6G BRAINS Deliverable D4.3

AI driven Resource Allocation for High Dynamic Ultra-dense D2D Network

Editors:	Israel Koffman, RunEL and Prof. Huiyu Zhou, ULEIC
Deliverable nature:	Report
Dissemination level: (Confidentiality)	Public PU
Contractual delivery date:	30.09.2023 (M33)
Actual delivery date:	12.10.2023
Suggested readers:	Public
Version:	1.0
Total number of pages:	74
Keywords:	IAB, Reinforcement Learning, Cell Free, Grant free, Device to Device, AI

Abstract:

This document updates the progress made in the different research areas related to 6G Access Architecture performed in the 6G BRAINS project. The updated research areas in this document include:

- Artificial Intelligence (AI) based scheduler for a 5G Cell-Free (CF) networks with Integrated Access and Back-haul (IAB)
- AI based Grant-free NOMA for massive Machine-Type Communication Integrated with Intelligent beamforming
- Hybrid user terminal modelling and Resource Allocation for the cell-free D2D network
- Intelligent IAB with Beam Steering based on User Location
- Advanced Test and Simulation Tools supporting 6G BRAINS research

Disclaimer

This document contains material, which is the copyright of certain 6G BRAINS consortium parties, and may not be reproduced or copied without permission.

In case of Public (PU):

All 6G BRAINS consortium parties have agreed to full publication of this document.

In case of Restricted to Programme (PP):

All 6G BRAINS consortium parties have agreed to make this document available on request to other framework programme participants.

In case of Restricted to Group (RE):

All 6G BRAINS consortium parties have agreed to full publication of this document. However this document is written for being used by <organisation / other project / company etc.> as <a contribution to standardisation / material for consideration in product development etc.>.

In case of Consortium confidential (CO):

The information contained in this document is the proprietary confidential information of the 6G BRAINS consortium and may not be disclosed except in accordance with the consortium agreement.

The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the 6G BRAINS consortium as a whole, nor a certain part of the 6G BRAINS consortium warrant that the information contained in this document is capable of use, nor that use of the information is free from risk, accepting no liability for loss or damage suffered by any person using this information.

The EC flag in this document is owned by the European Commission and the 5G PPP logo is owned by the 5G PPP initiative. The use of the EC flag and the 5G PPP logo reflects that 6G BRAINS receives funding from the European Commission, integrated in its 5G PPP initiative. Apart from this, the European Commission and the 5G PPP initiative have no responsibility for the content of this document.

The research leading to these results has received funding from the European Union Horizon 2020 Programme under grant agreement number 101017226 *– 6G BRAINS – H2020-ICT-2020-2. The content of this document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.*

Impressum

[Full project title]	Bring Reinforcement-learning Into Radio Light Network for Massive Connections
[Short project title]	6G BRAINS
[Number and title of work-package]	WP4: AI Driven High Dynamic Ultra-dense D2D Cell Free Network Architecture
[Number and title of task(s)]	T4.2: Multi-spectrum Links enabled Intelligent IAB and AI driven Beam Scheduler (Leader-REL) T4.3: AI Driven Grant-free NOMA Coordinated with OMA Medium Control with Underlaid D2D Clusters (Leader: ULEIC) T4.4: AI Enabled User Terminal Modelling for the D2D-enhanced Resource Allocation Approach (Leader: ULEIC)
[Document title]	D4.3: AI driven Resource Allocation for High Dynamic Ultra-dense D2D Network
[Editor: Name, company]	Israel Koffman, RunEL and Prof. Huiyu Zhou, ULEIC
[Work-package leader: Name, company]	Israel Koffman, RunEL
[Estimation of PM spent on the Deliverable]	10 PM

Executive summary

6G-BRAINS project explored different architectures and essential modules for B5G and 6G ultra dense and dynamic cellular networks. In the document we describe the design of an Artificial Intelligence (AI) based scheduler for a 5G Cell-Free (CF) Networks with Integrated Access and Backhaul (IAB) and intelligent beam steering. The ultra-dense dynamic CF Network includes a cluster of Device to Device (D2D) User terminals (UE) and human-centric control interfaces.

This document reviews, summarizes and concludes the progress made in the research areas related to B5G and 6G access architecture which were achieved in the final 6G BRAINS project reporting period.

The updated items in this document include:

- Artificial Intelligence (AI) based scheduler for a 5G Cell-Free (CF) networks with Integrated Access and Back-haul (IAB)
- AI based Grant-free NOMA for massive Machine-Type Communication Integrated with Intelligent beamforming.
- Hybrid user terminal modelling and resource allocation for cell-free D2D network
- Intelligent IAB with Beam Steering based on User Location
- Advanced Test and Simulation Tools supporting 6G BRAINS research.

The Cell free Integrated Access and Backhaul scheduler is subdivided into two modules, namely: (1) Resources Allocation module, (2) Routing module.

The IAB Resource Allocation finds the optimal way to divide the spectrum between the wireless backhaul and the access requirements of the different Donors and the Nodes in the Network using supervised learning AI method with CQI, DL and UL profile and connected base-station for input parameters and the rate each link must support and its efficiency for cost function. AI based Scheduler considers two approaches, namely: Cloud-based (centralized) and distributed based. The research seeks to determine and quantify the trade-off between the benefits of the two approaches. The Supervised Learning model architecture uses 2 sub-models, namely: Big model: is implemented in the IAB-donor (cloud based) and divides the total bandwidth resources between all IAB (including itself) components (IAB Donor and Nodes); Small model: is implemented at the IAB components and allocates the bandwidth resources between the access (UL and DL links of the UEs) and the backhaul (UL and DL links to the served Nodes). This Supervised Learning AI Model then requires training.

The Routing part is based on Multi-Agent Deep Reinforcement Learning (MA-DRL) for a fully synchronized time-slotted wireless network with the objective to find the optimal route from each BS (Donor or Nodes) to each user in terms of: Packet Error Probability (PEP) for the whole packet trajectory; Maintenance of Quality of Service (QoS) requirements; Network Congestion Management including Queue management and Fairness. Two approaches have been examined in our previous report (D4.1), namely: a centralized model using the Dijkstra's algorithm; a decentralized solution using Q-Routing algorithm; to determine their benefits. Here we designed a decentralized routing algorithm called MA-RAC that learns from experience while interacting with a simulated 5G environment that outperforms the Q Routing scheduler previously introduced. This deliverable also examines how network load affects the performance of different algorithms in various mobility scenarios. We evaluate each algorithm's resilience under different loads for various mobility scenarios. Following this, we conclude that an increase in load / mobility affects network performance significantly for most of the proposed routing algorithms.

Efficient users clustering and power assignment is presented for a dynamic cell-free network that can cluster NOMA users over the limited spectrum resources and allocate optimal transmit power to each user in the cluster. The clustering finds a group of NOMA-based D2D users that can be scheduled on the same resource blocks forming a cluster. Parent points are distributed following a homogeneous Poisson Point Process, where every parent point (RmUE) is uniformly distributed in a certain area and then offspring points (mUEs) around one parent point to form a cluster.

Deep Reinforcement Learning (DRL) is used to predict the power, mobile users' frequency, users' location or beam trajectories, where the environment provides a reward to the agent for every interaction and the agent aims to select the right action for the next interaction to maximize the discounted reward over a time horizon. The goal is to design a DRL system that jointly optimizes the clustering of UEs and the beamforming vectors to obtain the optimal beamforming vector that maximizes the throughput and minimizes loss.

Artificial Intelligence (AI) is incorporated with beamforming and millimetre Wave (mmWave) enabling intelligent beam steering based on user locations.

Deep Learning Integrated Reinforcement Learning (DLIRL) algorithm was carried out for the beamforming solution to overcome the problems associated with mmWave such as blockage impacting the coverage, reliability of highly mobile links, latency overheads associated with the highly mobile users in dense mmWave deployments requiring frequent hand-offs.

The DLIRL beamforming or beam steering tracks the user locations who move at a speed of 25 km/hour and based on its location changing in terms of latitude and longitude to perform beam steering towards the moving user. The beam steering experiment was implemented using MATLAB 2021a and data for DNN training was obtained via ray tracing with MATLAB's Site Viewer. The training for DLIRL is based on the impulse response of the received signal at the coordinated Base Stations from isotropic transmissions from user equipment, which is jointly received at the coordinating IABs, and was used for performing training out of the useful information on the surrounding environment from its interaction with the surrounding environment. The beamforming power control and interference coordination is jointly carried out at the IAB donor and nodes to enhance the performance of the 5G network.

We are aware that telecommunications service providers have an urgent need to reduce operational costs (OPEX) while supporting the rapid introduction of new services and products and identifying and leveraging monetization opportunities. AI/ML is a powerful technology to support these needs.

6G BRAINS recognizes that transition to an industrialization phase and enabling mass adoption of AI/ML, dedicated AI/ML research in specific Cellular Network modules is further required.

The research results and conclusions provided in this document show the potential benefits of using these technologies in 6G building blocks including Schedulers, Integrated Access and Backhaul, Device to Device Networks, Beam Forming and Human/Machine Interfaces.

Furthermore, it is our opinion that AI/ML should be adopted at additional levels of the network architecture.

List of authors

Company	Author	Contribution
RunEL	Israel Koffman, Baruch Globen	Editor, Reviewer Chapter 1 and Chapter 2
Brunel University	Geoffrey Eappen, Ali Mahbas, John Cosmas, Raj Nilavalan	Chapter 5
Leicester University	Guoqing Xia, Bohan Li and Huiyu Zhou	Editor, Chapter 3, Chapter 4 and Chapter 7
VIAVI	Ata Sattarzadeh Hashemi, Chi-ming Leung	Chapter 6
Eurescom	Anastasius Gavras	Summary, Review

Contents

Abbreviations

5G	Fifth Generation (mobile/cellular networks)
5G NR	5G New Radio
5G PPP	5G Infrastructure Public Private Partnership
6G BRAINS	Bring Reinforcement-learning Into Radio Light Network for Massive Connections
AI	Artificial Intelligence
AP	Access Point
AWGN	Additive White Gaussian Noise
B5G	Beyond 5 G
BS	Base Station
CAPEX	Capital Expenditure
CF	Cell-free
CQI	Channel Quality Information
CQT	Constant Q Transform
CSI	Channel state information
CSP	Communication Service Provider
D2D	Device to Device
DDPG	Deep Deterministic Policy Gradient
DL and UL	Down Link and UP Link
DRL and DLIRL	Deep Reinforcement Learning and Deep Learning Integrated RL
DQL & DDQL	Deep Q-Learning and Double DQL
DSP	Digital Signal Processor
FPGA	Field Programmable Gate Array
gNB	5G Base Station
GPU	Graphics Processing Unit
IAB	Integrated Access and Backhaul

IoT	Internet of Things
KPI	Key Performance Indicator
LCM	Life Cycle Management
LDA	Linear Discriminant Analysis
M2M	Machine to Machine
MA-DRL	Multi Agent Deep Reinforcement Learning
MA-RAC	Multi Agent Relational Actor Critic
MDP	Markov decision process
MLP	Multilayer Perception
mMTC	Massive Machine Type Communication
MUD	Multiuser Detection
MSE	Mean Square Error
MVC	Model View Controller
NOMA	Non-Orthogonal Multiple Access
OPEX	Operational Expenditure
PCP	Poisson Cluster Process
PDSCH	Physical Downlink Shared Chanel
PEP	Packet Error Probability
QoE	Quality of Experience
QoS	Quality of Service
R&D	Research and Development
RAN	Radio Access Network
RmUE & mUE	Remote mobile UE and mobile UE
SARSA	State- Action-Reward-State-Action
SDN	Software Defined Networks
SINR	Signal to Interference + Noise Ratio

STFT	Short Time Fourier Transform
UE	5G User Equipment
UL & DL	Uplink and Downlink

1 Introduction

1.1 Introduction and Objective of this document

Communications Service Providers (CSPs) strive for relentless efficiency and business agility to address new revenue opportunities, and to meet or exceed customer expectations.

5G programmable networks enable new revenue-creating opportunities through both enhanced user experience as well as the tailoring of telecommunications networks to provide differential services for both existing and new types of enterprise customers (e.g., Industry 4.0, Automotive, Fixed Wireless etc.).

The introduction of new technologies as well as additional novel services for the customers, the densification of networks to support macro and micro coverage, and the need to ensure services with differing requirements significantly increases the network complexity.

Artificial intelligence (AI) technologies have already matured to the point where CSPs have been applying them to their networks, often starting with non-time-critical processes, and are now applying them to the sensitive parts of their networks that directly impact user experience. The increased complexity of networks due to more services, new network technologies, and massive network densification further necessitates the application of AI in telecommunications networks as operations become more complex.

AI technologies make many CSPs' system functions more capable as well as enable new system functions and approaches. Some example applications include:

- Improving network performance though better scheduling optimized to match the operator business model and users' requirements.
- Improving network optimization and use of existing resources, such as spectrum, transport, cloud infrastructure and network functionality
- improving experience management through both increased customers understanding as well as increased tailoring of the offered experience
- improving product and service definition, design, planning and offerings
- improving network and performance planning (such as radio, data centre location and transport)

Consequently, CSPs are facing today several challenges regarding which standards to follow, which aspects of open source should be utilized directly or via vendors, how to increase industry alignment for scale while simultaneously allowing for differentiation, how to leverage the scale of public cloud providers, how to collect and manage data, and how to support the Life Cycle Management (LCM) of AI models.

The objective of this document is to update and summarize the progress performed in the work package 4 Research of the 6G BRAINS project in innovative Access Network research and concepts that are investigated by the WP-4 team beneficiaries including:

- Artificial Intelligence (AI) based scheduler for a 5G Cell-Free (CF) Networks with Integrated Access and Backhaul (IAB)
- Grant-free NOMA for massive Machine-Type Communication

- Hybrid user terminal modelling for the D2D enabled cooperative network.
- Intelligent IAB with Beam Steering based on User Location
- Advanced Test and Simulation Tools supporting 6G BRAINS research.

As for the scheduler part, in modern cellular networks the RAN scheduler challenges are very demanding in capacity, response time and reliability. These requirements need are especially demanding at Cell Free Network solutions. The RAN architecture includes also an IAB (Integrated Access Backhaul) capability making the solution extra flexible, efficient with low CAPEX.

The scheduler task is to allocate the frequency, time and space resources to each user's end to end link (UL and DL). Allocation considers users priority, traffic load, required resources, session response time, available resources, air-link performance, interference and more.

The scheduler solution was divided into two modules: The Resource Allocation Module and the Routing Module.

NOMA (non-orthogonal multiple access) for IoT applications and D2D (device-to-device) cell free solutions were also researched.

To cope with the expected challenges an AI based solution was researched.

Herein we describe the research items details and the related progress achieved in the in the final 6G Brains project reporting period.

During the parties research many experiments were performed on a wide variety of models with different Neural Networks architectures and methods.

The AI based scheduler was compared to a simple scheduler according to the suggested benchmark It was shown that our AI based scheduler is better than the simple one.

A novel architecture for the integration of the intelligent beamforming and compressed sensing technique is proposed for the grant-free NOMA in a MIMO cellular network. The spectral efficiency is improved by supporting more users in the same frequency spectrum.

In Hybrid user terminal modelling for the D2D work an advanced strategy relying on heterogeneous graph neural network. The CFD-HetGNN consists of aggregation plus combination layers and DNN modules, and it was shown to significantly outperform traditional MF method with equal power allocation.

The Intelligent IAB with Beam Steering based on User Location research yielded a novel DLIRL (Deep Learning Integrated Reinforcement Learning) - based beam-steering scheme, capable to steer the beam with the user movement. Using this technique, the average spectral efficiency of wireless systems can be enhanced.

The Advanced Test and Simulation Tools which were researched can generate data for testing AI-based RAN control applications (xApps). With these advanced tools operators can optimize and launch new services and ease network congestion problems.

The following chapters review the project research, including research details, results and conclusions.

2 AI Based Cell Free Scheduler with IAB

2.1 Previous Research Summary

In previous published reports D4.1 (Chapter 2) and D4.2 (Chapter 2) we presented the scheduler challenges in modern cellular networks and especially in Cell Free Networks with Integrated Access and Backhaul (IAB). We divided the scheduler solution into two sub-schedulers modules: The *Resource Allocation* Module described in D4.1 document and the *Routing* Module described in D4.2 document.

In the first portion of this chapter (Paragraph 2.2 to 2.6) we describe the progress in the **Resource Allocation** scheduler module [1] research achieved in the final 6G Brains project reporting period and in the second portion of this chapter (Paragraph 2.7 to 2.10) we describe the progress in the **Routing** scheduler module [2] research achieved in the same period.

2.2 Resource Allocation Scheduler Module

In general, the scheduler main task is to allocate the frequency resource to each user's link (UL and DL) without collision. To formulate the problem we define two input vector feature as x^{UE}, x^{IAB}, target scheduling function as Φ and the proportional bandwidth allocation relative to the total bandwidth as η^{-} . The input vector feature (UE and IAB input feature respectively) contains the desire throughput, efficiency of each link and the division between those two. Now we can write:

$$\bar{\eta} = \Phi\left(concatenate(\mathbf{x}^{UE}, \mathbf{x}^{IAB})\right)$$

In term of neural network, we concatenate x^{UE} and x^{IAB} to be one input layer and then forward it through Deep Neural Network model Φ. η^{-} is a vector, each index i correspond to single link and mark its proportional bandwidth allocation η_i.

The solution approach can be discussed in two angles: Cloud-based or Distributed-based. Cloud-based would be the optimal approach which an "artificial brain" placed in a single point (IAB-donor) oversees all tasks related to resource management across the whole network or at least a network segment. All available data should be collected and stored in this artificial brain which is tasked with executing all required computations and with feeding back the resulting optimal resource allocation policy to all other network terminals. Unfortunately, such a centralized approach is not always compatible with future wireless networks due to at least three major reasons: latency, privacy and connectivity [3]. There is a distinct trade-off between wireless networks complexity and simplicity of the algorithm.

Second approach is the Distributed-based where each device in the network is a rational and independent decision-maker, which acquires its own local dataset and uses it to build its own local Neural Network model. This technique does not require any interaction between the network infrastructure and the edge users, as far as data sharing and processing are concerned, and has the potential of enabling the 6G vision of distributed, self-managing networks. This approach is preferred to lower wireless networks [3] throughput and restrictions.

5G-NR IAB network can be divided into two hierarchies, network's access and networks backhaul. Access is the bottom part of the hierarchies and handle Uu interface (connection to the UE) while the backhaul is the top part of the hierarchies and handle the F1 interface (connection between the IAB-nodes). Because of those division in the network, we suggest dividing the problem into 2 sub-problem (two hierarchies), each one will use a different model:

- **Big Model** - implement in the IAB-donor (cloud base) and divided the total bandwidth between all IAB (including itself).
- **Small Model** - implementing all IAB (the same model, distributed base) and allocated the frequency between the access (UL an DL links of the UEs) and the backhaul (UL and DL links to the served Base station)

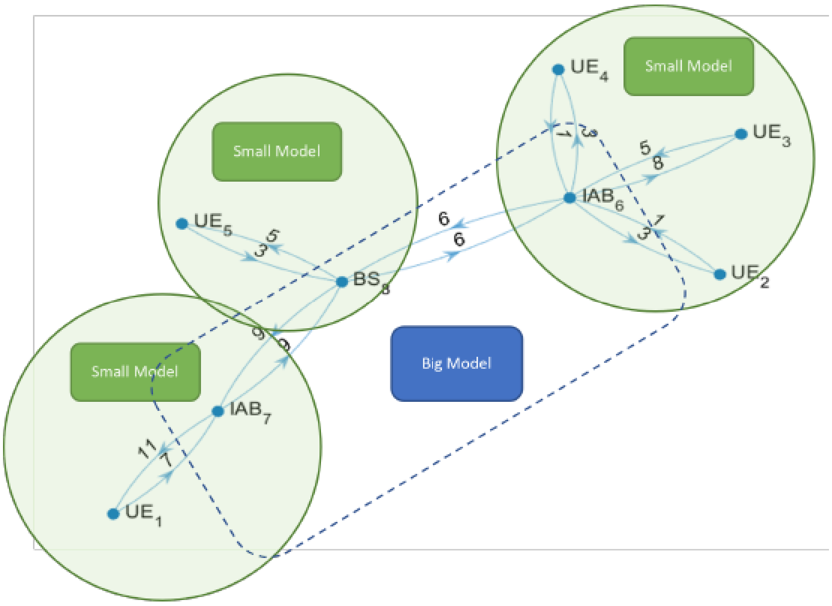
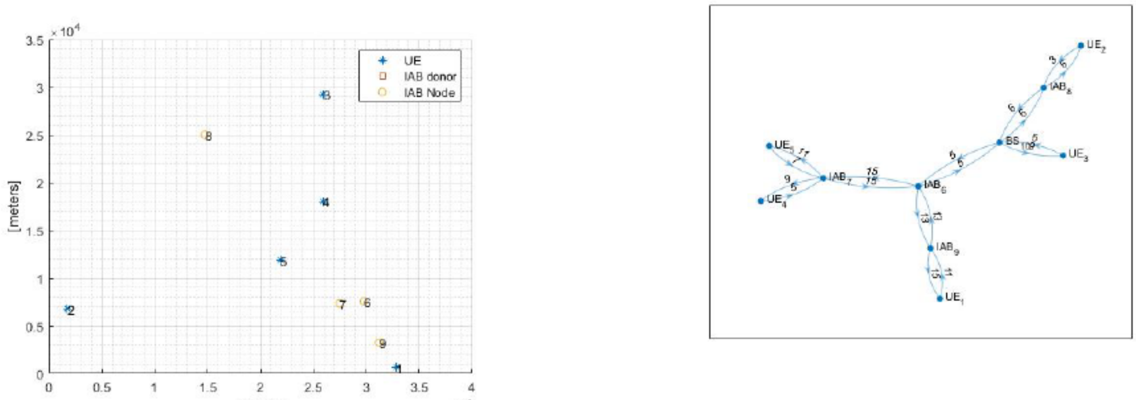


Figure 1: ”Big” and ”Small” model segmentation on the network topology

Figure 1 above describe the segmentation of the models on the network topology illustration describe in Figure 2 below.



(a) Unit deployment in defined space of 40x40km

(b) Network topology - node are the communication units, edges are the RF link between them andm the weights are the CQI

Figure 2: Example of standard unit deployment (a) and network topology established (b) according to the simulation environment.

Due to the problem-splitting, the output of the Big Model become an input feature for the Small Model. Therefore we update the problem formulation:

$$\bar{\eta} = \Phi^{small}\left(\Phi^{big}(x^{IAB}), x^{UE}\right)$$

Where $\Phi^{big}(x^{IAB})$ calculate the bandwidth allocated to each gNB, and Φ^{small} split this bandwidth between the access and the backhaul of the IAB-Node. Φ^{big} and Φ^{small} are multi-layer perceptions (MLP) contain 3 layers with GeLU (Gaussian error Linear Units) activation [4] and Softmax output layer. The output represent the proportional bandwidth allocation relative to the total bandwidth. ”Small model” output is a resource allocation for every UE and one backhaul - each connection correspond to two output (UL and DL). Additionally, we added embedded layer ψ to each model in order to extract more features. Finally we get the formulated model as:

$$\bar{\eta} = \Phi^{small}\left(\Phi^{big}\left(\psi(\mathbf{x}^{IAB})\right), \psi(\mathbf{x}^{UE})\right)$$

2.3 Updated Research Results

2.3.1 Handling dynamic number of UE by the Scheduler

When implementing Deep Neural Network, the number of the input and output layers should properly define with a fixed number. It can be done for the big model, because we fixed the number of IAB-nodes and IAB-donors. However, although the number of UE is also fix, in every Monte-Carlo iteration each gNB serve different number of UE, so the ”Small model” need to handle with changing input size. The changing input directly effect on the output - for example, for input size feature correspond to 10 UEs, the output should allocate resources for 10 UEs (UL and DL) in addition to the backhaul allocation. Our solution contains two stages:

1. Determine maximum UE per base-station - this is the maximum UE each gNB can support, it also consistent with real limitation of gNB hardware. Sets the maximum number of entries to the input layer.
2. Block and normalize the output - because the number of UE are known to the gNB and to the ”Small model”, the model block all the output (multiply by zero) besides those who corresponds to the inputs UE features.

For maximum UE per base-station set to m_{ue} and number for connected UEs equal to c_{ue}, the first 10 output allocations (each UE get two allocations, for UL and DL) are multiplied by ones (pass with-out change). All the other output’s allocations corresponding to the $m_{ue} - c_{ue}$ UEs which did not connect are multiplied by zero. Output’s allocations corresponding to the backhaul does not change. Such intervention in information harms the Softmax attribute (sum over the output equal to one), so we divide each output in the summation over all the new outputs as describe in the equation below.

$$\Phi_{small}(x) = \frac{I^{c_{ue}} MLP_{small}(x)}{\sum_i I^{c_{ue}} MLP_{small}(x)}$$

Note that we implement element-wise multiplication in the numerator (output a vector) and scalar product in the denominator. I^{Cue} serves as indicator row vector with index $i = 0, 1, \ldots, m_{ue} - 1$ where:

$$I^{c_{ue}} = \begin{cases} 1, i < c_{ue} \\ 0, elsewhere \end{cases}$$

Full ”Small model” illustration describe in the Figure 3 below including the embedding, Softmax and normalzation layers.

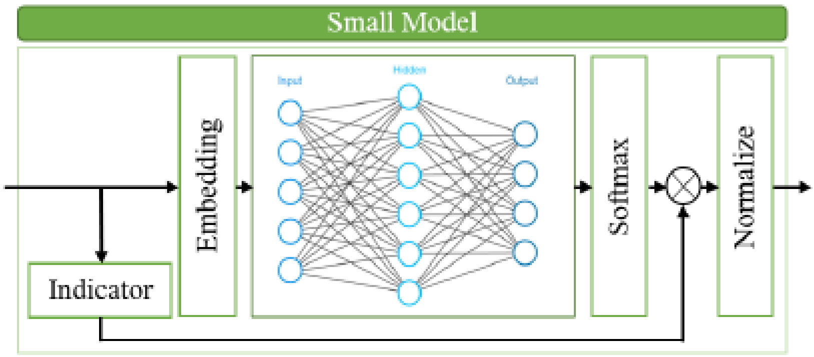


Figure 3: Small model structure

2.3.2 Model Evaluation

The model predicts the proportional bandwidth allocation for every link in the network topology relative to the total bandwidth. All the links needed a precise bandwidth allocation according to the CQI and the desire throughput (capacity). The allocation that supports the link capacity can be calculated - and used in a similar way as ”label”, but not exactly, thus we called it auxiliary label. However, the current allocation is not necessary the auxiliary label, it depends on the total resources of the network. If there isn’t enough bandwidth, the model can’t support the needed allocations (the auxiliary labels) and if there is an excessive amount of bandwidth, we can allocate more than is needed without harming the network. Therefore, we cannot determine the true allocation base on the auxiliary labels alone - an unsupervised learning problem. Hence, we implemented custom cost function that punishes the model if it didn’t allocate enough to the links, and do not punish the model if it allocates more than the link required. The cost function is a combination of cost on every single link of the 5G network. The function ”punish” the link i if the bandwidth allocate to the link did not support the throughput demand. It can be described in the following equations:

$$C^{allocate} - C^{request} < 0$$

$$\eta_i \cdot eff_i(CQI) \cdot \Omega - C_i^{request} \cdot 10^6 < 0$$

$$\eta_i - \frac{C_i^{request} \cdot 10^6}{eff_i(CQI) \cdot \Omega} < 0$$

Where C_i and $eff_i(CQI)$ are the *required capacity* and the *calculated efficiency* of link *i*. The efficiency are depends only from the channel quality indicator (CQI) which is function of the link’s signal to noise ratio (SNR) [5]. Ω is the total bandwidth of the network.

Now we can define αi as the auxiliary label of link i as:

$$\alpha_i = \frac{C_i^{request} \cdot 10^6}{eff_i(CQI) \cdot \Omega}$$

And the cost function of the model is:

$$Cost_{topology} =| \sum_{i \in alllinks} \begin{cases} 0, & (\eta_i - \alpha_i) > 0 \\ \eta_i - \alpha_i, & (\eta_i - \alpha_i) < 0 \end{cases}$$

$Cost_{topology}$ use the difference between the require and the model result (capacity the model did not able to allocate) for the cost calculation. If the model allocates more bandwidth then it needed the cost not increase. However, this implementation of the cost function may not change bandwidth allocation to links with excess bandwidth (each link is greedy), so we added small regularize hyperparameter λ to the cost:

$$Cost_{topology} = \sum_{i \in alllinks} \begin{cases} |\lambda(\eta_i - \alpha_i)|, & (\eta_i - \alpha_i) > 0 \\ (\eta_i - \alpha_i), & (\eta_i - \alpha_i) < 0 \end{cases}$$

We also use another evaluation metric to evaluate how much capacity the model did not able to allocate, this metric called $Performance$ and defined as:

$$Performance = \sum_{i \in alllinks} \begin{cases} 0, & (\eta_i - \alpha_i) > 0 \\ (C_i - \eta_i \cdot eff_i(CQI) \cdot \Omega), & (\eta_i - \alpha_i) < 0 \end{cases}$$

Although the model was separated into ”Big” and ”Small”, the network trained end-to-end. Hence, the models train simultaneously with a single cost function by concatenation them into one model as presented in the equations above. Figure 4 below show to full forward procedure of the models. The motivation behind that is to simplify the training without mixing the models.

