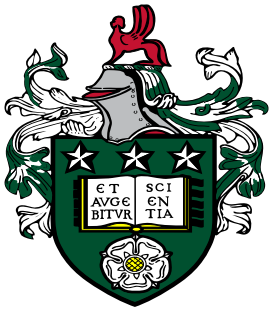
Optimisation of Intelligent Reflecting Surfaces in 5G & Beyond Access and Backhaul Networks

Noman Aftab



Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

School of Electronic and Electrical Engineering

University of Leeds, UK

(September, 2021)

Supervisors

Dr Syed Ali Raza Zaidi

Dr Ahmed Lawey

Dr Des McLernon

Date of submission

30.09.2021

I dedicate this work to my parents, my wife, and my children

Acknowledgements

First and foremost, I would like to thank Almighty Allah for all His blessings. For whatever I achieve in life is due to His grace.

Secondly, I am most grateful to my supervisor, Dr Ali Raza Zaidi, for his continuous guidance during my study at the University of Leeds. It is due to his immense support that I have been successful in completing my thesis. Furthermore, his insights and valuable advice paved the road toward accomplishing the goals of this work. Finally, I would like to thank Dr Ahmed Lawey for giving me many helpful suggestions and assistance for my work. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research.

I would also like to acknowledge my thesis advisor, Dr Des McLernon, for his support and guidance in refining and improving my work. During my studies in Leeds, I have been fortunate to work with various members of the IRASS group, with special mention to Naveed Nawaz, for providing me with valuable comments; Sharjeel Afridi for creating a friendly and cooperative atmosphere at work.

I want to thank the management of the University of Engineering & Technology, Lahore, for providing me with the opportunity to pursue my higher education at the University of Leeds.

Finally, my deepest thanks go to my family and friends. Thanks to my parents for their unconditional love and support all these years, and my wife for the countless efforts, and my kids Areesh and Eshal for their love, support, patience, and understanding during my course PhD.

My sincere appreciation goes to my family who always encouraged and motivated me to reach this point in my life; to my parents who taught me the alphabet; to my

wife for her companionship, support, care, endless patience, constant reassurance and steadfast belief; to my brothers and sisters for their belief and support. Finally, I would like to thank all my colleagues and friends who made my time at Leeds university such a pleasure. Thank you so much!

Copyright

The right of Noman Aftab to be identified as author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

Declaration

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that the work submitted is his own and that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in chapter 4 is to be submitted in IEEE Transactions on Green Communications and Networking for publication as follows

- Optimisation of IRS in 5G and Beyond Backhaul Networks

The work in chapter 5 is to be submitted in in journal as follows.

- Joint Power and Latency Minimization of Mobile Edge Computing in Future Access Networks Assisted by Intelligent Reflecting Surfaces

Contents

List of Figures

List of Tables

Nomenclature

HetNet	Heterogeneous Network
AIMMS	Advanced Multidimensional Interactive Modeling System
AWGN	Additive White Gaussian Noise
BS	Base station
BS	base station
CCC	Centralized Cloud computing
E2ED	end-to-end delay
EB	Exabyte
ETSI	European Telecommunications Standards Institute
GA	Genetic Algorithm
IRS	Intelligent Reflecting Surface
KPI	key performance indicators
MCC	Mobile Cloud computing
MEC	Mobile Edge Computing
MEMS	Micro Electrical Mechanical Systems
MILP	Mixed Integer Linear Programming

MM	Majorisation-Minimisation
mmWave	Millimeter wave
OFDM	Orthogonal Frequency Division Multiplexing
OPL	Optimised Programming Language
SBS	Small Base station
SC	Small Cell
SINR	signal-to-interference-plus-noise ratio
SISO	Single-Input Siingle-Output

Abstract

With 5G and beyond networks, energy consumption and operational costs go up because of a focus on network densification. New technologies are needed for better resource management. In spite of the access network's ability to handle large amounts of traffic, the backhaul network has slowed cellular network growth and efficiency. Even though passive fiber-optic networks are generally available in the backhaul network, capital expenditures and indirect operating costs hamper ultra-dense deployments. The use of mmWave links as a backhaul management option in the light of network softwarization (Software Defined Networking, SDN) may be one way to combat the above issue, but this may not be the best solution since there is a possibility of NLOS communication in the mmWave backhaul link due to various obstacles, such as densely situated base stations and transceivers located at low heights, such as street lights. Essentially, the above-mentioned issue can be resolved by using Intelligent Reflecting Surfaces (IRS), also known as software-controlled metasurfaces, on the backhaul side, which will allow signals to travel from the source base station to the destination base station even in the absence of LOS links (mmwave links). Part I of this thesis examined the benefits of adding IRS channels along with backhaul mmWave channels. The MILP optimisation model is designed to minimise backhaul networks' total power consumption and to guarantee maximum service to their users. Based on the MILP results, using IRS channels together with mmWave channels can reduce both static and dynamic backhaul power if a certain number of mmWave channels is blocked simultaneously. In our study range and input parameters, the number of active elements within an IRS can further optimize dynamic backhaul power. The IRS can also be deployed in minimum numbers through mmWave channels to optimize the number of users served.

Furthermore, IRS provides the advantages of edge computing in a multi-channel, multi-user downlink communication system in which the necessary bandwidth is guaranteed to all users. The objective of IRS is to minimize latency (including process delay and transmission delay) and power consumption. On the basis of the results of the model, it has been demonstrated that by modifying the weights of the objective function, which reflect the operator preferences, the system can minimize IRS latency, IRS power consumption, or both.

Chapter 1

Introduction

As outlined in this chapter, the motivation for studying the use of Intelligent Reflecting Surface (IRS) elements in a future backhaul network to overcome the challenges of intermittent connectivity, outages, and rapidly changing channel conditions is described. As mobile devices become more densely deployed, latency becomes a critical issue and we establish how IRS can be helpful in minimizing the latency associated with mobile access. The last section of this chapter discusses the limitations of the related work, presents our contribution to using IRS to improve Signal-to-Noise Ratio (SNR) in backhaul networks, to minimize latency in Mobile Edge Computing on the access side, and concludes by summarizing how the rest of the thesis is organized.

1.1 Motivation

It is estimated that modern wireless communication networks such as 5G and the future 6G networks will support more than 6 billion users. However, a significant proportion of applications require high data rates, [5]-[6], which will make more congestion on the current available spectrum. In addition, the demand for data services continues to grow, as revealed by the latest wireless network statistics, which show that in 2019, the data traffic of mobile users was 68% more than the previous year, i.e., 38 Exabyte (EB) as compared to the 27 EB in 2018 , [7]. As a result, Ericsson predicts a growth of 27% per year in mobile data traffic until 2025[5]. Today's 5G [8],[9] focuses on a user-centric, comprehensive, concrete system design, resulting in a vision of more than 10 Gbps per user peak data rate with an less than 1ms of end-to-end delay (E2ED) [10].

In addition, 5G user density is supposed to be around 300 users per kilometre [8] , [11], which will increase load on the capacities offered by current wireless networks. Therefore, new schemes such as MIMO (Multi Input Multi Output) are implemented with the latest modulation techniques and further spectrum exploitation to satisfy the demands. Also, in 5G, cell sizes are expected to be small; therefore, mass deployment of small cells (SC) is one solution to offer limited capacity on demand. Furthermore, Umbrella eNodeBs also known as macro-base stations are used to provide coverage for small base stations (SBS) in SCs on demand [12].

These small cells (SC) can take advantage of reusing frequencies, thus improving the area spectrum efficiency. The purpose of the backhaul network is to provide a connection between the core network and base stations (BSs), and normally, fibre optic is used for this purpose. However, providing fibre optic connections for all the BSs in small cells may result in higher implementation and maintenance costs

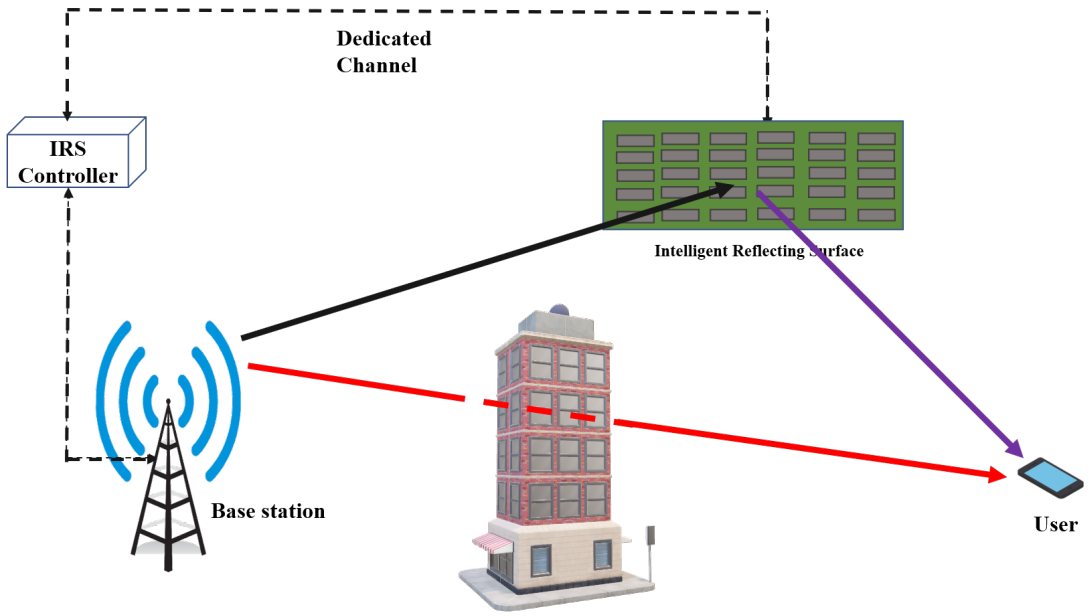


Figure 1.1: IRS used to provide virtual path to the user.

for the telecom operators. Under those circumstances, researchers have come up with wireless backhaul (BH) solutions, which relay the user traffic towards the operator's core network using the wireless interface to provide the required capacity with limited costs. Due to the massive spectrum available in the 60, 70, or 80 GHz bands, millimetre Wave (mmWave) technology is considered an exciting candidate for future 5G backhaul networks. Although, due to Friis law, omnidirectional free path loss increases at these frequencies [13], beamforming techniques can be applied to compensate for the path loss [14]. However, shadowing is one of the vital factors for mmWave signals causing outages, intermittent connectivity, and rapidly varying channel conditions [15]. Thus the deploying mmWave in the backhaul link has its limitations, i.e., as the base stations are nearly deployed, there is a chance of blockage due to some obstacle that will obstruct the mmWave link, causing the re-routing of the backhaul traffic.

A new emerging hardware technology known as the Intelligent Reflecting surface has just arrived; it reflects the impinging plane wave in the shape of a beam, thus acting

as a reconfigurable reflectarray by employing Micro-Electrical-Mechanical Systems (MEMS) or varactor diodes [16], [17] and [18] to achieve a desired objective. The objective includes but is not limited to increasing the received power of the wanted signal, decreasing the interference power, or improving the relationship between these two. IRS can accomplish these goals via beam-focussing, which changes the amplitude of the radiation pattern in a particular direction or changes the direction of the beam called beam steering. In Fig.1.1, IRS enhances coverage by creating virtual links for users with blocked direct link to the BS. There are many use cases where IRS can be employed [19], which includes: unmanned aerial vehicles (UAV) communications, mmWave coverage extension, wireless information and power transfer, physical layer security, etc.

In Mobile Edge Computing(MEC), mobile devices with resource-incentive applications can make use of computational offloading [20]. Nevertheless, when the communication link used to offload the computational tasks is not favourable, the benefits brought by MEC cannot be fully exploited. Fortunately,by leveraging the benefits of IRS, MEC paradigm can also be employed here to mitigate the propagation-induced impairments.

1.2 Current Work and Limitations

IRS has recently drawn the attention of many researchers. For example, the authors in [21],[22] presented a detailed description of IRS technology and discussed advanced solutions and theoretical performance shortcomings. Energy efficiency approaches for phase transition of IRS elements and power allocation are presented in [23], where an exact model of IRS power consumption is exhibited. A practical application in outdoor environments shows that the IRS power allocation techniques

[24] can produce up to 300% energy efficiency compared to amplify-and-forward and multi-antenna systems. Joint Active and passive beamforming is mentioned in [16], where some recommendations are provided for optimal deployment of IRS. In [19], the far-field path loss model for IRS-based connection is derived by making use of techniques associated with optical physics where each reflective element acts as a diffuse scatterer. There is a practical IRS application described in [25], that uses 256 positive intrinsic negative diodes (PINs) with 2-bit phase transition to achieve a 21.7 dBi antenna gain at 2.3 GHz, and a 19.1 dBi antenna gain at millimetre wave (mmWave) frequency of 28.5 GHz.

1.2.1 Backhaul optimisation using IRS elements

The majority of the studies examined the use of IRS in future Access networks to accomplish various objectives; however, only one of the studies [26], employed IRS to provide wireless multi-hop backhauling services for multiple base stations (BSs) connected in a mesh topology. According to the authors, the authors analysed the performance of the proposed architecture from the perspective of Rician fading channels. This was done in order to calculate the probability of error and outage. In the simulation, it was determined that the IRS supported mesh backhauling architecture has several desirable characteristics that can be utilized to overcome the backhauling challenges. The proposed system model does not consider any backhaul traffic routing or the power consumption of IRS elements.

In the first part of this thesis, we developed an optimization model to determine the amount of backhaul power saved by introducing IRS channels along mmWave channels while guaranteeing maximum connectivity. This scheme shows the optimum value of active elements in an IRS-supported backhaul channel under the assumption that renewable sources, for example, solar panels, are utilised to activate these

elements.

1.2.2 Latency minimisation in Mobile Edge Computing using Intelligent Reflecting Surface

Mobile Edge Computing implements cloud computing capabilities near to the mobile users, i.e., at the edge of the mobile network. The main intentions are reducing the latency, ensuring highly efficient network operation, and improving the user experience. In [27], authors formed a joint user association and computation offloading problem to minimise mobile users' energy consumption and MEC servers. They disintegrated it into two sub-problems i.e., user association and computation offloading. As the name suggests, the user-association subproblem decides whether a mobile user can be served by a specific base station. In contrast, the computation offloading sub-problem jointly optimises the computational resources, transmission power, and offloading schemes. To reduce the computational latency, an IRS was introduced in [28] in MEC systems and a latency minimisation problem was formulated based on this model, subject to the constraints on IRS phase shifts and edge computing capability constraints. The authors developed algorithms to optimise the communication and computation settings. However, their work did not consider any impact on power consumption imposed by introducing IRS elements in the model. Against this background, in this part of the thesis, we examine the advantages of IRS in a mobile edge computing setting in a multi-channel-based multi-user downlink communication system by guaranteeing that all the users are served. A MILP model with objective function and constraints is developed. The objective function is to jointly minimize the latency (including process delay and transmission delay) and the power consumed by the IRS elements.

Figure 1.2: Thesis Organization

1.3 Thesis Outline

The thesis is organized into 7 chapters, as shown in figure 1.2, and the remaining part of it is outlined as follows:

- **Chapter 2:** In this chapter, we present background on the established work related to this thesis. We present a detailed description of the IRS construction, its channel model and novel idea of using IRS for frequency selectivity is

elaborated. Then, the use cases and key requirements for 5G are addressed. It then sheds light on small cell networks and the associated challenges as a key enabler to 5G. We then presented a brief overview of mmWave backhaul along with its its channel model, and architecture. Finally, a brief background of the Mobile Edge Computing (MEC) paradigm along with the issue of latency is exhibited.

- **Chapter 3:** In this chapter, the basics of optimisation problems and their formulation in MILP is exhibited. We present a node-link formulation to design our network models that consider each link's capacity. A piecewise Linearization technique is developed to linearise the non-linear objective functions and constraints. Moreover, a brief review of genetic algorithms to solve the optimisation problems is also presented.

- **Chapter 4:** This chapter presented the idea of using the IRS in the mmWave backhaul links to overcome the issue of blockages and how the IRS can provide savings in the backhaul power consumption. We developed an optimisation model based on MILP to minimise the total backhaul power while guaranteeing the maximum users are served. We evaluate the suggested architecture under different scenarios related to the type of blockages in backhaul, the availability of IRS channels by measuring the power consumed by base stations for forwarding the backhaul traffic and the power consumed by IRS elements.

- **Chapter 5:** In this chapter, an IRS is proposed for employment in Mobile Edge Computing (MEC) systems to reduce the computational latency. A latency-minimisation problem is formulated based on the MILP model, subject to practical constraints on the total edge computing capability. The benefits of using IRSs in the MEC system were evaluated under various simulation environments.

- **Chapter 6:** In this chapter, a heuristic based on a genetic algorithm is developed to validate the results obtained from the optimisation problems. We specified binary chromosomes resembling our decision variable. A unique selection and mutation process for the generated population using those chromosomes is also presented. We also devised a repair function to check the validity of generated chromosomes and repair according to certain constraints. Different fitness functions are evaluated related to our optimisation problems. The results obtained from heuristics are comparable with those results obtained earlier with the MILP model.

- **Chapter 7:**This chapter concludes the thesis and discusses possible future directions for research.

1.4 List of Publications

- Aftab, N., Zaidi, S. A. R., McLernon, D. C., & Lawey, A. "optimisation of Intelligent Reflecting Surfaces in 5G and Beyond Millimeter Wave Backhaul Networks", (to be submitted to IEEE Transactions on Green Communications and Networking).

- Aftab, N., Zaidi, S. A. R., McLernon, D. C., & Lawey, A. "Joint Power and Latency minimisation of Mobile Edge Computing in Future Access Networks Assisted by Intelligent Reflecting Surfaces", (submitted to Journal).

- Aftab, N., Zaidi, S. A. R., McLernon, D. C., & Lawey, A. "Intelligent Reflecting Surface Assisted Future Access Networks", (to be submitted to Conference).

1.5 Conclusions

This chapter describes the rationale for exploring the use of Intelligent Reflecting Surface (IRS) elements to prevent intermittent connectivity, outages, and rapidly changing channel conditions in a future backhaul network. With more mobile devices being deployed, latency becomes a critical concern, and we discuss how IRS can help minimize latency associated with mobile access. In addition to discussing the limitations of related work, we presented our contribution to improving the Signal-to-Noise Ratio (SNR) of wireless backhaul networks, as well as reducing latency on the access side of Mobile Edge Computing and the chapter is concluded by summarizing how the rest of the thesis is organized.

Chapter 2

Background Theory

In this chapter, we are going to provide the background knowledge with respect to backhaul communications using mmWave based on IRS and how to minimize latency using IRS in a MEC system. An in-depth description of the IRS, its configuration, and the different advantages IRS offers in 5G and future 6G networks is presented. We discuss key performance indicators (KPIs) for 6G and small cell deployments for 5G. A study of mmWave backhaul communication with the implementation of IRS is presented. Towards the end of the chapter, we describe mobile edge computing in terms of its communication and computation models.

2.1 Intelligent Reflecting surface(IRS)

Recently, IRS, also called metasurfaces, has been developed to control the propagation medium, thus enhancing the quality of service (QoS) to expand the wireless

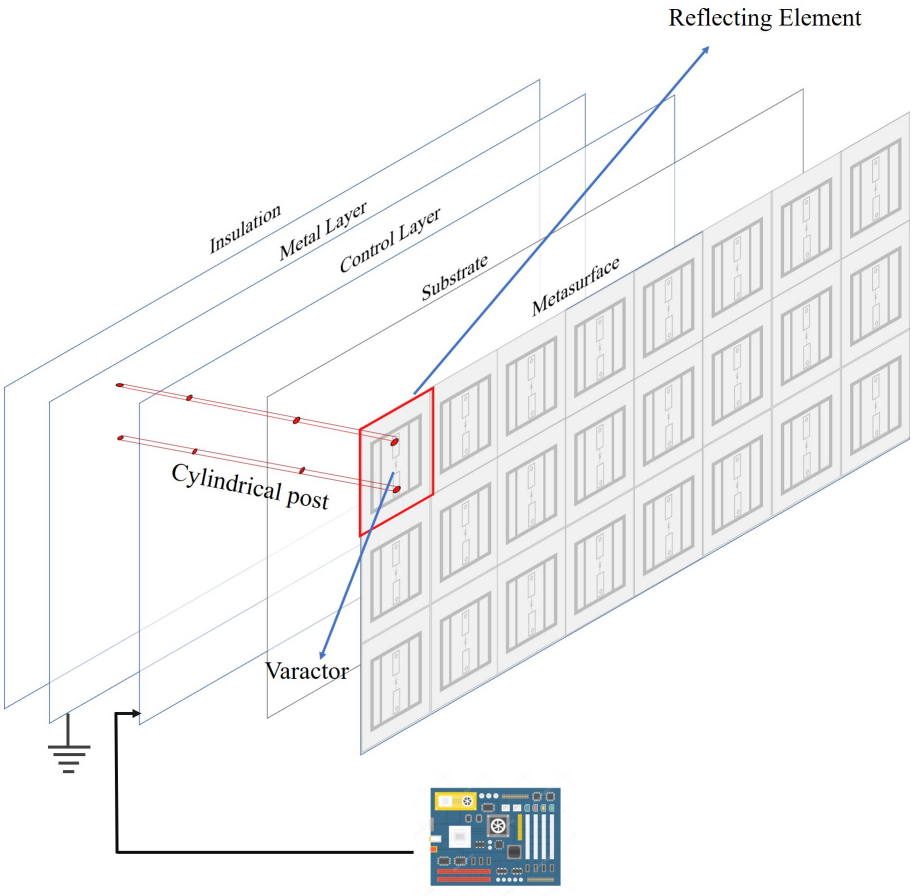


Figure 2.1: The architecture of IRS, where a single reflecting element is marked in red [1].

networks energy and spectral efficiencies. This technology is expected to play an essential role in the energy and spectral efficiency of 5G and beyond future wireless networks. IRS employs nearly passive antenna elements, introducing phase shifts to the received signals and reflecting them to the destination. Thus, to get an efficient transmission, multiple reflectors are implemented, and the reflected signals add coherently after the proper selection of introduced phase shifts. Thus, the signal-to-noise ratio (SNR) improves dramatically and hence, enhances the spectrum efficiency [29], [30]. The following sub-section explained the configuration of the IRS.

2.1.1 IRS configuration

The IRS is normally embedded on a printed circuit board (PCB), and the metasurfaces are located on the top of a dielectric substrate, having an extensive number of metallic patches printed periodically, as shown in Fig.2.1. In a reflecting element, two different metallic patches are connected through a common varactor diode, each of which is connected with the conductive cylindrical post, thus forming a parallel connection between the varactor and the cylindrical post[1]‘. Those cylindrical posts passes through a dielectric substrate and a control layer and ends at the metal layer, whereas, the metal layer performs the role of a ground plane. Thus, the substrate's thickness and the length of the cylindrical post determine the circuit's inductance. In the control layer, a DC voltage is applied by the FPGA-based controller, and the capacitance of each varactor varies depending on the individually controlled bias voltages. An insulation layer eliminates electromagnetic leakage from the metal layer. When an RF signal reaches the metallic patch, the current flows through the metallic patch is divided into two different paths, i.e., through the cylindrical post and the varactor. The distributed current meets on the opposite side of the metallic patch and hence the RF signal is reflected back by that metallic patch.

2.1.2 IRS path Loss

The potential performance gains brought by adopting IRS can be confirmed by formulating a path loss model for received power and SNR performance. Based on wireless communication's conventional two-way channel model, the authors in [21] specified a simplified path loss model. As shown in Fig.2.2, in addition to the straight path from the transmitter to the receiver, each reflective element in the IRS provides the second path from transmitter to receiver, thus establishing a two-way

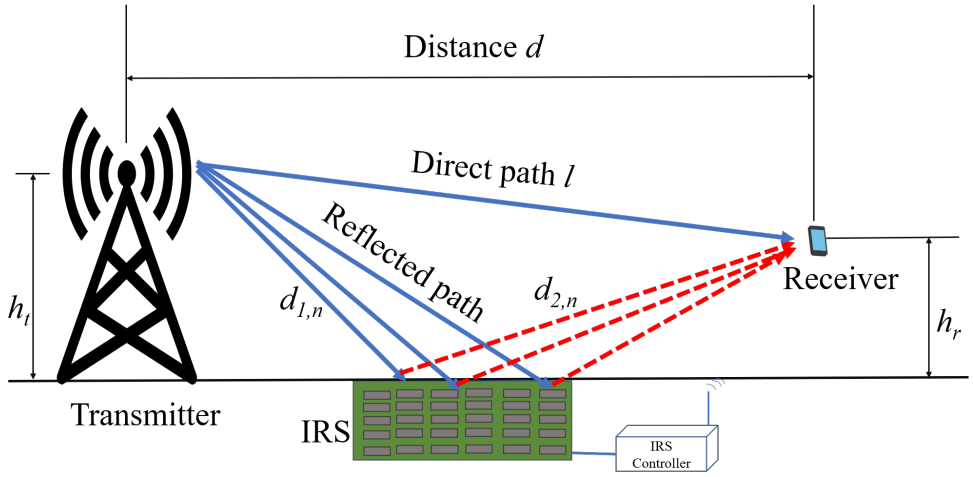


Figure 2.2: An IRS-aided two-way channel model for wireless communications where IRS controller can configure each IRS element

signal propagation model. The signal power at the receiver can be determined by the combination of all the signals from all the paths, as follows

$$P_{Rx}(d) = P_{Tx}\left(\frac{\lambda}{4\pi}\right)^2 \left|\frac{1}{l} + \sum_{n=1}^{N} \frac{\kappa_n e^{-j\triangle\phi_n}}{d_{1,n} + d_{2,n}}\right|^2, \tag{2.1}$$

where, P_{Tx} and P_{Rx} represents the transmit and receive power of signals, respectively, the number of elements in IRS are denoted by N, and l represents the length of the direct path, from which the distance d between transceivers can be approximated. Through each n-th reflecting element, the distance between the transmitter and IRS is denoted by $d_{1,n}$, and the distance between IRS and receiver is denoted by $d_{2,n}$. In Eq.(2.1), the summation term represents the signal reflections through different paths due to IRS elements. The distance between each direct and reflected path through the n-th element determines the phase difference, $\triangle\phi_n$. Whereas, κ_n, represents the reflection coefficient that depends on the reflected object's EM properties, which is traditionally uncontrolled without using the IRS. However, in case of IRS, the phase shift of each IRS element can be controlled in a way such that $\kappa_n = e^{j\triangle\phi_n}$; thus, both the phase of the direct path and reflected signal through N

elements are coherently aligned.[31]. Usually, an assumption is made for distances, i.e., $l \approx d_{1,n} + d_{2,n} \approx d$ for all n, which approximates the signal strength obtained in Eq.(2.1), as

$$P_{Rx}(d) \propto (N+1)^2 P_{Tx} \left(\frac{\lambda}{4\pi d}\right)^2. \tag{2.2}$$

Thus, in the absence of direct link or with a large number of reflecting elements, N, the above path loss can be simply rewritten as

$$P_{Rx}(d) \approx N^2 \left(\frac{\lambda}{4\pi d}\right)^2. \tag{2.3}$$

The N^2 term in IRS supported path loss Eq.(2.3) introduces some extra gain to the received signal power compared to the received signal power in free space path loss.

2.1.3 Frequency Selectivity using IRS

The IRS does not only provide the gain in the path loss, but it also provides frequency selectivity, a unique feature of IRS, which can be determined by the equivalent circuit of each reflecting element of an IRS as shown in Fig.2.3, where, L_1 represents the inductance resulting from two conductive cylindrical posts, and L_2, R and C are passive elements of the the varactor, which can be determined according to the varactor model in [32]. By choosing the suitable capacitance C, the amplitude and phase shift of the reflected signal can be configured. To be explicit, for a signal of frequency f, the impedance of the reflect element circuit can be written as

$$Z(C,f) = \frac{j2\pi f L_1 \left(j2\pi f L_2 + \frac{1}{j2\pi f C} + R\right)}{j2\pi f L_1 + \left(j2\pi f L_2 + \frac{1}{j2\pi f C} + R\right)}, \tag{2.4}$$

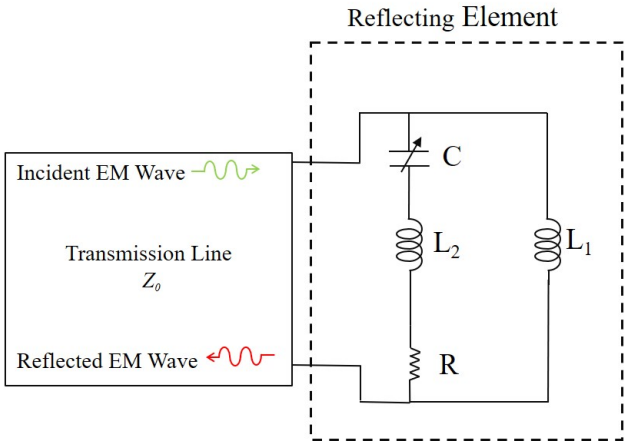


Figure 2.3: The equivalent circuit of an IRS element(in dotted box)

thus the reflection coefficient of each IRS element is given by

$$\phi\left(C, f\right) = \frac{Z(C, f) - Z_0}{Z(C, f) + Z_0}. \tag{2.5}$$

From Eq.(2.5), with determined circuit parameters, signals of different frequencies can generate different responses (reflection coefficients) through the same IRS element.

2.1.3.1 Comparison of IRS with other related Technologies

The proposed IRS differs remarkably from related technologies like amplify-and-forward (AF) relay and the backscatter communication. First, opposed to the AF relay, which amplifies and regenerates signals to assist source-destination transmission. IRS as a nearly passive array only reflects the signal, without any transmitter module, consuming comparatively less transmit power than AF relay. Second, unlike the traditional radio frequency identification (RFID) tag of the backscatter communication, which interacts with the receiver by reflecting the signal sent by a reader, IRS is primarily used to improve the communication efficiency of an existing link

Table 2.1: A comprehensive comparison between IRS and other related technologies

Technology	Functional mechanism	No. of transmitter RF Chains needed	Duplex	Energy consumption	Hardware cost	Role
MIMO relay	Active, receive and transmit	N	Half/full duplex	High	High	Helper
Backscatter	Passive, reflect	0	Full duplex	Very low (10 μW) [33]	Very low	Source
IRS	Nearly passive, reflect	0	Full duplex	Low (5 mW) [24]	Low	Helper

rather than producing its own information by reflection. Thus, in backscatter communication, the direct-path signal (i.e., from reader to the receiver) is unwanted interference and therefore receiver must suppress/cancel this interference. However, in IRS-enabled communications, the same information is conveyed by both the direct and reflected path signals. A further comprehensive comparison between the IRS and other technologies is compiled in table 2.1.

