AI-fuelled Dimensioning and Optimal Resource Allocation of 5G/6G Wireless Communication Networks

Conference Paper · August 2024

DOI: 10.1109/MeditCom61057.2024.10621426

AI-fuelled Dimensioning and Optimal Resource Allocation of 5G/6G Wireless Communication Networks

P. Papaioannou*, I. Pastellas†, C. Tranoris*, S. Karagiorgou†, S. Denazis*

*Electrical & Computer Eng. Dept., University of Patras, Greece, †UBITECH LTD, Thessalias 8 and Etolias 10, Chalandri, Athens, PC 15231, Greece
{papajohn, sdena}@upatras.gr, {ipastellas, skaragiorgou}@ubitech.eu, tranoris@ece.upatras.gr

***Abstract*—The advent of 5G/6G broadband wireless networks brings several challenges with respect to optimal resource planning and allocation. In a heavily interconnected network of wireless devices, and users along with their equipment, all compete for scarce resources which further emphasizes the importance of fair and efficient allocation of those resources for the proper functioning of the networks. This paper tackles a crucial and timely topic, i.e., understand the various factors involved for optimizing network performance and ensuring fair access for different users, applications and devices. Integrating Machine Learning (ML) and Artificial Intelligence (AI) for predictive dimensioning and pattern mining over the network traffic can enable dynamic and intelligent resource allocation, increase network capacity, enhance the underlying capabilities between users and core network, and better correlate the Quality of Service (QoS). The scientific contribution of this paper entails novel AI models harvesting data from real-world 5G/6G testbeds offered through the AI as a Service (AIaaS) paradigm to enable model reuse and seamless exploitation for different 5G/6G application requirements and learning tasks.**

I. Introduction

A key characteristic of 5G/6G networks is their diverse and ever-changing environment, encompassing several different elements including varying frequency, time, space, power, and user dynamics. This complexity presents a significant hurdle in terms of resource distribution and task scheduling, necessitating a careful trade-off analysis across these factors, along with considerations about the network's condition, traffic behaviors, and user inclinations. As a result, there is a pressing need for optimisation approaches that span multiple dimensions to identify solutions that are either optimal or close to optimal, harmonising the competing goals and limitations inherent to various network components and situations. The ultra-low latency and high-speed capabilities of 5G and the emerging 6G networks enable the collection of large datasets from various sources like Internet of Things (IoT) equipment, sensors, and mobile devices, in real-time. This data is invaluable for training and validating Machine Learning (ML) models, especially those requiring immediate data processing, like optimal resource planning and predictive dimensioning derived from real-time analytics for prompt decision-making. The unique characteristics of 5G, such as network slicing, also facilitate the testing of new algorithms under various network conditions, ensuring robustness and efficiency in diverse application contexts. As we transition to 6G, these capabilities will be further enhanced, offering even greater opportunities for innovative data-driven solutions.

This paper explores how data from advanced experiments performed over a real-world 5G/6G testbed can be used to assess resource usage and plan for optimal resource allocation and optimisation. 5G testbeds play a pivotal role in data generation, acting as rich sources for real-time, high-volume, and diverse data derived by user applications, which are essential for the development and testing of advanced ML models, methods, and algorithms. Alongside real data, 5G/6G testbeds can be used to generate synthetic data, which is particularly useful when real data is scarce, sensitive, or costly to obtain. This synthetic data, crafted through algorithms or artificial environments (e.g., digital twins or simulations), can help in training ML models without compromising privacy or incurring high costs. By providing both real and synthetic data under specific scenarios, 5G/6G testbeds are a valuable tool for Artificial Intelligence (AI), since scenario-based data collection can be performed not only via real equipment but also via different real-world application scenarios simulation, such as urban traffic patterns, over-the-top (OTT) media services, industrial automation processes, or smart city applications. By doing so, they can generate and collect data that mirrors real-life situations, allowing for the analysis of complex, dynamic and heterogeneous systems. What is also important is that testbeds can generate annotated data and patterns that sometimes are really difficult to collect and identify on an operator's production network.