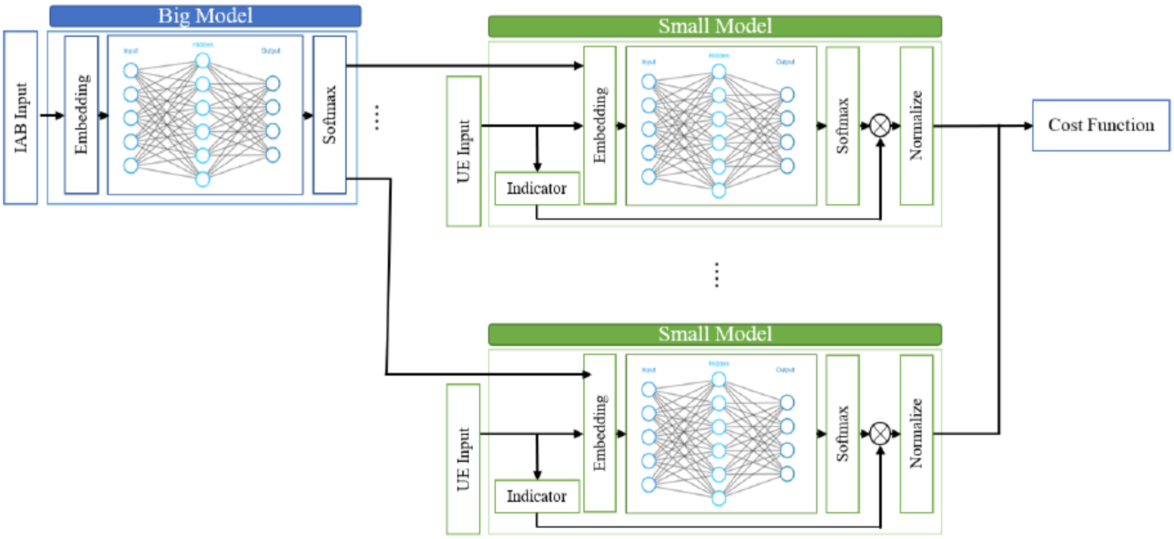
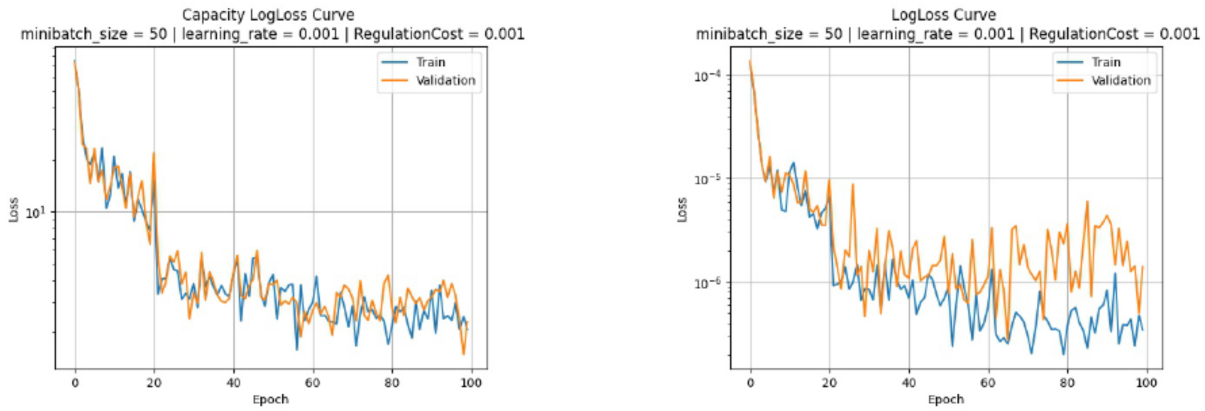


Figure 4: Deep Neural Network architecture - combination between cloud base model (Big Model)and distributed base model (Small Model)



(a) (b)

Figure 5: (a) Performance log-loss (b) Topology cost log-loss

2.4 Research Results

The proposed deep neural network architecture converge into local minimum for both Cost topology and *Performance* as shown in Figure 5. The hyperparameter that have the best results were: minibatch *size* = 50, l*earning rate* = 10^{-3} and $\lambda = 10^{-3}$

We proved that the architecture can be used for scheduling, the remain question is how good the scheduler is? In order to understand it we use other proposed scheduling method:

1. Equal resource allocation (ERA) [6] [7] - Divide equally between all the links.
2. Split Spectrum (SS) [6] -
    (a) Divide the spectrum between the access and the backhaul by factor 0.5.
    (b) Divide equally between all the access links and the backhaul links.
3. 3. Split Spectrum Backhaul Aware (SS-BA) [8] -
    (a) Divide the spectrum between the access and the backhaul by factor 0.5.
    (b) Divide backhaul links proportionally to the IAB's demand.
    (c) Divide equally between all the access links.
4. Fair Access and Backhaul (FAnB) -
    (a) Divide the spectrum bandwidth proportionally to the IAB's demand
    (b) Divide it between the access and the backhaul by factor 0.5.
    (c) (c) Divide the access bandwidth proportionally to the user's demand

The optimal solution, giving centralized solution, should allocate the bandwidth properly. We compare the result of it with all the other schedulers.

2.5 Research Results Analysis

During the work progress, many experiments were performed on a wide variety of models with different hyperparameters, different Neural Networks architectures and methods. The goal was clear - allocation of the available 5G network resources in order to maximize the performance. Judging how well the model performed only by the loss measures is not enough, since those measures don't show the actual allocation.

The amount of information is enormous, 100 UEs with two links (UL and DL) to 10 different IABs which also has two links for backhaul. The result analysis requires visualization tools. Therefore, this section presents in Figure 6 and Figure 7 below the outcomes of our model, both statistically and graphically.

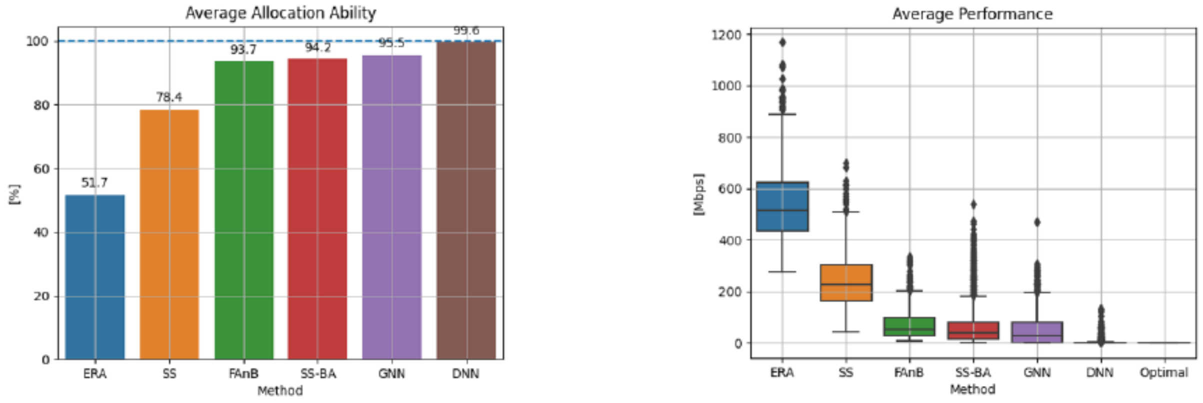
In order to perform a proper analysis and compassion, we define the following benchmarks:

- **Average allocation ability** - the capacity that the scheduler succeeds to allocate compared to the optimal allocation in percentage.
- **Performance** - the difference between the require capacities and the actual capacity given the scheduler bandwidth allocation. Can be calculate from equation below.

$$\frac{1}{N}\sum_{n=1}^{N}[C_i^{requested}(n)| - C_i^{allocation}(n)]^+$$

Where C_i is the capacity over link *i*

- **Average unfulfilled** - The average number of links (UL and DL) that the model did not allocate properly.
- **Average Loss per link** - The average [Mbps] loss for every link that didn't allocate correctly.
- **Bandwidth Usage** - The total bandwidth that the scheduler allocated for the throughput.

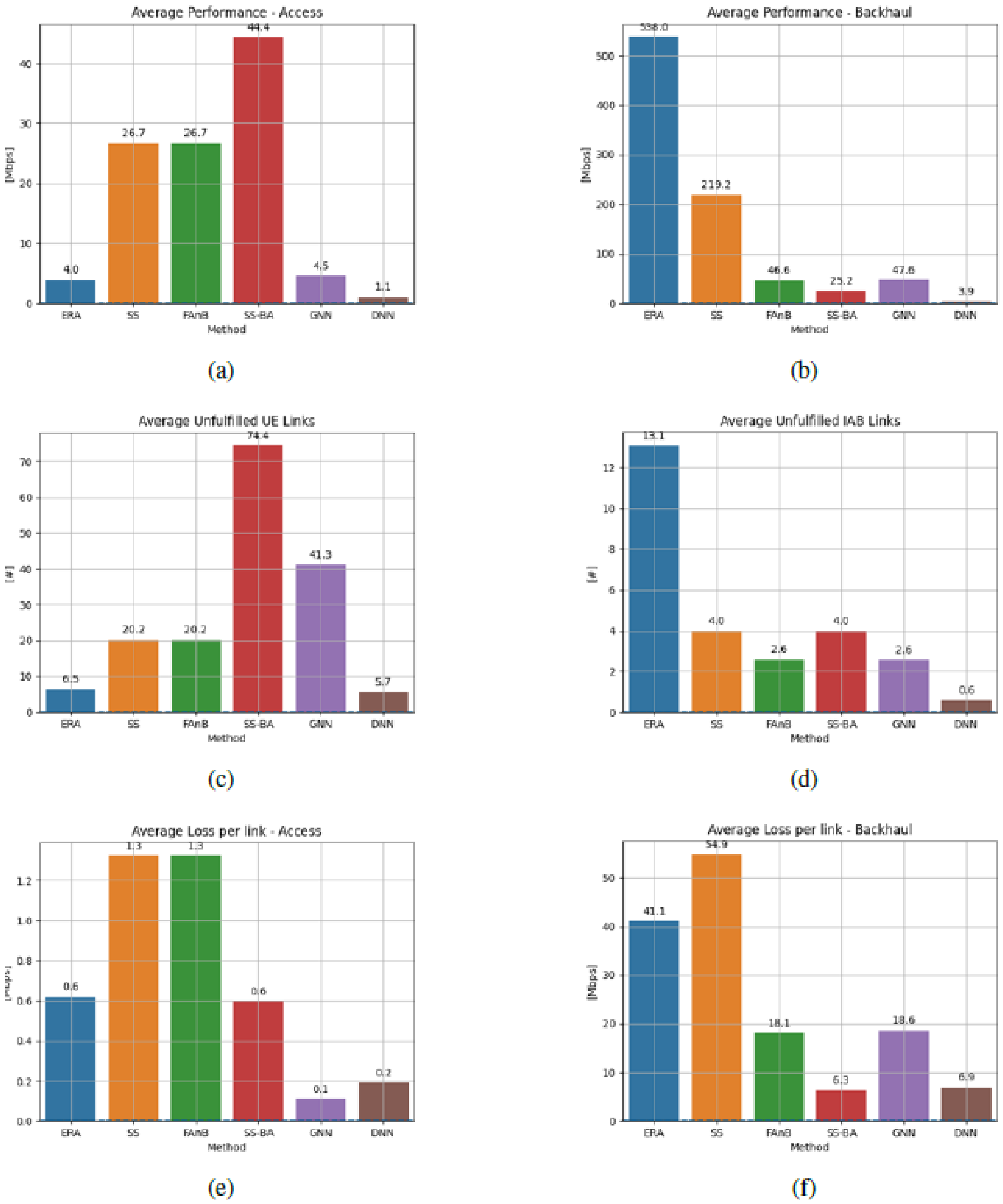


(a) (b)

Figure 6: Allocation ability and performance of the scheduling methods. (a) par plot of allocation ability (b) box plot of performance

2.6 Conclusions and recommendations

In this chapter, we present the 5G-NR IAB network and suggest sophisticated scheduler to allocate bandwidth resource properly. We develop the scheduler according to data that generate in 5G environment. We split the scheduler task into two hierarchies in order to develop practical solution - making the algorithm part cloud-base and part distributed-base. In the end, we compare our scheduler to a simple scheduler according to benchmark that we suggested - Performance Average unfulfilled (soft and hard decision respectively) and Average loss per link. We showed that our scheduler is better than the simple one.



(a) (b)

(c) (d)

(e) (f)

Figure 7: Simulation Results depicting Performance, Unfulfilled and loss per link of the Access and the Backhaul, (a) Access Performance (b) Backhaul Performance (c) Access Unfulfilled (d) Backhaul Unfulfilled (e) Access loss per link (f) Backhaul loss per link

2.7 Routing Scheduler Module

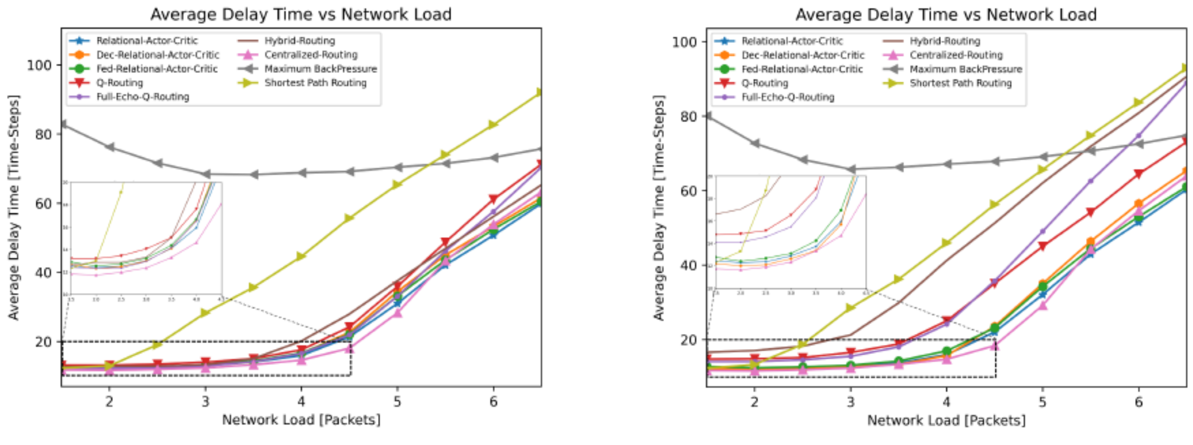
In chapter 2 of Deliverable D4.2 we presented an advanced decentralized routing algorithm for 5G-NR IAB is proposed in this document to reduce congestion in the network. Our study showed that network routing is critical to reduce congestion in such an IAB network. Further, we designed a decentralized routing algorithm called Multiple Agent-Relational Actor Critic (MA-RAC) that learns from experience while interacting with a simulated 5G environment. We benchmark our algorithm against the traditional reinforcement learning approach (Q-Routing) and a lower bound that we derive from a centralized solution. We concluded that the presented routing algorithm outperforms the traditional Q-Routing algorithm in different network scenarios, and it achieves results that are similar to those of a centralized solution. In this paragraph and on we report the results obtained I this research by investigating the influence of the load on the scheduler performance

2.8 The Influence of Network Load

This section examines how network load influences the performance of different algorithms in various mobility scenarios. That is, we evaluate each algorithm’s resilience under different loads for various mobility scenarios. Following this, we conclude that an increase in load / mobility influences the network performance significantly for most of the proposed routing algorithms. As a next step, we will describe in detail our experiments and their corresponding empirical results.

To change the network load, the parameter λ of the Poisson distribution has been modified. This parameter indicates the average number of packets generated in each time-slot by the IABs. To modify λ, we scanned various loads successively from bottom-to-top, and then from top-to-bottom. The results presented are an average of 10 different measurements for each load across five different network topologies.

First, we evaluate the performance of our algorithms over a static topology, which means that users cannot change their base-station association or location. Next, we increase the user’s speed to $3_{\text{m/sec,}}$ which means that users can change their location and change their association with a base station at any time. We then refer to this scenario as a dynamic topology and evaluate our algorithms’ performance. In addition, we investigate how the speed of UEs affects the performance of the different algorithms. Furthermore, due to the lack of further information provided by the arrival ratio, we present only the average delay metric. The following sections provide results and insights from those experiments.



(a) Static Topology. (b) Dynamic Topology.

Figure 8: Static topology and dynamic topology - average delay performance for different routing algorithms under different loads.

2.8.1 Static Topology

This experiment evaluated the algorithm’s performance under a static network topology and changing load. In Figure 8a, we present the performance of the algorithms. Based on these results, we can conclude that all three versions of the Relational Advantage Actor Critic (**A2C**) algorithm outperformed other algorithms, despite acting in a decentralized manner. Moreover, these results also revealed that using exploration in conjunction with exploitation (Full-Echo Q-Routing) greatly improved performance at low to medium loads when compared with Q-Routing, whilst they achieved similar performances for higher loads.

As a further interesting phenomenon, we observe that Hybrid Routing improves performance when compared with traditional Q-Routing under low loads but degrades performance under higher loads. We explain this phenomenon by stating that higher loads require better coordination among agents, as making an incorrect routing decision while the network is already congested will probably result in packet loss. Our conclusion is that since each agent learns an independent stochastic policy as part of the algorithm, it is less effective at handling non-stationary problems than traditional Q-Routing algorithms. In our next experiment, we allow UEs to travel and change their base-station association, which will further increase the non-stationariness of the problem.

Additionally, as network load increases, we observed performance degradation of the Shortest-Path algorithm. Our explanation for this phenomenon is that queue delay does not have a significant influence on performance for low loads, but as the load increases, it becomes more significant.

2.8.2 Dynamic Topology

This experiment evaluated the algorithm’s performance under a dynamic network topology and changing load. Following this, based on the results illustrated in Figure 8b we can determine that although acting in a decentralized manner, all versions of Relational A2C’s algorithm managed to achieve superior performance than the other algorithms for this scenario. Further, Hybrid Routing does not perform well in this non-stationary environment, as we observe further degradation in performance when compared to a static scenario, and at low load it is even less effective than Q-Routing.

Moreover, when compared with the static scenario, Full-Echo Q-Routing exhibits performance degradation at medium to high loads. It performs worse in this range than the traditional Q-Routing. Our explanation for this phenomenon is that when using Full Echo Q-Routing, frequent topology changes may cause multiple modifications to the routing policy, which will result in increased instability and longer routing times as loads increase. In addition, this phenomenon coincides with the phenomena observed in [7], in which the authors applied those algorithms to a grid topology.

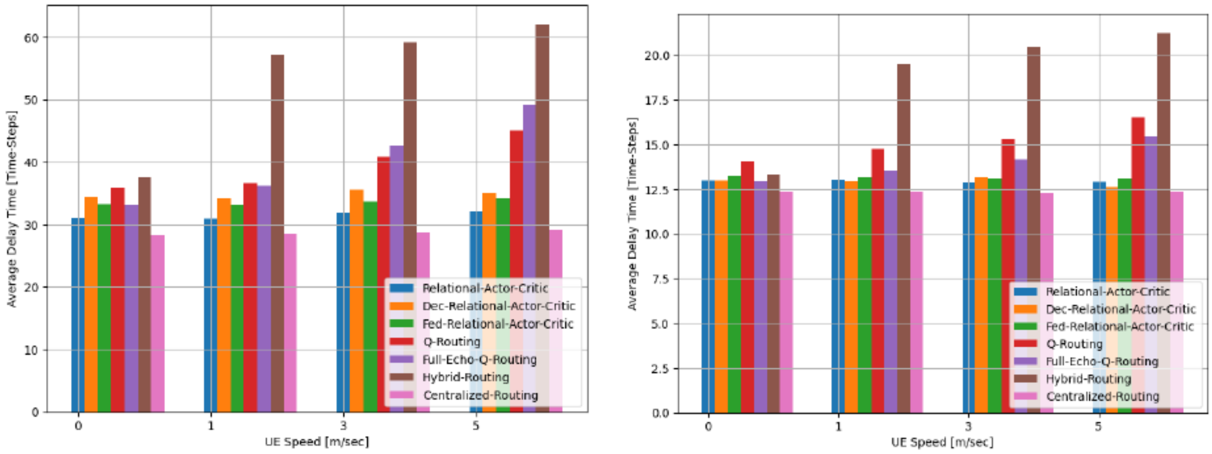
The significant performance gap between all versions of Relational A2C and the traditional hybrid algorithm supports our claim that by following our proposed algorithms, agents can cooperate more effectively and increase the stability of our routing policy, proving that agents that cooperate achieve better outcomes than agents who act selfishly. Thus, these results demonstrate that optimizing the joint goal and solving the coordination problem between agents significantly improves performance for both static and dynamic topologies.

When comparing the performance of static versus dynamic topologies as illustrated in Figure 8, we observe that all tabular RL baselines suffer from performance degradation when users are permitted to move and change their base-station association. Based on these results, which indicate a significant difference between static and dynamic topologies, we propose the following experiment to examine the effect of UE movement on algorithm performance.

2.8.3 The Influence of Dynamic Topology Changes

In this simulation, we evaluated the algorithm’s robustness to dynamic topology changes under constant loads. For this purpose, we scanned various UE speeds and measured the arrival ratio and average delay with the different algorithms. We evaluate the algorithms’ performance for two possible loads, medium load (λ = 3) and high load (λ = 5), presented in Figure 9a and Figure 9b respectively.

Figure 9a and Figure 9b demonstrate that increasing the UE speed will not have a significant impact on the Relational A2C algorithm’s performance for varying network loads. Also, in medium loads, Full Echo Q-Routing achieves superior performance to Q-Routing, with both suffering from similar performance degradation because of their different UE speeds. Furthermore, we observe that the combination of higher load with increased speed results in more severe performance degradation for Full Echo Q-Routing compared to traditional Q-Routing. This observation is consistent with the results of our previous experiments. Moreover, Hybrid Routing shows the greatest degradation among all Q-Routing algorithms, further demonstrating that it is less able to cope with this non-stationary setting than traditional Q-Routing algorithms. Aside from this, we observe that all other RL-based algorithms suffer from performance degradation when UE speed is increased, which serves as an additional indication of our proposed approach’s superiority.



(a) Average Delay For High Load $\lambda = 5$.

(b) Average Delay For Medium Load $\lambda = 3$.

Figure 9: Average network delay for different routing algorithms under different UE speed.

2.9 Experiment Results for Online Changes

The purpose of this section is to examine how online changes impact algorithm performance in a variety of scenarios. Our study indicates that Relational A2C-based approaches are superior for handling such online changes when compared to other algorithms. As a next step, we will describe the online scenarios that we have studied. In the first case, we analysed the algorithm’s response to bursts of traffic. In the second case, we investigate how the algorithm responds to node failures and recovery situations. Also, in these experiments we capture those online changes by measuring the algorithm’s performance using a sliding window with a length of 100 time slots.

2.9.1 Experiment Results for Bursts of Traffic

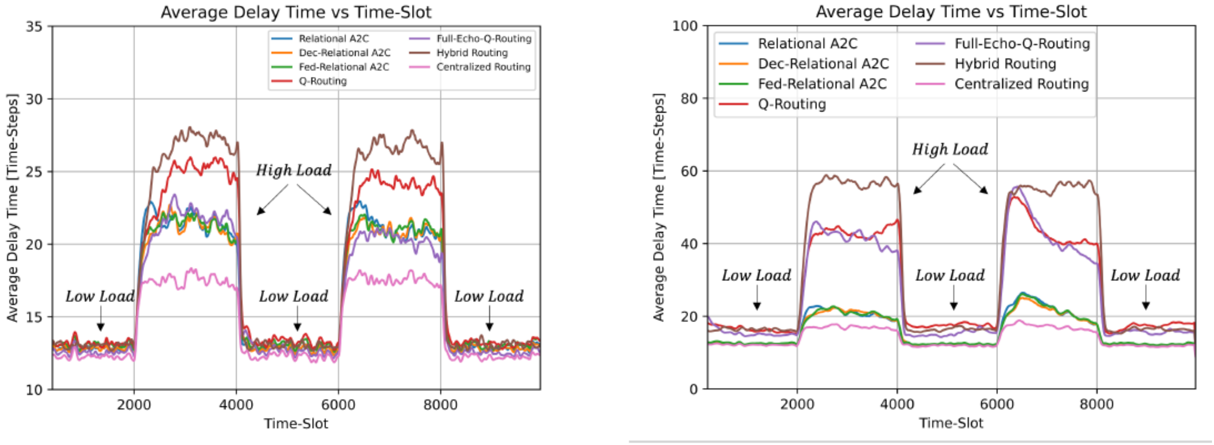
During this experiment we measure the algorithm’s response to bursts of traffic. This was accomplished by evaluating the algorithm performance while changing the network load. Our experiment involves changing the network load rapidly from low (λ = 2) to high (λ = 5) and then back to low. The results presented here are an average of 10 measurements across five different topologies of the network. Furthermore, we conducted this experiment for both static and dynamic topologies, as depicted in Figure 9a and 9b, respectively. Further, due to the absence of further information offered by the arrival ratio, we present only the average delay metric measured through time for this experiment.

In both static and dynamic scenarios, it is evident that the Relational A2C versions achieve superior performance than the traditional RL-algorithms. Additionally, all algorithms exhibit similar reaction times when measuring the impact of changes in load from low to high and from high to low. When comparing between static and dynamic scenarios, we can see that the Relational A2C algorithms did not suffer from any performance degradation, while the tabular methods greatly suffer.

Further, we observe that Full Echo Q-Routing is superior to traditional Q-Routing in static scenarios, while they achieve similar performance in dynamic scenarios. Additionally, Hybrid Routing suffers from the greatest degradation of all Q-Routing algorithms when working with high loads, emphasizing its inability to cope with such a non-stationary scenario. Our conclusions are consistent with the results of our previous experiments, which further verifies the results.

2.9.2 Experiment Results for Node Failure

In this experiment, we evaluate how the algorithms respond to a scenario of node failure and recovery. To achieve this, the algorithm performance was evaluated while removing a random base station from the network for a specified period of time. Since we insert dynamic changes through node failures, we now consider only static topologies, in which users cannot move. The following results in Figure 10 are based on an average of 30 consecutive experiments with random base-station failure in each experiment.

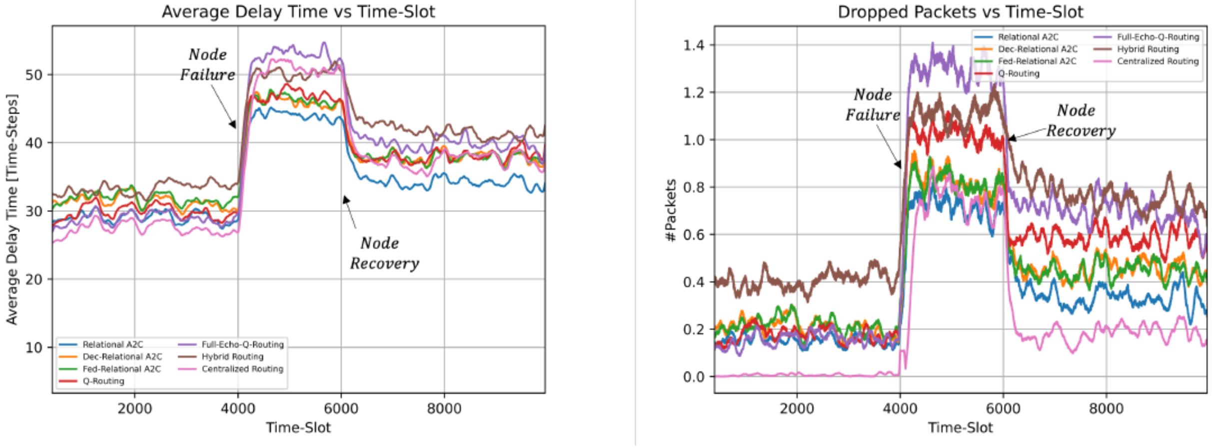


(a) Average Delay For Static Topology.

(b) Average Delay For Dynamic Topology.

Figure 10 Static topology and dynamic topology - average delay performance for different routing algorithms under bursts of traffic. low load λ= 2, high load λ = 5

This experiment was conducted for a high load as shown in Figure 10. To illustrate the arrival ratio metric in such an online setting, we measured the number of packets that dropped within our sliding window. From these results, it is evident that all algorithms can recover from the node failure situation, although we observe that the average delay does not return to its initial value after the node has recovered. In our opinion, this is because the other base stations' queues are already congested at this point. This results in a more congested network than before the node failure. In conjunction with the fact that the arrival rate process does not change over time, performance is degraded under higher loads.



(a) Average Delay Through Time.

(b) Dropped Packets Through Time.

Figure 11: Performance for different routing algorithms in case of node failure for high load λ = 5.

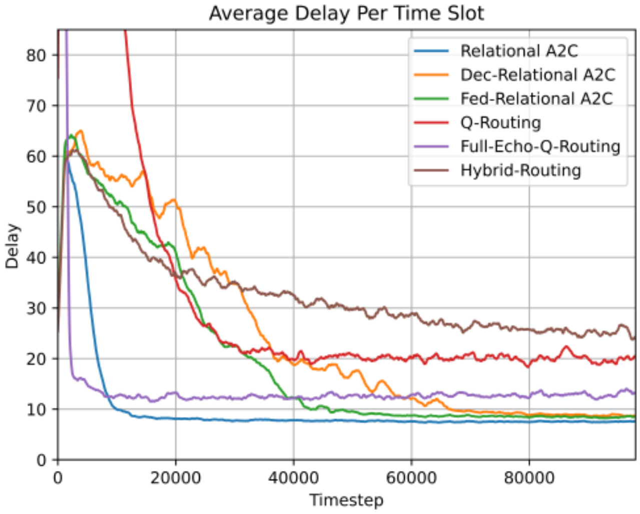


Figure 12: Illustration of the performance of different routing algorithms through their training procedures.

