance of the TR model under different channel conditions. After the AWGN was introduced, Fast Fourier Transform (FFT) of the IQ samples was used to train and validate a Convolutional Neural Network (CNN) based TR model. Using FFT of the IQ samples as input leads to higher accuracy due to the more distinguishable frequency domain features of the signal. This makes the FFT preprocessing essential for ITS band spectrum sharing, where high accuracy is required to identify and protect safety critical information transmitted by the V2X users.

The structure of the CNN model and the implementation details are described in [5]. The CNN in [5] was modified to take into account five distinct classes, namely WiFi, LTE, 5G NR, C-V2X PC5 and ITS-G5, and the new dataset was collected at the 5.9 GHz ITS band. The classification accuracy of the proposed technology recognition model is computed to evaluate the performance of the CNN model as it identifies the captured samples from the considered wireless technologies. The classification accuracy indicates the fraction of correctly identified outcomes from the total predictions. To assess the model's effectiveness in different SNR conditions, various levels of noise ranging from -15 to 30 dB are introduced. Figure 5 shows the confusion matrix of the proposed technology recognition models with different sampling rates at 0 dB SNR. Each column in this figure represents the true label, and each row represents the predicted labels. The accuracy of the CNN-predicted tags and true labels indicates that the classification probability of proposed CNN models increases as the sampling rate used increases.

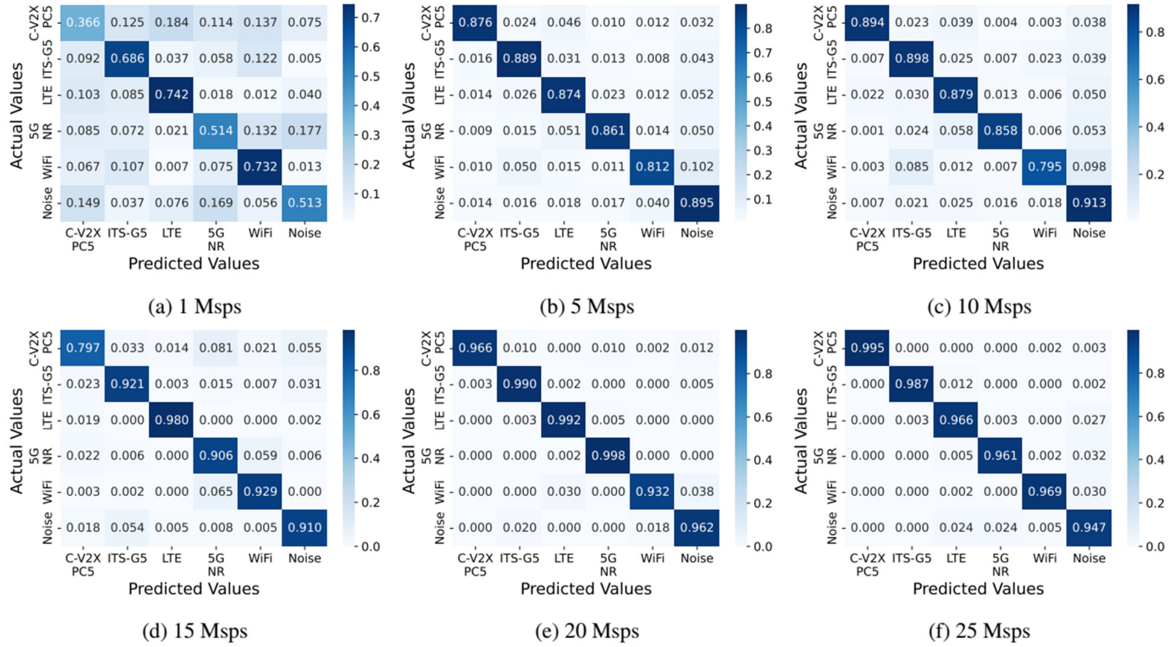


Figure 5: Confusion matrices for CNN models using different sampling rates at 0 dB SNR.

Figure 6 shows the accuracy of the proposed technology recognition model at different sampling rates. The classification accuracy results are obtained by assessing the model's performance on 30% of the total dataset samples that are used to validate and test the CNN model. Figure 6 shows the classification accuracy of the technology recognition models at different sampling rates as the SNR of the received signal varies from -15 to 30 dB. The figure illustrates that the CNN model's classification accuracy increases when a higher sampling rate is used. For an SNR of 0 dB, we can clearly observe that setting the sampling rate to 15 Msps or higher leads to a classification accuracy of more than 90%. More precisely, using the CNN model that uses the highest considered sampling rate (25 Msps) offers an excellent accuracy of 97.5% as compared to the lower classification accuracy of 48.5% for the CNN model that employs the lowest considered sampling rate (1 Msps).

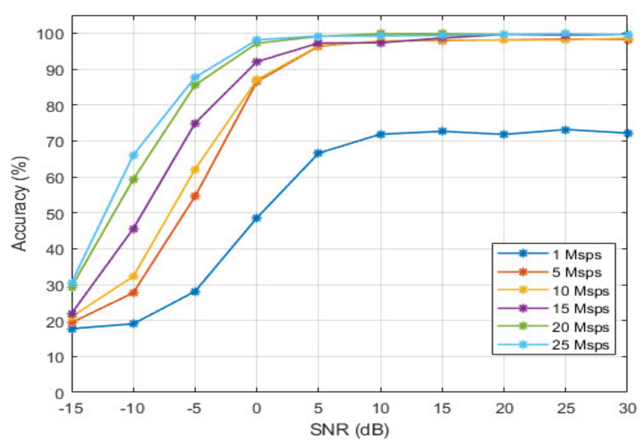


Figure 6: Classification accuracy in relation to SNR for the proposed technology recognition model using different sampling rates.

For SNR values higher than 5 dB, the classification accuracy can reach up to 73.2% when the CNN model is trained and validated with 1 Msps sampling rate. For the rest of the CNN models, the classification accuracy obtained is higher than 98%. However, the classification accuracy of the CNN model with 1 Msps sampling rate drops to 28% if the SNR of the received signal drops to -5 dB, while a classification accuracy of 86% is achieved with the CNN model that employs a 20 Msps sampling rate. For SNR values higher than 10 dB, an excellent classification accuracy (higher than 90%) is achieved with a lower sampling rate of 5 Msps.

3.1.2 Traffic Characterization

The characterization process starts with the determination of the transmission pattern of each identified signal. The transmission pattern of an identified signal represents the T_{ON} and T_{OFF} duration statistics within each characterization window. The T_{ON} indicates the time duration where an identified technology occupies the channel continuously, whereas T_{OFF} represents the silent duration in between the transmissions. For ITS-G5 and C-V2X PC5 technologies, the minimum possible frame duration (minimum possible T_{ON}) is greater than two consecutive resolution windows. Hence, a resolution window identified as a different technology in between two resolution windows identified as ITS-G5 or C-V2X PC5 is changed to match the bounding resolution windows. This process was introduced to reduce the effect of misclassification on the frame characterization process for the high priority safety critical ITS-G5 and C-V2X PC5 technologies. Figure 7 shows an example of the transmission pattern of actual received signal from the ITS-G5 transmitter with 50 pps for 0.15 s duration. The figure also shows the characterized ITS-G5 transmission pattern using technology recognition.

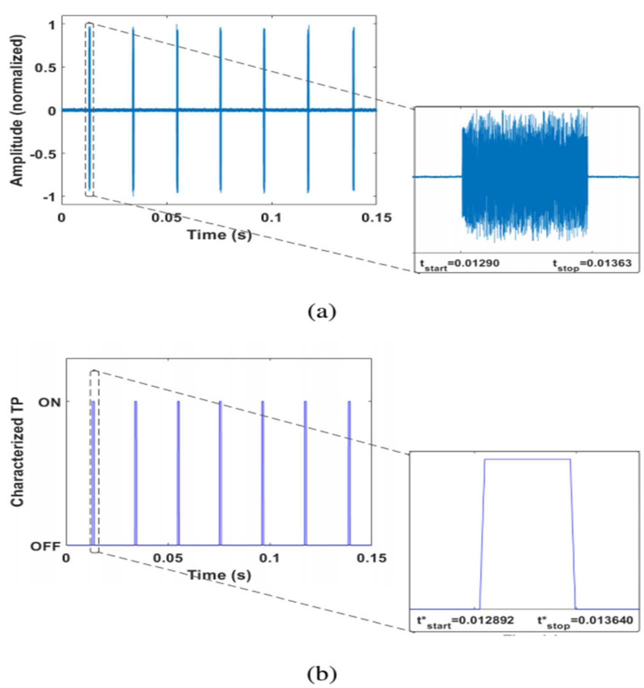


Figure 7: a) Actual transmitted Infrastructure to Vehicle Information (IVI) frames from ITS-G5 transmitter (with 50 pps transmission) in 0.15 s duration b) Corresponding Transmission Pattern (TP) characterized using the proposed technology

Generally, it can be observed that the proposed technology recognition model followed by the traffic characterization process can be used to accurately classify and characterize the traffic from each technology. The estimated traffic characteristics can be used to develop spectrum sharing schemes in future 6G networks.

3.2 Device Discovery

With the advent of autonomous vehicles, it will be necessary to develop strategies to recognize, identify and count the number of people and vehicles present in a given place of interest for road traffic.

One approach is to detect the devices present in an area of interest so that the number of people present can be extrapolated with the aim of providing context awareness to optimize the behaviour of autonomous vehicles and thus enhance road safety.

Device/person recognition cannot be addressed by a single technological solution. This knowledge must be built through the use of different sensing and recognition technologies that allow to draw a global map of the situation.

In the case of IoT sensing nodes, the information collected from the devices detected in the range of each IoT sensing node is made available to DEDICAT 6G platform as shown in Figure 8 to increase the knowledge of the environment and assist in enabling dynamic coverage management and connectivity extension.

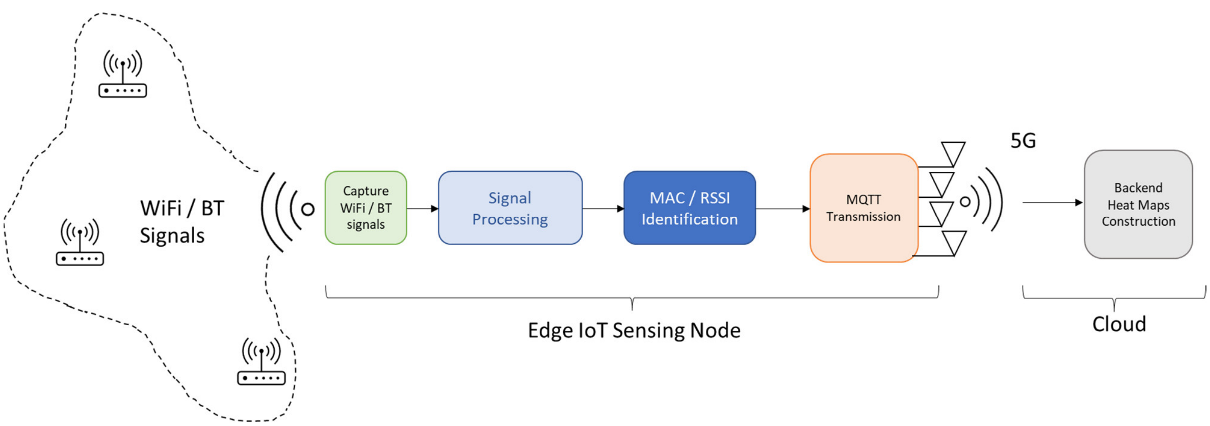


Figure 8: IoT Sensing Node Architecture

3.2.1 IoT Sensing Nodes

IoT Sensing Nodes have been designed to address the problem of having an estimation of people and vehicles (VRUs) connected to the DEDICAT 6G system or not present in the area of interest through the analysis of general-purpose signals, WiFi and Bluetooth (BT). By developing a device with the ability to scan the signals that are in the coverage area of each IoT sensing node, the system is able to process the signals captured at the edge, extract the information identifying the device's MAC and their Received Signal Strength Indicator (RSSI) in real time and send the information via Message Queuing Telemetry Transport (MQTT) through low latency communication links to be stored in a cloud database for further processing.

To ensure that the system is able to handle the volume of data sent, the device has a high capacity 5G communications module that allows data to be sent to the server with low latency, ensuring that no information collected is lost and processed data from detected devices is available in the DB in real time.

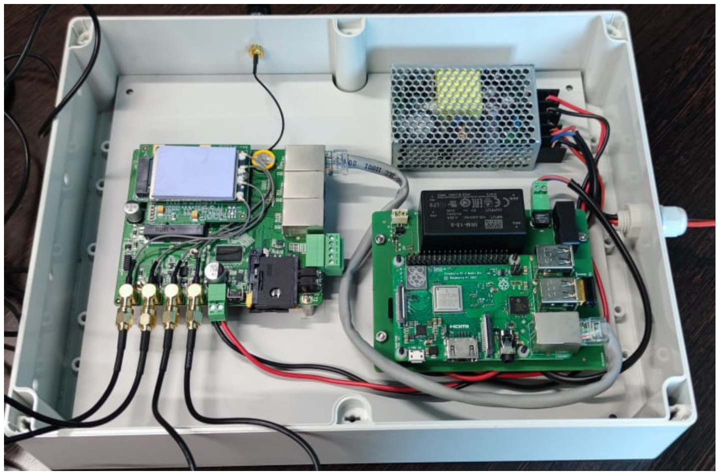


Figure 9: IoT Sensing Node modules detail

Figure 9 shows the different modules which integrate the IoT Sensing Node. The IoT Sensing Node provides a 5G communication link with dual purpose. The main function is to ensure that the information collected by the device is sent to feed the heat map micro service. With this information and the subsequent processing, it is possible to provide context awareness of the system's coverage area so that the information can be used to improve the creation of Local Dynamic Maps (LDM). In addition, as a second function, it serves as a backup communications access point for nearby connected DEDICAT 6G devices, VRUs registered in the system, which need to send their location information to the DEDICAT 6G platform to collaborate in the generation of the global LDM. Figure 10 shows details about the 5G antennas which are integrated in the IoT Sensing Node.

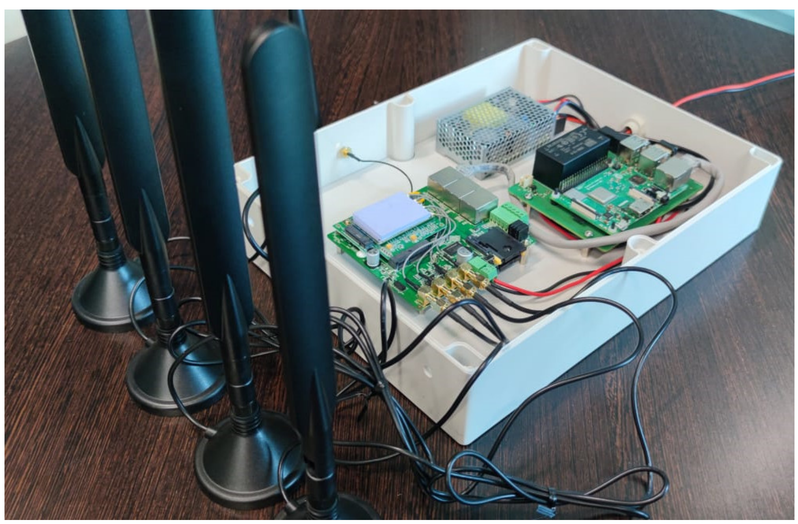


Figure 10: Detail of IoT Sensing Node 5G antennas

For connected users to be able to make opportunistic use of the IoT sensing node communications resources, two conditions must be met: these users must be registered in the IoT sensing node micro service and one of the IoT sensing nodes must have detected the device in the coverage area. For this, through the micro service dashboard, the MAC of the connected user is registered and the micro service backend checks that the device is in the coverage area of one of the nodes, including the VRU in the system whitelist. Figure 11 illustrates the procedure for registering in the system whitelist.

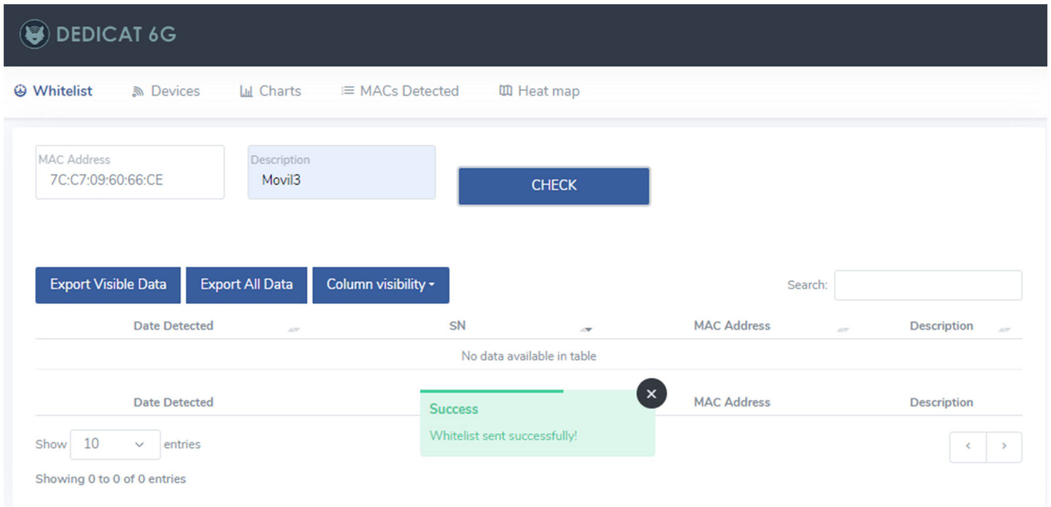


Figure 11: Device whitelist registration

Figure 12 presents an example of MAC devices detected by the sensing nodes for the micro service dashboard.

Editar	Copiar	Borrar	211582	2022-08-18 13:23:55	00000000cb872da4	WIFI	1E:A9:5B:55:AE:CB	-83
Editar	Copiar	Borrar	211583	2022-08-18 13:23:55	000000008d8df1f	WIFI	3E:AD:B6:43:F1:95	-85
Editar	Copiar	Borrar	211580	2022-08-18 13:23:53	000000008d8df1f	WIFI	1E:1F:55:04:D8:5E	-45
Editar	Copiar	Borrar	211579	2022-08-18 13:23:53	000000008d8df1f	BT	FC:A5:D0:06:E4:04	-67
Editar	Copiar	Borrar	211581	2022-08-18 13:23:53	00000000cb872da4	WIFI	CA:A4:61:87:C7:14	-43
Editar	Copiar	Borrar	211578	2022-08-18 13:23:51	000000008d8df1f	WIFI	36:B7:B9:D2:BB:84	-87
Editar	Copiar	Borrar	211576	2022-08-18 13:23:51	000000008d8df1f	BT	10:4E:89:E0:E7:8A	-85
Editar	Copiar	Borrar	211577	2022-08-18 13:23:51	00000000cb872da4	WIFI	C2:DE:95:EE:71:66	-83
Editar	Copiar	Borrar	211575	2022-08-18 13:23:49	00000000cb872da4	BT	D5:BB:FF:22:11:94	-67

Figure 12: Sensing Nodes DB MACs detected

Thereafter, the connected device (VRU) can make use of the IoT Sensing Node's communication resources to send its position information to the DEDICAT 6G system. When the device is out of range of the IoT sensing nodes coverage area, the connectivity permission is removed, reserving the IoT sensing node communication capability for another user who needs it.

Figure 13 presents an example of the DEDICAT 6G micro service dashboard showing the list of detected devices identifying the type of signal (WiFI or Bluetooth), the MAC address and the RSSI level.

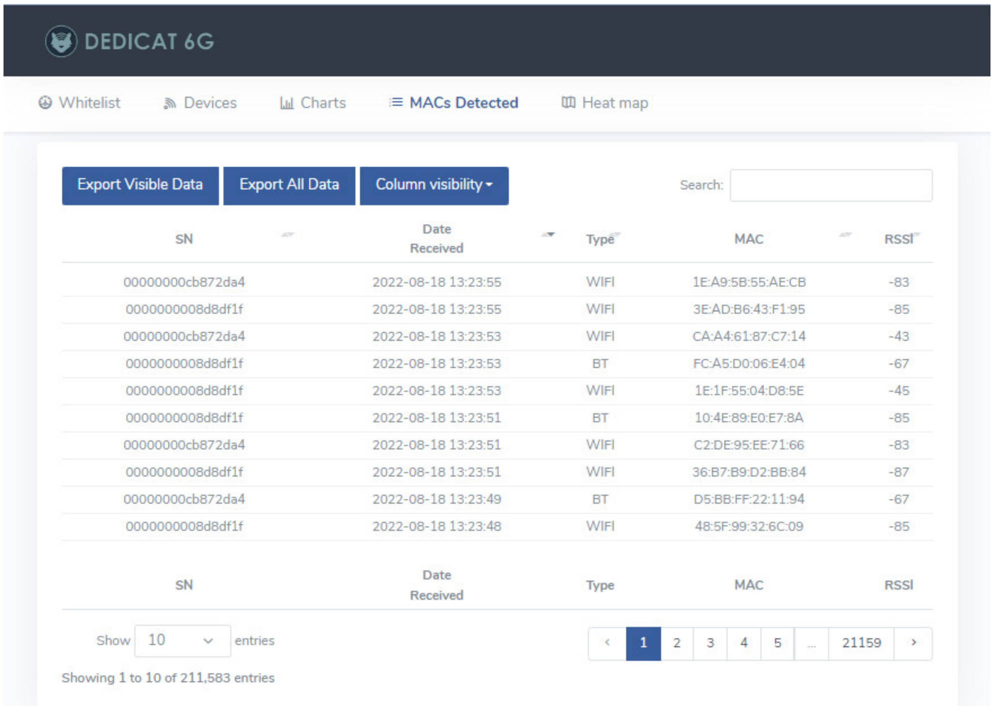


Figure 13: Detail of MAC Detected Dashboard

3.2.2 Heat Maps and detection charts

Thanks to the information collected by the IoT sensing nodes, the backend of the micro service generates, on the one hand, detection charts in which the history of device detections that have been performed by each of the nodes and on the other hand, real-time heat maps of the occupation by connected and non-connected users present in the area of interest on the road can be consulted. This information can be sent to collaborate in the construction of a global map of VRUs and cars (LDM) present in a given area that generates context awareness that facilitates coverage extension tasks.

As can be seen in Figure 14, through the IoT sensing node micro service dashboard, charts with the history of detected devices of each node can be displayed and specific searches can be performed.

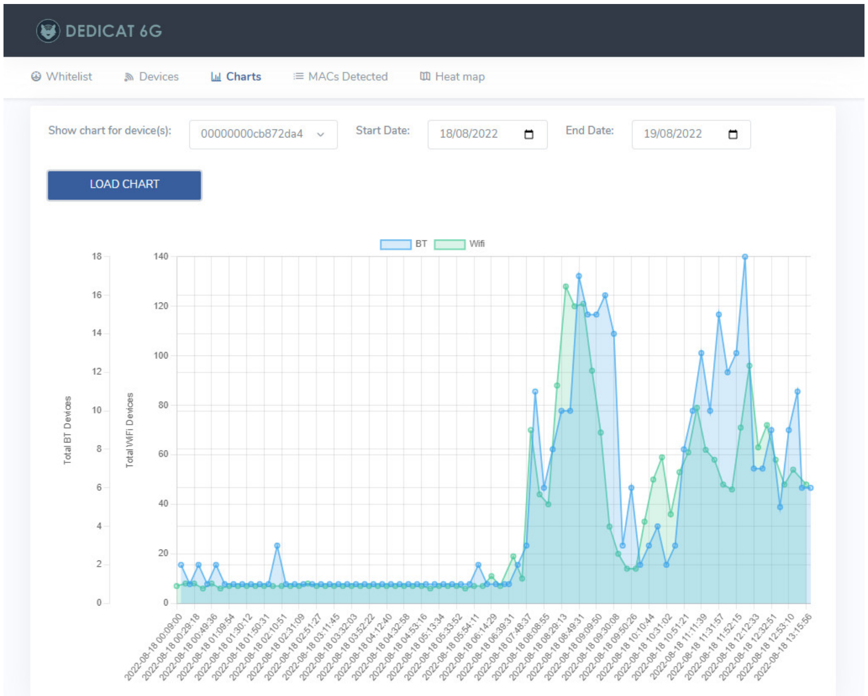


Figure 14: Device detection graphs

The information collected in real time by each of the IoT sensing node is sent over a 5G communications link to the micro service server hosted in the cloud. In the backend of the micro service, the information is processed to build a heat map around each of the nodes that show an estimation of the VRUs occupancy of the coverage area of the node.

```
var datos = JSON.parse(data);

// Eliminamos los resultados que no tengan coordenada GPS o estén sin definir
Object.keys(datos).forEach(k => (!datos[k].lat && datos[k].lat !== undefined) && delete datos[k]);

for (var i = 0; i < Object.keys(datos).length; i++) {

    result[i] = Object.values(datos)[i];

    var marker = {
        lat: result[i].lat,
        lng: result[i].lng,
        sn: result[i].sn,
        count: result[i].count
    }

    markers.push(marker);

}

testData = {
    max: Object.keys(datos).length,
    data: result
};

heatmapLayer = create_heat(110);

if (map && map.remove) {
    map.off();
    map.remove();
}
```

Figure 15: Data processing of device detections

To adjust the display of the heat maps, the developed software (Figure 16 and Figure 17) has a configuration layer that allows adjusting the system estimation through parameters such as detection radius, scale and display parameters such as colour range, etc. so that a map can be displayed and any user can visualize the occupancy estimation of a certain area at a glance.

```
heatmapLayer.setData(testData);

for (var marker_counter = 0; marker_counter < markers.length; marker_counter += 1) {

    var marker = L.marker([markers[marker_counter]['lat'], markers[marker_counter]['lng']], {
            draggable: false,
        })
        .addTo(map)
        .bindTooltip(
            '<b>SN: </b>' + markers[marker_counter]['sn'] +
            '<br><b>Lat: </b>' + markers[marker_counter]['lat'] +
            '<br><b>Lng: </b>' + markers[marker_counter]['lng'] +
            '<br><b>Datetime: </b> Last hour received' +
            '<br><b>Total devices: </b>' + markers[marker_counter]['count']).on('click', function(e) {
            var latlng = this.getLatLng();

            alertInfo("Coordinates", latlng.lat + ", " + latlng.lng);

        });
}

L.control.scale().addTo(map);

// make accessible for debugging
layer = heatmapLayer;
```

Figure 16: Configuration for Heat map visualization

```
function create_heat(new_radius) {

    heat = new HeatmapOverlay({
        "radius": new_radius,
        "maxOpacity": .8,
        // scales the radius based on map zoom
        "scaleRadius": false,
        // if set to false the heatmap uses the global maximum for colorization
        // if activated: uses the data maximum within the current map boundaries
        //   (there will always be a red spot with useLocalExtremas true)
        "useLocalExtrema": true,
        // which field name in your data represents the latitude - default "lat"
        latField: 'lat',
        // which field name in your data represents the longitude - default "lng"
        lngField: 'lng',
        // which field name in your data represents the data value - default "value"
        valueField: 'count'
    });
    return heat;
}
```

Figure 17: Configuration of Heat map nodes

Figure 18 shows an example of the visualization of the heat maps generated by the system. This map has been obtained during the laboratory tests carried out to validate the developments made. In these tests, three nodes have been deployed in three different buildings and it can be seen that most of the detections are made in the main building of the company and therefore the maximum occupancy detected occurs around that node. The third node further away, is in a building with low occupancy and the detections can be seen to have a very low RSSI (low intensity and blue colour), indicating that most of them are produced by the passage of people in the vicinity of the building or on the adjacent road, as opposed to the other two devices, which are de-detecting the continuous presence of workers in the buildings (red/yellow colour and high intensity).