This work injects the AI methods and mechanisms to 5G testbeds to perform:

- Optimal resource planning, since the resources of a production network are limited and usually more expensive and more advanced than a 5G testbed, where the latter also allows for various degrees of experimental freedom;
- Real-time data analytics derived by predictive and prescriptive learning tasks, which is crucial for applications requiring immediate response;

979-8-3503-0948-5/24/$31.00 ©2024 IEEE

- Data correlation and multi-domain integration, since a 5G/6G testbed can facilitate the fusion of data from multiple application domains, enabling a holistic view of complex ecosystems. This integration can reveal correlations and interdependencies that are not apparent when data silos exist; and
- Customisable network slicing for diverse requirements, thus allowing the creation of multiple virtual networks with different characteristics, allowing parallel testing and experimentation.

This work presents a 5G testbed, logging monitoring metrics and feeding an analytics pipeline for deriving insights along with our early benchmarking results. The analytics pipeline is exposed with the as a Service model, thus as an AI as a Service (AIaaS) capability, to enable reuse and seamless exploitation of the provided AI mechanisms.

The rest of the paper is organized, as follows: Section II provides the current literature review around 5G networks dimensioning, optimisation and resource allocation using AI and ML models. Section III presents the 5G testbed setup, configuration and data generation. Section IV, presents our analytic pipeline architecture, while Section V discusses our first experimental results, and Section VI concludes the paper and provides research ideas to be pursued in the future.

II. Related work on 5G Communication Networks Dimensioning and Optimal Resource Allocation

Resource planning for 5G networks has been a subject for research since the first deployments of such networks. The ability to fully take advantage of the available resources is a major interest for providers so that the maximum customer satisfaction can be achieved while minimising CApital expenditure (CAPEX) and OPerating EXpenses (OPEX). The research as expected is more oriented to commercial network deployments and not targeted for research testbeds. In this regard, Waleed et al. [1] identify and present solutions regarding the resource allocation task of 5G networks by satisfying often competing objectives and constraints (i.e., throughout, CPU usage and memory use), while maintaining the Quality of Service (QoS) required by the network users. This is all presented in the context of Cloud Radio Access Network (CRAN), an architecture that separates base station functions into two main components: the Centralized Processing unit (CP) and the distributed remote radio unit (RRU), which allows for centralized control and coordination leading to improved resource utilisation and efficient network management. A similar survey on commercial CRAN networks has been performed in [2] focusing on the power consumption and harvesting the baseband unit (BBU) computational resource, capacity, and wavelength parameters.

Kamal et al. [3] performed a literature review on the 5G resource allocation problem trying to get insights on questions ranging from 5G challenges and resource allocation importance to resource allocation algorithms and metrics used in such approaches. The authors state that resource allocation is an important aspect on 5G systems, since it allows systems to be more dynamic and satisfy diverse users' requirements. They analyze work from other authors to find the common ground on algorithms and parameters used. On a different approach, the authors in [4] have a more hands-on approach, by defining the resource allocation problem through mathematical terms considering three kinds of resources (i.e., bandwidth, cpu usage and memory), while trying to accommodate three different kinds of services requiring either bandwidth as in enhanced mobile broadband (eMBB), latency as in ultra-reliable and low-latency communications (uRLLC), and finally massive machine-type communications (mMTC). They also performed simulation tests and presented their algorithms and results.

Moscholios et al. [5] presented a special issue to bring together the state-of-the-art research contributions addressing the challenges of contemporary 5G communication networks design, dimensioning, and optimisation, computing resources, and services.

Jayaraman et al. [6] proposed a modified Resource Allocation (RA) scheme using a learning-based Resource Segmentation (RS) algorithm. Their algorithm uses a modified Random Forest Algorithm (RFA) with Signal Interference and Noise Ratio (SINR) and Position Coordinates (PC) to obtain the location of end-users. It further predicts Modulation and Coding Schemes (MCS) for establishing a connection between the end-user device and the Remote Radio Head (RRH).

Abdellatif et al. [7] proposed a dynamic network slicing and resource allocation framework that maintains high-level network operational performance, while fulfilling diverse services' requirements and Key Performance Indicators (KPIs), e.g., availability, reliability, and data quality. They introduced a novel methodology and resource allocation schemes, that enable high-quality selection of radio access points, resource allocation, and data routing from end users to the cloud.

