Optimizing Sum Rate in Satellite-HAP Communication Systems Using Machine Learning

Submitted in partial fulfilment of the requirements for the award of degree of

Bachelor of Technology

In

Electronics and Communication Engineering

Submitted by

Challa Gopala Krishna Reddy (21ECB0B09)

Shrikar Kaveti (21ECB0B21)

Kalikineedi Durga Krishna Thanmay (21ECB0B24)

Supervisor

Prof. Dr. Sourabh Solanki

Assistant Professor

DEPARTMENT OF ELECTRONICS AND COMMUNICATION
NATIONAL INSTITUTE OF TECHNOLOGY WARANGAL

2021-2025

CERTIFICATE

This is to certify that the project work entitled "**Optimizing Sum Rate in Satellite-HAP Communication Systems Using Machine Learning**" is a bonafide record of work carried out by **Challa Gopala Krishna Reddy (21ECB0B09), Shrikar Kaveti (21ECB0B21), Kalikineedi Durga Krishna Thanmay (21ECB0B24)** submitted to the faculty of the Department of Electronics and Communication Engineering in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Electronics and Communication Engineering at National Institute of Technology, Warangal for the academic year 2024-2025.

Head of the Department
Prof. Dr. D. Vakula
Professor
Dept. of Electronics and
Communication Engineering,
NITW

Project Superior
Prof. Dr. Sourabh Solanki
Assistant Professor
Dept. of Electronics and
Communication Engineering,
NITW

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honestyand integrity and have not misrepresented or fabricated or falsified any idea / data / fact / source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from thesources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature of Student)

Name - Challa Gopala Krishna Reddy

Roll No – 21ECB0B09

Date – 21 / 04 / 2025

(Signature of Student)

Name – Shrikar Kaveti

Roll No – 21ECB0B21

Date – 21 / 04 / 2025

(Signature of Student)

Name - Kalikineedi Durga Krishna Thanmay

Roll No – 21ECB0B24

Date – 21 / 04 / 2025

ACKNOWLEDGEMENT

We wholeheartedly thank our guide **Prof. Dr. Sourabh Solanki**, **Assistant Professor**, Department of Electronics and Communication Engineering, for being a source of motivation, constant supervision and invaluable encouragement, and for guiding us with perfection in doing work.

We are grateful to the **Electronics and Communication Engineering Department**, for providing us with all the facilities to carry out this project work. We are immensely thankful to **Prof. Dr. D. Vakula**, Head of Department of Electronics and Communication Engineering for his moral support during the total span of the project.

We are very thankful to **Prof. Dr. Farukh Hashmi**, Project Coordinator for his continuous support and guidance during the project and to the Project Evaluation Committee, for their strenuous efforts to evaluate our project.

List Of Figures

Table of Contents

Abbreviations

URLLC - Ultra-Reliable Low Latency Communication

eMBB - Enhanced Mobile Broadband

HAP - High Altitude Platform

RAN - Radio Access Network

GEO - Low Earth Orbit

NTN - Non-Terrestrial Networks

IAB - Integrated Access and Backhaul

OMA - Orthogonal Multiple Access

NOMA - Non-Orthogonal Multiple Access

Abstract

The demand for communication in remote and underserved regions underscores the need for satellite and HAP-based networks as a practical solution. This project investigates the optimization of uplink communication rates within a GEO-HAP-user network, focusing on maximizing the sum rate from users. With URLLC standards facilitating GEO-to-HAP connections and eMBB managing HAP-to-user links, our work defines a system model that incorporates communication constraints, framing an optimization problem to enhance user data rates effectively.

Machine learning techniques are applied to manage and allocate resources in this complex communication setup. Synthetic data generated from realistic scenarios—considering parameters like signal-to-noise ratio (SNR), bandwidth allocation, and user distribution—enables development of a machine learning model trained to optimize key variables, such as power allocation and bandwidth. The model's aim is maximization of sum rate to users under varying network conditions, offering a flexible and adaptive approach to uplink optimization.

Evaluating the model's performance across key metrics like sum rate, latency, and resource utilization, we compare the results with those from classical optimization methods. Findings demonstrate the potential of ML-driven optimization for high-efficiency, satellite-based networks, presenting a scalable solution to enhance communication in areas lacking terrestrial infrastructure and providing a foundation for future advancements in connectivity technology.

Chapter 1 - Introduction

1.1 Problem Statement

In present times, Terrestrial Communication is the dominant mode of communication dues to its low cost and practicality for implementation in crowded and urban landscape but there are some limitations for the terrestrial network as some area does not have plain area and some areas very remote because of which transportation and installation of terrestrial terminal are very expensive. Another major problem is the areas damaged due to disaster.

For such areas, Satellite - User Terminal Communication can be very practical and recent studies have suggested that the Satellite - HAP - User Network can be very practical. In our project, we aim to use a Satellite - HAP - User Terminal Network which utilizes Ultra-Reliable Low-Latency Communication (URLLC) between Satellite and HAP in order to control the HAP and Enhanced Mobile Broadband (eMMB) between HAP and User to transfer data between the users and HAP. we are focusing on optimizing the bandwidth, power utilization by HAP for each user and satisfy the eMMB threshold for all the users.

1.2 Objective

1. Develop a communication system using GEO satellite and High-Altitude Platforms (HAPs) to support communication in remote or under-served areas.

2. Formulate and implement machine learning models to optimize the sum rate for end users in a GEO-HAP-User communication network.

3. Evaluate the performance of the proposed machine learning framework using metrics - sum rate, latency, and other relevant metrics in URLLC and eMMB scenarios.

4. Compare the performance of the proposed machine learning framework with iterative optimization algorithm.

1.3 Motivation

Traditional Terrestrial Networks works efficient and cost-effective way in urban and clustered areas but fails in the remote and parsley populated areas. For this problem, now the telecommunication industry is adapting to use the satellite (GEO and GEO) for the under-served regions. recent papers have suggested that instead of using direct satellite to user terminal communication, Satellite-HAP-User can be a more practical and cost-effective solution for remote areas and now the industry is moving towards the 6G communication, the requirement for high throughput and reliable communication service is more than ever. By using the eMMB to provide high throughput to each user and URLLC to control the HAP the requirements for the 6G can be satisfied. The challenge for such a system is to get a optimized solution with low latency. Hence, instead of using the iterative approach to calculate the power distribution and bandwidth distribution. we propose using Machine Learning model to calculate the optimized solution with low latency in a cost-effective way. Thereby ensuring to achieve an optimal and scalable resource allocation for robust and efficient communication.

1.4 Background

1. **Geostationary Earth Orbit (GEO) Satellites**

   Geostationary Earth Orbit (GEO) satellites are satellites that are present about 35,000 in height above the earth. They stay in fixed position w.r.t Earth surface. Due to this, they do not have to perform handovers frequently. They are used mainly for broadcasting, weather monitoring and long-distance communications.

2. **High Altitude Platforms (HAPs)**

   HAPs, typically deployed in the stratosphere at altitudes between 17 to 22 kilometres, serve as intermediaries between satellites and end users. These platforms can provide broad coverage, especially over remote or underserved regions, and can dynamically adapt to traffic demand by repositioning themselves as needed. HAPs are integral to the GEO-HAP-user communication architecture, supporting both URLLC and eMBB applications by handling large data transfers from GEO satellites and distributing them to ground users. Challenges include power constraints, atmospheric conditions, and regulatory considerations, all of which are critical to ensuring consistent and reliable communication performance.

3. **Ultra-Reliable Low-Latency Communication (URLLC)**

   URLLC is a core component of next-generation communication standards, including 6G, designed to meet the stringent requirements of mission-critical scenarios. In the GEO-HAP-user communication model, URLLC is utilized between the GEO satellite and HAP to ensure high reliability and minimal delay, enabling seamless data transfer from the satellite to the platform, even in challenging environments.

4. **Enhanced Mobile Broadband (eMBB)**

   eMBB goal is delivering large user throughput and capacity, essential for applications like video streaming, virtual reality, and massive IoT. In the context of the GEO-HAP-user communication system, eMBB operates between the HAP and end users, facilitating high-throughput, reliable data transmission over large areas. The eMBB link between HAP and users must handle variations in demand and network conditions, necessitating intelligent resource management strategies to optimize performance, especially in high-demand areas or during peak times.

5. **Machine Learning in GEO-HAP Communication Systems**

   Machine learning (ML) plays a vital role in optimizing resource allocation in complex and dynamic network environments like the GEO-HAP-user system. ML techniques can analyse network parameters, user demand, and channel conditions to maximize the sum rate for users by adjusting resources in real-time. In this context, ML-based approaches are beneficial for managing handovers, minimizing latency, and dynamically allocating bandwidth, improving overall system performance compared to traditional allocation techniques. By integrating ML, the system can adapt to changing conditions and maintain high-quality service, particularly in environments where conventional resource management falls short.

1.5 Applications

1. **Maritime and Aviation Communication**

   Ships and aircraft operating in remote or international waters and airspaces often lack consistent connectivity. The GEO-HAP-user system can support high-speed data transfer in these areas, enhancing safety and operational efficiency in sectors like commercial shipping, fishing, and aviation.

2. **Massive IoT Connectivity**

   For large-scale IoT applications that require widespread, low-latency connections—such as environmental monitoring, agriculture, and supply chain logistics—a GEO-HAP-user network can provide connectivity to distributed sensors and devices, enabling real-time data collection and processing.

1.6 Merits

1. **Scalable Coverage**

   A network of GEO satellites and HAPs can provide scalable coverage, giving good connectivity.

2. **Resilience and Redundancy**

   The dual-layer architecture, with GEO satellites working alongside HAPs, enhances network resilience, as failures or load imbalances in one layer can be compensated for by the other. This design ensures continuous connectivity even in scenarios where parts of the network may experience outages.

3. **Flexibility and Adaptability with ML Integration**

   Machine learning optimizes resource allocation and channel management helping for real-time solutions.

1.7 Demerits

1. **High Deployment and Maintenance Costs**

   Establishing a satellite is a costly process and development and deployment of HAPs involve significant costs. Bad weather can affect HAPs, due to which maintenance is difficult.

2. **Frequent Satellite Handover and Tracking Requirements**

   GEO satellites move quickly across the sky, necessitating frequent handovers and complex tracking systems to maintain continuous connectivity. This need for high-precision tracking adds to system complexity and may introduce delays or disruptions in service if not managed effectively.

3. **Power and Capacity Limitations of HAPs**

   The power in a HAP is limited and cannot be changed easily, resulting in restriction of their capacity.

4. **Privacy and Security Challenges**

   Aerial communication networks like GEO satellite or HAPs are highly prone to data interception and attacks from hackers. Good protocols mut be in place to avoid problems.