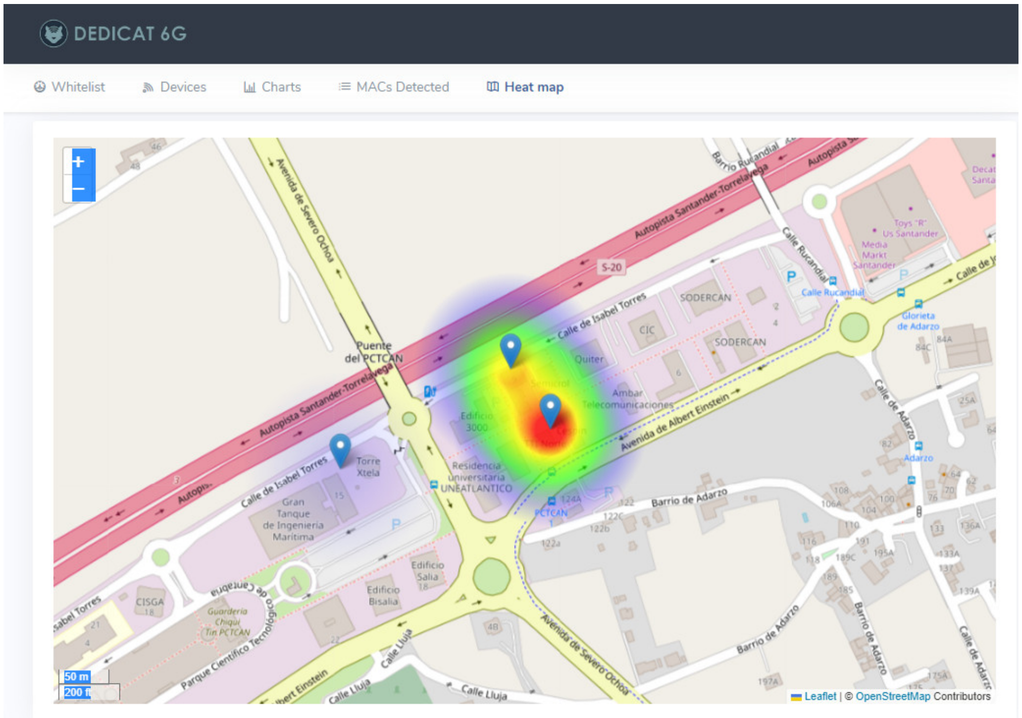


Figure 18: Heat Map Visualization of device occupancy

3.3 Strategy for tracking VRUs with RSUs

A large percentage of all fatal traffic accidents occur at intersections [10][11]. These accidents are partly because in many densely populated areas, obstacles such as buildings, fences, and plants obscure the perimeter of intersections. Therefore, accurate information about the location and movement path tracking of hidden or obscure VRUs is very important to reduce road accidents. One way to obtain information about obscured subjects is to per-form real-time object detection at a roadside unit (RSU) located at an intersection and transmit the collected information to the driver of a nearby vehicle using wireless vehicle-based infrastructure (V2I) communication. In other words, a typical approach for these applications is to track the movement of VRUs on the road in real-time through camera sensors installed in the RSU and warn the driver of a hazardous situation via vehicular-to-vehicular (V2V) communication. The VRU tracking strategy based on image processing from the camera sensor installed in the RSU is shown in Figure 19.

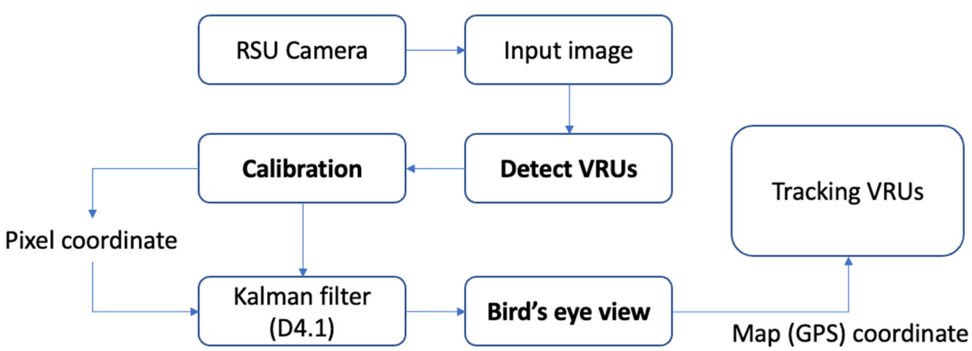


Figure 19: VRU tracking strategy

3.3.1 Detect VRUs

Cameras are the most widely used in intelligent vehicles and ITS for target classification. It offers the possibility to detect rich objects such as lanes, traffic signs, etc. Compared to in-vehicle cameras, the RSU's cameras are mounted higher, allowing for wider road coverage. The input image passes through the initial image form, and then processes segmentation, object detection and classification to detect VRUs. We applied the YoloV4[12] framework, a target detection algorithm that performs image-based target detection and localization, on NVIDIA Jetson AGX Orin, which serves as the edge computing platform for our RSU system.

The VRU detection results are the bounding box and classification information of targets. The position information of the detected VRU can be obtained by converting the pixel coordinate system to the geodetic coordinate system through camera calibration.

3.3.2 Calibration

To summarize the calibration process, the calibration parameters of the camera can be divided into an internal parameter matrix and an external parameter matrix, and the mapping relationship is as follows.

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \cdot \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & \gamma & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

where K is the camera's intrinsic parameter matrix, $\alpha_x = f \cdot m_x$ and $\alpha_y = f \cdot m_y$ represent focal length in terms of pixels, m_x and m_y are the inverses of the width and height of a pixel on the projection plane, f is the focal length in terms of distance, γ represents the skew coefficient between the x and y axis, u_0 and v_0 represent the optical center of the image, R and T are the extrinsic parameters (the rotation matrix and translation matric, respectively) which denote the coordinate system transformations from 3D world coordinates to 3D camera coordinates. Figure 20 shows the coordinate mapping from the world coordinate system to the pixel coordinate systems.

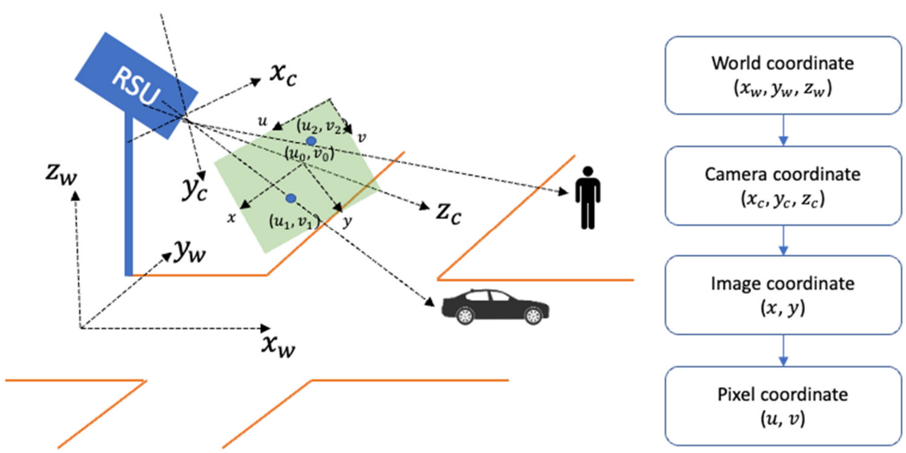


Figure 20: Coordinate mapping from RSU's camera

3.3.3 Bird's eye view

VRU detection and tracking is pixel coordinate information from RSU camera. For sharing VRU tracking information through V2X communication, it needs to do a perspective transformation from a perspective view to Bird's eye view, which is a basically a top-down representation of a scene. In other words, by calculating GPS coordinates from pixel coordinates, VRU information estimated by RSU can be easily utilized in a vehicle connected to V2X. Figure 21 illustrates the perspective transformation.

The first step is to select 4 points (p_1, p_2, p_3, p_4) from the original image that will be the vertices of the plan to be transformed. It noted that these points would be better to form at least two opposite, parallel rectangles to make the proportions equal when the transformation occurs.

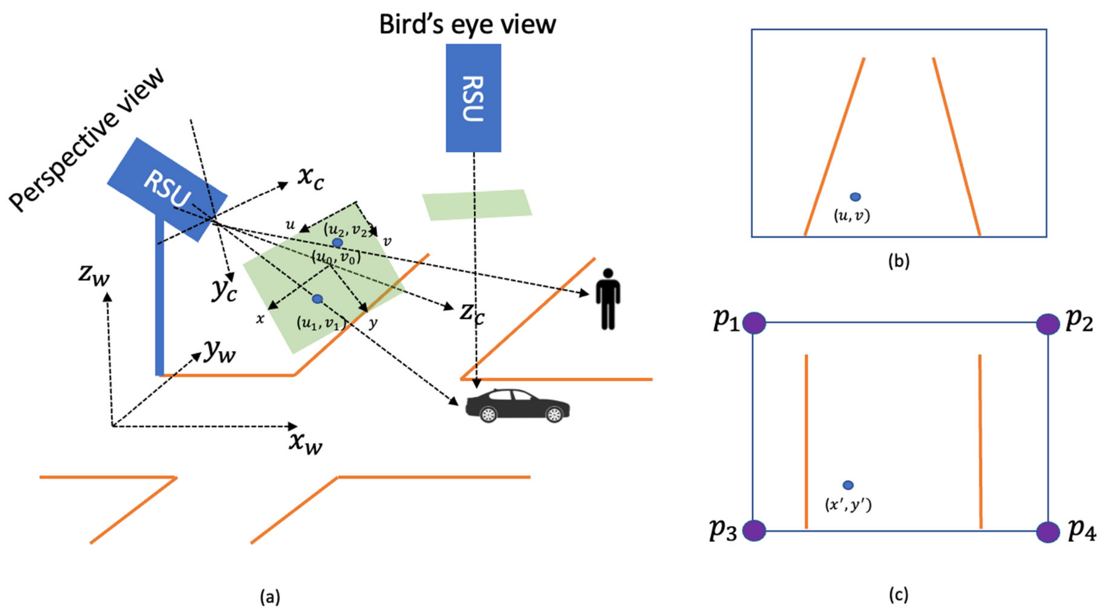


Figure 21: Perspective transformation (a) Perspective transformation (b) Perspective view image (c) Bird's eye view image

Then, the GPS coordinated (x', y') can be calculated by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

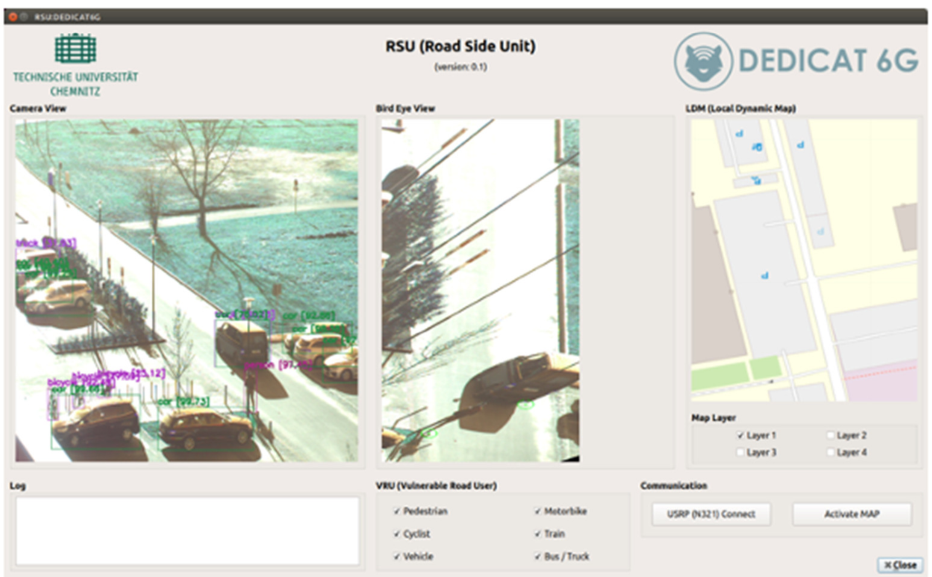
where a is the x-coordinate of the center of the upper left pixel, b is the pixel size in the x-direction, c is the rotation of x-axis, d is the y-coordinate of the center of the upper left pixel, e is the rotation of y-axis, and f is the pixel size in the y-direction.

3.3.4 RSU application for tracking VRUs

The RSU application for VRU detection and tracking (initial version, to be updated in D6.2) has been developed as shown in Figure 22, and there is a demo video at the link below.

https://www.dropbox.com/s/sq70bd0hntigsbq/TUC_RSU_Demo_Ver0.1.mov?dl=0

(a)



(b)

Figure 22: RSU application (a) VRU detection and tracking (b) GUI integration

4 Decision making mechanisms for dynamic coverage and connectivity extension

In this section, we describe and evaluates several network operation decision making functionalities for MAP/UE association:

- a centralized UE association and MAP placement making a trade-off between the network cost minimization (number of MAPs) and user utility (spectral efficiency) maximization,
- a distributed UE association and MAP placement deciding the number and the optimal location of MAPs, which maximizes the throughput and ratio of well-deserved users while minimizing the number of drones deployed
- a heterogeneous MAP-assisted networks with machine learning to maximize the QoS satisfaction level and the energy efficiency

We also describe coverage extension decision making (i.e. MAP operation managing mobility and swarm operation) and assistance to Intelligence distribution in the coverage extension.

4.1 Mechanisms to manage vehicular based MAPs

Within DEDICAT 6G, vehicles acting as MAPs that offer multi-RAT capabilities are investigated. Multiple RAT technologies can be used for the deployment of vehicular based MAPs including LTE, 5G-NR, and Wi-Fi. Towards 6G, it is envisioned that spectrum management and utilization will become more flexible and dynamic compared to today's static and conservative approaches. Hence, a RAT used by a vehicular MAP might be able to operate in any frequency band (as long as this is supported by the frontend and allowed by the regulators) offering coexistence with other technologies and protecting potential incumbents.

For our proof of concept and without loss of generality, we assume that the MAPs operate in any channel of the 5.9 GHz ITS band, in which short-range V2X communications take place today. As such, the deployed MAPs should be able to coexist with other MAPs operating in those channels based on LTE, 5G-NR, Wi-Fi technologies, and protect transmissions from incumbents such as C-V2X PC5, and ITS-G5. Hence, efficient knowledge building, RAT selection and configuration schemes are needed to satisfy the application requirements, achieving the targeted coverage and QoS, as well as enabling harmless coexistence. Figure 23 shows the main procedures for the management and configuration of vehicular based MAPs. Initially, the Knowledge Building process takes place aiming to estimate the characteristics of the wireless environment. In this phase, active wireless technologies are identified and characterized. This technology recognition and traffic characterization process is presented in Section 3.1.

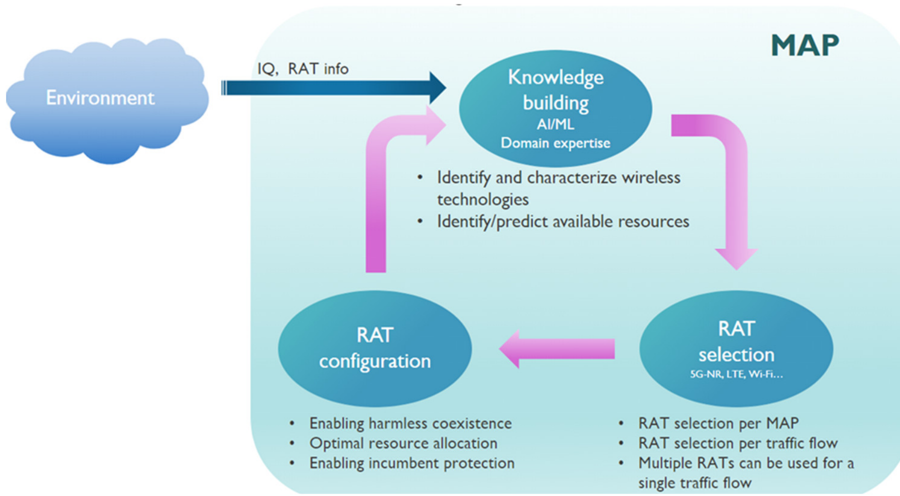


Figure 23: RAT management mechanism in Vehicular MAPs.

4.1.1 RAT Selection

After the knowledge building process, the RAT Selection process follows. A newly activated MAP selects the most suitable RAT(s) based on the RAT selection scheme, which considers the characteristics of the wireless environment obtained in the knowledge building. From the information collected in the technology recognition and characterization process, a newly activated MAP will be able to determine the type of active co-located wireless technology and estimate their corresponding traffic load and transmission pattern. The RAT selection process must consider other RATs within the MAP's coverage, potentially used by other co-located vehicles, roadside infrastructures, wireless local area networks, private networks, etc. For this reason, a MAP should be capable to select a RAT based on the information obtained in the knowledge building phase.

Along with the available resource based on the information derived from the knowledge-building process, the RAT selection process will also consider the traffic load of the newly activated MAP itself. RAT selection process should also consider several factors such as the capabilities of the MAP, the traffic demands of the applications/services and the identified characteristics of the wireless environment.

As the traffic load and the characteristics of wireless networks change dynamically, MAPs can activate their RATs based on the statistics obtained during the RAT activation. However, the traffic load of each MAP can arbitrarily change after some time leading to inefficient resource utilization and higher interference. In such scenarios, load-balancing through multiple RATs can be used to improve the quality of service. Therefore, the RAT selection process is performed continuously to adapt to the traffic load of the host MAP and the available wireless resources.

On the other hand, a MAP can have traffic contents that require higher reliability. Hence, multiple RATs can be used by a MAP for redundant transmissions on a traffic flow basis. Such redundant transmissions can be triggered based on signal quality indicators used in different RATs of a MAP. Once the need for redundant transmission is determined for a traffic flow, the MAP can activate a new RAT enhancing the redundancy of that traffic flow.

In general, the activated RATs then can be used in different ways to serve the applications. For instance, a single RAT can be used for a specific MAP, a RAT can be selected per traffic flow or multiple RATs can be used for a single traffic flow for redundancy or optimization purposes. For instance, some technologies such as LTE/5G-NR can be selected when no or little other traffic is identified, while Wi-Fi could be a better choice in more busy channels because of its inherent coexistence mechanisms (CSMA/CA).

4.1.2 RAT Configuration

When a MAP is initialized, it should try to avoid an occupied channel and operate in a free one. If this is not possible, then it should use the less busy channel, providing coexistence to other MAPs and, if required, incumbent protection to C-V2X PC5 and ITS-G5. In the RAT configuration phase, the main aim is to optimize the selection and management of the available wireless resources in order to provide harmless coexistence, incumbent protection and increase the overall performance of the MAP.

Similar to the RAT selection process, the RAT configuration process is also based on the statistics collected in the knowledge building phase. The RAT configuration scheme will be implemented in such a way that the MAP uses transmission patterns that lead to optimal resources utilization while maintaining harmonious coexistence with other wireless technologies. In the RAT selection process, the best RATs will be selected based on the traffic load of the MAP and the information collected from the environment. However, the selected RATs must be configured properly to achieve optimal resource utilization, while maintaining harmonious coexistence with other co-located transmissions. The RAT configuration scheme considers the wireless environment characteristics gathered during knowledge building and the traffic characteristics of the active RATs in the host MAP. However, the traffic load of each MAP will change over time. Hence, the RAT configuration scheme is executed continuously so that the MAP constantly reconfigures the resources used by each active RAT.

4.1.3 Preliminary implementation

A knowledge building solution that includes technology recognition and traffic characterization is presented in Section 3.1. As an initial step, we have started the implementation of a Dynamic Spectrum Sharing (DSS) based RAT configuration solution for vehicular MAPs that deploy 4G LTE and 5G NR RATs. As a significant range of frequency channels are already occupied by LTE network, it was essential to discover a solution to implement 5G network on the existing bands. DSS is a method that enables the utilization of spectrum from 4G LTE to 5G NR by implementing the co-existence of the two RATs in the same network without requiring each radio to monopolize a dedicated frequency band. Consequently, DSS has been one of the factors for the accelerated roll-out enabling seamless transition from 4G to 5G by integrating 5G systems in existing 4G bands without hard/static spectrum re-farming [13].

In LTE-NR DSS, Multi-Broadcast Single-Frequency Network (MBSFN) subframes are utilized to enable coexistence [13]. MBSFN is a feature utilized for LTE point-to-multipoint transmissions such as Evolved Multimedia Broadcast Multicast Services (eMBMS). A frame is 10 ms in duration and consists of 10 subframes. Within multicast frame, up to six subframes are reserved for MBSFN transmissions whereas the rest of them can be used for unicast and paging information. In other words, multicast and broadcast traffic is limited to a maximum of six subframes, out of the 10 subframes of each LTE frame. The remaining subframes are designated for unicast traffic transmissions, as well as control and synchronization information. In each MBSFN subframes of the LTE frame, the first two OFDM symbols are used for unicast control information while the remaining OFDM symbols of each subframe are used for multicast traffic and control channels [14]. In DSS, the OFDM symbols assigned for multicast traffic and control information are completely muted, so that the duration reserved for these symbols are used for 5G NR transmission rather than eMBMS. The first two OFDM symbols of each empty MBSFN subframe are used for LTE Physical Downlink Control cHannel (PDCCH) and Cell specific Reference symbols (CRS), while 5G NR signals can be transmitted in the remaining symbol duration. Figure 24 shows the frame structure of LTE in MBSFN based DSS. The figure shows that a frame with empty MBSFN subframes occurs every MBSFN scheduling period (M). As per standard [14], the number of multicast subframes per MBSFN frame can be N subframes where N ϵ {1,2,3,4,5,6} and periodicity of such frames can be every M frames where M ϵ {1,2,4,8,16,32}.

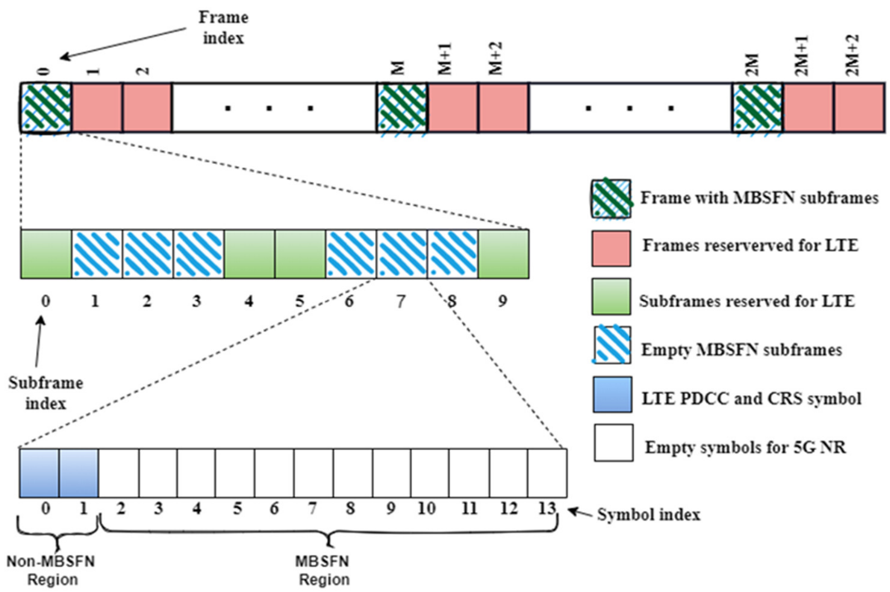


Figure 24: LTE frame structure in DSS mode.