To study the benefits of the IRS in wireless communications, extensive research works have been made to evaluate its channel estimation [34], capacity analysis [35] and modelling of practical reflecting phase shifts [36], along with the associated phase shifts design [37]. In particular, joint design of the IRS phase shift and precoding at the BS is introduced in [37]. To reduce the transmit power while the received signal-to-interference-plus-noise ratio (SINR) is maintained. These studies are then expanded to the most efficient discrete phase shift settings [38]. Nevertheless, the developed algorithm in [37] has an extreme computational complexity which prevents its applicability in large-scale IRS. Guo et al. [39], suggested a low-complexity algorithms to decrease the complexity, whereas Pan et al. [40], presented majorisation-minimisation (MM) algorithms for the case of multi-cell scenarios. Moreover, to overcome the overhead due to the IRS channel estimation, authors in [41] grouped the IRS elements, so the same phase shift coefficient is shared by each group. As a result, they optimised the phase shift and power allocation in ortho-

gonal frequency division multiplexing (OFDM) systems. Moreover, the use of IRS is also explored in simultaneous wireless information and power transfer (SWIPT) [42], as well as in enhancing the physical-layer security [43]. These remarkable research contributions encouraged us to investigate further the benefits of IRS in the future 5G networks.

2.2 Small Cell Infrastructure in 5G Network

This section presents a collection of feature 6G Key Performance Indicators (KPIs) featured in various surveys. It then proposes one of the crucial technologies for improving the capacity of 5G networks, namely small cells, and then presents the related deployment challenges of small cells.

2.2.1 Key Performance Indicators in 6G

From the 1G network to the 5G network, and now to the future 6G network, the key performance indicators have largely stayed the same. However, with progression of each generation the minimum requirements have become more stringent. There is one exception, the energy efficiency KPIs, which were introduced in 5G, but without specifying specific targets.

According to many analysts, 6G is likely to have a similar set of KPIs as previous generations, but with much greater ambitions as shown in Fig.2.4. Even though the KPIs were largely independent in 5G (although less strict at high mobility and for large coverage areas), a cross-relationship between the KPIs is desirable in 6G through a definition of groups. There should be no discrepancy in the requirements for different indicators in a group, but they should be completed at the same time.

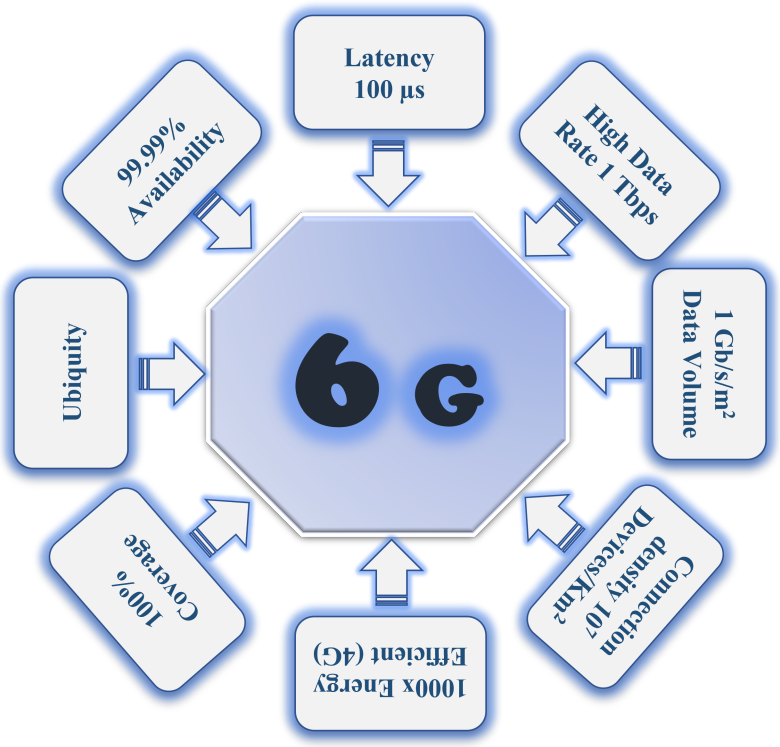


Figure 2.4: 6G networks requirements

As a result, we will go from a situation where broadband connectivity is delivered in a single way to a situation where the specifications of different broadband applications become too specialised to be unified simultaneously. Therefore, 6G will need to be configurable in real-time to cover these different groups.

The following are the desirable KPIs for 6G Wireless Access.

Extreme Data Rates: Both indoor and outdoor connections are expected to reach peak data rates of 1 Tbps. A minimum 1 Gbps data rate is intended to be experienced by 95% of users in every location.

Enhanced Spectral efficiency and Coverage: Through the use of improved MIMO technology and modulation schemes, peak spectral efficiency could be increased to 60 bits per second. However, the greatest expected improvement will come from the uniformity of the spectral efficiency throughout the coverage area.

The spectral efficiency should be 3 bits per second per Hz from the perspective of the end-user. Furthermore, it is necessary to develop new PHY layer algorithms to enable broadband connectivity in scenarios with high mobility and more broadly those for which legacy wireless network generations do not fully meet the needs.

Extra Wide Bandwidths: In order to support high peak rates, the maximum bandwidth must be increased significantly. A mmWave band can have bandwidths of up to 10 GHz, while THz and visible light bands can reach 100 GHz.

Enhanced Energy Efficiency: As part of the focus on sustainable development, 6G technologies are expected to strive for greater energy efficiency, both in terms of the total power consumption per device and the transmission efficiency. Accordingly, efficiency should reach up to one terabit per joule in the latter case. A core element of 6G is the development of energy-efficient communication strategies.

Ultra-Low Latency: By using bandwidths greater than 10 GHz, we can achieve a latency of 0.1 ms and a jitter of 1 micro second, resulting in extremely precise timing.

Extremely High Reliability: The requirement for mission-critical applications up to 10^{-9} reliability is prevalent in some new use cases. It is unlikely that all of these requirements will be simultaneously supported, but different use cases will have different sets of KPIs, where of only some reach the maximum requirements mentioned above.

2.2.2 5G Dense Small Cells Deployments

In 5G, the deployment of small cells to provide the local coverage and capacity using the low transmit power base stations (BSs), will be a critical enabler for increasing the future network's capacity and coverage, particularly in city centres and areas

with high traffic [8]. The typical coverage range of small cells is form ten to hundreds of meters [44]. They are classified as femtocells, picocells and microcells, subject to the coverage distance and transmit power, as shown in table 2.2, where a macro-cell is just included for comparison. These small cells can be used both indoors and outdoors environments.

Table 2.2: Classification of Cellular cells

Cell type	Typical cell size	BS transmit power	Number of users
Femtocell	10-20 (m)	$\leq$ 100 (mW)	Few
Picocell	4-200(m)	0.25-2 (W)	20-40
Microcell	200(m) -2 (km)	$\leq$ 2 (W)	100
Macrocell	30-35 (km)	5-40(W)	Many

A combination of radio technologies and different small cells results in a Heterogeneous Network (HetNet), as shown in Fig. 2.5. Due to the multi-tier architecture of HetNet networks, it is supposed to be one of the significant technologies in 5G networks. Specifically, the base stations of different tiers have distinct characteristics such as access technology, coverage area and transmit power [45]. Therefore, despite the gains such as localise coverage and high capacity, brought by small cells, many challenges like interference management, user association and ackhaul bottleneck need to be solved. A brief description of one of the challenge related to our thesis, i.e., user association and backhaul bottleneck, is given below.

User Association and Backhaul Bottleneck

In wireless communication research, user association issue has received much attention. Already developed user association strategies employ a mixture of criteria such as the available bandwidth, link quality (quality of reference signal received), and traffic type to improve the network efficiency with a lower probability of ser-

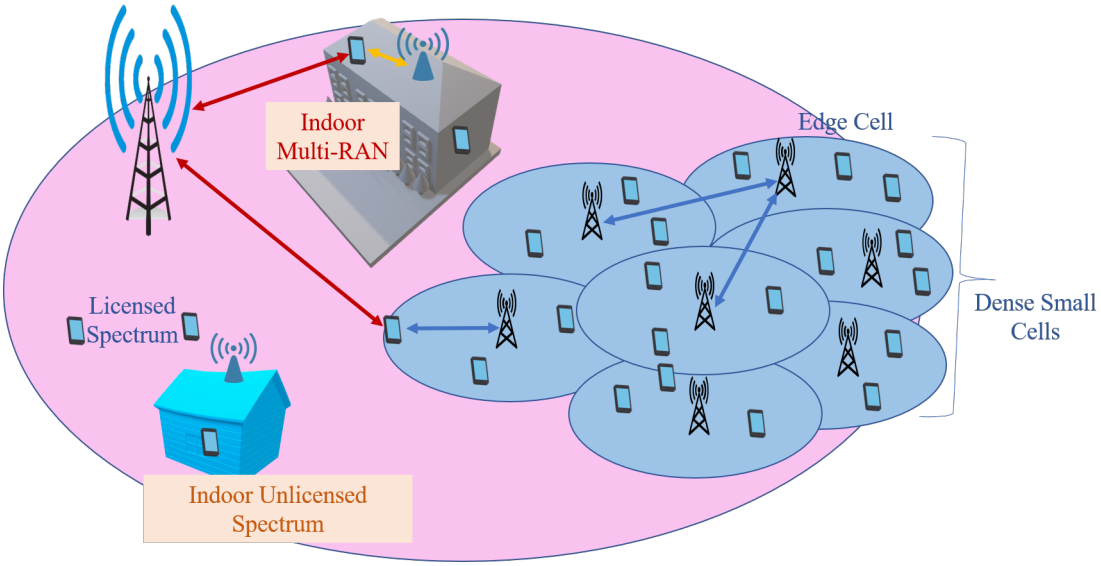


Figure 2.5: HetNet principle

vice interruption [46], [47]. However, one of the critical factors generally overlooked is backhaul bottleneck, where the role of the backhaul network is to connect the access networks to the core networks using wired medium like fibre optics or wireless interface(e.g. microwaves and mmWave). In the next section, we presented an idea of preferring the wireless technologies over the wired technologies for small cells backhaul.

2.3 Solutions to small-cell backhauling

Consider a backhaul network with a single macro-cell region, where the macro-cell BS is connected with the core network via fibre connection, and will serve as the gateway node to the core network for a set of small-cell BSs. Fig.2.6, portrays the conceptual view of small-cell backchaining. For the small cell backhaul connection, optical fibre is supposed to be the ideal candidate to accommodate endless capacity and high reliability; however, due to increased operational challenges, and higher

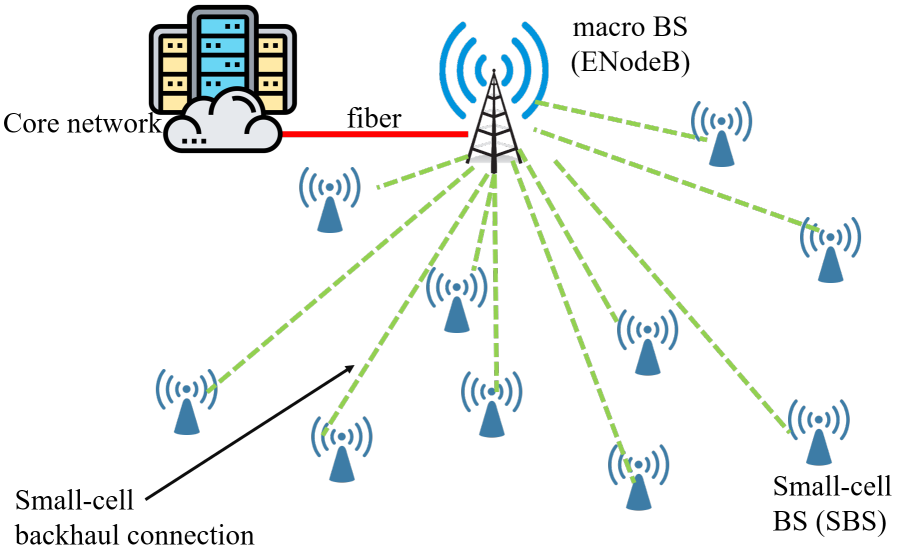


Figure 2.6: The conceptual idea of small-cell backhauling

associated costs, it is not reasonable to install adequate fibres in each small cell for future networks [48],[49]. Thus, wireless backhaul technologies have become a preferable solution for small cell backhaul due to their cost-efficiency and flexibility of deployment. In the next section, we will present a comparative analysis of wired mediums like optical fibre with other wireless interfaces.

2.3.1 Millimetre-wave band

Millimetre wave bands (millimetre waves) at 60 GHz and 70-80 GHz are preferable choices for line-of-sight (LOS) wireless backhaul solutions for future 5G networks due to the availability of high spectrum and hence, the higher data rates. In addition, these bands benefit from “light license” or “license exemption”, which allows for efficient frequency reuse, resulting in a much lower spectrum cost [15],[50]. Moving to the mmWave band, the ITU has released some other bands at 28–30 GHz, 38-40GHz and 57-64 GHz (the free-licensed band), which is extended to 71 GHz, thus giving a contiguous band of 14 GHz. Furthermore, in the E-band, 71–76 GHz,

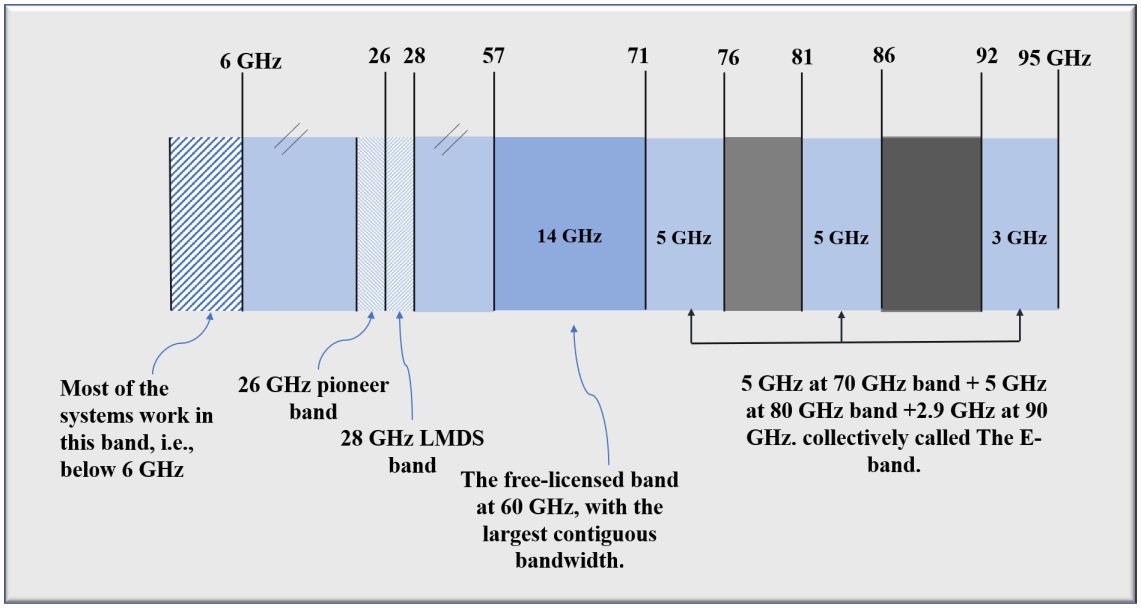


Figure 2.7: Different candidates mmWave bands for 5G

81–86 GHz, and 92–95 GHz bands are allocated for mmWave bands [51]. Ofcom in the UK has released one of the pioneer mmWave bands, centred at 26 GHz for 5G networks, ranging between 24.25 – 27.5 GHz [52]. Fig. 2.7, shows the mmWave candidate bands for 5G.

2.4 Analysis of mmWave wireless communications

The analysis on mmWave wireless communications can be tracked back to the 1960s when researchers began to study the propagation aspects of wireless signals on the mmWave band [53]. Some of them also investigated applications of mmWave wireless transmissions in the the radar and satellite communication systems [54], [55]. With the fast development in antenna technologies and electronic circuits, industries and vendors start implementing mmWave wireless communications in the current 5G cellular systems. Consequently, the research on mmWave wireless communications has freshly attracted much attention from both the academia and industry. In the

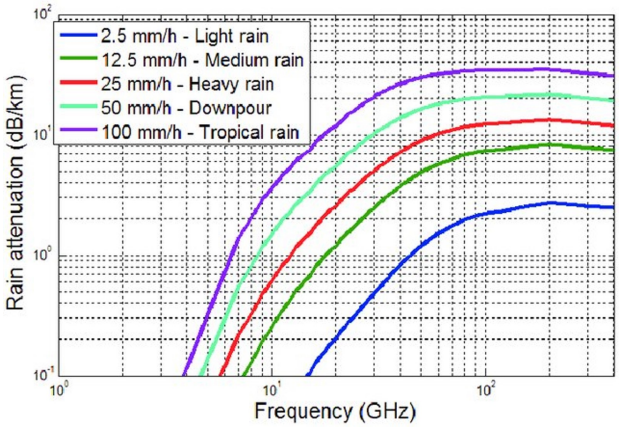


Figure 2.8: Rain attenuation of mmWave signals [2]

following section, we will focus on the latest research outcomes on the characteristics of mmWave wireless communications.

2.4.1 Propagation Characteristics of mmWave Wireless Signals

As one of the important technologies of 5G cellular systems, mmWave communications can make use of more than giga-Hertz (GHz) available frequency spectrum from 30 GHz to 300 GHz. Because the mmWave frequency range is very wide and different mmWave sub-bands have different propagation characteristics, only a few mmWave sub-bands, as mentioned in section 2.3.1, will be employed in 5G cellular.

2.4.1.1 Rain attenuation and atmospheric absorption at mmWave frequencies

Compared to the wireless signals in the sub-6 GHz band, mmWave signals suffer from huge propagation loss. One of the main reasons is that the propagation loss of wireless signals is quadratic to the signal frequency. Another significant reason is the weather condition may contribute significantly to the path loss of mmWave signals. It is well known that the atmospheric and molecular absorption and rain attenuation limit the range of mmWave communications [3, 56, 2]. Rain attenuation, also known

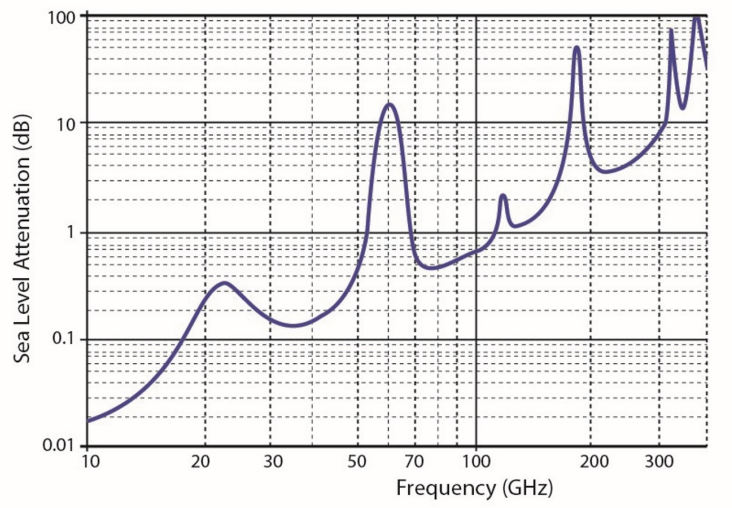


Figure 2.9: Atmospheric and molecular absorption of mmWave signals[3]

as Specific attenuation γ_R (dB/km) is depends on the rain rate R^{α} (mm/h), and is given by [56],

$$\gamma_R = kR^{\alpha}, \tag{2.6}$$

the values for the coefficients k , and α are determined as afunction of f (GHz) are given in [56, table 5]. In the heavy rain, mmWave signals experience significant attenuation because raindrops are comparable to wavelengths of mmWave and therefore causing a high signal scattering. The attenuation due to rain at mmWave frequencies is shown in Fig.2.8, whereas the attenuation due to the atmospheric and molecular absorption is shown in Fig.2.9. It can be deduced from Fig.2.8 and Fig.2.9, the propagation range in the mmWave bands is more affected by rain attenuation at 71-86 GHz, while, 60 GHz (V-band) suffers most from oxygen absorption. However, at sea level, the total attenuation by atmospheric gases at 71 -86 GHz band lies between 0.37 to 0.5 dB/km, which is insignificant considering the deployment density of small cells [57]. In the UK, the rainfall rate that surpasses 0.01% of the average year, is 35 mm/hr as specified by ITU-R P.837 and is designated as a heavy

rain. According to [56], the rain attenuation at this rainfall rate is estimated as 14 – 15 dB/km. Hence, for a mmWave backhaul link of 100m, the attenuation due to rain is almost 1.5 dB, which is comparatively lower than its path loss over the same distance. Therefore, mmWave communications are primarily used for small cell backhaul with cell sizes on the order of 300 m [58].

2.4.1.2 MmWave Channel Models and Path Loss calculation for Backhaul Networks

3GPP has recommended TR 38.901 and TR 38.900 channel models for above 6 GHz frequency, providing specifications on the channel's path gain, spatial and delay characteristics [59]. A lot of research work also concentrates on estimating mmWave channels. Here we focus on some relevant work based on the outdoor scenario. From the measurements in [58],[60], it is observed that the transmission range of mmWave signals can go beyond even 200 m in outdoor environments, which matches the conclusion above. Moreover, Haneda et al. [61] developed mmWave channels for frequencies range of 6 GHz to 100 GHz and examined several factors, like path loss, LOS probability, penetration, shadowing and blockage models. They classified blockages into two classes; one is the fixed blockage, i.e., due to the buildings and other structures, and the other one is the dynamic blockage which is due to humans, cars and trucks, etc. Note, that most of the above mmWave channel measurement and modelling concentrates on the access tier, which portrays the communication channel between end users and base stations, whereas our work concentrates on the backhaul tier of the mm Wave communication, which is considerably different from the access tier. For example, the transmit power is relatively high in the backhaul network, so more powerful devices can be deployed. Moreover, in the absence of dynamic blockage, the backhaul links are usually line-of-sight. Therefore,

to simplify our analysis for path loss modelling of backhaul network, the conventional assumption of additive white Gaussian noise (AWGN) channels is considered. The capacity of the mmWave backhaul channel is given by Shannon's capacity as

$$C = B \log_2(1 + \text{SINR}), \tag{2.7}$$

where B is the bandwidth of mmWave backhaul channel, SINR is signal to interference plus noise ratio, which is defined as

$$\text{SINR} = \frac{P_{Rx}}{\sigma^2 + I}, \tag{2.8}$$

where, P_{Rx}, denotes the power of the desired transmitter's signal at the receiver, σ^2 represents the power of thermal noise and I denotes the combined power of all the interfering signal from other transmitters. Nevertheless, due to the higher signal attenuation at mmWave frequencies, interference among the adjacent mmWave channels can be ignored. As a result of the static nature of the backhaul network, simple techniques for allocating frequency can be implemented at the beginning. Now, the Friis transmission equation is used for the calculation of P_{Rx} and is given as

$$P_{Rx}(d) = P_{Tx} \times G_{Tx} \times G_{Rx} \times \left(\frac{\lambda}{4\pi d}\right)^{\eta}, \tag{2.9}$$

where, the transmit power, antenna gains at the transmitter and receiver are denoted by P_{Tx}, G_{Tx}, and G_{Rx}, respectively. λ, denotes the wavelength of the signal, whereas the propagation distance and path loss exponent are represented by d and η , respectively. Eq.(2.9) governs the LOS propagation in free space and reveals that with the increase in the frequency, the path loss also increases. Hence, mmWaves are supposed to be the best candidate for short-distance communication links like

small cells backhaul. The ratio of power transmitted P_{Tx} to power received P_{Rx}, is known as free space path loss PL_{free} , and by assuming the isotropic antennas at the transmitter and receiver i.e., $G_{Tx} = G_{Rx} = 1$, and $\eta = 2$ for free space, PL_{free} is given as

$$\begin{aligned} PL_{free} &= \frac{P_{Tx}}{P_{Rx}}, \\ &= \left(\frac{4\pi d}{\lambda}\right)^2, \\ &= \left(\frac{4\pi df}{c}\right)^2. \end{aligned} \tag{2.10}$$

By using Eq.(2.10) the free space path loss in dBs [62], and is given by

$$P_{L_{free_{(dB)}}} = 20\log_{10}\left(f_{BH(MHz)}\right) + 32.45 + 20\log_{10}\left(d_{(km)}\right). \tag{2.11}$$

In. Eq.(2.11), f_{BH} denotes the frequency of mmWave backhaul link in MHz. Let the signal attenuation due to atmospheric gases (i.e., vapour and oxygen) and rain at distance, d are expressed as

$$PL_{d(dB)} = d_{(km)}(L_{vap} + L_{O_2} + L_R)(dB/km), \tag{2.12}$$

where, L_{vap}, L_{O_2} and $L_R\,(\gamma_R)$ represent the signal attenuation due to vapour, oxygen [3] and rain [56], respectively. Thus, the total path loss $PL_{(dB)}$ can be expressed as the sum of the free space path loss (PL_{free}) and the signal attenuation due to atmospheric gases and rain $PL_{d(dB)}$ as

$$PL_{(dB)} = PL_{free_{(dB)}} + PL_{d_{(dB)}}. \tag{2.13}$$

Hence, the signal to noise ratio $SNR_{(dB)}$, for mmWave channel [14], and is given as

$$SNR_{(dB)} = P_{Tx_{(dBm)}} + G_{Tx_{(dBi)}} + G_{Rx_{(dBi)}} - N_{th_{(dBm)}} - NF_{(dB)} \\ - PL_{(dB)} - T_{x_{loss(dB)}} - R_{x_{loss(dB)}} - L_{margin} \quad , \tag{2.14}$$

where, $P_{Tx_{(dBm)}}$ is the transmit power, $G_{Tx_{(dBi)}}$ and $G_{Rx_{(dBi)}}$ are the transmitter and receiver antenna gains, respectively. $N_{th_{(dBm)}}$ stands for the thermal noise, $NF_{(dB)}$ denotes the noise figure, $PL_{(dB)}$ is the total path loss, $T_{x_{loss(dB)}}$ and $R_{x_{loss(dB)}}$ represent the transmitter and receiver losses, receptively and LM^{BH} is the link margin. The thermal noise $N_{th_{(dBm)}}$ is given by

$$N_{th} = 10 \log_{10} \left(k_B T_{(K)} B_{(Hz)} \right), \tag{2.15}$$

where, k_B denotes the Boltzmann's constant (joules per kelvin), T is the absolute temperature (kelvin) and B denotes the channel bandwidth (Hz).

2.4.1.3 Blockage in mmWave backhaul

It is commonly assumed that the wireless signals cannot diffract easily around those obstacles whose size is considerably larger than the wavelengths of wireless signals. Consequently, backhaul links in the 60 GHz band with smaller wavelengths are prone to blockage by obstacles in outdoor and indoor environments. For instance, there are large numbers of trees, buildings and billboards in the dense urban outdoor area as well as moving big vehicles. Thus, for a partially blocked mmWave backhaul link the path loss in Eq.(2.13), become

$$PL_{(dB)} = PL_{free_{(dB)}} + PL_{d_{(dB)}} + BL_{(dB)} \tag{2.16}$$

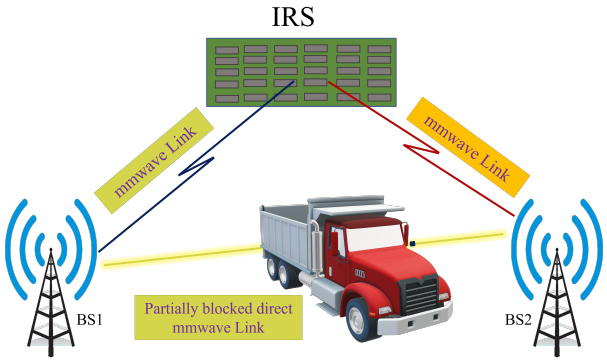


Figure 2.10: IRS supported mmWave backhaul due to dynamic blockage [4]

where $BL_{(dB)}$, is the random blockage parameter in the range of 20-50 dBs to incorporate dynamic blockages due to the bigger obstacles like cars or truck in mmWave backhaul. As already mentioned in section 2.1.2, introducing IRS in the wireless system enhanced the gain by a factor of N^2, so to tackle the issue of dynamic blockage in mmWave backhaul, we introduce an IRS along the mmWave channel as shown in Fig.2.10. In the next section, we will provide some background information related to the architecture of mmWave backhaul networks.

2.4.2 Architecture of mmWave backhaul

To achieve better performance, reliability, cost-effective solution and high backhaul throughput, different mmWave backhaul topologies are employed by the research community and industry. Two famous topologies are discussed here.

2.4.3 Star topology

This approach assumes a centrally located macro base station and the small cell base stations are distributed in the vicinity of that macro base station as shown in Fig.2.11. All the small base stations have the same coverage and transmission

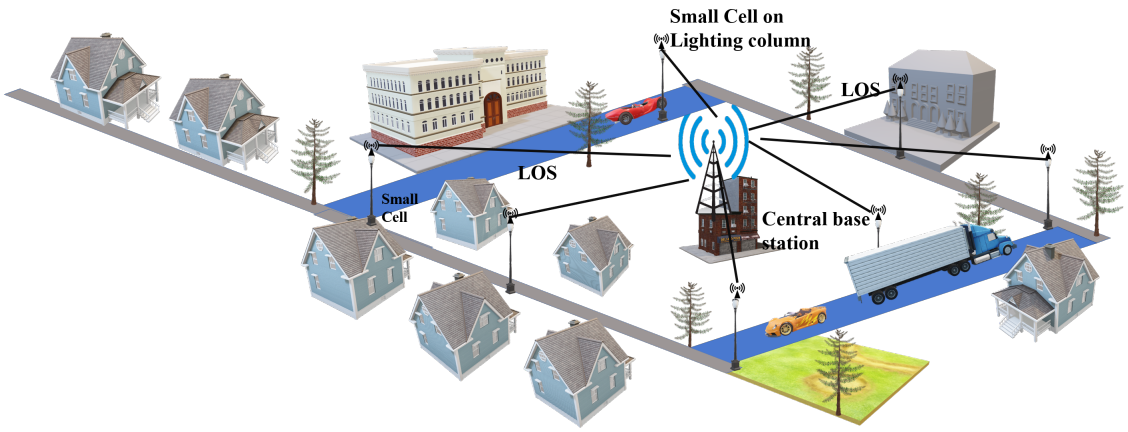


Figure 2.11: Star topology of Small cell wireless backhaul

power. Traffic from small cells is aggregated at the macro base station; then, it is transmitted to the core network through the existing fibre connections [49].

2.4.4 Mesh Topology

In a mesh topology, traffic from one small cell is layered with the help of other small cells to the macro base stations, which is connected to the core network via a fibre link [49], as shown in Fig.2.12. It has certain benefits over star topology, like it can provide greater resilience to traffic fluctuations and availability due to the path redundancy between the small cells. Nevertheless, these advantages come with related costs such as scalability issues and complexity of topology management, although they can be optimised by applying proper planning. In our backhaul scenario, we used mesh topology for mmWave backhaul.