Zhao [8] proposed an energy-efficient resource allocation method by using a Deep Reinforcement Learning (DRL) method, which enhances the network performance and reduces the network operational cost. The method considers the connection relationship between base stations, users and the transmission power allocated by base stations to users as decision variables, maximising the overall energy efficiency, and taking the needs of mobile users as constraints, to guarantee advanced QoS.

Munaye et al. [9] proposed a method which enables the improvement of the IoT connectivity related with the network slicing concepts. The allocation of resources is based on individual network slices specified as audio, texting, video, and browsing. Then, to maximize the average resource allocation performance, they trained a DRL to optimally allocate network slices to the users, balancing resource blocks, and support QoS for fair resource allocation.

Laboni et al. [10] formulate the resource allocation task in 5G mobile edge computing (MEC) as a multi-objective problem through a mixed-integer non-linear programming problem. They adopt a hyper-heuristic algorithm by leveraging the combined powers of Sine-Cosine, and other optimisation algorithms. Their algorithm works at the higher level, and it exploits one of the three lower-level heuristics in each iteration to efficiently capture the dynamically varying environmental parameters, and thereby solve the resource allocation problem. This helps to achieve a global optimum in allocating resources of a 5G MEC network.

Pamba et al. [11] develop a resource allocation model using a novel deep learning algorithm for optimal resource allocation. Their algorithm is formulated using the constraints associated with optimal radio resource allocation. The objective function design aims at reducing the system delay. Their study predicts the traffic in a complex environment and allocates resources accordingly. Their results show an improved rate of allocation compared to the other methods.

Existing works are limited either by addressing medium allocation characteristics, or computational resources allocation. Compared to the above-mentioned approaches, the scientific contribution of this work is the proposition of a resource allocation predictive model using Deep Learning to correlate traffic sources from network applications with the optimal resources to be allocated to enhance the core network scheduling capabilities.

III. 5G Testbed Setup and Conducted Experiments

Our experiments have been performed at the University of Patras 5G facility, an academic isolated non-public 5G infrastructure. This facility offers end-to-end support to various application verticals using the lab for experimentation.

A. Testbed Description

The Network Slice as a Service (NSaaS) delivery model is supported, where custom network slices are provided for evaluation of both equipment and software solutions, as well as KPIs. Entry point to this facility is the web portal [12] of the open-source Operations Support System (OSS) Openslice [13]. Telemetry and monitoring is supported to all layers of the experiment flow to offer comprehensive metrics ranging from Radio Access Network (RAN) measurements to cloud-related metrics. Prometheus [14] is used for storage and data persistence, and Grafana [15] for dashboards and graphical monitoring of the results. Custom made data collectors are also deployed in the lab. On the 5G network front, various installations and configurations are deployed and supported. For further information, please refer to the Patras 5G Wiki [16].

B. Experiments Description

We present the initial experimental results performed in Q4 2023 and Q1 2024 in the above mentioned 5G Testbed. The experiments were of three distinct categories, as detailed below.

1) Common Aspects of Experiments: As shown in Figure 1, some aspects are common in all experiments. The basic setup for each experiment was the deployment of a 5G network with core and RAN elements. For each experiment the corresponding Network Application (NetApp) has been deployed on a Virtual Machine (VM), whilst the 5G core network ensured the connectivity between the devices used in the experiments and

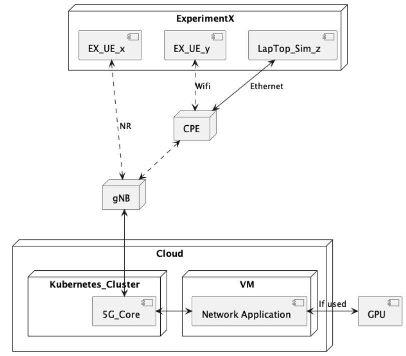

Fig. 1. Testbed Overview and Setup

the corresponding NetApps. Depending on the NetApp and the experiment requirements, GPU access was also provided per case.