One of the most important issues in introducing new features in existing RATs is considering compatibility of the solution with the standard UEs. DSS based RAT configuration is complaint with standard LTE UE and NR UE. NR UE typically acquires scheduling information from NR System Information 1(SIB1). The LTE UE on the other hand decodes the LTE SIB2 to determine the MBSFN subframes.

This DSS based RAT configuration feature will be implemented in SDR based end-to-end LTE network. As an initial step, the implementation of empty MBSFN subframes is introduced in the Software Radio System RAN (srsRAN) [15] SDR solution. This includes implementation of empty MBSFN subframes based on the queue length of the LTE traffic and modifications in eNB-UE signaling to enable adaptive MBSFN resource allocation. The eNB-UE signaling modification includes implementation of periodic LTE SIB update feature in MBSFN SDR solution for periodically conveying modified MBSFN parameters to the UE side. These initial implementations were made by introducing changes to the srsRAN SDR source code. The modifications are made on both eNB and UE sides. On the eNB side, a flexible empty MBSFN allocation mechanism is implemented. Similarly, the SIB generation process is modified in such a way that the SIB2 and SIB13 messages are updated in every M frames. On the other hand, the periodic SIB decoding feature is implemented on the UE side. This feature enables the UE to decode the SIBs every 160 ms and adapts according to the received resource configuration information. Figure 25 shows the spectrogram of an LTE frames when 6 empty MBSFN subframes are configured in every frame.

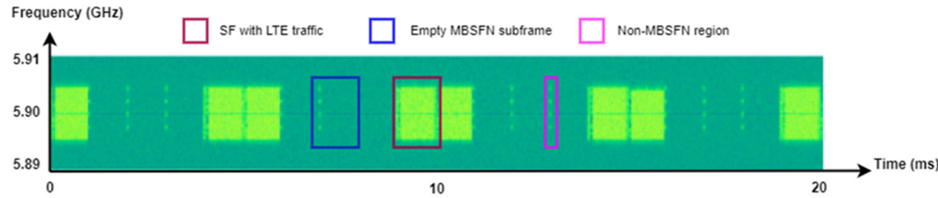


Figure 25: Spectrogram of LTE frames configured in DSS mode.

The DSS based RAT configuration solution will be extended by integrating the outcome of technology recognition and traffic characterization obtained in knowledge building phase. In the future step, the technology recognition and traffic characterization process will be used to determine the channel occupancy of 5G NR. After that, the LTE eNB decides the number of empty MBSFN subframes based on the characterized 5G NR channel occupancy and LTE traffic queue. Furthermore, MAPs with single RAT capability i.e. either LTE or 5G NR are considered in this preliminary implementation. Hence, a RAT initiation is considered without RAT selection procedure. As a future step, we plan to include RAT selection algorithm considering MAPs with multi-RAT capability. The DSS based RAT configuration will also be extended considering multiple RATs and different QoS requirements.

4.2 Cost efficient and QoS Aware user association and 3D MAP placement

6G networks require a flexible infrastructure to dynamically provide ubiquitous network coverage. Mobile Access Points (MAP) deployment is a promising solution. In this section, we formulate the joint 3D MAP deployment and user association problem over a dynamic network under interference and mobility constraints. First, we propose an iterative algorithm to optimize the deployment of MAPs. Our solution efficiently and quickly determines the number, position and configuration of MAPs for highly dynamic scenarios. MAPs provide appropriate Quality of Service (QoS) connectivity to mobile ground user in mm-wave or sub-6GHz bands and find their optimal positions in a 3D grid. Each MAP also implies an energy cost (e.g. for travel) to be minimized. Once all MAPs deployed, a deep multi-agent reinforcement learning algorithm is proposed to associate multiple users to multiple MAPs under interference constraint. Each user acts as an independent agent that operates in a fully distributed architecture and maximizes the network sum-rate.

4.2.1 System model and problem formulation

4.2.1.1 System model

We consider a downlink network where *K* MAPs, (e.g., UAVs), are jointly deployed with a Macro Base Station (MBS) to provide ubiquitous coverage to *P* UEs. We define $\mathcal{A}$ the set of APs and $\mathcal{U} = \{u_0, u_1, \dots, u_{P-1}\}$ the set of UEs. We assume that each UE *i* is equipped with two antennas and can communicate at sub-6GHz and mm-wave frequencies with the MBS and MAPs, respectively. In this network, we focus on the optimization of MAP locations jointly with Radio Resource Management (RRM). In this context, let $l_{i,k} = \{x_{i,k}, y_{i,k}, z_{i,k}\}$ denote *k*-th possible location of the MAP *i* represented as a coordinate in 3D dimensional space as represented in Figure 26.

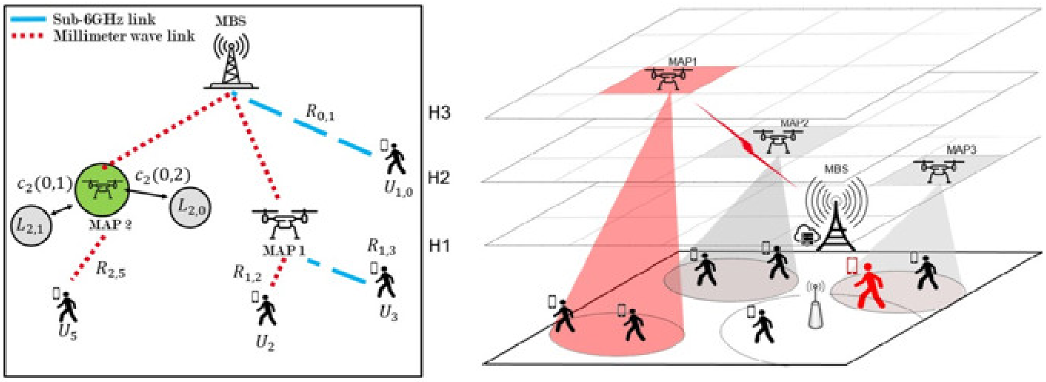


Figure 26: Cell architecture and 3D cell configuration

We denote with $\mathcal{L}_i = \{l_{i,k}\}_{k=0,\dots,L_{i-1}}$ the set of all possible locations of MAP *i*, and let denote $\mathcal{L} = \cup_{i\in\mathcal{A}} \mathcal{L}_i$ the set of all possible MAPs locations, which can be determined a priori using path-planning or defined as the set of possible safe locations of MAPs e.g., in urban area. We define $\ell_{i,k}$ as the binary variable which equals 1 if the MAP *i* is effectively deployed on its *k*-th location and 0 otherwise. Moving a MAP from one location to another incurs a certain cost either in terms of energy consumption, network operation or renting. We consider this aspect by defining $c_i(k,p)$ as the cost associated to moving a MAP *i* from location $k \in \mathcal{L}_i(t)$ to location $p \in \mathcal{L}_i(t+1)$:

$$c_i(k,p) = e_i(k,p)E_c + c_{i,0}$$

where E_c is the cost of a unit of energy, $e_i(k,p)$ is the energy consumed by MAP *i* to move from *k* to *p*, which is a function of the distance [16], and $c_{i,0}$ is a fixed cost due to, e.g., the renting of MAP *i*. Hence, we can define the total cost $C(t)$ incurred by the deployment of all MAPs as follows:

$$C(t) = \sum_{i\in\mathcal{A}} \sum_{k\in\mathcal{L}_i(t)} \sum_{p\in\mathcal{L}_i(t+1)} c_i(k,p)\ell_{i,k}(t)\ell_{i,p}(t+1)$$

Such a cost function may also vary depending on the number of targeted UEs, which will be served by the MAP. Therefore, our first objective focuses on determining the optimal subset $\mathcal{D} \subset \mathcal{L}$ of MAPs locations that minimize $C(t)$ by jointly minimizing the number of deployed MAPs and optimizing their deployment w.r.t. UEs QoS. Accordingly, our second objective focuses on the user association problem. This is because the optimal assignment of UEs to APs improves the network spectral efficiency and the perceived QoS of UEs [7]. Hence, let us denote with $x_{i,j}(t)$ the binary association variable, which equals 1 if UE *j* is associated with AP $i \in \mathcal{A}$ at time t, and 0 otherwise. We assume that all APs perform beam training in advance, so that they are able to set up an appropriate beam when a connection is established between AP *i* and UE *j*. We denote with $R_{i,j}(t)$ the corresponding communication rate, which is given by the Shannon capacity:

$$R_{i,j}(t) = B_{i,j}\log_2(1 + SINR_{i,j}(t))$$

where $B_{i,j}$ is the bandwidth allocated by AP *i* to UE *j* and $SINR_{i,j}$ the signal-to- interference-plus-noise ratio between AP *i* and UE *j*, which comprises intra-cell and inter-cell interference of both grounded and mobile APs. Then, given the data demand of UE *j*, $D_j(t)$ we define its QoS's satisfaction $\kappa_j(t) \in [0,1]$ as follows:

$$\kappa_j(t) = \sum_{i\in\mathcal{A}} x_{i,j}(t)\min\left(1, \frac{R_{i,j}(t)}{D_j(t)}\right)$$

Accordingly, we say that the QoS is fully satisfied when $\kappa_j(t) = 1$. Finally, to account with fairness in the association, we define the total network utility function as:

$$R_\alpha = \sum_{i\in\mathcal{A}} \sum_{j\in\mathcal{U}} x_{i,j}(t) U_\alpha(\min(R_{i,j}(t), D_j(t))$$

Where $U_\alpha(.)$ is the α-fair utility function given in [17] as:

$$U_\alpha(x) = = \begin{cases} (1-\alpha)^{-1} x^{1-\alpha}, & \forall \alpha \geq 0 \; and \; \alpha \neq 1 \\ \log x, & if \; \alpha = 1 \end{cases}$$

4.2.1.2 Channel model

The channel model varies according to several factors such as the radio environment (i.e. suburban, urban, dense urban, high rise building), the communication band (i.e. sub-6GHz and mm-wave), and the type of communication (i.e. ground-to-ground or ground-to-air). In general, the channel path loss PL_T for any communication link can be defined on the basis of the Line of Sight (LoS) conditions as follows:

$$PL_T = pPL_{LoS} + (1-p)PL_{NLoS}$$

where p is the LoS probability, PL_{LoS}, and PL_{NLoS} are the LoS and NLoS path loss, respectively.

Air/ground sub-6GHz path-loss.

Following [18], we define the LoS probability as a function of the elevation angle θ:

$$p(\theta) = c(\theta - \theta_0)^d$$

where $\theta_0 = 15°$ is the lowest possible angle and c, d are environmental parameters, and we compute the frequency-dependent path loss model as a function of the link type $l \in \{\text{LoS}, \text{NLoS}\}$:

$$PL_l = 20\log_{10}(d) + 20\log_{10}(f) - 27.55 + \chi_{\sigma_l}$$

Here, d the distance between the transmitter and the receiver, f is the carrier frequency in MHz, and χ_{σ_l} is the shadowing coefficient, which follows a normal distribution with a mean μ_l and a standard deviation σ_l, whose values are given in [18].

Air/ground mm-wave path-loss.

Here, we define the LoS probability as a function of the height of the transmitter (h_t) and receiver (h_r) and some environmental parameters [19]:

$$p(d) = \prod_{n=0}^{\max(0,\gamma(d))} 1 - e^{\left(-\frac{\gamma(d)\max(h_t,h_r)-(n+1/2)(h_t-h_r)^2}{2\varepsilon^2\gamma(d)^2}\right)}$$

Where $\gamma(d)$ represents the average number of buildings crossing the link between the transmitter and the receiver separated by a distance d and ε represents the random height of each obstacle, which follows a Rayleigh distribution. Hence, the distance-dependent path loss model is [20]:

$$PL_l = \alpha_l + 10\beta_l \log_{10}(d) + \chi_{\sigma_l}; \quad l \in \{\text{LoS}, \text{NLoS}\}$$

where, α_l, β_l depend on the radio environment.

Ground/ground sub-6GHz or mm-wave path-loss.

Here, the path loss model can be defined without considering the LoS probability [21]:

$$PL_T(d) = 10\alpha\log_{10}(d) + \beta + 10\gamma\log_{10}(f) + \chi_{\sigma_l}$$

Where d is the distance between the transmitter and the receiver, f the carrier frequency and χ_{σ_l} the shadowing effect.

4.2.1.3 Formulation of MAP deployment Problem

After the above definitions, we formulate the MAP deployment problem to minimize the total deployment cost as:

$$
\begin{aligned}
\underset{x_{i,j},\ell_{i,k}}{\text{minimize}} \quad & \frac{1}{T}\sum_{t=0}^{T-1} C(t) && (\mathcal{P}_1)\\
\text{s.t.} \quad & x_{i,j}(t),\ \ell_{i,k}(t) \in \{0,1\}, && \forall i,j,k,t \ \ (\mathcal{C}_1)\\
& C_i(t) \le C_{\max}, && \forall i \in \mathcal{A}\backslash\{0\} \ \ (\mathcal{C}_2)\\
& \sum_{j\in\mathcal{U}} x_{i,j}(t) \le N_i, && \forall i,t, \ \ (\mathcal{C}_3)\\
& \sum_{i\in\mathcal{A}} x_{i,j}(t) = 1, && \forall j,t, \ \ (\mathcal{C}_4)\\
& \kappa_j(t) \ge \mathcal{Q}_j, && \forall j,t, \ \ (\mathcal{C}_5)\\
& \sum_{k\in\mathcal{L}_i(t)} \ell_{i,k}(t) \le 1, && \forall t, i \in \mathcal{A}\backslash\{0\}, \ \ (\mathcal{C}_6)\\
& \sum_{i\in\mathcal{A}\backslash\{0\}} \sum_{k\in\mathcal{L}_i(t)} \ell_{i,k}(t) \le K_{\max}, && \forall t, \ \ (\mathcal{C}_7)
\end{aligned}
$$

The constraint $(\mathcal{C}_1)$ defines $x_{i,j}$ and $\ell_{i,k}$ as binary variables. The constraint $(\mathcal{C}_2)$ ensures that the deployment cost of a MAP is lower than the maximum cost C_{max}. The constraints $(\mathcal{C}_3)$ and $(\mathcal{C}_4)$ ensure that each AP *i* serves at most N_i UEs and that each UE is associated to exactly one AP. The constraint $(\mathcal{C}_5)$ guarantees the QoS satisfaction of each UE. Finally, the constraints $(\mathcal{C}_6)$-$(\mathcal{C}_7)$ guarantee that a MAP is deployed to at most one location at a time and that the total number of deployed MAP does not exceed K_{max}. It is worth noting that Problem $(\mathcal{P}_1)$ is non-convex and NP-hard, thus difficult to solve with classical optimizations techniques.

4.2.2 Proposed solution

Our proposed solution for deploying MAPs jointly considers the deployment cost, UEs' mobility, co-channel interference, and traffic request dynamic. One key challenge is that the optimal MAP deployment strategy strongly depends on UEs' traffic requests and the co-channel interference that will be generated, which is not known until UEs are fully associated. At the same time, the optimal association of UEs also depends on the MAP deployment. This ping-pong effect makes the problem very complex and difficult to solve. To limit such complexity, we first define a 3D grid of positions for MAPs. The discretization of the 3D space gives a finite number of solutions for the problem $(\mathcal{P}_1)$. However, the search space of possible solutions remains large, prohibiting any exhaustive search approach. Thus, we design SIMBA, a Scalable Iterative Monte-Carlo Based Algorithm, with low-complexity, which explores the search space to find (sub)-optimal solutions as illustrated in Figure 27.

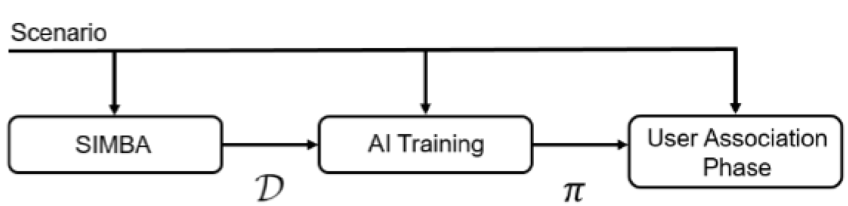


Figure 27: Proposed solution architecture

SIMBA, first performs Monte-Carlo explorations of MAPs deployment strategies and then exploits the best solution by adopting a standard user association algorithm based on maximum Signal-to-noise-ratio (MAX-SNR) to find the sub-optimal MAP deployment with low complexity. Next, based on SIMBA output, we apply our previously proposed Multi-Agent Reinforcement Learning (MARL) framework [7] to train a user association, which in contrast to the MAX-SNR algorithm, considers co-channel interference. We show that this approach is able to compensate the sub-optimality of SIMBA.

4.2.2.1 SIMBA: scalable iterative monte-carlo based algorithm

This section describes SIMBA, a low-complexity iterative algorithm (see algorithm in Figure 28), which finds near-optimal MAPs deployment solution (in terms of deployment cost and UEs' QoS satisfaction). SIMBA alternates a Monte-Carlo exploration and exploitation phases. During an exploration, SIMBA randomly samples a set $\mathcal{M}_k$ of k locations on which MAPs are deployed. Each time it deploys a MAP *i* on a location *p*, it updates a score associated to this location. Let $score_p^{(i)}(t)$ be such a score:

$$score_p^{(i)}(t) = \frac{1}{|\mathcal{U}_i|} \sum_{j \in \mathcal{U}_i} \kappa_j(t)$$

Where $\mathcal{U}_i$ is the number of UEs served by MAP *i*. As we consider interference and mobility, the score of a position is computed based on UEs' perceived QoS. Thus, higher the score, better the QoS of UEs served by a MAP deployed at that position. In the exploitation phase, a MAP is deployed at the location with the highest score. This location is no longer sampled until a solution of $(\mathcal{P}_1)$ has been found. Then we iterate over T episodes at the end of which we select the best solution of MAPs deployment $\mathcal{D}$. Eventually, given $\mathcal{D}$, we learn the optimal user association using a multi-agent reinforcement learning based approach.

Algorithm 1: SIMBA MAP Deployment Algorithm

Input: Define the set of MAPs possible locations $\mathcal{L}$ and the number of Monte-Carlo iterations M.
Initialize the score of locations: $\text{score}_p = 0, \ \forall p \in \mathcal{L}$.
Initialize the set of deployed MAP locations $\mathcal{D} = \emptyset$.
Set $C_{\min} = \infty$ and $k = K_{\max}$ (max. number of UAVs)
for $t = 1, \dots, T$ **do**
  Initialize an empty set of locations $\mathcal{D}_t = \emptyset$.
  for $s = 1 \dots K_{\max}$ **do**
    Step 1: *Monte-Carlo exploration*
    for $m = 1, \dots, M$ **do**
      Randomly sample k locations $\mathcal{M}_k(m) \sim \mathcal{L}$.
      Deploy a UAV to location p, $\forall p \in \mathcal{M}_k(m)$.
      Perform user association procedure[1].
      for $p \in \mathcal{M}_k(m)$ **do**
        Compute the score of p using Eq. (13).
        Update score_p of location p.
      end
    end
    Step 2: *Monte-Carlo exploitation*
    Store location $i = \arg\max_i \{\text{score}_i\}$ in $\mathcal{D}_t$.
    Deploy UAVs into the locations in $\mathcal{D}_t$.
    Compute $C(t)$, and $\kappa_j(t)$, $\forall j$.
    if $(\mathcal{C}_1)$-$(\mathcal{C}_7)$ *are guaranteed* **then**
      Break.
    else
      Remove i from $\mathcal{L}$ and set $k = k - 1$.
    end
  end
  Step 3: *Test Monte-Carlo solution*
  if $C(t) < C_{\min}$ **then**
    $C_{\min} = C(t)$
    Save current deployment $\mathcal{D} = \mathcal{D}_t$
  end
end
Output: $\mathcal{D}$ set of locations of deployed MAPs

[1]**Note**: Here, we adopt the MAX-SNR algorithm to limit complexity.

Figure 28: SIMBA MAP deployment algorithm

4.2.2.2 User association

In this section, we describe the proposed MARL algorithm for user association. In the proposed framework, we model each UE as an agent, which cooperatively learns with its teammates a common user association policy through interaction with the shared radio environment. To this end, agents learn to map their local and global observations $o_j(t)$ of the radio environment to actions $a_j(t)$ corresponding to connection requests towards MAPs. Following our previous work [7], let $o_j^l(t) = \{R_\alpha(t), RSS_j(t), AoA_j(t), R_{a_j(t),j}(t), D_j(t)\}$ denote the local observation, which comprises the received signal strength , RSS_j, and the associated angle of arrival AoA_j w.r.t to all MAPs. Here, $R_{a_j(t),j}$ represents UE *j*'s perceived rate and D_j the UE j data demand. Moreover let $o_j^g(t) = \Big\{(x_k(t), y_k(t), z_k(t), a_{k,}(\mathrm{t}-1)\Big), \forall \mathrm{k} \in \mathcal{N}_j(\mathrm{t})\}$ denote the global observations of UE *j*, where $(x_k(t), y_k(t), z_k(t))$ is the location of its *k*-th neighbours, $a_{k,}(\mathrm{t}-1)$ is the connection request in previous time slot, and $\mathcal{N}_j(\mathrm{t})$ denotes the UE neighbourhood. These global observations represent UE *j* perception of its surrounding environment. The goal of the learning procedure is to define the user association policy π_w, with learnable weights w which outputs the association probability vector $p_j(t) = \pi\left(o_j(t)\right) \in \mathbb{R}^{|\mathcal{A}|}$ that maximizes the sum of γ-discounted rewards over a time horizon T_e:

$$G_j(t) = \sum_{\tau=t+1}^{T_e} \gamma^{\tau-t-1} R_\alpha(\tau)$$

where γ is the discount factor such that $0 < \gamma < 1$. Finally, we construct the policy π_w using an actor-critic module, which is optimized via proximal policy optimization [7]. In particular, our proposed solution is specifically conceived to handle dynamic networks with varying number and position of UEs.

4.2.3 Results

Here, we assess the performance of our proposed MAP deployment method on dynamic scenarios at different scales. The first scenario, named **SmallScale**, is a small scale deployment of 10 UEs randomly and 4 MAPs moving through 12 positions at 3 different altitudes (i.e. 15, 35 and 50m) in a 100m by 100m area. In this scenario, we can easily compare our method with brute force mechanism, named Exhaustive, without too high computational time. The **MediumScale** scenario is made with 40 UEs and 10 MAPs moving through 27 positions at 3 altitudes in a 200m by 200m area. The UEs are deployed with uniform probability and move randomly over the cell. The UEs' traffic follows a Poisson distribution bit/s and simulates dynamic rate demand. Both scenarios include a baseline random algorithm, named Random, where the deployment decision and the chosen location follow a uniform law $\mathcal{U}(0,1)$. Thus each UAV chooses a random location and when a better combination is found, the solution is updated.

Concerning MAPs, each UAV has a coverage range defined by the aperture angle of its antenna to ground UEs and its altitude. At most K=10 UAVs can be deployed and each UAV can connect to at most $N_i = 10$ UEs. We fix the rent cost $c_{i,0} = c_0 = 1, \forall i,$ so that we can omit it from the optimization in cost equation. Table 1 gives an overview of the simulation parameters.

Table 1: Simulation parameters

Scenario Parameters	Small Scale	Medium Scale
Cell size	100×100 m	200×200 m
Number of UEs	10	40
Number of positions	$4 \times 3 = 12$	$9 \times 3 = 27$
UE Mobility	Random Walk	
Avg. traffic demand $D_j(t)$	200 Mbps	
Channel Parameters	**MBS**	**BS/UAV**
Carrier Frequency f_c	2 GHz	28 GHz
Bandwidth	10 MHz	500 MHz
Thermal Noise N_0	-174 dBm/ Hz	
Shadowing power σ^2	9 dB	12 dB
Transmit Power	46 dBm	20 dBm
Antenna Gain	17 dBi	Directive [19]
UAV Aperture	120 deg	
Altitude	$[10, 35, 50]$ m	

4.2.3.1 Drone deployment

We set M=10 and T=100 in SIMBA and average the results over 30 Monte-Carlo simulations. We conveniently fix the QoS's target of $(\mathcal{C}_5)$ to $Q_j = 100\%$ for **SmallScale** and $Q_j = 85\%$ for **MediumScale** due to limited radio resources. We first assess the deployment cost of our proposed solution compared to the two benchmarks as shown in Figure 29. Our algorithm converges faster than a naive random approach to a close optimal-solution. In a high mobility context, it is important to obtain a flexible algorithm that finds a solution faster than the network changes.