Chapter 2 - Literature Review

1. **Title -** eMBB-URLLC Resource Slicing: A Risk-Sensitive Approach
   Year – 2019
   Type of Publication – IEEE COMMUNICATIONS LETTERS, VOL. 23, NO. 4, APRIL 2019
   Authors – Madyam Alsenwi, Nguyen H. Tran, Mehdi Bennis, Anupam Kumar Bairagi, Choong Seon Hong

   ## Summary
   URLLC and eMBB are 2 of 3 core services the 5G aims to provide. In general, eMBB traffic is continuous whereas URLLC is infrequent but has tight constraints regarding latency so they should be addressed without fail. This paper proposes a risk-sensitive approach where resources are allocated to URLLC traffic while maintaining a minimum data for eMBB users.

   ## Methodology
   In order to solve this problem, tit is divided into 2 subproblems:
   eMBB Scheduling: This is done to optimize the data rates for eMBB-based devices. This problem generally uses integer programming, which is computationally intense. By relaxing integer constraints, the problem becomes a convex problem, which helps us achieve solutions efficiently.

   URLLC Placement: In order to satisfy URLLC demands, the authors use a chance-constraint approach which is relaxed using Markov's Inequality. Conditional Value at Risk (CVaR) is used to manage risk of eMBB users when URLLC traffic is incoming.

   These problems are solved iteratively, allowing the algorithm to reach a convergent solution that satisfies both requirements.

   ## Result
   The methodology shows that dynamic multiplexing of URLLC traffic, achieved by selectively puncturing eMBB resources, can maximize eMBB data rate while making sure URLLC requirements are satisfied. The Integration of CVaR as a risk measure improves system performance by stabilizing data rates for eMBB traffic despite the URLLC traffic.

   ## Conclusion

   The study presents a practical framework for balancing 5G service requirements by using a risk-sensitive resource allocation strategy. By breaking down the complex problem into 2 subproblems, we maximize eMBB data throughput and at same time maintaining URLLC reliability. The risk-sensitive strategy by using CVaR ensures optimized resource allocation without undermining performance for any other type of service.

2. **Title -** A RAN Resource Slicing Mechanism for Multiplexing of eMBB and URLLC Services in OFDMA Based 5G Wireless Networks
   Year – 2023
   Type of Publication - 2023 IEEE 11th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)
   Authors – Praveen Kumar Korrai, Eva Lagunas, Shree Krishna Sharma, Symeon Chatzinotas, Ashok Bandi, Bjorn Ottersten

Summary

It is difficult to handle resource management when both eMBB and URLLC traffic are involved. Network Slicing is used to address this challenge by creating independent, logical units/slices within the same resource. This helps in dedicated resource allocation for some special requirements. In this study, the allocation problem of resources is used as an maximization problem focussing on total sum rate to meet the needs of both services, while giving good performance.

Methodology

The approach proposed in the papers consists of slice-aware RAN resource allocation mechanism, which makes use of Adaptive Modulation and Coding (AMC) to maximize the sum rate. This problem has following constraints:

Orthogonality Constraint - Maintains isolations among users, preventing interference.
Latency Constraint - Ensures URLLC traffic meets strict latency requirements.
Minimum Rate Constraint - Guarantees minimum data rates for consistent service quality.
Reliability Constraint - Ensures reliable communication, considering channel conditions.

Certain assumptions and simplifications are made in order to reduce the complexity introduced due to AMC. To reduce computation time, a heuristic method has also been suggested which is of relatively less complexity. This combinatorial mixed-integer non-linear optimization problem is then formulated into linear program.

Result

The simulations provided indicate that optimization-based scheduling algorithm helps in meeting the minimum data-rate requirements of the eMBB users, while heuristic-based algorithm satisfies URLLC latency requirements efficiently.

Conclusion

This work presents a slice-aware resource allocation framework for multiplexing eMBB and URLLC users on shared RAN resources. The framework balances data rate and URLLC's latency by simplifying AMC_related complexity. The results show us that both constraints are met properly, making it effective for real-time applications.

3. **Title -** Joint HAP Access and GEO Satellite Backhaul in 6G: Matching Game-Based Approaches
Year – 2021
Type of Publication - IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 39, NO. 4, APRIL 2021
Authors – Ziye Jia, Min Sheng, Jiandong Li, Di Zhou, Zhu Han

Summary

Geostationary Earth Orbit (GEO) satellites and High-Altitude Platforms (HAPs), which are integrated to form Space-air-ground networks, provide good connectivity for remote areas which lack terrestrial base station coverage. GEO satellites offer global coverage while HAPs provide stable, low energy connections. This work aims to maximize revenue for GEO satellites within a 6G Space-air-ground network framework by addressing resource allocation issues among HAP and user. The authors make use of matching game-based approach called Restricted Three-Sided Matching (R-TMSC).

Methodology

To maximize GEO satellite revenue, the model is rewritten into a mixed integer nonlinear programming (MINLP) model. The study introduces a satellite-oriented R-TMSC algorithm to optimize link in three groups: users, HAPs, and satellites. The methodology includes two core matching components:

Gale-Shapley-Based Matching (Users and HAPs) - A stable matching algorithm, where users and HAPs are paired based on preference lists. This approach iteratively finds the most stable pairings.

Pairwise-Stable Matching (HAPs and Satellites) - To address dynamic satellite-HAP connections due to satellite movement, a pairwise-stable matching is achieved using random path trials, followed by adjustments to maintain stability.

Result

Simulation results show that the significantly enhance connectivity and stability in space-air-ground networks.

R-TMSC Algorithm - Delivers effective revenue maximization for satellites and optimizes the number of users served, achieving stable results across multiple time slots with moderate execution time.

GS+RPPS Algorithm - Random Path to Pairwise-Stable Matching) - Excels in dynamic scenarios requiring rapid updates across consecutive time slots.

Each algorithm's performance adapts well to different network scenarios, providing options for selecting the most suitable approach based on specific network requirements.

Conclusion

This study proposes a three-sided matching approach to optimize GEO satellite and HAP resource allocation within 6G space-air-ground networks. The algorithms proposed simplify the complexity of the problem, efficiently handling dynamic, high-connection variability environments, maximizing GEO satellite revenue and service stability.

4. **Title** - Integrating GEO Satellite and UAV Relaying via Reinforcement Learning for Non-Terrestrial Networks
Year – 2020
Type of Publication – GLOBECOM 2020 - 2020 IEEE Global Communications Conference
Authors – Ju-Hyung Lee, Jihong Park, Mehdi Bennis, Young-Chai Ko

Summary

This study discusses about integration of GEO (low-earth orbit) satellite and HAPs (high-altitude platforms), like UAVs (unmanned aerial vehicles) to build advanced non-terrestrial network for beyond 5G systems. The main focus is to optimize packet forwarding between distant users via GEO satellite and mobile HAP, to maximize E2E (End to End) data rate. The dynamic nature of the system is a huge challenge in real-time optimization which is addressed using the DRL (Deep Reinforcement Learning) approach with action dimension reduction.

Methodology

This optimization has been re-formulated as a DRL task aimed to maximize the E2E data rate for the system. The proposed approach includes:

DQN (Deep Q-Network) Framework: The DQN Framework is utilized to optimize the system association and HAP's location for effective transfer

Action Dimension Reduction: This approach reduces the action space by limiting the association candidates to only those GEO satellites proximate to the source terminal simplifying the decision-making process.

Results

Results tell us th DRL based method massively improves the E2E data rate, achieving up to 5.74 times higher rates compared to baseline method.

Key findings are as follows:
E2E Data Rate Gains- The optimized Src-SAT-HAP-Dst link achieved substantial increases in data rate, showcasing the efficiency of DRL in dynamic environments.

Scalability Potential - The DQN-based model, while successful in a single-link case, indicates scalability for more complex scenarios, such as multi-link networks.

Conclusion

The study gives a solution to solve the problem of maximizing E2E data rate in the mentioned system within a space-ai-ground network by jointly optimizing satellite and HAP positioning using DRL. The author suggests about extending this method to a multi-link scenario with a distributed DRL setup, potentially increasing scalability and efficiency in large-scale, non-terrestrial networks.

5. **Title -** Integrated Access and Backhaul via Satellites
   Year – 2023
   Type of Publication - 2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)
   Authors – Zaid Abdullah, Steven Kisseleff, Eva Lagunas, Vu Nguyen Ha, Frank Zeppenfeldt, Symeon Chatzinotas

Summary

The paper studies the topic of Integrated Access and Backhaul in the non-terrestrial networks (NTNs) using various networks, which helps in the 5G networks. The general Integrated access and Backhaul is used in the terrestrial networks only but this paper analyses the possibility of using this in GEO satellites. This mainly focusses on the efficient solution to optimize the signal and data rates in order for the network to provide the signal to the remote areas where the access of the signal is not reachable.

Methodology

The paper analyses the scope of compatibility of IAB standards of 3GPP for thew new radio networks for 5G which can be used for 5G networks. This enhances the system to be used in the non-terrestrial networks (NTN). This resource analyses the allocation of resources efficiently without any disruptions. The case study is conducted to analyse the primary methods of the backhaul links and the separating access. This can be achieved by using frequency and time division duplexing for optimization.

Results

The final simulation and research tell that the integration of IAB networks for GEO satellites is efficient and they work for the non-terrestrial networks with good efficiency. They also made a conclusion which states that the frequency division duplexing technique is more efficient than the time division duplexing.

Conclusion

The research paper finally concludes that the usage of IAB for the GEO satellites and the integration of them is promising and leads to high data rates and helps in providing the data signal to the rural remote areas which needs basic signal access efficiently. This study also concludes that the Orthogonal allocation of the bandwidth leads to efficient results.

6. **Title -** Intelligent Resource Slicing for eMBB and URLLC Coexistence in 5G and Beyond: A Deep Reinforcement Learning Based Approach
   Year – 2021
   Type of Publication - IEEE Transactions on Wireless Communications (Volume: 20, Issue: 7, July 2021)
   Authors - Madyan Alsenwi, Nguyen H. Tran, Mehdi Bennis,Shashi Raj Pandey, Anupam Kumar Bairagi, Choong Seon Hong

Summary

The researchers faced hurdles in the resource sharing in 5G network among the eMBB and URLLC networks. These two require high data rate and high reliability and low latency rate. They've used a DRL (Deep Reinforcement Learning) approach to tackle this issue. This maximizes the data rate for eMBB and increases the reliability in the URLLC.