To access the network, various UEs were used, such as 5G-enabled mobile phones. For cases where the end user device was not 5G ready, a Customer Premise Equipment (CPE) was used as a gateway by providing 5G back-haul access and allowing legacy devices to connect, either through Ethernet or WiFi networks. Finally, in cases where simulation was required, either laptops or properly configured VMs were deployed that used the above mentioned CPE, as a gateway to use the 5G core network.

The 5G core network deployed can be configured to satisfy various requirements of the 5G tested scenarios. By modifying the configuration parameters, the deployed network can be oriented in downlink configuration. This means that more resources are allocated for the downlink path (gNb to UE), and therefore the downlink bandwidth of the network can be maximised. If the experiment requires the UE to send data as fast as possible (via the uplink), the correct configuration can be applied to satisfy this, as well. Finally, a low latency configuration is also available, if required, depending on the configuration of the 5G network. For this purpose, different bandwidth measurements have been recorded, as shown in table I:

Bandwidth (Mbps)		
	Downlink Config	Uplink Config
Downlink bw	700	200
Uplink bw	30	150

TABLE I
5G Testbed Experiments

2) Experiment 1: It has focused on a Public Protection and Disaster Relief (PPDR) scenario where the end user devices stream video via mobile phones. This experiment used Real-time Transport (RTP) protocol and the stream was directed to the corresponding NetApp. An uplink gNodeB configuration was used and no GPU was used. Various tests were performed using 1, 2 and 3 mobiles phones simultaneously with various configurations on the end user device application side.

3) Experiment 2: It has focused on a PPDR scenario, where the NetApp streams video to the end user devices over UDP

Fig. 2. Experiment 3 - Sample Results

Fig. 3. Monitored Data Overview

protocol. The device was not 5G enabled and it used the CPE to connect to the 5G network. Downlink gNodeB configuration was used for this experiment and various tests were performed by modifying the configuration of the NetApp, which made use of the available GPU. Simulation tests were also performed for this experiment by substituting the end user device with a laptop configured to receive the stream and connected with the CPE.

4) Experiment 3: It has focused on a PPDR scenario, where the end user device sends data to the NetApp using the TCP protocol. Simulation tests were performed by configuring a VM to act as the end user device and transmitting the required data through the CPE. Uplink gNodeB configuration was used, and the NetApp made use of the available GPU. Figure 2 shows some results obtained during testing.

5) Experiments Overview: In Table II, we enlist the basic characteristics of the performed experiments. For each experiment, we present the traffic type and the expected direction of the traffic according to the UC. Regarding the 5G network configuration, the tested gNb configuration is stated, as well as whether various gNb configurations were used in the experiments. Finally, the type of UEs used are listed (real or simulated).

Experiments Description						
Experiment			Configuration		User Device	
ID	Type	UL/DL	Config	Multiple	Phys	Sim
1	RTP	UL	UL	Yes	Yes	No
2	UDP	DL	DL	No	Yes	Yes
3	TCP	UL	UL/DL	No	No	Yes

TABLE II
PERFORMED EXPERIMENTS

C. Data Description

The monitoring functionality of Patras facility allows for gathering and storing of various metrics during the experiments execution. In this work, we focus on specific metrics gathered from the 5G cell, as well as metrics reported by the used devices. The monitoring dashboard is depicted in Figure 3.

1) Cell-related Metrics: We gather and analyse the bitrate reported from the gNb in both the uplink and downlink directions. This is reported in bits per second and measured in the physical layer of the 5G connection. Both values are gathered regardless of the traffic direction each experiment is focusing on. It is worth mentioning that none of these experiments reached the maximum capacity of the deployed 5G network.

2) UE-related Metrics: In 5G connections, the end user devices report various metrics that show the connection status between the device and the gNb. The metrics gathered and taken into consideration in this work are, the: (i) uplink and downlink bitrate; and (ii) MCS. MCS defines the number of bits carried by every symbol during transmission and is an indicator of the signal quality. In general, high MCS means high quality and higher bitrates. Signal to Noise Ratio (SNR) for the uplink direction is also taken into account, as the Channel Quality Indicator (CQI) reported by the UE to the gNb indicating the quality of the connection between them from the UE perspective.

It should be noted that in the experiments where multiple 5G devices have been deployed, the metrics from a single UE have been used as an indicator.