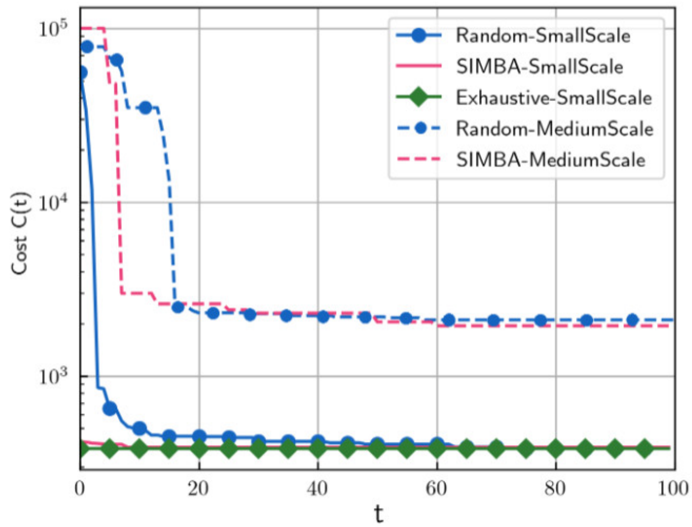


Figure 29: MAP deployment cost for each method and for both scenarios as a function of the number of iterations

Moreover, our algorithm guarantees not to deploy more drones than needed, which may imply a high cost. Meanwhile, as shown in Figure 30.a, our solution ensures and guarantees the targeted QoS for UEs in both small and medium scale scenarios with less MAPs compared to a naive approach as illustrate Figure 30.b. Finally, as we added more potential locations for UAVs in **MediumScale** scenario, our solution is better at identifying the best combinations than a random naive approach and faster than an exhaustive solution, thus, guaranteeing a near-optimal solution with 4.22 UAVs deployed in average.

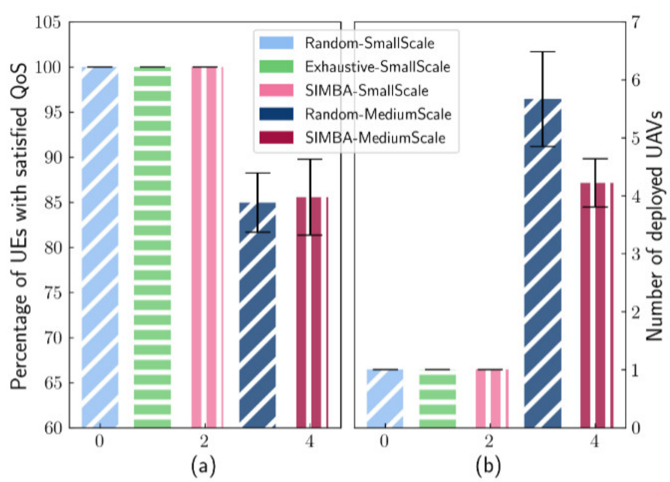


Figure 30: (a) Average percentage of UEs with satisfied QoS and (b) average number of UAVs deployed for both scenarios and for each method.

4.2.3.2 User association

To train the user association policy, we use the MARL framework described in [7]. All simulation results are plotted for a learning rate $\mu = 10^{-4}$ and a discount factor γ=0.6. The MARL agents are trained for $T_e = 3000$ episodes. Agents are trained with $\alpha = 1$ for the α-fair utility function, meaning that the agents are trained on a fair sum-rate setting. Note that all the hyperparameters were determined empirically.

We perform several deployments and trainings with increasing number of MAPs and compare the proposed solution to the MAX-SNR-based approach for **MediumScale** scenario. Here, results are averaged over 300 simulations with T=150 iterations.

Figure 31 shows the impact of increasing the number of UAVs on the network performance. We observe that, for a number of deployed UAVs greater than tree, our proposed approach increases the log network sum-rate by 1.5%, implying a network sum-rate enhancement by nearly 30% compared to MAX-SNR algorithm.

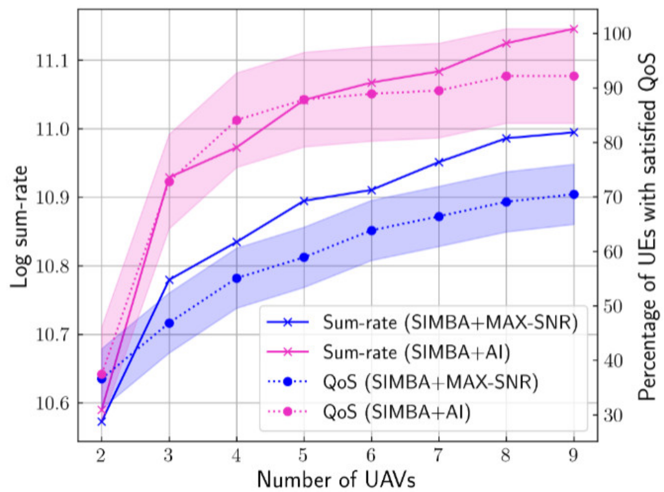


Figure 31: Average network log sum-rate and average percentage of UEs with QoS satisfaction as a function of the number of deployed UAVs for MediumScale scenario for both UE association algorithms

Moreover, for the given scenario, Figure 31 illustrates the trade-off between drone deployment and UE QoS. With more than 4 UAVs deployed, the log sum-rate barely varies. This result confirms that when interference is taken into account, increasing the number of access points does not necessarily imply better UE's QoS at the risk of increasing the deployment cost.

Next, in the **MediumScale** scenario, we increase the user density $\lambda(UEs/m^2)$ to show the effectiveness of our solution in this complex setting. Figure 32 compares the average handover frequency and the network log sum-rate as a function of user density λ. The increase of λ ultimately increases the number of handovers frequency for the MAX-SNR algorithm as multiple UEs compete for the same resources. In contrast, our proposed solution guarantees stable performance due to its capability to balance the network load, especially in dense deployment scenarios.

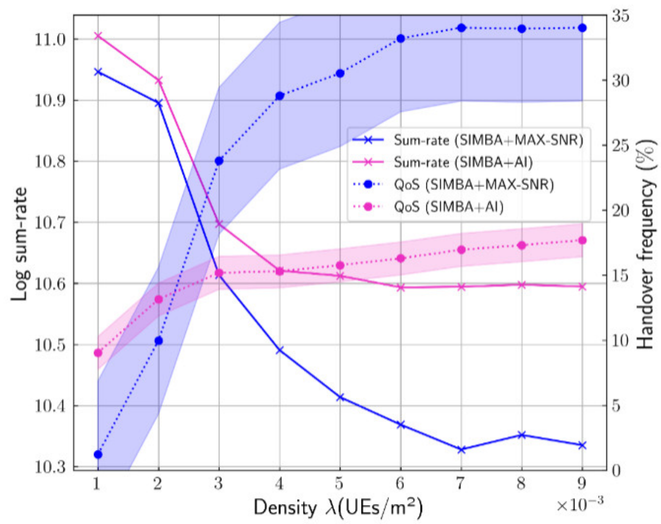


Figure 32: Average network log sum-rate and average handover frequency as a function of the user density

As shown in Figure 32, our proposed solution improves the log network sum-rate by 4%, which implies an increase in network sum-rate by 60% compared to a MAX-SNR algorithm, in particular for dense deployment scenario (e.g. $\lambda = 9 \times 10^{-3}\ UEs/m^2$).

4.2.4 Conclusions

In this section, we proposed an algorithm to solve the joint problem of MAP deployment and user association while taking into account UE mobility, UE demand and network interference. The algorithm finds the MAP positions in a 3D space and optimizes the user association for a given network configuration. The proposed algorithm is the first step for more complex scenarios. We will consider path planning optimization to exploit MAP connectivity while moving within the network, include the backhaul constraint and optimization and extend our study to more realistic configurations with aerial and ground base stations.

4.3 QoS aware and energy efficient user association in a MAP-assisted network

In this subsection, we study the QoS-aware and energy efficient user association. When mobile APs (e.g. Drones) are deployed to expand coverage or extend connectivity, the efficient energy utilization becomes important. While mobile Aps are operated with the embedded batteries, the battery power would be used for mobile APs' operation including the movement control and the signal transmission for communication to the ground BS (gBS) and to UEs. In addition, considering that the mobile APs are connected to gBS, the backhaul capacity is also imposed as a constraint in a MAP-assisted network. When multiple UEs are distributed, the services requiring the different level of QoS would be provided to UEs. For UEs capable of multi-connectivity, UEs can be connected to multiple BSs (gBS and/or MAPs). In this study, we focus on the capacity expansion perspective by using MAPs and the multi-connectivity features. Then, considering imposed constraints, the efficient user association mechanism is proposed.

4.3.1 System model and Problem Formulation

Our goal of user association is to maximize the QoS satisfaction level of UEs and the energy efficiency while considering constraints of the backhaul capacity and fixed transmission power. As Figure 33 is illustrated, there is a ground BS (gBS) and mobile APs (e.g. drone) can be deployed to support the connectivity to more UEs. While mobile APs are connected to the gBS with the wireless backhaul, MAPs are assumed to serve UEs, thus MAPs are considered as aerial BSs (aBSs) from the UE perspective.

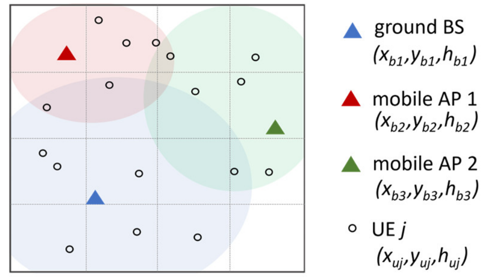


Figure 33: The system model

The channel between the gBS and UEs is modelled by the macro cell propagation channel model in 3GPP TR 36.942. Specifically, considering a typical gBS antenna height of 15 meters above the rooftop level, the path loss between the gBS m and UE n is formulated as

$$\emptyset_{mn} = 37.6 \log_{10}\left(\frac{d_{mn}}{1000}\right) + 21 \log_{10}(f_c) + 58.8 + PL_{sd} \tag{1}$$

where d_{mn} is the 3D distance between gBS and UE n. PL_{sd} is a random variable conforming to a Gaussian distribution of $N(0, \sigma_0^2)$ due to the shadow fading.

In the air-to-ground channel model between aBSs and UEs [36], there exist two propagation groups, Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) connections. The channels between
UEs and aBSs are modelled by randomly selecting a group of propagation model from two propagation groups, LoS and NLoS. The probability of having LoS connectivity between UE n and aBS m is affected by the locations of the aBS and the UE and environment. It can be calculated as

$$\rho_{mn} = \frac{1}{1 + \alpha \exp\left(-\beta\left(\frac{180}{\pi} \arctan\left(\frac{h_{mn}}{r_{mn}}\right) - \alpha\right)\right)} \tag{2}$$

where α and β are the environmental parameters indicating the characteristics of environment where the user is located (e.g., rural, urban, etc.). h_{mn} is height distance and r_{mn} is the horizontal distance between aBS m and UE n. The average pathloss in dB between aBS m and UE n can be obtained as

$$\emptyset_{mn} = 20 \log_{10}\left(\frac{4\pi f_c d_{mn}}{c}\right) + \zeta_{mn}. \tag{3}$$

The first term indicates the free space pathloss where f_c is the carrier frequency and the second term ζ_{mn} is the average additional pathloss of LoS and NLoS connections and can be calculated as follows.

$$\zeta_{mn} = \rho_{mn}\zeta_{mn}^{los} + (1 - \rho_{mn})\zeta_{mn}^{nlos}. \tag{4}$$

ζ_{mn}^{los} and ζ_{mn}^{nlos} denote the associated excessive pathloss in dB with probabilities ρ_{mn} and $1 - \rho_{mn}$, respectively.

Considering the different level of QoS, the utility function is introduced to incorporate the data rate requirements. The utility function of UE n wanting the data rate R_n^{req} is represented by normalised sigmoidal-like function as follows [35].

$$\Phi_n(R_n) = \log\left(\frac{q_n}{1+e^{-c_n(R_n-e_n)}} - q_n k_n\right) \ , \qquad \text{for n} \leq \text{N} \tag{5}$$

where $c_n > 0$, $e_n = R_n^{req}$, $q_n = (1+e^{c_n e_n})/e^{c_n e_n}$ and $k_n = 1/e^{c_n e_n}$. Let us form two vectors as follows.

$$\boldsymbol{\alpha_n} = [\alpha_{0n}, \alpha_{1n}, \cdots, \alpha_{mn}, \cdots, \alpha_{Mn}]^T, \quad \boldsymbol{R_n} = [\mathrm{R}_{0n}, \mathrm{R}_{1n}, \cdots, \mathrm{R}_{mn}, \cdots, \mathrm{R}_{Mn}]^T, \ \text{ for n} \leq \text{N}, \tag{6}$$

where $\boldsymbol{\alpha_n}$ is the binary user association variables vector and $\boldsymbol{R_n}$ is the vectorized data rate. Then, the utility of UE n can be calculated as

$$\mathrm{U}_n(\boldsymbol{\alpha_n}) \triangleq \Phi_n(\mathbf{R}_{\boldsymbol{n}}{}^T \cdot \boldsymbol{\alpha_n}), \ \text{ for n} \leq \text{N}, \tag{7}$$

where $\mathbf{R}_{\boldsymbol{n}}{}^T \cdot \boldsymbol{\alpha_n}$ calculate UE n's aggregated data rate over connected BS(s). We can calculate the overall utility value $\varphi_{ut} = \Sigma_n \log(\mathrm{U}_n(\boldsymbol{\alpha_n}))$.

Then the problem to associate users to multiple BSs, P can be formulated by using an optimisation approach as follows.

$$P: \quad \max_{\alpha_{mn}, \mathrm{p}_{mn}} \{\varphi_{ut}, \eta_{pe}\} \tag{8}$$

$$\text{s.t.} \quad \sum_{n \leq N} \alpha_{mn} p_{mn} \leq P_m^{max}, \qquad \forall\, m \leq M, \tag{9}$$

$$\sum_{n \leq N} \alpha_{mn} R_{mn} \leq C_m, \qquad \forall\, m \leq M, \tag{10}$$

$$\sum_{m \leq M} \alpha_{mn} \leq q_n, \qquad \forall\, n \leq N, \tag{11}$$

$$\sum_{n \leq N} \alpha_{mn} \leq q_m, \qquad \forall\, m \leq M, \tag{12}$$

$$\alpha_{mn} = \{0,1\}, \qquad \forall\, m \leq M, \forall\, n \leq N, \tag{13}$$

$$\mathrm{p}_{mn} \geq 0, \qquad \forall\, m \leq M, \forall\, n \leq N. \tag{14}$$

In this formulation, α_{mn} is the binary user association variable to indicate whether UE n is associated with BS m and p_{mn} denotes the transmit power from BS m to UE n. The constraints (9) and (10) indicate the transmit power limit and the backhaul capacity limit, respectively. UE n can be associated with multiple BSs (at most q_n BSs) simultaneously depending on its multi-connectivity capability in (11) while BS can associate multiple UEs according to (12). In the objective, φ_{ut} indicates the utility value reflecting the satisfied QoS level and η_{pe} denotes the power efficiency. While the total achieved rate of all users and the total transmit power of all BSs can be expressed as follows, respectively.

$$R_\Sigma = \sum_{m=0}^{M}\sum_{n=1}^{N} \alpha_{mn} R_{mn}, \quad P_\Sigma = \sum_{m=0}^{M}\sum_{n=1}^{N} \alpha_{mn} p_{mn}, \tag{15}$$

Then, the overall power efficiency is calculated by $\eta_{pe} = R_\Sigma / P_\Sigma$.

4.3.2 Proposed algorithm

The formulated problem P is a complicated mixed integer nonlinear programming (MINLP) type, non-convex and NP-hard. The complexity can increase exponentially depending on the scale of networks. Therefore, we propose a sub-optimal algorithm adopting the matching game theory that can provide a low complexity distributed solution.

In this study, the matching is defined as the assignment of UEs in N to BSs in M. The sets of BSs M and UEs N are modelled as two teams player. Each player in M and N tends to match the most preferred partners in the opposite team while behaving independently, thus this matching based approach can be worked in a distributed manner. The matching π is expressed to mapping players $n \in N$ and $m \in M$ under the following conditions of the matching $\pi: M \cup N \Rightarrow 2^{M \cup N}$ such that,

1) $\pi(m)$Í N such that $|\pi(m)| \leq q_m$, $\forall m, \forall n$,
2) $\pi(m)$Í M such that $|\pi(n)| \leq q_n$, $\forall m, \forall n$,
3) $n \in \pi(m)$ if and only if $m \in \pi(n)$, $\forall m, \forall n$.

$\pi(m)$ indicates the set of matched partners of player m under the matching relation π. $|\pi(\cdot)|$ is the cardinality of $\pi(\cdot)$ and q_m and q_n denote the maximum capacity of player m and n. Depending on the capacity of BSs and the multi-capability capability of UEs, a value of q_m and q_n can be decided. m and n manage each accepted list A_m and A_n, respectively. When n becomes to map to m, m adds n to A_m and n adds m to A_n.

Whilst the preference relationship is expressed with $\succ$, the expression $m \succ_n m'$ indicates the UE n prefers BS m over m' (where $m \neq m'$). A same notation is used for BSs for their preference. While the preference level over candidate partners is managed with the preference list l_m and l_n.

From the BS side, for the better resource use, the reference signal received power (RSRP) level based on measurement report feedback by UEs is assumed to be considered for its preference decision. When RSRP at UE n is larger than RSRP at UE n', the rank or preference $\succ_m$ in the preference list l_m is decided, given as,

$$n \succ_m n' \Leftrightarrow \mathrm{RSRP}_n \succ_m \mathrm{RSRP}_n. \quad (16)$$

From the UE side, for the first connection setup, UEs prefer to matching to the BS of the best channel quality (like the BS side) to set up the most robust connection. From the second favourable BSs, UEs consider the network load of BSs with the number of attached UEs. That is, for BS m, let the number of associated UEs to BS m denote N_m. Then, the preference relation $\succ_n$ of UE n is defined as,

$$m \succ_n m' \Leftrightarrow \mathrm{N}_m \succ_n. \quad (17)$$

Considering the quote of UEs and BSs with constraints in backhaul capacity, the UEs will play the matching process to find proper connection setups. Although UEs become to be associated with multiple BSs, the required QoS would not be satisfied. For this case, power allocation adjustment is followed by the user association procedure. During the user association matching procedure, the equal power allocation is assumed, then the power allocation adjustment process will be activated only for UEs which require more data rates. The details of the proposed algorithm are illustrated in the following table.

Table 2: Pseudo code of the proposed user association algorithm

Algorithm 1. Matching based User Association algorithm	
1	**Input:** The preference list l_m & l_n where $m \in M$ & $n \in N$, $C_m = backhaul\ capacity\ of\ BS\ m$
2	**Initialise:** $A_m = \emptyset, A_n = \emptyset, J_n = \emptyset,\ t \leftarrow 0, p_{mn} = P_m/q_m$
3	**Repeat:**
4	$t \leftarrow t + 1$
5	Each n proposes the most preferred *m,* where $m \in l_n^{(t)}$.
6	The indices of selected *m* are deleted from $l_n^{(t)}$.
7	**while** $l_n^{(t)} \neq \emptyset$ and $\exists m \in l_n^{(t)}$ **do**
8	Calculate the UE's utility $U_n^{(t)}$ by using the established connections $A_n^{(t)}$
9	**if** $U_n^{(t)} < 1$ and $\lvert A_n^{(t)} \rvert < q_n$ and C_m
10	$m \leftarrow l_n^{(t)}(1)$
11	**if** $\lvert A_m^{(t)} \rvert < q_m$ and $\exists n \in l_m^{(t)}$ and $C_m^{(t)}$
12	$\pi(n)^{(t)} \leftarrow \pi(n)^{(t)} \cup m\ , A_n^{(t)} \leftarrow A_n^{(t)} \cup m,\ A_m^{(t)} \leftarrow A_m^{(t)} \cup n, C_m^{(t)} = C_m^{(t)} - R_{mn}$
13	update $l_n^{(t)}, l_m^{(t)}$ ($l_n^{(t)} \leftarrow l_n^{(t)} \backslash m\ ,\ l_m^{(t)} \leftarrow l_m^{(t)} \backslash n$)
14	**else if** $\lvert A_m^{(t)} \rvert = q_m$ and $\exists n \in l_m^{(t)}$
15	**for** $n' \in A_m^{(t)}$
16	**if** $(n, \pi_m^{(t)}) \succ_m (n', \pi_m^{(t)})$
17	$J_{n'}^{(t)} \leftarrow J_{n'}^{(t)} \cup m\ , \pi(m)^{(t)} \leftarrow \pi(m)^{(t)} \cup n,\ C_m^{(t)} = C_m^{(t)} + R_{mn'} - R_{mn}$
18	$A_m^{(t)} \leftarrow A_m^{(t)} \backslash n',\ A_{n'}^{(t)} \leftarrow A_{n'}^{(t)} \backslash m,\ A_n^{(t)} \leftarrow A_n^{(t)} \cup m,\ A_m^{(t)} \leftarrow A_m^{(t)} \cup n$
19	update $l_n^{(t)},\ l_m^{(t)}$
20	**else**
21	$J_n^{(t)} \leftarrow J_n^{(t)} \cup m,$ update $l_n^{(t)},\ l_m^{(t)}$
22	**else**
23	$J_n^{(t)} \leftarrow J_n^{(t)} \cup m,$ update $l_n^{(t)},\ l_m^{(t)}$
24	**until** $\pi_m^{(t)} \neq \pi_m^{(t-1)}$
25	**procedure** Utility Fitting
26	$\bar{N}^{(t)} \leftarrow$ sort (n, ascending) for $U_n^{(t)} < 1$
27	**while** $\bar{N}^{(t)} \neq \emptyset$ and $\exists n \in \bar{N}^{(t)}$ **do**
28	select *m* where $m \in A_n^{(t)}$ and $\lvert A_m^{(t)} \rvert < \lvert A_{m'}^{(t)} \rvert$
29	**if** $C_m^{(t)}$ and $\bar{P}_m^{(t)} < P_m$
30	$p_{mn}^{(t)} \leftarrow p_{mn}^{(t)} + \min(p_{mn}, \overline{p_{mn}})$
31	$\bar{P}_m^{(t)} \leftarrow \bar{P}_m^{(t)} + p_{mn}^{(t)}$
32	Update $U_n^{(t)}, \bar{N}^{(t)}, C_m^{(t)}$
33	**end** procedure

In the proposed algorithm, the preference lists l_m and l_n of both set of players, BS *m* and UE *n*, are assumed to be provided. The list of accepted UEs for BS *m*, A_m, the list of accepted BSs for UE *n*, A_n and the list of rejected BSs for UE *n*, J_n are initialized. The proposed matching operates by applying the following procedure successively.

At iteration *t*, each UE *n* makes a proposal to the most preferred BS *m* included in its preference list l_n but not included in the rejection list J_n. After UE's proposal, the selected index *m* is deleted from UE's preference list l_n. While BS *m* receives UE *n*'s proposal, BS *m* considers the proposal only from *n* included in its preference list l_m,.

When BS *m* has the capability to accommodate a new UE (i.e., considering its quota q_m and the backhaul capacity C_m), BS *m* holds the proposal from *n*. (lines 11-13). In case BS *m* *has* no quota to accommodate any new proposal, BS m checks the accepted list A_m and finds any least favourable matching partner n'. If there exists a least favourable matching $\pi_{mn'}$, BS m rejects n' and removes the existing matching $\pi_{mn'}$ from A_m. Then, the new matching π_{mn} is arranged. In the case that any matching instances are created or rejected, the corresponding preference lists, rejection lists, and residual capacity of BS m are updated accordingly.

After terminating this iterative matching process, there could be UEs whose required QoS level are not satisfied. To complement those UEs, the utility fitting procedure to adjust the power allocation is followed. Considering the achieved utility level, UEs are sorted, and from UEs of the lowest utility value, the power adjustment procedure is started. UE n negotiates with the connected BSs to increase the transmit power. In this step, the BS m of the least traffic load (i.e., connected to the least number of UEs) are firstly contacted since having the least traffic load indicates the more power budget from BS m. Once BS m receives the power negotiation request from UE n, it calculates the allowed max power $\overline{p_{mn}}$ derived by its capacity C_m and total power budget P_m and the power adjustment is decided.

Let us consider the complexity of the proposed algorithm. Preference lists of BSs and UEs are assumed to be provided as inputs. For each UE and BS, the complexity of creating their preference list using standard sorting algorithm is $O(M \log M)$ and $O(N \log N)$ [37]. Therefore, the input length is calculated as $\sum_{m \le M} |l_m| + \sum_{n \le N} |l_n| = 2NM$ where $|l_m|$ and $|l_n|$ are the length of the preference lists of UE m and BS n. The matching procedure will terminate after a finite number of repetitions, and for the worst case, the complexity becomes linear with the size of the preference lists $O(MN)$. In the utility fitting procedure, the time complexity is calculated as $O(MN \log MN)$.

4.3.3 Results

We consider cases of different numbers of UEs. While the different UEs want the different QoS level modelled by the required data rate, the UE's required rate is designed to be chosen from a set $R_{req} = \{10, 5, 3\}$[Mbps]. The required rate of each UE is also chosen randomly with the location at the beginning of simulation. To compare the performance of the proposed scheme, the greedy algorithm is considered as the reference scheme. In the reference algorithm, UEs tend to act in a greedy manner, i.e. UEs try to connect to all available BSs by using the multi-connectivity capability (i.e., all BSs which constraints in the backhaul capacity and minimum SINR level are satisfied).