Methodology

The paper divides the issue into two parts: one is to allocate the resources for eMBB and the data scheduling for the URLLC In the first part it optimizes the efficiency, data rate, and the allocation frequency band for non-URLLC users. The second method includes the PGACL (Policy Gradient Actor Critic Learning) and also uses a DRL (Deep Reinforcement Algorithm) to manage the URLLC data traffic without the interference of eMBB. This two-part method enhances the system efficiency by managing the data traffic for eMBB and also manages the URLLC efficiently.

Results

The researchers on implementing the algorithms concluded that by using the DRL approach the efficiency has been increased to 90% in managing both URLLC and eMBB by dynamically managing both the networks and managing the disruptions.

Conclusion

In order to conclude, this DRL method which is being used in managing the URLLC and eMBB for optimization of the 5G networks has been successful and help in the better data rate and reliability and also in allocating the bandwidth for the resources involved. This helps in aiding the networks which require dynamic requirement of power and other resources.

7. **Title -** Network Slicing for URLLC and eMBB With Max-Matching Diversity Channel Allocation
Year – 2020
Type of Publication - IEEE Communications Letters (Volume: 24, Issue: 3, March 2020)
Authors - Elco Joao Dos Santos, Jr., Richard Demo Souza, João Luiz Rebelatto, Hirley Alves

Summary

In this research paper the researchers try to optimize the problem in the sharing of resource for the URLLC and eMBB which needs low latency and more reliability and also requires high data speed respectively. The researchers try to solve this problem by using algorithms like Max-Matching Diversity (MMD) algorithm which is used in allocating the bandwidths to the eMBB users. This includes in analysing the Orthogonal Multiple Access (OMA) and also Non-Orthogonal Multiple Access (NOMA).

Methodology

The researchers change the present analysis of NOMA and OMA by applying the MMD algorithm for the combined eMBB and URLLC. Finally hoping to optimize the required data rate of eMBB and the high reliability of the URLLC networks.

Results

The researchers found that this methodology optimizes throughput and the reliable nature of both eMBB and URLLC respectively. The MMD algorithm used in this procedure is more helpful and efficient. For OMA this enhances the diversity by reducing the outages and also aids NOMA by reducing the interference and noise between eMBB and URLLC.

Conclusion

In order to conclude, the researchers found great benefit in using the MMD algorithm which enhances the efficiency, data rate and the reliability The MMD algorithm used in the NOMA and OMA enhances the performance of the eMBB and URLLC to a significant extent which can be used in the 5G communications.

8. **Title -** UAV-GEO Integrated Backbone: A Ubiquitous Data Collection Approach for B5G Internet of Remote Things Networks
Year – 2021
Type of Publication - IEEE Transactions on Wireless Communications (Volume: 20, Issue: 7, July 2021)
Authors - Ting Ma, Haibo Zhou, Bo Qian,Nan Cheng, Xuemin Shen, Xiang Chen, Bo Bai

Summary

The research paper basically combines the functionality of both Unmanned Aerial Vehicles (UAV) and the Low Earth Orbit (GEO) satellite. This is done to support the data acquiring for Internet of Remote Things (IoRT) and the data gathering in Beyond 5G(B5G) networks. The main challenges faced by using both UAVs and GEOs is their high mobility and their less availability in remote areas. The research paper suggests a way to generate a system of two layers where UAVs take packets IoRT equipment and send to GEO satellites.

Methodology

The research paper discusses the division of the data gathering into two parts: IoT to UAV data transmission and UAV to GEO transmission. The paper optimizes the 3-D trajectory of UAVs, transmission power of UAVs, selection of GEO satellites and the allocation of bandwidth for the IoT devices. This uses Block Coordinate Descent techniques (BCD) and Successive Convex Approximation. These methods increase the data transfer and decrease the errors and power consumption.

Results

After the test simulation is done, this research paper suggests that the optimization techniques used, work well in acquiring the data, allocating the bandwidth and minimizing the power consumption. This paper also suggests that if we optimize the UAV path and the GEO satellite selection process. There will be a more scope of more data transfer with less errors and less power consumption.

Conclusion

To conclude, this paper suggests the best way to optimize the integration of UAVs and GEO satellites with data acquisition of more efficiency in B5G in remote areas. This also allocates the bandwidth and coordinate the IoT devices which increases the efficiency of data transfer and reduces the power consumption.

9. **Title -** Coexistence Mechanism Between eMMB and URLLC in 5G Wireless Network
Year – 2021
Type of Publication - IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 69, NO. 3, MARCH 2021
Authors – Anupam Kumar Bairagi, Shirajum Munir, Madyan Alsenwi, Nguyen H. Tran, Sultan S. Alsharani, Mehedi Masud, Zhu Han, Choong Seon Hong

Summary

This paper investigates about combined scheduling of eMMB and URLLC traffic by making use of puncturing. The main aim was maximizing of MEAR of the eMBB users as well as maintaining URLLC services using the PSUM (Penalty Successive Upper Bound Minimization). Evaluation is done based on fariness scores and MEAR of the eMBB users.

Methodology

Author decomposes the multiplexing of eMMB and URLLC traffic based using the puncturing into two subproblems which are eMBB and URLLC scheduling, while making sure main objectives are satisfied. The resources are shared by splitting time zone into main slots and mini slots.

Results

The proposed approach is evaluated using MEAR and fairness score of eMMB UEs. The fairness score decreases with an increase in payload size for lower standard deviation for a mini-slot. fairness score increases with the increasing payload when standard deviation of incoming URLLC traffic in mini-slot is high. The average URLLC latency has no relation with the value of standard deviation of URLLC traffic in mini-slot and payload size.

Conclusion

The Author proposed method for coexistence of URLLC and eMMB traffic for enabling 5G systems. Author has formulated coexisting problem into maximization of the MEAR value of eMMB UEs meanwhile attending the URLLC traffic by decomposing the problem into two problems of eMMB scheduling using PSUM based algorithm and URLLC scheduling using optimal transportation model.

Chapter 3 - Workflow

1. Comprehensive Literature Review

a. A comprehensive review of the URLLC and eMBB communication standards is performed, focusing on existing optimization frameworks in satellite-HAP networks.
b. Apart from communication standards, an examination is done on ML methods/approaches being used to solve the complex optimization problems within these systems.

2. System Model Development

a. A system model is created based around the satellite, HAP and the user nodes.
b. The key objectives and also the key constraints such as channel capacity, bandwidth and power limitations are defined.
c. The key steps are as follows
   - Channel Modelling
   - Resource Allocation
   - Performance Metrics

3. Formulation of optimization problem

Using the model/channel information, we create an optimization problem in accordance to our problem statement

4. Generation of Dataset

a. Using the optimization equation derived from above step, a dataset is generated with realistic parameter configurations.
b. This dataset provides the basis for model training and testing.

5. Machine Learning model development

a. Based on the requirements, a suitable machine learning model is selected for solving the optimization problem.
b. Pre-processing of data is done according to requirement of the model and is trained.

6. Comparison of both approaches

a. Results from directly solving the optimization equation and from ML model are compared. For easier comparison, we can visualize the results.
b. Apart from the results, we also compare them based on various metrics to see which one is more computationally efficient.

Chapter 4 - Work Done

1. Channel Capacity Constraints

Since bandwidth constraints are present

- H2S (HAP to Satellite) Link – Limit the channel capacity to B_2.
- H2U (HAP to Users) Link – Limit the channel capacity to B_1.

The channel capacity for each link can be defined using the Shannon-Hartley theorem

$$C = B \cdot \log_2 \left(1 + \frac{P}{N}\right)$$

where,

- C refers to capacity of channel
- B refers to the bandwidth
- P refers to the Signal Power
- N refers to Noise Power

Thus, for H2S link and H2U link

$$C_{H2S} \leq B_2 \cdot \log_2 \left(1 + \frac{P_{S2H}}{N_{S2H}}\right)$$

$$C_{H2U} \leq B_1 \cdot \log_2 \left(1 + \frac{P_{H2U}}{N_{H2U}}\right)$$

2. Latency and Reliability Constraints

H2S (HAP to Satellite) Link – Since this link follows URLLC requirements, it must adhere to ultra-reliable and low-latency standards. This means

a. **Low Latency Constraint** – The total delay incurred for S2H link must not cross a threshold, say D_{max}.
b. **High Reliability Constraint** – Reliability can be represented as a high probability of successful transmission, say $R_{H2S} \geq 1 - 10^{-9}$.

Therefore, for the H2S link, the latency constraint can be $D_{H2S} \leq D_{max}$ and the reliability constraint can be $P(successful\ transmission\ on\ H2S) \geq 1 - 10^{-9}$.

H2U (HAP to Users) Link – This link follows eMMB requirements, focusing on high data rates. Here, reliability can be relaxed compared to URLLC, latency does not need to be as strict. However, the Sum Rate R_{sum} for users should be maximized. This can be represented as

$$R_{sum} = \sum_{i=1}^{N} C_{H2U_i}$$

Where N refers to number of users connected to the HAP, and each user has a channel capacity C_{H2U} dependent on the available bandwidth B_1 and the SNR (Signal to Noise Ratio).

3. Environmental and Physical Constraints

For the H2S link, we assume an AWGN (Additive White Gaussian Noise) channel model, suitable for the quasi-vacuum environment, which allows for minimal interference and high channel capacity.

H2S Environmental Constraints – Given that the H2S link operates in a quasi-vacuum, the communication channel is primarily affected by Gaussian noise. Hence, we take no account of capacity restrictions for the H2S link and assume it can support a high data rate up to its bandwidth limit B_2. Therefore, the effective channel model for H2S is

$$C_{S2H} \approx B_2 \cdot \log_2\left(1 + \frac{P_{S2H}}{N_{S2H}}\right)$$

For the H2U link, Free Space Path Loss Model is used.

Free Space Path Loss

$$L_{f\,sl} = 20\log_{10}(d) + 20\log_{10}(f) + 20\log_{10}\left(\frac{4\pi}{c}\right)$$

where,

- L is Path Loss
- d is Distance Between HAP and UE
- f is Frequency of Transmitted Signal
- c is Speed of Light in Vacuum

The total path loss L_{total} for the H2U link would then be

$$L_{total} = L_{fsl} + L_{rain}$$

The SNR for each user is affected by this loss, which impacts C_{H2U_i}

$$C_{H2U_i} = B_1 \cdot \log_2\left(1 + \frac{P_{H2U}}{N_{H2U} + L_{total}}\right)$$

System Model Design

1. System Description

- This system consists of ground users, a high-altitude platform (HAP), and a GEO satellite.
- The users are present on the ground and use emBB (Enhanced Mobile Broadband), which provides high-speed data and network capacity. eMBB is designed to improve the user experience for mobile broadband applications. The users expect a minimum data-rate level so as to remain functional.
- The HAP also sends information to users directly using eMBB link. The channel used to send data to users is in S-band and is divided into orthogonal channels. The HAP allocates power to users and optimizes its location such that sum-rate is maximized.
- The GEO satellite sends and receives control information from HAP. URLLC is used to ensure reliability and to meet latency requirements to send control-based information.