IV. Data Analytics Pipeline Architecture

An essential part of learning over the monitored data is the Data Analytics Pipeline. In this section, we provide the phases of the data analysis. The analytics pipeline serves two purposes: (i) harvest the telemetry data to predict workloads and decision making related to resource allocation, slicing and service orchestration; and (ii) expose Application Programming Interfaces (APIs) to support the AIaaS paradigm. By exposing the learnt patterns via AIaaS APIs enables to adapt the configuration of the 5G testbed to optimise its functions. The Data Analytics Pipeline is presented in Figure 4. As it is depicted, we harvest telemetry 5G data and we further feed a data exploration and analytics pipeline to learn patterns about the experimental NetApps and their workload class. Last, the trained AI models are being exposed via dedicated AIaaS APIs to serve the learnt patterns, and thus support a set of predictive tasks.

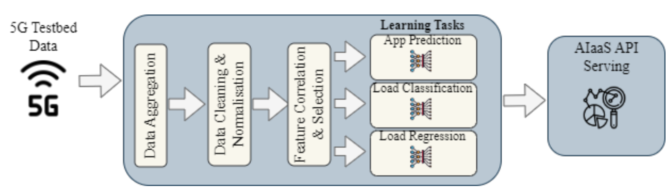

Fig. 4. AIaaS Pipeline (training and serving phase)

A. Data Aggregation

We aggregate all the datasets from the experiments into a single data structure, and add a new column specifying the type of experiment. This new column is then used as the label for the analysis in Section V.

B. Data Cleaning and Normalisation

Towards data cleaning, we replace all NaN values with zero (0), since all of them are numeric values that vary at time. Also, we remove irrelevant columns / features (metadata of Prometheus, features where there is only one value in the entire dataset, etc.).

We then normalize the data by using the standard scaling method (i.e., subtract by the mean value and divide by the standard deviation), and in some specific cases (such as in the Deep Neural Network (DNN) modeling), we normalize by the Min Max normalisation method.

C. Feature Correlation and Selection

Only features related to the cell uplink / downlink bitrates, UE devices and details about the gNb service state are kept. Feature correlation is then used as a statistical measure to understand the relationship between the 5G network features. This measure expresses how much two features change together. A positive correlation means that as one feature increases, the other feature also increases. A negative correlation means that as one feature increases, the other feature decreases. A correlation of zero means that there is no relationship between the two features. Feature correlation is used to ease feature selection and extract the most important features for the next step, which is data modelling.

D. Data Modelling and Predictive Tasks

We train three (3) AI models. The first supports predictive workloads and forecasts future values of the 5G network load; the second one supports binary classification to forecast the class of the load (i.e., currently low-high classes are supported); and the third one supports NetApps category prediction based on the experiments type.

The predictive workloads facilitate to optimally allocate the right resources based on the future workload values. The load classification facilitates the decisions related with the service orchestration, optimal slicing, and vertical scaling when needed according to the load classes. Last, the NetApps category prediction enables to forecast the expected type of application in order to allocate resources according to its requirements linked with the necessary QoS.

We split the data into training (80%) and test sets (20%) to evaluate the performance accuracy of the models. To predict workload values on future 5G network load, we trained several Machine Learning models from conventional ML models (i.e., Random Forest, Logistic Regression, XGBBoost) to Deep Learning models (a feed-forward Deep Neural Network (DNN), and a Long Short-Term Memory (LSTM) neural network followed by a DNN). While for all models we use a 1D vector as input X, in the case of LSTM we use a 2D input that are windows of features in previous steps. The LSTM model takes as input a window of features X and produces output Y, which is the prediction of the next value of Y. To forecast the class of the load, we trained an LSTM neural network that takes as input a window of features X and produces output Y, which is the prediction of the class of Y. Last, to forecast the type of application, we trained an LSTM neural network that takes as input a window of features X and produces output Y, which is the prediction Y of the NetApp category. The size of the time window used in all AI models training is 30s, having as input a 2D array with the length of data in 30s frames. Other models, receive as input 1D array (with the number of features as dimension) and outputs the class of load, the class (i.e., category) of traffic type, or the actual UL/DL bitrate depending on the task.