2.5 Mobile Edge Computing Paradigm

Cloud computing as an efficient computing platform has enjoyed rapid development over the last few decades, mainly driven by ever-growing computing and processing

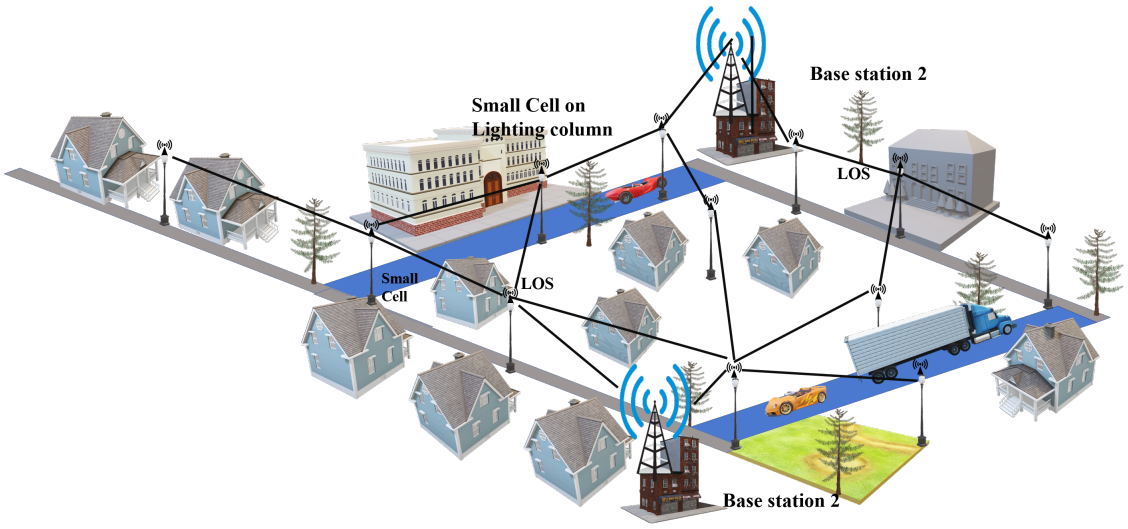


Figure 2.12: Mesh topology of Small cell wireless backhaul

demands of client devices. Recently, a brightly new concept of Mobile Edge Computing(MEC) gained significant attention from both academia and industry, which is promising to provide computing services with high bandwidth, ultralow latency and real-time access through shifting cloud computing from the centralized data centres located remotely to the edge of mobile networks proximate to end-users [63]. In this section, we present the basic background of Mobile Cloud computing (MCC), Mobile Edge Computing (MEC) and forces that driven the shift of cloud computing from the centre to the edge of the networks.

2.5.1 Mobile Cloud Computing

As the most effective and convenient communication tools, mobile devices like smartphones have become part and parcel of our daily lives. Due to the limited size and resources of these devices, they cannot effectively handle the computation-intensive or latency-critical tasks. Therefore, the concept of Centralized Cloud Computing CCC [64] appeared to offload the ever-increasing computation-intensive tasks to remote powerful data centres for computing, also called central clouds. Mobile Cloud

Computing MCC is a refined concept in which CCC is integrated into the mobile environment, thus enables the mobile users to explore the benefit of cloud resources [65]. MCC can be interpreted as a combination of mobile networks and CCC [66], and it can be considered as one of the most common tools to access applications and services on the internet by mobile users. The primary function of CCC/MCC is computation offloading, i.e., shifting intensive computation from resource-limited UEs to powerful central cloud data centres.

Although the CCC / MCC can provide mobile users with high-performance computing services, the location of central clouds is still far away from users. Therefore, excessive transmission latency occurred while accessing the CCC/MCC services. Thus, the MEC systems are developed to deal with UEs' computation-intensive latency-critical tasks.

2.5.2 Mobile Edge Computing

The concept of MEC was first proposed by European Telecommunications Standards Institute, ETSI in 2014 as a new platform for providing cloud computing capabilities near to mobile users [67]. Thus, the reason behind introducing MEC is that UEs' computation-intensive and latency-critical tasks can be offloaded at the edge of wireless networks by using edge cloud servers for computational processing. These cloud servers, also known as MEC servers, are deployed at the wireless network's base stations to release the pressure from resource-limited UEs to do massive computing workloads and hence increase their battery lifetime. MEC servers are typically data centres on a small-scale, used by telecom or cloud computing operators, and they are located in conjunction with base stations, as shown in Fig.2.13. In this way, the MEC allows the BSs to have the ability of storage and processing, and thus guarantee that the UEs can be directly connected to the edge clouds.

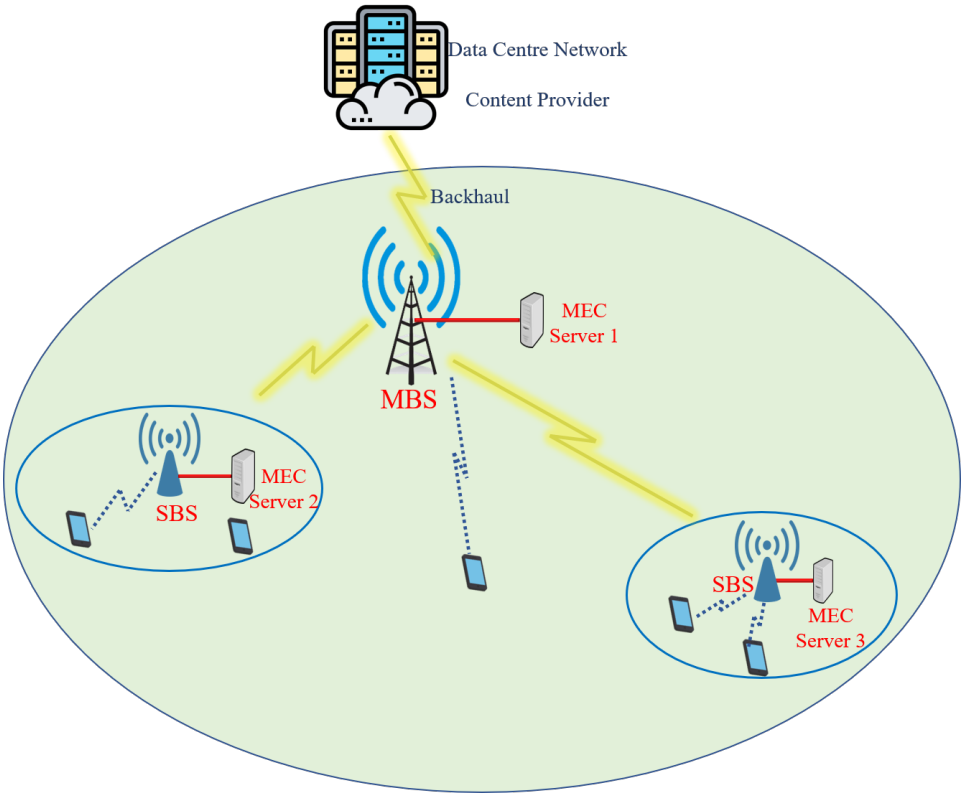


Figure 2.13: A simplified Mobile Edge Commuting (MEC) architecture of mobile network

In comparison with the MCC, the MEC has four main advantages in the aspects of latency reduction, energy saving, context awareness, and privacy/security enhancement as they are available iclose to the users. The attractive advantages of MEC lead to the fact that it has been seen as one of the key enablers in shaping future advanced wireless networks. Recently, both academia and the industry gave significant attention to the MEC [68] , [69], by identifying a large number of use cases for MEC. Some use cases, including but not limited to, are intelligent video acceleration and massive machine-type communications (mMTC). [20]. Currently, the full potential of this MEC paradigm has not been fully realized, primarily due to the low availability of the computation off-loading link. Devices at the edge of the cell, for example, have a low success rate of off-loading and/or may have a longer latency if they off-load their computations. Therefore, these devices must rely on their own

computing resources. Unfortunately, these resources are frequently inadequate to handle resource-intensive applications. Accordingly, from a communications perspective, it is important to improve the efficiency of MEC.

2.6 Computation Task Model

The computation tasks can be affected by various parameters such as task size, computation intensity, latency, bandwidth utilization, context awareness, scalability, and generality, etc., and thus developing accurate computation task models is highly sophisticated. It is known that latency used for communication and computation, have been widely considered as one of the important performance metrics for MEC systems, and the objective is to complete the UEs' computation in tensive latency-critical tasks with low latency. Hence, to properly describe the properties related to e latency, we adopt a reasonable and mathematically tractable computation task model in this thesis, which has also been widely used in the existing MEC literature. For a given computation task with fixed computation task size, it can be fully characterized by a positive parameter $[L, c]$. Here, L denotes request size (in bits) of the computation task-input data, c is the required computational resources to compute 1-bit of task-input data, also called computation workload/intensity.

2.6.1 Computation Offloading Modes

According to the structural characteristics of various applications or computation tasks, different computation offloading modes should be leveraged to deal with different computation tasks [70]. In this subsection, we introduce computation offloading modes, which are local offloading, complete offloading and partial offloading mode.

- **Local computing**: all task data is processed locally at the node without passing it to the MEC server. Factors affecting the effectiveness of local computing are node's data read speed and CPU performance.
- **Complete offloading**: all task data is transmitted to the MEC server, where the MEC server processes these tasks using its computing resources. Factors that affect the efficiency of complete offloading are the task size, computing power of the MEC server, channel condition and channel capacity.
- **Partial offloading**: According to the offloading algorithm, the data task is divided into two components, some of the data is processed locally, while the remaining data is processed in the MEC server. Partial offloading algorithms consider factors like data transmission, computing resource allocation and energy consumption before taking offloading decisions. Therefore, partial offloading is comparatively considered one of the most complex models because it must integrate different influencing factors.

In this thesis, we selected complete offloading scheme for MEC by assuming that all users will have limited resources available to them.

2.6.2 Communications in MEC Systems

In MEC systems, communications act as an essential part for completing users' computation tasks, which typically happen between UEs and BSs (with co-located MEC servers) through wireless channels. For computation tasks, communications mainly correspond to the computation offloading from UEs to the MEC servers (at BSs) in the uplink mode, whereas in downlink mode, communications is related to download the computation results from the MEC servers the UE. In fact, BSs not only provide wireless interfaces for MEC servers, but also allow access to remote

central clouds (large-scale data centres) through backhaul links, thus assisting the MEC servers to further offload some computation-intensive tasks to enjoy the more powerful computing capabilities at the central clouds.

Next, we will analyse the communications in MEC systems for the latency perceptive. The maximum achievable communication rate (in bits per second), i.e., the channel capacity, R is given with the help of the Shannon-Hartley theorem [71], as

$$R = B \log_2 \left(1 + \frac{S}{\mathcal{N}_0}\right), \tag{2.17}$$

where B denotes the bandwidth of the wireless channel in Hertz (Hz), S indicates the average received signal power in watts (W), $\mathcal{N}_0$ denotes the average power of the noise and interference over the bandwidth. Based on the computation task model, $[L, c]$, the communication latency, t_{trans} (also called transmission latency), for offloading L bits of computation task-input data from a MEC server to the UE in downlink can be calculated as

$$t_{trans} = \frac{L}{R}, \tag{2.18}$$

where, R is the capacity for computation offloading given by Eq.(2.17). Fig.2.14, shows an IRS is introduced in the MEC setting to reduce the transmission delay by offering more capacity (data rate) to a single user in downlink mode.

2.6.3 Computation in MEC system

Computation also plays an important role in MEC systems for completing the UEs' computation tasks. Similarly, in this part, we mainly pay attention to the analysis of the latency related to computation in MEC systems. Based on the computation

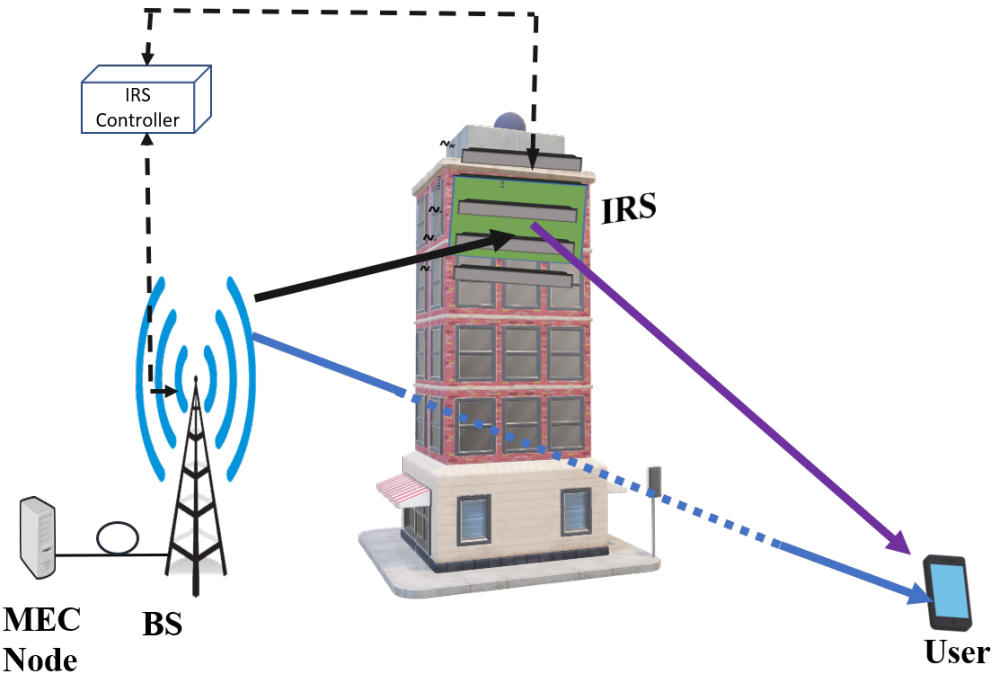


Figure 2.14: An IRS-supported single user downlink communication system with the provision of MEC

task model presented earlier, i.e., the computation latency for completing the task $[L, c]$ by the MEC server is given as

$$t_{proc} = \frac{Lc}{f}, \tag{2.19}$$

where f denotes the CPU Cycles allocated by MEC server for each user task.

2.7 Conclusions

The purpose of this chapter was to provide an overview of backhaul communications using mmWave based on IRS and to discuss how to minimize latency by using IRS in a MEC system. Detailed information is provided on the IRS, its configuration, and the numerous advantages it offers in 5G and future 6G networks. Moreover this

chapter discusses the key performance indicators for the deployment of small cells in 5G and 6G networks. MmWave backhaul communication is analysed in conjunction with the implementation of IRS. Towards the end of the chapter, mobile edge computing is described in terms of its communications and computation models.

Chapter 3

Mathematical Modelling and Solution Tools

In this chapter we will provide background information related to the optimization problem modelling which will help us to understand and solve it. The focus of this chapter is to describe linear programming and the AMPL model, the optimization tool we used to solve optimization problems. We will model a small wireless network using the node-link formulation. The remainder of this chapter explains genetic algorithms that can be used to verify the results presented.

Taking a look at the tools that will be utilized for our optimization problem, we will now lay the foundation for our optimization problem.

3.1 Mixed Integer Linear Programming (MILP)

Linear Programming (LP), mathematical modelling technique and a particular example of mathematical programming [72]. LP is considered as is one of the three

constrained optimisation classes (linear, nonlinear and integer), having all of its mathematical expressions (equations and inequalities) linear, as its name suggests [73]. The standard form of linear programming model is based on four components:

- The objective function characterises optimisation outcome, and the optimisation problem aims to minimise or maximise this objective function based on whether the model's result is a cost or a reward.
- Provide a model for the variables where their values represent the feasible solution of the objective function when all constraints are satisfied. On the other hand, these variables are the best objective values if they lead to the optimal solution [74]. When one of these variables is a non-integer, the optimisation technique is known as Mixed Integer Linear Programming (MILP).
- Constraints are a set of linear mathematical expressions (including both equations and inequalities), which will shape the feasible region of the solution for a particular optimisation problem (polyhedron region).
- Boundaries of variables that are used to restrain the lower and upper limits of all the variables in the model.

In order to properly understand the MILP programming, consider the following problem that will use all these components together to formulate the problem [75].

Objective function

$$y = c^T.x, \tag{3.1}$$

with subject to the following constraint

$$A \centerdot x \geq b, \tag{3.2}$$

and a constraint for non-negativity, i.e.,

$$x \geq 0, \tag{3.3}$$

where y is the objective function and the transposed vector of cost / reward coefficients vector is represented by c^T as

$$c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \tag{3.4}$$

the vector of decision variables is represented by x and is given as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \tag{3.5}$$

A denotes the constraint matrix which is

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \tag{3.6}$$

in Eq.(3.2), b is the minimum requirement that is must be met and is given by

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \tag{3.7}$$

In spite of the fact that several approaches can be utilised to solve constrained optimisation problems, systematic approaches are almost always preferred for solving linear programming problems because of their convexity. One of the most common methods among them is the Branch-and-Bound strategy [76]. In order to conduct an intelligent and repeated search of all feasible solutions, B&B breaks down the entire polyhedron region into smaller subsets. An upper and lower bound is determined for each subset, and it excludes any subset exceeding the cost/reward ratio of the feasible solution. Once a feasible solution is discovered, whose cost/reward is within the bounds of any subset, this partitioning process ends. It is important to note that B&B is also referred to as implicit enumeration, separation and evaluation, or divide and conquer.

3.1.1 Network Modelling Problem

A network optimization problem can be formulated using a variety of approaches in linear programming. In this thesis, network models are designed using the node-link formulation. A network is formulated by assuming that both the demands and the links are generally directed, as well as considering total link flow at each link. With the exception of the source and destination nodes, the total traffic leaving a particular node plus the total traffic entering that node must equal zero [77]. Between each of the two end nodes of a considered demand, intermediate or transition nodes exist

through which traffic enters and leaves the node via its incoming and outgoing links. The total flow at each intermediate node of the considered demand equals the total flow at all the outgoing links for each intermediate node. This is known as the flow conservation law. In the event that the consider demand is generated at the node (source node), then the demand volume equals the total outgoing flow; otherwise, when the node is a sink node, the demand volume equals the total incoming flow.

To illustrate the node-link formulation, consider a three node network shown in Fig.3.1, having three demand volumes, i.e., Λ_{13}, which is from node 1 to node 3, Λ_{12} which is from node 1 to node 2 and Λ_{23} which is from node 2 to node 3. By implementing the flow conservation at node 1 for demand volume and employing the rule that anything entering into the node is considered as negative and anything leaving that node is considered as positive, gives us the flow equation as

$$-\Lambda_{13} + \Lambda_{13,13} + \Lambda_{13,12} - \Lambda_{13,31} - \Lambda_{13,21} = 0, \tag{3.8}$$

where:

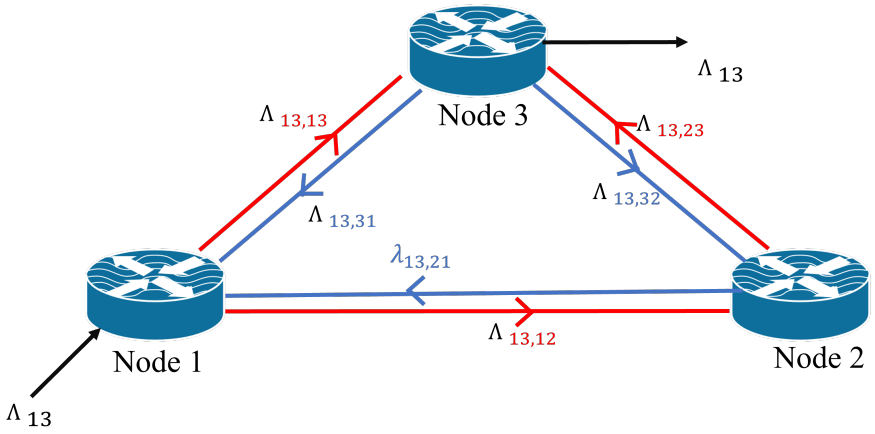
$\Lambda_{13,13}$ is the volume of the demand with node 1 as a source node and node 3 as a destination node, which is flowing through the link from node 1 to node 3;

$\Lambda_{13,12}$ is the volume of the demand with node 1 as a source node and node 3 as a destination node, which is flowing through the link from node 1 to node 2;

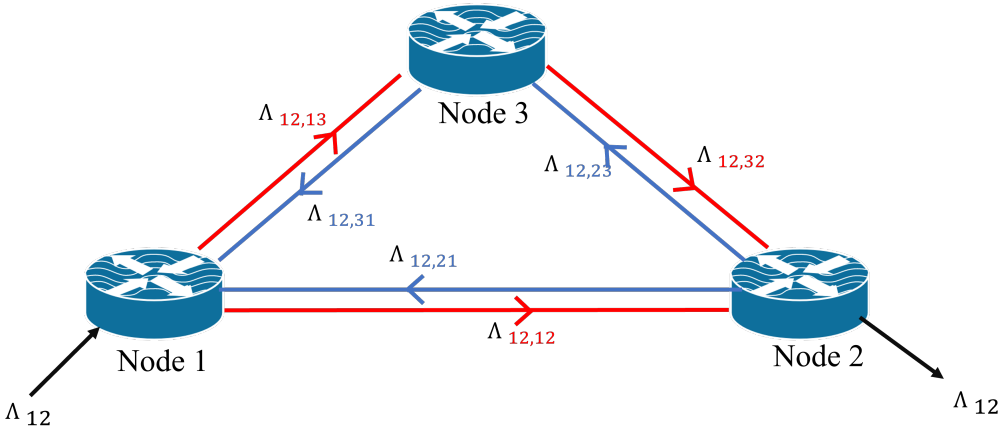
$\Lambda_{13,31}$ is the volume of the demand with node 1 as a source node and node 3 as a destination node, which is flowing through the link from node 3 to node 1;

$\Lambda_{13,21}$ is the volume of the demand with node 1 as a source node and node 3 as a destination node, which is flowing through the link from node 1 to node 3;

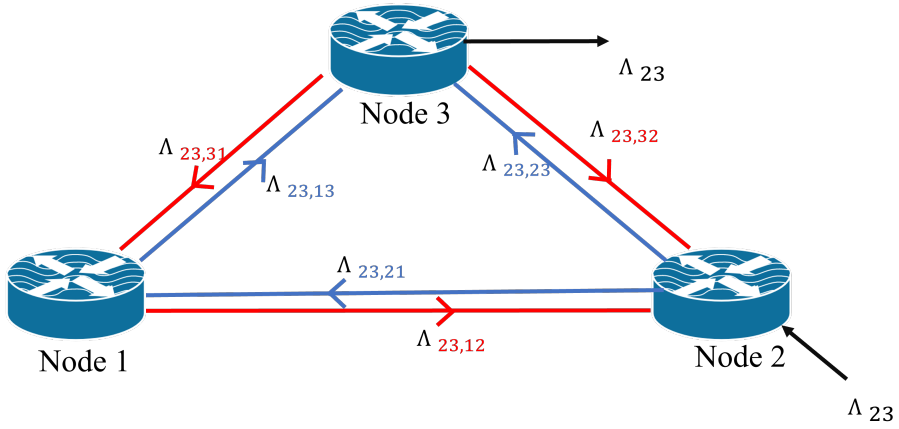
Λ_{13} is the total volume of the demand with node 1 as a source node and node 3 as a destination node.



(a) Demand flow from node 1 to node 3



(b) Demand flow from node 1 to node 2



(c) Demand flow from node 2 to node 3

Figure 3.1: Example of demand flows in three nodes network

The following equations are obtained by employing flow conservation at nodes 2 and 3 for the demand volume Λ_{13}, with the same assumptions as before. Thus at node 2

$$-\Lambda_{13,12} - \Lambda_{13,32} + \Lambda_{13,23} + \Lambda_{13,21} = 0, \tag{3.9}$$

also, at node 3

$$-\Lambda_{13,13} - \Lambda_{13,23} + \Lambda_{13} + \Lambda_{13,31} + \Lambda_{13,32} = 0. \tag{3.10}$$

For the demand volume λ_{13}, it is essential to mention that node 1, 2 and 3 act as a source node, an intermediate node and a sink node, respectively. Thus

$$\Lambda_{13,31} = \Lambda_{13,32} = \Lambda_{13,21} = 0, \tag{3.11}$$

thus, the former flow conservation equations for the demand volume λ_{13}, for each of the nodes 1, 2, and 3 will become:

$$\begin{aligned} \Lambda_{13,12} + \Lambda_{13,13} + \quad &= \Lambda_{13} \\ -\Lambda_{13,12} \quad + \Lambda_{13,23} &= 0 \\ -\Lambda_{13,13} - \Lambda_{13,23} &= -\Lambda_{13}. \end{aligned} \tag{3.12}$$

By the same analogy, for the demand volume λ_{12}, the system equations can be written as

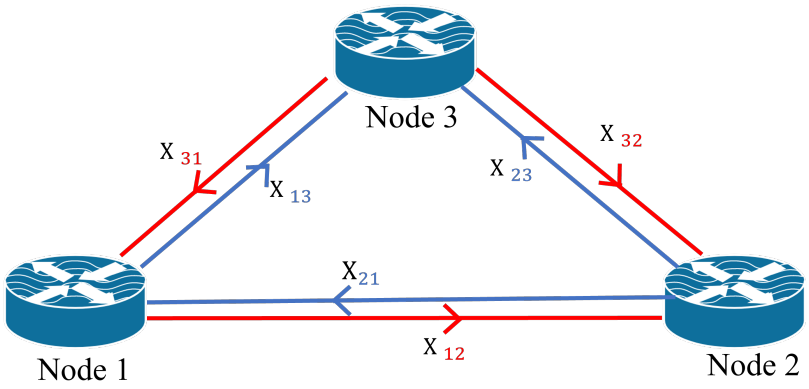


Figure 3.2: Link capacity example in three nodes network

$$\begin{aligned} \Lambda_{12,12} + \Lambda_{12,13} &= \Lambda_{12} \\ -\Lambda_{12,13} + \Lambda_{13,32} &= 0 \\ -\Lambda_{12,12} - \Lambda_{12,32} &= -\Lambda_{12}. \end{aligned} \tag{3.13}$$

Similarly for the demand volume, Λ_{23}

$$\begin{aligned} \Lambda_{23,21} + \Lambda_{23,23} &= \Lambda_{23} \\ \Lambda_{23,13} - \Lambda_{23,21} &= 0 \\ -\Lambda_{23,13} - \Lambda_{23,23} &= -\Lambda_{23}. \end{aligned} \tag{3.14}$$

Now by assuming that each of the undirected link can be represented by two arcs and at some instant, the demand can flow in one of these arcs only. Therefore, some extra constraints can be added into the model to consider the capacity of each link as shown in Fig. 3.2Here, if χ denotes the capacity of a link from node 1 to 2, the corresponding capacity constraint can be written as

$$\Lambda_{12,12} + \Lambda_{13,12} \leq \chi_{12}. \tag{3.15}$$

Now by using the same analogy for all the other links present in the network and combining all of them gives us the system model as

$$O = \Lambda_{12,12} + \Lambda_{12,13} + \Lambda_{12,32} + \Lambda_{13,12} + \Lambda_{13,13} + \Lambda_{13,23} + \Lambda_{23,21} + \Lambda_{23,13} + \Lambda_{23,21} + \Lambda_{23,23} \tag{3.16}$$

In Eq.(3.16), O denotes the objective function of the model, which we want to minimize. In particular, here, the cost occurs due to the routing of the traffic in the whole network.

subject to

$$\begin{array}{cccccccccccc}
\Lambda_{12,12} & +\Lambda_{12,13} & & & & & & & & & = & \Lambda_{12} \\
 & -\Lambda_{12,13} & +\Lambda_{12,32} & & & & & & & & = & 0 \\
-\Lambda_{12,12} & & -\Lambda_{12,32} & & & & & & & & = & -\Lambda_{12} \\
 & & & \Lambda_{13,12} & +\Lambda_{13,13} & & & & & & = & \Lambda_{13} \\
 & & & -\Lambda_{13,12} & & +\Lambda_{13,23} & & & & & = & 0 \\
 & & & & -\Lambda_{13,13} & -\Lambda_{13,23} & & & & & = & -\Lambda_{13} \\
 & & & & & & \Lambda_{23,21} & & +\Lambda_{23,23} & = & \Lambda_{23} \\
 & & & & & & -\Lambda_{23,21} & +\Lambda_{23,13} & & = & 0 \\
 & & & & & & & -\Lambda_{23,13} & +\Lambda_{23,23} & = & -\Lambda_{23} \\
\Lambda_{12,12} & & & +\Lambda_{13,12} & & & & & & \leq & \chi_{12} \\
 & & & & & & \Lambda_{23,21} & & & \leq & \chi_{21} \\
 & \Lambda_{12,13} & & & +\Lambda_{13,13} & & & +\Lambda_{23,13} & & \leq & \chi_{13} \\
 & & & & & \Lambda_{13,23} & & & +\Lambda_{23,23} & \leq & \chi_{23} \\
 & & \Lambda_{12,32} & & & & & & & \leq & \chi_{32}
\end{array}$$

this only holds for all non negative Λ.

To model and solve the optimisation problems, several programming languages are utilised nowadays. Over the years, various modelling and optimisation languages

have been developed by modelling language providers, like Optimised Programming Language (OPL) and the (AIMMS). In [78], a detailed description of some of the modelling languages with their providers is presented. We have chosen "A Mathematical Programming Language" (AMPL) to model our optimisation problems due to the likeliness of its statements with the algebraic notations and generality of its syntax. [79]. The user of the AMPL program enters their codes and data into the software, where these codes are transformed into an intermediate file, that a solver can understand. A mathematical program that works by reading the intermediate file and implementing a suitable algorithm is a solver. There are various choices of solvers like Coin-or Branch and Cut (CBC) [80], GNU Linear Programming Kit (GLPK) [81], and the generally accepted IBM CPLEX[82]. In this work, CPLEX is adopted as the optimisation problem solver.

3.2 Non-Linear objective function or constraints

Most of the optimization problems in wireless communication can be expressed in terms of linear constraints. Flow rate or limits on data rate use, for example, are generally linear functions. However, many objective functions tends to be non-linear. Design problems for which the objective contains some non-linear constraints which are function of some non-linear functions become non-linear, e.g.,

- In backhaul optimisation problem, the data flow rate/capacity C flowing through any node (BS) is related to the SNR (γ) as

$$C = B \log_2(1 + \gamma), \tag{3.17}$$

  the $\log_2$ function in Eq.(3.17) is non-linear.

- In MEC paradigm, one of the objective function is to minimise transmission

delay, which in turns depends on a capacity$C$ of the channel.

$$t_{trans} = \frac{L}{C}, \tag{3.18}$$

as L is a constant so $\frac{1}{C}$, make this objective function non-linear.

Various methods exist for solving non-linear problems. One of these is to divide the non-linear functions into different linear sections (piecewise Linearization). The benefit of this technique is that we then have a linear problem to which any MILP algorithm, such as CPLEX, can be used. The detailed description is as follows.