A further, but equally significant, conclusion we can draw from these measurements is that while Relational A2C has a lower delay than Centralized Routing in Figure 11a its packet loss rate is higher in Figure 11b. This demonstrates the importance of using both arrival rate and average delay in our analysis.

2.10 Algorithm Convergence

In this section, we examine the algorithm's convergence time under a constant load, starting with a random initialization point. According to our results, all versions of our algorithm were able to converge to a similar solution, as opposed to the other baselines that suffer from performance degradation. Thus, we conclude that all Relational A2C-based algorithms were able to achieve a better routing solution than traditional algorithms. The results presented in Figure 12 are the average of those measurements.

Additionally, we found that centralized training achieved the fastest convergence among the different training paradigms of our proposed method. In fact, this convergence gap arises because the different agents share their experience implicitly through the mutual updates of both the actor and the critic among all the base stations. Furthermore, we conclude that Fed-Relational A2C performs better than Dec-Relational A2C in terms of convergence. In general, this difference can be understood intuitively since federated learning approaches manage the trade-o_ between a fully decentralized training paradigm and a centralized training paradigm. Another interesting phenomenon is that Full Echo Q-Routing appears to be the fastest algorithm that converges to a stable solution (an approximated point of equilibrium). This rapid convergence is explained by the fact that Full Echo Q-Routing receives all the rewards available to it, regardless of the chosen action. As a result, it can reduce the number of interactions required for convergence with the environment.

2.11 Conclusion

In this research, we examined the problem of routing in an IAB network in which multiple IAB nodes operate simultaneously within the same network in order to route multiple packets towards their destinations. A successful joint routing policy must be developed by the IAB nodes in order to route packets effectively without causing network congestion. Due to physical limitations, the IAB Node is only able to exchange limited information with its neighbours, and thus does not know the current status of the entire network. This raises the interesting and unanswered question of which hops should be selected at each time step so that the network arrival ratio achieved by matching routing policies is maximal and the average packet delay is minimal.

To identify the joint routing policy that maximizes the network's arrival ratio while minimizing the average packet delay, we developed a novel Relational A2C algorithm, which aims to determine the best joint routing strategy, based on observations collected by the IABs nodes via online learning. To support decentralized training, we developed two different approaches which achieve similar performance to the centralized training approach. For various different scenarios, we compared the arrival ratio and average delay of the proposed Relational A2C algorithms with those of six other algorithms. It was found that Relational A2C performed better than baseline algorithms in all cases and achieved performance similar to that of a centralized approach. Further, we demonstrated that network routing is crucial to reducing congestion and maximizing the utilization of network resources in IAB networks. These results clearly demonstrate the ability of Relational A2C based algorithms to learn near-centralized policies and the overall superiority of the proposed approach over existing methods.

2.12 References

[1] -A. Moses, H. Permuter from Ben Gurion University, Beer Sheeva, Israel: “5G Integrated Access Backhaul scheduling algorithm base Deep Learning” April 2022.

[2] - S. Yamin, H. Permuter from Ben Gurion University, Beer Sheeva, Israel: “Multi-Agent Reinforcement Learning for Network Routing in Integrated Access Backhaul Networks, May 2023

[3] - A. Zappone, M. Di Renzo, and M. Debbah, “Wireless networks design in the era of deep learning: Model-based, ai-based, or both,” IEEE Transactions on Communications, vol. 67, no. 10,pp. 7331–7376, 2019.

[4] D. Hendrycks and K. Gimpel, “Gaussian error linear units (gelus),” arXiv preprint arXiv:1606.08415, 2016.

[5] Y. Wang, W. Liu, and L. Fang, “Adaptive modulation and coding technology in 5g system,” in 2020 International Wireless Communications and Mobile Computing (IWCMC), pp. 159–164, IEEE, 2020.

[6] R. I. Rony, E. Lopez-Aguilera, and E. Garcia-Villegas, “Cooperative spectrum sharing in 5G access and backhaul networks,” in 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 239–246, IEEE, 2018.

[7] S. Zhang, X. Xu, M. Sun, X. Tao, and C. Liu, “Joint spectrum and power allocation in 5g integrated access and backhaul networks at mmwave band,” in 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1–7, IEEE, 2020.

[8] C. Saha, M. Afshang, and H. S. Dhillon, “Bandwidth partitioning and downlink analysis in millimeter wave integrated access and backhaul for 5g,” IEEE Transactions on Wireless communications, vol. 17, no. 12, pp. 8195–8210, 2018.

3 Machine learning based Grant-free NOMA for Massive Machine-Type Communication Integrated with Intelligent beamforming.

3.1 Related works and motivation

The massive machine type communication (mMTC), proposed in the fifth-generation mobile communication (5G), will still play an critical role in the future beyond 5G and even the sixth generation mobile communication (6G). In mMTC scenarios, e.g., the internet of things (IoT), a cellular base station (BS) usually needs to serve a large number of users. Non-orthogonal multiple access (NOMA) has been verified as an effective method to support the massive user access with limited channel resources [1-5]. Another characteristic of the mMTC scenario is that at any given time only a small fraction of potential users are active and transmit small data packets, i.e., the user activity is sporadic [6-9]. In this case, still using the conventional grant-based NOMA techniques would cause the large access delay and signalling overhead. Therefore, an efficient communication paradigm shift is necessary to enable the low-latency and high-reliability mMTC applications.

3.1.1 Related works

Recently, the uplink grant-free NOMA methods have been proposed by researchers as feasible solutions. Roughly speaking, in the grant-free access, the active users could directly transmit data using the available channel resources that the BS broadcasts periodically, without needing the complicated channel access request and granting process [10]. Thus, the grant-free access is effective in the mMTC applications, in reducing the access delay and signalling overhead due to the sporadic and small-scale data transmission. In the meantime an issue in the grant-free access occurs, that is, the BS could not identify the active users before data transmission. Therefore, blind user activity detection is necessary at the BS by using the sum received signal of the active users.

Due to the sporadic transmission property in mMTC, the received signal can be sparse with high probability. The compressed sensing (CS) techniques are promising in recovering the sparse signals from the far fewer samples than those required by the classic Nyquist sampling [11–13]. Accordingly, the number of necessary resource elements used for data transmission can be reduced when considering the CS-based receiver. The integration of the CS and grant-free MOMA depends on the transceiver design. At the transmitter, the active users modulate the information bits into modulated symbols, and then spread them onto the specific subcarriers for multiplexing transmission. The widely used coding techniques include the low density signature (LDS) [1], the sparse code multiple access (SCMA) [2, 3, 14, 15] and the index modulation (IM) [16, 17], etc.. At the receiver, the received signals on different subcarriers are sampled simultaneously which are used for the user activity detection and data recovery by mean of the CS techniques [10–13, 18].

Extensive CS-based sparse signal recovery methods have been proposed, including the orthogonal matching pursuit (OMP) [19], compressed sampling matching pursuit (CoSaMP) [12], subspace pursuit (SP) [13] and approximate message passing (AMP) method [18], etc. These methods require prior knowledge of the user sparsity level, which is often impractical in engineering applications. Further considering the data transmission in consecutive slots usually in mMTC scenarios, the time correlation has been utilized in existing methods for the enhanced user activity detection and signal recovery in grant-free NOMA systems [20–27].

The assumption on the time correlation of the user activity is divided into two categories. The first one is that the user activity stays constant in one frame, called frame wise (block) sparsity. Based on this, the modified AMP [20], SP [21] and block-coordinate-descent (BCD) [22] methods were respectively developed for the joint user detection and data recovery in grant-free NOMA. These methods do not require the prior user sparsity level but to estimate it based on the prior noise power. In particular, to avoid using the prior information of the noise level, the authors in [21] also proposed a cross-validation-based method to determine the user sparsity level in light of the residual energy computed by the measurement samples selected in advance. The authors in [23] considered an orthogonal approximate message passing (OAMP)-multiple measurement vector (MMV) algorithm with simplified structure learning (SSL) and accurate structure learning (ASL), termed as OAMP-MMV-SSL and OAMPMMV-ASL, respectively. These two methods can iteratively estimate the user sparsity ratio and the noise variance based on the expectation maximization [23].

The second assumption is called dynamic user sparsity where the user activity can be different in consecutive slots. A dynamic CS method [24] and a modified SP method [25] were proposed respectively, by utilizing the time correlation of adjacent user support sets for enhanced user detection performance. In addition, the weighted $l_{2,1}$ minimization model-based method was proposed with further enhanced detection performance [26]. The aided single bit transmission with value 0 or 1 was also applied to detect whether the active user has data to transmit in current slot in scenarios with the dynamic sparsity [27]. All of them require the noise level as the prior information.

3.1.2 Motivation

The aforementioned CS-based methods for grant-free NOMA system are usually based on the single-antenna BS. Recently, [28] demonstrated that, both the missed device detection and the false alarm probabilities for activity detection can always converge to zero by utilizing the vector AMP algorithm [18], in the asymptotic massive multiple-input multiple output regime. A joint spatial-temporal-structured adaptive SP method was proposed for grant-free NOMA to jointly estimate channels and detect users by considering the block sparsity over multiple slots and multiple antennas [29]. The dynamic sparsity in grant-free mMTC is also investigated in terms of the multiple antenna reception [27].

Current CS-based grant-free NOMA techniques still require a large number of channel resources for the accurate sparse signal recovery for massive connectivity, even though they can realise the system overloading to some extent. The spatial division multiple access (SDMA) technique characterized by the multiple-antenna BS has been verified effective to support massive user connectivity, especially when integrating with the power-domain NOMA techniques [30–36]. Therefore, it is a promising solution to integrate the SDMA with the CS-based grant-free NOMA technique in mMTC applications for improved spectral efficiency. However, to our best knowledge, there is no work in open literature that has taken this into consideration.

3.2 Model definition

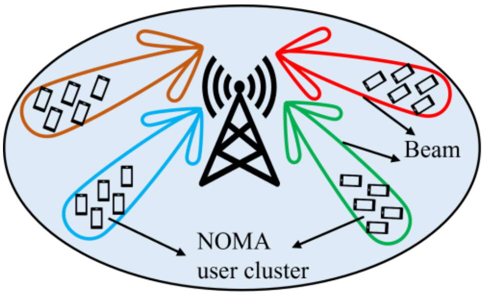


Figure 13: System architecture for the integration of beamforming and grant-free NOMA

In this work, we study the multiuser detection (MUD) and data recovery (DR) for the uplink grant-free NOMA to a multiple-antenna BS. We consider i) the first temporal correlation assumption, i.e., the frame-wise block sparsity for each user; ii) the coherent grant-free access type with the channel information estimated using pilots before the data transmission [20-26].

We consider the spreading-based grant-free NOMA in a multiple-antenna cellular system to support the mMTC with limited channel resources. As shown in Figure 13, *NQ* users (devices) are grouped into *N* clusters according to their channel correlation by using common clustering methods, such as *K*-means [34, 37, 38]. The main steps of K-means clustering are as follows:

- Step 1 Given the number of clusters *N*
- Step 2 Select *N* data objects as the initial cluster centres randomly
- Step 3 Calculate the distance of each data object
- Step 4 Classify the data object into the cluster
- Step 5 Adjust the new cluster and recalculate the centre of the new cluster
- Step 6 Repeat the above three steps to find if the centre converges

Without loss of generality, the equal-size clusters are assumed, e.g., *Q* users in each cluster *n = 1, 2, ... ,N*. The BS is equipped with a uniform linear array with *M* antenna elements while all users are with a single antenna. All user clusters employ the same frequency resources, i.e., *K* subcarriers, for simultaneous communication with the BS. To support mMTC, we consider an overloaded system with *K < NQ*.

3.2.1 Signal model

A slow-fading channel is assumed, which remains unchanged within a coherence time interval (longer than the frame length of the mMTC). Assuming the line-of-sight transmission only, the angle of arrival (AoA) from user *q* can be denoted as $\theta_{n,q}$ and the steering vector is defined as, $\boldsymbol{a}_{n,q} = \left[1, e^{j2\pi\, d \sin\theta_{n,q}/\lambda}, \dots, e^{j2\pi\,(M-1) d \sin\theta_{n,q}/\lambda}\right]^{\mathrm{T}}$, where *e* is the Euler's number, λ is the carrier wavelength and *d* is the distance between the adjacent antenna elements, usually set to be a half wavelength $\lambda/2$. The channel gain vector $\boldsymbol{g}_{n,q,k}$ between the user *q* and the multiple-antenna BS using subcarrier *k* can be modelled as the product of the channel fading and the steering vector, defined as $\boldsymbol{g}_{n,q,k} = f_{n,q,k}\boldsymbol{a}_{n,q}$, where the channel fading $f_{n,q,k}$ consists of the large-scale fading, including the path loss and shadowing fading, and the small-scale random fading following the standard complex Gaussian distribution.

The received signal at the BS on subcarrier *k* and at slot *t* can be formulated as,

$$\boldsymbol{y}_{k,t} = \sum_{n\in\aleph} \widetilde{\boldsymbol{G}}_{n,k}\boldsymbol{x}_{n,t} + \boldsymbol{v}_{k,t},$$

where $\aleph$ is the index set of all clusters, $\boldsymbol{x}_{n,t}$ is the transmitted signal vector with its *q*th entry being $x_{n,q,t}$, $\boldsymbol{v}_{k,t}$ is is the additive Gaussian noise vector and $\widetilde{\boldsymbol{G}}_{n,k} \triangleq [\widetilde{\boldsymbol{g}}_{n,1,k}, \widetilde{\boldsymbol{g}}_{n,2,k}, \dots, \widetilde{\boldsymbol{g}}_{n,Q,k}]$ with the equivalent channel gain vector $\widetilde{\boldsymbol{g}}_{n,q,k} \triangleq s_{n,q,k}\boldsymbol{g}_{n,q,k}$. Note that $s_{n,q,k}$ is the spreading factor for user *q* on subcarrier *k*. The cascaded received signal vector $\boldsymbol{y}_t$ is given by,

$$\boldsymbol{y}_t = [\boldsymbol{y}_{1,t}^{\mathrm{T}}, \boldsymbol{y}_{2,t}^{\mathrm{T}}, \dots, \boldsymbol{y}_{K,t}^{\mathrm{T}}]^{\mathrm{T}} = \sum_{n\in\aleph} \widetilde{\boldsymbol{G}}_n \boldsymbol{x}_{n,t} + \boldsymbol{v}_t$$

with the equivalent channel matrix $\widetilde{\boldsymbol{G}}_n = [\widetilde{\boldsymbol{G}}_{n,1}^{\mathrm{T}}, \widetilde{\boldsymbol{G}}_{n,2}^{\mathrm{T}}, \dots, \widetilde{\boldsymbol{G}}_{n,K}^{\mathrm{T}}]^{\mathrm{T}}$ and the noise vector $\boldsymbol{v}_t = [\boldsymbol{v}_{1,t}^{\mathrm{T}}, \boldsymbol{v}_{2,t}^{\mathrm{T}}, \dots, \boldsymbol{v}_{K,t}^{\mathrm{T}}]^{\mathrm{T}}$.

For any cluster *n = 1, 2, ... , N*, the multi-antenna received signal is combined by beamforming, i.e.,

$$\boldsymbol{y}_{n,t} = (\boldsymbol{I}_K * \boldsymbol{b}_n)^{\mathrm{H}} \boldsymbol{y}_t$$

where $*$ denotes the Kronecker product, $\boldsymbol{I}_K$ is a $K \times K$ identity matrix and $\boldsymbol{b}_n$ is the beamforming weight vector for cluster *n*.

By defining $\boldsymbol{B}_{n,l} = (\boldsymbol{I}_K * \boldsymbol{b}_n)^{\mathrm{H}} \widetilde{\boldsymbol{G}}_l$, $\boldsymbol{y}_{n,t}$ can be rewritten as,

$$\boldsymbol{y}_{n,t} = \boldsymbol{B}_{n,n}\boldsymbol{x}_{n,t} + \sum_{l\in\aleph\setminus n} \boldsymbol{B}_{n,l}\boldsymbol{x}_{l,t} + (\boldsymbol{I}_K * \boldsymbol{b}_n)^{\mathrm{H}} \boldsymbol{v}_t,$$

where the first term on the right-hand side is the desired signal for cluster *n*, the second is the sum signal of the inter-cluster interferences, and the last is the noise term.

3.3 Problem formulation

The second grant-free access type is considered, i.e., the channel gains are a priori estimated. Non-sparse spreading signatures are considered, e.g., the Zadoff-Chu sequences. With the known channel information and spreading signatures, one can obtain the equivalent channel gain matrix $\widetilde{\boldsymbol{G}}_n$. This section aims at developing an algorithm for optimising the beamforming weight and signal estimate jointly at the BS.

Define the transmitted signal matrix for cluster n as $\boldsymbol{X}_n = [\boldsymbol{x}_{n,1}, \boldsymbol{x}_{n,2}, \dots, \boldsymbol{x}_{n,T}]$, with T denoting the number of slots in one frame. The least-squares (LS) error function for MUD and DR is given by, $\xi_{LS}(\boldsymbol{b}_n, \boldsymbol{X}_n) = \sum_{t=1}^{T} \left\|\boldsymbol{y}_{n,t} - \boldsymbol{B}_{n,n}\boldsymbol{x}_{n,t}\right\|_2^2$. To optimize the signal estimation, it is important to constrain the beamforming mainlobe towards the desired user cluster, i.e., $\boldsymbol{b}_n^{\mathrm{H}}\overline{\boldsymbol{a}}_n = 1$ where $\overline{\boldsymbol{a}}_n = 1/Q \sum_{q=1}^{Q} \boldsymbol{a}_{n,q}$ is the average of the steering vectors of the users in cluster *n*. The joint optimisation problem can be formulated as,

$$\underset{\boldsymbol{b}_n, \boldsymbol{X}_n}{\operatorname{argmin}} \, \xi_{LS}(\boldsymbol{b}_n, \boldsymbol{X}_n)$$

$$s.t. \left\|\boldsymbol{x}_{n,t}\right\|_0 \le \bar{s}, \boldsymbol{b}_n^{\mathrm{H}}\overline{\boldsymbol{a}}_n = 1$$

where $\bar{s}$ is the maximum user sparsity level.

3.4 Proposed method

3.4.1 Beamforming schemes

On the one hand, the LS error in (8) can be converted into the mean squared error (MSE) when three conditions satisfy, i.e., 1) the number of slots (samples) is large enough, 2) the transmitted signals follow stationary distributions and 3) the channel states stay unchanged within a frame. By taking the mathematical expectation of $\left\|\boldsymbol{y}_{n,t} - \boldsymbol{B}_{n,n}\boldsymbol{x}_{n,t}\right\|_2^2$, the MSE cost function can be obtained. With the mainlobe constraint $\boldsymbol{b}_n^{\mathrm{H}}\overline{\boldsymbol{a}}_n = 1$, a statistical beamforming (SBF) scheme is proposed, i.e.,

$$\boldsymbol{b}_n^{\mathrm{SBF}} = \frac{\left(\sum_{l\in\aleph\backslash n}\alpha_l\sigma_l^2\sum_{k=1}^{K}\widetilde{\boldsymbol{G}}_{l,k}\widetilde{\boldsymbol{G}}_{l,k}^{\mathrm{H}} + K\sigma_v^2\boldsymbol{I}_M\right)^{-1}\overline{\boldsymbol{a}}_n}{\overline{\boldsymbol{a}}_n^{\mathrm{H}}\left(\sum_{l\in\aleph\backslash n}\alpha_l\sigma_l^2\sum_{k=1}^{K}\widetilde{\boldsymbol{G}}_{l,k}\widetilde{\boldsymbol{G}}_{l,k}^{\mathrm{H}} + K\sigma_v^2\boldsymbol{I}_M\right)^{-1}\overline{\boldsymbol{a}}_n}$$

where α_l is the user sparsity ratio in cluster l, σ_l^2 is the transmission power of each user in cluster l and σ_v^2 is the power of the additive noise. Practically, we can select an empirical SNR (ESNR) σ_l^2/σ_v^2 and a rough α_l from (0, 1] without needing their exact values since the relative value of these two factors results in a trade-off between the suppression of the noise and interference.

On the other hand, the small data sample per user is insufficient to match the statistics in MSE cost function by using the sample variance. In addition, the inaccurate ESNRs and user activity probabilities also influence the trade-offs between the interference suppression and noise suppression to some extent. Thus, an LS-based dynamic beamforming (DBF) scheme is proposed, i.e.,

$$\boldsymbol{b}_n^{\mathrm{DBF}} = \frac{(\boldsymbol{R}_n + \epsilon\boldsymbol{I}_M)^{-1}\overline{\boldsymbol{a}}_n}{\overline{\boldsymbol{a}}_n^{\mathrm{H}}(\boldsymbol{R}_n + \epsilon\boldsymbol{I}_M)^{-1}\overline{\boldsymbol{a}}_n}$$

where $\boldsymbol{R}_n = 1/(KT)\sum_{k=1}^{K}\sum_{t=1}^{T}\boldsymbol{i}_{n,k,t}\boldsymbol{i}_{n,k,t}^{\mathrm{H}}$ can be seen as the auto-correlation matrix of the interference plus noise component (IpNC) $\boldsymbol{i}_{n,k,t} \triangleq \boldsymbol{y}_{k,t} - \widetilde{\boldsymbol{G}}_{n,k}\boldsymbol{x}_{n,t}$, and ϵ is a diagonal loading factor.

3.4.2 Integration of beamforming and compressed sensing

The received signal over a frame can be represented in matrix form by,

$$Y = \sum\nolimits_{n\in\aleph}\widetilde{\boldsymbol{G}}_n\boldsymbol{X}_n + \boldsymbol{V}$$

where the tth column vector of $\boldsymbol{X}_n$ is $\boldsymbol{x}_{n,t}$ and the $t$thcolumn of $\boldsymbol{V}$ is $\boldsymbol{v}_t$. Similarly, extending $\boldsymbol{y}_{n,t}$ in one frame yields,

$$\boldsymbol{Y}_n = (\boldsymbol{I}_K \ast \boldsymbol{b}_n)^{\mathrm{H}}\boldsymbol{Y} = \boldsymbol{B}_{n,n}\boldsymbol{X}_n + \sum\nolimits_{l\in\aleph\backslash n}\boldsymbol{B}_{n,l}\boldsymbol{X}_l + (\boldsymbol{I}_K \ast \boldsymbol{b}_n)^{\mathrm{H}}\boldsymbol{V}.$$

To use the frame-wise sparsity of the user activity, we have

$$\boldsymbol{\eta}_n = \boldsymbol{D}_n\boldsymbol{c}_n + \boldsymbol{z}_n$$

where $\boldsymbol{\eta}_n = \mathrm{vec}\{\boldsymbol{Y}_n^{\mathrm{T}}\}$, $\boldsymbol{D}_n = \boldsymbol{B}_{n,n} \ast \boldsymbol{I}_T$ and $\boldsymbol{c}_n = \mathrm{vec}\{\boldsymbol{X}_n^{\mathrm{T}}\}$. $\boldsymbol{z}_n$ is the IpNC under beamforming.

A general framework for the integration of adaptive beamforming (ABF) and CS for uplink grant-free access for any user cluster n is illustrated by Figure 14. The block-sparsity based adaptive SP (ASP) method is considered in the CS module.

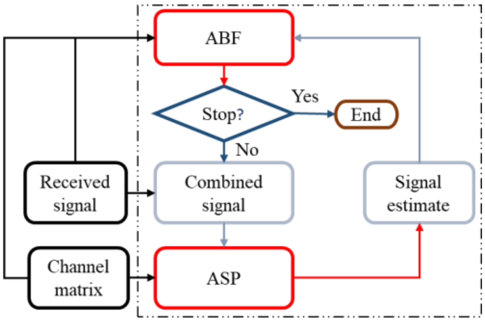
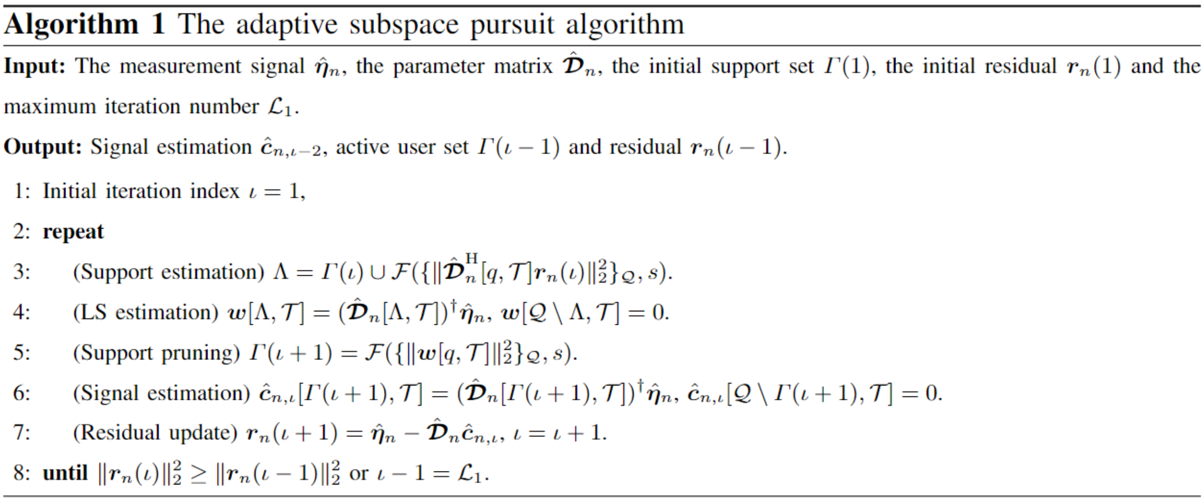


Figure 14: Block diagram of the joint adaptive beamforming and subspace pursuit method (J-ABF-SP): ABF (adaptive beamforming), ASP (adaptive subspace pursuit)



Algorithm 1 The adaptive subspace pursuit algorithm

Input: The measurement signal $\hat{\boldsymbol{\eta}}_n$, the parameter matrix $\hat{\boldsymbol{\mathcal{D}}}_n$, the initial support set $\mathit{\Gamma}(1)$, the initial residual $\boldsymbol{r}_n(1)$ and the maximum iteration number $\mathcal{L}_1$.