V. Experimental Results

This section presents the experimental results on the telemetry data collected by the 5G testbed. We evaluate the performance of the AI models by means of accuracy, precision, recall and Mean Squared Error (MSE) according to the learning task.

A. Preliminary Analysis on 5G Datasets

By using the telemetry datasets coming from the 5G testbed, we use the Data Analytics Pipeline to train the AI models and expose the AIaaS predictive tasks as APIs. We also benchmark over different ML and DL methods to measure their performance by means of accuracy, precision, recall, and MSE based on the learning task. More specifically, we use different ML and DL models to achieve three tasks. A regression task, where we predict the future value of the 5G cell uplink / downlink bitrate towards QoS prediction. A binary classification, where we predict the class of the load (i.e., currently low-high classes are supported). Finally, a classification task, where we predict the type of the traffic / application (i.e., NetApp) that allocated resources in each experiment to understand its behavioural pattern.

B. Analytic Tasks and Results

In table III, we present the evaluation measures of the above-mentioned algorithms. The results are derived by using the predictions of each model to assess the performance (i.e., in the classification or regression tasks) of the test set. We also measure the training time required to let the model converge to a high accuracy score. We observe that most of the algorithms perform very well on predicting the NetApp type using the 5G dataset. More specifically, Table III shows that Random Forest, XGBBoost, and LSTM with DNN achieve higher than 95% accuracy along with precision and recall. Random Forest achieves the best overall performance (96.5%), and also has the lowest training time.

The same approach was adopted to gather the experimental results for the CPU load classification related with the Experiment 1. As it is depicted in Table IV, Random Forest achieves the highest performance with 97.7%.

Classification (Traffic/Application Type)				
Algorithms	Accuracy	Precision	Recall	Time
Random Forest	96.5	96.6	96.4	1.01
Logistic Regression	73.7	75.2	73.7	0.58
XGBBoost	95.4	95.5	95.4	3.65
DNN	82.0	82.0	82.0	56.23
LSTM+DNN	95.11	95.13	95.11	61.01

TABLE III
APPLICATION CLASSIFICATION - RESULTS IN % AND SECS

Classification (CPU Load for Experiment 1)				
Algorithms	Accuracy	Precision	Recall	Time
Random Forest	97.7	97.7	97.7	0.37
Logistic Regression	71.0	71.6	71.0	0.05
XGBBoost	94.0	94.0	94.0	1.27
DNN	86.5	87.0	86.5	20.68

TABLE IV
LOAD CLASSIFICATION - RESULTS IN % AND SECS

The tables that are following present the MSE of the test set for predicting the Uplink and Downlink bitrate for each of the experiments, respectively. Again by following a similar approach, we evaluate the different algorithms by predicting the next value of the bitrate in every direction. While all algorithms perform well (i.e., with very low MSE), LSTM+DNN model has the best overall performance, especially in cases where other models struggle to predict the next value. For instance, in Experiment 3 of Table V, the next value of the Uplink bitrate is predicted with very low error.

Bitrate Regression (Experiment 1)		
Algorithm	Uplink	Downlink
Random Forest	0.00016	0.00017
DNN	0.002	0.0009
XGBBoost	0.0007	0.0004
LSTM+DNN	0.00021	0.00008
Bitrate Regression (Experiment 2)		
Algorithm	Uplink	Downlink
Random Forest	0.00084	0.00084
DNN	0.00085	0.00085
XGBBoost	0.00084	0.00084
LSTM+DNN	0.000095	0.00016
Bitrate Regression (Experiment 3)		
Algorithm	Uplink	Downlink
Random Forest	0.022	0.00036
DNN	0.022	0.00053
XGBBoost	0.022	0.00039
LSTM+DNN	0.0043	0.00014

TABLE V
BITRATE NEXT VALUE - RESULTS IN MSE

VI. CONCLUSION

This paper presents a resource allocation predictive model using Machine Learning and Deep Learning to correlate traffic monitored from NetApps with the optimal resources to be allocated to enhance the core network scheduling capabilities. It presents the 5G testbed, the experiments conducted and the datasets, as well as the Data Analytics Pipeline accompanied by a set of learning tasks. We also present the AI models that we trained to collect the first experimental results as an AI as a Service (AIaaS) capability. The aim is to increase the reuse of the network resources, to optimally allocate resources based on the NetApp category and to support the seamless exploitation of the provided AI mechanisms.