3.2.1 Piecewise Linearization technique

This technique can be used to linearise both concave and convex functions. First we applied this technique to linearise, Eq.(3.17), which is a concave function, by dividing the SNR between its maximum and minimum value given by the model into small regions (or segments) as

$$\bar{\gamma}_r = r \left(\frac{\gamma_{(max)} - \gamma_{(min)}}{NR} + \gamma_{(min)} \right), \forall r \in NR, \tag{3.19}$$

where NR is the total number of segments. Now the capacity $\bar{C}_r$ in each segment against $\bar{\gamma}_r$, can be given as

$$\bar{C}_r = B \log_2(1 + \bar{\gamma}_r) \quad \forall r \in NR. \tag{3.20}$$

The plot of $\bar{C}_r$ against $\bar{\gamma}_r$ is shown in Fig.3.3 segments value of capacity. Now as the given capacity γ can lie in any region between $\bar{\gamma}_{r-1} - \bar{\gamma}_r$, so we introduce a set

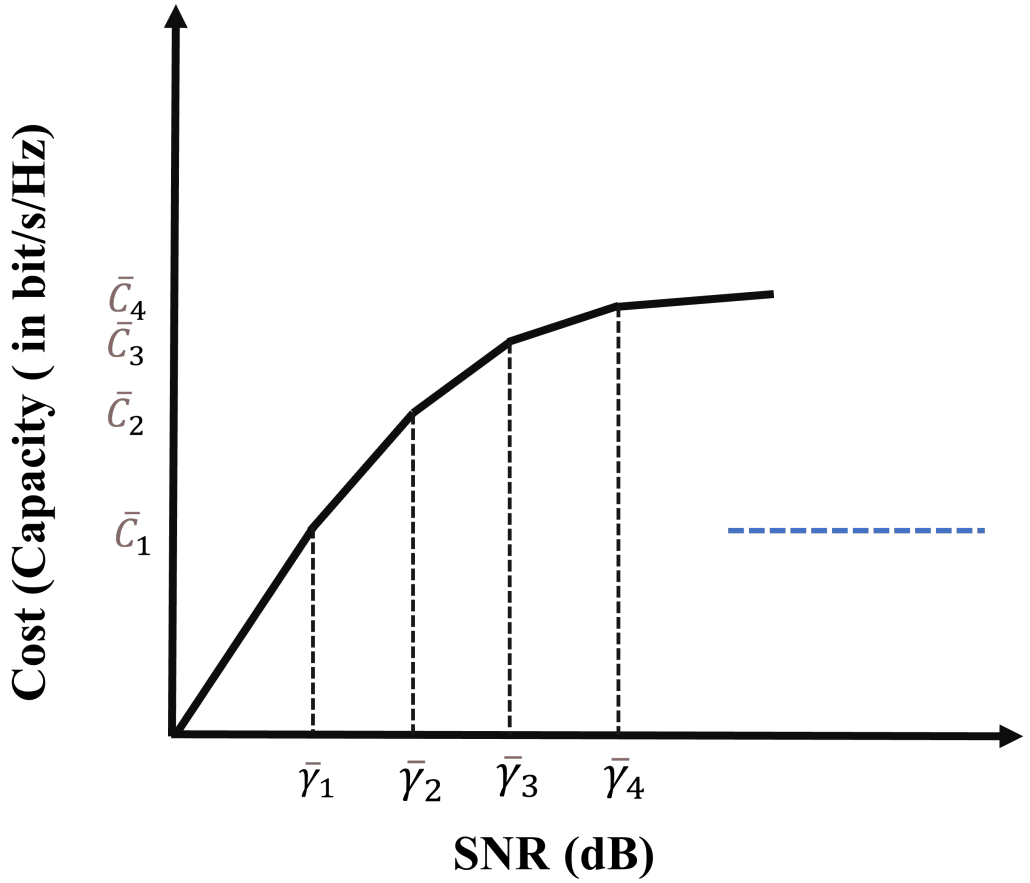


Figure 3.3: Linearization of Capacity-SNR Eq.(3.17)

γ_r having each element equals to γ .

$$\gamma = \sum_{r=1}^{NR} \gamma_r, \quad \forall r. \tag{3.21}$$

Now for each segment r, we check whether the γ_r lies in that segment, by introducing following constraints

$$\bar{\gamma}_{r-1} Z_r \leq \gamma_r \leq \bar{\gamma}_r Z_r, \quad \forall r, \tag{3.22}$$

where, Z_r is a binary variable which is defined as

$$\sum_{r=1}^{NR} Z_r = 1, \qquad \forall r, \tag{3.23}$$

The above set of constraints (Eq.(3.21)- Eq.(3.23) ensure that the variable Z_r should only be one for the value of r, where the given SNR γ_r exactly lies in the range specified by $\bar{\gamma}_{r-1} - \bar{\gamma}_r$. Now we will find the capacity, C by summing over all the segments r

$$C = \sum_{r=1}^{NR} \left[\bar{C}_{r-1} Z_r + (\gamma_r - \bar{\gamma}_{r-1} Z_r) \frac{\bar{C}_r - \bar{C}_{r-1}}{\bar{\gamma}_r - \bar{\gamma}_{r-1}} \right], \qquad \forall r, \tag{3.24}$$

as Z_r is only 1 for a single value of r, so it will find the approximated value of capacity, C against that SNR, γ in that particular region r.

3.3 Genetic Algorithm

Genetic algorithms are a family of optimization algorithms which are applied to find the optimal solution (s) to a computational problem by maximizing or minimizing a particular function. These algorithms come under the category of evolutionary computation [83]. They mimic the biological processes of natural selection and reproduction to determine "test solutions" [84]. There are many processes of genetic algorithms which are random, as in evolution. However, one can use this s optimization technique to establish the level of control and randomization [84]. Therefore, these algorithms are much more robust and effective than the exhaustive search and the random search algorithms [83], because no extra information relating to the given problem is needed. This unique characteristic of a genetic algorithm enables it to find the solutions to those problems, which cannot be handled by other optimiza-

tion techniques owing to the absence of continuity, linearity, derivatives or different other components.

3.3.1 Components, Structure, & Terminology

Because genetic algorithms simulate a biological process, much of the relevant terminology comes from biology. However, the entities of this terminology in genetic algorithms are much simpler than their biological equivalents [85]. The standard essential components of almost all genetic algorithms are:

- An initial population of chromosomes.
- Fitness function for optimisation.
- Selection of chromosomes that will reproduce.
- Crossover for producing next-generation of chromosomes
- Mutation of chromosomes in the new generation

3.4 Conclusions

In this chapter, we have provided background information related to optimization problem modelling that will assist us in understanding and solving the problem. We discussed linear programming and the AMPL model in this chapter, which was used to solve optimization problems. The node-link formulation was used to model a small wireless network. Throughout the remainder of this chapter we will describe genetic algorithms that can be employed to test the results presented.

Chapter 4

Optimisation of IRS in 5G and Beyond Backhaul Networks

In this chapter we present a general framework for developing policies that optimize user association, the routing of user traffic across backhaul channels to the core network, and which base stations or backhaul channels should be switched off to minimize energy usage. Furthermore, the problem of optimization is extended to examine the impact of the power consumption of active elements within the IRS on the overall backhaul power. Further, what impact will the availability of IRS channels for blocked mmWave channels have on the number of users served?

4.1 Introduction

With 7.1 billion people connected to mobile services by 2021, the demand for connectivity continues to rise. In consequence, by 2025, there will be 7.49 billion unique

mobile subscribers, which is equivalent to 90% of the world's population [86]. The deployment of small cells (SCs) can provide limited capacity on demand to meet the demands of a large number of users in 5G. In addition, umbrella eNodeBs (also called macro base stations) are deployed to provide coverage to small Base Stations (BS)[12]. The backhaul connects the SBs and MBS to the core network; typically, fibre optic cable is used as a media for backhaul. The provision of optical fibre connections to each BS of a SC, however, could lead to increased network operator deployment and maintenance costs. As a result, researchers have proposed wireless backhaul (BH) solutions, which route user traffic towards the core network of an operator via a wireless interface to provide the required capacity at a reasonable cost. Due to the large amount of spectrum available in the 60, 70, or 80 GHz range, millimeter-wave technology is a promising candidate for backhaul solutions. Even though the omnidirectional free path loss increases at those frequencies due to Friis Law [13], it can be compensated by beamforming, which can result in high antenna gains [14]. However, shadowing is one of the greatest threats for mmWave signals, causing outages, intermittent connectivity, and rapidly varying channel conditions [15]. Thus, the utilization of mmWave in the backhaul link has its limitations, i.e., as the base stations are deployed, there is a possibility of the mmWave link being blocked, causing the rerouting of the backhaul traffic.

New emerging hardware technology known as the Intelligent Reflecting surface, also called software-controlled metasurfaces, has just arrived; it reflects the impinging plane wave in the shape of a beam, thus acting as a reconfigurable reflectarray by employing Micro-Electrical-Mechanical Systems (MEMS), PIN or varactor diodes [37], [17] and [18]. In this work, we deployed IRS in the backhaul links along the mmWave channel to overcome the blockage in mmWave links as shown in Fig. 4.1. To the best of our knowledge, most of the researchers focused on implementing IRS

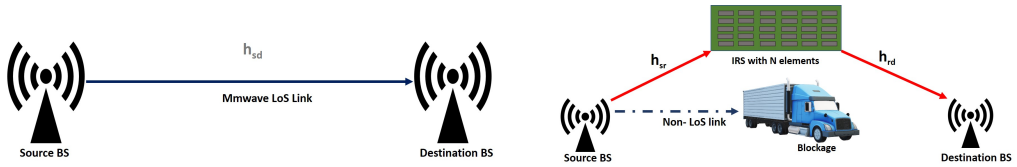
on the access side. Still, no one had explored the feasibility of using the IRS in the backhaul transmission, i.e., from a source base station to a destination base station by adapting the propagation environment along the mmWave channel. Therefore, the study focuses on (i) saving in static BH power due to the IRS channels along the mmWave channels when several mmWave channels get entirely blocked simultaneously in a grid topology (ii) saving in dynamic and overall BH power due to the partial blockage of mmWave channels in a grid topology. Moreover, by plotting the dynamic backhaul power against the number of IRS elements, the optimum value of IRS elements can be achieved within the studies range of values for those IRS elements and the input parameters of our problem (iii) optimization of a number of users by deploying a minimum number of IRS.

The remainder of this chapter is organized as follows. In section II, expressions for the capacity and power in the access and backhaul networks are derived. In Section III, a MILP objective function is introduced with all the constraints. In Section IV, all the results are discussed. Finally, Section V concludes the chapter.

Notation: For ease of reference, Table 4.1 , 4.2 , 4.3 and 4.4 summarizes the main parameters, variables, indices and input parameters respectively, which will be used throughout this chapter. Furthermore, vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. Finally, upper-case calligraphic letters represent the sets.

4.2 Optimization of Intelligent Reflecting Surfaces in 5G and Beyond Backhaul Networks

Several studies have investigated enhancing the link efficiency of mmWave communication in the presence of static blockages. However, dynamic blockages e.g. parked

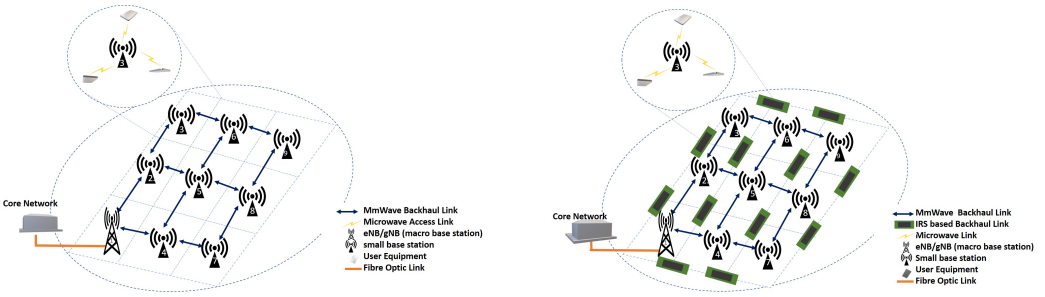


(a) mmWave communication between two BSs **(b)** A parked truck between two BSs blocking LoS mmWave path in an urban environment, so an IRS aided transmission is used for the backhaul traffic.

Figure 4.1: An IRS is deployed along the mmWave channel for backhaul transmission

cars/truck see Fig.4.1(b), causes user traffic to be interrupted in the backhaul link. To alleviate this problem, most studies propose re-routing of user traffic through other backhaul link or may be served through completely different access link if required, such a solution, unfortunately incurs penalty in terms complex re-routing algorithms.

Our main focus is on optimization of future 5G networks with IRS supported mmWave mesh backhaul model. We consider a HetNet scenario, where both MBS and SBS are deployed to serve the UEs. We focus on a single MBS fixed at the origin, and a set of BSs is deployed forming a grid (3x3) as shown in the Fig. 4.2(a). The BSs are connected to each other as well as the MBS through mmWave links, thus forming a mesh backhaul network. MBS has a direct fiber-optic connection to the core network, thus all the user traffic pass through the MBS. As discussed earlier, due to the inevitable blockage between BSs, mmWave communication may not be accomplished in the backhaul link, hence, IRS are deployed along with the mmWave links so that the traffic can still be routed through the reflecting IRSs. We assume that the users are uniformly scattered around the service area. Each user u receives at a rate captured by R_u, in Mbps. We focus on downlink transmissions, with the MBS serving as a source node and the UEs are the sink nodes.



(a) mmWave backhaul transmission **(b)** IRS supported backhaul transmission

Figure 4.2: System Model for backhaul transmission

For the access network, we assume a set of microwave links between the user equipment UE and the serving base stations. In our analysis, only the small base station can provide the access links, so MBS does not provide access links to users. We use a constant power allocation, i.e., the maximum transmitted power of each cell is divided equally into its physical resource blocks (PRBs). Assuming each BS has a total bandwidth, B^{AN} in the access side, while, B^{BH} is the bandwidth for each backhaul link in the backhaul side.

According to the specific user data requirement, a base station can assign any number of its resource blocks to that particular user but their is a limit on number of maximum resource blocks a base station can assign according to the demand of user. We assume that if there is no UE anchored to a BS, the RF transceivers at the BS are switched off to save energy.

For the backhaul traffic, the mmWave transceivers installed on the BS will be switched off if there is no incoming or outgoing traffic being serviced by the BS. The distance between adjacent BSs is set to 100 meters [10], while the IRS are deployed between the two adjacent BS in such a way that the distance between IRS and any of the BS is always less than the distance between adjacent BS. For the Backhaul traffic, we consider mmWave communication from a source base station with a single antenna transmitter to a single-antenna destination base station, as

we leave the multi antenna analysis for future work. A V-band (60 GHz) is chosen for mmWave backhaul transmission although it has high path loss and atmospheric attenuation which can be compensated by highly directional antennas. Moreover, as the BS are deployed in a range of 100-200 meters causing the atmospheric attenuation lies within the acceptable range. In our setup, an IRS with N elements is configured to reflect the signal from transmitting base station towards the destination base station, as illustrated in Fig.4.2(b). Now by using the [87, eq(12)], the capacity of IRS supported SISO system is given by

$$R_{\mathrm{IRS}}(N) = \log_2 \left(1 + \frac{P \left(N \alpha \sqrt{\beta_{\mathrm{IRS}}} \right)^2}{\sigma^2} \right), \tag{4.1}$$

where, P is the transmit power of the transmitting base station, $\alpha \in (0,1]$ denotes the fixed amplitude reflection coefficient and $\sqrt{\beta_{IRS}}$ is the product of β_{sr} and β_{rd}, representing the LOS channel gains between the source BS to the IRS and IRS to the destination BS respectively and σ^2 is the power of white Gaussian noise. Moreover, unlike the work in [87], we don't consider the direct path with gain β_{sd} in the IRS channel, this is because we assume the direct path is either blocked or very weak, and Tx and Rx base stations may use the antenna on a mechanical base to align the signal between them. As the channel capacity depends upon the transmitter power, number of reflecting elements, gains and the noise associated with the backhaul channel, we, therefore, study the impact of increasing the number of reflecting elements in our scheme. Without losing any generality, and for the sake of simplicity, we exclude fibre optic links from our study by assuming very high capacity links (e.g., 40, or 100 Gbps) with negligible power consumption [88]. Next, we find the expressions for the capacity and power in the access and the backhaul networks, which will be used in our MILP model.

Table 4.1: ENERGY-EFFICIENT MILP MODEL PARAMETERS

Parameters	Comments	Parameters	Comments
MBS	Set of macro base station nodes	P_{bs}^{maxBH}	Maximum BH power of a base station
SBS	Set of small base station nodes	B^{BH}	mmWave bandwidth in Backhaul link
BS	Set of all base station nodes	$P_{0_{i,j}}^{BH}$	Minimum non-zero output power of the
UE	Set of user nodes		BH transceiver $\forall i,j \in BS, i \neq j$
PRB	Set of physical resource blocks	$d_{i,j}^{BH}$	Distance between two base stations i and j
CH	Set of all the channels (denotes CH1 and CH2 the		in Backhaul, $\forall i,j \in BS, i \neq j$
	direct mmWave link and IRS link respectively)	Δ^{BH}	Slope of the BH load-dependent power consumption
N_m^n	Set of neighbours of node m in the network ,$\forall m \in N$	f^{BH}	Frequency at which Backhaul link is operating
$d_{bs,\mathrm{u}}^{AN}$	Distance of user from base station in meters	$\upsilon_{i,j,ch}^{BH}$	Path loss in dBs when there is a link between
$\epsilon_{bs,u}^{AN}$	Range parameter in meters i.e.,		two base stations over each channel at backhaul link
	if a user lies in certain range of a BS	ϑ_{Tx}^{BH}	Transmitted antenna Gain in backhaul link
ϖ_{rb}	Bandwidth of each PRB	ϑ_{Rx}^{BH}	Received antenna Gain in backhaul link
$\upsilon_{bs,u}^{AN}$	Path loss in dBs of each user from the base station	χ_{Tx}^{BH}	Losses at the transmitter side in backhaul link
	in the access side	$\chi_{\mathrm{Rx}}^{\mathrm{BH}}$	Losses at the receiver side in backhaul link
$P_{0_{bs}}^{AN}$	Minimum non-zero output power of the transceiver	NF^{BH}	Noise Figure for backhaul link
	in the access side	N_{th}^{BH}	Thermal noise in backhaul
Δ_{bs}^{AN}	Slope of the load-dependent power consumption	LM^{BH}	Link Margin of the backhaul link
	in the AN link	β^{BH}	Blockage weightage
$\Psi_{bs,u}^{AN}$	Linear path gain of each user from the base station	$\Gamma_{i,j}^{BH}$	Binary random parameter for backhaul links
	in the access side		$\forall i,j \in BS, i \neq j$
$\kappa_{bs,u}^{AN}$	Large scale fading in dBs of each user from the	N	Number of active elements in an IRS
	base station in the access side	$\wp$	power of single active element in IRS
$\iota_{bs,u}^{AN}$	Linear large scale fading of each user from the	ϵ_{vap+O_2}	Signal attenuation due to vapour water and oxygen
	base station in the access side		at the Backhaul link
$\varsigma_{\mathrm{bs,u,rb}}^{\mathrm{AN}}$	Linear small scale fading of each resource block consumed	ϵ_{Rain}	Signal attenuation due to rain at the Backhaul link
	by each user from the base station in the access side	$\upsilon_{i,j,ch}^{BH}$	Path loss in dBs of each backhaul link
$G_{bs,u,rb}^{AN}$	Gain of each resource block consumed by each user		$\forall i,j \in BS, i \neq j$ at each channel
	from the base station in the access side	$\rho_{i,j,ch}^{BH}$	Total gain of each backhaul link $\forall i,j \in BS, i \neq j$
B^{AN}	LTE bandwidth in the access side		at each channel
R_{u}^{AN}	Demand of each user in Mbps	$L_{i,j,ch}^{BH}$	Linear path loss of each backhaul link
η^{AN}	Efficiency of amplifier in the access side		$\forall i,j \in BS, i \neq j$ at each channel
P_{bs}^{maxAN}	Maximum power of a base station in the access side	w_1	Weightage parameter for the IRS power
P_{bsrb}^{AN}	Transmitted power of each base station for each PRB	w_2	Weightage parameter for the users
σ^{AN}	Noise variance in access side	M	Very large +ve number
$\gamma_{bs,u,rb}^{AN}$	SINR of each PRB associated with each user u assigned		
	from the base station		

4.2.0.1 Access Network

In the access network, only those users are able to associate to a BS if those users are within a certain range, $\epsilon_{bs,u}^{AN}$ (in meters) of the BS. The total path gain $G_{bs,u,rb}^{AN}$ achieved by each user u from the base station bs for each resource block rb is the product of the path gain $\Psi_{bs,u}^{AN}$, large scale fading $\iota_{bs,u}^{AN}$, and small scale fading $\varsigma_{\mathrm{bs,u,rb}}^{\mathrm{AN}}$, and in linear scale is given as

$$G_{bs,u,rb}^{AN} = \Psi_{bs,u}^{AN} \iota_{bs,u}^{AN} \varsigma_{\mathrm{bs,u,rb}}^{\mathrm{AN}}, \forall bs \in BS, u \in UE, rb \in PRB. \tag{4.2}$$

The path loss in the access side which is the function of the distance $d_{bs,u}$, [89] in dBs, is given by

$$\upsilon_{bs,u}^{AN} = 38 + 30 \log_{10}(d_{bs,u}), \quad \forall\, bs \in BS, u \in UE, \tag{4.3}$$

$$\upsilon_{bs,u}^{AN} = 128.1 + 37.6 \log_{10}(d_{bs,u}), \quad \forall\, bs \in MBS, u \in UE,$$

where Eq (4.3) corresponds to the path loss equation for BSs. The linear path gain is given by

$$\Psi_{bs,u}^{AN} = 10^{\left(\frac{-\upsilon_{bs,u}^{AN}}{10}\right)}, \forall\, bs \in BS, u \in UE. \tag{4.4}$$

Large scale fading $\kappa_{bs,u}^{AN}$(dB) of each user from the BS (pico cell) is incorporated into the path gain Eq.(4.2), by the addition of a zero-mean Gaussian random variable [90], with standard deviation 6 (for SBS) and 10 (for MBS) is given as

$$\kappa_{bs,u}^{AN} = \mathcal{N}(0,6), \quad \forall\, bs \in SBS, u \in UE, \tag{4.5}$$
$$\kappa_{bs,u}^{AN} = \mathcal{N}(0,10), \quad \forall\, bs \in mBS, u \in UE.$$

Table 4.2: ENERGY-EFFICIENT MILP MODEL VARIABLES‘

Variables	Comments
$\sigma_{bs,u,rb}^{AN}$	Binary indicator set to 1 if a PRB is assigned by a base station bs to the user u
$\varrho_{bs,u}^{AN}$	Binary indicator set to 1 if the user u is served by the base station bs
P_{bs}^{AN}	RF output power of each base station in the access link
$P_{T_{bs}}^{AN}$	Total power consumed by each base station in the access link
$C_{bs,u}^{AN}$	Capacity of the access link between a base station bs and the user u
φ_{bs}^{AN}	Binary indicator set to 1 if the base station is associated with any user.
φ_{u}^{AN}	Binary indicator set to 1 the demand $R_{\mathrm{u}}^{AN}$of a user u is satisfied or otherwise 0
$p_{i,j,ch}^{BH}$	Power consumed by each backhaul link on each channel
$\sigma_{i,j,ch}^{BH}$	Binary indicator set to 1 if any traffic is flowing through backhaul link otherwise 0
$\xi_{s,u,i,j}$	Traffic flow variable from the source node s to the destination node u passing through the intermediate links i and j $\quad \forall s \in MBS, u \in UE, i, j \in N : i \neq j$ and j in $N_m[i]$
$P_{bs,ch}^{BH}$	Power consumed by each base station for backhaul traffic on each channel
$R_{i,j,ch}^{BH}$	Rate of each backhaul link on each Channel
$\Upsilon_{i,j,ch}^{BH}$	SNR of each Backhaul link on each channel

The corresponding linear large scale fading $\iota_{bs,u}^{AN}$ is given by

$$\iota_{bs,u}^{AN} = 10^{\left(\frac{\kappa_{bs,u}^{AN}}{10}\right)}, \quad \forall\, bs \in BS, u \in UE. \tag{4.6}$$

Small scale fading is associated with the resource blocks having different frequency bands, so a Rayleigh fading channel is used to incorporate the effect of small scale fading $\varsigma_{\mathrm{bs,u,rb}}^{\mathrm{AN}}$, in the linear scale. In Rayleigh fading, the received power will follow the negative exponential distribution, which can be generated by summing the squares of two Gaussian random numbers, each with zero mean and 0.5 variance, and is given by

$$\varsigma_{\mathrm{bs,u,rb}}^{\mathrm{AN}} = \left[\mathcal{N}(0, \frac{1}{\sqrt{2}})\right]^2 + \left[\mathcal{N}(0, \frac{1}{\sqrt{2}})\right]^2, \tag{4.7}$$
$$\forall\, bs \in BS, u \in UE, rb \in PRB.$$

The maximum transmitted power of each base station ρ_{bs}^{maxAN} in the access side is given in Table (4.4). As we assume that BS transmits constant power to each allocated resource block, thus the required transmit power P_{bs}^{AN} (in milliwatts) of each PRB in any BS is

$$P_{bsrb}^{AN} = \frac{P_{bs}^{\text{maxAN}}}{\text{card}\,(PRB)},\ \forall bs \in BS, \tag{4.8}$$

in Eq.(4.8), card (PRB) denotes the cardinality (total number of PRBs a base station can handle) of set PRB. Thus, the SINR in the access side is given as

$$\Upsilon_{bs,u,rb}^{AN} = \frac{P_{bsrb}^{AN} G_{bs,u,rb}^{AN}}{\left(\eta^{AN} \sum_{l=0,l\neq 0}^{BS} P G_{l,u,rb}^{AN} \epsilon_{l,u}^{AN}\right) + \sigma^{AN}}\ , \tag{4.9}$$
$$\forall bs \in BS, u \in U, rb \in R,$$

where σ^{AN} is the noise variance in the access side, [90], and calculated in Table (4.4)and η^{AN} is the scaling factor meaning that only 50% of PRBs can interfere at any time with that particular PRB, thus the corresponding Channel Capacity at the access side is given as

$$C_{bs,u}^{AN} = \sum_{rb\in PRB} \sigma_{bs,u,rb}^{AN} \varpi_{rb} \log_2(1 + \Upsilon_{bs,u,rb}^{AN}), \forall bs \in BS, u \in UE. \tag{4.10}$$

Now, the load dependent RF power of each base station in the access side is given by multiplying the total number of allocated PRBs, $\sigma_{bs,u,rb}^{AN}$, with the power of each PRB, P_{bs}^{AN} as

$$P_{bs}^{AN} = \sum_{u\in U} \sum_{rb\in PRB} \sigma_{bs,u,rb}^{AN} P_{bsrb}^{AN}, \qquad \forall bs \in BS. \tag{4.11}$$

Table 4.3: ENERGY-EFFICIENT MILP MODEL INDICES

Indices	Comments
bs	Index of the base stations
i, j	Indices of two nodes in the model, on which the user traffic transverse form one node to another
u	Index of user nodes
rb	Index of the physical resource block

In Eq.(4.10), ϖ_{rb} is the bandwidth of a single resource block and $\sigma^{AN}_{bs,u,rb}$ is a binary indicator that takes 1 when the rbth resource block from a base station bs is assigned to a user u, or it takes 0 otherwise. In order to find the total power consumption by each base station in the access link, linear approximation in [91] is used, where the variable RF output power, P^{AN}_{bs} and the power consumed by each base station is related as

$$P^{AN}_{T_{bs}} = \left(\varphi^{AN}_{bs} P^{AN}_{0_{bs}} + \Delta^{AN}_{bs} P^{AN}_{bs}\right), \forall bs \in BS, \tag{4.12}$$

where, φ^{AN}_{bs} is binary variable which is 1 if the demand of the user is anchored by a base station by assigning single/multiple resource blocks to that user or otherwise 0. Eq.(4.10) and Eq.(4.12) give the channel capacity, $C^{AN}_{bs,u}$ and the total power consumed, $P^{AN}_{T_{bs}}$ by each small base station in the access side. Similarly, in the next section, we will define the expressions for capacity and the total power consumed by each small base station in the backhaul network.

4.2.0.2 Backhaul Network

Let $\upsilon^{BH}_{i,j,ch}$ denotes the path loss in dB, between two base stations i and j over channel, ch, whereas $ch = 1$, corresponds to the mmWave channel, while $ch = 2$, refers to the IRS channel. Now, by using [91, Eqs.(6)-(8)], $\upsilon^{BH}_{i,j,ch}$, for mmWave channel is given

as

$$\begin{aligned} \upsilon_{i,j,ch}^{BH} &= 20\log_{10}(f^{BH}) + 32.45 + 20\log_{10}(d_{i,j}^{BH}/1000) + \\ &(d_{i,j}^{BH}/1000) * (\epsilon_{vap+O_2} + \epsilon_{Rain}) + \Gamma_{i,j}^{BH}\beta^{BH}, \\ &\forall i,j \in BS, ch = 1:j \text{ in } N_{m_i} \text{ and } j \neq i. \end{aligned} \tag{4.13}$$

In Eq.(4.13), f^{BH} denotes the frequency of backhaul link and $d_{i,j}^{BH}$ is the distance between two base stations. ϵ_{vap+O_2} ans ϵ_{Rain} represent the signal attenuation due to atmospheric gases and rain respectively. The addition of β^{BH} in Eq.(4.13), is to incorporate the effect of blockage in the mmWave channel, while $\Gamma_{i,j}^{BH}$, denotes the random binary variable. For the IRS supported BH channel, the path loss, $\upsilon_{i,j,ch}^{BH}$ in dB is given as

$$\begin{aligned} \upsilon_{i,j,ch}^{BH} &= 20\log_{10}(f^{BH}) + 32.45 + 20\log_{10}(d_{i,j}^{HS}/1000) + \\ &(d_{i,j}^{HS}/1000)(\epsilon_{vap+O_2} + \epsilon_{Rain}) + 20\log_{10}(f^{BH}) + \\ &32.45 + 20\log_{10}(d_{i,j}^{HS}/1000) + \\ &(d_{i,j}^{HS}/1000)(\epsilon_{vap+O_2} + \epsilon_{Rain}) - 20\log_{10}(N), \\ &\forall i,j \in BS, ch = 2:j \text{ in } N_{m_i} \text{ and } j \neq i. \end{aligned} \tag{4.14}$$

The pathloss, $\upsilon_{i,j,ch}^{BH}$ in Eq.(4.13) and Eq.(4.14), is used to specify a parameter $\rho_{i,j,ch}$, which can be determined by subtracting from the total losses the gains of the transmitter and the receiver of the BH link[14], and is given in dBm as

$$\begin{aligned} \rho_{i,j,ch(\text{dbm})} &= \left(\chi_{Tx}^{\text{BH}} + \chi_{Rx}^{\text{BH}} + \upsilon_{i,j,ch}^{BH} + LM^{BH} + NF^{BH}\right)_{(\text{dB})} \\ &+ N_{th}^{BH}{}_{(\text{dBm})} - \vartheta_{Tx}^{BH}{}_{(\text{dBi})} - \vartheta_{Rx(\text{dBi})}^{BH}, \\ &\forall i,j \in BS, ch \in CH : j \text{ in } N_{m_i} \text{ and } j \neq i, \end{aligned} \tag{4.15}$$

where, the parameters χ_{Tx}^{BH} and χ_{Rx}^{BH} represent the transmitter and receiver losses, receptively, LM^{BH} is the link margin. NF^{BH} denotes the noise figure, N_{th}^{BH} stands for the thermal noise in dBm, while $\vartheta_{Tx\ (\mathrm{dBi})}^{BH}$ and $\vartheta_{Rx(\mathrm{dBi})}^{BH}$ are the transmitter and receiver antenna gains respectively. Thus, in linear scale, $\rho_{i,j,ch(\mathrm{dbm})}$ can be represented by $L_{i,j,ch}^{BH}$, as

$$L_{i,j,ch}^{BH} = 10^{\frac{\rho_{i,j,ch}}{10}}, \forall i,j \in BS, ch \in CH : j \text{ in } N_{m_i} \text{ and } j \neq i. \tag{4.16}$$

Let $\Upsilon_{i,j,ch}^{BH}$ is the SNR of the backhaul link which is related to backhaul link's capacity $R_{i,j,ch}^{BH}$ and the backhaul bandwidth, B^{BH} by the Shannon's capacity theorem as

$$\begin{aligned} R_{i,j,ch}^{BH} &= B^{BH} \log_2 \left(1 + \Upsilon_{i,j,ch}^{BH}\right), \\ \Upsilon_{i,j,ch}^{BH} &= 2^{\frac{R_{i,j,ch}^{BH}}{B^{BH}}} - 1, \\ &\forall i,j \in BS, ch \in CH : j \text{ in } N_{m_i} \text{ and } j \neq i. \end{aligned} \tag{4.17}$$

As Eq. (4.17) is non-linear, a piecewise linear approximation techniques is utilized [92] for its approximation. Now by using Eq.(4.16) and Eq.(4.17), the Tx power consumed by the BH transceiver of link (i,j), [13], can be found as

$$p_{i,j,ch}^{BH} = \Upsilon_{i,j,ch}^{BH} L_{i,j,ch}^{BH}, \forall i,j \in BS, ch \in CH : j \text{ in } N_{m_i} \text{ and } j \neq i. \tag{4.18}$$

The backhaul power, P_{i}^{BH} consumed by each base station i, is sum of the power consumed by all transceivers on the backhaul links associated with that particular

base station i, and is given as

$$P_{i}^{BH} = \sum_{j \in N_{m_i}} \sum_{ch \in CH} \sigma_{i,j,ch}^{BH} P_{0_{i,j}}^{BH} + \sum_{j \in N_{m_i}} \sum_{ch \in CH} \Delta^{BH} p_{i,j,ch}^{BH}, \tag{4.19}$$
$$\forall i, j \in BS: \ j \neq i,$$

where, $\sigma_{i,j,ch}^{BH}$ is a binary variable, which is 1, when the BH link (i, j) on channel ch is active and 0 otherwise, $P_{0_{i,j}}^{BH}$ denotes the minimum non-zero output power of the BH transceiver of the link,and $\Delta_{i,j}^{BH}$ represents the slope of the load-dependent power consumption on the BH link. In Eq.(4.19), the first term denotes the static BH power, while the last term denotes the dynamic BH power. Therefore, the total backhaul power $P_{T_i}^{BH}$ is the sum of backhaul power consumed by each BS and the power consumed by active elements of IRS if they are used is

$$P_{T_i}^{BH} = P_{i}^{BH} + w_1 \sum_{i \in B} \sum_{j \in N_m[i]: j \, in \, B} \wp \, \sigma_{i,j,2}^{BH} \, N, \tag{4.20}$$
$$\forall i, j \in BS: \ j \neq i,$$

where, $\wp$ is the power consumed by a single IRS element and w_1 is the weightage parameter; having a value less than 1 represents that the IRS are powered by the solar panels or their efficiency will significantly increase in near future. The path loss given by equations (4.13) and (4.14) of mmWave and IRS channel respectively, is the deciding factor to route the backhaul traffic through the backhaul links. For our setup, as the path loss of unblocked mmWave channel is always better than the path loss of IRS channel, so the backhaul traffic tends to pass through the mmWave channels. But the occurrence of blockage in the mmWave channel would have a severe impact on its path loss and make it comparable with the path loss of IRS. Moreover It is important to mention that the last term in Eq.(4.14) is the

improvement in the path loss of IRS channel due to the reflecting elements in IRS, N [93]. In addition, due to the higher signal attenuation at mmWave frequencies, we ignore the interference among the adjacent BH channels. This is a plausible assumption, as low-complexity frequency allocation techniques can be executed at an initial stage due to the static nature of the BH network.