For the different numbers of UEs N, three cases $N = \{10,15,25\}$ are selected. Figure 34 illustrates the achieved average data rate performance of the proposed scheme (labelled 'Pro') and the reference greedy scheme (labelled 'Ref'). In Figure 35, the average numbers of connected BSs from UEs are shown. That is, by associating with multiple BSs shown in Figure 35, UEs could achieve the data rates illustrated in Figure 34. For each case (different *N* and different scheme), UEs' performance is grouped as three categories depending on required rate levels.

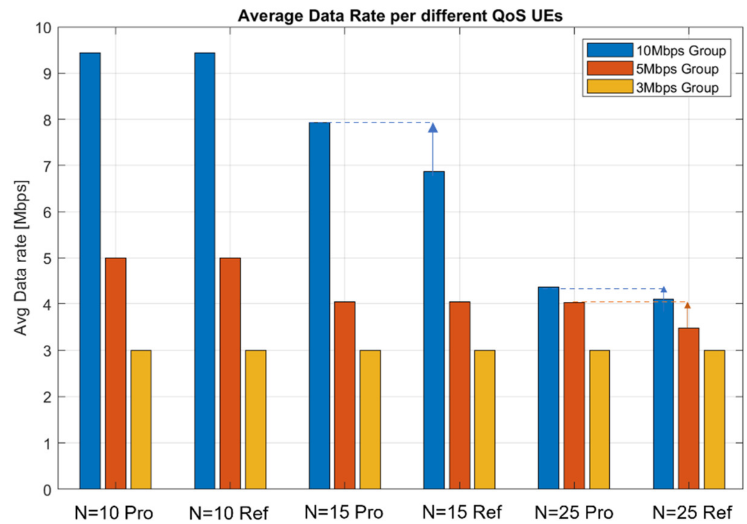


Figure 34: Comparison of the average data rate performance

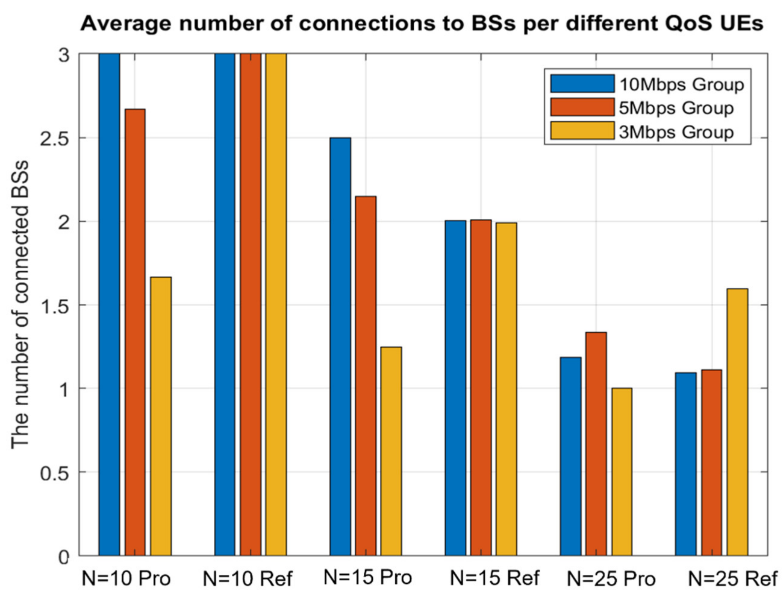


Figure 35: Comparison of the number of connections to BSs from different UEs

For cases of a small number of UEs ($N < 10$), UEs could be associated with multiple BSs and their requirements could be satisfied relatively easily since there is no serious competition between UEs. For the case of 10 UEs ($N = 10$), while UEs requiring 5 and 3 Mbps could be satisfied successfully by connecting to multiple BSs, UEs requiring higher data rate (10Mbps) achieve could not achieve 10 Mbps rate. Regarding the multi-connectivity capability, the limit of maximum number of links are introduced to promote the fair resource utilization between UEs. In this simulation, the maximum number of links connected to BSs from one UE is set to 3. For the case $N = 10$ of the proposed scheme ('N=10 Pro'), UEs of 10 Mbps group could receive data with the rate close to 10 Mbps by associating to three BSs in average. For UEs of 5Mbps group, they tried to connect to around 2.6 BSs in average and the required rate 5 Mbps could be satisfied. Even UEs of 3 Mbps group are connected to more than 1 BS for the required QoS. However, in the reference scheme case ('N=10 Ref'), all UEs access to 3 BSs in the greedy manner. Even the group of UEs requiring 3Mbps is connected to 3 BSs. It leads to waste of radio resource. While UEs wants links for 3Mbps rate, the connection setup is established for much higher rate, thus the radio resource is used inefficiently. The effect of resource waste is shown clearly when the traffic load increases. For the case of N=15, in the reference scheme, the UEs in 3Mbps group are connected to 2 BSs, which is the same number which UEs in 10 Mbps group are connected. It can be derived that this greedy manner leads to lower satisfaction level of UEs in 10 Mbps group. In the proposed algorithm case, by associating UEs to BSs considering their service requirements, it tends to support higher data rate requirements (10 Mbps Group) better than the reference scheme. When the node number increases further, UEs become more competitive for radio resources and UE's satisfied QoS level becomes decreased. Nevertheless, it is shown that the proposed algorithm generates better QoS level for N=25 shown in Figure 34. Due to the complexity, the comparison with the optimal algorithm (implemented by exhaustive search) could not been carried out for all cases of different numbers of UEs. However, for the case of the smaller number of UEs ($N = 5$), the proposed algorithm is compared with the optimal algorithm and it is shown that the proposed algorithm produces the optimal solution found by the optimal algorithm.

Table 3: Comparison of power efficiency for cases of different UEs

No. of UEs	N=10	N=15	N=25
Proposal	1.2009	1.2736	1.5179
Reference	1.0018	1.1428	1.3588

Table 3 shows the power efficiency performance of two algorithms. For all cases, by associating to BSs considering the traffic load of BSs and the required QoS level of UEs, the proposed algorithm produces the better power efficiency compared to the reference greedy method.

4.3.4 Conclusion and future work

In this subsection we presented the QoS-aware and energy-efficient user association algorithm in a MAP-assisted network. When mobile APs are available, UEs are capable of connecting to multiple APs, then even UEs requiring higher data rates can be satisfied its required QoS level. In the proposed algorithm, which is a distributed mechanism based on matching theory, UEs tries to associate to proper BSs considering their QoS level and the network traffic loads under the given constraints in backhaul capacity and total transmit power. As a two-step approach, the power adjustment procedure is followed by user association and contribute to improving UE's QoS level.

4.4 Bi-objective trade-off analysis for joint user association and MAP placement method

In wireless communications involving mobile access points, a number of joint optimization problems have been investigated in the past [22]. The problems can be classified based on the number and type of subobjectives, parameters involved in the target optimization function, the centralization degree of the control architecture, and whether the search space is discrete or continuous.

In this section, we focus on the joint user association and MAP placement problem. In particular, we deploy a bi-objective optimization problem which address the trade-off between the network cost minimization and user utility maximization in a single framework.

Network cost minimization for the target minimum user utility is important especially when the network operator has a limited number of MAPs to be deployed to support a number of temporary events distributed to a larger area. On the other hand, some of the events may involve user traffic, such as file transfers, which may benefit from maximizing the user utility also from the network operator perspective, so that the resources are released as soon as possible for the next user requesting service.

In the following, the target constraint bi-objective problem and used solution framework are first briefly described. We focus on a centralized control approach which is more suitable for a network cost minimization subobjective. Then the simulation system and results are de-scribed. The simulation results are produced to illustrate the potential effect of the bi-objective approach on improving the performance as function of different optimization constraints that rise from the user utility targets, traffic requirements, as well as network infrastructure constraints.

4.4.1 Problem

We consider *N* users deployed in a network which involves *K* fixed access points (FAPs) that are further assisted by a swarm of aerial MAPs which can be placed in *L* candidate discrete 3D locations and *M* = *K* + *L*. The AP placement binary decision variable of the *m*th location is denoted as y_m which is 1 if the AP is placed and 0, otherwise. Then, user *n* is associated with an AP at location *m* if user association variable x_{nm} is 1, and 0, otherwise.

Let $y_{\max}$, C_m, $U_{\min}$, $d_{\max}$, d_m be the maximum number of available MAPs, maximum capacity of the *m*th AP, minimum required user utility, maximum allowed backhaul distance, and distance of the *m*th AP to its closest AP, respectively. Let U_{nm} be the user utility of the *n*th user associated to *m*th AP which is selected to be the spectral efficiency in bit/s/Hz using the Shannon capacity law. The bi-objective constrained optimization problem at hand is given by

$$\min_{y_m, x_{nm}} \left\{ \overbrace{\left(\sum_{m=1}^{M} y_m\right)}^{AP\ placement}, \overbrace{\left(-\sum_{m=1}^{M}\sum_{n=1}^{N} U_{nm} x_{nm}\right)}^{User\ association} \right\}$$

s.t.

$x_{nm}, y_m \in \{0,1\}$ (binary decision variables)

$x_{nm} \leq y_m, \forall n, m$ (association vs. placement)

$\sum_{m=1}^{M} x_{nm} \leq 1, \forall n$ (association to max one AP)

$\sum_{m=1}^{M}\sum_{n=1}^{N} x_{nm} \geq \beta N$ (network coverage)

$y_m = 1, m \leq K$ (fixed AP locations)

$\sum_{n=1}^{N} x_{nm} \leq C_m, \forall m$ (max AP capacity)

$\sum_{m=K+1}^{M} y_m \leq y_{\max}$ (max number of MAPs)

$U_{nm} \geq x_{nm} U_{\min}, \forall n, m$ (utility constraint)

$y_m d_m \leq d_{max}, m > K$ (backhaul constraint)

The above bi-objective approach aims in general at balancing between minimizing the number of access points and maximizing the network utility via joint placement and association decision. The applied constraints are explained in the respective parentheses.

4.4.2 Applied solution framework

In order to resolve the problem, we first apply the scalarization approach to transfer the bi-objective optimization problem into a single-objective optimization problem. That is, the above objective function is transferred into

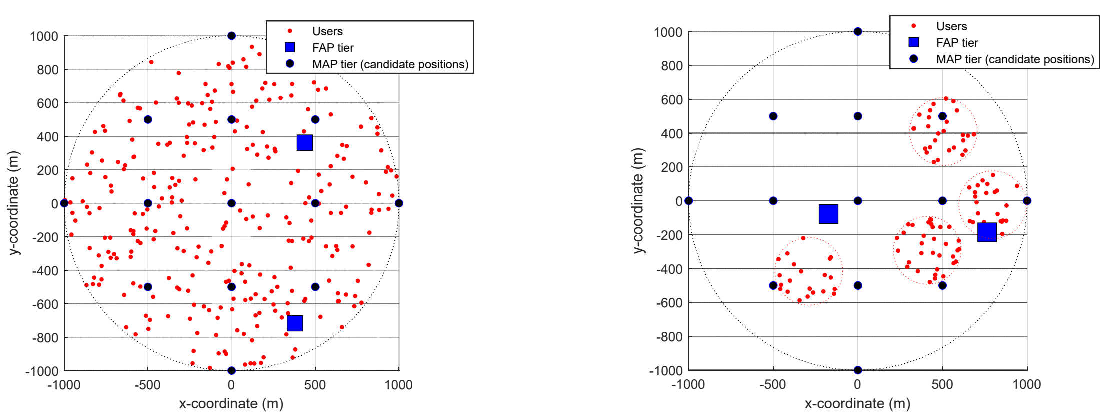
$$\min_{y_m, x_{nm}} \left\{ w_1 \sum\nolimits_{m=1}^{M} y_m - w_2 \sum\nolimits_{m=1}^{M} \sum\nolimits_{n=1}^{N} U_{nm} x_{nm} \right\}$$

where (w_1, w_2) are relative weights between the cost minimization and the sum user utility maximization. The constraints remain the same as in the original problem. The above objective is a two-parameter integer linear programming problem. It can be solved using the branch and bound search algorithm which is based on recursively splitting the problem into subproblems and using upper and lower bounds of the optimal solution (cf. [23]).

The weight selection for the bi-objective approach can be used to select the preference of the network operator. In general, it can force to simultaneously use minimum MAPs and then the best user utility. In the next subsection, we are mainly interested in the following three special cases where we aim at selecting the weights to either i) to minimize the number of MAPs only, ii) to maximize user utility only, or iii) to first minimize the number of MAPs and then maximize the user utility using the minimum number of MAPs from the first phase.

4.4.3 Results

The main aim of the simulation study is to reveal the different trade-offs and relative performance between the aforementioned weighting strategies of the bi-objective function. We consider the downlink transmission of randomly deployed users in a circular area with radius of 1000 m. We apply a terrestrial FAP tier with two randomly deployed FAPs. The aerial MAP tier at altitude of 30 m is deployed with 13 candidate positions where the maximum number of MAPs is 10. For the path loss model, we assume the ITU-R UMi model [25] and urban channel model [24] for terrestrial and aerial links, respectively. The spatial traffic is modelled as a Poisson traffic as well as clustered Poisson model where the clusters represent separate subevents and can be controlled by specifying the average number of subevents, subevent density, and subevent radius. Matlab toolboxes are utilized where possible. Illustrations of the snapshots of the traffic models are shown in Figure 36.



a) Poisson traffic b) Clustered Poisson traffic

Figure 36: Illustration of snapshots of spatial traffic models.

We first wish to examine the relationship between the coverage probability, normalized user utility, and AP deployment ratio for the three different objective weighting scenarios (minimum cost, maximum user utility, bi-objective case) introduced in the previous section using the Poisson spatial traffic model. Here the utility is normalized by the achievable utility without any constraints presented in the previous subsection restricting the association. The AP is assumed to be deployed if it is both placed in a candidate position and it has at least one user associated. The connectivity of the FAP tier is in this case not satisfying well the user utility constraints due to channel attenuation of the terrestrial links to illustrate a situation how the aerial MAP tier can be helpful.

It is seen from Figure 37 that the coverage probability of all three optimization cases is similar due to the same user utility constraint on the x-axis. However, there are significant differences in Figure 38 and Figure 39 where the normalized network utility and deployment ratio are presented, respectively. For the maximum utility case, the network utility increases as the minimum spectrum efficiency constraint to connect users is reduced. However, for the minimum cost case and bi-objective cases, the curves are non-monotonic. This is because the ability of the network to support very high spectral efficiencies start to diminish the network utility which becomes a dominating feature for all approaches. Moreover, the utilization of less MAPs will reduce the achievable network utility for other approaches than the one that only maximizes the network utility. It is further seen that the bi-objective case is able to provide better utility than the case which only maximizing the utility while it has similar performance with the maximum utility case, for certain range of spectrum efficiency constraint. Finally, in Figure 40, we deploy the clustered traffic model and illustrate how the density of fixed and mobile network infrastructure affects the coverage probability as a function of maximum allowed backhaul distance between a deployed MAP and the closest FAP when using the bi-objective approach. Clearly, also for the clustered model, the coverage performance can be improved by increasing either the FAP or MAP density.

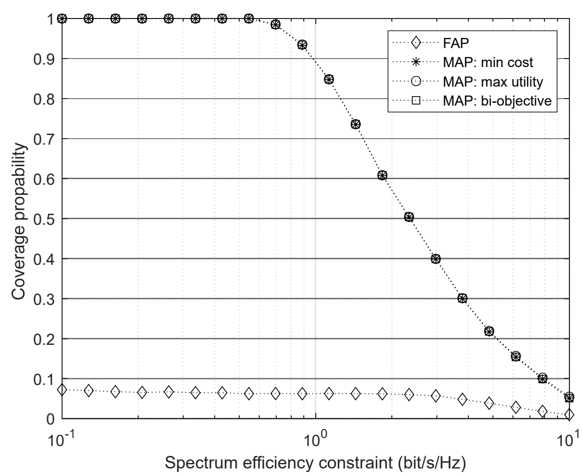


Figure 37: Simulated coverage probabilities versus minimum constraint for spectral efficiency.

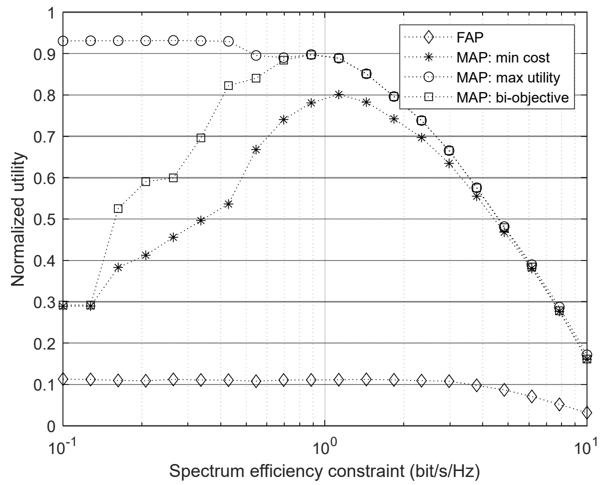


Figure 38: Simulated user utilities versus minimum constraint for spectral efficiency.

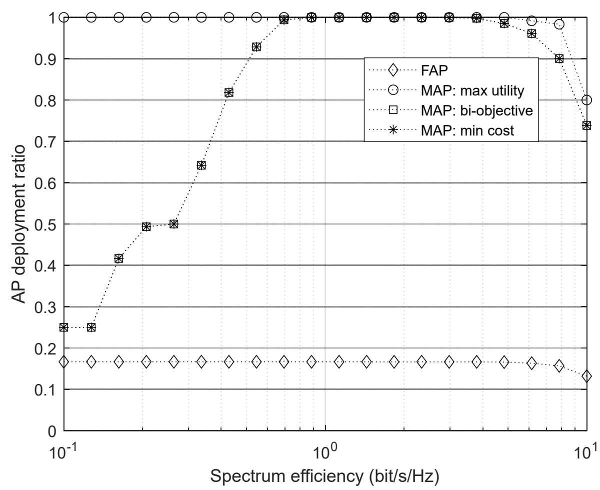
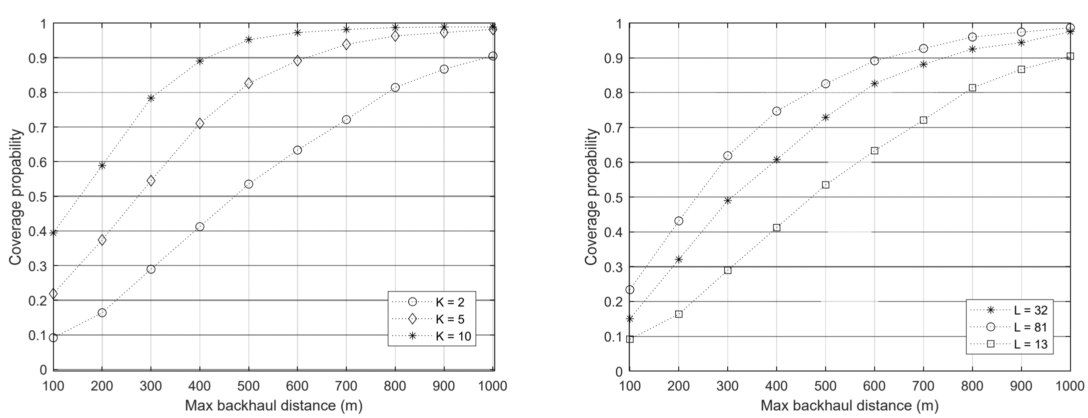


Figure 39: Simulated AP deployment ratios versus minimum constraint for spectral efficiency.



a) Effect of number of FAPs

b) Effect of number of candidate MAP points.

Figure 40: Simulated coverage probabilities versus maximum backhaul distance for different FAP and MAP densities with clustered traffic using the bi-objective approach.

4.4.4 Conclusions

We address joint user association and MAP placement problem by deploying a bi-objective optimization approach which address the trade-off between the network cost minimization and user utility maximization in a single framework. It is observed that the target spectrum efficiency constraint significantly affects the relative performance between different optimizers. Moreover, it is demonstrated how the coverage probability performance is affected by the backhaul distance constraint and density of fixed and mobile network infrastructure for the clustered traffic model. In future work, distributed association approaches could be studied to further simplify the overall control structure.

4.5 Mechanism managing robot based MAPs

In this subsection we study coverage extension in mobile communication networks in challenging situations such as areas where infrastructure has become unavailable due to natural disasters, areas with difficult morphology where it is difficult to set up infrastructure and more. For this reason, it is considered a remote area where wireless infrastructure is not widely available, but it comprises some MAPs and APs that can offer access to remote users taking into consideration among others the capacity of each network entities, the transmission range. This subsection provides the updated formulation of the problem, the initial proposed solution approach with some preliminary results, the status of connecting the algorithm with the architecture, the visualization dashboard developed and some future work.

4.5.1 Updated Problem Statement/Formulation

It is assumed a set of the MAP entities, M, a set of the APs, A, a set of users (with mobile clients), U and a set of docking stations, D, e.g., for charging drones. Moreover, it is assumed a set of locations at which the network elements (MAPs, APs) can be placed, L. Each MAP $m \in M$ has a transmission range $range_m$ and a capacity cap_m reflecting the maximum number of users (or other MAPs) that can be served by a specific MAP. To avoid congestion issues (e.g., a lot of users are served at the same time by one only MAP), we set low values of capacity to MAPs. Additionally, each AP $a \in A$ has a transmission range $range_a$ and a capacity cap_a reflecting the maximum number of users and MAPs that can be served.

The decision variables considered for describing the problem are:

$$W_m = \begin{cases} 1, if\ MAP\ m \in M\ is\ utilized \\ 0, otherwise \end{cases}$$

$$B_{a,u} = \begin{cases} 1, if\ AP\ a \in A\ is\ connected\ with\ user\ u \in U \\ 0, otherwise \end{cases}$$

$$X_{m,i} = \begin{cases} 1, if\ MAP\ m \in M\ is\ located\ in\ i \in L \\ 0, otherwise \end{cases}$$

$$Y_{m,n} = \begin{cases} 1, if\ MAP\ m \in M\ is\ connected\ with\ MAP\ n \in M \\ 0, otherwise \end{cases}$$

$$Z_{u,m} = \begin{cases} 1, if\ user\ u \in U\ is\ connected\ with\ MAP\ m \in M \\ 0, otherwise \end{cases}$$

$$Q_{m,a} = \begin{cases} 1, if\ MAP\ m \in M\ is\ connected\ with\ AP\ a \in A \\ 0, otherwise \end{cases}$$

$$D_{m,d} = \begin{cases} 1, if\ MAP\ m \in M\ is\ connected\ with\ docking\ station\ d \in D \\ 0, otherwise \end{cases}$$

Furthermore, it is introduced a communication cost, ct, among two connected entities (MAPs, APs, users, locations) related to the distance of the entities and the frequency used. As a result, ct values can be high for entities that are far away between each other, to let the algorithm determine closer entities (if available). Also, the movement of a MAP to a docking station for charging results in a moving cost mc. In general, this cost depends on the distance of the MAP from the current position to the docking station.

Accordingly, the overall optimization problem can be formulated as the minimization of the following objective function (OF):

$$min\ OF = \min_{W,X,Y,Q,D} w_1 \sum_{m \in M} W_m + w_2 \sum_{m \in M} \sum_{i \in L} \left(X_{m,i} * ct(m,i)\right) + w_3 \sum_{m \in M} \sum_{i \in L} \sum_{a \in A} (X_{m,i} * Q_{m,a} * ct(a,m)) + w_4 \sum_{m \in M} \sum_{i \in L} \sum_{d \in D} (X_{m,i} * D_{m,d} * mc(d,m))$$

Subject to (constraints):

$\sum_{i \in L} X_{m,i} = 1, \forall\, m \in M$ every MAP can be placed at one only location

$\sum_{m \in M} X_{m,i} \leq 1, \forall\, i \in L$ one MAP at most can be placed at each location

$X_{m,i} \geq 1 - W_m, \forall\, m \in M,\ where\ i\ represents\ the\ initial\ location\ of\ MAP\ m,$ this constraint ensures that the MAPs that are not utilized ($W_m = 0$) stay at their initial location

$B_{a,u} = 1, when\ Distance(a,u) \leq range_a, \forall\, a \in A, u \in U,$ it ensures that the users that are inside APs range, they are connected with that AP

$Y_{m,n} = Y_{m,n}, \forall\, m,n \in M,$ it denotes that a MAP connected with another MAP, is a connection of that MAP as well

$\sum_{a \in A} Q_{m,a} + \sum_{a \in A} Q_{n,a} \geq Y_{m,n}, \forall\, m,n \in M,$ it denotes that at least one of the two connected MAPs should be connected with an AP (assumption for reducing a bit the complexity).

$\sum_{a \in A} Q_{m,a} \leq 1, \forall\, m \in M,$ it denotes that each MAP can be connected with at most one AP.