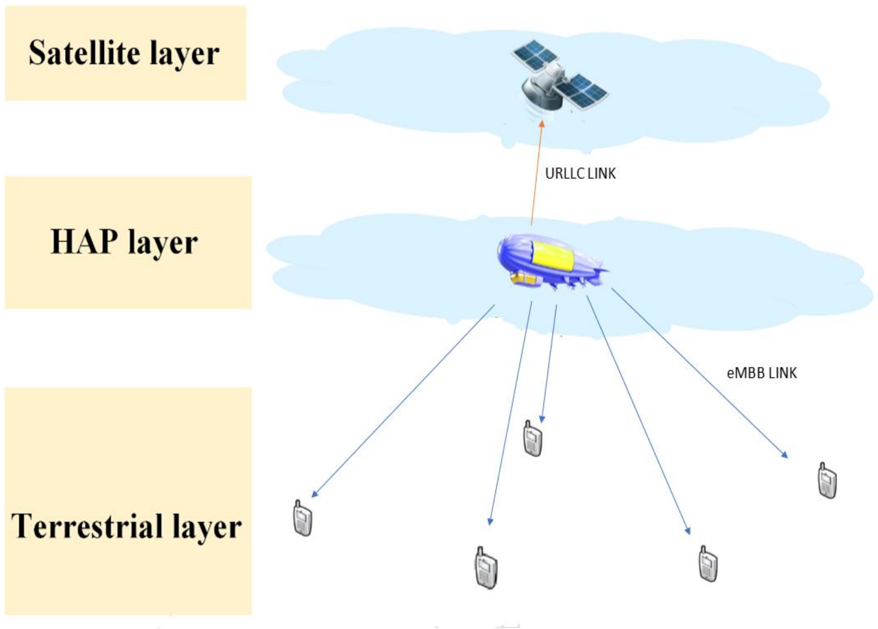

Fig. 1. System Model

The HAP in the system deals with HAP-satellite link and the HAP-user link. HAP-satellite link is a URLLC link whereas the HAP-user link is an eMBB link. In this scenario, we assume K number of eMBB users and N URLLC users (our case N = 1). Let B correspond to the total number of Resource Blocks (RB) in the system. Each Resource Block that is considered occupies 12 sub-carrier in frequency.

Typically, eMBB transmissions take place for a long period of time in contrast to URLLC, which are infrequent and happen in a short period of time. To make sure they are able to send required message in short period of time, they are allowed to span multiple frequency resources. To make efficient use of our resources, we make use of the resource slicing between the eMBB and URLLC links. Thus, when URLLC arrive to be sent while eMBB transmission is happening, we use puncturing mechanism in order to make sure URLLC packets are able to meet their constraints. Here, URLLC traffic is observed in mini-slot whereas eMBB is scheduled in the main time slot. Let ‘M’ represent the total mini-slots in a main slot t.

1. a) Channel Modeling for HAP-User link:

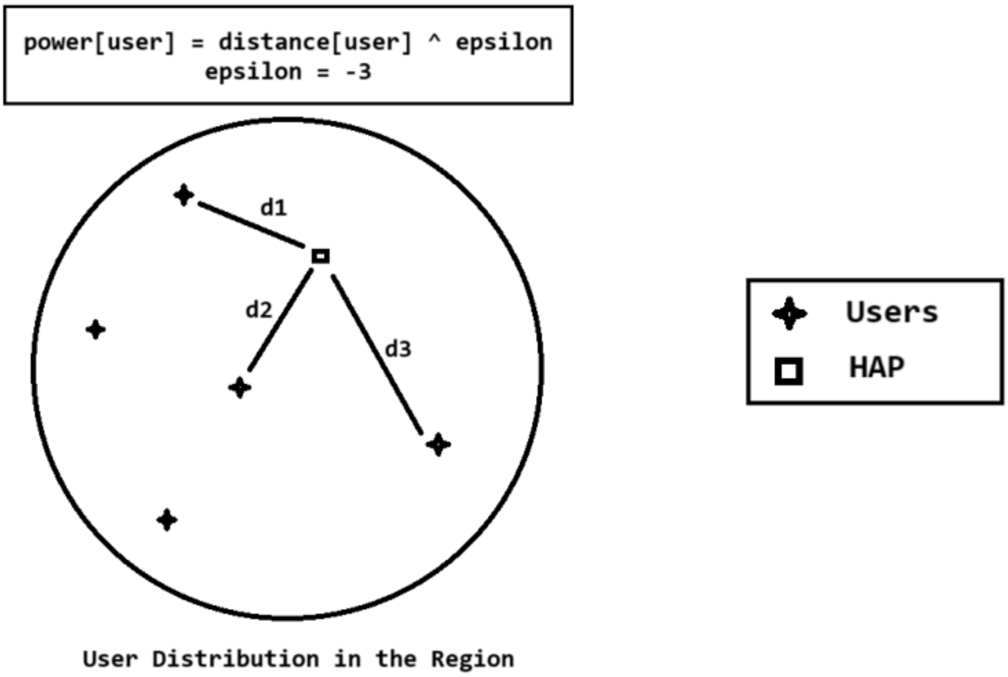

Fig. 2. Channel Coefficients Calculations

To model the channel between HAP and users, we make use of large-scale fading (or distance-dependent path loss modelling). If h_{kb} represents the channel coefficient between HAP and user k for a Resource Block b, it can be defined as follows:

$$|h_{kb}|^2 = (d_{kb})^{-\epsilon} \tag{1}$$

where ϵ refers to the path loss exponent (generally in the range [2,4]) and d_{kb} is the distance between HAP and user k.

Parameter	Value
Total bandwidth (W)	20 MHz
Number of Sub-carriers	12
Sub-carrier spacing	15Khz
Noise PSD (N_0)	-174 dBm/Hz
Satellite antenna gain (GGEO)	30 dBi
HAP antenna gain (GHAP)	15 dBi
Carrier frequency (fc)	3 GHz (FR1)
GEO satellite altitude	35,786 km
HAP altitude	20 km
URLLC traffic model	Poission process with rate λu

Table. 1. System Model Parameters

1a) Channel Modeling for HAP-satellite link:

The channel modeling of HAP-satellite link is done by assuming Line of Sight (LoS) between the Satellite and HAP. Free Space Path Loss (PL) model is used for channel modelling.

$$\beta_i = \left(\frac{G_{sat} G_{HAP}}{PL_i}\right) \psi(\theta_{HAP}) \tag{2}$$

where G_{sat} and G_{HAP} refer to satellite and HAP antenna gains respectively. Here PL_i is defined as follows:

$$PL_i = (4\pi f_c d_{S2H}/c)^2 \tag{3}$$

refers to the Path Loss in free space between HAP and the satellite with f_c being carrier frequency and d_{S2H} being the distance between satellite and HAP and c is the speed of light.

$\psi(\theta_{HAP})$ is the normalized antenna gain pattern at satellite w.r.t HAP, such that $\psi(\theta_{HAP})$ is one if (θ_{HAP}) is zero. Otherwise, it is defined as follows:

$$\left|\frac{J_1(k.a.sin(\theta_{HAP}))}{k.a.sin(\theta_{HAP})}\right| \tag{4}$$

where (θ_{HAP}) refers to boresight angle from satellite's node to HAP. The parameter k is defined as $k = \frac{2\pi f_c}{c}$. The parameter a in above equation refers to radius of satellite's circular antenna apreture. $J_1(\quad)$ is the Bessel function of the first kind and first order.

2. eMBB data Rate Calculation:

The Modified eMBB data rate considering puncturing can be written as follows:

$$r_{kb}^{e}(t) = f_b\left(1 - \frac{z_{kb}(t)}{M}\right)\log_2\left(1 + \frac{p_{kb}(t)h_{kb}(t)}{\sigma^2}\right) \tag{5}$$

where,

f_b refers to the bandwidth of RB,

$h_{kb}(t)$ is channel coefficient

$p_{kb}(t)$ refers to the power allocated by HAP to transmit message to user k over RB b in a time slot t.

Using this, total data rate from one single user is::

$$r_k^e(t) = \sum_{b\in\mathcal{B}} x_{kb}(t) r_{kb}^e(t) \tag{6}$$

Where $x_{kb}(t)$ tells us which resource block is allocated to which user at a time slot t. $x_{kb}(t)$ is one if user k has access to RB b at a time slot t. In all other cases, it is zero.

3. URLLC data rate

Since the URLLC packets size is very small, we cannot calculate the data rate and transmission error using Shannon's capacity theorem. Hence, , we make use of the Finite block-length channel coding regime. The equation for data rate is:

$$r_n^u(t) = \sum_{k\in\mathcal{K}}\sum_{b\in\mathcal{B}} f_b \frac{x_k^b(t) z_{kb}(t)}{M \times N} \log\left(1 + \frac{p_{nb}^u(t) h_{nb}^u(t)}{\sigma^2}\right) - \sqrt{\frac{D_{nb}^u(t)}{c_{nb}^u(t)}} Q^{-1}(\vartheta) \tag{7}$$

where,

$Q^{-1}()$ refers to inverse of Gaussian Q-function, refers to transmission error probability, and $D_{nb}^u(t)$ represents channel dispersion and is defined as follows:

$$D_{nb}^u(t) = 1 - \left(\frac{1}{1+\frac{p_n^u(t) h_n^u(t)}{\sigma^2}}\right)^2. \tag{8}$$

4. Optimization Problem Formulation

At the start of each time slot, all the RBs are allocated to eMBB users. During the time slot, if URLLC traffic request arrives, some of the mini-slots of RBs are punctured and are given to deal with the URLLC traffic. While puncturing, it is important to make sure that the reliability of eMBB transmission is not greatly compromised. So, while going about resource allocation the following things have to be kept in mind:

1. Maximization of eMBB throughput
2. Making sure eMBB transmission is reliable
3. Satisfying URLLC constraints

In order to deal with eMBB reliability, we try to make sure variance between data rates is as less as possible. A function is defined, which takes into account the eMBB data rate sum and variance, as follows

$$\mathcal{F}(x, p, z) = \sum_{k=1}^{K} E_h\left[\frac{1}{T}\sum_{t=0}^{T} r_k^e(t)\right] - \beta\, \mathrm{Var}_h[r_k^e(t)] \tag{9}$$

In the above equation, beta refers to variance weight, Var is the variance.