4.3 MILP model

The model objective is to minimize the total BH power power consumption of the topology given by Eq.(4.20), while guaranteeing that the maximum users, σ_u^{AN} are to be served by the network.

$$\text{Minimize } P_{T_i}^{BH} - w_2 \sum_{u \in U} \sigma_u^{AN}, \ \forall i \in BS. \tag{4.21}$$

Subject to the following constraints:

$$\sum_{j \in N_{m_i} : j \, in \, BS \cup \{u\}} \xi_{s,u,i,j} - \sum_{j \in N_{m_i} : j \, in \, BS \cup \{u\}} \xi_{s,u,j,i} = \begin{cases} \sigma_u^{AN} R_{\mathrm{u}} & if \ i = s \\ -\sigma_u^{AN} R_{\mathrm{u}} & if \ i = u \\ 0 & otherwise \end{cases}, \tag{4.22}$$

$$\forall s \in MBS, u \in UE, i \in N.$$

$$\begin{cases} \sum_{s \in MBS} M \xi_{s,u,bs,u} & \geq \sum_{rb \in PRB} \sigma_{bs,u,rb}^{AN} \\ \sum_{s \in MBS} \xi_{s,u,bs,u} & \leq M \sum_{rb \in PRB} \sigma_{bs,u,rb}^{AN} \end{cases}, \forall bs \in BS, u \in N_{m_{bs}} : u \neq bs, \tag{4.23}$$

$$\begin{cases} \sum\limits_{rb \in PRB} \sigma^{AN}_{bs,u,rb} & \geq \varrho^{AN}_{bs,u} \\ \sum\limits_{rb \in PRB} \sigma^{AN}_{bs,u,rb} & \leq M \varrho^{AN}_{bs,u} \end{cases}, \forall bs \in BS, u \in UE. \tag{4.24}$$

$$\begin{cases} \sum\limits_{bs \in BS} \sum\limits_{rb \in PRB} \sigma^{AN}_{bs,u,rb} & \geq \varphi^{AN}_{u} \\ \sum\limits_{bs \in BS} \sum\limits_{rb \in PRB} \sigma^{AN}_{bs,u,rb} & \leq M \varphi^{AN}_{u} \end{cases}, \forall u \in UE. \tag{4.25}$$

$$C^{AN}_{bs,u} \leq \sum_{bs \in BS} 1.5 \; R^{AN}_{u} \varphi^{AN}_{u}, \; \forall u \in N_{m_{bs}} : u \neq bs. \tag{4.26}$$

$$\varrho^{AN}_{bs,u} \leq \epsilon^{AN}_{bs,u} \qquad \forall bs \in BS, u \in U. \tag{4.27}$$

$$\sum_{bs \in SBS} \varrho^{AN}_{bs,u} \leq 1, \quad \forall u \in UE. \tag{4.28}$$

$$\sum_{u \in U} \sigma^{AN}_{bs,u,rb} \leq 1, \qquad \forall bs \in BS, \; rb \in PRB. \tag{4.29}$$

$$\begin{aligned} &\sum_{u \in U} \varrho^{AN}_{bs,u} \geq \varphi^{AN}_{bs} \\ &\sum_{u \in U} \varrho^{AN}_{bs,u} \geq \mathrm{M} \varphi^{AN}_{bs}, \qquad \forall bs \in BS, \end{aligned} \tag{4.30}$$

$$P_{bs}^{AN} \leq P_{bs}^{\text{maxAN}}, \qquad \forall\, bs \in BS, \tag{4.31}$$

$$\sum_{s\in M}\sum_{u\in U}\xi_{s,u,i,j} = \sum_{ch\in CH} R_{i,j,ch,}^{BH}\ \forall i,j \in BS : i \neq j \text{ and } j \,\text{in}\, N_{m_i}. \tag{4.32}$$

$$\begin{cases} MR_{i,j,ch}^{BH} & \geq \sigma_{i,j,ch}^{BH} \\ R_{i,j,ch}^{BH} & \leq M\sigma_{i,j,ch}^{BH} \end{cases}, \forall i,j \in BS, ch \in CH : j \,\text{in}\, N_{m_i} \text{ and } j \neq i, \tag{4.33}$$

$$\sum_{ch\in CH} \sigma_{i,j,ch}^{BH} \leq 1,\ \forall i,j \in BS : i \neq j \text{ and } j \text{ in } N_{m_i}. \tag{4.34}$$

$$\sum_{ch\in CH} P_{i,j,ch}^{BH} \leq P_{i,j}^{maxBH}, \qquad \forall\ i,j \in BS, :\ j \neq i. \tag{4.35}$$

The minus sign in objective function, Eq.(4.21), refers to the maximization of number of users, σ_u^{AN}. Moreover, the total BH power, $P_{T_i}^{BH}$ at each base station, is the sum of static BH power, dynamic BH power and the power consumed by IRS elements if they are used. The constraints ensure that the user traffic demand is served, the traffic routing is only done on active BSs and there is no violation of maximum link capacity. The decisions variables specify the user traffic demands R_u, the users associated with each BS $\varrho_{bs,u}^{AN}$, for each flow $\xi_{s,u,i,j}$, the total traffic sent on each link $R_{i,j}$ and the on/off switch state of transceivers of all the BSs ($\varrho_{bs,u}^{AN}$ and $\sigma_{i,j,ch}^{BH}$). Constraint (4.22) represents the flow conservation for the access and

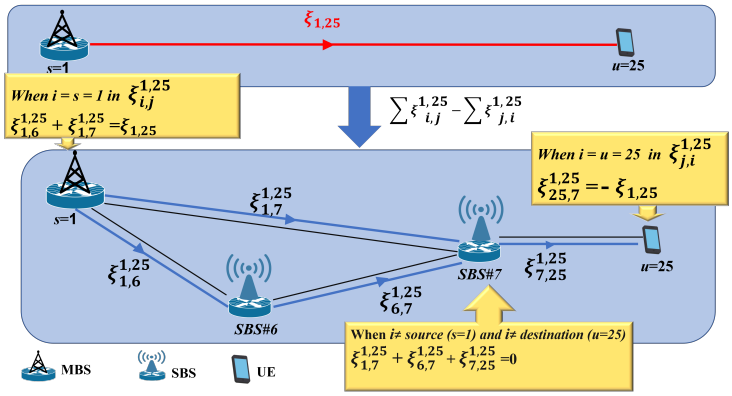


Figure 4.3: Flow conservation principle

backhaul traffic. For better understanding, Fig. 4.3 illustrates the principle of flow conservation. The multiplication of rate variable, R_{u} by binary indicator, σ_u^{AN} is to make it sure either that user is served by a base station or not.

The constraint (4.23) ensures that the resource blocks can only be assigned by a base station to the user if there is a flow on the access link between that particular base station-user pair. The binary variable $\sigma_{bs,u,rb}^{AN}$, is used in Eq.(4.11) to calculate the load dependent RF power of each base station in the access side. Constraint (4.24) ensures if there is a single or multiple resource blocks are assigned to a user from a base station, then the binary indicator $\varrho_{bs,u}^{AN}$ is set to 1 thus showing that the transceivers in that base station are switched ON in the access link . Constraint (4.25) ensures the user connectivity i.e., if the demand of the user is anchored by a base station by assigning single/multiple resource blocks to that user thus, making φ_u^{AN} equals to 1. Constraint (4.26) ensures that the capacity of each user from a base station is less than 1.5 times the total demand required by that user. In constraint (4.26), demand of each user, R_u^{AN} is multiplied with the the binary variable, φ_u^{AN}, to ensure that the capacity constraint (4.26) should only considers active user demands. It is important to mention here that constraint (4.26) is only employed for the first two sets of results where we are forcing the model to serve all the users. Hence,

in those scenarios, we can't allow the model to allocate all the resource blocks to the users, which will impact access power which is not included in our objective function; thus, we introduce a limit on the maximum capacity offered to each user.

Constraint (4.27) ensures that the user is only served by that particular base station; if the user lies with in the range, $\epsilon_{bs,u}^{AN}$ of the base station, Constraint (4.28) ensures that the user is allowed to be served by a single base station in the access side. There is also a possibility that a particular user is not served by any base station i.e., user blocking is allowed. Constraint (4.29) ensures that if a resource block is to be assigned by a base station to the user, in that case, it only assigns that resource block to a single user. Constraint (4.30) ensures the switching of the base station, i.e., φ_{bs}^{AN} is set to 1, if the base station is anchored to any user in the access side, otherwise 0. The binary variable φ_{bs}^{AN} is used in Eq.(4.12) to consider the static power of only serving BSs in calculating the total access power. The constraint (4.31) ensures that the power consumed by each base station, P_{bs}^{AN} in eq.(4.11) for forwarding the access link traffic is less than the maximum access link power P_{bs}^{maxAN} of each base station, the value of $P_{bs}^{\text{maxAN}}$is given in Table 4.4.

This constraint (4.32) ensures that if there is a flow between two base stations in the backhaul link then it will be equal to the rate of that backhaul link. The constraint (4.33) ensures if there is any traffic on the backhaul link using any channel then it will switch rate indicator $\sigma_{i,j,ch}^{BH}$ to 1. The binary variable $\sigma_{i,j,ch}^{BH}$ is used in Eq.(4.19) to calculate the load dependent RF power of each base station in the Backhaul side. Constraint (4.34) ensures the traffic through each backhaul link is passed through only one channel, i.e, it can pass through either the mmWave or the IRS channel. As each backhaul link is supported by two channels i.e., mmWave or IRS channel, so splitting of BH traffic between the these two channels is not allowed. Constraint (4.35) ensures that the power consumed by each base station for each backhaul link

is less than the maximum backhaul power P_{bs}^{maxBH} of the base station.

Therefore, form Eq.(4.32), the flow variable, $\xi_{s,u,i,j}$ in our optimization problem, gives us the rate (capacity) of each backhaul link on each channel. From the rate variable $R_{i,j,ch}^{BH}$, the model will find the associated SNR, $\Upsilon_{i,j,ch}^{BH}$ of the each backhaul link by using Eq.(4.17). Thus , the model will find the power consumed by each backhaul link, $p_{i,j,ch}^{BH}$ and power consumed by each base station, P_i^{BH} in serving all the attached BH links By using Eqs.(4.19) , (4.18).

4.4 Results

As mentioned earlier, IRS channels are deployed along the mmWave channels to provide the alternative path to the BH traffic for each blocked mmWave channel. Although, the number of active elements in each IRS also provides a gain in the path loss, which is much better than the path loss of the blocked mmWave channel but still these active elements of each IRS channel (if it is used) also consumed power. Hence, introducing the IRS channels along the mmWave channels in the network added the burden on the BH power. The user association at the access side is accomplished by constraint (4.26). Moreover, in Eq.(4.19), the total backhaul power includes the static BH power, the load-dependent dynamic BH power, and the power consumed by active elements of all the used IRS channels. The results have been divided into three sections. The first section assumes that some of the mmWave channels are blocked simultaneously in a grid topology where users density is continuously varied. Hence, the impact on the static BH power is investigated by introducing the IRS channels along the mmWave channels. Then how can we achieve the maximum saving in the dynamic and total backhaul power explored in section two. Finally, in the last section, our focus is on increasing the number of

Table 4.4: ENERGY-EFFICIENT MILP MODEL INPUT PARAMETERS

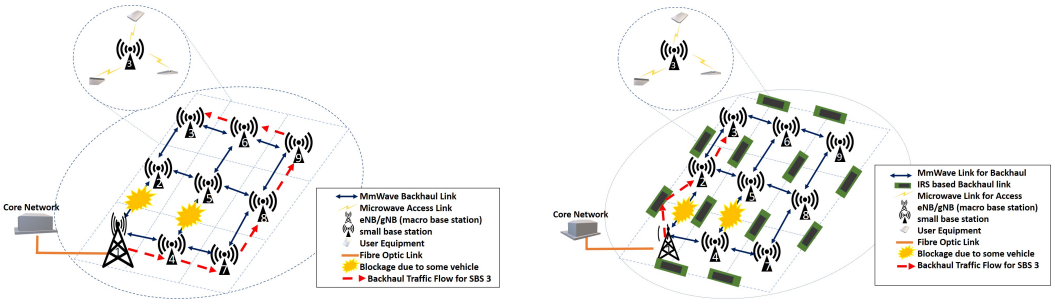
Parameters	Comments
Range parameter in meters i.e., if a user lies in certain range of a BS $\epsilon_{bs,u}^{AN}$	50-300 (m) [90]
Bandwidth of each resource block ϖ_{rb}	180 (KHz) [90]
Demand of each user R_{u}^{AN}	1-100 Mbps[94]
LTE bandwidth in the access side B^{AN}	20 (MHz) [90]
Noise variance in the access side σ^{AN}	$\left(10^{-\frac{174}{10}}\right) B^{AN}$[90]
Maximum transmitted power of small base station in access side $P_{bs}^{\mathrm{maxAN}}\ \forall\, bs \in SBS$	$10^{\frac{24}{10}}$ (mW) [90]
Maximum transmitted power of small base station in access side $P_{bs}^{\mathrm{maxAN}}\ \forall\, bs \in MBS$	$10^{\frac{46}{10}}$ (mW) [90]
Scaling factor in SINR calculation of each PRB η^{AN}	0.5
Range parameter in the access side, $\epsilon_{bs,u}^{AN}$	80-200 (m)
Minimum non-zero output power of small base station in access side $P_{0_{bs}}^{AN}\ \forall\, bs \in SBS$	130 (W) [90]
Minimum non-zero output power of small base station in access side $P_{0_{bs}}^{AN}\ \forall\, bs \in SBS$	6.8 (W) [90]
Set of physical resource blocks, PRB	100 [90]
Slope of the load-dependent power consumption in the access side $\Delta_{bs}^{AN}\ \forall\, bs \in SBS$	4.0 [90]
Slope of the load-dependent power consumption in the access side $\Delta_{bs}^{AN}\ \forall\, bs \in SBS$	4.7 [90]
Number of transceiver chains in the BH link $\Pi_{i,j}^{BH},\ \forall i,j \in BS, i \neq j$	1
Maximum transmitted power of a base station in the backhaul side $P_{i,j}^{maxBH}\ \forall i,j \in BS, i \neq j$	224 (mW) [95]
Operating frequency of backhaul link f^{BH}	60 (GHz) [96]
mmWave bandwidth in backhaul link B^{BH}	500 (MHz) [50]
Minimum non-zero output power of the BH transceiver ,$P_{0_{i,j}}^{BH}$	3900 (mW) [95]
Slope of the load-dependent power consumption in the BH link $\Delta_{i,j}^{BH}$ (treated as microwave BS)	4 -40[97]
Number of active elements in each IRS, N	1000-5000
Power of single active element in IRS, $\wp$	5 (mW) [24]
Very large +ve number, M	10000
Signal attenuation due to vapour water and oxygen at the Backhaul linkϵ_{vap+O_2}	16 (dB) [3]
Signal attenuation due to rain at the backhaul link ϵ_{Rain}	2 (dB) [56]
Transmitted antenna gain in backhaul link $\vartheta_{Tx}^{\mathrm{BH}}$	30 (dBi) [98]
Received antenna gain in backhaul link ϑ_{Rx}^{BH}	30 (dBi) [98]
Losses at the transmitter side in backhaul link χ_{Tx}^{BH}	5(dB) [99]
Losses at the receiver side in backhaul link $\chi_{\mathrm{Rx}}^{\mathrm{BH}}$	5 (dB)[99]
Noise Figure for the backhaul link NF^{BH}	6 (dB) [100]
Link Margin of backhaul link LM^{BH}	9 (dB)[98]
thermal noise the backhaul link N_{th}^{BH}(dBm)	$-174+\ 10\log_{10}\left(B^{BH}10^{6}\right)$[14]
Number of Users	10-30
Blockage weightage β^{BH}	40 (dB)
Binary random parameter for backhaul links $\Gamma_{i,j}^{BH},\ \forall i,j \in BS, i \neq j$	10-100%
Weightage parameter 1 for objective function w_1	0-1
Weightage parameter for the users, w_2	0/1

served users by deploying the minimum number of IRS channels in a tree topology. simplify the analysis, the results do not take the access power consumption explicitly, but it is restricted by the fact that capacity is limited by constraint (4.26). the model objective can be easily extended to include explicit access power minimization.

4.4.1 Static BH Power Analysis

To study the impact of introducing IRS channels on the static component of backhaul power, a grid topology containing nine base stations is considered, as shown in Fig.4.4. Initially, the users are distributed in the vicinity of only one BS, i.e., BS#3. The objective function here, is to only minimize the total BH power, and this can be done by setting $w_2 = 0$, in the objective function (4.21). Thus the capacity constraint (4.26) in the access side is removed from the optimization problem and forces the model to serve all the users. In Eq.(4.20) the total BH power, $P_{T_i}^{BH}$ also contains the dynamic component of BH power, which is included only to get the flow conservation working. Generally, in the absence of any blockage in the mmWave channels, the backhaul traffic would take the shortest route, i.e., it is generated from gNB and routed via BS#2 to BS#3 using only the mmWave channels. Now consider a blockage scenario, in which two of the mmWave backhaul channels got blocked simultaneously, i.e., the BH channel between BS#1 and BS#2 and the BH channel between BS#4 and BS#5. As discussed earlier, due to the inevitable blockage between BSs, mmWave communication may not be accomplished in these backhaul links. Thus, the backhaul traffic takes more hops to reach BS#3, as shown in Fig.4.4(a), increasing the static power of backhaul links, as more BSs are involved in routing the backhaul traffic for the users in BS#3.

In Fig.4.4(b), an IRS channel is introduced along each mmWave channel. Now with the same blockage scenario as mentioned in the last paragraph, the model would



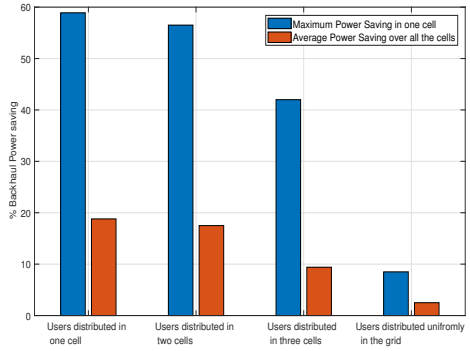
(a) In the absence of IRS, the backhaul traffic will take more multi-hops by using the unblocked mmWave channels to reach BS#3

(b) In the presence of IRS, the backhaul traffic will take less hops by also using the IRS channels along with the unblocked mmWave channels to reach BS#3

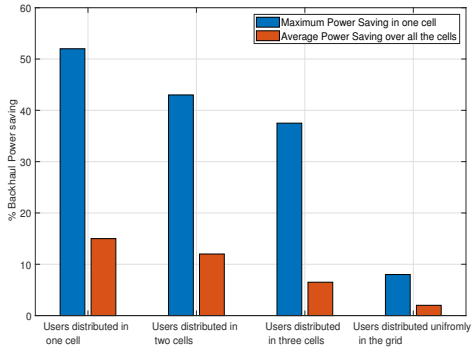
Figure 4.4: Backhaul traffic route for users of BS#3, when some of the mmWave links got blocked

decide whether to take more hops using only the unblocked mmWave channels or lesser hops via a combination of unblocked mmWave channels and the IRS channels. The first choice consumes more static BH power but zero IRS power, while the second choice consumes less static power but non-zero IRS power. For the particular case of users distributed in BS#3, the model picked a later option by preferring fewer hops via a combination of IRS unblocked mmWave channels, thus consuming less BH power.

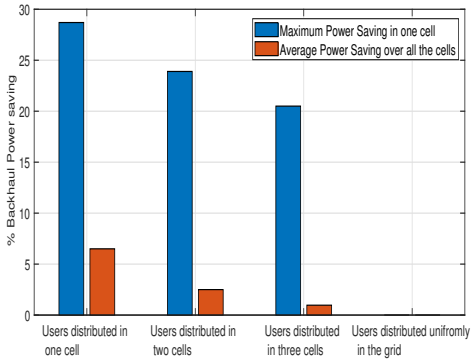
The impact on the BH power is further investigated by distributing the users in different scenarios, while considering those two mmWave channels are still blocked. In the first scenario, users are successively distributed in the vicinity of a first, second and third Bs, and so on, while the same approach is followed in the second scenario, where the users are now distributed in two BSs each time. Similarly, in the third scenario, users are distributed in the vicinity of three BSs, while in the last scenario, users are distributed uniformly in the whole grid. In each scenario, the maximum BH power saving is determined by the saving occurred due to a specific BS/BSs user distribution, while the average BH power-saving would take the average of all the savings in BH power among all the BS/BSs user distributions. The simulation



(a) When the number of active elements of IRS N, is 700



(b) When the number of active elements of IRS N, is 1000



(c) When the number of active elements of IRS N, is 2000

Figure 4.5: %age of average and maximum total Backhaul power saving for various users distributions,

is repeated 100 times for each scenario, and the percentage saving in maximum and average BH power is plotted for different values of active elements in then IRS, as shown in Fig.4.5. For $N = 700$, it is observed that highest saving in both the the maximum and average BH power occurred, when the users are distributed in the first scenario (i.e., in the vicinity of a single BS) as shown in Fig.4.5 (a), while the least savings are observed when the users are distributed uniformly in the whole grid. The gradual decrease in backhaul power savings is because more BH links are used for the users, distributed in more than one cell. When N, is further increased to 1000, again, the highest savings are observed for a single cell user distribution

scenario, as shown in Fig.4.5(b). However, these savings are comparatively less than the case when N, was 700, as there is more burden of power consumed by active elements of IRS. With the further increase in N to 2000, the savings in BH power keep on decreasing from single-cell to uniformly distributed scenario, where they ultimately become zero, as shown in Fig.4.5(c). Form the above results, the maximum saving in static power is observed for N=700. It is noteworthy to mention that the percentage saving in total BH power is mostly coming from the static power as dynamic BH power is almost 3% of the static BH power.

The above investigation is taken into account the failure of two fixed mmWave channels which occurred near the source BS (gNodeB), the future work will be the investigation of the saving in BH power due to the impact of location of failure in mmWave channels, i.e., either it will be near the gNB or far away from it, which will give another degree of freedom.

4.4.2 Dynamic BH Power Analysis

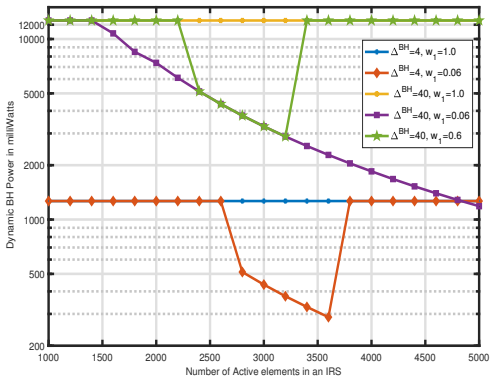
The last observation inferred that deploying the IRS channels along the mmWave channels gave us some savings in the total BH power. Specifically, the static component of BH power is significantly reduced due to fewer hops, whereas the dynamic component of BH power is relatively unchanged. As mentioned earlier, the path loss of IRS channel in Eq.(4.14), also depends on N, which is related to dynamic backhaul power in Eqs. (4.17, 4.18). Therefore, we studied the impact of N on the dynamic and the total BH power.

In the first setup, named as fixed blockage, all the mmWave channels are assumed to be partially blocked simultaneously as shown in Fig.4.2, by adding a fixed blockage parameter of 40 dBs in Eq.(4.13), thus making the path loss of mmWave channel severely bad but still usable. The users with specific demand rates are uniformly

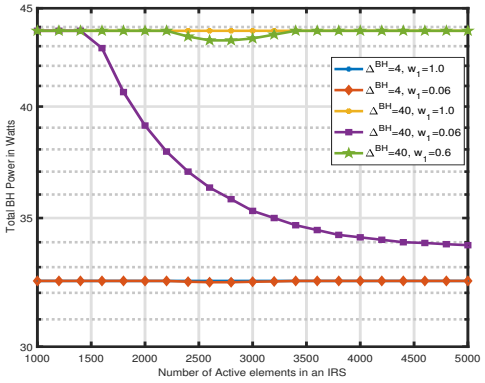
distributed in the grid. The N is varied in the range of 1000 to 5000 elements, by ignoring the values less than 1000, as they would not give us the enough gain in the path loss of IRS channel as compared to the path loss of partially blocked mmWave channel. Moreover, selecting such a large value for N in an IRS is only possible in the mmWave backhaul networks because each IRS can be configured "offline". The objective function here is again to minimize the backhaul power by setting $w_2 = 0$ in the eq.(4.25), so the capacity constraint (4.26) in the access side is removed from the optimization problem and forces the model to serve all the users. The dynamic BH power, the total BH power, which includes the static and dynamic BH power plus the power consumed by IRS elements and the number of active BH channels, are plotted against N in Fig.4.7, for different values of load-dependent slope of dynamic BH power, Δ^{BH}, and weightage parameter for IRS power, w_1 as given in Eq.(4.19).

Initially, when $\Delta_{BH} = 4$ and $w_1 = 1$, all the BH traffic is routed through the mmWave backhaul channels as the IRS component in BH power is large enough to dominate. Keeping Δ_{BH} =4, and reducing the impact of IRS power component by making w_1=0.06. Now, initially varying the N from 1000 and 1800, all the BH traffic is routed through the partially blocked mmWave backhaul channels due to negligible improvement in path loss of the IRS channel Eq.(4.14), and hence the total backhaul power includes the power consumed by the mmWave channels. When N is further increased between 1800 to 4800, the path loss offered by the IRS is better than the path loss of partially blocked mmWave channels.