$\frac{\sum_{a \in A} Q_{m,a}}{A} \leq W_m, \forall\, m \in M,$ it ensures that an AP is connected with MAPs that are utilized

$\sum_{m \in M} Z_{u,m} + \sum_{a \in A} B_{a,u} = 1, \forall\, u \in U,$ it ensures that each user must be connected with only one MAP or AP

$\frac{\sum_{u \in U} Z_{u,m}}{U} \leq W_m, \forall\, m \in M,$ it ensures that a MAP can connect with users only if it is utilized ($W_m = 1$)

$\sum_{d \in D} D_{m,d} = 1, \forall\, m \in M,$ each MAP should be assigned (potentially connected) with only one docking station

$\Phi_m \leq cap_m, where\ \Phi_m = \sum_{u \in U} Z_{u,m} + \sum_{n \in M} Y_{m,n} + \sum_{a \in A} Q_{m,a}, \forall m \in M,$ the capacity of each MAP is respected

$\Phi_a \leq cap_a, where\ \Phi_a = \sum_{u \in U} B_{a,u} + \sum_{m \in M} Q_{m,a}, \forall a \in A,$ the capacity of each AP is respected

$X_{m,i} * Z_{u,m} * Distance(u,m) \leq range_m, \forall\, m \in M, i \in L, u \in U,$ the users connected with a MAP should be located inside the MAP's range

$X_{m,i} * X_{n,j} * Distance(m,n) \leq range_m, \forall\, m,n \in M, i,j \in L,$ the MAP connected with a MAP should be located inside the MAP's range

$X_{m,i} * Q_{m,a} * Distance(a,m) \leq range_a, \forall\, m \in M, i \in L, a \in A,$ the MAPs connected with an AP should be located inside the AP's range

$\sum_{a\in A} Q_{m,a} + \sum_{n\in M, m\neq n} Y_{m,n} \geq W_m, \forall\, m \in M$, each MAP should be connected with an AP or a MAP connected with an AP.

$\sum_{m\in M} \sum_{a\in A} Q_{m,a} \geq \frac{\sum_{m\in M} W_m}{M}$, there should be at least one connected AP with a utilized MAP.

The *objective function (OF)* monitors the location of all MAPs and calculates the total communication costs that are related from these locations, as well as the finding of the nearest docking station for charging (of robot MAPs). The first term of the function illustrates the cost of utilizing a MAP (ensures that we utilize the minimum needed). The second term of the function illustrates the communication cost related to the MAPs location. The third term depicts the communication cost of connecting a MAP with an AP and the fourth term depicts the cost related to the docking station where the drones should go for charging. All terms are weighted (w_1, w_2, w_3, w_4) depending on the use case. As it is observed the formula is not linear, but we can linearize it by replacing the products of binary variables e.g., $\prod_{i=1}^{n} x_i$ with a binary variable z respecting the following constraints: a) $z \leq x_i \forall i = 1, \dots, n$ and b) $z \geq \sum_{i=1}^{n} x_i - (n-1)$.

4.5.2 Initial proposed solution and preliminary results

The above problem is initially solved with the use of a Mixed Integer Programming (MIP) python solver called Python-MIP (https://www.python-mip.com/). We first linearized the formula and the constraints, and we tried various scenarios to test its performance. MIP solvers are known to provide the optimal solution but are computationally intractable, especially for large scale experimentation. For this reason, we are working on developing a metaheuristic algorithm with close to optimal solutions which will be described in the following deliverable.

As we are in a process of approaching the problem with a different algorithm, we are presenting here a scenario with its suggested solution of the MIP model and the execution time. We assumed a 2D space of 5x6 squares depicting the 30 available locations (L0, L1, …, L29). We assume 2 AP, 8 users, 2 docking stations and 4 MAPs initially located as it is shown in Figure 41. All ct and mc appearing in OF are distances between the relevant network elements. The weights utilized for our example are all equal apart from w_1 which is a bit larger so that we ensure that it is utilized the minimum MAPs. The capacity of each MAP is $cap_m = 5$ and the capacity of each AP is $cap_a = 6$. The maximum range of each AP $range_a$ is two squares far and the maximum range of each MAP $range_m$ is one square far. The output of this scenario can be seen in the terminal in Figure 42 as well as it can be seen schematically in the right side of Figure 41. As it is shown AP0 is connected with users U0, U2 and MAP0, and AP1 is connected with MAP2. MAP3 is not utilized so it remains to its initial position. Moreover, we can see that MAP1 is connected with users U1 and U3, MAP2 is connected with MAP0 and MAP0 is connected with users U4, U5, U6 and U7. Finally, docking station D1 is assigned to MAP0 and MAP2 which are the closest utilized MAPs and docking station D0 is assigned to MAP1. The execution time of this example was 109.65 sec.

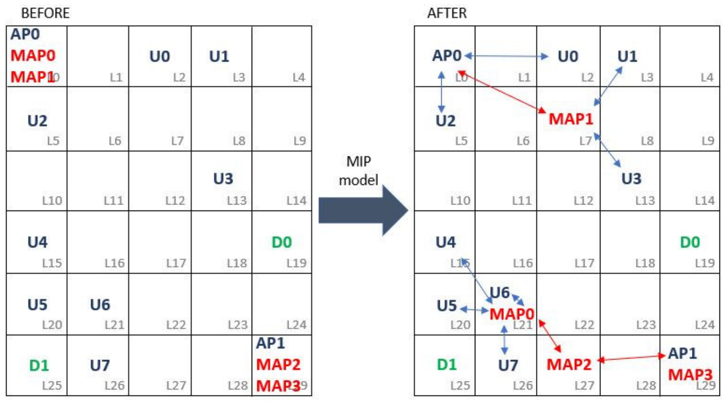
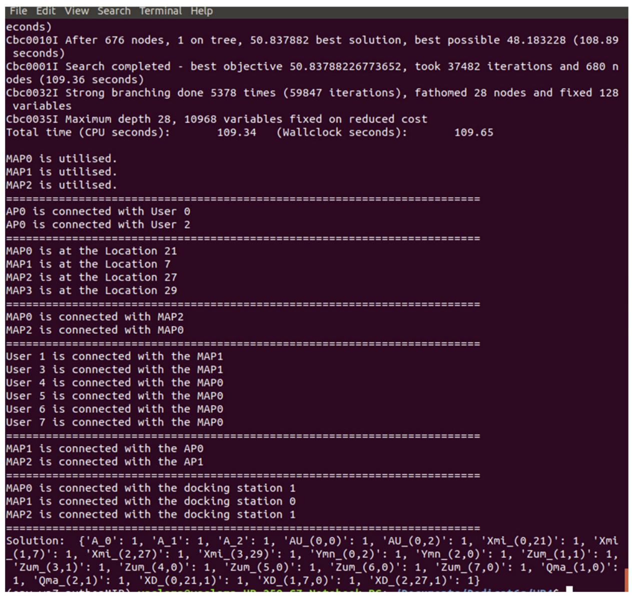


Figure 41: Schematically represented output of MIP mechanism managing robot based MAPs model when having a 2D area of 2 AP, 4 MAPs, 8 users and 2 docking stations.



```
File Edit View Search Terminal Help
econds)
Cbc0010I After 676 nodes, 1 on tree, 50.837882 best solution, best possible 48.183228 (108.89
 seconds)
Cbc0001I Search completed - best objective 50.83788226773652, took 37482 iterations and 680 n
odes (109.36 seconds)
Cbc0032I Strong branching done 5378 times (59847 iterations), fathomed 28 nodes and fixed 128
 variables
Cbc0035I Maximum depth 28, 10968 variables fixed on reduced cost
Total time (CPU seconds):       109.34   (Wallclock seconds):       109.65

MAP0 is utilised.
MAP1 is utilised.
MAP2 is utilised.
========================================================================
AP0 is connected with User 0
AP0 is connected with User 2
========================================================================
MAP0 is at the Location 21
MAP1 is at the Location 7
MAP2 is at the Location 27
MAP3 is at the Location 29
========================================================================
MAP0 is connected with MAP2
MAP2 is connected with MAP0
========================================================================
User 1 is connected with the MAP1
User 3 is connected with the MAP1
User 4 is connected with the MAP0
User 5 is connected with the MAP0
User 6 is connected with the MAP0
User 7 is connected with the MAP0
========================================================================
MAP1 is connected with the AP0
MAP2 is connected with the AP1
========================================================================
MAP0 is connected with the docking station 1
MAP1 is connected with the docking station 0
MAP2 is connected with the docking station 1
========================================================================
Solution:  {'A_0': 1, 'A_1': 1, 'A_2': 1, 'AU_(0,0)': 1, 'AU_(0,2)': 1, 'Xmi_(0,21)': 1, 'Xmi
_(1,7)': 1, 'Xmi_(2,27)': 1, 'Xmi_(3,29)': 1, 'Ymn_(0,2)': 1, 'Ymn_(2,0)': 1, 'Zum_(1,1)': 1,
 'Zum_(3,1)': 1, 'Zum_(4,0)': 1, 'Zum_(5,0)': 1, 'Zum_(6,0)': 1, 'Zum_(7,0)': 1, 'Qma_(1,0)':
 1, 'Qma_(2,1)': 1, 'XD_(0,21,1)': 1, 'XD_(1,7,0)': 1, 'XD_(2,27,1)': 1}
```

Figure 42: Terminal python output of MIP mechanism managing robot based MAPs model when having a 2D area of 2 AP, 4 MAPs, 8 users and 2 docking stations.

In the scope of testing coverage extension a Clearpath Robotics Jackal Unmanned Ground Vehicle/AGV will be used as a MAP to reach a specific location with the aim of providing users with opportunistic wireless network for coverage extension. For demonstration purposes this is accompanied by a visualization dashboard that is connected to the robot (Figure 43).

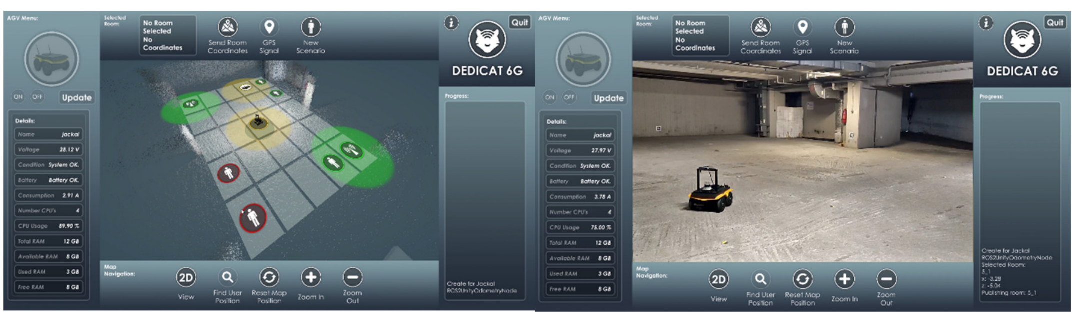


Figure 43: Coverage extension visualization dashboard

4.5.3 Conclusion and future work

In this subsection we presented the preliminary work done related with the mechanism managing robot based MAPs coverage extension algorithm. We provided the updated formulation of the problem, the solution explored till now and some preliminary results as well as some pictures of the VR app developed for integrating our coverage extension work with available robots. For the next deliverable, it is planned to explore another model for succeeding faster execution time for large scale experimentation. We plan to provide evaluations and more progress of the VR app and platform with utilizing real robots.

4.6 NFV Orchestration

As we introduced in D4.1, NFV is a key technology in the B5G ecosystem and an important topic to be taken into account when extending the coverage in a dynamic 5G-based manner. Thus, the DEDICAT 6G platform will command the NFV Orchestrator (NFV-O) to correctly instantiate the required VNFs linked to a specific network service/slice to enable ad-hoc connectivity when extending or moving the connectivity area. The concrete actor to perform such functionality within the DEDICAT 6G approach is the Orchestration Engine (OE).

The orchestration engine is part of the service orchestration FC and serves as a link between the decision-making FCs and the NFV orchestrator to correctly instantiate the ad hoc network slices, thus enabling the provisioning of a unique 5G/B5G-based virtualized network space to run vertical's apps under the requested conditions. Moreover, the orchestrator engine will assist the Decision Making FG in the application of the instantiation, scaling, or migration network service/slices orchestration procedures according to the outputs generated by the algorithms in this FG.

To play such role within the DEDICAT 6G platform, the orchestration engine must receive instructions and concrete information coming from several internal FCs and with the external NFV Orchestrator. The overall functional overview and their interactions are represented in Figure 44 and summarized as follows:

- **CEDM FC:** It is the FC responsible to take decisions about coverage extension. Its output will feed the orchestration engine with information about the existing MAPs deployment where the network provisioning is required.
- **IDDM FC:** The IDDM will send to the orchestration engine some recommendations to ensure optimality in the management of the computational resources consumed by the NFV-related intelligence (e.g., VNFs).
- **NODM FC:** This FC will assist the orchestration engine to properly configure the network slices by providing network-related information.
- **µS/FC Registry FC:** It can serve information about the computational requirements of to either µS or FC when needed.
- **µS/FC Repository FC:** The orchestration engine can access this repository when information about the container/VM images of the FC/µS is required.
- **NFV Orchestrator:** This external entity is the one responsible to manage the NFV infrastructure (NFVI) resources to establish network slices. The orchestration engine will command the NFV-O according to the decisions made within the DEDICAT 6G platform.

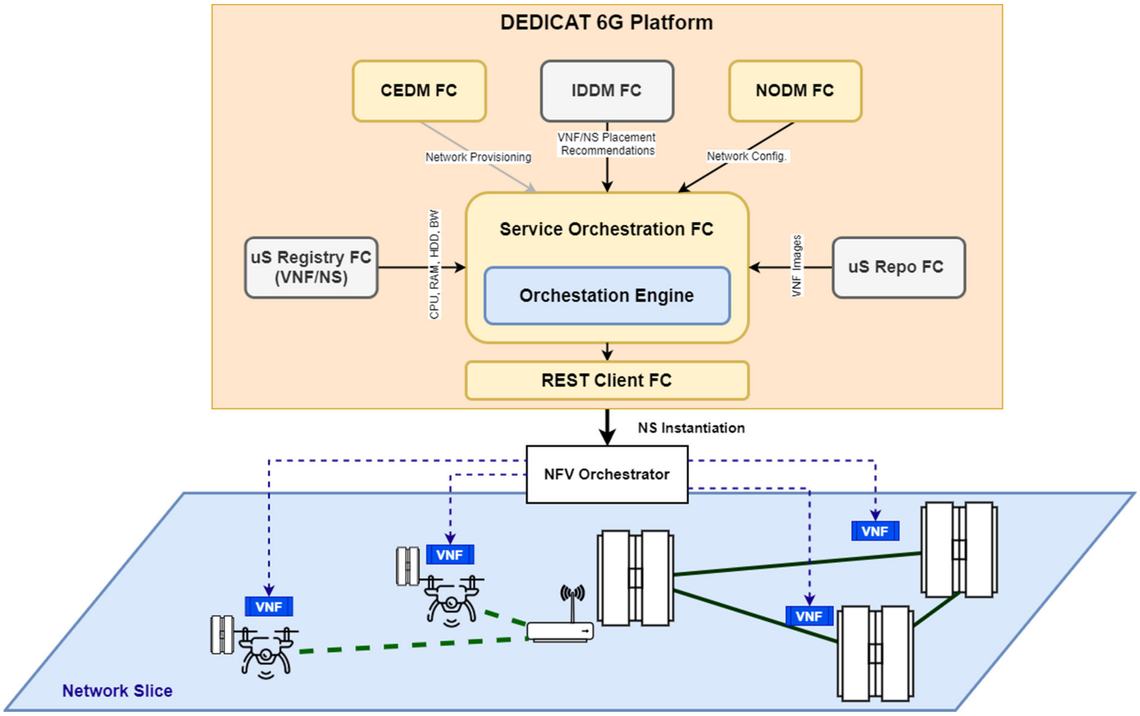


Figure 44: Orchestration engine schema in the DEDICAT 6G platform.

It is worth mentioning that the nature of the orchestration engine component tackles both WP3 and WP4 domains, thus, some parts in D3.2 and D4.2 are common in both documents to clearly understand this work in a standalone way. For the interest of WP4, we will put the light, first, on the internal interactions among the CEDM, the NODM and the Orchestration Engine, and also, on the interaction between the OE and the NFV-O. This section complements the vision presented in D3.2, where the focus is on the analysis of the computational impact on the IT resources available at the edge nodes derived from the instantiation of network slices.

4.6.1 Orchestration Engine Design

In this work, we present a cloud-native modular design for the orchestration engine component ready to work in a distributed micro service environment. This design is focused to be implemented in a lightweight and scalable implementation suitable for deployment both in the cloud and at the edge domains. It is worth mentioning that this design is targeted to communicate with standards and trending external tools.

The Orchestration Engine is focused on assisting the Decision Making FG to translate and apply recommendations in the external 5G/B5G system, mainly focused on the NFV side by means of the NFV-O. Our design (shown in Figure 44) assumes the ETSI Open-Source MANO (OSM) as NFV Orchestrator. In this iteration, the orchestration engine design comprises 5 main sub-modules:

Engine Manager: This module works with all the other components to provide the main functionalities of the orchestration engine and as a link of the rest of the modules. It is in charge of managing all internal modules in a coordinated way according to the requests coming from the Decision Making FCs.

Blueprint/Descriptor creator: This entity is responsible for automatically creating or modifying descriptors based on incoming requests. A descriptor defines what is un-

derstood as a VNF, a network service and a network slice. These templates are based on the ETSI SOL006 data model [26], which are supported by OSM.
OSM Client: This component is in charge of implementing OSM Client to be able to communicate with OSM through the native OSM API. It also implements some new features that are not currently available in OSM Client, for instance, post processing messages received by OSM to extract more valuable and concrete information to manage the instances which are running.
gRPC Server: It exports the set of functionalities offered by the orchestration engine to the rest of FCs in the DEDICAT 6G platform. Also, the gRPC-based server to enable rapid and efficient communications channels among the different FCs. More information about gRPC and its benefits for micro service communications in the next subsection.
Cache DB: This database will store temporal and persistent data, like information about the network and received messages, necessary for the orchestration engine tasks.

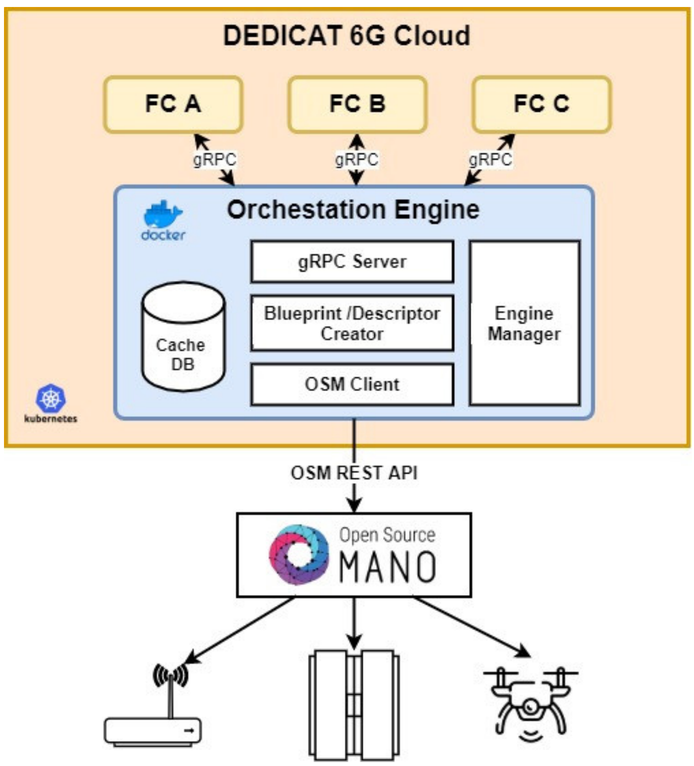


Figure 45: Orchestration Engine design

4.6.2 Experiment Set-up

For the purpose of implementing, testing and validating the functionalities of the orchestration engine, details of a preliminary set of experiments aimed at that end are shown here.

4.6.2.1 Particular WP4 Goals

As exposed before, the orchestration engine is a component that implements some key functionalities of the WP2-defined Service Orchestration FC. The main objective of this experiment is to demonstrate the key functionalities of such component within a realistic B5G-driven environment under the DEDICAT 6G umbrella. Despite, the main objective can be shared between WP3 and WP4, here we have defined a set of WP4-oriented sub-goals to be shown in this work:

- Demonstrate an agile and efficient microservice-based communication among the CEDM FC, NODM FC and the orchestration engine.
- Show a preliminary working implementation of some WP4-related FCs suitable to be potentially extended to the rest of the DEDICAT 6G platform.
- Preliminary implementation of the IDDM FC output WP2-defined data model.
- Automatic creation of VNF, network service and network slice descriptors/templates according to the outputs provided by the FCs of the Decision Making FG following the specifications defined in [26].
- Show how the OE instructs the NFV-O to instantiate network slices according to the recommendations given by the FCs of the Decision Making FG.

4.6.2.2 Related Technologies

We now introduce the set of open-source software tools employed to perform the corresponding functions presented in the orchestration engine design chosen to meet the above-mentioned objectives:

- **Docker [27]:** Docker is an open-source tool that automates the deployment of applications running inside software containers, thereby bringing an additional layer of abstraction and automation to the virtualisation of applications. It is currently the most widely used tool in microservices-based deployments. In this work, Docker is used to create the containers where the different components are implemented and deployed.
- **Kubernetes [28]:** is an open-source platform for managing container deployment, as already presented in D3.1. The role of Kubernetes in this implementation is to manage and host the Docker containers related to the DEDICAT 6G FCs and also to host the OSM instance.
- **gRPC [29]:** it is a high-performance, lightweight, and RPC-based communication protocol used in microservices environments and hosted by the Cloud Native Computing Foundation. A prominent feature of gRPC is that its data is structured using strict and lightweight rules defined by Protocol Buffers [30] , a framework for serving structured data. The data in Protobuf, at the same time as in gRPC, is specified and defined by an Interface Description Language (IDL). It can create language-independent interfaces to support communication between servers or clients written in different languages. The present work generates gRPC libraries to enable the communication among the DEDICAT 6G FCs. Additionally, the messages of the DM-related FCs (outputs) have been implemented by using the data models defined in D2.4.
- **SQLite [31]:** this lightweight and fast SQL-driven database is the chosen one to represent the role of the cache DB in the orchestration engine.
- **OSM [32]:** Open Source MANO is the ETSI-hosted NFV orchestrator fully aligned with ETSI NFV standards. As explained extensively in D4.1, it is capable of managing and orchestrating virtualised resources in NFV-based ecosystems to enable the instantiation of 5G/B5G network slices. This is the tool selected for these purposes as NFV-O in the present work.
- **OpenStack [33]:** is one of the most popular open-source cloud infrastructures used to manage and host third-party services in virtual machines, bare metal and containers. In this implementation, OpenStack plays the role of a Virtual Infrastructure Manager (VIM) that represents an edge node with the computational resources to host the VMs of the VNFs in a network instantiation. In our case, we consider two different Open-Stacks as follows in the next sub-section.

4.6.2.3 Implementation Schema

The trial has been carried out on the ATOS Telecommunications testbed, following the implementation scheme described in Figure 46. The testbed features a Kubernetes cluster (K8s) representing the cloud domain intended to host and control the Docker containers, in this case, the DEDICAT 6G FCs and the orchestration engine, all of them written in Python. In addition, in this same cluster, we have deployed an instance of OSM, playing the role of NFV-O. Finally, the edge domain is modelled with two Open Stacks (VIM) instances, physically separated, to consider two different edge nodes. We assume both VIMs as part of the NFV infrastructure to enable the establishment of multi-site network slices.

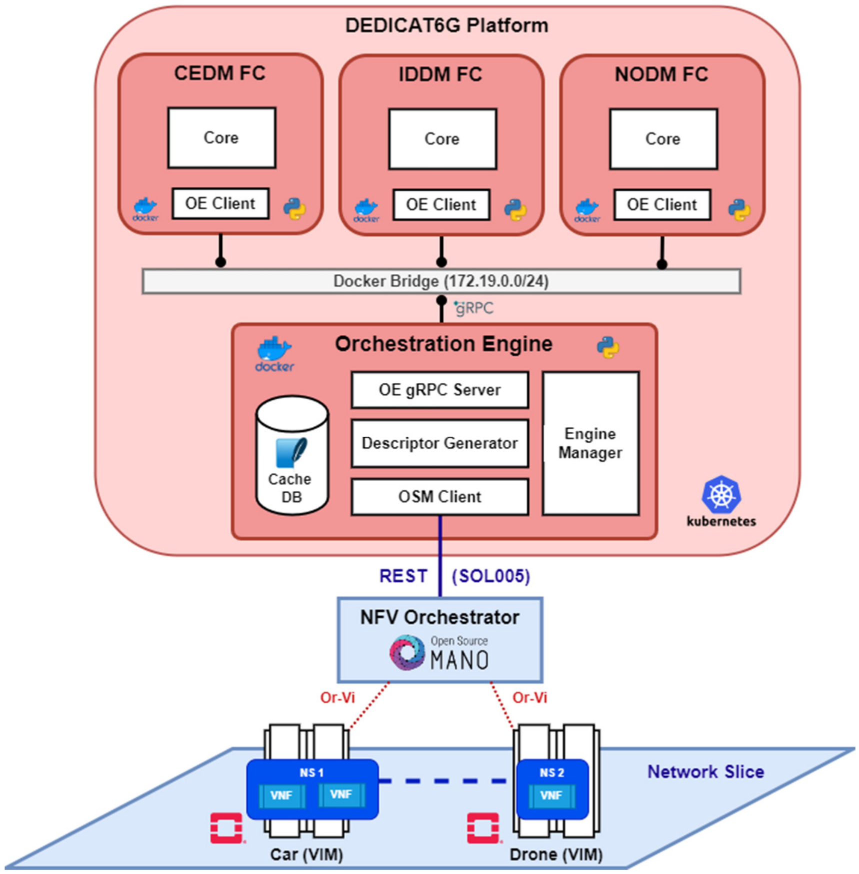


Figure 46: Orchestration Engine Implementation Schema

Regarding connectivity among existing entities, we consider two types of communication, internal to DEDICAT 6G and external to the rest of the system. A Docker bridge (a virtual network) has been configured to allow communication within the DEDICAT 6G platform. This bridge is used to enable the establishment of gRPC channels, which is the microservices-based communication chosen for data exchange between the FCs and the orchestration engine. In this work, the orchestration engine is considered as the anchor of the DEDICAT 6G platform to link to external entities. In this case, the orchestration engine has an OSM client to allow leveraging the by-default OSM REST interface. It is noteworthy that this interface is defined in the ETSI GS NFV-SOL 005 specification. [34]. Finally, OSM recognizes the two Open Stacks as VIMs, and uses the Or-Vi interface for the instantiation of the VNFs in an automatic manner [Or-Vi].