In order to satisfy URLLC constraints, we have to make sure that the outage probability is less than a threshold. If we consider $L_m(t)$ as URLLC traffic at a mini-slot m, then the total URLLC packets that arrive during a given time slot is defined as $L(t) = \sum_{m\in\mathcal{M}} L_m(t)$

Using this, URLLC reliability is defined as follows:

$$\Pr\left[\sum_{n\in\mathcal{N}} r_n^u(t) \leq \zeta L(t)\right] \leq \Theta^{max}, \tag{10}$$

Where zeta is URLLC packet size.

The eMBB/URLLC joint resource allocation problem is defined as follows:

$$\text{maximize}_{x,p,z} \sum_{k=1}^{K} E_{\boldsymbol{h}} \left[\frac{1}{T} \sum_{t=0}^{T} r_k^e(t) \right] - \beta \, \mathrm{Var}_{\mathrm{h}}[r_k^e(t)] \tag{11a}$$

$$\text{subject to} \Pr \left[\sum_{n=1}^{N} r_n^u(t) \le \zeta L(t) \right] \le \Theta^{max}, \tag{11b}$$

$$\sum_{k=1}^{K} \sum_{b=1}^{B} p_{kb}(t) \le P_{total}, \tag{11c}$$

$$p_{kb}(t) \ge 0, \quad \forall\, k \in \mathcal{K},\, b \in \mathcal{B} \tag{11d}$$

$$\sum_{k=1}^{K} x_{kb}(t) \le 1, \quad \forall\, b \in \mathcal{B}, \tag{11e}$$

$$p_{kb}(t) \ge 0, \quad \forall\, k \in \mathcal{K},\, b \in \mathcal{B} \tag{11f}$$

5. Modification of Optimization Problem

The optimization problem defined in (8a) is a mixed-integer nonlinear programming (MINLP) type problem which is difficult to solve. So, in order to simplify this problem, we divide this main problem in few subproblems which are convex in nature.

Firstly, we convert the simplify the main optimization function (utility function) into a simple exponential function that deals with both mean and variance:

$$\mathcal{G}(x,p,z) = \frac{1}{\mu} \log E_h \left[\exp(\mu \textstyle\sum_{k=1}^{K} r_k^e(t))\right] \tag{12}$$

Here the parameter μ is used to manage the risk-sensitivity. As the value of μ becomes more negative, the function becomes more risk-averse. The taylor series expansion of above equation around μ = 0 can be written as :

$$\mathcal{G}(x,p,z) = E_h \left[\sum_{k=1}^{K} r_k^e(t) \right] + \frac{\mu}{2} \mathrm{Var} \left[\sum_{k=1}^{K} r_k^e(t) \right] + \mathcal{O}(\mu^2) \tag{13}$$

Using this change, we can redefine the allocation problem as follows:

$$\boldsymbol{P}\text{: maximize}_{x,p,z} \frac{1}{\mu} \log E_h \left[\exp\left(\mu \sum_{k=1}^{K} r_k^e(t) \right) \right] \tag{14a}$$

$$\text{subject to} \Pr \left[\sum_{n=1}^{N} r_n^u(t) \le \zeta L(t) \right] \le \Theta^{max}, \tag{14b}$$

$$\sum_{k=1}^{K}\sum_{b=1}^{B} p_{kb}(t) \leq P_{total}, \tag{14c}$$

$$p_{kb}(t) \geq 0, \quad \forall\, k \in \mathcal{K},\; b \in \mathcal{B} \tag{14d}$$

$$\sum_{k=1}^{K} x_{kb}(t) \leq 1, \quad \forall\, b \in \mathcal{B}, \tag{14e}$$

$$p_{kb}(t) \geq 0, \quad \forall\, k \in \mathcal{K},\; b \in \mathcal{B} \tag{14f}$$

To solve this redefined problem, we make use of an algorithm called as DDRA (Decomposition and Relaxation based Resource Allocation). In this algorithm, first the main problem P is split into 3 smaller problems. The first problem deals with eMBB RB allocation, The second problem deals with eMBB power allocation. The third problem deals with the URLLC scheduling. These 3 problems are iteratively solved until convergence is reached. Finally, we make some binary conversion to get the final required answer. All the problems defined are convex optimization problems.

A. eMBB RB allocation:

For a fixed eMBB power allocation "p" and URLLC scheduling "z", the problem can be written as follows:

$$\boldsymbol{P1}: \quad \text{maximize}_{\boldsymbol{x},\boldsymbol{p},\boldsymbol{z}} \frac{1}{\mu} \log E_h \left[\exp\left(\mu \sum_{k=1}^{K} r_k^e(t) \right) \right] \tag{15a}$$

$$\text{subject to} \sum_{k=1}^{K} x_{kb}(t) \leq 1, \quad \forall b \in \mathcal{B}, \tag{15b}$$

$$x_{kb}(t) \in \{0,1\}, \quad \forall k \in \mathcal{K} \text{ and } b \in \mathcal{B}. \tag{15c}$$

This problem is a MINLP problem which can be relaxed by letting values of x take fractional values while running the algorithm. Once convergence happens, we can easily round-off to get desired answer. This can be written as follows:

$$\widetilde{\boldsymbol{P1}}: \quad \text{maximize}_{\tilde{x}} \frac{1}{\mu} \log E_h \left[\exp\left(\mu \sum_{k=1}^{K} r_k^e(t) \right) \right] \tag{16a}$$

$$\text{subject to} \sum_{k=1}^{K} \widetilde{x_{kb}}(t) \leq 1, \quad \forall b \in \mathcal{B}, \tag{16b}$$

$$0 \leq \widetilde{x_{kb}}(t) \leq 1, \quad \forall k \in \mathcal{K},\; b \in \mathcal{B}. \tag{16c}$$

B. **eMBB Power Allocation Problem:**

For a pre-defined eMBB RB allocation "x" and URLLC scheduling "z", we can find the optimal power distribution as follows:

$$\boldsymbol{P2}: \quad \text{maximize}_{\boldsymbol{p}} \frac{1}{\mu} \log E_{\boldsymbol{h}} \left[\exp\left(\mu \sum_{k=1}^{K} r_k^e(t) \right) \right] \tag{17a}$$

$$\text{subject to} \sum_{k=1}^{K} \sum_{b=1}^{B} p_{kb}(t) \leq P_{max}. \tag{17b}$$

C. **URLLC Scheduling Problem:**

Given the eMBB RB allocation "x" and eMBB power allocation "p", the ULRLC scheduling is defined as follows:

$$\boldsymbol{P3}: \quad \text{maximize}_{\boldsymbol{z}} \frac{1}{\mu} \log E_{\boldsymbol{h}} \left[\exp\left(\mu \sum_{k=1}^{K} r_k^e(t) \right) \right] \tag{18a}$$

$$\text{subject to, } \Pr\left(\sum_{n=1}^{N} r_n^u(t) \leq \zeta L(t) \right) \leq \Theta^*, \tag{18b}$$

$$z_{kb}(t) \in \{0,1, \dots, M\}, \quad \forall\, k \in \mathcal{K},\ b \in \mathcal{B}. \tag{18c}$$

The problem defined above is a combinatorial optimization problem which happens to be a problem where getting convergence is complex. In order to make this problem simpler, we replace the interger matrix "z" with a weighting matrix w where each element lies in the range [0,1]. We can approximate eMBB data rate is:

$$r_{kb}^e(t) = f_b\big(1 - w_{kb}(t)\big) \log_2\left(1 + \frac{p_{kb}(t) h_{kb}(t)}{\sigma^2}\right) \tag{19}$$

Then we can calculate URLLC data rate as follows:

$$r_n^u(t) = \sum_{k \in \mathcal{K}} \sum_{b \in \mathcal{B}} f_b \frac{x_k^b(t) w_{kb}(t)}{M \times N} \log\left(1 + \frac{p_{nb}^u(t) h_{nb}^u(t)}{\sigma^2}\right) - \sqrt{\frac{D_{nb}^u(t)}{c_{nb}^u(t)}}\, Q^{-1}(\vartheta), \tag{20}$$

We make use of Markov's Inequality in order to represent chance constraint as linear constraint:

$$\Pr\left[\sum_{n \in \mathcal{N}} r_n^u(t) \leq \zeta L(t) \right] \leq \frac{\zeta E[L]}{\sum_{n \in \mathcal{N}} r_n^u(t)}. \tag{21}$$

The URLLC problem is reformulated as follows:

$$\widetilde{\boldsymbol{P}}\mathbf{3}: \quad \text{maximize}_{\boldsymbol{w}} \frac{1}{\mu} \log E_{\boldsymbol{h}} \left[\exp \left(\mu \sum_{k=1}^{K} \widetilde{r_k^e}(t) \right) \right] \tag{22a}$$

$$\text{Subject to, } \sum_{n \in \mathcal{N}} \widetilde{r_n^u}(\boldsymbol{w}) \geq \frac{\zeta E[L]}{\Theta^*}. \tag{22b}$$

$$0 \leq w_{kb} \leq 1, \forall k \in \mathcal{K}, b \in \mathcal{B} \tag{22c}$$

6. Using DRL for URLLC Scheduling

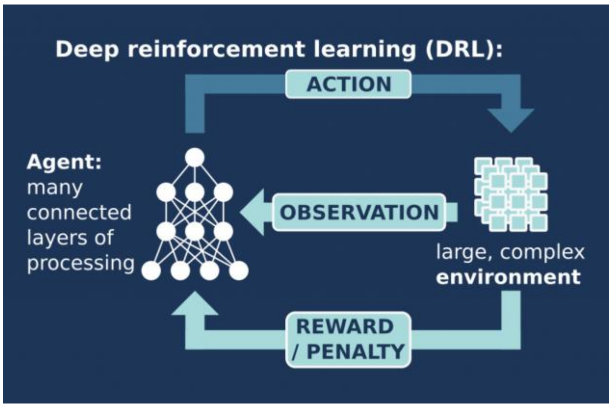

Fig. 3. DRL Architecture

The URLLC scheduling solution that is given by DRRA algorithm in some cases can violate URLLC reliability constraint in worst cases due to the relaxation applied. The real time URLLC traffic is unpredictable which makes it essential to have an algorithm that takes this in to account. Thus, a DRL-based algorithm is used to tackle this problem. We make use of results from traditional approach that are obtained in the DRRA algorithm initially to improve convergence time. By combining both the DRRA algorithm and DRL-based algorithm, we achieve a solution that is highly reliable and efficient.