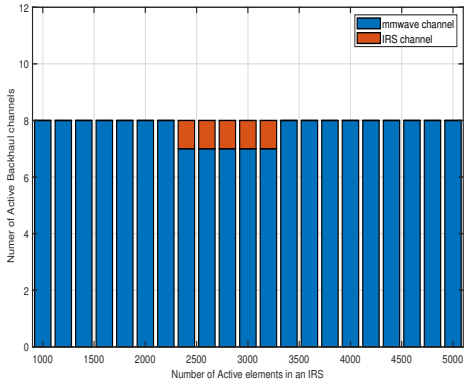
Therefore, the model prefers IRS channels for routing for some of the BH traffic of blocked mmWave channels, while other BH traffic is still routing through the partially blocked mmWave channels, as shown in Fig.4.7(c). This diversion causes a decrease in the dynamic and the total BH power, as shown in Fig.4.7 (a) and (b)



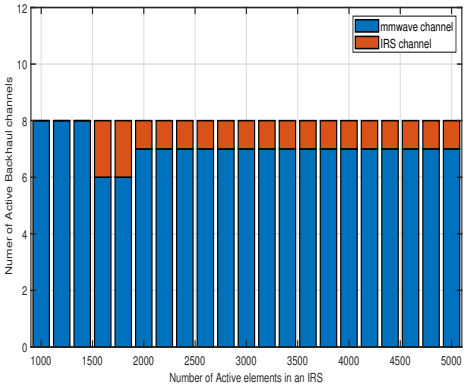
(a) Semi-log plot of the dynamic backhaul power



(b) Total backhaul power



(c) Number of active backhaul channels for $\Delta^{BH} = 4$ and $w_1 = 0.06$



(d) Number of active backhaul channels when $\Delta^{BH} = 40$ and $w_1 = 0.06$

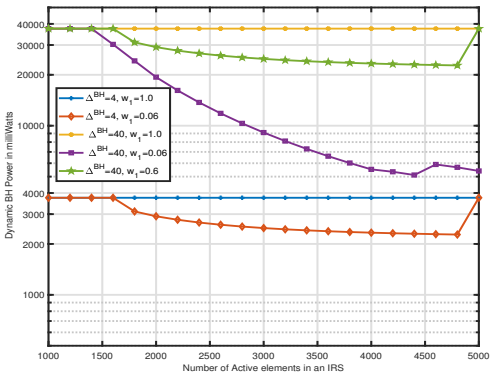
Figure 4.6: Impact on the BH power due to the number of active elements (N) in an IRS, when different mmWave links undergo random blockage, i.e., either 0 dB or 40dBs

respectively. The IRS power component in total BH power in Eq.(4.19) is comparatively small during this range.With the further increase in N beyond 4800, results in the further reduction of the path loss of the IRS channel; on the contrary, it also significantly increases the power consumption of IRS channels, hence abstaining the model to prefer IRS channels over the partially blocked mmWave channels. Therefore, the model switches OFF all the IRS channels, and the traffic again starts flowing from the mmWave channels, and the dynamic and total BH power curves

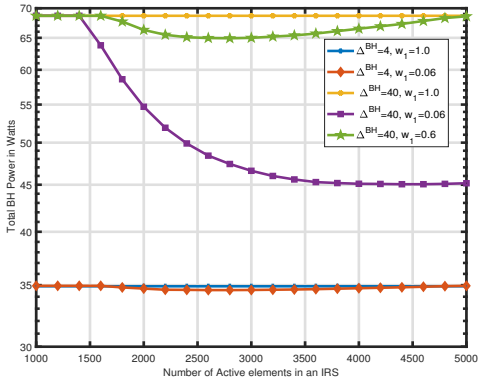
return to the initial value. Now by changing Δ^{BH} from 4 to 40 and w_1 to 1, significantly increased the dynamic and total BH power for each value of N. However, the larger value of IRS power component abstain the model to prefer IRS channels over partially blocked mmWave channels. Therefore, w_1 is reduced to 0.06, and the dynamic and total BH power starts decreasing with N, also as the IRS power component is too low as compared to the BH power component; more number of IRS channels are used as shown in Fig.4.7 (d) as compared to the Fig.4.7 (c). Thus, helping the model to use IRS channels for a longer range of N. Now by fixing Δ^{BH} to 40 and changing w_1 to 0.6, both the dynamic and total BH power decreases with increase in N,then after a certain value of N, they started to increase. At when N=4600, an increase in dynamic BH power is observed, where the model switched off one of the IRS channels and an already serving base station in the BH to save the BH power. Thus, all the BH traffic is rerouting with users still associated with previous BSs, causing more hopes to reach the BSs in the backhaul. With a further increase in N, beyond N=4600, the both dynamic BH power again starts decreasing again.

The above proposition is our baseline scenario for all future analyses. In the above analysis, all the mmWave channels are assumed to be partially blocked; this phenomenon is scarce. Thus a random blockage parameter of 40 dBs, is chosen so that nearly half of the mmWave channel undergo blockage at a time, whereas reaming settings remain unchanged. The dynamic and the total power is plotted against N in Fig.4.6. Again the dynamic and total backhaul power follow the same trends for the particular values of Δ_{BH} and w_1 as was shown in Fig.4.7.

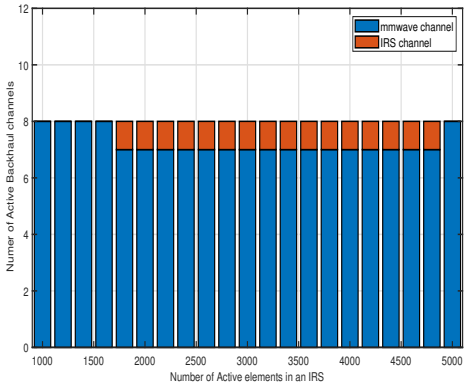
By comparing the results of fixed blockage with those of random blockage, we see that in the case of fixed blockage, the network used IRS channels to pass some of the partially blocked mmWave channels, however, the remaining BH traffic was forced



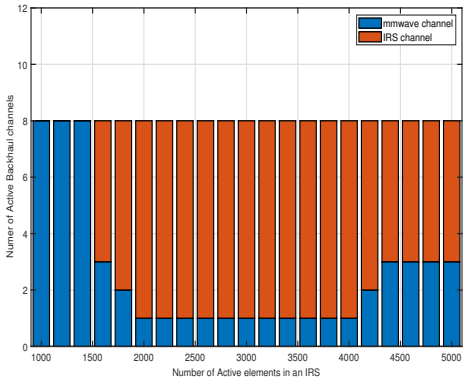
(a) Semi-log plot of the dynamic backhaul power



(b) Total backhaul power



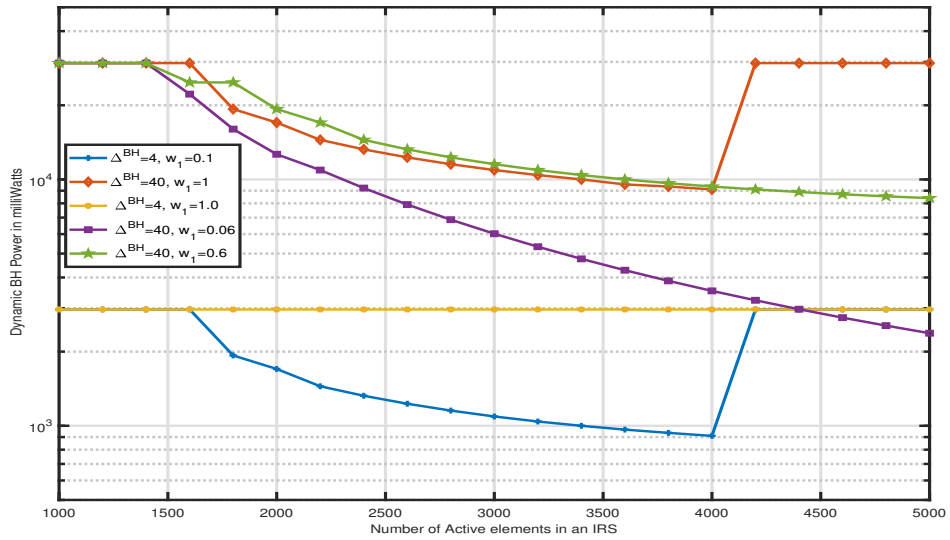
(c) Number of active backhaul channels for $\Delta^{BH} = 4$ and $w_1 = 0.06$



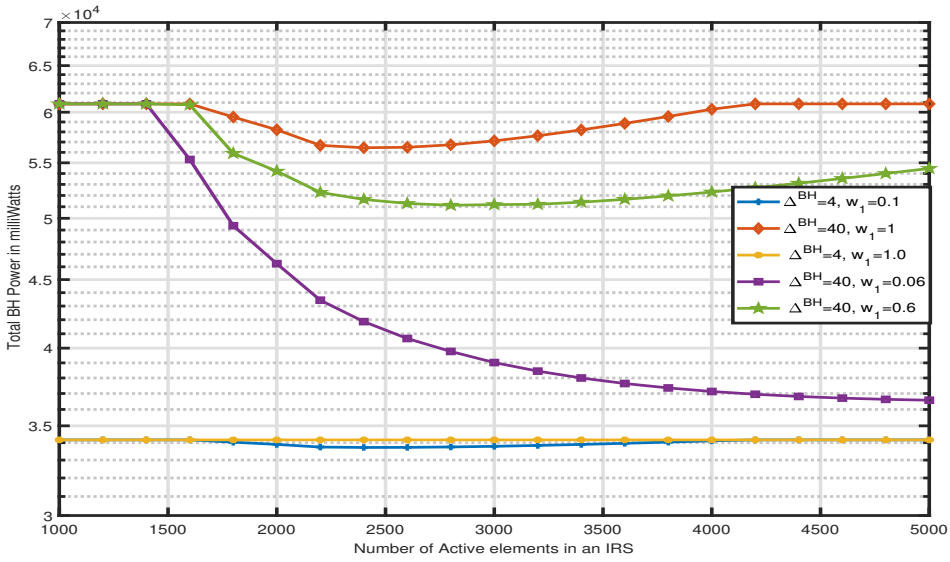
(d) Number of active backhaul channels when $\Delta^{BH} = 40$ and $w_1 = 0.06$

Figure 4.7: Impact on the BH power due to the number of active elements (N) in an IRS, when all the mmWave links undergo fixed blockage of 40 dB

to go through the partially blocked mmWave channels with greater path loss, since there was no alternative. However, for the random blockage, the model optimizes the traffic of partially blocked mmWave channels that are diverted through unblocked mmWave channels and IRS channels, thus reducing the number of IRS channels used, as illustrated in Fig.4.6(c) and (d). It is also observed that in the case of random blockage, the specific value of both the dynamic and total BH power against each value of N Fig.4.7 (a) and (b) is less then the corresponding values in fixed blockage case as shown in Fig.4.6(a) and (b), as we have a number of good links



(a)



(b)

Figure 4.8: Impact on the BH power due to the number of active elements (N) in an IRS, when different mmWave links undergo random blockage, i.e., either 0 dB or 40dBs, considering the capacity constraint at the access side

with lower power requirements. Also we got more saving in dynamic BH power for random blockage as compared to fixed blockage (for each value of Δ^{BH} and w_1),

as the model has the opportunity to use good links that help in minimizing the dynamic BH power. From the previous discussion, it can be proclaimed that using the IRS channels along the mmWave channels can provide savings in the static BH power (depending on the users distribution) as well as the dynamic BH power. The results also reveal the optimum number of IRS elements values for maximum saving in dynamic and backhaul power within the studies range of values for those elements and the input parameters of our problem. It is also observed that the power consumed by IRS is one of the limiting factors for minimizing the backhaul power, especially when the same network provider is responsible for providing power to IRS along with the mmWave BSs.

4.4.2.1 Considering the Access power

In this section we consider the capacity constraint on the access side, only those users are allowed to be served by the model, whose demand R_u is less than the capacity $C^{AN}_{bs,u}$ offered to them by associated base station and is given as

$$C^{AN}_{bs,u} \geq \sum_{bs \in SBS} R^{AN}_u \varphi^{AN}_u, \; \forall u \in N_{m_{bs}} : u \neq bs. \tag{4.36}$$

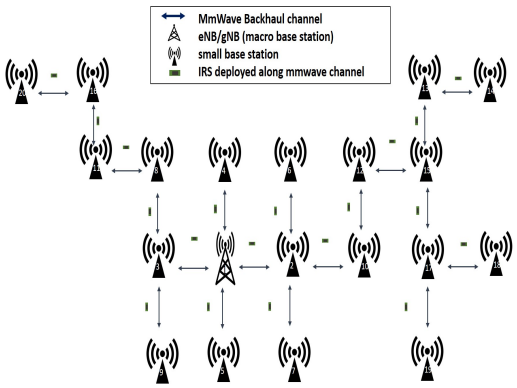
The other objective function and constraints remains same.The Fig.4.8, shows the dynamic and backhaul power for IRS elements, and it can be observed that the system follows the same trends but with different weightage parameters, which shows that the system is very sensitive to weightage parameters.

4.4.3 Users Analysis

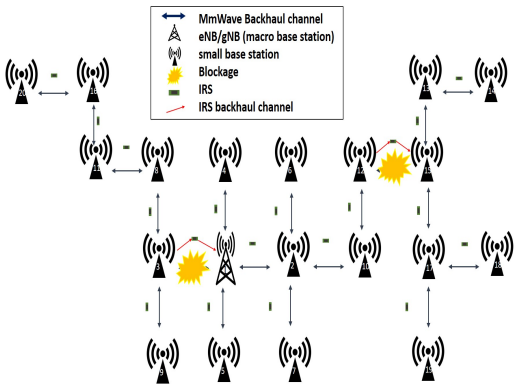
The question arises on how introducing the minimum number of IRS still serves the maximum user association with optimum backhaul power. Consider a network

topology comprising macro and small base stations connected via mmWave channels, as shown in Fig.4.9(a), and thirty users are uniformly distributed in the vicinity of those BSs. User association is classified into two categories, i.e., in the first category, named as fixed users association; users can only be served by the nearest BS while, in the second category termed as optimized user association, a user can be served by one of the nearby BSs, provided it lies in the range of that particular BS. Consider a case in which two of the mmWave channels are blocked simultaneously, as shown in Fig.4.9(a), i.e.,channels between BS#1-BS#3 and the channel between BS#12-BS#15 are blocked. The impact on user association due to blockage is studied under those three scenarios, based on the availability of the IRS channel. The objective function here is to maximize the number of users (by setting $w_2 = 1$), while minimizing the dynamic BH power, which is only needed to smooth the BH traffic flow. In the first scenario, only mmWave channels are used for routing BH traffic; it is revealed that the total number of served users decreases in the absence of any IRS channel for both fixed and optimized user associations.

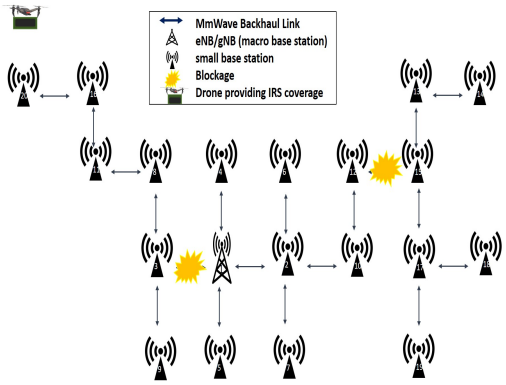
Moreover, in the case of optimized user association, an increase in the range parameter also caused an increase in the number of served users. In the second scenario, as it is not cost-effective to deploy the IRS channel with each mmWave channel, a novel approach uses a drone to provide the IRS channel for one of the blocked mmWave channels. As the two mmWave channels are assumed to be blocked, so the model will decide which one of the blocked mmWave channels will be served by the IRS; hence, the resulted number of users for this scenario is shown in Fig.4.9(d). The numbers of served users in this scenario increase significantly for both the optimized and fixed user association compared to the 1st scenario due to the IRS availability for blocked mmWave channels. Furthermore, in this scenario, the number of served users kept on increasing as the range parameter increases till the point, ($\epsilon^{AN}_{bs,u}$=200



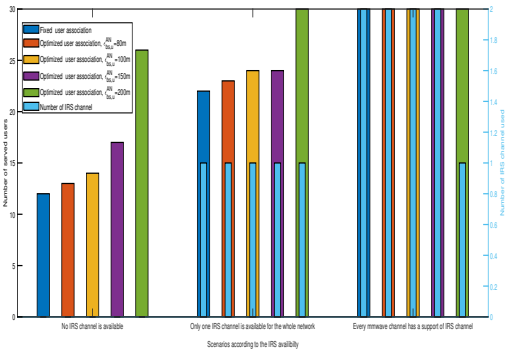
(a) Tree topology for mmWave backhaul network where IRS channel is deployed along each mmWave channel



(b) When backhaul traffic of each blocked mmWave channel is rerouted trough alternative IRS channel



(c) when there is limited number of IRS, so the model has to optimized the number of served users accordingly

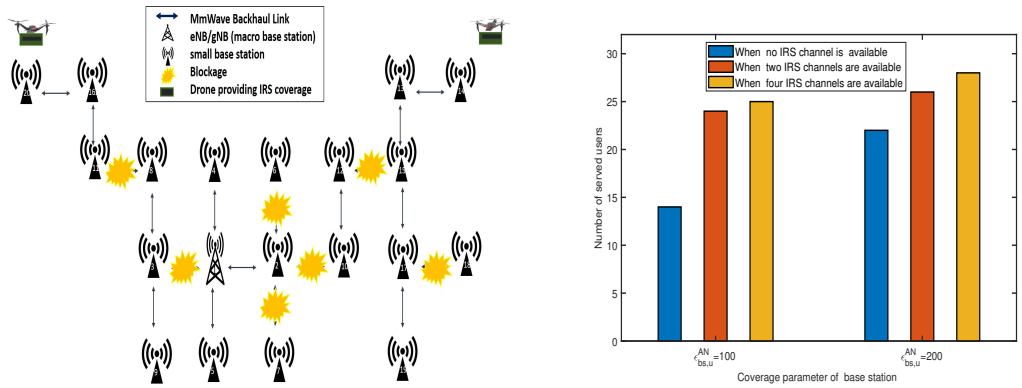


(d) Number of served users according to the IRS availability

Figure 4.9: Impact on the number of served users due to the blockage of only two fixed mmWave channels with/without the support of IRS

meters), the network can serve all the users with only one IRS channel as one of the blocked mmWave channel traffic is diverted through other mmWave channels through the farthest BSs.

It is further noted that when $\epsilon^{AN}_{bs,u}$=100, users are connected with nine BSs, whereas for $\epsilon^{AN}_{bs,u}$=200, only seven BSs are used to serves users. For $\epsilon^{AN}_{bs,u}$=100, user#33 was the only user connected to the BS#7, but when $\epsilon^{AN}_{bs,u}$ is increased to 200, user#33 is now connected with one of the neighbouring BS, i.e., BS#1, therefore enabling the model to switch OFF BS#7. In the third scenario, every mmWave channel has the



(a) Several mmWave channels blocked simultaneously with limited IRS availability

(b) Impact on the number of served users due to the blockage of several random mmWave channels with/without the support of IRS

Figure 4.10: Impact on the number of served users due to blockage of several mmWave channels

support of the IRS channel in the BH link to ensure the maximum user association as shown in Fig.4.9(a).

It is observed that by using the IRS channels for those two blocked mmWave channels, the network can serve all the users for both fixed and optimized users association. In the case of optimized users with smaller range parameters, all the alternative IRS channels for the blocked mmWave channels are used. But, with an increase in $\epsilon^{AN}_{bs,u}$, the network starts associating users to the farthest BS instead of the nearest one to save BH power in a manner that the minimum number of BH links are used, e.g. for $\epsilon^{AN}_{bs,u}$ =200 meters; several mmWave channels, including one of the blocked mmWave channels, are no longer needed; hence, the BH traffic required only one IRS channel for the other blocked mmWave channel, which is still in use, thus consuming less dynamic BH power. Consequently, we can use this model to reduce the cost of deploying IRS even though we do not explicitly have this target in the objective function.

The above investigation took into account the blockage of only two fixed mmWave channels, but due to the random nature of blockage in mmWave channels, the impact

on the number of served users by blocking random mmWave channels as shown in Fig4.10(a), is studied, in which seven random mmWave channels are blocked simultaneously for the same number of users, i.e., 30. In the first scenario, no IRS channel is available to provide the alternative link to the blocked mmWave channel, whereas, in the second scenario, two IRS provided by the drones, which are present to provide the alternative path to the blocked mmWave channels, and in the third scenario, four IRS channels are available.

The simulation is repeated 100 times for different sets of blocked mmWave channels, and the average number of served users is plotted against the coverage parameter $\epsilon_{bs,u}^{AN}$, as shown in Fig.4.10. It is revealed that the number of served users increases with the introduction of IRS channels along with the mmWave channel. However, increasing the IRS channels from two to four has little impact on the average number of served users. Furthermore, by comparing Fig.4.9 and Fig.4.10, it is also observed that the average number of served users decreases with an increase in the number of blocked mmWave channels for the same number of IRS. Furthermore, it is also found that the increase in the coverage parameter resulted in more served users along with a lesser number of IRS channels. It is important to mention that these results obtained without capacity constraints at the access side and without considering access side power consumption, this is to simplify the analysis and focus on the mmWave vs IRS supported backhaul.

4.5 Conclusions

Throughout this chapter, we discussed how IRS could be incorporated into the backhaul links of future wireless networks in order to reduce their power consumption in the event of blockage or to maximize their user capacity. In order to minimize

backhaul power while maintaining service to the maximum number of users, a MILP optimisation model was developed. When certain mmWave channels are simultaneously blocked, introducing IRS along the mmWave channels can save both static and dynamic backhaul power in comparison to rerouting to mmWave backhaul links only. Using the optimal number of IRS along mmWave channels can also maximize the number of users served. Furthermore, it is necessary to reduce the power consumption of each IRS element in order to make IRS blocks (which contain many elements) comparable with backhaul connections in terms of power consumption.

Chapter 5

Latency Minimisation using IRS in MEC

In this chapter, we examine the advantages of IRS in an edge computing setting in the context of a multi-channel-based, multi-user, downlink communication system by guaranteeing service to all the users. We first introduce the system model which is the foundation of our optimization framework. We then introduce the MINLP objective function with all the constraints. We then present a linearisation technique in order to make our model linear so that it can be solved by a linear solver such as CPLEX. Using the linear objective function and constraints, we propose a MILP model. It is our objective to jointly minimise both the latency (including the process delay) and the power consumption of the IRS elements. By controlling the weights of the objective function, which reflects the preferences of the operator, the IRS elements can be tailored to minimise latency, IRS power consumption, or both.

5.1 Introduction

New advances in programmable metamaterials [101], are fostering the emergence of intelligent reflective surfaces (IRS) that are designed to improve the energy and spectral efficiency of wireless communication [22]. IRSs are composed of an IRS controller and many near-to-passive reflection elements which are controlled by the IRS controller. The IRS controller determines the phase shift of these elements. At the receiver, the signals reflected from the IRS will combine coherently with those reflected from other paths, resulting in an amplification of the received signals [29], [30]. IRSs are normally embedded on printed circuit boards (PCBs) and the metasurfaces are located on top of a dielectric substrate, as illustrated in Fig.5.1. Two metallic patches are attached to each reflecting element by a common varactor/pin diode, which is connected to the cylindrical conductive post, thereby forming a parallel connection between the diode and the cylindrical post. These cylindrical posts pass through a dielectric substrate, a control layer, and end at a metal layer, while the metal layer serves as a ground plane. In this case, the inductance of the circuit is determined by the thickness of the substrate and the length of the cylindrical post. By using an FPGA-based controller, a DC voltage is applied to each varactor, causing its capacitance to vary as a function of the bias voltage. An insulation layer is used to prevent electromagnetic leakage. Following reception of the RF signal, current starts flowing through the metallic patch and then splits into two paths: via the varactor and cylindrical post. Consequently, the RF signal is reflected by that metallic patch on the opposite side, where the distributed current meets. In addition, it is important to note that each IRS element is tuned by a varactor/pin diode, which requires a power supply circuit [102].

To investigate the advantages of IRSs in wireless communications, comprehensive research efforts have been spent into their channel estimation [34] and capacity

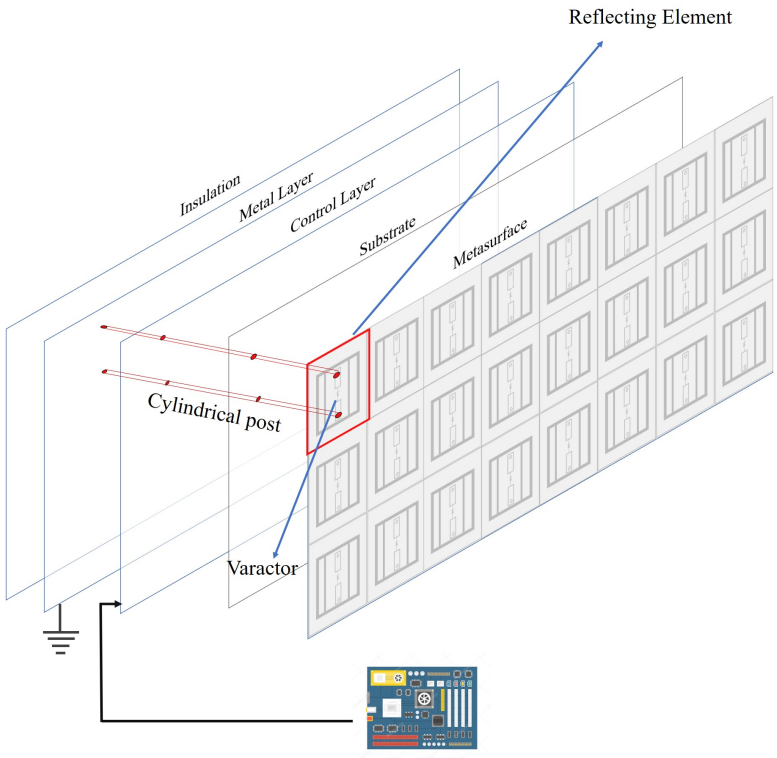


Figure 5.1: The architecture of IRS, where a single reflecting element is marked in red.

analysis [35]. In [41], IRS elements are grouped together, where the same phase change coefficient is shared by all the elements of each group.

According to their study [103], different bias voltages were applied to a single reflecting element to achieve frequency selectivity over a particular frequency range. In the study, the authors considered a frequency range of 5.5 GHz to 6 GHz. They varied the bias voltage from 0 to 19 volts, which produced a phase shift of almost 180 degrees. Through their work, they laid the foundation for frequency selectivity using the IRS, where users won't receive interference from other users.

In today's 5G and future 6G era, billions of sensors and machines are anticipated to be connected. [104]. Due to the limited processing capabilities of these devices, they are unable to accommodate time critical and resource-intensive applications.

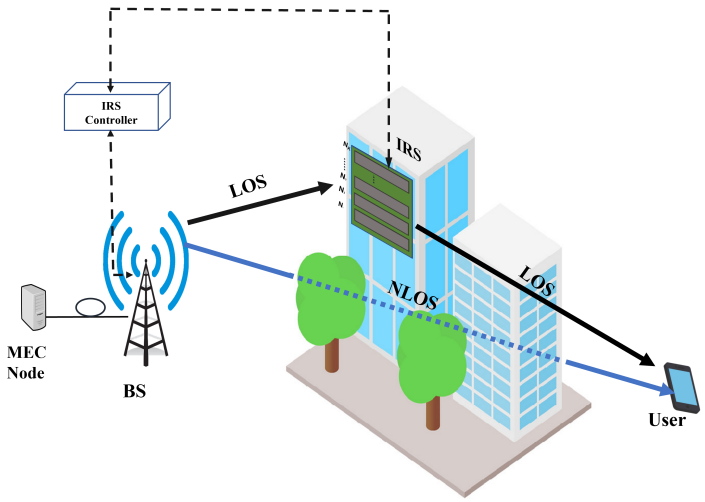


Figure 5.2: IRS aided single user downlink system with MEC server

This problem can be addressed by network operators deploying strong computing nodes adjacent to the network's edge, generally at the base stations (BSs) [105]. In this manner, devices with low computational power can benefit from the computing capabilities of edge devices by offloading their resource-intensive tasks to these devices. It is known as Mobile Edge Computing (MEC), and nodes with higher processing power are referred to as MES servers. Nevertheless, devices that are typically located at the cell edge may have a higher computational offloading latency than if they were to perform their computation locally if the link employed to offload the computation is not optimal [106]. Therefore, these devices are forced to use their own computing resources, with the limitation that resource-intensive applications cannot be supported. It is therefore imperative to improve communication performance between devices that support MEC.

As part of the MEC paradigm, a variety of techniques have been proposed to address the issue of higher latency due to unfavourable channel conditions. In [68], for example, it was suggested that if a channel between the MEC server and the

user is not suitable to support an adequate transmission rate, then deferring the computational offloading will be a better option until a more appropriate channel can be created or by switching completely to another frequency/spatial channel that offers better performance and quality. Furthermore, the authors propose that mobile devices that cannot communicate directly with MEC servers due to the unfavourable wireless channel make use of D2D communication with the nearby devices in order to transmit their computational tasks to the MEC server. Moreover, the authors in [28], presented a concept of employing IRS in the MEC setup in order to decrease the computation latency. Iterative algorithms are used to optimize the framework's computing and communication settings to achieve low latency. The authors, however, did not account for the power consumption of the IRS infrastructure when the same network provider is also responsible for powering the MEC servers.

To further investigate IRS's role in the MEC paradigm, we have formulated an optimization problem to minimise the computational latency of MEC and the power consumed by IRS by exploiting its frequency selectivity. As shown in Fig.5.2, an IRS-aided single user MEC scenario is illustrated. The user is unable to receive a favourable channel from the base station as a result of blocking, and the IRS-aided channel helps to minimise the latency by enhancing the capacity and rate offered to the user.

5.2 System Model and Problem Formulation

This section introduces the communications and computational models and defines a minimisation problem focusing on the computational latency and power consumed by IRS elements; details are as follows.

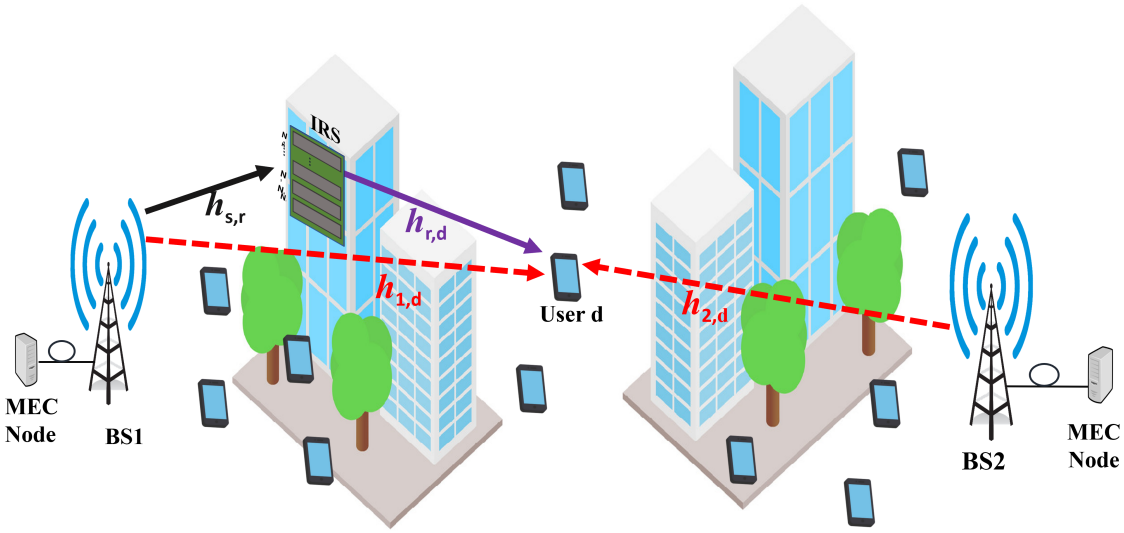


Figure 5.3: An IRS-enhanced multi-user downlink communication system with MEC nodes

5.2.1 Communication Model

The proposed IRS-assisted downlink multi-user system is illustrated in Fig.5.3, considering two base stations, $BS1$ and $BS2$. The users are uniformly distributed in a grid, and each user demands a specific data rate captured by R_d, in Mbps. Each user acts as a sink node and is severed by any of these BSs acting as a source node. Moreover, $BS1$ is supported by an IRS to enhance the receive SNR offered to the user, thus the improved SNR provides better coverage and less latency. By employing the 3GPP Urban Micro (UMi) environment from [107, Table B.1.2.1-1], having a carrier frequency of 3 GHz, the channel gains between the BSs, IRS and users are modelled . We use the LOS path between the BS-IRS and IRS-user, whereas an NLOS path is considered between the BS-user. These paths are explicitly defined for distances $\geq$ 10 m [107, Table B.1.2.1-1]. N_{tot} denotes the total number of elements

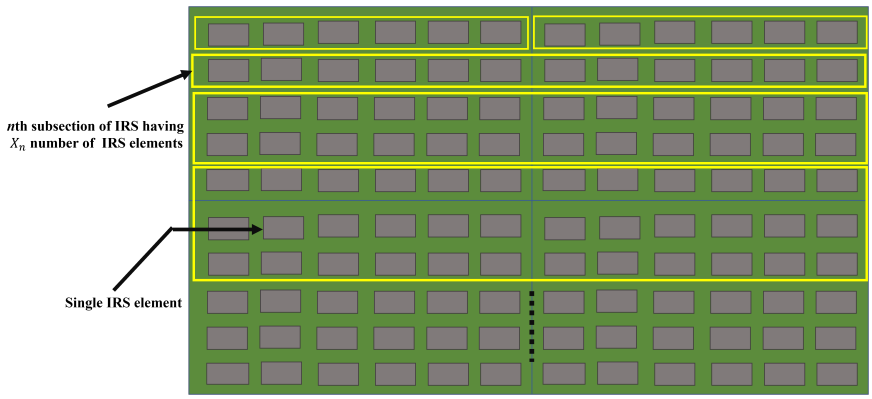


Figure 5.4: A diagram illustrating different IRS sub-sections with elements within each row controlled by a controller

in the IRS, which are divided into n sub-sections having X_n, number of elements as shown in Fig.5.4. Each user can be allocated a single or no IRS section n, according to its demand R_d. The benefit of sectioning of IRS elements is given by the fact that each sub-section of the IRS is assumed to reflect a specific channel by using frequency selectivity as previously mentioned. Each section of IRS is assigned to a particular channel , α_f so that the corresponding reflective element of that section is 1. The reflection coefficient of other elements in other sections that are not assigned to that channel is assumed to be zero. We assume 5 dBi antennas at the base station and IRS, and each user has an omnidirectional antenna with 0 dBi.