4.6.2.4 Workflow

Here, in Figure 47, we present the workflow followed in this work. It is important to clarify that this experiment starts with the interaction between the CEDM FC and the orchestration engine. Thus, we assume that the CEDM and the NODM have previously calculated their output. In this work, we focused on the orchestration engine side, not in the optimization provided by the algorithms of the Decision Making FG. Also, due to the relation to coverage extension mechanisms, the parallel interactions among the orchestration engine and the IDDM are shown in D3.2. The steps in this experiment are as follows:

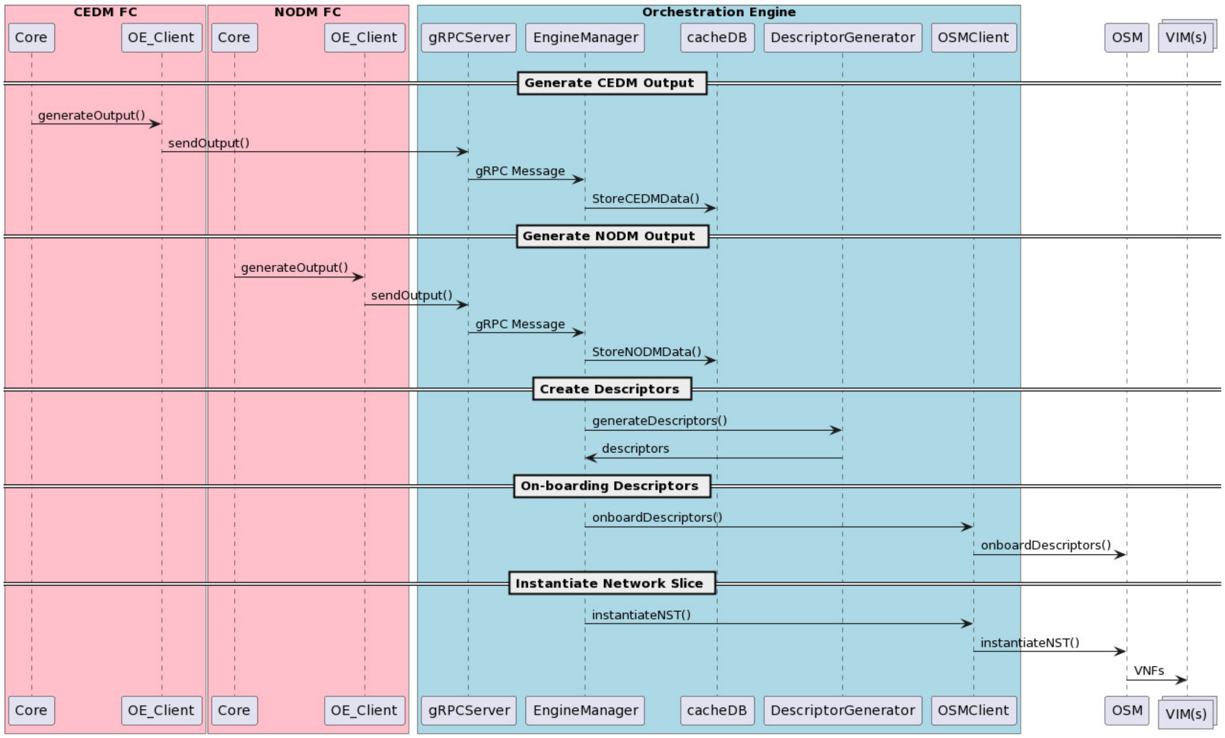


Figure 47: Sequence Diagram of the workflow

1. The Core of the CEDM FC generates the output by using the structure defined in the corresponding WP2 data model, and encapsulate this info in the shape required and sends it to its OE_Client
2. The OE_Client, opens a gRPC channel to send the generated message (CEDM output) to the Orchestration Engine gRPCServer by using a concrete rpc.
3. Now, the IDDM Output is in the orchestration engine domain, and the gRPC forwards the message to the EngineManager
4. Then, the EngineManager is in charge of parsing the data and structure it according to the internal structure of the cacheDB, and store it
5. Steps 1 to 4 are repeated for the NODM FC in an analogous manner.
6. Once the CEDM and NODM data is stored and processed, the EngineManager commands the DescriptorGenerator to automatically create the corresponding descriptors (VNFDs, NSDs and NSTD) with the CEDM and NODM information in the correct fields according to the ETSI SOL006 data model.
7. After this stage, the network slice is properly modelled in this set of descriptors, and they are onboarded in OSM to be ready for their instantiation
8. Then the network slice is ready to be instantiated. Thus, the EngineManager requests the OSMClient to instantiate the NST (Network Slice Template in ETSI SOL006 terminology).

9. The OSMClient sends a request by using the OSM Representational State Transfer (REST) interface to instantiate a network slice based on the descriptor attached in that request.
10. Finally, OSM process this request and orchestrate the instantiation of the network slice in the VIMs.
11. After a few seconds, the network slice attached to the extension of the coverage is properly configured and available to be used, creating a dedicated 5G network space.

4.6.3 Preliminary Results

As can be seen in Figure 48, four different Docker containers have been built and instantiated, three that emulate the CEDM, IDDM and NODM FCs, and an additional one that contains the orchestration engine. Each of these containers, have pre-configured ports open to enable gRPC communication by using the Docker bridge, thus, allowing the transmission and reception of messages compliant with the data models defined in D2.4.

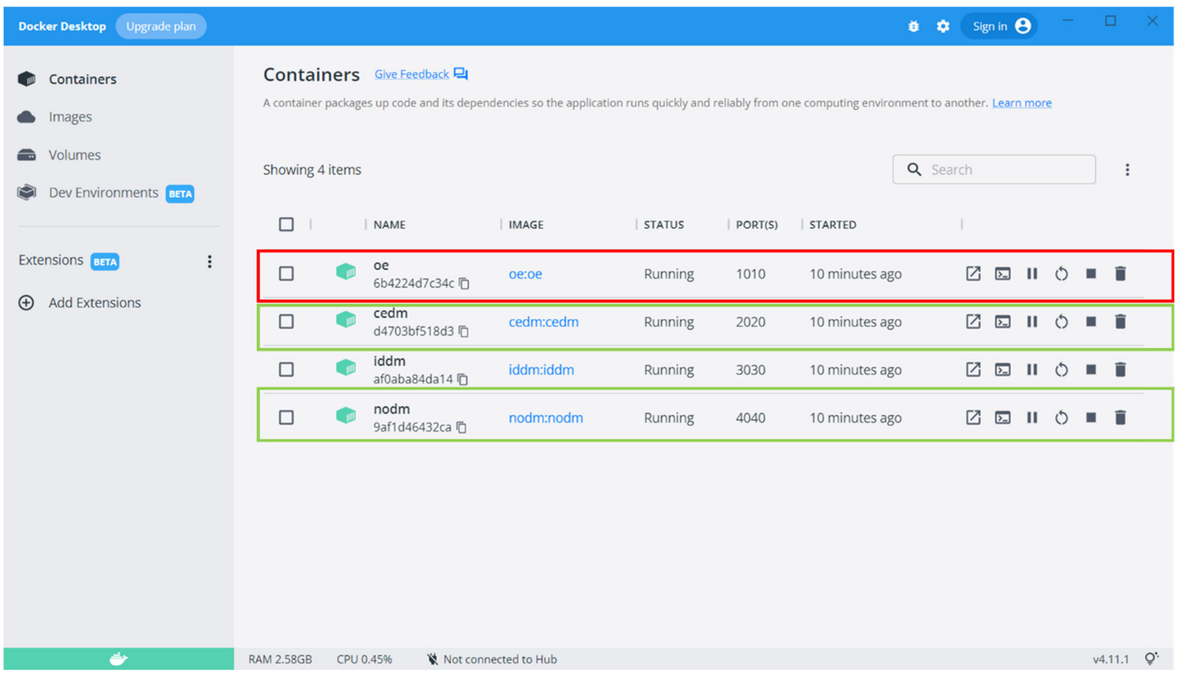


Figure 48: Snapshot of the Docker environment with all the containers up and running

In this scenario, our main target is to instantiate a network slice, called *NST001*, which includes two NSs, with two and one VNFs respectively, more information about the NST structure below. The first step once all containers are up and running, is to start the gRPC server of the orchestration engine, as can be seen in Figure 49a. Once the gRPC server is ready, the FCs can send all output messages, by using a specific gRPC Orchestrator Engine client that will be received and stored by the OE. In this case, we will focus on the CEDM and NODM side, as IDDM is further detailed in D3.2.

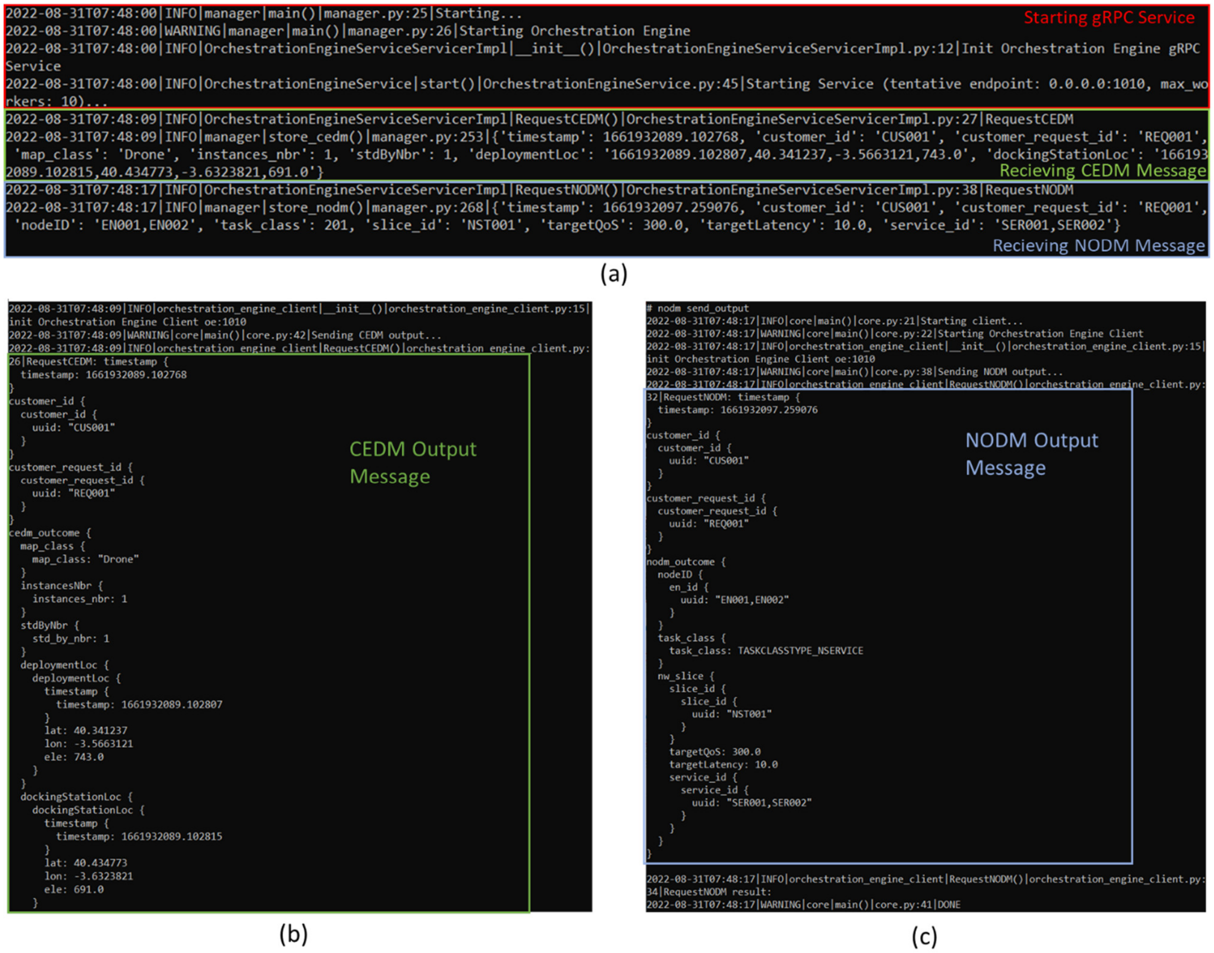


Figure 49: Terminal snapshot with logs messages of the containers corresponding to: a) Orchestrator Engine; b) CEDM; and c) NODM.

The core of the CEDM FC and NODM FC build the messages that are being received using internal algorithms, here we assume the algorithm have been previously executed. It is worth mentioning that the messages follow the structure defined in D2.4 for the output generated by the CEDM and NODM FCs and implemented by using Protocol Buffer. Then these messages are received by the OE gRPC server and stored in the cacheDB in the orchestration engine as can be seen in Figure 50. In this figure, it can be appreciated how the CEDM indicates the OE about how the coverage extension associated to the request "REQ001" should take into place. On one hand, the CEDM asks the drone to move to a location, to act as MAP. In this case, we assume that the car will act as an access point to the drone, enabling 5G connectivity linked to the backhaul side. On the other hand, the NODM message indicates a network slice should be instantiated, including two network services "SER001" and "SER002", which are the drone_ns and car_ns, respectively. This slice will allow the data exchange between the drone and the car, guaranteeing the connectivity and enabling the coverage extension.

CEDM table

	cedmRequestID	timestamp	customer_id	customer_reques	map_class	instances_nbr	stdByNbr	deploymentLoc	dockingStationLoc
1	1	1661932089.102768	CUS001	REQ001	Drone	1	1	1661932089.102807,40.341237,-3.5663121,743.0	1661932089.102815,40.434773,-3.6323821,691.0

NODM table

	nodmRequestID	customer_id	timestamp	customer_reques	nodeID	TaskClass	slice_id	targetQoS	targetLatency	service_id
1	1	CUS001	1661932097.259076	REQ001	EN001,EN002	201	NST001	300	10	SER001,SER002

Figure 50: Snapshot of the SQLite GUI showing the IDDM information stored

When the OE receives enough information to process the new request, including some information sent by IDDM, the OE starts processing the information to instantiate a network slice containing the necessary NSs for the drone and the car. These NSs, are associated with one or several VNF images, also including resource information like RAM, Virtual CPU (VCPU) and Storage to be consumed by the VNFs. As can be seen in Figure 51, this mapping between the NSs and VNFs is stored in the cache DB (and managed by the Engine Manger), this information can be obtained via the µS/FC registry and repository FCs. In our experiment, this functionality is modelled by the internal registries of OpenStack.

VNF table

	vnf_id	name	image	memory	vcpu	storage
1	1	car_cirros_1	cirros-0.3.5-x86_64-disk.img	2	2	20
2	2	car_cirros_2	cirros-0.3.5-x86_64-disk.img	3	3	30
3	3	drone_cirros_1	cirros-0.3.5-x86_64-disk.img	1	2	20

NS table

	ns_id	ns_name	instances_nbr	vnf_id
1	1	car_ns	1	1
2	2	car_ns	1	2
3	3	drone_ns	1	3

Figure 51: VNF internal mapping to associate to the NSs in OE cache DB.

Once the information regarding VNFs for every NS is extracted, the next step is to prepare the necessary descriptors to translate the information received by the FCs, to a Data model understood by OSM (ETSI SOL006). Three types of descriptors are generated, including network slice templates (NST), network service descriptors (NSD) and VNF descriptors (VNFD). For a better understanding, Figure 52 shows the structure defined for the network slice, assuming the two NSs with its corresponding VNFs.

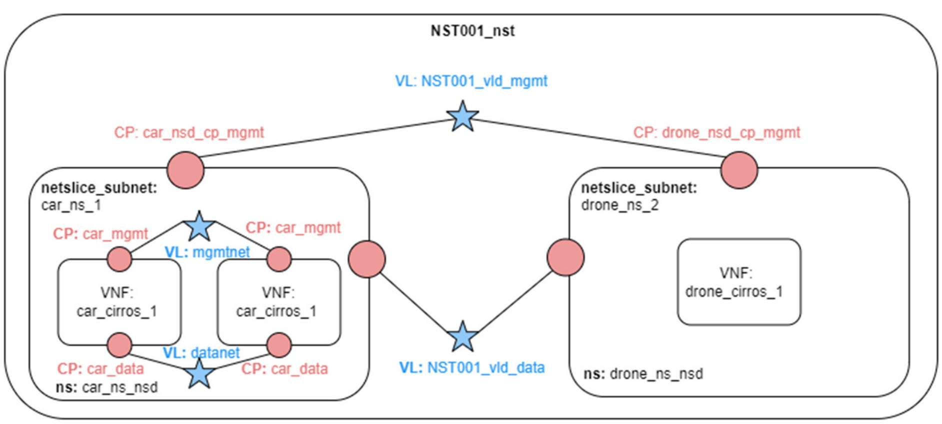


Figure 52: Network slice schema with two network services.

The first NS called *drone_ns*, emulates a NS to be instantiated in the drone IT resources, and the latter called *car_ns*, with a similar meaning. The *drone_ns* is composed by one VNF which will be instantiated in the VIM associated to the drone, while the *car_ns* has two VNFs which will be instantiated in the VIM corresponding to the car. The NST assumes a set of virtual links to interconnect its services and VNFs to provide end-to-end connectivity. Figure 52 depicts the links between all the components of the slice, which allow interconnectivity of management information and data interfaces between all VNFs in the slice.


```yaml
vnfd:
  id: car_cirros_1_vnfd
  product-name: car_cirros_1_vnfd
  description: Generated by DEDICAT6G's orchestration engine
  provider: ATOS
  version: '1.0'
  mgmt-cp: vnf-cp0-ext
  virtual-storage-desc:
  - id: car_cirros_1_vnfd-VM-storage
    size-of-storage: 20
  virtual-compute-desc:
  - id: car_cirros_1_vnfd-VM-compute
    virtual-cpu:
      num-virtual-cpu: 2
    virtual-memory:
      size: 2
  sw-image-desc:
  - id: "cirros-0.3.5-x86_64-disk.img"
    name: "cirros-0.3.5-x86_64-disk.img"
    image: "cirros-0.3.5-x86_64-disk.img"
  df:
  - id: default-df
    instantiation-level:
    - id: default-instantiation-level
      vdu-level:
      - vdu-id: car_cirros_1_vnfd-VM
        number-of-instances: 1
```


(a)


```yaml
nsd:
  nsd:
  - id: car_ns_nsd
    name: car_ns_nsd
    designer: None
    description: Generated by DEDICAT6G's orchestration engine
    version: '1.0'
    vnfd-id:
    - car_cirros_1_vnfd
    - car_cirros_2_vnfd
    df:
    - id: default-df
      vnf-profile:
      - id: car_cirros_1
        vnfd-id: car_cirros_1_vnfd
        virtual-link-connectivity:
        - virtual-link-profile-id: car_ns_nsd_vld0
          constituent-cpd-id:
          - constituent-base-element-id: "1"
            constituent-cpd-id: vnf-cp0-ext
      - id: car_cirros_2
        vnfd-id: car_cirros_2_vnfd
        virtual-link-connectivity:
        - virtual-link-profile-id: car_ns_nsd_vld0
          constituent-cpd-id:
          - constituent-base-element-id: "1"
            constituent-cpd-id: vnf-cp0-ext
```


(b)


```yaml
nst:
-   id: NST001_nst
    name: NST001_nst
    # Slice Service Type [eMBB, mMTC, URLLC]
    SNSSAI-identifier:
        slice-service-type: eMBB
    # Quality of service identifier
    quality-of-service:
        id: 1
    # Netslice subnets (Network services)
    netslice-subnet:
    -   id: drone_ns_nsd
        is-shared-nss: 'false'
        description: NetSlice Subnet (service) drone_ns_nsd
        nsd-ref: drone_ns_nsd
    -   id: car_ns_nsd
        is-shared-nss: 'false'
        description: NetSlice Subnet (service) car_ns_nsd
        nsd-ref: car_ns_nsd
```


(c)

Figure 53: Examples of some of the generated descriptors of: a) network slice (NST); b) network service (NSD) and c) VNF (VNFD).

An example of each type of descriptors can be seen in Figure 53, showing the hierarchy of each one, from the NST relating the NS, to the car VNF specifying the needed image and computational resources. Focusing on the VNF descriptor, we can see that the resources requested in the VNF table in Figure 51 the first row are mapped to the corresponding fields in the descriptor, including the RAM, VCPU, storage and also the VNF image, which in this case is a very lightweight version of Linux called cirrOS with some basic functionalities. Furthermore, a configuration file is created to map the multi-site information received by the FCs, making it possible to instantiate different NS that belongs to the same network slice in different VIMs.

Once the descriptors are generated, they are automatically packed by the Descriptor generator and onboarded into OSM by means of the OSMClient sub-module. An example of this on-boarding phase can be seen in Figure 54, where the three VNFDs have been successfully onboarded.

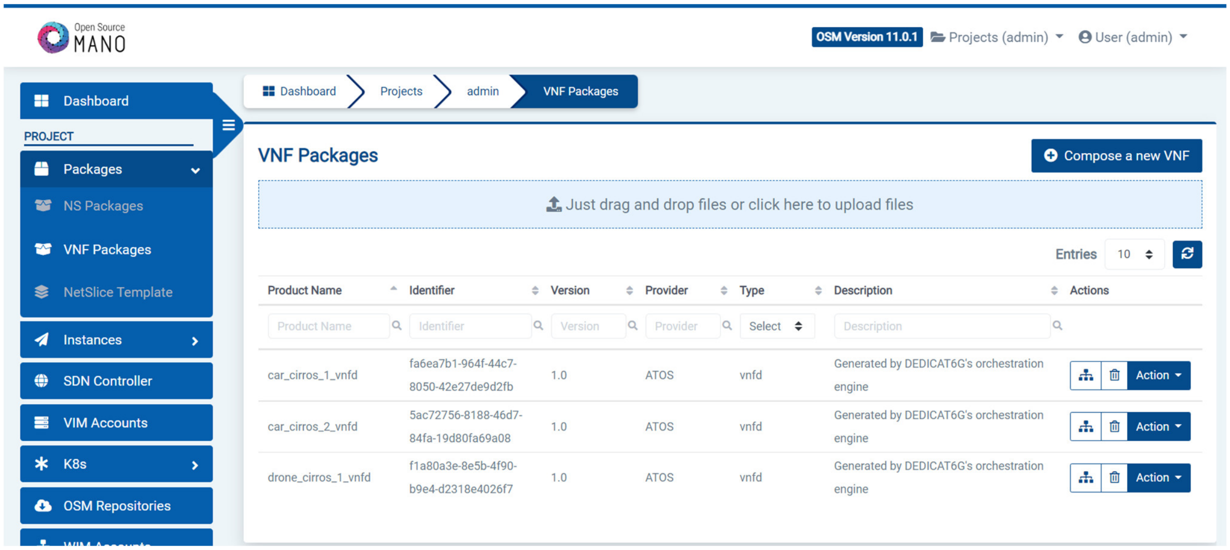


Figure 54: VNF descriptors onboarded in OSM.

After the descriptors' onboarding, the OE commands OSM to instantiate the required slice. These actions are executed via the OSM Client. Figure 55 shows the NST001 network slice successfully running in OSM dashboard.