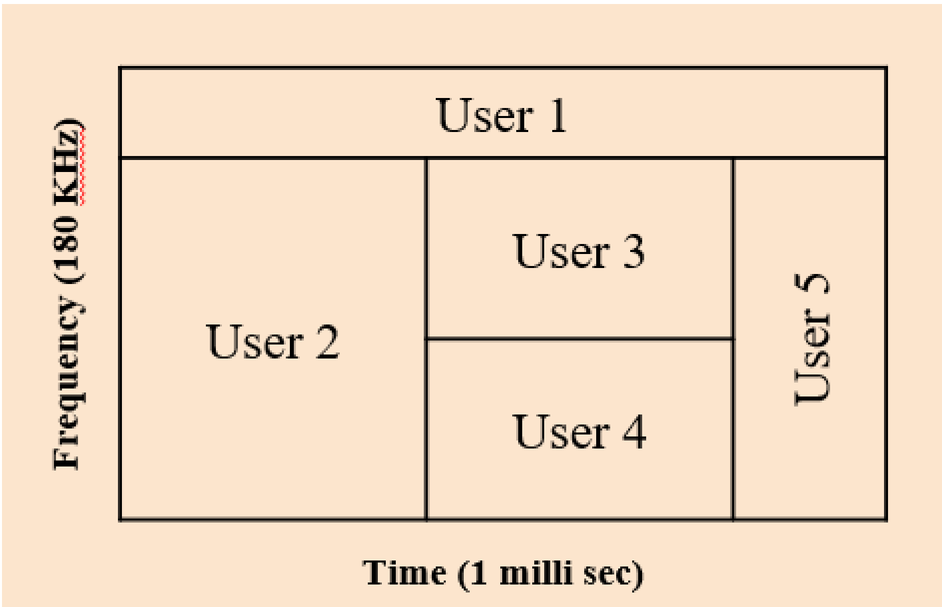

Fig. 4. Resource Block is the Environment for DRL Agent

A reinforcement learning model consists of following components: action space, state space and reward. The algorithm takes some action for each state and receives a reward.

1. State Space: The state space is used to define state of each eMBB and URLLC user. For a time slot, we can define state space s(t) = {x(t), p(t), h^e(t), L(t), h^u(t)}. To reduce the size of state space, we define $\widehat{r_k^e}(t)$ as data rate of eMBB without puncturing as follows:

$$\widehat{r_k^e}(t) = \sum_{b \in \mathcal{B}} x_{kb}(t) f_b \log_2\left(1 + \frac{p_{kb}(t) h_{kb}(t)}{\sigma^2}\right) \tag{23}$$

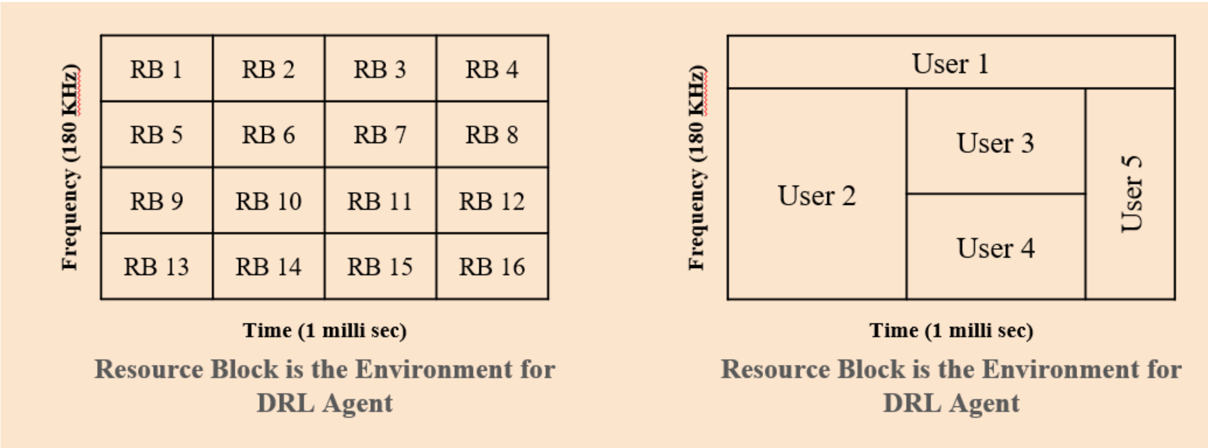

Fig. 5. Resource Block Allocation

The state space is collapsed to s(t) = {$\hat{r}^e$(t), h^u(t), L(t)}.

2. Action Space: Action space in our scenario refers to the number of punctured mini-slots of each RB,

$$a(t) = \{z_{kb}, \forall\ b \in B,\ k \in K\},$$

Which is a B x M puncturing matrix.

3. Reward: The reward function is calculated by considering both eMBB and URLLC requirements as follows:

$$R\big(a(t), s(t)\big) = s(t) + \phi(t) E\left[\sum_{n=1}^{N} r_n^u(t) - \zeta L(t)\right] \tag{24}$$

With

$$\phi(t+1) = \max\{\phi(t) + \Theta(t) - \Theta^{max}, 0\} \tag{25}$$

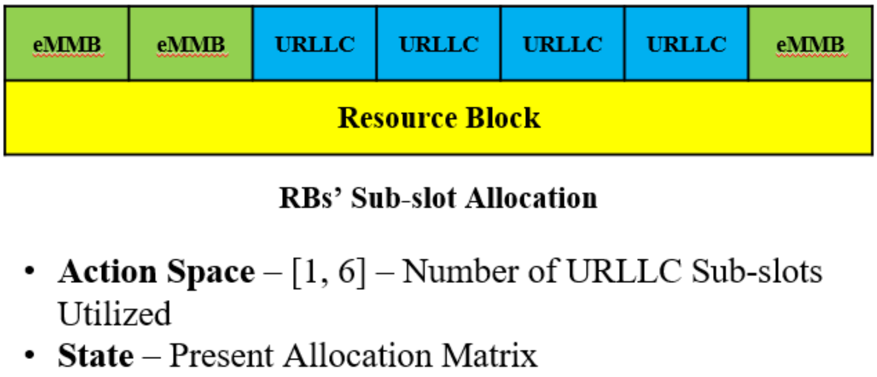

Fig. 6. Sub-slot Allocation

Using above information, we create an environment which defines all the required variables like z and helps in taking steps to reach our required target. In order to understand whether the step we've taken is in the right or in the wrong direction, we make use of the reward function.

The aim of DRL algorithm is to select a policy π(a,s)= $\{\pi_m^b, \forall\ b \in B, m \in M\}$, where π_m^b represents the probability of puncturing m mini-slots from resource block b, based on the network state s(t). The agent observes the current network state s(t) and decides how many mini-slots to puncture per mini-block. After action is taken, reward is calculated using the equation mentioned and updates the network state with new information. The agent then adjusts the policy for the decision cycle.

The cumulative discounted reward with a given policy π is as follows:

$$Q^{\pi}(s,a) = E\left[\sum_{t=0}^{\infty} \gamma(t) R(s_t, a_t) \mid s_0 = s, \pi\right] \tag{26}$$

This is calculated using Bellman Equation as follows:

$$Q^{\pi}(s,a) = E\left[R\big(s(t), a(t)\big) + Q^{\pi}\big(s(t+1), a(t+1)\big)\right] \tag{27}$$

The function J(π) can be considered as network objective reward value which is as follows:

$$J(\pi) = E[Q^{\pi}(s,a)] = \int_{\mathcal{S}} \int_{\mathcal{A}} \pi(s,a) Q^{\pi}(s,a)\ da\, ds. \tag{28}$$

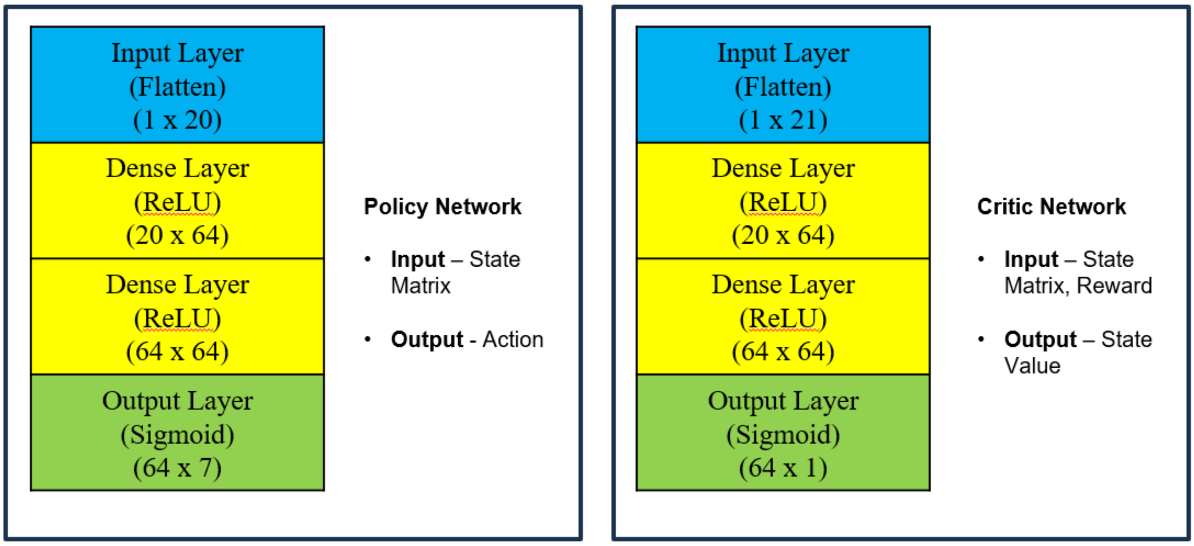

Fig. 7. Actor and Critic Network Architecture

We need to maximize J(π) by choosing a policy. To achieve this, techniques like Q-learning and policy gradient methods can be used. However, Q-learning often struggles to find the best policy in real-time due to its slow learning rate. In contrast, policy gradient methods offer faster convergence. For reducing computation time and improving convergence time, we make use of PGACL algorithm (Policy Gradient-based Actor Critic Learning).

PGACL is technique of DRL Algorithms which works better on complex environments as it used a neural network for policy making and decision making. The Neural Network in the PGACL allows it to understand complex environments and its complex reward function making it simple to tune the model.