Assuming both the base station, s and user, d have single antenna, then the rate in (bit/sec/Hz) of the single-input single-output (SISO) channel without IRS is given by (12 in [87]), as

$$R_{\text{SISO}} = \log_2\left(1 + \frac{P\,|h_{\text{sd}}|^2}{\sigma^2}\right), \tag{5.1}$$

where P, is the transmit power assigned to each user by a BS, assuming fixed and equal power allocation for all the users. The deterministic flat-fading channel between base station s and the user d is denoted by $\mathrm{h}_{\text{sd}} \in \mathbb{C}$, and σ^2 represents the

Table 5.1: ENERGY-EFFICIENT MILP MODEL PARAMETERS

Parameters	Comments	Parameters	Comments
BS	Set of base stations	$P_s^{\max}$	Maximum power of a base station
U	Set of all users	P_{srb}	Transmitted power of each base station for each PRB
s	Index of the base stations	σ^2	Noise power
d	Index of user nodes	$\beta_{s,d}$	Path loss in dBs of each user from the base station
n	Index of the section of IRS elements	$\beta_{s,r}$	Path loss in dBs between the IRS and the base station
N	set of IRS sections	$\beta_{r,d}$	Path loss in dBs of each user from the IRS
X_n	Number of elements in nth IRS section,	G_{TX}	Transmitter antenna gain
N_{tot}	Total number of IRS elements	M	Very large +ve number
$d_{s,\mathrm{d}}$	Distance of user from the base station in meters	α_d	Number of requests generated by each user per second
PRB	Set of physical resource blocks	f_s	CPU computational capability of each base station s
$d_{s,r}$	Distance between IRS and the base station	c	Number of CPU Cycles required to process a single bit,
$d_{r,d}$	Distance between the IRS and the user	L	Request size in bits
B	LTE bandwidth	μ_s	Service request rate
α	Reflection Coefficient for each IRS element	w_1	Weightage parameter for the transmission delay
ϖ_{ch}	Bandwidth of each channel	w_2	Weightage parameter for the process delay
R_d	Demand of each user in Mbps	w_3	Weightage parameter for the power
ψ_s	Binary parameter that assign		consumed by the IRS elements
	IRS to only first base station,ψ_s		

receiver noise power. Furthermore, all the channels use different frequencies, so we only need to rely on SNR rather than SINR. Interestingly, the expression (5.1) only depends on the amplitude of the channel but not on the phase, therefore $|h_{\mathrm{sd}}|$ in Eq.(5.1) can be replaced with $\sqrt{\beta_{\mathrm{sd}}}$, thus

$$R_{\mathrm{SISO}} = \log_2\left(1 + \frac{P\beta_{sd}}{\sigma^2}\right), \tag{5.2}$$

where, $\beta_{s,d}$ denotes the NLOS channel gain between the base station s and user d. To get a deterministic model, we neglect the effect of shadow fading. Thus, $\beta_{s,d}$, as

Table 5.2: ENERGY-EFFICIENT MILP MODEL VARIABLES‘

Variables	Comments
$V_{s,d,n}$	Binary indicator set to 1 if the base station s uses section n in its nearby IRS to serve the user d, otherwise 0
$Y_{s,d}$	Binary indicator set to 1 if s serves a user d, otherwise 0
$C_{s,d,n}$	Capacity of each user d from the base station s over nth section of active elements of IRS
$\kappa_{s,d,n}$	Transmission delay of each user d from the base station bs over, nth section of active elements of IRS.
$\Lambda_{s,d}$	Requests of each user d, anchored by, each base station bs
λ_s	Request arrival rate from all the users served by each bs
$Z_{bs,d}$	Binary indicator set to 1 if the requests of user d are anchored by the base station bs
$\Upsilon_{s,d,n}$	Received SNR at user d which is served by a base station bs, and aided by the nth section of nearby IRS

a function of distance d_{sd} (in linear scale), is given by [107], as

$$\beta_{sd} = 10^{\left(\frac{G_{TX} - 35.1 - 36.7\log_{10}(d_{sd})}{10}\right)}, \tag{5.3}$$

where, G_{TX} represents the transmitter antenna gains at the BS. According to [87], the capacity between the base station s and user d through the n-th sub-section of IRS is given as

$$C_{s,d,n} = V_{s,d,n}\varpi_{ch}\log_2\left(1 + \left(\frac{P\left(\sqrt{\beta_{sd}} + X_n\alpha\sqrt{\beta_{IRS}}\right)^2}{\sigma^2}\right)\right), \tag{5.4}$$
$$\forall s \in BS, d \in U, n \in N,$$

where ϖ_{ch} is the bandwidth of a single channel, σ^2 is the power of white Gaussian noise, $\alpha \in (0, 1]$ denotes the fixed amplitude reflection coefficient and $\sqrt{\beta_{IRS}}$ is the product of β_{sr} and β_{rd}, representing the LOS channel gains between the BS-IRS and IRS-user respectively and are (in linear scale) given by

$$\beta_{sr} = 10^{\left(\frac{G_{TX} - 37.5 - 22\log_{10}(d_{sr})}{10}\right)}, \tag{5.5}$$

$$\beta_{rd} = 10^{\left(\frac{G_{TX} - 37.5 - 22\log_{10}(d_{rd})}{10}\right)}, \tag{5.6}$$

Note, in Eq.(5.4), n is indexed over each configuration of IRS elements denoted by N, whereas X_n refers to number of elements in each nth configuration of IRS. In Eq.(5.4), a binary variable, $V_{s,d,n}$ is generated, which is set to 1, if the system reserved a capacity $C_{s,d,n}$ between a base station s and user d over the nth configuration, otherwise it is is zero. Let $\Upsilon_{s,d,n}$ denotes the SNR between base station s and user d through the nth IRS sub-section, then from Eq.(5.4), it is given as

$$\Upsilon_{s,d,n} = \left(\frac{P\left(\sqrt{\beta_{s,d}} + X_n\alpha\sqrt{\beta_{IRS}}\right)^2}{\sigma^2}\right), \tag{5.7}$$
$$\forall s \in BS, d \in U, n \in N.$$

Thus, as the $\Upsilon_{s,d,n}$ depends upon the transmitter power, channel gains, noise and number of elements in each sub-section of IRS, we, therefore, study the impact of sectioning of IRS elements in our scheme.

5.2.2 Computational Model

In this model, assuming the limited computing capabilities of users, a complete offloading scheme is investigated where all the users tasks are loaded into the MEC server's and processed using the MEC server computing resources [70]. In complete offloading, the latency is the summation of transmission delay and process delay. Factors affecting the effectiveness of complete offloading are channel capacity and computing power of MEC server. At first, latency calculations due to the transmission and process delays in the downlink are established.

The transmission delay, $\kappa_{s,d,n}$(in seconds) is given as

$$\kappa_{s,d,n} = \begin{cases} \frac{L}{C_{s,d,n}\ 10^6} & \text{if} \quad C_{s,d,n} > 0 \\ 0 & \text{elsewise} \end{cases}, \tag{5.8}$$

$$\forall\ s \in BS, d \in U, n \in N,$$

where, L denotes the request size in bits and capacity $C_{s,d,n}$ is in Mbps. For the process delay, we assume a MEC server in each BS, and each server follows the M/M/1 queuing model, so under the existence of steady state, the process delay, ι_s(in seconds) at each base station, s can be written as

$$\iota_s = \frac{1}{\mu_s - \lambda_s}, \forall s \in BS, \tag{5.9}$$

where, μ_s and λ_s are the service request rate and request arrival rate respectively. The service request rate μ_s(requests per second), is a function of CPU processing speed in Hz (denoted by f_s), number of CPU cycles required to process a single bit (denoted by c) and L [28], and is given as

$$\mu_s = \text{floor}\left(\frac{f_s}{c\ L}\right), \forall s \in BS. \tag{5.10}$$

The request arrival rate, λ_s(requests per second), from all the users served by each base station s, and is given as

$$\lambda_s = \sum_{d \in U} \Lambda_{s,d}\ , \forall s \in BS, \tag{5.11}$$

where, $\Lambda_{s,d}$ associates all the requests of each user d to each base station s, and is

given as

$$\sum_{s \in BS} \Lambda_{s,d} = \alpha_d \ , \forall d \in U, \tag{5.12}$$

here, α_d denotes the number of requests generated by each user in each second.

5.2.3 Problem Formulation

Due to the non-linearity of Eq.(5.9) and Eq.((5.8), first, we formulate the non-linear model with the objective function and constraints in P1. Then a Linearisation technique is introduced to make the model linear, which will help us in defining the linear model, as given in P2.

P1. Non Linear Model

Our optimization problem's primary goal is to minimise the latency including transmission and process delays, and power consumed by IRS elements. In addition, due to the limited number of IRS elements (sections), the model will optimize IRS sections according to the user's demand. The problem is formulated as an MINLP, whose objective function give as

$$\text{Minimize} \quad w_1 \sum_{s \in BS, d \in U, n \in N} \kappa_{s,d,n} + w_2 \sum_{s \in BS} \iota_s + w_3 \sum_{s \in BS} \sum_{d \in U} \sum_{n \in N} V_{s,d,n} \ \chi_n. \tag{5.13}$$

with subject to the following constraints:

$$\begin{aligned} &\sum_{s \in BS} \sum_{n \in N} V_{s,d,n} \geq \varphi_d \\ &\sum_{s \in BS} \sum_{n \in N} V_{s,d,n} \leq M \ \varphi_d \ , \\ &\quad \forall \ d \in U, \end{aligned} \tag{5.14}$$

$$R_d \varphi_d \leq \sum_{s \in BS} \sum_{n \in N} C_{s,d,n}, \quad \forall\, d \in U, \tag{5.15}$$

$$\sum_{d \in U} \varphi_d = \mathrm{card}(U), \quad \forall\, d \in U, \tag{5.16}$$

$$\sum_{n \in N} V_{s,d,n} \geq Y_{s,d} \tag{5.17}$$

$$\sum_{n \in N} V_{s,d,n} \leq M\, Y_{s,d}\ ,$$

$$\forall\ s \in BS, d \in U,$$

$$\lambda_s \leq 0.9 * \mu_s \quad \forall s \in BS \tag{5.18}$$

$$\sum_{s \in BS} Y_{s,d} \leq 1, \quad \forall\, d \in U, \tag{5.19}$$

$$\sum_{n \in N} V_{s,d,n} \leq 1, \quad \forall\, s \in BS, d \in U, \tag{5.20}$$

$$\sum_{d \in U} \sum_{n \in N} \chi_n\, V_{s,d,n} \leq N_{tot}, \tag{5.21}$$

$$\forall s \in BS.$$

$$\Lambda_{s,d} \geq Y_{s,d} \tag{5.22}$$
$$\Lambda \leq M\, Y_{s,d}\ ,$$
$$\forall\ s \in BS, d \in U,$$

In constraint (5.14), the binary association variable is generated from the variable $V_{s,d,n}$, means if there at least one base station, s and one IRS section n reserved to a particular user d, then φ_d should be set to 1 for that user which indicates that the user is served. It is important to mention that when the users are served by either $BS1$(without using IRS) or by $BS2$, $V_{s,d,n}$ can still be 1 due to the fact that $n = 1$ corresponds to the scenario when zero IRS element is used. In Constraint (5.16), demand of each user, R_{d} is multiplied with the the binary variable, φ_d, to ensure that the capacity constraint holds for active user demands. Constraints (5.16) ensures all the users are served by the BSs. In Constraint (5.17), the variable $Y_{s,d}$ is generated using the variable $V_{s,d,n}$, which means that if there at least one IRS section assigned to a user d associated to BS, then $Y_{s,d}$ is set to 1, otherwise it is zero. Constraint(5.18) ensure the validity of Eq.(5.9). Constraint (5.19) ensures that each user is served by only one BS. Constraint (5.20) ensures that each user d is allocated one IRS configuration, recall that a configuration of an IRS could be to use a particular section (n>1) or not to use the IRS at all (n=1). Constraint (5.22) ensures that a user d only connects to a bastion bs that receives that user requests rate, $\Lambda_{s,d}$. Finally, Constraint (5.21) ensures that sum of IRS elements in the utilized sections of the IRS do not exceed total number of elements of that IRS.

P1.1 Linearisation of non-Linear equations

Suppose we want to find the value of delay $\kappa_{s,d,n}$ against a certain value of capacity $C_{s,d,n}$ in Eq.(5.8). First, we divide the capacity between its maximum and

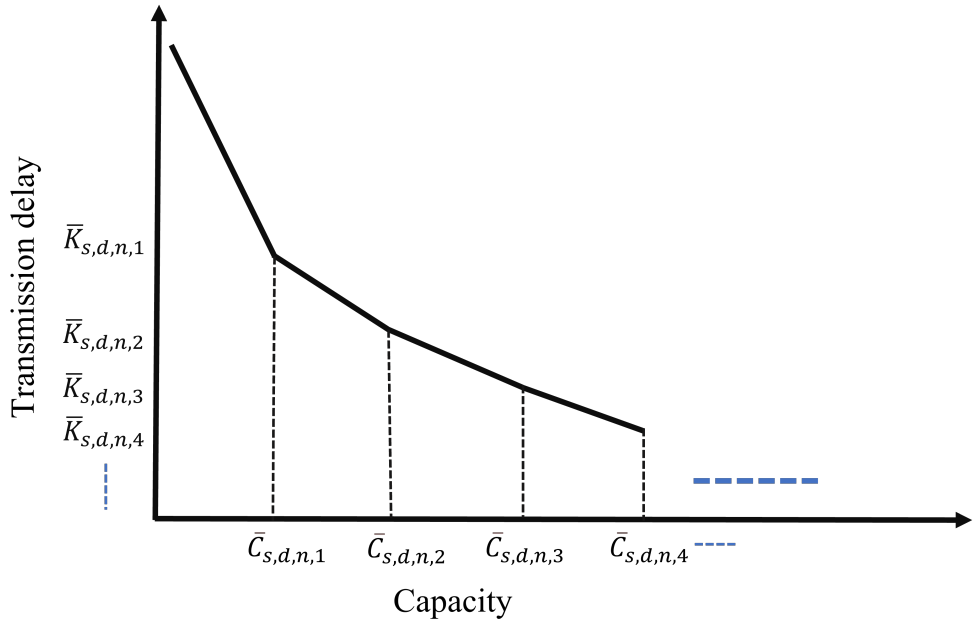


Figure 5.5: Linearisation of Transmission delay - capacity Eq.(5.8).

minimum value given by the model into small regions (or segments) as

$$\bar{C}_{s,d,n,i} = i \left(\frac{C_{s,d,n(max)} - C_{s,d,n(min)}}{I} + C_{s,d,n(min)} \right), \quad \forall \ s \in BS, d \in U, n \in N, i \in I, \tag{5.23}$$

where, I is the total number of segments. Now the transmission delay $\bar{\kappa}_{s,d,n,i}$ in each segment against $\bar{C}_{s,d,n,i}$, can be given as

$$\bar{\kappa}_{s,d,n,i} = \begin{cases} \frac{L}{\bar{C}_{s,d,n,i}\ 10^6} & \text{if} \quad \bar{C}_{s,d,n,i} > 0 \\ 0 & \text{elsewise} \end{cases}, \tag{5.24}$$

$$\forall\ s \in BS, d \in U, n \in N, i \in I.$$

The plot of $\bar{\kappa}_{s,d,n,i}$ against $\bar{C}_{s,d,n,i}$ is shown in Fig.5.5. Now as the given capacity $C_{s,d,n}$, can lie in any region between $\bar{C}_{s,d,n,i-1} - \bar{C}_{s,d,n,i}$, so we introduce a variable

$C_{s,d,n,i}$ whose value is given by

$$C_{s,d,n} = \sum_{i=1}^{I} [C_{s,d,n,i}] \ \forall \ s \in BS, d \in U, n \in N, i \in I \tag{5.25}$$

Now for each segment i, we check whether the $C_{s,d,n,i}$ lies in that segment by introducing following constraints

$$\begin{aligned} &\bar{C}_{s,d,n,i-1} Z_i \leq C_{s,d,n,i} \leq \bar{C}_{s,d,n,i} Z_i, \\ &\forall \ s \in BS, d \in U, n \in N, i \in I, \end{aligned} \tag{5.26}$$

where Z_i is a binary variable which is defined as

$$\sum_{i=1}^{I} Z_{s,d,n,i} = 1 \ \forall \ s \in BS, d \in U, n \in N, i \in I \tag{5.27}$$

The above set of constraints (Eq.(5.25)-Eq.(5.27), ensure that the variable Z_i should only be one for the value of i, where the given capacity $C_{s,d,n}$ exactly lies in the range specified by $\bar{C}_{s,d,n,i-1} - \bar{C}_{s,d,n,i}$. Now we will find the delay, $\kappa_{s,d,n}$ by summing over all the segments i as

$$\begin{aligned} \kappa_{s,d,n} = \sum_{i=1}^{I} [&\bar{\kappa}_{s,d,n,i-1} Z_{s,d,n,i} + \\ &\left(C_{s,d,n,i} - \bar{C}_{s,d,n,i-1} Z_{s,d,n,i}\right) \frac{\bar{\kappa}_{s,d,n,i} - \bar{\kappa}_{s,d,n,i-1}}{\bar{C}_{s,d,n,i} - \bar{C}_{s,d,n,i-1}}\Big] \\ &\forall \ s \in BS, d \in U, n \in N, i \in I \end{aligned} \tag{5.28}$$

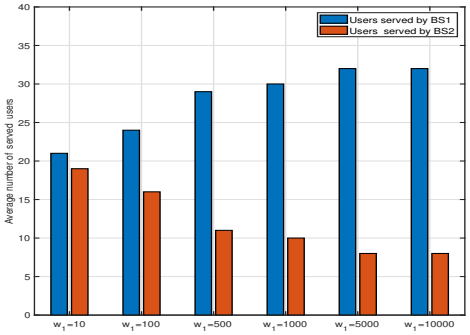
From Eq.(5.27), as $Z_{s,d,n,i}$ is only 1 for a single value of i, so it will find the approximated value of delay $\kappa_{s,d,n}$ against that capacity $C_{s,d,n}$ in that particular region i. Similar set of constants can be defined to linearise Eq.(5.9) by assuming $L = 1$ and $\mu_s - \lambda_s$ as C just for the case of reference.

Table 5.3: ENERGY-EFFICIENT MILP MODEL INPUT PARAMETERS

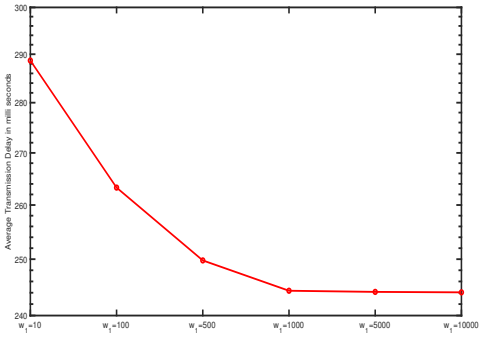
Parameters	Comments
Set of all the users, U (uniformly distributed)	40
Set of physical resource blocks, PRB	100
set of IRS sections, N	1-5
Number of elements in each nth IRS section, X_n	0,25,50,100,150
Total number of elements in an IRS, N_{tot}	1500-5000
Distance of user from the base station in meters, $d_{s,d}^{AN}$	10-100 meters
Distance between IRS and the 1st base station, $d_{s,r}^{AN}$	20 meters
Distance between the IRS and the users, $d_{r,d}^{AN}$	10-50 meters
LTE bandwidth , B	20 (MHz) [90]
Reflection Coefficient for each IRS section, α	1
Bandwidth of each channel, ϖ_{ch}	180 (kHz) [90]
Demand of each user, R_{d}	0.1-0.2 Mbps[94]
Maximum power of a base station, $P_s^{\max}$	200 mWatts [90]
Transmitted power of each base station for each PRB, P	$\frac{P_s^{\max}}{PRB}$ mWatts
Power of single active element in IRS, $\wp$	5 (mW)[24]
Noise variance, σ^2	$\left(10^{-\frac{174}{10}}\right) B$[90]
Transmitter Antenna Gain ,G_{TX}	5 dBi
Number of requests generated by each user, α_d	1-3 requests/seconds
Number of CPU Cycles required to process a single bit, c	750 cycles/sec
Number of CPU Cycles allocated by MEC server in each BS to a single user, f_s	5×10^8cycles/sec[108]
Request size in bits, L	300 Kb[109]
Service request rate, μ_s (in requests/sec)	floor($\frac{5\times10^8}{750\ 300\times10^3}$)
Binary parameter that assign IRS to only first base station,ψ_s	=1 if s=1
Weightage parameter for the transmission delay w_1	10-1000
Weightage parameter for the process delay, w_2	500-1000000
Weightage parameter for the power consumed by IRS elements, w_3	0-1000
Very large +ve number, M	10000

P2. Linear Model

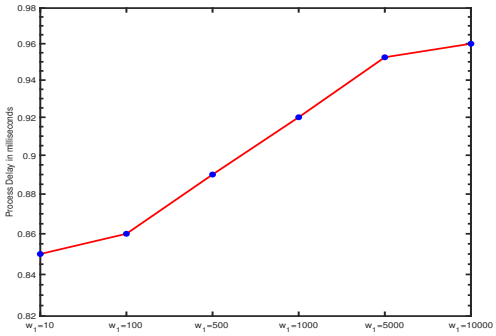
The above set of equations will linearise variables defined in Eqs.(5.9) and Eq.(5.8). Therefore, we can express our MILP model using the objective function in Eq.(5.13) and all constraints mentioned in Eq.(5.14)-Eq.(5.14).



(a) Average number of served users by each BS



(b) Average transmission delay plotted on semi-log scale



(c) Average process delay plotted on semi-log scale

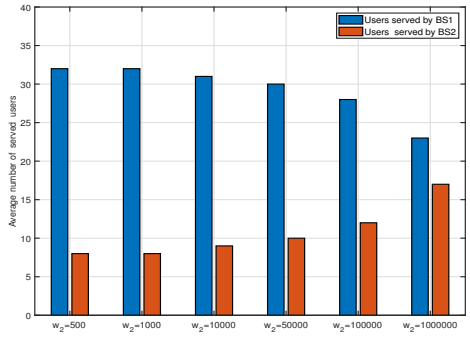
Figure 5.6: Average number of served users, transmission delay and process delay, are plotted against the weightage parameter, w_1 of the transmission delay

5.3 Numerical Results

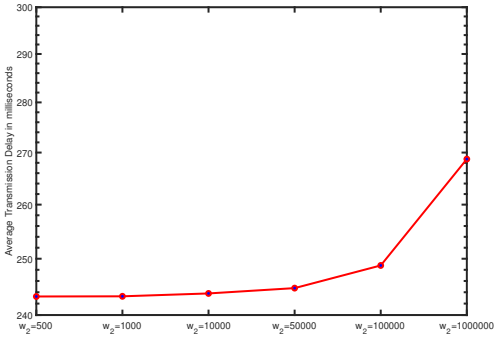
As mentioned earlier that the IRS elements are introduced to enhance spectral efficiency which would yield an improvement in downlink capacity offered by $BS1$. For the 1st set of results, we studied the impact of varying the weightage of the transmission delay in the objective function, as the service provider is more interested in minimising the transmission delay, as they don't have direct or minimal control over the IRS elements or the MEC server. In Fig.5.6, the weightage parameter, w_1 of the transmission delay in Eq.(5.13), is varied from 10 to 10000, by fixing weightage parameter for process delay, $w_2 = 100000$ and making w_3=1. The average

number of users served by $BS1$ and $BS2$, along with the average transmission and process delays are plotted for 100 simulations while guaranteeing that all the users demands are satisfied. In each simulation, users are distributed randomly with random demands and requests. It can be inferred from Fig.5.6 (a), that with an increase in w_1, users which were initially served by $BS2$ are now served by $BS1$, due to the enhanced capacity offered by $BS1$, by making use of more IRS sections. The increase in overall capacity also results in less transmission delay as shown in Fig.5.6 (b). In Fig.5.6 (c) with higher, w_1 the model focuses more on reducing the average transmission delay thus an increase in average process delay is observed. These results can be expanded to large number of BSs, but we restrict our analysis to two BSs, as the model becomes tractable and still representative of the overall system.

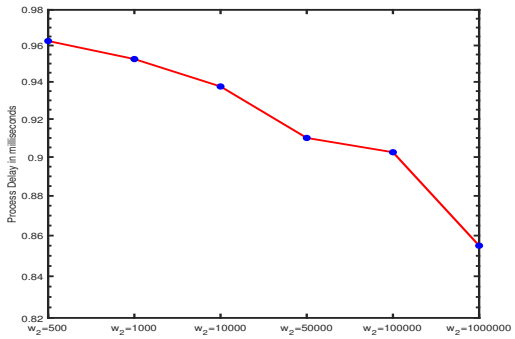
From the Fig.5.6, it has been found that the change in w_1 switched users from $BS2$ to $BS1$, but as each MEC server follows the M/M/1 queuing model, it will impact the process delay at each BS. Therefore, by fixing the weightage parameter for transmission delay, $w_1 = 1$ and varying the weightage parameter, w_2 of the process delay in Eq.(5.13), from 1 to 5000, impact on the average number of served users by both the BSs and transmission and process delays are investigated, w_3 is equal to 1 . In Fig.5.7, the average number of users served by $BS1$ and $BS2$ along with total delays are plotted for 100 simulations, while guaranteeing that demands of all the users are satisfied. It is evident that with increase in w_2, those users which $BS1$ initially served are now served by BS2 because more the users served by a single BS, the more would be the process delay as shown in Fig.5.7,(a). Hence, the model minimised the process delay by balancing the number of users served by each BS as shown in Fig.5.7(b). In Fig.5.6(c) with higher w_2 the model pays more attention on minimising the average process delay, thus an increase in average transmission



(a) Average number of served users by each BS



(b) Average transmission delay plotted on semi-log scale

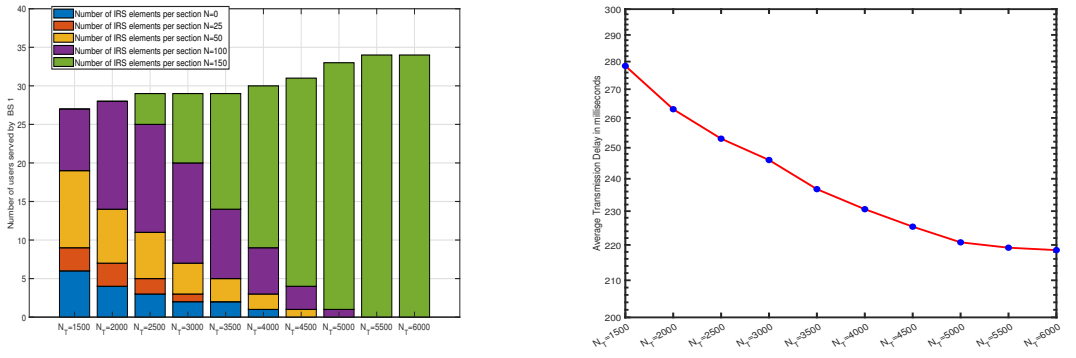


(c) Average process delay plotted on semi-log scale

Figure 5.7: Average number of served users, transmission delay and process delay, are plotted against the weightage parameter, w_2 of the process delay

delay is observed.

From the previous result it can be inferred that by the introducing IRS helped the $BS1$ to increase the capacity to its served users. In Fig.5.8, impact on the average number of served users and transmission delay by varying the total number of IRS elements N_T is investigated. In this setting, the weightage parameters are as $w_1 = 1000$, $w_2 = 50000$ and $w_3 = 1$. From Fig.5.8(a), it can be inferred that initially when N_T =1500, those users which are served by $BS1$ are using IRS sections having less number of active elements, hence, offering less capacity to the users. But with the increase in value of N_T, $BS1$ served the users with enhanced capacity and



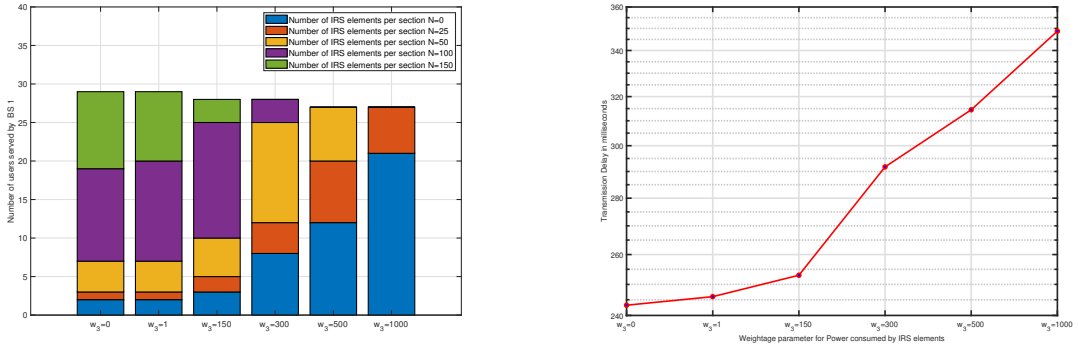
(a) Average number of served users served by $BS1$, utilizing different sections of IRS element

(b) Average transmission delay plotted on semi-log scale

Figure 5.8: Impact of total number of IRS elements on average number of served users and transmission delay

less transmission delay as shown in Fig.5.8(b) by allocating IRS sections with more elements. When N_t reaches 5500, $BS1$ allocated highest section of IRS elements(i.e., X_n=150) to all of its served users and, moreover, at this point, maximum users are served.