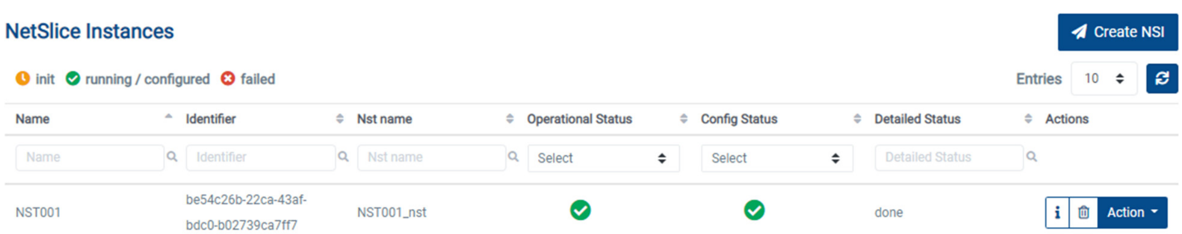
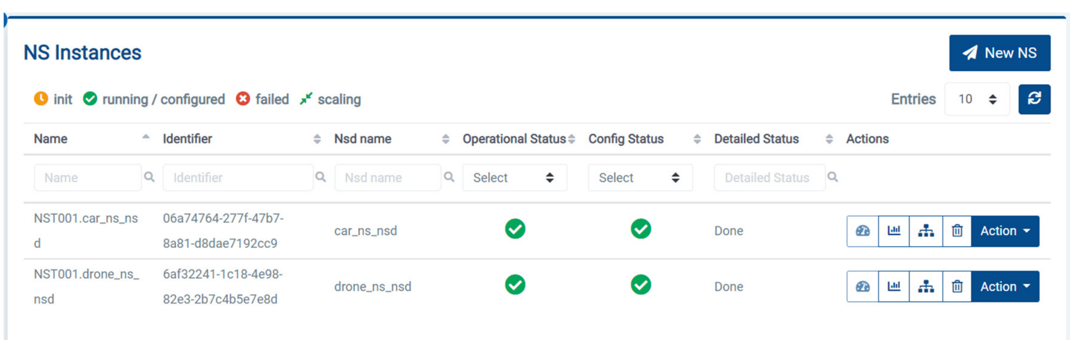
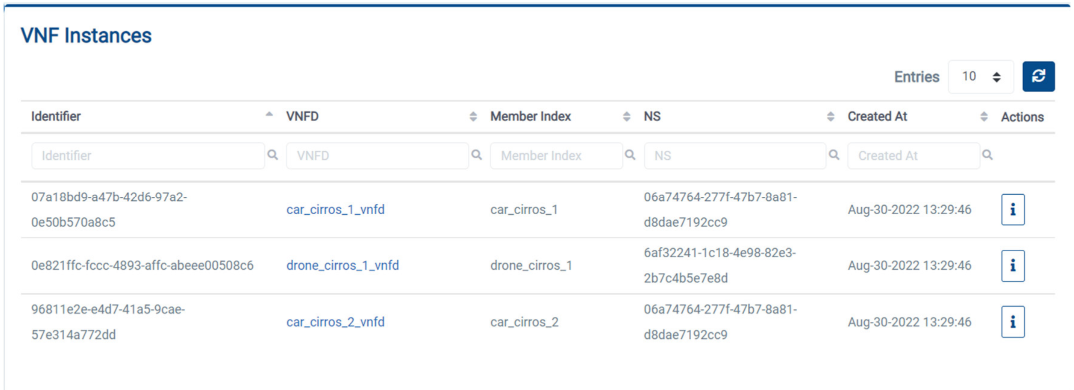


Figure 55: Snapshot of OSM dashboard with the network slice instance commanded by the Orchestrator Engine

Associated with this slice, we have the two network services successfully instantiated, representing the car_ns and the drone_ns. Finally, we can see the running VNF instances in Figure 56 which as we previously saw in Figure 51 and Figure 53, we have the two NSs up and running and the two VNF instances for the car_ns, and one for the drone_ns.





VNF Instances

Entries 10

Identifier	VNFD	Member Index	NS	Created At	Actions
07a18bd9-a47b-42d6-97a2-0e50b570a8c5	car_cirros_1_vnfd	car_cirros_1	06a74764-277f-47b7-8a81-d8dae7192cc9	Aug-30-2022 13:29:46	i
0e821ffc-fccc-4893-affc-abeee00508c6	drone_cirros_1_vnfd	drone_cirros_1	6af32241-1c18-4e98-82e3-2b7c4b5e7e8d	Aug-30-2022 13:29:46	i
96811e2e-e4d7-41a5-9cae-57e314a772dd	car_cirros_2_vnfd	car_cirros_2	06a74764-277f-47b7-8a81-d8dae7192cc9	Aug-30-2022 13:29:46	i

Figure 56: NS and VNF instances.

These network links allow the transfer of data between the drone and the car in the multi-site network slice instantiation. To demonstrate this, Figure 57 shows that a ping can be successfully made between a VNF running in the drone VIM, and one running in the car VIM.

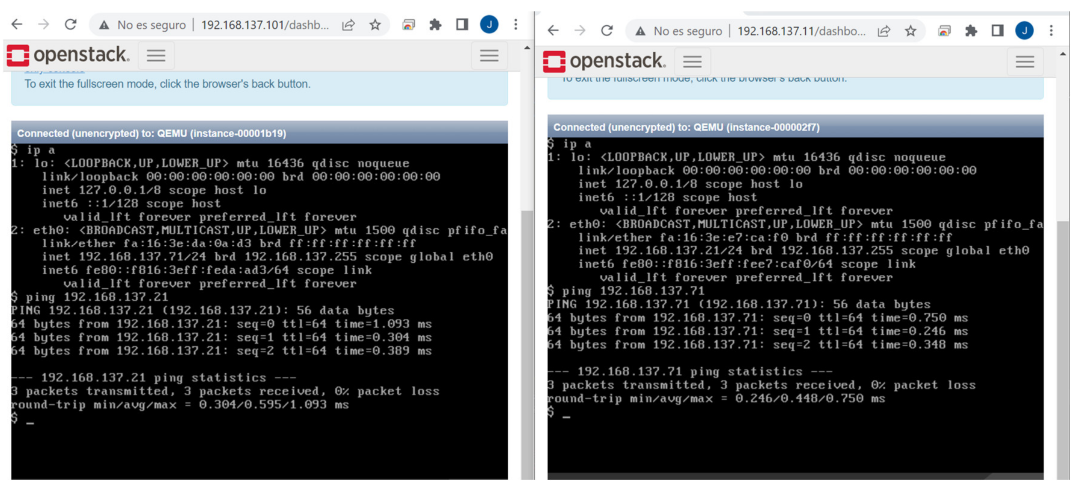


Figure 57: Car VNF and Drone VNF consoles showing successful connectivity.

In this work, we have demonstrated the performance of the orchestration engine in helping to meet the overall and particular objectives defined in this section. Here we have seen how the orchestration engine acts as a key element in the DEDICAT 6G system to enable the correct provisioning of ad-hoc network slices, which can be leveraged to enable 5G/B5G connectivity in actions related to coverage extension. Moreover, this proof-of-concept followed a cloud-based microservice implementation for the DEDICAT 6G platform, where each FC is implemented in lightweight Docker containers, under a gRPC-based communication scheme. This approach brings multiple advantages, such as the ability to be used in the cloud or in edge domains interchangeably with the potential for a high level of scalability and performance. These features can be particularly useful in scenarios where coverage extension is needed in contexts with limited computational and network resources.

4.7 Strategy to Edge Mission Critical Services

4.7.1 Objectives

The objective of AIRBUS is to introduce Mission Critical services into a split environment in order to make critical communication, in the sense of 3GPP MCS standards, available at the Edge plane and delivered "as a service" to enhance reliability of voice communication (MC-PTT), video (MC-Video) and multimedia data exchange (MC-Data) for First Responders.

As described in D4.1, the Mission Critical services will apply on top of decision-making strategy of DEDICAT 6G to deliver appropriate services to First Responders and PPDR users during the response to a crisis.

To respond to a crisis, Public Safety users count on coverage and connectivity for their communication to collaborate. Those are also at the forensic of their safety protection.

The actual solution, even offered through a Cloud infrastructure, is based on a client – server infrastructure. Following section describes how the Mission Critical will be delivered on the Edge and the strategy employed through DEDICAT 6G to make critical communications reliable and efficient.

4.7.2 Mission Critical on the Edge

The future of critical communications will apply on technology and solutions which eases the information sharing and leveraging situational awareness. Those solutions are based on cloud infrastructure available with an Internet connection and novel devices such smartphones, smart glasses, etc.

The usual architecture is described below:

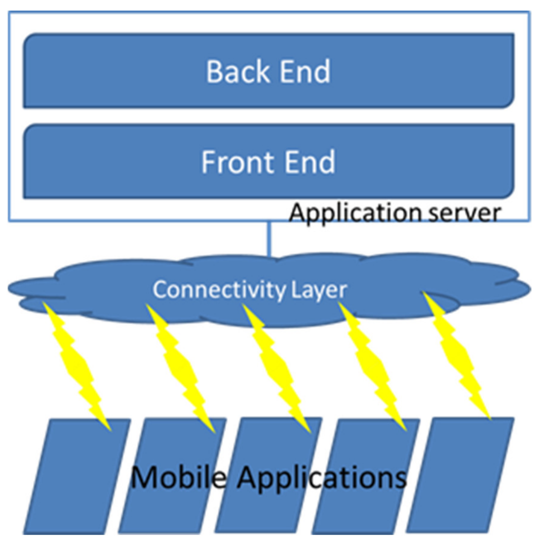


Figure 58: Usual architecture for client-server exchange.

4.7.3 Implementing a full MCS service

The first hypothesis to implement MCS on the Edge was to build a container with all the services embedded and making it available to DEDICAT 6G platform for the deployment at the Edge. Such solution increases the payload of the services for a limited number of subscribers. Some services will be deployed as part of an "all-in-one" solution but not all will be used so, it will impact badly the energy consumption which is not compliant with Energy objective of the project.

The solution considered to implement MCS into DEDICAT 6G is to split each services into functions in order to make only the needed services deployable depending on the situation.

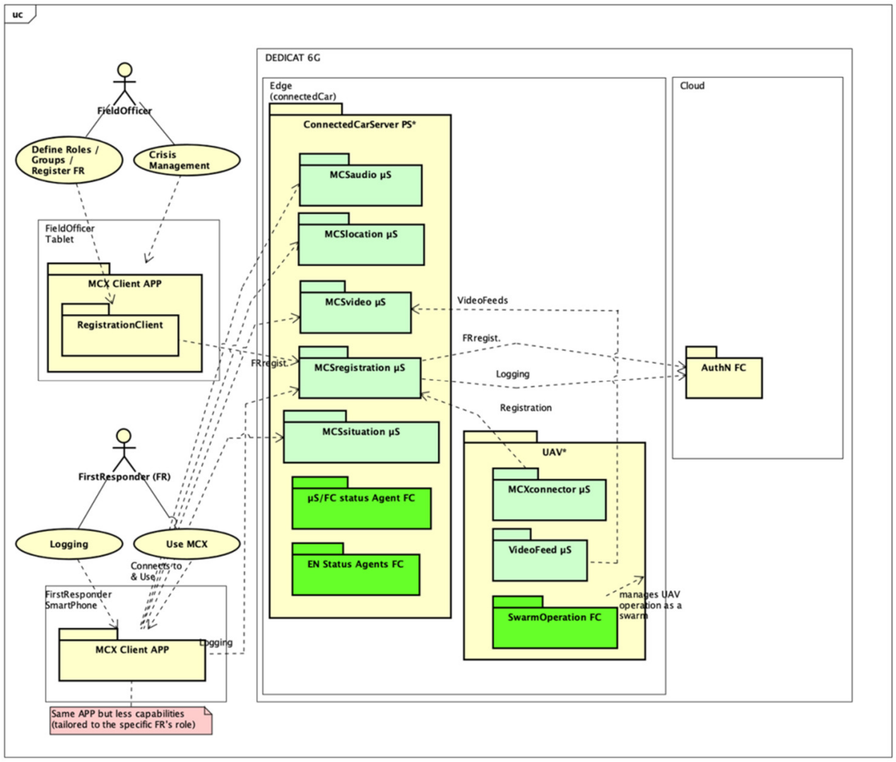


Figure 59: Example of deployment of MCS services during disaster operations.

According to deliverables D2.1 to D2.3, AIRBUS has defined the following main functions split into multiple services:

- Audio function for supporting MC-PTT:

  The audio services support the features to allow push-to-talk, audio broadcast to a group of users.

  - Audio signalization service;
  - Audio media stream service.
- Video function for supporting MC-VIDEO:

  The video services support the real-time video broadcast to a group of users including audio (video stream or video group call).

  - Video signalization service;
  - Video media stream service.
- Data functions for supporting MC-DATA:

The data services support the features to share operational status of each user (available, tasked, "en route", etc.) and mapping with users' positioning.

- Situation service;
- Location service.

- Administration function to support MCS management features:

The administration services support the users' subscription.

  - Registration service.

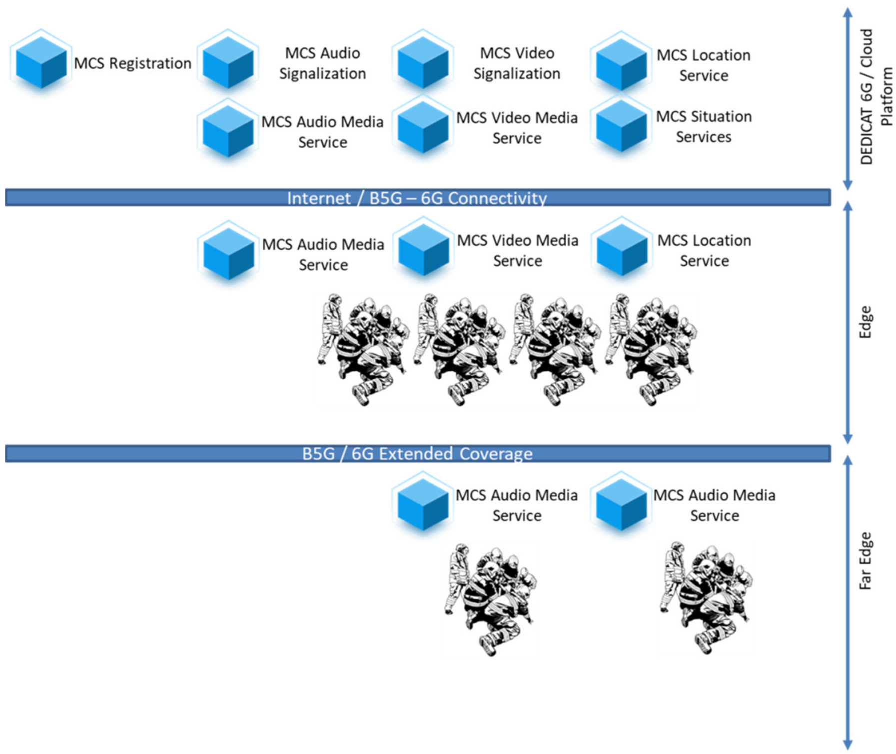


Figure 60: Example of containers deployment depending on needed MCS services at the Far Edge.

The services are provided through Docker containers. Based on the orchestrator strategies, MCS services which has been detected "in failure" or "on limitation" by the DEDICAT 6G Agent Notification, can be deployed by the DEDICAT 6G platform at the Edge or Far-Edge to support First Responders' critical communications during mission management.

AIRBUS has started the evaluation of the deployment based on AWS / OVH. Containers have been deployed from Docker Hub to AWS (or OVH) host.

Client applications have subscribed to MCS Services.

A failure has been simulated on Audio services by killing the MCS Audio Media Service. The Audio service has been automatically deployed by the orchestrator at AWS host and audio connection between client applications has been maintained.

First evaluation and measures will be continued and reported under WP6 work in D6.2.

5 Conclusions

This deliverable reports the activity conducted in the work package 4 on the design and development of mechanisms for the dynamic coverage and connectivity extension through the exploitation of innovative devices (e.g., drones, robots, connected cars, other mobile assets like forklifts in a warehouse, etc.). D4.2 provides a second release of the mechanisms and components for dynamic coverage and connectivity extension, including knowledge building. This deliverable is an extension of D4.1. These innovative components are fuelling the lab integration reported in WP6.

First, D4.2 deliverable describes the DEDICAT 6G functional architecture for Coverage Extension as a Service (CEaaS), highlighting the interactions between the DEDICAT 6G platform and a legacy 5G network infrastructure and provides an overview of the functionalities that the DEDICAT 6G platform will offer for CEaaS. One important feature is the context awareness, the ability to monitor what is happening in the network and the system as a whole, to constantly deduce the current status and whether decision making is required. Another important aspect is related to decision making components, to enable the system to make decisions on the MAPs deployment strategies. This includes making decisions about intelligence distribution (i.e. service placement), network operation (i.e. MAP-UE association and RAT selection) and coverage extension (e.g. MAP and swarm operation).

Second, D4.2 describes in detail several context and situation awareness functionalities (i.e. device discovery, technology recognition and VRUs tracking). These modules provide essential information for making decisions exploiting MAPs.

- Device discovery module is based on a new type of enhanced IoT devices estimating the number of users in a coverage area by short-range signal scanning techniques. Thanks to the information monitored and collected by the IoT sensing nodes, the backend of the micro-service generates detection chart reporting the history of device detections that have been performed by each of the nodes and real-time heat maps of the occupation by connected and non-connected users present in the area of interest.
- Technology recognition module can recognize a wide range of wireless technologies operating in the 5.9GHz ITS band (LTE, 5G NR, WiFi, C-V2X PC5 and ITS-G5) and can use the statistics of each identified technology to estimate and predict the traffic characteristics of each technology and to implement flexible and dynamic spectrum management and utilization.
- VRUs tracking module can track the movement of VRUs on the road or obscured subjects in real-time through camera sensors installed in the roadside unit located at an intersection and warns the driver of a hazardous situation via wireless vehicle-based infrastructure communication.

Third, D4.2 describes and evaluates several network operation decision making functionalities for MAP/UE association. The first option in a centralized UE association and MAP placement making a trade-off between the network cost minimization (number of MAPs) and user utility (spectral efficiency) maximization in a single framework where user utility is both an optimization constraint and optimization sub-objective. The second option is a distributed UE association and MAP placement deciding the number and the optimal location of MAPs, which maximizes the throughput and ratio of well-deserved users while minimizing the number of drones deployed and the execution time in sub6GHz and mm-wave bands. The last option provides heterogeneous MAP-assisted networks with machine learning to maximize the QoS satisfaction level and the energy efficiency by jointly optimizing user association and power allocation under wireless backhaul link capacity constraint in highly crowded areas with heavy traffic loads. Another network operation decision making functionality described in D4.2 is the management and configuration of vehicular based MAPs (i.e. RAT selection and configuration based on MAP capabilities, applications/services traffic demands, and/or identified characteristics of the wireless environment).

Finally, D4.2 also describes coverage extension decision making (i.e. MAP operation managing mobility and swarm operation) and assistance to Intelligence distribution in the coverage extension. Intelligence is every software-based part that can consume computational resources available on the nodes of a B5G network scheme. Three types of intelligence can be defined: (i) DEDICAT 6G FC instances assisting the placement optimization and deploying concrete FCs instances along the network depending on the needs, (ii) NFV Orchestration of VNFs in the instantiation of network slices to enable end-to-end 5G/B5G connectivity and (iii) verticals apps impacting the coverage extension to guarantee the service. Concerning the MAP mobility management, D4.2 proposes an algorithm for robot based MAPs, finding the path or trajectory that each MAP should follow to reach the target position and the selection of nearby docking/charging stations.

In future work, DEDICAT 6G framework will be extended with various types of approaches to address CEaaS, in particular machine learning and AI-based solutions for different use-cases (e.g., smart highway, smart warehousing, public safety and enhanced experience of temporary events). A particular attention will be paid to the management of dynamic networks. System level simulations will provide performance evaluation of coverage and connectivity extension mechanisms. Moreover, an experimental proof of concept will be developed to validate the technology enablers. This software implementation will be tested in a lab environment. This proof of concept can be used to showcase the results and feed them into WP6 for final implementation and integration.

References

[1] P. U. T. Taher Issoufaly, "BLEB : Bluetooth Low Energy Botnet for large scale individual tracking ," International Conference on Next Generation Computing Applications (NextComp), 2017.

[2] Y. C. Naipeng Li, Yuchun Guo, "An Effective Tracking System Based on Wi-Fi Trajectory Matching ," 9th International Conference on Anti-counterfeiting, Security, and Identification (ASID), 2015.

[3] H. Li, E. C. L. Chan, X. Guo, J. Xiao, K. Wu, and L. M. Ni, "Wi-Counter: Smartphone-Based People Counter Using Crowdsourced Wi-Fi Signal Data," IEEE Transactions on Human-Machine Systems, vol. 45, no. 4, pp. 442–452, 2015.

[4] C. J. Bernardos and M. A. Uusitalo, "European Vision for the 6G Network Ecosystem," June 2021.

[5] V. Maglogiannis, A. Shahid, D. Naudts, E. De Poorter, and I. Moerman, ''Enhancing the coexistence of LTE and Wi-Fi in unlicensed spectrum through convolutional neural networks,'' IEEE Access, vol. 7, pp. 28464–28477, 2019.

[6] M. Girmay, A. Shahid, V. Maglogiannis, D. Naudts and I. Moerman, "Machine Learning Enabled Wi-Fi Saturation Sensing for Fair Coexistence in Unlicensed Spectrum," in IEEE Access, vol. 9, pp. 42959-42974, 2021

[7] M. Sana, A. De Domenico, W. Yu, Y. Lostanlen, and E. Calvanese Strinati, "Multi-agent reinforcement learning for adaptive user association in dynamic mmwave networks," IEEE Transactions on Wireless Communications, vol. 19, no. 10, pp. 6520–6534, 2020

[8] E. Catte, M. Sana, and M. Maman, "Cost-Efficient and QoS-Aware User Association and 3D Placement of 6G Aerial Mobile Access Points," in 2022 Joint European Conference on Networks and Communications and 6G Summit (EuCNC/6G Summit), 2022.

[9] G. Gur, "Expansive networks: Exploiting spectrum sharing for capacity boost and 6G vision," Journal of Communications and Networks, vol. 22, no. 6, pp. 444–454, 2020.

[10] F. Provenzan. (2015). Traffic Control—Connected Technologies. Econlite. Accessed: Nov. 27, 2018.

[11] European Road Safety Observatory, European Commission. (2017). Annual Accident Report 2017.

[12] Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. Available online: https://arxiv.org/abs/2004.10934

[13] MEDIATEK, "5G NR and 4G LTE Coexistence: A Comprehensive Deployment Guide to Dynamic Spectrum Sharing," White Paper. [Online]. Available: https://d86o2zu8ugzlg.cloudfront.net/mediatek-craft/documents/Dynamic-Spectrum-Sharing-WhitePaper-PDFDSSWP-031320.pdf.

[14] ETSI, TR 26.849 V12.1.0 Technical Specification Group Services and System Aspects; "Multimedia Broadcast/Multicast Service MBMS improvements, MBMS operation on demand", 2015.

[15] srsRAN. https://www.srslte.com.

[16] H. V. Abeywickrama et al., "Comprehensive Energy Consumption Model for Unmanned Aerial Vehicles, Based on Empirical Studies of Battery Performance," IEEE Access, vol. 6, pp. 58383–58394, 2018.

[17] E. Altman, K. Avrachenkov, and A. Garnaev, "Generalized a-fair resource allocation in wireless networks," in Proc. IEEE Conference on Decision and Control, pp. 2414–2419, 2008

[18] "Al-Hourani, S.Kandeepan, J. Jamalipour, Modeling air-to-ground path loss for low altitude platforms in urban environments, 2014, 2014 IEEE Global Communications Conference

[19] W. Yi, Y. Liu, E. Bodanese, A. Nallanathan and G. K. Karagiannidis, "A Unified Spatial Framework for UAV-Aided MmWave Networks," in *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8801-8817, Dec. 2019, doi: 10.1109/TCOMM.2019.2945332.
[20] H. Shakhatreh, W. Malkawi, A. Sawalmeh, M. Almutiry and A. Alenez, Modeling Ground-to-Air Path Loss for Millimeter Wave UAV Networks, arXiv, 2021, https://arxiv.org/abs/2101.12024
[21] S. Sun *et al*., "Propagation Path Loss Models for 5G Urban Micro- and Macro-Cellular Scenarios," *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, 2016, pp. 1-6, doi: 10.1109/VTCSpring.2016.7504435.
[22] I. Elnabty et al., "A survey on UAV placement optimization for UAV-assisted communication in 5G and beyond networks," Physical Communications, 2022.
[23] M. Conforti et al., Integer Programming, Springer, 2014.
[24] A. Al-Hourani et al., "Optimal LAP altitude for maximum coverage," IEEE Wireless Commun. Lett., Dec. 2014.
[25] ITU-R, "Guidelines for evaluation of radio interface technologies for IMT-2020, " ITU-R Re-port M.2412-0, 2017.
[26] ETSI GS NFV-SOL 006, "NFV descriptors based on YANG Specification". Available Online: https://www.etsi.org/deliver/etsi_gs/NFV-SOL/001_099/006/03.06.01_60/gs_nfv-sol006v030601p.pdf
[27] https://www.docker.com
[28] https://kubernetes.io
[29] https://grpc.io
[30] https://developers.google.com/protocol-buffers
[31] https://www.sqlite.org
[32] https://osm.etsi.org/
[33] https://www.openstack.org/
[34] ETSI GS NFV-SOL 005, "RESTful protocols specification for the Os-Ma-nfvo Reference Point". Available Online: https://www.etsi.org/deliver/etsi_gs/NFV-SOL/001_099/005/03.06.01_60/gs_nfv-sol005v030601p.pdf
[35] H. Lee, S. Vahid, and K. Moessner, "Traffic-Aware Resource Allocation with Aggregation in Hetero-geneous Networks with WLANs," in2018 European Conference onNetworks and Communications (EuCNC), June 2018.
[36] C. T. Cicek, H. Gultekin, B. Tavli, and H. Yanikomeroglu, "Backhaul-Aware Optimization of UAV Base Station Location and Bandwidth Allocation for Profit Maximization," IEEE Access, vol. 8, pp. 154 573–154 588, Aug. 2020
[37] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. The MIT Press, 2001.