Output: Signal estimation $\hat{\boldsymbol{c}}_{n,\iota-2}$, active user set $\mathit{\Gamma}(\iota-1)$ and residual $\boldsymbol{r}_n(\iota-1)$.

```
Initial iteration index $\iota = 1$,
repeat
    (Support estimation) $\Lambda = \mathit{\Gamma}(\iota) \cup \mathcal{F}(\{\|\hat{\boldsymbol{\mathcal{D}}}_n^{\mathrm{H}}[q,\mathcal{T}]\boldsymbol{r}_n(\iota)\|_2^2\}_{\mathcal{Q}}, s)$.
    (LS estimation) $\boldsymbol{w}[\Lambda,\mathcal{T}] = (\hat{\boldsymbol{\mathcal{D}}}_n[\Lambda,\mathcal{T}])^{\dagger}\hat{\boldsymbol{\eta}}_n$, $\boldsymbol{w}[\mathcal{Q}\setminus\Lambda,\mathcal{T}] = 0$.
    (Support pruning) $\mathit{\Gamma}(\iota+1) = \mathcal{F}(\{\|\boldsymbol{w}[q,\mathcal{T}]\|_2^2\}_{\mathcal{Q}}, s)$.
    (Signal estimation) $\hat{\boldsymbol{c}}_{n,\iota}[\mathit{\Gamma}(\iota+1),\mathcal{T}] = (\hat{\boldsymbol{\mathcal{D}}}_n[\mathit{\Gamma}(\iota+1),\mathcal{T}])^{\dagger}\hat{\boldsymbol{\eta}}_n$, $\hat{\boldsymbol{c}}_{n,\iota}[\mathcal{Q}\setminus\mathit{\Gamma}(\iota+1),\mathcal{T}] = 0$.
    (Residual update) $\boldsymbol{r}_n(\iota+1) = \hat{\boldsymbol{\eta}}_n - \hat{\boldsymbol{\mathcal{D}}}_n\hat{\boldsymbol{c}}_{n,\iota}$, $\iota = \iota + 1$.
until $\|\boldsymbol{r}_n(\iota)\|_2^2 \geq \|\boldsymbol{r}_n(\iota-1)\|_2^2$ or $\iota - 1 = \mathcal{L}_1$.
```

Based on the beamforming weight $\hat{\boldsymbol{b}}_n$ which is initialised by the SBF weight $\boldsymbol{b}_n^{\mathrm{SBF}}$ before the first iteration, the measurements (combined signals) for the ASP are generated by (step 12 in Algorithm 2),

$$\hat{\boldsymbol{Y}}_n = (\boldsymbol{I}_K * \hat{\boldsymbol{b}}_n)^{\mathrm{H}}\boldsymbol{Y},\ \ \hat{\boldsymbol{\eta}}_n = \mathrm{vec}\{\hat{\boldsymbol{Y}}_n^{\mathrm{T}}\}$$

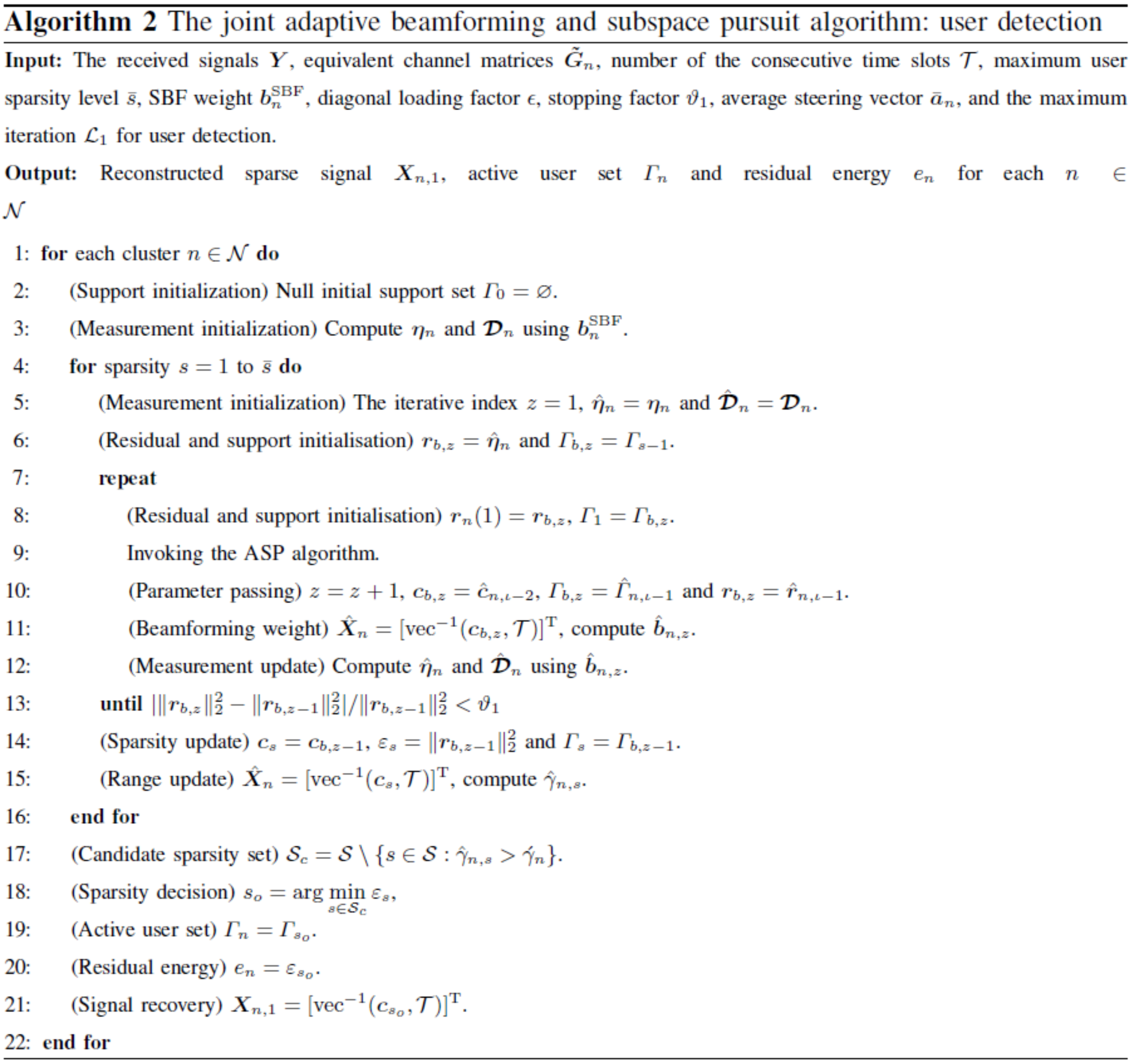
We also have (step 12 in Algorithm 2 and step 8 in Algorithm 3),

$$\hat{\boldsymbol{B}}_{n,n} = (\boldsymbol{I}_K * \hat{\boldsymbol{b}}_n)^{\mathrm{H}}\tilde{\boldsymbol{G}}_n,\ \ \hat{\boldsymbol{D}}_n = \hat{\boldsymbol{B}}_{n,n} * \boldsymbol{I}_T.$$

With the estimated active user set $\hat{\boldsymbol{\Gamma}}_{n,\iota+1}$ at the ιth iteration, the signal is estimated by,

$$\hat{\boldsymbol{c}}_n\big[\hat{\boldsymbol{\Gamma}}_{n,\iota+1}, T\big] = \big(\hat{\boldsymbol{D}}_n\big[\hat{\boldsymbol{\Gamma}}_{n,\iota+1}, T\big]\big)^{+}\hat{\boldsymbol{\eta}}_n$$

$$\hat{\boldsymbol{c}}_n\big[Q\backslash\hat{\boldsymbol{\Gamma}}_{n,\iota+1}, T\big] = 0$$



Algorithm 2 The joint adaptive beamforming and subspace pursuit algorithm: user detection

Input: The received signals $\boldsymbol{Y}$, equivalent channel matrices $\tilde{\boldsymbol{G}}_n$, number of the consecutive time slots $\mathcal{T}$, maximum user sparsity level $\bar{s}$, SBF weight b_n^{SBF}, diagonal loading factor ϵ, stopping factor ϑ_1, average steering vector $\bar{a}_n$, and the maximum iteration $\mathcal{L}_1$ for user detection.

Output: Reconstructed sparse signal $\boldsymbol{X}_{n,1}$, active user set $\mathit{\Gamma}_n$ and residual energy e_n for each $n \in \mathcal{N}$

```
1:  for each cluster $n \in \mathcal{N}$ do
2:      (Support initialization) Null initial support set $\mathit{\Gamma}_0 = \varnothing$.
3:      (Measurement initialization) Compute $\boldsymbol{\eta}_n$ and $\boldsymbol{\mathcal{D}}_n$ using $b_n^{\mathrm{SBF}}$.
4:      for sparsity $s = 1$ to $\bar{s}$ do
5:          (Measurement initialization) The iterative index $z = 1$, $\hat{\boldsymbol{\eta}}_n = \boldsymbol{\eta}_n$ and $\hat{\boldsymbol{\mathcal{D}}}_n = \boldsymbol{\mathcal{D}}_n$.
6:          (Residual and support initialisation) $\boldsymbol{r}_{b,z} = \hat{\boldsymbol{\eta}}_n$ and $\mathit{\Gamma}_{b,z} = \mathit{\Gamma}_{s-1}$.
7:          repeat
8:              (Residual and support initialisation) $\boldsymbol{r}_n(1) = \boldsymbol{r}_{b,z}$, $\mathit{\Gamma}_1 = \mathit{\Gamma}_{b,z}$.
9:              Invoking the ASP algorithm.
10:             (Parameter passing) $z = z + 1$, $\boldsymbol{c}_{b,z} = \hat{\boldsymbol{c}}_{n,\iota-2}$, $\mathit{\Gamma}_{b,z} = \hat{\mathit{\Gamma}}_{n,\iota-1}$ and $\boldsymbol{r}_{b,z} = \hat{\boldsymbol{r}}_{n,\iota-1}$.
11:             (Beamforming weight) $\hat{\boldsymbol{X}}_n = [\mathrm{vec}^{-1}(\boldsymbol{c}_{b,z}, \mathcal{T})]^{\mathrm{T}}$, compute $\hat{\boldsymbol{b}}_{n,z}$.
12:             (Measurement update) Compute $\hat{\boldsymbol{\eta}}_n$ and $\hat{\boldsymbol{\mathcal{D}}}_n$ using $\hat{\boldsymbol{b}}_{n,z}$.
13:         until $|\|\boldsymbol{r}_{b,z}\|_2^2 - \|\boldsymbol{r}_{b,z-1}\|_2^2| / \|\boldsymbol{r}_{b,z-1}\|_2^2 < \vartheta_1$
14:         (Sparsity update) $\boldsymbol{c}_s = \boldsymbol{c}_{b,z-1}$, $\varepsilon_s = \|\boldsymbol{r}_{b,z-1}\|_2^2$ and $\mathit{\Gamma}_s = \mathit{\Gamma}_{b,z-1}$.
15:         (Range update) $\hat{\boldsymbol{X}}_n = [\mathrm{vec}^{-1}(\boldsymbol{c}_s, \mathcal{T})]^{\mathrm{T}}$, compute $\hat{\gamma}_{n,s}$.
16:     end for
17:     (Candidate sparsity set) $\mathcal{S}_c = \mathcal{S} \setminus \{s \in \mathcal{S} : \hat{\gamma}_{n,s} > \acute{\gamma}_n\}$.
18:     (Sparsity decision) $s_o = \arg\min_{s \in \mathcal{S}_c} \varepsilon_s$,
19:     (Active user set) $\mathit{\Gamma}_n = \mathit{\Gamma}_{s_o}$.
20:     (Residual energy) $e_n = \varepsilon_{s_o}$.
21:     (Signal recovery) $\boldsymbol{X}_{n,1} = [\mathrm{vec}^{-1}(\boldsymbol{c}_{s_o}, \mathcal{T})]^{\mathrm{T}}$.
22: end for
```

where Q is the set of user indices for any cluster and $(\cdot)^{+}$ denotes the Moore–Penrose inverse. The vector $\hat{\boldsymbol{c}}_n[q,T]$ as the qth $T \times 1$ vector block of $\hat{\boldsymbol{c}}_n$ and denote the matrix $\widehat{\boldsymbol{D}}_n[q,T]$ as the matrix block of $\widehat{\boldsymbol{D}}_n$ constitued by consecutive columns from index $(q-1)T+1$ to index qT. Furthermore, $\hat{\boldsymbol{c}}_n\left[\widehat{\boldsymbol{\Gamma}}_{n,\iota+1}, T\right]$ and $\widehat{\boldsymbol{D}}_n\left[\widehat{\boldsymbol{\Gamma}}_{n,\iota+1}, T\right]$ denote the sub-vector and sub-matrix by selecting their respective blocks according to the indices from the set $\widehat{\boldsymbol{\Gamma}}_{n,\iota+1}$. To sum up, the detailed steps of the ASP algorithm are summarised in Algorithm 1. The finding function $F(V,\zeta)$ selects the indices of the first ζ largest elements of an ordered set/vector V.

Subsequently, with the output $\widehat{\boldsymbol{X}}_n = \mathrm{vec}^{-\mathbf{1}}\{\hat{\boldsymbol{c}}_n, T\}$ of the ASP, the IpNC is estimated by $\hat{\boldsymbol{\imath}}_{n,k,t} \triangleq \boldsymbol{y}_{k,t} - \widetilde{\boldsymbol{G}}_{n,k}\widehat{\boldsymbol{x}}_{n,t}$ where $\widehat{\boldsymbol{x}}_{n,t}$ is the tth column of $\widehat{\boldsymbol{X}}_n$. The beamforming weight is accordingly updated by (step 11 in Algorithm 2),

$$\widehat{\boldsymbol{b}}_n = \frac{(\widehat{\boldsymbol{R}}_n + \epsilon \boldsymbol{I}_M)^{-1}\overline{\boldsymbol{a}}_n}{\overline{\boldsymbol{a}}_n^{\mathrm{H}}(\widehat{\boldsymbol{R}}_n + \epsilon \boldsymbol{I}_M)^{-1}\overline{\boldsymbol{a}}_n}$$

where $\widehat{\boldsymbol{R}}_n = 1/(KT)\sum_{k=1}^{K}\sum_{t=1}^{T}\hat{\boldsymbol{\imath}}_{n,k,t}\hat{\boldsymbol{\imath}}_{n,k,t}^{\mathrm{H}}$.

Algorithm 3 The interference cancellation enhanced signal recovery

```
Input: The received signals $\boldsymbol{Y}$, equivalent channel matrices $\tilde{\boldsymbol{G}}_n$, number of the consecutive time slots $\mathcal{T}$, diagonal loading factor $\epsilon$, average steering vector $\bar{a}_n$, maximum number of iterations $\mathcal{L}_2$ and $\mathcal{L}_3$, active user set $\Gamma_n$, initial error $e_n$ and initial signal estiamtion $\boldsymbol{X}_{n,1}$.
Output: Reconstructed sparse signal $\boldsymbol{X}_n$
(Weight initialisation) For each cluster $n$, $\hat{\boldsymbol{X}}_n = \boldsymbol{X}_{n,1}$, $\hat{\boldsymbol{\imath}}_{n,k,t} = \boldsymbol{y}_{k,t} - \tilde{\boldsymbol{G}}_{n,k}\hat{\boldsymbol{x}}_{n,t}$, compute $\hat{\boldsymbol{b}}_n$.
(Error initialisation) For each cluster $n$, $\tilde{e}_{1,n} = e_n$.
for Iteration $\iota_2 = 1$ to $\mathcal{L}_2$ do
    for Cluster $n = 1$ to $N$ do
        (Interference reconstruction) construct the received interference signal $\boldsymbol{Y}_{\imath,n} = \sum_{l=1,l\neq n}^{N} \tilde{\boldsymbol{G}}_l \boldsymbol{X}_{l,\iota_2}$.
        (Interference cancellation) $\mathcal{Y}_n = \boldsymbol{Y} - \boldsymbol{Y}_{\imath,n}$.
        for Iteration $\iota_3 = 1$ to $\mathcal{L}_3$ do
            (Measurement update) Compute $\hat{\eta}_n$ and $\hat{\boldsymbol{\mathcal{D}}}_n$ using $\hat{\boldsymbol{b}}_n$.
            (Signal estimation) $\hat{c}_n[\Gamma_n,\mathcal{T}] = (\hat{\boldsymbol{\mathcal{D}}}_n[\Gamma_n,\mathcal{T}])^{\dagger}\hat{\eta}_n$, $\hat{c}_n[\mathcal{Q}\setminus\Gamma_n,\mathcal{T}] = 0$.
            (Residual update) $\tilde{e}_{\iota_3+1,n} = \|\hat{\eta}_n - \hat{\boldsymbol{\mathcal{D}}}_n\hat{c}_n\|_2^2$.
            if $\tilde{e}_{\iota_3+1,n} < \tilde{e}_{\iota_3,n}$ and $\iota_3 < \mathcal{L}_3$ then
                (Beamforming weight) $\hat{\boldsymbol{X}}_n = [\mathrm{vec}^{-1}(\hat{c}_n,\mathcal{T})]^{\mathrm{T}}$, $\hat{\boldsymbol{\imath}}_{n,k,t} = \boldsymbol{y}_{k,t} - \tilde{\boldsymbol{G}}_{n,k}\hat{\boldsymbol{x}}_{n,t}$, and then compute $\hat{\boldsymbol{b}}_n$.
            else
                (Residual modification) $\tilde{e}_{1,n} = \tilde{e}_{\iota_3,n}$.
                break.
            end if
        end for
        (Signal update) $\boldsymbol{X}_{n,\iota_2+1} = \hat{\boldsymbol{X}}_n$.
    end for
end for
(Signal recovery) $\boldsymbol{X}_n = \boldsymbol{X}_{n,\mathcal{L}_2+1}$.
```

A method with two steps is proposed for determining the user sparsity level.

1) The candidate sparsity set $S_c = S\backslash\{s \in S: \hat{\gamma}_{n,s} > \acute{\gamma}_n\}$ with $S = \{1,2,\dots,\bar{s}\}$.

2) The sparsity is given by $s_o = \underset{s\in S_c}{\operatorname{argmin}}\left\|\hat{\boldsymbol{\eta}}_n - \hat{\boldsymbol{D}}_n\hat{\boldsymbol{c}}_n\right\|_2^2$

Note that $\hat{\gamma}_{n,s}$ is the temporal power ratio of the estimated transmitted signals with given sparsity level *s*, defined as (step 15 in Algorithm 2),

$$\hat{\gamma}_{n,s} \triangleq \frac{\max_{q\in\Gamma_s}\left\|\hat{\boldsymbol{x}}_{n,q}\right\|_2^2}{\min_{q\in\Gamma_s}\left\|\hat{\boldsymbol{x}}_{n,q}\right\|_2^2}$$

where Γ_s is the active user set estimate with given sparsity level *s and* $\hat{\boldsymbol{x}}_{n,q} = \hat{\boldsymbol{c}}_n[q,T]$ is a block vector of the above signal estimate. $\acute{\gamma}_n$ is a threshold which can be an empirical value larger than 1 when considering the same transmission power for the active users in the same cluster.

To sum up, a joint adaptive beamforming and subspace pursuit algorithm (J-ABF-SP) is presented in Algorithm 2 [39].

An interference cancellation (IC) scheme is also developed in Algorithm 3 to improve the signal estimation based on the active user set and the initial signal estimates from the J-ABF-SP algorithm.

With the active user set and initial signal estimates from the J-ABF-SP algorithm, the received signal from each cluster n can be reconstructed as $\widetilde{\boldsymbol{G}}_n \boldsymbol{X}_{n,\iota}$, where $\boldsymbol{X}_{n,\iota+1}$ is the signal estimate after the ιth IC. Then, we can obtain the IC-enabling received signal for cluster *n*, i.e.,

$$\boldsymbol{F}_n = Y - \sum_{l \in \aleph \backslash n} \widetilde{\boldsymbol{G}}_l \boldsymbol{X}_{l,\iota}.$$

Then, the new measurements are generated by (step 8 in Algorithm 3),

$$\widehat{\boldsymbol{Y}}_n = (\boldsymbol{I}_K * \widehat{\boldsymbol{b}}_n)^{\mathrm{H}} \boldsymbol{F}_n, \;\; \widehat{\boldsymbol{\eta}}_n = \mathrm{vec}\{\widehat{\boldsymbol{Y}}_n^{\mathrm{T}}\}$$

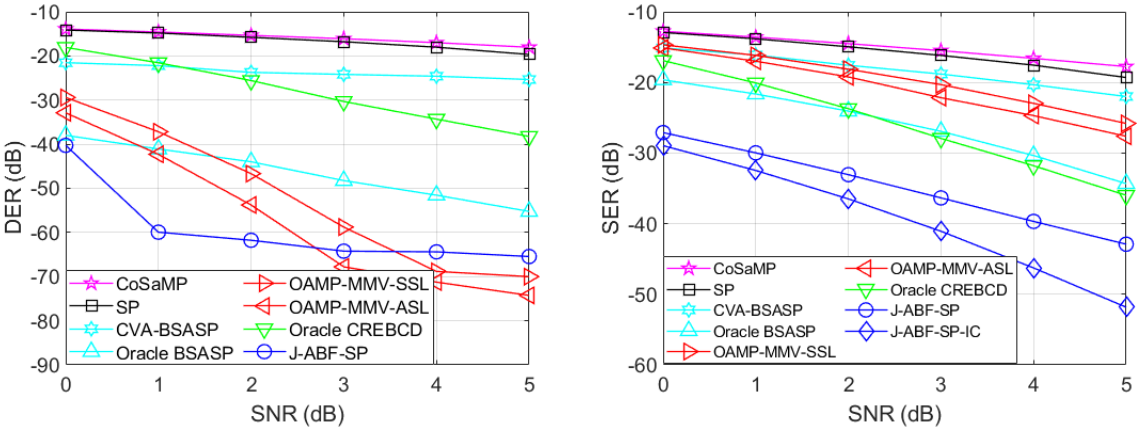
The algorithms 2 and 3 are referred to as the IC-enhanced joint adaptive beamforming and subspace pursuit algorithm (J-ABF-SP-IC).

3.5 Simulation implementation

In this section, the MUD and DR performance of the proposed J-ABF-SP algorithms will be assessed through simulations, in comparison with some benchmark methods, including the CoSaMP, SP, CVA-BSASP, Oracle-BSASP, Oracle-CREBCD, OAMP-MMV-SSL and the OAMP-MMV-ASL methods. In particular, the CoSaMP, SP and the Oracle methods are evaluated with known user sparsity levels. Without loss of generality, we consider the transmitted symbols randomly generated from 16QAM constellation for all the users. For the benchmark algorithms, we consider the single-antenna (e.g., the first antenna) reception of any one user cluster as the received signal. For the proposed algorithms, we select 3 as the sparsity decision threshold for each cluster *n*.

A BS with M antenna elements is considered, serving massive users simultaneously. The users are assumed to be grouped based on the channel correlation into $N \leq M$ clusters with Q users in each cluster n, n = 1, 2,... ,N. Assume the AoAs of the users in each cluster are randomly distributed over an angle range with a width of 5 degrees. Without loss of generality, we consider N = 3 and Q = 40, with the central angles of the three clusters being -30, -10 and 10 degrees, respectively. All users employ the common K = 20 subcarriers with the same spreading signatures utilised in all the N clusters. In this case, the frequency-domain system overloading factor is $\frac{NQ}{K} = 600\%$, which increases linearly with the number of user clusters. We consider the user activity rate to be $\alpha = 10\%$, i.e., the number of active users so = 4 in each cluster, which is far less than the number of the total users. Each frame consists of $T = 7$ continuous symbol durations, following the LTE-Advanced standard.

The detection error rate (DER) and the symbol error rate (SER) are considered as performance metrics. For any cluster n, the DER is defined as $p_{d,n} = (f_n + m_n)/Q$ where f_n and m_n denote the number of falsely detected inactive users and the number of missed active users, respectively. The SER is defined as $p_{s,n} = p_{d,n} + S_{e,n}/(QT)$ where $S_{e,n}$ denotes the number of error symbols of detected active users. Both the DER and SER are calculated over a large number of independent trials. In the following, we consider the same input SNR for each user cluster n and present the average values of the DERs or SERs of the N clusters, unless noted otherwise. The ESNR = 13 dB for the SBF and the number of antennas *M* = 5, unless specified otherwise.

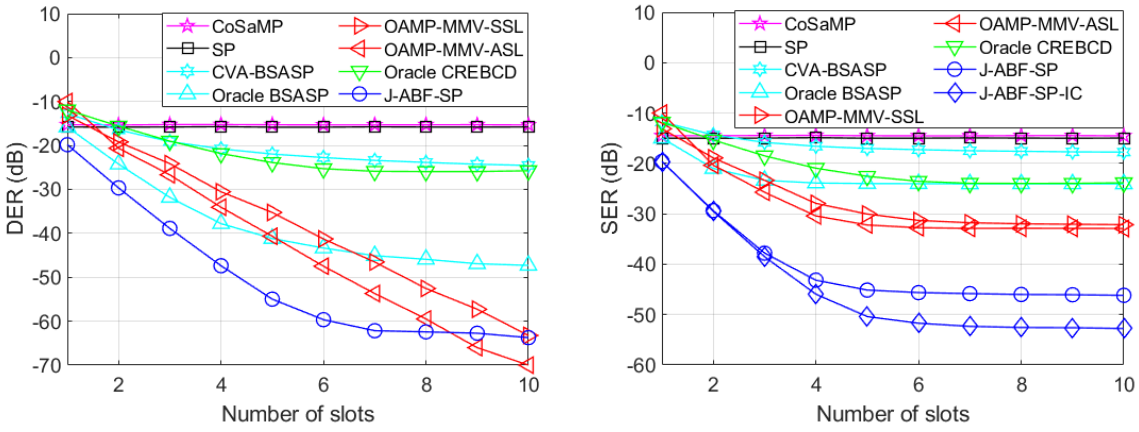


(a) (b)

Figure 15: MUD performance with the input SNR: (a) DER; (b) SER

Figure 15(a) shows the DERs regarding the input SNRs for different MUD methods. The proposed J-ABF-SP algorithm performs better in user detection than the Oracle-BSASP algorithm even though the latter knows the user sparsity level a priori. This is because both the SBF and ABF used in the J-ABF-SP can suppress the IpNC contained in the received signal, leading to a higher receiver signal-to-interference-plus-noise ratio (SINR) than that of the Oracle-BSASP. We also observe that the J-ABF-SP algorithm can achieve extremely low DERs even at low SNRs, e.g., -60 dB DER under the 1 dB SNR. In this regard, it does not matter that the J-ABF-SP presents a slightly higher DER than that of the OAMP-MMV algorithms as the SNR increases to a certain value, e.g., 4 dB. Additionally, the results show that the J-ABF-SP can achieve a more than 25 dB gain in DER performance in comparison with the other benchmark algorithms.

Figure 15(b) plots the SERs regarding the input SNRs for different MUD methods. It shows that the proposed J-ABF-SP algorithm presents a more than 8 dB gain in SER performance in comparison to the OAMP-MMV algorithms when the SNR is higher than 0 dB and performs even much better than the other benchmark algorithms. In addition, the J-ABF-SP-IC algorithm outperforms the J-ABF-SP due to the IC improving the SINR at the receiver.

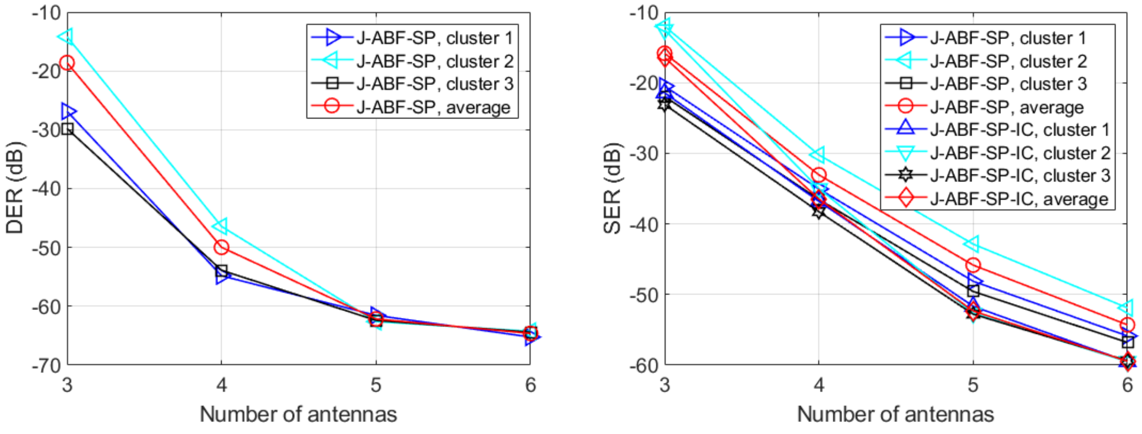


(a) (b)

Figure 16: MUD performance with the number of slots: (a) DER; (b) SER

Figure 16 illustrates the DERs and the SERs with respect to the number of slots. We consider SNR=2 dB for each cluster with other conditions unchanged. The CoSaMP and SP algorithms perform the

MUD and DR by slot, so their performance remains nearly unchanged with the number of slots. With only one slot, the Oracle-BSASP algorithm degenerates into the SP algorithm, resulting in the same DER and SER performance. The proposed algorithms achieve significantly low DERs and SERs compared to the benchmark algorithms, even with only one slot in a frame. Moreover, the performance enhancement by the proposed algorithms tends to increase with the number of slots and eventually converges. In particular, compared with the OAMP-MMV algorithms, the J-ABF-SP algorithm shows slightly inferior DER performance when the number of slots increases to 9, but demonstrates remarkable superiority in SER performance.



(a) (b)

Figure 17: MUD performance with the number of antennas: (a) DER; (b) SER

The impact of the number of antennas M on the performance of the proposed algorithms is studied below. SNR=2 dB for every cluster and 7 slots in a frame are considered, with other conditions unchanged. Figure 17 illustrate the DER and SER of each cluster with respect to the number of antennas, respectively. The DERs of all clusters gradually decrease with the number of antennas. Specifically, the DER of cluster 2 is initially higher than those of the other two clusters with a small number of antennas, but approaches a similar value with the increased number of antennas. This is because cluster 2 is located spatially between the other two clusters and thus suffers from larger interferences, but this impact is mitigated with the enhanced beamforming gain and spatial resolution provided by the increased number of antennas. Similarly, more antennas result in better SERs and smaller SER differences among different clusters. In addition, J-ABF-SP-IC outperforms J-ABF-SP in SER performance. Moreover, the SER performance is enhanced by more than 20 dB by increasing the number of antennas from 4 to 6, indicating a promising prospect for the integration of SDMA and CS for uplink grant-free communication.

3.6 Conclusions

A novel architecture for the integration of the intelligent beamforming and compressed sensing technique is proposed for the grant-free NOMA in a MIMO cellular network. In particular, a joint adaptive beamforming and subspace pursuit method is developed for the blind multiuser detection and data recovery with the help of a novel user sparsity decision method. With the beamforming, the spectral efficiency is improved by supporting more users in the same frequency spectrum. In addition, both the detection error rate and the symbol error rate can be significantly reduced with the integration and data interaction between the beamforming and subspace pursuit methods.

3.7 References

[1] R. Hoshyar, F. P. Wathan, and R. Tafazolli, “Novel low-density signature for synchronous CDMA systems over AWGN channel,” IEEE Trans. Signal Process., vol. 56, no. 4, pp. 1616–1626, 2008.

[2] H. Nikopour and H. Baligh, “Sparse code multiple access,” in 2013 IEEE 24th Ann. Int. Symp. Personal, Indoor, and Mobile Radio Commun. (PIMRC), 2013, pp. 332–336.

[3] S. Zhang, X. Xu, L. Lu, Y. Wu, G. He, and Y. Chen, “Sparse code multiple access: An energy efficient uplink approach for 5G wireless systems,” in 2014 IEEE Global Commun. Conf., 2014, pp. 4782–4787.

[4] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-s. Kwak, “Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges,” IEEE Commun. Surveys & Tutorials, vol. 19, no. 2, pp. 721–742, 2017.

[5] S. Kusaladharma, W.-P. Zhu, W. Ajib, and G. A. A. Baduge, “Achievable rate characterization of NOMA-aided cell-free massive mimo with imperfect successive interference cancellation,” IEEE Trans. Commun., vol. 69, no. 5, pp. 3054–3066, 2021.