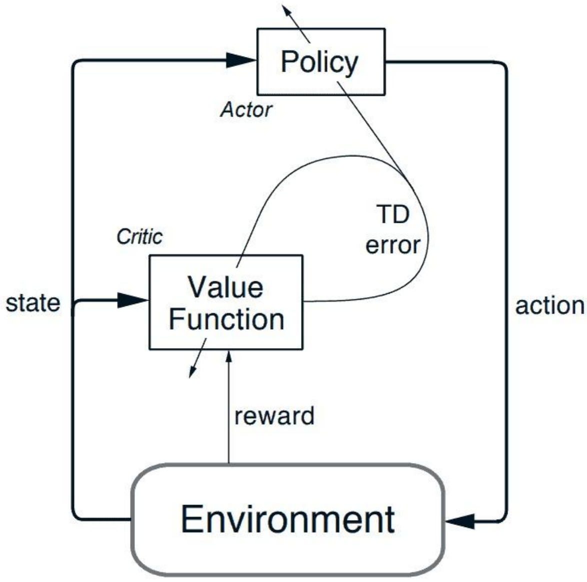

Fig. 8. PGACL Architecture

In PGACL, as seen in the above figure. A random action is performed in the environment changing the state of the environment and based on the change in state and performed action a reward is given to the agent using the reward function after which policy network makes a action based on the previous interaction with the environment. The goal of the PGACL is to minimize the Temporal difference between the Policy Network and the Critic Network.