In the near future, we plan to enrich the variety of predictions with more complex 5G NetApps and learning tasks. Enriching the AI models to learn additional patterns from the 5G applications besides PPDR scenarios, like streaming scenarios either in the traditional streaming scenario (downlink traffic) or User Generated Content (UCG) scenarios that focus on uplink traffic, will enable to update them with near real-time resource allocation characteristics.

ACKNOWLEDGMENT

The work of the authors has been partially supported by the Horizon Europe Projects ACROSS (Grant Agreement No. 101097122), FIDAL (Grant Agreement No. 101096146), and INCODE (Grant Agreement No. 101093069).

REFERENCES

[1] W. Ejaz, S. K. Sharma, S. Saadat, M. Naeem, A. Anpalagan, and N. A. Chughtai, “A comprehensive survey on resource allocation for cran in 5g and beyond networks,” *Journal of Network and Computer Applications*, vol. 160, p. 102638, 2020.

[2] L. O. Chenke, C. Nche, É. M. D. Djomadji, and E. T. Bety, “Review of 5g c-ran resource allocation,” *EAI Endorsed Transactions on Mobile Communications and Applications*, vol. 7, no. 4, pp. e5–e5, 2022.

[3] M. A. Kamal, H. W. Raza, M. M. Alam, M. M. Su’ud, and A. b. A. B. Sajak, “Resource allocation schemes for 5g network: A systematic review,” *Sensors*, vol. 21, no. 19, p. 6588, 2021.

[4] Y. Fu, X. Wang, and F. Fang, “Multi-objective multi-dimensional resource allocation for categorized qos provisioning in beyond 5g and 6g radio access networks,” *IEEE Transactions on Communications*, 2023.

[5] I. D. Moscholios, M. Głabowski, P. G. Sarigiannidis, and M. D. Logothetis, “A special issue on modeling, dimensioning, and optimization of 5g communication networks, resources, and services,” p. 1859, 2022.

[6] R. Jayaraman, B. Manickam, S. Annamalai, M. Kumar, A. Mishra, and R. Shrestha, “Effective resource allocation technique to improve qos in 5g wireless network,” *Electronics*, vol. 12, no. 2, p. 451, 2023.

[7] A. A. Abdellatif, A. Mohamed, A. Erbad, and M. Guizani, “Dynamic network slicing and resource allocation for 5g-and-beyond networks,” in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2022, pp. 262–267.

[8] S. Zhao, “Energy efficient resource allocation method for 5g access network based on reinforcement learning algorithm,” *Sustainable Energy Technologies and Assessments*, vol. 56, p. 103020, 2023.

[9] Y. Y. Munaye, R.-T. Juang, H.-P. Lin, G. B. Tarekegn, D.-B. Lin, and S.-S. Jeng, “Radio resource allocation for 5g networks using deep reinforcement learning,” in *2021 30th Wireless and Optical Communications Conference (WOCC)*. IEEE, 2021, pp. 66–69.

[10] N. M. Laboni, S. J. Safa, S. Sharmin, M. A. Razzaque, M. M. Rahman, and M. M. Hassan, “A hyper heuristic algorithm for efficient resource allocation in 5g mobile edge clouds,” *IEEE Transactions on Mobile Computing*, vol. 23, pp. 29–41, 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:252825649

[11] R. V. Pamba, R. Bhandari, A. Asha, and A. S. Bist, “An optimal resource allocation in 5g environment using novel deep learning approach,” *J. Mobile Multimedia*, vol. 19, pp. 1331–1356, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:260946132

[12] Patras 5g. [Online]. Available: https://patras5g.eu

[13] Etsi software development group for openslice (sdg osl). [Online]. Available: https://osl.etsi.org

[14] Prometheus-monitoring system. [Online]. Available: https://prometheus.io/

[15] Grafana: The open observability platform — grafana labs. [Online]. Available: https://grafana.com/

[16] Patras 5g wiki. [Online]. Available: https://wiki.patras5g.eu