[6] H. Shariatmadari, R. Ratasuk, S. Iraji, A. Laya, T. Taleb, R. J¨antti, and A. Ghosh, “Machine-type communications: current status and future perspectives toward 5G systems,” IEEE Commun. Mag., vol. 53, no. 9, pp. 10–17, 2015.

[7] Lenovo, “Uplink grant-free access for 5G mMTC,” 3GPP document R1-1609398, TSG-RAN WG1 Meeting #86b, October 2016.

[8] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, “Toward massive machine type cellular communications,” IEEE Wireless Commun., vol. 24, no. 1, pp. 120–128, 2017.

[9] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic, and E. de Carvalho, “Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the internet of things,” IEEE Signal Process. Mag., vol. 35, no. 5, pp. 88–99, 2018.

[10] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, “Grant-free non-orthogonal multiple access for IoT: A survey,” IEEE Commun. Surveys and Tutorials, vol. 22, no. 3, pp. 1805–1838, 2020.

[11] L. Qiao, J. Zhang, Z. Gao, D. Zheng, M. J. Hossain, Y. Gao, D. W. K. Ng, and M. Di Renzo, “Joint activity and blind information detection for UAV-assisted massive IoT access,” IEEE J. Sel. Areas Commun., vol. 40, no. 5, pp. 1489–1508, 2022.

[12] K. Senel and E. G. Larsson, “Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach,” IEEE Trans. Commun., vol. 66, no. 12, pp. 6164–6175, 2018.

[13] L. Liu and W. Yu, “Massive connectivity with massive mimo—part i: Device activity detection and channel estimation,” IEEE Trans. Signal Process., vol. 66, no. 11, pp. 2933–2946, 2018.

[14] J. Ahn, B. Shim, and K. B. Lee, “EP-based joint active user detection and channel estimation for massive machine-type communications,” IEEE Trans. Commun., vol. 67, no. 7, pp. 5178–5189, 2019.

[15] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, “Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO,” IEEE Trans. Signal Process., vol. 68, pp. 764–779, 2020.

[16] C. Wei, H. Liu, Z. Zhang, J. Dang, and L. Wu, “Approximate message passing-based joint user activity and data detection for NOMA,” IEEE Commun. Lett., vol. 21, no. 3, pp. 640–643, 2017.

[17] Y. Du, C. Cheng, B. Dong, Z. Chen, X. Wang, J. Fang, and S. Li, “Block-sparsity-based multiuser detection for uplink grant-free NOMA,” IEEE Trans. Wireless Commun., vol. 17, no. 12, pp. 7894–7909, 2018.

[18] P. Gao, Z. Liu, P. Xiao, C. H. Foh, and J. Zhang, “Low-complexity block coordinate descend based multiuser detection for uplink grant-free NOMA,” IEEE Trans. Veh. Technology, pp. 1–1, 2022.

[19] Y. Mei, Z. Gao, Y. Wu, W. Chen, J. Zhang, D. W. K. Ng, and M. Di Renzo, “Compressive sensing-based joint activity and data detection for grant-free massive IoT access,” IEEE Trans. Wireless Commun., vol. 21, no. 3, pp. 1851–1869, 2022.

[20] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” IEEE Trans. Inf. Theory, vol. 53, no. 12, pp. 4655–4666, 2007.

[21] D. Guo and C.-c. Wang, “Multiuser detection of sparsely spread CDMA,” IEEE J. Sel. Areas Commun., vol. 26, no. 3, pp. 421–431, 2008.

[22] D. Needell and J. Tropp, “CoSaMP: iterative signal recovery from incomplete and inaccurate samples,” Communications of the ACM, vol. 53, no. 12, pp. 93–100, 2010.

[23] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” IEEE Trans. Inf. Theory, vol. 55, no. 5, pp. 2230–2249, 2009.

[24] D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” Proceedings of the National Academy of Sciences, vol. 106, no. 45, pp. 18 914–18 919, 2009.

[25] Q. Luo, Z. Liu, G. Chen, Y. Ma, and P. Xiao, “A novel multitask learning empowered codebook design for downlink SCMA networks,” IEEE Wireless Commun. Lett., vol. 11, no. 6, pp. 1268–1272, 2022.

[26] Q. Luo, H. Wen, G. Chen, Z. Liu, P. Xiao, Y. Ma, and A. Maaref, “A novel non-coherent SCMA with massive MIMO,” IEEE Wireless Commun. Lett., pp. 1–1, 2022.

[27] B. Wang, L. Dai, Y. Zhang, T. Mir, and J. Li, “Dynamic compressive sensing-based multi-user detection for uplink grant-free NOMA,” IEEE Commun. Lett., vol. 20, no. 11, pp. 2320–2323, 2016.

[28] Y. Du, B. Dong, Z. Chen, X. Wang, Z. Liu, P. Gao, and S. Li, “Efficient multi-user detection for uplink grant-free NOMA: Prior-information aided adaptive compressive sensing perspective,” IEEE J. Sel. Areas Commun., vol. 35, no. 12, pp. 2812–2828, 2017.

[29] T. Li, J. Zhang, Z. Yang, Z. L. Yu, Z. Gu, and Y. Li, “Dynamic user activity and data detection for grant-free NOMA via weighted $l_{2,1}$ minimization,” IEEE Trans. Wireless Commun., vol. 21, no. 3, pp. 1638–1651, 2022.

[30] L. Wu, Z. Wang, P. Sun, and Y. Yang, “Temporal correlation enhanced sparse activity detection in MIMO enabled grant-free noma,” IEEE Trans. Veh. Technology, vol. 71, no. 3, pp. 2887–2899, 2022.

[31] L. Wu, P. Sun, Z. Wang, and Y. Yang, “Joint user activity identification and channel estimation for grant-free NOMA: A spatial–temporal structure-enhanced approach,” IEEE Internet of Things J., vol. 8, no. 15, pp. 12 339–12 349, 2021.

[32] Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, and Z. Zhong, “Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems,” IEEE Trans. Signal Process., vol. 65, no. 18, pp. 4874–4886, 2017.

[33] J. Cui, Z. Ding, P. Fan, and N. Al-Dhahir, “Unsupervised machine learning-based user clustering in millimeter wave-NOMA systems,” IEEE Trans. Wireless Commun., vol. 17, no. 11, pp. 7425–7440, 2018.

[34] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X.-G. Xia, “Joint Tx-Rx beamforming and power allocation for 5G millimeter-wave non-orthogonal multiple access networks,” IEEE Trans. Commun., vol. 67, no. 7, pp. 5114–5125, 2019.

[35] K. Senel, H. V. Cheng, E. Bj¨ornson, and E. G. Larsson, “What role can NOMA play in massive MIMO?” IEEE J. Sel. Topics Signal Process., vol. 13, no. 3, pp. 597–611, 2019.

[36] H. Zhang, H. Zhang, W. Liu, K. Long, J. Dong, and V. C. M. Leung, “Energy efficient user clustering, hybrid precoding and power optimization in terahertz MIMO-NOMA systems,” IEEE J. Select. Areas Commun., vol. 38, no. 9, pp. 2074–2085, 2020.

[37] G. Xia, Y. Zhang, L. Ge, and H. Zhou, “Deep reinforcement learning based dynamic power allocation for uplink device-to-device enabled cell-free network,” in 2022 IEEE Int. Symp. Broadband Multimedia Syst. and Broadcast. (BMSB), 2022, pp. 01–06.

[38] Q. N. Le, V.-D. Nguyen, O. A. Dobre, N.-P. Nguyen, R. Zhao, and S. Chatzinotas, “Learning-assisted user clustering in cell-free massive MIMO-NOMA networks,” IEEE Trans. Veh. Technol., vol. 70, no. 12, pp. 12 872–12 887, 2021.

[39] G. Xia, P. Xiao, B. Li, et al. Joint beamforming and compressed sensing for uplink grant-free access[J]. 2023.

4 AI Resource Allocation for the Cell-free Network with Underlaid D2D

4.1 Related Works and Motivation

4.1.1 Related Works

Cell-free massive MIMO (CF-mMIMO) is a novel network architecture that can overcome the boundary limitation of traditional cell-based networks and provide consistent and robust communication service for all users within a coverage area [1]. However, CF-mMIMO involves many distributed access points (APs), which makes the resource allocation (including power allocation and beamforming design) optimization challenging, especially when CF-mMIMO coexists with underlaid D2D communications.

Machine learning (ML) has emerged as a promising technique for solving the resource allocation problem in wireless communications, compared with the convex optimization-based methods. For example, in [2], deep reinforcement learning (DRL)-based power allocation was employed to solve the max-min and spectral efficiency (SE) maximization problems in CF-mMIMO. In [3], deep CNN was utilized to maximize SE in CF-mMIMO, and it was shown that the learning-based method can outperform the well-known use-and-then-forget-based PA method. In [4], a clustered deep neural network (DNN) model that leverages the large-scale fading information was proposed to implement power allocation in CF-mMIMO, and its performance was comparable to that of the weighted minimum mean square error (WMMSE) -assisted alternating direction method.

The learning-based resource allocation methods mentioned above have demonstrated their superiority to the traditional methods in terms of performance and complexity. However, none of them have exploited the structures of wireless networks in the optimization process. Therefore, they cannot handle the unseen scenarios, such as the networks with varying sizes. To address this issue, the graph neural network (GNN) that offers the benefits of scalability, generalization and parallel execution has attracted considerable research interests in wireless communications. Specifically, in [5], a wireless channel graph convolution network (WCGCN) was developed for solving the power allocation and beamforming problem in device-to-device (D2D) communication systems. It was shown that the WCGCN trained on small-sized systems can be generalized to optimize large systems with higher density of devices and larger cell size. Moreover, in [6] and [7], the GNN-assisted power allocation in ad hoc networks was studied and the proposed methods significantly outperformed the conventional methods due to the GNN's generalizability. However, the GNN models for ad hoc or D2D networks mentioned above are all homogeneous, where only one type of nodes, e.g., transmitter-receiver pairs, exist and they share the same contents as node features. Therefore, these GNN models are not applicable to the heterogeneous networks that comprise various types of base stations (BSs), APs and MSs, which differ in their node features such as transmit power constraint, beamforming vector, duplexing mode, and so on.

To adapt GNN to heterogeneous graphs, heterogeneous Graph Neural Network (HGNN) was first introduced in [8] for graph mining tasks (e.g., link prediction and recommendation), and recently applied to tackle the optimization problems in wireless communications field. Different from [5], the authors of [9] proposed an HGNN-based model to handle a heterogeneous D2D system, where the transmitter-receiver pairs have different numbers of antennas at transmitter. An HGNN model was also proposed in [10] for the intelligent reflecting surface (IRS)-assisted multi-user multiple-input single-output (MU-MISO) systems. In this case, a single IRS and multiple DL MSs form a simple heterogeneous graph, which is used to learn precoders and the reflective coefficients of the IRS.

However, the authors only considered the scalability with respect to MSs, while neglecting the cases of multiple reflecting elements. In [11], an HGNN model, named bipartite graph neural network (BGNN), was introduced to design precoders for MU-MISO systems, where MSs and BS antennas are considered as two types of nodes. Although the authors demonstrated that BGNN can be adapted to CF networks, it cannot cope with the interference among BSs in multi-cell heterogeneous networks or among APs in CF-mMIMO when each BS/AP only serves some MSs. Furthermore, in [12], an HGNN-based model was used for PA in multi-cell heterogeneous networks where BSs and MSs are categorized into two types of nodes showing that HGNN can outperform conventional GNN.

4.1.2 Motivation

The existing HGNN-related work has made great breakthroughs, but none of them can handle CF-mMIMO with underlaid D2D, mainly due to the following obvious reasons: (i) In CF-mMIMO, each AP with multiple antennas can simultaneously transmit signals to multiple MSs. Moreover, each MS is usually served by a cluster of APs. Therefore, the node features of APs and MSs may change with the sizes of CF-mMIMO networks. (ii) The CF-mMIMO with underlaid D2D systems have three types of nodes, i.e., APs, MSs and D2D pairs. Therefore, each AP node can connect with the nearby MS nodes via the DL communication paths, and with the nearby D2D nodes via interference paths. Therefore, a CF-mMIMO with underlaid D2D network is depicted as a very complex heterogeneous graph, which cannot be directly solved by the off-the-shelf HGNN models designed for the cell-based networks. To address these challenges, in this work, we propose a novel CF-HGNN model specifically designed for solving the beamforming design and power allocation problem in the CF-mMIMO with underlaid D2D networks.

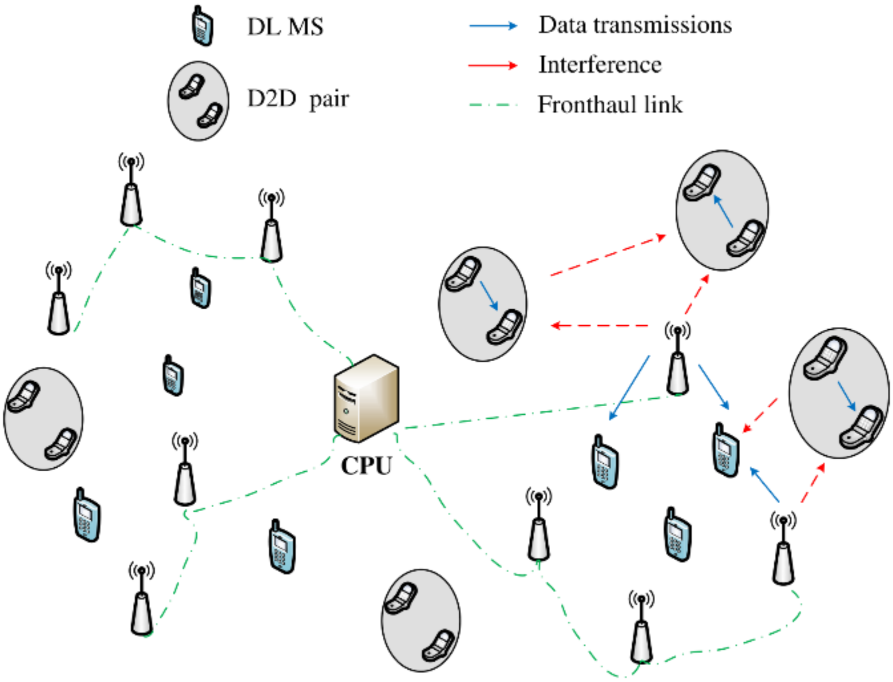


Figure 18: CF-mMIMO with underlaid D2D network

4.2 System Model

We study a scenario of CF-mMIMO with underlaid D2D communications, as depicted in Figure 18. Specifically, the CPU is assumed to take charge of channel estimation and beamforming implementation, while L APs of each with N_{A} antennas transmit the coded DL signal to U single-antenna MSs. The system also contains D D2D transceiver pairs, where the transmitter equipped with N_{D} antennas communicate with single-antenna receiver. We assume AP-MS and D2D transmissions occur over the same time-frequency resource block, and hence there exists interference between these two types of communications.

Let $\boldsymbol{h}_{l,u} \in \mathbb{C}^{N_{\mathrm{A}}}$ denote the channel response between the l-th AP and the u-th MS, which is modelled as $\boldsymbol{h}_{l,u} \sim \mathcal{CN}(0, \beta_{l,u}\boldsymbol{I}_{N_{\mathrm{A}}})$, and $\boldsymbol{h}_d \in \mathbb{C}^{N_{\mathrm{D}}}$ denotes the channel response between the transmitter and receiver at the d-th D2D pair, following the distribution $\boldsymbol{h}_d \sim \mathcal{CN}(0, \beta_d \boldsymbol{I}_{N_{\mathrm{D}}})$. We further denote the interference channels from the l-th AP to the receiver of the d-th D2D pair, from the transmitter of the d-th D2D pair to the u-th MS and from the transmitter of the d-th D2D pair to the receiver of the d'-th D2D pair, by $\boldsymbol{h}_{l,d} \sim \mathcal{CN}(0, \beta_{l,d}\boldsymbol{I}_{N_{\mathrm{A}}})$, $\boldsymbol{h}_{d,u} \sim \mathcal{CN}(0, \beta_{d,u}\boldsymbol{I}_{N_{\mathrm{D}}})$ and $\boldsymbol{h}_{d,d'} \sim \mathcal{CN}(0, \beta_{d,d'}\boldsymbol{I}_{N_{\mathrm{D}}})$, respectively. The $\beta \in \{\beta_{l,u}, \beta_d, \beta_{l,d}, \beta_{u,d}, \beta_{d,d'}\}$ is the large-scale fading factor accounting for path loss and shadowing.

4.2.1 Transmission Model

The data transmitted at the l-th AP can be written as

$$\boldsymbol{s}_l = \sum_{i \in \mathcal{U}_l} \boldsymbol{f}_{l,i} x_i, \tag{1}$$

where $\mathcal{U}_l$ is a set of MSs served by AP l determined by user clustering, x_i is the data for the i-th MS satisfying $\mathbb{E}\{|x_i|^2\} = 1$, $\boldsymbol{f}_{l,i} \in \mathbb{C}^{N_{\mathrm{A}}}$ is the precoding vector under the constraint of $\sum_{i \in \mathcal{U}_l} \|\boldsymbol{f}_{l,i}\|^2 \leq P_l$, where P_l is the transmit power budget at AP l. The data transmitted from the transmitter to the receiver within D2D pair d is given as $\boldsymbol{s}_d = \boldsymbol{w}_d x_d$, where $\mathbb{E}\{|x_d|^2\} = 1$, $\boldsymbol{w}_d \in \mathbb{C}^{N_{\mathrm{D}}}$ is the precoding vector satisfying $\|\boldsymbol{w}_d\|^2 \leq P_d$ and P_d is the power budget at D2D transmitter.

Then, the signal received at the u-th MS can be expressed as

$$y_u = \sum_{l=1}^{L} \boldsymbol{h}_{l,u}^H \boldsymbol{s}_l + \sum_{d=1}^{D} \boldsymbol{h}_{d,u}^H \boldsymbol{s}_d + n_u, \tag{2}$$

where $\mathcal{L}_u$ is a set of APs serving MS u, $n_u \sim \mathcal{CN}(0, \sigma_u^2)$ denotes the additive white Gaussian noise. Furthermore, the signal received at the d-th D2D receiver is

$$y_d = \boldsymbol{h}_d^H \boldsymbol{s}_d + \sum_{d'=1, d' \neq d}^{D} \boldsymbol{h}_{d',d}^H \boldsymbol{s}_{d'} + \sum_{l=1}^{L} \boldsymbol{h}_{l,d}^H \boldsymbol{s}_l + n_d, \tag{3}$$

where $n_d \sim \mathcal{CN}(0, \sigma_d^2)$ is the additive white Gaussian noise.

Accordingly, the SINR obtained by the u-th MS can be expressed as

$$\mathrm{SINR}_u = \frac{\left|\tilde{\boldsymbol{h}}_u^H \tilde{\boldsymbol{f}}_u\right|^2}{\sum_{l=1}^{L} \sum_{i=1, i \neq u}^{U} |\boldsymbol{h}_{l,u}^H \boldsymbol{f}_{l,i}|^2 + \sum_{d=1}^{D} |\boldsymbol{h}_{d,u}^H \boldsymbol{w}_d|^2 + \sigma_u^2}, \tag{4}$$

where $\tilde{\boldsymbol{h}}_u = \left[\boldsymbol{h}^H_{l_1,u}, \ldots, \boldsymbol{h}^H_{l_{|\mathcal{L}_u|},u}\right]^H \in \mathbb{C}^{|\mathcal{L}_u|N_A}$, $\tilde{\boldsymbol{f}}_u = \left[\boldsymbol{f}^H_{l_1,u}, \ldots, \boldsymbol{f}^H_{l_{|\mathcal{L}_u|},u}\right]^H \in \mathbb{C}^{|\mathcal{L}_u|N_A}$. Furthermore, the SINR obtained at the receiver of the d-th D2D pair is given as

$$\mathrm{SINR}_d = \frac{|\boldsymbol{h}_d^H \boldsymbol{w}_d|^2}{\sum_{d'=1,d'\neq d}^{D} |\boldsymbol{h}_{d',d}^H \boldsymbol{w}_{d'}|^2 + \sum_{l=1}^{L} \sum_{i\in\mathcal{U}_l} |\boldsymbol{h}_{l,d}^H \boldsymbol{f}_{l,i}|^2 + \sigma_d^2}. \tag{5}$$

4.2.2 Problem Formulation

In light of the maximization of total spectral efficiency and minimum QoS requirements of both AP-MS and D2D communications, the overall optimization problem can be formulated as

$$\begin{aligned}
&\max_{\{\boldsymbol{f}_{l,u}\},\{\boldsymbol{w}_d\},\mathcal{L}_u,\mathcal{U}_l} \Lambda_{\mathrm{SE}} = \sum_{u=1}^{U} \alpha_u \log(1+\mathrm{SINR}_u) + \sum_{d=1}^{D} \alpha_d \log(1+\mathrm{SINR}_d) \\
&\text{s.t. } \log(1+\mathrm{SINR}_u) \geq \chi_a, \ \forall u \\
&\log(1+\mathrm{SINR}_d) \geq \chi_b, \ \forall d \\
&|\mathcal{U}_l| \leq \lambda, \ \forall l \\
&\sum_{u\in\mathcal{U}_l} \|\boldsymbol{f}_{l,u}\|^2 \leq P_l, \ \forall l \\
&\|\boldsymbol{w}_d\|^2 \leq P_d, \ \forall d,
\end{aligned} \tag{6}$$

where χ_a and χ_b guarantee the normal operation of AP-MS and D2D communications, respectively. λ limits the size of user clusters for the consideration of system complexity and spectral efficiency. As the optimization problem (6) is nonconvex and NP-hard to be dealt with, we resort to graph neural network, which will be demonstrated to exhibit low complexity and high scalability for solving resource allocation in CF-like wireless systems.

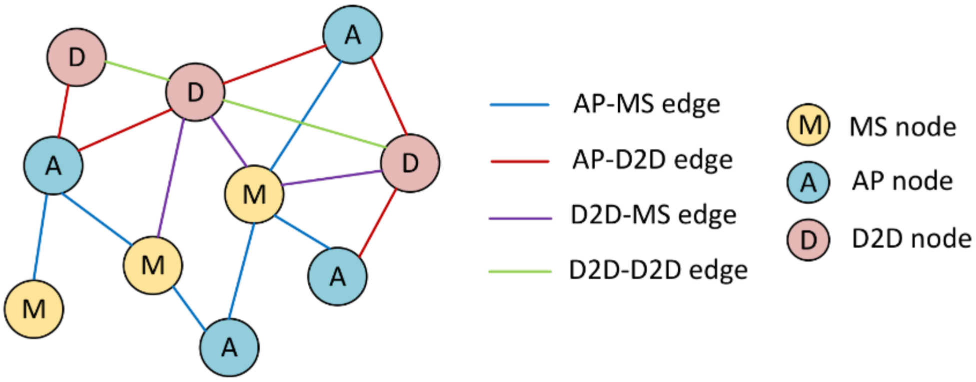


Figure 1 An example of heterogeneous graph mapped by a CF-mMIMO with underlaid D2D network.

4.3 From the Proposed System to a Heterogeneous Graph

As shown in Figure 19, it is straightforward to depict the proposed system as an heterogeneous graph, where AP, MS and D2D pair are deemed as three types of nodes and there also exist four types of edges, i.e., AP-MS, AP-D2D, MS-D2D and D2D edge. To be more specific, we give an example of heterogeneous graph mapped by a CF-mMIMO with underlaid D2D network in Figure 20. To mathematically describe the graph in this figure, we represent it as $G=(\mathcal{V},\mathcal{E})$, where $\mathcal{V}$ denotes the set of nodes, and $\mathcal{E}$ denotes the set of edges. This heterogeneous graph has a node type mapping function $\phi\colon \mathcal{V}\longrightarrow \mathcal{Q}$ and an edge type mapping function $\psi\colon \mathcal{E}\longrightarrow \mathcal{P}$, where $\mathcal{Q}$ and $\mathcal{P}$ are the sets of predefined node types and edge types. Specifically, we have $\mathcal{Q}=\{Q_1,Q_2,Q_3\}$ and $\mathcal{P}=\{P_1,P_2,P_3,P_4\}$, where Q_i and P_j denote the i-th node type and j-th edge type, respectively. Then, a node with F_v-dimensional features is denoted as $\boldsymbol{v}_i\in\mathbb{R}^{F_v\times 1}$, while an edge between $\boldsymbol{v}_j$ and $\boldsymbol{v}_i$ is denoted as $\boldsymbol{e}_{i,j}=(\boldsymbol{v}_i,\boldsymbol{v}_j)\in\mathbb{R}^{F_e\times 1}$, which has F_e-dimensional features.

Given the mapping function ϕ and ψ, each node belongs to a particular node type of $\phi(\boldsymbol{v})\in\mathcal{Q}$ and each edge belongs to a specific relation of $\psi(\boldsymbol{e})\in\mathcal{P}$. The neighbourhood of a node $\boldsymbol{v}_i$ is defined as $\mathcal{N}_i=\{\boldsymbol{v}_j\in\mathcal{V}|\boldsymbol{e}_{i,j}\in\mathcal{E}\}$.

4.3.1 The Construction of Heterogeneous Graph for The Proposed System

4.3.1.1 Sparse Graph Construction

Owing to the constraint of (6d), the heterogeneous graph of our proposed system is a sparse graph. That is to say with the aid of user clustering, each AP only connects with a group of MSs rather than all of them. The specific user clustering approach used in this work is given as follows:

(i) First, we initialize the sets $\mathcal{L}_u=\emptyset,\forall u$ and $\mathcal{U}_l=\emptyset,\forall l$. Then, we sequentially allocate a master AP l to each MS u following the principle of

$$l=\mathrm{argmax}_l\beta_{l,u},\quad \mathcal{L}_u=\{l\},\quad \mathcal{U}_l=\{d\} \tag{7}$$

(ii) Next, considering the fairness among all the MSs, we find the MS u having the smallest large-scale channel gain, i.e., $\sum_{l\in\mathcal{L}_u}\beta_{l,u}<\sum_{l\in\mathcal{L}_{u'}}\beta_{l,u'},\forall u',u\neq u'$. Then, we allow it to first select the serving AP. This process is continuously implemented until $|\mathcal{U}_l|=\lambda,\ \forall l$.

After user clustering, the connections between APs and MSs are determined. Then, to figure out the connections of interference path, we introduce a threshold distance d_{thr}. If the distance between two interfering nodes is larger than d_{thr}, these two nodes have no edge with each other. Otherwise, they are connected together. Accordingly, based on the above-mentioned heuristic methods, given a CF-mMIMO with underlaid D2D network, we can transform it into a sparse heterogeneous graph.

4.3.1.2 Node and Edge Features

The efficient graph training requires the node and edge features to be appropriately configured. To this end, the node feature of AP is given by

$$\boldsymbol{v}_l=[\boldsymbol{h}_{l,u_1}^H,\ldots,\boldsymbol{h}_{l,u_\lambda}^H,P_l]^H\in\mathbb{C}^{N_{\mathrm{A}}\lambda+1},\ u\in\mathcal{U}_l \tag{8}$$

and the node feature of D2D pair is given by

$$\boldsymbol{v}_d=[\boldsymbol{h}_d^H,P_d]^H\in\mathbb{C}^{N_{\mathrm{D}}+1},\ \forall d. \tag{9}$$

Note that MSs have no feature and their node features can be initialized to zero vectors, i.e., $\boldsymbol{v}_u=\boldsymbol{0},\forall u$. Moreover, the feature of edge is defined as $e_{i,j}=\beta_{i,j}$.

4.3.1.3 Data Normalization

As shown in Section 4.3.1.2, the data involved in the node and edge features, i.e., CSI, power budget and Euclidean distance, have non-identical distributions with different scales. In this regard, the data normalization is necessary in this case. Specifically, we use the Z-score normalization method to separately pre-process the CSI in node feature and Euclidean distance in edge feature. Additionally, the data of power budget are pre-processed by max-min normalization.

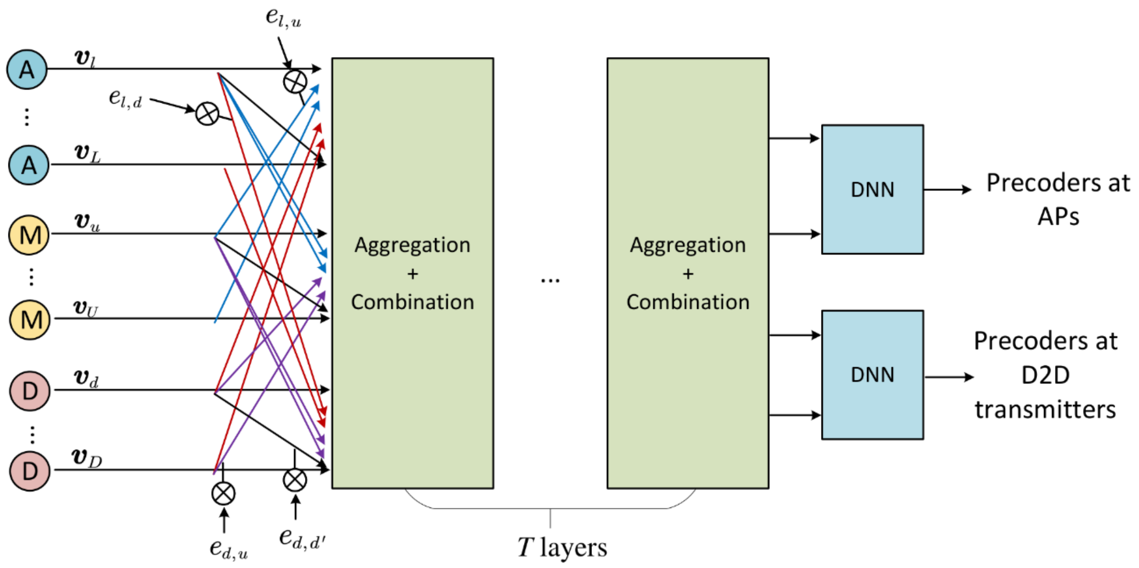


Figure 20: The architecture of CFD-HetGNN

4.3.2 HetGNN-Relied Approach for Beamforming Design

Based on the construction of heterogeneous graph, we propose a HetGNN-relied on approach for beamforming design, namely CFD-HetGNN. Its overall architecture is shown in Figure 20, which consists of two parts, i.e., aggregation plus combination layers and DNN modules.