Chapter 5 – Results

DRL Results

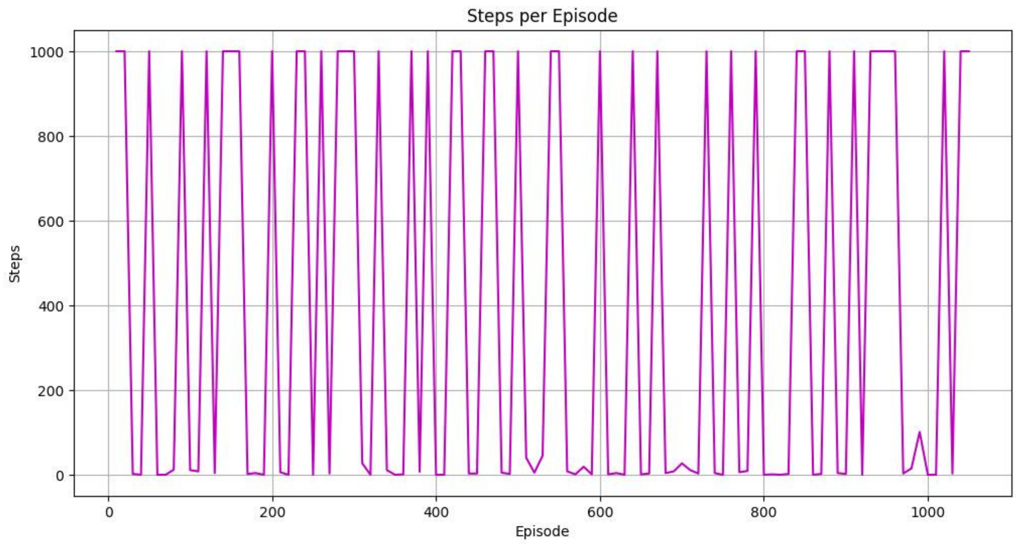

Fig. 9. Number of Steps taken Per Episode

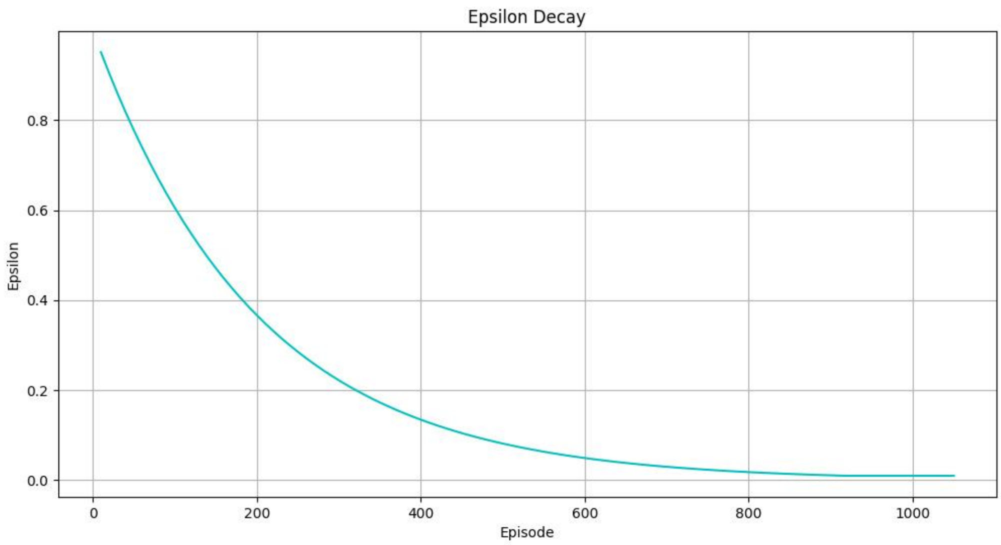

Fig. 10. Epsilon Decay per Episode

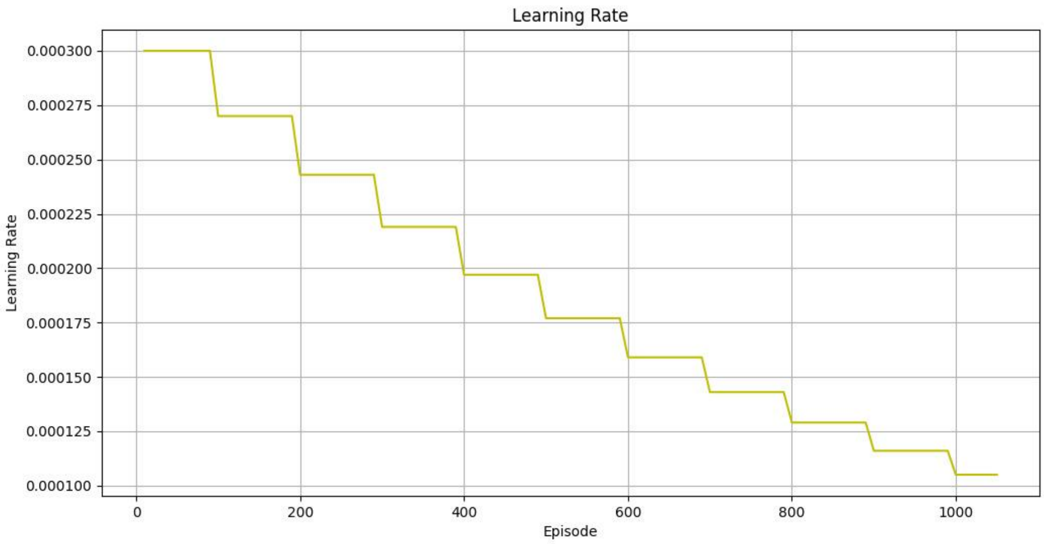

Fig. 11. Learning Rate per Episode

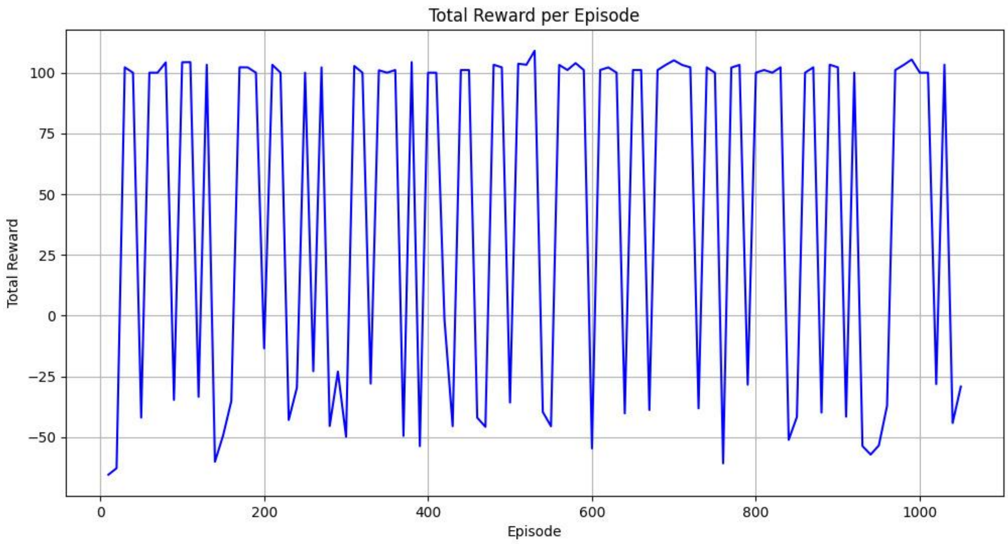

Fig. 12. Total Reward per Episode

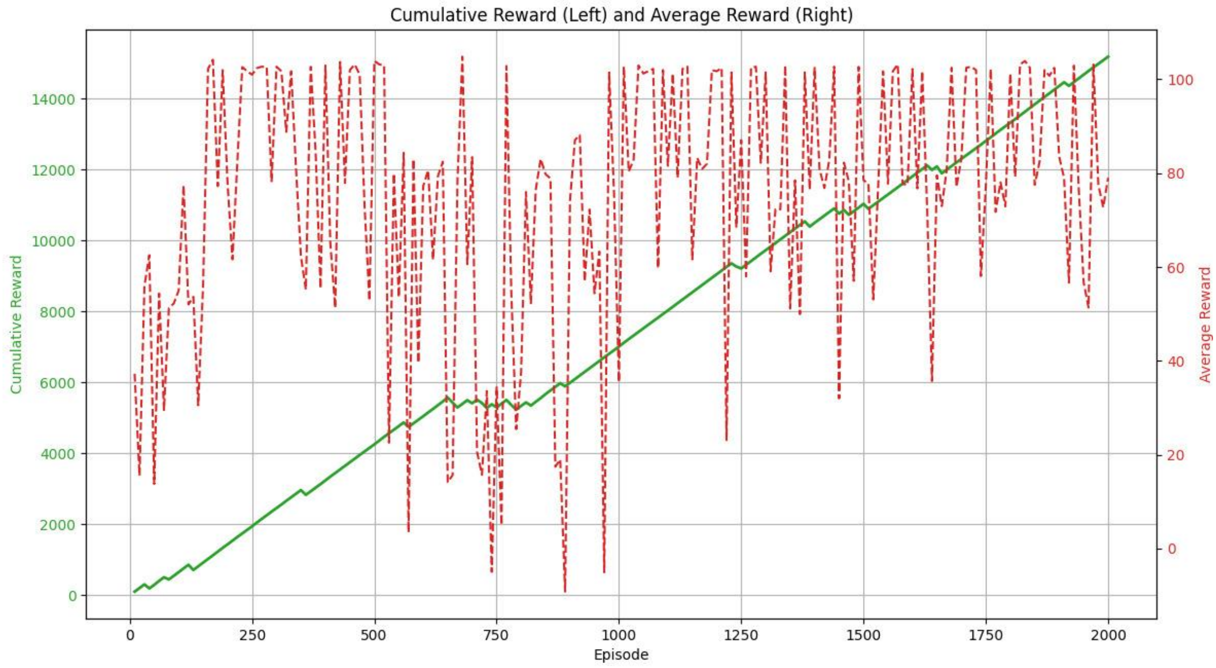

Fig. 13. Cumulative Reward and Average Reward Per Episode

Traditional Approach Results

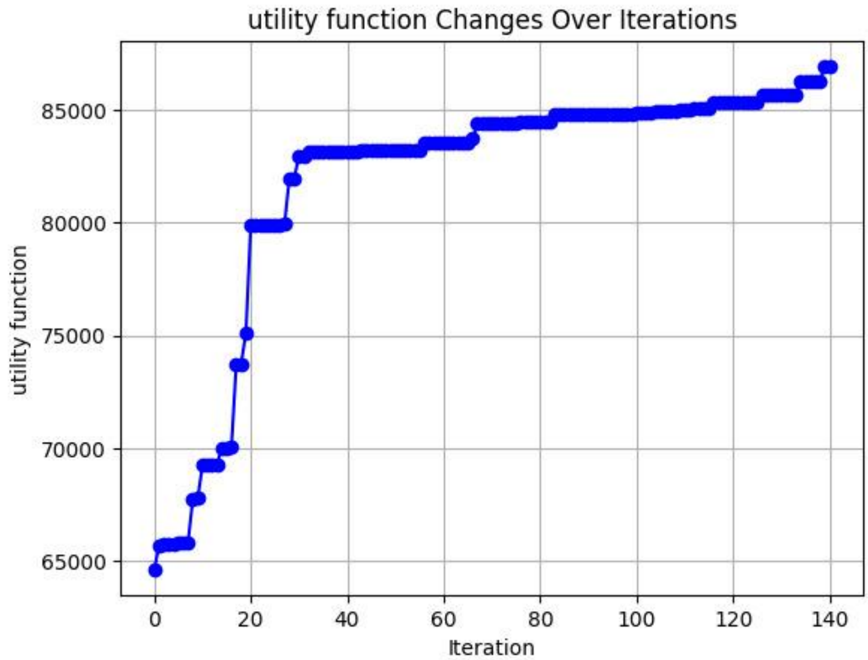

Fig. 14. Utility Function for Each Iteration

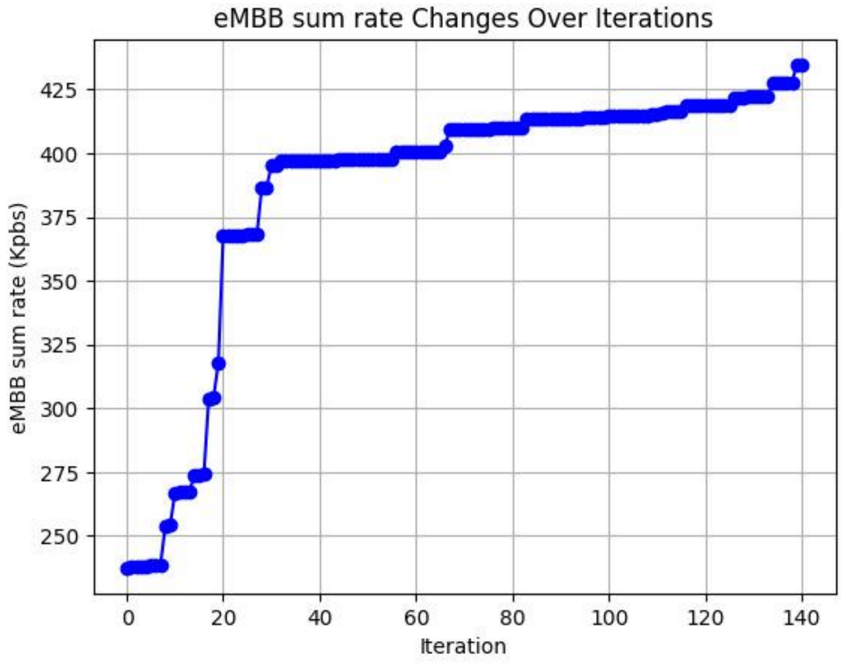

Fig. 15. Total eMBB Sum-Rate for Each Iteration

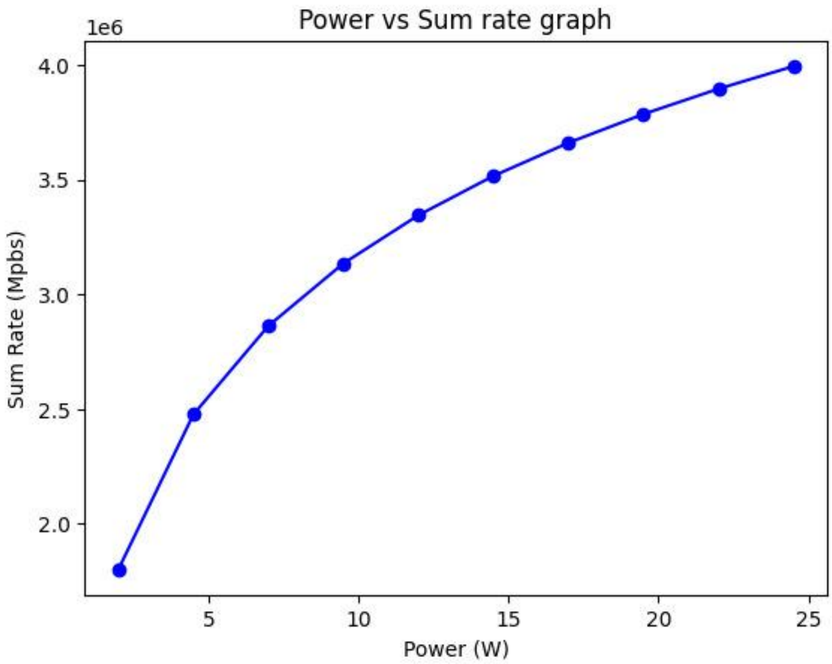

Fig. 16. Power vs eMBB Sum-Rate

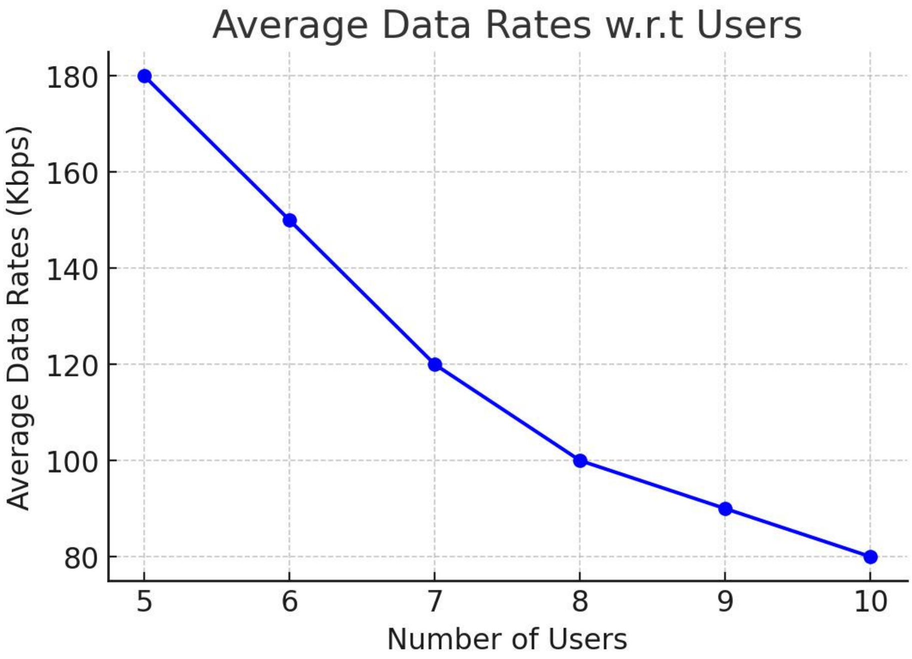

Fig. 17. Average Data-Rate vs Number of Users

Chapter 6 - References

[1] W. Lee, M. Kim and D. -H. Cho, "Deep Power Control: Transmit Power Control Scheme Based on Convolutional Neural Network," in *IEEE Communications Letters*, vol. 22, no. 6, pp. 1276-1279, June 2018

[2] M. Alsenwi, N. H. Tran, M. Bennis, A. Kumar Bairagi and C. S. Hong, "eMBB-URLLC Resource Slicing: A Risk-Sensitive Approach," in *IEEE Communications Letters*, vol. 23, no. 4, pp. 740-743, April 2019

[3] P. Korrai, E. Lagunas, S. K. Sharma, S. Chatzinotas, A. Bandi and B. Ottersten, "A RAN Resource Slicing Mechanism for Multiplexing of eMBB and URLLC Services in OFDMA Based 5G Wireless Networks," in *IEEE Access*, vol. 8, pp. 45674-45688, 2020

[4] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi and C. S. Hong, "Intelligent Resource Slicing for eMBB and URLLC Coexistence in 5G and Beyond: A Deep Reinforcement Learning Based Approach," in *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4585-4600, July 2021

[5] E. J. dos Santos, R. D. Souza, J. L. Rebelatto and H. Alves, "Network Slicing for URLLC and eMBB With Max-Matching Diversity Channel Allocation," in *IEEE Communications Letters*, vol. 24, no. 3, pp. 658-661, March 2020

[6] A. K. Bairagi *et al*., "Coexistence Mechanism Between eMBB and uRLLC in 5G Wireless Networks," in *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1736-1749, March 2021

[7] P. Yang, X. Xi, K. Guo, T. Q. S. Quek, J. Chen and X. Cao, "Proactive UAV Network Slicing for URLLC and Mobile Broadband Service Multiplexing," in *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 10, pp. 3225-3244, Oct. 2021

[8] Z. Jia, M. Sheng, J. Li, D. Zhou and Z. Han, "Joint HAP Access and LEO Satellite Backhaul in 6G: Matching Game-Based Approaches," in *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1147-1159, April 2021

[9] J. -H. Lee, J. Park, M. Bennis and Y. -C. Ko, "Integrating LEO Satellite and UAV Relaying via Reinforcement Learning for Non-Terrestrial Networks," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, Taipei, Taiwan, 2020, pp. 1-6

[10] T. Ma *et al*., "UAV-LEO Integrated Backbone: A Ubiquitous Data Collection Approach for B5G Internet of Remote Things Networks," in *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 11, pp. 3491-3505, Nov. 2021