As introducing the IRS element not only provided better gain but also reduces the transmission delay, but it will have a burden on power consumed by active IRS elements. In Fig. 5.9, average number of served users by $BS1$ and average transmission delay are plotted against different values of weightage parameter of the power consumed by IRS elements (w_3) in the objective function (Eq.(5.13)). By fixing the other two weightage parameters for transmission and process delays as $w_1 = 1000$, $w_2 = 50000$ with N=3000. It is noted that initially for smaller values of w_3, $BS1$ served its users with IRS sections having more number of elements, thus, enhancing the capacity offered to its users resulting in less transmission delay as shown in 5.9(a) and Fig.5.9(b) respectively. But, with an increase in w_3, $BS1$ served most of its users with IRS sections having a few elements or without IRS (having zero element) offering less capacity and hence, an increase in transmission



(a) Average number of served users served by $BS1$, utilizing different sections of IRS

(b) Average Transmission delay plotted on semi-log scale

Figure 5.9: Impact of weightage parameter w_3 on average number of served users and transmission delay and transmission delay

delay is observed.

5.4 Conclusions

We have examined the problem of computational offloading in 5G/6G wireless networks using IRS to minimise latency components such as transmission and process delays. In addition, we have examined the power consumed by IRS wireless networks to meet the maximum demands of end users. As a result, the introduction of IRS elements into wireless networks minimised latency by increasing the capacity available to subscribers. Additionally, the optimal number of IRS elements is determined in order to serve as many users as possible.

Chapter 6

Genetic Algorithm Implementation

The goal of this chapter is to design a heuristic based on a genetic algorithm in order to validate the results obtained from the optimisation problems. Our decision variables were represented by binary chromosomes. We also describes a unique selection and mutation process for the population generated from those chromosomes. Furthermore, we developed a repair function that was used to verify the validity of generated chromosomes and repair according to some constraints. Various fitness functions were evaluated based on the optimization problems. We obtain similar results from heuristics as we did from the MILP model.

6.1 Introduction

In this section, an alternative method of validating the results of the MILP model is presented using a binary genetic algorithm (GA). Using a genetic algorithm, a

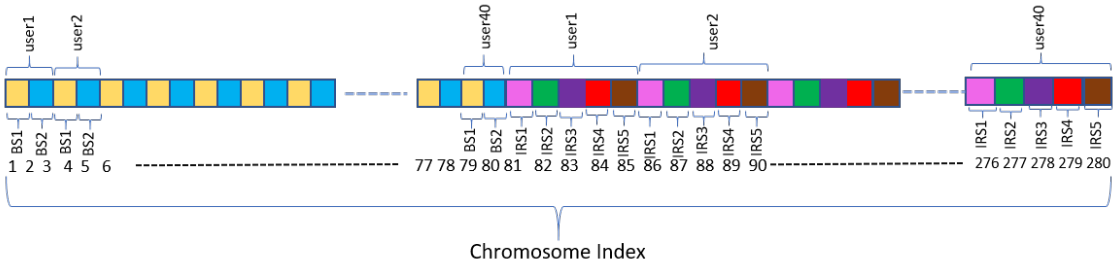


Figure 6.1: Chromosome structure.

population of candidates, or individuals, can evolve over several generations in order to identify the individuals with the greatest fitness. In this study, optimal fitness is defined as the weighted minimum of transmission delay, process delay, and power consumption of IRS elements, while candidates represent the distribution of users, BSs and IRSs in a network. As part of the first step of every GA, chromosomes are defined for each problem under consideration.

6.2 Genetic Model

We define a chromosome composed of 280 genes that corresponds to the values of $V_{bs,d,n}$, introduced in the last chapter. Each chromosome is divided into two sections, where the first sub-section corresponds to the BS-user association and the second sub-section corresponds to the user-IRS association. For the first 80 genes, each pair of gene corresponds to a single user where first and the second gene in each pair represents the user served by $BS1$ and $BS2$ respectively as shown in Fig6.1. In order to hold the user association constraint(5.19), a constraint function is formulated ensuring that only one of the gene in each pair will be equal to 1. For the last 200 genes, indexing from 81 to 280, each (quintuplets) of 5 genes represents the sections of IRS, so if a BS allocated a specific configuration of IRS (including IRS section without any elements) to a user d,the specific gene will be equal to 1

Algorithm 6.1 Genetic algorithm for minimizing Cost Function

Input: Input Parameters

Output: Cost function

1. Create the initial parent population
2. Check the constraint functions on the population
3. Repair the population where necessary
4. Evaluate the fitness function of population
5. Set a counter $i = 1$, for the number of generation
6. Select the parents
7. Perform Crossover
8. Perform Mutation
9. Get the children population
10. Check again the constraints on the population
11. Repair the population where necessary
12. Evaluate the fitness function of population
13. Merge and sort the parent and children populations
14. Create new parent population by removing extra individuals
15. if the termination criteria is not satisfied
16. $i = i + 1$
17. go to 6
18. else
19. Optimal Cost
20. EXIT
21. end if

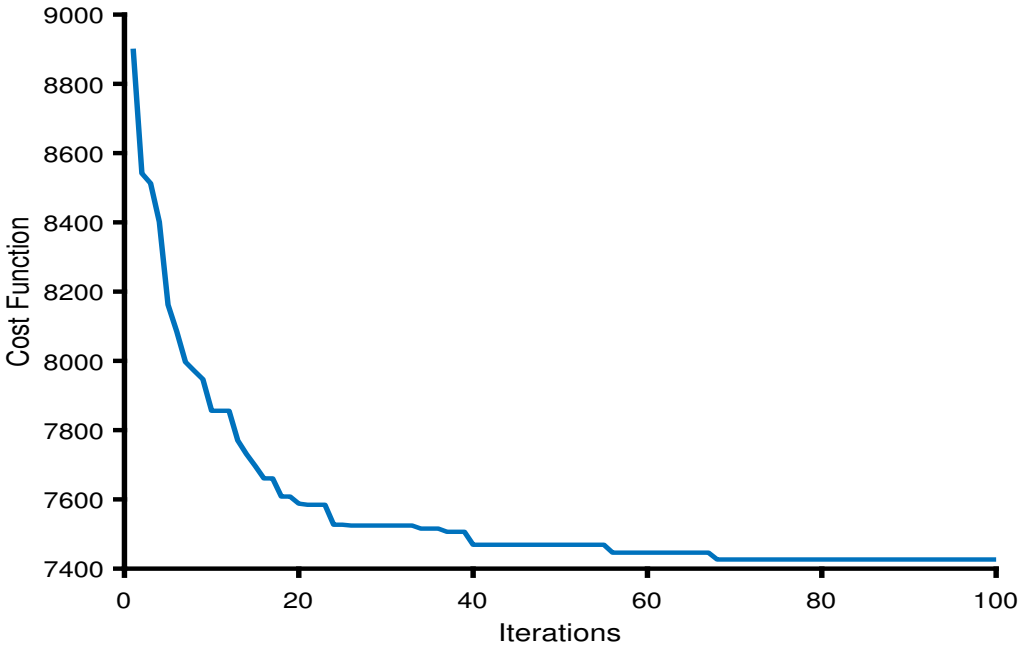


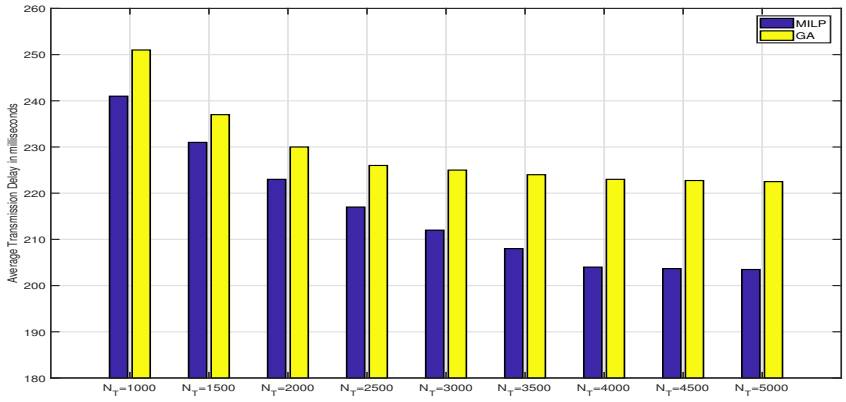
Figure 6.2: GA Objective Function is plotted over total number of iterations.

for that user. Again the constraint function for last 200 genes is formulated which ensures that only one gene in each quintuplet will be 1 to hold the constraint(5.20).

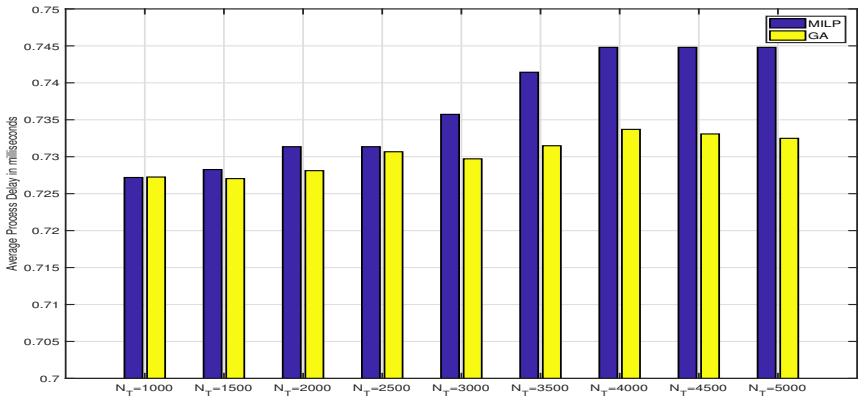
The objective of problem under consideration is to jointly minimize the transmission delay, process delay and power consumed by IRS elements (Eq.4.21), resulting in a multiple-objective optimization problem . So by assigning weights to each individual objective function and combining them into a single composite function[110], a fitness function is formulated. The crossover and mutation operations are performed after each selection operation, as exhibited in Algorithm 6.1.

6.3 GA SETUP AND RESULTS

A total of 1000 chromosomes are believed to evolve through 100 generations, with crossovers and mutations having probabilities of 99% and 0.001%, respectively. With the use of the double point crossover function, the first crossover point is



(a) Transmission Delay of GA and MILP

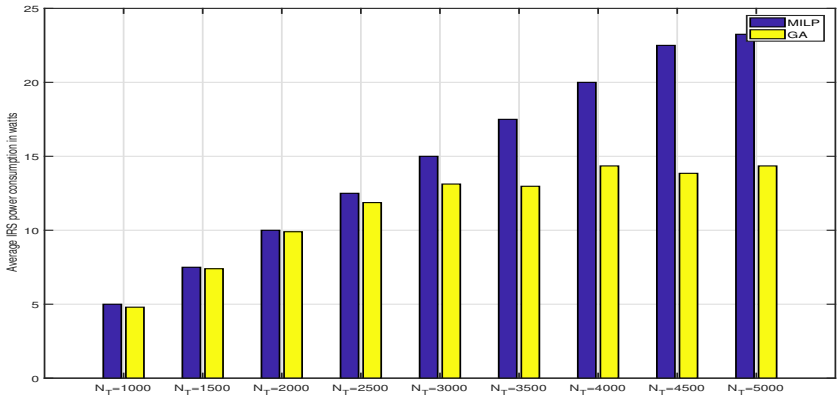


(b) Process delay of GA and MILP

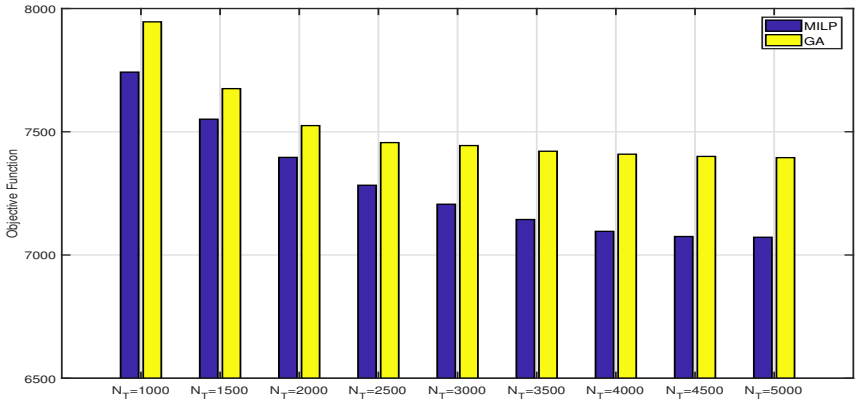
Figure 6.3: Transmission delay and process delay of GA are compared to the MILP model, plotted over different number of reflective elements N_T

randomly selected from 1-80 genes, whereas the second crossover point is randomly selected between 81-280 genes. Thus allowing only the BS-user genes of parent#1 to be crossed over with BS-user genes of parent#2 and user-IRS genes of parent#1 to be crossed over with user-IRS genes of parent#2. Children were able to preserve the same genetic information as their parents under this scheme, as illustrated in Fig.6.1.

In Fig. 6.2, the objective function of GA is plotted against number of iterations with same wights and $N_T = 3000$, and it can be inferred that the solution has been



(a) IRS power consumption of GA and MILP



(b) Objective function of GA and MILP

Figure 6.4: IRS power consumption and overall cost function of GA are compared to the MILP model, plotted over different number of reflective elements N_T

found beyond 68 generations, as no further change is observed after that point. In Fig.6.3 and Fig. 6.4 transmission delay, process delay, IRS power consumption and objective function of MILP is compared with the average transmission delay, average process delay, average IRS power consumption and average objective function of GA (each average is taken over 5 simulations), which are plotted against number of elements, N_T in IRS, with $w_1 = 500$, $w_2 = 10000$,and $w_3 = 1$ for the same setting of users distributions, demands and request rates, and it can be inferred that they are comparable.

6.4 Conclusions

In this chapter, we designed a heuristic based on a genetic algorithm to validate the results obtained from the optimisation problems. We represented our decision variables as binary chromosomes. The population generated from those chromosomes was also subjected to a unique selection and mutation process. Moreover, we have developed a repair function that can be used to verify the validity of generated chromosomes and to perform repairs in accordance with some constraints. Numerous fitness functions were examined based on optimization problems. The heuristics derived from the MILP model resulted in similar results.

Chapter 7

Conclusion and Future Work

In this chapter, a review of the main results of this thesis is presented and important conclusions are highlighted. Moreover, the possible extension and future directions of this work are also presented here.

7.1 Conclusion

7.1.1 Optimisation of Intelligent Reflecting Surfaces in 5G and Beyond Backhaul Networks

In this part of the thesis, we discussed the idea of incorporating IRS into the backhaul links for future wireless networks to reduce the backhaul power consumption of wireless networks in the event of blockage or to maximize the number of users served. A MILP optimisation model was developed to minimize backhaul power while ensuring that a maximum number of users are served. We found that introdu-

cing IRS along mmWave channels in a situation where specific mmWave channels are blocked simultaneously can save static and dynamic backhaul power compared with re-routing through only mmWave backhaul links. A further benefit of deploying the optimal number of IRS along mmWave channels is that the number of users served can be maximized. In addition, it is necessary to decrease the power consumed by an IRS element to make the power of each IRS block (which contains several elements) comparable to that of backhaul connections.

7.1.2 Joint Power and Latency Minimization of Mobile Edge Computing in Future Access Networks Assisted by Intelligent Reflecting Surfaces

This part of the thesis focuses on the computational offloading problem of utilizing IRS techniques in future 5G/6G wireless networks to minimize latency components such as transmission and process delays, as well as the power consumed by IRS wireless networks to satisfy the maximum demands from end-users. Accordingly, introducing IRS elements into a wireless network minimized the latency by enhancing the capacity offered to subscribers. The optimal value of the number of IRS elements is also evaluated in order to serve as many users as possible. In order to validate the results obtained from the MILP model, a genetic algorithm has been introduced. It can be concluded that both models yield similar outcomes.

7.2 Future Work

Future work will include for the optimisation of the backhaul network, an optimal multi-antenna scenario will be developed. Moreover a heuristic that will aid in the

analysis of possible savings in BH power will be developed. Analyse of mmWave channels to determine the location of blockages can be another interesting work.

In order to optimize both the access network and the backhaul network, a multichannel optimization model will be developed.

Bibliography

[1] M. Jung, W. Saad, M. Debbah, and C. S. Hong, "On the optimality of reconfigurable intelligent surfaces (riss): Passive beamforming, modulation, and resource allocation," *IEEE Transactions on Wireless Communications*, 2021.

[2] Z. Qingling and J. Li, "Rain attenuation in millimeter wave ranges," in *2006 7th International Symposium on Antennas, Propagation & EM Theory.* IEEE, 2006, pp. 1–4.

[3] ((Accessed 29 March 2019)) Attenuation by atmospheric gases. Available:https://www.itu.int/dms pubrec/.

[4] Y. Liu and D. M. Blough, "Blockage tolerance in roadside millimeter-wave backhaul networks," *Computer Networks*, vol. 198, p. 108377, 2021.

[5] Eriscon, "Mobility report 2019," Accessed 29 September, 2020, [Online] Available: https://www.ericsson.com/4acd7e/assets/local/mobility-report/documents/2019/emr-november-2019.pdf.

[6] V. Ziegler and S. Yrjola, "6g indicators of value and performance," in *2020 2nd 6G wireless summit (6G SUMMIT).* IEEE, 2020, pp. 1–5.

[7] Cisco, "2020 annual internet report," Accessed 18 September, 2020, [Online] Available: https://www.cisco.com/c/en/us/solutions/collateral/executiveperspectives/annual-internet-report/white-paper-c11-741490.html.

[8] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dom-

inant theme for wireless evolution into 5g," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, 2014.

[9] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka *et al.*, "Scenarios for 5g mobile and wireless communications: the vision of the metis project," *IEEE communications magazine*, vol. 52, no. 5, pp. 26–35, 2014.

[10] X. Ge, S. Tu, G. Mao, C.-X. Wang, and T. Han, "5g ultra-dense cellular networks," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 72–79, 2016.

[11] D. López-Pérez, M. Ding, H. Claussen, and A. H. Jafari, "Towards 1 gbps/ue in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2078–2101, 2015.

[12] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[13] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. prentice hall PTR New Jersey, 1996, vol. 2.

[14] A. Mesodiakaki, A. Kassler, E. Zola, M. Ferndahl, and T. Cai, "Energy efficient line-of-sight millimeter wave small cell backhaul: 60, 70, 80 or 140 ghz?" in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. IEEE, 2016, pp. 1–9.

[15] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE journal on selected areas in communications*, vol. 32, no. 6, pp. 1164–1179, 2014.

[16] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Transactions on Communications*, 2021.

[17] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 162–169, 2018.

[18] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive mimo is a reality what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, 2019.

[19] E. Björnson, Ö. Özdogan, and E. G. Larsson, "Reconfigurable intelligent surfaces: Three myths and two critical questions," *IEEE Communications Magazine*, vol. 58, no. 12, pp. 90–96, 2020.

[20] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing key technology towards 5g," *ETSI white paper*, vol. 11, no. 11, pp. 1–16, 2015.

[21] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE access*, vol. 7, pp. 116 753–116 773, 2019.

[22] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable ai metasurfaces: An idea whose time has come," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–20, 2019.

[23] J. Huang, "Reflectarray antenna," *Encyclopedia of RF and Microwave Engineering*, 2005.

[24] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

[25] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. Di Renzo *et al.*, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45 913–45 923, 2020.

[26] M. Al-Jarrah, E. Alsusa, A. Al-Dweik, and M.-S. Alouini, "Performance analysis of wireless mesh backhauling using intelligent reflecting surfaces," *IEEE Transactions on Wireless Communications*, 2021.

[27] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint optimization of radio and computational resources for multicell mobile-edge computing," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 1, no. 2, pp. 89–103, 2015.

[28] T. Bai, C. Pan, Y. Deng, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Latency minimization for intelligent reflecting surface aided mobile edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2666–2682, 2020.

[29] J. Zhao, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless networks," 1905.

[30] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy

communication: Is artificial noise helpful or not?" *IEEE Wireless Communications Letters*, vol. 9, no. 6, pp. 778–782, 2020.

[31] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2283–2314, 2020.

[32] W. Cai, H. Li, M. Li, and Q. Liu, "Practical modeling and beamforming for intelligent reflecting surface aided wideband systems," *IEEE Communications Letters*, vol. 24, no. 7, pp. 1568–1571, 2020.

[33] U. Karthaus and M. Fischer, "Fully integrated passive uhf rfid transponder ic with 16.7-μw minimum rf input power," *IEEE Journal of solid-state circuits*, vol. 38, no. 10, pp. 1602–1608, 2003.

[34] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced ofdm: Channel estimation and reflection optimization," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 518–522, 2019.

[35] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical csi," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8238–8242, 2019.

[36] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5849–5863, 2020.

[37] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[38] ——, "Beamforming optimization for intelligent reflecting surface with dis-

crete phase shifts," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 7830–7833.

[39] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate optimization for intelligent reflecting surface enhanced wireless networks," *arXiv preprint arXiv:1905.07920*, 2019.

[40] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Intelligent reflecting surface for multicell mimo communications," *arXiv preprint arXiv:1907.10864*, 2019.

[41] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets ofdm: Protocol design and rate maximization," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4522–4535, 2020.

[42] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided swipt," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 586–590, 2019.

[43] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1410–1414, 2019.

[44] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE Journal on Selected Areas in communications*, vol. 30, no. 3, pp. 497–508, 2012.

[45] S. Sun, K. Adachi, P. H. Tan, Y. Zhou, J. Joung, and C. K. Ho, "Heterogeneous network: An evolutionary path to 5g," in *2015 21st Asia-Pacific Conference on Communications (APCC)*. IEEE, 2015, pp. 174–178.

[46] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji,

"Cell association and interference coordination in heterogeneous lte-a cellular networks," *IEEE Journal on selected areas in communications*, vol. 28, no. 9, pp. 1479–1489, 2010.

[47] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5g networks: A survey and an outlook," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1018–1044, 2016.

[48] A. Ulvan, R. Bestak, and M. Ulvan, "Handover procedure and decision strategy in lte-based femtocell network," *Telecommunication systems*, vol. 52, no. 4, pp. 2733–2748, 2013.

[49] X. Ge, H. Cheng, M. Guizani, and T. Han, "5g wireless backhaul networks: challenges and research advances," *IEEE network*, vol. 28, no. 6, pp. 6–11, 2014.

[50] H. Lehpamer, "Millimeter-wave radios in backhaul networks," *Communication Inrastructure Corporation*, 2008.

[51] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges," *Wireless networks*, vol. 21, no. 8, pp. 2657–2676, 2015.

[52] U. Ofcom, "Update on 5g spectrum in the uk," 2017.

[53] D. Hogg, "Path diversity in propagation of millimeter waves through rain," *IEEE Transactions on Antennas and Propagation*, vol. 15, no. 3, pp. 410–415, 1967.

[54] S. L. Johnston, "Millimeter wave radar," *Dedham*, 1980.

[55] R. Crane, "Propagation phenomena affecting satellite communication systems operating in the centimeter and millimeter wavelength bands," *Proceedings of the IEEE*, vol. 59, no. 2, pp. 173–188, 1971.

[56] ((Accessed 22 March 2019)) Specific attenuation model for rain for use in prediction methods,p.838-3. Available:https://www.itu.int/rec/R.

[57] I. Union, "Attenuation by atmospheric gases," *Recommendation ITU-R P. 676*, vol. 10, 2013.

[58] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *IEEE access*, vol. 1, pp. 335–349, 2013.

[59] T. S. Rappaport, S. Sun, and M. Shafi, "5g channel model with improved accuracy and efficiency in mmwave bands," *IEEE 5G Tech Focus*, vol. 1, no. 1, pp. 1–6, 2017.

[60] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE transactions on antennas and propagation*, vol. 61, no. 4, pp. 1850–1859, 2012.

[61] K. Haneda, J. Zhang, L. Tan, G. Liu, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li *et al.*, "5g 3gpp-like channel models for outdoor urban microcellular and macrocellular environments," in *2016 IEEE 83rd vehicular technology conference (VTC spring)*. IEEE, 2016, pp. 1–7.

[62] J. D. Parsons, *The mobile radio propagation channel*. Wiley, 2000.

[63] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 450–465, 2017.

[64] M. Othman, S. A. Madani, S. U. Khan *et al.*, "A survey of mobile cloud

computing application models," *IEEE communications surveys & tutorials*, vol. 16, no. 1, pp. 393–413, 2013.

[65] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches," *Wireless communications and mobile computing*, vol. 13, no. 18, pp. 1587–1611, 2013.

[66] L. Liu, R. Moulic, and D. Shea, "Cloud service portal for mobile device management," in *2010 IEEE 7th International Conference on E-Business Engineering*. IEEE, 2010, pp. 474–478.

[67] M. ETSI, "Mobile edge computing-introductory technical white paper," *etsi2014mobile, no. Issue*, 2014.

[68] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.

[69] S. Barbarossa, S. Sardellitti, and P. Di Lorenzo, "Communicating while computing: Distributed mobile cloud computing over 5g heterogeneous networks," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 45–55, 2014.

[70] X. Shan, H. Zhi, P. Li, and Z. Han, "A survey on computation offloading for mobile edge computing information," in *2018 IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing,(HPSC) and IEEE International Conference on Intelligent Data and Security (IDS)*. IEEE, 2018, pp. 248–251.

[71] N. Shan, Y. Li, and X. Cui, "A multilevel optimization framework for computation offloading in mobile edge computing," *Mathematical Problems in Engineering*, vol. 2020, 2020.

[72] E. Oki, *Linear programming and algorithms for communication networks: a practical guide to network design, control, and management.* CRC Press, 2019.

[73] J. Bisschop, *AIMMS optimization modeling.* Lulu. com, 2006.

[74] J. Lee, "A first course in linear optimization," 2016.

[75] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali, *Linear programming and network flows.* John Wiley & Sons, 2008.

[76] E. L. Lawler and D. E. Wood, "Branch-and-bound methods: A survey," *Operations research*, vol. 14, no. 4, pp. 699–719, 1966.

[77] M. Pióro and D. Medhi, *Routing, flow, and capacity design in communication and computer networks.* Elsevier, 2004.

[78] J. Kallrath, *Modeling languages in mathematical optimization.* Springer Science & Business Media, 2013, vol. 88.

[79] R. Fourer, D. M. Gay, and B. W. Kernighan, "A modeling language for mathematical programming," *Management Science*, vol. 36, no. 5, pp. 519–554, 1990.

[80] J. Forrest. Cbc (coin-or branch and cut) open-source mixed integer programming solver. [Online]. Available: https://projects.coin-or.org/Cbc, 2012

[81] V. Conitzer, "Lecture notes 3: Solving linear and integer programs using the gnu linear programming kit."

[82] IBM, "12.2 user manual for cplex."

[83] K. E. Kinnear Jr, "A perspective on the work in this book," *Advances in genetic programming*, pp. 3–19, 1994.

[84] D. Goldberg, "Genetic algorithms in search, optimization, and machine learning, addison-wesley, reading, ma, 1989," *NN Schraudolph and J*, vol. 3, no. 1, 1989.

[85] M. Mitchell, "Genetic algorithms: An overview." in *Complex.*, vol. 1, no. 1. Citeseer, 1995, pp. 31–39.

[86] G. Association *et al.*, "The mobile economy 2018," *London: GSM Association*, 2016.

[87] E. Björnson, Ö. Özdogan, and E. G. Larsson, "Intelligent reflecting surface vs. decode-and-forward: How large surfaces are needed to beat relaying?" *arXiv preprint arXiv:1906.03949*, 2019.

[88] F. Farias, M. Fiorani, S. Tombaz, M. Mahloo, L. Wosinska, J. C. Costa, and P. Monti, "Cost-and energy-efficient backhaul options for heterogeneous mobile network deployments," *Photonic Network Communications*, vol. 32, no. 3, pp. 422–437, 2016.

[89] L. ETSI, "Evolved universal terrestrial radio access (e-utra)," *User Equipment (UE) radio access capabilities (3GPP TS 36.306 version 10.4. 0 Release 10)*, 2012.

[90] E. U. T. R. Access, "Study on small cell enhancements for eutra and eutran higher layer aspects (release 12)„" 2013.

[91] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume *et al.*, "How much energy is needed to run a wireless network?" *IEEE wireless communications*, vol. 18, no. 5, pp. 40–49, 2011.

[92] ((Accessed 19 July 2019)) Piecewise linear functions in mip models. Available:

http://yetanothermathprogrammingconsultant.blogspot.com/2015/10/piecewise-linear-functions-in-mip-models.html.

[93] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 581–585, 2019.

[94] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5g backhaul challenges and emerging research directions: A survey," *IEEE access*, vol. 4, pp. 1743–1766, 2016.

[95] E. Zola, A. J. Kassler, and W. Kim, "Joint user association and energy aware routing for green small cell mmwave backhaul networks," in *2017 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2017, pp. 1–6.

[96] P. Wang, Y. Li, L. Song, and B. Vucetic, "Multi-gigabit millimeter wave wireless communications for 5g: From fixed access to cellular networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 168–178, 2015.

[97] A. Mesodiakaki, E. Zola, and A. Kassler, "Joint user association and backhaul routing for green 5g mesh millimeter wave backhaul networks," in *Proceedings of the 20th ACM International Conference on Modelling, Analysis and Simulation of Wireless and Mobile Systems*. ACM, 2017, pp. 179–186.

[98] H. M. Corp, "Highly integrated 60 ghz radio transceiver chipset," 2012. [Online]. Available: https://www.microwavejournal.com/articles/17871-highly-integrated-60-ghz-radio-transceiver-chipset

[99] ((Accessed 29 March 2019)) Stq-cw-vfvf025-f1. Available:https://www.sagemillimeter.com/content/datasheets/STQ-CW-VFVF025-F1.pdf.

[100] "A 18mw, 3.3db nf, 60ghz lna in 32nm soi cmos technology with autonomic nf calibration," (Accessed 09 July 2019), available:https://users.ece.cmu.edu/xinli/papers/.

[101] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light: Science & Applications*, vol. 3, no. 10, pp. e218–e218, 2014.

[102] N. Kaina, M. Dupré, G. Lerosey, and M. Fink, "Shaping complex microwave fields in reverberating media with binary tunable metasurfaces," *Scientific reports*, vol. 4, no. 1, pp. 1–8, 2014.

[103] X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, and E. Björnson, "Ris-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials," *arXiv preprint arXiv:2103.00534*, 2021.

[104] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, "Internet of things in the 5g era: Enablers, architecture, and business models," *IEEE journal on selected areas in communications*, vol. 34, no. 3, pp. 510–527, 2016.

[105] D. Computing, "Storage and radio resource allocation over cooperative femtocells (tropic)," 2012.

[106] Y. Wang, M. Sheng, X. Wang, L. Wang, and J. Li, "Mobile-edge computing: Partial computation offloading using dynamic voltage scaling," *IEEE Transactions on Communications*, vol. 64, no. 10, pp. 4268–4282, 2016.

[107] E. U. T. R. Access, "Further advancements for e-utra physical layer aspects (release 9), 3gpp," *TS*, vol. 36, p. V9, 2010.

[108] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offload-

ing for mobile-edge cloud computing," *IEEE/ACM Transactions on Networking*, vol. 24, no. 5, pp. 2795–2808, 2015.

[109] H. Guo, J. Zhang, J. Liu, and H. Zhang, "Energy-aware computation offloading and transmit power allocation in ultradense iot networks," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4317–4329, 2018.

[110] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliability engineering & system safety*, vol. 91, no. 9, pp. 992–1007, 2006.