4.3.2.1 Aggregation Plus Combination Layers

Once the data have been well-prepared by data normalization, the message passing process relying on aggregation plus combination layers commences. During the process, each node collects the information from the neighbouring nodes, and combine them with its own feature. Then, the updated feature is taken as the input for the next round of aggregation and combination. The details are presented as follows:

(i) **Aggregation:** The node feature of all the nodes connected to node i and the feature of interconnection edges are first aggregated and then averaged to obtain the neighbouring information of node i, which can be expressed as

$$\boldsymbol{v}_{i,j}^{(t-1)} = \frac{1}{|\mathcal{N}_i|}\Xi_{\text{agg}}\Big(\sum_{j\in\mathcal{N}_i}\boldsymbol{v}_j^{(t-1)}\cdot e_{i,j}\Big) \tag{10}$$

where Ξ_{agg} is an aggregator composed of the fully-connected and activation layers, $\boldsymbol{v}_j^{(t-1)}$ denotes the node feature at the $(t-1)$ layer.

(ii) **Combination:** Then, the node combines its own information with the aggregated neighbouring information. Additionally, in order to maintain the original information of each node during the process of multi-layer combination, we add a simple residual block, which integrates the post-combined and the original node features together by element-wise Mean$(\cdot)$ operator. The above-mentioned process at node i can be written as

$$\boldsymbol{v}_i^{(t)} = \mathrm{Mean}\left(\Xi_{\mathrm{com}}\left(\boldsymbol{v}_i^{(t-1)} \oplus \boldsymbol{v}_{i,j}^{(t-1)}\right) + \boldsymbol{v}_{\mathrm{Res}}\right) \tag{11}$$

where Ξ_{com} is a combiner having the same structure with Ξ_{agg}.

We should mention that the parameters of aggregator and combiner are not shared among different types of nodes.

4.3.2.2 DNN Modules

After T-layer message passing, the final representation of nodes are fed into DNN modules, which are leveraged to calculate the beamforming vectors for both APs and D2D pairs. The computation process can be expressed as

$$\begin{aligned} \tilde{\boldsymbol{f}}_l &= \Xi_{\mathrm{AP}}\left(\boldsymbol{v}_l^{(T)}\right),\ \forall l \\ \boldsymbol{w}_d &= \Xi_{\mathrm{D2D}}\left(\boldsymbol{v}_d^{(T)}\right),\ \forall d, \end{aligned} \tag{12}$$

where $\tilde{\boldsymbol{f}}_l = [\boldsymbol{f}_{l,u_1}^H, \ldots, \boldsymbol{f}_{l,u_\lambda}^H]^H \in \mathbb{C}^{N_{\mathrm{A}}\lambda}, u \in \mathcal{U}_l$.

Then, to perform the training of CFD-HetGNN in an unsupervised way without any ground truth, the loss function is defined as follows:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) = &-\Lambda_{\mathrm{SE}} + \kappa_a \mathrm{Relu}\left[\sum_{u}^{U}\left(\chi_a - \log\left(1 + \mathrm{SINR}_u\right)\right)\right] + \kappa_b \mathrm{Relu}\left[\sum_{d}^{D}\left(\chi_b - \log\left(1 + \mathrm{SINR}_d\right)\right)\right] \\ &+ \kappa_c \mathrm{Relu}\left[\sum_{l}^{L}\sum_{u \in \mathcal{U}_l}\|\boldsymbol{f}_{l,u}\|^2 - P_l\right] + \kappa_d \mathrm{Relu}\left[\sum_{d}^{D}\|\boldsymbol{w}_d\|^2 - P_d\right] \end{aligned} \tag{13}$$

where $\boldsymbol{\theta}$ denotes all the trainable parameters of CFD-HetGNN model, and $\kappa_i, i \in \{a,b,c,d\}$ are weighted factors to enforce the training process towards meeting the demand of constraints in optimization problem (6).

4.4 Simulation Results

We consider a CF-mMIMO network within an area of $400 \times 400 m^2$, where the number of APs, MSs and D2D pairs are set to 24, 6, and 6 respectively. The number of antennas equipped at APs and D2D transmitter are assumed to be 6 and 4, respectively. Furthermore, the power budget of APs is 40 dBm, while that of D2D pairs is 30 dBm. To guarantee the basic QoSs for AP-MS and D2D communications, both χ_a and χ_b are set to 0.1 bits/s/Hz. For the consideration of system complexity as well as the overhead of fronthaul links, the maximum numbers of MSs that each AP can concurrently serve is set to 2, (i.e., $\lambda = 2$). The large-scale fading model adopted during simulation is given by

$$\beta[\mathrm{dB}] = -30.5 - 36.7\log_{10}(d) + \sigma_{\mathrm{sh}} z, \tag{14}$$

where d denotes the distance between any two nodes, $\sigma_{\mathrm{sh}} z$ is the shadowing fading with a standard deviation of $\sigma_{\mathrm{sh}} = 4$ dB and $z \sim \mathcal{N}(0,1)$.

As for the configuration of neural network, a two-layer CFD-HetGNN is applied on the basis of Pytorch Geometric [13]. To optimize the CFD-HetGNN, we leverage the Adam optimizer with a learning rate of 0.001. The structures of the involved MLPs for Aggregation&Combination and DNN modules are listed in Table I. Note that the MLPs for Aggregation&Combination include the LeakyRelu activation layer, batch normalization layer and Dropout layer, while that for D2D modules only include LeakyRelu activation layer. Furthermore, we empirically set κ_i in (14) as {1, 0.1, 0.1, 0.1} during training. Moreover, we randomly generate 10000 and 1000 CF network layout samples for training and testing, respectively, under the assumption that APs, MSs and D2D pairs are uniformly distributed within the area. The batch size for training is 64, and the network parameters are only updated during training, which stay constant during testing. The CFD-HetGNN is run on a GeForce GTX laptop 3080Ti.

Table 1 Architecture of involved MLPs

MLPs	Layers
Ξ_{agg}	$24 \times 256 \times 512$
Ξ_{com}	$536 \times 1024 \times 24$
Ξ_{AP}	$24 \times 256 \times 512 \times 24$
Ξ_{D2D}	$24 \times 256 \times 8$

In order to verify the performance of CFD-HetGNN, we introduce a benchmark beamforming strategy, namely matched filtering with equal power allocation (MF-EPA), where both AP and D2D transmitter employ MF precoding method and the power allocation at each AP is equally distributed to the MSs served. Note that the MF-EPA does not consider the QoSs of AP-MS and D2D communications.

The training convergence of our proposed CFD-HetGNN is shown in Figure 21. The model is trained with 200 epochs and unsupervised training is implemented without any ground truth. It can be observed that 100 epochs are sufficient to achieve a satisfactory performance.

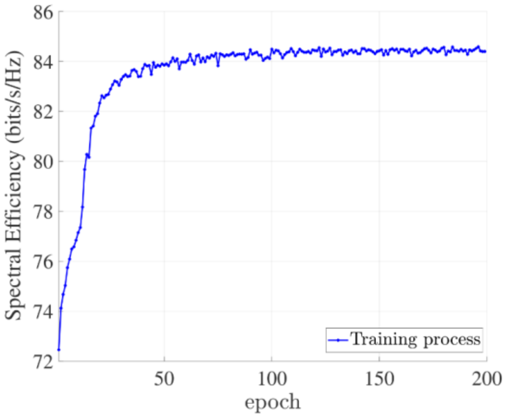
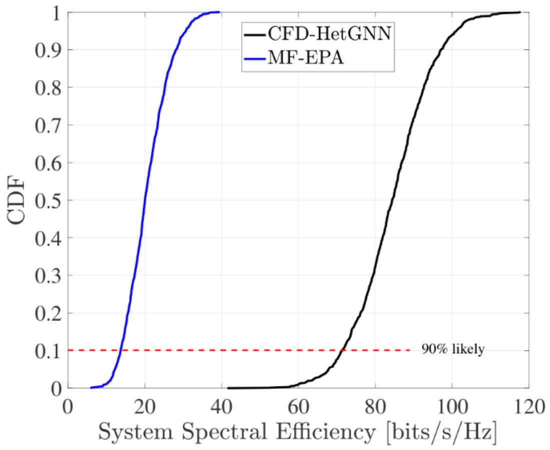


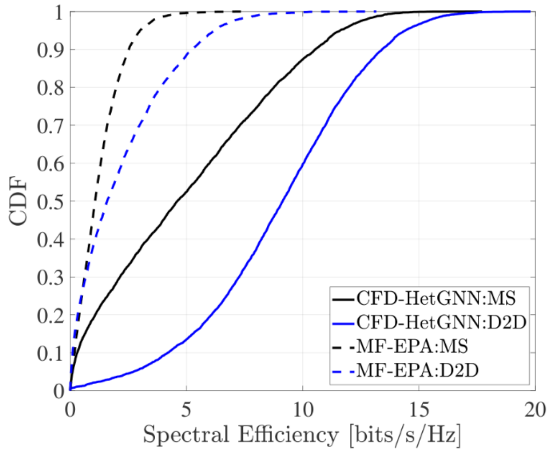
Figure 21: Testing spectral efficiency vs. training epochs

We further compare the performance of CFD-HetGNN and MF-EPA in terms of spectral efficiency. As shown in Figure 22(a), the system SE performance gaps between the CFD-HetGNN and MF-EPA with regard to 1000 testing network layouts are depicted. It can be observed that CFD-HetGNN can significantly outperform MF-EPA. In particular, CFD-HetGNN achieves six times higher 90% likely system SE than MF-EPA method. The reason behind is that with the aid of graph representation, CFD-HetGNN is able to fully exploit the structure of CF-mMIMO with underlaid D2D network, and hence the efficient interference management is relatively easy to be realized. By contrast, due to the limit number of antennas at both APs and D2D transmitter, MF-based method fails to manage the complicated interference among APs, MSs and D2D pairs. Therefore, although there is no QoS requirements during the implementation of MF-EPA, its performance lags far behind CFD-HetGNN.

We also look deeper into the performance achieved by MSs and D2D pairs under two methods, as shown in Figure 5(b). We can see that in general D2D pairs have better performance than MSs, since the distance between transmitter and receiver within a D2D pair is relatively close compared with that between APs and MSs. Another possible reason is that, as APs work in a cooperative way, it can efficiently mitigate the interference signal arriving at each D2D receiver. On contrary, D2D transmitter with few antennas can hardly suppress the D2D-MS interference. In addition, from Figure 22(b), CFD-HetGNN achieves better performance in terms of both MSs and D2D pairs SE.



(a) Cumulative distribution of system SE



(b) Cumulative distribution of MS/D2D SE

Figure 22: Performance comparison between CFD-HetGNN and MF-EPA

4.5 Conclusions

In this work, in order to efficiently design beamforming along with power allocation for CF-mMIMO with underlaid D2D systems, we proposed an advanced strategy relying on heterogeneous graph neural network. The CFD-HetGNN consists of aggregation plus combination layers and DNN modules, and it was shown to significantly outperform traditional MF method with equal power allocation. In near future, we plan to equip our proposed wireless network with the capability of sensing, that is the integrated sensing and communication for CF-mMIMO with underlaid D2D. To address the complicated resource allocation and beamforming design in such a scenario, we will further improve the CFD-HetGNN model.

4.6 References

[1] Ö. T. Demir, E. Björnson, L. Sanguinetti et al., "Foundations of user-centric cell-free massive MIMO," Foundations and Trends® in Signal Proc., vol. 14, no. 3-4, pp. 162–472, 2021.

[2] L. Luo, et.al, "Downlink power control for cell-free massive MIMO with deep reinforcement learning," IEEE Trans. on Veh. Tech., 2022.

[3] M. Bashar, et.al, "Deep learning-aided finite-capacity fronthaul cell-free massive MIMO with zero forcing," in ICC'20. IEEE, 2020, pp. 1–6.

[4] M. Zaher, Ö. T. Demir, E. Björnson, and M. Petrova, "Learning-based downlink power allocation in cell-free massive MIMO systems," arXiv preprint arXiv:2109.03128, 2021.

[5] Y. Shen, et.al, "Graph neural networks for scalable radio resource management: Architecture design and theoretical analysis," IEEE J. on Selected Areas in Com., vol. 39, no. 1, pp.101–115, 2020.

[6] A. Chowdhury, G. Verma, C. Rao, A. Swami, and S. Segarra, "Unfolding WMMSE using graph neural networks for efficient power allocation," IEEE Trans. on Wireless Com., vol. 20, no. 9, pp. 6004–6017, 2021.

[7] M. Eisen and A. Ribeiro, "Optimal wireless resource allocation with random edge graph neural networks," IEEE Trans. on Signal Proc., vol. 68, pp. 2977–2991, 2020.

[8] C. Zhang, et.al, "Heterogeneous graph neural network," in Proc. of the 25th ACM SIGKDD, 2019, pp. 793–803.

[9] X. Zhang, et.al, "Scalable power control/beamforming in heterogeneous wireless networks with graph neural networks," in 2021 IEEE GLOBECOM'21. IEEE, 2021, pp. 01–06.

[10] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," IEEE J. on Selected Areas in Comm., vol. 39, no. 7, pp. 1931–1945, 2021.

[11] J. Kim, H. Lee, S.-E. Hong, and S.-H. Park, "A bipartite graph neural network approach for scalable beamforming optimization," IEEE Trans. on Wireless Com., 2022.

[12] J. Guo and C. Yang, "Learning power allocation for multi-cell-multi-user systems with heterogeneous graph neural networks," IEEE Trans. on Wireless Com., vol. 21, no. 2, pp. 884–897, 2021.

[13] M. Fey and J. E. Lenssen, "Fast graph representation learning with pytorch geometric," arXiv:1903.02428, 2019.

5 Intelligent IAB with Beam Steering based on User Location

5.1 Aims and Objectives

High mobility is considered as one of the important aspects in the Fifth Generation (5G) and beyond systems to support number of applications such as high-speed rail (HSR). It is anticipated that supporting speeds greater than 1000 Km/h and high data rates are essential to enhance the user experience in the future wireless systems and meet the future requirements [1]. However, it is widely believed that the current systems cannot meet these requirements due to Doppler effect and beam misalignment. For instance, Orthogonal frequency division multiplexing (OFDM) suffers in high mobility due to Doppler effect [2]. Furthermore, the beamforming becomes more challenging in the future systems due to the large number of antenna elements and high mobility scenarios [3]. Minimizing the impact of Doppler in addition to proposing an efficient beamforming scheme tackling the current challenges such as complexity and training overhead can boost the system performance and meet the future requirements. Regarding the Doppler issue, some solutions have been proposed to minimize its impact and boost the system performance. For instance [4] tackled the issues caused by the high mobility in Millimetre Wave (mmWave) systems in order to meet the future requirements. This work proposed new changes to the current system to minimize the impact of Doppler effect. The proposed changes include shortening the subframe to minimize the interval between reception of channel quality feedback and the corresponding transmissions and performing feedback more frequently. However, these changes can cause a few issues such as system complexity and signal control overhead. Recently Orthogonal Time Frequency Space Modulation (OTFS) has been proposed as an alternative to OFDM due to its stability for high mobility communications [5]. OFDM is a well-established modulation technique used in many wireless communication standards such as Wi-Fi (IEEE 802.11), Fourth Generation (4G), and 5G mobile. On the other hand, OTFS is a relatively newer modulation scheme that aims to overcome some of the limitations of OFDM. OTFS achieves a higher resilience to Doppler and delay spread compared to OFDM. It offers improved performance in scenarios with fast-moving users, such as HSR or vehicles on motorways. However, one drawback of OTFS is its higher computational complexity compared to OFDM. The processing required to transform the signal into the delay-Doppler domain and perform the inverse transformation can be computationally intensive. This complexity can limit the practical implementation of OTFS in some systems [6]. We believe that using both OTFS and OFDM in the future system can provide flexibility, support older devices and sensors, enhance affordability (devices equipped with OFDM system might stay cheaper for some time due to its maturity and popularity) and avoid the unnecessary complexity. In this work, we consider both systems by considering the user speed and system complexity as it will be shown later in this work. On the other hand, the beamforming has also attracted the interest of many researchers due to its importance [3], [7], [9]–[13]. The traditional algorithms are designed to estimate the direction of arrival (DOA) of the signal and optimize the weights or phases of the antenna elements accordingly. However, these solutions are considered challenging in the dynamic environments as the weights need to be computed more frequently. This issue becomes more serious and computationally expensive when the number of antenna elements increases. It is widely believed that the traditional solutions suffer from a significant trade-off between the complexity and system performance. Furthermore, Deep Learning (DL) solution is considered as one of the potential solutions in beamforming design to boost the system performance [8], [9]. Some studies have addressed the beamforming and proposed DL algorithms to ease some of the challenges in the conventional algorithms such as the amount of training overhead [10]–[12]. In [12], the joint design of beamforming, power control and interference coordination is modelled as a non-convex optimization problem to maximize the signal to interference plus noise ratio (SINR) which is solved by using Reinforcement Learning (RL) and Q-learning. One of the main drawbacks of this work is that the user locations and optimized beamforming vectors are assumed to be known. However, this information is not always available in real scenarios. Furthermore, DL approach has been used to propose efficient beamforming schemes and improve the system performance [10], [11]. Although, the DL is considered as one of the important approaches for fast and effective beamforming [9], [13], the existing machine learning network has some challenges such as the significant amount of data required for the training and its complexity (these networks are designed to learn and extract meaningful patterns and representations which makes it difficult to interpret and explain the conclusions made) [14]. The contribution of this research is summarized as follows:

- A novel DLIRL-based beam steering scheme is implemented, which is capable to steer the beam with the user movement.
- The proposed scheme combines the performance of the DL and RL. The DL is used for preparing the optimized beamform codebook and RL is used for selecting the best beam out of the optimized beamform codebook based on the user movements.
- The training of the DLIRL is carried out based on the channel information without requiring the user's location. Thus, reducing the training overheads.
- A novel way of combining the DL with the RL and employing it for beam steering applications.
- The proposed scheme is trained in the offline mode for a particular environment simulated using the MATLAB site viewer. The trained model is the employed successfully for the beam steering. Furthermore, the proposed scheme is sufficiently flexible for getting trained to any provided environment.

5.2 Model definition

The proposed work is designed for mmWave-based wireless communication networks. The considered system model comprises of I number of IAB nodes termed as IABs serving a UE fitted with the single isotropic antenna as shown in Figure 23. The IAB node is connected to the B5G core via an IAB donor. The B5G core carries the brain for the intelligent and adaptive beamforming in the form of a DLIRL network. Each IABs are equipped with NT antennas communicating information symbol $s^K \in \mathbb{C}$ for the K^{th} subcarrier here K = 1…., k. Each IAB has baseband precoder vector for each K^{th} subcarrier $f_{BB}^{K} \in \mathbb{C}^{I\times 1}$, and for RF precoder $f_{RF} \in \mathbb{C}^{N^T\times 1}$. The RF precoder for the l^{th} IAB at m^{th} antenna element can be modelled as a phase shifter network which is mathematically represented as $[f_{RF}^{l,m}] = \frac{1}{\sqrt{N^T} e^{j\phi^{l,m}}}$, here $\phi^{l,m}$ is quantized angle [15].

The downlink transmission for the transmitted data symbol can be represented as $y = f_{RF}^{l} f_{BB}^{K} s^K$. Here $\mathbb{E}[s^K (s^K)^H] = \frac{P^K}{k}$, P^K is the power associated with the K^{th} subcarrier and k is the total number of subcarriers. The constraint in the total transmission power of IAB should be $\|F_{RF} f_{BB}^{K}\|_F^2 = 1 \;\; K = 1, 2, \cdots, k$.

Here, $F_{RF} = blkdiag(f_{RF}^{1}, \cdots, f_{RF}^{I}) \in \mathbb{C}^{IN^T\times I}$. The channel vector between the l^{th} IAB and the UE for the K^{th} subcarrier is denoted as $h_{l,K} \in \mathbb{C}^{N^T\times 1}$. Then the received data stream $x_{l,K}$ by the UE from the l^{th} IAB during the downlink at K^{th} subcarrier can be written as in Eq. (1)

$$x_{l,K} = \sqrt{P_{avg}}\, h_{l,K}^{H}\, f_{RF}^{l,K}\, f_{BB}^{l,K}\, s^K + n_{K,l} \tag{1}$$

where P_{avg} is the average received power at the UE. The $n_{K,l} \sim N_c(0, \sigma^2)$ is the K^{th} subcarrier.
noise at the UE.

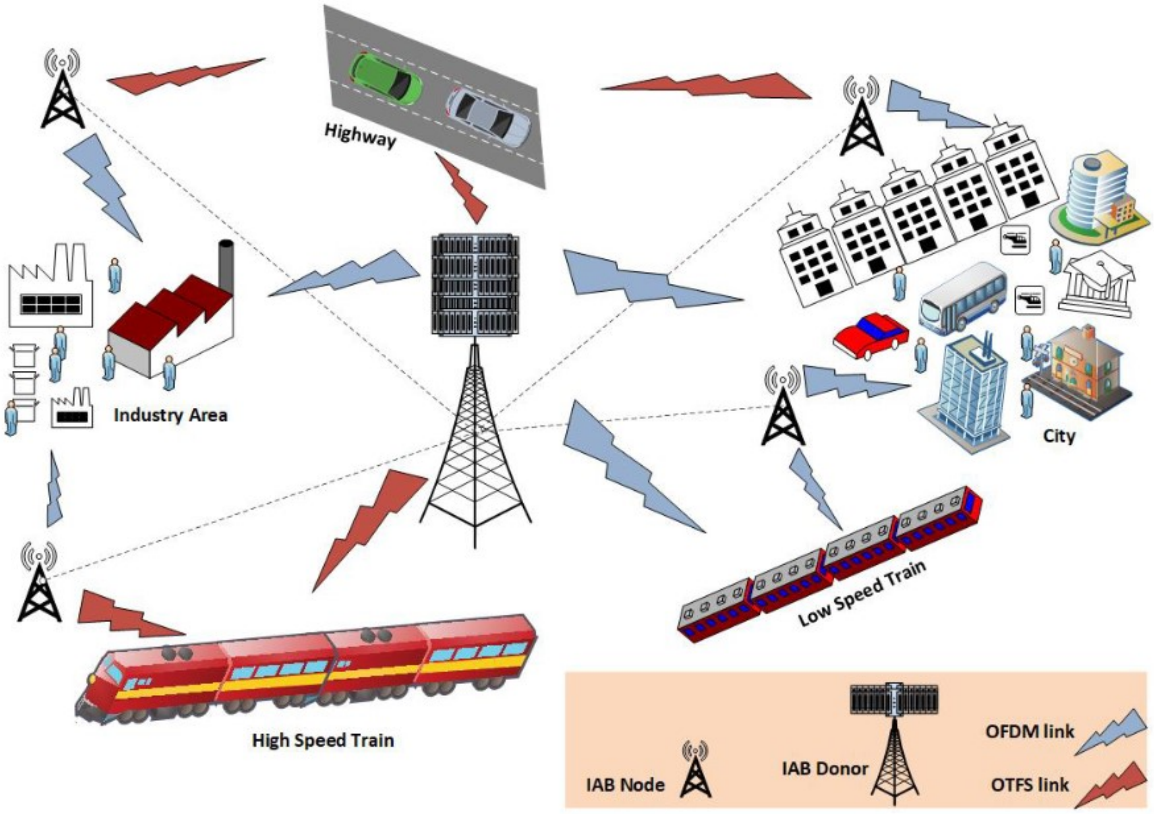


Figure 23: B5G System Model

System Model Symbol Notations	
I	Number of IAB nodes
N_T	Number transmit antenna elements at an IAB
s_K	Information symbol for K^{th} subcarrier
$\mathbf{f}_{\mathbf{BB}}^{K}$	Baseband precoder matrix for each K^{th} subcarrier
f_{RF}	RF precoder vector
P_{avg}	Average received power
$\boldsymbol{x}_{l,K}$	The received data stream by the UE from the l^{th} IAB during the downlink at K^{th}
$\mathbf{h}_{l,K}$	Channel vector between the l^{th} IAB and the UE for the K^{th} subcarrier

5.2.1 Channel Modelling

The channel between IABs and the UE is considered as a wideband mmWave channel with c = 1,…C clusters contributing to one ray of time delay as τ_c, AoA as ϕc and ϑc for azimuth and elevation angles respectively. The channel path loss between UE and the l^{th} IAB over K^{th} is represented as $L_{l,K}$. The antenna array response vector between l^{th} IAB and the UE for the K^{th} subcarrier is represented as $a_{l,K}(\theta_c, \phi_c)$. Mathematically the channel can be modelled for the K^{th} subcarrier as shown in Eq. 2 [14]:

$$h_{l,K} = \sqrt{\frac{N^T}{L_{l,K}}} \sum_{c=1}^{C} \beta_c \, a_{l,K}(\theta_c, \phi_c) \tag{2}$$

Here β_c is the complex gain associated with the resolvable path *c*. The considered channel model is assumed to remain constant over coherence time T_c as it is a block fading channel.

5.2.2 Problem Formulation

The challenges associated with 5G standards are to reach the goals of higher data rate, lower latency, better coverage, and high mobility. To achieve this, it is important to have the most flexible and controlled beamforming scheme. The existing beamforming techniques [1-6] tried to achieve flexibility and control via dedicated transmit/receive for each element. Considering massive MIMO-based wireless communication systems, building this type of architecture is highly difficult due to extensive cost, power, and space-based limitations [12], thus, hindering the design budget. With this motivation, this work is dedicated to utilising an effective AI training scheme based on DL and RL combinedly termed DLIRL for the beamforming strategy. To quantify the goals achieved by the proposed scheme, the performance is evaluated by considering the beamforming system that can maximize effective spectral efficiency for mmWave-based wireless communication applications at UE. For that purpose, the achievable rate at UE for the considered hybrid beamformer, system, and channel model are evaluated as per Eq. (3):

$$S_e = \frac{1}{k} \sum_{K=1}^{k} \log_2 \left[1 + SNR \left| \sum_{l=1}^{I} h_{l,K}^H f_{RF}^l \mathrm{f}_{BB}^{l,K} \right|^2 \right] \tag{3}$$

Here, *I* is the total IABs considered. The SNR represents the signal to noise ratio at UE denoted as $\frac{P_{avg}}{k\sigma^2}$. The objective of this work is to create an effective beamforming design and channel training so as to maximize the achievable rate at UE. The final optimization problem can be deduced as:

$$\begin{aligned} S_e(f_{RF}^l, \mathrm{f}_{BB}^{l,K}) &= \arg\max \frac{1}{k} \sum_{K=1}^{k} \log_2 \left[1 + SNR \left| \sum_{l=1}^{I} h_{l,K}^H f_{RF}^l \mathrm{f}_{BB}^{l,K} \right|^2 \right] \\ \text{s.t.} \| f_{RF}^l, \mathrm{f}_{BB}^{l,K} \|^2 &= 1, \qquad \forall l, \end{aligned} \tag{4}$$

5.3 System Design

The proposed scheme comprises deep learning integrated with reinforcement learning for the beamforming design. The employed techniques make use of the learning strategy by performing mapping between beamforming weights and the environmental setup, channel estimates, AoA, and Time Difference of Arrival (TDoA). The proposed DLIRL employs a pilot signal transmitted from UE to the IAB's to learn about the channel condition and predict the optimal beamform weights. The pilot signals received at IABs are the result of the interaction of the signal with the environment during its propagation. These reflected and diffracted waves are jointly received at different IABs and carries the thumbprint of the environmental factors and channel conditions. The Environmental Thumbprint (ET) carried by these pilot signals is employed for training. The DLIRL has two periods: working and training. The DLIRL initiates with the training period. During this, DLIRL receives pilot sequences transmitted from the predefined UEs positions. The UE transmission is omnidirectional, and it carries ET. The DLIRL then maps the received sequence with the training process and learns it. The working period marks the prediction scheme of the DLIRL and performs prediction of the optimal beamform without any need for additional training. Multiple advantages achieved via the proposed scheme are that it does not require additional resources for learning during the working period. Moreover, the proposed scheme employs an omnidirectional UE pilot sequence for the training period. Therefore, it has minimal overheads during the training period. Also, the proposed scheme can be trained for any environment including LoS and Non-Line of Sight (NLoS).

5.3.1 Training Period

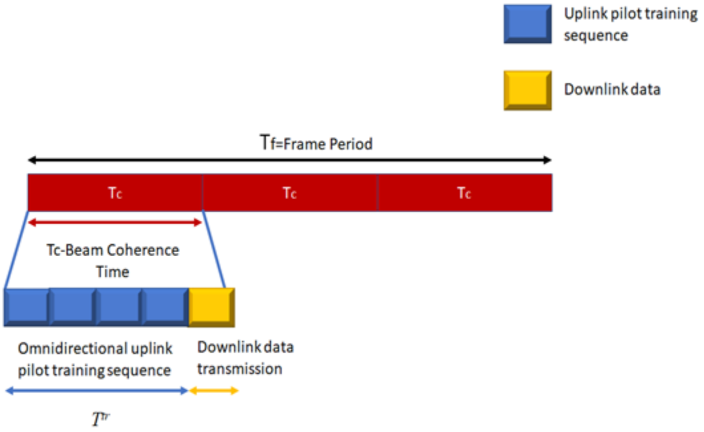


Figure 24: Beam coherence time interval during training process

In this period DNN observes the environment and trains the deep neural network. We consider a beam coherence time T_c, a concept introduced in [15] for mmWave-based wireless communication systems also shown in Figure 24 It can be defined as the period over which the beams are unchanging. Considering T^{tr} as the channel training period of the first T^{tr} time instants of the Tc, then Eq. (4) can be re-written as in Eq. (5).

$$S_e(f^l_{RF}, \mathrm{f}^{l,K}_{BB}) = \left(1 - \frac{T^{tr}}{T_c}\right) \arg\max \frac{1}{k} \sum_{K=1}^{k} \log_2 \left[1 + SNR \left| \sum_{l=1}^{I} h^H_{l,K} f^l_{RF} \mathrm{f}^{l,K}_{BB} \right|^2 \right] \qquad (5)$$

Figure 24 depicts the training period for deep neural network-based learning. In each T_c UE transmits pilot matrix $\mathcal{S}^K_p \in \mathcal{C}^{N^T \times I}$ repeatedly, here K = 1, 2, 3,… , k. The received pilot sequence at l^{th} IAB is as in Eq. (6):

$$\mathrm{x}^p_{l,K} = h_{l,K}\, \mathcal{S}^K_p + n_{K,l} \qquad (6)$$

The combined beamforming strategy starts with feeding the received pilot signals from all IABs to the Fusion Centre (FC) comprising DLIRL. The FC first selects the RF beamform vectors for the IABs downlink as:

$$f^l_{RF} = \underset{f^l_{RF} \in K_{RF}, \forall l}{\arg\max} \sum_{K=1}^{k} \log_2 \left[1 + SNR \left| \sum_{l=1}^{I} h^H_{l,K} \right|^2 \right] \qquad (7)$$

Here, K_{RF} is the RF beamform codebook. The FC then applies baseband precoder calculated as [14]:

$$f^*_{BB} = \frac{\left(h^H_K F_{RF}\right)^H}{h^H_K F_{RF}}, \qquad \forall K \qquad (8)$$

The DNN is fed with the pilot symbols $x_{l,K}^{p}$ as the input for the training model. The spectral efficiency for every RF and baseband beamform vector acts as the output shown in Eq. (5). The DNN is trained during this period and learns the correlation between received pilot symbols at each IABs and the ET. Post-training the DNN moves to the working model where it performs the prediction. The detailed DNN architecture is explained in section 5.3.2.

5.3.2 DNN Architecture

We propose a multilayer DNN structure for training the proposed model and getting adapted to the scenario. The inputs to the model are the Channel Impulse Response (CIR), AoA, and TDoA. The CIR, AoA, and TDoA which are obtained via received pilot symbols at *l* IABs. The DNN structure comprises 4 hidden layers. The first layer has 12 nodes, and the remaining three layers have 10 nodes each. The DNN is trained end to end as a supervised learning class [16]. The DNN input is normalized based on the training dataset. The DNN architecture employed Rectifier Linear Unit (ReLU) as the activation function [17-18].

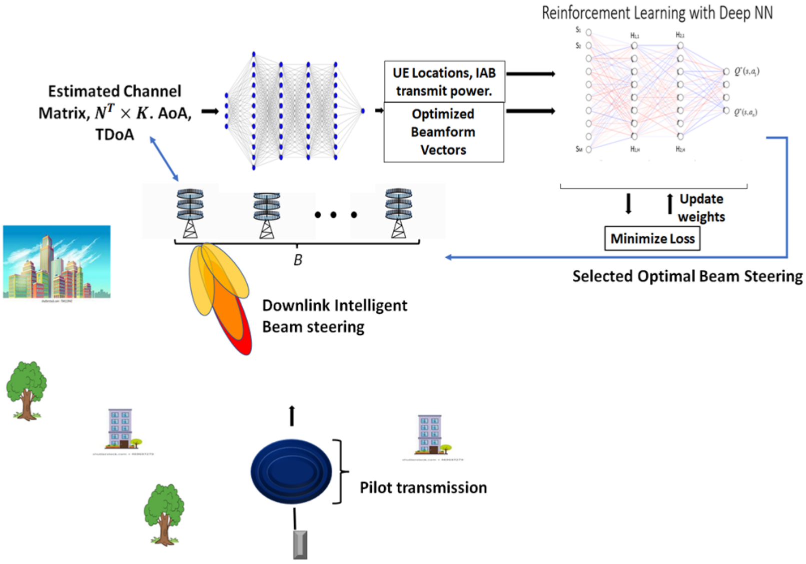


Figure 25. The DLIRL Beamforming System Model

5.3.3 Working Period

Based on the estimated channel, the RF and the Baseband precoding beamform codebook k_{RF} , k_{BB} is respectively formed using Eq. (7) and Eq. (8). The objective of the DNN is to maximize $S_e(f_{RF}^l f_{BB}^l)$ S for each IABs. The regression-based learning model is adapted such that for each l^{th} IAB, the error difference between the DNNs predicted output $S_e^{pred,p,l}$ and the desired $S_e^{des,p,l}$ output is minimum. Here, *p* = 1, 2,… , *Train*, the *Train* is the number of RF and the baseband beamforming vectors. Mathematically DNN is trained to minimize the error function *e(w)* for the set of different weight values of the DNN. The *e(w)* for the *lth* IAB can be written as in Eq. (9):

$$e(\theta) = \sum_{p=1}^{T_{rain}} M_{se}\left(S_e^{pred,p,l}, S_e^{des,p,l}\right) \tag{9}$$

here $M_{se}\left(S_e^{pred,p,l}, S_e^{des,p,l}\right)$ is the mean square error (MSE) between the predicted output $S_e^{pred,p,l}$ and the desired $S_e^{des,p,l}$ output. The optimal RF and the baseband precoding vectors get updated in the beamform codebook k_{RF}, k_{BB} respectively. From this stage onwards, the work of the Integrated RL (IRL) begins. The IRL employs the deep Q network for fine-tuning the beamform. The IRL goes through the optimized RF and baseband beamform codebook to further fine-tune the beamform vectors and steer the beam more precisely to the user location. The deep RL network employed in this work is shown in Figure 25. The objective of the IRL network is to maximize the received UEs SNR for the given beamform vectors. The received power at UE for the time instant t from the l^{th} IAB can be denoted as:

$$Pow_{UE}^{l}(t) = Pow_{IAB}^{l}(t)\left|h_{l,K}^{H}\, f_{RF}^{l}\, f_{BB}^{l}\right|^{2} \tag{10}$$

Here $Pow_{IAB}^{l}(t)$ is the transmitted power associated with the l^{th} IAB at time instant *t*. Now the received UEs SNR for the signal from the l^{th} IAB can be written as:

$$SNR^{l}(t) = \frac{Pow_{IAB}^{l}(t)\left|h_{l,K}^{H}(t)\, f_{RF}^{l}(t)\, f_{BB}^{l}(t)\right|^{2}}{\sigma_n^2} \tag{11}$$

The maximization problem for the IRL is as shown in Eq. (12) such that the transmission power of each IAB should be within the total permissible power of each IAB i.e. $Total_{power}$.

$$\begin{aligned} &\max_{Pow_{IAB}^{l}(t) f_{RF}^{l}(t)\, f_{BB}^{l}(t)} \sum_{l=1,2,3,\cdots,N} SNR^{l}(t) \\ &\text{s.t.} \quad Pow_{IAB}^{l}(t) \in Total_{power},\, f_{RF}^{l}(t)\, f_{BB}^{l}(t) \in k_{RF}(t)\, k_{BB}(t) \end{aligned} \tag{12}$$

In this research, an effort is made to fine-tune beam steering via DRL fed with the DNN optimized beamform vectors, UE location, and IAB transmit power to jointly control them to maximize Eq. (12). The IRL is an efficient deep learning Q network. The state-space τ associated with the IRL are $state_1^l$ indicating the transmission power of the l^{th} IAB, $state_2^l$ which is the beamform vectors associated with the l^{th} IAB. The $state_3^l$ indicates the deployed UE locations employed in the training. Action space A involves regulation in the transmission power and beamform vectors from the codebook k_{RF}, k_{BB} of the respective IABs.

5.3.4 Deep IRL Model

The deep IRL involves training of the Q Neural Network (NN) as shown in Figure 25. For the policy *po* the value of the state s and action a is given as $Q_{p_o}(s,a)$. For converging to the optimal state-action value $Q_{po}^{*}(s,a)$ we employed the NN architecture. The NN is defined by its weight values as we_t for each time step *t*. The $vec(we_t)$ is represented as we_t. The state action value in terms of NN weights can be represented as $Q_{po}(s,a : we_t)$ as shown in Algorithm 1 below.

The NN architecture of the deep IRL has the activator in the form of sigmoid function as $\gamma \rightarrow \frac{1}{1+\exp^{-\gamma}}$ [10]. Here the objective function for the NN of the IRL is to minimize the Mean Square Error (MSE) represented as:

$$\min_{we_t} Err(we_t) = \epsilon_{s,a}\left[\left(o_t - Q_{po}(s,a : we_t)\right)^{2}\right] \tag{13}$$

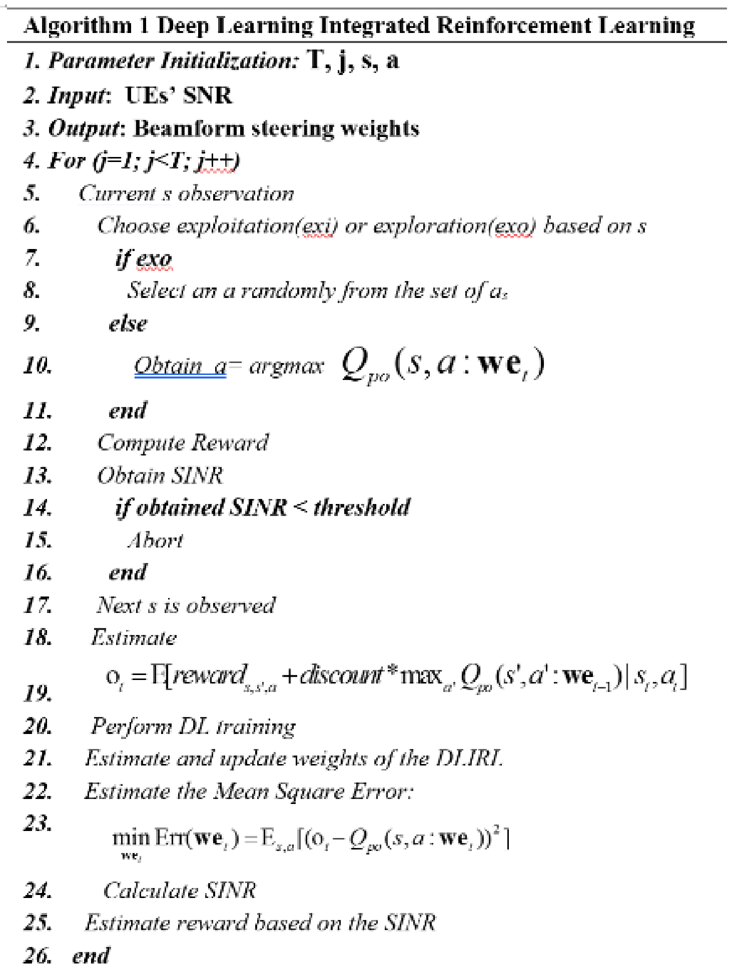
Here

$$o_t = \epsilon\Big[reward_{s,s',a} + discount * \max_{d'} Q_{po}(s', a' : we_{t-1}) \,|\, s_t, a_t\Big] \qquad (14)$$

The $reward_{s,s',a}$ is the reward for the agent post taking the action *a* and moving from state *s* at time *t* to state *s′*. The *a′* represents the next action to be taken. The weights of the NN are updated based on the gradient descent as:

$$we_{t+1} = we_t - step \bigtriangledown Err(we_t) \qquad (15)$$

Here, *step* indicates the step size having a value between 0 and 1. In the proposed IRL, the value $Q^*_{po}(s,a)$ is estimated based on the approximation $Q_{po}(s,a : we_t)$ that minimizes the MSE $Err(we_t)$. The proposed scheme is implemented corresponding to the downlink scenario. Here, IABs are separated at an estimated distance, and user equipment is positioned at a specific geographical location within the coverage of IABs. Also, these UEs move at a particular velocity. In this work, it is fixed at 20 mph. The reward function for the IRL is estimated based on the performance of an action on meeting the threshold *SINR*. The maximum reward at a unit time step is assigned to the agent having the best performance.



```
Algorithm 1 Deep Learning Integrated Reinforcement Learning
Parameter Initialization: T, j, s, a
Input: UEs' SNR
Output: Beamform steering weights
For (j=1; j<T; j++)
    Current s observation
    Choose exploitation(exi) or exploration(exo) based on s
        if exo
            Select an a randomly from the set of a_s
        else
            Obtain a= argmax $Q_{po}(s,a:\mathbf{we}_t)$
        end
    Compute Reward
    Obtain SINR
        if obtained SINR < threshold
            Abort
        end
    Next s is observed
    Estimate
        $o_t = E[reward_{s,s',a} + discount * \max_{a'} Q_{po}(s',a':\mathbf{we}_{t-1}) \,|\, s_t, a_t]$
    Perform DL training
    Estimate and update weights of the DLIRL.
    Estimate the Mean Square Error:
        $\min_{\mathbf{we}_t} Err(\mathbf{we}_t) = E_{s,a}[(o_t - Q_{po}(s,a:\mathbf{we}_t))^2]$
    Calculate SINR
    Estimate reward based on the SINR
end
```

5.4 Results

In this result analysis, we will examine the performance of Orthogonal Time Frequency Space (OTFS) signals compared to Orthogonal Frequency Division Multiplexing (OFDM) signals in terms of Bit Error

Rate (BER) as a function of Signal-to-Noise Ratio (SNR). The objective is to understand the conditions under which OTFS signals exhibit better performance than OFDM signals. OTFS is a relatively new modulation and coding scheme that has gained attention for its ability to combat severe multipath fading channels, which are common in wireless communication systems. OFDM, on the other hand, is a widely used modulation scheme known for its spectral efficiency. The goal of this analysis is to determine under which SNR conditions OTFS signals demonstrate superior BER performance compared to OFDM.

5.4.1 Experimental Setup:

Signal Generation: Simulated OTFS and OFDM signals with comparable parameters such as bandwidth and symbol rate.

Channel Model: Modelled a fading channel with severe multipath effects to stress-test both modulation schemes.

SNR Range: Varied the SNR over a wide range of values.

Error Detection: Employed suitable error detection and correction mechanisms, such as Viterbi decoding for OTFS and FFT-based equalization for OFDM.

Data Collection: Collected BER data points at each SNR level.

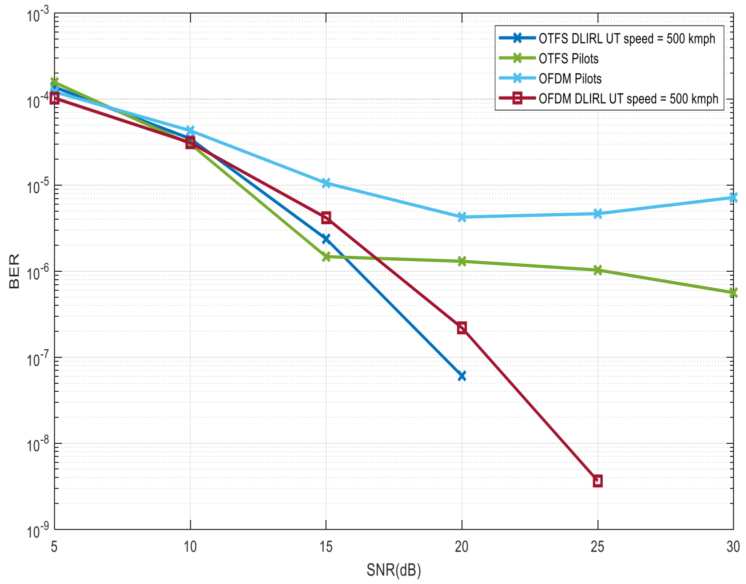


Figure 26: BER against SNR for OTFS and OFDM at 200 miles per hour

5.4.2 Results

The BER vs. SNR curves for both OTFS and OFDM signals were plotted and analysed. The following observations were made and shown in Figures 26-28:

Performance in Severe Fading Channels due to high speed: OTFS demonstrates significantly better BER performance than OFDM in environments with severe multipath fading. As SNR decreases and multipath effects become more pronounced, OTFS maintains a much lower BER compared to OFDM.

Robustness to Delay Spread: OTFS's performance advantage can be attributed to its inherent ability to combat delay spread effectively. It exploits the full spatial-temporal diversity of the channel, resulting in improved symbol recovery even in highly dispersive channels.

OFDM's Spectral Efficiency: In scenarios with high SNR and minimal multipath fading, OFDM may exhibit better spectral efficiency due to its higher data rate capabilities. However, this advantage diminishes as the channel conditions degrade due to higher speed.

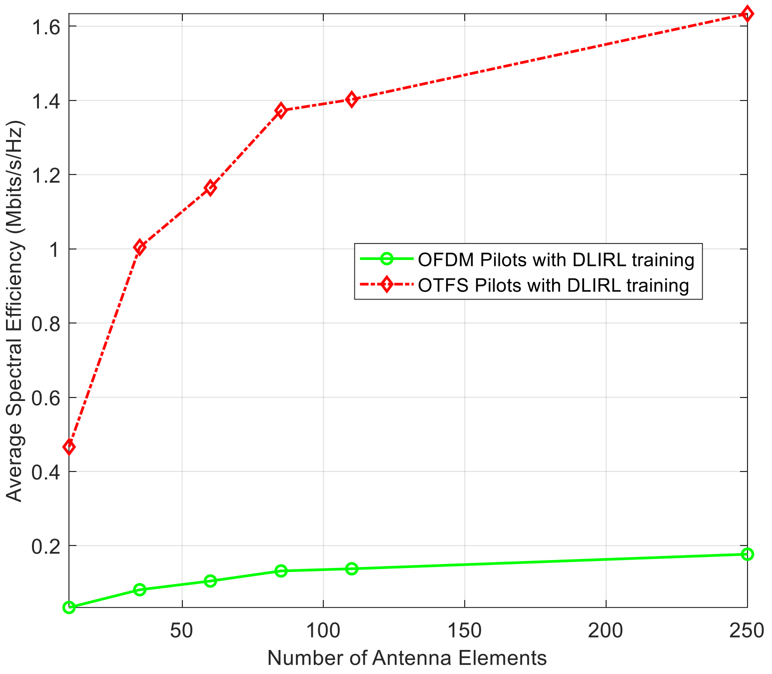


Figure 27: Average Spectral Efficiency against Number of Antenna Elements for OTFS and OFDM at 200 miles per hour

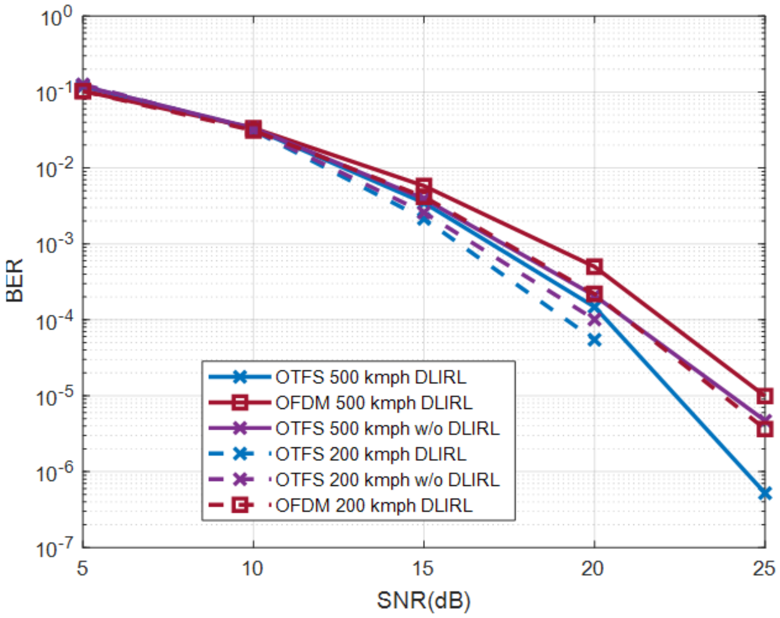


Figure 28: BER vs. SNR

5.5 Conclusion

The results of the BER vs. SNR analysis clearly indicate that OTFS signals outperform OFDM signals in channels characterized by severe multipath fading and delay spread. OTFS's unique spatial-temporal processing capabilities make it a compelling choice for wireless communication systems operating in challenging environments, such as urban or indoor scenarios. However, it's essential to consider the specific requirements and constraints of the communication system when choosing between OTFS and OFDM, as OFDM may still be preferred in scenarios with ample SNR and available bandwidth.

In summary, OTFS modulation offers a promising solution to address the challenges of multipath fading, and its superior BER performance at low SNR levels makes it a strong candidate for future wireless communication systems in demanding propagation conditions. Further research and practical implementations are needed to harness the full potential of OTFS in real-world applications.

5.6 References

[1] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, “On the road to 6g: Visions, requirements, key technologies, and testbeds,” IEEE Communications Surveys and Tutorials, vol. 25, no. 2, pp. 905–974, 2023.

[2] J. Moreno, J. M. Riera, L. d. Haro, and C. Rodriguez, “A survey on future railway radio communications services: challenges and opportunities,” IEEE Communications Magazine, vol. 53, no. 10, pp. 62–68, 2015.

[3] X. Liu and D. Qiao, “Location-fair beamforming for high speed railway communication systems,” IEEE Access, vol. 6, pp. 28632–28642, 2018.

[4] H. Song, X. Fang, and Y. Fang, “Millimeter-wave network architectures for future high-speed railway communications: Challenges and solutions,” IEEE Wireless Communications, vol. 23, no. 6, pp. 114–122, 2016.

[5] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, “Orthogonal time frequency space modulation,” in 2017 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6, 2017.

[6] L. Gaudio, G. Colavolpe, and G. Caire, “Otfs vs. ofdm in the presence of sparsity: A fair comparison,” IEEE Transactions on Wireless Communications, vol. 21, no. 6, pp. 4410–4423, 2022.

[7] R. Jiang, J. Zhao, Y. Xu, X. Wang, and L. Zhang, “Low-complexity beam selection scheme for high speed railway communications,” IEEE Access, vol. 8, pp. 16022–16032, 2020.

[8] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, “Fast beamforming design via deep learning,” IEEE Transactions on Vehicular Technology, vol. 69, no. 1, pp. 1065–1069, 2020.

[9] X. Zhu, F. Qi, and Y. Feng, “Deep-learning-based multiple beamforming for 5g uav iot networks,” IEEE Network, vol. 34, no. 5, pp. 32–38, 2020.

[10] C. Liu, W. Yuan, Z. Wei, X. Liu, and D. W. K. Ng, “Location-aware predictive beamforming for uav communications: A deep learning approach,” IEEE Wireless Communications Letters, vol. 10, no. 3, pp. 668–672, 2021.

[11] F. B. Mismar, B. L. Evans, and A. Alkhateeb, “Deep reinforcement learning for 5g networks: Joint beamforming, power control, and interference coordination,” IEEE Transactions on Communications, vol. 68, no. 3, pp. 1581–1592, 2020.

[12] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, “Deep learning coordinated beamforming for highly-mobile millimeter wave systems,” IEEE Access, vol. 6, pp. 37328–37348, 2018.

[13] H. A. Kassir, Z. D. Zaharis, P. I. Lazaridis, N. V. Kantartzis, T. V. Yioultsis, and T. D. Xenos, “A review of the state of the art and future challenges of deep learning-based beamforming,” IEEE Access, vol. 10, pp. 80869–80882, 2022.

[14] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, “Enet: A deep neural network architecture for real-time semantic segmentation,” 2016.

6 Advanced Test and Simulation Tools

Train and test of any AI driven algorithms in the future network requires proper data sets. The accuracy of the developed algorithms is dependent on the accuracy of these data sets and therefore, having access to them can be a valuable asset. Unfortunately, there are not many datasets available to be used in the public and the condition that those data sets have be recorded are fixed and can't be modified. The TM500 developed at Viavi along with its user interface is capable of producing this kind of data sets that might be used in development and train and test of the AI algorithms.

The TM500 Family of products are used for functional, system integration, capacity and stress testing in the lab and emulate up 1000's of mobile devices, across multiple cells to set-up and test 4G and 5G Base Stations.

We should pay attention that although in this report we are just looking into user feedback in the 5G networks, but we can provide a wider range of information that can be also used in xApps design and implementation in the RIC. RIC stands for RAN Intelligent Controller and is an open platform designed to host RAN control applications (called xApps) which will be developed by specialist software developers – sometimes external to the RIC vendor.

RIC enables the optimization of RAN resources through near real-time analytic processing and provides adaption recommendations. The RIC is usually cloud native, and a central component of an open and virtualized RAN network. The RIC is an O-RAN Alliance defined network component that aligns with 3GPP Release 15 and beyond, it supports network slicing, mobile broadband, mission critical communications etc. The data generated using the Viavi test equipment is also perfect to test the AI-based RAN control applications (xApps).

It helps operators to optimize and launch new services by allowing them to make the most of network resources. It also helps operators to ease network congestion.

6.1 Channel Model

In this part we will briefly look into the channel model specifications which is used in the user emulator. First, we should mention that the emulator basically do all the functions of a real user, including synchronization, transmitting RACH, attachment to the network with the related signalling and then providing the feedback according to the instructions of the gNB. The CSI feedback which is defined in 3GPP documents are quite complicated and depending on the different setting in the network may vary a lot. Therefore, the current UE emulator can provide interesting logs based on the settings of the gNB. The channel modelling includes large and small scale fading and A general multi antenna transmitter at the base station is considered. The model of the transmitter is the same as our previous work published in D4.2. In our model, the base station (BS) is equipped with N transmit antenna elements, which are arranged in a $N1 \times N2$ size panel of cross-polarised antenna pairs. Hence, $N = 2N1N2$, where $N1$ is the number of columns, $N2$ is the number of rows and 2 is the number of polarisations. The set of precoding matrices, also known as codebook, are set up and each precoding matrix is formed from basic beamforming vectors. Each beamforming vector is a discrete Fourier transform (DFT) vector of length $N1N2$ with oversampling factors $O1$ and $O2$, where $O1 = 4$ and Thus, there are $K = O1O2N1N2$ beamforming vectors in total. $O_1O_2N_1N_2$ beams point to $O_1O_2N_1N_2$ distinct directions. Depend on antenna structure at BS antenna panel, the angle width in azimuth and zenith dimensions, such as in the bounded area of directions, are not lower than the gain of the boresight direction subtracts 3dB, which is also called 3dB azimuth and zenith beamwidths. The 3dB azimuth bandwidth, denoted by $\emptyset_{3\mathrm{dB}}$, and 3dB zenith beamwidth denoted by $\theta_{3\mathrm{dB}}$ are the model parameters specified by end users. Furthermore, the gain to a direction is a composition of two factors, the gain due to directional antenna elements and the gain from a beamforming technique. This composite gain is represented by a single input by end users, G, which takes the value of $10\log_{10}(N_1N_2)$dB by default. Meanwhile, the model also takes the spherical coordinate of the UE, $(r, \emptyset, \theta)$, as the model input, where r is the distance between the BS and the UE, $\emptyset$ is the UE azimuth angle, and θ is the UE zenith angle. Provided that the Cartesian coordinate of the UE relative to the BS is (x, y, z), we have:

$$r = \sqrt{x^2 + y^2 + z^2};$$

$$\emptyset = \arg{(x + jy)};$$

$$\theta = \arccos{(\frac{z}{\sqrt{x^2+y^2+z^2}})},$$

where arg $(\cdot)$ denotes the argument (phase) of a complex number and arccos $(\cdot)$ is the inverse cosine function.

The developed model models the large scale (i.e., slow fading pathloss) and small scale (i.e., multipath fast fading) radio propagation effects. Slow fading pathloss model of each link follows the simple free space pathloss (FSPL) model, where the pathloss exponent is set to 2 and no shadowing effect is considered. The reason we only support pathloss exponent of 2 is because there is little benefit/effect for the modelling feature. Meanwhile, we do not introduce shadowing, aiming to avoid difficulty in a case that one wants to test a behaviour requiring a deterministic outcome. Fast fading, which is based on a stochastic model, should bring in sufficient degree of uncertainty for a scattering environment, and it will be used to reflect through the CSI and hybrid automatic repeat request (HARQ) acknowledgement (ACK)/negative-acknowledgement (NACK) feedback reporting sent to the BS. In combination with the channel SNR, channel gain coefficients are used to generate instantaneous SNR, which is then used directly to map to RI, CQI and used to populate HARQ ACK/NACK.

In our model, fast fading component adopts the tapped delay line (TDL) models following the 3GPP specification. Each TDL model is associated with a channel delay profile, namely TDL-A, TDL-B, TDL-C, TDL-D or TDL-E. When a channel delay profile is specified to a link between the BS antenna and a UE's antennas, the channel is time dispersive and its impulse response has several random components, each is a Jakes process, at different delays. The variances of those random components are allocated according to the delay profile and the sum of variances is KdB lower than the deterministic component cause by LoS path. Since each channel delay profile specified in the 3GPP standard only defines the powers at normalised delays, a parameter called delay scale needs to be set to derive the actual delays for all the taps.

The tool performs a system level emulation for the users and is constantly evolving to capture more complicated channel types and improve the accuracy of the modelling. The current Mobility and Traffic Simulation (MTS) provides the ability to model large scale fading (via path loss) and small-scale fading (i.e., multi-path fast fading).

The fast-fading model effect is reflected in:

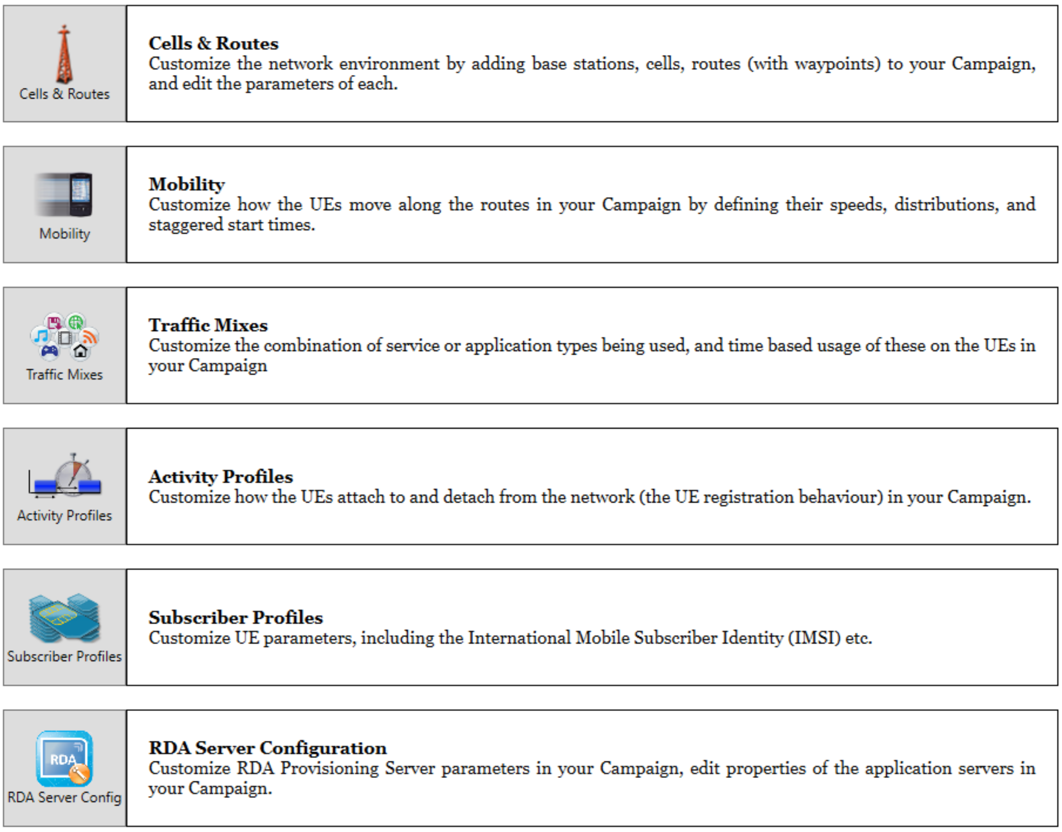
- The CSI report (i.e., RI and CQI)
- The PDSCH ACK/NACK reported.
- The PUSCH frequency response

For the fast fading and beam selection, some simplifying assumptions are made to make it feasible to model the channel of thousands of users in the hardware. Otherwise, calculations of the CQI and PMI for the users will be prohibitively complicated and impossible to be carried on the available servers.

Certain state of the art algorithms have been developed to keep the emulator practical and make it possible to emulate the behaviours of several UEs in networks with different settings.

6.2 The main GUI and its parameters

The GUI is a convenient way to define different basic parameters of the system and set up the scenario. The main panel of the GUI is shown in the following figure 29.



Cells & Routes
Customize the network environment by adding base stations, cells, routes (with waypoints) to your Campaign, and edit the parameters of each.

Mobility
Customize how the UEs move along the routes in your Campaign by defining their speeds, distributions, and staggered start times.

Traffic Mixes
Customize the combination of service or application types being used, and time based usage of these on the UEs in your Campaign

Activity Profiles
Customize how the UEs attach to and detach from the network (the UE registration behaviour) in your Campaign.

Subscriber Profiles
Customize UE parameters, including the International Mobile Subscriber Identity (IMSI) etc.

RDA Server Configuration
Customize RDA Provisioning Server parameters in your Campaign, edit properties of the application servers in your Campaign.

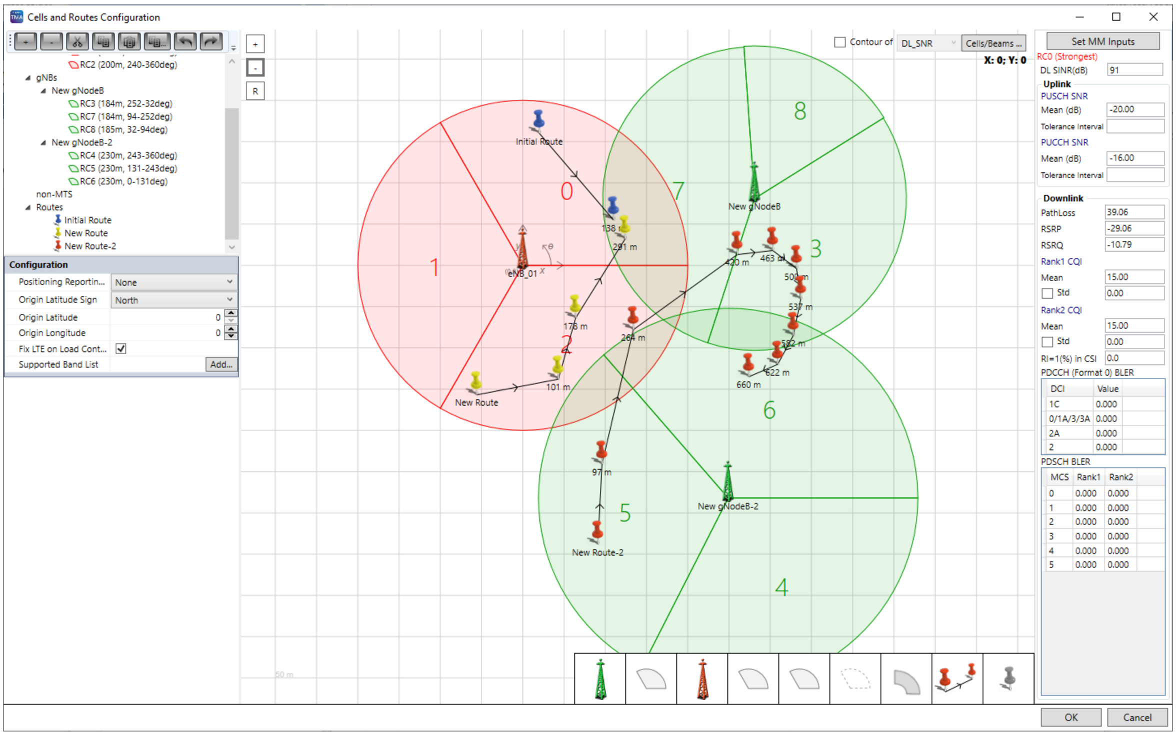
Figure 1 The main settings in the UE emulator GUI

Each new scenario is called a campaign. In each campaign we can define a new setup with a different setting. The available settings are:

- Cells and routes
- Mobility
- Traffic mixes
- Activity profile
- Subscriber profile
- RDA server configuration

Among these, the first four setting are more relevant to our project.

According to Figure 29, the layout of the network is defined in the "cell & routes". By clicking it, we will see the following window shown below in Figure 30.



- RC2 (200m, 240-360deg)
- gNBs
  - New gNodeB
    - RC3 (184m, 252-32deg)
    - RC7 (184m, 94-252deg)
    - RC8 (185m, 32-94deg)
  - New gNodeB-2
    - RC4 (230m, 243-360deg)
    - RC5 (230m, 131-243deg)
    - RC6 (230m, 0-131deg)
- non-MTS
- Routes
  - Initial Route
  - New Route
  - New Route-2

Configuration

Positioning Reportin...	None
Origin Latitude Sign	North
Origin Latitude	0
Origin Longitude	0
Fix LTE on Load Cont...	☑
Supported Band List	Add...

RC0 (Strongest)

DL SINR(dB): 91

Uplink

PUSCH SNR
Mean (dB): -20.00
Tolerance Interval:

PUCCH SNR
Mean (dB): -16.00
Tolerance Interval:

Downlink

PathLoss: 39.06
RSRP: -29.06
RSRQ: -10.79

Rank1 CQI
Mean: 15.00
☐ Std: 0.00

Rank2 CQI
Mean: 15.00
☐ Std: 0.00

RI=1(%) in CSI: 0.0

PDCCH (Format 0) BLER

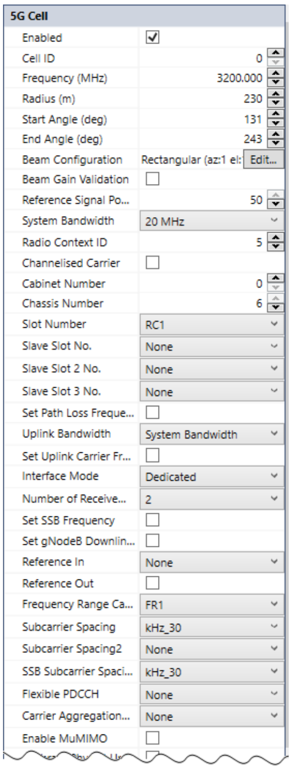
DCI	Value
1C	0.000
0/1A/3/3A	0.000
2A	0.000
2	0.000

PDSCH BLER

MCS	Rank1	Rank2
0	0.000	0.000
1	0.000	0.000
2	0.000	0.000
3	0.000	0.000
4	0.000	0.000
5	0.000	0.000

Figure 2 Cells and routes configurations

In the cells and routes configurations, we can define the layout of the network. This include the location of the base stations, the type of the base station (4G/5G), number of sectors, the radius and ... For example, In the shown set up, we have 3 different cells that one cell, shown in red is a 4G cell and the other two which are shown in light green, are the 5G cells. Each cell can have a separate dedicated setting (See Figure 31 below).

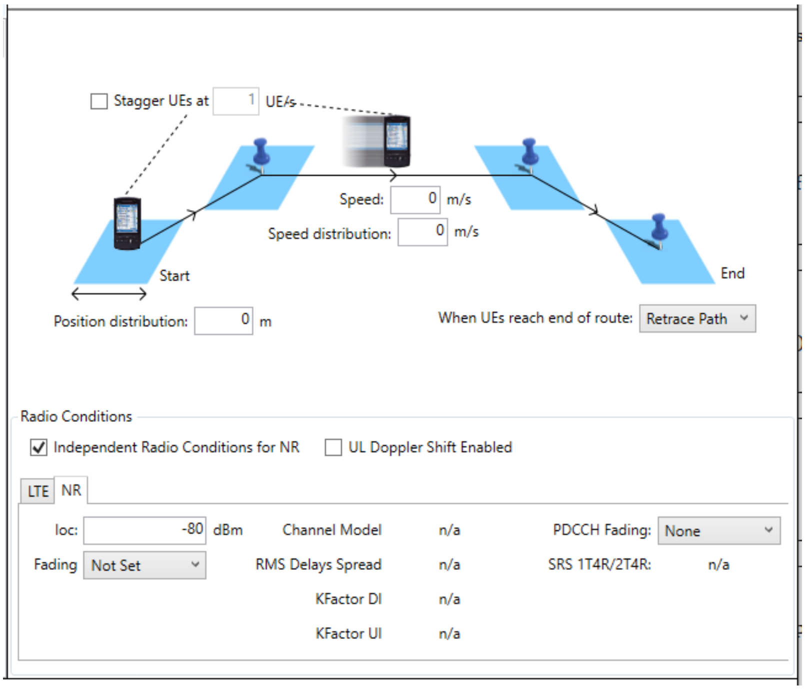


5G Cell

Enabled	☑
Cell ID	0
Frequency (MHz)	3200.000
Radius (m)	230
Start Angle (deg)	131
End Angle (deg)	243
Beam Configuration	Rectangular (az:1 el: Edit...
Beam Gain Validation	☐
Reference Signal Po...	50
System Bandwidth	20 MHz
Radio Context ID	5
Channelised Carrier	☐
Cabinet Number	0
Chassis Number	6
Slot Number	RC1
Slave Slot No.	None
Slave Slot 2 No.	None
Slave Slot 3 No.	None
Set Path Loss Freque...	☐
Uplink Bandwidth	System Bandwidth
Set Uplink Carrier Fr...	☐
Interface Mode	Dedicated
Number of Receive...	2
Set SSB Frequency	☐
Set gNodeB Downlin...	☐
Reference In	None
Reference Out	☐
Frequency Range Ca...	FR1
Subcarrier Spacing	kHz_30
Subcarrier Spacing2	None
SSB Subcarrier Spaci...	kHz_30
Flexible PDCCH	None
Carrier Aggregation...	None
Enable MuMIMO	☐

Figure 3. 5G NR gNB settings

In addition to the layout of the network, we can define different routes in the network along which users will travel. We can define the routes by setting waypoints in a trajectory. Then using the Mobility tab, we can customize how the UEs move along the routes in your Campaign by defining their speeds, distributions, and staggered start times. A snapshot of this tab and some of the settings is shown in Figure 4. According to the inputs in this window, the velocity of the users is calculated between the waypoints and then then is converted to radial velocity and therefore, the doppler values for different users are calculated. From the input parameters, the mobility model will run online and produce a series of instantaneous CSI reports based on the user's configuration, which include UE's location, UE specified RI, wideband PMI, wideband and sub-band CQI, and HARQ ACK/NACK status when requested.



☐ Stagger UEs at 1 UE/s

Speed: 0 m/s

Speed distribution: 0 m/s

Position distribution: 0 m

When UEs reach end of route: Retrace Path

Radio Conditions

☑ Independent Radio Conditions for NR ☐ UL Doppler Shift Enabled

Ioc: -80 dBm

Fading Not Set

Channel Model n/a

RMS Delays Spread n/a

KFactor Dl n/a

KFactor Ul n/a

PDCCH Fading: None

SRS 1T4R/2T4R: n/a

Figure 4. UE Mobility setting

We can further change the traffic type of the UEs and also define their internal capabilities, but these settings are out of the scope of our specific task.

6.3 Simulation implementation

The purpose of this report is to show how we can log user generated feedback information in the current commercial UE emulator and use the information to train different AI based algorithms using them. We should mention that to do so, we have made some changes in the settings to enable logging the channel state information for all UEs in the network. We get the log directly from the emulator and in fact we do not transmit them in the uplink. Transmitting to the uplink will be time consuming and needs extra effort to analyse and extract the required info. A sample of the generated log for a network with following settings is generated as depicted in the picture.

Simulation settings:

Number of NR cells, : 2 [cell id= 151 , 4.8GHz carrier freq]

Number of NR sectors: 4 (sec1: 0-90 deg, sec2: 90-180deg; sec=180-270deg; sec4=270-360deg)

Antenna panel spec: 32 ports

Number of Users (UEs): 64

Trajectory of the users : (cartesian plane) gNB position: (0,0), (500,0)

UE position: start from (200, 200) to (500, 500) and back in loop between those way points

CSI setting:

cri-RI-PMI-CQI

cqi-FormatIndicator widebandCQI

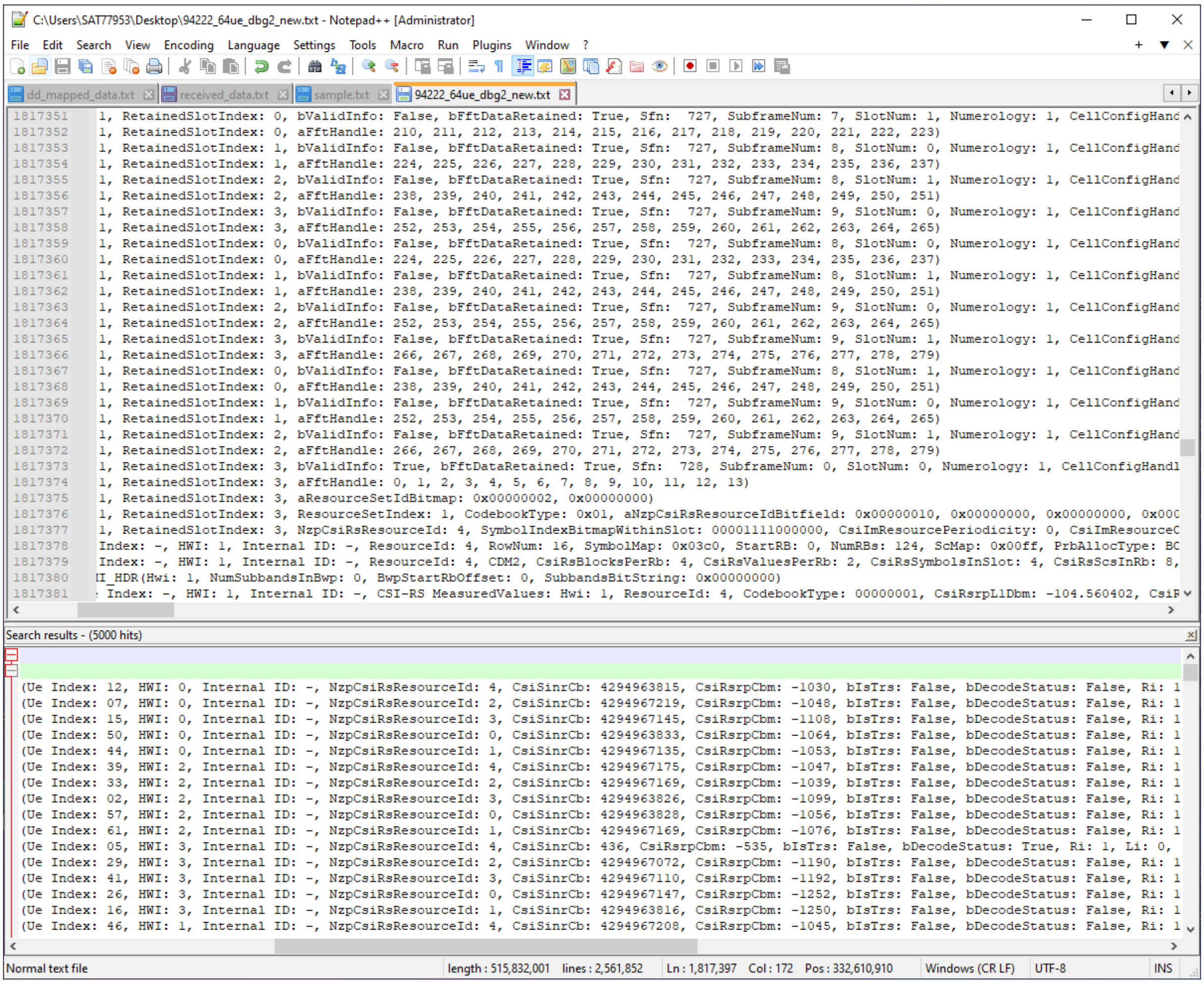
pmi-FormatIndicator widebandPMI

n1-n2 eight-two-TypeI-SinglePanel

codebookMode 1

aperiodicTriggerStateList setup

Mobility profile: AWGN Downlink slow fading enabled in UEs.



```
1, RetainedSlotIndex: 0, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 7, SlotNum: 1, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 0, aFftHandle: 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223)
1, RetainedSlotIndex: 1, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 8, SlotNum: 0, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 1, aFftHandle: 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237)
1, RetainedSlotIndex: 2, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 8, SlotNum: 1, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 2, aFftHandle: 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251)
1, RetainedSlotIndex: 3, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 9, SlotNum: 0, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 3, aFftHandle: 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265)
1, RetainedSlotIndex: 0, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 8, SlotNum: 0, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 0, aFftHandle: 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237)
1, RetainedSlotIndex: 1, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 8, SlotNum: 1, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 1, aFftHandle: 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251)
1, RetainedSlotIndex: 2, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 9, SlotNum: 0, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 2, aFftHandle: 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265)
1, RetainedSlotIndex: 3, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 9, SlotNum: 1, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 3, aFftHandle: 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279)
1, RetainedSlotIndex: 0, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 8, SlotNum: 1, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 0, aFftHandle: 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251)
1, RetainedSlotIndex: 1, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 9, SlotNum: 0, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 1, aFftHandle: 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265)
1, RetainedSlotIndex: 2, bValidInfo: False, bFftDataRetained: True, Sfn:  727, SubframeNum: 9, SlotNum: 1, Numerology: 1, CellConfigHand
1, RetainedSlotIndex: 2, aFftHandle: 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279)
1, RetainedSlotIndex: 3, bValidInfo: True, bFftDataRetained: True, Sfn:  728, SubframeNum: 0, SlotNum: 0, Numerology: 1, CellConfigHandl
1, RetainedSlotIndex: 3, aFftHandle: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13)
1, RetainedSlotIndex: 3, aResourceSetIdBitmap: 0x00000002, 0x00000000)
1, RetainedSlotIndex: 3, ResourceSetIndex: 1, CodebookType: 0x01, aNzpCsiRsResourceIdBitfield: 0x00000010, 0x00000000, 0x00000000, 0x000
1, RetainedSlotIndex: 3, NzpCsiRsResourceId: 4, SymbolIndexBitmapWithinSlot: 00001111000000, CsiImResourcePeriodicity: 0, CsiImResourceC
Index: -, HWI: 1, Internal ID: -, ResourceId: 4, RowNum: 16, SymbolMap: 0x03c0, StartRB: 0, NumRBs: 124, ScMap: 0x00ff, PrbAllocType: BC
Index: -, HWI: 1, Internal ID: -, ResourceId: 4, CDM2, CsiRsBlocksPerRb: 4, CsiRsValuesPerRb: 2, CsiRsSymbolsInSlot: 4, CsiRsScsInRb: 8,
I_HDR(Hwi: 1, NumSubbandsInBwp: 0, BwpStartRbOffset: 0, SubbandsBitString: 0x00000000)
 Index: -, HWI: 1, Internal ID: -, CSI-RS MeasuredValues: Hwi: 1, ResourceId: 4, CodebookType: 00000001, CsiRsrpL1Dbm: -104.560402, CsiR
```

```
(Ue Index: 12, HWI: 0, Internal ID: -, NzpCsiRsResourceId: 4, CsiSinrCb: 4294963815, CsiRsrpCbm: -1030, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 07, HWI: 0, Internal ID: -, NzpCsiRsResourceId: 2, CsiSinrCb: 4294967219, CsiRsrpCbm: -1048, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 15, HWI: 0, Internal ID: -, NzpCsiRsResourceId: 3, CsiSinrCb: 4294967145, CsiRsrpCbm: -1108, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 50, HWI: 0, Internal ID: -, NzpCsiRsResourceId: 0, CsiSinrCb: 4294963833, CsiRsrpCbm: -1064, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 44, HWI: 0, Internal ID: -, NzpCsiRsResourceId: 1, CsiSinrCb: 4294967135, CsiRsrpCbm: -1053, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 39, HWI: 2, Internal ID: -, NzpCsiRsResourceId: 4, CsiSinrCb: 4294967175, CsiRsrpCbm: -1047, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 33, HWI: 2, Internal ID: -, NzpCsiRsResourceId: 2, CsiSinrCb: 4294967169, CsiRsrpCbm: -1039, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 02, HWI: 2, Internal ID: -, NzpCsiRsResourceId: 3, CsiSinrCb: 4294963826, CsiRsrpCbm: -1099, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 57, HWI: 2, Internal ID: -, NzpCsiRsResourceId: 0, CsiSinrCb: 4294963828, CsiRsrpCbm: -1056, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 61, HWI: 2, Internal ID: -, NzpCsiRsResourceId: 1, CsiSinrCb: 4294967169, CsiRsrpCbm: -1076, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 05, HWI: 3, Internal ID: -, NzpCsiRsResourceId: 4, CsiSinrCb: 436, CsiRsrpCbm: -535, bIsTrs: False, bDecodeStatus: True, Ri: 1, Li: 0,
(Ue Index: 29, HWI: 3, Internal ID: -, NzpCsiRsResourceId: 2, CsiSinrCb: 4294967072, CsiRsrpCbm: -1190, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 41, HWI: 3, Internal ID: -, NzpCsiRsResourceId: 3, CsiSinrCb: 4294967110, CsiRsrpCbm: -1192, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 26, HWI: 3, Internal ID: -, NzpCsiRsResourceId: 0, CsiSinrCb: 4294967147, CsiRsrpCbm: -1252, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 16, HWI: 3, Internal ID: -, NzpCsiRsResourceId: 1, CsiSinrCb: 4294963816, CsiRsrpCbm: -1250, bIsTrs: False, bDecodeStatus: False, Ri: 1
(Ue Index: 46, HWI: 1, Internal ID: -, NzpCsiRsResourceId: 4, CsiSinrCb: 4294967208, CsiRsrpCbm: -1045, bIsTrs: False, bDecodeStatus: False, Ri: 1
```

Figure 5: UE log generated by the UE emulator servers.

7 Summary

These deliverable details the research performed in Work Package 4 during the second period of the 6G Brains project. The research topics included modelling and analysis of the different 6G potential architectures, access technologies and AI based modules as follows:

1. The intelligent cell-free IAB scheduler for traffic routing to maximize the network's arrival ratio while minimizing the average packet delay,
2. The grant-free NOMA for massive machine-type communication integrated with intelligent beamforming,
3. The AI based hybrid user terminal modelling and resource allocation for the cell-free network with underlaid D2D,
4. Intelligent IAB with beam steering based on user location and
5. the advanced test and simulation tools for multi-UE mobility behaviours modelling and channel evaluation.

The main content of the deliverable is described below.

1. In this deliverable, the problem of routing in an IAB network is examined in which multiple IABs nodes operate simultaneously within the same network in order to route multiple packets towards their destinations. To identify the joint routing policy that maximizes the network's arrival ratio while minimizing the average packet delay, a novel Relational A2C algorithm is developed to determine the best joint routing strategy, based on observations collected by the IABs nodes via online learning. To support decentralized training, two different approaches are developed, which achieve similar performance to the centralized training approach. For various scenarios, the arrival ratio and average delay of the proposed Relational A2C algorithms are compared with six baseline algorithms. It was found that Relational A2C performed better than baseline algorithms in all cases and achieved performance similar to that of a centralized approach. Further, it is demonstrated that network routing is crucial to reducing congestion and maximizing the utilization of network resources in IAB networks. These results clearly demonstrate the ability of Relational A2C based algorithms to learn near-centralized policies and the overall superiority of the proposed approach over existing methods.
2. A novel architecture for the integration of the intelligent beamforming and compressed sensing technique is proposed for the grant-free NOMA in a MIMO cellular network. K-means method is employed for the user clustering for the spatial division multiple access for the spatially distributed machine-type users. In particular, a joint adaptive beamforming and subspace pursuit method is developed for the blind multiuser detection and data recovery with the help of a novel user sparsity decision method. With the beamforming, the spectral efficiency is improved by supporting more users in the same frequency spectrum. In addition, both the detection error rate and the symbol error rate can be significantly reduced with the integration and data interaction between the beamforming and subspace pursuit methods.
3. In order to efficiently design beamforming along with power allocation for CF-mMIMO with underlaid D2D systems, we proposed an advanced strategy relying on heterogeneous graph neural network. The CFD-HetGNN consists of aggregation plus combination layers and DNN modules, and it was shown to significantly outperform traditional MF method with equal

power allocation. In near future, we plan to equip our proposed wireless network with the capability of sensing, that is the integrated sensing and communication for CF-mMIMO with underlaid D2D. To address the complicated resource allocation and beamforming design in such a scenario, we will further improve the CFD-HetGNN model.

4. A novel DLIRL-based beam steering scheme is implemented, which is capable to steer the beamform with the user movement. The proposed scheme combines the performance of the DL and RL. The DL is used for preparing the optimized beamforming codebook and RL is used for selecting the best beam out of the optimized beamforming codebook based on the user movements. The training of the DLIRL is carried out based on the channel information without requiring the user's location. Thus, reducing the training overheads. A novel way of combining the DL with the RL and employing it for beam steering applications. The proposed scheme is trained in the off-line mode for a particular environment simulated using the MATLAB site viewer. The trained model is the employed successfully for the beam steering. Furthermore, the proposed scheme is sufficiently flexible for getting trained to any provided environment.
5. The TM500 Test and Simulation tool was developed along with its user interface, capable of producing this kind of data sets that might be used in development and train and test of the AI algorithms. The TM500 Family of products are used for functional, system integration, capacity and stress testing in the lab and emulate up 1000's of mobile devices, across multiple cells to set-up and test 4G and 5G Base Stations. In this report, it looks into user feedback in the 5G networks, but a wider range of information can be provided, that can be also used in xApps design and implementation in the RIC (RAN Intelligent Controller). The data generated using the test equipment is also perfect to test the AI-based RAN control applications (xApps). It helps operators to optimize and launch new services by allowing them to make the most of network resources. It also helps operators to ease network congestion.